CrossMark

BIT

Additive Schwarz preconditioner for the finite volume element discretization of symmetric elliptic problems

L. Marcinkowski[1] · **T. Rahman[2]** · **A. Loneland[2,3]** ·
J. Valdman[4,5]

Abstract A symmetric and a nonsymmetric variant of the additive Schwarz pre-
conditioner are proposed for the solution of a class of finite volume element
discretization of the symmetric elliptic problem in two dimensions, with large jumps
in the entries of the coefficient matrices across subdomains. It is shown that the conver-

✉ L. Marcinkowski
  lmarcin@mimuw.edu.pl

  T. Rahman
  Talal.Rahman@hib.no

  A. Loneland
  Atle.Loneland@ii.uib.no

  J. Valdman
  jvaldman@prf.jcu.cz

[1] Faculty of Mathematics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland

[2] Department of Computing, Mathematics and Physics, Bergen University College,
     5020 Bergen, Norway

[3] Department of Informatics, University of Bergen, 5020 Bergen, Norway

[4] Institute of Mathematics and Biomathematics, University of South Bohemia, Branišovská 31,
     37005 České Budějovice, Czech Republic

[5] Institute of Information Theory and Automation of the ASCR, Pod Vodárenskou věží 4,
     18208 Prague, Czech Republic

gence of the preconditioned generalized minimal residual iteration using the proposed preconditioners depends polylogarithmically, in other words weakly, on the mesh parameters, and that they are robust with respect to the jumps in the coefficients.

1 Introduction

We consider the classical finite volume element discretization of the second-order elliptic partial differential equation, where we seek for the discrete solution in the space of standard P_1 conforming finite element functions, or the space of continuous and piecewise linear functions, cf. [17,20]. We further consider the second-order elliptic partial differential equation to have coefficients that may have large jumps across subdomains. Due to the finite volume discretization, the resulting systems are in general nonsymmetric, which become increasingly nonsymmetric for coefficients varying increasingly inside the finite elements. All of these make the problem hard to solve in a reasonable amount of time, particularly in large scale computations. Robust and efficient algorithms for the numerical solution of such systems are therefore needed. The design of such algorithms is often challenging, particularly their analysis which is not as well understood as it is for the symmetric system. The main objective of this paper is to design and analyze a class of robust and scalable preconditioners based on the additive Schwarz domain decomposition methodology, cf. [30,32], using them in the preconditioned generalized minimal residual (GMRES) iteration based on minimizing the energy norm of the residual, cf. [9,29], for solving the system.

The finite volume element method or the FVE method, also known in the literature as the control volume finite element method or the CVFE, provides a systematic approach to construct a finite volume or a control volume discretization of the differential equations using a finite element approximation of the discrete solution. The method has drawn much interest in the scientific communities over the years. It has both the flexibility of a finite element method to easily adapt to any geometry, and the property of a finite volume discretization of being locally conservative, like conservation of mass, conservation of energy, etc. For a quick overview of the existing research on the methodology, we refer to [1–4,11,12,17,20,23,34].

Additive Schwarz methods have been studied extensively in the literature, see [30,32] and the references therein. When it comes to solving second-order elliptic problems, the general focus has been on solving symmetric systems resulting from the finite element discretization of the problem. Despite the growing interest for finite volume elements, the research on fast methods for the numerical solution of nonsymmetric systems resulting from the finite volume element discretization has been very limited. In particular methods like the domain decomposition, which

are considered among the most powerful methods for large scale computation, have rarely been tested on finite volume elements. Among the few existing works which can be found in the literature, are the works of [13,35] based on the overlapping partition of the domain, and the works of [27,33] based on the nonoverlapping partition of the domain. The latter ones are given without the convergence analysis.

In this paper, we propose nonoverlapping domain decomposition methods for the finite volume element, which are based on substructuring, and formulate them as additive Schwarz methods. We show that their convergence is robust with respect to jumps in the coefficients across subdomains, and depends poly-logarithmically on the mesh parameter when used as preconditioners in the GMRES iteration. For an overview of important classes of iterative substructuring methods we refer to [32, chapters 4–6] for symmetric and positive definite systems, and [32, Chapter 11.3] for their extension to nonsymmetric systems. We restrict ourselves to problems in 2D; extension to 3D is possible and will be investigated in the future. Also, since the present work is an attempt to develop a first analysis of iterative substructuring type domain decomposition methods for the finite volume element, we limit ourselves to coefficients that may have jumps only across subdomain boundaries. However, by modifying the coarse space, e.g. using oscillatory boundary conditions instead of the linear boundary conditions for the basis functions, cf. [18,19,27], or by enriching the coarse space with eigenfunctions from local eigenvalue problems, cf. [10,14,21,31], it will be possible to extend the methods to effectively deal with highly heterogeneous coefficients.

For the general purpose of designing additive Schwarz preconditioners for a finite volume element discretization in this paper, we have formulated an abstract framework which is then later used in the analysis of the preconditioners proposed in the paper. The framework borrows the basic ingredients of the abstract Schwarz framework for additive Schwarz methods, cf. [30,32], while the analysis follows the work of [9] where additive Schwarz methods were considered for the advection-diffusion problem. The framework has already been used in two of authors recent papers on the Crouzeix–Raviart finite volume element, cf. [24,25], demonstrating its usefulness in the design and analysis of new and effective preconditioners for the finite volume element discretization of elliptic problems.

For further information on domain decomposition methods for nonsymmetric problems in general, we refer to [26,30,32] and references therein.

The paper is organized as follows: in Sect. 2, we present the differential problem, and in Sect. 3, its finite volume element discretization. In Sect. 4, we present the two variants of the additive Schwarz preconditioners and the two main results, Theorems 4.1 and 4.2. The complete analysis is provided in the next two sections, the abstract framework in Sect. 5, and the required estimates in Sect. 6. Finally, numerical results are provided in Sect. 7.

Throughout this paper, we use the following notations: $x \lesssim y$ and $w \gtrsim z$ denote that there exist positive constants c, C independent of mesh parameters and the jump of coefficients such that $x \leq cy$ and $w \geq Cz$, respectively.

2 The differential problem

Given Ω, a polygonal domain in the plane, and $f \in L^2(\Omega)$, the purpose is to solve the following differential equation,

$$-\nabla \cdot (A(x)\nabla u)(x) = f(x), \quad x \in \Omega,$$
$$u(s) = 0, \quad s \in \partial\Omega,$$

where $A \in (L^\infty(\Omega))^{2\times 2}$ is a symmetric matrix valued function satisfying the uniform ellipticity as follows,

$$\exists \alpha > 0 \quad \text{such that} \quad \xi^T A(x)\xi \geq \alpha |\xi|_2^2 \quad \forall x \in \Omega \quad \text{and} \quad \forall \xi \in \mathbb{R}^2,$$

where $|\xi|_2^2 = \xi_1^2 + \xi_2^2$. Further we consider α equal to 1 which can be always obtained by scaling the original problem by α^{-1}. We assume that Ω is decomposed into a set of disjoint polygonal subdomains $\{D_j\}$ such that, in each subdomain D_j, $A(x)$ is continuous and smooth in the sense that

$$|A|_{W^{1,\infty}(D_i)} \leq C_\Omega, \tag{2.1}$$

where C_Ω is a positive constant. We also assume that

$$\exists \lambda_j > 0 \quad \text{such that} \quad \xi^T A(x)\xi \geq \lambda_j |\xi|_2^2 \geq |\xi|_2^2 \quad \forall x \in D_j \quad \text{and} \quad \forall \xi \in \mathbb{R}^2.$$

Due to $A \in (L^\infty(D_j))^{2\times 2}$, we have the following,

$$\exists \Lambda_j > 0 \text{ such that} \quad |v^T A(x)\xi| \leq \Lambda_j |v|_2 |\xi|_2 \quad \forall x \in D_j \quad \text{and} \quad \forall \xi, v \in \mathbb{R}^2.$$

We also assume that $\Lambda_j \lesssim \lambda_j$. We then have

$$\lambda_j |u|_{H^1(D_j)}^2 \leq \int_{D_j} \nabla u^T A(x)\nabla u \, dx \leq \Lambda_j |u|_{H^1(D_j)}^2 \quad \forall u \in H^1(D_j). \tag{2.2}$$

In the weak formulation, the differential problem is then to find $u \in H_0^1(\Omega)$ such that

$$a(u, v) = f(v) \quad \forall v \in H_0^1(\Omega), \tag{2.3}$$

where

$$a(u, v) = \int_\Omega \nabla u^T A(x)\nabla v \, dx \quad \text{and} \quad f(v) = \int_\Omega fv \, dx.$$

3 The discrete problem

For the discretization, we use a finite volume element discretization where the Eq. (2.3) is discretized using the standard finite volume method on a mesh which is dual to the primal mesh, and where the finite element space or the solution space is defined on the primal mesh, see for instance in [16,17,20]; for an overview of finite volume element methods we refer to [23].

Let $T_h = T_h(\Omega)$ be a shape regular triangulation of Ω, cf. [8] or [6], hereon referred to as the primal mesh consisting of triangles $\{\tau\}$ with the size parameter $h = \max_{\tau \in T_h} \operatorname{diam}(\tau)$, and let Ω_h, $\partial \Omega_h$, and $\overline{\Omega}_h$ be the sets of triangle vertices corresponding to $\Omega, \partial\Omega$, and $\overline{\Omega}$, respectively. We assume that each $\tau \in T_h$ is contained in one of D_j.

