Czechoslovak Mathematical Journal

Jan Brousek; Pavla Fraňková; Petr Vaněk

Improved convergence bounds for smoothed aggregation method: linear dependence of the convergence rate on the number of levels

Czechoslovak Mathematical Journal, Vol. 66 (2016), No. 3, 829–845

Persistent URL: http://dml.cz/dmlcz/145874

Czechoslovak Mathematical Journal, 66 (141) (2016), 829–845

IMPROVED CONVERGENCE BOUNDS FOR SMOOTHED AGGREGATION METHOD: LINEAR DEPENDENCE OF THE CONVERGENCE RATE ON THE NUMBER OF LEVELS

Jan Brousek, Pavla Fraňková, Petr Vaněk, Plzeň

(Received December 26, 2015)

Dedicated to the memory of Miroslav Fiedler

Abstract. The smoothed aggregation method has became a widely used tool for solving the linear systems arising by the discretization of elliptic partial differential equations and their singular perturbations. The smoothed aggregation method is an algebraic multigrid technique where the prolongators are constructed in two steps. First, the tentative prolongator is constructed by the aggregation (or, the generalized aggregation) method. Then, the range of the tentative prolongator is smoothed by a sparse linear prolongator smoother. The tentative prolongator is responsible for the approximation, while the prolongator smoother enforces the smoothness of the coarse-level basis functions.

Keywords: smoothed aggregation; improved convergence bound

MSC 2010: 65N55, 65F10, 65N12

1. Introduction

In [7], we established the convergence bound for the smoothed aggregation method that depends polylogarithmically on the meshsize. In particular, we proved that the convergence rate estimate is dependent on the third power of the number of levels (it takes $\log_3 h^{-1}$ of levels to reach $O(1)$ degrees of freedom on the coarsest level). To be more precise, we have shown that the condition number of the stiffness matrix preconditioned by the symmetric multigrid iteration grows as $O(L^3)$, where L is the number of levels. In this paper, we prove an improved convergence bound that depends on the first power of the number of levels (that is, the condition number of the preconditioned system grows as $O(L) = O(\log_3 h^{-1})$). Thus, the dependence of the condition number of the preconditioned system on the meshsize is logarithmical.

Our assumptions are basically similar to those of [7]; the difference is that while in [7] we needed to control the growth of the diameter of the aggregates (and the overlaps of the balls circumscribed to the aggregates), here, we need to control the growth of the diameter of the supports of the smoothed aggregation coarse-level basis functions and their overlaps. It is anyway needed to control the growth of overlaps of the basis-functions supports (the sparsity of the coarse-level matrices) for the computational complexity reasons.

The experimental material published in [9], [2] suggests that the convergence rate is bounded uniformly with respect to the number of levels. In the theoretical front, for 10 levels, the estimate of the condition number of the stiffness matrix preconditioned by the symmetric smoothed aggregation method grows 1000 times according to the old theory and only 10 times according to the theory presented here.

In [7], we verified the assumptions of the regularity-free abstract convergence theory of [1] based solely on the weak approximation property of the disaggregated functions. Here, we verify the assumptions of the theory of [1] based on the weak approximation property of the disaggregated functions and the stability of the smoothed-aggregation based interpolation operator in the energy norm. The first goal is to avoid the need to prove directly the weak approximation property of the smoothed aggregation-based prolongators, because it invokes the need of the equivalence of the discrete and the continuous L_2-norms on the coarse-levels. The proof of such equivalence is very difficult and is so far restricted to the cases with model geometry, see [3]. We avoid the need of this equivalence on the operator level by a simple trick. As a result, we will need only the equivalence of the discrete and continuous L_2-norms to hold for purely disaggregated functions that holds trivially because the disaggregation operators are orthogonal matrices. The convergence bound is established for the second order scalar elliptic problems. Both the method and the presented theory can be easily extended to the case of the three-dimensional linear elasticity. For the generalized aggregation method suitable for constructing the tentative prolongator, when solving the systems of partial differential equations, we refer to [7].

2. Standard variational multigrid algorithm

We solve the system of linear algebraic equations

$$A\mathbf{x} = \mathbf{f} \tag{2.1}$$

with a symmetric, positive definite $n \times n$ matrix A and a right-hand side $\mathbf{f} \in \mathbb{R}^n$. To define the multigrid algorithm, we need the system of linear prolongators

$$I^l_{l+1}\colon\ \mathbb{R}^{n_{l+1}} \to \mathbb{R}^{n_l}, \quad n_1 = n,\ n_{l+1} < n_l,\ l = 1, \ldots, L-1$$

and the smoothing iterative procedures $\mathcal{S}_l(\cdot,\cdot)\colon \mathbb{R}^{n_l} \times \mathbb{R}^{n_l} \to \mathbb{R}^{n_l}$ on all levels $l = 1,\dots,L-1$. The multigrid algorithm is defined as follows:

Algorithm 1. Given the system (2.1), the prolongators I^l_{l+1}, $l = 1,\dots,L-1$, the smoothers $\mathcal{S}_l(\cdot,\cdot)$, $l = 1,\dots,L-1$, the right-hand side $\mathbf{f} \in \mathbb{R}^n$ and a smoothing and cycle parameters $\nu, \gamma > 0$, set $A_1 = A$, $A_{l+1} = (I^l_{l+1})^{\mathrm{T}} A_l I^l_{l+1}$, $l = 1,\dots,L-1$ and $\mathbf{f}^1 = \mathbf{f}$.

For a given input iterate $\mathbf{x} \in \mathbb{R}^n$, perform the iteration $\mathbf{x} \leftarrow MG(\mathbf{x},\mathbf{f})$ as follows: Set $MG(\cdot,\cdot) = MG_1(\cdot,\cdot)$ and $\mathbf{x}^1 = \mathbf{x}$ where $MG_l(\cdot,\cdot)$, $l = 1,\dots,L-1$ is given by:

- ▷ Pre-smoothing: Perform ν iterations of $\mathbf{x}^l \leftarrow \mathcal{S}_l(\mathbf{x}^l,\mathbf{f}^l)$.
- ▷ Coarse-level correction:
  - – Set $\mathbf{f}^{l+1} = (I^l_{l+1})^{\mathrm{T}}(\mathbf{f}^l - A_l\mathbf{x}^l)$,
  - – if $l = L-1$, solve directly $A_{l+1}\mathbf{x}^{l+1} = \mathbf{f}^{l+1}$, otherwise set $\mathbf{x}^{l+1} = \mathbf{0}$ and perform γ iterations of $\mathbf{x}^{l+1} = MG_{l+1}(\mathbf{x}^{l+1},\mathbf{f}^{l+1})$,
  - – correct the approximation on the level l by $\mathbf{x}^l = \mathbf{x}^l + I^l_{l+1}\mathbf{x}^{l+1}$.
- ▷ Post-smoothing: Perform ν iterations of $\mathbf{x}^l \leftarrow \mathcal{S}_l(\mathbf{f}^l,\mathbf{x}^l)$.

