96-29

ON SOME APPLICATIONS OF THE H^σ–STABLE WAVELET–LIKE HIERARCHICAL FINITE ELEMENT SPACE DECOMPOSITIONS

PANAYOT S. VASSILEVSKI

ABSTRACT. In this paper we first review the construction of stable Riesz bases for finite element spaces with respect to Sobolev norms. Then, we construct optimal order multilevel preconditioners for the matrices in the normal form of the equations arising in the finite element discretization of non–symmetric second order elliptic equations. The optimality of the AMLI methods is proven under H^2–regularity assumption on underlined elliptic problem. A second application is in the area of domain embedding methods that utilizes H^1–bounded extension operators of data, finite element functions defined on the boundary across which the given (irregular) domain is embedded into a more regular (e.g., parallelepiped) one.

1. INTRODUCTION

In this paper we are concerned with the construction of efficient numerical methods for matrix problems arising from finite element methods for elliptic partial differential equations. In practical computation, the standard nodal basis for the finite element space is often chosen as the computational basis and the resulting matrices are ill-conditioned. In Vassilevski and Wang [VW97] the objective was to seek a substitution for the standard nodal basis so that the stiffness matrix arising from the new basis is well–conditioned, preserving the two major properties required for a computationally feasible basis: (a) the basis functions must be computable and (b) they must also have local support, hence the resulting stiffness matrix is sparse. This paper will first review the construction of local projections operators by the wavelet-like method proposed in [VW96a]. The latter wavelet–like projections operators have a main application in the construction of a stable Riesz basis with the above mentioned features for the finite element application to elliptic problems.

Attempts in the search of a stable Riesz basis with some restrictions, either on the mesh or on the analysis, have been made in Griebel and Oswald [GO94], Kotyczka and Oswald [KO95], and Stevenson [Ste95a], [Ste95b]. For a comparative study on the construction of economical Riesz bases for Sobolev spaces we refer to Lorentz and Oswald [LO96]. The method from Vassilevski and Wang [VW96a], [VW97] is general and provides a satisfactory answer for most of the elliptic equations. It is based on modifying the existing (unstable) hierarchical basis by using operators which are approximations of the L^2–projections onto coarse finite element spaces. For more

Date: September 9, 1996.

1991 *Mathematics Subject Classification.* 65F10, 65N20, 65N30.

Key words and phrases. wavelet–like Riesz basis, least squares, multilevel preconditioning, domain embedding, preconditioning, domain decomposition, finite elements, second order elliptic problems.

This work of the author was partially supported by the Bulgarian Ministry for Education, Science and Technology under grant I–504, 1995 and by the U.S. National Science Foundation grant INT-95-06184.

details we refer to [VW97]. The construction and the main properties of the stable Riesz basis are reviewed in Section 2.

In the present paper we stress upon two applications of the stable wavelet–like local projection operators (denoted further by π_k) which are the main ingredient in the construction of stable Riesz bases for finite element spaces.

The first application, presented in Section 3, is in the construction of AMLI (Algebraic Multi–Level Iteration) preconditioners for matrices in normal form (i.e., A^TA) in the case of convection–diffusion (non–symmetric) finite element elliptic equations. We prove, under H^2–regularity assumption (commonly imposed when studying convergence of multilevel methods in L^2–norm, cf. e.g. Bank and Dupont [BD81] and Goldstein, Manteuffel and Parter [GMP93]), that in three space dimensions the AMLI method of hybrid type (see further Definition 3.2) is both of optimal order and spectrally equivalent to $B = A^TA$. This method may be an alternative to the classical W–cycle multigrid with sufficiently many smoothing iterations.

The second main application is based on the H^σ–boundedness of the local projection operators π_k. In Section 4, we use this fact for the construction of approximate harmonic extension operators. This techniques has already been used in Bramble and Vassilevski [BV96] for constructing preconditioners in the interface domain decomposition (or DD) technique that allows for inexact subdomain solvers. Here, we present another application in the dual to the DD method; namely, the domain embedding technique. For the latter we refer to Nepomnyaschikh [Nep91b], [Nep91a], see also Proskurowski and Vassilevski [PV94] and the references given therein. The application of the H^σ–boundedness of the local projection operators in the domain embedding context was possible due to an algebraic fact, that a strengthened Cauchy–Schwarz inequality for the matrix implies the same strengthened Cauchy–Schwarz inequality for the inverse matrix. The detailed presentation is given in the last Section 4.

For practical aspects of the wavelet–like modification of the classical HB (hierarchical basis) or equivalently, of the bounded local projection operators π_k, we refer to Vassilevski and Wang [VW96b], [VW97] and Bramble and Vassilevski [BV96].

2. A Stable Riesz Basis by Wavelet Method

In this section we review the construction of local projections π_k which are H^σ–stable, $\sigma \in (0,1]$, and provide computationally feasible Riesz basis for the finite element space $V = V_J$. The bilinear form of main interest is the one from the second-order elliptic problems. The method to be presented here was proposed by Vassilevski and Wang in [VW96a]. It relies on the fundamental estimate due to Oswald [Osw94] which characterizes the Sobolev space norms $\|.\|_\sigma$, $\sigma \in [0,1]$ for finite element spaces:

$$\sum_{j=1}^{k} h_j^{-2\sigma} \|(Q_j - Q_{j-1})v\|_0^2 \le \sigma_N \|v\|_\sigma^2 \quad \text{for all } v \in V_k \cdot \tag{2.1}$$

Here, $V_0 \subset V_1 \subset \cdots \subset V_J$ is a sequence of nested conforming finite element spaces contained in $H_0^1 = H_0^1(\Omega)$ obtained by $J \ge 1$ successive steps of uniform refinement of an initial coarse triangulation $\mathcal{T}_0$ of the polygonal domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$. The kth level meshsize is denoted by h_k and we assume that $h_k = \frac{1}{2}h_{k-1} = 2^{-k}h_0$. Also, the kth level triangulation (the set of elements at level k) will be denoted by $\mathcal{T}_k$. Finally,

$Q_k : L^2(\Omega) \to V_k$ is the standard L^2-projection operator. Here and in what follows by $\| \cdot \|_\sigma$, $\sigma \in [0,1]$ we will denote the norm in the Sobolev space $H_0^\sigma(\Omega)$. (The space H_0^σ is obtained by interpolation between the spaces H_0^1 and L^2, see, e.g., Bramble [Bra95].)

2.1. **On the basis construction.** Define the L^2-projection operators $Q_k : L^2(\Omega) \to V_k$ by the identity,

$$(Q_k v, \psi) = (v, \psi) \quad \text{for all } \psi \in V_k \cdot$$

Here and in what follows by $(.,.)$ we mean the standard L^2-inner product. The L^2-norm will be denoted by $\|.\|_0$.
Next, assume that there are computationally feasible approximations $Q_k^a : L^2(\Omega) \to V_k$ to Q_k such that for some small tolerance $\tau > 0$ the following estimate holds:

$$\|(Q_k - Q_k^a)v\|_0 \le \tau \|Q_k v\|_0 \quad \text{for all } v \in L^2(\Omega). \tag{2.2}$$

We also need the nodal interpolation operators $I_k : \mathcal{C}(\Omega) \to V_k$ defined for any continuous function ψ by $I_k\psi = \sum_{i=1}^{n_k} \psi(x_i)\phi_i^{(k)}$. Here, $\{\phi_i^{(k)},\ i = 1, \ldots, n_k\}$ is the nodal basis of V_k. That is, $\phi_i^{(k)}(x_j) = \delta_{i,j}$–the Kronecker symbol, when x_j runs over all the nodal degrees of freedom $\mathcal{N}_k$ of (the node set at kth discretization level) of V_k. Note that $\{\phi_i^{(k-1)},\ i = 1, \ldots, n_{k-1}\} \cup \{\phi_i^{(k)},\ i = n_{k-1}+1, \ldots, n_k\}$ also forms a basis, called the two–level hierarchical basis of V_k.

Definition 2.1 (Wavelet–like local projection operators). *The projection operators of major interest are defined as follows:*

$$\pi_k = \prod_{j=k}^{J-1} (I_j + Q_j^a(I_{j+1} - I_j)), \tag{2.3}$$

with $\pi_J = I$.

It is clear that $\pi_k\psi = \psi$ if $\psi \in V_k$ since $I_j\psi + Q_j^a(I_{j+1} - I_j)\psi = I_j\psi = \psi$ for $j \ge k$ based on $(I_{j+1} - I_j)\psi = 0$ and $I_j\psi = \psi$. This also implies that $\pi_k^2 = \pi_k$.

Note also that $\pi_{k-1}(I_k - I_{k-1})\phi = Q_{k-1}^a(I_k - I_{k-1})\phi$ and $\pi_k - \pi_{k-1} = (I - Q_{k-1}^a)(I_k - I_{k-1})\pi_k$. Then, the components in the definition for the wavelet–like multilevel hierarchical basis read as follows:

$$\{\phi_i^{(0)}, i = 1, \ldots, n_0\} \bigcup_{j=1}^{k} \{(I - Q_{j-1}^a)\phi_i^{(j)},\ i = n_{j-1}+1, \ldots, n_j\} \cdot \tag{2.4}$$

The above components $\{(I - Q_{j-1}^a)\phi_i^{(j)},\ i = n_{j-1}+1, \ldots, n_j\}$ can be seen as a modification of the classical hierarchical basis components based on the interpolation operator I_k since $(I - Q_{j-1}^a)\phi_i^{(j)} = (I - Q_{j-1}^a)(I_j - I_{j-1})\phi_i^{(j)}$; the modification of the classical hierarchical basis components $\{(I_j - I_{j-1})\phi_i^{(j)},\ i = n_{j-1}+1, \ldots, n_j\}$ comes from the additional term $Q_{j-1}^a(I_j - I_{j-1})\phi_i^{(j)}$. In other words, the modification was made by subtracting from each nodal hierarchical basis function $\phi_i^{(j)}$ its approximate L^2-projection $Q_{k-1}^a\phi_i^{(j)}$ onto the coarse level $k-1$. The modified hierarchical basis functions are close relatives of the known Battle–Lemarié wavelets [Dau92].

Observe that in the limit case when $Q_k^a = Q_k$ (i.e., $\tau = 0$ in (2.2)), we get

$$\pi_k v = Q_k I_{k+1} Q_{k+1} I_{k+2} \ldots Q_{J-1} I_J v = Q_k Q_{k+1} \ldots Q_{J-1} v = Q_k v.$$

Therefore, $\pi_k v = Q_k v$. I.e., π_k reduces to the exact L^2–projection Q_k. As is well–known, the L^2–projection operators are bounded both in $H_0^1(\Omega)$ and $L^2(\Omega)$. Note finally that in this case we also have $\pi_k - \pi_{k-1} = Q_k - Q_{k-1}$. The decomposition of V based on the operators $Q_k - Q_{k-1}$ was a starting point in the construction of the BPX–method in Bramble, Pasciak and Xu [BPX90] (see also the survey Xu [Xu92]), and its development later in Zhang [Zha92] and Oswald [Osw94]. This gives us a hope that the hierarchical multilevel basis corresponding to the above choice of the operators π_k may yield a stable Riesz basis if τ is sufficiently small.

2.2. **Preliminary estimates.** For an analysis of the multilevel basis (2.4) we need some auxiliary estimates already presented in Vassilevski and Wang [VW96a].