Define V_h as the conforming linear finite element space consisting of functions which are continuous piecewise linear over the triangulation T_h, and are equal to zero on $\partial\Omega$.

Now, let $T_h^* = T_h^*(\Omega)$ be the dual mesh corresponding to T_h. For simplicity we use the so called Donald mesh for the dual mesh. For each triangle $\tau \in T^h$, let c_τ be the centroid, x_j, $j = 1, 2, 3$ the three vertices, and $m_{kl} = m_{lk}$, $k, l = 1, 2, 3$ the three edge midpoints. Divide each triangle τ into three polygonal regions inside the triangle by connecting its edge midpoints $m_{kl} = m_{lk}$ to its centroid c_τ with straight lines. One such polygonal region $\omega_{\tau,x_1} \subset \tau$, associated with the vertex x_1, as illustrated in Fig. 1, is the region which is enclosed by the line segments $\overline{c_\tau m_{13}}, \overline{m_{13}x_1}, \overline{x_1 m_{12}}$, and $\overline{m_{12}c_\tau}$, and whose vertices are c_τ, m_{13}, x_1, and m_{12}. Now let ω_{x_k} be the control volume associated with the vertex x_k, which is the sum of all such polygonal regions associated with the vertex x_k, i.e.

$$\omega_{x_k} = \bigcup_{\{\tau \in T_h : x_k \text{ is a vertex of } \tau\}} \omega_{\tau,x_k}$$

The set of all such control volumes form our dual mesh, i.e. $T_h^* = T_h^*(\Omega) = \{\omega_x\}_{x \in \overline{\Omega}_h}$. A control volume ω_{x_k} is called a boundary control volume if $x_k \in \partial\Omega_h$.

Let V_h^* be the space of piecewise constant functions over the dual mesh T_h^*, which have values equal to zero on $\partial\Omega_h$. We let the nodal basis of V_h be $\{\phi_x\}_{x \in \Omega_h}$, where ϕ_x is the standard finite element basis function which is equal to one at the vertex x and zero at all other vertices. Analogously, the nodal basis of V_h^* is $\{\psi_x\}_{x \in \Omega_h}$ where ψ_x is a piecewise constant function which is equal to one over the control volume ω_x associated with the vertex x, and is zero elsewhere.

Fig. 1 Showing ω_{τ,x_1} (*shaded region*) which is part of the control volume ω_{x_1} restricted to the triangle τ. The control volume is associated with the vertex x_1
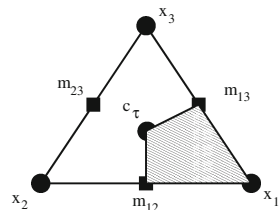
The two interpolation operators, I_h and I_h^*, are defined as follows. $I_h : C(\Omega) + V_h^* \to V_h$ and $I_h^* : C(\Omega) \to V_h^*$ are given respectively as

$$I_h^* v = \sum_{x \in \Omega_h} v(x) \psi_x \quad \text{and} \quad I_h v = \sum_{x \in \Omega_h} v(x) \phi_x.$$

We note here that $I_h I_h^* v = v$ for $v \in V_h$, as well as $I_h^* I_h u = u$ for $u \in V_h^*$.

Let the finite volume bilinear form be defined on $V_h \times V_h^*$ as $a_{FV} : V_h \times V_h^* \to \mathbb{R}$ such that

$$a_{FV}(u, v) = - \sum_{x_i \in \Omega_h} v(x_i) \int_{\partial \omega_{x_i}} A \nabla u \mathbf{n} ds \quad u \in V_h, \ v \in V_h^*,$$

or equivalently on $V_h \times V_h$ as $a_h : V_h \times V_h \to \mathbb{R}$ such that

$$a_h(u, v) = a_{FV}(u, I_h^* v) \tag{3.1}$$

for $u, v \in V_h$. We note here that $a_h(\cdot, \cdot)$ is a nonsymmetric bilinear form in general, while $a(\cdot, \cdot)$ is a symmetric bilinear form. The discrete problem is then to find $u_h \in V_h$ such that

$$a_{FV}(u_h, v) = f(v) \quad \forall v \in V_h^*, \tag{3.2}$$

or equivalently $a_h(u_h, v) = f(I_h^* v) \quad \forall v \in V_h$. The problem has a unique solution for h sufficiently small, which follows from the fact that (5.2) holds, cf. Proposition 6.1, leading to the fact that $a_h(u, u)$ is V_h-elliptic, cf. Lemma 5.2. An error estimate in the norm induced by $a(u, v)$ can also be given, cf. e.g. [17, § 3],

We close this section with the following remark. Note that in some cases, the bilinear form $a_h(\cdot, \cdot)$ may equal the symmetric bilinear form $a(\cdot, \cdot)$, as for instance in the case when the matrix A is piecewise constant over the subdomains $\{D_j\}$. This may not be true if we choose to use a different dual mesh or if A is not piecewise constant. In this paper, we only consider the case when the bilinear form $a_h(\cdot, \cdot)$ is nonsymmetric.

4 An edge based additive Schwarz method

We propose in this section two variants of the edge based additive Schwarz method (ASM) for the discrete finite volume element problem of the previous section. They are constructed in the same way as it is done in the standard case using the classical abstract framework of additive Schwarz methods, cf. [26,30,32]. The convergence analysis of the methods is however based on the abstract framework which will be developed in Sect. 5. The new framework is an extension of the classical framework, and is based on three key assumptions. We validate those assumptions for the proposed methods in Sect. 6, thereby completing the analysis.

We assume that we have a partition of Ω into the set of N nonoverlapping polygonal subdomains $\{\Omega_j\}_{j=1}^N$, such that they form a shape regular coarse partition of Ω, as described in [7]. Accordingly, we assume that the intersection between two subdomains

is either a vertex, an edge, or the empty set, and that there is a fixed number (M) of reference polygonal domains E_k $(1 \leq k \leq M)$ of diameter one each, such that for each subdomain Ω_j there is an invertible affine map $\mu_j(x) = B_j x + b_j$ which maps Ω_j to one of the reference domains $E_{k(j)}$ $(1 \leq k(j) \leq M)$. Here B_j is a 2×2 nonsingular matrix and b_j a 2×1 vector, and

$$\|B_j\|_\infty \lesssim H_j^{-1}, \qquad \|B_j^{-1}\|_\infty \lesssim H_j,$$

where $H_j = \operatorname{diam}(\Omega_j)$.

We assume that each subdomain Ω_k lies in exactly one of the polygonal subdomains $\{D_j\}$ described earlier, and that none of their boundaries cross each other. The interface

$$\Gamma = \bigcup_{k=1}^N \partial \Omega_k \backslash \partial \Omega,$$

which is the sum of all subdomain edges and subdomain vertices not lying on the boundary $\partial \Omega$, or crosspoints, plays a crucial role in the design of our preconditioner. We also assume that the primal mesh T_h is perfectly aligned with the partitioning of Ω, in other words, no edges of the primal mesh cross any edge of the coarse substructure. As a consequence, the coefficient matrix $A(x)$ restricted to a subdomain Ω_k is in $(W^{1,\infty}(\Omega_k))^{2\times2}$, and hence (cf. (2.1))

$$|A|_{W^{1,\infty}(\Omega_k)} \leq |A|_{W^{1,\infty}(D_j)} \leq C_\Omega.$$

Each subdomain Ω_k inherits its own local triangulation from the T_h, denote it by $T_h(\Omega_k) = \{\tau \in T_h : \tau \subset \Omega_k\}$. Let $V_h(\Omega_k)$ be the space of continuous and piecewise linear functions over the triangulation $T_h(\Omega_k)$, which are zero on $\partial \Omega \cap \partial \Omega_k$, and let $V_{h,0}(\Omega_k) := V_h(\Omega_k) \cap H_0^1(\Omega_k)$. The local spaces are equipped with the bilinear form

$$a_k(u, v) = \int_{\Omega_k} \nabla u^T A(x) \nabla v \, dx.$$

We define the local projection operator $\mathscr{P}_k : V_h(\Omega_k) \to V_{h,0}(\Omega_k)$ such that

$$a_k(\mathscr{P}_k u, v) = a_k(u, v) \qquad \forall v \in V_{h,0}(\Omega_k)$$

and the local discrete harmonic extension operator $\mathscr{H}_k : V_h(\Omega_k) \to V_h(\Omega_k)$ such that

$$\mathscr{H}_k u = u - \mathscr{P}_k u.$$

Note that $\mathscr{H}_k u$ is equal to u on the boundary $\partial \Omega_k$, and discrete harmonic inside Ω_k in the sense that

$$a_k(\mathscr{H}_k u, v) = 0 \quad \forall v \in V_{h,0}(\Omega_k).$$

The local and global spaces of discrete harmonic functions are then defined as

$$W_k = \mathcal{H}_k V_h(\Omega_k) \quad \text{and} \quad W = \mathcal{H} V_h = \{u \in V_h : u_{|\Omega_k} = \mathcal{H}_k u_{|\Omega_k}\},$$

respectively.

