ELSEVIER

Applied Numerical Mathematics 28 (1998) 17–35

APPLIED NUMERICAL MATHEMATICS

$\mathcal{V}$-cycle Galerkin–multigrid methods for nonconforming methods for nonsymmetric and indefinite problems

Zhangxin Chen [a,*,1], Do Y. Kwak [b,2]

[a] Department of Mathematics, Southern Methodist University, P.O. Box 750156, Dallas, TX 75275, USA

[b] Department of Mathematics, Korea Advanced Institute of Science and Technology, Taejon, South Korea 305-701

Abstract

In this paper we analyze a class of $\mathcal{V}$-cycle multigrid methods for discretizations of second-order nonsymmetric and/or indefinite elliptic problems using nonconforming P_1 and rotated Q_1 finite elements. These multigrid methods are based on the so-called Galerkin approach where the quadratic forms over coarse grids are constructed from the quadratic form on the finest grid and iterated coarse-to-fine grid operators. The analysis shows that these $\mathcal{V}$-cycle multigrid iterations with one smoothing on each level converge at a uniform rate provided that the coarsest level in the multilevel iterations is sufficiently fine (but independent of the number of multigrid levels). Various types of smoothers for the nonsymmetric and indefinite problems are considered and analyzed. The theory presented here also applies to mixed finite element methods for the nonsymmetric and indefinite problems.

Keywords: Mixed method; Nonconforming method; Finite elements; Multigrid method; Convergence; Nonsymmetric and/or indefinite problems

1. Introduction

This is the third paper of a series where we develop and analyze a class of multigrid methods for discretizations of partial differential problems using nonconforming and mixed finite elements. This new class of multigrid methods, introduced in the first paper [13], is based on the so-called Galerkin approach where the quadratic forms over coarse grids are constructed from the quadratic form on the finest grid and iterated coarse-to-fine grid operators. Its convergence for both the $\mathcal{V}$- and $\mathcal{W}$-cycle multigrid methods with one smoothing iteration on each level has been shown. In the second paper [18], the convergence of these new multigrid methods for partial differential problems without regularity assumptions has

* Corresponding author. E-mail: zchen@dragon.math.smu.edu.

1 Partly supported by National Science Foundation grant DMS-9626179.

2 This research was accomplished in part during the second author's stay at the Southern Methodist University. The second author is also in part supported by KOSEF contract 971-0106-034-2. E-mail: dykwak@math.kaist.ac.kr.

PII: S0168-9274(98)00036-1

been obtained. In contrast, the usual multigrid methods for nonconforming finite elements use discrete equations on all levels which are defined by the same discretization. Furthermore, for these usual methods only the $\mathcal{W}$-cycle multigrid methods have been shown to converge under the assumption that the number of smoothing iterations on all levels is sufficiently large (see the extensive references in [13]).

In this paper we analyze this new class of $\mathcal{V}$-cycle multigrid methods for discretizations of second-order nonsymmetric and/or indefinite elliptic problems using nonconforming P_1 and rotated Q_1 finite elements. Multigrid methods for solving nonsymmetric and/or indefinite problems by nonconforming finite elements have been first introduced and analyzed in [19], but these multigrid iterations use conforming coarse-grid corrections. While the convergence of the multigrid $\mathcal{V}$-cycle methods with one smoothing has been shown for this conforming coarse-grid correction approach, the analysis only applies to the P_1-nonconforming finite element [19]. The reason for this is that only the nonconforming P_1 element contains the conforming P_1 element as a subspace over the same triangulation, which is used in the coarse-grid corrections; other nonconforming elements do not contain any reasonable conforming subspaces over the same triangulation. We here prove convergence estimates for the $\mathcal{V}$-cycle multigrid methods for the new approach for the nonsymmetric and indefinite problems under rather weak assumptions (e.g., the domain need not be convex and problems need not have regularity assumptions) for both the P_1 and rotated Q_1 finite elements. The rotated Q_1 finite element has applications to the Stokes problem [26], the problem related to the deformation of martensitic crystals with microstructure [23], and semi-conductor modeling [11].

A variety of smoothers are considered and analyzed here. These smoothers are the variants of those for conforming finite element methods [7]. One type of smoothers is defined in terms of the corresponding symmetric problem, and the other type is entirely based on the original nonsymmetric and indefinite problem. These two types of smoothers include point and line Jacobi and Gauss–Seidel iterations. The analysis here assumes that the nonsymmetric/indefinite terms are a "compact perturbation"; the convection-dominated problems are not studied here. Also, due to the equivalence between nonconforming and mixed finite element methods (see [2,3,10,12,17] for symmetric problems and [19] for nonsymmetric problems), all the analysis throughout this paper directly applies to the mixed methods. Finally, we mention that there has been intensive research on multigrid methods for nonsymmetric and indefinite problems using conforming finite elements (see the references in [7]).

The rest of the paper is organized as follows. In Section 2 we state the continuous problem and its corresponding discrete system. Then, in Section 3 we describe multigrid methods for nonconforming methods and carry out the convergence analysis. Finally, in Sections 4 and 5 we apply the theory to the P_1 and rotated Q_1 finite elements, respectively. We mention that there are extensive numerical results available for discretizations of nonsymmetric problems using nonconforming and mixed finite elements by means of the present and usual approaches [13,19]. These numerical results have shown convergence of these approaches. That is part of the reason that we are interested in the theoretical proof. However, with the usual approach we are not able to prove the convergence. With the present approach, we can show it here for both the P_1 and rotated Q_1 nonconforming elements for second-order nonsymmetric and indefinite problems. Also, the problem under consideration has many practical applications such as those to flow of fluids in porous media [15].

2. Preliminaries

In this section we consider as our model problem the following equation:

$$\begin{aligned} -\nabla\cdot(\mathcal{A}\nabla u)+\mathcal{B}\cdot\nabla u+cu&=f \quad \text{in } \Omega,\\ u&=0 \quad \text{on } \partial\Omega, \end{aligned} \tag{2.1}$$

where $\Omega\subset\mathbb{R}^2$ is a simply connected bounded polygonal domain with the boundary $\partial\Omega$, $f\in L^2(\Omega)$, the coefficient $\mathcal{A}\in(L^\infty(\Omega))^{2\times2}$ satisfies the uniformly positive definite condition

$$\xi^{\mathrm T}\mathcal{A}(x)\xi\geqslant a_0\xi^{\mathrm T}\xi,\quad x\in\Omega,\ \xi\in\mathbb{R}^n, \tag{2.2}$$

and the coefficients $\mathcal{B}$ and c are bounded. Other conditions on $\mathcal{A}$ and $\mathcal{B}$ will be stated later. Finally, we assume that (2.1) has a unique solution.

Problem (2.1) is recast in weak form as follows. The bilinear form $a(\cdot,\cdot)$ is given by

$$a(v,w)=(\mathcal{A}\nabla v,\nabla w)+(\mathcal{B}\cdot\nabla v,w)+(cv,w),\quad v,w\in H^1(\Omega),$$

where $(\cdot,\cdot)$ denotes the $L^2(\Omega)$ or $(L^2(\Omega))^2$ inner product, as appropriate. The solution $u\in H_0^1(\Omega)$ of (2.1) then satisfies

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

Associated with $a(\cdot,\cdot)$, we also introduce the symmetric positive definite form $\widehat{a}(\cdot,\cdot)$ by

$$\widehat{a}(v,w)=(\mathcal{A}\nabla v,\nabla w)+(v,w),\quad v,w\in H^1(\Omega).$$

The difference form is indicated by

$$D(v,w)=a(v,w)-\widehat{a}(v,w). \tag{2.4}$$

For $0<h<1$, let $\mathcal{E}_h$ be a partition of Ω into triangles or rectangles of size h, and define V_h to be the space of the nonconforming P_1 [21] (see the definition in Section 4) or rotated Q_1 [2,11,26] (see Section 5) finite elements. Associated with V_h, we define a mesh-dependent form $a_h(\cdot,\cdot)$ by

$$a_h(v,w)=\sum_{E\in\mathcal{E}_h}\{(\mathcal{A}\nabla v,\nabla w)_E+(\mathcal{B}\cdot\nabla v,w)_E\}+(cv,w),\quad v,w\in V_h\cup H_0^1(\Omega),$$

where $(\cdot,\cdot)_E$ is the $L^2(E)$ inner product. The corresponding symmetric form is denoted by $\widehat{a}_h(\cdot,\cdot)$. The nonconforming finite element solution $u_h\in V_h$ of (2.1) is given by

$$a_h(u_h,v)=(f,v)\quad\forall v\in V_h. \tag{2.5}$$

The norm induced by $(\widehat{a}_h(v,v))^{1/2}$ for $v\in V_h\cup H_0^1(\Omega)$ is equivalent to the norm

$$\left(\sum_{E\in\mathcal{E}_h}\|\nabla v\|^2_{L^2(E)}+\|v\|^2_{L^2(\Omega)}\right)^{1/2}.$$

Thus, we define

$$\|v\|_{\mathcal{E},h}=\widehat{a}_h(v,v)^{1/2}\quad\forall v\in V_h\cup H_0^1(\Omega).$$

Let us note the inequality

$$|a_h(v,w)|\leqslant C\|v\|_{\mathcal{E},h}\|w\|_{\mathcal{E},h}\quad\forall v,w\in V_h\cup H_0^1(\Omega), \tag{2.6}$$

where (and below) C, with or without a subscript, denotes a generic constant independent on h.

