Convergence of Algebraic Multigrid Based on Smoothed Aggregation II: Extension to a Petrov-Galerkin Method

Petr Vanek, Ales Janka, Hervé Guillard

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Convergence of Algebraic Multigrid Based on Smoothed Aggregation II: Extension to a Petrov-Galerkin Method

Petr Vaněk — Aleš Janka — Hervé Guillard

N° 3683

Mai 1999

THÈME 4

Rapport de recherche

ISSN 0249-6399

Convergence of Algebraic Multigrid Based on Smoothed Aggregation II: Extension to a Petrov-Galerkin Method

Petr Vaněk* , Aleš Janka† , Hervé Guillard‡

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet Sinus

Rapport de recherche n° 3683 — Mai 1999 — 22 pages

Abstract: We give a convergence estimate for a Petrov-Galerkin Algebraic Multigrid method. In this method, the prolongations are defined using the concept of smoothed aggregation while the restrictions are simple aggregation operators. The analysis is carried out by showing that these methods can be interpreted as variational Ritz-Galerkin ones using modified transfer and smoothing operators. The estimate depends only on a weak approximation property for the aggregation operators. For a scalar second order elliptic problem using linear elements, this assumption is shown to hold using simple geometrical arguments on the aggregates.

Key-words: multigrid, finite elements, finite volumes, Aggregation methods, agglomeration methods

* Dpt of Math, University of California, Los Angeles, CA
† INRIA, BP. 93, 06902 Sophia Antipolis Cedex, France
‡ INRIA, BP. 93, 06902 Sophia Antipolis Cedex, France

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, B.P. 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : 04 92 38 77 77 - International : +33 4 92 38 77 77 —Fax : 04 92 38 77 65 - International : +33 4 92 38 77 65

Convergence d'une méthode multigrille algébrique basée sur le concept d'aggrégation lissée
II: Extension à une formulation de type Petrov Galerkin

Résumé : Nous donnons une borne sur la vitesse de convergence d'une méthode multigrille algébrique de type Petrov-Galerkin. Dans cette méthode, les opérateurs de prolongements sont construits en utilisant le concept d'aggrégation lissée tandis que les restrictions sont de simples opérateurs d'aggrégation. L'analyse de convergence est réalisée en montrant que ces méthodes peuvent être interprétées comme des méthodes de Ritz Galerkin en redéfinissant les opérateurs de prolongements et de lissage. L'estimation du taux de convergence dépend seulement d'une propriété d'approximation faible valable pour les opérateurs d'aggrégation. Pour une équation elliptique scalaire du second-ordre, on montre que cette propriété est vérifiée en utilisant des hypothèses simples sur la géométrie des aggrégats.

Mots-clés : Multigrille, éléments finis, volumes finis, aggrégation, agglomération

1 Introduction

Smoothed aggregation algebraic iterative methods, introduced in [12, 13] and further developed in [14, 15, 16, 17], have proved to be very efficient iterative methods for solving symmetric positive definite linear systems arizing from the finite element discretization of elliptic boundary value problems.

The smoothed aggregation based coarsening is performed in two stages. First we form small disjoint clusters of the fine-level degrees of freedom called *aggregates*. In the simplest case, each aggregate gives rise to a one column of the tentative prolongator; the column is created by restricting the vector of ones onto the aggregate. Such a procedure can be viewed as a piecewise constant coarsening in a discrete sense. The resulting prolongators have reasonable approximation properties, but their range contains high-energy vectors. For this reason, the first stage is followed by smoothing the range of prolongator by a Jacobi-type smoother.

Abstract convergence bounds for the variational smoothed aggregation multigrid method have been established in [17]. The variational framework however, calls for coarsening by a factor of 3 in each spatial direction. If the more usual coarsening by 2 is performed, the fill-in of coarse-level matrices gradually increases and the method becomes expensive both in terms of storage and computational complexity.

For a certain class of problems, such a limitation on the coarsening ratio is an inconvenient obstacle ([4, 9, 5, 6, 8].) Moreover, on the technical front, many of existing software packages use integrated multigrid solvers that perform standard coarsening by the factor 2. This makes difficult to change these solvers by more efficient smoothed aggregation ones.

In [4], we show that the smoothed aggregation method can be extended to any coarsening ratio, including even ones, by the choice of carefully selected smoothing polynomials. However, the resulting methods do not fit in the classical variational framework. The objective of this paper is to provide a convergence theory for these methods. This is done by showing that these methods can be interpreted as Ritz-Galerkin methods using modified transfer and smoothing operators.

For the sake of simplicity, we restrict our considerations to scalar second order elliptic problems discretized so that all degrees of freedom have the meaning of nodal function values. The abstract results (Theorem 3.1 and Lemma 4.3) however, do not rest on such a restriction. Treating more general elliptic problems and discretizations require more sophisticated tentative prolongators as described in [17].

The paper is organized as follows: In section 2, we present the iterative algorithms that we are to analyze, show that they can be interpreted as Petrov-Galerkin methods and define the *tentative prolongators* that will be used in the sequel. Section 3 reinterprets the Petrov-Galerkin algorithms as Ritz-Galerkin ones and shows that the modified smoothing operators used in this interpretation keep the smoothing properties of the original ones. Section 4 defines the smoothing polynomials used in the definition of the final prolongation operators, while in Section 5, we verify the assumptions on the tentative prolongator for discrete problems coming from the discretization of scalar elliptic problems.

2 Petrov-Galerkin MG method

We consider solving a linear system

$$A\mathbf{x} = \mathbf{b} \tag{1}$$

where A is a symmetric positive definite $n_1 \times n_1$ matrix and $\mathbf{b}$ is a vector of $I\!R^{n_1}$. We call $I\!R^{n_1}$ the finest level and pose $A_1 = A$ .The generation of the coarse levels $l = 2, \dots, L$ is done by first specifying full rank prolongation matrices P^l_{l+1} of dimension $n_l \times n_{l+1}$ and positive semidefinite prolongator smoothers $S_l = s_l(A_l)$ where s_l is a non-negative polynomial on the spectrum $\sigma(A_l)$ of A_l. For the time being, we do not specify the definition of the prolongator smoother S_l, we just note that its role will be to enforce the smoothness of the coarse-space functions. The coarse level matrices are defined by:

$$A_{l+1} = (P^l_{l+1})^T A_l S_l P^l_{l+1}, \ A_1 \equiv A. \tag{2}$$

On each level l, we assume the existence of linear smoothing operators or *preconditioners* $R_l : I\!R^{n_l} \to I\!R^{n_l}$ that transform an approximate solution $\mathbf{x}^l$ into an improved one by:

$$\mathbf{x}^l \leftarrow (I - R_l A_l)\mathbf{x}^l + R_l \mathbf{b}^l$$

and we call $K_l = I - R_l A_l$ the associated *error propagation operator*. Inversely, given S_l an error propagation operator, we will associate to it the preconditioner N_l defined by $N_l = (I - S_l)A_l^{-1}$, if A_l is invertible or $N_l = (I - S_l)A_l^{+}$ with A_l^{+} being a pseudoinverse of A_l if this matrix is not invertible.

Using this notation, a multilevel method can be written down as follows:

Algorithm 2.1 *Perform* $\mathbf{x} \leftarrow MG^{PG}(\mathbf{x}, \mathbf{b})$, *where* $MG^{PG} = MG_1^{PG}$ *and* $MG_l^{PG}(\cdot,\cdot)$, $l = 1, \dots L-1$ *is defined by:*

Pre-smoothing: *Perform* $\mathbf{x}^l \leftarrow (I - R_l A_l)\mathbf{x}^l + R_l \mathbf{b}^l$, *where* $R_l : I\!R^{n_l} \to I\!R^{n_l}$ *is the given smoother preconditioner.*

Additional pre-smoothing: *Perform one iteration of the smoother with the error propagation operator* S_l.

Coarse grid correction:

- *Set* $\mathbf{b}^{l+1} = (P^l_{l+1})^T(\mathbf{b}^l - A_l \mathbf{x}^l)$,
- *if* $l + 1 = L$, *solve* $A_{l+1}\mathbf{x}^{l+1} = \mathbf{b}^{l+1}$ *by a direct method, otherwise set* $\mathbf{x}^{l+1} = \mathbf{0}$ *and perform* γ *iterations of* $\mathbf{x}^{l+1} \leftarrow MG^{PG}_{l+1}(\mathbf{x}^{l+1}, \mathbf{b}^{l+1})$, *where* $\gamma > 0$ *is a given cycle parameter,*
- *correct the solution on level* l *by* $\mathbf{x}^l \leftarrow \mathbf{x}^l + S_l P^l_{l+1}\mathbf{x}^{l+1}$.

