MULTIGRID AND MULTILEVEL METHODS FOR NONCONFORMING Q_1 ELEMENTS

ZHANGXIN CHEN AND PETER OSWALD

ABSTRACT. In this paper we study theoretical properties of multigrid algorithms and multilevel preconditioners for discretizations of second-order elliptic problems using nonconforming *rotated* Q_1 finite elements in two space dimensions. In particular, for the case of square partitions and the Laplacian we derive properties of the associated intergrid transfer operators which allow us to prove convergence of the $\mathcal{W}$-cycle with any number of smoothing steps and close-to-optimal condition number estimates for $\mathcal{V}$-cycle preconditioners. This is in contrast to most of the other nonconforming finite element discretizations where only results for $\mathcal{W}$-cycles with a sufficiently large number of smoothing steps and variable $\mathcal{V}$-cycle multigrid preconditioners are available. Some numerical tests, including also a comparison with a preconditioner obtained by switching from the nonconforming *rotated* Q_1 discretization to a discretization by conforming bilinear elements on the same partition, illustrate the theory.

1. Introduction

In recent years there have been analyses and applications of the nonconforming *rotated* (NR) Q_1 finite elements for the numerical solution of partial differential problems. These nonconforming elements were first proposed and analyzed in [24] for numerically solving the Stokes problem; they provide the simplest example of discretely divergence-free nonconforming elements on quadrilaterals. More results on these Stokes elements can be found in [26]. There also exist n-dimensional counterparts of these elements, with analogous properties [25]. Then the NR Q_1 elements were used to simulate the deformation of martensitic crystals with microstructure [18] due to their simplicity. Conforming finite element methods can be used to approximate the microstructure with layers which are oriented with respect to meshes, while nonconforming finite element methods allow the microstructure to be approximated on meshes which are not aligned with the microstructure (see, e.g., [18] for the references).

Independently, the NR Q_1 elements have been derived within the framework of mixed finite element methods [11, 1]. It has been shown that the nonconforming method using these elements is equivalent to the mixed method utilizing the lowest-order Raviart-Thomas mixed elements on rectangles (respectively, rectangular parallelepipeds) [25]. Based on this equivalence theory, both the NR Q_1 and the

Raviart-Thomas mixed methods have been applied to model semiconductor devices [11]; they have been effectively employed to compute the electric potential equation with a doping profile which has a sharp junction.

Error estimates of the NR Q_1 elements can be derived by the classical finite element analysis [24, 17]. They can also be obtained from the known results on the mixed method based on the equivalence between these two methods [1]. It has been shown that the so-called "nonparametric" *rotated* Q_1 elements produce optimal-order error estimates. As a special case of the nonparametric families, the optimal-order error estimates can be obtained for partitions into rectangles (respectively, rectangular parallelepipeds) oriented along the coordinate axes. Finally, superconvergence results have been obtained in [1, 17].

Unlike the simplest triangular nonconforming elements, i.e., the nonconforming P_1 elements, the NR Q_1 elements do not have any reasonable conforming subspace. Consequently, there are differences between these two types of nonconforming elements. The NR Q_1 elements can be defined on quadrilaterals with degrees of freedom given by the values at the midpoints of edges of the quadrilaterals, or by the averages over the edges of the quadrilaterals. While these two versions lead to the same definition for the nonconforming P_1 elements, they produce different results in terms of implementation for the NR Q_1 elements. With the second version of the NR Q_1 elements, we are able to prove all the theoretical results for the multigrid algorithms and multilevel additive and multiplicative Schwarz methods considered in this paper. However, we are unable to obtain these results with their first version. In particular, as numerical tests in [23] indicate, the energy norm of the iterates of the usual intergrid transfer operators, which enters both upper and lower bounds for the condition number of preconditioned systems, deteriorates with the number of grid levels for the first version. But it is bounded independently of the number of grid levels for the second version, as shown here for square partitions.

Since the nonconforming P_1 finite element space contains the conforming P_1 elements (with respect to the same triangulation), the convergence of the standard $\mathcal{V}$-cycle algorithm for the nonconforming P_1 elements can be shown when the coarse-grid correction steps of this algorithm are established on the conforming P_1 spaces [29, 19, 12]. Such an approach to establishing $\mathcal{V}$-cycle results fails for the NR Q_1 elements. On the other hand, within the context of the nonconforming methods, i.e., when the coarse-grid correction steps are defined on the nonconforming P_1 spaces themselves, the convergence of the $\mathcal{V}$-cycle algorithm has not been shown, and the $\mathcal{W}$-cycle algorithm has been proven to converge only under the assumption that the number of smoothing steps is sufficiently large [7, 8, 3, 4, 27, 1, 12, 15].

Multigrid algorithms for the NR Q_1 discretizations of a second-order elliptic boundary value problem were first developed and analyzed in [1], and further discussed in [12] and [9]. The second version of these elements was used in [1] and [12], while their first version was exploited in [9]. Moreover, the analysis in [9] was given for elliptic boundary value problems which are not required to have full elliptic regularity. However, in all these three papers, only the $\mathcal{W}$-cycle algorithm with a sufficiently large number of smoothing steps was shown to converge using the standard proof of convergence of multigrid algorithms for conforming finite element methods [2]. We finally mention that the study of the NR Q_1 elements in the context of domain decomposition methods has been given in [13, 14].

This paper should be viewed mainly as a contribution to the theory of multigrid methods for nonconforming finite element discretizations. In Section 2, we derive

some new properties of intergrid transfer operators associated with the second version of the NR Q_1 elements. The crucial estimates (2.13) (second inequality) and (2.15) are shown for the Laplace operator with Dirichlet boundary conditions on the unit square in $\mathbb{R}^2$, and for sequences of uniform square partitions for (2.15). Consequently, most of the new results on multigrid methods and multilevel preconditioners proved in the subsequent sections are restricted to this model case. Throughout the paper, we make some comments on extending the results to more general elliptic problems, domains, and partition types.

In Section 3, we show optimal, level-independent convergence rates for the $\mathcal{W}$-cycle algorithm to hold with any number of smoothing steps. The NR Q_1 elements have so far been the first type of nonconforming elements which are shown to possess this feature. The question of establishing level-independent convergence rates for the standard $\mathcal{V}$-cycle still remains open.

Multilevel preconditioners of hierarchical basis and BPX type for the NR Q_1 elements are studied in Section 4. Following [23], we develop a convergence theory for the multilevel additive Schwarz methods and their related multiplicative $\mathcal{V}$-cycle algorithms. A key ingredient in the analysis is to control the energy norm growth of the iterated coarse-to-fine grid operators, which enters both upper and lower bounds for the condition number of preconditioned systems. So far, the energy norm of the iterated intergrid transfer operators has been shown to be bounded independently of grid levels solely for the nonconforming P_1 elements [20]. In this paper, we prove this property for the NR Q_1 elements (see Lemma 2.4). As a consequence, we obtain a suboptimality result for the multilevel preconditioners of hierarchical basis and BPX type for the NR Q_1 elements.

In Section 5, we apply ideas of [22] and study the problem of switching the NR Q_1 discretization to a spectrally equivalent discretization for which optimal preconditioners are already available. For square partitions, the conforming bilinear finite element space is a suitable candidate. The switching approach leads to optimal preconditioning results for the NR Q_1 elements.

Thanks to the equivalence between the *rotated* Q_1 nonconforming method and the lowest-order Raviart-Thomas mixed rectangular method, all the results derived here carry over directly to the latter method [1, 12]. An extension to the corresponding discretely divergence-free NR Q_1 Stokes discretization is not straightforward since the standard intergrid transfer operators for the scalar case do not preserve the solenoidality constraint (see [26] for intergrid transfer operators for the Stokes case and related multigrid results).

Finally, in Section 6 we present some numerical results on convergence rates and condition numbers which confirm the theoretical findings.

2. Preliminary results

Let $H^s(\Omega)$ and $L^2(\Omega) = H^0(\Omega)$ be the usual Sobolev spaces with the norm

$$||v||_s = \left(\int_\Omega \sum_{|\alpha| \leq s} |D^\alpha v|^2 \, \mathrm{d}x \right)^{1/2},$$

where s is a nonnegative integer, and Ω is a two-dimensional domain. Also, let $(\cdot, \cdot)$ denote the $L^2(\Omega)$ or $(L^2(\Omega))^2$ inner product, as appropriate. The $L^2(\Omega)$ norm is

indicated by $|| \cdot ||$. Finally,

$$H_0^1(\Omega) = \{v \in H^1(\Omega) : v|_\Gamma = 0\},$$

where $\Gamma = \partial\Omega$.

From now on, let Ω be the unit square $(0,1)^2$ (extensions will be mentioned separately). Let h_1 and $\mathcal{E}_{h_1} = \mathcal{E}_1$ be given, where $\mathcal{E}_{h_1}$ is a partition of Ω into uniform squares with length h_1 and oriented along the coordinate axes (in the simplest dyadic case, one would take $h_1 = 1/2$). For each integer $2 \le k \le K$, 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 squares in $\mathcal{E}_{k-1}$, and let $\mathcal{E}_h = \mathcal{E}_K$ be the finest grid. Also, let $\partial\mathcal{E}_k$ be the set of all interior edges in $\mathcal{E}_k$. In this and the following sections, we replace subscript h_k simply by subscript k.

For each k, we introduce the NR Q_1 space

$$V_k = \Big\{ v \in L^2(\Omega) : v|_E = a_E^1 + a_E^2 x + a_E^3 y + a_E^4(x^2 - y^2),\ a_E^i \in \mathbb{R},\ \forall E \in \mathcal{E}_k;$$

$$\text{if } E_1 \text{ and } E_2 \text{ share an edge } e, \text{ then } \int_e \xi|_{\partial E_1}\, \mathrm{d}s = \int_e \xi|_{\partial E_2}\, \mathrm{d}s;$$

$$\text{and } \textstyle\int_{\partial E \cap \Gamma} \xi|_\Gamma\, \mathrm{d}s = 0 \Big\}.$$

Note that $V_k \not\subset H_0^1(\Omega)$ and $V_{k-1} \not\subset V_k$, $k \ge 2$. We introduce the space

$$\hat{V}_k = \sum_{l=1}^{k} V_l \supset V_k,$$

the discrete energy scalar product on $\hat{V}_k \oplus H_0^1(\Omega)$ by

$$(v, w)_{\mathcal{E},k} = \sum_{E \in \mathcal{E}_k} (\nabla v, \nabla w)_E, \qquad v,\, w \in \hat{V}_k \oplus H_0^1(\Omega),$$

and the discrete norm on $\hat{V}_k \oplus H_0^1(\Omega)$ by

$$||v||_{\mathcal{E},k} = \sqrt{(v,v)_{\mathcal{E},k}}, \qquad v \in \hat{V}_k \oplus H_0^1(\Omega).$$

We introduce two sets of intergrid transfer operators $I_k : V_{k-1} \to V_k$ and $P_{k-1} : V_k \to V_{k-1}$ as follows. Following [1, 12], if $v \in V_{k-1}$ and e is an edge of a square in $\mathcal{E}_k$, then $I_k v \in V_k$ is defined by

$$\frac{1}{|e|} \int_e I_k v\, \mathrm{d}s = \begin{cases} 0 & \text{if } e \subset \partial\Omega, \\[2mm] \dfrac{1}{|e|} \displaystyle\int_e v\, \mathrm{d}s & \text{if } e \text{ is interior to some } E \in \mathcal{E}_{k-1}, \\[3mm] \dfrac{1}{2|e|} \displaystyle\int_e (v|_{E_1} + v|_{E_2})\, \mathrm{d}s & \text{if } e \subset \partial E_1 \cap \partial E_2 \\ & \qquad \text{for some } E_1, E_2 \in \mathcal{E}_{k-1}. \end{cases}$$

If $v \in V_k$ and e is an edge of an element in $\partial\mathcal{E}_{k-1}$, then $P_{k-1}v \in V_{k-1}$ is given by

$$\frac{1}{|e|} \int_e P_{k-1}v\, \mathrm{d}s = \frac{1}{2}\left\{ \frac{1}{|e_1|} \int_{e_1} v\, \mathrm{d}s + \frac{1}{|e_2|} \int_{e_2} v\, \mathrm{d}s \right\},$$

where e_1 and e_2 in $\partial\mathcal{E}_k$ form the edge $e \in \partial\mathcal{E}_{k-1}$. Note that the definition of P_{k-1} automatically preserves the zero average values on boundary edges. Also, it can be

seen that

(2.1) $$P_{k-1}I_k v = v, \qquad v \in V_{k-1},\, k \geq 1.$$

That is, $P_{k-1}I_k$ is the identity operator Id_{k-1} on V_{k-1}. This relation is not satisfied when the NR Q_1 elements are defined with degrees of freedom given by the values at the midpoints of edges of elements.

For future use, note that I_k (as well as P_{k-1}) can be extended to the larger spaces $\hat{V}_k$ in a natural way. For example, the definition of $\hat{I}_k v \in V_k$ by

$$\frac{1}{|e|}\int_e \hat{I}_k v \, \mathrm{d}s = \begin{cases} 0 & \text{if } e \subset \partial\Omega, \\[2mm] \dfrac{1}{2|e|}\displaystyle\int_e (v|_{E_1} + v|_{E_2})\,\mathrm{d}s & \text{if } e = \partial E_1 \cap \partial E_2 \\[2mm] & \qquad\text{for some } E_1, E_2 \in \mathcal{E}_k, \end{cases}$$

is meaningful for any $v \in \hat{V}_k$, and satisfies $\hat{I}_k|_{V_{k-1}} = I_k$ as well as $\hat{I}_k|_{V_k} = \mathrm{Id}_k$.

