Numerische Mathematik manuscript No.
(will be inserted by the editor)

Additive Schwarz Methods for Elliptic Mortar Finite Element Problems

Petter E. Bjørstad[1⋆], **Maksymilian Dryja**[2⋆⋆], **Talal Rahman**[3]

[1] Institute for Informatics, University of Bergen, N-5020 Bergen, Norway.
e-mail: `petter@ii.uib.no`

[2] Department of Mathematics, Warsaw University, Banacha 2, 02-097 Warsaw, Poland.
e-mail: `dryja@mimuw.edu.pl`

[3] Institute for Mathematics, University of Augsburg, Universitätsstr. 14, D-86159 Augsburg, Germany. e-mail: `talal.rahman@math.uni-augsburg.de`

Summary Two variants of the additive Schwarz method for solving linear systems arising from the mortar finite element discretization on nonmatching meshes of second order elliptic problems with discontinuous coefficients are designed and analyzed. The methods are defined on subdomains without overlap, and they use special coarse spaces, resulting in algorithms that are well suited for parallel computation. The condition number estimate for the preconditioned system in each method is proportional to the ratio H/h, where H and h are the mesh sizes, and it is independent of discontinuous jumps of the coefficients. For one of the methods presented the choice of the mortar (nonmortar) side is independent of the coefficients.

1 Introduction

In this paper we discuss two variants of the additive Schwarz method for solving systems of algebraic equations resulting from the approximation of second order elliptic problems with discontinuous coefficients. The discretization is obtained by using the mortar finite element method on nonmatching meshes, a technique which was first introduced by Bernardi, Maday, and Patera in [3]. The main characteristic of the method is that the

⋆ This work has been supported in part by the Norwegian Research Council, grant 113492/420.

⋆⋆ This work has been supported in part by the National Science Foundation, grant NSF-CCR-9732208 and in part by the Polish Science Foundation, grant 2P03A02116.

meshes on adjacent subdomains can be nonmatching across subdomain interfaces, allowing for the mesh on each subdomain to be generated independently of the rest of the domain. The coupling between adjacent subdomain meshes is done by imposing a weak continuity constraint on the finite element functions across subdomain interfaces. Several iterative solvers have been developed for the mortar element since it was first introduced, see for example [1,2,6–8,10–12,14,17–19], and the references therein.

The first method we introduce is an extension of the original average Schwarz method of Bjørstad, Dryja, and Vainikko [5] for second order elliptic problems using conforming finite elements, to the mortar case. This method has good convergence behavior, but uses a computationally expensive coarse space. We improve this situation by replacing the coarse space with two special coarse spaces, resulting in a method which is more parallel and yet retains the essential convergence behavior of the first method. The basic principle behind the construction of the second method is based on the idea used for the additive Schwarz method proposed by Bjørstad and Dryja in [4].

Since the methods are defined on nonoverlapping subdomains and since they contain special coarse spaces, they are very well suited for parallel computations. This is because, in each iterative step, the algorithms reduce to solving several independent subproblems, and this requires minimal communication for the computation of the residual. The rate of convergence of these methods when used as preconditioners for the conjugate gradient method, is proportional to $(H/h)^{1/2}$, where H and h are the parameters of the coarse and the fine triangulation, respectively. It should be pointed out here that the convergence is independent of the jumps of the coefficients. Similarly, in the first method the choice of mortar and nonmortar sides does not depend on the coefficients of the elliptic problem.

The remainder of this paper is organized as follows. In Section 2 we recall the mortar finite element method for our elliptic problem. We introduce our additive Schwarz methods in Section 3-4, and in Section 5 we touch on some implementation issues. Finally, some numerical examples are presented in Section 6.

2 The discrete problem

Let $\overline{\Omega} = \cup_{i=1}^{N} \overline{\Omega}_i$ be the partition of our computational domain in two dimensions, where each Ω_i is a polygonal subdomain, and the subdomains are nonoverlapping. We consider the following differential problem: Find $u^* \in H_0^1(\Omega)$ such that

$$a(u^*, v) = f(v), \; v \in H_0^1(\Omega), \tag{1}$$

where

$$a(u,v) = \sum_{i=1}^{N} a_i(u,v) = \sum_{i=1}^{N} \rho_i (\nabla u, \nabla v)_{L^2(\Omega_i)},$$

and

$$f(v) = \int_{\Omega} fv\ dx = \sum_{i=1}^{N} \int_{\Omega_i} fv\ dx,$$

with ρ_i being positive and constant. The coefficients ρ_i can however have large jumps across subdomain interfaces. We remark that the proposed methods can be used as preconditioners also for the problems where the coefficients ρ_i depend on x and are discontinuous only across subdomain interfaces. In this case the constant ρ_i can be taken as an average of $\rho_i(x)$ over the subdomain Ω_i.

We consider only the geometrically conforming case, i.e., the intersection between the closure of two different subdomains is either empty, a vertex, or a whole edge. The subdomains form a coarse triangulation of the whole domain Ω with the mesh parameter $H = \max_i H_i$, where H_i is the diameter of Ω_i. In each subdomain Ω_i, we use triangular elements to triangulate. We assume that the triangles touching the subdomain boundary $\partial\Omega_i$ are quasi-uniform, having a mesh size of order h_i. We do not put such a restriction on the interior triangles. We also assume that the coarse triangulation and the fine triangulation in each Ω_i are shape regular in the sense of [9]. The resulting triangulation can be nonmatching across subdomain interfaces.

Let $X_i(\Omega_i)$ be the finite element space of piecewise linear continuous functions defined on the triangulation of Ω_i and vanishing on $\partial\Omega_i \cap \partial\Omega$, and let

$$X^h(\Omega) = X_1(\Omega_1) \times X_2(\Omega_2) \cdots \times X_N(\Omega_N).$$

In order to describe our discrete problem, we need the following auxiliary notations and finite element spaces. Let Γ_{ij} be an open edge common to Ω_i and Ω_j, i.e., $\overline{\Gamma}_{ij} = \overline{\Omega}_i \cap \overline{\Omega}_j$, and let $W^{h_i}(\Gamma_{ij})$ and $W^{h_j}(\Gamma_{ij})$ be the restrictions of $X_i(\Omega_i)$ and $X_j(\Omega_j)$ onto Γ_{ij}, respectively. Note that each interface Γ_{ij} inherits two different discretizations from its two sides. We select one side of Γ_{ij} as the master side, called the mortar, and the other side as the slave side, called the nonmortar. Define the skeleton $\mathcal{S} = (\cup\partial\Omega_i) \setminus \partial\Omega$ as follows:

$$\overline{\mathcal{S}} = \cup_m \overline{\gamma}_m, \text{ with } \gamma_m \cap \gamma_n = \emptyset \text{ if } m \neq n,$$

where each γ_m denotes an open mortar edge. We write γ_m as $\gamma_{m(i)}$ if it is an edge of Ω_i, i.e., $\gamma_{m(i)} \subset \partial\Omega_i$. Let $\delta_m = \delta_{m(j)} \subset \partial\Omega_j$ be the corresponding

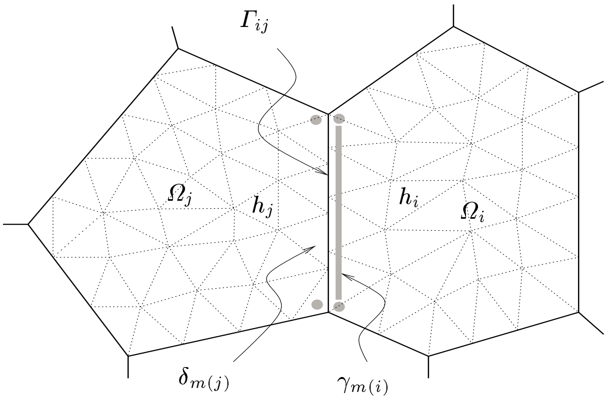

Fig. 1. A mortar and a nonmortar side of a subdomain interface with nonmatching grid on both sides.

open nonmortar edge of Ω_j that occupies the same geometrical space as $\gamma_{m(i)}$, i.e., $\gamma_{m(i)} = \Gamma_{ij} = \delta_{m(j)}$. See Figure 1 for an illustration.

In each figure of this paper a thick line is drawn on the mortar side of an interface, which represents the open mortar edge of that side. The corresponding nonmortar edge lies on the other side of the interface. The nonmortar edges are never shown explicitly in the figures. The end points of a mortar or a nonmortar are represented using thick dots. Note that each such point occupies the same geometrical space as one of the vertices of the subdomain to which its edge belongs. Further, we say that the function values (at the discretization points) on a mortar are nonzero if the edge is black and zero if the edge is light gray. The same applies to the end points. In addition "nz" (short form of nonzero) will be used to specify nonzeros on open nonmortar edges.

We denote the set of all vertices in Ω_i by ν_i. As a general rule for choosing the mortars and the nonmortars, we let $\gamma_{m(i)}$ be the mortar and $\delta_{m(j)}$ the corresponding nonmortar if $\rho_i \geq \rho_j$. This is, however, not necessary for our first method to be presented in the next section.

Since the triangulations on Ω_i and Ω_j may not match on their common interface Γ_{ij}, see Figure 1, the functions in $X^h(\Omega)$ can be discontinuous across the interface Γ_{ij}. We therefore impose a weak continuity condition across the interface called the mortar condition, see [3]. Let $u_h \in X^h$, where $u_h = \{u_i\}_{i=1}^N$. A function $u_h \in X^h$ satisfies the mortar condition on $\delta_{m(j)}$, if, for all functions $\psi \in M^{h_j}(\delta_{m(j)})$ $(\gamma_{m(i)} = \delta_{m(j)} = \Gamma_{ij})$,

$$\int_{\delta_{m(j)}} (u_{i|\gamma_{m(i)}} - u_{j|\delta_{m(j)}})\psi \, ds = 0. \tag{2}$$

Here the space $M^{h_j}(\delta_{m(j)})$ is a subspace of $W^{h_j}(\delta_{m(j)})$, with functions having constant values on elements touching $\partial\delta_{m(j)}$. The condition can be rewritten as follows. Let $\Pi_m(u_i, Tr\ v_j) : L^2(\delta_{m(j)}) \to W^{h_j}(\delta_{m(j)})$ denote a projection defined by

$$\int_{\delta_{m(j)}} \Pi_m(u_i, Tr\ v_j)\psi\ ds = \int_{\delta_{m(j)}} u_{i|\delta_{m(j)}}\psi\ ds, \tag{3}$$

for all functions $\psi \in M^{h_j}(\delta_{m(j)})$, and

$$\Pi_m(u_i, Tr\ v_j)_{|\partial\delta_{m(j)}} = v_{j|\partial\delta_{m(j)}}. \tag{4}$$

Here $Tr\ v_j$ denotes the trace of v_j. Thus, $u_{j|\delta_{m(j)}} = \Pi_m(u_i, Tr\ v_j)$ if $v_j = u_j$ on $\partial\delta_{m(j)}$.

V^h is a subspace of X^h of functions which satisfy the mortar condition for all $\delta_m \subset \mathcal{S}$. The discrete problem has the form: Find $u_h^* = \{u_i\}_{i=1}^N \in V^h$ such that

$$a(u_h^*, v_h) = f(v_h), \quad \forall v_h \in V^h \tag{5}$$

where

$$a(u_h, v_h) = \sum_{i=1}^N a_i(u_i, v_i) = \sum_{i=1}^N \rho_i(\nabla u_i, \nabla v_i)_{L^2(\Omega_i)},$$

and $v_h = \{v_i\}_{i=1}^N \in V^h$. V^h is a Hilbert space with an inner product defined by $a(u_h, v_h)$. The problem has a unique solution and its error bound is known, see [3].

The basis functions of the finite element space V^h

The finite element space V^h can be expressed as the span of its basis functions, i.e., $V^h = span\{\phi_k\}$, where each basis function ϕ_k is associated with a node x_k from the sets Ω_{ih}, $\gamma_{m(i)h}$ ($\gamma_{m(i)} \subset \partial\Omega_i$) and ν_i, for $i = 1, \cdots, N$, except those on $\partial\Omega$. Ω_{ih} and $\gamma_{m(i)h}$ denote the sets of nodal points of Ω_i and $\gamma_{m(i)}$, respectively. Let $x_k^{(i)}$ be the local representation of the node x_k, indicating that the node belongs to $\overline{\Omega}_i$. Let $\varphi_k^{(i)}$ denote the standard nodal basis function associated with the node $x_k^{(i)}$. The basis functions are defined as follows.

If $x_k \in \Omega_{ih}$, then $\phi_k(x)$ is exactly equal to $\varphi_k^{(i)}(x)$. If $x_k \in \gamma_{m(i)h}$, then $\phi_k(x) = \varphi_k^{(i)}(x)$ on $\overline{\Omega}_i$, while on $\overline{\delta}_{m(j)}$, where $\gamma_{m(i)} = \delta_{m(j)}$, $\phi_k(x) = \Pi_m(\varphi_k^{(i)}, 0)(x)$. ϕ_k is zero at the remaining nodal points of $\overline{\Omega}_j$, and zero everywhere on the remaining subdomains. For $x_k \in \nu_i$, ϕ_k takes its form depending on whether x_k is a point of intersection between two mortars, two nonmortars, or a mortar and a nonmortar. We define the function ϕ_k for $k \in \nu_i$ below.