Post-smoothing: *Perform* $\mathbf{x}^l \leftarrow (I - R_l^T A_l)\mathbf{x}^l + R_l^T \mathbf{b}^l$.

Assume the matrix $A_1 \equiv A$ has been obtained using the conforming finite element discretization of the following scalar elliptic problem:

$$\text{find } u \in V: \ a(u,v) = f(v) \quad \text{for all } v \in V, \tag{3}$$

where $a(\cdot,\cdot)$ is a bilinear form on the function space V and $f(\cdot) \in V^{-1}$.

We show now that Algorithm 2.1 corresponds to a Petrov-Galerkin method. To this end, let us denote by Π the one-to-one mapping that associates any $\mathbf{u} \in I\!R^{n_1}$ to a finite element function defined by $u = \sum_{i=1}^{n_1} u_i \varphi_i^1$, where $\{\varphi_i\}$ is a finite element basis. For all levels $l > 1$, let us introduce the spaces $\mathcal{M}_l \subset \mathcal{M}_1 \equiv \operatorname{span}\{\varphi_i^1\}$ and $\mathcal{N}_l \subset \mathcal{M}_1$ defined by:

$$\begin{aligned} \mathcal{M}_l &= \textbf{Range}\, \Pi S_1 P_2^1 S_2 P_3^2 \ldots S_{l-1} P_l^{l-1}, \\ \mathcal{N}_l &= \textbf{Range}\, \Pi P_2^1 P_3^2 \ldots P_l^{l-1}. \end{aligned} \tag{4}$$

It is easy to see that the bases of these spaces are recursively generated by the relations:

$$\begin{bmatrix} \varphi_1^{l+1} \\ \vdots \\ \varphi_{n_{l+1}}^{l+1} \end{bmatrix} = (P_{l+1}^l)^T S_l \begin{bmatrix} \varphi_1^l \\ \vdots \\ \varphi_{n_l}^l \end{bmatrix} \tag{5}$$

and

$$\begin{bmatrix} \psi_1^{l+1} \\ \vdots \\ \psi_{n_{l+1}}^{l+1} \end{bmatrix} = (P_{l+1}^l)^T \begin{bmatrix} \psi_1^l \\ \vdots \\ \psi_{n_l}^l \end{bmatrix} \tag{6}$$

where we have put $\psi_j^1 = \varphi_j^1; j = 1, \ldots, n_1$. Further, by comparing (5) and (6) with (2), it follows that for every level,

$$A_l = \{a(\varphi_i^l, \psi_j^l)\}_{i,j=1}^{n_l}. \tag{7}$$

Hence, one can see that the correction step in Algorithm 2.1 can be alternatively written as:

$$\text{find } \ u_{l+1} \in \mathcal{M}_{l+1}: \ a(u_{l+1}, v_{l+1}) = (b_{l+1}, v_{l+1}) \quad \text{for all } v_{l+1} \in \mathcal{N}_{l+1}, \tag{8}$$

where $b_{l+1} = \Pi P_2^1 \ldots P_{l+1}^l \mathbf{b}^{l+1}$ and thus corresponds to a Petrov-Galerkin approximation with basis functions $\varphi_j^{l+1}; j = 1, \ldots, n_{l+1}$ and test functions $\psi_j^{l+1}; j = 1, \ldots, n_{l+1}$.

As the prolongator smoothers are by assumption, only positive semidefinite, the shape functions $\varphi_j^l; j = 1, \ldots, n_l$ are not guaranteed to be linearly independent. Therefore the coarse-level matrices A_l, $l > 1$, can be singular. On the other hand, by (2), the nullspaces of $S_1 P_2^1 \ldots S_{l-1} P_l^{l-1}$ and A_l coincide and therefore (8) has a unique solution

$$\Pi S_1 P_2^1 \ldots S_l P_{l+1}^l A_{l+1}^+ \mathbf{b}^{l+1}$$

independent of the choice of a pseudoinverse A_{l+1}^+.

Remark 2.1 *The additional pre-smoothing step allows the algorithm to be rewritten as a variational (Ritz-Galerkin) multigrid method, which is easier to analyze (see Sect. 3.). Since* $S_l = s_l(A_l)$ *where* $s_l(x)$ *is a polynomial verifying* $s_l(0) = 1$ *(see Sect. 4), the action of the preconditioner* $N_l = (I - S_l)A_l^+$ *can be implemented without requiring the computation of the pseudoinverse* A_l^+.

The definition of the tentative prolongator P_{l+1}^l uses a partition (disjoint covering) of the set $\{1, 2, \ldots, n_l\}$ of degrees of freedom on the level l. For the time being, assume the aggregates $\{C_i^l\}_{i=1}^{n_{l+1}}$ are simply disjoint clusters of degrees of freedom that are close in a certain sense (e.g. coupled by nonzero entry of the matrix A_l^α, α being a small integer.) More specific assumptions on the aggregates will be given in Sect. 5.

Once the partition $\{C_j^l\}_{j=1}^{n_{l+1}}$ of the set of degrees of freedom on the level l is specified, we can create the $n_l \times n_{l+1}$ *tentative* prolongation matrix by

$$(P_{l+1}^l)_{ij} = \begin{cases} 1 \ \text{ if } \ i \in C_j^l \\ \text{else} \quad 0. \end{cases} \tag{9}$$

The abstract convergence estimate in [17] depends on the range of *composite* tentative prolongators

$$P_l^1 = P_2^1 \ldots P_l^{l-1}, \ I\!R^{n_l} \to I\!R^{n_1} \tag{10}$$

and it deteriorates as the maximum of condition numbers $\operatorname{cond}((P_l^1)^T P_l^1)$ increases. To avoid this problem, we will modify the construction of the tentative prolongators in such a way that $(P_l^1)^T P_l^1 = I$ for every level $l > 1$. More general construction of tentative prolongators can be found in [17].

Since the columns of P_{l+1}^l are orthogonal, we can orthonormalize them without changing the range of the resulting composite prolongators using the following diagonal scaling procedure:

For $l = 1, \ldots, L-1$ **do**

- *create the diagonal* $n_{l+1} \times n_{l+1}$ *matrix* $D = (P_{l+1}^l)^T P_{l+1}^l$,
- *set* $P_{l+1}^l \leftarrow P_{l+1}^l D^{-1/2}$ *and, if the level* $l+2$ *exists, set* $P_{l+2}^{l+1} \leftarrow D^{1/2} P_{l+2}^{l+1}$

end do.

The scaling procedure as it is written above requires all unscaled tentative prolongators to be available at the same time. To avoid this need, (9) and the scaling can be reorganized as follows:

Algorithm 2.2 (Scaled tentative prolongator) *Let* $\mathbf{k}^1 \in I\!R^{n_1}$ *be a vector of ones or, more generally,* $\mathbf{k}^1$ *to be a discrete representation of a unit function aside from the essential boundary conditions (see Remark 2.2.)*
For $l = 1, \ldots, L-1$ **do**

1. *Create the partitioning* $\{C_j^l\}$ *of the set* $\{1,\ldots,n_l\}$ *and set* $n_{l+1} = \mathbf{card}\{C_j^l\}$.
2. *create the* $n_l \times n_{l+1}$ *matrix* P_{l+1}^l *by*

$$(P_{l+1}^l)_{ij} = \begin{cases} k_i^l & \text{if} \quad i \in C_j^l \\ \text{else} & 0, \end{cases}$$

3. *create the diagonal* $n_{l+1} \times n_{l+1}$ *matrix* $D = (P_{l+1}^l)^T P_{l+1}^l$,
4. *set* $P_{l+1}^l = P_{l+1}^l D^{-1/2}$, $\mathbf{k}^{l+1} = \mathbf{diag} D^{1/2}$.

end do

As the product of orthogonal matrices is an orthogonal matrix, it holds that

$$(P_l^1)^T P_l^1 = (P_l^{l-1})^T \ldots (P_2^1)^T P_2^1 \ldots P_l^{l-1} = (P_l^{l-1})^T P_l^{l-1} = I \;\forall l = 2,\ldots,L. \tag{11}$$

