UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

Detection of Strong Coupling in Algebraic Multigrid Solvers

Tony F. Chan
Petr Vanek

March 2000

CAM Report 00-10

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90095-1555

http://www.math.ucla.edu/applied/cam/index.html

Detection of Strong Coupling in Algebraic Multigrid Solvers.[1]

Tony F Chan and Petr Vanek

Dept of Mathematics, UCLA

Abstract. Based on abstract convergence theory for the smoothed aggregation multigrid method [6], we present a new method for detecting strong connections (couplings) in matrices obtained by discretization (and subsequent coarsening) of elliptic problems. Although the coupling evaluation given here has been derived in connection with smoothed aggregation algorithm, it is fully applicable in any AMG method, providing zero-energy modes are available in the solver.

1 Introduction

Algebraic multigrid methods have become quite popular[5,16,15,6,3] recently. Based on convergence results for one of these methods, namely the smoothed aggregation multigrid method [6], we present a new method for detecting strong connections (couplings) in matrices obtained by discretization (and subsequent coarsening) of elliptic problems. Although the coupling evaluation given here has been derived using smoothed aggregation abstract convergence bounds established in [7,8,14], it is fully applicable in any AMG method, providing zero-energy modes are available in the solver.

The ideal way of detecting strong connections rests on analyzing the local stiffness matrices. If i, j are two degrees of freedom associated with one element T, their coupling can be reliably evaluated using simple formula of [5]

$$\text{coup}(i,j) = \frac{|a_{ij}|}{\sqrt{a_{ii}\, a_{jj}}}, \tag{1}$$

where $A_T = \{a_{ij}\}$ is a local stiffness matrix corresponding to the element T. Since (1) defines the energy cosine of i-th and $j-$th basis function, the criterion can be easily extended for evaluating the nodal coupling in case of nonscalar problems (e.g. using energy cosine of nodal spaces.) Providing the local stiffness matrices are available, (1) evaluates reliably strong connections on the finest level. For coarser levels however, the notion of local stiffness matrix can be reproduced only under strong geometrical restrictions on the coarsening that are often violated by algebraic methods. For this reason, finding a reliable way for detecting strong connections using the global (or coarsened) stiffness matrix remains a challenging problem.

[1] Supported by NSF under Contract ASC-9720257, NASA Ames under contract NAG2-1238 and Sandia Nat'l Lab under contract LG-4440.

Our approach is based on a heuristic interpretation of the algebraic assumptions of the smoothed aggregation abstract convergence theory [14]. In this method, the final prolongator I_{l+1}^l ($l+1$ is the coarser level) is constructed as a product $I_{l+1}^l = S_l P_l$ of a so-called *prolongator smoother* S_l and a *tentative prolongator* P_{l+1}^l, where both components S_l and P_l are determined by the identified strong coupling adjacency. If the identified strong connection adjacency is too sparse, it results in a powerful (but expensive) tentative prolongator P_l, (identity in an extreme case) but an inefficient prolongator smoother S_l. The opposite extreme gives S_l being a polynomial in A_l (the ideal case) but a tentative prolongator P_{l+1}^l unable to reflect any anisotropic behavior. Our proposed technique establishes a "trade-off" between a weak approximation condition for P_{l+1}^l and a filtering condition on S_l. Under certain restrictions, constants in both conditions can be evaluated computationally, given any proposed strong connection adjacency.

2 Coupling detection algorithm

We describe a strong adjacency detection based on abstract convergence estimates for smoothed aggregations methods established in [14]. Although the coupling evaluation given here has been derived using smoothed aggregation abstract convergence bounds established in [7,8,14], it is fully applicable in any AMG method, providing zero-energy modes are available in the solver. Before formulating the algorithm suitable for solving general nonscalar elliptic problems, we first give its simplified version suitable for scalar problems.

2.1 Coupling detection for scalar problems

This section provides an algorithm for detecting strong connections for matrices obtained by discretizing scalar problems with one dimensional space of zero-energy modes. In other words, we assume that each row of the matrix A corresponds to one node and zero-energy modes are given by a *single vector* $\mathbf{b} \in \Re^n$.

