LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-JRNL-744422

Modifying AMG Coarse Spaces with Weak Approximation Property to Exhibit Approximation in Energy Norm

X. Hu, P. S. Vassilevski

January 11, 2018

SIAM Journal on Matrix Analysis and Applications

MODIFYING AMG COARSE SPACES WITH WEAK APPROXIMATION PROPERTY TO EXHIBIT APPROXIMATION IN ENERGY NORM*

XIAOZHE HU † AND PANAYOT S. VASSILEVSKI‡

Abstract. Algebraic multigrid (AMG) coarse spaces are commonly constructed so that they exhibit the so-called weak approximation (WAP) property which is necessary and sufficient condition for uniform two-grid convergence. This paper studies a modification of such coarse spaces so that the modified ones provide approximation in energy norm. Our modification is based on the projection in energy norm onto an orthogonal complement of original coarse space. This generally leads to dense modified coarse space matrices which is hence computationally infeasible. To remedy this, based on the fact that the projection involves inverse of a well-conditioned matrix, we use polynomials to approximate the projection and, therefore, obtain a practical, sparse modified coarse matrix and prove that the modified coarse space maintains computationally feasible approximation in energy norm. We present some numerical results for both, PDE discretization matrices as well as graph Laplacian ones, which are in accordance with our theoretical results.

Key words. AMG, weak approximation property, strong approximation property

AMS subject classifications. 65F10, 65N20, 65N30

1. Introduction.

Algebraic multigrid is one of the most successful methods for solving large-scale sparse systems of linear equations $A\mathbf{u} = \mathbf{f}$ with symmetric positive definite (SPD) matrix A, especially for the case when A comes from finite element discretization of second order elliptic equations. AMG has also been extended to matrices arising from much broader classes of discretized PDEs (e.g., [24], the AMS and ADS solvers in [14], [15]) and even for non-PDE matrices (using adaptive AMG, see e.g., [4, 8]), including ones coming from network simulations (e.g. graph Laplacian, [20]). For an overview of some AMG methods, we refer to [28] and more recently to [29].

Another important aspect of AMG, which is the main focus of this work, is that it provides a hierarchy of coarse spaces, which are natural candidates for dimension reduction, sometimes referred to as numerical *upscaling*. There are quite a few literature on multigrid-based upscaling techniques, e.g., [11, 21, 23], and domain-decomposition-based upscaling approaches, e.g., [18, 26]. However, one difficulty, which needs to be overcome with such an approach, is that the traditional AMG coarse spaces can not guarantee the required approximation accuracy. More precisely, by the construction, traditional AMG coarse spaces only guarantee to possess a so-called *weak approximation property* (WAP), i.e., for any vector $\mathbf{u} \in \mathbb{R}^n$, there exists a vector $\mathbf{u}_c$ belonging to the coarse space, such that $\|\mathbf{u} - P\mathbf{u}_c\|_D \le \eta_w \|\mathbf{u}\|_A$, where $\|\mathbf{u}\|_A := \sqrt{\mathbf{u}^T A\mathbf{u}}$ is the so-called energy norm and $\|\mathbf{u}\|_D := \sqrt{\mathbf{u}^T D\mathbf{u}}$ is the (weighted) ℓ_2-norm induced by a proper chosen SPD matrix D. The WAP is known to be necessary and sufficient for the uniform convergence of the two-level AMG methods (cf., e.g., [28]). However, to

*This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. The work of the second author was partially supported by NSF under grant DMS-1619640. The work of Hu was partially supported by NSF grant DMS-1620063.

†Department of Mathematics, Tufts University, Medford, MA 02155. (xiaozhe.hu@tufts.edu)

‡Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, P.O. Box 808, L-561, Livermore, CA 94551, U.S.A. and Department of Mathematics, Portland State University, Portland, OR 97201 (panayot@llnl.gov, panayot@pdx.edu).

use the same coarse space for dimension reduction, we need that the Galerkin projection (projection with respect to energy norm $\|\cdot\|_A$) onto the coarse space exhibit some approximation property. A sufficient condition is that the coarse spaces satisfy the so-called *strong approximation property* (SAP), i.e., the coarse-level solution should approximate the original (fine-level) solution with some guaranteed accuracy in energy norm. Mathematically, the SAP means that, for any vector $\mathbf{u} \in \mathbb{R}^n$, there exists a vector $\mathbf{u}_c$ belonging to the coarse space, such that $\|A\|\|\mathbf{u} - P\mathbf{u}_c\|_A^2 \leq \eta_s \|A\mathbf{u}\|^2$. Although the AMG coarse spaces do have approximation properties (by construction, in a weighted ℓ_2-norm), the coarse-level solution (i.e., the computationally feasible Galerkin projection) does not generally possess that, neither in (weighted) ℓ_2-norm nor in energy norm $\|\cdot\|_A$. To the best of our knowledge, none of the existing multigrid- and domain decomposition-based upscaling techniques have the desired SAP property with provable satisfactory bound on the resulting constant η_s.

In this paper, we address the issue that the usual AMG coarse spaces do not satisfy the SAP with provable satisfactory bound on the resulting constant η_s and develop an approach by extending a construction originated in [22] to our more general AMG upscaling setting. Our main contribution, which distinguish our result from all the existing results, is that our modified coarse space satisfies the SAP with provable satisfactory bound on the resulting constant η_s, which provides computable approximation to the fine-level solution in both the energy norm and (weighted) ℓ_2-norm. The proposed method simply modifies the AMG coarse space Range(P) (P is the prolongation matrix which satisfies the WAP by construction/assumption) to Range($(I - \pi_f)P$) where π_f is a projection onto the A-orthogonal complement of Range(P) (i.e., orthogonal complement of Range(P) with respect to the A inner product $\mathbf{u}^T A\mathbf{v}$). We show that such modified coarse space provides a two-level A-orthogonal decomposition of the original fine-level solution $\mathbf{u}$ and, thereby, energy error estimate of the coarse solution. Moreover, the SAP of the coarse space can be derived based on such decomposition as well. Details of the construction of π_f will be presented in Section 3. Because the definition of π_f involves the inverse of a matrix (see Section 2.3 for details), such modification typically leads to dense coarse matrices which is mostly of theoretical interest. In order to design a more practical approach, we take advantage of the fact that A is well-conditioned on the A-orthogonal complement of Range(P) (which we prove holds for P satisfying the WAP) and, therefore, modifying the coarse spaces based on polynomial approximations to control the sparsity of the respective coarse matrices is feasible. That is, we are able to modify the coarse space so that both, the SAP (hence the error estimate in the energy norm) and the sparsity of the coarse matrix, are satisfied. The energy error estimate improves when the polynomial degree increases (with the expense of increased matrix density). We present numerical results illustrating the effectiveness of the proposed method. We would like to point out that other computationally feasible AMG-type upscaling approaches were presented in [1] and [13] for problems that can be formulated in a mixed (saddle-point) form.

The remainder of the paper is structured as follows. In Section 2, we introduce the WAP and formulate some properties of the matrices arising from the unsmoothed aggregation AMG. It provides the motivation for the construction of the improved coarse spaces which is presented in Section 3. The error analysis in the computationally infeasible case with exact projections is presented in that section as well. The computationally feasible case with approximate projections, giving rise to the improved coarse space satisfying the SAP and with guaranteed approximation properties is presented in Section 4. The case of elliptic problems with high contrast coefficients is briefly discussed in Section 5. The numerical illustration of the presented methods for both, PDE-type matrices and graph Laplacian ones, can be found in Section 6. Finally, some conclusions are drawn in the last Section 7.

2. Weak approximation property in AMG.

In this section, we recall the two-grid method and the weak approximation property that is widely used to prove the convergence of two-grid methods. We point out that the prolongation matrices P constructed in various AMG methods usually satisfy the WAP (a notion intoduced already in the original AMG paper, [2].

We consider a SPD matrix $A \in \mathbb{R}^{n\times n}$ and let D be another SPD matrix such that,

$$\mathbf{v}^T A\mathbf{v} \leq \mathbf{v}^T D\mathbf{v}. \tag{2.1}$$

A typical choice of D is the diagonal of A with proper scaling, i.e. $D = \omega^{-1}\operatorname{diag}(A)$, $\omega \in \mathbb{R}$ or the so-called "ℓ_1-*smoother*" (cf. e.g., [5]). We denote the norms induced by A and D by $\|\cdot\|_A$ and $\|\cdot\|_D$, respectively.

2.1. The two-grid method.

First, we briefly recall the standard two-grid method. Assume we have a smoother M such as Jacobi, Gauss-Seidel, etc., a prolongation P, and the coarse-grid problem $A_c = P^T AP$. Based on these standard components, we define the standard (symmetrized) two-grid method in Algorithm 2.1.

Algorithm 2.1 Two-grid method

For a current iterate $\mathbf{u}$, we perform:

1: Presmoothing: $\mathbf{u} \leftarrow \mathbf{u} + M^{-1}(\mathbf{f} - A\mathbf{u})$
2: Restriction: $\mathbf{r}_c \leftarrow P^T(\mathbf{f} - A\mathbf{u})$
3: Coarse-grid correction: $\mathbf{e}_c = A_c^{-1}\mathbf{r}_c$
4: Prolongation: $\mathbf{u} \leftarrow \mathbf{u} + P\mathbf{e}_c$
5: Postsmoothing: $\mathbf{u} \leftarrow \mathbf{u} + M^{-T}(\mathbf{f} - A\mathbf{u})$

It is well-known that the two-grid method (Algorithm 2.1) leads to the composite iteration matrix E_{TG} based on which we define the two-grid operator B_{TG} as follow,

$$I - B_{TG}^{-1}A = E_{TG} = (I - M^{-T}A)(I - PA_c^{-1}P^TA)(I - M^{-1}A).$$

For the convergence rate of the two-grid method, we have the following two-grid estimates which can be found in Theorem 4.3, [10].

THEOREM 2.1. *For B_{TG} and the two-grid error propagation operator E_{TG}, we have the sharp estimates*

$$\mathbf{v}^T A\mathbf{v} \leq \mathbf{v}^T B_{TG}\mathbf{v} \leq K_{TG}\mathbf{v}^T A\mathbf{v} \quad \textit{or equivalently} \quad \|E_{TG}\|_A = \rho_{TG} := 1 - \frac{1}{K_{TG}},$$

where

$$K_{TG} = \max_{\mathbf{v}} \min_{\mathbf{v}_c} \frac{\|\mathbf{v} - P\mathbf{v}_c\|^2_{\widetilde{M}}}{\|\mathbf{v}\|^2_A},$$

and $\widetilde{M} := M^T(M^T + M - A)^{-1}M$ is the symmetrized smoother (starting with M^T).