We also define the iterates of I_k and P_{k-1} by

$$R_k^K = I_K \cdots I_{k+1} : V_k \to V_K,$$
$$Q_k^K = P_k \cdots P_{K-1} : V_K \to V_k.$$

Finally, the discrete energy scalar product on the space $\hat{V}_K$ is defined by restriction:

$$(v, w)_{\mathcal{E}} = (v, w)_{\mathcal{E},K}, \quad v,\, w \in \hat{V}_K.$$

Obviously, we have the inverse inequality

(2.2) $$\|v\|_{\mathcal{E}} \leq C 2^k \|v\|, \quad v \in \hat{V}_k,\, 1 \leq k \leq K,$$

(here and later, by C, c,... we denote generic positive constants which are independent of k, K, and the functions involved).

In this section we collect some basic properties of the intergrid transfer operators P_{k-1} (respectively, I_k) and their iterates Q_k^K (respectively, R_k^K). The crucial results are the boundedness of the operators I_k with constant $\sqrt{2}$ (Lemma 2.3) and the uniform boundedness of the operators R_k^K with respect to the discrete energy norm $\|\cdot\|_{\mathcal{E}}$ (Lemma 2.4).

Lemma 2.1. *It holds that P_{k-1} ($2 \leq k \leq K$) is an orthogonal projection with respect to the energy scalar product, i.e., for any $v \in V_k$,*

(2.3)
$$(v - P_{k-1}v, w)_{\mathcal{E}} = 0, \qquad \forall w \in V_{k-1},$$
$$\|v\|_{\mathcal{E}}^2 = \|v - P_{k-1}v\|_{\mathcal{E}}^2 + \|P_{k-1}v\|_{\mathcal{E}}^2.$$

Moreover, there are constants C and c, independent of v, such that the difference $\hat{v} = v - P_{k-1}v \in \hat{V}_k$ satisfies

(2.4) $$c 2^k \|\hat{v}\| \leq \|\hat{v}\|_{\mathcal{E}} \leq C 2^k \|\hat{v}\|.$$

Proof. For any $E \in \mathcal{E}_{k-1}$ with the four subsquares $E_i \in \mathcal{E}_k$ ($i = 1, \ldots, 4$, see Figure 1), an application of Green's formula implies that

(2.5)
$$\begin{aligned}(\nabla[v - P_{k-1}v], \nabla w)_E &= \textstyle\sum_{i=1}^{4}(\nabla[v - P_{k-1}v], \nabla w)_{E_i} \\ &= \textstyle\sum_{i=1}^{4}\sum_{j=1}^{4} \frac{\partial w}{\partial \nu_{E_i}^j}\big|_{e_{E_i}^j} \int_{e_{E_i}^j}(v - P_{k-1}v)|_{E_i}\,\mathrm{d}s,\end{aligned}$$

where $e_{E_i}^j$ are the four edges of E_i with the outer unit normals $\nu_{E_i}^j$, $i = 1, \ldots, 4$. Note that in (2.5) the line integrals over edges interior to $E \in \mathcal{E}_{k-1}$ cancel by
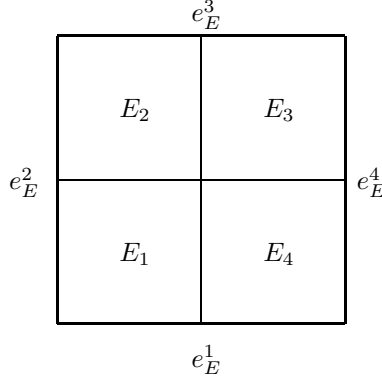
FIGURE 1. Edges and subsquares of $E \in \mathcal{E}_{k-1}$.

continuity of $P_{k-1}v$ in the interior of E. Also, if $e^j_{E_i}$ and $e^{\hat{j}}_{E_{\hat{i}}}$ form an edge of E, it follows by the definition of P_{k-1} that

$$\int_{e^j_{E_i}} (v - P_{k-1}v)|_{E_i} \, \mathrm{d}s + \int_{e^{\hat{j}}_{E_{\hat{i}}}} (v - P_{k-1}v)|_{E_{\hat{i}}} \, \mathrm{d}s = 0,$$

and that

$$\frac{\partial w}{\partial \nu^j_{E_i}} \big|_{e^j_{E_i}} = \frac{\partial w}{\partial \nu^{\hat{j}}_{E_{\hat{i}}}} \big|_{e^{\hat{j}}_{E_{\hat{i}}}}$$

is constant. Then, by (2.5), we see that

$$(\nabla[v - P_{k-1}v], \nabla w)_E = 0.$$

Now, sum over all $E \in \mathcal{E}_{k-1}$ to derive the orthogonality relations in (2.3).

The upper estimate in (2.4) directly follows from (2.2). The lower bound can be easily obtained from a direct calculation of the energy norms of $v - P_{K-1}v$ on all $E \in \mathcal{E}_{k-1}$. This completes the proof. $\square$

Before we start with the investigation of the prolongations I_k, it will be useful to collect some formulas. For $E \in \mathcal{E}_{k-1}$ and any $v \in V_{k-1}$, define

$$b^i_E = \frac{1}{|e^i_E|} \int_{e^i_E} v \, \mathrm{d}s,$$

(see Figure 1 for the notation), and set

$$s_E = b^1_E + b^2_E + b^3_E + b^4_E, \qquad \triangle^1_E = b^3_E - b^1_E,$$
$$\theta^0_E = b^1_E + b^3_E - b^2_E - b^4_E, \qquad \triangle^2_E = b^4_E - b^2_E.$$

Then, with the subscript E omitted, we have the next lemma.

Lemma 2.2. *It holds that*

(2.6)
$$\|v\|^2_{L^2(E)} = h^2_{k-1} \left(\frac{1}{16} s^2 + \frac{1}{12} \{ (\triangle^1)^2 + (\triangle^2)^2 \} + \frac{1}{40} (\theta^0)^2 \right),$$
$$\|\nabla v\|^2_{L^2(E)} = (\triangle^1)^2 + (\triangle^2)^2 + \frac{3}{2} (\theta^0)^2,$$
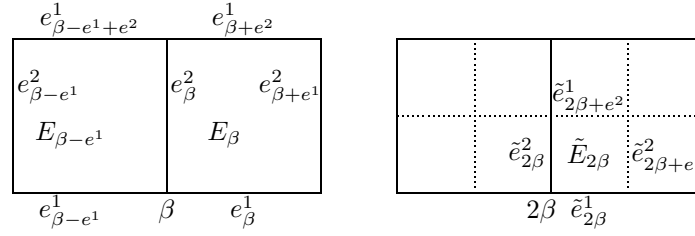
FIGURE 2. An illustration for Lemma 2.3.

and

$$(2.7) \qquad \begin{aligned} \tfrac{h_{k-1}^2}{10} \left\{ (b^1)^2 + (b^2)^2 + (b^3)^2 + (b^4)^2 \right\} &\le ||v||_{L^2(E)}^2 \\ &\le \tfrac{h_{k-1}^2}{4} \left\{ (b^1)^2 + (b^2)^2 + (b^3)^2 + (b^4)^2 \right\}. \end{aligned}$$

Proof. Using the affine invariance of the local interpolation problem connecting v with its edge averages b^i, it suffices to prove (2.6) and (2.7) for the master square $E = (-1,1)^2$. A straightforward calculation gives

$$(2.8) \qquad v = v(x,y) = \frac{1}{4}s + \frac{\triangle^2}{2}x + \frac{\triangle^1}{2}y - \frac{3}{8}\theta^0(x^2 - y^2).$$

Now direct integration yields the desired results in (2.6). Also, (2.7) follows from the first equation of (2.6) by computing the eigenvalues of the symmetric 4×4 matrix $\mathcal{T}^t D \mathcal{T}$, where $D = \mathrm{diag}(1/16, 1/12, 1/12, 1/40)$, $\mathcal{T}$ stands for the transformation matrix from the vector (b^1, b^3, b^2, b^4) to $(s, \triangle^1, \triangle^2, \theta^0)$, and $\mathcal{T}^t$ is the transpose of $\mathcal{T}$. These eigenvalues are $1/10$, $1/6$, $1/6$, and $1/4$, which implies (2.7). $\square$

Lemma 2.2 is the basis for computing all the discrete energy and L^2 norms needed in the sequel. The formula (2.8) valid for the master square can be used to derive explicit expressions for the edge averages of $I_k v$ and $I_k v - v$. Toward this end, we first compute the corresponding values for the master square, and then use the invariance of the local interpolation problem for v under affine transformations (for the square triangulations under consideration, these transformations are just dilation and translation) to return to the notation on each $E \in \mathcal{E}_{k-1}$.

Note that for the square partitions under consideration, vertices, subsquares, and horizontal/vertical edges can be labeled by multi-indices $\beta \in \mathbb{Z}^2$ in a natural way. The origin and the square attached to it are labeled by $\beta = (0,0)$, for all k. See Figure 2 for the further conventions. The left picture shows two squares $E_{\beta-e^1}$, E_β from $\mathcal{E}_{k-1}$, and the right the same two squares (together with their subsquares) as part of $\mathcal{E}_k$. We use the notation $e^1 = (1,0)$, and $e^2 = (0,1)$ for the unit vectors. Horizontal and vertical edges are distinguished by the superscripts 1 and 2, respectively; e.g., e_β^1 in the left picture is the horizontal edge in $\partial\mathcal{E}_{k-1}$ emanating from the vertex with index β and belonging to the element with index E_β of $\mathcal{E}_{k-1}$. The same vertex is labeled by 2β if considered as vertex in $\mathcal{E}_k$, and so on.

Given an arbitrary $v \in V_{k-1}$, let b_β^1 and b_β^2 denote its averages over the horizontal and vertical edges e_β^1 and e_β^2 in $\partial\mathcal{E}_{k-1}$, respectively. The corresponding quantities

for $I_k v \in V_k$ are indicated by a_α^j, $j = 1, 2$. Now, introduce the auxiliary quantities

$$\hat{\theta}_\beta^1 = b_\beta^1 + b_{\beta-e^1}^1 - b_{\beta+e^2}^1 - b_{\beta+e^2-e^1}^1,$$

$$\hat{\theta}_\beta^2 = b_\beta^2 + b_{\beta-e^2}^2 - b_{\beta+e^1}^2 - b_{\beta+e^1-e^2}^2.$$

With this notation at hand, it follows from the definition of I_k that the edge averages of $I_k v$ can be written as follows:

(2.9)
$$a_{2\beta}^1 = b_\beta^1 + \tfrac{1}{8}\hat{\theta}_\beta^2,$$
$$a_{2\beta+e^1}^1 = b_\beta^1 - \tfrac{1}{8}\hat{\theta}_\beta^2,$$
$$a_{2\beta+e^2}^1 = \tfrac{5}{8}b_\beta^2 + \tfrac{1}{8}b_{\beta+e^1}^2 + \tfrac{1}{8}b_\beta^1 + \tfrac{1}{8}b_{\beta+e^2}^1,$$
$$a_{2\beta+e^2+e^1}^1 = \tfrac{5}{8}b_{\beta+e^1}^2 + \tfrac{1}{8}b_\beta^2 + \tfrac{1}{8}b_\beta^1 + \tfrac{1}{8}b_{\beta+e^2}^1,$$

and

(2.10)
$$a_{2\beta}^2 = b_\beta^2 + \tfrac{1}{8}\hat{\theta}_\beta^1,$$
$$a_{2\beta+e^2}^2 = b_\beta^2 - \tfrac{1}{8}\hat{\theta}_\beta^1,$$
$$a_{2\beta+e^1}^2 = \tfrac{5}{8}b_\beta^1 + \tfrac{1}{8}b_{\beta+e^2}^1 + \tfrac{1}{8}b_\beta^2 + \tfrac{1}{8}b_{\beta+e^1}^2,$$
$$a_{2\beta+e^2+e^1}^2 = \tfrac{5}{8}b_{\beta+e^2}^1 + \tfrac{1}{8}b_\beta^1 + \tfrac{1}{8}b_\beta^2 + \tfrac{1}{8}b_{\beta+e^1}^2.$$

These formulas are valid for interior edges. Whenever the edge average a_α^j is associated with a boundary edge of $\mathcal{E}_k$, this value has to be replaced by zero. We give the elementary argument for the first and third formulas in (2.9); the others follow by symmetry arguments. Let us start with the edge $\tilde{e}_{2\beta+e^2}^1$ in $\partial\mathcal{E}_k$. Since it is interior to E_β (see Figure 2), we have

$$a_{2\beta+e^2}^1 = \frac{1}{|\tilde{e}_{2\beta+e^2}^1|} \int_{\tilde{e}_{2\beta+e^2}^1} v\, \mathrm{d}s \,.$$

Using the dilation invariance, this integral can be computed by transferring E_β to the master square and using (2.8). This leads us to integrating the expression in (2.8) along the path $-1 \leq x \leq 0$, $y = 0$, and substituting the values for the parameters obtained from the corresponding edge averages of v:

$$s = b_\beta^1 + b_{\beta+e^2}^1 + b_\beta^2 + b_{\beta+e^1}^2, \quad \Delta^1 = b_{\beta+e^2}^1 - b_\beta^1,$$
$$\Delta^2 = b_{\beta+e^1}^2 - b_\beta^2, \quad \theta^0 = b_\beta^1 + b_{\beta+e^2}^1 - b_\beta^2 - b_{\beta+e^1}^2 \,.$$

As a result, we have

$$a_{2\beta+e^2}^1 = \frac{s}{4} - \frac{\Delta^2}{4} - \frac{\theta^0}{8} = \frac{1}{8}(5b_\beta^2 + b_{\beta+e^1}^2 + b_\beta^1 + b_{\beta+e^2}^1) \,,$$

which is the third formula in (2.9).

