www.oeaw.ac.at

ÖAW AUSTRIAN ACADEMY OF SCIENCES

Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

RICAM JOHANN-RADON-INSTITUTE FOR COMPUTATIONAL AND APPLIED MATHEMATICS

Overlapping Additive Schwarz preconditioners for degenerated elliptic problems: Part I- isotropic problems

S. Beuchler, S.V. Nepomnyaschikh

RICAM-Report 2006-32

www.ricam.oeaw.ac.at

Overlapping Additive Schwarz preconditioners for degenerated elliptic problems: Part I isotropic problems

Sven Beuchler
Institute for Computational Mathematics
University of Linz
Altenberger Strasse 69
A-4040 Linz, Austria
sven.beuchler@jku.at

Sergey V. Nepomnyaschikh
Institute for Computational Mathematics
and Computational Geophysics
SD Russian Academy of Sciences
Novosibirsk, Russia
svnep@oapmg.sscc.ru

November 8, 2006

Abstract

In this paper, we consider the degenerated isotropic boundary value problem $-\nabla(\omega^2(x)\nabla u(x,y)) = f(x,y)$ on the unit square $(0,1)^2$. The weight function is assumed to be of the form $\omega^2(\xi) = \xi^\alpha$, where $\alpha \geq 0$. This problem is discretized by piecewise linear finite elements on a triangular mesh of isosceles right-angled triangles. The system of linear algebraic equations is solved by a preconditioned gradient method using a domain decomposition preconditioner with overlap. Two different preconditioners are presented and the optimality of the condition number for the preconditioned system is proved for $\alpha \neq 1$. The preoconditioning operation requires $\mathcal{O}(N)$ operations, where N is the number of unknowns. Several numerical experiments show the preformance of the proposed method.

1 Introduction

In this paper, we investigate the degenerated and isotropic boundary value problem

$$\begin{aligned} -(\omega^2(x)u_x)_x - (\omega^2(x)u_y)_y &= f, \quad \text{in } \Omega = (0,1)^2 \\ u &= 0, \quad \text{on } \partial\Omega \end{aligned} \tag{1.1}$$

with some strongly monotonic increasing and bounded weight function $\omega : [0,1] \mapsto \mathbb{R}$ satisfying $\omega(0) = 0$. In the past, degenerated problems have been considered relatively rarely. One reason is the unphysical behavior of the partial differential equation (pde) which is quite unusual in technical applications. One work focusing on this type of partial differential equation is the book of Kufner and Sändig [17]. Nowadays, problems of this type become more and more popular because there are stochastic pde's of a similar structure. An example of an isotropic degenerated stochastic pde is the Black-Scholes pde, [21].

Moreover, there are examples of locally anisotropic degenerated elliptic problems. One of them is the solver related to the problem of the sub-domains for the p-version of the finite element method using quadrilateral elements. This matrix can be interpreted as h-version fem-discretization matrix of the problem $-y^2u_{xx} - x^2u_{yy} = f$. We refer to [1], [2] for more details.
The discretization of (1.1) using the h-version of the finite element method (fem) leads to a linear system of algebraic equations

$$\mathcal{K}\underline{u} = \underline{f}. \tag{1.2}$$

It is well known from the literature that preconditioned conjugate gradient-methods (pcg-methods) with domain decomposition preconditioners are among the most efficient iterative solvers for systems of the type (1.2), see e.g. [7], [8], [9], [10], [23], [18]. In this paper, we will propose and analyze overlapping Domain Decomposition (DD) preconditioners.
The type of overlapping DD-preconditioners presented in this paper is originally developed for problems with jumping coefficients in [20], see also [13], [22] for the case of highly variiying coefficients. In a second paper [3], we will analyze these overlapping DD preconditioners for a locally anisotropic degenerated problems. Here, we adapt the techniques of [20] to problem (1.1). To keep the notation and the proofs simple, we will prove the optimality of this method only for tensor product discretizations in two dimensions. The generalization of the method to three dimensional tensor product discretizations is straightforward. Moreover, this method can be extended to more general h-version fem discretizations, using the ficticous space lemma, [19].
Only a limited number of papers have investigated fast solvers for degenerated elliptic problems. The paper [6] deals with the Laplacian in $2D$ in polar coordinates. In the paper [12], multigrid methods for some other types of degenerated problems are proposed. Multigrid solvers for FE-discretizations of the problems in [3] have been investigated in [1], see also [5] and [16]. The paper [4] proposes wavelet methods for several classes of degenerated elliptic problems on the unit square. One of them is problem (1.1) under the restriction $\lim_{\xi\to 0+} \frac{\xi}{\omega^2(\xi)} = 0$ to the weight functions. Moreover, a fast direct solver based on eigenvalue computations combined with fast Fourier transform can be designed if a tensor product discretization is used.
The remaining part of this paper is organized as follows. In Section 2, we introduce the reader into our problem and into our notation. The preconditioners are defined in Section 3. Moreover, the main theorems with the condition number estimates are stated. The efficient solution of the preconditioned systems will be presented in Section 4. In Section 5, we formulate some auxiliary results from the Additive Schwarz Method (ASM), which are required for the proofs of our main theorems given in Section 6. In Section 7, we present some numerical experiments which show the performance of the presented methods. Finally, we present some concluding remarks and generalizations to a general domain using the Ficticous space lemma.
Throughout this paper, the integer k denotes the level number. For two real symmetric and positive definite $n \times n$ matrices A, B, the relation $A \preceq B$ means that $A - cB$ is negative definite, where $c > 0$ is a constant independent of n. The relation $A \sim B$ means $A \preceq B$ and $B \preceq A$, i.e. the matrices A and B are spectrally equivalent. The parameter c denotes a generic constant. The isomorphism between a function $u = \sum_i u_i\psi_i \in L^2$ and the corresponding vector of coefficients $\underline{u} = [u_i]_i$ in the basis $[\Psi] = [\psi_1, \psi_2, \ldots]$ is denoted by $u = [\Psi]\underline{u}$.

2 Setting of the problem

In this paper, we investigate the following boundary value problem: Let $\Omega = (0,1)^2$. Find $u \in \mathbb{H}_{\omega,0} := \{u \in L_2(\Omega) : \int_\Omega \omega^2(x)(\nabla u)^T \nabla u \,\mathrm{d}(x,y) < \infty, u\mid_{\partial\Omega} = 0\}$ such that

$$a(u,v) := \int_\Omega \omega^2(x)(\nabla v)^T(x,y)\nabla u(x,y)\,\mathrm{d}(x,y) = (f,v) \quad \forall v \in \mathbb{H}_{\omega,0}. \tag{2.1}$$

We point out that the diffusion matrix $\mathcal{D} = \omega^2(x)I$ of (2.1), where I denoted the unity matrix, is not necessarily uniformly positive definite in $\overline{\Omega}$.
To be specific, we consider the weight function $\omega^2(x) = x^\alpha$, $\alpha > 0$.

Lemma 2.1. *The function $\omega : [0,1] \mapsto \mathbb{R}$ given by $\omega^2(x) = x^\alpha$, $\alpha > 0$ satisfies the following assertions:*

- *the function ω is monotonic increasing,*
- *the function ω is continuous,*
- *the estimate*

$$\omega(2\xi) \le c_\omega \omega(\xi) \quad \forall \xi \in \left(0, \frac{1}{2}\right] \tag{2.2}$$

holds with some constant $c_\omega = 2^{-\alpha/2} > 0$.

Problems of the type (2.1) are called degenerated problems. In the past, degenerated problems have been considered relatively rarely. One reason is the unphysical behavior of the partial differential equation which is quite unusual in technical applications. Nowadays, problems of this type become more and more popular because there are stochastic pde's which have a similar structure. Setting $\omega(\xi) = \xi$, one obtains a degenerated stochastic partial differential equation, i.e. the Black-Scholes partial differential equation, [21].
We discretize problem (2.1) by piecewise linear finite elements on the regular Cartesian grid consisting of congruent, isosceles, right-angled triangles. For this purpose, some notation is introduced. Let k be the level of approximation and $n = 2^k$. Let $x_{ij}^k = (\frac{i}{n}, \frac{j}{n})$, where $i, j = 0, \ldots, n$. The domain Ω is divided into congruent, isosceles, right-angled triangles $\tau_{ij}^{s,k}$, where $0 \le i, j < n$ and $s = 1, 2$, see Figure 1. The triangle $\tau_{ij}^{1,k}$ has the three vertices $x_{ij}^k, x_{i+1,j+1}^k$ and $x_{i,j+1}^k$, $\tau_{ij}^{2,k}$ has the three vertices $x_{ij}^k, x_{i+1,j+1}^k$ and $x_{i+1,j}^k$, see Figure 1. Piecewise linear finite elements are used on the mesh $T_k = \{\tau_{ij}^{s,k}\}_{i=0,j=0,s=1}^{n-1,n-1,2}$. The

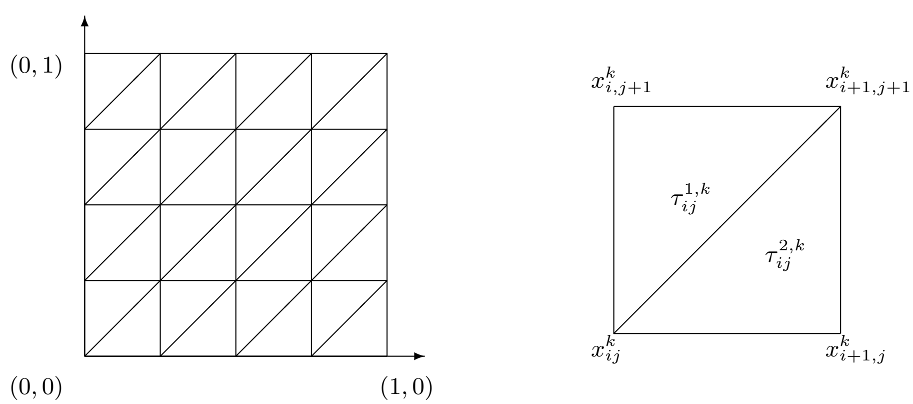

Figure 1: Mesh for the finite element method (left), Notation within a macro-element $\mathcal{E}_{ij}^k$ (right).

subspace of piecewise linear functions ϕ_{ij}^k with

$$\phi_{ij}^k \in H_0^1(\Omega),\ \phi_{ij}^k \mid_{\tau_{lm}^{s,k}} \in \mathbb{P}_1(\tau_{lm}^{s,k})$$

is denoted by $\mathbb{V}_k$, where $\mathbb{P}_1$ is the space of polynomials of degree ≤ 1. A basis of $\mathbb{V}_k$ is the system of the usual hat-functions $\Phi_k = \{\phi_{ij}^k\}_{i,j=1}^{n-1}$ uniquely defined by

$$\phi_{ij}^k(x_{lm}^k) = \delta_{il}\delta_{jm}$$

and $\phi_{ij}^k \in \mathbb{V}_k$, where δ_{il} is the Kronecker delta. Now, we can formulate the discretized problem. Find $u^k \in \mathbb{V}_k$ such that

$$a(u^k, v^k) = (f, v^k) \quad \forall v^k \in \mathbb{V}_k \tag{2.3}$$

holds. Problem (2.3) is equivalent to solving the system of linear algebraic equations

$$K_k \underline{u}_k = \underline{f}_k, \tag{2.4}$$

where $K_k = \left[a(\phi_{ij}^k, \phi_{lm}^k)\right]_{i,j,l,m=1}^{n-1}$, $\underline{u}_k = [u_{ij}]_{i,j=1}^{n-1}$ and $\underline{f}_k = \left[\langle f, \phi_{lm}^k \rangle\right]_{l,m=1}^{n-1}$. The size of the matrix K_k is $N \times N$ with $N = (n-1)^2$.

3 Definition of the preconditioners

In this section, we define the preconditioners for the matrix K_k (2.3). We introduce the following notation. Let

- $\Omega_{i,x} = \left\{(x, y) \in \mathbb{R}^2, 2^{-1-i} < x < 2^{-i}, 0 < y < 1\right\}$, $i = 0, \ldots, k-2$,
- $\Omega_{k-1,x} = \left\{(x, y) \in \mathbb{R}^2, 0 < x < 2^{-k}, 0 < y < 1\right\}$,
- $\Gamma_{i,x} = \left\{(x, y) \in \mathbb{R}^2, x = 2^{-i}, 0 < y < 1\right\}$, $i = 1, \ldots, k-1$,
- $\tilde{\Omega}_{j,x} = \text{int}\left(\bigcup_{i=j}^{k-1} \overline{\Omega_{i,x}}\right)$, and
- $n_j = 2^{k-j} - 1$ be the number of interior grid points in $\tilde{\Omega}_{j,x}$ in x-direction and $N_j = (n-1)n_j$ be the total number of interior grid points.
- Moreover, let
$$\varepsilon_j = \omega^2 \left(2^{-j}\right).$$

Figure 2 displays a sketch with the notation in the case $k = 4$. On $\tilde{\Omega}_{j,x}$, we introduce the bilinear form

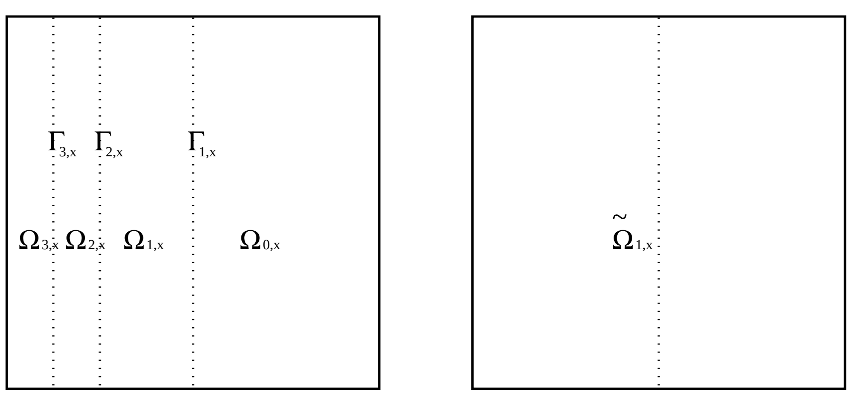

Figure 2: Notation for $k = 4$.

$$a_j(u, v) = \int_{\tilde{\Omega}_{j,x}} \nabla u \cdot \nabla v, \quad j = 0, \ldots, k-1.$$

Moreover, let

$$\begin{aligned} C_{j,D} &= [a_j(\phi_{ii'},\phi_{ll'})]_{i,l=1;\, i',l'=1}^{n_j,n_0}, \quad j=0,\dots,k-1, \quad \text{and} \\ C_{j,N} &= [a_j(\phi_{ii'},\phi_{ll'})]_{i,l=n_{j+1}+1; i',l'=1}^{n_j,n_0}, \quad j=0,\dots,k-2. \end{aligned}$$

These matrices correspond to the Laplacian on $\tilde{\Omega}_{j,x}$ with Dirchlet boundary conditions at the left boundary $x=0$ and to the Laplacian on $\Omega_{j,x}$ with Neumann boundary conditions at the left boundary $\Gamma_{j+1,x}$. At the remaining three edges, we have Dirichlet boundary conditions. Finally, let

$$\Delta_{j,D} = \begin{bmatrix} C_{j,D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{N-N_j} \end{bmatrix} \in \mathbb{R}^{N\times N} \quad \text{and} \quad \Delta_{j,N} = \begin{bmatrix} \mathbf{0}_{N_{j+1}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_{j,N} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0}_{N-N_j} \end{bmatrix} \in \mathbb{R}^{N\times N} \tag{3.1}$$

be the global assembled stiffness matrices.
Then, we define a first preconditioner

$$C^{-1} = \sum_{j=0}^{k-1} \varepsilon_j^{-1} \Delta_{j,D}^{+}, \tag{3.2}$$

where B^+ denotes the pseudo-inverse inverse of a matrix B. Then, we can prove the following result.