2.2. The Weak Approximation Property. In AMG, we construct a prolongation $P \in \mathbb{R}^{n \times n_c}$ and the corresponding coarse space Range(P) which exhibits the WAP. We note that the WAP is a necessary and sufficient condition for uniform two-level AMG convergence (e.g., [28]) and can be stated as, for any vector $\mathbf{v} \in \mathbb{R}^n$, there is a coarse vector $\mathbf{v}_c \in \mathbb{R}^{n_c}$, such that

$$\|\mathbf{v} - P\mathbf{v}_c\|_D \leq \eta_w \, \|\mathbf{v}\|_A, \tag{2.2}$$

where η_w is the so-called WAP constant. By requiring that the smoother is spectrally equivalent to D, which can be verified for standard smoothers such as Gauss-Seidel and Jacobi, we can estimate the two-grid constant K_{TG} based on the WAP. More precisely, we have $K_{TG} \leq c\eta_w^2$ where the constant c here measures the spectral equivalence between $\widetilde{M}$ and D. This implies that $\rho_{TG} \leq 1 - \frac{1}{c\eta_w^2}$, i.e., the corresponding two-grid method converges uniformly.

In order to have a computationally feasible approach (which will become clear later on), in this paper, we follow [5, 27] and assume that P is constructed based on aggregation-based approach (without smoothing). Roughly speaking, we first form a set of aggregates $\{\mathcal{A}_i\}_{i=1}^{n_a}$, which is a nonoverlapping partitioning of the index set $\{1,\ 2,\ \ldots,\ n\}$, i.e., $\cup_{i=1}^{n_a} \mathcal{A}_i = \{1,\ 2,\ \ldots,\ n\}$ and $\mathcal{A}_i \cap \mathcal{A}_j = \emptyset$, if $i \neq j$. Moreover, we denote the size of $\mathcal{A}_i$ by $n_{\mathcal{A}_i}$ which is defined by the cardinality of $\mathcal{A}_i$. We then solve certain (generalized) eigenvalue problems locally to obtain the local basis $\{\mathbf{q}^c_{\mathcal{A}_i,j}\}_{j=1}^{n_i^c}$ for each aggregate $\mathcal{A}_i$. The overall prolongation is defined as

$$P = \begin{pmatrix} P_{\mathcal{A}_1} & 0 & \cdots & 0 \\ 0 & P_{\mathcal{A}_2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & P_{\mathcal{A}_{n_a}} \end{pmatrix} \text{ with } P_{\mathcal{A}_i} = (\mathbf{q}^c_{\mathcal{A}_i,1}, \cdots, \mathbf{q}^c_{\mathcal{A}_i,n_i^c}), \tag{2.3}$$

and, naturally, the coarse space is just Range(P). The WAP (2.2) can be shown by the properties of the local eigenvalue problems. We refer to [5, 27] for the details. We note that such local spectral construction of P (2.3) dated back to [3] and is also possible for graph Laplacian matrices (see, e.g., [12]).

As already mentioned, a WAP of the above form is a necessary condition for uniform two-level AMG convergence, so we assume (2.2) to hold for a block-diagonal P and a diagonal D (scaled as in (2.1)).

The assumptions on P and D imply that the matrix P^TDP is sparse, actually it is block diagonal with each diagonal block corresponding to an aggregate $\mathcal{A}_i$. Hence, it is easily invertible and the projection $\pi_D = P(P^TDP)^{-1}P^TD$ is sparse, hence computationally feasible. Taking $\mathbf{v}_c = (P^TDP)^{-1}P^TD\mathbf{v}$ in (2.2), we arrive at the following estimate, which is another way to present the WAP of the coarse space using the projection π_D,

$$\|\mathbf{v} - \pi_D\mathbf{v}\|_D \leq \eta_w \, \|\mathbf{v}\|_A. \tag{2.4}$$

As already mentioned, the WAP plays an important role in the convergence analysis of AMG methods. For example, we can derive two-level convergence rate directly from the WAP. However, in this paper, our goal is to take advantage of the WAP and modify the coarse space such that the modified one satisfies not only the WAP but also the so-called strong approximation property. Coarse spaces that satisfy the SAP with provable satisfactory bound on the constant can provide a coarse-level solution which approximates the fine-level solution with guaranteed accuracy in energy norm, and, therefore, are important both theoretically (e.g., in the V-cycle convergence analysis) and practically (e.g., for upscaling).

2.3. The A-Orthogonal Complement to $\mathrm{Range}(I-\pi_D)$**.** Our modification of the coarse spaces (which will be presented in the next section) uses information from the orthogonal complement $\mathrm{Range}(I-\pi_D)$. Therefore, in this subsection, we introduce how to construct a sparse linearly independent basis of the space $\mathrm{Range}(I-\pi_D)$ and how to project a coarse vector onto it.

The construction of the basis of the space $\mathrm{Range}(I-\pi_D)$ is, of course, not unique. Here, we are looking for a sparse (locally supported) basis due to computational complexity considerations. In the case of aggregation-based AMG, this can be done as follows. On each aggregate $\mathcal{A}_i$, we select n_i^f vectors, $\{\mathbf{q}^f_{\mathcal{A}_i,j}\}_{j=1}^{n_i^f}$, which are orthonormal with respect to $D_{\mathcal{A}_i} := D|_{\mathcal{A}_i}$ and span the $D_{\mathcal{A}_i}$-orthogonal complement of $\mathrm{Range}(P_{\mathcal{A}_i})$. Recall that $n_{\mathcal{A}_i}$ is the size of the aggregates $\mathcal{A}_i$ and n_i^c be the number of columns of $P_{\mathcal{A}_i}$, we choose n_i^f such that $n_{\mathcal{A}_i} = n_i^c + n_i^f$. It is clear that the vectors $\mathbf{q}^f_{\mathcal{A}_i,j}$ extended by zero outside $\mathcal{A}_i$ form a basis of $\mathrm{Range}(I-\pi_D)$. Introducing the matrix $P_\perp$ with the vectors $\mathbf{q}^f_{\mathcal{A}_i,j}$ as its columns, then we have,

$$P_\perp^T D P = 0 \text{ and } P_\perp^T D P_\perp = I. \tag{2.5}$$

Exploiting the local basis of $\mathrm{Range}(I-\pi_D)$, we project any given vector $P\mathbf{v}_c \in \mathrm{Range}(P)$ onto the A-orthogonal complement of $\mathrm{Range}(I-\pi_D)$ by solving the following problem: find $\mathbf{v}_f \in \mathrm{Range}(I-\pi_D)$, such that

$$(\mathbf{w}_f)^T A \mathbf{v}_f = (\mathbf{w}_f)^T A P \mathbf{v}_c, \quad \forall\, \mathbf{w}_f \in \mathrm{Range}(I-\pi_D). \tag{2.6}$$

Since we have a sparse (computable) basis of $\mathrm{Range}(I-\pi_D)$ represented by $P_\perp$, we can rewrite (2.6) as the following linear system of equations,

$$A_f \overline{\mathbf{v}}_f = P_\perp^T A P \mathbf{v}_c, \tag{2.7}$$

where $A_f = P_\perp^T A P_\perp$ and $\mathbf{v}_f = P_\perp \overline{\mathbf{v}}_f$. By solving (2.7), we compute the projection $\mathbf{v}_f = \pi_f P \mathbf{v}_c$. In fact, the matrix representation of π_f is given by $\pi_f = P_\perp A_f^{-1} P_\perp^T A$. Note the inverse of A_f is involved in the definition of π_f.

We next study the conditioning of A_f with the goal to derive computationally feasible (sparse) approximations to its inverse within reasonable computational cost. We have the following main result.

THEOREM 2.2. *If the coarse space* $\mathrm{Range}(P)$ *satisfies the WAP with constant* η_w*, then the condition number* $\kappa(A_f)$ *of* A_f *satisfies* $\kappa(A_f) \le \eta_w^2$*.*

Proof. Choose $\mathbf{v} = \mathbf{v}_f := (I-\pi_D)\mathbf{v}$ in (2.4) and (2.1), which leads to the following spectral equivalence relations,

$$\frac{1}{\eta_w^2}\, \mathbf{v}_f^T D \mathbf{v}_f \le \mathbf{v}_f^T A \mathbf{v}_f \le \mathbf{v}_f^T D \mathbf{v}_f, \quad \forall\, \mathbf{v}_f \in \mathrm{Range}(I-\pi_D).$$

Equivalently, letting $\mathbf{v}_f = P_\perp \overline{\mathbf{v}}_f$, using properties (2.5), we have

$$\frac{1}{\eta_w^2}\, \overline{\mathbf{v}}_f^T \overline{\mathbf{v}}_f \le \overline{\mathbf{v}}_f^T A_f \overline{\mathbf{v}}_f \le \overline{\mathbf{v}}_f^T \overline{\mathbf{v}}_f, \quad \forall\, \overline{\mathbf{v}}_f, \tag{2.8}$$

which implies that the condition number $\kappa(A_f)$ of A_f, satisfies $\kappa(A_f) \le \eta_w^2$, which is the desired result. ∎

Remark 2.3. When the WAP constant η_w is bounded, especially independent of problem size, then Theorem 2.2 implies that A_f is well-conditioned.

Since A_f is well-conditioned, A_f^{-1} can be accurately approximated by a matrix polynomial $q_\nu(A_f)$ of degree ν. Therefore, to approximate the solution of $A_f\mathbf{x}_f = \mathbf{f}_f$, we can use the representation

$$\mathbf{x}_f = A_f^{-1}\mathbf{f}_f = \left[\left(A_f^{-1} - q_\nu(A_f)\right)\mathbf{f}_f + q_\nu(A_f)\mathbf{f}_f\right].$$

The first term on the right hand side above can be made as small as we want by choosing q_ν appropriately. More precisely, it can be made of order $\epsilon \ll 1$ if we choose the polynomial degree $\nu = \mathcal{O}(\log \epsilon^{-1})$ (cf., e.g., [9], or [28], p. 413). We give specific examples of polynomials $q_\nu(t)$ in Section 4. By dropping the first term, we get the approximation

$$\mathbf{x}_f \approx q_\nu(A_f)\mathbf{f}_f. \tag{2.9}$$

An important observation is that, if $\mathbf{f}_f$ is locally supported (sparse), the above approximation can be kept reasonably sparse. In particular, consider (2.7), i.e., $\mathbf{f}_f = P_\perp^T AP\mathbf{v}_c$, and let $\mathbf{v}_c$ be one of the unit coordinate vectors, then $P\mathbf{v}_c$ is a column of P and has local support represented by a corresponding aggregate $\mathcal{A}$. Thus, such $\mathbf{f}_f$ is locally support on $\mathcal{A}$ and its immediate neighbors. In this case, the approximation $q_\nu(A_f)\mathbf{f}_f$ is supported locally. More precisely, its suport depends on the sparsity of A_f^ν, hence the diameter of the non-zero pattern of $q_\nu(A_f)\mathbf{f}_f$ can be estimated to be of order ν times the size of the neighborhood of $\mathcal{A}$ and, therefore, can be kept under control when ν is kept small.

