[4] S. Brenner, An optimal-order nonconforming multigrid method for the biharmonic equation, *SIAM J. Numer. Anal.* **26** (1989) 1124–1138.

[5] S. Brenner, Multigrid methods for nonconforming finite elements, in: J. Mandel et al., eds., *Proceedings Fourth Copper Mountain Conference on Multigrid Methods* (SIAM, Philadelphia, 1989) 54–65.

[6] S. Brenner, Two-level additive Schwarz preconditioners for nonconforming finite elements, in: D. E. Keyes, J. Xu, eds., *Proceedings Seventh International Conference on Domain Decomposition* (Contemporary Mathematics, vol. 180, AMS, Providence, 1995) 9–14.

[7] P. Ciarlet, *The finite element method for elliptic problems* (North-Holland, New York, 1978).

[8] Z. Chen, P. Oswald, Multigrid and multilevel methods for nonconforming rotated Q1 elements, *Math. Comp.* (submitted).

[9] L. C. Cowsar, Domain decomposition methods for nonconforming finite element spaces of Lagrange type, in: N. D. Melson et al., eds., *Proceedings Sixth Copper Mountain Conference on Multigrid Methods* (NASA Conference Publications, vol. 3224, 1993) 93–109.

[10] W. Dörfler, Hierarchical bases for elliptic problems, *Math. Comp.* **58** (1992) 513–529.

[11] W. Dörfler, A note on the preconditioner by Bramble, Pasciak, and Xu, *Preprint* (Inst. für Angew. Math., Univ. Zürich, 1991).

[12] W. Hackbusch, *Iterative solution of large sparse systems of equations* (Springer, New York, 1994).

[13] M. R. Hanisch, Multigrid preconditioning for the biharmonic Dirichlet problem, *SIAM J. Numer. Anal.* **30** (1993) 184–214.

[14] M. E. G. Ong, Uniform refinement of a tetrahedron, *SIAM J. Sci. Comput.* **15** (1994) 1134-1144.

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

[16] P. Oswald, *Multilevel finite element approximation : theory and application* (Teubner Skripten zur Numerik, Teubner, Stuttgart, 1994).

[17] P. Oswald, Preconditioners for nonconforming elements, *Math. Comp.* (1996) (to appear).

[18] M. Sarkis, Two-level Schwarz methods for nonconforming finite elements and discontinuous coefficients, in: N.D.Melson et al., eds., *Proceedings Sixth Copper Mountain Conference on Multigrid Methods* (NASA Conference Publications, vol. 3224, 1993) 543–565.

[19] J. Xu, Iterative methods by space decomposition and subspace correction, *SIAM Review* **34** (1992) 581–613.

Table 3
Results for the Morley element

	type (a)		type (b)	
J	τ_J	τ_J/τ_{J-1}	τ_J	τ_J/τ_{J-1}
1	1.88	1.88	1.41	1.41
2	3.33	1.77	1.92	1.37
3	6.05	1.81	2.83	1.47
4	11.68	1.93	4.64	1.67
5	23.90	2.05	8.85	1.88
6	51.01	2.13	18.09	2.05

Table 4
Results for the Zienkiewicz element

	type (a)		type (b)	
J	τ_J	τ_J/τ_{J-1}	τ_J	τ_J/τ_{J-1}
1	1.550	1.550	1.174	1.174
2	1.882	1.214	1.309	1.115
3	2.044	1.086	1.379	1.054
4	2.115	1.035	1.411	1.023
5	2.145	1.014	1.425	1.010
6	2.157	1.006	1.430	1.004

O(J^3)-behavior (or better) for the condition numbers of the hierarchical resp. BPX-type preconditioners associated with the sets $\{P_j\}, \{R_j\}$ as described above. However, this requires further investigation.

References

[1] J. H. Bramble, *Multigrid methods* (Pitman Research Notes in Mathematical Sciences 294, Longman Sci.&Techn., Harlow, Essex, 1993).

[2] J. H. Bramble, J. E. Pasciak and J. Xu, Parallel multilevel preconditioners, *Math. Comp.* **55** (1990) 1–22.

[3] J. H. Bramble, J. E. Pasciak and J. Xu, The analysis of multigrid algorithms with nonnested spaces or noniherited quadratic forms, *Math. Comp.* **56** (1991) 1–34.

Table 1
Results for the P1 element

	type (a)		type (b)	
J	τ_J	τ_J/τ_{J-1}	τ_J	τ_J/τ_{J-1}
1	1.250	1.25	1.625	1.63
2	1.422	1.14	2.016	1.24
3	1.535	1.08	2.277	1.13
4	1.612	1.05	2.453	1.08
5	1.664	1.03	2.571	1.05
6	1.700	1.02	2.650	1.03

Table 2
Results for the rotated Q1 element

	type (a)		type (b)	
J	τ_J	τ_J/τ_{J-1}	τ_J	τ_J/τ_{J-1}
1	1.66	1.66	1.13	1.13
2	3.06	1.84	1.23	1.09
3	5.40	1.76	1.29	1.05
4	9.43	1.75	1.32	1.03
5	16.58	1.76	1.34	1.01
6	29.36	1.77	1.35	1.01

the definition for the P1 element but lead to different results). In both cases, there is only one generic nodal basis function $N_{0,i}$. The results show that (b) is potentially better for constructing multilevel preconditioners. Details on this particular element can be found in [8].

Morley element. The energy scalar product is as in (38). On the three-directional mesh, there are only three generic basis functions: $N_{0,i}$ associated (a) with vertices (function value interpolation), (b) with edges in diagonal direction resp. (c) with edges parallel to the axes (normal derivative interpolation). We provide numerical values for (a) and (b). These indicate exponential growth of order $\asymp 2^J$, at least in the range interesting for computations.

Zienkiewicz element. The relevant basis functions are associated with (a) function value interpolation or (b) interpolation of the partial derivative in x_1- or x_2-direction at the vertices. In view of the discussion at the end of subsection 3.1, the boundedness of the values τ_J makes it likely to expect an

$\{\tau^J_{j,i} \equiv \|P_J \dots P_{j+1} N_{j,i}\|^2_E / (\alpha_j \|N_{j,i}\|^2_H)\}$, *since* j *is small one could have neglected the denominator as well. In the typical examples, if at least one of the sequences* $\{\tau^J_{j,i}\ ,\ J > j\}$ *computed for the selected set of basis functions is geometrically growing then so do the lower and upper bounds for the condition numbers in Theorem 4. Analogous statements hold for Theorem 6.*

Concluding the paper we present some numerical experiments. We have computed the first few values of

$$\tau_J \equiv \frac{\|P_J \dots P_1 N_{0,i}\|^2_E}{\|N_{0,i}\|^2_E}\ , \qquad J \geq 1\ , \tag{44}$$

for the typical basis functions for some low order elements on uniform grids. The partition $\mathcal{T}_0$ was a three-directional uniform triangulation of sidelength 1 for triangular elements resp. a partition into squares of sidelength 1, the domain Ω was a large enough square and $N_{0,i}$ was chosen such that no boundary effects could come in (the generic interior basis functions). Experiments with nodal basis functions close to the boundary showed quite analogous or even better behavior and are not reported on here. In all cases, the prolongation operators $\{P_j\}$ are defined by nodal value averaging procedures as explained above.

Under these circumstances, according to Lemma 9, the values τ_J provide reasonable insight into the expected behavior of $\{\tau^*_k\}$ and, thus, of the condition numbers for the preconditioners described in subsection 2.2. E.g., exponential growth of any of these sequences indicates exponential growth of the condition numbers as well.

Experiments of this type can help to select elements and intergrid transfer operators suitable for preconditioning. They show also for which situations we should expect negative results. Difficulties with V-cycle methods for the Morley element seem to be known ([13], however, this paper discusses only multigrid (multiplicative) preconditioners and methods). The numbers for the Zienkiewicz element are much more encouraging which makes this element interesting for preconditioning finite element discretization matrices for fourth order problems.

Triangular P1 element. The two generic basis functions $N_{0,i}$ are associated with (a) edges in diagonal direction and (b) edges parallel to the axes. For this and the following example, the energy scalar product is defined by (37).

Rectangular rotated Q1 element. The element can be characterized by the local polynomial space span$\{1, x_1, x_2, x_1^2 - x_2^2\}$ with the local interpolation problems composed of (a) function value interpolation at the edge midpoints or, alternatively, (b) integral averages along the edges (both choices extend positive constant C, this would lead to

$$\frac{\|P_{j'+k}\dots P_{j'+1}N_{j',i'}\|_E^2}{\alpha_{j'}\|N_{j',i'}\|_H^2} = \frac{\|P_{j+k}\dots P_{j+1}N_{j,i}\|_E^2}{\alpha_j\|N_{j,i}\|_H^2} , \quad k \geq 1 , \tag{41}$$

whenever the two sequences are translation/dilation invariant. The proof of the following lemma shows that (41) could be further relaxed but we do not wish to complicate the exposition.

