TWO-LEVEL METHOD FOR SOLIDS ON UNSTRUCTURED MESHES*

RADEK TEZAUR , PETR VANĚK AND MARIAN BREZINA†

December 1995

Abstract. We propose an abstract two-level algorithm with convergence independent of the coarse-space size. The abstract algorithm is applied to problems of solids discretized on unstructured meshes. With no regularity assumptions we prove uniform convergence with respect to coarse-space size, domain, essential boundary conditions and jumps in Young modulus. Numerical experiments confirm the theory and show that the method works well even if some assumptions of the theory are violated.

Key words. Convergence independent of the coarse space size, unstructured meshes, black-box solver, smoothed tentative prolongator, 3D elasticity.

AMS(MOS) subject classifications. 65F10,65N55

1. Introduction. The standard two-level method suffers from the dependence of the rate of convergence on the size of the coarse space. When solving an elliptic problem of order $2k$, assuming no regularity, the rate of convergence can be estimated by $1 - C(\frac{h}{H})^{2k}$. Here H, h denote the characteristic meshsizes on the coarse and fine level, respectively. Consequently, in order to achieve the optimal convergence result, one is limited to using a coarse space only a few times smaller than the fine level one.

Today's two-level domain decomposition preconditioners typically offer an improvement by yielding condition numbers bounded polylogarithmically in $\frac{H}{h}$. The direct local solvers used by these methods, however, result in large computational complexity, while a straightforward application of approximate subdomain solvers is problematic.

We suggest a method with coarse-space size independent convergence diminishing significantly these difficulties. Our algorithm is based on a concept of smoothed tentative prolongator introduced in [10] and further investigated in [12, 8]. Our tentative coarse space is derived from a system of nonoverlapping subdomains and supports rigid body modes in a local sense. Coarse-space basis has very small overlaps of supports; the tentative basis functions associated with two different subdomains are orthogonal in a discrete l^2-product. A special polynomial smoother is applied to the tentative prologator, which extends the supports but only so much that their intersections remain bounded. Our choice of the prolongator smoother, pre- and post-smoother enables the proof of coarse-space size independent convergence. The price in terms of computational complexity we have to pay for having a coarse-space of modest size is much smaller in our case than it is for standard domain decomposition methods with direct subdomain solvers. Denoting the average number of nodes per subdomain by N_d, the recommended degree of the prolongator smoothers is $\lfloor \frac{1}{2}(N_d^{1/3} - 1) \rfloor$. Consequently, the asymptotic

* This research was supported by NSF grants ASC-9121431 and ASC-9217394.

† Center for Computational Mathematics, University of Colorado at Denver, Denver, CO 80217-3364.

coputational cost associated with treating one subdomain is $O(N_d^{4/3})$. This contrasts with $O(N_d^{7/3})$ operations needed during the factorization of the local problem when using direct subdomain solvers.

We find it convenient to divide the analysis into two parts: an abstract theory that yields a coarse-space size independent convergence, provided that assumptions on the components of the method are satisfied, and verification of the assumptions for the case of 3D solids. Based on our experience with industrial models, it is desirable to design methods with the convergence independent of boundary conditions. In designing robust methods for solving elasticity problems, one must not rely solely on the coercivity of the bilinear form guaranteed by the boundary conditions. The reason for this is that the constant in the Korn's inequality is often very small. Moreover, arguments using the "global" coercivity are usually difficult to use in the case of nonscalar equations with jumps in coefficients. For these reasons, we employ a coarse space supporting local kernels of the bilinear form. The analysis of this coarse space can be carried out using local arguments on the factorspace modulo rigid body modes and the assumptions on the essential boundary conditions are no longer needed. Furthermore, based on empirical observations, we suggest to solve the problem with a matrix obtained by block-diagonal scaling of the original stiffness matrix. This requirement is not essential to the analysis and slightly complicates the theory, but it is of a considerable computational value. Qualitatively the same result can be obtained assuming only the diagonal scaling, but we deem the scaling by block diagonal more natural for treating nonscalar equations.

For problems with large jumps in Young modulus, the theory requires each subdomain to be a union of a bounded number of star-shaped domains with large overlaps. This is more restrictive than the shape-regularity requirement formulated using a locally Lipschitz one-to-one mapping in the scalar case. However, in the case of uniformly bounded coefficients such an assumption is not necessary.

The abstract theory presented here is based on the unpublished technical report [11]; extended abstract of this paper has been published in [13].

The paper is organized as follows. The abstract algorithm and analysis are presented in Section 2. Section 3 deals with the construction of the tentative coarse space suitable for real-life problems of solids discretized on unstructured meshes. Requirements on the system of subdomains for our construction of the tentative prolongator as well as theoretical tools needed for the application of the abstract theory are the subject of Section 4. Section 5 presents the convergence result independent of the coarse space size, domain, essential boundary conditions and jumps in Young modulus provided the subdomains are aligned with the jumps. The estimate of computational complexity $O(n^{7/6})$ for the case of the coarse space of the optimal size is also given here. Finally, numerical experiments of Section 6 confirm the theoretical results and show robustness of the method even if applied to problems not in the scope of the current theory.

2. Abstract framework. In this section we develop and analyze a simple two-level abstract method for the numerical solution of a system of linear algebraic equations

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

with a symmetric positive definite $n \times n$ matrix A. The method is determined by two components : a full-rank *tentative prolongator* $P : \mathbb{R}^m \to \mathbb{R}^n$ ($m << n$) and a symmetric *prolongator smoother* $M \in [\mathbb{R}^n]$ that commutes with A. The proposed method is, up to a post-processing step, the standard variational two-level scheme with a smoothed prolongator MP and special smoothing procedures derived from M. The construction of such smoothers is motivated by an attempt to reduce the error components which cannot be represented in the coarse-space.

Let us set

$$A_M = M^2 A, \quad M' = I - \frac{\omega}{\bar{\varrho}(A_M)} A_M, \tag{2}$$

where $\omega \in (0, 2)$ is a given parameter and $\bar{\varrho}(A_M)$ is an upper bound of the spectral radius $\varrho(A_M)$. We define the pre-smoother $\mathbf{x} \leftarrow \mathcal{S}_M(\mathbf{x}, \mathbf{b})$ and the post-smoother $\mathbf{x} \leftarrow \mathcal{S}_{M'}(\mathbf{x}, \mathbf{b})$ to be linear iterative methods consistent with (1) such that their error propagation operators are matrices M and M' respectively. Note that such methods can be easily constructed if M is a polynomial in A. The algorithm can be written down as follows:

ALGORITHM 2.1. *For the given initial guess* $\mathbf{x} \in \mathbb{R}^n$,

repeat

1. $\mathbf{x} \leftarrow \mathcal{S}_M(\mathbf{x}, \mathbf{b})$,
2. *solve* $P^T M A M P \mathbf{v} = P^T M (A\mathbf{x} - \mathbf{b})$,
3. $\mathbf{x} \leftarrow \mathbf{x} - MP\mathbf{v}$,
4. $\mathbf{x} \leftarrow \mathcal{S}_{M'}(\mathbf{x}, \mathbf{b})$

until convergence;

5. *Post process* $\mathbf{x} \leftarrow \mathcal{S}_M(\mathbf{x}, \mathbf{b})$.

Steps 1–4 of the algorithm above form the standard two-level method given by prolongator MP and smoothers $\mathcal{S}_M$, $\mathcal{S}_{M'}$. For such an algorithm, the theory presented here gives the error estimate in the A_M-seminorm. The practical value of this estimate is undermined by the fact that A_M depends on the coarse space, as using a smaller coarse-space calls for a more powerful smoother M to get the optimal convergence result. The seemingly unnecessary postprocessing step 5 will enable us to prove the same result in the energy norm of the original problem (1).

Throughout the paper, we shall use $\langle \cdot, \cdot \rangle$ to denote the Euclidean scalar product. Let $\| \cdot \|$ be the corresponding norm. If B is a symmetric positive (semi)definite matrix, we also use the (semi)norm derived from this matrix, i.e. $\|\mathbf{x}\|_B = \langle B\mathbf{x}, \mathbf{x} \rangle^{1/2}$. We denote the induced operator matrix norms by the same symbols.

The following assumption specifies requirements on M, P and $\bar{\varrho}(A_M)$ needed for proving the coarse-space size independent convergence.

ASSUMPTION 2.2. *There exist positive constants* C_1, C_2 *independent of* m, n *and there exists a positive constant* $C_D(m, n)$ *such that:*

1. *There is a linear mapping* $Q_C : \mathbb{R}^n \to \text{Range}\,(P)$ *such that*

$$\|(I - Q_C)\mathbf{u}\| \leq C_1 C_D(m,n)\varrho^{-1/2}(A)\|\mathbf{u}\|_A \quad \forall u \in \mathbb{R}^n. \tag{3}$$

2. *The prolongator smoother* M *is symmetric, commutes with* A*, satisfies* $\varrho(M) \leq 1$ *and the Hackbusch's smoothing property in the form*

$$\varrho(M^2A) \leq \bar{\varrho}(M^2A) \leq C_2^2 C_D^{-2}(m,n)\varrho(A). \tag{4}$$

A suitable choice of the prolongator smoother M and of the corresponding estimate $\bar{\varrho}(M^2A)$ will be given at the end of this section.

