Optimized Schwarz Methods for the Advection-Diffusion Equation and for Problems with Discontinuous Coefficients

Olivier Dubois

Department of Mathematics and Statistics
McGill University, Montréal
Québec, Canada

June, 2007

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Your file Votre référence
ISBN: 978-0-494-38582-1
Our file Notre référence
ISBN: 978-0-494-38582-1

Abstract

Optimized Schwarz methods are iterative domain decomposition procedures with greatly improved convergence properties, for solving second order elliptic boundary value problems. The enhanced convergence is obtained by replacing the Dirichlet transmission conditions in the classical Schwarz iteration with more general conditions that are optimized for performance. The convergence is optimized through the solution of a min-max problem. The theoretical study of the min-max problems gives explicit formulas or characterizations for the optimized transmission conditions for practical use, and it permits the analysis of the asymptotic behavior of the convergence.

In the first part of this work, we continue the study of optimized transmission conditions for advection-diffusion problems with smooth coefficients. We derive asymptotic formulas for the optimized parameters for small mesh sizes, in the overlapping and non-overlapping cases, and show that these formulas are accurate when the component of the advection tangential to the interface is not too large.

In a second part, we consider a diffusion problem with a discontinuous coefficient and non-overlapping domain decompositions. We derive several choices of optimized transmission conditions by thoroughly solving the associated min-max problems. We show in particular that the convergence of optimized Schwarz methods improves as the jump in the coefficient increases, if an appropriate scaling of the transmission conditions is used. Moreover, we prove that optimized two-sided Robin conditions lead to mesh-independent convergence. Numerical experiments with two subdomains are presented to verify the analysis. We also report the results of experiments using

the decomposition of a rectangle into many vertical strips; some additional analysis is carried out to improve the optimized transmission conditions in that case.

On a third topic, we experiment with different coarse space corrections for the Schwarz method in a simple one-dimensional setting, for both overlapping and non-overlapping subdomains. The goal is to obtain a convergence that does not deteriorate as we increase the number of subdomains. We design a coarse space correction for the Schwarz method with Robin transmission conditions by considering an augmented linear system, which avoids merging the local approximations in overlapping regions. With numerical experiments, we demonstrate that the best Robin conditions are very different for the Schwarz iteration with, and without coarse correction.

Résumé

Les méthodes de Schwarz optimisées sont des procédures itératives de décomposition de domaine qui possèdent une convergence grandement améliorée, pour résoudre des problèmes elliptiques de deuxième ordre avec conditions aux limites. Cette convergence accélérée est obtenue en remplaçant les conditions de transmission de type Dirichlet dans les méthodes classiques de Schwarz par des conditions plus générales qui sont optimisées pour une meilleure performance. L'optimisation du taux de convergence est formulé à l'aide d'un problème de type min-max. L'étude théorique de ces problèmes min-max nous amène à des formules explicites ou à des caractérisations simples pour les conditions de transmission optimisées à être utilisées en pratique, et nous permet également l'analyse du comportement asymptotique de la convergence.

Dans une première partie, nous continuons l'étude des conditions de transmission optimisées pour les problèmes d'advection-diffusion avec coefficients continus. Nous obtenons des formules asymptotiques pour les paramètres optimisés, valides pour de petits pas de maillage, avec ou sans recouvrement des sous-domaines. Nous démontrons que ces formules donnent de bonnes approximations lorsque la composante de l'advection tangentielle à l'interface n'est pas trop grande.

Dans une deuxième partie de ce travail, nous considérons un problème de diffusion avec un coefficient discontinu, et une décomposition en sous-domaines sans recouvrement. Nous dérivons plusieurs choix de conditions de transmission optimisées en trouvant la solution précise des problèmes min-max qui y sont associés. Nous montrons en particulier que la convergence des méthodes de Schwarz optimisées s'améliore lorsque le saut du coefficient augmente, si on écrit les conditions sous une forme adéquate.

De plus, nous prouvons que les conditions de Robin optimisées avec deux paramètres libres produisent une convergence indépendante du pas de maillage. Des calculs numériques avec deux sous-domaines sont présentés pour vérifier l'analyse théorique. Nous rapportons également les résultats de tests utilisant la décomposition d'un rectangle en plusieurs bandes verticales; une analyse supplémentaire est performée afin d'améliorer la convergence des méthodes de Schwarz optimisées dans ce cas.

Sur un troisième sujet, nous expérimentons avec différentes techniques de correction sur un espace grossier pour les méthodes de Schwarz, pour un problème en une dimension, avec et sans recouvrement des sous-domaines. Le but est d'obtenir une convergence qui ne se détériore pas lorsque l'on augmente le nombre de sous-domaines. Nous proposons une correction sur un espace grossier pour la méthode de Schwarz avec conditions de type Robin, en considérant un système linéaire augmenté, ce qui permet d'éviter d'avoir à combiner plusieurs approximations locales dans les régions de recouvrement. À l'aide de calculs numériques, nous démontrons aussi que les meilleures conditions de Robin pour l'itération de Schwarz sont très différentes que l'on inclut ou non une correction sur un espace grossier.

Acknowledgments

First and foremost, I would like to thank my supervisor Prof. Martin J. Gander for his guidance, his availability, his great enthusiasm about research. The many discussions we had were always illuminating. Many thanks also go to Prof. Nilima Nigam, who served as my co-supervisor through many administrative procedures, but more importantly she was an indispensable motivator and career advisor. She went out of her way several times to help me and provided useful advice on countless occasions. I would also like to sincerely thank Prof. Ernst Hairer, who helped me arrange a visit at the Université de Genève, and provided funding in the form of a teaching assistantship. Part for this work was done during the productive periods of time I spent in Switzerland.

The applied mathematics group at McGill University provided a highly motivating environment for research. I had several helpful discussions with many people: Prof. Paul Tupper, Mélanie Beck, Mohammad Al-Khaleel, my office mates David Cottrell, Joel Phillips and Neil Olver, and finally Simon Gemmrich, who was kind enough to read parts of this thesis and supplied many useful comments. Thank you guys, and long live the Applied Math Working Seminar!

J'aimerais tout spécialement remercier ma famille pour leur support inconditionnel et leurs encouragements incessants, depuis le tout début de mes études. J'ai également une pensée pour mes grands-parents, qui voient finalement aboutir mes longues années d'études. Mes amis ont également eu un rôle très important: ils m'ont permis de décrocher des mathématiques et de me changer les idées à de nombreuses occasions. Merci à vous tous d'être là pour moi, même si je n'ai jamais vraiment réussi à vous

expliquer le sujet de ma thèse!

This work was supported by graduate scholarships from the Natural Sciences and Engineering Research Council of Canada (NSERC), the Fonds québécois de recherche sur la nature et les technologies (FQRNT), and McGill University.

Finally, I would like to thank the external examiner for reading this document, and submitting a report in a very short amount of time.

Table of Contents

Introduction

The discretization of second order elliptic boundary value problems, using methods such as finite differences, finite volume or finite elements, leads to sparse linear systems

$$A\mathbf{u} = \mathbf{b}.$$

In many applications, to achieve a certain accuracy in the discretization process, we need to use a very large number of nodes to mesh the domain, either because the domain of computation is physically big, or because we want to resolve small features in the solution. This means that the linear system can be of very large dimension. Examples of such important applications include weather prediction, aerospace engineering and petroleum reservoir simulations.

For solving large sparse linear systems, direct methods are often too costly, and we must rely on iterative methods instead. Moreover, in some cases, the matrix itself may be too large to fit in the memory of a single desktop computer. For these reasons, clusters of processors are used to share the data and the computational work. Domain decomposition is a very natural idea to efficiently compute in parallel the solution of such problems: the computational domain is decomposed into smaller subdomains, overlapping or non-overlapping, and each subdomain is assigned to a specific processor. At each step of the iterative method, local problems on the subdomains are solved in parallel, and then data is exchanged between neighboring subdomains. This is repeated until we get an acceptable approximation for the solution of the global linear system $A\mathbf{u} = \mathbf{b}$.

Most research on domain decomposition methods is aimed at obtaining *fast* convergence of this iterative process; we want to solve the local subproblems only a few

times. In particular, we do not want the convergence of the iteration to deteriorate when

- the mesh size h gets smaller (i.e. when increasing the problem size),
- the coefficients of the elliptic operator have large discontinuities,
- the number of subdomains increases.

The last criterion is required in order to obtain a *scalable* parallel method: ideally, when doubling the number of subdomains (processors), the running time of the solver should be cut by a factor of 2.

Optimized Schwarz methods are domain decomposition methods in which the transmission conditions (what type of data is exchanged between neighboring subdomains) are optimized to get fast convergence. This work concerns the study of optimized transmission conditions and the analysis of the convergence properties of the resulting methods in practical situations. This thesis is organized as follows. First, in Chapter 1, we introduce in detail the optimized Schwarz methods, review why and how they were invented, and present a brief literature overview of recent developments. In Chapter 2, we consider an advection-diffusion equation with constant coefficients, and we derive asymptotic formulas for optimized transmission conditions that are valid when the mesh size h is small. In Chapter 3, for a diffusion problem with a discontinuous coefficient, we thoroughly solve the optimization problems corresponding to several choices of transmission conditions. The asymptotic convergence properties of the resulting methods are analyzed, and we present numerical experiments with two subdomains. This is followed, in Chapter 4, with numerical experiments for the decomposition of a rectangle into many vertical strips. Some additional analysis shows how to adapt the optimized transmission conditions to improve the convergence in that situation. Finally, in Chapter 5, we explore the effectiveness of coarse space corrections for the Schwarz iteration with Robin transmission conditions, when applied to a simple one-dimensional problem.

Chapter 1

Introduction to Optimized Schwarz Methods (OSM)

In this chapter, we introduce the main ideas behind Optimized Schwarz Methods (OSM). We present a short overview of their historical development as well as some of the recent advances, and also review the concept of *equioscillation*, which is used extensively in this work.

1.1 Historical Review and Recent Developments

The origin of domain decomposition methods dates back to 1870, when H. A. Schwarz introduced an iterative procedure for obtaining the solution of a second order elliptic boundary value problem [63]. He considered the potential problem

$$\begin{cases} -\Delta u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{cases} \tag{1.1}$$

where Ω is the union of a disc and a rectangle, as pictured in Figure 1.1. At that time, it was known how to solve the potential problem on a disc or on a rectangle, but not on more complicated geometries. So, to show that the problem (1.1) can be solved, Schwarz proposed an iterative algorithm that involves solving problems on the disc Ω_1 and on the rectangle Ω_2 separately: given an initial guess u_2^0 on Ω_2, the

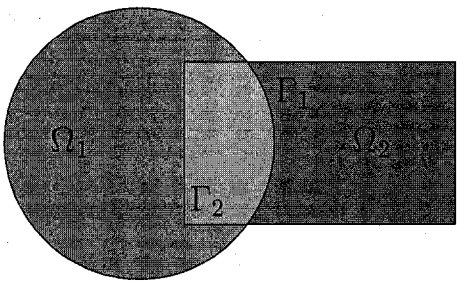

Figure 1.1: The first domain decomposition.

method consists in solving the subproblems

$$\begin{cases} -\Delta u_1^{n+1} &= f & \text{in } \Omega_1, \\ u_1^{n+1} &= 0 & \text{on } \partial\Omega_1 \cap \partial\Omega, \\ u_1^{n+1} &= u_2^n & \text{on } \Gamma_1 := \partial\Omega_1 \cap \Omega_2, \end{cases} \tag{1.2}$$

$$\begin{cases} -\Delta u_2^{n+1} &= f & \text{in } \Omega_2, \\ u_2^{n+1} &= 0 & \text{on } \partial\Omega_2 \cap \partial\Omega, \\ u_2^{n+1} &= u_1^{n+1} & \text{on } \Gamma_2 := \partial\Omega_2 \cap \Omega_1, \end{cases} \tag{1.3}$$

for $n \geq 0$. This is called the *Schwarz alternating method*, or the *Gauss-Seidel Schwarz method*. Schwarz proved convergence of this iteration to a solution of the global problem in the sense that

$$\lim_{n\to\infty} u_j^n = u|_{\Omega_j}, \quad \text{for } j = 1, 2,$$

and used this to show that there exists a unique solution to the problem (1.1), which was the aim of his work.

More than a century later, starting in the mid 1980's, people recognized the potential of Schwarz's idea for parallelizing numerical solvers for boundary value problems, since the iteration can be generalized to many subdomains and involves solving subproblems of smaller sizes. However, to make it a *parallel* algorithm easily extensible to many subdomains, we first need to change one boundary condition in the Schwarz alternating method (1.3) to

$$u_2^{n+1} = u_1^n \qquad \text{on } \Gamma_2,$$

which allows us to solve in both subdomains at the same time to obtain u_1^{n+1} and u_2^{n+1} (see [46]). We will call this new iteration the *parallel Schwarz method* or *Jacobi Schwarz method.* The terminology Gauss-Seidel / Jacobi Schwarz method is sometimes used to point out the natural connection these methods have with classical iterative methods for solving linear systems. Because we are now interested in using this iteration to compute an approximation for the solution of the boundary value problem, we want the iterates u_j^n to approach the solution quickly. The convergence properties of these classical Schwarz methods (alternating and parallel) were thoroughly studied and are now well understood, see the books [64], [60], [65], and references therein. Three main problems arise with these methods:

- they converge only when the subdomains *overlap*,
- for acoustics (e.g. for problems of Helmholtz type), they do not converge, even when there is overlap,
- the convergence is slow when the size of the overlap is small.

These issues stimulated further developments.

In the parallel Schwarz method, Dirichlet conditions are used to communicate data between the subdomains. In general, one can change these to more general transmission conditions,

$$\begin{cases} -\Delta u_1^{n+1} &= f & \text{in } \Omega_1, \\ u_1^{n+1} &= 0 & \text{on } \partial\Omega_1 \cap \partial\Omega, \\ \mathcal{B}_1(u_1^{n+1}) &= \mathcal{B}_1(u_2^n) & \text{on } \Gamma_1, \end{cases}$$

$$\begin{cases} -\Delta u_2^{n+1} &= f & \text{in } \Omega_2, \\ u_2^{n+1} &= 0 & \text{on } \partial\Omega_2 \cap \partial\Omega, \\ \mathcal{B}_2(u_2^{n+1}) &= \mathcal{B}_2(u_1^n) & \text{on } \Gamma_2, \end{cases}$$

where $\mathcal{B}_j$ are transmission operators. Such a modification first appeared in [47], in which Lions proposed a variant of the Schwarz method for non-overlapping subdomains, using the Robin conditions

$$\mathcal{B}_1(u) := \frac{\partial u}{\partial n} + \lambda u, \qquad \mathcal{B}_2(u) := \frac{\partial u}{\partial n} - \lambda u,$$

where $\Gamma := \Gamma_1 = \Gamma_2$, $\frac{\partial}{\partial n}$ is the exterior normal derivative on Γ with respect to Ω_1, and λ is a positive scalar. He proved that this algorithm is convergent for positive definite problems, and also noted that the constant in the Robin conditions could be replaced by a function or by a local or nonlocal operator. For problems of Helmholtz type, Després introduced in [16] a Schwarz iteration using radiation conditions at the interfaces and also proved convergence of the iteration for non-overlapping domain decompositions.

A natural question then comes to mind: how do we choose transmission conditions in the Schwarz method to get *fast* convergence of the iteration? It turns out there exist *optimal* transmission operators for the Schwarz iteration in general, as shown in [55] for a two-dimensional advection-diffusion problem, however they turn out to be non-local operators, which are not convenient and costly to implement in practice. Thus, local approximations for these optimal operators are needed. For this, one idea is to expand the Fourier symbols of the optimal operators with respect to a small viscosity parameter, as proposed in [11], or to expand the symbols with respect to small frequencies to obtain Taylor approximations, see for example [55].

A different strategy to find good choices of local transmission conditions is by *optimizing* the performance of the Schwarz iteration. The main idea is to fix a certain class of local transmission conditions $\mathcal{C}$, and optimize the convergence factor of the iteration over this class. To do this, we need to have an explicit expression for the convergence factor ρ, which is usually defined as a ratio of consecutive errors. Although it is hard to estimate in general, the convergence can be fully analyzed when considering a simple model problem (on the infinite plane, with constant coefficients and two subdomains, for example) and using a Fourier transform to get a convergence factor $\rho(k)$ as a function of the transform variable k, corresponding to frequency components on the interface. We then wish to uniformly minimize this convergence factor over a range of relevant frequencies, i.e. solve the optimization problem

$$\min_{\mathcal{B}_j \in \mathcal{C}} \left(\max_{k \in [k_1, k_2]} \rho(k; \mathcal{B}_j) \right). \tag{1.4}$$

By solving this min-max problem and using the resulting transmission conditions, we

get *Optimized Schwarz Methods* (OSM). This strategy was first introduced in [40] for the advection-diffusion equation, where a certain subclass of second order conditions were optimized for non-overlapping subdomains, see also [41], [45] and [43].

Remark 1.1 (Relevant range of frequencies). *Although the convergence factor ρ is explicitely computed only for a continuous model problem on the infinite plane, we impose bounds on the frequency range by incorporating information about the actual problem we are interested in solving numerically.*

(i) If we wish to compute a solution over a bounded domain with homogeneous Dirichlet boundary conditions, then the minimal frequency component of the solution can be estimated by $k_1 = \frac{\pi}{H}$ where H is the diameter of the subdomain. The lower bound k_1 could also be taken to be the smallest frequency not resolved by the coarse grid, if one is used.

(ii) If the solution is computed numerically on a uniform grid with grid spacing h on the interface, then the maximum frequency which can be represented on this grid is typically estimated by $k_2 = \frac{\pi}{h}$.

Remark 1.2 (Continuous vs. discrete set). *Note that it is also possible to optimize the convergence factor over an appropriate* discrete *set of frequencies between k_1 and k_2; this makes the min-max problem harder to solve analytically, without significantly improving the convergence of the iteration for fine meshes. Thus, most often we will only consider the optimization problem* (1.4)*, set over a continuous range of frequency parameters $[k_1, k_2]$.*

For self-adjoint positive definite problems, optimized Schwarz methods are studied in detail in [32] for both overlapping and non-overlapping decompositions, and the convergence is shown to be greatly improved compared to the classical Schwarz method or to Taylor approximations of the optimal operators. This improvement comes with no additional computational cost per iteration. First, when the overlap size $L > 0$ is constant (i.e. it does not decrease as the mesh is refined), then the

convergence of optimized Schwarz methods is independent of the mesh size h. For non-overlapping subdomains however, the convergence does depend on the mesh size h when choosing the maximum frequency to be $k_2 = \frac{\pi}{h}$. Moreover, when using overlap, it is also relevant to consider the case of a minimal overlap of size h; note that in this case, the parallel Schwarz iteration with Dirichlet transmission conditions has an asymptotic convergence factor of the form $1 - O(h)$ for small h. By analyzing the optimized convergence factor ρ^* for small h, we get the following asymptotic performances (see [32] for proofs):

- Optimized Robin conditions (1 free parameter)

$$\begin{aligned} \text{without overlap:} \quad & \max_{k_1 \leq k \leq \frac{\pi}{h}} \rho^*(k) = 1 - O(h^{\frac{1}{2}}), \\ \text{with overlap of size } h\text{:} \quad & \max_{k_1 \leq k \leq \frac{\pi}{h}} \rho^*(k) = 1 - O(h^{\frac{1}{3}}), \end{aligned}$$

- Optimized two-sided Robin conditions (2 free parameters)

$$\begin{aligned} \text{without overlap:} \quad & \max_{k_1 \leq k \leq \frac{\pi}{h}} \rho^*(k) = 1 - O(h^{\frac{1}{4}}), \\ \text{with overlap of size } h\text{:} \quad & \max_{k_1 \leq k \leq \frac{\pi}{h}} \rho^*(k) = 1 - O(h^{\frac{1}{5}}), \end{aligned}$$

- Optimized second order conditions (2 free parameters)

$$\begin{aligned} \text{without overlap:} \quad & \max_{k_1 \leq k \leq \frac{\pi}{h}} \rho^*(k) = 1 - O(h^{\frac{1}{4}}), \\ \text{with overlap of size } h\text{:} \quad & \max_{k_1 \leq k \leq \frac{\pi}{h}} \rho^*(k) = 1 - O(h^{\frac{1}{5}}). \end{aligned}$$

Thus, the convergence of optimized Schwarz methods is not always mesh independent, however it exhibits only a weak dependence on the mesh size h. Moreover, when using a Schwarz method as a preconditioner for a Krylov iteration, the dependence on h is attenuated further by an additional square root. Also note that, even by using a minimal overlap of size h, the convergence is significantly improved compared to the non-overlapping version, without significant additional computational cost.

For the Helmholtz equation, $-\Delta u - \omega^2 u = f$, the convergence factor of the Schwarz iteration is always 1 for the special frequency $k = \omega$, regardless of the choice of transmission conditions. To optimize the convergence factor in that case, a strategy is to consider a range of frequencies that excludes this special mode, and solve the problem

$$\min_{\mathcal{B}_j \in \mathcal{C}} \left(\max_{k \in [k_1, \omega_-) \cup (\omega_+, k_2]} \rho(k; \mathcal{B}_j) \right).$$

This was first proposed in [12], and the optimization problem was fully solved in [36] for Robin conditions, and in [31] and [34] for second order conditions in the cases without and with overlap respectively.

Remark 1.3. *Recall that the optimized transmission conditions are derived for a model problem with constant coefficients. For a problem that has smoothly varying coefficients in the domain, optimized transmission conditions can be computed using a "frozen coefficient" approach: at each grid point on the interface, we use the local value of the coefficients to compute optimized parameters in the transmission conditions. Hence, this amounts to solving a min-max problem for* each *grid point on the interface.*

As described above, optimized Schwarz methods are typically analyzed by using a Fourier transform or Fourier series, and thus the convergence analysis only applies to rectangular subdomains with a straight interface. In [49], the Schwarz method with Robin conditions, for two non-overlapping subdomains of general shape, is analyzed using the theory of Steklov-Poincaré operators. A sharp optimal estimate is derived for the condition number of the preconditioned operator; it is of the form $O(h^{-\frac{1}{2}})$ when the Robin parameter is chosen to be $O(h^{\frac{1}{2}})$, which agrees with the analysis of optimized Schwarz methods on rectangular subdomains. This general theory indicates that, by using the optimized parameters computed for a model problem (say on a rectangular geometry with a straight interface) in a Schwarz iteration applied to a more general problem with a curved interface, the asymptotic convergence factor of the iteration for small mesh sizes h will still be in the same form predicted by the analysis on the model problem. In addition, in [48], a similar analysis is performed

for a type of second order conditions with two free parameters, in which case an optimal condition number estimate of the form $O(h^{-\frac{1}{4}})$ is proved. For the case of many non-overlapping subdomains, in [59] the authors analyze the Schwarz method with Robin conditions and prove an estimate for the convergence factor of the form $1 - O(h^{\frac{1}{2}}H^{-\frac{1}{2}})$ when the constant Robin parameter is chosen to be $O(h^{-\frac{1}{2}}H^{-\frac{1}{2}})$ (H denotes the diameter of a subdomain). However, this estimate only holds under some specific restrictive assumptions about the domain decomposition. More precisely, if we define the distance from a subdomain to the boundary as the minimum number of subdomains that we need to travel through to reach the boundary, then this distance has to be bounded for all subdomains as the mesh and domain decomposition are refined.

Optimized Schwarz methods with Robin or more general transmission conditions can also be implemented for non-conforming grids; see [1] and [62] for the finite volume error analysis, and [35] for the finite element case.

When solving an elliptic problem on a non-convex polygonal domain, or when the domain decomposition introduces non-convex polygonal subdomains, there are singularities occurring in the solution at reentrant corners. Optimized Schwarz methods for such problems are studied in [13], and adapted interface conditions are proposed to be used near these corners.

The idea of optimized Schwarz methods can be applied for more complex problems, such as the compressible Euler equations [19, 20], the Maxwell equations [18] or the shallow-water equations on the sphere [58]. Optimized interface conditions for problems in highly heterogeneous media are also of interest, and will be the subject of Chapter 3.

1.2 The Equioscillation Property

To compute optimized transmission conditions, we want to solve problems of the form (1.4). These types of min-max problems are closely related to the Chebyshev best approximation problem, which is stated as follows: given a function $f \in C([a,b])$, find

the polynomial $p_n^*(x)$ of degree at most n, of best approximation in the L^∞ norm,

$$\max_{x\in[a,b]} |p_n^*(x) - f(x)| = \inf_{p_n\in\mathbb{P}_n} \left(\max_{x\in[a,b]} |p_n(x) - f(x)| \right).$$

Theorem 1.1 (Chebyshev Equioscillation Theorem). *Given $f \in C([a,b])$, there exists a unique polynomial of best approximation p_n^*, and it is characterized by an* equioscillation property*: there are $n+2$ distinct points $x_1, x_2, \ldots, x_{n+2} \in [a,b]$ such that*

$$|p_n^*(x_j) - f(x_j)| = \max_{x\in[a,b]} |p_n^*(x) - f(x)|, \quad \text{for } j = 1, 2, \ldots, n+2. \tag{1.5}$$

For a detailed exposition of this classical result of approximation theory, including numerical methods for computing the polynomial of best approximation, see [52].

For optimized one-sided transmission conditions, the min-max problem can be written in the form

$$\inf_{p\in\mathbb{P}_n} \left(\max_{k\in[k_1,k_2]} \left| \frac{p(k) - f(k)}{p(k) + f(k)} e^{-Lf(k)} \right| \right), \tag{1.6}$$

where $f(k)$ is a given function depending on the differential equation, $n = 0$ for Robin conditions and $n = 2$ for second order conditions. This has two main differences with the Chebyshev best approximation problem. First, we are not minimizing the difference $|p(k) - f(k)|$ but a rational expression weighted by an exponential which depends on $f(k)$. Secondly, the function $f(k)$ can be a complex-valued function, as in the case of the advection-diffusion equation. Even with these important differences, when solving analytically problems like (1.6) for optimized transmission conditions, we often find that the solution satisfies an equioscillation property analogous to (1.5) (see for example the theorems in [32]).

Definition 1.2. We say that the solution of the min-max problem (1.4) satisfies an *equioscillation property* if the optimized convergence factor attains its maximum at exactly $m+1$ points, where m is the number of free parameters in the transmission conditions.

The equioscillation property is illustrated in Figure 1.2, where the optimized convergence factor is shown in several cases. Some work has been done to try to prove a

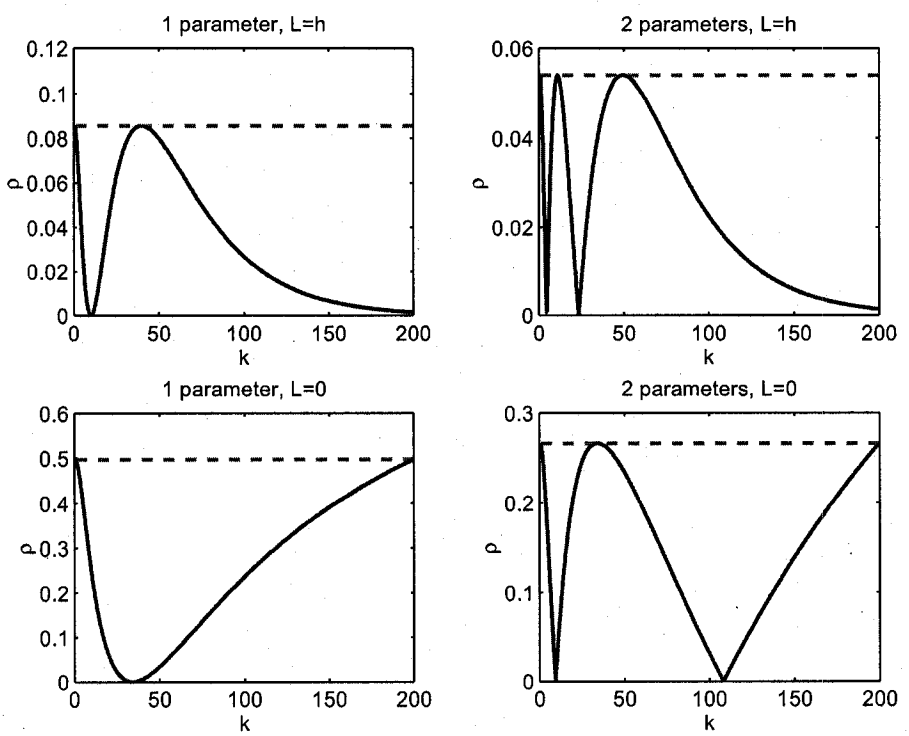

Figure 1.2: The equioscillation property for the optimized convergence factor, when $k_1 = 1$, $k_2 = \frac{\pi}{h} = 200$, with and without overlap.

generalization of the Chebyshev equioscillation theorem, for example in [6], where the results are successfully applied to computing optimized transmission conditions for the Schwarz waveform relaxation method, for parabolic problems. However, it should be noted that the solution of min-max problems like (1.6) is *not always* characterized by an equioscillation property: a counterexample is given by optimized Robin conditions for the advection-diffusion equation in [22]. Thus, in general, the equioscillation property should be proved (or disproved) before using it as a basis for computing optimized parameter values.

In addition of establishing when the equioscillation property holds or not, there are several reasons to be studying the min-max problems analytically:

(i) We need to prove the existence and uniqueness of a global minimizer for the min-max problem.

(ii) We want to study analytically the asymptotic behavior of the optimized convergence factor for small mesh sizes h or when varying the coefficients.

(iii) Fully solving the min-max problems leads to explicit formulas or simple characterizations for the optimized parameters that are easy to implement in solvers. This also permits more efficient calculations when the coefficients of the differential equation are varying, since many min-max problems need to be solved in that case, see Remark 1.3.

(iv) When solving the min-max problems *numerically*, there are some pitfalls. If we use a method for unconstrained minimization such as Nelder-Mead [57], there is no guarantee of convergence in higher dimensions, and the algorithm may converge to a local, non-global minimum depending on the initial guess.

Chapter 2

Optimized Schwarz Methods for Advection-Diffusion Problems

In this chapter, we consider a model advection-diffusion problem with constant coefficients. We first review the known results concerning optimized Schwarz methods, and we demonstrate why it is difficult to fully solve the min-max problems. We then provide asymptotic formulas for the optimized parameters, which are valid when the mesh size h is small enough, and the component of the advection tangential to the interface is not too large.

2.1 Preliminaries

Consider the advection-diffusion equation with constant coefficients in two dimensions, on the infinite plane,

$$\begin{cases} \mathcal{L}(u) := -\nu\Delta u + \mathbf{a}\cdot\nabla u + cu &= f \quad \text{on } \Omega = \mathbb{R}^2, \\ |u| &< \infty \quad \text{as } |\mathbf{x}| \to \infty, \end{cases} \tag{2.1}$$

where $\nu > 0$, $\mathbf{a} := (a, b)$, $c \geq 0$ are real constants. The domain is decomposed into two subdomains

$$\Omega_1 = (-\infty, L) \times \mathbb{R} \quad \text{and} \quad \Omega_2 = (0, \infty) \times \mathbb{R},$$

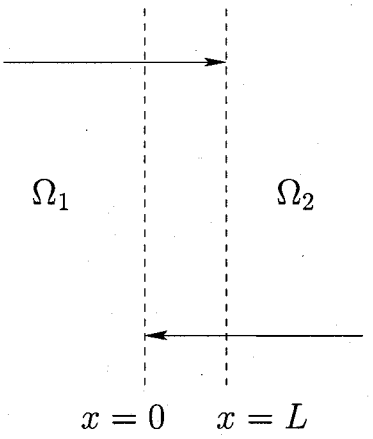

Figure 2.1: Domain decomposition for the model problem.

with an overlap of size $L \geq 0$. We analyze the convergence of the general Schwarz iteration

$$\begin{cases} \mathcal{L}(u_1^{n+1}) = f & \text{in } \Omega_1, \\ |u_1^{n+1}| < \infty & \text{at infinity}, \\ \left(\nu\dfrac{\partial u_1^{n+1}}{\partial x} - \dfrac{a}{2}u_1^{n+1}\right) + \dfrac{1}{2}\mathcal{S}_1(u_1^{n+1}) = \left(\nu\dfrac{\partial u_2^{n}}{\partial x} - \dfrac{a}{2}u_2^{n}\right) + \dfrac{1}{2}\mathcal{S}_1(u_2^{n}) & \text{at } x = L, \end{cases} \tag{2.2}$$

$$\begin{cases} \mathcal{L}(u_2^{n+1}) = f & \text{in } \Omega_2, \\ |u_2^{n+1}| < \infty & \text{at infinity}, \\ \left(\nu\dfrac{\partial u_2^{n+1}}{\partial x} - \dfrac{a}{2}u_2^{n+1}\right) - \dfrac{1}{2}\mathcal{S}_2(u_2^{n+1}) = \left(\nu\dfrac{\partial u_1^{n}}{\partial x} - \dfrac{a}{2}u_1^{n}\right) - \dfrac{1}{2}\mathcal{S}_2(u_1^{n}) & \text{at } x = 0, \end{cases} \tag{2.3}$$

where $\mathcal{S}_j$ are linear operators acting in the tangential direction to the interface Γ_j. It is sufficient to consider the homogeneous problem $f \equiv 0$ for the convergence analysis. For this simple model problem, we can take a Fourier transform in the variable y,

$$\mathcal{F}_y(u(x,y)) := \int_{\mathbb{R}} u(x,y)e^{-iyk}\,dy = \hat{u}(x,k),$$

and we assume that the operators $\mathcal{S}_j$ have Fourier symbols $\sigma_j(k)$, namely that

$$\mathcal{F}_y\left(\mathcal{S}_j(u)\right) = \sigma_j(k)\hat{u}(x,k).$$

In Fourier space, the differential equation becomes

$$\hat{\mathcal{L}}(\hat{u}) = -\nu\frac{\partial^2\hat{u}}{\partial x^2} + a\frac{\partial\hat{u}}{\partial x} + \left(c + \nu k^2 - ibk\right)\hat{u} = 0.$$

Solving this ODE in x for each subdomain and using the boundedness conditions at infinity (u_1 is bounded as $x \to -\infty$ and u_2 is bounded as $x \to \infty$), we find that the solutions must be in the form

$$u_1^n(x,k) = A^n(k)e^{\lambda_+(k)(x-L)}, \quad u_2^n(x,k) = B^n(k)e^{\lambda_-(k)x}, \quad \text{for } n = 1, 2, \ldots$$

where the characteristic roots $\lambda_\pm(k)$ are defined by

$$\lambda_\pm(k) := \frac{-a \pm \sqrt{a^2 + 4\nu c + 4\nu^2k^2 - 4i\nu bk}}{2\nu}.$$

From now on, we will use the notation

$$z(k) := \sqrt{a^2 + 4\nu c + 4\nu^2k^2 - 4i\nu bk}, \quad A := a^2 + 4\nu c,$$

$$\xi(k) := \text{real}(z(k)), \quad \eta(k) := \text{imag}(z(k)).$$

In this context, we define the convergence factor of the iteration, in Fourier space, as

$$\rho(k,\, \sigma_j) := \left|\frac{u_1^{n+1}(L,k)}{u_1^{n-1}(L,k)}\right| = \left|\frac{A^{n+1}(k)}{A^{n-1}(k)}\right| = \left|\frac{u_2^{n+1}(0,k)}{u_2^{n-1}(0,k)}\right| = \left|\frac{B^{n+1}(k)}{B^{n-1}(k)}\right|,$$

for any $n \geq 1$. By applying the transmission conditions of the general Schwarz method (2.2)-(2.3), we can derive an expression for the convergence factor, namely

$$\rho(k,\, \sigma_j) = \left|\frac{\sigma_1(k) - z(k)}{\sigma_1(k) + z(k)} \cdot \frac{\sigma_2(k) - z(k)}{\sigma_2(k) + z(k)} e^{-\frac{L}{\nu}z(k)}\right|. \tag{2.4}$$

2.1.1 Previous Work on Optimized Schwarz Methods

It is clear by inspection of (2.4) that there is an optimal choice of transmission conditions making the convergence factor vanish for all k, namely

$$\sigma_1^{opt}(k) = \sigma_2^{opt}(k) = z(k).$$

However, because $z(k)$ is not a polynomial in k, the corresponding operators in real space $\mathcal{S}_j^{opt}$ are non-local in y (they are Dirichlet-to-Neumann operators), and thus are not convenient for implementation, see [55]. Also in [55], the authors propose *local* transmission conditions by using low order Taylor approximations of the function $z(k)$ around $k = 0$. For example, the zeroth order and second order Taylor approximations give respectively

$$\sigma_1^{(0)}(k) = \sigma_2^{(0)}(k) = \sqrt{A},$$
$$\sigma_1^{(2)}(k) = \sigma_2^{(2)}(k) = \sqrt{A} - \frac{b}{\sqrt{A}}(ik) + \frac{\nu}{\sqrt{A}}\left(1 + \frac{b^2}{A}\right)k^2.$$

However, this strategy will not be effective for all frequencies k. For both of these choices of Taylor approximations, the asymptotic convergence factor (for small mesh size h) is of the form $1 - O(h)$ without overlap, and $1 - O(h^{\frac{1}{2}})$ with an overlap of size $O(h)$.

In her thesis [40], Japhet proposed to use second order transmission conditions for non-overlapping subdomains ($L = 0$), which correspond to the choice of operators

$$\mathcal{S}_1 = \mathcal{S}_2 = p + q\frac{\partial}{\partial y} - r\frac{\partial^2}{\partial y^2},$$

where $p, q, r > 0$ are real parameters. The full optimization of the convergence factor on the three free parameters was not solved in the work of Japhet, because of its complexity. Instead, the first parameter is fixed to the value $p = \sqrt{A}$, and an extra relation between q and r is assumed in order to reduce the optimization to only one free parameter (this relation is based on the optimization result in the case of a purely normal advection). The asymptotic convergence factor of the resulting Schwarz method is of the form $1 - O(h^{\frac{1}{3}})$, which is much better than the performance obtained from Taylor approximations of the optimal symbol $z(k)$

More recently, in [21], simpler Robin transmission conditions were studied, corresponding to the choice $\mathcal{S}_1 = \mathcal{S}_2 = p \in \mathbb{R}$. To optimize the value of the parameter p, we wish to solve the min-max problem

$$\min_{p\in\mathbb{R}} \left(\max_{k_1 \le k \le k_2} \left| \frac{p - z(k)}{p + z(k)} \right|^2 e^{-\frac{L}{\nu}\xi(k)} \right). \tag{2.5}$$

For the non-overlapping case, this min-max problem was fully solved.

Theorem 2.1 (Optimized Robin parameter, without overlap). *If there is no overlap ($L = 0$), the unique solution p^* of the min-max problem (2.5) is given by*

$$p^* = \begin{cases} |z(k_1)| & \text{if } p_c < |z(k_1)|, \\ p_c & \text{if } |z(k_1)| \leq p_c \leq |z(k_2)|, \\ |z(k_2)| & \text{if } p_c > |z(k_2)|, \end{cases}$$

$$\text{where } p_c := \sqrt{\frac{\xi(k_1)|z(k_2)|^2 - \xi(k_2)|z(k_1)|^2}{\xi(k_2) - \xi(k_1)}}.$$

Proof. The complete proof can be found in [21]. The first step consists in writing the imaginary part $\eta(k)$ of $z(k)$ as a function of the real part $\xi(k)$, by eliminating k, which gives

$$\eta^2 = b^2 \frac{\xi^2 - A}{\xi^2 + b^2},$$

and then writing the convergence factor as a function of ξ instead of k,

$$\rho(\xi, p) = \frac{(\xi^2 + b^2)(p - \xi)^2 + b^2(\xi^2 - A)}{(\xi^2 + b^2)(p + \xi)^2 + b^2(\xi^2 - A)}.$$

Then, the hardest step of the proof is showing that there is at most one critical point of ρ as a function of ξ, for $\xi \geq \xi(k_1)$. This reduces to showing that the cubic polynomial

$$P_3(X) := X^3 + (b^2 - p^2)X^2 + 2b^2\left(b^2 + \frac{3}{2}A - p^2\right)X + b^4(A - p^2)$$

has a unique real root for $X \geq \xi(k_1)^2$. □

This theorem also shows that the equioscillation property, in the sense of Definition 1.2, does not always hold. When $p_c \notin [\,|z(k_1)|, |z(k_2)|\,]$ (which does occur for some values of k_1 and k_2) the maximum of the optimized convergence factor is attained at only one of the endpoints, k_1 or k_2.

2.2 Analysis for Optimized Robin Conditions with Overlap

For the overlapping case, in [21] an approximation for the optimized Robin parameter is proposed: the imaginary part $\eta(k)$ is replaced by the constant $-b$ (which is a good approximation for large k). This gives an upper bound for the convergence factor, and this upper bound can be uniformly minimized instead to obtain a Robin parameter. However, this approximation for the optimized Robin parameter is not very good when the component of the advection tangential to the interface, b, is large. Thus, it is desired to fully solve the min-max problem (2.5) for a general advection and a general overlap size $L > 0$. In this section, we present some partial analysis of the min-max problem, which supports a conjecture we make concerning the optimized Robin parameter.

Suppose we choose $k_1 = 0$ and $k_2 = \infty$ for simplicity. If we want to imitate the proof of Theorem 2.1 to get the optimized Robin parameter, we first rewrite the min-max problem in terms of ξ instead of k,

$$\min_{p\in\mathbb{R}} \left(\max_{\sqrt{A}\leq\xi<\infty} \rho(\xi, p) \right) \tag{2.6}$$

where

$$\rho(\xi, p) := \frac{(\xi^2 + b^2)(p - \xi)^2 + b^2(\xi^2 - A)}{(\xi^2 + b^2)(p + \xi)^2 + b^2(\xi^2 - A)} e^{-\frac{L}{\nu}\xi}.$$

Conjecture 2.2.1. For $L > 0$ small enough, the unique minimizer p^* of problem (2.6) satisfies the equioscillation property

$$\rho(\sqrt{A}, p^*) = \rho(\xi_c(p^*), p^*),$$

where $\xi_c(p)$ is the location of the unique local maximum of $\rho(\xi, p)$ (when it exists) as a function of ξ.

The two main difficulties in proving this conjecture are:

(1) the equioscillation only holds when L is *small enough*, and how small L needs to be depends on b for example,

(2) it is hard to obtain an explicit formula for the critical point $\xi_c(p)$.

In the following, we present what we know about problem (2.6) to provide some evidence in favor of Conjecture 2.2.1.

Restriction of parameter space

First, we can observe that $\rho(\xi,p) < \rho(\xi,-p)$ for any $p > 0$. Thus we can immediately rule out the possibility that the optimized Robin parameter is negative. Now, taking a derivative of $\rho(\xi,p)$ with respect to the parameter p, we find that

$$\begin{aligned} \frac{\partial \rho}{\partial p}(\xi,p) &< 0 \qquad \text{for } 0 \le p < |z(k)| = \sqrt{\xi^2 + \eta(\xi)^2}, \\ \frac{\partial \rho}{\partial p}(\xi,p) &> 0 \qquad \text{for } p > |z(k)|. \end{aligned}$$

Note that $|z(k)|$ is an increasing function of k, and $|z(k)| \ge \sqrt{A}$ for $k \ge 0$. So, if $0 \le p < \sqrt{A}$, then the convergence factor can be uniformly improved for $\xi \in [\sqrt{A}, \infty)$ by increasing p. Hence, we can restrict the range of p in the min-max problem to the interval

$$p \in [\sqrt{A}, \infty).$$

In the general case when $k_1 > 0$ and $k_2 < \infty$, the search for the optimized Robin parameter can be narrowed to the interval $[\,|z(k_1)|, |z(k_2)|\,]$.

Search for critical points in ξ

Now, we would like to find the critical points of the function $\rho(\xi,p)$ with respect to the variable ξ. Taking a partial derivative of the convergence factor in ξ, and setting it to 0, we find that

$$\frac{\partial \rho}{\partial \xi}(\xi,p) = 0 \quad \Longleftrightarrow \quad Q(\xi^2) = 0,$$

where $Q(Y)$ is a polynomial of degree 4 given by

$$\begin{aligned} Q(Y) := & -LY^4 + \left(2Lp^2 - 4Lb^2 + 4\nu p\right) Y^3 \\ & + \left(-Lp^4 + 2Lp^2b^2 + 4\nu pb^2 - 4\nu p^3 - 4Lb^4 + 2LAb^2\right) Y^2 \\ & + 2b^2\left(2LAb^2 + 4\nu pb^2 + 6\nu pA - Lp^4 + Lp^2A - 4\nu p^3\right) Y \\ & - \left(p^2 - A\right) b^4\left(-AL + 4\nu p + Lp^2\right). \end{aligned} \tag{2.7}$$

Writing down formulas for the roots of this quartic is not a simple task, since the coefficients are fairly complex and depend on many parameters. In the end, we would like to prove a result resembling the following.

Conjecture 2.2.2. The polynomial $Q(Y)$ has at most two real roots in the interval $Y = \xi^2 \geq A$, for any allowed choice of parameter values (i.e. $p \geq \sqrt{A}$). The larger root, denoted by $Y_c(p) = \xi_c^2(p)$, corresponds to a local maximum of the convergence factor, whereas the other corresponds to a local minimum (when present).

This conjecture would be very useful in that it would allow us to write the maximum of the convergence factor as

$$\max_{\sqrt{A} \leq \xi < \infty} \rho(\xi, p) = \max\left(\rho(\sqrt{A}, p),\ \rho(\xi_c^2(p), p)\right). \tag{2.8}$$

In an attempt to prove this conjecture, we can simplify the expression for the quartic polynomial (2.7) a little by rescaling the variables appropriately with b, in order to reduce the number of free parameters. Let

$$X := \frac{\xi^2}{b^2}, \qquad \alpha^2 := \frac{A}{b^2}, \qquad P := \frac{p}{|b|}, \qquad \beta := \frac{|b|L}{\nu}.$$

We obtain that

$$\frac{\partial \rho}{\partial \xi}(\xi, p) = 0 \quad \Longleftrightarrow \quad \Pi(X) = 0,$$

where the polynomial $\Pi(X)$ is now

$$\begin{aligned} \Pi(X) := & -\beta X^4 + \left(4P - 4\beta + 2\beta P^2\right) X^3 \\ & + \left(-4\beta - \beta P^4 + 2\beta\alpha^2 - 4P^3 + 4P + 2\beta P^2\right) X^2 \\ & + \left(-8P^3 + 4\beta\alpha^2 + 8P - 2\beta P^4 + 2\beta P^2\alpha^2 + 12P\alpha^2\right) X \\ & - \left(P - \alpha\right)\left(P + \alpha\right)\left(\beta P^2 + 4P - \beta\alpha^2\right). \end{aligned}$$

This rescaling of the coefficients has reduced the 5 parameters (ν, a, b, c, L), to only 2 nonnegative parameters $(\alpha \geq 0, \beta \geq 0)$.

Behavior at $P = \alpha$

We start with a simple situation, when the Robin parameter is at its minimum value, i.e. $P = \alpha$. In this case, the polynomial $\Pi(X)$ reduces to

$$\begin{aligned}\Pi(X) = &-\beta X^4 + (4\alpha - 4\beta + 2\beta\alpha^2)X^3 \\ &+ (-4\beta - \beta\alpha^4 + 4\beta\alpha^2 - 4\alpha^3 + 4\alpha)X^2 \\ &+ 4\alpha(\alpha^2 + \alpha\beta + 2)X.\end{aligned}$$

In this section, we will prove that Conjecture 2.2.2 holds when $P = \alpha$. The strategy is to deduce the number of roots for $X \geq \alpha^2$ by only looking at the sign of each derivative of Π at the point $X = \alpha^2$,

$$\begin{aligned}\Pi(\alpha^2) &= 8\alpha^5 + 8\alpha^3 > 0, \qquad \forall\, \alpha > 0, \\ \Pi'(\alpha^2) &= 4\alpha(\alpha^2 + 1)(\alpha^2 - \alpha\beta + 2), \\ \Pi''(\alpha^2) &= -2\beta\alpha^4 + 16\alpha^3 - 16\beta\alpha^2 - 8\beta + 8\alpha, \\ \Pi'''(\alpha^2) &= -12\beta\alpha^2 + 24\alpha - 24\beta.\end{aligned}$$

We want to get information about the sign of each derivative $\Pi^{(j)}(\alpha^2)$ for any values of α and β. To do so, we look at the equations $\Pi^{(j)}(\alpha^2) = 0$ to find the locations where the sign could change, and we observe that all the equations are *linear* in β. Hence, we can solve these equations for β in terms of α, and obtain curves $\beta_j(\alpha)$ in the (α, β) plane such that

$$\begin{aligned}\Pi^{(j)}(\alpha^2) = 0 \quad &\Longleftrightarrow \quad \beta = \beta_j(\alpha), \\ \Pi^{(j)}(\alpha^2) > 0 \quad &\Longleftrightarrow \quad \beta < \beta_j(\alpha), \\ \Pi^{(j)}(\alpha^2) < 0 \quad &\Longleftrightarrow \quad \beta > \beta_j(\alpha).\end{aligned}$$

This gives the three curves

$$\beta_1(\alpha) = \alpha + \frac{2}{\alpha}, \qquad \beta_2(\alpha) = \frac{4\alpha(1 + 2\alpha^2)}{\alpha^4 + 8\alpha^2 + 4}, \qquad \beta_3(\alpha) = \frac{2\alpha}{\alpha^2 + 2}.$$

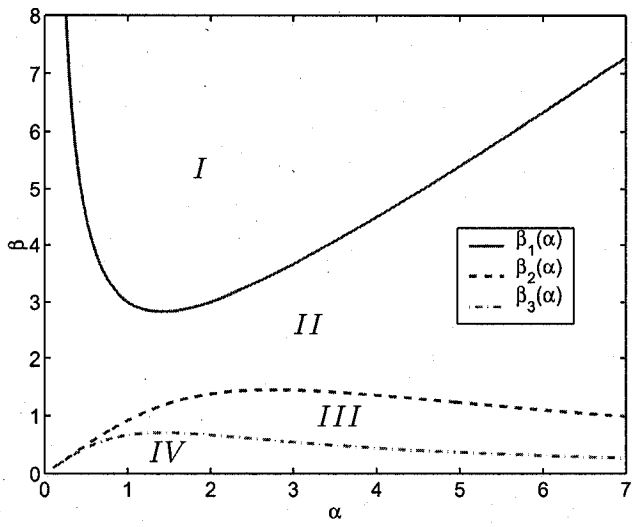

Figure 2.2: The curves $\beta_j(\alpha)$ when $P = \alpha$.

Figure 2.2 shows a plot of these functions for $\alpha > 0$. It seems that the three curves don't intersect; this is confirmed by the next proposition.

Proposition 2.1.

$$\beta_1(\alpha) > \beta_2(\alpha) > \beta_3(\alpha) > 0 \quad \forall\, \alpha > 0.$$

Proof. First, note that all the functions are smooth for $\alpha > 0$ (they have no real poles) and that they are all positive. Now, $\beta_1(\alpha)$ goes to infinity as α approaches infinity, and as α approaches 0^+. The minimum of β_1 is attained at $\alpha = \sqrt{2}$ and has the value $2\sqrt{2}$. Now the equation

$$2\sqrt{2} - \beta_2(\alpha) = 0$$

leads to a polynomial of degree 4 with no real roots. Also, $2\sqrt{2} > \beta_2(0) = 0$, thus we conclude that

$$\beta_1(\alpha) \geq 2\sqrt{2} > \beta_2(\alpha).$$

For the last inequality, we look at the difference $\beta_2(\alpha) - \beta_3(\alpha)$, and by putting everything on a common denominator, we obtain

$$\beta_2(\alpha) - \beta_3(\alpha) = \frac{2\alpha^3(3\alpha^2 + 2)}{(\alpha^2 + 2)(\alpha^4 + 8\alpha^2 + 4)} > 0, \quad \forall\, \alpha > 0.$$

Region	sign of $\Pi(\alpha^2)$	sign of $\Pi'(\alpha^2)$	sign of $\Pi''(\alpha^2)$	sign of $\Pi'''(\alpha^2)$
I	+	-	-	-
II	+	+	-	-
III	+	+	+	-
IV	+	+	+	+

Table 2.1: Sign of the derivatives of Π evaluated at α^2.

Therefore, $\beta_2(\alpha) > \beta_3(\alpha)$ for all $\alpha > 0$. □

This proposition shows that the curves β_j divide the quadrant $\{\alpha > 0, \beta > 0\}$ into 4 regions, as illustrated on Figure 2.2. The sign of each derivative at α^2 in each region is shown in Table 2.1. With this information, we are able to prove Conjecture 2.2.2 for the specific parameter value $P = \alpha$. More precisely, we show

Proposition 2.2. *When $P = \alpha$, $\Pi(X)$ has exactly one real root for $X > \alpha^2$, and it represents a local maximum of the convergence factor ρ.*

Proof. First note that $\Pi(\alpha^2) > 0$, and $\Pi(X) \to -\infty$ as $X \to \pm\infty$. Thus, $\Pi(X)$ has at least 2 real roots on $\mathbb{R}$, one on each side of $X = \alpha^2$. If $\Pi(X)$ has a pair of complex roots, we are done. Otherwise, it has 4 real roots. Now, Π''' is linear and

$$\lim_{X\to-\infty} \Pi'''(X) = \infty, \quad \lim_{X\to\infty} \Pi'''(X) = -\infty,$$

so the third derivative changes sign once, from positive to negative.

Consider the following two scenarios. Suppose $X^* \geq \alpha^2$ is such that

- *Scenario A*

$$\Pi(X^*) > 0, \quad \Pi'(X^*) \leq 0, \quad \Pi''(X^*) < 0, \quad \Pi'''(X^*) < 0.$$

- *Scenario B*

$$\Pi(X^*) > 0, \quad \Pi'(X^*) > 0, \quad \Pi'''(X^*) < 0.$$

In Scenario A, the third derivative stays negative for all $X \geq X^*$, and thus the second and first derivatives also remain negative. Hence, $\Pi(X)$ is strictly decreasing, and it has exactly one root for $X \geq X^*$.

In Scenario B, $\Pi'(X^*) > 0$, so we let X increase until we reach an X^{**} such that

$$\Pi'(X^{**}) = 0 \quad (\text{then we can deduce} \quad \Pi'(X) < 0 \quad \forall X > X^{**}).$$

In the interval $[X^*, X^{**}]$, $\Pi(X)$ is increasing so we do not encounter a root, and $\Pi'''(X)$ remains negative. Because Π' goes from positive to negative at X^{**}, we have that $\Pi''(X^{**}) < 0$. In summary, at the point X^{**}, we have

$$\Pi(X^{**}) > 0, \quad \Pi'(X^{**}) \leq 0, \quad \Pi''(X^{**}) < 0, \quad \Pi'''(X^{**}) < 0,$$

i.e. we fall into the Scenario A at $X = X^{**}$. Thus we have exactly one root for $X \geq X^{**} > X^*$.

