



HAL
open science

An Optimal Schwarz Preconditioner for a Class of Parallel Adaptive Finite Elements

Sébastien Loisel, Hieu Nguyen

► **To cite this version:**

Sébastien Loisel, Hieu Nguyen. An Optimal Schwarz Preconditioner for a Class of Parallel Adaptive Finite Elements. 2016. hal-01337957

HAL Id: hal-01337957

<https://hal.science/hal-01337957>

Preprint submitted on 27 Jun 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Optimal Schwarz Preconditioner for a Class of Parallel Adaptive Finite Elements

Sébastien Loisel^a, Hieu Nguyen^{a,1,*}

^a *Department of Mathematics, Heriot-Watt University, Riccarton, Edinburgh, EH14 4AS, United Kingdom*

Abstract

A Schwarz-type preconditioner is formulated for a class of parallel adaptive finite elements where the local meshes cover the whole domain. With this preconditioner, the convergence rate of Krylov methods is shown to depend only on the ratio of the second largest and smallest eigenvalues of the preconditioned system. These eigenvalues can be bounded independently of the mesh sizes and the number of subdomains, which proves the proposed preconditioner is optimal. Numerical results are provided to support the theoretical findings.

Keywords: Domain decomposition, preconditioner, Bank-Holst paradigm, two-grid discretizations, parallel adaptivity

2010 MSC: 65N55, 65N22, 65F08

1. Introduction

Adaptive finite element method (AFEM) has been a very popular method for solving partial differential equations in science and engineering [2]. AFEM automatically refines or coarsens meshes to adapt to the computed solutions, thus offering great reliability, robustness and efficiency. Recently, there has been a great demand to use AFEM on parallel distributed supercomputers with many processors to tackle large-scale problems. In order to improve the scalability of AFEM on supercomputers, it is usually combined with a domain decomposition method (DDM). In DDM, the domain is partitioned into a number of subdomains and smaller problems on these subdomains are solved in parallel to determine the overall solution [30, 34].

Combining AFEM with DDM, however, introduces challenges that are not present in the traditional version of AFEM. One of the notable challenges is that AFEM builds its meshes gradually and global or near-neighbour information is usually needed. The information can be approximated solutions, error estimates on intermediate meshes or mesh information utilised in adaptive meshing procedures. Since communication costs are high on distributed supercomputers, one wants to avoid communicating as much as possible. This can be achieved when each processor has a mesh of the whole domain and its adaptive enrichment is performed almost independently with those of other processors. In general, the adaptive enrichment on each processor focus mainly on its subdomain. Consequently, after the adaptive enrichment phase, each processor has a composite mesh of the whole domain, which is fine in its subdomain and much coarser elsewhere. The final global mesh is the union of the refined submesh provided by each processor. Figure 1 shows an example of the meshes before and after adaptive enrichment, and the final global mesh.

The initial idea of using local meshes of the whole domain was first introduced by Mitchell for a parallel multigrid method [28]. Then it was further developed into parallel adaptive algorithms. The notable ones include the *Bank-Holst algorithm* [10, 11] and the *local and parallel algorithms based on two-grid discretizations* [38, 39, 24]. Several

*Corresponding author

Email addresses: S.Loisel@hw.ac.uk (Sébastien Loisel), hnguyen@cimne.upc.edu (Hieu Nguyen)

¹ *Current address:* CIMNE - Centre Internacional de Metodes Numerics en Enginyeria, Universitat Politècnica de Catalunya, Barcelona, Spain

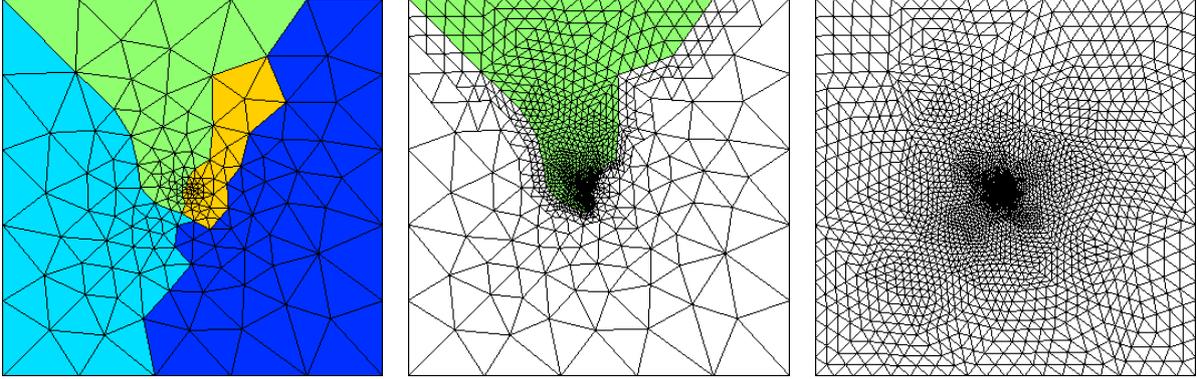


Figure 1: A coarse mesh with its partition (left), a local mesh on a processor after adaptive enrichment (middle), and the global fine mesh.

variants of these algorithms are studied in [12, 8, 36, 19, 40]. The two algorithms and their variants have been demonstrated to work well for many problems in both science and engineering [10, 29, 5, 6, 11, 3, 4, 36, 15, 33, 17, 32].

Different components contribute to their success. For discussions on how to obtain a suitable partition, where each subdomain contributes roughly the same amount of error, we refer to [10, 11]. For how to regularise the local meshes to make the global fine mesh conforming, we refer the readers to [16]. In this paper, we focus on solving the final global linear system. There is no restriction in the type of solvers can be used. However, it would be ideal if the solver can take advantage of the special formulation of the algorithms. In [14], Bank and Lu developed a dedicated domain decomposition solver for the Bank-Holst algorithm. The solver is empirically shown to be robust and efficient for many problems [11, 14, 15, 17]. However, its theoretical convergence can only be fully analysed for a special case where the global interface system is completely presented on all processors [18]. For this to happen, all elements attached to the interface, including ones that are far away from the considered subdomain, are required to be refined to the same level of the corresponding elements in the global fine mesh. In addition, the global iteration matrix of the solver is not symmetric, even if all of the local matrices are symmetric. Consequently, conjugate gradient acceleration can not be used.

In this paper, we propose a novel Additive Schwarz (AS) preconditioner that can be combined with Krylov methods, such as CG, to efficiently solve the global linear system in these parallel adaptive algorithms. Our preconditioner is formulated using the local meshes after adaptive enrichment. We recall that these are meshes of the whole domain. They are fine and identical with the global fine mesh in their corresponding subdomains, but generally much coarser elsewhere. If the adaptive meshes are nested, all the finite element spaces associated with the local meshes contain the coarse space associated with the starting coarse mesh. Therefore, there is no need to explicitly add a coarse space as in the traditional two-level AS. However, having the coarse space contained in every subspace introduces the number of subdomains as the largest eigenvalue, which might damages the scability of the preconditioner. Fortunately, we can show that this largest eigenvalue is isolated and the convergence rate of the CG method can be bounded by a quantity that depends only on the ratio of the second largest eigenvalue and the smallest eigenvalue. The ratio is called the *effective condition number*. Our main theoretical results lies in the analysis of these eigenvalues.

The estimate for the second largest eigenvalue is obtained by establishing a comparison to the largest eigenvalue in a related AS method. Our estimate takes advantage of the strengthened Cauchy-Schwarz inequality for the hierarchical decomposition of local subspaces into a low frequency component and a high frequency component. For estimating the smallest eigenvalue, we follow the subspace correction framework proposed by Xu [37] and prove the existence of a stable decomposition associated with the local meshes. Since these meshes are generally very different from one another and with the global fine mesh outside of their associated subdomains, the classical analysis of AS method (cf. [21, 34]) does not apply. Our analysis requires new sophisticated interpolation operators based on the work of Scott and Zhang [31]. These operators are defined in conjunction with a colouring scheme in order to construct the stable decomposition recursively.

In case exact solvers are employed on all local subspaces, our analysis of the eigenvalues shows that the effective condition number of the preconditioned system does not depend on the coarse mesh size H , the fine mesh size h and the

number of subdomains N ; thus our method is optimal. Roughly speaking, the proposed method performs comparable to a traditional AS method with an extremely thick overlap ($\delta \approx H$). With proper programming, it delivers superior rate of convergence while demanding about the same amount of computation as with traditional AS methods with a small overlap.

In some aspects, our result is related to the work of Bank et al. [13]. However, our preconditioner is very different as we use local subspaces associated with meshes of the whole domain and there is no explicit coarse component.

The rest of this paper is organised as follows. We first state the model problem and introduce key notations in section 2. The formulation of the preconditioner is presented in section 3. The analysis of the convergence of the CG method applied to the preconditioned system, as well as the estimates for the second largest and smallest eigenvalue are carried out in section 4. In section 5, we present some numerical experiments to verify our theoretical results.

2. Preliminaries

For simplicity of exposition, we confine our discussions to Poisson's equation with homogeneous Dirichlet condition:

$$\begin{aligned} -\Delta u(x) &= f(x) & \text{in } \Omega, \\ u(x) &= 0 & \text{on } \partial\Omega. \end{aligned} \quad (1)$$

Here Ω is a bounded domain with polygonal boundary in \mathbb{R}^d , $d = 2, 3$.

Let $\{\Omega_i\}_{i=1}^N$ be the subdomains in the partition of Ω . We assume that this is a non-overlapping partition, namely $\bar{\Omega} = \cup_{i=1}^N \bar{\Omega}_i$ and $\Omega_i \cap \Omega_j = \emptyset$ if $i \neq j$.

In this study, we will use several finite element meshes. The mesh \mathcal{T}_H of size H will be the shape regular and conforming coarse mesh provided to each processor at the beginning. We further assume that each Ω_i is a union of elements in \mathcal{T}_H . The meshes \mathcal{T}_i , $1 \leq i \leq N$ are local meshes on each processor at the end of the adaptive enrichment phase. They are *meshes of the whole domain* which are fine with elements of size $h \ll H$ within Ω_i , but coarser and largely coincide with \mathcal{T}_H elsewhere. The mesh \mathcal{T}_i is required to be conforming inside $\bar{\Omega}_i$. However, it can have hanging nodes outside of $\bar{\Omega}_i$. In addition, we assume that \mathcal{T}_i are aligned along their fine interface, namely if Ω_i and Ω_j are neighbouring subdomains then \mathcal{T}_i and \mathcal{T}_j are matched along the part of interface sharing between Ω_i and Ω_j .