REMARK 2.1. Let us consider the case of Assumption 2.2 satisfied with reasonable constants C_1, C_2, but still $\varrho(M^2A) << \bar{\varrho}(M^2A)$. In other words, our estimate $\bar{\varrho}(M^2A)$ is small enough to satisfy the Assumption 2.2, but it is too pessimistic. In this case, M' defined by (2) is not an efficient smoother and one would expect a negative impact on the convergence. However, we will show that under Assumption 2.2 the convergence rate can be bounded in terms of constants C_1 and C_2, which means that it is independent of the ratio $\bar{\varrho}(M^2A)/\varrho(M^2A)$. This behavior can be explained heuristically. If $\varrho(M^2A) << \bar{\varrho}(M^2A)$ then, owing to (4), M is a very powerful smoother which compensates for the deficiency of M'. □

Typically, regularity-free multilevel theories require two kinds of properties on coarse spaces: a weak approximation condition similar to (3), and the smoothness of the coarse spaces, which is usually formulated as the energetic stability of the coarse space interpolator Q_C ([2, 14, 15]). Our assumptions can be interpreted similarly. The smoothing property (4) enforces the low energy of coarse-space functions $\langle AMP\mathbf{x}, MP\mathbf{x}\rangle$. The classical two-level theory ([6, 3]) ensures convergence polynomially dependent on the constant $C_D(m,n)$ if (3) is satisfied. Employment of the prolongator smoother satisfying the additional requirement (4) enables us to eliminate this dependence and obtain a result independent of the coarse-space size. The pivotal argument here is that our choice of the components of the method makes it possible to view the algorithm as a method for solving a problem with a reduced energy that is given by the A_M-seminorm. Similar idea can be found in most of two-level domain decomposition algorithms (e.g., [9]), where the smoothness is enforced through eliminating the unknowns associated with subdomain interiors. For example, when dealing with an H^1-equivalent problem, the elimination of interior degrees of freedom results in suppressing the original energy in H^1 down to that in $H^{1/2}$ norm on the interfaces.

The most practical choice of the prolongator smoother M is an appropriate polynomial in A. We will construct M so that $\varrho(M^2A) \leq C\text{deg}(M)^{-2}\varrho(A)$, where deg denotes the degree of the polynomial. In order to illustrate implications of this choice, let us consider a second order elliptic problem discretized on a quasiuniform finite element mesh of characteristic meshsize h. Furthermore, let the supports of our tentative coarse-space basis functions be shape regular domains of characteristic size H with bounded intersections. In such a case, we will use the polynomial M of degree about H/h. Therefore, in order to satisfy the Assumption 2.2, we need the tentative

prolongator P that satisfies (3) with the constant $C_D(m,n) = H/h$. This approximation property can be easily proven if P approximates at least constant functions aside from the essential boundary conditions. Moreover, $\deg(M) \approx H/h$ assures sparsity of the coarse level matrix A_M. Detailed arguments will be given in Section 5.

Let us recall that ω is the damping parameter from the definition (2) of M'. The following theorem provides the abstract convergence estimate.

THEOREM 2.3. *Let $\mathbf{e}_i$ denote the error after i iterations given by steps 1–4 of Algorithm 2.1 , and let $\mathbf{e}_i^M$ be the error $\mathbf{e}_i$ smoothed by step 5. Then, under Assumption 2.2, it holds that*

$$\|\mathbf{e}_{i+1}\|_{A_M}^2 \leq (1 - C_3)\|\mathbf{e}_i\|_{A_M}^2, \tag{5}$$

and

$$\|\mathbf{e}_i^M\|_A^2 \leq (1 - C_3)^i \|\mathbf{e}_0\|_A^2, \tag{6}$$

where

$$C_3(\omega) = \frac{(C_1C_2)^{-2}\omega(2-\omega)}{1 + (C_1C_2)^{-2}\omega(2-\omega)} \leq C_3(1) = \frac{(C_1C_2)^{-2}}{1 + (C_1C_2)^{-2}}, \quad \omega \in (0,2). \tag{7}$$

Proof. Since $\varrho(M) \leq 1$ and $\mathbf{e}_i^M = M\mathbf{e}_i$, we have $\|\mathbf{e}_i^M\|_A = \|\mathbf{e}_i\|_{A_M}$ and $\|\mathbf{e}_0\|_{A_M} \leq \|\mathbf{e}_0\|_A$. Therefore, (6) follows from (5).

Note that M may be singular. We define $Q_M : \mathbb{R}^n \to \mathrm{Ker}\,(M)^\perp$ to be the orthogonal projection with respect to the Euclidean scalar product. Since M commutes with A, the eigenvectors of M and A_M coincide. Consequently, Q_M is A_M-symmetric and $\|Q_M\|_{A_M} = 1$. It is obvious that the linear part of the steps 1–4 is given by

$$M'[I - MP(P^TA_MP)^+P^TMA]M = M'MQ_M[I - P(P^TA_MP)^+P^TA_M]Q_M,$$

where $(P^TA_MP)^+$ is a pseudo-inverse of P^TA_MP. Since $\mathrm{Ker}\,(P^TA_MP) = \{\mathbf{x} \in \mathbb{R}^m : P\mathbf{x} \in \mathrm{Ker}\,(M)\}$, the algorithm is independent of a particular choice of the pseudo-inverse. Thus, the method can be viewed as a standard two-level method for solving a problem with matrix A_M (in place of A) and prolongator P (in place of MP).

Let us define $T = \mathrm{Ker}\,(P^TA_M) \cap \mathrm{Ker}\,(M)^\perp$ and consider the A_M-orthogonal decomposition

$$\mathrm{Ker}\,(M)^\perp = T \oplus \mathrm{Range}\,(Q_MP).$$

It can be easily seen that

$$Q_M[I - P(P^TA_MP)^+P^TA_M] : \mathrm{Ker}\,(M)^\perp \to T$$

is the A_M-orthogonal projection. Further, as M commutes with A, M' commutes with M. Therefore, taking into account that $\varrho(M) \leq 1$ and $\varrho(M') \leq 1$, we have

$$\|M'MQ_M[I - P(P^TA_MP)^+P^TA_M]Q_M\|_{A_M}^2 \leq \sup_{\substack{\mathbf{x}\in T\\ \mathbf{x}\neq 0}} \min\left\{\frac{\|M\mathbf{x}\|_{A_M}^2}{\|\mathbf{x}\|_{A_M}^2}, \frac{\|M'\mathbf{x}\|_{A_M}^2}{\|\mathbf{x}\|_{A_M}^2}\right\}. \tag{8}$$

In the rest of the proof, we will show that at least one of the expressions in the minimum above is bounded by $1 - C_3$ for any $\mathbf{x} \in T \setminus \{0\}$, with C_3 defined by (7).

We first express $\|M'\mathbf{x}\|_{A_M}/\|\mathbf{x}\|_{A_M}$ in terms of $\|M\mathbf{x}\|_{A_M}/\|\mathbf{x}\|_{A_M}$. Let us define

$$\beta = \frac{\bar{\varrho}(A_M)}{\varrho(A_M)}.$$

By Assumption 2.2, $\beta \geq 1$. Then, it is easy to see that

$$\frac{\|A_M\mathbf{x}\|^2}{\|\mathbf{x}\|^2_{A_M}} \geq \hat{C}(\mathbf{x})\varrho(A_M) \quad \text{implies} \quad \frac{\|M'\mathbf{x}\|^2_{A_M}}{\|\mathbf{x}\|^2_{A_M}} \leq 1 - \hat{C}(\mathbf{x})\frac{\omega}{\beta}\left(2 - \frac{\omega}{\beta}\right). \tag{9}$$

Let us recall that $A_M = AM^2$, and A and M commute. Hence $\|M^2\mathbf{x}\|^2_A = \|M\mathbf{x}\|^2_{A_M}$, and

$$\frac{\|A_M\mathbf{x}\|^2}{\|\mathbf{x}\|^2_{A_M}} = \frac{\|A_M\mathbf{x}\|^2}{\|M^2\mathbf{x}\|^2_A}\frac{\|M^2\mathbf{x}\|^2_A}{\|\mathbf{x}\|^2_{A_M}} = \frac{\|AM^2\mathbf{x}\|^2}{\|M^2\mathbf{x}\|^2_A}\frac{\|M\mathbf{x}\|^2_{A_M}}{\|\mathbf{x}\|^2_{A_M}}. \tag{10}$$

