LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-PROC-432896

Coarse Spaces by Algebraic Multigrid: Multigrid Convergence and Upscaled Error Estimates

P.S. Vassilevski

May 21, 2010

Sparse Representation of Multiscale Data and Images
Singapore, Singapore
December 14, 2009 through December 17, 2009

Disclaimer

COARSE SPACES BY ALGEBRAIC MULTIGRID: MULTIGRID CONVERGENCE AND UPSCALED ERROR ESTIMATES

PANAYOT S. VASSILEVSKI

Abstract. We give an overview of a number of algebraic multigrid methods targeting finite element discretization problems. The focus is on the properties of the constructed hierarchy of coarse spaces that guarantee (two-grid) convergence. In particular, a necessary condition known as "weak approximation property", and a sufficient one, referred to as "strong approximation property" are discussed. Their role in proving convergence of the TG method (as iterative method) and also on the approximation properties of the AMG coarse spaces if used as discretization tool is pointed out. Some preliminary numerical results illustrating the latter aspect are also reported.

1. Introduction

Consider our problem of main interest

$$A\mathbf{u} = \mathbf{f},$$

where A is a sparse $n \times n$ symmetric positive definite(or s.p.d.) matrix. Typically, A comes from a discretized partial differential equation (or PDE) on a very fine mesh $\mathcal{T}_h$. Multigrid methods (or MG) are becoming the method of choice for solving finite element discretization problems due to their potential for optimal order $\mathcal{O}(n)$ of complexity. This is definitely the case for solving elliptic PDEs with various extensions (e.g., to $H(\text{curl})$ and $H(\text{div})$ problems). Traditionally (or historically) MG is used when a hierarchy of discretization problems corresponding to a respective sequence of finite element discretization spaces obtained by some refinement procedure is available. When this is not the case, in order to apply MG, the needed hierarchy has to be constructed in some problem dependent "algebraic" way (as opposed to the traditional "geometric" mesh refinement way). The latter generally leads to a class of "algebraic" MG (of AMG) method. The AMG concept goes back to the papers by Brandt, McCormick and Ruge ([BMR82]-[BMR84]). Since then much progress has been made and a number of AMG methods have developed that may or may not utilize or assume additional fine-grid information about the problem at hand. To construct an AMG hierarchy is an "inverse problem": there are many hierarchies that can be constructed leading to equally good (or bad) MG performance. One main area of AMG research is to identify necessary and sufficient conditions for the needed

Date: March 27, 2010–beginning; Today is April 5, 2010.

1991 *Mathematics Subject Classification.* 65F10, 65N20, 65N30.

Key words and phrases. algebraic multigrid, upscaling, approximation properties.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

hierarchy of spaces that guarantee at least TG (two–grid) convergence with mesh-independent bounds (when applied to discretization problems). In that respect, two conditions known in the literature, namely a "weak approximation property" and a "strong approximation property" play important role in the TG and MG convergence analysis (see, e.g., [V08]). These conditions are reviewed in more details in § 2. In the present paper, we also point out their role in proving error estimates when solving the discretized problem on a coarser level (using the respective AMG coarse space). Note that the coarse AMG spaces are problem (or operator) dependent and therefore differ from the traditional piecewise polynomial finite element spaces (when available). We show, in § 6, that AMG offers a natural tool for solving PDEs on computationally feasible coarse meshes, commonly referred to as "upscaling". In practice, we need high enough accuracy from the upscaled (coarse) solution. We show that this is the case if the respective coarse spaces exhibit "strong approximation property". Most of the more traditional AMG spaces do possess a two-grid weak approximation property which also implies two-grid convergence (as an iterative method). We illustrate a spectral AMGe method and an AMG with constrained energy minimization basis as upscaling discretization tool. These two methods are in fact a "scale" of methods in the sense that they can become more and more accurate by enriching the coarse basis. We show some preliminary examples that illustrate their potential for improving the quality of the upscaled coarse solution when enriching the respective coarse spaces.

The only known to us AMG method with strong approximation property, the "window-based" spectral AMG [FVZ05], [V08], has too expensive setup cost if meant to be used as an iterative method. The spaces with strong approximation property however are more interesting from upscaling discretization point of view since they imply energy error estimates between the fine–grid solution and the respective AMG coarse space solution without assuming any additional regularity of the underlined PDE. We briefly review the method and suggest some strategies that can be used to reduce the currently high setup cost of generating the coarse space hierarchies in § 6. We also outline, in § 7, a new version of the window-based spectral AMG method that utilizes element matrices. that provides strong approximation properties.

2. Necessary and sufficient conditions for TG convergence

To define a two–grid (TG) method, we need:

- Convergent in A–norm "smoothers" M and M^T with $\mathcal{O}(n)$ cost of implementing one inverse action of M and M^T. Typical examples are forward and backward Gauss-Seidel.
- "Interpolation" and "restriction" matrices: P and P^T with $\mathcal{O}(n)$ cost to implement their actions;
- "Coarse–grid" matrix: $A_c = P^T A P$ of size $n_c \times n_c$.

We want:

- P and hence A_c to be sparse so that recursion can be applied.
- It is also desirable that $\frac{n_c}{n} < 1$, i.e., the problem size reduction be by a factor greater than one. The latter property ensures, when we apply the method

recursively, that the cost of one multilevel cycle is of order $\mathcal{O}(n)$. In geometric MG, typical factors are four or eight, in 2- or 3–space dimensions, respectively.

Once having the TG tools a typical TG algorithms takes the following standard form:

Algorithm 2.1 (TG algorithm)**.** *Let* $\mathbf{x} = 0$*, or any other given initial iterate* $\mathbf{x}$ *for solving* $A\mathbf{x} = \mathbf{b}$*.*

Given a current iterate $\mathbf{x}$*, perform the following steps:*

- *"pre–smoothing": solve* $M\mathbf{y} = \mathbf{b} - A\mathbf{x}$ *and compute the intermediate iterate*
$$\mathbf{x} := \mathbf{x} + \mathbf{y} = \mathbf{x} + M^{-1}(\mathbf{b} - A\mathbf{x}).$$
- *restrict the residual, i.e., compute*
$$\mathbf{r}_c = P^T(\mathbf{b} - A\mathbf{x}).$$
- *solve for a coarse–grid correction,*
$$A_c\mathbf{x}_c = \mathbf{r}_c.$$
- *interpolate and compute next intermediate iterate* $\mathbf{x} := \mathbf{x} + P\mathbf{x}_c$*.*
- *"post–smoothing": solve* $M^T\mathbf{z} = \mathbf{b} - A\mathbf{x}$*, and compute the next two–grid iterate*
$$\mathbf{x}_{TG} = \mathbf{x} + \mathbf{z} = \mathbf{x} + M^{-T}(\mathbf{b} - A\mathbf{x}).$$

The TG algorithm with zero initial iterate provides a mapping (or rather the inverse mapping)

$$\mathbf{b} \mapsto B_{TG}^{-1}\mathbf{b} = \mathbf{x}_{TG}.$$

To have an optimal TG iterative method, we need the spectral equivalence relations

$$\mathbf{v}^T A\mathbf{v} \le \mathbf{v}^T B_{TG}\mathbf{v} \le K_{TG}\ \mathbf{v}^T A\mathbf{v}, \tag{2.1}$$

to hold for a mesh-independent constant $K_{TG} = \frac{1}{1-\varrho_{TG}}$ where

$$\varrho_{TG} = \|I - B_{TG}^{-1}A\|_A$$

is the TG convergence factor. The lower bound in (2.1) is one if M and M^T are convergent in A-norm (which we assume).

