Applications of Mathematics

Petr Vaněk; Marian Brezina

Nearly optimal convergence result for multigrid with aggressive coarsening and polynomial smoothing

Applications of Mathematics, Vol. 58 (2013), No. 4, 369–388

Persistent URL: http://dml.cz/dmlcz/143335

NEARLY OPTIMAL CONVERGENCE RESULT FOR MULTIGRID WITH AGGRESSIVE COARSENING AND POLYNOMIAL SMOOTHING

Petr Vaněk, Plzeň, Marian Brezina, Boulder

(Received January 24, 2011)

Abstract. We analyze a general multigrid method with aggressive coarsening and polynomial smoothing. We use a special polynomial smoother that originates in the context of the smoothed aggregation method. Assuming the degree of the smoothing polynomial is, on each level k, at least Ch_{k+1}/h_k, we prove a convergence result independent of h_{k+1}/h_k. The suggested smoother is cheaper than the overlapping Schwarz method that allows to prove the same result. Moreover, unlike in the case of the overlapping Schwarz method, analysis of our smoother is completely algebraic and independent of geometry of the problem and prolongators (the geometry of coarse spaces).

Keywords: multigrid, aggressive coarsening, optimal convergence result

MSC 2010: 65F10, 65M55

1. Introduction

This paper is concerned with convergence of a multigrid method featuring aggressive coarsening. We analyze a general (abstract) multigrid algorithm with a special polynomial smoother that allows to prove a convergence bound independent of the relative size of the spaces on subsequent levels.

Assuming that the resolution on level k can be characterized by a meshsize h_k, and employing a carefully designed polynomial smoother as a multigrid relaxation

This work was sponsored by the TAČR (Technologická Agentura České Republiky) grant TA01020352, ITI (Institut Teoretické Informatiky) grant 1M0545, US Department of Energy under grant numbers DE-FG02-03ER25574 and DE-FC02-06ER25784, Lawrence Livermore National Laboratory under contract number B588587, the National Science Foundation under grant numbers DMS-0621199, DMS-0749317 and DMS-0811275 and Department of the Navy Grant N62909-11-1-7032, USA.

process, we prove a convergence result independent of the ratio h_{k+1}/h_k, provided that the degree of our smoother is greater than or equal to Ch_{k+1}/h_k, where C is a positive constant influencing convergence, and h_k and h_{k+1} are the characteristic resolutions of finer and coarser level, respectively (throughout the paper l denotes the coarsest level, and 0 the finest level). Thus, we allow the coarse space to be dramatically smaller than the preceding fine space and still obtain a multilevel convergence result not influenced by the ratio of their sizes. Here the aggressive coarsening is compensated for by a more powerful multigrid relaxation that consists of a sequence of Richardson type sweeps whose number is at least Ch_{k+1}/h_k, $C > 0$. We note that the assumption of existence of characteristic meshsize on each level is a sufficient condition, and the abstract convergence result presented in Theorem 4.1 does not depend on this assumption. The assumption will, however, allow us to verify the prerequisites of the abstract convergence result for the model problem considered. We stress that the abstract theory (Theorem 4.1) is not restricted to the quasiuniform case.

The smoother we use was originally derived from a prolongator smoother in the context of the smoothed aggregation method [9], [7], [4], and the previously proved theory [9], [8] was also limited to that context. Recent improvements of the convergence theory in [4], establishing the same convergence result as presented here, were also restricted to the smoothed aggregation method.

The regularity-free theory of [2] is known to derive no theoretical benefit from the use of more than $O(1)$ smoothing steps. Thus, until recently, the authors believed that the current result was possible only within the framework of smoothed aggregation, that is, smoothing the prolongator was deemed essential to establishing the result. The earlier works on this topic [6], [9], [7], depend crucially on this argument. In this paper, we prove a nearly optimal multilevel convergence result for this smoother used in general multigrid, and show that with a special choice of iteration parameters, Ch_{k+1}/h_k smoothing steps suffice to prove convergence independent of how aggressive the coarsening is. The near optimality of the convergence estimate is understood in the sense of the regularity-free theory [2], i.e., the convergence bound has a linear dependence on the number of levels.

Note that a similar convergence result can be proved for the overlapping Schwarz smoother. However, the relevant analysis requires verification of geometry-dependent assumptions on the overlapping subdomains which are tied to the geometrical properties of the coarse-level basis. In contrast, all assumptions on our smoother are strictly algebraic, and the analysis of the smoother is therefore independent of the geometry of the problem or the particular choice of prolongation operators (geometrical properties of coarse-levels). For our smoother, we only need the assumption that its degree is sufficiently large, which in the quasiuniform case means, greater than or equal to Ch_{k+1}/h_k, $C > 0$. Moreover, our smoother is cheaper than the overlapping Schwarz method, while its polynomial nature allows for easy and efficient parallel implementation whenever a highly tuned parallel matrix-vector multiply subroutine is available.

The paper is organized as follows: Section 2 presents a convergence result for the multigrid method in an abstract setting. As usual, we require that the multigrid relaxation process satisfies a smoothing condition, and that the hierarchy of coarse spaces (and the associated prolongators) satisfies a weak approximation property. By the very nature of aggressive coarsening, the smoothing procedure needs to do some of the work done under normal circumstances by the coarse-grid correction process. Therefore we have a weaker approximation condition and correspondingly a stronger smoothing condition. We see the introduction of this relaxed weak approximation condition, (2.14), as the main contribution of this paper. The corresponding stronger smoothing condition, (2.16), is shown to be satisfied by our choice of the polynomial smoother.

Section 3 is devoted to the analysis of an appropriate polynomial smoother sufficient to satisfy the smoothing condition (2.16), needed to establish the convergence bound. The smoother analysis presented here essentially follows [4] and presents a minor generalization.

Section 4 summarizes the results presented in Sections 2 and 3 in the form of a final abstract convergence estimate.

We conclude the paper by considering a model example in Section 5. Here we demonstrate how the abstract result can be applied to obtain a convergence result independent of the coarsening aggressivity in the case of a model example of a geometric multigrid method with aggressive coarsening for a simple H_0^1-equivalent model problem discretized over a quasiuniform mesh.

2. Multigrid algorithm and abstract estimates

We are solving a problem

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

with a symmetric positive definite (s.p.d.) matrix A of order n. We set $A_0 = A$ and $n_0 = n$. We assume injective prolongators

$$P_{k+1}^k \colon \mathbb{R}^{n_{k+1}} \to \mathbb{R}^{n_k}, \quad n_{k+1} < n_k,\ k = 0, \dots, l-1,$$

where l is the number of levels, are given. Define a *composite* prolongator

$$P_k^0 = P_1^0 \dots P_k^{k-1},$$

and assume that the coarse-level matrices are defined by the usual variational (Galerkin) formula

$$A_{k+1} = (P_{k+1}^k)^T A_k P_{k+1}^k = (P_{k+1}^0)^T A P_{k+1}^0. \tag{2.1}$$

To define a standard V-cycle multigrid, in addition to the hierarchy of matrices $\{A_k\}$ and prolongators $\{P_{k+1}^k\}$, we also need a multigrid relaxation, defined here on level k as an iterative process with an error propagation operator $I - M_k^{-1} A_k$. We assume that the smoothing matrices M_k are such that the relaxation process is an A_k-convergent iterative method, which is equivalent to $M_k^T + M_k - A_k$ being a positive definite matrix. We, in fact, assume that there is a constant $\alpha > 0$, uniform with respect to $k \geqslant 0$, such that,

$$\mathbf{v}_k^T (M_k^T + M_k - A_k)\mathbf{v}_k \geqslant \alpha \mathbf{v}_k^T A_k \mathbf{v}_k \quad \text{for all } \mathbf{v}_k \in \mathbb{R}^{n_k}. \tag{2.2}$$

We denote by $\overline{M}_k$ the symmetrized smoother

$$\overline{M}_k = M_k (M_k^T + M_k - A_k)^{-1} M_k^T. \tag{2.3}$$

It can be defined implicitly from the relation

$$I - \overline{M}_k^{-1} A_k = (I - M_k^{-T} A_k)(I - M_k^{-1} A_k). \tag{2.4}$$

Based on a given choice of P_{k+1}^k, M_k (that is A_k-convergent) for $0 \leqslant k \leqslant l-1$, and A_k obtained variationally from A_{k-1} for $1 \leqslant k \leqslant l$, starting with $B_l = A_l$, for $k = l-1, \ldots, 1, 0$, we recursively define a V-cycle preconditioner (a s.p.d. matrix) B_k in the following standard way:

$$I - B_k^{-1} A_k = (I - M_k^{-T} A_k)\,(I - P_{k+1}^k B_{k+1}^{-1} (P_{k+1}^k)^T A_k)\,(I - M_k^{-1} A_k).$$

Letting $B = B_0$, we are concerned in what occurs with the (upper) bound K_* in the estimate

$$\mathbf{v}^T A \mathbf{v} \leqslant \mathbf{v}^T B \mathbf{v} \leqslant K_\star \mathbf{v}^T A \mathbf{v} \tag{2.5}$$

(the lower bound holds because our algorithm is a variational multigrid). In what follows, $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ denote the Euclidean norm and the inner product in the relevant vector space. Further, for a symmetric positive definite matrix B, we define $\langle \cdot, \cdot \rangle_B = \langle B \cdot, \cdot \rangle$ and $\|\cdot\|_B = \langle \cdot, \cdot \rangle_B^{1/2}$.