Our theory uses the abstract result of [1]. We define the coarse-space $(U_l, \|\cdot\|_l)$ by

$$U_l = \mathrm{Range}(I^1_l), \quad \|\cdot\|_l\colon \mathbf{x} \in U_l \mapsto \inf_{\mathbf{y}\colon I^1_l\mathbf{y}=\mathbf{x}} \|\mathbf{y}\|, \tag{2.2}$$

where I^1_l is a *composite prolongator* given by $I^1_l = I^1_2 \dots I^{l-1}_l$, $l = 1,\dots,L$. Clearly

$$\|I^1_l\mathbf{x}\|_l \leqslant \|\mathbf{x}\| \tag{2.3}$$

for all $\mathbf{x} \in \mathbb{R}^{n_l}$.

Theorem 2.1 (Bramble, Pasciak, Wang, Xu [1], Theorem 1)**.** *Assume there are linear mappings $Q_l\colon U_1 \to U_l$, $l = 1,\dots,L$, $Q_1 = I$ such that for every $\mathbf{u} \in U_1$, the conditions*

$$\|(Q_l - Q_{l+1})\mathbf{u}\|_l \leqslant \frac{C_1}{\sqrt{\lambda_{\max}(A_l)}}\|\mathbf{u}\|_A \tag{2.4}$$

for all levels $l = 1,\dots,L-1$ and

$$\|Q_l\|_A \leqslant C_2 \tag{2.5}$$

for all levels $l = 1,\dots,L$ hold with positive constants C_1, C_2 independent of l and L. In addition, assume that the smoothers

$$\text{for } i = 1,\dots,\nu\colon\ \mathbf{x}^l \leftarrow \mathcal{S}_l(\mathbf{x}^l,\mathbf{f}^l)$$

in Algorithm 1 have the form $\mathbf{x}^l \leftarrow (I - R_l A_l)\mathbf{x}^l + R_l \mathbf{f}^l$*, where* R_l *are symmetric positive definite matrices that satisfy*

$$\lambda_{\min}(I - R_l A_l) \geqslant 0 \quad \textit{and} \quad \lambda_{\min}(R_l) \geqslant \frac{1}{C_3^2 \lambda_{\max}(A_l)} \tag{2.6}$$

for all $l = 1, \ldots, L-1$*, with constant* $C_3 > 0$ *independent of* l *and* L*. Then Algorithm 1 satisfies*

$$\|A^{-1}\mathbf{f} - MG(\mathbf{x}, \mathbf{f})\|_A \leqslant \Big(1 - \frac{1}{(1 + C_2 + C_1 C_3)^2 (L-1)}\Big) \|A^{-1}\mathbf{f} - \mathbf{x}\|_A.$$

The preconditioner $P\colon\ \mathbf{x} \mapsto MG(\mathbf{0}, \mathbf{x})$ *satisfies*

$$\operatorname{cond}(A, P) \leqslant (1 + C_2 + C_1 C_3)^2 (L-1).$$

3. Smoothed aggregation method

In the smoothed aggregation method ([5], [4], [6], [7], [9]), the prolongators $I_{l+1}^l\colon \mathbb{R}^{n_{l+1}} \to \mathbb{R}^{n_l}$ are constructed as the products

$$I_{l+1}^l = S_l P_{l+1}^l, \tag{3.1}$$

where $P_{l+1}^l\colon\ \mathbb{R}^{n_{l+1}} \to \mathbb{R}^{n_l}$ is a *tentative prolongator* obtained by generalized aggregation [8] and $S_l\colon\ \mathbb{R}^{n_l} \to \mathbb{R}^{n_l}$ is a sparse linear *prolongator smoother*. The tentative prolongator P_l is an orthogonal matrix with a disjoint column (or block column) structure responsible for the approximation. The prolongator smoother S_l enforces the smoothness of the coarse-space basis functions, or equivalently, reduces the coarse-level matrix spectral bounds

$$\varrho(A_{l+1}) = \varrho((I_{l+1}^l)^{\mathrm{T}} A_l I_{l+1}^l) = \varrho((P_{l+1}^l)^{\mathrm{T}} S_l^{\mathrm{T}} A_l S_l P_{l+1}^l) \leqslant \varrho(S_l^{\mathrm{T}} A_l S_l) \tag{3.2}$$

(the last inequality holds because P_{l+1}^l is an orthogonal matrix). The minimisation of the spectral bounds of the coarse-level matrices is desirable because of the key assumption (2.4) of the multi-level convergence theory [1] (Theorem 2.1). The smaller $\lambda_{\max}(A_l)$, the easier it becomes to satisfy (2.4) with a good (small) constant C_1. So generally, a more complex smoothing procedure S_l leads to better convergence [2], but increases the fill-in of the coarse-level matrices. In our multilevel method [7], we use the tentative prolongator given by generalized aggregations obtained by a coarsening by a factor of about three in all spatial directions, and S_l being the error propagation operator of a single Richardson sweep. This leads to sparse coarse-level matrices and guarantees a nearly optimal convergence bound for second-order elliptic problems. Thus, we use the prolongator smoother

$$S_l = I - \frac{\omega}{\lambda_l} A_l \tag{3.3}$$

with ω chosen so that it minimizes $\varrho(S_l^2 A_l)$, and λ_l being an available upper bound of $\varrho(A_l)$. This, in turn, reduces the spectral bound $\varrho(A_{l+1})$ in (3.2). To be more precise, choosing $\omega = 4/3$, (3.2), the spectral mapping theorem and substitution $\xi = t/\lambda_l$ yield (the symbol $\sigma(A_l)$ denotes the spectrum of A_l)

$$\begin{aligned}\varrho(A_{l+1}) \leqslant \varrho(S_l^2 A_l) &= \max_{t\in\sigma(A_l)} \Big(1 - \frac{4}{3}\frac{1}{\lambda_l} t\Big)^2 t \leqslant \max_{t\in[0,\lambda_l]} \Big(1 - \frac{4}{3}\frac{1}{\lambda_l} t\Big)^2 t \\ &= \lambda_l \max_{t\in[0,\lambda_l]} \Big(1 - \frac{4}{3}\frac{1}{\lambda_l} t\Big)^2 \frac{t}{\lambda_l} = \lambda_l \max_{\xi\in[0,1]} \Big(1 - \frac{4}{3}\xi\Big)^2 \xi = \frac{1}{9}\lambda_l.\end{aligned}$$

Thus, as λ_{l+1} we can take

$$\lambda_{l+1} = \min\Big\{\frac{\lambda_l}{9}, \widehat{\lambda}_{l+1}\Big\}, \tag{3.4}$$

where $\widehat{\lambda}_{l+1}$ is an upper bound of $\varrho(A_{l+1})$ obtained computationally. Then

$$\Big(\frac{1}{9}\Big)^{l-1} \lambda_1 \geqslant \lambda_l \geqslant \varrho(A_l). \tag{3.5}$$

We continue with the description of a simple aggregation procedure suitable for solving scalar elliptic problems. For more general form of aggregations that can be used for solving systems of partial differential equations (e.g. elasticity), we refer to [7].

