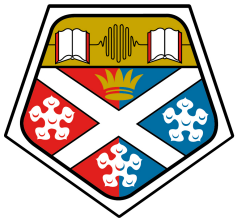


University of Strathclyde Glasgow

A thesis submitted to the
Department of Mathematics and Statistics
University of Strathclyde
in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

Robust Spectral Coarse Spaces for Indefinite Helmholtz-Type Problems

Author:
Mark Fry

Supervisor:
Matthias Langer
Victorita Dolean

April 24, 2026

Department of Mathematics and Statistics

Signed:

Date:

Acknowledgments

This thesis would not have been possible without the guidance, encouragement and support of many people. I am extremely grateful of my supervisors, Dr. Matthias Langer and Prof. Victorita Dolean, for their advice and constant patience throughout this work. Their high standards shaped both the direction and quality of this thesis.

I have also benefited greatly from discussions and collaboration with Ivan G. Graham (University of Bath), Pierre–Henri Tournier (Sorbonne Université, CNRS), Émile Parolin (Sorbonne Université, Inria), and Théophile Chaumont–Frelet (Inria). Their technical insights, constructive criticism, and generosity with their time have been invaluable.

Finally, I owe a debt of gratitude to my family and friends for their unwavering encouragement and understanding.

Abstract

High–frequency Helmholtz problems remain exceptionally difficult for both direct and iterative solvers. Finite element discretisations produce large, indefinite and often non-normal linear systems; standard preconditioners typically lose robustness as the wavenumber and the number of subdomains increase. Domain Decomposition Methods (DDMs) address these issues by splitting the global problem into local subproblems solved in parallel by direct methods. While DDMs may be used as iterative solvers, they have proven to be effective preconditioners within Krylov methods, where the use of an appropriate coarse space can significantly affect convergence and scalability.

This thesis develops scalable two–level preconditioners tailored to the Helmholtz problem. After establishing the mathematical background, we overview classical DDMs. The core contributions then focus on spectral coarse spaces. First, we revisit the GenEO coarse space (Δ-GenEO) and introduce an augmented variant, Δ_k-GenEO, obtained by modifying the local generalised eigenproblems to incorporate frequency information. We analyse eigenvalue selection strategies and the trade–off between coarse dimension and robustness. Second, we propose a new Helmholtz adapted space, H_k-GenEO, constructed from the indefinite Helmholtz problem. As previous methods have been based on positive (semi-) definite problems, we first had to develop the underlying theory to prove the required bounds for the method. We use numerical simulations to demonstrate the methods' effectiveness. We finish this work with a systematic comparative numerical study of newly developed coarse space designs in the GenEO family on various homogeneous and heterogeneous problems.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Wave propagation phenomena play a central role in science and engineering, appearing in acoustics, electromagnetics, seismology, and elasticity. Their mathematical description often reduces, under time-harmonic assumptions, to the Helmholtz equation. The numerical approximation of this equation represents one of the major challenges in scientific computing: the operator is indefinite, highly oscillatory at large frequencies, and leads to large, poorly conditioned linear systems when discretised. Throughout this thesis, $k > 0$ denotes the wavenumber, which quantifies the oscillatory nature of the solution and is proportional to the frequency of the underlying wave. Classical solvers, including multigrid and sweeping methods, tend to lose robustness as the wavenumber grows, making the design of scalable iterative solvers a topic of enduring research interest.

Domain decomposition (DD) methods have emerged as one of the most promising frameworks to address these challenges. By splitting the global domain into overlapping subdomains, DD solvers enable parallelism and local adaptivity. We write H for the characteristic diameter of a subdomain, that is, the typical size of the subdomains in the decomposition. However, for indefinite problems such as the Helmholtz equation, one-level domain decomposition methods suffer from a severe lack of scalability: information is exchanged only between neighbouring subdomains, and convergence deteriorates as the number of subdomains increases or as the frequency rises. To restore scalability, two-level methods introduce a coarse space that mediates global coupling across the domain. The design of such coarse spaces—particularly those that remain robust with respect to the wavenumber k, the subdomain diameter H, and coefficient heterogeneity—is the central focus of this thesis.

Spectral coarse spaces and the GenEO framework. Over the past decade, spectral coarse spaces have been recognised as an effective way to achieve robustness in domain decomposition preconditioners. They are constructed from local generalised eigenvalue problems that capture low-energy or slowly decaying error components that are otherwise poorly reduced by local solvers. The Generalised Eigenproblems in the Overlap (GenEO) framework provides a systematic and theoretically justified methodology for this construction in symmetric positive definite (SPD) settings. For elliptic problems, GenEO-based preconditioners yield convergence rates independent of the number of subdomains and of coefficient contrasts.

Extending these ideas to indefinite and non-self-adjoint problems, such as the Helmholtz equation, is far from straightforward. Early work introduced the so-called Δ-GenEO coarse space, which replaces the indefinite operator by a nearby SPD surrogate, allowing one to retain the existing theory at the price of pessimistic bounds on the subdomain size and eigenvalue threshold. In spectral coarse space constructions, we denote by $\tau > 0$ the eigenvalue threshold used to select the local eigenvectors retained in the coarse space. Building on these developments, the present work introduces a new generation of coarse spaces which comes in two flavours. A first refinement, named the Δ_k-GenEO variant, relaxed these constraints, showing that robust GMRES convergence can be obtained under much milder k-explicit conditions. These analyses

clarified the surprising practical efficiency of SPD-based spectral coarse spaces even for indefinite Helmholtz problems. A later addition is called the H_k-GenEO method—and constructs the coarse basis directly from the indefinite Helmholtz operator, using k-weighted norms and decoupled fine and coarse decompositions. This formulation bridges the gap between the heuristic efficiency of previous SPD-based approaches and the need for truly operator-consistent spectral constructions.

Throughout this work, we write $a \lesssim b$ to mean that there exists a constant $C > 0$, independent of key parameters (e.g., k, H, τ), such that $a \leq Cb$. Similarly, $a \gtrsim b$ means $b \lesssim a$, and $a \sim b$ means both $a \lesssim b$ and $a \gtrsim b$.

Contributions and organisation of the thesis. The thesis is organised into six chapters, described as follows.

- **Chapter 2: Mathematical background.** This chapter establishes the theoretical setting of the Helmholtz problem, from its derivation and boundary conditions to its weak formulation, finite element discretisation, and the numerical difficulties that arise at high frequency. It reviews key phenomena such as the pollution effect, the loss of coercivity, and the resulting indefiniteness of the discrete system, providing the analytical basis for subsequent solver design.

- **Chapter 3: Overview of domain decomposition methods.** The foundations of Schwarz methods and their modern additive variants are recalled, followed by an exposition of one-level and two-level domain decomposition preconditioners. The limitations of classical one-level schemes in terms of scalability and robustness are illustrated, motivating the introduction of coarse spaces. The chapter concludes with a survey of existing coarse-space families, including geometric, partition-of-unity, and spectral constructions such as GenEO and Δ-GenEO.

- **Chapter 4: The $\boldsymbol{\Delta_k}$-GenEO method.** This chapter develops and analyses the Δ_k-GenEO coarse space, an SPD-based variant of the GenEO framework tailored to indefinite Helmholtz problems. It extends the earlier Δ-GenEO approach by introducing a wavenumber-dependent right-hand side in the local eigenvalue problems, tightening the connection between the theoretical analysis and the empirical behaviour of spectral coarse spaces. The main theoretical contribution establishes sharper k-explicit conditions for GMRES convergence in two-level additive Schwarz preconditioners, improving the scaling from $H \sim k^{-2}$ to $H \sim k^{-1}$ and from $\tau \sim k^8$ to $\tau \sim k^2$. This result clarifies why coarse spaces built from SPD surrogates of indefinite operators can remain robust for moderate frequencies, despite their formal limitations.

- **Chapter 5: A new coarse space for Helmholtz — the H_k-GenEO method.** This is the central theoretical contribution of the thesis. It defines the H_k-GenEO coarse space through local generalised eigenvalue problems involving the full indefinite Helmholtz operator. Rigorous estimates are derived for the dependence of GMRES convergence on the wavenumber, the subdomain diameter, and the spectral threshold. The main result proves that robustness is achieved under the sharp conditions $H_\ell \lesssim k^{-1}$ and $\tau \gtrsim (1 + C_{\text{stab}})^2 k^2$, thus improving the previous Δ-GenEO bounds. The chapter also analyses the influence of using separate decompositions for local and coarse levels, demonstrating improved flexibility and scalability.

- **Chapter 6: Comparison of spectral coarse spaces.** A detailed numerical assessment compares several spectral coarse spaces, including DtN, Δ-GenEO, H-GenEO, and H_k-GenEO, on homogeneous and heterogeneous Helmholtz test cases. The results confirm that H_k-GenEO exhibits a robust performance, though this can deteriorate for larger

complex problems. Applications range from canonical two-dimensional benchmarks to three-dimensional configurations relevant to medical imaging and geophysical modelling.

Summary. In summary, this thesis advances the theory and practice of spectral coarse spaces for indefinite Helmholtz-type problems. It clarifies the range of validity of SPD-based constructions through the Δ_k-GenEO analysis and introduces the H_k-GenEO framework, which achieves full wavenumber robustness within a rigorous theoretical foundation. Together, these results provide new insights into the spectral mechanisms governing domain decomposition preconditioners for wave problems and offer practical guidelines for their scalable implementation in large-scale simulations.

Related works. During the course of this work, we have had three papers submitted for publication, and one which is currently in preparation.

- In preparation, "Spectral coarse spaces based on indefinite operators: the H_k-GenEO method", in collaboration with Victorita Dolean, Matthias Langer, Ivan G. Graham and Théophile Chaumont-Frelet.
- "Can Symmetric Positive Definite (SPD) coarse spaces perform well for Helmholtz problems?" in collaboration with Victorita Dolean and Matthias Langer. This has been accepted for publication in the Journal of Computational and Applied Mathematics [25].
- "Achieving wavenumber robustness in domain decomposition for heterogeneous Helmholtz equation: an overview of spectral coarse spaces" in collaboration with Victorita Dolean, Matthias Langer, Pierre-Henri Tournier and Emile Parolin. A preprint [27] can be found on the arXiv, and it has been submitted for publication in the journal Computers and Mathematics with Applications.
- "Wavenumber explicit estimates of the Schwarz preconditioner with Δ-GenEO coarse space for the indefinite Helmholtz problem" in collaboration with Victorita Dolean and Matthias Langer. This has been accepted for publication in the Proceedings of the 28th International Conference on Domain Decomposition Methods [26].

Chapter 2

Mathematical Background

This chapter lays the mathematical foundations for the study of high-frequency wave propagation via the Helmholtz equation. We begin by introducing the Helmholtz problem and its derivation from the time-harmonic wave equation, highlighting its central role in modelling oscillatory physical phenomena. We then discuss key types of boundary conditions—Dirichlet and impedance—and their implications for both analysis and numerical simulation. With this context, we define the interior heterogeneous Helmholtz problem and recast it in variational form, emphasising the challenges posed by its indefinite and non-coercive nature. The chapter proceeds to discretise the problem using the finite element method (FEM), leading to large, indefinite linear systems whose solution requires specialised techniques. We analyse the algebraic structure of the discrete system and describe the phenomena of pollution error and numerical dispersion, which significantly impact accuracy at high frequencies. Finally, we turn to iterative solvers—especially GMRES—and the vital role of preconditioning. Various preconditioning strategies are surveyed, including multigrid methods, shifted Laplacian approaches, sweeping strategies, and domain decomposition methods, setting the stage for the coarse space constructions explored in subsequent chapters.

2.1 The Helmholtz Problem

This section introduces the *Helmholtz equation*, a fundamental elliptic partial differential equation (PDE) used to model time-harmonic wave propagation in various physical contexts. The Helmholtz equation appears in many areas of science and engineering, including acoustics, electromagnetics, and elasticity, and forms the mathematical core of this work. We consider the general form of the Helmholtz equation in a domain $\Omega \subset \mathbb{R}^d$:

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

where:

- Δ is the Laplacian operator,
- u is the unknown field (e.g., pressure, electromagnetic potential),
- f is the source term, and
- k is the wavenumber, which may be constant or spatially varying.

Physically, the wavenumber k is related to the angular frequency ω and local wave speed $c(\mathbf{x})$ by

$$k(\mathbf{x}) = \frac{\omega}{c(\mathbf{x})}.$$

A more general version of the Helmholtz equation accounts for material heterogeneities through a symmetric positive-definite matrix $A(\mathbf{x})$:

$$-\nabla \cdot (A(\mathbf{x})\nabla u) - k^2 u = f.$$

Here, the matrix $A(\mathbf{x})$ modifies the diffusive behaviour of the operator, and its variation reflects the local physical properties of the medium.

Derivation from the Wave Equation In order to give some context to the importance of the Helmholtz equation, and how it can arise naturally, we shall use the wave equation. The wave equation is used to describe the propagation of waves through a medium. In some situations, the Helmholtz equation can be considered as a special case of the wave equation. Using a constant wave speed, c, the wave equation can be given as

$$\Delta U(\mathbf{x}, t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} U(\mathbf{x}, t) = 0,$$

where $U(\mathbf{x}, t)$ is the wave function. Assuming a harmonic time dependence, we then split the wave function into a spatial and time-dependent component,

$$U(\mathbf{x}, t) = u(\mathbf{x}) e^{-i\omega t}.$$

Substituting $U(\mathbf{x}, t)$ back into the wave equation we obtain

$$\Delta U(\mathbf{x}, t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} U(\mathbf{x}, t) = \left(\Delta u(\mathbf{x}) + \frac{\omega^2}{c^2} u(\mathbf{x})\right) e^{-i\omega t} = 0.$$

As the exponential term is non-zero, this gives the homogeneous Helmholtz equation,

$$\Delta u(\mathbf{x}) + \frac{\omega^2}{c^2} u(\mathbf{x}) = 0.$$

Under certain circumstances, it is also possible for Maxwell's equations and the time-harmonic elastic wave equation (often called the Navier equation) to be reduced to the Helmholtz equation. Given the central role the Helmholtz equation plays in wave propagation problems, a lot of research effort has been applied to studying the properties of its solutions, and designing methods for computing them efficiently.

2.1.1 Boundary Conditions

To ensure that the Helmholtz problem is *well posed*—that is, it has a unique and stable solution—appropriate boundary conditions must be specified. These conditions are essential both for theoretical analysis and numerical simulation.

In this work, we focus on two types of boundary conditions: Dirichlet and impedance. The selection of appropriate boundary conditions can significantly affect the problem definition and the results obtained.

Dirichlet Boundary Conditions

Dirichlet boundary conditions prescribe the value of the solution u on the boundary $\partial\Omega$ of a bounded domain $\Omega \subset \mathbb{R}^d$. For the Helmholtz equation, the inhomogeneous Dirichlet boundary condition takes the form

$$u(\mathbf{x}) = g(\mathbf{x}) \quad \forall \mathbf{x} \in \partial\Omega,$$

where $g \in H^{1/2}(\partial\Omega)$ is a boundary function specifying the trace of the solution on the boundary.

The space $H^{1/2}(\partial\Omega)$ is a Sobolev space defined on the boundary $\partial\Omega$. It arises naturally from the *trace theorem*, which states that if $u \in H^1(\Omega)$, then the restriction (or trace) of u to the boundary $\partial\Omega$ is well defined in $H^{1/2}(\partial\Omega)$. More precisely, there exists a continuous and surjective linear operator

$$\gamma : H^1(\Omega) \to H^{1/2}(\partial\Omega), \qquad \gamma(u) = u|_{\partial\Omega},$$

such that

$$\|\gamma(u)\|_{H^{1/2}(\partial\Omega)} \leq C\|u\|_{H^1(\Omega)}.$$

Thus, $H^{1/2}(\partial\Omega)$ represents precisely the class of boundary data that can be realized as the trace of an $H^1(\Omega)$ function. In particular, requiring $g \in H^{1/2}(\partial\Omega)$ guarantees that the boundary condition is compatible with the variational formulation of the Helmholtz problem, where the solution space is $H^1(\Omega)$. The homogeneous case corresponds to $g(\mathbf{x}) = 0$, enforcing that the field vanishes on $\partial\Omega$.

Physical interpretation In physical terms, Dirichlet boundary conditions are used to model scenarios in which the wave field is fixed or externally controlled on the boundary. For instance, in acoustics, specifying u on $\partial\Omega$ corresponds to prescribing the acoustic pressure, while in electromagnetic scattering, it may represent a perfectly conducting boundary in scalar approximations. This type of boundary condition is essential in the variational formulation, as it directly constrains the admissible solution space.

Analytical benefits For the majority of the theoretical developments in this work, we impose Dirichlet boundary conditions on the computational domain. This choice is motivated by several analytical simplifications it affords. In particular, Dirichlet boundary conditions eliminate the complex-valued boundary contributions that arise in more general settings, such as those with impedance or Robin-type conditions. As a result, the solution to the corresponding boundary value problem remains purely real when the source term is real, and the wave number is considered in its real-valued, constant form.

This restriction simplifies the analytic treatment of the problem, particularly when deriving variational formulations, establishing a priori estimates, and studying the spectral properties of the underlying operators. Moreover, the Dirichlet boundary condition ensures that all complexities arise from the indefiniteness of the Helmholtz operator itself, rather than from non-selfadjoint boundary effects. This allows for a more focused investigation of the challenges inherent in indefinite, non-coercive partial differential equations without additional complications arising from complex boundary interactions.

Impedance Boundary Conditions

Impedance boundary conditions are a specific example of the more general Robin boundary conditions. They provide a way to approximate radiation behaviour in wave propagation problems by imposing a linear relation between the field and its normal derivative on the boundary. For a finite computational domain, these boundary conditions are typically applied on an artificial boundary, $\partial\Omega$, that truncates an unbounded physical domain. The inhomogeneous impedance boundary condition takes the form

$$\frac{\partial u(\mathbf{x})}{\partial n} + iku(\mathbf{x}) = g(\mathbf{x}) \quad \forall \mathbf{x} \in \partial\Omega,$$

where $k > 0$, $\partial/\partial n$ denotes the outward normal derivative on $\partial\Omega$, and $g \in H^{-1/2}(\partial\Omega)$ is a given boundary function.

The space $H^{-1/2}(\partial\Omega)$ is the *dual space* of $H^{1/2}(\partial\Omega)$. We can recall that $H^{1/2}(\partial\Omega)$ is the trace space of $H^1(\Omega)$ functions, i.e. the space of all boundary values that can arise from interior

H^1 functions. The dual space $H^{-1/2}(\partial\Omega)$ therefore consists of all continuous linear functionals acting on $H^{1/2}(\partial\Omega)$.

Elements of $H^{-1/2}(\partial\Omega)$ should be interpreted as *generalised boundary data* that can be paired with $H^{1/2}(\partial\Omega)$ functions through the duality pairing

$$\langle g, v\rangle_{H^{-1/2},\, H^{1/2}}, \qquad g \in H^{-1/2}(\partial\Omega),\ v \in H^{1/2}(\partial\Omega).$$

This functional-analytic framework is essential for the weak formulation of boundary value problems. In particular:

- The term $u|_{\partial\Omega}$ belongs to $H^{1/2}(\partial\Omega)$ whenever $u \in H^1(\Omega)$ by the trace theorem.
- The normal derivative $\partial u/\partial n$ is only well defined in a weak sense for $u \in H^1(\Omega)$; in fact, $\partial u/\partial n \in H^{-1/2}(\partial\Omega)$.

Thus, requiring the boundary data g to lie in $H^{-1/2}(\partial\Omega)$ ensures that the right-hand side of the impedance condition can be consistently interpreted in the variational formulation. The case $g(\mathbf{x}) = 0$ corresponds to the homogeneous impedance boundary condition, which is commonly used as an absorbing boundary condition to simulate outgoing waves.

Physical interpretation Physically, the impedance condition approximates the Sommerfeld radiation condition at a finite distance and is often interpreted as modelling a partially absorbing boundary. The inhomogeneous form allows for the incorporation of incoming wave components or externally imposed sources on the boundary. This type of condition is especially relevant in scattering problems, where it is used to model the effect of an incident field impinging on the domain through the artificial boundary.

Numerical advantages Impedance boundary conditions will be employed exclusively in the numerical experiments presented in this work. This choice reflects their practical significance in computational wave propagation problems, particularly in simulating open or unbounded domains. By incorporating an absorbing term on the boundary, impedance conditions approximate the effect of radiation to infinity, thereby reducing artificial reflections that would otherwise contaminate the solution due to truncation of the physical domain. From a numerical perspective, these conditions are preferable in finite-domain simulations of scattering and resonance problems, where realistic wave behaviour near artificial boundaries is essential.

Analytical challenges However, the boundary conditions introduce additional analytical complications not present under Dirichlet conditions. In particular, the impedance boundary condition involves a complex-valued term proportional to the solution on the boundary, rendering the problem non-selfadjoint even when the source and wave number are real. This complex boundary interaction leads to complex-valued solutions and makes the spectral analysis of the associated operators more intricate. For this reason, we refrain from using impedance boundary conditions in the theoretical sections of this work. By reserving them for numerical experiments, we are able to maintain a sharper analytical focus on the indefinite and non-coercive nature of the Helmholtz operator itself, while still demonstrating the practical computational settings where realistic boundary modelling is essential.

Perfectly Matched Layer (PML) Boundary Treatment

Perfectly matched layers (PMLs) are a widely used technique for approximating wave propagation in unbounded domains. Strictly speaking, a PML is not a boundary condition on the physical boundary $\partial\Omega$. Instead, it consists of surrounding the physical domain Ω by an artificial absorbing layer Ω_{PML}, thereby enlarging the computational domain to

$$\widetilde{\Omega} = \Omega \cup \Omega_{\mathrm{PML}}.$$

Within this layer, the Helmholtz equation is modified through a complex coordinate stretching designed so that outgoing waves are attenuated rapidly without producing reflections at the interface with the physical domain.

The resulting problem is therefore posed on $\widetilde{\Omega}$, rather than only on Ω, and takes the form of a Helmholtz-type equation with complex-valued coefficients,

$$-\nabla \cdot \big(A(\mathbf{x})\nabla u(\mathbf{x})\big) - k^2 B(\mathbf{x})u(\mathbf{x}) = f(\mathbf{x}) \qquad \forall \mathbf{x} \in \widetilde{\Omega},$$

where the coefficients $A(\mathbf{x})$ and $B(\mathbf{x})$ coincide with those of the original problem inside Ω, and are modified only in the absorbing layer Ω_{PML}. The formulation is then closed by imposing a simple boundary condition on the outer boundary $\partial\widetilde{\Omega}$, most commonly

$$u(\mathbf{x}) = 0 \qquad \forall \mathbf{x} \in \partial\widetilde{\Omega}.$$

Since the PML problem is posed on the enlarged domain, the natural solution space is $H^1(\widetilde{\Omega})$, or $H^1_0(\widetilde{\Omega})$ in the homogeneous Dirichlet case on $\partial\widetilde{\Omega}$. Thus, unlike Dirichlet and impedance boundary conditions, the PML formulation is not expressed in terms of boundary trace data on $\partial\Omega$, but rather through a modification of the differential operator in an exterior layer.

Physical interpretation Physically, the PML acts as an artificial medium that absorbs outgoing waves. Its defining feature is that, at least in the ideal continuous setting, waves pass from the physical domain into the absorbing layer without reflection, and then decay exponentially inside the layer. In this way, the PML mimics radiation to infinity while retaining a finite computational domain.

Numerical advantages PMLs are often more effective than local absorbing boundary conditions when high accuracy is required in finite-domain simulations of open wave problems. They can significantly reduce spurious reflections from artificial boundaries and are therefore widely used in scattering, resonance, and exterior propagation problems.

Analytical challenges The PML formulation introduces complex-valued and, in general, spatially varying coefficients, making the resulting operator non-selfadjoint. In addition, the quality of the approximation depends on the thickness of the layer and the choice of damping profile. Consequently, the analysis is more involved than for standard Dirichlet problems, and the PML should be regarded as a numerical domain truncation strategy rather than as a classical boundary condition.

2.2 Problem Definition — Interior Helmholtz Problem

Having introduced the Helmholtz equation and discussed the role of boundary conditions, we are now ready to formally state the model problem that will serve as the foundation for the analysis and computations in this work. Our focus will be on a *heterogeneous interior Helmholtz problem*—a version of the equation posed on a bounded domain, where the medium may have spatially varying properties. We adopt the strong formulation of the problem as a starting point, and then transition to its variational (weak) form to enable the use of functional analysis and finite element discretisation. In doing so, we make precise assumptions on the coefficients and data that guarantee the existence and uniqueness of the solution under appropriate conditions. This model problem forms the analytical backbone of the methods and results developed in the remainder of the thesis.

Strong Formulation Let $\Omega \subset \mathbb{R}^d$, with $d = 2$ or 3, be a bounded Lipschitz domain. The model boundary value problem is: find a function $u : \Omega \to \mathbb{R}$ such that

$$-\nabla \cdot (A\nabla u) - k^2 u = f \qquad \text{in } \Omega, \tag{2.2.1a}$$

$$u = 0 \qquad \text{on } \partial\Omega. \tag{2.2.1b}$$

Here:

- $A : \Omega \to \mathbb{R}^{d\times d}$ is a matrix-valued function modelling spatial heterogeneity,
- $k > 0$ is the (constant) wavenumber,
- $f : \Omega \to \mathbb{R}$ is a prescribed source term.

The homogeneous Dirichlet condition (2.2.1b) imposes that the field vanishes on the boundary. Physically, this models a perfectly absorbing or sound-soft interface. Mathematically, it ensures that the associated differential operator is self-adjoint under suitable assumptions.

Remark 2.2.1. *It is also possible to write* (2.2.1) *as*

$$-\nabla \cdot (A\nabla u) - k^2 n u = f \qquad \textit{in } \Omega,$$

$$u = 0 \qquad \textit{on } \partial\Omega.$$

Here, n is an almost everywhere positive $L^\infty(\Omega)$ function. This addition allows for more generalised applications of (2.2.1).

Terminology. In wave propagation and scattering, "Helmholtz" often refers to unbounded-domain problems with Sommerfeld radiation conditions (or absorbing conditions on truncated domains). Here the model is posed on a bounded domain with homogeneous Dirichlet boundary conditions. We retain the Helmholtz terminology to emphasise the indefinite structure and the associated solver difficulties at high frequency.

Assumption 2.2.2 (Regularity and Ellipticity)**.** *The matrix A and the source f satisfy:*

(i) *$A(\mathbf{x})$ is symmetric and satisfies the uniform ellipticity condition: there exist constants $0 < a_{\min} \le a_{\max} < \infty$ such that*

$$a_{\min}|\boldsymbol{\xi}|^2 \le A(\mathbf{x})\boldsymbol{\xi} \cdot \boldsymbol{\xi} \le a_{\max}|\boldsymbol{\xi}|^2 \qquad \textit{for almost every } \mathbf{x} \in \Omega,\ \forall \boldsymbol{\xi} \in \mathbb{R}^d.$$

(ii) *The source term $f \in L^2(\Omega)$.*

(iii) *Without loss of generality, we assume $a_{\min} = 1$, and that the domain diameter $D_\Omega := \sup_{\mathbf{x},\mathbf{y}\in\Omega} |\mathbf{x} - \mathbf{y}|$ satisfies $D_\Omega \le 1$. This can always be achieved via rescaling.*

Variational (Weak) Formulation The problem is currently stated in the strong form, and will need to be recast into a variational formulation, otherwise known as the weak formulation. To derive the variational form, we introduce the solution space

$$V = \{v \in H^1(\Omega) \mid v = 0 \text{ on } \partial\Omega\} = H_0^1(\Omega).$$

Testing equation (2.2.1) against an arbitrary function $v \in V$, we obtain

$$\int_\Omega \left(-\nabla \cdot (A\nabla u) - k^2 u\right) \overline{v}\, d\mathbf{x} = \int_\Omega f\overline{v}\, d\mathbf{x}.$$

Applying integration by parts to the first term and using the boundary condition $u = 0$ on $\partial\Omega$, we obtain the variational form:

$$\int_\Omega A\nabla u \cdot \nabla \overline{v}\, d\mathbf{x} - k^2 \int_\Omega u\overline{v}\, d\mathbf{x} = \int_\Omega f\overline{v}\, d\mathbf{x} \quad \forall v \in V.$$

Defining the bilinear form $b : V \times V \to \mathbb{R}$ and linear functional $F : V \to \mathbb{R}$ as

$$b(u,v) := \int_\Omega A\nabla u \cdot \nabla \overline{v} - k^2 u\overline{v}\, d\mathbf{x}, \qquad F(v) := \int_\Omega f\overline{v}\, d\mathbf{x}, \tag{2.2.3}$$

the variational problem becomes:

$$\begin{aligned} &\text{Find } u \in V \text{ such that} \\ &b(u,v) = F(v) \quad \forall v \in V. \end{aligned} \tag{2.2.4}$$

We also introduce the elliptic bilinear form $a : V \times V \to \mathbb{R}$:

$$a(u,v) := \int_\Omega A\nabla u \cdot \nabla \overline{v}\, d\mathbf{x},$$

and define the k-weighted inner product and norm:

$$(u,v)_{1,k} := a(u,v) + k^2(u,v), \qquad \|u\|_{1,k} := \sqrt{(u,u)_{1,k}}.$$

Well-Posedness and Spectral Stability When using Dirichlet boundary conditions, the well-posedness of the boundary value problem depends on the value of k^2, and its relation to the eigenvalues λ_ℓ of $a(\cdot,\cdot)$. Since the embedding of $H_0^1(\Omega)$ in $L^2(\Omega)$ is compact, the associated elliptic eigenvalue problem admits a sequence of eigenpairs $\{(p_\ell, \lambda_\ell)\}_{\ell\in\mathbb{N}}$, with $p_\ell \in H_0^1(\Omega)$, $p_\ell \neq 0$, and $\lambda_\ell \in \mathbb{R}$, satisfying

$$a(p_\ell, v) = \lambda_\ell (p_\ell, v) \qquad \forall v \in H_0^1(\Omega).$$

The eigenfunctions $\{p_\ell\}_{\ell\in\mathbb{N}}$ form an orthogonal basis of $L^2(\Omega)$, and also provide a basis of $H_0^1(\Omega)$ in the usual spectral sense. Moreover, the eigenvalues are of finite multiplicity and satisfy $\lambda_\ell \to +\infty$ as $\ell \to \infty$. The eigenfunctions can be chosen to be orthonormal with respect to $(\cdot,\cdot)$ and the eigenvalues ordered such that

$$\lambda_1 \leq \lambda_2 \leq \cdots .$$

Since $\{p_\ell\}_{\ell\in\mathbb{N}}$ is a basis of $H_0^1(\Omega)$, equation (2.2.4) is equivalent to

$$a(u, p_\ell) - k^2(u, p_\ell) = (f, p_\ell) \qquad \text{for all } \ell \in \mathbb{N}. \tag{2.2.5}$$

Let us expand $u \in H_0^1(\Omega)$:

$$u = \sum_{j=1}^{\infty} \alpha_j p_j \qquad \text{with} \quad \alpha_j \in \mathbb{R} \quad \text{such that} \quad \|u\|^2 = \sum_{j=1}^{\infty} \alpha_j^2 < \infty. \tag{2.2.6}$$

Then the left-hand side of (2.2.5) can be rewritten as

$$\begin{aligned} a(u,p_\ell) - k^2(u,p_\ell) &= \sum_{j=1}^{\infty} \alpha_j a(p_j,p_\ell) - k^2 \sum_{j=1}^{\infty} \alpha_j (p_j,p_\ell) \\ &= \sum_{j=1}^{\infty} \alpha_j \lambda_j (p_j,p_\ell) - k^2 \sum_{j=1}^{\infty} \alpha_j (p_j,p_\ell) = (\lambda_\ell - k^2)\alpha_\ell. \end{aligned}$$

Hence (2.2.5) is equivalent to

$$(\lambda_\ell - k^2)\alpha_\ell = (f, p_\ell) \qquad \text{for all } \ell \in \mathbb{N}. \tag{2.2.7}$$

If $k^2 = \lambda_{\ell_0}$ for some $\ell_0 \in \mathbb{N}$, then (2.2.7) has no solution if $(f, p_{\ell_0}) \neq 0$ (the latter is the case, e.g. when $f = p_{\ell_0}$) and hence (2.2.4) has no solution (moreover, if $(f, p_{\ell_0}) = 0$ and (2.2.4) has a solution, it is not unique). On the other hand, if $k^2 \neq \lambda_\ell$ for all $\ell \in \mathbb{N}$, then we can solve (2.2.7) uniquely for α_ℓ, and therefore (2.2.4) has a unique solution. Let us consider this situation in more detail in the following proposition. This is a common situation, as discussed in [42, Section 6] and [5]. However, because the required conditions depend on the specific problem being considered, the constant $C_{\text{stab}} > 0$ and the associated proof have been adapted to the present setting.

Proposition 2.2.3 (Stability Estimate)**.** *For any $f \in L^2(\Omega)$, assume that $k^2 \notin \{\lambda_\ell : \ell \in \mathbb{N}\}$. Then*

$$C_{\text{stab}} := \sup_{\ell \in \mathbb{N}} \frac{\sqrt{\lambda_\ell + k^2}}{|\lambda_\ell - k^2|} < \infty, \tag{2.2.8}$$

and the unique solution $u \in H_0^1(\Omega)$ to the variational problem (2.2.4) satisfies

$$\|u\|_{1,k} \leq C_{\text{stab}} \|f\|. \tag{2.2.9}$$

Proof. Let $\ell_0 \in \mathbb{N} \cup \{0\}$ be such that $k^2 \in (\lambda_{\ell_0}, \lambda_{\ell_0+1})$, where we set $\lambda_0 := 0$ for convenience. The function

$$g(x) := \frac{\sqrt{x + k^2}}{|x - k^2|}, \qquad x \in [0, \infty) \setminus \{k^2\},$$

is increasing on $[0, k^2)$ and decreasing on (k^2, ∞) and hence

$$C_{\text{stab}} = \sup_{\ell \in \mathbb{N}} \frac{\sqrt{\lambda_\ell + k^2}}{|\lambda_\ell - k^2|} \leq \sup_{x \in [0, \lambda_{\ell_0}] \cup [\lambda_{\ell_0+1}, \infty)} g(x) = \max\{g(\lambda_{\ell_0}), g(\lambda_{\ell_0+1})\} < \infty.$$

Now let u be the solution of (2.2.4) and expand u as in (2.2.6). Then $\alpha_\ell = (f, p_\ell)/(\lambda_\ell - k^2)$ by (2.2.7) and hence

$$\begin{aligned}
\|u\|_{1,k}^2 &= a(u,u) + k^2(u,u) = a\left(\sum_{\ell=1}^{\infty} \alpha_\ell p_\ell, \sum_{j=1}^{\infty} \alpha_j p_j\right) + k^2\left(\sum_{\ell=1}^{\infty} \alpha_\ell p_\ell, \sum_{j=1}^{\infty} \alpha_j p_j\right) \\
&= \sum_{\ell=1}^{\infty} (\lambda_\ell + k^2)\alpha_\ell^2 = \sum_{\ell=1}^{\infty} (\lambda_\ell + k^2) \frac{(f, p_\ell)^2}{(\lambda_\ell - k^2)^2} \\
&\leq \sup_{j \in \mathbb{N}} \left[\frac{\lambda_j + k^2}{(\lambda_j - k^2)^2}\right] \sum_{\ell=1}^{\infty} (f, p_\ell)^2 = C_{\text{stab}}^2 \|f\|^2,
\end{aligned}$$

which proves (2.2.9). □

We see from (2.2.8) that $C_{\text{stab}} \to \infty$ as k^2 approaches an eigenvalue λ_ℓ.

Now that we have the weak formulation of the problem, we are able to derive the finite element discretisation to carry out the numerical calculations.

2.3 Finite Element Discretisation and Solution Methods

The variational formulation developed in the previous section provides a rigorous mathematical setting for the Helmholtz equation. However, to approximate solutions in practice, we must discretise this formulation using numerical methods. Among various options, the finite element method (FEM) is especially well suited for problems posed on complex geometries and with heterogeneous coefficients. This section presents the finite element discretisation of the Helmholtz problem and discusses the structure and numerical challenges of the resulting linear system.

2.3.1 Discretisation via Finite Elements

To discretise the variational problem (2.2.4), we first construct a finite-dimensional subspace $V^h \subset V = H_0^1(\Omega)$ tailored to the geometry of the domain. Let $\Omega \subset \mathbb{R}^d$ be a polygonal (or polyhedral) domain, and let $\mathcal{T}_h$ be a conforming, shape-regular, quasi-uniform mesh of Ω, consisting of N_K elements K_i and N nodes. The mesh parameter h denotes the maximum diameter of the elements:

$$h = \max_{K \in \mathcal{T}_h} \operatorname{diam}(K).$$

We briefly recall the meaning of the mesh assumptions used above.

- **Conforming mesh:** The mesh $\mathcal{T}_h$ is called conforming if the intersection of any two elements $K_i, K_j \in \mathcal{T}_h$ is either empty, a common vertex, a common edge (in 2D), or a common face (in 3D). In other words, there are no "hanging nodes" or partial overlaps, so that the finite element functions are globally continuous, ensuring that $V^h \subset H_0^1(\Omega)$.

- **Shape-regular mesh:** Shape regularity requires that all elements of $\mathcal{T}_h$ avoid degeneracy as $h \to 0$. Formally, there exists a constant $C_K > 0$, independent of h, such that for every element $K \in \mathcal{T}_h$,

  $$\frac{\rho_K}{h_K} \geq C_K,$$

  where $h_K = \operatorname{diam}(K)$ and ρ_K is the diameter of the largest ball contained in K. This condition prevents elements from becoming arbitrarily thin or flat, which is a requirement for stability and approximation estimates.

- **Quasi-uniform mesh:** A mesh $\mathcal{T}_h$ is quasi-uniform if there exists a constant $C_Q > 0$ such that

  $$\max_{K \in \mathcal{T}_h} h_K \leq C_Q \min_{K \in \mathcal{T}_h} h_K.$$

  In other words, the element sizes are comparable across the mesh, with no extreme variations.

We define the discrete space V^h as:

$$V^h = \{v_h \in C^0(\overline{\Omega}) \mid v_h|_K \in \mathbb{P}_p(K) \text{ for all } K \in \mathcal{T}_h\},$$

where $\mathbb{P}_p(K)$ denotes the space of polynomials of degree at most $p \geq 1$ on each element K and $C^0(\overline{\Omega})$ is the space of continuous functions over $\overline{\Omega}$. To respect the homogeneous Dirichlet condition imposed in the continuous space $V = H_0^1(\Omega)$, the discrete space must also vanish on the boundary. This is achieved by restricting V^h to the subspace V_0^h, defined as

$$V_0^h = \{v_h \in V^h \mid v_h|_{\partial\Omega} = 0\}.$$

It is then the V_0^h space that is used as the solution and test space.

Let $\{\varphi_i\}_{i=1}^N$ be a basis of V_0^h. We seek an approximate solution $u_h \in V_0^h$ of the form:

$$u_h = \sum_{i=1}^{N} \mathbf{u}_i \varphi_i,$$

where the coefficients $\mathbf{u} = (\mathbf{u}_1, \ldots, \mathbf{u}_N)^\top$ solve a discrete linear system. The discrete version of the variational formulation (2.2.4) is: Find $u_h \in V_0^h$ such that

$$b(u_h, v_h) = F(v_h), \quad \forall v_h \in V_0^h. \tag{2.3.1}$$

The discrete formulation (2.3.1), also known as a Galerkin equation, can be reformulated into a system of linear equations.

Matrix Formulation Substituting u_h and the test functions $v_h = \varphi_j$ into the bilinear form (2.2.3) yields the linear system:

$$\mathbf{B}\mathbf{u} = \mathbf{f}, \tag{2.3.2}$$

where:

- $\mathbf{B} = \mathbf{A} - k^2\mathbf{S}$ is the Helmholtz system matrix,
- $\mathbf{A}_{ij} = a(\varphi_j, \varphi_i) = \int_\Omega A\nabla\varphi_j \cdot \nabla\overline{\varphi_i}\, d\mathbf{x}$ are the entries of the stiffness matrix,
- $\mathbf{S}_{ij} = (\varphi_j, \varphi_i) = \int_\Omega \varphi_j\overline{\varphi_i}\, d\mathbf{x}$ are the entries of the mass matrix,
- $\mathbf{f}_i = \int_\Omega f\overline{\varphi_i}\, d\mathbf{x}$ denote the entries of the right-hand side vector.

It is also important to note that the matrix $\mathbf{B}$ is sparse. The entries

$$\mathbf{A}_{ij} = a(\varphi_j, \varphi_i) \quad \text{and} \quad \mathbf{S}_{ij} = (\varphi_j, \varphi_i)$$

are non-zero only when the supports of φ_i and φ_j overlap. Since standard finite element basis functions have local support, each basis function interacts with only a small number of neighbouring basis functions. As a result, each row of $\mathbf{B}$ contains only a limited number of non-zero entries. In particular, for a fixed polynomial degree on a shape-regular mesh, the number of non-zero entries is typically $\mathcal{O}(N)$.

The linear system (2.3.2) is difficult to solve because the matrix $\mathbf{B}$ is indefinite and has the potential to be large. In what follows we will discuss each of these properties in turn, outlining why the linear system has these properties, and how these properties affect the numerical solution.

2.3.2 Numerical Challenges of Solving the Discrete System

The matrix system (2.3.2) presents two primary challenges: it is both *indefinite* and *potentially large*—issues that are exacerbated at high frequencies.

Indefiniteness. The system matrix $\mathbf{B}$ arises from subtracting a scaled mass matrix from the stiffness matrix. While both $\mathbf{A}$ and $\mathbf{S}$ are symmetric positive definite, the combination $\mathbf{B} = \mathbf{A} - k^2\mathbf{S}$ may become indefinite or even singular for certain values of k. As k increases, the positive eigenvalues of $\mathbf{A}$ get shifted by the k^2 to the left. This shift causes the eigenvalues of $\mathbf{B}$ to become negative or approach zero, reflecting a loss of coercivity in the discrete system—an echo of the spectral behaviour described in Proposition 2.2.3. The larger the k term, the greater the shift to the left, the more negative eigenvalues are contained in the spectrum.

System Size and Resolution Requirements. The number of degrees of freedom N depends on the mesh size h and the desired resolution. Accurately resolving a wave field with wavelength $\lambda = 2\pi/k$ requires a minimum number of points per wavelength, typically $n_{\text{ppw}} \approx 10$:

$$n_{\text{ppw}} = \frac{2\pi}{hk} \approx \text{constant}.$$

This implies that $h = \mathcal{O}(1/k)$. High-frequency problems therefore require very fine meshes and lead to very large linear systems.

An illustration demonstrating the importance of having adequate points per wavelength can be seen in Figure 2.1, where we see the interpolation error grows if the number of degrees of freedom is not increased with k. For piecewise linear finite elements, the choice $n_{\text{ppw}} = 10$ has been recommended in [8, 63] as a suitable condition. In particular, as k increases, the required mesh refinement (i.e., decrease in h) leads to a rapidly growing number of degrees of freedom This leads to the accurate interpolation of an oscillatory function, or to a good approximation in the finite element space space V^h, meaning that the error between the true and approximate solutions is controlled by n_{ppw}. This in itself can lead to large systems for high values of k.

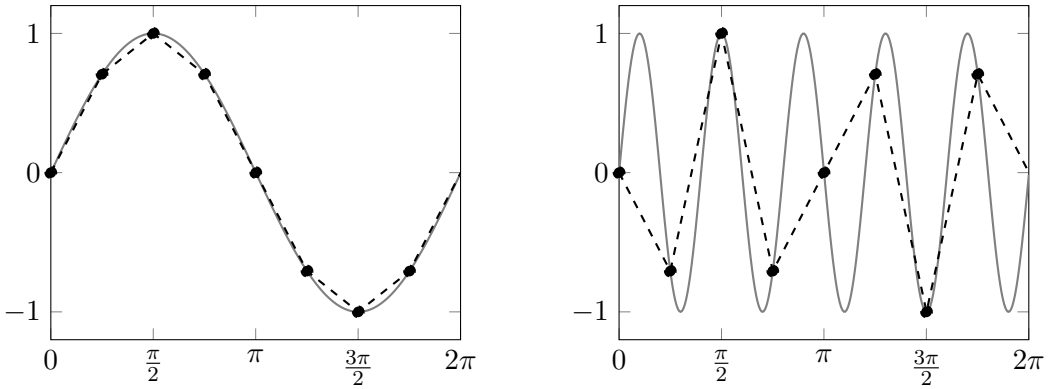


Figure 2.1: The interpolation error increases if the mesh is not refined with increasing frequency. The left-hand plot shows the interpolant of $\sin(x)$ on a mesh with 8 points per wavelength. The right-hand plot shows the interpolant of $\sin(5x)$ on the same mesh.

The Pollution Effect. When solving the Helmholtz equation numerically, the classical mesh resolution criterion, requiring a fixed number of points per wavelength, is not sufficient to guarantee an accurate finite element approximation as $k \to \infty$. This inadequacy arises from the so-called pollution effect, which manifests as a phase error that grows with increasing k, even when the mesh is refined according to the standard resolution condition. The key issue is that the finite element solution $u_h \in V^h$ must not only be close to the exact solution u, but must also be close to the best approximation in the finite element space V^h. The finite element solution and the best approximation in the finite element space V^h are connected by Céa's lemma [16]. It provides an estimate of the form

$$\|u - u_h\| \leq C \inf_{v_h \in V^h} \|u - v_h\|$$

where u is the exact solution, u_h is the finite element solution and $\|\cdot\|$ denotes the standard $L^2(\Omega)$ norm. For coercive elliptic problems, such as Poisson's equation, the constant C is independent of the mesh size h and of any parameters in the PDE. However, for the indefinite Helmholtz problem, C depends on the wave number k, and in fact grows with k even if kh is kept constant. This loss of quasi-optimality was first rigorously analysed in [2], where it was shown that for standard $\mathbb{P}_1$ finite elements, the finite element solution suffers from significant dispersion error—a discrepancy in the phase velocity of the numerical wave compared to the exact solution. The total finite element error may be decomposed as:

$$\|u - u_h\| \leq \underbrace{\|u - u^*\|}_{\text{Approximation error}} + \underbrace{\|u^* - u_h\|}_{\text{Pollution error}}$$

where $u^* \in V^h$ is the best approximation to u in the discrete space. The pollution error is not controlled solely by mesh resolution (i.e., number of points per wavelength), but by deeper properties of the discrete operator—most notably, its numerical dispersion relation. In the one-dimensional setting, a classical result (cf. [2, 63]) shows that the pollution error behaves like:

$$\|u - u_h\| \leq C_1 kh + C_2 k^3 h^2, \quad \text{for } kh < 1$$

where C_1, $C_2 > 0$ independent of k and h. The k^3h^2 term dominates in the high-frequency regime and represents the polluting term due to numerical dispersion. To ensure a bounded total error as $k \to \infty$, it is necessary to enforce:

$$k^3 h^2 \leq C,$$

for some constant $C > 0$. In other words, one has to choose h such that $h \lesssim k^{-\frac{3}{2}}$, which shows that the mesh must be refined more aggressively than the standard $kh = \text{const}$ condition would suggest. For higher finite elements, i.e. $\mathbb{P}_p$, we would need $h^p k^{p+1} = \text{const}$ as shown in [73]. An illustration of the pollution effect can be seen in Figure 2.2. The pollution effect is a significant difference from traditional elliptic error estimates and highlights the special challenges posed by Helmholtz-type problems.

The pollution effect is fundamentally tied to the dispersive nature of the discrete Helmholtz operator. The continuous Helmholtz equation admits plane wave solutions with a fixed phase velocity $c = \omega/k$. However, the discrete system yields solutions with wave numbers $k_h \neq k$, which leads to a numerical phase velocity $c_h = \omega/k_h$ differing from the physical one. The error $\|k_h - k\|$ reflects the numerical dispersion, and this deviation introduces a cumulative phase shift that grows with k, manifesting as the pollution effect. In higher-dimensional settings (2D, 3D), the pollution effect persists and often worsens due to more complex geometric dispersion.

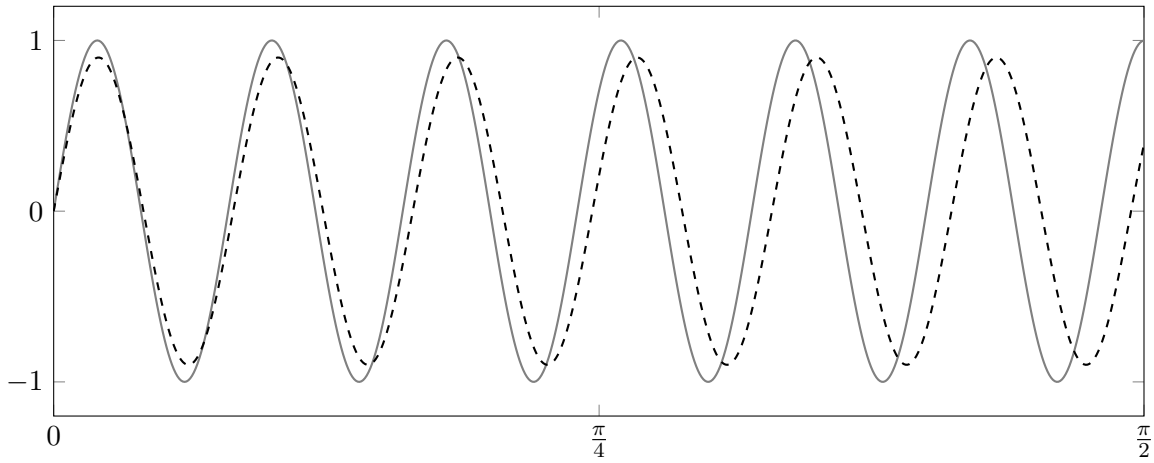


Figure 2.2: An example of the pollution effect. The mesh is fine enough to eliminate the interpolation error, but the phase of the solution is out of sync with the true solution due to the pollution effect.

Another effect restricting the size of h comes from the fact that, due to the indefiniteness of (2.3.1), the solvability of the discrete system is not guaranteed unless h is small enough. The following lemma from [82, Theorem 2] provides a sufficient condition for the solvability of the discrete system.

Lemma 2.3.1 (Schatz and Wang, 1996)**.** *Let Assumption 2.2.2 and (2.2.9) from Proposition 2.2.3 hold. Then there exists an $h_0 > 0$ such that, for each h with $0 < h < h_0$, the problem (2.3.1) has a unique solution $u_h \in V_H$. Moreover, let u be the unique solution of (2.2.4). Then, for every $\varepsilon > 0$ there exists $h_1 = h_1(\varepsilon) > 0$ such that, for every $h \in (0, h_1)$,*

$$\|u - u_h\| \leq \varepsilon \|u - u_h\|_{H^1(\Omega)} \tag{2.3.3}$$

and

$$\|u - u_h\|_{H^1(\Omega)} \leq \varepsilon \|f\|. \tag{2.3.4}$$

In summary, numerically solving the Helmholtz equation with a suitable level of accuracy requires a large (size at least $\sim k^d$) linear system, where the size increases as k increases.

2.3.3 Solving the Discrete System

The linear system (2.3.2) must be solved efficiently despite its size and indefiniteness. Two general classes of solvers are commonly considered:

Direct Solvers. Methods such as Gaussian elimination, multifrontal solvers, and nested dissection [34, 51] provide exact solutions (up to machine precision). However, their cost scales poorly with system size. For dense matrices, this is typically $\mathcal{O}(N^3)$. The discrete systems arising in this work produce sparse matrices, whose storage cost is governed by the number of non-zero entries rather than by the total number of matrix entries, so depending on the space dimension this cost can be reduced. Even so, the fill-in generated during factorisation leads to significant growth in both storage and computational cost. As the problem size increases, memory requirements can still become prohibitive at high frequencies. Techniques based on hierarchical matrix formats (e.g., $\mathcal{H}$-matrices or HSS) [93, 67] offer improvements but are limited in scalability and applicability. Moreover, these methods run into difficulties for higher values of k and have domain specific requirements that need to be met. Therefore, due to the size of the linear system, direct solvers are not considered feasible for many situations.

Iterative Solvers. In this work, we focus primarily on solving the Helmholtz problem using iterative methods. An iterative method to solve the $N \times N$ linear system (2.3.2) is a method that generates a sequence of approximations $\mathbf{u}^0, \mathbf{u}^1, \mathbf{u}^2, \ldots$ to the solution $\mathbf{u}$, with the aim that this sequence converges to $\mathbf{u}$. Not all iterative methods converge, however, and even convergent methods may do so too slowly to be practical. Each $\mathbf{u}^i$ is obtained from the previous iterate(s) by some multiplications with $\mathbf{B}$. Since the coefficient matrix is sparse, it is desirable to use methods that exploit this structure. In particular, sparse matrix–vector multiplication is one of the least expensive fundamental operations available for such matrices, both in terms of storage access and arithmetic cost. It would therefore be advantages to utilise methods that take advantage of such operations. Krylov subspace methods such as GMRES are the best candidates for the solution of large sparse systems, particularly when the matrix-vector product can be computed efficiently. However, convergence may be slow or stagnate entirely without suitable *preconditioning*—especially for highly indefinite problems like Helmholtz.

In this work, we focus on iterative methods paired with problem-specific preconditioners, which are essential for practical large-scale simulations. The next section provides a detailed overview of Krylov methods and preconditioning strategies for the Helmholtz equation.

2.4 Krylov Iterative Methods and Preconditioning

This section is going to briefly outline the general Krylov subspace methods, before moving onto the GMRES method specifically.

Krylov subspace methods Let $\mathbf{C} \in \mathbb{C}^{N\times N}$ be a non-singular matrix, $\mathbf{b} \in \mathbb{C}^N$ be a given vector and $\mathbf{x} \in \mathbb{C}^N$ be such that we have the linear system

$$\mathbf{C}\mathbf{x} = \mathbf{b}.$$

Krylov subspace methods are non-stationary iterative methods that aim to construct a sequence of vectors $\mathbf{x}_\ell$ that approximate the solution $\mathbf{x}$. The methods start with an initial guess, $\mathbf{x}_0$; then the subsequent approximation vectors are given as $\mathbf{x}_{\ell+1} = \mathbf{x}_\ell + \alpha_\ell \mathbf{p}_\ell$ with search direction $\mathbf{p}_\ell$ and constant $\alpha_\ell > 0$. This means that $\mathbf{x}_{\ell+1} - \mathbf{x}_0 \in \text{span}\{\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_\ell\}$, where $\text{span}\{\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_\ell\}$ is known as the search space. What differentiates Krylov subspace methods from other iterative methods, is the way that the search space is defined. Many Krylov methods construct a basis of this space and then impose an optimality or orthogonality condition to determine the next iterate; however, not all Krylov methods arise from a residual minimisation principle.

Definition 2.4.1 (Krylov subspace)**.** *For the initial guess $\mathbf{x}_0$, let the initial residual be given by $\mathbf{r}_0 = \mathbf{b} - \mathbf{C}\mathbf{x}_0$. The Krylov subspace of order ℓ is then defined as*

$$\mathcal{K}_\ell(\mathbf{C}, \mathbf{r}_0) = \text{span}\{\mathbf{r}_0, \mathbf{C}\mathbf{r}_0, \ldots, \mathbf{C}^{\ell-1}\mathbf{r}_0\}.$$

The sequence of solution vectors satisfy $\mathbf{x}_\ell \in \mathbf{x}_0 + \mathcal{K}_\ell(\mathbf{C}, \mathbf{r}_0)$. From Definition 2.4.1, it is possible to characterize each element of the Krylov subspace as a polynomial of $\mathbf{C}$ being applied to $\mathbf{r}_0$. Krylov subspace methods may be considered as a method of finding a low-dimensional (low compared to N) polynomial whose action approximates the action of $\mathbf{C}^{-1}$ on the vector $\mathbf{b}$ (or equivalently on the residual $\mathbf{r}_0$).

Consider a polynomial space $\mathcal{P}_\ell$ given by

$$\mathcal{P}_\ell = \{\phi(x) = \sigma_0 + \sigma_1 x + \ldots + \sigma_\ell x^\ell \mid \sigma_0, \sigma_1, \ldots, \sigma_\ell \in \mathbb{C}\};$$

then the action of $\mathbf{C}^{-1}$ on $\mathbf{b}$ can be approximated through

$$\mathbf{C}^{-1}\mathbf{b} \approx \phi(\mathbf{C})\mathbf{b}, \quad \text{where } \phi \in \mathcal{P}_{\ell-1}.$$

The Krylov subspace can be written as

$$\mathcal{K}_\ell(\mathbf{C}, \mathbf{r}_0) = \{\phi(\mathbf{C})\mathbf{r}_0 \mid \phi \in \mathcal{P}_{\ell-1}\},$$

giving the solution vector $\mathbf{x}_\ell = \mathbf{x}_0 + \phi(\mathbf{C})\mathbf{r}_0$. In the special case where $\phi(\mathbf{C})\mathbf{r}_0 = \mathbf{C}^{-1}\mathbf{r}_0$, the exact solution is recovered; however, in practice the exact solution may be obtained earlier if the residual lies in a low-dimensional invariant subspace (for example, if $\mathbf{b}$ is a combination of only a few eigenvectors of $\mathbf{C}$).

To demonstrate how to successfully determine the polynomial $\phi(\mathbf{C})$, it is beneficial to look at the operation of the Krylov method in terms of the residual $\mathbf{r}_\ell$. If $\mathbf{v}_\ell \in \mathcal{K}_\ell(\mathbf{C}, \mathbf{r}_0)$, then $\mathbf{v}_\ell$ can be written as $\mathbf{v}_\ell = q_{\ell-1}(\mathbf{C})\mathbf{r}_0$ for a polynomial $q_{\ell-1} \in \mathcal{P}_{\ell-1}$. This means the ℓ-th residual, $\mathbf{r}_\ell$, can be given by

$$\mathbf{r}_\ell = \mathbf{b} - \mathbf{C}\mathbf{x}_\ell = (I - \mathbf{C}q_{\ell-1}(\mathbf{C}))\mathbf{r}_0 = p_\ell(\mathbf{C})\mathbf{r}_0,$$

with $p_\ell(x) = 1 - xq_{\ell-1}(x) \in \mathcal{P}_\ell$ and $p_\ell(0) = 1$. Some Krylov methods are designed so that the residual norm is minimised at each iteration, while others enforce alternative projection or orthogonality conditions. Krylov subspace methods therefore often attempt to construct a polynomial p_ℓ which reduces the residual $\mathbf{r}_\ell$.

GMRES The *Generalized Minimal Residual method* (GMRES), introduced by Saad and Schultz [79], is a widely used Krylov subspace method designed for solving non-symmetric and indefinite linear systems. This makes it particularly suitable for the Helmholtz equation, where standard methods often struggle due to indefiniteness and lack of coercivity.

GMRES builds approximate solutions $\mathbf{x}_\ell$ by searching for the vector in the Krylov subspace $\mathcal{K}_\ell(\mathbf{C}, \mathbf{r}_0)$ that minimizes the residual norm. The algorithm can be conceptually divided into two main steps:

1. **Construction of an orthonormal basis for the Krylov subspace** $\mathcal{K}_\ell(\mathbf{C}, \mathbf{r}_0)$, using the Arnoldi process, which is a modified Gram–Schmidt orthogonalisation applied to the vectors $\mathbf{r}_0, \mathbf{C}\mathbf{r}_0, \ldots$.
2. **Minimisation of the residual** over the affine space $\mathbf{x}_0 + \mathcal{K}_\ell(\mathbf{C}, \mathbf{r}_0)$ at each iteration.

The approximate solution at iteration $\ell + 1$ is given by solving the minimisation problem:

$$\mathbf{x}_{\ell+1} = \mathbf{x}_0 + \underset{\mathbf{y} \in \mathcal{K}_{\ell+1}(\mathbf{C}, \mathbf{r}_0)}{\operatorname{argmin}} \|\mathbf{b} - \mathbf{C}(\mathbf{x}_0 + \mathbf{y})\| = \underset{\mathbf{x} \in \mathbf{x}_0 + \mathcal{K}_{\ell+1}}{\operatorname{argmin}} \|\mathbf{b} - \mathbf{C}\mathbf{x}\|.$$

In this formulation, the new residual is

$$\mathbf{r}_{\ell+1} = \mathbf{b} - \mathbf{C}\mathbf{x}_{\ell+1} = p_{\ell+1}(\mathbf{C})\mathbf{r}_0,$$

for some polynomial $p_{\ell+1} \in \mathcal{P}_{\ell+1}$ satisfying $p_{\ell+1}(0) = 1$.

Polynomial Interpretation. GMRES can be viewed as searching for a polynomial $p_{\ell+1}$ of degree $\ell+1$ such that

$$\mathbf{r}_{\ell+1} = p_{\ell+1}(\mathbf{C})\mathbf{r}_0, \quad \text{where} \quad \|\mathbf{r}_{\ell+1}\| = \min_{p\in\mathcal{P}_{\ell+1},\, p(0)=1} \|p(\mathbf{C})\mathbf{r}_0\|.$$

This formulation highlights the close connection between the convergence behaviour and the spectrum or numerical range of $\mathbf{C}$.

Convergence Analysis. A common way to study GMRES convergence is through the estimate

$$\frac{\|\mathbf{r}_{\ell+1}\|}{\|\mathbf{r}_0\|} \leq \min_{p\in\mathcal{P}_{\ell+1},\, p(0)=1} \|p(\mathbf{C})\|,$$

which provides a bound on the residual reduction and illustrates how convergence is influenced by the behaviour of the polynomial $p(\mathbf{C})$ on the spectrum or numerical range of $\mathbf{C}$.

The convergence of GMRES depends heavily on the distribution of the eigenvalues and, more generally, on the field of values of $\mathbf{C}$. For further background and detailed convergence analyses, we refer the reader to [78, 37].

Given our problem definition, it would also be possible to use alternative Krylov subspace iteration methods, such as MINRES. However, even though MINRES and GMRES are equivalent in exact arithmetic for symmetric indefinite matrices, GMRES is more numerically stable as demonstrated in [84]. On the other hand, note that due to the equivalence of MINRES and GMRES when using exact arithmetic under a symmetric positive definite preconditioner, the theory developed here for GMRES is equally applicable to MINRES.

2.4.1 Preconditioning for Krylov Subspace Methods

While Krylov subspace methods such as GMRES are powerful for solving large linear systems, their convergence can be slow—especially when the system matrix $\mathbf{C}$ is ill-conditioned or indefinite. This is particularly true in Helmholtz-type problems, where the indefiniteness grows with the wavenumber k.

To address this, one commonly employs a *preconditioner*: an auxiliary matrix $\mathbf{M}$ that is chosen so that the preconditioned system has more favourable spectral or numerical range properties and for which the action of $\mathbf{M}^{-1}$ can be applied efficiently.

Given the linear system

$$\mathbf{C}\mathbf{x} = \mathbf{b},$$

preconditioning may be applied either on the left or on the right. The *left-preconditioned* system is

$$\mathbf{M}^{-1}\mathbf{C}\mathbf{x} = \mathbf{M}^{-1}\mathbf{b},$$

and a Krylov method applied to this formulation generates approximations from subspaces of the form

$$\mathcal{K}_\ell(\mathbf{M}^{-1}\mathbf{C}, \mathbf{M}^{-1}\mathbf{r}_0),$$

where $\mathbf{r}_0 = \mathbf{b} - \mathbf{C}\mathbf{x}_0$. Alternatively, using *right-preconditioning* gives the system

$$\mathbf{C}\mathbf{M}^{-1}\mathbf{y} = \mathbf{b}, \qquad \mathbf{x} = \mathbf{M}^{-1}\mathbf{y}.$$

The Krylov method is then applied to the system with coefficient matrix $\mathbf{C}\mathbf{M}^{-1}$, generating iterates in the subspace

$$\mathcal{K}_\ell(\mathbf{C}\mathbf{M}^{-1}, \mathbf{r}_0).$$

The preconditioner $\mathbf{M}$ should satisfy two essential properties:

- The preconditioned system should converge in significantly fewer iterations than the original system.
- Applying $\mathbf{M}^{-1}$ (or an approximation thereof) should be computationally inexpensive.

In many situations $\mathbf{M}$ is chosen to approximate $\mathbf{C}$, but this is not strictly necessary. In practice it is often sufficient that the preconditioned operator has favourable spectral properties, such as clustered eigenvalues or a numerical range bounded away from the origin, even if $\mathbf{M}$ does not resemble $\mathbf{C}$ directly.

Preconditioning and the Helmholtz Problem. Specialised preconditioners are required when solving the Helmholtz equation, especially with high wavenumbers, k. Without them, the number of iterations and the cost of computation increases significantly to the point where computing the solution is effectively impossible. There are many subtle reasons why iterative methods, including GMRES, perform poorly when applied to the Helmholtz problem. A comprehensive review of different methods can be found in [40]. One common explanation for the poor convergence when using standard iterative methods is related to the lack of coercivity in the standard variational formulation of the Helmholtz problem [22]. When k is sufficiently large, and taking the bilinear form given by the discretisation of problem (2.3.1), from [44, Lemma 6.2] it is known that there exists $v \in H_0^1(\Omega)$ such that $b(v, v) = 0$. This means that the numerical range could contain the origin. The aim of the preconditioner is to ensure that the linear system is altered so that the origin is no longer contained within the numerical range.

2.4.2 Elman Theory for GMRES Convergence

The convergence of GMRES in the preconditioned setting is often analysed using what is known as *Elman's estimate* [35], which has been adapted to Helmholtz-type problems in [89].

Let $\mathbf{M}$ be a preconditioner and consider the preconditioned matrix $\mathbf{C}' = \mathbf{M}^{-1}\mathbf{C}$. Suppose there exist constants $c_p > 0$ and $C_p > 0$ such that, for all $\mathbf{x} \neq 0$,

$$c_p \leq \frac{\Re(\mathbf{x}^* \mathbf{C}' \mathbf{x})}{\|\mathbf{x}\|^2}, \qquad \text{and} \qquad \|\mathbf{C}'\mathbf{x}\| \leq C_p \|\mathbf{x}\|.$$

These bounds measure how far the field of values of $\mathbf{C}'$ lies from the origin and how spread-out the spectrum is.

Lemma 2.4.2 (Elman Estimate)**.** *Let $c_p > 0$ and $C_p > 0$ be as defined above. Then after ℓ GMRES iterations, the residual satisfies*

$$\|\mathbf{r}_\ell\| \leq \left(1 - \frac{c_p^2}{C_p^2}\right)^{\ell/2} \|\mathbf{r}_0\|.$$

Interpretation. This estimate implies that convergence improves as c_p^2/C_p^2 increases. In other words, GMRES converges more rapidly when the numerical range of the preconditioned operator is bounded away from zero. The main takeaway is that the distance of the numerical range of $\mathbf{M}^{-1}\mathbf{C}$ to the origin and the clustering of its spectrum, will determine the effectiveness of the GMRES methods. Designing preconditioners that achieve this property is central to the robust and scalable solution of Helmholtz problems.

In the next section, we explore several such solver and preconditioning strategies developed specifically for the Helmholtz equation, including multigrid-based and domain decomposition approaches.

Remark 2.4.3 (Left- and right-preconditioning)**.** *In the theoretical analysis, we use left-preconditioning as it is a more natural setting for field-of-values estimates of the type used in Elman's estimate, which fits conveniently with the usual domain decomposition framework. In the numerical experiments, however, we use right-preconditioning, since this is often more convenient in practice: GMRES measures the residual of the original linear system, so the stopping criterion is easier to interpret and implement.*

Although the analysis is written for the left-preconditioned formulation, this does not substantially change the relevance of the bounds for the right-preconditioned method. For additive Schwarz preconditioners, the left- and right-preconditioned formulations are closely related, and the theory developed for the left-preconditioned problem can be transferred to the right-preconditioned setting, at least in an equivalent weighted norm and, more generally, with the same asymptotic convergence behaviour up to norm-equivalence constants. For further discussion, see [81, 86].

2.5 Solver Techniques for the Helmholtz Equation

As discussed in the previous section, preconditioning is essential for the efficient and robust solution of linear systems arising from the Helmholtz equation, especially at high frequencies. Without preconditioning, Krylov subspace methods such as GMRES often suffer from slow or stagnating convergence due to the indefiniteness and poor spectral properties of the discretised operator.

While simple elliptic preconditioners—such as those based on the Laplacian—can offer limited improvements, they are typically inadequate for Helmholtz problems at moderate or high wavenumbers. As a result, a range of more sophisticated strategies have been developed to address the specific challenges of this setting. These approaches aim to balance practical considerations (such as ease of implementation and cost per iteration) with desirable theoretical properties (such as wavenumber-independent convergence).

In the following subsections, we review several classes of solvers that have been proposed for Helmholtz problems:

- multigrid-based approaches,
- complex-shifted Laplacians,
- sweeping strategies, and
- domain decomposition methods.

Each method addresses different aspects of the Helmholtz equation's difficulties and offers distinct trade-offs in terms of robustness, scalability, and generality.

2.5.1 Multigrid Methods

Multigrid methods are a well-established family of solvers for symmetric positive definite (SPD) problems. They combine relaxation on a fine grid with correction steps on a hierarchy of coarser grids. For Helmholtz problems, however, the indefinite and non-Hermitian nature of the operator causes standard multigrid to fail: certain oscillatory error components are not effectively reduced and may even be amplified.

The classical multigrid cycle consists of:

1. *Smoothing*, which dampens high-frequency error components,
2. *Coarse grid correction*, which handles the smooth components.

In Helmholtz problems, the distinction between smooth and oscillatory modes depends on the mesh size h relative to the wavelength $\lambda \sim 1/k$. When $kh \sim 1$, smoothing fails to reduce critical error components, and coarse grids cannot resolve the wave accurately, leading to phase errors and instability. Moreover, the eigenvalues of the fine- and coarse-grid operators may differ significantly in magnitude and sign, impairing the accuracy and direction of the correction.