Denote \mathcal{T}_h the union of \mathcal{T}_i restricted on $\bar{\Omega}_i$: $\mathcal{T}_h = \cup_{i=1}^N (\mathcal{T}_i|_{\bar{\Omega}_i})$. This mesh is the globally refined, shape regular and conforming mesh of size h of Ω . We assume the following nesting property holds

$$\mathcal{T}_H \subset \mathcal{T}_i \subset \mathcal{T}_h, \quad \text{for } 1 \leq i \leq N.$$

Now, we extend each Ω_i to a larger region Ω_i^\dagger so that all elements of \mathcal{T}_i that are outside of Ω_i^\dagger belong to \mathcal{T}_H (i.e. there is no refinement in \mathcal{T}_i outside of Ω_i^\dagger). We also require that $\partial\Omega_i^\dagger$ does not cut through any elements in \mathcal{T}_h or any elements in \mathcal{T}_i . The extension can be obtained by repeatedly adding to Ω_i layers of elements in \mathcal{T}_i . Since the adaptive meshing on processor i mainly focuses on the inside of the subdomain Ω_i , we can assume that only few layers of elements in \mathcal{T}_H outside of Ω_i get refined in creating \mathcal{T}_i . More specifically, we assume that the width of the regions $\Omega_i^\dagger \setminus \Omega_i$ are of size H (in case there is barely any refinement outside Ω_i , some elements in $\mathcal{T}_H|_{\Omega_i}$ might need to be included in Ω_i^\dagger). Figure 2 shows an example of a subdomain Ω_i and its extension Ω_i^\dagger . Lastly, we assume that the (overlapping) partition $\{\Omega_i^\dagger\}_{i=1}^N$ of Ω can be coloured using at most N^c colours, in such a way that if Ω_i^\dagger and Ω_j^\dagger are of the same colour and i is different from j , then $\Omega_i^\dagger \cap \Omega_j^\dagger = \emptyset$.

Let V_0 , V_i , and V_h be the linear finite element spaces (of piecewise linear polynomials) associated with \mathcal{T}_H , \mathcal{T}_i and \mathcal{T}_h respectively, i.e.

$$\begin{aligned} V_0 &= \{u_H(x) \in H_0^1(\Omega) \mid u_H(x)|_T \in \mathbb{P}_1(T), \forall T \in \mathcal{T}_H\}, \\ V_i &= \{u_h(x) \in H_0^1(\Omega) \mid u_h(x)|_T \in \mathbb{P}_1(T), \forall T \in \mathcal{T}_i\}, \\ V_h &= \{u_h(x) \in H_0^1(\Omega) \mid u_h(x)|_T \in \mathbb{P}_1(T), \forall T \in \mathcal{T}_h\}. \end{aligned}$$

where $\mathbb{P}_1(T)$ is the set of linear polynomials defined on element T .

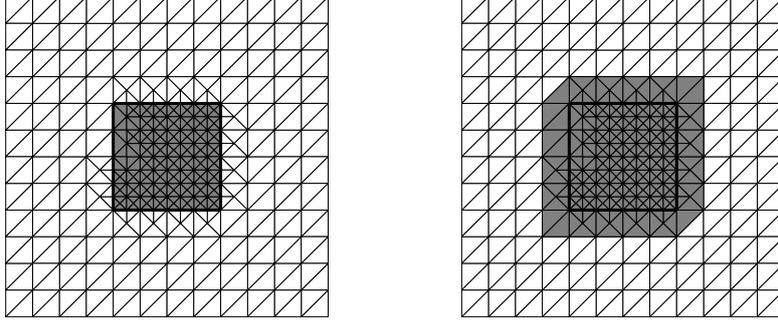


Figure 2: Subdomain Ω_i (left) and its extension Ω_i^\dagger (right) on their associated local mesh \mathcal{T}_i .

Also let $\{\psi_j(x)\}_{j=1}^n$ and $\{\psi_j^{(i)}(x)\}_{j=1}^{n_i}$ be the sets of linear nodal basis function associated with \mathcal{T}_h and \mathcal{T}_i , $i = 0, 1, \dots, N$. Correspondingly, denote $\{x_j\}_{j=1}^n$ and $\{x_j^{(i)}\}_{j=1}^{n_i}$ be the sets of nodal points of \mathcal{T}_h and \mathcal{T}_i , $i = 0, 1, \dots, N$. Here, for convenience, we use \mathcal{T}_0 to refer to \mathcal{T}_H .

The finite element approximation $u_h(x) \in V_h$ of $u(x)$ is the solution of the following problem: find $u_h(x) \in V_h$ such that

$$a(u_h, v_h) = \int_{\Omega} f(x) v_h(x) dx, \quad \text{for all } v_h(x) \in V_h, \quad (2)$$

where $a(u_h, v_h) = \int_{\Omega} (\nabla u_h \cdot \nabla v_h) dx$.

For $u_h(x) \in V_h$, denote $u \in \mathbb{R}^n$ its coordinate vector, i.e., $u_h(x) = \sum_{j=1}^n u(j) \psi_j(x)$. Then the problem (2) becomes

$$Au = f, \quad (3)$$

where $A \in \mathbb{R}^{n \times n}$, $A(k, j) = a(\psi_j, \psi_k)$, and $f \in \mathbb{R}^n$, $f(k) = \int_{\Omega} f(x) \psi_k(x) dx$. Clearly, A is symmetric positive definite and $a(u_h, v_h) = v^T Au =: (u, v)_A$.

3. Preconditioner formulation

We define $R_i^T \in \mathbb{R}^{n \times n_i}$ as follows

$$R_i^T = \begin{bmatrix} \psi_1^{(i)}(x_1) & \psi_2^{(i)}(x_1) & \cdots & \psi_{n_i}^{(i)}(x_1) \\ \psi_1^{(i)}(x_2) & \psi_2^{(i)}(x_2) & \cdots & \psi_{n_i}^{(i)}(x_2) \\ \vdots & \vdots & \cdots & \vdots \\ \psi_1^{(i)}(x_n) & \psi_2^{(i)}(x_n) & \cdots & \psi_{n_i}^{(i)}(x_n) \end{bmatrix}. \quad (4)$$

We note that R_i^T is the matrix representation of the point-wise interpolation operator from V_i , a coarser mesh with the basis $(\psi_1^{(i)}(x), \dots, \psi_{n_i}^{(i)}(x))$, to V_h , the fine mesh with the basis $(\psi_1(x), \dots, \psi_n(x))$. Unlike the traditional AS method, the matrix R_i^T does not consist of just 0 and 1 entries. For the columns associated with the nodal points outside Ω_i , there could be multiple nonzero entries belong to $(0, 1)$. However, for other columns (the majority), there is only one nonzero entry (1); and this entry corresponds to a nodal point inside Ω_i .

Now we introduce the local stiffness matrix $A_i \in \mathbb{R}^{n_i \times n_i}$ associated with the bilinear form $a(\cdot, \cdot)$ restricted on the subspace V_i , as follows

$$A_i(k, j) = a(\psi_j^{(i)}, \psi_k^{(i)}) = a \left(\sum_{l_1=1}^n R_i^T(l_1, j) \psi_{l_1}, \sum_{l_2=1}^n R_i^T(l_2, k) \psi_{l_2} \right) = \sum_{l_2, l_1=1}^n R_i(k, l_2) A_{l_2, l_1} R_i^T(l_1, j).$$

This implies that

$$A_i = R_i A R_i^T. \quad (5)$$

Clearly, A_i is symmetric and positive definite.

Next we define $P_i = R_i^T A_i^{-1} R_i A$. Since $P_i A = A P_i$ and $P_i^2 = P_i$, we see that P_i is an A -orthogonal projection onto the range of R_i^T . Since R_i^T represent the basis functions of V_i , cf. (4), P_i corresponds to a projection operator which is onto V_i .

Now we define our symmetric positive definite preconditioner

$$P^{-1} = \sum_{i=1}^N R_i^T A_i^{-1} R_i.$$

Then the preconditioned system can be written as

$$P^{-1} A = \sum_{i=1}^N P_i = \sum_{i=1}^N R_i^T A_i^{-1} R_i A.$$

Remark 1. Although the formulation of P_i and P^{-1} largely resemble that of the traditional AS methods, we emphasise that there is a fundamental difference in the subspaces V_i in use. In the current approach, V_i are the finite element spaces associated with local meshes (\mathcal{T}_i) of the whole domain Ω ; while in traditional AS methods, V_i are finite element spaces associated with the fine meshes ($\mathcal{T}_h|_{\Omega_i^\dagger}$) of subdomains (Ω_i^\dagger) slightly larger than Ω_i (see [34, p. 59]). In addition, in the current approach, the coarse space V_0 is contained in each V_i and there is no explicit coarse component in P^{-1} . For more information about traditional AS methods, we refer the reader to [34, 30] and references therein.

Remark 2. An advantage of P^{-1} over traditional AS preconditioners is the local matrix A_i can be assembled locally on each processor. Consequently, the global matrix A does not need to be assembled (to use in (5)). This is valuable in real-life applications where the system size is large.

Remark 3. Each restriction matrix R_i has more rows than its counterpart in the traditional two-level AS preconditioner \tilde{P}_{AS}^{-1} associated with the partitioning $\{\Omega_i^\dagger\}_{i=1}^N$ and the coarse space V_0 . In addition, the rows of R_i associated with the coarse degrees of freedom (dofs) outside Ω_i^\dagger and the corresponding rows of \tilde{R}_0 in \tilde{P}_{AS}^{-1} are exactly the same. This suggests an efficient way of computing R_i as follows. Each processor independently computes rows of R_i associated with dofs in $\tilde{\Omega}_i^\dagger$ and part of \tilde{R}_0 associated with its subdomain. Then the complete R_i can be obtained after an MPI-Alltoall communication that exchanges the information of \tilde{R}_0 . With this implementation, the cost of evaluating R_i , $i = 1, 2, \dots, N$ in P^{-1} is comparable with the cost of computing \tilde{R}_i , $i = 0, 1, \dots, N$ in \tilde{P}_{AS}^{-1} .

The preconditioner P^{-1} can be used to accelerate Krylov methods in solving the systems (3). Since P^{-1} and A are both symmetric positive definite the obvious choice is CG, the conjugate gradient method [23, 35].

In the next section, we will study the convergence of the CG method preconditioned by the proposed preconditioner P^{-1} .

4. Convergence analysis

In the first phase of our analysis, we will formulate Euclidean orthogonal projections Q_i corresponding to P_i and study the spectrum of the preconditioned system $P^{-1} A = \sum_{i=1}^N P_i$ via that of $\sum_{i=1}^N Q_i$.

Let $\{\phi_1(x), \phi_2(x), \dots, \phi_n(x)\}$ be an $a(\cdot, \cdot)$ -orthonormal basis of V_h . Without loss of generality, we can assume that $\{\phi_1(x), \phi_2(x), \dots, \phi_{n_0}(x)\}$ is an $a(\cdot, \cdot)$ -orthonormal basis of V_0 . Denote

$$U = \begin{bmatrix} \phi_1(x_1) & \cdots & \phi_n(x_1) \\ \vdots & \cdots & \vdots \\ \phi_1(x_n) & \cdots & \phi_n(x_n) \end{bmatrix}, \quad U_0 = \begin{bmatrix} \phi_1(x_1) & \cdots & \phi_{n_0}(x_1) \\ \vdots & \cdots & \vdots \\ \phi_1(x_n) & \cdots & \phi_{n_0}(x_n) \end{bmatrix},$$

It follows that $U^T A U = I_n$, $U_0^T A U_0 = I_{n_0}$.

Lemma 1. Let $Q_i = U^T A P_i U = U^{-1} P_i U$. Then Q_i is an Euclidean orthogonal projection and it has block diagonal structure $Q_i = \text{diag}(I_{n_0}, \hat{Q}_i)$, where $\hat{Q}_i \in \mathbb{R}^{(n_i - n_0) \times (n_i - n_0)}$ is also an Euclidean orthogonal projections. In addition,

$$\sigma(P^{-1}A) = \sigma\left(\sum_{i=1}^N Q_i\right) = \{N\} \cup \sigma\left(\sum_{i=1}^N \hat{Q}_i\right). \quad (6)$$

where $\sigma(\cdot)$ denotes the spectrum.