The above approximation is the main motivation for our work. Roughly speaking, such approximation allows us to modify the original coarse space (with the WAP) so that the modified one satisfies the SAP while keeping the sparsity of the modified prolongation under control. In the next two sections, we first introduce the SAP result in the case of exact A_f^{-1} and then present the coarse space modification based on the computationally feasible polynomial approximation.

3. The modified coarse space exhibiting the SAP.

In this section we define the modified coarse space. The construction presented here goes back to [22]. In this paper, we adopt a matrix-vector presentation and motivate the applicability of the construction in [22] to our setting of aggregation-based AMG exploiting the well-conditionedness of A_f proven in Theorem 2.2. Thereby, we extend the analysis in [22] to our more general (algebraic) setting by showing that the modified coarse spaces satisfy the SAP with provable satisfactory bound on the resulting constant η_s. In the following section, we extend these results to the case of approximate inverses.

3.1. Modification of the Coarse Space.

We first recall the projection $\pi_f = P_\perp A_f^{-1} P_\perp^T A$ which plays an important role in the construction of the modified coarse space. We also recall the original coarse space given by $\text{Range}(P) = \text{Range}(\pi_D)$. The modified coarse space of our main interest is simply $\text{Range}((I - \pi_f)\pi_D)$, or equivalently $\text{Range}((I - \pi_f)P)$. Naturally, the modified prolongation matrix takes the form $(I - \pi_f)P$.

Next, we show that we can obtain an A-orthogonal decomposition of any given vector $\mathbf{u}$ based on the modified coarse space, which in turn implies the SAP of our main interest. To this end, we first present some properties of the two projections π_D and π_f summarized in the following lemma.

LEMMA 3.1. *The projections π_D and π_f satisfy $\pi_D\pi_f = 0$. In addition, we have that $(I - \pi_f)\pi_D$ is also a projection.*

Proof. $\pi_D\pi_f = 0$ can be directly verified by $\pi_D = P(P^TDP)^{-1}P^TD$ and $\pi_f = P_\perp A_f^{-1}P_\perp^T A$. Together with properties (2.5), we have

$$\pi_D\pi_f = P(P^TDP)^{-1}\underbrace{(P^TDP_\perp)}_{=0}A_f^{-1}P_\perp^T A = 0.$$

On the other hand, using $\pi_D\pi_f = 0$ and also the fact that $\pi_D^2 = \pi_D$, we have

$$\begin{aligned}((I - \pi_f)\pi_D)^2 &= (\pi_D - \pi_f\pi_D)(\pi_D - \pi_f\pi_D) \\ &= \pi_D^2 - \pi_f\pi_D^2 - (\pi_D\pi_f)\pi_D + \pi_f(\pi_D\pi_f)\pi_D \\ &= \pi_D - \pi_f\pi_D = (I - \pi_f)\pi_D,\end{aligned}$$

which implies that $(I - \pi_f)\pi_D$ is a projection. □

We are now ready to derive our main two-level A-orthogonal decomposition.

THEOREM 3.2. *For a given $\mathbf{u}$, there exists a $\mathbf{v}$, such that*

$$\mathbf{u} = (I - \pi_D)\mathbf{v} + (I - \pi_f)\pi_D\mathbf{u}. \tag{3.1}$$

Also, the two components in the above decomposition are A-orthogonal.

Proof. We begin with the following A-orthogonal decomposition

$$\mathbf{u} = (I - \pi_D)\mathbf{v} + \boldsymbol{\xi}, \text{ where } \boldsymbol{\xi} \in (\mathrm{Range}(I - \pi_D))^{\perp_A}. \tag{3.2}$$

Given a $\mathbf{v}_c$, from the definition of $\mathbf{v}_f = \pi_f P\mathbf{v}_c$ in (2.6), we have

$$\mathbf{w}_f^T A(I - \pi_f)P\mathbf{v}_c = 0, \text{ for all } \mathbf{w}_f \in \mathrm{Range}(I - \pi_D).$$

The latter identity implies that the A-orthogonal complement $(\mathrm{Range}(I - \pi_D))^{\perp_A}$ of $\mathrm{Range}(I - \pi_D)$ satisfies the relations

$$(\mathrm{Range}(I - \pi_D))^{\perp_A} = \mathrm{Range}\left((I - \pi_f)P\right) = \mathrm{Range}\left((I - \pi_f)\pi_D\right). \tag{3.3}$$

This means that in (3.2), $\boldsymbol{\xi} = (I - \pi_f)\pi_D\mathbf{w}$ for some $\mathbf{w}$, hence the A-orthogonal decomposition (3.2) can be rewritten as follows,

$$\mathbf{u} = (I - \pi_D)\mathbf{v} + (I - \pi_f)\pi_D\mathbf{w}.$$

Finally, using Lemma 3.1 we have $\pi_D\mathbf{u} = \pi_D(I - \pi_f)\pi_D\mathbf{w} = \pi_D^2\mathbf{w} = \pi_D\mathbf{w}$, which shows (3.1). □

The above A-orthogonal decomposition (3.1) basically provides an energy stable decomposition since

$$\|\mathbf{u}\|_A^2 = \|(I - \pi_D)\mathbf{v}\|_A^2 + \|(I - \pi_f)\pi_D\mathbf{w}\|_A^2.$$

This is essential in multilevel analysis. In the following subsections, we prove the SAP for the modified coarse space $\mathrm{Range}((I - \pi_f)P)$ and also establish our first main error estimates, all based on this decomposition.

3.2. The Strong Approximation Property. In this subsection, we show that the modified coarse space $\mathrm{Range}((I-\pi_f)P)$ satisfies the SAP with provable satisfactory bound on the constant η_s. To this end, for given $\mathbf{f}$, we consider the solution $\mathbf{u}$ of the following linear system,

$$A\mathbf{u} = \mathbf{f}. \tag{3.4}$$

The corresponding modified coarse problem (also known as the upscaled problem) reads

$$P^T(I-\pi_f)^T A(I-\pi_f)P\mathbf{u}_c = P^T(I-\pi_f)^T\mathbf{f}. \tag{3.5}$$

In order to show the SAP, we are interested in estimating the error $\mathbf{e} = \mathbf{u}-(I-\pi_f)P\mathbf{u}_c$ in the energy norm $\|\cdot\|_A$, more precisely, the estimate of $\|\mathbf{u}-(I-\pi_f)P\mathbf{u}_c\|_A$ in terms of $\|\mathbf{f}\| = \|A\mathbf{u}\|$. The main result is formulated in the following theorem.

THEOREM 3.3. *Assume the WAP (2.2) holds. Let $\mathbf{e} = \mathbf{u}-(I-\pi_f)P\mathbf{u}_c$ be the error between the fine-level solution $\mathbf{u}$ of problem (3.4) and the upscaled (coarse) solution $\overline{\mathbf{u}}_c = (I-\pi_f)P\mathbf{u}_c$ of (3.5). Then, the following energy error estimate holds:*

$$\|\mathbf{e}\|_A \le \eta_w \|D^{-\frac{1}{2}}A\mathbf{u}\|. \tag{3.6}$$

Proof. By the property of the Galerkin projection, we have that $\mathbf{e} = \mathbf{u}-\overline{\mathbf{u}}_c$ is A-orthogonal to $\mathrm{Range}((I-\pi_f)P) = \mathrm{Range}((I-\pi_f)\pi_D) = (\mathrm{Range}(I-\pi_D))^{\perp_A}$. Therefore, using the decomposition (3.1), we have

$$\mathbf{e} = \mathbf{u}-\overline{\mathbf{u}}_c = (I-\pi_D)\mathbf{v}, \text{ for some } \mathbf{v}. \tag{3.7}$$

Since $\overline{\mathbf{u}}_c \in \mathrm{Range}((I-\pi_f)P) = (\mathrm{Range}(I-\pi_D))^{\perp_A}$, we also have

$$\|\mathbf{e}\|_A^2 = (\mathbf{u}-\overline{\mathbf{u}}_c)^T A(I-\pi_D)\mathbf{v} = (A\mathbf{u})^T(I-\pi_D)\mathbf{v} \le \|D^{-\frac{1}{2}}\mathbf{f}\|\|(I-\pi_D)\mathbf{v}\|_D.$$

Using the weak approximation property (2.4) for $\mathbf{v} := (I-\pi_D)\mathbf{v} = \mathbf{e}$, we then obtain

$$\|\mathbf{e}\|_A^2 \le \|D^{-\frac{1}{2}}\mathbf{f}\|\eta_w\|(I-\pi_D)\mathbf{v}\|_A = \eta_w\|D^{-\frac{1}{2}}\mathbf{f}\|\|\mathbf{e}\|_A,$$

which implies (3.6). ∎

From the energy error estimate (3.6), assuming that D is well-conditioned, we have the following corollary also known as *strong approximation property*.

COROLLARY 3.4 (Strong Approximation Property). *We have the following estimate*

$$\|A\|\|\mathbf{u}-\overline{\mathbf{u}}_c\|_A^2 \le \|D\|\|\mathbf{u}-\overline{\mathbf{u}}_c\|_A^2 \le \eta_s\ \|A\mathbf{u}\|^2, \tag{3.8}$$

where $\eta_s \le \|D\|\|D^{-1}\|\eta_w^2$, which is referred to as the SAP constant. If D is well-conditioned, then η_s is bounded from above by a constant.

As we have shown, the modified coarse space $\mathrm{Range}(I-\pi_f)P$ satisfies the SAP with provable satisfactory bound on the constant η_s. However, we want to point out that, the practical usage of this modified coarse space is limited since π_f involves A_f^{-1} which is dense in general. In Section 4, we discuss how to use the polynomial approximation (2.9) to modify the coarse space which can be used in practice with the SAP approximately satisfied.