To mitigate these issues, several adaptations have been proposed:

- **Modified wavenumbers** [41]: Coarse-grid operators use a corrected wavenumber k_c to reduce dispersion.
- **GMRES smoothing** [36]: GMRES replaces traditional relaxation to enhance robustness on coarse levels.

These variants improve performance in some cases but still struggle in general, especially in higher dimensions or for large k.

2.5.2 Shifted Laplacian Preconditioning

A widely used strategy is the *shifted Laplacian preconditioner*, which modifies the Helmholtz operator by adding artificial complex damping:

$$-\Delta u - (k^2 + i\varepsilon)u = f, \qquad \varepsilon > 0.$$

The added imaginary part improves the spectral properties, making the system strongly non-Hermitian and more amenable to multigrid inversion. The modified operator

$$A_\varepsilon := -\Delta - (k^2 + i\varepsilon)$$

is typically used as a left or right preconditioner for the original problem.

Choosing $\varepsilon = \mathcal{O}(k)$ yields nearly k-independent GMRES convergence rates when combined with appropriate multigrid solvers [47]. However, practical implementation involves trade-offs:

- Larger ε improves solver robustness but reduces the quality of the preconditioner.
- Smaller ε improves spectral similarity but can degrade multigrid performance.

Additionally, each application of the preconditioner requires solving a complex-valued PDE, which increases memory usage and implementation complexity. Nevertheless, the approach is conceptually simple and leverages existing multigrid tools.

2.5.3 Sweeping Preconditioners

Sweeping preconditioners fall into the broader class of incomplete factorisation methods. Traditional techniques like incomplete LU (ILU) are not well-suited to Helmholtz problems due to their sensitivity to indefiniteness. Instead, *directional sweeping* has emerged as an effective alternative.

Originally proposed in [38, 39], sweeping preconditioners factorize the Helmholtz operator in a block LDL^T form by sequentially eliminating variables layer by layer. An absorbing boundary condition is imposed on one end of the domain, and unknowns are eliminated in a forward or backward sweep. Key features include:

- Use of hierarchical matrices ($\mathcal{H}$-matrices) to compress Schur complements [4],
- Approximation of Schur complements via localized auxiliary problems with absorbing boundary layers.

Later developments include:

- **Additive sweeping** [69]: Enables parallel computation by decomposing the domain into thin layers with transmission conditions.
- **Multipreconditioned sweeps** [15]: Combine sweeps in multiple directions to enhance robustness and scalability.

These methods offer nearly linear computational complexity and k-independent iteration counts. However, they remain challenging to implement, particularly in three dimensions, where the sequential nature of sweeps can limit parallel scalability.

2.5.4 Conclusion and Motivation for Domain Decomposition

Domain decomposition (DD) preconditioners divide the global problem into overlapping or non-overlapping subdomains and solve local problems in parallel. The subdomain solutions are then combined using interface conditions and a coarse global correction.

Each of the preconditioning strategies discussed above addresses important aspects of the numerical challenges posed by the Helmholtz equation. Multigrid methods are fast and mature but struggle with robustness at high frequencies. Shifted Laplacian preconditioners offer simplicity and generality but require careful tuning and additional complex solves. Sweeping preconditioners are highly effective and nearly k-independent but remain difficult to scale and generalise, particularly in three dimensions.

In contrast, domain decomposition methods offer a compelling combination of theoretical scalability, flexibility across heterogeneous media, and compatibility with parallel architectures. They allow localised treatment of high-frequency effects, can be naturally enriched with problem-adapted coarse spaces, and support modular implementation strategies. Crucially, they can accommodate a range of boundary conditions and discretisations while still enabling the construction of preconditioners with controlled k-dependence.

For these reasons, domain decomposition serves as the foundation for the preconditioning strategies developed in this work. The next chapter presents these methods in depth, focusing on their construction, analysis, and performance in the context of high-frequency Helmholtz problems.

Outline of what follows. In the next chapter, we present domain decomposition methods in detail, including their theoretical framework, algorithmic implementation, and adaptation to high-frequency Helmholtz problems.

Chapter 3

Overview of Domain Decomposition Methods

Building upon the solver strategies introduced in the previous chapter, we now turn our attention to domain decomposition methods—an increasingly vital class of techniques for large-scale PDE problems. Their inherent parallelism and reduced memory footprint make them particularly appealing for high-performance computing. However, their success depends critically on the effectiveness of their design, especially in the context of complex geometries, heterogeneous media, or high-frequency wave propagation. This chapter begins with a historical and conceptual overview of classical Schwarz methods and their modern adaptations, which laid the foundation for iterative substructuring techniques. We then focus on the use of domain decomposition as preconditioners for Krylov methods, highlighting the limitations of one-level approaches in terms of scalability and robustness. To overcome these issues, we introduce the two-level framework, where global coupling is achieved through the construction of a coarse space. The chapter concludes with a detailed survey of state-of-the-art coarse space designs, including geometric, spectral, and multiscale approaches, with a particular emphasis on recent developments tailored to the Helmholtz equation. These developments provide the foundation for the theoretical and numerical advances that will be presented in the following chapters.

3.1 Classical Domain Decomposition

This section provides historical context and a conceptual foundation for domain decomposition methods by revisiting their classical roots. A comprehensive historical account is available in [46], but we briefly recall the original motivations and ideas that underpin many modern preconditioners and parallel solvers.

The method of domain decomposition was introduced by H. Schwarz to overcome the problem of irregular domains when attempting to prove existence and uniqueness of the solution of the homogeneous Poisson problem,

$$-\Delta u = f \qquad \text{in } \Omega, \tag{3.1.1a}$$

$$u = 0 \qquad \text{on } \partial\Omega, \tag{3.1.1b}$$

where $\Omega = \Omega_1 \cup \Omega_2$ is decomposed into two overlapping subdomains. In Schwarz's case, the domain in question was the configuration of an overlapping circle and rectangle, as shown in Figure 3.1. Schwarz proposed an iterative procedure that alternately solves the problem on Ω_1 and Ω_2, using boundary data from the neighbouring subdomain in the overlap region. This process ensures global convergence even when classical methods struggle on irregular geometries.

The solution over simple geometric domains, such as circles, rectangles, etc. was known at the time by using Fourier series, but these tools could not be used for more complex domains.

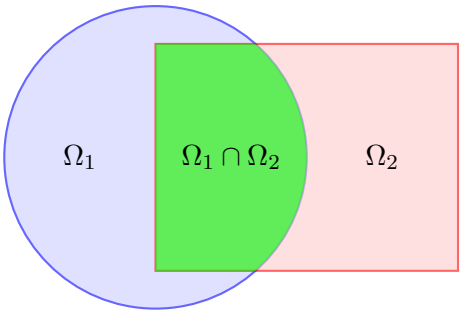


Figure 3.1: Original irregular domain.

For problem (3.1.1) to be solved on the domain $\Omega = \Omega_1 \cup \Omega_2$, Schwarz constructed an iterative process. He proposed a procedure that alternately solves the problem on Ω_1 and Ω_2, using boundary data from the neighbouring subdomain in the overlap region, ensuring convergence to a global solution on irregular geometries. The resulting scheme, now known as the classical Schwarz method, is defined by:

$$\begin{cases} -\Delta u_1^{n+1} = f & \text{in } \Omega_1 \\ u_1^{n+1} = 0 & \text{on } \partial\Omega \cap \partial\Omega_1 \\ u_1^{n+1} = u_2^n & \text{on } \partial\Omega_1 \setminus \partial\Omega \end{cases} \qquad \begin{cases} -\Delta u_2^{n+1} = f & \text{in } \Omega_2 \\ u_2^{n+1} = 0 & \text{on } \partial\Omega \cap \partial\Omega_2 \\ u_2^{n+1} = u_1^{n+1} & \text{on } \partial\Omega_2 \setminus \partial\Omega. \end{cases} \tag{3.1.2}$$

The iterative method in (3.1.2) demonstrates how the exchange of information is accomplished between the neighbouring subdomains, and that the solution on Ω_2 for iteration $n+1$ depends on the n solution on Ω_1. The alternating nature of this method means that it is not possible to run this in parallel, causing convergence to be slow. Another limitation of this method is that it requires an overlap between the domain Ω_1 and Ω_2; this method will not work if there is no overlap.

Building on the pioneering work carried out by Schwarz, P.L Lions modified the classical Schwarz method (3.1.2), developing a method that allowed for solutions to be found in each domain in parallel,

$$\begin{cases} -\Delta u_1^{n+1} = f & \text{in } \Omega_1 \\ u_1^{n+1} = 0 & \text{on } \partial\Omega \cap \partial\Omega_1 \\ u_1^{n+1} = u_2^n & \text{on } \partial\Omega_1 \setminus \partial\Omega \end{cases} \qquad \begin{cases} -\Delta u_2^{n+1} = f & \text{in } \Omega_2 \\ u_2^{n+1} = 0 & \text{on } \partial\Omega \cap \partial\Omega_2 \\ u_2^{n+1} = u_1^n & \text{on } \partial\Omega_2 \setminus \partial\Omega. \end{cases}$$

This method starts from the initial guess (u_1^0, u_2^0), solves the problem on domains Ω_1 and Ω_2 in parallel before moving to the next iteration. This method still exchanges information in the same manner, and still requires the domains to be overlapping. Although convergence remains slow, this represents a significant step forward in exploiting parallelisation to solve the problem. It was also demonstrated that increasing the size of the overlap could improve the convergence.

The alternating Schwarz method has been adapted and developed into fully parallel methods with far reaching applications, as can be seen in [89, 29]. Whilst the alternating and parallel Schwarz methods were originally proposed as iterative processes, it was soon established that these methods had potential as preconditioners to accelerate convergence in iterative methods, such as the Krylov iterative methods.

The evolution from Schwarz's original method to parallel formulations marks a key conceptual shift: from a theoretical tool for proving existence on complex domains to a practical foundation for scalable numerical solvers. Their reinterpretation as preconditioners for Krylov subspace methods laid the groundwork for more sophisticated domain decomposition approaches. The next sections build on this legacy, focusing on the formulation of Schwarz methods as preconditioners, and on the challenges of scalability and robustness that arise in modern large-scale and high-frequency problems.

3.2 One-Level Additive Schwarz Preconditioner for the Helmholtz Problem

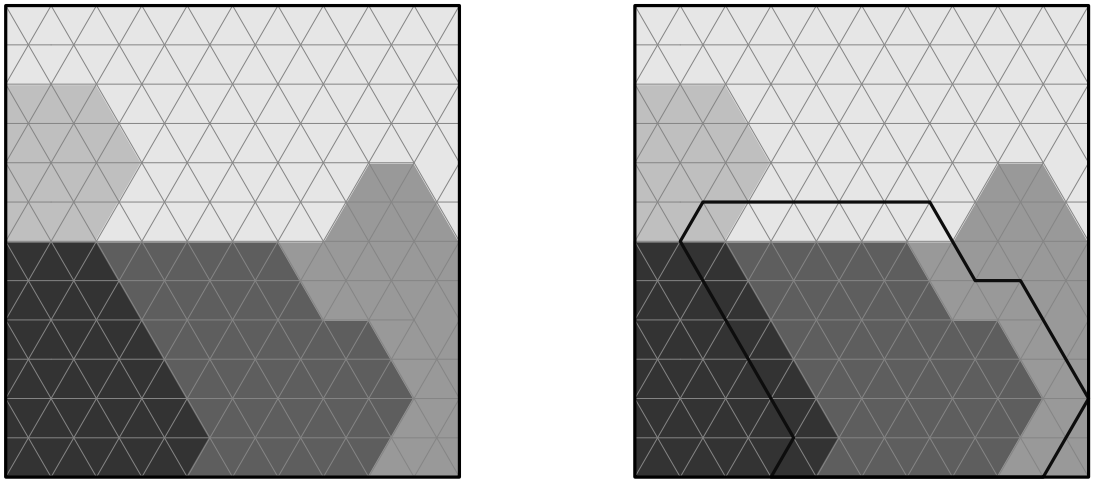
Building on the classical domain decomposition techniques introduced in the previous section, we now turn to their practical use as preconditioners for solving large-scale linear systems. In the context of the Helmholtz problem, achieving fast convergence with Krylov subspace methods, such as GMRES, often requires an appropriate preconditioner. Domain decomposition methods offer an attractive framework for designing parallel and scalable preconditioners. One of the simplest and most widely used among these is the one-level Additive Schwarz (AS) method, which we introduce in this section, together with its optimised variant, ORAS.

3.2.1 Discretisation and Domain Partitioning

We consider the discretised Helmholtz problem given in the form

$$\mathbf{Bu} = \mathbf{f},$$

where $\mathbf{B}$ is the discrete Helmholtz system matrix over the global domain Ω, $\mathbf{u}$ is the vector of unknowns, and $\mathbf{f}$ is the discrete forcing term.



(a) Example partitioning into 5 subdomains.

(b) Expansion of one subdomain by one layer of elements

Figure 3.2: Example partitioning of the global domain and expansion by one layer of finite elements.

Let $\mathcal{T}_h$ be a conforming triangulation of the global computational domain Ω, and let $V^h \subset H_0^1(\Omega)$ be a conforming finite element space associated with $\mathcal{T}_h$. The global domain is then decomposed into a set of N non-overlapping subdomains, $\{\Omega_i{}'\}_{i=1}^N$, which are resolved by the triangulation $\mathcal{T}_h$.

Definition 3.2.1 (Resolved by the triangulation)**.** *We say that the subdomain Ω_i' is* resolved by $\mathcal{T}_h$ *if there exists a subcollection*

$$\mathcal{T}_h(\Omega_i') \;:=\; \{K \in \mathcal{T}_h :\; K \subset \overline{\Omega_i'}\}$$

such that

$$\overline{\Omega_i'} \;=\; \bigcup_{K \in \mathcal{T}_h(\Omega_i')} \overline{K}.$$

Equivalently, the interface $\partial\Omega_i' \cap \Omega$ is a union of mesh faces/edges; no element of $\mathcal{T}_h$ is cut by $\partial\Omega_i'$.

In particular, Definition 3.2.1 implies the subdomain interface is aligned with the element facets of $\mathcal{T}_h$. If the decomposition is not resolved (i.e. the interfaces cut elements), one needs unfitted techniques or element subdivision to obtain a conforming discretization.

The splitting can be obtained by a user-defined method, or by using more automated processes based on the geometry of the domain, such as Metis [65] or Scotch [20], as seen in Figure 3.2a. These methods are based on the connectivity graph of $\mathbf{B}$, returning a labelling of the graph vertices into N parts, which we interpret as subdomains. From the input system matrix, $\mathbf{B}$, an undirected connectivity graph $G(\mathbf{B}) = (\mathcal{N}, \mathcal{C})$ is constructed:

$$\mathcal{N} = \{1, \dots, n\}, \qquad \{i, j\} \in \mathcal{C} \iff i \neq j \quad \text{and} \quad (B_{ij} \neq 0 \text{ or } B_{ji} \neq 0),$$

where $\mathcal{N}$ is the set of indices and $\mathcal{C}$ is the connectivity. If $\mathbf{B}$ is nonsymmetric, the graph is usually *symmetrized*. Algorithms that determine the partitioning of the vertices are then utilised. The distribution of vertices achieved must be done so that the number of elements assigned to each processor is roughly the same, and the number of adjacent elements assigned to different processors is minimized (graph cuts). The aim is to satisfy two goals:

1. Each partition carries approximately the same work, balancing computational cost and memory per process.

2. The total number of edges crossing partitions is minimized, since every cut edge indicates a data dependency between distinct processes.

These goals assist with the parallel implementation of the problem. If partition Ω_i' is assigned to processor p, the balance constraint spreads the computational work evenly, while the cut objective reduces the number of neighbouring processes.

In this work we employ mesh-resolved partitions, i.e. $\Omega_i' = \mathrm{Int}\big(\bigcup_{K \in \mathcal{T}_h(\Omega_i')} \overline{K}\big)$, so that the interface $\Gamma = \bigcup_{i \neq j} (\partial\Omega_i' \cap \partial\Omega_j')$ is a union of mesh faces. We use the notation $\mathrm{Int}(\cdot)$ for the interior of a domain. This choice simplifies the construction of Boolean restriction/prolongation operators, the definition of local spaces $V^h(\Omega_i')$, and the assembly of subdomain operators, while retaining good balance and small interfacial cuts.

The subdomains $\Omega_i{}'$ are then extended by one or more layers of elements, as in Figure 3.2b, in the following sense,

Definition 3.2.2. *Given a subdomain $D' \subset \Omega$, which is resolved by the chosen mesh, the extension of D' by a layer of elements is*

$$D = \mathrm{Int}\left(\bigcup_{\{\ell \mid \mathrm{supp}(\phi_\ell) \cap D' \neq \emptyset\}} \mathrm{supp}(\phi_\ell) \right)$$

Extension by multiple layers can then be achieved by applying this recursively.

With the use of Definition 3.2.2, the set of overlapping subdomains $\{\Omega_i\}_{i=1}^N$ can be constructed. For the domains Ω_i, we define the spaces,

$$\widetilde{V}_i := \left\{ v|_{\Omega_i} : v \in V^h \right\} \subset H^1(\Omega_i) \qquad \text{and} \qquad V_i := \left\{ v \in \widetilde{V}_i : v|_{\partial\Omega_i} = 0 \right\} \subset H_0^1(\Omega_i),$$

The spaces are well-defined by Definition 3.2.2. Let $\mathcal{N}$ be the set of indices of degrees of freedom of the global finite element space V^h, with $\mathcal{N}_i$ the set of indices of degrees of freedom of the local finite element space V_i. We now define the restriction operator, from the global space, V^h, to the local space $\widetilde{V}_i$,

$$R_i : V^h \to \widetilde{V}_i.$$

At the discrete level, this is a rectangular $|\mathcal{N}_i| \times |\mathcal{N}|$ matrix such that if $\mathbf{v}$ is the vector of degrees of freedom of $v \in V^h$, then $\mathbf{R}_i\mathbf{v}$ is the vector with degrees of freedom of $v|_{\Omega_i} \in \widetilde{V}_i$. For any $v \in V_i$, let $E_i v$ denote its zero extension to the whole of the global domain, Ω. Then,

$$E_i : V_i \to V^h, i = 1, \ldots, N.$$

Again, at the discrete level, this is given by a rectangular $|\mathcal{N}| \times |\mathcal{N}_i|$ matrix, $\mathbf{E}_i$. The space V_h is assumed to decompose

$$V^h = \sum_{i=1}^{N} E_i V_i.$$

Note that this sum is not necessarily a *direct* sum. That is, a function in V^h may be expressible as a sum of functions from the local spaces V_i, but the representation may not be unique. This is due to the fact that the subdomains Ω_i typically overlap, which implies that the local spaces V_i are not mutually orthogonal and do not form subspaces in the algebraic sense. Therefore, while we can reconstruct global functions from the local contributions, we cannot uniquely attribute portions of a function in V^h to a single V_i. For this reason, the V_i are often called 'local spaces' rather than true subspaces of V^h.

We also require a partition of unity operator that appropriately handles the unknowns in the overlapping regions. This operator consists of continuous piecewise linear functions associated with each subdomain, satisfying the following conditions:

Definition 3.2.3 (Partition of unity)**.** *A partition of unity subordinate to the subdomains* $(\Omega_i)_{i=1}^N$ *is a set of functions* $(\Xi_i)_{i=1}^N$ *such that:*

1. $\operatorname{supp}(\Xi_i) \subset \overline{\Omega_i}, \quad i = 1, \ldots, N,$
2. $0 \leq \Xi_i(x) \leq 1, \quad$ *for all* $x \in \Omega_i,$
3. $\sum_{i=1}^N \Xi_i(x) = 1, \quad$ *for all* $x \in \Omega.$

These functions ensure that each point in the global domain is represented as a convex combination of contributions from overlapping subdomains, enabling smooth transitions across interfaces.

In the discrete setting, the partition of unity is implemented via diagonal matrices $\mathbf{D}_i$ of size $|\mathcal{N}_i| \times |\mathcal{N}_i|$, where $\mathcal{N}_i$ denotes the set of degrees of freedom in subdomain Ω_i. These matrices satisfy the identity:

$$\mathbf{I} = \sum_{i=1}^{N} \mathbf{E}_i \mathbf{D}_i \mathbf{R}_i,$$

where $\mathbf{E}_i$ and $\mathbf{R}_i$ are the extension and restriction matrices associated with subdomain Ω_i, and $\mathbf{I}$ is the global identity matrix.

Using the previously defined restriction and extension matrices $\mathbf{R}_i$ and $\mathbf{E}_i$, we now define local approximations of the global matrix $\mathbf{B}$ on each subdomain Ω_i. Specifically, we introduce the **local discrete Helmholtz system matrix** $\mathbf{B}_i$ as

$$\mathbf{B}_i = \mathbf{R}_i \mathbf{B} \mathbf{E}_i, \quad \text{for } i = 1, 2, \ldots, N.$$

This matrix represents the discrete Helmholtz operator restricted to the local degrees of freedom in subdomain Ω_i. Conceptually, $\mathbf{B}_i$ captures how $\mathbf{B}$ acts within Ω_i, isolating the local portion of the global problem. Next, we define the corresponding **local solution matrix**

$$\mathbf{S}_i = \mathbf{E}_i \mathbf{B}_i^{-1} \mathbf{R}_i, \quad \text{for } i = 1, 2, \ldots, N.$$

This matrix represents solving the local problem on Ω_i and embedding the result back into the global space.

With these local operators in hand, the **one-level Additive Schwarz (AS) preconditioner** is given by the sum

$$\mathbf{M}_{\text{AS}}^{-1} = \sum_{i=1}^{N} \mathbf{S}_i = \sum_{i=1}^{N} \mathbf{E}_i \mathbf{B}_i^{-1} \mathbf{R}_i.$$

This construction was first introduced in [32] and [92]. It can be interpreted as an approximation to the inverse of $\mathbf{B}$, obtained by solving smaller problems on overlapping subdomains and combining the solutions.

Although effective and parallelisable, the one-level AS method comes with important limitations:

- **Redundant computation in overlaps:** Each overlapping region contributes to multiple subdomain solves, increasing the overall cost.
- **Inflexibility in heterogeneous domains:** When the domain contains distinct material regions, a decomposition aligned with material interfaces would be preferable—but this is often incompatible with overlap-based constructions.

In [18], an alternative to the one-level AS method was introduced that requires fewer communications between subdomains was introduced. This method is known as the one-level **Restricted Additive Schwarz (RAS) preconditioner**:

$$\mathbf{M}_{\text{RAS}}^{-1} = \sum_{i=1}^{N} \mathbf{A}_i = \sum_{i=1}^{N} \mathbf{E}_i' \mathbf{B}_i^{-1} \mathbf{R}_i.$$

We recall that we have previously defined the restriction operator $\mathbf{R}_i$, which restricts from the global space to the overlapping local spaces, and the corresponding extension operator, $\mathbf{E}_i$ that extends back to the global space. In the RAS preconditioner, we now introduce the the restriction operator $\mathbf{R}_i'$, which now carries out the restriction associated with the non-overlapping domains, Ω_i', whilst also making the information associated with the degrees of freedom in the overlap, $\Omega_i \backslash \Omega_i'$, vanish. The associated extension operator, $\mathbf{E}_i'$, will then extend back to the global space. Thus we have

$$\sum_{i=1}^{N} \mathbf{E}_i' \mathbf{R}_i = \mathbf{I}.$$

A simple way to generalise the RAS preconditioner is to replace $\mathbf{E}_i'$, which corresponds to the non-overlapping partition, with $\mathbf{E}_i' = \mathbf{E}_i \mathbf{D}_i$, where $\mathbf{D}_i$ are the partitions of unity such that $\mathbf{I} = \sum_{i=1}^{N} \mathbf{E}_i \mathbf{D}_i \mathbf{R}_i$ as defined previously. The effect of the RAS preconditioner is to eliminate the redundant computation that occurs in the overlaps.

In our numerical experiments, we will also be using another method, namely the **Optimised RAS (ORAS)** preconditioner which is based on local boundary value problems with Robin boundary conditions (impedance boundary conditions). In this case, we let $\mathbf{B}_i'$ be the stiffness matrix associated with the local Robin problem

$$\begin{aligned} -\nabla \cdot (A \nabla u_i) - k^2 u_i &= f \qquad \text{in } \Omega_i, \\ u_i &= 0 \qquad \text{on } \partial\Omega \cup \partial\Omega_i, \\ \frac{\partial u_i}{\partial n_i} + iku_i &= 0 \qquad \text{on } \partial\Omega_i \backslash \partial\Omega. \end{aligned}$$

The ORAS preconditioner is then defined as

$$\mathbf{M}_{\text{ORAS}}^{-1} = \sum_{i=1}^{N} \mathbf{A}_i = \sum_{i=1}^{N} \mathbf{E}_i' \mathbf{B}_i'^{-1} \mathbf{R}_i.$$

A significant drawback of the one-level methods is **scalability**: as the number of subdomains increases or the wavenumber grows (as is typical in Helmholtz problems), the convergence of the preconditioned system can deteriorate markedly. These limitations motivate the introduction of more robust and scalable strategies, such as two-level methods, which will be the subject of the next section.

3.2.2 Lack of Scalability and Robustness

One of the key properties desired in a preconditioner is *scalability*—the ability to maintain fast convergence even as the problem size or computational resources increase. In the context of domain decomposition, we adopt the following definition from [89]:

Definition 3.2.4 (Scalability)**.** *A domain decomposition iterative method for solving a linear system is said to be scalable if its rate of convergence does not deteriorate as the number of subdomains increases. Equivalently, convergence should remain stable as the characteristic subdomain size H becomes small.*

It is well known that the convergence of the one-level preconditioner deteriorates with increasing number of subdomains in the decomposition, see [29] and the references therein. An example of this issue can be seen in Table 3.1, which shows the iteration count over a simple domain where $\Omega = (0,1)^2$ with Dirichlet boundary conditions for a increasing wavenumber k and number of subdomains N. Here the iteration count grows as the number of subdomains increases. This effect is increasingly problematic as k increases.

	N				
k	16	36	64	100	144
20	56	106	123	136	147
40	115	245	195	384	390
60	153	235	225	451	332
80	224	322	> 500	> 500	> 500
100	257	380	> 500	> 500	> 500

Table 3.1: Iteration count when using the one-level AS preconditioner for different k and number of subdomains N.

In other words, the method is not scalable with respect to the number of subdomains. A central reason for this behaviour is the *limited communication* between subdomains: information is only exchanged between adjacent neighbours, and no mechanism exists for global coordination. To illustrate this issue, consider a decomposition into N_x subdomains along the x-direction. Information originating in one subdomain at the left boundary must propagate iteratively through N_x neighbors before it can affect the solution at the right boundary, as illustrated by Figure 3.3. This results in slow information transfer across the domain, making the method poorly suited as a scalable preconditioner.

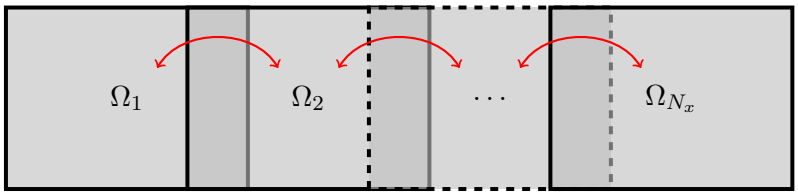


Figure 3.3: Illustration for the lack of communication between domains.

To distinguish between different scaling behaviours, it is helpful to introduce the notions of weak and strong scalability:

Definition 3.2.5 (Weak scalability)**.** *A domain decomposition method is weakly scalable if its convergence rate remains constant as both the problem size and the number of subdomains increase proportionally.*

Definition 3.2.6 (Strong scalability)**.** *An algorithm is strongly scalable if its runtime decreases proportionally with the number of processors used to solve a fixed-size problem.*

Strong scalability aims to reduce solution time through parallelism on a fixed problem. Ideally, doubling the number of processors halves the computation time, but in practice, overhead from communication and synchronisation often limits this speedup. Weak scalability, on the other hand, keeps the computational load per processor constant while increasing both problem size and resources; it is especially relevant for memory-bound problems where the full problem cannot fit on a single compute node.

Beyond scalability, **robustness** is another critical requirement for preconditioners. In the Helmholtz setting, robustness means that performance should not degrade significantly with increasing the wavenumber k or in the presence of strong heterogeneities in the medium. The one-level AS method often fails to meet this standard: both high-frequency oscillations and material discontinuities can severely hinder convergence.

To address these challenges, more advanced preconditioners are needed. A natural extension is the *two-level method*, which augments the one-level approach with a *global coarse problem*. This additional component enables communication across the full domain by solving an approximate problem on a lower-dimensional space, thereby improving both scalability and robustness. The next section is devoted to the construction and analysis of two-level domain decomposition preconditioners.

3.3 Two-Level Domain Decomposition Methods

As highlighted in the previous sections, one-level domain decomposition methods suffer from limited scalability due to the absence of global communication between subdomains. This manifests as a deterioration in convergence as the number of subdomains increases. To overcome this limitation, a second level of resolution—a coarse space—is introduced, providing a mechanism for global information exchange across the entire computational domain. The idea is to complement the local subdomain solves with a global correction step, in which a coarse problem, defined over a low-dimensional space, is solved using a direct method. This hierarchical approach is referred to as a two-level domain decomposition method. Pioneering contributions in this direction include the BPS preconditioner of Bramble, Pasciak, and Schatz [17], and the two-level overlapping Schwarz method of Dryja and Widlund [33].

Assume we have a space denoted $V_H \subset V_h$, that can be defined later, defined on a global domain. Then we can define the natural embedding operator, E_0,

$$E_0 : V_H \to V_h,$$

and let R_0 be the L^2-adjoint of E_0 (their matrix counterparts are denoted by $\mathbf{E}_0$ and $\mathbf{R}_0$). Using the definitions of the one-level preconditioners given in the previous section, a two-level preconditioner is constructed by the addition of a global coarse correction to the local contributions of the one-level preconditioners. This introduction of the coarse space can effectively eliminate the dependency of the two-level convergence rate on the number of subdomains.

The two-level preconditioner that forms the basis of the theoretical work in this thesis is the two-level additive Schwarz (AS) preconditioner, which takes the form

$$\mathbf{M}_{AS,2}^{-1} = \mathbf{E}_0\mathbf{B}_0^{-1}\mathbf{R}_0 + \sum_{i=1}^{N} \mathbf{E}_i\mathbf{B}_i^{-1}\mathbf{R}_i,$$

where $\mathbf{B}_0$ is known as the coarse problem, given by $\mathbf{B}_0 = \mathbf{R}_0\mathbf{B}\mathbf{E}_0$. It is also possible to use the RAS or ORAS one-level preconditioner in combination with the coarse space to get their respective two-level methods.

Whilst the two-level AS method is one of simpler applications of the coarse space, there are other methods that can be more suitable. One of these methods, which we will be utilising in the numerical experiments, is the adapted deflation method, commonly known as ADEF1 [88, 87]. The two-level method using the ADEF1 coarse correction is given as

$$\mathbf{M}_2^{-1} = \mathbf{M}^{-1}(\mathbf{I} - \mathbf{B}\mathbf{E}_0\mathbf{B}_0^{-1}\mathbf{R}_0) + \mathbf{E}_0\mathbf{B}_0^{-1}\mathbf{R}_0,$$

where $\mathbf{M}^{-1}$ is a one-level preconditioner of choice.

The inclusion of a well-designed coarse space ensures that the convergence of the method becomes independent of the number of subdomains, provided that the coarse space captures the global low-frequency error modes effectively. In what follows, we review early classical constructions of coarse spaces that motivated further research in this area.

3.3.1 Classical Coarse Space Constructions

The Nicolaides Coarse Space One of the earliest and simplest coarse spaces was proposed by Nicolaides in [76]. His idea was to construct a space of piecewise constant functions over the subdomains:

$$V_H^{\text{NICO}} = \text{span}\left\{\mathbf{I}_{\Omega_1}, \ldots, \mathbf{I}_{\Omega_N}\right\},$$

where $\mathbf{I}_{\Omega_i}$ is the indicator function on subdomain Ω_i. This coarse space is appealing due to its extremely low computational cost and its ability to provide minimal global coupling between subdomains.

However, due to the piecewise constant nature of its basis functions, the gradients are discontinuous across interfaces, leading to poor energy-norm approximation properties. In particular, the coarse space norm scales like H/h, where H is the subdomain diameter and h the fine mesh size. While this scales with respect to the number of subdomains, condition number estimates often depend on the jumps on the physical parameters of the problem leading to a lack of robustness with respect to those especially in the case of multiscale problems.

Despite its limitations, the Nicolaides space laid the groundwork for more advanced constructions that improve global coupling while preserving computational efficiency.

Partition of Unity Coarse Space Another early example of a coarse space that demonstrated improved performance over the Nicolaides space is the partition of unity (PoU) coarse space, introduced in [80]. This construction leverages overlapping domain decomposition and introduces smooth, continuous functions to mediate the interaction between overlapping subdomains as defined in Definition 3.2.3.

The corresponding coarse space is then constructed as the span of PoU-weighted basis functions. These are typically piecewise constant in the interior of each subdomain and decay linearly to zero across the overlapping regions, thus enforcing smooth global coupling.

For the Poisson problem, it has been shown [89] that such a coarse space enables convergence that is independent of both the mesh size and the number of subdomains, under certain geometric constraints on the decomposition. This improved behaviour is a result of the scaling and quotient space arguments (see [89, Section 3.4]), which justify the necessity of using piecewise constants in combination with partition of unity weights. Furthermore, the structure of the PoU functions facilitates the application of Poincaré-type inequalities, which play a critical role in the convergence analysis.

Geometric Grid Coarse Space A natural and widely used choice for constructing a coarse space is to base it on a *coarser discretisation* of the computational domain—commonly referred to as a *coarse grid*. One of the earliest instances of this approach appears in [92]. Since then, the idea has been extensively developed and adapted to more complex problems, such as the *absorptive Helmholtz equation* [56] and Helmholtz problems with *impedance (Robin) boundary conditions* [58]. Today, geometric coarse grids form a cornerstone of many robust domain decomposition preconditioners.

The central idea is straightforward: introduce a *coarse mesh* $\mathcal{T}_{\text{coarse}}$ over the global domain Ω, with a mesh size h_{coarse}. Let $V_0 \subset V$ denote the associated coarse finite element space. To couple this coarse space with the fine space V^h, one uses a *nodal interpolation operator* $I_0 : V^{h_{\text{coarse}}} \to V^h$, which maps coarse functions onto the fine mesh. This operator has a corresponding matrix representation $\mathbf{E}_0$. The *coarse problem matrix* is then defined as $\mathbf{B}_0 = \mathbf{R}_0\mathbf{B}\mathbf{E}_0$. For *symmetric positive definite (SPD)* problems, geometric coarse spaces have been shown to yield *robust convergence*, provided that the coarse mesh adequately resolves variations in material properties or coefficients [19]. This theory was further generalised in [55], where it was demonstrated that, under suitable combinations of subdomain decomposition and coarse basis functions, robust convergence can still be achieved *without* requiring the heterogeneities to be mild or slowly varying.

However, most of the existing theoretical results apply to SPD problems. The extension to *indefinite problems*, such as the Helmholtz equation, poses additional challenges. Progress has been made for the *absorptive Helmholtz equation*, where the addition of a small imaginary component (representing absorption) renders the problem *coercive*. In this context, [56, 58] established *explicit convergence bounds for GMRES*, expressed in terms of the wavenumber, absorption coefficient, coarse mesh diameter, subdomain size, and overlap width. Under specific parameter choices, they even demonstrated conditions under which convergence becomes *independent of the wavenumber*.

Numerical experiments presented in [14] confirmed that geometric coarse spaces can be *effective even in highly heterogeneous and high-frequency regimes*. Nonetheless, this approach comes with practical challenges. Since the performance of the coarse grid depends strongly on *problem-specific physical parameters* (notably the wavenumber), achieving optimal performance often requires *a priori knowledge* of these parameters. Furthermore, when the coarse mesh must resolve fine-scale physical features, it can become *computationally expensive*, undermining some of the original efficiency gains.

3.3.2 Spectral Coarse Spaces

Despite the success of classical coarse space constructions such as the Nicolaides and geometric grid spaces, designing efficient and robust coarse spaces remains one of the central challenges in domain decomposition methods. The main difficulty lies in balancing two often competing objectives: maintaining low computational cost while ensuring rapid convergence of the iterative method.

Over the past two decades, this challenge has driven a surge of innovation, leading to the development of increasingly sophisticated approaches to coarse space design. While the constructions introduced so far lay important foundations, they are often insufficient for problems involving strong heterogeneities, high frequencies, or complex geometries—conditions under which classical methods may fail to deliver scalable performance.

This section provides an overview of some of the most prominent modern strategies for constructing coarse spaces. Although not exhaustive, this survey highlights key developments and theoretical ideas that have shaped the current landscape. Particular emphasis is placed on spectral coarse spaces, which form the theoretical and practical focus of this thesis.

Spectral coarse spaces represent a class of coarse space constructions designed to enhance the convergence properties of domain decomposition methods—especially for problems involving

highly heterogeneous coefficients, complex geometries, or indefinite operators.

The central idea is to use spectral information to identify and incorporate global modes that are not well captured by local subdomain solvers. These typically correspond to low-energy error modes, which are slow to decay and thus dominate the residual when using one-level methods. By enriching the coarse space with these critical modes, the resulting preconditioner becomes more effective at reducing global error components and improving convergence rates.

This is achieved by solving a family of generalised eigenvalue problems, which take the form:

$$\mathbf{Bx} = \lambda \mathbf{Cx}, \quad \text{with } \mathbf{B}, \mathbf{C} \in \mathbb{C}^{|\mathcal{N}| \times |\mathcal{N}|},$$

where the specific choices of the bilinear forms represented by $\mathbf{B}$ and $\mathbf{C}$ depend on the problem and the construction strategy. In this framework, eigenvectors associated with small eigenvalues are interpreted as global error modes that are poorly handled by the local preconditioners and must therefore be included in the coarse space.

A direct solution of this eigenproblem on the global domain is computationally prohibitive, as it is often more expensive than solving the original PDE itself. To mitigate this, modern spectral coarse space methods formulate and solve local eigenproblems independently on each subdomain, typically in parallel. This strategy allows the coarse space to capture both local and global features of the solution without incurring a prohibitive computational cost.

In recent years, significant progress has been made in refining spectral coarse spaces for a wide variety of problems, particularly those with highly variable material parameters. The different methods mainly diverge in their choice of the bilinear form used on the right-hand side of the eigenproblem, reflecting different theoretical motivations and target applications.

GenEO-type Coarse Spaces As discussed above, spectral coarse spaces are built by enriching the coarse space with global modes that are responsible for slow convergence in domain decomposition methods. One of the most influential and practically successful realisations of this idea is the *Generalised Eigenproblems in the Overlap (GenEO)* approach, introduced in [85].

Developed for symmetric positive definite (SPD) problems with strong coefficient heterogeneities, the GenEO method provides a *robust and scalable* coarse space framework that has proven particularly effective for elliptic PDEs arising from variational formulations. Traditional domain decomposition methods often struggle with heterogeneous coefficients, leading to poor convergence. GenEO addresses this by explicitly constructing coarse basis functions associated with low-energy modes that are difficult to eliminate locally. A full description of the original GenEO-type coarse space can be found in Section 6.4.2.

Δ-GenEO Coarse Space With the basic GenEO definition as the base, extending this definition to be used with the Helmholtz equation has proven to be a non-trivial task. This is in no small part due to the indefiniteness and the lack of self-adjointness found in the Helmholtz problem. However, significant progress has been made in [13], where it was demonstrated that convergence in GMRES could be proved given that certain conditions on the size of the subdomains in the decomposition and the eigenvalue threshold value are met. It was argued in [13], that using a nearby symmetric positive definite variational form, the problems caused by the indefiniteness and the lack of self-adjointness can be avoided. For the case of the Helmholtz problem, this is realised by setting the wavenumber $k = 0$ and using the Laplacian operator. This new method was called the Δ-GenEO. A full description of the Δ-GenEO method can be found in Section 6.4.2.

Further GenEO Developments To address the shortcomings of Δ-GenEO at high frequencies, we have developed the Δ_k-GenEO, which is an adaptation of the Δ-GenEO to couple the original nearby SPD problem with a k-weighted norm. The development of the Δ_k-GenEO is a key contribution to later chapters. Whilst we developed the Δ_k-GenEO, the authors of

[14] proposed their own alternative to the Δ-GenEO: the *H-GenEO coarse space*, in which the full *Helmholtz operator* is used in the left-hand side of the eigenproblem. This leads to the formulation:

$$b_{\Omega_i}(p, v) = \lambda\, a_{\Omega_i}(\Xi_i p, \Xi_i v) \quad \forall v \in V_h(\Omega_i),$$

where $b_{\Omega_i}(\cdot,\cdot)$ is the bilinear form for the Helmholtz operator, and $a_{\Omega_i}(\cdot,\cdot)$ remains the Laplacian form. Numerical results from [14, 13] demonstrate improved performance at higher frequencies compared to Δ-GenEO. However, no rigorous theory exists yet for H-GenEO, and its convergence properties remain an open question.

In this thesis, we develop the theoretical foundations of the H-GenEO method, introducing the *H_k-GenEO* variant. This enhanced version couples the full Helmholtz problem with a *k-weighted norm*, and employs *distinct decompositions* for the one-level subdomains and for the local eigenvalue problems. These refinements aim to combine the robustness of GenEO with the frequency resolution needed for high-frequency Helmholtz problems, forming a central contribution of the forthcoming chapters.

Harmonic Coarse Spaces A broad family of modern coarse space methods—including the MS-GFEM-based approaches and the extended harmonic coarse space—derive their effectiveness from exploiting harmonic information within local or oversampled subdomains. These methods solve local generalised eigenvalue problems within carefully defined harmonic spaces to identify low-energy modes that are challenging to eliminate through standard one-level methods. In MS-GFEM, this is achieved through the construction of optimal local approximation spaces defined on generalised harmonic spaces using local spectral problems on oversampled domains. The extended harmonic coarse space similarly relies on discrete harmonic extensions and the use of extended subdomains, incorporating spectral filtering based on local Neumann solvers. Both approaches share the goal of enhancing scalability and robustness, especially for problems with complex local behaviour or high-frequency content. While promising, these methods often lead to large coarse spaces due to the inclusion of fine-scale harmonic features. A comparative evaluation of these harmonic-based strategies with other spectral coarse spaces will be conducted in Chapter 6.

Conclusion This chapter has presented the foundations and evolution of domain decomposition solvers, with particular attention to the design and role of coarse spaces in achieving scalable and robust performance. From classical constructions such as the Nicolaides and geometric coarse grid spaces, to advanced spectral approaches like GenEO and its variants, and finally to harmonic-based formulations including MS-GFEM and extended harmonic coarse spaces, we have traced the development of strategies that address the critical challenge of global error reduction. While each method brings its own strengths and trade-offs—ranging from ease of implementation to spectral accuracy—ongoing research continues to refine these tools to accommodate increasingly demanding problems, such as high-frequency wave propagation in heterogeneous media. The comparative performance, theoretical guarantees, and computational costs of these approaches will be further examined in the subsequent chapters, where their practical implications are assessed through both analysis and large-scale numerical experiments.

Chapter 4

Can Symmetric Positive Definite (SPD) Coarse Spaces Perform Well for Indefinite Helmholtz Problems?

This chapter was developed in collaboration with Victorita Dolean and Matthias Langer. The corresponding publication appeared in the Journal of Computational and Applied Mathematics [25]. It represents a further development of the proceedings paper entitled 'Wavenumber explicit estimates of the Schwarz preconditioner with Δ-GenEO coarse space for the indefinite Helmholtz problem" [26].

4.1 Introduction

This chapter studies overlapping two-level domain decomposition preconditioners for the heterogeneous Helmholtz problem

$$\begin{aligned} -\nabla \cdot (A\nabla u) - k^2 u &= f \qquad \text{in } \Omega, \\ u &= 0 \qquad \text{on } \partial\Omega, \end{aligned} \tag{4.1.1}$$

where $\Omega \subseteq \mathbb{R}^d$ $(d = 2, 3)$ is polygonal or polyhedral and Lipschitz, $A(x)$ is symmetric and uniformly positive definite (possibly high-contrast), $f \in L^2(\Omega)$, and $k > 0$ is the wavenumber. We assume (4.1.1) admits a unique weak solution $u \in H_0^1(\Omega)$ for all $f \in L^2(\Omega)$.

From Δ-GenEO to Δ_k-GenEO. A rigorous GenEO-based convergence analysis for indefinite and non-self-adjoint PDEs was developed in [13] via the Δ-GenEO method, which constructs the coarse space through local generalised eigenvalue problems derived from a nearby SPD operator. While this provided the first robust GMRES theory in an indefinite setting, the resulting sufficient conditions are overly pessimistic for Helmholtz-type problems: Theorem 4.1 of [13] guarantees robustness only if

$$H \lesssim k^{-2}, \qquad (1 + C_{\text{stab}})^2 \, k^8 \lesssim \tau, \tag{4.1.2}$$

where C_{stab} is the stability constant of the adjoint problem and τ is the eigenvalue tolerance that controls the coarse-space dimension.

In this chapter we introduce the Δ_k-*GenEO* method, a modification of Δ-GenEO in which the right-hand side of the local generalised eigenvalue problem is replaced by a k-weighted inner product. This adjustment preserves the computational simplicity of the SPD-based GenEO construction while aligning the theory more closely with Helmholtz behaviour. Our refined analysis yields substantially milder sufficient conditions for robust GMRES convergence:

$$H \lesssim k^{-1}, \qquad (1 + C_{\text{stab}})^2 \, k^2 \lesssim \tau. \tag{4.1.3}$$

Compared with (4.1.2), the dependence of H on k improves from quadratic to linear, and the lower bound on τ improves from order k^8 to quadratic. Although still conservative, these bounds clarify both the practical effectiveness and the intrinsic limitations of SPD-based spectral coarse spaces for indefinite Helmholtz problems.

Positioning relative to indefinite-operator coarse spaces. Recent Helmholtz-adapted spectral coarse spaces such as H-GenEO and H_k-GenEO [14, 24] formulate local eigenproblems directly using the indefinite operator and exhibit striking robustness in practice, but their complete theoretical analysis remains challenging. The Δ_k-GenEO framework plays a complementary role: it tightens the SPD-based analysis, provides a transparent benchmark for what can be achieved without abandoning SPD surrogates, and thereby helps delineate when genuinely indefinite-operator coarse spaces become necessary.

Contributions. The main contributions of this chapter are:

- a sharper k-explicit GMRES robustness theory for SPD-based GenEO coarse spaces, establishing (4.1.3) and thereby substantially improving the known sufficient conditions from [13];
- a refined stability/projection framework that identifies the structural sources of suboptimality in SPD-based coarse constructions for Helmholtz;
- numerical experiments demonstrating that Δ_k-GenEO often performs far better in practice than the conservative bounds predict, with moderate coarse dimensions and iteration counts over a broad parameter range.

Outline. Section 4.2 introduces the functional framework, weak formulation, discretisation, and the Δ_k-GenEO eigenproblems. Section 4.3 states the main GMRES convergence theorem and develops the key stability and approximation estimates. Section 4.4 contains the proof. Section 5.5 reports numerical experiments assessing robustness with respect to frequency, heterogeneity, and scalability.

4.2 Functional Framework and Discretisation

We now lay out the functional framework, assumptions, and discretisation strategies that will underpin our analysis. In particular, we describe the weak formulation of the Helmholtz problem, introduce the associated bilinear forms, and detail the finite element discretisation used in constructing and analysing the preconditioners. This section also recalls key regularity and stability assumptions and highlights analytical tools—such as the Friedrichs inequality—that play a central role in deriving later estimates.

The model problem is the Helmholtz equation with heterogeneous coefficients as defined in (4.1.1). In what follows we will state the main assumptions and technical lemmas. For this work, we will be assuming that Assumption 2.2.2 and (2.2.9) from Proposition 2.2.3 are satisfied. We will also be utilising the Schatz–Wang lemma, Lemma 2.3.1, throughout.

The weak formulation of (4.1.1) is: find $u \in H_0^1(\Omega)$ such that

$$b(u, v) = (f, v) \qquad \text{for all } v \in H_0^1(\Omega), \tag{4.2.1}$$

where the bilinear form $b(\cdot, \cdot) : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ is given by

$$b(u, v) = \int_\Omega \big(A\nabla u \cdot \nabla v - k^2 uv\big)\, \mathrm{d}\mathbf{x}.$$

Notation 4.2.1. *For any subdomain $\Omega' \subseteq \Omega$ we use $(\cdot,\cdot)_{\Omega'}$ to denote the $L^2(\Omega')$ inner product and $\|\cdot\|_{\Omega'}$ to denote the corresponding norm. We also introduce the bilinear forms:*

$$a_{\Omega'}(u,v) := \int_{\Omega'} A\nabla u \cdot \nabla v \,\mathrm{d}\mathbf{x}, \qquad b_{\Omega'}(u,v) := a_{\Omega'}(u,v) - k^2(u,v)_{\Omega'} \quad \text{for } u,v \in H^1(\Omega'),$$

and define the semi-norm induced by a, $\|u\|_{a,\Omega'} := \sqrt{a_{\Omega'}(u,u)}$. An important role is played by the k-weighted inner product:

$$(u,v)_{1,k,\Omega'} := a_{\Omega'}(u,v) + k^2(u,v)_{\Omega'},$$

and we denote the induced k-norm by $\|u\|_{1,k,\Omega'}$. When $\Omega' = \Omega$ we abandon the subscript Ω from these notations.

The form $a(\cdot,\cdot)$ is symmetric semi-positive definite on $H^1(\Omega)$, while $b(\cdot,\cdot)$ is symmetric but indefinite.

Let $\mathcal{T}_h$ be a shape-regular triangulation of Ω into simplices of maximal diameter h, and let $V^h \subseteq H^1_0(\Omega)$ be a conforming finite element space. The Galerkin discretisation of (4.2.1) is: find $u_h \in V^h$ such that

$$b(u_h, v) = (f, v) \qquad \text{for all } v \in V^h. \tag{4.2.2}$$

Let $\{\phi_i\}_{i=1}^n$ be a basis for V^h. Then (4.2.2) yields the linear system

$$\mathbf{B}\mathbf{u} = \mathbf{f}, \tag{4.2.3}$$

with entries $\mathbf{B}_{ij} = b(\phi_j, \phi_i)$, $\mathbf{u}_i = (u_h, \phi_i)$ and $\mathbf{f}_i = (f, \phi_i)$. Analogously, the stiffness matrix $\mathbf{A}$ is defined by $\mathbf{A}_{ij} = a(\phi_j, \phi_i)$.

Remark 4.2.2.

(a) *Lemma 2.3.1, the Schatz–Wang lemma, ensures discrete well-posedness despite indefiniteness. It will be repeatedly used in our stability and approximation estimates.*

(b) *Combining (2.3.3) and (2.3.4) we obtain*

$$\|u - u_h\|_{1,k} \le \varepsilon \|f\|$$

for small enough h.

A key tool in our analysis is the Friedrichs inequality.

Lemma 4.2.3 (Friedrichs inequality [68, Theorem 13.19])**.** *Let $\Omega' \subseteq \mathbb{R}^d$ be an open set lying between two parallel hyperplanes at distance L. Then, for all $u \in H^1_0(\Omega')$,*

$$\|u\|_{\Omega'} \le \tfrac{L}{\sqrt{2}} \|\nabla u\|_{\Omega'}.$$

Remark 4.2.4. *Combined with Assumption 2.2.2, Friedrichs' inequality yields: for any $\Omega' \subseteq \Omega$ of diameter H,*

$$\|u\|_{\Omega'} \le \tfrac{H}{\sqrt{2}} \|\nabla u\|_{\Omega'} \le \tfrac{H}{\sqrt{2}} \|u\|_a \le \tfrac{H}{\sqrt{2}} \|u\|_{1,k,\Omega'} \qquad \text{for all } u \in H^1_0(\Omega'). \tag{4.2.4}$$

In particular, for $\Omega' = \Omega$ we have

$$\|u\|_a \ge \sqrt{2}\, \|u\| \ge \|u\| \qquad \text{for all } u \in H^1_0(\Omega) \tag{4.2.5}$$

since $D_\Omega \le 1$ by Assumption 2.2.2.

4.2.1 Domain Decomposition Framework

We now introduce the domain decomposition framework needed for constructing one- and two-level preconditioners. The core idea is to partition the global domain into overlapping subdomains and then build local and global correction operators that accelerate convergence by improving information transfer across the entire domain.

We first partition Ω into N non-overlapping subdomains $\{\Omega_i'\}_{i=1}^N$, each resolved by the global finite element mesh $\mathcal{T}_h$. To introduce overlap, each Ω_i' is enlarged by one or more layers of mesh elements: For each overlapping Ω_i, as defined in Definition 3.2.2, we introduce the local finite element spaces

$$\widetilde{V}_i := \{v|_{\Omega_i} : v \in V^h\} \subseteq H^1(\Omega_i), \qquad V_i := \{v \in \widetilde{V}_i : v|_{\partial\Omega_i} = 0\} \subseteq H_0^1(\Omega_i).$$

Moreover, we set $H_i := \operatorname{diam}(\Omega_i)$ and $H := \max_i H_i$.

Remark 4.2.5. *The form b_{Ω_i} is generally indefinite. However, for sufficiently small H, coercivity is restored locally; see Lemma 4.3.14.*

For each $i = 1, \dots, N$, let $E_i : V_i \to V^h$ be the zero-extension operator, and let $R_i : V^h \to V_i$ be its L^2-adjoint. These operators preserve bilinear forms:

$$a_{\Omega_i}(u, v) = a(E_i u, E_i v), \qquad b_{\Omega_i}(u, v) = b(E_i u, E_i v), \qquad (u, v)_{\Omega_i} = (E_i u, E_i v).$$

In matrix form, the one-level additive Schwarz preconditioner reads

$$\mathbf{M}_{AS,1}^{-1} = \sum_{i=1}^{N} \mathbf{E}_i \mathbf{B}_i^{-1} \mathbf{R}_i, \qquad \mathbf{B}_i = \mathbf{R}_i \mathbf{B} \mathbf{E}_i,$$

where $\mathbf{E}_i, \mathbf{R}_i$ represent E_i, R_i in the global finite element basis.

To improve scalability, a coarse space $V_0 \subseteq V^h$ is added, with embedding $E_0 : V_0 \to V^h$ and adjoint R_0. The two-level preconditioner is then

$$\mathbf{M}_{AS,2}^{-1} = \sum_{i=0}^{N} \mathbf{E}_i \mathbf{B}_i^{-1} \mathbf{R}_i, \qquad \mathbf{B}_i = \mathbf{R}_i \mathbf{B} \mathbf{E}_i. \tag{4.2.6}$$

The preconditioned system takes the form

$$\mathbf{M}_{AS,2}^{-1} \mathbf{B} \mathbf{u} = \mathbf{M}_{AS,2}^{-1} \mathbf{f}.$$

For analysis, we define local projection-like operators $T_i : V^h \to V_i$ by

$$b_{\Omega_i}(T_i u, v) = b(u, E_i v) \qquad \text{for all } v \in V_i, \tag{4.2.7}$$

where we set $\Omega_0 := \Omega$, and the global operator

$$T := \sum_{i=0}^{N} E_i T_i.$$

Proposition 4.2.6 (Bootland et al. [13])**.** *For any $u, v \in V^h$, with corresponding nodal vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$,*

$$\langle \mathbf{M}_{AS,2}^{-1} \mathbf{B} \mathbf{u}, \mathbf{v} \rangle_{\mathbf{D}_k} = (Tu, v)_{1,k},$$

where $\langle \cdot, \cdot \rangle_{\mathbf{D}_k}$ is the $\mathbf{D}_k$-inner product on $\mathbb{R}^n$ and the matrix $\mathbf{D}_k$ given by

$$\mathbf{D}_k := \mathbf{A} + k^2 \mathbf{S},$$

where $\mathbf{S}_{ij} = (\phi_j, \phi_i)$; the inner product and the corresponding norm are defined by

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{D}_k} := \mathbf{y}^T \mathbf{D}_k \mathbf{x}, \qquad \|\mathbf{x}\|_{\mathbf{D}_k} := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_{\mathbf{D}_k}}. \tag{4.2.8}$$

4.2.2 The Δ_k-GenEO Coarse Space

We now specify the coarse space V_0 in (4.2.6). The Δ_k-GenEO method, introduced here as a variant of the GenEO approach [85], uses local spectral information from the overlap to capture global low-energy error modes.

Definition 4.2.7 (Degrees of freedom, cf. [85]). *For a subdomain Ω_i, define*

$$\overline{\mathrm{dof}}(\Omega_i) := \{\ell : \mathrm{supp}(\phi_\ell) \cap \Omega_i \neq \emptyset\}, \qquad \mathrm{dof}(\Omega_i) := \{\ell : \mathrm{supp}(\phi_\ell) \subseteq \overline{\Omega_i}\}.$$

Definition 4.2.8 (Partition of unity). *For each global degree of freedom j, define its multiplicity*

$$\mu_j = \#\{i : j \in \mathrm{dof}(\Omega_i)\}.$$

The local partition of unity operator $\Xi_i : \widetilde{V}_i \to V_i$ is given by

$$\Xi_i v = \sum_{j \in \mathrm{dof}(\Omega_i)} \frac{1}{\mu_j} v_j \phi_j^i \qquad \text{where} \quad v = \sum_{j \in \overline{\mathrm{dof}}(\Omega_i)} v_j \phi_j^i.$$

On each subdomain Ω_i we consider the following eigenvalue problem, which is used for the construction of the coarse space.

Definition 4.2.9 (Local spectral problem). *For each $i \in \{1, \dots, N\}$ find either $(p, \lambda) \in (\widetilde{V}_i \setminus \{0\}) \times \mathbb{R}$ such that*

$$a_{\Omega_i}(p, v) = \lambda \big(\Xi_i p, \Xi_i v\big)_{1,k,\Omega_i} \qquad \text{for all } v \in \widetilde{V}_i$$

or $(p, \lambda) \in (\ker \Xi_i) \times \{\infty\}$. We call λ and p eigenvalues and eigenfunctions respectively.

The next lemma shows that the eigenvalue problem in Definition 4.2.9 has sufficiently many eigenvectors.

Lemma 4.2.10. *Let $i \in \{1, \dots, N\}$ and set $s_i := \dim V_i$ and $n_i := \widetilde{V}_i$. There exist eigenvalues*

$$\lambda_1^i \leq \cdots \leq \lambda_{s_i}^i < \lambda_{s_i+1}^i = \cdots = \lambda_{n_i}^i = \infty \tag{4.2.9}$$

and corresponding eigenvectors $p_1^i, \dots, p_{n_i}^i$ such that $\{p_1^i, \dots, p_{n_i}^i\}$ is a basis of $\widetilde{V}_i$. The eigenvectors $p_1^i, \dots, p_{s_i}^i$ can be normalised such that

$$\big(\Xi_i p_\ell^i, \Xi_i p_m^i\big) = \delta_{\ell m}, \qquad \ell, m \in \{1, \dots, s_i\}. \tag{4.2.10}$$

Proof. Set $c_i(v, w) := (\Xi_i v, \Xi_i w)_{1,k,\Omega_i}$, $v, w \in \widetilde{V}_i$. We show that

$$\ker a_{\Omega_i} \cap \ker c_i = \{0\}, \tag{4.2.11}$$

where, e.g. $\ker a_{\Omega_i} = \{v \in \widetilde{V}_i : a_{\Omega_i}(v, w) = 0 \text{ for all } w \in \widetilde{V}_i\}$. Any function in $\ker a_{\Omega_i}$ must be constant. Since $(\cdot, \cdot)_{1,k,\Omega_i}$ is positive definite, a function in $\ker c_i$ must be in $\ker \Xi_i$. It follows from the definition of Ξ_i that

$$\ker \Xi_i = \mathrm{span}\{\phi_j^i : j \in \mathrm{dof}(\Omega_i) \setminus \mathrm{dof}(\Omega_i)\}, \tag{4.2.12}$$

and hence a function is in $\ker \Xi_i$ if and only if it vanishes on the interior nodes. As a consequence, a function in $\ker a_{\Omega_i} \cap \ker c_i$ must be the zero function. This shows that (4.2.11) holds. It also follows from (4.2.12) that $\dim \ker \Xi_i = n_i - s_i$. The assertions of the lemma now follow from [3, Lemma 3.14]. □

Remark 4.2.11. *The local spectral problem defined in [13], known as the Δ-GenEO method, is defined as,*

$$a_{\Omega_i}(p, v) = \lambda a_{\Omega_i}\big(\Xi_i p, \Xi_i v\big) \qquad \forall v \in \widetilde{V}_i$$

The Δ_k-GenEO method that we are introducing differs from this method by utilising the form $(\cdot,\cdot)_{1,k,\Omega_j}$ on the right-hand side of the problem instead of $a_{\Omega_j}(\cdot,\cdot)$.

We now introduce the Δ_k-GenEO coarse space.

Definition 4.2.12 (Δ_k-GenEO coarse space)**.** *Let the notation be as in Lemma 4.2.10 and let $(p_\ell^i, \lambda_\ell^i)$, $\ell = 1, \dots, n_i$, be the eigenpairs of the eigenvalue problem in Definition 4.2.9 satisfying (4.2.9) and (4.2.10), and let $m_i \in \{1, \dots, s_i\}$. The global coarse space is defined as*

$$V_0 := \operatorname{span}\big\{ E_i \Xi_i p_\ell^i : \ell = 1, \dots, m_i,\ i = 1, \dots, N \big\}.$$

Two quantities will play a central role:

$$\Lambda := \max_{T \in \mathcal{T}_h} \#\{\Omega_i : T \subseteq \Omega_i\}, \qquad \tau := \min_{1 \le i \le N} \lambda_{m_i+1}^i. \tag{4.2.13}$$

Here, Λ measures the maximal overlap multiplicity, while τ is the first unused eigenvalue, i.e. the spectral threshold defining the coarse space.

4.3 Statement of the Main Result and Theoretical Tools

Having introduced the two-level additive Schwarz preconditioner and the construction of the Δ_k-GenEO coarse space, we now state our main GMRES convergence result for the preconditioned Helmholtz system. Throughout this section we write

$$\Theta := \frac{1}{\tau}, \tag{4.3.1}$$

where τ is the spectral threshold defining the coarse space (see (4.2.13)). We apply GMRES with the $\mathbf{D}_k$-inner product, defined in (4.2.8), and use Elman's field-of-values framework.

Theorem 4.3.1 (GMRES convergence for the two-level preconditioned system)**.** *Assume that Assumption 2.2.2 and (2.2.9) from Proposition 2.2.3 hold, and that one has discrete well-posedness with $h < h_1$ from Lemma 2.3.1. Let H be the maximal subdomain diameter, Λ the overlap multiplicity, and τ the coarse-space threshold and let $k > 0$. Suppose that*

$$\begin{aligned} s &:= 8\Lambda\big(2 + 3\Lambda^2\Theta\big)(1 + C_{\mathrm{stab}})\, k\Theta^{1/2} < 1, \\ t &:= 6\sqrt{2}\,\Lambda\big(2 + 3\Lambda^2\Theta\big) Hk < 1. \end{aligned} \tag{4.3.2}$$

Then, for GMRES applied in the $\langle\cdot,\cdot\rangle_{\mathbf{D}_k}$ inner product to $\mathbf{M}_{AS,2}^{-1}\mathbf{B}\mathbf{u} = \mathbf{M}_{AS,2}^{-1}\mathbf{f}$, the residuals satisfy

$$\|\mathbf{r}^{(m)}\|_{\mathbf{D}_k}^2 \le \left(1 - \frac{c_1^2}{c_2^2}\right)^m \|\mathbf{r}^{(0)}\|_{\mathbf{D}_k}^2, \tag{4.3.3}$$

with

$$c_1 := \frac{1 - \max\{s, t\}}{(2 + 3\Lambda^2\Theta)}, \qquad c_2 := 18 + 8\Lambda^2. \tag{4.3.4}$$

Remark 4.3.2 (On the assumptions in Theorem 4.3.1)**.** *The assumptions of Theorem 4.3.1 are standard in the finite element analysis of Helmholtz problems. Assumption 2.2.2 is the usual uniform ellipticity and data-regularity hypothesis, while the normalisations $a_{\min} = 1$ and $D_\Omega \le 1$*

are harmless, since they may be obtained by rescaling. The stability bound (2.2.9) is the key well-posedness input. Although C_{stab} is finite whenever $k^2 \notin \{\lambda_\ell : \ell \in \mathbb{N}\}$, it may become large when k^2 lies close to the spectrum. Consequently, the theorem applies in a broad range of non-resonant situations, but the resulting parameter restrictions become more demanding when the problem is close to resonance. Finally, Lemma 2.3.1 requires the mesh to be sufficiently fine, which is the standard discretisation requirement for indefinite Helmholtz problems.

Remark 4.3.3 (On the proof ingredients)**.** *The proof follows Elman's FoV estimate, combining: (i) a lower bound on $(Tu, u)_{1,k,\Omega_i}$ in terms of $\|u\|^2_{1,k,\Omega_i}$ (via stable decomposition and coarse-space approximation with constants depending on Λ and Θ), and (ii) an upper bound on $\|Tu\|_{1,k,\Omega_i}$ (via stability of the local projectors and the coarse projector). The refined local/coarse stability and approximation estimates—proved in the next subsections—produce the improved k-dependences reflected in (4.3.2).*

Corollary 4.3.4 (Simplified k-explicit conditions)**.** *Let $k_0 > 0$ and assume that $k \geq k_0$. If the relations in (4.3.2) hold, then*

$$H \lesssim k^{-1}, \qquad (1 + C_{\mathrm{stab}})^2 \, k^2 \lesssim \tau. \tag{4.3.5}$$

Conversely, if (4.3.5) holds with small enough implicit constants in $\lesssim$, then (4.3.2) is satisfied and the convergence rate constants in (4.3.4) are independent of k, coefficient heterogeneity, and h (for $h < h_1$); so GMRES is mesh-independent and wavenumber-robust.

Proof. First assume that (4.3.2) holds. Clearly, $2 + 3\Lambda^2\Theta \geq 1$ and hence

$$(1 + C_{\mathrm{stab}})k\tau^{-1/2} = (1 + C_{\mathrm{stab}})k\Theta^{1/2} \leq (2 + 3\Lambda^2\Theta)\Lambda(1 + C_{\mathrm{stab}})k\Theta^{1/2} < \tfrac{1}{8},$$

which yields the second inequality in (4.3.5). Since $\Lambda \geq 1$, we have $Hk \leq \big(2+3\Lambda^2\Theta\big)\Lambda Hk < \frac{1}{6\sqrt{2}}$, which implies $H \lesssim k^{-1}$.

Now assume that

$$H \leq C_1 k^{-1}, \qquad (1 + C_{\mathrm{stab}})^2 \, k^2 \leq C_2 \tau$$

with $C_1, C_2 > 0$. The second relation, together with $k \geq k_0$, implies that $\Theta = \frac{1}{\tau} \leq C_2 k_0^{-2}$. We therefore have

$$s \leq 8\Lambda(2 + 3\Lambda^2 C_2 k_0^{-2})C_2^{1/2}, \tag{4.3.6}$$

$$t \leq 6\sqrt{2}\Lambda(2 + 3\Lambda^2 C_2 k_0^{-2})\Lambda C_1. \tag{4.3.7}$$

If C_1 and C_2 are small enough, then the right-hand sides of (4.3.6) and (4.3.7) are strictly less than 1, which implies that (4.3.2) holds. □

Remark 4.3.5 (Interpretation and scaling)**.** *The conditions stated in (4.3.2) are sufficient conditions obtained from a field-of-values analysis and are therefore conservative. Their purpose is to guarantee k-robust convergence of GMRES in a worst-case theoretical setting, uniformly with respect to coefficient heterogeneity and mesh size. This behaviour reflects the gap between sufficient theoretical conditions and observed performance, and is consistent with previous experience for spectral coarse spaces for indefinite problems. That said, from a practical perspective, the condition $H \lesssim k^{-1}$ corresponds to resolving each subdomain by a fixed number of points per wavelength, which is standard for Helmholtz problems (even if this condition can be relaxed in practice). This is the easier condition to interpret and enforce, since it depends linearly on Hk. Similarly, the lower bound on the spectral threshold τ ensures that coarse modes associated with low-energy components in the k-weighted norm are captured.*

Choice of spectral threshold τ The theory depends on the stability constant C_{stab} and may require increasing τ, especially when C_{stab} is large, for example when k^2 lies close to the Dirichlet spectrum. In practice, however, τ should be interpreted as a tuning parameter that controls the trade-off between coarse-space size and robustness. The local generalised eigenvalue problems typically exhibit favourable spectral separation, with a small number of low-energy modes capturing the dominant global components. As a result, robust convergence is often achieved for values of τ that are significantly smaller than those suggested by the sufficient theoretical bounds.

Roadmap of the estimates. The remainder of this section develops the stability and approximation tools used in the proof of Theorem 4.3.1:

1. *Coarse-space approximation and stable decomposition* (Lemmas 2.3.1, 4.2.3 and the Δ_k-GenEO projection bounds): yields the lower FoV distance via $(Tu,u)_{1,k,\Omega_i} \gtrsim \|u\|^2_{1,k,\Omega_i}/(1+\Lambda^2\Theta)$.
2. *Local and coarse projector stability*: provides $\|Tu\|_{1,k,\Omega_i} \lesssim (1+\Lambda^2)\|u\|_{1,k,\Omega_i}$ with the k-dependence isolated in s and t.
3. *FoV–GMRES contraction* (Elman): combining (1)–(2) gives (4.3.3) with c_1, c_2 as in (4.3.4).

4.3.1 Coarse Space Approximation and Stable Decomposition

We analyse the accuracy of the local spectral projection that maps $v \in \widetilde{V}_i$ onto the span of the m_i dominant eigenfunctions produced by the Δ_k-GenEO construction.

