MULTIGRID ALGORITHMS FOR NONCONFORMING AND MIXED METHODS FOR SYMMETRIC AND NONSYMMETRIC PROBLEMS

ZHANGXIN CHEN, DO Y. KWAK AND YOON J. YON*

Abstract. In this paper we consider multigrid algorithms for nonconforming and mixed finite element methods for symmetric and nonsymmetric second order elliptic problems. We prove optimal convergence properties of the $\mathcal{W}$-cycle multigrid algorithm and uniform condition number estimates for the variable $\mathcal{V}$-cycle preconditioner for the symmetric problem. For the nonsymmetric and/or indefinite problem, we show that a simple $\mathcal{V}$-cycle multigrid iteration converges at a uniform rate provided that the coarsest level in the multilevel iteration is sufficiently fine (but independent on the number of multigrid levels). Various types of smoothers for the nonsymmetric and indefinite problem are discussed. Extensive numerical results for both symmetric and nonsymmetric problems are given to illustrate the present theories.

Key words. mixed method, nonconforming method, finite elements, multigrid algorithm, convergence, symmetric problems, nonsymmetric and/or indefinite problems

AMS(MOS) subject classifications. 65N30, 65N22, 65F10

1. Introduction.

In this paper we study some multigrid algorithms for second order elliptic problems including nonsymmetric and/or indefinite problems. We consider the solution of the discrete systems which arises from the application of nonconforming and mixed finite element methods. For the nonsymmetric and indefinite problem, we assume that the nonsymmetric/indefinite terms are a "compact perturbation"; the convection-dominated problems are not studied here.

In the second section we study multigrid algorithms for solving the symmetric problem by the standard P_1 nonconforming finite elements. The P_1 nonconforming multigrid algorithms have been studied in the past few years. There have been two types of multigrid algorithms for solving the symmetric problem. The first one exploits the nonconforming finite elements in both smoothing iterations and coarse-grid corrections in the multilevel iteration. For this type of multigrid algorithm, only the $\mathcal{W}$-cycle algorithms were proven to be convergent under the assumption that the number of smoothing iterations on all levels is big enough [8], [9], [12]. The arguments in these earlier papers follow the standard proof of convergence of multigrid algorithms for conforming finite element methods [2], and do not apply to the $\mathcal{V}$-cycle algorithm.

The second type of multigrid algorithm uses the nonconforming finite elements in the smoothing iterations on the finest level, but the P_1 conforming finite elements in the coarse-grid corrections in the multilevel iteration. For this approach, uniform iterative convergence estimates for the $\mathcal{V}$-cycle multigrid algorithm with one smoothing step have been obtained for the symmetric problem [12], [19].

In this paper we re-examine the first type of multigrid algorithm for the symmetric problem. We present a different convergence analysis for these multigrid algorithms than those in [8], [9], and [12]. This analysis applies to both the $\mathcal{W}$-cycle and the variable $\mathcal{V}$-cycle. The variable $\mathcal{V}$-cycle algorithm is one in which the number of smoothing steps increase exponentially as the number of grid decreases; there is no requirement on the number of smoothings on the finest level for the variable $\mathcal{V}$-cycle. We prove

* Z. Chen, Department of Mathematics and Institute for Scientific Computation, Texas A&M University, College Station, TX 77843–3404, email: zchen@isc.tamu.edu. Partly supported by the Department of Energy under contract DE-ACOS-840R21400. D. Kwak and Yon, Department of Mathematics, Korea Advanced Institute of Science and Technology, Taejon, Korea 305–701, email: dykwak@math.kaist.ac.kr, yjyon@math.kaist.ac.kr.

optimal convergence properties of the $\mathcal{W}$-cycle multigrid algorithm and uniform condition number estimates for the variable $\mathcal{V}$-cycle preconditioner. Furthermore, explicit bounds for the convergence rate are determined, which have not been given in the earlier papers.

In the third section we study multigrid algorithms for solving the nonsymmetric and/or indefinite problem by the nonconforming finite elements. We first consider the problem of existence and uniqueness of the solution of the discrete systems arising from application of the nonconforming method to the nonsymmetric and/or indefinite problem. We prove that the discrete systems have a unique solution and produce optimal order error estimates provided that the size of meshes is sufficiently small. We then provide a convergence analysis for the multigrid algorithms. There has been intensive research on the multigrid algorithms for the nonsymmetric and indefinite problem using the conforming finite elements. Paper [5] has a good survey in its introduction, and is most closely related to the subject of this paper. For the nonconforming multigrid algorithm of the nonsymmetric and indefinite problem under consideration, we only analyze the second type of multigrid algorithm mentioned above. We show that the result for the symmetric problem can carry over to the nonsymmetric and indefinite case. Namely, we prove uniform iterative convergence estimates for the $\mathcal{V}$-cycle multigrid algorithm for the nonsymmetric and indefinite problem under rather weak assumptions (e.g., the domain need not be convex). We mention that the present argument does not cover the first type of multigrid algorithm for the nonsymmetric and indefinite problem since the $\mathcal{V}$-cycle algorithm for the symmetric problem in this case has not been proven to converge.

A variety of smoothers are considered here. 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.

Not only is the analysis of multigrid algorithms for nonconforming finite element methods of interest for their own sake (see, e.g., [15], [17], [18], [21] and the bibliographies therein), but it has great application to mixed finite element methods. It has been shown [12], [14], [15] that the linear system arising from the mixed methods of the symmetric problem can be algebraically condensed to a symmetric, positive definite system for Lagrange multipliers. This linear system is identical to the system arising from the nonconforming finite element methods. Hence the analysis of multigrid algorithms for the nonconforming methods can carry over directly to the mixed methods. The analyses of multigrid algorithms for the mixed methods of the symmetric and nonsymmetric and indefinite problems are given in §2.2 and §3.3, respectively. According to the knowledge of the authors, the multigrid algorithms for the nonsymmetric and indefinite problem by the nonconforming and mixed methods are analyzed here for the first time.

In the final section extensive numerical results for both symmetric and nonsymmetric problems are given to illustrate the present theories and compare the nonconforming multigrid methods with standard conforming finite element and finite difference multigrid methods. Special attention is paid to the latter case. In particular, for the nonsymmetric and indefinite problem, both types of multigrid methods mentioned above are tested for the first time. The later analysis is carried out for the two-dimensional, triangular case; it works for the three-dimensional case without substantial changes as noticed in [12], [14], [15]. Also, rectangular finite elements can be similarly considered.

2. The Symmetric Problem. In this section we develop multigrid algorithms for the symmetric problem. The nonconforming finite elements are considered in §2.1, and the mixed methods are described in §2.2.

2.1. The nonconforming multigrid algorithm. We consider as our model problem the following symmetric equation

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

where $\Omega \subset \mathbb{R}^n$, $n = 2, 3$ is a simply connected bounded polygonal domain with the boundary $\partial\Omega$, $f \in L^2(\Omega)$, and the coefficients $\mathcal{A} \in (L^\infty(\Omega))^{n\times n}$, $c \in L^\infty(\Omega)$ satisfy

$$\xi^t \mathcal{A}(x) \xi \geq a_0 \xi^t \xi, \qquad x \in \Omega,\ \xi \in \mathbb{R}^n, \tag{2.2a}$$

and

$$0 \leq c_1 \leq c(x) \leq c_2, \qquad x \in \Omega, \tag{2.2b}$$

with fixed constants $a_0 > 0, c_1, c_2$.

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

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

where $(\cdot,\cdot)$ denotes the $L^2(\Omega)$ or $(L^2(\Omega))^n$ inner product, as appropriate. Then the weak form of (2.1) for the solution $u \in H_0^1(\Omega)$ is

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

For $0 < h < 1$, let $\mathcal{E}_h$ be a triangulation of Ω into triangles of size h, and define the P_1-nonconforming finite element space

$$\begin{aligned} V_h = \{v \in L^2(\Omega) : {} & v|_E \text{ is linear for all } E \in \mathcal{E}_h, v \text{ is continuous} \\ & \text{at the barycenters of interior edges and} \\ & \text{vanishes at the barycenters of edges on } \partial\Omega\}. \end{aligned}$$

Associated with V_h, we introduce a bilinear form on $V_h \oplus H_0^1(\Omega)$ by

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

where $(\cdot,\cdot)_E$ is the $L^2(E)$ inner product. Then the P_1-nonconforming finite element discretization of (2.1) is to find $u_h \in V_h$ such that

$$a_h(u_h, v) = (f, v), \qquad \forall\ v \in V_h. \tag{2.4}$$

To develop a multigrid algorithm for (2.4), we need to assume a structure to our family of partitions. Let h_1 and $\mathcal{E}_{h_1} = \mathcal{E}_1$ be given. For each integer $1 < k \leq K$, let $h_k = 2^{1-k}h_1$ and $\mathcal{E}_{h_k} = \mathcal{E}_k$ be constructed by connecting the midpoints of the edges of the triangle in $\mathcal{E}_{k-1}$, 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.