The following characterization holds (cf., [FVZ05] or [V08]):

$$K_{TG} = \max_{\mathbf{v}} \frac{\min\limits_{\mathbf{v}_c} \|\mathbf{v} - P\mathbf{v}_c\|^2_{\widetilde{M}}}{\mathbf{v}^T A\mathbf{v}},$$

where

$$\widetilde{M} = M^T(M + M^T - A)^{-1}M \text{ is the symmetrized smoother}$$

(such as symmetric Gauss–Seidel).

Typically $\widetilde{M} \simeq D_A$, the diagonal of A. In the case of M being (forward) Gauss–Seidel, the symmetric Gauss–Seidel matrix $\widetilde{M}$ is known to be spectrally equivalent to D_A. More specifically, we have (cf., e.g., Proposition 6.12 in [V08])

$$\frac{1}{4}\ \mathbf{v}^T D_A\mathbf{v} \le \mathbf{v}^T\widetilde{M}\mathbf{v} \le \delta^2\ \mathbf{v}^T D_A\mathbf{v},$$

where δ is bounded by the maximal number of nonzero entries per row of A. In summary, TG convergence implies the following *weak approximation property*:

$$\|\mathbf{v} - P\mathbf{v}_c\|_{D_A}^2 \simeq \|\mathbf{v} - P\mathbf{v}_c\|_{\widetilde{M}}^2 \leq K_{TG}\ \mathbf{v}^T A\mathbf{v}.$$

In the simplest case, we may assume that $D_A \simeq \|A\|\ I$, which leads to the more familiar form of the weak approximation property stated in a matrix-vector form:

$$\|A\|^{\frac{1}{2}}\ \|\mathbf{v} - P\mathbf{v}_c\| \leq \eta_w\ \|\mathbf{v}\|_A. \tag{2.2}$$

In a finite element setting, when A comes from a bilinear form $a(.,.)$ and a fine-grid f.e. space S_h, using a coarse space S_H on a mesh $H \simeq h$, the above result translates to the following approximation property of the coarse space

$$\inf_{v_H \in S_H} \|v_h - v_H\|_{0,\,\rho} \leq CH\ \sqrt{a(v_h,\ v_h)}. \tag{2.3}$$

The left–hand side is a ρ–weighted L_2–norm (the weight ρ comes from the diagonal of A). This is seen from (2.2), using the equivalence

$$\|v_h\|_0^2 \simeq h^d \|\mathbf{v}\|^2$$

between the integral L_2-norm of a finite element function v_h and the ℓ_2-norm of its coefficient vector $\mathbf{v}$ (corresponding to a nodal basis of S_h) and the well-known fact that $\|A\| \simeq h^{d-2}$ (cf. e.g., Proposition 1.3. in [V08]). In the case when the coarse mesh H is not comparable in size with h, a similar result can be shown for smoothers M coming from overlapping Schwarz methods with subdomains that have diameter $\mathcal{O}(H)$ and similar-size overlap.

In summary, we have the following result.

Proposition 2.1. *Consider the coarse space S_H corresponding to the range of an interpolation mapping P, coming from a convergent TG method with smoother M such that its symmetrized version $\widetilde{M} = M^T(M + M^T - A)^{-1}M$ is spectrally equivalent to D_A (the diagonal of A). Then, S_H exhibits the L_2–approximation property* (2.3).

We comment that generally speaking the coarse-grid interpolant $u_H \in S_H$, i.e., the best approximation defined in (2.3) that corresponds to the solution $u_h \in S_h$ of the fine–grid problem $A\mathbf{u} = \mathbf{f}$, is not computable (without computing u_h having coefficient vector $\mathbf{u}$ at the first place). That is, the coarse space corresponding to a convergent TG method does have an approximation property in L_2, but we do not generally have access to the best approximation u_H. We cannot actually estimate the error between u_h and the coarse-grid solution $\overline{u}_H$ that has coefficient vector $P\mathbf{u}_c$ and $\mathbf{u}_c$ is the solution to the Galerkin coarse–grid problem $P^T AP\mathbf{u}_c = P^T\mathbf{f}$. Our numerical evidence though does suggest that for the AMG methods we consider next the Galerkin coarse–grid problems do possess L_2–approximation property. Since we are also interested to use the AMG generated coarse spaces as discretization (upscaling) tool, which means we want to avoid solving the problem on extremely fine meshes and solve instead only the upscaled (coarse-grid) problem. We show in what follows that if the coarse-space possesses a strong–approximation property, the error estimation in energy norm is possible with the computable Galerkin coarse–grid solution.

In the MG literature the strong approximation property is referred to one of the following estimates:

$$\|A\| \, \|\mathbf{v} - P\mathbf{v}_c\| \leq \eta_s \, \|A\mathbf{v}\|, \tag{2.4}$$

or

$$\|A\| \|\mathbf{v} - P\mathbf{v}_c\|_A^2 \leq \eta_s^2 \|A\mathbf{v}\|^2. \tag{2.5}$$

Note that the first estimate implies the second one. The strong approximation property is known to be only a sufficient condition for (V–cycle) MG convergence (any number of levels, not only two levels). Moreover, a classical result by Braess and Hackbusch (1983) ([BH83], see also [V08]) shows that the convergence improves with the number of smoothing steps $m \geq 1$, i.e.,

$$K_{MG} \leq 1 + \frac{C}{m}.$$

Here K_{MG} is the counterpart of K_{TG} in the multilevel case. For geometric MG, strong approximation property is verified under the assumption of full H^2–regularity of the underlined elliptic PDE associated with the given f.e. bilinear form $a(.,.)$.

For one class of AMG methods, namely the window-based spectral AMG method it can be proved (cf., [V08] and earlier in [FVZ05]) that the stronger first estimate (2.4) holds which is a purely algebraic result (i.e., without any additional assumptions on regularity of the underlined PDE if applied to discretization problems). We study this method and corresponding estimates in more detail in a following section.

3. Element agglomeration AMG methods

If we use fine–grid finite element information, a subclass of AMG is the so–called element based AMG, or AMGe, proposed in [AMGe]. If we generate coarse counterparts of elements and element matrices by recursively agglomerating fine-grid elements, we end up with the element agglomeration AMGe proposed in [JV01].