We start with a simple example.

Example 1. Consider the simplest tentative prolongator P_2^1 for the 1D Laplace equation discretized on a mesh consisting of $n_1 = 3n_2$ nodes:

$$
P_2^1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & & \cdot & \\ 1 & & \cdot & \\ 1 & & \cdot & \\ & 1 & \cdot & \\ & 1 & \cdot & \\ & 1 & \cdot & \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ & & \cdot & 1 \\ & & \cdot & 1 \\ & & \cdot & 1 \end{bmatrix}. \tag{3.6}
$$

The columns of P_2^1 are (up to the scaling) 0/1 vectors with disjoint nonzero structure. Each column corresponds to disaggregation of one $\mathbb{R}^{n_2}$ variable into three $\mathbb{R}^{n_1}$ variables; the nonzero structure of the jth column corresponds to the jth *aggregate* of the system

$$\{\{1, 2, 3\}, \{4, 5, 6\}, \ldots, \{n_1 - 2, n_1 - 1, n_1\}\}$$

forming the disjoint covering of the set $\{1, \ldots, n_1\}$. So, P_2^1 can be thought of as a piece-wise constant interpolation in a discrete sense. Obviously, P_2^1 is an orthogonal matrix.

Let $\{\mathcal{A}_i^l\}$ be the system of aggregates forming a disjoint covering of the set $\{1, \ldots, n_l\}$. As in Example 1 we strive to form the aggregates so that each aggregate contains closely coupled degrees of freedom. The algorithm that generates the aggregates $\{\mathcal{A}_i^l\}$ that correspond to the coarsening by a factor of about 3 in each spatial direction, using the nonzero structure of the matrix A_l, is given in [9]. We set n_{l+1} to be the number of aggregates $\mathcal{A}_i^l$. As in Example 1, we first construct the tentative prolongator by

$$
(P_{l+1}^l)_{ij} = \begin{cases} 1 & \text{for } i \in \mathcal{A}_j^l, \\ 0 & \text{otherwise.} \end{cases} \tag{3.7}
$$

To maintain the orthogonality of P_{l+1}^l, we use the following scaling procedure: Set $\mathbf{b}^1 = (1, 1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^n$.

Algorithm 2. Given the system of aggregates $\{\mathcal{A}_i^l\}_{i=1}^{n_{l+1}}$ and the vector $\mathbf{b}^l \in \mathbb{R}^{n_l}$, construct the tentative prolongator $P_{l+1}^l \colon \mathbb{R}^{n_{l+1}} \to \mathbb{R}^{n_l}$ and the vector $\mathbf{b}^{l+1} \in \mathbb{R}^{n_{l+1}}$ so that $P_{l+1}^l \mathbf{b}^{l+1} = \mathbf{b}^l$ as follows:

- ▷ Construct P_{l+1}^l given by (3.7),
- ▷ set $P_{l+1}^l \leftarrow \operatorname{diag}(\mathbf{b}^l) P_{l+1}$,
- ▷ construct the diagonal $n_{l+1} \times n_{l+1}$ matrix $B_{l+1} = (P_{l+1}^l)^{\mathrm{T}} P_{l+1}^l$,
- ▷ set $P_{l+1}^l \leftarrow P_{l+1}^l B_{l+1}^{-1/2}$,
- ▷ set $\mathbf{b}^{l+1}$ to be a vector consisting of the diagonal entries of the matrix $B_{l+1}^{1/2}$,

$$\mathbf{b}^{l+1} = ((B_{l+1})_{11}^{1/2}, (B_{l+1})_{22}^{1/2}, \ldots, (B_{l+1})_{n_{l+1} n_{l+1}}^{1/2})^{\mathrm{T}}.$$

Let $P_l^1 = P_2^1 \ldots P_l^{l-1}$. Clearly, the resulting prolongators P_{l+1}^l and P_l^1 are orthogonal matrices,

$$P_{l+1}^l \mathbf{b}^{l+1} = \mathbf{b}^l \quad \text{and} \quad P_l^1 \mathbf{b}^l = (1, 1, \ldots, 1)^{\mathrm{T}}. \tag{3.8}$$

4. Avoiding the need of the equivalence of the discrete and continuous L_2-norms on the coarse levels

The direct use of Theorem 2.1 invokes the need of the equivalence of the discrete and the continuous L_2-norms on the coarse-levels. To be more precise, to verify (2.4) directly, one needs to establish the equivalence

$$\|I_l^1 \mathbf{x}\|_l \equiv \sqrt{\mathbf{x}^{\mathrm{T}} \mathbf{x}} \approx \text{scaling} \, \|\pi_h I_l^1 \mathbf{x}\|_{L_2}, \tag{4.1}$$

where π_h is the finest level finite element interpolation operator that takes a finest level vector and returns the corresponding finite element function. (The proof then proceeds by proving the interpolation in the L_2-norm and using the above equivalence to get the estimate for $\|\cdot\|_l$.) It is very difficult to establish such an equivalence and its proof is so far restricted to the cases with a model geometry, see [3].

In this short section we avoid the need for this equivalence by a simple trick on the operator level. As a result, we will need only the equivalence of the discrete and the continuous L_2-norms for purely disaggregated functions that holds trivially, because the disaggregation operators are (unlike the prolongators I_l^1 in (4.1)) orthogonal matrices. In other words, due to the orthogonality of P_l^1 and the fact that $\|\pi_h \mathbf{x}\|_{L_2} \approx$ scaling$\sqrt{\mathbf{x}^{\mathrm{T}} \mathbf{x}}$, the equivalence (4.1) with I_l^1 replaced by P_l^1 holds trivially.

Let l be a level. We choose linear operators $\widetilde{Q}_l\colon \mathbb{R}^{n_1} \to \mathbb{R}^{n_l}$, $\widetilde{Q}_1 = I$, $l = 1, \ldots, L$ and define the mappings

$$(4.2) \qquad Q_l\colon \mathbf{u} \in U_1 \mapsto I_l^1 \widetilde{Q}_l \mathbf{u} \in U_l \quad \text{and} \quad Q_l^P\colon \mathbf{u} \in U_1 \mapsto P_l^1 \widetilde{Q}_l \mathbf{u}.$$

Theorem 4.1. *Let S be given by (3.3) with $\omega \in (0,2)$, let P_{j+1}^j, $j = 1, \ldots, L-1$ be orthogonal matrices and $\lambda_l \geqslant \lambda_{\max}(A_l)$, $l = 1, \ldots, L$. Assume the operators $\widetilde{Q}_l$ are chosen so that $\widetilde{Q}_1 = I$ and the corresponding mappings Q_l and Q_l^P defined by (4.2) satisfy*

$$(4.3) \qquad \forall\, l = 1, \ldots, L-1, \quad \mathbf{u} \in U_1\colon \ \|(Q_l^P - Q_{l+1}^P)\mathbf{u}\| \leqslant \frac{C_P}{\sqrt{\lambda_l}} \|\mathbf{u}\|_A$$

and

$$(4.4) \qquad \forall\, l = 1, \ldots, L\colon \ \|Q_l\|_A \leqslant C_E$$

with constants C_P and C_E independent of l and L. Then (2.4) and (2.5) are satisfied with $C_1 = C_P + \omega C_E$ and $C_2 = C_E$.