We now define the subspaces required for the ASM preconditioner, cf. [26,30,32]. For each subdomain Ω_k, the local subspace $V_k \subset V_h$ is defined as $V_{h,0}(\Omega_k)$ extending it by zero to the rest of the subdomains, i.e.

$$V_k = \{v \in V_h : v_{|\overline{\Omega}_k} \in V_{h,0}(\Omega_k) \quad \text{and} \quad v_{|\Omega \setminus \overline{\Omega}_k} = 0\}$$

The coarse space $V_0 \subset W$ is defined as the space of discrete harmonic functions which are piecewise linear over the subdomain edges (i.e. linear along each subdomain edge and discrete harmonic inside each subdomain). Its degrees of freedom are associated with the subdomain vertices lying inside the domain Ω, i.e. the crosspoints, and hence the dimension equals the cardinality of $\mathcal{V} = \bigcup_k \mathcal{V}_k$, where $\mathcal{V}_k$ is the set of all subdomain vertices which are not on the boundary $\partial \Omega$. For an analogous coarse space, we refer to [32, §5.4] where 3D substructuring methods are considered. Using their ideas it is possible to extend the method to 3D.

Finally, the local edge based subspaces are defined as follows. For each subdomain edge Γ_{kl}, which is the interface between Ω_k and Ω_l, we let $V_{kl} \subset W$ be the local edge based subspace consisting of functions which may be nonzero inside Γ_{kl}, but zero on the rest of the interface Γ, and discrete harmonic in the subdomains. It is not difficult to see that the support of V_{kl} is contained in $\overline{\Omega}_k \cup \overline{\Omega}_l$.

We have the following decomposition of the finite element spaces W and V_h. Both the symmetric and the nonsymmetric variant of the preconditioner use the same decomposition:

$$
\begin{aligned}
W &= V_0 + \sum_{\Gamma_{kl} \subset \Gamma} V_{kl}, \\
V_h &= W + \sum_{k=1}^{N} V_k = V_0 + \sum_{\Gamma_{kl} \subset \Gamma} V_{kl} + \sum_{k=1}^{N} V_k.
\end{aligned}
\tag{4.1}
$$

Note that the subspaces of W, i.e. V_0 and V_{kl}, $\Gamma_{kl} \subset \Gamma$, are orthogonal to the subspaces V_k, $k = 1, \ldots, N$, with respect to the bilinear form $a(\cdot, \cdot)$.

Remark 4.1 If $A(x)$ is the identity matrix, and $\{\Omega_j\}_{1 \leq j \leq N}$ form a coarse triangulation of the domain Ω, then the coarse space V_0 will consist of functions that are continuous and piecewise linear over the coarse triangulation, cf. [32, §5.4.1].

In the following, we describe the two preconditioners: the symmetric and the nonsymmetric preconditioner. We will see later in the numerical experiments section that the two are very similar in their performance. It is not difficult to see that they also have similar computational complexities.

4.1 Symmetric preconditioner

For the symmetric variant of the preconditioner, we define the coarse space operator and the local subspace operators, $T_k : V_h \rightarrow V_k, k = 0, 1, \ldots, N$, as

$$a(T_k u, v) = a_h(u, v) \qquad \forall v \in V_k,$$

and the local edge based subspace operators $T_{kl} : V_h \rightarrow V_{kl}, \Gamma_{kl} \subset \Gamma$, as

$$a(T_{kl} u, v) = a_h(u, v) \qquad \forall v \in V_{kl}.$$

Note that the bilinear form used to calculate each of the operations above, that is on the left hand side of the equation, is the symmetric bilinear form $a(\cdot, \cdot)$. Now, defining our additive Schwarz operator T as the sum of the operators, that is

$$T = T_0 + \sum_{k=1}^{N} T_k + \sum_{\Gamma_{kl} \subset \Gamma} T_{kl},$$

we can replace the discrete problem (3.2) by the following equivalent preconditioned system of equations, in the operator form, cf. [30, Chapter 5]:

$$T u_h = g, \tag{4.2}$$

where $g = g_0 + \sum_{k=1}^{N} g_k + \sum_{\Gamma_{kl} \subset \Gamma} g_{kl}, g_0 = T_0 u_h, g_k = T_k u_h, k = 1, \ldots, N$, and $g_{kl} = T_{kl} u_h, \Gamma_{kl} \subset \Gamma$, are the right hand sides with u_h being the exact solution. Note that the right hand sides can be calculated without knowing the exact solution, cf. [30].

The GMRES iteration is used to solve the preconditioned system (4.2). The convergence rate is given by the standard GMRES convergence (Theorem 5.1) with estimates of its two parameters using Theorem 4.1.

Theorem 4.1 *There exists h_1 such that, if $h \leq h_1$, then for any $u \in V_h$*

$$a(Tu, Tu) \lesssim a(u, u), \qquad a(Tu, u) \gtrsim \left(1 + \log\left(\frac{H}{h}\right)\right)^{-2} a(u, u),$$

where $H = \max_k(H_k)$ and $H_k = \mathrm{diam}(\Omega_k)$.

The theorem is proved using the abstract framework given in Sect. 5, i.e. by using Theorem 5.2, and the three propositions in Sect. 6, Propositions 6.1, 6.2, and 6.3, verifying the three assumptions required by the framework.

4.2 Nonsymmetric preconditioner

Analogously, for the nonsymmetric variant of the preconditioner, the coarse space operator and the local subspace operators, $S_k : V_h \rightarrow V_k, k = 0, 1, \ldots, N$, are defined as

$$a_h(S_k u, v) = a_h(u, v) \qquad \forall v \in V_k,$$

and the local edge based subspace operators, $S_{kl} : V_h \rightarrow V_{kl}, \Gamma_{kl} \subset \Gamma$, as

$$a_h(S_{kl} u, v) = a_h(u, v) \qquad \forall v \in V_{kl}.$$

The bilinear form used to calculate each of the operations above is the nonsymmetric bilinear form $a_h(\cdot, \cdot)$. Again, denoting the additive Schwarz operator by S, where

$$S = S_0 + \sum_{k=1}^{N} S_k + \sum_{\Gamma_{kl} \subset \Gamma} S_{kl},$$

the discrete problem (3.2) can be replaced by the equivalent preconditioned system of equations, in the operator form:

$$S u_h = \hat{g}, \tag{4.3}$$

where $\hat{g} = \hat{g}_0 + \sum_{k=1}^{N} \hat{g}_k + \sum_{\Gamma_{kl} \subset \Gamma} \hat{g}_{kl}$, $\hat{g}_k = S_k u_h$, $k = 0, 1, \ldots, N$, and $\hat{g}_{kl} = S_{kl} u_h$, $\Gamma_{kl} \subset \Gamma$, are the right hand sides with u_h being the exact solution. As in the symmetric case, the right hand sides can be calculated without knowing the exact solution.

Again, the GMRES iteration is used to solve the preconditioned system (4.3), whose convergence rate is given by the standard GMRES convergence (Theorem 5.1) with estimates of its two parameters using Theorem 4.2.

Theorem 4.2 *There exists an h_1 such that, if $h \leq h_1$, then for any $u \in V_h$*

$$a(Su, Su) \lesssim a(u, u), \quad a(Su, u) \gtrsim \left(1 + \log\left(\frac{H}{h}\right)\right)^{-2} a(u, u),$$

where $H = \max_k(H_k)$ and $H_k = \mathrm{diam}(\Omega_k)$.

The theorem is again proved using the abstract framework of Sect. 5, i.e. by using Theorem 5.3, and the three propositions in Sect. 6, Propositions 6.1, 6.2, and 6.3, verifying the three assumptions required by the framework.

5 The abstract framework

In this section, we formulate an abstract framework for the convergence analysis of additive Schwarz methods for a class of finite volume elements, using the methods as preconditioners in the GMRES iteration. The framework is based on three key assumptions which need to be verified every time a convergence analysis is to be performed. Once the assumptions are verified, the framework can be used to derive estimates for the convergence of the method under consideration. For two very recent applications of the framework, where the Crouzeix–Raviart finite volume element has been considered, we refer to [24,25].

We consider a family of finite dimensional subspaces V_h indexed by the parameter h, an inner product $a(\cdot, \cdot)$ and its induced norm $\| \cdot \|_a := \sqrt{a(\cdot, \cdot)}$, and a family of discrete problems: Find $u_h \in V_h$

$$a_h(u_h, v) = f(v) \quad \forall v \in V_h,$$

where $a_h(u, v)$ is a nonsymmetric bilinear form.

We start by stating the convergence result of the GMRES iteration (cf. [28]) for solving a system of equations, in the operator form,

$$Pu = b,$$

where the pair P and b can be either the pair T and g or the pair S and $\hat{g}$, in our case. The theorem was originally developed using the l_2 norm, cf. e.g. [15], it however extends to any Hilbert norm, cf. [9,29]. The convergence rate is based on the two parameters: the smallest eigenvalue of the symmetric part of the operator and the norm of the operator, respectively

$$\beta_1 = \inf_{u \neq 0} \frac{a(Pu, u)}{\|u\|_a^2} \quad \text{and} \quad \beta_2 = \sup_{u \neq 0} \frac{\|Pu\|_a}{\|u\|_a}. \tag{5.1}$$

A standard convergence rate of the GMRES iteration, in the $\| \cdot \|_a$ norm, is given in the following theorem.

Theorem 5.1 (Eisenstat et al. [15]) *If $\beta_1 > 0$, then the GMRES method for solving the linear system* (5.9) *converges for any starting value $u_0 \in V_h$ with the following estimate*:

$$\|g - Pu_m\|_a \leq \left(1 - \frac{\beta_1^2}{\beta_2^2}\right)^{m/2} \|g - Pu_0\|_a,$$

where u_m is the m-th iterate of the GMRES method.