Theorem 3.1. *Let C be defined via* (3.2) *and let $\omega^2(\xi) = \xi^\alpha$. If $\alpha > 0$, then we have $K_k \preceq C$. If $0 \le \alpha < \frac{1}{2}$, then we have also $C \preceq K_k$.*

Proof. A detailed proof is presented in subsection 6.4. □

Since Theorem 3.1 can be proved only for $\alpha < \frac{1}{2}$ directly, we introduce a second preconditioner. Let

$$\hat{C}_{j,D} = \left[\int_{\overline{\Omega}_{j,x}\cup\overline{\Omega}_{j+1,x}} \nabla\phi_{ii'} \cdot \nabla\phi_{ll'} \right]_{i,j=n_{j+2}+2;\, i',l'=1}^{n_j,n_0}$$

be the Laplacian on $\overline{\Omega_{j+1,x}} \cup \overline{\Omega_{j,x}}$ with Dirichlet boundary conditions at all edges and

$$\hat{\Delta}_{j,D} = \begin{bmatrix} \mathbf{0}_{N_{j+2}+n_0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \hat{C}_{j,D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0}_{N-N_j} \end{bmatrix} \in \mathbb{R}^{N\times N}, \quad j=0,\dots,k-2, \tag{3.3}$$

be the corresponding assembled matrix. Then, we introduce a second overlapping preconditioner for K_k as

$$C_{mod}^{-1} = \sum_{j=0}^{k-2} \varepsilon_j^{-1} \hat{\Delta}_{j,D}^{+} + \varepsilon_{k-1}^{-1} \Delta_{k-1,D}^{+}. \tag{3.4}$$

Theorem 3.2. *Let C_{mod} be defined via* (3.4)*. Let $\omega^2(\xi) = \xi^\alpha$ with $\alpha \neq 1$. Then, the matrix C_{mod} is symmetric positive definite and satisfies $K_k \sim C_{mod}$.*

Proof. A detailed proof is given in subsection 6.1 for $\alpha > 1$ and in subsection 6.2 for $\alpha < 1$. □

Remark 3.3. *From the definition of the preconditioners, the relation $C_{mod} \le C$ follows directly. Combining Theorem 3.1 and Theorem 3.2, the estimate $C \sim K_k$ holds if $\alpha \neq 1$ and $\alpha > 0$. In the case $\alpha = 1$, we are not able to prove an optimal result. Here, only the weaker estimate $k^{-2}C \preceq K_k \preceq C$ can be proved. This behavior can also be seen in the numerical experiments of section 7.*

4 Computational aspects

In this section, we investigate the preconditioning operation $C^{-1}\underline{w}$ for the two preconditioners of proceeding section. We present algorithms to perform this preconditioning operation in optimal arithmetical complexity.
We have developed the preconditioners

$$C^{-1} = \sum_{j=0}^{k-1} \varepsilon_j^{-1} \Delta_{j,D}^{+},$$

see (3.2) and

$$C_{mod}^{-1} = \sum_{j=0}^{k-2} \varepsilon_j^{-1} \hat{\Delta}_{j,D}^{+} + \varepsilon_{k-1}^{-1} \Delta_{k-1,D}^{+},$$

see (3.4). For the operation $C^{-1}\underline{w}$, solvers for the Laplacian with Dirichlet boundary conditions on the domains $\tilde{\Omega}_{j,x}$, $j = 0, \ldots, k-1$ are required. The corresponding domains are displayed in Figure 3 for $k = 4$.
For the operation C_{mod}^{-1}, we need solvers for the Laplacian on the domains $\overline{\Omega}_{j,x} \cup \overline{\Omega}_{j+1,x}$, $j = 0, \ldots, k-2$, see Figure 4 for $k = 4$. In the case of nested triangulations, several optimal solution methods for the discretization of the Laplacian are known in the literature. Examples are Multigrid methods, see e.g. [14] and the references therein, pcg-methods with BPX-preconditioners, see [11], [25], or, multigrid preconditioners, [15]. In 2D, also a pcg-method with a hierarchical basis preconditioner is possible, [24].

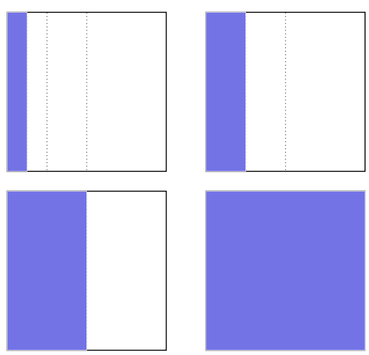

Figure 3: Computational domains for C (3.2): $\Delta_{3,D}$ and $\Delta_{2,D}$ above, $\Delta_{1,D}$ and $\Delta_{0,D}$ below.

Let $\mathcal{W}_j$ be the arithmetical cost for the solution of $\Delta_{j,D}\underline{w} = \underline{r}$ and $\mathcal{W}$ be the arithmetical cost for the solution of $C\underline{w} = \underline{r}$. Using one of the proposed methods mentioned above, we have

$$\mathcal{W}_j \leq c(n_0 + 1)(n_j + 1)$$

with some constant c which is independent of j and n. Then, we can estimate

$$\mathcal{W} = \sum_{j=0}^{k-1} \mathcal{W}_j \leq c(n_0 + 1) \sum_{j=0}^{k-1} (n_j + 1) = c(n_0 + 1) \sum_{j=0}^{k-1} 2^{k-j} \leq c(n_0 + 1)2^{k+1} \leq 2c(n_0 + 1)^2$$

using the geometric series. So, the cost of the preconditioning operation $C^{-1}\underline{w}$ is proportionally to the number of unknowns.
A similar result can be shown for the preconditioner C_{mod}.

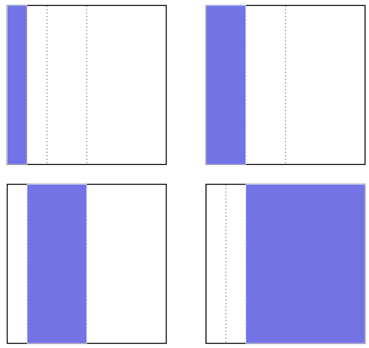

Figure 4: Computational domains for C_{mod} (3.4): $D\hat{e}lta_{3,D}$ and $\hat{\Delta}_{2,D}$ above, $\hat{\Delta}_{1,D}$ and $\hat{\Delta}_{0,D}$ below.

5 Preliminaries

In this section, we will formulate some auxiliary results.

5.1 Preliminaries from the Additive Schwarz Method

We start this subsection with the formulation of two results about the additive Schwarz method with inexact subproblem solvers. These results are developed in [19].

Lemma 5.1. *Let $\mathbb{H}$ be a Hilbert space with the scalar product $(\cdot,\cdot)$. Moreover, let $\mathbb{H}_i$, $i = 1,\ldots,m$ subspaces of $\mathbb{H}$ such that*

$$\mathbb{H} = \mathbb{H}_1 + \mathbb{H}_2 + \ldots + \mathbb{H}_m.$$

Let $\mathcal{A} : \mathbb{H} \mapsto \mathbb{H}$ be a linear, self adjoint, bounded and positive definite operator and let

$$(u,v)_{\mathcal{A}} = (\mathcal{A}u, v) \quad \forall u,v \in \mathbb{H}.$$

We denote by P_i, $i = 1,\ldots,m$, the orthogonal projection operators from $\mathbb{H}$ onto $\mathbb{H}_i$ with respect to the scalar product $(\cdot,\cdot)_{\mathcal{A}}$. We assume that for any $u \in \mathbb{H}$ there exists a decomposition $u = u_1 + \ldots + u_m$ such that

$$c_1 \sum_{i=1}^{m} (u_i,u_i)_{\mathcal{A}} \le (u,u)_{\mathcal{A}} \tag{5.1}$$

with a positive constant c_1. Moreover, let c_2 some positive constant such that

$$\sum_{i=1}^{m} (P_i u, u)_{\mathcal{A}} \le c_2\,(u,u)_{\mathcal{A}} \quad \forall u \in \mathbb{H}. \tag{5.2}$$

Also, let $\mathcal{B}_i : \mathbb{H} \mapsto \mathbb{H}_i$, $i = 1,\ldots,m$ be some selfadjoint operators such that

$$c_3\,(\mathcal{B}_i u_i, u_i) \le (\mathcal{A}u_i, u_i) \le c_4\,(\mathcal{B}_i u_i, u_i)\,, \quad \forall u_i \in \mathbb{H}_i,\ i = 1,\ldots,m. \tag{5.3}$$

Let $\mathcal{B}^{-1} = \mathcal{B}_1^+ + \ldots + \mathcal{B}_m^+$, where $\mathcal{B}_i^+$ denotes the pseudo-inverse operator for $\mathcal{B}_i$. Then,

$$c_1 c_3 \left(\mathcal{A}^{-1}u, u\right) \le \left(\mathcal{B}^{-1}u, u\right) \le c_2 c_4 \left(\mathcal{A}^{-1}u, u\right) \quad \forall u \in \mathbb{H}.$$

Lemma 5.2. *Let $\mathbb{V}$ and $\mathbb{W}$ be two Hilbert spaces with scalar products $(\cdot,\cdot)_{\mathbb{V}}$ and $(\cdot,\cdot)_{\mathbb{W}}$. Moreover, let Σ and S be selfadjoint, positive definite operators in $\mathbb{V}$ and $\mathbb{W}$, respectively. We denote by*

$$(\phi,\psi)_{\Sigma} = (\Sigma\phi,\psi)_{\mathbb{V}} \quad \textit{and} \quad (u,v)_{S} = (Su,v)_{\mathbb{W}}$$

the scalar products in $\mathbb{V}$ and $\mathbb{W}$ generated by the operators Σ and S. Let $\mathcal{E}:\mathbb{V}\mapsto\mathbb{W}$ be a linear operator such that

$$\alpha\,(\phi,\phi)_{\Sigma} \le (\mathcal{E}\phi,\mathcal{E}\phi)_{S} \le \beta\,(\phi,\phi)_{\Sigma} \quad \forall\phi\in\mathbb{V}.$$

Finally, we set

$$C^{+} = \mathcal{E}\Sigma^{-1}\mathcal{E}^{*},$$

where $\mathcal{E}^{*}$ is the adjoint to the operator with respect to the scalar products $(\cdot,\cdot)_{\mathbb{V}}$ and $(\cdot,\cdot)_{\mathbb{W}}$. Then,

$$\alpha\,(Cu,u)_{\mathbb{W}} \le (Su,u)_{\mathbb{W}} \le \beta\,(Cu,u)_{\mathbb{W}} \quad \forall u\in Im(\mathcal{E}) := \{u\in\mathbb{W};\exists v\in\mathbb{V}\ :\ u=\mathcal{E}v\}.$$

5.2 Algebraic Analysis of an overlapping DD-preconditioner

In this subsection, we prove an auxiliary result for an overlapping domain decomposition preconditioner in which the domain Ω is decomposed into stripes Ω_i. We consider the following situation:

- Let
$$\overline{\Omega} = \bigcup_{j=0}^{k-1}\overline{\Omega}_j$$
be a domain Ω which is decomposed into stripes Ω_i, i.e.
$$\Omega_i\cap\Omega_j=\emptyset \quad \text{for } i\neq j, \quad \overline{\Omega}_i\cap\overline{\Omega}_j = \begin{cases} \Gamma_i & i=j+1 \\ \Gamma_j & i=j-1 \\ \overline{\Omega}_i & i=j \\ \emptyset & |i-j|\ge 2 \end{cases}$$
and let $\overline{\Omega}_{k-1}\cap\partial\Omega = \Gamma_k$.
- Let τ_k be a triangulation of Ω which is admissible with the decomposition of Ω into Ω_i.
- Let $\Phi_k = [\phi_i]_{i=1}^{N}$ be the basis of hat functions according to the triangulation τ_k and $\mathbb{V}_k = \mathrm{span}\Phi_k$ be the corresponding finite element space.
- Let $a(\cdot,\cdot):\mathbb{V}_k\times\mathbb{V}_k\mapsto\mathbb{R}$ be a symmetric and positive definite bilinear form and let
$$\| u \|_{a,\Omega} = a(u,u)$$
be the energetic norm. In the same way, let
$$\| u \|_{a,\tilde{\Omega}} = a\,|_{\tilde{\Omega}}\,(u,u)$$
be the restriction of the norm onto a subdomain $\tilde{\Omega}\subset\Omega$.
- For $j=0,\ldots,k-2$, let $\mathbb{Y}_j = \{u\in\mathbb{V}_k : \mathrm{supp}\ u\subset\overline{\Omega}_j\cup\overline{\Omega}_{j+1}\}$ be the restriction of the finite element space $\mathbb{V}_k$ onto $\overline{\Omega}_j\cup\overline{\Omega}_{j+1}$ with Dirichlet boundary conditions at the boundaries Γ_j and Γ_{j+2}. For $j=k-1$, we set $\mathbb{Y}_{k-1} = \{u\in\mathbb{V}_k : \mathrm{supp}\ u\subset\overline{\Omega}_{k-1}\}$.

- Let

$$\| w \|^2_{\Gamma_j,\text{left}} = \min_{\substack{u \in \mathbb{V}_k \\ u|_{\Gamma_j} = w \\ u|_{\Gamma_{j+1}} = 0}} \| u \|^2_{a,\Omega_j} \quad \text{and} \quad \| w \|^2_{\Gamma_j,\text{right}} = \min_{\substack{u \in \mathbb{V}_k \\ u|_{\Gamma_j} = w \\ u|_{\Gamma_{j-1}} = 0}} \| u \|^2_{a,\Omega_{j-1}} \tag{5.4}$$

be the left and right trace norm on Γ_j.