Lemma 9 *Let the above assumptions of a translation/dilation-invariant setting hold. Set*

$$\tau_k^* = \max \frac{\|P_{j+k}\dots P_{j+1}N_{j,i}\|_E^2}{\alpha_j\|N_{j,i}\|_H^2} , \quad \tau_0^* = 1 , \tag{42}$$

where the maximum is taken over the finite set of selected basis functions mentioned above. Then

$$\tau_j^J \leq C\tau_{J-j}^* , \qquad \sum_{j=0}^J \tau_j^J \leq C\sum_{k=0}^J \tau_k^* . \tag{43}$$

If in addition (27) holds then in the definition of τ_k^* *the denominators can be replaced by* $\|N_{j,i}\|_E^2$.

PROOF. For fixed $J > j$, consider any $u_j = \sum_i x_{j,i}N_{j,i} \in V_j$ and

$$\tilde{P}_j u_j = \sum_i x_{j,i}(P_J\dots P_{j+1}N_{j,i}) \equiv \sum_i x_{j,i}u_{j,i}^J .$$

According to the above assumptions on $\{P_j\}$, the set $\{\operatorname{supp} u_{j,i}^J\}$ has locally finite overlap, hence

$$\|\tilde{P}_j u_j\|_E^2 \leq C\sum_i x_{j,i}^2\|u_{j,i}^J\|_E^2 .$$

For each index pair in this finite sum, there is a match in the fixed set from the assumptions and a brief inspection of (41), (42) gives $\|u_{j,i}^J\|_E^2 \leq \alpha_j\tau_{J-j}^*\|N_{j,i}\|_H^2$. We can finish the argument by using the upper bound of (24), and looking at the definition (10) of the τ_j^J.

Remark 10 *If one uses Lemma 9 in connection with Theorem 4 then the second inequality in (43) serves the upper condition number estimate. Lower estimates can be obtained by looking at the growth of each of the sequences* the corresponding R_{j-1}:

$$\begin{aligned}\|P_j u_{j-1}\|_E^2 &= \|R_{j-1}(P_j u_{j-1}) - (Id - R_{j-1})(P_j u_{j-1})\|_E^2 \\ &= \|u_{j-1}\|_E^2 + \|P_j u_{j-1} - u_{j-1}\|_E^2 ,\end{aligned}$$

now take any function u_{j-1} of local support such that the last term does not vanish.

Thus, one can not expect to handle the estimation of $\|\tilde{P}_j\|_E$ by assumptions on $\|P_j\|_E$ since this would lead to too pessimistic results, compare [10] for attempts in this direction. Experience shows that in many cases $\|\tilde{P}_j\|_E$ is much smaller than the corresponding product $\|P_{j+1}\|_E \cdots \|P_J\|_E$. E.g., Lemma 5 of [15] states for the two-dimensional P1 element and the above choice of $P_j = P_j^{ext}|_{V_{j-1}}$ that

$$\tau_j^J = \mathrm{O}(1) , \quad j, J \to \infty . \tag{40}$$

under certain conditions on $\mathcal{T}_0$. For the rotated Q1 element, analogous results have recently been proved in [8]. However, even for dyadically refined uniform grids of $\mathbf{R}^d$, in a shift-invariant setting, the investigation of the energy norms of the iterated prolongation operators $\{\tilde{P}_j\}$ by means of Fourier analysis tools leads to rather nonstandard, unsolved questions. Instead of going this direction, we propose the following, experimental approach.

We assume again that the bases $\{N_{j,i}\}$ are local and H-stable (24). Suppose that the operators P_j act in a local and translation/dilation-invariant fashion, i.e., the support of $P_{j+1}N_{j,i}$ is contained in the set $K_{j,i}$ given as the the union of all cells in $\mathcal{T}_j$ intersecting or touching the support of the basis function $N_{j,i}$, and if $K_{j'-1,i'}$ is obtained from $K_{j-1,i}$ by a coordinate transformation T containing translation and dilation, then we have $N_{j',i'}(T\cdot) = N_{j,i}(\cdot)$ and $(P_{j'+1}N_{j',i'})(T\cdot) = (P_{j+1}N_{j,i})(\cdot)$. These assumptions are fulfilled for all finite element examples we have checked so far (note that in some situations one could even require full affine-invariance). Thus, if we start with a finite partition $\mathcal{T}_0$ ($V_0 \neq \{0\}$) and use regular dyadic refinement only, we can find a set of finitely many $N_{j,i}$ such that each sequence $\{P_{j'+k} \ldots P_{j'+1}N_{j',i'},\ k \geq 1\}$ is translation/dilation equivalent to the same sequence corresponding to one of the selected basis functions (it is easy to check that in general a selection of all basis functions of levels $j \leq 3$ into this set would do, for the Morley resp. P1 element, even $j \leq 2$ would suffice). Finally, suppose that the norms $\|\cdot\|_E^2$ and $\|\cdot\|_H^2$ are additive with respect to the underlying domain and translation/dilation invariant, with the obvious scaling involved. For the choices stated in (37) resp. (38), together with $\alpha_j = C2^{2j}$ resp. $\alpha_j = C2^{4j}$ for some

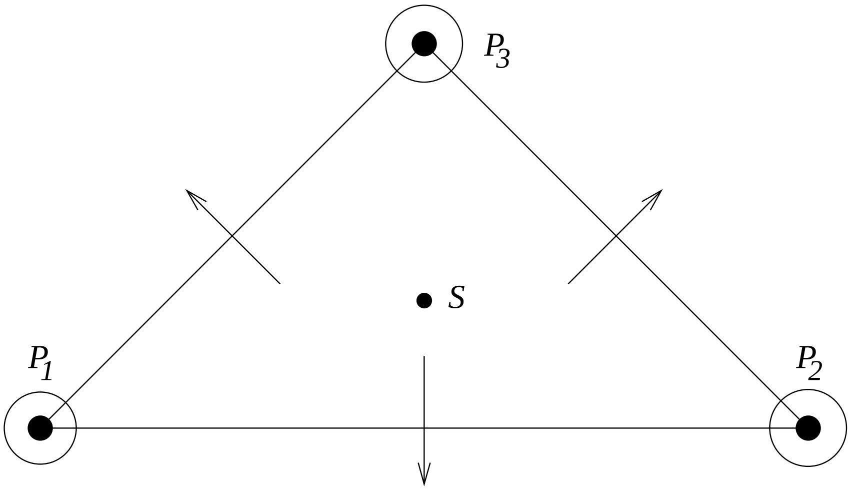

Fig. 4. Zienkiewicz element

This construction guarantees that quadratic functions are exactly reproduced. Since on a dyadically refined $\{\mathcal{T}_j\}$ the interpolation conditions are preserved with increasing j , a simple candidate for the restriction R_{j-1} is the nodal interpolation operator with respect to $\mathcal{T}_{j-1}$ while $\{P_j^{ext}\}$ can be defined again by averaging (actually, due to the properties of this element, only at edge midpoints M_e of edges in $\mathcal{T}_{j-1}$ averaging is necessary for the derivative in the direction n_e, all other interpolation values associated with $\mathcal{T}_j$ are well-defined for functions from $V_{j-1} + V_j$). We set $P_j = P_j^{ext}|_{V_{j-1}}$. The definition of the E and H scalar products is the same as for the Morley element, see (38) resp. (37). These choices satisfy the condition (34) of Lemma 7 as well as (7) and (9), where $\alpha_j = C2^{4j}$ is appropriate. Along the lines of [10], one can prove (35) with moderately growing constants $\hat{c}_J = \mathrm{O}(J^2)$ (this is expected since (35) corresponds to the hierarchical basis decomposition where this is the typical growth in two dimensions). Thus, for the Zienkiewicz element we get with this straightforward approach a still reasonable estimate of

$$c_J \leq C(J+1)^2 , \qquad J \geq 0 , \tag{39}$$

for the natural choice of the prolongations P_j. With a different set of $\{R_j\}$ one might further improve the estimate for c_J (compare the methods of [11]), however, we do not attempt to do this here.