We split the algorithm into two separate parts: a) evaluation function, b) the detection of strong connections itself. When solving anisotropic problems by smoothed aggregation method, the aggregates must be formed in accordance with strong connections. To prevent an excessive fill-in of coarse-level matrices, the prolongator smoothers are carefully *filtered* so that the nonzero structure of the prolongator smoother is given by the adjacency of strong connections that is used for generating the aggregates (see Sect. 4.) In the context of this method, the evaluation function measures "how drastic" the filtering of the $i-$th row of the prolongator smoother corresponds to each particular list of strongly coupled neighbours of i. The case of $\mathcal{N}_i = \{j : a_{ij} \neq 0\}$ corresponds to no filtering at all, and the evaluation function returns zero.

The detection of strong connections itself is then a simple algorithm that makes decisions based on results of evaluation function for candidate lists of nodes strongly connected to a center.

Algorithm 1 (EVALUATION FUNCTION). *Given an $n \times n$ matrix A, vector of zero-energy modes $\mathbf{b} \in \Re^n$, row number $i \in 1, \ldots, n$ and a candidate list of strongly coupled neighbors*

$$\{i\} \subset \mathcal{N} \subset \{j : a_{ij} \neq 0\}, \tag{2}$$

return

$$E(i, \mathcal{N}) = \frac{\sum_{j \in \mathcal{N}} a_{ij} b_j}{\left(\sum_{j \in \mathcal{N}} b_j\right)^{1/2}}. \tag{3}$$

Aside from the essential boundary conditions, the local kernel $\mathbf{b}$ satisfies $\sum_{j=1}^{n} a_{ij} b_j = 0$. The evaluation function (3) then indicates how close $\mathbf{b}$ is to the kernel of matrix A with all entries corresponding to weak connections replaced by zeroes.

Using the above evaluation function, the strong adjacency can be obtained as follows:

Algorithm 2 (STRONG ADJACENCY CONSTRUCTION). *Given an $n \times n$ matrix A, an upper bound $\bar{\lambda} \geq \varrho(A)$, vector of zero-energy modes $b \in \Re^n$ and a threshold $\alpha \in [0, 1]$, find strong connection adjacency organized as a $0/1$ matrix*

$$N = \{n_{ij}\}, \quad n_{ij} = \begin{cases} 1 \text{ if } j \text{ is strongly coupled to } i, \\ 0 \text{ otherwise} \end{cases}$$

as follows:
for $i = 1, \ldots, n$
1. *Search among all lists satisfying (2); find the smallest list $\mathcal{N}$ such that*

$$E(i, \mathcal{N}) \leq \alpha \bar{\lambda}. \tag{4}$$

and denote the resulting list by $\hat{\mathcal{N}}_i$.
2. *Create the $i-th$ row of N by*

$$n_{ij} = \begin{cases} 1 \text{ if } j \in \hat{\mathcal{N}}_i, \\ 0 \text{ otherwise} \end{cases}$$

end for

Step 1. of Alg. 2 consists of the loop over all subsets of the list of all nonzeroes in the row i that contain i. The result of this inner loop is the list of nodes that are strongly coupled to the node i. It requires an access to $i-$th row of the matrix A and to zero-energy modes vecor b. The outer loop can be

performed in parallel.

Outside the smoothed aggregation framework, the above detection method can be interpreted as a systematic way of seeking weak connections hidden in combinations of positive and negative off-diagonal entries using actions of the stiffness matrix on zero-energy modes. The denominator in (3) and the upper bound $\bar{\lambda} \geq \varrho(A)$ on the right-hand side of (4) assure a scaling invariance with respect to both A and $\mathbf{b}$. Fig. 1 shows the situation where criterion (1) is very sensitive to the magnitude of the threshold since it considerers one off-diagonal entry only. Assume a) zero-energy modes are formed by a vector of ones, i.e. $\sum_j a_{ij} = 0$, b) all off-diagonal entries are large in magnitude, c) sum of off-diagonal entries corresponding to connections of the center with vertices located outside the ellipse is close to zero.