Now, consider $\mathbf{x} \in T \subset \mathrm{Ker}\,(P^TAM^2)$. Then, setting $\mathbf{u} = M^2\mathbf{x}$, we have $\mathbf{u} \in \mathrm{Ker}\,(P^TA) = \mathrm{Range}\,(P)^{\perp_A}$, where $\perp_A$ denotes the A-orthogonal complement of a set. We estimate the ratio $\|AM^2\mathbf{x}\|^2/\|M^2\mathbf{x}\|^2_A$ using the standard orthogonality argument: For Q_C satisfying the weak approximation condition (3), we obtain

$$\begin{aligned} \|\mathbf{u}\|^2_A &= \langle A\mathbf{u}, \mathbf{u}\rangle = \langle A\mathbf{u}, \mathbf{u} - Q_C\mathbf{u}\rangle \leq \|A\mathbf{u}\|\,\|\mathbf{u} - Q_C\mathbf{u}\| \\ &\leq C_1C_D(m,n)\varrho^{-1/2}(A)\|A\mathbf{u}\|\,\|\mathbf{u}\|_A. \end{aligned}$$

Therefore,

$$\frac{\|A\mathbf{u}\|}{\|\mathbf{u}\|_A} \geq C_1^{-1}C_D^{-1}(m,n)\varrho^{1/2}(A).$$

Substituting this estimate into (10), using Assumption 2.2 and the definition of β, we get

$$\frac{\|A_M\mathbf{x}\|^2}{\|\mathbf{x}\|^2_{A_M}} \geq C_1^{-2}C_D^{-2}(m,n)\varrho(A)\frac{\|M\mathbf{x}\|^2_{A_M}}{\|\mathbf{x}\|^2_{A_M}} \geq (C_1C_2)^{-2}\beta\varrho(A_M)\frac{\|M\mathbf{x}\|^2_{A_M}}{\|\mathbf{x}\|^2_{A_M}}. \tag{11}$$

Thus, by (9) with $\hat{C}(\mathbf{x}) = (C_1C_2)^{-1}\beta\|Mx\|^2_{A_M}/\|x\|^2_{A_M}$,

$$\frac{\|M'\mathbf{x}\|^2_{A_M}}{\|\mathbf{x}\|^2_{A_M}} \leq 1 - \frac{\|M\mathbf{x}\|^2_{A_M}}{\|\mathbf{x}\|^2_{A_M}}(C_1C_2)^{-2}\omega\left(2 - \frac{\omega}{\beta}\right) \leq 1 - \frac{\|M\mathbf{x}\|^2_{A_M}}{\|\mathbf{x}\|^2_{A_M}}(C_1C_2)^{-2}\omega(2-\omega)$$

as $\beta \geq 1$. Owing to (8) and $\|M\mathbf{x}\|^2_{A_M}/\|\mathbf{x}\|^2_{A_M} \leq 1$, we may finally write

$$\|M'MQ_M[I - P(P^TA_MP)^+P^TA_M]Q_M\|^2_{A_M} \leq \sup_{\alpha\in[0,1]} \min\left\{\alpha, 1 - \alpha(C_1C_2)^{-2}\omega(2-\omega)\right\}.$$

The expression on the right hand side is bounded by $[1 + (C_1C_2)^{-2}\omega(2-\omega)]^{-1}$ which completes the proof. □

In the rest of this section we construct the smoother M. The algorithm has the following input: the desired degree d_M of M and an estimate $\hat{\varrho}$ of $\varrho(A)$ such that

$$\varrho(A) \le \hat{\varrho} \le C_\varrho \varrho(A). \tag{12}$$

The recommended choice of d_M will be given in Remark 5.1. For any positive integer i, we set

$$\hat{\varrho}_i = \frac{\hat{\varrho}}{9^i}. \tag{13}$$

Let $A_0 = A$; for any $i > 0, j \ge 0$ we recurrently define

$$M_i = \prod_{j=0}^{i-1} \left(I - \frac{4}{3}\hat{\varrho}_j^{-1} A_j \right), \quad A_j = \left(I - \frac{4}{3}\hat{\varrho}_{j-1}^{-1} A_{j-1} \right)^2 A_{j-1}. \tag{14}$$

It is easy to see that $\deg(M_i) = (3^i - 1)/2$. We choose the prolongator smoother

$$M = M_K \tag{15}$$

for K such that

$$\deg(M_{K+1}) > d_M \ge \deg(M_K). \tag{16}$$

In the definition (2), we take

$$\bar{\varrho}(AM^2) = \hat{\varrho}_K.$$

LEMMA 2.4. *Let the prolongator smoother M be defined by* (12)–(16). *Then,*

$$\varrho(AM^2) \le \bar{\varrho}(AM^2) = \frac{C_\varrho}{(2\deg(M)+1)^2}\varrho(A) \le \frac{C_\varrho}{4}\deg(M)^{-2}\varrho(A). \tag{17}$$

Proof. First, by induction, we prove $\varrho(A_i) \le \hat{\varrho}_i$: For $i = 0$, the inequality holds by (12); assume it holds for $j \le i$. Then, by (14) and the inductive assumption

$$\varrho(A_{i+1}) = \max_{t \in \sigma(A_i)} (1 - \frac{4}{3}\hat{\varrho}_i^{-1} t)^2 t \le \max_{t \in [0, \hat{\varrho}_i]} (1 - \frac{4}{3}\hat{\varrho}_i^{-1} t)^2 t \le \frac{1}{9}\hat{\varrho}_i = \hat{\varrho}_{i+1}.$$

Hence, $\varrho(A_K) \le 9^{-K}\hat{\varrho}$. Now, from (14), we find $AM^2 = A_K$. Finally, we obtain (17) from (12) and (13) by simple manipulations. □

REMARK 2.2. If the prolongator smoother M is defined by (12)–(16), with (15) replaced with the choice of $M = M_{K+1}$, we obtain acceleration in the following sense. Assuming $\omega = 1$ in (4), in order to achieve the error reduction of one step of the method with $M = M_{K+1}$, one has to perform on the average at least α_M steps of the method with $M = M_K$, where

$$\alpha_M = \frac{\log(1 - \frac{9(C_1C_2)^{-2}}{1+9(C_1C_2)^{-2}})}{\log(1 - \frac{(C_1C_2)^{-2}}{1+(C_1C_2)^{-2}})}$$

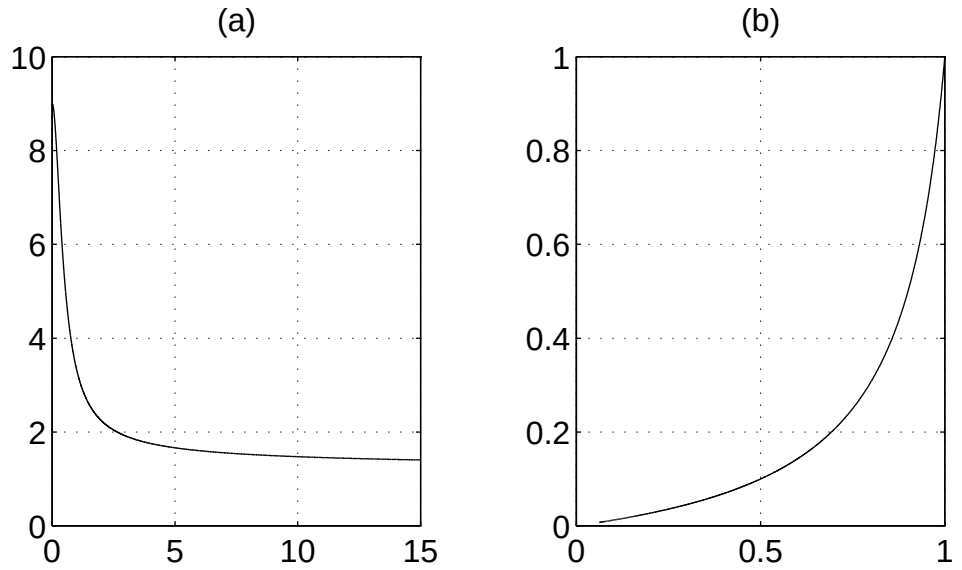

FIG. 1. *a) Dependence of α_M on $(C_1C_2)^{-1}$. b) Dependence of the convergence rate of the method with $M = M_{K+1}$ on that of the $M = M_K$.*

This expression is an increasing function of C_1C_2 with

$$\lim_{(C_1C_2)\to\infty} \alpha_M(C_1C_2) = 9, \lim_{(C_1C_2)\to 0+} \alpha_M(C_1C_2) = 1.$$

Thus for very large values of C_1C_2, we achieve large acceleration (up to 9 times fewer steps need to be taken), while for small C_1C_2 the speed-up is negligible. In other words, the gain from using more expensive prolongator smoother is most significant where the convergence with the less expensive smoother is slow. The dependence of α_M on the value of $(C_1C_2)^{-1}$ is shown in Figure 2 a). Figure 2 b) depicts the dependence of the rate of convergence bound for the $M = M_{K+1}$ on that of $M = M_K$. Computational complexity aspects will be discussed in Section 5. □

REMARK 2.3. (Preconditioning). In order to use the Algorithm 2.1 as a preconditioner of the conjugate gradients method, one needs to assure that the error propagation operator of the iterative method is $A-$symmetric. As M, M' and A commute and are symmetric, M and M' are $A-$symmetric. Therefore the method employing the same pre- and post-smoother defined in two steps by $\mathbf{y} \leftarrow \mathcal{S}_M(\mathbf{x}, \mathbf{b})$; $\mathbf{x} \leftarrow \mathcal{S}_{M'}(\mathbf{y}, \mathbf{b})$ has an $A-$symmetric positive semidefinite error propagation operator $MM'[I - MP(P^TA_MP)^+P^TMA]MM'$. Since $\varrho(M') \leq 1$, Theorem 2.3 yields

$$\|MM'[I - MP(P^TA_MP)^+P^TMA]MM'\|_A \leq (1 - C_3)^{1/2}, \tag{18}$$

where C_3 is given by (7). Let B denote one step of the iterative method consistent with (1), having the error propagation operator (18), started from zero approximation. Then it follows from (18) by well-known arguments that the relative condition number $\mathrm{cond}(A, B)$ is bounded by $\frac{1+\sqrt{1-C_3}}{C_3}$. □

3. Auxiliary coarse space for solids. In this section we describe an auxiliary coarse space for problems of isotropic solids with large jumps in Young modulus.