Analogously, we integrate in (2.8) along $-1 \leq x \leq 0$, $y = -1$, and obtain

$$\frac{1}{|\tilde{e}_{2\beta}^1|} \int_{\tilde{e}_{2\beta}^1} v|_{E_\beta}\, \mathrm{d}s = \frac{s}{4} - \frac{\Delta^1}{2} - \frac{\Delta^2}{4} + \frac{\theta^0}{4}$$
$$= b_\beta^1 + \tfrac{1}{4}(b_\beta^2 - b_{\beta+e^1}^2) \,.$$

To obtain the average of $v|_{E_{\beta-e^2}}$ along the same edge, we integrate in (2.8) along the path $-1 \leq x \leq 0, y = 1$, and apply an index shift by $-e^2$. This yields that

$$\frac{1}{|\tilde{e}_{2\beta}^1|} \int_{\tilde{e}_{2\beta}^1} v|_{E_{\beta-e^2}}\, \mathrm{d}s = b_\beta^1 + \tfrac{1}{4}(b_{\beta-e^2}^2 - b_{\beta+e^1-e^2}^2) \,.$$

Combining the last two formulas results in the first formula in (2.9) since according to the definition of $I_k v$

$$a_{2\beta}^1 = \frac{1}{2|\tilde{e}_{2\beta}^1|} \int_{\tilde{e}_{2\beta}^1} (v|_{E_\beta} + v|_{E_{\beta-e^2}}) \, ds \ .$$

To obtain the main results of this section which concern the behavior of I_k and its iterates with respect to the energy norm, we need to deal with the difference $I_k v - v$ which is an element of $\hat{V}_k$ for any $v \in V_k$. This is particularly useful since we have

(2.11) $$||I_k v||_{\mathcal{E}}^2 - ||v||_{\mathcal{E}}^2 = ||I_k v - v||_{\mathcal{E}}^2, \qquad v \in V_{k-1}.$$

The relation (2.11) follows from Lemma 2.1 (replace there v by $I_k v$ and w by v and use $P_{k-1} I_k = \mathrm{Id}_{k-1}$). It shows that I_k is expanding in the energy norm, and that the energy norm growth is intimately connected with the energy norm of $I_k v - v$.

The edge averages related to $I_k v - v$ have simple expressions if we introduce the following notation. Set

$$\theta_\beta^1 = b_{\beta+e^2}^1 - b_\beta^1 + b_{\beta-e^1}^1 - b_{\beta+e^2-e^1}^1,$$
$$\theta_\beta^2 = b_{\beta+e^1}^2 - b_\beta^2 + b_{\beta-e^2}^2 - b_{\beta+e^1-e^2}^2,$$

if e_β^1 resp. e_β^2 are interior edges in $\partial\mathcal{E}_{k-1}$. For boundary edges, they need to be modified. For example, if e_β^1 is a boundary edge in $\partial\mathcal{E}_{k-1}$, we define

$$\theta_\beta^2 = 2(b_{\beta+e^1}^2 - b_\beta^2),$$

analogously for vertical boundary edges.

Denote temporarily $w = I_k v - v$. We have essentially two cases. If an edge $e \in \partial\mathcal{E}_k$ belongs to the interior of some square in $\mathcal{E}_{k-1}$, then the edge averages of w (taken from the restrictions to either of the two squares in $\mathcal{E}_k$ attached to e) vanish by definition of I_k. What remains are edges e that belong to a (boundary or interior) edge of the partition $\mathcal{E}_{k-1}$. We give the result for the case that e coincides with $\tilde{e}_{2\beta}^1$ or $\tilde{e}_{2\beta+e^1}^1$, i.e., belongs to the edge $e_\beta^1 \in \partial\mathcal{E}_{k-1}$ (see Figure 2 for the notation):

(2.12)
$$\frac{1}{|\tilde{e}_{2\beta}^1|} \int_{\tilde{e}_{2\beta}^1} w|_{\tilde{E}_{2\beta}} ds = \frac{1}{|\tilde{e}_{2\beta+e^1}^1|} \int_{\tilde{e}_{2\beta+e^1}^1} w|_{\tilde{E}_{2\beta+e^1-e^2}} ds = \tfrac{1}{8}\theta_\beta^2,$$
$$\frac{1}{|\tilde{e}_{2\beta}^1|} \int_{\tilde{e}_{2\beta}^1} w|_{\tilde{E}_{2\beta-e^2}} ds = \frac{1}{|\tilde{e}_{2\beta+e^1}^1|} \int_{\tilde{e}_{2\beta+e^1}^1} w|_{\tilde{E}_{2\beta+e^1}} ds = -\tfrac{1}{8}\theta_\beta^2.$$

The averages of $w = I_k v - v$ on other edges (including those on the boundary of Ω) are given similarly. The derivation of (2.12) is left upon the reader (just recall the above calculations which led to the proof of the first inequality in (2.9)).

From (2.9)–(2.12) and Lemma 2.2, we immediately have the next lemma. Below the notation $\approx$ stands for two-sided inequalities with constants independent of k.

Lemma 2.3. *It holds that*

(2.13)
$$||\hat{I}_k v|| \leq \sqrt{\tfrac{5}{2}}||v||, \qquad \forall v \in \hat{V}_k,$$
$$||I_k v||_{\mathcal{E}} \leq \sqrt{2}||v||_{\mathcal{E}}, \qquad \forall v \in V_{k-1},$$

and

(2.14) $$2^k ||I_k v - v|| \approx ||I_k v - v||_{\mathcal{E}} \approx \left(\sum_\beta \{(\theta_\beta^1)^2 + (\theta_\beta^2)^2\}\right)^{1/2}, \qquad \forall v \in V_{k-1}.$$

Proof. Relation (2.14) is obvious from (2.12) and Lemma 2.2 (note that, for any element E in $\mathcal{E}_k$, the values Δ_E^1 resp. Δ_E^2 coincide with $\pm\frac{1}{8}\theta_\beta^2$ resp. $\pm\frac{1}{8}\theta_{\beta'}^2$ for the corresponding β, β'). The first estimate in (2.13) follows from a purely local argument and holds for any function v which is piecewise (on each square $E \in \mathcal{E}_k$) in the *rotated* Q_1 space span$\{1, x, y, x^2 - y^2\}$. Indeed, if the two edge averages corresponding to an edge $e \in \partial\mathcal{E}_k$ of such a function are denoted by b_e and b_e', then $\hat{I}_k v$ has edge average $(b_e + b_e')/2$ for this e (and 0 for boundary edges). Thus, by (2.7),

$$\|\hat{I}_k(v)\|^2 \leq \frac{h_k^2}{4}\sum_{e\in\partial\mathcal{E}_k} 2\left(\frac{b_e + b_e'}{2}\right)^2 \leq \frac{h_k^2}{4}\sum_{E\in\mathcal{E}_k}\sum_{e\subset\partial E}\tilde{b}_e^2 \leq 5/2\|v\|^2\ ;$$

in the last summation $\tilde{b}_e$ is either b_e or b_e' depending on E.

The most important result is the second inequality in (2.13). According to (2.11), it is enough to establish that

$$\|I_k v - v\|_\mathcal{E}^2 \leq \|v\|_\mathcal{E}^2\ , \qquad \forall\, v \in V_{k-1}\ .$$

Going back to the above description of the edge averages of $I_k v - v$, we see that for each square in $\mathcal{E}_k$ two of them are zero, and the other two equal $\pm\frac{1}{8}\theta_\beta^1$ resp. $\pm\frac{1}{8}\theta_{\beta'}^2$ for appropriate multi-indices β, β'. Using the second equality in (2.6) gives

$$\|I_k v - v\|_{L^2(E)}^2 = \frac{1}{64}((\theta_\beta^1)^2 + (\theta_{\beta'}^2)^2 + \frac{3}{2}(\theta_\beta^1 - \theta_{\beta'}^2)^2) \leq \frac{1}{16}((\theta_\beta^1)^2 + (\theta_{\beta'}^2)^2)\ ,$$

and carefully adding all local estimates, we arrive at

$$\|I_k v - v\|_\mathcal{E}^2 \leq \frac{1}{4}\sum_{\text{interior}}((\theta_\beta^1)^2 + (\theta_{\beta'}^2)^2) + \frac{1}{8}\sum_{\text{boundary}}((\theta_\beta^1)^2 + (\theta_{\beta'}^2)^2),$$

(note that terms corresponding to interior edges occur four times while terms corresponding to boundary edges only twice). Now, recall that $\theta_\beta^1 = \Delta_\beta^1 - \Delta_{\beta-e^2}^1$ for interior e_β^1 (analogously for interior e_β^2) while $\theta_\beta^1 = 2\Delta_\beta^1$ for a e_β^1 on the lower boundary edge of the unit square (analogously for other boundary edges). Thus, by using again the crude estimate $(a + b)^2 \leq 2(a^2 + b^2)$ and regrouping the $(\Delta_\beta^1)^2$ and $(\Delta_\beta^1)^2$ terms with respect to the squares in $\mathcal{E}_{k-1}$, we obtain

$$\|I_k v - v\|_\mathcal{E}^2 \leq \sum_{E\in\mathcal{E}_{k-1}}((\Delta_E^1)^2 + (\Delta_E^2)^2)\ .$$

A second application of (2.6) gives the desired result. Lemma 2.3 is established. $\square$

In the remainder of this section we prove the following property of the iterated coarse-to-fine intergrid transfer operators R_k^K.

Lemma 2.4. *It holds that*

(2.15) $\|R_k^K v\|_\mathcal{E} \leq C\|v\|_\mathcal{E}, \qquad \forall v \in V_k, 1 \leq k \leq K.$

Proof. The proof is technical; it follows the idea of the proof of an analogous statement for the P_1 nonconforming elements [20]. First, we consider the case of $\Omega = \mathbb{R}^2$. That is, we assume that all our definitions are extended to infinite square partitions of $\mathbb{R}^2$; due to the local character of all constructions, this is easy to do. We keep the same notation for the extended partitions $\mathcal{E}_k$, edges $e_\alpha^j \in \partial\mathcal{E}_k$, squares $E \in \mathcal{E}_k$, etc. In order to guarantee the finiteness of all norm expressions, we restrict our

attention to functions $v \in V_k$ with finite support. By the construction of I_k, this property is preserved when applying the operators I_k and R_k^K.

After the extension to the shift-invariant setting of $\mathbb{R}^2$, it is clear that it suffices to consider the case of $k = 1$. Set, for simplicity, $\tilde{R}^k = R_1^k$, $k = 1, \ldots, K$. Our main observation from numerical experiments [23] was that the sequence

$$\{||\tilde{R}^k v - \tilde{R}^{k-1} v||_{\mathcal{E}}^2, \ k = 2, \ldots, K\}$$

decays geometrically. What we want to prove next is the mathematical counterpart to this observation. To formulate the technical result, introduce

$$\sigma_j = \sum_{\alpha \in \mathbf{Z}^2} (\theta_\alpha^j)^2, \qquad j = 0, 1, 2,$$

where the quantities θ_α^j are determined from the edge averages of $v \in V_1$ by the same formulas as above. The corresponding quantities computed for $\tilde{v} = I_2 v \in V_2$ are denoted by $\tilde{\theta}_\alpha^j$ and $\tilde{\sigma}_j$, $j = 0, 1, 2$. From (2.14) in Lemma 2.3, we see that

$$\sigma_1 + \sigma_2 \approx ||\tilde{R}^2 v - v||_{\mathcal{E}}^2 \quad \text{and} \quad \tilde{\sigma}_1 + \tilde{\sigma}_2 \approx ||\tilde{R}^3 v - \tilde{R}^2 v||_{\mathcal{E}}^2;$$

moreover, we can iterate this construction. Thus, if we can prove that

$$(2.16) \qquad \tilde{\sigma} \equiv c^* \tilde{\sigma}_0 + \tilde{\sigma}_1 + \tilde{\sigma}_2 \leq \gamma^* \sigma \equiv \gamma^* (c^* \sigma_0 + \sigma_1 + \sigma_2),$$

where $0 < \gamma^* < 1$ and $c^* > 0$ are constants independent of v, then, by Lemmas 2.2 and 2.3,

$$(2.17) \qquad \begin{aligned} ||R_1^K v||_{\mathcal{E}} \ &\leq ||v||_{\mathcal{E}} + \sum_{k=2}^{K} ||\tilde{R}^k v - \tilde{R}^{k-1} v||_{\mathcal{E}} \\ &\leq ||v||_{\mathcal{E}} + C \sum_{k=1}^{K-1} \sqrt{(\gamma^*)^k} \sqrt{\sigma} \\ &\leq C ||v||_{\mathcal{E}}. \end{aligned}$$

Since this gives the desired boundedness of R_k^K (for $\mathbb{R}^2$) via dilation, we concentrate on (2.16).