The following result on estimating the error $e_j = (\pi_j - Q_j)v$ will play an important role in our analysis:

Lemma 2.1. *Let C_R be a mesh-independent upper bound of the L^2-norm for the operator $I_s - I_{s-1} : V_s \to V_s$. Then,*

$$\|R_{s-1} v\|_0 \le C_R \tau \|v\|_0 \quad \textit{for all } v \in V_s \cdot \tag{2.5}$$

For a given $\sigma \in (0, 1]$, assume that $\tau > 0$ is sufficiently small such that,

$$(1 + C_R \tau) \frac{1}{2^\sigma} \le q = \textit{Const} < 1 \cdot \tag{2.6}$$

Then, there exits an absolute constant C such that for $e_j = (\pi_j - Q_j)v$, $v \in V_k$ there holds:

$$\sum_{j=1}^{k} h_j^{-2\sigma} \|e_j\|_0^2 \le C\tau^2 \sum_{j=1}^{k} h_j^{-2\sigma} \|(Q_j - Q_{j-1})v\|_0^2 \le C\tau^2 \sigma_N \|v\|_\sigma^2 \quad \forall v \in V_k \cdot \tag{2.7}$$

Remark 2.1. *Note that in order to have L^2–stability of the deviations one has to assume that the tolerance τ is level dependent, i.e., one needs a tolerance $\tau = \mathcal{O}(J^{-1})$. Then*

$$\sum_{s=1}^{k} \|e_{s-1}\|_0^2 \le C \|v\|_0^2, \quad \textit{for all } v \in V_k \cdot \tag{2.8}$$

Lemma 2.2. *Let $V_k^1 = (I - M_{k-1}) V_k^{(1)}$, with $V_k^{(1)} \equiv (I_k - I_{k-1}) V_k$, be the modified hierarchical subspace of level k for any given L^2–bounded operators M_j. Then, there are positive constants c_1 and c_2 independent of k such that for any $\psi_1 = (I - M_{k-1})\phi_1 1 \in V_k^1$, with $\phi_1 \in V_k^{(1)}$,*

$$c_1 \|\phi_1\|_r^2 \le \|\psi_1\|_r^2 \le c_2 \|\phi_1\|_r^2, \qquad r = 0, 1. \tag{2.9}$$

We recall, that $\|.\|_1$ stands for the norm in the Sobolev space $H_0^1(\Omega)$ and $\|.\|_0$ denotes the $L^2(\Omega)$–norm.

Lemma 2.3. *Given v and let $v_1^{(k)} = (\pi_k - \pi_{k-1})v$. There exists a sufficiently small constant $\tau_0 > 0$ such that if the approximate projections Q_k^a satisfy (2.2) with $\tau \in (0, \tau_0)$ (see (2.6)), then*

$$\|v\|_1^2 \simeq \sum_{k=0}^{J} h_k^{-2} \|v_1^{(k)}\|_0^2.$$

For proofs of Lemmas 2.1, 2.2, and 2.3, see Vassilevski and Wang [VW97].

2.3. **Stability analysis.** Here we study the Riesz property of the wavelet–like multilevel hierarchical basis defined in (2.4); that is, we show the H^1-stability of the approximate wavelet basis defined in (2.4).

For any $v \in V$ let

$$v = \sum_{x_i \in \mathcal{N}_0} c_{0,i} \phi_i^{(0)} + \sum_{k=1}^{J} \sum_{x_i \in \mathcal{N}_k^{(1)}} c_{k,i} (I - Q_{k-1}^a) \phi_i^{(k)} \tag{2.10}$$

be its representation with respect to the approximate wavelet basis. The corresponding coefficient norm of v is given by

$$|||v||| = \left(h_0^{d-2} \sum_{x_i \in \mathcal{N}_0} c_{0,i}^2 + \sum_{k=1}^{J} h_k^{d-2} \sum_{x_i \in \mathcal{N}_k^{(1)}} c_{k,i}^2 \right)^{1/2}, \quad (d = 2,\ 3). \tag{2.11}$$

Our main result in this section is the following norm equivalence:

Theorem 2.1. *There exists a small (but fixed) $\tau_0 > 0$ such that if the approximate projections Q_k^a satisfy (2.2) with $\tau \in (0, \tau_0)$, then there are positive constants c_1 and c_2 satisfying*

$$c_1 |||v|||^2 \le \|v\|_1^2 \le c_2 |||v|||^2 \qquad \forall v \in V. \tag{2.12}$$

In other words, the modified hierarchical basis is a stable Riesz basis for the second order elliptic and Stokes problems. The equivalence relation (2.12) shall be abbreviated as $|||v|||^2 \simeq \|v\|_1^2$.

Proof. We first rewrite (2.10) as follows:

$$v = \sum_{k=0}^{J} v_1^{(k)}, \tag{2.13}$$

where, with $Q_{-1}^a = 0$,

$$v_1^{(k)} = \sum_{x_i \in \mathcal{N}_k^{(1)}} c_{k,i} (I - Q_{k-1}^a) \phi_i^{(k)} \in V_k^1. \tag{2.14}$$

Furthermore, by letting $\phi^{(k)} = \sum_{x_i \in \mathcal{N}_k^{(1)}} c_{k,i} \phi_i^{(k)} \in V_k^{(1)}$ we see that $v_1^{(k)} = (I - Q_{k-1}^a)\phi^{(k)}$.
Thus, by using (2.9) in Lemma 2.2 (with $r = 0$ and $M_{k-1} = Q_{k-1}^a$) we obtain

$$\|\phi^{(k)}\|_0^2 \simeq \|v_1^{(k)}\|_0^2. \tag{2.15}$$

Since $\phi^{(k)} \in V_k^{(1)}$, then

$$\|\phi^{(k)}\|_0^2 \simeq h_k^d \sum_{x_i \in \mathcal{N}_k^{(1)}} c_{k,i}^2.$$

Combining the above with (2.15) yields

$$|\!|\!| v |\!|\!|^2 \simeq \sum_{k=0}^{J} h_k^{-2} \|v_1^{(k)}\|_0^2.$$

This, together with Lemma 2.3, completes the proof of the theorem. □

Corollary 2.1. *For any fixed $\sigma \in (0,1]$, denote*

$$|\!|\!| v |\!|\!|_\sigma^2 = \sum_{k=0}^{J} h_k^{-2\sigma} \|v_1^{(k)}\|_0^2, \quad v_1^{(k)} = (\pi_k - \pi_{k-1})v \cdot$$

Then,

$$|\!|\!| v |\!|\!|_\sigma \simeq \|v\|_\sigma,$$

for all $v \in H_0^\sigma(\Omega)$ (the latter space is defined by interpolation between $H_0^1(\Omega)$ and $L^2(\Omega)$, see e.g. Bramble [Bra95]*) restricted to the finite element space $V = V_J$. The constants in the norm equivalence depend on σ as indicated in (2.6). For $\sigma = 0$ as mentioned in Remark 2.1, we have to assume a tolerance $\tau = \mathcal{O}(J^{-1})$.*

3. Algebraic Multi–Level Iteration (AMLI) preconditioners for A^TA

In this section we construct optimal order preconditioners for solving systems of linear algebraic equations $A\mathbf{x} = \mathbf{b}$ transformed in the normal form $A^TA\mathbf{x} = A^T\mathbf{b}$ which may be useful for solving convection–diffusion finite element elliptic equations and also in certain domain embedding methods (when A is symmetric).

Consider the homogeneous Dirichlet boundary value problem for the following second-order elliptic equation:

$$\mathcal{L}(u) \equiv -\nabla \cdot (a(x)\nabla u) + \mathbf{b}(x) \cdot \nabla u + c(x)u = f(x), \quad x \in \Omega, \tag{3.1}$$

where $a = a(x)$ is a symmetric and positive definite matrix with bounded and measurable entries; $\mathbf{b} = \mathbf{b}(x)$ and $c = c(x)$ are given bounded functions; $f = f(x)$ is a function in $H^{-1}(\Omega)$.

Note that we do not have here in mind the singularly perturbed case of convection dominated problems of the form (3.1).

A weak form for the Dirichlet problem of (3.1) seeks $u \in H_0^1(\Omega)$ satisfying

$$a(u,v) = f(v) \qquad \forall v \in H_0^1(\Omega), \tag{3.2}$$

where

$$a(u,v) = \int_\Omega (a(x)\nabla u \cdot \nabla v + \mathbf{b}(x) \cdot \nabla u \; v + c(x)uv)\, dx$$

and $f(v)$ is the action of the linear functional f on v.

Let us approximate (3.2) by using the Galerkin method with, say, continuous piecewise linear polynomials. If $V = V_h$ denotes the finite element space associated with a prescribed triangulation of Ω with mesh size h, then the Galerkin approximation is given as the solution of the following problem: *Find* $u_h \in V_h$ *satisfying*

$$a(u_h, \phi) = f(\phi) \qquad \forall \phi \in V_h. \tag{3.3}$$

It has been shown that the discrete problem (3.3) has a unique solution when the mesh size h is sufficiently small. Details can be found in [Sch74], [SW96].

To shorten the exposition, we assume that $a(\cdot,\cdot)$ is H_0^1–coercive.

We summarize the assumptions on $a(.,.)$:

- $H_0^1 \times H_0^1$*–boundedness*, i.e., there exists a positive constant γ_2 such that

$$a(v,w) \le \gamma_2 (a_0(v,v))^{\frac{1}{2}} (a_0(w,w))^{\frac{1}{2}}, \quad \text{for all } v,\ w \in H_0^1(\Omega); \tag{3.4}$$

- H_0^1*–coercivity*, i.e., there exists a constant γ_1 such that

$$a(v,v) \ge \gamma_1^2 a_0(v,v) \quad \text{for all } v \in H_0^1(\Omega); \tag{3.5}$$

Here, $a_0(.,.)$ is, for example, the principal symmetric and positive definite part of $a(.,.)$.

3.1. Bounds of the local projection operators.

We consider any local projection operator $\pi_k : \mathcal{C} \to V_k$ where $\mathcal{C}$ is for example space of continuous functions that in particular contains $V = V_J$ the finite element space under consideration.

Let

$$a_0(\pi_k v, \pi_k v) \le \eta^2(k_0) a_0(\pi_{k+k_0} v, \pi_{k+k_0} v) \quad \text{for all } v \in \mathcal{C}. \tag{3.6}$$

We assume that the norm–bound η of π_k in energy–norm may only depend on the level difference $k_0 \ge 0$. Our main application, though will be when $\eta(k_0)$ is independent of k_0, i.e., for stable local projection operators π_k.

Of main interest are the discretization operators $A_k : V_k \to V_k$ defined by the identity $(A_k\varphi, \psi) = a(\varphi, \psi)$ for all φ and $\psi \in V_k$.

Let P_k be the associated with $a(.,.)$ (non–symmetric) elliptic projection operator $P_k : H_0^1 \to V_k$, i.e., $a(P_k v, \psi) = a(v, \psi)$ for all $\psi \in V_k$. Similarly, define $P_k^\star : H_0^1 \to V_k$ by $a(\psi, P_k^\star v) = a(\psi, v)$ for all $\psi \in V_k$.

Let also λ_k be the largest eigenvalue of the symmetric positive definite operator associated with $a_0(.,.)$ restricted to V_k, i.e, $\lambda_k = \sup\limits_{v \in V_k} \frac{a_0(v,v)}{\|v\|_0^2}$.