Theorem 5.3. *Let all assumptions be satisfied. Then, for all decompositions of u into u_j, the assertion*

$$a(u,u) \leq \ 2\sum_{j=0}^{k-1} a(u_j,u_j) \quad \forall u = \sum_{j=0}^{k-1} u_j, \quad \textit{where } u_j \in \mathbb{Y}_j$$

holds.

Proof. The proof is simple. Due to the construction of the spaces $\mathbb{Y}_j$, we have

$$a(u,v) = 0 \quad \forall u \in \mathbb{Y}_j, v \in \mathbb{Y}_{j'}, \ |j-j'| > 1.$$

Using the Cauchy inequality and the arithmetical-geometrical mean value, we can conclude that

$$\begin{aligned} a(u,u) = \sum_{j,j'=1}^{k-1} a(u_j,u_{j'}) &= \sum_{j=0}^{k-1} a(u_j,u_j) + 2\sum_{j=0}^{k-2} a(u_j,u_{j+1}) \\ &\leq \sum_{j=0}^{k-1} a(u_j,u_j) + 2\sum_{j=0}^{k-2} \sqrt{\| u_j \|^2_{a,\Omega_{j+1}} \| u_{j+1} \|^2_{a,\Omega_{j+1}}} \\ &\leq \sum_{j=0}^{k-1} a(u_j,u_j) + \sum_{j=0}^{k-2} (\| u_j \|^2_{a,\Omega_{j+1}} + \| u_{j+1} \|^2_{a,\Omega_{j+1}}) \\ &\leq 2\sum_{j=0}^{k-1} a(u_j,u_j). \end{aligned}$$

This proves the assertion. □

Theorem 5.4. *In addition to the above assumptions, let us assume the following:*
There exists a integer j_0 such that

- *There exists a constant $\gamma < 1$ which is independent of the discretization parameter and j such that*

$$a(u,v) \leq \gamma \| u \|_{a,\Omega_{j+1}} \| v \|_{a,\Omega_{j+1}} \quad \forall j = 0,\ldots,j_0, \quad \forall u \in \mathbb{Y}_j, \forall v \in \mathbb{Y}_{j+1}. \tag{5.5}$$

- *There exists a constant $q < 1$ and a constant c_2 which are independent of j and the discretization parameter such that*

$$q^{-1} \| w \|^2_{\Gamma_j,\text{left}} \leq \| w \|^2_{\Gamma_j,\text{right}} \leq c_2 \| w \|^2_{\Gamma_j,\text{left}} \quad \forall w, \quad j = j_0+1,\ldots,k-1. \tag{5.6}$$

- *There exists a constant c_1 which is independent of discretization parameter such that*

$$c_1^{-1} \| w \|^2_{\Gamma_j,\text{left}} \leq \| w \|^2_{\Gamma_j,\text{right}} \leq c_2 \| w \|^2_{\Gamma_j,\text{left}} \quad \forall w, \quad j = j_0. \tag{5.7}$$

Then, there exists a decomposition $u = \sum_{j=0}^{k-1} u_j$ *with* $u_j \in \mathbb{Y}_j$ *such that*

$$c_L^2 \sum_{j=0}^{k-1} a(u_j, u_j) \le a(u,u) \quad \forall u \in \mathbb{V}_k.$$

The constant $c_L > 0$ *depends only on* γ, c_1, c_2 *and* q.

Proof. We construct an explicit decomposition using extension operators $\mathcal{T}_{j,\mathrm{left/right}}$ and start the proof with the definition of extension operators $\mathcal{T}_{j,\mathrm{left}}$. For a given function $w \in \mathbb{V}_k\,|_{\Gamma_j}$, let $\mathcal{T}_{j,\mathrm{left}} : \mathbb{V}_k\,|_{\Gamma_j} \mapsto \mathbb{V}_k\,|_{\Omega_j}$ be defined by the following conditions:

$$\begin{aligned} &\mathcal{T}_{j,\mathrm{left}} w \in \mathbb{V}_k\,|_{\overline{\Omega}_j}, && \mathcal{T}_{j,\mathrm{left}} w\,|_{\Gamma_j} = w \\ &\mathrm{supp}\ \mathcal{T}_{j,\mathrm{left}} w \subset \overline{\Omega}_j && a(\mathcal{T}_{j,\mathrm{left}} w, v) = 0 \quad \forall v \in \mathbb{V}_k, \mathrm{supp}\ v \subset \overline{\Omega}_j. \end{aligned} \tag{5.8}$$

Due to this definition, the operator $\mathcal{T}_{j,\mathrm{left}}$ is the discrete energetic extension from Γ_{j+1} to Ω_{j+1}. Moreover, let $\mathcal{T}_{j,\mathrm{right}} : \mathbb{V}_k\,|_{\Gamma_j} \mapsto \mathbb{V}_k\,|_{\Omega_{j-1}}$ be defined by the following conditions:

$$\begin{aligned} &\mathcal{T}_{j,\mathrm{right}} w \in \mathbb{V}_k\,|_{\overline{\Omega}_{j-1}}, && \mathcal{T}_{j,\mathrm{right}} w\,|_{\Gamma_j} = w \\ &\mathrm{supp}\ \mathcal{T}_{j,\mathrm{right}} w \subset \overline{\Omega}_{j-1} && a(\mathcal{T}_{j,\mathrm{right}} w, v) = 0 \quad \forall v \in \mathbb{V}_k, \mathrm{supp}\ v \subset \overline{\Omega}_j. \end{aligned} \tag{5.9}$$

Due to this definition, the operator $\mathcal{T}_{j,\mathrm{right}}$ is the discrete energetic extension from Γ_j to Ω_{j-1}.
We decompose a given $u \in \mathbb{V}_k$ into the functions $u_j \in \mathbb{Y}_j$ into the following way:

$$\begin{aligned} u_0 &= \left\{ \begin{array}{ccc} u & \text{in} & \overline{\Omega}_0 \\ \mathcal{T}_{0,\mathrm{left}} w_1 & \text{in} & \overline{\Omega}_1 \\ 0 & & \text{else} \end{array} \right. , \\ u_j &= \left\{ \begin{array}{ccc} u - u_{j-1} & \text{in} & \overline{\Omega}_j \\ \mathcal{T}_{j,\mathrm{left}} w_{j+1} & \text{in} & \overline{\Omega}_{j+1} \\ 0 & & \text{else} \end{array} \right. , j = 1, \ldots, k-2, \\ u_{k-1} &= \left\{ \begin{array}{ccc} u - u_{k-2} & \text{in} & \overline{\Omega}_{k-1} \\ 0 & & \text{else} \end{array} \right. \end{aligned} \tag{5.10}$$

with

$$w_j = u\,|_{\Gamma_j}, \quad j = 1, \ldots, k-1, \quad \text{and} \quad w_k = 0.$$

Due to the construction of the functions u_j, the function u_j belongs to $\mathbb{Y}_j$. We consider now the strengthened Cauchy inequality between the spaces $\mathbb{Y}_j$ and $\mathbb{Y}_{j+1}$, i.e.

$$\tilde{\gamma}_j^2 = \max_{\substack{u_j \in \mathbb{Y}_j \\ u_{j+1} \in \mathbb{Y}_{j+1} \\ u_j, u_{j+1} \neq 0}} \frac{a^2(u_j, u_{j+1})}{a(u_j, u_j)\, a(u_{j+1}, u_{j+1})}. \tag{5.11}$$

We will prove that it is possible to restrict ourselves to the traces of the functions u_j on Γ_{j+1}. For a given trace function w_j, let

$$g_j = \left\{ \begin{array}{ccc} \mathcal{T}_{j,\mathrm{left}} w_j & \text{in} & \Omega_j \\ \mathcal{T}_{j,\mathrm{right}} w_j & \text{in} & \Omega_{j-1} \\ 0 & \text{else} & \end{array} \right\} \quad \in \mathbb{Y}_{j-1},\ j = 1, \ldots, k,$$

with $\mathcal{T}_{j,\text{right}}$ of (5.9). Due to the construction of the function g_j, we can conclude that

$$\| g_j \|_{a,\Omega_j} = \| w_j \|_{\Gamma_j,\text{left}} \quad \text{and} \quad \| g_j \|_{a,\Omega_{j-1}} = \| w_j \|_{\Gamma_j,\text{right}}, j = 1, \dots, k. \tag{5.12}$$

Moreover,

$$\begin{array}{ll} u_{j+1} = u_{j+1,I} + g_{j+2}, & \text{where} \\ \quad u_{j+1,I} \mid_{\Gamma_{j+2}} = 0, \quad u_{j+1} \mid_{\Gamma_{j+2}} = g_{j+2}, & a(g_{j+2}, u_{j+1,I}) = 0, \quad j = 0, \dots, k-2. \end{array}$$

A direct consequence of $a(g_{j+2}, u_{j+1,I}) = 0$ and (5.12) is the relation

$$\begin{aligned} a(u_{j+1}, u_{j+1}) &= a(g_{j+2}, g_{j+2}) + a(u_{j+1,I}, u_{j+1,I}) \\ &= \| g_{j+2} \|^2_{a,\Omega_{j+2}} + \| g_{j+2} \|^2_{a,\Omega_{j+1}} + a(u_{j+1,I}, u_{j+1,I}) \\ &= \| w_{j+2} \|^2_{\Gamma_{j+2},\text{left}} + \| w_{j+2} \|^2_{\Gamma_{j+2},\text{right}} + a(u_{j+1,I}, u_{j+1,I}), \quad j = 0, \dots, k-2. \end{aligned} \tag{5.13}$$

With the same arguments, we obtain

$$\begin{aligned} a(u_j, u_j) &= a(g_{j+1}, g_{j+1}) + a(u_{j,I}, u_{j,I}) \\ &= \| w_{j+1} \|^2_{\Gamma_{j+1},\text{left}} + \| w_{j+1} \|^2_{\Gamma_{j+1},\text{right}} + a(u_{j,I}, u_{j,I}), \quad j = 0, \dots, k-1. \end{aligned} \tag{5.14}$$

Since $a(\mathcal{T}_{j+1} w, u_{j+1,I}) = 0$, we can conclude with the strengthened Cauchy inequality and (5.12) that

$$\begin{aligned} 2|a(u_j, u_{j+1})| &= 2\left|a\mid_{\Omega_{j+1}}(u_j, u_{j+1})\right| = 2\left|a\mid_{\Omega_{j+1}}(\mathcal{T}_{j+1} w_j, u_{j+1})\right| \\ &= 2\left|a\mid_{\Omega_{j+1}}(g_{j+1}, g_{j+2} + u_{j+1,I})\right| \\ &= 2\left|a\mid_{\Omega_{j+1}}(g_{j+1}, g_{j+2})\right| \\ &\le 2\gamma_j \| g_{j+1} \|_{a,\Omega_{j+1}} \| g_{j+2} \|_{a,\Omega_{j+1}} \\ &= 2\gamma_j \sqrt{\beta_j} \| w_{j+1} \|_{\Gamma_{j+1},\text{left}} \sqrt{\beta_j^{-1}} \| w_{j+2} \|_{\Gamma_{j+2},\text{right}} \\ &\le \beta_j\gamma_j \| w_{j+1} \|^2_{\Gamma_{j+1},\text{left}} + \beta_j^{-1}\gamma_j \| w_{j+2} \|^2_{\Gamma_{j+2},\text{right}}, \quad j = 0, \dots, k-2 \end{aligned} \tag{5.15}$$

with some positive parameters β_j specified later. The constant γ_j denotes the constant of strengthened Cauchy inequality, which satisfies the estimate $\gamma_j \le 1$, $j = 0, \dots, k-2$.
We use now the estimates (5.13), (5.14) and (5.15) and can conclude

$$\begin{aligned} a(u,u) &= \sum_{j=0}^{k-1} a(u_j, u_j) + 2\sum_{j=0}^{k-2} a(u_j, u_{j+1}) \\ &\ge \sum_{j=0}^{k-1} a(u_{j,I}, u_{j,I}) + \sum_{j=0}^{k-2} \| w_{j+1} \|^2_{\Gamma_{j+1},\text{left}} + \| w_{j+1} \|^2_{\Gamma_{j+1},\text{right}} \\ &\quad - \left(\sum_{j=0}^{k-2} \beta_j\gamma_j \| w_{j+1} \|^2_{\Gamma_{j+1},\text{left}} + \beta_j^{-1}\gamma_j \| w_{j+2} \|^2_{\Gamma_{j+2},\text{right}} \right) \\ &\ge \sum_{j=0}^{k-2} (1 - \beta_j\gamma_j) \| w_{j+1} \|^2_{\Gamma_{j+1},\text{left}} + (1 - \beta_{j-1}^{-1}\gamma_{j-1}) \| w_{j+1} \|^2_{\Gamma_{j+1},\text{right}} + \sum_{j=0}^{k-1} a(u_{j,I}, u_{j,I}). \end{aligned} \tag{5.16}$$

In the above estimate, we have set $\gamma_{-1} = 0$ per definition and used $w_k = 0$. Let

$$
\begin{aligned}
s_0 &:= \sum_{j=0}^{j_0-1} (1-\beta_j\gamma_j) \parallel w_{j+1} \parallel^2_{\Gamma_{j+1},\text{left}} + (1-\beta_{j-1}^{-1}\gamma_{j-1}) \parallel w_{j+1} \parallel^2_{\Gamma_{j+1},\text{right}}, \\
s_1 &:= (1-\beta_{j_0}\gamma_{j_0}) \parallel w_{j_0+1} \parallel^2_{\Gamma_{j_0+1},\text{left}} + (1-\beta_{j_0-1}^{-1}\gamma_{j_0-1}) \parallel w_{j_0+1} \parallel^2_{\Gamma_{j_0+1},\text{right}}, \\
s_2 &:= \sum_{j=j_0+1}^{k-2} (1-\beta_j\gamma_j) \parallel w_{j+1} \parallel^2_{\Gamma_{j+1},\text{left}} + (1-\beta_{j-1}^{-1}\gamma_{j-1}) \parallel w_{j+1} \parallel^2_{\Gamma_{j+1},\text{right}} .
\end{aligned} \tag{5.17}
$$

We define now

$$
\beta_j = \begin{cases} 1 & j \leq j_0 - 1 \\ \beta & j \geq j_0 \end{cases} \qquad \text{with} \quad \beta = \min\left\{ \sqrt{q^{-1}}, 1 + \frac{1-\gamma}{2c_1} \right\}.
$$

The assumption (5.5) implies $\gamma_j \leq \gamma < 1$ for $j \leq j_0$. Hence, we obtain the estimate

$$
s_0 \geq (1-\gamma) \sum_{j=0}^{j_0-1} \parallel w_{j+1} \parallel^2_{\Gamma_{j+1},\text{left}} + \parallel w_{j+1} \parallel^2_{\Gamma_{j+1},\text{right}} . \tag{5.18}
$$