We first consider when x_k is a point of intersection between a mortar $\gamma_{m(i)}$ and a nonmortar $\delta_{n(i)}$, i.e., $x_k \in \partial\gamma_{m(i)} \bigcap \partial\delta_{n(i)}$. Then, $\phi_k(x) = \varphi_k^{(i)}(x)$ on $\overline{\gamma}_{m(i)}$, while $\phi_k(x) = \Pi_m(\varphi_k^{(i)}, 0)(x)$ on $\overline{\delta}_{m(j)}$ ($\gamma_{m(i)} = \delta_{m(j)}$), and $\phi_k(x) = \Pi_n(0, Tr\ \varphi_k^{(i)})(x)$ on $\overline{\delta}_{n(i)}$. Finally, ϕ_k is zero at the remaining nodal points of $\overline{\Omega}_i$ and $\overline{\Omega}_j$ ($\gamma_{m(i)} = \delta_{m(j)}$), and zero everywhere on the remaining subdomains. In the similar way, ϕ_k is defined when x_k is a point of intersection between two mortars ($\gamma_{m(i)}$ and $\gamma_{n(i)}$) or two nonmortars ($\delta_{m(i)}$ and $\delta_{n(i)}$).

Using the basis functions of V^h, the problem (5) can now be rewritten in the matrix form as

$$\mathbf{Au}^* = \mathbf{f}, \tag{6}$$

where $\mathbf{u}^*$ is a vector of nodal values of u_h^*. The matrix $\mathbf{A}$ is symmetric and positive definite, and its condition number is of the same order as that of a conforming finite element method provided that h_i are all of the same order.

For the rest of this paper we make the following changes in notation. We drop the subscript h for functions in X^h and V^h. We use the function name (e.g. v_j) instead of the whole expression containing Tr followed by the function name (i.e., $Tr\ v_j$) as the second argument to $\Pi_m(.,.)$, in other words, we simply use $\Pi_m(u_i, v_j)$ to denote $\Pi_m(u_i, Tr\ v_j)$.

3 The additive average Schwarz method

In this section, we design an additive Schwarz method for the problem (5), which is inspired by the additive average Schwarz method of Bjørstad, Dryja, and Vainikko in [5]. The method is defined using the general framework for the additive Schwarz methods, see [16], i.e., in terms of a decomposition of the global space V^h into subspaces and the bilinear forms defined on these subspaces. We decompose V^h into smaller subspaces $V^{(i)}$ as

$$V^h = V^{(0)} + \sum_{i=1}^{N} V^{(i)}. \tag{7}$$

For $i = 1, \cdots, N$, $V^{(i)}$ is the restriction of V^h to Ω_i with zero on $\partial\Omega_i$ and on the remaining subdomains. $V^{(0)}$ is defined as a range of a linear operator I_A, i.e.,

$$V^{(0)} = \{I_A v :\ v \in V^h\}. \tag{8}$$

The operator I_A is defined as follows. For $u = \{u_i\}_{i=1}^N \in V^h$, $I_A u$, in $\overline{\Omega}_i$, is given by

$$(I_A u_i)(x) = \begin{cases} u_i(x), \, x \in \partial\Omega_{ih} \\ \overline{u}_i, \qquad x \in \Omega_{ih} \end{cases}, \tag{9}$$

where

$$\overline{u}_i = \frac{1}{|\nu_i|}\left(\sum_{\gamma_{m(i)} \subset \partial\Omega_i} \overline{u}_{\gamma_{m(i)}} + \sum_{\delta_{m(i)} \subset \partial\Omega_i, \, \gamma_{m(j)} = \delta_{m(i)}} \overline{u}_{\gamma_{m(j)}} \right). \tag{10}$$

Here Ω_{ih}, as defined earlier, and $\partial\Omega_{ih}$ are the sets of nodal points of Ω_i and $\partial\Omega_i$, respectively,

$$\overline{u}_{\gamma_{m(i)}} = \frac{1}{|\gamma_{m(i)}|} \int_{\gamma_{m(i)}} u_i \, ds \tag{11}$$

is the average of u_i over the mortar side $\gamma_{m(i)}$, $|\nu_i|$ is the number of vertices of Ω_i, and $|\gamma_{m(i)}|$ is the length of $\gamma_{m(i)}$. Note that, with $\delta_{m(i)} = \gamma_{m(j)}$ ($\gamma_{m(j)} \subset \partial\Omega_j$), the average of u_i over the nonmortar side $\delta_{m(i)}$, $\overline{u}_{\delta_{m(i)}}$, is equal to $\overline{u}_{\gamma_{m(j)}}$ by the mortar condition. On $\overline{\delta}_{m(i)}$, we also have

$$u_i(x) = \Pi_m(0, u_i) + \Pi_m(u_j, 0). \tag{12}$$

We use the exact bilinear form for all our subproblems, i.e., for $i = 0, \cdots, N$ and $u, v \in V^{(i)}$, we define $b^{(i)}(.,.) : V^{(i)} \times V^{(i)} \to \Re$ as

$$b^{(i)}(u, v) = a(u, v). \tag{13}$$

For $i = 1, \cdots, N$, any function $u \in V^{(i)}$ is of the form $u = \{u_j\}_{j=1}^N$ where $u_j = 0$ for $j \neq i$. Hence, for $i = 1, \cdots, N$, we have

$$a(u, v) = a_i(u_i, v_i), \tag{14}$$

where $u = \{u_j\}_{j=1}^N \in V^{(i)}$ and $v = \{v_j\}_{j=1}^N \in V^{(i)}$.

The projection like operators $T^{(i)} : V^h \to V^{(i)}$ are defined in the standard way, i.e., for $i = 0, \cdots, N$ and $u \in V^h$, $T^{(i)}u \in V^{(i)}$ is the solution of

$$b^{(i)}(T^{(i)}u, v) = a(u, v), \; v \in V^{(i)}. \tag{15}$$

Note that, for $i = 1, \cdots, N$, the problem (15) reduces to a Dirichlet problem on Ω_i with homogeneous boundary condition. To compute the corresponding stiffness matrices we need the standard nodal basis functions $\varphi_k^{(i)}$ associated with $x_k \in \Omega_{ih}$ (interior) only. For $i = 0$, the problem (15) is a coarse space problem, see the implementation section for the construction of the stiffness matrix. Let

$$T = T^{(0)} + T^{(1)} + \cdots + T^{(N)}.$$

The problem (5) and (6) are now replaced by the following preconditioned system,

$$Tu^* = g, \tag{16}$$

where $g = \sum_{i=0}^{N} T^{(i)} u^*$. Note that $T^{(i)}u^*$ can be calculated without knowing u^*, the solution of (5), see [16].

3.1 Analysis

Theorem 1 *For $u \in V^h$,*

$$c_0 \frac{h}{H} a(u,u) \leq a(Tu,u) \leq c_1 a(u,u), \tag{17}$$

where both c_0 and c_1 are positive constants independent of the mesh sizes $h = \inf_i h_i$ and $H = \max_i H_i$ and of the jumps of the coefficients ρ_i.

Proof We use the general Schwarz framework to prove the theorem, see Chapter 5 of the book [16], where we only need to verify three key assumptions.

Assumption 2. $V^{(0)}$ is the coarse space here, and since $\{V^{(i)}, i = 1, \cdots, N\}$ are orthogonal with respect to the bilinear form $a(.,.)$, we have $\rho(\mathcal{E}) = 1$.

Assumption 3. For $b^{(i)}, i = 0, \cdots, N$, we have $\omega = 1$ since only the exact bilinear forms are used, see (13).

Assumption 1. For $u \in V^h$, let $u^{(0)} = I_A u \in V^{(0)}$ and $w = u - u_0 = \{w_i\}_{i=1}^N$. It follows immediately that $w_i = 0$ on $\partial\Omega_i$. We set, for $i = 1, \cdots, N$, $u^{(i)} = (0, \cdots, 0, w_i, 0 \cdots, 0)$, where $u^{(i)}$ now belongs to $V^{(i)}$. Clearly, we have

$$u = u^{(0)} + \sum_{i=1}^{N} u^{(i)}, \tag{18}$$

with $u^{(0)} \in V^{(0)}$ and $u^{(i)} \in V^{(i)}, i = 1, \cdots, N$. We need to prove that

$$b^{(0)}(u^{(0)}, u^{(0)}) + \sum_{i=1}^{N} b^{(i)}(u^{(i)}, u^{(i)}) \leq c \frac{H}{h} a(u,u). \tag{19}$$

For $i = 1, \cdots, N$, note that $b^{(i)}(u^{(i)}, u^{(i)}) = a_i(w_i, w_i) = a_i(u_i - u_i^{(0)}, u_i - u_i^{(0)})$, where $u^{(0)} = \{u_i^{(0)}\}_{i=1}^N \in V^{(0)}$. Hence, by the triangle inequality, we have

$$\sum_{i=1}^{N} b^{(i)}(u^{(i)}, u^{(i)}) \leq 2a(u, u) + 2a(u^{(0)}, u^{(0)}). \tag{20}$$

For $u^{(0)}$, we have

$$b^{(0)}(u^{(0)}, u^{(0)}) = a(u^{(0)}, u^{(0)}) = \sum_{i=1}^{N} a_i(u_i^{(0)}, u_i^{(0)}). \tag{21}$$

Let n_i be the number of nodes on $\partial\Omega_i$. By the inverse inequality and the properties of I_A, we have on each Ω_i

$$\begin{aligned} a_i(u_i^{(0)}, u_i^{(0)}) &= \rho_i |u_i^{(0)} - \overline{u}_i|^2_{H^1(\Omega_i)} \\ &\leq c\rho_i \sum_{K \subset \overline{\Omega}_i^B} h_K^{-2} \parallel u_i^{(0)} - \overline{u}_i \parallel^2_{L^2(K)} \\ &\leq c\rho_i \sum_{x \in \partial\Omega_{ih}} (u_i(x) - \overline{u}_i)^2 \\ &\leq c\rho_i \left\{ h_i^{-1} \parallel u_i \parallel^2_{L^2(\partial\Omega_i)} + n_i \overline{u}_i^2 \right\}. \end{aligned} \tag{22}$$

The summation in the second line is taken over all elements K touching the boundary $\partial\Omega_i$, where Ω_i^B represents the union of all such elements in Ω_i. Since the subdomains are assumed to be shape regular, we can estimate the second term inside the curly brackets as follows.

$$n_i \overline{u}_i^2 \ \leq \ c n_i \frac{1}{|\nu_i|} \sum_{\Gamma \subset \partial\Omega_i} \frac{1}{|\Gamma|} \parallel u_i \parallel^2_{L^2(\Gamma)} \ \leq \ c h_i^{-1} \parallel u_i \parallel^2_{L^2(\partial\Omega_i)}. \tag{23}$$

Here Γ in the summation is $\gamma_{m(i)}$ and $\delta_{m(i)}$. We have also used the fact that $\overline{u}_{\delta_{m(i)}} = \overline{u}_{\gamma_{m(j)}}$ on $\delta_{m(i)} = \gamma_{m(j)}$. Again, for any constant α, we note that

$$u_i - \overline{u}_i = (u_i - \alpha) - \overline{(u_i - \alpha)}.$$

Using this, a simple trace inequality and the Poincaré inequality (choose α to be the average of u_i over the subdomain Ω_i,i.e., $\alpha = \frac{1}{|\Omega_i|} \int_{\Omega_i} u_i dx$), we have from (22)

$$\begin{aligned} a_i(u_i^{(0)}, u_i^{(0)}) &\leq c\rho_i h_i^{-1} \parallel u_i \parallel^2_{L^2(\partial\Omega_i)} \\ &\leq c\rho_i h_i^{-1} \left\{ H_i |u_i|^2_{H^1(\Omega_i)} + H_i^{-1} \parallel u_i \parallel^2_{L^2(\Omega_i)} \right\} \\ &\leq c \frac{H_i}{h_i} a_i(u_i, u_i). \end{aligned} \tag{24}$$

Summing (24) over $i = 1, \cdots, N$, and using the sum in (20) we finally get (19). The proof of the theorem follows.

4 Coarse reformulation

In this section we present our second method which is obtained from the additive average Schwarz method by simply replacing its coarse space (call it $V^{(0)}_{avg}$) with two new coarse spaces, $V^{(-1)}$ and $V^{(0)}$, which can be treated independently. $V^{(-1)}$ is a space of functions given by their values on $\mathcal{S}$, and $V^{(0)}$ is a special space having a dimension equal to the number of subdomains. The sum of these two new spaces will contain the space $V^{(0)}_{avg}$, see Remark 2. The basic principle for this approach is based on the idea of Bjørstad and Dryja in [4].