The abstract framework is based on three assumptions leading to the two main results of the framework, namely Theorems 5.2 and 5.3 respectively for the symmetric and the nonsymmetric case.

In the first assumption, we assume that the nonsymmetric bilinear form $a_h(\cdot, \cdot)$ is a small perturbation of the symmetric bilinear form $a(\cdot, \cdot)$.

Assumption (1) For all $h < h_0$, where h_0 is a constant,

$$E_h(u, v) := a_h(u, v) - a(u, v)$$

converges to zero as h tends to zero satisfying the following uniform bound.

$$\exists\, C_E > 0 : \ \forall h < h_0 \quad \text{and} \quad \forall u, v \in V_h, \quad |E_h(u, v)| \leq C_E h \|u\|_a \|v\|_a, \tag{5.2}$$

where C_E is a constant independent of h.

The above assumption is important, also because, it leads to the continuity and the coerciveness of the bilinear form $a_h(\cdot, \cdot)$, as shown in the following lemmas.

Lemma 5.1 *For any $M \in (1, 2)$ there exists $h_1 \leq h_0$ such that if $h < h_1$ then the bilinear form $a_h(u, v)$ is uniformly bounded in the $\| \cdot \|_a$-norm, i.e.*

$$\forall u, v \in V_h \quad |a_h(u, v)| \leq M \|u\|_a \|v\|_a. \tag{5.3}$$

Proof It follows from (5.2) [Assumption (5)] that

$$a_h(u, v) = a(u, v) + E_h(u, v) \leq (1 + C_E h)\|u\|_a \|v\|_a \leq (1 + C_E h_1)\|u\|_a \|v\|_a.$$

Taking $h_1 = \min((M - 1)/C_E, h_0)$ ends the proof.

Lemma 5.2 *For any $\alpha \in (0, 1)$, there exists $h_1 \leq h_0$ such that if $h < h_1$ then the bilinear form $a_h(u, v)$ is uniformly V_h-elliptic in the $\| \cdot \|_a$-norm, i.e.*

$$\forall u \in V_h \quad a_h(u, u) \geq \alpha \|u\|_a^2. \tag{5.4}$$

Proof By assumption (5.2) [Assumption (1)], we have

$$a(u, u) \leq a_h(u, u) + |E_h(u, u)| \leq a_h(u, u) + C_E h a(u, u).$$

If $h < h_1 \leq h_0$ and $C_E h_1 \leq 1 - \alpha$, then

$$a_h(u, u) \geq (1 - C_E h_1) a(u, u) \geq \alpha a(u, u),$$

and the proof follows.

The following two assumptions are associated with the domain decomposition, where the function space V_h is decomposed into its subspaces as follows,

$$V_h = Z_0 + \sum_{k=1}^{N} Z_k,$$

where $Z_k \subset V_h$ for $k = 0, 1, \ldots, N$, with Z_0 being the coarse space. These are the same assumptions used in the abstract Schwarz framework for symmetric positive definite systems, cf. [30,32], the first one ensuring a stable splitting of the function space, while the second one ensuring a strengthened Cauchy–Schwarz inequality between the subspaces excluding the coarse space.

Assumption (2) For any $u \in V_h$, there are functions $u_k \in Z_k$, $k = 0, 1, \ldots, N$, such that the following holds. There exists a positive constant C_0 (which may depend on the mesh parameters) such that

$$u = u_0 + \sum_{k=1}^{N} u_k \tag{5.5}$$

and

$$\sum_{k=0}^{N} a(u_k, u_k) \leq C_0^2 a(u, u).$$ (5.6)

Assumption (3) For any $k, l = 1, \ldots, N$, let ϵ_{kl} be the minimal nonnegative constants such that

$$a(u_k, u_l) \leq \epsilon_{kl} \|u_k\|_a \|u_l\|_a \qquad u_k \in Z_k, u_l \in Z_l.$$ (5.7)

Let $\rho(\mathcal{E})$ be the spectral radius of the $N \times N$ symmetric matrix $\mathcal{E} = (\epsilon_{kl})_{k,l=1}^{N}$.

5.1 Symmetric preconditioner

For $k = 0, 1, \ldots, N$, we define the projection operator $T_k : V_h \to Z_k$ as

$$a(T_k u, v) = a_h(u, v) \qquad \forall v \in Z_k.$$ (5.8)

Note that the bilinear form $a(u, v)$ is an inner product in V_h, hence T_k is a well defined linear operator. Let the additive Schwarz preconditioned operator $T : V_h \to V_h$ be given as $T = T_0 + \sum_{k=1}^{N} T_k$, and the original problem be replaced by the preconditioned one:

$$T u_h = g,$$ (5.9)

where $g = g_0 + \sum_{k=1}^{N} g_k$ with $g_k = T_k u_h$ for $k = 0, 1, \ldots, N$. Note that T_k and T are in general nonsymmetric.

The following lemma is needed for the two main theorems of the framework, namely Theorems 5.2 and 5.3.

Lemma 5.3 *Let $u_k \in Z_k$ for $k = 0, 1, \ldots, N$. Then*

$$\left\| \sum_{k=0}^{N} u_k \right\|_a^2 \leq 2(1 + \rho(\mathcal{E})) \sum_{k=0}^{N} \|u_k\|_a^2,$$

where $\rho(\mathcal{E})$ is the spectral radius of the matrix $\mathcal{E} = (\epsilon_{kl})_{k,l=1}^{N}$.

Proof We see that

$$\left\| \sum_{k=0}^{N} u_k \right\|_a^2 \leq 2\|u_0\|_a^2 + 2 \left\| \sum_{k=1}^{N} u_k \right\|_a^2.$$

Using (5.7) and a Schwarz inequality in the l_2-norm we get

$$\left\| \sum_{k=1}^{N} u_k \right\|_a^2 = \sum_{k,l=1}^{N} a(u_k, u_l) \leq \sum_{k,l=1}^{N} \epsilon_{kl} \|u_k\|_a \|u_l\|_a$$

$$\leq \rho(\mathscr{E}) \sqrt{\sum_{k=1}^{N} \|u_k\|_a^2} \sqrt{\sum_{l=1}^{N} \|u_l\|_a^2} = \rho(\mathscr{E}) \sum_{k=1}^{N} \|u_k\|_a^2,$$

and the proof follows.

The first main theorem of the framework giving estimates of the two parameters of the GMRES convergence, cf. Theorem 5.1, for the symmetric preconditioner, is stated in the following.

Theorem 5.2 *There exists $h_1 \leq h_0$ such that if $h < h_1$ then*

$$a(Tu, Tu) \leq \beta_2^2 a(u, u), \tag{5.10}$$

$$a(Tu, u) \geq \beta_1 a(u, u), \tag{5.11}$$

where $\beta_2 = 2M(1 + \rho(\mathscr{E}))$ and $\beta_1 = \alpha^2 C_0^{-2} - \beta_2 C_E h$.

Proof It follows from Lemma 5.3 that

$$a(Tu, Tu) = \left\| \sum_{k=0}^{N} T_k u \right\|_a^2 \leq 2(1 + \rho(\mathscr{E})) \sum_{k=0}^{N} \|T_k u\|_a^2.$$

By (5.3) and (5.8), we get

$$\sum_{k=0}^{N} \|T_k u\|_a^2 = \sum_{k=0}^{N} a(T_k u, T_k u) = \sum_{k=0}^{N} a_h(u, T_k u) = a_h(u, Tu) \leq M \|u\|_a \|Tu\|_a.$$

$$\tag{5.12}$$

The upper bound (5.10) then follows with $\beta_2 = 2M(1 + \rho(\mathscr{E}))$.

To prove the lower bound, cf. (5.11), we start with the splitting of $u \in V_h$, cf. (5.5), such that (5.6) holds. Then using (5.4), (5.8), a Schwarz inequality, and (5.6), we get

$$\alpha a(u, u) \leq a_h(u, u) = \sum_{k=0}^{N} a_h(u, u_k) = \sum_{k=0}^{N} a(T_k u, u_k)$$

$$\leq \sum_{k=0}^{N} \|T_k u\|_a \|u_k\|_a \leq \sqrt{\sum_{k=0}^{N} \|T_k u\|_a^2} \sqrt{\sum_{k=0}^{N} \|u_k\|_a^2} \leq C_0 \sqrt{\sum_{k=0}^{N} \|T_k u\|_a^2} \|u\|_a.$$

This and (5.12), then yield

$$\alpha^2 a(u, u) \leq C_0^2 \sum_{k=0}^{N} \|T_k u\|_a^2 = C_0^2 a_h(u, Tu).$$

Finally, from the assumption (5.2) [Assumption (1)] and the upper bound (5.10), we get

$$a_h(u, Tu) = a(u, Tu) + E_h(u, Tu) \leq a(u, Tu) + C_E h \|u\|_a \|Tu\|_a$$
$$\leq a(u, Tu) + \beta_2 C_E h \|u\|_a^2.$$

Hence,

$$a(Tu, u) \geq (\alpha^2 C_0^{-2} - \beta_2 C_E h) a(u, u).$$

Taking $\beta_1 = (\alpha^2 C_0^{-2} - \beta_2 C_E h)$ we get the lower bound in (5.11).

Remark 5.2 In some cases the constant C_0 in (5.6) may depend on h ($C_0 = C_0(h)$) then $C_0(h)$ cannot grow too fast with decreasing h, otherwise β_1 may become negative and our theory would not work, e.g. if $\rho(\mathscr{E})$ is independent of h which is usually the case in ASM methods, then it would be sufficient if $\lim_{h \to 0} C_0^2(h) h = 0$, because then there exists an $h_1 \leq h_0$ such that β_1 is positive for any $h < h_1$.