P r o o f. We estimate using $I_l^1 = I_2^1 \ldots I_l^{l-1}$, (3.1), (2.3), (3.3), $\varrho(S_l) < 1$, triangle inequality, $P_l^1 = P_2^1 \ldots P_l^1$, the orthogonality of P_l^1 and Galerkin isometry $\|I_l^1 \mathbf{x}\|_A = \|\mathbf{x}\|_{A_l}$:

$$(4.5) \qquad \begin{aligned} \|(Q_l - Q_{l+1})\mathbf{u}\|_l &= \|I_l^1(\widetilde{Q}_l - I_{l+1}^l \widetilde{Q}_{l+1})\mathbf{u}\|_l \\ &= \|I_l^1(\widetilde{Q}_l - S_l P_{l+1}^l \widetilde{Q}_{l+1})\mathbf{u}\|_l \\ &\leqslant \|(\widetilde{Q}_l - S_l P_{l+1}^l \widetilde{Q}_{l+1})\mathbf{u}\| \\ &= \Bigl\| S_l(\widetilde{Q}_l - P_{l+1}^l \widetilde{Q}_{l+1})\mathbf{u} + \frac{\omega}{\lambda_l} A_l \widetilde{Q}_l \mathbf{u} \Bigr\| \\ &\leqslant \|S_l(\widetilde{Q}_l - P_{l+1}^l \widetilde{Q}_{l+1})\mathbf{u}\| + \frac{\omega}{\lambda_l} \|A_l \widetilde{Q}_l \mathbf{u}\| \\ &\leqslant \|(\widetilde{Q}_l - P_{l+1}^l \widetilde{Q}_{l+1})\mathbf{u}\| + \frac{\omega}{\sqrt{\lambda_l}} \|\widetilde{Q}_l \mathbf{u}\|_{A_l} \\ &= \|P_l^1(\widetilde{Q}_l - P_{l+1}^l \widetilde{Q}_{l+1})\mathbf{u}\| + \frac{\omega}{\sqrt{\lambda_l}} \|Q_l \mathbf{u}\|_A \\ &= \|(Q_l^P - Q_{l+1}^P)\mathbf{u}\| + \frac{\omega}{\sqrt{\lambda_l}} \|Q_l \mathbf{u}\|_A. \end{aligned}$$

The estimate (2.4) with $C_1 = C_P + \omega C_E$ now follows by (4.3) and (4.4). The estimate (2.5) with $C_2 = C_E$ follows immediately from (4.4). □

Remark 4.2. The abstract theory of [7] uses condition (4.3) to verify (4.4). This is the reason why the convergence result depends on the power of three of L. Here, we verify (4.4) directly for smoothed aggregation functions and get a convergence result dependent on the first power of L. Certainly, by doing so, we loose the elegance of the abstract assumption of [7], but in our opinion, very little on the strength of the convergence result. As becomes clear in the next section, we will need to control the growth of the diameter of the supports of the coarse-level basis functions (that is needed anyway for the computational complexity reasons), while in [7] we needed to control the diameter of the aggregates.

5. Verification of the assumptions of Theorem 4.1 for scalar elliptic problem

Let $\Omega \subset \mathbb{R}^d$, $d = 2$ or $d = 3$ be a polygon or polytope. Assume $\Gamma_D \subset \partial\Omega$ and $\mu(\Gamma_D) > 0$. (In case Γ_D is not connected, we assume all connected fragments have a positive measure.) Consider a variational problem

$$\text{find } u \in H^1_{0,\Gamma_D}(\Omega)\colon\ a(u,v) = f(v), \quad v \in H^1_{0,\Gamma_D}(\Omega). \tag{5.1}$$

Here, $a(\cdot,\cdot)$ is a symmetric bilinear form coercive and continuous on

$$H^1_{0,\Gamma_D}(\Omega) = \{u \in H^1(\Omega)\colon\ \operatorname{tr} u = 0 \text{ on } \Gamma_D,\ |\cdot|_{H^1(\Omega)}\}$$

and $f(\cdot) \in (H^1_{0,\Gamma_D}(\Omega))^{-1}$. We consider a quasi-uniform triangulation τ_h of Ω with characteristic mesh-size h and boundaries of elements aligned with Γ_D. Let $V_h = \operatorname{span}\{\varphi_i\}_{i=1}^n \subset H^1_{0,\Gamma_D}(\Omega)$ be the corresponding $P1$ finite element space with the finite element basis functions scaled so that $\|\varphi_i\|_{L_\infty(\Omega)} = 1$. We assume the linear system (2.1) arose by the standard conforming finite element discretization of (5.1), that is, by replacing $H^1_{0,\Gamma_D}(\Omega)$ in (5.1) by V_h.

Let $\pi_h\colon\ \mathbf{x} \in \mathbb{R}^n \mapsto \sum_i x_i\varphi_i$ and let $\mathbf{e}_i$ be the i-th canonical basis vector of $\mathbb{R}^{n_l}$. We define the basis on the level l by

$$\varphi_i^l = \pi_h I_l^1 \mathbf{e}_i, \quad i = 1,\dots,n_l,\ l = 1,\dots,L. \tag{5.2}$$

The following is our key assumption on the geometry of coarse spaces.

Assumption 5.1. For every basis function φ_i^l there is a domain $B_i^l \subset \Omega$ being an intersection of a ball, with Ω satisfying $B_i^l \supset \operatorname{supp}\varphi_i^l$ such that for each level (an integer $l \in [1,\ L]$), the domains B_i^l, $i = 1,\dots,n_l$, satisfy

1. $\mathrm{diam}(B_i^l) \leqslant C3^{l-1}h$ with C independent of l and i,
2. there is an integer N independent of l such that for each level l, each point $\mathbf{x} \in \Omega$ belongs to at most N domains B_i^l.