The decomposition of the finite element space V^h takes the form

$$V^h \;=\; V^{(-1)} \;+\; V^{(0)} \;+\; \sum_{i=1}^{N} V^{(i)}, \tag{25}$$

where $V^{(i)}, i = 1, \cdots, N$, are the local subspaces defined in the previous section. The space $V^{(-1)}$ is defined as follows.

$$V^{(-1)} = \left\{ v \in V^h : v(x) = 0,\; x \in \overline{\Omega}_h \setminus (\cup_i \partial\Omega_{ih}) \right\}. \tag{26}$$

The coarse space $V^{(0)}$ is defined in the following paragraphs.

Definition 1 *We say that a node x_k is connected to the subdomain Ω_i if $x_k \in \overline{\Omega}_{ih}$. If the node $x_k \in \overline{\gamma}_{m(i)h}$ ($x_k \in \overline{\delta}_{m(i)h}$) then x_k is said to be connected to both Ω_i and Ω_j if $\gamma_{m(i)} = \delta_{m(j)}$ ($\delta_{m(i)} = \gamma_{m(j)}$), in other words a node belonging to a subdomain is connected to a neighbor subdomain only through the mortars or the nonmortars the node belongs to. As an example see Figure 3(c), where the node $x^{(k)}_a$ is connected to exactly three subdomains, Ω_k, Ω_i and Ω_j only. Accordingly, for 2-D the maximum number of subdomains a node can be connected to is three.*

Let χ_i, associated with the subdomain Ω_i, be the piecewise linear continuous function on the triangulation of Ω_i, defined by its nodal values at $x \in \overline{\Omega}_{ih}$. For each such node x,

$$\chi_i(x) = \frac{1}{\sum_j \rho_j(x)},$$

where the sum is taken over the subdomains that x is connected to, in the sense of Definition 1. Note that for $\rho_i = \rho_j = 1$, χ_i is 1 for $x \in \Omega_{ih}$, $\frac{1}{2}$ for $x \in (\partial\Omega_{ih} \setminus \nu_i)$ and $\frac{1}{3}$ for $x \in \nu_i$.

We partition the set of subdomains $\Omega_i, i = 1, \cdots, N$, into two subsets, the set of interior subdomains N_I, and the set of boundary subdomains N_B. Subdomains belonging to N_B intersect the boundary $\partial\Omega$ in at least one point, while those of the interior set N_I do not. N_B is partitioned further into two new subsets, N_{BE} and N_{BV}, where N_{BE} is the set of subdomains that share at least an edge with $\partial\Omega$, while N_{BV} contains subdomains that share only a vertex with $\partial\Omega$.

We associate with each subdomain Ω_i the sets G_i and Q_i containing the indices of its neighboring subdomains defined as follows. G_i contains the index of a neighbor Ω_j if it shares an edge Γ_{ij} ($\overline{\Gamma}_{ij} = \overline{\Omega}_i \cap \overline{\Omega}_j$) with Ω_i. Q_i contains the index of a neighbor Ω_j if $\overline{\Omega}_i \cap \overline{\Omega}_j$ is a crosspoint, there is a subdomain Ω_k such that Γ_{ki} ($\overline{\Gamma}_{ki} = \overline{\Omega}_k \cap \overline{\Omega}_i$) and Γ_{jk} ($\overline{\Gamma}_{jk} = \overline{\Omega}_j \cap \overline{\Omega}_k$) are the two edges of Ω_k which intersect at that crosspoint, and Γ_{ki} is a mortar in Ω_k, cf. Figure 3(c). Two other sets, G_i' and Q_i', are also defined in this connection. G_i' is the subset of G_i, containing the index of Ω_j for which the edge Γ_{ij} touches $\partial\Omega$. Q_i' is the subset of Q_i, containing the index of Ω_j for which the corresponding crosspoint lies on $\partial\Omega$.

We are now ready to define the coarse space $V^{(0)}$ which is given as the span of its basis functions, $\Phi_i, i = 1, \cdots, N$, i.e.,

$$V^{(0)} = span\{\Phi_i : i = 1, \cdots, N\}. \tag{27}$$

Each function Φ_i, associated with the subdomain Ω_i, is a function in the finite element space V^h.

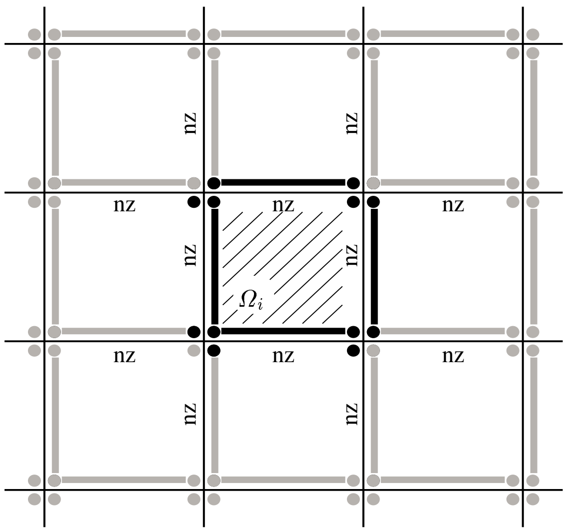

Fig. 2. Illustration of a basis function Φ_i associated with an interior subdomain Ω_i, indicating places, like the subdomain interior (shaded area), the mortars (dark thick line), the nonmortars (marked "nz"), and the end points (dark thick dots), where the function takes nonzero values.

For an interior subdomain Ω_i ($\Omega_i \in N_I$), see Figure 2, the function Φ_i is constructed in three steps. We define Φ_i first (*i*) on $\overline{\Omega}_i$, then (*ii*) on $\overline{\Omega}_j$ for $i \in G_j$, and then (*iii*) on $\overline{\Omega}_j$ for $i \in Q_j$.

(*i*) Φ_i on $\overline{\Omega}_i$ (cf. Figure 2) is given as

$$\Phi_i(x) = \begin{cases} 1, & x \in \Omega_{ih}, \\ \rho_i \chi_i(x), & x \in \gamma_{m(i)h} \cup \nu_i, \\ \rho_i \Pi_m(\chi_j, \chi_i)(x), & x \in \delta_{m(i)h}, \delta_{m(i)} = \gamma_{m(j)}. \end{cases} \tag{28}$$

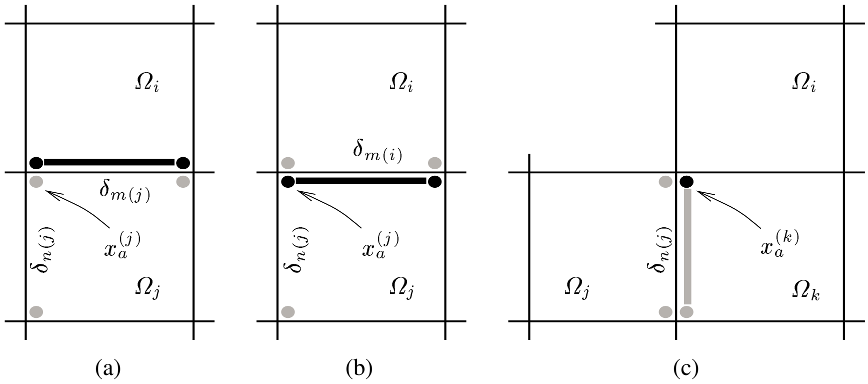

Fig. 3. Illustrating Φ_i (associated with an interior subdomain Ω_i) on $\overline{\Omega}_j$, where $i \in G_j$ ((a) and (b)) and $i \in Q_j$ ((c)).

(*ii*) Φ_i on $\overline{\Omega}_j$, where $i \in G_j$, we have two cases to consider. For the first case, let $\Gamma_{ij} = \delta_{m(j)} = \gamma_{m(i)}$, see Figure 3(a). Then, on $\overline{\Omega}_j$,

$$\Phi_i(x) = \begin{cases} \rho_i \Pi_m(\chi_i, \chi_j)(x), & x \in \overline{\delta}_{m(j)h},\ \delta_{m(j)} = \gamma_{m(i)}, \\ \Phi(x), & x \in \overline{\delta}_{n(j)h},\ \partial\delta_{n(j)} \cap \partial\delta_{m(j)} \neq \emptyset, \\ 0, & \text{at all other } x \text{ in } \overline{\Omega}_{jh}. \end{cases} \tag{29}$$

For the second case, let $\Gamma_{ij} = \gamma_{m(j)} = \delta_{m(i)}$, see Figure 3(b). Φ_i on $\overline{\Omega}_j$ is then given as

$$\Phi_i(x) = \begin{cases} \rho_i \chi_j(x), & x \in \overline{\gamma}_{m(j)h},\ \gamma_{m(j)} = \delta_{m(i)}, \\ \Phi(x), & x \in \overline{\delta}_{n(j)h},\ \partial\delta_{n(j)} \cap \partial\gamma_{m(j)} \neq \emptyset, \\ 0, & \text{at all other } x \text{ in } \overline{\Omega}_{jh}. \end{cases} \tag{30}$$

The function $\Phi(x)$ in both (29) and (30) depend on the configurations of Ω_i and Ω_j. Let $\Gamma_{jk} = \delta_{n(j)} = \gamma_{n(k)}$ for some k. Both in (29) and (30), $\partial\Gamma_{ij} \cap \partial\Gamma_{jk}$ is nonempty. There are two different situations we need to consider for $\Phi(x)$. For the first, let Γ_{ki} be the edge such that $\partial\Gamma_{ki} \cap \partial\Gamma_{jk}$ is nonempty, see Figure 4 for an illustration. Then we have

$$\Phi(x) = \rho_i \chi_j(x_a^{(j)}) \Pi_n(0, \varphi_a^{(j)}) + \rho_i \chi_k(x_a^{(k)}) \Pi_n(\varphi_a^{(k)}, 0), \qquad (31)$$

where $x_a^{(j)} = \partial\Gamma_{ij} \cap \partial\Gamma_{jk} \in \nu_j$ and $x_a^{(k)} = \partial\Gamma_{ki} \cap \partial\Gamma_{jk} \in \nu_k$. Note that $x_a^{(j)} = x_a^{(k)}$ (geometrically occupying the same space) since we only consider the geometrically conforming case. For the other situation, where there exists no such Γ_{ki} (cf. figures 3a-3b), we simply have

$$\Phi(x) = \rho_i \chi_j(x_a^{(j)}) \Pi_n(0, \varphi_a^{(j)}). \qquad (32)$$

Note that the first term in (31) is the same as the expression (32), and the second term reappears in step (*iii*) below. So, without any loss of generality, we can assume, for the sake of simplicity, that the situation illustrated in Figure 4 never occurs. The expression for $\Phi(x)$ is then given by (32) only. We note also that the sets G_i and Q_i are now disjoint.

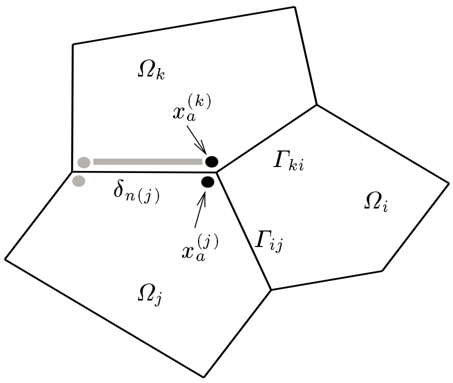

Fig. 4. Illustrating the situation for (31), which results due to having a vertex of Ω_i as a cross point of exactly three subdomains.

(*iii*) Φ_i on $\overline{\Omega}_j$, where $i \in Q_j$, is given as follows. Let Γ_{ki} and Γ_{jk} be the two edges such that $\Gamma_{jk} = \delta_{n(j)} = \gamma_{n(k)}$ and $x_a^{(k)} = \partial\Gamma_{ki} \cap \partial\Gamma_{jk} \in \nu_k$, see Figure 3(c). Then

$$\Phi_i(x) = \begin{cases} \rho_i \chi_k(x_a^{(k)}) \Pi_n(\varphi_a^{(k)}, 0), & x \in \overline{\delta}_{n(j)h}, \\ 0, & \text{at all other } x \text{ in } \overline{\Omega}_{jh}. \end{cases} \qquad (33)$$

On the remaining subdomains, $\Phi_i = 0$. This completes the definition of Φ_i for $\Omega_i \in N_I$, see Figure 2 for an example.

If Ω_i is a boundary subdomain ($\Omega_i \in N_B$) then the function Φ_i is defined as above but by imposing $\chi_j(x) = 0$ at $x \in \partial\Omega_{jh} \cap \partial\Omega_h$ for all

$\Omega_j \in N_B$. The values of Φ_i on some nonmortar edges (cf. Figure 5) touching $\partial\Omega$ will be different and therefore needs to be redefined. Below, we define the function on those edges.

Let $\{\hat{\Phi}_i\}_{\Omega_i \in N_B}$ be the set of functions constructed in exactly the same way as the functions associated with the interior subdomains, i.e., by following the steps (*i*)-(*iii*) above. We use the following rule: Any edge of $\Omega_i \in N_{BE}$ which lies on the boundary $\partial\Omega$ is a mortar in Ω_i. The function Φ_i, for $\Omega_i \in N_B$, on nonmortar edges (cf. Figure 5) touching the boundary $\partial\Omega$ can now be obtained from the corresponding $\hat{\Phi}_i$.