From (2.9) and (2.10) we find the following formulas for $\tilde{\theta}_\alpha^j$:

$$\tilde{\theta}_{2\beta}^0 = -\tfrac{1}{8}\theta_\beta^1 + \tfrac{1}{8}\theta_\beta^2 + \tfrac{1}{4}\theta_\beta^0,$$

$$\tilde{\theta}_{2\beta+e^1}^0 = \tfrac{1}{8}\theta_{\beta+e^1}^1 - \tfrac{1}{8}\theta_\beta^2 + \tfrac{1}{4}\theta_\beta^0,$$

$$\tilde{\theta}_{2\beta+e^2}^0 = \tfrac{1}{8}\theta_\beta^1 - \tfrac{1}{8}\theta_{\beta+e^2}^2 + \tfrac{1}{4}\theta_\beta^0,$$

$$\tilde{\theta}_{2\beta+e^1+e^2}^0 = -\tfrac{1}{8}\theta_{\beta+e^1}^1 + \tfrac{1}{8}\theta_{\beta+e^2}^2 + \tfrac{1}{4}\theta_\beta^0,$$

$$\tilde{\theta}_{2\beta}^1 = \tfrac{1}{2}\theta_\beta^1 - \tfrac{1}{8}(\theta_\beta^2 + \theta_{\beta-e^1}^2) - \tfrac{3}{8}(\theta_\beta^0 - \theta_{\beta-e^1}^0),$$

$$\tilde{\theta}_{2\beta+e^1}^1 = \tfrac{1}{4}\theta_\beta^2,$$

$$\tilde{\theta}_{2\beta+e^2}^1 = \tfrac{1}{2}\theta_\beta^1 - \tfrac{1}{8}(\theta_{\beta+e^2}^2 + \theta_{\beta-e^1+e^2}^2) + \tfrac{3}{8}(\theta_\beta^0 - \theta_{\beta-e^1}^0),$$

$$\tilde{\theta}_{2\beta+e^1+e^2}^1 = \tfrac{1}{4}\theta_{\beta+e^2}^2,$$

$$\tilde{\theta}_{2\beta}^2 = \tfrac{1}{2}\theta_\beta^2 - \tfrac{1}{8}(\theta_\beta^1 + \theta_{\beta-e^2}^1) + \tfrac{3}{8}(\theta_\beta^0 - \theta_{\beta-e^2}^0),$$

$$\tilde{\theta}_{2\beta+e^2}^2 = \tfrac{1}{4}\theta_\beta^1,$$

$$\tilde{\theta}_{2\beta+e^1}^2 = \tfrac{1}{2}\theta_\beta^2 - \tfrac{1}{8}(\theta_{\beta+e^1}^1 + \theta_{\beta-e^2+e^1}^1) - \tfrac{3}{8}(\theta_\beta^0 - \theta_{\beta-e^2}^0),$$

$$\tilde{\theta}_{2\beta+e^1+e^2}^2 = \tfrac{1}{4}\theta_{\beta+e^1}^1.$$

It is elementary but tedious to verify all these expressions, we give details for the first, fifth and sixth equations, and all others are similar and follow by some kind of symmetry argument:

$$
\begin{aligned}
\tilde{\theta}^0_{2\beta} &= a^1_{2\beta} + a^1_{2\beta+e^2} - a^2_{2\beta} - a^2_{2\beta+e^1} \\
&= b^1_\beta + \tfrac{1}{8}(b^2_\beta + b^2_{\beta-e^2} - b^2_{\beta+e^1} - b^2_{\beta+e^1-e^2}) + \tfrac{1}{8}(b^1_\beta + b^1_{\beta+e^2} + 5b^2_\beta + b^2_{\beta+e^1}) \\
&\quad - b^2_\beta - \tfrac{1}{8}(b^1_\beta + b^1_{\beta-e^1} - b^1_{\beta+e^2} - b^1_{\beta-e^1+e^2}) + \tfrac{1}{8}(5b^1_\beta + b^1_{\beta+e^2} + b^2_\beta + b^2_{\beta+e^1}) \\
&= \tfrac{1}{8}(3b^1_\beta - b^1_{\beta-e^1} + b^1_{\beta+e^2} + b^1_{\beta-e^1+e^2}) - \tfrac{1}{8}(3b^2_\beta - b^2_{\beta-e^2} + b^2_{\beta+e^1} b^2_{\beta+e^1-e^2}) \\
&= -\tfrac{1}{8}\theta^1_\beta + \tfrac{1}{8}\theta^2_\beta + \tfrac{1}{4}\theta^0_\beta \ , \\[1em]
\tilde{\theta}^1_{2\beta} &= a^1_{2\beta+e^2} - a^1_{2\beta} + a^1_{2\beta-e^1} - a^1_{2\beta-e^1+e^2} \\
&= \tfrac{1}{8}(b^1_\beta + b^1_{\beta+e^2} + 5b^2_\beta + b^2_{\beta+e^1}) - \tfrac{1}{8}(b^1_{\beta-e^1} + b^1_{\beta-e^1+e^2} + 5b^2_\beta + b^2_{\beta-e^1}) \\
&\quad + b^1_{\beta-e^1} + \tfrac{1}{8}(b^2_\beta + b^2_{\beta-e^2} - b^2_{\beta-e^1} - b^2_{\beta-e^1-e^2}) \\
&\quad - b^1_\beta - \tfrac{1}{8}(b^2_\beta + b^2_{\beta-e^2} - b^2_{\beta+e^1} - b^2_{\beta+e^1-e^2}) \\
&= \tfrac{7}{8}(b^1_{\beta-e^1} - b^1_\beta) + \tfrac{1}{8}(b^1_{\beta+e^2} - b^1_{\beta-e^1+e^2}) \\
&\quad + \tfrac{1}{4}(b^2_{\beta+e^1} - b^2_{\beta-e^1}) + \tfrac{1}{8}(b^2_{\beta+e^1-e^2} - b^2_{\beta-e^1-e^2}) \\
&= \tfrac{1}{2}\theta^1_\beta - \tfrac{1}{8}(\theta^2_\beta + \theta^2_{\beta-e^1}) - \tfrac{3}{8}(\theta^0_\beta - \theta^0_{\beta-e^1}) \ ,
\end{aligned}
$$

and

$$
\begin{aligned}
\tilde{\theta}^1_{2\beta+e^1} &= (a^1_{2\beta+e^1+e^2} - a^1_{2\beta+e^2}) + (a^1_{2\beta^1} - a^1_{2\beta+e^1}) \\
&= \tfrac{1}{2}(b^2_{\beta+e^1} - b^2_\beta) + \tfrac{1}{4}(b^2_\beta + b^2_{\beta-e^2} - b^2_{\beta+e^1} - b^2_{\beta+e^1-e^2}) \\
&= \tfrac{1}{4}(-b^2_\beta + b^2_{\beta-e^2} + b^2_{\beta+e^1} - b^2_{\beta+e^1-e^2}) = \tfrac{1}{4}\theta^2_\beta \ .
\end{aligned}
$$

These formulas are used to compute the quantities $\tilde{\sigma}_j$. In order to present the calculations in reasonably short form, we introduce the notation

$$
\sigma^{\beta^*}_j = \sum_{\beta\in\mathbf{Z}^2} \theta^j_\beta \theta^j_{\beta+\beta^*}, \quad \sigma^{\beta^*}_{jl} = \sum_{\beta\in\mathbf{Z}^2} \theta^j_\beta \theta^l_{\beta+\beta^*}, \quad k,l = 0,1,2 \ (j \neq l);
$$

if $\beta^* \in \mathbb{Z}^2$ is the null vector, it is omitted in this notation. With them, we see, by carefully evaluating all squares, that

$$
\begin{aligned}
\tilde{\sigma}_0 &= \sum_\alpha (\tilde{\theta}^0_\alpha)^2 = \sum_\beta \left((\tilde{\theta}^0_{2\beta})^2 + (\tilde{\theta}^0_{2\beta+e^1})^2 + (\tilde{\theta}^0_{2\beta+e^2})^2 + (\tilde{\theta}^0_{2\beta+e^1+e^2})^2 \right) \\
&= \tfrac{1}{16}\sigma_0 + \tfrac{1}{64}(\sigma_1 + \sigma_2) + \tfrac{1}{16}(-\sigma_{01} + \sigma_{02}) - \tfrac{1}{32}\sigma_{12} \\
&\quad + \tfrac{1}{16}\sigma_0 + \tfrac{1}{64}(\sigma_1 + \sigma_2) + \tfrac{1}{16}(\sigma^{e^1}_{01} - \sigma_{02}) - \tfrac{1}{32}\sigma^{-e^1}_{12} \\
&\quad + \tfrac{1}{16}\sigma_0 + \tfrac{1}{64}(\sigma_1 + \sigma_2) + \tfrac{1}{16}(\sigma_{01} - \sigma^{e^2}_{02}) - \tfrac{1}{32}\sigma^{e^2}_{12} \\
&\quad + \tfrac{1}{16}\sigma_0 + \tfrac{1}{64}(\sigma_1 + \sigma_2) + \tfrac{1}{16}(-\sigma^{e^1}_{01} + \sigma^{e^2}_{02}) - \tfrac{1}{32}\sigma^{-e^1+e^2}_{12} \\
&= \tfrac{1}{4}\sigma_0 + \tfrac{1}{16}(\sigma_1 + \sigma_2) - \tfrac{1}{32}\underbrace{(\sigma_{12} + \sigma^{e^2}_{12} + \sigma^{-e^1}_{12} + \sigma^{e^2-e^1}_{12})}_{\equiv \sigma^*} .
\end{aligned}
$$

Analogously,

$$
\begin{aligned}
\tilde{\sigma}_1 &= \sum_\beta \left((\tilde{\theta}_{2\beta}^1)^2 + (\tilde{\theta}_{2\beta+e^1}^1)^2 + (\tilde{\theta}_{2\beta+e^2}^1)^2 + (\tilde{\theta}_{2\beta+e^1+e^2}^1)^2 \right) \\
&= \left(\tfrac{9}{32}(\sigma_0 - \sigma_0^{e^1}) + \tfrac{1}{4}\sigma_1 + \tfrac{1}{32}(\sigma_2 + \sigma_2^{e^1}) - \tfrac{3}{8}(\sigma_{01} - \sigma_{01}^{e^1}) \right. \\
&\quad \left. + \tfrac{3}{32}(\sigma_{02}^{-e^1} - \sigma_{02}^{e^1}) - \tfrac{1}{8}(\sigma_{12} + \sigma_{12}^{-e^1}) \right) + \tfrac{1}{16}\sigma_2 \\
&\quad + \left(\tfrac{9}{32}(\sigma_0 - \sigma_0^{e^1}) + \tfrac{1}{4}\sigma_1 + \tfrac{1}{32}(\sigma_2 + \sigma_2^{e^1}) + \tfrac{3}{8}(\sigma_{01} - \sigma_{01}^{e^1}) \right. \\
&\quad \left. + \tfrac{3}{32}(\sigma_{02}^{e^1+e^2} - \sigma_{02}^{-e^1+e^2}) - \tfrac{1}{8}(\sigma_{12}^{e^2} + \sigma_{12}^{-e^1+e^2}) \right) + \tfrac{1}{16}\sigma_2 \\
&= \tfrac{9}{16}\sigma_0 + \tfrac{1}{2}\sigma_1 + \tfrac{3}{16}\sigma_2 - \tfrac{9}{16}\sigma_0^{e^1} + \tfrac{1}{16}\sigma_2^{e^1} - \tfrac{1}{8}\sigma^* \\
&\quad - \tfrac{3}{32}(\sigma_{02}^{e^1} + \sigma_{02}^{-e^1+e^2} - \sigma_{02}^{-e^1} - \sigma_{02}^{e^1+e^2}), \\
\tilde{\sigma}_2 &= \tfrac{9}{16}\sigma_0 + \tfrac{3}{16}\sigma_1 + \tfrac{1}{2}\sigma_2 - \tfrac{9}{16}\sigma_0^{e^2} + \tfrac{1}{16}\sigma_1^{e^2} - \tfrac{1}{8}\sigma^* \\
&\quad - \tfrac{3}{32}(\sigma_{01}^{-e^2} + \sigma_{01}^{e^1+e^2} - \sigma_{01}^{e^2} - \sigma_{01}^{e^1-e^2}).
\end{aligned}
$$