Our analysis is based on the XZ-identity ([11]), formulated here in its matrix-vector form suitable for our purposes as follows. Given multigrid smoothers defined by M_j such that $M_j^T + M_j - A_j$ are the s.p.d., interpolation matrices P_{j+1}^j, and the coarse matrices defined as $A_{j+1} = (P_{j+1}^j)^T A_j P_{j+1}^j$, the following XZ-identity holds (cf. [10]):

$$
\begin{aligned}
(2.6)\qquad \mathbf{v}^T A\mathbf{v} &\leqslant \mathbf{v}^T B\mathbf{v} \\
&= \inf_{\{\mathbf{v}_k\}} \left\{ \|\mathbf{v}_l\|_{A_l}^2 + \sum_{j=0}^{l-1} \|M_j^T \mathbf{v}_j^f + A_j P_{j+1}^j \mathbf{v}_{j+1}\|^2_{(M_j^T+M_j-A_j)^{-1}} \right\}, \\
\mathbf{v}_0 = \mathbf{v}, &\quad \mathbf{v}_k^f \equiv \mathbf{v}_k - P_{k+1}^k \mathbf{v}_{k+1}.
\end{aligned}
$$

The infimum here is taken over the components $\{\mathbf{v}_k\}$ of all possible decompositions of $\mathbf{v}$ obtained as follows: Starting with $\mathbf{v}_0 = \mathbf{v}$, for $k \geqslant 0$, $\mathbf{v}_k = \mathbf{v}_k^f + P_{k+1}^k \mathbf{v}_{k+1}$, i.e., choosing $\mathbf{v}_{k+1} \in \mathbb{R}^{n_{k+1}}$ arbitrary, we then let $\mathbf{v}_k^f = \mathbf{v}_k - P_{k+1}^k \mathbf{v}_{k+1}$.

We observe that applying the triangle inequality together with the trivial inequality $(a+b)^2 \leqslant 2(a^2+b^2)$, $a, b \in \mathbb{R}$ and using the property (2.2), results in the estimate

$$
\begin{aligned}
(2.7)\qquad &\sum_{j=0}^{l-1} \|M_j^T \mathbf{v}_j^f + A_j P_{j+1}^j \mathbf{v}_{j+1}\|^2_{(M_j^T+M_j-A_j)^{-1}} \\
&\leqslant \sum_{j=0}^{l-1} 2(\|M_j \mathbf{v}_j^f\|^2_{(M_j^T+M_j-A_j)^{-1}} + \|A_j P_{j+1}^j \mathbf{v}_{j+1}\|^2_{(M_j^T+M_j-A_j)^{-1}}) \\
&= 2\sum_{j=0}^{l-1} \|M_j \mathbf{v}_j^f\|^2_{(M_j^T+M_j-A_j)^{-1}} + 2\sum_{j=0}^{l-1} \|A_j P_{j+1}^j \mathbf{v}_{j+1}\|^2_{(M_j^T+M_j-A_j)^{-1}} \\
&\leqslant 2\sum_{j=0}^{l-1} \|\mathbf{v}_j^f\|^2_{\overline{M}_j} + \frac{2}{\alpha}\sum_{j=0}^{l-1} \|A_j P_{j+1}^j \mathbf{v}_{j+1}\|^2_{A_j^{-1}} \\
&= 2\sum_{j=0}^{l-1} \|\mathbf{v}_j^f\|^2_{\overline{M}_j} + \frac{2}{\alpha}\sum_{j=0}^{l-1} \|P_{j+1}^j \mathbf{v}_{j+1}\|^2_{A_j} \\
&= 2\sum_{j=0}^{l-1} \|\mathbf{v}_j^f\|^2_{\overline{M}_j} + \frac{2}{\alpha}\sum_{j=0}^{l-1} \|\mathbf{v}_{j+1}\|^2_{A_{j+1}} \\
&= 2\sum_{j=0}^{l-1} \|\mathbf{v}_j^f\|^2_{\overline{M}_j} + \frac{2}{\alpha}\sum_{j=1}^{l} \|\mathbf{v}_j\|^2_{A_j}.
\end{aligned}
$$

From here, we see that in order to bound the relative condition number of the V-cycle preconditioner B with respect to A based on estimates (2.6) and (2.7), it is sufficient to bound the expressions below in terms of $\|\mathbf{v}\|_A^2$ for some particular choice of $\{\mathbf{v}_k\}$:

$$\sum_{k=0}^{l-1} \|\mathbf{v}_k^f\|_{\overline{M}_k}^2 \leqslant C_1 \|\mathbf{v}\|_A^2, \tag{2.8}$$

$$\sum_{k=1}^{l} \|\mathbf{v}_k\|_{A_k}^2 \leqslant C_2 \|\mathbf{v}\|_A^2, \tag{2.9}$$

and

$$\|\mathbf{v}_l\|_{A_l}^2 \leqslant C_3 \|\mathbf{v}\|_A^2. \tag{2.10}$$

Indeed, the estimates (2.6) and (2.7) give

$$\begin{aligned}\mathbf{v}^T A\mathbf{v} \leqslant \mathbf{v}^T B\mathbf{v} &\leqslant \|\mathbf{v}_l\|_{A_l}^2 + 2\sum_{j=0}^{l-1} \|\mathbf{v}_j^f\|_{\overline{M}_j}^2 + \frac{2}{\alpha}\sum_{j=1}^{l} \|\mathbf{v}_j\|_{A_j}^2 \\ &\leqslant \Big(C_3 + 2C_1 + \frac{2}{\alpha}C_2\Big)\mathbf{v}^T A\mathbf{v}.\end{aligned} \tag{2.11}$$

Note that (2.10) follows from (2.9) with $C_3 = C_2$.

We define coarse-spaces V_k and associated norms $\|\cdot\|_k$ by

$$\begin{aligned} V_k &= \operatorname{Rng}(P_k^0), \\ \|\cdot\|_k\colon\ P_k^0\mathbf{x} &\mapsto \|\mathbf{x}\| \equiv \sqrt{\mathbf{x}^T\mathbf{x}}, \quad k = 0,\dots,l\ (P_0^0 = I).\end{aligned} \tag{2.12}$$

Further, we define

$$\lambda_{k,j} = \sup_{\mathbf{x}\in\mathbb{R}^{n_k}\setminus\{\mathbf{0}\}} \frac{\langle AP_k^0\mathbf{x}, P_k^0\mathbf{x}\rangle}{\|P_k^0\mathbf{x}\|_j^2}, \quad k = 0,\dots,l,\ 0 \leqslant j \leqslant k. \tag{2.13}$$

Note that $\lambda_{k,j} \leqslant \varrho(A_j)$ and $\lambda_{k,k} = \varrho(A_k)$.