Let $V^{\mathcal{T}}$ be a conforming finite element space of displacement field associated with a mesh $\mathcal{T}$ on a Lipschitz domain Ω, and $\{\varphi_i\}_{i=1}^n$ be its basis. We assume there are three degrees of freedom (displacement components) per node, and the degrees of freedom are located in vertices of elements only. Further, we assume that elements support at least linear functions, the basis functions are scaled so that $\|\varphi_i\|_{L^\infty} = 1$, $|\varphi_i|_{H^1} \approx h_i$, and the standard equivalence of discrete and continuous norms $\|\sum \alpha_i \varphi_i\|_{L_2}^2 \approx \sum \alpha_i^2 h_i^3$ holds, where h_i denotes the local meshsize. It is obvious that standard P1 and Q1 elements conform to these assumptions.

Let $V_0^{\mathcal{T}} \subset V^{\mathcal{T}}$ be the subspace of all finite element displacement fields subject to general single-point constraints. We may impose up to three linearly independent constraints per boundary node, in the form

$$\sum_{l=1}^{3} \alpha_{kl} u_l(v) = 0, \qquad \alpha_{kl} \in \mathbb{R}, \quad k = 1, \ldots, 3, \tag{19}$$

where $u_l(v)$ are components of the displacement field, and $v \in \partial\Omega$. We are interested in the numerical solution of the following linear elasticity problem: Find a displacement field $u \in V_0^{\mathcal{T}}$ such that

$$a(u, v) = f(v) \qquad \forall v \in V_0^{\mathcal{T}},$$

where

$$a(u, v) = \int_\Omega (\lambda \operatorname{div} u \operatorname{div} v + 2\mu \sum_{i,j} \varepsilon_{ij}(u)\varepsilon_{ij}(v))\, d\mathbf{x}. \tag{20}$$

Here ε is the strain tensor, $\varepsilon_{ij}(u) = \frac{1}{2}(\partial_i u_j + \partial_j u_i)$, and $\lambda > 0$, $\mu > 0$ are the Lamé coefficients. The Lamé coefficients can be expressed in terms of the Poisson ratio ν and the Young modulus E as follows

$$\mu = \frac{E}{1+\nu} \qquad \text{and} \qquad \lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}. \tag{21}$$

The standard finite element discretization of the above bilinear form leads to a positive definite global stiffness matrix $A_{\mathcal{T},0}$. This matrix is obtained from the unconstrained global stiffness matrix $A_{\mathcal{T}} = \{a(\varphi_i, \varphi_j)\}_{i,j=1}^n$ by eliminating some degrees of freedom using (19). We assume that the constrained degrees of freedom are chosen so that the following is satisfied. Each constrained degree of freedom on a node can be expressed as a linear combination of free degrees of freedom with coefficients bounded uniformly with respect to node index. In other words, denoting $S(v), F(v)$ the index set of constrained and free degrees of freedom on node v, respectively, we assume that the single point constraints (19) are transformed into

$$u_i = \sum_{j \in F(v)} \beta_{ij} u_j, \quad i \in S(v), \tag{22}$$

with coefficients $|\beta_{ij}| \leq C_\beta$. Here i and j are indices in the global numbering of degrees of freedom of the matrix $A_{\mathcal{T}}$. This can be done, for example, as follows. First, for each node where the boundary condition (19) is imposed, the Gauss-Jordan elimination with row pivoting is applied to the matrix $\{\alpha_{kl}\}_{k,l=1}^{3}$ of constraints (19). Then, degrees of freedom associated with pivots are chosen to be the constrained degrees of freedom and are eliminated from the global stiffness matrix. In this case $C_\beta \leq 2$.

Let S denote the set of all constrained degrees of freedom, $S = \cup_{v \in \mathcal{T}} S(v)$, and similarly let F be the set of all free degrees of freedom $F = \cup_{v \in \mathcal{T}} F(v)$. We introduce a one-to-one mapping $N : F \to \{1, ..., n_f = \mathrm{card}(F)\}$ that establishes the index correspondence between the free degrees of freedom of $A_{\mathcal{T}}$ and the unknowns of $A_{\mathcal{T},0}$. Let us note that the constrained stiffness matrix $A_{\mathcal{T},0}$ can be interpreted as a Gram matrix $\{a(\phi_i, \phi_j)\}_{i,j=1}^{n_f}$ with a modified finite element basis

$$\phi_{N(j)} = \varphi_j + \sum_{i \in S(v)} \beta_{ij} \varphi_i, \quad j \in F(v) \tag{23}$$

where coefficients β_{ij} are given by (22). This basis is obtained from the requirement that the basis from (22) satisfies

$$\sum_{i=1}^{3} v_i \varphi_i = \sum_{i \in F} v_i \phi_i.$$

We define matrix D as follows

$$D_{ij} = \begin{cases} (A_{\mathcal{T},0})_{ij} & \text{if i and j are associated with the same node, i.e.} \\ & i, j \in N(F(v)) \text{ for some } v \in \mathcal{T} \\ 0 & \text{otherwise.} \end{cases} \tag{24}$$

This matrix is block diagonal when a suitable ordering of degrees of freedom is used. We will formulate the method for the following scaled matrix

$$A = D^{-1/2} A_{\mathcal{T},0} D^{-1/2}. \tag{25}$$

This is motivated by the need to accommodate variations of coefficients in the equation, and makes it possible to achieve invariance of the algorithm with respect to a change of spatial coordinates in (20).

Let $\Omega_i, i = 1, \ldots, n_s$ be a system of closed disjoint subdomains of Ω such that each subdomain is a simply connected aggregate of elements, and every node of the mesh belongs to one subdomain only. In other words, the union of Ω_i's cover the set of all nodes, but it does not cover the whole domain – there is a layer one element wide between the neighboring subdomains. We use F_i, S_i to denote the set of all free and constrained degrees of freedom on subdomain Ω_i, respectively.

On each subdomain Ω_i, we define a space of rigid body modes

$$R_i = \{u(\mathbf{x}) = C\mathbf{x} + \mathbf{c}, \quad \mathbf{x} \in \Omega_i\} \tag{26}$$

where C is an arbitrary antisymmetric 3×3 matrix and $\mathbf{c} \in \mathbb{R}^3$ a constant vector. In order to construct the auxiliary coarse space, we need a basis of R_i. As a

computationally convenient example of a basis in R_i, we suggest $\{f^{ij}(\mathbf{x}),\ \mathbf{x} = [x_1, x_2, x_3] \in \Omega_i\}_{j=1}^6$ given by

$$[f^{i1}, f^{i2}, f^{i3}, f^{i4}, f^{i5}, f^{i6}] = \begin{bmatrix} -x_2 + x_2^0 & -x_3 + x_3^0 & 0 & 1 & 0 & 0 \\ -x_1 + x_1^0 & 0 & -x_3 + x_3^0 & 0 & 1 & 0 \\ 0 & -x_1 + x_1^0 & x_2 - x_2^0 & 0 & 0 & 1 \end{bmatrix} \tag{27}$$

where $\mathbf{x}^0 = [x_1^0, x_2^0, x_3^0]$ is some fixed point in Ω_i.

ALGORITHM 3.1. (Tentative prolongator)

For every domain Ω_i and for $j = 1, \ldots 6$:

1. *Express f^{ij} as a linear combination of the basis of the unconstrained space $V^{\mathcal{T}}$ by*

$$f^{ij}(\mathbf{x}) = \sum_k c_k^{ij} (\varphi_k|_{\Omega_i}).$$

2. *Generate $\mathbf{d}^{ij} \in \mathbb{R}^{n_f}$ by $d^{ij}_{N(k)} = c_k^{ij} \quad \forall k \in F_i$.*
3. *Interpret the vector $D^{1/2}\mathbf{d}^{ij}$ as $6(i-1)+j$-th column of the tentative prolongator P.*

Note that under our assumptions on the basis $\{\varphi_i\}$ the coefficients c_k^{ij} are nodal values of the vector function f^{ij} in grid points. The algorithm in this form can be used to treat more general bases (e.g., unscaled bases or high order elements). Results given in the next section can be easily extended to such cases.