Following [8], [9], [12], the coarse-to-fine intergrid transfer operator $I_k : V_{k-1} \to V_k$ for $k = 2, \dots, K$ 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

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

Let $A_k : V_k \to V_k$ be the discretization operator on level k given by

$$(A_k v, w) = a_k(v, w), \qquad \forall\, w \in V_k. \tag{2.5}$$

The operator A_k is clearly symmetric (in both the $a_k(\cdot,\cdot)$ and $(\cdot,\cdot)$ inner products) and positive definite. Also, we define the operators $P_{k-1} : V_k \to V_{k-1}$ and $P^0_{k-1} : V_k \to V_{k-1}$ by

$$a_{k-1}(P_{k-1} v, w) = a_k(v, I_k w), \qquad \forall\, w \in V_{k-1},$$

and

$$\left(P^0_{k-1} v, w\right) = (v, I_k w), \qquad \forall\, w \in V_{k-1}.$$

It is easy to see that $I_k P_{k-1}$ is a symmetric operator wtih respect to the a_k form. Note that neither P^0_k nor P_k is a projection in the nonconforming case. Finally, the multigrid algorithm which we shall consider also requires linear smoothing operators $R_k : V_k \to V_k$ for $k = 1, \dots, K$. Let R^t_k denote the adjoint of R_k with respect to the $(\cdot,\cdot)$ inner product and define

$$R^{(l)}_k = \begin{cases} R_k & \text{if } l \text{ is odd,} \\ R^t_k & \text{if } l \text{ is even.} \end{cases}$$

Multigrid algorithms can be represented in variational and non-variational forms. Both approaches are equivalent and connected in the sense that the multigrid processes result in a linear iterative scheme with a reduction operator equal to $I - B_K A_K$, where $B_K : V_K \to V_K$ is the multigrid operator. In this paper, following [7], we define the multigrid operator $B_k : V_k \to V_k$ in terms of an iterative process as follows.

Multigrid Algorithm 2.1. Let $1 < k \le K$ and p be a positive integer. Set $B_1 = A_1^{-1}$. Assume that B_{k-1} has been defined and define $B_k g$ for $g \in V_k$ as follows:

1. Set $x^0 = 0$ and $q^0 = 0$.
2. Define x^l for $l = 1, \dots, m(k)$ by
$$x^l = x^{l-1} + R_k^{(l+m(k))}(g - A_k x^{l-1}).$$
3. Define $y^{m(k)} = x^{m(k)} + I_k q^p$, where q^i for $i = 1, \dots, p$ is defined by
$$q^i = q^{i-1} + B_{k-1}\left[P^0_{k-1}\left(g - A_k x^{m(k)}\right) - A_{k-1} q^{i-1} \right].$$
4. Define y^l for $l = m(k)+1, \dots, 2m(k)$ by
$$y^l = y^{l-1} + R_k^{(l+m(k))}\left(g - A_k y^{l-1}\right).$$
5. Set $B_k g = y^{2m(k)}$.

In Algorithm 2.1, $m(k)$ gives the number of pre- and post-smoothing iterations and can vary as a function of k. If $p=1$, we have a $\mathcal{V}$-cycle multigrid algorithm. If $p=2$, we have a $\mathcal{W}$-cycle algorithm. As mentioned in the introduction, a variable $\mathcal{V}$-cycle algorithm is one in which the number of smoothings $m(k)$ increase exponentially as k decreases (i.e., $p=1$ and $m(k)=2^{K-k}$). The smoothings are alternated following [7] and are put together so that the resulting multigrid preconditioner B_k is symmetric in the $L^2(\Omega)$ inner product for each k.

We now apply the theory developed in [7] to analyze Algorithm 2.1. Toward that end, we require appropriate conditions for the smoother, need to establish the stability property of the intergrid transfer operator, and show the so-called "regularity and approximation" inequality. Below we use $\|\cdot\|$ to indicate the standard $L^2(\Omega)$ norm.

The construction of smoothers for the nonconforming method fits into the general theory of [7]. We can use point and line Jacobi and Gauss-Seidel smoothing procedures to define R_k [12], for example. The smoothing estimates are a consequence of the general smoothing theory developed in [6]. Hence, it suffices to consider the remaining two conditions, which are defined as follows:

$$a_k(I_kv, I_kv) \leq C_* a_{k-1}(v,v), \qquad \forall\ v \in V_{k-1}, \tag{2.6}$$

and

$$|a_k\left((I - I_kP_{k-1})\, v, v\right)| \leq C_\alpha \left(\frac{\|A_kv\|^2}{\lambda_k}\right)^\alpha a_k(v,v)^{1-\alpha}, \qquad \forall\ v \in V_k, \tag{2.7}$$

for $k=2,\dots,K$, where C_* and C_α are constants independent of k, λ_k is the largest eigenvalue of A_k, $I : V_k \to V_k$ is the identity operator, and $0 < \alpha \leq 1$. Here and throughout the paper, C, with or without subscript, denotes a generic positive constant.

We first prove (2.6). We remark that the constant C_* in (2.6) is in general bigger than two, as shown in the following example. Thus the general theory in [7] does not apply to the standard $\mathcal{V}$-cycle, rather to the variable $\mathcal{V}$-cycle and the $\mathcal{W}$-cycle.

To construct the example, we rewrite the definition of I_k as follows. From Fig. 1 it is not difficult to see that for $v \in V_{k-1}$,

$$(I_kv)(x_A) = \frac{1}{2}(v(x_1) + v(x_2)), \tag{2.8a}$$

and

$$(I_kv)(x_B) = v(x_1) + \frac{1}{4}\{v(x_2) + v(x_4) - v(x_3) - v(x_5)\}. \tag{2.8b}$$

Example 1. Let Ω be given in Fig. 2 and v be in V_{k-1} with the nodal values determined in Fig. 2. Then with $\mathcal{A} = I$ and $c = 0$ it is easy to check that

$$a_{k-1}(v,v) = 16,$$

and

$$a_k(I_kv, I_kv) = 32.5.$$

Consequently, we see that

$$a_k(I_kv, I_kv) > 2a_{k-1}(v,v).$$

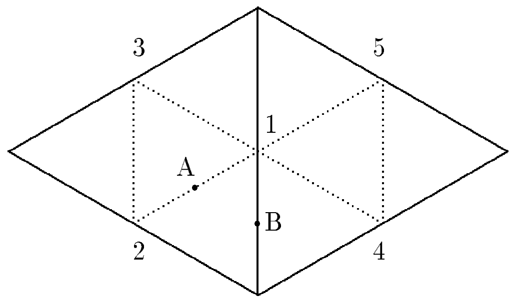

FIG. 1. *Illustration of the definition of I_k.*

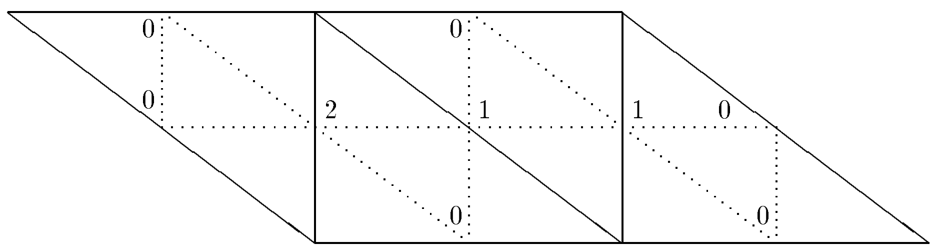

FIG. 2. *The definition of the function v in Example 1.*

We now turn to condition (2.6).

Proposition 2.1. *There is a constant C_* independent of k such that* (2.6) *is satisfied.*
PROOF. Note that for every $v \in V_{k-1}$, $a_{k-1}(v,v)$ is a norm in V_{k-1} equivalent to

$$\sum_{E\in\mathcal{E}_{k-1}} \sum_{i,j=1}^{3} \left((v(\overline{x}_i) - v(\overline{x}_j))^2 + h_{k-1}^2 v^2(\overline{x}_i) \right),$$

where the $\overline{x}_i$ are the midpoints of the edges of E. A similar result holds for every $v \in V_k$. Then the result (2.6) easily follows from the definition of I_k in (2.8). □

To prove condition (2.7), for simplicity we assume that the solution to (2.1) satisfies

$$\|u\|_{H^2(\Omega)} \le C\|f\|. \tag{2.9}$$

We point out that the present results can be extended to nonconforming methods for problems on domains which are non-convex. Ω can be the L-shaped domain, for example. For such domains, the full regularity result (2.9) does not hold. However, the analysis below can be modified to show that the regularity and approximation property is still satisfied but for a smaller value of α; see [7] and the next section for more details in this direction.

We now turn to assumption (2.7). We need the following three lemmas.