PROOF. Since $Q_i^2 = Q_i$ and $Q_i^T = Q_i$, Q_i is an Euclidean orthogonal projection. In addition, as $V_0 \subset V_i$ and the columns of U_0 and R_i^T represent basis functions of V_0 and V_i respectively, we see that $\text{range}(U_0) \subset \text{range}(R_i^T) = \text{range}(P_i)$. Therefore, we can write $P_i U = P_i [U_0 \ *] = [P_i U_0 \ *] = [U_0 \ *]$ and

$$Q_i = U^T A P_i U = \begin{bmatrix} U_0^T \\ * \end{bmatrix} A [U_0 \ *] = \begin{bmatrix} U_0^T A U_0 & * \\ * & * \end{bmatrix} = \begin{bmatrix} I_{n_0} & Z_i \\ Z_i^T & \hat{Q}_i \end{bmatrix}.$$

Since $Q_i^2 = Q_i$, it implies that $Z_i Z_i^T = 0$, or $Z_i = 0$. Therefore, $Q_i = \text{diag}(I_{n_0}, \hat{Q}_i)$. As Q_i is an orthogonal Euclidean projection, \hat{Q}_i is also an orthogonal Euclidean projection. The first part of (6) follows from the fact that

$$\sum_{i=1}^N Q_i = U^{-1} \left(\sum_{i=1}^N P_i \right) U = U^{-1} (P^{-1}A) U.$$

The second part of (6) is a consequence of $\sum_{i=1}^N Q_i = \text{diag}(N I_{n_0}, \sum_{i=1}^N \hat{Q}_i)$.

Lemma 2. Let $\hat{\lambda}_{\min}$ and $\hat{\lambda}_{\max}$ be the smallest and largest eigenvalues of $\sum_{i=1}^N \hat{Q}_i$ respectively. Then

$$\sigma_A(P^{-1}A) \subset [\hat{\lambda}_{\min}, \hat{\lambda}_{\max}] \cup \{N\}, \text{ where } 0 < \hat{\lambda}_{\min} \leq \hat{\lambda}_{\max} \leq N. \quad (7)$$

PROOF. Since \hat{Q}_i is a projection, $\sigma(\hat{Q}_i) = \{0, 1\}$ and $\sigma(\sum_{i=1}^N \hat{Q}_i) \subset [0, N]$. Because P^{-1} and A are both positive definite, $\hat{\lambda}_{\min} > 0$. Then (7) follows from (6).

Remark 4. The result presented in (7) indicates that $\hat{\lambda}_{\min}$ and $\hat{\lambda}_{\max}$ are actually the smallest and the second largest eigenvalues of the preconditioned system $P^{-1}A$. The eigenvalue $\hat{\lambda}_{\max}$ equals N if and only if the local subspace V_i has common subset strictly larger than V_0 . This only happens when N is small and local meshes are structured. In general, $N > \hat{\lambda}_{\max}$ and N is an isolated eigenvalue of $P^{-1}A$.

In the next step, we will take advantage of the special spectrum decomposition in (6) to study the convergence of the CG method applied to the preconditioned system $P^{-1}A$. But first, we quote from [1] the following result

$$\frac{\|e_k\|_A}{\|e_0\|_A} = \inf_{q \in P_k} \frac{\|q(P^{-1}A)e_0\|_A}{\|e_0\|_A} \leq \inf_{q \in P_k} \max_{\lambda \in \sigma(P^{-1}A)} |q(\lambda)|. \quad (8)$$

Here $e_k = u_k - u$ is the exact error at the step n of the CG method, $\sigma(P^{-1}A)$ denotes the spectrum of $P^{-1}A$, and P_k is the set of polynomials q of degree k or less, with $q(0) = 1$. More details about the CG method can be found in [35, 23] and the references therein.

Theorem 3. The error of the CG method applied to equation (3) when it is left-preconditioned by P^{-1} satisfies

$$\frac{\|e_k\|_A}{\|e_0\|_A} \leq \frac{2(N - \hat{\lambda}_{\min})}{N} \left(\frac{\sqrt{\hat{k}} - 1}{\sqrt{\hat{k}} + 1} \right)^{k-1} < 2 \left(\frac{\sqrt{\hat{k}} - 1}{\sqrt{\hat{k}} + 1} \right)^{k-1}, \quad (9)$$

where $\hat{k} = \hat{\lambda}_{\max}/\hat{\lambda}_{\min}$ is called the **effective condition number** of $P^{-1}A$.

PROOF. By (8), it is sufficient to find a polynomial $q(x) \in P_k$ whose maximum value for $x \in [\hat{\lambda}_{\min}, \hat{\lambda}_{\max}]$ is the second quantity in (9). Consider the polynomial

$$q(x) = \frac{T_{k-1}(\gamma - \frac{2x}{\hat{\lambda}_{\max} - \hat{\lambda}_{\min}})(N-x)}{NT_{k-1}(\gamma)}, \quad (10)$$

where $\gamma = (\hat{\lambda}_{\max} + \hat{\lambda}_{\min})/(\hat{\lambda}_{\max} - \hat{\lambda}_{\min}) > 1$ and $T_{k-1}(x)$ is the Chebyshev polynomial of degree $k-1$. More information about Chebyshev polynomials can be found in [27]. Clearly, q has degree k and $q(0) = 1$.

For $x \in [\hat{\lambda}_{\min}, \hat{\lambda}_{\max}]$, the quantity $\gamma - \frac{2x}{\hat{\lambda}_{\max} - \hat{\lambda}_{\min}}$ belongs to $[-1, 1]$ and $|N-x| \leq N - \hat{\lambda}_{\min}$. It follows that

$$\left| T_{k-1} \left(\gamma - \frac{2x}{\hat{\lambda}_{\max} - \hat{\lambda}_{\min}} \right) (N-x) \right| \leq N - \hat{\lambda}_{\min}. \quad (11)$$

We use the standard estimate for $T_{k-1}(x)$:

$$T_{k-1}(\gamma) = \frac{1}{2} \left[\left(\frac{\sqrt{k}+1}{\sqrt{k}-1} \right)^{k-1} + \left(\frac{\sqrt{k}+1}{\sqrt{k}-1} \right)^{-(k-1)} \right] \geq \frac{1}{2} \left(\frac{\sqrt{k}+1}{\sqrt{k}-1} \right)^{k-1}. \quad (12)$$

More details can be found in [35, p. 300]. The inequalities (9) then follow immediately from (11) and (12).

We have shown in Theorem 3 that the convergence of the CG method with preconditioner P^{-1} can be bounded by quantities mainly depend on the ratio of $\hat{\lambda}_{\min}$ and $\hat{\lambda}_{\max}$, the second largest and smallest eigenvalues of $P^{-1}A$. In the next step, we present estimates for these eigenvalues.

4.1. Second largest eigenvalue estimate

Our plan to estimate $\hat{\lambda}_{\max}$ is to seek an explicit formula for \hat{Q}_i and compare the largest eigenvalue of $\sum_{i=1}^N \hat{Q}_i$ with that of the related traditional AS method. We begin with some preparation.

Let \hat{V}_i be the subspace of V_i spanned by nodal basis functions associated with nodal points which are in \mathcal{T}_i but are not in \mathcal{T}_H . With a slight abuse of notation we can write

$$\hat{V}_i = \text{span} \{ \psi_j^{(i)}(x), \forall j \text{ s.t } x_j \notin \mathcal{T}_H \}$$

Clearly, $V_i = V_0 \oplus \hat{V}_i$. This is a hierarchical decomposition of V_i into subspace V_0 of coarse basis functions and subspace \hat{V}_i of fine basis functions. We quote from [7] (see also [22]) the following well-known result of the strengthened Cauchy-Schwarz inequality for hierarchical bases.

Lemma 4. *Given the finite element hierarchical decomposition $V_i = V_0 \oplus \hat{V}_i$. Then for all $v_0(x) \in V_0$ and all $\hat{v}_i(x) \in \hat{V}_i$:*

$$|a(v_0, \hat{v}_i)| \leq \gamma \|v_0\|_A \|\hat{v}_i\|_A, \quad i = 1, \dots, N. \quad (13)$$

Here the constant γ , $0 < \gamma < 1$, (the maximum of all the constants associated with local meshes \mathcal{T}_i) depends on the shape regularity quality of the meshes \mathcal{T}_H , \mathcal{T}_i , but is otherwise independent of the mesh sizes h and H .

Now let $m_i = n_i - n_0$ and $\{\omega_1^{(i)}(x), \dots, \omega_{m_i}^{(i)}(x)\}$ be an $a(\cdot, \cdot)$ -orthonormal basis of \hat{V}_i . Denote

$$W_i = \begin{bmatrix} \omega_1^{(i)}(x_1) & \cdots & \omega_{m_i}^{(i)}(x_1) \\ \vdots & \cdots & \vdots \\ \omega_1^{(i)}(x_n) & \cdots & \omega_{m_i}^{(i)}(x_n) \end{bmatrix}.$$

We note that the columns of U_0 and the columns of W_i represent bases of V_0 and \hat{V}_i respectively. Therefore, $\text{range}(P_i) = \text{range}(R_i^T) = \text{range}([U_0 \ W_i])$ since $V_0 \oplus \hat{V}_i = V_i$.

Lemma 5. *Let $U^T A W_i = [X_i^T \ Y_i^T]^T$, where $X_i \in \mathbb{R}^{n_0 \times m_i}$, $Y_i \in \mathbb{R}^{n-n_0 \times m_i}$. Then $\hat{Q}_i = Y_i(Y_i^T Y_i)^{-1} Y_i^T$, for $i = 1, \dots, N$.*

PROOF. Since $Q_i = U^T A P_i U$ and U is non-singular, we have

$$\begin{aligned} \text{range}(Q_i) &= U^T A(\text{range}(P_i)) = U^T A(\text{range}([U_0 \ W_i])) = \text{range}(U^T A[U_0 \ W_i]) \\ &= \text{range}\left(\begin{bmatrix} I & X_i \\ 0 & Y_i \end{bmatrix}\right) = \text{range}\left(\begin{bmatrix} I & 0 \\ 0 & Y_i \end{bmatrix}\right) = \text{range}(E_i), \end{aligned} \quad (14)$$

where $E_i = \text{diag}(I, Y_i)$. So Q_i is an projection onto the range of E_i . In addition,

$$n_i = \dim(V_i) = \text{rank}([U_0 \ W_i]) = \text{rank}\left(\begin{bmatrix} I & X_i \\ 0 & Y_i \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} I & 0 \\ 0 & Y_i \end{bmatrix}\right).$$

Therefore, $\text{rank}(Y_i) = n_i - n_0 = m_i$. In other words, the matrix Y_i has full rank. It follows that the columns of E_i are linearly independent. This together with (14) imply

$$Q_i = E_i(E_i^T E_i)^{-1} E_i^T = \begin{bmatrix} I & 0 \\ 0 & Y_i(Y_i^T Y_i)^{-1} Y_i^T \end{bmatrix}.$$

Then the desired equality follows from the fact that $Q_i = \text{diag}(I_{n_0}, \hat{Q}_i)$.