We assume:

- *Assumption I:* (FULL ELLIPTIC REGULARITY) There hold the following optimal L^2–error estimates:

$$\begin{aligned} \lambda_k \|v - P_k v\|_0^2 &\le \sigma_R a_0(v - P_k v, v - P_k v), \quad \text{for all } v \in H_0^1, \\ \lambda_k \|v - P_k^\star v\|_0^2 &\le \sigma_R a_0(v - P_k^\star v, v - P_k^\star v), \quad \text{for all } v \in H_0^1. \end{aligned} \tag{3.7}$$

The first error estimate in (3.7) implies:

$$\lambda_k a_0(v - P_k v, v - P_k v) \le \sigma_A^2 \|Av\|_0^2, \quad \text{for all } v \in V\cdot \tag{3.8}$$

We also assume the following standard inverse estimate:

- *Assumption II:* (INVERSE ESTIMATE)

$$\lambda_m \simeq h_m^{-2} \simeq 2^{2m} h_0^{-1}. \tag{3.9}$$

A main result of this section is the following estimate:

$$\|A_k \pi_k v\|_0 \le \sigma_E \eta(k_0)\, \|A_{k+k_0} v\|_0 \quad \text{for all } v \in V_{k+k_0}. \tag{3.10}$$

Based on the boundedness estimate for $a(.,.)$ one has:

$$\begin{aligned} \|A_k \psi\|_0^2 &= a(\psi, A_k\psi) \\ &\le \gamma_2 \sqrt{a_0(\psi,\psi)}\sqrt{a_0(A_k\psi, A_k\psi)} \\ &\le \gamma_2 \sqrt{\lambda_k}\sqrt{a_0(\psi,\psi)}\|A_k\psi\|_0, \end{aligned}$$

which shows

$$\begin{aligned} \|A_k \psi\|_0 &\le \gamma_2 \sqrt{\lambda_k}\sqrt{a_0(\psi,\psi)} \\ &\le \tfrac{\gamma_2}{\gamma_1}\sqrt{\lambda_k}\sqrt{a(\psi,\psi)} \quad \text{for all } \psi \in V_k. \end{aligned} \tag{3.11}$$

Consider the estimates:

$$\begin{aligned} \|A_k\pi_k v\|_0 &\le \|A_k\pi_k(v - P_k v)\|_0 + \|A_k P_k v\|_0 \\ &\le \gamma_2 \lambda_k^{\frac{1}{2}} \left(a_0(\pi_k(v-P_kv), \pi_k(v-P_kv))\right)^{\frac{1}{2}} + \|A_kP_kv\|_0 \\ &\le \gamma_2\eta(k_0)\ \left(\lambda_k a_0(v-P_kv, v-P_kv)\right)^{\frac{1}{2}} + \|A_kP_kv\|_0 \\ &\le \gamma_2\eta(k_0)\sigma_A\|Av\|_0 + \|A_kP_kv\|_0. \end{aligned} \tag{3.12}$$

Now, noting that

$$\|A_kP_kv\|_0^2 = a(P_kv, A_kP_kv) = a(v, A_kP_kv) = (Av, A_kP_kv) \le \|Av\|_0\|A_kP_kv\|_0,$$

we get

$$\|A_kP_kv\|_0 \le \|Av\|_0. \tag{3.13}$$

Combining estimates (3.12–3.13), estimate (3.10) then follows letting $\sigma_E = \gamma_2\sigma_A + 1$ and $A = A_{k+k_0}$.

3.2. **Spectral equivalence estimates for the two–level preconditioner.** In this susbsection we derive some spectral relations and define a two–level preconditioner based on the direct space decomposition $V_{k+k_0} = (I - \pi_k)V_{k+k_0} \oplus V_k$.

One then immediately gets the following spectral equivalence result: Let $(Sv_2, v_2) \equiv \inf_{v \in V_{k+k_0}:\, \pi_k v = v_2} \|A_{k+k_0}v\|_0^2$, $v \in V_{k+k_0}$, define the Schur complement of $B_{k+k_0} \equiv A_{k+k_0}^T A_{k+k_0}$ with respect to the direct decomposition $V_{k+k_0} = (I-\pi_k)V_{k+k_0} \oplus V_k$. I.e., let

$$A_{k+k_0} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{matrix} \} \ (I-\pi_k)V_{k+k_0} \\ \} \ V_k \end{matrix} . \tag{3.14}$$

Note that $A_{22} = A_k$.
Then $S = B_{22} - B_{21}B_{11}^{-1}B_{12}$ where $B_{k+k_0} = \{B_{rs}\}_{r,s=1}^2$ $(= A_{k+k_0}^T A_{k+k_0})$. That is,

$$\begin{aligned} B_{11} &= A_{11}^TA_{11} + A_{21}^TA_{21} \\ B_{12} &= A_{11}^TA_{12} + A_{21}^TA_{22} \\ B_{21} &= A_{12}^TA_{11} + A_{22}^TA_{21} \\ B_{22} &= A_{22}^TA_{22} + A_{12}^TA_{12} = B_k + A_{12}^TA_{12}. \end{aligned}$$

Noting then that $(A_{12}v_2, A_{12}v_2) = a(v_2, A_{12}v_2) = a(v_2, P_k^\star A_{12}v_2) = (A_kv_2, P_k^\star A_{12}v_2) \le \|P_k^\star\|_0\|A_{12}v_2\|_0\|A_kv_2\|_0 \le \sigma(k_0)\|A_{12}v_2\|_0\|A_kv_2\|_0$, i.e.,

$$\|A_{12}v_2\|_0 \le \sigma(k_0)\ \|A_kv_2\|_0, \quad \text{where } v_2 = \pi_kv,\ v_1 = (I-\pi_k)v,$$

one gets the estimates, from (3.10) and $(Sv_2, v_2) \leq (B_k v_2, v_2) + \|A_{12}v_2\|_0^2$,

$$(B_k v_2, v_2) \leq \sigma_E^2 \eta^2(k_0)(Sv_2, v_2), \ (Sv_2, v_2) \leq (1 + \sigma^2(k_0))(B_k v_2, v_2)$$

for all $v_2 \in V_k$.
Therefore, we proved the following result.

Lemma 3.1. *The Schur complement* $S : V_k \to V_k$ *of* $B_{k+k_0} = A_{k+k_0}^T A_{k+k_0}$ *with respect to the direct decomposition* $V_{k+k_0} = (I - \pi_k)V_{k+k_0} \oplus V_k$ *and the operator* $B_k = A_k^T A_k$ *satisfy the spectral estimates:*

$$\frac{1}{1+\sigma^2(k_0)}(Sv_2, v_2) \leq (B_k v_2, v_2) \leq \sigma_E^2 \eta^2(k_0)\ (Sv_2, v_2) \quad \textit{for all } v_2 \in V_k.$$

The constant $\sigma(k_0)$ *is defined from the following norm bound:*

$$\|P_k^\star v\|_0 \leq \sigma(k_0)\|v\|_0, \quad \textit{for all } v \in V_{k+k_0},$$

and admits the following asymptotic behavior with respect to $k_0 \to \infty$*:*

$$\sigma(k_0) \leq C h_k / h_{k+k_0} \simeq 2^{k_0} \cdot \tag{3.15}$$

Proof. Estimate (3.15) is seen by duality, i.e., from *Assumption I* and the behavior of $\lambda_m \simeq h_m^{-2}$, i.e., *Assumption II* (which is actually, standard inverse estimate). To see that, consider the estimates:

$$\begin{aligned} a_0(v - P_k^\star v, v - P_k^\star v) \ &\leq \gamma_1^{-2} a(v - P_k^\star v, v - P_k^\star v) \\ &= \gamma_1^{-2} a(v, v - P_k^\star v) \\ &\leq \gamma_2 \gamma_1^{-2} \left[a_0(v,v)\right]^{\frac{1}{2}} \left[a_0(v - P_k^\star v, v - P_k^\star v)\right]^{\frac{1}{2}}, \end{aligned}$$

which imply,

$$a_0(v - P_k^\star v, v - P_k^\star v) \leq \gamma_2^2/\gamma_1^4 \ a_0(v,v) \quad \text{for all } v \cdot \tag{3.16}$$

Consider then the estimates (based on *Assumption I* and (3.16)):

$$\begin{aligned} \|v - P_k^\star v\|_0 \ &\leq C h_k \ \left(a_0(v - P_k^\star v, v - P_k^\star v)\right)^{\frac{1}{2}} \\ &\leq C h_k \gamma_2/\gamma_1^2 \ (a_0(v,v))^{\frac{1}{2}} \\ &\leq C h_k \sqrt{\lambda_{k+k_0}} \ \|v\|_0 \\ &\leq C h_k / h_{k+k_0} \ \|v\|_0, \end{aligned}$$

which shows the desired asymptotic behavior of $\sigma(k_0)$ since $\|P_k^\star v\|_0 \leq \|v\|_0 + \|v - P_k^\star v\|_0 \simeq \frac{h_k}{h_{k+k_0}} \|v\|_0$. □

As a corollary, one may formulate the following two-level method for B_{k+k_0}.

Definition 3.1. *The operator*

$$M_{TL} = \begin{bmatrix} \widehat{B}_{11} & 0 \\ B_{21} & \widehat{B}_k \end{bmatrix} \begin{bmatrix} I & \widehat{B}_{11}^{-1} B_{12} \\ 0 & I \end{bmatrix}$$

defines a two-level preconditioner to B_{k+k_0}. *Here* $\widehat{B}_{11}$ *is symmetric positive definite approximation to* B_{11} *and similarly,* $\widehat{B}_k$ *is symmetric positive definite approximation to* B_k *or to* $S_D \equiv B_{22} - B_{21}\widehat{B}_{11}^{-1}B_{12}$.

The analysis of the two–level preconditioner is based on Lemma 3.1 and in the case of $\widehat{B} = S_D$, on the estimate

$$(3.17) \qquad \frac{1}{1+\sigma^2(k_0)}(S_D v_2, v_2) \leq (B_k v_2, v_2) \leq \sigma_E^2 \eta^2(k_0)\ (S_D v_2, v_2) \quad \text{for all } v_2 \in V_k,$$

which is proved in the same way as Lemma 3.1 using the inequalities:

$$\begin{aligned}(S_D v_2, v_2) &\leq (B_{22} v_2, v_2)\\ &\leq (1+\sigma^2(k_0))(B_k v_2, v_2)\\ &\leq (1+\sigma^2(k_0))\sigma_E^2\eta^2(k_0)(S v_2, v_2)\\ &\leq (1+\sigma^2(k_0))\sigma_E^2\eta^2(k_0)(S_D v_2, v_2).\end{aligned}$$

Here, we have also assumed that $\widehat{B}_{11}$ is properly scaled, namely, that $(B_{11}v_1, v_1) \leq (\widehat{B}_{11}v_1, v_1)$ for all v_1, which implies the inequality $(Sv_2, v_2) \leq (S_D v_2, v_2)$, used in the last line above. One can also show that $(B_{11}v_1, v_1) \leq (1+\sigma^2(k_0))\sigma_E^2\eta^2(k_0)(Bv, v)$, based on Lemma 3.1 and inequality $(B_{22}v_2, v_2) \leq (1+\sigma^2(k_0))\sigma_E^2\eta^2(k_0)(Sv_2, v_2)$. Thus, B_{11} and B_{22} allow for approximations and the relative condition number of the two–level preconditioner M_{TL} with respect to B, can be estimated in terms of $\mathcal{H}(k_0) \equiv (1+\sigma^2(k_0))\sigma_E^2\eta^2(k_0)$ and the constants involved in the spectral equivalence relations between $\widehat{B}_{11}$ and B_{11} and between $\widehat{B}_k$ and B_k. More details are found from Vassilevski [Vas97], see also Bank [Ban96] and the analysis in the next subsection.