Remark 2.2 *Assume the matrix* A_1 *has arized through the discretization of a scalar elliptic problem using* $P1$ *or* $Q1$ *finite element basis functions* $\{\varphi_i\}$ *that are scaled so that* $\|\varphi_i\|_{L^\infty} = 1$. *Then the vector of ones represents a unit function in the sense that*

$$\varphi_1 + \ldots + \varphi_{n_1} = 1 \quad on \quad \Omega^{int},$$

where Ω^{int} *consists of all elements not adjacent to the part of the boundary where an essential boundary condition is imposed. In general, the Algorithm 2.2 requires the vectors* $\mathbf{k}^1$ *such that*

$$\sum_{i=1}^{n_1} k_i^1 \varphi_i = 1 \quad on \quad \Omega^{int}. \tag{12}$$

For problems with highly varying coefficients or problems discretized on strongly irregular meshes, it usually pays off to use a scaled matrix

$$A_1' = D_1^{-1/2} A_1 D_1^{-1/2}, \quad D_1 = \mathbf{diag} A_1$$

in the place of the original stiffness matrix A_1. *As the scaled matrix* A_1' *corresponds to a scaled finite element basis*

$$\varphi_i' = a_{ii}^{-1/2} \varphi_i, \quad i = 1,\ldots,n_1, \tag{13}$$

one has to replace $\mathbf{k}^1$ *satisfying (12) by a scaled vector*

$$(\mathbf{k}^1)' = (a_{11}^{-1/2} k_1^1,\; a_{22}^{-1/2} k_2^1, \ldots, a_{n_1 n_1}^{-1/2}\, k_{n_1}^1)^T$$

to enforce (12) for $\{\varphi_i'\}$ *and* $(\mathbf{k}^1)'$.

3 Abstract estimates

We use prolongator smoothers S_l that are positive semidefinite polynomials in A_l and therefore

$$A_{l+1} \equiv (P^l_{l+1})^T A_l S_l P^l_{l+1} = (S^{1/2} P_{l+1})^T A_l (S^{1/2} P_{l+1}),\ l = 1, \dots, L-1. \tag{14}$$

Hence, the procedure (2) can be viewed as a variational coarsening given by smoothed prolongator :

$$M_l P_{l+1},\ M_l = S_l^{1/2}. \tag{15}$$

First, we show that the entire Algorithm 2.1 can be perceived as a variational multigrid with prolongators $M_l P^l_{l+1}$, $l = 1, \dots, L-1$, as follows:

Algorithm 3.1 *Perform* $\mathbf{x} \leftarrow MG^{RG}(\mathbf{x}, \mathbf{b})$, *where* $MG^{RG} = MG^{RG}_1$ *and* $MG^{RG}_l(\cdot, \cdot)$, $l = 1, \dots L-1$ *is defined by:*

Pre-smoothing:

Perform $\mathbf{x}^l \leftarrow (I - R_l A_l)\mathbf{x}^l + R_l \mathbf{b}^l$ *followed by one iteration of the smoother with error propagation operator* M_l *(see Remark 3.1.)*

Coarse grid correction:

- *Set* $\mathbf{b}^{l+1} = (M_l P^l_{l+1})^T (\mathbf{b}^l - A_l \mathbf{x}^l)$,
- *if* $l + 1 = L$, *solve* $A_{l+1}\mathbf{x}^{l+1} = \mathbf{b}^{l+1}$ *by a direct method, otherwise set* $\mathbf{x}^{l+1} = \mathbf{0}$ *and perform* γ *iterations of* $\mathbf{x}^{l+1} \leftarrow MG^{RG}_{l+1}(\mathbf{x}^{l+1}, \mathbf{b}^{l+1})$, *where* $\gamma > 0$ *is a given cycle parameter,*
- *correct the solution on level* l *by* $\mathbf{x}^l \leftarrow \mathbf{x}^l + M_l P^l_{l+1} \mathbf{x}^{l+1}$.

Post-smoothing:

Perform one iteration of the smoother with error propagation operator M_l *followed by* $\mathbf{x}^l \leftarrow (I - R_l^T A_l)\mathbf{x}^l + R_l^T \mathbf{b}^l$.

Remark 3.1 *More precisely, Algorithm 3.1 uses a pre-smoother of the form*

$$\mathbf{x}^l \leftarrow (I - R_l A_l)\mathbf{x}^l + R_l \mathbf{b}^l, \quad \mathbf{x}^l \leftarrow M_l \mathbf{x}^l + (I - M_l) A_l^+ \mathbf{b}^l,$$

where the symbol $^+$ *denotes the pseudoinverse. The above smoother can be rewritten as*

$$\begin{aligned} \mathbf{x}^l &\leftarrow K_l' \mathbf{x}^l + R_l' \mathbf{b}^l, \quad \textit{where} && (16) \\ K_l' &= M_l (I - R_l A_l), \quad R_l' = (I - K_l') A_l^+. && (17) \end{aligned}$$

Lemma 3.1 *Algorithms 2.1 and 3.1 are equivalent.*

Proof: *First, consider the case where the matrices A_l are invertible. By well-known arguments, the error propagation operators E_l^{PG} and E_l^{RG} of the algorithms MG_l^{PG} and MG_l^{RG} are*

$$\begin{aligned} E_l^{PG} &= (I - R_l^T A_l)\{I - S_l P_{l+1}^l [I - (E_{l+1}^{PG})^\gamma] A_{l+1}^{-1} (P_{l+1}^l)^T A_l\} S_l (I - R_l A_l), \\ E_l^{RG} &= (I - R_l^T A_l) M_l \{I - M_l P_{l+1}^l [I - (E_{l+1}^{RG})^\gamma] A_{l+1}^{-1} (M_l P_{l+1}^l)^T A_l\} M_l (I - R_l A_l) \end{aligned}$$

for $l = 1, \ldots, L-1$ and

$$E_L^{PG} = E_L^{RG} = 0. \tag{18}$$

Assume $E_{l+1}^{PG} = E_{l+1}^{RG}$ for some positive $l < L$. Using $M_l = S_l^{1/2}$, elementary manipulations give $E_l^{PG} = E_l^{RG}$. By induction, it follows that $E_l^{PG} = E_l^{RG}$ on every level, which completes the proof.

Now let us consider the case where the matrices A_l, $l = 2, \ldots, L$ are singular. If both Algorithms 2.1 and 3.1 use the Penrose-Moore pseudoinverse on the coarsest level, it is sufficient to realize that

$$\textbf{Range}\, A_{l+1} = \textbf{Range}\, (M_l P_{l+1}^l)^T A_l = \textbf{Range}\, (P_{l+1}^l)^T A_l S_l.$$

Then the inverses of A_{l+1} in E_l^{PG}, E_l^{RG} can be understood as inverses on **Range** A_{l+1} *and the proof holds without any further change.*

If the coarsest solvers are general pseudoinversions, (18) has to be replaced by

$$\textbf{Range}\, (E_l^{PG} - E_l^{RG}) \subset \textbf{Null} A_l \tag{19}$$

for $l = L$. Since for $\mathbf{n}_{l+1} \in \textbf{Null} A_{l+1}$

$$S_l P_{l+1}^l \mathbf{n}_{l+1},\ M_l P_{l+1}^l \mathbf{n}_{l+1} \in \textbf{Null} A_l,$$

(19) holds by induction on every level. Since A_1 is regular, $E_1^{PG} = E_1^{RG}$ and the proof is completed.