We now state the next theorem [19].

Theorem 2.1. *Let V_h be the space of the nonconforming P_1 or rotated Q_1 finite elements. Then problem* (2.5) *has a unique solution for h sufficiently small.*

3. The multigrid method

To develop a multigrid method for (2.5), we need to assume a structure to our family of partitions. Let h_0 and $\mathcal{E}_{h_0} = \mathcal{E}_0$ be given. For each integer $0 < k \leqslant K$, let $h_k = 2^{-k}h_0$ and $\mathcal{E}_{h_k} = \mathcal{E}_k$ be constructed by connecting the midpoints of the edges of the element in $\mathcal{E}_{k-1}$, and let $\mathcal{E}_h = \mathcal{E}_K$ be the finest grid. In this and the following sections, we replace subscript h_k simply by subscript k.

Let the mesh size of $\mathcal{E}_0$ be d_0; then, by similarity, the mesh size of $\mathcal{E}_k$ is $2^{-k}d_0$. From Theorem 2.1, for (2.5) to be well behaved, the approximation grid must be sufficient fine. As in the conforming case [7], we shall require that the coarsest grid in the multilevel method be sufficient fine. Toward that end, let the coarsest grid size be determined by an integer L. Then the space V_k has a mesh size of $h_k = 2^{-L-k}d_0 = 2^{-k}h_0$, $k = 0, \ldots, K$.

3.1. Notation

Let V_k be the space of either the nonconforming P_1 or rotated Q_1 finite elements on each level $k = 0, \ldots, K$, and let $a_k(\cdot,\cdot)$ be the quadratic form on $V_k \times V_k$, as defined in the last section. The corresponding symmetric form is indicated by $\widehat{a}_k(\cdot,\cdot)$, $k = 0, \ldots, K$. Also, for the nonconforming P_1 or rotated Q_1 finite elements, let $(\cdot,\cdot)_k$ be the usual discrete L^2 inner product, and $I_k : V_{k-1} \to V_k$ be the standard averaging coarse-to-fine grid operator (see Sections 4 and 5).

We now introduce the iterates of I_k [13,20,25]

$$H_k^K = I_K \cdots I_{k+1} : V_k \to V_K, \quad k = 0, \ldots, K,$$

with $H_K^K = I$ (the identity operator), and the quadratic form $b_k(\cdot,\cdot)$ on $V_k \times V_k$:

$$b_k(v, w) = a_K\left(H_k^K v, H_k^K w\right) \quad \forall v, w \in V_k, \qquad k = 0, \ldots, K.$$

The symmetric positive definite quadratic form $\widehat{b}_k(\cdot,\cdot)$ on $V_k \times V_k$ is similarly defined by

$$\widehat{b}_k(v, w) = \widehat{a}_K\left(H_k^K v, H_k^K w\right) \quad \forall v, w \in V_k, \qquad k = 0, \ldots, K.$$

The norms corresponding to $(\cdot,\cdot)_k$, $\widehat{a}_k(\cdot,\cdot)$, and $\widehat{b}_k(\cdot,\cdot)$ will be denoted by $\|\cdot\|_k$, $\|\cdot\|_{\mathcal{E},k}$, and $\|\cdot\|_{1,k}$, respectively. It follows [13,20,25] that there are positive constants C_1 and C_2 such that

$$C_1\|v\|_{1,k} \leqslant \|v\|_{\mathcal{E},k} \leqslant C_2\|v\|_{1,k} \quad \forall v \in V_k, \qquad k = 0, \ldots, K, \tag{3.1}$$

and

$$C_1\|v\|_{\mathcal{E},k} \leqslant \left\|H_k^K v\right\|_{\mathcal{E},K} \leqslant C_2\|v\|_{\mathcal{E},k} \quad \forall v \in V_k, \qquad k = 0, \ldots, K. \tag{3.2}$$

As in the previous section, we denote the difference form by

$$D_k(v, w) = a_k(v, w) - \widehat{a}_k(v, w) \quad \forall v, w \in V_k.$$

Also, define

$$d_k(v,w) = D_K\big(H_k^K v, H_k^K w\big) \quad \forall v, w \in V_k, \qquad k = 0,\dots,K,$$

i.e.,

$$d_k(v,w) = b_k(v,w) - \widehat{b}_k(v,w) \quad \forall v, w \in V_k.$$

We now assume that there are positive constants C such that

$$\big|D_k(v,w)\big| \leqslant C\|v\|_{\mathcal{E},k}\|w\|_k \quad \forall v, w \in V_k, \qquad k = 0,\dots,K, \tag{3.3}$$

and

$$\big|D_k(v,w)\big| \leqslant C\|v\|_k\|w\|_{\mathcal{E},k} \quad \forall v, w \in V_k, \qquad k = 0,\dots,K. \tag{3.4}$$

Note that (3.3) directly follows from the definition of $D_k(v,w)$:

$$D_k(v,w) = \sum_{E\in\mathcal{E}_k} (\mathcal{B}\cdot\nabla v, w)_E + \big((c-1)v, w\big).$$

While (3.4) is trivial for conforming finite elements, it is not so straightforward for nonconforming elements. It has been shown in [19] for the P_1 nonconforming element. For the rotated Q_1 nonconforming element, it will be proven in Section 5. Due to (3.1), (3.3) and (3.4) also hold for the norm $\|\cdot\|_{1,k}$ in place of $\|\cdot\|_{\mathcal{E},k}$. Finally, if (3.3) and (3.4) hold for D_k, so do they for d_k by (3.1) and (3.2).

The following operator $\Pi_K^k : V_K \to V_k$ plays a crucial role in the subsequent analysis:

$$b_k\big(\Pi_K^k v, w\big) = b_K\big(v, H_k^K w\big), \quad v \in V_K,\ w \in V_k,$$

for $k = 0,\dots,K$; i.e., Π_K^k is the adjoint operator of H_k^K with respect to $b_k(\cdot,\cdot)$. If the solution of (2.1) satisfies regularity estimates of the form

$$\|u\|_{1+\alpha} \leqslant C\|f\|_{-1+\alpha}, \quad 0 < \alpha \leqslant 1, \tag{3.5}$$

then it follows [13,14] that there exists a constant $\overline{h}$ such that for $h_k \leqslant \overline{h}$ $(k = 0,\dots,K)$,

$$\big\|(I - H_k^K \Pi_K^k)v\big\|_K \leqslant C h_k^\alpha \big\|(I - H_k^K \Pi_K^k)v\big\|_{1,K} \quad \forall v \in V_K, \tag{3.6}$$

and

$$\big\|H_k^K \Pi_K^k v\big\|_{1,K} \leqslant C\|v\|_{1,K} \quad \forall v \in V_K, \qquad k = 0,\dots,K. \tag{3.7}$$

In the case where regularity estimates of the form of (3.6) are not known to hold, it can be shown by combining the techniques in [27] for conforming finite elements and the ideas in [13] for handling $H_k^K \Pi_K^k$ for nonconforming elements that, given $\varepsilon > 0$, there exists an $\overline{h}(\varepsilon) > 0$ such that for $0 < h_k \leqslant \overline{h}$,

$$\big\|(I - H_k^K \Pi_K^k)v\big\|_K \leqslant \varepsilon \big\|(I - H_k^K \Pi_K^k)v\big\|_{1,K} \quad \forall v \in V_K, \qquad k = 0,\dots,K. \tag{3.8}$$

The above ε will appear in our later convergence result. We observe that ε can be taken arbitrarily small if L is sufficiently large. Hence L will be sufficiently large so that Theorem 2.1, (3.7), and (3.8) hold.

We remark that the hard part in these inequalities (3.1)–(3.8) is the verification of the upper bound in (3.2). That is why we are here restricted to the P_1 and rotated Q_1 nonconforming elements for which it was verified under some conditions on $\mathcal{E}_0$ [20,25] (see also Sections 4 and 5). This upper bound for other nonconforming elements was discussed in [13].

3.2. The multigrid method

Note that while on the finest level

$$b_K(v,w)=a_K(v,w)\quad \forall v,w\in V_K,$$

$b_k(\cdot,\cdot)$ are different from $a_k(\cdot,\cdot)$ on all the lower levels. The multigrid method below will be defined in terms of $b_k(\cdot,\cdot)$ instead of $a_k(\cdot,\cdot)$. Hence this approach differs from the usual one in that the usual nonconforming multigrid methods are defined in terms of $a_k(\cdot,\cdot)$ [19]. This idea has been exploited for a long time in the context of finite difference methods (see the references in [22,24]). For conforming finite elements, these two quadratic forms are the same.