This situation can be observed in the case of the Poisson equation discretized using bilinear quadrilateral elements on the rectangular grid. The rectangular elements stretched to a $1 : 10$ aspect ratio yield the coefficient stencil

$$A = \begin{bmatrix} -1 & 1.9 & -1 \\ -3.9 & 8 & -3.9 \\ -1 & 1.9 & -1 \end{bmatrix}, \tag{5}$$

where the coefficients -3.9 correspond to a side of relative length 1 while 1.9 corresponds to a side of length 10 [15].

The criterion described in this section correctly selects nodes contained in the ellipse as strongly coupled neighborhood for a wide range of thresholds.

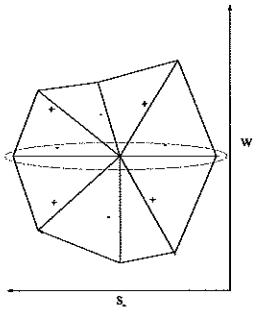

Fig. 1.

2.2 Generalization for noscalar problems

In this section, we generalize Alg. 2 for nonscalar problems with multidimensional space of zero-energy modes.

Throughout this section, we assume that the basis vectors of zero-energy modes are supplied as columns of matrix the $n \times r$ matrix B ($n = ord(A)$), and both the matrix A and the zero-energy modes B are organized in blocks corresponding to N nodes as follows:

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \dots & A_{1N} \\ A_{21} & A_{22} & A_{23} & \dots & A_{2N} \\ & & \dots & & \\ & & \dots & & \\ A_{N1} & A_{N2} & A_{N3} & \dots & A_{NN} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \\ \cdot \\ \cdot \\ B_N \end{bmatrix}. \tag{6}$$

Since the notion of "node" in the geometrical sense is not always well-defined for coarse-level matrices obtained by algebraic coarsening (see Alg. 4), nodes are understood here as small clusters of degrees of freedom forming a disjoint covering of the set of all degrees of freedom $\{1, \dots, n\}$. For simplicity, we assume that each of the nodes consists of the same number of degrees of freedom n_f. Hence the blocks A_{ij} are $n_f \times n_f$ square matrices and the blocks B_i are $n_f \times r$ rectangular matrices, where r is the dimension of the space of zero-energy modes.

Using this notation, the evaluation function Alg. 1 can be generalized as follows:

Algorithm 3 (EVALUATION FUNCTION). *Given an $n \times n$ matrix A, $n \times r$ matrix of zero-energy modes B organized as in (6), node number $i \in 1, \dots, N$ and a candidate list of strongly coupled nodes $\mathcal{N}$ satisfying*

$$\{i\} \subset \mathcal{N} \subset \{j : A_{ij} \neq 0\}, \tag{7}$$

return $E(i, \mathcal{N}) \geq 0$ evaluated as follows:

1. *Set m to be the number of nodes in $\mathcal{N}$ and create the the block selections*

$$U_A = \begin{bmatrix} A_{j_1,i} \\ A_{j_1,i} \\ \cdot \\ \cdot \\ A_{j_m,i} \end{bmatrix}, \quad U_B = \begin{bmatrix} B_{j_1} \\ B_{j_1} \\ \cdot \\ \cdot \\ B_{j_m} \end{bmatrix}, \quad j_k \in \mathcal{N}.$$

2. *Orthonormalize columns of matrix U_B, store the result in $\tilde{U}_B$.*
3. *Evaluate $E(i, \mathcal{N}) = \varrho(U_A^T \tilde{U}_B)$.*

Note that the orthonormalization step can be conveniently performed using QR decomposition.

Using the above evaluation function, the strong connection adjacency can be obtained by Alg. 2 with n replaced by N and (2) by (7).