Remark 2.1. Definition (2.13) allows the following interpretation: The spectral bound

$$\varrho(A_k) = \sup_{\mathbf{x}\in\mathbb{R}^{n_k}\setminus\{\mathbf{0}\}} \frac{\langle A_k\mathbf{x},\mathbf{x}\rangle}{\|\mathbf{x}\|^2} = \sup_{\mathbf{x}\in\mathbb{R}^{n_k}\setminus\{\mathbf{0}\}} \frac{\langle AP_k^0\mathbf{x}, P_k^0\mathbf{x}\rangle}{\|P_k^0\mathbf{x}\|_k^2}$$

indicates the smoothness of the space V_k with respect to the norm $\|\cdot\|_k$. The quantity

$$\lambda_{k,j} = \sup_{\mathbf{x}\in\mathbb{R}^{n_k}\setminus\{\mathbf{0}\}} \frac{\langle AP_k^0\mathbf{x}, P_k^0\mathbf{x}\rangle}{\|P_k^0\mathbf{x}\|_j^2} = \sup_{\mathbf{x}\in\mathbb{R}^{n_k}\setminus\{\mathbf{0}\}} \frac{\langle AP_k^0\mathbf{x}, P_k^0\mathbf{x}\rangle}{\|P_j^0P_k^j\mathbf{x}\|_j^2} = \sup_{\mathbf{x}\in\mathbb{R}^{n_k}\setminus\{\mathbf{0}\}} \frac{\langle AP_k^0\mathbf{x}, P_k^0\mathbf{x}\rangle}{\|P_k^j\mathbf{x}\|^2},$$

$j < k$, indicates the smoothness of the space V_k with respect to the finer space norm $\|\cdot\|_j$.

We now formulate our abstract convergence estimate,

Theorem 2.1. *Let $\bar{\lambda}_{k+1,k} \geqslant \lambda_{k+1,k}$, $k = 0, \ldots, l-1$ be upper bounds. We assume the existence of linear mappings $Q_k\colon V_0 \to V_k$, $Q_0 = I$, satisfying*

$$\|(Q_k - Q_{k+1})\mathbf{v}\|_k \leqslant \frac{C_a}{\sqrt{\bar{\lambda}_{k+1,k}}}\|\mathbf{v}\|_A \quad \forall \mathbf{v} \in V_0,\ k = 0, \ldots, l-1, \tag{2.14}$$

and

$$\|Q_k\|_A \leqslant C_s, \quad k = 0, \ldots, l. \tag{2.15}$$

Further, we assume that our smoothers, M_k, satisfy (2.2) *and the symmetrized smoothers $\overline{M}_k$ satisfy*

$$\|\mathbf{v}\|^2_{\overline{M}_k} \leqslant \beta(\bar{\lambda}_{k+1,k}\|\mathbf{v}\|^2 + \|\mathbf{v}\|^2_{A_k}) \quad \forall \mathbf{v} \in \mathbb{R}^{n_k},\ k = 0, \ldots, l-1. \tag{2.16}$$

Then the resulting multigrid operator B is nearly spectrally equivalent to A, more precisely,

$$\mathbf{v}^T A\mathbf{v} \leqslant \mathbf{v}^T B\mathbf{v} \leqslant \Big[C_s^2 + 2l\Big(\beta(C_a^2 + 4C_s^2) + \frac{1}{\alpha}C_s^2\Big)\Big]\mathbf{v}^T A\mathbf{v} \quad \forall \mathbf{v} \in \mathbb{R}^{n_0}. \tag{2.17}$$

Remark 2.2. The difference from the results previously obtained based on the theory in [2] is in our use of the weak approximation condition (2.14). The original theory relied instead on the condition

$$\|(Q_k - Q_{k+1})\mathbf{v}\|_k \leqslant \frac{C_a}{\sqrt{\varrho(A_k)}}\|\mathbf{v}\|_A, \tag{2.18}$$

and the approximation properties of the space V_{k+1} were thus measured against the smoothness of the space V_k (because of $\varrho(A_k)$). In typical applications, the approximation on the left-hand side of (2.18) is guided by h_{k+1}, while the spectral bound of A_k and the scaling of the $\|\cdot\|_k$-norm are guided by h_k. To prove (2.18), the ratio h_{k+1}/h_k has to be bounded, and the resolutions of spaces V_k and V_{k+1} have to be comparable.

In our case, the approximation properties of the space V_{k+1} are measured against (the upper bound of) $\lambda_{k+1,k} \equiv \sup_{\mathbf{x}\in \mathrm{Rng}(P^k_{k+1})\setminus\{\mathbf{0}\}} \langle A_k\mathbf{x}, \mathbf{x}\rangle/\|\mathbf{x}\|^2 \leqslant \varrho(A_k)$, that is, against the smoothness of the space V_{k+1} (measured with respect to the norm $\|\cdot\|_k$ used on the left-hand side of (2.14)), and therefore the resolutions of the spaces V_k and V_{k+1} do not have to be comparable. The current estimate thus allows us to prove a convergence result independent of the coarsening ratio. The cost of the uniform convergence result, when the coarsening ratio becomes large ($\lambda_{k+1,k} \ll \varrho(A_k)$), is in the increasing demand on the smoother that arises through the smoothing condition (2.16).

P r o o f. We write linear mappings $Q_k\colon V_0 \to V_k \equiv \operatorname{Rng}(P_k^0)$, $k = 0, \ldots, l$, in the form

$$Q_k = P_k^0 \widetilde{Q}_k, \quad \widetilde{Q}_k\colon V_0 = \mathbb{R}^{n_0} \to \mathbb{R}^{n_k}.$$

In the XZ-identity (2.6), we choose

$$\mathbf{v}_k = \widetilde{Q}_k \mathbf{v}, \quad k = 1, \ldots, l.$$

Therefore (see (2.6)),

$$\mathbf{v}_k^f = \mathbf{v}_k - P_{k+1}^k \mathbf{v}_{k+1} = (\widetilde{Q}_k - P_{k+1}^k \widetilde{Q}_{k+1})\mathbf{v}.$$

Thus, to prove our theorem means to verify the inequalities (2.8), (2.9), and (2.10) for the above particular decomposition of $\mathbf{v}$.

To prove (2.8), we estimate using the assumptions (2.16), (2.14) and (2.15), the definition (2.12) of $\|\cdot\|_k$, $Q_k = P_k^0 \widetilde{Q}_k$ for $k = 0, \ldots, l$ and the triangle inequality:

$$\begin{aligned}
\text{(2.19)} \qquad \|\mathbf{v}_k^f\|_{M_k}^2 &= \|(\widetilde{Q}_k - P_{k+1}^k \widetilde{Q}_{k+1})\mathbf{v}\|_{M_k}^2 \\
&\leqslant \beta(\bar{\lambda}_{k+1,k} \|(\widetilde{Q}_k - P_{k+1}^k \widetilde{Q}_{k+1})\mathbf{v}\|^2 + \|(\widetilde{Q}_k - P_{k+1}^k \widetilde{Q}_{k+1})\mathbf{v}\|_{A_k}^2) \\
&= \beta(\bar{\lambda}_{k+1,k} \|P_k^0(\widetilde{Q}_k - P_{k+1}^k \widetilde{Q}_{k+1})\mathbf{v}\|_k^2 + \|P_k^0(\widetilde{Q}_k - P_{k+1}^k \widetilde{Q}_{k+1})\mathbf{v}\|_A^2) \\
&= \beta(\bar{\lambda}_{k+1,k} \|(Q_k - Q_{k+1})\mathbf{v}\|_k^2 + \|(Q_k - Q_{k+1})\mathbf{v}\|_A^2) \\
&\leqslant \beta(C_a^2 \|\mathbf{v}\|_A^2 + 2(\|Q_k \mathbf{v}\|_A^2 + \|Q_{k+1}\mathbf{v}\|_A^2)) \\
&\leqslant \beta(C_a^2 + 4C_s^2)\|\mathbf{v}\|_A^2.
\end{aligned}$$

Thus,

$$\sum_{k=0}^{l-1} \|\mathbf{v}_k^f\|_{M_k}^2 \leqslant l\beta(C_a^2 + 4C_s^2)\|\mathbf{v}\|_A^2,$$

proving (2.8) with a constant

$$C_1 = l\beta(C_a^2 + 4C_s^2).$$

To prove (2.9) and (2.10), we realize that

$$\|\mathbf{v}_k\|_{A_k}^2 = \|\widetilde{Q}_k \mathbf{v}\|_{A_k}^2 = \|Q_k \mathbf{v}\|_A^2 \leqslant C_s^2 \|\mathbf{v}\|_A^2,$$

hence (2.9) immediately follows with a constant

$$C_2 = C_s^2 l$$

and (2.10) with a constant

$$C_3 = C_s^2.$$

The estimate (2.17) now follows by (2.11). □

3. Polynomial smoother

In this section, we investigate a polynomial smoother with the error propagation operator

$$I - M_k^{-T} A_k = I - M_k^{-1} A_k = S_k^{\gamma}\Big(I - \frac{1}{\bar\lambda_{S_k^2 A_k}} S_k^2 A_k\Big), \tag{3.1}$$

where S_k is a polynomial in A_k such that $\varrho(S_k) \leqslant 1$, $\bar\lambda_{S_k^2 A_k} \geqslant \varrho(S_k^2 A_k)$ and γ is a positive integer. The particular cases of interest are $\gamma = 1$ and $\gamma = 2$.