Now it is clear that if (α, β) is in region I (see Table 2.1), then we are in Scenario A at $X^* = \alpha^2$. If (α, β) is in region II or III, then we are in Scenario B at $X^* = \alpha^2$. So for these three regions, the above argument proves the proposition.

It remains to deal with region IV. In that case, it is clear that as we increase X from α^2, Π, Π' and Π'' will remain positive until the third derivative changes sign. So we can increase X to a value X^* such that

$$\Pi(X^*) > 0, \quad \Pi'(X^*) > 0, \quad \Pi''(X^*) > 0, \quad \Pi'''(X^*) < 0,$$

without finding any root of Π. Then we fall into Scenario B, and can conclude that we have exactly one root for $X \geq \alpha^2$. □

Increasing $P > \alpha$

The previous arguments dealt with the special situation when $P = \alpha$, but we wish to prove Conjecture 2.2.2 for all $P \geq \alpha$. We ask the question: what can happen as we increase P from its minimum value α? To observe the behavior as P increases, we write

$$P = \alpha e, \quad \text{for } e \geq 1.$$

The polynomial can now be rewritten as

$$\begin{aligned}\Pi(X) = &-\beta X^4 + \left(4\,e\alpha - 4\,\beta + 2\,\beta\,e^2\alpha^2\right) X^3 \\ &+ \left(-4\,\beta - \beta\,e^4\alpha^4 + 2\,\beta\,\alpha^2 - 4\,e^3\alpha^3 + 4\,e\alpha + 2\,\beta\,e^2\alpha^2\right) X^2 \\ &- 2\,\alpha\left(\alpha^3\beta\,e^4 - \alpha^3\beta\,e^2 + 4\,\alpha^2 e^3 - 6\,e\alpha^2 - 2\,\beta\,\alpha - 4\,e\right) X \\ &- \alpha^3\,(e-1)\,(e+1)\left(\beta\,\alpha\,e^2 + 4\,e - \beta\,\alpha\right).\end{aligned}$$

We adopt a similar strategy as for the case $P = \alpha$. Seeing as all the coefficients of $\Pi(X)$ are linear in β, we can find curves $\beta_j(\alpha;\, e)$ such that

$$\begin{aligned}\Pi^{(j)}(\alpha^2) = 0 \quad &\Longleftrightarrow \quad \beta = \beta_j(\alpha;\, e), \\ \Pi^{(j)}(\alpha^2) > 0 \quad &\Longleftrightarrow \quad \beta < \beta_j(\alpha;\, e), \\ \Pi^{(j)}(\alpha^2) < 0 \quad &\Longleftrightarrow \quad \beta > \beta_j(\alpha;\, e).\end{aligned}$$

This yields the following functions

$$\begin{aligned}\beta_0(\alpha;\, e) &= \frac{-4e\left[\alpha^2(e^2-1) + e^2 - 3\right]}{\alpha\,(\alpha^2+1)\,(e^2-1)^2}, \\ \beta_1(\alpha;\, e) &= \frac{-2e\,(2\,e^2\alpha^2 - 3\,\alpha^2 - 2)}{\alpha\,[\alpha^2(e^2-1)(e^2-2) + 2]}, \\ \beta_2(\alpha;\, e) &= \frac{-4e\alpha\,\left[\alpha^2(e^2-3) - 1\right]}{e^4\alpha^4 - 6\,e^2\alpha^4 + 6\,\alpha^4 - 2\,e^2\alpha^2 + 10\,\alpha^2 + 4}, \\ \beta_3(\alpha;\, e) &= \frac{-2e\alpha}{\alpha^2(e^2-2) - 2}.\end{aligned}$$

When $P = e\alpha$, these curves in the (α, β)-plane have various crossings between them, and sometimes have real poles. Moreover, the α-coordinates of these crossings and poles move around when increasing e. If we want to determine the number of roots larger than α^2 based solely on the signs of the derivatives of the polynomial at α^2, as in the case $P = \alpha$, we now have to consider all the cases listed in Table 2.2. Some more analysis of the curves $\beta_j(\alpha;\, e)$ allows us to eliminate many cases. Unfortunately, this analysis does not permit us to eliminate the cases V and XIV which represent relatively small regions in the (α, β)-plane, see Figure 2.3. For values of α and β in these regions, we cannot conclude that the quartic polynomial $\Pi(X)$

Region	$\Pi(\alpha^2)$	$\Pi'(\alpha^2)$	$\Pi''(\alpha^2)$	$\Pi'''(\alpha^2)$	number of roots
I	+	-	-	-	1
II	+	+	-	-	1
II	+	+	+	-	1
IV	+	+	+	+	1
V	+	+	-	+	1 or 3
VI	+	-	-	+	1 or 3
VII	+	-	+	-	1 or 3
VIII	+	-	+	+	1 or 3
IX	-	-	-	-	0
X	-	-	-	+	0 or 2
XI	-	-	+	-	0 or 2
XII	-	-	+	+	0 or 2
XIII	-	+	-	-	0 or 2
XIV	-	+	-	+	0, 2 or 4
XV	-	+	+	-	0 or 2
XVI	-	+	+	+	0 or 2

Table 2.2: Sign of the derivatives of $\Pi(X)$ evaluated at α^2, and the associated possible number of roots for $X > \alpha^2$, for a general quartic polynomial $\Pi(X)$.

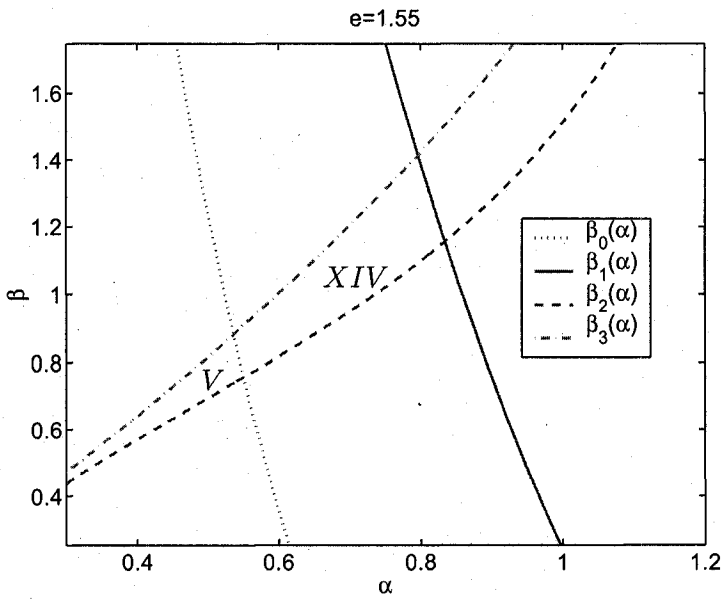

Figure 2.3: The curves when $P = 1.55\alpha$.

has at most two real roots larger than α^2 by solely taking into account the signs of the derivatives at α^2.

On the other hand, by scanning these small regions and computing numerically the roots of the quartic for many values for α and β, we couldn't find a case when we have more than two real roots larger α^2. It is possible to get four positive real roots, but two of them are always less than α^2. So, it would seem that Conjecture 2.2.2 holds, but the analysis presented here could not completely prove it. And without a result such as (2.8), the proof of the equioscillation property that we expect in Conjecture 2.2.1 cannot be completed.

We also include here the following proposition, which could potentially be useful: it states that exactly one real root enters the interval $X \geq \alpha^2$ by traveling on the real axis for $e > 1$. However, this results does not account for pairs of real roots appearing from the merging of two complex roots.

Proposition 2.3. *If we look at the value of the polynomial at $X = \alpha^2$ as a function of e, then $\Pi(\alpha^2; e)$ is a polynomial of degree 4 in e, and has exactly one real root for*

$e > 1$, *independently of the choice of* $\alpha > 0$, $\beta > 0$.

Proof. We can directly show that

$$\Pi(\alpha^2; e)\big|_{e=1} > 0, \qquad \frac{\partial \Pi}{\partial e}(\alpha^2; e)\bigg|_{e=1} < 0,$$

$$\frac{\partial^2 \Pi}{\partial e^2}(\alpha^2; e)\bigg|_{e=1} < 0, \qquad \frac{\partial^3 \Pi}{\partial e^3}(\alpha^2; e)\bigg|_{e=1} < 0.$$

This means that $\Pi(\alpha^2; e)$, as a function of e, is positive at $e = 1$ and is strictly decreasing for all $e \geq 1$. □

2.3 Asymptotic Formulas for Optimized One-Sided Robin Conditions

Thus, from the previous we see that it can be hard for a general advection-diffusion equation to fully solve the min-max problem to get optimized transmission conditions. In particular, we observe numerically that when b (the component of the advection tangential to the interface) is large, the optimized convergence factor does not satisfy the equioscillation property until h is very small. In the current and subsequent sections, we derive asymptotic formulas for various optimized transmission conditions that are valid *for h small enough.* In all cases, we proceed as follows.

1. First, we solve the min-max problem numerically for several values of the coefficients and for small values of h. A simple Matlab function that was used for this is included in Appendix A.1.

2. We look at the behavior of the optimized parameters and local maxima of the convergence factor as functions of h. By computing the slope in a logarithmic plot, we guess the leading order asymptotic behavior of these quantities for small h.

3. Then, *assuming* that the equioscillation property holds for h small enough (based on what is observed numerically), we match leading order terms in the

asymptotic expansion of the convergence factor. This leads to equations for the unknown constants in the asymptotic formulas for the optimized parameters.

4. Finally, we show in practice how close these asymptotic formulas are, compared to the (numerically) optimized parameters, in cases when b is small and large.

Please note that this process does not *prove* anything about the solution of the min-max problem, and that it is strongly based on observations made from numerical experimentation. For simplicity, we will choose the minimum and maximum frequencies to be $k_1 = 0$, $k_2 = \frac{\pi}{h}$ in the non-overlapping case, and $k_1 = 0$, $k_2 = \infty$ for the overlapping case (instead of the choices prescribed by Remark 1.1). Note that in the overlapping case, the dependence of the convergence factor on the mesh size h only comes from choosing a minimal overlap of the form $L = C_l h$.

The Overlapping Case

Consider first the case of Robin transmission conditions with one free parameter,

$$\sigma_1(k) = \sigma_2(k) = p \in \mathbb{R},$$

with an overlap of size proportional to h, $L = C_l h$. The convergence factor can be written as

$$\rho(k,p) = \frac{[p - \xi(k)]^2 + \eta(k)^2}{[p + \xi(k)]^2 + \eta(k)^2} e^{-\frac{L}{\nu}\xi(k)},$$

and the optimized Robin parameter is obtained by solving the min-max problem

$$\min_{p \in \mathbb{R}} \left(\max_{0 \leq k \leq \infty} \rho(k,p) \right).$$

By solving this optimization *numerically* for various values of h, we find that the optimized convergence factor has an interior local maximum at k_c that grows like $O(h^{-\frac{2}{3}})$, and the optimized Robin parameter exhibits a growth of $O(h^{-\frac{1}{3}})$. In addition, for h small enough, the equioscillation property holds at the frequencies k_1 and k_c.

Thus, from these observations, we make the ansatz

$$k_c = C_c h^{-\frac{2}{3}}, \qquad p^* = C_p h^{-\frac{1}{3}},$$

and we assume that the equioscillation $\rho(k_1, p^*) = \rho(k_c, p^*)$ holds for h small enough. By matching the leading order terms, we can derive two equations for the coefficients C_c and C_p.

- k_c is a local maximum of the convergence factor: $\frac{\partial \rho}{\partial k}(k_c, p^*) = 0$

$$\Rightarrow \quad C_p = \nu C_l C_c^2.$$

- Equioscillation property at k_1 and k_c: $\rho(0, p^*) = \rho(k_c, p^*)$

$$\Rightarrow \quad C_p^2 + \nu C_l C_c^2 C_p = 2\nu\sqrt{A} C_c.$$

Solving these two equations for the two unknown coefficients, we get

$$k_c = \left(\frac{\sqrt{A}}{\nu C_l^2}\right)^{\frac{1}{3}} h^{-\frac{2}{3}}, \qquad \boxed{p^* = \left(\frac{\nu A}{C_l}\right)^{\frac{1}{3}} h^{-\frac{1}{3}}}. \tag{2.9}$$

Finally, expanding the convergence factor at $k = 0$ for small h, we find the asymptotic convergence factor

$$\max_{0 \le k < \infty} \rho(k, p^*) = 1 - 4\left(\frac{C_l\sqrt{A}}{\nu}\right)^{\frac{1}{3}} h^{\frac{1}{3}} + O(h^{\frac{2}{3}}).$$

In Figure 2.4, we compare the Robin parameter optimized numerically, with the Robin parameter given by the asymptotic formula (2.9), for a small and a large value of b. We see that, even when $b = 25$, the convergence factor obtained by using the asymptotic formula for the Robin parameter is not very far from the optimized (best) convergence factor. In these plots, and for the ones to follow, please take note of the scale of the y-axis; it changes significantly from figure to figure, affecting the comparison.

2.4 Asymptotic Formulas for Optimized Two-Sided Robin Conditions

Let us consider now two-sided Robin transmission conditions,

$$\sigma_1(k) = p_1, \quad \sigma_2(k) = p_2, \qquad \text{where } p_1, p_2 > 0.$$

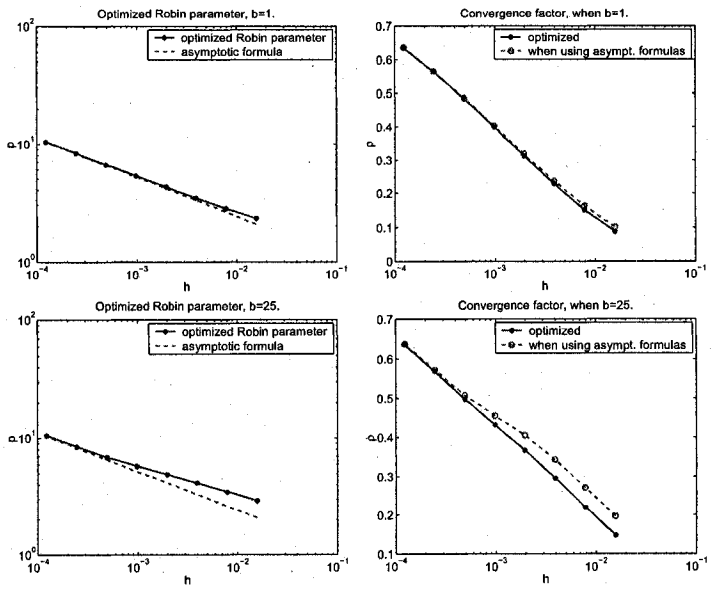

Figure 2.4: Comparison of the **optimized Robin parameter** (obtained numerically) with the asymptotic formula valid for small h, when $\nu = 0.1$, $a = c = 1$ and $\mathbf{L} = \mathbf{h}$.

The Non-Overlapping Case

In the non-overlapping case, $L = 0$, we set $k_1 = 0$ and $k_2 = \frac{\pi}{h}$. To simplify the calculations, we use the *square* of the convergence factor, which can be written as

$$\rho(k, p_1, p_2)^2 = \frac{[p_1 - \xi(k)]^2 + \eta(k)^2}{[p_1 + \xi(k)]^2 + \eta(k)^2} \cdot \frac{[p_2 - \xi(k)]^2 + \eta(k)^2}{[p_2 + \xi(k)]^2 + \eta(k)^2}.$$

The optimized parameter values (p_1^*, p_2^*) are obtained via the solution of the min-max problem

$$\min_{p_1, p_2 > 0} \left(\max_{0 \le k \le \frac{\pi}{h}} \rho(k, p_1, p_2)^2 \right). \tag{2.10}$$

By solving the min-max problem (2.10) *numerically* for various coefficient values and various mesh sizes h, we make the following observations:

- $\rho(k, p_1^*, p_2^*)$ has one interior local maximum at k_c, where k_c grows asymptotically like $h^{-\frac{1}{2}}$.
- The optimized parameters p_1^* and p_2^* seem to grow asymptotically like $h^{-\frac{1}{4}}$ and $h^{-\frac{3}{4}}$ respectively.
- The optimized convergence factor satisfies an equioscillation property at the frequencies k_1, k_c and k_2, for h small enough.

Hence, from these observations, which agree with the results of [32] for self-adjoint problems, we make the educated assumptions that, for small h,

$$k_c = C_c h^{-\frac{1}{2}}, \quad p_1^* = C_1 h^{-\frac{1}{4}}, \quad p_2^* = C_2 h^{-\frac{3}{4}},$$

$$\rho(k_1, p_1^*, p_2^*) = \rho(k_c, p_1^*, p_2^*) = \rho(k_2, p_1^*, p_2^*).$$

By looking at the leading order term of the different equations that should be satisfied, we derive relations between the unknown coefficients C_c, C_1, C_2:

- k_c is a local maximum of the convergence factor: $\frac{\partial \rho}{\partial k}(k_c, p_1^*, p_2^*) = 0$

$$\Rightarrow \quad 4\nu^2 C_c^2 = C_1 C_2.$$

- Equioscillation property at k_1 and k_2: $\rho(0, p_1^*, p_2^*) = \rho(\frac{\pi}{h}, p_1^*, p_2^*)$

$$\Rightarrow \quad 4C_1\nu^2C_c^2 + C_1^2C_2 = 2\nu\sqrt{A}C_2C_c.$$

- Equioscillation property at k_1 and k_c: $\rho(0, p_1^*, p_2^*) = \rho(k_c, p_1^*, p_2^*)$

$$\Rightarrow \quad C_1C_2 = 2\nu\pi\sqrt{A}.$$

Solving these three equations for the three coefficients C_1, C_2 and C_c, we get

$$k_c = \left(\frac{\pi\sqrt{A}}{2\nu}\right)^{\frac{1}{2}} h^{-\frac{1}{2}},$$

$$\boxed{p_1^* = \left(\frac{\nu\pi A^{\frac{3}{2}}}{2}\right)^{\frac{1}{4}} h^{-\frac{1}{4}}}, \quad \boxed{p_2^* = \left(2^5\nu^3\pi^3\sqrt{A}\right)^{\frac{1}{4}} h^{-\frac{3}{4}}}. \tag{2.11}$$

Finally, expanding the convergence factor $\rho(0, p_1^*, p_2^*)$ for small h, we obtain the asymptotic performance

$$\max_{0\le k\le\frac{\pi}{h}} \rho(k, p_1^*, p_2^*) = 1 - 2\left(\frac{2\sqrt{A}}{\pi\nu}\right)^{\frac{1}{4}} h^{\frac{1}{4}} + O(h^{\frac{1}{2}}).$$

In Figure 2.5, we compare the two-sided Robin parameters optimized numerically, with the Robin parameters given by the asymptotic formulas (2.11), for a small and a large value of b. We again see that, although the asymptotic formulas are getting worse as b increases, when $b = 25$ the convergence factor obtained by using the asymptotic formulas is still not very far from the optimized (i.e. best) convergence factor.

The Overlapping Case

For the overlapping case, we choose an overlap size proportional to the mesh size, $L = C_l h$, and set $k_1 = 0$, $k_2 = \infty$ for simplicity. The *square* of the convergence factor can be written as

$$\rho(k, p_1, p_2)^2 = \frac{[p_1 - \xi(k)]^2 + \eta(k)^2}{[p_1 + \xi(k)]^2 + \eta(k)^2} \cdot \frac{[p_2 - \xi(k)]^2 + \eta(k)^2}{[p_2 + \xi(k)]^2 + \eta(k)^2} e^{-\frac{2C_l h}{\nu}\xi(k)}.$$

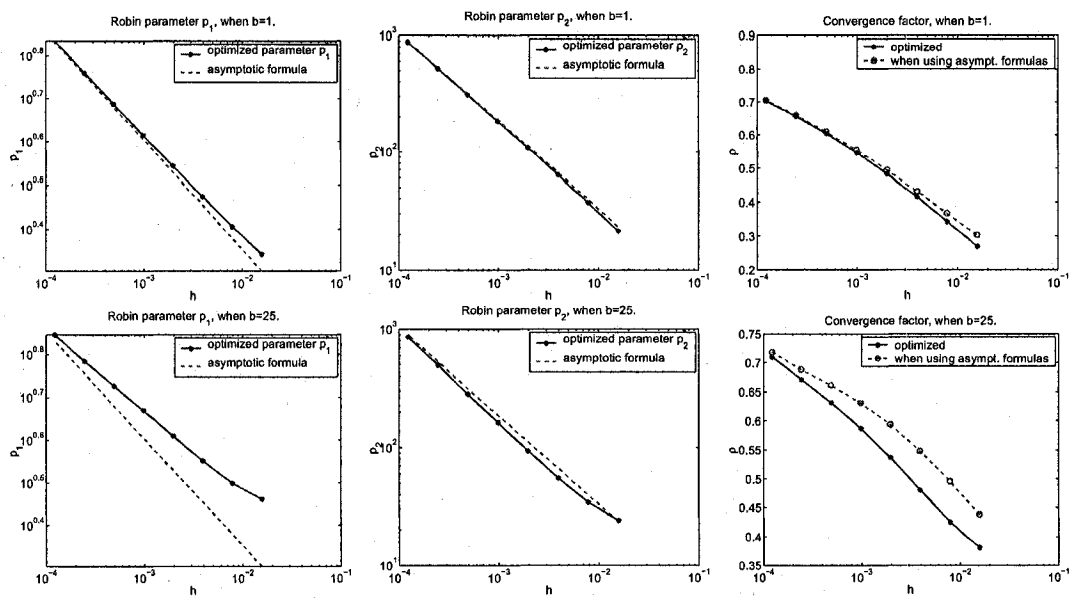

Figure 2.5: Comparison of the **optimized two-sided Robin parameters** (obtained numerically) with the asymptotic formulas valid for small h, when $\nu = 0.1$, $a = c = 1$ and $\mathbf{L} = \mathbf{0}$.

The optimized parameter values (p_1^*, p_2^*) are obtained via the solution of the min-max problem

$$\min_{p_1,p_2>0} \left(\max_{0\leq k\leq\infty} \rho(k, p_1, p_2)^2 \right). \tag{2.12}$$

By solving the min-max problem (2.12) *numerically* for various coefficient values and various mesh sizes h, we can make a few observations:

- $\rho(k, p_1^*, p_2^*)$ has two interior local maxima, at k_a and k_b, where k_a grows asymptotically like $h^{-\frac{2}{5}}$ and k_b grows asymptotically like $h^{-\frac{4}{5}}$.
- The optimized parameters p_1^* and p_2^* seem to grow like $h^{-\frac{1}{5}}$ and $h^{-\frac{3}{5}}$ respectively.
- The optimized convergence factor satisfies an equioscillation property at the frequencies $k_1 = 0$, k_a and k_b, for h small enough.

Hence, from these observations, which agree with the results of [32] for self-adjoint problems, we make the assumptions that, for small h,

$$k_a = C_a h^{-\frac{2}{5}}, \quad k_b = C_b h^{-\frac{4}{5}}, \quad p_1^* = C_1 h^{-\frac{1}{5}}, \quad p_2^* = C_2 h^{-\frac{3}{5}},$$

$$\rho(k_1, p_1^*, p_2^*) = \rho(k_a, p_1^*, p_2^*) = \rho(k_b, p_1^*, p_2^*).$$

By looking at the leading order term of the different equations that should be satisfied, we derive relations between the unknown coefficients C_a, C_b, C_1 and C_2:

- k_a is a local maximum of the optimized convergence factor: $\frac{\partial \rho}{\partial k}(k_a, p_1^*, p_2^*) = 0$

$$\Rightarrow \quad C_1 C_2 = 4\nu^2 C_a^2.$$

- k_b is a local maximum of the optimized convergence factor: $\frac{\partial \rho}{\partial k}(k_b, p_1^*, p_2^*) = 0$

$$\Rightarrow \quad C_2 = 2\nu C_l C_b^2.$$

- Equioscillation property at k_1 and k_a: $\rho(0, p_1^*, p_2^*) = \rho(k_a, p_1^*, p_2^*)$

$$\Rightarrow \quad 4C_1\nu^2 C_a^2 + C_1^2 C_2 = 2\nu\sqrt{A} C_2 C_a.$$

- Equioscillation property at k_1 and k_b: $\rho(0, p_1^*, p_2^*) = \rho(k_b, p_1^*, p_2^*)$

$$\Rightarrow \quad 2\nu C_l C_b^2 C_1 + C_1 C_2 = 2\nu\sqrt{A} C_b.$$

Solving these four equations for the four coefficients C_a, C_b, C_1 and C_2, we get

$$k_a = \left(\frac{A^{\frac{3}{2}}}{2^6\nu^3 C_l^2}\right)^{\frac{1}{5}} h^{-\frac{2}{5}}, \quad k_b = \left(\frac{\sqrt{A}}{4\nu C_l^4}\right)^{\frac{1}{5}} h^{-\frac{4}{5}},$$

$$\boxed{p_1^* = \left(\frac{\nu A^2}{8C_l}\right)^{\frac{1}{5}} h^{-\frac{1}{5}}}, \quad \boxed{p_2^* = \left(\frac{2\nu^3 A}{C_l^3}\right)^{\frac{1}{5}} h^{-\frac{3}{5}}}. \tag{2.13}$$

Finally, expanding the convergence factor $\rho(0, p_1^*, p_2^*)$ for small h we obtain the asymptotic performance

$$\max_{0\le k\le\infty} \rho(k, p_1^*, p_2^*) = 1 - 2\left(\frac{8C_l\sqrt{A}}{\nu}\right)^{\frac{1}{5}} h^{\frac{1}{5}} + O(h^{\frac{2}{5}}).$$

In Figure 2.6, we compare the two-sided Robin parameters optimized numerically, with the Robin parameters given by the asymptotic formulas (2.13) for a small and a large value of b. Once again, the asymptotic formulas are not as good when b is large, but they still give a convergence factor close to the best one.

2.5 Asymptotic Formulas for Optimized Second Order Conditions

We consider now (one-sided) second order transmission conditions of the form

$$\mathcal{S}_1 = \mathcal{S}_2 = p + q\left(b\frac{\partial}{\partial y} - \nu\frac{\partial^2}{\partial y^2}\right),$$

with $p, q > 0$. (Recall that our general Schwarz iteration is given by (2.2) and (2.3)). The corresponding Fourier symbol is

$$\sigma_1(k) = \sigma_2(k) = p + q(-ibk + \nu k^2).$$

The same process used in Section 2.4 can again be applied for these transmission conditions to derive asymptotic formulas for the optimized parameters p^* and q^*.

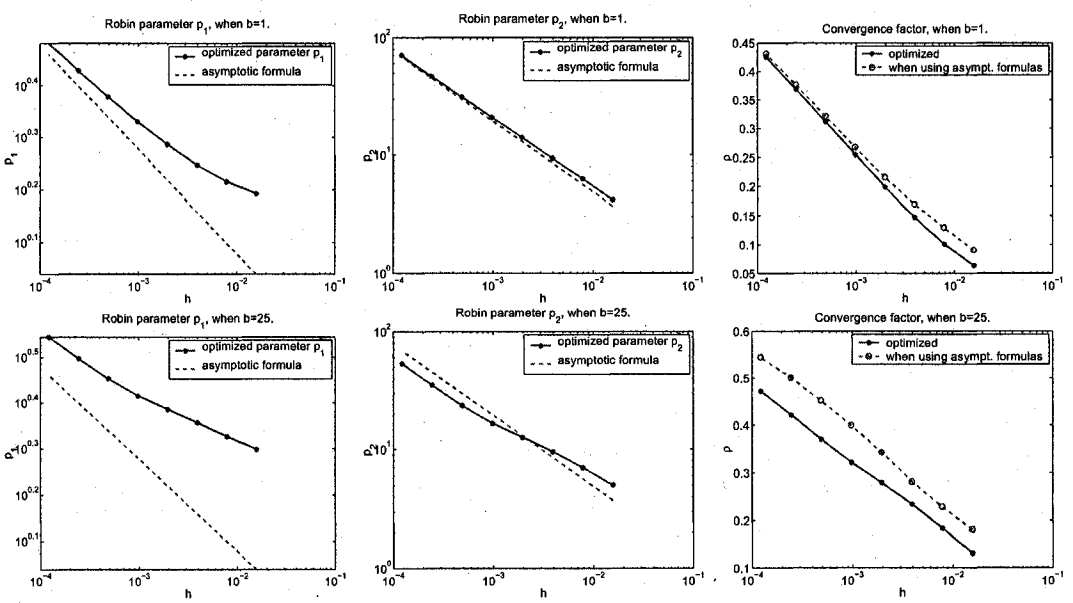

Figure 2.6: Comparison of the **optimized two-sided Robin parameters** (obtained numerically) with the asymptotic formulas valid for small h, when $\nu = 0.1$, $a = c = 1$ and $\mathbf{L} = \mathbf{h}$.

The Non-Overlapping Case

In the non-overlapping case, $L = 0$, we set $k_1 = 0$ and $k_2 = \frac{\pi}{h}$. The convergence factor can be written as

$$\rho(k,p,q) = \frac{[p + \nu q k^2 - \xi(k)]^2 + [-bqk - \eta(k)]^2}{[p + \nu q k^2 + \xi(k)]^2 + [-bqk + \eta(k)]^2}.$$

The optimized parameter values (p^*, q^*) are obtained via the solution of the min-max problem

$$\min_{p,q>0} \left(\max_{0 \le k \le \frac{\pi}{h}} \rho(k,p,q) \right). \tag{2.14}$$

By solving the min-max problem (2.14) *numerically* for various coefficient values and various mesh sizes h, we make a few observations:

- $\rho(k, p^*, q^*)$ has one interior local maximum at k_c, where k_c grows asymptotically like $h^{-\frac{1}{2}}$.
- The optimized parameter p^* seems to grow like $h^{-\frac{1}{4}}$ and q^* seems to decay like $h^{\frac{3}{4}}$.
- The optimized convergence factor satisfies an equioscillation property at the frequencies k_1, k_c and k_2, for h small enough.

Hence, from these observations, which once again agree with the results of [32] for self-adjoint problems, we make the assumptions that, for small h,

$$k_c = C_c h^{-\frac{1}{2}}, \quad p^* = C_p h^{-\frac{1}{4}}, \quad q^* = C_q h^{\frac{3}{4}},$$

$$\rho(k_1, p^*, q^*) = \rho(k_c, p^*, q^*) = \rho(k_2, p^*, q^*).$$

By looking at the leading order term of the different equations that should be satisfied, we derive relations between the unknown coefficients C_c, C_p, C_q:

- k_c is a local maximum of the convergence factor: $\frac{\partial \rho}{\partial k}(k_c, p^*, q^*) = 0$

$$\Rightarrow \quad C_p = \nu C_c^2 C_q.$$

- Equioscillation property at k_1 and k_2: $\rho(0,p^*,q^*) = \rho(\frac{\pi}{h},p^*,q^*)$

$$\Rightarrow \quad C_p(C_p + \nu C_q C_c^2) = 2\nu\sqrt{A}C_c.$$

- Equioscillation property at k_1 and k_c: $\rho(0,p^*,q^*) = \rho(k_c,p^*,q^*)$

$$\Rightarrow \quad 2C_p = \pi\sqrt{A}C_q.$$

Solving these three equations for the three coefficients C_p, C_q and C_c, we get

$$k_c = \left(\frac{\pi\sqrt{A}}{2\nu}\right)^{\frac{1}{2}} h^{-\frac{1}{2}},$$

$$\boxed{p^* = \left(\frac{\nu\pi A^{\frac{3}{2}}}{2}\right)^{\frac{1}{4}} h^{-\frac{1}{4}}}, \quad \boxed{q^* = \left(\frac{8\nu}{\pi^3\sqrt{A}}\right)^{\frac{1}{4}} h^{\frac{3}{4}}}. \tag{2.15}$$

Finally, expanding the convergence factor $\rho(0,p^*,q^*)$ for small h we obtain the asymptotic performance

$$\max_{0\le k\le\frac{\pi}{h}} \rho(k,p^*,q^*) = 1 - 4\left(\frac{2\sqrt{A}}{\pi\nu}\right)^{\frac{1}{4}} h^{\frac{1}{4}} + O(h^{\frac{1}{2}}).$$

In Figure 2.7, we compare the second order parameters optimized numerically, with the parameters given by the asymptotic formulas (2.15) for a small and a large value of b. Note in particular the extremely good approximation that the asymptotic formulas give when b is not large.

The Overlapping Case

For the overlapping case, we choose an overlap size proportional to the mesh size, $L = C_l h$, and set $k_1 = 0$, $k_2 = \infty$ for simplicity. The convergence factor for our choice of second order conditions can be written as

$$\rho(k,p,q) = \frac{[p+\nu qk^2 - \xi(k)]^2 + [-bqk - \eta(k)]^2}{[p+\nu qk^2 + \xi(k)]^2 + [-bqk + \eta(k)]^2} \, exp\left(-\frac{C_l h}{\nu}\xi(k)\right).$$

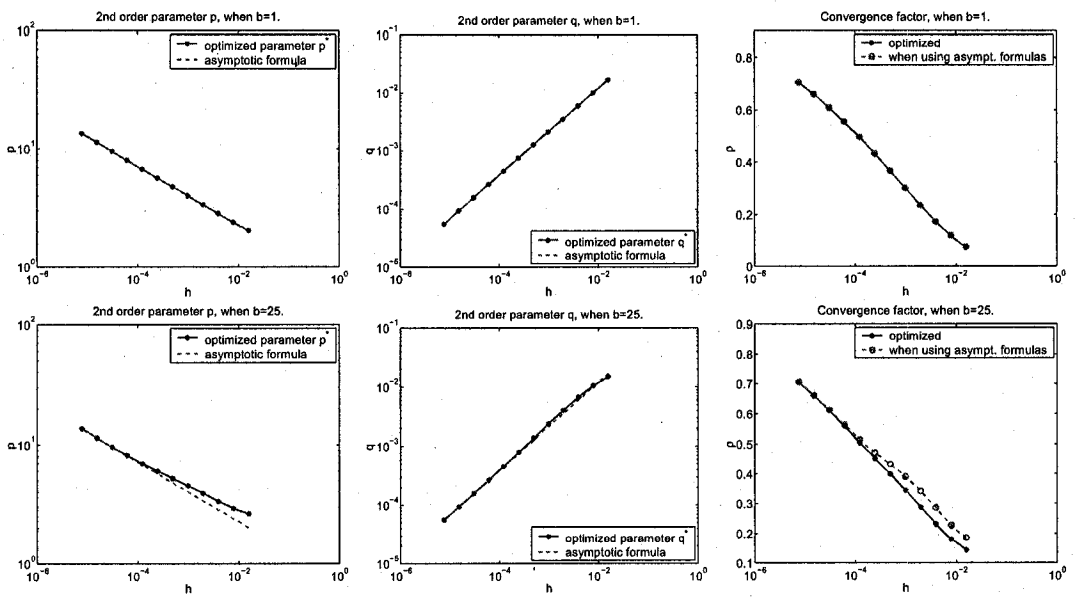

Figure 2.7: Comparison of the **optimized second order parameters** (obtained numerically) with the asymptotic formulas valid for small h, when $\nu = 0.1$, $a = c = 1$ and $\mathbf{L} = \mathbf{0}$.

The optimized parameter values (p^*, q^*) are obtained via the solution of the min-max problem

$$\min_{p,q>0} \left(\max_{0 \leq k \leq \infty} \rho(k, p, q) \right). \tag{2.16}$$

By solving the min-max problem (2.16) *numerically* for various coefficient values and various mesh sizes h, we make a few observations:

- $\rho(k, p^*, q^*)$ has two interior local maxima, at k_a and k_b, where k_a grows asymptotically like $h^{-\frac{2}{5}}$ and k_b grows asymptotically like $h^{-\frac{4}{5}}$.
- The optimized parameter p^* seem to grow like $h^{-\frac{1}{5}}$ and q^* seem to decay like $h^{\frac{3}{5}}$.
- The optimized convergence factor satisfies an equioscillation property at the frequencies $k_1 = 0$, k_a and k_b, for h small enough.

Hence, from these observations, which also agree with the results of [32] for self-adjoint problems, we make the educated assumptions that, for small h,

$$k_a = C_a h^{-\frac{2}{5}}, \quad k_b = C_b h^{-\frac{4}{5}}, \quad p^* = C_p h^{-\frac{1}{5}}, \quad q^* = C_q h^{\frac{3}{5}},$$

$$\rho(k_1, p^*, q^*) = \rho(k_a, p^*, q^*) = \rho(k_b, p^*, q^*).$$

By looking at the leading order term of the different equations that should be satisfied, we derive relations between the unknown coefficients C_a, C_b, C_p and C_q:

- k_a is a local maximum of the optimized convergence factor: $\frac{\partial \rho}{\partial k}(k_a, p^*, q^*) = 0$

$$\Rightarrow \quad C_p = \nu C_q C_a^2.$$

- k_b is a local maximum of the optimized convergence factor: $\frac{\partial \rho}{\partial k}(k_b, p^*, q^*) = 0$

$$\Rightarrow \quad 4 = C_l C_q C_b^2.$$

- Equioscillation property at k_1 and k_a: $\rho(0, p^*, q^*) = \rho(k_a, p^*, q^*)$

$$\Rightarrow \quad C_p(C_p + \nu C_q C_a^2) = 2\nu\sqrt{A} C_a.$$

- Equioscillation property at k_1 and k_b: $\rho(0, p^*, q^*) = \rho(k_b, p^*, q^*)$

$$\Rightarrow \quad C_p(C_l C_q C_b^2 + 4) = 2\sqrt{A} C_q C_b.$$

Solving these four equations for the four coefficients C_a, C_b, C_p and C_q, we get

$$k_a = \left(\frac{A^{\frac{3}{2}}}{2^4 \nu^3 C_l^2}\right)^{\frac{1}{5}} h^{-\frac{2}{5}}, \quad k_b = \left(\frac{4\sqrt{A}}{\nu C_l^4}\right)^{\frac{1}{5}} h^{-\frac{4}{5}},$$

$$\boxed{p^* = \left(\frac{\nu A^2}{4C_l}\right)^{\frac{1}{5}} h^{-\frac{1}{5}}}, \quad \boxed{q^* = \left(\frac{2^6 \nu^2 C_l^3}{A}\right)^{\frac{1}{5}} h^{\frac{3}{5}}}. \tag{2.17}$$

Finally, expanding the $\rho(0, p^*, q^*)$ for small h we obtain the asymptotic performance

$$\max_{0 \le k \le \infty} \rho(k, p^*, q^*) = 1 - 4\left(\frac{4C_l\sqrt{A}}{\nu}\right)^{\frac{1}{5}} h^{\frac{1}{5}} + O(h^{\frac{2}{5}}).$$

In Figure 2.8, once more we compare the second order parameters optimized numerically, with the parameters given by the asymptotic formulas (2.17) for a small and a large value of b.

2.6 Additional Remarks

First, note that all the asymptotic formulas we have derived for h small are *independent of* b. This is due to the fact that, for k large enough,

$$z(k) = \sqrt{A + 4\nu^2 k^2 - 4i\nu bk} \approx 2\nu k,$$

and that $z(k_1) = z(0) = \sqrt{A}$, which is independent of b. Thus, the influence of the tangential component of the advection only appears in lower order terms in the asymptotic expansion with respect to small h. But, by making b large enough with respect to $k_2 = \frac{\pi}{h}$, we will get

$$z(k) \approx \sqrt{-4i\nu bk},$$

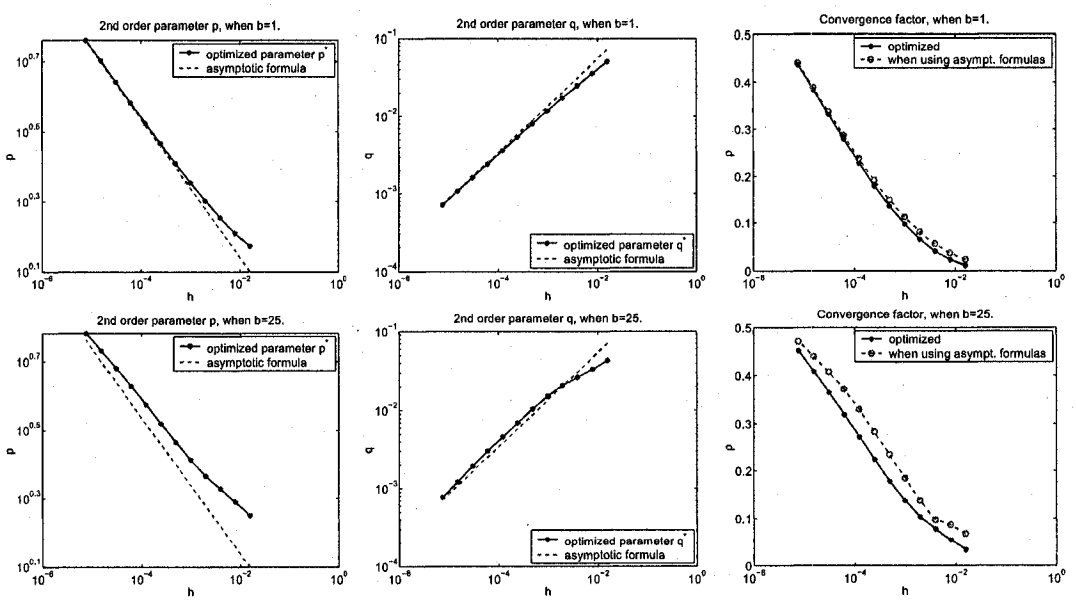

Figure 2.8: Comparison of the **optimized second order parameters** (obtained numerically) with the asymptotic formulas valid for small h, when $\nu = 0.1$, $a = c = 1$ and $\mathbf{L} = \mathbf{h}$.

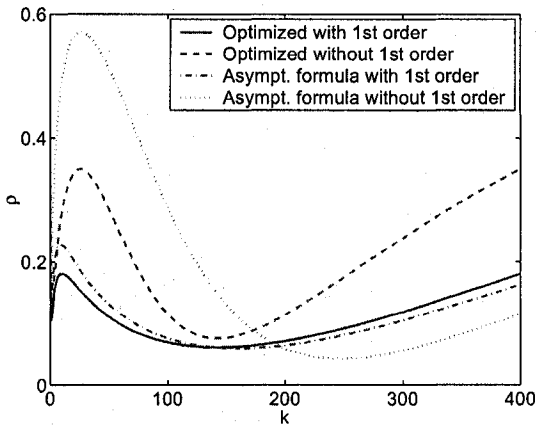

Figure 2.9: Comparison of the convergence factors obtained when including the first order term or not, without overlap, for $h = \pi/400$ and the coefficients $\nu = 0.1$, $a = c = 1$, $b = 25$.

and our asymptotic formulas will no longer be valid. Also, it can be seen numerically that for b large enough, the equioscillation property will not hold, hence our asymptotic formulas cannot be valid since they were derived by *assuming* this property.

For the second order transmission conditions of Section 2.5, we have included a first order term in the tangential direction to the interface, scaled by b,

$$\mathcal{S}_j = p + q\left(b\frac{\partial}{\partial y} - \nu\frac{\partial^2}{\partial y^2}\right).$$

Since we have observed above that the asymptotic formulas for p^* and q^* are independent of b, then the *same* asymptotic formulas are also obtained when we do not include the first order derivative in the transmission conditions. One may wonder if this first order term makes any significant difference. Figure 2.9 shows that it does have a big impact. When b is relatively large (here $b = 25$), the convergence factor optimized with the first order term is much smaller than the convergence factor optimized without it.

2.7 Numerical Experiments with Two Subdomains

To illustrate and compare the convergence of various optimized Schwarz methods, we consider here an advection-diffusion problem with constant coefficients, on a square, with homogeneous Dirichlet boundary conditions

$$\begin{cases} -\nu\Delta u + (a,b)\cdot\nabla u + cu &= 1 \quad \text{in } \Omega = (0,\pi)\times(0,\pi), \\ u &= 0 \quad \text{on } \partial\Omega. \end{cases} \tag{2.18}$$

In all the experiments that follow, we choose the coefficients $\nu = 0.1$, $a = 1$, $c = 1$, and show the results with two different values of b, namely $b = 1$ and $b = 25$. The square domain is decomposed into two non-overlapping subdomains

$$\Omega_1 = \left(0, \frac{\pi}{2}\right) \times (0,\pi), \quad \Omega_2 = \left(\frac{\pi}{2}, \pi\right) \times (0,\pi).$$

We use a finite volume discretization on a uniform grid with grid size h. At each iteration of the Schwarz method, to compute the error, we first glue the two subdomain approximations $\mathbf{u}_j^n$ together by taking an average in the overlapping region (or on the interface when $L = 0$), to get a global approximation $\mathbf{U}^n$. Then, we compute the ℓ^∞-error between this approximation $\mathbf{U}^n$ and the discrete solution of the global problem (2.18) discretized on the same grid. We use the "backslash" operator of Matlab for solving the local and global linear systems.

First, let us fix the grid size to be $h = \frac{\pi}{400}$. The convergence of different Schwarz methods is shown in Figure 2.10 for the case $b = 1$, and in Figure 2.11 for the case $b = 25$. In these figures, the error is plotted at every second iteration, to obtain a straight line. Please note that the x-axis is different for the non-overlapping and overlapping decompositions. The optimized parameters for the transmission conditions are computed by solving the min-max problem numerically (using the code of Appendix A.1) when no theoretical result is available. From these results, we make the following remarks:

(i) The optimized second order conditions yield the fastest convergence among the conditions we tested, with great improvement when compared to Dirichlet

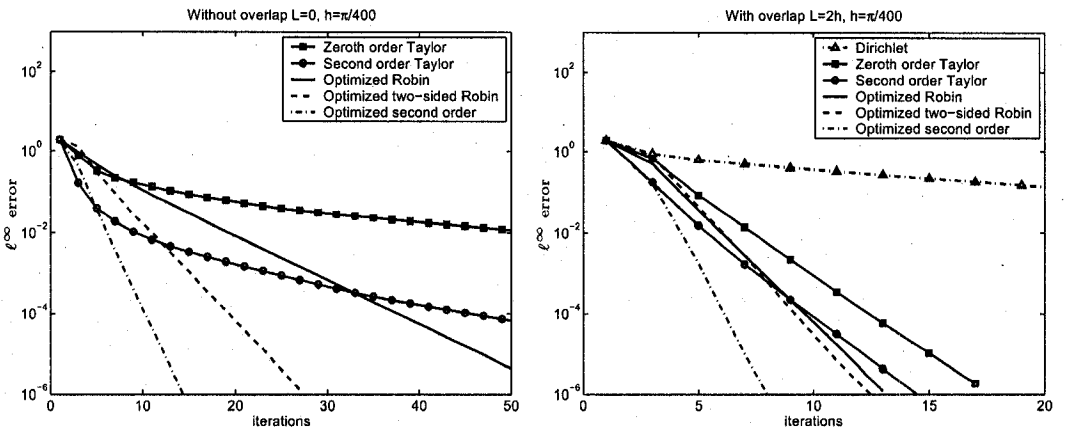

Figure 2.10: Convergence of the Schwarz iteration with various choices of transmission conditions, in the non-overlapping case on the left, and in the overlapping case ($L = 2h$) on the right, with the coefficients $\nu = 0.1$, $a = c = 1$, $\mathbf{b} = \mathbf{1}$, and $h = \frac{\pi}{400}$.

conditions and Taylor approximations. Moreover, note that the convergence is the same for $b = 1$ and $b = 25$, it doesn't seem to be affected when varying b. This is likely due to the inclusion of the first order term tangential to the interface, scaled by b, in the transmission conditions.

(ii) In the non-overlapping case, the optimized one-sided Robin conditions give faster convergence when b is large. Note that the convergence for optimized one-sided Robin conditions is the same as for the optimized two-sided Robin conditions when $b = 25$ (the two lines are indistinguishable for $L = 2h$).

(iii) The performance of the second order Taylor approximations quickly deteriorates for large b; the Schwarz iteration with these conditions does not even converge in the case $b = 25$ and $L = 0$.

Now, let us look at the asymptotic convergence of the different methods when h is small, and also at the comparison between the optimized transmission conditions and the asymptotic formulas we derived in this chapter. For this purpose, we compare in the same table several choices of transmission conditions: Taylor approximations,

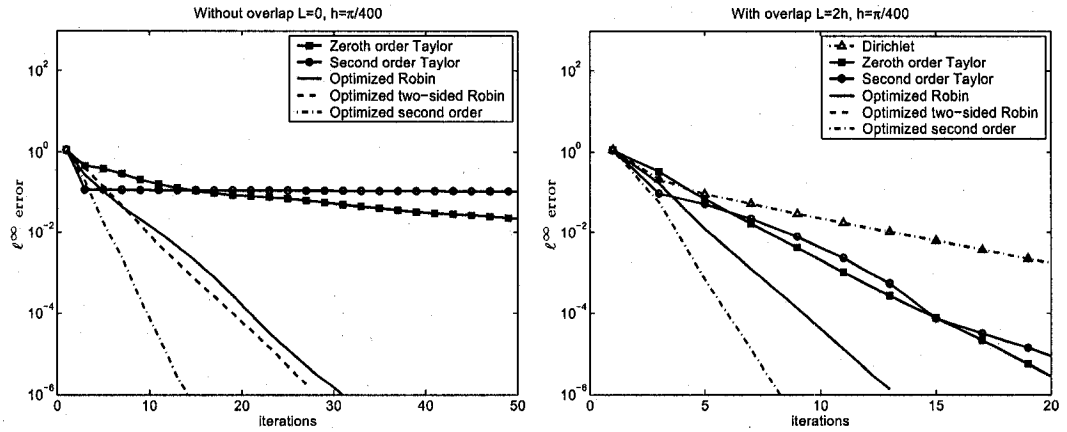

Figure 2.11: Convergence of the Schwarz iteration with various choices of transmission conditions, in the non-overlapping case on the left, and in the overlapping case ($L = 2h$) on the right, with the coefficients $\nu = 0.1$, $a = c = 1$, $\mathbf{b} = \mathbf{25}$, and $h = \frac{\pi}{400}$.

Robin and second order conditions optimized numerically, and also Robin and second order conditions where the parameters are calculated by using the asymptotic formulas. We will use the notation:

- D: Dirichlet transmission conditions
- T0: Taylor approximation of zeroth order
- T2: Taylor approximation of second order
- R1: optimized Robin conditions (second version with the appropriate scaling)
- R2: optimized two-sided conditions
- R2a: asymptotic formulas for the optimized two-sided parameters
- S: optimized second order conditions
- Sa: asymptotic formulas for the optimized second order parameters

Tables 2.3 and 2.4 give the number of iterations that were needed to reach a tolerance of 10^{-6} (using our ℓ^∞ error definition) for the different choices of transmission conditions, for $b = 1$ and $b = 25$ respectively. From these results, several relevant observations can be made:

(i) In the overlapping case $L = 2h$, the number of iterations for the Schwarz iteration with Dirichlet transmission conditions doubles every time the mesh size is cut by a factor of 2. This confirms the expected asymptotic convergence factor of the form $1 - O(h)$.

(ii) Optimized two-sided Robin and optimized second order conditions exhibit the slowest growth in the number of iterations as h decreases.

(iii) The asymptotic formulas for the optimized parameters give a convergence which is very close to the convergence obtained by using the fully optimized transmission conditions. Even when $b = 25$ and we use the asymptotic formulas for the second order parameters (Sa), the Schwarz iteration only takes at most 2 additional iterations to reach the same accuracy, compared to the fully optimized second order conditions (S).

(iv) When $b = 25$, for all the Robin transmission conditions (T0, R1, R2, R2a), the iteration number decreases with h at the beginning, before starting to grow as expected for small h. For the optimized Robin conditions, this indicates that the equioscillation property does not hold yet for those intermediate values of h, and that the asymptotic convergence factor of the form $1 - O(h^\alpha)$ is not yet valid; it becomes valid only for smaller values of h.

The convergence can be accelerated by using the optimized Schwarz method as a preconditioner for a Krylov subspace method; precisions on how this can be done are given in Section 3.6.1.

	D	T0	T2	R1	R2	R2a	S	Sa
h	Non-overlapping case, $L = 0$							
$\frac{\pi}{50}$	-	46	16	18	17	14	9	9
$\frac{\pi}{100}$	-	80	30	27	19	17	10	10
$\frac{\pi}{200}$	-	158	54	40	22	21	12	12
$\frac{\pi}{400}$	-	314	98	55	27	27	14	14
$\frac{\pi}{800}$	-	633	182	78	33	33	17	17
	Overlapping case, $L = 2h$							
$\frac{\pi}{50}$	18	7	5	6	6	8	5	7
$\frac{\pi}{100}$	36	9	7	8	8	9	5	7
$\frac{\pi}{200}$	72	12	10	10	9	10	6	7
$\frac{\pi}{400}$	142	17	14	13	12	13	7	8
$\frac{\pi}{800}$	284	24	19	16	13	15	9	9

Table 2.3: Number of iterations to reach a tolerance of 10^{-6}, for small values of h, for the coefficients $\nu = 0.1$, $a = c = 1$ and $\mathbf{b} = \mathbf{1}$.

	D	T0	T2	R1	R2	R2a	S	Sa
h	Non-overlapping case, $L = 0$							
$\frac{\pi}{50}$	-	438	> 1000	84	54	104	19	10
$\frac{\pi}{100}$	-	288	> 1000	48	42	56	13	12
$\frac{\pi}{200}$	-	246	> 1000	31	33	35	12	14
$\frac{\pi}{400}$	-	324	> 1000	30	27	36	14	18
$\frac{\pi}{800}$	-	637	> 1000	42	32	46	17	22
	Overlapping case, $L = 2h$							
$\frac{\pi}{50}$	7	6	6	5	5	6	4	5
$\frac{\pi}{100}$	13	9	9	7	7	9	5	7
$\frac{\pi}{200}$	22	13	14	10	10	12	6	8
$\frac{\pi}{400}$	45	21	23	13	13	16	8	10
$\frac{\pi}{800}$	84	29	39	16	16	19	9	11

Table 2.4: Number of iterations to reach a tolerance of 10^{-6}, for small values of h, for the coefficients $\nu = 0.1$, $a = c = 1$ and $\mathbf{b} = \mathbf{25}$.

Chapter 3

Optimized Schwarz Methods for a Diffusion Problem with Discontinuous Coefficient

3.1 Introduction

The simulation of flow in heterogeneous porous media is an important problem that arises in many engineering applications: oil recovery, earthquake prediction, underground disposal of nuclear waste, etc. In these applications, the computational domain often consists of several regions with different materials, having (very) different physical properties. In the mathematical model, this translates into a partial differential equation with discontinuous coefficients. For instance, the steady-state one phase flow in porous media can be described, in its simplest form, by the equation

$$-\nabla \cdot (\nu \nabla p) = f,$$

where p is the pressure and ν is the permeability. This equation is derived from mass conservation and Darcy's law [15], which expresses the conservation of momentum. In practice, the permeability coefficient k can differ by several orders of magnitude in different layers of the media.