We define the discretization operator $A_k: V_k \to V_k$ on level k given by

$$(A_k v, w)_k = b_k(v,w),\quad v,w\in V_k,\qquad k=0,\dots,K.$$

Also, define the operators $P_{k-1}: V_k \to V_{k-1}$ and $P^0_{k-1}: V_k\to V_{k-1}$ by

$$b_{k-1}(P_{k-1}v,w)=b_k(v,I_kw),\quad w\in V_{k-1},\qquad k=1,\dots,K,$$

and

$$\left(P^0_{k-1}v,w\right)_{k-1}=(v,I_kw)_k,\quad w\in V_{k-1},\qquad k=1,\dots,K.$$

The operators $\widehat{A}_k$ and $\widehat{P}_{k-1}$ are similarly defined in terms of the $\widehat{b}_k$ form in place of b_k. Finally, let $R_k: V_k\to V_k$ for $k=1,\dots,K$ be linear operators; examples of R_k will be given in Sections 3.3 and 3.4 below. On V_0, let $R_0=A_0^{-1}$; i.e., we solve exactly on the coarsest level. Now we define a simplest $\mathcal{V}$-cycle multigrid method only with pre-smoothing. Other types of methods with just post-smoothing or both pre- and post-smoothing can be analyzed analogously. Also, the analysis of the $\mathcal{W}$-cycle can be given similarly.

The following method iteratively defines a multigrid operator $B_k: V_k\to V_k$:

Multigrid Method (MG). Set $B_0=A_0^{-1}$. For $0<k\leqslant K$, assume that B_{k-1} has been defined and define B_kg for $g\in V_k$ by

(1) Set $x_k=R_kg$.

(2) Define $B_kg=x_k+I_kq$, where $q\in V_{k-1}$ is given by

$$q=B_{k-1}P^0_{k-1}(g-A_kx_k).$$

We shall now write the multigrid operator B_k in a product form, which is a fundamental ingredient in the subsequent convergence analysis.

Let $g=A_kx$. It is clear from MG that

$$q=B_{k-1}P^0_{k-1}A_k(I-R_kA_k)x,$$

so that, by the facts that $P^0_{k-1}A_k=A_{k-1}P_{k-1}$ and $P_{k-1}I_k=I$ on V_{k-1} [13,18], we have

$$q=B_{k-1}A_{k-1}P_{k-1}(I-R_kA_k)x.$$

That is,

$$I-B_kA_k=(I-I_kB_{k-1}A_{k-1}P_{k-1})J_k, \tag{3.9}$$

where $J_k=I-R_kA_k$. Also, by the definition of Π^k_K we see that

$$\Pi^{k-1}_K=P_{k-1}\Pi^k_K\quad\text{and}\quad \Pi^{k-1}_KH^K_k=P_{k-1}.$$

Then it follows from (3.9) that

$$\begin{aligned} I - H_k^K B_k A_k \Pi_K^k &= I - H_k^K \Pi_K^k + H_k^K (I - B_k A_k) \Pi_K^k \\ &= I - H_k^K \Pi_K^k + H_k^K (I - I_k B_{k-1} A_{k-1} P_{k-1}) J_k \Pi_K^k \\ &= (I - H_k^K I_k B_{k-1} A_{k-1} \Pi_K^{k-1})(I - H_k^K \Pi_K^k + H_k^K J_k \Pi_K^k). \end{aligned}$$

Now, set $S^k = H_k^K R_k A_k \Pi_K^k$. Note that S^k maps V_K into V_K, but exploits the coarse space V_k. That is why we here use the superscript k to differ from the subscript used in other operators such as I_k, which have a range in V_k. Then we obtain

$$I - H_k^K B_k A_k \Pi_K^k = (I - H_{k-1}^K B_{k-1} A_{k-1} \Pi_K^{k-1})(I - S^k).$$

Finally, let $E^k = I - H_k^K B_k A_k \Pi_K^k$ and $E = E^K$, so $E^k = E^{k-1}(I - S^k)$ and

$$E = (I - S^0) \cdots (I - S^K). \tag{3.10}$$

The same remark for the notation on S^k applies to E^k as well. A product form for symmetric problems has been described in [13,18].

3.3. Smoothers based on $\widehat{A}_k$

The smoothers presented in this and next subsections are the variants of those for the conforming finite element method (see, e.g., the references in [7]). In this subsection we describe three smoothers denoted by $\widehat{R}_k$, which are based on the symmetric problem. We list three conditions on these smoothing operators, and then we give convergence estimates for MG with $R_k = \widehat{R}_k$.

Let $\widehat{\Pi}_K^k$ be the adjoint operator of H_k^K with respect to the $\widehat{b}_k$ form; i.e.,

$$\widehat{b}_k(\widehat{\Pi}_K^k v, w) = \widehat{b}_K(v, H_k^K w) \quad \forall w \in V_k.$$

Accordingly, set $\widehat{S}^k = H_k^K \widehat{R}_k \widehat{A}_k \widehat{\Pi}_K^k$. Now, the first assumption is standard:

$$\frac{(v, v)_k}{\lambda_k} \leqslant C_R (\overline{R}_k v, v)_k \quad \forall v \in V_k, \tag{3.11}$$

where the constant C_R is independent of k, $\overline{R}_k = (I - \widehat{J}_k^* \widehat{J}_k)\widehat{A}_k^{-1}$ with $\widehat{J}_k = I - \widehat{R}_k \widehat{A}_k$ and $*$ being the adjoint with respect to the inner product $\widehat{b}_k(\cdot,\cdot)$, and λ_k is the largest eigenvalue of $\widehat{A}_k$. The second assumption is also standard:

$$\widehat{b}_K(\widehat{S}^k v, \widehat{S}^k v) \leqslant \theta \widehat{b}_K(\widehat{S}^k v, v) \quad \forall v \in V_K, \tag{3.12}$$

where the constant θ (independent of k) is required to be less than two. Note that if (3.12) holds, then for $v \in V_K$,

$$\begin{aligned} \widehat{b}_K((I - \widehat{S}^k)v, (I - \widehat{S}^k)v) &= \widehat{b}_K(v, v) - 2\widehat{b}_K(\widehat{S}^k v, v) + \widehat{b}_K(\widehat{S}^k v, \widehat{S}^k v) \\ &\leqslant \widehat{b}_K(v, v) - (2 - \theta)\widehat{b}_K(\widehat{S}^k v, v) \\ &\leqslant \widehat{b}_K(v, v), \end{aligned} \tag{3.13}$$

so that the operator norm of $I - \widehat{S}^k$ in terms of $\widehat{b}_K$ is bounded by one. The last assumption is that for $k > 0$, there exists a constant C_S such that

$$(\widehat{S}^k v, \widehat{S}^k v)_K \leqslant C_S \lambda_k^{-1} \widehat{b}_K(\widehat{S}^k v, v) \quad \forall v \in V_K. \tag{3.14}$$

Note that both (3.12) and (3.14) hold on V_K. Also, we remark that if assumptions (3.11), (3.12), and (3.14) are satisfied for a smoother R_k, so are they for its adjoint R_k^{T} with respect to the inner product $(\cdot,\cdot)_k$. This implies that assumption (3.11) is satisfied for $\overline{R}_k = (I - \widehat{J}_k \widehat{J}_k^*)\widehat{A}_k^{-1}$ and assumptions (3.12) and (3.14) for $(\widehat{S}^k)^*$.

Example 1. The simplest smoother is given in this example:

$$\widehat{R}_k = \lambda_k^{-1} I,$$

where λ_k is defined as in (3.11). For this example, it is trivial to see that (3.11), (3.12), and (3.14) hold with $C_R = \theta = C_S = 1$. Obviously, (3.11), (3.12), and (3.14) also hold with any $\overline{\lambda}_k$ replacing λ_k in this example provided that it satisfies that $\lambda_k \leqslant \overline{\lambda}_k \leqslant C\lambda_k$. In this case, (3.11) is valid with $C_R = \overline{\lambda}_k/\lambda_k$ and (3.14) holds with $C_S = \lambda_k/\overline{\lambda}_k$.