3.3. A Weighted ℓ_2–Error Estimate. The estimate (3.6) allows us to prove an ℓ_2–error estimate, which is a direct application of the Aubin-Nitsche argument. Let $\mathbf{e} = \mathbf{u} - \overline{\mathbf{u}}_c$ be the error and consider the following linear system,

$$A\mathbf{w} = D\mathbf{e}.$$

We have,

$$\|\mathbf{e}\|_D^2 = \mathbf{e}^T(D\mathbf{e}) = \mathbf{e}^T A\mathbf{w}.$$

Since $\mathbf{e}$ is A-orthogonal to the modified coarse space Range$(I - \pi_f)\pi_D$, we have, for $\overline{\mathbf{w}}_c = (I - \pi_f)\pi_D\mathbf{w}$,

$$\|\mathbf{e}\|_D^2 = \mathbf{e}^T A(\mathbf{w} - \overline{\mathbf{w}}_c) \leq \|\mathbf{e}\|_A \|\mathbf{w} - \overline{\mathbf{w}}_c\|_A.$$

Applying estimate (3.6) to the error $\mathbf{e}_\mathbf{w} := \mathbf{w} - \overline{\mathbf{w}}_c$ leads to

$$\|\mathbf{e}\|_D^2 \leq \|\mathbf{e}\|_A \eta_w \|D^{-\frac{1}{2}} A\mathbf{w}\| = \eta_w \|\mathbf{e}\|_A \|D^{\frac{1}{2}}\mathbf{e}\| = \eta_w \|\mathbf{e}\|_A \|\mathbf{e}\|_D.$$

This implies the desired weighted ℓ_2-error estimate stated below.

THEOREM 3.5. *Let $\mathbf{e} = \mathbf{u} - (I - \pi_f)\pi_D\mathbf{u}$ be the error between the solutions of the original fine-level problem (3.4) and the upscaled one (3.5). Then, the following weighted ℓ_2–error estimate holds:*

$$\|\mathbf{e}\|_D \leq \eta_w \|\mathbf{e}\|_A \leq \eta_w^2 \|D^{-\frac{1}{2}} A\mathbf{u}\|. \tag{3.9}$$

4. Modified coarse space using approximate inverses.

In this section, we discuss how to use approximations to make the modified coarse spaces more practical. The basic idea is based on the well-conditioning of A_f as shown in Theorem 2.2 which allows for uniform polynomial approximation (2.9). We argue that such an approximation keeps the sparsity of the modified prolongation matrix under control while maintaining the approximation properties of the modified coarse space reasonably well. These are properties that make the resulting modified coarse spaces appropriate for upscaling as well for efficient use in multigrid methods in practice.

4.1. Modification via Polynomial Approximation. We begin with one possible choice of polynomial approximation. Recall that according to (2.8), the spectrum of A_f is contained in $[1/\eta_w^2, 1] \subset (0, 1]$. Therefore, we want to chose a polynomial p_ν of degree $\nu \geq 1$, such that $p_\nu(0) = 1$ and $tp_\nu^2(t)$ has a small maximum norm over the interval $t \in [0, 1]$. One choice is the polynomial used in the smoothed aggregation algebraic multigrid (SA-AMG). It is defined via the Chebyshev polynomials of odd degree, $T_{2\nu+1}$, as follows:

$$p_\nu(t) = \frac{(-1)^\nu}{2\nu + 1} \frac{T_{2\nu+1}(\sqrt{t})}{\sqrt{t}}. \tag{4.1}$$

As is well-known (e.g., shown in [6, 28, 12]), this polynomial has the following property

$$\max_{t \in (0,1]} \sqrt{t} |p_\nu(t)| = \frac{1}{2\nu + 1}. \tag{4.2}$$

Since $p_\nu(0) = 1$, $p_\nu(t) = 1 - tq_{\nu-1}(t)$, where $q_{\nu-1}$ is a polynomial of degree $\nu - 1$. We actually use $q_{\nu-1}(t)$ to approximate A_f^{-1}, namely

$$A_f^{-1} \approx \widetilde{A}_f^{-1} \equiv q_{\nu-1}(A_f). \tag{4.3}$$

By rewriting (4.3), we get

$$I - \widetilde{A}_f^{-1} A_f = I - q_{\nu-1}(A_f) A_f = p_\nu(A_f).$$

The A_f-norm of this matrix can be made arbitrarily small as $\nu \to \infty$ by the property (4.2).

Letting $\widetilde{\pi}_f := P_\perp \widetilde{A}_f^{-1} P_\perp^T A$, we define the modified prolongation matrix $\widetilde{P}$ as follows,

$$\widetilde{P} := (I - \widetilde{\pi}_f) P. \tag{4.4}$$

The corresponding modified coarse space is $\mathrm{Range}((I - \widetilde{\pi}_f)P) = \mathrm{Range}((I - \widetilde{\pi}_f)\pi_D)$. Note that, if we choose ν properly (sufficiently large but fixed), the modified prolongation matrix $\widetilde{P}$ stays reasonably sparse and can be used in practice with nearly optimal computational cost.

We notice that, the formula $\widetilde{P} = (I - P_\perp q_{\nu-1}(A_f) P_\perp^T A)P$, somewhat resembles the construction of prolongation matrices used in SA-AMG. More specifically, in SA-AMG, we have $\widetilde{P} := p_\nu(D^{-1}A)P$. This observation offers the possibility to construct new SA-AMG methods by choosing simple $P_\perp$ (for example, not necessarily spanning the entire complement of $\mathrm{Range}(P)$) so that $A_f := P_\perp^T A P_\perp$ and hence the resulting $\widetilde{P}$ and respective modified coarse level matrix $\widetilde{P}^T A \widetilde{P}$ be reasonably sparse.

Remark 4.1. We may also note that $\widetilde{P} = (I - P_\perp q_{\nu-1}(A_f) P_\perp^T A)P$ resembles the so-called *approximate wavelet modified hierarchical basis (AWMHB) method* where $P_\perp$ (corresponding to the HB) is modified by polynomially based approximate L_2-projections to exhibit better energy stability (cf. [25] or [28]).

With the approximate modified coarse space, the two-level decomposition can be rewritten in the following perturbation form

$$\mathbf{u} = (I - \pi_D)\mathbf{v} + (I - \widetilde{\pi}_f)\pi_D \mathbf{u} + (\widetilde{\pi}_f - \pi_f)\pi_D \mathbf{u}. \tag{4.5}$$

Obviously, we do not have A-orthogonality anymore. However, as we show later, the first two terms of the decomposition (4.5) are approximately A-orthogonal whereas the last term can be made small, which leads to the desired error estimates.

4.2. Approximate Orthogonality. To show that the first two terms of the decomposition (4.5) are approximately A-orthogonal, we prove that the two spaces $\mathrm{Range}(P_\perp)$ $(= \mathrm{Range}(I - \pi_D))$ and $\mathrm{Range}(\widetilde{P})$ $(= \mathrm{Range}((I - \widetilde{\pi}_f)\pi_D))$ are approximately A-orthogonal. To this end, we first establish some properties of π_D and $\widetilde{\pi}_f$ summarized in the following lemma.

LEMMA 4.2. *We have $\pi_D \widetilde{\pi}_f = 0$ and that $(I - \widetilde{\pi}_f)\pi_D$ is a projection.*

Proof. The proof is the same as the proof of Lemma 3.1. ∎

Remark 4.3. Lemma 4.2 actually holds for $\widetilde{\pi}_f$ obtained by approximating A_f^{-1} with any $\widetilde{A}_f^{-1}$ in the definition of π_f. Therefore, this allows us to use, for example, other polynomials, i.e., not only the SA polynomial (4.1).

Next, we estimate the cosine of the abstract angle between the two spaces. For any vectors $\mathbf{v}_f$ and $\mathbf{v}_c$ and use the property (4.2) of the SA polynomial (4.1), we have

$$
\begin{aligned}
\mathbf{v}_f^T P_\perp^T A \widetilde{P} \mathbf{v}_c &= \mathbf{v}_f^T P_\perp^T A \left(I - P_\perp \widetilde{A}_f^{-1} P_\perp^T A \right) P \mathbf{v}_c \\
&= \left((I - A P_\perp \widetilde{A}_f^{-1} P_\perp^T) A P_\perp \mathbf{v}_f \right)^T P \mathbf{v}_c \\
&= \left(A P_\perp (I - \widetilde{A}_f^{-1} P_\perp^T A P_\perp) \mathbf{v}_f \right)^T P \mathbf{v}_c \\
&= \left(P_\perp (I - \widetilde{A}_f^{-1} A_f) \mathbf{v}_f \right)^T A P \mathbf{v}_c \\
&= \left(P_\perp p_\nu(A_f) \mathbf{v}_f \right)^T A P \mathbf{v}_c \\
&\le \sqrt{\mathbf{v}_f^T A_f p_\nu^2(A_f) \mathbf{v}_f} \sqrt{\mathbf{v}_c^T P^T A P \mathbf{v}_c} \\
&\le \max_{t \in (0,1]} \sqrt{t} |p_\nu(t)| \, \|\mathbf{v}_f\| \|P \mathbf{v}_c\|_A \\
&\le \tfrac{1}{2\nu+1} \|\mathbf{v}_f\| \|P \mathbf{v}_c\|_A.
\end{aligned}
\tag{4.6}
$$

Given $\mathbf{w}$ and $\mathbf{v}$, consider $P\mathbf{v}_c = \pi_D \mathbf{w}$ and $P_\perp \mathbf{v}_f = (I - \pi_D)\mathbf{v}$. Then, from (4.6) and use the facts that $\|\mathbf{v}_f\| = \|(I - \pi_D)\mathbf{v}\|_D$, $\|P\mathbf{v}_c\|_A = \|\pi_D \mathbf{w}\|_A$, and $\widetilde{P}\mathbf{v}_c = (I - \widetilde{\pi}_f)\pi_D \mathbf{w}$, to obtain

$$
((I - \pi_D)\mathbf{v})^T A (I - \widetilde{\pi}_f) \pi_D \mathbf{w} \le \frac{1}{2\nu + 1} \|(I - \pi_D)\mathbf{v}\|_D \|\pi_D \mathbf{w}\|_A. \tag{4.7}
$$

From (2.1), the WAP (2.4), we have

$$
\|(I - \pi_D)\mathbf{v}\|_A \le \|(I - \pi_D)\mathbf{v}\|_D \le \eta_w \|\mathbf{v}\|_A \text{ and}
$$

hence by Kato's Lemma ([28]),

$$
\|\pi_D\|_A = \|I - \pi_D\|_A \le \eta_w. \tag{4.8}
$$

The latter estimates together with (4.7) imply,

$$
((I - \pi_D)\mathbf{v})^T A (I - \widetilde{\pi}_f) \pi_D \mathbf{w} \le \frac{\eta_w^2}{2\nu + 1} \|\mathbf{v}\|_A \|\mathbf{w}\|_A. \tag{4.9}
$$

This gives us the desired approximate A-orthogonality result stated below.