For large numerical computations with those problems, domain decomposition is a natural idea; a non-overlapping decomposition is directly suggested by the different materials. Classical Schwarz methods using Dirichlet transmission conditions require overlap to converge (see for example the books [64], [65]) and thus are not convenient in this context.

Substructuring

For non-overlapping domain decompositions, a popular strategy is *substructuring* (see [65] for an extensive review): the linear system is condensed on the interfaces by eliminating the unknowns in the interior of the subdomains. Consider the case of two subdomains Ω_1 and Ω_2 separated by an interface Γ. The Steklov-Poincaré operator, which takes Dirichlet data and returns Neumann data, is defined as

$$\mathcal{S}u := \frac{\partial v_1}{\partial n} - \frac{\partial v_2}{\partial n} \quad \text{on } \Gamma,$$

where v_1 and v_2 are the solutions of the Dirichlet subproblems

$$\begin{cases} \mathcal{L}(v_j) &= 0 \quad \text{in } \Omega_j, \\ v_j &= 0 \quad \text{on } \partial\Omega_j \cap \partial\Omega, \\ v_j &= u \quad \text{on } \Gamma. \end{cases}$$

A dual operator, which starts from Neumann data and returns Dirichlet data, can be defined via

$$\mathcal{F}\lambda := u_1 - u_2 \quad \text{on } \Gamma,$$

where u_1 and u_2 are now the solutions of the Neumann subproblems

$$\begin{cases} \mathcal{L}(u_j) &= 0 \quad \text{in } \Omega_j, \\ u_j &= 0 \quad \text{on } \partial\Omega_j \cap \partial\Omega, \\ \dfrac{\partial u_j}{\partial n} &= \lambda \quad \text{on } \Gamma. \end{cases}$$

The primal Schur method is a preconditioned Krylov subspace method applied to solve the linear system $Su = f$, where S is the discrete version of the operator $\mathcal{S}$. It is called *primal* since it enforces the continuity of the solution (the primal variables)

on the interface at each iteration. In the dual Schur method (better known as Finite Element Tearing and Interconnecting or FETI [29]), the continuity of dual variables is enforced instead, and a preconditioned Krylov subspace method is applied to the linear system $F\lambda = g$, where F is the discrete version of $\mathcal{F}$.

Recently, the FETI-DP [26] method was introduced, where continuity of the primal variables is imposed at the cross-points between subdomains, and continuity of the dual variables elsewhere on the interface. This has the nice feature of including a natural coarse space. The analysis of such methods produces condition number estimates of the form

$$\langle P\lambda, \lambda\rangle \leq \langle F\lambda, \lambda\rangle \leq C(1 + \log(H/h))^2 \langle P\lambda, \lambda\rangle, \quad \forall \lambda \in V,$$

where P is an appropriate preconditioner, h is the mesh size, H is the diameter of the subdomains, and C is constant *independent* of h, H and of the jump in the coefficients of the problem. Hence, such techniques are shown to be robust for strongly heterogeneous problems. On the other hand, these methods only involve the solution of subproblems with either Dirichlet and Neumann boundary conditions. Optimized transmission conditions (e.g. Robin) recently started to be incorporated in this setting, for example in the FETI-H method [27]. In [37], a Robin-Robin preconditioner is proposed for advection-diffusion problems having discontinuous viscosity coefficients, and an estimate for the convergence rate is derived for a model problem.

Optimized transmission conditions

For problems in heterogeneous media, a few domain decomposition methods with optimized Robin transmission conditions have recently appeared in the literature. First, in [24] and [10], optimized Robin conditions are mentioned but the optimization problem is not fully solved; instead, an approximate choice is made for the Robin parameters. In [38], [39] and [30], optimized Schwarz methods are designed at the algebraic level, where the problem is discretized in the tangential direction(s) to the interface. For advection-diffusion problems, the proposed Robin transmission

conditions are of the form

$$\mathcal{T}_1 = C\frac{\partial}{\partial x} - \frac{B}{2} + \alpha D, \quad \mathcal{T}_2 = -C\frac{\partial}{\partial x} + \frac{B}{2} + \alpha D,$$

where C and B are matrices corresponding to the discretized diffusion and advection coefficients of the problem, and D is an approximation of the diagonal of the discrete Dirichlet-to-Neumann maps Λ_j. An optimized Robin parameter α^* is then obtained by uniformly minimizing the convergence factor over a discrete spectrum, i.e. solving

$$\min_{\alpha>0}\left(\max_{\lambda\in\sigma(\Lambda_1 D^{-1})\cup\sigma(\Lambda_2 D^{-1})}\left|\frac{\lambda-\alpha}{\lambda+\alpha}\right|^2\right).$$

Note that the Robin parameter is the same along the entire interface (whether the coefficients are constant or not). Some work has also been developed on optimized Schwarz waveform relaxation for parabolic problems with discontinuous coefficients, for example in [33] and [8].

More recently, Maday and Magoulès have studied optimized Schwarz methods for a diffusion problem with discontinuous coefficient. In [50], they considered two choices of optimized Robin conditions; for the first one, they did not solve the min-max problem, but still noticed numerically that the solution is not always unique. For their second choice, which they call "one and half parameter based Robin conditions", the min-max problem was solved under the assumption that the Robin parameter p lies in the same range as the frequency k. In [51], two-sided Robin conditions are also considered. The associated optimization problem is not fully solved (the equioscillation property is not proved), but an asymptotic analysis is shown, and the authors observe that the optimized convergence rate improves as we increase the jump in the coefficients.

In this chapter, we consider the same model problem as Maday and Magoulès, i.e. a simple diffusion problem with discontinuous coefficient, and derive optimized Schwarz methods by thoroughly solving the associated min-max problems. For each choice of transmission conditions, we obtain precise formulas for the parameters that give the fastest convergence for the Schwarz method, and then analyze the asymptotic performance. We proceed as follows. First, in Section 3.2, we present an adapted

version of a general convergence proof for the Schwarz method with Robin transmission conditions. In Section 3.3, we introduce the model problem and the Schwarz iteration we consider, perform a Fourier convergence analysis, and discuss optimal coupling conditions. In Section 3.4, we review the basic ideas behind optimized Schwarz methods, and we then proceed to analyze several choices of optimized transmission conditions for our model problem. The asymptotic performance is studied for the cases of small mesh sizes and strong heterogeneity. For comparison purposes, Section 3.5 contains a short discussion and analysis of the Dirichlet-Neumann method. In Section 3.6, numerical experiments are presented to verify our convergence analysis. Finally, in Section 3.7, we discuss generalizations to diffusion-reaction problems, anisotropic diffusions, and problems in 3D.

3.2 A Convergence Proof For A Non-Overlapping Schwarz Iteration

In this section, we present a simple proof of convergence for a Schwarz iteration with a specific choice of Robin transmission conditions, using energy estimates. Such a proof first appeared in [47], where only weak convergence of the iterates in H^1 was shown. The technique was refined and strong convergence was established, see [17, 5, 14] where the Helmholtz equation was considered. Robin conditions can also be adapted for advection-diffusion problems, and the convergence of the method was shown for two subdomains in [2]. We include here a version of the proof based on [4] and adapted for a problem with discontinuous coefficients. A more general proof that includes the use of second order transmission conditions can be found in [54].

We consider the problem

$$\begin{cases} -\nabla \cdot (\nu(\mathbf{x})\nabla u) + \eta(\mathbf{x})u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{cases} \tag{3.1}$$

where Ω is a bounded domain in $\mathbb{R}^2$ or $\mathbb{R}^3$, with Lipschitz boundary. The coefficients

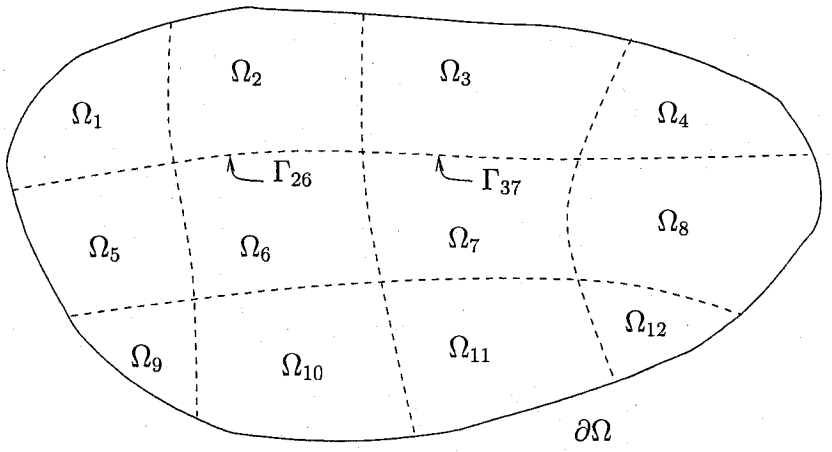

Figure 3.1: A general decomposition into M non-overlapping subdomains.

are assumed to be piecewise constant

$$\nu(\mathbf{x}) = \nu_j > 0 \quad \text{and} \quad \eta(\mathbf{x}) = \eta_j > 0 \quad \text{for } \mathbf{x} \in \Omega_j, \quad j = 1, 2, ..., M.$$

The regions $\{\Omega_j\}_{j=1}^M$ form a natural decomposition into non-overlapping subdomains, and we denote the interior interfaces by $\Gamma_{ij} := \partial\Omega_i \cap \partial\Omega_j$.

The continuous problem (3.1) can be rewritten in a variational formulation: we wish to find $u \in H_0^1(\Omega)$ such that

$$\int_\Omega (\nu(\mathbf{x})\nabla u \cdot \nabla v + \eta(\mathbf{x})uv)\, d\mathbf{x} = \int_\Omega fv\, d\mathbf{x} \quad \forall\, v \in H_0^1(\Omega). \tag{3.2}$$

Assuming that $f \in L^2(\Omega)$, there exists a unique solution to this problem, by the Lax-Milgram lemma.

For solving problem (3.1) with a decomposition of the domain, we consider a Schwarz iteration with Robin transmission conditions

$$\begin{cases} -\nu_i \Delta u_i^{n+1} + \eta_i u_i^{n+1} &= f & \text{in } \Omega_i, \\ u_i &= 0 & \text{on } \partial\Omega_i \cap \partial\Omega, \\ \nu_i \frac{\partial u_i^{n+1}}{\partial n_i} + \gamma_{ij} u_i^{n+1} &= \nu_j \frac{\partial u_j^n}{\partial n_i} + \gamma_{ij} u_j^n & \text{on } \Gamma_{ij}, \text{ for } j \in N(i), \end{cases}$$

where $\frac{\partial}{\partial n_i}$ is the normal derivation in the outward direction with respect to Ω_i, $N(i)$ contains the indices j such that $\text{meas}(\Gamma_{ij}) > 0$, and γ_{ij} are real constants with $\gamma_{ij} \geq$

$\gamma^0 > 0$. We also assume that $\gamma_{ij} = \gamma_{ji}$; this assumption is critical for the proof of convergence and we will comment on it later. In this case, if we denote the Robin data to be transmitted by

$$\lambda_{ij}^n := \nu_j \frac{\partial u_j^n}{\partial n_i} + \gamma_{ij} u_j^n,$$

then by applying the transmission conditions twice, we find that we can update this Robin data by using the simple relation

$$\lambda_{ij}^{n+1} = 2\gamma_{ij}(u_j^{n+1} - u_i^n) + \lambda_{ij}^{n-1},$$

and thus one does not need to extract a normal derivative when doing computation. In addition, if we start with $\lambda_{ij}^0 \in L^2(\Gamma_{ij})$, then $\lambda_{ij}^n \in L^2(\Gamma_{ij})$.

We can therefore write the variational form of the subproblems as:
find $u_i^{n+1} \in H^1_{\partial\Omega}(\Omega_i) := \{v \in H^1(\Omega_i) \,|\, v = 0 \text{ on } \partial\Omega_i \cap \partial\Omega\}$ such that

$$\int_{\Omega_i} (\nu_i \nabla u_i^{n+1} \cdot \nabla v + \eta_i u_i^{n+1} v)\, d\mathbf{x} + \sum_{j \in N(i)} \int_{\Gamma_{ij}} \gamma_{ij} u_i^{n+1} v \, ds \tag{3.3}$$

$$= \int_{\Omega_i} f v \, d\mathbf{x} + \sum_{j \in N(i)} \int_{\Gamma_{ij}} \lambda_{ij}^n v \, ds \quad \forall v \in H^1_{\partial\Omega}(\Omega_i).$$

These subproblems are well-posed, given that the constants γ_{ij} are positive. Let us denote the error in each subdomain, at each iteration, by $e_i^n := u_i^n - u$ where u_i^n solves (3.3) and u solves (3.2). It is clear by linearity that the error e_i^n solves the subproblem (3.3) with $f \equiv 0$.

Theorem 3.1 (Convergence of the Schwarz iteration). *Suppose the Robin conditions satisfy $\gamma_{ij} = \gamma_{ji} \geq \gamma^0 > 0$. Then,*

$$e_i^n \to 0 \quad \text{in } H^1(\Omega_i) \quad \text{as } n \to \infty.$$

Proof. We consider a *pseudoenergy* defined by

$$E^{n+1} := \sum_{i=1}^{M} \sum_{j \in N(i)} \int_{\Gamma_{ij}} \left| \nu_i \frac{\partial e_i^{n+1}}{\partial n_i} \right|^2 + \left| \gamma_{ij} e_i^{n+1} \right|^2 ds.$$

We will show that this quantity is decreasing. We first rewrite the integrand by completing the square and using the transmission conditions of the Schwarz iteration

$$\begin{aligned}\left|\nu_i\frac{\partial e_i^{n+1}}{\partial n_i}\right|^2+\left|\gamma_{ij}e_i^{n+1}\right|^2&=\left(\nu_i\frac{\partial e_i^{n+1}}{\partial n_i}+\gamma_{ij}e_i^{n+1}\right)^2-2\nu_i\frac{\partial e_i^{n+1}}{\partial n_i}\gamma_{ij}e_i^{n+1},\\&=\left(\nu_j\frac{\partial e_j^{n}}{\partial n_i}+\gamma_{ij}e_j^{n}\right)^2-2\nu_i\frac{\partial e_i^{n+1}}{\partial n_i}\gamma_{ij}e_i^{n+1},\\&=\left|\nu_j\frac{\partial e_j^{n}}{\partial n_i}\right|^2+\left|\gamma_{ij}e_j^{n}\right|^2-2\nu_i\frac{\partial e_i^{n+1}}{\partial n_i}\gamma_{ij}e_i^{n+1}+2\nu_j\frac{\partial e_j^{n}}{\partial n_i}\gamma_{ij}e_j^{n},\\&=\left|\nu_j\frac{\partial e_j^{n}}{\partial n_j}\right|^2+\left|\gamma_{ji}e_j^{n}\right|^2-2\nu_i\frac{\partial e_i^{n+1}}{\partial n_i}\gamma_{ij}e_i^{n+1}-2\nu_j\frac{\partial e_j^{n}}{\partial n_j}\gamma_{ji}e_j^{n}.\end{aligned}$$

By integrating over Γ_{ij} and summing over i and j, we find that

$$\begin{aligned}E^{n+1}=E^n&-2\sum_{i=1}^{m}\sum_{j\in N(i)}\int_{\Gamma_{ij}}\nu_i\frac{\partial e_i^{n+1}}{\partial n_i}\gamma_{ij}e_i^{n+1}\,ds\\&-2\sum_{j=1}^{m}\sum_{i\in N(j)}\int_{\Gamma_{ji}}\nu_j\frac{\partial e_j^{n}}{\partial n_j}\gamma_{ji}e_j^{n}\,ds.\end{aligned}\tag{3.4}$$

Each error e_i^n satisfies the partial differential equation $-\nu_i\Delta u+\eta_i u=0$ weakly on Ω_i, namely

$$\int_{\Omega_i}\nu_i\nabla e_i^{n+1}\cdot\nabla v+\eta_i e_i^{n+1}v\,d\mathbf{x}=\sum_{j\in N(i)}\int_{\Gamma_{ij}}\nu_i\frac{\partial e_i^{n+1}}{\partial n_i}v\,ds,$$

for all $v\in H^1_{\partial\Omega}(\Omega_i)$. Choosing $v=\gamma_{ij}e_i^{n+1}$ in the above formula, we obtain

$$\begin{aligned}\sum_{j\in N(i)}\int_{\Gamma_{ij}}\nu_i\frac{\partial e_i^{n+1}}{\partial n_i}\gamma_{ij}e_i^{n+1}&=\int_{\Omega_i}\gamma_{ij}\left(\nu_i|\nabla e_i^{n+1}|^2+\eta_i|e_i^{n+1}|^2\right)\,d\mathbf{x},\\&\geq\gamma^0\left(\nu_i|e_i^{n+1}|^2_{H^1(\Omega_i)}+\eta_i\|e_i^{n+1}\|^2_{L^2(\Omega_i)}\right),\\&\geq C\|e_i^{n+1}\|^2_{H^1(\Omega_i)},\end{aligned}$$

where $C:=\gamma^0\min_i(\nu_i,\eta_i)$. Inserting this back into equation (3.4), we get

$$E^{n+1}\leq E^n-2C\sum_{i=1}^{M}\left(\|e_i^{n+1}\|^2_{H^1(\Omega_i)}+\|e_i^{n}\|^2_{H^1(\Omega_i)}\right),$$

which shows that the pseudoenergy E^n is decreasing. Moreover, summing this inequality from $n = 0$ to $n = N$, we find that

$$E^{N+1} + 2C \sum_{n=0}^{N} \sum_{i=1}^{M} \left(\|e_i^{n+1}\|^2_{H^1(\Omega_i)} + \|e_i^n\|^2_{H^1(\Omega_i)} \right) \leq E^0.$$

The initial energy E^0 is a constant and all the terms on the left hand side of this inequality are positive, hence in particular we can deduce that $\sum_{n=0}^{N} \|e_i^n\|^2_{H^1(\Omega_i)}$ is bounded as a function of N, for each i. Therefore, the infinite series converges,

$$\sum_{n=0}^{\infty} \|e_i^n\|^2_{H^1(\Omega_i)} < \infty.$$

This strong statement implies that the general term in the series must go to 0 as a function of n, i.e.

$$\|e_i^n\|_{H^1(\Omega_i)} \to 0 \quad \text{as } n \to \infty.$$

□

Note that we have assumed that $\eta_i > 0$ in the proof, for simplicity. When $\eta(\mathbf{x}) \equiv 0$, the above argument only leads us to conclude that $|e_i^n|_{H^1(\Omega_i)} \to 0$ as $n \to \infty$. In that case, to get the convergence in the H^1 norm, we can first look at subdomains touching the boundary of Ω ($\text{meas}(\partial\Omega_i \cap \partial\Omega) > 0$): for those subdomains, the Dirichlet boundary condition gives us a Poincaré-Friedrich inequality, and the H^1 seminorm is equivalent to the H^1 norm. Then, we can proceed towards interior subdomains that are not touching $\partial\Omega$. For details, see for example [17].

Several remarks can be made about this proof.

1. The result also shows the convergence of the Schwarz iteration when the subproblems are discretized using a conforming finite element method.

2. The proof does not give us any information about the *rate of convergence* of the method.

3. It is not known whether this technique of proof can be extended to the case of overlapping domain decompositions (useful for continuous coefficients) or to the

case when $\gamma_{ij} \neq \gamma_{ji}$. Finding a general convergence proof for those situations remains an open problem. We will see in Section 3.4.3 (in particular Corollary 3.1), when $\gamma_{ij} \neq \gamma_{ji}$, that extra conditions need to be imposed on the function γ_{ij}, in addition to the positivity, to guarantee convergence when the coefficients are discontinuous.

In practice, it is useful to consider more general transmission conditions, giving us more freedom in our choice (and hence has the potential for faster convergence), even if a convergence proof in the general setting has not yet been established.

3.3 A Diffusion Model Problem

For a general bounded domain Ω, general coefficients, general domain decompositions, it is very hard to estimate the convergence rate of a Schwarz method as a function of the parameters in the transmission conditions (or even to establish that it converges). The strategy we adopt in the remainder of this chapter is to consider a simple model problem, for which a complete analysis of the Schwarz iteration can be carried out, and observe to which extent the results are effective in more general situations.

We start with the case of a well-posed steady-state diffusion problem

$$\begin{cases} -\nabla \cdot (\nu(\mathbf{x}) \nabla u) &= f \quad \text{in } \Omega \subseteq \mathbb{R}^2, \\ \mathcal{B}(u) &= g \quad \text{on } \partial\Omega, \end{cases} \tag{3.5}$$

where the scalar diffusion coefficient $\nu(\mathbf{x})$ is the piecewise constant function

$$\nu(\mathbf{x}) = \begin{cases} \nu_1 & \text{for } \mathbf{x} \in \Omega_1, \\ \nu_2 & \text{for } \mathbf{x} \in \Omega_2. \end{cases}$$

Here, Ω_1 and Ω_2 form a natural decomposition of the domain Ω into non-overlapping subdomains, and we denote by Γ the interface between the subdomains. Using physical coupling conditions between the subdomains, the problem can also be written equivalently in a multi-domain formulation

$$\left\{ \begin{array}{rcll} -\nu_1 \Delta u_1 &=& f & \text{in } \Omega_1, \\ u_1 &=& u_2 & \text{on } \Gamma, \end{array} \right. \quad \left. \begin{array}{rcll} -\nu_2 \Delta u_2 &=& f & \text{in } \Omega_2, \\ \nu_1 \frac{\partial u_1}{\partial n} &=& \nu_2 \frac{\partial u_2}{\partial n} & \text{on } \Gamma, \end{array} \right\} \tag{3.6}$$

also sometimes called a transmission problem. The matching conditions impose the continuity of the solution and of its flux across the interface Γ. It can be shown that the variational formulation of problems (3.5) and (3.6) are equivalent.

As usual, for the derivation of optimized Schwarz methods, we consider a model problem on the infinite plane $\Omega = \mathbb{R}^2$, with the subdomains

$$\Omega_1 = (-\infty, 0) \times \mathbb{R}, \quad \Omega_2 = (0, \infty) \times \mathbb{R},$$

and the condition that the solution remains bounded at infinity (this condition will be implicitly assumed in all that follows). To be consistent with the physical coupling conditions in (3.6), we use a parallel Schwarz iteration with general transmission conditions of the form

$$\begin{cases} -\nu_1 \Delta u_1^{n+1} &= f & \text{in } \Omega_1, \\ (\nu_1 \partial_x + \mathcal{S}_1) u_1^{n+1}(0, y) &= (\nu_2 \partial_x + \mathcal{S}_1) u_2^n(0, y) & \text{for } y \in \mathbb{R}, \\ -\nu_2 \Delta u_2^{n+1} &= f & \text{in } \Omega_2, \\ (\nu_2 \partial_x - \mathcal{S}_2) u_2^{n+1}(0, y) &= (\nu_1 \partial_x - \mathcal{S}_2) u_1^n(0, y) & \text{for } y \in \mathbb{R}, \end{cases} \tag{3.7}$$

where $\mathcal{S}_j$ are linear operators acting in the y direction only. We require that the operators $\mathcal{S}_j$ be chosen so that the subproblems in (3.7) are well-posed, and also such that the limit of the iteration (if it exists) is the solution of the coupled problem (3.6). By linearity, it will be sufficient to consider only the homogeneous case, $f \equiv 0$, in the convergence analysis.

3.3.1 Fourier Analysis

Our simple model problem allows us to use a Fourier transform in the y variable,

$$\mathcal{F}_y(u(x, y)) = \hat{u}(x, k) := \int_{-\infty}^{\infty} u(x, y) e^{-iyk} \, dy,$$

to analyze the convergence of the Schwarz method (3.7). Suppose the operators $\mathcal{S}_j$ have Fourier symbols $\sigma_j(k)$,

$$\mathcal{F}_y(\mathcal{S}_j u(x, y)) = \sigma_j(k) \hat{u}(x, k), \quad \text{for } j = 1, 2.$$

In Fourier space, the partial differential equation in Ω_j becomes a second order ODE in x (for each fixed k),

$$\begin{cases} -\nu_1(\partial_{xx} - k^2)\hat{u}_1^{n+1} &= 0 & \text{for } x < 0,\ k \in \mathbb{R}, \\ (\nu_1\partial_x + \sigma_1(k))\hat{u}_1^{n+1}(0,k) &= (\nu_2\partial_x + \sigma_1(k))\hat{u}_2^{n}(0,k) & \text{for } k \in \mathbb{R}, \end{cases}$$
$$\begin{cases} -\nu_2(\partial_{xx} - k^2)\hat{u}_2^{n+1} &= 0 & \text{for } x > 0,\ k \in \mathbb{R}, \\ (\nu_2\partial_x - \sigma_2(k))\hat{u}_2^{n+1}(0,k) &= (\nu_1\partial_x - \sigma_2(k))\hat{u}_1^{n}(0,k) & \text{for } k \in \mathbb{R}, \end{cases}$$

with characteristic roots $\lambda_\pm(k) := \pm|k|$. By requiring that the solution in each subdomain be bounded at infinity, we find solutions of the form

$$\hat{u}_1^n(x,k) = A^n(k)e^{|k|x}, \quad \hat{u}_2^n(x,k) = B^n(k)e^{-|k|x}.$$

Applying the transmission conditions coupling the two subdomains, we obtain two relations for the coefficients,

$$(\nu_1|k| + \sigma_1(k))A^{n+1}(k) = (-\nu_2|k| + \sigma_1(k))B^n(k),$$
$$(-\nu_2|k| - \sigma_2(k))B^{n+1}(k) = (+\nu_1|k| - \sigma_2(k))A^n(k).$$

Combining these two equations, we get the convergence factor for a double step of the Schwarz iteration

$$\rho(k,\sigma_1,\sigma_2) := \left|\frac{\hat{u}_j^{n+1}(0,k)}{\hat{u}_j^{n-1}(0,k)}\right| = \left|\frac{\sigma_1(k) - \nu_2|k|}{\sigma_1(k) + \nu_1|k|} \cdot \frac{\sigma_2(k) - \nu_1|k|}{\sigma_2(k) + \nu_2|k|}\right|. \tag{3.8}$$

Note that at $k = 0$, the convergence factor equals 1, for any choice of σ_j, i.e. the frequency $k = 0$ can never converge. Of course, this makes sense because the solution to the global model problem is only unique up to a constant. This will not cause any difficulty: we will only consider strictly positive frequency components, $|k| \geq k_1 > 0$, because in practice the domain of computation will be bounded and the problem we wish to solve numerically will be well-posed, see Section 3.4.

3.3.2 Optimal Operators

The operators $\mathcal{S}_j$ (or equivalently the symbols σ_j) are still free to be chosen at this point. It is straightforward, by inspection of (3.8), to see that we get optimal conver-

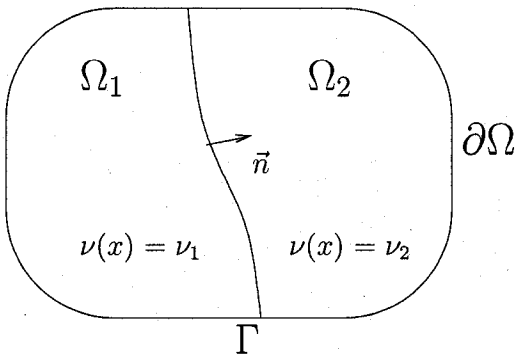

Figure 3.2: A general geometry for the definition of optimal operators.

gence in two iterations, if we choose the symbols

$$\sigma_1^{opt}(k) = \nu_2|k|, \quad \sigma_2^{opt}(k) = \nu_1|k|.$$

These symbols are not polynomials in k, and thus the corresponding operators in real space,

$$\mathcal{S}_j^{opt}(u(0,y)) = \mathcal{F}_y^{-1}(\sigma_j^{opt}(k)\hat{u}(0,k)),$$

are nonlocal in y. In fact, the operators $\mathcal{S}_j^{opt}$ are Dirichlet-to-Neumann maps, which involve solving a problem in the adjacent subdomain.

We can define these operators in general as follows. Consider a bounded domain Ω with a non-overlapping decomposition $\{\Omega_1, \Omega_2\}$, let Γ denote the interface between the subdomains and let $\frac{\partial}{\partial n}$ denote the normal derivative across Γ in the outward direction with respect to Ω_1, as depicted in Figure 3.2. We define the operator $\mathcal{S}_1^{opt} : H_{00}^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$ via $\mathcal{S}_1^{opt}(u) := -\nu_2 \frac{\partial w}{\partial n}$ where w solves the problem

$$\begin{cases} -\nu_2 \Delta w &= f & \text{in } \Omega_2, \\ w &= 0 & \text{on } \partial\Omega_2 \cap \partial\Omega, \\ w &= u & \text{on } \Gamma. \end{cases}$$

Similarly, we can define the operator $\mathcal{S}_2^{opt} : H_{00}^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$ via $\mathcal{S}_2^{opt}(u) := \nu_1 \frac{\partial v}{\partial n}$

where v solves the problem

$$\begin{cases} -\nu_1 \Delta v &= f & \text{in } \Omega_1, \\ v &= 0 & \text{on } \partial\Omega_1 \cap \partial\Omega, \\ v &= u & \text{on } \Gamma. \end{cases}$$

It is easy to verify that, using these nonlocal operators, the Schwarz iteration converges in two iterations, for any initial guesses u_j^0, $j = 1, 2$.

These Dirichlet-to-Neumann operators can be generalized to the case of a domain decomposition into M strips, and yield an optimal convergence in M iterations, see [56]. This result is also discussed in Section 4.4.

The optimal operators are easy to characterize in one dimension. Consider a bounded interval Ω, with homogeneous Dirichlet boundary conditions at the endpoints, and a general diffusion coefficient $\nu(x) > 0$. Consider a partition of Ω into two non-overlapping subintervals Ω_1 and Ω_2. In this case, the optimal transmission conditions are simple Robin conditions

$$\mathcal{S}_1^{opt}(u) = \left[\int_{\Omega_2} \nu(x)^{-1} \, dx \right]^{-1} u,$$
$$\mathcal{S}_2^{opt}(u) = \left[\int_{\Omega_1} \nu(x)^{-1} \, dx \right]^{-1} u,$$

as was pointed out in [47].

3.4 Optimized Transmission Conditions

Being nonlocal, the optimal operators are not convenient for implementation; they are computationally costly since an application of the optimal operator involves solving a problem in the complement of the subdomain. Instead, we would like to find good *local* transmission conditions that still yield very fast convergence of the Schwarz iteration (3.7). The idea is to fix a class $\mathcal{C}$ of convenient transmission conditions to consider, and uniformly optimize the convergence factor over a range of relevant frequencies for our problem. This leads to a min-max problem of the form

$$\min_{\sigma_j \in \mathcal{C}} \left(\max_{k_1 \leq k \leq k_2} \rho(k, \sigma_1, \sigma_2) \right). \tag{3.9}$$

Although the convergence factor ρ was computed for a continuous model problem on the infinite plane, we impose bounds on the frequency range by incorporating information about the actual problem we intend to solve, as discussed in Remark 1.1. For example, if we wish to compute a solution over a bounded domain with homogeneous Dirichlet boundary conditions, then the minimal frequency component of the solution can be estimated by $k_1 = \frac{\pi}{H}$ where H is the diameter of the subdomain. The lower bound k_1 could also be taken to be the smallest frequency not resolved by the coarse grid, if one is used. Moreover, if the solution is computed numerically on a grid with grid spacing h on the interface, then the maximum frequency which can be represented on this grid is typically estimated by $k_2 = \frac{\pi}{h}$.

Before analyzing several classes of transmission conditions and solving the associated min-max problem, let us state sufficient conditions on σ_1 and σ_2 for convergence of the Schwarz iteration.

Theorem 3.2. *Under the conditions*

$$\begin{cases} 0 < \sigma_2(k) \leq \sigma_1(k) & \text{if } \nu_1 < \nu_2, \\ 0 < \sigma_1(k) \leq \sigma_2(k) & \text{if } \nu_2 < \nu_1, \end{cases} \tag{3.10}$$

for all $k \neq 0$, the Schwarz iteration (3.7) *converges for all nonzero frequencies,*

$$\rho(k, \sigma_1, \sigma_2) < 1, \quad \text{for } k \neq 0.$$

Proof. For convergence, we want

$$\left| \frac{\sigma_1 - \nu_2|k|}{\sigma_1 + \nu_1|k|} \cdot \frac{\sigma_2 - \nu_1|k|}{\sigma_2 + \nu_2|k|} \right| < 1,$$

which can be written as two separate inequalities

$$-(\sigma_1 + \nu_1|k|)(\sigma_2 + \nu_2|k|) < (\sigma_1 - \nu_2|k|)(\sigma_2 - \nu_1|k|) < (\sigma_1 + \nu_1|k|)(\sigma_2 + \nu_2|k|).$$

By expanding the products, we find that the inequality on the right holds if and only if $\sigma_1 + \sigma_2 > 0$. For the inequality on the left, we want

$$0 < \sigma_1\sigma_2 + |k|(\nu_1 - \nu_2)(\sigma_2 - \sigma_1) + \nu_1\nu_2k^2.$$

Simple sufficient conditions for this inequality to hold are

$$\sigma_1\sigma_2 \geq 0 \quad \text{and} \quad (\nu_1 - \nu_2)(\sigma_2 - \sigma_1) \geq 0$$

(these are clearly not *necessary* conditions). Combined with the previous conclusion that $\sigma_1 + \sigma_2 > 0$, these reduce to the conditions (3.10). □

In the following subsections, we consider four choices of transmission conditions and their corresponding min-max problems: three different types of Robin conditions (which can be viewed as zeroth order approximations of σ_j^{opt}), and one choice of second order conditions. In all cases, the conditions of the Theorem 3.2 will eventually turn out to be satisfied, thus guaranteeing convergence when using optimized conditions for the model problem. Moreover, for a straight interface, when $\sigma_j(k) > 0$ the subproblems will be well-posed.

In the following, when solving optimization problems of the type (3.9), we are interested in characterizing *all* the global minimizers.

Definition 3.3. We say that two min-max problems are ***equivalent*** if they have the same minimum value *and* they have the same set of global minimizers.

3.4.1 Optimized Robin Conditions: First Version

Let us first derive optimized Robin transmission conditions with only one free parameter. The most obvious way to reduce σ_1 and σ_2 to one degree of freedom is to choose

$$\sigma_1(k) = \sigma_2(k) = p, \quad \text{with } p \in \mathbb{R}.$$

We can already see that this may not be the best choice, by noticing that the optimal symbols σ_j^{opt} are scaled differently in terms of ν_1 and ν_2, and appropriate conditions should probably imitate this scaling (see Section 3.4.2). Nonetheless, we still fully analyze this case for completeness and comparison purposes.

For this choice, we can write the convergence factor (3.8) as

$$\rho(k,p) = \left| \frac{(p - \nu_-|k|)(p - \nu_+|k|)}{(p + \nu_-|k|)(p + \nu_+|k|)} \right|,$$

where we introduce the notation

$$\nu_- := \min\{\nu_1, \nu_2\}, \quad \nu_+ := \max\{\nu_1, \nu_2\}.$$

We now wish to find the *best* value for the parameter p in the following sense: it should minimize the convergence factor uniformly over a relevant range of frequencies $[k_1, k_2]$. This is stated as the min-max problem

$$\min_{p\in\mathbb{R}} \left(\max_{k_1\le k\le k_2} \rho(k,p) \right). \tag{3.11}$$

We only look at a positive range of frequencies, $k_1 > 0$, since the convergence factor is an even function of k. This means we can replace $|k|$ by k for simplicity in the analysis. As a first step towards solving the optimization problem (3.11), we show how to reduce the search range for p.

Lemma 3.1 (Restricting the range for p). *The min-max problem* (3.11) *is equivalent to the problem where we minimize over p only in the interval $[\nu_- k_1, \nu_+ k_2]$.*

Proof. First, we can restrict ourselves to the case $p > 0$, by noticing that

$$|\rho(k,p)| < |\rho(k,-p)|, \quad \text{whenever } p > 0, \forall\, k > 0.$$

From Theorem 3.2, assuming that $p > 0$ ensures convergence of the iteration at this point. Suppose now that $p \notin [\nu_- k_1, \nu_+ k_2]$. Then, we have

$$(p - \nu_- k)(p - \nu_+ k) > 0 \text{ for all } k \in [k_1, k_2],$$

and so we can write the convergence factor without absolute values

$$\rho(k,p) = \frac{(p - \nu_- k)(p - \nu_+ k)}{(p + \nu_- k)(p + \nu_+ k)}.$$

Taking a derivative with respect to p, we find

$$\frac{\partial \rho}{\partial p}(k,p) = \frac{2k(\nu_- + \nu_+)(p^2 - \nu_-\nu_+ k^2)}{(p + \nu_- k)^2 (p + \nu_+ k)^2}.$$

In the case $p < \nu_- k_1$, we have

$$p^2 < \nu_-^2 k_1^2 \le \nu_-\nu_+ k^2, \quad \text{which implies that} \quad \frac{\partial \rho}{\partial p} < 0, \quad \forall\, k \in [k_1, k_2].$$

Thus, increasing p will make the convergence factor decrease uniformly on $[k_1, k_2]$. Similarly, in the case $p > \nu_+ k_2$, we have

$$p^2 > \nu_+^2 k_2^2 \geq \nu_- \nu_+ k^2, \quad \text{which implies that} \quad \frac{\partial \rho}{\partial p} > 0, \quad \forall\, k \in [k_1, k_2].$$

Thus, decreasing p will make the convergence factor decrease uniformly over k. This shows that any minimizer p solving problem (3.11) cannot be less than $\nu_- k_1$, nor can it be greater than $\nu_+ k_2$. □

We now turn to the behavior of $\rho(k, p)$ as a function of k.

Lemma 3.2 (Local maxima in k). *Let $k_c(p) := (\nu_- \nu_+)^{-\frac{1}{2}} p$. For fixed p, we can write the maximum value of $\rho(k, p)$ as*

$$\max_{k_1 \leq k \leq k_2} \rho(k, p) = \begin{cases} \max\{\, \rho(k_1, p),\ \rho(k_c(p), p),\ \rho(k_2, p)\,\} & \text{if } k_c(p) \in [k_1, k_2], \\ \max\{\, \rho(k_1, p),\ \rho(k_2, p)\,\} & \text{if } k_c(p) \notin [k_1, k_2]. \end{cases}$$

Proof. By differentiating $\rho(k, p)$ with respect to k (ignoring the points where ρ vanishes and is not differentiable), we find that

$$\frac{\partial \rho}{\partial k} = 0 \quad \text{if and only if} \quad k = k_c(p) = \frac{p}{\sqrt{\nu_- \nu_+}}.$$

It is easy to check that this critical point k_c is a local maximum of ρ. In fact, by looking at the sign of $\frac{\partial \rho}{\partial k}$ for a given $p > 0$, we observe that, with respect to k,

$$\rho(k, p) \text{ is } \begin{cases} \text{strictly decreasing for} & k \in \left[0, \frac{p}{\nu_+}\right) \cup \left(\frac{p}{\sqrt{\nu_- \nu_+}}, \frac{p}{\nu_-}\right), \\ \text{strictly increasing for} & k \in \left(\frac{p}{\nu_+}, \frac{p}{\sqrt{\nu_- \nu_+}}\right) \cup \left(\frac{p}{\nu_-}, \infty\right). \end{cases}$$

By observing that the local maximum $k_c(p)$ does not always lie in the interval $[k_1, k_2]$ for some values of p, we can express the maximum magnitude of the convergence factor as stated in the lemma. □

There are not enough degrees of freedom to be able to (always) match the value at the three possible local maxima; this would impose 2 conditions on 1 free parameter.

In fact, the value of the convergence factor at the interior local maximum $k_c(p)$ simplifies to

$$\rho(k_c(p), p) = \frac{(\sqrt{\nu_+} - \sqrt{\nu_-})^2}{(\sqrt{\nu_+} + \sqrt{\nu_-})^2} =: R_c,$$

which is, surprisingly, independent of the parameter p. In other words, R_c is a fixed value that cannot be improved by varying the parameter p.

The following theorem provides a complete characterization and formulas for the minimizers of the min-max problem (3.11). Although this is the first min-max problem we study here, the result and its proof are arguably the most complex of this chapter. Depending on the case, there can be a unique global minimizer, two distinct ones, or a full interval of minimizers. We make use of the notation

$$\mu := \frac{\nu_+}{\nu_-}, \qquad k_r = \frac{k_2}{k_1}.$$

Theorem 3.4 (Optimized Robin parameter: first version). *Let*

$$f(\mu) := \left[(\mu+1)^2 + (\mu-1)\sqrt{\mu^2+6\mu+1}\right](4\mu)^{-1}.$$

(i) If $k_r \geq f(\mu)$, then one value of p minimizing the convergence factor is $p^* = \sqrt{\nu_-\nu_+ k_1 k_2}$. This minimizer p^* is unique when $\rho(k_1, p^*) \geq R_c$. Otherwise, the minimum is also attained for any p chosen in a closed interval around p^*.

(ii) If $k_r < f(\mu)$, then two distinct values of p attain the minimum; they can be obtained by solving

$$\rho(k_1, p^*) = \rho(k_2, p^*)$$

in the intervals $[\nu_- k_1, \sqrt{\nu_-\nu_+}k_1]$ and $[\sqrt{\nu_-\nu_+}k_2, \nu_+ k_2]$ respectively. More precisely, these two distinct minimizers are the two positive roots of the biquadratic polynomial

$$p^4 + \left[\nu_-\nu_+(k_1^2+k_2^2) - k_1k_2(\nu_- + \nu_+)^2\right]p^2 + (\nu_-\nu_+k_1k_2)^2. \quad .$$

Proof. It was already shown that p can be restricted to the interval $[\nu_- k_1, \nu_+ k_2]$ in Lemma 3.1. The main idea is to look at three different subintervals for p,

$$I_l := [\nu_- k_1,\ \sqrt{\nu_-\nu_+}k_1], \quad I_c := [\sqrt{\nu_-\nu_+}k_1,\ \sqrt{\nu_-\nu_+}k_2], \quad I_r := [\sqrt{\nu_-\nu_+}k_2,\ \nu_+ k_2],$$

and find the best value of the parameter in each subinterval separately.

First, suppose that $p \in I_c$. We have that $k_c(p) = (\nu_-\nu_+)^{-\frac{1}{2}}p$ (the interior local maximum) lies in the interval $[k_1, k_2]$. Thus, from Lemma 3.2, the maximum value of the convergence factor is given by

$$\max_{k_1 \le k \le k_2} \rho(k,p) = \max\{\, \rho(k_1,p),\ R_c,\ \rho(k_2,p)\,\}.$$

To see how the endpoint values of the convergence factor behave as functions of p, we look at the partial derivative

$$\frac{\partial \rho}{\partial p}(k,p) = \operatorname{sign}((p-\nu_- k)(p-\nu_+ k))\frac{2k(\nu_- + \nu_+)(p^2 - \nu_-\nu_+ k^2)}{(p+\nu_- k)^2(p+\nu_+ k)^2}.$$

Analyzing the sign of this expression, we find, first for $k = k_1$.

$$\rho(k_1,p) \text{ is } \begin{cases} \text{decreasing} & \text{for } p < \nu_2 k_1, \\ \text{increasing} & \text{for } p > \nu_2 k_1, \end{cases}$$

and for $k = k_2$

$$\rho(k_2,p) \text{ is } \begin{cases} \text{decreasing} & \text{for } p < \nu_1 k_2, \\ \text{increasing} & \text{for } p > \nu_1 k_2. \end{cases}$$

Hence, the smallest value we can obtain by looking at the endpoints only is given by the equioscillation property $\rho(k_1,p) = \rho(k_2,p)$, leading to the parameter $p_c = \sqrt{\nu_-\nu_+ k_1 k_2}$. Note that the value at the interior local maximum, R_c, can be greater than the value at the endpoints, when $p = p_c$. In that case, we see that moving the parameter p in an interval around p_c still yields the same maximum value of the convergence factor (given by R_c).

Now, it remains to look at values of p in the intervals I_l and I_r, and compare the result with p_c. The situation in these two intervals is perfectly symmetric and it is

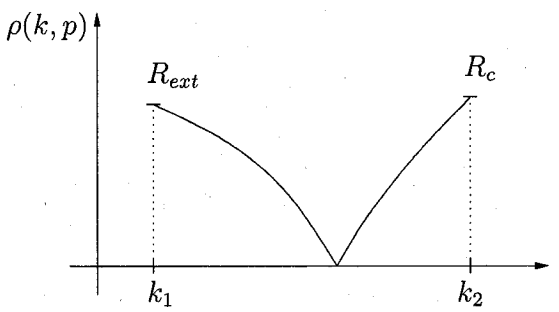

Figure 3.3: An illustration of the convergence factor when $p = \sqrt{\nu_-\nu_+}k_2$.

sufficient to consider only one of those, say $p \in I_r = [\sqrt{\nu_-\nu_+}k_2,\ \nu_+k_2]$. In this case, the local maximum $k_c(p)$ lies outside the relevant interval for k, and hence

$$\max_{k_1 \le k \le k_2} \rho(k,p) = \max\{\, \rho(k_1,p),\ \rho(k_2,p)\,\}.$$

When $p = \sqrt{\nu_-\nu_+}k_2$, we have $k_c(p) = k_2$. For this value of p, the convergence factor at the endpoints is

$$\rho(k_1,p) = \frac{(\sqrt{\mu}k_r - 1)|k_r - \sqrt{\mu}|}{(\sqrt{\mu}k_r + 1)(k_r + \sqrt{\mu})} =: R_{ext}, \qquad \rho(k_2,p) = \rho(k_c(p),p) = R_c.$$

We want to compare these two values. We now show that $R_{ext} < R_c$ if and only if $k_r < f(\mu)$. First, when $k_r < \sqrt{\mu} < f(\mu)$, we have

$$R_c - R_{ext} = \frac{2\sqrt{\mu}(\mu+1)(k_r-1)^2}{(\sqrt{\mu}+1)^2(\sqrt{\mu}k_r+1)(k_r+\sqrt{\mu})} > 0.$$

Now, if $k_r > \sqrt{\mu}$, by solving the equation $R_c = R_{ext}$ we get two roots

$$k_r^{\pm}(\mu) = \left[(\mu+1)^2 \pm (\mu-1)\sqrt{\mu^2+6\mu+1}\right](4\mu)^{-1}.$$

By looking more closely at the smallest root, we find

$$\begin{aligned} k_r^-(\mu) &= \left[(\mu+1)^2 - (\mu-1)\sqrt{\mu^2+6\mu+1}\right](4\mu)^{-1} \\ &\le \left[(\mu+1)^2 - (\mu-1)\sqrt{(\mu+1)^2}\right](4\mu)^{-1} \\ &\le \frac{\mu+1}{2\mu} < 1, \end{aligned}$$

since $\mu > 1$ by definition. Thus, we can disregard this root because we always have $k_r > 1$, and we are left with $k_r^+(\mu) := f(\mu)$. Finally, by inspection, we have $R_{ext} = 0$ when $k_r = \sqrt{\mu}$ and that $R_{ext} \to 1$ as $k_r \to \infty$, so we can conclude that

$$\begin{cases} R_{ext} < R_c & \text{when } k_r < f(\mu), \\ R_{ext} = R_c & \text{when } k_r = f(\mu), \\ R_{ext} > R_c & \text{when } k_r > f(\mu). \end{cases}$$

Now that this is established, the argument splits into two separate cases.

(i) If $k_r \geq f(\mu) > \sqrt{\mu}$, we have that $R_{ext} \geq R_c$. The value $\rho(k_1, p)$ increases as we increase p, so we cannot improve the convergence factor for $p > \sqrt{\nu_- \nu_+}k_2$. And because $R_{ext} \geq R_c \geq \max_k \rho(k, p_c)$, $p = p_c$ as defined above is a global minimizer.

(ii) If $k_r < f(\mu)$, we get $R_{ext} < R_c$. In this case, the value $\rho(k_2, p)$ is strictly decreasing and reaches 0 when $p = \nu_2 k_2$. The value $\rho(k_1, p)$ is eventually growing as a function of p, so we can find a unique parameter value $p_r \in I_r$ such that

$$\rho(k_1, p_r) = \rho(k_2, p_r) < R_c,$$

which beats the best convergence factor that can be obtained for $p \in I_c$. By the same argument, we can show there is also a value $p_l \in I_l$ that satisfies the above equation (with $\rho(k_1, p_l) = \rho(k_1, p_r)$). So, we get two distinct minimizers p_r and p_l, which can be computed by finding the two positive roots of the biquadratic polynomial shown in the theorem statement. Explicit formulas for p_r and p_l can easily be written down.

We considered all possible scenarios, so this completes the proof. □

Figure 3.4 shows four possible behaviors of the optimized convergence factor, depending on how the ratio of coefficients compares to the ratio of frequencies. First, in the upper-left figure, the minimizer p_c is unique. In the upper-right and bottom-left graphs, any p in an interval attains the same minimum, because R_c is fixed. In the bottom-right graph, we are in the case when two distinct values of p (p_l and p_r) minimize the convergence factor.

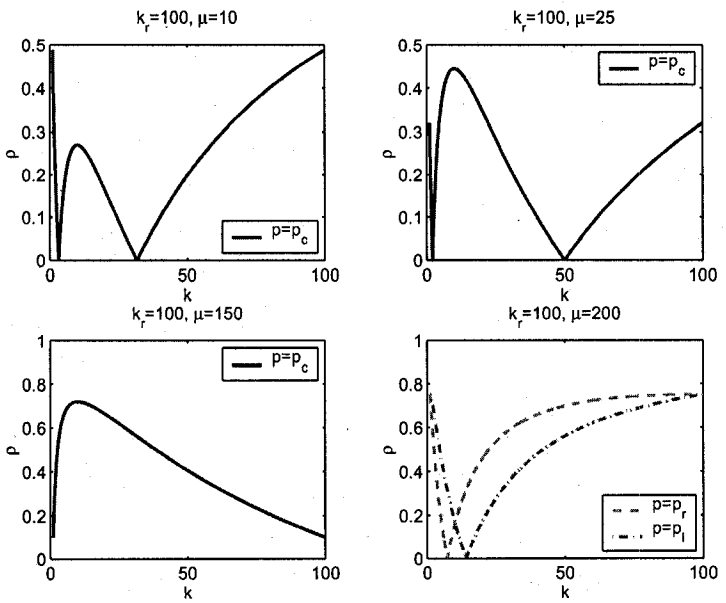

Figure 3.4: Optimized convergence factor for different ratios of coefficients.

Theorem 3.5 (Asymptotic performance). *When ν_1 and ν_2 are kept constant, $k_2 = \frac{\pi}{h}$ and h is small enough, the optimized Robin parameter given by Theorem 3.4 is $p^* = \sqrt{\nu_- \nu_+ k_1 \pi} h^{-1/2}$, and the asymptotic convergence factor of the Schwarz method is*

$$\max_{k_1 \le k \le \pi/h} \rho(k, p^*) = 1 - 2\left(\sqrt{\mu} + \frac{1}{\sqrt{\mu}}\right)\left(\frac{k_1 h}{\pi}\right)^{\frac{1}{2}} + O(h). \tag{3.12}$$

Proof. For k_2 large enough, we get $k_r > f(\mu)$, and so $p^* = p_c = \sqrt{\nu_- \nu_+ k_1 \pi} h^{-1/2}$ is a minimizer of the convergence factor. Also, $\rho(k_1, p^*)$ approaches 1 as $h \to 0$, and R_c is a constant. Thus, for h small enough, we have $\rho(k_1, p^*) > R_c$ and the minimizer p^* becomes unique. Finally, expanding $\rho(k_1, p^*)$ for small h gives the stated result. □

By setting $\nu_1 = \nu_2 =: \nu$ in the above two theorems, we recover the results in the case of constant coefficient, in agreement with [32], namely that

$$p^* = \nu\sqrt{k_1 k_2}, \qquad \max_{k_1 \le k \le \pi/h} \rho(k, p^*) = 1 - 4\sqrt{k_1}\pi h^{\frac{1}{2}} + O(h). \tag{3.13}$$

3.4.2 Optimized Robin Conditions: Second Version

From Figure 3.4, it can be noticed that the optimized convergence factor is not very good when the ratio of coefficients μ is large. This can be justified by the fact that it is not appropriate to approximate the optimal symbols $\sigma_1^{opt} = \nu_2|k|$ and $\sigma_2^{opt} = \nu_1|k|$ with the *same* parameter value when ν_1 and ν_2 are very different.

There is a second choice that we can make, which is directly driven by the form of the optimal symbols. To be consistent with these symbols, we now use a different scaling of the Robin parameters,

$$\sigma_1(k) = \nu_2 q, \quad \sigma_2(k) = \nu_1 q, \quad \text{for } q \in \mathbb{R}.$$

In this case, the convergence factor can be written, for $k > 0$, as

$$\rho(k, q) = \frac{(q-k)^2}{(q+\mu k)(q+k/\mu)},$$

where $\mu := \frac{\nu_+}{\nu_-}$ as before. The condition $q > 0$ will again be sufficient to ensure convergence, by Theorem 3.2, but this condition is not assumed a priori. The min-max problem we wish to solve to find the optimized value for q is

$$\min_{q \in \mathbb{R}} \left(\max_{k_1 \le k \le k_2} \rho(k, q) \right). \tag{3.14}$$

It turns out that this optimization problem is a lot easier to solve than the previous one, (3.11), and more importantly it leads to a better convergence factor.

Theorem 3.6 (Optimized Robin parameter: second version). *The unique optimized Robin parameter q^* solving the min-max problem* (3.14) *is given by the formula $q^* = \sqrt{k_1 k_2}$.*

Proof. By taking partial derivatives of the convergence factor with respect to q and k respectively we find

$$\operatorname{sign}\left(\frac{\partial \rho}{\partial q}\right) = \operatorname{sign}(q-k), \quad \operatorname{sign}\left(\frac{\partial \rho}{\partial k}\right) = \operatorname{sign}(k-q).$$

From these facts, we deduce the following properties.

(i) We can restrict the range of q to the interval $[k_1, k_2]$, otherwise the convergence factor is uniformly improved by appropriately increasing or decreasing q.

(ii) The convergence factor, as a function of k, is strictly decreasing for $k \in (k_1, q)$, and strictly increasing for $k \in (q, k_2)$.

(iii) $\rho(k_1, q)$ is increasing with respect to q, and $\rho(k_1, k_1) = 0$.

(iv) $\rho(k_2, q)$ is decreasing with respect to q, and $\rho(k_2, k_2) = 0$.