Lemma 2.2. *There is a constant C independent of k such that*

$$\|v - I_k v\|^2 \le C h_k^2 \sum_{E\in\mathcal{E}_{k-1}} \|\nabla v\|_{L^2(E)}^2, \qquad v \in V_{k-1}. \tag{2.10}$$

PROOF. (2.10) follows from the definition of I_k since I_k can be regarded as a nodal point interpolation into V_k. □

For $k = 2, \dots, K$, let $\bar{P}_{k-1}$ be the elliptic projection into the P_1-conforming finite element space U_{k-1}, i.e.,

$$a_{k-1}(\bar{P}_{k-1}v, w) = a_k(v, w), \qquad \forall\ w \in U_{k-1}. \tag{2.11}$$

Note that $U_{k-1} = V_{k-1} \cap H_0^1(\Omega)$ and $a_{k-1}(\cdot,\cdot) = a(\cdot,\cdot)$ on U_{k-1}.

Lemma 2.3. *There is a constant C independent of k such that*

$$\|\bar{P}_{k-1}v - v\|^2 \le Ch_k^2 \sum_{E\in\mathcal{E}_k} \|\nabla v\|^2_{L^2(E)}, \qquad v \in V_k. \tag{2.12}$$

PROOF. Let $f_0 \in V_k$ be defined by

$$(f_0, w) = a_k(v, w), \qquad \forall\ w \in V_k. \tag{2.13}$$

Then it follows from (2.11) that

$$a(\bar{P}_{k-1}v, w) = (f_0, w), \qquad \forall\ w \in U_{k-1}, \tag{2.14}$$

since $U_{k-1} \subset V_k$. Also, let $z \in H_0^1(\Omega)$ satisfy

$$a(z, w) = (f_0, w), \qquad \forall\ w \in H_0^1(\Omega). \tag{2.15}$$

Note that, by (2.13)–(2.15), we see that v and $\bar{P}_{k-1}v$ are the nonconforming and conforming finite element solutions of z, respectively. From the standard error estimates [16] we find that

$$\|z - v\| \le Ch_k^2 \|z\|_{H^2(\Omega)} \quad \text{and} \quad \|z - \bar{P}_{k-1}v\| \le Ch_k^2\|z\|_{H^2(\Omega)},$$

so that, by (2.9),

$$\|v - \bar{P}_{k-1}v\| \le \|z - v\| + \|z - \bar{P}_{k-1}v\| \le Ch_k^2\|f_0\|. \tag{2.16}$$

Take $w = f_0$ in (2.13) to see that

$$\|f_0\|^2 = a_k(v, f_0) \le Ch_k^{-1}\|f_0\| \left(\sum_{E\in\mathcal{E}_k} \|\nabla v\|^2_{L^2(E)}\right)^{1/2}, \tag{2.17}$$

by an inverse inequality and the Poincaré inequality. Combine (2.16) and (2.17) to yield (2.12). □

Lemma 2.4. *It holds that*

$$\left\|(P_{k-1} - \bar{P}_{k-1})v\right\|^2 \le Ch_k^2 \sum_{E\in\mathcal{E}_k} \|\nabla v\|^2_{L^2(E)}, \qquad v \in V_k, \tag{2.18}$$

where C is independent of k.

PROOF. Note that, by the definitions of P_{k-1} and $\bar{P}_{k-1}$,

$$a_{k-1}(P_{k-1}v, w) = a_k(v, I_k w), \qquad \forall\ w \in V_{k-1}, \tag{2.19}$$

$$a_{k-1}(\bar{P}_{k-1}v, w) = a_k(v, I_k w), \qquad \forall\ w \in U_{k-1}. \tag{2.20}$$

Let $f_0 \in V_{k-1}$ satisfy

$$(f_0, w) = a_k(v, I_k w), \qquad \forall\ w \in V_{k-1}. \tag{2.21}$$

Then, by an inverse inequality, Proposition 2.1, and the Poincaré inequality, we see that

$$\begin{aligned}\|f_0\|^2 &= a_k(v, I_k f_0)\\ &\le C\left(\sum_{E\in\mathcal{E}_k}\|\nabla v\|^2_{L^2(E)}\right)^{1/2}\left(\sum_{E\in\mathcal{E}_k}\|\nabla(I_k f_0)\|^2_{L^2(E)}\right)^{1/2}\\ &\le Ch_k^{-1}\left(\sum_{E\in\mathcal{E}_k}\|\nabla v\|^2_{L^2(E)}\right)^{1/2}\|f_0\|.\end{aligned}$$

That is,

$$\|f_0\| \le Ch_k^{-1}\left(\sum_{E\in\mathcal{E}_k}\|\nabla v\|^2_{L^2(E)}\right)^{1/2}. \tag{2.22}$$

Let $z \in H_0^1(\Omega)$ satisfy (2.15). Then we see that $P_{k-1}v$ and $\bar{P}_{k-1}v$ are the finite element approximations to z in V_{k-1} and U_{k-1}, respectively. Thus (2.18) follows as in Lemma 2.3. □

Proposition 2.5. *Condition* (2.7) *is satisfied with* $\alpha = 1/2$.
Proof. By the definition of A_k, we have, for $v \in V_k$,

$$a_k\left((I - I_k P_{k-1})v, v\right) = (A_k v, (I - I_k P_{k-1})v) \le C\|A_k v\|\ \|(I - I_k P_{k-1})v\|. \tag{2.23}$$

Note that, by the triangle inequality,

$$\|(I - I_k P_{k-1})v\| \le \left\|(I - \bar{P}_{k-1})v\right\| + \|(I - I_k)P_{k-1}v\| + \left\|(P_{k-1} - \bar{P}_{k-1})v\right\|. \tag{2.24}$$

Since by definition, P_{k-1} is the adjoint of I_k, condition (2.6) implies that there is constant C independent on k such that

$$a_{k-1}(P_{k-1}v, P_{k-1}v) \le Ca_k(v, v), \qquad \forall\ v \in V_k. \tag{2.25}$$

Then, by (2.23)–(2.25) and Lemmas 2.2–2.4, we see that

$$|a_k\left((I - I_k P_{k-1})v, v\right)| \le Ch_k\|A_k v\| a_k(v, v)^{1/2}, \qquad v \in V_k. \tag{2.26}$$

By Gerschgorin's theorem [9], [12],

$$\lambda_k \le Ch_k^{-2}. \tag{2.27}$$

The regularity and approximation condition (2.7) with $\alpha = 1/2$ follows from (2.26) and (2.27). □

The convergence rate for the multigrid algorithm 2.1 on the kth level is measured by a convergence factor δ_k satisfying

$$|a_k\left((I - B_k A_k)v, v\right)| \le \delta_k a_k(v, v), \qquad \forall\ v \in V_k. \tag{2.28}$$

Theorem 2.6. (i) *Define B_k by $p = 2$ for all k in Algorithm* 2.1. *Then there exists $C > 0$ independent on k such that for m big enough, but independent on k*

$$\delta_k \le \delta \equiv \frac{C}{C + \sqrt{m}}. \tag{2.29}$$

(ii) *Define B_k by $p = 1$ and $m(k) = 2^{K-k}$ for $k = 2, \dots, K$ in Algorithm* 2.1. *Then there are η_0, $\eta_1 > 0$, independent on k, such that*

$$\eta_0 a_k(v, v) \le a_k(B_k A_k v, v) \le \eta_1 a_k(v, v), \qquad \forall v \in V_k, \tag{2.30}$$

with $\eta_0 \ge m(k)^{1/2}/(C + m(k)^{1/2})$ and $\eta_1 \le (C + m(k)^{1/2})/m(k)^{1/2}$.

The proof of this theorem follows from Propositions 2.1 and 2.5 and Theorems 6 and 7 in [7]. The constant C in Theorem 2.6 depends only on C_α in (2.7) and the constant appearing in the smoothing estimates [7]. From Theorem 2.6, we have an optimal convergence property of the $\mathcal{W}$-cycle and a uniform condition number estimate for the variable $\mathcal{V}$-cycle preconditioner.