The agglomerated elements provide a natural partition of the degrees of freedom (element vertices) into nonoverlapping classes (called intersection sets) that are partially ordered, see Fig. 1. One class precedes another one, if the former belongs to (or is a subset of) more agglomerated elements. In particular, there are maximal sets (of single points in Fig. 1) which can naturally be identified as vertices of the agglomerated elements. The partial ordering of the intersection sets can be used to define a boundary of a given class; namely, the classes that precedes it. To define an interpolation mapping associated with a given set of agglomerated elements we have several choices. One is to select coarse degrees of freedom as subset of the fine degrees of freedom (i.e., from the fine-grid element vertices). Then the minimal coarse grid consists of the vertices of the agglomerated elements. However, more coarse degrees of freedom can be selected (cf., [KV06], [VZ04], and [LV08]). Then a P is constructed by moving from a boundary of a class F into itself. At the coarse degrees of freedom P is identity. For any F and its boundary ∂F, we use the local matrix A_{Ω_F} assembled from the fine grid element matrices A_τ for all elements τ that touch (or have a

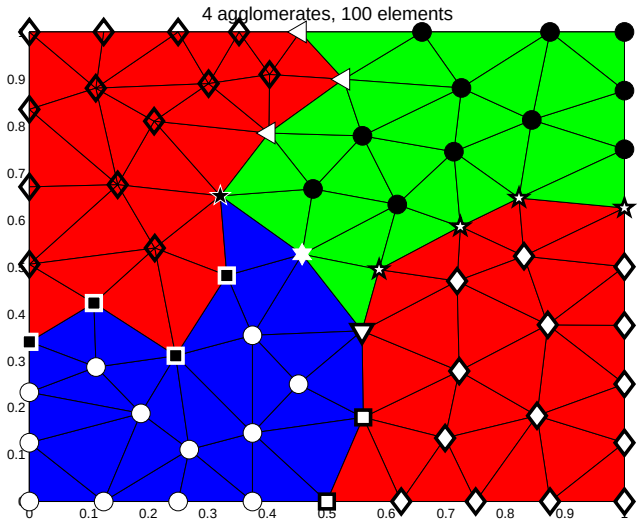


FIGURE 1. Classes of intersection sets

common fine-grid vertex with) $F \cup \partial F$. Given a vector $\mathbf{v}_{\partial F}$, we use minimal energy extension by solving

$$\mathbf{v}^T A_{\Omega_F} \mathbf{v} \mapsto \min,$$

subject to $\mathbf{v}|_{\partial F} = \mathbf{v}_{\partial F}$. Then, to define $P\mathbf{v}_c$ for a given coarse–grid vector $\mathbf{v}_c$, assuming (by induction) that $\mathbf{v}_{\partial F} = (P\mathbf{v}_c)|_{\partial F}$ has been already defined, we extend P to F as $(P\mathbf{v}_c)|_F = \mathbf{v}|_F$ where $\mathbf{v}$ is the above minimal energy extension of $\mathbf{v}_{\partial F}$.

Another alternative in AMGe is to use *spectral* degrees of freedom.

In Chartier et al. [ρAMGe] and in [Ch07] the following spectral choice of coarse degrees of freedom was proposed.

From A_{Ω_F} (defined previously), compute its Schur complement S_F by eliminating all degrees of freedom outside F. Solve the generalized eigenvalue problem for S_F and $D_F = \text{diag}\,(A)|_F$:

$$S_F \mathbf{q}_k = \lambda_k \, D_F \mathbf{q}_k.$$

Choose a portion $m_F < m$ of $\{\mathbf{q}_k\}_{k=1}^m$ in the lower part of the spectrum as "coarse degrees of freedom" and form a local interpolation matrix

$$P_F = [\mathbf{q}_1, \, \ldots, \, \mathbf{q}_{m_F}].$$

One approach to construct the global P is to use harmonic extension of P_F, $F \subset \partial T$ in Ts interior for any agglomerated element T. Other options are also possible. For more details we refer to [LV08].

A new two–level version of a spectral AMGe was proposed recently by J. Galvis and Y. Efendiev in [GE09]. It uses overlapping sets F for the local eigenproblems. In [GEV] a version of that method was proposed that allows for multilevel recursion.

In what follows, we show the performance of the spectral AMGe method from [Ch07] as an upscaling discretization tool.

Consider the 2D diffusion equation posed in $\Omega = (0,1)^2$

$$-\mathrm{div}\rho\nabla u = f, \tag{3.1}$$

where $u = 0$ on $\partial\Omega$ and $f = -1$. Here,

$$\rho = \frac{(2 + 1.8\sin(2\Pi x))(2 + 1.8\sin(2\Pi y))}{(2 - 1.8\sin(2\Pi x))(2 - 1.8\sin(2\Pi y))}, \tag{3.2}$$

A piecewise constant approximation of the diffusion coefficient (3.2) is illustrated in Fig. 2. The hierarchy set of agglomerated elements used in the test is shown in Fig. 3.

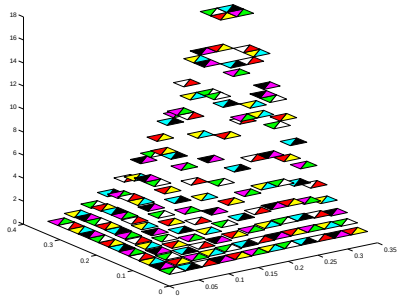


FIGURE 2. Piecewise constant approximation of ρ; $h = 1/36$.

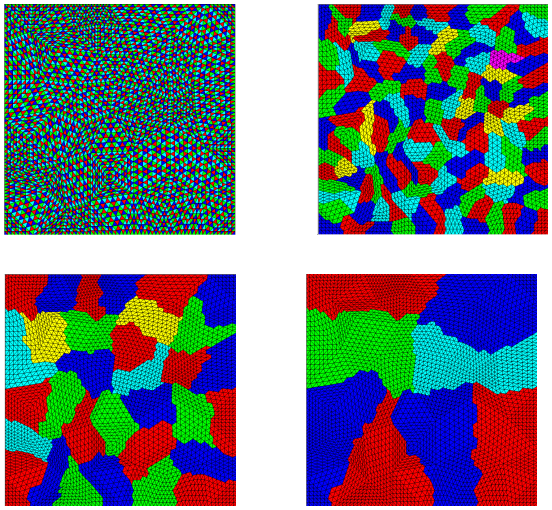


FIGURE 3. Fine grid and level 3, 4 and 5 coarse grids.

The computed coarse grid upscaled finite element solutions u_H are seen in Fig. 4. Finally, the L_2–error $\|u_h - u_H\|_0$ is listed in Table 1. A clear conclusion from that table is that the coarse spaces produced by the spectral AMGe offer reasonable accuracy.