Our first goal is to verify (4.4). First we prove that apart from the essential boundary conditions, the coarse-level basis functions form a decomposition of unity. Define

$$\mathcal{I}(\Omega_{D,l}) = \{i\colon\ \overline{B}_i^l \cap \Gamma_D \neq \emptyset\} \quad \text{and} \quad \Omega_{D,l} = \bigcup_{i \in \mathcal{I}(\Omega_{D,l})} B_i^l. \tag{5.3}$$

Lemma 5.2. *For vector $\mathbf{b}^l$ created by Algorithm 2 and the basis $\{\varphi_i^l\}_{i=1}^{n_l}$ on the level l we have*

$$\sum_{i=1}^{n_l} b_i^l \varphi_i^l = 1 \quad \text{on } \Omega \setminus \Omega_{D,l}. \tag{5.4}$$

P r o o f. Aside from the essential boundary conditions, the basis functions on any level satisfy (5.4) and $\mathbf{b}^l \in \ker(A_l)$. Indeed, assume we solve the pure Neumann problem ($\Gamma_D = \emptyset$). Since the unit function belongs to the kernel of $H^1(\Omega)$-seminorm, $\|\mathbf{x}\|_A \approx |\pi_h \mathbf{x}|_{H^1(\Omega)}$ and the finest level basis functions φ_i^1 satisfy $\|\varphi_i^1\|_{L_\infty(\Omega)} = 1$, we have $\mathbf{b}^1 = (1, 1, \ldots, 1)^{\mathrm{T}} \in \ker(A_1)$. Assume $\mathbf{b}^l \in \ker(A_l)$. By (3.8), $P_{l+1}^l \mathbf{b}^{l+1} = \mathbf{b}^l$. Since $\mathbf{b}^l \in \ker(A_l)$, we have

$$I_{l+1}^l \mathbf{b}^{l+1} = (I - \alpha A_l) P_{l+1}^l \mathbf{b}^{l+1} = (I - \alpha A_l)\mathbf{b}^l = \mathbf{b}^l. \tag{5.5}$$

Hence $A_{l+1}\mathbf{b}^{l+1} = (I_{l+1}^l)^{\mathrm{T}} A_l I_{l+1}^l \mathbf{b}^{l+1} = (I_{l+1}^l)^{\mathrm{T}} A_l \mathbf{b}_l = \mathbf{0}$. Thus, $\mathbf{b}^{l+1} \in \ker(A_{l+1})$. Hence by induction, $\mathbf{b}^l \in \ker(A_l)$ holds for all levels l and as a consequence, (5.5) holds for all levels $l = 1, \ldots, L-1$. By (5.5), $\sum_{i=1}^{n_l} b_i^l \varphi_i^l = \sum_{i=1}^{n_l} \pi_h I_l^1 b_i^l \mathbf{e}_i = \pi_h I_l^1 \mathbf{b}^l = \pi_h \mathbf{b}^1 = 1$.

Let $\Gamma_D \neq \emptyset$ again. If (5.4) is violated at the point $\mathbf{x} \in \Omega$, there must be a basis function φ_j^l such that $\mathbf{x} \in \operatorname{supp} \varphi_j^l$ and φ_j^l is influenced by zero value on Γ_D. This means that on some level $1 \leqslant k < l$ there is an aggregate $\mathcal{A}_p^k$, $p \in \operatorname{supp} I_l^{k+1} \mathbf{e}_j$ (support of the vector is understood as the list of indices of its nonzero entries) that contains a degree of freedom q directly adjacent to Γ_D in the sense that $\partial \operatorname{supp} \varphi_q^k \cap \Gamma_D \neq \emptyset$. Clearly, $q \in \operatorname{supp} I_l^{k+1} \mathbf{e}_j$ and

$$\operatorname{supp} \varphi_j^l = \bigcup_{i \in \operatorname{supp} I_l^k \mathbf{e}_j} \operatorname{supp} \varphi_i^k \supset \operatorname{supp} \varphi_q^k.$$

Hence, $\partial \operatorname{supp} \varphi_j^l \cap \Gamma_D \neq \emptyset$. Thus, we proved that if (5.4) is violated at $\mathbf{x} \in \Omega$, then there is a basis function φ_j^l such that $\partial \operatorname{supp} \varphi_j^l \cap \Gamma_D \neq \emptyset$, hence $\overline{B}_i^l \cap \Gamma_D \neq \emptyset$ and (5.4) follows. □

Assume the system of aggregates $\{\mathcal{A}_i^l\}_{i=1}^{n_l}$, $l=1,\dots,L-1$ is given. We introduce *composite aggregates* $\widetilde{\mathcal{A}}_i^l$ to be the aggregates $\mathcal{A}_i^l$ understood as the corresponding sets of degrees of freedom on the level 1. Formally, the composite aggregates are defined recursively as

$$\widetilde{\mathcal{A}}_i^l \equiv \widetilde{\mathcal{A}}_i^{l,1} \quad \text{with } \widetilde{\mathcal{A}}_i^{l,j-1} = \bigcup_{k\in\widetilde{\mathcal{A}}_i^{j,l}} \mathcal{A}_k^{j-1}. \tag{5.6}$$

Further, define the discrete l_2-(semi)norm of the vector $\mathbf{x}\in\mathbb{R}^n$ by

$$\|\mathbf{x}\|_{l_2(\widetilde{\mathcal{A}}_i^l)} = \Bigl(\sum_{j\in\widetilde{\mathcal{A}}_i^l} x_j^2\Bigr)^{1/2}.$$

Clearly, the composite tentative prolongator P_l^1 has a disjoint nonzero column structure corresponding to the aggregates $\widetilde{\mathcal{A}}_i^l$; for the i-th column $P_l^1\mathbf{e}_i$ of P_l^1, we have (denoting by $\mathbf{e}_i$ the i-th canonical basis vector of $\mathbb{R}^{n_l}$)

$$\operatorname{supp} P_l^1\mathbf{e}_i = \widetilde{\mathcal{A}}_i^{l-1}, \quad \operatorname{supp} P_l^1\mathbf{e}_i \cap \operatorname{supp} P_l^1\mathbf{e}_j = \emptyset \quad \text{for } i\neq j,$$

as the aggregates $\widetilde{\mathcal{A}}_i^{l-1}$ and $\widetilde{\mathcal{A}}_j^{l-1}$ are disjoint.

For every domain B_i^l, define an index set $\mathcal{I}(\widehat{B}_i^l)=\{j\colon\ B_j^l\cap B_i^l\neq\emptyset\}$ and consider a domain $\widehat{B}_i^l$: $\Omega\supset\widehat{B}_i^l\supset \bigcup_{j\in\mathcal{I}(\widehat{B}_i^l)} B_j^l$. From Assumption 5.1 it follows that it is possible to choose domains $\widehat{B}_i^l$ so that $\operatorname{diam}(\widehat{B}_i^l)\leqslant C3^{l-1}h$ and each $\mathbf{x}\in\mathbb{R}^d$ is contained in at most N domains $\widehat{B}_i^l$ (with C and N different from the ones in Assumption 5.1). For all $i\notin\mathcal{I}(\Omega_{D,l})$, define the local interpolation operators Π_i^l: $H^1(\widehat{B}_i^l)\to H^1(B_i^l)$ by

$$\Pi_i^l u = \sum_{j\in\mathcal{I}(\widehat{B}_i^l)} \Bigl(\frac{\operatorname{card}(\widetilde{\mathcal{A}}_j^{l-1})^{1/2}}{\mu(B_j^l)}\int_{B_j^l} u\,\mathrm{d}\Omega\Bigr)\varphi_i^l \tag{5.7}$$

and the global interpolation operator $\Pi^l\colon\ H^1_{0,\Gamma_D}(\Omega)\to\operatorname{span}\{\varphi_{\mathrm{i}}^{\mathrm{l}}\}\subset \mathrm{H}^1_{0,\Gamma_\mathrm{D}}(\Omega)$

$$\Pi^l u = \sum_{i=1}^{n_l}\Bigl(\frac{\operatorname{card}(\widetilde{\mathcal{A}}_i^{l-1})^{1/2}}{\mu(B_i^l)}\int_{B_i^l} u\,\mathrm{d}\Omega\Bigr)\varphi_i^l. \tag{5.8}$$