2.2. The mixed multigrid algorithm. In this section we consider multigrid algorithms for a mixed finite element method for (2.1). The Raviart-Thomas spaces [22] over triangles is given by

$$\begin{aligned}
\Lambda_h &= \left\{ v \in (L^2(\Omega))^2 : v|_E = \left(a_E^1 + a_E^2 x,\ a_E^3 + a_E^2 y \right),\ a_E^i \in \mathbb{R},\ E \in \mathcal{E}_h \right\}, \\
W_h &= \left\{ w \in L^2(\Omega) : w|_E \text{ is constant for all } E \in \mathcal{E}_h \right\}, \\
L_h &= \left\{ \mu \in L^2(\partial\mathcal{E}_h) : \mu|_e \text{ is constant},\ e \in \partial\mathcal{E}_h;\ \mu|_e = 0,\ e \subset \partial\Omega \right\},
\end{aligned}$$

where $\partial\mathcal{E}_h$ denotes the set of all interior edges. Then the hybrid form of the mixed method for (2.1) is to seek $(\sigma_h, u_h, \lambda_h) \in \Lambda_h \times W_h \times L_h$ such that

$$\begin{aligned}
\sum_{E \in \mathcal{E}_h} (\nabla \cdot \sigma_h, w)_E + (c_h u_h, w) &= (f, w), && \forall\ w \in W_h, \\
(\mathcal{X}_h \sigma_h, v) - \sum_{E \in \mathcal{E}_h} \left[(u_h, \nabla \cdot v)_E - (\lambda_h, v \cdot \nu_E)_{\partial E} \right] &= 0, && \forall\ v \in \Lambda_h, \\
\sum_{E \in \mathcal{E}_h} (\sigma_h \cdot \nu_E, \mu)_{\partial E} &= 0, && \forall\ \mu \in L_h,
\end{aligned} \tag{2.31}$$

where ν_E denotes the unit outer normal to E, $\mathcal{X}_h = \mathcal{Q}_h \mathcal{A}^{-1}$ (component-wisely), $c_h = \mathcal{Q}_h c$, and $\mathcal{Q}_h$ denotes the $L^2(\Omega)$ projection operator onto W_h. The solution σ_h is introduced to approximate the vector field

$$\sigma = -\mathcal{A} \nabla u,$$

which is the variable of primary interest in many applications. Since σ lies in the space

$$H(\mathrm{div}; \Omega) = \left\{ v \in (L^2(\Omega))^2 : \nabla \cdot v \in L^2(\Omega) \right\},$$

and we do not require that Λ_h be a subspace of $H(\mathrm{div}; \Omega)$, the last equation in (2.31) is used to enforce that the normal components of σ_h are continuous across the interior edges in $\partial\mathcal{E}_h$, so in fact $\sigma_h \in H(\mathrm{div}; \Omega)$. Also, the projection of the coefficients $\mathcal{A}^{-1}, c$ into the space W_h is introduced in (2.31). The projection of coefficients gives us considerable computational savings, without any loss of accuracy [13]. Furthermore, it can be used to establish an equivalence between the triangular nonconforming method and the mixed method (2.31). This equivalence has been obtained in [12], [14], [15] in the case of $c = 0$. We now extend it to the present situation.

There is no continuity requirement on the spaces Λ_h and W_h, so σ_h and u_h can be locally (element by element) eliminated from (2.31). In fact, applying the ideas in [12], (2.31) can be algebraically condensed to the symmetric, positive definite system for the Lagrange multiplier λ_h:

$$M\lambda = F, \tag{2.32}$$

where the contributions of the triangle E to the stiffness matrix M and the right-hand side F are

$$m_{ij}^E = \overline{\nu}_E^i \beta^E \overline{\nu}_E^j + \frac{(c,1)_E}{3}\delta_{ij}, \quad F_i^E = -\frac{(J_E^f, \overline{\nu}_E^i)_E}{|E|} + (J_E^f, \nu_E^i)_{e_E^i},$$

where ν_E^i denotes the outer unit normal to the edge e_E^i (E has three edges), $\overline{\nu}_E^i = |e_E^i|\nu_E^i$, $|e_E^i|$ is the length of e_E^i, $\beta^E = (((\mathcal{X}_h)_{ij}, 1)_E)^{-1}$, $J_E^f = (f,1)_E(x,y)/(2|E|)$, δ_{ij} is the Kronecker symbol, and $|E|$ denotes the area of E. After the computation of λ_h, (2.31) can be used to recover σ_h and u_h on each element.

As shown in [12] for the case where $c = 0$, the system (2.32) corresponds to the system arising from the triangular nonconforming finite element method. This is proven below.

Theorem 2.7. *Let $f_h = \mathcal{Q}_h f$. Then* (2.32) *corresponds to the linear system produced by the problem: Find $\psi_h \in V_h$ such that*

$$\tilde{a}_h(\psi_h, \varphi) = (f_h, \varphi), \quad \forall \varphi \in V_h, \tag{2.33}$$

where $\tilde{a}_h(\psi_h, \varphi) = \sum_{E\in\mathcal{E}_h} (\mathcal{X}_h^{-1}\nabla\psi_h, \nabla\varphi)_E + (c_h\psi_h, \varphi)$.

Proof. Let $\{\psi_1^h, \dots, \psi_{m_h}^h\}$ be the basis of V_h such that each ψ_i^h equals 1 at exactly one midpoint and equals 0 at all other midpoints. Then for each $E \in \mathcal{E}_h$ we have

$$\psi_i^h|_E = \frac{1}{|E|}\overline{\nu}_E^i \cdot ((x,y) - p_l), \quad i \neq l,$$

for some midpoint p_l. Also, note that for any linear functions ψ and ϕ on a triangle $E \in \mathcal{E}_h$,

$$(\psi, \phi)_E = \frac{1}{3}|E|\sum_{s=1}^{3}\psi(p_s)\phi(p_s),$$

where the p_s's are the midpoints of the edges of E. Then we see that

$$(\mathcal{X}_h^{-1}\nabla\psi_i^h, \nabla\psi_j^h)_E + (c_h\psi_i^h, \psi_j^h)_E = \overline{\nu}_E^i\beta^E\overline{\nu}_E^j + \frac{(c,1)_E}{3}\delta_{ij},$$

which is m_{ij}^E, and

$$\begin{aligned} F_i^E &= -\tfrac{(J_E^f, \overline{\nu}_E^i)_E}{|E|} + (J_E^f, \nu_E^i)_{e_E^i} \\ &= -\tfrac{(f,1)_E}{2|E|}(1, \psi_i^h)_E + \tfrac{(f,1)_E}{2|e_E^i|}(\psi_i^k, 1)_{e_E^i} \\ &= \tfrac{(f,1)_E}{|E|}(1, \psi_i^h)_E, \end{aligned}$$

which is $(f_h, \psi_i^h)_E$. □

It follows from this theorem that Algorithm 2.1 can be used to solve (2.32), i.e., to solve the mixed method (2.31). Moreover, the previous results on the $\mathcal{W}$-cycle and variable $\mathcal{V}$-cycle algorithms hold.

3. The Nonsymmetric and/or Indefinite Problem. In this section we develop multigrid algorithms for the nonsymmetric and indefinite problem. In §3.1, we consider the problem of existence and uniqueness of the solution to the discrete system. The nonconforming multigrid algorithm is analyzed in §3.2, and the mixed methods are described in §3.3.

3.1. Preliminaries. In this subsection we consider the nonconforming finite element method applied to the nonsymmetric and indefinite problem. We consider as our model problem the following equation:

$$\begin{aligned} -\nabla\cdot(\mathcal{A}\nabla u) + \mathcal{B}\cdot\nabla u + cu &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{3.1}$$

with the same notation as in (2.1). We assume that the symmetric coefficient $\mathcal{A}$ satisfies (2.2a), $\mathcal{B}$ is continuously differentiable on $\overline{\Omega}$ and piecewisely C^2 with the sum of the second-order derivatives over pieces being bounded, and $|c|$ is bounded (need not satisfy (2.2b)). Further, we assume that the solution of (3.1) exists.

The bilinear form $A(\cdot,\cdot)$ is now given by

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

The solution $u \in H_0^1(\Omega)$ of (3.1) then satisfies

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

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

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

which corresponds to $a(\cdot,\cdot)$ with $c = 1$ of the second section. The difference form is indicated by

$$D(v,w) = A(v,w) - \widehat{A}(v,w). \tag{3.3}$$

With the same V_h as before, we define a mesh-dependent form $A_h(\cdot,\cdot)$ by

$$A_h(v,w) = \sum_{E\in\mathcal{E}_h} \{(\mathcal{A}\nabla v, \nabla w)_E + (\mathcal{B}\cdot\nabla v, w)_E\} + (cv,w), \qquad v,w \in V_h \oplus H_0^1(\Omega).$$

The corresponding symmetric form is denoted by $\widehat{A}_h(\cdot,\cdot)$. The nonconforming finite element solution $u_h \in V_h$ of (3.1) is given by

$$A_h(u_h, v) = (f, v), \qquad \forall\ v \in V_h. \tag{3.4}$$

The norm induced by $\left(\widehat{A}_h(v,v)\right)^{1/2}$ for $v \in V_h \oplus H_0^1(\Omega)$ is equivalent to the norm $\left(\sum_{E\in\mathcal{E}_h} ||\nabla v||^2_{L^2(E)} + ||v||^2\right)^{1/2}$. Thus, we define

$$||v||_h = \widehat{A}_h(v,v)^{1/2}, \qquad \forall v \in V_h \oplus H_0^1(\Omega).$$

Let us note the inequality

$$|A_h(v,w)| \le C||v||_h||w||_h, \qquad \forall v, w \in V_h \oplus H_0^1(\Omega). \tag{3.5}$$

It is not hard to show the Garding inequality

$$C_1||v||_h^2 - C_2||v||^2 \le |A_h(v,v)|, \qquad \forall v \in V_h \oplus H_0^1(\Omega). \tag{3.6}$$