The estimates $q < 1$, $\gamma < 1$ and $c_1 > 0$ imply $\beta > 1$ and

$$
(1-\gamma) + c_1(1-\beta) \geq \frac{1}{2}(1-\gamma).
$$

Using assumption (5.7), we can estimate

$$
\begin{aligned}
s_1 &\geq \left((1-\beta_{j_0}\gamma_{j_0}) + c_1^{-1}(1-\beta_{j_0-1}^{-1}\gamma_{j_0-1})\right) \parallel w_{j_0+1} \parallel^2_{\Gamma_{j_0+1},\text{left}} \\
&\geq \frac{1}{c_1}\left((1-\gamma) + c_1(1-\beta)\right) \parallel w_{j_0+1} \parallel^2_{\Gamma_{j_0+1},\text{left}} \\
&\geq \frac{1-\gamma}{2c_1} \parallel w_{j_0+1} \parallel^2_{\Gamma_{j_0+1},\text{left}} \geq \frac{1-\gamma}{2c_1}\frac{1}{1+c_2}\left(\parallel w_{j_0+1} \parallel^2_{\Gamma_{j_0+1},\text{left}} + \parallel w_{j_0+1} \parallel^2_{\Gamma_{j_0+1},\text{right}} \right).
\end{aligned}
$$

Finally, we estimate s_2. The assumption (5.6), $\beta > 1$ and $\gamma_j \leq 1$ imply

$$
s_2 \quad \geq \quad \left(1 - \beta + (1-\beta^{-1})\frac{1}{q}\right) \sum_{j=j_0+1}^{k-2} \parallel w_{j+1} \parallel^2_{\Gamma_{j+1},\text{left}} . \tag{5.19}
$$

The estimate $\beta \leq \sqrt{q^{-1}}$ implies $\frac{1}{q} \geq \beta^2 \geq 1$. Hence, we can estimate

$$
\left(1 - \beta + (1-\beta^{-1})\frac{1}{q}\right) \quad \geq \quad (1-\beta) + (1-\beta^{-1})\beta^2 = (1-\beta)^2 > 0.
$$

Inserting this estimate into (5.19) and using (5.7) yields

$$
s_2 \geq \frac{(1-\beta)^2}{1+c_2} \sum_{j=j_0+1}^{k-2} \left(\parallel w_{j+1} \parallel^2_{\Gamma_{j+1},\text{left}} + \parallel w_{j+1} \parallel^2_{\Gamma_{j+1},\text{right}} \right). \tag{5.20}
$$

Now, we insert (5.18), (5.19) and (5.20) into (5.16) and obtain

$$\begin{aligned} a(u,u) &\geq \min\left\{\frac{(1-\beta)^2}{1+c_2}, \frac{1-\gamma}{2c_1}\frac{1}{1+c_2}\right\}\sum_{j=0}^{k-2}\Big(\| w_{j+1} \|^2_{\Gamma_{j+1},\mathrm{left}} + \| w_{j+1} \|^2_{\Gamma_{j+1},\mathrm{right}}\Big) \\ &\quad + \sum_{j=0}^{k-1} a(u_{j,I},u_{j,I}) \\ &\geq \min\left\{\frac{(1-\beta)^2}{1+c_2}, \frac{1-\gamma}{2c_1}\frac{1}{1+c_2}, 1\right\}\sum_{j=0}^{k-1}\left(a(g_j,g_j)+a(u_{j,I},u_{j,I})\right) \\ &= \min\left\{\frac{(1-\beta)^2}{1+c_2}, \frac{1-\gamma}{2c_1}\frac{1}{1+c_2}, 1\right\}\sum_{j=0}^{k-1} a(u_j,u_j). \end{aligned}$$

This proves the theorem. □

Remark 5.5. *The proof shows that the constant c_L can be chosen as*

$$c_L^2 = \min\left\{\frac{\left(1-\min\left\{\sqrt{q^{-1}}, 1+\frac{1-\gamma}{2c_1}\right\}\right)^2}{1+c_2}, \frac{1-\gamma}{2c_1}\frac{1}{1+c_2}, 1\right\}.$$

We will finish this subsection with two generalizations of Theorem 5.4 in which we have replaced the assumptions (5.5)-(5.7) by another ones.

Remark 5.6. *The result of Theorem 5.4 remains valid if the assumptions* (5.5)-(5.7) *are replaced by the following assumption. There exists a constant $q<1$ and a constant c_2 which are independent of j and the discretization parameter such that*

$$q^{-1} \| w \|^2_{\Gamma_j,\mathrm{right}} \leq \| w \|^2_{\Gamma_j,\mathrm{left}} \leq c_2 \| w \|^2_{\Gamma_j,\mathrm{right}} \quad \forall w, \quad j=1,\ldots,k-1. \tag{5.21}$$

Proof. The only difference in the proof is the estimate for s_2 in (5.20). Using the definition of s_2 given (5.17), we can conclude that

$$s_2 = \sum_{j=1}^{k-2}(1-\beta_j) \| w_{j+1} \|^2_{\Gamma_{j+1},\mathrm{left}} + (1-\beta_{j-1}^{-1}) \| w_{j+1} \|^2_{\Gamma_{j+1},\mathrm{right}} .$$

In contrast to (5.19), we choose now $\beta_j=\beta<1$ and estimate $\| w_{j+1} \|^2_{\Gamma_{j+1},\mathrm{left}}$ by the left inequality of (5.21). This gives

$$s_2 \geq (1-\beta)q^{-1} + 1 - \beta^{-1}) \| w_{j+1} \|^2_{\Gamma_{j+1},\mathrm{right}} .$$

We choose now $\beta=\sqrt{q}$ and obtain

$$s_2 \geq (1-\beta)^2 \| w_{j+1} \|^2_{\Gamma_{j+1},\mathrm{right}} .$$

This result is similar to (5.20) and finishes the proof. □

Remark 5.7. *Let us assume that instead of assumption* (5.6) *and* (5.7) *the following assumption holds. Let $a_2(\cdot,\cdot)$ be another bilinear form which satisfies the relation*

$$c_3\, a_2 |_{\Omega_j} (u,u) \leq a |_{\Omega_j} (u,u) \leq c_4\, a_2 |_{\Omega_j} (u,u) \quad \forall u \in \mathbb{X}_n, j=j_0+1,\ldots,k-1. \tag{5.22}$$

Moreover, there exists a decomposition of $u = \sum_{j=j_0}^{k-1} \tilde{u}_j$, $\tilde{u}_j \in \mathbb{Y}_j$ *such that*

$$a_2 |_{\Omega_l} (u,u) = c_5^2 \sum_{j=j_0}^{k-1} a_2 |_{\Omega_l} (\tilde{u}_j, \tilde{u}_j) \quad \forall u \in \mathbb{X}_n \tag{5.23}$$

with $\Omega_L = \bigcup_{j=j_0+1}^{k-1} \Omega_j$. *The constants* $c_3 > 0$, $c_4 > 0$ *and* $c_5 > 0$ *do not depend on the discretization parameter or* j. *Then, there exists a decomposition* $u = \sum_{j=0}^{k-1} u_j$ *with* $u_j \in \mathbb{Y}_j$ *such that*

$$c_L^2 \sum_{j=0}^{k-1} a(u_j, u_j) \le a(u,u) \quad \forall u \in \mathbb{V}_k.$$

The constant $c_L > 0$ *depends only on* γ, c_3, c_4 *and* c_5.

Proof. Instead of the decomposition $u = \sum_{j=0}^{k-1} u_j$ given in (5.10), we take the decomposition

$$u = \sum_{j=0}^{j_0-1} u_j + \sum_{j=j_0}^{k-1} \tilde{u}_j$$

with $\tilde{u}_j$ be defined via (5.23). Using (5.5), we can estimate

$$a(u,u) \ge (1-\gamma) \sum_{j=0}^{j_0-2} a(u_j,u_j) + (1-\gamma) a |_{\Omega_{j_0}} (u_{j_0}, u_{j_0}) + a |_{\Omega_L} (\hat{u}, \hat{u}). \tag{5.24}$$

Since $\hat{u} \in \mathbb{Y}_{j_0+1} \oplus \ldots \oplus \mathbb{Y}_{k-1}$, we can estimate the right sum in (5.24) and obtain

$$\begin{aligned} a |_{\Omega_L} (\hat{u}, \hat{u}) \ge c_3 a_2 |_{\Omega_L} (\hat{u}, \hat{u}) \quad &\ge \quad c_3 c_5^2 \left(\sum_{j=j_0+1}^{k-1} a_2(\tilde{u}_j, \tilde{u}_j) + a_2 |_{\Omega_{j_0+1}} (u_{j_0}, u_{j_0}) \right) \\ &\ge \quad c_3 c_5^2 c_4^{-1} \left(\sum_{j=j_0+1}^{k-1} a(\tilde{u}_j, \tilde{u}_j) + a |_{\Omega_{j_0+1}} (u_{j_0}, u_{j_0}) \right) \end{aligned} \tag{5.25}$$

by using the left inequality of (5.22), (5.23) and the right inequality of (5.22). Inserting (5.25) into (5.24) proves the assertion with $c_L^2 = \min\{1-\gamma, c_3 c_5^2 c_4^{-1}\}$. □

5.3 Transformation of the continuous bilinear form to a bilinear form with piecewise constant coefficients

We define now a bilinear form $a_p(\cdot,\cdot)$ with piecewise constant coefficients. The energetic norm of this bilinear form will define a spectrally equivalent norm to the energetic norm of the original bilinear form $a(\cdot,\cdot)$ (2.1). Let

$$\kappa^2(\xi) = \{ \ \varepsilon_j, \quad \xi \in (2^{-j-1}, 2^{-j}) \qquad \text{with} \quad \varepsilon_j := \omega^2(2^{-j}) \tag{5.26}$$

be an piecewise constant coefficient function and

$$a_p(u,v) := \int_\Omega (\nabla v)^T \cdot \begin{bmatrix} \kappa^2(x) & 0 \\ 0 & \kappa^2(x) \end{bmatrix} \nabla u \tag{5.27}$$

be the corresponding bilinear form. Moreover, let

$$\| u \|_p^2 := a_p(u,u) \quad \forall u \in \mathbb{V}_k \tag{5.28}$$

be the energetic norm with respect to the bilinear form $a_p(\cdot,\cdot)$. The stiffness matrix with respect to the basis Φ_k is denoted by $K_{k,p}$, i.e.

$$K_{k,p} = [a_p(\phi_{il}, \phi_{i'l'})]_{i,l=1;\, i',l'=1}^{n_0,n_0} . \tag{5.29}$$

Lemma 5.8. *Let us assume that the weight function ω satisfies Assumption 2.1. Then, we have*

$$a(u,u) \le a_p(u,u) \le 2c_\omega^2 a(u,u) \quad \forall u \in \mathbb{V}_l.$$

The constant c_ω is from (2.4).

Proof. Using the definition of the weight function κ and the monotonicity of ω, we have

$$\omega(\xi) \le \kappa(\xi) \quad \forall \xi \in (0,1).$$

This gives the lower estimate. To prove the upper estimate, we can conclude

$$\kappa(2\xi) \le c_\omega \kappa(\xi) \le c_\omega \omega(2\xi) \tag{5.30}$$

from the monotonicity of ω and (2.2). This gives the upper estimate with $c = c_\omega^2$ for $x > 2^{-k}$.
On the left strip with $x < 2^{-k}$, all functions $u \in \mathbb{V}_k$ are linear. Hence, the gradient is constant. We use the estimate (5.30) again and obtain

$$a\mid_{\{x<2^{-k}\}}(u,u) \ge a\mid_{\{2^{-k-1}<x<2^{-k}\}}(u,u) \ge c_\omega^{-2} a_p\mid_{\{2^{-k-1}<x<2^{-k}\}}(u,u) \ge \frac{1}{2}c_\omega^{-2} a_p\mid_{\{x<2^{-k}\}}(u,u).$$

This finishes the proof.

□

Remark 5.9.

- *In the case of the weight function $\omega^2(\xi) = \xi^\alpha$ with $\alpha > 0$, we have $c_\omega^2 = 2^\alpha$.*
- *A direct consequence of* (2.4), (5.29) *and Lemma 5.8 is the eigenvalue estimate*

$$\frac{1}{2}c_\omega^{-2} K_{k,p} \le K_k \le K_{k,p}.$$

5.4 A result for tridiagonal matrices

Finally, an estimate for tridiagonal matrices with constant main- and subdiagonals is required. For a fixed integer m and some positive parameter κ, we introduce

$$\begin{aligned} F_m &= m \begin{bmatrix} 2+\kappa & -1 & & & \mathbf{0} \\ -1 & 2+\kappa & -1 & & \\ & \ddots & \ddots & \ddots & \\ \mathbf{0} & & -1 & 2+\kappa & -1 \\ & & & -1 & 2+\kappa \end{bmatrix} \in \mathbb{R}^{m-1\times m-1} \\ \tilde{F}_m &= m \begin{bmatrix} 1+\frac{\kappa}{2} & e_1^T \\ e_1 & F_m \end{bmatrix} \in \mathbb{R}^{m\times m}, e_1 = (1,0\ldots,0)^T \in \mathbb{R}^{m-1\times 1}. \end{aligned} \tag{5.31}$$

Then, the following result holds.

Lemma 5.10. *Let F_m and $\tilde{F}_m$ be defined via* (5.31)*. Let $s_m = 1 + \frac{\kappa}{2} - e_1^T F_m^{-1} e_1$ be the Schur complement of $\tilde{F}_m$ with respect to the first row and column and $\hat{s}_m = e_1^T F_m^{-1} e_m$, $e_m = (0, \ldots, 0, 1)^T$ and $\gamma_m = \frac{|\hat{s}_m|}{s_m}$. For $m \geq \{\frac{1}{\sqrt{\kappa}}, 2\}$, the estimate*

$$\gamma_m \leq \frac{20}{21} \tag{5.32}$$

is valid.

Proof. The proof is elementary and given in [3]. □

6 Condition number estimates

In this section, we will prove the main results. We start with the proof of Theorem 3.2 in subsections 6.1 and 6.2 for $\alpha > 1$ and $\alpha < 1$, respectively.
After that, we prove Theorem 3.1. Here, we introduce a nonoverlapping preconditioner C_{non} for K_k and prove $C_{non} \sim K_k$ in subsection 6.3. In subsection 6.4, we simplify this preconditioner and obtain the main result.
All results will be proved for the matrix $K_{k,p}$ (5.29). By Lemma 5.8, the result follows for the matrix K_k.