In the rest of this section, we prove an abstract convergence estimate for Algorithm 3.1. Define the smoothed composite prolongator $I_l^1 : \mathbb{R}^{n_l} \to \mathbb{R}^{n_1}$ by

$$I_l^1 = M_1 P_2^1 \ldots M_{l-1} P_l^{l-1}, \quad I_1^1 = I, \tag{20}$$

and the hierarchy of coarse Hilbert spaces $V_L \subset V_{L-1} \subset \ldots \subset V_1 \equiv \mathbb{R}^{n_1}$ by

$$V_l = \left(\textbf{Range}\, I_l^1,\ \|\cdot\|_l : \mathbf{u} \mapsto \min\{\|\mathbf{x}\|_{\mathbb{R}^{n_l}} \quad \mathbf{u} = I_l^1 \mathbf{x}\} \right). \tag{21}$$

Note, that from (14) and (15) it follows that $A_l = (I_l^1)^T A I_l^1$, and

$$\begin{aligned} \|I_l^1 \mathbf{x}\|_A &= \|\mathbf{x}\|_{A_l} \quad \text{for all } \mathbf{x} \in I\!R^{n_l}, && (22)\\ \max_{\mathbf{u}\in V_l} \frac{\|\mathbf{u}\|_A}{\|\mathbf{u}\|_l} &= \max_{\mathbf{x}\in I\!R^{n_l}} \frac{\|I_l^1\mathbf{x}\|_A}{\|\mathbf{x}\|_{I\!R^{n_l}}} = \sqrt{\varrho(A_l)}. && (23)\end{aligned}$$

We define the symmetrized smoother preconditioners

$$\begin{aligned} \bar{R}_l' &= (I - K_l'^* K_l')A_l^+, && (24)\\ \bar{R}_l &= (I - K_l^* K_l)A_l^+, && (25)\\ &= (I - (I - R_l^T A_l)(I - R_l A_l))A_l^+, && (26)\end{aligned}$$

where * denotes the adjoint operator with respect to A_l-scalar product. Note that for $\bar{R}_l^{'}$ understood as a mapping from $V_l \to V_l$, i.e.

$$\bar{R}_{V_l}^{'} : I_l^1\mathbf{x},\ \mathbf{x} \in I\!R^{n_l} \ \mapsto\ I_l^1 \bar{R}_l^{'}\mathbf{x},$$

the definition of $\|\cdot\|_l$ gives

$$\min_{\mathbf{u}\in V_l} \frac{\|\bar{R}_{V_l}^{'}\mathbf{u}\|_l}{\|\mathbf{u}\|_l} = \lambda_{\min}(\bar{R}_l^{'}|_{\textbf{ Range } A_l}), \tag{27}$$

where the symbol $|$ denotes the restriction.

Our estimates use an abstract convergence result as proved in the monography [1]. Using (23), (27) it can be written in our notation as follows:

Lemma 3.2 *(Bramble [1], Theorem 3.3). Assume there are linear mappings $Q_l : V_1 \to V_l$, $Q_1 = I$ and constants $c_1,\ c_2 > 0$ such that*

$$\begin{aligned} \|Q_l\mathbf{u}\|_A &\le c_1\|\mathbf{u}\|_A\ \forall \mathbf{u}\in V_1,\ l = 1,\dots,L, && (28)\\ \|(Q_l - Q_{l+1})\mathbf{u}\|_l &\le \frac{c_2}{\sqrt{\varrho(A_l)}}\|\mathbf{u}\|_A\ \forall\mathbf{u}\in V_1,\ l = 1,\dots,L{-}1. && (29)\end{aligned}$$

Further, assume that there are constants $C_R > 0$ and $\theta \in [0, 2)$, independent of the level, such that

$$\begin{aligned} \lambda_{\min}(\bar{R}_l^{'}|_{\textbf{ Range } A_l}) &\ge \frac{1}{c_R^2\ \varrho(A_l)}, && (30)\\ (R_l^{'}A_l\mathbf{u}, R_l^{'}A_l\mathbf{u})_{A_l} &\le \theta(R_l^{'}A_l\mathbf{u},\mathbf{u})_{A_l}\ \forall\mathbf{u}\in I\!R^{n_l}. && (31)\end{aligned}$$

Then Algorithm 3.1 satisfies

$$\|A^{-1}\mathbf{b} - MG(\mathbf{x},\mathbf{b})\|_A \le \left(1 - \frac{1}{c_0}\right)\|A^{-1}\mathbf{b} - \mathbf{x}\|_A \quad \forall\mathbf{x}\in V_l,$$

where

$$c_0 = \left(1 + c_1 + c_2 c_R \sqrt{\frac{\theta}{2-\theta}}\right)^2 \frac{L-1}{2-\theta}. \tag{32}$$

The direct application of the previous lemmma to our AMG method consisting in treating the spaces V_l as analogues of finite element spaces is very problematic for the following reason: The natural basis of our coarse space $V_l \equiv$ **Range** I_l^1 is $\{I_l^1 \mathbf{e}_i^l\}_{i=1}^{n_l}$, where $\mathbf{e}_i^l$ is the $i-$th canonical basis vector of $I\!R^{n_l}$. As the matrices $M_l = S_l^{1/2}$ are dense in general, I_l^1 is dense as well. Thus, the support of the continuous analogue of the basis function $I_l^1 \mathbf{e}_i^l$ is the entire computational domain. The verification of (28) and (29) is usually performed by using the standard Finite Element approximation theory and relies heavily on local geometrical properties of the finite element bases. These properties are here difficult to establish due to the non-local nature of the coarse space basis functions.

In contrast to this, the geometrical properties of the spaces of *disaggregated* vectors $P_l^1 \mathbf{x}$, $\mathbf{x} \in I\!R^{n_l}$ are very clear due to the fact that the aggregates are disjoint on each level. The following abstract result proved in [17] allows to verify (28) and (29) using separately the conditions on the ranges of the composite tentative prolongators P_l^1 and on the prolongator smoothers S_l. For the scaled tentative prolongators P_l^1 and prolongator smoothers being polynomials in the matrices A_l, it can be written down as follows:

Assumption 3.1 *Let $S_l = s_l(A_l)$, $l = 1, \ldots, L-1$ be polynomials in the matrices A_l such that $0 \le S_l \le I$ and $(P_l^1)^T P_l^1 = I$ on every level. Further, let $\bar{\lambda}_l \ge \varrho(A_l)$ and*

$$\tilde{Q}_l : V_1 \to I\!R^{n_l}, \quad l = 2, \ldots, L$$

be given linear operators. Assume for some $C_1, C_2, C_S > 0$ and all
$l = 1, \ldots, L-1$

$$\|\mathbf{u} - P_{l+1}^1 \tilde{Q}_{l+1} \mathbf{u}\|_{I\!R^{n_1}}^2 \le \frac{C_1^2}{\bar{\lambda}_l} \|\mathbf{u}\|_A^2, \quad \forall \mathbf{u} \in V_1 \tag{33}$$

$$\|(I - S_l^{1/2})\mathbf{x}\|_{I\!R^{n_l}}^2 \le \frac{C_2^2}{\varrho(A_l)} \|\mathbf{x}\|_{A_l}^2, \quad \forall \mathbf{x} \in I\!R^{n_l} \tag{34}$$

$$\varrho(A_l S_l) \le C_S^2 \bar{\lambda}_l. \tag{35}$$

Note that the prolongator smoothers enter the assumption through the scaling by $\bar{\lambda}_l$ on the right-hand side of (33). By making $\bar{\lambda}_l$ small, the prolongator smoothers make the approximation condition (33) easier to satisfy.

Lemma 3.3 *Under the Assumption 3.1, for every $\mathbf{u} \in V_1$, the mappings*

$$Q_1 = I, \quad Q_l = I_l^1 \tilde{Q}_l, \quad l = 2, \ldots, L$$

satisfy

$$\|Q_l\mathbf{u}\|_A \le c_1\|\mathbf{u}\|_A, \quad \forall l = 1,\dots,L, \tag{36}$$

with $c_1 = 1 + C_S C_1(l-1)$, *and*

$$\|(Q_l - Q_{l+1})\mathbf{u}\|_l \le c_2\varrho(A_l)^{-1/2}\|\mathbf{u}\|_A, \quad \forall l = 1,\dots,L-1 \tag{37}$$

with $c_2 = C_1 + C_2\|Q_l\|_A \le C_1 + C_2 c_1$.

