# Analysis of an Algebraic Petrov-Galerkin Smoothed Aggregation Multigrid Method


Hervé Guillard [a], Aleš Janka [b] and Petr Vaněk [c]

[a] *INRIA, BP. 93, F-06902 Sophia Antipolis Cedex, France*

[b] *Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland*

[c] *Formerly, Dpt of Math, University of California, Los Angeles, CA-90024-1555*


---


**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, Algebraic multigrid, Smoothed Aggregation, Agglomeration multigrid
*2000 MSC:* 65B99, 65N22, 65N55, 65F10, 65F50, 65Y20


---

## 1 Introduction

Finite volume multigrid methods for inviscid fluid dynamics uses very simple constant prolongation and restriction operators [3]. However for the computation of $2^{nd}$ order elliptic equations or viscous flows, more sophisticated (at least linear) interpolation operators have to be used. A rule of thumb investigated by Hemker in [2] is that the sum of the orders of the prolongation and of the restriction should at least be equal to the order of the differential equation solved. However, for agglomeration methods on non-structured finite element


*Email addresses:* herve.guillard@sophia.inria.fr (Hervé Guillard), ales.janka@epfl.ch (Aleš Janka), vanek@math.ucla.edu (Petr Vaněk).



*Preprint submitted to Elsevier Science* *23 August 2006*

meshes, the construction of linear interpolation operators is a difficult task due to the lack of geometrical regularity of the aggregates. One way to deal with this problem is to use instead of linear prolongators the concept of smoothed aggregation [7,8]. 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 range of the resulting prolongators 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 [10]. 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. In [6], 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. 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 $\mathbb{R}^{n_1}$. We call $\mathbb{R}^{n_1}$ the finest level and pose $A_1 = A$ .The generation of the coarse levels $l = 2, \ldots, L$ is done by first specifying full rank prolongation matrices

$P^\ell_{\ell+1}$ of dimension $n_\ell \times n_{\ell+1}$ and positive semidefinite prolongator smoothers $S_\ell = s_\ell(A_\ell)$ where $s_\ell$ is a non-negative polynomial on the spectrum $\sigma(A_\ell)$ of $A_\ell$. For the time being, we do not specify the definition of the prolongator smoother $S_\ell$, we just note that its role will be to enforce the smoothness of the coarse-space functions. The coarse level matrices are defined by:

$$A_{\ell+1} = (P^\ell_{\ell+1})^T A_\ell S_\ell P^\ell_{\ell+1},\ A_1 \equiv A. \tag{2}$$

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

$$\mathbf{x}^\ell \leftarrow (I - R_\ell A_\ell)\mathbf{x}^\ell + R_\ell \mathbf{b}^\ell$$

and we call $K_\ell = I - R_\ell A_\ell$ the associated *error propagation operator*. Inversely, given $S_\ell$ an error propagation operator, we will associate to it the preconditioner $N_\ell$ defined by $N_\ell = (I - S_\ell)A_\ell^{-1}$, if $A_\ell$ is invertible or $N_\ell = (I - S_\ell)A_\ell^+$ with $A_\ell^+$ being a pseudoinverse of $A_\ell$ if this matrix is not invertible.

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

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

(1) *(Pre-smoothing) Perform* $\mathbf{x}^\ell \leftarrow (I - R_\ell A_\ell)\mathbf{x}^\ell + R_\ell \mathbf{b}^\ell$, *where* $R_\ell : \mathbb{R}^{n_\ell} \to \mathbb{R}^{n_\ell}$ *is the given* smoother preconditioner.

(2) *(Additional pre-smoothing) Perform one iteration of the smoother with the error propagation operator* $S_\ell$.

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

(4) *(Post-smoothing) Perform* $\mathbf{x}^\ell \leftarrow (I - R_\ell^T A_\ell)\mathbf{x}^\ell + R_\ell^T \mathbf{b}^\ell$.

Let us now comment on the practical implementation of Algorithm 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 $S_\ell = s_\ell(A_\ell)$ where $s_\ell(x)$ is a polynomial verifying $s_\ell(0) = 1$ (see Sect. 4), the action of the preconditioner $N_\ell = (I - S_\ell)A_\ell^+$ in step (2) of algorithm 1 can be implemented without requiring the computation of the pseudoinverse

$A_\ell^+$. Actually, the concatenation of step (2) and the projection of the residual on the coarse space can be written : $\mathbf{b}^{\ell+1} = (P_{\ell+1}^\ell)^T S_\ell(\mathbf{b}^\ell - A_\ell \mathbf{x}^\ell)$ and thus the computation of $A_\ell^{-1}$ or $A_\ell^+$ is of course never required.

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 1 corresponds to a Petrov-Galerkin method. To this end, let us denote by $\Pi$ the one-to-one mapping that associates any $\mathbf{u} \in \mathbb{R}^{n_1}$ 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 $\mathcal{M}_\ell \subset \mathcal{M}_1 \equiv \text{span}\{\varphi_i^1\}$ and $\mathcal{N}_\ell \subset \mathcal{M}_1$ defined by:

$$\mathcal{M}_\ell = \text{Rng}\,\Pi S_1 P_2^1 S_2 P_3^2 \dots S_{l-1} P_\ell^{l-1}, \mathcal{N}_\ell = \text{Rng}\,\Pi P_2^1 P_3^2 \dots P_\ell^{l-1}. \tag{4}$$

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

$$\begin{bmatrix} \varphi_1^{\ell+1} \\ \vdots \\ \varphi_{n_{\ell+1}}^{\ell+1} \end{bmatrix} = (P_{\ell+1}^\ell)^T S_\ell \begin{bmatrix} \varphi_1^l \\ \vdots \\ \varphi_{n_l}^l \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \psi_1^{\ell+1} \\ \vdots \\ \psi_{n_{\ell+1}}^{\ell+1} \end{bmatrix} = (P_{\ell+1}^\ell)^T \begin{bmatrix} \psi_1^l \\ \vdots \\ \psi_{n_l}^l \end{bmatrix}, \tag{5}$$

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

$$A_\ell = \{a(\varphi_i^\ell, \psi_j^\ell)\}_{i,j=1}^{n_\ell}. \tag{6}$$

Hence, one can see that the correction step in Algorithm 1 can be alternatively written as:

$$\text{find } \; u_{\ell+1} \in \mathcal{M}_{\ell+1} : \; a(u_{\ell+1}, v_{\ell+1}) = (b_{\ell+1}, v_{\ell+1}) \quad \text{for all } v_{\ell+1} \in \mathcal{N}_{\ell+1}, \tag{7}$$

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

As the prolongator smoothers are by assumption, only positive semidefinite, the shape functions $\varphi^l_j; j = 1, \ldots, n_l$ are not guaranteed to be linearly independent. Therefore the coarse-level matrices $A_\ell$, $l > 1$, can be singular. On the other hand, by (2), the nullspaces of $S_1 P^1_2 \ldots S_{l-1} P^{l-1}_l$ and $A_l$ coincide and therefore (7) has a unique solution $\Pi S_1 P^1_2 \ldots S_\ell P^\ell_{\ell+1} A^+_{\ell+1} \mathbf{b}^{\ell+1}$ independent of the choice of a pseudoinverse $A^+_{\ell+1}$.