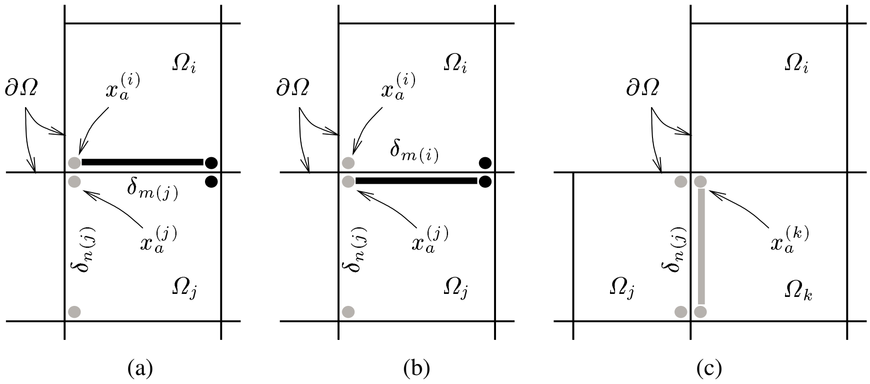

Fig. 5. Illustrating Φ_i associated with a boundary subdomain Ω_i. The function takes zero values at the nodes on the outer boundary $\partial\Omega$.

Let $x_a^{(i)} \in \nu_i \cap \partial\Omega_h$ and $x_a^{(j)} \in \nu_j \cap \partial\Omega_h$ be the end points (geometrically occupying the same space) of $\delta_{m(i)}$ and $\gamma_{m(j)}$, respectively, such that $\delta_{m(i)} = \gamma_{m(j)} \subset \partial\Omega_j$, see Figure 5(b). At $x \in \overline{\delta}_{m(i)h}$, we have (cf. (28))

$$\Phi_i = \hat{\Phi}_i - \rho_i\chi_i(x_a^{(i)})\Pi_m(0, \varphi_a^{(i)}) - \rho_i\chi_j(x_a^{(j)})\Pi_m(\varphi_a^{(j)}, 0). \quad (34)$$

The last two terms above are simply due to $\varphi_a^{(i)}$ and $\varphi_a^{(j)}$, respectively.

For nonmortar edges in the neighboring subdomains Ω_j, we need to consider only for $i \in G'_j$ and $i \in Q'_j$ since in the other cases there will be no change.

Consider $i \in G'_j$. Let $x_a^{(i)} \in \nu_i \cap \partial\Omega_h$ and $x_a^{(j)} \in \nu_j \cap \partial\Omega_h$ be the end points (geometrically occupying the same space) of $\gamma_{m(i)}$ and $\delta_{m(j)}$, respectively, such that $\gamma_{m(i)} = \delta_{m(j)}$, see Figure 5(a). At $x \in \overline{\delta}_{m(j)h}$, we have (cf. (29))

$$\Phi_i = \hat{\Phi}_i - \rho_i\chi_i(x_a^{(i)})\Pi_m(\varphi_a^{(i)}, 0) - \rho_i\chi_j(x_a^{(j)})\Pi_m(0, \varphi_a^{(j)}). \quad (35)$$

Let $\delta_{n(j)}$ be the nonmortar referred to in (29) and (30). $\Phi(x)$, in this case, is given by (32) only. Now let $x_a^{(j)} \in \nu_j \cap \partial\Omega$, see Figures 5(a)-(b). At

$x \in \overline{\delta}_{n(j)h}$, we then have $\Phi_i = \hat{\Phi}_i - \rho_i \chi_j(x_a^{(j)}) \Pi_n(0, \varphi_a^{(j)}) = 0$ which follows from (32).

Consider now $i \in Q'_j$. Let $\delta_{n(j)}$ be the nonmortar referred to in (33) with $x_a^{(k)} \in \nu_k \cap \partial\Omega$, see Figure 5(c). From (33), it follows that $\Phi_i = \hat{\Phi}_i - \rho_i \chi_j(x_a^{(k)}) \Pi_n(\varphi_a^{(k)}, 0) = 0$ at $x \in \overline{\delta}_{n(j)h}$. Consequently, Φ_i is zero everywhere on $\overline{\Omega}_j$.

A somewhat similar but simpler coarse space defined in terms of discrete harmonic functions in the context of substructuring algorithms for the mortar finite element problem has been used by Dryja in [11]. Similar spaces for additive Schwarz methods using conforming elements can be found in [4, 13].

Remark 1 The functions Φ_i have the property that for $\Omega_i \in N_I$, $\Phi_i(x) + \sum_{j \neq i} \Phi_j(x) = 1$ at all $x \in \overline{\Omega}_i$, and for $\Omega_i \in N_B$ this equality is true everywhere in $\overline{\Omega}_i$ except for the triangles touching the boundary $\partial\Omega$ in at least one point.

Remark 2 Let $V_{avg}^{(0)}$ be the coarse space defined in the previous section and let $\psi = \{\psi_i\}_{i=1}^N$ be any function in that space. Define $\psi^{(0)} = \sum_{i=1}^N \overline{\psi}_i \Phi_i$, where $\overline{\psi}_i$ is the average of ψ_i on Ω_i defined by (10), and $\psi^{(-1)} = \psi - \psi^{(0)}$. Clearly, $\psi^{(0)} \in V^{(0)}$ and $\psi^{(-1)} \in V^{(-1)}$. The space $V^{(-1)} + V^{(0)}$ has a higher dimension than $V_{avg}^{(0)}$, and hence $V_{avg}^{(0)} \subset V^{(-1)} + V^{(0)}$.

As in our first method, we use the exact bilinear form for all our subproblems here. For $i = -1, \cdots, N$, and $u, v \in V^{(i)}$, $b^{(i)}(.,.) : V^{(i)} \times V^{(i)} \to \Re$ are the bilinear forms defined as in (13). Let

$$T = T^{(-1)} + T^{(0)} + \cdots + T^{(N)},$$

be the additive operator, where $T^{(i)} : V^h \to V^{(i)}$ is the orthogonal projection in the sense of scalar product $a(.,.)$, and $T^{(i)}u \in V^{(i)}$ is the solution of (15) for $i = -1, \cdots, N$.

4.1 Analysis

Theorem 2 *For $u \in V^h$,*

$$c_0 \frac{h}{H} a(u,u) \leq a(Tu,u) \leq c_1 a(u,u), \tag{36}$$

where both c_0 and c_1 are positive constants independent of the mesh sizes $h = \inf_i h_i$ and $H = \max_i H_i$ and of the jumps of the coefficients ρ_i.

In order to prove the theorem we will need some estimates involving the coarse space $V^{(0)}$, which we now show in Lemma 1.

Lemma 1 *For* $u = \{u_i\}_{i=1}^N \in V^h$ *let* $u^{(0)} = \sum_{i=1,\cdots,N} \overline{u}_i \Phi_i$*, where* $\overline{u}_i$ *is now the average of* u_i *over* $\partial\Omega_i$*, i.e.,* $\overline{u}_i = \frac{1}{|\partial\Omega_i|}\int_{\partial\Omega_i} u_i \, ds$*. We have*

$$a(u^{(0)}, u^{(0)}) \leq c\frac{H}{h} a(u,u) \tag{37}$$

and

$$\sum_{i=1}^{N} \rho_i h_i^{-1} \parallel u^{(0)} - \overline{u}_i \parallel_{L^2(\Omega_i)}^2 \leq cHa(u,u), \tag{38}$$

where c *is a positive constant independent of the mesh sizes* h *and* H *and of the jumps of the coefficients* ρ_i*.*

Proof We prove the first inequality and then show how the second inequality can be derived from the first one. For the proof, we look at each subdomain at a time, and also, treat the interior and the boundary subdomains separately.

Consider first the case when Ω_i is an interior subdomain ($\Omega_i \in N_I$). $u^{(0)}$ on $\overline{\Omega}_i$ can be written as

$$u^{(0)} = \overline{u}_i \Phi_i + \sum_{j \in G_i \cup Q_i} \overline{u}_j \Phi_j .$$

It follows from the definition of $u^{(0)}$ and Remark 1 that, on $\overline{\Omega}_i$,

$$u^{(0)} - \overline{u}_i = \sum_{j \in G_i \cup Q_i} (\overline{u}_j - \overline{u}_i) \Phi_j .$$

Using this and the triangle inequality we find

$$\begin{aligned} a_i(u^{(0)}, u^{(0)}) &= \rho_i |u^{(0)} - \overline{u}_i|_{H^1(\Omega_i)}^2 \\ &\leq c \sum_{j \in G_i \cup Q_i} \rho_i (\overline{u}_j - \overline{u}_i)^2 |\Phi_j|_{H^1(\Omega_i)}^2 . \end{aligned} \tag{39}$$

We estimate the above sum in two steps.

(i) *Estimation of the factor* $(\overline{u}_j - \overline{u}_i)^2$ *in (39).* We estimate the factor $(\overline{u}_j - \overline{u}_i)^2$ for $j \in G_i$ and $j \in Q_i$ separately.

Let $j \in G_i$, and we consider only $\delta_{m(i)} = \gamma_{m(j)}$ ($\gamma_{m(i)} = \delta_{m(j)}$ gives the same estimate which can be derived using the same technique). Let $\overline{u}_{\gamma_{m(j)}}$ and $\overline{u}_{\delta_{m(i)}}$ be the averages over the mortar side $\gamma_{m(j)}$ and the nonmortar side $\delta_{m(i)}$, respectively, as defined in Section 3. By the mortar condition, we know that $\overline{u}_{\gamma_{m(j)}} = \overline{u}_{\delta_{m(i)}}$. We thus have by the triangle inequality

$$(\overline{u}_j - \overline{u}_i)^2 \leq 2\left\{(\overline{u}_j - \overline{u}_{\gamma_{m(j)}})^2 + (\overline{u}_i - \overline{u}_{\delta_{m(i)}})^2\right\} . \tag{40}$$

Note also that

$$\overline{u}_j - \overline{u}_{\gamma_{m(j)}} = -\overline{(u_j - \overline{u}_j)}_{\gamma_{m(j)}}.$$

Using this and the Cauchy-Schwarz inequality, we estimate the first term inside the curly brackets of (40) as

$$\begin{aligned}(\overline{u}_j - \overline{u}_{\gamma_{m(j)}})^2 &= \left\{ \frac{1}{|\gamma_{m(j)}|} \int_{\gamma_{m(j)}} (u_j - \overline{u}_j)\, ds \right\}^2 \\ &\leq cH_j^{-1} \parallel u_j - \overline{u}_j \parallel^2_{L^2(\gamma_{m(j)})} \\ &\leq c|u_j|^2_{H^1(\Omega_j)}. \end{aligned} \tag{41}$$

The last inequality above follows from (42) below, which we show now. For any constant α, we note that $\alpha - \overline{u}_j = -\overline{(u_j - \alpha)}_j$. Hence, by the Cauchy-Schwarz inequality,

$$(\alpha - \overline{u}_j)^2 = \left\{ \frac{1}{|\partial\Omega_j|} \int_{\partial\Omega_j} (u_j - \alpha)\, ds \right\}^2 \leq cH_j^{-1} \parallel u_j - \alpha \parallel^2_{L^2(\partial\Omega_j)} .$$

By the triangle inequality, the above estimate, a simple trace inequality and the Poincaré inequality (choose $\alpha = \frac{1}{|\Omega_j|} \int_{\Omega_j} u_j dx$), we have

$$\begin{aligned}\parallel u_j - \overline{u}_j \parallel^2_{L^2(\partial\Omega_j)} &\leq 2 \left\{ \parallel u_j - \alpha \parallel^2_{L^2(\partial\Omega_j)} + \parallel \alpha - \overline{u}_j \parallel^2_{L^2(\partial\Omega_j)} \right\} \\ &\leq c \parallel u_j - \alpha \parallel^2_{L^2(\partial\Omega_j)} \\ &\leq cH_j |u_j|^2_{H^1(\Omega_j)}. \end{aligned} \tag{42}$$

The second term on the right hand side of (40) can be estimated in exactly the same way, but this time we take the average over $\delta_{m(i)}$ instead of $\gamma_{m(j)}$. Hence, for $j \in G_i$, we obtain

$$(\overline{u}_j - \overline{u}_i)^2 \leq c \left\{ |u_i|^2_{H^1(\Omega_i)} + |u_j|^2_{H^1(\Omega_j)} \right\}. \tag{43}$$

Now consider $j \in Q_i$. An estimate similar to (43) can be derived. Let $x_a^{(k)}$ be the vertex of Ω_k such that the two edges of Ω_k, Γ_{jk} and Γ_{ki}, intersect at $x_a^{(k)}$, see Figure 3(c). By the triangle inequality,

$$\begin{aligned}(\overline{u}_j - \overline{u}_i)^2 \leq c \Big\{ &(\overline{u}_j - \overline{u}_{\Gamma_{jk}})^2 + (\overline{u}_k - \overline{u}_{\Gamma_{jk}})^2 \\ &+ (\overline{u}_k - \overline{u}_{\Gamma_{ki}})^2 + (\overline{u}_i - \overline{u}_{\Gamma_{ki}})^2 \Big\}, \end{aligned} \tag{44}$$

where we have used the same fact which we have used before, i.e., the averages over an interface from its both sides are the same. Applying the same steps as in (41) on each term of (44), we obtain

$$(\overline{u}_j - \overline{u}_i)^2 \leq c \left\{ |u_i|^2_{H^1(\Omega_i)} + |u_k|^2_{H^1(\Omega_k)} + |u_j|^2_{H^1(\Omega_j)} \right\}. \tag{45}$$

(2) *Estimation of the factor* $\rho_i|\Phi_j|^2_{H^1(\Omega_i)}$ *in (39).* By the inverse inequality and the discrete L^2-norm, we have

$$\rho_i|\Phi_j|^2_{H^1(\Omega_i)} \leq c\rho_i \sum_{K\subset\overline{\Omega}_i^B} h_K^{-2} \parallel \Phi_j \parallel^2_{L^2(K)} \leq c\rho_i \sum_{x\in\partial\Omega_{ih}} \Phi_j^2(x) \quad (46)$$

since Φ_j is zero at all $x \in \Omega_{ih}$. Hence, in order to estimate the sum in (46) we need to consider the function Φ_j only at places, along $\partial\Omega_i$, where it is nonzero. Again, we estimate the sum (factor) for $j \in G_i$ and $j \in Q_i$ separately.