Proof: *To make this paper self consistent, we give below a version of the proof of lemma 3.3. A proof using more general assumptions than assumptions 3.1 can be found in [17]. Throughout the proof we use the symbol* $\mathbf{u}$ *to denote an arbitrary finest level vector and we omit the quantification "*$\forall\mathbf{u} \in I\!R^{n_1}$*".*

If (33) holds for some linear operators $\tilde{Q}_l$, *it holds also for* $\tilde{Q}_l$ *such that* $P_l^1\tilde{Q}_l$, $l = 2,\dots,L$ *are projections onto* **Range** P_l^1 *orthogonal with respect to* $I\!R^{n_1}$*−inner product. Hence, we can assume that* $P_l^1\tilde{Q}_l$ *are such projections without loosing the generality. Then, setting* $\tilde{Q}_1 = I$ *and* $P_1^1 = I$, *the assumption (33) gives*

$$\begin{aligned}\|(P_l^1\tilde{Q}_l - P_{l+1}^1\tilde{Q}_{l+1})\mathbf{u}\|^2_{I\!R^{n_1}} &= \|(I - P_{l+1}^1\tilde{Q}_{l+1})\mathbf{u}\|^2_{I\!R^{n_1}} - \|(I - P_l^1\tilde{Q}_l)\mathbf{u}\|^2_{I\!R^{n_1}} &(38)\\ &\le \tfrac{C_1^2}{\bar{\lambda}_l}\|\mathbf{u}\|_A^2, &(39)\end{aligned}$$

as the decomposition

$$(I - P_{l+1}^1\tilde{Q}_{l+1})\mathbf{u} = (I - P_l^1\tilde{Q}_l)\mathbf{u} + (P_l^1\tilde{Q}_l - P_{l+1}^1\tilde{Q}_{l+1})\mathbf{u}$$

is orthogonal. Since $A_{l+1} = (P_{l+1}^l)^T S_l A_l P_{l+1}^l$, $I_{l+1}^1 = I_l^1 S_l^{1/2} P_{l+1}^l$ *and* S_l *commutes with* A_l, *the assumptions* $\varrho(S_l) \le 1$ *and* $\varrho(S_l A_l) \le C_S^2\bar{\lambda}_l$ *give:*

$$\begin{aligned}\|Q_{l+1}\mathbf{u}\|_A &= \|I_{l+1}^1\tilde{Q}_{l+1}\mathbf{u}\|_A = \|S_l^{1/2}P_{l+1}^l\tilde{Q}_{l+1}\mathbf{u}\|_{A_l} &(40)\\ &= \|S_l^{1/2}(P_{l+1}^l\tilde{Q}_{l+1} - \tilde{Q}_l)\mathbf{u} + S_l^{1/2}\tilde{Q}_l\mathbf{u}\|_{A_l}\\ &\le \|S_l^{1/2}(P_{l+1}^l\tilde{Q}_{l+1} - \tilde{Q}_l)\mathbf{u}\|_{A_l} + \|S_l^{1/2}\tilde{Q}_l\mathbf{u}\|_{A_l}\\ &\le C_S\bar{\lambda}_l^{1/2}\|(\tilde{Q}_l - P_{l+1}^l\tilde{Q}_{l+1})\mathbf{u}\|_{I\!R^{n_l}} + \|\tilde{Q}_l\mathbf{u}\|_{A_l}.\end{aligned}$$

As $(P_l^1)^T P_l^1 = I$, *we have* $\|P_l^1\mathbf{u}\|_{I\!R^{n_1}} = \|\mathbf{u}\|_{I\!R^{n_l}}$ *and the previous inequality together with (38) and (22) yields*

$$\|Q_{l+1}\mathbf{u}\|_A \le C_S\bar{\lambda}_l^{1/2}\|(P_l^1\tilde{Q}_l - P_{l+1}^l\tilde{Q}_{l+1})\mathbf{u}\|_{I\!R^{n_1}} + \|\tilde{Q}_l\mathbf{u}\|_{A_l} \le C_S C_1\|\mathbf{u}\|_A + \|Q_l\mathbf{u}\|_A$$

Now (36) follows by induction using $Q_1 \equiv I$.

Let us prove (37). From the definition of Q_l, the identity $I_{l+1}^1 = I_l^1 S_l^{1/2} P_{l+1}^l$, the definition (21) and the assumption (34) we have

$$\begin{aligned} \|(Q_l - Q_{l+1})\mathbf{u}\|_l &= \|I_l^1(\tilde{Q}_l - S_l^{1/2} P_{l+1}^l \tilde{Q}_{l+1})\mathbf{u}\|_l \leq \|(\tilde{Q}_l - S_l^{1/2} P_{l+1}^l \tilde{Q}_{l+1})\mathbf{u}\|_{I\!R^{n_l}} \\ &= \|S_l^{1/2}(\tilde{Q}_l - P_{l+1}^l \tilde{Q}_{l+1})\mathbf{u} + (I - S_l^{1/2})\tilde{Q}_l \mathbf{u}\|_{I\!R^{n_l}} \\ &\leq \|S_l^{1/2}(\tilde{Q}_l - P_{l+1}^l \tilde{Q}_{l+1})\mathbf{u}\|_{I\!R^{n_l}} + \|(I - S_l^{1/2})\tilde{Q}_l \mathbf{u}\|_{I\!R^{n_l}} \\ &\leq \|S_l^{1/2}(\tilde{Q}_l - P_{l+1}^l \tilde{Q}_{l+1})\mathbf{u}\|_{I\!R^{n_l}} + C_2 \varrho(A_l)^{-1/2} \|Q_l \mathbf{u}\|_A. \end{aligned} \tag{41}$$

Using $(P_l^1)^T P_l^1 = I$ and $\varrho(S_l) \leq 1$ again, gives

$$\|S_l^{1/2}(\tilde{Q}_l - P_{l+1}^l \tilde{Q}_{l+1})\mathbf{u}\|_{I\!R^{n_l}} \leq \|(\tilde{Q}_l - P_{l+1}^l \tilde{Q}_{l+1})\mathbf{u}\|_{I\!R^{n_l}} = \|P_l^1(\tilde{Q}_l - P_{l+1}^l \tilde{Q}_{l+1})\mathbf{u}\|_{I\!R^{n_1}}.$$

The last inequality together with (41), (36) and (38) completes the proof of (37).

The following lemma translates the assumptions (30), (31) on the pre-smoothers of the Ritz-Galerkin Algorithm 3.1 into the requirements on the simpler pre-smoothers of the original Petrov-Galerkin Algorithm 2.1.

Lemma 3.4 *For every level $l = 1, \ldots, L-1$ it holds that*

$$\lambda_{\min}(\bar{R}_l'|_{\,\textbf{Range}\ A_l}) \geq \min\left\{\frac{1}{\varrho(A_l)}, \lambda_{\min}(\bar{R}_l)\right\}. \tag{42}$$

If, in addition, R_l is symmetric positive definite, commuting with A_l, the inequality in (31) holds with

$$\theta = \max\{1, \varrho(R_l A_l)\}. \tag{43}$$

Proof: *We start with proving (42). Using $S_l = M_l^2$, the commutativity of A_l and M_l, definition (24) and remark 3.1, we get*

$$\begin{aligned} \bar{R}_l' &= [I - M_l(I - R_l^T A_l)(I - R_l A_l)M_l]A_l^+ \\ &= (I - S_l)A_l^+ + M_l[I - (I - R_l^T A_l)(I - R_l A_l)]A_l^+ M_l \\ &= (I - S_l)A_l^+ + M_l \bar{R}_l M_l \;\geq\; (I - S_l)A_l^+ + \lambda_{\min}(\bar{R}_l) S_l \end{aligned}$$

where for two symmetric positive definite matrices A and B, the notation $A \geq B$ means that $A - B$ is a symmetric positive semi-definite matrix. Hence, by the spectral mapping theorem and using $S_l = s_l(A_l)$, where s_l is a polynomial,

$$\begin{aligned} \lambda_{\min}(\bar{R}_l'|_{\,\textbf{Range}\ A_l}) &\geq \min_{\mathbf{x} \in\ \textbf{Range}\ A_l} \left(\frac{\mathbf{x}^T (I - S_l) A_l^+ \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \lambda_{\min}(\bar{R}_l) \frac{\mathbf{x}^T S_l \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right) \\ &\geq \min\left\{ \frac{1 - s_l(t)}{t} + s_l(t)\lambda_{\min}(\bar{R}_l) \;\middle|\; t \in \sigma(A_l),\ t \neq 0 \right\} \\ &\geq \min\left\{ \frac{1 - s_l(t)}{\varrho(A_l)} + s_l(t)\lambda_{\min}(\bar{R}_l) \;\middle|\; t \in \sigma(A_l),\ t \neq 0 \right\}. \end{aligned}$$

As $s_l(t) \in [0,1]$ for $t \in [0, \varrho(A_l)]$ (see the forthcoming lemma 4.2), the last estimate is a convex combination of $\varrho(A_l)^{-1}$ and $\lambda_{\min}(\bar{R}_l)$, completing the proof of (42).