The definition of the tentative prolongator $P^\ell_{\ell+1}$ uses a partition (disjoint covering) of the set $\{1, 2, \ldots, n_\ell\}$ of degrees of freedom on the level $l$. For the time being, assume the aggregates $\{\mathcal{A}^\ell_i\}^{n_{\ell+1}}_{i=1}$ are simply disjoint clusters of degrees of freedom that are close in a certain sense. Once the partition $\{\mathcal{A}^\ell_j\}^{n_{\ell+1}}_{j=1}$ of the set of degrees of freedom on the level $l$ is specified, we can create the $n_\ell \times n_{\ell+1}$ *tentative* prolongation matrix by

$$(P^\ell_{\ell+1})_{ij} = \begin{cases} 1 & \text{if} \quad i \in \mathcal{A}^\ell_j \\ \text{else} & 0. \end{cases} \tag{8}$$

The abstract convergence estimate in [10] depends on the range of *composite* tentative prolongators

$$P^1_l = P^1_2 \ldots P^{l-1}_l, \ \mathbb{R}^{n_\ell} \rightarrow \mathbb{R}^{n_1} \tag{9}$$

and it deteriorates as the maximum of condition numbers $\text{cond}((P^1_\ell)^T P^1_\ell)$ increases. To avoid this problem, we will modify the construction of the tentative prolongators in such a way that $(P^1_\ell)^T P^1_\ell = I$ for every level $l > 1$.

Since the columns of $P^\ell_{\ell+1}$ 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 $n_{\ell+1} \times n_{\ell+1}$ matrix $D = (P^\ell_{\ell+1})^T P^l_{\ell+1}$,
- set $P^\ell_{\ell+1} \leftarrow P^\ell_{\ell+1} D^{-1/2}$ and, if the level $l+2$ exists, set $P^{\ell+1}_{l+2} \leftarrow D^{1/2} P^{\ell+1}_{l+2}$

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, (8) and the scaling can be reorganized as follows:

**Algorithm 2 (Scaled tentative prolongator)** *Let $\mathbf{k}^1 \in \mathbb{R}^{n_1}$ 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.*
*For $l = 1, \ldots, L-1$ do*

*(1) Create the partitioning* $\{\mathcal{A}_j^\ell\}$ *of the set* $\{1,\ldots,n_\ell\}$ *and set* $n_{\ell+1}=\mathrm{card}\{\mathcal{A}_j^\ell\}$.

*(2) Create the* $n_l \times n_{\ell+1}$ *matrix* $P_{\ell+1}^l$ *by*

$$(P_{\ell+1}^\ell)_{ij} = \begin{cases} k_i^\ell \ \textit{if}\ i\in\mathcal{A}_j^\ell \\ 0 \quad \textit{otherwise}, \end{cases}$$

*(3) create the diagonal* $n_{\ell+1}\times n_{\ell+1}$ *matrix* $D=(P_{\ell+1}^\ell)^T P_{\ell+1}^\ell$,

*(4) set* $P_{\ell+1}^\ell = P_{\ell+1}^\ell D^{-1/2}$, $\mathbf{k}^{\ell+1}=\mathrm{diag}(D^{1/2})$.

*end do.*

As the product of orthogonal matrices is an orthogonal matrix, it holds for all $l=2,\ldots,L$ that

$$(P_\ell^1)^T P_\ell^1 = (P_\ell^{l-1})^T \ldots (P_2^1)^T P_2^1 \ldots P_\ell^{l-1} = (P_\ell^{l-1})^T P_\ell^{l-1} = I. \tag{10}$$

## 3 Abstract estimates

We use prolongator smoothers $S_\ell$ that are positive semidefinite polynomials in $A_\ell$ and therefore

$$A_{\ell+1} \equiv (P_{\ell+1}^\ell)^T A_\ell S_\ell P_{\ell+1}^\ell = (S_\ell^{1/2} P_{\ell+1})^T A_\ell (S_\ell^{1/2} P_{\ell+1}),\ l=1,\ldots,L-1. \tag{11}$$

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

$$M_\ell P_{\ell+1},\ M_\ell = S_\ell^{1/2}. \tag{12}$$

First, we show that the entire Algorithm 1 can be perceived as a variational multigrid with prolongators $M_\ell P_{\ell+1}^\ell$, $l=1,\ldots,L-1$, as follows:

**Algorithm 3 (Ritz-Galerkin)** *Perform* $\mathbf{x}\leftarrow MG^{RG}(\mathbf{x},\mathbf{b})$, *where* $MG^{RG}=MG_1^{RG}$ *and* $MG_\ell^{RG}(\cdot,\cdot)$, $l=1,\ldots L-1$ *is defined by:*

*(1) (Pre-smoothing) Perform* $\mathbf{x}^\ell \leftarrow (I-R_\ell A_\ell)\mathbf{x}^\ell + R_\ell \mathbf{b}^\ell$ *followed by one iteration of the smoother with error propagation operator* $M_\ell$ *(see Remark 1).*

*(2) (Coarse grid correction)*

- *Set* $\mathbf{b}^{\ell+1} = (M_\ell P_{\ell+1}^\ell)^T(\mathbf{b}^\ell - A_\ell \mathbf{x}^\ell)$,

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

*(3) (Post-smoothing) Perform one iteration of the smoother with error propagation operator* $M_\ell$ *followed by* $\mathbf{x}^{\ell} \leftarrow (I - R^T_\ell A_\ell)\mathbf{x}^{\ell} + R^T_\ell \mathbf{b}^{\ell}$.