We first consider $j \in G_i$. Let Γ_{ij}, the edge shared between Ω_i and Ω_j, be a mortar $\gamma_{m(i)}$ in Ω_i. Then, by the definition of the basis functions (cf. (30)), the function Φ_j on $\overline{\gamma}_{m(i)}$ is

$$\Phi_j = \rho_j\chi_i. \quad (47)$$

Similarly, let Γ_{ij} be a nonmortar $\delta_{m(i)}$ in Ω_i. Again, by the definition of the basis functions (cf. (29)) and (12), Φ_j on $\overline{\delta}_{m(i)}$ is

$$\begin{aligned}\Phi_j &= \rho_j\Pi_m(\chi_j,0)\\ &\quad + \rho_j\chi_i(x_a^{(i)})\Pi_m(0,\varphi_a^{(i)}) + \rho_j\chi_i(x_b^{(i)})\Pi_m(0,\varphi_b^{(i)}),\end{aligned} \quad (48)$$

where $x_a^{(i)} \in \nu_i$ and $x_b^{(i)} \in \nu_i$ are the two end points of $\delta_{m(i)}$.

Now let $\gamma_{n(i)}$ or $\delta_{n(i)}$ be another edge in Ω_i that intersects Γ_{ij} ($\gamma_{m(i)}$ or $\delta_{m(i)}$) at $x_a^{(i)}$. On $\overline{\gamma}_{n(i)}$, Φ_j is represented by the nodal values: $\Phi_j(x) = 0$ at all $x \in \overline{\gamma}_{n(i)h}$ except at $x_a^{(i)}$, where it is equal to $\rho_j\chi_i(x_a^{(i)})$. Φ_j on $\overline{\delta}_{n(i)}$ will simply be (cf. (32))

$$\Phi_j = \rho_j\chi_i(x_a^{(i)})\Pi_n(0,\varphi_a^{(i)}). \quad (49)$$

Let $\gamma_{o(i)}$ ($\delta_{o(i)}$) be the third edge in Ω_i that intersects Γ_{ij} ($\gamma_{m(i)}$ or $\delta_{m(i)}$) at $x_b^{(i)}$. Φ_j on $\gamma_{o(i)}$ ($\delta_{o(i)}$) has the similar representation as for $\gamma_{n(i)}$ ($\delta_{n(i)}$) above. Φ_j is zero everywhere else on $\partial\Omega_i$.

As we now know how to express Φ_j ($j \in G_i$) (using (47)-(49)) anywhere on $\partial\Omega_i$, we estimate the sum in (46) here. To make our calculation simple, we assume that $\gamma_{o(i)}$ (not $\delta_{o(i)}$) intersects Γ_{ij} at $x_b^{(i)}$. There are four combinations we need to investigate: $\gamma_{m(i)}$ with $\gamma_{n(i)}$, $\delta_{m(i)}$ with $\delta_{n(i)}$, $\gamma_{m(i)}$ with $\delta_{n(i)}$, and $\delta_{m(i)}$ with $\gamma_{n(i)}$.

Let $\gamma_{m(i)}$ and $\gamma_{n(i)}$ intersect at $x_a^{(i)}$. Using (47), the triangle inequality, and the fact that $\rho_j\chi_i(x) \leq 1$ for all $x \in \overline{\gamma}_{m(i)}$, we obtain

$$\begin{aligned}\rho_i \sum_{x\in\partial\Omega_{ih}} \Phi_j^2(x) = \rho_i \sum_{x\in\overline{\gamma}_{m(i)h}} \Phi_j^2(x) &\leq c\rho_i h_i^{-1} \parallel \rho_j\chi_i \parallel^2_{L^2(\gamma_{m(i)})}\\ &\leq c\rho_i\frac{H_i}{h_i}.\end{aligned} \quad (50)$$

Note that

$$\rho_i(\rho_j^2\chi_i^2(x)) = \rho_j(\rho_i\rho_j\chi_i^2(x)) \leq \rho_j, \quad \forall x \in \overline{\gamma}_{m(i)},$$

is true since $\rho_i\rho_j\chi_i^2(x) \leq 1$ for all $x \in \overline{\gamma}_{m(i)}$, and hence it is possible to replace ρ_i with ρ_j in the estimate (50).

Let $\delta_{m(i)}$ and $\delta_{n(i)}$ intersect at $x_a^{(i)}$. Note that $\rho_j\chi_i(x) \leq 1$ for all $x \in \overline{\delta}_{m(i)}$, $\rho_j\chi_j(x) \leq 1$ for all $x \in \overline{\gamma}_{m(j)}$, and $\rho_j\chi_k(x_a^{(k)}) \leq 1$. Combining (48)-(49), and using the triangle inequality, we obtain

$$\begin{aligned}
\rho_i \sum_{x\in\partial\Omega_{ih}} \Phi_j^2(x) &= \rho_i \sum_{x\in\overline{\delta}_{m(i)h}} \Phi_j^2(x) + \rho_i \sum_{x\in\delta_{n(i)h}} \Phi_j^2(x) \\
&\leq c\rho_i h_i^{-1} \Big\{ \rho_j^2\chi_i^2(x_a^{(i)}) \parallel \Pi_m(0,\varphi_a^{(i)}) \parallel^2_{L^2(\delta_{m(i)})} \\
&\qquad + \rho_j^2\chi_i^2(x_b^{(i)}) \parallel \Pi_m(0,\varphi_b^{(i)}) \parallel^2_{L^2(\delta_{m(i)})} \\
&\qquad + \parallel \rho_j\Pi_m(\chi_j,0) \parallel^2_{L^2(\delta_{m(i)})} \\
&\qquad + \rho_j^2\chi_i^2(x_a^{(i)}) \parallel \Pi_n(0,\varphi_a^{(i)}) \parallel^2_{L^2(\delta_{n(i)})} \Big\} \\
&\leq c\rho_i h_i^{-1}\{h_i + h_i + H_i + h_i + h_i\} \\
&\leq c\rho_i \frac{H_i}{h_i}.
\end{aligned} \tag{51}$$

Here we have used, for the first, second and fourth terms, the fact that $\parallel \Pi_n(0,\varphi_a^{(i)}) \parallel^2_{L^2(\delta_{n(i)})} \leq ch_i$, which follows from Lemma 4 in [8]. And, finally, for the third term, we have used the fact that

$$\parallel \Pi_m(\chi_j,0) \parallel_{L^2(\delta_{m(i)})} \leq c \parallel \chi_j \parallel_{L^2(\gamma_{m(j)})},$$

which follows from the L^2-stability of mortar projections (Lemma 1 in [10]). Again, we note that

$$\rho_i(\rho_j^2\chi_k^2(x_a^{(k)})) = \rho_j(\rho_i\rho_j\chi_k^2(x_a^{(k)})) \leq \rho_j$$

since $\rho_i\rho_j\chi_k^2(x_a^{(k)}) \leq 1$, and that

$$\rho_i(\rho_j^2\chi_i^2(x)) = \rho_j(\rho_i\rho_j\chi_i^2(x)) \leq \rho_j, \quad \forall x \in \overline{\delta}_{m(i)},$$

and

$$\rho_i(\rho_j^2\chi_j^2(x)) = \rho_j(\rho_i\rho_j\chi_j^2(x)) \leq \rho_j, \quad \forall x \in \overline{\gamma}_{m(j)}$$

since $\rho_i\rho_j\chi_i^2(x) \leq 1$ on $\overline{\delta}_{m(i)}$ and $\rho_i\rho_j\chi_j^2(x) \leq 1$ on $\overline{\gamma}_{m(j)}$ ($\gamma_{m(j)} = \delta_{m(i)}$). Hence, it is possible to replace ρ_i with ρ_j in the estimate (51).

Finally, the case when $x_a^{(i)}$ is the point of intersection between $\gamma_{m(i)}$ and $\delta_{n(i)}$, or $\delta_{m(i)}$ and $\gamma_{n(i)}$, can be treated using a combination of the results from the first two cases. This will yield only similar estimates.

We now estimate the sum in (46) for $j \in Q_i$. Let $x_a^{(k)}$ be a vertex of Ω_k such that the two edges of Ω_k, Γ_{jk} and Γ_{ki}, intersect at $x_a^{(k)}$, and that Γ_{ki} is a mortar ($\gamma_{n(k)}$) in Ω_k. Let $\delta_{n(i)}$ be the corresponding nonmortar in Ω_i. We use (33) to represent Φ_j on $\delta_{n(i)}$ as

$$\Phi_j = \rho_j \chi_k(x_a^{(k)}) \Pi_n(\varphi_a^{(k)}, 0). \tag{52}$$

Note that $\rho_j \chi_k(x_a^{(k)}) \le 1$. By Lemma 5 in [8], we thus have

$$\begin{aligned}
\rho_i \sum_{x \in \partial\Omega_{ih}} \Phi_j^2(x) &= \rho_i \sum_{x \in \delta_{n(i)h}} \Phi_j^2(x) \\
&\le c\rho_i h_i^{-1} \rho_j^2 \chi_k^2(x_a^{(k)}) \parallel \Pi_n(\varphi_a^{(k)}, 0) \parallel^2_{L^2(\delta_{n(i)})} \\
&\le c\rho_i h_i^{-1} h_i \\
&\le c\rho_i.
\end{aligned} \tag{53}$$

In the same way as before, since $\rho_i \rho_j \chi_k^2(x_a^{(k)}) \le 1$, we can easily write $\rho_i \rho_j^2 \chi_k^2(x_a^{(k)}) \le \rho_j$, which allows us to replace ρ_i in (53) with ρ_j. Note that we also need to be able to replace ρ_i in (53) with ρ_k, see (45). This is possible since $\gamma_{n(k)} = \Gamma_{ki}$ is a mortar in Ω_k, which implies that $\rho_i \le \rho_k$.

Since the subdomains are shape regular, the maximal number of neighbors in $G_i \cup Q_i$ can be bounded independently of the total number of subdomains, N. Now, combining (39)-(53), for $\Omega_i \in N_I$, we obtain

$$a_i(u^{(0)}, u^{(0)}) \le c\frac{H}{h} \left\{ a_i(u_i, u_i) + \sum_{j \in G_i \cup Q_i} a_j(u_j, u_j) \right\}. \tag{54}$$

The above estimate applies also to the case when Ω_i is a boundary subdomain ($\Omega_i \in N_B$), which we show now. We start by introducing an auxiliary function $\hat{u}^{(0)} = \sum_{\Omega_i \in N_I} \overline{u}_i \Phi_i + \sum_{\Omega_i \in N_B} \overline{u}_i \hat{\Phi}_i$. The function $\hat{\Phi}_i$ for $\Omega_i \in N_B$ has been defined before, see on page 14. On $\overline{\Omega}_i$, where $\Omega_i \in N_B$, the function $\hat{u}^{(0)}$ takes the form

$$\hat{u}^{(0)} = \overline{u}_i \Phi_i + \sum_{j \in G_i \cup Q_i,\ \Omega_j \in N_I} \overline{u}_j \Phi_j + \sum_{j \in G_i \cup Q_i,\ \Omega_j \in N_B} \overline{u}_j \hat{\Phi}_j.$$

By the triangle inequality, we find

$$a_i(u^{(0)}, u^{(0)}) \le 2 \left\{ a_i(u^{(0)} - \hat{u}^{(0)}, u^{(0)} - \hat{u}^{(0)}) \ + \ a_i(\hat{u}^{(0)}, \hat{u}^{(0)}) \right\}. \tag{55}$$

The function $\hat{u}^{(0)}$ on $\overline{\Omega}_i$ has the same form as $u^{(0)}$ on an interior subdomain. The second term has, therefore, the same bound as in (54). For the first term we note that

$$(u^{(0)} - \hat{u}^{(0)}) = \overline{u}_i(\Phi_i - \hat{\Phi}_i) + \sum_{j \in G_i' \cup Q_i'} \overline{u}_j(\Phi_j - \hat{\Phi}_j)$$

on $\overline{\Omega}_i$. Using this relation and the triangle inequality we have

$$a_i(u^{(0)} - \hat{u}^{(0)}, u^{(0)} - \hat{u}^{(0)}) = \rho_i |u^{(0)} - \hat{u}^{(0)}|^2_{H^1(\Omega_i)} \leq c \Big\{ \rho_i \overline{u}_i^2 |\Phi_i - \hat{\Phi}_i|^2_{H^1(\Omega_i)} + \sum_{j \in G_i' \cup Q_i'} \rho_i \overline{u}_j^2 |\Phi_j - \hat{\Phi}_j|^2_{H^1(\Omega_i)} \Big\}. \quad (56)$$

We estimate the right hand side in two steps again.