Lemma 3.1. *Problem* (3.4) *has a unique solution for h sufficiently small.*
PROOF. Let U_h be the same as before, i.e., the P_1 conforming finite element space, and let $z_h \in U_h$ satisfy

$$A_h(z_h, v) = (f, v), \qquad \forall v \in U_h. \tag{3.7}$$

Then it follows from (3.4)–(3.6) that

$$\begin{aligned} C_1&||u_h - z_h||_h^2 - C_2||u_h - z_h||^2 \\ &\le |A_h(u_h - z_h, u_h - z_h)| \\ &\le |A_h(u - z_h, u_h - z_h)| + |A_h(u_h - u, u_h - z_h)| \\ &= |A_h(u - z_h, u_h - z_h)| + |f(u_h - z_h) - A_h(u, u_h - z_h)| \\ &\le C_3||u - z_h||_h||u_h - z_h||_h + |f(u_h - z_h) - A_h(u, u_h - z_h)|, \end{aligned}$$

so that, by dividing through by $C_1||u_h - z_h||_h$, we have

$$\begin{aligned} ||u_h - z_h||_h &- \tfrac{C_2}{C_1}||u_h - z_h|| \\ &\le \tfrac{C_3}{C_1}||u - z_h||_h + \tfrac{1}{C_1}\sup_{v\in V_h} \tfrac{|f(v) - A_h(u,v)|}{||v||_h}. \end{aligned} \tag{3.8}$$

Note that, since (3.1) and the associated adjoint problem are assumed to be uniquely solvable, a duality argument [20] can be used to show the estimate

$$||u_h - z_h|| \le C_4 h||u_h - z_h||_h. \tag{3.9}$$

Thus, by (3.8) and (3.9), we see that, if $h < C_1/(C_2C_4)$,

$$||u_h - z_h||_h \le C\left(||u - z_h||_h + \sup_{v\in V_h} \frac{|f(v) - A_h(u,v)|}{||v||_h}\right). \tag{3.10}$$

From a known result for the conforming method (3.7) [23], there exists $h_0 > 0$ such that $z_h = 0$ is the only solution corresponding to $u = 0$ for $h < h_0$. Therefore, for $h < \min(h_0, C_1/(C_2C_4))$, it follows from (3.10) that the homogeneous nonconforming equation (3.4) has a unique solution $u_h = 0$. Since V_h is finite dimensional, this also implies existence. □

Define the projection operator $\mathcal{P}_h : H_0^1(\Omega) \to V_h$ by

$$A_h(\mathcal{P}_h u, v) = A_h(u, v), \qquad \forall v \in V_h.$$

It follows in the usual way from (3.9), (3.10), and the corresponding result for the conforming finite elements that, if the solution of (3.1) satisfies regularity estimates of the form

$$||u||_{1+\alpha} \le C||f||_{-1+\alpha}, \tag{3.11}$$

then

$$||u - \mathcal{P}_h u|| \le Ch^{\alpha}||u - \mathcal{P}_h u||_h, \tag{3.12}$$

and

$$||\mathcal{P}_h u||_h \le C||u||_h. \tag{3.13}$$

In the case where regularity estimates of the form of (3.11) are not known to hold, it can be shown as in the conforming case [24] that, given $\epsilon > 0$, there exists an $h_0(\epsilon) > 0$ such that for $0 < h \le h_0$,

$$||u - \mathcal{P}_h u|| \le \epsilon||u - \mathcal{P}_h u||_h, \tag{3.14}$$

and (3.13) is satisfied. The above ϵ will appear in our later convergence result.

3.2. The nonconforming multigrid algorithm. The family of partitions $\{\mathcal{E}_k\}_{k=1}^K$ is constructed in the same manner as before. Let the mesh size of $\mathcal{E}_1$ be d_1; then, by similarity, the mesh size of $\mathcal{E}_k$ is $2^{1-k}d_1$. From Lemma 3.1, for (3.4) to be well behaved, the approximation grid must be sufficient fine. As in the conforming case [5], we shall require that the coarsest grid in the multilevel algorithm be sufficient fine. Toward that end, let the coarse mesh size be denoted by an integer L. Then the space V_k has a mesh size of $h_k = 2^{1-L-k}d_1 = 2^{1-k}h_1$, $k = 1, \ldots, K$.

As noted in [5] and demonstrated in our experiments in §4, in practice, the coarse grid can be taken considerably coarser than the solution grid. The reason for this is that we can only expect that the discrete errors depend monotonically on the grid sizes; consequently, if the fine grid approximation is reasonably accurate, we expect that there exists a sequence of coarser grids whose approximations are well defined.

For $k = 2, \ldots, K$, we define the projection operators $P_{k-1} : V_k \oplus H^1(\Omega) \to U_{k-1}$ and $P^0_{k-1} : L^2(\Omega) \to U_{k-1}$ by

$$A_{k-1}(P_{k-1}v, w) = A_k(v, w), \qquad \forall w \in U_{k-1},$$

and

$$(P^0_{k-1}v, w) = (v, w), \qquad \forall w \in U_{k-1}.$$

Also, for each $k = 1, \ldots, K$, we introduce the conforming discretization operator $M_k : U_k \to U_k$ by

$$(M_k v, w) = A_k(v, w), \qquad \forall w \in U_k.$$

The nonconforming discretization operator on V_k is still indicated by A_k, for expositional convenience.

We first describe a simplest $\mathcal{V}$-cycle multigrid algorithm for iteratively computing the solution of the conforming method (3.7). The next two algorithms are slightly different from Algorithm 2.1. Specifically, we smooth only as we proceed to coarser grids. So they are a special case of Algorithm 2.1. Alternatively, we could consider a multigrid algorithm with just post-smoothing or both pre- and post-smoothing, as in Algorithm 2.1. These algorithms can be analyzed analogously, and are not considered here.

The following algorithm iteratively defines a multigrid operator $N_k : U_k \to U_k$. The operator $R_k : U_k \to U_k$ is a linear smoothing operator. A variety of examples for R_k has been given in [5]; we do not repeat these examples in the paper.

Multigrid Algorithm 3.1. Set $N_1 = M_1^{-1}$. For $1 < k \le K$, assume that N_{k-1} has been defined and define $N_k g$ for $g \in U_k$ by

1. Set $x_k = R_k g$.
2. Define $N_k g = x_k + q$, where $q \in U_{k-1}$ is given by

$$q = N_{k-1} P^0_{k-1} (g - M_k x_k).$$

We now define the $\mathcal{V}$-cycle algorithm for the nonconforming method (3.4), which determines a multigrid operator $B_K : V_K \to V_K$. The operator $Q_K : V_K \to V_K$ below is a linear smoothing operator. Examples of this operator will be given in §3.2.1.

MULTIGRID ALGORITHM 3.2. If $K = 1$, set $B_1 = A_1^{-1}$. If $K > 1$, define $B_K g$ for $g \in V_K$ by

1. Set $x_K = Q_K g$.
2. Define $B_K g = x_K + q$, where $q \in U_{K-1}$ is given by

$$q = N_{K-1} P^0_{K-1} (g - A_K x_K).$$

We remark that the coarse-grid correction in Algorithm 3.2 is defined on the conforming finite element spaces. That is, it is of the second type of multigrid algorithm, mentioned in the introduction. It will be analyzed in §3.2.2.

3.2.1. Smoothers. The smoothers presented in this subsection are the variants of those for the conforming finite element method (see, e.g., [5]). We first describe three smoothers which are based on the symmetric problem, and then three smoothers which correspond to the original nonsymmetric and indefinite problem.

The simplest smoother is given in the next example.

Example 2. We define

$$Q_K = \lambda_K^{-1} I,$$

where λ_K is the largest eigenvalue of $\widehat{A}_K$.

The following two smoothers are defined in terms of subspace decompositions. To this end, let

$$V_K = \sum_{j=1}^{l(K)} V_{j,K},$$

where $V_{j,K}$ is the one-dimensional subspace spanned by a nodal basis function or the one spanned by the nodal basis functions along a line, and $l(K)$ is the number of such spaces. The smoothers in Examples 3 and 4 below are additive and multiplicative, respectively.

Example 3. We define

$$Q_K = \gamma \sum_{j=1}^{l(K)} \widehat{A}_{j,K}^{-1} Q_{j,K},$$

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

$$(\widehat{A}_{j,K} v, \varphi) = \widehat{A}_K(v, \varphi), \qquad \forall \varphi \in V_{j,K},$$

$Q_{j,K} : V_{j,K} \to V_{j,K}$ is the projection operator on $V_{j,K}$ with respect to the L^2 inner product $(\cdot, \cdot)$, and the constant γ is a scaling factor which is chosen to ensure that the smoothing property is satisfied [5].

Example 4. Given $g \in V_K$, we define

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

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

3. Set $Q_K g = x_{l(K)}$.

The following example corresponds to the first example, and the later two examples are closely related to Examples 3 and 4.

Example 5. We define

$$Q_K = \lambda_K^{-2} A_K^t,$$

where λ_K is as in Example 2 and A_K^t is the adjoint operator of A_K with respect to the L^2 inner product $(\cdot, \cdot)$.

Example 6. We define

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

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

$$(A_{j,K} v, \varphi) = A_K(v, \varphi), \qquad \forall \varphi \in V_{j,K},$$

and $Q_{j,K} : V_{j,K} \to V_{j,K}$ and γ are as in Example 3.