From (2.4) and the fact that the error propagation operator corresponding to the symmetrized smoother $\overline{M}_k$ is

$$I - \overline{M}_k^{-1} A_k = (I - M_k^{-T} A_k)(I - M_k^{-1} A_k) = S_k^{2\gamma}\Big(I - \frac{1}{\bar\lambda_{S_k^2 A_k}} S_k^2 A_k\Big)^2,$$

it follows that the corresponding symmetrized smoother $\overline{M}_k$ is given by

$$\overline{M}_k^{-1} = A_k^{-1}\Big[I - \Big(I - \frac{1}{\bar\lambda_{S_k^2 A_k}} S_k^2 A_k\Big)^2 S_k^{2\gamma}\Big]. \tag{3.2}$$

Lemma 3.1. *We assume that S_k is a polynomial in A_k such that $\varrho(S_k) \leqslant 1$, $\bar\lambda_{S_k^2 A_k} \geqslant \varrho(S_k^2 A_k)$, and γ is a positive integer. Let $\{v_i\}$ be the eigenvectors of A_k and $\lambda_i(S_k)$ the corresponding eigenvalues of S_k. For a given parameter, $q \in (0,1)$, define*

$$U_1 = \{\operatorname{span}\{v_i\}\colon\ |\lambda_i(S_k)| \leqslant q\} \quad \text{and} \quad U_2 = \{\operatorname{span}\{v_i\}\colon\ |\lambda_i(S_k)| > q\}.$$

Then the symmetrized smoother $\overline{M}_k$ in (3.2) is positive definite and satisfies

$$\|\mathbf{x}\|^2_{\overline{M}_k} \leqslant \frac{1}{1-q^{2\gamma}} \|\mathbf{x}\|^2_{A_k}\ \forall \mathbf{x} \in U_1, \quad \|\mathbf{x}\|^2_{\overline{M}_k} \leqslant \frac{\bar\lambda_{S_k^2 A_k}}{q^2} \|\mathbf{x}\|^2\ \forall \mathbf{x} \in U_2,$$

and

$$\|\mathbf{x}\|^2_{\overline{M}_k} \leqslant \frac{1}{1-q^{2\gamma}} \|\mathbf{x}\|^2_{A_k} + \frac{\bar\lambda_{S_k^2 A_k}}{q^2} \|\mathbf{x}\|^2 \quad \forall \mathbf{x} \in \mathbb{R}^{n_k},\ q \in (0,1). \tag{3.3}$$

Remark 3.1. Our goal is to satisfy the smoothing condition (2.16). Therefore, in view of (3.3), the property the smoother (3.1) needs to satisfy is

$$\bar\lambda_{S_k^2 A_k} \leqslant C \bar\lambda_{k+1,k} \tag{3.4}$$

(with $\bar{\lambda}_{k+1,k} \ll \varrho(A_k)$ for aggressive coarsening). Indeed, from (3.4) and (3.3), it follows that

$$\begin{aligned}\|\mathbf{x}\|^2_{\overline{M}_k} &\leqslant \max\Big\{\frac{1}{1-q^{2\gamma}}, \frac{1}{q^2}\Big\}(\|\mathbf{x}\|^2_{A_k} + \bar{\lambda}_{S_k^2 A_k}\|\mathbf{x}\|^2)\\ &\leqslant \max\Big\{\frac{1}{1-q^{2\gamma}}, \frac{1}{q^2}\Big\}\cdot\max\{1, C\}\ (\|\mathbf{x}\|^2_{A_k} + \bar{\lambda}_{k+1,k}\|\mathbf{x}\|^2) \quad \forall \mathbf{x} \in \mathbb{R}^{n_k}.\end{aligned}$$

Here $q \in (0,1)$ is a parameter we choose. Thus, (2.16) follows from (3.4) and (3.3) with

$$\beta = \min_{q\in(0,1)} \max\Big\{\frac{1}{1-q^{2\gamma}}, \frac{1}{q^2}\Big\}\cdot\max\{1, C\}. \tag{3.5}$$

The role of the smoothing polynomial, $S_k = p(A_k)$, is therefore to minimize $\varrho(S_k^2 A_k)$ (to attain the same order of magnitude as $\bar{\lambda}_{k+1,k}$), subject to the constraint that S_k is an error propagation operator of an A_k-non-divergent smoother, that is, $p(0) = 1$ and $\varrho(S_k) \leqslant 1$. Let $\bar{\lambda}_k \geqslant \varrho(A_k)$ be an available upper bound. The polynomial p of a given degree, N_k, satisfying the above constraints and minimizing the right-hand side of the inequality

$$\varrho(S_k^2 A_k) = \varrho(p^2(A_k)A_k) = \max_{t\in\sigma(A_k)} p^2(t)t \leqslant \max_{t\in[0,\bar{\lambda}_k]} p^2(t)t$$

will be given in Lemma 3.2.

Remark 3.2. For $\gamma = 1$, using the minimizer $\hat{q} = 1/\sqrt{2}$, we have (see (3.5))

$$\min_{q\in(0,1)} \max\Big\{\frac{1}{1-q^{2\gamma}}, \frac{1}{q^2}\Big\} = 2.$$

Similarly, for $\gamma = 2$, using the minimizer $\hat{q} = \sqrt{\frac{1}{2}(-1+\sqrt{5})}$, we get

$$\min_{q\in(0,1)} \max\Big\{\frac{1}{1-q^{2\gamma}}, \frac{1}{q^2}\Big\} = \frac{2}{-1+\sqrt{5}} \doteq 1.618034.$$

Proof. The proof given here is a generalization of the one given in [4].

Recall that both S_k and $I - \bar{\lambda}^{-1}_{S_k^2 A_k} S_k^2 A_k$ are polynomials in A_k, hence all these matrices have common eigenvectors, mutually commute and U_1 and U_2 are their common invariant subspaces. Further, $\varrho(S_k) \leqslant 1$ and $\varrho(I - \bar{\lambda}^{-1}_{S_k^2 A_k} S_k^2 A_k) \leqslant 1$.

To prove

$$\langle \overline{M}_k \mathbf{x}, \mathbf{x}\rangle \leqslant \frac{1}{1-q^{2\gamma}}\langle A_k \mathbf{x}, \mathbf{x}\rangle \text{ on } U_1, \tag{3.6}$$

($\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product in $\mathbb{R}^{n_k}$), we use (3.2) and estimate for $\mathbf{x} \in U_1$:

$$\begin{aligned}\langle \overline{M}_k^{-1}\mathbf{x},\mathbf{x}\rangle &= \langle A_k^{-1}\mathbf{x},\mathbf{x}\rangle - \Big\langle A_k^{-1}\Big(I - \frac{1}{\bar{\lambda}_{S_k^2A_k}}S_k^2A_k\Big)^2 S_k^{2\gamma}\mathbf{x},\mathbf{x}\Big\rangle \\ &\geqslant \langle A_k^{-1}\mathbf{x},\mathbf{x}\rangle - \langle A_k^{-1}S_k^{2\gamma}\mathbf{x},\mathbf{x}\rangle \\ &\geqslant \langle A_k^{-1}\mathbf{x},\mathbf{x}\rangle - q^{2\gamma}\langle A_k^{-1}\mathbf{x},\mathbf{x}\rangle = (1-q^{2\gamma})\langle A_k^{-1}\mathbf{x},\mathbf{x}\rangle.\end{aligned}$$

Since U_1 is an invariant subspace of both $\overline{M}_k$ and A_k, both $\overline{M}_k^{-1}$ and $\overline{M}_k$ are symmetric, positive definite on U_1 and the statement (3.6) follows.