Let us prove (43). By (17), it follows that

$$R_l' A_l = I - M_l + M_l R_l A_l.$$

By assumption, R_l is symmetric positive definite and it commutes with A_l. Hence, taking into account that $M_l = s_l^{1/2}(A_l)$, the product $R_l' A_l$ is both symmetric and A_l−symmetric. As $s_l(t) \in [0,1]$ for $t \in \sigma(A_l)$, $R_l^{'} A_l$ is also positive semi-definite and (31) is satisfied with

$$\theta = \varrho(R_l' A_l).$$

Further,

$$R_l' A_l \leq I - M_l + \varrho(R_l A_l) M_l,$$

and by the spectral mapping theorem

$$\varrho(R_l' A_l) \leq \min_{t \in [0, \varrho(A_l)]} \{1 - s_l^{1/2}(t) + s_l^{1/2}(t) \varrho(R_l A_l)\}.$$

Now, the proof of (43) follows from the same convex combination argument as the proof of (42).

Now we are ready to prove the abstract convergence theorem for the Algorithm 2.1.

Assumption 3.2 (Smoothing property) *Assume the preconditioners $R_l,\ l = 1, \dots, L$–1 of the pre-smoothers in the Algorithm 2.1 are chosen so that*

$$\lambda_{\min}(\bar{R}_l) \geq \frac{1}{C_R^2 \varrho(A_l)} \tag{44}$$

and either

1. *each $R_l,\ l = 1, \dots, L-1$ is a symmetric positive semi-definite matrix commuting with A_l such that $\varrho(R_l A_l) \leq \theta < 2$, or*
2. *(31) is satisfied for*

$$R_l' = (I - K_l') A_l^+, \quad K_l' = (I - R_l A_l) S_l^{1/2}.$$

Theorem 3.1 Under the Assumptions 3.1 and 3.2, it holds for the Algorithm 2.1 that

$$\|A^{-1}\mathbf{b} - MG^{PG}(\mathbf{x}, \mathbf{b})\|_A \leq \left(1 - \frac{1}{C}\right) \|A^{-1}\mathbf{b} - \mathbf{x}\|_A \quad \forall \mathbf{x} \in V_l,$$

where

$$C=\left[2+C_SC_1(L-1)+\max\{1,C_R\}(C_1+C_2+C_1C_2C_S(L-1))\sqrt{\frac{\theta}{2-\theta}}\right]^2\frac{L-1}{2-\theta}=O(L^3).$$

Here, C_1, C_2, C_S, C_R and θ are the constants from the Assumptions 3.1 and 3.2.

Proof: *In the view of Lemma 3.1, it is sufficient to verify the assumptions (28), (29), and (30), (31). The assumptions (28) and (29) are verified in Lemma 3.3, the smoothing properties (30) and (31) follow from Lemma 3.4.*

4 Prolongator smoother

This section specifies suitable smoothers S_l using the Assumption 3.1 as a guideline. Our objective is to construct the smoothers that

a) minimize the available estimates $\bar{\lambda}_l \geq \varrho(A_l)$ on the right-hand side of (33) in order to make (33) easier to satisfy,

b) satisfy the constraint (34) with a reasonable constant C_2 independent of $\sigma(A_l)$.

Because of b), we restrict our considerations to the case of polynomials s_l such that $s_l(0)=1$. The following lemma clarifies the dependence of the spectral radii $\varrho(A_l)$ on the prolongator smoothers and gives the upper bounds $\bar{\lambda}_l \geq \varrho(A_l)$.

Lemma 4.1 *Let $\bar{\lambda}_1 \geq \varrho(A_1)$. Then,*

$$\bar{\lambda}_l := \bar{\lambda}_1 C_S^{2(l-1)} \geq \varrho(A_l), \quad l=1,\ldots,L,$$

where C_S is defined as in (35).

Proof: *The statement holds by definition for $l=1$. Assume it holds for some $l \geq 1$. Then by (2), (35) and (11),*

$$\varrho(A_{l+1}) = \max_{\mathbf{x}\in I\!R^{n_{l+1}}} \frac{(A_l S_l P^l_{l+1}\mathbf{x}, P^l_{l+1}\mathbf{x})_{I\!R^{n_l}}}{\|\mathbf{x}\|^2_{I\!R^{n_{l+1}}}} \leq C_S^2\bar{\lambda}_l \frac{\|P^l_{l+1}\mathbf{x}\|^2_{I\!R^{n_l}}}{\|\mathbf{x}\|^2_{I\!R^{n_{l+1}}}} = C_S^2\bar{\lambda}_l.$$

As $C_S^2\bar{\lambda}_l = \bar{\lambda}_{l+1}$, the proof follows by induction.

With Lemma 4.1 in mind, we choose $S_l = s(\bar{\lambda}_l^{-1}A_l)$, where s is a polynomial of a given degree minimizing

$$\max_{x\in[0,1]} p(x)x \text{ subject to } p(0)=1. \tag{45}$$

More precisely, in section 5, we will show that the tentative prolongators verify :

$$\|\mathbf{u} - P_l^1 \tilde{Q}_l \mathbf{u}\|^2_{I\!R^{n_1}} \leq \frac{C_A^2 (d+1)^{2l}}{\bar{\lambda}_1} \|\mathbf{u}\|^2_A,$$

and thus the possibility to verify assumption 33 will rely on the fact that we can construct a prolongator smoother such that $\bar{\lambda}_l \sim \bar{\lambda}_1/(d+1)^{2(l-1)}$. The following lemma will show that it is indeed possible. For polynomial smoothers of even degree, the results (49) and (50) have already been established in [16].

Lemma 4.2 *Let $\mathcal{P}_n$ be a set of polynomials of degree n such that $p(0) = 1$ for all $p(x) \in \mathcal{P}_n$. Then for any integer $n > 0$, there is a unique polynomial $s(x) \in \mathcal{P}_n$, such that*

$$\min_{p \in \mathcal{P}_n} \left(\max_{x \in [0,1]} p(x)x \right) = \max_{x \in [0,1]} s(x)x. \tag{46}$$

The polynomial $s(x)$ is given by

$$s(x) = \begin{cases} \prod_{k=1}^{r} \left(1 - \dfrac{x}{r_k}\right)^2 \cdot (1-x) & \textit{for odd } n = 2r+1 \\ \prod_{k=1}^{r} \left(1 - \dfrac{x}{r_k}\right)^2 & \textit{for even } n = 2r \end{cases} \tag{47}$$

where the roots r_k of $s(x)$ are in both cases given by

$$r_k = \frac{1}{2}\left(1 - \cos\frac{2k\pi}{n+1}\right) = \sin^2\frac{k\pi}{n+1}, \qquad k = 1, \ldots, r \tag{48}$$

In addition, the polynomial $s(x)$ satisfies

$$\max_{x \in [0,1]} s(x)x \quad = \quad \frac{1}{(n+1)^2} \tag{49}$$

$$0 \leq s(x) \quad \leq \quad 1, \qquad x \in [0,1] \tag{50}$$

$$[1 - s^{1/2}(x)]^2 \quad \leq \quad \left(\frac{1}{2} + \frac{\pi^2}{12}\right)(n+1)^2 x, \qquad \forall x \in [0,1] \tag{51}$$

Proof: *Let us consider a polynomial w_{n+1} of degree $(n+1)$ such that*

$$\max_{x \in [-1,1]} w_{n+1}(x) = -\min_{x \in [-1,1]} w_{n+1}(x) = w_{n+1}(1),$$

Then, denoting $\|w_{n+1}\|_\infty = \max_{x \in [-1,1]} |w_{n+1}(x)|$, we can write the following:

$$s(x)x = q_{n+1}(x) = \frac{c}{2} \frac{\left(\|w_{n+1}\|_\infty - w_{n+1}(1-2x)\right)}{\|w_{n+1}\|_\infty}\ , \tag{52}$$

where $c = \max_{x\in[0,1]} q_{n+1}(x)$. *Because of the constraint* $s(0) = 1$, *we also have*

$$1 = q'_{n+1}(0) = c\frac{w'_{n+1}(1)}{\|w_{n+1}\|_\infty}$$

And hence

$$c = \frac{\|w_{n+1}\|_\infty}{w'_{n+1}(1)}.$$

If we want to minimize c *to satisfy (46), we have to choose* $w_{n+1}(x)$ *to be an arbitrary multiple of a Chebyshev polynomial of degree* $(n+1)$. *Let us choose such a multiple that* $\|w_{n+1}\|_\infty = 1$, *then*

$$s(x)x = q_{n+1}(x) = \frac{c}{2}\big(1 - w_{n+1}(1-2x)\big), \tag{53}$$

where

$$c = \max_{x\in[0,1]} q_{n+1}(x) = \frac{1}{w'_{n+1}(1)} = \frac{1}{(n+1)^2}.$$

This proves (46) and (49).