The next two smoothers are defined in terms of subspace decompositions. Toward that end, we define

$$V_k = \sum_{j=1}^{l(k)} V_{j,k},$$

where $V_{j,k}$ is the one-dimensional subspace spanned by a nodal (respectively, edge) basis function for the P_1 element (respectively, for the rotated Q_1 element) or the one spanned by the nodal (respectively, edge) basis functions along a line, and $l(k)$ is the number of such spaces. These spaces satisfy the following property:

$$\|v\|_k \leqslant Ch_k \|v\|_{1,k} \quad \forall v \in V_{i,k}. \tag{3.15}$$

Example 2. This example defines an additive smoother:

$$\widehat{R}_k = \gamma \sum_{j=1}^{l(k)} \widehat{A}_{j,k}^{-1} Q_{j,k},$$

where $\widehat{A}_{j,k} : V_{j,k} \to V_{j,k}$ is the symmetric discretization operator on $V_{j,k}$ given by

$$(\widehat{A}_{j,k} v, \varphi) = \widehat{b}_k(v,\varphi) \quad \forall \varphi \in V_{j,k},$$

$Q_{j,k} : V_k \to V_{j,k}$ is the projection operator on $V_{j,k}$ with respect to the inner product $(\cdot,\cdot)_k$, and the constant γ is a scaling factor which is chosen to ensure that the smoothing property (3.12) is satisfied [6]. Note that $\widehat{R}_k$ is symmetric with respect to the inner product $(\cdot,\cdot)_k$.

Example 3. The final example in this subsection determines a multiplicative smoother. Given $g \in V_k$, we define

(1) Set $x_0 = 0$.
(2) Determine x_i, for $i = 1, \dots, l(k)$, by

$$x_i = x_{i-1} + \widehat{A}_{j,k}^{-1} Q_{j,k}\big(g - \widehat{A}_k x_{i-1}\big).$$

(3) Set $\widehat{R}_k g = x_{l(k)}$.

Applying the support properties of the basis functions, the subspaces $V_{j,k}$ satisfy a so-called limited interaction property with respect to $b_k(\cdot,\cdot)$ for each j, as seen in [13]. Thus, assumptions (3.11) and (3.12) can be shown in a standard way as in the conforming case [5,6]. Also, the proof of (3.14) in [7] for the conforming case just uses the same limited interaction property with respect to the inner product $(\cdot,\cdot)_k$. Again, this property obviously holds for the nonconforming P_1 and rotated Q_1 elements, so (3.14) can be proven as in [7].

3.4. Smoothers based on A_k

In this subsection we give three examples of smoothers directly based on A_k. Example 4 below corresponds to the first example, and Examples 5 and 6 are closely related to Examples 2 and 3, respectively.

Example 4. We define

$$R_k = \bar{\lambda}_k^{-2} A_k^{\mathrm{T}},$$

where $\bar{\lambda}_k$ is given as in Example 1 and A_k^{T} is the adjoint operator of A_k with respect to the inner product $(\cdot,\cdot)_k$. This smoother was originally analyzed in [4].

Example 5. We define

$$R_k = \gamma \sum_{j=1}^{l(k)} A_{j,k}^{-1} Q_{j,k},$$

where $A_{j,k} : V_{j,k} \to V_{j,k}$ is the discretization operator on $V_{j,k}$ given by

$$(A_{j,k} v, \varphi) = b_k(v, \varphi) \quad \forall \varphi \in V_{j,k}, \tag{3.16}$$

and $Q_{j,k} : V_k \to V_{j,k}$ and γ are as in Example 2. For $A_{j,k}$ to be invertible, we need h_k to be small enough (e.g., $0 < h_k \leqslant \bar{h}$, as in Section 3.1) so that (3.16) is well defined. For the subsequent analysis, we shall also use the projection operator $P_{j,k} : V_k \to V_{j,k}$ satisfying

$$b_k(P_{j,k} v, w) = b_k(v, w) \quad \forall w \in V_{j,k}.$$

As in (3.7), $P_{j,k}$ satisfies

$$\|P_{j,k} v\|_{1,k} \leqslant C \|v\|_{1,\Omega_{j,k}}, \quad v \in V_k, \tag{3.17}$$

for $0 < h_k \leqslant \bar{h}$, where the subdomain $\Omega_{j,k}$ is the support of functions in $V_{j,k}$.

Example 6. Given $g \in V_k$, we define

(1) Set $x_0 = 0$.

(2) Determine x_i, for $i = 1, \ldots, l(k)$, by

$$x_i = x_{i-1} + A_{j,k}^{-1} Q_{j,k}(g - A_k x_{i-1}).$$

(3) Set $R_k g = x_{l(k)}$.

3.5. Analysis of the multigrid method

We now carry out a convergence analysis for MG with the smoothers provided in Examples 1–6 in the framework of [7]. The analysis is based on the product form (3.10) and perturbation from the convergence estimate for the multigrid method applied to the symmetric problem.

We first state a result from [13,14,18] on the convergence rate for the application of MG to the symmetric problem. For this, we need an additional assumption on the coefficient $\mathcal{A}$. In the case of $\widehat{\delta} < 1$ (independent of k) in Theorem 3.1 below, we assume that the elements of $\mathcal{A}$ are in the Sobolev space $W_{r,q}(\Omega)$ for $r > 2/q$ (see [1] for the definition of $W_{r,q}(\Omega)$) in addition to (2.2). In the other case which does not require any elliptic regularity (i.e., in the case of $\widehat{\delta} = 1 - 1/(CK)$ in Theorem 3.1 below), we just assume (2.2) and the boundedness of the elements of $\mathcal{A}$. With this and the definition of $\widehat{E}^K$

$$\widehat{E}^K = (I - \widehat{S}^0) \cdots (I - \widehat{S}^K), \tag{3.18}$$

we have the following theorem [13,14,18]:

Theorem 3.1. *For $k > 0$, let $\widehat{R}_k$ be given by any of Examples* 1–3. *Then there exists a positive constant $\widehat{\delta}$ such that*

$$\widehat{b}_K(\widehat{E}^K v, \widehat{E}^K v) \leqslant \widehat{\delta}^2 \widehat{b}_K(v, v) \quad \forall v \in V_K,$$

where $\widehat{\delta} < 1$ is independent of K if $\mathcal{A} \in (W_{r,q}(\Omega))^{2\times 2}$ $(r > 2/q)$, and $\widehat{\delta} = 1 - 1/(CK)$ for some positive constant C otherwise.

The proof of Theorem 3.1 for the symmetric problem with full elliptic regularity (i.e., $\alpha = 1$ in (3.5)) or without any elliptic regularity was carried out in detail in [13] and [18], respectively. The case of less than full elliptic regularity was mentioned in [14]. To estimate E^K, we need the next lemma.

Lemma 3.2. *For $k > 0$, let $\widehat{R}_k$ be defined by any of Examples* 1–3. *Then with $Z^k = S^k - \widehat{S}^k$,*

$$\widehat{b}_K(Z^k v, w) = D_K\big(v, (\widehat{S}^k)^* w\big) \quad \forall v, w \in V_K, \tag{3.19}$$

where $(\widehat{S}^k)^* = H_k^K \widehat{R}_k^{\mathrm{T}} \widehat{A}_k \widehat{\Pi}_K^k$. For $k = 0$, we have

$$\widehat{b}_K(Z^0 v, w) = D_K\big((I - H_0^K \Pi_K^0) v, H_0^K \widehat{\Pi}_K^0 w\big) \quad \forall v, w \in V_K. \tag{3.20}$$

Proof. For $k > 0$, with $R_k = \widehat{R}_k$ we see that

$$\begin{aligned}
\widehat{b}_K(S^k v, w) &= \widehat{b}_K(H_k^K \widehat{R}_k A_k \Pi_K^k v, w) \\
&= \widehat{b}_k(\widehat{R}_k A_k \Pi_K^k v, \widehat{\Pi}_K^k w) \\
&= b_k(\Pi_K^k v, \widehat{R}_k^{\mathrm{T}} \widehat{A}_k \widehat{\Pi}_K^k w) \\
&= b_K(v, H_k^K \widehat{R}_k^{\mathrm{T}} \widehat{A}_k \widehat{\Pi}_K^k w) \\
&= b_K\big(v, (\widehat{S}^k)^* w\big) \\
&= \widehat{b}_K\big(v, (\widehat{S}^k)^* w\big) + D_K\big(v, (\widehat{S}^k)^* w\big) \\
&= \widehat{b}_K(\widehat{S}^k v, w) + D_K\big(v, (\widehat{S}^k)^* w\big).
\end{aligned}$$

For $k=0$, we have

$$\begin{aligned}\widehat{b}_K(H_0^K\Pi_K^0 v, w) &= \widehat{b}_0(\Pi_K^0 v, \widehat{\Pi}_K^0 w)\\ &= b_0(\Pi_K^0 v, \widehat{\Pi}_K^0 w) - d_0(\Pi_K^0 v, \widehat{\Pi}_K^0 w)\\ &= b_K(v, H_0^K\widehat{\Pi}_K^0 w) - D_K(H_0^K\Pi_K^0 v, H_0^K\widehat{\Pi}_K^0 w)\\ &= \widehat{b}_K(v, H_0^K\widehat{\Pi}_K^0 w) + D_K(v, H_0^K\widehat{\Pi}_K^0 w) - D_K(H_0^K\Pi_K^0 v, H_0^K\widehat{\Pi}_K^0 w)\\ &= \widehat{b}_K(H_0^K\widehat{\Pi}_K^0 v, w) + D_K((I - H_0^K\Pi_K^0)v, H_0^K\widehat{\Pi}_K^0 w).\end{aligned}$$

This completes the proof. □

To use the perturbation analysis, we need to estimate Z^k. In turn, it follows from Lemma 3.2 that we need to bound D_K. That is why we have assumed (3.3) and (3.4) for the difference form $d_k(\cdot,\cdot)$. To show (3.4), we require an assumption on the coefficient $\mathcal{B}$. That is, we assume that $\mathcal{B}$ is continuously differentiable on $\overline{\Omega}$ and piecewise C^2 with the sum of the second-order derivatives over pieces being bounded (see Lemma 3.3 in [19] and Lemma 5.1 in Section 5). With this, we have the next convergence result when $R_k = \widehat{R}_k$ given by Examples 1–3.