Thus, introducing $\mathcal{A} = \sigma_1 + \sigma_2$ and $\tilde{\mathcal{A}} = \tilde{\sigma}_1 + \tilde{\sigma}_2$, we have

$$
\begin{aligned}
\tilde{\sigma}_0 &= \tfrac{1}{4}\sigma_0 + \tfrac{1}{16}\mathcal{A} - \tfrac{1}{32}\sigma^*, \\
\tilde{\mathcal{A}} &= \tfrac{9}{8}\sigma_0 + \tfrac{11}{16}\mathcal{A} - \tfrac{9}{16}(\sigma_0^{e^1} + \sigma_0^{e^2}) + \tfrac{1}{16}(\sigma_1^{e^2} + \sigma_2^{e^1}) - \tfrac{1}{4}\sigma^* - \tfrac{3}{32}\sigma^{**},
\end{aligned}
\tag{2.18}
$$

where

$$
\sigma^{**} = \sigma_{01}^{-e^2} + \sigma_{01}^{e^1+e^2} + \sigma_{02}^{e^1} + \sigma_{02}^{-e^1+e^2} - \sigma_{01}^{e^2} - \sigma_{01}^{e^1-e^2} - \sigma_{02}^{-e^1} - \sigma_{02}^{e^1+e^2}.
$$

Next, we simplify σ^* and σ^{**}. Note that

$$
\begin{aligned}
\sigma^* - 2\sigma_1^{e^2} &= \sum_\beta \theta_\beta^1 (\theta_\beta^2 + \theta_{\beta+e^2}^2 + \theta_{\beta-e^1+e^2}^2 + \theta_{\beta-e^1}^2 - \theta_{\beta+e^2}^1 - \theta_{\beta-e^2}^1) \\
&= \sum_\beta \theta_\beta^1 (\theta_{\beta+e^2-e^1}^0 + \theta_{\beta-e^2}^0 - \theta_{\beta-e^1-e^2}^0 - \theta_{\beta+e^2}^0 + 2\theta_\beta^1), \\
\sigma^* - 2\sigma_2^{e^1} &= \sum_\beta \theta_\beta^2 (\theta_\beta^1 + \theta_{\beta-e^2}^1 + \theta_{\beta+e^1-e^2}^1 + \theta_{\beta+e^1}^1 - \theta_{\beta+e^1}^2 - \theta_{\beta-e^1}^2) \\
&= \sum_\beta \theta_\beta^2 (\theta_{\beta-e^2-e^1}^0 + \theta_{\beta+e^1}^0 - \theta_{\beta+e^1-e^2}^0 - \theta_{\beta-e^1}^0 + 2\theta_\beta^2),
\end{aligned}
$$

so that

$$
\sigma^* = \sigma_1^{e^2} + \sigma_2^{e^1} + \mathcal{A} - \frac{1}{2}\sigma^{**}.
$$

Analogously, we can simplify σ^{**} as follows:

$$
\begin{aligned}
\sigma^{**} &= \sum_\beta \theta_\beta^0 \big((\theta_{\beta+e^1+e^2}^1 + \theta_{\beta-e^2}^1 + \theta_{\beta+e^1}^2 + \theta_{\beta-e^1+e^2}^2) \\
&\quad - (\theta_{\beta+e^2}^1 + \theta_{\beta+e^1-e^2}^1 + \theta_{\beta-e^1}^2 + \theta_{\beta+e^1+e^2}^2) \big) \\
&= \sum_\beta \theta_\beta^0 \big((\theta_{\beta+e^1+e^2}^0 + \theta_{\beta+e^1-e^2}^0 + \theta_{\beta-e^1-e^2}^0 + \theta_{\beta-e^1+e^2}^0) \\
&\quad - 2(\theta_{\beta+e^1}^0 + \theta_{\beta+e^2}^0 + \theta_{\beta-e^1}^0 + \theta_{\beta-e^2}^0) + 4\theta_\beta^0) \big) \\
&= 2(\sigma_0^{e^1+e^2} + \sigma_0^{e^1-e^2}) - 4(\sigma_0^{e^1} + \sigma_0^{e^2}) + 4\sigma_0.
\end{aligned}
$$

In these calculations, the identity

$$
\theta_\beta^1 + \theta_{\beta-e^2}^1 - \theta_\beta^2 - \theta_{\beta-e^1}^2 = \theta_\beta^0 - \theta_{\beta-e^2}^0 - \theta_{\beta-e^1}^0 + \theta_{\beta-e^1-e^2}^0
$$

which is valid for arbitrary $\beta \in \mathbb{Z}^2$ and shows that the sequences $\{\theta_\beta^j\}$ are not completely independent, has been used several times.

Substitution of the expressions for σ^* and σ^{**} into (2.18) leads to

$$
\begin{aligned}
\tilde{\sigma}_0 &= \tfrac{1}{4}\sigma_0 + \tfrac{1}{32}\mathcal{A} - \tfrac{1}{32}(\sigma_1^{e^2} + \sigma_2^{e^2}) + \tfrac{1}{64}\sigma^{**} \\
(2.19) \qquad &= \tfrac{5}{16}\sigma_0 + \tfrac{1}{32}\mathcal{A} + \tfrac{1}{32}\left(\sigma_0^{e^1-e^2} + \sigma_0^{e^1+e^2} - 2\sigma_0^{e^1} - 2\sigma_0^{e^2} - \sigma_1^{e^2} - \sigma_2^{e^1}\right) \\
&\leq \tfrac{1}{2}\sigma_0 + \tfrac{1}{16}\mathcal{A},
\end{aligned}
$$

where we have used the fact that $|\sigma_j^{\beta^*}| \leq \sigma_j$, $j = 0, 1, 2$, which is valid for arbitrary β^*. With the same argument, we see that

$$
\begin{aligned}
\tilde{\mathcal{A}} &= \tfrac{9}{8}\sigma_0 + \tfrac{7}{16}\mathcal{A} - \tfrac{9}{16}\left(\sigma_0^{e^1} + \sigma_0^{e^2}\right) - \tfrac{3}{16}\left(\sigma_1^{e^2} + \sigma_2^{e^2}\right) + \tfrac{1}{32}\sigma^{**} \\
(2.20) \qquad &= \tfrac{5}{4}\sigma_0 + \tfrac{7}{16}\mathcal{A} + \tfrac{1}{16}\left(\sigma_0^{e^1-e^2} + \sigma_0^{e^1+e^2}\right) \\
&\quad - \tfrac{11}{16}\left(\sigma_0^{e^1} + \sigma_0^{e^2}\right) - \tfrac{3}{16}\left(\sigma_1^{e^2} + \sigma_2^{e^1}\right) \\
&\leq \tfrac{11}{4}\sigma_0 + \tfrac{5}{8}\mathcal{A}.
\end{aligned}
$$

Now, set $\mathcal{B} = c\sigma_0$ and $\tilde{\mathcal{B}} = c\tilde{\sigma}_0$. Then it follows from (2.18) and (2.19) that

$$
\tilde{\mathcal{A}} \leq \frac{5}{8} + \frac{11}{4c}\mathcal{B}, \qquad \tilde{\mathcal{B}} \leq \frac{c}{16}\mathcal{A} + \frac{1}{2}\mathcal{B},
$$

and

$$
(\tilde{\mathcal{A}} + \tilde{\mathcal{B}}) \leq \max\left(\frac{5}{8} + \frac{c}{16}, \frac{11}{4c} + \frac{1}{2}\right)(\mathcal{A} + \mathcal{B}).
$$

Let $c = c^* \equiv 3\sqrt{5} - 1$, so we see that (2.16) holds with

$$
\gamma^* = \frac{5}{8} + \frac{c^*}{16} = \frac{11}{4c^*} + \frac{1}{2} = \frac{3\sqrt{5} + 9}{16} < 1.
$$

It remains to reduce the assertion of Lemma 2.4 to the shift-invariant situation just considered. To this end, starting with any $v \in V_k$ on the unit square, we repeatedly use an odd extension. Namely, set $\hat{v} = v$ on $[0,1]^2$ and

$$
\hat{v}(x,y) = -\hat{v}(-x,y), \qquad (x,y) \in [-1,0) \times [0,1];
$$

after this, define

$$
\hat{v}(x,y) = -\hat{v}(x,-y), \qquad (x,y) \in [-1,1] \times [-1,0),
$$

and continue this extension process with the unit square replaced by $[-1,1]^2$ such that after the next two steps $\hat{v}$ is defined on $[-1,3]^2$. Outside this larger square we continue by zero. Clearly, $\|\hat{v}\|_{\mathcal{E}}^2 = 16\|v\|_{\mathcal{E}}^2$, where the norms for $\hat{v}$ and v are taken with respect to $\Re^2$ and the unit square, respectively.

It is not difficult to check by induction that on $[0,1]^2$ the functions $R_k^K \hat{v}$ (obtained by the repeated application of the prolongations defined on $\Re^2$) and $R_k^K v$ (as defined above with respect to $[0,1]^2$) coincide. Also, the values of $I_{k+1}\hat{v}$ on $[-2^{-(k+1)}, 1 + 2^{-(k+1)}]^2$ depend solely on the values of $\hat{v}$ on the square $[-2^{-k}, 1 + 2^{-k}]^2$, and on this enlarged square $I_{k+1}\hat{v}$ coincides with its odd extension from $[0,1]^2$. Finally, the zero edge averages are automatically reproduced along the boundary of $[0,1]^2$ from the above extension procedure. Therefore, by (2.17) and the dilation argument, we obtain

$$
\|R_k^K v\|_{\mathcal{E}}^2 \leq \|R_k^K \hat{v}\|_{\mathcal{E}}^2 \leq C\|\hat{v}\|_{\mathcal{E}}^2 = 16C\|v\|_{\mathcal{E}}^2,
$$

which finishes the proof of Lemma 2.4. $\qquad\square$

Let us conclude this section with the following remark. All proofs given so far are valid for the unit square $\Omega = (0,1)^2$ and sequences of uniform square partitions $\mathcal{E}_k$ as indicated above. The energy norm is the one corresponding to the Dirichlet problem for the Laplace equation. Since constants (such as in the second estimate of (2.13)) are sometimes crucial for what follows, one has to be careful with generalizations. For instance, the second relation in (2.13) is valid whenever $\mathcal{E}_{k-1}$ is a collection of equally sized squares. This covers certain L-shaped domains Ω. However, replacing the $H_0^1(\Omega)$ norm by more general energy norms seems to be problematic. Lemma 2.4 can be extended to polygonal domains if they are equipped with an initial partition $\tilde{\mathcal{E}}_1$ (into quadrilaterals) which is topologically equivalent to the above considered square partition $\mathcal{E}_1$ of the unit square. Then, the sequence $\{\tilde{\mathcal{E}}_k\}$ can then be inherited from $\{\mathcal{E}_k\}$. Using the parametric version of the NR Q_1 elements [25], one easily sees that intergrid operators associated with $\{\tilde{\mathcal{E}}_k\}$ are just copies of the I_k and R_k^K considered above. The result of Lemma 2.4 then carries over by spectral equivalence of the norms (note that (2.15) is insensitive to replacing $(\cdot, \cdot)_\mathcal{E}$ by spectrally equivalent forms; i.e., the uniform boundedness assertion remains valid for more general second-order uniformly elliptic problems than the Poisson equation). We did not check any details for the n-dimensional counterparts ($n \geq 3$) of these elements as defined in [25].

We will not discuss any further the possible extensions of the above properties of intergrid operators. Below we will indicate which of the algorithms can be justified to converge for larger classes of domains, partition sequences and second-order elliptic boundary value problems, respectively.

3. MULTIGRID ALGORITHMS

In this section and the next section we consider multigrid algorithms and multilevel preconditioners for the numerical solution of the second-order elliptic problem

$$\text{(3.1)} \qquad \begin{aligned} -\nabla \cdot (\mathbf{A}\nabla u) &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \Gamma, \end{aligned}$$

where $\Omega \subset \Re^2$ is a simply connected bounded polygonal domain with the boundary Γ, $f \in L^2(\Omega)$, and the symmetric coefficient matrix $\mathbf{A} \in (L^\infty(\Omega))^{2 \times 2}$ satisfies

$$\text{(3.2)} \qquad \alpha_1 \xi^t \xi \geq \xi^t \mathbf{A}(x,y)\xi \geq \alpha_0 \xi^t \xi, \qquad (x,y) \in \Omega, \ \xi \in \Re^2,$$

with fixed constants $\alpha_1, \alpha_0 > 0$. This guarantees that the energy norm related to (3.1) is spectrally equivalent to that for the homogeneous Dirichlet problem for the Poisson equation, i.e., to the $H_0^1(\Omega)$ norm. The condition number of preconditioned linear systems to be analyzed later depends on the ratio α_1/α_0. However, some of the multigrid results below are only valid in the case of the Laplace operator, i.e. if $\mathbf{A}(x,y) \equiv \alpha_0 \mathbf{I}$.

Problem (3.1) is recast in weak form as follows. The bilinear form $a(\cdot, \cdot)$ is defined by

$$a(v,w) = (\mathbf{A}\nabla v, \nabla w), \qquad v, \, w \in H^1(\Omega).$$

Then the weak form of (3.1) for the solution $u \in H_0^1(\Omega)$ is

$$\text{(3.3)} \qquad a(u,v) = (f,v), \qquad \forall \, v \in H_0^1(\Omega).$$

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

$$a_k(v, w) = \sum_{E \in \mathcal{E}_k} (\mathbf{A} \nabla v, \nabla w)_E, \qquad v, w \in V_k \oplus H_0^1(\Omega).$$

The NR Q_1 finite element discretization of (3.1) is to find $u_K \in V_K$ such that

(3.4) $a_K(u_K, v) = (f, v), \qquad \forall\, v \in V_K.$

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

(3.5) $(A_k v, w) = a_k(v, w), \qquad \forall\, w \in V_k.$

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 $R_{k-1} : V_k \to V_{k-1}$ and $R_{k-1}^0 : V_k \to V_{k-1}$ by

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

and

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

It is easy to see that $I_k R_{k-1}$ is a symmetric operator with respect to the a_k form. Note that neither R_k^0 nor R_k is a projection in the nonconforming case. Finally, let Λ_k dominate the spectral radius of A_k.