3 Smoothed aggregation multigrid method

The smoothed aggregation multigrid method (proposed by Vaněk in [10,9] and further developed in [6,11,7]), builds the prolongator in the form

$$I_{l+1}^l = S_l P_{l+1}^l, \tag{8}$$

where P_{l+1}^l is a very simple *tentative prolongator* satisfying

$$B^l = P_{l+1}^l B^{l+1} \tag{9}$$

where B^l is a matrix formed by columns of zero-energy modes basis vectors and S_l is a Richardson-type *prolongator smoother* derived from matrix A_l (or its perturbation), e.g.,

$$S_l = I - \frac{\omega}{\varrho(A_l)} A_l. \tag{10}$$

Note that on the finest level, the zero-energy modes B^1 must be given. To satisfy (9), we build simultaneously P_2^1 and B^2 so that $B^1 = P_2^1 B^2$. Then using B^2, we construct P_3^2 and B^3, etc. For details, see Alg. 4.

The purpose of the prolongator smoother S_l is to minimize $\varrho(A_{l+1})$, where $A_{l+1} = (I_{l+1}^l)^T A_l I_{l+1}^l$. Simple pointwise smoothers eliminate efficiently high-energy errors. The prolongator smoother (10) is an error propagation operator of a Richardson-type iteration. By applying it to the range of P_{l+1}^l, we suppress high-energy vectors, which in turn reduces the "maximal energy measure" $\varrho(A_{l+1})$.

It remains to specify the tentative prolongator P_{l+1}^l. Assume B^l is available and denote the number of its columns by r. Our goal is to create the tentative prolongator P_{l+1}^l and the coarse-level representation B^{l+1} of B^l satisfying (9).

Our construction is based on the supernodes aggregation concept. On each level, degrees of freedom are organized in small disjoint clusters called supernodes. On the finest level, these clusters have to be specified, e.g., as the sets of degrees of freedom associated with the finite element vertices, the coarse level supernodes are then created by our aggregation algorithm. The prolongator P_{l+1}^l is constructed from a given system of aggregates $\{\mathcal{A}_i^l\}_{i=1}^{N_l}$ that forms a disjoint covering of level l supernodes. A simple greedy algorithm for generating aggregates based on the structure of the matrix A_l is given in [6]. The property (9) is enforced aggregate by aggregate; columns of P_{l+1}^l associated with the aggregate $\mathcal{A}_i^l$ are formed by orthonormalized restrictions of the columns of B^l onto the aggregate $\mathcal{A}_i^l$. For each aggregate, such a construction gives rise to r degrees of freedom on the coarse level forming a coarse level supernode.

The detailed algorithm follows. For ease of presentation, we assume that the fine level supernodes are numbered by consecutive numbers within each aggregate. This assumption can be easily avoided by renumbering.

Algorithm 4. *For the given system of aggregates* $\{\mathcal{A}_i^l\}_{i=1}^{N_l}$ *and the* $n_l \times r$ *matrix* B^l *satisfying* $P_l^1 B^l = B^1$, *we create a prolongator* P_{l+1}^l, *a matrix* B^{l+1} *satisfying (9) and supernodes on level* $l+1$ *as follows:*

1. *Let* d_i *denote the number of degrees of freedom associated with aggregate* $\mathcal{A}_i^l$. *Partition the* $n_l \times r$ *matrix* B^l *into blocks* B_i^l *of size* $d_i \times r$, $i = 1, \ldots, N_l$, *each corresponding to the set of degrees of freedom on an aggregate* $\mathcal{A}_i^l$ *(see Fig. 2).*
2. *Decompose* $B_i^l = Q_i^l R_i^l$, *where* Q_i^l *is an* $d_i \times r$ *orthogonal matrix, and* R_i^l *is an* $r \times r$ *upper triangular matrix.*
3. *Using the blocks* Q_i^l, $i = 1, \ldots, N_l$, *create the prolongator* P_{l+1}^l *as shown by Fig. 2.*
4. *Create* B^{l+1} *consisting of the blocks* R_i^l, $i = 1, \ldots, N_l$, *(see Fig. 2.)*
5. *For each aggregate* $\mathcal{A}_i^l$, *the coarsening gives rise to* r *degrees of freedom on the coarse level (the* $i-th$ *block column of* P_{l+1}^l*). These degrees of freedom define the* $i-th$ *coarse level supernode.*