We can thus conclude that we minimize the convergence factor uniformly when the values at k_1 and k_2 are equal, i.e. we have the equioscillation property

$$\rho(k_1, q^*) = \rho(k_2, q^*).$$

Solving this equation for q^*, we find $q^* = \sqrt{k_1 k_2}$. □

Theorem 3.7 (Asymptotic performance). *When ν_1 and ν_2 are kept constant, $k_2 = \frac{\pi}{h}$ and h goes to 0, the optimized Robin parameter given by Theorem 3.6 is $q^* = \sqrt{k_1 \pi} h^{-\frac{1}{2}}$ and the asymptotic convergence factor of the Schwarz method is*

$$\max_{k_1 \le k \le \frac{\pi}{h}} \rho(k, q^*) = 1 - \frac{(\mu+1)^2}{\mu} \left(\frac{k_1 h}{\pi}\right)^{\frac{1}{2}} + O(h). \tag{3.15}$$

Proof. The result is obtained by simply expanding $\rho(k_1, q^*)$ for small values of h. □

Again, when we set $\nu_1 = \nu_2 =: \nu$ in the above two theorems, we get the appropriate optimized conditions for the case of continuous diffusion coefficient, exactly as in (3.13).

We now proceed to show that, asymptotically, this second choice of scaling of the Robin parameter is better than the first one studied in Section 3.4.1. By comparing (3.12) and (3.15), we see that both choices give similar asymptotic convergence factors, in the form $1 - Ch^{1/2}$, but with different constants C. For comparison, the constants for the first and second choice of Robin conditions are respectively

$$C_1 = 2\frac{(\nu_1 + \nu_2)}{\sqrt{\nu_1 \nu_2}} \sqrt{\frac{k_1}{\pi}} \qquad \text{and} \qquad C_2 = \frac{(\nu_1 + \nu_2)^2}{\nu_1 \nu_2} \sqrt{\frac{k_1}{\pi}}.$$

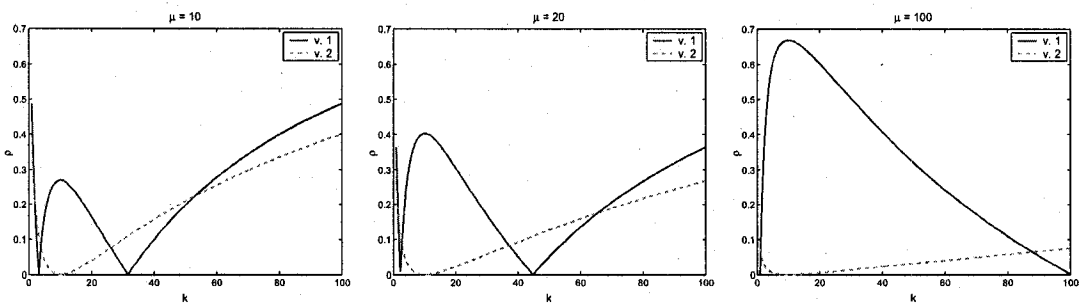

Figure 3.5: Comparison of the convergence factors for the two versions of optimized Robin conditions. Here, $\nu_2 = k_1 = 1$, $k_2 = 100$, and only ν_1 is varied.

Noting that $2x \leq x^2$ when $x \geq 2$, and that in this case $x = \frac{(\nu_1+\nu_2)}{\sqrt{\nu_1\nu_2}} \geq 2$ for any $\nu_1, \nu_2 > 0$, we have $C_1 \leq C_2$, with equality only if $\nu_1 = \nu_2$. Thus, this shows that the *second* version of optimized Robin conditions yields a better (i.e. larger) constant in the asymptotic convergence factor. In practice, it seems that the optimized convergence factor for the second version is *always* smaller than for the first version, not only asymptotically.

Figure 3.5 shows a comparison of the optimized convergence factors, when using the two choices of Robin conditions studied so far. When the ratio of coefficients is large, the second version yields a much smaller convergence factor. We will provide theoretical justification of this observation in Section 3.4.6, and confirm it with numerical experiments as well.

3.4.3 Optimized Two-Sided Robin Conditions

Let us now consider the most general Robin conditions, with two free parameters

$$\sigma_1(k) = \nu_2 p_1, \quad \sigma_2(k) = \nu_1 p_2, \quad \text{with } p_1, p_2 \in \mathbb{R}.$$

Again, we have scaled σ_j consistently with the optimal symbols, but here it is only to simplify the calculations; it does not matter since p_1 and p_2 are two independent free parameters.

Let $\lambda := \frac{\nu_1}{\nu_2}$ (not to be confused with μ, which is always greater than 1). The convergence factor can be written as

$$\rho(k, p_1, p_2) = \left| \frac{(p_1 - k)(p_2 - k)}{(p_1 + \lambda k)(p_2 + k/\lambda)} \right|.$$

Before finding optimized values, let us first state sufficient (but not necessary) conditions on p_1 and p_2 to obtain convergence of the iteration, as a corollary of Theorem 3.2.

Corollary 3.1. *Suppose the Robin parameters p_1, $p_2 \in \mathbb{R}$ satisfy the orderings*

$$\begin{cases} 0 < p_2 \leq p_1 & \text{when } \nu_1 < \nu_2, \\ 0 < p_1 \leq p_2 & \text{when } \nu_2 < \nu_1. \end{cases}$$

Then, we have $\rho(k, p_1, p_2) < 1$ for all $k > 0$.

The min-max problem to solve for optimized two-sided Robin conditions is

$$\min_{p_1, p_2 \in \mathbb{R}} \left(\max_{k_1 \leq k \leq k_2} \rho(k, p_1, p_2) \right). \tag{3.16}$$

We will solve this optimization problem and show that the unique optimized parameters p_1^* and p_2^* satisfy an equioscillation property of the convergence factor. In the process, we will show that the optimized parameters satisfy the conditions of Corollary 3.1, hence guaranteeing convergence. We break down the proof into a sequence of lemmas, but the main steps are the same as in the previous sections, namely

(1) restrict the range for the parameters,

(2) locate potential candidates for local maxima in k,

(3) analyze how these local maxima behave when varying the parameters.

Lemma 3.3. *It is sufficient to consider only positive values for the parameters p_1 and p_2, i.e. the min-max problem* (3.16) *is equivalent to the problem*

$$\min_{p_1, p_2 > 0} \left(\max_{k_1 \leq k \leq k_2} \rho(k, p_1, p_2) \right).$$

Proof. For $p_1 < 0$, it is easy to see that $\rho(k, p_1, p_2) > \rho(k, -p_1, p_2)$. Similarly, for $p_2 < 0$, $\rho(k, p_1, p_2) > \rho(k, p_1, -p_2)$. Moreover, $\frac{\partial \rho}{\partial p_1}(k, 0, p_2) < 0$ and $\frac{\partial \rho}{\partial p_2}(k, p_1, 0) < 0$ for all $k > 0$, thus excluding also the values $p_1 = 0$ and $p_2 = 0$. □

Lemma 3.4. *If $\lambda > 1$, then* (3.16) *is equivalent to the problem*

$$\min_{0<p_1\leq p_2} \left(\max_{k_1\leq k\leq k_2} \rho(k, p_1, p_2) \right).$$

If $\lambda < 1$, then (3.16) *is equivalent to the problem*

$$\min_{0<p_2\leq p_1} \left(\max_{k_1\leq k\leq k_2} \rho(k, p_1, p_2) \right).$$

Without loss of generality, we can look only at the min-max problem in the case $\lambda > 1$: the other case ($\lambda < 1$) reduces to the first one by interchanging p_1 and p_2 and replacing λ by $1/\lambda$.

Proof. Suppose $\lambda > 1$ and $p_1 > p_2$. Then,

$$\rho(k, p_1, p_2) = \frac{|k^2 - (p_1 + p_2)k + p_1 p_2|}{|k^2 + (p_1/\lambda + \lambda p_2)k + p_1 p_2|}.$$

By interchanging the values of p_1 and p_2, the numerator of the fraction is unchanged, but the denominator is increased, since

$$\frac{p_1}{\lambda} + \lambda p_2 < \frac{p_2}{\lambda} + \lambda p_1 \quad \text{when } p_1 < p_2 \text{ and } \lambda > 1.$$

Hence, in this case, $\rho(k, p_1, p_2) > \rho(k, p_2, p_1)$, i.e. the convergence factor is uniformly improved by interchanging p_1 and p_2. Thus, it is sufficient to consider only pairs of parameters such that $p_1 \leq p_2$. The case when $\lambda < 1$ is treated the same way. □

Lemma 3.4 implies that the smaller of the two parameters is associated with the subdomain having the larger diffusion coefficient (e.g. $p_1 < p_2$ when $\nu_2 < \nu_1$). At this point, note that the conditions of Corollary 3.1 are now satisfied. From now on, we assume without loss of generality that $\lambda > 1$, and consequently that $p_1 \leq p_2$.

Lemma 3.5. *We can restrict the range of the parameters p_1 and p_2 to the interval $[k_1, k_2]$: when $\lambda > 1$, the min-max problem (3.16) is equivalent to the problem*

$$\min_{k_1 \le p_1 \le p_2 \le k_2} \left(\max_{k_1 \le k \le k_2} \rho(k, p_1, p_2) \right).$$

Proof. Taking partial derivatives with respect to the parameters, and analyzing the sign, we find

$$\operatorname{sign}\left(\frac{\partial \rho}{\partial p_1}\right) = \begin{cases} \text{strictly positive} & \text{when } k < p_1, \\ \text{strictly negative} & \text{when } k > p_1, \end{cases}$$

$$\operatorname{sign}\left(\frac{\partial \rho}{\partial p_2}\right) = \begin{cases} \text{strictly positive} & \text{when } k < p_2, \\ \text{strictly negative} & \text{when } k > p_2. \end{cases}$$

Thus, when $p_1 < k_1$, increasing p uniformly improves the convergence factor. When $p_1 > k_2$, decreasing p uniformly improves the convergence factor. A similar argument holds for p_2 as well. □

Lemma 3.6 (Local maxima in k). *The maximum of the convergence factor on the interval $[k_1, k_2]$ can be computed by looking at three points only,*

$$\max_{k_1 \le k \le k_2} \rho(k, p_1, p_2) = \max\{\ \rho(k_1, p_1, p_2),\ \rho(\sqrt{p_1 p_2}, p_1, p_2),\ \rho(k_2, p_1, p_2)\ \}. \tag{3.17}$$

Proof. Again, this is straightforward. Analyzing the derivative of ρ with respect to k, we find

$$\operatorname{sign}\left(\frac{\partial \rho}{\partial k}\right) = \begin{cases} -\operatorname{sign}(p_1 p_2 - k^2) & \text{when } k \notin [p_1, p_2], \\ \operatorname{sign}(p_1 p_2 - k^2) & \text{when } k \in [p_1, p_2], \end{cases}$$

which implies that

$$\rho(k, p_1, p_2) \text{ is } \begin{cases} \text{strictly decreasing} & \text{for } k \in [k_1, p_1) \cup (\sqrt{p_1 p_2}, p_2), \\ \text{strictly increasing} & \text{for } k \in (p_1, \sqrt{p_1 p_2}) \cup (p_2, k_2]. \end{cases}$$

Thus, the convergence factor $\rho(k, p_1, p_2)$ has a local maximum at $k = \sqrt{p_1 p_2}$ (unless $p_1 = p_2$). □

Figure 3.6 illustrates this result. Our aim is to show that the optimized parameters p_1^* and p_2^* are obtained by an equioscillation of these three local maxima. We first start by showing the equioscillation property with the two endpoints.

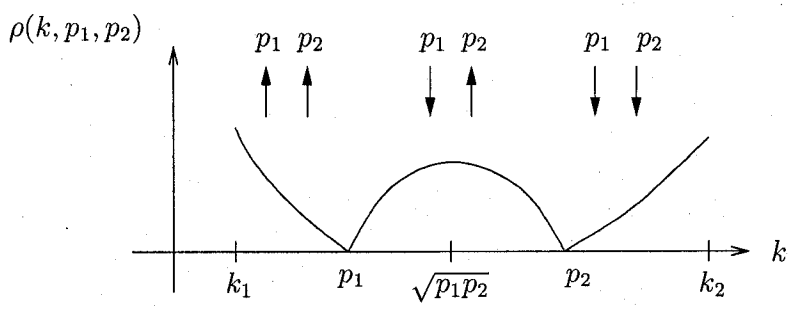

Figure 3.6: A sketch of how the convergence factor behaves as a function of k.

Lemma 3.7 (Equioscillation with the endpoints). *The optimized convergence factor $\rho(k, p_1^*, p_2^*)$ must satisfy the equioscillation property at the endpoints, i.e.*

$$\rho(k_1, p_1^*, p_2^*) = \rho(k_2, p_1^*, p_2^*),$$

which holds if and only if $p_1^* p_2^* = k_1 k_2$.

Proof. We look again at the behavior of $\rho(k, p_1, p_2)$ as we vary the parameters. The partial derivatives were already computed in the proof of Lemma 3.5. From those we can deduce the behavior shown in Figure 3.6: the arrows below p_1 and p_2 indicate the change in $\rho(k, p_1, p_2)$ (whether it is increasing or decreasing) when p_1 and p_2 are *increased*, independently.

Let us compare the values of the convergence factor at the endpoints. Suppose first that $\rho(k_1, p_1, p_2) < \rho(k_2, p_1, p_2)$. Then, from the Figure 3.6 we can observe that increasing p_1 uniformly improves ρ. In the other case, when $\rho(k_1, p_1, p_2) > \rho(k_2, p_1, p_2)$, decreasing p_2 uniformly improves the convergence factor. One can also easily check that these operations can always be done while staying inside the allowed range for the parameters. Thus, at the optimized parameters p_1^* and p_2^*, we must have $\rho(k_1, p_1^*, p_2^*) = \rho(k_2, p_1^*, p_2^*)$. Some additional algebraic manipulations of this identity lead to the equation $p_1^* p_2^* = k_1 k_2$. □

As a consequence of this theorem, the optimized parameters must satisfy

$$k_1 \leq p_1^* \leq \sqrt{k_1 k_2} \leq p_2^* \leq k_2.$$

We now have enough tools to complete the proof of the main result.

Theorem 3.8 (Optimized two-sided Robin parameters). *When $\lambda > 1$, the unique minimizing pair (p_1^*, p_2^*) of problem (3.16) is the unique solution of the system of equations*

$$p_1^* p_2^* = k_1 k_2, \tag{3.18}$$

$$\rho(k_1, p_1^*, p_2^*) = \rho(\sqrt{p_1^* p_2^*}, p_1^*, p_2^*), \tag{3.19}$$

such that $p_1^* \leq p_2^*$.

Proof. Lemma 3.7 directly implies equation (3.18), so we set $p_2 = k_1 k_2 p_1^{-1}$ and note that in this situation $\rho(k_1, p_1, p_2) = \rho(k_2, p_1, p_2)$ is automatically satisfied. Thus we have reduced the optimization problem (3.16) to solving a one-parameter min-max problem

$$\min_{k_1 \leq p_1 \leq \sqrt{k_1 k_2}} \left(\max\{R_1(p_1), R_c(p_1)\} \right), \tag{3.20}$$

where we use the notation

$$R_1(p_1) := \rho\left(k_1,\, p_1,\, \frac{k_1 k_2}{p_1}\right), \quad R_c(p_1) := \rho\left(\sqrt{k_1 k_2},\, p_1,\, \frac{k_1 k_2}{p_1}\right).$$

It remains to show that the unique solution of problem (3.20) occurs when $R_1(p_1^*) = R_c(p_1^*)$. By direct computation, we find that

$$\operatorname{sign}\left(\frac{dR_c}{dp_1}\right) = \operatorname{sign}(p_1 - \sqrt{k_1 k_2}) < 0, \quad \forall p_1 \in [k_1, \sqrt{k_1 k_2}),$$

so the value of $R_c(p_1)$ is strictly decreasing in p_1.

Now, expanding the expression for $R_1(p_1)$, we can write it as

$$R_1(p_1) = \frac{k_1^2 + k_1 k_2 - \left(p_1 + \frac{k_1 k_2}{p_1}\right) k_1}{k_1^2 + k_1 k_2 + \left(\frac{p_1}{\lambda} + \lambda \frac{k_1 k_2}{p_1}\right) k_1} =: \frac{[N]}{[D]},$$

where the numerator $[N]$ and denominator $[D]$ of the fraction are nonnegative. Taking a derivative in p_1 (using the quotient rule), we get

$$\frac{dR_1}{dp_1} = \frac{k_1}{p_1^2}\left[\left(k_1 k_2 - p_1^2\right)[D] + \frac{1}{\lambda}\left(\lambda^2 k_1 k_2 - p_1^2\right)[N]\right] / [D]^2.$$

Examining both terms in the numerator of this fraction, we see they are both non-negative since, using the assumption that $\lambda > 1$,

$$\lambda^2 k_1 k_2 > k_1 k_2 = p_1 p_2 \geq p_1^2.$$

Thus, $R_1(p_1)$ is strictly increasing in p_1 on the interval $[k_1, \sqrt{k_1 k_2})$.

In addition, at the extremal values of p_1, we have

$$\begin{aligned} 0 = R_1(k_1) &< R_c(k_1), \\ R_1(\sqrt{k_1 k_2}) &> R_c(\sqrt{k_1 k_2}) = 0. \end{aligned}$$

These facts are sufficient to conclude that there exists a unique $p_1^* \in (k_1, \sqrt{k_1, k_2})$ such that $R_1(p_1^*) = R_c(p_1^*)$, giving the unique solution to the optimization problem (3.16). □

Theorem 3.8 states that the unique optimized pair of parameter (p_1^*, p_2^*) can be found by solving the system of nonlinear equations (3.18)-(3.19). However, some extra algebraic manipulations allow us to show that computing p_1^* can be reduced to the task of finding the unique real root, in the interval $(k_1, \sqrt{k_1 k_2})$, of the quartic

$$(p_1 + \lambda k_1)(p_1 + \lambda k_2)(\sqrt{k_1 k_2} - p_1)^2 - (p_1 - k_1)(k_2 - p_1)(p_1 + \lambda\sqrt{k_1 k_2})^2. \qquad (3.21)$$

Hence, there are explicit formulas for computing p_1^* and p_2^* (which we omit for simplicity); there is no need to use a nonlinear solver for (3.18)-(3.19).

Theorem 3.9 (Asymptotic performance). *When $\lambda = \mu > 1$ is fixed, $k_2 = \frac{\pi}{h}$ and h goes to 0, the optimized two-sided Robin parameters are*

$$p_1^* \sim \frac{2\mu k_1}{\mu - 1} - \frac{4\mu k_1^{3/2}}{\sqrt{\pi}} \frac{(\mu+1)^2}{(\mu-1)^3} h^{\frac{1}{2}}, \quad p_2^* \sim \frac{\pi(\mu-1)}{2\mu} h^{-1} + \sqrt{k_1 \pi} \frac{(\mu+1)^2}{\mu(\mu-1)} h^{-\frac{1}{2}},$$

and the asymptotic convergence factor of the Schwarz method is

$$\max_{k_1 \leq k \leq \frac{\pi}{h}} \rho(k, p_1^*, p_2^*) = \frac{1}{\mu} - \frac{4(\mu+1)}{\mu(\mu-1)} \sqrt{\frac{k_1}{\pi}} h^{\frac{1}{2}} + O(h). \qquad (3.22)$$

Proof. We first make the ansatz $p_1^* = Ah^\alpha$ for some $\alpha > -\frac{1}{2}$, since we know that $p_1^* < \sqrt{k_1 k_2} = O(h^{-\frac{1}{2}})$. Examining the quartic polynomial (3.21) and keeping only the leading order term, we get

$$(Ah^\alpha + \mu k_1) - \mu(Ah^\alpha - k_1) = 0,$$

and so we must have $\alpha = 0$ and $A = \frac{2\mu}{\mu - 1}$. In order to get the first two terms in the asymptotic performance of the Schwarz method, we must also compute a second term in the expansion for p_1^*. So, we now make the ansatz $p_1^* = A + Bh^\beta$ for some $\beta > 0$. Looking at the second leading order term when expanding (3.21) we find

$$\beta = \frac{1}{2}, \quad B = -\frac{4k_1^{\frac{3}{2}}\mu(\mu+1)^2}{\sqrt{\pi}(\mu-1)^3}.$$

The asymptotic formula for p_2^* is then obtained using the relation $p_2^* = k_1 k_2 / p_1^*$. Finally, expanding $R_1(p_1^*)$ for small h with the above values, we find the expansion (3.22). □

Theorems 3.8 and 3.9 are stated for the case $\lambda > 1$ only. When $\lambda < 1$, Lemma 3.4 shows that, by simply interchanging the roles of p_1 and p_2, and replacing λ by $1/\lambda$, we can still apply these results. More precisely, if $\lambda < 1$, we can first compute the optimized parameters $\tilde{p}_1^*$ and $\tilde{p}_2^*$ corresponding to $\tilde{\lambda} = 1/\lambda$ using Theorem 3.8, and then interchange the parameters, namely use $p_1^* = \tilde{p}_2^*$ and $p_2^* = \tilde{p}_1^*$. The asymptotic expansion of the convergence factor (3.22) still holds in that case, with $\mu = 1/\lambda$.

When $\nu_1 = \nu_2$ (or as $\mu \to 1$), Theorem 3.8 does not nicely reduce to the results obtained for the case of continuous diffusion coefficient. Moreover, the asymptotic convergence factor (3.22) is no longer valid (it degenerates) as $\mu \to 1$. In fact, this asymptotic convergence factor is somewhat surprising. From the known results for continuous coefficient (see [32]), we would have expected an asymptotic expansion of the form $1 - O(h^{1/4})$. Instead, for discontinuous coefficients, we find that, when using two-sided optimized Robin transmission conditions, we get a mesh independent convergence: the convergence factor is bounded away from 1, asymptotically as $h \to 0$.

3.4.4 Optimized Second Order Conditions

It is also possible to use transmission conditions that include tangential derivatives along the interface. For example, we now consider the following second order transmission conditions

$$\left[\nu_1\frac{\partial}{\partial x}+\nu_2\left(p-q\frac{\partial^2}{\partial y^2}\right)\right]u_1^{n+1}(0,y)=\left[\nu_2\frac{\partial}{\partial x}+\nu_2\left(p-q\frac{\partial^2}{\partial y^2}\right)\right]u_2^{n}(0,y),$$
$$\left[-\nu_2\frac{\partial}{\partial x}+\nu_1\left(p-q\frac{\partial^2}{\partial y^2}\right)\right]u_2^{n+1}(0,y)=\left[-\nu_1\frac{\partial}{\partial x}+\nu_1\left(p-q\frac{\partial^2}{\partial y^2}\right)\right]u_1^{n}(0,y),$$

on the interface $x=0$. These are *one-sided* transmission conditions in the sense that the same parameters are used in both directions across the interface, and we could easily imagine a *two-sided* version with four free parameters. Also note that we have properly scaled these conditions in view of the optimal operators, as in Section 3.4.2 and 3.4.3. In Fourier space, the corresponding symbols are

$$\sigma_1(k)=\nu_2(p+qk^2),\quad \sigma_2(k)=\nu_1(p+qk^2).$$

This is equivalent to approximating the optimal symbols σ_j^{opt} with second order polynomials in k. We do not include a first order term, since σ_j^{opt} are even functions of k (the underlying differential operator is self-adjoint). The convergence factor in this situation is

$$\rho(k,p,q)=\frac{(p+qk^2-k)^2}{(p+qk^2+\mu k)(p+qk^2+k/\mu)}.$$

We will assume a priori that the parameters are strictly positive, for simplicity, and look to solve the associated optimization problem

$$\min_{p,q>0}\left(\max_{k_1\le k\le k_2}\rho(k,p,q)\right). \tag{3.23}$$

Lemma 3.8. *Any minimizing pair (p^*,q^*) of problem (3.23) must satisfy the inequality $p^*q^*\le\frac{1}{4}$.*

Proof. Computing partial derivatives with respect to the parameters, we find that

$$\operatorname{sign}\left(\frac{\partial\rho}{\partial p}\right)=\operatorname{sign}\left(\frac{\partial\rho}{\partial q}\right)=\operatorname{sign}(p+qk^2-k).$$

When $pq > \frac{1}{4}$, the roots of the quadratic $p + qk^2 - k$ are imaginary, hence implying that $\frac{\partial \rho}{\partial p} > 0$ and $\frac{\partial \rho}{\partial q} > 0$ for all $k > 0$. This means that the convergence factor can be uniformly reduced by decreasing p or q or both, until $pq \leq \frac{1}{4}$. □

When the coefficients are continuous, it is known that the optimization of two-sided Robin conditions and one-sided second order conditions are closely related problems, see [32]. The proof of the following theorem is inspired by this relation, but yet it leads to a very different result when compared to Section 3.4.3.

Theorem 3.10 (Optimized second order parameters). *The min-max problem* (3.23) *has a unique minimizing pair* (p^*, q^*)*, given by the equioscillation of the convergence factor at the frequencies* k_1*,* k_2 *and* $k_c = \sqrt{p/q}$*. This gives the formulas*

$$p^* = \frac{(k_1 k_2)^{\frac{3}{4}}}{\sqrt{2(k_1 + k_2)}}, \quad q^* = \frac{1}{\sqrt{2(k_1 + k_2)}(k_1 k_2)^{\frac{1}{4}}}. \tag{3.24}$$

Proof. It will be easier to show the equioscillation property for a transformed problem. Let us change the parameters (p, q) to $(\tilde{p}_1, \tilde{p}_2)$, where

$$\tilde{p}_1 := \frac{1 - \sqrt{1 - 4pq}}{2q}, \quad \tilde{p}_2 := \frac{1 + \sqrt{1 - 4pq}}{2q}.$$

The new parameters $\tilde{p}_1$ and $\tilde{p}_2$ are simply the zeros of the convergence factor ρ. This change of variable is well-defined when $p > 0$, $q > 0$ and $pq \leq \frac{1}{4}$; the latter inequality was established in Lemma 3.8. The inverse transformation of the parameters is given by

$$p = \frac{\tilde{p}_1 \tilde{p}_2}{\tilde{p}_1 + \tilde{p}_2}, \quad q = \frac{1}{\tilde{p}_1 + \tilde{p}_2}.$$

The convergence factor in terms of $\tilde{p}_1$ and $\tilde{p}_2$ becomes

$$\rho(k, p, q) = R(k, \tilde{p}_1, \tilde{p}_2) := \frac{(\tilde{p}_1 - k)^2 (\tilde{p}_2 - k)^2}{[\,\tilde{p}_1 \tilde{p}_2 + k^2 + \mu(\tilde{p}_1 + \tilde{p}_2)\,]\,[\,\tilde{p}_1 \tilde{p}_2 + k^2 + \frac{1}{\mu}(\tilde{p}_1 + \tilde{p}_2)\,]},$$

and the associated min-max problem is

$$\min_{0 < \tilde{p}_1 \leq \tilde{p}_2} \left(\max_{k_1 \leq k \leq k_2} R(k, \tilde{p}_1, \tilde{p}_2) \right).$$

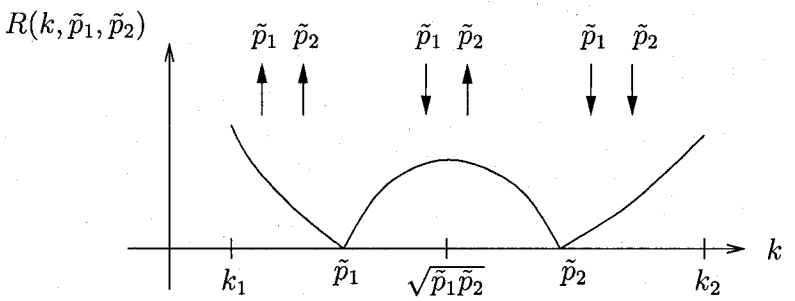

Figure 3.7: A sketch of how the convergence factor $R(k, \tilde{p}_1, \tilde{p}_2)$ behaves as a function of the frequency k.

This optimization problem looks similar to (3.16), for finding optimized two-sided Robin conditions, but it is in fact significantly different when $\mu \neq 1$. Solving this modified problem with the usual sequence of arguments will prove easier. Looking at the sign of the various derivatives, we find

$$\operatorname{sign}\left(\frac{\partial R}{\partial \tilde{p}_1}\right) = \operatorname{sign}(\tilde{p}_1 - k), \quad \operatorname{sign}\left(\frac{\partial R}{\partial \tilde{p}_2}\right) = \operatorname{sign}(\tilde{p}_2 - k),$$
$$\operatorname{sign}\left(\frac{\partial R}{\partial k}\right) = \operatorname{sign}\left[(\tilde{p}_1 - k)(k - \tilde{p}_2)(\tilde{p}_1\tilde{p}_2 - k^2)\right].$$

The derivatives with respect to the parameters lead us to conclude the bounds

$$k_1 \leq \tilde{p}_1 \leq \tilde{p}_2 \leq k_2. \tag{3.25}$$

Furthermore, taking into account the derivative with respect to k, we are able to deduce the behavior shown in Figure 3.7, which is exactly the same as the one we obtained for two-sided Robin conditions, in Figure 3.6.

Thus, using the same argument as in the proof of Lemma 3.7, the optimized parameters must satisfy an equioscillation at the endpoints, namely

$$R(k_1, \tilde{p}_1^*, \tilde{p}_2^*) = R(k_2, \tilde{p}_1^*, \tilde{p}_2^*).$$

Solving this algebraic equation for $\tilde{p}_2$ as a function of $\tilde{p}_1$, we get three possible formulas:

(1) $\tilde{p}_2 = f_1(\tilde{p}_1) := \frac{k_1 k_2}{\tilde{p}_1}$,

(2) $\tilde{p}_2 = f_2(\tilde{p}_1) := x(\tilde{p}_1) + \sqrt{x(\tilde{p}_1)^2 + k_1 k_2}$ if $\tilde{p}_1 < \sqrt{k_1 k_2}$,

(3) $\tilde{p}_2 = f_3(\tilde{p}_1) := \sqrt{x(\tilde{p}_1)^2 + k_1 k_2} - x(\tilde{p}_1)$ if $\tilde{p}_1 > \sqrt{k_1 k_2}$,

where

$$x(\tilde{p}_1) := \frac{\tilde{p}_1 (k_2 - k_1)^2}{2|\tilde{p}_1^2 - k_1 k_2|}.$$

We now proceed to show that the last two formulas lead to parameter values outside the bounds (3.25) established earlier.

Suppose first that $k_1 < \tilde{p}_1 < \sqrt{k_1 k_2}$. Then we have the estimate

$$x(\tilde{p}_1) = \frac{\tilde{p}_1 (k_2 - k_1)^2}{2(k_1 k_2 - \tilde{p}_1^2)} > \frac{(k_2 - k_1)}{2}.$$

Substituting this lower bound for $x(\tilde{p}_1)$ into the second formula above, we find that $f_2(\tilde{p}_1) > k_2$, hence $(\tilde{p}_1,\, f_2(\tilde{p}_1))$ cannot be a minimizing pair.

Now suppose that $\sqrt{k_1 k_2} < \tilde{p}_1 < k_2$ and consider $f_3(\tilde{p}_1)$, given by the third formula. Then, we find

$$\frac{\partial x}{\partial \tilde{p}_1} = \frac{-(k_2 - k_1)^2 (1 + \frac{k_1 k_2}{\tilde{p}_1^2})}{2(\tilde{p}_1 - \frac{k_1 k_2}{\tilde{p}_1})^2} < 0, \quad \frac{\partial f_3}{\partial x} = \frac{x}{\sqrt{x^2 + k_1 k_2}} - 1 < 0.$$

Using the chain rule, we get that $\frac{\partial f_3}{\partial \tilde{p}_1} = \frac{\partial f_3}{\partial x} \cdot \frac{\partial x}{\partial \tilde{p}_1} > 0$. Therefore, we have the bound $\tilde{p}_2 = f_3(\tilde{p}_1) < f_3(k_2) = k_1$, and so this third equation cannot give minimizers of the min-max problem either.

The only valid equation remaining is the first one, $\tilde{p}_2 = f_1(\tilde{p}_1) = \frac{k_1 k_2}{\tilde{p}_1}$. Substituting for $\tilde{p}_2$ in the convergence factor, we are left with a function of k and $\tilde{p}_1$ only,

$$S(k,\, \tilde{p}_1) := R(k,\, \tilde{p}_1,\, k_1 k_2 / \tilde{p}_1),$$

and we wish to solve the problem

$$\min_{k_1 \le \tilde{p}_1 \le \sqrt{k_1 k_2}} \left(\max \left\{ S(k_1,\, \tilde{p}_1)\,,\, S(\sqrt{k_1 k_2},\, \tilde{p}_1) \right\} \right). \tag{3.26}$$

It remains to show that this problem is solved by equioscillation of the two values. Analyzing the partial derivative of each of those function values, we get

$$\begin{aligned} \operatorname{sign}\left(\tfrac{\partial S}{\partial \tilde{p}_1}(k_1,\tilde{p}_1)\right) &= \operatorname{sign}((\tilde{p}_1-k_1)(\tilde{p}_1-k_2)(\tilde{p}_1^2-k_1k_2)) &> 0, \\ \operatorname{sign}\left(\tfrac{\partial S}{\partial \tilde{p}_1}(\sqrt{k_1k_2},\tilde{p}_1)\right) &= \operatorname{sign}(\tilde{p}_1^2-k_1k_2) &< 0. \end{aligned}$$

Moreover, when inserting the extremal values for the parameter $\tilde{p}_1$, we get

$$\begin{aligned} &0 = S(k_1,k_1) < S(k_1,\sqrt{k_1k_2}), \\ &S(\sqrt{k_1k_2},k_1) > S(\sqrt{k_1k_2},\sqrt{k_1k_2}) = 0. \end{aligned}$$

Hence, the unique minimizer of problem (3.26) is given by the unique parameter $\tilde{p}_1^* \in (k_1,\sqrt{k_1k_2})$ such that $S(k_1,\tilde{p}_1^*) = S(\sqrt{k_1k_2},\tilde{p}_1^*)$. Solving this algebraic equation, we find that the unique minimizing pair $(\tilde{p}_1^*,\tilde{q}^*)$ for the transformed problem (3.4.4) is given by the direct formulas

$$\begin{aligned} \tilde{p}_1^* &= \frac{(k_1k_2)^{\frac{1}{4}}}{\sqrt{2}}\left[\sqrt{k_1+k_2}-(\sqrt{k_2}-\sqrt{k_1})\right], \\ \tilde{p}_2^* &= \frac{(k_1k_2)^{\frac{1}{4}}}{\sqrt{2}}\left[\sqrt{k_1+k_2}+(\sqrt{k_2}-\sqrt{k_1})\right]. \end{aligned}$$

Transforming back to the original parameter space (p,q), we obtain the formulas (3.24). Note that equioscillation of the transformed convergence factor at the frequencies k_1, $\sqrt{\tilde{p}_1\tilde{p}_2}$, and k_2 is equivalent to the equioscillation of the original convergence factor at k_1, $\sqrt{\frac{p}{q}}$ and k_2. □

Theorem 3.11 (Asymptotic performance). *When ν_1 and ν_2 are fixed, $k_2 = \frac{\pi}{h}$, and h goes to 0, the optimized second order parameters are*

$$p^* = \frac{k_1^{\frac{3}{4}}\pi^{\frac{1}{4}}}{\sqrt{2}}h^{-\frac{1}{4}} + O(h^{\frac{3}{4}}), \quad q^* = \frac{1}{\sqrt{2}\pi^{\frac{3}{4}}k_1^{\frac{1}{4}}}h^{\frac{3}{4}} + O(h^{\frac{7}{4}}),$$

and the asymptotic convergence factor of the Schwarz method is

$$\max_{k_1\le k\le \pi/h}\rho(k,p^*,q^*) = 1-\sqrt{2}\left(2+\mu+\frac{1}{\mu}\right)\left(\frac{k_1}{\pi}\right)^{\frac{1}{4}}h^{\frac{1}{4}} + O(h^{\frac{1}{2}}).$$

Proof. These are obtained by directly expanding the formulas (3.24) for p^* and q^*, and subsequently the convergence factor $\rho(k_1, p^*, q^*)$ for small h, when $k_2 = \frac{\pi}{h}$. □

It is interesting to note that the formulas for the optimized parameters do not depend on the diffusion coefficients at all, only the resulting convergence factor does. In fact, the optimized parameters are the same as those obtained when the coefficient is continuous, i.e. when solving the problem $-\Delta u = f$ (see [32]). This is also true for the one-sided Robin conditions of Section 3.4.2. Now, when the coefficient is *continuous*, the asymptotic convergence factor of optimized second order conditions and optimized two-sided Robin conditions are comparable: both are of the form $1-O(h^{\frac{1}{4}})$. Moreover, the second order conditions actually give a better asymptotic constant in this expansion (see [32]), so in practice we would recommend using the second order conditions over the two-sided Robin conditions. However, here we showed that when the coefficient is discontinuous, the optimized two-sided Robin conditions yield a mesh independent asymptotic convergence factor, whereas the optimized second order conditions do not.

3.4.5 Comparison of the Convergence Factors

Figure 3.8 shows a comparison of the convergence factors for the four different choices of optimized transmission conditions analyzed in this chapter. For the mesh size $h = \frac{\pi}{200}$, which is not very small, the performance of the two-sided Robin conditions and second order conditions are comparable: the optimized convergence factors are less than 0.08 when $\mu = 10$, and less than 0.02 when $\mu = 50$, producing very rapid error reduction for both methods.

3.4.6 Asymptotics for Strong Heterogeneity

All the asymptotic convergence factors derived so far were always with respect to a small grid spacing h while the coefficients ν_1, ν_2 are held constant. It is also interesting and relevant to consider the case when the ratio of coefficients is very large (corresponding to strong heterogeneity in the material), and how this affects the

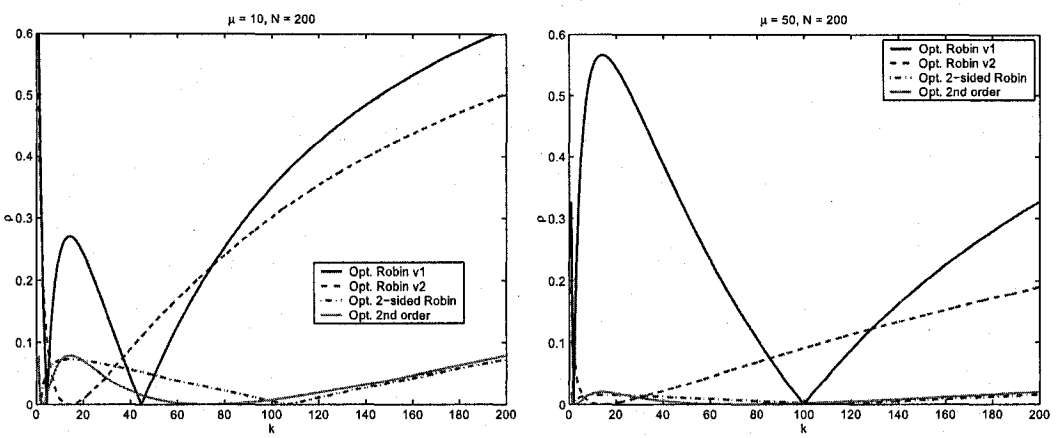

Figure 3.8: Convergence factors for $\mu = 10$ on the left and $\mu = 50$ on the right, in the case $h = \pi/200$.

asymptotics. For this purpose, suppose that $\mu := \max(\nu_1, \nu_2)/\min(\nu_1, \nu_2) \gg 1$ is large, and assume that h is small but fixed. For applications of flow in heterogeneous porous media, the coefficients can jump by several orders of magnitude across interfaces between subdomains.

Theorem 3.12 (Asymptotics for a large jump in diffusion). *For the different optimized conditions, we find different behaviors as $\mu \to \infty$, for fixed small h. The following expressions are obtained by first expanding the optimized convergence factor asymptotically for large μ, and then keeping only the dominant term when h is small.*

- *Optimized Robin conditions, version 1:*

$$\max_{k_1 \le k \le \pi/h} \rho(k, p^*) = \frac{\sqrt{k_2} - \sqrt{k_1}}{\sqrt{k_2} + \sqrt{k_1}} + O\left(\frac{1}{\mu}\right) \qquad \approx 1 - 2\sqrt{\frac{k_1 h}{\pi}}.$$

- *Optimized Robin conditions, version 2:*

$$\max_{k_1 \le k \le \pi/h} \rho(k, q^*) = \frac{\left(\sqrt{k_2} - \sqrt{k_1}\right)^2}{\sqrt{k_1 k_2}} \cdot \frac{1}{\mu} + O\left(\frac{1}{\mu^2}\right) \qquad \approx \sqrt{\frac{\pi}{k_1 h}} \left(\frac{1}{\mu}\right).$$

- *Optimized two-sided Robin conditions: let* $p_1^0 := \lim_{\mu\to\infty} p_1^*$ *and* $p_2^0 := \lim_{\mu\to\infty} p_2^*$

$$, \max_{k_1\le k\le \pi/h} \rho(k, p_1^*, p_2^*) = \frac{(p_1^0 - k_1)(p_2^0 - k_1)}{k_1 p_2^0} \cdot \frac{1}{\mu} + O\left(\frac{1}{\mu^2}\right) \quad \approx \frac{1}{\mu}.$$

- *Optimized second order conditions:*

$$\max_{k_1\le k\le \pi/h} \rho(k, p^*, q^*) = \frac{(p^* + q^* k_1^2 - k_1)^2}{k_1(p^* + q^* k_1^2)} \cdot \frac{1}{\mu} + O\left(\frac{1}{\mu^2}\right) \quad \approx \left(\frac{\pi}{4k_1 h}\right)^{\frac{1}{4}} \left(\frac{1}{\mu}\right).$$

Proof. For the second version of optimized Robin conditions and for optimized second order conditions, the formulas for the parameters are independent of μ, so it is straightforward to expand the convergence factor as $\mu \to \infty$.

For the first version of optimized Robin conditions, for μ large enough we will have $k_r < f(\mu)$, and thus fall in the case of two distinct minimizers (see Theorem 3.4). Solving the associated biquadratic, the two roots have the expansions

$$p_l^* = \sqrt{k_1 k_2} + (k_2 - k_1)^2 \mu^{-1} + O(\mu^{-2}),$$
$$p_r^* = \sqrt{k_1 k_2}\mu - \frac{(k_2 - k_1)^2}{2\sqrt{k_1 k_2}} + O(\mu^{-1}).$$

When expanding $\rho(k_1, p^*(\mu))$ for large μ, we get the same result using each of the two parameter values.

Finally, for the optimized two-sided Robin conditions, we can deduce that the optimized parameters p_1^* and p_2^* must be constant with respect to μ, to leading order, since both parameters lie in the interval $[k_1, k_2]$. In fact, by using the leading order term in the polynomial (3.21), we find

$$p_1^0 = \lim_{\mu\to\infty} p_1^* = \frac{1}{4}\left(\sqrt{k_1} + \sqrt{k_2}\right) - \frac{1}{4}\sqrt{\left(\sqrt{k_1} + \sqrt{k_2}\right)^4 - 16k_1k_2}\,,$$
$$p_2^0 = \lim_{\mu\to\infty} p_2^* = \frac{1}{4}\left(\sqrt{k_1} + \sqrt{k_2}\right) + \frac{1}{4}\sqrt{\left(\sqrt{k_1} + \sqrt{k_2}\right)^4 - 16k_1k_2}\,.$$

The result is then obtained by expanding $\rho(k_1, p_1^0, p_2^0)$. □

The results of this theorem also allow us to classify the performance of the four methods when μ is very large. For the first version of optimized Robin conditions,

the convergence factor doesn't seem to depend on μ asymptotically: it approaches a constant close to 1. For the other three optimized methods that we analyzed, the convergence improves asymptotically for large μ, it behaves like $\rho \approx C(h)\frac{1}{\mu}$. The two-sided Robin conditions have the best constant in front of the leading order term, with $C(h) = 1$, whereas the one-sided conditions (Robin and second order) have an asymptotic constant $C(h)$ that grows (i.e. gets worse) for small values of h.

The fact that the convergence *improves* when increasing the jump in the coefficients can be counter-intuitive: an a priori harder problem (with discontinuous coefficients) is actually easier to solve (we get faster convergence) than the corresponding problem with continuous coefficients.

Now, we can also look at asymptotics when both the jump in the diffusion coefficient is large and the mesh size is small *simultaneously*. This is relevant, for example, when there are boundary layers to resolve in the solution, in which case we might be forced to choose the mesh size h as a function of the coefficient ratio, for example we may have a restriction of the form $h \leq C\mu^{-\alpha}$. We show only a specific example of such a combined asymptotic expansion. Suppose that $\nu_2 = 1$ is kept constant and that $\nu_1 = \varepsilon$ with $\varepsilon \ll 1$. For the second version of optimized Robin conditions, the following proposition describes precisely the transition between the regime when μ is large (in which case ρ is small) and the regime when h is small (in which case $\rho \to 1$).

Proposition 3.1 (Asymptotics in h and ε). *For the second version of optimized Robin conditions, we get three separate cases.*

- *If* $\varepsilon h^{-1/2} \to 0$ *(i.e. h goes to 0 slower than ε^2),*

$$\max_{k_1 \leq k \leq \pi/h} |\rho(k, q^*)| \approx \sqrt{\frac{\pi}{k_1}} \varepsilon h^{-1/2}.$$

- *If* $\varepsilon h^{-1/2} \approx C$,

$$\max_{k_1 \leq k \leq \pi/h} |\rho(k, q^*)| = \frac{1}{1 + C\sqrt{\frac{k_1}{\pi}}} - O(h^{1/2}).$$

- *If* $\varepsilon h^{-1/2} \to \infty$ *(i.e. h goes to 0 faster than* ε^2*),*

$$\max_{k_1 \le k \le \pi/h} |\rho(k, q^*)| = 1 - \sqrt{\frac{k_1}{\pi}} \varepsilon^{-1} h^{1/2} + O(h).$$

3.5 The Dirichlet-Neumann Method

Let us briefly look at the Dirichlet-Neumann iterative method, and analyze it in the same framework as for the optimized Schwarz methods. In this section, we will consider a diffusion-reaction equation with discontinous coeffcients as a model problem,

$$\mathcal{L}(u) := -\nabla \cdot (\nu(\mathbf{x}) \nabla u) + \eta(\mathbf{x}) u = 0 \quad \text{in } \Omega = \mathbb{R}^2,$$

where

$$\nu(\mathbf{x}) = \begin{cases} \nu_1 & \text{for } \mathbf{x} \in \Omega_1, \\ \nu_2 & \text{for } \mathbf{x} \in \Omega_2, \end{cases} \qquad \eta(\mathbf{x}) = \begin{cases} \eta_1 & \text{for } \mathbf{x} \in \Omega_1, \\ \eta_2 & \text{for } \mathbf{x} \in \Omega_2, \end{cases}$$

with the subdomains $\Omega_1 = (-\infty, 0) \times \mathbb{R}$ and $\Omega_2 = (0, \infty) \times \mathbb{R}$.

A first version of the Dirichlet-Neumann method, that we will denote (DN), can be written as

$$\left\{ \begin{array}{rcll} \mathcal{L}_1 u_1^{n+1} & = & 0 & \text{in } \Omega_1, \\ u_1^{n+1} & = & \mu^n & \text{at } x = 0, \end{array} \right\} \quad \left\{ \begin{array}{rcll} \mathcal{L}_2 u_2^{n+1} & = & 0 & \text{in } \Omega_2, \\ \nu_2 \frac{\partial u_2^{n+1}}{\partial x} & = & \nu_1 \frac{\partial u_1^{n+1}}{\partial x} & \text{at } x = 0, \end{array} \right\}$$

$$\mu^{n+1} = \theta u_2^{n+1} + (1 - \theta) \mu^n \quad \text{at } x = 0,$$

where θ is a positive relaxation parameter. A second version, to be denoted (ND), is obtained by imposing the Dirichlet condition in the subdomain Ω_2 instead

$$\left\{ \begin{array}{rcll} \mathcal{L}_1 u_1^{n+1} & = & 0 & \text{in } \Omega_1, \\ \nu_1 \frac{\partial u_1^{n+1}}{\partial x} & = & \nu_2 \frac{\partial u_2^{n+1}}{\partial x} & \text{at } x = 0, \end{array} \right\} \quad \left\{ \begin{array}{rcll} \mathcal{L}_2 u_2^{n+1} & = & 0 & \text{in } \Omega_2, \\ u_2^{n+1} & = & \mu^n & \text{at } x = 0, \end{array} \right\}$$

$$\mu^{n+1} = \theta u_1^{n+1} + (1 - \theta) \mu^n \quad \text{at } x = 0.$$

It is also possible to write Dirichlet-Neumann methods in which we relax on the Neumann data instead of the Dirichlet data, but it can be easily shown that we obtain the same convergence factors as for the algorithms (DN) and (ND). Note that we wrote *sequential* versions of the methods.

For this model problem, with the help of the Fourier transform, we find the following convergence factors for the two methods

$$\rho_{DN}(k,\theta) := \left|\frac{\hat{u}_1^{n+1}(0,k)}{\hat{u}_1^{n}(0,k)}\right| = \left|\frac{\hat{u}_2^{n+1}(0,k)}{\hat{u}_2^{n}(0,k)}\right| = \left|1-\theta\left[1+\frac{\nu_1\lambda^{(1)}(k)}{\nu_2\lambda^{(2)}(k)}\right]\right|,$$
$$\rho_{ND}(k,\theta) = \left|1-\theta\left[1+\frac{\nu_2\lambda^{(2)}(k)}{\nu_1\lambda^{(1)}(k)}\right]\right|,$$

where $\lambda^{(j)}$ are the characteristic roots of the ordinary differential equation in Fourier space, for each subdomain Ω_j, namely $\lambda^{(j)}(k) := \sqrt{k^2+\frac{\eta_j}{\nu_j}}$.

It is then natural to ask what is a good value for the relaxation parameter θ to use. We look for the optimized value by solving the min-max problem

$$\min_{\theta>0}\left(\max_{k_1\le k\le k_2}|\rho_{DN}(k,\theta)|\right), \tag{3.27}$$

and a similar problem for the method (ND). First define the functions

$$F(k) := \frac{\nu_1}{\nu_2}\sqrt{\frac{k^2+\frac{\eta_1}{\nu_1}}{k^2+\frac{\eta_2}{\nu_2}}}, \quad G(k) := \frac{\nu_2}{\nu_1}\sqrt{\frac{k^2+\frac{\eta_2}{\nu_2}}{k^2+\frac{\eta_1}{\nu_1}}}.$$

Theorem 3.13 (Optimized relaxation parameter). *The unique minimizers of problem* (3.27) *for the methods (DN) and (ND) are, respectively,*

$$\theta^*_{DN} = \frac{2}{2+F(k_1)+F(k_2)}, \qquad \theta^*_{ND} = \frac{2}{2+G(k_1)+G(k_2)}.$$

Proof. The functions $F(k)$ and $G(k)$ are monotonically increasing for all k or monotonically decreasing for all k, depending on the sign of $\frac{\eta_1}{\nu_1}-\frac{\eta_2}{\nu_2}$. So, the maximum of the convergence factor is found by looking at the frequencies $k=k_1$ and $k=k_2$. Also, the convergence factor, without the absolute values, is a linear function of the relaxation parameter θ. Thus, the best convergence factor is obtained by equioscillation $\rho(k_1,\theta^*)=\rho(k_2,\theta^*)$. For the method (DN), this leads to equation

$$1-\theta^*_{DN}\left[1+F(k_1)\right] = \theta^*_{DN}\left[1+F(k_2)\right]-1,$$

which gives the formula for the optimized relaxation parameter. We get θ^*_{ND} in the same way. □

Theorem 3.14 (Asymptotic performance). *If $\lambda := \frac{\nu_1}{\nu_2}$ is fixed, $k_2 = \frac{\pi}{h}$ and h goes to 0, the asymptotic convergence factor of the Dirichlet-Neumann methods with optimized relaxation parameter is*

- *for method (DN),*

$$\max_{k_1 \le k \le \pi/h} |\rho_{DN}(k, \theta^*_{DN})| = \left| \frac{\lambda - F(k_1)}{2 + \lambda + F(k_1)} \right| + O(h^2),$$

- *for method (ND)*

$$\max_{k_1 \le k \le \pi/h} |\rho_{ND}(k, \theta^*_{ND})| = \left| \frac{\frac{1}{\lambda} - G(k_1)}{2 + \frac{1}{\lambda} + G(k_1)} \right| + O(h^2).$$

Thus, the optimized Dirichlet-Neumann methods have mesh independent convergence: the convergence factor approaches a constant strictly less than 1 as $h \to 0$. Furthermore, when the coefficients are continuous, or when $\eta_1 = \eta_2 = 0$, the optimized relaxation parameters are actually *optimal*: the convergence factor is uniformly 0 (for both methods). However, as soon as we use more subdomains, the optimized relaxation parameters are hard to find; it is not known how to compute them in general. Also, the symmetry of the domain decomposition is essential for the optimality of Dirichlet-Neumann methods.

When using the Dirichlet-Neumann method without relaxation ($\theta = 1$), one version of the method converges and the other one doesn't, when $\eta_j = 0$, since the convergence factors become respectively

$$\rho_{DN}(k, 1) = \frac{\nu_1}{\nu_2}, \qquad \rho_{ND}(k, 1) = \frac{\nu_2}{\nu_1}.$$

In that situation, to guarantee convergence, we need to impose the Dirichlet transmission condition for the subproblem that has the smaller diffusion coefficient.

3.6 Numerical Experiments With Two Subdomains

The numerical experiments shown in this section are carried out on the problem

$$\begin{cases} -\nabla \cdot (\nu(\mathbf{x})\nabla u) &= 1 \quad \text{in } \Omega = (0,\pi) \times (0,\pi), \\ u &= 0 \quad \text{on } \partial\Omega. \end{cases} \tag{3.28}$$

The square domain is decomposed into two non-overlapping subdomains

$$\Omega_1 = \left(0, \frac{\pi}{2}\right) \times (0,\pi), \quad \Omega_2 = \left(\frac{\pi}{2}, \pi\right) \times (0,\pi).$$

We use a finite volume discretization on a uniform grid with grid size h. In all the experiments, $\nu_1 = \frac{1}{\mu}$ and $\nu_2 = 1$ where μ is taken to be larger than 1. Vectors of random values between -1 and 1 are fed as initial conditions for the Schwarz iteration so that the initial error includes all possible frequency components. When using the Schwarz method as an iterative solver, exactly as in (3.7), we use the ℓ^∞ error

$$e^n := \|\mathbf{U} - \mathbf{u}^n\|_\infty,$$

where $\mathbf{U}$ is the discrete solution of the global problem (3.28) and $\mathbf{u}^n$ is obtained by gluing together the subdomain solutions at iteration n, taking an average of $\mathbf{u}_1^n$ and $\mathbf{u}_2^n$ on the interface $x = \frac{\pi}{2}$.

3.6.1 Krylov Acceleration

The convergence can be accelerated by using the Schwarz method as a preconditioner for a Krylov subspace method. There are several ways to formulate this; the standard method is to write the Schwarz method as a stationary iterative method for a linear system condensed on the interface. In this subsection, we present a different way which is particularly convenient for implementation.