4. Verification of the weak approximation property. For the tentative coarse space given by Algorithm 3.1 we will show the weak approximation condition (3) with a constant independent of the boundary conditions, the domain Ω and the magnitude of the jumps in E, provided the subdomains Ω_i are suitably aligned with the jumps. We need to introduce additional notation and assumptions on the triangulation and subdomain decomposition.

DEFINITION 4.1. ([4]) *Suppose a domain ω has diameter d and is star-shaped with respect to a ball B, i.e. for all x in Ω, the closed convex hull of $\{x\} \cup B$ is a subset of Ω. Let $\varrho_{max} = \sup\{\varrho : \omega$ is star-shaped with respect to a ball of radius $\varrho\}$. Then the chunkiness parameter of ω is defined by*

$$\gamma(\omega) = \frac{d}{\varrho_{max}}.$$

ASSUMPTION 4.2.

1. *Each Ω_i can be mapped onto a unit cube $\hat{\Omega}$ by a mapping $F_i : \hat{\Omega} \to \Omega_i$ that satisfies*

$$c_F H_i \le \|\partial F_i(x)\| \le C_F H_i \quad \forall x \in \Omega_i \tag{28}$$

where $\|\cdot\|$ is a 3×3 operator matrix norm. In other words, Ω_i is a shape regular domain of characteristic size H_i.

2. *There is about the same number of elements in each subdomain. More specifically, denoting the characteristic number of elements per subdomain by N_{es}, we require*

$$c_H N_{es}^{1/3} \leq H_i/h_i \leq C_H N_{es}^{1/3}. \tag{29}$$

A bounded number of subdomains may violate this requirement in the sense that the number of elements they contain may be smaller.

3. *Each subdomain Ω_i is a union of at most k_Ω star shaped domains ω_j^i with uniformly bounded chunkiness parameters $\gamma(\omega_j^i) < \gamma_0$ such that*

$$\mu(\omega_{k+1} \cap (\cup_{j=1}^k \omega_j)) \geq c_0 \min(\mu(\omega_{k+1}), \mu(\cup_{j=1}^k \omega_j)), \quad k < k_\Omega.$$

4. *The tringulation $\mathcal{T}$ is shape-regular, and is quasi-uniform with a characteristic element size h_i on each subdomain Ω_i.*
5. *The Poisson ratio satisfies $0 < \nu_1 \leq \nu(x) \leq \nu_2 < 0.5$ and the Young modulus E is almost constant on each subdomain in the sense that there exist constants positive c_E, C_E independent of the subdomain index and a constant $\bar{E}_i > 0$ such that on $\Omega_i \quad c_E \bar{E}_i \leq E \leq C_E \bar{E}_i$.*
6. *If Ω_i and Ω_j are neighboring subdomains (in the sense that for some element $T \in \mathcal{T}$ it holds $\partial T \cap \partial\Omega_i \neq \emptyset$ and $\partial T \cap \partial\Omega_j \neq \emptyset$) and $\bar{E}_i >> \bar{E}_j$, we assume that there is a jump in E that occurs at $\partial\Omega_i$. In other words, the discontinuity is located on the boundary of the subdomain with a larger Young modulus.*

The purpose of assumption 4.2.6. is to ensure for the basis function φ_j associated with a node $v \in \Omega_i$ to satisfy

$$a(\varphi_j, \varphi_j) \approx h_i \bar{E}_i. \tag{30}$$

If assumption 4.2.6. were not satisfied, the basis functions corresponding to the nodes on $\partial\Omega_j$ adjacent to a subdomain Ω_i with $\bar{E}_i >> \bar{E}_j$ could violate (30).

The following lemma will be useful in proving the weak approximation property later in this section.

LEMMA 4.3. *Under the assumptions 4.2.1,3. above, it holds that*

$$\inf_{w \in R_i} \|u - w\|^2_{[L^2(\Omega_i)]^3} \leq C H_i^2 \int_{\Omega_i} \sum_{i,j=1}^3 \varepsilon_{ij}(u)^2 dx \qquad \forall u \in [H^1(\Omega_i)]^3$$

where C depends on γ_0, c_0, k_Ω.

Proof. It is well known [7] that for a star shaped domain ω Korn's constant $K(\omega)$ on the factor space modulo rigid body modes R,

$$K(\omega)^{-1} = \inf_u \frac{\int_\omega \sum_{i,j=1}^3 \varepsilon_{ij}(u)^2 dx}{|u|_{[H^1(\omega)]^3/R}},$$

can be controlled in terms of the chunkiness parameter $\gamma(\omega)$. Furthermore, for two Lipschitz domains ω_1 and ω_2, Korn's constant on $\omega_1 \cup \omega_2$ can be estimated [5] as

$$K(\omega_1 \cup \omega_2) \geq K(\omega_1) + K(\omega_2) + \frac{\min(\mu(\omega_1), \mu(\omega_2))}{\mu(\omega_1 \cap \omega_2)} (\sqrt{K(\omega_1)} + \sqrt{K(\omega_2)})^2,$$

where μ denotes the Lebesgue measure. This, together with the assumptions, implies that $K(\Omega_i)$ depends only on γ_0, c_0 and k_Ω. The scaled Poincaré inequality completes the proof. □

We will use a_Δ to denote the bilinear form as in (20) in which the integral is taken over $\Delta \subset \Omega$.

LEMMA 4.4. *Let D be the matrix defined by (24), $D_1 = \mathrm{diag}\,(A_{T,0})$ and $D_2 = \mathrm{diag}\,(\{a(\varphi_{N^{-1}(i)}, \varphi_{N^{-1}(j)})\}_{i,j=1}^{n_f})$ Then, assuming 4.2.3–4, D, D_1 and D_2 are spectrally equivalent with constants of equivalence depending only on aspect ratios of elements.*

Proof. As every matrix in question is either diagonal or block diagonal (after a suitable reordering), we can prove the spectral equivalence node by node (i.e. block by block). For convenience, we omit node indices (that is, we use symbol D, D_1 and D_2 for blocks corresponding to a node instead of whole matrices, and $F \equiv F(v), S \equiv S(v)$ throughout this lemma). Thus, our blocks corresponding to node v are defined by $D = \{a(\phi_i, \phi_j)\}_{i,j \in N(F)}$, $\quad D_1 = \mathrm{diag}\,(D)$, $\quad D_2 = \mathrm{diag}\,(\{a(\varphi_i, \varphi_j)\}_{i,j \in F})$, see (23). Let $\mathcal{S}$ denote the (shared) support of basis functions $\phi_{N(i)}, \varphi_i$, $\quad i \in F$, and $T \in \mathcal{S}$ an element adjacent to the node v.

Denoting the characteristic Young modulus on T by E_T, Korn's inequality together with boundedness of $a(\cdot,\cdot)$ gives:

$$a_T(\sum_{i \in F \cup S} c_i \varphi_i, \sum_{i \in F \cup S} c_j \varphi_j) \approx E_T \sum_{i \in F \cup S} c_i^2 |\varphi_i|^2_{H^1(T)} \approx \sum_{i \in F \cup S} c_i^2 a_T(\varphi_i, \varphi_i)$$

where the constants of equivalence depend on the shape of elements only. Summing over all elements adjacent to node v, we obtain

$$a_{\mathcal{S}}(\sum_{i \in S \cup F} c_i \varphi_i, \sum_{i \in S \cup F} c_j \varphi_j) \approx \sum_{i \in S \cup F} c_i^2 a_{\mathcal{S}}(\varphi_i, \varphi_i) \approx E_{\mathcal{S}}^2 h \sum_{i \in S \cup F} c_i^2, \tag{31}$$

where h is the local mesh size, and $E_{\mathcal{S}}^2 = \max_{T \in \mathcal{S}} E_T^2$. By definition of D and (23), for $\mathbf{w} = \{w_{N(i)}\}_{i \in F}$

$$\mathbf{w}^T D \mathbf{w} = a\left(\sum_{j \in F} w_{N(j)} \varphi_j + \sum_{i \in S} (\sum_{j \in F} w_{N(j)} \beta_{ij}) \varphi_i, \sum_{j \in F} w_{N(j)} \varphi_j + \sum_{i \in S} (\sum_{j \in F} w_{N(j)} \beta_{ij}) \varphi_i\right).$$

Thus, denoting $\delta_i = w_{N(i)}$ for $i \in F$ and $\delta_i = \sum_{j \in F} w_{N(j)} \beta_{ij}$ for $i \in S$, we have

$$\mathbf{w}^T D \mathbf{w} \approx \sum_{i \in F \cup S} E_{\mathcal{S}}^2 h \delta_i^2.$$

As $0 \le |\beta_{ij}| \le \mathrm{card}(S)$, $\quad \mathbf{w}^T D \mathbf{w} \approx E_{\mathcal{S}}^2 h \mathbf{w}^T \mathbf{w} \approx \mathbf{w}^T D_2 \mathbf{w}$.