Theorem 3.3. *Let $R_k = \widehat{R}_k$ be one of the smoothers defined in Examples* 1–3 *and $\mathcal{B}$ satisfy the above assumption. Then, given $\varepsilon > 0$, there exists an $\overline{h} > 0$ such that for $0 < h_k \leqslant \overline{h}$,*

$$\widehat{b}_K(Ev, Ev) \leqslant \delta^2 \widehat{b}_K(v, v) \quad \forall v \in V_K,$$

where $E = E^K$ is given as in (3.10), *$\delta = \widehat{\delta} + C(h_1 + \varepsilon)$, and $\widehat{\delta}$ is determined by Theorem* 3.1.

Proof. For an arbitrary operator $\mathcal{O}: V_K \to V_K$, its operator norm is defined by

$$\|\mathcal{O}\|_{\widehat{b}} = \sup_{v,w\in V_K} \frac{\widehat{b}_K(\mathcal{O}v, w)}{\widehat{b}_K(v,v)^{1/2}\widehat{b}_K(w,w)^{1/2}}.$$

First, for $k=0$, it follows from an application of (3.4), (3.1), (3.8), and (3.7) to (3.20) that

$$\widehat{b}_K(Z^0 v, w) \leqslant C\varepsilon \big\|(I - H_0^K\Pi_K^0)v\big\|_{1,K} \big\|H_0^K\widehat{\Pi}_K^0 w\big\|_{1,K} \leqslant C\varepsilon \|v\|_{1,K}\|w\|_{1,K}.$$

Next, for $k>0$, apply (3.3), (3.1), (3.14), and the remark following (3.14) to (3.19) to see that

$$\big|\widehat{b}_K(Z^k v, w)\big| \leqslant C\|v\|_{1,K}\big\|(S^k)^* w\big\|_K \leqslant Ch_k\|v\|_{1,K}\|w\|_{1,K}.$$

Since $I - S^k = I - \widehat{S}^k - Z^k$, the operator norm of $I - S^k$ is less than $1 + Ch_k$ by (3.13). Hence we obtain

$$\|E^k\|_{\widehat{b}} \leqslant (1 + C\varepsilon)\prod_{i=1}^{k}(1 + Ch_i) \leqslant C. \tag{3.21}$$

Now, from the definition of E^k and $\widehat{E}^k$ we see that

$$E^k - \widehat{E}^k = (E^{k-1} - \widehat{E}^{k-1})(I - \widehat{S}^k) - E^{k-1}Z^k.$$

Then, by (3.13) and (3.21), we have for $k>0$,

$$\begin{aligned}\|E^k - \widehat{E}^k\|_{\widehat{b}} &\leqslant \|E^{k-1} - \widehat{E}^{k-1}\|_{\widehat{b}}\|I - \widehat{S}^k\|_{\widehat{b}} + \|E^{k-1}\|_{\widehat{b}}\|Z^k\|_{\widehat{b}}\\ &\leqslant \|E^{k-1} - \widehat{E}^{k-1}\|_{\widehat{b}} + Ch_k.\end{aligned} \tag{3.22}$$

By iterating (3.22) and the inequality $\|E^0 - \widehat{E}^0\|_{\widehat{b}} = \|Z^0\|_{\widehat{b}} \leqslant C\varepsilon$, we find that

$$\|E^k - \widehat{E}^k\|_{\widehat{b}} \leqslant C\left(\varepsilon + \sum_{k=1}^{K} h_k\right) \leqslant C(h_0 + \varepsilon).$$

Finally, the desired result follows from Theorem 3.1 and the triangle inequality. □

We remark that ε can be made arbitrarily small by taking h_0 sufficiently small. Thus Theorem 3.3 means that MG for (2.1) converges (with a rate which can be independent of K) provided that the coarsest grid is sufficiently fine. The coarsest grid mesh size can be taken to be independent of K.

We next discuss Example 4. For this, we first consider the multigrid method for the symmetric problem which uses the smoother

$$\widehat{R}_k = \overline{\lambda}_k^{-2} \widehat{A}_k. \tag{3.23}$$

Specifically, we replace A_k by $\widehat{A}_k$ and R_k by $\widehat{R}_k$ in MG for the symmetric problem. While this smoother does not satisfy (3.11), we can still show that the error reduction operator $\widehat{E}^K$ resulting from this smoother satisfies the estimate in Theorem 3.1 by applying the arguments in [14,18]. We now show the convergence rate for MG applied to (2.1) using the smoother given in Example 4. To this end, we need the next lemma.

Lemma 3.4. *For $k > 0$, let R_k be defined by Example 4. Then we have*

$$\|Z^k\|_{\widehat{b}} \leqslant C h_k,$$

where the constant C is independent of k.

Proof. Note that

$$Z^k = \overline{\lambda}_k^{-2} H_k^K \left(A_k^{\mathrm{T}} A_k \Pi_K^k - \widehat{A}_k^2 \widehat{\Pi}_K^k\right) = \overline{\lambda}_k^{-2} H_k^K \left(A_k^{\mathrm{T}}(A_k \Pi_K^k - \widehat{A}_k \widehat{\Pi}_K^k) + (A_k^{\mathrm{T}} - \widehat{A}_k)\widehat{A}_k \widehat{\Pi}_K^k\right). \tag{3.24}$$

Next, observe that

$$\begin{aligned}
\widehat{b}_K(H_k^K A_k \Pi_K^k v, w) &= \widehat{b}_k(A_k \Pi_K^k v, \widehat{\Pi}_K^k w)\\
&= b_k(\Pi_K^k v, \widehat{A}_k \widehat{\Pi}_K^k w)\\
&= b_K(v, H_k^K \widehat{A}_k \widehat{\Pi}_K^k w)\\
&= \widehat{b}_K(v, H_k^K \widehat{A}_k \widehat{\Pi}_K^k w) + D_K(v, H_k^K \widehat{A}_k \widehat{\Pi}_K^k w)\\
&= \widehat{b}_k(\widehat{\Pi}_K^k v, \widehat{A}_k \widehat{\Pi}_K^k w) + D_K(v, H_k^K \widehat{A}_k \widehat{\Pi}_K^k w)\\
&= \widehat{b}_K(H_k^K \widehat{A}_k \widehat{\Pi}_K^k v, w) + D_K(v, H_k^K \widehat{A}_k \widehat{\Pi}_K^k w).
\end{aligned}$$

Consequently, by (3.3), (3.2), and the definition of $\overline{\lambda}_k$ we see that

$$\begin{aligned}
\overline{\lambda}_k^{-1} \widehat{b}_K\left(H_k^K (A_k \Pi_K^k - \widehat{A}_k \widehat{\Pi}_K^k) v, w\right) &= \overline{\lambda}_k^{-1} D_K(v, H_k^K \widehat{A}_k \widehat{\Pi}_K^k w)\\
&\leqslant C \overline{\lambda}_k^{-1} \|v\|_{1,K} \|H_k^K \widehat{A}_k \widehat{\Pi}_K^k w\|_K\\
&\leqslant C \overline{\lambda}_k^{-1} \|v\|_{1,K} \|\widehat{A}_k \widehat{\Pi}_K^k w\|_k\\
&\leqslant C \overline{\lambda}_k^{-1/2} \|v\|_{1,K} (\widehat{A}_k \widehat{\Pi}_K^k w, \widehat{\Pi}_K^k w)_k^{1/2}\\
&\leqslant C h_k \|v\|_{1,K} \|w\|_{1,K}.
\end{aligned} \tag{3.25}$$

In an analogous manner, we have

$$\|\overline{\lambda}_k^{-1} H_k^K (A_k^{\mathrm{T}} - \widehat{A}_k)\widehat{\Pi}_K^k\|_{\widehat{b}} \leqslant C h_k. \tag{3.26}$$

Finally, it can be easily shown that

$$\|\overline{\lambda}_k^{-1} H_k^K A_k^{\mathrm{T}} \widehat{\Pi}_K^k\|_{\widehat{b}} \leqslant C. \tag{3.27}$$

Now, combine (3.24)–(3.27) and the fact [18] that

$$\widehat{\Pi}_K^k H_k^K = I \quad \text{on } V_k, \tag{3.28}$$

to obtain

$$\begin{aligned}
\|Z^k\|_{\widehat{b}} &\leqslant \|\overline{\lambda}_k^{-1} H_k^K A_k^{\mathrm{T}} \widehat{\Pi}_K^k\|_{\widehat{b}} \|\overline{\lambda}_k^{-1} H_k^K (A_k \Pi_K^k - \widehat{A}_k \widehat{\Pi}_K^k)\|_{\widehat{b}} \\
&\quad + \|\overline{\lambda}_k^{-1} H_k^K (A_k^{\mathrm{T}} - \widehat{A}_k)\widehat{\Pi}_K^k\|_{\widehat{b}} \|\overline{\lambda}_k^{-1} H_k^K \widehat{A}_k \widehat{\Pi}_K^k\|_{\widehat{b}} \\
&\leqslant C h_k.
\end{aligned}$$

This completes the proof of the lemma. □

Now, as in the proof of Theorem 3.3, we have the next convergence estimate for Example 4.