Lemma 4.3.6 (Local spectral projection: stability and spectral tail bound)**.** *Let the notation be as in Lemma 4.2.10 and fix $i \in \{1,\dots,N\}$. Let $(p^i_\ell, \lambda^i_\ell)$, $\ell = 1,\dots,n_i$, be the eigenpairs of Definition 4.2.9 satisfying (4.2.9) and (4.2.10). Let $m_i \in \{1\dots,s_i-1\}$ and define the projector $\Pi^i_{m_i} : \widetilde{V}_i \to \widetilde{V}_i$ by*

$$\Pi^i_{m_i} v := \sum_{\ell=1}^{m_i} \big(\Xi_i v, \Xi_i p^i_\ell\big)_{1,k,\Omega_i}\, p^i_\ell.$$

Then, for $w := v - \Pi^i_{m_i} v$, we have

$$\|w\|^2_{a_{\Omega_i}} \le \|v\|^2_{1,k,\Omega_i}, \qquad \|\Xi_i w\|^2_{1,k,\Omega_i} \le \frac{1}{\lambda^i_{m_i+1}} \|w\|^2_{a_{\Omega_i}}. \tag{4.3.8}$$

Moreover, $\Pi^i_{m_i}$ is the $(\Xi_i\cdot,\Xi_i\cdot)_{1,k,\Omega_i}$-orthogonal projector onto $\operatorname{span}\{p^i_1,\dots,p^i_{m_i}\}$, hence minimises $\|\Xi_i(v-z)\|_{1,k,\Omega_i}$ over all z in that span.

Proof. The proof of Lemma 4.2.10 shows that we can apply [3, Lemma 3.15], which yields

$$\|w\|^2_{a_{\Omega_i}} \le \|v\|^2_{a_{\Omega_i}} \le \|v\|^2_{1,k,\Omega_i}$$

and the second relation in (4.3.8). □

We often use the following standard overlap stability relations (see, eg. [58, Lemma 3.6] and [58, eq:(2.10)]):

$$\Big\| \sum_i E_i q_i \Big\|^2_{1,k} \le \Lambda \sum_i \|q_i\|^2_{1,k,\Omega_i} \qquad \text{and} \qquad \sum_i \|v|_{\Omega_i}\|^2_{1,k,\Omega_i} \le \Lambda \|v\|^2_{1,k}. \tag{4.3.9}$$

Lemma 4.3.7 (Global approximation by the Δ_k-GenEO coarse space)**.** *For $v \in V^h$ define*

$$z_0 := \sum_{i=1}^{N} E_i \, \Xi_i \big(\Pi^i_{m_i}(v|_{\Omega_i}) \big) \ \in V_0.$$

Then

$$\inf_{z \in V_0} \|v - z\|^2_{1,k} \ \leq \ \|v - z_0\|^2_{1,k} \leq \Lambda^2 \Theta \|v\|^2_{1,k}, \tag{4.3.10}$$

where Θ is as in (4.3.1).

Proof. Set $w_i := v|_{\Omega_i} - \Pi^i_{m_i}(v|_{\Omega_i})$ so that

$$v - z_0 = \sum_i E_i \, \Xi_i w_i.$$

It follows from Lemma 4.3.6 that

$$\|\Xi_i w_i\|^2_{1,k,\Omega_i} \leq \frac{1}{\lambda^i_{m_i+1}} \|w_i\|^2_{a_{\Omega_i}} \leq \Theta \|w_i\|^2_{a_{\Omega_i}} \leq \Theta \|v|_{\Omega_i}\|^2_{1,k,\Omega_i};$$

note that

$$\Theta := \max_{1 \leq i \leq N} \frac{1}{\lambda^i_{m_i+1}}.$$

Summing over i and using the second overlap stability bound in (4.3.9) we obtain

$$\sum_{i=1}^{N} \|\Xi_i w_i\|^2_{1,k,\Omega_i} \leq \Theta \sum_{i=1}^{N} \|v|_{\Omega_i}\|^2_{1,k,\Omega_i} \leq \Lambda \Theta \|v\|^2_{1,k}. \tag{4.3.11}$$

Using (4.3.9) again we arrive at

$$\|v - z_0\|^2_{1,k} = \Big\| \sum_{i=1}^{N} E_i \Xi_i w_i \Big\|^2_{1,k} \leq \Lambda \sum_{i=1}^{N} \|\Xi_i w_i\|^2_{1,k,\Omega_i} \leq \Lambda^2 \Theta \|v\|^2_{1,k},$$

which proves (4.3.10). □

Lemma 4.3.8 (Stable decomposition)**.** *Let z_0 be as in Lemma 4.3.7 and define, for $i = 1, \ldots, N$,*

$$z_i := \Xi_i \big(v|_{\Omega_i} - \Pi^i_{m_i}(v|_{\Omega_i}) \big) \ \in V_i.$$

Then

$$v = \sum_{i=0}^{N} E_i z_i, \qquad \|z_0\|^2_{1,k,\Omega_0} + \sum_{i=1}^{N} \|z_i\|^2_{1,k,\Omega_i} \leq \big(2 + 3\Lambda^2 \Theta\big) \|v\|^2_{1,k}. \tag{4.3.12}$$

Proof. Let w_i be as in the proof of Lemma 4.3.7; then $z_i = \Xi_i w_i$. The first relation in (4.3.12) follows easily from the definition of z_i. Using $(x+y)^2 \leq 2x^2 + 2y^2$ and Lemma 4.3.7 we obtain

$$\|z_0\|^2_{1,k} \leq \big(\|v\|_{1,k} + \|v - z_0\|_{1,k} \big)^2 \leq 2\|v\|^2_{1,k} + 2\|v - z_0\|^2_{1,k} \leq \big(2 + 2\Lambda^2 \Theta\big) \|v\|^2_{1,k}.$$

Adding $\sum_{i=1}^{N} \|z_i\|^2_{1,k,\Omega_i}$ on both sides and using (4.3.11) and $\Lambda \geq 1$ we arrive at the second relation in (4.3.12). □

Remark 4.3.9 (Why stable decomposition matters)**.** *The second inequality of* (4.3.12) *says that the energy of v splits stably between coarse and local components. In two-level additive Schwarz this ensures: (1) the coarse space approximates global modes well (Lemma 4.3.7), and (2)* any *v can be represented as a coarse part plus local corrections without blowing up the energy. In the SPD case these two facts lead to tight condition-number bounds; here, they lead to lower/upper bounds on the* field of values *of the preconditioned Helmholtz operator, which drive GMRES convergence.*

Remark 4.3.10 (Classical GenEO estimates)**.** *The local spectral projection bounds (Lemma 4.3.6), the global approximation estimate (Lemma 4.3.7), and the stable decomposition (Lemma 4.3.8) are classical within the GenEO framework; see [85] and the indefinite extension in [13]. We re-state them here for completeness and to fix notation in the Δ_k-GenEO setting.*

4.3.2 Local Projector: Solvability, Basic Identities, and Stability

To facilitate the analysis, we introduce, for each $i = 1, \dots, N$, the local *projector map* $P_i : V^h \to V_i$ defined by

$$(P_i u, v)_{1,k,\Omega_i} = (u, E_i v)_{1,k}, \qquad v \in V_i, \tag{4.3.13}$$

and the coarse projector $P_0 : V^h \to V_0$ defined by

$$(P_0 u, v_0)_{1,k} = (u, v_0)_{1,k}, \qquad v_0 \in V_0. \tag{4.3.14}$$

The associated global operator is

$$P := \sum_{i=0}^{N} E_i P_i : \ V^h \to V^h.$$

Lemma 4.3.11 (Well-posedness and algebraic identities)**.** *Under Assumption 2.2.2 and* (2.2.9) *from Proposition 2.2.3, the bilinear forms a_{Ω_i} and a are SPD on V_i and V_0, respectively; hence P_i $(i \geq 1)$ and P_0 are well defined by Lax–Milgram. Moreover, the following statements hold.*

1. *(Locality) For $v \in V_i$, $(u, E_i v)_{1,k} = (u|_{\Omega_i}, v)_{1,k,\Omega_i}$; so $P_i u$ is a projection of $u|_{\Omega_i}$ onto V_i.*
2. *(Orthogonal projector) P_0 is the $(\cdot,\cdot)_{1,k}$-orthogonal projector onto V_0; in particular, $\|P_0 u\|_{1,k} \leq \|u\|_{1,k}$.*
3. *(Identity) For all $u \in V^h$,*

$$(Pu, u)_{1,k} = \|P_0 u\|^2_{1,k,\Omega_0} + \sum_{i=1}^{N} \|P_i u\|^2_{1,k,\Omega_i}. \tag{4.3.15}$$

Proof (sketch). SPD and Lax–Milgram are standard. The first statement follows from the zero-extension property of E_i and the definition of $(\cdot,\cdot)_{1,k,\Omega_i}$. The second claim is immediate from (4.3.14) and standard properties of orthogonal projectors in Hilbert spaces. For the third statement, use (4.3.13) and (4.3.14) with $v = P_i u$ or $v_0 = P_0 u$ and sum over $i = 0, \dots, N$. □

The next result links the global weighted k-norm of u to the action of P and controls the size of the local images.

Proposition 4.3.12 (Inequalities for the local operators)**.** *Let $u \in V^h$. With the stable decomposition of Lemma 4.3.8 and its constant $C_{\mathrm{sd}} := 2 + 3\Lambda^2\Theta$, we have*

$$\|u\|^2_{1,k} \leq C_{\mathrm{sd}}\, (Pu, u)_{1,k} \tag{4.3.16}$$

and

$$\sum_{i=0}^{N} \|P_i u\|^2_{1,k,\Omega_i} \leq (\Lambda + 1)\|u\|^2_{1,k}. \tag{4.3.17}$$

Proof. Let $u = \sum_{i=0}^{N} E_i z_i$ be the decomposition from Lemma 4.3.8. Then, by (4.3.13) and (4.3.14),

$$\|u\|_{1,k}^2 = \left(u, \sum_{i=0}^{N} E_i z_i\right)_{1,k} = \sum_{i=0}^{N} (u, E_i z_i)_{1,k} = \sum_{i=0}^{N} (P_i u, z_i)_{1,k,\Omega_i}.$$

Cauchy–Schwarz (applied twice) and Lemma 4.3.8 give

$$\begin{aligned}\|u\|_{1,k}^2 &\le \sum_{i=0}^{N} \|P_i u\|_{1,k,\Omega_i} \|z_i\|_{1,k,\Omega_i} \le \left(\sum_{i=0}^{N} \|P_i u\|_{1,k,\Omega_i}^2\right)^{1/2} \left(\|z_0\|_{1,k}^2 + \sum_{i=1}^{N} \|z_i\|_{1,k,\Omega_i}^2\right)^{1/2} \\ &\le C_{\mathrm{sd}}^{1/2} \left(\sum_{i=0}^{N} \|P_i u\|_{1,k,\Omega_i}^2\right)^{1/2} \|u\|_{1,k}.\end{aligned}$$

Dividing by $\|u\|_{1,k}$, using the identity (4.3.15) to replace the sum by $(Pu, u)_{1,k,\Omega_i}$ and squaring the inequality we obtain (4.3.16).

For (4.3.17), we first bound the coarse part by contraction of P_0: $\|P_0 u\|_{1,k}^2 \le \|u\|_{1,k}^2$. For the overlapping parts we use (4.3.9) to obtain

$$\begin{aligned}\sum_{i=1}^{N} \|P_i u\|_{1,k,\Omega_i}^2 = \sum_{i=1}^{N} (u, E_i P_i u)_{1,k} = \left(u, \sum_{i=1}^{N} E_i P_i u\right)_{1,k} &\le \|u\|_{1,k} \left\| \sum_{i=1}^{N} E_i P_i u \right\|_{1,k} \\ &\le \Lambda^{1/2} \|u\|_{1,k} \left(\sum_{i=1}^{N} \|P_i u\|_{1,k,\Omega_i}^2\right)^{1/2},\end{aligned}$$

which implies

$$\sum_{i=1}^{N} \|P_i u\|_{1,k,\Omega_i}^2 \le \Lambda \|u\|_{1,k}^2. \tag{4.3.18}$$

Adding the coarse term we arrive at (4.3.17). □

Remark 4.3.13 (Classical facts)**.** *The definition and properties of P_i and P_0 are standard in domain decomposition; see, e.g. [89, Sec. 2.2]. We include them here to keep the presentation self-contained and to fix notation.*

We now return to the Helmholtz-specific projectors T_i defined by the indefinite form $b_{\Omega_i}(\cdot,\cdot)$ in (4.2.7). The next two lemmas give solvability (local coercivity) and a uniform $(\cdot,\cdot)_{1,k,\Omega_i}$-stability bound.

Lemma 4.3.14 (Solvability of T_i)**.** *If $Hk < \sqrt{2}$, then b_{Ω_i} is coercive on V_i and the operators $T_i : V^h \to V_i$ defined by*

$$b_{\Omega_i}(T_i u, v) = b(u, E_i v) \qquad \text{for all } v \in V_i$$

are well posed for $i = 1, \dots, N$.

Proof. By Friedrichs' inequality (4.2.4) we have $\|u\|_{\Omega_i} \le \frac{H}{\sqrt{2}} \|u\|_{a_{\Omega_i}}$ and hence

$$b_{\Omega_i}(u,u) = a_{\Omega_i}(u,u) - k^2 \|u\|_{\Omega_i}^2 \ge \left(1 - \tfrac{k^2 H^2}{2}\right) \|u\|_{a_{\Omega_i}}^2.$$

Thus b_{Ω_i} is coercive for $Hk < \sqrt{2}$, and Lax–Milgram applies. □

Before we proceed, we state the following boundedness property of $b_{\Omega'}$, which is a simple consequence of the Cauchy–Schwarz inequality.

Lemma 4.3.15.

$$|b_{\Omega'}(u,v)| \le \|u\|_{1,k,\Omega'}\|v\|_{1,k,\Omega'}. \tag{4.3.19}$$

Proof. Using the definitions of $a_{\Omega'}$ and $b_{\Omega'}$ and the Cauchy–Schwarz inequality we obtain

$$\begin{aligned}|b_{\Omega'}(u,v)| &= |a_{\Omega'}(u,v) - k^2(u,v)_{\Omega'}| \le \|u\|_{a_{\Omega'}}\|v\|_{a_{\Omega'}} + k^2\|u\|_{\Omega'}\|v\|_{\Omega'} \\ &\le \sqrt{\|u\|^2_{a_{\Omega'}} + k^2\|u\|^2_{\Omega'}}\sqrt{\|v\|^2_{a_{\Omega'}} + k^2\|v\|^2_{\Omega'}} = \|u\|_{1,k,\Omega'}\|v\|_{1,k,\Omega'},\end{aligned}$$

which proves (4.3.19). □

Lemma 4.3.16 (Stability of T_i)**.** *If $Hk \le 1/\sqrt{2}$, then, for all $u \in V^h$ and $i = 1,\dots,N$,*

$$\|T_i u\|_{1,k,\Omega_i} \le 2\|u\|_{1,k,\Omega_i}.$$

Proof. It follows from (4.3.19) and (4.2.4) that

$$\begin{aligned}\|T_i u\|^2_{1,k,\Omega_i} &= b_{\Omega_i}(T_i u, T_i u) + 2k^2\|T_i u\|^2_{\Omega_i} = b(u, E_i T_i u) + 2k^2\|T_i u\|^2_{\Omega_i} \\ &= b_{\Omega_i}(u|_{\Omega_i}, T_i u) + 2k^2\|T_i u\|^2_{\Omega_i} \\ &\le \|u|_{\Omega_i}\|_{1,k,\Omega_i}\|T_i u\|_{1,k,\Omega_i} + k^2H^2\|T_i u\|^2_{1,k,\Omega_i}.\end{aligned}$$

Rearranging this inequality we obtain

$$\left(1 - k^2H^2\right)\|T_i u\|_{1,k,\Omega_i} \le \|u\|_{1,k,\Omega_i},$$

which, together with the assumption $Hk \le \frac{1}{\sqrt{2}}$, yields (4.3.12). □

4.3.3 Coarse Projector T_0: Solvability and Stability

We define the coarse (global) Helmholtz projector $T_0 : V^h \to V_0$ by

$$b(T_0 u, v_0) = b(u, v_0) \qquad \text{for all } v_0 \in V_0. \tag{4.3.20}$$

Thus $e_0 := T_0 u - u$ satisfies $b(e_0, v_0) = 0$ for all $v_0 \in V_0$.

Lemma 4.3.17 (Solvability of T_0)**.** *Suppose that the spectral threshold Θ satisfies*

$$\sqrt{2}k\Lambda\Theta^{1/2}(1 + C_{\mathrm{stab}}) < 1. \tag{4.3.21}$$

Then there exists $h_1 > 0$ such that, for all $h < h_1$, the problem (4.3.20) is well posed, i.e. T_0 is uniquely defined on V^h.

Proof. We argue by contradiction. Assume that there exists $w_0 \in V_0 \setminus \{0\}$ with $b(w_0, z) = 0$ for all $z \in V_0$. Let $w \in H^1_0(\Omega)$ solve $b(w, v) = (w_0, v)$ for all $v \in H^1_0(\Omega)$ (such a w exists by (2.2.9)) and let $w_h \in V^h$ solve $b(w_h, v) = (w_0, v)$ for all $v \in V^h$. Taking $v = w_0$ in the latter relation and using Lemma 4.3.15 we obtain, for all $z \in V_0$,

$$\|w_0\|^2 = b(w_h, w_0) = b(w_h - z, w_0) \le \|w_h - z\|_{1,k}\|w_0\|_{1,k}.$$

Minimising over $z \in V_0$ and using the global Δ_k-GenEO approximation (Lemma 4.3.7) with $v = w_h$ we get

$$\|w_0\|^2 \le \Lambda\Theta^{1/2}\|w_0\|_{1,k}\|w_h\|_{1,k}. \tag{4.3.22}$$

Pick $z = w_0$ in $b(w_0, z) = 0$ to obtain $(w_0, w_0)_{1,k} = 2k^2\|w_0\|^2$, which implies $\|w_0\|_{1,k} = \sqrt{2}k\|w_0\|$. (2.2.9) and Remark 4.2.2 yield

$$\|w_h\|_{1,k} \le \|w\|_{1,k} + \|w - w_h\|_{1,k} \le C_{\mathrm{stab}}\|w_0\| + \|w - w_h\|_{1,k} \le (1 + C_{\mathrm{stab}})\|w_0\|. \tag{4.3.23}$$

Altogether, we have

$$\|w_0\|^2 \le \sqrt{2}k\Lambda\Theta^{1/2}(1 + C_{\mathrm{stab}})\|w_0\|^2,$$

which contradicts (4.3.21) since $w_0 \neq 0$. Hence T_0 is well defined. □

Lemma 4.3.18 (Stability of T_0)**.** *Assume that* (4.3.21) *holds and that $h < h_1$ as above. Then, for all $u \in V^h$,*

$$\|T_0 u - u\| \leq \Lambda \Theta^{1/2}(1 + C_{\text{stab}})\|T_0 u - u\|_{1,k}. \tag{4.3.24}$$

Moreover, if

$$2k\Lambda\Theta^{1/2}(1 + C_{\text{stab}}) \leq \tfrac{1}{2}, \tag{4.3.25}$$

then

$$\|T_0 u - u\|_{1,k} \leq 2\|u\|_{1,k}. \tag{4.3.26}$$

Proof. Let $e_0 := u - T_0 u$. Solve $b(w_h, v) = (e_0, v)$ for $w_h \in V^h$. Since $b(e_0, z) = 0$ for $z \in V_0$ by the definition of T_0, we have

$$\|e_0\|^2 = b(w_h, e_0) = b(w_h - z, e_0) \qquad \text{for all } z \in V_0.$$

We proceed exactly as in the proof of Lemma 4.3.17 (with w_0 replaced by e_0, cf. (4.3.22)) to obtain

$$\|e_0\|^2 \leq \Lambda\Theta^{1/2}\|w_h\|_{1,k}\,\|e_0\|_{1,k}.$$

Using (4.3.23) (again with w_0 replaced by e_0) we find that

$$\|e_0\| \leq \Lambda\Theta^{1/2}(1 + C_{\text{stab}})\|e_0\|_{1,k},$$

which is (4.3.24).

To prove (4.3.26), let P_0 be as in (4.3.14). Since $P_0 u - T_0 u \in V_0$, we have $b(e_0, P_0 u - T_0 u) = 0$, which, together with the fact that P_0 is orthogonal with respect to $(\cdot,\cdot)_{1,k}$, yields

$$\begin{aligned}
\|e_0\|_{1,k}^2 &= b(e_0, e_0) + 2k^2(e_0, e_0) = b(e_0, u - T_0 u) - b(e_0, P_0 u - T_0 u) + 2k^2(e_0, e_0) \\
&= b(e_0, u - P_0 u) + 2k^2(e_0, u - T_0 u) \\
&= (e_0, u - P_0 u)_{1,k} - 2k^2(e_0, u - P_0 u) + 2k^2(e_0, u - T_0 u) \\
&= (e_0, u - P_0 u)_{1,k} + 2k^2(e_0, P_0 u - T_0 u) \\
&\leq \|e_0\|_{1,k}\|u - P_0 u\|_{1,k} + 2k^2\|e_0\|\,\|P_0 u - T_0 u\| = \|e_0\|_{1,k}\|u - P_0 u\|_{1,k} + 2k^2\|e_0\|\,\|P_0 e_0\| \\
&\leq \|e_0\|_{1,k}\|u - P_0 u\|_{1,k} + 2k\|e_0\|\,\|P_0 e_0\|_{1,k} \leq \|e_0\|_{1,k}\|u\|_{1,k} + 2k\|e_0\|\,\|e_0\|_{1,k}.
\end{aligned}$$

We can now use (4.3.24) and (4.3.25) to obtain

$$\|e_0\|_{1,k} \leq \|u\|_{1,k} + 2k\Lambda\Theta^{1/2}(1 + C_{\text{stab}})\|e_0\|_{1,k} \leq \|u\|_{1,k} + \tfrac{1}{2}\|e_0\|_{1,k},$$

which, in turn, implies (4.3.26). □

Remark 4.3.19 (Comparison and provenance)**.** *Conditions* (4.3.21) *and* (4.3.25) *strengthen the coarse-level assumptions in [13] by leveraging the sharper GenEO approximation (Lemma 4.3.7). The proof strategy is classical: a contradiction argument for solvability (coercivity of b on V_0) and a Petrov–Galerkin stability estimate for* (4.3.24).

Corollary 4.3.20 (τ-form of the coarse conditions)**.** *Let $\Theta = 1/\tau$ and $\Lambda \geq 1$. Then condition* (4.3.25) *is equivalent to*

$$\tau \geq 16\Lambda^2(1 + C_{\text{stab}})^2 k^2.$$

Remark 4.3.21 (Interpretation of the coarse-level condition)**.** *Conditions* (4.3.21) *and* (4.3.25) *arise from the requirement that the coarse-level Helmholtz projector T_0 be well defined and stable. They ensure coercivity of the bilinear form $b(\cdot,\cdot)$ on the coarse space V_0, and are therefore necessary ingredients in the field-of-values analysis underpinning Theorem 4.3.1.Note that the*

first inequality in (4.3.2) *implies* (4.3.25)*, which, in turn, implies* (4.3.21)*. All three conditions impose constraints on the spectral threshold* τ *through the quantity* $\Theta = 1/\tau$ *and ultimately stem from the same coarse-space approximation estimate provided by Lemma 4.3.7. Corollary 4.3.20 makes this connection explicit by showing that* (4.3.21) *is equivalent to a quadratic lower bound on* τ*, namely* $\tau \gtrsim (1 + C_{\mathrm{stab}})^2 k^2$*, which is consistent with the simplified* k*-explicit conditions in* (4.3.5)*. From a practical perspective, condition* (4.3.21) *should again be interpreted as a sufficient condition. It guarantees that all coarse modes responsible for potential instabilities are captured, but it does not require that the bound on* τ *be saturated in practice. As observed in the numerical experiments, stable and efficient convergence is typically achieved for values of* τ *that are significantly smaller than the worst-case theoretical bound suggested by* (4.3.21)*.*

On the size of the coarse space and the τ-bound The conditions in Corollary 4.3.20 constrain the *threshold* τ, not the number of kept local modes m_i directly. Translating a bound on τ into a bound on m_i requires information about the spectral density $\{\lambda_\ell^i\}_{\ell \geq 1}$ of the local problems, which depends on geometry, overlap, H, and coefficients. In the absence of such structure, the sufficient bound $\tau \gtrsim (1 + C_{\mathrm{stab}})^2 k^2$ can be read as 'include all local modes whose Rayleigh quotients are $\lesssim k^2$', which may *appear* to force a rapidly growing coarse space as k increases. In practice, two points mitigate this pessimism:

- **Sufficiency vs. necessity.** The k^2 scaling is a *worst-case sufficient* condition for field-of-values robustness. Empirically, with $H \sim c/k$ and moderate overlap, the spectrum of $(\Xi_i \cdot, \Xi_i \cdot)_{1,k,\Omega_i}$ typically exhibits an early cluster/gap, so only a modest number of modes per subdomain are needed to stabilise GMRES. Thus, m_i often grows much more slowly than what a literal reading of the τ-inequality would suggest.
- **What matters for GMRES.** The theory guarantees that unresolved 'bad' components (those with $\lambda_\ell^i \leq \tau$) are transferred to the coarse level. Once the coarse space captures these, further enrichment has diminishing returns. For higher k, local coercivity is already ensured by $H \sim c/k$ (cf. Lemma 4.3.14), and the residual high-frequency components are effectively damped by the local solves.

Remark 4.3.22 (Limits of SPD coarse spaces at very high frequency)**.** *The present analysis deliberately uses an SPD coarse space for an indefinite problem. While our bounds certify robustness under* $H \sim k^{-1}$ *and* $\tau \gtrsim (1 + C_{\mathrm{stab}})^2 k^2$*, they are not intended to be* tight *at very large* k*. Beyond some regime, SPD coarse spaces may cease to be the most effective choice; shifted/indefinite coarse operators or multilevel variants are natural successors. Our results should therefore be read as clarifying the range over which SPD GenEO-type spaces remain effective, and as a baseline for future indefinite coarse spaces.*

For ease of reference, we summarise in Table 4.1 the sufficient k-explicit conditions on the subdomain diameter H and the spectral threshold τ required to guarantee robust GMRES convergence for the original Δ-GenEO method and for the proposed Δ_k-GenEO variant. The table highlights the substantial reduction in wavenumber dependence achieved by the k-dependent formulation, while emphasising that all bounds remain sufficient rather than necessary.

Table 4.1: Comparison of sufficient k-explicit conditions for robust GMRES convergence using Δ-GenEO and Δ_k-GenEO coarse spaces.

	Δ-GenEO	Δ_k-GenEO
Subdomain diameter H	$H \lesssim k^{-2}$	$H \lesssim k^{-1}$
Spectral threshold τ	$\tau \gtrsim k^8$	$\tau \gtrsim k^2$

While our bounds certify robustness under $H \sim k^{-1}$ and $\tau \gtrsim (1 + C_{\mathrm{stab}})^2 k^2$, they are not intended to be *tight* for very large k. Beyond some regime, SPD coarse spaces may cease to be the most effective choice; shifted/indefinite coarse operators or multilevel variants are natural successors. Our results should therefore be read as clarifying the range over which SPD GenEO-type spaces remain effective, and as a baseline for future indefinite coarse spaces.

4.4 Proof of the Main Result

Before proving the main result we need an additional technical lemma.

Lemma 4.4.1 (FoV lower bound and operator-norm bound)**.** *Under the assumptions of Theorem 4.3.1 and the constraints (4.3.2), the projection operator $T = \sum_{i=0}^{N} E_i T_i$ satisfies, for all $u \in V^h$,*

$$c_1 \|u\|_{1,k}^2 \leq (Tu, u)_{1,k} \tag{4.4.1}$$

and

$$\|Tu\|_{1,k} \leq \sqrt{c_2}\, \|u\|_{1,k}, \tag{4.4.2}$$

where c_1, c_2 are as in (4.3.4).

Proof. Step 1 (FoV template: relate $\|u\|_{1,k}^2$ to $(Tu, u)_{1,k}$ plus two error terms). With $C_{\mathrm{sd}} = 2 + 3\Lambda^2\Theta$ we obtain from Proposition 4.3.12 that

$$\|u\|_{1,k}^2 \leq C_{\mathrm{sd}}\, (Pu, u)_{1,k} = C_{\mathrm{sd}} \sum_{i=0}^{N} (E_i P_i u, u)_{1,k}.$$

For each $i \in \{0, \ldots, N\}$ we use the relation $b(\cdot,\cdot) = a(\cdot,\cdot) - k^2(\cdot,\cdot)$ and the definitions of T_i and P_i to obtain

$$\begin{aligned}
(E_i P_i u, u)_{1,k} &= (u, E_i P_i u)_{1,k} = b(u, E_i P_i u) + 2k^2 (u, E_i P_i u) = b_{\Omega_i}(T_i u, P_i u) + 2k^2 (u, E_i P_i u) \\
&= (T_i u, P_i u)_{1,k,\Omega_i} - 2k^2 (T_i u, P_i u)_{\Omega_i} + 2k^2 (u, E_i P_i u) \\
&= (T_i u, P_i u)_{1,k,\Omega_i} - 2k^2 (E_i T_i u, E_i P_i u) + 2k^2 (u, E_i P_i u) \\
&= (E_i T_i u, u)_{1,k} + 2k^2 (u - E_i T_i u, E_i P_i u),
\end{aligned}$$

which yields

$$\|u\|_{1,k}^2 \leq C_{\mathrm{sd}} \left((Tu, u)_{1,k} + 2k^2 (u - T_0 u, P_0 u) + 2k^2 \sum_{i=1}^{N} (u - E_i T_i u, E_i P_i u) \right). \tag{4.4.3}$$

We now bound the two perturbation terms.

Step 2 (Coarse term bound via T_0 stability). It follows from Cauchy–Schwarz, (4.2.5) and Lemma 4.3.18 that

$$\begin{aligned}
|2k^2 (u - T_0 u, P_0 u)| &\leq 2k^2 \|T_0 u - u\|\, \|P_0 u\| \leq 2k \|T_0 u - u\|\, \|P_0 u\|_{1,k} \\
&\leq 2k \Lambda \Theta^{1/2} (1 + C_{\mathrm{stab}}) \|T_0 u - u\|_{1,k} \|P_0 u\|_{1,k}.
\end{aligned}$$

The first inequality in (4.3.2) implies that (4.3.25) holds. Also we have the fact that P_0 is an orthogonal projection with respect to $(\cdot,\cdot)_{1,k}$, and hence (4.3.26) yields

$$|2k^2 (u - T_0 u, P_0 u)| \leq 4k \Lambda \Theta^{1/2} (1 + C_{\mathrm{stab}}) \|u\|_{1,k} \|P_0\|_{1,k} \leq 4k \Lambda \Theta^{1/2} (1 + C_{\mathrm{stab}}) \|u\|_{1,k}^2. \tag{4.4.4}$$

Step 3 (Local terms bound via T_i stability and overlap). For $i \geq 1$, we have, by Cauchy–Schwarz, Friedrichs on Ω_i, and Lemma 4.3.16,

$$\begin{aligned}
|2k^2(u - E_iT_iu, E_iP_iu)| &\leq 2k^2 \|u - T_iu\|_{\Omega_i} \|P_iu\|_{\Omega_i} \leq 2k \|u - T_iu\|_{1,k,\Omega_i} \|P_iu\|_{\Omega_i} \\
&\leq 2k\Big(\|u\|_{1,k,\Omega_i} + \|T_iu\|_{1,k,\Omega_i}\Big) \|P_iu\|_{\Omega_i} \\
&\leq Hk\sqrt{2}\Big(\|u\|_{1,k,\Omega_i} + \|T_iu\|_{1,k,\Omega_i}\Big)\|P_iu\|_{1,k,\Omega_i} \\
&\leq Hk\sqrt{2}\Big(\|u\|_{1,k,\Omega_i} + 2\|u\|_{1,k,\Omega_i}\Big)\|P_iu\|_{1,k,\Omega_i} \\
&= 3Hk\sqrt{2}\,\|u\|_{1,k,\Omega_i}\|P_iu\|_{1,k,\Omega_i}.
\end{aligned}$$

Summing over $i \in \{1, \dots, N\}$, applying Cauchy–Schwarz and using (4.3.9), (4.3.18) we obtain

$$\begin{aligned}
\left|2k^2 \sum_{i=1}^{N} (u - E_iT_iu, E_iP_iu)\right| &\leq 3Hk\sqrt{2} \sum_{i=1}^{N} \|u\|_{1,k,\Omega_i}\|P_iu\|_{1,k,\Omega_i} \\
&\leq 3Hk\sqrt{2} \left(\sum_{i=1}^{N} \|u\|^2_{1,k,\Omega_i}\right)^{\frac{1}{2}} \left(\sum_{i=1}^{N} \|P_iu\|^2_{1,k,\Omega_i}\right)^{\frac{1}{2}} \\
&= 3\Lambda Hk\sqrt{2}\,\|u\|^2_{1,k}. \qquad (4.4.5)
\end{aligned}$$

Step 4 (Combine and absorb). Inserting (4.4.4) and (4.4.5) into (4.4.3) we obtain

$$\begin{aligned}
\|u\|^2_{1,k} &\leq C_{\mathrm{sd}}\Big((Tu, u)_{1,k} + 4k\Lambda\Theta^{1/2}(1 + C_{\mathrm{stab}})\|u\|^2_{1,k} + 3\Lambda Hk\sqrt{2}\,\|u\|^2_{1,k}\Big) \\
&= C_{\mathrm{sd}}(Tu, u)_{1,k} + \tfrac{1}{2}(s + t)\|u\|^2_{1,k} \leq C_{\mathrm{sd}}(Tu, u)_{1,k} + \max\{s, t\}\|u\|^2_{1,k}.
\end{aligned}$$

Since $s, t < 1$, this implies (4.4.1) with $c_1 = \dfrac{1 - \max\{s, t\}}{C_{\mathrm{sd}}}$.

Operator-norm bound. We write $Tu = T_0u + \sum_{i=1}^{N} E_iT_iu$ and use the overlap stability (4.3.9) to find

$$\|Tu\|^2_{1,k} \leq 2\Big(\|T_0u\|^2_{1,k} + \Big\|\sum_{i=1}^{N} E_iT_iu\Big\|^2_{1,k}\Big) \leq 2\|T_0u\|^2_{1,k} + 2\Lambda \sum_{i=1}^{N} \|T_iu\|^2_{1,k,\Omega_i}. \qquad (4.4.6)$$

For the first term on the right-hand side, we use Cauchy–Schwarz and Lemma 4.3.18, which yield

$$\begin{aligned}
\|T_0u\|^2_{1,k} &= (T_0u, T_0u)_{1,k} = (T_0u - u, T_0u)_{1,k} + (u, T_0u)_{1,k} \\
&\leq \|T_0u - u\|_{1,k}\|T_0u\|_{1,k} + \|u\|_{1,k}\|T_0u\|_{1,k} \\
&\leq 2\|u\|_{1,k}\|T_0u\|_{1,k} + \|u\|_{1,k}\|T_0u\|_{1,k} = 3\|u\|_{1,k}\|T_0u\|_{1,k},
\end{aligned}$$

and hence

$$\|T_0u\|_{1,k} \leq 3\|u\|_{1,k}. \qquad (4.4.7)$$

For the second term on the right-hand we use Lemma 4.3.16 to obtain

$$\sum_{i=1} \|T_iu\|^2_{1,k,\Omega_i} \leq \sum_{i=1} 4\big\|u|_{\Omega_i}\big\|^2_{1,k,\Omega_i} \leq 4\Lambda\|u\|^2_{1,k}. \qquad (4.4.8)$$

Combining (4.4.7) and (4.4.8) with (4.4.6) we arrive at

$$\|Tu\|^2_{1,k} \leq 2 \times 9\|u\|^2_{1,k} + 2 \times 4\Lambda^2\|u\|^2_{1,k} = (18 + 8\Lambda^2)\|u\|^2_{1,k},$$

which is (4.4.2). □

We are now able to finish the proof of Theorem 4.3.1.

Proof of Theorem 4.3.1. Let $\mathcal{M} := \mathbf{M}_{AS,2}^{-1}\mathbf{B}$ and endow $\mathbb{R}^n$ with the $\mathbf{D}_k$-inner product $\langle \cdot, \cdot \rangle_{\mathbf{D}_k}$ and norm $\| \cdot \|_{\mathbf{D}_k}$. By Proposition 4.2.6, for all $u, v \in V^h$ with coefficient vectors $\mathbf{u}, \mathbf{v}$,

$$(Tu, v)_{1,k} = \langle \mathcal{M}\mathbf{u}, \mathbf{v} \rangle_{\mathbf{D}_k}.$$

Taking $v = u$ and using (4.4.1) we obtain the FoV distance (from the origin)

$$\inf_{\|\mathbf{u}\|_{\mathbf{D}_k}=1} \operatorname{Re} \langle \mathcal{M}\mathbf{u}, \mathbf{u} \rangle_{\mathbf{D}_k} \geq c_1.$$

Next, (4.4.2) gives

$$\|\mathcal{M}\mathbf{u}\|_{\mathbf{D}_k}^2 = \langle \mathcal{M}\mathbf{u}, \mathcal{M}\mathbf{u} \rangle_{\mathbf{D}_k} = (Tu, Tu)_{1,k} = \|Tu\|_{1,k}^2 \leq c_2 \|u\|_{1,k}^2 = c_2 \|\mathbf{u}\|_{\mathbf{D}_k}^2;$$

hence the operator norm satisfies $\|\mathcal{M}\|_{\mathbf{D}_k} \leq \sqrt{c_2}$.

By the Elman field-of-values estimate for preconditioned GMRES [35], the residuals satisfy

$$\|\mathbf{r}^{(m)}\|_{\mathbf{D}_k}^2 \leq \left(1 - \left(\frac{\delta}{\beta}\right)^2\right)^m \|\mathbf{r}^{(0)}\|_{\mathbf{D}_k}^2, \qquad \delta := \inf_{\|\mathbf{u}\|_{\mathbf{D}_k}=1} \operatorname{Re} \langle \mathcal{M}\mathbf{u}, \mathbf{u} \rangle_{\mathbf{D}_k}, \quad \beta := \|\mathcal{M}\|_{\mathbf{D}_k}.$$

Using $\delta \geq c_1$ and $\beta^2 \leq c_2$ we obtain

$$\|\mathbf{r}^{(m)}\|_{\mathbf{D}_k}^2 \leq \left(1 - \frac{c_1^2}{c_2}\right)^m \|\mathbf{r}^{(0)}\|_{\mathbf{D}_k}^2,$$

which is the claimed convergence bound (with the stated c_1 and c_2). □

4.5 Numerical Results

The aim of this section is to provide a comprehensive assessment of the performance and robustness of the Δ_k-GenEO coarse space when combined with the two-level additive Schwarz (AS) preconditioner, and to compare its behaviour with that of the classical one-level AS method. While theoretical analysis provides insight into the spectral properties of these preconditioners, the actual performance in practice is influenced by a number of factors that are difficult to quantify a priori, such as the frequency, the domain partitioning, the heterogeneity of the medium, and the chosen eigenvalue threshold for coarse space construction.

The numerical experiments presented in this section are not designed to strictly enforce the sufficient conditions obtained in the theoretical results. Instead, they deliberately explore a range of regimes that extend beyond these conservative bounds in order to assess the practical robustness of the proposed Δ_k-GenEO method. Many experiments intentionally violate one or more of the theoretical assumptions. The resulting performance illustrates that the method remains effective well beyond the strict regime guaranteed by the analysis, thereby highlighting the gap between sufficient theoretical conditions and observed behaviour in practice.

In what follows, we conduct a series of numerical experiments. Our tests are restricted to two representative cases:

1. **Homogeneous test case.** A benchmark posed on the unit square with homogeneous Dirichlet boundary conditions and a single interior point source. This case allows for the performance of the preconditioners to be tested in a controlled setting.

2. **Heterogeneous test case.** A layered medium with strong contrasts, designed to probe the robustness of the method under medium heterogeneity, and to examine how the coarse space size and convergence are affected in more challenging settings.

Both test cases allow us to study how iteration counts and coarse space dimension scale with respect to the number of subdomains and the frequency. The influence of the following parameters is reported:

- subdomain diameter measured in wavelengths,
- global and per-subdomain coarse space dimension,
- eigenvalue threshold τ for coarse space construction.

4.5.1 Implementation Details

All computations are performed in `FreeFEM` using the `ffddm` framework [91], which provides highly flexible Message Passing Interface (MPI)-parallel scripts for domain decomposition methods.

Domain partition and overlap The computational domain Ω is partitioned into N non-overlapping square subdomains $\{\Omega_i'\}_{i=1}^N$. Unless otherwise stated, an overlap is introduced by extending each Ω_i' by a minimum number of layers of finite elements, yielding overlapping subdomains $\{\Omega_i\}_{i=1}^N$.

Setup: one-level preconditioner In the one-level AS method, we perform the following

1. Assemble the local stiffness and mass matrices on Ω_i corresponding to the bilinear forms a_{Ω_i} and b_{Ω_i}.
2. Compute and store a sparse LU factorization of the resulting local system using `MUMPS`.

Setup: two-level (Δ_k-GenEO) coarse space The two-level method augments the one-level setup with a global coarse correction space. On each MPI rank i, that is, on each parallel process, we:

1. Assemble the local generalised eigenproblem where Ξ_i is the partition of unity operator.
2. Solve the eigenproblem using `SLEPc`, and select the first m_i eigenmodes according to a fixed spectral threshold ($\lambda_\ell^i \leq \tau$).
3. Form the corresponding coarse basis vectors $E_i \Xi_i(p_\ell^i)$ and contribute to the global coarse operator $\mathbf{B}_0$, defined as the $b(\cdot,\cdot)$-form restricted to V_0.

Software stack For the numerical linear algebra components, we employ the sparse direct solver `MUMPS` [1] to factorise both local and coarse matrices. The local spectral problems arising in the construction of the Δ_k-GenEO coarse space are solved using the `SLEPc` library [61], which provides scalable eigensolvers for large sparse eigenvalue problems. Iterative solutions are obtained with GMRES. Unless otherwise stated, GMRES is not restarted; iterations are terminated once the relative residual is reduced below 10^{-6}, or the number of iterations reaches 200.

Preconditioner application (solve phase) Each iteration of the preconditioned GMRES method involves:

1. *Local solves:* one forward/backward substitution with the stored LU factorization on each Ω_i, executed in parallel across all MPI ranks.

2. *Coarse solve:* one forward/backward substitution with the factorised coarse matrix $\mathbf{B}_0$, distributed over the P coarse ranks, accompanied by the necessary global communication (scatter and gather).

Algorithmic complexity. From a parallel complexity perspective, the implementation naturally separates into a *setup phase* and an *application phase*, whose scaling behaviour is qualitatively different. During the setup phase, all local tasks—assembly of local matrices, solution of the local generalised eigenvalue problems, and factorisations—are embarrassingly parallel across subdomains. The only global communication required at this stage is the assembly of the coarse operator $\mathbf{B}_0$, whose size scales with the coarse space dimension $\dim V_0$.

During the application phase, each GMRES iteration consists of fully parallel local solves on the subdomains and a single coarse-level solve, accompanied by standard global scatter/gather communication. This structure is typical of two-level additive Schwarz methods. As a result, the dominant communication costs are associated with the coarse solve and are governed by the size of the coarse space rather than the fine-grid problem size.

We emphasise that the present implementation is not optimised and does not aim to fully exploit all available parallelisation, communication, or solver-tuning strategies. Consequently, wall-clock timings, communication overheads, and strong or weak scaling results would be highly implementation- and hardware-dependent and would not be representative of the intrinsic capabilities of the method. For this reason, the numerical study focuses on algorithmic complexity and qualitative scaling behaviour—such as iteration counts, coarse space growth, and robustness with respect to frequency and heterogeneity—which are the primary indicators of performance for spectral coarse spaces.

It should be noted that due to hardware limitations, we have not been able to utilise full parallelisation. Due to the restricted number of CPU cores available, we enabled oversubscription, which allows more MPI ranks than physical CPU cores, meaning multiple ranks time-share the same core. This preserves functionality, but can reduce performance.

Homogeneous Problem

In this part we base the numerical experiments on the model problem (4.1.1) in 2D, defined on the unit square $\Omega = (0,1)^2$ and with $A(\mathbf{x}) = 1$. We impose Dirichlet boundary conditions on all sides of the domain. A point source is located in the centre of the domain at $(\frac{1}{2}, \frac{1}{2})$ and provides the forcing function f. The point source is numerically modelled by a Gaussian function: $f(x,y) = 10^4 \exp(-10^3[(x-\frac{1}{2})^2 + (y-\frac{1}{2})^2])$. A schematic of this model problem is found in Figure 4.1. We note that the specific choice of a Gaussian approximation for the point source is made for numerical convenience. Since the construction of the Δ_k-GenEO coarse space and the associated preconditioner are independent of the right-hand side, we do not expect qualitatively different behaviour for other smooth, localised, or volume source terms.

To discretise the problem, we triangulate Ω using a Cartesian grid with spacing h and alternating diagonals to form a simplicial mesh. The local wavenumber, k is constant and the wavelength used to measure geometrical parameters is $\lambda = 2\pi/k$. The discrete problem (4.2.3) is assembled using a $\mathbb{P}_1$ finite element approximation on this mesh. To mitigate the *pollution effect*, we choose the wavenumber k and the mesh size h simultaneously so that the dimensionless quantity kh remains sufficiently small. In practice, this is enforced by fixing a minimum number of grid points per wavelength λ. Here, we ensure that $kh \lesssim 0.1$, which corresponds to at least 25π points per wavelength λ. This is widely considered a very fine mesh, with most real world applications accepting 10 points per wave length. However, this value has been chosen to ensure an accurate solution whilst maintaining a solvable problem size.

In Table 4.2 we show the dimension of the fine mesh (n), the diameter of the domain measured in wavelengths (L), the diameter of the sub-domains measured in wavelengths (H), the number of local degrees of freedom (n_{loc}), the GMRES iteration count (It) for the one-level and two level

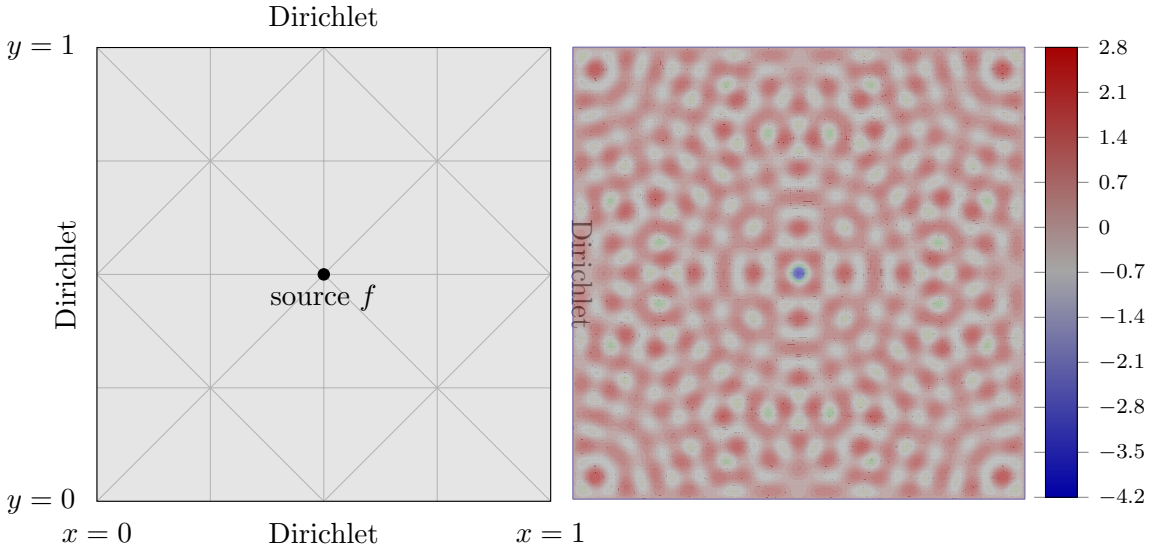


Figure 4.1: Schematic of the homogeneous problem with example triangular mesh (left). The solution for the homogeneous problem with $k = 100$ (right).

							One Level	Δ-GenEO			Δ_k-GenEO		
N	k	h^{-1}	n	L	H	n_{loc}	It	It	CS	CS_{loc}	It	CS	CS_{loc}
16	20	240	58081	3.18	0.82	3969	41	36	400	25	22	272	17
	40	480	231361	6.37	1.62	15129	107	92	804	50	73	580	36
	60	720	519841	9.55	2.41	33489	153	156	1232	77	78	976	61
	80	960	923521	12.73	3.21	59049	>200	>200	1640	102	155	1424	89
	100	1200	1442401	15.92	4.01	91809	>200	>200	2052	128	170	1932	121
36	20	240	58081	3.18	0.56	1849	79	23	684	19	23	428	12
	40	480	231361	6.37	1.09	6889	>200	130	1356	38	112	860	24
	60	720	519841	9.55	1.62	15129	>200	>200	2020	56	144	1444	40
	80	960	923521	12.73	2.15	26569	>200	>200	2740	76	>200	2028	56
	100	1200	1442401	15.92	2.68	41209	>200	>200	3428	95	>200	2752	76
64	20	240	58081	3.18	0.42	1089	92	22	896	14	23	608	10
	40	480	231361	6.37	0.82	3969	175	133	1856	29	57	1248	20
	60	720	519841	9.55	1.22	8649	>200	>200	2848	44	181	1884	29
	80	960	923521	12.73	1.62	15129	>200	>200	3780	59	>200	2692	42
	100	1200	1442401	15.92	2.02	23409	>200	>200	4768	74	>200	3488	54
100	20	240	58081	3.18	0.34	729	96	21	1176	12	22	684	7
	40	480	231361	6.37	0.66	2601	>200	108	2420	24	105	1472	15
	60	720	519841	9.55	0.98	5625	>200	>200	3600	36	124	2360	24
	80	960	923521	12.73	1.30	9801	>200	>200	4940	49	192	3440	34
	100	1200	1442401	15.92	1.62	15129	>200	>200	6084	61	>200	4324	43
144	20	240	58081	3.18	0.29	529	104	18	1440	10	19	1012	7
	40	480	231361	6.37	0.56	1849	>200	67	3024	21	63	1928	13
	60	720	519841	9.55	0.82	3969	>200	>200	4368	30	78	2928	20
	80	960	923521	12.73	1.09	6889	>200	>200	5952	41	>200	3800	26
	100	1200	1442401	15.92	1.35	10609	>200	>200	7392	51	>200	5140	36

Table 4.2: Homogeneous media test case, using $\tau = 0.7$ and $\tau = 0.5$ for Δ-GenEO and Δ_k-GenEO respectively.

methods using the Δ-GenEO and Δ_k-GenEO coarse space, the dimension of the coarse space (CS) and averaged number of contributions to the coarse space per subdomain (CS_{loc}) for the Δ-GenEO and Δ_k-GenEO methods.

Heterogenous Problem

We now consider the same experiments as in the homogeneous case, but with a heterogeneous diffusion coefficient $A(\mathbf{x})$. This aims to study the robustness of the method with respect to medium heterogeneity. For this purpose, we consider a layered medium within the same unit square domain. More precisely, we use an alternating layer configuration in which the hetero-

								One Level	Δ-GenEO			Δ_k-GenEO		
N	k	h^{-1}	$a_{\max}$	n	L	H	n_{loc}	It	It	CS	CS_{loc}	It	CS	CS_{loc}
16	20	240	10	58081	3.18	0.82	3969	134	20	400	25	21	258	16
	60	720	10	519841	9.55	2.41	33489	>200	121	1234	77	91	842	53
	100	1200	10	1442401	15.92	4.01	91809	>200	>200	2062	129	146	1608	100
16	20	240	100	58081	3.18	0.82	3969	121	19	400	25	21	262	16
	60	720	100	519841	9.55	2.41	33489	>200	76	1234	77	61	826	52
	100	1200	100	1442401	15.92	4.01	91809	>200	196	2064	129	134	1566	98
16	20	240	1000	58081	3.18	0.82	3969	90	15	400	25	18	262	16
	60	720	1000	519841	9.55	2.41	33489	>200	66	1234	77	58	820	51
	100	1200	1000	1442401	15.92	4.01	91809	>200	179	2064	129	111	1568	98
36	20	240	10	58081	3.18	0.56	1849	174	19	672	19	20	420	12
	60	720	10	519841	9.55	1.62	15129	>200	151	2042	57	125	1328	37
	100	1200	10	1442401	15.92	2.68	41209	>200	>200	3426	95	195	2410	67
36	20	240	100	58081	3.18	0.56	1849	147	19	666	18	21	422	12
	60	720	100	519841	9.55	1.62	15129	>200	108	2050	57	73	1320	37
	100	1200	100	1442401	15.92	2.68	41209	>200	>200	3426	95	194	2356	65
36	20	240	1000	58081	3.18	0.56	1849	107	16	666	18	18	422	12
	60	720	1000	519841	9.55	1.62	15129	>200	85	2050	57	67	1320	37
	100	1200	1000	1442401	15.92	2.68	41209	>200	>200	3426	95	176	2356	65
64	20	240	10	58081	3.18	0.42	1089	200	18	900	14	20	584	9
	60	720	10	519841	9.55	1.22	8649	>200	177	2862	45	129	1804	28
	100	1200	10	1442401	15.92	2.02	23409	>200	>200	4778	75	>200	3200	50
64	20	240	100	58081	3.18	0.42	1089	168	19	904	14	21	586	9
	60	720	100	519841	9.55	1.22	8649	>200	130	2850	45	84	1794	28
	100	1200	100	1442401	15.92	2.02	23409	>200	>200	4768	74	>200	3168	50
64	20	240	1000	58081	3.18	0.42	1089	127	16	904	14	18	586	9
	60	720	1000	519841	9.55	1.22	8649	>200	119	2850	45	80	1794	28
	100	1200	1000	1442401	15.92	2.02	23409	>200	>200	4768	74	>200	3168	50
100	20	240	10	58081	3.18	0.34	729	179	18	1176	12	19	752	8
	60	720	10	519841	9.55	0.98	5625	>200	162	3640	36	112	2388	24
	100	1200	10	1442401	15.92	1.62	15129	>200	>200	6134	61	>200	4070	41
100	20	240	100	58081	3.18	0.34	729	109	19	1168	12	19	792	8
	60	720	100	519841	9.55	0.98	5625	>200	94	3664	37	64	2390	24
	100	1200	100	1442401	15.92	1.62	15129	>200	>200	6174	62	171	4160	42
100	20	240	1000	58081	3.18	0.34	729	68	16	1168	12	17	792	8
	60	720	1000	519841	9.55	0.98	5625	>200	68	3664	37	56	2382	24
	100	1200	1000	1442401	15.92	1.62	15129	>200	170	6174	62	113	4192	42
144	20	240	10	58081	3.18	0.29	529	>200	17	1436	10	19	938	7
	60	720	10	519841	9.55	0.82	3969	>200	174	4424	31	126	2766	19
	100	1200	10	1442401	15.92	1.35	10609	>200	>200	7422	52	>200	4838	34
144	20	240	100	58081	3.18	0.29	529	198	19	1412	10	20	934	6
	60	720	100	519841	9.55	0.82	3969	>200	135	4428	31	89	2772	19
	100	1200	100	1442401	15.92	1.35	10609	>200	>200	7466	52	>200	4834	34
144	20	240	1000	58081	3.18	0.29	529	130	16	1412	10	18	934	6
	60	720	1000	519841	9.55	0.82	3969	>200	130	4428	31	90	2772	19
	100	1200	1000	1442401	15.92	1.35	10609	>200	>200	7466	52	>200	4810	33

Table 4.3: Results showing the dimension of the fine mesh (n), the diameter of the domain measured in wavelengths (L), the diameter of the sub-domains measured in wavelengths (H), the number of local degrees of freedom (n_{loc}), the GMRES iteration count (It) for the one-level and two level methods using the Δ-GenEO and Δ_k-GenEO coarse space, the dimension of the coarse space (CS) and averaged number of contributions to the coarse space per subdomain (CS_{loc}) for the Δ-GenEO and Δ_k-GenEO methods. These results are for the heterogeneous media test case, using $\tau = 0.7$ and $\tau = 0.5$ for Δ-GenEO and Δ_k-GenEO respectively.

geneity is introduced through the material parameter $a(\mathbf{x})$ where

$$a(\mathbf{x}) = \begin{cases} a_{\max} & \text{if } y \in [0, 0.1) \cup [0.2, 0.3) \cup [0.4, 0.5) \cup [0.6, 0.7) \cup [0.8, 0.9), \\ a_{\min} & \text{otherwise,} \end{cases}$$

with $a_{\max} > 1$ and $a_{\min} = 1$; see Figure 4.6. The heterogeneous coefficient is given by

$$A(\mathbf{x}) = a(\mathbf{x})I.$$

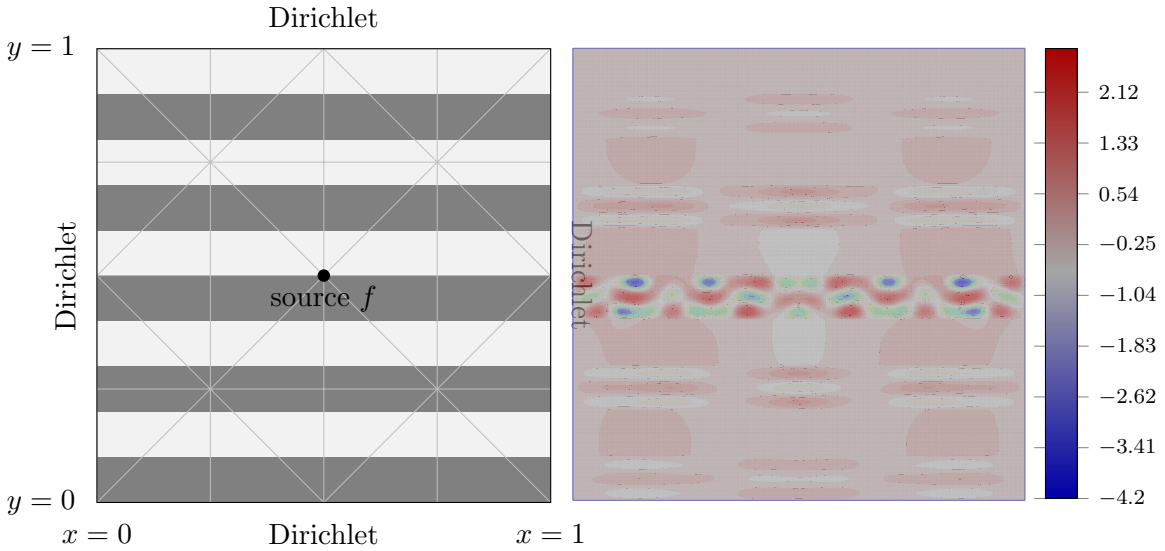


Figure 4.6: The heterogeneous function $a(\mathbf{x})$ within the alternating layers. The shading gives the value of $a(\mathbf{x})$ with the darkest shade being $a(\mathbf{x}) = a_{\max}$, where $a_{\max} > 1$ is a parameter, and the white taking the value $a_{\min} = 1$ (left). The solution for the heterogeneous problem with $k = 100$ and $a_{\max} = 50$ (right).

4.5.2 Results

In Tables 4.2 and 4.3, we display the results comparing the one-level AS and the two-level AS with the Δ-GenEO from [13] and the newly proposed Δ_k-GenEO. The indefiniteness of the problem is controlled by changing k with different subdomain decompositions N. The entries are the number of global degrees of freedom, n, the diameter of the domain, L, measured in the smallest wavelength $\lambda_{\min}$, the diameter of the subdomains, H, measured in the smallest wavelength, the number of local degrees of freedom, n_{loc}, the iterations required to reach the GMRES tolerance, It, for the one-level, Δ-GenEO and Δ_k-GenEO method, the total size of the Δ-GenEO and Δ_k-GenEO coarse space, CS, and the averaged number of contributions to the coarse space per subdomain, CS_{loc}. For the Δ-GenEO and Δ_k-GenEO method τ has been selected for each to give the nearest iteration counts across the two methods. In Table 4.3 we also show the effect from changing the $a_{\max}$ being used.

Further to this, using the smallest and largest values of k tested, we show the number of GMRES iterations required to reach the desired tolerance as a function of the total coarse space size and the threshold value selection in Figures 4.2 and 4.4. The number of GMRES iterations and coarse space size as a function of the subdomain diameter in wavelengths is shown in Figures 4.3 and 4.5, with the heterogeneous results in Figures 4.4 and 4.5 having $a_{\max} = 10$.

4.5.3 Analyses and Comparison

We now compare the results for the homogeneous and heterogeneous test cases across the different methods. In this study, we deliberately focus on iteration counts and coarse-space dimensions

rather than reporting CPU timings. While wall-clock times can be informative, they are highly sensitive to implementation details, solver tolerances, hardware architecture, and the degree of optimisation of the software stack, and therefore do not provide a robust basis for comparison across platforms.

For two-level domain decomposition methods, iteration counts and coarse-space size are widely accepted as reliable proxies for performance. In particular, reduced iteration counts directly decrease the number of local and coarse solves, while smaller coarse spaces lead to cheaper coarse-level factorisations and solves, as well as reduced communication overhead. Consequently, the consistently lower iteration counts and more compact coarse spaces observed for the Δ_k-GenEO method translate into a lower time-to-solution in practical, optimised implementations, even though explicit timings are not reported here.

From both the Tables 4.2 and 4.3, and Figures 4.2 and 4.4 we can identify several points of interest:

Two-level is essential. Δ_k-GenEO dominates Across all tested N and k, the one-level method quickly becomes non-viable beyond small k or a large N. As demonstrated when $k = 40$, the iteration count exceeds the 200 iteration limit for $N \geq 36$. When $k \geq 60$, the one-level method fails for the majority of N being tested. In contrast, both two-level variants are robust at low to moderate k, with Δ_k-GenEO consistently achieving the same or lower iteration count than the Δ-GenEO for a smaller coarse space (CS).

Frequency sensitivity and thresholding For $k \leq 40$, both two-level methods can yield nearly N independent iteration counts, provided that an adequate sized coarse space is used. The Δ_k-GenEO method is shown to require $\approx$ 25–35% fewer basis vectors in the coarse space when compared to the Δ-GenEO method. As k increases, acceptable performance requires increasing the eigenvalue threshold τ, subsequently increasing the CS. Using a $\tau \in [0.4, 0.7]$ is sufficient for $k \lesssim 40$, while $\tau \geq 0.6$–0.7 is required for $k \in \{80, 100\}$ to bring iterations down from the iteration limit to 40–80 iterations when using the Δ_k-GenEO coarse space. However, when using the Δ-GenEO coarse space, increasing τ fails to bring the iteration count below the upper limit for the high k values. Beyond these ranges, further increases in τ let the CS grow substantially but can start to yield diminishing iteration gains. This effect is more prominent when using the Δ-GenEO than for Δ_k-GenEO. Beyond both methods' inherent limitations for higher frequencies, numerical experiments provide insight into how the spectral threshold τ may be chosen in practice. Across all tested configurations, robust convergence of the Δ_k-GenEO method is obtained for values of τ less than 1 well below the worst-case bounds predicted by the theory, with only a modest increase in coarse-space dimension as τ is increased.

Scaling with N, H, and k With fixed spectral threshold τ, the global coarse-space dimension grows approximately linearly with the number of subdomains N and increases with the wavenumber k. At the subdomain level, the number of local coarse modes, denoted by CS_{loc}, increases with the subdomain diameter measured in wavelengths, H, since fewer and wider subdomains support a larger number of low-energy local modes. While the theory suggests a fairly conservative restriction on the size subdomain size, we can observe from the results that we are able to achieve convergence with substantially larger subdomains then suggested.

Consequently, for fixed k and τ, increasing H reduces the number of subdomains and therefore decreases the total coarse-space dimension CS, while simultaneously increasing the average number of local contributions CS_{loc}. For low to moderate frequencies, the GMRES iteration counts are fairly insensitive to changes in H for both the Δ-GenEO and Δ_k-GenEO methods, indicating that the coarse spaces remain sufficiently rich to capture the relevant global modes. The total coarse-space size also differs only mildly between the two methods in this regime. At

higher frequencies, however, the Δ_k-GenEO method is more robust, reflecting its ability to retain additional modes required to stabilise the iteration in more indefinite regimes. As a general rule, achieving scalability with respect to the number of domains and frequency happens at the expense of a larger coarse space.

Effect of heterogeneity (contrast $a_{\max}$) Introducing heterogeneity produces mild changes in CS size and, in many cases, slightly improves iterations for the two-level methods. For example, at $k = 60$ and fixed N, increasing $a_{\max}$ often decreases the iteration counts for both coarse spaces, with Δ_k-GenEO maintaining a clear advantage in both iterations and CS size. Overall, the behaviours observed in the homogeneous case persist under heterogeneity, with only weak dependence on $a_{\max}$. The one-level method is adversely effected by the heterogeneity, obtaining higher iteration counts even with low k and N, where the one-level method performed more favourably in the homogeneous case.

High-frequency regime At $k \geq 60$, Δ_k-GenEO remains viable across the tested N provided τ is large enough (typically ≥ 0.6–0.7). In comparison, Δ-GenEO often reaches the iteration limit even as τ increases.

One-level baseline The one-level method should be avoided except for very small k and N, and with no heterogeneity. Its iteration counts grow with N even at $k = 20$ and generally hit the iteration limit for $k \geq 40$.

From these results, we can see that for both the homogeneous and heterogeneous tests, the Δ_k-GenEO is significantly more robust and scalable than the one-level method. It can remain effective at high k provided the eigenvalue threshold τ is increased sufficiently. Whilst the Δ-GenEO is effective for moderate k, it requires a larger coarse space in order to see the same reduction in iteration counts, and still fails to converge within the iteration limit when k is high. The one-level method on the other hand, is not competitive except at the very lowest frequencies. It has also been noticed that with an increase in the heterogeneity, it was possible to obtain fewer iteration counts. The lower iteration counts for some of the heterogeneous problems may stem from the fact that in some parts of the domain the heterogeneity in fact reduces the strength of the indefiniteness. For example, if we let $\Omega_{\max}$ be a section of the domain where $a(\mathbf{x})$ is at its maximum value, $a_{\max}$, then if we restrict problem (4.1.1) to $\Omega_{\max}$, we see that (4.1.1) can be reduced to

$$\Delta u - a_{\max}^{-1} k^2 u = a_{\max}^{-1} f.$$

Thus the restriction acts to reduce the effective indefiniteness in this region. If using the Δ_k-GenEO, then the following operational guidance should be followed,

- For low to moderate frequency ($k \lesssim 40$): choose $\tau \approx 0.4-0.7$ to get $\sim 20-50$ iterations with moderate CS.
- For high frequency (e.g. $k \geq 80$): set $\tau \approx 0.6-0.7$ to keep iterations $\lesssim 50-80$ across N.
- Expect $CS \propto N$ and CS_{loc} to grow with H and k.

4.5.4 Comparison with H_k-GenEO

To further clarify the positioning of the proposed Δ_k-GenEO method, we include a brief numerical comparison with the H_k-GenEO coarse space briefly mentioned in the introduction, where the local spectral problem involves the full (indefinite) operator. That is, for each $i \in \{1, \ldots, N\}$ we solve the local eigenvalue problem consisting in finding the eigenpairs (p, λ) such that

$$b_{\Omega_i}(p, v) = \lambda \big(\Xi_i p, \Xi_i v\big)_{1,k,\Omega_i} \qquad \text{for all } v \in \widetilde{V}_i$$

The discrete versions of eigenfunctions p are added then to the coarse space in the same fashion as for the Δ_k-GenEO method.

The problem being studied for the comparison between the Δ_k-GenEO and H_k-GenEO coarse space is that of the homogeneous problem as outlined in Section 4.5.1. Table 4.4 displays the results comparing the one-level AS and the two-level AS with the H_k-GenEO and the Δ_k-GenEO. The indefiniteness of the problem is controlled by k, and the number of subdomains is given by N. All other entries are as previously defined. Figure 4.7 shows the number of iterations to achieve the required GMRES tolerance as a function of the total coarse space size, and the selected threshold. Figure 4.8 shows the number of GMRES iterations and coarse space size as a function of the subdomain diameter in wavelengths.

We notice that for low to moderate frequencies, both approaches exhibit robust two-level convergence, with iteration counts that depend only weakly on the number of subdomains once a moderate spectral threshold is used. In this regime, Δ_k-GenEO remains comparable to H_k-GenEO, although it typically requires a slightly larger coarse space to achieve similar iteration counts. As the frequency increases, however, the behaviour of the two methods diverges more clearly: while H_k-GenEO maintains fast convergence for moderate thresholds, Δ_k-GenEO becomes increasingly sensitive to the wavenumber and the subdomain partition, requiring larger coarse spaces and higher iteration counts to converge.

This highlights an intrinsic limitation of purely SPD-based coarse spaces in the high-frequency, strongly indefinite regime. Nevertheless, an important advantage of Δ_k-GenEO lies in its conceptual and theoretical simplicity: the associated local spectral problems are symmetric positive definite and admit a relatively complete and transparent analysis. In contrast, the fully indefinite spectral problems underlying H_k-GenEO are considerably more involved, and a comprehensive theoretical understanding of their robustness and scalability remains, to date, far from established. In Table 4.4 we show the dimension of the fine mesh (n), the diameter of the domain measured in wavelengths (L), the diameter of the sub-domains measured in wavelengths (H), the number of local degrees of freedom (n_{loc}), the GMRES iteration count (It.) for the one-level and two level methods using the H_k-GenEO and Δ_k-GenEO coarse space, the dimension of the coarse space (CS) and averaged number of contributions to the coarse space per subdomain (CS_{loc}) for the H_k-GenEO and Δ_k-GenEO.