**Remark 1** *More precisely, Algorithm 3 uses a pre-smoother of the form*

$$\mathbf{x}^{\ell} \leftarrow (I - R_\ell A_\ell)\mathbf{x}^{\ell} + R_\ell \mathbf{b}^{\ell}, \quad \mathbf{x}^{\ell} \leftarrow M_\ell \mathbf{x}^{\ell} + (I - M_\ell)A^{+}_\ell \mathbf{b}^{\ell},$$

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

$$\mathbf{x}^{\ell} \leftarrow K'_\ell \mathbf{x}^{\ell} + R'_\ell \mathbf{b}^{\ell}, where K'_\ell = M_\ell(I - R_\ell A_\ell), R'_\ell = (I - K'_\ell)A^{+}_\ell. \tag{13}$$

**Lemma 2** *Algorithms 1 and 3 are equivalent.*

**PROOF.** First, consider the case where the matrices $A_\ell$ are invertible. By well-known arguments, the error propagation operators $E^{PG}_\ell$ and $E^{RG}_\ell$ of the algorithms $MG^{PG}_\ell$ and $MG^{RG}_\ell$ are

$$\begin{aligned} E^{PG}_\ell &= (I - R^T_\ell A_\ell)\{I - S_\ell P^{\ell}_{\ell+1}[I - (E^{PG}_{\ell+1})^{\gamma}]A^{-1}_{\ell+1}(P^{l}_{\ell+1})^T A_\ell\} S_\ell (I - R_\ell A_\ell), \\ E^{RG}_\ell &= (I - R^T_\ell A_\ell) M_\ell \{I - M_\ell P^{\ell}_{\ell+1}[I - (E^{RG}_{\ell+1})^{\gamma}]A^{-1}_{\ell+1}(M_\ell P^{l}_{\ell+1})^T A_\ell\} M_\ell (I - R_\ell A_\ell) \end{aligned}$$

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

$$E^{PG}_\ell = E^{RG}_\ell = 0. \tag{14}$$

Assume $E^{PG}_{\ell+1} = E^{RG}_{\ell+1}$ for some positive $l < L$. Using $M_\ell = S^{1/2}_\ell$, elementary manipulations give $E^{PG}_\ell = E^{RG}_\ell$. By induction, it follows that $E^{PG}_\ell = E^{RG}_\ell$ on every level, which completes the proof.
Now let us consider the case where the matrices $A_\ell$, $l = 2, \ldots, L$ are singular. If both Algorithms 1 and 3 use the Penrose-Moore pseudoinverse on the coarsest level, it is sufficient to realize that

$$\mathrm{Rng} A_{\ell+1} = \mathrm{Rng}(M_\ell P^{\ell}_{\ell+1})^T A_\ell = \mathrm{Rng}(P^{\ell}_{\ell+1})^T A_\ell S_\ell.$$

Then the inverses of $A_{\ell+1}$ in $E^{PG}_\ell, E^{RG}_\ell$ can be understood as inverses on $\mathrm{Rng} A_{\ell+1}$ and the proof holds without any further change.
If the coarsest solvers are general pseudoinversions, (14) has to be replaced by

$$\mathrm{Rng}(E^{PG}_\ell - E^{RG}_\ell) \subset \mathrm{Ker}(A_\ell) \tag{15}$$

for $l = L$. Since for $\mathbf{n}_{\ell+1} \in \mathrm{Ker}(A_{\ell+1})$ $S_\ell P^\ell_{\ell+1}\mathbf{n}_{\ell+1},\ M_\ell P^\ell_{\ell+1}\mathbf{n}_{\ell+1} \in \mathrm{Ker}(A_\ell)$, (15) holds by induction on every level. Since $A_1$ is regular, $E_1^{PG} = E_1^{RG}$ and the proof is completed.

We prove now, an abstract convergence estimate for Algorithm 3. Define the smoothed composite prolongator $I^1_\ell : \mathbb{R}^{n_\ell} \to \mathbb{R}^{n_1}$ by

$$I^1_\ell = M_1 P^1_2 \ldots M_{l-1} P^{l-1}_\ell, \quad I^1_1 = I, \tag{16}$$

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

$$V_\ell = \left(\mathrm{Rng} I^1_\ell,\ \|\cdot\|_\ell : \mathbf{u} \mapsto \min\{\|\mathbf{x}\|_{\mathbb{R}^{n_\ell}} : \mathbf{u} = I^1_\ell \mathbf{x}\}\right). \tag{17}$$

Note, that from (11) and (12) it follows that $A_\ell = (I^1_\ell)^T A I^1_\ell$, and

$$\|I^1_\ell \mathbf{x}\|_A = \|\mathbf{x}\|_{A_\ell} \quad \text{for all } \mathbf{x} \in \mathbb{R}^{n_\ell}, \tag{18}$$

$$\max_{\mathbf{u}\in V_\ell} \frac{\|\mathbf{u}\|_A}{\|\mathbf{u}\|_\ell} = \max_{\mathbf{x}\in\mathbb{R}^{n_\ell}} \frac{\|I^1_\ell \mathbf{x}\|_A}{\|\mathbf{x}\|_{\mathbb{R}^{n_\ell}}} = \sqrt{\varrho(A_\ell)}. \tag{19}$$

We define the symmetrized smoother preconditioners

$$\begin{aligned} \bar{R}'_\ell &= (I - K'^*_\ell K'_\ell) A^+_\ell, \\ \bar{R}_\ell &= (I - K^*_\ell K_\ell) A^+_\ell = (I - (I - R^T_\ell A_\ell)(I - R_\ell A_\ell)) A^+_\ell, \end{aligned} \tag{20}$$

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

$$\bar{R}'_{V_\ell} : I^1_\ell \mathbf{x},\ \mathbf{x} \in \mathbb{R}^{n_\ell} \ \mapsto\ I^1_\ell \bar{R}'_\ell \mathbf{x},$$

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

$$\min_{\mathbf{u}\in V_\ell} \frac{\|\bar{R}'_{V_\ell}\mathbf{u}\|_\ell}{\|\mathbf{u}\|_l} = \lambda_{\min}(\bar{R}'_\ell|_{\mathrm{Rng}A_\ell}), \tag{21}$$

where the symbol $|$ denotes the restriction.
Our estimates use an abstract convergence result as proved in the monography [1]. Using (19), (21) it can be written in our notation as follows:

**Lemma 3** *(Bramble [1], Theorem 3.3). Assume there are linear mappings $Q_\ell : V_1 \to V_\ell,\ Q_1 = I$ and constants $c_1,\ c_2 > 0$ such that*

$$\|Q_\ell \mathbf{u}\|_A \le c_1 \|\mathbf{u}\|_A \ \forall \mathbf{u} \in V_1, \ l = 1, \dots, L, \tag{22}$$

$$\|(Q_\ell - Q_{\ell+1})\mathbf{u}\|_\ell \le \frac{c_2}{\sqrt{\varrho(A_\ell)}} \|\mathbf{u}\|_A \ \forall \mathbf{u} \in V_1, \ l = 1, \dots, L-1. \tag{23}$$

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

$$\lambda_{\min}(\bar{R}'_\ell|_{\mathrm{Rng} A_\ell}) \ge \frac{1}{c_R^2 \ \varrho(A_\ell)}, \tag{24}$$

$$(R'_\ell A_\ell \mathbf{u}, R'_\ell A_\ell \mathbf{u})_{A_\ell} \le \theta (R'_\ell A_\ell \mathbf{u}, \mathbf{u})_{A_\ell} \ \forall \mathbf{u} \in \mathbb{R}^{n_\ell}. \tag{25}$$

*Then Algorithm 3 satisfies*

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

*where*

$$c_0 = \left(1 + c_1 + c_2 c_R \sqrt{\frac{\theta}{2-\theta}}\right)^2 \frac{L-1}{2-\theta}. \tag{26}$$

The application of this lemma to our method consisting in treating the spaces $V_\ell$ as finite element spaces is difficult because the natural basis of our coarse space $V_\ell \equiv \mathrm{Rng} I_\ell^1$ is $\{I_\ell^1 \mathbf{e}_i^\ell\}_{i=1}^{n_\ell}$, where $\mathbf{e}_i^\ell$ is the $i-$th canonical basis vector of $\mathbb{R}^{n_\ell}$. As the matrices $M_\ell = S_\ell^{1/2}$ are dense, $I_\ell^1$ is dense as well. Thus, the support of the basis function $I_\ell^1 \mathbf{e}_i^\ell$ is the entire computational domain. The verification of (22) and (23) is usually performed by using the standard Finite Element approximation theory and relies 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. However, the geometrical properties of the spaces of *disaggregated* vectors $P_\ell^1 \mathbf{x}$, $\mathbf{x} \in \mathbb{R}^{n_\ell}$ are very simple due to the fact that the aggregates are disjoint on each level. The following abstract result proved in [10] allows to verify (22) and (23) using separately the conditions on the ranges of the composite tentative prolongators $P_\ell^1$ and on the prolongator smoothers $S_\ell$. For the scaled tentative prolongators $P_\ell^1$ and prolongator smoothers being polynomials in the matrices $A_\ell$, it can be written down as follows:

**Assumption 3.1** *Let $S_\ell = s_\ell(A_\ell)$, $l = 1, \dots, L-1$ be polynomials in the matrices $A_\ell$ such that $0 \le S_\ell \le I$ and $(P_\ell^1)^T P_\ell^1 = I$ on every level. Further, let $\bar{\lambda}_\ell \ge \varrho(A_\ell)$ and*

$$\tilde{Q}_\ell : V_1 \to \mathbb{R}^{n_\ell}, \quad l = 2, \dots, L$$

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

$$\|\mathbf{u} - P^1_{\ell+1}\tilde{Q}_{\ell+1}\mathbf{u}\|^2_{\mathbb{R}^{n_1}} \leq \frac{C_1^2}{\bar{\lambda}_\ell}\|\mathbf{u}\|^2_A, \quad \forall \mathbf{u} \in V_1 \tag{27}$$

$$\|(I - S_\ell^{1/2})\mathbf{x}\|^2_{\mathbb{R}^{n_\ell}} \leq \frac{C_2^2}{\varrho(A_\ell)}\|\mathbf{x}\|^2_{A_\ell}, \quad \forall \mathbf{x} \in \mathbb{R}^{n_\ell} \tag{28}$$

$$\varrho(A_\ell S_\ell) \leq C_S^2 \bar{\lambda}_\ell. \tag{29}$$

Note that the prolongator smoothers enter the assumption through the scaling by $\bar{\lambda}_\ell$ on the right-hand side of (27). By making $\bar{\lambda}_\ell$ small, the prolongator smoothers make the approximation condition (27) easier to satisfy.

**Lemma 4** *Under the Assumption 3.1, for every* $\mathbf{u} \in V_1$*, the mappings*

$$Q_1 = I, \quad Q_\ell = I^1_\ell \tilde{Q}_\ell, \quad l = 2, \ldots, L$$

*satisfy*

$$\|Q_\ell \mathbf{u}\|_A \leq c_1 \|\mathbf{u}\|_A, \quad \forall l = 1, \ldots, L, \tag{30}$$

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

$$\|(Q_\ell - Q_{\ell+1})\mathbf{u}\|_l \leq c_2 \varrho(A_\ell)^{-1/2}\|\mathbf{u}\|_A, \quad \forall l = 1, \ldots, L-1 \tag{31}$$

*with* $c_2 = C_1 + C_2\|Q_\ell\|_A \leq C_1 + C_2 c_1$*.*

**PROOF.** To make this paper self consistent, we give below a version of the proof of Lemma 4. A proof using more general assumptions than assumptions 3.1 can be found in [10]. Throughout the proof we use the symbol $\mathbf{u}$ to denote an arbitrary finest level vector.
If (27) holds for some linear operators $\tilde{Q}_\ell$, it holds also for $\tilde{Q}_\ell$ such that $P^1_\ell \tilde{Q}_\ell$, $l = 2, \ldots, L$ are projections onto $\mathrm{Rng} P^1_\ell$ orthogonal with respect to $\mathbb{R}^{n_1}-$inner product. Hence, we can assume that $P^1_\ell \tilde{Q}_\ell$ are such projections without loosing generality. Then, setting $\tilde{Q}_1 = I$ and $P^1_1 = I$, the assumption (27) gives