(i) *Estimation of* $\overline{u}_i^2$ *and* $\overline{u}_j^2$ *in (56).* Both Ω_i and Ω_j here are boundary subdomains. We derive estimates for only $\overline{u}_i^2$ since, for $\overline{u}_j^2$, the estimates are then straightforward.

Ω_i can either be in N_{BE} or in N_{BV}. For $\Omega_i \in N_{BE}$, we use a simple trace inequality and Friedrichs inequality (u_i is zero on a boundary edge of Ω_i) to get

$$\overline{u}_i^2 = \left\{ \frac{1}{|\partial\Omega_i|} \int_{\partial\Omega_i} u_i \, ds \right\}^2 \leq c H_i^{-1} \parallel u_i \parallel^2_{L^2(\partial\Omega_i)} \leq c |u_i|^2_{H^1(\Omega_i)}. \quad (57)$$

For $\Omega_i \in N_{BV}$, we get an extra $\frac{H_i}{h_i}$ factor in the estimate. Note that u_i is zero at least at one point on $\partial\Omega_i$, and hence, for any constant α, $|\alpha| \leq \max_{x \in \partial\Omega_i} |u_i(x) - \alpha|$. Using this, a simple trace inequality and the Poincaré inequality (choose $\alpha = \frac{1}{|\Omega_i|} \int_{\Omega_i} u_i \, dx$), we obtain

$$\overline{u}_i^2 \leq 4 \max_{x \in \partial\Omega_i} |u_i(x) - \alpha|^2 \leq c h_i^{-1} \parallel u_i - \alpha \parallel^2_{L^2(\partial\Omega_i)} \leq c \frac{H_i}{h_i} |u_i|^2_{H^1(\Omega_i)}. \quad (58)$$

(2) *Estimation of the two seminorms in (56).* We estimate the first and the second seminorms separately. By the inverse inequality and the discrete L^2-norm, for the first seminorm, we have

$$\begin{aligned} \rho_i |\Phi_i - \hat{\Phi}_i|^2_{H^1(\Omega_i)} &\leq c\rho_i \sum_{K \subset \overline{\Omega}_i^B} h_K^{-2} \parallel \Phi_i - \hat{\Phi}_i \parallel^2_{L^2(K)} \\ &\leq c\rho_i \sum_{x \in \partial\Omega_{ih}} (\Phi_i(x) - \hat{\Phi}_i(x))^2 \\ &= c \Big\{ \rho_i \sum_{x \in \partial\Omega_{ih} \cap \partial\Omega_h} \hat{\Phi}_i^2(x) \\ &\quad + \sum_{\delta_{m(i)} \subset \partial\Omega_i} \rho_i \sum_{x \in \delta_{m(i)h}} (\Phi_i(x) - \hat{\Phi}_i(x))^2 \Big\}. \end{aligned} \quad (59)$$

The outer sum in the second term above is taken over those nonmortar sides of Ω_i, that touch the boundary $\partial\Omega$ at most at two points. We have used the fact that $\Phi_i(x) = \hat{\Phi}_i(x)$ at all $x \in \Omega_{ih}$, $x \in \overline{\gamma}_{m(i)h} \setminus \partial\Omega_h$ and $x \in \overline{\Gamma}_{ij}$ where Ω_j is an interior subdomain (i.e., $\Omega_j \in N_I$).

For simplicity, we assume that $\delta_{m(i)}$ in (59) touches $\partial\Omega$ at only one point, since the estimates for the case where $\delta_{m(i)}$ touches the boundary at two points will simply contribute a factor 2 to the constants in the estimates.

For $\Omega_i \in N_{BE}$ and $\Omega_i \in N_{BV}$, the first term in (59) can be bounded by $c\rho_i \frac{H_i}{h_i}$ and $c\rho_i$, respectively. In either case, for the second term, we use (34), the triangle inequality, and Lemma 4 and Lemma 5 in [8], to get

$$\begin{aligned} \rho_i \textstyle\sum_{x\in\delta_{m(i)h}} & (\Phi_i(x) - \hat{\Phi}_i(x))^2 \\ \leq \quad & c\rho_i h_i^{-1} \Big\{ \rho_i^2 \chi_i^2(x_a^{(i)}) \parallel \Pi_m(0, \varphi_a^{(i)}) \parallel^2_{L^2(\delta_{m(i)})} \\ & \qquad + \rho_i^2 \chi_j^2(x_a^{(j)}) \parallel \Pi_m(\varphi_a^{(j)}, 0) \parallel^2_{L^2(\delta_{m(i)})} \Big\} \\ \leq \quad & c\rho_i h_i^{-1} \{h_i + h_i\} \\ \leq \quad & c\rho_i. \end{aligned} \tag{60}$$

Here $x_a^{(i)} \in \nu_i \cap \partial\Omega_h$ and $x_a^{(j)} \in \nu_j \cap \partial\Omega_h$ (geometrically occupying the same space) are the end points of $\delta_{m(i)}$ and $\gamma_{m(j)}$, respectively. We have used the fact that $\rho_i\chi_i(x_a^{(i)}) \leq 1$ and $\rho_i\chi_j(x_a^{(j)}) \leq 1$.

For the second seminorm in (56), we derive estimates for only $\Omega_i \in N_{BV}$ since, for $\Omega_i \in N_{BE}$, it is straightforward to see that the same estimates hold. By the inverse inequality and the discrete L^2-norm, we have

$$\begin{aligned} \rho_i |\Phi_j - \hat{\Phi}_j|^2_{H^1(\Omega_i)} &\leq c\rho_i \sum_{K\subset\overline{\Omega}_i^B} \parallel \Phi_j - \hat{\Phi}_j \parallel^2_{L^2(K)} \\ &\leq c\rho_i \sum_{x\in\partial\Omega_{ih}} (\Phi_j(x) - \hat{\Phi}_j(x))^2 \end{aligned} \tag{61}$$

since $\Phi_j(x) - \hat{\Phi}_j(x) = 0$ at all $x \in \Omega_{ih}$. Again, we estimate the sum for $j \in G_i'$ and $j \in Q_i'$ separately.

Consider first $j \in G_i'$. Let $x_a^{(i)} \in \nu_i \cap \partial\Omega_h$ and $x_a^{(j)} \in \nu_j \cap \partial\Omega_h$ (geometrically occupying the same space) be the end points of $\delta_{m(i)}$ and $\gamma_{m(j)}$, respectively, such that $\delta_{m(i)} = \gamma_{m(j)}$, or the end points of $\gamma_{m(i)}$ and $\delta_{m(j)}$, respectively, such that $\gamma_{m(i)} = \delta_{m(j)}$. We assume again that they touch $\partial\Omega$ only at one point. Let $\delta_{n(i)}$ or $\gamma_{n(i)}$ be the other edge of Ω_i that intersects $\delta_{m(i)}$ or $\gamma_{m(i)}$ at $x_a^{(i)}$. As before, there are four cases we need to investigate depending on whether the two edges of Ω_i are mortar or nonmortar.

As the first case, let $x_a^{(i)} \in \nu_i \cap \partial\Omega_h$ be the point of intersection between the two mortars $\gamma_{m(i)}$ and $\gamma_{n(i)}$. $\hat{\Phi}_j(x) - \Phi_j(x) = 0$ at all $x \in \overline{\Omega}_{ih}$ except at $x_a^{(i)}$, where it is equal to $\hat{\Phi}_j(x_a^{(i)}) = \rho_j \chi_i(x_a^{(i)})$. Hence

$$\rho_i \sum_{x \in \partial\Omega_{ih}} (\Phi_j(x) - \hat{\Phi}_j(x))^2 = \rho_i \rho_j^2 \chi_i^2(x_a^{(i)}) \leq \rho_j \tag{62}$$

using the fact that $\rho_i \rho_j \chi_i^2(x_a^{(i)}) \leq 1$.

As the second case, let $x_a^{(i)} \in \nu_i \cap \partial\Omega_h$ be the point of intersection between the two nonmortars $\delta_{m(i)}$ and $\delta_{n(i)}$. $\hat{\Phi}_j(x) - \Phi_j(x) = 0$ at all $x \in \overline{\Omega}_{ih}$ except at $x \in (\delta_{m(i)h} \cup \delta_{n(i)h})$ and $x_a^{(i)}$. At $x \in \delta_{m(i)h}$, following (35), $\hat{\Phi}_j - \Phi_j$ can be written as

$$\hat{\Phi}_j - \Phi_j = \rho_j \chi_j(x_a^{(j)}) \Pi_m(\varphi_a^{(j)}, 0) + \rho_j \chi_i(x_a^{(i)}) \Pi_m(0, \varphi_a^{(i)}),$$

and at $x \in \delta_{n(i)h}$, $\hat{\Phi}_j - \Phi_j$, by definition (cf. (32)), has the form

$$\hat{\Phi}_j - \Phi_j = \hat{\Phi}_j = \rho_j \chi_i(x_a^{(i)}) \Pi_n(0, \varphi_a^{(i)}).$$

It follows from the triangle inequality, Lemma 4 and Lemma 5 in [8], that

$$\begin{aligned}
&\rho_i \textstyle\sum_{x \in \partial\Omega_{ih}} (\Phi_j(x) - \hat{\Phi}_j(x))^2 \\
&= \quad \rho_i \rho_j^2 \chi_i^2(x_a^{(i)}) + \rho_i \left\{ \sum_{x \in \delta_{m(i)h} \cup \delta_{n(i)h}} (\Phi_j(x) - \hat{\Phi}_j(x))^2 \right\} \\
&\leq \quad \rho_j + c\rho_i h_i^{-1} \Big\{ \rho_j^2 \chi_j^2(x_a^{(j)}) \parallel \Pi_m(\varphi_a^{(j)}, 0) \parallel^2_{L^2(\delta_{m(i)})} \\
&\qquad + \rho_j^2 \chi_i^2(x_a^{(i)}) \parallel \Pi_m(0, \varphi_a^{(i)}) \parallel^2_{L^2(\delta_{m(i)})} \\
&\qquad + \rho_j^2 \chi_i^2(x_a^{(i)}) \parallel \Pi_n(0, \varphi_a^{(i)}) \parallel^2_{L^2(\delta_{n(i)})} \Big\} \\
&\leq \quad \rho_j + c\rho_j h_i^{-1} \{h_i + h_i + h_i\} \\
&\leq \quad c\rho_j .
\end{aligned} \tag{63}$$

We have used that $\rho_i \rho_j \chi_j^2(x_a^{(j)}) \leq 1$ and $\rho_i \rho_j \chi_i^2(x_a^{(i)}) \leq 1$.

For the other case, where $x_a^{(i)} \in \nu_i \cap \partial\Omega_h$ is a point of intersection between either $\gamma_{m(i)}$ and $\delta_{n(i)}$, or $\delta_{m(i)}$ and $\gamma_{n(i)}$, similar estimates can be derived by simply combining the results from the first two cases.

We now estimate the sum in (61) for $j \in Q_i'$. Let $x_a^{(k)} \in \nu_k \cap \partial\Omega$ be the point of intersection between the two edges Γ_{jk} and Γ_{ki} of Ω_k, for some k, such that $\Gamma_{ki} = \delta_{n(i)} = \gamma_{n(k)}$. $\hat{\Phi}_j(x) - \Phi_j(x) = 0$ at all $x \in \overline{\Omega}_{ih}$ except at

$x \in \delta_{n(i)h}$. At $x \in \delta_{n(i)h}$, $\hat{\Phi}_j(x) - \Phi_j(x)$, by definition (cf. (33)), has the form

$$\hat{\Phi}_j - \Phi_j = \hat{\Phi}_j = \rho_j \chi_k(x_a^{(k)}) \Pi_n(\varphi_a^{(k)}, 0).$$

By the inverse inequality, the discrete L^2-norm, Lemma 5 in [8] and the fact that $\rho_i \rho_j \chi_k^2(x_a^{(k)}) \leq 1$, we then obtain

$$\begin{aligned}
\rho_i \sum_{x \in \partial\Omega_{ih}} (\Phi_j(x) - \hat{\Phi}_j(x))^2 &= \rho_i \sum_{x \in \delta_{n(i)h}} (\Phi_j(x) - \hat{\Phi}_j(x))^2 \\
&\leq c\rho_i h_i^{-1} \rho_j^2 \chi_k^2(x_a^{(k)}) \parallel \Pi_n(\varphi_a^{(k)}, 0) \parallel^2_{L^2(\delta_{n(i)})} \\
&\leq c\rho_j h_i^{-1} h_i \\
&= c\rho_j. \qquad (64)
\end{aligned}$$

Now, combining the inequalities (56)-(64), and replacing the resulting estimate together with the estimate (the same as in (54)) for $a_i(\hat{u}^{0)}, \hat{u}^{0)})$ in (55), we obtain the same estimate as (54). The proof of (37) now follows by summing (54) over all subdomains.