THEOREM 4.4. *Assume the SA polynomial (4.1) is used to define $\widetilde{\pi}_f$, then the approximate modified coarse space* $\mathrm{Range}(\widetilde{P})$ *(=* $\mathrm{Range}((I - \widetilde{\pi}_f)\pi_D)$*) and the hierarchical complement* $\mathrm{Range}(P_\perp)$ *(=* $\mathrm{Range}(I - \pi_D)$*) of the original coarse space* $\mathrm{Range}(P)$ *are almost A-orthogonal in the following sense,*

$$
((I - \pi_D)\mathbf{v})^T A (I - \widetilde{\pi}_f) \pi_D \mathbf{w} \le \frac{\eta_w^2}{2\nu + 1} \|(I - \pi_D)\mathbf{v}\|_A \|(I - \widetilde{\pi}_f)\pi_D \mathbf{w}\|_A. \tag{4.10}
$$

Proof. Apply (4.9) for $\mathbf{v} := (I - \pi_D)\mathbf{v}$ and $\mathbf{w} := (I - \widetilde{\pi}_f)\pi_D \mathbf{w}$ and use the facts that both π_D and $(I - \widetilde{\pi}_f)\pi_D$ are projections. ∎

4.3. Energy Error Estimate. The second result we prove is an energy error estimate using the approximate modified coarse space $\mathrm{Range}(\widetilde{P})$. We start with the following lemma which shows that the third term in the perturbed decomposition (4.5) is small.

LEMMA 4.5. *Assume the SA polynomial (4.1) is used to define $\widetilde{\pi}_f$, then we have*

$$
\|(\widetilde{\pi}_f - \pi_f)\pi_D \mathbf{u}\|_A \le \frac{\eta_w^2}{2\nu + 1} \|\mathbf{u}\|_A \tag{4.11}
$$

Proof. Let $\pi_D \mathbf{u} = P\mathbf{u}_c$, and consider the deviation term

$$
\begin{aligned}
\|(\widetilde{\pi}_f - \pi_f)\pi_D \mathbf{u}\|_A &= \|P_\perp(\widetilde{A}_f^{-1} - A_f^{-1})P_\perp^T A P \mathbf{u}_c\|_A \\
&= \|A_f^{\frac{1}{2}}\left((I - p_\nu(A_f))A_f^{-1} - A_f^{-1}\right) P_\perp^T A P \mathbf{u}_c\| \\
&= \|p_\nu(A_f)A_f^{-\frac{1}{2}} P_\perp^T A P \mathbf{u}_c\| \\
&\le \|p_\nu(A_f)A_f^{-\frac{1}{2}} P_\perp^T A^{\frac{1}{2}}\| \|P\mathbf{u}_c\|_A \\
&= \|A^{\frac{1}{2}} P_\perp A_f^{-\frac{1}{2}} p_\nu(A_f)\| \|P\mathbf{u}_c\|_A \\
&= \|p_\nu(A_f)\| \|P\mathbf{u}_c\|_A.
\end{aligned} \tag{4.12}
$$

For the SA polynomial (4.1), using the fact that $\lambda_{\min}(A_f) \ge \frac{1}{\eta_w^2}$ (see (2.8)) and $\|\pi_D\|_A \le \eta_w$, (4.8), we have

$$
\begin{aligned}
\|(\widetilde{\pi}_f - \pi_f)\pi_D \mathbf{u}\|_A &\le \frac{1}{\sqrt{\lambda_{\min}(A_f)}} \max_{t\in[0,1]} \sqrt{t}|p_\nu(t)| \; \|P\mathbf{u}_c\|_A \\
&\le \frac{\eta_w}{2\nu+1} \; \|P\mathbf{u}_c\|_A \\
&= \frac{\eta_w}{2\nu+1} \; \|\pi_D \mathbf{u}\|_A \\
&\le \frac{\eta_w^2}{2\nu+1} \; \|\mathbf{u}\|_A,
\end{aligned}
$$

which completes the proof. $\square$

Consider the modified coarse problem based on the approximate inverse $\widetilde{A}_f^{-1}$ in $\widetilde{P}$, as follows

$$
\widetilde{P}^T A \widetilde{P} \widetilde{\mathbf{u}}_c = \widetilde{P}^T \mathbf{f}.
$$

Let $\widetilde{\mathbf{u}} = \widetilde{P}\widetilde{\mathbf{u}}_c \in \mathrm{Range}(\widetilde{P})$ be the respective coarse (upscaled) solution. We have the following energy error estimate which is an extension of energy error estimate (3.6).

THEOREM 4.6. *If p_ν is the SA polynomial* (4.1)*, then the following energy error estimate holds*

$$
\|\mathbf{u} - \widetilde{P}\widetilde{\mathbf{u}}_c\|_A \le \|\mathbf{e}\|_A + \|(\widetilde{\pi}_f - \pi_f)\pi_D \mathbf{u}\|_A \le \eta_w \|D^{-\frac{1}{2}} A \mathbf{u}\| + \frac{\eta_w^2}{2\nu + 1} \; \|\mathbf{u}\|_A, \tag{4.13}
$$

with the perturbation term (last term on the right hand side) exhibiting linear decay in ν.

Proof. Since the coarse solution is the best approximation to the solution $\mathbf{u}$ of the original linear system (3.4) from the modified coarse space in the A-norm, we have

$$
\|\mathbf{u} - \widetilde{P}\widetilde{\mathbf{u}}_c\|_A = \min_{\mathbf{v}_c} \; \|\mathbf{u} - \widetilde{P}\mathbf{v}_c\|_A.
$$

Note that $(I - \widetilde{\pi}_f)\pi_D \mathbf{u} \in \mathrm{Range}(\widetilde{P})$, then we have

$$
\|\mathbf{u} - \widetilde{P}\widetilde{\mathbf{u}}_c\|_A = \min_{\mathbf{v}_c} \; \|\mathbf{u} - \widetilde{P}\mathbf{v}_c\|_A \le \|\mathbf{u} - (I - \widetilde{\pi}_f)\pi_D \mathbf{u}\|_A.
$$

Hence, according to the decomposition (4.5) and $\mathbf{e} = (I - \pi_D)\mathbf{v}$ in (3.7), we have

$$
\|\mathbf{u} - \widetilde{P}\widetilde{\mathbf{u}}_c\|_A \le \|\mathbf{u} - (I - \widetilde{\pi}_f)\pi_D \mathbf{u}\|_A = \|\mathbf{e} + (\widetilde{\pi}_f - \pi_f)\pi_D \mathbf{u}\|_A \le \|\mathbf{e}\|_A + \|(\widetilde{\pi}_f - \pi_f)\pi_D \mathbf{u}\|_A.
$$

Then apply Theorem 3.3 to the first term and Lemma 4.5 to the second term, to arrive at (4.13). $\square$

Further, assume that D is well-conditioned, we have the following approximate the SAP, which is a perturbation of Corollary (3.4).

COROLLARY 4.7. *If p_ν is the SA polynomial (4.1), then*

$$\|A\|^{\frac{1}{2}}\|\mathbf{u}-\widetilde{P}\widetilde{\mathbf{u}}_c\|_A \le \|D\|^{\frac{1}{2}}\|\mathbf{u}-\widetilde{P}\widetilde{\mathbf{u}}_c\|_A \le \eta_s^{\frac{1}{2}}\|A\mathbf{u}\| + \frac{\eta_w\|D\|^{\frac{1}{2}}}{2\nu+1}\|\mathbf{u}\|_A.$$

where $\eta_s \le \|D\|\|D^{-1}\|\eta_w^2$. If D is well-conditioned, then η_s is bounded above by a constant.

4.4. Other Approximations. The SA polynomial (4.1) is just one possible choice for approximating A_f^{-1}. There are other possible choices as well. In this subsection, we briefly discuss other possibilities.

If we have the WAP constant η_w explicitly available, that is, we have explicit eigenvalue bounds, $\lambda(A_f) \in [\alpha,\ \beta] \subset [\frac{1}{\eta_w^2},\ 1]$, we can use the (best) Chebyshev polynomial

$$p_\nu(t) = \frac{T_\nu\left(\frac{\beta+\alpha-2t}{\beta-\alpha}\right)}{T_\nu\left(\frac{\beta+\alpha}{\beta-\alpha}\right)}. \tag{4.14}$$

Then due to the optimality property of Chebyshev polynomial,

$$\|p_\nu(A_f)\| \le \frac{2q^\nu}{1+q^{2\nu}}, \quad q = \frac{\eta_w - 1}{\eta_w + 1},$$

together with the identity (4.12), we end up with the following error estimate.

THEOREM 4.8. *If p_ν is the Chebyshev polynomial (4.14) used to define the approximate modified coarse space Range$(\widetilde{P})$, then the following energy error estimate holds*

$$\|\mathbf{u}-\widetilde{P}\widetilde{\mathbf{u}}_c\|_A \le \eta_w\|D^{-\frac{1}{2}}A\mathbf{u}\| + \frac{2q^\nu\eta_w}{1+q^{2\nu}}\ \|\mathbf{u}\|_A, \tag{4.15}$$

where now the perturbation term exhibiting geometric decay in ν.

It is clear that error estimate (4.15) is much better than (4.13). We note that in the spectral AMGe method in the form presented in [5], explicit bounds of η_w are available. Therefore, the Chebyshev polynomial (4.14) can be used to modify the coarse space in the spectral AMGe setting.