Lemma 6. For X_i, Y_i defined in Lemma 5, we have

$$(1 - \gamma^2)I \preceq Y_i^T Y_i, \quad (15)$$

where $0 < \gamma < 1$ is the constant introduced in Lemma 4. The notation \preceq denotes the positive semi-definite ordering (cf. [25]). In addition,

$$\sum_{i=1}^N Y_i Y_i^T \preceq N_c I_{n-n_0}. \quad (16)$$

PROOF. Using the definitions of X_i, Y_i and the fact that W_i has A-orthonormal columns, we have

$$X_i^T X_i + Y_i^T Y_i = [X_i^T \ Y_i^T] \begin{bmatrix} X_i \\ Y_i \end{bmatrix} = W_i^T A U U^T A W_i = W_i^T A W_i = I_{m_i}. \quad (17)$$

Therefore, in order to show (15) we will bound $X_i^T X_i$ from above.

For $v_0(x) \in V_0$ and $\hat{v}_i(x) \in \hat{V}_i$, their coordinate vectors are of the following forms

$$v_0 = U \begin{bmatrix} y \\ 0 \end{bmatrix}, \quad \hat{v}_i = [U_0 \ W_i] \begin{bmatrix} 0 \\ z \end{bmatrix}, \quad y \in \mathbb{R}^{n_0}, \quad z \in \mathbb{R}^{m_i}.$$

Now the inequality (13) can be written in the matrix form as follows

$$[y^T \ 0] U^T A [U_0 \ W_i] \begin{bmatrix} 0 \\ z \end{bmatrix} \leq \gamma \left([y^T \ 0] U^T A U \begin{bmatrix} y \\ 0 \end{bmatrix} \right) \left([0 \ z^T] \begin{bmatrix} U_0^T \\ W_i^T \end{bmatrix} A [U_0 \ W_i] \begin{bmatrix} 0 \\ z \end{bmatrix} \right).$$

Equivalently for any $y \in \mathbb{R}^{n_0}$ and $z \in \mathbb{R}^{m_i}$: $[y^T \ 0] \begin{bmatrix} I & X_i \\ 0 & Y_i \end{bmatrix} \begin{bmatrix} 0 \\ z \end{bmatrix} = y^T X_i z \leq \gamma \|y\|_2 \|z\|_2$. This implies that $\|X_i\|_2 \leq \gamma$ and $\|X_i^T X_i\|_2 \leq \gamma^2$. In other words, $X_i^T X_i \preceq \gamma^2 I_{m_i}$. Then (15) follows immediately from (17).

Next we are going to prove (16). Let $V_i^\dagger = V_h|_{\Omega_i^\dagger}$, $i = 1, \dots, N$. We note that V_i^\dagger are the local spaces in the related traditional AS method (see [34, p. 59]). Since all elements in \mathcal{T}_i that are outside of Ω_i^\dagger belong to \mathcal{T}_H , \hat{V}_i is a subset of V_i^\dagger . Consequently, there is an orthonormal basis of V_i^\dagger in the form of $\{\omega_1^{(i)}, \dots, \omega_{m_i}^{(i)}, \omega_{m_i+1}^{(i)}, \dots, \omega_{\tilde{m}_i}^{(i)}\}$. Let $\tilde{W}_i \in \mathbb{R}^{n \times \tilde{m}_i}$ be defined as follows

$$\tilde{W}_i = \begin{bmatrix} \omega_1^{(i)}(x_1) & \cdots & \omega_{\tilde{m}_i}^{(i)}(x_1) \\ \vdots & \cdots & \vdots \\ \omega_1^{(i)}(x_n) & \cdots & \omega_{\tilde{m}_i}^{(i)}(x_n) \end{bmatrix}.$$

Denote $[\widetilde{X}_i^T \ \widetilde{Y}_i^T] = U^T A \widetilde{W}_i$, where $\widetilde{X}_i \in \mathbb{R}^{n_0 \times \widetilde{m}_i}$, $Y_i \in \mathbb{R}^{n-n_0 \times \widetilde{m}_i}$. Then the first m_i columns of \widetilde{Y}_i form Y_i . Assume $Y_i = [y_1^i \ \cdots \ y_{m_i}^i]$ and $\widetilde{Y}_i = [y_1^i \ \cdots \ y_{m_i}^i \ y_{m_i+1}^i \ \cdots \ y_{\widetilde{m}_i}^i]$. For any $z \in \mathbb{R}^{n-n_0}$ we have

$$z^T \left(\sum_{i=1}^N Y_i Y_i^T \right) z = \sum_{i=1}^N \sum_{j=1}^{m_i} (y_j^i)^T z)^2 \leq \sum_{i=1}^N \sum_{j=1}^{\widetilde{m}_i} (y_j^i)^T z)^2 = z^T \left(\sum_{i=1}^N \widetilde{Y}_i \widetilde{Y}_i^T \right) z. \quad (18)$$

Therefore,

$$\sum_{i=1}^N Y_i Y_i^T \leq \sum_{i=1}^N \widetilde{Y}_i \widetilde{Y}_i^T. \quad (19)$$

Now let \widetilde{Q}_i be the Euclidean orthogonal projection corresponding to the Schwarz projection \widetilde{P}_i associated with Ω_i^\dagger in the traditional AS method (see [34, chapter 2]). Similar to (14), we have $\text{range}(\widetilde{Q}_i) = \text{range}(U^T A \widetilde{W}_i)$. In addition, $F_i = U^T A \widetilde{W}_i = [\widetilde{X}_i^T \ \widetilde{Y}_i^T]$ has orthonormal columns. Thus the projection \widetilde{Q}_i can be written as

$$\widetilde{Q}_i = F_i F_i^T = \begin{bmatrix} \widetilde{X}_i \widetilde{X}_i^T & \widetilde{X}_i \widetilde{Y}_i^T \\ \widetilde{Y}_i \widetilde{X}_i^T & \widetilde{Y}_i \widetilde{Y}_i^T \end{bmatrix}.$$

Therefore, for any $z \in \mathbb{R}^{n-n_0}$

$$z^T \sum_{i=1}^N \widetilde{Y}_i \widetilde{Y}_i^T z = [0 \ z^T] \sum_{i=1}^N \widetilde{Q}_i \begin{bmatrix} 0 \\ z \end{bmatrix} \leq \rho \left(\sum_{i=1}^N \widetilde{Q}_i \right) z^T z = \rho \left(\sum_{i=1}^N \widetilde{P}_i \right) z^T z,$$

where ρ denote the spectral radius. On the other hand, according to [21, Theorem 4.1], $\rho(\sum_{i=1}^N \widetilde{P}_i) \leq N_c$. Consequently,

$$\widetilde{Y}_i \widetilde{Y}_i^T \leq N_c I_{n-n_0}. \quad (20)$$

The ordering (16) then follows from (19) and (20).

We now present one of our main results, the estimate for the second largest eigenvalue.

Theorem 7. *The second largest eigenvalue of the preconditioned system $P^{-1}A$ is bounded as follows*

$$\hat{\lambda}_{\max} \leq \frac{N_c}{(1-\gamma^2)}. \quad (21)$$

PROOF. From (5), we have $\hat{\lambda}_{\max} = \rho(\sum_{i=1}^N \hat{Q}_i) = \rho(\sum_{i=1}^N Y_i (Y_i^T Y_i)^{-1} Y_i^T)$. On the other hand, it follows from (16) and (15) that

$$\sum_{i=1}^N Y_i (Y_i^T Y_i)^{-1} Y_i^T \leq \frac{1}{(1-\gamma^2)} \sum_{i=1}^N Y_i Y_i^T \leq \frac{N_c}{(1-\gamma^2)} I_{n-n_0}.$$

Then the equality (21) follows immediately.

4.2. Smallest eigenvalue estimate

Our estimate of $\hat{\lambda}_{\min}$ follows the standard approach where a stable decomposition is constructed [37, 21, 34]. However, as the local meshes \mathcal{T}_i are meshes of the whole domain and they are very different from one another and from the global fine mesh \mathcal{T}_h outside of their associated subdomains, the stable decomposition in [21, 34] is no longer valid. In order to adapt to the situation, we build our stable decomposition inductively on the colouring defined in section 2. In our construction, the partition of unity is replaced by a set of cut-off functions, and the point-wise interpolation is replaced by a special interpolation inspired by [31].

Cut-off functions. Denote C_k the set of indices of subdomains coloured by colour c_k , $1 \leq c_k \leq N^c$. Then for each subdomain Ω_i , $i \in C_k$, we define the cut-off function $\theta_i^{(c_k)}(x)$ as follows:

$$\theta_i^{(c_k)}(x) = \begin{cases} 1 & \text{if } x \in \bar{\Omega}_i \\ 0 & \text{if } x \notin \bar{\Omega}_i^\dagger \\ \frac{\text{dist}(x, \partial\Omega_i^\dagger \setminus \partial\Omega)}{\text{dist}(x, \partial\Omega_i^\dagger \setminus \partial\Omega) + \text{dist}(x, \partial\Omega_i \setminus \partial\Omega)} & \text{if } x \in \Omega_i^\dagger \setminus \Omega_i, \end{cases} \quad (22)$$

Clearly, $\theta_i^{(c_k)}$ is well-defined, continuous on $\bar{\Omega}$ and satisfies

$$0 \leq \theta_i^{(c_k)}(x) \leq 1, \quad \text{for all } x \in \bar{\Omega}. \quad (23)$$

In addition,

$$\text{supp}(\theta_i^{(c_k)}) \subset \bar{\Omega}_i^\dagger, \quad \text{supp}(\theta_i^{(c_k)}) \cap \text{supp}(\theta_j^{(c_k)}) = \emptyset, \quad i, j \in C_k, \quad i \neq j. \quad (24)$$

Since the width of $\Omega_i^\dagger \setminus \Omega_i$ is of size H , according to [34, Lemma 3.4], there exists constant C^θ does not depend on i and H such that

$$\|\nabla \theta_i^{(c_k)}\|_\infty \leq C^\theta / H. \quad (25)$$

In the next step, we present the framework to construct the modified Lagrange type interpolation operator introduced by Scott and Zhang in [31]. Some stability properties for this type of interpolation will also be provided for later use.