$$\begin{aligned}\|(p^1_\ell \tilde{Q}_\ell - P^1_{\ell+1}\tilde{Q}_{\ell+1})\mathbf{u}\|^2_{\mathbb{R}^{n_1}} &= \|(I - P^1_{\ell+1}\tilde{Q}_{\ell+1})\mathbf{u}\|^2_{\mathbb{R}^{n_1}} - \|(I - P^1_l \tilde{Q}_\ell)\mathbf{u}\|^2_{\mathbb{R}^{n_1}} \\ &\leq \frac{C_1^2}{\bar{\lambda}_\ell}\|\mathbf{u}\|^2_A,\end{aligned} \tag{32}$$

as the decomposition

$$(I - P^1_{\ell+1}\tilde{Q}_{\ell+1})\mathbf{u} = (I - P^1_l\tilde{Q}_\ell)\mathbf{u} + (P^1_l\tilde{Q}_\ell - P^1_{\ell+1}\tilde{Q}_{\ell+1})\mathbf{u}$$

is orthogonal. Since $A_{\ell+1} = (P^\ell_{\ell+1})^T S_\ell A_\ell P^\ell_{\ell+1}$, $I^1_{\ell+1} = I^1_\ell S^{1/2}_\ell P^\ell_{\ell+1}$ and $S_\ell$ commutes with $A_\ell$, the assumptions $\varrho(S_\ell) \le 1$ and $\varrho(S_\ell A_\ell) \le C_S^2\bar{\lambda}_\ell$ give:

$$\begin{aligned}
\|Q_{\ell+1}\mathbf{u}\|_A &= \|I^1_{\ell+1}\tilde{Q}_{\ell+1}\mathbf{u}\|_A = \|S^{1/2}_\ell P^\ell_{\ell+1}\tilde{Q}_{\ell+1}\mathbf{u}\|_{A_\ell} \qquad (33)\\
&= \|S^{1/2}_\ell(P^\ell_{\ell+1}\tilde{Q}_{\ell+1} - \tilde{Q}_\ell)\mathbf{u} + S^{1/2}_\ell\tilde{Q}_\ell\mathbf{u}\|_{A_\ell}\\
&\le \|S^{1/2}_\ell(P^\ell_{\ell+1}\tilde{Q}_{\ell+1} - \tilde{Q}_\ell)\mathbf{u}\|_{A_\ell} + \|S^{1/2}_\ell\tilde{Q}_\ell\mathbf{u}\|_{A_\ell}\\
&\le C_S\bar{\lambda}^{1/2}_\ell\|(\tilde{Q}_\ell - P^\ell_{\ell+1}\tilde{Q}_{\ell+1})\mathbf{u}\|_{\mathbb{R}^{n_\ell}} + \|\tilde{Q}_\ell\mathbf{u}\|_{A_\ell}.
\end{aligned}$$

As $(P^1_l)^T P^1_\ell = I$, we have $\|P^1_\ell\mathbf{u}\|_{\mathbb{R}^{n_1}} = \|\mathbf{u}\|_{\mathbb{R}^{n_\ell}}$ and the previous inequality together with (32) and (18) yields

$$\|Q_{\ell+1}\mathbf{u}\|_A \le C_S\bar{\lambda}^{1/2}_\ell\|(P^1_\ell\tilde{Q}_\ell - P^\ell_{\ell+1}\tilde{Q}_{\ell+1})\mathbf{u}\|_{\mathbb{R}^{n_1}} + \|\tilde{Q}_\ell\mathbf{u}\|_{A_\ell} \le C_S C_1\|\mathbf{u}\|_A + \|Q_\ell\mathbf{u}\|_A$$

Now (30) follows by induction using $Q_1 \equiv I$.

Let us prove (31). From the definition of $Q_\ell$, the identity $I^1_{\ell+1} = I^1_\ell S^{1/2}_\ell P^\ell_{\ell+1}$, the definition (17) and the assumption (28) we have

$$\begin{aligned}
\|(Q_\ell - Q_{\ell+1})\mathbf{u}\|_\ell &= \|I^1_\ell(\tilde{Q}_\ell - S^{1/2}_\ell P^\ell_{\ell+1}\tilde{Q}_{\ell+1})\mathbf{u}\|_\ell \le \|(\tilde{Q}_\ell - S^{1/2}_\ell P^\ell_{\ell+1}\tilde{Q}_{\ell+1})\mathbf{u}\|_{\mathbb{R}^{n_\ell}}\\
&= \|S^{1/2}_\ell(\tilde{Q}_\ell - P^\ell_{\ell+1}\tilde{Q}_{\ell+1})\mathbf{u} + (I - S^{1/2}_\ell)\tilde{Q}_\ell\mathbf{u}\|_{\mathbb{R}^{n_\ell}}\\
&\le \|S^{1/2}_\ell(\tilde{Q}_\ell - P^\ell_{\ell+1}\tilde{Q}_{\ell+1})\mathbf{u}\|_{\mathbb{R}^{n_\ell}} + \|(I - S^{1/2}_\ell)\tilde{Q}_\ell\mathbf{u}\|_{\mathbb{R}^{n_\ell}}\\
&\le \|S^{1/2}_\ell(\tilde{Q}_\ell - P^\ell_{\ell+1}\tilde{Q}_{\ell+1})\mathbf{u}\|_{\mathbb{R}^{n_\ell}} + C_2\varrho(A_\ell)^{-1/2}\|Q_\ell\mathbf{u}\|_A. \qquad (34)
\end{aligned}$$

Using $(P^1_\ell)^T P^1_\ell = I$ and $\varrho(S_\ell) \le 1$ again, gives

$$\|S^{1/2}_\ell(\tilde{Q}_\ell - P^\ell_{\ell+1}\tilde{Q}_{\ell+1})\mathbf{u}\|_{\mathbb{R}^{n_\ell}} \le \|(\tilde{Q}_\ell - P^\ell_{\ell+1}\tilde{Q}_{\ell+1})\mathbf{u}\|_{\mathbb{R}^{n_\ell}} = \|P^1_\ell(\tilde{Q}_\ell - P^\ell_{\ell+1}\tilde{Q}_{\ell+1})\mathbf{u}\|_{\mathbb{R}^{n_1}}.$$

The last inequality together with (34), (30) and (32) completes the proof of (31).