4.6 Conclusion and Discussion

In this work, we have introduced the Δ_k-GenEO coarse space. We then derived and rigorously analysed the required conditions to achieve robust convergence of GMRES when applied to the Helmholtz problem preconditioned by the Δ_k-GenEO coarse space. Our main theoretical results deliver a refinement of the bounds on the subdomain diameter, H, and the eigenvalue tolerance, τ, when compared to the Δ-GenEO method, reducing the wavenumber dependence from $H \lesssim k^{-2}$ and $\tau \gtrsim k^{-8}$ to $H \lesssim k^{-1}$ and $\tau \gtrsim k^{-2}$, respectively. These improvements stem from strengthened estimates on the projection and stability operators used in the spectral construction of the coarse space. The theoretical improvements are also reflected in the numerical results. It has been demonstrated that the Δ_k-GenEO method outperforms the Δ-GenEO method across all tested scenarios. Not only has the Δ_k-GenEO shown to have improved robustness with respect to more indefinite problems, it is also able to achieve convergence in fewer iterations and using a smaller coarse space over that of the Δ-GenEO method. While the improved conditions of the Δ_k-GenEO are still conservative, they represent a meaningful step towards more practical and scalable preconditioners for indefinite wave problems.

Beyond the specific bounds obtained, this work provides insight into why GenEO-type coarse spaces based on symmetric positive definite operators can remain effective for indefinite Helmholtz problems. By introducing a k-dependent modification of the GenEO eigenvalue problem, we clarify how much of the observed robustness can be explained within an SPD-based

							One Level	H_k-GenEO			Δ_k-GenEO		
N	k	h^{-1}	n	L	H	n_{loc}	It	It	CS	CS_{loc}	It	CS	CS_{loc}
16	20	240	58081	3.18	0.82	3969	41	16	240	18	21	340	21
	40	480	231361	6.37	1.62	15129	107	20	628	49	29	784	49
	60	720	519841	9.55	2.41	33489	153	17	1160	86	31	1312	82
	80	960	923521	12.73	3.21	59049	>200	27	1876	132	75	1936	121
	100	1200	1442401	15.92	4.01	91809	>200	36	2660	190	68	2672	167
36	20	240	58081	3.18	0.56	1849	79	19	344	12	22	516	14
	40	480	231361	6.37	1.09	6889	>200	31	844	28	56	1168	32
	60	720	519841	9.55	1.62	15129	>200	18	1528	49	34	1932	54
	80	960	923521	12.73	2.15	26569	>200	23	2312	72	68	2740	76
	100	1200	1442401	15.92	2.68	41209	>200	51	3260	99	90	3684	102
64	20	240	58081	3.18	0.42	1089	92	19	448	8	23	736	12
	40	480	231361	6.37	0.82	3969	175	24	1056	18	56	1604	25
	60	720	519841	9.55	1.22	8649	>200	18	1924	34	72	2460	38
	80	960	923521	12.73	1.62	15129	>200	23	2820	49	76	3584	56
	100	1200	1442401	15.92	2.02	23409	>200	21	3840	65	88	4616	72
100	20	240	58081	3.18	0.34	729	96	18	684	8	21	980	10
	40	480	231361	6.37	0.66	2601	>200	24	1440	16	87	1936	19
	60	720	519841	9.55	0.98	5625	>200	18	2200	24	104	3020	30
	80	960	923521	12.73	1.30	9801	>200	24	3224	34	144	4324	43
	100	1200	1442401	15.92	1.62	15129	>200	24	4504	49	80	5740	57
144	20	240	58081	3.18	0.29	529	104	21	672	5	19	1056	7
	40	480	231361	6.37	0.56	1849	>200	30	1544	12	64	2256	16
	60	720	519841	9.55	0.82	3969	>200	23	2448	18	77	3796	26
	80	960	923521	12.73	1.09	6889	>200	45	3700	28	162	5140	36
	100	1200	1442401	15.92	1.35	10609	>200	25	5232	39	197	6576	46

Table 4.4: Threshold values for the coarse space are $\tau = 0.4$ and $\tau = 0.6$ for H_k-GenEO and Δ_k-GenEO respectively.

framework, and where its fundamental limitations lie. In particular, the analysis helps to narrow the long-standing gap between conservative theoretical guarantees and the much milder coarse-space growth and iteration counts observed in practice.

Several directions for future research naturally follow from these results. These include the development and rigorous analysis of coarse spaces built directly from indefinite operators, multilevel extensions of the Δ_k-GenEO framework, and optimised implementations suitable for large-scale parallel simulations. The present work provides a theoretical and numerical baseline against which such developments can be assessed.

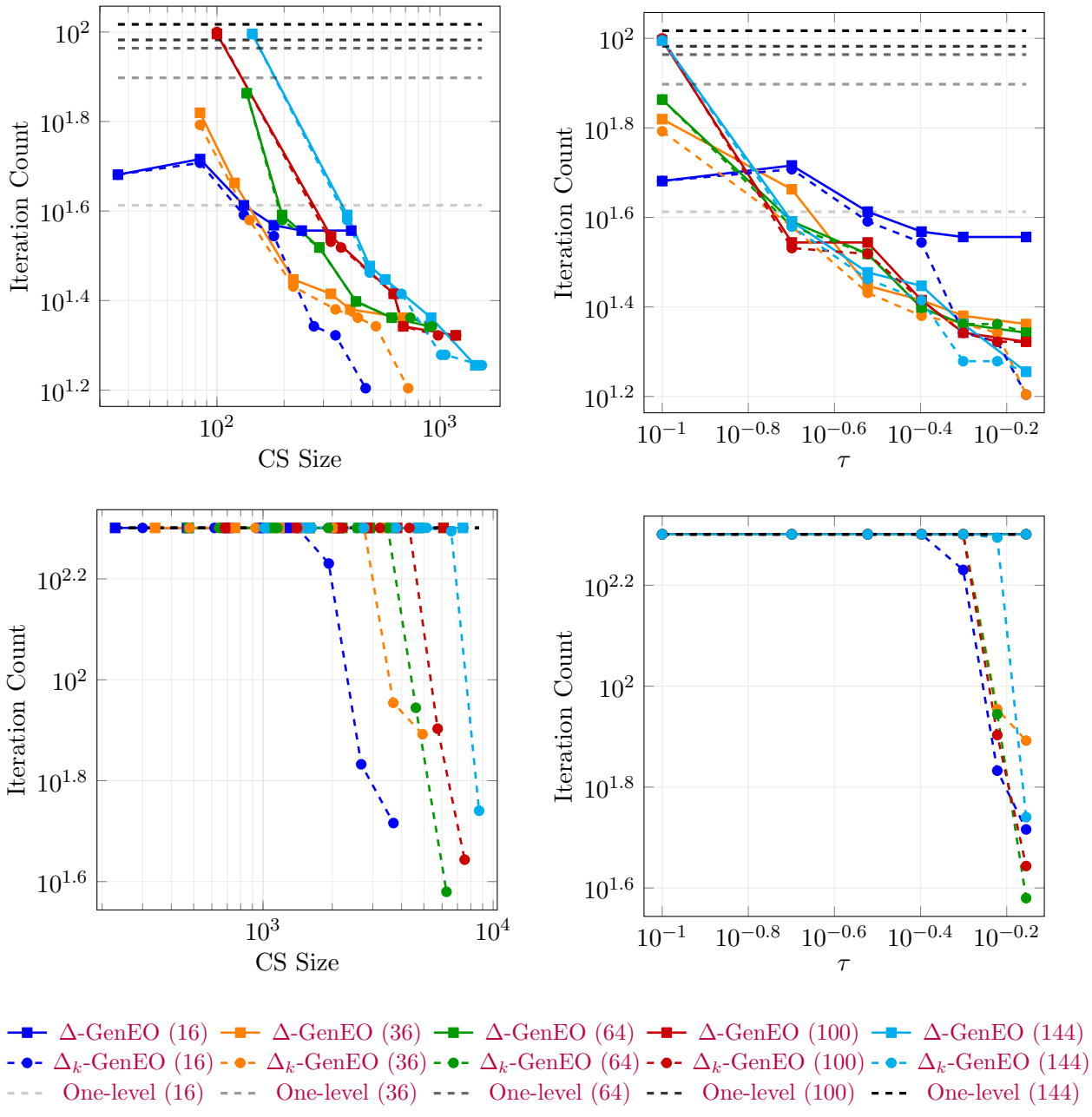


Figure 4.2: Influence of the coarse space size (left) and threshold choice (right) on the iteration count for the homogeneous media test case with $k = 20$ (top) and $k = 100$ (bottom). The number in brackets indicates the number of subdomains.

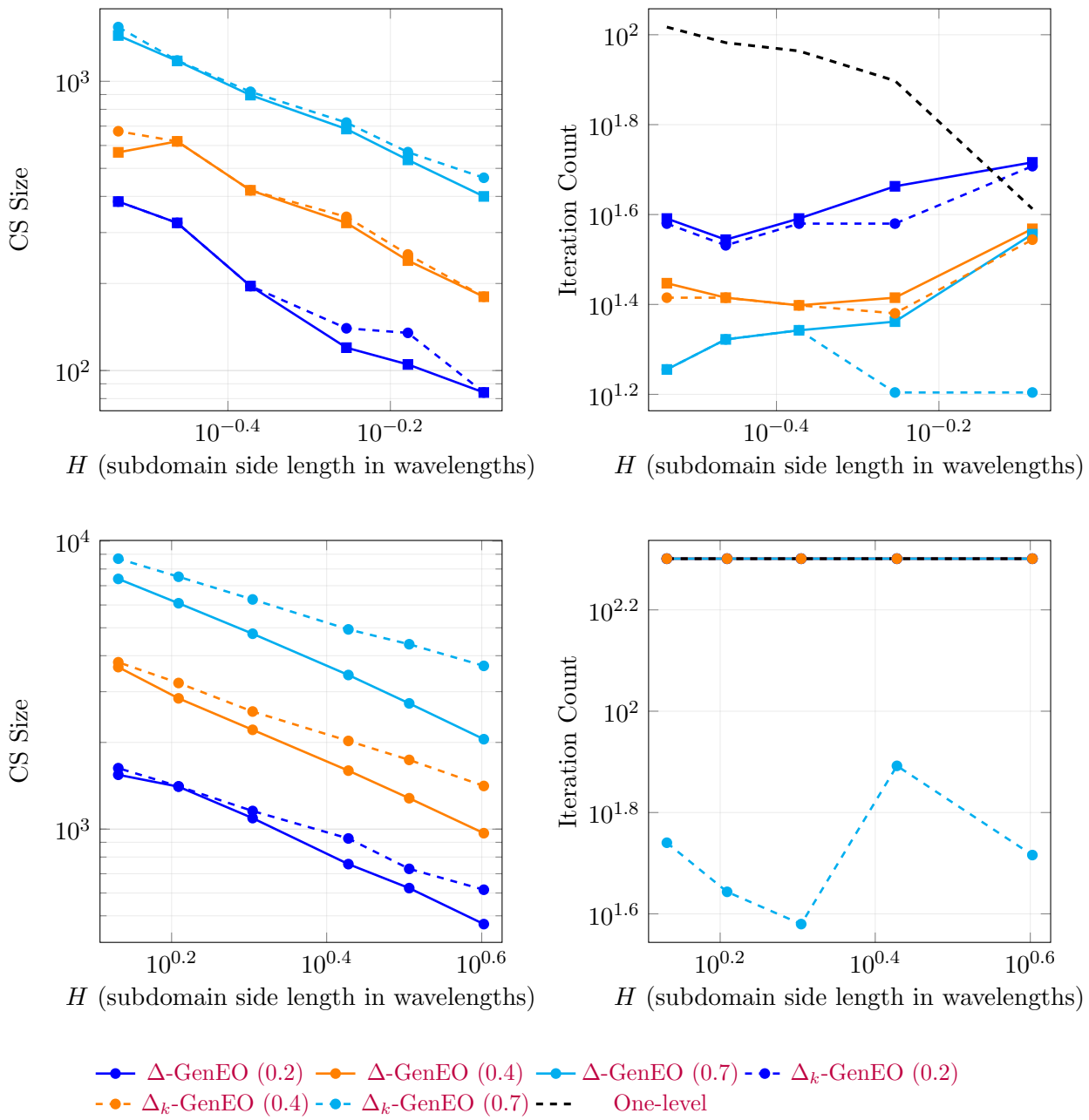


Figure 4.3: Influence of the subdomain diameter on the coarse space size (left) and iteration count (right) for the homogeneous media test case $k = 20$ (top) and $k = 100$ (bottom). The number in brackets indicates τ used.

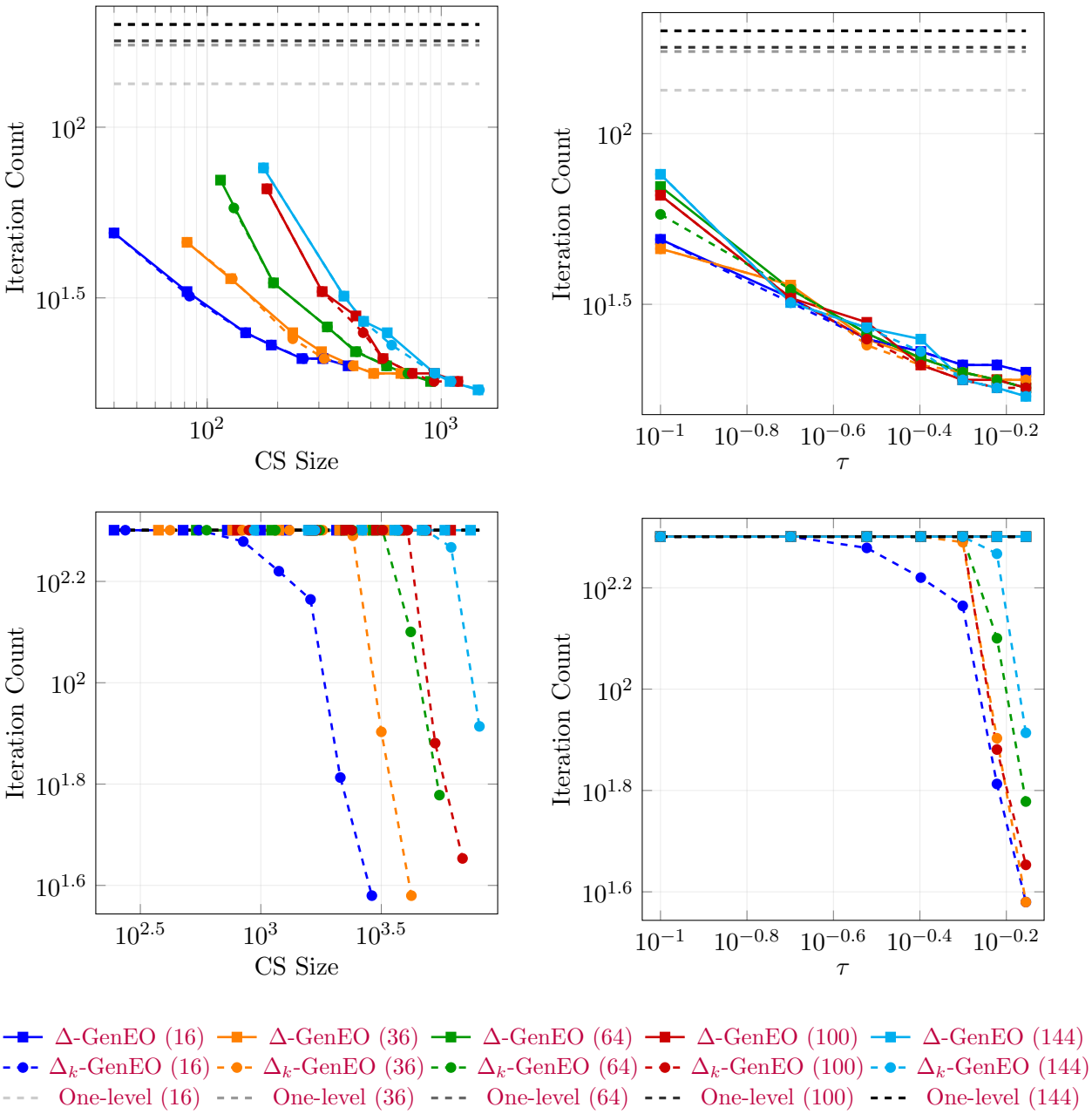


Figure 4.4: Influence of the coarse space size (left) and threshold choice (right) on the iteration count for the heterogeneous media test case with $k = 20$ (top) and $k = 100$ (bottom), all with $a_{\max}(\mathbf{x}) = 10$. The number in brackets indicates the number of subdomains.

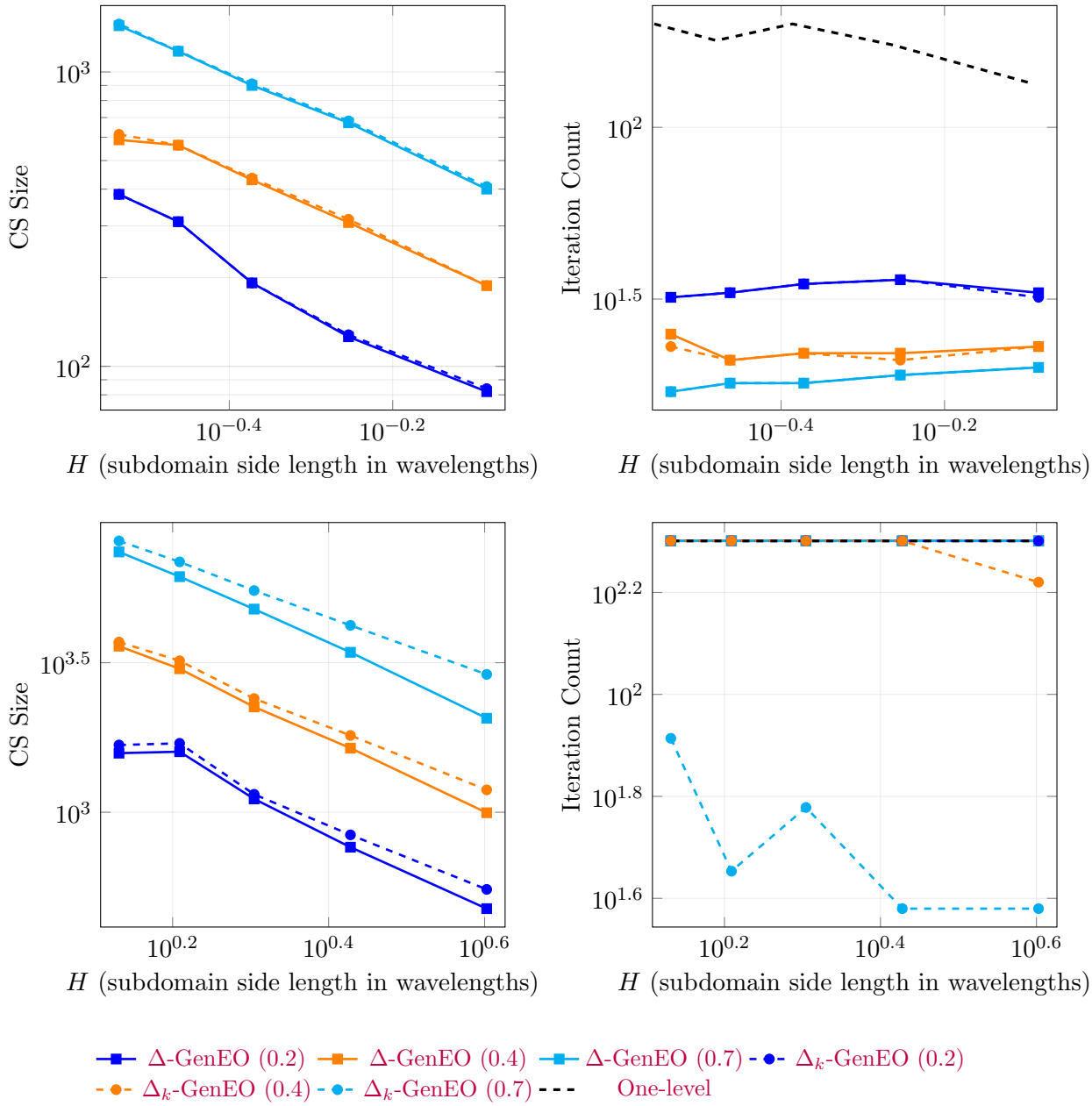


Figure 4.5: Influence of the subdomain diameter on the coarse space size (left) and iteration count (right) for the heterogeneous media test case $k = 20$ (top) and $k = 100$ (bottom) , all with $a_{\max}(\mathbf{x}) = 10$. The number in brackets indicates τ used.

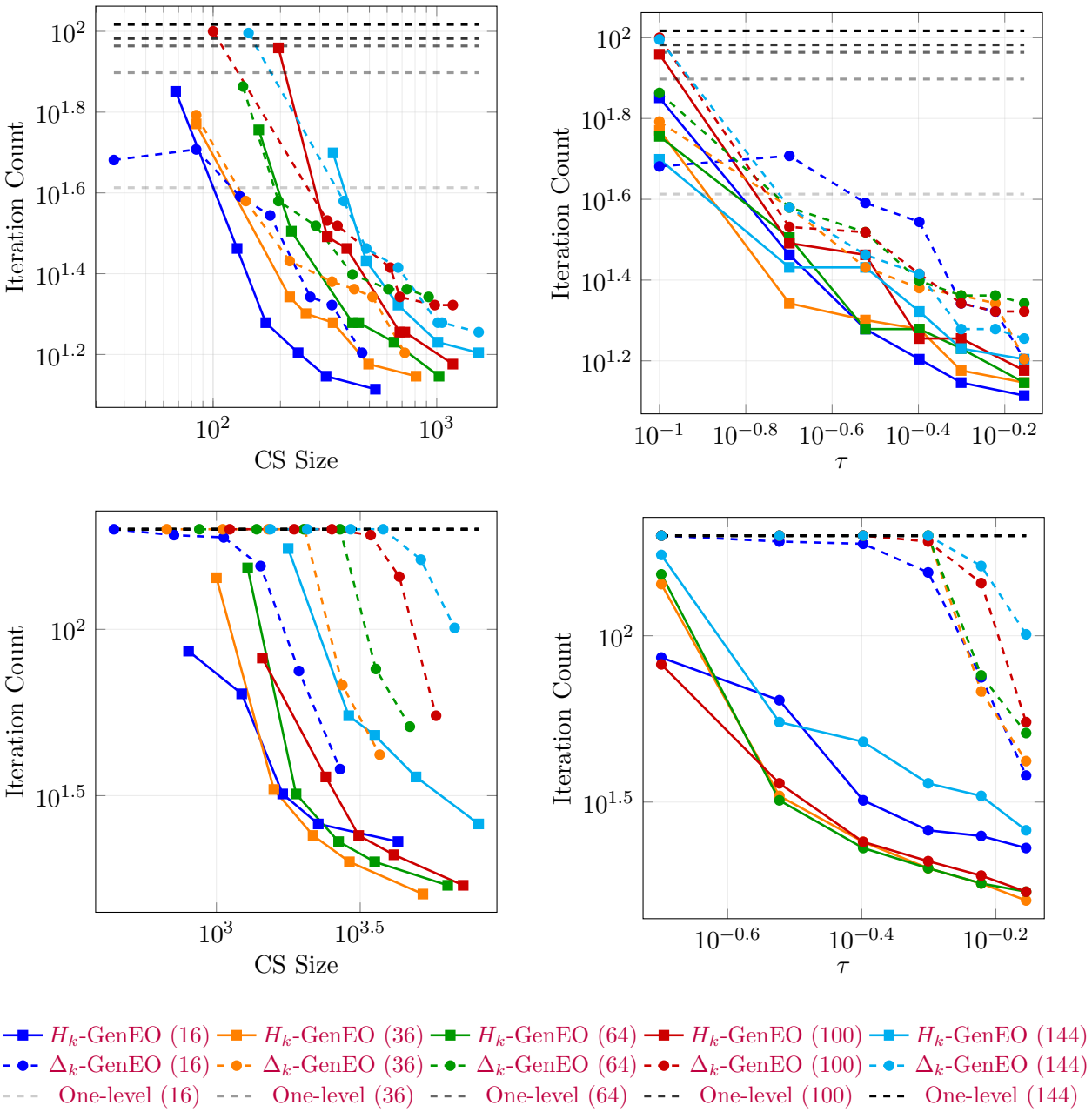


Figure 4.7: Influence of the coarse space size (left) and threshold choice (right) on the iteration count for the homogeneous media test case with $k = 20$ (top) and $k = 100$ (bottom). The number in brackets indicates the number of subdomains.

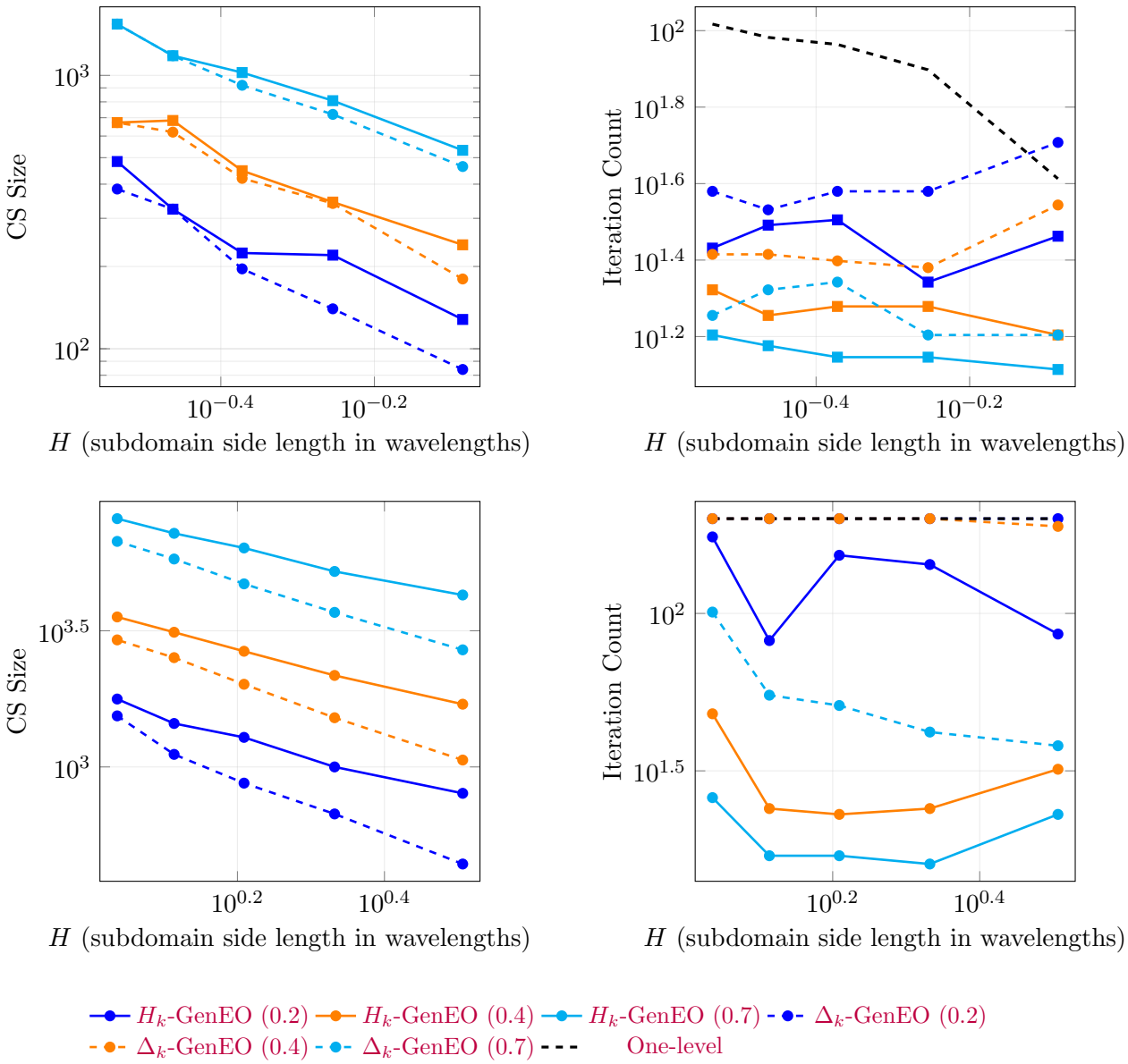


Figure 4.8: Influence of the subdomain diameter on the iteration count (left) and coarse space size (right) for the homogeneous media test case $k = 20$ (top) and $k = 100$ (bottom). The number in brackets indicates τ used.

Chapter 5

Spectral Coarse Spaces Based on Indefinite Operators: The H_k-GenEO Method

This chapter has been done in collaboration with Victorita Dolean, Matthias Langer, Ivan G. Graham and Théophile Chaumont-Frelet. It will form part of a paper in preparation titled "Spectral coarse spaces based on indefinite operators: the H_k-GenEO method".

5.1 Introduction

This chapter develops the theory and implementation of two-level overlapping domain decomposition preconditioners for finite element systems arising from indefinite Helmholtz-type boundary value problems with heterogeneous coefficients:

$$-\operatorname{div}(A\nabla u) - k^2 n\, u = f \qquad \text{in } \Omega, \tag{5.1.1a}$$

$$u = 0 \qquad \text{on } \partial\Omega, \tag{5.1.1b}$$

where $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ is polygonal/polyhedral and Lipschitz, $A \in (L^\infty(\Omega))^{d\times d}$ is uniformly elliptic, and $n \in L^\infty(\Omega)$ is almost everywhere positive. We assume (5.1.1) admits a unique weak solution for all $f \in L^2(\Omega)$. As k increases the operator becomes more indefinite, and robust iterative solution requires a preconditioner whose performance degrades minimally with k.

Aim: coarse spaces built from the full indefinite operator. Earlier GenEO extensions to indefinite/non-self-adjoint problems, notably Δ-GenEO [13] and its k-weighted refinement Δ_k-GenEO (Chapter 4, [25]), still define local eigenvalue problems using SPD surrogate operators. While effective over a substantial parameter range, their performance becomes suboptimal at higher frequencies, reflecting limitations inherent in SPD-based coarse constructions.

The goal of this chapter is to develop a GenEO-type coarse space that is defined *directly* in terms of the full indefinite Helmholtz operator. We consider the H_k-GenEO construction, inspired by the H-GenEO approach of [14], in which the local generalised eigenvalue problems involve localisations of the original indefinite bilinear form (together with a k-weighted norm). This choice yields local spectra that can contain both positive and negative components, mirroring the global operator but invalidating classical SPD spectral arguments. Consequently, the analysis must proceed by different techniques.

Two decompositions: fine and coarse scales. A further structural feature of the proposed method is the use of two distinct decompositions: a fine-scale decomposition with characteristic diameter H_ℓ driving the local Schwarz solves, and a coarse-scale decomposition with diameter H_c defining the assembly of the global coarse correction. Previous GenEO analyses typically

assume a single scale $H_\ell = H_c = H$. Decoupling these scales provides additional flexibility and leads to sharper bounds, in particular by reducing the extent to which the coarse-space dimension is tied to the coarse mesh diameter. After finite element discretisation, the resulting system is symmetric but indefinite, so we employ GMRES. The convergence theory relies on Elman's field-of-values framework, as discussed in Section 2.4.2.

Main result. Theorem 5.3.1 establishes *explicit k-dependent conditions* on H_ℓ and the eigenvalue tolerance τ that guarantee convergence of GMRES. The behaviour of the method is primarily governed by:

- the fine subdomain size H_ℓ, which determines the ability of local solves to capture oscillatory behaviour,
- the eigenvalue tolerance
$$\tau = \min_{i=1}^{N} \lambda^i_{m_i+1},$$
where $\lambda^i_{m_i+1}$ is the first unused eigenvalue in subdomain Ω_i, and which controls the dimension of the coarse space.

While these parameters are often chosen empirically, we derive rigorous k-dependent bounds, improving substantially upon previous results. Let $C_{\text{stab}} > 0$ denote the stability constant from (2.2.9), which becomes large as k^2 approaches a Dirichlet eigenvalue of $-\operatorname{div}(A\nabla u)$.

Table 5.1 provides a side-by-side comparison of the sufficient conditions for GMRES robustness across three approaches:

Method	**Subdomain size**	**Eigenvalue tolerance τ**
Δ-GenEO [13]	$H \lesssim k^{-2}$	$(1 + C_{\text{stab}})^2\, k^8 \lesssim \tau$
Δ_k-GenEO (Chapter 4)	$H \lesssim k^{-1}$	$(1 + C_{\text{stab}})^2\, k^2 \lesssim \tau$
H_k-GenEO (this chapter)	$H_\ell \lesssim k^{-1}$	$(1 + C_{\text{stab}})^2\, k^2 \lesssim \tau$

Table 5.1: Comparison of robustness conditions for different GenEO-type coarse spaces.

Although these results indicate that H_k-GenEO and Δ_k-GenEO satisfy the same theoretical conditions, the theory developed for H_k-GenEO offers new insight into the treatment of the purely indefinite Helmholtz problem. Moreover, the numerical results reveal that the two methods exhibit markedly different practical behavior. Moreover, the separation of fine and coarse decompositions decouples the coarse space dimension from the coarse mesh diameter H_c, eliminating a restrictive assumption of classical GenEO methods. Finally, we remark that the theoretical conditions in the earlier bounds are likely overly pessimistic in practice. As shown in the numerical experiments of [13], robust convergence is often achieved with coarse spaces of moderate dimension.

5.2 Useful Background

In this section we introduce the variational formulation of (5.1.1) and its finite element discretisation. We then present an abstract finite-dimensional framework for variational generalised eigenvalue problems with indefinite forms. Finally, we define the additive Schwarz domain decomposition preconditioner whose novel feature is the H_k-GenEO coarse space, constructed from modes of local *indefinite* generalised eigenvalue problems.

5.2.1 Problem Formulation and Finite Element Discretisation

The weak formulation of (5.1.1) is: find $u \in H_0^1(\Omega)$ such that

$$b(u,v) = (f,v) \qquad \text{for all } v \in H_0^1(\Omega), \tag{5.2.1}$$

where $f \in L^2(\Omega)$ and $b(\cdot,\cdot) : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ is defined by

$$b(u,v) = \int_\Omega \left(A\nabla u \cdot \nabla v - k^2 n\, uv \right) \mathrm{d}x.$$

The following mild regularity and scaling assumptions will be used throughout.

Assumption 5.2.1. *The coefficients A and n and the domain Ω in (5.1.1) satisfy:*

(i) *$A : \Omega \to \mathbb{R}^{d\times d}$ and $n : \Omega \to \mathbb{R}$ are measurable, A is symmetric, and there exist $0 < a_{\min} \le a_{\max}$ and $0 < n_{\min} \le n_{\max}$ such that*

$$\begin{aligned} a_{\min}|\boldsymbol{\xi}|^2 \le A(x)\boldsymbol{\xi}\cdot\boldsymbol{\xi} \le a_{\max}|\boldsymbol{\xi}|^2 \qquad & \text{for a.e. } x \in \Omega \text{ and all } \boldsymbol{\xi} \in \mathbb{R}^d, \\ \text{and} \qquad n_{\min} \le n(x) \le n_{\max} \qquad & \text{for a.e. } x \in \Omega. \end{aligned}$$

(ii) *Since (5.1.1) can be equivalently rewritten as*

$$-\nabla\cdot\left(\frac{A}{a_{\min}}\nabla u\right) - \left(\frac{k^2 n_{\max}}{a_{\min}}\right)\left(\frac{n}{n_{\max}}\right)u = \frac{f}{a_{\min}},$$

we will assume, without loss of generality, that $a_{\min} = 1 = n_{\max}$ and that the diameter D_Ω of Ω satisfies $D_\Omega \le 1$.

These assumptions are extensions of Assumptions 2.2.2 for the more general problem definition (5.1.1)

Notation 5.2.2. *For any subdomain $\Omega' \subseteq \Omega$, let $(\cdot,\cdot)_{\Omega'}$ denote the $L^2(\Omega')$ inner product and $\|\cdot\|_{\Omega'}$ its norm. We introduce the bilinear forms*

$$a_{\Omega'}(u,v) := \int_{\Omega'} A\nabla u\cdot\nabla v\,\mathrm{d}x, \qquad b_{\Omega'}(u,v) := a_{\Omega'}(u,v) - k^2(nu,v)_{\Omega'}, \quad u,v \in H^1(\Omega'), \tag{5.2.2}$$

and the seminorm $|u|_{a,\Omega'} := \sqrt{a_{\Omega'}(u,u)}$. We also use the k-weighted (A- and n-dependent) inner product

$$(u,v)_{1,k,\Omega'} := a_{\Omega'}(u,v) + k^2(nu,v)_{\Omega'},$$

with induced norm $\|u\|_{1,k,\Omega'}$. When $\Omega' = \Omega$, subscripts Ω are omitted.

The forms $a(\cdot,\cdot)$ and $(\cdot,\cdot)_{1,k}$ are symmetric and positive definite on $H_0^1(\Omega)$; $b(\cdot,\cdot)$ is symmetric but, in general, indefinite. We assume (2.2.9) for the solvability of (5.2.1). Such a constant C_{stab} exists whenever k^2 is not a Dirichlet eigenvalue of the operator $-\frac{1}{n}\operatorname{div}(A\nabla\cdot)$.

Discretisation Let $\mathcal{T}_h$ be a family of shape-regular simplicial meshes on Ω, parametrised by the maximum element diameter h. We denote by $\sigma_{\mathcal{T}_h}$ the shape-regularity parameter of $\mathcal{T}_h$. Let $V^h \subset H_0^1(\Omega)$ be the H^1-conforming Lagrange finite element space of degree $r \ge 1$ with nodal basis $\{\phi_j\}_{j=1}^n$ (the indices j are the *degrees of freedom*). The Galerkin approximation of (5.2.1) is: find $u_h \in V^h$ such that

$$b(u_h, v_h) = (f, v_h) \qquad \text{for all } v_h \in V^h. \tag{5.2.3}$$

This is equivalent to the linear system

$$\mathbf{B}\mathbf{u} = \mathbf{f}, \tag{5.2.4}$$

where $(\mathbf{B})_{ij} := b(\phi_j, \phi_i)$ and $(\mathbf{f})_i := (f, \phi_i)$. We will also use the matrices $\mathbf{A}$ and $\mathbf{S}$ corresponding to $a(\cdot,\cdot)$ and $n(\cdot,\cdot)$, namely $(\mathbf{A})_{ij} := a(\phi_j, \phi_i)$ and $(\mathbf{S})_{ij} := (n\phi_j, \phi_i)$.

The solvability of (5.2.3) for sufficiently fine meshes and/or large degree r is ensured by the next lemma, proved for $r = 1$ in [82, Theorem 2]. Since we could not find a reference for general $r \geq 1$, we include a proof.

Lemma 5.2.3. *Let Assumptions 5.2.1 and (2.2.9) hold, and set*

$$\beta := k \max_{\substack{\phi \in L^2(\Omega) \\ \|\phi\|=1}} \min_{v_h \in V^h} \|u_\phi - v_h\|_{1,k},$$

where u_ϕ is the unique element of $H^1_0(\Omega)$ such that $b(u_\phi, w) = (n\phi, w)$ for all $w \in H^1_0(\Omega)$. Then there exists a function $\theta \in C^0((0,\infty),(0,\infty))$ depending only on Ω, A, and k, with $\lim_{\tau\to 0} \theta(\tau) = 0$, such that

$$\beta \leq C(\sigma_{\mathcal{T}_h})\theta(h/r). \tag{5.2.5}$$

In addition, if $\beta < \sqrt{2}/2$, then the linear system (5.2.3) admits a unique solution u_h, and

$$\|u - u_h\|_{1,k} \leq \frac{1}{\sqrt{1-2\beta^2}} \min_{v_h \in V^h} \|u - v_h\|_{1,k} \leq \frac{\beta}{\sqrt{1-2\beta^2}} \frac{1}{k} \|f\| \tag{5.2.6}$$

for all $f \in L^2(\Omega)$.

Proof. The definition of β is meaningful by (2.2.9). The uniqueness of u_h and the estimates in (5.2.6) for $\beta < \sqrt{2}/2$ constitute a standard result often referred to as the "Schatz argument" [83, 82]. For completeness, we recall the brief proof.

Let $f \in L^2(\Omega)$ and let $u_h \in V^h$ be any solution of (5.2.3). Further, let $u \in H^1_0(\Omega)$ be the unique solution of (5.2.1), which exists by (2.2.9). Again by (2.2.9), there exists a unique $\xi \in H^1_0(\Omega)$ such that $b(\xi, w) = k(u - u_h, w)$ for all $w \in H^1_0(\Omega)$. By the definition of β, there is $\xi_h \in V^h$ with

$$\|\xi - \xi_h\|_{1,k} \leq \beta \|u - u_h\|.$$

Taking $w = u - u_h$ in the definition of ξ yields

$$k\|u - u_h\|^2 = b(u - u_h, \xi) = b(u - u_h, \xi - \xi_h) \leq \|u - u_h\|_{1,k} \, \|\xi - \xi_h\|_{1,k},$$

where we used that u_h solves (5.2.3). Hence

$$k\|u - u_h\| \leq \beta \|u - u_h\|_{1,k}.$$

Therefore,

$$(1 - 2\beta^2)\|u - u_h\|^2_{1,k} \leq \|u - u_h\|^2_{1,k} - 2k^2\|u - u_h\|^2 = b(u - u_h, u - u_h) = b(u - u_h, u - v_h)$$

for all $v_h \in V^h$. Assuming $2\beta^2 < 1$ we obtain the first estimate in (5.2.6); the second follows from the definition of β.

So far (5.2.6) holds for any solution, but existence is not yet guaranteed. Since u_h is defined by a finite-dimensional linear system, stability implies uniqueness, which in turn implies existence.

We now prove (5.2.5). By the Fredholm alternative, (2.2.9) ensures that the mapping $S : H^{-1}(\Omega) \to H^1_0(\Omega)$ defined by

$$b(S\psi, v) = \langle \psi, v \rangle$$

for all $\psi \in H^{-1}(\Omega)$ and $v \in H^1_0(\Omega)$ is linear and continuous. Since the embedding $L^2(\Omega) \hookrightarrow H^{-1}(\Omega)$ is compact, $K := S(B_{L^2})$, where B_{L^2} is the closed unit ball of $L^2(\Omega)$, is compact in $H^1_0(\Omega)$.

Observe that

$$\beta = k \max_{v\in K} \min_{v_h\in V^h} \|v - v_h\|_{1,k}.$$

Fix $\varepsilon > 0$. By compactness of K, there exist $\{v^j\}_{j=1}^N \subset K$ such that for all $v \in K$ there is $j \in \{1,\dots,N\}$ with $\|v - v^j\|_{1,k} \le \varepsilon/(3k)$. By density of $C_c^\infty(\Omega)$ in $H_0^1(\Omega)$, for each j there exists $\widetilde{v}^j \in C_c^\infty(\Omega)$ with $\|v^j - \widetilde{v}^j\|_{1,k} \le \varepsilon/(3k)$. By [72], there exists $v_h^j \in V^h$ such that

$$k\|\widetilde{v}^j - v_h^j\|_{1,k} \le C(\sigma_{\mathcal{T}_h})\left(1 + \frac{kh}{r}\right)\frac{kh}{r}|\widetilde{v}^j|_{H^2(\Omega)} \le C(\sigma_{\mathcal{T}_h})\left(1 + \frac{kh}{r}\right)\frac{kh}{r}\left(\max_{1\le j\le N}|\widetilde{v}^j|_{H^2(\Omega)}\right)$$

for h/r sufficiently small. A triangle inequality then shows that

$$k \min_{v_h\in V^h} \|v - v_h\|_{1,k} \le \frac{2\varepsilon}{3} + C(\sigma_{\mathcal{T}_h})M(\varepsilon)\left(1 + \frac{kh}{r}\right)\frac{kh}{r},$$

for all $v \in K$, where $M(\varepsilon) := \max_{1\le j\le N}|\widetilde{v}^j|_{H^2(\Omega)}$. Hence $\beta \le \varepsilon$. If kh is small enough so that

$$\left(1 + \frac{kh}{r}\right)\frac{kh}{r} \le C(\sigma_{\mathcal{T}_h})^{-1}M(\varepsilon)^{-1}\frac{\varepsilon}{3},$$

then $\beta \le \varepsilon$. Since $\varepsilon > 0$ was arbitrary, the existence of θ with the desired properties follows. □

Before proceeding, we recall Friedrichs' inequality (e.g., [68, Theorem 13.19]).

Lemma 5.2.4 (Friedrichs' inequality)**.** *Let $\Omega' \subset \mathbb{R}^d$ be an open set that lies between two parallel hyperplanes separated by a distance L. Then, for all $u \in H_0^1(\Omega')$,*

$$\|u\|_{\Omega'} \le \frac{L}{\sqrt{2}}\|\nabla u\|_{\Omega'}. \tag{5.2.7}$$

Combining (5.2.7) with Assumption 5.2.1(ii), we obtain that for any subdomain $\Omega' \subset \Omega$ with diameter H,

$$\|u\|_{\Omega'} \le \frac{H}{\sqrt{2}}\|\nabla u\|_{\Omega'} \le \frac{H}{\sqrt{2}}\|u\|_{1,k,\Omega'} \qquad \text{for all } u \in H_0^1(\Omega'). \tag{5.2.8}$$

5.2.2 Abstract Spectral Theory for Indefinite Generalised Eigenvalue Problems

We shall build our coarse space by solving local generalised eigenvalue problems, which may be indefinite (see (5.2.32)). So in this section we derive, in an abstract setting, the required properties of such problems. First we derive some properties of a positive semidefinite bilinear form (such as appears on the right-hand side of (5.2.32)).

Proposition 5.2.5. *Suppose $\widetilde{V}$ is a real vector space with $\dim \widetilde{V} = n$. Let c be a positive semidefinite bilinear form on $\widetilde{V}$, set $\ker \mathsf{c} := \{v \in \widetilde{V} : \mathsf{c}(v,w) = 0, \text{ for all } w \in \widetilde{V}\}$ and suppose $\dim\ker \mathsf{c} = s < n$. Further, let $\widehat{V} \subset \widetilde{V}$ be an $(n-s)$-dimensional subspace such that $\widehat{V} \cap \ker \mathsf{c} = \{0\}$. Then $\widetilde{V} = \widehat{V} \oplus \ker \mathsf{c}$, where $\oplus$ denotes a direct sum, and c is positive definite on $\widehat{V}$.*

Proof. The relation $\widetilde{V} = \widehat{V} \oplus \ker \mathsf{c}$ is clear since $\dim \widehat{V} + \dim\ker \mathsf{c} = n$ and $\widehat{V} \cap \ker \mathsf{c} = \{0\}$. The form $\mathsf{c}|_{\widehat{V}\times\widehat{V}}$ is symmetric and satisfies $\mathsf{c}(v,v) > 0$ for $v \in \widehat{V} \setminus \{0\}$ by the assumption on $\widehat{V}$. Hence $\mathsf{c}|_{\widehat{V}\times\widehat{V}}$ is positive definite. □

The following is our main abstract theorem.

Theorem 5.2.6. *Let $\widetilde{V}$, $\mathtt{c}$ and s be as in Proposition 5.2.5 and suppose also that $\mathtt{b}$ is a symmetric (possibly indefinite) bilinear form on $\widetilde{V}$. Consider the generalised eigenvalue problem: find $\lambda \in \mathbb{R}$ and $p \in \widetilde{V}\backslash\{0\}$ such that*

$$\mathtt{b}(p,v) = \lambda \mathtt{c}(p,v) \quad \text{for all} \quad v \in \widetilde{V}. \tag{5.2.9}$$

(i) *If $s \geq 1$, suppose also that $\mathtt{b}$ is positive definite on $\ker \mathtt{c}$. Then $\widetilde{V}$ has a basis $\{p_j : j = 1,\dots,n\}$, in which $\{p_j : j = 1,\dots,n-s\}$ are eigenvectors of (5.2.9) which can be chosen to be orthonormal with respect to $\mathtt{c}$ and the corresponding finite eigenvalues can be ordered*

$$-\infty < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{n-s} < \infty. \tag{5.2.10}$$

The remaining basis vectors $\{p_j : j = n-s+1,\dots,n\}$ form a basis of $\ker \mathtt{c}$. In this sense we say that they correspond to infinite eigenvalues of (5.2.9). Moreover,

$$\mathtt{b}(p_j, p_{j'}) = \lambda_j \delta_{j,j'}, \qquad j \in \{1,\dots,n-s\},\ j' \in \{1,\dots,n\}; \tag{5.2.11}$$
$$\mathtt{c}(p_j, p_{j'}) = \delta_{j,j'}, \qquad j \in \{1,\dots,n-s\},\ j' \in \{1,\dots,n\}. \tag{5.2.12}$$

(ii) *If $s = 0$, the same conclusion holds but the positive definiteness condition on $\mathtt{b}$ is not required, and all eigenvalues are finite.*

Proof. Suppose $s \geq 1$ and $\mathtt{b}$ is positive definite on $\ker \mathtt{c}$. If we choose $\widehat{V} \subset \widetilde{V}$ so that $\widetilde{V} = \widehat{V} \oplus \ker \mathtt{c}$ (as in Proposition 5.2.5), then every $p \in \widetilde{V}$ has a unique decomposition $p = \widehat{p} + p_{\mathtt{c}}$ with $\widehat{p} \in \widehat{V}$ and $p_{\mathtt{c}} \in \ker \mathtt{c}$. Thus the GEVP (5.2.9) is equivalent to seeking $\lambda \in \mathbb{R}$, $\widehat{p} \in \widehat{V}$ and $p_{\mathtt{c}} \in \ker \mathtt{c}$, such that $\widehat{p} + p_{\mathtt{c}} \neq 0$ and

$$\mathtt{b}(\widehat{p} + p_{\mathtt{c}}, v) = \lambda \mathtt{c}(\widehat{p}, v) \qquad \text{for all} \quad v \in \widetilde{V}.$$

This, in turn, is equivalent to solving the coupled system

$$\mathtt{b}(\widehat{p}, v_{\mathtt{c}}) + \mathtt{b}(p_{\mathtt{c}}, v_{\mathtt{c}}) = 0 \qquad \text{for all} \quad v_{\mathtt{c}} \in \ker \mathtt{c}, \tag{5.2.13}$$
$$\mathtt{b}(\widehat{p}, \widehat{v}) + \mathtt{b}(p_{\mathtt{c}}, \widehat{v}) = \lambda \mathtt{c}(\widehat{p}, \widehat{v}) \qquad \text{for all} \quad \widehat{v} \in \widehat{V}, \tag{5.2.14}$$

for $\lambda \in \mathbb{R}$ and $\widehat{p} \in \widehat{V}$ and $p_{\mathtt{c}} \in \ker \mathtt{c}$, such that $\widehat{p} + p_{\mathtt{c}} \neq 0$.

Now, by assumption on $\mathtt{b}$, for each $\widehat{w} \in \widehat{V}$, there exists a unique $w_{\mathtt{c}} \in \ker \mathtt{c}$ such that

$$\mathtt{b}(w_{\mathtt{c}}, v_{\mathtt{c}}) = -\mathtt{b}(\widehat{w}, v_{\mathtt{c}}) \qquad \text{for all} \quad v_{\mathtt{c}} \in \ker \mathtt{c}. \tag{5.2.15}$$

Denoting the solution to (5.2.15) as $w_{\mathtt{c}} = \mathcal{S}\widehat{w}$, this defines a linear map $\mathcal{S} : \widehat{V} \to \ker \mathtt{c}$ with the property

$$\mathtt{b}(\mathcal{S}\widehat{w}, v_{\mathtt{c}}) = -\mathtt{b}(\widehat{w}, v_{\mathtt{c}}) \qquad \text{for all} \quad \widehat{w} \in \widehat{V} \quad \text{and} \quad v_{\mathtt{c}} \in \ker \mathtt{c},$$

or, equivalently,

$$\mathtt{b}\big((I + \mathcal{S})\widehat{w}, v_{\mathtt{c}}\big) = 0 \qquad \text{for all} \quad \widehat{w} \in \widehat{V} \quad \text{and} \quad v_{\mathtt{c}} \in \ker \mathtt{c}. \tag{5.2.16}$$

We now consider the GEVP: find $\lambda \in \mathbb{C}$ and $\widehat{p} \in \widehat{V}\backslash\{0\}$ such that

$$\widetilde{\mathtt{b}}(\widehat{p}, \widehat{v}) := \mathtt{b}\big((I + \mathcal{S})\widehat{p}, \widehat{v}\big) = \lambda \mathtt{c}(\widehat{p}, \widehat{v}) \qquad \text{for all} \quad \widehat{v} \in \widehat{V}. \tag{5.2.17}$$

(This is the Schur complement of (5.2.13), (5.2.14).)

Note that $\mathcal{S}$ is self-adjoint with respect to $\mathtt{b}$. This is because, using the symmetry of $\mathtt{b}$ (at the first and third steps) and the definition of $\mathcal{S}$ (at the second and fourth steps),

$$\mathtt{b}(\mathcal{S}\widehat{p}, \widehat{v}) = \mathtt{b}(\widehat{v}, \mathcal{S}\widehat{p}) = -\mathtt{b}(\mathcal{S}\widehat{v}, \mathcal{S}\widehat{p}) = -\mathtt{b}(\mathcal{S}\widehat{p}, \mathcal{S}\widehat{v}) = \mathtt{b}(\widehat{p}, \mathcal{S}\widehat{v}).$$

Thus the bilinear form $\widetilde{\mathsf{b}}$ on the left-hand side of (5.2.17) is symmetric on $\widehat{V}$. Now c is symmetric and (by Proposition 5.2.5) positive definite on $\widehat{V}$; so the GEVP (5.2.17) has $n-s$ eigenpairs, which we denote $(\lambda_j, \widehat{p}_j)$, $j = 1, \ldots, n-s$, with real finite eigenvalues λ_j, which can be ordered as in (5.2.10). The $\widehat{p}_j \in \widehat{V}$ can be chosen orthonormal with respect to c and form a basis of $\widehat{V}$. Now we define

$$p_j = (I + \mathcal{S})\widehat{p}_j, \qquad j = 1, \ldots, n-s. \tag{5.2.18}$$

By (5.2.18), (5.2.17) and since $\mathcal{S}\widehat{p}_j \in \ker \mathsf{c}$, we have

$$\mathsf{b}(p_j, \widehat{v}) = \widetilde{\mathsf{b}}(\widehat{p}_j, \widehat{v}) = \lambda_j \mathsf{c}(\widehat{p}_j, \widehat{v}) = \lambda_j \mathsf{c}(p_j, \widehat{v}) \qquad \text{for all } \widehat{v} \in \widehat{V}.$$

Also, by (5.2.18) and (5.2.16) we have $\mathsf{b}(p_j, v_\mathsf{c}) = 0$, for all $v_\mathsf{c} \in \ker \mathsf{c}$ and so (see (5.2.13) and (5.2.14)), (λ_j, p_j), $j = 1, \ldots, n-s$ are eigenpairs of the original GEVP (5.2.9). Choosing an arbitrary basis, $\{p_{n-s+1}, \ldots, p_n\}$, of $\widehat{V}$ we obtain a basis of $\widetilde{V}$.

For $j \in \{1, \ldots, n-s\}$, relation (5.2.12) follows for $j' \in \{1, \ldots, n-s\}$ because $\mathcal{S}\widehat{p}_j, \mathcal{S}\widehat{p}_{j'} \in \ker \mathsf{c}$, and so $\mathsf{c}(p_j, p_{j'}) = \mathsf{c}(\widehat{p}_j, \widehat{p}_{j'}) = \delta_{m,m'}$ since the $\{\widehat{p}_j : j = 1, \ldots, n-s\}$ were chosen orthonormal with respect to c. For $j' \in \{n-s+1, \ldots, n\}$, (5.2.12) is trivial because then $p_{j'} \in \ker \mathsf{c}$.

Relation (5.2.11) then follows directly from (5.2.12) because, for $j \in \{1, \ldots, n-s\}$ and any $j' \in \{1, \ldots, n\}$, $\mathsf{b}(p_j, p_{j'}) = \lambda_j \mathsf{c}(p_j, p_{j'})$.

This completes the proof when $s \geq 1$, while the case $s = 0$ is trivial since (5.2.9) can then be reduced to a standard eigenvalue problem. □

Remark 5.2.7. *1. The requirement that* b *is positive definite on* $\ker \mathsf{c}$ *in part* (i) *of Theorem 5.2.6 could be weakened to simply requiring that problem* (5.2.15) *has a unique solution for all* $\widehat{w} \in \widehat{V}$*, where* $\widehat{V}$ *is as in Proposition 5.2.5. However, since we shall need the positive definiteness in Lemma 5.2.9 anyway, we avoid this extra generality.*

2. Note that the eigenvectors p_j, $j = 1, \ldots, n-s$ *defined via* (5.2.18) *are, in general, not elements of* $\widehat{V}$ *because* $\mathcal{S}$ *maps into* $\ker \mathsf{c}$.

The following corollary then follows directly from Theorem 5.2.6.

Corollary 5.2.8. *Under the conditions of Theorem 5.2.6, every* $v \in V$ *can be expressed as*

$$v = v_0 + \sum_{j=1}^{n-s} \mathsf{c}(v, p_j) p_j \quad \textit{with} \quad v_0 \in \ker \mathsf{c}, \tag{5.2.19}$$

If $\ker \mathsf{c} = \{0\}$ *then* $s = 0$ *and* $v_0 = 0$.

We now define a projection operator mapping V onto the span of the first m eigenvectors of (5.2.9), where m is chosen so that all eigenvectors corresponding to negative eigenvalues are included (recall the ordering (5.2.10)). This operator is crucial in the analysis of the proposed preconditioner.

Lemma 5.2.9. *Under the conditions of Theorem 5.2.6, suppose that there exists* $m \in \{1, \ldots, n-s-1\}$ *such that* $\lambda_{m+1} > 0$*, and define the projector*

$$\Pi v = \sum_{j=1}^{m} \mathsf{c}(v, p_j) p_j.$$

Then

$$\mathsf{c}(v - \Pi v, v - \Pi v) \leq \frac{1}{\lambda_{m+1}} \mathsf{b}(v - \Pi v, v - \Pi v).$$

In order to later apply this lemma we will show (Theorem 5.3.8) that such an m exists, under a modest assumption.

Proof. Let $v \in \widetilde{V}$ be expressed as in (5.2.19) and set $\alpha_l := \mathtt{c}(v, p_l)$, $l \in \{1, \dots, n-s\}$. We use the properties from Theorem 5.2.6 to obtain,

$$\begin{aligned}\mathtt{b}(v - \Pi v, v - \Pi v) &= \mathtt{b}\left(v_0 + \sum_{j=m+1}^{n-s} \alpha_j p_j \,,\, v_0 + \sum_{j=m+1}^{n-s} \alpha_j p_j\right) \\ &= \mathtt{b}(v_0, v_0) + 2\sum_{j=m+1}^{n-s} \alpha_j \mathtt{b}(p_j, v_0) + \sum_{j=m+1}^{n-s} \lambda_j \alpha_j^2 \\ &= \mathtt{b}(v_0, v_0) + 2\sum_{j=m+1}^{n-s} \alpha_j \lambda_j \mathtt{c}(p_j, v_0) + \sum_{j=m+1}^{n-s} \lambda_j \alpha_j^2\end{aligned}$$

Then we use the the fact that $v_0 \in \ker \mathtt{c}$ and the positive definiteness of $\mathtt{b}$ on $\ker \mathtt{c}$ to obtain

$$\begin{aligned}\mathtt{b}(v - \Pi v, v - \Pi v) &= \mathtt{b}(v_0, v_0) + \sum_{j=m+1}^{n-s} \lambda_j \alpha_j^2 \\ &\geq \sum_{j=m+1}^{n-s} \lambda_j \alpha_j^2 \geq \lambda_{m+1} \sum_{j=m+1}^{n-s} \alpha_j^2. \end{aligned} \tag{5.2.20}$$

Using an almost identical argument and the results from Theorem 5.2.6 we obtain

$$\mathtt{c}(v - \Pi v, v - \Pi v) = \sum_{j=m+1}^{n-s} \alpha_j^2. \tag{5.2.21}$$

The result follows from (5.2.20) and (5.2.21). □

This lemma shows that, although the bilinear form $\mathtt{b}(\cdot, \cdot)$ may be indefinite on the full space, it becomes positive definite on the complement of the subspace spanned by the first m eigenvectors, provided all discarded eigenvalues λ_j (for $j > m$) are strictly positive. In the context of coarse space construction, this implies that if we retain all eigenfunctions associated with non-positive or small eigenvalues, then the remainder is well-controlled by $\mathtt{b}(\cdot, \cdot)$, leading to stability estimates essential for robust preconditioning.

5.2.3 Domain Decomposition

Notation 5.2.10. *Let Ω' be any subdomain of Ω, composed of a union of elements of the mesh $\mathcal{T}_h$. We introduce the finite element spaces:*

$$\widetilde{V}_{\Omega'} := \left\{v|_{\Omega'} : v \in V^h\right\} \subset H^1(\Omega') \qquad \text{and} \qquad V_{\Omega'} := \left\{v \in \widetilde{V}_{\Omega'} : v|_{\partial\Omega'} = 0\right\} \subset H_0^1(\Omega').$$

For any $v_{\Omega'} \in V_{\Omega'}$, we let $E_{\Omega'} v_{\Omega'} \in V^h$ denote its zero extension to the whole domain Ω, and we define the restriction operator $R_{\Omega'} : V^h \to V_{\Omega'}$ to be the $L^2(\Omega)$ adjoint of $E_{\Omega'}$, i.e.,

$$(R_{\Omega'} v, w_{\Omega'})_{\Omega'} := (v, E_{\Omega'} w_{\Omega'}) \quad \text{for all} \quad v \in V^h \quad \text{and} \quad w_{\Omega'} \in V_{\Omega'}.$$

Remark 5.2.11. *Although $b_{\Omega'}(\cdot, \cdot)$ (see Notation 5.2.2) is generally indefinite, for small enough diameter of D it becomes positive definite. This is discussed in Lemma 5.3.16.*

In order to construct the two-level Schwarz preconditioner, we choose two overlapping covers of Ω. The first consists of 'local subdomains' on which restrictions of (5.2.4) will be solved. The second consists of 'coarse-space subdomains' on which generalised eigenvalue problems will be solved to build the coarse space. There is no need for these covers to coincide and in practical algorithms it may be convenient for them to be different. Both covers can be constructed by first choosing a non-overlapping partition and then expanding each subdomain by adding one or more layers of elements; see, e.g. [89, Section 3.2].

The local and coarse space subdomains will be denoted $\{\Omega_j^\ell\}_{j=1}^Q$ and $\{\Omega_i^c\}_{i=1}^N$ respectively, with all subdomains consisting of unions of mesh elements $T \in \mathcal{T}_h$, and such that

$$\Omega = \bigcup_{i=1}^N \Omega_i^c, \qquad \Omega = \bigcup_{j=1}^Q \Omega_j^\ell.$$

For each, Ω_i^c and Ω_j^ℓ, we denote their diameters by $H_{c,i}$, $H_{\ell,j}$ respectively, and we set $H_c := \max_{i=1}^N H_{c,i}$ and $H_\ell := \max_{j=1}^Q H_{\ell,j}$.

We define the corresponding finite element spaces $V_{\Omega_i^c}$, $\widetilde{V}_{\Omega_i^c}$ and $V_{\Omega_j^\ell}$, $\widetilde{V}_{\Omega_j^\ell}$ as in Notation 5.2.10.

Standard properties for such overlapping covers are (see, e.g., [58, Eq. (2.10)]):

$$\sum_{i=1}^N \|v|_{\Omega_i^c}\|^2_{1,k,\Omega_i^c} \le \Lambda_c \|v\|^2_{1,k}, \qquad \sum_{j=1}^Q \|v|_{\Omega_j^\ell}\|^2_{1,k,\Omega_j^\ell} \le \Lambda_\ell \|v\|^2_{1,k} \qquad \text{for all } v \in V^h, \tag{5.2.22}$$

where

$$\Lambda_\circ := \max_{T\in\mathcal{T}_h} \big(\#\{\Omega_i^\circ \mid 1 \le i \le N,\, T \subseteq \Omega_i^\circ\}\big) \qquad \text{for } \circ \in \{c, \ell\}.$$

In addition ([58, Lemma 3.6]), for any set of functions $v_j \in H_0^1(\Omega)$ with $\operatorname{supp} v_j \subseteq \overline{\Omega_j^\ell}$ for each $j = 1, \dots, Q$, we have the estimate

$$\left\| \sum_{j=1}^Q v_j \right\|^2_{1,k} \le \Lambda_\ell \sum_{j=1}^Q \|v_j|_{\Omega_j^\ell}\|^2_{1,k,\Omega_j^\ell}, \tag{5.2.23}$$

and, similarly,

$$\left\| \sum_{i=1}^N v_i \right\|^2_{1,k} \le \Lambda_c \sum_{i=1}^N \|v_i|_{\Omega_i^c}\|^2_{1,k,\Omega_i^c},$$

for functions $v_i \in H_0^1(\Omega)$ with $\operatorname{supp} v_i \subseteq \overline{\Omega_i^c}$, for each $i = 1, \dots, N$.

The one-level additive Schwarz preconditioner can now be given in matrix form as

$$\mathbf{M}_{AS,1}^{-1} = \sum_{j=1}^Q \mathbf{E}_j^\ell (\mathbf{B}_j^\ell)^{-1} \mathbf{R}_j^\ell, \qquad \text{where } \mathbf{B}_j^\ell = \mathbf{R}_j^\ell \mathbf{B} \mathbf{E}_j^\ell;$$

here, $\mathbf{E}_j^\ell$ and $\mathbf{R}_j^\ell$ denote the matrix representations of

$$E_j^\ell := E_{\Omega_j^\ell} \quad \text{and} \quad R_j^\ell := R_{\Omega_j^\ell},$$

respectively, with respect to the basis functions $\{\phi_i\}_{i=1}^n$ of V^h.

In order to improve the preconditioner, a global coarse space is added.

Let $V_0 \subset V^h$ be such a coarse space, let $E_0 : V_0 \to V^h$ be the natural embedding, and let R_0 be the L^2 adjoint of E_0,

$$(R_0 w, v_0) = (w, E_0 v_0) \qquad \text{for all } w \in V^h, v_0 \in V_0.$$

The two-level additive Schwarz preconditioner is then

$$\mathbf{M}_{AS,2}^{-1} = \mathbf{E}_0\mathbf{B}_0^{-1}\mathbf{R}_0 + \mathbf{M}_{AS,1}^{-1}, \quad \text{where } \mathbf{B}_0 := \mathbf{R}_0\mathbf{B}\mathbf{E}_0$$

(with $\mathbf{E}_0$ and $\mathbf{R}_0$ denoting matrix representations of E_0 and R_0). The preconditioned version of (5.2.4) is

$$\mathbf{M}_{AS,2}^{-1}\mathbf{B}\mathbf{u} = \mathbf{M}_{AS,2}^{-1}\mathbf{f}. \tag{5.2.24}$$

For the analysis we need certain projection operators defined as follows. For each $j = 1, \ldots, Q$, we define $T_j^\ell : V^h \to V_{\Omega_j^\ell}$, by,

$$b_{\Omega_j^\ell}(T_j^\ell u, v) = b(u, E_j^\ell v) \qquad \text{for all } u \in V^h \text{ and } v \in V_{\Omega_j^\ell}, \tag{5.2.25}$$

and $T_0 : V^h \to V_0$, by

$$b(T_0 u, v) = b(u, E_0 v) = b(u, v) \qquad \text{for all } u \in V^h \text{ and } v \in V_0. \tag{5.2.26}$$

Sufficient conditions for the existence of T_j^ℓ and T_0 are given in Lemmas 5.3.16 and 5.3.19. Given the operators T_j^ℓ and T_0, we define $T : V^h \to V^h$ by

$$T = E_0 T_0 + \sum_{j=1}^{Q} E_j^\ell T_j^\ell. \tag{5.2.27}$$

Then the preconditioned system (5.2.24) is related to T via the following result (see, e.g., [56, Theorem 5.5]).

Proposition 5.2.12. *For any $u, v \in V^h$, with corresponding nodal vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$,*

$$\langle \mathbf{M}_{AS,2}^{-1}\mathbf{B}\mathbf{u}, \mathbf{v}\rangle_{\mathbf{D}_k} = (Tu, v)_{1,k}, \tag{5.2.28}$$

where $\langle \cdot, \cdot \rangle_{\mathbf{D}_k}$ is the $\mathbf{D}_k$-weighted inner product on $\mathbb{R}^n$, with $\mathbf{D}_k$ given by

$$\mathbf{D}_k := \mathbf{A} + k^2\mathbf{S}. \tag{5.2.29}$$

5.2.4 The H_k-GenEO Coarse Space

In order to define the H_k-GenEO coarse space, we need some definitions from [85].

Definition 5.2.13 ([85] Definition 3.2)**.** *Given any $\Omega' \subseteq \Omega$, formed from a union of elements $T \in \mathcal{T}_h$, let*

$$\overline{\text{dof}}(\Omega') := \{j \mid 1 \le j \le n \text{ and } \text{supp}(\phi_j) \cap \Omega' \neq \emptyset\}$$

denote the set of degrees of freedom that are active in the subdomain Ω', including freedoms on the boundary. In a similar manner, let

$$\text{dof}(\Omega') := \{j \mid 1 \le j \le n \text{ and } \text{supp}(\phi_j) \subset \overline{\Omega'}\}$$

denote the internal degrees of freedom.

Definition 5.2.14 (Multiplicative POU on the local subdomains)**.** *We choose a partition of unity consisting of continuous real-valued functions $\{\Phi_j^\ell\}_{j=1}^Q$ on Ω with $0 \le \Phi_j^\ell \le 1$ and $\text{supp}(\Phi_j^\ell) \subseteq \overline{\Omega_j^\ell}$, for each j, and, in addition, $\sum_{j=1}^Q \Phi_j^\ell(x) = 1$ on Ω. Moreover*

$$\|\nabla \Phi_j^\ell\|_{L^\infty(\Omega)} \lesssim \frac{1}{\delta} \quad \text{for all } j,$$

where δ is proportional to the width of the overlap between neighbouring subdomains of the cover $\{\Omega_j^\ell\}_{j=1}^Q$. Such a POU with each Φ_j^ℓ continuous piecewise linear (with respect to the mesh $\mathcal{T}_h$) is explicitly constructed in [89, Section 3.2].