Modified Lagrange interpolations. Let \mathcal{T}° be a finite element mesh of Ω with its set of nodal points $\mathcal{N}^\circ = \{x_j^\circ\}_{j=1}^{n^\circ}$. Denote V° the finite element space associated with \mathcal{T}° and let $\{\psi_j^\circ\}_{j=1}^{n^\circ}$ be the set of linear nodal basis functions of V° corresponding to \mathcal{N}° . For any node x_j° , we fix an edge e_j° in \mathcal{T}° that has x_j° as one of its vertex. Let $\{x_{j,k}^\circ\}_{k=1}^2$ be the two nodal points in \mathcal{N}° associated with e_j° . Without loss of generality, we choose $x_{j,1}^\circ = x_j^\circ$. For the nodal basis $\{\psi_{j,k}^\circ\}_{k=1}^2$ associated with $\{x_{j,k}^\circ\}_{k=1}^2$, we have an $L^2(e_j^\circ)$ -dual basis $\{\eta_{j,k}^\circ\}_{k=1}^2$ defined by $\int_{e_j^\circ} \eta_{j,k}^\circ \psi_{j,l}^\circ = \delta_{kl}$, $k, l = 1, 2$, where $\delta_{k,l}$ is the Kronecker delta. For simplicity, we let $\eta_j^\circ \equiv \eta_{j,1}^\circ$, for $x_j^\circ \in \mathcal{N}_i$. Then, we have

$$\int_{e_j^\circ} \eta_j^\circ \psi_k^\circ = \delta_{jk}, \quad k, j = 1, 2, \dots, n^\circ. \quad (26)$$

Now we can define the interpolation operator,

$$I^\circ = I_{\mathcal{T}^\circ}^{\{e_j^\circ\}} : H^1(\Omega) \rightarrow V^\circ, \\ I^\circ u(x) = \sum_{j=1}^{n_i} \psi_j^\circ(x) \int_{e_j^\circ} \eta_j^\circ(\xi) u(\xi) d\xi. \quad (27)$$

Here, the notation $I_{\mathcal{T}^\circ}^{\{e_j^\circ\}}$ is used to emphasise that the interpolation operator depends on the mesh \mathcal{T}° and the choice of edges $\{e_j^\circ\}_{j=1}^{n^\circ}$. However, for simplicity I° is used in other places.

The following Lemma is useful when we want to consider $I^\circ u$ on a subset of Ω .

Lemma 8. *Let u be a function in $H^1(\Omega)$ and Ω^s be a subset of Ω . Assume that Ω^s is also an union of elements in \mathcal{T}° . Then following statement holds*

$$I^\circ u(x) = \sum_{j, x_j^\circ \in \Omega^s} \psi_j^\circ(x) \int_{e_j^\circ} \eta_j^\circ(\xi) u(\xi) d\xi, \quad \text{for all } x \in \bar{\Omega}^s.$$

PROOF. The proof is obvious as the basis functions $\psi_j^\circ(x)$ associated with $x_j^\circ \notin \bar{\Omega}^s$ vanish in $\bar{\Omega}^s$.

Let $\{x_j^{(i)}\}_{j=1}^{n_i}$ be the set of nodal points of the finite element mesh \mathcal{T}_i , $0 \leq i \leq N$. For each mesh \mathcal{T}_i , $0 \leq i \leq N$ we will choose a set of edges $\{e_j^{(i)}\}_{j=1}^{n_i}$ in \mathcal{T}_i corresponding to $\{x_j^{(i)}\}_{j=1}^{n_i}$ that satisfies the following conditions:

- (i) $e_j^{(i)}$ contains $x_j^{(i)}$
- (ii) $e_j^{(i)} \in \partial\Omega$, if $x_j^{(i)} \in \partial\Omega$
- (iii) $e_j^{(i)} \in \partial\Omega_i \setminus \partial\Omega$, if $x_j^{(i)} \in \partial\Omega_i \setminus \partial\Omega$, $i \neq 0$
- (iv) $e_j^{(i)} \in \partial\Omega_k$, if $x_j^{(i)} \notin \partial\Omega \cup \partial\Omega_i$ is shared by two or more subdomains in the partition $\{\Omega_l\}_{l=1}^N$. Here Ω_k is the subdomain with smallest colour that contains $x_j^{(i)}$.

For each mesh \mathcal{T}_i , we fix a choice of edges $\{e_j^{(i)}\}_{j=1}^{n_i}$ satisfying the four conditions above. Then we let

$$I_i^{h,H} = I_{\mathcal{T}_i}^{\{e_j^{(i)}\}} : H^1(\Omega) \rightarrow V_i, \quad 1 \leq i \leq N$$

$$I^H = I_{\mathcal{T}_0}^{\{e_j^{(0)}\}} : H^1(\Omega) \rightarrow V_0,$$

be the modified Lagrange interpolation operators associate with \mathcal{T}_i and $\{e_j^{(i)}\}_{j=1}^{n_i}$, and with \mathcal{T}_0 and $\{e_j^{(0)}\}_{j=1}^{n_0}$ respectively. According to [31], there exist a constant C^I depend only on the shape regularity of the associated meshes such that

$$\|I_i^{h,H} u\|_{H^1(K)} \leq C^I |u|_{H^1(\omega_K)}, \quad K, \omega_K \in \mathcal{T}_i, \quad (28)$$

$$\|u - I^H u\|_{L^2(K)} \leq C^I H |u|_{H^1(\omega_K)}, \quad K, \omega_K \in \mathcal{T}_H, \quad (29)$$

$$\|I^H u\|_{H^1(K)} \leq C^I |u|_{H^1(\omega_K)}, \quad K, \omega_K \in \mathcal{T}_H. \quad (30)$$

where $\omega_K = \text{interior}\left(\bigcup\{\bar{K}_j \mid \bar{K}_j \cap \bar{K} \neq \emptyset, K_i \in \mathcal{T}^e\}\right)$.

Lemma 9. *The interpolation operator $I_i^{h,H}$ preserves fine functions in the regions where the mesh \mathcal{T}_i is fine. In other words,*

$$I_i^{h,H} u|_{\bar{\Omega}_i} = u|_{\bar{\Omega}_i},$$

for any function $u(x)$ satisfies $u(x)|_{\bar{\Omega}_i} \in V_h|_{\bar{\Omega}_i}$.

PROOF. Let $x_j^{(i)}$ be a nodal point of \mathcal{T}_i , $x_j^{(i)} \in \bar{\Omega}_i$. Since $\mathcal{T}_i|_{\Omega_i} \equiv \mathcal{T}_h|_{\Omega_i}$, this nodal point also presents in \mathcal{T}_h . In addition, the two nodal basis functions associated with $x_j^{(i)}$ in V_i and V_h are identical on $\bar{\Omega}_i$, namely

$$\psi_j^{(i)}|_{\Omega_i} = \psi_j|_{\Omega_i}. \quad (31)$$

On the other hand, by (iii) the chosen edge $e_j^{(i)} \in \mathcal{T}_i$ for the nodal point $x_j^{(i)}$ should also be an edge in \mathcal{T}_h if $x_j^{(i)} \in \bar{\Omega}_i$. Therefore, by (26) we have

$$\int_{e_j^{(i)}} \eta_j(\xi) u(\xi) d\xi = u(x_j), \quad \text{for all } x_j^{(i)} \in \bar{\Omega}_i. \quad (32)$$

Using (27), Lemma 8, (32) and (31), we have

$$I_i^{h,H} u(x) = \sum_{j=1}^{n_i} \psi_j^{(i)}(x) \int_{e_j^{(i)}} \eta_j^{(i)}(\xi) u(\xi) d\xi = \sum_{j, x_j \in \bar{\Omega}_i} \psi_j^{(i)}(x) \int_{e_j^{(i)}} \eta_j^{(i)}(\xi) u(\xi) d\xi = \sum_{j, x_j \in \bar{\Omega}_i} \psi_j^{(i)}(x) u(x_j) = u(x).$$

We are now in a position to estimate the smallest eigenvalue of the preconditioned system $P^{-1}A$. The idea is to construct local functions colour by colour. The proposed interpolations will ensure that residual functions vanish on all considered subdomains, and stay zero there in later induction steps. The following Lemma lays the foundation for our construction of local functions in a stable decomposition.

Lemma 10. Assume $u(x) \in V_h$. Let $u^{(0)}(x) := u(x)$. Then our inductive construction of residual functions $u^{(k)}(x)$ is as follows

$$w^{(k)} = I^H u^{(k-1)}, \quad (w^{(k)} \in V_H) \quad (33)$$

$$v^{(k)} = u^{(k-1)} - w^{(k)}, \quad (v^{(k)} \in V_h) \quad (34)$$

$$v_i^{(k)} = I_i^{h,H} \theta_i^{(c_k)} v^{(k)}, \quad (v_i^{(k)} \in V_i). \quad (35)$$

$$u^{(k)} = v^{(k)} - \sum_{i \in C_k} v_i^{(k)} = v^{(k)} - \sum_{i \in C_k} I_i^{h,H} \theta_i^{(c_k)} v^{(k)}, \quad (u^{(k)} \in V_h) \quad (36)$$

where $k = 1, 2, \dots, N_c$. Then the following equalities hold

$$u^{(k)}|_{\bar{\Omega}_i} \equiv 0, \quad \text{for all } i \in C_k, k_i \leq k, \quad (37)$$

$$u = \sum_{k=0}^{N_c-1} w^{(k)} + \sum_{k=1}^{N_c} \sum_{i \in C_k} v_i^{(k)}, \quad (38)$$

$$\left| \sum_{i \in C_k} v_i^{(k)} \right|_{H^1(\Omega)}^2 = \sum_{i \in C_k} |v_i^{(k)}|_{H^1(\Omega)}^2. \quad (39)$$

PROOF. Substituting $k = 1$ into (36) gives $u^{(1)} = v^{(1)} - \sum_{i \in C_1} I_i^{h,H} \theta_i^{(c_1)} v^{(1)}$. For $i, j \in C_1$, $i \neq j$, according to (22), $\theta_i^{(c_1)} = 1$ on $\bar{\Omega}_i$, and $\theta_j^{(c_1)} = 0$ on $\bar{\Omega}_i$. Therefore, $I_i^{h,H} \theta_i^{(c_1)} v^{(1)} = I_i^{h,H} v^{(1)} = v^{(1)}$ on $\bar{\Omega}_i$, $i \in C_1$ as a consequence of Lemma 9. In addition, $I_j^{h,H} \theta_j^{(c_1)} v^{(1)} \equiv I_j^{h,H} 0 = 0$ on $\bar{\Omega}_i$. Combining these together, we have

$$u^{(1)}|_{\bar{\Omega}_i} \equiv 0, \quad \text{for all } i \in C_1. \quad (40)$$

For any $x \in \bar{\Omega}_i$, $i \in C_1$ from (33) and Lemma 8, it follows that

$$w^{(2)}(x) = I^H u^{(1)}(x) = \sum_{j, x_j^{(0)} \in \bar{\Omega}_i} \psi_j^{(0)}(x) \int_{e_j^{(0)}} \eta_j^{(0)}(\xi) u^{(1)}(\xi) d\xi \quad (41)$$

By condition (iv), $e_j^{(0)} \in \bar{\Omega}_i$ for all $x_j^{(0)} \in \bar{\Omega}_i$, $i \in C_1$. This together with (40) imply

$$w^{(2)}|_{\bar{\Omega}_i} \equiv 0, \quad \text{for all } i \in C_1. \quad (42)$$

Then from (34), (40) and (42), it follows that

$$v^{(2)}|_{\bar{\Omega}_i} \equiv 0, \quad \text{for all } i \in C_1. \quad (43)$$

Substituting $k = 2$ into (36), we obtain $u^{(2)} = v^{(2)} - \sum_{i \in C_2} I_i^{h,H} \theta_i^{(c_2)} v^{(2)}$. Similarly, we have

$$u^{(2)}|_{\bar{\Omega}_i} \equiv 0, \quad \text{for all } i \in C_2. \quad (44)$$

Now assume $l \in C_1$. For any $x \in \bar{\Omega}_l$, $i \in C_2$ according to Lemma 8,

$$I_i^{h,H} \theta_i^{(c_2)} v^{(2)}(x) = \sum_{j, x_j^{(0)} \in \bar{\Omega}_i} \psi_j^{(i)}(x) \int_{e_j^{(i)}} \eta_j^{(i)}(\xi) (\theta_i^{(c_2)} v^{(2)})(\xi) d\xi. \quad (45)$$