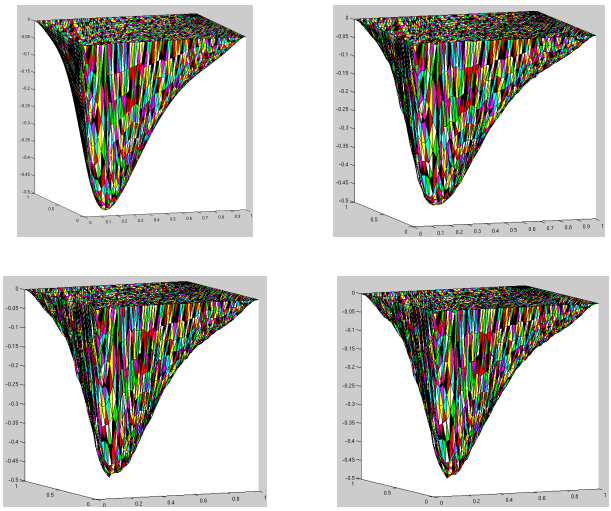


FIGURE 4. Fine grid solution and level 3, 4 and 5 coarse–grid solutions.

level	# elements	# dofs	# nnz	$\|u_h - u_H\|_0$
0	6400	3321	22761	
3	128	732	14874	1.303002e-02
4	32	248	5444	1.861606e-02
5	8	92	2180	2.098479e-02

TABLE 1. L_2- error of spectral AMGe upscaled solutions

4. AMG coarse spaces by constrained energy minimization

In this section, we briefly summarize an AMG method proposed recently in [V10]. Our goal is to illustrate its potential as an upscaling discretization tool. Its performance as a MG method was demonstrated in [V10]. For previous work on coarse spaces by constrained energy minimization we refer, for example, to [Wag96], [WCS00], [XZ04], [KV06], [VZ04], and [vLSG].

We are given a set of agglomerates $\{T\}$ (see e.g. Fig. 5). To construct a basis of locally supported functions for the coarse space associated with the set of agglomerated elements, we proceed as follows. For each T, in the simplest case, we construct a single basis φ_T locally supported (in Ω_T - the union of all neighbors $T^{'}$ of T) by solving the constrained minimization problem:

$$J(\varphi_T) = \frac{1}{2}\, a(\varphi_T,\ \varphi_T) \mapsto \min, \tag{4.1}$$

subject to the constraints (in this case, prescribed averages):

$$\int_{T'} \varphi_T\, dx = \delta_{T,T'} \int_{T} 1\, dx \text{ for all } T^{'} \subset \Omega_T. \tag{4.2}$$

By construction, the basis provides approximate partition of unity:

$$\sum_T \varphi_T \approx \mathbf{1}$$

The method allows for fitting arbitrary set of given functions $\{v_k\}_{k=1}^m$, The above simple case corresponded to $m = 1$ and v_1 being the constant function 1. In the multiple function case, we assign to each T and $k \geq 1$, a basis function $\varphi_T^{(k)}$ that solves the following local problem

$$J(\varphi_T^{(k)}) = \frac{1}{2}\, a(\varphi_T^{(k)},\ \varphi_T^{(k)}) \mapsto \min,$$

subject to the constraints

$$\int_{T'} v_l \varphi_T^{(k)}\, dx = \delta_{T,T'} \int_T v_l v_k\, dx \text{ for all } l = 1,\ \ldots,\ m.$$

By construction, we have, for $1 \leq k \leq m$,

$$\sum_T \varphi_T^{(k)} \approx v_k.$$

That is, the coarse space contains approximately the given functions $\{v_k\}_{k=1}^m$,

If we minimize the global functional

$$a(\sum_T \varphi_T^{(k)} - v_k,\ \sum_T \varphi_T^{(k)} - v_k) \mapsto \min,$$

subject to the same integral moments constraints, then we can ensure that the coarse space contains as accurately as we want the given functions v_k. Since, by solving the local problems, we already have

$$\sum_T \varphi_T^{(k)} \approx v_k,$$

the above global minimization requires only few "smoothing" overlapping Schwarz iterations.

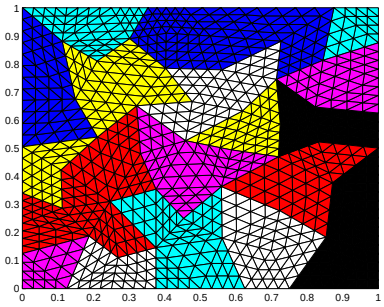


FIGURE 5. Agglomerated elements T obtained by one level of coarsening.

In Fig. 6 we show typical examples of coarse basis functions corresponding to the constant function.

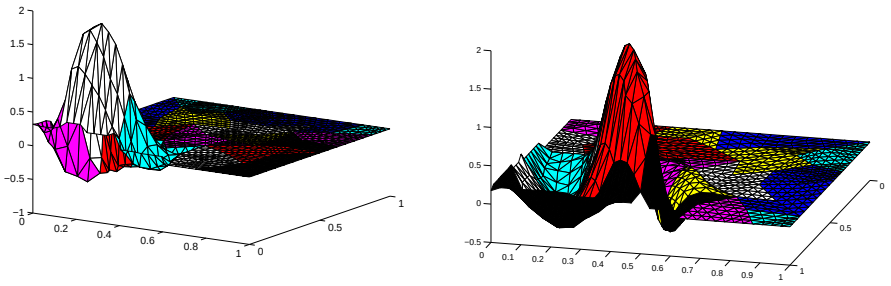


FIGURE 6. Typical coarse basis functions based on fitting one (constant) function.

The case of multiple functions $v_k = \sin(\Pi k_x x)\sin(\Pi k_y y)$, $k = (k_x, k_y)$, $k_x,\ k_y = 1, \ldots, \sqrt{m}$, $(m = 4)$ is illustrated in Fig. 7 for the Laplace operator.

The influence of the diffusion operator $-\mathrm{div}\rho\nabla u$, with coefficient (3.2) (see Fig. 2) on the recovery is seen in Fig. 8.

Finally, we demonstrate the upscaling properties of the method applied to the diffusion operator (used previously) $-\mathrm{div}\rho\nabla u = f$, $u = 0$ on $\partial\Omega$. $\rho = \frac{(2+1.8\sin(2\Pi x))(2+1.8\sin(2\Pi y))}{(2-1.8\sin(2\Pi x))(2-1.8\sin(2\Pi y))}$, and $f = -1$.

In Fig. 9 and Fig. 10, the fine–grid solution and respective coarse-grid solutions are shown, for two cases of coarse spaces (one based on fitting four bilinear functions and another one based on fitting one sin function). The corresponding errors are plotted in Fig. 11 and 12, respectively. It is clear that the richer coarse space provide better approximation.

5. The strong approximation property

In this section, we discuss the strong approximation property mentioned earlier.

We are given a s.p.d. $n \times n$ sparse matrix A and let $P:\ \mathbb{R}^{n_c} \mapsto \mathbb{R}^n$, $n_c < n$, be a given (rectangular) interpolation matrix.