Remark 5.2.15. *By [58, estimate (3.2)], for any $v \in H^1_0(\Omega)$,*

$$\sum_{j=1}^{Q} \|\Phi_j^\ell v\|^2_{1,k,\Omega_j^\ell} \lesssim \Lambda_\ell \left(1 + \frac{1}{(k\delta)^2}\right) \|v\|^2_{1,k}, \tag{5.2.30}$$

We use a different POU on the coarse subdomains.

Definition 5.2.16 (Operator POU on the coarse subdomains)**.** *For any $l \in \operatorname{dof}\Omega$, let μ_l denote the number of subdomains for which l is an internal degree of freedom, i.e.*

$$\mu_l := \#\big\{i \mid 1 \le i \le L,\, l \in \operatorname{dof}(\Omega_i^c)\big\}.$$

Then, for each $i = 1, \dots, N$, we define the coarse space partition of unity operator, $\Xi_i : V \to V_{\Omega_i^c}$ as a weighted combination of degrees of freedom v_l

$$\Xi_i^c(v) := \sum_{l \in \operatorname{dof}(\Omega_i^c)} \frac{1}{\mu_l} v_l \phi_l^i, \quad \text{where} \quad \phi_l^i := \phi_l|_{\Omega_i^c}.$$

It is easy to verify that this is also a partition of unity, in the sense that

$$\sum_{i=1}^{N} E_i^c \Xi_i^c(v|_{\Omega_i^c}) = v \qquad \text{for all } \ v \in V^h.$$

Remark 5.2.17 (Why two different partitions of unity?)**.** *The smooth partition of unity $\{\Phi_j^\ell\}$ on the local subdomains is used in the stable decomposition and analysis of the additive Schwarz operator, where gradient bounds are essential to control the energy norm of localized components. In contrast, the algebraic partition $\{\Xi_i^c\}$ defined over the coarse subdomains is used in constructing the coarse space and assembling local eigenvalue problems. Its role is purely algebraic and leverages finite element basis functions; smoothness is not required since it operates on discrete spaces. Having two distinct partitions provides the necessary flexibility to decouple the requirements of stability and coarse space construction.*

The local generalised eigenvalue problem that is going to form the basis of the coarse space can now be introduced.

Definition 5.2.18 (Local generalised eigenvalue problem)**.** *For each $i \in \{1, \dots, N\}$, define $b_{\Omega_i^c}$ as in (5.2.2) and set*

$$c_{\Omega_i^c}(w, v) := \big(\Xi_i^c(w), \Xi_i^c(v)\big)_{1,k,\Omega_i^c} \qquad \text{for all } w, v \in \widetilde{V}_{\Omega_i^c}. \tag{5.2.31}$$

The generalised eigenvalue problem is then to find $p^i \in \widetilde{V}_{\Omega_i^c} \setminus \{0\}$ and $\lambda^i \in \mathbb{R}$, such that

$$b_{\Omega_i^c}(p^i, v) = \lambda^i \, c_{\Omega_i^c}(p^i, v) \qquad \text{for all } v \in \widetilde{V}_{\Omega_i^c}. \tag{5.2.32}$$

Definition 5.2.19 (H_k-GenEO Coarse space)**.** *Let (p_m^i, λ_m^i) be eigenpairs for (5.2.32) with $\lambda_1^i, \lambda_2^i, \dots$ chosen in non-decreasing order. For each $i = 1, \dots, N$, let m_i be such that $\lambda_{m_i+1}^i > 0$. (A mild sufficient condition for the existence of such m_i is given in Theorem 5.3.8 below.)*

The coarse space, V_0, is then given by

$$V_0 := \operatorname{span}\big\{E_i^c \Xi_i^c(p_m^i) : \ m = 1, \dots, m_i \ \text{ and } i = 1, \dots, N\big\}.$$

Remark 5.2.20. *In contrast to previous works on the GenEO construction (e.g., [85], [13] and [3]) (where the forms are always at least semi-definite), the left-hand side of (5.2.32) can here be indefinite and also the right-hand side in (5.2.32) is based on a k-weighted scalar product. However, as $k \to 0$ the GEVP (5.2.31) approaches the GEVP solved in the classical GenEO method, e.g. [85].*

In the proof of our main result we also need the following estimates for $\|\Xi_i^c(v)\|_{\Omega_i^c}$.

Lemma 5.2.21.

$$\|\Xi_i^c(v)\|_{\Omega_i^c} \lesssim \|v\|_{\Omega_i^c} \qquad \text{for all } v \in \widetilde{V}_i, \tag{5.2.33}$$

$$\|\Xi_i^c(v)\|_{\Omega_i^c} \gtrsim \Lambda_c^{-1}\|v\|_{\Omega_i^c} \qquad \text{for all } v \in V_i. \tag{5.2.34}$$

Proof. Due to the assumption that we are working with Lagrange elements, for any $T \in \mathcal{T}_h$, there is an affine map F_T mapping the unit simplex $\widehat{T}$ to T. Then (see, e.g. [21, Theorem 3.1.2])), for any $w \in V^h$, setting $\widehat{w} = w \circ F_T$, we have

$$\|w\|_T^2 \;\sim\; h_T^d\|\widehat{w}\|_{\widehat{T}}^2 \;\sim\; h_T^d \sum_{j\in\mathrm{dof}(T)} w_j^2, \tag{5.2.35}$$

where w_j is the value of w at $j \in \mathrm{dof}(T)$ and the hidden constants are independent of the mesh and w. (The second relation in (5.2.35) uses equivalence of norms on the space of polynomials of degree r.) Hence, by Definition 5.2.16 of $\Xi_i^c(v)$ and since degrees of freedom of $\Xi_i^c(v)$ are all interior to Ω_i^c, we have:

$$\begin{aligned}\|\Xi_i^c(v)\|_{\Omega_i^c}^2 &= \sum_{T\subset\overline{\Omega_i^c}} \|\Xi_i^c(v)\|_T^2 \sim \sum_{T\subset\overline{\Omega_i^c}} h_T^d \sum_{j\in\mathrm{dof}(T)} (\Xi_i^c(v))_j^2 = \sum_{T\subset\overline{\Omega_i^c}} h_T^d \sum_{j\in\mathrm{dof}(T)} \mu_j^{-2} v_j^2 \\ &\geq \Lambda_c^{-2} \sum_{T\subset\overline{\Omega_i^c}} h_T^d \sum_{j\in\mathrm{dof}(T)} v_j^2 \sim \Lambda_c^{-2} \sum_{T\subset\overline{\Omega_i^c}} \|v\|_T^2 = \Lambda_c^{-2}\|v\|_{\Omega_i^c}^2,\end{aligned}$$

proving (5.2.34). The proof of (5.2.33) is analogous but simpler. □

5.3 Statement of the Main Result and Theoretical Tools

In this section we state our main result and develop several technical tools needed for its proof. (The proof of the main result is given in the following section.) Throughout, let Λ_ℓ and Λ_c denote the maximal overlap multiplicities of the fine and coarse coverings, respectively.

We introduce the notation

$$\Theta := \frac{1}{\min_{1\leq i\leq N} \lambda_{m_i+1}^i} = \frac{1}{\tau}, \tag{5.3.1}$$

where, recalling the definition from the introduction,

$$\tau := \min_{1\leq i\leq N} \lambda_{m_i+1}^i,$$

and $\lambda_{m_i+1}^i$ is the first *discarded* local eigenvalue on subdomain Ω_i in the H_k-GenEO selection.

Theorem 5.3.1 (GMRES convergence of the two-level preconditioned system)**.** *Let Assumptions 5.2.1 and (2.2.9) be satisfied and let $k > 0$. Then there exist $h_\star > 0$ and $C_1 > 0$ such that, for given $C_2 \in (0, C_1]$, there exists $C_3 > 0$ so that the following statements hold for all $h \in (0, h_\star)$ satisfying $C_2 \leq (hk)^2 \leq C_1$. Assume H_ℓ and τ are chosen so that*

$$s := 2\big(1 + C_3\Lambda_\ell\Lambda_c^4\Theta\big)\Big(2k\sqrt{C_3}\Lambda_c^2\Theta^{\frac{1}{2}}(1 + C_{\mathrm{stab}}) + 3k\Lambda_\ell H_\ell\Big) < 1. \tag{5.3.2}$$

Consider GMRES with the $\langle\cdot,\cdot\rangle_{\mathbf{D}_k}$ inner product (where $\mathbf{D}_k$ is defined in (5.2.29)), applied to the preconditioned system (5.2.24). Then, after m iterations, the residual $\mathbf{r}^{(m)}$ satisfies

$$\|\mathbf{r}^{(m)}\|_{\mathbf{D}_k}^2 \leq \big(1 - \gamma^2\big)^m \|\mathbf{r}^{(0)}\|_{\mathbf{D}_k}^2,$$

where

$$\gamma := \frac{1 - s}{\big(1 + C_3\Lambda_\ell\Lambda_c^4\Theta\big)\big(18 + 8\Lambda_\ell^2\big)}. \tag{5.3.3}$$

Corollary 5.3.2. *Assume that $k \geq 1$. If (5.3.2) holds, then there exists $C > 0$, independent of all parameters, such that*

$$kH_\ell \leq C \qquad \text{and} \qquad (1 + C_{\text{stab}})^2 k^2 \Theta \leq C. \tag{5.3.4}$$

Conversely, if (5.3.4) holds for $C > 0$ small enough to satisfy

$$2\big(1 + C_3 \Lambda_\ell \Lambda_c^4 C\big)\big(2\Lambda_c^2 \sqrt{C_3}\sqrt{C} + 3\Lambda_\ell C\big) < 1, \tag{5.3.5}$$

then (5.3.2) is satisfied and γ from (5.3.3) is bounded below by a positive constant depending only on Λ_ℓ, Λ_c, and C (but not on A, n or k). In particular, GMRES convergence is robust with respect to k and A.

Proof. Assume first that (5.3.2) is satisfied. Since $\Lambda_\ell \geq 1$ and $\Lambda_c \geq 1$ we have

$$2k\Theta^{\frac{1}{2}} \sqrt{C_3}(1 + C_{\text{stab}}) + 3kH_\ell \leq \frac{1}{2},$$

hence

$$kH_\ell \leq \frac{1}{6} \qquad \text{and} \qquad (1 + C_{\text{stab}})^2 k^2 \Theta \leq \frac{1}{C_3 16},$$

which proves (5.3.4).

Conversely, assume that (5.3.4) holds with $C > 0$ such that (5.3.5) is satisfied. Since $k \geq 1$,

$$\Theta \leq \frac{C}{(1 + C_{\text{stab}})^2 k^2} \leq C.$$

Using this bound, together with (5.3.4) and (5.3.5), we obtain

$$s = 2\big(1+C_3\Lambda_\ell\Lambda_c^4\Theta\big)\Big(2k\sqrt{C_3}\Lambda_c^2\Theta^{\frac{1}{2}}(1+C_{\text{stab}})+3k\Lambda_\ell H_\ell\Big) \leq 2\big(1+C_3\Lambda_\ell\Lambda_c^4 C\big)\big(2\sqrt{C_3}\Lambda_c^2\sqrt{C}+3\Lambda_\ell C\big) < 1,$$

i.e. (5.3.2) holds. Moreover,

$$\gamma = \frac{1 - s}{\big(1 + C_3\Lambda_\ell\Lambda_c^4\Theta\big)(18 + 8\Lambda_\ell^2)} \geq \frac{1 - 2\big(1 + C_3\Lambda_\ell\Lambda_c^4 C\big)\big(2\sqrt{C_3}\Lambda_c^2\sqrt{C} + 3\Lambda_\ell C\big)}{\big(1 + C_3\Lambda_\ell\Lambda_c^4\Theta\big)(18 + 8\Lambda_\ell^2)},$$

which yields the stated lower bound for γ. □

In practice, conditions (5.3.4) constrain the fine subdomain diameter H_ℓ and the number of local eigenmodes retained in the coarse space (through τ), both depending on k.

Remark 5.3.3 (Big picture and interpretation)**.**

1. Two levers for robustness. *The bound (5.3.4) shows that robustness is controlled by: (i) a resolution condition $kH_\ell \lesssim 1$ (the fine subdomains must resolve a fixed fraction of a wavelength), and (ii) a spectral separation condition $(1 + C_{\text{stab}})^2 k^2 \Theta \lesssim 1$, i.e. τ is not too small. Operationally, (ii) means we include in the coarse space all local eigenmodes with* non-positive or small positive *eigenvalues.*

2. Role of τ (and $\Theta = 1/\tau$). *The parameter $\tau = \min_i \lambda^i_{m_i+1}$ is the first* discarded *local eigenvalue. A larger τ (smaller Θ) corresponds to a* richer *coarse space. Our analysis shows that choosing $\tau \gtrsim (1 + C_{\text{stab}})^2 k^2$ suffices for k-robust convergence; cf. the k^4 and k^8 requirements in earlier Δ-GenEO bounds.*

3. Overlap multiplicities. *The factors Λ_ℓ and Λ_c enter only through mild algebraic prefactors in s and γ. Thus, for fixed coverings with bounded multiplicities, the iteration count is independent of k, A, n, and the number of subdomains.*

4. Behaviour near resonances. *The stability constant C_{stab} grows as k^2 approaches a Dirichlet eigenvalue of $-\frac{1}{n}\operatorname{div}(A\nabla\cdot)$. Accordingly, the requirement on τ becomes more stringent near such singular points, which matches the expected deterioration for Dirichlet Helmholtz problems.*

5. Decoupling of scales. *Note that H_c (the coarse partition diameter) does not appear explicitly in (5.3.4); its influence is indirect via Λ_c. This reflects the advantage of the H_k-GenEO construction: the coarse space dimension is governed by the spectral threshold τ, not by the coarse mesh size.*

5.3.1 Properties of the H_k-GenEO Coarse Space

In this subsection we apply the abstract spectral theory from Section 5.2.2 to the local generalised eigenvalue problem (5.2.32). We work on a fixed coarse subdomain Ω_i^c and (in this subsection only) simplify notation as follows.

Notation 5.3.4 (Notation used in Section 5.3.1)**.** *For a given i we set*

$$\widetilde{V}_i := \widetilde{V}_{\Omega_i^c}, \quad V_i := V_{\Omega_i^c}, \quad b_i := b_{\Omega_i^c}, \quad c_i := c_{\Omega_i^c},$$

and

$$n_i := \dim \widetilde{V}_i, \qquad s_i := \dim \ker c_i.$$

The constants appearing below are independent of H_i^c and h, and (unless explicitly stated) independent of the coefficients A and n; we write $\lesssim, \gtrsim, \sim$ with the usual meaning.

To apply Theorem 5.2.6 and Lemma 5.2.9, we must verify that b_i is positive definite on $\ker c_i$. The next lemma establishes this under a mild resolution condition.

Lemma 5.3.5. *We have*

$$\ker c_i \;=\; \operatorname{span}\big\{\, \phi_l^i : l \in \overline{\operatorname{dof}}(\Omega_i^c) \setminus \operatorname{dof}(\Omega_i^c) \,\big\}, \qquad \phi_l^i := \phi_l|_{\Omega_i^c}, \tag{5.3.6}$$

where ϕ_l denotes the global nodal basis function with $(\phi_l)_{l'} = \delta_{l,l'}$. Moreover, there exists $C_1 > 0$ such that, whenever $(kh)^2 \le C_1$,

$$b_i(v,v) \ge \frac{1}{4}\|v\|^2_{1,k,\Omega_i^c} \qquad \text{for all } v \in \ker c_i. \tag{5.3.7}$$

Proof. Characterisation of $\ker c_i$. By Definition 5.2.16, if $v \in \widetilde{V}_i$ satisfies $v_l = 0$ for all $l \in \operatorname{dof}(\Omega_i^c)$ then its local extension $\Xi_i(v) \equiv 0$, hence $c_i(v,w) = 0$ for all $w \in \widetilde{V}_i$, so $v \in \ker c_i$. Conversely, if $c_i(v,w) = 0$ for all $w \in \widetilde{V}_i$, then, using Assumption 5.2.1 (with the scaling $a_{\min} = 1 = n_{\max}$),

$$0 = c_i(v,v) = \int_{\Omega_i^c} \big(\nabla \Xi_i^c(v) \cdot A \nabla \Xi_i^c(v) + k^2 n\, \Xi_i^c(v)^2\big)\,\mathrm{d}x \;\;\Rightarrow\;\; \Xi_i(v) \equiv 0,$$

which implies $v_l = 0$ for all $l \in \operatorname{dof}(\Omega_i^c)$. This yields (5.3.6). See also [85, Lemma 3.11].

Positivity of b_i *on* $\ker c_i$. For any $v \in \widetilde{V}_i$,

$$b_i(v,v) = \int_{\Omega_i^c} \big(\nabla v \cdot (A\nabla v) - k^2 n\, v^2\big)\,\mathrm{d}x \;\ge\; \|\nabla v\|^2_{\Omega_i^c} - k^2 \|v\|^2_{\Omega_i^c}, \tag{5.3.8}$$

using $a_{\min} = 1$ and $n \le n_{\max} = 1$.

Let $v \in \ker c_i$. By (5.3.6), $\operatorname{supp} v$ is contained in the union of fine elements touching $\partial\Omega_i^c$. For any $T \in \mathcal{T}_h$ with $T \subseteq \Omega_i^c$, define the patch

$$\mathcal{E}(T) \;=\; \bigcup \{\, \overline{T'} :\; T' \in \mathcal{T}_h,\; T' \subseteq \Omega_i^c,\; \overline{T'} \cap \overline{T} \neq \emptyset \,\}.$$

The family $\{\mathcal{E}(T)\}$ covers Ω_i^c with uniformly bounded overlap, and $\operatorname{diam}\mathcal{E}(T) \sim h$ (shape-regularity). Since v vanishes on a portion of $\partial\mathcal{E}(T)$ of measure $\gtrsim h^{d-1}$ (because $\operatorname{supp} v$ hugs the subdomain boundary), the generalised Friedrichs' (or Poincaré–trace) inequality (e.g., [89, Lemma A.15]) yields

$$\|v\|_{\mathcal{E}(T)} \lesssim h\,\|\nabla v\|_{\mathcal{E}(T)}.$$

Summing over T and using the bounded overlap,

$$\|v\|_{\Omega_i^c}^2 \lesssim h^2\,\|\nabla v\|_{\Omega_i^c}^2 \qquad \Rightarrow \qquad k^2\|v\|_{\Omega_i^c}^2 \lesssim (kh)^2\,\|\nabla v\|_{\Omega_i^c}^2. \tag{5.3.9}$$

Combining (5.3.8) and (5.3.9) gives

$$b_i(v,v) \geq \big(1 - C'(kh)^2\big)\,\|\nabla v\|_{\Omega_i^c}^2.$$

Moreover,

$$\|v\|_{1,k,\Omega_i^c}^2 = \|\nabla v\|_{\Omega_i^c}^2 + k^2\|v\|_{\Omega_i^c}^2 \leq \big(1 + C''(kh)^2\big)\,\|\nabla v\|_{\Omega_i^c}^2.$$

Hence, for $(kh)^2 \leq \min\{(2C')^{-1}, (2C'')^{-1}\}$,

$$b_i(v,v) \geq \tfrac{1}{2}\,\|\nabla v\|_{\Omega_i^c}^2 \geq \tfrac{1}{4}\,\|v\|_{1,k,\Omega_i^c}^2,$$

which is (5.3.7). □

Remark 5.3.6 (Intuition for Lemma 5.3.5)**.** *The form $c_i(\cdot,\cdot)$ measures a k-weighted energy of a local extension $\Xi_i(\cdot)$; its kernel consists of functions whose interior (subdomain) DOFs vanish, i.e., traces supported near $\partial\Omega_i^c$. On this boundary layer, a patchwise Poincaré/Friedrichs' inequality shows $k^2\|v\|^2 \ll \|\nabla v\|^2$ when kh is small. Thus the negative $-k^2\|v\|^2$ part in $b_i(v,v)$ cannot outweigh the gradient term, and b_i is strictly positive on $\ker c_i$. This is precisely the hypothesis needed to invoke the abstract spectral Lemma 5.2.9.*

We now apply the theory in Section 5.2.2 with $\mathtt{b} := b_i$ (defined in (5.2.2)), $\mathtt{c} := c_i$ (defined in (5.2.31)), $s := s_i$, $n := n_i$, and $m := m_i$. The corresponding eigenvalues (5.2.10) are denoted

$$-\infty < \lambda_1^i \leq \lambda_2^i \leq \cdots \leq \lambda_{n_i - s_i}^i < \infty, \tag{5.3.10}$$

and the corresponding eigenfunctions are $\{p_m^i : m = 1, \ldots, n_i - s_i\} \subseteq \widetilde{V}_i$.

Our next main result is Theorem 5.3.7, which provides a lower bound for the minimal (possibly negative) eigenvalue λ_1^i in (5.3.10). We will use this bound later in the proof of Lemma 5.3.9.

To prepare for this, choose a complement of $\ker c_i$ as

$$\widehat{V}_i := \operatorname{span}\big\{\phi_l^i : l \in \operatorname{dof}(\Omega_i^c)\big\} = V_i \subseteq H_0^1(\Omega_i^c), \tag{5.3.11}$$

which is $(n_i - s_i)$-dimensional and satisfies $\widehat{V}_i \cap \ker c_i = \{0\}$; hence

$$\widetilde{V}_i = \widehat{V}_i \oplus \ker c_i,$$

so the assumptions of Proposition 5.2.5 hold.

We will also use the standard inverse estimate (e.g., [89, Lemma B.27]): for each $i = 1, \ldots, N$, there exists $C_{\mathrm{inv}} > 0$ (independent of h) such that

$$\|\nabla v_i\|_{\Omega_i^c} \leq C_{\mathrm{inv}} h^{-1} \|v_i\|_{\Omega_i^c}, \qquad \forall\, v_i \in \widetilde{V}_i. \tag{5.3.12}$$

Theorem 5.3.7 (Estimate of λ_1^i from below)**.** *Let C_1 and C_{inv} be as in Lemma 5.3.5 and (5.3.12). If $\lambda_1^i < 0$, then*

$$\lambda_1^i \gtrsim \Lambda_c^2\left(\frac{2}{(H_i^c k)^2} - 1 - \frac{4}{n_{\min}}\left(\frac{a_{\max}C_{\mathrm{inv}}^2}{(hk)^2}+1\right)^2\right). \tag{5.3.13}$$

Moreover, for any $C_2 \in (0, C_1]$ and for $(hk)^2$ satisfying $C_2 \le (hk)^2 \le C_1$, we have the uniform bound

$$\lambda_1^i \gtrsim -\Lambda_c^2\left(1 + \frac{4}{n_{\min}}\left(\frac{a_{\max}C_{\mathrm{inv}}^2}{C_2}+1\right)^2\right). \tag{5.3.14}$$

Proof. From the proof of Theorem 5.2.6 (see (5.2.17)), λ_1^i is the minimum eigenvalue of

$$\widetilde{b}_i(\widehat{p}^i, \widehat{v}^i) := b_i\big((I+\mathcal{S}_i)\widehat{p}^i, \widehat{v}^i\big) = \lambda^i\, c_i(\widehat{p}^i, \widehat{v}^i) \qquad \forall\, \widehat{v}^i \in \widehat{V}_i,$$

where $\mathcal{S}_i : \widehat{V}_i \to \ker c_i$ is defined by

$$b_i(\mathcal{S}_i\widehat{v}^i, w_{c_i}) = -b_i(\widehat{v}^i, w_{c_i}) \qquad \forall\, w_{c_i} \in \ker c_i. \tag{5.3.15}$$

Since c_i is positive definite on $\widehat{V}_i$ (Proposition 5.2.5), Rayleigh's principle gives

$$\lambda_1^i = \min_{0\neq\widehat{v}^i\in\widehat{V}_i} \frac{\widetilde{b}_i(\widehat{v}^i,\widehat{v}^i)}{c_i(\widehat{v}^i,\widehat{v}^i)}. \tag{5.3.16}$$

Step 1: control of $\mathcal{S}_i$. Using Lemma 5.3.5 (coercivity of b_i on $\ker c_i$), Assumption 5.2.1, the definition (5.3.15) of $\mathcal{S}_i$, and the inverse estimate (5.3.12), for any $\widehat{v}^i \in \widehat{V}_i$ we obtain

$$\begin{aligned}
\frac{n_{\min}}{4}k^2\|\mathcal{S}_i\widehat{v}^i\|_{\Omega_i^c}^2 &\le b_i(\mathcal{S}_i\widehat{v}^i, \mathcal{S}_i\widehat{v}^i) = -\,b_i(\widehat{v}^i, \mathcal{S}_i\widehat{v}^i)\\
&\le a_{\max}\|\nabla(\mathcal{S}_i\widehat{v}^i)\|_{\Omega_i^c}\,\|\nabla\widehat{v}^i\|_{\Omega_i^c} + k^2\|\mathcal{S}_i\widehat{v}^i\|_{\Omega_i^c}\,\|\widehat{v}^i\|_{\Omega_i^c}\\
&\le (a_{\max}C_{\mathrm{inv}}^2h^{-2}+k^2)\|\mathcal{S}_i\widehat{v}^i\|_{\Omega_i^c}\,\|\widehat{v}^i\|_{\Omega_i^c},
\end{aligned}$$

and thus

$$\|\mathcal{S}_i\widehat{v}^i\|_{\Omega_i^c} \le \frac{4}{n_{\min}}\left(\frac{a_{\max}C_{\mathrm{inv}}^2}{(hk)^2}+1\right)\|\widehat{v}^i\|_{\Omega_i^c}. \tag{5.3.17}$$

Step 2: bounding the cross term. By Cauchy–Schwarz, the inverse estimate, and (5.3.17),

$$\begin{aligned}
\big|b_i(\mathcal{S}_i\widehat{v}^i, \widehat{v}^i)\big| &\le a_{\max}\|\nabla(\mathcal{S}_i\widehat{v}^i)\|\,\|\nabla\widehat{v}^i\| + k^2\|\mathcal{S}_i\widehat{v}^i\|\,\|\widehat{v}^i\|\\
&\le \big(a_{\max}C_{\mathrm{inv}}^2h^{-2}+k^2\big)\,\|\mathcal{S}_i\widehat{v}^i\|\,\|\widehat{v}^i\|\\
&\le \frac{4}{n_{\min}}k^2\left(\frac{a_{\max}C_{\mathrm{inv}}^2}{(hk)^2}+1\right)^2\|\widehat{v}^i\|_{\Omega_i^c}^2.
\end{aligned} \tag{5.3.18}$$

Step 3: lower bound for $b_i(\widehat{v}^i,\widehat{v}^i)$. Since $\widehat{V}_i \subset H_0^1(\Omega_i^c)$ (by (5.3.11)), Friedrichs' inequality (5.2.7) with $L \sim H_i^c$ yields

$$\|\widehat{v}^i\|_{\Omega_i^c}^2 \le \frac{(H_i^c)^2}{2}\,\|\nabla\widehat{v}^i\|_{\Omega_i^c}^2.$$

Therefore,

$$b_i(\widehat{v}^i,\widehat{v}^i) \ge \|\nabla\widehat{v}^i\|_{\Omega_i^c}^2 - k^2\|\widehat{v}^i\|_{\Omega_i^c}^2 \ge \Big(2(H_i^c)^{-2} - k^2\Big)\,\|\widehat{v}^i\|_{\Omega_i^c}^2. \tag{5.3.19}$$

Step 4: combine to bound $\widetilde{b}_i$. Using $\widetilde{b}_i(\widehat{v}^i,\widehat{v}^i) = b_i(\widehat{v}^i,\widehat{v}^i) + b_i(\mathcal{S}_i\widehat{v}^i,\widehat{v}^i)$, together with (5.3.18) and (5.3.19),

$$\widetilde{b}_i(\widehat{v}^i,\widehat{v}^i) \ge \left(2(H_i^c)^{-2} - k^2 - \frac{4}{n_{\min}}k^2\left(\frac{a_{\max}C_{\mathrm{inv}}^2}{(hk)^2}+1\right)^2\right)\|\widehat{v}^i\|_{\Omega_i^c}^2. \tag{5.3.20}$$

Step 5: lower bound for $c_i(\widehat{v}^i, \widehat{v}^i)$. By Lemma 5.2.21 (stability of the local extension) and the definition of c_i,

$$c_i(\widehat{v}^i, \widehat{v}^i) \;\geq\; k^2 \,\|\Xi_i(\widehat{v}^i)\|^2_{\Omega_i^c} \;\gtrsim\; k^2 \Lambda_c^{-2} \,\|\widehat{v}^i\|^2_{\Omega_i^c}. \tag{5.3.21}$$

Step 6: Rayleigh quotient. Assume that $\lambda_1^i < 0$. Then $\min_{\widehat{w}^i \neq 0} \widetilde{b}_i(\widehat{w}^i, \widehat{w}^i) < 0$, and we obtain from (5.3.20) and (5.3.21) that, for every $\widehat{v}^i \in \widehat{V}_i$,

$$\begin{aligned}
\frac{\widetilde{b}_i(\widehat{v}^i, \widehat{v}^i)}{c_i(\widehat{v}^i, \widehat{v}^i)} &\geq \frac{\min_{0 \neq \widehat{w}^i \in \widehat{V}_i} \widetilde{b}_i(\widehat{w}^i, \widehat{w}^i)}{c_i(\widehat{v}^i, \widehat{v}^i)} \geq \frac{\min_{0 \neq \widehat{w}^i \in \widehat{V}_i} \widetilde{b}_i(\widehat{w}^i, \widehat{w}^i)}{\min_{0 \neq \widehat{w}^i \in \widehat{V}_i} c_i(\widehat{w}^i, \widehat{w}^i)} \\
&\geq \Lambda_c^2 \left(\frac{2}{(H_i^c k)^2} - 1 - \frac{4}{n_{\min}} \left(\frac{a_{\max} C_{\mathrm{inv}}^2}{(hk)^2} + 1 \right)^2 \right),
\end{aligned}$$

which, together with (5.3.16), proves (5.3.13). The uniform bound (5.3.14) follows by replacing $(hk)^2$ in the bracket by any value in $[C_2, C_1]$; and if $\lambda_1^i \geq 0$, (5.3.14) holds trivially. □

The following theorem shows that, under a very mild assumption, there exist positive eigenvalues for each subdomain, which implies that we can apply Lemma 5.2.9.

Theorem 5.3.8. *For each $i \in \{1, \dots, N\}$, assume there exists a set $\mathcal{D}_i$ of $s_i + 1$ degrees of freedom in $\mathrm{dof}(\Omega_i^c)$ such that $\operatorname{supp} \phi_l^i \cap \operatorname{supp} \phi_{l'}^i$ has zero d-dimensional measure for each pair $l \neq l'$ in $\mathcal{D}_i$. Then there exists $C_\star > 0$ such that, if $(hk)^2 \leq C_\star$, then for each i there is an index m_i with $\lambda_{m_i+1}^i > 0$.*

Proof. Fix i and enumerate $\mathcal{D}_i = \{l_1, \dots, l_{s_i+1}\}$. Since $l \in \mathrm{dof}(\Omega_i^c)$ implies $\phi_l^i \in V_i \subseteq \widetilde{V}_i$, each $\phi_{l_j}^i$ admits the expansion (by (5.2.19))

$$\phi_{l_j}^i \;=\; v_0^{i,j} + \sum_{m=1}^{n_i - s_i} \alpha_m^{i,j} p_m^i, \qquad v_0^{i,j} \in \ker c_i, \quad \alpha_m^{i,j} \in \mathbb{R}. \tag{5.3.22}$$

Since $\dim \ker c_i = s_i$, there exist coefficients $\beta_1^i, \dots, \beta_{s_i+1}^i$, not all zero, such that $\sum_{j=1}^{s_i+1} \beta_j^i v_0^{i,j} = 0$. Define

$$\psi^i \;:=\; \sum_{j=1}^{s_i+1} \beta_j^i \, \phi_{l_j}^i.$$

Multiplying (5.3.22) by β_j^i and summing over j gives

$$\psi^i \;=\; \sum_{m=1}^{n_i - s_i} \alpha_m^i p_m^i \tag{5.3.23}$$

for suitable coefficients α_m^i (the kernel part cancels by construction).

Step 1 (positivity of $b(\psi^i, \psi^i)$). Let $d_{l_j}^i := \mathrm{diam}(\operatorname{supp} \phi_{l_j}^i)$. By Friedrichs' inequality (5.2.8) on each support patch and $a_{\min} = 1$, $n_{\max} = 1$, for each j,

$$\begin{aligned}
b_{\Omega_i^c}(\phi_{l_j}^i, \phi_{l_j}^i) = \|\phi_{l_j}^i\|^2_{a,\Omega_i^c} - k^2 \|\sqrt{n}\, \phi_{l_j}^i\|^2_{\Omega_i^c} \;&\geq\; \left(\tfrac{2}{(d_{l_j}^i)^2} - k^2 \right) \|\phi_{l_j}^i\|^2_{\Omega_i^c} \\
&= \frac{1}{h^2} \left(2 \Big(\frac{h}{d_{l_j}^i} \Big)^2 - (hk)^2 \right) \|\phi_{l_j}^i\|^2_{\Omega_i^c}.
\end{aligned} \tag{5.3.24}$$

Shape-regular quasi-uniform meshes imply $h/d_{l_j}^i \geq c > 0$, so for $(hk)^2$ small enough the right-hand side of (5.3.24) is positive for all j. Moreover, the pairwise supports of $\phi_{l_j}^i$ have zero measure overlap by assumption, hence $b_{\Omega_i^c}(\phi_{l_j}^i, \phi_{l_{j'}}^i) = 0$ for $j \neq j'$. Therefore $b_{\Omega_i^c}(\psi^i, \psi^i) > 0$.

Step 2 (*contradiction if all eigenvalues are non-positive*). If $\lambda_m^i \le 0$ for all $m = 1, \dots, n_i - s_i$, then from (5.3.23) and (5.2.11),

$$b_{\Omega_i^c}(\psi^i, \psi^i) = \sum_{m=1}^{n_i - s_i} \lambda_m^i \, (\alpha_m^i)^2 \;\le\; 0,$$

contradicting Step 1. Hence at least one eigenvalue is strictly positive. If m_i is chosen so that $\lambda_{m_i}^i \le 0 < \lambda_{m_i+1}^i$, the claim follows. □

Lemma 5.3.9 below is a generalisation of [13, Lemma 3.1], [85, Lemma 2.11] and [3, Lemma 3.14] to the indefinite generalised eigenvalue problem (5.2.32). It gives an error estimate for a local projection operator, when it is used to approximate a function $v \in \widetilde{V}_i$ in the space being spanned by the eigenfunctions from Definition 5.2.18.

Lemma 5.3.9. *Let $C_2 \le C_1$ and C_{inv} be as in Theorem 5.3.7, and assume:*

(i) $C_2 \le (hk)^2 \le C_1$;

(ii) $m_i \in \{1, \dots, n_i - s_i - 1\}$ *is such that* $\lambda_{m_i+1}^i > 0$.

Define the local projector

$$\Pi_{m_i} v := \sum_{m=1}^{m_i} c_i(v, p_m^i)\, p_m^i, \qquad v \in \widetilde{V}_i. \tag{5.3.25}$$

Then there exists a constant $C_3 > 0$ (depending on C_1, C_2, C_{inv} and hidden mesh/overlap constants, and at most cubically on $a_{\max}$ and quadratically on $n_{\min}^{-1}$) such that, with $w := v - \Pi_{m_i} v$,

$$0 \;\le\; b_i(w, w) \;\le\; C_3\, \Lambda_c^2\, \|v\|_{1,k,\Omega_i^c}^2, \tag{5.3.26}$$

$$c_i(w, w) \;\le\; \frac{1}{\lambda_{m_i+1}^i}\, b_i(w, w). \tag{5.3.27}$$

Proof. (a) *Decomposition and non-negativity.* Write $v = v_0 + \sum_{m=1}^{n_i - s_i} \alpha_m p_m^i$ as in (5.2.19), with $v_0 \in \ker c_i$. By the definition (5.3.25), Π_{m_i} keeps the first m_i c_i-orthonormal modes, hence

$$w \;=\; v - \Pi_{m_i} v \;=\; v_0 \;+\; \sum_{m=m_i+1}^{n_i - s_i} \alpha_m\, p_m^i.$$

Because $\lambda_{m_i+1}^i > 0$ by assumption and b_i is positive definite on $\ker c_i$ (Lemma 5.3.5),

$$b_i(w, w) = b_i(v_0, v_0) \;+\; \sum_{m=m_i+1}^{n_i - s_i} \lambda_m^i\, \alpha_m^2 \;\ge\; 0,$$

which proves the left inequality in (5.3.26).

(b) *Upper bound for $b_i(w, w)$.* To obtain the second inequality in (5.3.26), we have to work a little harder. Noting that $b_i(w, \Pi_i v) = 0$, we have

$$b_i(w, w) = b_i(v, v) - b_i(\Pi_{m_i} v, \Pi_{m_i} v) = a_{\Omega_i^c}(v, v) - k^2 \|\sqrt{n} v\|_{\Omega_i^c}^2 - b_i(\Pi_{m_i} v, \Pi_{m_i} v). \tag{5.3.28}$$

Now, to estimate the last term in (5.3.28) we use the orthogonality properties (5.2.11), the ordering (5.3.10), (5.2.12) and finally Theorem 5.3.7 to obtain

$$\begin{aligned} -b_i(\Pi_{m_i} v, \Pi_{m_i} v) = -\sum_{m=1}^{m_i} |\alpha_m^i|^2\, \lambda_m^i &\le -\lambda_1^i \sum_{m=1}^{m_i} |\alpha_m^i|^2 \le \max\{-\lambda_1^i, 0\} \sum_{m=1}^{m_i} |\alpha_m^i|^2 \\ &\le \max\{-\lambda_1^i, 0\} \sum_{m=1}^{n_i - s_i} |\alpha_m^i|^2 = \max\{-\lambda_1^i, 0\} c_i(v, v). \end{aligned} \tag{5.3.29}$$

Hence combining (5.3.28) and (5.3.29) we obtain

$$b_i(w,w) \le a_{\Omega_i^c}(v,v) + \max\{-\lambda_1^i, 0\} c_i(v,v),$$

and the second inequality in (5.3.26) follows directly if $\lambda_1^i \ge 0$.

On the other hand if $\lambda_1^i < 0$ we use (5.3.12) and then (5.2.33) to obtain

$$\begin{aligned} c_i(v,v) &\le a_{\max}\|\nabla \Xi_i(v)\|^2_{\Omega_i^c} + k^2\|\Xi_i(v)\|^2_{\Omega_i^c} \le (a_{\max}C_{\mathrm{inv}}^2 h^{-2} + k^2)\|\Xi_i^c(v)\|^2_{\Omega_i^c} \\ &\le \big(a_{\max}C_{\mathrm{inv}}^2 C_2^{-1} + 1\big)k^2\|\Xi_i^c(v)\|^2_{\Omega_i^c} \lesssim \big(a_{\max}C_{\mathrm{inv}}^2 C_2^{-1} + 1\big)k^2\|v\|^2_{\Omega_i^c} \end{aligned} \tag{5.3.30}$$

Combining this with (5.3.29), (5.3.28) and the lower bound on λ_1^i given in (5.3.14) we obtain

$$b_i(w,w) \le a_{\Omega_i^c}(v,v) + \Lambda_c^2\big[1 + 4n_{\min}^{-1}\big(a_{\max}C_{\mathrm{inv}}^2 C_2^{-1} + 1\big)^2\big]C^*\big(a_{\max}C_{\mathrm{inv}}^2 C_2^{-1} + 1\big)k^2\|v\|^2_{\Omega_i^c},$$

where C^* is the hidden constant in (5.3.30). This proves the second inequality in (5.3.27) when $\lambda_1^i < 0$.

Since $b_i(u,u) = a(u,u) - k^2(nu,u) \le a(u,u) \le a_{\max}\,\|\nabla u\|^2_{\Omega_i^c} \le a_{\max}\,\|u\|^2_{1,k,\Omega_i^c}$,

$$b_i(w,w) \;\le\; a_{\max}\,\|w\|^2_{1,k,\Omega_i^c}. \tag{5.3.31}$$

We now control $\|w\|_{1,k,\Omega_i^c}$ in terms of $\|v\|_{1,k,\Omega_i^c}$. Using the inverse estimate on $\widetilde{V}_i$ and the lower bound $c_i(u,u) \gtrsim k^2\Lambda_c^{-2}\|u\|^2_{\Omega_i^c}$ (Lemma 5.2.21), for any $u \in \widetilde{V}_i$ and $C_2 \le (hk)^2 \le C_1$,

$$\|u\|^2_{1,k,\Omega_i^c} \;\le\; \big(h^{-2} + k^2\big)\|u\|^2_{\Omega_i^c} \;\lesssim\; \left(1 + \frac{1}{(hk)^2}\right)\Lambda_c^2\, c_i(u,u) \;\lesssim\; \frac{\Lambda_c^2}{C_2}\, c_i(u,u). \tag{5.3.32}$$

Since Π_{m_i} is the c_i-orthogonal projector (onto $\mathrm{span}\{p_1^i, \dots, p_{m_i}^i\}$),

$$c_i(w,w) \;=\; c_i(v - \Pi_{m_i}v,\, v - \Pi_{m_i}v) \;\le\; c_i(v,v).$$

Combining this with (5.3.32) for $u = w$ and $u = v$ yields

$$\|w\|^2_{1,k,\Omega_i^c} \;\lesssim\; \frac{\Lambda_c^2}{C_2}\, c_i(w,w) \;\le\; \frac{\Lambda_c^2}{C_2}\, c_i(v,v) \;\lesssim\; \Lambda_c^2\,\|v\|^2_{1,k,\Omega_i^c},$$

where in the last step we used the stability of the local extension (standard in overlapping DDM): $c_i(v,v) = \|\,\Xi_i v\,\|^2_{1,k,\Omega_i^c} \lesssim \Lambda_c^2\,\|v\|^2_{1,k,\Omega_i^c}$. Substituting this into (5.3.31) proves the right inequality in (5.3.26) with a constant C_3 depending on $C_1, C_2, C_{\mathrm{inv}}$ (through the hidden constants in (5.3.32)) and at most linearly on $a_{\max}$.

(c) *The estimate* (5.3.27). Since $\lambda_{m_i+1}^i > 0$, Lemma 5.2.9 (applied with $\mathtt{b} = b_i$, $\mathtt{c} = c_i$ and $m = m_i$) gives

$$c_i(w,w) \;\le\; \frac{1}{\lambda_{m_i+1}^i}\, b_i(w,w),$$

which is exactly (5.3.27). □

Remark 5.3.10 (What these bounds say)**.**

- *Theorem 5.3.8 guarantees a* strictly positive spectral gap *locally: after including all non-positive modes into the coarse space (up to m_i), the next mode has $\lambda_{m_i+1}^i > 0$. This is the key hypothesis that triggers the abstract projection estimate (Lemma 5.2.9).*
- *Lemma 5.3.9 then shows the H_k-GenEO local projector discards only a component w whose b_i-energy is controlled by the $(1,k)$-norm of v (up to Λ_c^2), and, crucially, w satisfies a* local inf–sup *estimate (5.3.27) with constant $\lambda_{m_i+1}^i$. Intuitively: we keep all "dangerous" (non-positive/small) modes; what remains is safely controlled.*

- *The existence of strictly positive local eigenvalues required in Assumption* (ii) *of Lemma 5.3.9 is ensured (under a mild geometric assumption on local DOFs) by Theorem 5.3.8* (proved in the appendix).
- *The projector Π_{m_i} collects all "dangerous" local modes—those with non-positive or very small b_i-energy relative to c_i. The remainder w therefore lives in a subspace where b_i is coercive (by Lemma 5.3.5 together with $\lambda^i_{m_i+1} > 0$), so its c_i-energy is controlled by its b_i-energy with constant $1/\lambda^i_{m_i+1}$—this is exactly (5.3.27). The upper bound (5.3.26) then follows by comparing b_i to the $(1,k)$-norm and using the stability of the local extension (the Λ_c factor).*

From Lemma 5.3.9, we can now pass from a *local* spectral approximation to a *global* one by patching subdomain contributions via stable extensions and a partition of unity. This yields the central *global approximation property* of the H_k-GenEO coarse space.

Lemma 5.3.11 (Global approximation property)**.** *Under the same conditions as in Lemma 5.3.9, define Θ as in (5.3.1). Further, let $v \in V^h$ and set*

$$z_0 := \sum_{i=1}^{N} E_i^c \, \Xi_i^c\big(\Pi_{m_i}^c \, v|_{\Omega_i^c}\big). \tag{5.3.33}$$

Then $z_0 \in V_0$ and

$$\inf_{z \in V_0} \|v - z\|_{1,k}^2 \;\le\; \|v - z_0\|_{1,k}^2 \;\le\; C_3 \, \Lambda_c^4 \, \Theta \, \|v\|_{1,k}^2. \tag{5.3.34}$$

Proof. By construction, $\Pi_{m_i}^c(v|_{\Omega_i^c}) \in \operatorname{span}\{p_1^i, \dots, p_{m_i}^i\}$, hence $z_0 \in V_0$. Using the partition-of-unity stability, the overlap bounds, and Lemma 5.3.9, we obtain

$$\begin{aligned}
\|v - z_0\|_{1,k}^2 &= \Big\| \sum_{i=1}^{N} E_i^c \Xi_i^c\big(v|_{\Omega_i^c}\big) - \sum_{i=1}^{N} E_i^c \Xi_i^c\big(\Pi_{m_i}^c v|_{\Omega_i^c}\big) \Big\|_{1,k}^2 \\
&\le \Lambda_c \sum_{i=1}^{N} \big\| \Xi_i^c\big(v|_{\Omega_i^c} - \Pi_{m_i}^c v|_{\Omega_i^c}\big) \big\|_{1,k,\Omega_i^c}^2 \quad \text{(partition-of-unity stability)} \\
&= \Lambda_c \sum_{i=1}^{N} c_i\big(v|_{\Omega_i^c} - \Pi_{m_i}^c v|_{\Omega_i^c},\, v|_{\Omega_i^c} - \Pi_{m_i}^c v|_{\Omega_i^c}\big) \\
&\le \Lambda_c \sum_{i=1}^{N} \frac{1}{\lambda^i_{m_i+1}} \, b_i\big(v|_{\Omega_i^c} - \Pi_{m_i}^c v|_{\Omega_i^c},\, v|_{\Omega_i^c} - \Pi_{m_i}^c v|_{\Omega_i^c}\big) \quad \text{(Lemma 5.3.9, (5.3.27))} \\
&\le C_3 \, \Lambda_c^3 \sum_{i=1}^{N} \frac{1}{\lambda^i_{m_i+1}} \, \|v|_{\Omega_i^c}\|_{1,k,\Omega_i^c}^2 \quad \text{(Lemma 5.3.9, (5.3.26))} \\
&\le C_3 \Lambda_c^3 \Theta \sum_{i=1}^{N} \|v|_{\Omega_i^c}\|_{1,k,\Omega_i^c}^2 \;\le\; C_3 \Lambda_c^4 \Theta \, \|v\|_{1,k}^2,
\end{aligned}$$

which proves the second inequality in (5.3.34). Since $z_0 \in V_0$, the first inequality follows. □

Remark 5.3.12 (Why (5.3.34) matters in domain decomposition)**.** *The bound* (5.3.34) *is the* approximation property *of the coarse space: for any fine function v, there is a coarse representative z_0 whose error in the $(1,k)$-norm is controlled by Θ (the reciprocal of the first discarded local eigenvalue) and the overlap factor Λ_c. This is the first pillar of two-level methods: the coarse space captures the long-range (global) components efficiently.*

The next result shows that the H_k-GenEO coarse space together with the local subdomain spaces yields a *stable decomposition* of any $v \in V^h$. In the SPD setting, this property is the cornerstone for condition-number bounds of two-level Schwarz [85]. In our indefinite setting, it plays the analogous role for bounding the *field of values* of the preconditioned operator.

Lemma 5.3.13 (Stable decomposition)**.** *Under the same conditions as in Lemma 5.3.9, let $v \in V^h$, define $z_0 \in V_0$ as in (5.3.33), and define, for each $j = 1, \dots, Q$,*

$$z_j^\ell := \Phi_j^\ell \, (v - z_0)\big|_{\Omega_j^\ell} \in V_{\Omega_j^\ell}.$$

Then:

(i) $\quad v = z_0 + \sum_{j=1}^{Q} E_j^\ell z_j^\ell,$

(ii) $\quad \|z_0\|_{1,k}^2 + \sum_{j=1}^{Q} \|z_j^\ell\|_{1,k,\Omega_j^\ell}^2 \lesssim \big(1 + C_3 \Lambda_\ell \Lambda_c^4 \Theta\big) \, \|v\|_{1,k}^2.$

Proof. **(i)** Using $\sum_{j=1}^{Q} \Phi_j^\ell \equiv 1$ and $\operatorname{supp}(\Phi_j^\ell) \subseteq \overline{\Omega_j^\ell}$,

$$\sum_{j=1}^{Q} E_j^\ell z_j^\ell = \sum_{j=1}^{Q} E_j^\ell \big(\Phi_j^\ell (v - z_0)\big) = (v - z_0) \sum_{j=1}^{Q} \Phi_j^\ell = v - z_0,$$

which proves (i).

(ii) By Lemma 5.3.11,

$$\|z_0\|_{1,k}^2 \;\le\; 2\big(\|z_0 - v\|_{1,k}^2 + \|v\|_{1,k}^2\big) \;\le\; 2\big(1 + C_3 \Lambda_c^4 \Theta\big) \|v\|_{1,k}^2.$$

Moreover, by the multiplicative partition-of-unity estimate (cf. (5.2.30)),

$$\sum_{j=1}^{Q} \|z_j^\ell\|_{1,k,\Omega_j^\ell}^2 = \sum_{j=1}^{Q} \|\Phi_j^\ell (v - z_0)\|_{1,k,\Omega_j^\ell}^2 \;\lesssim\; \Lambda_\ell \Big(1 + \frac{1}{(k\delta)^2}\Big) \, \|v - z_0\|_{1,k}^2 \;\lesssim\; \Lambda_\ell \, \|v - z_0\|_{1,k}^2,$$

where in the last step we used $k\delta \gtrsim kh \ge \sqrt{C_2}$. Combining with the bound on z_0 yields

$$\|z_0\|_{1,k}^2 + \sum_{j=1}^{Q} \|z_j^\ell\|_{1,k,\Omega_j^\ell}^2 \;\lesssim\; \big(1 + C_3 \Lambda_\ell \Lambda_c^4 \Theta\big) \, \|v\|_{1,k}^2.$$

□

It is convenient to introduce the orthogonal projections onto the local spaces $P_j^\ell : V^h \to V_{\Omega_j^\ell}$, with respect to the inner product $(\cdot,\cdot)_{1,k}$, by requiring that, for each $j = 1, \dots, Q$,

$$(P_j^\ell v, w)_{1,k,\Omega_j^\ell} = (v, E_j^\ell w)_{1,k} \qquad \text{for all } w \in V_{\Omega_j^\ell},\ v \in V^h. \tag{5.3.35}$$

Similarly we define the orthogonal projection onto the coarse space $P_0 : V^h \to V_0$ by

$$(P_0 v, w)_{1,k} = (v, w)_{1,k} = (v, E_0 w)_{1,k}, \qquad \text{for all } w \in V_0. \tag{5.3.36}$$

Using these we define the operator $P : V^h \to V^h$ by

$$P := E_0 P_0 + \sum_{j=1}^{Q} E_j^\ell \, P_j^\ell. \tag{5.3.37}$$

The operator P is an easier-to-analyse proxy for the operator T representing the preconditioned matrix (see (5.2.27) and Proposition 5.2.12): T is defined analogously but with the *indefinite* bilinear form $b(\cdot,\cdot)$ on the left-hand side of (5.3.35)–(5.3.36) instead of the $(\cdot,\cdot)_{1,k}$ inner product. In the next result we estimate the field of values of P; this lifts to an estimate of the field of values of T (Lemma 5.4.1), which in turn yields our main GMRES convergence theorem (Theorem 5.3.1).

Proposition 5.3.14 (Energy coercivity of P)**.** *Under the assumptions of Lemma 5.3.9, any $v \in V^h$ satisfies*

$$\|v\|_{1,k}^2 \lesssim \big(1 + C_3\,\Lambda_\ell\,\Lambda_c^4\,\Theta\big)\,(Pv, v)_{1,k}. \tag{5.3.38}$$

Proof. By Lemma 5.3.13(i) we have the stable splitting $v = z_0 + \sum_{j=1}^{Q} E_j^\ell z_j^\ell$. Therefore,

$$\begin{aligned}
\|v\|_{1,k}^2 &= (v,\, z_0)_{1,k} + \sum_{j=1}^{Q} (v,\, E_j^\ell z_j^\ell)_{1,k} \\
&= (P_0 v,\, z_0)_{1,k} + \sum_{j=1}^{Q} \big(P_j^\ell v,\, z_j^\ell\big)_{1,k,\Omega_j^\ell} \qquad \text{(by the defining relations of } P_0 \text{ and } P_j^\ell) \\
&\le \|P_0 v\|_{1,k}\,\|z_0\|_{1,k} \;+\; \sum_{j=1}^{Q} \|P_j^\ell v\|_{1,k,\Omega_j^\ell}\,\|z_j^\ell\|_{1,k,\Omega_j^\ell} \quad \text{(Cauchy–Schwarz)} \\
&\le \Big(\|P_0 v\|_{1,k}^2 + \sum_{j=1}^{Q} \|P_j^\ell v\|_{1,k,\Omega_j^\ell}^2\Big)^{\frac12} \Big(\|z_0\|_{1,k}^2 + \sum_{j=1}^{Q} \|z_j^\ell\|_{1,k,\Omega_j^\ell}^2\Big)^{\frac12}.
\end{aligned}$$

By Lemma 5.3.13(ii),

$$\|z_0\|_{1,k}^2 + \sum_{j=1}^{Q} \|z_j^\ell\|_{1,k,\Omega_j^\ell}^2 \lesssim \big(1 + C_3\,\Lambda_\ell\,\Lambda_c^4\,\Theta\big)\,\|v\|_{1,k}^2.$$

For the first factor, take $w = P_0 v$ in (5.3.36) and $w = P_j^\ell v$ in (5.3.35) to get

$$\|P_0 v\|_{1,k}^2 = (v, E_0 P_0 v)_{1,k}, \qquad \|P_j^\ell v\|_{1,k,\Omega_j^\ell}^2 = (v, E_j^\ell P_j^\ell v)_{1,k}.$$

Hence

$$\|P_0 v\|_{1,k}^2 + \sum_{j=1}^{Q} \|P_j^\ell v\|_{1,k,\Omega_j^\ell}^2 = \big(E_0 P_0 v + \sum_{j=1}^{Q} E_j^\ell P_j^\ell v,\, v\big)_{1,k} = (Pv, v)_{1,k}.$$

Combining the two bounds and dividing by $\|v\|_{1,k}$ yields (5.3.38). □

Remark 5.3.15 (Interpretation and role in the two-level framework)**.**

- What (5.3.38) says. *The operator P is the SPD "proxy" additive Schwarz operator (built with the $(\cdot,\cdot)_{1,k}$ inner product). Inequality (5.3.38) is a* coercivity *(or energy lower bound): the quadratic form $(Pv, v)_{1,k}$ controls the full energy $\|v\|_{1,k}^2$ up to the factor $\big(1+C_3\Lambda_\ell\Lambda_c^4\Theta\big)$.*
- Why it holds. *The proof uses results from Lemma 5.3.11 and 5.3.13:* (a) *the coarse space gives a good* global *approximation (z_0), and* (b) *the remainder splits into* local *pieces with stable energy. Because P_0 and P_j^ℓ are the $(1,k)$-orthogonal projections, $(Pv, v)_{1,k}$ is exactly the squared norm of the projected pieces.*
- Dependence on parameters. *The only "inflation" in the constant comes from overlap multiplicities $(\Lambda_\ell, \Lambda_c)$ and the spectral threshold $\Theta = 1/\tau$: better overlap design and a richer coarse space (larger τ) tighten the bound.*

5.3.2 Existence and Stability of T_j^ℓ, $j = 1, \ldots, Q$

The next lemma shows that T_j^ℓ is well defined for $j = 1, \ldots, Q$ if the subdomains Ω_j^ℓ are small enough with respect to k.

Lemma 5.3.16. *If $H_\ell k < \sqrt{2}$, then, for each $j = 1, \ldots, Q$, $b_{\Omega_j^\ell}(\cdot, \cdot)$ is positive definite and the operators T_j^ℓ are well defined.*

Proof. Using the definition of $b_{\Omega_j^\ell}(\cdot, \cdot)$, Friedrichs' inequality and the assumption that $a_{\min} = 1 = n_{\max}$, we obtain

$$b_{\Omega_j^\ell}(u,u) = a_{\Omega_j^\ell}(u,u) - k^2(nu,u)_{\Omega_j^\ell} \geq \frac{2}{(H_\ell)^2}\|u\|^2_{\Omega_j^\ell} - k^2\|u\|^2_{\Omega_j^\ell} = \frac{2-(H_\ell)^2k^2}{(H_\ell)^2}\|u\|^2_{\Omega_j^\ell},$$

and hence $b_{\Omega_j^\ell}(\cdot, \cdot)$ is positive definite. The system (5.2.25) is then positive definite and the result follows. □

Remark 5.3.17. *Whilst this is a sufficient condition for T_j^ℓ to be well defined, it is not necessary. In general, the T_j^ℓ will be well-defined when the linear system corresponding to (5.2.25) is non-singular. However, it seems hard to guarantee good stability properties for T_j^ℓ unless the above condition is satisfied.*

Before we proceed, we state the following well-known boundedness property of $b_{\Omega'}$, which is a simple consequence of the Cauchy–Schwarz inequality.

$$|b_{\Omega'}(u,v)| \leq \|u\|_{1,k,\Omega'}\|v\|_{1,k,\Omega'}. \tag{5.3.39}$$

Lemma 5.3.18 (Stability of T_j^ℓ, $j = 1, \ldots, Q$)**.** *Suppose that $\sqrt{2}H_\ell k \leq 1$. Then*

$$\|T_j^\ell v\|_{1,k,\Omega_j^\ell} \leq 2\big\|v|_{\Omega_j^\ell}\big\|_{1,k,\Omega_j^\ell} \quad \text{for all} \quad v \in V^h.$$

Proof. Using (5.2.25), (5.3.39) and the Friedrichs' inequality (5.2.8) we obtain

$$\begin{aligned}
\|T_j^\ell v\|^2_{1,k,\Omega_j^\ell} &= b_{\Omega_j^\ell}(T_j^\ell v, T_j^\ell v) + 2k^2\|\sqrt{n}T_j^\ell v\|^2_{\Omega_i^\ell} = b_{\Omega_i^\ell}\big(v|_{\Omega_j^\ell}, T_j^\ell v\big) + 2k^2\|\sqrt{n}T_j^\ell v\|^2_{\Omega_j^\ell} \\
&\leq \big\|v|_{\Omega_j^\ell}\big\|_{1,k,\Omega_j^\ell}\|T_j^\ell v\|_{1,k,\Omega_j^\ell} + 2k^2\|\sqrt{n}\,T_j^\ell v\|^2_{\Omega_j^\ell} \\
&\leq \big\|v|_{\Omega_j^\ell}\big\|_{1,k,\Omega_j^\ell}\|T_j^\ell v\|_{1,k,\Omega_j^\ell} + k^2(H_\ell)^2\|T_j^\ell v\|^2_{1,k,\Omega_j^\ell},
\end{aligned}$$

which implies

$$\big(1 - k^2(H_\ell)^2\big)\|T_j^\ell v\|_{1,k,\Omega_j^\ell} \leq \big\|v|_{\Omega_j^\ell}\big\|_{1,k,\Omega_j^\ell}.$$

Using the condition $H_\ell k \leq 1/\sqrt{2}$, we can deduce the required bound. □

5.3.3 Existence and Stability of T_0

To ensure that T_0 is well defined, a condition on Θ (defined in (5.3.1)) is required.

Lemma 5.3.19 (T_0 is well-defined)**.** *Suppose that Θ is chosen small enough so that*

$$\sqrt{2C_3}\,k\Lambda_c^2\Theta^{\frac{1}{2}}(1 + C_{\text{stab}}) < 1. \tag{5.3.40}$$

Then there exists $h_1 > 0$ such that, for all $h \in (0, h_1)$, the operator T_0 is well defined.

Proof. Assume (for a contradiction) that there exists a $w_0 \in V_0 \backslash \{0\} \subseteq V^h$, such that

$$b(w_0, z) = 0 \qquad \text{for all } z \in V_0. \tag{5.3.41}$$

Then let $w \in H_0^1(\Omega)$ be the solution of

$$b(w, v) = (w_0, v) \qquad \text{for all } v \in H_0^1(\Omega),$$

which exists and is unique by (2.2.9). Then Lemma 5.2.3 implies that, for all $\varepsilon > 0$, there exists an $h_1 = h_1(\varepsilon) > 0$ such that, for all $h \in (0, h_1)$, there is a solution $w_h \in V^h$ of

$$b(w_h, v) = (w_0, v) \qquad \text{for all } v \in V^h.$$

This holds, in particular, for $v = w_0$. This, together with (5.3.41) and (5.3.39), yields that, for all $z \in V_0$,

$$\|w_0\|^2 = b(w_h, w_0) = b(w_h, w_0) - b(z, w_0) = b(w_h - z, w_0) \leq \|w_h - z\|_{1,k} \|w_0\|_{1,k}.$$

As this is true for all $z \in V_0$, we obtain, by Lemma 5.3.11,

$$\|w_0\|^2 \leq \|w_0\|_{1,k} \inf_{z \in V_0} \|w_h - z\|_{1,k} \leq \|w_0\|_{1,k} \sqrt{C_3} \Lambda_c^2 \Theta^{\frac{1}{2}} \|w_h\|_{1,k}. \tag{5.3.42}$$

Equation (5.3.41) with $z = w_0$ implies that

$$0 = b(w_0, w_0) = \|w_0\|_{1,k}^2 - 2k^2 \|w_0\|^2,$$

which, together with (5.3.42), yields

$$\|w_0\|^2 \leq \sqrt{2C_3}\, k \|w_0\| \Lambda_c^2 \Theta^{\frac{1}{2}} \|w_h\|_{1,k}. \tag{5.3.43}$$

Then we use the triangle inequality, (2.2.9) and (5.2.6) to obtain

$$\|w_h\|_{1,k} \leq \|w\|_{1,k} + \|w - w_h\|_{1,k} \leq (C_{\text{stab}} + \varepsilon) \|w_0\| \leq (C_{\text{stab}} + 1) \|w_0\| \tag{5.3.44}$$

where $\varepsilon := \beta/(1 - 2\beta^2)$ is chosen small enough. With (5.3.44) used in (5.3.43), we arrive at

$$\|w_0\|^2 \leq \sqrt{2C_3}\, k \Lambda_c^2 \Theta^{\frac{1}{2}} (1 + C_{\text{stab}}) \|w_0\|^2.$$

Together with the condition (5.3.40), this leads to a contradiction. □

The next lemma obtains a condition for the stability of T_0.

Lemma 5.3.20 (Stability of T_0)**.** *Assume that* (5.3.40) *is satisfied. Then there exists* $h_1 > 0$ *such that, for* $h \in (0, h_1)$*,*

$$\|T_0 u - u\| \leq \sqrt{C_3} \Lambda_c^2 \Theta^{\frac{1}{2}} (1 + C_{\text{stab}}) \|T_0 u - u\|_{1,k} \qquad \textit{for all } u \in V^h. \tag{5.3.45}$$

Suppose, in addition that,

$$2k \sqrt{C_3} \Lambda_c^2 \Theta^{\frac{1}{2}} (1 + C_{\text{stab}}) \leq \frac{1}{2}; \tag{5.3.46}$$

then

$$\|u - T_0 u\|_{1,k} \leq 2 \|u\|_{1,k} \qquad \textit{for all } u \in V^h. \tag{5.3.47}$$

Proof. Under condition (5.3.40), Lemma 5.3.19 ensures the existence of $h_1 > 0$ such that, for $h \in (0, h_1)$, the operator $T_0 : V^h \to V_0$ is well defined. Let $u \in V^h$ and consider now the following auxiliary problem:

$$\text{Find } w_h \in V^h \text{ such that } \quad b(w_h, v) = \big(T_0 u - u, v\big) \quad \text{for all } v \in V^h. \tag{5.3.48}$$

The proof of Lemma 5.3.19 shows that there is a unique solution w_h of (5.3.48) for $h \in (0, h_1)$ and that

$$\|w_h\|_{1,k} \le (1 + C_{\text{stab}})\|T_0 u - u\|. \tag{5.3.49}$$

Now, by the definition (5.3.36) of T_0, we have $b(T_0 u - u, z) = 0$ for all $z \in V_0$. We can use $v = T_0 u - u \in V^h$ in (5.3.48) and (5.3.39) to obtain, for every $z \in V_0$,

$$\begin{aligned}\|T_0 u - u\|^2 &= b(w_h, T_0 u - u) = b(w_h, T_0 u - u) - b(z, T_0 u - u)\\ &= b(w_h - z, T_0 u - u) \le \|w_h - z\|_{1,k}\, \|T_0 u - u\|_{1,k}.\end{aligned}$$

Combining this with Lemma 5.3.11 we obtain

$$\|T_0 u - u\|^2 \le \|T_0 u - u\|_{1,k} \inf_{z \in V_0} \|w_h - z\|_{1,k} \le \|T_0 u - u\|_{1,k} \sqrt{C_3}\Lambda_c^2 \Theta^{\frac{1}{2}} \|w_h\|_{1,k}.$$

which, together with (5.3.49), proves (5.3.45).

We now focus on the proof of (5.3.47). By the definition of P_0 in (5.3.36), we have

$$(P_0 u - u, v)_{1,k} = 0, \qquad \text{for all} \quad u \in V^h, v \in V_0.$$

Using $v = T_0 u \in V_0$ in this relation we obtain $(T_0 u, P_0 u - u)_{1,k} = 0$. Since $P_0 u - T_0 u \in V_0$, the definition of T_0 yields

$$b(u - T_0 u, P_0 u - T_0 u) = 0. \tag{5.3.50}$$

Now we can use (5.3.50), the link between the bilinear forms $(\cdot,\cdot)_{1,k}$ and b, and the Cauchy–Schwarz inequality to obtain

$$\begin{aligned}\|u - T_0 u\|_{1,k}^2 &= (u - T_0 u, u - T_0 u)_{1,k} = b(u - T_0 u, u - T_0 u) + 2k^2(u - T_0 u, u - T_0 u)\\ &= b(u - T_0 u, u - T_0 u) - b(u - T_0 u, P_0 u - T_0 u) + 2k^2(u - T_0 u, u - T_0 u)\\ &= b(u - T_0 u, u - P_0 u) + 2k^2(u - T_0 u, u - T_0 u)\\ &= (u - T_0 u, u - P_0 u)_{1,k} - 2k^2(u - T_0 u, u - P_0 u) + 2k^2(u - T_0 u, u - T_0 u)\\ &= (u - T_0 u, u - P_0 u)_{1,k} + 2k^2(u - T_0 u, P_0 u - T_0 u)\\ &\le \|u - T_0 u\|_{1,k}\|u - P_0 u\|_{1,k} + 2k^2\|u - T_0 u\|\|T_0 u - P_0 u\|\\ &\le \|u - T_0 u\|_{1,k}\|u - P_0 u\|_{1,k} + 2k\|u - T_0 u\|\|T_0 u - P_0 u\|_{1,k}\\ &= \|u - T_0 u\|_{1,k}\|u - P_0 u\|_{1,k} + 2k\|u - T_0 u\|\|P_0(T_0 u - u)\|_{1,k}\\ &\le \|u - T_0 u\|_{1,k}\|u\|_{1,k} + 2k\|u - T_0 u\|\|T_0 u - u\|_{1,k},\end{aligned}$$

where in the last step we used the fact that P_0 is the orthogonal projection onto V_0 with respect to $(\cdot,\cdot)_{1,k}$. Dividing both sides by $\|u - T_0 u\|_{1,k}$ and then using (5.3.45) and the assumption (5.3.46) we arrive at

$$\begin{aligned}\|u - T_0 u\|_{1,k} \le \|u\|_{1,k} + 2k\|u - T_0 u\| &\le \|u\|_{1,k} + 2k\sqrt{C_3}\Lambda_c^2\Theta^{\frac{1}{2}}(1 + C_{\text{stab}})\|u - T_0 u\|_{1,k}\\ &\le \|u\|_{1,k} + \frac{1}{2}\|u - T_0 u\|_{1,k},\end{aligned}$$

which proves (5.3.47). □

5.4 Proof of the Main Result

Strategy at a glance. The key steps in proving the theoretical results consist in a comparison of the easy-to-analyse SPD operator P with the true operator T:

(i) relate $(Pu,u)_{1,k}$ to $(Tu,u)_{1,k}$ up to a remainder R;

(ii) bound R using the stability of local/coarse solves and the overlap estimates;

(iii) combine with the coercivity of P (Proposition 5.3.14) to obtain a field-of-values lower bound for T; and

(iv) bound $\|Tu\|_{1,k}$ from above. Elman's theory then yields GMRES convergence.

These results are stated and shown in the following lemma.

Lemma 5.4.1. *Let the assumptions in Theorem 5.3.1 be satisfied and set*

$$c_1 := \frac{1-s}{1+C_3\Lambda_\ell\Lambda_c^4\Theta}, \qquad c_2 := 18+8{\Lambda_\ell}^2.$$

Then, for all $u \in V^h$,

$$c_1\|u\|_{1,k}^2 \le (Tu,u)_{1,k}, \tag{5.4.1}$$

and

$$\|Tu\|_{1,k}^2 \le c_2\|u\|_{1,k}^2. \tag{5.4.2}$$

Proof. Let $u \in V^h$. We proceed in several steps.