3.2 Investigation of τ_j^J

Typically, the prolongation operators P_j for nonconforming elements which have been used in the literature lead to an increase of the energy norm, i.e., usually one has $\|P_j\|_E \geq c$ for some constant $c > 1$. For the P1 element (or the Morley element) this easily follows from (7) and the orthogonality property of

These are the standard choices for the Morley element (see [4,1]) which satisfy the condition (34) of Lemma 7 as well as (7) and (9), where $\alpha_j = C2^{4j}$ is appropriate. The verification of the assumptions of Lemma 8 parallels the case of the P1 element. It suffices to look at the case $j = J$. By Green's formula, if u and v are quadratic polynomials on a certain triangle, then

$$\int_{\Delta} \nabla^2 u : \nabla^2 v \;\mathrm{d}x = \sum_i \sum_{\tilde{e} \subset \bar{\Delta}_i} \frac{\partial \nabla v}{\partial n_{\tilde{e}}} \int_{\tilde{e}} \nabla u|_{\Delta_i} \;\mathrm{d}s$$

$$= \sum_{\tilde{e} \subset \mathrm{int}(\Delta)} |\tilde{e}| \left[\frac{\partial \nabla v}{\partial n_{\tilde{e}}} \right] \cdot \nabla u(M_{\tilde{e}}) + \sum_{\tilde{e} \subset \partial\Delta} |\tilde{e}| \frac{\partial \nabla v}{\partial n_{\tilde{e}}} \cdot \nabla u(M_{\tilde{e}}) \;,$$

notice that ∇u is a linear and $\frac{\partial \nabla v}{\partial n_{\tilde{e}}}$ a constant vector function along $\tilde{e}$. Now, as above apply this to $u = u_J - R_{J-1}u_J$ and $v = v_{J-1}$ on the four triangles Δ_i forming a triangle Δ from $\mathcal{T}_{J-1}$, and check that due to the definition of the elements and R_{J-1} all terms in the last sum cancel. This gives the orthogonality property of the restrictions.

To see (36), fix any Δ from $\mathcal{T}_{J-1}$. Since $R = R_{J-1}$ reproduces quadratic polynomials on Δ, we have for $u = u_J$

$$\begin{aligned} \|u - Ru\|^2_{L_2(\Delta)} &= \inf_{\deg(p) \le 1} \|(u-p) - R(u-p)\|^2_{L_2(\Delta)} \\ &\le C \inf_{\deg(p)\le 1} \|u - p\|^2_{L_2(\Delta)} \le C(\mathrm{diam}\,\Delta)^4 \|u\|^2_{E,\Delta} . \end{aligned}$$

The local L_2 boundedness of R as well as the approximation result used in the last step are left to the reader (e.g., take p such that $p = u$ on Δ_4). Now, apply this local estimate to $u - Ru$ instead of u, multiply by $\alpha_J \asymp (\mathrm{diam}\,\Delta)^{-4}$, and sum over all triangles Δ.

For more complicated, higher order nonconforming elements orthogonal projections (with respect to $(\cdot,\cdot)_E$) can not be made explicit by local considerations. Thus, their use as restriction operators is impossible. For certain types of nonconforming elements, [10,11] contain some results on proving estimates like (8) leading to the lower estimate in (13). As an example, consider the Zienkiewicz element which is used in connection with plate bending and is defined by a relatively small number of interpolation conditions per triangle. The local interpolation problem is indicated in Figure 4: on each triangle it consists of determining a reduced cubic polynomial p by interpolating function and gradient values at the three vertices and one linearly independent condition fixing the value at the barycenter:

$$p(S) = \frac{1}{3} \sum_{i=1}^{3} p(P_i) - \frac{1}{6} \sum_{i=1}^{3} \nabla p(P_i) \cdot (S - P_i) \;.$$

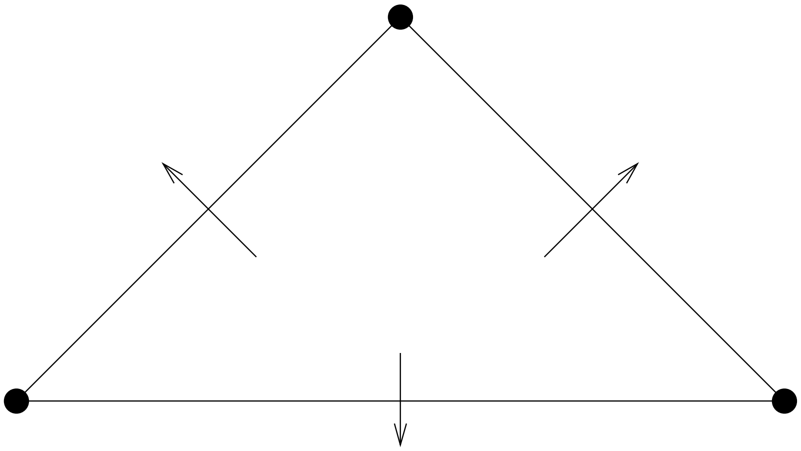

Fig. 3. Morley element

Similarly to the P1 case, the restriction R_{j-1} is defined by preserving the values of u_j at the vertices of $\mathcal{T}_{j-1}$ and by setting

$$\frac{\partial R_{j-1}u_j}{\partial n_e}(M_e) = \frac{1}{2}(\frac{\partial u_j}{\partial n_e}(M_1) + \frac{\partial u_j}{\partial n_e}(M_2))$$

for the values of the normal derivatives. Again, the prolongations $\{P_j^{ext}\}$ can be defined by averaging. Specifically, this concerns function values at the midpoints M_e and normal derivatives at M_1 and M_2 for edges e in $\mathcal{T}_{j-1}$ which may have jumps for functions $\hat{u}_j \in V_{j-1} + V_j$ while at all other points usual nodal interpolation is well-defined. E.g.,

$$(P_j^{ext}\hat{u}_j)(M_e) = \begin{cases} 0 & \text{if } e \subset \partial\Omega \\ \frac{1}{2}(\hat{u}_j(M_e+) + \hat{u}_j(M_e-)) & \text{elsewhere} \end{cases} ,$$

analogously

$$\frac{\partial P_j^{ext}\hat{u}_j}{\partial n_e}(M_e) = \begin{cases} 0 & \text{if } e \subset \partial\Omega \\ \frac{1}{2}(\frac{\partial \hat{u}_j}{\partial n_e}(M_e+) + \frac{\partial \hat{u}_j}{\partial n_e}(M_e-)) & \text{elsewhere} \end{cases} .$$

We set $P_j = P_j^{ext}|_{V_{j-1}}$. The definition of the energy product has to be changed to its counterpart for fourth–order problems

$$(u,v)_E = \sum_{\Delta \in \mathcal{T}_J} \int_{\Delta} \nabla^2 u : \nabla^2 v \; \mathrm{d}x \qquad \forall\, u, v \in \sum_{j=0}^{J} V_j \;, \tag{38}$$

while $H = L_2(\Omega)$ is preserved.

Substituting this into the above sums, we see that all terms or vanish, or cancel. Summing over all Δ from $\mathcal{T}_{J-1}$, the desired orthogonality property comes out. The three-dimensional case is completely analogous.

The inequality (36) is more or less obvious, e.g., it can be seen from observing that $p - Rp$ vanishes on Δ whenever p is a linear function on Δ (local preservation of polynomials of degree 1). By definition of R and some obvious properties of nonconforming P1-elements, this gives

$$\begin{aligned}\|u - Ru\|^2_{L_2(\Delta)} &= \inf_{\deg(p)\le 1} \|(u-p) - R(u-p)\|^2_{L_2(\Delta)} \\ &\le C \inf_{\deg(p)\le 1} \|u-p\|^2_{L_2(\Delta)} \le C(\operatorname{diam}\Delta)^2 \|u\|^2_{E,\Delta}.\end{aligned}$$

(To see the last step, simply take p to coincide with u on Δ_4. This makes the values of $u - p$ vanish at the nodal points interior to Δ, at the remaining nodal points $u - p$ can be expressed by (second) differences of u-values from neighbored nodal points which in turn can be estimated by local H^1-seminorms.) It remains to multiply by $\alpha_J \asymp (\operatorname{diam}\Delta)^{-2}$, to apply this inequality to a $\tilde{u} \in V_J$ which extends $u - Ru$ outside Δ in an arbitrary way (thus, $R\tilde{u} = 0$ on Δ!), and to sum the inequalities with respect to Δ.