To prove

$$\langle \overline{M}_k\mathbf{x},\mathbf{x}\rangle \leqslant \frac{\bar{\lambda}_{S_k^2A_k}}{q^2}\|\mathbf{x}\|^2 \quad \text{on } U_2, \tag{3.7}$$

we estimate for $\mathbf{x} \in U_2$:

$$\begin{aligned}\langle \overline{M}_k^{-1}\mathbf{x},\mathbf{x}\rangle &= \langle A_k^{-1}\mathbf{x},\mathbf{x}\rangle - \Big\langle A_k^{-1}\Big(I - \frac{1}{\bar{\lambda}_{S_k^2A_k}}S_k^2A_k\Big)^2 S_k^{2\gamma}\mathbf{x},\mathbf{x}\Big\rangle \\ &\geqslant \langle A_k^{-1}\mathbf{x},\mathbf{x}\rangle - \Big\langle A_k^{-1}\Big(I - \frac{1}{\bar{\lambda}_{S_k^2A_k}}S_k^2A_k\Big)\mathbf{x},\mathbf{x}\Big\rangle \\ &= \frac{1}{\bar{\lambda}_{S_k^2A_k}}\langle S_k^2\mathbf{x},\mathbf{x}\rangle \\ &> \frac{q^2}{\bar{\lambda}_{S_k^2A_k}}\|\mathbf{x}\|^2.\end{aligned}$$

Since U_2 is an invariant subspace of $\overline{M}_k$, both $\overline{M}_k^{-1}$ and $\overline{M}_k$ are symmetric, positive definite on U_2 and the statement (3.7) follows.

Let us consider the decomposition of $\mathbf{x} \in \mathbb{R}^{n_k} \setminus \{\mathbf{0}\}$,

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2, \quad \mathbf{x}_1 \in U_1,\ \mathbf{x}_2 \in U_2.$$

From the definition of the spaces U_1, U_2 it follows that the spaces U_1 and U_2 are orthogonal, that is,

$$\langle \mathbf{x}_1,\mathbf{x}_2\rangle = 0.$$

Since U_1 and U_2 are invariant subspaces of both A_k and $\overline{M}_k$, it also follows that

$$\langle A_k\mathbf{x}_1,\mathbf{x}_2\rangle = \langle \overline{M}_k\mathbf{x}_1,\mathbf{x}_2\rangle = 0.$$

Therefore, since $\overline{M}_k$ is symmetric, positive definite on both U_1 and U_2, it follows that

$$\langle \overline{M}_k \mathbf{x}, \mathbf{x}\rangle = \langle \overline{M}_k \mathbf{x}_1, \mathbf{x}_1\rangle + \langle \overline{M}_k \mathbf{x}_2, \mathbf{x}_2\rangle > 0,$$

hence $\overline{M}_k$ is symmetric, positive definite on $\mathbb{R}^{n_k}$. Thus, the spaces U_1 and U_2 form a decomposition of $\mathbb{R}^{n_k}$ that is orthogonal with respect to the norms $\|\cdot\|$, $\|\cdot\|_{A_k}$ and $\|\cdot\|_{\overline{M}_k}$. Then (3.6) and (3.7) give

$$\begin{aligned}\|\mathbf{x}\|^2_{\overline{M}_k} &= \|\mathbf{x}_1\|^2_{\overline{M}_k} + \|\mathbf{x}_2\|^2_{\overline{M}_k}\\ &\leqslant \frac{1}{1-q^{2\gamma}}\|\mathbf{x}_1\|^2_{A_k} + \frac{\bar{\lambda}_{S_k^2 A_k}}{q^2}\|\mathbf{x}_2\|^2 \leqslant \frac{1}{1-q^{2\gamma}}\|\mathbf{x}\|^2_{A_k} + \frac{\bar{\lambda}_{S_k^2 A_k}}{q^2}\|\mathbf{x}\|^2,\end{aligned}$$

proving (3.3). □

While the validity of property (2.16) is addressed by Lemma 3.1, we still need to verify that inequality (2.2) is satisfied for our choice of the smoother. To this end, the smoother S_k is introduced in the next lemma.

Lemma 3.2. *For any $\lambda > 0$ and integer $N > 0$ there is a unique polynomial $p_{\lambda,N}$ of degree N such that*

$$\max_{0\leqslant t\leqslant\lambda} p^2_{\lambda,N}(t)t$$

is minimal under the constraint $p_{\lambda,N}(0) = 1$. The polynomial p is given by

$$p_{\lambda,N}(t) = \Big(1-\frac{t}{r_1}\Big)\dots\Big(1-\frac{t}{r_N}\Big), \quad r_k = \frac{\lambda}{2}\Big(1-\cos\Big(\frac{2k\pi}{2N+1}\Big)\Big), \tag{3.8}$$

$k = 1,\dots,N$. The polynomial $p_{\lambda,N}$ satisfies

$$\max_{0\leqslant t\leqslant\lambda} p^2_{\lambda,N}(t)t = \frac{\lambda}{(2N+1)^2} \tag{3.9}$$

and

$$\max_{0\leqslant t\leqslant\lambda} |p_{\lambda,N}(t)| = 1. \tag{3.10}$$

The polynomial $p_{\lambda,N}$ is the transformed Chebyshev polynomial

$$p_{\lambda,N}(t) = (-1)^N \frac{1}{2N+1}\frac{\sqrt{\lambda}}{\sqrt{t}}T_{2N+1}\Big(\frac{\sqrt{t}}{\sqrt{\lambda}}\Big),$$

where T_k is a Chebyshev polynomial of degree k, that is $T_0(t) = 1$, $T_1(t) = t$, and $T_{k+1}(t) = 2tT_k(t) - T_{k-1}(t)$ for $k \geqslant 1$.

P r o o f. Proof of the lemma in this form can be found in [3]. The analysis of Chebyshev polynomials can be found in [1], see also [10]. □

Let $\bar{\lambda}_k$ be an available upper bound of $\varrho(A_k)$ and let the integer N_k be a given degree of the smoothing polynomial. We choose

$$S_k = p_{\bar{\lambda}_k, N_k}(A_k), \tag{3.11}$$

where $p_{\lambda,N}$ is given by (3.8). Further, we set

$$\bar{\lambda}_{S_k^2 A_k} = \frac{\bar{\lambda}_k}{(2N_k+1)^2}.$$

Then, by Lemma 3.2 and the spectral mapping theorem, we have

$$\begin{aligned} \varrho(S_k^2 A_k) &= \max_{t \in \sigma(A_k)} p^2_{\bar{\lambda}_k, N_k}(t)t \leqslant \max_{t \in [0, \bar{\lambda}_k]} p^2_{\bar{\lambda}_k, N_k}(t)t \\ &= \bar{\lambda}_{S_k^2 A_k} \equiv \frac{\bar{\lambda}_k}{(2\deg(S_k)+1)^2}, \quad \varrho(S_k) \leqslant 1. \end{aligned} \tag{3.12}$$

Lemma 3.3. *For the smoother* (3.1), *with* $\gamma = 1$ *and* S_k *given by* (3.11) *and* (3.8), *the inequality* (2.2) *holds with*

$$\alpha = \frac{\delta_0}{2-\delta_0}, \quad \delta_0 = 1 - \frac{2}{3\sqrt{3}} \in (0,1).$$

Further, for $\gamma > 0$ *that is even, and* S_k *being a polynomial in* A_k *satisfying* $\varrho(S_k) \leqslant 1$, *the inequality* (2.2) *holds with* $\alpha = 1$. *(That is, for even* $\gamma > 0$, *we do not have to assume that* S_k *is given by* (3.11) *and* (3.8), *we only need* S_k *to be a polynomial in* A_k *such that* $\varrho(S_k) \leqslant 1$.*)*

Proof. For the proof in the case of $\gamma = 1$ and S_k given by (3.11) and (3.8), see [4], Lemma 6.2 and Proposition 7.3.