We define $\mathcal{I}(\widehat{\Omega}_{D,l})$ to be the index set of all domains B_i^l intersecting $\Omega_{D,l}$ and consider a domain $\widehat{\Omega}_{D,l}$: $\Omega\supset\widehat{\Omega}_{D,l}\supset\bigcup_{i\in\mathcal{I}(\widehat{\Omega}_{D,l})} B_i^l$. Clearly, by Assumption 5.1, it is possible to choose $\widehat{\Omega}_{D,l}$ so that $\operatorname{dist}(\mathbf{x},\Gamma_D) \leqslant Ch3^{l-1}$ for all $\mathbf{x} \in \widehat{\Omega}_{D,l}$ and we have by the Friedrichs inequality

$$\|u\|_{L_2(\widehat{\Omega}_{D,l})} \leqslant Ch_l|u|_{H^1(\widehat{\Omega}_{D,l})}, \quad h_l = 3^{l-1}h \tag{5.9}$$

for all $u \in H^1_{0,\Gamma_D}(\Omega)$.

Next we prove that Π^l is H^1-seminorm stable on a boundary layer adjacent to Γ_D.

Lemma 5.3. *Assume λ_1 has been obtained by the Gershgorin theorem. Then there is a constant $C>0$ independent of h, l and L such that*

$$\forall\, u \in H^1_{0,\Gamma_D}(\Omega)\colon\ |\Pi^l u|_{H^1(\Omega_{D,l})} \leqslant C|u|_{H^1(\widehat{\Omega}_{D,l})}. \tag{5.10}$$

Proof. By bounded intersections of domains B^l_i and the Cauchy-Schwarz inequality we get

$$\begin{aligned}
|\Pi^l u|^2_{H^1(\Omega_{D,l})} &\leqslant C \sum_{i\in\mathcal{I}(\Omega_{D,l})} |\Pi^l_i u|^2_{H^1(B^l_i)} \\
&= C \sum_{i\in\mathcal{I}(\Omega_{D,l})} \left| \sum_{j\in\mathcal{I}(\widehat{B}^l_i)} \left(\frac{\operatorname{card}(\widetilde{\mathcal{A}}^{l-1}_j)^{1/2}}{\mu(B^l_j)} \int_{B^l_j} u\,\mathrm{d}\Omega \right) \varphi^l_j \right|^2_{H^1(B^l_i)} \\
&\leqslant C \sum_{i\in\mathcal{I}(\Omega_{D,l})} \sum_{j\in\mathcal{I}(\widehat{B}^l_i)} \left(\frac{\operatorname{card}(\widetilde{\mathcal{A}}^{l-1}_j)^{1/2}}{\mu(B^l_j)} \int_{B^l_j} u\,\mathrm{d}\Omega \right)^2 |\varphi^l_j|^2_{H^1(\Omega)} \\
&\leqslant C \sum_{i\in\mathcal{I}(\widehat{\Omega}_{D,l})} \left(\frac{\operatorname{card}(\widetilde{\mathcal{A}}^{l-1}_i)^{1/2}}{\mu(B^l_i)} \int_{B^l_i} u\,\mathrm{d}\Omega \right)^2 |\varphi^l_i|^2_{H^1(\Omega)} \\
&= C \sum_{i\in\mathcal{I}(\widehat{\Omega}_{D,l})} \left(\frac{\operatorname{card}(\widetilde{\mathcal{A}}^{l-1}_i)^{1/2}}{\mu(B^l_i)} (u,1)_{L_2(B^l_i)} \right)^2 |\varphi^l_i|^2_{H^1(\Omega)} \\
&\leqslant C \sum_{i\in\mathcal{I}(\widehat{\Omega}_{D,l})} \left(\frac{\operatorname{card}(\widetilde{\mathcal{A}}^{l-1}_i)^{1/2}}{\mu(B^l_i)} \|u\|_{L_2(B^l_i)} \|1\|_{L_2(B^l_i)} \right)^2 |\varphi^l_i|^2_{H^1(\Omega)} \\
&\leqslant C \sum_{i\in\mathcal{I}(\widehat{\Omega}_{D,l})} \frac{\operatorname{card}(\widetilde{\mathcal{A}}^{l-1}_i)}{\mu(B^l_i)} \|u\|^2_{L_2(B^l_i)} |\varphi^l_i|^2_{H^1(\Omega)}.
\end{aligned} \tag{5.11}$$

Further, taking the bound obtained by the Gershgorin theorem for λ_1, we have

$$\lambda_1 \leqslant Ch^{d-2}$$

and by (3.5) we get

$$|\varphi_i^l|^2_{H^1(\Omega)} \leqslant \varrho(A_l) \leqslant \Big(\frac{1}{9}\Big)^{l-1} \lambda_1 \leqslant C\Big(\frac{1}{9}\Big)^{l-1} h^{d-2}. \tag{5.12}$$

In addition, since $\widetilde{\mathcal{A}}_i^l \subset B_i^l$ and the finest level mesh is quasiuniform, it also holds that

$$\frac{\operatorname{card}(\widetilde{\mathcal{A}}_i^{l-1})}{\mu(B_i^l)} \leqslant Ch^{-d}. \tag{5.13}$$

Substituting (5.13) and (5.12) into (5.11) and using $h_l = 3^{l-1}h$ and (5.9) yields

$$|\Pi^l u|^2_{H^1(\Omega_{D,l})} \leqslant Ch_l^{-2} \sum_{i\in\mathcal{I}(\widehat{\Omega}_{D,l})} \|u\|^2_{L_2(B_i^l)} \leqslant Ch_l^{-2}\|u\|^2_{L_2(\widehat{\Omega}_{D,l})} \leqslant C|u|^2_{H^1(\widehat{\Omega}_{D,l})}.$$

□

Next we prove that the local interpolation operator Π_i^l, $i \notin \mathcal{I}(\Omega_{D,l})$ preserves a constant.

Lemma 5.4. *Let $i \notin \mathcal{I}(\Omega_{D,l})$ and let c be a constant function on $\widehat{B}_i^l$. Then*

$$\Pi_i^l c = c \quad \text{on } B_i^l. \tag{5.14}$$

(Note: it is irrelevant that potentially, c, being a constant function on $\widehat{B}_i^l$, does not belong to $H^1_{0,\Gamma}(\Omega)$, since Π_i^l is understood as a mapping from $H^1(\widehat{B}_i^l)$ to $H^1(B_i^l)$.)