We are interested in the following *strong approximation property*:

For any fine-grid vector $\mathbf{u} \in \mathbb{R}^n$ there is a coarse interpolant $P\mathbf{u}_c$ such that

$$\|A\|\|\mathbf{u} - P\mathbf{u}_c\|_A^2 \le C_A\ \|A\mathbf{u}\|^2. \tag{5.1}$$

If the problem of our main interest

$$A\mathbf{u} = \mathbf{f},$$

comes from a finite element discretization of a PDE on a domain $\Omega \subset \mathbb{R}^d$ ($d = 2$ or 3), then $\mathbf{f} = (f_i)$ comes from a given r.h.s. function $f(\mathbf{x}) \in L_2(\Omega)$, where the entries f_i are computed as the following integral moments

$$f_i = (f,\ \varphi_i) \equiv \int_\Omega f(\mathbf{x})\varphi_i\, d\mathbf{x}.$$

Above, φ_i runs over a basis of the fine–grid finite element space V_h associated with a triangulation of Ω with characteristic fine-grid mesh size h. For a nodal (Lagrangian)

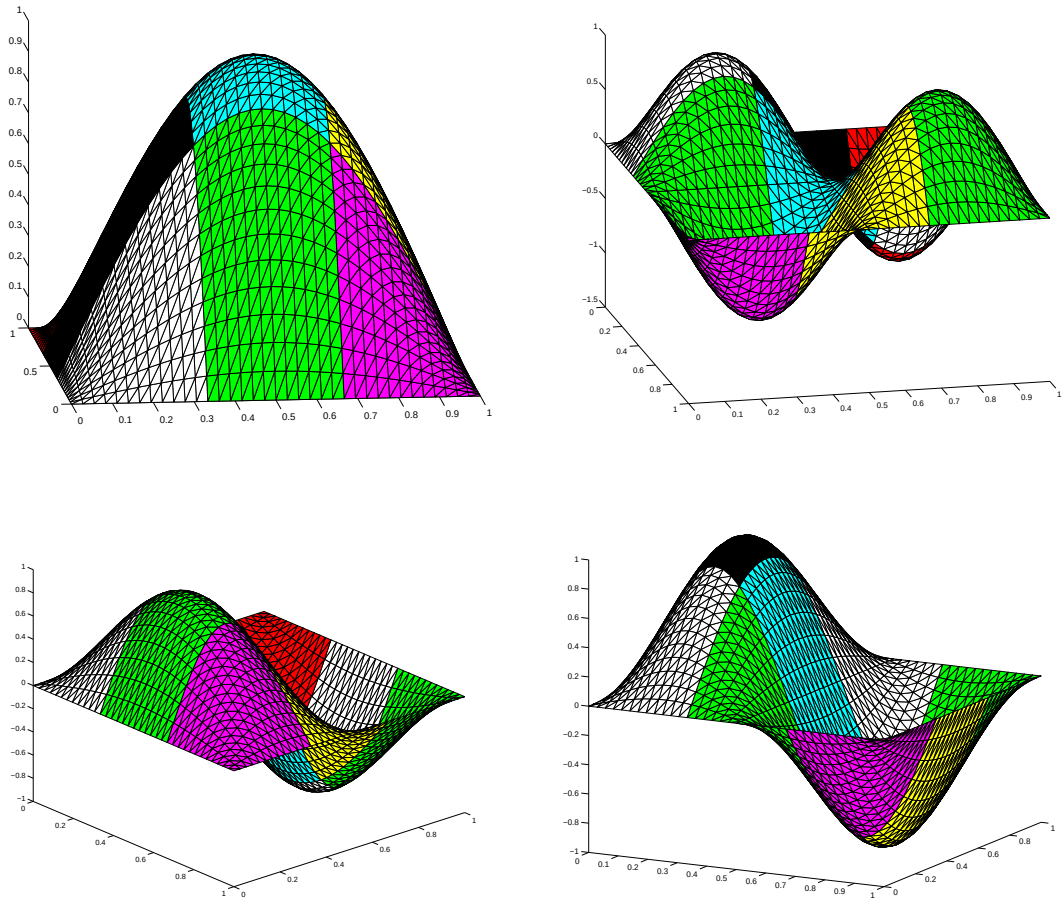


FIGURE 7. $\sum_T \Phi_T^{(k)}$ based on fitting four sin functions v_k on a 3×3 coarse mesh $(H = 1/3)$; $h = 1/36$.

basis, the index "i" runs over the set of fine degrees of freedom $x_i \in \mathcal{N}_h$. The unknown $\mathbf{u}$ stands for the coefficient vector of the finite element, Galerkin, approximation u_h to the solution of the underlined PDE posed variationally, i.e., $u_h \in V_h$ solves the discretized PDE in a variational form associated with a given bilinear form $a(.,\ .)$, stated as follows

$$a(u_h,\ \varphi) = (f,\ \varphi) \text{ for all } \varphi \in V_h.$$

As an example, we consider a second order self–adjoint elliptic bilinear form $a(u,\ \varphi) = \int_\Omega k(\mathbf{x})\ \nabla u \cdot \nabla \varphi\ d\mathbf{x}$ for $u,\ \varphi \in H_0^1(\Omega)$ and a given positive coefficient function $k = k(\mathbf{x}),\ \mathbf{x} \in \Omega$, the given polygonal/polyhedral domain in $\mathbb{R}^d$, $d = 2$ or 3. Using a standard piecewise linear conforming finite element space V_h on a quasiuniform triangulation $\mathcal{T}_h$, as it is well-known, the stiffness matrix $A = (a(\varphi_j,\ \varphi_i))$ computed from a nodal Lagrangian basis $\{\varphi_i\}_{\mathbf{x}_i \in \mathcal{N}_h}$ of V_h satisfies

$$\|A\| \simeq h^{d-2}. \tag{5.2}$$

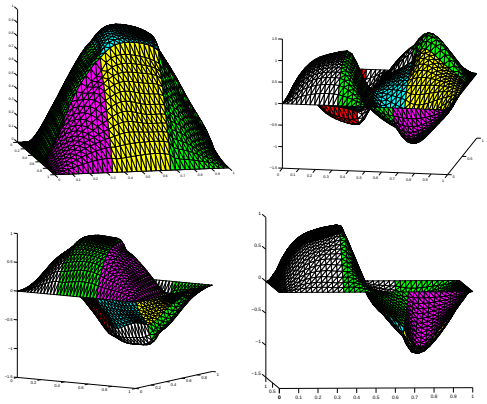
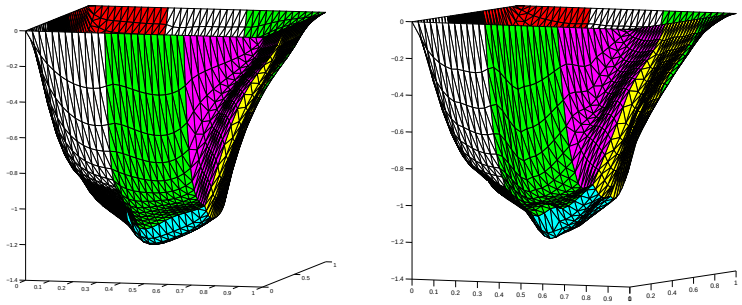
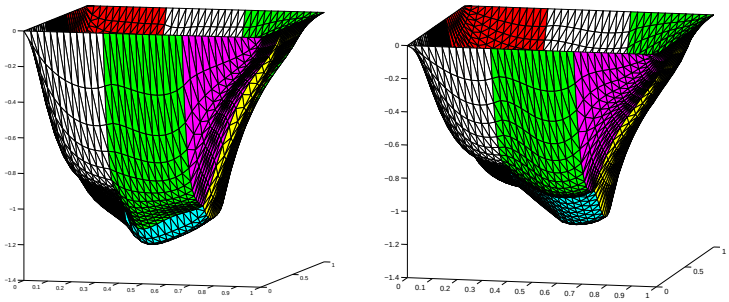