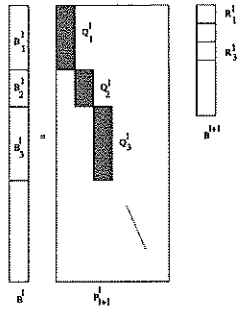

Fig. 2. The construction of a tentative prolongator P_{l+1}^l and B^{l+1}

4 Smoothed aggregation multigrid with perturbed prolongator smoothers

Solving anisotropic problems calls for a coarsening scheme that follows strong connections (semicoarsening.) In [7], we analyzed the algorithm that uses prolongator smoothers S_l being the error propagation operators of Jacobi-type smoothers derived from the matrices A_l. Such prolongator smoothers, when used together with tentative prolongators that follow strong connections, tend to cause excessive fill-in of coarse-level matrices. To eliminate this drawback, we use ([14]) prolongator smoothers (10) with A_l^F in the place of A_l, where $A_l^F \approx A_l$ is a *filtered matrix.*

In what follows, we give a simplified version of a filtering algorithm that creates a filtered matrix A^F of a given nonzero structure N_l. Throughout this section, we assume that the zero-energy modes $\mathbf{b}$ are given by a single vector. The more general algorithm is given in [14].

Given a list $\mathcal{N} \subset \{1, \ldots, n\}$, define a "filter" $F(\mathcal{N}) : \Re^n \to \Re^n$ that drops all nonzeroes at positions $i \notin \mathcal{N}$, i.e.,

$$F_{\mathcal{N}} : \mathbf{x} \in \Re^n \mapsto \mathbf{y} \in \Re^n, \quad y_i = \begin{cases} x_i & \text{if } i \in \mathcal{N} \\ 0 & \text{otherwise} \end{cases}, \tag{11}$$

and the projection $Q_{\mathcal{N}}$ onto $\{\alpha F_{\mathcal{N}}\mathbf{b},\ \alpha \in \Re^1\}$ by

$$Q_{\mathcal{N}} = I - F_{\mathcal{N}}\mathbf{b}\left((F_{\mathcal{N}}\mathbf{b})^T(F_{\mathcal{N}}\mathbf{b})\right)^{-1}(F_{\mathcal{N}}\mathbf{b})^T. \tag{12}$$

Using above definitions, the filtering algorithm can be written down as follows:

Algorithm 5 (FILTERING). *Given the matrix A and the strong connection adjacency $\mathcal{N}$, create a filtered matrix A^F as follows:*
For all nodes $i = 1$ to n do
1. *Define the list $\mathcal{N} = \{j : N_{ij} = 1\}$,*
2. *filter $i-$th row of the matrix A:* $\mathbf{row}_i^T(A^F) \leftarrow Q_{\mathcal{N}}F_{\mathcal{N}}\,\mathbf{row}_i^T(A)$

end for

The operator $Q_{\mathcal{N}}$ in Step 2 ensures that the zero-energy mode is preserved for the filtered matrix A^F.

The abstract convergence bounds for smoothed aggregation method with perturbed prolongator smoothers are established in [14]. In our notation, a simplified version of a key abstract result can be written down as follows:

Lemma 6. *Let $\bar{\lambda}_l \geq \varrho(A_l)$, $l = 1, \ldots, L$. We assume that:*

1. *There is a sequence of mappings $\tilde{Q}_l : \Re^{n_1} \to \Re^{n_l}$, $l = 2, \ldots, L$, and a positive constant C_A such that for all levels $l < L$ and all finest-level vectors $\mathbf{u} \in \Re^{n_1}$,*

$$\|(I - P_{l+1}^1\tilde{Q}_{l+1})\mathbf{u}\|_{\Re^{n_1}} \leq \frac{C_A}{\sqrt{\bar{\lambda}_l}}\|\mathbf{u}\|_A. \tag{13}$$

2. *The matrix $D_l = A_l - A_l^F$ satisfies $D_l P_L^l = 0$, and*

$$\sum_{j=l+1}^{L-1}\left(\frac{1}{\bar{\lambda}_j}\right)^{1/2}\|D_l P_j^l\| \leq C_F. \tag{14}$$

Here, $\|D_l P_j^l\| = \sqrt{\lambda_{\max}\left((P_j^l)^T D_l^T D_l P_j^l\right)}$ denotes the operator matrix norm induced by Euclidean vector norms.