The polynomial $q_{n+1}(x)$ *vanishes at the points* x *where* $w_{n+1}(1-2x) = 1$, *that is where* $1 - 2x = \cos(2k\pi/n+1)$. *In the case* $n = 2r$, *the value* $k = 0$ *gives the simple root of* q_{n+1} $r_0 = 0$, *while* $k = 1, \ldots, r$ *yield double roots given by (48), whereas in the case* $n = 2r+1$, *there are two simple roots of* q_{n+1}, $r_0 = 0$ *and* $r_{r+1} = 1$, *for* $k = 0$ *and* $k = r+1$, *respectively. The values of* $k = 1, \ldots, r$ *give again double roots as in (48).*

For proving (50) we use w_{n+1} *in (52); we will show that*

$$0 \le \frac{1 - w_{n+1}(1-2x)}{2xw'_{n+1}(1)} \le 1, \qquad \forall x \in [0,1].$$

Since $w_{n+1}(x) \in [-1,1]$ *for* $x \in [-1,1]$, *the lower bound is obvious. After a substitution* $1 - 2x = \tilde{x}$ *the upper bound becomes*

$$w_{n+1}(\tilde{x}) \ge 1 + w'_{n+1}(1)(\tilde{x} - 1), \quad \forall \tilde{x} \in [-1,1],$$

which is the well known fact that a graph of the Chebyshev polynomial lies above its tangent at $\tilde{x} = 1$.

Now, we will prove (51). First, define

$$\tilde{s}(x) = \begin{cases} s^{1/2}(x)(1-x)^{1/2} & \text{for odd } n = 2r+1 \\ s^{1/2}(x) & \text{for even } n = 2r \end{cases}.$$

In the proof of (51) we will need the fact that for any integer $n > 0$ *the graph of* $\tilde{s}(x)$ *for* $x \in [0,1]$ *lies above its tangent at* $x = 0$, $\tilde{s}(0) = 1$, *ie:*

$$1 + \tilde{s}'(0)x \le \tilde{s}(x), \qquad x \in [0,1]. \tag{54}$$

Note, that $\tilde{s}(x)$ behaves polynomially for $x \in [0, r_1]$, where r_1 is the smallest root of $s(x)$. Also, $\tilde{s}(x) \geq 0$ for $x \in [0,1]$ and

$$\tilde{s}(x) = \left| \prod_{k=1}^{\tilde{r}} \left(1 - \frac{x}{r_k}\right) \right|,$$

where $\tilde{r} = r + 1$ for odd n, $\tilde{r} = r$ for even n, and the additional root $r_{r+1} = 1$.

The argument in (54) then follows from the fact that $\tilde{s}''(x) > 0$ for $x \in [0, r_1]$, ie. $\tilde{s}(x)$ is in this interval convex.

By differentiating $\tilde{s}(x)$ we get $\tilde{s}'(0) = -\sum_{k=1}^{\tilde{r}} r_k^{-1}$. Employing the estimate $\sin x \geq \frac{2}{\pi}x$, $x \in [0, \frac{\pi}{2}]$, the standard sum $\sum_{k=1}^{\infty} k^{-2} = \frac{\pi^2}{6}$, and using (54) and (50) we can prove (51):

$$[1 - s^{1/2}(x)]^2 \leq 2[1 - s^{1/2}(x)] \leq 2[1 - \tilde{s}(x)] \leq -2\tilde{s}'(0)x$$

continued in case when $n = 2r + 1$ by

$$\begin{aligned} -2\tilde{s}'(0)x &= \left(2 + \sum_{k=1}^{n} \frac{2}{\sin^2 \frac{k\pi}{n+1}}\right) x \leq \left(\frac{1}{2} + \frac{1}{2}\sum_{k=1}^{n} \frac{1}{k^2}\right)(n+1)^2 x \\ &\leq \left(\frac{1}{2} + \frac{1}{2}\sum_{k=1}^{\infty} \frac{1}{k^2}\right)(n+1)^2 x \leq \left(\frac{1}{2} + \frac{\pi^2}{12}\right)(n+1)^2 x, \end{aligned}$$

while for $n = 2r$ we continue in the same way by

$$-2s'(0)x = \left(\sum_{k=1}^{n} \frac{2}{\sin^2 \frac{k\pi}{n+1}}\right) x \leq \frac{1}{2}\left(\sum_{k=1}^{n} \frac{1}{k^2}\right)(n+1)^2 x \leq \frac{\pi^2}{12}(n+1)^2 x,$$

which concludes the proof.

The construction of the prolongator smoother S_l is summed up in the following algorithm.

Algorithm 4.1 (Prolongator smoothers) *For a given level l, the matrix A_l, the estimate $\bar{\lambda}_1 \geq \varrho(A_1)$ and the degree of the prolongator smoother d define the prolongator smoother S_l as follows:*

1. *Get the estimate $\bar{\varrho}(A_l) \geq \varrho(A_l)$ and set*

$$\bar{\lambda}_l = \min\left\{\bar{\varrho}(A_l), \frac{\bar{\lambda}_1}{(d+1)^{2(l-1)}}\right\},$$ [1]

2. *Set $S_l = s(\bar{\lambda}_l^{-1} A_l)$, where s is the polynomial of the degree d given by (47).*

[1] $\bar{\lambda}_1/(d+1)^{2(l-1)}$ is the estimate provided by Lemma 4.1, 4.2.

The proof of the following statement follows directly from Lemma 4.1 and 4.2.

Assumption 4.1 *Let the tentative prolongators P^l_{l+1} be constructed by the Algorithm 2.2 and the prolongator smoothers S_l are polynomials of the degree d defined by Algorithm 4.1. We assume that there exist linear mappings*

$$\tilde{Q}_l : I\!R^{n_1} \to I\!R^{n_l},\ l = 2, \ldots, L$$

such that

$$\|\mathbf{u} - P^1_l \tilde{Q}_l \mathbf{u}\|_{I\!R^{n_1}} \leq \frac{C_A(d+1)^{l-1}}{\sqrt{\bar{\lambda}_1}} \|\mathbf{u}\|_A \ \forall \mathbf{u} \in I\!R^{n_1},\ l = 2, \ldots, L \tag{55}$$

holds with a constant $C_A > 1$ independent of l.

Lemma 4.3 *Under the Assumption 4.1, the Assumption 3.1 holds with constants*

$$C_1 = C_A, \quad C_2 = \sqrt{\frac{1}{2} + \frac{\pi^2}{12}}\,(d+1) \quad \text{and} \quad C_S = \frac{1}{d+1}.$$

5 Model example

Let $\Omega \subset I\!R^D$, $D = 2, 3$ be a bounded domain, τ_h a quasiuniform finite element mesh on Ω, and V_h a P1 or Q1 finite element space associated with τ_h. For convenience, we assume that the zero Dirichlet boundary condition has been imposed in all boundary nodes for functions in V_h and the scaling $\|\phi_i\|_{L^\infty} = 1,\ i = 1, \ldots, n_1$.

Our goal is to solve a second order scalar elliptic problem

$$\text{find } u \in V_h \text{ such that } a(u, v) = (f, v)_{L^2(\Omega)} \text{ for every } v \in V_h,$$

where $a(\cdot, \cdot)$ is a bilinear form on $H^1(\Omega)$ satisfying

$$c|u|^2_{H^1(\Omega)} \leq a(u, u) \leq C|u|^2_{H^1(\Omega)} \tag{56}$$

for every $u \in H^1(\Omega)$.

Assume the tentative prolongators have been constructed by Algorithm 2.2, the prolongator smoothers are polynomials of the degree d defined by Algorithm 4.1 and the pre-smoother preconditioners R_l satisfy Assumption 3.2 (e.g. $R_l = \varrho(A_l)^{-1} I$.)