Theorem 3.5. *Let R_k be defined by Example 4. Then, given $\varepsilon > 0$, there exists an $\overline{h} > 0$ such that for $0 < h_k \leqslant \overline{h}$,*

$$\widehat{b}_K(Ev, Ev) \leqslant \delta^2 \widehat{b}_K(v, v) \quad \forall v \in V_K,$$

where $\delta = \widehat{\delta} + C(h_1 + \varepsilon)$ and $\widehat{\delta}$ is given by Theorem 3.1 applied to $\widehat{R}_k$ given in (3.23).

We now consider Example 5. The perturbation analysis is based on the multigrid method for $\widehat{A}_k$ with $\widehat{R}_k$ as a smoother given by Example 2. Theorem 3.1 provides an estimate for the operator norm of $\widehat{E}^K$.

Theorem 3.6. *Let R_k be given by Example 5. Then the result in Theorem 3.5 remains valid, with $\widehat{\delta}$ being determined by Theorem 3.1 applied to $\widehat{R}_k$ defined in Example 2.*

Proof. It follows from Examples 2 and 5 and the definition of $P_{j,k}$ and $\widehat{P}_{j,k}$ that the perturbation operator Z^k takes the form

$$Z^k = \gamma \sum_{j=1}^{l(k)} H_k^K (P_{j,k} \Pi_K^k - \widehat{P}_{j,k} \widehat{\Pi}_K^k) = \gamma \sum_{j=1}^{l(k)} H_k^K \{ (P_{j,k} - \widehat{P}_{j,k}) \widehat{\Pi}_K^k + P_{j,k} (\Pi_K^k - \widehat{\Pi}_K^k) \}. \tag{3.29}$$

For the first term of (3.29), by the the definition of $P_{j,k}$ and $\widehat{P}_{j,k}$ again we have

$$\begin{aligned}
\widehat{b}_K(H_k^K P_{j,k} \widehat{\Pi}_K^k v, w) &= \widehat{b}_k(P_{j,k} \widehat{\Pi}_K^k v, \widehat{\Pi}_K^k w) \\
&= \widehat{b}_k(P_{j,k} \widehat{\Pi}_K^k v, \widehat{P}_{j,k} \widehat{\Pi}_K^k w) \\
&= b_k(P_{j,k} \widehat{\Pi}_K^k v, \widehat{P}_{j,k} \widehat{\Pi}_K^k w) - d_k(P_{j,k} \widehat{\Pi}_K^k v, \widehat{P}_{j,k} \widehat{\Pi}_K^k w) \\
&= \widehat{b}_k(\widehat{\Pi}_K^k v, \widehat{P}_{j,k} \widehat{\Pi}_K^k w) + d_k((I - P_{j,k}) \widehat{\Pi}_K^k v, \widehat{P}_{j,k} \widehat{\Pi}_K^k w) \\
&= \widehat{b}_K(H_k^K \widehat{P}_{j,k} \widehat{\Pi}_K^k v, w) + d_k((I - P_{j,k}) \widehat{\Pi}_K^k v, \widehat{P}_{j,k} \widehat{\Pi}_K^k w),
\end{aligned}$$

so that

$$\widehat{b}_K\big(H_k^K(P_{j,k}-\widehat{P}_{j,k})\widehat{\Pi}_K^k v, w\big)=d_k\big((I-P_{j,k})\widehat{\Pi}_K^k v, \widehat{P}_{j,k}\widehat{\Pi}_K^k w\big).$$

Applying (3.3) for d_k, the remark following (3.4), (3.17), and (3.15) yields that

$$\widehat{b}_K\big(H_k^K(P_{j,k}-\widehat{P}_{j,k})\widehat{\Pi}_K^k v, w\big)\leqslant Ch_k\|v\|_{1,\Omega_{j,k}}\|w\|_{1,\Omega_{j,k}}. \tag{3.30}$$

Hence, we see that

$$\widehat{b}_K\left(\sum_{j=1}^{l(k)} H_k^K(P_{j,k}-\widehat{P}_{j,k})\widehat{\Pi}_K^k v, w\right)\leqslant Ch_k\sum_{j=1}^{l(k)}\|v\|_{1,\Omega_{j,k}}\|w\|_{1,\Omega_{j,k}}.$$

By a limited overlap property of the subdomains $\Omega_{j,k}$, we have

$$\left\|\sum_{j=1}^{l(k)} H_k^K(P_{j,k}-\widehat{P}_{j,k})\widehat{\Pi}_K^k\right\|_{\widehat{b}}\leqslant Ch_k. \tag{3.31}$$

The second term of (3.29) can be estimated in a similar fashion:

$$\left\|\sum_{j=1}^{l(k)} H_k^K P_{j,k}(\Pi_K^k-\widehat{\Pi}_K^k)\right\|_{\widehat{b}}\leqslant Ch_k,$$

which, together with (3.29), (3.31), and the proof of Theorem 3.3, implies the desired result. □

Finally, we consider Example 6. Perturbation is based on the multigrid method for $\widehat{A}_k$ with $\widehat{R}_k$ as a smoother given by Example 3. Theorem 3.1 provides an estimate for the resulting operator $\widehat{E}^K$.

Theorem 3.7. *Let R_k be given by Example 6. Then the result in Theorem 3.5 holds, with $\widehat{\delta}$ being determined by Theorem 3.1 applied to $\widehat{R}_k$ defined in Example 3.*

Proof. From the definition of $\widehat{R}_k$ and R_k in Examples 3 and 6, we see that

$$S^k=H_k^K(I-Q_{l(k)})\Pi_K^k \quad\text{and}\quad \widehat{S}^k=H_k^K(I-\widehat{Q}_{l(k)})\widehat{\Pi}_K^k,$$

where

$$Q_j=(I-P_{j,k})(I-P_{j-1,k})\cdots(I-P_{0,k}),\quad j=0,\ldots,l(k),$$
$$\widehat{Q}_j=(I-\widehat{P}_{j,k})(I-\widehat{P}_{j-1,k})\cdots(I-\widehat{P}_{0,k}),\quad j=0,\ldots,l(k).$$

Consequently, the perturbation operator is given by

$$\begin{aligned}Z^k&=H_k^K(I-Q_{l(k)})\Pi_K^k-H_k^K(I-\widehat{Q}_{l(k)})\widehat{\Pi}_K^k\\&=H_k^K(\widehat{Q}_{l(k)}-Q_{l(k)})\widehat{\Pi}_K^k+H_k^K(I-Q_{l(k)})(\Pi_K^k-\widehat{\Pi}_K^k).\end{aligned} \tag{3.32}$$

To estimate the first term of (3.32), note that

$$H_k^K(\widehat{Q}_j-Q_j)\widehat{\Pi}_K^k=H_k^K(I-\widehat{P}_{j,k})(\widehat{Q}_{j-1}-Q_{j-1})\widehat{\Pi}_K^k-H_k^K(\widehat{P}_{j,k}-P_{j,k})Q_{j-1}\widehat{\Pi}_K^k.$$

Since the last two terms are orthogonal with respect to $\widehat{b}_K(\cdot,\cdot)$, we find that

$$\begin{aligned}&\|H_k^K(\widehat{Q}_j - Q_j)\widehat{\Pi}_K^k\|_{\widehat{b}}^2\\ &\quad = \|H_k^K(I - \widehat{P}_{j,k})(\widehat{Q}_{j-1} - Q_{j-1})\widehat{\Pi}_K^k\|_{\widehat{b}}^2 + \|H_k^K(\widehat{P}_{j,k} - P_{j,k})Q_{j-1}\widehat{\Pi}_K^k\|_{\widehat{b}}^2.\end{aligned}$$

It follows from (3.28) and the fact that the operator norm of $H_k^K(I - \widehat{P}_{j,k})\widehat{\Pi}_K^k$ is bounded by one that

$$\|H_k^K(\widehat{Q}_j - Q_j)\widehat{\Pi}_K^k\|_{\widehat{b}}^2 \leqslant \|H_k^K(\widehat{Q}_{j-1} - Q_{j-1})\widehat{\Pi}_K^k\|_{\widehat{b}}^2 + \|H_k^K(\widehat{P}_{j,k} - P_{j,k})Q_{j-1}\widehat{\Pi}_K^k\|_{\widehat{b}}^2.$$