Step 1 **(Relation between $(Tu,u)_{1,k}$ and $(Pu,u)_{1,k}$)** . We start by employing (5.3.37) and the definition of b to obtain

$$\begin{aligned}(Pu,u)_{1,k} = (u,Pu)_{1,k} &= \left(u, E_0P_0u + \sum_{j=1}^{Q} E_j^\ell P_j^\ell u\right)_{1,k} \\ &= b(u,E_0P_0u) + 2k^2(u,E_0P_0u) + \sum_{j=1}^{Q}\left[b(u,E_j^\ell P_j^\ell u) + 2k^2(u,E_j^\ell P_j^\ell u)\right].\end{aligned}$$

Then, using (5.2.25) and (5.2.26), we obtain

$$\begin{aligned}(Pu,u)_{1,k} &= b(T_0u,P_0u) + 2k^2(u,E_0P_0u) + \sum_{j=1}^{Q}\left[b_{\Omega_j^\ell}(T_j^\ell u,P_j^\ell u) + 2k^2(u,E_j^\ell P_j^\ell u)\right] \\ &= (T_0u,P_0u)_{1,k} - 2k^2(T_0u,P_0u) + 2k^2(u,E_0P_0u) \\ &\quad + \sum_{j=1}^{Q}\left[(T_j^\ell u,P_j^\ell u)_{1,k,\Omega_j^\ell} - 2k^2(T_j^\ell u,P_j^\ell u)_{\Omega_j^\ell} + 2k^2(u,E_j^\ell P_j^\ell u)\right].\end{aligned}$$

Now using the definitions (5.3.35) and (5.3.36) of P_j^ℓ and P_0, we have

$$\begin{aligned}(Pu,u)_{1,k} &= (E_0T_0u,u)_{1,k} - 2k^2(E_0T_0u,E_0P_0u) + 2k^2(u,E_0P_0u) \\ &\quad + \sum_{j=1}^{Q}\left[(E_j^\ell T_j^\ell u,u)_{1,k} - 2k^2(E_j^\ell T_j^\ell u,E_j^\ell P_j^\ell u) + 2k^2(u,E_j^\ell P_j^\ell u)\right].\end{aligned}$$

We now rearrange the right-hand side to obtain:

$$\begin{aligned}(Pu,u)_{1,k} &= (E_0T_0u,u)_{1,k} + \sum_{j=1}^{Q}(E_j^\ell T_j^\ell u,u)_{1,k} \\ &\quad - 2k^2\left((E_0T_0u-u,E_0P_0u) + \sum_{j=1}^{Q}(E_j^\ell T_j^\ell u-u,E_j^\ell P_j^\ell u)\right) \\ &= (Tu,u)_{1,k} - R,\end{aligned}$$

where we used the definition (5.2.27) of T and we have introduced the remainder:

$$R := 2k^2\left((E_0T_0u-u,E_0P_0u) + \sum_{j=1}^{Q}(E_j^\ell T_j^\ell u-u,E_j^\ell P_j^\ell u)\right), \tag{5.4.3}$$

Step 2. **(Bounding the first term in R.)** Recalling that $E_0w_0 = w_0$ for all $w_0 \in V_0$, and noting that the assumption in (5.3.2) implies that (5.3.46) is satisfied, we can use (5.3.45), (5.3.47) and the fact that P_0 is the orthogonal projection with respect to $(\cdot,\cdot)_{1,k}$, to obtain

$$\begin{aligned}k^2(E_0T_0u-u,E_0P_0u) &\le k^2\|E_0T_0u-u\|\|E_0P_0u\| = k^2\|T_0u-u\|\|P_0u\| \\ &\le k\|T_0u-u\|\|P_0u\|_{1,k} \le \sqrt{C_3}k\Lambda_c^2\Theta^{\frac12}(1+C_{\mathrm{stab}})\|T_0u-u\|_{1,k}\|P_0u\|_{1,k} \\ &\le 2k\sqrt{C_3}\Lambda_c^2\Theta^{\frac12}(1+C_{\mathrm{stab}})\|u\|_{1,k}\|P_0u\|_{1,k} \\ &\le 2k\sqrt{C_3}\Lambda_c^2\Theta^{\frac12}(1+C_{\mathrm{stab}})\|u\|_{1,k}^2.\end{aligned} \tag{5.4.4}$$

Step 3. **(Bounding the second term in R.)** Using Friedrichs' inequality, we obtain, for each $j = 1,\dots,Q$,

$$\begin{aligned}k^2(E_j^\ell T_j^\ell u-u,E_j^\ell P_j^\ell u) &= k^2\big(T_j^\ell u-u|_{\Omega_j^\ell},P_j^\ell u\big)_{\Omega_j^\ell} \le k^2\big\|T_j^\ell u-u|_{\Omega_j^\ell}\big\|_{\Omega_j^\ell}\|P_j^\ell u\|_{\Omega_j^\ell} \\ &\le k\big\|T_j^\ell u-u|_{\Omega_j^\ell}\big\|_{1,k,\Omega_j^\ell}\|P_j^\ell u\|_{\Omega_j^\ell} \\ &\le k\big\|T_j^\ell u-u|_{\Omega_j^\ell}\big\|_{1,k,\Omega_j^\ell}\frac{H_\ell}{\sqrt{2}}\|P_j^\ell u\|_{1,k,\Omega_j^\ell} \\ &\le \frac{kH_\ell}{\sqrt{2}}\Big(\|T_j^\ell u\|_{1,k,\Omega_j^\ell} + \big\|u|_{\Omega_j^\ell}\big\|_{1,k,\Omega_j^\ell}\Big)\|P_j^\ell u\|_{1,k,\Omega_j^\ell}.\end{aligned} \tag{5.4.5}$$

Now, note that the hypothesis (5.3.2) implies that $kH_\ell < 1/(3\Lambda_\ell) \le 1/3 < \sqrt{2}/2$. Hence we can apply Lemma 5.3.18 to obtain, from (5.4.5),

$$k^2(E_j^\ell T_j^\ell u-u,E_j^\ell P_j^\ell u) \le \frac{3kH_\ell}{\sqrt{2}}\big\|u|_{\Omega_j^\ell}\big\|_{1,k,\Omega_j^\ell}\|P_j^\ell u\|_{1,k,\Omega_j^\ell} < 3kH_\ell\big\|u|_{\Omega_j^\ell}\big\|_{1,k,\Omega_j^\ell}\|P_j^\ell u\|_{1,k,\Omega_j^\ell}.$$

Taking the sum over $j = 1,\dots,Q$, applying the Cauchy–Schwarz inequality and the overlap property (5.2.22), we arrive at

$$\begin{aligned}k^2\sum_{j=1}^{Q}(E_j^\ell T_j^\ell u-u,E_j^\ell P_j^\ell u) &\le 3kH_\ell\sum_{j=1}^{Q}\big\|u|_{\Omega_j^\ell}\big\|_{1,k,\Omega_j^\ell}\|P_j^\ell u\|_{1,k,\Omega_j^\ell} \\ &\le 3kH_\ell\left(\sum_{j=1}^{Q}\big\|u|_{\Omega_j^\ell}\big\|_{1,k,\Omega_j^\ell}^2\right)^{\frac12}\left(\sum_{j=1}^{Q}\|P_j^\ell u\|_{1,k,\Omega_j^\ell}^2\right)^{\frac12} \\ &\le 3kH_\ell\Lambda_\ell^{1/2}\|u\|_{1,k}\left(\sum_{j=1}^{Q}\|P_j^\ell u\|_{1,k,\Omega_j^\ell}^2\right)^{\frac12}.\end{aligned} \tag{5.4.6}$$

To estimate the sum on the right-hand side of (5.4.6), note that

$$\begin{aligned}\|E_j^\ell P_j^\ell u\|_{1,k}^2 = (E_j^\ell P_j^\ell u, E_j^\ell P_j^\ell u)_{1,k} &= (P_j^\ell u, P_j^\ell u)_{1,k,\Omega_j^\ell} \\ &= (P_j^\ell u, u)_{1,k,\Omega_j^\ell} = (E_j^\ell P_j^\ell u, u)_{1,k},\end{aligned}$$

and so, using a standard property of overlapping covers (5.2.23), we have

$$\begin{aligned}\left\| \sum_{j=1}^Q E_j^\ell P_j^\ell u \right\|_{1,k}^2 &\le \Lambda_\ell \sum_{j=1}^Q \|E_j^\ell P_j^\ell u\|_{1,k}^2 = \Lambda_\ell \sum_{j=1}^Q (E_j^\ell P_j^\ell u, u)_{1,k} \\ &= \Lambda_\ell \left(\sum_{j=1}^Q E_j^\ell P_j^\ell u,\, u \right)_{1,k} \le \Lambda_\ell \left\| \sum_{j=1}^Q E_j^\ell P_j^\ell u \right\|_{1,k} \|u\|_{1,k}.\end{aligned}$$

Hence

$$\left\| \sum_{j=1}^Q E_j^\ell P_j^\ell u \right\|_{1,k} \le \Lambda_\ell \|u\|_{1,k},$$

which then implies

$$\begin{aligned}\sum_{j=1}^Q \|P_j^\ell u\|_{1,k,\Omega_j^\ell}^2 = \sum_{j=1}^Q (E_j^\ell P_j^\ell u, u)_{1,k} &= \left(\sum_{j=1}^Q E_j^\ell P_j^\ell u,\, u \right)_{1,k} \\ &\le \left\| \sum_{j=1}^Q E_j^\ell P_j^\ell u \right\|_{1,k} \|u\|_{1,k} \le \Lambda_\ell \|u\|_{1,k}^2.\end{aligned}$$

This, together with (5.4.6), leads to

$$k^2 \sum_{j=1}^Q (E_j^\ell T_j^\ell u - u, E_j^\ell P_j^\ell u) \le 3k H_\ell \Lambda_\ell \|u\|_{1,k}^2. \tag{5.4.7}$$

Step 4 **(Obtaining (5.4.1).)** We can now combine (5.4.3), (5.4.4) and (5.4.7) to obtain

$$|R| \le 2\left(2k\sqrt{C_3}\Lambda_c^2 \Theta^{1/2}(1 + C_{\text{stab}}) + 3kH_\ell\Lambda_\ell\right)\|u\|_{1,k}^2 = \frac{s}{1 + C_3\Lambda_\ell\Lambda_c^4\Theta}\|u\|_{1,k}^2,$$

with s as defined in (5.3.2). Hence, using this and Proposition 5.3.14, we obtain

$$\begin{aligned}(Tu, u)_{1,k} \ge (Pu, u)_{1,k} + R &\ge (1 + C_3\Lambda_\ell\Lambda_c^4\Theta)^{-1}\|u\|_{1,k}^2 - |R| \\ &\ge (1 + C_3\Lambda_\ell\Lambda_c^4\Theta)^{-1}(1 - s)\|u\|_{1,k}^2,\end{aligned}$$

which yields the result.

Step 5. **(Obtaining (5.4.2).)** Since $E_0 T_0 u = T_0 u$, we have

$$\|Tu\|_{1,k}^2 = \left\| T_0 u + \sum_{j=1}^Q E_j^\ell T_j^\ell u \right\|_{1,k}^2 \le 2\|T_0 u\|_{1,k}^2 + 2\left\| \sum_{j=1}^Q E_j^\ell T_j^\ell u \right\|_{1,k}^2. \tag{5.4.8}$$

For the first term on the right-hand side of (5.4.8), we use Cauchy–Schwarz and (5.3.47) to obtain

$$\begin{aligned}\|T_0 u\|_{1,k}^2 = (T_0 u, T_0 u)_{1,k} &= (T_0 u - u, T_0 u)_{1,k} + (u, T_0 u)_{1,k} \\ &\le \|T_0 u - u\|_{1,k}\|T_0 u\|_{1,k} + \|u\|_{1,k}\|T_0 u\|_{1,k} \\ &\le 2\|u\|_{1,k}\|T_0 u\|_{1,k} + \|u\|_{1,k}\|T_0 u\|_{1,k} = 3\|u\|_{1,k}\|T_0 u\|_{1,k},\end{aligned}$$

and hence

$$\|T_0 u\|_{1,k} \leq 3\|u\|_{1,k}. \tag{5.4.9}$$

For the second term on the right-hand side of (5.4.8) we use again Lemma 5.3.18 and (5.2.23) to obtain

$$\left\| \sum_{j=1}^{Q} E_j^\ell T_j^\ell u \right\|_{1,k}^2 \leq \Lambda_\ell \sum_{j=1}^{Q} \|T_j^\ell u\|_{1,k,\Omega_j^\ell}^2 \leq \Lambda_\ell \sum_{j=1}^{Q} 4 \left\|u|_{\Omega_i}\right\|_{1,k,\Omega_j^\ell}^2 \leq 4\Lambda_\ell^2 \|u\|_{1,k}^2. \tag{5.4.10}$$

Combining (5.4.9) and (5.4.10) with (5.4.8) we arrive at

$$\|Tu\|_{1,k}^2 \leq 2 \times 9\|u\|_{1,k}^2 + 2 \times 4\Lambda_\ell^2 \|u\|_{1,k}^2 = (18 + 8\Lambda_\ell^2)\|u\|_{1,k},$$

which proves (5.4.2). □

We now complete the proof of Theorem 5.3.1.

Proof of Theorem 5.3.1. Using (5.4.1) with (5.2.28) we obtain

$$c_1 \|u\|_{1,k}^2 \leq (Tu, u)_{1,k} = \left\langle \mathbf{M}_{AS,2}^{-1} \mathbf{B}\mathbf{u}, \mathbf{u} \right\rangle_{\mathbf{D}_k},$$

which can be written as

$$c_1 \leq \frac{\left\langle \mathbf{M}_{AS,2}^{-1} \mathbf{B}\mathbf{u}, \mathbf{u} \right\rangle_{\mathbf{D}_k}}{\|\mathbf{u}\|_{\mathbf{D}_k}^2}.$$

Also, combining (5.4.2) with (5.2.28), we get

$$\|\mathbf{M}_{AS,2}^{-1} \mathbf{B}\mathbf{u}\|_{\mathbf{D}_k}^2 = \|Tu\|_{1,k}^2 \leq c_2 \|u\|_{1,k}^2 = c_2 \|\mathbf{u}\|_{\mathbf{D}_k}^2,$$

which implies

$$\|\mathbf{M}_{AS,2}^{-1} \mathbf{B}\|_{\mathbf{D}_k}^2 \leq c_2.$$

The result on GMRES convergence now follows from the Elman theory [35] with $\gamma = \frac{c_1}{c_2}$. □

Remark 5.4.2 (Takeaways before the numerical results)**.**

- *The lower bound* (5.4.1) *shows that the two-level operator T has its field of values bounded away from* 0 *by a constant proportional to* $(1-s)/(1+C_3\Lambda_\ell\Lambda_c^4\Theta)$*:* the coarse space quality (*via* Θ) *and* the overlaps $(\Lambda_c, \Lambda_\ell)$ *directly govern robustness.*
- *The upper bound* (5.4.2) *ensures no blow-up of the preconditioned norm, with a mild dependence on* Λ_ℓ*.*
- *Together these yield a k-robust GMRES rate as soon as* $kH_\ell \lesssim 1$ *and* $(1 + C_{\text{stab}})^2 k^2 \Theta \lesssim 1$ *(Corollary 5.3.2), highlighting the key design levers for H_k-GenEO:* (a) *pick* $H_\ell \sim k^{-1}$, (b) *include local modes until* $\tau = \min_i \lambda_{m_i+1}^i$ *reaches* $\tau \gtrsim (1 + C_{\text{stab}})^2 k^2$*.*

These predictions are borne out by the experiments in the next section.

5.5 Numerical Results

This section aims to assess the effectiveness of the two–level overlapping Schwarz preconditioner equipped with the H_k–GenEO coarse space, using the Δ_k–GenEO as a comparison. The experiments are designed to test wavenumber robustness, scalability with respect to the number of subdomains and sensitivity to the spectral threshold τ, which controls the dimension of the coarse space. Two representative test cases are being considered.

1. **Homogeneous case.** We will solve the Helmholtz equation on the unit square with homogeneous Dirichlet boundary conditions and a single interior point source. This controlled setting isolates frequency and decomposition effects.

2. **Heterogeneous case.** A layered medium on the same domain, alternating between $a_{\min} = 1$ and $a_{\max} > 1$.

For many of the numerical experiments in this section, the results are reported using different values of τ when comparing the H_k-GenEO and Δ_k-GenEO. This is an attempt to keep the coarse space a comparatively similar size in order to highlight the effectiveness of the H_k-GenEO method.

More extensive test cases are carried out in Chapter 6 in comparison with other spectral coarse spaces. The numerical tests in this section are solving the problem as defined in the theory, whereas the numerical tests in Chapter 6 are using variations of the Helmholtz problem in different settings.

5.5.1 Implementation Details

We refer the reader to Section 4.5.1 for a complete description of the implementation details. We are using the same methodology with H_k-GenEO used as the coarse space.

5.5.2 Homogeneous Problem

We begin by testing the H_k-GenEO method on a simple yet representative model: the 2D Helmholtz equation on the unit square $\Omega = (0, 1)^2$ with homogeneous Dirichlet boundary conditions on all sides. A point source is located in the centre of the domain at $(\frac{1}{2}, \frac{1}{2})$ and provides the forcing function f. The point source is numerically modelled by a Gaussian function: $f(x, y) = 10^4 \exp(-10^3[(x - \frac{1}{2})^2 + (y - \frac{1}{2})^2])$. A schematic of this model problem is found in Figure 4.1.

We discretise the domain using a uniform Cartesian grid with spacing h and use alternating diagonals to create a conforming simplicial mesh. Unless otherwise stated, the coarse space is computed on the same decomposition as the local subdomains, i.e., $H_\ell = H_c$, and the overlap is one layer of elements. The main goal is to assess the robustness of the H_k-GenEO preconditioner as the wavenumber k increases and as the number of subdomains N varies. The coarse space is constructed by including all local eigenfunctions with eigenvalue below a fixed threshold τ, with three values tested: $\tau = 0.2.0.4, 0.6$. Here we have used the definition $\tau = \min_{i=1}^{N} \lambda_{m_i+1}^i$, where $\lambda_{m_i+1}^i$ is the first unused eigenvalue in subdomain Ω_i.

				H_k-GenEO					Δ_k-GenEO		
N	k	h^{-1}	Dim(Fine)	Max.Num. Neg.	$\lambda_{\min}$	Dim(Loc)	Dim(CS)	It. Count	Dim(Loc)	Dim(CS)	It. Count
16	20	240	58081	4.0	-0.324293	18	240	16	29	464	16
	40	480	231361	13.0	-0.703879	49	628	20	65	1040	24
	60	720	519841	22.0	-0.849475	86	1160	17	110	1768	21
	80	960	923521	41.0	-0.910659	132	1876	27	168	2692	38
	100	1200	1442401	58.0	-0.941228	190	2660	36	230	3688	52
36	20	240	58081	3.0	-0.178013	12	344	19	20	720	16
	40	480	231361	6.0	-0.484958	28	844	31	43	1540	45
	60	720	519841	13.0	-0.703879	49	1528	18	71	2544	25
	80	960	923521	20.0	-0.815187	72	2312	23	103	3696	42
	100	1200	1442401	28.0	-0.875285	99	3260	51	137	4936	78
64	20	240	58081	1.0	-0.125924	8	448	19	14	920	22
	40	480	231361	4.0	-0.324293	18	1056	24	34	2144	29
	60	720	519841	8.0	-0.553713	34	1924	18	51	3268	32
	80	960	923521	13.0	-0.703879	49	2820	23	74	4704	51
	100	1200	1442401	17.0	-0.793739	65	3840	21	98	6272	38
100	20	240	58081	1.0	-0.100482	8	684	18	12	1180	21
	40	480	231361	3.0	-0.230414	16	1440	24	26	2584	44
	60	720	519841	4.0	-0.423257	24	2200	18	40	4004	47
	80	960	923521	8.0	-0.590168	34	3224	24	58	5804	55
	100	1200	1442401	13.0	-0.703879	49	4504	24	75	7520	44
144	20	240	58081	1.0	-0.0852859	5	672	21	11	1540	18
	40	480	231361	3.0	-0.178013	12	1544	30	22	3168	34
	60	720	519841	4.0	-0.324293	18	2448	23	35	5040	32
	80	960	923521	6.0	-0.484958	28	3700	45	47	6736	101
	100	1200	1442401	8.0	-0.612531	39	5232	25	60	8688	55

Table 5.2: Results for the homogeneous problem showing the dimension of the fine mesh (dim(Fine)), the maximum number of negative eigenvalues per subdomain (Max. Num. Neg), the minimum eigenvalue ($\lambda_{\min}$), the maximum number of modes per domain used in the coarse space (dim(Loc)), the dimension of the coarse space (dim(CS)), and the GMRES iteration count for a varying k. The threshold value is set to $\tau = 0.4$ for H_k-GenEO, with $\tau = 0.7$ for the Δ-GenEO.

				H_k-GenEO					Δ_k-GenEO		
N	k	h^{-1}	Dim(Fine)	Max.Num. Neg.	$\lambda_{\min}$	Dim(Loc)	Dim(CS)	It. Count	Dim(Loc)	Dim(CS)	It. Count
16	20	240	58081	4.0	-0.324293	12	144	21	21	340	21
	40	480	231361	13.0	-0.703879	26	356	27	49	784	29
	60	720	519841	22.0	-0.849475	49	668	23	82	1312	31
	80	960	923521	41.0	-0.910659	81	1108	65	121	1936	75
	100	1200	1442401	58.0	-0.941228	112	1612	52	167	2672	68
36	20	240	58081	3.0	-0.178013	8	220	22	14	516	22
	40	480	231361	6.0	-0.484958	14	456	77	32	1168	56
	60	720	519841	13.0	-0.703879	26	844	27	54	1932	34
	80	960	923521	20.0	-0.815187	41	1288	44	76	2740	68
	100	1200	1442401	28.0	-0.875285	58	1876	68	102	3684	90
64	20	240	58081	1.0	-0.125924	4	224	32	12	736	23
	40	480	231361	4.0	-0.324293	12	672	32	25	1604	56
	60	720	519841	8.0	-0.553713	19	1060	32	38	2460	72
	80	960	923521	13.0	-0.703879	26	1540	41	56	3584	76
	100	1200	1442401	17.0	-0.793739	39	2276	33	72	4616	88
100	20	240	58081	1.0	-0.100482	4	324	31	10	980	21
	40	480	231361	3.0	-0.230414	8	684	60	19	1936	87
	60	720	519841	4.0	-0.423257	13	1144	37	30	3020	104
	80	960	923521	8.0	-0.590168	19	1744	56	43	4324	144
	100	1200	1442401	13.0	-0.703879	26	2444	37	57	5740	80
144	20	240	58081	1.0	-0.0852859	4	484	27	7	1056	19
	40	480	231361	3.0	-0.178013	8	1012	37	16	2256	64
	60	720	519841	4.0	-0.324293	12	1584	29	26	3796	77
	80	960	923521	6.0	-0.484958	14	1920	163	36	5140	162
	100	1200	1442401	8.0	-0.612531	19	2632	56	46	6576	197

Table 5.3: Results for the homogeneous problem showing the dimension of the fine mesh (dim(Fine)), the maximum number of negative eigenvalues per subdomain (Max. Num. Neg), the minimum eigenvalue ($\lambda_{\min}$), the maximum number of modes per domain used in the coarse space (dim(Loc)), the dimension of the coarse space ((dim(CS)), and the GMRES iteration count for a varying k. The threshold value is set to $\tau = 0.2$ for the H_k-GenEO, with $\tau = 0.4$ for the Δ_k-GenEO.

				H_k-GenEO					Δ_k-GenEO		
N	k	h^{-1}	Dim(Fine)	Max.Num. Neg.	$\lambda_{\min}$	Dim(Loc)	Dim(CS)	It. Count	Dim(Loc)	Dim(CS)	It. Count
16	20	240	58081	4.0	-0.324293	32	392	14	29	464	16
	40	480	231361	13.0	-0.703879	81	1068	17	65	1040	24
	60	720	519841	22.0	-0.849475	145	1984	15	110	1768	21
	80	960	923521	41.0	-0.910659	225	3152	25	168	2692	38
	100	1200	1442401	58.0	-0.941228	321	4572	29	230	3688	52
36	20	240	58081	3.0	-0.178013	20	600	15	20	720	16
	40	480	231361	6.0	-0.484958	45	1404	25	43	1540	45
	60	720	519841	13.0	-0.703879	81	2576	15	71	2544	25
	80	960	923521	20.0	-0.815187	121	3908	18	103	3696	42
	100	1200	1442401	28.0	-0.875285	171	5584	56	137	4936	78
64	20	240	58081	1.0	-0.125924	13	740	15	14	920	22
	40	480	231361	4.0	-0.324293	32	1800	18	34	2144	29
	60	720	519841	8.0	-0.553713	56	3232	15	51	3268	32
	80	960	923521	13.0	-0.703879	81	4732	17	74	4704	51
	100	1200	1442401	17.0	-0.793739	111	6508	15	98	6272	38
100	20	240	58081	1.0	-0.100482	12	1044	17	12	1180	21
	40	480	231361	3.0	-0.230414	26	2360	20	26	2584	44
	60	720	519841	4.0	-0.423257	41	3748	16	40	4004	47
	80	960	923521	8.0	-0.590168	58	5360	19	58	5804	55
	100	1200	1442401	13.0	-0.703879	81	7536	15	75	7520	44
144	20	240	58081	1.0	-0.0852859	8	1096	17	11	1540	18
	40	480	231361	3.0	-0.178013	20	2640	20	22	3168	34
	60	720	519841	4.0	-0.324293	32	4232	16	35	5040	32
	80	960	923521	6.0	-0.484958	45	6048	33	47	6736	101
	100	1200	1442401	8.0	-0.612531	63	8500	16	60	8688	55

Table 5.4: Results for the homogeneous problem showing the dimension of the fine mesh (dim(Fine)), the maximum number of negative eigenvalues per subdomain (Max. Num. Neg), the minimum eigenvalue ($\lambda_{\min}$), the maximum number of modes per domain used in the coarse space (dim(Loc)), the dimension of the coarse space ((dim(CS)), and the GMRES iteration count for a varying k. The threshold value is set to $\tau = 0.6$ for H_k-GenEO, and $\tau = 0.7$ for Δ_k-GenEO.

5.5.3 Heterogenous Problem

We now consider the same experiments as in the homogeneous case, but with the diffusion coefficient $A(\mathbf{x})$ giving a heterogeneous domain. This aims to study the robustness of the method with respect to medium heterogeneity. For this purpose, we consider a layered medium within the same unit square domain. We use an alternating layer configuration in which the heterogeneity is introduced through the material parameter $a(\mathbf{x})$ with $a_{\min}(\boldsymbol{x}) = 1$ and introduce heterogeneity through altering $a_{\max}(\mathbf{x})$, as shown in Figure 4.6. The heterogeneous coefficient is given by

$$A(\mathbf{x}) = a(\mathbf{x})I.$$

Choice of the spectral threshold τ. In all experiments the parameter τ controls the richness of the coarse space, since all local eigenfunctions with eigenvalue $\lambda \leq \tau$ are retained. Accordingly, increasing τ improves robustness but also increases cost. The setup phase becomes more expensive because more local eigenvectors must be computed and assembled into the coarse basis, while the solve phase becomes more expensive because the coarse matrix is larger, its factorisation is costlier, and each GMRES iteration requires a more expensive coarse correction. For this reason we report results for $\tau \in \{0.2, 0.4, 0.6\}$: $\tau = 0.2$ represents an aggressive truncation with a cheap but relatively small coarse space, $\tau = 0.6$ represents a robustness-oriented choice with a noticeably larger coarse problem, and $\tau = 0.4$ serves as a practical compromise between these two extremes. This also explains why, when comparing with Δ_k-GenEO, we sometimes use different numerical values of τ: the two methods generate different local spectra, so slightly different thresholds are required in order to compare methods at broadly similar coarse-space dimensions rather than at identical threshold values.

Remark 5.5.1. *In this study, we deliberately emphasise iteration counts and coarse-space dimensions rather than reported CPU timings. While wall-clock timings can certainly be informative, they are highly sensitive to implementation details, stopping criteria, hardware architecture and the extent to which the software has been optimised. They therefore do not provide a sufficiently reproducible or platform-independent basis for comparison.*

In the context of two-level domain decomposition methods, iteration counts and coarse-space dimension are standard and reliable proxies for computational performance. A reduction in iteration count directly decreases the number of preconditioner applications, and hence the number of local and coarse solves required, while a smaller coarse space lowers the cost of coarse-level factorisation and solves and can also reduce communication costs. It follows that the consistently lower iteration counts and smaller coarse spaces observed for H_k-GenEO provide strong evidence that, in an efficient implementation, the method would also yield a lower overall time-to-solution, even though such timings are not reported here.

5.5.4 Analyses and Comparison

We can now draw several conclusions from the homogeneous and heterogeneous tests carried out for the different threshold values used.

Before discussing the individual threshold values, it is useful to compare the numerical observations with the sufficient conditions from Corollary 5.3.2. The theory identifies two main requirements for robust convergence: a geometric condition $kH_\ell \lesssim 1$, and a spectral condition $(1 + C_{\text{stab}})^2 k^2 \Theta \lesssim 1$, equivalently that τ is not too small. The numerical results support both trends qualitatively. When N increases, the subdomain size H_ℓ decreases, the local problems become less indefinite, and the tables show fewer negative eigenvalues together with a generally flatter iteration history. Likewise, when τ is increased, more negative and weakly positive local modes are retained in the coarse space, and the iteration counts become substantially more stable. At the same time, the numerical results are clearly less restrictive than the theorem in a quantitative sense: acceptable convergence is still observed in several cases where the subdomain size, measured in wavelengths, is larger than one wavelength. This is not a contradiction, since the theory gives a sufficient condition rather than a sharp threshold, and in practice a richer coarse space can partially compensate for a larger value of H_ℓ.

Results for $\tau = 0.4$. Table 5.2 and Table 5.5 reports performance metrics for varying k and number of subdomains N, while keeping the fine mesh fixed with $h^{-1} = k \times 12$, ensuring adequate resolution for the homogeneous and heterogenous test cases respectively. The following trends are evident:

- As k increases, the number of negative eigenvalues per subdomain (Max. Num. Neg) increases, and the smallest local eigenvalue becomes more negative. This reflects the increasing indefiniteness of the problem. A feature specific to heterogeneous media is the contrast-induced reduction of the negative eigenvalues. It is observed that the most negative eigenvalue becomes less negative for the higher contrast. This coincides with fewer negative eigenvalues being taken per-subdomain. The effect is less noticeable as the size of the domains decreases (for the larger N), as the problem is naturally becoming more positive.

- Despite this, the GMRES iteration count remains relatively stable (between 14 and 25), demonstrating that the H_k-GenEO coarse space successfully captures the problematic modes. However, the Δ_k-GenEO has a more obvious variation in GMRES iterations, indicating an instability in the method. Whilst the Δ_k-GenEO method is comparable for the low to moderate k, it is unable to achieve the same low GMRES iteration for higher k.

- The coarse space dimension increases with both k and N, as expected, but this growth is moderate. Even at $k = 100$, the iteration counts remain acceptable up to $N = 144$.

Overall, the $\tau = 0.4$ results are the closest to the theoretical design principle in Section 5.3: the threshold is large enough to recover stable convergence over a wide range of k and N, but the resulting coarse space is still moderate in size. In this sense, $\tau = 0.4$ acts as a useful baseline choice, balancing robustness and coarse-level cost.

Results for $\tau = 0.2$. Table 5.3 and Table 5.6 presents the same homogeneous and heterogeneous experiments with a lower threshold $\tau = 0.2$. This tighter cutoff includes fewer eigenfunctions in the coarse space, leading to smaller coarse space dimensions but also higher iteration counts:

- For low k, the H_k-GenEO performance is acceptable, but for higher values (e.g., $k = 80$ and $N = 144$), GMRES iteration counts grow significantly. When this is compared to the Δ_k-GenEO results, the GMRES iteration count is comparable for low k, but the difference becomes more significant with moderate to high k. In the heterogenous test cases, iteration counts remain acceptable across N and $a_{\max}$, mirroring the homogeneous case. At moderate and high frequency, the iteration count is affected more significantly by the heterogeneity. For $N = 16$, $k = 60$, iterations drop from 68 with $a_{\max} = 10$ to 32 when $a_{\max} = 1000$. And when $N = 64$, $k = 60$, the iteration count drops from 56 to 35. Whilst the Δ_k-GenEO does see a decrease in iteration counts with an increase in $a_{\max}$, it still remains considerably higher then the H_k-GenEO.

- Compared to $\tau = 0.4$, we observe that reducing the coarse space too aggressively leads to a loss of robustness, especially as k increases.

- This validates the theoretical prediction: a smaller eigenvalue threshold requires a correspondingly finer decomposition or larger coarse space to maintain stability.

Results for $\tau = 0.6$. Table 5.4 and Table 5.7 reports the homogeneous and heterogenous results for a higher eigenvalue threshold $\tau = 0.6$, meaning that more eigenfunctions are included in the coarse space. Compared to the previous results, we see a consistent trend:

- Coarse space dimensions grow substantially, with roughly 1.5–2$\times$ more basis functions than in the $\tau = 0.4$ case, especially for higher k, displaying only a mild dependence on $a_{\max}$. The corresponding Δ_k-GenEO coarse spaces remain typically larger than H_k-GenEO in order to obtain comparable iteration counts.

- GMRES iteration counts, however, drop significantly—typically to 11–16 iterations, nearly independent of both k and N. This trend is also noticible in the heterogenous test cases accross the different $a_{\max}$ values, although with a slightly higher range, typically 15–20. When this is compared to the Δ_k-GenEO, with a comparable coarse space dimension, the iteration count remains higher for the low k, with a steady increase in iterations as k increases. This demonstrates the benefit of using the H_k-GenEO method, which utilises the indefiniteness of the problem.

- Even for the most challenging cases (e.g., $k = 100$, $N = 100$), convergence is maintained.

Taken together, Tables 5.2–5.7 show that H_k-GenEO maintains scalability and robustness with respect to both the number of subdomains N and the wavenumber k, remaining effective even as indefiniteness grows and with an increase in heterogeneity. Across $\tau \in \{0.2, 0.4, 0.6\}$ and in heterogeneous media, the H_k-GenEO method is the more iteration stable choice, while Δ_k-GenEO typically requires larger coarse spaces and exhibits higher, more variable iteration

counts. Increasing τ enlarges the coarse space and improves robustness. At $\tau = 0.2$ the reduced coarse space size leads to a loss of stability as k and N grow. With $\tau = 0.4$ the trade–off between coarse space size and GMRES iteration counts is broadly favourable, and then with $\tau = 0.6$ convergence is nearly k and N independent with only rare outliers. The clear benefit of a more generous threshold is a preconditioner that is more robust and less sensitive to frequency, with flattened iteration counts across parameter ranges.

The very large iteration counts that appear for some of the higher-frequency cases are therefore not unexpected. When τ is small, some negative and near-zero positive eigenmodes are no longer represented in the coarse space, so the preconditioner loses the very components needed to control the global low-energy error. This effect is strongest when k is large and N is small, since these are precisely the cases where the local problems are most indefinite and the number of problematic modes is largest. In addition, because the problem has Dirichlet boundary conditions, the stability constant C_{stab} deteriorates when k^2 approaches a Dirichlet eigenvalue of the continuous operator. Hence one should expect isolated spikes in iteration count for certain wavenumbers, rather than a perfectly monotone dependence on k. These outliers become particularly visible for $\tau = 0.2$, because the coarse space then has little redundancy and therefore less margin to absorb resonance-related deterioration.

From the perspective of the theory, the numerical experiments therefore confirm the qualitative role of both parameters: smaller H_ℓ and larger τ both improve robustness, while large iteration counts occur when one or both of these mechanisms are weakened. The agreement is strongest at the level of trends rather than sharp thresholds. In particular, the computations suggest that the sufficient conditions from Corollary 5.3.2 are conservative: convergence is often still acceptable outside the nominal regime $kH_\ell \lesssim 1$, provided the spectral enrichment is strong enough. In Tables 5.5, 5.6 and 5.7 we show the dimension of the fine mesh (dim(Fine)), the maximum number of negative eigenvalues per subdomain (Max. Num. Neg), the minimum eigenvalue ($\lambda_{\min}$), the maximum number of modes per domain used in the coarse space (dim(Loc)), the dimension of the coarse space ((dim(CS)), and the GMRES iteration count for a varying k for a given threshold τ.

5.5.5 Spectral Complexity of Local Eigenproblems

Table 5.8 presents the number of local eigenvalues per subdomain lying below the threshold $\tau = 0.4$, separated into negative and positive contributions. These values are crucial for determining the size and effectiveness of the H_k-GenEO coarse space, as all eigenfunctions corresponding to eigenvalues below τ are retained. The table reveals several key trends:

- **Spectral shift with k:** For fixed N, both the number of negative eigenvalues (Max.Neg) and the number of small positive eigenvalues (Max.Pos) increase as k increases. This is consistent with the growing indefiniteness of the operator and reflects the fact that more unstable and near-resonant modes emerge locally.
- **Effect of domain decomposition size (N):** At fixed k, increasing N (i.e., using smaller subdomains) leads to fewer negative and small positive eigenvalues. For example, at $k = 60$, the Max.Neg count drops from 22 (for $N = 16$) to only 4 (for $N = 144$) for the homogeneous problem. This aligns with theoretical predictions: when subdomains are small enough (i.e., $H_\ell \lesssim k^{-1}$), local problems tend to become positive definite or only weakly indefinite.

In summary, this table provides concrete spectral evidence for the theoretical bounds derived in Section 5.3: to ensure robustness, the coarse space must include all modes with eigenvalues $\leq \tau$, including both negative and small positive ones. Moreover, it illustrates the advantage of fine subdomain resolution (large N): the number of problematic modes per patch decreases significantly, allowing for more efficient coarse spaces without loss of robustness.

					H_k-GenEO					Δ_k-GenEO		
N	$a_{\max}$	k	h^{-1}	Dim(Fine)	Max.Num. Neg.	$\lambda_{\min}$	Dim(Loc)	Dim(CS)	It. Count	Dim(Loc)	Dim(CS)	It. Count
16	10	20	240	58081	2.0	-0.0795278	18	196	19	26	408	20
		60	720	519841	15.0	-0.620619	74	866	19	94	1500	24
		100	1200	1442401	35.0	-0.840064	155	1860	20	181	2890	38
16	1000	20	240	58081	2.0	-0.0010342	18	190	19	25	404	15
		60	720	519841	12.0	-0.523291	73	826	20	90	1446	21
		100	1200	1442401	31.0	-0.796087	151	1758	18	174	2780	32
36	10	20	240	58081	2.0	-0.0444599	13	322	19	19	682	19
		60	720	519841	7.0	-0.512559	45	1220	19	63	2274	38
		100	1200	1442401	17.0	-0.785888	89	2456	19	117	4216	38
36	1000	20	240	58081	2.0	-0.000860171	13	320	19	19	678	16
		60	720	519841	7.0	-0.473817	45	1180	20	62	2224	27
		100	1200	1442401	15.0	-0.767375	89	2382	49	115	4126	35
64	10	20	240	58081	1.0	-0.0381661	9	446	19	14	912	18
		60	720	519841	6.0	-0.437747	33	1580	50	48	3052	76
		100	1200	1442401	13.0	-0.738717	65	3100	18	86	5512	60
64	1000	20	240	58081	1.0	-0.00392729	9	458	20	14	916	16
		60	720	519841	5.0	-0.417451	34	1578	20	47	3026	26
		100	1200	1442401	11.0	-0.730374	66	3012	21	85	5432	55
100	10	20	240	58081	1.0	-0.0364662	8	572	19	12	1178	18
		60	720	519841	4.0	-0.37355	29	2094	19	38	3832	42
		100	1200	1442401	12.0	-0.689838	55	3882	19	69	6874	45
100	1000	20	240	58081	1.0	-0.000587616	8	548	18	12	1168	16
		60	720	519841	3.0	-0.361218	29	2122	18	39	3864	26
		100	1200	1442401	10.0	-0.687193	57	3872	19	68	6824	26
144	10	20	240	58081	1.0	-0.0513371	6	646	21	10	1456	17
		60	720	519841	4.0	-0.295028	20	2250	23	32	4628	59
		100	1200	1442401	8.0	-0.601849	41	4328	21	56	8078	82
144	1000	20	240	58081	1.0	-0.016502	7	660	20	10	1436	16
		60	720	519841	3.0	-0.288366	22	2268	21	32	4608	36
		100	1200	1442401	8.0	-0.599871	42	4266	119	56	8034	71

Table 5.5: The threshold value: $\tau = 0.4$ for the H_k-GenEO and $\tau = 0.6$ for the Δ_k-GenEO.

5.5.6 Threshold Selection

One of the key features for the spectral coarse space, is the selection of an adequate threshold values, τ, to ensure a suitable coarse space size. This has been identified in the numerical results as an area that can significantly effect the robustness of the method. In previous spectral coarse spaces that have been based on positive (semi-)definite eigenvalue problems, the threshold value has been selected in order to capture the eigenfunctions that correspond to the smallest, and most problematic, eigenvalues. As the eigenvalue problem (5.2.18) is indefinite, setting a maximum threshold value means that the entire negative part of the spectrum is also captured. We now explore the effect of taking a lower bound threshold value. In Table 5.9, we have a comparison between taking $\tau = 0.6$ as the standard upper threshold limit, and then comparing this with taking an upper threshold of $\tau_+ = 0.6$, and a lower threshold of $\tau_- = -0.6$. In Table 5.10, we have a comparison between $\tau = 0.2$ as the standard upper threshold limit, with $\tau = +/- 0.2$ as the upper and lower threshold limit. Both Table 5.9 and Table 5.10 are for the homogeneous test case.

We can see from Table 5.9, that setting the lower threshold limit has little effect on the results. Whilst there is a reduction in the coarse space dimension, this appears to have little effect on the iteration counts. For the high N, where the subdomains problem is less indefinite with the smaller subdomain diameter, the reduction in the number of negative eigenvalues being admitted to the coarse space is insignificant. With a decrease of at most one eigenvalue per-subdomain for the largest k. This reduction has little to no effect on the iteration counts obtained. However, when we look at the results for the lower N, where the problem remains more indefinite, the effect of having the lower threshold value starts to become more noticeable. When the larger k values are being used, there is a drop of at most 8 eigenfunctions per-subdomain being admitted into the coarse space. This causes a mild reduction in the size of the coarse space, but again

					H_k-GenEO					Δ_k-GenEO		
N	$a_{\max}$	k	h^{-1}	Dim(Fine)	Max.Num. Neg.	$\lambda_{\min}$	Dim(Loc)	Dim(CS)	It. Count	Dim(Loc)	Dim(CS)	It. Count
16	10	20	240	58081	2.0	-0.0795278	11	112	25	12	188	23
		60	720	519841	15.0	-0.620619	42	474	68	40	642	107
		100	1200	1442401	35.0	-0.840064	89	1058	35	74	1188	166
16	1000	20	240	58081	2.0	-0.0010342	12	112	25	12	186	19
		60	720	519841	12.0	-0.523291	43	452	32	40	638	62
		100	1200	1442401	31.0	-0.796087	85	974	26	74	1178	137
36	10	20	240	58081	2.0	-0.0444599	7	180	28	9	316	21
		60	720	519841	7.0	-0.512559	25	650	30	28	1002	156
		100	1200	1442401	17.0	-0.785888	49	1350	35	50	1806	> 200
36	1000	20	240	58081	2.0	-0.000860171	7	178	26	9	312	19
		60	720	519841	7.0	-0.473817	26	640	28	28	1016	85
		100	1200	1442401	15.0	-0.767375	49	1288	62	50	1786	> 200
64	10	20	240	58081	1.0	-0.0381661	5	220	31	7	436	22
		60	720	519841	6.0	-0.437747	19	854	56	21	1366	167
		100	1200	1442401	13.0	-0.738717	37	1658	34	38	2400	> 200
64	1000	20	240	58081	1.0	-0.00392729	7	218	27	7	430	21
		60	720	519841	5.0	-0.417451	20	830	35	22	1386	118
		100	1200	1442401	11.0	-0.730374	38	1634	34	38	2402	> 200
100	10	20	240	58081	1.0	-0.0364662	4	318	27	6	564	21
		60	720	519841	4.0	-0.37355	17	1130	32	19	1904	113
		100	1200	1442401	12.0	-0.689838	33	2142	37	32	3210	> 200
100	1000	20	240	58081	1.0	-0.000587616	5	360	23	5	548	19
		60	720	519841	3.0	-0.361218	19	1194	28	20	2012	57
		100	1200	1442401	10.0	-0.687193	35	2232	25	33	3282	167
144	10	20	240	58081	1.0	-0.0513371	4	410	30	4	614	23
		60	720	519841	4.0	-0.295028	13	1202	82	14	2070	168
		100	1200	1442401	8.0	-0.601849	23	2198	53	26	3708	> 200
144	1000	20	240	58081	1.0	-0.016502	4	408	25	4	632	20
		60	720	519841	3.0	-0.288366	14	1222	68	15	2104	124
		100	1200	1442401	8.0	-0.599871	25	2162	114	26	3694	> 200

Table 5.6: The threshold value is set to $\tau = 0.2$.

has little to no effect on the number of iterations.

When the threshold is reduced, Table 5.10 shows that this limitation starts to affect the robustness of the method. Across all the N and k being tested, we can see a considerable drop in the number of negative eigenvalues for each sub-domain. This decrease in coarse space size leads to an increase in the number of iterations, generally showing an increase of 10–30, with the higher values of k showing the largest increase. This indicates that the lower threshold limit reduces the number of problematic eigenfunctions admitted into the coarse space, resulting in slower convergence. Since the purpose of the coarse space is to capture the highly oscillatory eigenfunctions, lowering the threshold can prevent this. This is indicated by Figure 5.1, which shows the eigenplots for a floating subdomain (one that shares no edge with the global boundary). When we use $\tau_- = -0.6$, we can see that the eigenfunction in Figure 5.2d is still mildly oscillatory, but the effect becomes less significant as the eigenvalues get smaller, suggesting that admitting these smaller eigenfunctions into the coarse space would have minimal effect. When we go to $\tau_- = -0.2$, we can see that the eigenfunction in Figure 5.2g is still highly oscillatory, with the smaller eigenfunctions remaining highly oscillatory. This demonstrates that using an aggressive lower bound can reduce the number of highly oscillatory eigenfunctions included in the coarse space, thereby impeding the effectiveness of the coarse space. The same effect can be seen in the eigenfunctions for the heterogeneous test case in Figure 5.2.

					H_k-GenEO					Δ_k-GenEO		
N	$a_{\max}$	k	h^{-1}	Dim(Fine)	Max.Num. Neg.	$\lambda_{\min}$	Dim(Loc)	Dim(CS)	It. Count	Dim(Loc)	Dim(CS)	It. Count
16	10	20	240	58081	2.0	-0.0795278	29	340	16	26	408	20
		60	720	519841	15.0	-0.620619	119	1468	15	94	1500	24
		100	1200	1442401	35.0	-0.840064	248	3132	18	181	2890	38
16	1000	20	240	58081	2.0	-0.0010342	29	332	16	25	404	15
		60	720	519841	12.0	-0.523291	115	1392	17	90	1446	21
		100	1200	1442401	31.0	-0.796087	241	2974	15	174	2780	32
36	10	20	240	58081	2.0	-0.0444599	20	556	16	19	682	19
		60	720	519841	7.0	-0.512559	71	2066	15	63	2274	38
		100	1200	1442401	17.0	-0.785888	142	4182	16	117	4216	38
36	1000	20	240	58081	2.0	-0.000860171	19	544	17	19	678	16
		60	720	519841	7.0	-0.473817	69	2018	16	62	2224	27
		100	1200	1442401	15.0	-0.767375	141	4026	22	115	4126	35
64	10	20	240	58081	1.0	-0.0381661	14	732	17	14	912	18
		60	720	519841	6.0	-0.437747	52	2660	46	48	3052	76
		100	1200	1442401	13.0	-0.738717	103	5160	16	86	5512	60
64	1000	20	240	58081	1.0	-0.00392729	14	736	17	14	916	16
		60	720	519841	5.0	-0.417451	52	2638	17	47	3026	26
		100	1200	1442401	11.0	-0.730374	103	5000	17	85	5432	55
100	10	20	240	58081	1.0	-0.0364662	12	932	17	12	1178	18
		60	720	519841	4.0	-0.37355	44	3392	16	38	3832	42
		100	1200	1442401	12.0	-0.689838	84	6242	16	69	6874	45
100	1000	20	240	58081	1.0	-0.000587616	12	932	16	12	1168	16
		60	720	519841	3.0	-0.361218	44	3352	16	39	3864	26
		100	1200	1442401	10.0	-0.687193	85	6200	16	68	6824	26
144	10	20	240	58081	1.0	-0.0513371	9	1120	17	10	1456	17
		60	720	519841	4.0	-0.295028	33	3894	18	32	4628	59
		100	1200	1442401	8.0	-0.601849	65	7116	16	56	8078	82
144	1000	20	240	58081	1.0	-0.016502	10	1150	17	10	1436	16
		60	720	519841	3.0	-0.288366	32	3856	17	32	4608	36
		100	1200	1442401	8.0	-0.599871	65	7032	22	56	8034	71

Table 5.7: The threshold value is set to $\tau = 0.6$.

				Homogeneous		Heterogenous	
N	k	h^{-1}	Dim(Fine)	Max.Neg.	Max.Pos	Max.Neg.	Max.Pos
16	20	240	58081	4.0	14.0	2.0	16.0
	40	480	231361	13.0	36.0	7.0	35.0
	60	720	519841	22.0	64.0	12.0	61.0
	80	960	923521	41.0	91.0	18.0	90.0
	100	1200	1442401	58.0	132.0	31.0	120.0
36	20	240	58081	3.0	9.0	2.0	12.0
	40	480	231361	6.0	22.0	3.0	24.0
	60	720	519841	13.0	36.0	7.0	40.0
	80	960	923521	20.0	52.0	9.0	55.0
	100	1200	1442401	28.0	71.0	15.0	75.0
64	20	240	58081	1.0	7.0	1.0	8.0
	40	480	231361	4.0	14.0	3.0	17.0
	60	720	519841	8.0	26.0	5.0	29.0
	80	960	923521	13.0	36.0	8.0	40.0
	100	1200	1442401	17.0	48.0	11.0	55.0
100	20	240	58081	1.0	7.0	1.0	7.0
	40	480	231361	3.0	13.0	2.0	17.0
	60	720	519841	4.0	20.0	3.0	26.0
	80	960	923521	8.0	26.0	6.0	38.0
	100	1200	1442401	13.0	36.0	10.0	47.0
144	20	240	58081	1.0	4.0	1.0	6.0
	40	480	231361	3.0	9.0	2.0	11.0
	60	720	519841	4.0	14.0	3.0	19.0
	80	960	923521	6.0	22.0	6.0	24.0
	100	1200	1442401	8.0	31.0	8.0	34.0

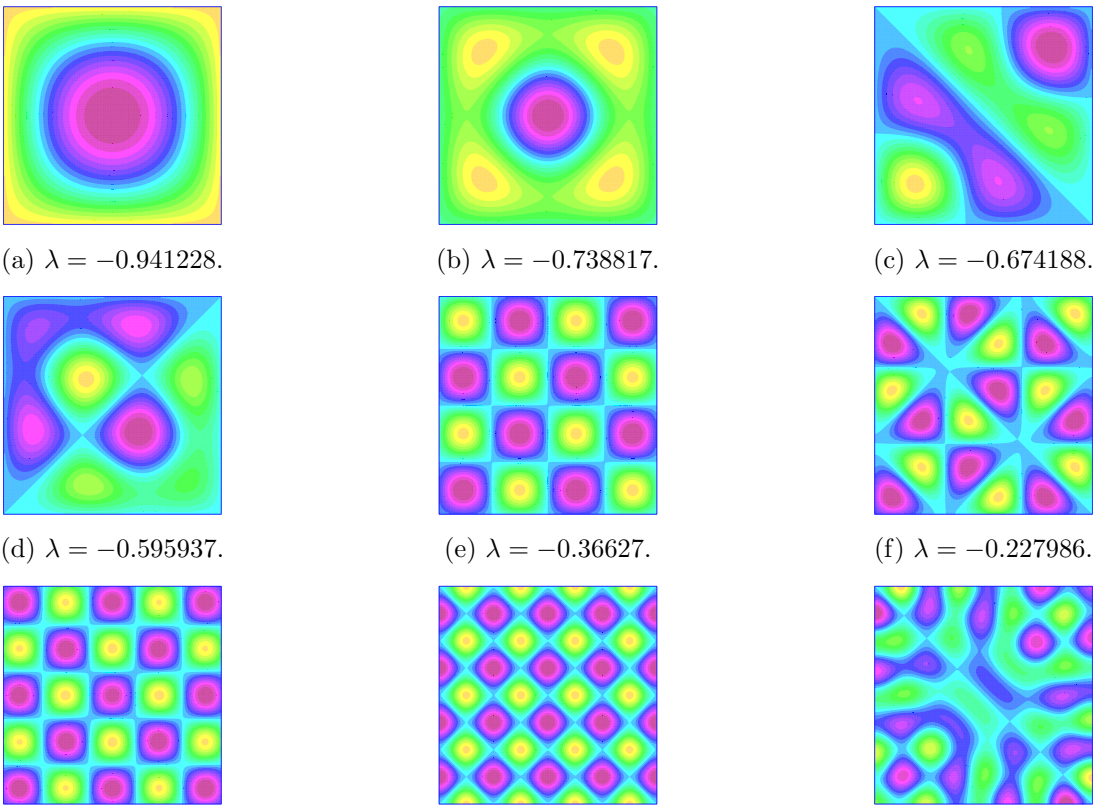
Table 5.8: Results showing the dimension of the fine mesh (dim(Fine), the maximum number of negative eigenvalues per subdomain (Max. Num. Neg) and the maximum number of positive eigenvalues per subdomain (Max. Num. Pos). The threshold value is set to $\tau = 0.4$, with $a_{\max} = 1000$ used for the heterogenous results.

				$\tau = +\backslash -0.6$				$\tau = 0.6$			
N	k	h^{-1}	Dim (Fine)	Max.Num. Neg.	$\lambda_{\min}$	Dim (CS)	It. Count	Max.Num. Neg.	$\lambda_{\min}$	Dim (CS)	It. Count
16	20	240	58081	4	-0.324293	392	14	4.0	-0.324293	392	14
	60	720	519841	19	-0.52068	1936	15	22.0	-0.849475	1984	15
	100	1200	1442401	50	-0.595937	4444	30	58.0	-0.941228	4572	29
36	20	240	58081	3	-0.178013	600	15	3.0	-0.178013	600	15
	60	720	519841	12	-0.419254	2540	15	13.0	-0.703879	2576	15
	100	1200	1442401	25	-0.587053	5476	54	28.0	-0.875285	5584	56
64	20	240	58081	1	-0.125924	740	15	1.0	-0.125924	740	15
	60	720	519841	8	-0.553713	3232	15	8.0	-0.553713	3232	15
	100	1200	1442401	16	-0.562646	6444	15	17.0	-0.793739	6508	15
100	20	240	58081	1	-0.100482	1044	17	1.0	-0.100482	1044	17
	60	720	519841	4	-0.423257	3748	16	4.0	-0.423257	3748	16
	100	1200	1442401	12	-0.419254	7436	16	13.0	-0.703879	7536	15
144	20	240	58081	1	-0.0852859	1096	17	1.0	-0.0852859	1096	17
	60	720	519841	4	-0.324293	4232	16	4.0	-0.324293	4232	16
	100	1200	1442401	7	-0.599849	8400	16	8.0	-0.612531	8500	16

Table 5.9: Results for the homogeneous problem showing the dimension of the fine mesh (dim(Fine)), the maximum number of negative eigenvalues per subdomain (Max. Num. Neg), the minimum eigenvalue in the coarse space ($\lambda_{\min}$), the maximum number of modes per domain used in the coarse space (dim(Loc)), the dimension of the coarse space ((dim(CS)), and the GMRES iteration count for a varying k. The threshold value is set to $\tau = +/ - 0.6$.

				$\tau = +\backslash -0.2$				$\tau = 0.2$			
N	k	h^{-1}	Dim (Fine)	Max.Num. Neg.	$\lambda_{\min}$	Dim (CS)	It. Count	Max.Num. Neg.	$\lambda_{\min}$	Dim (CS)	It. Count
16	20	240	58081	3	-0.0956413	128	29	4.0	-0.324293	144	21
	60	720	519841	11	-0.199637	512	33	22.0	-0.849475	668	23
	100	1200	1442401	28	-0.187746	1132	62	58.0	-0.941228	1612	52
36	20	240	58081	3	-0.178013	220	22	3.0	-0.178013	220	22
	60	720	519841	9	-0.176594	700	68	13.0	-0.703879	844	27
	100	1200	1442401	15	-0.170618	1408	78	28.0	-0.875285	1876	68
64	20	240	58081	1	-0.125924	224	32	1.0	-0.125924	224	32
	60	720	519841	5	-0.194464	900	104	8.0	-0.553713	1060	32
	100	1200	1442401	9	-0.187731	1796	62	17.0	-0.793739	2276	33
100	20	240	58081	1	-0.100482	324	31	1.0	-0.100482	324	31
	60	720	519841	3	-0.148092	1044	45	4.0	-0.423257	1144	37
	100	1200	1442401	9	-0.176594	2044	154	13.0	-0.703879	2444	37
144	20	240	58081	1	-0.0852859	484	27	1.0	-0.0852859	484	27
	60	720	519841	3	-0.0956413	1440	125	4.0	-0.324293	1584	29
	100	1200	1442401	5	-0.160377	2200	85	8.0	-0.612531	2632	56

Table 5.10: Results for the homogeneous problem showing the dimension of the fine mesh (dim(Fine)), the maximum number of negative eigenvalues per subdomain (Max. Num. Neg), the minimum eigenvalue in the coarse space ($\lambda_{\min}$), the maximum number of modes per domain used in the coarse space (dim(Loc)), the dimension of the coarse space (dim(CS)), and the GMRES iteration count for a varying k. The threshold value is set to $\tau = +/ - 0.2$.

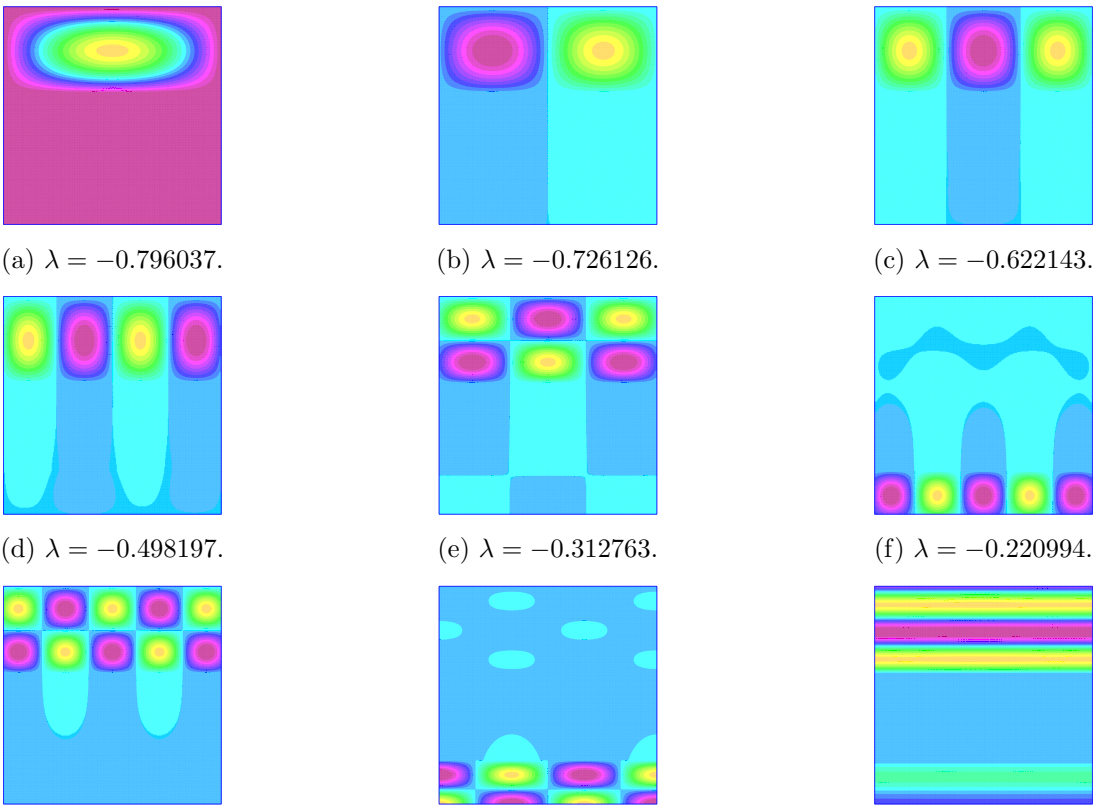


(a) $\lambda = -0.941228$. (b) $\lambda = -0.738817$. (c) $\lambda = -0.674188$.

(d) $\lambda = -0.595937$. (e) $\lambda = -0.36627$. (f) $\lambda = -0.227986$.

(g) $\lambda = -0.187746$. (h) $\lambda = -0.000630012$. (i) $\lambda = 0.00348611$.

Figure 5.1: Eigenvector plots for the homogeneous problem for $k = 100$ and $h^{-1} = 1200$.



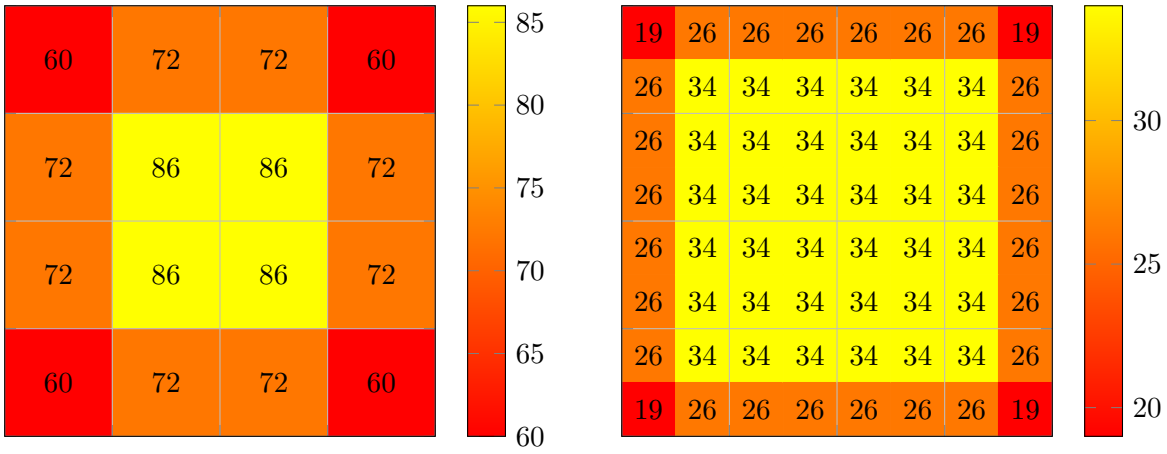
(a) $\lambda = -0.796037$. (b) $\lambda = -0.726126$. (c) $\lambda = -0.622143$.

(d) $\lambda = -0.498197$. (e) $\lambda = -0.312763$. (f) $\lambda = -0.220994$.

(g) $\lambda = -0.156157$. (h) $\lambda = -0.000110566$. (i) $\lambda = 0.000605787$.

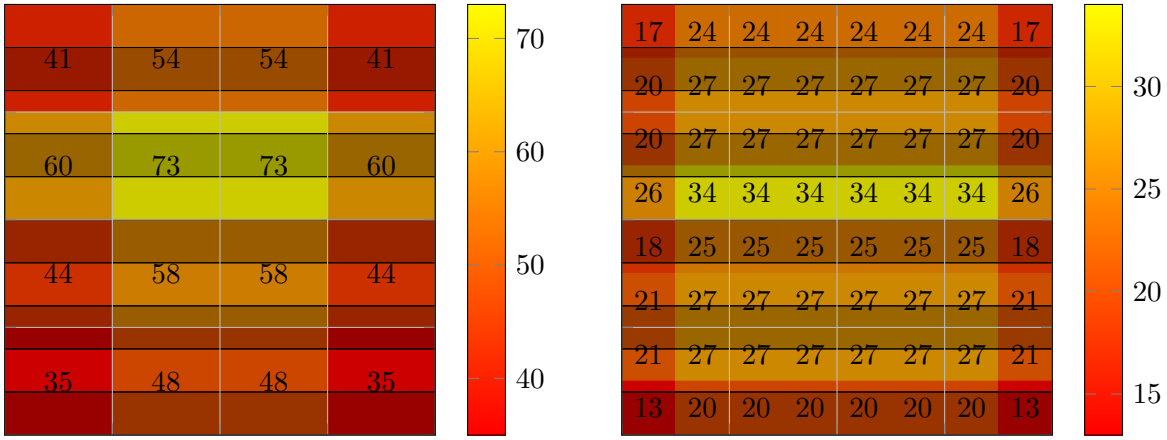
Figure 5.2: Eigenvector plots for the heterogeneous problem for $k = 100$ and $h^{-1} = 1200$.

We see in Figure 5.3 the number of eigenfunctions being selected per-subdomain for the homogeneous (top) and heterogeneous (bottom). We can see that the number of eigenfunctions being selected is influenced by the global boundary conditions and the heterogeneity. This suggests that there is the potential for dynamically adapting the size of the local coarse space on each subdomain. However, the additional computational cost in altering the selection criteria could outweigh the benefits.



(a) $k = 60$, $N = 16$, $\tau = 0.4$.

(b) $k = 60$, $N = 64$, $\tau = 0.4$.



(c) $k = 60$, $N = 16$, $\tau = 0.4$.

(d) $k = 60$, $N = 64$, $\tau = 0.4$.

Figure 5.3: Number of eigenvalues used in the coarse space from each subdomain for the homogeneous (top) and heterogenous (bottom) tests.

5.5.7 Decoupled Decomposition

We now analyse the effect of building the coarse space using separate decompositions for the one level method, using Q decompositions, and the coarse-level using N decompositions. The results in Table 5.11 show the coarse space size and iteration count when using $Q = 144$ and varying the coarse decomposition $N \in \{16, 25, 36, 64, 100, 144\}$ for the homogeneous problem. It is over these coarse decompositions that the generalised eigenvalue is solved. The following observations are obtained from the results:

- The iteration counts remain similar across the majority of k values whilst using significantly smaller coarse spaces. When using the coarsest decompositions, we experience a very mild increase in iteration counts of 1–4 iterations, but a reduction in coarse space size of 40–60%.
- When $k = 80$, we see that the method of using a coarser decomposition is not as effective. This could be an issue with using $Q = 144$ for the underlying one-level method, as the iterations are unusually high when $N = 144$ compared to the other k values. This effect can be seen in Tables 5.2–5.4 and is more prominent when a lower threshold value is used. This is believed to be a side effect of using Dirichlet boundary conditions and the possibility of having a k^2 that is approaching an eigenvalue of the negative Laplacian, as discussed in Chapter 2.

Reducing the number of subdomains in the coarse decomposition does reduce the size of the coarse space, but at the cost of a slight increase in iterations. If memory restrictions are a consideration, then reducing the coarse space size could be an attractive proposition.

			$N = 144$		$N = 100$		$N = 64$		$N = 36$		$N = 25$		$N = 16$	
k	h^{-1}	Dim(Fine)	CS	It.	CS	It.	CS	It.	CS	It.	CS	It.	CS	It.
20	240	58081	1096	17	1044	20	740	24	600	19	530	18	392	20
40	480	231361	2640	20	2360	23	1800	24	1404	25	1230	21	1068	23
60	720	519841	4232	16	3748	23	3232	20	2576	20	2285	17	1984	20
80	960	923521	6048	33	5360	52	4732	45	3908	42	3527	44	3152	41
100	1200	1442401	8500	16	7536	19	6508	20	5584	19	5065	18	4572	19

Table 5.11: Results showing the coarse space size (CS) and iteration count (It.) for a one-level decomposition of $Q = 144$ subdomains and varying coarse level decompositions (N) for $\tau = 0.6$ in the homogeneous domain.

We now examine the results in Table 5.12 for the heterogeneous case with $a_{\max} = 1000$. In contrast to the homogeneous test, using different coarse and one-level decompositions yields less favourable results. Although the coarse-space dimension is reduced, the GMRES iteration count increases with both k and the subdomain size. This decrease in performance is attributed to the interaction between the heterogeneity and the subdomain size and location. Figure 5.2 shows that the heterogeneity significantly influences the eigenvectors, while Figure 5.3 demonstrates that the number of selected eigenvectors per subdomain is also effected by the size and position of the subdomain relative to the heterogeneity. Increasing the subdomain size changes the portion of heterogeneity contained in each subdomain, thereby altering the spectra of the local generalized eigenvalue problems. As a consequence, the eigenmodes selected for the coarse space will no longer correspond with the modes required to accelerate convergence. This effect can be reduced by careful selection of the coarse decomposition. In Table 5.13, we have altered the coarse decomposition so it has the same dimension as the one-level decomposition perpendicular to the layers, with altering dimensions parallel to the layers.

			$N = 144$		$N = 100$		$N = 64$		$N = 36$		$N = 25$		$N = 16$	
k	h^{-1}	Dim(Fine)	CS	It.	CS	It.	CS	It.	CS	It.	CS	It.	CS	It.
20	240	58081	1150	17	932	24	736	27	544	35	437	31	332	42
40	480	231361	2448	16	2104	25	1618	37	1206	49	1032	42	802	60
60	720	519841	3856	17	3352	89	2638	130	2018	62	1757	117	1446	128
80	960	923521	5370	20	4670	44	3774	63	2952	80	2597	78	2144	97
100	1200	1442401	7032	22	6200	79	5000	> 200	4026	193	3563	171	2974	197

Table 5.12: Results showing the coarse space size (CS) and iteration count (It.) for a one-level decomposition of $Q = 144$ subdomains and varying coarse level decompositions (N) for $\tau = 0.6$ in the heterogeneous domain.