3.3. The AMLI preconditioner. In this section we define the AMLI preconditioner and present the analysis of its relative condition number with respect to $B = B_{k+k_0}$.

One can generalize the preconditioner from Definition 3.1 to the multi–level case in a standard way, cf. Axelsson and Vassilevski [AV90], Vassilevski [Vas92] or the survey paper by Vassilevski [Vas97].

Definition 3.2. *Let $M_0 = A_0$. For $s = 1, 2, \ldots, \left[\frac{J}{k_0}\right]$ and $m = min\{J, sk_0\}$, $k = (s-1)k_0$ consider $B_m = \{B_{r,s}\}_{r,s=1}^2$. The operator*

$$M_m = \begin{bmatrix} \widehat{B}_{11} & 0 \\ B_{21} & \widetilde{S}_k \end{bmatrix} \begin{bmatrix} I & \widehat{B}_{11}^{-1}B_{12} \\ 0 & I \end{bmatrix}$$

defines the Algebraic Multi–Level Iteration (AMLI) preconditioner to B_m. Here, $\widehat{B}_{11}$ is symmetric positive definite approximation to B_{11}, and

$$(3.18) \qquad \widetilde{S}_k^{-1} \equiv \left[I - p_\nu(M_k^{-1}S_D)\right] S_D^{-1},\ S_D \equiv B_{22} - B_{21}\widehat{B}_{11}^{-1}B_{12}$$

where

$$(3.19) \qquad p_\nu(t) = \frac{1+T_\nu\left(\frac{\alpha_1+\alpha_2-2t}{\alpha_2-\alpha_1}\right)}{1+T_\nu\left(\frac{\alpha_1+\alpha_2}{\alpha_2-\alpha_1}\right)},$$

and $T_\nu(t)$ is the Chebyshev polynomial of first kind and degree $\nu = \nu_{k_0} > 1$. Also, $\alpha_2 = 1+\sigma^2(k_0)$ and $\alpha \equiv \alpha_1\sigma_E^2\eta^2(k_0)$ is a lower bound of the minimum eigenvalue of $M_k^{-1}B_k$, which has to be estimated.

Remark 3.1. *To implement one action of M_m^{-1}, as readily seen, one performs two inverse actions of $\widehat{B}_{11}$, matrix–vector products with the blocks B_{21} and B_{12} of B as well as ν inverse actions of M_k (already defined by induction) and also $\nu - 1$ actions of S_D which is based on the actions of the blocks of B, B_{21}, B_{22} and B_{12} and one more action of $\widehat{B}_{11}^{-1}$ for each action of S_D.*

For the analysis of the AMLI preconditioner M_m the following estimate will be useful:

$$(Sv_2, v_2) \leq (S_D v_2, v_2) \text{ all } v_2, \tag{3.20}$$

provided $\widehat{B}_{11}$ is scaled such that

$$(B_{11}v_1, v_1) \leq (\widehat{B}_{11}v_1, v_1) \leq (1 + \beta(k_0))\,(B_{11}v_1, v_1) \quad \text{for all } v_1. \tag{3.21}$$

We also have:

$$(B_{22}v_2, v_2) \leq (1 + \sigma^2(k_0))(B_k v_2, v_2) \leq (1 + \sigma^2(k_0))\sigma_E^2 \eta^2(k_0)\,(Sv_2, v_2) \quad \text{for all } v_2.$$

This estimate, letting

$$\mathcal{H}(k_0) = (1 + \sigma^2(k_0))\sigma_E^2 \eta^2(k_0), \tag{3.22}$$

implies the strengthened Cauchy–Schwarz inequality

$$(B_{21}v_1, v_2) \leq \gamma(k_0)\,[(B_{11}v_1, v_1)]^{\frac{1}{2}}\,[(B_{22}v_2, v_2)]^{\frac{1}{2}} \quad \text{for all } v_1, v_2, \tag{3.23}$$

where $\gamma(k_0) = \sqrt{1 - \frac{1}{\mathcal{H}(k_0)}}$.

The latter inequality in turn implies the estimate,

$$(B_{11}v_1, v_1) \leq \mathcal{H}(k_0)\,(B_m v, v), \quad \text{for all } v = v_1 + v_2,\ v_2 = \pi_k v. \tag{3.24}$$

Choose now $\nu > \mathcal{H}(k_0)$. Then there exists a sufficiently small $\alpha > 0$ such that the following inequality holds:

$$1 + \beta(k_0)\mathcal{H}(k_0) + \mathcal{H}^2(k_0)\frac{(1-\widetilde{\alpha})^\nu}{\alpha\left[\sum_{l=1}^{\nu}(1+\sqrt{\widetilde{\alpha}})^{\nu-l}(1-\sqrt{\widetilde{\alpha}})^{l-1}\right]^2} \leq \frac{1}{\alpha},\ \widetilde{\alpha} = \frac{\alpha}{\mathcal{H}(k_0)}. \tag{3.25}$$

We note that (3.25) (after multiplying it by α) reduces to $\frac{1}{\nu^2}\mathcal{H}^2(k_0) \leq 1$ by letting $\alpha \to 0$.

We are now in a position to prove the main result concerning the spectral equivalence relations between the AMLI–preconditioner M_m and B_m.

Theorem 3.1. *Assume that $\nu > \mathcal{H}(k_0)$ and let α satisfy inequality (3.25). Then, the following estimates hold:*

$$(M_m v, v) \leq (B_m v, v) \leq \frac{1}{\alpha}\,(M_m v, v) \quad \textit{for all } v \in V_m. \tag{3.26}$$

Proof. The proof follows from a standard induction argument. We have $M_0 = B_0$, hence (3.26) holds for any $\alpha \leq 1$. Assume now, that for some $s \geq 1$ and $k = (s-1)k_0$

(3.26) is valid. Then for $((M_m - B_m)v, v) = ((\widehat{B}_{11} - B_{11})v_1, v_1) + ((\widetilde{S}_k - S_D)v_2, v_2)$ first the following estimate hold:

$$0 \leq ((M_m - B_m)v, v), \quad \text{for all } v,$$

due to the choice of the polynomial p_ν in (3.19) (and α as in (3.25)), the induction assumption, and (3.17).

We will use next the following technical fact,

$$\sup\left\{\frac{p_\nu(t)}{1-p_\nu(t)} : t \in \left[\alpha\frac{1}{\sigma_E^2\eta^2(k_0)},\ 1+\sigma^2(k_0)\right]\right\} = \frac{\sup\limits_{t\in[\alpha_1,\,\alpha_2]} p_\nu(t)}{1 - \sup\limits_{t\in[\alpha_1,\,\alpha_2]} p_\nu(t)}$$

and that $\sup\limits_{t\in[\alpha_1,\,\alpha_2]} p_\nu(t) = \frac{2}{1+T_\nu\left(\frac{\alpha_2+\alpha_1}{\alpha_2-\alpha_1}\right)}$, where $\alpha_1 = \frac{\alpha}{\sigma_E^2\eta^2(k_0)}$ and $\alpha_2 = 1+\sigma^2(k_0)$. Therefore, for $\widetilde{\alpha} = \frac{\alpha_1}{\alpha_2} = \frac{\alpha}{\mathcal{H}(k_0)}$, $\sup\limits_{t\in[\alpha_1,\,\alpha_2]} p_\nu(t) = \frac{4q^\nu}{(1+q^\nu)^2}$, $q = \frac{1-\sqrt{\widetilde{\alpha}}}{1+\sqrt{\widetilde{\alpha}}}$, which in particular implies that

$$\begin{aligned}\frac{\sup\limits_{t\in[\alpha_1,\,\alpha_2]} p_\nu(t)}{1-\sup\limits_{t\in[\alpha_1,\,\alpha_2]} p_\nu(t)} &= \frac{4q^\nu}{(1-q^\nu)^2} = \frac{4(1-\sqrt{\widetilde{\alpha}})^\nu(1+\sqrt{\widetilde{\alpha}})^\nu}{\left[(1+\sqrt{\widetilde{\alpha}})^\nu-(1-\sqrt{\widetilde{\alpha}})^\nu\right]^2}\\ &= \frac{(1-\widetilde{\alpha})^\nu}{\widetilde{\alpha}\left[\sum\limits_{l=1}^{\nu}(1+\sqrt{\widetilde{\alpha}})^{\nu-l}(1-\sqrt{\widetilde{\alpha}})^{l-1}\right]^2} \simeq \frac{\mathcal{H}(k_0)}{\alpha\nu^2}, \quad \alpha \to 0.\end{aligned}$$

Also, by the choice of $\widehat{B}_{11}$ (see (3.21)), the spectral relations (3.17) and inequality (3.25), we have:

$$\begin{aligned}&((M_m - B_m)v, v) \leq \beta(k_0)(B_{11}v_1, v_1)\\ &\qquad +(S_D(I - p_\nu(M_k^{-1}S_D))^{-1}v_2, v_2) - (S_Dv_2, v_2)\\ &\leq \beta(k_0)\mathcal{H}(k_0)(B_mv, v)\\ &\qquad +(S_Dv_2, v_2)\sup\left\{\tfrac{p_\nu(t)}{1-p_\nu(t)} : t \in \left[\alpha\lambda_{\min}(B_k^{-1}S_D),\ \lambda_{\max}(B_k^{-1}S_D)\right]\right\}\\ &\leq \beta(k_0)\mathcal{H}(k_0)(B_mv, v)\\ &\qquad +(B_{22}v_2, v_2)\sup\left\{\tfrac{p_\nu(t)}{1-p_\nu(t)} : t \in \left[\alpha\lambda_{\min}(B_k^{-1}S_D),\ \lambda_{\max}(B_k^{-1}S_D)\right]\right\}\\ &\leq \beta(k_0)\mathcal{H}(k_0)(B_mv, v)\\ &\qquad +(Sv_2, v_2)\mathcal{H}(k_0)\sup\left\{\tfrac{p_\nu(t)}{1-p_\nu(t)} : t \in \left[\alpha\tfrac{1}{\sigma_E^2\eta^2(k_0)},\ 1+\sigma^2(k_0)\right]\right\}\\ &\leq \left\{\beta(k_0)\mathcal{H}(k_0) + \frac{\mathcal{H}^2(k_0)}{\alpha}\frac{(1-\widetilde{\alpha})^\nu}{\left[\sum\limits_{l=1}^{\nu}(1+\sqrt{\widetilde{\alpha}})^{\nu-l}(1-\sqrt{\widetilde{\alpha}})^{l-1}\right]^2}\right\}(B_mv, v)\\ &\leq (\tfrac{1}{\alpha} - 1)(B_mv, v).\end{aligned}$$

The latter inequality completes the proof. □

Now, we emphasize the asymptotic behavior of $\mathcal{H}(k_0)$. Assume at this point that the local projection operator $\pi_k : V_{k+k_0} \to V_k$ is H_0^1–bounded, i.e., $\eta(k_0)$ is independent of k_0. Such operators are available based on approximate wavelet modification of the classical HB (hierarchical basis) as described in the previous section (see, Definition 2.1). Then, one has:

$$\mathcal{H}(k_0) \simeq \sigma^2(k_0) \simeq 2^{2k_0}, \ k_0 \to \infty. \tag{3.27}$$

From complexity requirement we also have the restriction on ν,

$$2^{dk_0} > \nu\,.$$

The latter inequality together with $\nu > \mathcal{H}(k_0) \simeq 2^{2k_0}$ show the following main result:

Theorem 3.2. *In three space dimensions, $d = 3$, the AMLI preconditioner M_m is both, of optimal complexity, and spectrally equivalent to B_m if k_0 is sufficiently large and the polynomial degree ν satisfies $2^{dk_0} > \nu > \mathcal{H}(k_0) \simeq 2^{2k_0}$.*

3.4. **Estimation of the condition number of B_{11}.** In this section we show that the condition number of the major block B_{11} on the diagonal of B_{k+k_0} grows at most like $\left(\frac{h_k}{h_{k+k_0}}\right)^3$, hence for k_0 fixed B_{11} is well-conditioned.