The multigrid processes below result in a linear iterative scheme with a reduction operator equal to $\mathrm{Id}_K - B_K A_K$, where $B_K : V_K \to V_K$ is the multigrid operator to be defined below.

Multigrid Algorithm 3.1. Let $2 \le k \le K$ 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 $l = 1, \ldots, m(k)$ by

$$x^l = x^{l-1} + S_k(g - A_k x^{l-1}).$$

3. Define $y^{m(k)} = x^{m(k)} + I_k q^p$, where q^i for $i = 1, \ldots, p$ is defined by

$$q^i = q^{i-1} + B_{k-1} \left[R_{k-1}^0 \left(g - A_k x^{m(k)} \right) - A_{k-1} q^{i-1} \right].$$

4. Define y^l for $l = m(k) + 1, \ldots, 2m(k)$ by

$$y^l = y^{l-1} + S_k \left(g - A_k y^{l-1} \right).$$

5. Set $B_k g = y^{2m(k)}$.

In Algorithm 3.1, $m(k)$ gives the number of pre- and post-smoothing iterations and can vary as a function of k. In this section, we set $S_k = (\Lambda_k)^{-1} \mathrm{Id}_k$ in the pre- and post-smoothing steps. If $p = 1$, we have a $\mathcal{V}$-cycle multigrid algorithm. If $p = 2$, we have a $\mathcal{W}$-cycle algorithm. 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 $m(k) = 2^{K-k}$).

We now follow the methodology developed in [6] to state convergence results for Algorithm 3.1. The two ingredients in their analysis are the regularity and approximation property and the boundedness of the intergrid transfer operator:

(3.6) $\left| a_k \left(v - I_k R_{k-1} v, v \right) \right| \le C \dfrac{\| A_k v \|}{\sqrt{\lambda_k}} \sqrt{a_k(v, v)}, \qquad \forall\, v \in V_k,$

and

$$(3.7) \qquad a_k(I_k v, I_k v) \le C a_{k-1}(v, v), \qquad \forall \, v \in V_{k-1},$$

for $k = 2, \dots, K$, where λ_k is the largest eigenvalue of A_k. The proof of (3.6) is standard under the full elliptic regularity assumption on the solution of (3.1); see the proof of a similar result for the P_1 nonconforming elements in [15]. Inequality (3.7) has been shown in [1] using the approximation property of the operator I_k. However, here we see that if $\mathbf{A} = \alpha_0 \mathbf{I}$ is a scalar multiple of the two-by-two identity matrix $\mathbf{I}$, by the second inequality in (2.13) in Lemma 2.3, we actually have

$$(3.8) \qquad a_k(I_k v, I_k v) \le 2 a_{k-1}(v, v), \qquad \forall \, v \in V_{k-1}.$$

This leads to the following main result of this section. Let the convergence rate for Algorithm 3.1 on the kth level be measured by the convergence factor δ_k satisfying

$$|a_k (v - B_k A_k v, v)| \le \delta_k a_k(v, v), \qquad \forall \, v \in V_k.$$

Theorem 3.2. *Define B_k by $p = 2$ and $m(k) = m$ for all k in Algorithm 3.1. Then, for $\Omega = (0, 1)^2$, if $\mathbf{A} = \alpha_0 \mathbf{I}$ is constant, there exists $C > 0$, independent of k, such that*

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

The proof of this theorem follows from (3.6), (3.8), and Theorem 7 in [6]. From Theorem 3.2, we have an optimal convergence property of the $\mathcal{W}$-cycle with one smoothing. While a uniform preconditioner result for the variable $\mathcal{V}$-cycle has been given for the first version of the NR Q_1 elements in [9], we see from (3.6) and (3.7) that the same result also holds for the second version even in the case of the variable coefficient $\mathbf{A}$. That is, defining B_k by $p = 1$ and $m(k) = 2^{K-k}$ for $k = 2, \dots, K$, there are $\eta_0, \eta_1 > 0$, independent of k, such that

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

with

$$\eta_0 \ge \sqrt{m(k)}/(C + \sqrt{m(k)}) \quad and \quad \eta_1 \le (C + \sqrt{m(k)})/\sqrt{m(k)}.$$

Finally, we mention that for a general $\mathbf{A}$ the convergence result for the $\mathcal{W}$-cycle can be theoretically established (e.g., by the theory of [6]) only for sufficiently many smoothing steps on each level, and that Theorem 3.2 is a first improvement for the model problem under consideration.

4. MULTILEVEL PRECONDITIONERS

In this section we discuss additive multilevel preconditioners of hierarchical basis and BPX type for (3.4). We assume that the reader is familiar with the theory of additive Schwarz methods as outlined in [16]; see also [21], [30], or [28]. Below we use the notation

$$\{V; a(\cdot, \cdot)\} = \sum_k R_k \{V_k; b_k(\cdot, \cdot)\},$$

which briefly expresses the following assumptions: V, V_k are finite-dimensional Hilbert spaces, equipped with their respective symmetric positive definite bilinear forms $a(\cdot, \cdot), b_k(\cdot, \cdot)$, $R_k : V_k \to V$ are linear mappings such that the space V is the (not necessarily direct) sum of its subspaces $R_k V_k$. Since in our applications $V_k \not\subset V$, the R_k are not just natural embeddings, their choice is a crucial ingredient

of the algorithms which are associated with the above space splitting. Roughly speaking, these algorithms aim at iteratively solving a variational problem on V governed by the bilinear form $a(\cdot, \cdot)$, by solving subproblems in V_k associated with the form $b_k(\cdot, \cdot)$. The transfer of information between V and the V_k is performed by the operators R_k and their adjoints. A small condition number of the space splitting (which is expressed by certain two-sided norm equivalencies; see below) guarantees good convergence rates of these algorithms. For details, see the above references.

We start with a theoretical result which follows from the material in Section 2 along the lines of [23]. Since we rely on Lemma 2.4, we assume that Ω is the unit square, and that $\{\mathcal{E}_k\}$ is a sequence of uniform square partitions (compare, however, the remark at the end of Section 2 about extensions of Lemma 2.4). More precisely, we derive the condition numbers of the additive space splittings

$$(4.1) \qquad \{V_K; (\cdot, \cdot)_{\mathcal{E}}\} = R_1^K \{V_1; (\cdot, \cdot)_{\mathcal{E}}\} + \sum_{k=2}^{K} R_k^K \{V_k; 2^{2k}(\cdot, \cdot)\},$$

and

$$(4.2) \qquad \{V_K; (\cdot, \cdot)_{\mathcal{E}}\} = R_1^K \{V_1; (\cdot, \cdot)_{\mathcal{E}}\} + \sum_{k=2}^{K} R_k^K \{(\mathrm{Id}_k - I_k P_{k-1})V_k; 2^{2k}(\cdot, \cdot)\}.$$

The condition number of (4.1) is given by [21]

$$(4.3) \qquad \kappa = \frac{\lambda_{\max}}{\lambda_{\min}}, \quad \lambda_{\max} = \sup_{v \in V_K} \frac{||v||_{\mathcal{E}}^2}{|||v|||^2}, \quad \lambda_{\min} = \inf_{v \in V_K} \frac{||v||_{\mathcal{E}}^2}{|||v|||^2},$$

where

$$|||v|||^2 = \inf_{v_k \in V_k} \left\{ ||v_1||_{\mathcal{E}}^2 + \sum_{k=2}^{K} 2^{2k} ||v_k||^2 \right\},$$

with $v = \sum_k R_k^K v_k$. A similar definition can be given for (4.2).

Theorem 4.1. *Under the above assumptions on Ω and $\{\mathcal{E}_k\}$, there are positive constants c and C, independent of K, such that*

$$(4.4) \qquad c \leq \frac{||v||_{\mathcal{E}}^2}{|||v|||^2} \leq CK, \qquad \forall v \in V_K,$$

and

$$(4.5) \qquad c \leq \frac{||v||_{\mathcal{E}}^2}{|||v|||^2} \leq CK, \qquad \forall v \in V_K,$$

where

$$|||v|||^2 = ||Q_1^K v||_{\mathcal{E}}^2 + \sum_{k=2}^{K} 2^{2k} ||(\mathrm{Id}_k - I_k P_{k-1})Q_k^K v||^2.$$

That is, the condition numbers of the additive space splittings (4.1) and (4.2) are bounded by $\mathrm{O}(K)$ as $K \to \infty$.

Proof. For $k = 2, \ldots, K$, it follows from the definitions of I_k, $\hat{I}_k$, and Q_k^K, (2.4), and the first inequality of (2.13) that

$$
\begin{aligned}
2^{2k}||(\mathrm{Id}_k - I_k P_{k-1})Q_k^K v||^2 &= 2^{2k}||\hat{I}_k(\mathrm{Id}_k - P_{k-1})Q_k^K v||^2 \\
&\leq \tfrac{5}{2}\, 2^{2k}||(\mathrm{Id}_k - P_{k-1})Q_k^K v||^2 \\
&\leq C||(\mathrm{Id}_k - P_{k-1})Q_k^K v||_{\mathcal{E}}^2 \\
&= C||Q_k^K v - Q_{k-1}^K v||_{\mathcal{E}}^2.
\end{aligned}
$$

Summing on j and using the orthogonality relations in (2.3), we see that

$$
\begin{aligned}
\inf_{v_k \in V_k} &\left\{ ||v_1||_{\mathcal{E}}^2 + \textstyle\sum_{k=2}^{K} 2^{2k}||v_k||^2 \right\} \\
&\leq ||Q_1^K v||_{\mathcal{E}}^2 + \textstyle\sum_{k=2}^{K} 2^{2k}||(\mathrm{Id}_k - I_k P_{k-1})Q_k^K v||^2 \\
&\leq C||v||_{\mathcal{E}}^2,
\end{aligned}
$$

with $v = \sum_k R_k^K v_k$, which implies the lower bounds in (4.4) and (4.5).

For the upper bounds, we consider an arbitrary decomposition $v = \sum_{k=1}^{K} R_k^K v_k$ with $v_k \in V_k$. Then we see, by Lemma 2.4, that

$$
||v||_{\mathcal{E}}^2 \leq \left(\sum_{k=1}^{K} ||R_k^K v_k||_{\mathcal{E}} \right)^2 \leq K \sum_{k=1}^{K} ||R_k^K v_k||_{\mathcal{E}}^2 \leq CK \sum_{k=1}^{K} ||v_k||_{\mathcal{E}}^2.
$$

Consequently, by (2.2), we have

$$
||v||_{\mathcal{E}}^2 \leq CK \left(||v_1||_{\mathcal{E}}^2 + \sum_{k=2}^{K} 2^{2k}||v_k||^2 \right).
$$

Now, taking the infimum with respect to all decompositions, we obtain

$$
\begin{aligned}
||v||_{\mathcal{E}}^2 &\leq CK \inf_{v_k \in V_k} \left\{ ||v_1||_{\mathcal{E}}^2 + \textstyle\sum_{k=2}^{K} 2^{2k}||v_k||^2 \right\} \\
&\leq CK \left(||Q_1^K v||_{\mathcal{E}}^2 + \textstyle\sum_{k=2}^{K} 2^{2k}||(\mathrm{Id}_k - I_k P_{k-1})Q_k^K v||^2 \right),
\end{aligned}
$$

with $v = \sum_k R_k^K v_k$, which finishes the proof of the theorem. $\square$

We now discuss the algorithmical consequences for the splittings (4.1) and (4.2). Theoretically, Theorem 4.1 already produces suitable preconditioners for the matrix A_K using (4.1) and (4.2). However, they are still complicated since they involve L^2-projections onto V_k, $1 < k < K$, which means to solve large linear systems within each preconditioning step. To get more practicable algorithms, we replace the L^2 norms in V_k and $W_k = (\mathrm{Id}_k - I_k P_{k-1})V_k \subset V_k$, $k = 2, \ldots, K$, by their suitable discrete counterparts. We first consider the splitting (4.1); (4.2) will be discussed later.