Derivation of the inequality (38). The second inequality of the lemma follows from the proof of the first part. Let $\Omega_i \in N_I$. We then have

$$\rho_i \parallel u^{(0)} - \overline{u}_i \parallel^2_{L^2(\Omega_i)} \leq c \sum_{j \in G_i \cup Q_i} \rho_i (\overline{u}_j - \overline{u}_i)^2 \parallel \Phi_j \parallel^2_{L^2(\Omega_i)},$$

where we have replaced the H^1-seminorm with the L^2-norm in (39). The estimate for $\parallel \Phi_j \parallel^2_{L^2(\Omega_i)}$ will gain a factor h_i^2 compared with $|\Phi_j|^2_{H^1(\Omega_i)}$, see (50), (51) and (53). This results in a factor $h_i H$ instead of $\frac{H}{h_i}$. If we now follow the same steps as in the proof of the first part of the lemma we obtain

$$\rho_i \parallel u^{(0)} - \overline{u}_i \parallel^2_{L^2(\Omega_i)} \leq c h_i H \left\{ a_i(u_i, u_i) + \sum_{j \in G_i \cup Q_i} a_j(u_j, u_j) \right\}. \qquad (65)$$

For $\Omega_i \in N_B$ we use exactly the same arguments, which again leads to the factor $h_i H$ instead of $\frac{H}{h_i}$, and follow the same steps as in the proof of the first part of the lemma for $\Omega_i \in N_B$ to get the same estimate as (65). The proof of (38) thus follows by summing (65) over all subdomains. The proof of Lemma 1 is now complete.

We are now ready to prove our theorem.

Proof (Theorem 2) We use the general Schwarz framework to prove the theorem, and, as before, we have three key assumptions to verify.

Assumption 2. $\rho(\mathcal{E}) = 1$ since the subspaces $V^{(0)}$ and $V^{(-1)}$ are the two coarse spaces, see [16].

Assumption 3. Exact bilinear forms are used for all subproblems, hence, $\omega = 1$ for $i = -1, \cdots, N$.

Assumption 1. For $u = \{u_i\}_{i=1}^N \in V^h$, we define $u^{(0)}$ as

$$u^{(0)} = \sum_{i=1,\cdots,N} \overline{u}_i \Phi_i,$$

where $\overline{u}_i = \frac{1}{|\partial\Omega_i|} \int_{\partial\Omega_i} u_i \, ds$ and $\Phi_i, i = 1, \cdots, N$ are the basis functions of $V^{(0)}$. Let $w = u - u^{(0)} = \{w_i\}_{i=1}^N$. Let $u^{(-1)} = \{u_i^{(-1)}\}_{i=1}^N \in V^{(-1)}$ be defined on each $\overline{\Omega}_i$ by the nodal values as follows

$$u_i^{(-1)}(x) = \begin{cases} w_i(x), \, x \in \partial\Omega_{ih}, \\ \quad 0, \quad\; x \in \Omega_{ih}. \end{cases}$$

Similarly, let $u^{(j)} = \{u_i^{(j)}\}_{i=1}^N \in V^{(j)}, j = 1, \cdots, N$ be defined by the nodal values as

$$u_i^{(i)} = \begin{cases} w_i(x), \, x \in \Omega_{ih}, \\ \quad 0, \quad\; x \in (\overline{\Omega}_h \setminus \Omega_{ih}), \end{cases}$$

with $u_i^{(j)} = 0$ whenever $j \neq i$. Clearly, $u^{(i)} \in V^{(i)}$ and $u = \sum_{i=-1}^N u^{(i)}$.

Using the above splitting of u, we need to show that

$$\sum_{i=-1}^{N} b^{(i)}(u^{(i)}, u^{(i)}) \leq c \frac{H}{h} a(u, u). \tag{66}$$

For $u^{(-1)} = \{u_i^{(-1)}\}_{i=1}^N \in V^{(-1)}$, we have

$$b^{(-1)}(u^{(-1)}, u^{(-1)}) = \sum_{i=1}^{N} a_i(u_i^{(-1)}, u_i^{(-1)}). \tag{67}$$

By the inverse inequality, the discrete L^2-norm, the triangle inequality and the definition of $u_i^{(-1)}$, we get

$$
\begin{aligned}
a_i(u_i^{(-1)}, u_i^{(-1)}) &= \rho_i |u_i^{(-1)}|^2_{H^1(\Omega_i)} \\
&\leq c\rho_i \sum_{K \subset \overline{\Omega}_i^B} h_K^{-2} \parallel u_i^{(-1)} \parallel^2_{L^2(K)} \\
&\leq c\rho_i \sum_{x \in \partial\Omega_{ih}} w_i^2(x) \\
&\leq c\rho_i h_i^{-1} \parallel u_i - u^{(0)} \parallel^2_{L^2(\partial\Omega_i)} \\
&\leq c\rho_i h_i^{-1} \Big\{ \parallel u_i - \overline{u}_i \parallel^2_{L^2(\partial\Omega_i)} \\
&\qquad + \parallel u^{(0)} - \overline{u}_i \parallel^2_{L^2(\partial\Omega_i)} \Big\}. \qquad (68)
\end{aligned}
$$

For the first term inside the curly brackets we have

$$\parallel u_i - \overline{u}_i \parallel^2_{L^2(\partial\Omega_i)} \leq cH_i |u_i|^2_{H^1(\Omega_i)}, \qquad (69)$$

which follows from (42). For the second term, we note, by the definition of $u^{(0)}$, that $u^{(0)}(x) = \overline{u}_i$ at every $x \in \Omega_{ih}$, and hence

$$
\begin{aligned}
\parallel u^{(0)} - \overline{u}_i \parallel^2_{L^2(\partial\Omega_i)} &\leq ch_i \sum_{x \in \partial\Omega_{ih}} (u^{(0)}(x) - \overline{u}_i)^2 \\
&\leq ch_i \sum_{K \subset \overline{\Omega}_i^B} h_K^{-2} \parallel u^{(0)} - \overline{u}_i \parallel^2_{L^2(K)} \\
&\leq ch_i^{-1} \parallel u^{(0)} - \overline{u}_i \parallel^2_{L^2(\Omega_i)} . \qquad (70)
\end{aligned}
$$

Now replacing (69) and (70) in (68), adding over $i = 1, \cdots, N$, and finally using Lemma 1 on the second term, we obtain

$$b^{(-1)}(u^{(-1)}, u^{(-1)}) \leq c\frac{H}{h} a(u,u). \qquad (71)$$

From Lemma 1 it follows immediately that

$$b^{(0)}(u^{(0)}, u^{(0)}) \leq c\frac{H}{h} a(u,u). \qquad (72)$$

Finally,

$$
\begin{aligned}
\sum_{i=1}^N b^{(i)}(u^{(i)}, u^{(i)}) &= \sum_{i=1}^N a_i(u_i^{(i)}, u_i^{(i)}) \\
&\leq c \sum_{i=1}^N \Big\{ a_i(u,u) + a_i(u_i^{(0)}, u_i^{(0)}) + a_i(u_i^{(-1)}, u_i^{(-1)}) \Big\} \\
&= c \Big\{ a(u,u) + b^{(0)}(u^{(0)}, u^{(0)}) + b^{(-1)}(u^{(-1)}, u^{(-1)}) \Big\} \\
&\leq c\frac{H}{h} a(u,u). \qquad (73)
\end{aligned}
$$

The inequality (66) now follows from (71), (72) and (73). This concludes the proof of Theorem 2.

5 Implementation issues

The subproblems associated with the subspaces $V^{(i)}, i = 1, \cdots, N$, are the standard local subproblems, it is well known how to implement their solvers. In this section we briefly discuss the coarse problems only.

The basis functions $\{\phi_k\}$ of the finite element space V^h are associated with the subdomain interior nodes Ω_{ih}, the mortar nodes $\gamma_{m(i)h}$ ($\gamma_{m(i)} \subset \partial\Omega_i$) and the subdomain vertices $\nu_i, i = 1, \cdots, N$, except those on $\partial\Omega$, see Section 2 for a description. Let $\mathbf{u}$ be the vector representation of $u \in V^h$ with respect to the basis of V^h, which contains the nodal values of u. The matrix $\mathbf{A}$ of the discrete problem can be split as follows. Let λ_i be the union of Ω_{ih}, $\gamma_{m(i)h}$ where $\gamma_{m(i)} \subset \partial\Omega_i$, ν_i and $\overline{\gamma}_{m(j)h}$ where $\gamma_{m(j)} = \delta_{m(i)} \subset \partial\Omega_i$, except those nodes on $\partial\Omega$. Let $\mathbf{R}_i$ be the restriction matrix containing only ones and zeros, which, if multiplied with $\mathbf{u}$ from the left, will return a vector of length $|\lambda_i|$ containing the nodal values of $\mathbf{u}$ only at the nodes of λ_i. Let $\mathbf{R}_i^T$ be the corresponding extension operator.

Now let $\mathbf{A}_i$ be the local stiffness matrix associated with the subdomain Ω_i, whose rows and columns correspond to the nodes of the set λ_i. Using the above three types of matrices we can split the global stiffness matrix as $\mathbf{A} = \sum_{i=1}^N \mathbf{R}_i^T \mathbf{A}_i \mathbf{R}_i$, where $\mathbf{A}_i = \{a_i(\phi_j, \phi_l)\}$ with $x_j, x_l \in \lambda_i$. Note that $\mathbf{A}_i, i = 1, \cdots, N$ can be constructed locally by the processor (virtual processor) responsible for Ω_i. We use these local matrices to construct our coarse stiffness matrices.

The method with the coarse reformulation

Consider first the method with the coarse reformulation. Let $V_{cr}^{(-1)}$ and $V_{cr}^{(0)}$ be the two coarse spaces as defined in the previous section. We use the mortar basis of V^h to represent the basis function Φ_i of $V_{cr}^{(0)}$ (cf. (27)) as $\Phi_i = \sum_k \Phi_i(x_k)\phi_k$, where $\Phi_i(x_k)$ is the nodal value of Φ_i at a node x_k. By the definition of Φ_i, the nodal value at x_k will be

$$\Phi_i(x_k) = \begin{cases} 1, & x_k \in \Omega_{ih}, \\ \rho_i \chi_i(x_k), & x_k \in \gamma_{m(i)h} \cup \nu_i, \\ \rho_i \chi_j(x_k), & x_k \in \overline{\gamma}_{m(j)h},\ \gamma_{m(j)} = \delta_{m(i)}, \\ \rho_i \chi_j(x_k), & x_k \in \partial\delta_{m(j)h},\ \delta_{m(j)} = \gamma_{m(i)}, \\ 0, & x_k \in \partial\Omega_h, \\ 0, & \text{otherwise}. \end{cases} \tag{74}$$

Let $\mathbf{v}_i$ be the vector representation of Φ_i with respect to the basis of V^h. Let $\mathbf{A}_{cr}^{(0)}$ be the corresponding stiffness matrix of size $N \times N$, then the element $(\mathbf{A}_{cr}^{(0)})_{jl}$ is given by $a(\Phi_j, \Phi_l) = \mathbf{v}_j^T \mathbf{A} \mathbf{v}_l = \sum_{i=1}^{N} (\mathbf{R}_i \mathbf{v}_j)^T \mathbf{A}_i (\mathbf{R}_i \mathbf{v}_l)$. Note that the term inside the summation is nonzero when both j and l belong to $G_i \cup Q_i \cup \{i\}$ as both Φ_j and Φ_l in that case will have nonzero support on Ω_i. The matrix $\mathbf{A}_{cr}^{(0)}$ can thus be assembled in the standard way from its element matrices $\mathbf{A}_i^{(0)}$, $i = 1, \cdots, N$, where $\mathbf{A}_i^{(0)}$ contains elements from the set $\{(\mathbf{R}_i \mathbf{v}_j)^T \mathbf{A}_i (\mathbf{R}_i \mathbf{v}_l)\}_{j,l}$, where $j, l \in G_i \cup Q_i \cup \{i\}$. Note here that the vectors $\mathbf{v}_j$ and $\mathbf{v}_l$ both have constant values as their elements corresponding to the interior nodes Ω_{ih}, i.e., one if the indices are equal to i and zero otherwise. This information can be used to calculate the elements cheaply, for instance, by treating the whole set of nodes Ω_{ih} as a single node in the calculation.