We begin, for any $v_1 = (I - \pi_k)v$, with the standard approximation property of the L^2-projection operator Q_k:

$$\|Q_k v_1 - v_1\|_0 \leq C h_k \sqrt{a_0(v_1, v_1)} \leq C h_k \gamma_1^{-1} \sqrt{(A_{11} v_1, v_1)}.$$

Then, based on the deviation estimate (2.7) (recall that π_k is close to the exact L^2-projection operator Q_k) and on the coercivity estimate (note that $\pi_k v_1 = 0$):

$$\|Q_k v_1\|_0 = \|(Q_k - \pi_k) v_1\|_0 \leq C h_k \sqrt{a_0(v_1, v_1)} \leq C h_k \gamma_1^{-1} \sqrt{(A_{11} v_1, v_1)},$$

one arrives at the following approximation estimate:

$$\|v_1\|_0^2 \leq (\|Q_k v_1\|_0 + \|v_1 - Q_k v_1\|_0)^2 \leq c_1^{-1} \lambda_k^{-1} (A_{11} v_1, v_1) \leq c_1^{-1} \lambda_k^{-1} \|A_{11} v_1\|_0 \|v_1\|_0.$$

Therefore, one gets

$$\lambda_k \|v_1\|_0 \leq c_1^{-1} \|A_{11} v_1\|_0. \tag{3.28}$$

We also have, using estimate (3.11) and the proven L^2-coercivity of A_{11}, (3.28),

$$\begin{aligned}
\|A_{21} v_1\|_0 &\leq \|A v_1\|_0 \\
&\leq \sqrt{\lambda_{k+k_0}}\, \tfrac{\gamma_2}{\gamma_1} (A v_1, v_1)^{\frac{1}{2}} \\
&= \sqrt{\lambda_{k+k_0}}\, \tfrac{\gamma_2}{\gamma_1} (A_{11} v_1, v_1)^{\frac{1}{2}} \\
&\leq \sqrt{\lambda_{k+k_0}} \tfrac{\gamma_2}{\gamma_1} (\|A_{11} v_1\|_0 \|v_1\|_0)^{\frac{1}{2}} \\
&\leq \sqrt{\tfrac{\lambda_{k+k_0}}{c_1 \lambda_k}} \tfrac{\gamma_2}{\gamma_1} \|A_{11} v_1\|_0.
\end{aligned}$$

Then, since $B_{11} = A_{11}^T A_{11} + A_{21}^T A_{21}$, using (3.11), the latter estimate, and the coercivity one (3.28), the required eigenvalue bounds stated in the next theorem are proven.

Theorem 3.3. *The following estimates hold:*

$$\begin{aligned}
c_1^2 \lambda_k^2 \leq \frac{\|A_{11} v_1\|_0^2}{\|v_1\|_0^2} \leq \frac{(B_{11} v_1, v_1)}{\|v_1\|_0^2} &\leq \left[1 + \left(\frac{\gamma_2}{\gamma_1}\right)^2 \frac{\lambda_{k+k_0}}{c_1 \lambda_k}\right] \frac{\|A_{11} v_1\|_0^2}{\|v_1\|_0^2} \\
&\leq \left[1 + \left(\frac{\gamma_2}{\gamma_1}\right)^2 \frac{\lambda_{k+k_0}}{c_1 \lambda_k}\right] \gamma_2^2 \lambda_{k+k_0}^2.
\end{aligned}$$

That is, the condition number of B_{11} grows at most like $\left(\frac{\lambda_{k+k_0}}{\lambda_k}\right)^3 \simeq 2^{6k_0}$. Therefore, for bounded k_0 (as required in our application), the block B_{11} will be well-conditioned and hence relatively easy to approximate (such as in (3.21)).

Remark 3.2. *Note that the results of Theorem 3.3 are proven without the full regularity Assumption I.*

4. Bounded extension operators

In this section we review one more application of the bounded local projection operators π_k to construct H^1–bounded extension operators, first used in Bramble and Vassilevski [BV96]. Here we emphasize on the application of such bounded extension operators in the case of constructing preconditioners for finite element elliptic equations by the domain embedding technique.

The main fact used in the presented domain embedding technique is that one can transform the given problem based on bounded extension operators to a more stable form that allows for approximate solutions (preconditioners) in the embedding (more regular) domain Ω. We also take advantage of an algebraic fact (perhaps first proven here) that a strengthened Cauchy–Schwarz inequality for the matrix implies a strengthened Cauchy–Schwarz inequality for the inverse matrix with the same constant.

4.1. **Wavelet–like extension operators.**

In this section we describe a general framework of constructing computationally feasible H^1–bounded extension operators $\mathcal{E}$ that extend data defined on a interface boundary Γ into the interior of subdomains $\{\Omega_i\}_{i=1}^p$ that compose the original domain Ω. In the domain embedding application (to be considered in the following section) we will have $p = 2$ and Ω_1 will be a given in some sense irregular domain. Ω will be the embedding, computationally more regular, domain (e.g., parallelepiped) which will contain Ω_1. Then $\Omega_2 = \Omega \setminus \overline{\Omega}_1$. Finally, the interface Γ will be (a part) of $\partial\Omega_1$ across which Ω_1 is embedded in Ω, i.e., $\Gamma = \partial\Omega_1 \cap \partial\Omega_2$,

Consider the bilinear form $a(u, \varphi) = \int\limits_\Omega a\nabla u \cdot \nabla\varphi \, dx$, u, $\varphi \in H_0^1(\Omega)$ and let $V = V_J$ be the finite element space of continuous piecewise linear functions corresponding to a triangulation $\mathcal{T}_h$ which we assume is obtained by $J \geq 1$ successive steps of uniform refinement of a an initial coarse triangulation $\mathcal{T}_0$ of Ω. We also assume that the non–overlapping domain partitioning of Ω

$$\Omega = \Gamma \cup \Omega_1 \cup \Omega_2 \cup \ldots \cup \Omega_p, \tag{4.1}$$

consists of subdomains Ω_i that are coarse–grid domains (that is each Ω_i is completely covered by elements from $\mathcal{T}_0$). This implies that the interface Γ is also completely covered by boundaries of elements from $\mathcal{T}_0$. Let h_0 be the meshsize corresponding to $\mathcal{T}_0$, then $h_k = 2^{-k}h_0$ will be the meshsize corresponding to $\mathcal{T}_k$ if we divide each edge of the elements of $\mathcal{T}_{k-1}$ in two equal parts to create the elements of $\mathcal{T}_k$.

The coefficient $a = a(x)$, $x \in \Omega$ which can be a $d \times d$ symmetric positive definite matrix, bounded uniformly in $\overline{\Omega}$. It is is assumed that a vary smoothly in each Ω_i but may have large jumps across Γ. For the purpose of constructing preconditioners for $a(.,.)$ it is sufficient to assume that a is piecewise constant with respect to the partition $\{\Omega_i\}$ of Ω. Let

$$\begin{aligned} \rho_{i,\max} &= \sup_{x\in\Omega_i} \max_{\underline{\xi}\in\mathbb{R}^d} \frac{\underline{\xi}^T a(x)\underline{\xi}}{\underline{\xi}^T\underline{\xi}} \\ \rho_{i,\min} &= \inf_{x\in\Omega_i} \min_{\underline{\xi}\in\mathbb{R}^d} \frac{\underline{\xi}^T a(x)\underline{\xi}}{\underline{\xi}^T\underline{\xi}}, \end{aligned} \tag{4.2}$$

be bound of the variation of the coefficient matrix $a(x)$ in each subdomain Ω_i, $i = 1, 2, \ldots, p$.

Definition 4.1 (Extension operator). *Define the following extension operator* $\mathcal{E}_0 : V|_\Gamma \to V$, *by*

$$\mathcal{E}_0\varphi|_{\Omega_i} = \sum_{k=1}^{J} E_k^0(q_k^i - q_{k-1}^i)\varphi_i, \quad \varphi_i = \varphi|_{\partial\Omega_i} \cdot$$

Here E_k^0 *is the trivial extension of any function* $\psi_i \in V_k|_{\partial\Omega_i}$ *by zero at the interior kth level nodes in* Ω_i *and* $q_k^i : L^2(\partial\Omega_i) \to V_k|_{\partial\Omega_i}$ *are the* L^2*–projection operators restricted to the boundaries of* Ω_i.

It is clear then that

$$a(\mathcal{E}_0\varphi, \mathcal{E}_0\varphi) \le \sum_{i=1}^{p} \rho_{i,\max} \int_{\Omega_i} \nabla\mathcal{E}_0\varphi_i \cdot \nabla\mathcal{E}_0\varphi_i dx \cdot \tag{4.3}$$

Using the following $|.|_1$–seminorm characterization of space $H^1(\Omega_i) \cap V_k|_{\Omega_i}$

$$|v|_{1,\Omega_i}^2 \simeq \inf_{v:\, v=\sum_{k=1}^{J} v_k,\, v_k\in V_k|_{\Omega_i}} \sum_{k=1}^{J} h_k^{-2}\|v_k\|_{0,\Omega_i}^2,$$

used for $v := \mathcal{E}_0\varphi_i$ and replacing v_k with $E_k^0(q_k^i - q_{k-1}^i)\varphi_i$ in (4.3), we end up with the following upper bound,

$$\begin{aligned} a(\mathcal{E}_0\varphi, \mathcal{E}_0\varphi) &\le \sum_{i=1}^{p} \rho_{i,\max} \sum_{k=1}^{J} h_k^{-2}\|E_k^0(q_k^i - q_{k-1}^i)\varphi_i\|_{0,\Omega_i}^2 \\ &\le \tau_1^{-1} \sum_{i=1}^{p} \rho_{i,\max} \sum_{k=1}^{J} h_k^{-1}\|(q_k^i - q_{k-1}^i)\varphi_i\|_{0,\partial\Omega_i}^2 \cdot \end{aligned} \tag{4.4}$$

Next, use the quasi–optimality of the $L^2(\partial\Omega_i)$ projections $\{q_k^i\}_{k=0}^J$; namely:

$$\sum_{k=1}^{J} h_k^{-1}\|(q_k^i - q_{k-1}^i)\varphi_i\|_{0,\partial\Omega_i}^2 \simeq \inf_{\varphi_i=\sum_{k=1}^{J}\varphi_i^{(k)},\, \varphi_i^{(k)}\in V_k|_{\partial\Omega_i}} \sum_{k=1}^{J} h_k^{-1}\|\varphi_i^{(k)}\|_{0,\partial\Omega_i}^2 \cdot \tag{4.5}$$

Given an arbitrary $v \in V$ such that $v|_{\partial\Omega_i} = \varphi_i$. consider the decomposition,

$$v|_{\Omega_i} = Q_0^i(v|_{\Omega_i}) + \sum_{k=1}^{J}(Q_k^i - Q_{k-1}^i)(v|_{\Omega_i}) \cdot$$

Here Q_k^i are the $L^2(\Omega_i)$ projections onto the spaces $V_k|_{\Omega_i}$.