In order to prove that $D_1 \approx D_2$, it suffices to show that

$$a(\varphi_j + \sum_{i \in S} \beta_{ji} \varphi_i, \varphi_j + \sum_{i \in S} \beta_{ji} \varphi_i) \approx E_{\mathcal{S}}^2 h, \quad \text{for } j \in F.$$

From (31) and $|\beta_{ij}| \le \mathrm{card}(S)$ it follows that

$$a\left(\varphi_j + \sum_{i \in S} \beta_{ij} \varphi_i, \varphi_j + \sum_{i \in S} \beta_{ij} \varphi_i\right) \approx a(\varphi_j, \varphi_j) + \sum_{i \in S} \beta_{ij}^2 a(\varphi_i, \varphi_i) \approx E_{\mathcal{S}}^2 h,$$

completing the proof. □

Before we turn to proving the weak approximation property, we introduce some additional notation. Recall that $n_f = \mathrm{card}(F)$ is the number of all free degrees of freedom, and $\{\varphi_i\}_{i=1}^{n}$ and $\{\phi_i\}_{i=1}^{n_f}$ is the finite element basis of the constrained and unconstrained finite element space, respectively. Consider the index set $\mathcal{I} \subset \{1, \ldots, n = \mathrm{card}(F \cup S)\}$.

For the vector $\mathbf{x} \in \mathbb{R}^n$ we define the discrete l^2-norm

$$\|\mathbf{x}\|_{l^2(\mathcal{I})} = \left(\sum_{i \in \mathcal{I}} x_i^2 \right)^{1/2}.$$

The subspace of all vectors $\mathbf{x} \in \mathbb{R}^n$ such that $x_i = 0$ for all $i \notin \mathcal{I}$ will be denoted by $V(\mathcal{I})$. Let us define finite element interpolators

$$\begin{aligned} \Pi_0 : \mathbf{u} \in \mathbb{R}^{n_f} &\rightarrow \sum_{i=1}^{n_f} u_i \phi_i, \\ \Pi_0^i : \mathbf{u} \in V(F_i) \subset \mathbb{R}^n &\rightarrow \sum_{i \in N(F_i)} u_i \phi_i, \end{aligned}$$

and

$$\begin{aligned} \Pi : \mathbf{u} \in \mathbb{R}^n &\rightarrow \sum_{i \in F \cup S} u_i \varphi_i, \\ \Pi^i : \mathbf{u} \in V(F_i \cup S_i) \subset \mathbb{R}^n &\rightarrow \sum_{i \in F_i \cup S_i} u_i \varphi_i. \end{aligned}$$

The inverse of these interpolators is well-defined on their range.

Let $P_{R_i} : [H^1(\Omega_i)]^3 \rightarrow R_i$ be the L^2-orthogonal projection onto the space of rigid body modes on Ω_i. We define the linear operator $Q_i : [H^1(\Omega_i)]^3 \rightarrow \mathrm{Range}\,(\Pi_0^i)$ as follows:

1. $u \in [H^1(\Omega_i)]^3 \rightarrow v = P_{R_i} u \in R_i$,
2. $v \rightarrow \mathbf{w} = (\Pi^i)^{-1} v \in V(F_i \cup S_i)$,
3. $\mathbf{w} \rightarrow \mathbf{w}^c \in V(F_i)$ by $w_j^c = w_j$ if $j \in F_i$, zero otherwise.
4. Set $Q_i u = \Pi_0^i \mathbf{w}_c$.

Finally, let us define the mapping $Q : \mathbb{R}^{n_f} \rightarrow \mathbb{R}^{n_f}$ by

$$Q = (\Pi_0)^{-1} (\sum_{i=1}^{n_s} Q_i) \Pi_0.$$

LEMMA 4.5 (WEAK APPROXIMATION PROPERTY). *Let $\|\cdot\|$ denote the Euclidean norm in $\mathbb{R}^{n_f}$. Under the assumptions 4.2 it holds that*

$$\|\mathbf{u} - D^{1/2} Q D^{-1/2} \mathbf{u}\| \leq C N_{es}^{1/3} \|\mathbf{u}\|_A \quad \forall \mathbf{u} \in \mathbb{R}^{n_f}, \tag{32}$$

where the constant C depends only on the constants from Assumption 4.2. Moreover, for the prolongator P generated by Algorithm 3.1, we have

$$\mathrm{Range}\,(D^{1/2} Q D^{-1/2}) = \mathrm{Range}\,(P). \tag{33}$$

Proof. The proof of (33) follows immediately from the construction of P and Q. Let us prove (32).

Set $\mathbf{u} = D^{1/2}\mathbf{v}$. Then

$$\|\mathbf{u} - D^{1/2}QD^{-1/2}\mathbf{u}\|^2 = \|D^{1/2}(I-Q)\mathbf{v}\|^2.$$

Due to Lemma 4.4, we may replace D by the spectrally equivalent matrix D_2. Consequently, owing to (30), we can write

$$\begin{aligned} \|\mathbf{u} - D^{1/2}QD^{-1/2}\mathbf{u}\|^2 &\le C\|D_2^{1/2}(I-Q)\mathbf{v}\|^2 \\ &\le C\sum_{i=1}^{n_s} \bar{E}_i h_i \|\mathbf{v} - (\Pi_0^i)^{-1}Q_i\Pi_0^i\mathbf{v}\|^2_{l^2(F_i)}. \end{aligned} \tag{34}$$

The operator $(\Pi_0^i)^{-1}Q_i\Pi_0^i\mathbf{v}$ returns the nodal values of $P_{R_i}\Pi_0^i\mathbf{v}$ in the unconstrained degrees of freedom of Ω_i (the index set F_i). For

$$\mathbf{x} = (\Pi_0^i)^{-1}Q_i\Pi_0^i\mathbf{v} \quad \text{and} \quad \mathbf{y} = (\Pi^i)^{-1}P_{R_i}\Pi_0^i\mathbf{v}$$

we have $y_j = x_j$ for all $j \in F_i$. For the same reason, $\|\mathbf{v}\|_{l^2(F_i)} = \|(\Pi^i)^{-1}\Pi_0^i\mathbf{v}\|_{l^2(F_i)}$. Thus,

$$\begin{aligned} \|\mathbf{v} - (\Pi_0^i)^{-1}Q_i\Pi_0^i\mathbf{v}\|^2_{l^2(F_i)} &= \|(\Pi^i)^{-1}(I-P_{R_i})\Pi_0^i\mathbf{v}\|^2_{l^2(F_i)} \\ \le \|(\Pi^i)^{-1}(I-P_{R_i})\Pi_0^i\mathbf{v}\|^2_{l^2(F_i\cup S_i)} &\le Ch_i^{-3}\|(I-P_{R_i})\Pi_0^i\mathbf{v}\|^2_{[L^2(\Omega_i)]^3}. \end{aligned}$$

Substituting the last estimate into (34), using the Lemma 4.3 and (29), we obtain

$$\begin{aligned} \|\mathbf{u} - D^{1/2}QD^{-1/2}\mathbf{u}\|^2 &\le C\sum_{i=1}^{n_s}\bar{E}_i\left(\frac{H_i}{h_i}\right)^2\int_{\Omega_i}\sum_{k,l=1}^{3}\varepsilon_{kl}(\Pi_0^i\mathbf{v})^2\,\mathrm{d}x \\ \le CN_{es}^{2/3}a(\Pi_0\mathbf{v},\Pi_0\mathbf{v}) &= CN_{es}^{2/3}\langle A_{\mathcal{T},0}\mathbf{v},\mathbf{v}\rangle = CN_{es}^{2/3}\langle A\mathbf{u},\mathbf{u}\rangle, \end{aligned} \tag{35}$$

which completes the proof. □

5. Main result. We are now ready to formulate and prove the convergence theorem.

THEOREM 5.1. *Let the tentative prolongator P be given by Algorithm 3.1 applied to a system of subdomains $\{\Omega_i\}_{i=1}^m$ satisfying Assumption 4.2. Let the prolongator smoother M be defined by* (12)–(16) *with d_M satisfying*

$$c_N N_{es}^{1/3} \le d_M \le C_N N_{es}^{1/3}. \tag{36}$$

Then the error $\mathbf{e}_i^M$ satisfies

$$\|\mathbf{e}_i^M\|_A^2 \le (1-C)^i\|\mathbf{e}_0\|_A^2. \tag{37}$$

Constant C depends on C_ϱ, maximal aspect ratio of elements, constants from Assumption 4.2, (36)*, and is independent of H_i, h_i, $\bar{E}_i$, Ω and boundary conditions.*

Moreover, the coarse-level matrix and the smoothed prolongator MP has a uniformly bounded number of nonzero entries per row.

Proof. Due to Lemma 4.5, the approximation property (3) is satisfied with $C(m,n) = N_{es}^{1/3}$. Lemma 2.4 shows that the smoothing property (4) holds with the same $C(m,n)$ if we choose the degree of the polynomial $\deg(M) \approx N_{es}^{1/3}$. The optimal convergence result now follows from Theorem 2.3.