If the prolongations P_j resp. P_j^{ext} satisfy (34) then the optimal result $c_J = \mathrm{O}(1)$, $J \to \infty$, is the output (note that in general $c_J \ge C > 0$ for a constant depending on (3) only, see the proof of Theorem 1). For the P1 element, the standard choice is to define $P_j^{ext}\hat{u}_j$, $\hat{u}_j \in V_{j-1} + V_j$, at the edge midpoints of $\mathcal{T}_j$ by

$$(P_j^{ext}\hat{u}_j)(M_e) = \begin{cases} 0 & \text{if } e \subset \partial\Omega \\ \frac{1}{2}(\hat{u}_j(M_e+) + \hat{u}_j(M_e-)) & \text{elsewhere} \end{cases}$$

where $+$ and $-$ indicate one-sided limits of $\hat{u}_j$ at M_e from the two triangles Δ^+ and Δ^- attached to e (note that for $\hat{u}_j = u_j \in V_j$ all M_e are continuity points of u_j, thus, the limits coincide and $P_j^{ext}u_j = u_j$ as expected). Analogous definitions apply to the three-dimensional case. Now (34) is obvious from local considerations. For the P1 element in two dimensions, these results have been established in [15] by different means.

The analogous arguments go through for the Morley element which seems to be the simplest element useful for fourth–order problems. The local interpolation problem for this nonconforming element is indicated in Figure 3: on each triangle a quadratic polynomial p will be determined by interpolating function values at the three vertices and normal derivatives at the midpoints of the edges.

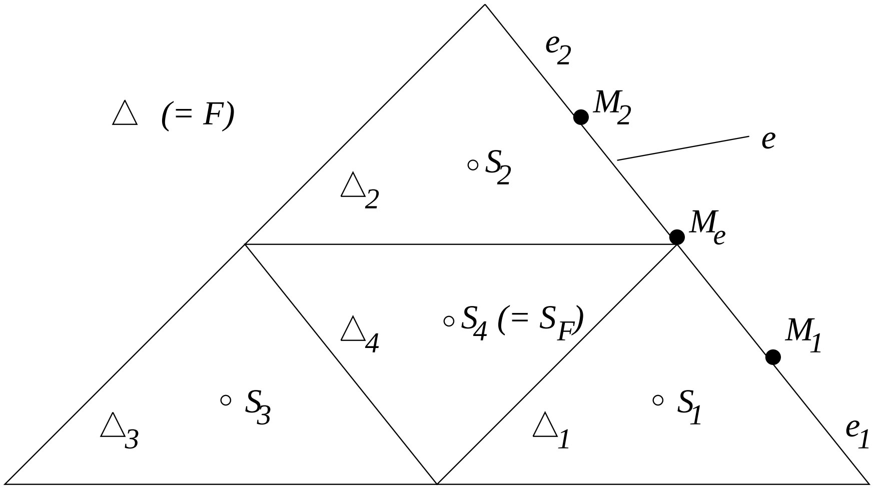

Fig. 2. Notation for the definition of R_{j-1}

The restrictions R_{j-1} are defined as follows: For any edge (face) of $\mathcal{T}_{j-1}$ we set

$$(R_{j-1}u_j)(M_e) = \frac{1}{2}\sum_{i=1}^{2} u_j(M_i) \qquad \left((R_{j-1}u_j)(S_F) = \frac{1}{4}\sum_{i=1}^{4} u_j(S_i)\right) \ ,$$

the notation is explained in Figure 2.

We verify the assumptions of Lemma 8. It suffices to consider $j = J$, for brevity set $u = u_J$, $v = v_{J-1}$, $R = R_{J-1}$, and let Δ be an arbitrary triangle (tetrahedron) of $\mathcal{T}_{J-1}$ decomposed into 4 triangles Δ_i (8 tetrahedra) of $\mathcal{T}_J$. The further considerations are different only in the notations, we consider the 2D case in detail. Using that u is linear on all Δ_i and (directional) first derivatives of v are constant on Δ, we get by Green's formula

$$\sum_{\Delta_i\in\Delta}\int\limits_{\Delta_i} \nabla(Ru-u)\cdot\nabla v \ \mathrm{d}x = \sum_i \sum_{\tilde e\subset\bar\Delta_i} \frac{\partial v}{\partial n_{\tilde e}} \int\limits_{\tilde e} (Ru-u)|_{\Delta_i} \ \mathrm{d}s$$

$$= \sum_{\tilde e\subset \mathrm{int}(\Delta)} \left[\frac{\partial v}{\partial n_{\tilde e}}\right] |\tilde e|(Ru(M_{\tilde e}) - u(M_{\tilde e})) + \sum_{\tilde e\subset\partial\Delta} \frac{\partial v}{\partial n_{\tilde e}} |\tilde e|(Ru(M_{\tilde e}) - u(M_{\tilde e}))$$

(all edges $\tilde e$ are from $\mathcal{T}_J$, and $|\tilde e|$ stands for the length of $\tilde e$). For edges $\tilde e$ interior to Δ the jump $[\frac{\partial v}{\partial n_{\tilde e}}]$ of the derivative of v in the direction $n_{\tilde e}$ normal to $\tilde e$ vanishes while for the pair e_1, e_2 of edges $\tilde e$ in $\partial\Delta$ forming an edge e in $\mathcal{T}_{J-1}$ (set $n_e = n_{e_1} = n_{e_2}$) we have by construction $|e_1| = |e_2| = |e|/2$ and

$$\frac{\partial v}{\partial n_{e_1}} = \frac{\partial v}{\partial n_{e_2}} = \frac{\partial v}{\partial n_e} \ , \quad \sum_{i=1}^{2} Ru(M_i) = 2Ru(M_e) = \sum_{i=1}^{2} u(M_i) \ .$$

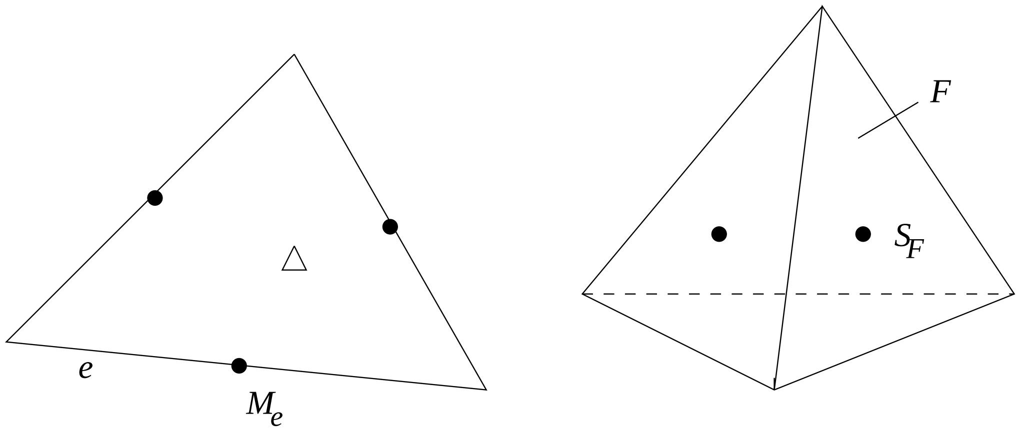

Fig. 1. Triangular and tetrahedral P1 elements

PROOF. If we set $u_j = \tilde{R}_j u_J$ ($u_{-1} = 0$), then we get from the orthogonality assumption

$$\sum_{j=0}^{J} \alpha_j \|\tilde{R}_j u_J - \tilde{R}_{j-1} u_J\|_H^2 = \sum_{j=0}^{J} \alpha_j \|u_j - R_{j-1} u_j\|_H^2$$

$$\leq C \sum_{j=0}^{J} \|u_j - R_{j-1} u_j\|_E^2 = C \sum_{j=0}^{J} (\|u_j\|_E^2 - \|u_{j-1}\|_E^2) = C \|u_J\|_E^2 \ .$$

This proves the assertion.

We illustrate the use of Lemma 7 by considering first the P1 element case (for more information on the finite element types considered in our examples, we refer to [7]). Let $\{\mathcal{T}_j\}$ be a sequence of dyadically, regularly refined partitions of a polygonal (polyhedral) domain Ω into triangles (tetrahedra) of size $\asymp 2^{-j}$ (for an appropriate version of such refinement in the three-dimensional case, see, e.g., [14]). Let $\{V_j\}$ be the sequence of nonconforming P1 elements which are piecewise defined by nodal interpolation in the edge midpoints M_e of the edges e (barycenters S_F of the faces F) of the triangles (tetrahedra), see Figure 1. Since we wish to consider the discretization of a $H_0^1(\Omega)$-elliptic problem, zero interpolation values are assigned to nodal points on $\partial\Omega$.