Letting $\varphi_i^{(k)} = (Q_k^i - Q_{k-1}^i)(v|_{\Omega_i})\big|_{\partial\Omega_i}$ we get $\varphi_i = Q_0^i(v|_{\Omega_i})\big|_{\partial\Omega_i} + \sum\limits_{k=1}^{J} \varphi_i^{(k)}$. Therefore, we can use this decomposition in (4.5) to get:

$$
\begin{aligned}
\sum_{k=1}^{J} h_k^{-1} \|(q_k^i - q_{k-1}^i)\varphi_i\|_{0,\,\partial\Omega_i}^2 &\leq C \sum_{k=1}^{J} h_k^{-1} \|\,(Q_k^i - Q_{k-1}^i)(v|_{\Omega_i})\big|_{\partial\Omega_i}\,\|_{0,\,\partial\Omega_i}^2 \\
&\leq C \sum_{k=1}^{J} h_k^{-2} \|(Q_k^i - Q_{k-1}^i)(v|_{\Omega_i})\|_{0,\,\Omega_i}^2 \\
&\leq C \|\nabla v\|_{0,\,\Omega_i}^2 \\
&\leq C \rho_{i,\min}^{-1} \int_{\Omega_i} a(x) \nabla v \cdot \nabla \, v dx \cdot
\end{aligned} \tag{4.6}
$$

Here, we used a standard inverse inequality to bound boundary integrals $\int_{\partial\Omega_i}$ by domain integrals $\int_{\Omega_i}$.

Combining (4.4) and (4.6), by summation over i, and taking infimum over $v \in V$ such that $v|_{\partial\Omega_i} = \varphi_i$, the following main norm–bound is proved:

Theorem 4.1. *The extension operator $\mathcal{E}_0$ defined in Definition 4.1 is H^1–bounded and satisfies an estimate of the form:*

$$a(\mathcal{E}_0\phi, \mathcal{E}_0\phi) \leq \eta_0 \inf_{v \in V:\ v|_\Gamma = \phi} a(v, v) \quad \textit{for any } \phi \in V|_\Gamma \cdot$$

Here, η_0 depends only on the local variation of the coefficient matrix $a(x)$ defined in (4.2), i.e., on $\max\limits_{1 \leq i \leq p} \frac{\rho_{i,max}}{\rho_{i,min}}$.

For practical computations, however we have to replace the exact $L^2(\partial\Omega_i)$–projection operators q_k^i by their computationally feasible approximations π_k^i as introduced in Section 2, Definition 2.1 (defined respectively for the finite element spaces restricted to the boundaries $\partial\Omega_i$).

Definition 4.2 (Wavelet–like extension operator). *Define the following extension operator $\mathcal{E}\ : V|_\Gamma \to V$, by*

$$\mathcal{E}\varphi|_{\Omega_i} = \sum_{k=1}^{J} E_k^0 (\pi_k^i - \pi_{k-1}^i)\varphi_i, \quad \varphi_i = \varphi|_{\partial\Omega_i} \cdot \tag{4.7}$$

Then using Corollary 2.1 (based on Lemma 2.1 for $\sigma = \frac{1}{2}$ and the norm characterization of $H^{\frac{1}{2}}(\partial\Omega_i)$) one has the following modification of Theorem 4.1.

Theorem 4.2. *The extension operator $\mathcal{E}$ defined in Definition 4.2 is H^1–bounded and satisfies an estimate of the form:*

$$a(\mathcal{E}\phi, \mathcal{E}\phi) \leq \eta \inf_{v \in V:\ v|_\Gamma = \phi} a(v, v) \quad \textit{for any } \phi \in V|_\Gamma \cdot \tag{4.8}$$

Here, η depends only on the local variation of the coefficient matrix $a(x)$ defined in (4.2), i.e., on $\max\limits_{1 \leq i \leq p} \frac{\rho_{i,max}}{\rho_{i,min}}$.

4.2. **Approximate harmonic extension operator and related strengthened Cauchy–Schwarz inequality.** Let A be a stiffness matrix coming from a finite element discretization of a second order elliptic problem on a domain Ω with Dirichlet boundary conditions. Let now Ω_1 be a subdomain of Ω which is exactly covered by elements of the given triangulation of Ω. In practice we are interested in a problem formulated in Ω_1 which for a reason of simplicity we then embed in a more standard computational domain Ω such as rectangle or parallelepiped in three dimensions. Let the interface boundary across which Ω_1 is embedded in Ω be Γ. We consider, for the original problem formulated in Ω_1 Dirichlet boundary conditions. In such a way, if we use domain decomposition ordering of the degrees of freedom in Ω, with respect to the partitioning $\Omega = \Omega_1 \cup \Gamma \cup \Omega_2$, $\Omega_2 \equiv \Omega \setminus \overline{\Omega}_1$, the stiffness matrix A_1 associated with the original problem will be a principal submatrix of A, the matrix associated with the problem in the extended domain Ω, i.e., we will have,

$$A = \begin{array}{c} \begin{bmatrix} A_1 & A_{10} & 0 \\ A_{01} & A_0 & A_{02} \\ 0 & A_{20} & A_2 \end{bmatrix} \end{array} \begin{array}{l} \}\ \Omega_1 \\ \}\ \Gamma \\ \}\ \Omega_2 \equiv \Omega \setminus \overline{\Omega}_1 \end{array} \tag{4.9}$$

The methods that we will be interested in for solving the problem

$$A_1 \mathbf{x}_1 = \mathbf{b}_1, \tag{4.10}$$

will be based on the existence of efficient preconditioners for the matrix A, i.e., for problems in the standard computational domain Ω.

For ease of presentation we will reorder the block matrix A into the following two-by-two block form:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{array}{l} \}\ \Omega_1 \cup \Omega_2 \\ \}\ \Gamma \end{array}. \tag{4.11}$$

Note that A_{11} is block diagonal with blocks on its main diagonal A_1 and A_2. Hence if we derive a preconditioner $\widehat{B}_{11}$ for A_{11}, it will induce in a natural way a preconditioner B_1 for the original matrix A_1 whose inverse actions can be computed via $B_1^{-1}\mathbf{v}_1 = \left(\widehat{B}_{11}^{-1} \begin{bmatrix} \mathbf{v}_1 \\ 0 \end{bmatrix} \begin{array}{l} \}\ \Omega_1 \\ \}\ \Omega_2 \end{array}\right)\Big|_{\Omega_1}$.

For the construction of the preconditioners for A_{11} we will need a mapping $\mathcal{E}$ that transforms data given on the interface boundary Γ to data in the interior of $\Omega \setminus \Gamma = \Omega_1 \cup \Omega_2$. The requirement that we impose on this mapping is to be bounded in the energy norm defined by the bilinear form $a(.,.)$ from which the stiffness matrix A is computed. In vector–matrix notation, we have a (rectangular) matrix $\begin{bmatrix} E_{12} \\ I \end{bmatrix} \begin{array}{l} \}\ \Omega \setminus \Gamma = \Omega_1 \cup \Omega_2 \\ \}\ \Gamma \end{array}$, such that the following estimate holds:

$$\left(\begin{bmatrix} E_{12} \\ I \end{bmatrix} \mathbf{v}_2 \right)^T A \left(\begin{bmatrix} E_{12} \\ I \end{bmatrix} \mathbf{v}_2 \right) \leq \eta \inf_{\mathbf{v}_1} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}^T A \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}. \tag{4.12}$$

Here η is a positive constant (≥ 1) independent of the mesh size. In other words, the mapping $\mathcal{E}$ that extends a given data on Γ into the interior of the subdomains Ω_1 and Ω_2 is approximate "harmonic" with respect to the given bilinear form $a(.,.)$. Using functions–bilinear form notation we can write (4.12) as

$$a(\mathcal{E}v_2, \mathcal{E}v_2) \leq \eta \inf_{v|_\Gamma = v_2} a(v, v),$$

i.e., the extension provided by the mapping $\mathcal{E}$ is quasi–optimal. The case $\eta = 1$ corresponds to the exact "harmonic" extension, i.e., $E_{12} = -A_{11}^{-1}A_{12}$. The latter mapping is impractical since it would require exact solutions in the subregions and this was our original intend (i.e., to solve a problem posed in Ω_1 with a coefficient matrix A_1; recall that $A_{11} = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{matrix} \}\,\Omega_1 \\ \}\,\Omega_2 \end{matrix}$). A possible choice of computationally feasible approximate harmonic extension mappings $\mathcal{E}$ is given in Definition 4.2 (see aslo Theorem 4.2).

Consider now the (square) transformation matrix,

$$E = \begin{bmatrix} I & E_{12} \\ 0 & I \end{bmatrix} \begin{matrix} \} & \Omega_1 \cup \Omega_2 \\ \} & \Gamma \end{matrix}$$

and then consider the following transformed matrix (for this approach we refer to Vassilevski and Axelsson [VA94] and Bramble and Vassilevski [BV96]),

$$\widehat{A} = E^T A E = \begin{bmatrix} \widehat{A}_{11} & \widehat{A}_{12} \\ \widehat{A}_{21} & \widehat{A}_{22} \end{bmatrix}. \tag{4.13}$$

One easily derives the relations:

$$\begin{aligned} \widehat{A}_{11} &= A_{11} \\ \widehat{A}_{12} &= A_{12} + A_{11}E_{12} \\ \widehat{A}_{21} &= A_{21} + E_{12}^T A_{11} \\ \widehat{A}_{22} &= \begin{bmatrix} E_{12} \\ I \end{bmatrix}^T A \begin{bmatrix} E_{12} \\ I \end{bmatrix}. \end{aligned} \tag{4.14}$$

The important observation is that the first block has not been changed after transforming A to $\widehat{A}$. Also notice the form of the block $\widehat{A}_{22}$ of $\widehat{A}$. On its basis the estimate (4.12) takes the equivalent form:

$$\mathbf{v}_2^T \widehat{A}_{22} \mathbf{v}_2 \le \eta \inf_{\widehat{\mathbf{v}}_1} \begin{bmatrix} \widehat{\mathbf{v}}_1 \\ \mathbf{v}_2 \end{bmatrix}^T \widehat{A} \begin{bmatrix} \widehat{\mathbf{v}}_1 \\ \mathbf{v}_2 \end{bmatrix} = \eta \inf_{\mathbf{v}_1} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}^T A \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} \quad (\widehat{\mathbf{v}}_1 = \mathbf{v}_1 - E_{12}\mathbf{v}_2)\,. \tag{4.15}$$

It will be demonstrated in what follows that the transformed form of A, $\widehat{A}$ is more stable in the sense that it allows for certain approximations of the blocks on the main diagonal of $\widehat{A}$ and also allows for approximations to its corresponding Schur complements and the same applies for the block–entries and respective Schur complements of the inverse of $\widehat{A}$. The main tool in proving these facts is the following strengthened Cauchy–Schwarz inequality valid for the two–by–two block structure of $\widehat{A}$:

$$\mathbf{v}_2^T \widehat{A}_{21} \widehat{\mathbf{v}}_1 \le \gamma \sqrt{\widehat{\mathbf{v}}_1^T A_{11} \widehat{\mathbf{v}}_1} \sqrt{\mathbf{v}_2^T \widehat{A}_{22} \mathbf{v}_2}, \quad \text{for all } \widehat{\mathbf{v}}_1,\ \mathbf{v}_2, \tag{4.16}$$

where $\gamma = \sqrt{1 - \frac{1}{\eta}}$, i.e., $\gamma \in [0, 1)$ (strictly less than one). This follows from inequality (4.15) as shown in Vassilevski [Vas92] (see also the proof of Theorem 4.3).