Using either the SA polynomial (4.1) or the Chebyshev polynomial (4.14) basically provides an approximate solution to the linear system (2.7). Therefore, another way to solve (2.7) is via nonlinear iterative methods such as the conjugate gradient (CG) method. Using CG implicitly constructs a polynomial $p_\nu(t)$ which defines $\widetilde{\pi}_f$. The convergence analysis of CG can be used to estimate $\|(\widetilde{\pi}_f - \pi_f)\pi_D\mathbf{u}\|_A$. Denote the ν-th iteration of CG for solving $A_f\overline{\mathbf{u}}_c = P_\perp^T AP\mathbf{u}_c$ by $\overline{\mathbf{u}}_c^\nu$ with zero initial guess, then similarly to (4.12), we have

$$\begin{aligned}
\|(\widetilde{\pi}_f - \pi_f)\pi_D\mathbf{u}\|_A &= \|P_\perp(\widetilde{A}_f^{-1} - A_f^{-1})P_\perp^T AP\mathbf{u}_c\|_A = \|(\widetilde{A}_f^{-1} - A_f^{-1})P_\perp^T AP\mathbf{u}_c\|_{A_f} \\
&= \|\overline{\mathbf{u}}_c^\nu - \overline{\mathbf{u}}_c\|_{A_f} \le 2q^\nu\|\overline{\mathbf{u}}_c\|_{A_f} = 2q^\nu\|A_f^{-1}P_\perp^T AP\mathbf{u}_c\|_{A_f} \\
&\le 2q^\nu\|A_f^{-\frac{1}{2}}P_\perp^T A^{\frac{1}{2}}\|\|P\mathbf{u}_c\|_A = 2q^\nu\|A^{\frac{1}{2}}P_\perp A_f^{-\frac{1}{2}}\|\|P\mathbf{u}_c\|_A \\
&= 2q^\nu\|P\mathbf{u}_c\|_A \le 2q^\nu\eta_w\|\mathbf{u}\|_A.
\end{aligned}$$

Therefore, we have the following result.

THEOREM 4.9. *If p_ν is the polynomial generated by CG, then the following energy error estimate holds*

$$\|\mathbf{u} - \widetilde{P}\widetilde{\mathbf{u}}_c\|_A \leq \eta_w \|D^{-\frac{1}{2}} A\mathbf{u}\| + 2q^\nu \eta_w \, \|\mathbf{u}\|_A, \tag{4.16}$$

with the perturbation term exhibiting geometric decay in ν.

Remark 4.10. The error estimates (4.15) and (4.16) both have perturbation terms that decay geometrically with the same rate q, therefore, we can conclude that modifying the coarse space based on CG polynomial gives better estimates than the SA polynomial. Note that the CG approximation also, as in the SA case, does not need estimates for the spectrum of A_f, whereas these are needed in the Chebyshev polynomial case.

4.5. Example: Linear Finite Elements for Laplace Equation. As a simple example, we consider the Laplace equation, $-\Delta u = f$, discretized using piecewise linear finite elements. In this case, we have $\eta_w \simeq \frac{H}{h}$ (cf., [5]) where H stands for the diameter of the aggregates. This fact, combined with a simple argument relating the right hand side of the discrete problem, $\mathbf{f}$, and the L_2-norm $\|f\|_0$ of the right hand side function f (as shown in [27]), we conclude that the first term $\eta_w \|D^{-\frac{1}{2}} A\mathbf{u}\| \simeq H\|f\|_0$. If we want to balance the second term with the first one, we need to choose $2q^\nu \simeq H$ (assume Chebyshev polynomial or CG used). This implies that

$$\nu \log\left(1 + \frac{2}{\eta_w - 1}\right) \simeq \log \frac{1}{H},$$

and since $\log\left(1 + \frac{2}{\eta_w - 1}\right) \simeq \frac{2}{\eta_w - 1} \simeq \frac{h}{H}$, we have the following estimate for the polynomial degree (or the number of iterations used for CG)

$$\nu \simeq \frac{H}{h} \log \frac{1}{H}.$$

This ensures the error estimate,

$$\|\mathbf{u} - \widetilde{P}\widetilde{\mathbf{u}}_c\|_A \leq C \left(H\|f\|_0 + H\|\mathbf{u}\|_A\right).$$

Similar argument can also be applied to the SA polynomial case in order to get an estimate of the polynomial degree.

5. Remarks for Elliptic Problems with High Contrast Coefficients. We consider the case with exact projection π_f for simplicity in this section. In section 3, we showed that the second component of the two-level A-orthogonal decomposition

$$\mathbf{u} = (I - \pi_D)\mathbf{v} + (I - \pi_f)\pi_D \mathbf{u},$$

is actually the solution $\overline{\mathbf{u}}_c$ of the modified coarse problem (3.5). It is worth noticing the the first component above, $(I - \pi_D)\mathbf{v}$, is the A-orthogonal projection of $\mathbf{u}$ onto the space Range$(I - \pi_D)$. We already discussed the fact that the matrix of this problem is sparse and well-conditioned (after symmetric diagonal scaling of A). Thus it is computationally feasible to explicitly compute this component as well. Of course, this is not surprising since a two-grid AMG with the standard coarse space Range(P) and using D as a smoother is uniformly convergent, hence $\mathbf{u}$ can be approximated well by a few V-cycles. Note that such an AMG uses only sparse matrix-operations with much sparser matrices than the one of the upscaled problem (3.5) and A_f. Therefore, introducing the modified coarse space $\mathrm{Range}((I-\pi_f)\pi_D)$ and the resulting error estimate (3.6) (and its corollaries) are mostly of theoretical value. In the case of approximate projections, if we cannot control the sparsity of the coarse matrices so that the resulting method requires much less memory and computational cost than the original matrix A, then the upscaled problem is mostly of theoretical value only. With our numerical tests we demonstrate that in the PDE case, careful choice of the polynomial degree can lead to some savings in practice for the upscaled problems. The situation for graph Laplacian matrices is more challenging for graphs with irregular degree distribution.

One possible practical application of the presented method is the diffusion equation,

$$\begin{cases} -\operatorname{div}(\kappa\nabla u) = f, & \text{in}\,\Omega, \\ u = 0, & \text{on}\,\partial\Omega, \end{cases} \tag{5.1}$$

where $\Omega \subset \mathbb{R}^d$, $d = 2$ or $d = 3$, is a polygonal/polyhedral domain. Using H^1-conforming finite element space on a quasiuniform mesh $\mathcal{T}_h$, we end up with linear system of the form

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

By construction, we have $\|\mathbf{f}\| \simeq h^{\frac{d}{2}}\|f\|_0$. Let D be the diagonal of A. We have $D \simeq \mathrm{diag}(h^{d-2}\kappa_i)$, where $\kappa_i h^d$ are the diagonal entries of the weighted mass matrix corresponding to the κ-weighted L_2-bilinear form. Hence, we have the estimate

$$\|D^{-\frac{1}{2}}\mathbf{f}\| \le \eta_b\, h\|f\|_{0,\,\kappa^{-1}}$$

for a uniform constant η_b, which leads to the error estimate

$$\|u_h - u_H\|_{1,\,k} \le \eta_w\eta_b\, h\|f\|_{0,\,\kappa^{-1}}.$$

Here u_h is the finite element solution of the fine-grid problem and u_H is the finite element solution corresponding to the upscaled solution $\overline{\mathbf{u}}_c$ of (3.5). Note that, this error estimate is independent of the coefficient κ with the expense of the weighted norms involved. For $\kappa \simeq 1$, using the fact that $\eta_w \simeq H/h$, the last error estimate reads $\|u_h - u_H\|_1 \le CH\|f\|_0$ which is an analog to the one in [22].

6. Numerical Experiments.

In this section, we present numerical results illustrating the theory demonstrating the approximation properties of the modified coarse spaces. In all experiments, we use the AMGe method in the form proposed in [5, 27] to construct the original coarse space $\mathrm{Range}(P)$ so that it satisfies the WAP. More presicely, we use a greedy type algorithm to construct a set of aggregates and solve a generalized eigenvalue problem (see (13) in [5]) to construct the tentative prolongation as shown in (2.3). To assess the quality of the proposed approach in practice, we only consider two-grid method and the modified coarse space based on the polynomial approximations as discussed in Section 4. In fact, we use the CG polynomial in all our experiments as it gives the best error estimates (see Remark 4.10). The tests are run in Matlab using an AMG package developed by the authors.

Example 6.1. *Consider the diffusion problem* (5.1) *posed on* $\Omega = [0,1] \times [0,1]$ *with*

$$\kappa = \begin{cases} \epsilon, & \text{in } [0.25, 0.5] \times [0.25, 0.5] \cap [0.5, 0.75] \times [0.5, 0.75] \\ 1, & \text{otherwise.} \end{cases}$$

Our first example is diffusion problem (5.1) with discontinuous coefficient. As discussed in Section 5, the modified coarse space provides error estimates that are independent of the jumps. The results shown in Figure 1 support this theoretical results. Here, the fine level problems are all of size $4,225 \times 4,225$ on a uniform triangular mesh with $h = \frac{1}{64}$ and the coarse level matrices are all of size 302×302. We change the contrast of the diffusion coefficient, i.e., ϵ, and report how the SAP constant η_s changes with respect to the degree ν of the polynomial (since we use CG, the degree is equivalent to the number of iterations). For comparison, we also report the SAP constants when we modify the coarse space exactly by directly inverting A_f. As clearly seen, the SAP constant stays almost the same for different choices of ϵ for a fixed ν, and is indeed practically independent of the contrast ϵ. This is consistent with the theory and shows that the modified coarse spaces provide approximations in the energy norm that are robust with respect to the jumps. From Figure 1, we also observe that the SAP constant decreases to the SAP constant that corresponds to the modified coarse space with exact inverse, when ν increases with a rate that is almost the same for different ϵ. This is also consistent with the theoretical results presented in Section 4; namely, that the decay rate should depend on η_w which, in fact, in the present case depends on $\frac{H}{h}$. Our next numerical experiment further verifies this property; see the results shown in Figure 2. Since h is $N^{-1/2}$ and H is roughly $N_c^{-1/2}$, we present the results in terms of the ratio $\frac{N}{N_c}$, which is roughly $\left(\frac{H}{h}\right)^2$. More specifically, from Figure 2, we see that the SAP constant decreases when ν increases and the bigger the ratio $\frac{N}{N_c}$ is, the slower the decay rate is. But, the SAP constant converges to the SAP constant corresponding to the modified coarse space with exact A_f inverse, as expected.