On the right hand side of (45), if $x_j^{(i)} \in \bar{\Omega}_l \setminus \partial\Omega_i$ then by condition (iv), $e_j^{(i)} \in \partial\Omega_l \subset \bar{\Omega}_l$. This together with (43) imply $\int_{e_j^{(i)}} \eta_j^{(i)}(\xi) (\theta_i^{(c_2)} v^{(2)})(\xi) d\xi = 0$. If $x_j^{(i)} \in \bar{\Omega}_l \cap \partial\Omega_i$ then by condition (iii), $e_j^{(i)} \in \partial\Omega_i$. From (26), (22), (43) and the fact

that $x_j^{(i)} \in \bar{\Omega}_l$, we have $\int_{\rho^{(i)}} \eta_j^{(i)}(\xi) (\theta_i^{(c_2)} v^{(2)})(\xi) d\xi = \theta_i^{(c_2)} v^{(2)}(x_j^{(i)}) = v^{(2)}(x_j^{(i)}) = 0$. In summary, $I_i^{h,H} \theta_i^{(c_2)} v^{(2)} = 0$ on $\bar{\Omega}_l$, for all $l \in C_1$, $i \in C_2$. This together with (43) imply $u^{(2)}|_{\bar{\Omega}_i} \equiv 0$, for all $l \in C_1$. From (44), it follows that

$$u^{(2)}|_{\bar{\Omega}_i} \equiv 0, \quad \text{for all } i \in C_1 \cup C_2.$$

Continuing this process for $k = 3, \dots, N^c$, we obtain (37).

Since $\{\bar{\Omega}_i\}_{i=1}^N$ covers Ω , (37) implies $u^{(N^c)}|_{\Omega} \equiv 0$. Tracing backward, we have

$$\begin{aligned} 0 = u^{(N^c)} &= u^{(N^c-1)} - w^{(N^c)} - \sum_{i \in C_{N^c}} v_i^{(N^c)} = u^{(N^c-2)} - w^{(N^c-1)} - w^{(N^c)} - \sum_{i \in C_{N^c-1}} v_i^{(N^c-1)} - \sum_{i \in C_{N^c}} v_i^{(N^c)} \\ &= u^{(0)} - \sum_{k=1}^{N^c} w^{(k)} - \sum_{k=1}^{N^c} \sum_{i \in C_k} v_i^{(k)}. \end{aligned}$$

This implies (38) because $u^{(0)}(x) = u(x)$.

Since $\theta_i^{(c_k)}$ has support on $\bar{\Omega}_i^\dagger$, the functions $\theta_i^{(c_k)} v^{(k)}$ and consequently $v_i^{(k)} = I_i^{h,H} \theta_i^{(c_k)} v^{(k)}$ also have support on $\bar{\Omega}_i^\dagger$. Therefore, $v_i^{(k)}$ have disjoint supports, and (39) follows immediately.

Now we are ready to state the main result of this subsection.

Theorem 11. *For any $u(x) \in V_h$ there exists a decomposition*

$$u = \sum_{i=1}^N u_i, \quad u_i(x) \in V_i, \quad 1 \leq i \leq N,$$

that satisfies

$$\sum_{i=1}^N a(u_i, u_i) \leq C_m a(u, u),$$

where C_m is a constant independent of H , h and N but not N^c . In addition, the smallest eigenvalue of the preconditioned system $P^{-1}A$ can be bounded from below as follows

$$\hat{\lambda}_{\min} \geq C_m^{-1}.$$

PROOF. In this proof, for simplicity, we use $x \lesssim y$ to denote $x \leq C y$, where the constant C might depend on the interpolation constant, the constant in bounding the gradients of cut-off functions and the number of colours in the colouring (C^l , C^θ and N^c respectively) but does not depend on the mesh sizes (h , H) and the number of subdomains in the partition (N).

Based on (38) in Lemma 10, we define

$$u = \sum_{i=1}^N u_i, \quad \text{where } u_i = \begin{cases} w^{(k_i)} + v_i^{(k_i)}, & \text{if } i = \min(C_{k_i}) \\ v_i^{(k_i)}, & \text{otherwise} \end{cases}. \quad (46)$$

We will show that this is a stable decomposition.

First, from the definition of $w^{(k)}$ in (33) and the stability properties of I^H in (30), it follows that $|w^{(k)}|_{H^1(K)} \leq C^l |u^{(k-1)}|_{H^1(\omega_K)}$, for K and $\omega_K \in \mathcal{T}_0$. Squaring and summing over all $K \in \mathcal{T}_0$, we have

$$|w^{(k)}|_{H^1(\Omega)}^2 \lesssim |u^{(k-1)}|_{H^1(\Omega)}^2. \quad (47)$$

Then it follows from (34), Young's inequality, and (47) that

$$|v^{(k)}|_{H^1(\Omega)}^2 \leq 2 \left(|u^{(k-1)}|_{H^1(\Omega)}^2 + |w^{(k)}|_{H^1(\Omega)}^2 \right) \lesssim |u^{(k-1)}|_{H^1(\Omega)}^2. \quad (48)$$

On the other hand, from (34), (33) and (29), we have

$$\|v^{(k)}\|_{L^2(K)} = \|u^{(k-1)} - I^H u^{(k-1)}\|_{L^2(K)} \leq C^I H |u^{(k-1)}|_{H^1(\omega_K)}, \quad K, \omega_K \in \mathcal{T}_0.$$

Squaring and summing over all $K \in \mathcal{T}_0$, we obtain

$$\|v^{(k)}\|_{L^2(\Omega)}^2 \lesssim H^2 |u^{(k-1)}|_{H^1(\Omega)}^2. \quad (49)$$

For $i \in C_k$, from (35) and (28) it follows that

$$|v_i^{(k)}|_{H^1(K)} = |I_i^{h,H}(\theta_i^{(c_k)} v^{(k)})|_{H^1(K)} \leq C^I |\theta_i^{(c_k)} v^{(k)}|_{H^1(K)}, \quad K, K \in \mathcal{T}_i. \quad (50)$$

Squaring and summing over all $K \in \mathcal{T}_i$, we find that

$$|v_i^{(k)}|_{H^1(\Omega)}^2 \lesssim |\theta_i^{(c_k)} v^{(k)}|_{H^1(\Omega)}^2. \quad (51)$$

For $1 \leq k \leq N^c$, using (39), (51) and Young's inequality, we have

$$\begin{aligned} \left| \sum_{i \in C_k} v_i^{(k)} \right|_{H^1(\Omega)}^2 &= \sum_{i \in C_k} |v_i^{(k)}|_{H^1(\Omega)}^2 \lesssim \sum_{i \in C_k} |\theta_i^{(c_k)} v^{(k)}|_{H^1(\Omega)}^2 = \left| v^{(k)} \sum_{i \in C_k} \theta_i^{(c_k)} \right|_{H^1(\Omega)}^2 \\ &\leq 2 \int_{\Omega} \left(\nabla v^{(k)} \left(\sum_{i \in C_k} \theta_i^{(c_k)} \right) \right)^2 dx + 2 \int_{\Omega} \left(v^{(k)} \nabla \left(\sum_{i \in C_k} \theta_i^{(c_k)} \right) \right)^2 dx. \end{aligned} \quad (52)$$

The first term on the right hand side of (52) can be estimated using the fact that $\theta_i^{(c_k)}$ have disjoint supports (see (24)) and uniformly bounded by 1 (see (22)):

$$\int_{\Omega} \left(\nabla v^{(k)} \left(\sum_{i \in C_k} \theta_i^{(c_k)} \right) \right)^2 dx \leq \int_{\Omega} (\nabla v^{(k)})^2 dx = |v^{(k)}|_{H^1(\Omega)}^2. \quad (53)$$

The second term of (52) is estimated using the fact that $\nabla \theta_i^{(c_k)}$ are uniformly bounded (see (25)) and have disjoint supports (implied from (24)), and (49). :

$$\int_{\Omega} \left(v^{(k)} \nabla \left(\sum_{i \in C_k} \theta_i^{(c_k)} \right) \right)^2 dx \leq \left| \nabla \left(\sum_{i \in C_k} \theta_i^{(c_k)} \right) \right|_{\infty}^2 \int_{\Omega} (v^{(k)})^2 dx \lesssim H^{-2} \|v^{(k)}\|_{L^2(\Omega)}^2 \lesssim |v^{(k)}|_{H^1(\Omega)}^2. \quad (54)$$

Combining (52) and (53) and (54), we have

$$\left| \sum_{i \in C_k} v_i^{(k)} \right|_{H^1(\Omega)}^2 \lesssim |v^{(k)}|_{H^1(\Omega)}^2. \quad (55)$$

Using (36), Young's inequality, (55) and (48), it yields

$$|u^{(k)}|_{H^1(\Omega)}^2 = \left| v^{(k)} - \sum_{i \in C_k} v_i^{(k)} \right|_{H^1(\Omega)}^2 \leq 2 \left(|v^{(k)}|_{H^1(\Omega)}^2 + \left| \sum_{i \in C_k} v_i^{(k)} \right|_{H^1(\Omega)}^2 \right) \lesssim |v^{(k)}|_{H^1(\Omega)}^2 \lesssim |u^{(k-1)}|_{H^1(\Omega)}^2.$$

Consequently,

$$|u^{(k)}|_{H^1(\Omega)}^2 \lesssim |u^{(0)}|_{H^1(\Omega)}^2 = |u|_{H^1(\Omega)}^2. \quad (56)$$

Using (46), Young's inequality, (47), (39), (55) and (48), we have

$$\sum_{i=1}^N a(u_i, u_i) = \sum_{k=1}^{N^c} \sum_{i \in C_k} |u_i^{(k)}|_{H^1(\Omega)}^2 \leq 2 \sum_{k=1}^{N^c} \left(|w^{(k)}|_{H^1(\Omega)}^2 + \sum_{i \in C_k} |v_i^{(k)}|_{H^1(\Omega)}^2 \right) \lesssim |u^{(k-1)}|_{H^1(\Omega)}^2. \quad (57)$$

Then it follows from (57) and (56) that there exist C_m independent of h , H and N such that

$$\sum_{i=1}^N a(u_i, u_i) \leq C_m |u|_{H^1(\Omega)}^2 = C_m a(u, u). \quad (58)$$

The lower bound for the smallest eigenvalue $\hat{\lambda}_{\min}$ is followed immediately by [37] (see also [34, Chapter 2]).

Combining Theorems 7 and 11, we find a bound for the effective condition number of the proposed method.

Theorem 12. *In case exact solvers are employed on all subspaces, the effective condition number of the preconditioned system $P^{-1}A$ satisfies*

$$\hat{\kappa} = \frac{\hat{\lambda}_{\max}}{\hat{\lambda}_{\min}} \leq \frac{C_m N^c}{(1 - \gamma^2)}.$$

5. Numerical Experiments

In this section, we present two numerical experiments to support the theory formulated in section 4. A vectorised Matlab code which allows hanging nodes is developed for this study. The finite element codes iFEM [20] and PLTMG [9] are also used extensively in early testing.