The following lemma translates the assumptions (24), (25) on the pre-smoothers of the Ritz-Galerkin Algorithm 3 into the requirements on the simpler pre-smoothers of the original Petrov-Galerkin Algorithm 1.

**Lemma 5** *For every level $l = 1, \ldots, L-1$ it holds that*

$$\lambda_{\min}(\bar{R}'_\ell|_{\mathrm{Rng}A_\ell}) \geq \min\left\{\frac{1}{\varrho(A_\ell)}, \lambda_{\min}(\bar{R}_\ell)\right\}. \tag{35}$$

*If, in addition, $R_\ell$ is symmetric positive definite, commuting with $A_\ell$, the inequality in (25) holds with*

$$\theta = \max\{1, \varrho(R_\ell A_\ell)\}. \tag{36}$$

**PROOF.** We start with proving (35). Using $S_\ell = M_\ell^2$, the commutativity of $A_\ell$ and $M_\ell$, definition (20) and remark 3, we get

$$\begin{aligned}
\bar{R}'_\ell &= [I - M_\ell(I - R_\ell^T A_\ell)(I - R_\ell A_\ell)M_\ell]A_\ell^+ \\
&= (I - S_\ell)A_\ell^+ + M_\ell[I - (I - R_\ell^T A_\ell)(I - R_\ell A_\ell)]A_\ell^+ M_\ell \\
&= (I - S_\ell)A_\ell^+ + M_\ell \bar{R}_\ell M_\ell \;\geq\; (I - S_\ell)A_\ell^+ + \lambda_{\min}(\bar{R}_\ell)S_\ell
\end{aligned}$$

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 $S_\ell = s_\ell(A_\ell)$, where $s_\ell$ is a polynomial,

$$\begin{aligned}
\lambda_{\min}(\bar{R}'_\ell|_{\mathrm{Rng}A_\ell}) &\geq \min_{\mathbf{x}\in\mathrm{Rng}A_\ell}\left(\frac{\mathbf{x}^T(I - S_\ell)A_\ell^+\mathbf{x}}{\mathbf{x}^T\mathbf{x}} + \lambda_{\min}(\bar{R}_\ell)\frac{\mathbf{x}^T S_\ell \mathbf{x}}{\mathbf{x}^T\mathbf{x}}\right) \\
&\geq \min\left\{\frac{1 - s_\ell(t)}{t} + s_\ell(t)\lambda_{\min}(\bar{R}_\ell) \;\mid t \in \sigma(A_\ell),\ t \neq 0\right\} \\
&\geq \min\left\{\frac{1 - s_\ell(t)}{\varrho(A_\ell)} + s_\ell(t)\lambda_{\min}(\bar{R}_\ell) \;\mid t \in \sigma(A_\ell),\ t \neq 0\right\}.
\end{aligned}$$

As $s_\ell(t) \in [0,1]$ for $t \in [0, \varrho(A_\ell)]$ (see the forthcoming lemma 8), the last estimate is a convex combination of $\varrho(A_\ell)^{-1}$ and $\lambda_{\min}(\bar{R}_\ell)$, completing the proof of (35).
Let us prove (36). By (13), it follows that

$$R'_\ell A_\ell = I - M_\ell + M_\ell R_\ell A_\ell.$$

By assumption, $R_\ell$ is symmetric positive definite and it commutes with $A_\ell$. Hence, taking into account that $M_\ell = s_\ell^{1/2}(A_\ell)$, the product $R'_\ell A_\ell$ is both symmetric and $A_\ell$–symmetric. As $s_\ell(t) \in [0,1]$ for $t \in \sigma(A_\ell)$, $R'_\ell A_\ell$ is also positive semi-definite and (25) is satisfied with $\theta = \varrho(R'_\ell A_\ell)$. Further, $R'_\ell A_\ell \leq I - M_\ell + \varrho(R_\ell A_\ell) M_\ell$ , and by the spectral mapping theorem

$$\varrho(R'_\ell A_\ell) \leq \min_{t \in [0, \varrho(A_\ell)]} \{1 - s_\ell^{1/2}(t) + s_\ell^{1/2}(t) \varrho(R_\ell A_\ell)\}.$$

Now, the proof of (36) follows from the same convex combination argument as the proof of (35).

Now we are ready to prove the abstract convergence theorem for the Algorithm 1.

**Assumption 3.2 (Smoothing property)** *Assume the preconditioners $R_\ell$, $l = 1, \ldots, L-1$ of the pre-smoothers in the Algorithm 1 are chosen so that*

$$\lambda_{\min}(\bar{R}_\ell) \geq \frac{1}{C_R^2 \varrho(A_\ell)} \tag{37}$$

*and either*

*(1) each $R_\ell$, $l = 1, \ldots, L-1$ is a symmetric positive semi-definite matrix commuting with $A_\ell$ such that $\varrho(R_\ell A_\ell) \leq \theta < 2$, or*
*(2) (25) is satisfied for*

$$R'_\ell = (I - K'_\ell) A_\ell^+, \quad K'_\ell = (I - R_\ell A_\ell) S_\ell^{1/2}.$$

**Theorem 6** *Under the Assumptions 3.1 and 3.2, it holds for the Algorithm 1 that*

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

*where*

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

*Here, $C_1, C_2, C_S, C_R$ and $\theta$ are the constants from the Assumptions 3.1 and 3.2.*

**PROOF.** In the view of Lemma 2, it is sufficient to verify the assumptions (22), (23), and (24), (25). The assumptions (22) and (23) are verified in Lemma 4, the smoothing properties (24) and (25) follow from Lemma 5.

## 4 Prolongator smoother

This section specifies suitable smoothers $S_\ell$ using the Assumption 3.1 as a guideline. Our objective is to construct the smoothers that

a) minimize the available estimates $\bar{\lambda}_\ell \geq \varrho(A_\ell)$ on the right-hand side of (27) in order to make (27) easier to satisfy,
b) satisfy the constraint (28) with a reasonable constant $C_2$ independent of $\sigma(A_\ell)$.