6.1 The modified overlapping preconditioner for $\alpha > 1$

In this subsection, we give the proof of Theorem 3.2 for $\alpha > 1$. We will apply Lemma 5.1. Therefore, we have to verify the assumptions (5.1), (5.2) and (5.3).
In a first step, we introduce two trace norms for functions on $\Gamma_{j,x}$. Let

$$\| w \|^2_{\Gamma_{j,x},\text{left}} = \min_{\substack{u \in \mathbb{V}_k \\ u|_{\Gamma_{j,x}} = w \\ u|_{\Gamma_{j+1,x}} = 0}} |u|^2_{1,\Omega_{j,x}} \quad \text{and} \quad \| w \|^2_{\Gamma_{j,x},\text{right}} = \min_{\substack{u \in \mathbb{V}_k \\ u|_{\Gamma_{j,x}} = w \\ u|_{\Gamma_{j-1,x}} = 0}} |u|^2_{1,\Omega_{j-1,x}}. \tag{6.1}$$

Now, we prove the following result.

Lemma 6.1. *The spectral equivalence relations*

$$\| w \|^2_{\Gamma_{j,x},\text{left}} \leq 2 \| w \|^2_{\Gamma_{j,x},\text{right}} \leq 2 \| w \|^2_{\Gamma_{j,x},\text{left}} \quad \forall w \in \mathbb{V}_k |_{\Gamma_{j,x}} . \tag{6.2}$$

hold.

Proof. Let us start with the following observation. On the left domain $\Omega_{j,x}$, we have a layer of m triangles and on the right domain $\Omega_{j-1,x}$ we have a layer of $2m$ triangles. Let $\mathcal{T}_{\text{left}}$ and $\mathcal{T}_{\text{right}}$ be the discrete harmonic extension of a function on $\Gamma_{j,x}$ to $\Omega_{j,x}$ and $\Omega_{j-1,x}$ respectively. The function $u = \mathcal{T}_{\text{left}} w$ is uniquely defined by its values at the nodes $2^{-k}(i,s)$, $i = 2^{k-j-1}, \ldots, 2^{k-j}$, $s = 0, \ldots, 2^k$. We write simply u_{rs} with $r = 2^{k-j} - i'$, $r = 0, 1, \ldots, m$ for this value. So, the first index r corresponds to the distance of layers in x-direction to $\Gamma_{j,x}$. In the same way, we introduce v_{rs}, $r = 0, \ldots, 2m$, $s = 0, \ldots, 2^k$ for the nodal values which correspond to $v = \mathcal{T}_{\text{right}} w$. Again, the first index r corresponds to the distance of layers to $\Gamma_{j,x}$. Then, we can conclude

$$\begin{aligned} \| w \|^2_{\Gamma_{j,x},\text{left}} = |u|^2_{1,\Omega_{j,x}} &= 2^{-2k} \sum_{s=0}^{2^k-1} \sum_{r=0}^{m-1} (u_{r+1,s} - u_{r,s})^2 + (u_{r,s+1} - u_{r,s})^2 \\ &\leq 2^{-2k} \sum_{s=0}^{2^k-1} \sum_{r=0}^{m-1} (v_{2r+2,s} - v_{2r,s})^2 + (v_{2r,s+1} - v_{2r,s})^2 \end{aligned}$$

using the optimality of the extension $\mathcal{T}_{\text{left}}$. Next, we use the simple inequality $(a+b)^2 \leq 2a^2 + 2b^2$ for the first sum of the right hand side and obtain

$$\sum_{s=0}^{2^k-1}\sum_{r=0}^{m-1}(v_{2r+2,s}-v_{2r,s})^2 = \sum_{s=0}^{2^k-1}\sum_{r=0}^{m-1}(v_{2r+2,s}-v_{2r+1,s}+v_{2r+1}-v_{2r,s})^2 \leq 2\sum_{s=0}^{2^k-1}\sum_{r=0}^{2m-1}(v_{r+1,s}-v_{r,s})^2.$$

Hence, we can conclude that

$$\| w \|^2_{\Gamma_{j,x},\text{left}} \leq 2\; 2^{-2k}\sum_{s=0}^{2^k-1}\sum_{r=0}^{2m-1}(v_{r+1,s}-v_{r,s})^2 + (v_{2r,s+1}-v_{2r,s})^2 = 2 \| w \|^2_{\Gamma_{j,x},\text{right}} .$$

This proves the lower inequality in (6.2). The upper inequality is proved in the same way starting with the minimality for the H^1-seminorm of $\mathcal{T}_{\text{right}}$. □

Now, we able to prove Theorem 3.2 for $\alpha > 1$.

Proof. Using Lemma 5.8 it suffices to prove the result for the matrix $K_{k,p}$ (5.29). We apply Lemma 5.1 and verify the assumptions. For the weight function $\omega^2(\xi) = \xi^\alpha$, $\alpha > 1$, assumption (5.3) is valid with $\underline{c} = 2^{-\alpha}$ and $\overline{c} = 1$. Due to Lemma 5.3, assumption (5.2) is valid with $\beta = 2$.
Let $\omega^2(\xi) = \xi^\alpha$ with $\alpha > 1$. By (6.2), assumption (5.5) holds with $q = 2^{1-\alpha} < 1$ for all j. Therefore, we can apply Theorem 5.4, which gives us (5.1). This proves Theorem 3.2. □

6.2 The modified overlapping preconditioner for $\alpha < 1$.

In the case $\alpha < 1$, we have to modify the proof. This proof will use the tensor-product-structure directly and requires three steps,

1. the stability of a decomposition in the 1D-case,
2. the stability of a decomposition with possibly dominating mass term in the 1D case,
3. the proof of the two dimensional case based on tensor product arguments and 2.

In order to prove the two-dimensional result by tensor-product arguments, we have to investigate the following model problem: For $n = 2^k$, let $\tau_s^n = \left(\frac{s}{n}, \frac{s+1}{n}\right)$, $s = 0, \ldots, n-1$, be a partition of the interval $(0,1)$. Let $\mathbb{X}_n = \text{span}[\phi_s^n]_{s=1}^{n-1} = \text{span}[\Phi_1]$ be the basis of the one-dimensional hat functions on this partition given by

$$\phi_s^n(x) = \begin{cases} nx-(s-1) & \text{on } \tau_{s-1}^n, \\ (s+1)-nx & \text{on } \tau_s^n, \\ 0 & \text{otherwise,} \end{cases} \qquad s = 1, \ldots, n-1. \tag{6.3}$$

Moreover, let

$$a_{1,\lambda} = \int_0^1 \kappa^2(x)u'(x)v'(x)\,\mathrm{d}x + \lambda\, n\sum_{i=1}^{n-1}\rho_s u\left(\frac{s}{n}\right)v\left(\frac{s}{n}\right) \quad \text{and} \quad \| u \|_{1,\lambda} = a_{1,\lambda}(u,u) \tag{6.4}$$

with $\rho_s = \frac{1}{2}\left[\kappa^2(x)\mid_{\tau_{s-1}^n} + \kappa^2(x)\mid_{\tau_s^n}\right]$ and some nonnegative parameter λ be a bilinear form on $\mathbb{X}_n \times \mathbb{X}_n$, and the energetic norm, respectively. Due to $\kappa^2(x) > 0$ for $x \in (0,1)$, this bilinear form is symmetric and coercive.

For $j = 0, \ldots, k-2$, let $\Omega_j = \left(2^{-j+1}, 2^{-j}\right)$ and $\Omega_{k-1} = (0, 2^{k-1})$. Moreover, we introduce

$$\begin{aligned} \tilde{\mathbb{W}}_j &= \operatorname{span}\{\phi_i^n\}_{i=n_{j+1}+2}^{n_j}, \quad j = 0, \ldots, k-1, \quad \text{and} \\ \mathbb{W}_j &= \operatorname{span}\{\phi_i^n\}_{i=n_{j+2}+2}^{n_j}, \quad j = 0, \ldots, k-2, \quad \mathbb{W}_{k-1} = \tilde{\mathbb{W}}_{k-1}. \end{aligned}$$

Due to this definition, the spaces $\mathbb{W}_j$ and $\tilde{\mathbb{W}}_j$ are formed by those hat functions (6.3) which have a support in $\overline{\Omega}_{j+1} \cup \overline{\Omega}_j$, and $\overline{\Omega}_j$, respectively.
Now, we prove the following result for $\lambda = 0$. This result is key for the proof in the one-dimensional case.

Lemma 6.2. *There exists a decomposition $u = \sum_{j=0}^{k-1} u_j$ with $u_j \in \mathbb{W}_j$ such that*

$$a_{1,0}(u,u) \geq c^2 \sum_{j=0}^{k-1} a_{1,0}(u_j, u_j) \quad \forall u \in \mathbb{X}_n.$$

The constant $c^2 > 0$ is constant which does not depend on n.

Proof. We will use Theorem 5.4 and Remark 5.6. So we adapt the notation of this theorem, i.e. let $\Gamma_{j+1} = \overline{\Omega}_{j+1} \cup \overline{\Omega}_j$ and

$$\| w \|^2_{\Gamma_j,\text{left}} = \min_{\substack{u \in \mathbb{X}_n \\ u|_{\Gamma_j} = w \\ u|_{\Gamma_{j+1}} = 0}} \| u \|^2_{a_{1,0},\Omega_j} \quad \text{and} \quad \| w \|^2_{\Gamma_j,\text{right}} = \min_{\substack{u \in \mathbb{X}_n \\ u|_{\Gamma_j} = w \\ u|_{\Gamma_{j-1}} = 0}} \| u \|^2_{a_{1,0},\Omega_{j-1}} . \tag{6.5}$$

We will verify now assumption (5.21) Since $\kappa^2(x)\ |_{\Omega_j} = \varepsilon_j$, i.e. the coefficient function is constant, it is possible to compute the norms in (6.5) explicitly. A straightforward computation shows that

$$\| w \|^2_{\Gamma_j,\text{left}} = \varepsilon_j 2^k 2^{j-k+1} w^2 \quad \text{and} \quad \| w \|^2_{\Gamma_j,\text{right}} = \varepsilon_{j-1} 2^k 2^{j-k+1} w^2, \quad w \in \mathbb{R}.$$

Therefore,

$$\frac{\| w \|^2_{\Gamma_j,\text{left}}}{\| w \|^2_{\Gamma_j,\text{right}}} = 2\frac{\varepsilon_j}{\varepsilon_{j-1}} = 2^{1-\alpha} > 1, \quad \alpha < 1.$$

This gives (5.21) with $q = 2^{\alpha-1} < 1$ and $c_2 = q^{-1}$. □

With the help of this lemma, one can finish the proof of Theorem 3.2 in the 1D-case. For the two-dimensional case, this result is required for arbitrary $\lambda \in \mathbb{R}$. This will be done in

Lemma 6.3. *There exists a decomposition $u = \sum_{j=0}^{k-1} u_j$ with $u_j \in \mathbb{W}_j$ such that*

$$a_{1,\lambda}(u,u) \geq c^2 \sum_{j=0}^{k-1} a_{1,\lambda}(u_j, u_j) \quad \forall u \in \mathbb{X}_n, \lambda > 0.$$

The constant $c^2 > 0$ is a constant which does not depend on n and λ.

Proof. Again, we adapt the notation of Theorem 5.4.
Moreover, let $m_j = 2^{k-j+1}$ be the number of elements inside Ω_j. Then, the series $\{m_j\}_j$ is monotonic decreasing. Therefore is exists a j_0 such that $m_{j_0-1}^{-2} \leq \lambda \leq m_{j_0}^{-2}$. Now, we verify the assumptions (5.5), (5.22) and (5.23).

If $j \leq j_0$, we have $\lambda \geq m_j^{-2}$. Since the coefficient functions before mass and stiffness term of the bilinear form $a_{1,\lambda}$ (6.4) are constant inside Ω_j, we can use the results of Lemma 5.10. Due to the properties of the Schur-complement, we have

$$\| w \|^2_{\Gamma_j,\mathrm{left}} = \| w \|^2_{\Gamma_{j+1},\mathrm{right}} = w^2 s_{m_j} \quad \forall w \in \mathbb{R}$$

with s_{m_j} of Lemma 5.10. A simple computation shows

$$a_1(\mathcal{T}_{j,\mathrm{left}} u, \mathcal{T}_{j+1,\mathrm{right}} v) = u \hat{s}_{m_j} v \quad \forall u, v \in \mathbb{R}$$

with $\hat{s}_m$ of Lemma 5.10. Hence, we can conclude that

$$\gamma^2_{m_j} = \max_{\substack{u,v \in \mathbb{R} \\ u,v \neq 0}} \frac{a_1(\mathcal{T}_{j,\mathrm{left}} u, \mathcal{T}_{j+1,\mathrm{right}} v)}{\| u \|_{\Gamma_j,\mathrm{left}} \| v \|_{\Gamma_{j+1},\mathrm{right}}} = \frac{\hat{s}_{m_j}}{s_{m_j}} < \frac{20}{21}.$$

Then, we obtain

$$\begin{aligned} a_{1,\lambda}|_{\Omega_j}(u,u) \quad &\geq \quad (1-\gamma)\left[a_{1,\lambda}|_{\Omega_j}(u_j,u_j) + a_{1,\lambda}|_{\Omega_j}(u_{j-1},u_{j-1})\right], \\ & \qquad u = u_{j-1} + u_j, u_j \in \mathbb{W}_j, u_{j-1} \in \mathbb{W}_{j-1}, \lambda \geq m_j^{-2} \end{aligned} \tag{6.6}$$

with $\gamma = \frac{20}{21}$. This gives (5.5).
If $j \geq j_0$, we have $\lambda \leq m_j^{-2}$. We use the constant coefficients before both terms of the bilinear term again. Then, we obtain

$$\frac{3}{2} a_{1,0}|_{\Omega_j}(u,u) \geq a_{1,\lambda}|_{\Omega_j}(u,u) \geq a_{1,0}|_{\Omega_j}(u,u), \quad \forall u \in \mathbb{X}_n, \quad \forall \lambda \leq m_j^{-2} \tag{6.7}$$

by a simple explicit computation. This gives (5.22) with $a_2(\cdot,\cdot) = a_{1,0}(\cdot,\cdot)$. Relation (5.23) is a consequence of Lemma 6.2.
Using Theorem 5.4 in combination with Remark 5.7, the assertion follows. □

We define now an overlapping preconditioner of the type (3.4) for the stiffness matrix which corresponds to the bilinear form $a_{1,\lambda}(\cdot,\cdot)$ (6.4). This matrix is expressed by the relation

$$\underline{u}^T A_\lambda \underline{u} \quad = \quad a_{1,\lambda}([\Phi_1]\underline{u}, [\Phi_1]\underline{u}). \tag{6.8}$$