5.2 Nonsymmetric preconditioner

For $k = 0, 1, \ldots, N$, we define the projection operators $S_k : V_h \to Z_k$ as

$$a_h(S_k u, v) = a_h(u, v) \qquad \forall v \in Z_k. \tag{5.13}$$

Note that the bilinear form $a_h(u, v)$ is Z_k-elliptic, cf. (5.4), so S_k is a well defined linear operator. Now, introducing the additive Schwarz operator $S : V_h \to V_h$ as $S = S_0 + \sum_{k=1}^{N} S_k$, we replace the original problem with

$$S u_h = \hat{g},$$

where $\hat{g} = \hat{g}_0 + \sum_{k=1}^{N} \hat{g}_k$ with $\hat{g}_k = S_k u_h$ for $k = 0, 1, \ldots, N$.

The second main theorem of the framework giving estimates of the two parameters of the GMRES convergence, for the nonsymmetric preconditioner, is stated in the following.

Theorem 5.3 *There exists $h_1 \leq h_0$ such that for any $h < h_1$, the following bounds hold.*

$$a(Su, Su) \leq \beta_2^2 a(u, u),$$
$$a(Su, u) \geq \beta_1 a(u, u)$$

where $\beta_2 = \frac{2M}{\alpha}(1 + \rho(\mathscr{E}))$ and $\beta_1 = \frac{\alpha^3}{M^2 C_0^2} - \beta_2 C_E h$, and, as before, $\rho(\mathscr{E})$ is the spectral radius of the matrix $\mathscr{E} = (\epsilon_{kl})_{k,l=1}^N$.

Proof We follow the lines of proof of Theorem 5.2. For the upper bound, we use Lemma 5.3 to see that

$$a(Su, Su) \le 2(1 + \rho(\mathscr{E})) \sum_{k=0}^N \|S_k u\|_a^2.$$

Using (5.3), (5.4), and (5.13), we get

$$
\alpha \sum_{k=0}^N a(S_k u, S_k u) \le \sum_{k=0}^N a_h(S_k u, S_k u) = \sum_{k=0}^N a_h(u, S_k u)
$$
$$
= a_h(u, Su) \le M\|u\|_a\|Su\|_a. \quad (5.14)
$$

The upper bound then follows with $\beta_2 = \frac{2M}{\alpha}(1 + \rho(\mathscr{E}))$.

For the lower bound, again, we use the splitting (5.5) of $u \in V_h$ such that (5.6) holds. Next (5.3), (5.4), (5.5), (5.13), a Schwarz inequality, and (5.6) yield that

$$
\alpha a(u, u) \le a_h(u, u) = \sum_{k=0}^N a_h(u, u_k) = \sum_{k=0}^N a_h(S_k u, u_k) \le M \sum_{k=0}^N \|S_k u\|_a \|u_k\|_a
$$
$$
\le M \sqrt{\sum_{k=0}^N \|S_k u\|_a^2} \sqrt{\sum_{k=0}^N \|u_k\|_a^2} \le M\,C_0 \sqrt{\sum_{k=0}^N \|S_k u\|_a^2}\,\|u\|_a.
$$

Combining the estimate above with (5.14), we get

$$
\alpha^2 a(u, u) \le M^2 C_0^2 \sum_{k=0}^N \|S_k u\|_a^2 \le \frac{M^2 C_0^2}{\alpha} a_h(u, Su).
$$

Finally, using similar arguments as in the proof of Theorem 5.2, we can conclude that

$$
a(u, Su) \ge \left(\frac{\alpha^3}{M^2 C_0^2} - \beta_2 C_E h \right) a(u, u),
$$

and the lower bound follows with $\beta_1 = \frac{\alpha^3}{M^2 C_0^2} - \beta_2 C_E h$.

6 Technical tools

In this section, we present the technical results necessary for the proof of Theorems 4.1 and 4.2. We use the abstract framework developed in the previous section, based on

which, we first show (5.2) [verifying Assumption (1)], then show that $\rho(\mathscr{E})$ is bounded by a constant [verifying Assumption (3)], and finally give an estimate of C_0^2 such that (5.5)–(5.6) hold [verifying Assumption (2)], all of these being formulated below as Propositions 6.1, 6.2, and 6.3, respectively.

We start with the proposition which shows that (5.2) holds true for the two bilinear forms $a(\cdot, \cdot)$ and $a_h(\cdot, \cdot)$ of (2.3) and (3.1), respectively.

Proposition 6.1 *It holds that*

$$\exists\, C_E > 0 : \ \forall h < h_0 \quad and \quad u, v \in V_h, \quad |a_h(u, v) - a(u, v)| \leq C_E h \|u\|_a \|v\|_a,$$

where C_E is a constant independent of h and the jumps of the coefficients across $\partial D_j s$, but may depend on C_Ω in (2.1).

The statement of this proposition is formulated in [17, §3] without proof. However, as stated in the paper, it can be proved by following the lines of the proof of Lemma 3.1 in [16].

The next two lemmas are well known, and are given here without proofs. The first lemma is the extension theorem for discrete harmonic functions, cf. e.g. [5, Lemma 5.1]. The second lemma gives an estimate of the $H_{00}^{1/2}(\Gamma_{kl})$ norm of a finite element function which is zero on $\partial\Omega_k \backslash \Gamma_{kl}$ by its $H^{1/2}$ seminorm and L^∞ norm, cf. e.g. [22, Lemma 4.1]. Let $\Gamma_{kl,h}$ be the set of nodal points that are on the open edge Γ_{kl} common to Ω_k and Ω_l.

Lemma 6.1 (Discrete extension theorem) *Let $u \in W_k$, then*

$$|u|_{H^1(\Omega_k)} \lesssim |u|_{H^{1/2}(\partial\Omega_k)}.$$

Lemma 6.2 *Let $u, u_1 \in W_k$ be such that $u_1(x) = u(x)$ for $x \in \Gamma_{kl,h}$ and $u_{1|\partial\Omega_k \backslash \Gamma_{kl}} = 0$, then*

$$\|u_1\|_{H_{00}^{1/2}(\Gamma_{kl})}^2 \lesssim |u|_{H^{1/2}(\Gamma_{kl})}^2 + \left(1 + \log\left(\frac{H_k}{h}\right)\right) \|u\|_{L^\infty(\Gamma_{kl})}^2$$

In the following we present additional set of technical lemmas, they are given here with proofs. The first one is a simple result which will be used to estimate the H^1 seminorm of functions from the coarse space V_0.

Lemma 6.3 *For $u \in V_0$ and C being an arbitrary constant, the following holds, i.e.*

$$|u|_{H^1(\Omega_k)}^2 \lesssim \sum_{x \in \mathscr{V}_k} |u(x) - C|^2,$$

where $\mathscr{V}_k$ is the set of all vertices of Ω_k which are not on $\partial\Omega$.

Proof Note that $u_{|\Omega_k} = \sum_{x \in \mathcal{V}_k} u(x)\phi_{x|\Omega_k}$, where ϕ_x is a discrete harmonic function which is equal to one at x, zero at $\mathcal{V}_k \backslash \{x\}$, and linear along the edges $\Gamma_{kl} \subset \partial\Omega_k$. Thus, for any constant C, we have

$$|u|^2_{H^1(\Omega_k)} = |u - C|^2_{H^1(\Omega_k)} \lesssim \sum_{x \in \mathcal{V}_k} |u(x) - C|^2 |\phi_x|^2_{H^1(\Omega_k)} \lesssim \sum_{x \in \mathcal{V}_k} |u(x) - C|^2.$$

The last inequality follows from using the fact that a discrete harmonic function has the minimal energy of all functions taking the same values on the boundary, and then by applying the standard estimate of H^1-seminorm of the coarse nodal basis function.

Definition 6.1 Let $I_H : V_h \rightarrow V_0$ be a coarse interpolant defined using function values at the vertices $\mathcal{V}$ as follows. For $u \in V_h$, $I_H u \in V_0$ implies that

$$I_H u(x) = u(x) \quad x \in \mathcal{V}.$$

Lemma 6.4 *For any $u \in V_h$, the following holds, i.e.*

$$|I_H u|^2_{H^1(\Omega_k)} \lesssim \left(1 + \log\left(\frac{H_k}{h}\right)\right) |u|^2_{H^1(\Omega_k)}.$$

Proof From Lemmas 6.3 and the discrete Sobolev like inequality, cf. e.g. Lemma 7 in [30] or Lemma 4.15 in [32], we get

$$|I_H u|^2_{H^1(\Omega_k)} \lesssim \left(1 + \log\left(\frac{H_k}{h}\right)\right) \left(H_k^{-2} \|u - C\|^2_{L^2(\Omega_k)} + |u - C|^2_{H^1(\Omega_k)}\right),$$

for any constant C. A scaling argument and a quotient space argument complete the proof.