Example 7. 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 $Q_K g = x_{l(K)}$.

3.2.2. Analysis of the multigrid algorithm. We now provide a convergence analysis for Algorithm 3.2 with the smoothers given in Examples 2–7 in the framework of [5]. All of their analysis is based on perturbation from the uniform convergence estimate for the multigrid algorithm applied to the symmetric problem. Essential use in [5] is made of a product representation of the error operator and two properties of the difference form $D(\cdot, \cdot)$ (see (2.4) in [5]). In this section we shall show that our error operator has the same structure (see Lemma 3.3 below) and the form $D(\cdot, \cdot)$ satisfies the same properties (see Lemma 3.4 below). Thus the convergence analysis given in [5] carries over to Algorithm 3.2 since the uniform iterative convergence estimate for Algorithm 3.2 applied to the symmetric problem has been shown in [12] and [19].

Lemma 3.2. *It holds that*

$$B_K = Q_K + N_{K-1} P_{K-1}^0 (I - A_K Q_K),$$

and

$$N_k = R_k + N_{k-1} P_{k-1}^0 (I - M_k R_k), \qquad k = 2, \ldots, K.$$

This lemma can be easily seen from Algorithms 3.1 and 3.2.

Lemma 3.3. *Let $P_K = I$, $T_1 = P_1$, $T_k = R_k M_k P_k$, $k = 2, \dots, K-1$, and $T_K = Q_K A_K P_K$. Then*

$$I - B_K A_K = (I - T_1)(I - T_2) \dots (I - T_K). \tag{3.15}$$

Proof. From the definitions of P_{k-1} and P^0_{k-1}, we see that

$$\begin{aligned} &P^0_{K-1} A_K = M_{K-1} P_{K-1}, \\ &P^0_{k-1} M_k = M_{k-1} P_{k-1}, \qquad k = 2, \dots, K-1, \\ &P_{k-1} P_k = P_{k-1}, \qquad k = 2, \dots, K. \end{aligned}$$

Then it follows from Lemma 3.2 that

$$I - B_K A_K = (I - N_{K-1} M_{K-1} P_{K-1})(I - Q_K A_K),$$

and

$$\begin{aligned} &I - N_{K-1} M_{K-1} P_{K-1} \\ &\quad = I - P_{K-1} + (I - N_{K-1} M_{K-1}) P_{K-1} \\ &\quad = I - P_{K-1} + (I - N_{K-2} M_{K-2} P_{K-2})(I - R_{K-1} M_{K-1}) P_{K-1} \\ &\quad = (I - N_{K-2} M_{K-2} P_{K-2})(I - P_{K-1} + (I - R_{K-1} M_{K-1}) P_{K-1}) \\ &\quad = (I - N_{K-2} M_{K-2} P_{K-2})(I - T_{K-1}). \end{aligned}$$

Therefore, a straightforward mathematical induction argument shows the desired result (3.15) since $P_K = I$. □

The product representation of the error operator in Lemma 3.3 is a fundamental ingredient in the convergence analysis. The other important ingredients are the following properties of the difference operator $D(\cdot, \cdot)$. They are trivial in the conforming case; however, as shown below, the second property is not so straightforward in the nonconforming case.

Lemma 3.4. *Under the above assumption on the coefficient $\mathcal{B}$, there is a constant C independent on k such that*

$$|D(v, w)| \le C ||v||_k ||w||, \qquad \forall v, w \in V_k, \tag{3.16}$$

and

$$|D(v, w)| \le C ||w||_k ||v||, \qquad \forall v, w \in V_k. \tag{3.17}$$

Proof. (3.16) directly follows from the definition of $D(v, w)$:

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

To prove (3.17), we apply integration by parts on each finite element to see that

$$D(v, w) = \sum_{E \in \mathcal{E}_k} \{ (\mathcal{B} \cdot \nu_E v, w)_{\partial E} - (\nabla \cdot \mathcal{B} w + \mathcal{B} \cdot \nabla w, v)_E \} + ((c - 1) v, w). \tag{3.18}$$

Evidently, it suffices to estimate the terms over edges.

Let E_1, $E_2 \in \mathcal{E}_k$ share an edge e with midpoint m^k, and let e have the parametric representation $x = x(t)$, $y = y(t)$ with t as parameter. They are linear functions of t. Then, by the midpoint rule and the continuity at midpoints on the elements of V_k, we find that

$$\begin{aligned}
&\int_e (\mathcal{B}\cdot\nu_{E_1} vw)|_{E_1} ds + \int_e (\mathcal{B}\cdot\nu_{E_2} vw)|_{E_2} ds \\
&\quad = |e|\{(\mathcal{B}\cdot\nu_{E_1} vw)|_{E_1}(m^k) + (\mathcal{B}\cdot\nu_{E_2} vw)|_{E_2}(m^k)\} \\
&\qquad + \tfrac{|e|^3}{24}\left\{\left(\tfrac{dx}{dt}\right)^2 + \left(\tfrac{dy}{dt}\right)^2\right\}^{-1}\left\{\tfrac{d^2}{dt^2}(\mathcal{B}\cdot\nu_{E_1} vw)|_{E_1}(\xi_1^k) + \tfrac{d^2}{dt^2}(\mathcal{B}\cdot\nu_{E_2} vw)|_{E_2}(\xi_2^k)\right\} \\
&\quad = \tfrac{|e|^3}{24}\left\{\left(\tfrac{dx}{dt}\right)^2 + \left(\tfrac{dy}{dt}\right)^2\right\}^{-1}\left\{\tfrac{d^2}{dt^2}(\mathcal{B}\cdot\nu_{E_1} vw)|_{E_1}(\xi_1^k) + \tfrac{d^2}{dt^2}(\mathcal{B}\cdot\nu_{E_2} vw)|_{E_2}(\xi_2^k)\right\},
\end{aligned}$$

for some points ξ_1^k, $\xi_2^k \in e$. Note that, since v and w are piecewisely linear, for $i = 1, 2$,

$$\frac{d^2}{dt^2}(\mathcal{B}\cdot\nu_{E_i} vw) = \frac{d^2}{dt^2}(\mathcal{B}\cdot\nu_{E_i})vw + 2\frac{d}{dt}(\mathcal{B}\cdot\nu_{E_i})\frac{d}{dt}(vw) + 2(\mathcal{B}\cdot\nu_{E_i})\frac{dv}{dt}\frac{dw}{dt}.$$

Also, by the chain rule, we have with any function $g = g(x(t), y(t))$:

$$\frac{dg}{dt} = \frac{\partial g}{\partial x}\frac{dx}{dt} + \frac{\partial g}{\partial y}\frac{dy}{dt},$$

and

$$\frac{d^2 g}{dt^2} = \frac{\partial^2 g}{\partial x^2}\left(\frac{dx}{dt}\right)^2 + 2\frac{\partial^2 g}{\partial x\partial y}\frac{dx}{dt}\frac{dy}{dt} + \frac{\partial^2 g}{\partial y^2}\left(\frac{dy}{dt}\right)^2,$$

since e is a line segment. Consequently, we see that

$$\begin{aligned}
&\left|\int_e (\mathcal{B}\cdot\nu_{E_1} vw)|_{E_1} ds + \int_e (\mathcal{B}\cdot\nu_{E_2} vw)|_{E_2} ds\right| \\
&\quad \le \tfrac{C|e|^3}{24}\textstyle\sum_{i=1}^2 \left(|v| + |\tfrac{\partial v}{\partial x}| + |\tfrac{\partial v}{\partial y}|\right)\left(|w| + |\tfrac{\partial w}{\partial x}| + |\tfrac{\partial w}{\partial y}|\right)(\xi_i^k).
\end{aligned}$$

This, together with the Cauchy-Schwarz inequality, an inverse inequality, and the fact that v and w are piecewisely linear, implies that

$$\begin{aligned}
&\left|\textstyle\sum_{E\in\mathcal{E}_k}(\mathcal{B}\cdot\nu_E v, w)_{\partial E}\right| \\
&\quad \le Ch_k^3 \textstyle\sum_{e\in\partial\mathcal{E}_k}\left(|v| + |\tfrac{\partial v}{\partial x}| + |\tfrac{\partial v}{\partial y}|\right)\left(|w| + |\tfrac{\partial w}{\partial x}| + |\tfrac{\partial w}{\partial y}|\right)(\xi_e^k) \\
&\quad \le Ch_k(||v|| + ||v||_k)(||w|| + ||w||_k) \\
&\quad \le C||w||_k||v||,
\end{aligned}$$

which, by (3.18), yields the desired result (3.17). Thus the proof is complete. □

With Lemmas 3.3 and 3.4 and the arguments presented in Theorems 5.2–5.6 of [5], we have the following theorem.

Theorem 3.5. *Let Q_K be one of the smoothers defined in Examples 2–7. Then, given $\epsilon > 0$, there exists an $h_0 > 0$ such that for $h_1 \le h_0$,*

$$\widehat{A}_K(Ev, Ev) \le \delta^2 \widehat{A}_K(v, v), \qquad \forall v \in V_K,$$

where $E = I - B_K A_K$, $\delta = \widehat{\delta} + C(h_1 + \epsilon)$, *and* $\widehat{\delta}$ *is less than one and independent on* K.