Summing over j yields that

$$\|H_k^K(\widehat{Q}_{l(k)} - Q_{l(k)})\widehat{\Pi}_K^k\|_{\widehat{b}}^2 \leqslant \sum_{j=0}^{l(k)} \|H_k^K(\widehat{P}_{j,k} - P_{j,k})Q_{j-1}\widehat{\Pi}_K^k\|_{\widehat{b}}^2,$$

since $\widehat{Q}_{-1} = Q_{-1} = I$. Now, applying (3.30) and a limited interaction property (as in the proof of Theorem 5.6 in [7]), we obtain

$$\|H_k^K(\widehat{Q}_{l(k)} - Q_{l(k)})\widehat{\Pi}_K^k\|_{\widehat{b}} \leqslant Ch_k. \tag{3.33}$$

Also, using the relation

$$\widehat{b}_K(H_k^K(\Pi_K^k - \widehat{\Pi}_K^k)v, w) = D_K((I - H_k^K\Pi_K^k)v, H_k^K\widehat{\Pi}_K^k w),$$

it can be seen that the second term of (3.32) can be bounded by the same estimate:

$$\|H_k^K(I - Q_{l(k)})(\Pi_K^k - \widehat{\Pi}_K^k)\|_{\widehat{b}} \leqslant Ch_k.$$

This, together with (3.32), (3.33), and the proof of Theorem 3.3 yields the desired result. □

4. The P_1-nonconforming element

In this section we consider the nonconforming P_1 element. Let h_0 and $\mathcal{E}_{h_0} = \mathcal{E}_0$ be given. For each integer $1 \leqslant k \leqslant K$, let $h_k = 2^{-k}h_0$ and $\mathcal{E}_{h_k} = \mathcal{E}_k$ be constructed by connecting the midpoints of the edges of the triangle in $\mathcal{E}_{k-1}$. For each k, define the P_1-nonconforming finite element space [21]

$$V_k = \{v \in L^2(\Omega)\colon v|_E \text{ is linear for all } E \in \mathcal{E}_k,\ v \text{ is continuous at the midpoints of interior edges,}\\ \text{and } v \text{ vanishes at the midpoints of edges on } \partial\Omega\}.$$

For the P_1-nonconforming element, the inner product $(\cdot,\cdot)_k$ is defined by

$$(v, w)_k = h_k^2 \sum_q v(q)w(q), \quad v, w \in V_k,$$

where the summation is taken over all the midpoints q in $\mathcal{E}_k$.

Since $V_{k-1} \not\subset V_k$ (i.e., non-nested), we need to introduce intergrid transfer operators to connect them. Following [8,9], the coarse-to-fine intergrid transfer operator $I_k : V_{k-1} \to V_k$ for $k = 1, \ldots, K$ is defined as follows. For $v \in V_{k-1}$, let q be a midpoint of an edge of a triangle in $\mathcal{E}_k$; then we define $I_k v$ by

$$(I_k v)(q) = \begin{cases} 0 & \text{if } q \in \partial\Omega,\\ v(q) & \text{if } q \notin \partial E \text{ for any } E \in \mathcal{E}_{k-1},\\ \frac{1}{2}\{v|_{E_1}(q) + v|_{E_2}(q)\} & \text{if } q \in \partial E_1 \cap \partial E_2 \text{ for some } E_1, E_2 \in \mathcal{E}_{k-1}.\end{cases}$$

Now, for the P_1-nonconforming element, (3.1)–(3.4), (3.7), and (3.8) are satisfied [13,19,25]. The upper bound in (3.2) was shown under the condition that there meet at most six edges at each interior vertex in the initial triangulation $\mathcal{E}_0$ and four edges at each boundary vertex in $\mathcal{E}_0$ [25]; this condition on $\mathcal{E}_0$ is easily satisfied. Therefore, the convergence estimates for MG proven in Section 3.5 hold for this element.

5. The rotated Q_1-nonconforming element

In this section we consider the rotated Q_1-nonconforming element for (2.1). For this, let $\mathcal{E}_{h_0} = \mathcal{E}_0$ be a partition of Ω into rectangles having maximum diameter h_0 and oriented along the coordinate axes. For each integer $1 \leqslant k \leqslant K$, let $h_k = 2^{-k}h_0$ and $\mathcal{E}_{h_k} = \mathcal{E}_k$ be constructed by connecting the midpoints of the edges of the rectangle in $\mathcal{E}_{k-1}$, and let $\partial\mathcal{E}_k$ be the set of all interior edges in $\mathcal{E}_k$. The rotated Q_1 nonconforming space is defined by [11,26]

$$V_k = \Big\{ v \in L^2(\Omega)\colon v|_E = a_E^1 + a_E^2 x + a_E^3 y + a_E^4 (x^2 - y^2),\ a_E^i \in \mathbb{R}\ \forall E \in \mathcal{E}_k;\\ \text{if } E_1 \text{ and } E_2 \text{ share an edge } e, \text{ then } \int_e \xi|_{\partial E_1}\,\mathrm{d}s = \int_e \xi|_{\partial E_2}\,\mathrm{d}s;\\ \text{and } \int_{\partial E \cap \Gamma} \xi|_\Gamma\,\mathrm{d}s = 0 \Big\}.$$

For this element, the inner product $(\cdot,\cdot)_k$ is given as follows. Let $\{\phi_k^j\}$ be the basis functions of V_k such that the edge average of ϕ_k^j equals one at exactly one edge and zero at all other edges. Then each $v \in V_k$ has the representation

$$v = \sum_j v^j \phi_k^j.$$

Now, for $v,\ w \in V_k$ we define

$$(v, w)_k = h_k^2 \sum_j v^j w^j.$$

By the uniform L^2-stability of the basis functions [20], we can easily show that the norm induced by $(\cdot,\cdot)_k$ is equivalent to the standard $L^2(\Omega)$ norm.

Since $V_{k-1} \not\subset V_k$ again, following [2,11] we define the coarse-to-fine intergrid transfer operators $I_k\colon V_{k-1} \to V_k$ as follows. If $v \in V_{k-1}$ and e is an edge of a rectangle in $\mathcal{E}_k$, then $I_k v \in V_k$ is defined by

$$\int_e I_k v\,\mathrm{d}s = \begin{cases} 0 & \text{if } e \subset \partial\Omega,\\ \int_e v\,\mathrm{d}s & \text{if } e \not\subset \partial E \text{ for any } E \in \mathcal{E}_{k-1},\\ \frac{1}{2}\int_e (v|_{E_1} + v|_{E_2})\,\mathrm{d}s & \text{if } e \subset \partial E_1 \cap \partial E_2 \text{ for some } E_1, E_2 \in \mathcal{E}_{k-1}. \end{cases}$$

For the rotated Q_1-nonconforming element, (3.1)–(3.3), (3.7), and (3.8) have been shown in [13,20]. The upper bound in (3.2) was established for square partitions of a square for the rotated Q_1-nonconforming element [20]. Extensions to other domains and triangulations have been discussed in [20]; it holds for polygonal domains if their initial triangulation into quadrilaterals is topologically equivalent to a uniform square partition of $\Omega = (0,1)^2$, for example. It remains to prove (3.4), which is completed in the next lemma.

Lemma 5.1. *We assume that the coefficient $\mathcal{B}$ is continuously differentiable on $\overline{\Omega}$ and piecewise C^2 with the sum of the second-order derivatives over pieces being bounded. Then there is a constant C independent of k such that* (3.4) *is satisfied.*

Proof. To prove (3.4), we apply integration by parts on each finite element to see that

$$D_k(v,w) = \sum_{E\in\mathcal{E}_k} \{(\mathcal{B}\cdot\nu_E v, w)_{\partial E} - (\nabla\cdot\mathcal{B}w + \mathcal{B}\cdot\nabla w, v)_E\} + ((c-1)v, w). \tag{5.1}$$

Evidently, it suffices to estimate the terms over edges. Let $E_1, E_2 \in \mathcal{E}_k$ share a vertical edge e with midpoint m^k; a horizontal edge can be analyzed similarly. Then, by the midpoint rule we find that

$$\begin{aligned} &\int_e (\mathcal{B}\cdot\nu_{E_1} vw)|_{E_1}\,\mathrm{d}s + \int_e (\mathcal{B}\cdot\nu_{E_2} vw)|_{E_2}\,\mathrm{d}s \\ &\quad = |e|\{(\mathcal{B}\cdot\nu_{E_1} vw)|_{E_1}(m^k) + (\mathcal{B}\cdot\nu_{E_2} vw)|_{E_2}(m^k)\} \\ &\qquad + \frac{|e|^3}{24}\left\{\frac{\partial^2}{\partial y^2}(\mathcal{B}\cdot\nu_{E_1} vw)|_{E_1}(\xi_1^k) + \frac{\partial^2}{\partial y^2}(\mathcal{B}\cdot\nu_{E_2} vw)|_{E_2}(\xi_2^k)\right\}, \end{aligned} \tag{5.2}$$

for some points $\xi_1^k, \xi_2^k \in e$.