The results in Table 5.13 show that reducing the size of the coarse domains in one dimension has allowed for a reduction in the coarse space, whilst retaining acceptable GMRES iteration counts. Although the iteration counts are higher then if the same subdomain size is used for both the one-level and coarse decomposition, the trade off for having a smaller coarse space could be appealing. However, the method starts to lose robustness at the higher values of k, indicating the significant effect the heterogeneity can have. These results demonstrate that careful selection of the decompositions that complement the heterogeneities can improve performance.

			$N = 144$		$N = 120$		$N = 96$		$N = 72$		$N = 60$		$N = 48$	
k	h^{-1}	Dim(Fine)	CS	It.	CS	It.	CS	It.	CS	It.	CS	It.	CS	It.
20	240	58081	1150	17	990	21	908	21	814	22	742	21	712	22
40	480	231361	2448	16	2284	23	2018	25	1818	27	1732	27	1604	27
60	720	519841	3856	17	3552	31	3246	82	2900	33	2772	39	2626	41
80	960	923521	5370	20	5122	34	4540	44	4142	43	3954	43	3724	43
100	1200	1442401	7032	22	6706	52	6022	82	5534	56	5336	71	5002	74

Table 5.13: Results showing the coarse space size (CS) and iteration count (It.) for a one-level decomposition of $Q = 144$ subdomains and varying coarse level decompositions (N) for $\tau = 0.6$ in the heterogeneous domain.

5.6 Conclusion

This chapter developed and analysed a two–level domain decomposition preconditioner for indefinite Helmholtz type problems whose coarse space, H_k-GenEO, is built directly from local generalised eigenvalue problems involving the full indefinite operator. This is in contrast to the Δ-GenEO and Δ_k-GenEO, which rely on nearby SPD problems.

The main development of this chapter has been in the establishment of a theoretical framework for the indefinite eigenvalue problem that is central to the H_k-GenEO method. The key consequence of this is the establishment of simple, k–explicit sufficient conditions for robustness:

$$kH_\ell \lesssim 1, \qquad (1 + C_{\text{stab}})^2 \, k^2 \, \Theta \lesssim 1,$$

with constants independent of the subdomain count, coefficient heterogeneity, and mesh size. Practically, this means:

- the fine subdomain diameter H_ℓ should change with the wavelength,
- the spectral threshold $\tau = \min_i \lambda^i_{m_i+1}$ should be chosen so that it retains all negative and sufficiently small positive local modes.

A structural feature of the H_k-GenEO method is the decoupling the decompositions, meaning the coarse partition diameter H_c does not appear in the robustness conditions. This decoupling allows for increasing the one-level decomposition, benefiting from having smaller local problems to solve, whilst being able to use larger coarse decompositions in an attempt to have a smaller coarse space dimension, while maintaining practical GMRES iterations.

The numerical study on the unit-square for both homogeneous and layered heterogeneous media, using a range of wavenumbers, k, and decompositions, N, demonstrates that the proposed H_k-GenEO coarse space is a robust and scalable preconditioner for indefinite problems.

Across all tests, iteration counts remain stable with respect to both k and N once the eigenvalue threshold τ is chosen to admit a sufficient number of local modes. With the choice of $\tau = 0.4$, GMRES typically converges in $\approx$ 14–25 iterations over broad ranges of k and N, and this behaviour continues in heterogeneous media with only mild sensitivity to $a_{\max}$. Increasing the threshold to $\tau = 0.6$ enlarges the coarse space but further flattens the iteration counts to $\approx$ 11–16, giving convergence that is nearly independent of frequency and subdomain count. In contrast, an aggressive reduction to $\tau = 0.2$ shrinks the coarse space size but noticeably degrades robustness at higher k and larger N. Throughout, H_k-GenEO outperforms or matches Δ_k-GenEO at comparable coarse-space sizes, with the latter generally requiring larger coarse spaces while exhibiting higher and more variable iteration counts.

Experiments with two-sided thresholds show that imposing a negative lower bound has little benefit at moderate-to-generous τ (e.g. $\tau = \pm 0.6$). The coarse space becomes only marginally smaller and iteration counts are essentially unchanged, especially for larger N where local problems become more positive, so fewer negative eigenvalues are obtained. However, when the upper threshold is small (e.g. $\tau = \pm 0.2$), enforcing a negative lower bound can remove highly oscillatory negative modes and significantly affect the convergence.

From the results demonstrating the effect of taking different decompositions, it was shown that for the homogeneous tests it is possible to gain a reduction in the size of the coarse space whilst retaining comparative GMRES iterations. However, for heterogenous domains, it is necessary to adjust the decomposition based on the heterogeneity present.

By constructing the coarse space from the indefinite operator itself and proving k–explicit approximation and decomposition properties, H_k-GenEO delivers a practical and scalable preconditioner for Helmholtz-type systems. The analysis pinpoints two bounds, which can guarantee GMRES convergence rates that are stable with respect to frequency, heterogeneity, and subdomain count.

Chapter 6

Achieving Wavenumber Robustness in Domain Decomposition for Heterogeneous Helmholtz Equation: An Overview of Spectral Coarse Spaces

This chapter is the result of a collaboration with Victorita Dolean, Matthias Langer, Pierre-Henri Tournier and Emile Parolin. The corresponding submitted preprint [27] can be found on the arXiv. It has also been submitted to the journal Computers and Mathematics with Applications for publication.

6.1 Introduction

The central challenge addressed in this chapter is the computational complexity of frequency-domain simulations of wave propagation and scattering in heterogeneous media, as encountered in acoustics, electromagnetics, and seismic imaging. Finite element discretisations of such models yield large-scale linear systems that are typically indefinite and poorly conditioned; the difficulty intensifies in high-frequency regimes because accurate resolution requires mesh refinement at least proportional to the frequency, rendering direct solvers impractical. This motivates the development of scalable, parallel iterative methods, and in particular two-level domain decomposition preconditioners.

We consider a heterogeneous Helmholtz model posed on a bounded computational domain $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$, where the complex-value field $u(\mathbf{x})$ satisfies

$$-\nabla \cdot \big(a(\mathbf{x})\nabla u\big) - \omega^2\, m(\mathbf{x})\, u = f(\mathbf{x}) \qquad \text{in } \Omega, \tag{6.1.1}$$

$$\mathcal{C}(u) = 0 \qquad \text{on } \partial\Omega, \tag{6.1.2}$$

with $\mathcal{C}$ representing the imposed boundary conditions, $a(\mathbf{x}) > 0$ a spatially varying coefficient associated with material stiffness (or inverse density), $m(\mathbf{x}) > 0$ a spatially varying "mass" coefficient related to the squared slowness, $\omega > 0$ the angular frequency and $f(\mathbf{x})$ the source term. The associated local wavenumber is

$$k(\mathbf{x}) = \omega\sqrt{\frac{m(\mathbf{x})}{a(\mathbf{x})}},$$

which reduces to the familiar relation $k(\mathbf{x}) = \omega/c(\mathbf{x})$ when $a \equiv 1$ and $m(\mathbf{x}) = c^{-2}(\mathbf{x})$, with $c(\mathbf{x})$ being the wave speed. In heterogeneous media, spatial variations in a and m lead to additional variations in $k(\mathbf{x})$, making the high-frequency regime (large k) particularly challenging for numerical solvers.

The challenges of designing robust solvers for Helmholtz problems are well documented [40, 48], and despite substantial progress (theoretical: [57, 58, 52]; numerical: [31, 30, 90, 77]), no single approach consistently dominates in the high-frequency heterogeneous regime. Domain decomposition methods are attractive for large-scale settings [29], but their effectiveness depends critically on the choice of coarse space, which must transmit global information efficiently while respecting the nonlocal character of wave solutions.

This chapter focuses on *spectral* coarse spaces constructed from local generalised eigenvalue problems. Representative families include GenEO-type constructions [85, 59], interface-based DtN coarse spaces [23], and more recent Helmholtz-harmonic or extended harmonic coarse spaces inspired by multiscale spectral approximation [71, 62]. Earlier numerical comparisons in a common software environment were reported in [14], together with controlled studies contrasting DtN and GenEO-type Helmholtz variants [11]. Building on these foundations, we substantially broaden the scope to include the most recent coarse-space developments and to evaluate them across a wide range of controlled and realistic benchmarks.

Contributions. The main contributions of this chapter are:

- a unified and reproducible comparison of the major spectral coarse space families for heterogeneous, high-frequency Helmholtz problems, implemented within a common FreeFEM framework;
- large-scale numerical experiments on both idealised test cases and challenging 2D/3D benchmarks, providing new evidence on robustness, scalability, and computational cost;
- practical guidance on coarse-space selection, highlighting empirically observed trade-offs between iteration counts, coarse-space dimension, and setup cost.

Outline. Section 6.2 introduces the heterogeneous Helmholtz model and discretisation. Section 6.3 summarises the two-level domain decomposition framework. Section 6.4 describes the DtN, GenEO-type, and harmonic coarse-space families and their recent variants. Section 6.5 presents the comparative numerical studies, and Section 6.6 concludes.

6.2 Discretisation of the Heterogeneous Helmholtz Problem

We focus here on the interior heterogeneous Helmholtz problem (6.1.1) with suitable boundary conditions. In practical settings, the domain is truncated to a bounded computational domain Ω, and the far-field Sommerfeld radiation condition must be approximated on the artificial boundary of Ω to model wave behaviour appropriately. A widely used approximation is the Robin (impedance) condition, though alternatives include absorbing boundary conditions (ABCs) [94] or perfectly matched layers (PML) [60, 6]. We consider Dirichlet conditions on $\Gamma_D \subset \partial\Omega$ and Robin conditions on $\Gamma_R = \partial\Omega \setminus \Gamma_D$, solving the boundary value problem:

$$-\nabla \cdot \big(a(\boldsymbol{x})\nabla u\big) - \omega^2 m(\boldsymbol{x})\, u = f(\boldsymbol{x}) \qquad \text{in } \Omega, \tag{6.2.1a}$$

$$u = u_{\Gamma_D} \qquad \text{on } \Gamma_D, \tag{6.2.1b}$$

$$a(\boldsymbol{x})\,\frac{\partial u}{\partial \boldsymbol{n}} + i\,\omega\, Z(\boldsymbol{x})\, u = 0 \qquad \text{on } \Gamma_R, \tag{6.2.1c}$$

where $\boldsymbol{n}$ is the outward normal, and $Z(\boldsymbol{x})$ is an impedance coefficient typically chosen as $Z(\boldsymbol{x}) = \sqrt{a(\boldsymbol{x})\, m(\boldsymbol{x})}$ to approximate the Sommerfeld radiation condition. The problem is well posed

if $\Gamma_R \neq \emptyset$ but may become ill-posed for certain parameter choices when $\Gamma_R = \emptyset$. The local wavenumber $k(\boldsymbol{x})$ can vary within Ω and may have discontinuities across material interfaces, allowing $k \in L^\infty(\Omega)$.

For $\Omega' \subset \Omega$, the trial and test spaces are defined as

$$V(\Omega') = \{u \in H^1(\Omega') : u = u_{\Gamma_D} \text{ on } \Gamma_D \cap \partial\Omega'\}, \qquad V_0(\Omega') = \{u \in H^1(\Omega') : u = 0 \text{ on } \Gamma_D \cap \partial\Omega'\}. \tag{6.2.2}$$

The weak formulation seeks $u \in V(\Omega)$ satisfying:

$$a_\Omega(u, v) = F(v), \quad \forall v \in V_0(\Omega), \tag{6.2.3}$$

where, for any $\Omega' \subset \Omega$, we define:

$$a_{\Omega'}(u, v) = \int_{\Omega'} \left(a(\boldsymbol{x})\nabla u \cdot \nabla \bar{v} - \omega^2 m(\boldsymbol{x})\, u\, \bar{v}\right) \mathrm{d}\boldsymbol{x} + \int_{\Gamma_R \cap \partial\Omega'} i\,\omega\, Z(\boldsymbol{x})\, u\, \bar{v}\, \mathrm{d}s, \qquad F(v) = \int_\Omega f\, \bar{v}\, \mathrm{d}\boldsymbol{x}. \tag{6.2.4}$$

Using a simplicial mesh $\mathcal{T}^h$ of Ω with element diameter h, we use finite elements to discretise (6.2.3), as outlined in [14]. A Lagrange finite element approximation reduces the problem to solving the complex linear system:

$$A\mathbf{u} = \mathbf{f}, \tag{6.2.5}$$

where $A \in \mathbb{C}^{n\times n}$ and $\mathbf{f} \in \mathbb{C}^n$ are respectively derived from the sesquilinear and antilinear forms, and n denotes the total number of degrees of freedom in Ω. Accurate approximation requires fine discretisation, particularly as $k(\boldsymbol{x})$ increases, due to the pollution effect [2]. For piecewise linear ($\mathbb{P}_1$) elements, maintaining accuracy requires $k_{\max}^3 h^2$ to remain bounded, where $k_{\max} = \sup_{\mathbf{x}\in\Omega} k(\boldsymbol{x})$, implying $h = \mathcal{O}(k_{\max}^{-3/2})$. Higher-order elements relax this constraint, but interpolation properties degrade for very high orders. An alternative common strategy is to fix the number of points per wavelength, n_{ppwl}, with $h = \mathcal{O}(k_{\max}^{-1})$, where the wavelength is $\lambda(\boldsymbol{x}) = 2\pi/k(\boldsymbol{x})$. Practitioners often use 5 or 10 points per wavelength.

6.3 Domain Decomposition Methods

To solve the discrete heterogeneous Helmholtz problem (6.2.5), we employ GMRES with acceleration provided by a two-level overlapping domain decomposition preconditioner. The core one-level approach used is the optimised restricted additive Schwarz (ORAS) method [29]. To formulate the domain decomposition preconditioner, we begin by dividing the domain Ω into a set of N non-overlapping subdomains $\{\Omega'_s\}_{s=1}^N$, which align with the finite element mesh $\mathcal{T}^h$. Overlapping subdomains $\{\Omega_s\}_{s=1}^N$ are then created by expanding each Ω'_s with neighbouring mesh elements, defined as:

$$\Omega_s = \text{Int}\left(\bigcup_{\text{supp}(\phi_j)\cap\Omega'_s\neq\emptyset} \text{supp}(\phi_j)\right),$$

where $\text{Int}(\cdot)$ and $\text{supp}(\cdot)$ represent the interior and support of a function, respectively, and ϕ_j are the nodal basis functions of the finite element space. Additional layers can be recursively added to increase the overlap as needed.

With the overlapping decomposition, we define the preconditioner components. The restriction operator $R_s \in \mathbb{R}^{n_s\times n}$ restricts functions to Ω_s where n_s denotes the local number of degrees of freedom in Ω_s, while the transpose R_s^T extends by zero outside Ω_s. A partition of unity is represented by diagonal matrices $D_s \in \mathbb{R}^{n_s\times n_s}$, ensuring

$$\sum_{s=1}^N R_s^T D_s R_s = I.$$

ORAS also requires solving local heterogeneous Helmholtz–Robin problems:

$$-\nabla \cdot \big(a(\boldsymbol{x})\nabla w_s\big) - \omega^2 m(\boldsymbol{x})\, w_s = f \qquad \text{in } \Omega_s, \tag{6.3.1a}$$

$$a(\boldsymbol{x})\, \frac{\partial w_s}{\partial \boldsymbol{n}_s} + i\,\omega\, Z(\boldsymbol{x})\, w_s = 0 \qquad \text{on } \partial\Omega_s \setminus \partial\Omega, \tag{6.3.1b}$$

$$\mathcal{C}(w_s) = 0 \qquad \text{on } \partial\Omega_s \cap \partial\Omega, \tag{6.3.1c}$$

where $\mathcal{C}$ applies the boundary conditions (6.2.1b)–(6.2.1c), and $Z(\boldsymbol{x})$ is the local impedance, typically chosen as $Z(\boldsymbol{x}) = \sqrt{a(\boldsymbol{x})\, m(\boldsymbol{x})}$. The finite element discretisation of (6.3.1) yields a local stiffness matrix $\widehat{A}_s \in \mathbb{C}^{n_s \times n_s}$. The ORAS preconditioner is then assembled as:

$$M_{\text{ORAS}}^{-1} = \sum_{s=1}^{N} R_s^T D_s \widehat{A}_s^{-1} R_s,$$

where the local solves $\widehat{A}_s^{-1}$ are performed in parallel.

To ensure robustness and scalability, particularly for indefinite problems like the Helmholtz equation, the ORAS method incorporates a coarse space in a two-level framework. The coarse space, represented by a matrix Z with linearly independent column vectors, plays a key role in addressing scalability. Using deflation, a coarse operator $E = Z^\dagger A Z$ and correction operator $Q = Z E^{-1} Z^\dagger$ are defined, yielding the two-level ORAS method:

$$M_{\text{ORAS},2}^{-1} = M_{\text{ORAS}}^{-1}(I - AQ) + Q.$$

The choice of coarse space is critical, particularly for indefinite problems where its addition may not always improve performance [43]. We next discuss the selection of coarse spaces.

6.4 Spectral Coarse Spaces

In this study, we investigate spectral coarse spaces for the discrete Helmholtz problem (6.2.5), which are constructed by solving local eigenvalue problems on subdomains and assembling the resulting eigenfunctions into a global coarse space. While a broad family of such spectral coarse spaces exists, we focus on two representative classes that follow a similar construction philosophy: **GenEO-type spaces** and **harmonic-type spaces**.

We begin by reviewing the DtN coarse space introduced in [23], then present the Δ-GenEO [13, 26] and H_k-GenEO spaces [24], which are inspired by the GenEO methodology [85]. In parallel, we consider recently developed Helmholtz-harmonic coarse spaces from [62, 71, 74]. Finally, we establish connections between these approaches.

Although comparative analyses were conducted in [14, 11], our study builds upon and extends them by incorporating the latest developments in the field, along with a detailed complexity analysis on both model and benchmark problems. Note that in light of the same recent developments the DtN coarse space can be both considered as an ancestor of GenEO coarse spaces but also, by construction, falls into the category of harmonic coarse spaces.

Notation: Given a variational problem with system matrix B, we denote by B_s the local Dirichlet matrix on subdomain Ω_s. If Robin boundary conditions are imposed on internal interfaces, the corresponding local matrix is denoted by $\widehat{B}_s$, whereas for Neumann boundary conditions we write $\widetilde{B}_s$.

6.4.1 The DtN Coarse Space

We begin with the Dirichlet-to-Neumann (DtN) coarse space, originally introduced in [75] for elliptic problems and extended to the Helmholtz equation in [23]. This method constructs the coarse space from the low-frequency eigenmodes of a boundary operator defined via local DtN

maps, which are then extended harmonically inside each subdomain. From a broader perspective, the DtN approach can be viewed as an early instance of harmonic-type coarse spaces, and also as a precursor to GenEO-type constructions that build on similar principles of localised spectral analysis.

Let $\Gamma_s = \partial\Omega_s \setminus \partial\Omega$ denote the interface between subdomain Ω_s and its neighbours. Given Dirichlet data v_{Γ_s} on Γ_s, *the local heterogeneous Helmholtz extension* is the solution v of the boundary value problem:

$$\begin{aligned} -\nabla \cdot \big(a(\boldsymbol{x})\nabla v\big) - \omega^2 m(\boldsymbol{x})\, v = 0 \qquad & \text{in } \Omega_s, \\ v = v_{\Gamma_s} \qquad & \text{on } \Gamma_s, \\ \mathcal{C}(v) = 0 \qquad & \text{on } \partial\Omega_s \cap \partial\Omega, \end{aligned}$$

where $\mathcal{C}(v) = 0$ enforces the original problem boundary conditions. The DtN map then takes v_{Γ_s} to its corresponding Neumann data on Γ_s:

$$\mathrm{DtN}_{\Omega_s}(v_{\Gamma_s}) = a(\boldsymbol{x})\, \frac{\partial v}{\partial n}\bigg|_{\Gamma_s}.$$

The DtN eigenproblem for subdomain Ω_s is

$$\mathrm{DtN}_{\Omega_s}(u_{\Gamma_s}) = \lambda\, u_{\Gamma_s}, \tag{6.4.2}$$

where u_{Γ_s} are eigenfunctions with eigenvalues $\lambda \in \mathbb{C}$. The corresponding coarse space is constructed by taking the heterogeneous Helmholtz extension of u_{Γ_s} in Ω_s and extending it by zero across Ω via the partition of unity [23].

For the discrete setting, we introduce local Neumann matrices $\widetilde{A}_s$, where $\mathcal{C} = 0$ on $\partial\Omega_s \cap \partial\Omega$. Let Γ_s and I_s denote boundary and interior indices, respectively. Defining the mass matrix on the interface as $M_{\Gamma_s} = \left(\int_{\Gamma_s} \phi_j \phi_i\right)_{i,j\in\Gamma_s}$ the discrete DtN eigenproblem is given by

$$\left(\widetilde{A}_{\Gamma_s,\Gamma_s} - A_{\Gamma_s,I_s} A_{I_s,I_s}^{-1} A_{I_s,\Gamma_s}\right) \mathbf{u}_{\Gamma_s} = \lambda M_{\Gamma_s} \mathbf{u}_{\Gamma_s}, \tag{6.4.3}$$

where $A_{I_s,\Gamma_s}, A_{\Gamma_s,I_s}$ and $A_{I_s,I_s}, A_{\Gamma_s,\Gamma_s}$ are obvious notations for sub-blocks of matrix A corresponding to the set of indices in the interior I_s and on the boundary Γ_s. Because of the Robin boundary condition on Γ_R, the left-hand-side matrix in (6.4.3) is a priori non-Hermitian. The Helmholtz extension to interior degrees of freedom is obtained via $\mathbf{u}_{I_s} = -A_{I_s,I_s}^{-1} A_{I_s,\Gamma_s} \mathbf{u}_{\Gamma_s}$, and the global coarse basis in Z is formed as $R_s^T D_s \mathbf{u}_s$, where $\mathbf{u}_s$ is the full local vector computed from $\mathbf{u}_{\Gamma_s}$ and $\mathbf{u}_{I_s}$ [23]. We see that the discrete counterpart of (6.4.2) is (6.4.3) consisting in a Schur complement based generalised eigenvalue problem.

A key aspect is the selection of eigenvectors for the coarse space. In [23] a criterion has been identified empirically consisting in retaining the eigenvalues with the smallest real parts, imposing the threshold

$$\Re(\lambda) < \eta_{\max}.$$

While $\eta_{\max} = k_s$, where $k_s := \sup_{\boldsymbol{x}\in\Omega_s} k(\boldsymbol{x})$, was initially proposed, recent work [10, 11] suggests that $\eta_{\max} = k_s^{4/3}$ can improve robustness with respect to the wavenumber. Also in [11] this threshold was systematically studied and empirical k-dependent bounds were derived for the sizes of the corresponding coarse spaces.

Remark 6.4.1. *Unlike for Laplace's equation with self-adjoint boundary conditions where generalised eigenvalue problems are self-adjoint and positive definite, the DtN eigenproblems for Helmholtz are in general non-self-adjoint. As a consequence, there are currently no theoretical convergence guarantees for the DtN coarse space. Nevertheless, recent studies—including [14]*

and [11]—demonstrate that this approach performs remarkably well in practice, and its complexity can be quantified empirically. A significant advantage of this construction is that its coarse space dimension is inherently limited by the number of degrees of freedom on the subdomain interface Γ_s, as the eigenfunctions are defined on Γ_s and only extended into the interior. This allows for global modes with full support to be generated from a relatively compact spectral basis.

6.4.2 GenEO-Type Coarse Spaces: Δ-GenEO and H-GenEO

The Generalised Eigenproblems in the Overlap (GenEO) coarse space, introduced in [85], offers a robust approach for symmetric positive definite problems, even in heterogeneous settings. Note that originally, GenEO was formulated variationally under the assumption that the bilinear form $a_\Omega(\cdot,\cdot)$ is symmetric and coercive, an assumption that is not satisfied here. Given a subdomain Ω_s and its overlapping region Ω_s°, the local eigenproblem is defined as

$$a_{\Omega_s}(u,v) = \lambda a_{\Omega_s^\circ}(\Xi_s(u),\Xi_s(v)) \quad \forall v \in V(\Omega_s),$$

where Ξ_s is the partition of unity operator. Here, $a_{\Omega_s}(\cdot,\cdot)$ represents the variational problem restricted to domain Ω_s with problem boundary conditions on $\partial\Omega$ and natural conditions on $\partial\Omega_s \setminus \partial\Omega$. We can also define $a_{\Omega_s^\circ}(\cdot,\cdot)$ over the overlap region. The smallest eigenvalue λ not included in the GenEO space determines the preconditioned operator's condition number, making thresholding on $\lambda < \lambda_{\max}$ a natural choice. Alternatively, a fixed number of eigenvectors per subdomain can be selected for efficiency. The restriction to Ω_s° is not necessary, and a more practical variant replaces it with Ω_s, simplifying implementation [29]. This leads to the discrete eigenproblem

$$\widetilde{A}_s\mathbf{u} = \lambda D_s A_s D_s \mathbf{u}, \tag{6.4.4}$$

where $A_s = R_s A R_s^T$ is the local Dirichlet matrix. The coarse space basis in Z is then constructed as $R_s^T D_s \mathbf{u}$. This formulation will serve as the foundation for a Helmholtz-specific extension.

Δ-GenEO Coarse Space

Applying GenEO to the Helmholtz problem is challenging due to the loss of self-adjointness and definiteness. While theoretical advancements have been made [13, 12], a fully justified spectral coarse space for Helmholtz remains elusive. One workaround, Δ-GenEO, replaces the Helmholtz operator with the Laplacian (setting $\omega = 0$ in (6.2.1) and (6.2.4)), yielding the eigenproblem

$$\widetilde{L}_s\mathbf{u} = \lambda D_s L_s D_s \mathbf{u},$$

where L_s and $\widetilde{L}_s$ are the local Laplace Dirichlet and Neumann matrices, respectively. This provides a well-posed problem with real non-negative eigenvalues, but its effectiveness deteriorates as k increases since Laplace solutions diverge significantly from those of Helmholtz. For this coarse space theoretical bounds have been rigorously established in [13] and further improved in [26]. These bounds show, as expected, that this coarse space cannot perform very effectively at high frequencies and the size of the coarse space increases dramatically with the frequency.

The H_k-GenEO Coarse Space

To construct a more suitable spectral coarse space for the heterogeneous Helmholtz problem, it is essential to incorporate the Helmholtz operator itself. However, directly applying (6.4.4) with the heterogeneous operator leads to a non-self-adjoint problem with complex eigenvalues, which can cause solver instability. Instead, we propose a hybrid approach that links the indefinite heterogeneous Helmholtz operator to a positive-definite surrogate operator. The local eigenproblem is given by

$$\widetilde{A}_s\mathbf{u} = \lambda\, D_s\, P_s\, D_s \mathbf{u}, \tag{6.4.5}$$

where:

- $\widetilde{A}_s$ is the local Neumann matrix for the heterogeneous Helmholtz operator

$$-\nabla \cdot \big(a(\boldsymbol{x})\nabla u\big) - \omega^2 m(\boldsymbol{x})\, u,$$

- D_s is the local partition-of-unity weighting matrix,
- P_s is a symmetric positive-definite (SPD) matrix chosen to be "close" to A_s in a suitable sense.

A natural choice for P_s is L_s, the discretisation of the Laplace operator $-\Delta$ with heterogeneous coefficient $a(\boldsymbol{x})$. In practice, many SPD choices perform well, but another particularly convenient choice—allowing for simpler theoretical estimates—is $P_s = B_s$, where B_s is the discretisation of the *positive* heterogeneous Helmholtz operator

$$-\nabla \cdot \big(a(\boldsymbol{x})\nabla u\big) + \omega^2 m(\boldsymbol{x})\, u. \tag{6.4.6}$$

The corresponding method is called H_k*-GenEO*, and theoretical estimates have been derived for the indefinite heterogeneous Helmholtz problem in [24]. When Robin conditions are used, the eigenvalues of (6.4.5) are complex but tend to cluster near the real axis, enabling a stable selection criterion based on

$$\Re(\lambda) < \eta_{\max}.$$

A thorough numerical study performed in [11] shows that H_k-GenEO achieves wavenumber-independent GMRES convergence and empirical frequency scaling formulae were derived. A robust and commonly used choice is $\eta_{\max} = \frac{1}{2}$.

Remark 6.4.2. *Unlike the DtN coarse space, where the spectral information is extracted from the subdomain boundary and then harmonically extended into the interior, the GenEO family constructs eigenproblems that are defined* throughout each subdomain*, often including the overlap region. This leads to* potentially larger coarse spaces *since the spectral basis is built over the full subdomain volume.*

Although a full theoretical analysis of the GenEO methodology in the Helmholtz context is still incomplete, rigorous convergence estimates are available for certain variants. In particular, for Δ*-GenEO, such bounds were established in [13] and later refined in [26]. These results were recently extended to the* H_k*-GenEO construction [24]. However, these theoretical results tend to be conservative, often overestimating the coarse space size required in practice. Numerical experiments indicate that both GenEO variants exhibit significantly better performance than what the theory predicts.*

6.4.3 Helmholtz-Harmonic Coarse Spaces

We now describe the class of *Helmholtz-harmonic coarse spaces*, introduced independently in [71, 62] and later in a fully algebraic setting in [74]. These spaces are built from discrete solutions of the *heterogeneous* Helmholtz equation on local subdomains and then assembled into a global conforming approximation space using partition-of-unity weights.

Harmonic GenEO Coarse Space

The underlying philosophy is similar in spirit to the *GenEO family*: a local spectral decomposition is used to select relevant low-energy modes that define the coarse space. However, there is a key distinction: while GenEO formulates a spectral problem using the full operator (e.g. Laplacian or Helmholtz) and searches for dominant directions in the whole local space, *harmonic*

coarse spaces restrict the search space to local Helmholtz-harmonic functions. The spectral problem is then posed over this smaller subspace typically associated with the *positive* heterogeneous Helmholtz operator (6.4.6). This restriction makes the underlying spectral theory easier and the coarse space selection process *real-valued*, while still encoding the Helmholtz physics through the choice of constrained subspace.

To introduce the coarse spaces described in [71, 62], we begin by defining the *space of local Helmholtz solutions* on a subdomain Ω_s:

$$W(\Omega_s) := \{u \in V(\Omega_s) : \ a_{\Omega_s}(u, v) = 0, \ \forall v \in V_0(\Omega_s)\},$$

where $V(\Omega_s)$ and $V_0(\Omega_s)$ are defined in (6.2.2). The sesquilinear form $a_{\Omega_s}(\cdot, \cdot)$ corresponds to the Helmholtz operator defined in (6.2.4). Thus, $W(\Omega_s)$ consists of heterogeneous Helmholtz solutions with the appropriate boundary conditions.

To define a spectral selection on this space, we introduce the *positive heterogeneous Helmholtz sesquilinear form*:

$$b_{\Omega_s}(u, v) = \int_{\Omega_s} \left(a(\boldsymbol{x})\nabla u \cdot \nabla \bar{v} + \omega^2 m(\boldsymbol{x})\, u\bar{v}\right) \mathrm{d}\boldsymbol{x}.$$

We then solve the following eigenproblem (see [62, eq. (3.6)] and [71, eq. (2.20)]):

$$\text{find } u \in W(\Omega_s) \quad \text{such that} \quad b_{\Omega_s}(u, v) = \lambda\, b_{\Omega_s}(\Xi_s(u), \Xi_s(v)) \quad \forall v \in W(\Omega_s),$$

where Ξ_s is the partition-of-unity operator associated with the subdomain Ω_s. This problem is *self-adjoint and coercive*, and all eigenvalues λ are real and positive.

To handle the Helmholtz-harmonicity constraint in practice, the eigenvalue problem is recast as a saddle point problem [71, eq. (2.27)]: find $u \in V(\Omega_s)$, $p \in V_0(\Omega_s)$, $p = 0$ on $\partial\Omega_s \setminus \partial\Omega$ such that

$$\begin{aligned} b_{\Omega_s}(u, v) + \overline{a_{\Omega_s}(v, p)} &= \lambda b_{\Omega_s}(\Xi_s(u), \Xi_s(v)), && \forall v \in V(\Omega_s), \\ a_{\Omega_s}(u, q) &= 0, && \forall q \in V_0(\Omega_s), \\ p &= 0, && \text{on } \partial\Omega_s \setminus \partial\Omega. \end{aligned}$$

In matrix notations, the eigenproblem system is: find $\mathbf{u}$ and $\mathbf{p}$ such that

$$\begin{bmatrix} \widetilde{B}_s & A_{s,0}^* \\ A_{s,0} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \lambda \begin{bmatrix} D_s B_s D_s & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{p} \end{bmatrix},$$

where:

- $A_s = R_s A R_s^T$ is the local Dirichlet matrix associated with the heterogeneous Helmholtz sesquilinear form a_{Ω_s},
- $A_{s,0}$ is the restriction of A_s to rows not associated with boundary nodes,
- $B_s = R_s B R_s^T$ is the local Dirichlet matrix associated with the positive Helmholtz bilinear form b_{Ω_s},
- $\widetilde{B}_s$ is its Neumann counterpart,
- D_s is the partition of unity diagonal weight matrix.

For simplicity, the system above was written for homogeneous Dirichlet boundary conditions $u_{\Gamma_D} = 0$.

After solving the eigenproblem, the *coarse space basis* is formed by selecting all eigenvectors $\mathbf{u}$ associated with eigenvalues $\lambda < \tau$, where $\tau > 0$ is a user-defined threshold. These are then extended to the global space via:

$$Z = \left\{R_s^T D_s \mathbf{u}\right\}.$$

Each basis function is thus globally supported (via the partition of unity), locally Helmholtz-harmonic (by construction), and selected through a stable positive definite spectral problem.

Extended Harmonic GenEO Coarse Space

Building on the construction of Helmholtz-harmonic coarse spaces described above, we now present a fully algebraic variant introduced in [74]. This approach retains the same core philosophy—constructing coarse functions from locally Helmholtz-harmonic components—but introduces an *extended overlapping decomposition*. Contrary to the previous variant, the discrete Helmholtz-harmonic constraint is not enforced on the eigenvectors, but rather in the construction of the coarse space matrix Z thanks to a suitable operator (also present in the eigenproblem).

Specifically, for each original subdomain Ω_s, an enlarged subdomain $\check{\Omega}_s$ is defined by adding at least one layer of mesh elements along its interfaces with neighbouring subdomains. This auxiliary decomposition $\{\check{\Omega}_s\}_{s=1}^N$ is used solely for the construction of the coarse space and does *not* affect the first-level (fine-scale) preconditioner.

All quantities associated with the extended subdomains are denoted with a superscript $\check{\cdot}$. For example, $\check{R}_s$ is the restriction matrix to $\check{\Omega}_s$, and $\check{A}_s = \check{R}_s A \check{R}_s^T$ is the local Dirichlet matrix on $\check{\Omega}_s$.

The local eigenproblem reads: find $\check{\mathbf{u}}$ such that

$$\check{\widetilde{B}}_s \check{\mathbf{u}} = \lambda \left(R_s \check{R}_s^T - \widehat{A}_s^{-1} R_s \check{R}_s^T \check{A}_s \right)^\dagger D_s B_s D_s \left(R_s \check{R}_s^T - \widehat{A}_s^{-1} R_s \check{R}_s^T \check{A}_s \right) \check{\mathbf{u}},$$

where:

- $\check{A}_s = \check{R}_s A \check{R}_s^T$ is the local Dirichlet matrix associated with the heterogeneous Helmholtz sesquilinear form $a_{\check{\Omega}_s}$ on the enlarged subdomain $\check{\Omega}_s$,
- $B_s = R_s B R_s^T$ is the local Dirichlet matrix associated with the positive Helmholtz bilinear form b_{Ω_s},
- $\check{\widetilde{B}}_s$ is the local Neumann matrix associated with the positive Helmholtz bilinear form $b_{\check{\Omega}_s}$,
- $\widehat{A}_s^{-1}$ corresponds to the local Robin problem on Ω_s.

This eigenproblem is written under the assumption of no Dirichlet boundary conditions ($\Gamma_D = \emptyset$) for simplicity, but can accommodate them as long as the action of $\widehat{A}_s^{-1} R_s \check{R}_s^T \check{A}_s$ yields a solution satisfying the problem boundary conditions.

By construction, this eigenproblem is *self-adjoint and coercive*, so all eigenvalues λ are real and positive. Moreover, while B_s and $\check{\widetilde{B}}_s$ are typically associated with the positive Helmholtz matrix, other SPD matrices could also be used in their place. We point out that the matrix on the right-hand side does not have to be assembled in practice to solve the eigenproblems; it is enough to know the matrix-vector product.

The coarse space basis is formed by collecting all vectors of the form

$$Z = \left\{ R_s^T D_s \left(R_s \check{R}_s^T - \widehat{A}_s^{-1} R_s \check{R}_s^T \check{A}_s \right) \check{\mathbf{u}} \ : \ \lambda < \tau \right\}$$

for some user-defined threshold $\tau > 0$. By construction, each resulting vector lies in the global finite element space and corresponds to a *local discrete Helmholtz solution* on Ω_s thanks to the presence of the operator $R_s \check{R}_s^T - \widehat{A}_s^{-1} R_s \check{R}_s^T \check{A}_s$, and multiplied by the partition of unity D_s.

6.4.4 Summary and Comparison of Spectral Coarse Spaces

Having described the main families of spectral coarse spaces—namely the DtN, GenEO-type, and Helmholtz-harmonic constructions—we now highlight their connections and comparative features. A summary of this comparison is given in Table 6.1.

Conceptual links Despite differing technical constructions, all approaches share a common spectral philosophy: they extract low-energy components via localised eigenproblems, then assemble them into globally supported basis functions using a partition of unity.

The **DtN** coarse space is the most geometrically minimal, using interface-local eigenfunctions extended via Helmholtz harmonic continuation. In this sense, it belongs to the broader class of **Helmholtz-harmonic coarse spaces**. It may also be seen as a precursor to the **GenEO** family, which replaces boundary-local eigenproblems with volumetric ones defined over subdomains or overlapping regions. The **harmonic GenEO** spaces, in turn, reinterpret the GenEO framework by restricting the usual self-adjoint positive definite eigenproblems to the subspace of local Helmholtz solutions. The constraint to solve in the Helmholtz-harmonic subspace can be imposed using an augmented eigenproblem. Finally, the **extended-harmonic** coarse spaces involve self-adjoint positive definite eigenproblems similarly to the GenEO family, but posed on an extended subdomain and with a suitable operator weighting the matrix in the eigenproblem. The resulting eigenvectors are projected in the Helmholtz-harmonic subspace of the (non-extended) subdomain.

Complexity comparison The main costs of two-level domain decomposition preconditioners are: (i) construction of the coarse space (local eigenproblems and global assembly), (ii) factorisation of the coarse operator, and (iii) repeated local and coarse solves during GMRES iterations.

For discretisations with a fixed number of points per wavelength ($hk = \mathcal{O}(1)$) as often happens in engineering practice, the number of unknowns per subdomain scales like $\mathcal{O}(k^d)$ in d dimensions, while the number of unknowns on the subdomain boundary scales like $\mathcal{O}(k^{d-1})$. As a result the cost of constructing the coarse space and its dimension differ between the different methods:

- **DtN:** the eigenproblem is posed on the interface (although the operator features a local volume inverse of size $\mathcal{O}(k^d)$) and as a result, the number of basis functions is bounded by the number of interface unknowns in $\mathcal{O}(k^{d-1})$.
- **GenEO-type:** the eigenproblem is volumetric hence leads to $\mathcal{O}(k^d)$ candidate functions per subdomain. To ensure wavenumber-independent convergence, the coarse operator (although well below this limit in practice) can become large, leading to an important factorisation cost.
- **Harmonic:** The saddle-point eigenproblem is twice the number of unknowns in the subdomain, hence of size $\mathcal{O}(k^d)$, but the retained modes need to solve the Helmholtz problem, hence their number is necessarily bounded by the number of boundary unknowns in $\mathcal{O}(k^{d-1})$.
- **Extended harmonic:** The eigenproblem is posed in the extended volume, still in $\mathcal{O}(k^d)$, but the elements in the coarse space are Helmholtz solutions by construction (up to the partition of unity), hence their number is also necessarily bounded by the number of boundary unknowns in $\mathcal{O}(k^{d-1})$.

At runtime, each GMRES iteration requires local subdomain solves (local matrices are factored in the setup phase) typically via sparse direct solvers and a global coarse correction ($\mathcal{O}(N_c^2)$ after factorisation). In principle, these scalings suggest coarse space dimensions increase dramatically with the wavenumber, but in practice our numerical results show milder growth, with iteration counts nearly independent of k once a suitable coarse space is included.

Rationale for numerical comparison Given the subtle theoretical distinctions and the absence of definitive complexity bounds in the indefinite Helmholtz regime, it is difficult to draw

Table 6.1: Comparison of spectral coarse space families for the Helmholtz problem

Property	**DtN**	**GenEO-type**	**Harmonic**	**Extended harmonic**
Eigenfunctions	Interface Γ_s	Subdomain Ω_s (or overlap)	Subdomain Ω_s	Extended subdomain $\check{\Omega}_s$
Spectral domain	Interface	Full volume	Helmholtz-harmonic subspace	Full extended volume
Eigenvalues	complex	complex	real positive	real positive
Maximum dimension	$\mathcal{O}(\lvert\Gamma_s\rvert)$	$\mathcal{O}(\lvert\Omega_s\rvert)$	$\mathcal{O}(\lvert\Gamma_s\rvert)$	$\mathcal{O}(\lvert\Gamma_s\rvert)$
Theory	None (empirically)	Available for some variants	Available	Available
Adapted to high k	Yes	Yes (H_k-GenEO)	Yes	Yes
Notable references	[23]	[85], [13], [26], [24]	[62], [71]	[74]

conclusions purely from the analytical form of the coarse spaces. The actual performance depends on practical implementation details, the choice of eigenvalue threshold, and the frequency regime. Therefore, a numerical comparison is essential to:

- assess the robustness of each coarse space across a range of frequencies and heterogeneous media;
- quantify the trade-off between coarse space dimension and solver convergence;
- validate whether the observed empirical behaviour aligns with or diverges from existing theoretical predictions.

In the next section, we provide such a numerical study, focusing on both homogeneous and heterogeneous Helmholtz problems with increasing complexity.

6.5 Numerical Assessment and Comparison

The purpose of this section is to evaluate the performance and robustness of the spectral coarse spaces introduced in Section 4 in realistic simulation settings. While these methods differ in their mathematical formulation and spectral construction, it is often difficult to draw firm conclusions about their practical efficiency based solely on theoretical arguments. In particular, the coarse space dimension, conditioning effects, and preconditioner performance depend intricately on the frequency, domain geometry, heterogeneity, and the choice of eigenvalue threshold.

To overcome this, we conduct a series of numerical experiments designed to probe the behaviour of each coarse space under increasing complexity evaluated within a common two-level ORAS preconditioning framework, using GMRES as the iterative solver. In all cases, the GMRES solver is not restarted and the tolerance is 10^{-6} on the residual, with a right-preconditioning. The overlap is minimal (unless stated otherwise) with a symmetry constraint with respect to the interface, implying two layers of cells in the overlap. Each test case is analysed through a combination of strong and weak scaling experiments, studying the influence of:

- subdomain diameter (in wavelength),
- coarse space size (global and per subdomain),
- eigenvalue threshold τ,
- and overlap and partition of unity choices.

Our goal is to assess both efficiency (number of GMRES iterations) and scalability (coarse space dimension and distribution), and to offer a fair comparison between the different coarse space strategies under practically relevant conditions.

Implementation details The FreeFEM implementation of the benchmarks is done within the ffddm framework, a set of parallel FreeFEM scripts implementing Schwarz domain decomposition methods. The ffddm documentation is available on the FreeFem.org web page (see [91]). The ffddm framework relies on Message Passing Interface (MPI) parallelism. As is usually done in domain decomposition methods, we assign one subdomain per MPI process.

The first step is to decompose the computational domain into overlapping subdomains. Unless otherwise stated, the automatic graph partitioner Metis [66] is used, which produces a non-overlapping decomposition of the set of mesh elements while minimising interfaces between subdomains and conserving good load-balancing. The overlapping decomposition is then obtained by adding successive layers of elements to reach the desired size of overlap. The setup of the one-level preconditioner then consists in assembling and factorising the local matrices in each subdomain in parallel. This is done using the sparse direct solver MUMPS [1]. The two-level preconditioner for each method is then built by first solving the corresponding local eigenvalue problems in each subdomain with SLEPc [61], and finally assembling and factorising the coarse space operator in a distributed manner by MUMPS using a few cores (ranging from 1 to 144 depending on the size of the coarse space).

During the solution phase, each application of the preconditioner involves solving linear systems with local subdomain matrices (first level) and with the coarse space matrix (second level), which is done by forward-backward substitution using the factorisations computed during the setup phase.

A FreeFEM script comparing all methods for the homogeneous Helmholtz equation in a square is available at `https://github.com/FreeFem/FreeFem-sources/blob/develop/examples/ffddm/Helmholtz-2d-all.edp`. This was run on an Intel i7-8700 CPU with 64GB RAM running the Windows 11 operating system.

6.5.1 Numerical Simulations in a Square: Homogeneous and Heterogeneous Problems

The numerical tests in this section are structured as follows:

1. **Homogeneous test case.** A baseline benchmark on a unit square domain with Robin boundary conditions and a single interior point source. This allows us to measure the scaling of coarse space dimension and iteration count with respect to the number of subdomains and frequency, in a basic controlled setting.

2. **Heterogeneous test case.** A layered medium with strong contrasts in wave speed is introduced to study robustness under medium heterogeneity and its impact on coarse space size and convergence.

Homogeneous Problem

In this part we base the numerical experiments on the model problem in 2D, defined on the unit square $\Omega = (0,1)^2$. We impose Robin boundary conditions on all sides of the domain. A point source is located in the centre of the domain at $(\frac{1}{2}, \frac{1}{2})$ and provides the forcing function f. The point source is numerically modelled by a Gaussian function: $f(x,y) = 10^4 \exp(-10^3[(x-\frac{1}{2})^2 + (y-\frac{1}{2})^2])$. A schematic of this model problem is found in Figure 6.1.

To discretise the problem, we triangulate Ω using a Cartesian grid with spacing h and alternating diagonals to form a simplicial mesh. We consider constant coefficients $a(\boldsymbol{x}) \equiv 1$ and $m(\boldsymbol{x}) \equiv 1$. The local wavenumber $k = \omega\sqrt{\frac{m}{a}} = \omega$ is then constant and the wavelength used to measure geometrical parameters is $\lambda_m = 2\pi/k$. The discrete problem (6.2.5) is assembled using a $\mathbb{P}_2$ finite element approximation on this mesh. To mitigate the *pollution effect*, we choose the angular frequency ω (or equivalently the wavenumber k) and the mesh size h simultaneously so that the dimensionless quantity kh remains sufficiently small. In practice, this is enforced

by fixing a minimum number of grid points per wavelength λ_m; here, we ensure that $kh \lesssim 0.5$, which corresponds to at least 10 points per wavelength λ_m.

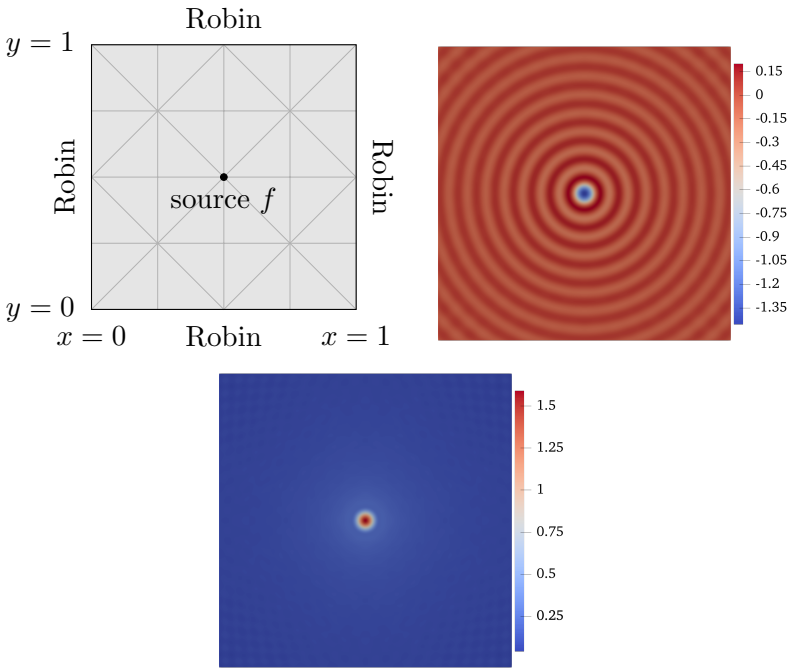


Figure 6.1: Schematic of the 2D wave guide model problem with example triangular mesh (top left). Real part (top right) and modulus (bottom) of the total field for the homogeneous media test case with $\omega = 100$.

Heterogeneous Problem

The next numerical experiments aim to study the robustness of the method with respect to medium heterogeneity. For this purpose, we consider a layered medium in the unit square. More precisely, we use an *alternating layer* configuration (see Figure 6.4) in which the heterogeneity is introduced through the material parameter $m(\boldsymbol{x})$. In the present setting, we keep $a(\boldsymbol{x}) \equiv 1$ everywhere and the mass coefficient is defined as $m(\boldsymbol{x}) = c(\boldsymbol{x})^{-2}$, where $c(\boldsymbol{x})$ is the spatially varying wave speed. The medium alternates between two constant wave speeds: $c(\boldsymbol{x}) \in \{1,\ \rho\}$, where $\rho > 1$ is a contrast parameter controlling the strength of the heterogeneity. The local wavenumber is then given by $k(\boldsymbol{x}) = \omega\sqrt{\frac{m(\boldsymbol{x})}{a(\boldsymbol{x})}} = \frac{\omega}{c(\boldsymbol{x})}$, with $\omega > 0$ the angular frequency. In our numerical experiments, we vary ω while keeping ρ fixed unless otherwise stated. Geometric parameters are measured using the minimal wavelength $\lambda_m = 2\pi/\omega$ corresponding to $c(\boldsymbol{x}) = 1$. An example of a numerical solution for the heterogeneous layered medium is shown in Figure 6.4.

Results

The full results are provided in Tables 6.2 and 6.3. The entries are the diameter of the domain L measured in the smallest wavelength $\lambda_m = 2\pi \min_{\boldsymbol{x}}(c(\boldsymbol{x}))/\omega$ present in the problem, the total number of degrees of freedom n, the averaged diameter of a subdomain H measured in wavelength λ_m, the averaged number n_s of degrees of freedom in a subdomain, the averaged number $n_s^{\partial\Omega_s}$ of degrees of freedom on the boundary of a subdomain, the number N of subdomains, the total size of the coarse space (CS), the averaged number of contributions of a subdomain to the coarse space (CS_s) and the number of GMRES iterations to reach the tolerance (It). The methods being compared are the baseline one-level method (1lvl): extended (ext) and har-

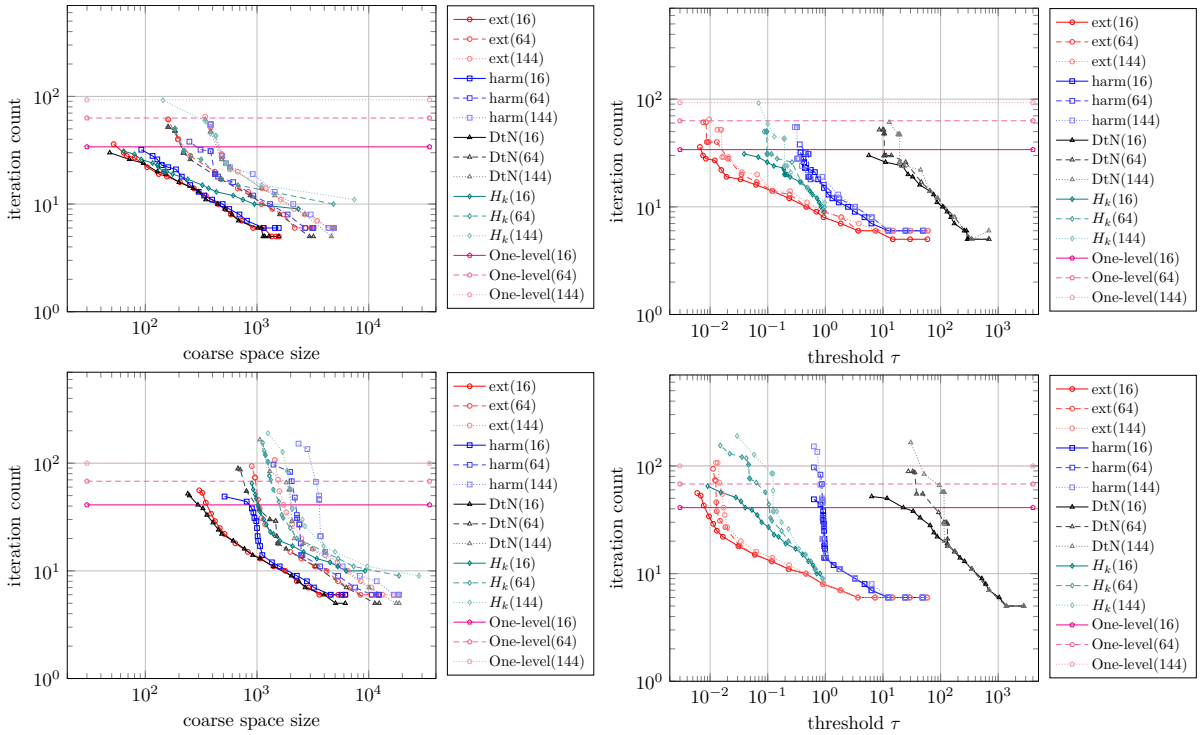


Figure 6.2: Influence of the coarse space size (left) and threshold choice (right) on the iteration count in strong scaling for the homogeneous media test case with $\omega = 20$ (top) and $\omega = 100$ (bottom). The number in brackets indicates the number of subdomains.

monic coarse spaces (harm) on one side with the classical the Dirichlet-to-Neumann (DtN) and H_k-GenEO methods (H_k) on the other side.

Besides, we report, for the smallest and largest frequencies tested, the number of GMRES iterations to reach the required tolerance as a function of the total coarse space size and choice of threshold parameter in Figures 6.2 and 6.5. The number of subdomains is provided in brackets in the legend entries. The results reported in the table correspond to the parameters with the minimum number of iterations obtained for each method.

The number of iterations of GMRES to reach the required tolerance as a function of the subdomain diameter (measured in wavelength) is reported Figures 6.3 and 6.6.

Assessment and Comparative Conclusions for the Square Test Cases

We now draw conclusions from the strong scaling tests performed on the 2D square homogeneous and heterogeneous test cases. From both Tables 6.2–6.3 and Figures 6.2–6.6, several trends emerge consistently across frequencies and levels of heterogeneity. General comments on one-level and two-level methods are:

- The performance of the one-level method deteriorates with increasing frequency or with heterogeneities.
- In contrast, all two-level methods are scalable and very robust with respect to the frequency or in the presence of heterogeneity. As a result, they dramatically outperform the one-level baseline.
- These properties are however achieved only for large enough coarse spaces, as the iteration count of two-level methods correlates strongly with coarse space size.
- In particular, robustness is achieved with a coarse space size that increases with increasing number of subdomains and increasing frequency.

						1lvl	ext			harm			DtN			H_k		
L $[\lambda_m]$	N	n	H $[\lambda_m]$	n_s	$n_s^{\partial\Omega_s}$	It	It	CS	CS_s	It	CS	CS_s	It	CS	CS_s	It	CS	CS_s
4.5	16	231361	1.1	15373	490	34	5	1352	84	6	1148	72	5	1132	71	9	2352	147
4.5	36	231361	0.8	7108	331	47	6	1528	42	6	1904	53	5	2102	58	10	3576	99
4.5	64	231361	0.6	4156	251	63	6	2176	34	6	2690	42	5	2946	46	10	4832	76
4.5	100	231361	0.5	2762	203	77	6	2784	28	6	3504	35	5	3802	38	11	6094	61
4.5	144	231361	0.4	1990	171	93	6	4368	30	6	4366	30	5	4636	32	11	7440	52
9.0	16	519841	2.3	33853	730	36	6	1360	85	6	1724	108	5	1906	119	9	3504	219
9.0	36	519841	1.5	15455	491	48	6	2298	64	6	2864	80	5	3116	87	10	5304	147
9.0	64	519841	1.1	8926	371	63	6	3200	50	6	3996	62	5	4342	68	10	7136	112
9.0	100	519841	0.9	5863	299	73	6	5140	51	6	5140	51	5	5572	56	10	8962	90
9.0	144	519841	0.8	4177	251	92	6	6480	45	6	6478	45	5	6834	47	10	10896	76
13.5	16	1442401	3.4	92413	1210	39	5	3346	209	6	2856	178	5	3156	197	10	3316	207
13.5	36	1442401	2.3	41748	811	52	6	4748	132	7	4768	132	6	4182	116	10	8760	243
13.5	64	1442401	1.7	23866	611	67	6	5316	83	6	6684	104	5	7164	112	10	11744	184
13.5	100	1442401	1.4	15520	491	82	6	6874	69	6	8576	86	5	9148	91	10	14760	148
13.5	144	1442401	1.1	10950	411	97	6	8448	59	6	10460	73	5	11186	78	10	17808	124
18.0	16	2076481	4.5	132493	1450	39	6	3404	213	6	4288	268	7	3782	236	10	6960	435
18.0	36	2076481	3.0	59695	971	51	6	5676	158	7	5728	159	6	5020	139	10	10488	291
18.0	64	2076481	2.3	34036	731	67	6	8000	125	7	8028	125	6	7044	110	9	14048	220
18.0	100	2076481	1.8	22077	587	81	6	8180	82	6	10344	103	5	10950	110	10	17596	176
18.0	144	2076481	1.5	15537	491	98	6	10074	70	6	13992	97	6	13340	93	10	21264	148
22.5	16	3690241	5.6	234253	1930	41	6	3616	226	6	4570	286	5	5042	315	10	6280	392
22.5	36	3690241	3.8	105188	1291	54	6	6036	168	6	7612	211	5	8198	228	9	13944	387
22.5	64	3690241	2.8	59776	971	68	6	8452	132	6	10692	167	5	11358	177	9	18656	292
22.5	100	3690241	2.3	38647	779	84	6	14940	149	7	13778	138	6	14526	145	9	23362	234
22.5	144	3690241	1.9	27110	651	100	6	13344	93	6	16796	117	5	17694	123	9	28176	196

Table 6.2: Strong scaling experiment for the homogeneous media test case.

Comparing the different two-level methods, we remark that:

- The smallest number of iterations is achieved with the DtN coarse spaces, closely followed by the harmonic and extended-harmonic coarse spaces, and then by the H_k-GenEO coarse spaces.
- To achieve this, the DtN coarse spaces are however slightly larger than the harmonic and extended-harmonic ones, while the H_k-GenEO coarse spaces are the largest.
- In a context with a tight coarse space size budget, the DtN and extended-harmonic coarse spaces are the most efficient, especially at large frequencies.
- A sensible choice for the eigenvalue threshold τ to ensure low iteration count is in the range $5 - 10$ for extended-harmonic, around 10 for harmonic, in the range $500 - 1000$ for DtN and in the range $0.8 - 1$ for H_k-GenEO.

6.5.2 A Test Case from Medical Imaging

Description of the problem We consider the problem of plane wave scattering by randomly positioned, penetrable micro-reflectors in two dimensions. This setting models acoustic wave interaction with soft biological tissues, as explored in recent quantitative ultrasound imaging studies [49, 50]. We fix a constant *reference medium* with coefficients $a_0 > 0$, $m_0 > 0$, and define the *reference wave speed* and the corresponding *reference wavenumber* and *reference wavelength*

$$c_0 = \sqrt{\frac{a_0}{m_0}}, \quad k_0 = \omega\sqrt{\frac{m_0}{a_0}} = \frac{\omega}{c_0}, \quad \lambda_0 = \frac{2\pi}{k_0}.$$

All geometric parameters in this test case will be expressed in units of λ_0.

The incident wave is a plane wave of the form $u_i(\mathbf{x}) = \exp(\imath\, k_0\, \mathbf{d}\cdot\mathbf{x})$, $\mathbf{d} \in \mathbb{R}^2$, $|\mathbf{d}| = 1$, where $\mathbf{d}$ is the propagation direction. The total field u satisfies the heterogeneous Helmholtz equation

$$-\nabla\cdot\big(a(\mathbf{x})\nabla u\big) - \omega^2 m(\mathbf{x})\, u = 0, \quad \text{in } \mathbb{R}^2,$$

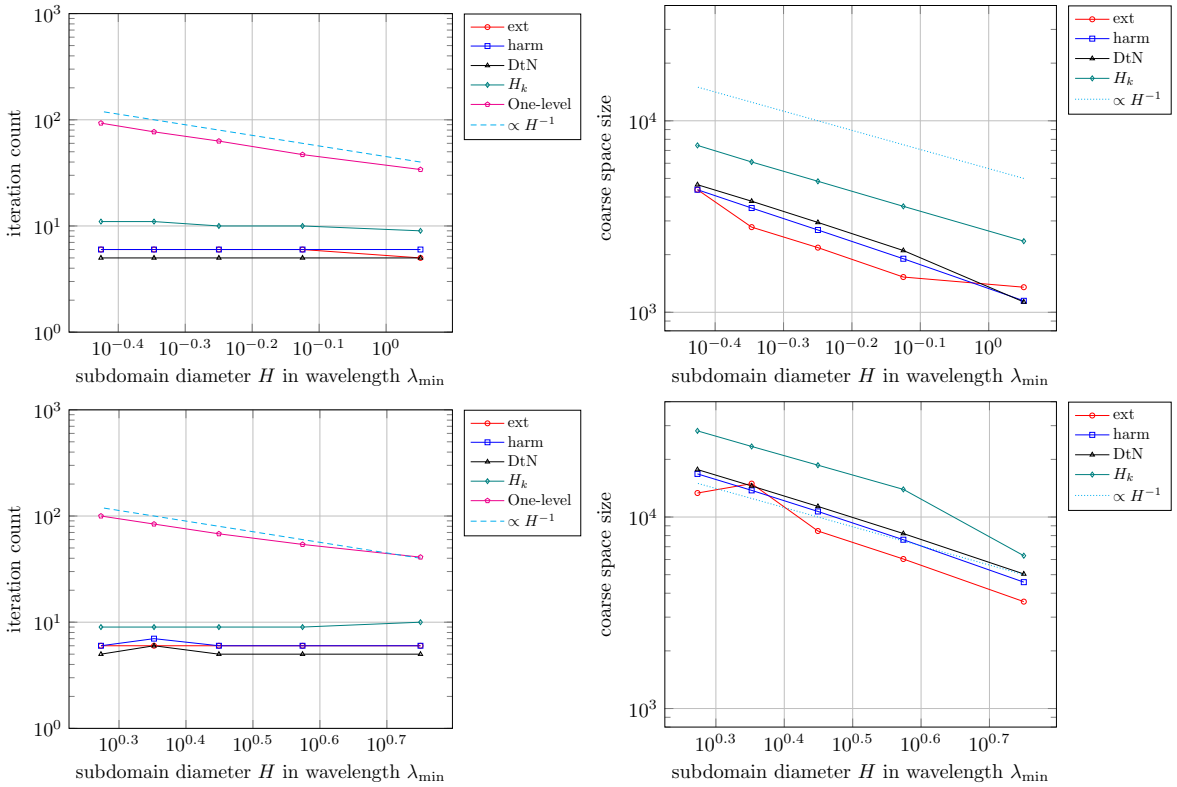


Figure 6.3: Influence of the subdomain diameter on the iteration count (left) and coarse space size (right) in strong scaling for the homogeneous media test case $\rho = 10$ with $\omega = 20$ (top) and $\omega = 100$ (bottom).

with an outgoing radiation condition on the scattered field $u_s = u - u_i$. The coefficients $a(\mathbf{x})$ and $m(\mathbf{x})$ describe the $N_r \in \mathbb{N}$ penetrable micro-reflectors:

$$a(\mathbf{x}) = \mathbf{1}_{\Omega \setminus \bigcup_{j=1}^{N_r} B_\epsilon(\mathbf{x}_j)} + \sum_{j=1}^{N_r} a_j \, \mathbf{1}_{B_\epsilon(\mathbf{x}_j)}, \quad c_0^2 m(\mathbf{x}) = \mathbf{1}_{\Omega \setminus \bigcup_{j=1}^{N_r} B_\epsilon(\mathbf{x}_j)} + \sum_{j=1}^{N_r} m_j \, \mathbf{1}_{B_\epsilon(\mathbf{x}_j)},$$

where $B_\epsilon(\mathbf{x}_j)$ is a disk of radius $\epsilon = \lambda_0/4$ centred at $\mathbf{x}_j$, a_j and m_j are drawn from a normal distribution with mean 1 and standard deviation $\sigma = 0.1$, the disk centres $\mathbf{x}_j$ are generated by a Matèrn hardcore spatial process. In the background medium (a_0, m_0) we have $k(\mathbf{x}) \equiv k_0$, while inside each inclusion $k(\mathbf{x})$ deviates due to the perturbed material parameters.

We truncate the unbounded domain using a perfectly matched layer (PML) following [45]. Let $\chi \in C^\infty_{\text{comp}}(\mathbb{R}^2)$ be a smooth cut-off function supported around the heterogeneous region that is identically one on an open connected set including the support of $(1 - m)(1 - a)$. We introduce the compactly supported unknown

$$\widetilde{u} := u - (1 - \chi) u_i,$$

which satisfies

$$-\nabla \cdot \big(a(\mathbf{x}) \nabla \widetilde{u}\big) - \omega^2 m(\mathbf{x})\, \widetilde{u} = 2\, \nabla \chi \cdot \nabla u_i + u_i\, \Delta \chi, \quad \text{in } \mathbb{R}^2.$$

We use a PML implementation adapted from [7].

A snapshot of the medium and computed fields is shown in Figure 6.7 for $N_r = 2040$ reflectors in a domain of size $47.8\lambda_0 \times 47.8\lambda_0$. The presence of the micro-reflectors imposes a strong constraint on the mesh size corresponding to about 22 points per wavelength λ_0. We use $\mathbb{P}_2$ Lagrange finite elements.

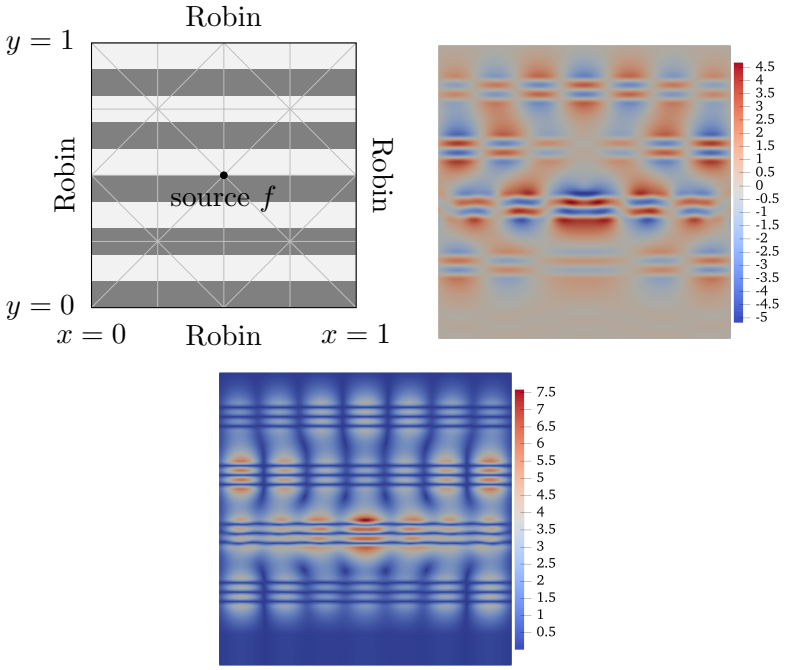


Figure 6.4: Piecewise constant layer profiles for the wave speed $c(\boldsymbol{x})$ (top left). For the darkest shade, $c(\boldsymbol{x}) = 1$, while for the lightest shade, $c(\boldsymbol{x}) = \rho$, with ρ being the contrast factor. Real part (top right) and modulus (bottom) of the total field for the heterogeneous media test case: $\rho = 10$ and $\omega = 100$.

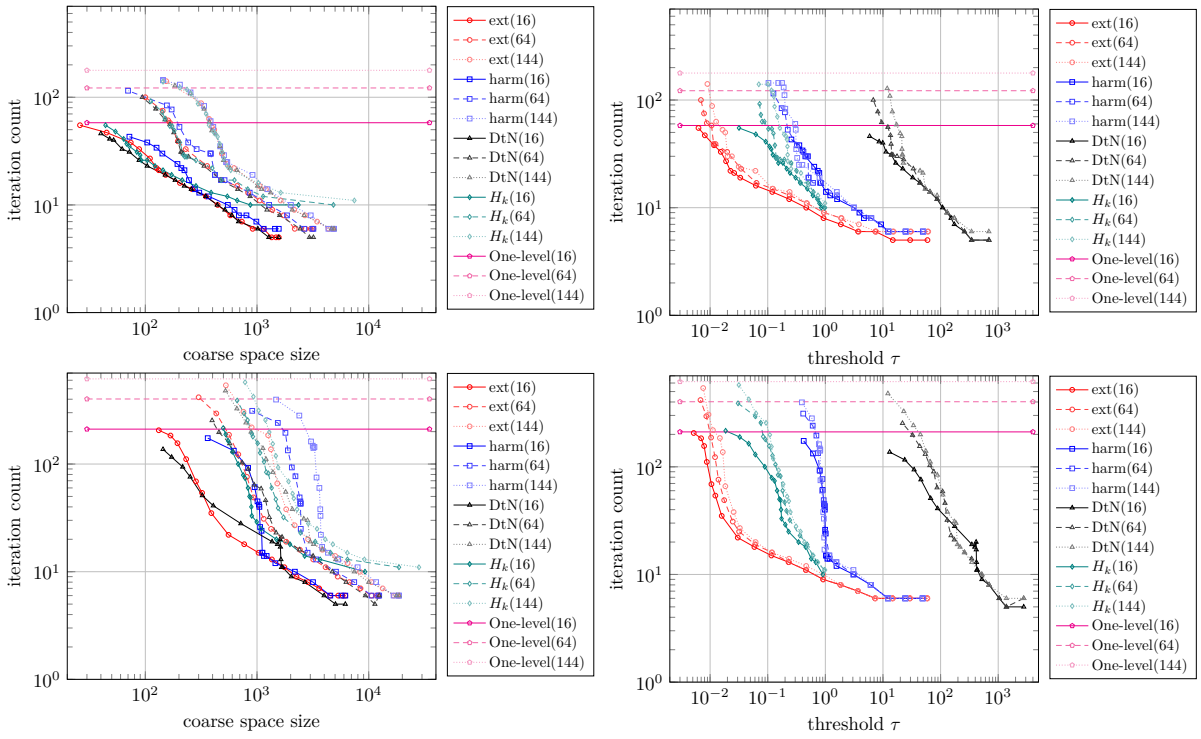


Figure 6.5: Influence of the coarse space size (left) and threshold choice (right) on the iteration count in strong scaling for the heterogeneous media test case $\rho = 10$ with $\omega = 20$ (top) and $\omega = 100$ (bottom).