FIGURE 8. $\sum_T \Phi_T^{(k)}$ from diffusion operator and 4 sin functions; 3×3 coarse mesh ($H = 1/3$); $h = 1/36$.



Fine–grid solution; $h = 1/36$ and upscaled solution on 3×3 coarse mesh, $H = 1/3$.

FIGURE 9. Four bilinear functions v_k used in the constraints.



Fine–grid solution; $h = 1/36$ and upscaled solution on 3×3 coarse mesh, $H = 1/3$.

FIGURE 10. One sin function used in the constraints.

The equivalence constants above generally depend on the variation $\frac{\max_{\mathbf{x}\in\Omega} k(\mathbf{x})}{\min_{\mathbf{x}\in\Omega} k(\mathbf{x})}$ but are mesh-independent.

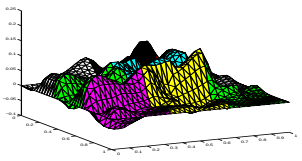


FIGURE 11. Pointwise error for function u_H obtained using four bilinear functions v_k; $\|u_h - u_H\|_0 = 0.0503$.

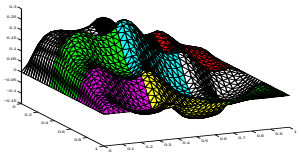


FIGURE 12. Pointwise error for function u_H obtained using one *sin* function v_k; $\|u_h - u_H\|_0 = 0.1113$.

Assume, that we have come up with a coarse space $V_H \subset V_h$ such that the coefficient vectors of functions in V_H viewed as elements of V_h can be represented as the range of an interpolation mapping P. We can define respective coarse basis functions by forming $P\mathbf{e}_{i_c}$ for each coarse coordinate vector $\mathbf{e}_{i_c} \in \mathbb{R}^{n_c}$ that has a single nonzero entry at the i_cth position. Then, consider the fine–grid function $\phi_{i_c}^{(H)}$ that has coefficient vector the i_cth column of P, i.e., equal to $P\mathbf{e}_{i_c}$. The set of functions $\{\phi_{i_c}^{(H)}\}$ forms the coarse basis of interest. The parameter H stands for characteristic size of the support of the coarse basis functions.

The above matrix–vector strong approximation property (5.1) admits the following finite element function form:

$$\|A\|\ a(u_h - u_H,\ u_h - u_H) \le C_A \sum_{\mathbf{x}_i \in \mathcal{N}_h} f_i^2 = C_A \sum_{\mathbf{x}_i \in \mathcal{N}_h} \left(\int_\Omega f(\mathbf{x}) \varphi_i \, d\mathbf{x} \right)^2 .$$

Using Cauchy-Schwarz inequality, we have

$$\sum_{\mathbf{x}_i \in \mathcal{N}_h} \left(\int_\Omega f(\mathbf{x}) \varphi_i \, d\mathbf{x} \right)^2 \le \sum_{\mathbf{x}_i \in \mathcal{N}_h} \int_{\text{support } (\varphi_i)} f^2(\mathbf{x}) \, d\mathbf{x} \int_\Omega \varphi_i^2 \, d\mathbf{x}.$$

For a fairly general class of basis functions (including piecewise linears) on a quasiuniform mesh, we have

$$\int_\Omega \varphi_i^2 \, d\mathbf{x} \simeq |\text{support } (\varphi_i)| \simeq h^d.$$

Due to the bounded overlap of the supports of the finite element basis functions, we also have

$$\sum_{\mathbf{x}_i\in\mathcal{N}_h}\int\limits_{\text{support }(\varphi_i)} f^2(\mathbf{x})\,d\mathbf{x}\le\kappa\,\|f\|_0^2.$$

Thus, we arrive at the energy error estimate of our main interest (using (5.2))

$$a(u_h-u_H,\ u_h-u_H)\le C_A\kappa\,\frac{h^d}{\|A\|}\|f\|_0^2\simeq C_A\kappa\,h^2\,\|f\|_0^2.$$

In practice, we typically have $C_A=\mathcal{O}(\left(\frac{H}{h}\right)^2)$ with a constant in the $\mathcal{O}$ symbol, independent of the two mesh sizes (h and H); see, e.g., Corollary 7.1 later on. Thus, we get the following final upscaling energy error estimate:

$$a(u_h-u_H,\ u_h-u_H)\le c_A\,H^2\|f\|_0^2. \tag{5.3}$$

In the remaining sections, we summarize a few AMG methods that provide *strong approximation property*.

6. AMG with strong approximation property: the "window"-based spectral AMG

In [FVZ05] (see also [V08]), the following AMG method was proposed that exhibits strong approximation property. The original version tends to lead to relatively large coarse spaces so that the resulting two (and multi)–level methods have unacceptably high complexities. In the present section, we propose several approaches in the attempt to reduce the complexity of the original method.

Given an overlapping partition $\{w\}$ of the set of indices $i=1,2,\ \ldots,\ n$, we extract the rows of a given $n\times n$ matrix A with indices from any given set (called window) w. The respective rectangular matrix is denoted by A_w. By proper reordering, A_w can be written as follows

$$A_w=[A_{ww},\ A_{w,\,\chi}]\,.$$

Here, A_{ww} is the principal submatrix of A (row and column indices from w) and $A_{w,\chi}$ is the submatrix of A with columns outside w (and row indices from w).