Moreover, we denote the mass and stiffness part of the bilinear form (6.4) by

$$\underline{u}^T T_\omega \underline{u} = \int_0^1 \kappa^2(x) u'(x) u'(x)\,\mathrm{d}x, \quad \underline{u}^T M_\omega \underline{u} = \frac{1}{n}\sum_{i=1}^{n-1} \rho_s u\left(\frac{s}{n}\right) u\left(\frac{s}{n}\right), \quad u = [\Phi_1]\underline{u}. \tag{6.9}$$

Then, we have

$$A_\lambda = \lambda n^2 M_\omega + T_\omega \tag{6.10}$$

In order to define the overlapping preconditioner for A_λ, we have to introduce some auxiliary matrices. Let $I_n \in \mathbb{R}^{n \times n}$ be the identity matrix and

$$T_{n-1} \quad = \quad \begin{bmatrix} 2 & -1 & & & \mathbf{0} \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ \mathbf{0} & & & -1 & 2 \end{bmatrix} \in \mathbb{R}^{n-1 \times n-1} \tag{6.11}$$

be the one-dimensional Laplacian.
For $j = 0, \dots, k-2$, let

$$\begin{aligned} M_j &= \begin{bmatrix} \mathbf{0}_{n_{j+2}+1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \varepsilon_j I_{n_j - n_{j+2}-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0}_{n_0 - n_j} \end{bmatrix} \in \mathbb{R}^{n_0 \times n_0},\ i = 1, 2, \\ \Delta_{j,1} &= \begin{bmatrix} \mathbf{0}_{n_{j+2}+1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \varepsilon_j T_{n_j - n_{j+2}-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0}_{n_0 - n_j} \end{bmatrix} \in \mathbb{R}^{n_0 \times n_0}, \end{aligned}$$

where ε_j is defined via (5.26). For $j = k-1$, we set

$$M_{k-1} = \begin{bmatrix} \varepsilon_{k-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{n_0-1} \end{bmatrix} \in \mathbb{R}^{n_0 \times n_0},\ i = 1, 2, \qquad \Delta_{k-1,1} = \begin{bmatrix} 2\varepsilon_{k-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{n_0-1} \end{bmatrix} \in \mathbb{R}^{n_0 \times n_0}.$$

Now, we can define

$$C_1^{-1} = \sum_{j=0}^{k-1} (\lambda M_j + \Delta_{j,1})^+ \tag{6.12}$$

as preconditioner for A_λ. Now, we are able to formulate a summarizing lemma.

Lemma 6.4. *For $\lambda > 0$, let A_λ and C_1 be defined via* (6.10) *and* (6.12), *respectively. Moreover, let $\omega^2(\xi) = \xi^\alpha$, $0 \le \alpha < 1$. Then, $c_1 C_1 \le A_\lambda \le c_2 C_1$. The constants do not depend on the parameter λ and the discretization parameter.*

Proof. We apply Lemma 5.1 with the bilinear form $(\cdot,\cdot)_{\mathcal{A}} = a_{1,\lambda}(\cdot,\cdot)$ and verify the assumptions (5.1), (5.2) and (5.3). The space splitting implies $\beta = 2$, cf. Theorem 5.4, which proves (5.2). Relation (5.1) follows from Lemma 6.3.
The bilinear form $a_{1,\lambda}(\cdot,\cdot)$ (6.4) is the sum of two terms, a stiffness term and a mass term. The coefficient before both terms are piecewise constant, i.e. ε_j on Ω_j. So, the maximum of the coefficients on $\overline{\Omega}_j \cup \overline{\Omega}_{j+1}$ is ε_j and the minimum is ε_{j+1}. In the preconditioner C_1 (6.12), the coefficient on $\overline{\Omega}_j \cup \overline{\Omega}_{j+1}$ is replaced by ε_j. Assumption 2.1 implies that the ratio of coefficients $\varepsilon_{j+1}^{-1}\varepsilon_j$ is bounded. This gives (5.3) and proves the lemma for the matrix C_1. □

Finally, we prove Theorem 3.2 for $\alpha < 1$.

Proof. Due to Lemma 5.8, it suffices to show the result for the matrix $K_{k,p}$ (5.29). A simple computation shows that

$$K_{k,p} = T_{n_0} \otimes M_\omega + I_{n_0} \otimes T_\omega,$$

where the matrices T_n, M_ω and T_ω are defined via (6.11) and (6.9). Since the matrix T_{n_0} is symmetric and positive definite, we have

$$T_{n_0} = Q^T \Lambda Q \quad \text{with} \quad Q^T Q = I_{n_0},\ \Lambda = \operatorname{diag}[\lambda_i]_i, \quad \lambda_i > 0$$

Hence,

$$\begin{aligned} K_{k,p} &= (Q^T \otimes I_{n_0})(\Lambda \otimes M_\omega + I_{n_0} \otimes T_\omega)(Q \otimes I_{n_0}) \\ &= (Q^T \otimes I_{n_0}) \text{blockdiag} \left[\lambda_i M_\omega + T_\omega\right]_i (Q \otimes I_{n_0}). \end{aligned}$$

We apply now Lemma 6.4 and obtain

$$
\begin{aligned}
K_{k,p}^{-1} &= (Q^T \otimes I_{n_0})\text{blockdiag}\left[(\lambda_i M_\omega + T_\omega)^{-1}\right]_i (Q \otimes I_{n_0}) \\
&\sim (Q^T \otimes I_{n_0})\text{blockdiag}\left[\sum_{j=0}^{k-1}(\lambda_i M_j + \Delta_{j,1})^+\right]_i (Q \otimes I_{n_0}) \\
&= (Q^T \otimes I_{n_0})\sum_{j=0}^{k-1}(\Lambda \otimes M_j + I_{n_0} \otimes \Delta_{j,1})^+ (Q \otimes I_{n_0}) \\
&= \sum_{j=0}^{k-1}\left((Q^T \otimes I_{n_0})(\Lambda \otimes M_j + I_{n_0} \otimes \Delta_{j,1})(Q \otimes I_{n_0})\right)^+ \\
&= \sum_{j=0}^{k-1}(T_{n_0} \otimes M_j + I_{n_0} \otimes \Delta_{j,1})^+ = C_{mod}^{-1},
\end{aligned}
$$

which proves the result. □

6.3 A nonoverlapping preconditioner

In a first step, we define an nonoverlapping preconditioner C_{nov}. For $j \leq k-1$, let

$$\tilde{\mathbb{V}}_j = \{u \in \mathbb{V}_k,\ u(x) = 0 \quad \forall x \notin \tilde{\Omega}_{j,x}\} \quad \text{and} \quad \mathbb{W}_j = \tilde{\mathbb{V}}_j \mid_{\Omega_{j,x}} .$$

Moreover, we introduce a discrete energetic extension operator $\mathcal{E}_j : \mathbb{W}_j \mapsto \tilde{\mathbb{V}}_j$

$$\mathcal{E}_j u = v \quad \forall u \in \mathbb{W}_j, \quad j = 0, \ldots, k-2, \tag{6.13}$$

such that

$$a_p(v, w) = 0 \quad \forall w \in \tilde{\mathbb{V}}_{j+1}. \tag{6.14}$$

The matrix representation of the extension operator $\mathcal{E}_j$ with respect to the canonical basis Φ_k is denoted by the matrix $E_j \in \mathbb{R}^{N \times N}$. The space of the discrete harmonic functions is denoted by $\mathbb{H}_j$, i.e.

$$\mathbb{H}_j = \mathcal{E}_j \mathbb{W}_j, \quad j = 0, \ldots, k-2 \quad \text{and} \quad \mathbb{H}_{k-1} = \tilde{\mathbb{V}}_{k-1}. \tag{6.15}$$

We investigate now the space splitting

$$\mathbb{V}_k = \mathbb{H}_0 + \mathbb{H}_1 + \ldots + \mathbb{H}_{k-1}. \tag{6.16}$$

Lemma 6.5. *The splitting* (6.16) *is an orthogonal splitting with respect to the bilinearform $a_p(\cdot,\cdot)$, i.e.*

$$\mathbb{V}_k = \mathbb{H}_0 \oplus \mathbb{H}_1 \oplus \ldots \oplus \mathbb{H}_{k-1}.$$

Moreover, there exists exactly one $u_j \in \mathbb{H}_j$ such that

$$u = \sum_{j=0}^{k-1} u_j \quad \textit{and} \quad \sum_{j=0}^{k-1} a_p(u_j, u_j) = a_p(u, u) \quad \forall u \in \mathbb{V}_k.$$

Proof. The orthogonality is a consequence of the construction of the operator $\mathcal{E}_j$ and the spaces $\mathbb{H}_j$, see (6.14) and (6.15). This gives the first assertion. The second assertion follows from the first one. □

Lemma 6.6. *The spectral equivalence relation*

$$\varepsilon_j |u|^2_{1,\Omega_{j,x}} \leq \| u \|_p^2 \leq 2\varepsilon_j |u|^2_{1,\Omega_{j,x}} \quad \forall u \in \mathbb{H}_j.$$

is valid.

Proof. We start with the first assertion. By (5.28), we have

$$\| u \|_p^2 = \sum_{m=j}^{k} \varepsilon_m |u|^2_{1,\Omega_{m,x}} = \varepsilon_j |u|^2_{1,\Omega_{j,x}} + \sum_{m=j+1}^{k} \varepsilon_m |u|^2_{1,\Omega_{m,x}} \quad \forall u \in \mathbb{H}_j. \tag{6.17}$$

This gives the lower estimate. By construction of the space $\mathbb{H}_j$, we can conclude

$$\sum_{m=j+1}^{k} \varepsilon_m |u|^2_{1,\Omega_{m,x}} \leq \sum_{m=j+1}^{k} \varepsilon_m |v|^2_{1,\Omega_{m,x}} \forall v \in \mathbb{V}_k,\ v(2^{-j-1}, y) = u(2^{-j-1}, y), 0 \leq y \leq 1.$$

Setting v the symmetric reflection, i.e. $v(2^{j-1} - x, y) = u(2^{-j-1} + x, y)$, we obtain

$$\sum_{m=j+1}^{k} \varepsilon_m |u|^2_{1,\Omega_{m,x}} \leq \sum_{m=j+1}^{k} \varepsilon_j |v|^2_{1,\Omega_{m,x}} \leq \varepsilon_j |v|^2_{1,\tilde{\Omega}_{j+1,x}} = \varepsilon_j |u|^2_{\Omega_{j,x}} \tag{6.18}$$

using the monotonicity of κ. Combining (6.17) and (6.18) gives the upper estimate. The proof of the second assertion uses the same arguments. □

Now, we are able to introduce a nonoverlapping preconditioner $\mathcal{C}_{non}$. Using the matrices (3.1), we introduce the matrices

$$\mathcal{B}_j = \varepsilon_j^{-1} E_j \Delta^+_{j,N} E_j^T, \quad j = 0, \ldots, k-2, \quad \text{and} \quad \mathcal{B}_{k-1} = \varepsilon_{k-1}^{-1} \Delta^+_{k-1,D}. \tag{6.19}$$

Then, we define the preconditioner

$$\mathcal{C}_{non}^{-1} = \mathcal{B}_0 + \mathcal{B}_1 + \ldots + \mathcal{B}_{k-1}. \tag{6.20}$$

To solve $\mathcal{C}_{non} \underline{w} = \underline{r}$, we have to solve systems with the matrix $\Delta_{j,N}$ and to multiply with the extension operator $\mathcal{E}_j \leftrightarrow E_j$. We note that

$$a_p(\phi_{il}, \phi_{i'l'}) = \varepsilon_j \int_{\Omega_{j,x}} \nabla \phi_{il} \cdot \nabla \phi_{i'l'}, \quad 2^{k-j-1} \leq i, i' \leq 2^{k-j} - 1 \tag{6.21}$$

by definition of the bilinear form $a_p(\cdot, \cdot)$. Thus, we have to solve systems with the Laplacian.

Theorem 6.7. *Let $\mathcal{C}_{non}$ be defined via* (6.20)*. Moreover, let $K_{k,p}$ be defined via* (5.29)*. Then, $K_{k,p} \sim \mathcal{C}_{non}$.*

Proof. The proof is a collection of the previous results.
We apply Lemma 5.1 for the space splitting (6.16). We verify now the assumptions (5.1), (5.2), (5.3). By Lemma 6.5, we can conclude that

$$\sum_{j=0}^{k-1} a_p(u_j, u_j) = a_p \left(\sum_{j=0}^{k-1} u_j, \sum_{j=0}^{k-1} u_j \right).$$

Thus, $c_1 = c_2 = 1$ in (5.1) and (5.2). By Lemma 6.6, 5.2 and (6.21), relation (5.3) is valid with $c_4 = 2$ and $c_3 = 1$. This proves the assertion. □

Summarizing, we have constructed an optimal preconditioner for the stiffness matrix $K_{k,p}$. To prove the optimality of C_{non}, we don not use tensor product product arguments. We can change the matrix $\Delta^+_{j,N}$ in (6.19) by any preconditioner for this matrix, but, however, we have to do multiplications with the discrete energetic extension operator $\mathcal{E}_j$. In the next subsection, we will investigate a preconditioner without discrete energetic extensions. This leads us to the overlapping preconditioner (3.2).