Proof. By (3.8) and from the nonzero structure and orthogonality of the composite tentative prolongator P_l^1, it follows that the vector $\mathbf{b}^l$ produced by Algorithm 2 satisfies $b_j^l = \operatorname{card}(\widetilde{\mathcal{A}}_j^{l-1})^{1/2}$. We use Lemma 5.2 and the fact that only the supports of the basis functions φ_j^l, $j \in \mathcal{I}(\widehat{B}_i^l)$ intersect B_i^l:

$$\begin{aligned}(\Pi_i^l c)|_{B_i^l} &= \sum_{j\in\mathcal{I}(\widehat{B}_i^l)} \frac{\operatorname{card}(\widetilde{\mathcal{A}}_j^{l-1})^{1/2}}{\mu(B_i^l)} \int_{B_j^l} c\,\mathrm{d}\Omega\, \varphi_j^l|_{B_i^l} = c \sum_{j\in\mathcal{I}(\widehat{B}_i^l)} \operatorname{card}(\widetilde{\mathcal{A}}_j^{l-1})^{1/2}\varphi_j^l|_{B_i^l} \\ &= c \sum_{j\in\mathcal{I}(\widehat{B}_i^l)} b_j^l\varphi_j^l|_{B_i^l} = c\sum_{j=1}^{n_l} \varphi_j^l|_{B_i^l} = c.\end{aligned}$$

□

In the next lemma we show that the local interpolation operator Π_i^l, $i \notin \mathcal{I}(\Omega_{D,l})$ is H^1-seminorm stable.

Lemma 5.5. *Let $i \notin \mathcal{I}(\Omega_{D,l})$. Then there is a constant $C > 0$ independent of h, i, l and L such that for every $u \in H^1_{0,\Gamma_D}(\Omega)$*

$$|\Pi_i^l u|_{H^1(B_i^l)} \leqslant C|u|_{H^1(\widehat{B}_i^l)}. \tag{5.15}$$

Proof. Set $c = \operatorname*{argmin}_{q\in\mathbb{R}} \|u-q\|_{L_2(\widehat{B}_i^l)}$ and $\hat{u} = u - c$. Since $\operatorname{diam}(\widehat{B}_i^l) \leqslant Ch_l$, $h_l = h3^{l-1}$, the Poincaré inequality yields

$$\|\hat{u}\|_{L_2(\widehat{B}_i^l)} \leqslant Ch_l|\hat{u}|_{H^1(\widehat{B}_i^l)}. \tag{5.16}$$

By Lemma 5.4 we get

$$|\Pi_i^l u|_{H^1(B_i^l)} = |\Pi_i^l u - c|_{H^1(B_i^l)} = |\Pi_i^l (u-c)|_{H^1(B_i^l)} = |\Pi_i^l \hat{u}|_{H^1(B_i^l)}. \tag{5.17}$$

Further, by the definition of Π_i^l, (5.12) (5.13), the Cauchy-Schwarz inequality and the Poincaré inequality (5.16) we get

$$\begin{aligned}
|\Pi_i^l \hat{u}|^2_{H^1(B_i^l)} &= \Bigg| \sum_{j\in\mathcal{I}(\widehat{B}_i^l)} \Bigg(\frac{\operatorname{card}(\widetilde{\mathcal{A}}_j^{l-1})^{1/2}}{\mu(B_j^l)} \int_{B_j^l} \hat{u}\,\mathrm{d}\Omega \Bigg)^2 \varphi_j^l \Bigg|^2_{H^1(B_i^l)} \\
&\leqslant C \sum_{j\in\mathcal{I}(\widehat{B}_i^l)} \Bigg(\frac{\operatorname{card}(\widetilde{\mathcal{A}}_j^{l-1})^{1/2}}{\mu(B_j^l)} \int_{B_j^l} \hat{u}\,\mathrm{d}\Omega \Bigg)^2 |\varphi_j^l|^2_{H^1(\Omega)} \\
&= C \sum_{j\in\mathcal{I}(\widehat{B}_i^l)} \Bigg(\frac{\operatorname{card}(\widetilde{\mathcal{A}}_j^{l-1})^{1/2}}{\mu(B_j^l)} (\hat{u}, 1)_{L_2(B_j^l)} \Bigg)^2 |\varphi_j^l|^2_{H^1(\Omega)} \\
&\leqslant C \sum_{j\in\mathcal{I}(\widehat{B}_i^l)} \Bigg(\frac{\operatorname{card}(\widetilde{\mathcal{A}}_j^{l-1})^{1/2}}{\mu(B_j^l)} \|\hat{u}\|_{L_2(B_j^l)} \|1\|_{L_2(B_j^l)} \Bigg)^2 |\varphi_j^l|^2_{H^1(\Omega)} \\
&\leqslant C \sum_{j\in\mathcal{I}(\widehat{B}_i^l)} \frac{\operatorname{card}(\widetilde{\mathcal{A}}_j^{l-1})}{\mu(B_j^l)} \|\hat{u}\|^2_{L_2(B_j^l)} |\varphi_j^l|^2_{H^1(\Omega)} \\
&\leqslant Ch_l^{-2} \sum_{j\in\mathcal{I}(\widehat{B}_i^l)} \|\hat{u}\|^2_{L_2(B_j^l)} \leqslant Ch_l^{-2} \|\hat{u}\|^2_{L_2(\widehat{B}_i^l)} \leqslant C|\hat{u}|^2_{H^1(\widehat{B}_i^l)} \\
&= C|u|^2_{H^1(\widehat{B}_i^l)}.
\end{aligned}$$

The statement (5.15) follows by the previous estimate together with (5.17). □

Define the operator $\widetilde{Q}_l\colon \mathbb{R}^{n_1} \to \mathbb{R}^{n_l}$ by

$$(\widetilde{Q}_l\mathbf{u})_i = \frac{\operatorname{card}(\widetilde{\mathcal{A}}_i^{l-1})^{1/2}}{\mu(B_i^l)} \int_{B_i^l} \pi_h \mathbf{u}\,\mathrm{d}\Omega, \quad i = 1,\dots,n_l. \tag{5.18}$$

Then the operator Q_l defined in (4.2) has a form $Q_l = \pi_h^{-1}\Pi^l\pi_h$, where Π^l is the global interpolation operator defined in (5.8).

In the next lemma we verify condition (4.4).

Lemma 5.6. *There is a constant $C > 0$ independent of h, l and L such that for every $u \in H^1_{0,\Gamma_D}(\Omega)$*

$$|\Pi^l\mathbf{u}|_{H^1(\Omega)} \leqslant C|u|_{H^1(\Omega)}. \tag{5.19}$$

As a consequence, the operator $Q_l = \pi_h^{-1}\Pi_l\pi_h$ satisfies (4.4).