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

**Lemma 7** *Let* $\bar{\lambda}_1 \geq \varrho(A_1)$. *Then,*

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

*where* $C_S$ *is defined as in (29).*

**PROOF.** The statement holds by definition for $l = 1$. Assume it holds for some $l \geq 1$. Then by (2), (29) and (10),

$$\varrho(A_{\ell+1}) = \max_{\mathbf{x} \in \mathbb{R}^{n_{\ell+1}}} \frac{(A_\ell S_\ell P^\ell_{\ell+1}\mathbf{x}, P^\ell_{\ell+1}\mathbf{x})_{\mathbb{R}^{n_\ell}}}{\|\mathbf{x}\|^2_{\mathbb{R}^{n_{\ell+1}}}} \leq C_S^2 \bar{\lambda}_\ell \frac{\|P^\ell_{\ell+1}\mathbf{x}\|^2_{\mathbb{R}^{n_\ell}}}{\|\mathbf{x}\|^2_{\mathbb{R}^{n_{\ell+1}}}} = C_S^2 \bar{\lambda}_l.$$

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

With Lemma 7 in mind, we choose $S_\ell = s(\bar{\lambda}_\ell^{-1} A_\ell)$, where $s$ is a polynomial of a given degree minimizing

$$\max_{x \in [0,1]} p(x)x \ \text{ subject to } p(0) = 1. \tag{38}$$

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

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

and thus the possibility to verify assumption (27) will rely on the fact that we can construct a prolongator smoother such that $\bar{\lambda}_\ell \sim \bar{\lambda}_1/(d+1)^{2(l-1)}$. The following lemma will show that it is indeed possible. For polynomial smoothers of even degree, the results (42) and (43) have already been established in [9].

**Lemma 8** *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*

$$\min_{p\in\mathcal{P}_n}\Big(\max_{x\in[0,1]} p(x)x\Big) = \max_{x\in[0,1]} s(x)x. \tag{39}$$

*The polynomial $s(x)$ is given by*

$$s(x) = \begin{cases} \prod_{k=1}^{r}\left(1 - \frac{x}{r_k}\right)^2 \cdot (1-x) & \text{for odd } n = 2r+1 \\ \prod_{k=1}^{r}\left(1 - \frac{x}{r_k}\right)^2 & \text{for even } n = 2r \end{cases} \tag{40}$$

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

$$r_k = \frac{1}{2}\left(1 - \cos\frac{2k\pi}{n+1}\right) = \sin^2\frac{k\pi}{n+1}, \qquad k = 1,\ldots,r \tag{41}$$

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

$$\max_{x\in[0,1]} s(x)x = \frac{1}{(n+1)^2} \tag{42}$$

$$0 \le s(x) \le 1, \qquad x \in [0,1] \tag{43}$$

$$[1 - s^{1/2}(x)]^2 \le \left(\frac{1}{2} + \frac{\pi^2}{12}\right)(n+1)^2 x, \qquad \forall x \in [0,1] \tag{44}$$

**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:

$$s(x)x = q_{n+1}(x) = \frac{c}{2}\frac{(\|w_{n+1}\|_\infty - w_{n+1}(1-2x))}{\|w_{n+1}\|_\infty}, \tag{45}$$

where $c = \max_{x\in[0,1]} q_{n+1}(x)$. Because of the constraint $s(0) = 1$, we also have $1 = q'_{n+1}(0) = cw'_{n+1}(1)/\|w_{n+1}\|_\infty$ and hence $c = \|w_{n+1}\|_\infty / w'_{n+1}(1)$. If we want to minimize $c$ to satisfy (39), 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

$$s(x)x = q_{n+1}(x) = \frac{c}{2}(1 - w_{n+1}(1-2x)), \tag{46}$$

where $c = \max_{x\in[0,1]} q_{n+1}(x) = (w'_{n+1}(1))^{-1} == 1/(n+1)^2$. This proves (39) and (42).

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 (41), 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 (41).

For proving (43) we use $w_{n+1}$ in (45); we will show that

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

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

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

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 (44). 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 (44) 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:

$$1 + \tilde{s}'(0)x \leq \tilde{s}(x), \qquad x \in [0,1]. \tag{47}$$

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 (47) 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 (47) and (43) we can prove (44):

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

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 $S_\ell$ is summed up in the following algorithm.

**Algorithm 4 (Prolongator smoothers)** *For a given level $l$, the matrix $A_\ell$, the estimate $\bar{\lambda}_1 \geq \varrho(A_1)$ and the degree of the prolongator smoother $d$ define the prolongator smoother $S_\ell$ as follows:*

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

$$\bar{\lambda}_\ell = \min \left\{ \bar{\varrho}(A_\ell), \frac{\bar{\lambda}_1}{(d+1)^{2(l-1)}} \right\},^{1}$$

*2. Set $S_\ell = s(\bar{\lambda}_\ell^{-1} A_\ell)$, where $s$ is the polynomial of the degree $d$ given by (40).*

The proof of the following statement follows directly from Lemma 7 and 8.

**Assumption 4.1** *Let the tentative prolongators $P_{\ell+1}^{\ell}$ be constructed by the Algorithm 2 and the prolongator smoothers $S_\ell$ are polynomials of the degree $d$ defined by Algorithm 4. We assume that there exist linear mappings*

$$\tilde{Q}_\ell : \mathbb{R}^{n_1} \to \mathbb{R}^{n_\ell},\ l = 2, \ldots, L$$

*such that*

$$\|\mathbf{u} - P_l^1 \tilde{Q}_l \mathbf{u}\|_{\mathbb{R}^{n_1}} \leq \frac{C_A (d+1)^{l-1}}{\sqrt{\bar{\lambda}_1}} \|\mathbf{u}\|_A \ \forall \mathbf{u} \in \mathbb{R}^{n_1},\ l = 2, \ldots, L \tag{48}$$

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

**Lemma 9** *Under the Assumption 4.1, the Assumption 3.1 holds with constants*

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

## 5 Model example

Let $\Omega \subset \mathbb{R}^D,\ D = 2, 3$ be a bounded domain, $\tau_h$ a quasiuniform finite element mesh on $\Omega$, and $V_h$ a P1 or Q1 finite element space associated with $\tau_h$. 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

$$c|u|^2_{H^1(\Omega)} \leq a(u, u) \leq C|u|^2_{H^1(\Omega)} \tag{49}$$

for every $u \in H^1(\Omega)$.
Assume the tentative prolongators have been constructed by Algorithm 2, the

prolongator smoothers are polynomials of the degree $d$ defined by Algorithm 4 and the pre-smoother preconditioners $R_\ell$ satisfy Assumption 3.2 (e.g. $R_\ell = \varrho(A_\ell)^{-1}I$.)
For each aggregate $\mathcal{A}_i^\ell$, define the *composite aggregate* $\hat{\mathcal{A}}_i^\ell$ to be the aggregate $\mathcal{A}_i^\ell$ understood as the set of degrees of freedom on the level 1, i.e.