We remark that $\widehat{\delta}$ comes from the uniform convergence estimate of Algorithm 3.2 applied to the symmetric problem [12], [19].

3.3. The mixed multigrid algorithm. As in the last section, we now consider a mixed finite element method for numerically solving (3.1). With the same spaces Λ_h, W_h, and L_h as in §2.2, the hybrid form of the mixed finite element solution to (3.1) is $(\sigma_h, u_h, \lambda_h) \in \Lambda_h \times W_h \times L_h$ satisfying

$$\begin{aligned} \sum_{E\in\mathcal{E}_h} (\nabla\cdot\sigma_h, w)_E - (\mathcal{Y}_h\cdot\sigma_h, w) + (c_h u_h, w) &= (f, w), && \forall\ w \in W_h, \\ (\mathcal{X}_h\sigma_h, v) - \sum_{E\in\mathcal{E}_h} [(u_h, \nabla\cdot v)_E - (\lambda_h, v\cdot\nu_E)_{\partial E}] &= 0, && \forall\ v \in \Lambda_h, \\ \sum_{E\in\mathcal{E}_h} (\sigma_h\cdot\nu_E, \mu)_{\partial E} &= 0, && \forall\ \mu \in L_h, \end{aligned} \tag{3.19}$$

where $\mathcal{Y}_h = \mathcal{X}_h(\mathcal{Q}_h\mathcal{B})$ and the other notation is the same as before. We recall that $\mathcal{Q}_h$ denotes the $L^2(\Omega)$ projection onto W_h.

Again, after an algebraical condensation, system (3.19) can be reduced to a linear system for the Lagrange multiplier λ_h:

$$M\lambda = F, \tag{3.20}$$

where the contributions of the element E to the stiffness matrix M and the right-hand side F are

$$\begin{aligned} m_{ij}^E &= \overline{\nu}_E^i \beta^E \overline{\nu}_E^j + \tfrac{1}{3}(\mathcal{Q}_h\mathcal{B})|_E\cdot\overline{\nu}_E^i + \tfrac{1}{3}(c,1)_E\delta_{ij}, \\ F_i^E &= -\tfrac{(J_E^f, \overline{\nu}_E^i)_E}{|E|} + (J_E^f, \nu_E^i)_{e_E^i}. \end{aligned}$$

Furthermore, with the same argument as in Theorem 2.7, we have the next result.

Theorem 3.6. *System* (3.20) *corresponds to the linear system arising from the nonconforming problem: Find* $\psi_h \in V_h$ *such that*

$$\tilde{A}_h(\psi_h, \varphi) = (f_h, \varphi), \quad \forall\varphi \in V_h,$$

where

$$\tilde{A}_h(\psi_h, \varphi) = \sum_{E\in\mathcal{E}_h} \left\{(\mathcal{X}_h^{-1}\nabla\psi_h, \nabla\varphi)_E + (\mathcal{Q}_h\mathcal{B}\cdot\nabla\psi_h, \varphi)_E\right\} + (c_h\psi_h, \varphi).$$

It thus follows that Algorithm 3.2 can be exploited to solve the system arising from the mixed method (3.19), and the convergence result in Theorem 3.5 is valid.

4. Numerical Examples. We report the results of a couple of numerical examples to illustrate the theory developed in the earlier sections and to show a comparison between the results obtained here and those generated by the well established conforming finite element and finite difference multigrid algorithms [4], [7]. We first compute a symmetric problem, and then a nonsymmetric and indefinite problem. Special efforts are made in the second example.

h_K	(κ_v, δ_v)	(κ_w, δ_w)	$(\kappa_{vv}, \delta_{vv})$
1/8	(1.48, .43)	(1.46, .40)	(1.45, .39)
1/16	(1.64, .46)	(1.47, .42)	(1.47, .41)
1/32	(1.81, .50)	(1.48, .43)	(1.48, .43)
1/64	(1.86, .51)	(1.48, .43)	(1.48, .43)
1/128	(1.96, .54)	(1.48, .43)	(1.48, .43)

Table 1. Convergence results for Example 8.

Example 8. In the first example we consider the following equation on the unit square:

$$\begin{aligned} -\nabla \cdot (\mathcal{A}\nabla u) &= f \quad \text{in } \Omega = (0,1)^2, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{4.1}$$

We approximate the solution to (4.1) using the triangular nonconforming method (2.4). In this example, conditions (2.6) and (2.7) are satisfied. The analysis of the second section guarantees that the condition number of $B_K A_K$ for the variable $\mathcal{V}$-cycle algorithm can be bounded independently on the number of levels and the $\mathcal{W}$-cycle algorithm has an optimal convergence property. Table 1 gives the condition number κ for the system $B_K A_K$ and the reduction factor for the system $I - B_K A_K$ as a function of the mesh size on the finest grid, where the $\mathcal{V}$-cycle, $\mathcal{W}$-cycle, and variable $\mathcal{V}$-cycle algorithms are indicated by (κ_v, δ_v), (κ_w, δ_w), and $(\kappa_{vv}, \delta_{vv})$, respectively. The $\mathcal{V}$-cycle and $\mathcal{W}$-cycle schemes use one smoothing step. (To see how the convergence rate depends upon the number of the smoothing steps, refer to [12].) For all of the runs, the coarse grid is of size $h_1 = 1/2$. Point Jacobi smoothing is applied. As noticed in the conforming case [7], the variable $\mathcal{V}$-cycle and the $\mathcal{W}$-cycle algorithms have essentially identical computational results. This is due to the fact that both algorithms have exactly the same number of total smoothings on each grid in the multi-level iteration. While there is no complete theory for the $\mathcal{V}$-cycle algorithm, it is of practical interest that the condition numbers for this cycle remain relatively small, but the convergence rate deteriorates with the mesh size. Finally, compared with the numerical results obtained in [4], [7], we see that the nonconforming multigrid algorithms in fact compare favorably with these standard multigrid algorithms.

Example 9. In the second example we consider the nonsymmetric and indefinite problem

$$\begin{aligned} -\nabla \cdot (\mathcal{A}\nabla u) + \mathcal{B} \cdot \nabla u + cu &= f \quad \text{in } \Omega = (0,1)^2, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{4.2}$$

In (4.2), the symmetric and positive definite part is taken as in (4.1), but three different choices for the constants $\mathcal{B}$ and c are made in our experiments:

$$\mathcal{B} = (c, c),$$

where $c = -5$, 10, and 15.

c	(h_K, h_1)	$\overline{\delta}_v$	$\|\|E\|\|$
-5	(1/16, 1/4)	0.639436	0.74
10	(1/16, 1/4)	0.634469	0.74
15	(1/16, 1/8)	NC	-
-5	(1/32, 1/4)	0.737174	0.77
10	(1/32, 1/4)	0.735298	0.77
15	(1/32, 1/8)	0.732282	0.76
-5	(1/64, 1/4)	0.658209	0.75
10	(1/64, 1/4)	0.657875	0.75
15	(1/64, 1/8)	0.657205	0.75

Table 2. Convergence results with one Jacobi pre-smoothing and with conforming corrections

c	(h_K, h_1)	$\overline{\delta}_v$	$\|\|E\|\|$
-5	(1/16, 1/4)	0.464180	0.56
10	(1/16, 1/4)	0.428082	0.53
15	(1/16, 1/8)	NC	-
-5	(1/32, 1/4)	0.551210	0.57
10	(1/32, 1/4)	0.533360	0.55
15	(1/32, 1/8)	0.523353	0.54
-5	(1/64, 1/4)	0.490357	0.56
10	(1/64, 1/4)	0.484002	0.55
15	(1/64, 1/8)	0.481043	0.55

Table 3. Convergence results with one Gauss-Seidel pre-smoothing and with conforming corrections

c	(h_K, h_1)	$\overline{\delta}_v$	$\|\|E\|\|$
-5	(1/16, 1/4)	0.486392	0.56
10	(1/16, 1/4)	0.613320	0.88
15	(1/16, 1/8)	0.476986	0.56
-5	(1/32, 1/4)	0.651377	0.77
10	(1/32, 1/4)	0.902975	1.25
15	(1/32, 1/8)	0.721914	1.00
-5	(1/64, 1/4)	0.726687	0.93
10	(1/64, 1/4)	0.978622	1.37
15	(1/64, 1/8)	0.967870	1.24

Table 4. Convergence results with one Jacobi pre-smoothing and with nonconforming corrections

We first report the results obtained by using Algorithm 3.2 with one (point) Jacobi and Gauss-Seidel pre-smoothing. They are shown in Tables 2 and 3, respectively, where (h_K, h_1) denotes the mesh sizes of the finest and coarsest grids, respectively,