Proof. The proof of (5.19) follows from Lemma 5.3 and Lemma 5.5 using bounded overlaps of the balls B_i^l and $\Omega_{D,l}$ and bounded overlaps of the balls $\widehat{B}_i^l$ and $\widehat{\Omega}_{D,l}$, by

$$\begin{aligned}|\Pi u|^2_{H^1(\Omega)} &\leqslant C\Bigg(|\Pi^l u|^2_{H^1(\Omega_{D,l})} + \sum_{i\notin\mathcal{I}(\Omega_{D,l})} |\Pi_i^l u|^2_{H^1(B_i^l)}\Bigg)\\ &\leqslant C\Bigg(|u|^2_{H^1(\widehat{\Omega}_{D,l})} + \sum_{i\notin\mathcal{I}(\Omega_{D,l})} |u|^2_{H^1(\widehat{B}_i^l)}\Bigg)\\ &\leqslant C|u|^2_{H^1(\Omega)}.\end{aligned}$$

The estimate (4.4) is a direct consequence of (5.19). Indeed,

$$\|Q_l\mathbf{u}\|_A = \|\pi_h^{-1}\Pi^l\pi_h\mathbf{u}\|_A \leqslant C|\Pi^l\pi_h\mathbf{u}|_{H^1(\Omega)} \leqslant C|\pi_h\mathbf{u}|_{H^1(\Omega)} \leqslant C\|\mathbf{u}\|_A.$$

□

Clearly, the operator Q_l^P defined in (4.2) returns a vector that is constant on each composite aggregate and on $\widehat{\mathcal{A}}_i^{l-1}$ has the value

$$(Q_l^P\mathbf{u})_j = \frac{1}{\mu(B_i^l)} \int_{B_i^l} \pi_h\mathbf{u}\,\mathrm{d}\Omega, \quad j \in \widetilde{\mathcal{A}}_i^{l-1}. \tag{5.20}$$

It remains to verify assumption (4.3).

Lemma 5.7. *There is a constant $C > 0$ independent of h, l and L such that* (4.3) *holds with $C_P = C$.*

P r o o f. We set q_i to be the value returned in (5.20) on the aggregate $\mathcal{A}_i^l$. Since $\operatorname{diam}(\widehat{B}_i^l) \leqslant Ch3^{l-1}$, we have by the Poincaré inequality

$$\|\pi_h \mathbf{u} - q_i\|_{L_2(\widehat{B}_i^l)} \leqslant Ch3^{l-1} |\pi_h \mathbf{u}|_{H^1(\widehat{B}_i^l)}.$$

We estimate using the above inequality, (5.20) and the fact that the composite aggregates form a disjoint covering of the set $\{1, \ldots, n\}$:

$$\begin{aligned}
\|\mathbf{u} - Q_l^P \mathbf{u}\|^2 &= \sum_{i=1}^{n_l} \|\mathbf{u} - Q_l^P \mathbf{u}\|^2_{l_2(\widehat{\mathcal{A}}_i^{l-1})} = \sum_{i=1}^{n_l} \|\mathbf{u} - q_i\|^2_{l_2(\widetilde{\mathcal{A}}_i^{l-1})} \\
&\leqslant Ch^{-d} \sum_{i=1}^{n_l} \|\pi_h \mathbf{u} - q_i\|^2_{L_2(\widehat{B}_i^l)} \leqslant C9^{l-1} h^{2-d} \sum_{i=1}^{n_l} |\pi_h \mathbf{u}|^2_{H^1(\widehat{B}_i^l)} \\
&\leqslant C9^{l-1} h^{2-d} |\pi_h|^2_{H^1(\Omega)} \leqslant \frac{C}{h^{d-2}/9^{l-1}} \|\mathbf{u}\|_A^2.
\end{aligned}$$

As λ_1 we take the estimate of $\varrho(A)$ obtained by the Gershgorin theorem, hence $\lambda_1 \leqslant Ch^{d-2}$. The estimate (4.3) now follows by (3.5). □

Now we are ready to formulate the final convergence theorem.

Theorem 5.8. *Let prolongators I_{l+1}^l be constructed by the smoothed aggregation method as described in Section* 3 *with the prolongator smoother given by* (3.3) *with $\omega = 4/3$ and λ_1 obtained by the Gershgorin theorem. Assume the multigrid smoothers satisfy* (2.6) *and the aggregates are such that Assumption* 5.1 *holds. Then Algorithm* 1 *converges with the rate of convergence*

$$\|A^{-1}\mathbf{f} - MG(\mathbf{x}, \mathbf{f})\|_A \leqslant \Big(1 - \frac{1}{(1+C)^2(L-1)}\Big) \|A^{-1}\mathbf{f} - \mathbf{x}\|_A.$$

The constant $C > 0$ is independent of h and L. In addition, the preconditioner $P\colon\ \mathbf{x} \mapsto MG(\mathbf{0}, \mathbf{x})$ satisfies $\operatorname{cond}(A, P) \leqslant (1 + C)^2(L - 1)$.

P r o o f. The proof follows directly from Theorem 2.1, Theorem 4.1 and Lemmas 5.7 and 5.6. □

References

[1] *J. H. Bramble, J. E. Pasciak, J. Wang, J. Xu*: Convergence estimates for multigrid algorithms without regularity assumptions. Math. Comput. *57* (1991), 23–45.

[2] *M. Brezina, P. Vaněk, P. S. Vassilevski*: An improved convergence analysis of smoothed aggregation algebraic multigrid. Numer. Linear Algebra Appl. *19* (2012), 441–469.

[3] *P. Fraňková, J. Mandel, P. Vaněk*: Model analysis of BPX preconditioner based on smoothed aggregations. Appl. Math., Praha *60* (2015), 219–250.

[4] *P. Vaněk*: Fast multigrid solver. Appl. Math., Praha *40* (1995), 1–20.

[5] *P. Vaněk*: Acceleration of convergence of a two-level algorithm by smoothing transfer operator. Appl. Math., Praha *37* (1992), 265–274.

[6] *P. Vaněk, M. Brezina*: Nearly optimal convergence result for multigrid with aggressive coarsening and polynomial smoothing. Appl. Math., Praha *58* (2013), 369–388.

[7] *P. Vaněk, M. Brezina, J. Mandel*: Convergence of algebraic multigrid based on smoothed aggregations. Numer. Math. *88* (2001), 559–579.

[8] *P. Vaněk, M. Brezina, R. Tezaur*: Two-grid method for linear elasticity on unstructured meshes. SIAM J. Sci Comput. *21* (1999), 900–923.

[9] *P. Vaněk, J. Mandel, R. Brezina*: Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems. Computing *56* (1996), 179–196.

Authors' address: Jan Brousek, Pavla Fraňková, Petr Vaněk, Department of mathematics, University of West Bohemia, Univerzitní 22, 306 14 Plzeň, Czech Republic, e-mail: brousek@kma.zcu.cz, frankova@ntis.zcu.cz, ptrvnk@kma.zcu.cz.