6.4 The overlapping preconditioner C for $\alpha < \frac{1}{2}$

Now, we will prove Theorem 3.1 for the preconditioner (3.2). The starting point is the preconditioner $\mathcal{C}_{non}$ (6.20) which will be simplified.
In a first step, we prove the stability of the energetic extension in H^1 for $\alpha < \frac{1}{2}$ on tensor product meshes. Two auxiliary results in one dimension are required for the proof of this result. The first one is a result about the local distribution of the energy of an extended function with minimal energy. This result might be of a particular interest. The second result is about the stability of the energetic extension in 1D for a weighted bilinear form with mass term. Let $a_{1,\lambda}(\cdot,\cdot)$ be the bilinear form (6.4) on $\mathbb{X}_n$ In addition, let

$$a_\lambda(u,v) = a_\lambda(\Phi_1\underline{u},\Phi_1\underline{v}) := \int_0^1 u'(x)v'(x)\,\mathrm{d}x + \lambda\, n \sum_{s=1}^{n-1} u\left(\frac{s}{n}\right) v\left(\frac{s}{n}\right) \quad \text{and} \quad \|\cdot\|_\lambda^2 = a_\lambda(\cdot,\cdot). \tag{6.22}$$

Lemma 6.8. *In addition to the above assumptions, let us assume that v_* is the solution of*

$$\min_{v\in\mathbb{X}_n \quad v(1)=g \quad v(0)=0} a_{1,\lambda}(v,v).$$

Moreover, let s and s' be two integer which satisfy $2^{j-1} < s \le 2^j$ and $2^{j'-1} < s' \le 2^{j'}$. Then,

$$a_\lambda\,|_{\tau_s^n}\,(v_*,v_*) \le 2^{\alpha(j-j')} a_\lambda\,|_{\tau_s'^n}\,(v_*,v_*) \quad \textit{for } j < j'.$$

Proof. The function $v_* = [\Phi_1]\underline{v}$ is expressed by the solution of the following system of equations

$$\begin{aligned} (2+\lambda)v_s - v_{s-1} - v_{s+1} &= 0 \quad \text{if } s \ne 2^j, \\ (q+1)(1+\frac{\lambda}{2})v_s - v_{s-1} - qv_{s+1} &= 0 \quad \text{if } s = 2^j. \end{aligned} \tag{6.23}$$

with $v_0 = 1$, $v_n = g$ and $q = \frac{\varepsilon_{k-j-1}}{\varepsilon_{k-j}} = 2^\alpha$.
Without loss of generality, let $g \ge 0$. A direct consequence of the minimal energy extension is the inequality chain

$$0 = v_0 \le v_1 \le v_2 \le \ldots \le v_{n-1} \le v_n = g. \tag{6.24}$$

Using (6.23), we can conclude that

$$v_{s+1} - v_s = v_s - v_{s-1} + \lambda v_s \ge v_s - v_{s-1} > 0 \quad \text{for } s \ne 2^j$$

and

$$v_{s+1} - v_s \ge q^{-1} v_s - v_{s-1} > 0 \quad \text{for } s = 2^j.$$

Hence, we can estimate the H^1-part of the norm $\|\cdot\|_\lambda$ (6.22) and obtain

$$\int_{\tau_s} (v_*')^2 \le q^{2j-2j'} \int_{\tau_{s'}} (v_*')^2 = 2^{2\alpha(j-j')} \int_{\tau_s'} (v_*')^2 \quad s \le s',$$

which is equivalent to

$$\int_{\tau_s} \omega^2(\xi)(v_*'(\xi))^2\,\mathrm{d}\xi \le 2^{\alpha(j-j')} \int_{\tau_{s'}} \omega^2(\xi)(v_*'(\xi))^2\,\mathrm{d}\xi$$

The result for the L_2-part of the norm $\|\cdot\|_\lambda$ (6.22) follows directly from (6.24). This proves the assertion.
□

Remark 6.9. *The proof shows that the estimates are sharp for $\lambda = 0$.*

The map $\mathcal{E}_1 : \mathbb{R} \mapsto \mathbb{X}_n$ given by $v_* = \mathcal{E}_1 g$ defines an energetic extension operator with respect to the energetic norm $\| \cdot \|_{1,\lambda}$ (6.4). Now, we investigate the stability of this extension in the norm $\| \cdot \|_\lambda$ (6.22).

Lemma 6.10. *Let $\omega^2(\xi) = \xi^\alpha$, $0 \le \alpha < \frac{1}{2}$. Moreover, let $\mathcal{E}_1$ be the discrete energetic extension operator defined above. Then, the operator $\mathcal{E}_1$ is stable in $\| \cdot \|_\lambda$ (6.22), i.e.*

$$\varepsilon_0 \parallel \mathcal{E}_1 u \parallel_\lambda^2 \le \frac{1}{1-2\alpha} \varepsilon_0 \min_{\substack{v \in \mathbb{X}_n, \\ v(1) = u, \\ v(0) = 0}} \parallel v \parallel_\lambda^2$$

Proof. Recall that $\Omega_j = (2^{-j-1}, 2^{-j})$ and $\Omega_{k-1} = (0, 2^{-k+1})$.
By summation of the result of Lemma 6.8 over all elements τ_s inside $\tilde{\Omega}_l$ and $\tilde{\Omega_0}$, we have

$$\varepsilon_l \parallel \mathcal{E}_1 u \parallel_{\lambda,\Omega_l}^2 \le 2^{(\alpha-1)l} \varepsilon_0 \parallel \mathcal{E}_1 u \parallel_{\lambda,\Omega_0}$$

or, equivalently,

$$\parallel \mathcal{E}_1 u \parallel_{\lambda,\Omega_l}^2 \le 2^{(2\alpha-1)l} \parallel \mathcal{E}_1 u \parallel_{\lambda,\Omega_0} .$$

In the case $\alpha < \frac{1}{2}$, we obtain

$$\varepsilon_0 \parallel \mathcal{E}_1 u \parallel_{\lambda,(0,1)}^2 = \varepsilon_0 \sum_{l=0}^{k-1} \parallel \mathcal{E}_1 u \parallel_{\lambda,\Omega_l}^2 \le \frac{1}{1-2\alpha} \varepsilon_0 \parallel \mathcal{E}_1 u \parallel_{\lambda,\Omega_0}^2 = \frac{1}{1-2\alpha} a_\lambda \mid_{\Omega_0} (\mathcal{E}_1 u, \mathcal{E}_1 u)$$

by the geometric series. By the definition of the operator $\mathcal{E}_1$,

$$a_{p,\lambda} \mid_{\Omega_0} (\mathcal{E}_1 u, \mathcal{E}_1 u) = \min_{\substack{v \in \mathbb{X}_n, \\ v(0) = 0, \\ v(1) = u}} a_{1,\lambda} \mid_{\Omega_0} (v, v).$$

By the monotonicity of the values ε_j, we can conclude that

$$a_{p,\lambda} \mid_{\Omega_0} (u, u) \le \varepsilon_0 \parallel u \parallel_\lambda^2 .$$

This gives

$$\varepsilon_{j+1} \parallel \mathcal{E}_1 u \parallel_\lambda^2 \le \frac{1}{1-2\alpha} \varepsilon_{j+1} \min_{\substack{v \in \mathbb{X}_n, \\ v(1) = u, \\ v(0) = 0}} \parallel v \parallel_\lambda^2,$$

which proves the assertion. □

Remark 6.11. *In the case $\omega^2(\xi) = \xi^\alpha$ with $\alpha = \frac{1}{2}$, one obtains a constant c which proportionally to the level number k. For $\alpha > \frac{1}{2}$, we have to use another estimate.*

Remark 6.12. *By a scaling argument, the result can be extended to*

$$\varepsilon_{j+1} \parallel \mathcal{E}_{1,j} u \parallel_{\lambda,\Omega_{j+1}}^2 \le \frac{1}{1-2\alpha} \varepsilon_{j+1} \min_{\substack{v \in \mathbb{X}_n, \\ v(2^{-j-1}) = u, \\ v(0) = 0}} \parallel v \parallel_{\lambda,\Omega_{j+1}}^2 .$$

In a second step, we consider now the corresponding two dimensional result.

Lemma 6.13. *Let* $\omega^2(\xi) = \xi^\alpha$ *with* $0 \le \alpha < \frac{1}{2}$. *Moreover, let* $\mathcal{E}_j$ *be the discrete energetic extension operator defined via* (6.13). *Then, the extension is stable in* H^1*, i.e.*

$$\varepsilon_{j+1}|\mathcal{E}_j u|^2_{H^1(\Omega_{j+1,x})} \le \frac{1}{1-2\alpha} \min_{\substack{v \in \mathbb{V}_k \\ v\,|_{\Gamma_{j+1,x}}}} |v|^2_{H^1(\Omega_{j+1,x})}$$

Proof. Since the weight function depends only on the x-direction, the y-direction does not depend on the weight. Moreover, we use a tensor-product discretization and transform the problem into the basis eigenfunctions v_r with respect to the y-direction. Hence, the extension (6.13) can be encoupled into the one-dimensional problems with the bilinear form $a_{1,\lambda_r}(\cdot,\cdot)$ (6.4), where $\lambda_r > 0$ denotes the corresponding eigenvalues. Now, the assertion is a direct consequence of the one dimensional result in Lemma 6.10. $\square$

A consequence of this result, Lemma 5.1 and Lemma 5.2 is the following

Corollary 6.14. *Let* $\Delta_{j,D}$ *and* $\Delta_{j,N}$ *be defined via* (3.1). *Let* E_j *be the matrix representation of the extension operator* $\mathcal{E}_j$ (6.13). *Moreover, let* $\omega^2(\xi) = \xi^\alpha$ *with* $0 \le \alpha < \frac{1}{2}$. *Then,*

$$\left(\varepsilon_j^{-1}\Delta^+_{j,D}\underline{v},\underline{v}\right) \le \left((\varepsilon_j^{-1}E_j\Delta^+_{j,N}E_j^T + \varepsilon_j^{-1}\Delta^+_{j+1,D})\underline{v},\underline{v}\right) \le c\left(\varepsilon_j^{-1}\Delta^+_{j,D}\underline{v},\underline{v}\right) \quad \forall \underline{v}.$$

The constant is independent of j *and the discretization parameter* h.

Proof. We consider the Additive Schwarz splitting for $\Delta_{j,D}$ into $\Delta^+_{j+1,D}$ and $E_j\Delta_{j,N}E_j^T$. Then, the proof is consequence of Lemma 6.13. $\square$

Now, we introduce an overlapping preconditioner

$$C_{ov}^{-1} = \sum_{j=0}^{k-2}\left(\varepsilon_j^{-1}E_j\Delta^+_{j,N}E_j^T + \varepsilon_j^{-1}\Delta^+_{j+1,D}\right) + \varepsilon_{k-1}^{-1}\Delta^+_{k-1,D}. \tag{6.25}$$

Using (6.20) and the positive semidefinitness of $\Delta^+_{j+1,D}$, we can estimate

$$\begin{aligned} C_{non}^{-1} &= \sum_{j=0}^{k-2}\varepsilon_j^{-1}E_j\Delta^+_{j,N}E_j^T + \varepsilon_{k-1}^{-1}\Delta^+_{k-1,D} \\ &\le \sum_{j=0}^{k-2}\varepsilon_j^{-1}E_j\Delta^+_{j,N}E_j^T + \varepsilon_{k-1}^{-1}\Delta^+_{k-1,D} + \sum_{j=0}^{k-2}\varepsilon_j^{-1}\Delta^+_{j+1,D} = C_{ov}^{-1} \end{aligned} \tag{6.26}$$

Moreover,

$$\begin{aligned} (C_{ov}^{-1}\underline{v},\underline{v}) &= \sum_{j=0}^{k-1}\varepsilon_j^{-1}\left(\Delta^{-1}_{j,N}E_j^T\underline{v},E_j^T\underline{v}\right) + \sum_{j=0}^{k-2}\varepsilon_j^{-1}\left(\Delta^+_{j+1,D}\underline{v},\underline{v}\right) + \varepsilon_{k-1}^{-1}\left(\Delta^+_{k-1,D}\underline{v},\underline{v}\right) \\ &= \sum_{j=0}^{k-2}u_j + \sum_{j=0}^{k-2}v_j + v_{k-1} \end{aligned}$$

with

$$u_j = \varepsilon_j^{-1}\left(\Delta^{-1}_{j,N}E_j^T\underline{v},E_j^T\underline{v}\right) \quad \text{and} \quad v_j = \varepsilon_j^{-1}\left(\Delta^+_{j+1,D}\underline{v},\underline{v}\right).$$

Applying Corollary 6.14 to the weight function $\omega(\xi) = \xi^\alpha$ with $\alpha \in (0, 0.5)$, we have the estimate

$$v_j \leq q(v_{j+1} + u_{j+1}) \quad \text{with } q = 2^{-\alpha}.$$

Hence,

$$\begin{aligned} \left(C_{ov}^{-1}\underline{v}, \underline{v}\right) &= \sum_{j=0}^{k-2}(u_j + v_j) + v_{k-1} \\ &\leq \sum_{j=0}^{k-2}\sum_{l=0}^{j} q^l u_j + \sum_{j=0}^{k-1} q^j v_{k-1} \\ &\leq \frac{1}{1-q}\sum_{j=0}^{k-2} u_j + v_{k-1} = \frac{1}{1-2^\alpha}\left(C_{non}^{-1}\underline{v}, \underline{v}\right) \quad \forall \underline{v}. \end{aligned} \tag{6.27}$$

Using (6.26) and (6.27), we have shown the following result.

Lemma 6.15. *Let* $\omega(\xi) = \xi^\alpha$ *with* $\alpha > 0$. *Let* C_{non} *and* C_{ov} *be defined via* (6.20) *and* (6.25). *Then,*

$$\left(C_{non}^{-1}\underline{v}, \underline{v}\right) \leq \left(C_{ov}^{-1}\underline{v}, \underline{v}\right) \leq \frac{1}{1-2^{-\alpha}}\left(C_{non}^{-1}\underline{v}, \underline{v}\right) \quad \forall \underline{v}.$$

Now, we are able to prove Theorem 3.1.

Proof. Due to Lemma 6.15, Lemma 5.8 and Theorem 6.7, it suffices to show

$$C_{ov} \sim C.$$

The relation $C^{-1} \leq C_{ov}^{-1}$ is trivial. This proves $K_k \leq C$. By Corollary 6.14, we can conclude that

$$C_{ov}^{-1} = \sum_{j=0}^{k-2}\left(\varepsilon_j^{-1}E_j\Delta_{j,N}^{+}E_j^T + \varepsilon_j^{-1}\Delta_{j+1,D}^{+}\right) + \varepsilon_{k-1}^{-1}\Delta_{k-1,D}^{+} \preceq \sum_{j=0}^{k-1}\varepsilon_j^{-1}\Delta_{j,D}^{+} = C,$$

which proves the second assertion of Theorem 3.1. □

7 Numerical Examples

In this section, we present some numerical examples.
In the first two examples, we consider the bilinearform $a_p(\cdot,\cdot)$ for $d = 1$. Figure 5 displays the maximal and minimal eigenvalue for the matrix $C_{mod}^{-1}K_{k,p}$ with the modified preconditioner C_{mod} (3.4) for different weight functions.
The minimal eigenvalue of the matrix $C_{mod}^{-1}K_{k,p}$ is bounded from below by a positive constant in the cases $\omega^2(\xi) = \xi^\alpha$ iff $\alpha \neq 1$. A logarithmic growth can be seen for $\alpha = 1$. The maximal eigenvalue is bounded from above by a constant of 2 for all investigated weight functions.
Moreover, we investigated the preconditioner C (3.2). Figure 6 displays the maximal and minimal eigenvalue for the matrix $C^{-1}K_{k,p}$ for different weight functions. The minimal eigenvalue of the matrix $C^{-1}K_{k,p}$ is bounded from below by a positive constant in the cases $\omega^2(\xi) = \xi^\alpha$ iff $\alpha \neq 1$. A logarithmic growth can be seen for $\alpha = 0$. So, the asymptotic behavior is similar for both preconditioners if $\omega^2(\xi) = \xi^\alpha$ with $\alpha > 0$. However, the minimal eigenvalue of $C^{-1}K_{k,p}$ is greater for the preconditioner C than for the preconditioner C_{mod}. So, the preconditioner C (3.2) should be preferred. The maximal eigenvalue in bounded from above for $\alpha > 0$ and growths logarithmically for $\alpha = 0$.