We are interested, for a proper nonnegative diagonal matrix D_w, in the normal matrices $A_w^TD_wA_w$. The diagonal matrices D_w provide a partition of unity, i.e., if I_w stands for extension by zero outside the set w, then $\sum\limits_w I_wD_wI_w^T=I$. This property ensures that

$$\sum_w\mathbf{v}^TA_w^TD_wA_w\mathbf{v}=\sum_w\mathbf{v}^TA^TI_wD_wI_w^TA\mathbf{v}=\|A\mathbf{v}\|^2,\qquad A_w=I_w^TA. \tag{6.1}$$

The method in question uses the symmetric semi-definite Schur complements S_w defined as follows:

$$\mathbf{v}_w^TS_w\mathbf{v}_w=\inf_{\mathbf{v}_\chi}\begin{bmatrix}\mathbf{v}_w\\ \mathbf{v}_\chi\end{bmatrix}^TA_w^TD_wA_w\begin{bmatrix}\mathbf{v}_w\\ \mathbf{v}_\chi\end{bmatrix}. \tag{6.2}$$

The original method utilizes the eigenvectors of the semidefinite Schur complements S_w,

$$S_w \mathbf{p}_k = \lambda_k \ \mathbf{p}_k, \ k = 1, \ \ldots, \ n_w. \tag{6.3}$$

For efficiency reason, for a given tentative interpolation matrix $\widetilde{P}$, we use in (6.3) instead the modified Schur complements

$$\mathbf{v}_w^T S_w \mathbf{v}_w = \inf_{\mathbf{v}_\chi \in I_\chi^T \mathrm{Range}(\widetilde{P})} \begin{bmatrix} \mathbf{v}_w \\ \mathbf{v}_\chi \end{bmatrix}^T A_w^T D_w A_w \begin{bmatrix} \mathbf{v}_w \\ \mathbf{v}_\chi \end{bmatrix}. \tag{6.4}$$

In what follows, we denote the exact window Schur complement with S_w^*.

We first form local interpolation matrices P_w by putting together the first $n_w^c \geq 1$ eigenvectors (in the lower part of the spectrum of S_w), i.e.,

$$P_w = \left[\mathbf{p}_1, \ \ldots, \ \mathbf{p}_{n_w^c}\right]. \tag{6.5}$$

The corresponding eigenvalues (ordered in an increasing order) are such that $\lambda_k \leq \mathrm{tol}\ \lambda_{\max}(S_w)$ for $k \leq n_w^c$ and

$$\lambda_k(S_w) > \mathrm{tol}\ \|S_w\| = \mathrm{tol}\ \lambda_{\max}(S_w) \ \text{ for } k > n_w^c. \tag{6.6}$$

Here, we have the freedom to choose the pre-selected tolerance "tol" (a number between zero and one) that may also vary with w.

The eigenvectors $\{\mathbf{p}_k\}_{k=1}^{n_w}$ are orthogonal and assumed normalized.

The global P is computed based on another partition of unity set of nonnegative $n_w \times n_w$ diagonal matrices $\{Q_w\}$ that satisfy

$$I = \sum_w I_w Q_w I_w^T.$$

Then P is defined as follows

$$P = \sum_w I_w Q_w \left[0, \ P_w, \ 0\right] = \sum_w I_w Q_w P_w (I_w^c)^T. \tag{6.7}$$

Here, I_w^c maps the local indices of the eigenvectors coming from the window w to their global indices expanding the result with zeros elsewhere. Thus we have defined a process that from a tentative $\widetilde{P}$ produces another one P. This can be iterated several times (by possibly changing the parameters such as $\{w\}$ and tol). In the next theorem, we formulate conditions ensuring that P admits a strong approximation property.

Theorem 6.1. *Consider the iterated window spectral AMG interpolation matrix P constructed on the basis of the modified window Schur complements using a $\widetilde{P}$ that satisfies the following estimate*

$$\sum_w \|D_w^{\frac{1}{2}} A_{w,\,\chi}(\mathbf{v}_\chi - I_\chi^T \widetilde{P} \mathbf{w}_c)\|^2 \leq \mu \ \|A\mathbf{v}\|^2. \tag{6.8}$$

That is, $\widetilde{P}$ is such that for any $\mathbf{v}$, when restricted to a complementary set χ, there is a coarse vector $\mathbf{w}_c$ (depending on $\mathbf{v}$ and the set χ) such that for a fixed number $\mu > 0$

(6.8) *holds. Then, if we choose tol* $= \frac{1}{\delta} \leq 1$ *in the two-level spectral decomposition defining the local* P_w *so that (see* (6.6)*)*

$$\|S_w\| \leq \delta \, \lambda_{m_w+1}(S_w),$$

and if we also assume the quasiuniformity of the windows, i.e., the estimate

$$\beta \, \|A\|^2 \leq \|S_w^*\|, \tag{6.9}$$

then, the following main strong approximation property holds for P

$$\|A\|^2 \|\mathbf{v} - P\mathbf{v}_c\|^2 \leq \eta \, \|A\mathbf{v}\|^2. \tag{6.10}$$

Here, $\eta = \frac{\delta}{\beta} \, (1 + \sqrt{\mu})^2$*, where* μ *is from* (6.8).

The proof of this theorem falls outside the scope of the present paper and can be found in a forthcoming report.

7. A new "window"-based spectral AMG method utilizing fine-grid element matrices

Here, we present a summary of a modified version of the "window"-based spectral AMG method applied to finite element matrices A. The method is studied and analyzed in more details in a forthcoming report.

The new ingredient is in the different eigenproblems that we use. Also, it utilizes a special partition of unity matrices. A main additional assumption is that the window sets are covered exactly by fine–grid elements and that we have access to the respective fine–grid element matrices so that we can assemble the semi–definite local matrices further denoted by Λ_w. Therefore, we have the identity $A = \sum_w I_w \Lambda_w I_w^T$.

We solve eigenproblems associated with the pair of matrices Λ_w and S_w, where S_w is the exact window-based Schur complement (as introduced before). The eigenproblems read (compare with (6.3)):

$$S_w \mathbf{p}_k = \lambda_k \, \Lambda_w \, \mathbf{p}_k, \; k = 1, \ldots, \; n_w, \tag{7.1}$$

where the eigenvalues are numbered in an increasing order and the eigenvectors are Λ_w–normalized.

Since, the matrices Λ_w can also be only semi–definite, to have real eigenvalues the nullspace of Λ_w should be contained in the nullspace of S_w, which is the case for finite elliptic matrices (Laplacian–like as well as elasticity).

It is clear that we can choose the eigenvectors $\mathbf{p}_k$ be orthogonal to the nullspace of Λ_w (and Λ_w–orthogonal to each other). Let the columns of $P_w^{(0)}$ span the nullspace of Λ_w. Then, we have $\mathbf{p}_k^T P_w^{(0)} = 0$.

Based on a preselected tolerance tol $\in [0, 1)$, we choose n_w^c such that $\lambda_k >$ tol λ_{n_w} for $k > n_w^c$. The local interpolation matrices are defined similarly as before (cf., (6.5)), now augmented with the nullspace, i.e.,

$$P_w = \left[P_w^{(0)}, \; \mathbf{p}_1, \; \ldots, \; \mathbf{p}_{n_w^c} \right]. \tag{7.2}$$

To define the global one, we use special diagonal matrices $\{Q_w\}_w$ with nonnegative entries that provide partition of unity, i.e., we have

$$I = \sum_w I_w Q_w I_w^T.$$

Each Q_w has entries on its diagonal $q_{w,\,i}$, $i \in w$, defined as follows:

$$q_{w,\,i} = \frac{\|\Lambda_w\|}{\sum\limits_{w'\colon\, i\in w'} \|\Lambda_{w'}\|}. \tag{7.3}$$

At the end we formulate our main result (without a proof) just to illustrate the potential of the method.