We will be needing the following corollaries of the strengthened Cauchy–Schwarz inequality (found already in Axelsson and Gustafsson [AG83]).

Lemma 4.1. *Let A be a symmetric positive definite, two–by–two block matrix,*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

that satisfies the strengthened Cauchy–Schwarz (or Cauchy–Bunyakowsky–Schwarz) inequality,

$$\mathbf{v}_2^T A_{21} \mathbf{v}_1 \leq \gamma \, \left[\mathbf{v}_1^T A_{11} \mathbf{v}_1\right]^{\frac{1}{2}} \, \left[\mathbf{v}_2^T A_{22} \mathbf{v}_2\right]^{\frac{1}{2}} \quad \textit{for all } \mathbf{v}_1, \ \mathbf{v}_2,$$

for some positive constant γ strictly less than one. Consider now the following Schur complements of A:

$$\begin{aligned} S_1 &= A_{11} - A_{12} A_{22}^{-1} A_{21} \ \textit{and} \\ S_2 &= A_{22} - A_{21} A_{11}^{-1} A_{12}. \end{aligned}$$

Then the following inequalities hold:

$$\begin{aligned} \mathbf{v}_1^T S_1 \mathbf{v}_1 &\geq (1-\gamma^2) \mathbf{v}_1^T A_{11} \mathbf{v}_1 \quad \textit{for all } \mathbf{v}_1 \\ \mathbf{v}_2^T S_2 \mathbf{v}_2 &\geq (1-\gamma^2) \mathbf{v}_2^T A_{22} \mathbf{v}_2 \quad \textit{for all } \mathbf{v}_2. \end{aligned} \tag{4.17}$$

4.3. The strengthened Cauchy–Schwarz inequality for the inverse matrix. In this section we prove the main result on which our construction of the domain embedding preconditioners is based. Namely, we show that given a two–by–two block matrix A, symmetric positive definite, for which a strengthened Cauchy–Schwarz inequality holds with a constant $\gamma \in [0,1)$ (as in Lemma 4.1), then for the induced two–by–two block structure of A^{-1} the same strengthened Cauchy–Schwarz inequality, with the same constant γ holds.

Theorem 4.3. *Let A be a symmetric positive definite, two–by–two block matrix,*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}. \tag{4.18}$$

We assume that there exists a positive constant γ, strictly less than one, such that the following strengthened Cauchy–Schwarz (or Cauchy–Bunyakowsky–Schwarz) inequality holds:

$$\mathbf{v}_2^T A_{21} \mathbf{v}_1 \leq \gamma \, \left[\mathbf{v}_1^T A_{11} \mathbf{v}_1\right]^{\frac{1}{2}} \, \left[\mathbf{v}_2^T A_{22} \mathbf{v}_2\right]^{\frac{1}{2}} \quad \textit{for all } \mathbf{v}_1, \ \mathbf{v}_2 \, . \tag{4.19}$$

Consider then $B = A^{-1} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$ partitioned according to the block partitioning of A as in (4.18). Then for the same constant $\gamma \in (0,1)$, the following strengthened Cauchy–Schwarz inequality holds:

$$\mathbf{v}_2^T B_{21} \mathbf{v}_1 \leq \gamma \, \left[\mathbf{v}_1^T B_{11} \mathbf{v}_1\right]^{\frac{1}{2}} \, \left[\mathbf{v}_2^T B_{22} \mathbf{v}_2\right]^{\frac{1}{2}} \quad \textit{for all } \mathbf{v}_1, \ \mathbf{v}_2 \, . \tag{4.20}$$

Or equivalently, we have:

$$\mathbf{v}_1^T B_{11} \mathbf{v}_1 \leq \frac{1}{1-\gamma^2} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}^T B \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}, \quad \textit{for all } \mathbf{v}_1, \ \mathbf{v}_2 \, . \tag{4.21}$$

Proof. Consider the following block–factorization form of A,

$$A = \begin{bmatrix} I & 0 \\ A_{21} A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & S_2 \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1} A_{12} \\ 0 & I \end{bmatrix},$$

where $S_2 = A_{22} - A_{21} A_{11}^{-1} A_{12}$ is the corresponding Schur complement. Note that S_2 is symmetric and positive definite. This block–factorization form of A implies the following block factorization of $B = A^{-1}$:

$$B = \begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & S_2^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix}.$$

This factorization in turn shows the following exact representation of $B = A^{-1}$:

$$B = \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}S_2^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}S_2^{-1} \\ -S_2^{-1}A_{21}A_{11}^{-1} & S_2^{-1} \end{bmatrix},$$

i.e., we have,

$$\begin{aligned} B_{11} &= A_{11}^{-1} + A_{11}^{-1}A_{12}S_2^{-1}A_{21}A_{11}^{-1}, \\ B_{12} &= -A_{11}^{-1}A_{12}S_2^{-1} \\ B_{21} &= -S_2^{-1}A_{21}A_{11}^{-1}, \text{ and} \\ B_{22} &= S_2^{-1}. \end{aligned} \tag{4.22}$$

As it is shown in Vassilevski [Vas92], the strengthened Cauchy–Schwarz inequality (4.20) and the inequality (4.21) are equivalent. This is seen by looking at the non-negative quadratic form

$$(\mathbf{v} + t\mathbf{w})^T B(\mathbf{v} + t\mathbf{w}) - (1-\gamma^2)\mathbf{v}_1^T B_{11}\mathbf{v}_1 \geq 0,$$

for any real t and any vectors $\mathbf{v}$, $\mathbf{w}$ of the form $\mathbf{v} = \left[\begin{smallmatrix} \mathbf{v}_1 \\ 0 \end{smallmatrix}\right]$ and $\mathbf{w} = \left[\begin{smallmatrix} 0 \\ \mathbf{v}_2 \end{smallmatrix}\right]$. I.e.,

$$\gamma^2 \mathbf{v}_1^T B_{11}\mathbf{v}_1 + 2\mathbf{v}_2^T B_{21}\mathbf{v}_1 t + \mathbf{v}_2^T B_{22}\mathbf{v}_2 t^2 \geq 0 \cdot$$

Its discriminant $D = (\mathbf{v}_2^T B_{21}\mathbf{v}_1)^2 - \gamma^2\mathbf{v}_1^T B_{11}\mathbf{v}_1\mathbf{v}_2^T B_{22}\mathbf{v}_2$ must be non–positive, and this is inequality (4.20). The converse is also true – the non–positivity of the discriminant D (which is the strengthened Cauchy–Schwarz inequality) implies that the above quadratic form is nonnegative, and this is (for $t = 1$) inequality (4.21).

To show (4.21), we use the formulas for the block entries of B given in (4.22).

The given strengthened Cauchy–Schwarz inequality (4.3) for A implies the following relations between matrix blocks of A (due to Lemma 4.1):

$$\mathbf{v}_2^T S_2\mathbf{v}_2 \geq (1-\gamma^2)\mathbf{v}_2^T A_{22}\mathbf{v}_2 \quad \text{for all } \mathbf{v}_2,$$

which implies

$$\mathbf{v}_2^T A_{22}^{-1}\mathbf{v}_2 \geq (1-\gamma^2)\mathbf{v}_2^T S_2^{-1}\mathbf{v}_2 \quad \text{for all } \mathbf{v}_2.$$

The latter inequality used for $\mathbf{v}_2 = A_{21}\mathbf{v}_1$ reads,

$$\mathbf{v}_1^T A_{12}A_{22}^{-1}A_{21}\mathbf{v}_1 \geq (1-\gamma^2)\mathbf{v}_1^T A_{12}S_2^{-1}A_{21}\mathbf{v}_1 \quad \text{for all } \mathbf{v}_1. \tag{4.23}$$

Similarly, for the other Schur complement $S_1 \equiv A_{11} - A_{12}A_{22}^{-1}A_{21}$, the original inequality also implies (due to Lemma 4.1),

$$\mathbf{v}_1^T(A_{11} - A_{12}A_{22}^{-1}A_{21})\mathbf{v}_1 = \mathbf{v}_1^T S_1\mathbf{v}_1 \geq (1-\gamma^2)\mathbf{v}_1^T A_{11}\mathbf{v}_1 \quad \text{for all } \mathbf{v}_1,$$

or which is the same,

$$\gamma^2\mathbf{v}_1^T A_{11}\mathbf{v}_1 \geq \mathbf{v}_1^T A_{12}A_{22}^{-1}A_{21}\mathbf{v}_1 \quad \text{for all } \mathbf{v}_1 \cdot$$

The last inequality, together with (4.23) imply

$$\frac{\gamma^2}{1-\gamma^2}\mathbf{v}_1^T A_{11}\mathbf{v}_1 \geq \mathbf{v}_1^T A_{12}S_2^{-1}A_{21}\mathbf{v}_1 \quad \text{for all } \mathbf{v}_1 \cdot$$

The latter inequality is equivalent to (by letting $\mathbf{v}_1 := A_{11}^{-1}\mathbf{v}_1$)

$$\frac{\gamma^2}{1-\gamma^2}\mathbf{v}_1^T A_{11}^{-1}\mathbf{v}_1 \geq \mathbf{v}_1^T A_{11}^{-1} A_{12} S_2^{-1} A_{21} A_{11}^{-1}\mathbf{v}_1 \quad \text{for all } \mathbf{v}_1 \cdot$$

Then we have,

$$\begin{aligned}\tfrac{1}{1-\gamma^2}\mathbf{v}_1^T A_{11}^{-1}\mathbf{v}_1 &= (1+\tfrac{\gamma^2}{1-\gamma^2})\mathbf{v}_1^T A_{11}^{-1}\mathbf{v}_1 \\ &\geq \mathbf{v}_1^T(A_{11}^{-1} + A_{11}^{-1}A_{12}S_2^{-1}A_{21}A_{11}^{-1})\mathbf{v}_1 \\ &= \mathbf{v}_1^T B_{11}\mathbf{v}_1, \quad \text{for all } \mathbf{v}_1\cdot\end{aligned}$$

The latter inequality shows, since $A_{11}^{-1} = B_{11} - B_{12}B_{22}^{-1}B_{21}$ is a Schur complement of the symmetric positive definite matrix B, that

$$\mathbf{v}_1^T B_{11}\mathbf{v}_1 \leq \frac{1}{1-\gamma^2}\inf_{\mathbf{v}_2}\left[\begin{smallmatrix}\mathbf{v}_1\\ \mathbf{v}_2\end{smallmatrix}\right]^T B \left[\begin{smallmatrix}\mathbf{v}_1\\ \mathbf{v}_2\end{smallmatrix}\right]\cdot$$

Thus the proof is complete. □

4.4. **The construction of preconditioners for A_{11} on the basis of available preconditioners for A.** In this section we present our approach of deriving preconditioners for the principal submatrix A_{11} of A (where A is partitioned in a two–by–two block form as in (4.18)) on the basis of any available preconditioner for A itself. Here we assume that either A itself satisfies a strengthened Cauchy–Schwarz inequality or that it can be transformed to $\widehat{A} = E^T A E$, with $E = \left[\begin{smallmatrix} I & E_{12}\\ 0 & I\end{smallmatrix}\right]$, for which a strengthened Cauchy–Schwarz inequality holds. The crucial part of the construction of the preconditioners is that the actions of E_{12} and E_{12}^T on vectors $\mathbf{v}_2$ and $\mathbf{v}_1$, respectively, be computable. For the domain embedding application E_{12} comes from the approximate harmonic extension mapping $\mathcal{E}$ (see (4.12)).