The next test illustrates the properties of the coarse matrices corresponding to the modified coarse spaces based on polynomial approximation. More specifically, we are interested in the sparsity of the modified prolongation matrix $\widetilde{P}$ (in terms of percentage w.r.t to the matrix size NN_c). We also are interested in the AMG operator complexity (OC) defined as the ratio between the total number of nonzeros of A plus the number of the nonzeros of the coarse-level matrix and the number of nonzeros of A. Note that $\nu = 0$ corresponds to the original prolongation P (and respective coarse matrix). From Table 1, as expected, we see that both the number of nonzeros and operator complexity grow when ν increases. The number of nonzeros of $\widetilde{P}$ grows faster when the ratio $\frac{N}{N_c}$ gets bigger whereas the operator complexity actually grows slower when $\frac{N}{N_c}$ gets larger. We note that in practice, for upscaling purposes, we need to have operator complexity less than two (then we use less memory to store the coarse matrix than the original fine-level one). Our results indicate that to achieve desired approximation accuracy for a reasonable computational cost can be a challenging task. In addition, we also use the modified coarse space in AMG iterative method and report number of iterations of the two-grid algorithms. Here, we choose $f = 1$ in the diffusion problem (5.1). In the two-grid algorithm, Gauss-Seidel relaxation is used, with zero initial guess and the stopping criterion is achieving a reduction of the ℓ_2 norm of the relative residual by 10^{-6}. As expected, the number of iterations

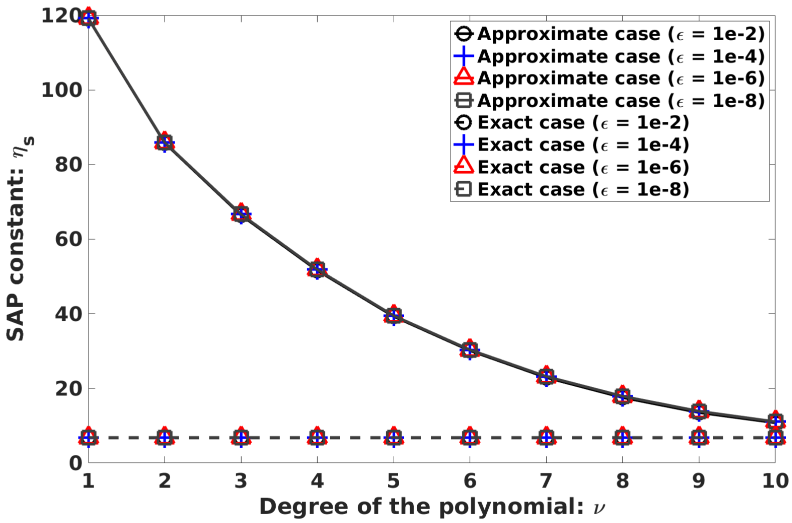

FIG. 1. *Example 6.1: the SAP constants for different ϵ ($h = 1/64$, $N = 4,225$ and $N_c = 302$)*

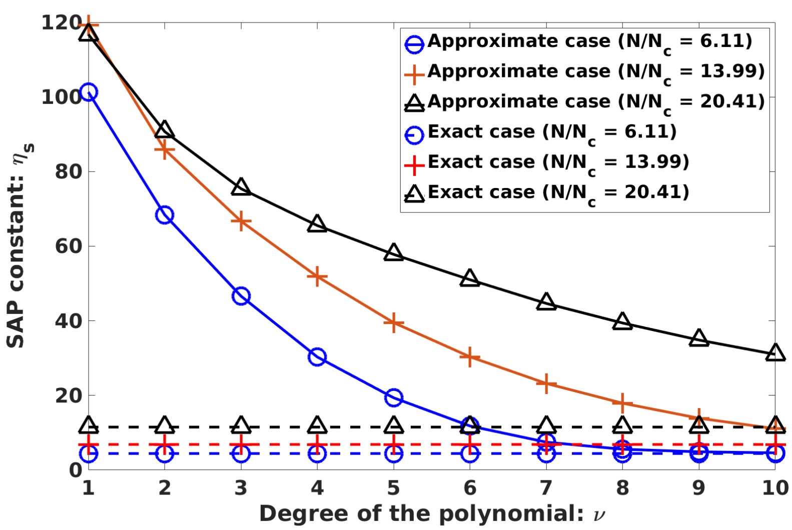

FIG. 2. *Example 6.1: the SAP constants for different N_c ($h = 1/64$, $N = 4,225$ and $\epsilon = 10^{-4}$)*

(Iter) decreases as ν increases. We note that in practice for solving linear systems, we need to consider the trade-off between the computational complexity and convergence behavior. The latter can also be a challenge in practice.

EXAMPLE 6.2. *To stress upon the fact that our approach is in fact purely algebraic, we apply our results to graph Laplacian systems corresponding to graphs listed in Table 2.*

In Figure 3, we present the SAP constants for the different graphs from Table 2. Here, we use a simple unsmoothed aggregation approach. In order to achieve aggres-

TABLE 1

Example 6.1: sparsity of the modified coarse space and performance of two-grid AMG method with different ν ($h = 1/64$, $N = 4,225$ and $\epsilon = 10^{-4}$)

	$N/N_c = 6.11$			$N/N_c = 13.99$			$N/N_c = 20.41$		
	nnz of $\widetilde{P}$	OC	Iter	nnz of $\widetilde{P}$	OC	Iter	nnz of $\widetilde{P}$	OC	Iter
$\nu = 0$	0.15%	1.22	43	0.43%	1.13	52	1.2%	1.15	60
$\nu = 1$	0.92%	2.19	30	2.75%	1.63	42	6.97%	1.65	55
$\nu = 2$	2.52%	3.92	23	7.27%	2.41	38	17.55%	2.24	50
$\nu = 3$	4.92%	6.62	19	13.78%	3.29	34	31.35%	2.73	48
$\nu = 4$	8.14%	8.89	16	21.87%	4.10	30	45.87%	2.98	46
$\nu = 5$	12.07%	11.71	14	30.96%	4.74	28	60.04%	3.10	41

TABLE 2

A set of networks from different real-world applications (first three graphs are from Stanford Large Network Dataset Collection [19] and the last graph is from SuiteSparse Matrix Collection [7]). For each graph, we show its number of vertices, number of edges, average vertex degree (ave. deg.) and maximal vertex degree (max. deg.)

	Vertices	Edges	ave. deg.	max. deg.	Description
bitcoin-alpha	3,775	14,120	7.48	510	Bitcoin Alpha web of trust network
ego-facebook	4,039	88,234	43.69	1045	Social circles from Facebook
ca-GrQc	4,158	13,425	6.46	81	Collaboration network of Arxiv
rw5151	5,151	15,248	5.92	7	Markov chain modeling

sive coarsening, the aggregates are built based on the sparsity pattern of L^2, where L corresponds to the graph Laplacian. The original coarse space (or respective interpolation matrix P) is constructed using the spectral AMGe method (as used in [12]). As we can see, although the ratio $\frac{N}{N_c}$ differs for the different graphs, if we use relatively accurate approximation (i.e. relatively large ν), the SAP constant stays small and is fairly similar for different graphs. This demonstrate that the modified coarse spaces are also robust for these real-world graphs.

In Figure 4 and 5, we illustrate the sparsity of the modified prolongations and respective coarse matrices. We notice that the nonzeros percentage of $\widetilde{P}$ grows fairly quickly, which suggests that in practice, only small ν makes sense. If the coarse level problem are meant to be used multiple times, due to reasonable operator complexity and good approximation property achieved by large ν, we could use more accurate approximated modified coarse spaces coming from relatively large ν. For graphs with irregular degree distribution, the challenge to maintain reasonable sparsity of the coarse matrices with good approximation properties is much more pronounced than in the discretized PDE case and it requires more specialized study.

7. Conclusions. In this paper, we investigate the use of certain AMG coarse spaces for the purpose of dimension reduction which in the present setting is referred to as numerical upscaling. As it is well-understood that although the traditional AMG coarse spaces do satisfy the WAP (weak approximation property), it is not sufficient for the purpose of upscaling because the coarse-level solutions do not necessarily approximate the fine-level solution with guaranteed accuracy. To remedy this, we follow the approach developed in [22] extending it to the presented AMG setting. The method exploits a projection π_f used to modify the original coarse space, which is assumed to possess a WAP, so that the resulting new, modified, coarse space satisfies a SAP (strong approximation property) with provable satisfactory bound on the resulting constant η_s. More specifically, the modified coarse space is one of the components

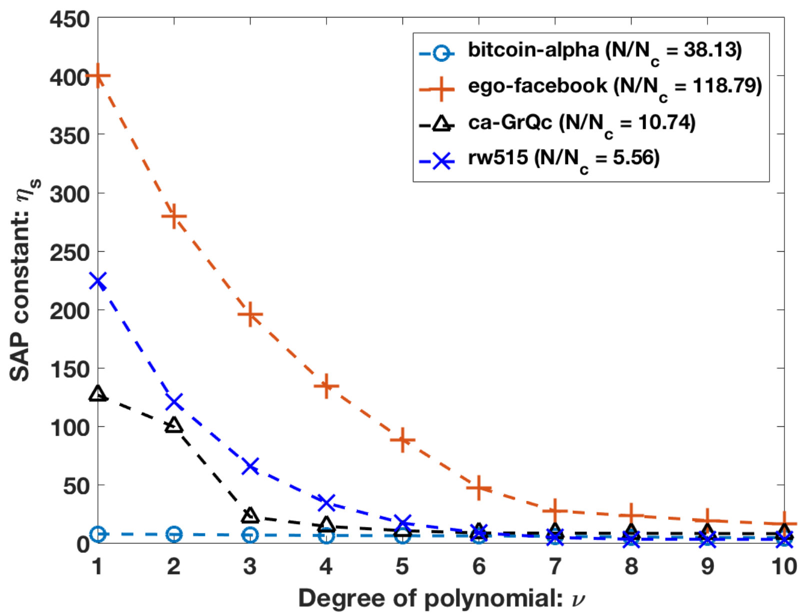

FIG. 3. *Example* 6.2*: the SAP constants for different* ν

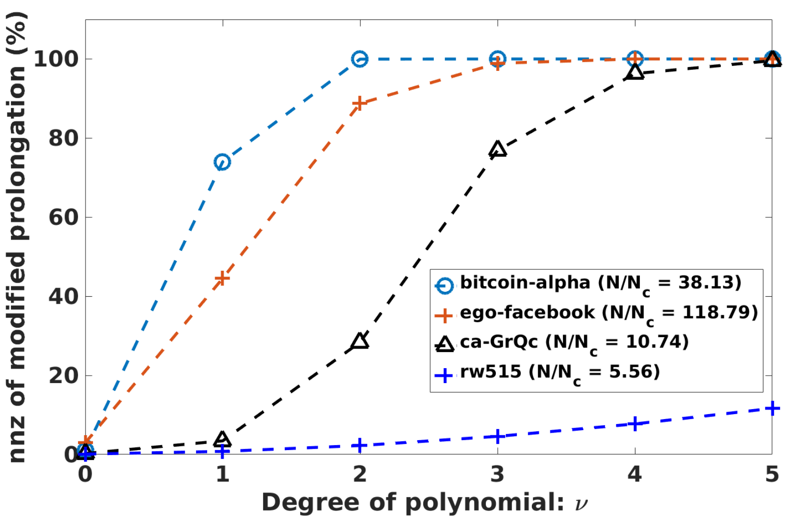

FIG. 4. *Example* 6.2*: number of nonzeros of* $\widetilde{P}$ *(in percentage) for different* ν

in a two-level A-orthogonal decomposition so that the corresponding coarse-level solution gives accurate approximation in energy norm. One main challenge with this approach is the fact that the matrix A_f^{-1} used in the definition of π_f, is dense even if A_f is sparse. Thus, modifying the original coarse space with exact π_f is computationally infeasible (for large-scale problems). In order to make such modification more practical, we use the fact (which we prove) that A_f is well-conditioned, allow-