Theorem 7.1. *Let A be a given finite element s.p.d. matrix. Consider a given set of windows $\{w\}$ where each window w is exactly covered by fine–grid elements. Assume also that the local finite element matrices Λ_w corresponding to the sets w are available. The nullspace (if nonempty) of the local matrices Λ_w is assumed known (explicitly computed). That is, let the nullspace of Λ_w be represented by the range of an explicitly available local matrix $P_w^{(0)}$. Assume that this nullspace is contained in the nullspace of the window Schur complement S_w (defined in* (6.2)*). The global interpolation matrix P is defined as in* (6.7) *based on the local interpolation matrices* (7.2) *and the weights $q_{w,\,i}$ (entries of the diagonal partition of unity matrices Q_w) are defined in* (7.3)*. Then, the following global strong approximation property holds*

$$\|\mathbf{v} - P\mathbf{v}_c\|_A^2 \le \kappa \ \max_w \ Cond^+(\Lambda_w) \ \max_w \ \left(\frac{1}{tol\ \lambda_{\max}(\Lambda_w^+ S_w)}\right) \ \|A\mathbf{v}\|^2.$$

Above, $\kappa \ge 1$ depends on the overlap of $\{w\}$, $tol \in (0,1]$ (in general depending on w) is the tolerance used to define the portion of the eigenvectors $\mathbf{p}_k$ in the lower part of the spectrum computed in (7.1) *used to define P_w, $Cond^+(\Lambda_w) = \frac{\|\Lambda_w\|}{\lambda^+_{\min}(\Lambda_w)}$ is the effective condition number of Λ_w computed in a subspace orthogonal to the nullspace of Λ_w. Finally, $\lambda_{\max}(\Lambda_w^+ S_w) = \max_k \lambda_k$ where λ_k are from* (7.1)*.*

Corollary 7.1. *For finite element s.p.d. matrices A coming from second order elliptic problems, the constructed finite element modification of the window-based spectral AMG method, the following strong approximation property holds*

$$\|\mathbf{v} - P\mathbf{v}_c\|_A^2 \le C\ \kappa\ \left(\frac{H}{h}\right)^2 \ \max_w \ \frac{1}{tol\ \|A_{w_0,w_0}\|}\ \|A\mathbf{v}\|^2.$$

Here, $w_0 \subset w$ is strictly in the interior of w in the sense that no other windows w' intersect it.

References

[BH83] D. Braess and W. Hackbusch, *A new convergence proof of the multigrid method including the V–cycle*, SIAM Journal on Numerical Analysis **20**(1983), pp. 967–975.

[BMR82] A. Brandt, S. McCormick and J.W. Ruge, "*Algebraic multigrid (AMG) for automatic multigrid solutions with application to geodetic computations,*" Report, Institute for Computational Studies, Fort Collins, Colorado, 1982.

[BMR84] A. Brandt, S. McCormick and J.W. Ruge, "*Algebraic multigrid (AMG) for sparse matrix equations,*" in Sparsity and its Applications (D.J. Evans, ed.), Cambridge University Press, 1984, pp. 257–284.

[AMGe] M. Brezina, A. J. Cleary, R. D. Falgout, V. E. Henson, J. E. Jones, T. A. Manteuffel, S. F. McCormick, and J. W. Ruge, "*Algebraic multigrid based on element interpolation (AMGe),*" SIAM Journal on Scientific Computing **22**(2000), pp. 1570-1592.

[ρAMGe] T. Chartier, R. Falgout, V. E. Henson, J. Jones, T. Manteuffel, S. McCormick, J. Ruge, and P. S. Vassilevski, "*Spectral AMGe (ρAMGe),*" SIAM Journal on Scientific Computing **25**(2003), pp. 1–26.

[Ch07] T. Chartier, R. Falgout, V. E. Henson, J. Jones, T. Manteuffel, S. McCormick, J. Ruge, and P. S. Vassilevski, "*Spectral element agglomerate AMGe,*" in Domain Decomposition Methods in Science and Engineering XVI, Lecture Notes in Computational Science and Engineering, Springer-Verlag, Berlin-Heidelberg, **55**(2007), pp. 515–524.

[FVZ05] R. Falgout, P. S. Vassilevski, and L. T. Zikatanov, "*On two–grid convergence estimates,*" Numerical Linear Algebra with Applications, **12**(2005), pp. 471–494.

[GE09] J. Galvis and Y. Efendiev, "*Domain decomposition preconditioners for multiscale flows in high contrast media,*" Preprint, submitted (2009).

[GEV] J. Galvis, Y. Efendiev and P. S. Vassilevski, *Spectral element agglomerate algebraic multigrid methods for elliptic problems with high contrast coefficients*, Proceedings of the 19th International Conference of Domain Decomposition Methods, 2009.

[JV01] J. E. Jones and P. S. Vassilevski, "*AMGe based on element agglomerations*", SIAM Journal on Scientific Computing **23**(2001), pp. 109-133.

[KV06] T. V. Kolev and P. S. Vassilevski, *AMG by element agglomeration and constrained energy minimization interpolation*, Numerical Linear Algebra with Applications **13**(2006), pp. 771–788.

[LV08] I. Lashuk and P. S. Vassilevski, *On some versions of the element agglomeration AMGe method*, Numerical Linear Algebra with Applications **15** (2008), pp. 595–620.

[vLSG] J. van Lent, R. Scheichl, and I.G. Graham, *Energy minimizing coarse spaces for two-level Schwarz methods for multiscale PDEs*. Numerical Linear Algebra with Applications **16**(2009), pp. 775-799.

[VZ04] P. S. Vassilevski and L. T. Zikatanov, *Multiple vector preserving interpolation mappings in algebraic multigrid*, SIAM Journal on Matrix Analysis and Applications, 27 (2006), pp. 1040–1055.

[V08] P. S. Vassilevski, "Multilevel Block–Factorization Preconditioners. Matrix-based Analysis and Algorithms for Solving Finite Element Equations", Springer, New York, 2008.

[V10] P. S. Vassilevski, "*General constrained energy minimization interpolation mappings for AMG,*" SIAM Journal on Scientific Computing **32**(2010), pp. 1–13.

[Wag96] C. Wagner, *On the algebraic construction of multilevel transfer operators*, Computing **65**(2000), pp. 73–95.

[WCS00] W. L. Wan, T. F. Chan, and B. Smith, *An energy-minimizing interpolation for robust multigrid methods*, SIAM Journal on Scientific Computing **21**(1999), pp. 1632–1649.

[XZ04] J. Xu and L. T. Zikatanov, *On an energy minimizing basis in algebraic multigrid methods*, Computing and Visualization in Science **7** (2004), pp. 121–127.

Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, P.O. Box 808, L-561, Livermore, CA 94551, U.S.A.

E-mail address: panayot@llnl.gov