Let $\{\phi_{\alpha,k}^j\}$ be the basis functions of V_k such that the edge average of $\phi_{\alpha,k}^j$ equals one at $e_{\alpha,k}^j$ and zero at all other edges. Then each $v \in V_k$ has the representation

$$
v = \sum_{j=1}^{2} \sum_{\alpha} a_\alpha^j \phi_{\alpha,k}^j.
$$

Thus, by the uniform L^2-stability of the bases, which follows from (2.7) in Lemma 2.2, we see that

$$(4.6) \qquad \frac{1}{5}2^{-2k}\sum_{j=1}^{2}\sum_{\alpha}(a_\alpha^j)^2 \le ||v||^2 \le \frac{1}{2}2^{-2k}\sum_{j=1}^{2}\sum_{\alpha}(a_\alpha^j)^2.$$

Note that (with the same argument as in Lemma 2.2)

$$(4.7) \qquad 2^{2k}||\phi_{\alpha,k}^j||^2 = \frac{41}{120}, \quad a_k(\phi_{\alpha,k}^j,\phi_{\alpha,k}^j) \approx ||\phi_{\alpha,k}^j||_{\mathcal{E}}^2 = 5,$$

so (4.6) can be interpreted as the two-sided inequality associated with the stability of any of the splittings

$$(4.8) \qquad \{V_k; 2^{2k}(\cdot,\cdot)\} = \sum_{j=1}^{2}\sum_{\alpha}\{V_{\alpha,k}^j; 2^{2k}(\cdot,\cdot)\},$$

$$(4.9) \qquad \{V_k; 2^{2k}(\cdot,\cdot)\} = \sum_{j=1}^{2}\sum_{\alpha}\{V_{\alpha,k}^j; (\cdot,\cdot)_{\mathcal{E}}\},$$

and

$$(4.10) \qquad \{V_k; 2^{2k}(\cdot,\cdot)\} = \sum_{j=1}^{2}\sum_{\alpha}\{V_{\alpha,k}^j; a_k(\cdot,\cdot)\},$$

into the direct sum of one-dimensional subspaces $V_{\alpha,k}^j$ spanned by the basis functions $\phi_{\alpha,k}^j$. Any of the splittings (4.8)–(4.10) can be used to refine (4.1). As we will see below, the difference is just in a diagonal scaling (i.e., a multiplication by a diagonal matrix) in the final algorithms. As example, we consider the splitting (4.10) in detail; the other two cases can be analyzed in the same fashion.

With (4.1) and (4.10), we have the splitting

$$(4.11) \qquad \{V_K; a_K(\cdot,\cdot)\} = R_1^K\{V_1; a_1(\cdot,\cdot)\} + \sum_{k=2}^{K}\sum_{j=1}^{2}\sum_{\alpha} R_k^K\{V_{\alpha,k}^j; a_k(\cdot,\cdot)\}.$$

It follows from (4.4), (4.6), and (4.7) that the condition number κ for (4.11) still behaves like O(K). Now, associated with this splitting we can explicitly state the additive Schwarz operator

$$(4.12) \qquad \mathcal{P}_K = R_1^K T_1 + \sum_{k=2}^{K}\sum_{j=1}^{2}\sum_{\alpha} R_k^K T_{\alpha,k}^j,$$

where

$$T_{\alpha,k}^j v = \frac{a_K(v, R_k^K \phi_{\alpha,k}^j)}{a_k(\phi_{\alpha,k}^j, \phi_{\alpha,k}^j)} \phi_{\alpha,k}^j,$$

and $T_1 v \in V_1$ solves the elliptic problem

$$a_1(T_1 v, w) = a_K(v, R_1^K w), \qquad \forall\, w \in V_1.$$

Thus the matrix representations of all operators with respect to the bases of the respective V_k are

$$T_k = \sum_{j=1}^{2} \sum_{\alpha} T_{\alpha,k}^j = S_k (R_k^K)^t A_K, \qquad S_k = \operatorname{diag}(a_j(\phi_{\alpha,k}^j, \phi_{\alpha,k}^j)^{-1}),$$

for $2 \le k \le K$, and

$$T_1 = A_1^{-1} (R_1^K)^t A_K,$$

where for convenience the same notation is used for operators and matrices. Hence it follows from (4.12) that

$$\mathcal{P}_K = \left(R_1^K A_1^{-1} (R_1^K)^t + \sum_{k=2}^{K} R_k^K S_k (R_k^K)^t \right) A_K \equiv C_K A_K,$$

which, together with the definition of $R_k^K = I_K \cdots I_{k+1}$, leads to the typical recursive structure for the preconditioner C_K

(4.13) $\qquad C_k = I_k C_{k-1} I_k^t + S_k, \quad k = K, \dots, 2, \quad S_1 = C_1 \equiv A_1^{-1}.$

Note that with these choices for S_k, the multiplication of a vector by C_K is formally a special case of Algorithm 3.1 if one sets $m(k) = 1$, $p = 1$, removes the post-smoothing step, and replaces A_k by a zero matrix for all $k \ge 2$.

From (4.13) and the definitions of I_k and S_k, we see that a multiplication by C_K only involves $\mathrm{O}(n_K + \ldots + n_2 + n_1^3) = \mathrm{O}(n_K)$ arithmetical operations, where $n_k \approx 2^{2k}$ is the dimension of V_k. This, together with (4.4), yields suboptimal work estimates for a preconditioned conjugate gradient method for (3.4) with the preconditioner C_K. That is, an error reduction by a factor ϵ in the preconditioned conjugate gradient algorithm can be achieved by $\mathrm{O}(n_K \sqrt{\log n_K} \log(\epsilon^{-1}))$ operations.

We now turn to the discussion of the algorithmical consequences for the splitting (4.2). To do this, we need to construct basis functions in W_k, $k = 2, \dots, K$. Starting with the bases $\{\phi_{\alpha,k}^j\}$ in V_k, to each interior edge $e_{\beta,k-1}^j \in \partial\mathcal{E}_{k-1}$, we replace the two associated basis functions $\phi_{2\beta,k}^j, \phi_{2\beta+e^j,k}^j$ with their linear combinations

$$\psi_{2\beta,k}^j = \phi_{2\beta,k}^j + \phi_{2\beta+e^j,k}^j, \quad \psi_{2\beta+e^j,k}^j = \phi_{2\beta,k}^j - \phi_{2\beta+e^j,k}^j, \quad j = 1,2,$$

where $e_{2\beta,k}^j$ and $e_{2\beta+e^j,k}^j \in \partial\mathcal{E}_k$ form the edge $e_{\beta,k-1}^j$. For all other interior edges $e_{\alpha,k}^j$, which do not belong to any edge in $\partial\mathcal{E}_{k-1}$, we set

$$\psi_{\alpha,k}^j = \phi_{\alpha,k}^j.$$

The new bases $\{\psi_{\alpha,k}^j\}$ in V_k are still L^2-stable; i.e., they satisfy an inequality analogous to (4.6). Moreover, if

$$v = \sum_{j=1}^{2} \sum_{\alpha} b_\alpha^j \psi_{\alpha,k}^j,$$

we have

$$P_{k-1} v = \sum_{j=1}^{2} \sum_{\beta} b_{2\beta}^j \phi_{\beta,k-1}^j,$$

and

$$(\mathrm{Id}_k - I_k P_{k-1})v = \sum_{j=1}^{2} \sum_{\alpha \neq 2\beta} c_\alpha^j \psi_{\alpha,k}^j,$$

since $\psi_{2\beta,k}^j - I_k \phi_{\beta,k-1}^j$ can be completely expressed by the functions $\psi_{\alpha,k}^l$ with $\alpha \neq 2\beta$ only. More precisely, we have

$$c_{2\beta+e^1}^1 = b_{2\beta+e^1}^1 - \tfrac{1}{8}(b_{2\beta}^2 + b_{2(\beta-e^2)}^2 - b_{2(\beta+e^1)}^2 - b_{2(\beta+e^1-e^2)}^2),$$
$$c_{2\beta+e^2}^1 = b_{2\beta+e^2}^1 - \tfrac{1}{8}(5b_{2\beta}^2 + b_{2\beta}^1 + b_{2(\beta+e^1)}^2 + b_{2(\beta+e^2)}^1),$$
$$c_{2\beta+e^1+e^2}^1 = b_{2\beta+e^2}^1 - \tfrac{1}{8}(5b_{2(\beta+e^1)}^2 + b_{2\beta}^1 + b_{2\beta}^2 + b_{2(\beta+e^2)}^1),$$

and similar relations hold for $j = 2$. Hence any function from W_k has a unique representation by linear combinations of $\{\psi_{\alpha,k}^j : \alpha \neq 2\beta\}$, and this basis system is L^2-stable. With this basis system, as in (4.11), we have the corresponding splitting

$$(4.14) \qquad \{V_K; a_K(\cdot,\cdot)\} = R_1^K\{V_1; a_1(\cdot,\cdot)\} + \sum_{k=2}^{K} \sum_{j=1}^{2} \sum_{\alpha \neq 2\beta} R_k^K\{W_{\alpha,k}^j; a_k(\cdot,\cdot)\}$$

into a direct sum of $R_1^K V_1$ and one-dimensional spaces $R_k^K W_{\alpha,k}^j$ induced by the basis functions $\psi_{\alpha,k}^j$. Then, with the same argument as for (4.13), we derive an additive preconditioner $\hat{C}_K$ for A_K recursively defined by

$$(4.15) \qquad \hat{C}_k = I_k \hat{C}_{k-1} I_k^t + \hat{I}_k \hat{S}_k \hat{I}_k^t, \quad k = K, \dots, 2, \quad \hat{C}_1 = \hat{S}_1 \equiv A_1^{-1},$$

where

$$\hat{S}_k = \mathrm{diag}\left(a_k(\psi_{\alpha,k}^j, \psi_{\alpha,k}^j)^{-1}, \alpha \neq 2\beta, j = 1, 2\right)$$

are diagonal matrices and $\hat{I}_k$ is the rectangular matrix corresponding to the natural embedding $W_k \subset V_k$ with respect to the bases $\{\psi_{\alpha,k}^j\}$ in W_k and $\{\phi_{\alpha,k}^j\}$ in V_k (one may use the bases $\{\psi_{\alpha,k}^j\}$ for all V_k, which would change the I_k representations, but keep $\hat{I}_k$ maximally simple). (4.15) has the same arithmetical complexity as before.

We now summarize the results in Theorem 4.1 and the above discussion in the next theorem.

Theorem 4.2. *Let Ω and $\{\mathcal{E}_k\}$ satisfy the above assumptions. Then the symmetric preconditioners C_K and $\hat{C}_K$ defined in (4.13) and (4.15) and associated with the multilevel splittings (4.11) and (4.14), respectively, have an $\mathrm{O}(n_K)$ operation count per matrix-vector multiplication and produce the following condition numbers:*

$$(4.16) \qquad \kappa(C_K A_K) \leq CK, \quad \kappa(\hat{C}_K A_K) \leq CK, \quad K \geq 1.$$

The splitting (4.11) can be viewed as the nodal basis preconditioner of BPX type [5], while the splitting (4.14) is analogous to the hierarchical basis preconditioner.

We now consider multiplicative algorithms for (3.4). One iteration step of a multiplicative algorithm corresponding to the splitting (4.11) takes the form

$$
\begin{aligned}
y^0 &= x_K^j, \\
(4.17) \qquad y^{l+1} &= y^l - \omega R_{K-l}^K S_{K-l} (R_{K-l}^K)^t (A_K y^l - f_K), \quad l = 0, \dots, K-1, \\
x_K^{j+1} &= y^K,
\end{aligned}
$$

where ω is a suitable relaxation parameter (the range of relaxation parameters for which the algorithm in (4.17) converges is determined mainly by the constant in the inverse inequality (2.2) [30, 28, 16]. The method (4.17) corresponds to a $\mathcal{V}$-cycle algorithm in Algorithm 3.1 with A_k replaced by $\tilde{A}_k = (R_k^K)^t A_K R_k^K$, one pre-smoothing and no post-smoothing steps.

The iteration matrix $M_{K,\omega}$ in (4.17) is given by

$$M_{K,\omega} = (\mathrm{Id}_K - \omega E_1) \cdots (\mathrm{Id}_K - \omega E_{K-1})(\mathrm{Id}_K - \omega E_K) , \quad E_k \equiv R_k^K S_k (R_k^K)^t A_K.$$

An analogous multiplicative algorithm for (3.4) corresponding to the splitting (4.14) can be defined.

From the general theory on multiplicative algorithms [30], [16], and by the same argument as for Theorem 4.2, we can show the following result.

Theorem 4.3. *Let Ω and $\{\mathcal{E}_k\}$ satisfy the above assumptions. For properly chosen relaxation parameter ω the multiplicative schemes corresponding to the splittings (4.11) and (4.14) possess the following upper bounds for the convergence rate:*

$$(4.18) \qquad \inf_{\omega} ||M_{K,\omega}||_{\mathcal{E}} \le 1 - \frac{C}{K}, \quad \inf_{\omega} ||\hat{M}_{K,\omega}||_{\mathcal{E}} \le 1 - \frac{C}{K}, \quad K \to \infty,$$

where $M_{K,\omega}$ and $\hat{M}_{K,\omega}$ denote the iteration matrices associated with (4.11) and (4.14), respectively.

We end with two remarks. First, one example for the choice of ω is that $\omega \approx K^{-1}$, which leads to the upper bounds in (4.18). Second, the diagonal matrices S_k and $\hat{S}_k$ in (4.13) and (4.15) can be replaced by any other spectrally equivalent symmetric matrices of their respective dimension.

5. Equivalent discretizations

As an alternative to the preconditioners described in Section 4 for which the estimates in Theorems 4.2 and 4.3 guarantee only suboptimal convergence rates, we propose now to switch from the NR Q_1 discretization (3.4) to a spectrally equivalent discretization for which optimal preconditioners are already available; see [22] for references and examples for other conforming elements. The most natural candidate for a switching procedure is the space of conforming bilinear elements

$$U_K = \left\{ \xi \in C^0(\overline{\Omega}) : \xi|_E \in Q_1(E), \ \forall E \in \mathcal{E}_k \text{ and } \xi|_\Gamma = 0 \right\},$$

on the same partition. For simplicity, we again assume that Ω is the unit square, and that the $\mathcal{E}_k$ are uniform square partitions. However, it is easy to realize that a switching procedure can be implemented also in the general case if, e.g., triangular linear elements are used as reference elements.