Let us write one iteration of the Schwarz method as

$$\mathbf{u}^{n+1} = \Phi(\mathbf{u}^n, \mathbf{f}),$$

where $\mathbf{u}^n$ is an augmented vector in which the unknowns on the interface appear more than once. In general for M subdomains,

$$\mathbf{u}^n = \begin{bmatrix} \mathbf{u}_1^n & \mathbf{u}_2^n & \cdots & \mathbf{u}_M^n \end{bmatrix}^T,$$

and $\mathbf{u}_j^n$ is a column vector listing the unknowns in $\overline{\Omega}_j$ under a given ordering. The function $\mathbf{\Phi}$ can be decomposed as

$$\mathbf{\Phi}(\mathbf{u}, \mathbf{f}) = \mathbf{\Phi}(\mathbf{u}, \mathbf{0}) + \mathbf{\Phi}(\mathbf{0}, \mathbf{f}),$$

and note that $\mathbf{\Phi}(\cdot, \mathbf{0})$ and $\mathbf{\Phi}(\mathbf{0}, \cdot)$ are both linear functions. Hence, we can write

$$\mathbf{\Phi}(\mathbf{u}, \mathbf{f}) = M\mathbf{u} + N\mathbf{f},$$

where the matrices M and N are defined by the linear applications

$$M\mathbf{u} := \mathbf{\Phi}(\mathbf{u}, \mathbf{0}), \quad N\mathbf{g} := \mathbf{\Phi}(\mathbf{0}, \mathbf{g}),$$

which are simple applications of a Schwarz iteration with zero right-hand side and zero initial guess respectively. The Schwarz iteration can then be formulated as

$$\mathbf{u}^{n+1} = \mathbf{u}^n + (N\mathbf{f} - (I - M)\mathbf{u}^n), \tag{3.29}$$

which is a simple stationary iterative method for the linear system

$$(I - M)\mathbf{u} = N\mathbf{f}. \tag{3.30}$$

To accelerate the convergence, we can use a Krylov subspace method to solve the non-symmetric linear system (3.30) instead. We use BiCGstab [66] instead of GMRES [61] in our experiments, so that we don't have to choose an appropriate restart parameter to avoid memory problems. Note that the stopping criterion in this case is based on the Euclidean norm of the relative residual in the system (3.30), namely

$$r^n = \frac{\|N\mathbf{f} - (I - M)\mathbf{u}^n\|_2}{\|N\mathbf{f}\|_2}.$$

It may be possible to design different, possibly more meaningful, stopping criteria, for example using backward error analysis, as in [3]; this could be the subject of future research.

Note that the iteration (3.29) can also be viewed, equivalently, as a *preconditioned* Richardson iteration, with preconditioner N, for the linear system

$$N^{-1}(I - M)\mathbf{u} = \mathbf{f}. \tag{3.31}$$

So one could use a preconditioned Krylov method to solve the linear system (3.31), with preconditioner N (this gives the same iterates as the method described above). The difference, in practice, is that this would require computing the action of the matrix $N^{-1}(I - M)$, which is not easily obtained from the function $\mathbf{\Phi}$ alone.

3.6.2 Comparisons

In Figure 3.9, the convergence of the various methods is shown for two different coefficient ratios. The two Dirichlet-Neumann methods, (DN) and (ND), with optimized relaxation parameters, are converging very quickly. This is because they are optimal methods when using two symmetric subdomains (see Section 3.5); however their convergence deteriorates very quickly when using more subdomains or when breaking the symmetry (see Section 4.3). The transmission conditions optimized over two free parameters perform much better than the one-sided Robin conditions, as expected. Also, the second version of the one-sided conditions converges faster than the first version with bad scaling, and the difference in performance increases as the ratio of coefficients increases, as predicted. The optimized Schwarz methods with properly scaled transmission conditions converge faster as the jump in the coefficient is increased.

Next, we fix $\mu = 10$ and vary the grid size h. Table 3.1 shows the number of iterations that were needed to reach a tolerance of 10^{-6} for the different optimized methods. The convergence of the optimized transmission conditions deteriorates as we decrease the mesh size h, except for the two-sided Robin conditions, for which the number of iterations appears to stop growing for small h, as the theoretical asymptotics of Section 3.4 indicated. It is very difficult to verify the asymptotic convergence factors more precisely, as it was done in [32], to confirm the exponent of the leading order terms in h. When the coefficient is discontinuous ($\mu > 1$), the expansions we

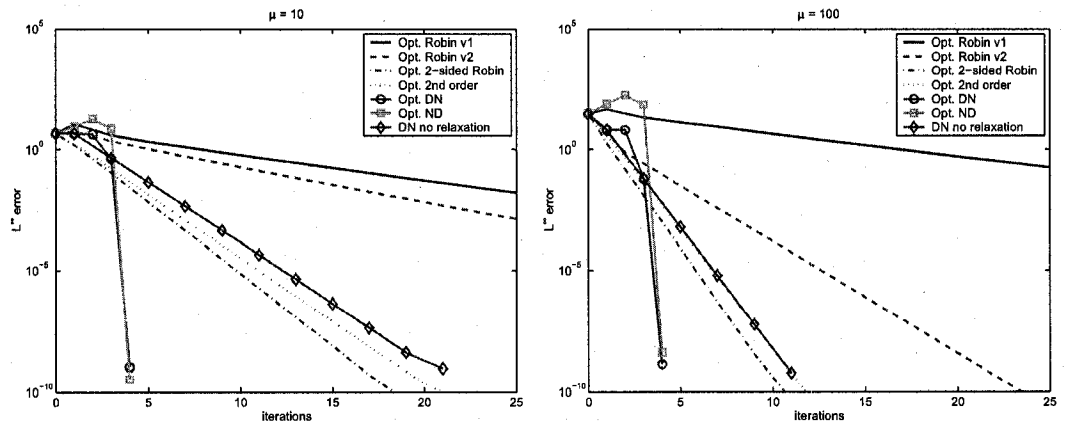

Figure 3.9: Convergence for $\mu = 10$ on the left and $\mu = 100$ on the right, in the case $n = 300$.

have derived become valid only for very small values of h, much smaller values than the ones we used in Table 3.1. To get some quantitative idea of how small h needs to be, let us look at a specific case. Suppose $k_1 = 1$ and $h = \pi/n$ for simplicity. For the optimized second order conditions, the asymptotic convergence factor is given by Theorem 3.11. For the expansion to be valid, h has to be at least small enough so that the second term in the expansion ($O(h^{1/4})$) is less than 1. This implies that

$$1 > \sqrt{2}\left(2 + \mu + \frac{1}{\mu}\right)\left(\frac{1}{n}\right)^{\frac{1}{4}},$$

which reduces to the condition that $n > 4(2+\mu+1/\mu)^4$. So, if $\mu = 10$ for example, we need to use *at least* $n = 85744$ grid points in y to observe the theoretical asymptotic convergence factor!

Now, let us fix $h = \frac{\pi}{300}$ and vary the heterogeneity ratio μ. Table 3.2 again shows the number of iterations needed to reach a tolerance of 10^{-6} for the different methods. These results show that the first version of the optimized Robin conditions is really not performing well for large discontinuities in the coefficient. On the other hand, all the other optimized conditions do show significant improvements in the convergence as we increase the ratio μ. This is also in agreement with the theoretical asymptotics

	Opt. Robin v.1	Opt. Robin v.2	Opt. 2-sided Robin	Opt. 2nd order
h	Optimized Schwarz as an iterative solver			
$\frac{\pi}{50}$	26	24	11	10
$\frac{\pi}{100}$	39	30	11	12
$\frac{\pi}{200}$	56	40	12	13
$\frac{\pi}{400}$	77	54	13	14
$\frac{\pi}{800}$	110	73	13	17
	Optimized Schwarz with Krylov acceleration			
$\frac{\pi}{50}$	19	13	9	7
$\frac{\pi}{100}$	21	14	9	8
$\frac{\pi}{200}$	25	17	10	10
$\frac{\pi}{400}$	28	21	10	10
$\frac{\pi}{800}$	37	26	11	11

Table 3.1: Number of iterations to reach a tolerance of 10^{-6}, for small values of h.

presented in Section 3.4.6. It may be noted, in Table 3.2, that the convergence of the first version of the optimized Robin conditions seems to show some improvement for large μ, when using Krylov acceleration. However, this is only due to the fact that the norm used to check convergence depends on the transmission conditions and its parameters. By looking at the convergence history more closely in Figure 3.10, we can see that the rate of convergence does not get better as we increase μ.

Recall that the optimized parameters for the transmission conditions are computed by analyzing a convergence factor that was derived for a *continuous model problem* on the infinite plane. We compare these optimized values with the parameters that yield the fastest convergence numerically for the discrete problem, in the case $\mu = 2$ and $h = \frac{\pi}{50}$. Figures 3.11 and 3.12 show the number iterations required to reach a tolerance of 10^{-6} for a range of parameter values around the optimized parameters. We see that the Fourier analysis performed on the continuous model problem predicts the best parameters very well. We should point out that there are extreme situations for which the continuous analysis will fail to produce very good parameter values, for

	Opt. Robin v.1	Opt. Robin v.2	Opt. 2-sided Robin	Opt. 2nd order
μ	Optimized Schwarz as an iterative solver			
10^1	67	48	13	14
10^2	79	16	7	8
10^3	206	9	5	6
10^4	234	7	5	5
10^5	234	5	5	5
	Optimized Schwarz with Krylov acceleration			
10^1	29	17	10	9
10^2	29	8	6	6
10^3	59	5	4	4
10^4	50	4	4	4
10^5	41	4	4	3

Table 3.2: Number of iterations to reach a tolerance of 10^{-6}, for large heterogeneity ratios μ.

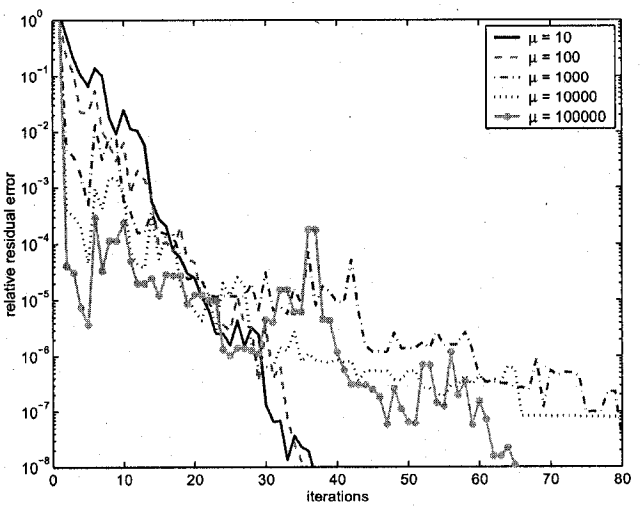

Figure 3.10: Convergence of the Schwarz method with optimized Robin conditions, version 1, when used as a preconditioner for BiCGstab, for the case $h = \frac{\pi}{300}$.

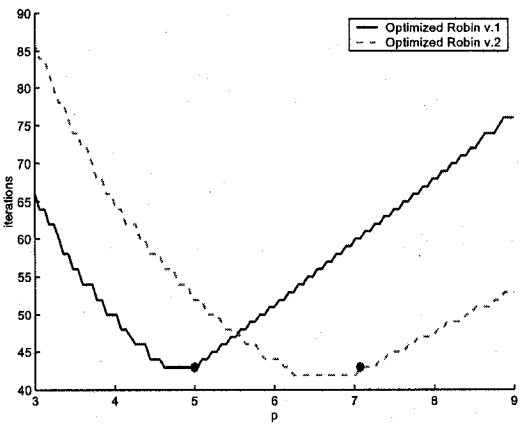

Figure 3.11: Optimized parameters (*) computed from the min-max problem (when $\mu = 2$), compared with the performance of other values of the parameters, for the two versions of one-sided Robin conditions.

example

- when the subdomains are very thin or of very different width,
- when the coefficients of the problem are varying a lot inside the subdomains,
- when the shape of the interfaces is very far from straight.

In such cases, one should solve modified min-max problems that are better adapted to these situations instead. Some of these situations can occur naturally for example when using many subdomains; we investigate this in more detail in Chapter 4.

3.7 Generalizations

In this section, we discuss several generalizations of the results we have presented. First, we show that, when considering a diffusion-reaction problem with discontinuous coefficients, we find similar asymptotic convergence factors for the optimized Schwarz methods, and we provide asymptotic formulas for the optimized parameters. Afterward, we prove that all the results can be easily extended to the case of a general

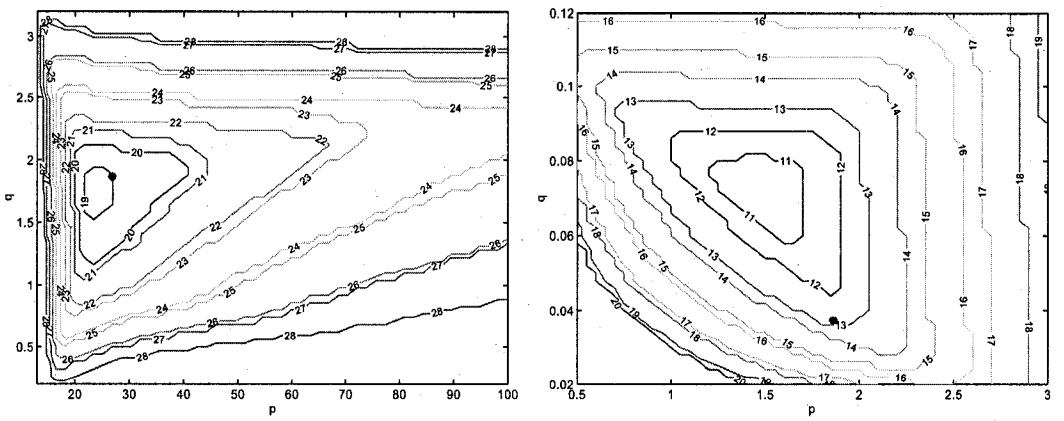

Figure 3.12: Optimized parameters (*) computed from the min-max problem (when $\mu = 2$), compared with the performance of other values of the parameters, on the left for the two-sided Robin conditions, and on the right for the second order conditions.

anisotropic diffusion. Finally, we note that the same min-max problems need to be solved for three dimensional problems.

3.7.1 Asymptotic Formulas for a Diffusion-Reaction Problem

Consider now the diffusion-reaction model problem

$$\begin{cases} -\nabla \cdot (\nu(\mathbf{x})\nabla u) + \eta(\mathbf{x})u &= 0 \quad \text{for } \mathbf{x} \in \Omega = \mathbb{R}^2, \\ |u| &< \infty \quad \text{as } \mathbf{x} \to \infty. \end{cases}$$

As before, the plane is divided into two non-overlapping subdomains,

$$\Omega_1 = (-\infty, 0) \times \mathbb{R}, \quad \Omega_2 = (0, \infty) \times \mathbb{R}.$$

Suppose the coefficients are piecewise constant functions,

$$\nu(\mathbf{x}) = \begin{cases} \nu_1 & \text{for } \mathbf{x} \in \Omega_1, \\ \nu_2 & \text{for } \mathbf{x} \in \Omega_2, \end{cases} \qquad \eta(\mathbf{x}) = \begin{cases} \eta_1 & \text{for } \mathbf{x} \in \Omega_1, \\ \eta_2 & \text{for } \mathbf{x} \in \Omega_2. \end{cases}$$

We wish to analyze the performance of the Schwarz iteration

$$\begin{cases} -\nu_1 \Delta u_1^n + \eta_1 u_1^n &= 0 & \text{in } \Omega_1, \\ (\nu_1 \partial_x + \mathcal{S}_1) u_1^n(0,y) &= (\nu_2 \partial_x + \mathcal{S}_1) u_2^{n-1}(0,y) & \text{for } y \in \mathbb{R}, \\ -\nu_2 \Delta u_2^n + \eta_2 u_2^n &= 0 & \text{in } \Omega_2, \\ (\nu_2 \partial_x - \mathcal{S}_2) u_2^n(0,y) &= (\nu_1 \partial_x - \mathcal{S}_2) u_1^{n-1}(0,y) & \text{for } y \in \mathbb{R}. \end{cases} \tag{3.32}$$

We adopt the same notation as before for the Fourier symbols corresponding to the operators $\mathcal{S}_j$, namely

$$\mathcal{F}_y(\mathcal{S}_j(u)) = \sigma_j(k)\hat{u}(x,k).$$

Let $c_i := \eta_i/\nu_i$, $\lambda := \nu_1/\nu_2$, and suppose that $\lambda > 1$ for simplicity. Exactly as in Section 3.3.1, we can use the Fourier transform to compute the convergence factor of the Schwarz iteration

$$\rho(k,\sigma_i) = \left| \frac{\hat{u}_j^{n+1}(0,k)}{\hat{u}_j^{n-1}(0,k)} \right| = \left| \frac{\sigma_1(k) - \nu_2\sqrt{k^2+c_2}}{\sigma_1(k) + \nu_1\sqrt{k^2+c_1}} \cdot \frac{\sigma_2(k) - \nu_1\sqrt{k^2+c_1}}{\sigma_2(k) + \nu_2\sqrt{k^2+c_2}} \right|.$$

In this case, we will not analytically solve the min-max problems associated with optimized transmission conditions; the level of complexity is significantly higher than for the simple diffusion problem. Instead, we will derive asymptotic formulas for the optimized parameters that are valid for small mesh sizes h, by using the same procedure as in Chapter 2 for the advection-diffusion equation.

Optimized Robin Conditions

Let us start with the one-sided Robin transmission conditions,

$$\sigma_1(k) = \nu_2 p, \quad \sigma_2(k) = \nu_1 p.$$

The optimized parameter value p^* is obtained by solving the min-max problem

$$\min_{p>0} \left(\max_{k_1 \le k \le \frac{\pi}{h}} \left| \frac{p - \sqrt{k^2+c_2}}{p + \lambda\sqrt{k^2+c_1}} \cdot \frac{p - \sqrt{k^2+c_1}}{p + \frac{1}{\lambda}\sqrt{k^2+c_2}} \right| \right). \tag{3.33}$$

We have solved this optimization problem numerically, using the Nelder-Mead algorithm [57], for various coefficient values and mesh sizes h (see Appendix A.3 for the

Matlab code). We observe that the optimized parameter seems to grow asymptotically like $h^{-\frac{1}{2}}$, and that the minimized convergence factor satisfy an equioscillation property at $k = k_1$ and $k = k_2$. This is consistent with the results of Section 3.4.2.

Thus, we make the educated guess that for small h, $p^* = C_p h^{-\frac{1}{2}}$ and that the equioscillation property holds. From the leading order term of the equation

$$\rho(k_1, p^*) = \rho(k_2, p^*),$$

we get a formula for the coefficient C_p, namely

$$p^* = C_p h^{-\frac{1}{2}} = \left(\frac{\pi(\lambda\sqrt{k_1^2 + c_1} + \sqrt{k_1^2 + c_2})}{\lambda + 1} \right)^{\frac{1}{2}} h^{-\frac{1}{2}}. \tag{3.34}$$

Then, expanding the convergence factor at $k = k_1$, we get the asymptotic convergence factor

$$\max_{k_1 \le k \le \frac{\pi}{h}} \rho(k, p^*) \sim 1 - \frac{(\lambda+1)^{\frac{3}{2}}}{\lambda} \left(\frac{\lambda\sqrt{k_1^2 + c_1} + \sqrt{k_1^2 + c_2}}{\pi} \right)^{\frac{1}{2}} h^{\frac{1}{2}}.$$

In Figure 3.13, we compare these asymptotic formulas we have derived with the optimized parameter and optimized convergence factor that are obtained numerically by solving (3.33) using the Nelder-Mead algorithm. Notice the especially good agreement of the formula for p^*. In fact, we conjecture that the asymptotic formula (3.34) is actually *exact*, i.e. that it gives exactly the best parameter value solving the min-max problem (3.33). Again, please be warned that in these plots, and in the ones to follow, the scale of the y-axis changes significantly from figure to figure, affecting the comparison.

Optimized Two-Sided Robin Conditions

Now let us consider two-sided Robin transmission conditions,

$$\sigma_1(k) = \nu_2 p_1, \quad \sigma_2(k) = \nu_1 p_2.$$

The optimized parameter values (p_1^*, p_2^*) are obtained via the solution of the min-max problem

$$\min_{p_1, p_2 > 0} \left(\max_{k_1 \le k \le \frac{\pi}{h}} \left| \frac{p_1 - \sqrt{k^2 + c_2}}{p_1 + \lambda\sqrt{k^2 + c_1}} \cdot \frac{p_2 - \sqrt{k^2 + c_1}}{p_2 + \frac{1}{\lambda}\sqrt{k^2 + c_2}} \right| \right). \tag{3.35}$$

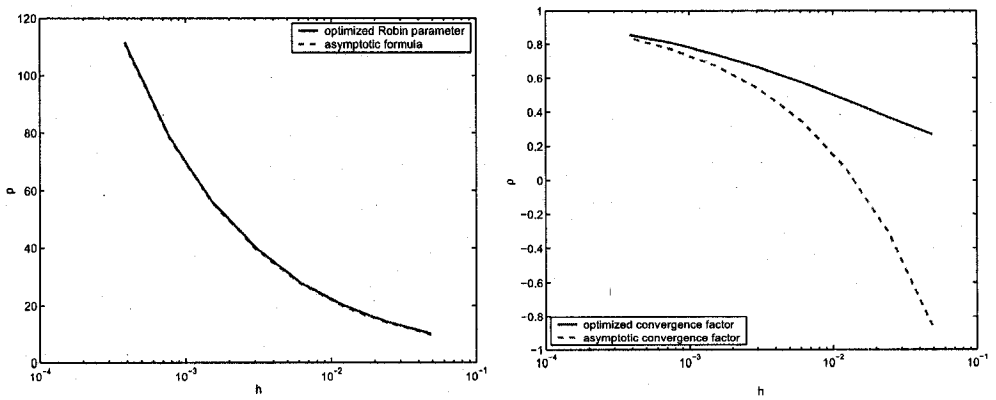

Figure 3.13: Comparison of the optimized Robin parameter (left) and optimized convergence factor (right) with the asymptotic formulas obtained for small h, in the case $\nu_1 = \eta_1 = 1$, $\nu_2 = 0.1$, $\eta_2 = 0.5$.

By solving this min-max problem (3.35) *numerically* for various coefficient values and various mesh sizes h, we make a few observations.

- $\rho(k, p_1^*, p_2^*)$ has one interior local maximum at k_c, where k_c grows asymptotically like $h^{-\frac{1}{2}}$.
- The optimized parameter p_1^* does not seem to grow or decay significantly as $h \to 0$, and p_2^* grows quickly like h^{-1}.
- The optimized convergence factor satisfies an equioscillation property at the frequencies k_1, k_c and k_2.

Hence, from these observations, which agree with our analysis of Section 3.4.2, we make the educated assumptions that, for small h,

$$k_c = C_c h^{-\frac{1}{2}}, \quad p_1^* = C_1, \quad p_2^* = C_2 h^{-1},$$

$$\rho(k_1, p_1^*, p_2^*) = \rho(k_c, p_1^*, p_2^*) = \rho(k_2, p_1^*, p_2^*).$$

By looking at the leading order term of the different equations that should be satisfied, we derive relations between the unknown coefficients C_c, C_1, C_2.

- k_c is a local maximum of the convergence factor: $\frac{\partial \rho}{\partial k}(k_c, p_1^*, p_2^*) = 0$

$$\Rightarrow \quad C_c = \sqrt{C_1 C_2}.$$

- Equioscillation property at k_1 and k_2: $\rho(k_1, p_1^*, p_2^*) = \rho(\pi/h, p_1^*, p_2^*)$

$$\Rightarrow \quad (\lambda+1)C_1C_2 + \lambda C_2\left(\sqrt{k_1^2+c_1} - \sqrt{k_1^2+c_2}\right) = \lambda\sqrt{k_1^2+c_1} + \sqrt{k_1^2+c_2}.$$

- Equioscillation property at k_1 and k_c: $\rho(k_1, p_1^*, p_2^*) = \rho(k_c, p_1^*, p_2^*)$

$$\Rightarrow \quad (\lambda-1)C_1 = \lambda(\sqrt{k_1^2+c_1} + \sqrt{k_1^2+c_2}).$$

Solving these three equations for the three coefficients C_1, C_2 and C_c, we get

$$\begin{aligned}
k_c = C_c h^{-\frac{1}{2}} &= \left(\frac{\pi}{2}(\sqrt{k_1^2+c_1} + \sqrt{k_1^2+c_2})\right)^{\frac{1}{2}} h^{-\frac{1}{2}}, \\
p_1^* = C_1 &= \frac{\lambda}{\lambda-1}\left(\sqrt{k_1^2+c_1} + \sqrt{k_1^2+c_2}\right), \\
p_2^* = C_2 h^{-1} &= \frac{(\lambda-1)\pi}{2\lambda} h^{-1}.
\end{aligned}$$

Finally, expanding the $\rho(k_1, p_1^*, p_2^*)$ for small h we obtain the asymptotic convergence factor

$$\max_{k_1 \le k \le \frac{\pi}{h}} \rho(k, p_1^*, p_2^*) = \frac{1}{\lambda} + O(h^{\frac{1}{2}}).$$

Figure 3.14 shows how these asymptotic formulas compare to the optimized parameters obtained through the numerical solution of the min-max problem (3.35). Note the different scales for the parameters p_1 and p_2; the asymptotic formula for p_1^* is not too far off.

Optimized Second Order Conditions

Finally, we also look at second order transmission conditions

$$\sigma_1(k) = \nu_2(p + qk^2), \quad \sigma_2(k) = \nu_1(p + qk^2).$$

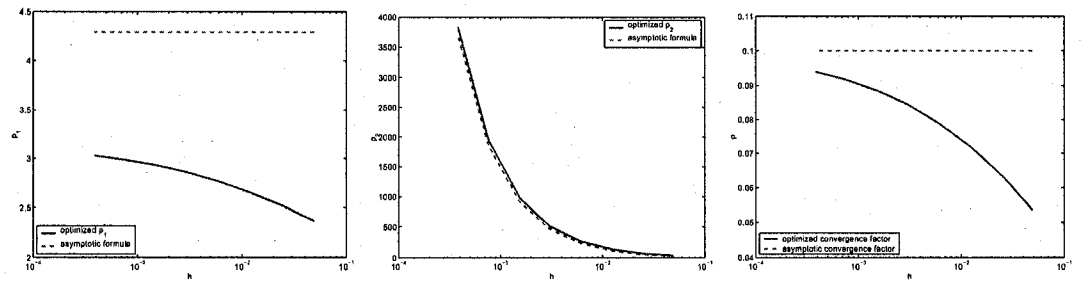

Figure 3.14: Comparison of the optimized two-sided Robin parameters (left and center) and associated optimized convergence factor (right) with the asymptotic formulas obtained for small h, in the case $\nu_1 = \eta_1 = 1$, $\nu_2 = 0.1$, $\eta_2 = 0.5$.

The optimized parameter values (p^*, q^*) are obtained by solving the min-max problem

$$\min_{p,q>0} \left(\max_{k_1 \leq k \leq \frac{\pi}{h}} \left| \frac{p + qk^2 - \sqrt{k^2 + c_2}}{p + qk^2 + \lambda\sqrt{k^2 + c_1}} \cdot \frac{p + qk^2 - \sqrt{k^2 + c_1}}{p + qk^2 + \frac{1}{\lambda}\sqrt{k^2 + c_2}} \right| \right). \tag{3.36}$$

By solving the min-max problem (3.36) numerically for various coefficient values and various mesh sizes h, we can make similar observations once again:

- $\rho(k, p^*, q^*)$ has one interior local maximum at k_c where k_c grows like $h^{-\frac{1}{2}}$.
- The optimized parameter p^* grows like $h^{-\frac{1}{4}}$ and q^* decays like $h^{\frac{3}{4}}$ asymptotically as $h \to 0$.
- The optimized convergence factor satisfies an equioscillation property at frequencies k_1, k_c and k_2.

These observations are in agreement with our analysis of Section 3.4.4, so we make the educated assumptions that, for small h,

$$k_c = C_c h^{-\frac{1}{2}}, \quad p^* = C_p h^{-\frac{1}{4}}, \quad q^* = C_q h^{\frac{3}{4}},$$

$$\rho(k_1, p^*, q^*) = \rho(k_c, p^*, q^*) = \rho(k_2, p^*, q^*).$$

By looking at the leading order term of the different equations that should be satisfied, we derive relations between the unknown coefficients.

- k_c is a local maximum of the convergence factor: $\frac{\partial \rho}{\partial k}(k_c, p^*, q^*) = 0$

$$\Rightarrow \quad C_c = \sqrt{\frac{C_p}{C_q}}.$$

- Equioscillation property at k_1 and k_2: $\rho(k_1, p^*, q^*) = \rho(\pi/h, p^*, q^*)$

$$\Rightarrow \quad (\lambda+1)C_p = \pi C_q \left(\lambda\sqrt{k_1^2+c_1} + \sqrt{k_1^2+c_2}\right).$$

- Equioscillation property at k_1 and k_c: $\rho(k_1, p^*, q^*) = \rho(k_c, p^*, q^*)$

$$\Rightarrow \quad \left(\lambda\sqrt{k_1^2+c_1} + \sqrt{k_1^2+c_2}\right) C_c = (1+\lambda)(C_p^2 + C_p C_q C_c^2).$$

Solving these three equations for the three coefficients C_p, C_q and C_c, we get

$$k_c = C_c h^{-\frac{1}{2}} = \left[\frac{\pi}{\lambda+1}\left(\lambda\sqrt{k_1^2+c_1} + \sqrt{k_1^2+c_2}\right)\right]^{\frac{1}{2}} h^{-\frac{1}{2}},$$

$$p^* = C_p h^{-\frac{1}{4}} = \left[\frac{\pi(\lambda\sqrt{k_1^2+c_1} + \sqrt{k_1^2+c_2})^3}{4(\lambda+1)^3}\right]^{\frac{1}{4}} h^{-\frac{1}{4}},$$

$$q^* = C_q h^{\frac{3}{4}} = \left[\frac{\lambda+1}{4\pi^3(\lambda\sqrt{k_1^2+c_1} + \sqrt{k_1^2+c_2})}\right]^{\frac{1}{4}} h^{\frac{3}{4}}.$$

Finally, expanding the $\rho(k_1, p^*, q^*)$ for small h we obtain the asymptotic convergence factor

$$\max_{k_1 \le k \le \frac{\pi}{h}} \rho(k, p^*, q^*) \approx 1 - \frac{\sqrt{2}(1+\lambda)^{\frac{7}{4}}}{\lambda}\left(\frac{\sqrt{k_1^2+c_1} + \sqrt{k_1^2+c_2}}{\pi}\right)^{\frac{1}{4}} h^{\frac{1}{4}}.$$

Figure 3.15 shows how these asymptotic formulas compare to the optimized parameters obtained through the numerical solution of the min-max problem (3.36). The asymptotic formulas give very good approximations for the optimized parameters, even if the asymptotic convergence factor seems to set in only for much smaller values of h (see Section 3.6.2 for a justification).

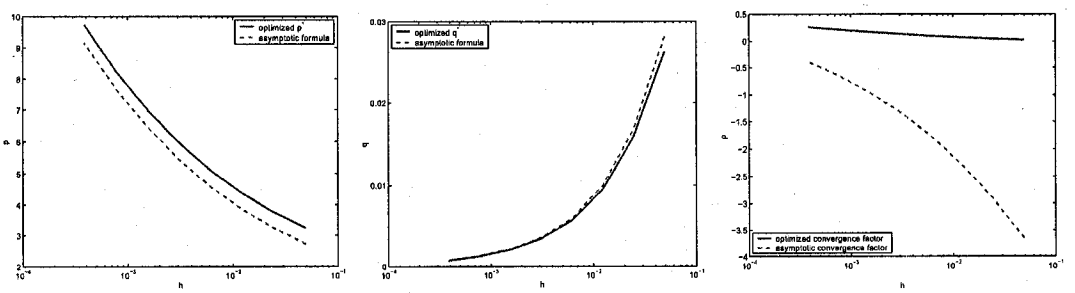

Figure 3.15: Comparison of the optimized 2nd order parameters (left and center) and associated optimized convergence factor (right) with the asymptotic formulas obtained for small h, in the case $\nu_1 = \eta_1 = 1$, $\nu_2 = 0.1$, $\eta_2 = 0.5$.

Numerical Comparison

In Figure 3.16, for a specific choice of coefficients, we show the convergence factor for the optimized transmissions conditions (found by solving the min-max problem numerically) when varying the mesh size h. We also included the convergence factor for 0^{th} and 2^{nd} order Taylor approximations of the optimal symbols, namely when using

$$\sigma_1^{(0)}(k) = \sqrt{\nu_2\eta_2}, \quad \sigma_2^{(0)}(k) = \sqrt{\nu_1\eta_1},$$

$$\sigma_1^{(2)}(k) = \sqrt{\nu_2\eta_2} + \nu_2\sqrt{\frac{\nu_2}{\eta_2}}k^2, \quad \sigma_2^{(2)}(k) = \sqrt{\nu_1\eta_1} + \nu_1\sqrt{\frac{\nu_1}{\eta_1}}k^2.$$

In particular, we see that the performance of the optimized two-sided Robin conditions appears to be mesh independent, and that $\rho \approx \frac{1}{\lambda} = 0.1$ in this case.

Refer to Appendix B.2 for some simple numerical experiments using the optimized transmission conditions and the asymptotic formulas we derived in this section, for two non-overlapping subdomains. These experiments also illustrate the asymptotic convergence properties of optimized Schwarz methods in this case.

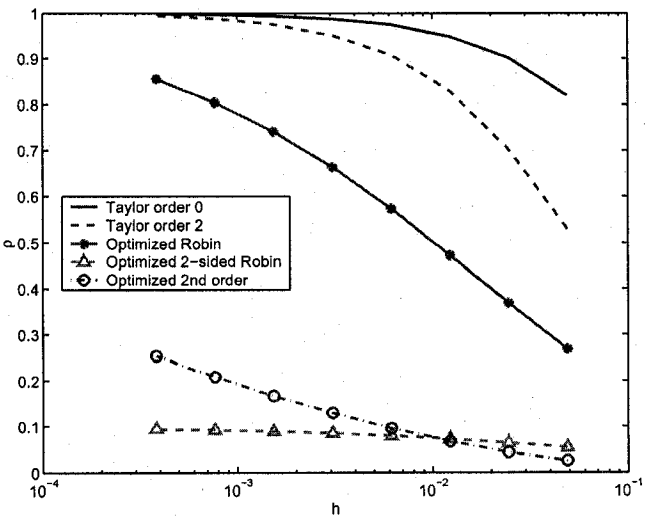

Figure 3.16: Maximum of the optimized convergence factor for different values of h, and for different choices of transmission conditions, in the case $\nu_1 = \eta_1 = 1$, $\nu_2 = 0.1$, $\eta_2 = 0.5$.

3.7.2 Anisotropic Diffusions

In this section, we consider the more general problem

$$\begin{cases} -\nabla \cdot (\mathbf{A}(\mathbf{x})\nabla u) + c(\mathbf{x})u &= f \quad \text{for } \mathbf{x} \in \Omega \subseteq \mathbb{R}^2, \\ \mathcal{B}(u) &= g \quad \text{for } \mathbf{x} \in \partial\Omega. \end{cases} \tag{3.37}$$

where $\mathbf{A}(\mathbf{x})$ is a symmetric positive definite matrix and $c(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \Omega$. This problem allows us to consider anisotropic media, where the diffusion might be different in different directions. We will show that optimized transmission conditions for the Schwarz method for this problem are obtained by solving the *same* min-max problems that were previously studied for the case of a scalar, isotropic diffusion coefficient $\nu(\mathbf{x})$.

Suppose the coefficients are piecewise constant on a non-overlapping domain de-

composition Ω_1, Ω_2, with interface Γ,

$$\mathbf{A}(\mathbf{x}) = \begin{cases} \mathbf{A}^{(1)} & \text{in } \Omega_1, \\ \mathbf{A}^{(2)} & \text{in } \Omega_2, \end{cases} \qquad c(\mathbf{x}) = \begin{cases} c^{(1)} & \text{in } \Omega_1, \\ c^{(2)} & \text{in } \Omega_2, \end{cases}$$

where

$$\mathbf{A}^{(j)} = \begin{bmatrix} a_{11}^{(j)} & a_{12}^{(j)} \\ a_{12}^{(j)} & a_{22}^{(j)} \end{bmatrix}$$

are constant, symmetric, positive definite matrices, for $j = 1, 2$. If we write $u_j := u|_{\Omega_j}$, then it can be shown (basically using integration by parts) that "natural" matching conditions across the interface are

$$\begin{aligned} u_1 &= u_2 && \text{on } \Gamma, \\ (\mathbf{A}^{(1)}\mathbf{n}) \cdot \nabla u_1 &= (\mathbf{A}^{(2)}\mathbf{n}) \cdot \nabla u_2 && \text{on } \Gamma, \end{aligned}$$

where $\mathbf{n}$ is the normal vector to Γ, pointing outward with respect to Ω_1. These are just the conditions of continuity of the solution and of its co-normal derivative (or flux).

We analyze a model problem on the domain $\Omega = \mathbb{R}^2$ decomposed into the two subdomains $\Omega_1 = (-\infty, 0) \times \mathbb{R}$ and $\Omega_2 = (0, \infty) \times \mathbb{R}$, with homogeneous right-hand side, $f \equiv 0$. The coupling conditions on the interface $\Gamma := \{x = 0\}$ translate into

$$\begin{aligned} u_1 &= u_2 && \text{at } x = 0, \\ (a_{11}^{(1)}\partial_x + a_{12}^{(1)}\partial_y)u_1 &= (a_{11}^{(2)}\partial_x + a_{12}^{(2)}\partial_y)u_2 && \text{at } x = 0. \end{aligned}$$

Using these, we write a consistent general Schwarz iteration in the form

$$\begin{aligned} &-\nabla \cdot (\mathbf{A}^{(j)} \nabla u_j^{n+1}) + c^{(j)} u_j^{n+1} = 0, \quad \text{in } \Omega_j, \quad j = 1, 2, \\ &\left(a_{11}^{(1)}\partial_x + a_{12}^{(1)}\partial_y + \mathcal{S}_1\right) u_1^{n+1} = \left(a_{11}^{(2)}\partial_x + a_{12}^{(2)}\partial_y + \mathcal{S}_1\right) u_2^{n} \quad \text{at } x = 0, \\ &\left(a_{11}^{(2)}\partial_x + a_{12}^{(2)}\partial_y - \mathcal{S}_2\right) u_2^{n+1} = \left(a_{11}^{(1)}\partial_x + a_{12}^{(1)}\partial_y - \mathcal{S}_2\right) u_1^{n} \quad \text{at } x = 0, \end{aligned} \tag{3.38}$$

where $\mathcal{S}_j$ are linear operators acting in the tangential direction to the interface, with corresponding Fourier symbols $\sigma_j(k)$. Written explicitly, the differential equation we wish to solve in each subdomain is

$$-a_{11}^{(j)} \frac{\partial^2 u_j}{\partial x^2} - a_{22}^{(j)} \frac{\partial^2 u_j}{\partial y^2} - 2a_{12}^{(j)} \frac{\partial^2 u_j}{\partial x \partial y} + c^{(j)} u_j = 0, \quad \text{in } \Omega_j, \quad \text{for } j = 1, 2.$$

When taking the Fourier transform of this equation, the characteristic roots we get for the ODE are

$$\lambda_{\pm}^{(j)}(k) = \frac{ika_{12}^{(j)} \pm \sqrt{\left[a_{11}^{(j)}a_{22}^{(j)} - (a_{12}^{(j)})^2\right]k^2 + a_{11}^{(j)}c^{(j)}}}{a_{11}^{(j)}}.$$

Note that $\mathrm{Re}(\lambda_{+}^{(j)}) > 0$ and $\mathrm{Re}(\lambda_{-}^{(j)}) < 0$. To simplify the notation, let

$$d^{(j)} := \det(\mathbf{A}^{(j)}), \quad \gamma^{(j)}(k) := \sqrt{d^{(j)}k^2 + a_{11}^{(j)}c^{(j)}}.$$

Using Fourier analysis, we find that the convergence factor for the general Schwarz iteration (3.38) can be written as

$$\rho(k,\sigma_1,\sigma_2) := \left|\frac{\hat{u}_j^{n+1}(0,k)}{\hat{u}_j^{n-1}(0,k)}\right| = \left|\frac{[\sigma_1(k) - \gamma^{(2)}(k)]}{[\sigma_1(k) + \gamma^{(1)}(k)]} \cdot \frac{[\sigma_2(k) - \gamma^{(1)}(k)]}{[\sigma_2(k) + \gamma^{(2)}(k)]}\right|. \tag{3.39}$$

Optimal and Optimized Schwarz Methods

From (3.39), it is clear that optimal symbols are given by

$$\sigma_1^{opt}(k) = \gamma^{(2)}(k), \quad \sigma_2^{opt}(k) = \gamma^{(1)}(k).$$

When there are positive reaction terms, $c^{(j)} > 0$, one could use Taylor approximations of order 2 of the optimal symbols in order to get local transmission conditions,

$$\gamma^{(j)}(k) \approx \sqrt{a_{11}^{(j)}c^{(j)}} + \frac{d^{(j)}}{2\sqrt{a_{11}^{(j)}c^{(j)}}}k^2,$$

but such approximations only give fast convergence for low frequencies.

Now, by inspection of the convergence factor (3.39), we can observe that with the change of variable

$$\nu_j := \sqrt{d^{(j)}}, \quad \eta_j := \frac{a_{11}^{(j)}c^{(j)}}{\sqrt{d^{(j)}}}, \tag{3.40}$$

we obtain the expression

$$\rho(k,\sigma_1,\sigma_2) = \left|\frac{\left[\sigma_1(k) - \nu_2\sqrt{k^2 + \frac{\eta_2}{\nu_2}}\right]}{\left[\sigma_1(k) + \nu_1\sqrt{k^2 + \frac{\eta_1}{\nu_1}}\right]} \cdot \frac{\left[\sigma_2(k) - \nu_1\sqrt{k^2 + \frac{\eta_1}{\nu_1}}\right]}{\left[\sigma_2(k) + \nu_2\sqrt{k^2 + \frac{\eta_2}{\nu_2}}\right]}\right|,$$

which is exactly the convergence factor of the general Schwarz iteration when applied to the equation

$$-\nabla \cdot (\nu(x)\nabla u) + \eta(x)u = 0, \tag{3.41}$$

with piecewise constant coefficients on the subdomains (see Section 3.7.1). Hence, any result on problem (3.41) can readily be used for the more general problem (3.37). For instance, the min-max problems we need to solve to compute optimized transmission conditions are the same, under the transformation (3.40). Note that the same conclusion can be reached through an appropriate change of coordinates in each subdomain.

In particular, when $c(x) = 0$, the results of Section 3.4 can be directly applied to compute optimized parameters for problem (3.37). As an example, if we consider the one-sided Robin conditions

$$\sigma_1(k) = \nu_2 p = \left[a_{11}^{(2)} a_{22}^{(2)} - (a_{12}^{(2)})^2\right]^{\frac{1}{2}} p, \qquad \sigma_2(k) = \nu_1 p = \left[a_{11}^{(1)} a_{22}^{(1)} - (a_{12}^{(1)})^2\right]^{\frac{1}{2}} p,$$

then the optimized parameter value is given by $p^* = \sqrt{k_1 k_2}$, using Theorem 3.6.

The same equivalence also applies for the case of continuous coefficients, including the possible use of an overlap, and thus all results found in [32] can be extended to anisotropic diffusions.

3.7.3 A Note About 3D Problems

For self-adjoint problems in three dimensions, the min-max problems associated to optimized transmission conditions are the same as in two dimensions, for a model problem. Consider a diffusion-reaction equation

$$\begin{cases} -\nabla \cdot (\nu(\mathbf{x})\nabla u) + \eta(\mathbf{x}) = 0 & \text{in } \Omega = \mathbb{R}^3, \\ |u| < \infty & \text{as } \mathbf{x} \to \infty. \end{cases}$$

The domain is divided into two non-overlapping subdomains (two half-spaces), $\Omega_1 = (-\infty, 0) \times \mathbb{R} \times \mathbb{R}$ and $\Omega_2 = (0, \infty) \times \mathbb{R} \times \mathbb{R}$, and we assume the coefficients are piecewise constant on the subdomains, as before. For the convergence analysis, we can use a

Fourier transform for both the y and z variables,

$$\mathcal{F}_{y,z}(u(x,y,z)) := \int_{\mathbb{R}}\int_{\mathbb{R}} u(x,y,z)e^{-i(my+nz)}dydz = \hat{u}(x,m,n).$$

The convergence factor of a general Schwarz iteration will then be

$$\rho(m,n;\sigma_j) = \left|\frac{\sigma_1(m,n) - \nu_2\sqrt{m^2+n^2+c_2}}{\sigma_1(m,n) + \nu_1\sqrt{m^2+n^2+c_1}} \cdot \frac{\sigma_2(m,n) - \nu_1\sqrt{m^2+n^2+c_1}}{\sigma_2(m,n) + \nu_2\sqrt{m^2+n^2+c_2}}\right|.$$

This is the same expression we have obtained previously in the 2D case, with $k^2 = m^2 + n^2$. So, if we choose σ_j to be constants or to be functions of $m^2 + n^2$ only, then the only thing that changes is the minimum and maximum frequency to set when solving the min-max problem. In 3D, we would choose

$$k_1 = \sqrt{m_1^2 + n_1^2} = \sqrt{\frac{\pi^2}{D_y^2} + \frac{\pi^2}{D_z^2}},$$
$$k_2 = \sqrt{m_2^2 + n_2^2} = \sqrt{\frac{\pi^2}{h_y^2} + \frac{\pi^2}{h_z^2}},$$

where D_y, D_z are the sizes of the domain and h_y, h_z are the mesh sizes in the y and z directions respectively.

Chapter 4

Behavior of Optimized Schwarz Methods with Many Subdomains

In this chapter, we expose and discuss the results of some numerical experiments when using optimized Schwarz method with many subdomains, but without the help of a coarse space correction. We restrict ourselves to the case of non-overlapping domain decompositions in vertical strips. We then investigate several possible explanations for the convergence behavior that we observe.

4.1 Problem Setup and Preliminaries

We consider a two-dimensional problem on a square with homogeneous Dirichlet boundary conditions, as in Section 3.6, namely

$$\begin{cases} -\nabla \cdot (\nu(\mathbf{x})\nabla u) &= 1 \quad \text{in } \Omega = (0,\pi) \times (0,\pi), \\ u &= 0 \quad \text{on } \partial\Omega. \end{cases} \tag{4.1}$$

The square is decomposed into M non-overlapping vertical strips of the same width,

$$\Omega_j = (x_{j-1}, x_j) \times (0,\pi), \quad \text{for } j = 1, 2, \ldots, M,$$

where $x_j = j\frac{\pi}{M}$. We denote the interfaces by $\Gamma_j := \{x = x_j\}$. We choose a piecewise constant diffusion coefficient which is alternating on the subdomains, i.e. $\nu(\mathbf{x}) = \nu_1$

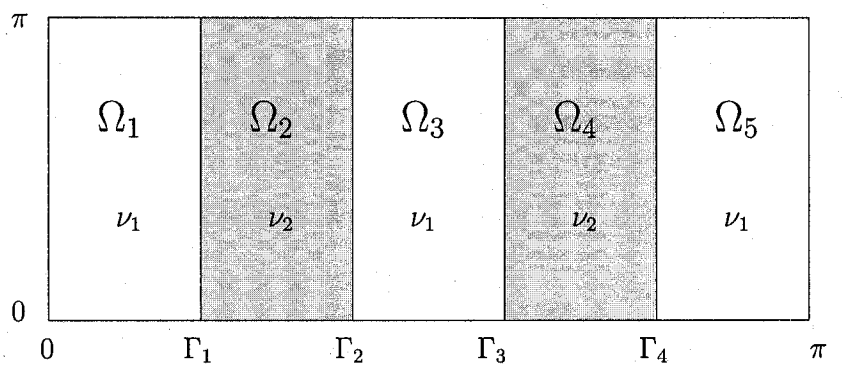

Figure 4.1: Model problem on a square decomposed in vertical strips, with alternating diffusion coefficient.

in Ω_{2j+1} and $\nu(\mathbf{x}) = \nu_2$ in Ω_{2j}, as illustrated by Figure 4.1. For the experiments, we set $\nu_1 = \frac{1}{\mu}$ and $\nu_2 = 1$, with $\mu > 1$.

It should be kept in mind that in these numerical experiments, we are effectively *changing* the problem we are solving as we increase the number of subdomains, since we are changing the diffusion coefficient, and thus making the problem harder.

We use the parallel Schwarz iteration

$$\begin{cases} -\nabla \cdot (\nu_j \nabla u_j^{n+1}) = 1 & \text{for } \mathbf{x} \in \Omega_j, \\ u_j^{n+1} = 0 & \text{on } \partial\Omega_j \cap \partial\Omega, \\ -\nu_j \dfrac{\partial u_j^{n+1}}{\partial x} + \mathcal{S}_j^{(l)} u_j^{n+1} = -\nu_{j-1} \dfrac{\partial u_{j-1}^{n}}{\partial x} + \mathcal{S}_j^{(l)} u_{j-1}^{n} & \text{at } x = x_{j-1} \text{ (left interface)}, \\ \nu_j \dfrac{\partial u_j^{n+1}}{\partial x} + \mathcal{S}_j^{(r)} u_j^{n+1} = \nu_{j+1} \dfrac{\partial u_{j+1}^{n}}{\partial x} + \mathcal{S}_j^{(r)} u_{j+1}^{n} & \text{at } x = x_j \text{ (right interface)}. \end{cases}$$

To compute the error at each iteration, we first construct a global approximation by taking an average on the interfaces,

$$U^n(\mathbf{x}) = \begin{cases} u_j^n(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega_j \\ \frac{1}{2}\left[u_j^n(\mathbf{x}) + u_{j+1}^n(\mathbf{x})\right] & \text{if } x = x_j,\ j = 1, 2, \ldots, M-1, \end{cases}$$

and then computing the maximum absolute difference (L^∞ norm) between U^n and the solution of the global problem (4.1), discretized using the same grid.

4.1.1 Choosing Optimized Transmission Conditions

We set up the optimized transmission conditions in a natural way, as follows. For each interface $\Gamma_j = \{x = x_j\}$, we look at the diffusion coefficient locally on the right and left side of the interface, and compute the optimized parameters for that case as if there are only 2 subdomains, using the results of Chapter 3. This gives the transmission operators $\mathcal{S}_{j-1}^{(r)}$ and $\mathcal{S}_j^{(l)}$ to be used on the interface Γ_j.

4.1.2 Dirichlet-Neumann Methods

For two subdomains, there are two versions of the Dirichlet-Neumann method, (DN) and (ND), depending on which side of the interface we impose the Dirichlet transmission condition (see Section 3.5). For the model problem on the infinite plane, both versions have an optimal relaxation parameter, leading to optimal convergence. However, we also noted that, without relaxation, one version of the method converges, the other one doesn't, when the diffusion coefficient is discontinuous. To ensure convergence, we must impose the Dirichlet boundary condition in the subdomain with smaller diffusion coefficient.

When using many subdomains, one strategy is the following: for each interface, we pick the "best" choice of Dirichlet or Neumann transmission conditions depending on the coefficients on each side, meaning the choice that converges even without relaxation. For example, if $\nu_j < \nu_{j+1}$, then we use the transmission conditions

$$\begin{aligned} u_j^{n+1} &= \mu_j^n && \text{for the subproblem in } \Omega_j, \\ \nu_{j+1}\frac{\partial u_{j+1}^{n+1}}{\partial n} &= \nu_j \frac{\partial u_j^n}{\partial n} && \text{for the subproblem in } \Omega_{j+1}, \end{aligned}$$

with the relaxation

$$\mu_j^{n+1} = \theta_j^* \, u_{j+1}^{n+1} + (1 - \theta_j^*) \, \mu_j^{n-1}, \tag{4.2}$$

where $\theta_j^* = (1 + \frac{\nu_j}{\nu_{j+1}})^{-1}$ is the relaxation parameter coming from the analysis on the model problem with two subdomains (see Section 3.5); however this relaxation parameter is no longer optimal when using more than two subdomains. In the case

$\nu_j > \nu_{j+1}$, then we interchange the Dirichlet and Neumann conditions and the relaxation is chosen to be $\theta_j^* = (1 + \frac{\nu_{j+1}}{\nu_j})^{-1}$ instead. We call this method "**DN best**". For our setup, when $\nu_1 = \frac{1}{\mu} < 1 = \nu_2$, this implies that the subdomains Ω_j with odd indices j have Dirichlet conditions on both sides, and subdomains with even indices have Neumann conditions on both sides. Note that here we have rewritten the transmission conditions and the relaxation formula in order to get a fully parallel iteration, whereas the Dirichlet-Neumann method was stated in its sequential formulation in Section 3.5.

It is not known how to choose the relaxation parameters appropriately for more than two subdomains, so we will also consider the Dirichlet-Neumann method, using the same transmission conditions as in "**DN best**", but without relaxation ($\theta_j^* = 1$). This will be referred in the numerical results as "**DN best no relaxation**". Note that a method without relaxation is easier to implement with the Krylov acceleration described in Section 3.6.1, since then $\mathbf{u}^{n+1}$ depends only on $\mathbf{u}^n$.

4.2 Results for a Decomposition in Vertical Strips

For all the experiments in this section, we give the number of iterations required to reach a tolerance of 10^{-6}, where the maximum number of iterations is 1000. We use a finite volume discretization on a uniform grid with 240 points in each direction. In the tables, the symbol **D** means that the iteration has ***diverged***; we stop the iteration when the L^∞ error increases over 10^6. We also give the iteration number we get when using the Schwarz method as a preconditioner for BiCGstab, as explained in Section 3.6.1. Tables 4.1 and 4.2 collect the results for several numbers of subdomains ranging from 2 to 16, in the case $\mu = 10$ and $\mu = 100$ respectively.

The first striking observation is the divergence of optimized Schwarz methods that can occur once we use many subdomains. The optimized two-sided Robin conditions seem to be the most affected, they do not converge as soon as we use 10 subdomains in the case $\mu = 10$. The only method which we know is guaranteed to converge when using multiple subdomains is the Schwarz method with the first version of optimized

Number of subdomains	2	4	6	8	10	12	16
	Method as an iterative solver						
Opt. Robin v.1	62	141	194	258	316	378	506
Opt. Robin v.2	44	84	117	152	187	223	294
Opt. 2-sided Robin	13	16	23	73	D	D	D
Opt. 2nd order	14	18	25	30	44	109	D
DN best	4	32	D	D	D	D	D
DN best no relaxation	14	37	D	D	D	D	D
	With Krylov acceleration (BiCGstab)						
Opt. Robin v.1	25	32	45	57	62	64	83
Opt. Robin v.2	16	21	29	33	33	39	44
Opt. 2-sided Robin	10	11	16	19	26	33	48
Opt. 2nd order	9	12	17	21	25	29	41
DN best no relaxation	6	15	22	25	49	55	108

Table 4.1: Number of iterations to reach a tolerance of 10^{-6}, for an increasing number of subdomains, when $\boldsymbol{\mu = 10}$.