Let now M be a given symmetric positive definite preconditioner to A and let the following spectral equivalence relations hold:

$$\gamma_1\mathbf{v}^T A\mathbf{v} \leq \mathbf{v}^T M\mathbf{v} \leq \gamma_2\mathbf{v}^T A\mathbf{v}, \quad \text{for all } \mathbf{v}, \tag{4.24}$$

for some positive constants γ_1, γ_2. Hence M^{-1} will be spectrally equivalent to $B \equiv A^{-1}$. Then $E^{-1}M^{-1}E^{-T}$ will be spectrally equivalent to $\widehat{B} \equiv \widehat{A}^{-1} = E^{-1}BE^{-T}$. As a corollary, we obtain that $\widehat{B}_{11} = [I,\ 0]\widehat{B}\left[\begin{smallmatrix} I\\ 0\end{smallmatrix}\right]$ will be spectrally equivalent to $[I,\ 0]E^{-1}M^{-1}E^{-T}\left[\begin{smallmatrix} I\\ 0\end{smallmatrix}\right] = [I,\ -E_{12}]M^{-1}\left[\begin{smallmatrix} I\\ -E_{12}^T\end{smallmatrix}\right]$. By Theorem 4.3, $\widehat{B}_{11}$ is spectrally equivalent to A_{11}^{-1} ($= \widehat{A}_{11}^{-1} = \widehat{B}_{11} - \widehat{B}_{12}\widehat{B}_{22}^{-1}\widehat{B}_{21}$, see (4.14) and (4.22)), hence

$$[I,\ 0]E^{-1}M^{-1}E^{-T}\left[\begin{smallmatrix} I\\ 0\end{smallmatrix}\right] = [I,\ -E_{12}]M^{-1}\left[\begin{smallmatrix} I\\ -E_{12}^T\end{smallmatrix}\right]$$

provides the inverse actions of a spectrally equivalent preconditioner to A_{11}.

More specifically, from (4.24) used in the following equivalent form

$$\gamma_2^{-1}\mathbf{v}^T\widehat{B}\mathbf{v} \leq \mathbf{v}^T E^{-1}M^{-1}E^{-T}\mathbf{v} \leq \gamma_1^{-1}\mathbf{v}^T\widehat{B}\mathbf{v}, \quad \text{for all } \mathbf{v},$$

we get as a corollary, letting $\mathbf{v} = \left[\begin{smallmatrix}\mathbf{v}_1\\ 0\end{smallmatrix}\right]$,

$$\gamma_2^{-1}\mathbf{v}_1^T\widehat{B}_{11}^{-1}\mathbf{v}_1 \leq \mathbf{v}_1^T[I,\ -E_{12}]M^{-1}\left[\begin{smallmatrix} I\\ -E_{12}^T\end{smallmatrix}\right]\mathbf{v}_1 \leq \gamma_1^{-1}\mathbf{v}_1^T\widehat{B}_{11}\mathbf{v}_1\cdot$$

The latter inequality combined with the inequality from Theorem 4.3,

$$\mathbf{v}_1^T A_{11}^{-1}\mathbf{v}_1 \leq \mathbf{v}_1^T \widehat{B}_{11}^{-1}\mathbf{v}_1 \leq \frac{1}{1-\gamma^2}\mathbf{v}_1^T A_{11}^{-1}\mathbf{v}_1, \quad \text{for all } \mathbf{v}_1,$$

imply the desired spectral equivalence relation:

(4.25)

$$\gamma_2^{-1}\mathbf{v}_1^T A_{11}^{-1}\mathbf{v}_1 \leq \mathbf{v}_1^T [I,\ -E_{12}]M^{-1}\begin{bmatrix} I \\ -E_{12}^T\end{bmatrix}\mathbf{v}_1 \leq \gamma_1^{-1}\frac{1}{1-\gamma^2}\mathbf{v}_1^T A_{11}^{-1}\mathbf{v}_1 \quad \text{for all } \mathbf{v}_1 \cdot$$

That is, we have:

Theorem 4.4. *Let M be a spectrally equivalent preconditioner for A and let $E = \begin{bmatrix} I & E_{12} \\ 0 & I\end{bmatrix}$ with computable actions of E_{12} and E_{12}^T, transform A to $\widehat{A} = E^T A E$, for which a strengthened Cauchy–Schwarz inequality with a constant $\gamma \in [0,1)$ holds. Then the mapping*

$$[I,\ -E_{12}]M^{-1}\begin{bmatrix} I \\ -E_{12}^T\end{bmatrix},$$

provides the inverse actions of a spectrally equivalent preconditioner for the principal submatrix A_{11} of A. More specifically, the spectral equivalence relations (4.25) hold, and hence we have the estimates:

$$Cond\left([I,\ -E_{12}]M^{-1}\begin{bmatrix} I \\ -E_{12}^T\end{bmatrix} A_{11}\right) \leq \frac{1}{1-\gamma^2} Cond\left(M^{-1}A\right) \leq \frac{1}{1-\gamma^2}\frac{\gamma_2}{\gamma_1},$$

where the constants γ_2 and γ_1 are from the spectral equivalence relation (4.24).

References

[AG83] Axelsson O. and Gustafsson I. (1983) Preconditioning and two–level multigrid methods of arbitrary degree of approximations. *Math. Comp.* 40: 219–242.

[AV90] Axelsson O. and Vassilevski P. S. (1990) Algebraic multilevel preconditioning methods, II. *SIAM J. Numer. Anal.* 27: 1569–1590.

[Ban96] Bank R. E. (1996) Hierarchical bases and the finite element method. *Acta Numerica* .

[BD81] Bank R. E. and Dupont T. (1981) An optimal order process for solving finite element equations. *Math. Comp.* 36: 35–51.

[BPX90] Bramble J. H., Pasciak J. E., and Xu J. (1990) Parallel multilevel preconditioners. *Math. Comp.* 55: 1–22.

[Bra95] Bramble J. H. (1995) *Multigrid methods.* Pitman Research Notes in Mathematical Series, # 234, Longman Scientific & Technical, (second edition), UK.

[BV96] Bramble J. H. and Vassilevski P. S. (1996) Wavelet–like extension operators in non–overlapping domain decomposition algorithms. Institute for Scientific Computation, Texas A & M University, College Station, TX (in preparation).

[Dau92] Daubechies I. (1992) *Ten Lectures on Wavelets.* SIAM, Philadelphia, PA.

[GMP93] Goldstein C. I., Manteuffel T. A., and Parter S. V. (1993) Preconditioning and boundary conditions without H^2 estimates: L^2 condition numbers and the distribution of the singular values. *SIAM J. Numer. Anal.* 30: 343–376.

[GO94] Griebel M. and Oswald P. (1994) Tensor product type subspace splittings and multilevel iterative methods for anisotropic problems. *Advances in Computational Mathematics* 4: 171–206.

[KO95] Kotyczka U. and Oswald P. (1995) Piecewise linear prewavelets of small support. In Chui C. and Schumaker L. L. (eds) *Approximation Theory VIII, vol. 2 (wavelets and multilevel approximation)*, pages 235–242. World Scientific, Singapore.

[LO96] Lorentz R. and Oswald P. (1996) Constructing economical Riesz bases for Sobolev spaces. talk, presented at the Domain Decomposition Conference held in Bergen, Norway, June 3–8, 1996, GMD, Bonn.

[Nep91a] Nepomnyaschikh S. V. (1991) Mesh theorems of traces, normalization of function traces and their inversion. *Soviet J. Numer. Anal. and Math. Modelling* 6: 223–242.

[Nep91b] Nepomnyaschikh S. V. (1991) Method of splitting into subspaces for solving boundary value problems in complex form domain. *Soviet J. Numer. Anal. and Math. Modelling* 6: 1–26.

[Osw94] Oswald P. (1994) *Multilevel Finite Element Approximation. Theory and Applications.* Teubner Skripten zur Numerik, Teubner, Stuttgart.

[PV94] Proskurowski W. and Vassilevski P. S. (1994) Preconditioning capacitance matrix problems in domain imbedding. *SIAM J. Sci. Comput.* 15: 77–88.

[Sch74] Schatz A. H. (1974) An observation concerning Ritz–Galerkin methods with indefinite bilinear forms. *Math. Comp.* 28: 959–962.

[Ste95a] Stevenson R. (1995) Robustness of the additive and multiplicative frequency decomposition multilevel method. *Computing* 54: 331–346.

[Ste95b] Stevenson R. (1995) A robust hierarchical basis preconditioner on general meshes. Report # 9533, Department of Mathematics, University of Nijmegen, Nijmegen, The Netherlands.

[SW96] Schatz A. H. and Wang J. (1996) Some new error estimates for Ritz-Galerkin methods with minimal regularity assumptions. *Math. Comp.* 65: 19–27.

[VA94] Vassilevski P. S. and Axelsson O. (1994) A two–level stabilizing framework for interface domain decomposition preconditioners. In Dimov I. T., Sendov B., and Vassilevski P. S. (eds) *Proceedings of the Third International Conference $O(h^3)$, Advances in Numerical Methods and Applications*, pages 196–202. World Scientific, Singapore, New Jersey, London, Hong Kong.

[Vas92] Vassilevski P. S. (1992) Hybrid V–cycle algebraic multilevel preconditioners. *Math. Comp.* 58: 489–512.

[Vas97] Vassilevski P. S. (1997) On two ways of stabilizing the hierarchical basis multilevel methods. *SIAM Review (to appear)* .

[VW96a] Vassilevski P. S. and Wang J. (1996) Stabilizing the hierarchical basis by approximate wavelets, I: Theory. *Numer. Linear Alg. Appl. (submitted)* .

[VW96b] Vassilevski P. S. and Wang J. (1996) Stabilizing the hierarchical basis by approximate wavelets, II: Implementation and Numerical Experiments. *SIAM J. Sci. Comput. (submitted)* .

[VW97] Vassilevski P. S. and Wang J. (1997) Wavelet–like methods in the design of efficient multilevel preconditioners for elliptic pdes. In Dahmen W., Kurdila A., and Oswald P. (eds) *Multiscale Wavelet Methods for PDEs, (Wavelet Analysis and its Applications Series).* Academic Press (to appear).

[Xu92] Xu J. (1992) Iterative methods by space decomposition and subspace correction. *SIAM Review* 34: 581–613.

[Zha92] Zhang X. (1992) Multilevel Schwarz methods. *Numer. Math.* 63: 521–539.

Center of Informatics and Computing Technology/CLPIP, Bulgarian Academy of Sciences, "Acad. G. Bontchev" street, Block 25 A, 1113 Sofia, Bulgaria

E-mail address: panayot@iscbg.acad.bg