$$\hat{\mathcal{A}}_i^\ell \equiv \{j \;:\; (P_\ell^1)_{ji} \neq 0\}, \quad l = 1,\dots,L-1,\ i = 1,\dots,n_\ell.$$

We assume that the composite aggregates satisfy the following geometrical assumptions: For each aggregate $\mathcal{A}_i^\ell$ there is a ball $U_i^\ell \subset \mathbb{R}^d$ such that:

(1) all degrees of freedom contained in composite aggregate $\tilde{\mathcal{A}}_i^l$ are located within $U_i^\ell$,
(2) $\mathrm{diam} U_i^\ell \le C(d+1)^{l-1}h$, where $h$ is the characteristic meshsize of $\tau_h$, $d$ 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 $U_i^\ell$. (Overlaps of balls $\{U_i^\ell\}_{i=1}^{n_\ell}$ are bounded.)

The analysis consists in verification of the Assumption 4.1 and uses Lemma 9 and Theorem 6 to carry out the convergence estimate.
The weak approximation property (48) depends only on the range of $P_\ell^1$ and the scaling performed inside the main loop of the Algorithm 2 does not change $\mathrm{Rng} P_\ell^1$. Therefore, we can carry out the estimates for the nonscaled tentative prolongators as defined in (8) without loosing generality.
The key tool for our verification is the following scaled Poincaré-Friedrichs inequality: For a ball $U \subset \mathbb{R}^D$, $U' \subset U$ and every function $u \in H^1(U)$ it holds that

$$\|u - p\|_{L^2(U')} \le C\ \mathrm{diam} U |u|_{H^1(U)}, \quad p = \int_U u\, d\mathbf{x}. \tag{50}$$

With (50) in mind, we define $\tilde{Q}_\ell,\ l = 1,\dots,L$ as follows: Let $\Pi : \mathbb{R}^{n_1} \mapsto H^1(\mathbb{R}^D)$ be the mapping that for $\mathbf{x} \in \mathbb{R}^{n_1}$ returns $\sum_{i=1}^{n_1} x_i\varphi_i$ extended by zero outside $\Omega$. We set

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

As $P_\ell^1$ is a nonscaled composite prolongator, $(P_\ell^1\mathbf{w}^\ell)_j = w_i^\ell$ for $j \in \hat{\mathcal{A}}_i^\ell$ and,

$$\|\mathbf{u} - P_\ell^1\tilde{Q}_\ell\mathbf{u}\|^2 = \sum_{i=1}^{n_\ell} \|\mathbf{u} - P_\ell^1\mathbf{w}^\ell\|^2_{l^2(\hat{\mathcal{A}}_i^\ell)} = \sum_{i=1}^{n_\ell} \|\mathbf{u} - w_i^\ell\|^2_{l^2(\hat{\mathcal{A}}_i^\ell)}, \tag{51}$$

where $\|\cdot\|_{l^2(\hat{\mathcal{A}}_i^\ell)} : \mathbf{x} \in \mathbb{R}^{n_1} \mapsto \sqrt{\sum_{k\in\hat{\mathcal{A}}_i^\ell} x_k^2}$. Using the quasiuniformity of $\tau_h$, the well-known equivalence of the discrete and continuous $L^2-$norms, (50) and the fact that $\Pi\mathbf{u}$ vanishes outside $\Omega$ we get

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

Substituting the last estimate into (51), using the bounded overlaps of $U_i^\ell$, $\mathrm{diam}U_i^\ell \leq C(d+1)^{l-1}h$, the uniform equivalence (49) and the well-known estimate $\varrho(A_1) \leq Ch^{D-2}$ we get (48) with $\bar{\lambda}_1 = Ch^{D-2}$.
Now, Lemma 9 together with Theorem 6 gives the estimate

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

with a constant $C$ independent of $L$, $h$, and $\Omega$.

## 6 Conclusion

On unstructured grids, the construction of coarse grids is a difficult task [4],[5]. For this reason, algebraic multigrid (AMG) is often used to create the coarse spaces from unstructured meshes. However, a rigourous convergence theory of the classical AMG method does not exist. The work in [8] introduces the concept of *smoothed aggregation* to construct the coarse spaces. This new AMG method possesses a rigourous convergence theory exposed in [10]. However, the smoothed aggregation AMG method of [8] is a Ritz-Galerkin method and implies a coarsening factor of $H/h = 3$ (instead of the usual coarsening factor of $H/h = 2$). In this paper, we describe a smoothed aggregation AMG method using a modified smoothed operator. The algorithm is based on a Petrov-Galerkin approach and can handle the usual coarsening factor of $H/h = 2$. As for the method of [8], we show in this paper that this method possesses a rigourous convergence theory.

## Acknowledgements

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] Hemker, P.W. On the order of prolongations and restrictions in multigrid procedures. J. Comput. Appl. Math 32 (1990) 423-429.

[3] Hemker, P.W., Koren, B., Lioen, W.M., Nool, M. and van der Maarel, H.T.M. Multigrid for steady gas dynamics problems. Computational Fluid Dynamics Review 1 (1995) 477-494.

[4] H. Guillard, Node-nested multigrid method with Delaunay Coarsening, *INRIA Rpt* No **1898** (May 1993).

[5] E. Morano, A. Dervieux, H. Guillard, M.P. Leclercq, B. Stoufflet, Faster Relaxations for non-structured MG with Voronoi coarsening in: Hirsch et al., eds. *Proceedings of CFD'92* **vol I** (Elsevier 1992), 69–74.

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

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

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

[9] P. Vaněk, J. Mandel, M. Brezina, Two Level Algebraic multigrid for the Helmholtz problem, in: J. Mandel, C. Farhat, X.C Cai, eds., *Comtemporary Mathematics* **218** (Proceedings of the $10^{th}$ International conference on domain decomposition, Boulder C0, 1997) (AMS Providence 1998) 349–356.

[10] P. Vaněk, M. Brezina, J. Mandel, Convergence of Algebraic Multigrid Based on Smoothed Aggregation, *Numer. Math.* **88** no. 3 (2001) 559–579.