For an even γ and S_k being a polynomial in A_k satisfying $\varrho(S_k) \leqslant 1$, we have

$$\langle M_k^{-1}\mathbf{x}, \mathbf{x}\rangle = \langle A_k^{-1}\mathbf{x}, \mathbf{x}\rangle - \Big\langle A_k^{-1}\Big(I - \frac{1}{\bar{\lambda}_{S_k^2 A_k}} S_k^2 A_k\Big) S_k^{\gamma}\mathbf{x}, \mathbf{x}\Big\rangle \leqslant \langle A_k^{-1}\mathbf{x}, \mathbf{x}\rangle.$$

Hence, $M_k \geqslant A_k$, and therefore assumption (2.2) on the smoother M_k holds trivially with $\alpha = 1$. □

Remark 3.3. The most natural way to implement the action of (3.1) for a given vector $\mathbf{x}$ is the following: To perform the iteration with the linear part S_k given by (3.11) and (3.8), we do for $i = 1, \ldots, N_k = \deg(S_k)$,

$$\mathbf{x} \leftarrow (I - \alpha_i A_k)\mathbf{x} + \alpha_i \mathbf{f}, \quad \alpha_i = \Big(\frac{\bar{\lambda}_k}{2}\Big(1 - \cos\Big(\frac{2i\pi}{2N_k+1}\Big)\Big)\Big)^{-1}, \quad \bar{\lambda}_k \geqslant \varrho(A_k).$$

To perform the iteration with the error propagation operator $I - \bar{\lambda}_{S_k^2 A_k}^{-1} S_k^2 A_k$, we do

$$\mathbf{x} \leftarrow \mathbf{x} - \frac{1}{\bar{\lambda}_{S_k^2 A_k}} S_k^2 (A_k \mathbf{x} - \mathbf{f}),$$

where the action of S_k is evaluated as the product

$$S_k \mathbf{x} = (I - \alpha_1 A_k) \dots (I - \alpha_{N_k} A_k) \mathbf{x}.$$

4. The final abstract result

In this section, we summarize the results proved in Sections 2 and 3 in the form of a theorem.

Theorem 4.1. *Let* $\bar{\lambda}_{k+1,k} \geqslant \lambda_{k+1,k}$ $(k = 0, \dots, l-1)$ *and* $\bar{\lambda}_k \geqslant \varrho(A_k)$ $(k = 0, \dots, l)$ *be upper bounds. We assume the existence of linear mappings (see* (2.12)*)* $Q_k \colon V_0 \to V_k$, $k = 0, \dots, l$, $Q_0 = I$, *satisfying* (2.14) *and* (2.15) *with positive constants* C_a *and* C_s*, independent of the level. Further, we assume that the linear part of both the pre- and post-smoother is given by* (3.1) *with* $S_k = p_{\bar{\lambda}_k, N_k}(A_k)$*, where the polynomial* $p_{\lambda,N}$ *is given by* (3.8) *and its degree,* N_k*, satisfies*

$$N_k \geqslant C_{\rm deg} \sqrt{\frac{\bar{\lambda}_k}{\bar{\lambda}_{k+1,k}}}, \quad k = 0, \dots, l-1, \tag{4.1}$$

with a constant $C_{\rm deg} > 0$ *independent of the level. We assume that* γ *in* (3.1) *is either even, or* $\gamma = 1$*. Then* (2.17) *is satisfied; that is,*

$$\mathbf{v}^T A \mathbf{v} \leqslant \mathbf{v}^T B \mathbf{v} \leqslant \Big[C_s^2 + 2l\Big(\beta(C_a^2 + 4C_s^2) + \frac{1}{\alpha} C_s^2\Big)\Big] \mathbf{v}^T A \mathbf{v} \quad \forall \mathbf{v} \in \mathbb{R}^{n_0}$$

holds with

$$\alpha = \begin{cases} 1 & \text{for } \gamma \text{ even}, \\ \dfrac{\delta_0}{2 - \delta_0}, \ \delta_0 = 1 - \dfrac{2}{3\sqrt{3}} & \text{for } \gamma = 1 \end{cases} \tag{4.2}$$

and

$$\beta = \min_{q \in (0,1)} \max \Big\{ \frac{1}{1 - q^{2\gamma}}, \frac{1}{q^2} \Big\} \cdot \max \Big\{ 1, \frac{1}{4C_{\rm deg}^2} \Big\}. \tag{4.3}$$

P r o o f. Statement (2.17) follows from Theorem 2.1 under assumptions (2.2) and (2.16) (inequalities (2.14) and (2.15) are assumptions of this theorem).

Assumption (2.2), with α given by (4.2), has been verified by Lemma 3.3.

According to Remark 3.1, (2.16) holds under assumption (3.4). Using Lemma 3.2, we estimate

$$\varrho(S_k^2 A_k) = \max_{t\in\sigma(A_k)} p^2_{\bar\lambda_k,N_k}(t)t \leqslant \bar\lambda_{S_k^2 A_k} \equiv \max_{t\in[0,\bar\lambda_k]} p^2_{\bar\lambda_k,N_k}(t)t = \frac{\bar\lambda_k}{(2\deg(S_k)+1)^2}.$$

Based on assumption (4.1), we further estimate

$$\bar\lambda_{S_k^2 A_k} \equiv \frac{\bar\lambda_k}{(2\deg(S_k)+1)^2} \leqslant \frac{\bar\lambda_k}{4\deg^2(S_k)} \leqslant \frac{\bar\lambda_k}{4C^2_{\deg}\frac{\bar\lambda_k}{\bar\lambda_{k+1,k}}} \leqslant \frac{1}{4C^2_{\deg}}\bar\lambda_{k+1,k},$$

thus proving (3.4) with the constant $C = 1/(4C^2_{\deg})$. Hence, inequality (2.16), with β given by (4.3), follows by Remark 3.1. Estimate (2.17), with α given by (4.2) and β given by (4.3), now follows by Theorem 2.1. □

5. Model example

We consider a model elliptic problem with H^1_0-equivalent form on a bounded polygonal or polyhedral domain $\Omega\subset\mathbb{R}^d$, $d=2$ or $d=3$, that is,

$$\text{find } u\in H^1_0(\Omega) \text{ such that } a(u,v) = (f,v)_{L_2(\Omega)}\ \forall v\in H^1_0(\Omega), \tag{5.1}$$

where $f\in L_2(\Omega)$ and

$$c|u|^2_{H^1(\Omega)} \leqslant a(u,u) \leqslant C|u|^2_{H^1(\Omega)} \quad \forall u\in H^1_0(\Omega). \tag{5.2}$$

Further, we consider a system of nested quasiuniform triangulations $\{\tau_{h_k}\}_{k=0}^{l}$ of Ω (τ_{h_k} being the refinement of $\tau_{h_{k+1}}$) and the corresponding piecewise linear ($P1$) finite element spaces

$$H^1_0(\Omega) \supset V_{h_0} \supset V_{h_1} \supset \ldots \supset V_{h_l}.$$

Here, h_k denotes a characteristic meshsize on level k. Note that the case of interest is $h_k \ll h_{k+1}$. We denote the standard $P1$ finite element basis of V_{h_k} by $\{\varphi_i^k,\ i=1,\ldots,n_k\}$, and define standard finite element interpolators in the usual way:

$$\Pi_{h_k}\colon\ \mathbf{x}\in\mathbb{R}^{n_k} \mapsto \sum_{i=1}^{n_k} x_i\varphi_i^k, \quad k=0,\ldots,l.$$

We assume that the matrix $A = A_0$ was obtained by the standard finite element discretization of (5.1) using the finite element basis $\{\varphi_i^0\}_{i=1}^{n_0}$, that is,

$$a(\Pi_{h_0}\mathbf{x}, \Pi_{h_0}\mathbf{x}) = \langle A_0\mathbf{x}, \mathbf{x}\rangle, \quad \mathbf{x} \in \mathbb{R}^{n_0}.$$

The multigrid prolongators are given by

$$P_{k+1}^k = \Pi_{h_k}^{-1}\Pi_{h_{k+1}}. \tag{5.3}$$

Note that P_{k+1}^k is an $n_k \times n_{k+1}$ matrix whose j-th column is the basis function φ_j^{k+1} represented in terms of the basis $\{\varphi_i^k\}$ of the immediately finer level. The coarse-level matrices are defined by (2.1), that is,

$$A_k = (P_k^{k-1})^T A_{k-1} P_k^{k-1} = (P_k^0)^T A P_k^0,$$
$$\langle A_k\mathbf{x}, \mathbf{x}\rangle = a(\Pi_{h_k}\mathbf{x}, \Pi_{k_k}\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^{n_k},\ k = 1, \dots, l.$$