5.1. L-shaped Domain

In this experiment, we consider the problem described in (1), where Ω is the L-shaped domain obtained from the unit square by removing the lower right quarter (see Figure 3).

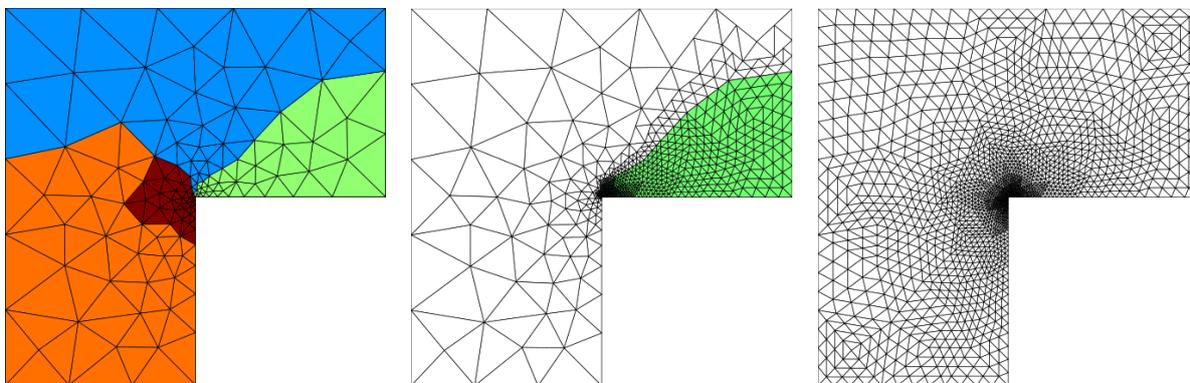


Figure 3: L-shaped domain: The coarse mesh with partition (left), a local mesh (middle) and the global mesh (right) when $L = 0$, $N = 4$ and $l = 2$

We will study the changes of $\hat{\lambda}_{\max}$ and $\hat{\lambda}_{\min}$ when h , H and N vary. According to Theorem 7 and Theorem 11, we expect $\hat{\lambda}_{\max}$ and $\hat{\lambda}_{\min}$ do not depend on h and H , but weakly depend on N .

By serial adaptivity, we create a conforming triangular mesh of 293 elements and 129 vertices. The mesh is partitioned into $N = 4, 8, \dots, 64$ subdomains using METIS [26]. Then, we perform $L = 0, 1, 2$ globally uniform refinements to create coarse meshes of different sizes (different H). These coarse meshes of the whole domain are then broadcast to N processors where they are adaptively refined by $l = 1, 2, 3$ cycles of adaptive refinements (vary h) in parallel. The adaptive refinement on each processor focuses on the associated local subdomain. However, there might be refinement outside to keep the final mesh on each processor conforming across the interface with the local meshes in the neighbouring subdomains. The global fine mesh is the union of the refined submeshes provided by all processors. The meshes at different stages of the experiment for $L = 0$, $N = 4$ and $l = 2$ is shown in Figure 3.

Table 1: Smallest and second largest eigenvalues ($\hat{\lambda}_{\min}$, $\hat{\lambda}_{\max}$) of the preconditioned system $P^{-1}A$ for meshes of different sizes and partitions.

	$L = 0$			$L = 1$			$L = 2$		
	$l = 1$	$l = 2$	$l = 3$	$l = 1$	$l = 2$	$l = 3$	$l = 1$	$l = 2$	$l = 3$
$N = 4$									
$\hat{\lambda}_{\min}$	0.980	0.962	0.945	0.973	0.948	0.935	0.973	0.948	0.935
$\hat{\lambda}_{\max}$	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000
$N = 8$									
$\hat{\lambda}_{\min}$	0.977	0.961	0.945	0.973	0.949	0.935	0.971	0.949	0.935
$\hat{\lambda}_{\max}$	3.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000
$N = 16$									
$\hat{\lambda}_{\min}$	0.980	0.963	0.945	0.973	0.948	0.935	0.971	0.948	0.935
$\hat{\lambda}_{\max}$	4.039	4.041	4.041	4.000	4.000	4.000	4.000	4.000	4.000
$N = 32$									
$\hat{\lambda}_{\min}$	0.981	0.967	0.950	0.975	0.951	0.937	0.972	0.951	0.937
$\hat{\lambda}_{\max}$	5.007	5.0079	5.007	5.000	5.000	5.000	5.000	5.000	5.000
$N = 64$									
$\hat{\lambda}_{\min}$	0.985	0.968	0.953	0.976	0.954	0.939	0.972	0.954	0.939
$\hat{\lambda}_{\max}$	6.033	6.035	6.035	6.000	6.000	6.000	6.000	6.000	6.000

From Table 1, we can see that there is little or no change in $\hat{\lambda}_{\min}$ when N and H (effectively L) vary. When h (effectively l) varies, there are changes in $\hat{\lambda}_{\min}$. However, these changes are not significant and we can safely say that $\hat{\lambda}_{\min}$ is bounded independently of N , H and h .

There are almost no change in $\hat{\lambda}_{\max}$ when H and h (effectively L and l) vary especially when the coarse meshes are sufficiently fine ($L = 1$ and $L = 2$). We do see that $\hat{\lambda}_{\max}$ increase consistently as N increases. However, the increase is modest and is at most proportional to the number of color N^c which is much smaller than the number of subdomain N when N is sufficiently large.

In conclusion, the behavior of $\hat{\lambda}_{\max}$ and $\hat{\lambda}_{\min}$ in this experiment agrees with the estimates in Theorem 7 and Theorem 11.

5.2. Seepage Under Dam

In the second experiment, we study seepage under a dam. The problem is modeled as the stationary state of groundwater flowing through porous media, which has the following governing equation:

$$\begin{aligned}
 -\nabla \cdot (k(x, y)\nabla h(x, y)) &= 0 && \text{in } \Omega, \\
 h &= g_D && \text{on } \partial_D\Omega, \\
 \frac{\partial h}{\partial n} &= g_N && \text{on } \partial_N\Omega,
 \end{aligned}$$

where $h(x, y)$ is the total hydraulic head and $k(x, y)$ is the hydraulic permeability coefficient. Here, we assume the hydraulic head is constant across the z direction and only work with two dimensional space.

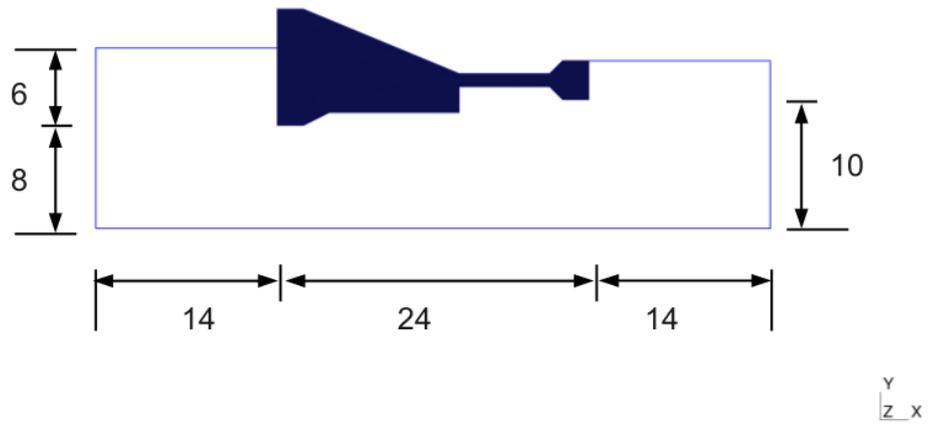


Figure 4: Geometry of the computational domain

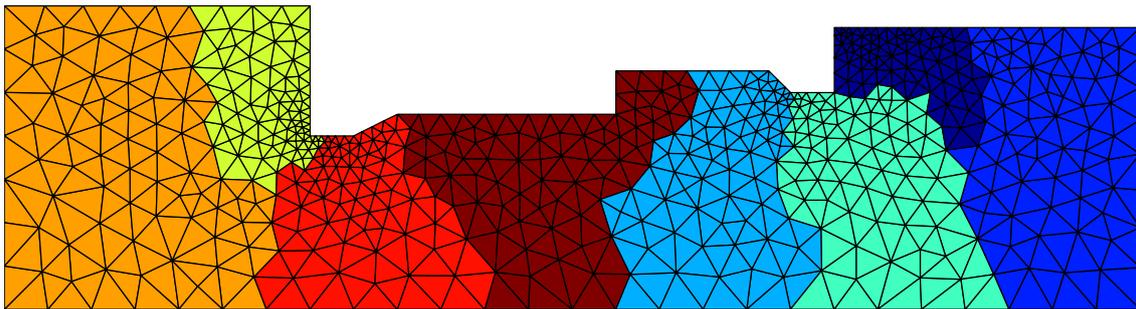


Figure 5: Seepage under a dam: a coarse mesh with partition

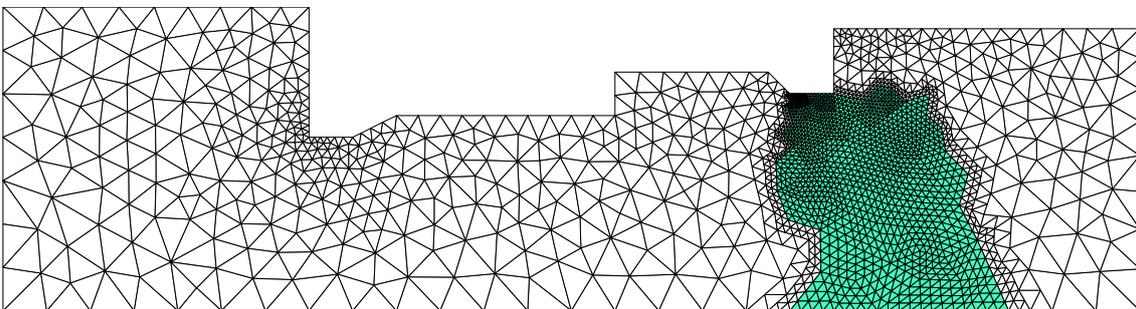


Figure 6: Seepage under a dam: a local adaptive mesh on a processor

Table 2: Numbers of iterations required to reduced the residual by a factor of 10^6 using the proposed preconditioner and two-level AS preconditioner (in parentheses).

N	$n_0 = 608$				$n_0 = 1042$			
	$l = 1$	$n = 2$	$l = 3$	$l = 4$	$l = 1$	$l = 2$	$l = 3$	$l = 4$
4	4(10)	4(10)	4(10)	4(10)	4(10)	4(10)	4(10)	4(10)
8	6(11)	7(11)	7(11)	7(13)	7(11)	7(11)	7(12)	6(15)
16	8(13)	9(13)	9(13)	8(13)	8(12)	9(11)	9(12)	8(14)
32	11(13)	12(13)	11(13)	11(17)	11(13)	11(12)	11(12)	11(15)
64	13(15)	13(13)	13(13)	12(16)	12(14)	13(13)	13(15)	13(17)

The simulation domain Ω , which is under the dam (in solid color), is illustrated in Figure 4. The upstream and downstream water level is assumed to be 10 and 1, respectively. This implies the Dirichlet boundary condition $h|_{D_1} = g_{D_1} = 10$ and $h|_{D_2} = g_{D_2} = 1$. In addition, we assume that the dam is made of impermeable material and there is no flow coming through the bottom, the left, the right boundary of the domain. Therefore, zero Neumann condition ($\partial h / \partial n = g_N = 0$) is imposed on associated boundary components. In our experiment, $k = 10^{-5}$.