Lemma 6.5 *Let $\Gamma_{kl} \subset \partial\Omega_k$ be an edge, and $u_{kl} \in W_k$ and $u \in V_h$ be functions such that $u_{kl}(x) = u(x) - I_H u(x)$ for $x \in \Gamma_{kl,h}$, and u_{kl} is zero on $\partial\Omega_k \backslash \Gamma_{kl}$. Then, we have*

$$|u_{kl}|^2_{H_{00}^{1/2}(\Gamma_{kl})} \lesssim \left(1 + \log\left(\frac{H_k}{h}\right)\right)^2 |u|^2_{H^1(\Omega_k)} \tag{6.1}$$

Proof Note that u_{kl} equals to $u - I_H u$ on Γ_{kl} and is zero on $\partial\Omega_k \backslash \Gamma_{kl}$, and hence by Lemma 6.2, we get

$$\|u_{kl}\|^2_{H_{00}^{1/2}(\Gamma_{kl})} \lesssim |u - I_H u|^2_{H^{1/2}(\Gamma_{kl})} + \left(1 + \log\left(\frac{H_k}{h}\right)\right) \|u - I_H u\|^2_{L^\infty(\Gamma_{kl})}. \tag{6.2}$$

The first term can be estimated using the standard trace theorem, a triangle inequality and Lemma 6.4 as follows,

$$|u - I_H u|^2_{H^{1/2}(\Gamma_{kl})} \lesssim |u|^2_{H^1(\Omega_k)} + |I_H u|^2_{H^1(\Omega_k)} \lesssim \left(1 + \log\left(\frac{H_k}{h}\right)\right) |u|^2_{H^1(\Omega_k)}. \tag{6.3}$$

For any constant C, we note that $u - I_H u = u - C - I_H(u - C)$ on Γ_{kl}, and since $I_H u$ is a linear function along Γ_{kl}, $\|I_H u\|_{L^\infty(\Gamma_{kl})} \leq \|u\|_{L^\infty(\Gamma_{kl})}$. Hence the L^∞ norm of $u - I_H u$ in (6.2) can be estimated as follows,

$$\|u - I_H u\|^2_{L^\infty(\Gamma_{kl})} \lesssim \|u - C\|^2_{L^\infty(\Gamma_{kl})} + \|I_H(u - C)\|^2_{L^\infty(\Gamma_{kl})} \lesssim \|u - C\|^2_{L^\infty(\Gamma_{kl})}$$
$$\lesssim \left(1 + \log\left(\frac{H_k}{h}\right)\right)\left(H_k^{-2}\|u - C\|^2_{L^2(\Omega_k)} + |u - C|^2_{H^1(\Omega_k)}\right),$$

where C is an arbitrary constant. The last inequality is due to the discrete Sobolev like inequality, cf. Lemma 7 in [30] or Lemma 4.15 in [32]. Finally, a scaling argument and a quotient space argument yield

$$\|u - I_H u\|^2_{L^\infty(\Gamma_{kl})} \lesssim \left(1 + \log\left(\frac{H_k}{h}\right)\right)|u|^2_{H^1(\Omega_k)}.$$

The above estimate together with the estimates (6.3) and (6.2), complete the proof. $\qquad \square$

A standard coloring argument bounds the spectral radius, and is given here in our second proposition.

Proposition 6.2 *Let $\mathscr{E}$ be the symmetric matrix of Cauchy–Schwarz coefficients, cf. (5.7), for the subspaces V_k, V_l, and V_{kl}, $k, l = 1, \ldots, N$, of the decomposition (4.1). Then,*

$$\rho(\mathscr{E}) \leq C,$$

where C is a positive constant independent of the coefficients and mesh parameters.

The third and final proposition gives an estimate of the C_0^2 such that (5.5)–(5.6) hold for any $u \in V_h$.

Proposition 6.3 *For any $u \in V_h$ there exists $u_k \in V_k$ $k = 0, 1, \ldots, N$ and $u_{kl} \in V_{kl}$ such that $u = u_0 + \sum_{k=1}^N u_k + \sum_{\Gamma_{kl} \subset \Gamma} u_{kl}$ and*

$$a(u_0, u_0) + \sum_{k=1}^N a(u_k, u_k) + \sum_{\Gamma_{kl} \subset \Gamma} a(u_{kl}, u_{kl}) \lesssim \left(1 + \log\left(\frac{H}{h}\right)\right)^2 a(u, u),$$

where $H = \max_{k=1}^N H_k$ with $H_k = \mathrm{diam}(\Omega_k)$.

Proof We first set $u_0 = I_H u \in V_0$, cf. Definition 6.1. Next, let $u_k \in V_k$ for $k = 1, \ldots, N$, be defined as $\mathscr{P}_k u_{|\overline{\Omega}_k}$ on Ω_k, be extended by zero to the rest of Ω.

Now define $w = u - u_0 - \sum_k u_k$. Note that w is discrete harmonic inside each subdomain Ω_k, since u_0 is discrete harmonic in the same way, and the sum

$$(w + u_0)_{|\overline{\Omega}_k} = u_{|\overline{\Omega}_k} - \mathscr{P}_k u_{|\overline{\Omega}_k} = \mathscr{H}_k u_{|\overline{\Omega}_k}$$

is in fact a function of W_k. Moreover,

$$w(x) = u(x) - I_H u(x) = 0 \quad x \in \mathcal{V}.$$

Consequently, w can be decomposed as follows,

$$w = \sum_{\Gamma_{kl} \subset \Gamma} u_{kl},$$

where $u_{kl} \in V_{kl}$, with $u_{|\Gamma_{kl}} = w_{|\Gamma_{kl}}$.

We now prove the inequality by considering each term at a time. For the first term, by Lemma 6.4, we see that

$$a(u_0, u_0) \lesssim \sum_{k=1}^{N} \Lambda_k |u_0|^2_{H^1(\Omega_k)} = \sum_{k=1}^{N} \Lambda_k |I_H u|^2_{H^1(\Omega_k)}$$
$$\lesssim \left(1 + \log\left(\frac{H}{h}\right)\right) \sum_{k=1}^{N} a_k(u, u) = \left(1 + \log\left(\frac{H}{h}\right)\right) a(u, u). \quad (6.4)$$

For the second term, since $\mathscr{P}_k$ is the orthogonal projection in $a_k(u, v)$, we get

$$\sum_{k=1}^{N} a(u_k, u_k) = \sum_{k=1}^{N} a_k(\mathscr{P}_k u_{|\overline{\Omega}_k}, \mathscr{P}_k u_{|\overline{\Omega}_k}) \leq \sum_{k=1}^{N} a_k(u_{|\overline{\Omega}_k}, u_{|\overline{\Omega}_k}) = a(u, u). \quad (6.5)$$

And, for the last term, let $\Gamma_{kl} \subset \Gamma$ be the edge which is common to both Ω_k and Ω_l. Note that $u_{kl} \in V_{kl}$ has support both in $\overline{\Omega}_k \cup \overline{\Omega}_l$. By Lemma 6.1, we note that

$$a(u_{kl}, u_{kl}) = \sum_{s=k,l} a_s(u_{kl}, u_{kl}) \lesssim \sum_{s=k,l} \Lambda_s |u_{kl}|^2_{H^1(\Omega_s)} \lesssim \left(\sum_{s=k,l} \Lambda_s\right) |u_{kl}|^2_{H^{1/2}_{00}(\Gamma_{kl})}.$$

Utilizing Lemma 6.5 for $s = k$ if $\Lambda_k \geq \Lambda_l$ (otherwise we take $s = l$), and (2.2), we get

$$\left(\sum_{s=k,l} \Lambda_s\right) |u_{kl}|^2_{H^{1/2}_{00}(\Gamma_{kl})} \lesssim \left(1 + \log\left(\frac{H}{h}\right)\right)^2 \Lambda_k |u|^2_{H^1(\Omega_k)}$$
$$\lesssim \left(1 + \log\left(\frac{H}{h}\right)\right)^2 a_k(u, u).$$

Combining the last two estimates, we get

$$
\sum_{\Gamma_{kl} \subset \Gamma} a(u_{kl}, u_{kl}) \lesssim \sum_{\Gamma_{kl} \subset \Gamma} \sum_{s=k,l} \left(1 + \log \left(\frac{H}{h} \right) \right)^2 a_s(u, u)
$$

$$
\lesssim \left(1 + \log \left(\frac{H}{h} \right) \right)^2 a(u, u). \tag{6.6}
$$

The proof then follows by summing (6.4), (6.5), and (6.6) together.

7 Numerical experiments

In this section, we present some numerical experiments showing the performance of the proposed methods. We consider our model problem to be defined on the domain $\Omega = [0, 1] \times [0, 1]$ with the right hand side f equal to 1, and apply the finite volume element discretization and the proposed additive Schwarz preconditioners for the solution. The domain is divided into $\frac{1}{H} \times \frac{1}{H}$ equally sized square subdomains of size $H \times H$, allowing equal number of subdomains in each direction, each of which is then further divided into $\frac{H}{h} \times \frac{H}{h}$ small equally sized squares of size $h \times h$, again with equal number of small squares in each direction. We get the primal mesh by slicing each small square into two triangles in a regular fashion, as shown in the figures.

All our numerical results have been obtained in Matlab, employing the preconditioned GMRES algorithm based on the $\| \cdot \|_a$-norm minimization, which is obtained by replacing the standard l_2 inner product with the $a(\cdot, \cdot)$ inner product in the original algorithm, see for instance [29]. For the stopping criteria, on the other hand, for simplicity, we only look at the l_2-norm of the residual. We let the algorithm to run until the l_2 norm of the initial residual is reduced by a factor of 10^6 in each test, i.e., until $\|r_i\|_2 / \|r_0\|_2 \leq 10^{-6}$ where r_0 and r_i respectively are the initial and the i-th residual vector.