Let $Q_{h_k}\colon H_0^1(\Omega) \to V_{h_k}$ be an $L_2(\Omega)$-orthogonal projection. We define $\widetilde{Q}_k\colon \mathbb{R}^{n_0} \to \mathbb{R}^{n_k}$ by

$$\Pi_{h_k}\widetilde{Q}_k = Q_{h_k}\Pi_{h_0}, \quad k = 0, \dots, l.$$

For $k = 0, \dots, l$, we set

$$Q_k = P_k^0\widetilde{Q}_k.$$

We will verify the assumptions of Theorem 4.1 for the above linear mappings Q_k. Namely, we need to verify assumptions (2.14) and (2.15) for our linear mappings Q_k and satisfy the assumption (4.1) for smoothers M_k whose error propagation operator is given by the polynomial (3.1), where S_k is chosen as in (3.11) and (3.8). We will show that our method converges uniformly with respect to the coarsening ratio if the polynomial $S_k = p_{\bar{\lambda}_k, N_k}(A_k)$ in (3.11) has a degree

$$N_k = \deg(S_k) \geqslant C\frac{h_{k+1}}{h_k}, \quad C > 0.$$

Note that the assumption $ch_{k+1}/h_k \leqslant \deg(S_k) \leqslant Ch_{k+1}/h_k$ is equivalent to

$$c\,\frac{h_{k+1}}{h_k} \leqslant \deg(I - M_k^{-1}A_k) = (2+\gamma)\deg(S_k) + 1 \leqslant C\frac{h_{k+1}}{h_k} \tag{5.4}$$

(with different constants $c, C > 0$). Again, we recall that the cases of practical interest are $\gamma = 1$ and $\gamma = 2$. In any case, we consider γ bounded. Thus, in what follows, we assume (5.4).

We will use the following well-known properties of the finite element functions ([5]):

$$\|(I - Q_{h_k})u\|_{L_2(\Omega)} \leqslant Ch_k|u|_{H^1(\Omega)} \quad \forall u \in H_0^1(\Omega), \tag{5.5}$$

$$|Q_{h_k}u|_{H^1(\Omega)} \leqslant C|u|_{H^1(\Omega)} \quad \forall u \in H_0^1(\Omega), \tag{5.6}$$

$$c\|\Pi_{h_k}\mathbf{x}\|^2_{L_2(\Omega)} \leqslant h_k^d\|\mathbf{x}\| \leqslant C\|\Pi_{h_k}\mathbf{x}\|^2_{L_2(\Omega)} \quad \forall \mathbf{x} \in \mathbb{R}^{n_k}, \tag{5.7}$$

$$\varrho(A_k) \leqslant C \max_{i=1,\dots,n_k} |\varphi_k^i|^2_{H^1(\Omega)} \leqslant Ch_k^{d-2}. \tag{5.8}$$

In the estimates to follow, C, c denote generic constants that will depend on the constants in (5.2), (5.5), (5.6), (5.7), (5.8), and (5.4).

First we estimate the value of $\lambda_{k+1,k}$ in (2.14):

$$\begin{aligned} \lambda_{k+1,k} &= \sup_{\mathbf{x}\in\mathbb{R}^{n_{k+1}}\setminus\{\mathbf{0}\}} \frac{\langle A_{k+1}\mathbf{x},\mathbf{x}\rangle}{\|P^0_{k+1}\mathbf{x}\|^2_k} \\ &= \sup_{\mathbf{x}\in\mathbb{R}^{n_{k+1}}\setminus\{\mathbf{0}\}} \Big(\frac{\langle A_{k+1}\mathbf{x},\mathbf{x}\rangle}{\|P^0_{k+1}\mathbf{x}\|^2_{k+1}} \cdot \frac{\|P^0_{k+1}\mathbf{x}\|^2_{k+1}}{\|P^0_{k+1}\mathbf{x}\|^2_k}\Big) \\ &\leqslant \sup_{\mathbf{x}\in\mathbb{R}^{n_{k+1}}\setminus\{\mathbf{0}\}} \frac{\langle A_{k+1}\mathbf{x},\mathbf{x}\rangle}{\|P^0_{k+1}\mathbf{x}\|^2_{k+1}} \cdot \sup_{\mathbf{x}\in\mathbb{R}^{n_{k+1}}\setminus\{\mathbf{0}\}} \frac{\|P^0_{k+1}\mathbf{x}\|^2_{k+1}}{\|P^0_{k+1}\mathbf{x}\|^2_k} \\ &= \sup_{\mathbf{x}\in\mathbb{R}^{n_{k+1}}\setminus\{\mathbf{0}\}} \frac{\langle A_{k+1}\mathbf{x},\mathbf{x}\rangle}{\|\mathbf{x}\|^2} \cdot \sup_{\mathbf{x}\in\mathbb{R}^{n_{k+1}}\setminus\{\mathbf{0}\}} \frac{\|\mathbf{x}\|^2}{\|P^0_k P^k_{k+1}\mathbf{x}\|^2_k} \\ &= \varrho(A_{k+1}) \sup_{\mathbf{x}\in\mathbb{R}^{n_{k+1}}\setminus\{\mathbf{0}\}} \frac{\|\mathbf{x}\|^2}{\|P^k_{k+1}\mathbf{x}\|^2}. \end{aligned} \tag{5.9}$$

Employing the equivalence (5.7) between the L_2-norm and the Euclidean norm, together with the definition $P^k_{k+1} = \Pi^{-1}_{h_k}\Pi_{h_{k+1}}$, we obtain

$$\|\Pi_{h_{k+1}}\mathbf{x}\|^2_{L_2(\Omega)} = \|\Pi_{h_k}P^k_{k+1}\mathbf{x}\|^2_{L_2(\Omega)} \approx h_k^d\|P^k_{k+1}\mathbf{x}\|^2.$$

From here and from (5.7), we have

$$\|P^k_{k+1}\mathbf{x}\|^2 \approx h_k^{-d}\|\Pi_{h_{k+1}}\mathbf{x}\|^2, \text{ and } \|\mathbf{x}\|^2 \approx h_{k+1}^{-d}\|\Pi_{h_{k+1}}\mathbf{x}\|^2.$$

The last two equivalences, together with (5.9) and (5.8), yield

$$\begin{aligned} \lambda_{k+1,k} &\leqslant Ch^{d-2}_{k+1} \sup_{\mathbf{x}\in\mathbb{R}^{n_{k+1}}\setminus\{\mathbf{0}\}} \frac{\|\mathbf{x}\|^2}{\|P^k_{k+1}\mathbf{x}\|^2} \\ &\leqslant Ch^{d-2}_{k+1} \sup_{\mathbf{x}\in\mathbb{R}^{n_{k+1}}\setminus\{\mathbf{0}\}} \frac{h^{-d}_{k+1}\|\Pi_{h_{k+1}}\mathbf{x}\|^2_{L_2(\Omega)}}{h_k^{-d}\|\Pi_{h_{k+1}}\mathbf{x}\|^2_{L_2(\Omega)}} \\ &\leqslant \bar{\lambda}_{k+1,k} \equiv C\frac{h_k^d}{h^2_{k+1}}. \end{aligned} \tag{5.10}$$