						1lvl	ext			harm			DtN			H_k		
L $[\lambda_m]$	N	n	H $[\lambda_m]$	n_s	$n_s^{\partial\Omega_s}$	It	It	CS	CS_s	It	CS	CS_s	It	CS	CS_s	It	CS	CS_s
4.5	16	231361	1.1	15373	490	58	5	1352	84	6	1148	72	5	1276	80	10	872	54
4.5	36	231361	0.8	7108	331	94	6	1528	42	6	1904	53	5	2100	58	10	3576	99
4.5	64	231361	0.6	4156	251	122	6	2176	34	6	2652	41	5	2946	46	10	4832	76
4.5	100	231361	0.5	2762	203	154	6	3500	35	6	3504	35	5	3802	38	11	6094	61
4.5	144	231361	0.4	1990	171	178	6	4368	30	6	4356	30	6	4636	32	11	7440	52
9.0	16	519841	2.3	33853	730	119	6	1710	107	6	1712	107	5	1906	119	10	3504	219
9.0	36	519841	1.5	15455	491	178	6	2860	79	6	2864	80	6	2538	70	11	5304	147
9.0	64	519841	1.1	8926	371	231	6	4032	63	6	4644	73	6	3572	56	11	7136	112
9.0	100	519841	0.9	5863	299	294	6	5140	51	6	5924	59	6	5572	56	11	8962	90
9.0	144	519841	0.8	4177	251	327	6	6409	45	6	7137	50	6	5764	40	12	10896	76
13.5	16	1442401	3.4	92413	1210	84	5	3345	209	6	2838	177	5	3154	197	10	5808	363
13.5	36	1442401	2.3	41748	811	117	5	5372	149	6	4732	131	6	4177	116	10	8760	243
13.5	64	1442401	1.7	23866	611	157	7	6657	104	8	6660	104	13	3956	62	10	11744	184
13.5	100	1442401	1.4	15520	491	205	6	6848	68	6	8560	86	5	9148	91	10	14760	148
13.5	144	1442401	1.1	10950	411	198	6	8388	58	6	10460	73	5	11186	78	10	17808	124
18.0	16	2076481	4.5	132493	1450	120	6	3395	212	6	3404	213	8	3778	236	10	6960	435
18.0	36	2076481	3.0	59695	971	160	6	5676	158	7	5676	158	9	6165	171	10	10488	291
18.0	64	2076481	2.3	34036	731	218	6	7964	124	7	7946	124	11	8562	134	10	14048	220
18.0	100	2076481	1.8	22077	587	294	6	10180	102	6	10196	102	5	10948	109	10	17596	176
18.0	144	2076481	1.5	15537	491	288	6	12561	87	6	13992	97	6	13340	93	10	21264	148
22.5	16	3690241	5.6	234253	1930	211	6	4520	282	6	4532	283	5	5039	315	10	9264	579
22.5	36	3690241	3.8	105188	1291	306	6	6014	167	6	7544	210	5	8198	228	10	13944	387
22.5	64	3690241	2.8	59776	971	402	6	10592	166	6	10592	166	5	11357	177	11	18656	292
22.5	100	3690241	2.3	38647	779	553	9	10826	108	8	15407	154	7	14526	145	10	23362	234
22.5	144	3690241	1.9	27110	651	618	6	16754	116	6	16704	116	6	17693	123	11	28176	196

Table 6.3: Strong scaling experiment for the heterogeneous media test case $\rho = 10$.

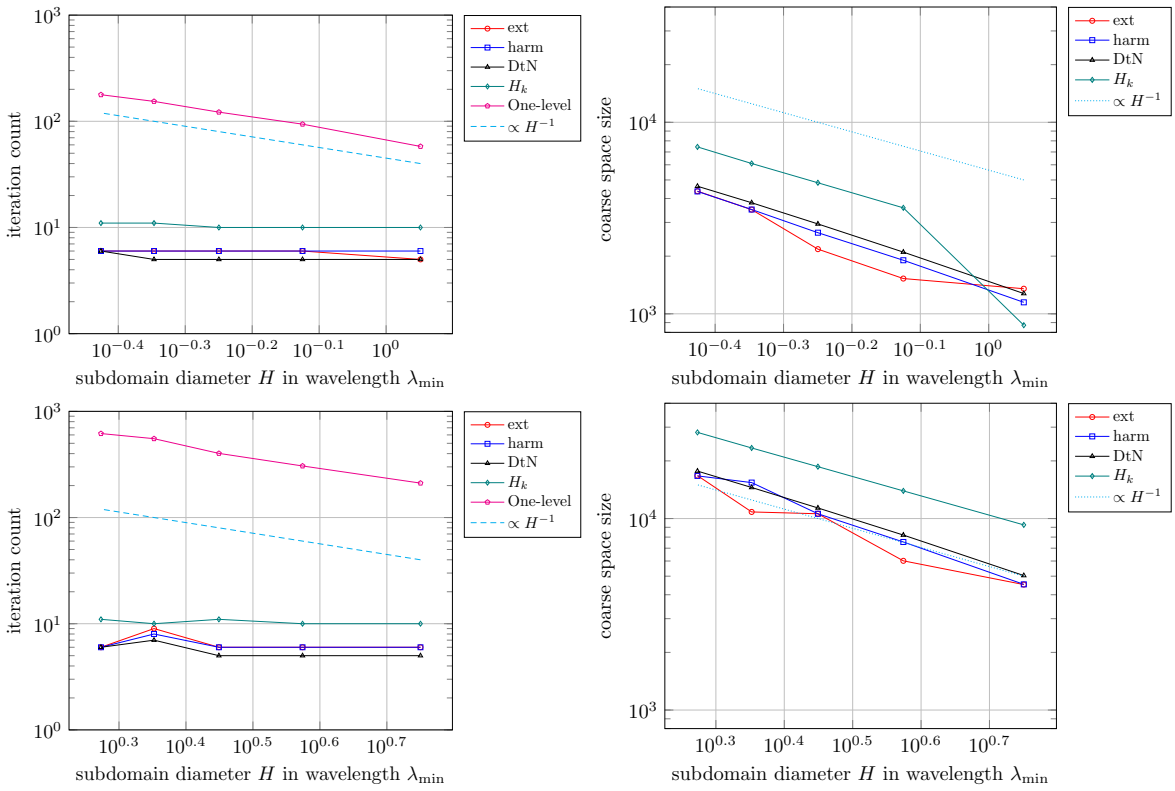


Figure 6.6: Influence of the subdomain diameter on the iteration count (left) and coarse space size (right) in strong scaling for the heterogeneous media test case $\rho = 10$ with $\omega = 20$ (top) and $\omega = 100$ (bottom).

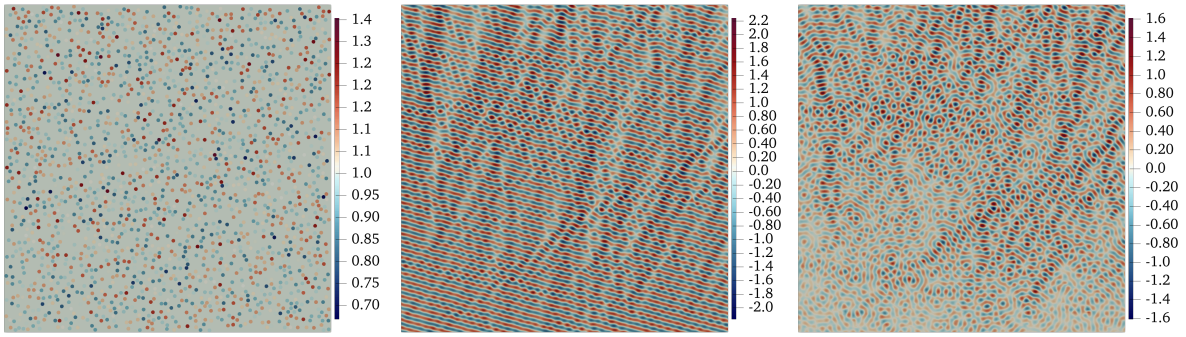


Figure 6.7: Values of the coefficient $c_0^2 m$ (left), real part of total field $\Re(u)$ (middle), and real part of scattered field $\Re(u_s)$ (right).

						1lvl	ext			harm			DtN			H_k		
L [λ_0]	N	n	H [λ_0]	n_s	$n_s^{\partial\Omega_s}$	It	It	CS	CS_s	It	CS	CS_s	It	CS	CS_s	It	CS	CS_s
33.8	8	1213761	11.9	154757	1664	37	6	2073	259	6	2371	296	12	2663	333	24	3894	487
33.8	16	1213761	8.4	78134	1121	46	7	2499	156	8	2907	182	16	1807	113	23	5380	336
33.8	32	1213761	6.0	39751	812	66	7	4456	139	9	5048	158	19	2652	83	21	7791	243
33.8	64	1213761	4.2	20308	568	83	8	6460	101	11	7098	111	24	2245	35	22	10906	170
33.8	128	1213761	3.0	10487	413	112	9	8198	64	13	9874	77	26	3564	28	23	15849	124
33.8	256	1213761	2.1	5454	286	145	10	14222	56	16	14608	57	31	4878	19	24	22005	86

Table 6.4: Strong scaling experiment for the imaging test case.

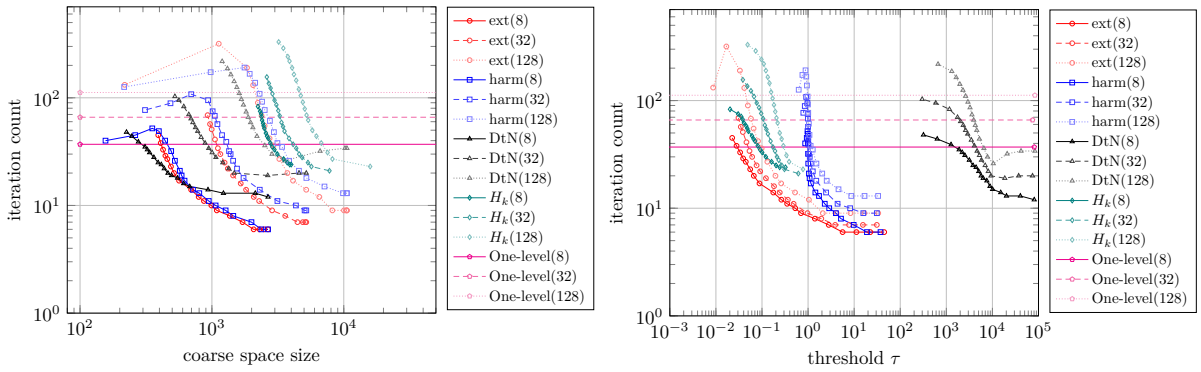


Figure 6.8: Influence of the coarse space size (left) and threshold choice (right) on the iteration count in strong scaling for the imaging test case.

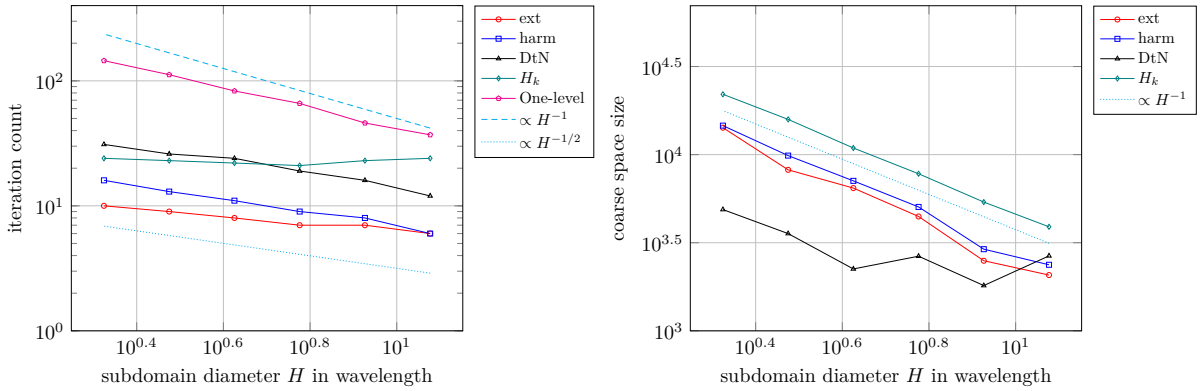


Figure 6.9: Influence of the subdomain diameter on the iteration count in strong scaling for the imaging test case.

Strong scaling test We first provide results for a strong scaling test, increasing the number of subdomains for a fixed problem size. The full domain has size $23.9\lambda_0 \times 23.9\lambda_0$ with 1213761 DOFs. The full results are provided in Table 6.4. The entries are the same as the ones previously described, except we use the reference wavelength λ_0 to measure lengths. Besides, we report the number of iterations of GMRES to reach the required tolerance as a function of the total coarse space size and choice of threshold parameter in Figure 6.8. The number of subdomains is provided in brackets in the legend entries. The results reported in the table correspond to the parameters with the minimum number of iterations obtained for each method, which explains the difference for the DtN approach. The number of iterations of GMRES to reach the required tolerance as a function of the subdomain diameter (measured in wavelength) is reported Figure 6.9.

						1lvl	ext			harm			DtN			H_k		
L [λ_0]	N	n	H [λ_0]	n_s	$n_s^{\partial\Omega_s}$	It	It	CS	CS_s	It	CS	CS_s	It	CS	CS_s	It	CS	CS_s
16.9	8	292873	6.0	38108	829	37	6	1007	126	6	1206	151	12	1060	132	20	1990	249
23.9	16	587129	6.0	38312	792	42	7	2042	128	7	2406	150	16	924	58	22	3804	238
33.8	32	1213761	6.0	39751	812	66	7	4456	139	9	5048	158	19	2652	83	21	7791	243
47.8	64	2366557	6.0	38856	798	86	8	9443	148	11	9288	145	23	5685	89	23	15327	239
67.5	128	4809133	6.0	39573	817	127	9	17250	135	14	19697	154	23	10340	81	23	31375	245
95.5	256	9520591	6.0	39238	811	187	11	39786	155	20	41337	161	29	13850	54	26	62267	243

Table 6.5: Weak scaling experiment for the imaging test case.

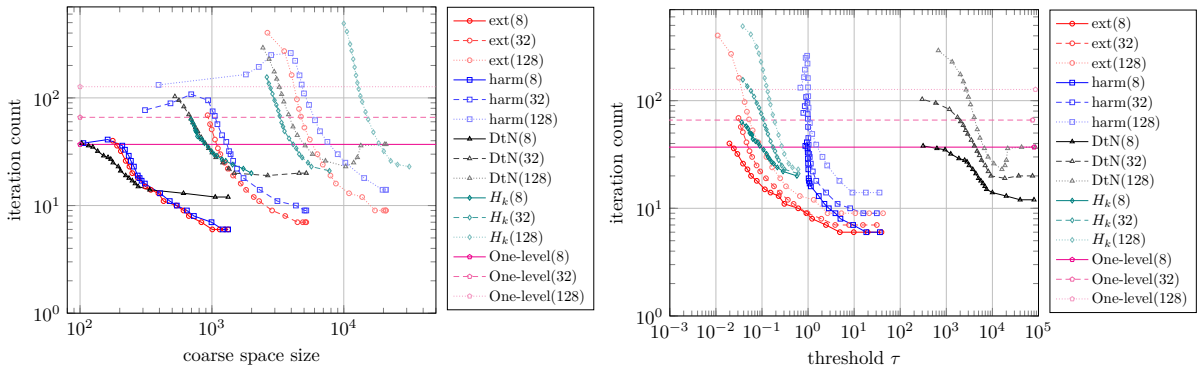


Figure 6.10: Influence of the coarse space size (left) and threshold choice (right) on the iteration count in weak scaling for the imaging test case.

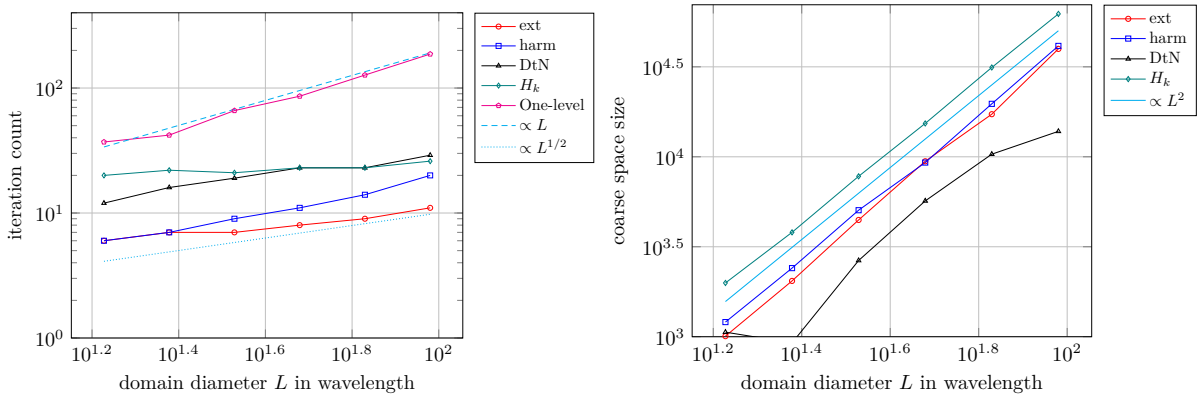


Figure 6.11: Influence of the full domain diameter on the iteration count in weak scaling for the imaging test case.

Weak scaling test We then perform a weak scaling test, increasing the number of subdomains for a fixed subdomain size. Each subdomain is in average of size $4.2\lambda_0 \times 4.2\lambda_0$ with about 39000 DOFs. The mesh size is fixed, so the pollution effect is not taken into account. The full results are provided in Table 6.5 with the same entries as for the strong scaling test case. The number of iterations of GMRES to reach the required tolerance as a function of the total coarse space size and choice of threshold parameter is reported Figure 6.10. The number of iterations of GMRES to reach the required tolerance as a function of the full domain diameter (measured in wavelength) is reported Figure 6.11.

Overlap and partition of unity function We turn to the study of the influence of the width of the overlap. So far the results for this test case are obtained for minimal overlap (with

a symmetry constraint with respect to the interface). The number of iterations of GMRES to reach the required tolerance as a function of the total coarse space size is reported in Figure 6.12. The size of the overlap measured as the number of cell layers from one side of the interface is indicated in the brackets of the legend entries.

Three different partition of unity functions, all defined as $\mathbb{P}_1$ functions hence piecewise affine, are used. The first one is the steepest, going from 1 to 0 in the smallest number of cells (typically two) close to the interface. The second one is the smoothest, going from 1 to 0 on the full size of the overlap. The third one is an intermediate choice, decaying smoothly on the overlap but vanishing one layer of cells before reaching the boundary, hence ensuring that its first derivative vanishes on the boundary (which makes sense in our case since we are using optimised boundary conditions as transmission conditions).

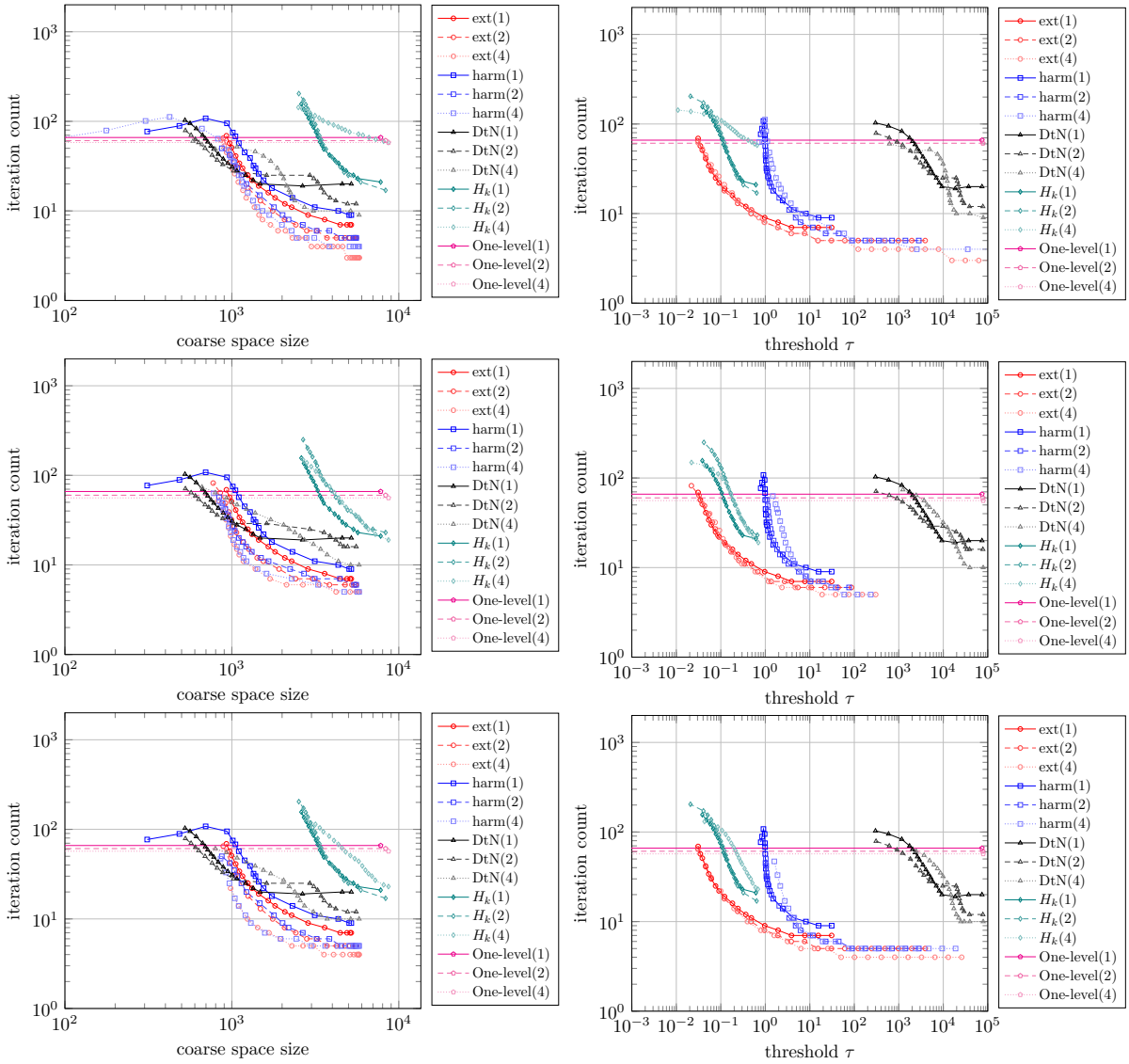


Figure 6.12: Influence of the overlap and the partition of unity function on the iteration count for the imaging test case. Steep (top), linearly decaying (middle) and with vanishing first derivative on the boundary (bottom).

Assessment and comparative conclusions for the medical imaging test case From these tests we can draw the following conclusions:

- The two-level methods remain robust in both weak and strong scalings, albeit with a

moderate increase in the iteration count with the number of subdomains, which is well-controlled by enlarging the coarse space.

- Harmonic and extended-harmonic coarse spaces yield the lowest iteration counts for a given coarse space size, with a slight advantage for the extended harmonic coarse space for larger number of subdomains.
- DtN coarse spaces remain competitive and compact, with a slightly worse iteration count but with the lowest size. We remark however some loss of robustness in the sense that the decrease in the number of iteration is not always monotonic with the size of the coarse space.
- H_k-GenEO coarse spaces are robust but more expensive.
- Sensible choices for the threshold τ are comparable to the previous test case, except for the DtN coarse spaces which required significantly larger values.
- Wider overlaps generally reduce iteration counts, especially for the harmonic and extended-harmonic coarse spaces.
- Smoother partition of unity functions (especially with vanishing derivative) yield slightly better convergence.

The imaging test case highlights the necessity of two-level methods heterogeneous for media with overall performance similar to the simple test on a square domain.

6.5.3 COBRA Cavity Test Case

Description of the problem We now move to a more sophisticated test case for spectral coarse spaces: the *COBRA cavity*, a three-dimensional, S-shaped waveguide originally designed for electromagnetic scattering studies by EADS Aerospatiale Matra Missiles as part of the EM-JINA 98 workshop (see [64, 70]). This geometry has since been adopted as a benchmark in domain decomposition research (e.g. [28, 9]) due to its challenging features for mid- to high-frequency wave propagation.

Unlike previous test cases with simpler geometries, the COBRA cavity introduces *geometric complexity* through its curvature, which can lead to *wave trapping effects*. Although the wavespeed is constant, the intricate shape of the cavity poses significant numerical challenges that go beyond those of straight waveguides.

The cross-section of the cavity measures 11 cm $\times$ 8.4 cm. The cavity walls are modelled as sound-soft, imposing Dirichlet boundary conditions. To reduce the unbounded scattering domain to a finite computational domain, we embed the cavity in a surrounding box. The material parameters are $a(\boldsymbol{x}) \equiv 1$ and $m(\boldsymbol{x}) \equiv 1$ and only the wavenumber k is varied. All geometrical parameters are measured in terms of the wavelength $\lambda = 2\pi/k$. The sides of the box are positioned 10 cm from the cavity in all directions, corresponding to between 1.3 and 5.7 wavelength λ depending on the frequency used in the weak scaling study presented below. Robin (impedance) boundary conditions are applied on all faces of the outer box.

A normally incident plane wave excites the cavity. The problem is discretised using $\mathbb{P}_2$ Lagrange finite elements, with a resolution of 8 points per wavelength to ensure adequate accuracy in the high-frequency regime.

Weak scaling test To assess scalability in more realistic and complex 3D geometries, we perform a *weak scaling test* on the COBRA cavity benchmark. In this experiment, the number of subdomains and total degrees of freedom increase proportionally with the wavenumber k, while keeping the average *subdomain size fixed* (approximately 1.6 wavelength λ in diameter

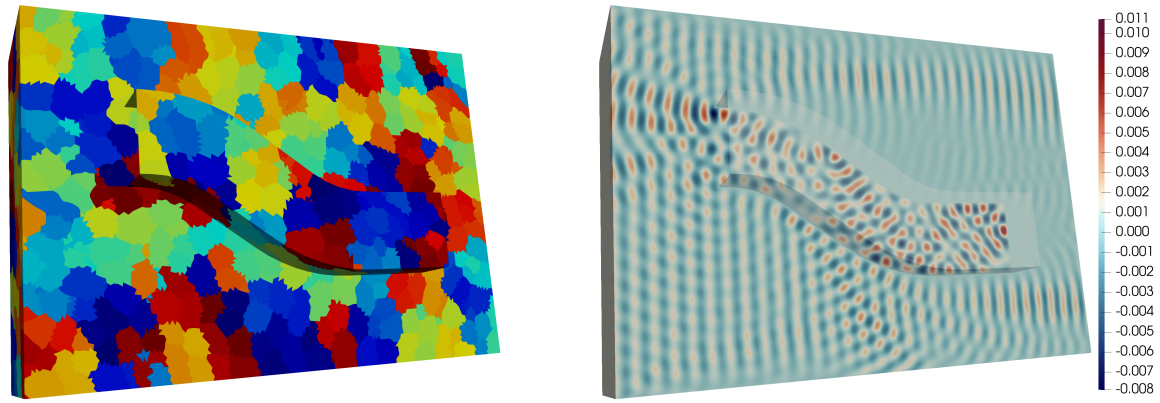


Figure 6.13: Partitioning of the COBRA cavity domain into 2916 subdomains (left) and real part of the total field (right) for $L = 22.9\lambda$ ($k = 360m^{-1}$).

					1lvl	ext			harm			DtN			H_k		
$L[\lambda]$	N	n	n_s	$n_s^{\partial\Omega_s}$	It	It	CS	CS_s	It	CS	CS_s	It	CS	CS_s	It	CS	CS_s
5.1	32	473004	22336	2986	51	8	6400	200	8	6400	200	13	6400	200	17	6400	200
7.6	108	1761164	25895	3824	157	10	21600	200	10	21600	200	25	21600	200	38	21600	200
10.2	256	4140366	26525	4144	359	14	51200	200	12	51200	200	44	51200	200	63	51200	200
12.7	500	8431281	27980	4468	432	16	100000	200	14	100000	200	64	100000	200	106	100000	200
15.3	864	13927097	27839	5035	942	34	172800	200	34	172800	200	97	172800	200	> 200	172800	200
17.8	1372	18879654	24589	4676	1055	44	274400	200	46	274400	200	120	274400	200	> 200	274400	200
20.4	2048	32848020	28143	5210	3711	60	409600	200	79	409600	200	162	409600	200	> 200	409600	200
22.9	2916	44520439	27138	5121	3398	74	583200	200	107	583200	200	> 200	583200	200	> 200	583200	200

Table 6.6: Weak scaling experiment for the COBRA cavity test case.

and around 27,000 degrees of freedom per subdomain). The number of subdomains N increases from 32 for $k = 80m^{-1}$ to 2916 for $k = 360m^{-1}$. Correspondingly, the diameter L of the global domain in wavelength goes from 5.1λ to 22.9λ. As an illustration, Figure 6.13 shows the partitioning of the domain as well as the real part of the solution for $L = 22.9\lambda$.

As shown in Table 6.6, the domain diameter measured in wavelengths increases from $L = 5.1\lambda$ (for $N = 32$) up to $L = 22.9\lambda$ (for $N = 2916$), the subdomain size measured in wavelengths is kept constant with $H = 1.6\lambda$. This setup preserves local resolution while increasing the global problem size, making it ideal to evaluate the robustness of coarse spaces and the scalability of the two-level ORAS preconditioner in the mid- to high-frequency regime. A constant average contribution from each subdomain to the coarse-space (CS_S) is also maintained, allowing a clear

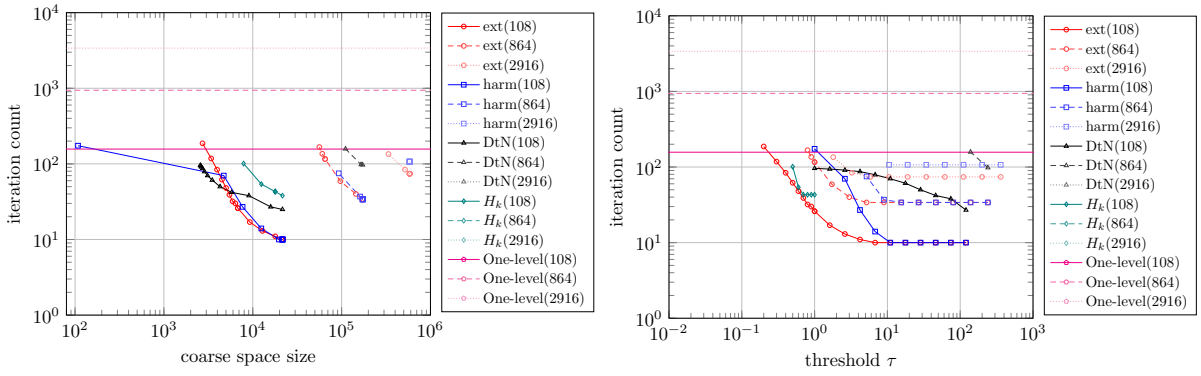


Figure 6.14: Influence of the coarse space size (left) and threshold choice (right) on the iteration count in weak scaling for the COBRA cavity test case.

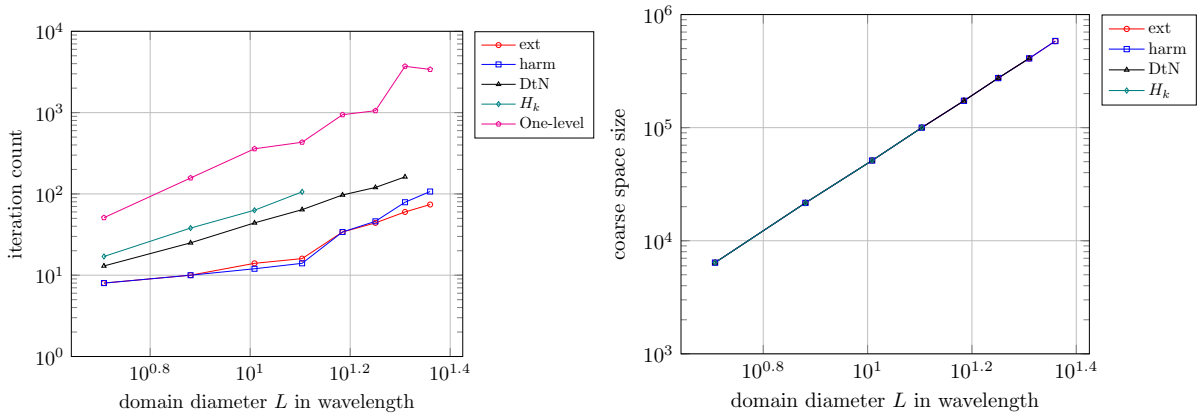


Figure 6.15: Influence of the full domain diameter on the iteration count in weak scaling for the COBRA cavity test case.

comparison of iteration counts.

In addition to that, we also report the number of GMRES iterations to reach the required tolerance as a function of the total coarse space size and choice of threshold parameter in Figure 6.14. The number of subdomains is provided in brackets in the legend entries.

The number of iterations of GMRES to reach the required tolerance as a function of the full domain diameter (measured in wavelength) is reported in Figure 6.15.

Assessment and comparative conclusions for the COBRA cavity test case

- GMRES iteration counts show that all two-level methods significantly outperform the one-level baseline, especially as the global domain size L increases.
- Harmonic and extended harmonic coarse spaces consistently achieve the lowest iteration counts, with only mild growth across scales.
- The DtN coarse space performs competitively at moderate scales but exhibits a slight increase in iteration count at higher L, and eventually breaks down for the largest problem.
- H_k-GenEO also performs competitively at moderate scales but breaks down beyond a certain coarse space size and for larger domains.

6.5.4 The GO_3D_OBS Crustal Geomodel

Description of the problem The `GO_3D_OBS` crustal geomodel is a high-resolution, three-dimensional synthetic model specifically designed to evaluate seismic imaging techniques, particularly for deep crustal and subduction zone exploration [53]. It represents a continental margin at a regional scale, incorporating realistic geological heterogeneities and sharp velocity contrasts that are characteristic of tectonic plate boundaries.

The model features a wide range of acoustic wavespeeds, $c(\boldsymbol{x})$ varies from $1500\,\mathrm{ms}^{-1}$ in the near-surface sedimentary layers and water column up to $8639\,\mathrm{ms}^{-1}$ in the lower crust and upper mantle. This range captures the physical complexity required to simulate realistic wave propagation in crustal-scale full-waveform inversion (FWI) settings.

Importantly, the geomodel includes the essential structural features of subduction zones, such as a dipping slab, accretionary prism, crust-mantle transitions, and bathymetry. It is inspired by a real-world FWI case study performed in the eastern Nankai Trough [54], and is designed to stress-test forward and inverse modelling tools in the presence of complex multiscale heterogeneity.

For the simulations considered in this work, we extract a target subregion of the full model, covering 20 km × 102 km × 28.3 km (Figure 6.16). The material coefficients of our model for this test case are set to $a \equiv 1$ and $m(\boldsymbol{x}) = c^{-2}(\boldsymbol{x})$. The geometrical parameters are measured according to the minimal wavelength $\lambda_m = 2\pi \min_{\boldsymbol{x}}(c(\boldsymbol{x}))/\omega$. The domain is discretised with an unstructured tetrahedral mesh, adapted to the local wavelength to ensure accurate resolution of the wavefield, with denser refinement near the bathymetry and across high-contrast interfaces.

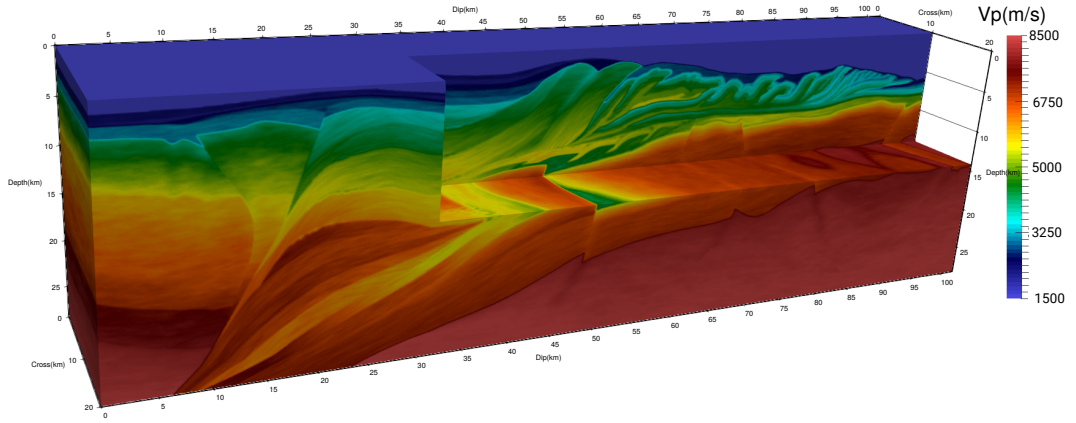


Figure 6.16: Target of the regional GO_3D_OBS model representing the crust of a subduction zone [53].

For this simulation, we use an unstructured mesh that complies with the bathymetry. Below the bathymetry, the size of the elements is set according to the local wavelength such that we have approximately 4 points per wavelength for the reference angular frequency $\omega = 3.75\pi\,\mathrm{rad\,s^{-1}}$ (Figure 6.17).

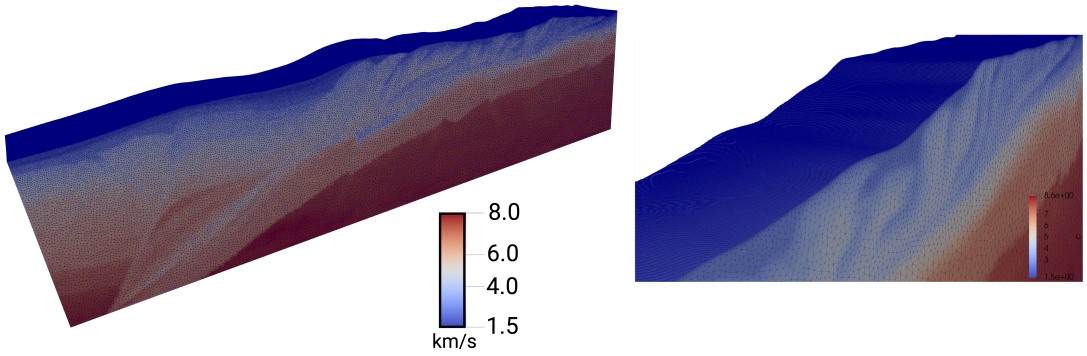


Figure 6.17: Tetrahedral mesh of the GO_3D_OBS model adapted to the local wavelength. Note that the mesh in the water layer is not shown in the Figure. Instead, the conformal meshing of the seabed is highlighted. The right panel shows a zoom.

Strong scaling test To further assess scalability in heterogeneous 3D media, we perform a *strong scaling test* on the crustal-geomodel benchmark. In this experiment, we consider two partitions with 1024 and 2048 subdomains, keeping $n = 22297073$ constant.

The results are reported in Table 6.7 and we also report the number of GMRES iterations to reach the required tolerance as a function of the total coarse space size and choice of threshold parameter in Figure 6.18. The number of subdomains is provided in brackets in the legend entries.

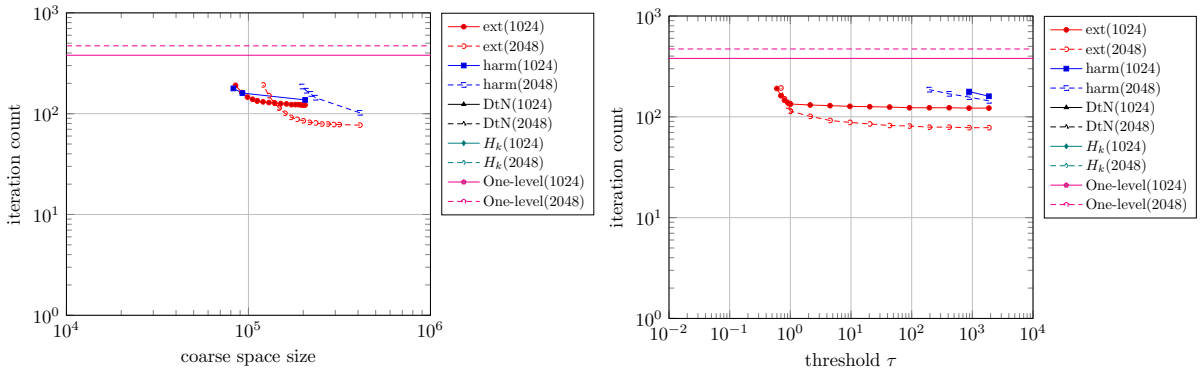


Figure 6.18: Influence of the coarse space size (left) and threshold choice (right) on the iteration count in a strong scaling experiment for the GO_3D_OBS test case.

					1lvl	ext			harm			DtN			H_k		
$L[\lambda_0]$	N	$H[\lambda_0]$	n_s	$n_s^{\partial\Omega_s}$	It	It	CS	CS_s	It	CS	CS_s	It	CS	CS_s	It	CS	CS_s
48.3	1024	4.8	35185	5416	380	122	204800	200	137	204800	200	> 200	204800	200	> 200	204800	200
48.3	2048	3.8	19864	3712	471	77	409600	200	102	409600	200	> 200	409600	200	> 200	409600	200

Table 6.7: Strong scaling experiment for the GO_3D_OBS test case.

Assessment and comparative conclusions for the crustal-geomodel test case

- Harmonic and extended-harmonic coarse spaces are able to provide a reduction in the iteration count compared to the one-level method. The smallest number of iterations is obtained for the extended-harmonic one, with more compact coarse spaces in the case with the largest number of subdomains.
- In contrast, both DtN and H_k-GenEO coarse spaces do not provide convergence for this challenging heterogeneous 3D test case.
- The harmonic coarse space required significantly larger values for the eigenvalue threshold parameter.

6.6 Conclusions

This study presents a comprehensive evaluation of spectral coarse spaces in two-level overlapping Schwarz methods for solving high-frequency Helmholtz problems in both homogeneous and heterogeneous media. The analysis spans a wide range of numerical settings, from idealised 2D test cases to large-scale 3D benchmarks such as the COBRA cavity and the GO_3D_OBS crustal geomodel.

A comparative overview of the key characteristics of each spectral coarse space is provided in Table 6.8. Our main findings can be summarised as follows:

- **Two-level domain decomposition is essential** for robust and scalable solvers at mid-to-high frequencies. All spectral coarse spaces studied outperform the one-level ORAS baseline by a wide margin.
- **Extended harmonic coarse spaces** offer the best trade-off between solver efficiency and coarse space size. They consistently deliver low GMRES iteration counts, even under weak scaling and in high-frequency 3D heterogeneous settings. We point out however that the eigenproblem is posed in an enlarged domain.

Problem	d	**Medium**	**B.C.**	**Scaling**	**Extended**	**Harmonic**	**DtN**	H_k**-GenEO**
Square	2D	homogeneous	Robin	strong	✓	✓	✓✓	✓
	2D	heterogeneous	Robin	strong	✓	✓	✓✓	✓
Medical imaging	2D	heterogeneous	PML	strong	✓✓	✓✓	✓	✓
	2D	heterogeneous	PML	weak	✓✓	✓✓	✓	✓
COBRA cavity	3D	homogeneous	Robin	weak	✓✓	✓✓	✓/✗	✓/✗
Crustal geomodel	3D	heterogeneous	PML	strong	✓✓	✓✓	✗	✗

Table 6.8: An overview of which coarse spaces perform well in the different problem scenarios tests. A ✓ indicates that the method performs well, with ✓✓ indicating this method was most favourable in a particular instance. A ✗ indicates that a method provided relatively little to no gain over the one-level method.

- **Harmonic coarse spaces** have a performance comparable to that of extended-harmonic coarse spaces. In some occurrences, in particular with a large number of subdomains, they require slightly larger coarse space sizes or a few more iterations compared to their extended counterpart.
- **DtN coarse spaces** are compact and competitive for moderate problem sizes but become less effective for larger domains and higher frequency, and breakdown can be observed in 3D. We remark a strong sensitivity with respect to the eigenvalue threshold parameter τ.
- $\boldsymbol{H_k}$**-GenEO** exhibits strong robustness in 2D but incurs significantly higher computational cost due to larger coarse spaces, and breakdown can be observed in 3D. It remains an attractive choice when memory constraints are less critical.
- Additional implementation aspects—such as overlap width, partition of unity smoothness, and eigenvalue threshold tuning—affect performance and should be carefully optimised in practice.

Overall, the results validate the use of physics-informed spectral coarse spaces—especially harmonic and extended harmonic types—as effective strategies for achieving scalability and wavenumber robustness in Helmholtz solvers across a wide spectrum of applications.

A comparison of these domain decomposition preconditioners in terms of setup and GMRES run times in addition to iteration count and coarse space sizes has been left outside of the scope of this study but is of practical importance and could be a perspective for future work.

Chapter 7

Conclusions and Future Development

7.1 Conclusions

A central theme of this thesis is the development of preconditioners for high-frequency Helmholtz problems that remain effective under increasing wavenumber and heterogeneous coefficients. Domain decomposition (DD) remains a widely used strategy for its theoretical scalability, the ease which it can be be run in parallel and flexibility in handling heterogeneous media and boundary conditions.

The main objective has been to find a coarse space that captures the global features of the Helmholtz problem which are not captured by the one-level method. While there are trade-offs to the methods discussed, it has been demonstrated that using a spectral coarse space can bring improved robustness at moderate coarse space dimensions for a range of k.

We introduced and analysed the Δ_k-GenEO coarse space as a refinement of Δ-GenEO. The main theoretical contribution is a tightening of the k explicit conditions which are sufficient for robustness. In particular, the wavenumber dependence of the subdomain diameter H and the eigenvalue threshold τ improves from

$$H \lesssim k^{-2}, \qquad \tau \gtrsim k^{-8}$$

to

$$H \lesssim k^{-1}, \qquad \tau \gtrsim k^{-2}.$$

Numerically, Δ_k-GenEO consistently outperforms Δ-GenEO across the tested regimes where it converges in fewer GMRES iterations and with smaller a coarse space.

We introduced and developed a two-level preconditioner based on the H_k-GenEO coarse space, constructed directly from local indefinite eigenvalue problems. This contrasts with Δ-GenEO and Δ_k-GenEO, which rely on nearby SPD problems. The core theoretical result is a clean pair of k explicit conditions for robustness:

$$k\, H_\ell \lesssim 1, \qquad (1 + C_{\text{stab}})^2\, k^2\, \Theta \lesssim 1,$$

with constants independent of subdomain count, heterogeneity, and mesh size.

On unit-square tests, homogeneous and layered heterogeneous, H_k-GenEO exhibits stable iteration counts once a sufficient coarse space size is used.

Across the spectral options examined, several conclusions can be drawn:

- **Two-level is essential.** All spectral coarse spaces decisively outperform a one-level baseline at mid-to-high frequencies.

- **Extended harmonic** coarse spaces deliver the most favourable trade-off between GRMRES iterations and coarse space size in challenging settings.
- **Harmonic** require slightly larger coarse spaces or a few additional iterations relative to extended-harmonic variants.
- **DtN** coarse spaces are compact and competitive at moderate sizes but become sensitive to τ and can degrade in large 3D problems.
- **H_k-GenEO** is robust in 2D, however, the associated coarse spaces are typically larger, making memory a primary concern in 3D.

This thesis advances the understanding of spectral coarse spaces in two ways. Firstly, we have introduced the Δ_k-GenEO which has refined the k explicit bounds that previously existed for the Δ-GenEO. Secondly, we have introduced an indefinite formulation using the H_k-GenEO and developed the theoretical framework and defined the k explicit bounds. With the comparison against such methods as the harmonic, extended-harmonic, and DtN spectral spaces, these developments provide an argument for the use of spectral coarse spaces in order to obtain robust domain decomposition solvers for Helmholtz problems.

7.2 Future Development

There are a few directions where it could be impactful to expand on the work carried out in this thesis. Firstly, the theoretical framework of the H_k-GenEO has been established for the real Helmholtz problem with Dirichlet boundary conditions, allowing the theory to be developed for the indefinite problem without the further complication of dealing with complex spaces. The next stage for the development of the H_k-GenEO, or other coarse spaces utilising the underlying indefinite nature, could be in developing the theory for more general problems, such as using absorbing boundary conditions. As the theory is using the AS method, future developments could be done in conjunction with alternative methods, such as ORAS, or with using an adapted deflation method, such as ADEF1. Given the current state, a suitable τ can only be established after a number of numerical experiments have been carried out with a level of sensitivity analysis. An interesting area to develop the theory would be a method of establishing a suitable τ that would guarantee a good convergence rate.

One observation from the extensive numerical simulations carried out in Chapter 6, is the improvements in GMRES iteration counts when the Helmholtz-harmonic coarse spaces are used. The Harmonic GenEO coarse space [71, 62], despite being based on a generalised eigenvalue problem similar to those used in Δ_k-GenEO and Δ-GenEO, performs significantly better than both and even surpasses H_k-GenEO. One of the noticeable differences between the two methods is that the generalised eigenvalue problem in the Helmholtz-harmonic coarse spaces is solved over the Helmholtz-extended space. It would be interesting to see if the H_k-GenEO coarse space could see the same benefits from using the Helmholtz-extended space, or indeed another space that could improve the iteration count, whilst minimising the size of the coarse space.

References

[1] P. Amestoy, I. Duff, J.-Y. L'Excellent, and J. Koster. A fully asynchronous multifrontal solver using distributed dynamic scheduling. *SIAM Journal on Matrix Analysis and Applications*, 23(1):15–41, 2001.

[2] I. M. Babuska and S. A. Sauter. Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers? *SIAM Journal on Numerical Analysis*, 34(6):2392–2423, 1997.

[3] P. Bastian, R. Scheichl, L. Seelinger, and A. Strehlow. Multilevel spectral domain decomposition. *SIAM Journal on Scientific Computing*, 45(3):S1–S26, 2023.

[4] M. Bebendorf. *Hierarchical matrices*. Springer, 2008.

[5] E. Beretta, M. V. De Hoop, F. Faucher, and O. Scherzer. Inverse boundary value problem for the Helmholtz equation: quantitative conditional Lipschitz stability estimates. *SIAM Journal on Mathematical Analysis*, 48(6):3962–3983, 2016.

[6] H. Beriot and A. Modave. An automatic perfectly matched layer for acoustic finite element simulations in convex domains of general shape. *International Journal for Numerical Methods in Engineering*, 122:1239–1261, 2021.

[7] A. Bermúdez, L. Hervella-Nieto, A. Prieto, and R. Rodríguez. An exact bounded PML for the Helmholtz equation. *C. R. Math. Acad. Sci. Paris*, 339(11):803–808, 2004.

[8] M. Bollhöfer, M. J. Grote, and O. Schenk. Algebraic multilevel preconditioner for the Helmholtz equation in heterogeneous media. *SIAM Journal on Scientific Computing*, 31(5):3781–3805, 2009.

[9] M. Bonazzoli, V. Dolean, I. G. Graham, E. A. Spence, and P.-H. Tournier. Domain decomposition preconditioning for the high-frequency time-harmonic maxwell equations with absorption. *Mathematics of Computation*, 88:2559–2604, 2019.

[10] N. Bootland and V. Dolean. On the Dirichlet-to-Neumann coarse space for solving the Helmholtz problem using domain decomposition. In *Numerical mathematics and advanced applications—ENUMATH 2019*, volume 139 of *Lecture Notes in Computational Science and Engineering*, pages 175–184. Springer, Cham, 2021.

[11] N. Bootland and V. Dolean. Can DtN and GenEO coarse spaces be sufficiently robust for heterogeneous Helmholtz problems? *Mathematical and Computational Applications*, 27(3):35, 2022.

[12] N. Bootland, V. Dolean, I. G. Graham, C. Ma, and R. Scheichl. GenEO coarse spaces for heterogeneous indefinite elliptic problems. In S. C. Brenner, E. Chung, A. Klawonn, F. Kwok, J. Xu, and J. Zou, editors, *Domain Decomposition Methods in Science and Engineering XXVI*, pages 117–125, Cham, 2022. Springer International Publishing.

[13] N. Bootland, V. Dolean, I. G. Graham, C. Ma, and R. Scheichl. Overlapping Schwarz methods with GenEO coarse spaces for indefinite and nonself-adjoint problems. *IMA Journal of Numerical Analysis*, 43(4):1899–1936, 08 2022.

[14] N. Bootland, V. Dolean, P. Jolivet, and P.-H. Tournier. A comparison of coarse spaces for Helmholtz problems in the high frequency regime. *Computers & Mathematics with Applications*, 98:239–253, 2021.

[15] N. Bootland and T. Rees. Multipreconditioning with directional sweeping methods for high-frequency Helmholtz problems. In *European Conference on Numerical Mathematics and Advanced Applications*, pages 168–176. Springer, 2023.

[16] D. Braess. *Finite elements: Theory, fast solvers, and applications in solid mechanics*. Cambridge University Press, 2001.

[17] J. H. Bramble, J. E. Pasciak, and A. H. Schatz. The construction of preconditioners for elliptic problems by substructuring. I. *Mathematics of Computation*, 47(175):103–134, 1986.

[18] X.-C. Cai and M. Sarkis. A restricted additive Schwarz preconditioner for general sparse linear systems. *SIAM Journal on Scientific Computing*, 21(2):792–797 (electronic), 1999.

[19] T. F. Chan and T. P. Mathew. Domain decomposition algorithms. *Acta Numerica*, 3:61–143, 1994.

[20] C. Chevalier and F. Pellegrini. Pt-scotch: A tool for efficient parallel graph ordering. *Parallel Computing*, 34(6):318–331, 2008. Parallel Matrix Algorithms and Applications.

[21] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. Society for Industrial and Applied Mathematics, 2002.

[22] P.-H. Cocquet and M. J. Gander. Analysis of the shifted Helmholtz expansion preconditioner for the Helmholtz equation. In P. E. Bjørstad, S. C. Brenner, L. Halpern, H. H. Kim, R. Kornhuber, T. Rahman, and O. B. Widlund, editors, *Domain Decomposition Methods in Science and Engineering XXIV*, pages 195–203, Cham, 2018. Springer International Publishing.

[23] L. Conen, V. Dolean, R. Krause, and F. Nataf. A coarse space for heterogeneous Helmholtz problems based on the Dirichlet-to-Neumann operator. *Journal of Computational and Applied Mathematics*, 271:83–99, 2014.

[24] V. Dolean, M. Fry, I. G. Graham, and M. Langer. Schwarz preconditioner with H_k-GenEO coarse space for the indefinite Helmholtz problem. *arXiv:2406.06283*, 2024.

[25] V. Dolean, M. Fry, and M. Langer. Can Symmetric Positive Definite (SPD) coarse spaces perform well for indefinite Helmholtz problems? *Journal of Computational and Applied Mathematics*, page 117403, 2026.

[26] V. Dolean, M. Fry, and M. Langer. Wavenumber explicit estimates of the Schwarz preconditioner with Δ-GenEO coarse space for the indefinite Helmholtz problem. In *Domain Decomposition Methods in Science and Engineering XXVIII*, volume 155 of *Lecture Notes in Computational Science and Engineering*, pages 287–294. Springer, 2026.

[27] V. Dolean, M. Fry, M. Langer, E. Parolin, and P.-H. Tournier. Achieving wavenumber robustness in domain decomposition for heterogeneous Helmholtz equation: an overview of spectral coarse spaces. *arXiv:2509.02131*, 2025.

[28] V. Dolean, M. J. Gander, S. Lanteri, J.-F. Lee, and Z. Peng. Effective transmission conditions for domain decomposition methods applied to the time-harmonic curl-curl Maxwell's equations. *Journal of Computational Physics*, 280:232–247, 2015.

[29] V. Dolean, P. Jolivet, and F. Nataf. *An introduction to domain decomposition methods*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2015. Algorithms, theory, and parallel implementation.

[30] V. Dolean, P. Jolivet, P. Tournier, and S. Operto. Iterative frequency-domain seismic wave solvers based on multi-level domain-decomposition preconditioners. In *EAGE 2020 Annual Conference & Exhibition Online*, pages 1–5. European Association of Geoscientists & Engineers, 2020.

[31] V. Dolean, P.-H. Tournier, P. Jolivet, and S. Operto. Large-scale frequency-domain seismic wave modeling on h-adaptive tetrahedral meshes with iterative solver and multi-level domain-decomposition preconditioners. In *SEG Technical Program Expanded Abstracts 2020*, pages 2683–2688. Society of Exploration Geophysicists, 2020.

[32] M. Dryja. An additive Schwarz algorithm for two-and three-dimensional finite element elliptic problems. *Domain decomposition methods*, pages 168–172, 1989.

[33] M. Dryja and O. B. Widlund. *Some domain decomposition algorithms for elliptic problems*, page 273–291. Academic Press Professional, Inc., USA, 1990.

[34] I. S. Duff, A. M. Erisman, and J. K. Reid. *Direct methods for sparse matrices*. Oxford University Press, 2017.

[35] S. C. Eisenstat, H. C. Elman, and M. H. Schultz. Variational iterative methods for nonsymmetric systems of linear equations. *SIAM Journal on Numerical Analysis*, 20(2):345–357, 1983.

[36] H. C. Elman, O. G. Ernst, and D. P. O'leary. A multigrid method enhanced by krylov subspace iteration for discrete Helmholtz equations. *SIAM Journal on Scientific Computing*, 23(4):1291–1315, 2001.

[37] M. Embree. How descriptive are GMRES convergence bounds? *arXiv:2209.01231*, 2022.

[38] B. Engquist and L. Ying. Sweeping preconditioner for the Helmholtz equation: Hierarchical matrix representation. *Communications on Pure and Applied Mathematics*, 64(5):697–735, 2011.

[39] B. Engquist and L. Ying. Sweeping preconditioner for the Helmholtz equation: Moving perfectly matched layers. *Multiscale Modeling & Simulation*, 9(2):686–710, 2011.

[40] O. G. Ernst and M. J. Gander. Why it is difficult to solve Helmholtz problems with classical iterative methods. In I. G. Graham, T. Y. Hou, O. Lakkis, and R. Scheichl, editors, *Numerical Analysis of Multiscale Problems*, pages 325–363. Springer, 2012.

[41] O. G. Ernst and M. J. Gander. Multigrid methods for Helmholtz problems: A convergent scheme in 1d using standard components. *Direct and Inverse Problems in Wave Propagation and Applications. De Gruyer*, pages 135–186, 2013.

[42] L. C. Evans. *Partial differential equations*, volume 19. American mathematical society, 2022.

[43] J. Fish and Y. Qu. Global-basis two-level method for indefinite systems. Part 1: convergence studies. *International Journal for Numerical Methods in Engineering*, 49(3):439–460, 2000.

[44] A. S. Fokas and B. Pelloni, editors. *Unified transform for boundary value problems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2015. Applications and advances.

[45] J. Galkowski, D. Lafontaine, E. A. Spence, and J. Wunsch. The *hp*-FEM applied to the Helmholtz equation with PML truncation does not suffer from the pollution effect. *Commun. Math. Sci.*, 22(7):1761–1816, 2024.

[46] M. J. Gander. Schwarz methods over the course of time. *Electronic Transactions on Numerical Analysis*, 31(5):228–255, 2008.

[47] M. J. Gander, I. G. Graham, and E. A. Spence. Applying GMRES to the Helmholtz equation with shifted laplacian preconditioning: what is the largest shift for which wavenumber-independent convergence is guaranteed? *Numerische Mathematik*, 131(3):567–614, 2015.

[48] M. J. Gander and H. Zhang. A class of iterative solvers for the Helmholtz equation: Factorizations, sweeping preconditioners, source transfer, single layer potentials, polarized traces, and optimized Schwarz methods. *SIAM Review*, 61(1):3–76, 2019.

[49] J. Garnier, L. Giovangigli, Q. Goepfert, and P. Millien. Scattered wavefield in the stochastic homogenization regime. *arXiv:2309.07777*, 2023.

[50] J. Garnier, L. Giovangigli, Q. Goepfert, and P. Millien. Probing the speckle to estimate the effective speed of sound, a first step towards quantitative ultrasound imaging. *arXiv:2505.07566*, 2025.

[51] A. George. Nested dissection of a regular finite element mesh. *SIAM Journal on Numerical Analysis*, 10(2):345–363, 1973.

[52] S. Gong, I. G. Graham, and E. A. Spence. Domain decomposition preconditioners for high-order discretizations of the heterogeneous Helmholtz equation. *IMA Journal of Numerical Analysis*, 41(3):2139–2185, 2021.

[53] A. Górszczyk and S. Operto. GO_3D_OBS: the multi-parameter benchmark geomodel for seismic imaging method assessment and next-generation 3D survey design (version 1.0). *Geoscientific Model Development*, 14(3):1773–1799, 2021.

[54] A. Górszczyk, S. Operto, and M. Malinowski. Toward a robust workflow for deep crustal imaging by FWI of OBS data: the eastern Nankai Trough revisited. *Journal of Geophysical Research: Solid Earth*, 122(6):4601–4630, 2017.

[55] I. Graham, P. Lechner, and R. Scheichl. Domain decomposition for multiscale PDEs. *Numerische Mathematik*, 106:589–626, 06 2007.

[56] I. G. Graham, E. A. Spence, and E. Vainikko. Domain decomposition preconditioning for high-frequency Helmholtz problems with absorption. *Mathematics of Computation*, 86(307):2089–2127, 2017.

[57] I. G. Graham, E. A. Spence, and E. Vainikko. Recent results on domain decomposition preconditioning for the high-frequency Helmholtz equation using absorption. *Lahaye D., Tang J., Vuik K. (eds) Modern Solvers for Helmholtz Problems. Geosystems Mathematics. Birkhäuser, Cham*, pages 3–26, 2017.

[58] I. G. Graham, E. A. Spence, and J. Zou. Domain decomposition with local impedance conditions for the Helmholtz equation with absorption. *SIAM Journal on Numerical Analysis*, 58(5):2515–2543, 2020.

[59] R. Haferssas, P. Jolivet, and F. Nataf. An additive Schwarz method type theory for Lions's algorithm and a symmetrized optimized restricted additive Schwarz method. *SIAM Journal on Scientific Computing*, 39(4):A1345–A1365, 2017.

[60] I. Harari, M. Slavutin, and E. Turkel. Analytical and numerical studies of a finite element PML for the Helmholtz equation. *Journal of Computational Acoustics*, 8:121–137, 2000.

[61] V. Hernandez, J. E. Roman, and V. Vidal. SLEPc: A scalable and flexible toolkit for the solution of eigenvalue problems. *ACM Transactions on Mathematical Software*, 31(3):351–362, 2005.

[62] Q. Hu and Z. Li. A novel coarse space applying to the weighted Schwarz method for Helmholtz equations. *arXiv:2402.06905*, 2024.

[63] F. Ihlenburg. *Finite element analysis of acoustic scattering*. Springer, 1998.

[64] J.-M. Jin. *The Finite Element Method in Electromagnetics*. John Wiley & Sons, 2015.

[65] G. Karypis and V. Kumar. A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM Journal on Scientific Computing*, 20(1):359–392, 1998.

[66] G. Karypis and V. Kumar. A software package for partitioning unstructured graphs, partitioning meshes, and computing fill-reducing orderings of sparse matrices. Technical report, University of Minnesota, Department of Computer Science and Engineering, Army HPC Research Center, Minneapolis, MN, 1998.

[67] S. Le Borne, L. Grasedyck, and R. Kriemann. Domain-decomposition based $\mathcal{H}$-LU preconditioners. In *Domain Decomposition Methods in Science and Engineering XVI*, pages 667–674. Springer, 2007.

[68] G. Leoni. *A First Course in Sobolev Spaces*. Graduate studies in mathematics. American Mathematical Society, 2017.

[69] F. Liu and L. Ying. Additive sweeping preconditioner for the Helmholtz equation. *Multiscale Modeling & Simulation*, 14(2):799–822, 2016.

[70] J. Liu and J.-M. Jin. Scattering analysis of a large body with deep cavities. *IEEE Transactions on Antennas and Propagation*, 51(6):1157–1167, 2003.

[71] C. Ma, C. Alber, R. Scheichl, and Y. Zhang. Two-level restricted additive Schwarz preconditioner based on multiscale spectral generalized FEM for heterogeneous Helmholtz problems. *arXiv:2409.06533*, 2025.

[72] J. M. Melenk and C. Rojik. On commuting p-version projection-based interpolation on tetrahedra. *Mathematics of Computation*, 89(321):45–87, 2020.

[73] J. M. Melenk and S. Sauter. Wavenumber explicit convergence analysis for Galerkin discretizations of the Helmholtz equation. *SIAM Journal on Numerical Analysis*, 49(3):1210–1243, 2011.

[74] F. Nataf and E. Parolin. Coarse spaces for non-symmetric two-level preconditioners based on local generalized eigenproblems. *arXiv:2404.02758*, 2024.

[75] F. Nataf, H. Xiang, V. Dolean, and N. Spillane. A coarse space construction based on local Dirichlet-to-Neumann maps. *SIAM Journal on Scientific Computing*, 33(4):1623–1642, 2011.

[76] R. A. Nicolaides. Deflation of conjugate gradients with applications to boundary value problems. *SIAM Journal on Numerical Analysis*, 24(2):355–365, 1987.

[77] S. Operto, P. Amestoy, H. Aghamiry, S. Beller, A. Buttari, L. Combe, V. Dolean, M. Gerest, G. Guo, P. Jolivet, et al. Is 3D frequency-domain FWI of full-azimuth/long-offset OBN data feasible? the Gorgon data FWI case study. *The Leading Edge*, 42(3):173–183, 2023.

[78] Y. Saad. *Iterative methods for sparse linear systems*. SIAM, 2003.

[79] Y. Saad and M. H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 7:856–869, 1986.

[80] M. Sarkis. Partition of unity coarse spaces: enhanced versions, discontinuous coefficients and applications to elasticity. *Domain decomposition methods in science and engineering*, pages 149–158, 2003.

[81] M. Sarkis and D. B. Szyld. Optimal left and right additive schwarz preconditioning for minimal residual methods with euclidean and energy norms. *Computer Methods in Applied Mechanics and Engineering*, 196(8):1612–1621, 2007.

[82] A. Schatz and J. Wang. Some new error estimates for Ritz–Galerkin methods with minimal regularity assumptions. *Mathematics of Computation*, 65(213):19–27, 1996.

[83] A. H. Schatz. An observation concerning Ritz-Galerkin methods with indefinite bilinear forms. *Mathematics of Computation*, 28:959–962, 1974.

[84] G. L. Sleijpen, H. A. Van der Vorst, and J. Modersitzki. Differences in the Effects of Rounding Errors in Krylov Solvers for Symmetric Indefinite Linear Systems. *SIAM Journal on Matrix Analysis and Applications*, 22(3):726–751, 2001.

[85] N. Spillane, V. Dolean, P. Hauret, F. Nataf, C. Pechstein, and R. Scheichl. Abstract robust coarse spaces for systems of PDEs via generalized eigenproblems in the overlaps. *Numerische Mathematik*, 126(4):741–770, 2014.

[86] N. Spillane, D. B. Szyld, and P. Matalon. Do you precondition on the left or on the right. A poster at DD29. In *29th International Conference on Domain Decomposition Methods (DD29)*, 2025.

[87] J. Tang, R. Nabben, C. Vuik, and Y. Erlangga. Theoretical and numerical comparison of various projection methods derived from deflation, domain decomposition and multigrid methods. *Reports of the Department of Applied Mathematical Analysis, 07-04*, 01 2007.

[88] J. M. Tang, R. Nabben, C. Vuik, and Y. A. Erlangga. Comparison of two-level preconditioners derived from deflation, domain decomposition and multigrid methods. *Journal of Scientific Computing*, 39(3):340–370, 2009.

[89] A. Toselli and O. Widlund. *Domain decomposition methods—algorithms and theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2005.

[90] P.-H. Tournier, P. Jolivet, V. Dolean, H. S. Aghamiry, S. Operto, and S. Riffo. 3d finite-difference and finite-element frequency-domain wave simulation with multilevel optimized additive Schwarz domain-decomposition preconditioner: A tool for full-waveform inversion of sparse node data sets. *GEOPHYSICS*, 87(5):T381–T402, 2022.

[91] P.-H. Tournier, P. Jolivet, and F. Nataf. FFDDM: FreeFem Domain Decomposition Method, 2019.

[92] O. Widlund and M. Dryja. *An additive variant of the Schwarz alternating method for the case of many subregions*. Technical Report 339, Ultracomputer Note 131. Department of Computer Science, Courant Institute, 1987.

[93] J. Xia, S. Chandrasekaran, M. Gu, and X. S. Li. Superfast multifrontal method for large structured linear systems of equations. *SIAM Journal on Matrix Analysis and Applications*, 31(3):1382–1411, 2010.

[94] A. Zarmi and E. Turkel. A general approach for high order absorbing boundary conditions for the Helmholtz equation. *Journal of Computational Physics*, 242:387–404, 2013.