The number of iterations to converge, and an estimate of the smallest eigenvalue of the symmetric part of the preconditioned operator P, that is the smallest eigenvalue of $\frac{1}{2} \left(P^t + P \right)$, which is also the most critical parameter of the two describing the GMRES convergence, are presented in the tables below. Our experiments have shown that the second parameter, which is the norm of the operator, is a constant independent of the mesh parameters H and h, and the coefficient A, which is in agreement with our analysis.

For the first numerical experiment we test the dependency of the iteration count and the smallest eigenvalue on the mesh parameters h and H, where the coefficient A is equal to $2 + \sin(\pi x) \sin(\pi y)$. The results are reported in Table 1 for the symmetric preconditioner. We observe a mild decrease in the value of the smallest eigenvalue, with a slow increase in the iteration count, as the subdomain problem size $\frac{H}{h}$ increases, suggesting a poly-logarithmic dependence as predicted in our theory.

In the second numerical experiment we perform similar tests, however, with a slightly faster varying A, namely $A = 2 + \sin(10\pi x) \sin(10\pi y)$. The results are reported in Tables 2 and 3 respectively for the symmetric and the nonsymmetric variant

Table 1 Iteration numbers and estimates of the smallest eigenvalue (in parentheses) for the symmetric preconditioner for different values of h and H

h	H					
	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$
$\frac{1}{8}$	7 (5.80e−1)					
$\frac{1}{16}$	9 (3.72e−1)	10 (5.60e−1)				
$\frac{1}{32}$	11 (2.48e−1)	13 (3.57e−1)	10 (5.56e−1)			
$\frac{1}{64}$	13 (1.76e−1)	16 (2.41e−1)	14 (3.53e−1)	10 (5.56e−1)		
$\frac{1}{128}$	15 (1.30e−1)	19 (1.72e−1)	17 (2.38e−1)	13 (3.53e−1)	10 (5.55e−1)	
$\frac{1}{256}$	16 (1.01e−1)	21 (1.28e−1)	20 (1.70e−1)	16 (2.38e−1)	13 (3.52e−1)	10 (5.54e−1)

Here $A = 2 + \sin(\pi x)\sin(\pi y)$

Table 2 Iteration numbers and estimates of the smallest eigenvalue (in parentheses) for the symmetric preconditioner for different values of h and H

h	H					
	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$
$\frac{1}{8}$	10 (5.31e−1)					
$\frac{1}{16}$	12 (3.07e−1)	13 (4.31e−1)				
$\frac{1}{32}$	14 (1.77e−1)	18 (2.42e−1)	14 (4.36e−1)			
$\frac{1}{64}$	15 (1.21e−1)	23 (1.61e−1)	18 (2.82e−1)	12 (5.20e−1)		
$\frac{1}{128}$	17 (8.93e−2)	27 (1.17e−1)	22 (1.94e−1)	16 (3.37e−1)	11 (5.53e−1)	
$\frac{1}{256}$	20 (6.94e−2)	31 (8.90e−2)	26 (1.41e−1)	20 (2.28e−1)	14 (3.57e−1)	11 (5.57e−1)

Here $A = 2 + \sin(10\pi x)\sin(10\pi y)$

Table 3 Iteration numbers and estimates of the smallest eigenvalue (in parentheses) for the nonsymmetric preconditioner for different values of h and H

| h	H					
	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$
$\frac{1}{8}$	10 (5.20e−1)					
$\frac{1}{16}$	12 (3.11e−1)	13 (4.25e−1)				
$\frac{1}{32}$	14 (1.79e−1)	18 (2.43e−1)	14 (4.44e−1)			
$\frac{1}{64}$	15 (1.21e−1)	23 (1.62e−1)	18 (2.84e−1)	12 (5.25e−1)		
$\frac{1}{128}$	17 (8.94e−2)	27 (1.17e−1)	22 (1.95e−1)	16 (3.38e−1)	11 (5.54e−1)	
$\frac{1}{256}$	20 (6.94e−2)	31 (8.90e−2)	26 (1.41e−1)	20(2.28e−1)	14 (3.57e−1)	11 (5.57e−1)

Here $A = 2 + \sin(10\pi x)\sin(10\pi y)$

of the preconditioner. In both cases, we observe convergence behavior which are similar to the one in the first experiment. We also note that the performances of the two variants of the preconditioner are almost identical.

Each diagonal (subdiagonal) in the Tables 1, 2 and 3, corresponds to a fixed subdomain problem size $\frac{H}{h}$. As we can see from the entries along each such diagonal (subdiagonal), that both the eigenvalue estimates and the iteration counts remain almost unchanged suggesting that both preconditioners are algorithmically scalable, in the sense that, for a fixed subdomain size $\frac{H}{h}$, the number of GMRES iterations is (asymptotically) independent of the number of subdomains (processors).

In the third numerical experiment, we consider an example where A is discontinuous across subdomains, which is given by $A = \alpha_1(2+\sin(10\pi x)\sin(10\pi y))$ with α_1 being a constant in each subdomain. The mesh parameters are $H = 1/8$ and $h = 1/64$. We assign the parameter α_1 in the coefficient A, the value 1 (white subdomain) or the value $\hat{\alpha}_1$ (red or shaded subdomain) in a checkerboard fashion as depicted in Fig. 2. Number of iterations and an estimates of the smallest eigenvalues for different values of $\hat{\alpha}_1$ (varying jumps) are reported in Table 4 showing that the convergence is independent of the jumps in the coefficient supporting our analysis. Again, we see an identical performance of the two variants of the algorithm.

In our fourth numerical experiment, we examine two different cases, where in the first the coefficients have jumps inside subdomains as depicted in Fig. 3, and in the second the jumps are along subdomain boundaries as depicted in Fig. 4. Their results
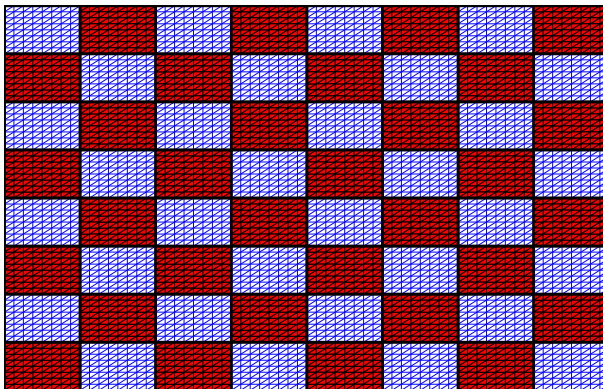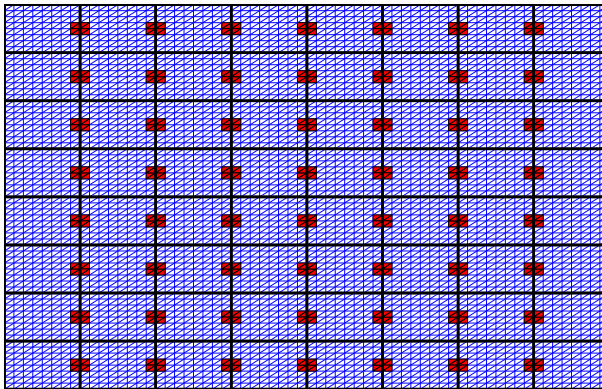

Fig. 2 Checkerboard distribution of A, with $A = \alpha_1(2 + \sin(10\pi x)\sin(10\pi y))$, where $\alpha_1 = \hat{\alpha}_1$ in the red (*shaded*) subdomains and 1 otherwise (color figure online)

Table 4 Corresponding to Fig. 2: iteration numbers and estimates of the smallest eigenvalue for different values of $\hat{\alpha}_1$, with fixed mesh sizes $h = 1/64$ and $H = 1/8$	$\hat{\alpha}_1$	Symmetric	Nonsymmetric
	10^0	23 (1.61e−1)	23 (1.62e−1)
	10^1	26 (1.61e−1)	26 (1.61e−1)
	10^2	27 (1.60e−1)	27 (1.60e−1)
	10^3	27 (1.60e−1)	27 (1.60e−1)
	10^4	27 (1.60e−1)	27 (1.60e−1)
	10^5	27 (1.60e−1)	27 (1.60e−1)
	10^6	27 (1.60e−1)	27 (1.60e−1)
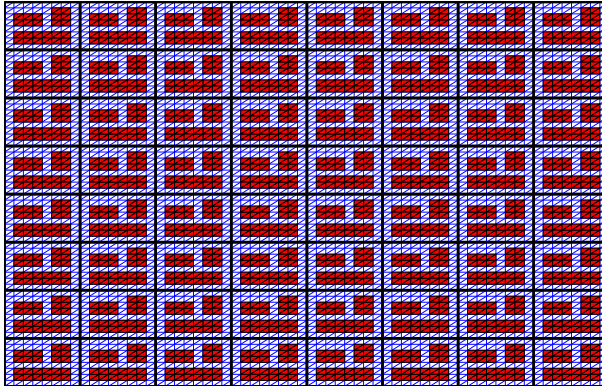
Fig. 3 Distribution of A with channels and inclusions in the interior of subdomains. $A = \alpha_1 (2 + \sin(10\pi x) \sin(10\pi y))$, where $\alpha_1 = \hat{\alpha}_1$ in the *red* (*shaded*) regions and 1 elsewhere (color figure online)

Fig. 4 Distribution of A with jumps along the subdomain boundaries. $A = \alpha_1 (2 + \sin(10\pi x) \sin(10\pi y))$, where $\alpha_1 = \hat{\alpha}_1$ in the *red* (*shaded*) regions and 1 elsewhere (color figure online)

are presented in Tables 5 and 6, respectively. As we can see from the table entries, both preconditioners are robust with respect to inclusions and channels inside subdomains, but not if the channels cross subdomain boundaries. This is not unexpected, as our coarse space include only functions that are linear along subdomain boundaries. Therefore any variations that are along the subdomain boundaries, the coarse problem cannot capture them that easy. For problems with inclusions and channels both inside and across subdomains, it may be enough to replace the coarse space with the so called multiscale coarse space, cf. [18,19,27], or in extreme cases to enrich it with eigenfunctions corresponding to bad eigen modes of some local eigenvalue problems, cf. [14,21,31]. These are subjects for future investigation.