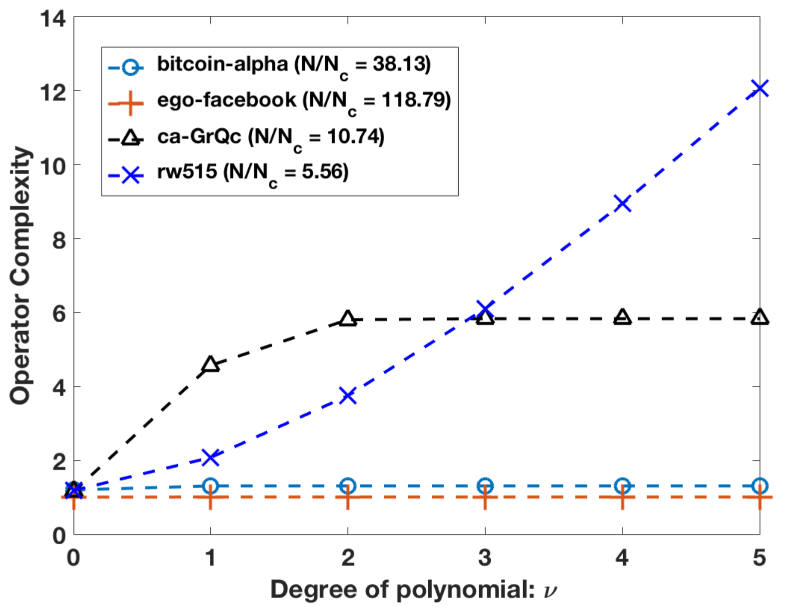

FIG. 5. *Example 6.2: operator complexity for different ν*

ing the use of polynomials to approximate its inverse, leading to an approximate π_f, which is used to define an approximate modified coarse space. Such approximation is computational feasible and also provides provable error estimates in energy norm. Moreover, the error estimates improve when increasing the degree of the polynomial used in the approximation.

We provide numerical results that illustrate the theory and demonstrate the accuracy and sparsity of the coarse problems coming from the approximately modified coarse space. The tests include both, examples of diffusion equation with high contrast coefficients as well as graph Laplacian matrices corresponding to some real-life applications.

As discussed, the use of such modified coarse spaces is of interest in dimension reduction which, as our model tests demonstrate, can be challenging for the present approach (in terms of maintaining reasonable sparsity of the coarse matrices). In the PDE case this challenge seems resolvable if large enough coarsening factor (H/h) is employed, whereas in the graph application for graphs with irregular degree distribution, in addition to high coarsening factor one may need to employ graph disaggregation (cf., [16]), which is left for a possible future study. Additionally, in the PDE case, it is of interest to extend the present results to other types of PDEs such as ones posed in $H(\text{curl})$ and $H(\text{div})$, which will provide alternatives to the existing AMGe upscaling methods (cf., [17], [13], and [1]).

REFERENCES

[1] A. Barker, C. S. Lee, and P. S. Vassilevski, "*Spectral Upscaling for Graph Laplacian Problems with Application to Reservoir Simulation,*" SIAM Journal on Scientific Computing **39**(5)(2017), pp. S323-S346.

[2] A. Brandt, S. McCormick, and J. Ruge, "*Algbraic Multigrid (AMG) for Sparse Matrix Equa-*

tions," in Sparsity and Its Applications, Edited by David J. Evans, Cambridge University Press, Cambridge, 1985, pp. 257–284.

[3] 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**(1)(2003), pp. 1-26.

[4] M. Brezina, R. Falgout, S. MacLachlan, T. Manteuffel, S. McCormick, and J. Ruge, "*Adaptive smoothed aggregation* (αSA) *multigrid,*" SIAM Rev., **47**(2)(2005), pp. 317-346.

[5] M. Brezina and P. Vassilevski, "*Smoothed aggregation spectral element agglomeration AMG: $SA-\rho AMGe$,*" in Large-Scale Scientific Computing, 8th International Conference, LSSC 2011, Sozopol, Bulgaria, June 6-10th, 2011. Revised Selected Papers. Lecture Notes in Computer Science, vol. 7116, Springer, 2012, pp. 3–15.

[6] M. Brezina, P. Vaněk, and P. S. Vassilevski, "*An Improved Convergence Analysis of Smoothed Aggregation Algebraic Multigrid,*" Numerical Linear Algebra with Applications **19**(3)(2012), pp. 441-469. (published online: 2 MAR 2011, DOI: 10.1002/nla.775).

[7] T. A. Davis and Y. Hu, "*The university of Florida sparse matrix collection,*" ACM Transactions on Mathematical Software, 38(1), 2011, pp. 1-25.

[8] P. D'Ambra and P. S. Vassilevski, "*Adaptive AMG with Coarsening Based on Compatible Weighted Matching,*" Computing and Visualization in Science 16 (2013), pp. 59-76.

[9] S. Demko, W.F. Moss, and P.W. Smith, "*Decay rates of inverses of band matrices*", Mathematics of Computation **43**(168)(1984), pp. 491-499.

[10] R. Falgout, P. S. Vassilevski, and L. T. Zikatanov, "*On Two-grid Convergence Estimates,*" Numerical Linear Algebra with Applications, 12(5-6), 2005, pp. 471-494.

[11] T. Grauschopf, M. Griebel, and H. Regler, "*Additive multilevel preconditioners based on bilinear interpolation, matrix-dependent geometric coarsening and algebraic multigrid coarsening for second-order elliptic PDEs*," Applied Numerical Mathematics, Multilevel Methods 23, 1997, pp. 6395

[12] X. Hu, P. S. Vassilevski, and J. Xu, "*A two-grid SA-AMG convergence bound that improves when increasing the polynomial degree,*" Numerical Linear Algebra with Applications **23**(4)(2016), pp. 746–771.

[13] D. Kalchev, C. S. Lee, U. Villa, Y. Efendiev, and P. S. Vassilevski, Upscaling of Mixed Finite Element Discretization Problems by the Spectral AMGe Method, SIAM Journal on Scientific Computing 38(5) (2016), pp. A2912-A2933.

[14] T. V. Kolev and P. S. Vassilevski, "*Parallel auxiliary space AMG for H(curl) problems*," Journal of Computational Mathematics **27**(2009), pp. 604–623.

[15] T. V. Kolev and P. S. Vassilevski, "*Parallel auxiliary space AMG for H(div) problems*,"S SIAM Journal on Scientific Computing **34**(2012), pp. A3079-A3098.

[16] V. Kuhlemann and P. S. Vassilevski, "*Improving the Communication Pattern in Mat-Vec Operations for Large Scale-free Graphs by Disaggregation,*" SIAM Journal on Scientific Computing **35**(5)(2013), pp. S465-S486.

[17] I. V. Lashuk and P. S. Vassilevski, "*The Construction of Coarse de Rham Complexes with Improved Approximation Properties,*" Computational Methods in Applied Mathematics **14**(2)(2014), pp. 257-303.

[18] J.V. Lent, R. Scheichl, I.G. Graham, "*Energy-minimizing coarse spaces for two-level Schwarz methods for multiscale PDEs*," Numerical Linear Algebra with Applications 16, 2009, pp. 775799.

[19] J. Leskovec and A. Krevl, "*SNAP Datasets: Stanford Large Network Dataset Collection,* http://snap.stanford.edu/data.

[20] Oren E. Livne and Achi Brandt, "*Lean Algebraic Multigrid (LAMG): Fast Graph Laplacian Linear Solver,*" SIAM Journal on Scientific Computing 34(4) (2012), pp. B499-B522

[21] S.P. MacLachlan and J.D. Moulton, "*Multilevel upscaling through variational coarsening*," Water Resources Research 42, 2006.

[22] A. Malqvist and D. Peterseim, "*Localization of elliptic multiscale problems,*" Math. Comp., **83**(290), pp. 2583-2603, 2014.

[23] J.D. Moulton, J.E. Dendy, and J.M. Hyman, "*The Black Box Multigrid Numerical Homogenization Algorithm.*," Journal of Computational Physics 142, 1998, pp. 80108.

[24] S. Reitzinger and J. Schöberl, "*An algebraic multigrid method for finite element discretizations with edge elements,*" Numerical Linear Algebra with Applications **9**(3) (2002), pp. 223-238.

[25] P. S. Vassilevski, "*On two ways of stabilizing the HB multilevel methods,*" SIAM Review 39(1997), 18–53.

[26] N. Spillane, V. Dolean, P. Hauret, F. Nataf, C. Pechstein, and R. Scheichl, "*Abstract robust coarse spaces for systems of PDEs via generalized eigenproblems in the overlaps*," Numer. Math. 126, 2014, pp. 741770.

[27] Panayot S. Vassilevski, "*Coarse Spaces by Algebraic Multigrid: Multigrid Convergence and Upscaling Error Estimates,*" Advances in Adaptive Data Analysis, **3**(1&2), pp. 229-249, 2011.

[28] Panayot S. Vassilevski, "Multilevel Block Factorization Preconditioners, Matrix-based Analysis and Algorithms for Solving Finite Element Equations," Springer, New York, 2008. 514 p.

[29] Jinchao Xu and Ludmil Zikatanov, "*Algebraic multigrid methods,*" Acta Numerica **26**(2017), pp. 591-721.