We will study the convergence of the CG method in solving (3) with the proposed preconditioner P^{-1} . According to Theorem 3 and Theorem 12, we expect that the number of iterations required to reduced the residual below a fixed relative tolerance does not depend on the mesh sizes (H, h) and the number of subdomains (N) but on the number of colours (N^c).

Our construction of the experiment is similar to that of the first experiment. We create, by serial adaptivity, two unstructured conforming triangular meshes of $n_0 = 608$ and $n_0 = 1042$ vertices. These two coarse meshes, which has different mesh sizes (different H), are partitioned into $N = 4, 8, \dots, 64$ subdomains. For each combination of meshes and partitions, we apply $l = 1, 2, 3, 4$ levels of refinement locally on each subdomain to create local adaptive meshes of the whole domain. These local meshes are used to formulate the proposed preconditioner P^{-1} . The traditional two-level AS preconditioner \tilde{P}_{AS}^{-1} with the minimal overlap ($\delta = h$) is also formulated for comparison. Table 2 reports the number of CG iterations required to reduce the residual by a factor of 10^6 using P^{-1} and \tilde{P}_{AS}^{-1} (in parentheses).

Except for a few cases where the two preconditioners require the same number of iterations, P^{-1} clearly outperforms \tilde{P}_{AS}^{-1} , especially for large problems (large number of refinements l). With P^{-1} , the iteration number barely changes when l (effectively h) or n_0 (effectively H) vary. However, when l increases the overlap in \tilde{P}_{AS}^{-1} ($\delta = h$) decreases while the size of the subdomains remains the same. As the conditioned number with the two-level AS is proportional to $(1 + H/\delta)$ [21, 34], we see increase in the iteration number associated with \tilde{P}_{AS}^{-1} when l increases. In order for \tilde{P}_{AS}^{-1} to be competitive with P^{-1} , the overlap δ need to be at least l layers of elements. This implies more computational and communication cost associated with the overlap when l increases.

For both preconditioners, the iteration number increases as N increases. However, the increase is modest and is at most proportional to N^c , the number of colours.

This experiment clearly demonstrates that our preconditioner is more efficient than the traditional two-level AS preconditioner, especially for large problems requiring many levels of refinement. The experiment also confirms that the performance of our preconditioner is independent of the mesh sizes and the number of subdomains.

6. Conclusion

We have introduced a new additive Schwarz preconditioner to use in parallel finite elements where local meshes are meshes of the whole domain. From a practical point of view, it has the advantage of being able to obtain local matrices through assembling without communication and without forming the global matrix. It is proved to be optimal in the sense that the effective condition number of the preconditioned system can be bounded independently of the

coarse mesh size, fine mesh size and the number of subdomains. In comparison with the traditional two-level AS preconditioner, it does not require extra cost associated with the overlap to maintain the optimal convergence when the mesh is refined. The theoretical findings have been confirmed by numerical experiments.

7. Acknowledgement

We would like to thank Waad Subber for suggesting the seepage problem and providing the geometry data of the problem. This work was supported by the Numerical Algorithms and Intelligent Software Centre funded by the UK EPSRC grant EP/G036136 and the Scottish Funding Council.

- [1] O. Axelsson, G. Lindskog, On the rate of convergence of the preconditioned conjugate gradient method, *Numer. Math.* 48 (1986) 499–523.
- [2] I. Babuška, O.C. Zienkiewicz, J. Gago, E.R. de A. Oliveira (Eds.), *Accuracy estimates and adaptive refinements in finite element computations*, Wiley Series in Numerical Methods in Engineering, John Wiley & Sons Ltd., Chichester, 1986.
- [3] N.A. Baker, Poisson–boltzmann methods for biomolecular electrostatics, *Methods in enzymology* 383 (2004) 94–118.
- [4] N.A. Baker, D. Bashford, D.A. Case, Implicit solvent electrostatics in biomolecular simulation, in: *New algorithms for macromolecular simulation*, Springer, 2006, pp. 263–295.
- [5] N.A. Baker, D. Sept, M.J. Holst, J.A. McCammon, The adaptive multilevel finite element solution of the poisson-boltzmann equation on massively parallel computers, *IBM Journal of Research and Development* 45 (2001) 427–438.
- [6] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, Electrostatics of nanosystems: application to microtubules and the ribosome, *Proceedings of the National Academy of Sciences* 98 (2001) 10037–10041.
- [7] R.E. Bank, Hierarchical bases and the finite element method, in: *Acta numerica*, 1996, volume 5 of *Acta Numer.*, Cambridge Univ. Press, Cambridge, 1996, pp. 1–43.
- [8] R.E. Bank, Some variants of the Bank-Holst parallel adaptive meshing paradigm, *Comput. Vis. Sci.* 9 (2006) 133–144.
- [9] R.E. Bank, PLTMG: A Software Package for Solving Elliptic Partial Differential Equations, Users’ Guide 11.0, Technical Report, Department of Mathematics, University of California at San Diego, 2011. URL: <http://ccom.ucsd.edu/~reb>.
- [10] R.E. Bank, M. Holst, A new paradigm for parallel adaptive meshing algorithms, *SIAM J. Sci. Comput.* 22 (2000) 1411–1443 (electronic).
- [11] R.E. Bank, M. Holst, A new paradigm for parallel adaptive meshing algorithms, *SIAM Rev.* 45 (2003) 291–323 (electronic). Reprinted from *SIAM J. Sci. Comput.* 22 (2000), no. 4, 1411–1443 [MR1797889].
- [12] R.E. Bank, P.K. Jimack, A new parallel domain decomposition method for the adaptive finite element solution of elliptic partial differential equations. *Concurrency and Computation: Practice and Experience* 13 (2001) 327–350.
- [13] R.E. Bank, P.K. Jimack, S.A. Nadeem, S.V. Nepomnyaschikh, A weakly overlapping domain decomposition preconditioner for the finite element solution of elliptic partial differential equations, *SIAM J. Sci. Comput.* 23 (2002) 1817–1841 (electronic).
- [14] R.E. Bank, S. Lu, A domain decomposition solver for a parallel adaptive meshing paradigm, *SIAM J. Sci. Comput.* 26 (2004) 105–127 (electronic).
- [15] R.E. Bank, H. Nguyen, Domain decomposition and *hp*-adaptive finite elements, in: Y. Huang, R. Kornhuber, O. Widlund, J. Xu (Eds.), *Domain Decomposition Methods in Science and Engineering XIX*, volume 78 of *Lecture Notes in Computational Science and Engineering*, Springer, 2011, pp. 3–13.
- [16] R.E. Bank, H. Nguyen, Mesh regularization in Bank-Holst parallel *hp*-adaptive meshing, in: O.W. Randolph E. Bank, Michael Holst, J. Xu (Eds.), *Domain Decomposition Methods in Science and Engineering XX*, volume 91 of *Lecture Notes in Computational Science and Engineering*, Springer-Verlag, 2013, pp. 103–110.
- [17] R.E. Bank, H. Nguyen, A parallel *hp*-adaptive finite element method, in: *Recent Advances in Scientific Computing and Applications*, volume 586 of *Contemporary Mathematics*, Amer. Math. Soc., Providence, RI, 2013, pp. 23–33.
- [18] R.E. Bank, P.S. Vassilevski, Convergence analysis of a domain decomposition paradigm, *Computing and Visualization in Science* 11 (2008) 333–350.
- [19] H. Bi, Y. Yang, H. Li, Local and parallel finite element discretizations for eigenvalue problems, *SIAM J. Sci. Comput.* 35 (2013) A2575–A2597.
- [20] L. Chen, *iFEM: an innovative finite element methods package in MATLAB*, Technical Report, Department of Mathematics, University of California at Irvine, 2009.
- [21] M. Dryja, O.B. Widlund, Domain decomposition algorithms with small overlap, *SIAM J. Sci. Comput.* 15 (1994) 604–620. Iterative methods in numerical linear algebra (Copper Mountain Resort, CO, 1992).
- [22] V. Eijkhout, P. Vassilevski, The role of the strengthened Cauchy–Buniakowski–Schwarz inequality in multilevel methods, *SIAM Rev.* 33 (1991) 405–419.
- [23] G.H. Golub, C.F. Van Loan, *Matrix computations*, Johns Hopkins Studies in the Mathematical Sciences, third ed., Johns Hopkins University Press, Baltimore, MD, 1996.
- [24] Y. He, J. Xu, A. Zhou, Local and parallel finite element algorithms for the Navier-Stokes problem, *J. Comput. Math.* 24 (2006) 227–238.
- [25] R.A. Horn, C.R. Johnson, *Matrix analysis*, second ed., Cambridge University Press, Cambridge, 2013.
- [26] G. Karypis, V. Kumar, A fast and high quality multilevel scheme for partitioning irregular graphs, *SIAM J. Sci. Comput.* 20 (1998) 359–392 (electronic).
- [27] J.C. Mason, D.C. Handscomb, *Chebyshev polynomials*, Chapman & Hall/CRC, 2002.
- [28] W.F. Mitchell, The full domain partition approach to distributing adaptive grids, *Applied Numerical Mathematics* 26 (1998) 265–275.
- [29] W.F. Mitchell, Adaptive grid refinement and multigrid on cluster computers, in: *Proceedings of the 15th International Parallel & Distributed Processing Symposium*, Citeseer, p. 119.
- [30] A. Quarteroni, A. Valli, *Domain decomposition methods for partial differential equations*, volume 10, Clarendon Press, 1999.

- [31] L.R. Scott, S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions, *Math. Comp.* 54 (1990) 483–493.
- [32] Y. Shang, A parallel finite element variational multiscale method based on fully overlapping domain decomposition for incompressible flows, *Numer. Methods Partial Differential Equations* 31 (2015) 856–875.
- [33] Y. Shang, Y. He, Parallel finite element algorithm based on full domain partition for stationary stokes equations, *Applied Mathematics and Mechanics* 31 (2010) 643–650.
- [34] A. Toselli, O. Widlund, Domain decomposition methods—algorithms and theory, volume 34 of *Springer Series in Computational Mathematics*, Springer-Verlag, Berlin, 2005.
- [35] L.N. Trefethen, D. Bau, III, Numerical linear algebra, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.
- [36] S. Vey, A. Voigt, Adaptive full domain covering meshes for parallel finite element computations, *Computing* 81 (2007) 53–75.
- [37] J. Xu, Iterative methods by space decomposition and subspace correction, *SIAM Rev.* 34 (1992) 581–613.
- [38] J. Xu, A. Zhou, Local and parallel finite element algorithms based on two-grid discretizations, *Math. Comp.* 69 (2000) 881–909.
- [39] J. Xu, A. Zhou, Local and parallel finite element algorithms based on two-grid discretizations for nonlinear problems, *Adv. Comput. Math.* 14 (2001) 293–327.
- [40] H. Zheng, J. Yu, F. Shi, Local and Parallel Finite Element Algorithm Based on the Partition of Unity for Incompressible Flows, *J. Sci. Comput.* 65 (2015) 512–532.