Table 5 Corresponding to Fig. 3: iteration numbers and estimates of the smallest eigenvalue for different values of $\hat{\alpha}_1$, with fixed mesh sizes $h = 1/64$ and $H = 1/8$

$\hat{\alpha}_1$	Symmetric	Nonsymmetric
10^0	23 (1.61e−1)	23 (1.62e−1)
10^1	27 (1.05e−1)	27 (1.05e−1)
10^2	30 (8.07e−2)	30 (8.07e−2)
10^3	31 (8.87e−2)	31 (8.87e−2)
10^4	31 (8.85e−2)	31 (8.85e−2)
10^5	31 (8.85e−2)	31 (8.85e−2)
10^6	31 (8.85e−2)	31 (8.85e−2)

Table 6 Corresponding to Fig. 4: iteration numbers and estimates of the smallest eigenvalue for different values of $\hat{\alpha}_1$, with fixed mesh sizes $h = 1/64$ and $H = 1/8$

$\hat{\alpha}_1$	Symmetric	Nonsymmetric
10^0	23 (1.61e−1)	23 (1.62e−1)
10^1	28 (1.21e−1)	28 (1.21e−1)
10^2	55 (3.54e−2)	55 (3.54e−2)
10^3	99 (4.25e−3)	98 (4.25e−3)
10^4	148 (4.33e−4)	145 (4.33e−4)
10^5	185 (4.34e−5)	178 (4.34e−5)
10^6	215 (4.34e−6)	210 (4.34e−6)

References

1. Bi, C., Chen, W.: Mortar finite volume element method with Crouzeix–Raviart element for parabolic problems. Appl. Numer. Math. **58**, 1642–1657 (2008)
2. Bi, C., Ginting, V.: A residual-type a posteriori error estimate of finite volume element method for a quasi-linear elliptic problem. Numer. Math. **114**, 107–132 (2009)
3. Bi, C., Ginting, V.: Finite-volume-element method for second-order quasilinear elliptic problems. IMA J. Numer. Anal. **31**(3), 1062–1089 (2011)
4. Bi, C., Rui, H.: Uniform convergence of finite volume element method with Crouzeix–Raviart element for non-selfadjoint and indefinite elliptic problems. J. Comput. Appl. Math. **200**, 555–565 (2007)
5. Bjørstad, P.E., Widlund, O.B.: Iterative methods for the solution of elliptic problems on regions partitioned into substructures. SIAM J. Numer. Anal. **23**(6), 1097–1120 (1986)
6. Braess, D.: Finite Elements, 3d edn. Cambridge University Press, Cambridge (2007). doi:10.1017/CBO9780511618635. Theory, fast solvers, and applications in elasticity theory. Translated from the German by Larry L. Schumaker
7. Brenner, S.C.: The condition number of the Schur complement in domain decomposition. Numer. Math. **83**(2), 187–203 (1999)
8. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods. Texts in Applied Mathematics, vol. 15, 3rd edn. Springer, New York (2008). doi:10.1007/978-0-387-75934-0

9. Cai, X.C., Widlund, O.B.: Domain decomposition algorithms for indefinite elliptic problems. SIAM J. Sci. Stat. Comput. **13**(1), 243–258 (1992). doi:10.1137/0913013

10. Chartier, T., Falgout, R.D., Henson, V.E., Jones, J., Manteuffel, T., McCormick, S., Ruge, J., Vassilevski, P.S.: Spectral AMGe (ρAMGe). SIAM J. Sci. Comput. **25**(1), 1–26 (2003). doi:10.1137/S106482750139892X

11. Chatzipantelidis, P.: A finite volume method based on the Crouzeix–Raviart element for elliptic PDE's in two dimensions. Numer. Math. **82**(3), 409–432 (1999). doi:10.1007/s002110050425

12. Chen, Z.: Domain decomposition algorithms for mixed methods for second-order elliptic problems. Netw. Heterog. Media **1**(4), 689–706 (2006)

13. Chou, S.H., Huang, J.: A domain decomposition algorithm for general covolume methods for elliptic problems. J. Numer. Math. **11**(3), 179–194 (2003). doi:10.1163/156939503322553072

14. Efendiev, Y., Galvis, J., Lazarov, R., Willems, J.: Robust domain decomposition preconditioners for abstract symmetric positive definite bilinear forms. ESAIM Math. Model. Numer. Anal. **46**(5), 1175–1199 (2012). doi:10.1051/m2an/2011073

15. Eisenstat, S.C., Elman, H.C., Schultz, M.H.: Variational iterative methods for nonsymmetric systems of linear equations. SIAM J. Numer. Anal. **20**(2), 345–357 (1983). doi:10.1137/0720023

16. Ewing, R.E., Lazarov, R., Lin, T., Lin, Y.: Mortar finite volume element approximations of second order elliptic problems. East-West J. Numer. Math. **8**(2), 93–110 (2000)

17. Ewing, R.E., Lin, T., Lin, Y.: On the accuracy of the finite volume element method based on piecewise linear polynomials. SIAM J. Numer. Anal. **39**(6), 1865–1888 (2002). doi:10.1137/S0036142900368873

18. Graham, I.G., Lechner, P.O., Scheichl, R.: Domain decomposition for multiscale PDEs. Numer. Math. **106**(4), 589–626 (2007). doi:10.1007/s00211-007-0074-1

19. Hajibeygi, H., Bonfigli, G., Hess, M., Jenny, P.: Iterative multiscale finite-volume method. J. Comput. Phys. **227**(19), 8604–8621 (2008)

20. Huang, J., Xi, S.: On the finite volume element method for general self-adjoint elliptic problems. SIAM J. Numer. Anal. **35**(5), 1762–1774 (1998). doi:10.1137/S0036142994264699

21. Klawonn, A., Radtke, P., Rheinbach, O.: FETI-DP methods with an adaptive coarse space. SIAM J. Numer. Anal. **53**(1), 297–320 (2015)

22. Le Tallec, P., Mandel, J., Vidrascu, M.: A Neumann–Neumann domain decomposition algorithm for solving plate and shell problems. SIAM J. Numer. Anal. **35**(2), 836–867 (1998)

23. Lin, Y., Liu, J., Yang, M.: Finite volume element methods: an overview on recent developments. Int. J. Numer. Anal. Mod. **4**(1), 14–34 (2013)

24. Loneland, A., Marcinkowski, L., Rahman, T.: Additive average Schwarz method for the Crouzeix–Raviart finite volume element discretization of elliptic problems. In second round of review (available online arXiv:1405.3494 [math.NA])

25. Loneland, A., Marcinkowski, L., Rahman, T.: Edge based Schwarz methods for the Crouzeix–Raviart finite volume element discretization of elliptic problems. Electron. Trans. Numer. Anal. **44**, 443–461 (2015)

26. Mathew, T.P.A.: Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations. Lecture Notes in Computational Science and Engineering, vol. 61. Springer-Verlag, Berlin (2008)

27. Nordbotten, J.M., Bjørstad, P.: On the relationship between the multiscale finite-volume method and domain decomposition preconditioners. Comput. Geosci. **12**, 367–376 (2008)

28. Saad, Y., Schultz, M.H.: GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM J. Sci. Stat. Comput. **7**(3), 856–869 (1986)

29. Sarkis, M., Szyld, D.B.: Optimal left and right additive Schwarz preconditioning for minimal residual methods with Euclidean and energy norms. Comput. Methods Appl. Mech. Eng. **196**, 1612–1623 (2007)

30. Smith, B.F., Bjørstad, P.E., Gropp, W.D.: Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations. Cambridge University Press, Cambridge (1996)

31. Spillane, N., Dolean, V., Hauret, P., Nataf, F., Pechstein, C., Scheichl, R.: Abstract robust coarse spaces for systems of PDEs via generalized eigenproblems in the overlaps. Numer. Math. **126**(4), 741–770 (2014). doi:10.1007/s00211-013-0576-y

32. Toselli, A., Widlund, O.: Domain Decomposition Methods—Algorithms and Theory. Springer Series in Computational Mathematics, vol. 34. Springer-Verlag, Berlin (2005)

33. Xie, H., Xu, X.: Mass conservative domain decomposition preconditioners for multiscale finite volume method. Multiscale Model. Simul. **12**(4), 1667–1690 (2014)
34. Yang, M.: A posteriori error analysis of nonconforming finite volume elements for general second-order elliptic PDEs. Numer. Methods PDEs **27**, 277–291 (2011)
35. Zhang, S.: On domain decomposition algorithms for covolume methods for elliptic problems. Comput. Methods Appl. Mech. Eng. **196**(1–3), 24–32 (2006). doi:10.1016/j.cma.2005.11.017