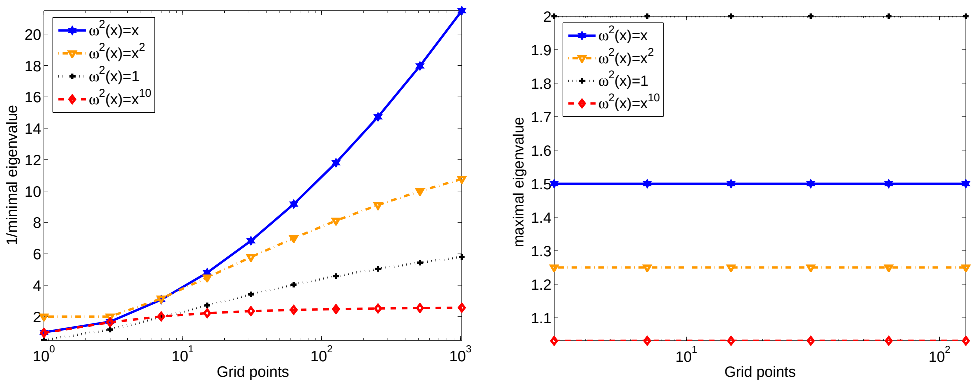

Figure 5: Eigenvalue bounds for $C_{mod}^{-1}K_{k,p}$ with the modified preconditioner (3.4): minimal eigenvalue left, maximal eigenvalue right for $d = 1$.

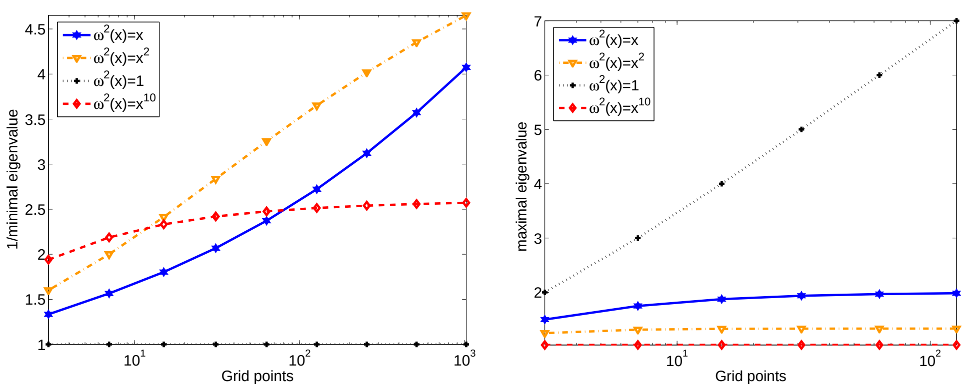

Figure 6: Eigenvalue bounds for $C^{-1}K_{k,p}$ with the preconditioner (3.2), minimal eigenvalue left, maximal eigenvalue right for $d = 1$.

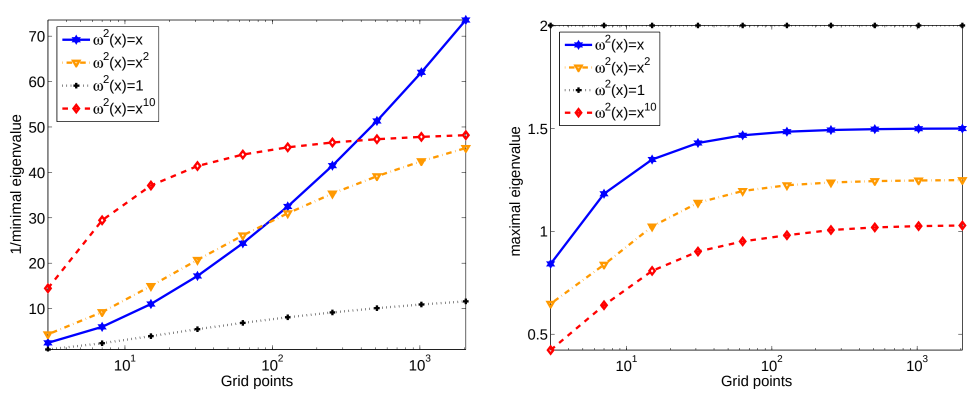

Figure 7: Eigenvalue bounds for $C^{-1}K_k$ with the modified preconditioner (3.4), minimal eigenvalue left, maximal eigenvalue right for $d = 1$.

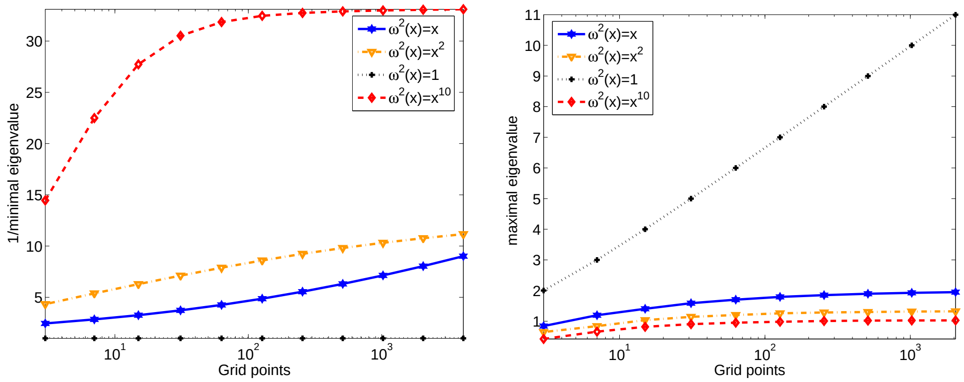

Figure 8: Eigenvalue bounds for $C^{-1}K_k$ with the modified preconditioner (3.2), minimal eigenvalue left, maximal eigenvalue right for $d = 1$.

In the next examples, we investigate the preconditioner for the original matrix K_k. Figure 7 displays eigenvalue bounds of $C_{mod}^{-1}K_k$ and Figure 8 displays eigenvalue bounds of $C^{-1}K_k$. The results are worse than for the matrix $K_{k,p}$. The maximal eigenvalues are about the same as for $C^{-1}K_{k,p}$, whereas the minimal eigenvalues get the additional factor 2^α of Lemma 5.8. In particular for the weight function $\omega^2(\xi) = \xi^{10}$, the results are not satisfying. Hence, the approximation of this weight function with a piecewise constant coefficient function, see the definition of the bilinear forms a_p (5.27) and (5.26), only in the interval $(2^{-j}, 2^{-j+1})$ might be the reason for these results. With a more accurate approximation of the weight function, the results could be improved.
In a next example, we investigate the quality of the preconditioner C for two dimensional problems. Figure 9 displays the maximal and minimal eigenvalue for the matrix $C^{-1}K_{k,p}$ with the modified preconditioner (3.2) for different weight functions. The results are slightly better than in the one-dimensional case. The

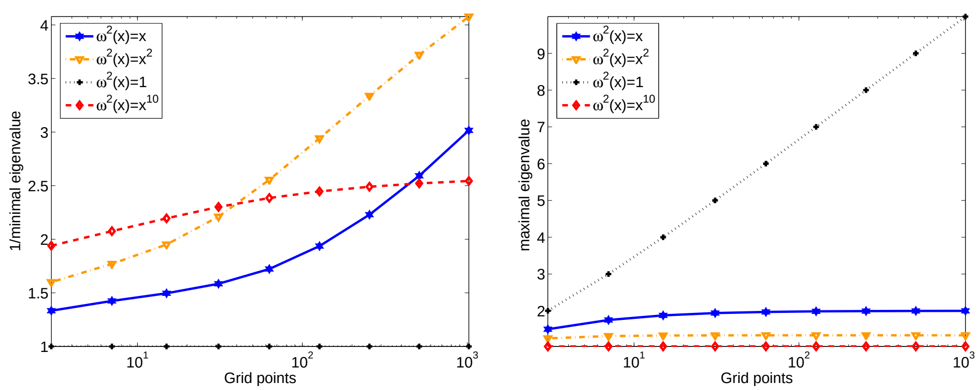

Figure 9: Eigenvalue bounds with the preconditioner (3.2), minimal eigenvalue left, maximal eigenvalue right for $d = 2$.

general behavior is similar to the 1D-case, cf. Figure 6.

8 Concluding remarks and possible generalizations

We will conclude the paper with the following remarks.

We analyzed overlapping DD-preconditioners for finite element discretizations for degenerated problems of the type $-\nabla \cdot (\omega^2(x)\nabla u) = f$ with $\omega^2(x) = x^\alpha$. The optimality of the preconditioner has been shown for $\alpha \neq 1$. The analysis is based on algebraic arguments, relation (6.2) for $\alpha > 1$ and tensor product arguments for $\alpha < 1$.
The proposed methods can be directly applied to the three-dimensional case, i.e. to the finite element discretization of degenerated elliptic boundary value problem in $\Omega = (0,1)^3$. The corresponding weak formulation is Find $u \in \mathbb{H}_{\omega,0} := \{u \in L_2(\Omega) : \int_\Omega (\nabla u)^T \omega^2(x) \nabla u \; \mathrm{d}(x,y,z) < \infty, u\mid_{\partial\Omega} = 0\}$ such that

$$a(u,v) := \int_\Omega (\nabla v)^T \cdot \omega^2(x) \nabla u \; \mathrm{d}(x,y,z) = (f,v) \quad \forall v \in \mathbb{H}_{\omega,0}. \tag{8.1}$$

In the case of a tensor product discretization, the presented proofs can directly be applied. For $\alpha > 1$, only relation (6.2) has to be verified. The proof of this relation can also be done in the three-dimensional case.
The presented proofs have been done for tensor-product discretizations on the unit square. On a general domain $\Omega \subset \mathbb{R}^2$, the corresponding problem reads

$$-\nabla \cdot (\omega^2(d(x,y))\nabla u) = f, \tag{8.2}$$

where $d(x,y)$ denotes the distance to the boundary of Ω, or the distance to one part of Ω. Since the weight function is continuous, we can apply the ficticous space lemma, [19]. We transfer the discretized problem to a tensor product discretization on the unit square Ω_f as decribed in Section 2. Here, not more than one node of the finite-element mesh of Ω is contained in one triangle of the finite element mesh of our ficticous domain Ω_f. For the problem on Ω_f, we apply the results of our paper. This gives us a fast solver for the discretized problem of the pde (8.2) in Ω.
Acknowlegdement: The paper was written during the Special Semester on Computational Mechanics in Linz 2005. The second author thanks the RICAM for the hospitality during his stay in Linz.

References

[1] S. Beuchler. Multi-grid solver for the inner problem in domain decomposition methods for p-FEM. *SIAM J. Numer. Anal.*, 40(3):928–944, 2002.

[2] S. Beuchler. AMLI preconditioner for the p-version of the FEM. *Num. Lin. Alg. Appl.*, 10(8):721–732, 2003.

[3] S. Beuchler and S. Nepomnyaschikh. Overlapping additiv schwarz preconditioners for degenerated elliptic problems: Part ii locally anisotropic problems. Technical report, RICAM, 2006.

[4] S. Beuchler, R. Schneider, and C. Schwab. Multiresolution weighted norm equivalences and applications. *Numer. Math.*, 98(1):67–97, 2004.

[5] Sven Beuchler. Multilevel solvers for a finite element discretization of a degenerate problem. *SIAM J. Numer. Anal.*, 42(3):1342–1356 (electronic), 2004.

[6] S. Börm and R. Hiptmair. Analysis of tensor product multigrid. *Numer. Algorithms*, 26(3):219–234, 2001.

[7] J. Bramble, J. Pasciak, and A. Schatz. The construction of preconditioners for elliptic problems by substructuring I. *Math. Comp.*, 47(175):103–134, 1986.

[8] J. Bramble, J. Pasciak, and A. Schatz. The construction of preconditioners for elliptic problems by substructuring II. *Math. Comp.*, 49(179):1–16, 1987.

[9] J. Bramble, J. Pasciak, and A. Schatz:. The construction of preconditioners for elliptic problems by substructuring III. *Math. Comp.*, 51(184):415–430, 1988.

[10] J. Bramble, J. Pasciak, and A. Schatz. The construction of preconditioners for elliptic problems by substructuring IV. *Math. Comp.*, 53(187):1–24, 1989.

[11] J. Bramble, J. Pasciak, and J. Xu. Parallel multilevel preconditioners. *Math. Comp.*, 55(191):1–22, 1991.

[12] J. Bramble and X. Zhang. Uniform convergence of the multigrid V-cycle for an anisotropic problem. *Math. Comp.*, 70(234):453–470, 2001.

[13] I. G. Graham, P. Lechner, and R. Scheichl. Domain decomposition for multiscale pdes. Technical report, University of Bath, 2006.

[14] W. Hackbusch. *Multigrid Methods and Applications*. Springer-Verlag. Heidelberg, 1985.

[15] M. Jung, U. Langer, A. Meyer, W. Queck, and M. Schneider. Multigrid preconditioners and their applications. Technical Report 03/89, Akad. Wiss. DDR, Karl-Weierstraß-Inst., 1989.

[16] V. G. Korneev. Почти оптимальныи метод решения задач Дирихле на подобластях декомпосиции иерархическои hp-версии. Дифферентиальные Уравнения, 37(7):1008–1018, 2001. An almost optimal method for Dirichlet problems on decomposition subdomains of the hierarchical hp-version.

[17] A. Kufner and A.M. Sändig. *Some applications of weighted Sobolev spaces*. B.G.Teubner Verlagsgesellschaft. Leipzig, 1987.

[18] A. M. Matsokin and S. V. Nepomnyaschikh. The Schwarz alternation method in a subspace. *Iz. VUZ Mat.*, 29(10):61–66, 1985.

[19] S. V. Nepomnyaschikh. Fictitious space method on unstructured meshes. *East-West J. Numer. Math.*, 3(1):71–79, 1995.

[20] S. V. Nepomnyaschikh. Preconditioning operators for elliptic problems with bad parameters. In *Eleventh International Conference on Domain Decomposition Methods (London, 1998)*, pages 82–88 (electronic). DDM.org, Augsburg, 1999.

[21] O. Pironneau and F. Hecht. Mesh adaption for the Black and Scholes equations. *East-West Journal of Numerical Mathematics*, 8(1):25–36, 2000.

[22] R. Scheichl and E. Vainikko. Additive schwarz and aggregation-based coarsening for elliptic problems with highly variable coefficients. Technical report, University of Bath, 2006.

[23] A. Toselli and O. Widlund. *Domain Decomposition Methods- Algorithms and Theory*. Springer, 2005.

[24] H. Yserentant. On the multi-level-splitting of the finite element spaces. *Numer. Math.*, 49:379–412, 1986.

[25] X. Zhang. Multilevel Schwarz methods. *Numer. Math.*, 63:521–539, 1992.