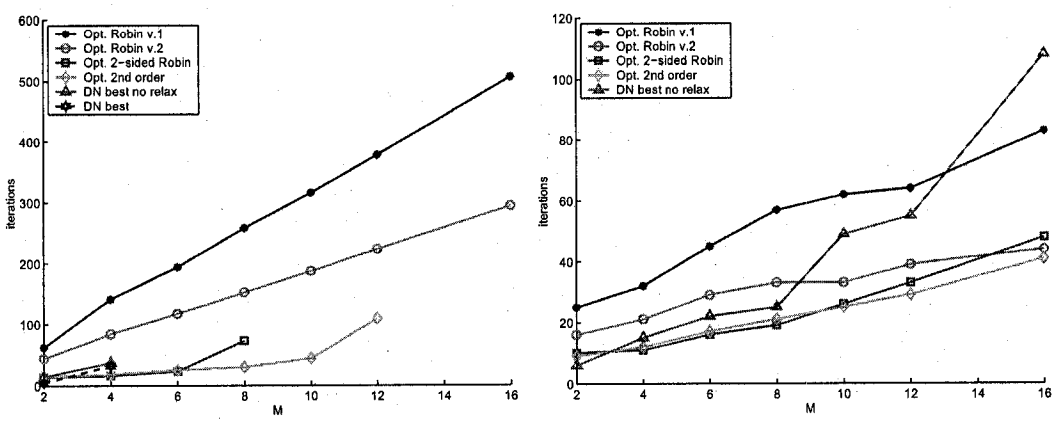

Figure 4.2: Growth of the number of iterations as a function of M, for the iterative solver on the left, and for the preconditioned Krylov subspace method on the right, in the case $\boldsymbol{\mu = 10}$, $N = 240$.

Number of subdomains	2	4	6	8	10	12	16
	Method as an iterative solver						
Opt. Robin v.1	72	74	86	114	141	170	230
Opt. Robin v.2	15	23	29	35	41	47	58
Opt. 2-sided Robin	7	9	10	15	17	25	71
Opt. 2nd order	8	10	11	13	15	20	40
DN best	4	11	15	23	35	51	> 1000
DN best no relaxation	8	11	13	19	25	40	> 1000
	With Krylov acceleration (BiCGstab)						
Opt. Robin v.1	25	34	57	116	254	432	790
Opt. Robin v.2	8	10	13	16	17	19	23
Opt. 2-sided Robin	6	8	8	11	15	15	23
Opt. 2nd order	5	7	10	11	15	21	23
DN best no relaxation	6	8	10	13	17	15	23

Table 4.2: Number of iterations to reach a tolerance of 10^{-6}, for an increasing number of subdomains, when $\boldsymbol{\mu = 100}$.

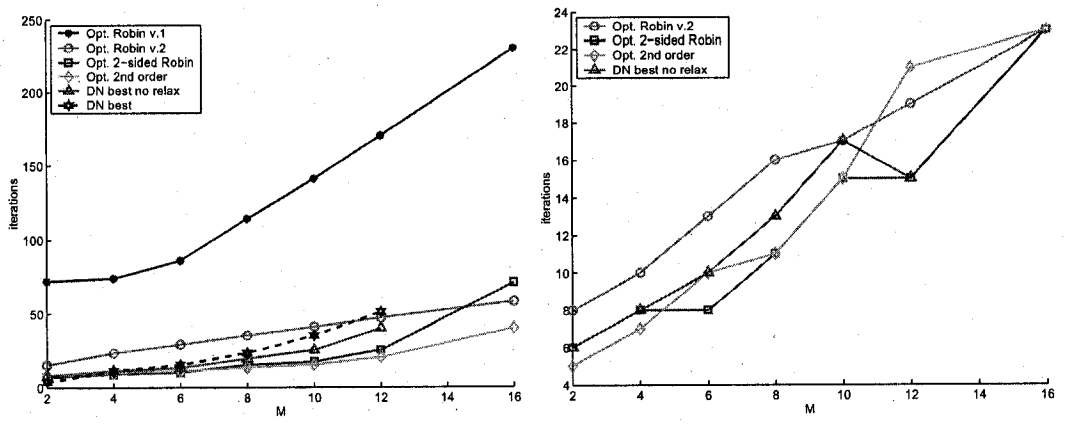

Figure 4.3: Growth of the number of iterations as a function of M, for the iterative solver on the left, and for the preconditioned Krylov subspace method on the right, in the case $\boldsymbol{\mu = 100}$, $N = 240$.

Robin conditions (as in Section 3.4.1), since in that case we have a general convergence proof, which was presented in Section 3.2. When increasing the ratio of coefficients, the convergence of optimized Schwarz methods is improved in general, but it seems to only delay the problem: when $\mu = 100$, no divergence occurs for up to 16 subdomains, but the iteration numbers for the optimized two-sided Robin conditions start showing the sign of a breakdown at $M = 16$. Under Krylov acceleration, naturally the iteration will never diverge, but it may not completely remove the problem, as there might still be a quick deterioration in convergence for larger numbers of subdomains. Also, note that this divergence phenomenon does not occur when the coefficient is constant over the entire domain, see Appendix B.1.

When plotting the iteration numbers as a function of the number of subdomains, as in Figure 4.2 and 4.3, we observe a linear growth in M, until the method potentially breaks down and starts diverging. A linear growth is unavoidable when using the Schwarz method by itself, since at least M iterations are required to globally communicate the information about the source term f across all subdomains. We also observe that the slope of the linear growth is not the same for all methods, and that the convergence of optimized transmission conditions with two free parameters is more robust with respect to the number of subdomains compared to the optimized one-sided Robin conditions, however they break down and diverge sooner for large M. Using Krylov acceleration significantly reduces this slope as well.

Also note that, when $\mu = 100$, the Krylov subspace method BiCGstab is having difficulties with the first version of optimized Robin conditions. This is a peculiar behavior that we do not fully understand.

In the next two sections, we will investigate two possible explanations for the divergence of optimized Schwarz methods, as observed in Table 4.1. The first one is the effect of solving on very thin subdomains or subdomains of very different width (Section 4.3). Another possible cause for the divergence is seen by looking at the optimal operators for the Schwarz iteration with many subdomains, and observing that these operators can sometimes be badly approximated by the optimized transmission conditions when the coefficients are varying in the domain (Section 4.4).

4.2.1 Red-black Coloring

We have also implemented a version of the Schwarz methods (including Dirichlet-Neumann without relaxation) which uses a *red-black* coloring. The idea is to color the subdomains so that two neighboring subdomains do not share the same color. In our case, with vertical strips, this is achieved by coloring *red* the subdomains with odd indices say, and the ones with even indices *black*. Then, one Schwarz iteration is performed in two steps: first, we solve in all the red subdomains (in parallel), and then in all the black subdomains (again in parallel) by using the most recent data from the red subdomains. This cuts the number of iterations by 2, and each iteration takes twice as long, so we have not gained much. However, such a strategy can be useful for example if we do not have enough processors to handle solving in all the subdomains in parallel.

In Table 4.3, we show the convergence results for the various methods implemented with red-black coloring when $\mu = 10$. By comparing the numbers in Table 4.1 and Table 4.3, the number of iteration is effectively cut by a factor of 2 when using red-black coloring, except, interestingly, for the Dirichlet-Neumann method without relaxation: the version with red-black coloring and accelerated with BiCGstab is converging significantly faster than the expected factor of 2.

4.3 Convergence Analysis for Bounded Rectangles

The optimized transmission conditions we used in the experiments are derived by minimizing the convergence factor of the Schwarz iteration for a model problem on the infinite plane. In practice, we use these optimized conditions for solving problems on bounded domains, so we don't expect to always get the *best* convergence, but remain fairly close. However, when the subdomains are very thin or have very different sizes, the model problem on the infinite plane is no longer an acceptable approximation to make. To make this more precise, in this section we derive the convergence factor

Number of subdomains	2	4	6	8	10	12	16
	Method as an iterative solver						
Opt. Robin v.1	23	61	89	121	144	177	235
Opt. Robin v.2	19	39	56	73	89	107	141
Opt. 2-sided Robin	6	8	10	32	D	D	D
Opt. 2nd order	6	9	11	15	22	57	D
DN best no relaxation	7	18	D	D	D	D	D
	With Krylov acceleration (BiCGstab)						
Opt. Robin v.1	12	19	22	30	29	31	36
Opt. Robin v.2	9	12	13	17	19	21	25
Opt. 2-sided Robin	5	6	7	9	11	13	16
Opt. 2nd order	5	7	8	12	11	15	20
DN best no relaxation	5	7	9	11	14	16	22

Table 4.3: Number of iterations needed by the Schwarz method with red-black coloring to reach a tolerance of 10^{-6}, for an increasing number of subdomains, when $\boldsymbol{\mu = 10}$, $N = 240$.

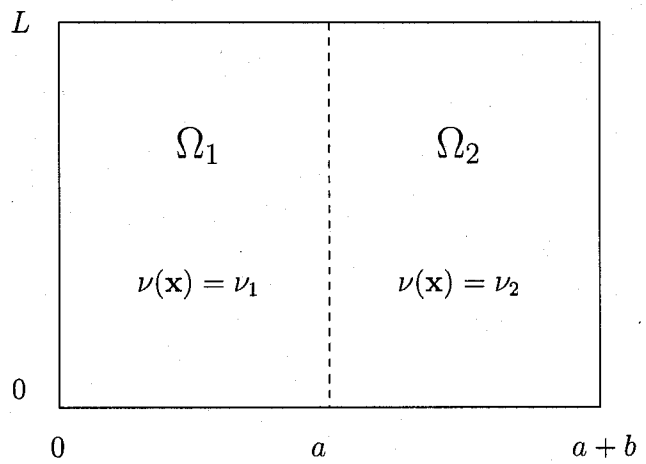

Figure 4.4: Model problem with two rectangular subdomains of width a and b.

corresponding to a different model problem on a bounded domain. This analysis is relevant for the problem we are solving in Section 4.1, since when M is large, the subdomains become very thin, and each vertical interface near the right or left boundary cuts the domain Ω into two very unequal parts.

Consider the model problem with homogeneous Dirichlet boundary conditions,

$$\begin{cases} -\nabla \cdot (\nu(\mathbf{x})\nabla u) &= 0 \quad \text{in } \Omega := (0, a+b) \times (0, L), \\ u &= 0 \quad \text{on } \partial\Omega. \end{cases} \tag{4.3}$$

As shown in Figure 4.4, we decompose the domain into two non-overlapping subdomains,

$$\Omega_1 := (0,\, a) \times (0,\, L), \quad \Omega_2 := (a,\, a+b) \times (0,\, L),$$

and assume that the diffusion coefficient is piecewise constant, $\nu(\mathbf{x}) = \nu_j$ on Ω_j. We first analyze the convergence of a Schwarz iteration with Robin transmission conditions

$$\begin{cases} -\nu_j \Delta u_j^{n+1} &= 0 & \text{in } \Omega_j, \\ u_j^{n+1} &= 0 & \text{on } \partial\Omega_j \cap \partial\Omega, \\ (\nu_1 \partial_x + \nu_2 p_1) u_1^{n+1} &= (\nu_2 \partial_x + \nu_2 p_1) u_2^n & \text{at } x = a, \\ (\nu_2 \partial_x - \nu_1 p_2) u_2^{n+1} &= (\nu_1 \partial_x - \nu_1 p_2) u_1^n & \text{at } x = a. \end{cases}$$

In this case, we can use separation of variables and Fourier series (instead of a continuous Fourier transform) to write down the solution of the subproblems. To simplify the notation, let $k_m := \frac{m\pi}{L}$. In Ω_1, we get a solution in the form

$$u_1^n(x,y) = \sum_{m=1}^{\infty} A_m^n \sinh(k_m x) \sin(k_m y),$$

and, in Ω_2, we can write the solution in the form

$$u_2^n(x,y) = \sum_{m=1}^{\infty} B_m^n \sinh(k_m(b+a-x)) \sin(k_m y).$$

Applying the first transmission condition $(\nu_1\partial_x + \nu_2 p_1)u_1^{n+1} = (\nu_2\partial_x + \nu_2 p_1)u_2^n$ at $x = a$, and matching the series term by term, we find for each m the equation

$$A_m^{n+1} = \frac{\nu_2 p_1 \sinh(bk_m) - \nu_2 k_m \cosh(bk_m)}{\nu_2 p_1 \sinh(ak_m) + \nu_1 k_m \cosh(ak_m)} B_m^n.$$

Similarly, applying the second transmission condition in the Schwarz iteration at $x = a$, $(\nu_2\partial_x - \nu_1 p_2)u_2^{n+1} = (\nu_1\partial_x - \nu_1 p_2)u_1^n$, and once again matching the series term by term, we get the equation

$$B_m^{n+1} = \frac{\nu_1 p_2 \sinh(ak_m) - \nu_1 k_m \cosh(ak_m)}{\nu_1 p_2 \sinh(bk_m) + \nu_2 k_m \sinh(bk_m)} A_m^n.$$

Combining these two equations gives the convergence factor for a double step of the Schwarz iteration

$$\begin{aligned}\rho_{R2}(k_m, p_1, p_2) &:= \left|\frac{A_m^{n+1}}{A_m^{n-1}}\right| = \left|\frac{B_m^{n+1}}{B_m^{n-1}}\right|, \\ &= \left|\frac{p_1 - k_m \coth(bk_m)}{p_1 + \lambda k_m \coth(ak_m)}\right| \cdot \left|\frac{p_2 - k_m \coth(ak_m)}{p_2 + \frac{1}{\lambda} k_m \coth(bk_m)}\right|,\end{aligned}$$

where $\lambda := \frac{\nu_1}{\nu_2}$. Thus, when considering the model problem (4.3) discretized using a uniform grid with N points in the y direction, better suited optimized two-sided Robin conditions are obtained by solving the min-max problem

$$\min_{p_1,p_2>0} \left(\max_{m=1,2,\ldots,N} \rho_{R2}(k_m, p_1, p_2) \right). \tag{4.4}$$

This optimization problem has two main differences when compared to the min-max problem solved in Section 3.4.3. First, in (4.4) the convergence factor is uniformly minimized over a discrete set of frequencies, instead of a continuous range (this justifies in part why we usually choose $k_1 = \frac{\pi}{L}$ and $k_2 = N\frac{\pi}{L}$). Secondly, the convergence factor itself differs with the presence of the terms $\coth(ak_m)$ and $\coth(bk_m)$. These two differences make the min-max problem (4.4) harder to solve analytically. When a and b are relatively large with respect to the height L, then the coth terms are exponentially close to 1. For example, if $a = b = L$, then

$$1 < \coth(Lk_m) \leq \coth(Lk_1) = \coth(\pi) \approx 1.004,$$

so in that case, using the convergence factor for the model problem on the infinite plane is a fairly good approximation. Problems arise when either a or b is small relative to L (i.e. at least one of the subdomains is thin).

With similar calculations we can derive the convergence factors for second order transmission conditions and Dirichlet-Neumann methods

$$\rho_{2nd}(k_m, p, q) = \left| \frac{p + qk_m^2 - k_m \coth(bk_m)}{p + qk_m^2 + \lambda k_m \coth(ak_m)} \cdot \frac{p + qk_m^2 - k_m \coth(ak_m)}{p + qk_m^2 + \frac{1}{\lambda} k_m \coth(bk_m)} \right|,$$

$$\rho_{DN}(k_m, \theta) = \left| 1 - \theta \left(1 + \lambda \frac{\tanh(bk_m)}{\tanh(ak_m)} \right) \right|,$$

$$\rho_{ND}(k_m, \theta) = \left| 1 - \theta \left(1 + \lambda \frac{\tanh(ak_m)}{\tanh(bk_m)} \right) \right|.$$

Now we look at the convergence of the different methods for problem (4.3), when the optimized parameters are chosen as in Chapter 3, not by solving min-max problems like (4.4). For the two-sided Robin and second order conditions, suppose we use the optimized parameters (p^*, q^*) from Sections 3.4.3 and 3.4.4 respectively, which are independent of a and b. Similarly, for the Dirichlet-Neumann methods, suppose we use the optimized relaxation parameter from the previous chapter

$$\theta^*_{DN} = \frac{1}{1+\lambda}, \quad \theta^*_{ND} = \frac{1}{1+\frac{1}{\lambda}}.$$

To simulate the extreme case of two very thin subdomains of the same width, let $a = b = \varepsilon$ with $\varepsilon \to 0$. First, we see that the Dirichlet-Neumann methods are always

optimal for symmetric subdomains, i.e. whenever $a = b$ we have

$$\rho_{DN}(k_m, \theta^*_{DN}) = \rho_{ND}(k_m, \theta^*_{ND}) \equiv 0.$$

On the other hand, the convergence of the optimized Schwarz methods progressively deteriorates,

$$\rho_{R2}(k_m, p_1^*, p_2^*) \to 1, \quad \rho_{2nd}(k_m, p^*, q^*) \to 1,$$

but without diverging.

Another extreme situation is the case of very unsymmetrical subdomains, when one subdomain is thin while the other one is wide. For example, let $a \to 0$ and $b \to \infty$. The asymptotic expansions of the convergence factor are

$$\begin{aligned}
\rho_{R2}(k_m, p_1^*, p_2^*) &= \left| \frac{p_1^* - k_m}{\lambda p_2^* + k_m} \right| + O(a), \\
\rho_{2nd}(k_m, p^*, q^*) &= \left| \frac{p^* + q^* k_m^2 - k_m}{\lambda(p^* + q^* k_m^2) + k_m} \right| + O(a), \\
\rho_{DN}(k_m, \theta^*_{DN}) &= \frac{\lambda}{1+\lambda} \left(\frac{1}{a} \right) + O(1), \\
\rho_{ND}(k_m, \theta^*_{ND}) &= \frac{1}{1+\lambda} + O(a).
\end{aligned}$$

It is clear that the Dirichlet-Neumann method (DN) will diverge quickly when a is small. For the two-sided Robin conditions, when $\lambda < 1$ the optimized parameters have the h-asymptotics $p_1^* = O(h^{-1})$ and $p_2^* = O(1)$, so we find that

$$\rho_{R2} = O\left(h^{-1}\right),$$

in the limit $a \to 0$, $b \to \infty$. Hence, the method will also diverge when a and h are small. Similarly, for the second order conditions, the optimized parameters have the h-asymptotics $p^* = O(h^{-\frac{1}{4}})$ and $q^* = O(h^{\frac{3}{4}})$, and in that case we get, in the limit,

$$\rho_{2nd} = O\left(h^{-\frac{1}{4}}\right).$$

Hence, asymptotically the second order conditions can also diverge when $\lambda < 1$, but they are less sensitive to different subdomain widths compared to the two-sided Robin conditions.

α	$\frac{1}{20}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{7}{8}$	$\frac{19}{20}$
Opt. Robin v.1	65	65	78	87	69	65	65
Opt. Robin v.2	63	61	74	84	67	63	63
Opt. 2-sided Robin	264	25	24	25	23	24	23
Opt. 2nd order	19	18	17	18	17	17	18
DN	D	52	18	4	14	20	20
ND	46	34	22	4	30	D	D

Table 4.4: Number of iterations needed by the Schwarz method to reach a tolerance of 10^{-6}, for different values of α, when $\boldsymbol{\mu = 2}$, $N = 200$.

For the case of two subdomains, let us verify numerically how the different methods behave when breaking the symmetry of the decomposition. For this purpose, consider the problem (4.3) with $L = \pi$, $a = \alpha\pi$ and $b = (1 - \alpha)\pi$. We use a uniform grid with 200 points in each direction. Tables 4.4 and 4.5 list the iteration numbers when varying α. We see that the Dirichlet-Neumann methods (with optimized relaxation parameters) depart very quickly from optimality when the subdomains do not have the same size. Also, these results confirm that the convergence of optimized two-sided Robin conditions is the most quickly affected among the optimized methods, and has a convergence that deteriorates when the subdomain with the smaller diffusion coefficient has a small width.

To see if the analysis exposed in this section can explain the divergence we observed, we reproduce the setup of the experiment corresponding to Table 4.1: let $\nu_1 = 0.1$, $\nu_2 = 1$, $k \in [1, \pi/240]$ and consider the optimized two-sided Robin parameters given in Section 3.4.3. Moreover, let us look at the first interface, which cuts the square domain into two subregions of widths $a = \frac{\pi}{M}$ and $b = \frac{(M-1)\pi}{M}$. Simple calculations then show that you need at least $M = 44$ to make the convergence factor ρ_{R2} bigger than 1, for the Schwarz iteration with two subdomains. So, the analysis of this section, although relevant for certain situations, does not seem to fully explain the divergence of the Schwarz method for numbers of subdomains as low as 10, as seen in Table 4.1.

α	$\frac{1}{20}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{7}{8}$	$\frac{19}{20}$
Opt. Robin v.1	41	51	55	54	41	39	41
Opt. Robin v.2	33	35	41	41	33	31	33
Opt. 2-sided Robin	29	13	13	13	13	13	13
Opt. 2nd order	16	13	13	13	13	13	13
DN	39	18	11	4	12	11	11
ND	108	56	30	4	44	D	D

Table 4.5: Number of iterations needed by the Schwarz method to reach a tolerance of 10^{-6}, for different values of α, when $\boldsymbol{\mu} = \mathbf{10}$, $N = 200$.

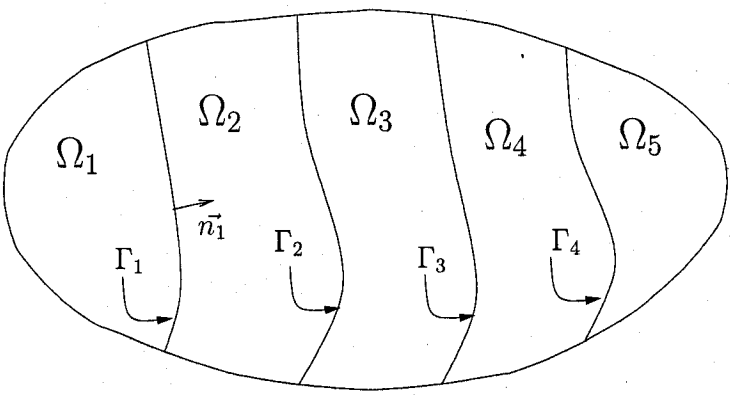

Figure 4.5: A general decomposition in strips.

4.4 Approximating the Optimal Operators

In the case of two subdomains, we showed in Section 3.3.2 that there are *optimal* operators for the Schwarz iteration that lead to convergence after 2 iterations. The optimized transmission conditions are trying to approximate (in a certain way) these optimal operators. In this section, we give the extension of these operators to more subdomains, and show that the optimized conditions we chose in Section 4.1.1 may not be good approximations for these operators.

Consider a bounded domain $\Omega \subset \mathbb{R}^2$, decomposed into M *strips*, as shown in Figure 4.5. Denote by Γ_j the interface to the right of subdomain Ω_j, and the associated

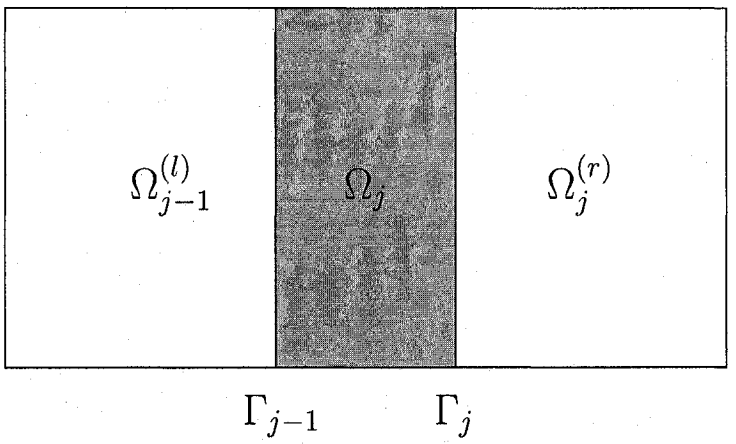

Figure 4.6: Notation for the definition of optimal operators.

normal vector directed outward relative to Ω_j by n_j. Also, denote by $\Omega_j^{(r)}$ and $\Omega_j^{(l)}$ the part of the domain Ω to the right and left of the interface Γ_j respectively, as shown in Figure 4.6. Thus,

$$\Omega = \Omega_j^{(l)} \cup \Gamma_j \cup \Omega_j^{(r)}.$$

We wish to solve the diffusion problem

$$\begin{cases} -\nabla \cdot (\nu(\mathbf{x})\nabla u) &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{cases} \tag{4.5}$$

where the diffusion coefficient is piecewise constant on the subdomains, $\nu(\mathbf{x}) = \nu_j$ for $\mathbf{x} \in \Omega_j$.

For each interface Γ_j, $j = 1, 2, \ldots, M-1$, we can define two Dirichlet-to-Neumann operators that we denote $\Lambda_j^{(r)}$ and $\Lambda_j^{(l)}$. For the right side of the interface,

$$\Lambda_j^{(r)}(u) := -\nu_{j+1}\frac{\partial u^{(r)}}{\partial n_j} \quad \text{on } \Gamma_j, \quad \text{where } u^{(r)} \text{ solves the problem}$$

$$\begin{cases} -\nabla \cdot \left(\nu(\mathbf{x})\nabla u^{(r)}\right) &= f \quad \text{in } \Omega_j^{(r)}, \\ u^{(r)} &= 0 \quad \text{on } \partial\Omega_j^{(r)} \cap \partial\Omega, \\ u^{(r)} &= u \quad \text{on } \Gamma_j. \end{cases}$$

Similarly, there is a Dirichlet-to-Neumann operator that involves solving to the left side of the interface,

$$\Lambda_j^{(l)}(u) := \nu_j \frac{\partial u^{(l)}}{\partial n_j} \quad \text{on } \Gamma_j, \quad \text{where } u^{(l)} \text{ solves the problem}$$

$$\begin{cases} -\nabla \cdot \left(\nu(\mathbf{x}) \nabla u^{(l)}\right) &= f & \text{in } \Omega_j^{(l)}, \\ u^{(l)} &= 0 & \text{on } \partial\Omega_j^{(l)} \cap \partial\Omega, \\ u^{(l)} &= u & \text{on } \Gamma_j. \end{cases}$$

Using these operators in the transmission conditions, we consider the Schwarz iteration

$$\begin{cases} -\nu_j \Delta u_j^{n+1} &= f & \text{in } \Omega_j, \\ u_j^{n+1} &= 0 & \text{on } \partial\Omega_j \cap \partial\Omega, \\ \nu_j \dfrac{\partial u_j^{n+1}}{\partial n_j} + \Lambda_j^{(r)} u_j^{n+1} &= \nu_{j+1} \dfrac{\partial u_{j+1}^{n}}{\partial n_j} + \Lambda_j^{(r)} u_{j+1}^{n} & \text{on } \Gamma_j \ (j < M), \\ -\nu_j \dfrac{\partial u_j^{n+1}}{\partial n_{j-1}} + \Lambda_{j-1}^{(l)} u_j^{n+1} &= -\nu_{j-1} \dfrac{\partial u_{j-1}^{n}}{\partial n_{j-1}} + \Lambda_{j-1}^{(l)} u_{j-1}^{n} & \text{on } \Gamma_{j-1} \ (j > 1). \end{cases} \tag{4.6}$$

These transmission conditions are sometimes called exact *absorbing* boundary conditions [42], since they are precisely the boundary conditions to be imposed when truncating the computational domain from Ω to Ω_j without changing the solution of the problem. Namely, if u is the solution of the global problem (4.5), then its restriction on Ω_j, $u_j := u|_{\Omega_j}$, solves the truncated problem

$$\begin{cases} -\nu_j \Delta u_j &= f & \text{in } \Omega_j, \\ u_j &= 0 & \text{on } \partial\Omega_j \cap \partial\Omega, \\ \nu_j \dfrac{\partial u_j}{\partial n_j} + \Lambda_j^{(r)} u_j &= 0 & \text{on } \Gamma_j \ (j < M), \\ -\nu_j \dfrac{\partial u_j}{\partial n_{j-1}} + \Lambda_{j-1}^{(l)} u_j &= 0 & \text{on } \Gamma_{j-1} \ (j > 1). \end{cases} \tag{4.7}$$

The following theorem is due to Nataf, Rogier and de Sturler [56].

Theorem 4.1 (Optimal convergence). *The Schwarz iteration* (4.6) *converges in exactly M iterations.*

Proof. After the first iteration, the solution u_1^1 solves the diffusion problem in Ω_1, which coincides with $\Omega_1^{(l)}$ by definition, and thus

$$-\nu_1 \frac{\partial u_1^1}{\partial n_1} + \Lambda_1^{(l)} u_1^1 = 0 \quad \text{on } \Gamma_1,$$

by definition of the Dirichlet-to-Neumann map $\Lambda_1^{(l)}$. So, at the second iteration, we solve a problem in Ω_2 using the boundary condition

$$-\nu_2 \frac{\partial u_2^2}{\partial n_1} + \Lambda_1^{(l)} u_2^2 = 0 \quad \text{on } \Gamma_1.$$

This boundary condition implies that u_2^2 can be extended to a function $\tilde{u}_2^2$ that solves the diffusion problem on the region $\Omega_2^{(l)} = \Omega_1 \cup \Gamma_1 \cup \Omega_2$. Hence, we find that

$$-\nu_2 \frac{\partial u_2^2}{\partial n_2} + \Lambda_2^{(l)} u_2^2 = 0, \quad \text{on } \Gamma_2,$$

which gives for u_3^3 the transmission condition

$$-\nu_3 \frac{\partial u_3^3}{\partial n_2} + \Lambda_2^{(l)} u_3^3 = 0, \quad \text{on } \Gamma_2.$$

This argument can be continued by propagating the homogeneous transmission condition from left to right. By induction, we can show that at iteration j,

$$-\nu_j \frac{\partial u_j^j}{\partial n_{j-1}} + \Lambda_{j-1}^{(l)} u_j^j = 0, \quad \text{on } \Gamma_{j-1}.$$

Moreover, note that once the transmission condition is 0, it remains 0 for all the iterations thereafter.

The same induction argument can be done by going from right to left instead, to show that

$$\nu_j \frac{\partial u_j^{M-j+1}}{\partial n_j} + \Lambda_j^{(r)} u_j^{M-j+1} = 0, \quad \text{on } \Gamma_j.$$

This implies that, at iteration M, the approximations in the subdomains all satisfy the boundary conditions

$$\begin{cases} \nu_j \dfrac{\partial u_j^M}{\partial n_j} + \Lambda_j^{(r)} u_j^M &= 0 \quad \text{on } \Gamma_j \ (j < M), \\ -\nu_j \dfrac{\partial u_j^M}{\partial n_{j-1}} + \Lambda_{j-1}^{(l)} u_j^M &= 0 \quad \text{on } \Gamma_{j-1} \ (j > 1), \end{cases}$$

i.e. u_j^M solves the problem (4.7), and therefore $u_j^M = u|_{\Omega_j}$ for all j. The iteration has thus converged after exactly M steps. □

Important observation: The Dirichlet-to-Neumann operator $\Lambda_j^{(r)}$ on the interface Γ_j does not depend only on ν_{j+1}, but on the diffusion coefficient $\nu(\mathbf{x})$ *everywhere* in $\Omega_j^{(r)}$. Thus, for example, if $\nu_{j+1} = 1$ but $\nu(\mathbf{x}) = 10^{-6}$ on most the domain $\Omega_j^{(r)}$, then our optimized transmission conditions, which take into consideration the diffusion coefficient only locally near the interface, might do a terrible job at approximating the operator $\Lambda_j^{(r)}$.

In the following subsections, we study in more detail these Dirichlet-to-Neumann operators, with the goal of computing optimized transmission conditions that represent better approximations for them.

4.4.1 Dirichlet-to-Neumann Maps in 1D

For a one-dimensional problem, with straightforward calculations we can compute the Dirichlet-to-Neumann operator for multiple subdomains. Consider the domain $\Omega = (a, b)$, the interfaces $x_i = iH$ where $H = \frac{b-a}{M}$, and the piecewise constant diffusion coefficient $\nu(x) = \nu_i$ for $x_{i-1} < x < x_i$. The Dirichlet-to-Neumann map at $x = a$ is defined by

$$\Lambda\alpha = -\nu_1 \frac{\mathrm{d}u}{\mathrm{d}x} \quad \text{at } x = a, \quad \text{where } u \text{ solves}$$

$$\begin{cases} -\dfrac{\mathrm{d}}{\mathrm{d}x}\left(\nu(x)\dfrac{\mathrm{d}u}{\mathrm{d}x}\right) = 0 & \text{for } x \in (a, b), \\ u(a) = \alpha, \\ u(b) = 0. \end{cases}$$

It is easy to show by direct calculations that

$$\Lambda\alpha = \left[\sum_{i=1}^{M} \frac{\nu_i^{-1}}{M}\right]^{-1} \frac{1}{b-a}\,\alpha. \tag{4.8}$$

In fact, this is simply a special case of the general formula

$$\Lambda\alpha = \left[\int_a^b \frac{1}{\nu(x)}\,\mathrm{d}x\right]^{-1} \alpha,$$

which appears in [47] as the optimal Robin parameter in 1D. Thus, in this case it is easy to use the optimal transmission conditions in practice to get convergence in M iterations, since they are simple Robin conditions.

For comparison purposes, when the diffusion coefficient is constant, the Dirichlet-to-Neumann operator is

$$\Lambda\alpha = \frac{\nu}{b-a}\alpha.$$

By comparing this with the formula (4.8), we see that the value

$$\tilde{\nu} := \left[\sum_{i=1}^{M} \frac{\nu_i^{-1}}{M}\right]^{-1} \tag{4.9}$$

is a *homogenized* diffusion coefficient for the domain. We use the term *homogenized* because this diffusion coefficient is closely related to the theory of homogenization of periodic structures, see for example [7] and [53]. Notice that this homogenized coefficient $\tilde{\nu}$ can be very different from the local coefficient value ν_1. For example, if $\nu_{2j+1} = 1$ and $\nu_{2j} = 0.1$ (and M is even), then $\tilde{\nu} = 0.\overline{18} \ll \nu_1 = 1$.

4.4.2 Dirichlet-to-Neumann Maps in 2D

The analysis in two dimensions is not quite as trivial. In the following, we show that the Dirichlet-to-Neumann (DtN) map for a model problem is also related to a homogenized coefficient value, but for low frequencies only in y. Consider the domain $\Omega = (0, L) \times \mathbb{R}$, and the decomposition into M infinite vertical strips

$$\Omega_j = (x_{j-1}, x_j) \times \mathbb{R}, \quad \text{where } x_j := j\frac{L}{M}.$$

Assume the diffusion coefficient is piecewise constant on the subdomains, $\nu(\mathbf{x}) = \nu_j$ in Ω_j, and denote the constant subdomain width by $H := \frac{L}{M}$. One can also consider vertical strips that are bounded in y with homogeneous Dirichlet conditions on the top and bottom boundaries, but the only change this would bring is we would have a discrete set of relevant frequencies instead of a continuous range (as in Section 4.3).

Consider the Dirichlet-to-Neumann map

$$\Lambda\alpha(y) := -\nu_1\frac{\partial u}{\partial x} \quad \text{at } x = 0, \quad \text{where } u \text{ solves the problem}$$

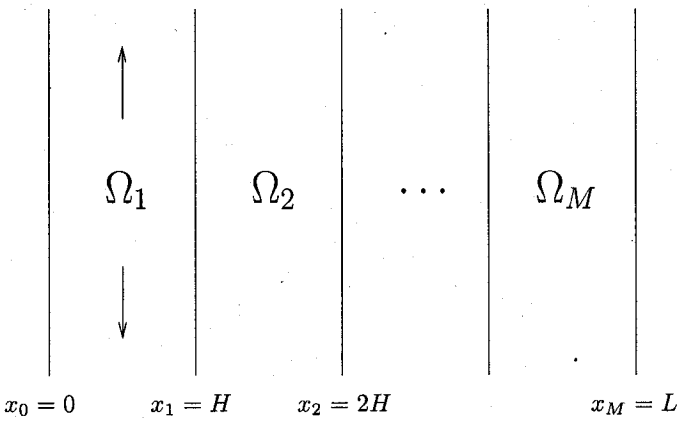

Figure 4.7: A decomposition into M infinite strips.

$$\begin{cases} -\nabla \cdot (\nu(\mathbf{x})\nabla u) = 0 & \text{for } \mathbf{x} \in \Omega, \\ |u(x,y)| < \infty & \text{as } |y| \to \infty, \\ u(0,y) = \alpha(y), \\ u(L,y) = 0. \end{cases} \tag{4.10}$$

Using a Fourier transform in the variable y, we get that the solution of problem (4.10) in each subdomain must be of the form

$$\hat{u}_j(x,k) = A_j(k)\cosh(|k|(x - x_{j-1})) + B_j(k)\sinh(|k|(x - x_{j-1})).$$

To simplify the notation, let

$$C := \cosh(H|k|), \quad S := \sinh(H|k|), \quad T := \tanh(H|k|).$$

In Fourier space, the solution to (4.10) must satisfy, at the interface $x = x_j$, the matching conditions

$$\hat{u}_j(x_j,k) = \hat{u}_{j+1}(x_j,k), \qquad \nu_j \frac{\partial \hat{u}_j}{\partial x}(x_j,k) = \nu_{j+1}\frac{\partial \hat{u}_{j+1}}{\partial x}(x_j,k),$$

which yield the following equations for the unknown coefficients $A_j(k)$ and $B_j(k)$, for $j = 1, ..., M-1$,

$$\begin{aligned} A_{j+1}(k) &= CA_j(k) + SB_j(k), \\ B_{j+1}(k) &= \frac{\nu_j}{\nu_{j+1}}\left[SA_j(k) + CB_j(k)\right]. \end{aligned}$$

In addition, the boundary conditions translate into the relations

$$A_1(k) = \hat{\alpha}(k), \qquad CA_M(k) + SB_M(k) = 0.$$

In total, we get $2M$ equations for the $2M$ unknown coefficients $A_j(k)$ and $B_j(k)$. We need to solve this linear system in order to compute $B_1(k)$, and then the Dirichlet-to-Neumann map in Fourier space is given by

$$(\hat{\Lambda\alpha})(k) = -\nu_1 \frac{\partial \hat{u}_1}{\partial x}(0,k) = -\nu_1 |k| B_1(k).$$

By directly solving the linear system for small numbers of subdomains, we get explicit formulas for the Dirichlet-to-Neumann map:

- for $M = 2$,

$$(\hat{\Lambda\alpha})(k) = \frac{1}{T} \frac{\nu_1 T^2 + \nu_2}{\nu_1 + \nu_2} \nu_1 |k| \hat{\alpha}(k),$$

- for $M = 3$,

$$(\hat{\Lambda\alpha})(k) = \frac{1}{T} \frac{(\nu_1\nu_2 + \nu_2^2 + \nu_1\nu_3)T^2 + \nu_2\nu_3}{\nu_2^2 T^2 + \nu_1\nu_2 + \nu_1\nu_3 + \nu_2\nu_3} \nu_1 |k| \hat{\alpha}(k),$$

- for $M = 4$,

$$(\hat{\Lambda\alpha})(k) = \frac{1}{T} \frac{(\nu_1\nu_3^2)T^4 + (\nu_1\nu_3\nu_4 + \nu_1\nu_2\nu_4 + \nu_2\nu_3^2 + \nu_2^2\nu_4 + \nu_1\nu_2\nu_3 + \nu_2^2\nu_3)T^2 + \nu_2\nu_3\nu_4}{(\nu_2^2\nu_4 + \nu_2\nu_3^2 + \nu_1\nu_3^2 + \nu_2^2\nu_3)T^2 + \nu_1\nu_2\nu_3 + \nu_1\nu_2\nu_4 + \nu_1\nu_3\nu_4 + \nu_2\nu_3\nu_4} \nu_1 |k| \hat{\alpha}(k).$$

These formulas quickly get messy as we increase M, and there does not seem to be an obvious pattern forming which we could generalize.

Evaluating the DtN map in general

We are interested in computing the Fourier symbol $\sigma_{dtn}(k)$ of the DtN map Λ for a general number of subdomains. For this purpose, let $\hat{\alpha}(k) \equiv 1$. In what follows, we will drop the explicit dependence of the coefficients on k to shorten the notation.

We rewrite the linear system for the coefficients A_j and B_j as

$$\begin{cases} A_1 &= 1, \\ \begin{bmatrix} A_{j+1} \\ B_{j+1} \end{bmatrix} &= \begin{bmatrix} C & S \\ \frac{\nu_j}{\nu_{j+1}} S & \frac{\nu_j}{\nu_{j+1}} C \end{bmatrix} \begin{bmatrix} A_j \\ B_j \end{bmatrix} \quad \text{for } j = 1, 2, \ldots, M-1, \\ A_M + TB_M &= 0. \end{cases} \tag{4.11}$$

If we define P_j to be the 2 by 2 matrix in the above system,

$$P_j := \begin{bmatrix} C & S \\ \frac{\nu_j}{\nu_{j+1}} S & \frac{\nu_j}{\nu_{j+1}} C \end{bmatrix},$$

then its inverse is

$$P_j^{-1} = \frac{1}{C^2 - S^2} \begin{bmatrix} C & -\frac{\nu_{j+1}}{\nu_j} S \\ -S & \frac{\nu_{j+1}}{\nu_j} C \end{bmatrix}.$$

Using these small matrices only, we can write down an equation relating B_1 in terms of B_M,

$$\begin{aligned} \begin{bmatrix} 1 \\ B_1 \end{bmatrix} &= P_1^{-1} P_2^{-1} \cdots P_{M-1}^{-1} \begin{bmatrix} A_M \\ B_M \end{bmatrix} \\ &= P_1^{-1} P_2^{-1} \cdots P_{M-1}^{-1} \begin{bmatrix} -T \\ 1 \end{bmatrix} B_M, \end{aligned} \tag{4.12}$$

using the boundary condition at $x = L$. So, if we first compute the vector

$$\begin{bmatrix} \xi \\ \eta \end{bmatrix} := P_1^{-1} P_2^{-1} \cdots P_{M-1}^{-1} \begin{bmatrix} -T \\ 1 \end{bmatrix},$$

then we get that $B_M = \frac{1}{\xi}$ and consequently

$$B_1(k) = \frac{\eta}{\xi}.$$

This gives us a strategy for calculating the Fourier symbol $\sigma_{dtn}(k)$ of the DtN map, where

$$(\hat{\Lambda\alpha})(k) = \sigma_{dtn}(k)\hat{\alpha}(k), \quad \text{and} \quad \sigma_{dtn}(k) = -\nu_1 |k| B_1(k).$$

For a given frequency k, computing $\sigma_{dtn}(k)$ basically only requires the multiplication of $M-1$ simple 2 by 2 matrices. Another important observation for the result to come is that we can multiply the right hand side of equation (4.12) by any scalar, without changing the value of the coefficient $B_1(k)$.

Asymptotics for small and large frequencies

Even if we could not produce a general formula for the Fourier symbol of the DtN map, we are still able to look at how the operator behaves for small and large frequencies asymptotically.

Theorem 4.2 (Asymptotics for the Fourier symbol of the DtN map). *For small and large frequencies, we have the following asymptotic behavior*

$$\begin{aligned} \sigma_{dtn}(k) &\approx \frac{1}{\tanh(L|k|)}\tilde{\nu}|k| && \text{for small } k, \\ \sigma_{dtn}(k) &\approx \nu_1|k| && \text{for large } k, \end{aligned}$$

where $\tilde{\nu}$ is the homogenized diffusion coefficient given by (4.9)*, from the 1D analysis.*

Proof. To obtain a more convenient equation to work with instead of (4.12), we multiply each matrix P_j^{-1} by a function of k and absorb the reciprocal into B_M, without affecting B_1. Let

$$\tilde{P}_j^{-1} := C(1-T^2)P_j^{-1} = \begin{bmatrix} 1 & -\frac{\nu_{j+1}}{\nu_j}T \\ -T & \frac{\nu_{j+1}}{\nu_j} \end{bmatrix}.$$

$$\tilde{B}_M = \frac{1}{C^{M-1}(1-T^2)^{M-1}}B_M.$$

With these changes, we get a modified equation for B_1

$$\begin{bmatrix} 1 \\ B_1 \end{bmatrix} = \tilde{P}_1^{-1}\tilde{P}_2^{-1}\cdots\tilde{P}_{M-1}^{-1}\begin{bmatrix} -T \\ 1 \end{bmatrix}\tilde{B}_M, \tag{4.13}$$

which is going to be better suited to the asymptotic analysis. The equation (4.13) also makes it clear that B_1 depends on k and H only through $T = \tanh(H|k|)$.

In the limit when $k \to \infty$, we have $T = \tanh(Hk) \to 1$ and the matrices $\tilde{P}_j^{-1}$ become singular,

$$\tilde{P}_j^{-1} = \begin{bmatrix} 1 & -\frac{\nu_{j+1}}{\nu_j} \\ -1 & \frac{\nu_{j+1}}{\nu_j} \end{bmatrix}.$$

Since each $\tilde{P}_j^{-1}$ is of the form $\begin{bmatrix} a & b \\ -a & -b \end{bmatrix}$, then their product will also be of the same form, because

$$\begin{bmatrix} a & b \\ -a & -b \end{bmatrix} \begin{bmatrix} c & d \\ -c & -d \end{bmatrix} = \begin{bmatrix} c(a-b) & d(a-b) \\ -c(a-b) & -d(a-b) \end{bmatrix}.$$

Therefore, we get, still in the limit $k \to \infty$,

$$\begin{bmatrix} \tilde{\xi} \\ \tilde{\eta} \end{bmatrix} := \prod_{j=1}^{M-1} \tilde{P}_j^{-1} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} a & b \\ -a & -b \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} b-a \\ a-b \end{bmatrix},$$

for some value of a and b. So,

$$B_1(k) = \frac{\tilde{\eta}}{\tilde{\xi}} = -1,$$

and we get the Fourier symbols $\sigma_{dtn}(k) = \nu_1|k|$.

In the limit when $k \to 0$, we have $T = \tanh(Hk) \to 0$. From the explicit formulas we have derived in the cases $M = 2$, 3, 4, we expect B_1 to be growing like $1/T$, and so we cannot take a direct limit in equation (4.13). We first claim that the product of the matrices $\tilde{P}_j^{-1}$ is given asymptotically by

$$\prod_{j=1}^{M-1} \tilde{P}_j^{-1} = \begin{bmatrix} 1 + O(T^2) & -\nu_M \sum_{j=1}^{M-1} \nu_j^{-1} T \\ O(T) & \frac{\nu_M}{\nu_1} + O(T^2) \end{bmatrix}. \tag{4.14}$$

This is proved by induction on M. First, when $M = 2$, equation (4.14) holds trivially by looking at the definition of $\tilde{P}_1^{-1}$. For the induction step, suppose that

$$\prod_{j=1}^{M-2} \tilde{P}_j^{-1} = \begin{bmatrix} 1 + O(T^2) & -\nu_{M-1} \sum_{j=1}^{M-2} \nu_j^{-1} T \\ O(T) & \frac{\nu_{M-1}}{\nu_1} + O(T^2) \end{bmatrix}.$$

Then, we get

$$\begin{aligned}
\prod_{j=1}^{M-1} \tilde{P}_j^{-1} &= \left(\prod_{j=1}^{M-2} \tilde{P}_j^{-1}\right) \tilde{P}_{M-1}^{-1} \\
&= \begin{bmatrix} 1+O(T^2) & -\nu_{M-1}\sum_{j=1}^{M-2}\nu_j^{-1}T \\ O(T) & \frac{\nu_{M-1}}{\nu_1}+O(T^2) \end{bmatrix} \begin{bmatrix} 1 & -\frac{\nu_M}{\nu_{M-1}}T \\ -T & \frac{\nu_M}{\nu_{M-1}} \end{bmatrix} \\
&= \begin{bmatrix} 1+O(T^2) & \left[-\frac{\nu_M}{\nu_{M-1}} - \nu_M\sum_{j=1}^{M-2}\nu_j^{-1}\right]T \\ O(T) & \frac{\nu_M}{\nu_1}+O(T^2) \end{bmatrix} \\
&= \begin{bmatrix} 1+O(T^2) & -\nu_M\sum_{j=1}^{M-1}\nu_j^{-1}T \\ O(T) & \frac{\nu_M}{\nu_1}+O(T^2) \end{bmatrix},
\end{aligned}$$

which proves our claim by induction. Using equation (4.14), we get

$$\begin{bmatrix} \tilde{\xi} \\ \tilde{\eta} \end{bmatrix} = \prod_{j=1}^{M-1} \tilde{P}_j^{-1} \begin{bmatrix} -T \\ 1 \end{bmatrix} = \begin{bmatrix} -\left(1+\nu_M\sum_{j=1}^{M-1}\nu_j^{-1}\right)T + O(T^3) \\ \frac{\nu_M}{\nu_1}+O(T^2) \end{bmatrix}.$$

Therefore, we get the asymptotic expansion

$$B_1(k) = \frac{-\nu_M/\nu_1 + O(T^2)}{T\left(1+\nu_M\sum_{j=1}^{M-1}\nu_j^{-1}\right)+O(T^3)} = \frac{-1}{\nu_1 T}\left[\sum_{j=1}^{M}\nu_j^{-1}\right]^{-1} + O(T),$$

$$\begin{aligned}
\Rightarrow \quad \sigma(k) = -\nu_1|k|B_1(k) &= \frac{1}{M\tanh(Hk)}\left[\frac{1}{M}\sum_{j=1}^{M}\nu_j^{-1}\right]^{-1}|k| + O(T) \\
&= \frac{1}{\tanh(Lk)}\tilde{\nu}|k| + O(T),
\end{aligned}$$

when k is small. In the last equality, we used the fact that

$$M\tanh(Hk) = \tanh(MHk) + O(k^2) = \tanh(Lk) + O(k^2).$$

□

This result shows that for highly oscillatory functions on the interface, the DtN map acts like the DtN map corresponding to a constant diffusion coefficient $\nu(\mathbf{x}) \equiv \nu_1$

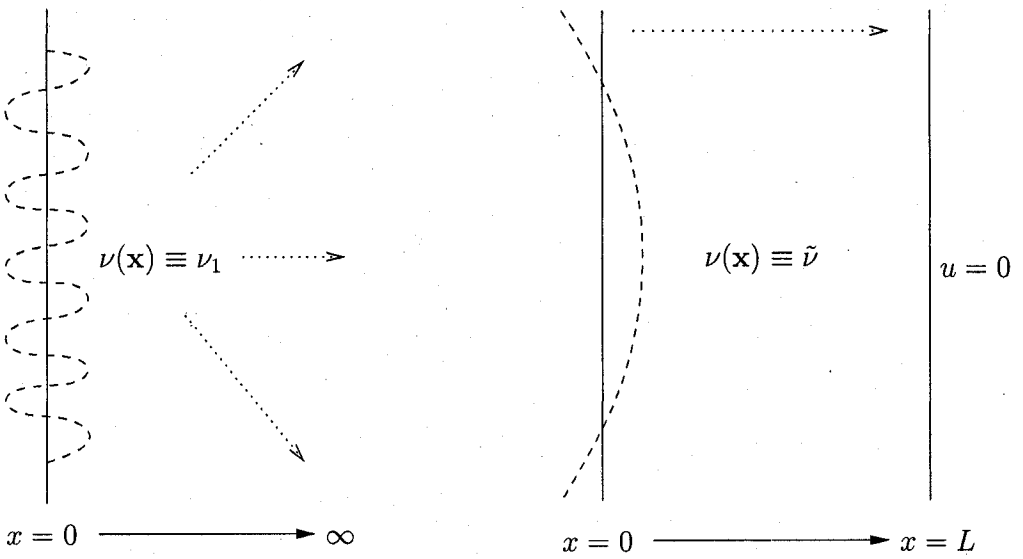

Figure 4.8: Equivalent problems defining the DtN map asymptotically for large frequencies (left) and small frequencies (right).

in a half-plane. In other words, the DtN map applied to large frequencies is not affected by the diffusion coefficient far from the interface nor by the boundary condition at $x = L$. Intuitively this is not surprising since the DtN operator is smoothing, high frequencies do not propagate very far into the domain (by the diffusive nature of the problem). At the other end, for low frequencies, the DtN map acts like the DtN map corresponding to a constant homogenized coefficient $\nu(\mathbf{x}) \equiv \tilde{\nu}$ on the entire domain, *with* the Dirichlet conditions taken into account. Again, intuitively, low frequencies have more global interaction over the entire domain.

So, with the choice of optimized transmission conditions described in Section 4.1.1, in which for every interface we look at the diffusion coefficient locally and use the optimized parameters corresponding to the model problem on two half-planes, we are making a good approximations for large frequencies but a bad one for low frequencies. In order to approximate the DtN operator more accurately for low frequencies, we try a different strategy, in which for the interface Γ_j we solve an optimization problem

using the convergence factor

$$\rho_{DtN}(k,\sigma_i) = \left| \frac{\sigma_1(k) - \sigma_{dtn}^{(r)}(k)}{\sigma_1(k) + \sigma_{dtn}^{(l)}(k)} \cdot \frac{\sigma_2(k) - \sigma_{dtn}^{(l)}(k)}{\sigma_2(k) + \sigma_{dtn}^{(r)}(k)} \right|, \tag{4.15}$$

where $\sigma_{dtn}^{(l)}$ and $\sigma_{dtn}^{(r)}$ are the Fourier symbols of the DtN maps corresponding the domains $\Omega_j^{(l)}$ and $\Omega_j^{(r)}$ respectively, and $\sigma_i(k)$ are given by the choice of transmission conditions in the Schwarz iteration (e.g. two-sided Robin, second order, etc.). We do not have explicit formulas for the Fourier symbols of the DtN maps, but we have described above an easy way to compute their value for each given value of k. Thus, we can solve the min-max problems with the convergence factor (4.15) numerically, and see how the resulting optimized conditions perform compared to those described in Section 4.1.1. For instance, for two-sided Robin conditions on interface Γ_j, we solve the min-max problem

$$\min_{p_1,p_2>0} \left(\max_{k\in K_n} \left| \frac{\nu_2 p_1 - \sigma_{dtn}^{(r)}(k)}{\nu_2 p_1 + \sigma_{dtn}^{(l)}(k)} \cdot \frac{\nu_1 p_2 - \sigma_{dtn}^{(l)}(k)}{\nu_1 p_2 + \sigma_{dtn}^{(r)}(k)} \right| \right), \tag{4.16}$$

where $K_n = \{\frac{j\pi}{L} \mid j = 1, 2, \ldots, n\}$ is the discrete set of frequencies relevant to the discrete problem, on a bounded domain of height L, with homogeneous Dirichlet boundary conditions, and n discretization points in y.