Define

$$(u,v)_E = \sum_{\Delta \in \mathcal{T}_J} \int_{\Delta} \nabla u \cdot \nabla v \ \mathrm{d}x \ , \qquad (u,v)_H = \int_{\Omega} uv \ \mathrm{d}x \ , \tag{37}$$

for all $u, v \in \sum_{j=0}^{J} V_j$ and assume that $a_J(u_J, v_J)$, the modified bilinear form on V_J, is symmetric and satisfies (3). Note that (9) is satisfied with $\alpha_j = C 2^{2j}$ for some constant C depending on $\mathcal{T}_0$.

compare (8), however, this choice does not necessarily give the minimal constant c_J (which stems from a much more complicated minimization with respect to $V_0 \times \ldots \times V_{J-1}$).

The following lemma might be helpful by separating the investigation of $\{P_j\}$ and $\{R_j\}$. Recall that by assumption the scalar products $(\cdot,\cdot)_E$ and $(\cdot,\cdot)_H$ are extended to $V_0 + V_1 + \ldots + V_J$.

Lemma 7 *Suppose that there are extensions* $P_j^{ext} : V_{j-1} + V_j \to V_j$ *of the prolongation operators* P_j *such that*

$$P_j^{ext}|_{V_{j-1}} = P_j \ , \quad P_j^{ext}|_{V_j} = Id \ , \quad \|P_j^{ext}\hat{v}_j\|_H^2 \le C\|\hat{v}_j\|_H^2 \ , \tag{34}$$

for all $\hat{v}_j \in V_{j-1} + V_j$ *and* $j > 0$. *Then the estimate*

$$\sum_{j=0}^{J} \alpha_j \|\tilde{R}_j u_J - \tilde{R}_{j-1} u_J\|_H^2 \le \hat{c}_J \|u_J\|_E^2 \qquad \forall\, u_J \in V_J \tag{35}$$

implies (8) with $c_J \le C\hat{c}_J$.

PROOF. This is obvious:

$$\sum_{j=0}^{J} \alpha_j \|(Id - P_j R_{j-1})\tilde{R}_j u_J\|_H^2 = \sum_{j=0}^{J} \alpha_j \|P_j^{ext}(\tilde{R}_j u_J - \tilde{R}_{j-1} u_J)\|_H^2$$

$$\le C \sum_{j=0}^{J} \alpha_j \|\tilde{R}_j u_J - \tilde{R}_{j-1} u_J\|_H^2 \le C\hat{c}_J \|u_J\|_E^2 \ .$$

The following simple lemma is useful for some low-order elements like the nonconforming P1 or Morley elements.

Lemma 8 *If* R_{j-1} *is the orthogonal projection with respect to the energy scalar product, i.e.,*

$$(R_{j-1}u_j, v_{j-1})_E = (u_j, v_{j-1})_E \qquad \forall\, v_{j-1} \in V_{j-1} \ ,$$

and satisfies

$$\alpha_j \|R_{j-1}u_j - u_j\|_H^2 \le C\|R_{j-1}u_j - u_j\|_E^2 \qquad \forall\, u_j \in V_j \ , \tag{36}$$

then $\hat{c}_J \le C$ *in (35).*

stiffness matrices A_j defined by

$$((A_j x_j, y_j)) \equiv a_j(u_j, v_j) \qquad \forall\, u_j, v_j \in V_j\ ,\ \ j = 0, \ldots, J-1\ ,$$

and obtain various V- and W-cycle multigrid results. See Chapter 4 of [1], especially Theorem 4.6, which shows an alternative construction of preconditioners by so-called variable V-cycles.

Our results on verifying the condition number estimates of Theorem 4 and 6 for the simpler additive preconditioners of BPX or hierarchical basis type in section 3 indicate the necessity of such alternatives in some cases like for the Morley element. Another alternative was presented in [17]: to avoid the detoriation of the condition numbers of the additive multilevel preconditioners for nonconforming elements (which comes mainly from the non–nestedness of the hierarchy of spaces), one may use a two-level method to switch to a conforming reference discretization for the nonconforming discretization on V_J which incorporates known (multilevel) preconditioners for the conforming "relative". The same idea is used in [6] in a domain decomposition framework, see also [9,18]. As a rule, this approach (as well as the use of variable V-cycle preconditioners) gives asymptotical optimal algorithms, in the sense, that the condition numbers remain bounded for $J \to \infty$, and the work per iteration is proportional to the number of unknowns.

3 Discussion and examples

3.1 Estimation of c_J

We do not have a general recipe for estimating the constant c_J in (8) for arbitrarily given sequences $\{P_j\}$ and $\{R_j\}$ satisfying (7). Note that in case of Theorem 4, since R_{j-1} does not enter the preconditioner itself we have some freedom of choice: The general form of the restrictions satisfying (7) is

$$R_{j-1} = P_j^+ + \hat{R}_{j-1} \qquad (\ P_j^+ = (P_j^* P_j)^{-1} P_j^*\ )\ ,$$

where $\hat{R}_{j-1} : V_j \to V_{j-1}$ is any linear operator vanishing on the range of P_j, and P_j^* denotes the adjoint of P_j with respect to the scalar product in H. The choice $R_{j-1} = P_j^+$ might seem to be most reasonable since the pseudoinverse P_j^+ satisfies the minimization property

$$\|u_j - P_j P_j^+ u_j\|_H = \min_{v_{j-1}} \|u_j - P_j v_{j-1}\|_H \qquad \forall\, u_j \in V_j\ ,$$

Remark 5 *Relying on Remark 3, we can derive a preconditioner C_J^{HB} which we call multilevel preconditioner of hierarchical basis type. Assume that there is a basis $\{H_{j,i} \,:\, i = 1, \ldots, m_j\}$ of $W_j \in V_j$ which satisfies the analog of (24). Recall that W_j was defined as the range of the operator $Id - I_j P_{j-1}$ which, according to (7), has dimension $m_j = n_j - n_{j-1}$, $j = 0, \ldots, J$ $(n_{-1} = 0)$. In analogy to the above setting, we require the $m_j \times m_j$ matrices $\hat{S}_j$ to satisfy*

$$0 < \hat{\gamma} \le \frac{((\hat{S}_j^{-1} x_j^{HB}, x_j^{HB}))}{\alpha_j \|w_j\|_H^2} \le \hat{\Gamma} < \infty \quad \forall\, w_j = \sum_{i=1}^{m_j} x_{j,i}^{HB} H_{j,i} \in W_j \tag{31}$$

(uniformly in $j = 0, 1, \ldots, J$). To make the comparison with C_J^{BPX} more transparent, denote by $\hat{P}_j$ the $n_j \times m_j$ matrix corresponding to a basis transformation which transfers the hierarchical basis coefficient vector x_j^{HB} of $w_j \in W_j \subset V_j$ into the nodal basis coefficient vector x_j of the same function with respect to $\{N_{j,i}\}$. As before, if the hierarchical basis functions are locally supported, and are stable with respect to $\|\cdot\|_H$, both $\hat{S}_j$ and $\hat{P}_j$ are expected to be sparse, with $\mathrm{O}(n_j)$ non-zero elements.

If we define $\hat{C}_J \equiv C_J^{HB}$ recursively by

$$\hat{C}_j = P_j \hat{C}_{j-1} P_j^T + \hat{P}_j \hat{S}_j \hat{P}_j^T\ , \quad j > 0\ ; \qquad \hat{C}_0 = \hat{P}_0 \hat{S}_0 \hat{P}_0^T \tag{32}$$

then we have

Theorem 6 *The matrix C_J^{HB} defined by the recursion (32) is a symmetric preconditioner for the linear system (1) satisfying*

$$c \cdot \frac{\hat{\gamma}}{\hat{\Gamma}} c_J^* \max_{j=0,\ldots,J} \hat{\tau}_j^J \le \kappa(A_J^{1/2} C_J^{HB} A_J^{1/2}) \le C \cdot \frac{\hat{\Gamma}}{\hat{\gamma}} c_J \sum_{j=0}^{J} \hat{\tau}_j^J\ . \tag{33}$$

Usually, if the upper estimates for the condition numbers in (30) resp. (33) are bounded independent of J or slowly growing, the above preconditioners are used in a pcg-method which is faster than applying the simpler extrapolated Richardson iteration to the preconditioned system (1). The latter is also called additive Schwarz method associated with the corresponding subspace splitting, see [19,1,16]. Multiplicative Schwarz methods based on the same splittings are known to be equivalent to specific multigrid V-cycle iterations, see [19,1,12]. If directly applied in the above framework, these algorithms involve coarse grid stiffness matrices of the form $\tilde{A}_j = \tilde{P}_j^T A_J \tilde{P}_j$, $j = 0, \ldots, J-1$, which in the general case can be generated only with $\mathrm{O}(J n_J)$ operations. We do not go into further details but refer to an alternative approach of [3] where, within a different set of assumptions, the authors directly work with the coarse-grid

We call preconditioners based on these choices for S_j multilevel preconditioners of BPX type as they are close to the corresponding method for conforming linear finite elements proposed in [2].