Let us show that the number of nonzero entries per row of the coarse-level matrix $A_c = (MP)^T A(MP)$ is bounded uniformly with respect to N_{es}. It is easy to see that $[A_c]_{ij}$ can be nonzero only if $\mathrm{supp}(\Pi_0 MP\mathbf{e}^i) \cap \mathrm{supp}(\Pi_0 MP\mathbf{e}^j) \neq \emptyset$, where $\mathbf{e}^i$ is the $i-$th canonical basis vector of $\mathbb{R}^m$. Clearly, $\mathrm{supp}(\Pi_0 P\mathbf{e}^i)$ is the domain Ω_i with one belt of surrounding elements added. Bounded overlaps of such supports are obvious. The smoother M adds at most $\deg(M) \approx N_{es}^{1/3} \approx H_i/h_i$ strips of elements. As subdomains Ω_i are shape-regular, each support has a nonempty intersection with only a bounded number of other supports. The sparsity of MP follows. ∎

THEOREM 5.2. *Let the assumptions of Theorem 5.1 be fulfilled and the Choleski factorization be used to solve the coarse-level problem. Then the optimal number of elements per subdomain is $N_{es} \approx n^{1/2}$ and the system (1) can be solved to the level of truncation error in $O(n^{7/6})$ operations.*

Proof. Operation counts in this proof make use of the fact that the matrices MP and $P^T M^2 AP$ are sparse and we have to perform $O(1)$ iterations. Components of the setup phase, and their respective costs in number of floating point operations are:

Construction of MP:	$O(\deg(M)n) = O(N_{es}^{1/3}n)$
Construction of coarse level matrix:	$O(n)$
Factorization of coarse level matrix:	$O((n/N_{es})^{7/3})$

The iterative phase consists of the following components:

Smoothing:	$O((\deg(M)n)) = O(N_{es}^{1/3}n)$
Prolongation, Restriction, Defect evaluation:	$O(n)$
Back substitution:	$O((n/N_{es})^{5/3})$

Summarizing, the total cost is:

$$\max\left(O(nN_{es}^{1/3}), O((n/N_{es})^{7/3})\right).$$

Minimizing, we find that the optimal subdomain size is $N_{es} = n^{1/2}$. In that case, the total cost is of the order of $n^{7/6}$. ∎

REMARK 5.1. (Choice of d_M in (16)). Let N_d denote the average number of nodes per subdomain. Note that $N_d \approx N_{es}$. For the computational purposes, we suggest to choose

$$d_M = \lfloor \frac{1}{2}(N_d^{1/3} - 1) \rfloor,$$

where $\lfloor \cdot \rfloor$ denotes the truncation to the nearest smaller integer. This choice is motivated as follows. Let $\mathbf{e}_i$ denote the $i-$th canonical basis vector of $\mathbb{R}^{6n_s}$. Assume we have a uniform grid of a meshsize h and our subdomains are cubes of size H. Then, $N_d^{1/3} - 1 = H/h$ and the choice of d_M assures, that supports of coarse space basis functions $\mathrm{supp}(\Pi_0 MP\mathbf{e}^i)$, $\mathrm{supp}(\Pi_0 MP\mathbf{e}^j)$ overlap only if these basis functions correspond to the same subdomain or two adjacent subdomains. $\square$

REMARK 5.2. (Case without jumps in E). If $E \approx \bar{E}$ on Ω, we can show that the weak approximation property (32) holds even for subdomains that do not satisfy Assumption 4.2.3. To this end, let Ω^{ext}, $\Omega \subset \Omega^{ext}$ be a sufficiently large domain. Let us consider the extension $u_{ext} \in [H^1(\Omega^{ext})]^3$ of $u \in [H^1(\Omega)]^3$ such that $u_{ext} = u$ on Ω and

$$a_{\Omega^{ext}}(u_{ext}, u_{ext}) \leq C(\Omega)\, a(u,u). \tag{38}$$

Such an extension exists owing to Korn's inequality on the factorspace modulo rigid body modes on Ω and the H^1 extension theorem ([1]). For each subdomain Ω_i let us consider a domain Ω_i^{ext}, $\Omega_i \subset \Omega_i^{ext}$ such that Ω_i^{ext} satisfies 4.2.3, H_i is the characteristic size of Ω_i^{ext} (in the sense of (28)) and overlaps of subdomains Ω_i^{ext} are bounded. For example, we can choose Ω_i^{ext} to be cubes. Then, following the proof of Lemma 4.3, we obtain

$$\bar{E} \inf_{w \in R_i} \|u - w\|^2_{L^2(\Omega_i)} \leq C H_i^2 a_{\Omega^{ext}}(u_{ext}, u_{ext}) \tag{39}$$

and the weak approximation property (32) can be proven with the constant C dependent on Ω, provided that the argument (35) in the proof of Lemma 4.5 is replaced with

$$\begin{aligned}
\|\mathbf{u} - D^{1/2} Q D^{-1/2}\mathbf{u}\|^2 &\leq C \sum_{i=1}^{n_s} \left(\frac{H_i}{h_i}\right)^2 a_{\Omega_i^{ext}}((\Pi_0\mathbf{v})_{ext}, (\Pi_0\mathbf{v})_{ext}) \\
\leq C N_{es}^{2/3} a_{\Omega^{ext}}((\Pi_0\mathbf{v})_{ext}, (\Pi_0\mathbf{v})_{ext}) &\leq C(\Omega) N_{es}^{2/3} a(\Pi_0\mathbf{v}, \Pi_0\mathbf{v}) \\
= C(\Omega) N_{es}^{2/3} \langle A_{\mathcal{T},0}\mathbf{v}, \mathbf{v}\rangle &= C(\Omega) N_{es}^{2/3} \langle A\mathbf{u}, \mathbf{u}\rangle.
\end{aligned}$$

The convergence result obtained this way depends on the constant $C(\Omega)$ from (38). This undesirable dependence can be eliminated if subdomains Ω_i and Ω_i^{ext} are chosen so that $\Omega_i^{ext} \subset \Omega, \quad i = 1, \ldots, n_s$. $\square$

6. Numerical experiments.

In this section, we present two sorts of numerical experiments for the elasticity problem (20). Model experiments on a uniformly discretized cube validate the theory above, i.e. uniformity of the convergence with respect to a coarse-space size, boundary conditions and jumps in coefficients. Secondly, we have used the algorithm to carry out numerical experiments with a real-life problem – model of a turbine, jumps in coefficients that are not aligned with subdomains and randomly chosen Young modulae. The purpose of these experiments is to illustrate the potential of the method in situations when some of the assumptions of the theory are not satisfied.

Linear elasticity problem (20) has been discretized using CHEXA UAI/NASTRAN elements in all of the experiments. UAI/NASTRAN has also been used to generate the

Cube: 9261 nodes, 27783 degrees of freedom, $E = \text{const}, \nu = 0.3$							
Remark	n_s	$\deg(M)$	Z	N_{it}	T_{CPU}	ρ	cond
Large subd.	8	4	8	5	30 + 37	0.07	1.30
Small subd.	343	1	23	5	33 + 20	0.09	1.37

TABLE 1
Uniformity of convergence with respect to coarse-space size

stiffness matrix and the rigid body modes. We have used the algorithm to precondition the conjugate gradient method, and stopped the iteration process once the relative preconditioned residual satisfied

$$\left[\frac{\langle AB\mathbf{e}_i, A\mathbf{e}_i\rangle}{\langle AB\mathbf{e}_0, A\mathbf{e}_0\rangle}\text{cond}(B,A)\right]^{1/2} \leq 5.10^{-5},$$

where B is the preconditioner and $\text{cond}(B,A)$ is a condition number estimate computed at run time. The computational domains have been decomposed using a greedy algorithm that tends to generate nearly cubic subdomains of about the same size for regular meshes.

Experiments in tables 6 6, 6, 6 have been performed on IBM RS 6000/360. Due to memory requirements of a mesh with more than 120000 degrees of freedom, we carried out the experiments with the turbine on a multiprocessor Silicon Graphics Power Challenge, using only one of its CPUs.

In the table headers, we use the following symbols:

n_s	the number of subdomains
$\deg(M)$	prolongator smoother degree
Z	fill-in of the coarse-level matrix, see the description below
N_{it}	number of iterations
ρ	rate of convergence = (relative residual after N_{it} iterations)$^{1/N_{it}}$
cond	condition number estimate
CPU	Setup CPU time + iteration CPU time

On the coarse level, there are 6 degrees of freedom associated with each subdomain. The coarse-level matrix is stored in a block-sparse format; the order of each block is 6. Block ij contains the information about the communication of $i-$th and $j-$th subdomain. The number Z reported in tables is the average number of nonzero blocks per column of the coarse-level matrix.

Uniformity of the convergence with respect to coarse-space size. In order to verify the coarse-space size independent convergence, we have run two experiments with quite different number of subdomains. The computational domain was a regularly discretized $[0,1] \times [0,1] \times [0,1]$ cube. We have prescribed the zero Dirichlet boundary condition on a face of the cube, and used Young modulus $E = 1$ and Poisson ratio $\nu = 0.3$. The results are shown in Table 6.