Here we repeat once again the setup of Section 4.2, where a uniform grid of $N = 240$ points in each direction is used, and $\nu_1 = 0.1$, $\nu_2 = 1$. Table 4.6 shows the iteration numbers obtained when the parameters are optimized numerically by uniformly minimizing the convergence factor (4.15) (for example, by solving problem (4.16) for two-sided Robin conditions). Simple Matlab code for this numerical optimization is included in A.2. When compared with Table 4.1, we observe a significant improvement in the convergence, yet we do not completely get rid of the divergence of optimized two-sided or second order conditions when M is large enough.

4.5 Other Remarks

The analysis and remarks that were made in Sections 4.3 and 4.4 were useful and relevant in many ways, but they do not fully account for the divergence of optimized

Number of subdomains	2	4	6	8	10	12	16
	Method as an iterative solver						
Opt. Robin v.1	63	129	167	211	256	303	400
Opt. Robin v.2	44	79	104	133	161	191	253
Opt. 2-sided Robin	13	19	24	35	77	> 1000	D
Opt. 2nd order	14	20	25	33	41	69	D
	With Krylov acceleration (BiCGstab)						
Opt. Robin v.1	25	35	45	50	61	59	91
Opt. Robin v.2	17	24	29	32	33	37	42
Opt. 2-sided Robin	10	14	16	18	25	30	43
Opt. 2nd order	10	14	14	19	24	29	41

Table 4.6: Number of iterations required to reach a tolerance of 10^{-6}, for an increasing number of subdomains, when $\boldsymbol{\mu = 10}$ and the optimized parameters are obtained by using the exact Dirichlet-to-Neumann Fourier symbol in the convergence factor.

Schwarz methods.

In the search for more insight, consider the problem with coefficient ratio $\mu = 10$, and 120 grid points in each direction (note that 240 grid points were used for Table 4.1). The Schwarz method with two-sided Robin conditions, where the parameters are optimized on the model problem in the infinite plane, is diverging as soon as we use 10 subdomains. As we increase the number of subdomains, we have looked at the convergence of the Schwarz iteration for parameter values around these optimized values; we get the results shown in Figure 4.9. We see that the convergence is not very sensitive on the parameter p_1, but there is a "divergence front" that moves up relative to p_2 and eventually makes the optimized conditions diverge. Hence, to avoid divergence, the Robin parameter p_2 should be increased linearly with respect to M. Based on Figure 4.9, we can pick the best value of p_2 for each M, to get a convergent method, as shown in Table 4.7: the number of iterations grows with M as expected, but without divergence.

We suspect that the lowest frequency in y is causing the iteration to diverge. This

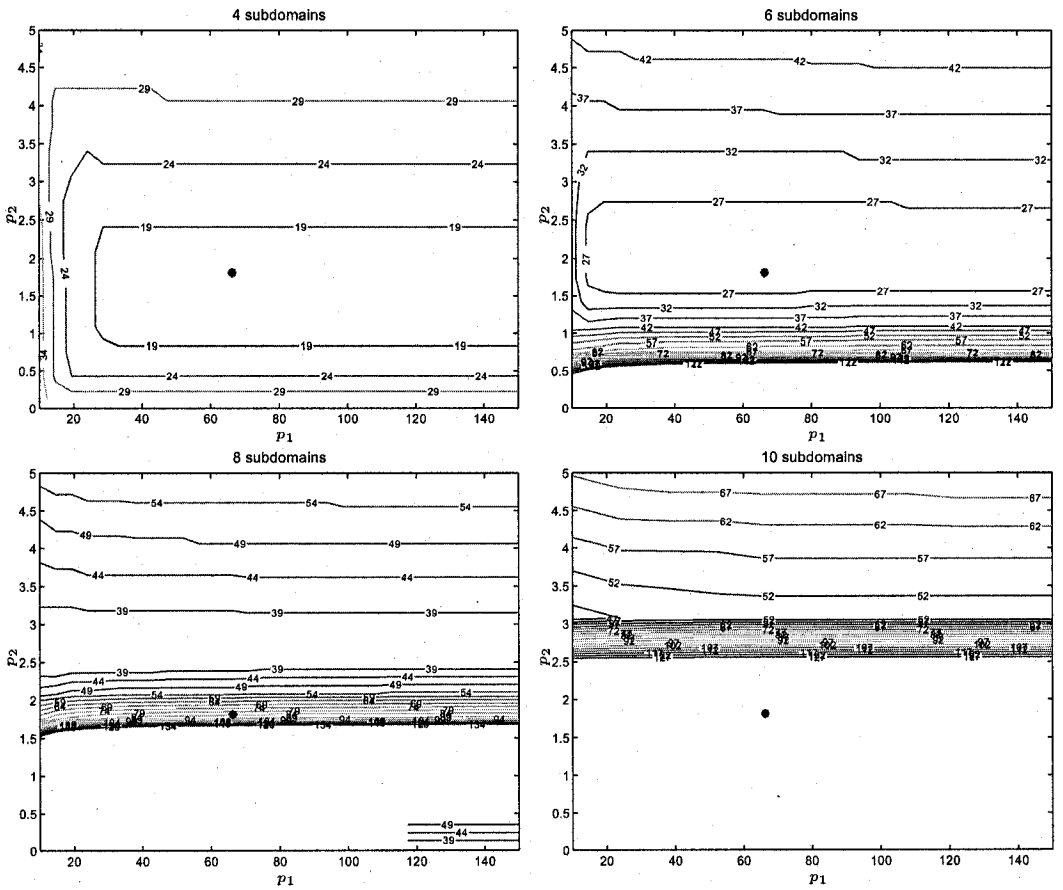

Figure 4.9: Convergence of two-sided Robin transmission conditions for choices of parameters close to the optimized values for the model problem. Here, $\mu = 10$ and $N = 120$.

	$M = 4$	$M = 6$	$M = 8$	$M = 10$
Two-sided Robin parameters (p_1, p_2)	(65, 1.5)	(65, 2)	(65, 2.75)	(65, 3.5)
Iteration number for tolerance 10^{-6}	14	21	35	51

Table 4.7: Convergence of the two-sided Robin conditions when increasing the parameter p_2 with the number of subdomains.

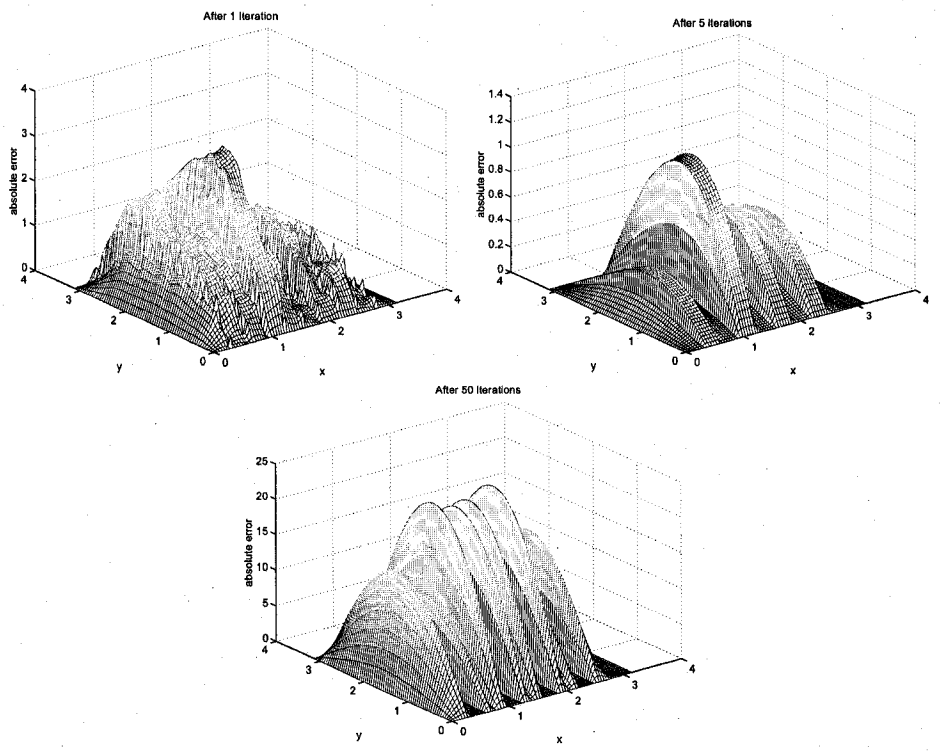

Figure 4.10: Absolute error after 1, 5 and 50 iterations of the Schwarz method with optimized two-sided Robin conditions, when $M = 10$, $\mu = 10$ and $N = 240$.

is confirmed in Figure 4.10, where the absolute error is shown at the 1^{st}, 5^{th} and 50^{th} iterations when using 10 subdomains ($\mu = 10$ and $N = 240$). The oscillatory modes are damped very quickly, and only the low frequencies remain and end up diverging.

Note that up to now, we have always optimized a convergence factor which is valid only for a Schwarz iteration on *two* subdomains. We believe that this convergence factor is not appropriate for low frequencies when using many subdomains. (For high frequencies, because they have a small range of influence, optimizing a two-subdomain convergence factor is good enough.) To resolve the divergence problem, either one should optimize the parameters by looking at a different convergence factor, or construct a coarse space correction that will handle the low frequencies globally and efficiently.

Chapter 5

Coarse Space Corrections for the Schwarz Method with Robin Conditions

For the parallel scalability of domain decomposition methods, it is necessary to have a convergence independent of the number of subdomains; ideally, if we double the number of subdomains (i.e. use twice as many processors), we would like the running time of the algorithm to be cut by a factor of two. However, for elliptic problems, if our domain decomposition method only exchanges information between neighboring subdomains at each step, we will need at least $\tilde{M}$ iterations to spread information about the source term globally across all the subdomains, where $\tilde{M}$ is the largest distance (length of shortest path) between two subdomains. Hence, we need to provide a mechanism that permits global communication across all the subdomains in order to get a convergence that does not deteriorate as the number of subdomains grows. One such mechanism consists in *correcting* our approximation, at each step, with a *coarse* function with global support. This can be viewed as a two-level method: we first apply a preconditioner on the fine space (in our case, we solve in parallel local problems in the subdomains), and then computing a correction by projecting the problem into a coarse space.

In this chapter, we present a few simple experiments for a simple one-dimensional problem, and investigate the performance of coarse space correction strategies for the Schwarz iteration with Robin transmission conditions. Among the questions that need addressing, we point out three in particular:

1. Can we design a coarse space correction that yields a convergence truly independent of the number of subdomains for the Schwarz iteration with Robin or more general transmission conditions?

2. When using overlap, are the optimized Schwarz methods combined with coarse space correction converging significantly faster than the classical two-level additive Schwarz preconditioner?

3. Should the parameters in the transmission conditions be optimized by taking into consideration the coarse correction?

We will provide partial answers to these questions through numerical experiments, and in this way motivate further analysis on this topic.

5.1 Preliminaries

Before experimenting with coarse space corrections for a simple one-dimensional problem, we first review some popular techniques for overlapping and non-overlapping domain decompositions.

5.1.1 Two-Level Additive Schwarz Preconditioner

To fix ideas, consider the problem

$$\begin{cases} -\Delta u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{cases}$$

where Ω is a bounded polygonal or polyhedral domain in $\mathbb{R}^2$ or $\mathbb{R}^3$. Let τ_h be a quasi-uniform shape-regular triangular or tetrahedral mesh, with mesh size h, and

let V_h be the space of continuous piecewise linear finite element functions on τ_h that vanish on $\partial\Omega$. We get a linear system to solve for the finite element solution $\mathbf{u}$,

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

where A is the stiffness matrix, and $\mathbf{b}$ is the product of the mass matrix and the discrete source term f.

Suppose we have an overlapping domain decomposition $\{\Omega_k\}_{k=1}^M$ conforming with the fine mesh, i.e. each Ω_k is the union of elements of τ_h. Then, we can define the local stiffness matrices A_k corresponding to the nodes lying in the interior of Ω_k. In addition, let R_k be the restriction operator from the global vector $\mathbf{u}$ to the degrees of freedom in the interior of Ω_k only. The (one-level) additive Schwarz preconditioner can be written as

$$M_{as}^{-1} = \sum_{k=1}^{M} R_k^T A_k^{-1} R_k.$$

To obtain a two-level preconditioner, we introduce a coarse mesh τ_H, where the coarse elements loosely correspond to the subdomains. More precisely, we assume that there exists a constant C such that

$$\operatorname{diam}(K) \leq C \operatorname{diam}(\Omega_i), \quad \text{whenever } K \in \tau_H \text{ and } K \cap \Omega_i \neq \emptyset.$$

Let V_H be the finite element space of continuous piecewise linear functions on the coarse mesh τ_H. Note that the fine mesh τ_h does not need to be a refinement of the coarse mesh τ_H. One only needs to define an interpolation operator $P_0 : V_H \to V_h$ from the coarse to the fine mesh. Then, let R_0 denote the adjoint of the operator P_0 in the ℓ^2-inner product (and not the matrix transpose). With these definitions, we can write the two-level additive Schwarz preconditioner as

$$M_{as2}^{-1} = P_0 A_0^{-1} R_0 + \sum_{k=1}^{M} R_k^T A_k^{-1} R_k,$$

where $A_0 := R_0 A P_0$. Note that for self-adjoint elliptic problems, M_{as2}^{-1} and A are symmetric positive definite matrices.

Theorem 5.1. *If δ denotes a measure of the overlap size between the subdomains, then we have the estimate for the spectral condition number*

$$\kappa(M_{as2}^{-1}A) := \frac{\lambda_{max}(M_{as2}^{-1}A)}{\lambda_{min}(M_{as2}^{-1}A)} \leq C\left(1+\frac{H}{\delta}\right),$$

where C is a constant independent of h, H and δ.

For a proof of this theorem and additional related results, see the survey in the books [64], [65] and references therein. If the overlap size is kept proportional to the subdomain diameter, $\delta = cH$, then the convergence of the conjugate gradient method applied to the preconditioned linear system

$$M_{as2}^{-1}A\mathbf{u} = M_{as2}^{-1}\mathbf{b}$$

is independent of the both the fine mesh size h and the number of subdomains.

5.1.2 The FETI method

We follow in this subsection the notation of [65]. For non-overlapping domain decompositions, the FETI method (Finite Element Tearing and Interconnecting) was first introduced in [29]. In this method, the substructures (subdomains) are completely "torn" apart and the approximations are allowed to not match on the interface (multiple valued); the continuity on the interface is enforced through separate constraints. Thus, if $A^{(k)}$ denotes the local stiffness matrix restricted to Ω_k, then we have M local problems

$$A^{(k)}\tilde{\mathbf{u}}_k^{full} = \tilde{\mathbf{f}}_k.$$

Here, the vector $\tilde{\mathbf{u}}_k^{full}$ includes all the unknowns in subdomain $\overline{\Omega}_k$, and we will use $\mathbf{u}_k$ to denote a vector regrouping the unknowns on the interface $\partial\Omega_k \backslash \partial\Omega$. Eliminating the degrees of freedom in the interior of the subdomain Ω_k, we get the Schur complement system on the interface $\partial\Omega_k \backslash \partial\Omega$

$$S^{(k)}\mathbf{u}_k = \mathbf{f}_k.$$

Define the vector $\mathbf{u} := (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_M)^T$ of interface degrees of freedom, and the block diagonal matrix

$$\tilde{S} := \begin{bmatrix} S^{(1)} & & & \\ & S^{(2)} & & \\ & & \ddots & \\ & & & S^{(M)} \end{bmatrix}.$$

The local Schur complement matrices $S^{(k)}$ are symmetric for self-adjoint elliptic operators. Then, we wish to solve the constrained minimization problem for $\mathbf{u}$,

$$\min_{B\mathbf{u}=0} \frac{1}{2}\langle \tilde{S}\mathbf{u}, \mathbf{u}\rangle - \langle \mathbf{f}, \mathbf{u}\rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the ℓ^2-inner product, and B is a matrix with entries equal to 0, -1 or 1 that enforces continuity on the interface. Introducing Lagrange multipliers $\boldsymbol{\lambda}$ for the constraints, we rewrite the problem in a saddle-point formulation

$$\begin{aligned} \tilde{S}\mathbf{u} + B^T\boldsymbol{\lambda} &= \mathbf{f}, \\ B\mathbf{u} &= 0. \end{aligned}$$

In the FETI method, we wish to eliminate $\mathbf{u}$ to obtain a linear system for the Lagrange multipliers $\boldsymbol{\lambda}$ only. However, for interior subdomains Ω_k that do not touch a Dirichlet boundary, the local Schur complement $S^{(k)}$ is not invertible (the inverse of $S^{(k)}$ would involve solving a pure Neumann problem in Ω_k). We can solve the equation for $\mathbf{u}$ if and only if

$$\mathbf{f} - B^T\boldsymbol{\lambda} \perp \text{kernel}(\tilde{S}). \tag{5.1}$$

This imposes a constraint on the Lagrange multipliers. Define the matrix

$$R := \begin{bmatrix} R^{(1)} & & & \\ & R^{(2)} & & \\ & & \ddots & \\ & & & R^{(M)} \end{bmatrix},$$

where the columns of $R^{(k)}$ form a basis for the kernel$(S^{(k)})$. Thus, range(R) = kernel(S). Moreover, introducing the matrices

$$\tilde{S}^\dagger := \text{a pseudo-inverse of } \tilde{S},$$

$$F := B\tilde{S}^{\dagger}B^T, \quad G := BR, \quad \mathbf{d} := B\tilde{S}^{\dagger}\mathbf{f}, \quad \mathbf{e} := R^T\mathbf{f},$$
$$P := I - G(G^TG)^{-1}G^T.$$

The matrix P is the orthogonal projection into range$(G)^{\perp}$, with respect to the inner product $\langle\cdot,\cdot\rangle$. After some manipulations to eliminate $\mathbf{u}$, we find that $\boldsymbol{\lambda}$ is a solution of

$$PF\boldsymbol{\lambda} = P\mathbf{d},$$
$$G^T\boldsymbol{\lambda} = \mathbf{e},$$

where the second equation is equivalent to the constraint (5.1). To solve this problem, the unknown is decomposed into $\boldsymbol{\lambda} = \boldsymbol{\lambda}_0 + P\tilde{\boldsymbol{\lambda}}$ where $G^T\boldsymbol{\lambda}_0 = \mathbf{e}$, and then we solve

$$PFP\tilde{\boldsymbol{\lambda}} = P\mathbf{d} - PF\boldsymbol{\lambda}_0, \tag{5.2}$$

using a preconditioned Krylov subspace method (a preconditioner is needed to remove the dependence on the mesh size h, see [28]). Note that G^TG is a square matrix of dimension equal to the number of columns in R, i.e. the dimension of kernel(S). Thus the action of P can be viewed as projecting out the coarse modes, and indeed it produces a scalable method.

This technique of imposing an additional constraint for the Lagrange multipliers was introduced in [28] for linear elasticity problems, in which case the kernel of $S^{(k)}$ is characterized by rigid body motions. The scalability of the method was demonstrated in [25]. However, note that this technique does not work, for example, when solving the Poisson problem and all the subdomains have a part of their boundary that coincides with $\partial\Omega$ where a Dirichlet boundary condition is imposed; in that case all the Schur complements are invertible.

5.1.3 Projected Interface System for General Transmission Conditions

The ideas of the FETI method can be extended and applied to non-symmetric problems and more general transmission conditions. For instance, in [40] and [44], an

advection-diffusion problem is studied and optimized second order conditions are applied; we briefly summarize here how a coarse correction can be applied in that case.

It is possible to write the Schwarz iteration with general transmission conditions as a stationary iterative method for a condensed linear system on the interface,

$$D\boldsymbol{\mu} = \mathbf{b},$$

where the vector $\boldsymbol{\mu}$ contains the transmission data for each interface. For instance, in the case of two non-overlapping subdomains and Robin transmission conditions, a Schwarz iteration can be written as

$$\begin{cases} \mathcal{L}u_j^{n+1} &= f & \text{in } \Omega_j, \\ \frac{\partial u_1^{n+1}}{\partial n} + pu_1^{n+1} &= \frac{\partial u_2^n}{\partial n} + pu_2^n & \text{on } \Gamma, \\ -\frac{\partial u_2^{n+1}}{\partial n} + qu_2^{n+1} &= -\frac{\partial u_1^n}{\partial n} + qu_1^n & \text{on } \Gamma. \end{cases}$$

In this case, the interface vector would be $\boldsymbol{\mu} = (\boldsymbol{\mu}_1,\ \boldsymbol{\mu}_2)^T$, where $\boldsymbol{\mu}_1$ is the discrete representation of $\frac{\partial u_1}{\partial n} + pu_1$ on Γ and $\boldsymbol{\mu}_2$ is the discrete representation of $-\frac{\partial u_2}{\partial n} + qu_2$ on Γ.

In [40], it is proposed to choose coarse basis vectors $\mathbf{c}_1$, $\mathbf{c}_2$, ..., $\mathbf{c}_p$, build a matrix C with the $\mathbf{c}_j$'s as columns, and consider the coarse space defined by the span of these coarse basis vectors, $\mathcal{C} := \text{range}(C) = \text{span}(\mathbf{c}_1, ...\mathbf{c}_p)$. Then, at each iteration, we require that the error $\mathbf{e}^n = \boldsymbol{\mu} - \boldsymbol{\mu}^n$ be orthogonal to the coarse space $\mathcal{C}$, with respect to the inner product $\langle D\cdot, D\cdot\rangle$, i.e. we want

$$(DC)^T(D\mathbf{e}^n) = (DC)^T(\mathbf{b} - D\boldsymbol{\mu}^n) = 0. \tag{5.3}$$

Let $G := DC$, and Q be the projection into the orthogonal complement of $\mathcal{C}$,

$$Q := I - C(G^TG)^{-1}G^TD.$$

The interface vector $\boldsymbol{\lambda}$ is decomposed into a component in $\mathcal{C}$ and a component in the orthogonal complement (always with respect to $\langle D\cdot, D\cdot\rangle$),

$$\boldsymbol{\mu} = C\boldsymbol{\alpha}^0 + Q\tilde{\boldsymbol{\alpha}}.$$

The constraint (5.3) gives the equation

$$G^T G\boldsymbol{\alpha}^0 = G^T\mathbf{b}$$

to be solved first for α^0, and then the interface system $D\boldsymbol{\mu} = \mathbf{b}$, when projected onto the orthogonal complement of the coarse space, gives the linear system

$$QDQ\tilde{\boldsymbol{\alpha}} = Q(\mathbf{b} - DC\boldsymbol{\alpha}^0). \tag{5.4}$$

In [40], the projected interface system (5.4) is solved by applying a projected Generalized Conjugate Residual (GCR) method.

It remains to fix a coarse space. Two choices are proposed in [40]:

(1) Choose, for each interface, a coarse basis vector which is 1 on the interface and 0 elsewhere,

(2) Choose, for each interface, a coarse basis vector that corresponds to the transmission condition for the function $u_i \equiv 1$ on the subdomain Ω_i, and 0 for all other interfaces.

This strategy was shown to work well numerically for symmetric problems, but not so well for problems with strong tangential advection on the interface.

A similar coarse correction strategy is implemented in [50] for a diffusion problem with discontinuous coefficient, and Robin transmission conditions. The numerical experiments show that the method is scalable (as the number of subdomains grows) only when the rectangular domain is decomposed into strips. The convergence still seems to deteriorate for more general two-dimensional decompositions.

5.1.4 The Dual-Primal FETI Method

In this subsection, we follow once again the notation of [65]. In the dual-primal FETI method, the substructures are not torn apart completely as in the FETI method: the vector $\mathbf{u}$ of unknowns in each subdomain is required to be continuous (have a unique value) at the cross-points between subdomains. Cross-points are mesh points on the

interface that join together more than two subdomains. The degrees of freedom are separated into three sets:

I :=unknowns in the interior of subdomains,

Π :=unknowns at the cross-points (*primal degrees of freedom*),

Δ :=unknowns in the interior of edges and faces on the interface (*dual degrees of freedom*).

In this case the stiffness matrix $\tilde{A}$ can be partitioned as

$$\tilde{A} = \begin{bmatrix} A_{II} & A_{I\Pi} & A_{I\Delta} \\ A_{\Pi I} & A_{\Pi\Pi} & A_{\Pi\Delta} \\ A_{\Delta I} & A_{\Delta\Pi} & A_{\Delta\Delta} \end{bmatrix}.$$

Eliminating the degrees of freedom from the sets I and Π, we find a Schur complement $\tilde{S}$ which is *not* block diagonal as in the FETI method, because of the coupling of the subdomains at the cross-points. On the other hand, $\tilde{S}$ is invertible and there is no need to add a constraint for solvability. We are lead to consider the saddle-point formulation

$$\begin{aligned} \tilde{S}\mathbf{u}_\Delta + B_\Delta^T \boldsymbol{\lambda} &= \mathbf{f}_\Delta \\ B_\Delta \mathbf{u}_\Delta &= 0. \end{aligned}$$

As mentioned there is no problem in eliminating $\mathbf{u}_\Delta$ since $\tilde{S}$ is invertible, and we get the linear system for the Lagrange multipliers

$$F\boldsymbol{\lambda} = \mathbf{d}, \tag{5.5}$$

where $F := B_\Delta \tilde{S}^{-1} B_\Delta^T$ and $\mathbf{d} := B_\Delta \tilde{S}^{-1} \mathbf{f}_\Delta$. The dual-primal FETI method is a preconditioned Krylov subspace method applied to the linear system (5.5) (again, a preconditioner needs to be used in order to remove the dependence on the fine mesh size h, see [65]). The coarse component of this method is intrinsic to the application of F. Note that the matrix-vector multiplication $\tilde{S}^{-1}\mathbf{v}$ is equivalent to computing

$\tilde{A}^{-1}\tilde{\mathbf{v}}$ where $\tilde{\mathbf{v}}$ is an extension by zero. If we define the set of indices $D := I \cup \Delta$, then the stiffness matrix $\tilde{A}$ (after reordering the unknowns) can be factorized by

$$\tilde{A} = \begin{bmatrix} A_{DD} & A_{D\Pi} \\ A_{\Pi D} & A_{\Pi\Pi} \end{bmatrix} = \begin{bmatrix} I & 0 \\ A_{\Pi D}A_{DD}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{DD} & 0 \\ 0 & S_\Pi \end{bmatrix} \begin{bmatrix} I & A_{DD}^{-1}A_{D\Pi} \\ 0 & I \end{bmatrix}.$$

where S_Π is a Schur complement matrix obtained by eliminating the unknowns from the set D. The matrix A_{DD} is block diagonal and hence can be factorized locally in parallel for each subdomain. It can be shown that S_Π is a sparse matrix of small dimension, and that computing $S_\Pi^{-1}\mathbf{g}$ solves a global coarse problem.

5.2 Coarse Space Corrections for the Linear System $A\mathbf{u} = \mathbf{b}$

We consider now the simple one-dimensional problem

$$\begin{cases} -u'' + \eta u = f(x) & \text{for } x \in (0, \pi), \\ u(0) = u(\pi) = 0, \end{cases} \tag{5.6}$$

with $\eta \geq 0$. Suppose this problem is discretized using finite differences on the uniform grid $x_j = jh$, with $h = \frac{\pi}{N}$. This leads to the linear system

$$A\mathbf{u} = \mathbf{b}, \tag{5.7}$$

where the global matrix A, of dimension $N-1$ by $N-1$, is given by

$$A := \frac{1}{h^2} \begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & \end{bmatrix} + \eta I, \tag{5.8}$$

where I is identity matrix. The discrete solution $\mathbf{u}$ is represented in the form

$$\mathbf{u} = (u_1, u_2, \ldots, u_{N-1})^T, \quad \text{where } u_j \approx u(x_j),$$

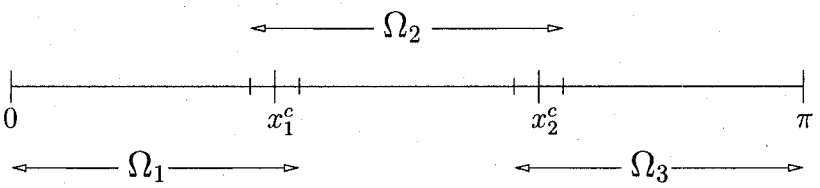

Figure 5.1: Decomposition of the interval into 3 overlapping subdomains.

and the right hand side vector is given by

$$\mathbf{b} = (b_1, b_2, \ldots, b_{N-1})^T, \quad \text{where } b_j = f(x_j).$$

Suppose the domain $\Omega = (0, \pi)$ is decomposed into M overlapping subdomains with an overlap of size L. First, define the coarse grid points $x_k^c := k\frac{\pi}{M}$. Then, the subdomains are chosen to be

$$\Omega_1 = \left(0,\ x_1^c + \frac{L}{2}\right), \quad \Omega_M = \left(x_{M-1}^c - \frac{L}{2},\ \pi\right),$$

$$\Omega_k = \left(x_{k-1}^c - \frac{L}{2},\ x_k^c + \frac{L}{2}\right), \quad k = 2, ..., M-1.$$

We assume that the subdomains are conforming with the grid, i.e. the endpoints of the subdomains are grid points. More precisely, we assume that N is a multiple of M and L is an even multiple of h.

On the subdomains, we can define local matrices A_k, which have the same entries as the matrix A, but are of smaller dimension $N_k - 1$ by $N_k - 1$, where $N_k - 1$ is the number of grid points in the *interior* of Ω_k. This local matrix A_k is the discretization of the problem on the interval Ω_k with *homogeneous* Dirichlet condition at both ends.

As in Section 5.1.1, the restriction matrix R_k simply retains only the unknowns at the grid points in the interior of the subdomain Ω_k. Then, the action of the transpose of this matrix, R_k^T, extends a vector by zeros. The additive Schwarz preconditioner is defined by

$$M_{as}^{-1} = \sum_{k=1}^{M} R_k^T A_k^{-1} R_k.$$

It is known that the condition number of the preconditioned matrix $M_{as}^{-1}A$ grows with the number of subdomains. To remedy this problem, we need to provide a mechanism for global communication across all the subdomains. The nodes $\{x_k^c\}_{k=1}^{M-1}$ forms a convenient coarse grid. We define the matrix P_c to be the linear interpolation from the coarse to the fine grid. An induced restriction operator, from the fine to the coarse grid, is $R_c := \frac{h}{H}P_c^T$; this restriction takes a weighted average of neighboring grid points. The coarse matrix can be defined from the prolongation and restriction operators, namely

$$A_c := R_c A P_c = \frac{1}{H^2}\begin{bmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \\ & & -1 & 2 \end{bmatrix} + \eta I\,, \tag{5.9}$$

which is a square matrix of dimension $M-1$. The two-level additive Schwarz preconditioner is then given by

$$M_{as2}^{-1} = P_c A_c^{-1} R_c + M_{as}^{-1}.$$

The application of this preconditioner involves solving homogeneous Dirichlet problems on the subdomains and a problem on the coarse grid; all of these problems can be solved in parallel.

The coarse correction can also be applied *sequentially* after having solved in parallèl the subdomain problems; this leads to a hybrid preconditioner (see [64]),

$$M_{hy}^{-1} = P_c A_c^{-1} R_c (I - A M_{as}^{-1}) + M_{as}^{-1}.$$

This modification improves the convergence, but makes the preconditioner no longer symmetric, whereas M_{as2}^{-1} is symmetric. In the numerical experiments, we will use the hybrid preconditioner M_{hy}^{-1}.

We test several methods, including the iterative realization of the Schwarz method without Krylov acceleration (i.e. running a stationary iterative method with the Schwarz preconditioner):

(i) Stationary iterative method with one-level additive Schwarz preconditioner:

$$\mathbf{u}^{n+1} = \mathbf{u}^n + M_{as}^{-1}(\mathbf{b} - A\mathbf{u}^n).$$

(ii) Stationary iterative method with two-level hybrid Schwarz preconditioner:

$$\mathbf{u}^{n+\frac{1}{2}} = \mathbf{u}^n + M_{as}^{-1}(\mathbf{b} - A\mathbf{u}^n),$$
$$\mathbf{u}^{n+1} = \mathbf{u}^{n+\frac{1}{2}} + P_c A_c^{-1} R_c(\mathbf{b} - A\mathbf{u}^{n+\frac{1}{2}}).$$

(iii) GMRES with one-level additive Schwarz preconditioner:

apply GMRES to the preconditioned linear system $M_{as}^{-1}A\mathbf{u} = M_{as}^{-1}\mathbf{b}$.

(iv) GMRES with two-level hybrid Schwarz preconditioner:

apply GMRES to the preconditioned linear system $M_{hy}^{-1}A\mathbf{u} = M_{hy}^{-1}\mathbf{b}$.

The convergence of the preconditioned stationary iterative methods is determined by the spectral radii of the matrices $I - M_{as}^{-1}A$ and $I - M_{hy}^{-1}A$ respectively. It is known that the stationary iterative method with one-level additive Schwarz preconditioner will not converge, since -1 is an eigenvalue of the matrix $I - M_{as}^{-1}A$, see [23] (this fact is illustrated in Figure 5.2). Intuitively, the problem is due to the fact that two different contributions are added in the overlapping regions.

To obtain convergence of the stationary iterative method (and to speed up convergence of the preconditioned GMRES iteration), we can modify the prolongation operators by replacing the matrices R_k^T with P_k, such that

$$\sum_{k=1}^{M} P_k R_k = I. \tag{5.10}$$

There are several ways to achieve this; here we consider three specific choices:

1. *Restricted prolongations*

   Choose the prolongation P_k that keeps only the entries in the non-overlapping

subdomain $[x_{k-1}^c, x_k^c)$, and then extend by zeros. This gives the *restricted additive Schwarz* preconditioner [9]. This preconditioner is extensively used nowadays. In particular, it significantly reduces the cost of communication: when merging approximations in an overlapping region, each processor needs from its neighbor his local approximation over *half* of the overlapping region (as opposed to the entire region in the additive Schwarz preconditioner).

2. *Partition of unity prolongations*
   Build P_k using a continuous partition of unity, i.e.

$$P_k \mathbf{u}_k := \text{diag}\left(\begin{bmatrix}\theta_k(x_1) & \theta_k(x_2) & \cdots & \theta_k(x_{N-1})\end{bmatrix}\right) R_k^T \mathbf{u}_k, \tag{5.11}$$

   where $\{\theta_k(x)\}_{k=1}^M$ is a (smooth) partition of unity, satisfying

$$\begin{aligned} \theta_k(x) &= 1 \quad \text{for } x \in [x_{k-1}^c, x_k^c], \\ \theta_k(x) &= 0 \quad \text{for } x \notin \Omega_k, \\ \sum_{k=1}^{M} \theta_k(x) &\equiv 1 \quad \text{for } x \in \overline{\Omega}. \end{aligned}$$

3. *Average prolongations*
   Choose the operators P_k such that $\sum_{k=1}^M P_k \mathbf{u}_k$ effectively takes the average of the two available approximations in overlapping regions, including the endpoints. We can define P_k by (5.11) with the choice of discontinuous partition of unity

$$\theta_k(x) := \begin{cases} \frac{1}{2} & \text{for } x \in \left[x_{k-1}^c - \frac{L}{2},\ x_{k-1}^c + \frac{L}{2}\right], \\ 1 & \text{for } x \in \left(x_{k-1}^c + \frac{L}{2},\ x_k^c - \frac{L}{2}\right), \\ \frac{1}{2} & \text{for } x \in \left[x_k^c - \frac{L}{2},\ x_k^c + \frac{L}{2}\right], \\ 0 & \text{elsewhere.} \end{cases}$$

Note that for these three choices, the property (5.10) holds. It then makes sense to use the preconditioners M_{as}^{-1} and M_{hy}^{-1} in a stationary iterative method.

We now show some numerical results when using these various methods for the problem (5.6) with $\eta = 1$ and $f(x) = \sin(x)$. We use a grid of $N = 1024$ points, and

No. of subdomains	2	4	8	16	32	64	128
	Stationary iterative method with M_{as}^{-1}						
restricted prolong.	35	97	351	1403	5581	21858	> 50000
partition of unity prolong.	35	98	377	1338	5513	22126	> 50000
average prolong.	70	185	726	2712	10759	43395	> 50000
	Preconditioned GMRES with M_{as}^{-1}						
additive prolong.	5	9	17	65	172	628	2079
restricted prolong.	4	8	16	64	197	634	2123
partition of unity prolong.	4	8	16	69	189	674	1468
average prolong.	4	8	16	71	162	588	2230

Table 5.1: Number of iterations for the one-level methods to reach a tolerance of 10^{-6}, for an increasing number of subdomains, and different strategies to combine two local approximations in the overlapping regions, when $\eta = 1$.

an overlap size proportional to the size of the subdomains, $L = \frac{\pi}{4M}$. The tolerance is set to 10^{-6}. We apply the GMRES method with a restart parameter of 25, and recall that the 2-norm of the relative residual in the preconditioned linear system is used for the stopping criterion. For the stationary iterative methods, we compute the ∞-norm of the difference between the approximation $\mathbf{u}^n$ and the discrete solution $\mathbf{u}$ of the global problem (5.7).

The iteration numbers are shown in Table 5.1 for the one-level Schwarz preconditioners, and in Table 5.2 for the two-level preconditioners. The one-level methods without the coarse grid component require a lot of iterations to reach the tolerance, even if the overlap size is large (it is a quarter of the subdomain size). When incorporating the coarse grid correction, the preconditioned GMRES iteration converges very quickly, and the number of iterations does not grow as we increase the number of subdomains, as expected from the theoretical analysis. Note that the stationary iterative method with two-level Schwarz preconditioner (using restricted or partition of unity prolongations) does not yield a convergence independent of the number of subdomains. However, when using the prolongations that combine by averaging, the

No. of subdomains	2	4	8	16	32	64	128
	Stationary iterative method with M_{hy}^{-1}						
restricted prolong.	28	98	281	839	1985	3230	4835
partition of unity prolong.	31	82	355	944	4356	16755	> 50000
average prolong.	12	10	10	10	10	10	10
	Preconditioned GMRES with M_{hy}^{-1}						
additive prolong.	6	9	9	9	9	9	8
restricted prolong.	4	7	8	8	9	9	8
partition of unity prolong.	4	7	8	8	8	8	8
average prolong.	4	7	7	7	7	7	7

Table 5.2: Number of iterations for the two-level methods to reach a tolerance of 10^{-6}, for an increasing number of subdomains, and different strategies to combine two local approximations in the overlapping regions.

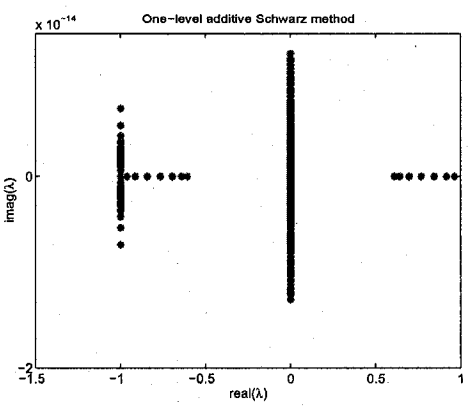

Figure 5.2: Spectrum of the iteration matrix $I - M_{as}^{-1}A$ for the one-level additive Schwarz method, in the case of 8 subdomains.

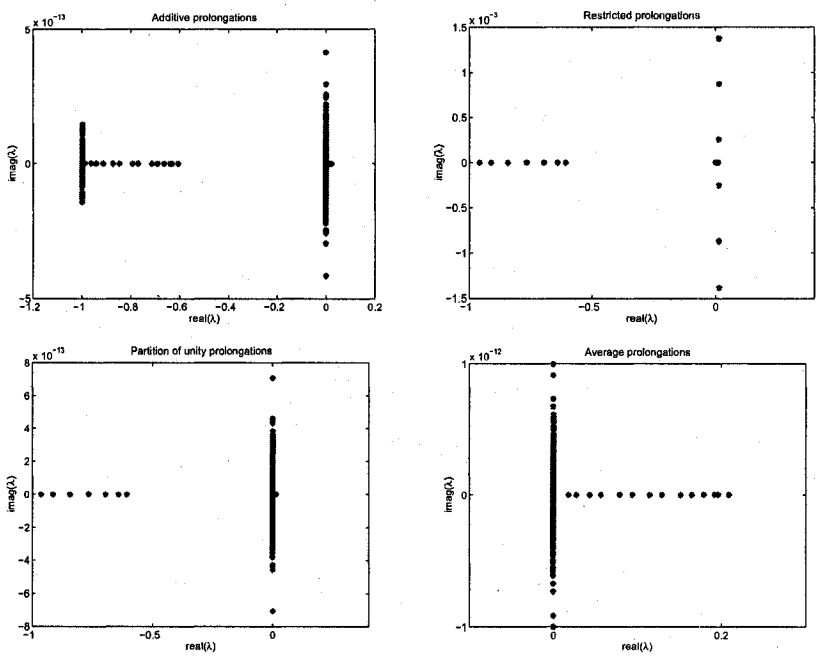

Figure 5.3: Spectrum of the iteration matrix $I - M_{hy}^{-1}A$ for the two-level additive Schwarz method, in the case of 8 subdomains, with different choices of prolongations.

stationary iterative method with two-level preconditioner converges after 10 iterations, and the convergence does not deteriorate with the number of subdomains. The average prolongations are also the choice that gives the fastest convergence of preconditioned GMRES iteration. Thus, taking an average to combine local approximations in the overlapping regions (including at the interfaces) appears to work extremely well with our choice of coarse correction. The spectrum of the iteration matrix $I - M_{hy}^{-1}A$ is shown in Figure 5.3, and indeed we observe that when taking averages in the overlap, we do not get negative eigenvalues close to -1. Note that the preconditioned matrix $M_{hy}^{-1}A$ appears to have *real* eigenvalues, except in the case of restricted prolongations.

We can extend these methods to the case of Robin transmission conditions as follows. Consider a Schwarz iteration with, in subdomain Ω_k, the transmission conditions

$$\begin{aligned}
\left(-\frac{\mathrm{d}}{\mathrm{d}x} + q_k\right) u_k^{n+1} &= \left(-\frac{\mathrm{d}}{\mathrm{d}x} + q_k\right) u_{k-1}^{n} \quad \text{at } x = x_{k-1}^c - \tfrac{L}{2}, \\
\left(\frac{\mathrm{d}}{\mathrm{d}x} + p_k\right) u_k^{n+1} &= \left(\frac{\mathrm{d}}{\mathrm{d}x} + p_k\right) u_{k+1}^{n} \quad \text{at } x = x_k^c + \tfrac{L}{2}.
\end{aligned}$$

Let $\hat{A}_k$ be the local matrix obtained by the discretization of the local subproblem with the Robin boundary conditions. Let $\hat{R}_j$ be the restriction matrix such that, given a global right hand side vector $\mathbf{f}$, the restriction $\hat{R}_j\mathbf{f}$ produces the appropriate local right hand side vector for solving the subproblem in Ω_k with zero Robin data ($u_{k-1}^n = u_{k+1}^n \equiv 0$). The matrix $\hat{R}_j$ differs from R_j in the fact that the source term $f(x)$ may be involved in the discretization of the Robin boundary condition. Using these modifications, we can define analogous Schwarz preconditioners

$$\begin{aligned}
\hat{M}_{as}^{-1} &= \sum_{k=1}^{M} P_k \hat{A}_k^{-1} \hat{R}_k, \\
\hat{M}_{hy}^{-1} &= P_c A_c^{-1} R_c (I - A\hat{M}_{as}^{-1}) + \hat{M}_{as}^{-1}.
\end{aligned}$$

In the following numerical results, we consider two choices of Robin transmission conditions: first the case of a constant Robin parameter $p_k = q_k = p = 1$, and secondly the case of optimal Robin conditions, given by

$$p_k^{opt} := \frac{\sqrt{\eta}}{\tanh(\sqrt{\eta}(\pi - x_k^c - \frac{L}{2}))}, \quad q_k^{opt} := \frac{\sqrt{\eta}}{\tanh(\sqrt{\eta}(x_{k-1}^c - \frac{L}{2} - 0))},$$

when $\eta > 0$ (these formulas for the optimal Robin parameters are obtained through straightforward direct calculations).

We use the same setting as for the experiments shown above, namely we consider the problem (5.6) with $\eta = 1$ and $f(x) = \sin(x)$, use a grid of $N = 1024$ points, and an overlap proportional to the size of the subdomains, $L = \frac{\pi}{4M}$. Table 5.3 contains the iteration numbers for this adapted method with Robin conditions, when the restricted prolongations are used. Note in this case that the stationary iterative method with two-level Schwarz preconditioner diverges iteratively for 8 and 16 subdomains. The results for the preconditioned GMRES method indicate that the convergence deteriorates as the number of subdomains grows; the coarse grid correction does not seem to fulfill its purpose. The exact same behavior is found when using the partition of unity prolongations. In the case of average prolongations, the stationary iterative method, without and with the coarse component, diverges in all cases when Robin conditions are used. So, in Table 5.4 we include only the iteration numbers for the preconditioned GMRES method. The convergence also suffers from a deterioration when increasing the number of subdomains, even for the two-level method. Thus, this strategy of coarse space correction is not effective when using Robin transmission conditions, it cannot compete with the results of Table 5.2.

5.3 Coarse Space Corrections for an Augmented Linear System $\tilde{A}\tilde{\mathbf{u}} = \tilde{\mathbf{b}}$

To avoid merging the local subdomain approximations in the overlapping regions (or on the interface when there is no overlap), a different strategy is to first rewrite the global linear system $A\mathbf{u} = \mathbf{b}$ by doubling the unknowns in the overlap to obtain an *augmented* linear system $\tilde{A}\tilde{\mathbf{u}} = \tilde{\mathbf{b}}$ which now depends on the transmission conditions to be used in the Schwarz iteration.

Let us consider the problem

$$\begin{cases} -u'' &= f(x) & \text{for } x \in (0,\pi), \\ u(0) &= u(\pi) = 0. \end{cases} \tag{5.12}$$

No. of subdomains	2	4	8	16
	Preconditioned stationary iterative method			
$p = 1$, one-level	8	21	45	96
$p = 1$, two-level	10	19	D	D
optimal Robin, one-level	4	7	14	25
optimal Robin, two-level	10	16	D	D
	Preconditioned GMRES method			
$p = 1$, one-level	4	8	16	96
$p = 1$, two-level	4	8	13	22
optimal Robin, one-level	4	6	14	23
optimal Robin, two-level	4	8	12	20

Table 5.3: Number of iterations to reach a tolerance of 10^{-6}, for an increasing number of subdomains, when using ***restricted*** prolongations to combine to local approximations in the overlapping regions.

No. of subdomains	2	4	8	16
	Preconditioned GMRES method			
$p = 1$, one-level	4	8	16	46
$p = 1$, two-level	4	8	16	26
optimal Robin, one-level	4	8	16	49
optimal Robin, two-level	4	8	16	32

Table 5.4: Number of iterations to reach a tolerance of 10^{-6}, for an increasing number of subdomains, when using ***average*** prolongations to combine to local approximations in the overlapping regions.

We use the same notation for the subdomains Ω_k as in Section 5.2. Let $\mathbf{u}_k$ denote a vector of unknowns in $\overline{\Omega}_k$, including the endpoints; this vector is of length $N_k + 1$. We will solve for an augmented vector

$$\tilde{\mathbf{u}} := (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_M)^T .$$

Define l_k to be the global index in the augmented vector $\tilde{\mathbf{u}}$ of the the first entry of $\tilde{\mathbf{u}}_k$. Similarly, define r_k to be the global index of the the last entry of $\tilde{\mathbf{u}}_k$

A transmission problem (or multi-domain formulation) for the continuous problem (5.12) can be written as

$$\left\{\begin{array}{rcll} u_1(0) & = & 0, & \\ -u_1'' & = & f & \text{in } \Omega_1, \\ \mathcal{B}_1^r(u_1) - \mathcal{B}_1^r(u_2) & = & 0 & \text{at } x = x_1^c + \frac{L}{2}, \\ \mathcal{B}_2^l(u_2) - \mathcal{B}_2^l(u_1) & = & 0 & \text{at } x = x_1^c - \frac{L}{2}, \\ -u_2'' & = & f & \text{in } \Omega_2, \\ \mathcal{B}_2^r(u_2) - \mathcal{B}_2^r(u_3) & = & 0 & \text{at } x = x_2^c + \frac{L}{2}, \\ \mathcal{B}_3^l(u_3) - \mathcal{B}_3^l(u_2) & = & 0 & \text{at } x = x_2^c - \frac{L}{2}, \\ -u_3'' & = & f & \text{in } \Omega_3, \\ \vdots & \vdots & \vdots & \vdots \\ -u_M'' & = & f & \text{in } \Omega_M, \\ u_M(\pi) & = & 0, & \end{array}\right\} \tag{5.13}$$

where $\mathcal{B}_k^{r,l}$ are transmission operators at the interfaces. Then, the augmented system $\tilde{A}\tilde{\mathbf{u}} = \tilde{\mathbf{b}}$ is the corresponding discrete version of the transmission problem (5.13) (respecting the particular ordering of the equations), which depends on the choice of operators $\mathcal{B}_k^{r,l}$.

For instance, in the overlapping case $L > 0$, we can choose Dirichlet transmission conditions, $\mathcal{B}_k^{r,l}(u) = u$. Suppose there are two subdomains and $L = 2h$, then the

augmented matrix is of the form

$$\tilde{A} = \begin{bmatrix} \frac{2}{h^2} & \frac{-1}{h^2} & & & & & & \\ \frac{-1}{h^2} & \frac{2}{h^2} & \frac{-1}{h^2} & & & & & \\ & \frac{-1}{h^2} & \frac{2}{h^2} & \frac{-1}{h^2} & & & & \\ & & & 1 & & -1 & & \\ & & -1 & & 1 & & & \\ & & & & \frac{-1}{h^2} & \frac{2}{h^2} & \frac{-1}{h^2} & \\ & & & & & \frac{-1}{h^2} & \frac{2}{h^2} & \frac{-1}{h^2} \\ & & & & & & \frac{-1}{h^2} & \frac{2}{h^2} \end{bmatrix}$$

In general, for one-dimensional problems, the discretization of the transmission problem (5.13) leads to an augmented matrix that has the tridiagonal block structure

$$\tilde{A} = \begin{bmatrix} A_1 & B_{12} & & & \\ B_{21} & A_2 & B_{23} & & \\ & B_{32} & A_3 & B_{34} & \\ & & \ddots & \ddots & \ddots \\ & & & B_{M,M-1} & A_M \end{bmatrix},$$

where A_k is the matrix discretizing the local problem on $\overline{\Omega}_k$ with boundary conditions given by the operators $\mathcal{B}_k^{r,l}$. The role of the off-diagonal matrices B_{ij} is to extract the transmission condition $-\mathcal{B}_i^{r,l}(u_j)$ from the neighboring subdomain.

Consider the Schwarz iteration corresponding to the transmission problem (5.13),

$$\begin{cases} -\dfrac{\mathrm{d}^2 u_k^{n+1}}{\mathrm{d}x^2} &= f & \text{in } \Omega_k, \\ \mathcal{B}_k^l(u_k^{n+1}) &= \mathcal{B}_k^l(u_{k-1}^n) & \text{at } x = x_{k-1}^c - \frac{L}{2}, \\ \mathcal{B}_k^r(u_k^{n+1}) &= \mathcal{B}_k^r(u_{k+1}^n) & \text{at } x = x_k^c + \frac{L}{2}, \end{cases} \tag{5.14}$$

where we define $\mathcal{B}_1^l = \mathcal{B}_M^r = I$ to match the boundary conditions of the global problem. Then, the discretized version of this Schwarz iteration can be written as a block-Jacobi iteration for the augmented system,

$$\tilde{\mathbf{u}}^{n+1} = \tilde{\mathbf{u}}^n + M_{BJ}^{-1}(\tilde{\mathbf{b}} - \tilde{A}\tilde{\mathbf{u}}^n),$$

where

$$M_{BJ}^{-1} := \begin{bmatrix} A_1^{-1} & & & \\ & A_2^{-1} & & \\ & & \ddots & \\ & & & A_M^{-1} \end{bmatrix}.$$

Suppose the augmented vector $\tilde{\mathbf{u}}$ has length $\tilde{N}$, then our fine space is

$$\tilde{\mathbf{u}} \in \mathcal{F} := \mathbb{R}^{\tilde{N}}.$$

To define a coarse space correction for the augmented system, we need the following ingredients:

- a coarse subspace $\mathcal{C} \subset \mathcal{F}$,
- a prolongation operator, $P_c : \mathcal{C} \to \mathcal{F}$,
- a restriction operator, $R_c : \mathcal{F} \to \mathcal{C}$,
- a matrix $\tilde{A}_c$ defining the problem on the coarse space.

Remark 5.1. *For the simple problem $-u'' = 0$, $u(0) = u(\pi) = 0$, after one iteration of the Schwarz method, the approximation $\tilde{\mathbf{u}}^1$ is piecewise linear on the subdomains and vanishes at the endpoints. Call the set of such functions $\mathcal{C}_l$. Every subsequent iteration of the Schwarz method yields approximations in that space, which has dimension $2(M-1)$. If we choose our coarse space to be $\mathcal{C}_l$, then after one coarse correction step, we will have obtained the exact solution of the problem. Note that the two choices of coarse spaces proposed by Japhet and described in Section 5.1.3 reduce to $\mathcal{C}_l$ for this one-dimensional problem. We wish to study coarse spaces $\mathcal{C}$ that have smaller dimension than $\mathcal{C}_l$.*

As the simplest choice, we introduce the coarse space $\mathcal{C}$ of functions that are piecewise constant on the subdomains Ω_2, Ω_3, ..., Ω_{M-1}, and 0 on the subdomains Ω_1 and Ω_M, because of the homogeneous Dirichlet boundary conditions at the endpoints, i.e.

$$\mathbf{u}_1 \equiv \mathbf{u}_M \equiv 0, \qquad \mathbf{u}_k \equiv c_k, \quad k = 2, 3, ..., M-1.$$

We represent these coarse functions by vectors of length $M-2$,

$$\mathbf{u}^{(c)} \in \mathcal{C} \quad \Rightarrow \quad \mathbf{u}^{(c)} = (c_2, c_3, \ldots, c_{M-1})^T.$$

Remark 5.2. *If we consider problem* (5.12) *with a Neumann boundary condition at* $x = \pi$ *instead,* $u'(\pi) = 0$, *then we would allow the coarse functions to be equal to a general constant* c_M *on* Ω_M.