Assumptions (9) and (27) usually follow by a trivial local consideration and are valid for the appropriate choice of the constants α_j (actually, together they determine $\{\alpha_j\}$ up to constants). In practice, if the coarsest-grid subspace V_0 is still relatively large, one replaces S_0 by exact solvers for the variational problem corresponding to (2) on V_0. Note finally that the locality assumption for $\{N_{j,i}\}$ is used in the proofs but also leads to sparse representations for $\{P_j\}$ and $\{R_j\}$.

Define the $n_J \times n_J$ matrix A_J in (1) by

$$((A_J x_J, y_J)) \equiv a_J(u_J, v_J) \qquad \forall\, u_J, v_J \in V_J$$

(y_J denotes the coefficient vector of v_J), and introduce $f_J \in \mathbf{R}^{n_J}$ with the components $(f, N_{J,i})_H$. We preserve the notations $P_j, \tilde{P}_j, R_j, \tilde{R}_j$ for the rectangular matrices corresponding to the intergrid transfer operators.

Define $C_J \equiv C_J^{BPX}$ recursively by

$$C_j = P_j C_{j-1} P_j^T + S_j\ , \quad j > 0\ ; \qquad C_0 = S_0\ . \tag{28}$$

In the typical applications, if P_j and S_j are sparse with $\mathrm{O}(n_j)$ non-zero elements, and the dimensions $\{n_j\}$ grow geometrically, $j = 0, 1, \ldots, J$, then a matrix-vector multiplication with C_J^{BPX} can be performed in $\mathrm{O}(n_J)$ operations. Moreover, a moment's reflection shows that $C_J^{BPX} A_J$ is the matrix representation of the additive Schwarz operator $\mathcal{P}_J^{BPX}$ associated with the splitting

$$\{V_J; a_J(\cdot,\cdot)\} = \sum_{j=0}^{J} \tilde{P}_j \{V_j; b_j(\cdot,\cdot)\} \tag{29}$$

which is obtained from the splitting (12) of Theorem 1 by replacing the scalar products $\alpha_j(u_j, v_j)_H$ on the subspaces V_j by their spectrally equivalent discrete counterparts $b_j(u_j, v_j) = ((S_j^{-1} x_j, y_j))$. As a direct consequence of Theorem 1, Lemma 2, and (23), we have

Theorem 4 *The matrix C_J^{BPX} defined by the recursion (28) is a symmetric preconditioner for the linear system (1) satisfying*

$$c \cdot \frac{\gamma}{\Gamma} \max_{j=0,\ldots,J} \tau_j^J \le \kappa(A_J^{1/2} C_J^{BPX} A_J^{1/2}) \le C \cdot \frac{\Gamma}{\gamma} c_J \sum_{j=0}^{J} \tau_j^J\ . \tag{30}$$

2.2 Preconditioners of BPX and hierarchical basis type

To get practical algorithms, we assume that all V_j are finite-dimensional and equipped with a designated basis $\{N_{j,i}\}$ such that each $u_j \in V_j$ has its unique representation

$$u_j = \sum_{i=0}^{n_j} x_{j,i} N_{j,i} \ .$$

By $x_j = (x_{j,i}) \in \mathbf{R}^{n_j}$ we denote the coefficient vector of u_j.

We fix some set of symmetric positive definite $n_j \times n_j$ matrices S_j such that for some $0 < \gamma < \Gamma < \infty$

$$\gamma \le \frac{((S_j^{-1} x_j, x_j))}{\alpha_j \|u_j\|_H^2} \le \Gamma \qquad \forall\, u_j \in V_j \ , \quad j = 0, 1, \ldots, J \ . \tag{23}$$

The notation $((x,y))$ stands for the Euclidean scalar product of two $\mathbf{R}^n$-vectors, the particular n will be clear from the context, here $n = n_j$ would be appropriate. In some model cases, if $N_{j,i}$ are the usual nodal basis functions for a given finite element type on partitions obtained by regular dyadic refinement, we may typically choose S_j as a multiple of an identity matrix or some other diagonal matrix. Indeed, assume that the basis $\{N_{j,i}\}$ for V_j is local and H-stable: The set $\{\mathrm{supp}\, N_{j,i}\}$ should locally have finite overlap, and

$$0 < c \le \frac{\sum_i x_{j,i}^2 \|N_{j,i}\|_H^2}{\|u_j\|_H^2} \le C < \infty \qquad \forall\, u_j = \sum_i x_{j,i} N_{j,i} \in V_j \ . \tag{24}$$

For finite element nodal basis functions (on regular partitions) this assumption easily follows from the definition and unique solvability of the local interpolation problems if $H = L_2(\Omega)$. Then

$$S_j = \mathrm{diag}\Big(\frac{1}{\alpha_j \|N_{j,i}\|_H^2}\Big) \ , \quad j \ge 0 \ , \tag{25}$$

satisfies (23) with $\gamma = c$ and $\Gamma = C$. Another popular choice is

$$S_j = \mathrm{diag}\Big(\frac{1}{(N_{j,i}, N_{j,i})_E}\Big) \ , \quad j \ge 0 \ . \tag{26}$$

which is justified if we have in addition to (9) the estimate

$$\alpha_j \|N_{j,i}\|_H^2 \le C \|N_{j,i}\|_E^2 \ , \quad \forall\, N_{j,i} \ . \tag{27}$$

Since, by our choice u_J is a_J-orthogonal to the range of P_J, and

$$u_J = v_J^* + P_J u_{J-1}^* \ , \qquad u_{J-1}^* = \sum_{j=0}^{J-1} P_{J-1} \dots P_{j+1} v_j^* \ ,$$

it follows that $a_J(u_J, u_J) \le a_J(v_J^*, v_J^*)$. By (3), (9), we arrive at

$$a_J(u_J, u_J) \le C\|v_J^*\|_E^2 \le C\alpha_j \|v_J^*\|_H^2 \le C|||u_J|||^2$$

for the above u_J, which in view of (19) gives

$$\|\mathcal{P}_J^{-1}\|_{A_0} = \left(\inf_{u_J \in V_J} \frac{a_J(u_J, u_J)}{|||u_J|||^2} \right)^{-1} = \sup_{u_J \in V_J} \frac{|||u_J|||^2}{a_J(u_J, u_J)} \ge c \ .$$

Combined with the previous estimate for $\|\mathcal{P}_J\|_{A_0}$, this gives the lower bound in (16). The proof of Theorem 1 is complete.

Remark 3 *The same considerations can be applied to the modified splitting*

$$\{V; a_J(\cdot,\cdot)\} = \sum_{j=0}^{J} \tilde{P}_j \{W_j; \alpha_j(\cdot,\cdot)_H\} \ , \tag{20}$$

where the subspaces $W_j \subset V_j$, $j \ge 0$, *are defined as the ranges of the operators* $Id - P_j R_{j-1}$ *(in particular,* $W_0 = V_0$*), and lead to the following estimates for the corresponding additive Schwarz operator:*

$$c \cdot c_J^* \left(\max_{j=0,\dots,J} \hat{\tau}_j^J \right) \le \kappa(\tilde{\mathcal{P}}_J) \le C \cdot c_J \sum_{j=0}^{J} \hat{\tau}_j^J \ , \tag{21}$$

where c_J^* *denotes the minimal value of* c_J *in (8), and*

$$\hat{\tau}_j^J = \sup\{\|\tilde{P}_j u_j\|_E^2 \ : \ \alpha_j \|u_j\|_H^2 \le 1 \, , \ u_j \in W_j\} \ , \quad j = 0, \dots, J \ . \tag{22}$$

Since, by definition of W_j *and (7), the splitting (20) is into a direct sum of subspaces, the proof is even simpler since the infimum can be dropped in the definition of the triple-bar norm. In general,* $\hat{\tau}_j^J \le \tau_j^J$ *since* $W_j \subset V_j$, *therefore the upper bound (21) is at least as good as in Theorem 1. In some cases the factor* c_J *which is closely tied to the use of the set of intergrid transfer operators (6) and, therefore, to the splitting (20) might be a too rough estimate in (13) resp. (16). Note again that* R_j *does not enter the implementation of the preconditioners associated with* $\mathcal{P}_J$ *and* $\tilde{\mathcal{P}}_J$.