$\overline{\delta}_v$ indicates the average error reduction factor in fifty iterations, and

$$||E|| = \sup_{v \in V_K} \widehat{A}_K(Ev, Ev)/\widehat{A}_K(v, v)$$

is the operator norm in the final iteration. In the cases where there is no convergence (denoted by NC in the tables), the coarsest levels in the multigrid iteration are not fine enough. This agrees with our earlier theory on the nonsymmetric and indefinite problem where the coarsest levels need to be sufficiently fine. Overally, in the case where there is convergence, the Gauss-Seidel smoothing performs better than the Jacobi smoothing, and $\overline{\delta}_v$ and $||E||$ are quite small for both smoothers. When $c = 15$, the coarsest level needs to be finer. This is the case where the convection term becomes 'bigger'.

c	(h_K, h_1)	$\overline{\delta}_v$	$\|\|E\|\|$
-5	(1/16, 1/4)	0.653408	0.76
10	(1/16, 1/4)	0.247183	0.30
15	(1/16, 1/8)	0.236150	0.27
-5	(1/32, 1/4)	0.750099	0.88
10	(1/32, 1/4)	0.299764	0.34
15	(1/32, 1/8)	0.289018	0.41
-5	(1/64, 1/4)	0.795464	0.95
10	(1/64, 1/4)	0.312528	0.39
15	(1/64, 1/8)	0.350614	0.52

Table 5. Convergence results with one Gauss-Seidel pre-smoothing and with nonconforming corrections

c	(h_K, h_1)	$\overline{\delta}_v$	$\|\|E\|\|$
-5	(1/16, 1/4)	0.340037	0.41
10	(1/16, 1/4)	0.337256	0.39
15	(1/16, 1/8)	0.332164	0.38
-5	(1/32, 1/4)	0.366170	0.47
10	(1/32, 1/4)	0.354337	0.41
15	(1/32, 1/8)	0.353695	0.41
-5	(1/64, 1/4)	0.374882	0.57
10	(1/64, 1/4)	0.388482	0.55
15	(1/64, 1/8)	0.342782	0.42

Table 6. Convergence results with one Jacobi pre- and post-smoothing and with nonconforming corrections

For comparison, we also demonstrate the results produced by using the first type of multigrid algorithm; i.e., all the coarse-grid corrections are defined on the nonconforming spaces instead of the conforming spaces. The results with one Jacobi and Gauss-Seidel pre-smoothing are presented in Tables 4 and 5, respectively. Evidently, the results with the Gauss-Seidel smoothing are much better than those with the Jacobi smoothing. As the finest level gets higher (e.g. $h_K = 1/128$, not reported here), we observed that the average error reduction factor approches 0.98 with the Jacobi

c	(h_K, h_1)	$\overline{\delta}_v$	$\|\|E\|\|$
-5	(1/16, 1/4)	0.165309	0.18
10	(1/16, 1/4)	0.130279	0.16
15	(1/16, 1/8)	0.122040	0.15
-5	(1/32, 1/4)	0.181391	0.23
10	(1/32, 1/4)	0.186105	0.23
15	(1/32, 1/8)	0.184373	0.24
-5	(1/64, 1/4)	0.208130	0.26
10	(1/64, 1/4)	0.219439	0.28
15	(1/64, 1/8)	0.215479	0.28

Table 7. Convergence results with Gauss-Seidel pre- and post-smoothing and with nonconforming corrections

c	(h_K, h_1)	$\overline{\delta}_w$	$\|\|E\|\|$
-5	(1/16, 1/4)	0.330665	0.39
10	(1/16, 1/4)	0.330144	0.39
15	(1/16, 1/8)	0.326979	0.38
-5	(1/32, 1/4)	0.345017	0.39
10	(1/32, 1/4)	0.344432	0.39
15	(1/32, 1/8)	0.343179	0.39
-5	(1/64, 1/4)	0.325344	0.38
10	(1/64, 1/4)	0.325438	0.38
15	(1/64, 1/8)	0.325509	0.38

Table 8. Convergence results of the $\mathcal{W}$-cycle with Jacobi pre- and post-smoothing and with nonconforming corrections

c	(h_K, h_1)	$\overline{\delta}_w$	$\|\|E\|\|$
-5	(1/16, 1/4)	0.133058	0.16
10	(1/16, 1/4)	0.137478	0.18
15	(1/16, 1/8)	0.146508	0.14
-5	(1/32, 1/4)	0.113530	0.15
10	(1/32, 1/4)	0.112054	0.15
15	(1/32, 1/8)	0.124663	0.18
-5	(1/64, 1/4)	0.116710	0.15
10	(1/64, 1/4)	0.117474	0.15
15	(1/64, 1/8)	0.123881	0.15

Table 9. Convergence results of the $\mathcal{W}$-cycle with Gauss-Seidel pre- and post-smoothing and with nonconforming corrections

smoothing. For this reason, we experimented with the first type of multigrid algorithm with one Jacobi and Gauss-Seidel both pre- and post-smoothing. The results are displayed in Tables 6 and 7, respectively. It appears that this type of algorithm needs at least two smoothing steps to have good results with the Jacobi smoothing. Finally, while there is no theoretical analysis for the $\mathcal{W}$-cycle algorithm for the nonsymmetric problem, we point out that the results generated by the $\mathcal{W}$-cycle algorithm are slightly better than those yielded by the $\mathcal{V}$-cycle algorithm, as shown in Tables 8 and 9.

REFERENCES

[1] T. Arbogast and Zhangxin Chen, On the implementation of mixed methods as nonconforming methods for second order elliptic problems, *Math. Comp.*, **64** (1995), to appear.

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

[3] J. Bramble, On the development of multigrid methods and their analysis, *Mathematics of Computation* 1943–1993, A Half-Century of Computational Mathematics, W. Gautschi, ed., Proceedings of Symposia in Applied Mathematics, AMS, Providence, **48** (1994), 5–19.

[4] J. Bramble, R. Ewing, J. Pasciak, and J. Shen, The analysis of multigrid algorithms for cell centered finite difference methods, Preprint, 1994.

[5] J. Bramble, Do 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

[6] J. Bramble and J. Pasciak, New convergence estimates for multigrid algorithms, *Math. Comp.*, **49** (1987), 311–329.

[7] J. Bramble, J. Pasciak, and J. Xu, The analysis of multigrid algorithms with non-nested spaces or non-inherited quadratic forms, *Math. Comp.*, **56** (1991), 1–34.

[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] S. Brenner, Multigrid methods for nonconforming finite elements, *Proceedings of Fourth Copper Mountain Conference on Multigrid Methods*, J. Mandel, et. als., eds., SIAM, Philadelphia, 1989, 54–65.

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

[12] Zhangxin Chen, Equivalence between and multigrid algorithms for mixed and nonconforming methods for second order elliptic problems, *IMA Preprint* #1218, 1994.

[13] Zhangxin Chen and J. Douglas, Jr., Approximation of coefficients in hybrid and mixed methods for nonlinear parabolic problems, *Mat. Aplic. Comp.*, **10** (1991), 137–160.

[14] Zhangxin Chen, R.E. Ewing, and R. Lazarov, Domain decomposition algorithms for mixed methods for second order elliptic problems, *Math. Comp.*, **65** (1996), to appear.

[15] Zhangxin Chen, R.E. Ewing, Y. Kuznetsov, R. Lazarov, and S. Maliassov, Multilevel preconditioners for mixed methods for second order elliptic problems, *ISC*-94-10-*MATH Technical Report, Texas A&M University*, submitted for publication.

[16] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, New York, Oxford, 1978.

[17] R.E. Ewing, Yu. Kuznetsov, R. Lazarov, and S. Maliassov, Preconditioning of nonconforming finite element approximations of second order elliptic problems, In *Proceedings of the Third International Conf. on Advances in Numer. Meth. and Appl.*, I. Dimov, Bl. Sendov, and P. Vassilevski, eds., World Scientific, London & Hong Kong (1994), 101–110.

[18] R.E. Ewing, Yu. Kuznetsov, R. Lazarov, and S. Maliassov, Substructuring preconditioning for finite element approximations of second order elliptic problems. I. Nonconforming linear elements for the Poisson equation in parallelepiped, ISC Preprint #2, TAMU, 1994.

[19] C. Lee, A nonconforming multigrid method using conforming subspaces, In *Proceedings of the Sixth Copper Mountain Conference on Multigrid Methods*, N. Melson et al., eds., NASA Conference Publication 3224, Part 1 (1993), 317–330.

[20] J. Nitsche, Lineare Spline-Funktionen und die Methode von Ritz für elliptische Randwertprobleme, *Arch. Rational Mech. Anal.*, **36** (1970), 348–355

[21] P. Oswald, On a hierarchical basis multilevel method with nonconforming P1 elements, *Numer. Math.*, **62** (1992), 189–212.

[22] P. Raviart and J. Thomas, A mixed finite element method for second order elliptic problems, In *Mathematical aspects of the FEM, Lecture Notes in Mathematics*, **606**, Springer-Verlag, Berlin & New York (1977), 292–315.

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

[24] A. Schatz and J. Wang, Some new error estimates in finite element methods, Preprint, 1993.