For each aggregate $\mathcal{C}^l_i$, define the *composite aggregate* $\hat{\mathcal{C}}^l_i$ to be the aggregate $\mathcal{C}^l_i$ understood as the set of degrees of freedom on the level 1, i.e.

$$\hat{\mathcal{C}}^l_i \equiv \{j \ : \ (P^1_l)_{ji} \neq 0\}, \quad l = 1, \ldots, L-1,\ i = 1, \ldots, n_l.$$

We assume that the composite aggregates satisfy the following geometrical assumptions: For each aggregate C^l_i there is a ball $U^l_i \subset I\!R^d$ such that:

1. all degrees of freedom contained in composite aggregate $\tilde{C}_i^l$ are located within U_i^l,
2. $\mathbf{diam}\ U_i^l \leq C(d+1)^{l-1}h$, where h is the characteristic meshsize of τ_h, d is the degree of the prolongator smoother and C is a positive constant independent of the level,
3. there is an integer constant N independent of the level such that every point $\mathbf{x} \in \Omega$ belongs to at most N balls U_i^l. (Overlaps of balls $\{U_i^l\}_{i=1}^{n_l}$ are bounded.)

The analysis consists in verification of the Assumption 4.1 and uses the Lemma 4.3 and Theorem 3.1 to carry out the convergence estimate.

The weak approximation property (55) depends only on the range of P_l^1 and the scaling performed inside the main loop of the Algorithm 2.2 does not change **Range** P_l^1. Therefore, we can carry out the estimates for the nonscaled tentative prolongators as defined in (9) without loosing generality.

The key tool for our verification is the following scaled Poincaré-Friedrichs inequality: For a ball $U \subset I\!R^D$, $U' \subset U$ and every function $u \in H^1(U)$ it holds that

$$\|u - p\|_{L^2(U')} \leq C\ \mathbf{diam}\ U|u|_{H^1(U)}, \quad p = \int_U u\, d\mathbf{x}. \tag{57}$$

With (57) in mind, we define $\tilde{Q}_l,\ l = 1, \dots, L$ as follows: Let $\Pi : I\!R^{n_1} \mapsto H^1(I\!R^D)$ be the mapping that for $\mathbf{x} \in I\!R^{n_1}$ returns $\sum_{i=1}^{n_1} x_i\varphi_i$ extended by zero outside Ω. We set

$$\tilde{Q}_l\mathbf{u} = \mathbf{w}^l, \quad \text{where} \quad w_i^l = \int_{U_i^l} \Pi\mathbf{u}\, d\mathbf{x}.$$

As P_l^1 is a nonscaled composite prolongator, $(P_l^1\mathbf{w}^l)_j = w_i^l$ for $j \in \hat{C}_i^l$ and,

$$\|\mathbf{u} - P_l^1\tilde{Q}_l\mathbf{u}\|^2 = \sum_{i=1}^{n_l} \|\mathbf{u} - P_l^1\mathbf{w}^l\|^2_{l^2(\hat{C}_i^l)} = \sum_{i=1}^{n_l} \|\mathbf{u} - w_i^l\|^2_{l^2(\hat{C}_i^l)}, \tag{58}$$

where $\|\cdot\|_{l^2(\hat{C}_i^l)} : \mathbf{x} \in I\!R^{n_1} \mapsto \sqrt{\sum_{k\in\hat{C}_i^l} x_k^2}$. Using the quasiuniformity of τ_h, the well-known equivalence of the discrete and continuous L^2-norms, (57) and the fact that $\Pi\mathbf{u}$ vanishes outside Ω we get

$$\begin{aligned}
\|\mathbf{u} - w_i^l\|^2_{l^2(\hat{C}_i^l)} &= Ch^{-D}\|\Pi\mathbf{u} - w_i^l\|^2_{L^2(U_i^l\cap\Omega)} \leq Ch^{-D}(\ \mathbf{diam}\ U_i^l)^2|\Pi\mathbf{u}|^2_{H^1(U_i^l)} \\
&= Ch^{-D}(\ \mathbf{diam}\ U_i^l)^2|\Pi\mathbf{u}|^2_{H^1(U_i^l\cap\Omega)}.
\end{aligned}$$

Substituting the last estimate into (58), using the bounded overlaps of U_i^l, $\mathbf{diam}\ U_i^l \leq C(d+1)^{l-1}h$, the uniform equivalence (56) and the well-known estimate $\varrho(A_1) \leq Ch^{D-2}$ we get (55) with $\bar{\lambda}_1 = Ch^{D-2}$.

Now, Lemma 4.3 together with Theorem 3.1 gives the estimate

$$\|A^{-1}\mathbf{b} - MG^{PG}(\mathbf{x}, \mathbf{b})\|_A \leq \left(1 - \frac{1}{CL^3}\right)\|A^{-1}\mathbf{b} - \mathbf{x}\|_A, \quad \text{for all} \quad \mathbf{x}, \mathbf{b} \in I\!R^{n_1}$$

with a constant C independent of L, h, and Ω.

Acknowledgments

The work of Petr Vaněk has been performed while he was enjoying the hospitality of INRIA Sophia-Antipolis. Part of his work has also been supported through the following grants: NSF ASC 97-20257, Sandia Lab LG-4440 and NASA Ames NAG2-1238.

References

[1] J. H. Bramble, **Multigrid methods**, Pitman research notes in mathematics series, Longman Scientific and Technical, 1993

[2] J. H. Bramble, J. E Pasciak, J. Wang and J. Xu, **Convergence estimates for multigrid algorithms without regularity assumptions**, Math. Comp., **57**, pp 23-45, 1991.

[3] D. Mavriplis, A. Jameson, L. Martinelli, **Multi-grid solution of the Navier-Stokes equations on triangular meshes**, ICASE Rpt No 89-11, 1989.

[4] H. Guillard, P. Vaněk, **An aggregation Multigrid solver for convection-diffusion problems on unstructured meshes**, UCD/CCM Report No 130, University of Colorado in Denver, 1998.

[5] M.H. Lallemand, A. Dervieux, **A multigrid finite element method for solving the two-dimensional Euler equations**, Proceedings of the third Copper Mountain conference on Multigrid methods, Lecture Notes in Pure and Applied mathematics, Ed S. F. Mc Cormick, Marcel Dekker Inc, April 87, pp 337-363.

[6] M.H. Lallemand, H. Steve, A. Dervieux, **Computers & Fluids**, **21**, (1992), pp397-433.

[7] J. Ruge and K. Stuben, **Algebraic Multigrid Methods**: in Multigrid Methods (S. Mc Cormick Ed), Frontiers in Applied Math, Vol 5, SIAM, Philadelphia.

[8] B. Koobus, M.H. Lallemand, A. Dervieux, **Unstructured Volume Agglomeration MG, Solution of the Poisson equation**, Int. J. for Numerical Meth. in Fluid, **18**, pp 27-42, 1994.

[9] P. Wesseling, **An introduction to Multigrid methods**, Wiley, 1992.

[10] G. Carré, A. Dervieux, **Unstructured multigrid for $k-\epsilon$ kernel, Workshop on efficient models for aeronautics**, Manchester, November 11-14, 1994.

[11] V. Venkatakrishnan, D.Mavriplis, **Unstructured multigrid for the 3-D Euler equations**, AIAA paper 94-0069, January 1994.

[12] P. Vaněk, **Acceleration of convergence of a two-level algorithm by smoothing trarnsfer operator** , Applications of Mathematics, 37 (1992), pp. 265–274.

[13] P. Vaněk, **Fast multigrid solver** . Applications of Mathematics, 1994

[14] P. Vaněk, J. Mandel, M. Brezina, **Algebraic Multigrid based on smoothed aggregation for second and fourth order problems**, Computing, **57**, pp 179-196, 1996.

[15] P. Vaněk, R. Tezaur, M. Brezina and J. Křižkova, **Two Level method with coarse space size independent convergence**, in Domain decomposition methods in sciences and engineering, 8^{th} International conference, Beijing, P.R. China, Wiley, pp 233-240, 1997.

[16] P. Vaněk, J. Mandel, M. Brezina, **Two Level Algebraic multigrid for the Helmholtz problem**, Comtemporary Mathematics, 218, pp 349-356. Proceedings of the 10^{th} International conference on domain decomposition, Boulder C0, 1997, J. Mandel, C. Farhat, X.C Cai Eds, AMS, Providence, 1998.

[17] P. Vaněk, M. Brezina, J. Mandel, **Convergence of Algebraic Multigrid Based on Smoothed Aggregation**, submitted to Num. Math.

Contents

1 Introduction **3**

2 Petrov-Galerkin MG method **4**

3 Abstract estimates **8**

4 Prolongator smoother **15**

5 Model example **19**

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - B.P. 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 Villers lès Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot St Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, B.P. 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399