Denoting $u_j = (Id - P_j R_{j-1})\tilde{R}_j u_J$, the lower estimate in (13) is now obvious from (11), (3), and (8):

$$|||u_J|||^2 \leq \sum_{j=0}^{J} \alpha_j \|u_j\|_H^2 \leq c_J \|u_J\|_E^2 \leq C c_J a_J(u_J, u_J) \, .$$

For the upper estimate, take an arbitrary decomposition $u_J = \sum_{j=0}^{J} \tilde{P}_j v_j$, and use (3), (10):

$$\begin{aligned} a_J(u_J, u_J) &\leq C \|u_J\|_E^2 \leq C\Big(\sum_{j=0}^{J} \sqrt{\tau_j^J} \frac{\|\tilde{P}_j v_j\|_E}{\sqrt{\tau_j^J}}\Big)^2 \\ &\leq C\Big(\sum_{j=0}^{J} \tau_j^J\Big) \sum_{j=0}^{J} \frac{\|\tilde{P}_j v_j\|_E^2}{\tau_j^J} \leq C\Big(\sum_{j=0}^{J} \tau_j^J\Big) \sum_{j=0}^{J} \alpha_j \|v_j\|_H^2 \, . \end{aligned}$$

Taking the infimum with respect to all decompositions of u_J, by (11) we get the upper estimate in (13).

Thus, the condition number estimate from above in (16) is established. The following examples show the lower bound. First we pick $u_J = \tilde{P}_j v_j$ where $v_j \in V_j$ is arbitrary (fix temporarily j). For all those u_J, by (3) and (11),

$$\frac{a_J(u_J, u_J)}{|||u_J|||^2} \geq c \frac{\|\tilde{P}_j v_j\|_E^2}{\alpha_j \|v_j\|_H^2}$$

Now taking the supremum with respect to v_j and j, we see (compare the definition of τ_j^J in (10), (19), and our above discussion) that

$$\|\mathcal{P}_J\|_{A_0} = \sup_{u_J \in V_J} \frac{a_J(u_J, u_J)}{|||u_J|||^2} \geq c \max_{j=0,1,\ldots,J} \tau_j^J \, .$$

On the other hand, fix any nontrivial $u_J \in V_J$ such that $a_J(u_J, P_J v_{J-1}) = 0$ for all $v_{J-1} \in V_{J-1}$ (the existence of such an element is guaranteed if $n_J \equiv \dim V_J > \dim V_{J-1}$, the exceptional case of $n_J = n_{J-1}$, which can not occur in practice, is left to the reader). Assume that $\{v_j^*\}$ is such that

$$u_J = \sum_{j=0}^{J} \tilde{P}_j v_j^* \, , \qquad |||u_J|||^2 = \sum_{j=0}^{J} \alpha_j \|v_j^*\|_H^2 \, .$$

introduce

$$|||u_0|||_0^2 = \inf_{u_1 \in H_1 \,:\, u_0 = Lu_1} \|u_1\|_{A_1}^2 \qquad \forall\, u_0 \in H_0 \; . \tag{17}$$

Lemma 2 *Assume that we have a two-sided estimate*

$$c_0|||u_0|||_0^2 \le \|u_0\|_{A_0}^2 \le C_0|||u_0|||_0^2 \qquad \forall\, u_0 \in H_0 \; . \tag{18}$$

Then the operator $\mathcal{P} = LA_1^{-1}L^*A_0$ is symmetric positive definite with respect to $(\cdot,\cdot)_{A_0}$, and

$$\|\mathcal{P}\|_{A_0} = \inf_{(18)} C_0 \; , \qquad \|\mathcal{P}^{-1}\|_{A_0} = (\sup_{(18)} c_0)^{-1} \; , \tag{19}$$

i.e., the spectral condition number $\kappa(\mathcal{P}) = \|\mathcal{P}\|_{A_0}\|\mathcal{P}^{-1}\|_{A_0}$ of $\mathcal{P}$ is given by the (optimal) constants c_0, C_0 in (18).

Note that $\mathcal{P} :: H_0 \to H_0$ itself does not depend on the particular choice of the scalar products $(\cdot,\cdot)_i$, $i = 0, 1$, however, the explicit representation $\mathcal{P} = LA_1^{-1}L^*A_0$ is often convenient to derive matrix representations for particular cases. To formally fit Theorem 1 into the framework of Lemma 2, set $H_0 = V_J$, with $(u_J, v_J)_{A_0} = a_J(u_J, v_J)$ as the proper choice for the scalar product, and introduce $H_1 = V_0 \times V_1 \times \ldots \times V_J$, with the elements $\tilde{u} \equiv (u_0, u_1, \ldots, u_J)$ and

$$(\tilde{u}, \tilde{v})_1 = \sum_{j=0}^{J} (u_j, v_j)_H \; , \qquad (\tilde{u}, \tilde{v})_{A_1} = \sum_{j=0}^{J} \alpha_j (u_j, v_j)_H \; .$$

Thus, the operator-matrix for A_1 is diagonal, with the identity operators on the V_j's multiplied by α_j in the diagonal. The operator L is given by $L\tilde{u} = \sum_{j=0}^{J} \tilde{P}_j u_j$, with $\tilde{P}_j$ defined in (5). Without making explicit A_0 and L^* (which depend on a choice for $(\cdot,\cdot)_0$), it is clear that the triple-bar norm in (11) coincides with the triple-bar norm in (17). Thus, by proving (13), we get estimates for c_0, C_0 in (18), and the remaining statements of Theorem 1 follow from Lemma 2 by observing that $\mathcal{P}_J = \mathcal{P}$ (the lower estimate in (16) requires a bit more work, see below).

We establish (13). By definition of the intergrid transfer operators, for any $u_J \in V_J$ we have

$$\sum_{j=0}^{J} \tilde{P}_j((Id - P_jR_{j-1})\tilde{R}_j u_J = \sum_{j=0}^{J} (\tilde{P}_j\tilde{R}_j u_J - \tilde{P}_{j-1}\tilde{R}_{j-1}u_J) = u_J \; .$$

one-dimensional subspaces $V_{j,i}$ are spanned by the nodal basis functions $N_{j,i}$, see subsection 2.2. Another modification of (12) leads to hierarchical preconditioners which depend on knowledge of a basis for the range of $Id - P_j R_{j-1}$ in V_j. Although the numerical performance of both variants might be different, the above theorem leads to similar lower and upper bounds for their asymptotic preconditioning power.

In section 3 we discuss the estimation of the constants c_J from (8) and τ_j^J defined in (10) which enter the condition number estimates. For some low order elements like the nonconforming P1 or Morley elements, $c_J = \mathrm{O}(1),\ J \to \infty$, can be shown by using orthoprojections R_j with respect to the energy scalar product $(\cdot,\cdot)_E$. For finite elements of higher degree, the techniques of [10,11] can be adopted. As a rule, with the natural choice of the intergrid transfer operators where the P_j, R_j are defined by nodal value averaging procedures, boundedness or moderate growth of the c_J will be observed.

What concerns the numbers τ_j^J, less is known. Only in two cases (triangular P1 element [15]), rotated Q1 rectangular element [8]) there is a theoretical estimate showing $\tau_j^J = \mathrm{O}(1)$ for $j, J \to \infty$. We show that estimates for the τ_j^J are numerically accessible in many cases, since in the situation of regular dyadic refinement it suffices to test the energy norm growth on nodal basis functions. We have tested a number of low order finite element types, the results are reported on in subsection 3.2. Unfortunately, for the standard sets of intergrid transfer operators (4) defined by nodal value averaging, the numbers for the Morley element are disappointing showing that the preconditioning power of the corresponding multilevel method deteriorates exponentially with the number of levels. The next simple nonconforming element for plate bending, the Zienkiewicz triangle, shows much better behavior in this respect.