(We take the final estimate as an upper bound $\bar{\lambda}_{k+1,k} \geqslant \lambda_{k+1,k}$, see Theorem 2.1.) To verify (2.14), we further estimate using (5.5), (5.7), (2.12), $P^k_{k+1} = \Pi^{-1}_{h_k}\Pi_{h_{k+1}}$, $\Pi_{h_k}\widetilde{Q}_k = Q_{h_k}\Pi_{h_0}$, $Q_k = P^0_k\widetilde{Q}_k$, the fact that $Q_{h_k}\colon H^1(\Omega) \to V_{h_k}$ is an $L_2(\Omega)$-orthogonal projection and $V_{h_{k+1}} \subset V_{h_k}$:

$$\begin{aligned}
\|(Q_k - Q_{k+1})\mathbf{v}\|_k^2 &= \|(P^0_k\widetilde{Q}_k - P^0_{k+1}\widetilde{Q}_{k+1})\mathbf{v}\|_k^2 \\
&= \|P^0_k(\widetilde{Q}_k - P^k_{k+1}\widetilde{Q}_{k+1})\mathbf{v}\|_k^2 \\
&= \|(\widetilde{Q}_k - P^k_{k+1}\widetilde{Q}_{k+1})\mathbf{v}\|^2 \\
&\approx h_k^{-d}\|\Pi_{h_k}(\widetilde{Q}_k - P^k_{k+1}\widetilde{Q}_{k+1})\mathbf{v}\|^2_{L_2(\Omega)} \\
&= h_k^{-d}\|(\Pi_{h_k}\widetilde{Q}_k - \Pi_{h_{k+1}}\widetilde{Q}_{k+1})\mathbf{v}\|^2_{L_2(\Omega)} \\
&= h_k^{-d}\|(Q_{h_k} - Q_{h_{k+1}})\Pi_{h_0}\mathbf{v}\|^2_{L_2(\Omega)} \\
&\leqslant h_k^{-d}(\|(I - Q_{h_k})\Pi_{h_0}\mathbf{v}\|^2_{L_2(\Omega)} + \|(Q_{h_k} - Q_{h_{k+1}})\Pi_{h_0}\mathbf{v}\|^2_{L_2(\Omega)}) \\
&= h_k^{-d}\|(I - Q_{h_{k+1}})\Pi_{h_0}\mathbf{v}\|^2_{L_2(\Omega)} \\
&\leqslant C\frac{h^2_{k+1}}{h^d_k}|\Pi_{h_0}\mathbf{v}|^2_{H^1(\Omega)}.
\end{aligned}$$

Since

$$|\Pi_{h_0}\mathbf{v}|^2_{H^1(\Omega)} \approx a(\Pi_{h_0}\mathbf{v}, \Pi_{h_0}\mathbf{v}) = \|\mathbf{v}\|^2_A,$$

and from (5.10), i.e.

$$\lambda_{k+1,k} \leqslant \bar{\lambda}_{k+1,k} \equiv C\frac{h^d_k}{h^2_{k+1}},$$

we obtain

$$\|(Q_k - Q_{k+1})\mathbf{v}\|_k^2 \leqslant \frac{C}{\bar{\lambda}_{k+1,k}}\|\mathbf{v}\|^2_A,$$

proving (2.14).

To verify (2.15), we use (5.6), (5.2) and $\Pi_{h_k}\widetilde{Q}_k = Q_{h_k}\Pi_{h_0}$ and observe that

$$\begin{aligned}
\|Q_k\mathbf{v}\|^2_A = \|\widetilde{Q}_k\mathbf{v}\|^2_{A_k} &= a(\Pi_{h_k}\widetilde{Q}_k\mathbf{v}, \Pi_{h_k}\widetilde{Q}_k\mathbf{v}) = a(Q_{h_k}\Pi_{h_0}\mathbf{v}, Q_{h_k}\Pi_{h_0}\mathbf{v}) \\
&\leqslant C|Q_{h_k}\Pi_{h_0}\mathbf{v}|^2_{H^1(\Omega)} \leqslant C|\Pi_{h_0}\mathbf{v}|^2_{H^1(\Omega)} \leqslant Ca(\Pi_{h_0}\mathbf{v}, \Pi_{h_0}\mathbf{v}) = C\|\mathbf{v}\|^2_A.
\end{aligned}$$

To satisfy (4.1), it is sufficient to use, in the definition (3.11) of S_k, a polynomial $p_{\bar{\lambda}_k,N_k}$ of sufficiently large degree. From (5.10), we have

$$\bar{\lambda}_{k+1,k} \equiv C\frac{h^d_k}{h^2_{k+1}} \geqslant \lambda_{k+1,k}.$$

Further, due to (5.8), we can take

$$\bar{\lambda}_k = Ch_k^{d-2} \geqslant \varrho(A_k).$$

Thus, we have

$$\frac{\bar{\lambda}_k}{\bar{\lambda}_{k+1,k}} = C\Big(\frac{h_{k+1}}{h_k}\Big)^2$$

and to satisfy (4.1), we need

$$\deg(S_k) \geqslant C\frac{h_{k+1}}{h_k},$$

which is guaranteed by (5.4). Thus the assumptions of Theorem 4.1 are verified whenever $\gamma = 1$ or γ is even.

We summarize the above results in the following theorem:

Theorem 5.1. *Consider the model elliptic problem and coarse spaces derived from nested quasiuniform triangulations as described in this section, with the intergrid transfer operators defined by the natural embedding of the spaces* (5.3). *Assume the error propagation operators of both the pre- and post-smoother are given on each level* $k = 0, \dots, l-1$ *by* (3.1), *with* S_k *defined by* (3.11), (3.8), *and either* $\gamma = 1$ *or* $\gamma > 0$ *even. We assume* γ *is bounded. In addition, we assume that the degree of the smoothing polynomial satisfies* (5.4). *Then the resulting multigrid operator,* B, *is nearly spectrally equivalent to* A, *that is,*

$$\mathbf{v}^T A\mathbf{v} \leqslant \mathbf{v}^T B\mathbf{v} \leqslant Cl\mathbf{v}^T A\mathbf{v} \quad \forall \mathbf{v} \in \mathbb{R}^{n_0},$$

where the constant C *is independent on the meshsizes* h_k *(and the coarsening ratio* h_{k+1}/h_k*).*

Acknowledgement. This work relates to Department of the Navy Grant N62909-11-1-7032 issued by Office of Naval Research Global. The United States Government has a royalty-free license throughout the world in all copyrightable material contained herein.

References

[1] *F. A. Bornemann, P. Deuflhard*: The cascadic multigrid method for elliptic problems. Numer. Math. *75* (1996), 135–152.
[2] *J. H. Bramble, J. E. Pasciak, J. Wang, J. Xu*: Convergence estimates for multigrid algorithms without regularity assumptions. Math. Comput. *57* (1991), 23–45.
[3] *M. Brezina, C. Heberton, J. Mandel, P. Vaněk*: An iterative method with convergence rate chosen a priori. Center for Computational Mathematics, University of Colorado at Denver (UCD/CCM) Report *140* (1999), http://ccm.ucdenver.edu/reports/rep140.pdf.
[4] *M. Brezina, P. Vaněk, P. S. Vassilevski*: An improved convergence analysis of smoothed aggregation algebraic multigrid. Numer. Linear Algebra Appl. *19* (2012), 441–469.

[5] *P. G. Ciarlet*: The Finite Element Method for Elliptic Problems. Studies in Mathematics and its Applications. North-Holland, Amsterdam, 1978.
[6] *J. Křížková, P. Vaněk*: Two-level preconditioner with small coarse grid appropriate for unstructured meshes. Numer. Linear Algebra Appl. *3* (1996), 255–274.
[7] *P. Vaněk*: Smoothed prolongation multigrid with rapid coarsening and massive smoothing. Appl. Math., Praha *57* (2012), 1–10.
[8] *P. Vaněk, M. Brezina, J. Mandel*: Convergence of algebraic multigrid based on smoothed aggregation. Numer. Math. *88* (2001), 559–579.
[9] *P. Vaněk, M. Brezina, R. Tezaur*: Two-grid method for linear elasticity on unstructured meshes. SIAM J. Sci. Comput. *21* (1999), 900–923.
[10] *P. S. Vassilevski*: Multilevel Block Factorization Preconditioners. Matrix-Based Analysis and Algorithms for Solving Finite Element Equations. Springer, New York, 2008.
[11] *J. Xu, L. Zikatanov*: The method of alternating projections and the method of subspace corrections in Hilbert space. J. Am. Math. Soc. *15* (2002), 573–597.

Authors' addresses: *Petr Vaněk*, Department of Mathematics, University of West Bohemia, Univerzitní 22, 306 14 Plzeň, Czech Republic, e-mail: `ptrvnk@kma.zcu.cz`; *Marian Brezina*, Department of Applied Mathematics, Campus Box 526, University of Colorado at Boulder, Boulder, CO 80309-0526, USA, e-mail: `marian.brezina@gmail.com`.