Now, for the space $V_{cr}^{(-1)}$, we note that it is simply $span\{\phi_l : x_l \in S_h\}$, where S_h is the set of all mortar nodes and the vertices of subdomains $\Omega_i, i = 1, \cdots, N$, except those on $\partial\Omega$, see (26). The construction of the corresponding stiffness matrix $\mathbf{A}_{cr}^{(-1)}$ is quite straightforward, and can be extracted from the small matrices $\mathbf{A}_i$ without any extra cost.

The average method

Consider now the average method. Let $V_{avg}^{(0)}$ be its coarse space. $V_{avg}^{(0)}$ can be expressed using a set of some basis functions, which we denote as $\{\Psi_l\}$. Note that the dimension of this coarse space is the same as that of $V_{cr}^{(-1)}$ above. We define the basis functions as follows: Associate each function Ψ_l to a node x_l from the set S_h by setting the function value to 1 at x_l and 0 at all other nodes in S_h, then, by the definition of $V_{avg}^{(0)}$, see (8)-(11), the nodal value at any node x_k is given as

$$\Psi_l(x_k) = \begin{cases} 1, & x_k = x_l, \\ 0, & x_k \in (S_h \setminus \{x_l\}), \\ \frac{1}{|\nu_i|} a_m, & x_k \in \Omega_{ih},\ x_l \in \gamma_{m(i)h}, \\ \frac{1}{|\nu_j|} a_m, & x_k \in \Omega_{jh},\ x_l \in \overline{\gamma}_{m(i)h},\ \gamma_{m(i)} = \delta_{m(j)}, \\ 0, & x_k \in \Omega_{ih},\ x_l = \partial\delta_{m(i)} \cap \partial\delta_{n(i)}, \\ \frac{1}{|\nu_i|} a_m, & x_k \in \Omega_{ih},\ x_l = \partial\gamma_{m(i)} \cap \partial\delta_{n(i)}, \\ \frac{1}{|\nu_i|}(a_m + a_n), & x_k \in \Omega_{ih},\ x_l = \partial\gamma_{m(i)} \cap \partial\gamma_{n(i)}, \\ 0, & \text{otherwise}. \end{cases} \tag{75}$$

Here a_m is the average over the edge $\gamma_{m(i)}$, see (11), of a function in $W^{h_i}(\gamma_{m(i)})$ having one as its value at the node $x_l \in \overline{\gamma}_{m(i)h}$ and zeros at the other nodes. Similarly, a_n defines the average over the edge $\gamma_{n(i)}$.

Ψ_l, in terms of the mortar basis, can be written as $\Psi_l = \sum_k \Psi_l(x_k)\phi_k$. For instance, for $x_l \in \gamma_{m(i)h}$, where $\gamma_{m(i)} \subset \partial\Omega_i$,

$$\Psi_l(x) = \phi_l(x) + \frac{1}{|\nu_i|} a_m \sum_{x_k \in \Omega_{ih}} \varphi_k(x) + \frac{1}{|\nu_j|} a_m \sum_{x_k \in \Omega_{jh}} \varphi_k(x).$$

We thus have $V_{avg}^{(0)} = span\{\Psi_l : x_l \in S_h\}$. The construction of the corresponding stiffness matrix $\mathbf{A}_{avg}^{(0)}$ is now similar to that of $\mathbf{A}_{cr}^{(0)}$.

Since the coarse spaces do not need any coarse grid discretization, the methods are easily applicable to unstructured grids. The method with the coarse reformulation, however, has some computational advantages over the average method. We remark that $\mathbf{A}_{avg}^{(0)}$ has a much larger stencil than $\mathbf{A}_{cr}^{(-1)}$, even though they have the same dimension, which makes $\mathbf{A}_{avg}^{(0)}$ a much denser matrix than $\mathbf{A}_{cr}^{(-1)}$. By a careful ordering of the nodes, it is possible to solve the $\mathbf{A}_{cr}^{(-1)}$ coarse problem by first eliminating the unknowns associated with all mortar nodes, which is done by solving a sequence of tridiagonal systems, and then by solving a Schur complement system on the subdomain vertices. In addition, $\mathbf{A}_{cr}^{(0)}$ has a minimal dimension. Solving the two coarse problems of the method with the coarse reformulation is therefore computationally cheaper than solving the coarse problem of the average method. Note, in addition, that the two coarse problems can be solved in parallel.

6 Numerical results

In this section, we present results from the numerical experiments carried out in Matlab using the two variants of the additive Schwarz methods, the average method and the method with the coarse reformulation, introduced in this paper. We have seen in the previous sections that the methods use only the exact bilinear forms for their subproblems. Two variants of the average method which use different approximate bilinear forms for their coarse spaces can be found in [15]. For the present paper, we take only the best approximate variant and compare it with the exact variants. The CG method is used to solve the preconditioned system (16). The iteration stops when the residual norm is reduced by the factor 10^{-6}.

We choose our model problem to be defined on the unit square and without loss of generality, we let the solution have zero boundary values. The force function f is chosen to have the form $f(x) = 2\pi^2 \sin(\pi x_1)\sin(\pi x_2)$. We use the following choice of mortar and nonmortar sides: If $\rho_i \geq \rho_j$ then choose the mortar side to be in Ω_i, i.e., $\gamma_{m(i)} = \delta_{m(j)}$ ($\gamma_{m(i)} \subset \partial\Omega_i$ and $\delta_{m(j)} \subset \partial\Omega_j$).

Table 1 shows iteration counts and condition number estimates for different partitions of the domain using the two average methods, the approximate and the exact variant, as well as the method with the coarse reformulation. In all tests the subdomains are triangulated using, in a checkerboard order, either 50 or 72 right angle triangles of the same size, so that any two neighboring subdomains get a nonmatching grid across their interface and the ratio H/h remains fixed. For a fixed partition, the coefficients ρ_i are picked uniformly from the interval $[10^{-1}, 10^3]$ and then distributed randomly among the subdomains.

Table 1. Comparison of the three variants of the additive Schwarz methods. Iteration counts and condition number estimates (in parentheses) are shown for a fixed ratio H/h.

Subdomains	Average with approx. bilinear form	Average with exact bilinear form	Coarse reformulation
4×4	37 (20.80)	27 (11.41)	29 (13.46)
8×8	39 (21.13)	29 (11.62)	32 (13.96)
16×16	41 (21.08)	30 (11.55)	34 (14.41)

As we can see from the table, the exact variant of the average method has the best condition number estimates, while the corresponding estimates for the approximate variant are about twice as big. It should be noted that the size of their coarse space problems are the same, and the corresponding stiffness matrices have the same density pattern. So unless an efficient algorithm is found for the approximate version this will not be an attractive method. The exact version, however, with its very high computational complexity, cannot be considered practical. The method with the coarse reformulation, on the other hand, has condition number estimates which are close to those of the exact variant of the average method. If we take into account the computational complexity and the parallel properties of the algorithms we can conclude that the method with the coarse reformulation is an attractive method.

The condition number estimates, as seen from Table 1, remain close to constant for a fixed H/h suggesting that the condition number depends only on the ratio H/h. In Table 2, we see that the condition number estimates become approximately doubled as we double the ratio H/h by moving from left to right in the table, implying that the dependence is linear. The subdomain size H and the mesh size h in the first column of this table correspond to the same partition and discretization which are used for the second row of Table 1, i.e. $8 \times 8 = 64$ subdomains each having either 50 or 72 triangles. The coefficients ρ_i are chosen in the same way as before, but due to a different (random) distribution of the coefficients the condition number estimates here are different from the corresponding estimates in Table 1.

Table 2. Iteration counts and condition number estimates (in parentheses) illustrating the linear dependence of the condition number on the ratio H/h.

Method	Subdomain size, Mesh size		
	$\{H,h\}$	$\{H,h/2\}$	$\{2H,h/2\}$
Avg., exact	29 (11.58)	46 (28.06)	66 (62.12)
Coarse ref.	32 (14.40)	48 (30.84)	67 (65.80)

An important feature of the methods is that they are all insensitive to jumps in the coefficient ρ_i across the subdomain boundaries. This can be seen from Table 3, where a checkerboard distribution of the coefficients is used. As we move from one column to another, in this table, the jump is increased or decreased by the factor 10^2, but the condition number estimates remain unaffected. 64 subdomains, each containing either 50 or 72 triangles, are used for the tests.

Table 3. Iteration counts and condition number estimates (in parentheses) illustrating the insensitiveness of the condition number to the coefficient jump.

Method	Coefficients ρ_i, checkerboard distribution		
	$\{1,10^0\}$	$\{1,10^2\}$	$\{1,10^4\}$
Avg., exact	23 (10.37)	26 (11.70)	26 (11.73)
Coarse ref.	27 (13.88)	29 (14.19)	30 (14.24)

The numerical results of this section simply confirm our theoretical results. Further experiments with these methods have confirmed that the average methods, both the exact- and the approximate- variant, show convergence which is independent of the choice of the mortar and the nonmortar sides, while the method with the coarse reformulation requires that $\gamma_{m(i)} = \delta_{m(j)}$ if $\rho_i \geq \rho_j$.

References

1. Y. ACHDOU AND Y. A. KUZNETSOV, *Substructuring preconditioners for finite element methods on nonmatching grids*, East-West J. Numer. Math., 3 (1995), pp. 1–28.

2. Y. ACHDOU, Y. MADAY, AND O. B. WIDLUND, *Iterative substructuring preconditioners for mortar element methods in two dimensions*, SIAM J. Numer. Anal., 36 (1999), pp. 551–580.
3. C. BERNARDI, Y. MADAY, AND A. T. PATERA, *A new non conforming approach to domain decomposition: The mortar element method*, in Collège de France Seminar, H. Brezis and J.-L. Lions, eds., Pitman, 1994. Appeared as a technical report already in 1989.
4. P. E. BJØRSTAD AND M. DRYJA, *A coarse space formulation with good parallel properties for an additive schwarz domain decomposition algorithm*, Submitted to Numerische Mathematik, (1999).
5. P. E. BJØRSTAD, M. DRYJA, AND E. VAINIKKO, *Additive Schwarz methods without subdomain overlap and with new coarse spaces*, in Domain Decomposition Methods in Sciences and Engineering, R. Glowinski, J. Périaux, Z. Shi, and O. B. Widlund, eds., John Wiley & Sons, 1996, pp. 141–157. Proceedings from the Eight International Conference on Domain Decomposition Methods, May 1995, Beijing.
6. D. BRAESS, W. DAHMEN, AND C. WIENERS, *A multigrid algorithm for the mortar finite element method*, SIAM J. Numer. Anal., 37 (2000), pp. 48–69.
7. X.-C. CAI, M. DRYJA, AND M. V. SARKIS, *Overlapping nonmatching grid mortar element methods for elliptic problems*, SIAM J. Numer. Anal., 36 (1999), pp. 581–606.
8. M. CASARIN AND O. B. WIDLUND, *A hierarchical preconditioner for the mortar finite element method*, ETNA, 4 (1996), pp. 75–88.
9. P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, North Holland Publishing Company, 2nd ed., 1979.
10. M. DRYJA, *Additive Schwarz methods for elliptic mortar finite element problems*, in Modeling and Optimization of Distributed Parameter Systems with Applications to Engineering, K. Malanowski, Z. Nahorski, and M. Peszynska, eds., IFIP, Chapman & Hall, London, 1996.
11. ———, *An iterative substructuring method for elliptic mortar finite element problems with a new coarse space*, East-West J. Numer. Math., 5 (1997), pp. 79–98.
12. M. DRYJA, M. V. SARKIS, AND O. B. WIDLUND, *Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions*, Numer. Math., 72 (1996), pp. 313–348.
13. M. DRYJA AND O. B. WIDLUND, *Schwarz methods of Neumann-Neumann type for three-dimensional elliptic finite element problems*, Comm. Pure Appl. Math., 48 (1995), pp. 121–155.
14. J. GOPALAKRISHNAN AND J. E. PASCIAK, *Multigrid for the mortar finite element method*, SIAM J. Numer. Anal., 37 (2000), pp. 1029–1052.
15. T. RAHMAN, *Additive Schwarz Methods for some Elliptic Problems*, PhD thesis, Dept. of Informatics, Univ. of Bergen, 2000.
16. B. F. SMITH, P. E. BJØRSTAD, AND W. GROPP, *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*, Cambridge University Press, 1996.
17. D. STEFANICA AND A. KLAWONN, *The FETI method for mortar finite elements*, in Proceedings of 11th International Conference on Domain Decomposition Methods, C.-H. Lai, P. Bjorstad, M. Cross, and O.Widlund, eds., DDM.org, 1999, pp. 121–129.
18. B. WOHLMUTH AND R. KRAUSE, *Multigrid methods based on the unconstrained product space arising from mortar finite element discretizations*, SIAM J. Numer. Anal., 39 (2001), pp. 192–213.
19. B. I. WOHLMUTH, *A multigrid for saddle point problems arrising from mortar finite element discretizations*, ETNA, 11 (2000), pp. 43–54.