2 Theory and preconditioners

2.1 Proof of Theorem 1

We rely on the abstract theory of additive Schwarz methods using Nepomnjashchikh's fictitious space lemma which we state in a form convenient for our purposes (see section 4.1, especially Theorem 17, of [16]). Let H_i be real Hilbert spaces with scalar products denoted by $(\cdot,\cdot)_i$, and $A_i : H_i \to H_i$ symmetric positive definite operators ($i = 0, 1$). By $(u_i, v_i)_{A_i} = (A_i u_i, v_i)_i$ (resp. $\|u_i\|_{A_i}$) we denote the associated bilinear forms (resp. norms). Let $L : H_1 \to H_0$ be linear, bounded, and surjective, denote by $L^* : H_0 \to H_1$ its adjoint, and

Theorem 1 *The discrete norm*

$$|||u_J|||^2 \equiv \inf_{v_j \in V_j \,:\, u_J = \sum_{j=0}^{J} \tilde{P}_j v_j} \sum_{j=0}^{J} \alpha_j \|v_j\|_H^2 \qquad (11)$$

associated with the additive subspace splitting

$$\{V_J; a_J(\cdot,\cdot)\} = \sum_{j=0}^{J} \tilde{P}_j \{V_j; \alpha_j(\cdot,\cdot)_H\} \qquad (12)$$

satisfies under the assumptions (3), (7), (8), (9) the two-sided inequality

$$\frac{c}{c_J} |||u_J|||^2 \leq a_J(u_J, u_J) \leq C(\sum_{j=0}^{J} \tau_j^J) |||u_J|||^2 , \quad u_J \in V_J , \qquad (13)$$

where the absolute constants $0 < c, C < \infty$ depend only on the constants in (3). Moreover, the additive Schwarz operator associated with (12)

$$\mathcal{P}_J = \sum_{j=0}^{J} \tilde{P}_j T_j , \qquad (14)$$

where $T_j : V_J \to V_j$ is determined by

$$\alpha_j (T_j u_J, v_j)_H = a_J(u_J, \tilde{P}_j v_j) \qquad \forall\, v_j \in V_j ,\ j = 0, 1, \ldots, J , \qquad (15)$$

possesses a two-sided condition number estimate

$$c \cdot \left(\max_{j=0,\ldots,J} \tau_j^J \right) \leq \kappa(\mathcal{P}_J) \leq C \cdot c_J \sum_{j=0}^{J} \tau_j^J . \qquad (16)$$

For the proof and details of the used terminology, see subsection 2.1. Theorem 1 gives an additional tool to investigate preconditioners of BPX type for some nonconforming situations, and allows us also to obtain negative results based on knowledge about the growth of the numbers (10) (see the estimate from below in (16)). Note that under optimal circumstances, if both c_J and τ_j^J are bounded, the upper estimate in (16) behaves like $\mathrm{O}(J)$ for $J \to \infty$. The numerical evidence for the P1 element [15] seems to indicate that this is the best we can hope for within the assumptions of Theorem 1 (however, there is no rigorous proof for this claim).

BPX-type preconditioners essentially coincide with the additive Schwarz operator for a subspace splitting similar to (12) by replacing V_j by $\sum_i V_{j,i}$ where the a multiplicative way and lead to relatively rough estimates. On the other hand, in the particular case of the nonconforming triangular P1 element, we observed in [15] that direct estimates for $\|\tilde{P}_j\|_E$, though harder to obtain, can be sometimes an order better than their trivial upper estimate $\|P_J\|_E\|P_{J-1}\|_E \ldots \|P_{j+1}\|_E$.

The considerations of [15] can be made more precise and general to cover other interesting situations. We present here a simplified version of this extension and discuss on its basis the behavior of multilevel preconditioners for some nonconforming elements. The assumptions are as follows. Let

$$R_j \;:\; V_{j+1} \to V_j \;, \quad \tilde{R}_j = R_j R_{j+1} \ldots R_{J-1} \;:\; V_J \to V_j \;, \tag{6}$$

$0 \le j < J$ (in addition, set $\tilde{R}_J = Id$, $R_{-1} = \tilde{R}_{-1} = 0$)), be some other set of intergrid transfer operators (restrictions) and their iterates such that

$$R_{j-1} P_j u_{j-1} = u_{j-1} \qquad \forall\, u_{j-1} \in V_{j-1} \;, \quad j \ge 1 \;, \tag{7}$$

and, for some constant c_J,

$$\sum_{j=0}^{J} \alpha_j \|(Id - P_j R_{j-1})\tilde{R}_j u_J\|_H^2 \le c_J \|u_J\|_E^2 \qquad \forall\, u_J \in V_J \;. \tag{8}$$

The positive numbers α_j are chosen such that the following inverse inequality holds:

$$\|u_j\|_E^2 \le \alpha_j \|u_j\|_H^2 \qquad \forall\, u_j \in V_j \;, \quad j = 0, \ldots, J \;. \tag{9}$$

It turns out that for some low order nonconforming elements (P1, Morley), with the standard choices for the energy product $(u, v)_E$, a good choice for the operators R_j are orthogonal projections (with respect to $(\cdot,\cdot)_E$) which can be found explicitly. Thus, in some cases, the estimate in (8) is easily available and not bound to additional regularity assumptions. We discuss the possibilities of proving (8) or replacing it by other conditions in subsection 3.1.

Introduce finally the numbers

$$\tau_j^J = \sup\{\|\tilde{P}_j u_j\|_E^2 \;:\; \alpha_j \|u_j\|_H^2 \le 1 \,,\; u_j \in V_j\} \;, \quad j = 0, \ldots, J \;, \tag{10}$$

which characterize the energy norm behavior of $\tilde{P}_j$ from (5) (i.e., of the intergrid transfer operators (4) iterated over many levels).

of nonconforming finite element spaces with respect to an increasing sequence of partitions obtained by regular or nested dyadic refinement. In the examples below, we mostly discuss the case of triangular elements in $\mathbf{R}^2$, but the theory developed also applies to simplicial resp. rectangular elements in $\mathbf{R}^d$. We are interested in preconditioners for the linear system

$$A_J x_J = f_J \ , \tag{1}$$

the so-called nodal basis discretization of a variational problem

$$\text{Find } \ u_J \in V_J \ \text{ such that} \qquad a_J(u_J, v_J) = (f, v_J)_H \quad \forall\, v_J \in V_J \tag{2}$$

in V_J, which are based upon simple subproblems associated with the coarse-grid spaces V_j. The bilinear form $a_J(u_J, v_J)$ will be assumed symmetric, continuous, and coercive with respect to a discrete energy scalar product $(u_J, v_J)_E$ on V_J. In particular,

$$c\|u_J\|_E^2 \le a_J(u_J, u_J) \le C\|u_J\|_E^2 \qquad \forall\, u_J \in V_J \tag{3}$$

(here and in the following, $0 < c, C < \infty$ denote generic constants which do not depend on J and may have different values at different places). The energy scalar product which may depend on J as well as the scalar product $(\cdot,\cdot)_H$ (in all our applications H coincides with $L_2(\Omega)$) should make sense on $V_0 + V_1 + \ldots + V_J$. Since the spaces V_j are not embedded into each other, one needs intergrid transfer operators (prolongations)

$$P_j \ : \ V_{j-1} \to V_j \ , \quad j = 1, \ldots, J, \tag{4}$$

and their iterates

$$\tilde{P}_j = P_J P_{J-1} \ldots P_{j+1} \ : \ V_j \to V_J \ , \quad j = 0, \ldots, J-1 \ \ (\tilde{P}_J = Id) \ , \tag{5}$$

to define subproblems of (2) associated with the coarse-grid spaces V_j (for notational convenience, set $V_{-1} = \{0\}$ and $P_0 = \tilde{P}_0 = 0$). This approach was started in the multigrid community, and basically led to the development of a general theory covering W-cycle convergence and optimal variable V-cycle preconditioners, see [5,3,1]. When it comes to simpler V-cycle preconditioners examples of which are multilevel methods based on further splitting the V_j with respect to their respective nodal bases $\{N_{j,i}\}$ or some hierarchical basis constructions, less is known.

We wish to mention a theory developed by Dörfler [10,11] where conditions in terms of $\|P_j\|_E$ are used. These norms entered the final results of [10,11] in

Intergrid transfer operators and multilevel preconditioners for nonconforming discretizations

Peter Oswald

Institute for Algorithms and Scientific Computing, GMD, D-53754 Sankt Augustin, Germany

We discuss multilevel preconditioners of hierarchical basis and BPX type for nonconforming discretizations of second and fourth order elliptic variational problems where the underlying subspace splitting of a nonconforming fine grid space is obtained from the natural sequence of nonconforming coarse grid spaces using appropriately designed intergrid transfer operators. We present a simple convergence theory which shows the importance of controlling the energy norm growth of the iterated coarse-to-fine-grid operators. It enters both upper and lower bounds for the condition number of the preconditioned linear system, and can be checked numerically in the case of regular dyadic refinement. For the standard sets of intergrid transfer operators (prolongations based on nodal value averaging), the numerical tests with some low order nonconforming elements on uniform grids indicate boundedness of these norms, with the exception of the Morley element where the condition numbers deteriorate exponentially with the number of levels. The results complement recent work by Bramble/Pasciak/Xu, Brenner, Dörfler, and the author.

Key words: Preconditioning, nonconforming finite elements.

1 Introduction

We consider multilevel methods of hierarchical basis or BPX type for nonconforming finite element discretizations of second or fourth order problems associated with the natural sequence

$$V_0 \rightarrow V_1 \rightarrow \ldots \rightarrow V_J$$