We introduce two linear operators $Y_K : U_K \to V_K$ and $\hat{Y}_K : V_K \to U_K$ as follows. If $\xi \in U_K$ and e is an edge of an element in $\mathcal{E}_K$, then $Y_K \xi \in V_K$ is given by

$$(5.1) \qquad\qquad \int_e Y_K \xi \, \mathrm{d}s = \int_e \xi \, \mathrm{d}s,$$

which preserves the zero average values on the boundary edges. If $v \in V_K$, we define $\hat{Y}_K v \in U_K$ by

$$(5.2) \qquad \begin{aligned} &(\hat{Y}_K v)(z) = 0 \quad \text{for all boundary vertices } z \text{ in } \mathcal{E}_K, \\ &(\hat{Y}_K v)(z) = \text{average of } v_j(z) \quad \text{for all internal vertices } z \text{ in } \mathcal{E}_K, \end{aligned}$$

where $v_j = v|_{E_j}$ and $E_j \in \mathcal{E}_K$ contains z as a vertex.

Another choice for U_K is the space of conforming P_1 elements

$$U_K = \left\{ \xi \in C^0(\overline{\Omega}) : \xi|_E \in P_1(E),\ \forall E \in \tilde{\mathcal{E}}_K \text{ and } \xi|_\Gamma = 0 \right\},$$

where $\tilde{\mathcal{E}}_K$ is the triangulation of Ω generated by connecting the two opposite vertices of the squares in $\mathcal{E}_K$. The two linear operators $Y_K : U_K \to V_K$ and $\hat{Y}_K : V_K \to U_K$ are defined as in (5.1) and (5.2), respectively. Moreover, for both the conforming bilinear elements and the conforming P_1 elements, it can be easily shown that there is a constant C, independent of K, such that

(5.3)
$$2^K \|\xi - Y_K\xi\| \leq C\|\xi\|_\varepsilon, \qquad \forall \xi \in U_K,$$
$$2^K \|v - \hat{Y}_K v\| \leq C\|v\|_\varepsilon, \qquad \forall v \in V_K.$$

Since optimal preconditioners exist for the discretization system $\overline{A}_K$ generated by the conforming bilinear elements (respectively, the conforming P_1 elements), the next result follows from (5.3) and the general switching theory in [22].

Theorem 5.1. *Let $\overline{C}_K$ be any optimal symmetric preconditioner for $\overline{A}_K$; i.e., we assume that a matrix-vector multiplication by $\overline{C}_K$ can be performed in $\mathrm{O}(n_K)$ arithmetical operations, and that $\kappa(\overline{C}_K\overline{A}_K) \leq C$, with constant independent of K. Let $S_K = 2^{2K}\mathrm{Id}_K$ (or $S_K = \mathrm{diag}(A_K)$ or any other spectrally equivalent symmetric matrix). Then*

(5.4)
$$\overline{C}_K^* = S_K + Y_K\overline{C}_K(Y_K)^t$$

is an optimal symmetric preconditioner for A_K.

6. Numerical experiments

In this section we present the results of numerical examples to illustrate the theories developed in the earlier sections. These numerical examples deal with the Laplace equation on the unit square:

(6.1)
$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega = (0,1)^2, \\ u &= 0 \quad \text{on } \Gamma, \end{aligned}$$

where $f \in L^2$. The NR Q_1 finite element method (3.4) is used to solve (6.1) with $\{\mathcal{E}_k\}_{k=1}^K$ being a sequence of dyadically, uniformly refined partitions of Ω into squares. The coarsest grid is of size $h_1 = 1/2$.

The first test concerns the convergence of Algorithm 3.1. The analysis of the third section guarantees the convergence of the $\mathcal{W}$-cycle algorithm with any number of smoothing steps and the uniform condition number property for the variable $\mathcal{V}$-cycle algorithm, but does not give any indication for the convergence of the standard $\mathcal{V}$-cycle algorithm, i.e., Algorithm 3.1 with $p = 1$ and $m(k) = 1$ for all k. The first two rows of Table 1 show the results for levels $K = 3, \ldots, 7$ for this symmetric V-cycle, where (κ_v, δ_v) denote the condition number for the system $B_K A_K$ and the reduction factor for the system $\mathrm{Id}_K - B_K A_K$ as a function of the mesh size on the finest grid h_K. While there is no complete theory for this $\mathcal{V}$-cycle algorithm, it is of practical interest that the condition numbers for this cycle remain relatively small.

For comparison, we run the same example by a symmetrized multilevel multiplicative Schwarz method corresponding to (4.17). One step of the symmetric version consists of two substeps, the first coinciding with (4.17) and the second repeating (4.17) in reverse order. The condition numbers κ_m for $\tilde{M}_{K,\omega} A_K$ with

TABLE 1. Numerical results for the multiplicative $\mathcal{V}$-cycles.

$1/h_K$	8	16	32	64	128
κ_v	1.54	1.70	1.84	1.96	2.06
δ_v	0.23	0.27	0.32	0.33	0.35
κ_m	1.75	1.81	1.84	1.85	1.85

$\omega \approx K^{-1}$ are presented in the third row of Table 1, where $\tilde{M}_{K,\omega} = M_{K,\omega}^t M_{K,\omega}$ is now symmetric. The results are better than expected from the upper bounds of Theorem 4.3 which seem to be only suboptimal.

In the second test we treat the above multigrid algorithm and symmetrized multilevel multiplicative method as preconditioners for the conjugate gradient method. In this test the problem (6.1) is assumed to have the exact solution

$$u(x,y) = x(1-x)y(1-y)e^{xy}.$$

Table 2 shows the number of iterations required to achieve the error reduction 10^{-6}, where the starting vector for the iteration is zero. The iteration numbers $(\text{iter}_v, \text{iter}_m)$ correspond to Algorithm 3.1 with $p = 1$ and $m(k) = 1$ for all k and the symmetrized multiplicative algorithm (4.17), respectively. Note that iter_v and iter_m remain almost constant when the step size increases.

TABLE 2. Iteration numbers for the pcg-iteration.

$1/h_K$	8	16	32	64	128
iter_v	8	8	9	9	10
iter_m	9	9	9	10	10

In the final test we report analogous numerical results (condition numbers and pcg-iteration count) for the additive preconditioner C_K associated with the splitting (4.11) (subscript a), and the preconditioner $\overline{C}_K^*$ (subscript s) which uses the switch from the system arising from (3.4) to the spectrally equivalent system generated by the conforming bilinear elements via the operators in (5.1) and (5.2). We have implemented the standard BPX-preconditioner [5], with diagonal scaling, as $\overline{C}_K$. These results are shown in Table 3. The numbers show the slight growth, which

TABLE 3. Results for the preconditioners C_K and $\overline{C}_K^*$.

$1/h_K$	8	16	32	64	128	256	512
κ_a	9.6	12.3	14.4	16.1	17.4	18.3	19.3
iter_a	18	22	24	26	27	28	28
κ_s	3.37	3.87	4.24	4.54	4.80	5.05	-
iter_s	10	11	13	13	14	15	-

is typical for most of the additive preconditioners and level numbers $K < 10$. The condition numbers κ_s for the switching procedure are practically identical to the condition numbers for $\overline{C}_K \overline{A}_K$ characterizing the BPX-preconditioner [5] in the conforming bilinear case. The switching procedure is clearly favorable as can be expected from the theoretical bounds of Theorems 4.2 and 5.1; however, the computations do not indicate whether the upper bound (4.16) is sharp or could be further improved.

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), 943–972. MR **95k:**65102
[2] R. Bank and T. Dupont, *An optimal order process for solving finite element equations*, Math. Comp. **36** (1981), 35–51. MR **82b:**65113
[3] D. Braess and R. Verfürth, *Multigrid methods for nonconforming finite element methods*, SIAM J. Numer. Anal. **27** (1990), 979–986. MR **91j:**65164
[4] J. Bramble, *Multigrid Methods*, Pitman Research Notes in Math., vol. 294, Longman, London, 1993. MR **95b:**65002
[5] J. Bramble, J. Pasciak, and J. Xu, *Parallel multilevel preconditioners*, Math. Comp. **55** (1991), 1-22. MR **90k:**65170
[6] 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. MR **91h:**65159
[7] S. Brenner, *An optimal-order multigrid method for P1 nonconforming finite elements*, Math. Comp. **52** (1989), 1–15. MR **89f:**65119
[8] S. Brenner, *Multigrid methods for nonconforming finite elements*, Proceedings of Fourth Copper Mountain Conference on Multigrid Methods, J. Mandel, et al., eds., SIAM, Philadelphia, 1989, pp. 54–65. MR **91h:**65189
[9] S. Brenner, *Convergence of nonconforming multigrid methods without full elliptic regularity*, Preprint, 1995, submitted.
[10] Zhangxin Chen, *Analysis of mixed methods using conforming and nonconforming finite element methods*, RAIRO Modèl. Math. Anal. Numer. **27** (1993), 9–34. MR **94c:**65132
[11] Zhangxin Chen, *Projection finite element methods for semiconductor device equations*, Computers Math. Applic. **25** (1993), 81–88. MR **93k:**65092

[12] Zhangxin Chen, *Equivalence between and multigrid algorithms for nonconforming and mixed methods for second order elliptic problems*, East-West J. Numer. Math. **4** (1996), 1–33. CMP 96:13

[13] Zhangxin Chen, R. E. Ewing, Y. Kuznetsov, R. Lazarov, and S. Maliassov, *Multilevel preconditioners for mixed methods for second order elliptic problems*, J. Numer. Lin. Alg. Appl. **30** (1996), 427–453. CMP 97:04

[14] Zhangxin Chen, R. Ewing, and R. Lazarov, *Domain decomposition algorithms for mixed methods for second order elliptic problems*, Math. Comp. **65** (1996), 467–490. MR **96g:**65117

[15] Zhangxin Chen, D. Y. Kwak, and Y. J. Yon, *Multigrid algorithms for nonconforming and mixed methods for nonsymmetric and indefinite problems*, IMA Preprint Series #1277, 1994, SIAM J. Scientific Computing, 1998 to appear.

[16] M. Griebel and P. Oswald, *On the abstract theory of additive and multiplicative Schwarz algorithms*, Numer. Math. **70** (1995), 163–180. MR **96a:**65164

[17] P. Kloucek, B. Li, and M. Luskin, *Analysis of a class of nonconforming finite elements for crystalline microstructures*, Math. Comp. **65** (1996), 1111–1135. MR **97a:**73076

[18] P. Kloucek and M. Luskin, *The computation of the dynamics of martensitic microstructure*, Continuum Mech. Thermodyn. **6** (1994), 209–240. MR **95d:**73009

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

[20] P. Oswald, *On a hierarchical basis multilevel method with nonconforming P1 elements*, Numer. Math. **62** (1992), 189–212. MR **93b:**65059

[21] P. Oswald, *Multilevel Finite Element Approximation : Theory and Application*, Teubner Skripten zur Numerik, Teubner, Stuttgart, 1994. MR **95k:**65110

[22] P. Oswald, *Preconditioners for nonconforming discritizations*, Math. Comp. **65** (1996), 923–941. MR **96j:**65056

[23] P. Oswald, *Intergrid transfer operators and multilevel preconditioners for nonconforming discretizations*, Appl. Numer. Math. **23** (1997), 139–158. CMP 97:09

[24] R. Rannacher and S. Turek, *Simple nonconforming quadrilateral Stokes element*, Numer. Meth. Partial Diff. Equations **8** (1992), 97–111. MR **92i:**65170

[25] P. Raviart and J. Thomas, *A mixed finite element method for second order elliptic problems*, Mathematical aspects of the FEM, Lecture Notes in Mathematics, **606**, Springer-Verlag, Berlin & New York (1977), pp. 292–315. MR **58:**3547

[26] S. Turek, *Multigrid techniques for a divergence-free finite element discretization*, East-West J. Numer. Math. **2** (1994), 229-255. MR **96c:**65195

[27] M. Wang, *The W-cycle multigrid method for finite elements with nonnested spaces*, Adv. in Math. **23** (1994), 238–250. MR **95e:**65118

[28] H. Yserentant, *Old and new convergence proofs for multigrid methods*, Acta Numerica, Cambr. Univ. Press, Cambridge, 1993, pp. 285–236. MR **94i:**65128

[29] J. Xu, *Convergence estimates for some multigrid algorithms*, Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. F. Chan et al., eds., SIAM, Philadelphia, 1990, pp. 174–187. MR **91f:**65194

[30] J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Review **34** (1992), 581–613. MR **93k:**65029

DEPARTMENT OF MATHEMATICS, BOX 156, SOUTHERN METHODIST UNIVERSITY, DALLAS, TEXAS 75275–0156
E-mail address: zchen@dragon.math.smu.edu

INSTITUTE OF ALGORITHMS AND SCIENTIFIC COMPUTING, GMD - GERMAN NATIONAL RESEARCH CENTER FOR INFORMATION TECHNOLOGY, SCHLOSS BIRLINGHOVEN, D-53754 SANKT AUGUSTIN, GERMANY
E-mail address: peter.oswald@gmd.de