To estimate the first term of (5.2), note that, by the midpoint rule again,

$$v|_{E_i}(m^k) = \overline{v}_e|_{E_i} - \frac{|e|^2}{24}\frac{\partial^2 v}{\partial y^2}|_{E_i}, \quad i = 1,2,$$

where

$$\overline{v}_e|_{E_i} = \frac{1}{|e|}\int_e v|_{E_i}\,\mathrm{d}y.$$

Then, by the continuity of the edge integrals of elements in V_k, we see that

$$\begin{aligned} &|e|\{(\mathcal{B}\cdot\nu_{E_1} vw)|_{E_1}(m^k) + (\mathcal{B}\cdot\nu_{E_2} vw)|_{E_2}(m^k)\} \\ &\quad = -\frac{|e|^3}{24}\sum_{i=1}^{2}\mathcal{B}\cdot\nu_{E_i}(m^k)\left\{\left(\overline{v}_e\frac{\partial^2 w}{\partial y^2}\right)\Big|_{E_i} + \left(\overline{w}_e\frac{\partial^2 v}{\partial y^2}\right)\Big|_{E_i} - \frac{|e|^2}{24}\left(\frac{\partial^2 v}{\partial y^2}\frac{\partial^2 w}{\partial y^2}\right)\Big|_{E_i}\right\}. \end{aligned}$$

Now, let e_1 be another vertical edge of E_1 with midpoint m_1^k. Applying the relation

$$\begin{aligned} &\mathcal{B}\cdot\nu_{E_1}(m^k)\overline{v}_e|_{E_1} + \mathcal{B}\cdot\nu_{E_1}(m_1^k)\overline{v}_{e_1}|_{E_1} \\ &\quad = (\mathcal{B}\cdot\nu_{E_1}(m^k) + \mathcal{B}\cdot\nu_{E_1}(m_1^k))\overline{v}_e|_{E_1} - \mathcal{B}\cdot\nu_{E_1}(m_1^k)(\overline{v}_e|_{E_1} - \overline{v}_{e_1}|_{E_1}), \end{aligned}$$

the definition of V_k, and inverse inequalities, we see that

$$\begin{aligned} &\left|\sum_{e\in\partial\mathcal{E}_k}|e|\{(\mathcal{B}\cdot\nu_{E_1} vw)|_{E_1}(m^k) + (\mathcal{B}\cdot\nu_{E_2} vw)|_{E_2}(m^k)\}\right| \\ &\quad \leqslant Ch_k(\|v\|_k + \|v\|_{1,k})(\|w\|_k + \|w\|_{1,k}) \leqslant C\|v\|_k\|w\|_{1,k}. \end{aligned} \tag{5.3}$$

Applying the same argument to

$$\frac{\partial^2}{\partial y^2}(\mathcal{B}\cdot\nu_{E_i}vw) = \frac{\partial^2}{\partial y^2}(\mathcal{B}\cdot\nu_{E_i})vw + 2\frac{\partial}{\partial y}(\mathcal{B}\cdot\nu_{E_i})\frac{\partial}{\partial y}(vw) + 2(\mathcal{B}\cdot\nu_{E_i})\frac{\partial v}{\partial y}\frac{\partial w}{\partial y} + (\mathcal{B}\cdot\nu_{E_i})\left(v\frac{\partial^2 w}{\partial y^2} + w\frac{\partial^2 v}{\partial y^2}\right), \quad i = 1, 2,$$

we can show that

$$\left|\sum_{e\in\partial\mathcal{E}_k}\frac{|e|^3}{24}\left\{\frac{\partial^2}{\partial y^2}(\mathcal{B}\cdot\nu_{E_1}vw)|_{E_1}(\xi_1^k) + \frac{\partial^2}{\partial y^2}(\mathcal{B}\cdot\nu_{E_2}vw)|_{E_2}(\xi_2^k)\right\}\right| \leqslant C\|v\|_k\|w\|_{1,k}.$$

Combine this, (5.2), and (5.3) to obtain the desired result. □

We conclude with a remark that extensions and generalizations of the techniques discussed in this paper are possible. These techniques can be applied to three-dimensional problems, other types of nonconforming finite elements, and more general boundary conditions, for example.

References

[1] R.A. Adams, *Sobolev Spaces* (Academic Press, New York, 1975).

[2] T. Arbogast and Z. Chen, On the implementation of mixed methods as nonconforming methods for second order elliptic problems, *Math. Comp.* 64 (1995) 943–972.

[3] D.N. Arnold and F. Brezzi, Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates, *RAIRO Model. Math. Anal. Numer.* 19 (1985) 7–32.

[4] R. Bank, A comparison of two multilevel iterative methods for nonsymmetric and indefinite elliptic finite element equations, *SIAM J. Numer. Anal.* 18 (1981) 724–743.

[5] J. Bramble, *Multigrid Methods*, Pitman Research Notes in Mathematics, Vol. 294 (Longman, London, 1993).

[6] J. Bramble and J. Pasciak, The analysis of smoothers for multigrid algorithms, *Math. Comp.* 58 (1992) 467–488.

[7] J. Bramble, D.Y. Kwak and J. Pasciak, Uniform convergence of multigrid V-cycle iterations for indefinite and nonsymmetric problems, *SIAM J. Numer. Anal.* 31 (1994) 1749–1763.

[8] D. Braess and R. Verfürth, Multigrid methods for nonconforming finite element methods, *SIAM J. Numer. Anal.* 27 (1990) 979–986.

[9] S. Brenner, An optimal-order multigrid method for P1 nonconforming finite elements, *Math. Comp.* 52 (1989) 1–15.

[10] Z. Chen, Analysis of mixed methods using conforming and nonconforming finite element methods, *RAIRO Math. Model. Numer. Anal.* 27 (1993) 9–34.

[11] Z. Chen, Projection finite element methods for semiconductor device equations, *Comput. Math. Appl.* 25 (1993) 81–88.

[12] Z. Chen, Equivalence between and multigrid algorithms for mixed and nonconforming methods for second order elliptic problems, *East–West J. Numer. Math.* 4 (1996) 1–33.

[13] Z. Chen, On the convergence of Galerkin–multigrid methods for nonconforming finite elements, Technical Report # 96-02, Southern Methodist University, TX (1996); submitted to *Math. Comp.*

[14] Z. Chen, The analysis of intergrid transfer operators and multigrid methods for nonconforming finite elements, *ETNA* 5 (1998) 78–96.

[15] Z. Chen and R.E. Ewing, From single-phase to compositional flow: applicability of mixed finite elements, *Transport in Porous Media* 27 (1997) 225–242.

[16] Z. Chen, R.E. Ewing and R. Lazarov, Domain decomposition algorithms for mixed methods for second order elliptic problems, *Math. Comp.* 65 (1996) 467–490.
[17] Z. Chen, R.E. Ewing, Y. Kuznetsov, R. Lazarov and S. Maliassov, Multilevel preconditioners for mixed methods for second order elliptic problems, *J. Numer. Linear Algebra Appl.* 30 (1996) 427–453.
[18] Z. Chen and D.Y. Kwak, Convergence of multigrid methods for nonconforming finite elements without regularity assumptions, *Comput. Appl. Math.*, to appear.
[19] Z. Chen, D.Y. Kwak and Y.J. Yon, Multigrid algorithms for nonconforming and mixed methods for nonsymmetric and indefinite problems, *SIAM J. Sci. Comput.* 19 (1998) 502–515.
[20] Z. Chen and P. Oswald, Multigrid and multilevel methods for nonconforming Q1 elements, *Math. Comp.* 67 (1998) 667–693.
[21] M. Crouzeix and P. Raviart, Conforming and nonconforming finite element methods for solving the stationary Stokes equations I, *RAIRO* 3 (1973) 33–75.
[22] W. Hackbusch, *Multigrid Methods and Applications* (Springer, Berlin, 1985).
[23] P. Kloucek and M. Luskin, The computation of the dynamics of martensitic microstructure, *Continuum Mech. Thermodyn.* 6 (1994) 209–240.
[24] S. McCormick, ed., *Multigrid Methods*, SIAM Frontiers in Applied Mathematics 3 (SIAM, Philadelphia, PA, 1987).
[25] P. Oswald, On a hierarchical basis multilevel method with nonconforming P1 elements, *Numer. Math.* 62 (1992) 189–212.
[26] R. Rannacher and S. Turek, Simple nonconforming quadrilateral Stokes element, *Numer. Methods Partial Differential Equations* 8 (1992) 97–111.
[27] A. Schatz and J. Wang, Some new error estimates in finite element methods, *Math. Comp.* 65 (1996) 19–27.