For this choice of coarse space, the natural prolongation is defined by

$$P_c \begin{bmatrix} c_2 \\ c_3 \\ \vdots \\ c_{M-1} \end{bmatrix} := \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_M \end{bmatrix},$$

where

$$\mathbf{u}_1 \equiv \mathbf{u}_M \equiv 0, \qquad \mathbf{u}_k \equiv c_k, \quad k = 2, 3, ..., M-1.$$

The restriction R_c can be defined as the operator that computes averages in the subdomains Ω_2, Ω_3, ... Ω_{M-1} while ignoring the subdomains Ω_1 and Ω_M,

$$R_c \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_M \end{bmatrix} := \begin{bmatrix} c_2 \\ c_3 \\ \vdots \\ c_{M-1} \end{bmatrix}, \quad \text{where } c_k = \text{average}(\mathbf{u}_k) \quad \text{for } k = 2, 3, ..., M-1.$$

Given a prolongation and a restriction operator, a problem matrix on the coarse space is induced by the formula

$$\tilde{A}_c := R_c \tilde{A} P_c.$$

Let us consider a general Schwarz iteration with Robin transmission conditions, corresponding to the choice of operators

$$\begin{aligned} \mathcal{B}_k^l(u) &= \left(-\frac{\mathrm{d}}{\mathrm{d}x} + q_k\right) u, \qquad k = 2, ..., M, \\ \mathcal{B}_k^r(u) &= \left(\frac{\mathrm{d}}{\mathrm{d}x} + p_k\right) u. \qquad k = 1, ..., M-1. \end{aligned}$$

In this case, we can explicitly derive the coarse matrix $\tilde{A}_c$ with simple calculations, and we find

$$\tilde{A}_c := R_c \tilde{A} P_c = \frac{1}{N_{sub}} \begin{bmatrix} (q_2+p_2) & -p_2 & & & & \\ -q_3 & (q_3+p_3) & -p_3 & & & \\ & -q_4 & (q_4+p_4) & -p_4 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -q_{M-1} & (q_{M-1}+p_{M-1}) \end{bmatrix},$$

where N_{sub} is the number of grid points in $\overline{\Omega}_k$, $k = 2, ..., M-1$. Although this coarse matrix $\tilde{A}_c$ vaguely resembles a discrete Laplacian, it looks significantly different compared to the coarse matrix A_c obtained in (5.9). It becomes symmetric only when $p_j = q_{j+1}$ for $j = 2, ..., M-2$, i.e. when using one-sided Robin transmission conditions between neighboring subdomains.

We can now define the algorithms we propose:

(i) Stationary iterative method with one-level Schwarz preconditioner for the augmented system:

$$\tilde{\mathbf{u}}^{n+1} = \tilde{\mathbf{u}}^n + M_{BJ}^{-1}(\tilde{\mathbf{b}} - \tilde{A}\tilde{\mathbf{u}}^n).$$

(ii) Stationary iterative method with two-level Schwarz preconditioner for the augmented system:

$$\tilde{\mathbf{u}}^{n+\frac{1}{2}} = \tilde{\mathbf{u}}^n + M_{BJ}^{-1}(\tilde{\mathbf{b}} - \tilde{A}\tilde{\mathbf{u}}^n),$$
$$\tilde{\mathbf{u}}^{n+1} = \tilde{\mathbf{u}}^{n+\frac{1}{2}} + P_c \tilde{A}_c^{-1} R_c(\tilde{\mathbf{b}} - A\tilde{\mathbf{u}}^{n+\frac{1}{2}}).$$

(iii) GMRES method with one-level Schwarz preconditioner:

apply GMRES to the preconditioned linear system $M_{BJ}^{-1}\tilde{A}\tilde{\mathbf{u}} = M_{BJ}^{-1}\tilde{\mathbf{b}}$.

(iv) GMRES method with two-level Schwarz preconditioner:

apply GMRES to the preconditioned linear system $M_{hy}^{-1}\tilde{A}\tilde{\mathbf{u}} = M_{hy}^{-1}\tilde{\mathbf{b}}$,

where

$$M_{hy}^{-1} := P_c \tilde{A}_c^{-1} R_c (I - \tilde{A} M_{BJ}^{-1}) + M_{BJ}^{-1}.$$

In addition to avoid the merging of subdomain solutions, another advantage of writing the Schwarz iteration using the augmented system formulation is that both the overlapping and non-overlapping cases are treated exactly the same way.

In the overlapping case, if we choose Dirichlet transmission operators, we find that the coarse matrix is given by

$$\tilde{A}_c := R_c \tilde{A} P_c = \frac{1}{N_{sub}} \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & \ddots & \ddots & \ddots \\ & & & -1 & 2 \end{bmatrix}.$$

Using the same ideas, we can also define a two-level Schwarz method for the augmented system in other situations, such as:

- for problems with discontinuous coefficients, e.g.

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(\nu(x)\frac{\mathrm{d}u}{\mathrm{d}x}\right) = f,$$

  with piecewise constant diffusion coefficient $\nu(x)$ and non-overlapping domain decomposition,

- for the positive definite Helmholtz equation $-u'' + \eta u = f$ (in this case the induced coarse matrix $\tilde{A}_c$ is quite different),

- by using a coarse space $\mathcal{C}$ of piecewise linear functions that vanish on the boundary (with appropriate choices of prolongations and restrictions).

5.3.1 Numerical Results

We first consider the problem (5.12) with $f(x) = \sin(x)$. We employ a finite difference scheme with $N = 512$ grid points. Again we check the error against a tolerance of 10^{-6}. We test the effectiveness of the coarse space correction for the augmented system using different choices of transmission conditions:

- Dirichlet conditions, corresponding to the classical Schwarz iteration,
- Robin conditions with constant Robin parameter,

$$p_k = q_k = p \quad \forall\, k,$$

- optimal Robin conditions,

$$p_k^{opt} = \frac{1}{1 - (x_k^c + \frac{L}{2})}, \quad q_k^{opt} = \frac{1}{(x_{k-1}^c - \frac{L}{2}) - 0}.$$

Figure 5.4 shows the convergence behavior as the number of subdomains increases (notice the different scales for the y axes). First, for the Dirichlet conditions, the stationary iterative method with one-level preconditioner reaches the maximum number of iterations that we allowed (5000), so the upper left graph in Figure 5.4 does not indicate that the convergence is independent of M. Now, for the Dirichlet conditions and constant Robin conditions, the preconditioned GMRES method with the two-level Schwarz preconditioner works very well, it converges in 5 iterations and the convergence does not deteriorate for larger M.

However, when using the optimal Robin conditions, although the methods (stationary and GMRES) with the one-level preconditioner converge after exactly M iterations (the best possible), the coarse space correction is not effective: the convergence is actually worse for the stationary iterative method. When running the preconditioned GMRES method, the convergence also deteriorates as M grows. This illustrates an important point: the transmission conditions that give fast convergence for the one-level method may not necessarily give fast convergence when adding a coarse correction. This is also seen by comparing the performance of the Robin transmission conditions with $p = 1$ and $p = 10$. The iteration numbers for the stationary iterative method are shown in Table 5.5. The Robin parameter $p = 1$ gives significantly faster convergence for the one-level method, whereas $p = 10$ is a much better parameter value for the two-level method.

To see even more clearly that the best Robin transmission conditions are very different for one-level and two-level methods, we ran the stationary iterative method

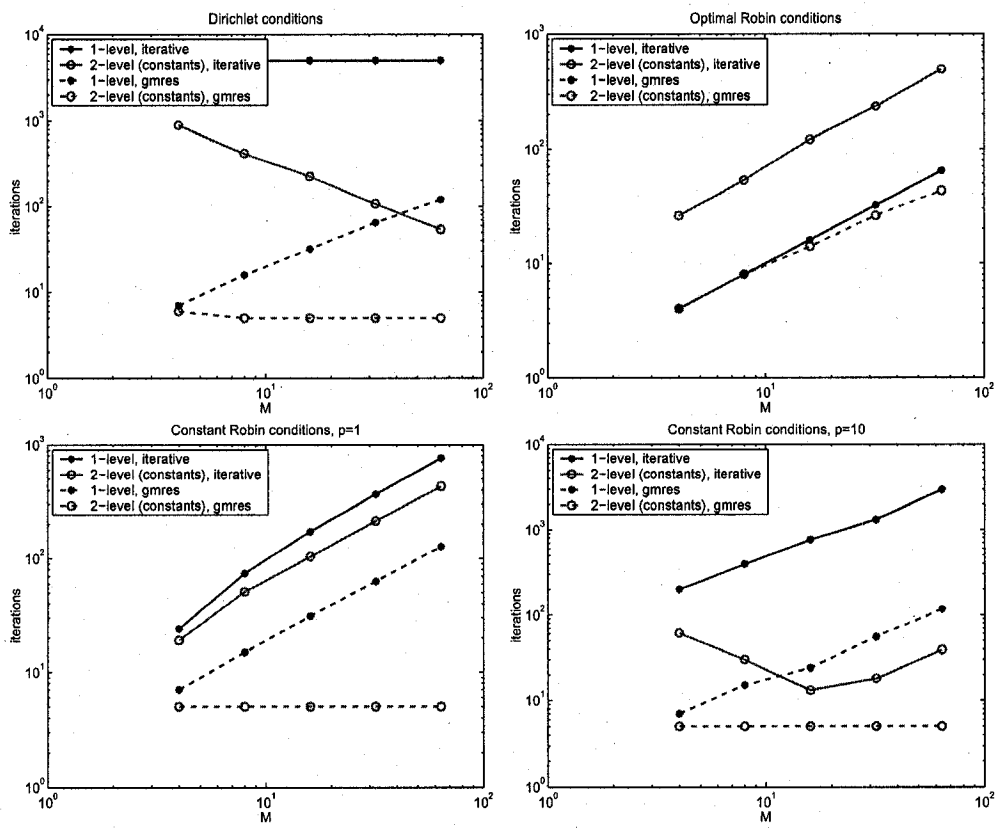

Figure 5.4: Number of iterations requires to reach a tolerance of 10^{-6} as the number of subdomains grows, in the overlapping case, with $L = 2h$.

No. of subdomains	4	8	16	32	64
1-level Schwarz, $p = 1$	24	74	172	366	763
1-level Schwarz, $p = 10$	200	394	761	1305	2929
2-level Schwarz, $p = 1$	19	51	104	213	430
2-level Schwarz, $p = 10$	61	30	13	18	39

Table 5.5: Number of iterations to reach a tolerance of 10^{-6}, for an increasing number of subdomains, for the stationary iterative method using the Schwarz preconditioners with Robin conditions.

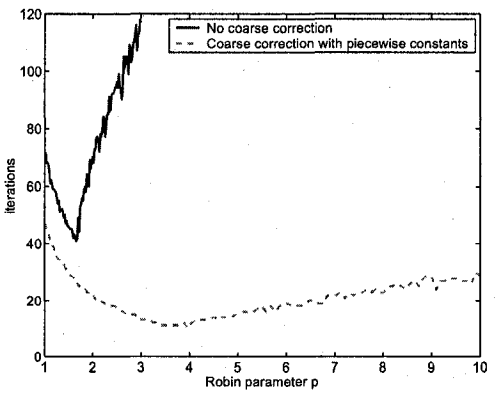

Figure 5.5: Convergence for different values of the Robin parameter p, in the case of 8 non-overlapping subdomains, and $N = 256$ grid points.

for many choices of p and found numerically the best value. For this, we used $N = 256$ grid points and no overlap. For the case of $M = 8$ subdomains, we get the curves shown in Figure 5.5, from which we can see the best parameters for the method with and without coarse corrections. For various numbers of the subdomains, the best parameter values along with their induced iteration number are presented in Table 5.6. The best p value for the one-level Schwarz preconditioner is close to 2 and does not significantly grow for an increasing number of subdomains. However, the resulting convergence deteriorates with M: the iteration number is doubled when M is doubled. On the other hand, the best p value for the two-level Schwarz method grow proportionally to M, $p^* \approx M/2$, and the resulting convergence seems to actually get better!

Comparison with preconditioners for $A\mathbf{u} = \mathbf{b}$

If we apply the two-level restricted additive Schwarz preconditioner, introduced in Section 5.2, for the same setup that was used for Figure 5.4 (i.e. $f(x) = \sin(x)$, $N = 512$ and $L = 2h$), we obtain the iteration numbers shown in Table 5.7. When we compare the results with the two-level preconditioners that we derived for the augmented system $\tilde{A}\tilde{\mathbf{u}} = \tilde{\mathbf{b}}$, we observe in particular that even by using Dirichlet

No. of subdomains	8	16	32	64
Best p for 1-level Schwarz	1.66	1.82	1.9	1.94
Best p for 2-level Schwarz	3.5	7.5	15	31
Iteration number for best p, 1-level	40	90	179	362
Iteration number for best p, 2-level	11	10	10	9

Table 5.6: Best value of the Robin parameter p for the one-level and two-level Schwarz preconditioner when used in a stationary iterative method, and their corresponding iteration numbers.

No. of subdomains	4	8	16	32	64
	Stationary iterative method				
Restricted additive Schwarz for $A\mathbf{u} = \mathbf{b}$	2875	5070	8894	9148	3574
Dirichlet conditions, for $\tilde{A}\tilde{\mathbf{u}} = \tilde{\mathbf{b}}$	890	417	223	107	54
Robin conditions, $p = 10$, for $\tilde{A}\tilde{\mathbf{u}} = \tilde{\mathbf{b}}$	61	30	13	18	39
	GMRES method				
Restricted additive Schwarz for $A\mathbf{u} = \mathbf{b}$	5	5	6	7	7
Dirichlet conditions, for $\tilde{A}\tilde{\mathbf{u}} = \tilde{\mathbf{b}}$	6	5	5	5	5
Robin conditions, $p = 10$, for $\tilde{A}\tilde{\mathbf{u}} = \tilde{\mathbf{b}}$	5	5	5	5	5

Table 5.7: Comparison of the two-level restricted additive Schwarz preconditioner with two-level preconditioners for the augmented system.

transmission conditions our proposed coarse correction gives better results, although the difference is much less apparent when GMRES is used.

Recall that in the two-level restricted additive Schwarz preconditioner, the coarse space consists of continuous piecewise linear functions, and our coarse space for the augmented system uses piecewise constant functions; the comparison is fair since both of these coarse spaces have almost the same dimension ($M - 1$ versus $M - 2$).

Positive definite Helmholtz equation

We consider now the positive definite Helmholtz problem

$$\begin{cases} -u'' + u &= \sin(x) \qquad \text{for } x \in (0, \pi), \\ u(0) &= u(\pi) = 0. \end{cases} \tag{5.15}$$

Suppose we use the same setup as for the results of Figure 5.4, and in addition we also experiment with a coarse space consisting of piecewise linear functions on each subdomain that vanish at the endpoints $x = 0$ and $x = \pi$ (this coarse space has dimension $2(M-1)$). The results are shown in Figure 5.6. We observe that the coarse correction with piecewise constant functions does not work, but when using piecewise linear functions, the number of iterations seems to stop growing when using Robin transmission conditions. The reason why our coarse corrections are not as effective as for the problem $-u'' = f(x)$ is explained as follows: the coarse functions are *not* solutions of the equation in the subdomains, $-u'' + u \neq 0$ for constant or linear functions. After the parallel step in which we solve in the subdomains, our approximation satisfies the equation $-u''+u = f(x)$ in the interior of each subdomain, but after correcting with constants or linears, we destroy this property and this may hinder the convergence.

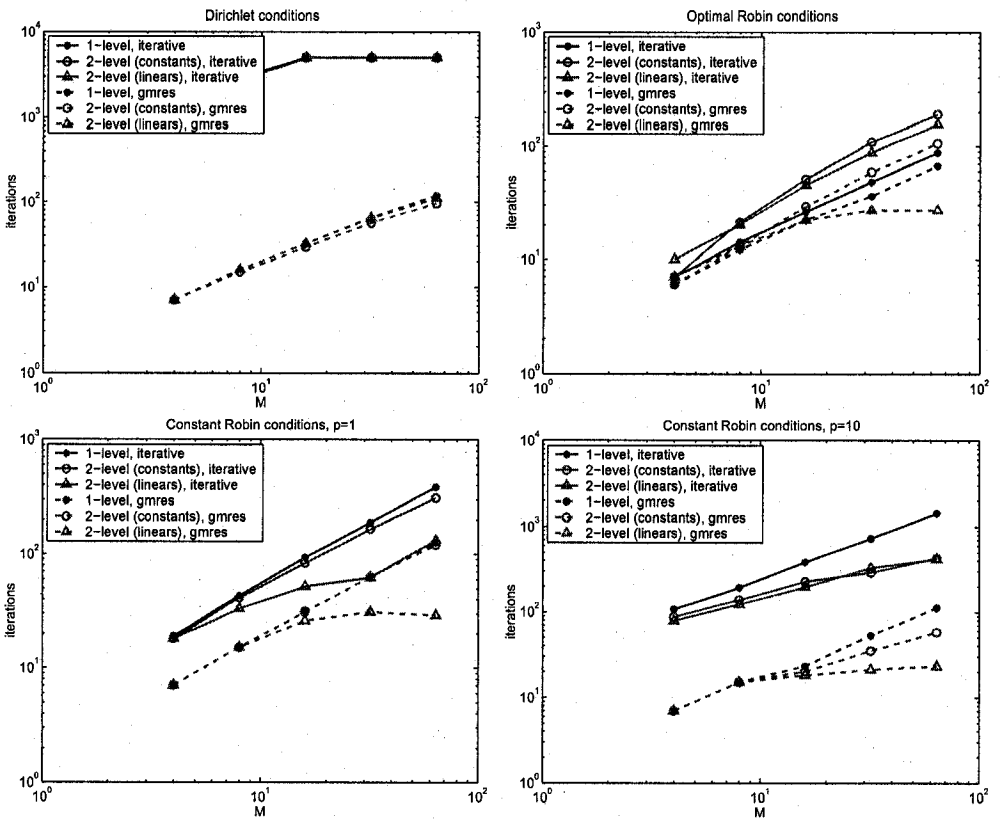

Figure 5.6: Number of iterations required to reach a tolerance of 10^{-6} as the number of subdomains grows.

5.4 Summary

We assemble here the main observations that we have made in this chapter, using numerical experiments with different coarse space corrections for a one-dimensional problem:

(1) When using *average prolongations* in a two-level additive Schwarz preconditioner (using Dirichlet transmission conditions), the preconditioned stationary iterative method converges surprisingly fast and the convergence is independent of the number of subdomains. When using other types of prolongations, the stationary iterative methods converges very slowly.

(2) A natural extension of the two-level additive Schwarz preconditioner to the case of Robin transmission conditions does not give a convergence independent of the number of subdomains, even under Krylov acceleration.

(3) Using an augmented system formulation, we designed a coarse space correction which is effective for the problem $-u'' = f$ when using Robin transmission conditions with a constant Robin parameter across all the subdomains. The same coarse space correction does not work for the optimal Robin conditions.

(4) The best Robin parameter for the one-level Schwarz iteration is very different from the best Robin parameter for the two-level Schwarz iteration.

We believe that the observations (2) and (4) described above also extend to problems in two or three dimensions.

Note that for a two-dimensional problem with a domain decomposition in strips and homogeneous Dirichlet conditions imposed on the whole boundary, a coarse space correction strategy using piecewise constant coarse basis functions (as in Section 5.3) is not appropriate, since every subdomain touches the boundary. In such a situation, we would need to design non-trivial coarse basis functions (e.g. using piecewise linear functions over each subdomans) that satisfy the boundary conditions.

Conclusion

For the advection-diffusion equation with smooth coefficients, we have derived asymptotic formulas for the optimized transmission conditions, valid for small mesh sizes. These formulas can be used in practice, but keeping in mind their limitations. Optimized two-sided Robin conditions and second order conditions significantly improve the convergence for both overlapping and non-overlapping domain decompositions. On this topic, it would be useful in future work to find formulas for the transmission conditions that are more accurate for strong advection tangential to the interface.

For the case of a diffusion problem with a discontinuous coefficient, we derived explicit formulas for the optimized parameters by fully solving the associated min-max problems, for several choices of transmission conditions. The analysis shows that the convergence of optimized Schwarz methods improves as the jump in the coefficient increases. This is a very desirable feature since, in many applications, the jumps in the coefficients can be very large. In addition, for optimized two-sided Robin transmission conditions, the convergence of the Schwarz method is mesh independent; this is a novel result because in the case of continuous coefficients, the convergence of non-overlapping optimized Schwarz methods always deteriorates for small mesh sizes (even if very slowly). As a next step for future consideration, these results need to be extended to the more general case of an advection-diffusion equation with discontinuous coefficients. In that case, jumps in the coefficients can cause boundary layers near the interface, thus imposing a restriction on the mesh size; it would be of interest to study the asymptotic convergence properties of optimized Schwarz methods in such a scenario.

Recall that using optimized transmission conditions in practice requires only minor modifications in existing implementations of domain decomposition methods, and that the greatly improved convergence comes without any additional computational cost per iteration. Also, note that optimized parameter values and the convergence analysis were studied for a model problem only, and the numerical experiments we presented involved rectangular subdomains with straight interfaces. However, in practice, by using the explicit formulas we derived for the optimized parameters when computing on subdomains of more general shapes, with curved interfaces, we still expect to obtain fast convergence of the Schwarz iteration. This should be demonstrated in future work with numerical experiments on non-rectangular subdomains.

Finally, for a one-dimensional problem, we proposed a two-level Schwarz method with Robin transmission conditions, using a coarse space correction for an augmented system. This strategy allows for both overlapping and non-overlapping domain decompositions. An important observation is made from the numerical results: the best transmission conditions for the one-level and two-level methods are very different, thus one should optimize the parameters by taking into account the coarse space correction. These results motivate further research:

- a theoretical analysis should be carried out to prove that the convergence is independent of the number of subdomains,
- similar coarse space corrections should be designed for two and three dimensional problems,
- methods need to be formulated for optimizing the transmission conditions as part of a two-level method.

Appendix A

Matlab Code to Compute Optimized Parameters

In this appendix, we include a few simple Matlab functions that were used to compute optimized parameters by numerically solving the corresponding min-max problem, when no formula is available. To make the code more elegant, we employ the feature of anonymous functions available in Matlab version 7. The maximum of the convergence factor is computed using a fine discrete grid of values of k, and the Nelder-Mead algorithm [57] is then applied to find the minimum. From our experience, this method seems to work well enough when optimizing on 1 or 2 free parameters, but not as well when considering 3 or more independent parameters.

A.1 Advection-Diffusion Equation

The following function computes the optimized parameters for Robin, two-sided Robin and second order transmission conditions for the advection-diffusion equation with constant coefficients. Note in particular that for computing the maximum of the convergence factor on the interval $[k_1, k_2]$, we sample the frequencies uniformly on a *logarithmic* scale; this captures better the interior local maxima since they are growing slower than $k_2 = \frac{\pi}{h}$ for h small.

```
function [Robin1,Robin2,Second]=OptimizedParameters_ADR(nu,a,b,c,L,k1,k2);
% [Robin1,Robin2,2ndorder]=OptimizedParameters_ADR(nu,a,b,c,L,k1,k2);
% Computes optimized parameters for the advection-diffusion equation
%     -nu Delta(u) + (a,b).grad(u) + c u
% by solving numerically the min-max problem
% with the frequency range [k1,k2], and overlap L.
% Robin1:  one-sided Robin parameter,   p
% Robin2:  two-sided Robin parameters, [p1; p2]
% Second:  second order parameters,    [ p;  q]

h = pi / k2;
N = 10000; % number of sampled frequencies

% uniform grid for k in [0,1]
if k1==0, kstart = 0:0.01:1; k1 = 1;
else,     kstart = [];  end;
% uniform grid on a logarithmic scale
k = log(k1):(log(k2)-log(k1))/N:log(k2);
k = exp(k); k = [kstart k];

% Optimized Robin parameter
F = @(p) max( ConvergenceFactor(k,[p 0;p 0],nu,a,b,c,L) );
[Robin1,fval,flag] = fminsearch(F,1);
if flag~=1, warning('fminsearch did not converge.'); end;

% Optimized two-sided Robin parameters
F = @(PQ) max( ConvergenceFactor(k,[PQ(1) 0;PQ(2) 0],nu,a,b,c,L) );
[Robin2,fval,flag] = fminsearch(F,[1;1]);
if flag~=1, warning('fminsearch did not converge.'); end;

% Optimized second order parameters
F = @(PQ) max( ConvergenceFactor(k,[PQ' ; PQ'],nu,a,b,c,L) );
[Second,fval,flag] = fminsearch(F,[1;1]);
if flag~=1, warning('fminsearch did not converge.'); end;
```

```
function rho=ConvergenceFactor(k,params,nu,a,b,c,L);
% rho=ConvergenceFactor(k,params,nu,a,b,c,L);
% Computes the convergence factor in Fourier space
% for the Schwarz iteration with parameters params,
% at the frequencies k.

z = sqrt(a^2+4*nu*c+4*nu^2*k.^2 -4*i*nu*b*k);
p1 = params(1,1); q1 = params(1,2);
p2 = params(2,1); q2 = params(2,2);
sigma1 = p1+q1*(-b*i*k + nu*k.^2);
sigma2 = p2+q2*(-b*i*k + nu*k.^2);

rho = abs( (sigma1-z)./(sigma1+z) .*(sigma2-z)./(sigma2+z) );
rho = rho.*abs( exp(-L*z/nu) ); % for L>0
```

A.2 Diffusion Problem with Discontinuous Coefficient

We consider here a diffusion problem with discontinuous coefficient and a decomposition of a rectangle into vertical strips. The following function computes optimized parameters for a given interface by minimizing the convergence factor given by (4.15), which uses the exact Dirichlet-to-Neumann maps for each side of the interface. The convergence factor is uniformly minimized over the *discrete* set of relevant frequencies, see (4.16). For each fixed frequency k, the Fourier symbol of the Dirichlet-to-Neumann map is computed using the method described in Section 4.4.2.

```
function [Robin1v1,Robin1v2,Robin2,Second]=OptimizedParameters_DtN(j,nus,H,k1,k2);
% [Robin1v1,Robin1v2,Robin2,Second]=OptimizedParameters_DtN(j,nus,H,k1,k2);
% Computes optimized parameters for the diffusion equation
%    -div( nu(x) grad(u) ) = f
% with nu(x) = nus(i) in Omega_i, and the subdomains
% are vertical strips of width H. The function
% returns the optimized parameters for the j_th interface
% by solving numerically the min-max problem
% on the discrete frequency range k1:1:k2 and using
% the Fourier symbols of the appropriate DtN maps.
% Robin1v1:  one-sided Robin parameter version 1,  p
% Robin1v2:  one-sided Robin parameter version 1,  p
% Robin2:    two-sided Robin parameters,          [p1; p2]
% Second:    second order parameters,             [ p;  q]

% assume k1=pi/L and k2=N*pi/L
k = k1:k2; % discrete set of relevant frequencies

% diffusion coefficients to the left of j_th interface
nus1 = nus(j:-1:1); nu1 = nus(j);
% diffusion coefficients to the right of j_th interface
nus2 = nus(j+1:end);  nu2 = nus(j+1);
% precompute Fourier symbols of DtN maps
dtn1=DtNmap(k,nus1,H);
dtn2=DtNmap(k,nus2,H);

% Optimized Robin parameter, v1
F = @(p) max( ConvergenceFactor(k,[p 0;p 0],dtn1,dtn2) );
[Robin1v1,fval,flag] = fminsearch(F,1);
if flag~=1, warning('fminsearch did not converge.'); end;

% Optimized Robin parameter, v2
F = @(p) max( ConvergenceFactor(k,[nu2*p 0;nu1*p 0],dtn1,dtn2) );
[Robin1v2,fval,flag] = fminsearch(F,1);
if flag~=1, warning('fminsearch did not converge.'); end;
```

```
% Optimized two-sided Robin parameters
F = @(PQ) max( ConvergenceFactor(k,[nu2*PQ(1) 0;nu1*PQ(2) 0],dtn1,dtn2) );
[Robin2,fval,flag] = fminsearch(F,[1;1]);
if flag~=1, warning('fminsearch did not converge.'); end;

% Optimized second order parameters
F = @(PQ) max( ConvergenceFactor(k,[nu2*PQ' ; nu1*PQ'],dtn1,dtn2) );
[Second,fval,flag] = fminsearch(F,[1;1]);
if flag~=1, warning('fminsearch did not converge.'); end;
```

```
function rho=ConvergenceFactor(k,params,dtn1,dtn2);
% rho=ConvergenceFactor(k,params,dtn1,dtn2);
% Computes the convergence factor in Fourier space
% for the Schwarz iteration with parameters params,
% at the frequencies k, where the Fourier symbols
% of the DtN maps are supplied, evaluated at k.

p1 = params(1,1); q1 = params(1,2);
p2 = params(2,1); q2 = params(2,2);
sigma1 = p1+q1*k.^2;
sigma2 = p2+q2*k.^2;

rho = abs( (sigma1-dtn2)./(sigma1+dtn1)...
         .*(sigma2-dtn1)./(sigma2+dtn2) );
```

```
function sigma=DtNmap(ks,nus,H);
% sigma=DtNmap(k,nus,H);
% Computes the Fourier symbol sigma(k)
% of the Dirichlet-to-Neumann map corresponding
% to M vertical strips of width H
% with diffusion coefficient nus(1), nus(2), etc.

sigma = [];
for k=ks,
  k = abs(k);
  if tanh(k*H)==1, % large frequencies
    sigma(end+1) = nus(1)*k;
  elseif k==0,     % low frequencies
    sigma(end+1) = 1/sum(1./nus) * 1/H;
  else,
    M = length(nus);
    nur = nus(2:end) ./ nus(1:end-1);
    C = cosh(k*H);  S = sinh(k*H);  T = S/C;

    % product of 2x2 matrices
    Pinv = eye(2);
    for j=M-1:-1:1,
      Pinv = [C  -nur(j)*S; -S nur(j)*C]*Pinv;
    end;
    v = Pinv*[-T; 1];
    B1 = v(2)/v(1); % eta/xi
    sigma(end+1) = -nus(1)*k*B1;
  end;
end;
```

A.3 Diffusion-Reaction Problem with Discontinuous Coefficients

Here we consider a diffusion-reaction problem with discontinuous coefficients and the decomposition of a square into two symmetric subdomains. The following function computes optimized parameters for the transmission conditions described in Section 3.7.1, by numerically solving the min-max problem in the same way as in Appendix A.1, again using a sampling of frequencies on a logarithmic scale.

```
function [Robin1,Robin2,Second]=OptimizedParameters_DR(nu1,nu2,eta1,eta2,k1,k2);
% [Robin1,Robin2,Second]=OptimizedParameters_DR(nu1,nu2,eta1,eta2,k1,k2);
% Computes optimized parameters for the diffusion-reaction problem
%    - div( nu(x) grad(u) ) + eta(x) u = f
% where nu(x)=nu_j and eta(x)=eta_j in Omega_j, j=1,2.
% by solving numerically the min-max problem
% with the frequency range [k1,k2].
% Robin1:  one-sided Robin parameter,   p
% Robin2:  two-sided Robin parameters, [p1; p2]
% Second:  second order parameters,    [ p;  q]

N = 10000; % number of sampled frequencies

% uniform grid for k in [0,1]
if k1==0, kstart = 0:0.01:1; k1 = 1;
else,     kstart = []; end;
% uniform grid on a logarithmic scale
k = log(k1):(log(k2)-log(k1))/N:log(k2);
k = exp(k); k = [kstart k];

% Optimized Robin parameter
F = @(p) max( ConvergenceFactor(k,[p 0;p 0],nu1,nu2,eta1,eta2) );
[Robin1,fval,flag] = fminsearch(F,1);
if flag~=1, warning('fminsearch did not converge.'); end;
```

```
% Optimized two-sided Robin parameters
F = @(PQ) max( ConvergenceFactor(k,[PQ(1) 0;PQ(2) 0],nu1,nu2,eta1,eta2) );
[Robin2,fval,flag] = fminsearch(F,[1;1]);
if flag~=1, warning('fminsearch did not converge.'); end;

% Optimized second order parameters
F = @(PQ) max( ConvergenceFactor(k,[PQ' ; PQ'],nu1,nu2,eta1,eta2) );
[Second,fval,flag] = fminsearch(F,[1;1]);
if flag~=1, warning('fminsearch did not converge.'); end;


function rho=ConvergenceFactor(k,params,nu1,nu2,eta1,eta2);
% rho=ConvergenceFactor(k,params,nu1,nu2,eta1,eta2);
% Computes the convergence factor in Fourier space
% for the Schwarz iteration with parameters params,
% at the frequencies k.

lambda1 = nu1*sqrt(k.^2 + eta1/nu1);
lambda2 = nu2*sqrt(k.^2 + eta2/nu2);
p1 = params(1,1); q1 = params(1,2);
p2 = params(2,1); q2 = params(2,2);
sigma1 = nu2*( p1+q1*k.^2 ); % appropriate scaling
sigma2 = nu1*( p2+q2*k.^2 );

rho = abs( (sigma1-lambda2)./(sigma1+lambda1) ...
         .*(sigma2-lambda1)./(sigma2+lambda2) );
```

Appendix B

Additional Numerical Experiments

B.1 Decomposition into Vertical Strips with Continuous Diffusion

In this section, we repeat the experiment of Section 4.2, but in the case of a continuous diffusion coefficient. We give the number of iterations required to reach a tolerance of 10^{-6}, and use a finite volume discretization on a uniform grid with 240 points in each direction. We also give the iteration number we get when using the Schwarz method as a preconditioner for BiCGstab, as explained in Section 3.6.1. Table B.1 collects the results for several numbers of subdomains ranging from 2 to 16, in the case $\nu_1 = \nu_2 = 0.1$, i.e. $\mu = 1$. Recall first that when $\nu_1 = \nu_2$, the two versions of optimized Robin conditions analyzed in Sections 3.4.1 and 3.4.2 are the same. Observe in particular that the optimized Schwarz methods always converge here, whereas divergence sometimes occurs when the diffusion coefficient is discontinuous (compare with Table 4.1).

Number of subdomains	2	4	6	8	10	12	16
	Schwarz method as an iterative solver						
Opt. Robin v.2	113	174	247	322	399	475	631
Opt. 2-sided Robin	41	60	85	110	136	160	215
Opt. 2nd order	21	30	41	53	64	74	99
DN best	4	53	117	202	320	452	806
	With Krylov acceleration (BiCGstab)						
Opt. Robin v.2	24	30	37	41	41	45	53
Opt. 2-sided Robin	20	26	27	31	35	41	50
Opt. 2nd order	13	14	17	21	27	37	47
DN best no relaxation	6	30	49	91	130	175	595

Table B.1: Number of iterations to reach a tolerance of 10^{-6}, for an increasing number of subdomains, when $\nu_1 = \nu_2 = 0.1$ ($\boldsymbol{\mu} = \mathbf{1}$), $N = 240$.

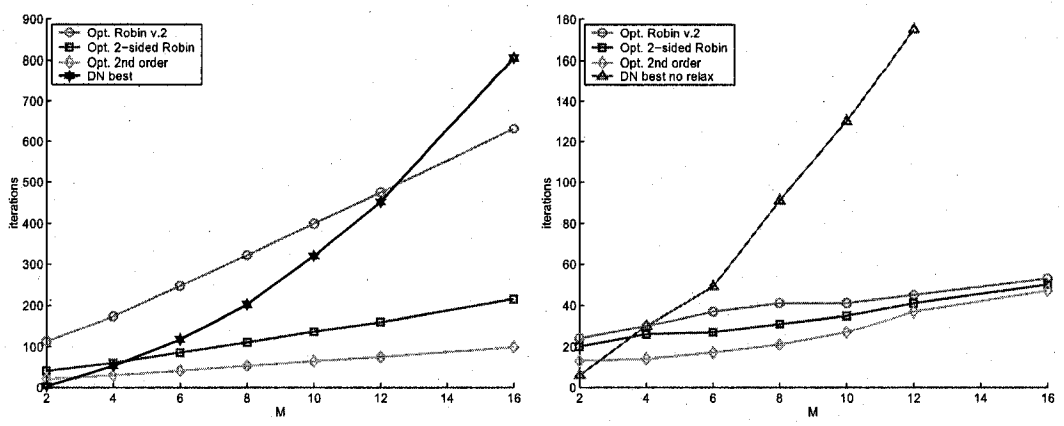

Figure B.1: Growth of the number of iterations as a function of M, for the iterative solver on the left, and for the preconditioned Krylov subspace method on the right, in the case $\boldsymbol{\mu} = \mathbf{1}$, $N = 240$.

B.2 A Diffusion-Reaction Problem with Discontinuous Diffusion

The numerical experiments shown in this section are carried out on the diffusion-reaction problem

$$\begin{cases} -\nabla \cdot (\nu(\mathbf{x})\nabla u) + u &= 1 \quad \text{in } \Omega = (0,\pi) \times (0,\pi), \\ u &= 0 \quad \text{on } \partial\Omega. \end{cases} \tag{B.1}$$

The square domain is decomposed into two non-overlapping subdomains

$$\Omega_1 = \left(0, \frac{\pi}{2}\right) \times (0,\pi), \quad \Omega_2 = \left(\frac{\pi}{2}, \pi\right) \times (0,\pi),$$

and $\nu(x) = \nu_j$ in Ω_j. We use a finite volume discretization on a uniform grid with grid spacing h.

We compare several choices of transmission conditions: Taylor approximations, Robin and second order conditions optimized numerically, and also Robin and second order conditions where the parameters are calculated by using the asymptotic formulas derived in Section 3.7.1.

- T0: Taylor approximation of zeroth order

$$\sigma_1(k) = \sqrt{\nu_2}, \qquad \sigma_2(k) = \sqrt{\nu_1}.$$

- T2: Taylor approximation of second order

$$\sigma_1(k) = \sqrt{\nu_2} + \frac{1}{2}\nu_2^{\frac{3}{2}}k^2, \qquad \sigma_2(k) = \sqrt{\nu_1} + \frac{1}{2}\nu_1^{\frac{3}{2}}k^2.$$

- R1: optimized Robin conditions (second version with the appropriate scaling)
- R1a: asymptotic formula for the optimized Robin parameter
- R2: optimized two-sided conditions
- R2a: asymptotic formulas for the optimized two-sided parameters
- S: optimized second order conditions

- Sa: asymptotic formulas for the optimized second order parameters

First, we fix $\mu = 10$ and vary the grid size h. Table B.2 shows the number of iterations that were needed to reach a tolerance of 10^{-6} for the different optimized methods. Note in particular the poor h-asymptotic convergence factor for the Taylor approximations, which is expected to be of the form $1 - O(h^{\frac{1}{2}})$. Also observe the very good performance obtained when using the asymptotic formulas for the optimized parameters in the one-sided Robin and second order conditions. For the two-sided Robin conditions, in this case the asymptotic formulas give significantly worse convergence when compared to the "exact" optimized parameters, however the iteration numbers do not grow as h decreases.

Next, we fix $h = \frac{\pi}{400}$ and vary the heterogeneity ratio μ. Table B.3 again shows the number of iterations needed to reach a tolerance of 10^{-6} for the different methods. The convergence is very fast when using the second order Taylor approximations for large μ, it even beats the performance of optimized transmission conditions. This μ-asymptotic behavior could probably be analyzed analytically by looking at the convergence factor for a model problem. Take note as well that in this case, the asymptotic formula for Robin parameter performs better than when using the fully optimized value, under Krylov acceleration.

	T0	T2	R1	R1a	R2	R2a	S	Sa
h	Optimized Schwarz method as an iterative solver							
$\frac{\pi}{50}$	86	25	19	18	11	22	8	8
$\frac{\pi}{100}$	166	42	24	24	11	22	10	10
$\frac{\pi}{200}$	328	76	32	32	11	22	12	12
$\frac{\pi}{400}$	662	145	44	44	12	23	13	14
$\frac{\pi}{800}$	1000	278	56	56	13	23	14	16
	Optimized Schwarz method with Krylov acceleration (BiCGstab)							
$\frac{\pi}{50}$	34	14	13	12	8	13	7	9
$\frac{\pi}{100}$	68	18	14	14	10	13	8	8
$\frac{\pi}{200}$	52	24	17	17	10	13	11	9
$\frac{\pi}{400}$	146	31	21	20	11	13	10	10
$\frac{\pi}{800}$	203	39	22	22	10	13	11	12

Table B.2: Number of iterations to reach a tolerance of 10^{-6}, for small values of h.

	T0	T2	R1	R1a	R2	R2a	S	Sa
μ	Optimized Schwarz method as an iterative solver							
10^1	662	145	44	44	12	23	13	14
10^2	96	14	14	14	10	11	10	9
10^3	20	6	6	8	8	8	7	8
10^4	8	4	6	6	6	6	6	6
10^5	6	4	6	6	6	6	5	6
	Optimized Schwarz method with Krylov acceleration (BiCGstab)							
10^1	146	31	21	20	11	13	10	10
10^2	34	9	10	10	8	9	9	9
10^3	18	6	6	6	5	6	7	7
10^4	15	4	7	6	5	5	6	6
10^5	11	4	8	5	5	5	5	6

Table B.3: Number of iterations to reach a tolerance of 10^{-6}, for large values of the heterogeneity ratio μ.

Bibliography

[1] Y. Achdou, C. Japhet, Y. Maday, and F. Nataf, *A new cement to glue non-conforming grids with Robin interface conditions: the finite volume case*, Numer. Math., 92 (2002), pp. 593–620.

[2] A. Alonso, R. L. Trotta, and A. Valli, *Coercive domain decomposition algorithms for advection-diffusion equations and systems*, J. Comput. Appl. Math., 96 (1998), pp. 51–76.

[3] M. Arioli, I. Duff, and D. Ruiz, *Stopping criteria for iterative solvers*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 138–144.

[4] J.-D. Benamou, *A domain decomposition method with coupled transmission conditions for the optimal control of systems governed by elliptic partial differential equations*, SIAM J. Numer. Anal., 33 (1996), pp. 2401–2416.

[5] J. D. Benamou and B. Després, *A domain decomposition method for the Helmholtz equation and related optimal control problems*, J. of Comp. Physics, 136 (1997), pp. 68–82.

[6] D. Bennequin, M. J. Gander, and L. Halpern, *A homographic best approximation problem with application to optimized Schwarz waveform relaxation*, (2006). Submitted.

[7] A. Bensoussan, J.-L. Lions, and G. Papanicolaou, *Asymptotic analysis for periodic structures*, vol. 5 of Studies in Mathematics and its Applications, North-Holland Publishing Co., Amsterdam, 1978.

[8] E. BLAYO, L. HALPERN, AND C. JAPHET, *Optimized Schwarz waveform relaxation algorithms with nonconforming time discretization for coupling convection-diffusion problems with discontinuous coefficients*, in Domain Decomposition Methods in Science and Engineering XVI, O. B. Widlund and D. E. Keyes, eds., vol. 55 of Lecture Notes in Computational Science and Engineering, Springer-Verlag, 2006, pp. 267–274.

[9] X.-C. CAI AND M. SARKIS, *A restricted additive Schwarz preconditioner for general sparse linear systems*, SIAM Journal on Scientific Computing, 21 (1999), pp. 239–247.

[10] D.-G. CALUGARU AND D. TROMEUR-DERVOUT, *Non-overlapping DDMs to solve flow in heterogeneous porous media*, in Proceedings of the 15th international domain decomposition conference, R. Kornhuber, R. H. W. Hoppe, J. Périaux, O. Pironneau, O. B. Widlund, and J. Xu, eds., Springer LNCSE, 2003, pp. 529–536.

[11] P. CHARTON, F. NATAF, AND F. ROGIER, *Méthode de décomposition de domaine pour l'équation d'advection-diffusion*, C. R. Acad. Sci., 313 (1991), pp. 623–626.

[12] P. CHEVALIER AND F. NATAF, *Symmetrized method with optimized second-order conditions for the Helmholtz equation*, in Domain decomposition methods, 10 (Boulder, CO, 1997), Amer. Math. Soc., Providence, RI, 1998, pp. 400–407.

[13] C. CHNITI, F. NATAF, AND F. NIER, *Improved interface conditions for a non-overlapping domain decomposition of a non-convex polygonal domain*, C. R. Math. Acad. Sci. Paris, 342 (2006), pp. 883–886.

[14] F. COLLINO, S. GHANEMI, AND P. JOLY, *Domain decomposition method for harmonic wave propagation: a general presentation*, Computer methods in applied mechanics and engineering, 184 (2000), pp. 171–211.

[15] H. DARCY, *Les Fontaines Publiques de la Ville de Dijon*, Dalmont, 1856.

[16] B. DESPRÉS, *Décomposition de domaine et problème de Helmholtz*, C.R. Acad. Sci. Paris, 1 (1990), pp. 313–316.

[17] B. DESPRÉS, *Méthodes de décomposition de domaines pour les problèms de propagation d'ondes en régime harmonique*, PhD thesis, Université Paris IX Dauphine, 1991.

[18] V. DOLEAN, M. J. GANDER, AND L. GERARDO-GIORDA, *Optimized Schwarz methods for Maxwell equations*, (2006). Preprint.

[19] V. DOLEAN, S. LANTERI, AND F. NATAF, *Optimized interface conditions for domain decomposition methods in fluid dynamics*, Int. J. Numer. Meth. Fluids, 40 (2002), pp. 1539–1550.

[20] V. DOLEAN AND F. NATAF, *A new domain decomposition method for the compressible Euler equations*, ESAIM-M2AN, 40 (2006), pp. 689–703.

[21] O. DUBOIS, *Optimized Schwarz methods for the advection-diffusion equation*, Master's thesis, McGill University, 2003.

[22] ——, *Optimized Schwarz methods with Robin conditions for the advection-diffusion equation*, in Domain Decomposition Methods in Science and Engineering XVI, O. B. Widlund and D. E. Keyes, eds., vol. 55 of Lecture Notes in Computational Science and Engineering, Springer-Verlag, 2006, pp. 181–188.

[23] E. EFSTATHIOU AND M. J. GANDER, *Why restricted additive Schwarz converges faster than additive Schwarz*, BIT, 43 (2003), pp. 945–959.

[24] I. FAILLE, E. FLAURAUD, F. NATAF, F. SCHNEIDER, AND F. WILLEN, *Optimized interface conditions for sedimentary basin modeling*, in Proceedings of the Thirteenth International Conference on Domain Decomposition Methods, N. Debit, M. Garbey, R. Hoppe, J. Périaux, D. Keyes, and Y. Kuznetsov, eds., 2001, pp. 453–460.

[25] C. FARHAT, P.-S. CHEN, J. MANDEL, AND F. X. ROUX, *The two-level FETI method. II. Extension to shell problems, parallel implementation and performance results*, Comput. Methods Appl. Mech. Engrg., 155 (1998), pp. 153–179.

[26] C. FARHAT, M. LESOINNE, P. LETALLEC, K. PIERSON, AND D. RIXEN, *FETI-DP: A dual-primal unified FETI method - part i: A faster alternative to the two-level FETI method*, Int. J. Numer. Meth. Engrg., 50 (2001), pp. 1523–1544.

[27] C. FARHAT, A. MACEDO, AND R. TEZAUR, *FETI-H: a scalable domain decomposition method for high frequency exterior Helmholtz problems*, in Eleventh International Conference on Domain Decomposition Method, C.-H. Lai, P. Bjørstad, M. Cross, and O. Widlund, eds., DDM.ORG, 1999, pp. 231–241.

[28] C. FARHAT, J. MANDEL, AND F.-X. ROUX, *Optimal convergence properties of the FETI domain decomposition method*, Comput. Methods Appl. Mech. Engrg., 115 (1994), pp. 367–388.

[29] C. FARHAT AND F.-X. ROUX, *A method of Finite Element Tearing and Interconnecting and its parallel solution algorithm*, Int. J. Numer. Meth. Engrg., 32 (1991), pp. 1205–1227.

[30] E. FLAURAUD, F. NATAF, AND F. WILLIEN, *Optimized interface conditions in domain decomposition methods for problems with extreme contrasts in the coefficients*, J. Comput. Appl. Math., 189 (2006), pp. 539–554.

[31] M. J. GANDER, *Optimized Schwarz methods for Helmholtz problems*, in Thirteenth international conference on domain decomposition, 2001, pp. 245–252.

[32] ——, *Optimized Schwarz methods*, SIAM J. Numer. Anal., 44 (2006), pp. 699–731.

[33] M. J. GANDER, L. HALPERN, AND M. KERN, *A Schwarz waveform relaxation method fo advection-diffusion-reaction problems with discontinuous coefficients*

and non-matching grids, in Domain Decomposition Methods in Science and Engineering XVI, O. B. Widlund and D. E. Keyes, eds., vol. 55 of Lecture Notes in Computational Science and Engineering, Springer-Verlag, 2006.

[34] M. J. GANDER, L. HALPERN, AND F. NATAF, *Optimized Schwarz methods*, in Twelfth International Conference on Domain Decomposition Methods, Chiba, Japan, T. Chan, T. Kako, H. Kawarada, and O. Pironneau, eds., Bergen, 2001, Domain Decomposition Press, pp. 15–28.

[35] M. J. GANDER, C. JAPHET, Y. MADAY, AND F. NATAF, *A new cement to glue nonconforming grids with Robin interface conditions: the finite element case*, in Domain decomposition methods in science and engineering, vol. 40 of Lect. Notes Comput. Sci. Eng., Springer, Berlin, 2005, pp. 259–266.

[36] M. J. GANDER, F. MAGOULÈS, AND F. NATAF, *Optimized Schwarz methods without overlap for the Helmholtz equation*, SIAM J. Sci. Comput., 24 (2002), pp. 38–60.

[37] L. GERARDO-GIORDA, P. LE TALLEC, AND F. NATAF, *A Robin-Robin preconditioner for advection-diffusion equations with discontinuous coefficients*, Comput. Methods Appl. Mech. Engrg., 193 (2004), pp. 745–764.

[38] L. GERARDO-GIORDA AND F. NATAF., *Optimized Schwarz methods for unsymmetric layered problems with strongly discontinuous and anisotropic coefficients*, J. Numer. Math., 13 (2005), pp. 265–294.

[39] L. GERARDO-GIORDA AND F. NATAF, *Optimized algebraic interface conditions in domain decomposition methods for strongly heterogeneous unsymmetric problems*, in Domain Decomposition Methods in Science and Engineering XVI, O. B. Widlund and D. E. Keyes, eds., vol. 55 of Lecture Notes in Computational Science and Engineering, Springer-Verlag, 2006.

[40] C. JAPHET, *Conditions aux limites artificielles et décomposition de domaine: Méthode OO2 (Optimisée d'Ordre 2). Application à la résolution de problèmes en mécanique des fluides*, Tech. Rep. 373, CMAP (Ecole Polytechnique), 1997.

[41] ——, *Optimized Krylov-Ventcell method. Application to convection-diffusion problems*, in Proceedings of the 9th international conference on domain decomposition methods, P. E. Bjørstad, M. S. Espedal, and D. E. Keyes, eds., ddm.org, 1998, pp. 382–389.

[42] C. JAPHET AND F. NATAF, *The best interface conditions for domain decomposition methods: absorbing boundary conditions*, in Absorbing boundaries and layers, domain decomposition methods, Nova Sci. Publ., Huntington, NY, 2001, pp. 348–373.

[43] C. JAPHET, F. NATAF, AND F. ROGIER, *The Optimized Order 2 method. Application to convection-diffusion problems*, Future Generation Computer Systems FUTURE, 18 (2001).

[44] C. JAPHET, F. NATAF, AND F.-X. ROUX, *Extension of a coarse grid preconditioner to non-symmetric problems*, in Domain decomposition methods, 10 (Boulder, CO, 1997), vol. 218 of Contemp. Math., Amer. Math. Soc., Providence, RI, 1998, pp. 279–286.

[45] C. JAPHET, F. NATAF, AND F.-X. ROUX, *The Optimized Order 2 Method with a coarse grid preconditioner. Application to convection-diffusion problems*, in Ninth International Conference on Domain Decompositon Methods in Science and Engineering, P. Bjorstad, M. Espedal, and D. Keyes, eds., John Wiley & Sons, 1998, pp. 382–389.

[46] P.-L. LIONS, *On the Schwarz alternating method. I.*, in First International Symposium on Domain Decomposition Methods for Partial Differential Equations, R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, eds., Philadelphia, PA, 1988, SIAM, pp. 1–42.

[47] ——, *On the Schwarz alternating method. III: a variant for nonoverlapping subdomains*, in Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, held in Houston, Texas, March 20-22, 1989, T. F. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., Philadelphia, PA, 1990, SIAM.

[48] S. H. LUI, *On the condition number of an optimized Schwarz method*, (2007). Preprint.

[49] ——, *Optimal Lions' nonoverlapping domain decomposition method for domains with an arbitrary interface*, (2007). Preprint.

[50] Y. MADAY AND F. MAGOULÈS, *Multilevel optimized Schwarz methods without overlap for highly heterogeneous media*, R05015, Laboratoire Jacques-Louis Lions, 2005.

[51] ——, *Improved ad hoc interface conditions for Schwarz solution procedure tuned to highly heterogenous media*, Applied Mathematical Modelling, 30 (2006), pp. 731–743.

[52] G. MEINARDUS, *Approximation of functions: theory and numerical methods*, Springer-Verlag, 1967.

[53] F. MURAT AND L. TARTAR, *Calculus of variations and homogenization*, in Topics in the mathematical modelling of composite materials, vol. 31 of Progr. Nonlinear Differential Equations Appl., Birkhäuser Boston, 1997, pp. 139–173.

[54] F. NATAF, *Interface connections in domain decomposition methods*, in Modern Methods in Scientific Computing and Applications, vol. 75 of NATO Advanced Study Institute, Université de Montréal, NATO Science Series II, 2001.

[55] F. NATAF AND F. ROGIER, *Factorization of the convection-diffusion operator and the Schwarz algorithm*, M^3AS, 5 (1995), pp. 67–93.

[56] F. NATAF, F. ROGIER, AND E. DE STURLER, *Optimal interface conditions for domain decomposition methods*, Tech. Rep. 301, CMAP (Ecole Polytechnique), 1994.

[57] J. NELDER AND R. MEAD, *A simplex method for function minimization*, The Computer Journal, 7 (1964), pp. 308–313.

[58] A. QADDOURI, L. LAAYOUNI, S. LOISEL, J. COTÉ, AND M. J. GANDER, *Optimized Schwarz methods with an overset grid for the shallow-water equations: preliminary results*, Applied Numerical Mathematics, (2007). In press, corrected proof, available online 15 February 2007.

[59] L. QIN AND X. XU, *On a parallel Robin-type nonoverlapping domain decomposition method*, SIAM J. Numer. Anal., 44 (2006), pp. 2539–2558 (electronic).

[60] A. QUARTERONI AND A. VALLI, *Domain Decomposition Methods for Partial Differential Equations*, Oxford Science Publications, 1999.

[61] Y. SAAD AND M. H. SCHULTZ, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comp., 7 (1986), pp. 856–869.

[62] L. SAAS, I. FAILLE, F. NATAF, AND F. WILLIEN, *Finite volume methods for domain decomposition on nonmatching grids with arbitrary interface conditions*, SIAM J. Numer. Anal., 43 (2005), pp. 860–890 (electronic).

[63] H. A. SCHWARZ, *Über einen Grenzübergang durch alternierendes Verfahren*, Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich, 15 (1870), pp. 272–286.

[64] B. F. SMITH, P. E. BJØRSTAD, AND W. GROPP, *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*, Cambridge University Press, 1996.

[65] A. TOSELLI AND O. WIDLUND, *Domain decomposition methods—algorithms and theory*, vol. 34 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 2005.

[66] H. A. VAN DER VORST, *Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 631–644.