Uniformity of the convergence with respect to jumps in Young modulus. In order to demonstrate the uniform boundedness of the convergence with respect to jumps in

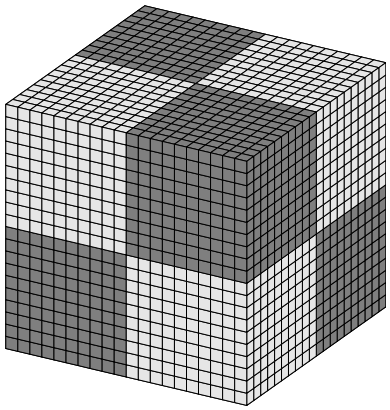

FIG. 2. *Jumps in E. Young modulus ($E = E_1$) on dark subdomains; $E = E_2 = 1 < E_1$ on light ones. All dark subdomains are visible.*

Cube: 9261 nodes, 27783 dfs, $\nu = 0.3$, for E see Fig. 6							
E_1/E_2	n_s	deg(M)	Z	N_{it}	T_{CPU}	ρ	cond
10	8	4	8	6	31 + 68	0.141	1.75
10*	27	4	18	7	47 + 32	0.196	2.30
10*	343	1	23	5	30 + 19	0.125	1.50
100	8	4	8	8	30 + 85	0.231	3.54
100*	27	4	18	9	47 + 40	0.296	3.91
100*	343	1	23	6	31 + 23	0.146	1.80
1000	8	4	8	8	32 + 87	0.258	4.46
1000*	27	4	18	10	48 + 45	0.328	4.96
1000*	343	1	23	6	31 + 22	0.150	1.81

TABLE 2

*Uniformity of convergence with respect to jumps in Young modulus. Symbol * indicates, that the assumptions of the theory are not valid.*

coefficients, the cube has been subdivided into 8 subdomains (see Fig. 6). There is a greater Young modulus $E = E_1$ on dark subdomains, while on the lighter subdomains $E = E_2 = 1$. Zero Dirichlet boundary condition has been imposed in all corners of the lower face and the adjacent surface points. Results of experiments are summed up in Table 6. Only the experiments with 8 computational subdomains conform to the Assumption 4.2, other cases violate 4.2.6.

Uniformity of the convergence with respect to boundary conditions. This set of experiments shown in Table 6 demonstrates the independence of the convergence of boundary conditions. The abbreviations of the names of experiments in the table stand for the following boundary conditions:

Face BC: Zero boundary condition on the lower face $[0,1] \times [0,1] \times \{0\}$ of the cube.

Local BC: Zero boundary condition in all corners of the lower face and adjacent surface nodes.

9261 nodes, 27783 degrees of freedom, $E = \text{const}, \nu = 0.3$							
Experiment	n_s	$\deg(M)$	Z	N_{it}	T_{CPU}	ρ	cond
Face BC	27	4	16	5	47 + 23	0.10	1.42
Face BC	343	1	23	5	33 + 20	0.09	1.37
Local BC	27	4	19	6	48 + 29	0.11	1.56
Local BC	343	1	24	6	31 + 22	0.10	1.48
Point BC	27	4	19	6	47 + 29	0.11	1.57
Point BC	343	1	24	6	32 + 22	0.10	1.46

TABLE 3
Uniformity of convergence with respect to boundary conditions

REFERENCES

[1] R. ADAMS, *Sobolev Spaces*, Academic Press, New York, 1975.

[2] J. H. BRAMBLE, J. E. PASCIAK, J. WANG, AND J. XU, *Convergence Estimates for Multigrid Algorithms without Regularity Assumptions*, Math. Comp., 57 (1991), pp. 23–45.

[3] A. BRANDT, *Algebraic multigrid theory: The symmetric case*, Appl. Math. Comput., 19 (1986), pp. 23–56.

[4] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 1994.

[5] C. M. DAFERMOS, *Some Remarks on Korn's Inequality*, ZAMP, 19 (1968), pp. 913–920.

[6] W. HACKBUSCH, *Multigrid Methods and Applications*, vol. 4 of Computational Mathematics, Springer–Verlag, Berlin, 1985.

[7] V. KONDRATIEV AND O. OLEINIK, *Hardy's and Korn's Type Inequalities and their Applications*, Rendiconti di Matematica, Serie VII, 10 (1990), pp. 641–666.

[8] J. KŘÍŽKOVÁ AND P. VANĚK, *Two–level preconditioner with small coarse grid appropriate for unstructured meshes*, Numerical Linear Algebra with Applications. to appear.

[9] J. MANDEL, *Domain Decomposition on Unstructured Subdomains*, in Domain Decomposition Methods in Science and Engineering, A. Quarteroni, J. Périaux, Y. A. Kuznetsov, and O. B. Widlund, eds., American Mathematical Society, Providence, RI, 1993, pp. 103–112. Proceedings of the 6th International Symposium on Domain Decomposition Methods, Como, Italy, 1992.

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

[11] P. VANĚK AND J. KŘÍŽKOVÁ, *Two–level Method on Unstructured Meshes with Convergence Rate Independent of the Coarse-space Size*, Tech. Report 35, Center of Computational Mathematics/UCD, 1995.

[12] P. VANĚK, J. MANDEL, AND M. BREZINA, *Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems*. To appear in Computing.

[13] P. VANĚK, R. TEZAUR, M. BREZINA, AND J. KŘÍŽKOVÁ, *Two–level Method with Coarse Space Size Independent Convergence*, in Domain Decomposition Methods in Sciences and Engineering, R. Glowinski, J. Périaux, Z. Shi, and O. Widlund, eds., John Wiley & Sons Ltd., New York, N.Y., 1995. to appear.

[14] J. WANG, *Convergence Analysis without Regularity Assumptions for Multigrid Algorithms Based on SOR Smoothing*, SIAM J. Numer. Anal., 29 (1992), pp. 987–1001.

[15] ——, *Convergence Analysis of Schwartz Algorithm and Multilevel Decomposition Iterative Methods II: Non-selfadjoint and Positive Indefinite Elliptic Problems*, SIAM J. Numer. Anal., 30 (1993), pp. 953–970.

9261 nodes, 27783 degrees of freedom, E = random, $\nu = 0.3$							
Experiment	n_s	$\deg(M)$	Z	N_{it}	T_{CPU}	ρ	cond
Random 100	27	4	20	9	49 + 45	0.26	2.88
Random 100	343	1	24	9	32 + 33	0.23	2.64
Random 1000	27	4	20	20	49 + 90	0.51	12.5
Random 1000	343	1	24	19	31 + 63	0.50	11.0

TABLE 4

Experimental results for randomly distributed Young modulus

Turbine (Fig 6), 41040 nodes, 123120 degrees of freedom							
Experiment	n_s	$\deg(M)$	Z	N_{it}	T_{CPU}	ρ	cond
Turbine	1730	1	22	12	85 + 240	0.39	16.0

TABLE 5

Results for turbine (Boundary conditions:Dirichlet, on a set of zero measure)

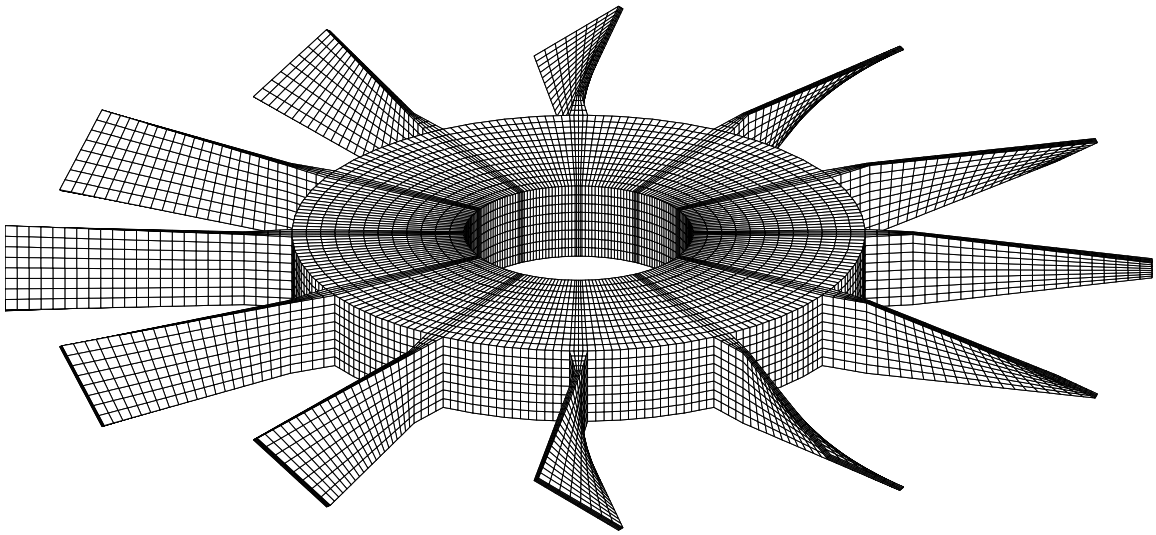

FIG. 3. *Turbine, 41040 nodes, 123120 degrees of freedom. Courtesy of Charbel Farhat, University of Colorado at Boulder.*