Then the A-norm convergence rate of the smoothed aggregation method is bounded by $1 - C/L^3$.

The assumption (13) is a weak approximation property for disaggregated vectors. The prolongator smoothers enter the condition (13) only through the scaling factor $1/\sqrt{\bar{\lambda}_l}$ on its right-hand side. The role of the prolongator smoothers is to enforce "smoothness" of the coarse spaces by making the values of $\bar{\lambda}_l$ small. Obviously, a smaller $\bar{\lambda}_l$ allows the approximation condition (13) to be satisfied with a smaller constant C_A.

The condition (14) specifies how much A_l can be perturbed while still preserving properties of a good prolongator smoother. It has a multiresolution character; the operator norm $\|D_l P_j^l\|$ measures the effect of a filtering of A_l on the coarser levels $j > l$. Obviously, the weight $1/\bar{\lambda}_j$ increases as j grows; hence (14) postulates that $\|D_l P_j^l\|$ should be small for large j.

The primary goal of detection strong connections is to keep the constant C_A in (13) small. For this reason, the most straightforward way of introducing the notion of strong coupling is based on (13); correct strong connections are simply such that aggregates following those connections result in small C_A. This approach naturally leads to an algorithm minimizing the constant in the discrete Poincaré inequality on each of aggregates. Since the Poincaré constant is inversely proportional to the smallest eigenvalue of a corresponding local stiffness matrix, such an approach is problematic if only a global matrix is available. Therefore another approach is needed.

The reason why (14) is difficult to satisfy for anisotropic problems stems from the fact that anisotropic problems contain low energy modes that are local in a graph sense. Different approaches for detecting strong connections can be therefore based on determining a part of the stencil (nonzero structure) of A_l where those local low energy modes are not local anymore. The adjacency given by the nonzero structure of A_l introduces a notion of distance on the set of nodes. By removing graph edges we enlarge the nodal distance; the correct strong connections then correspond to a *reduced adjacency* where high energy modes are still local but low energy modes that are local in the original adjacency are not local in the reduced one anymore. This leads to a splitting

$$A_l = A_l^F + D_l, \tag{15}$$

where A_l^F is as sparse as possible while retaining all high-energy modes of A_l. The condition (14) controls how much (and in which sense) A_l can be filtered and still give a reasonable prolongator smoother. Since properties of smoothers are determined by high-energy modes, the filtering condition (14) provides a tool for splitting (15).

In what follows, we clarify the relationship of Alg. 2 with the condition (14). More precisely, we show – though not fully rigorously – that for $j >> l$, Alg. 2 enforces:

$$\|D_l P_j^l\| \leq C\alpha\bar{\lambda}_l, \tag{16}$$

where α is a threshold used in (4). Note that the value of α we have in mind is quite small (e.g. 0.01.)

We start with proving a simple auxiliary Lemma.

Lemma 7. *Let A^F be the filtered matrix obtained by using Alg. 5 from the matrix A and a strong connection adjacency N. Then the difference $D = A - A^F$ satisfies*

$$\mathbf{row}_i(D)F_{\mathcal{N}_i}\mathbf{b} = \mathbf{row}_i(A)F_{\mathcal{N}_i}\mathbf{b}, \quad \textit{where } \mathcal{N}_i \equiv \{j : N_{ij} \neq 0\}.$$

Proof. The definition of D together with (12) give

$$\mathbf{row}_i(D)F_{\mathcal{N}_i}\mathbf{b} = \langle \mathbf{row}_i^T(A),\, F_{\mathcal{N}_i}\mathbf{b}\rangle \;-\; \langle Q_{\mathcal{N}_i}F_{\mathcal{N}_i}\mathbf{row}_i^T(A),\, F_{\mathcal{N}_i}\mathbf{b}\rangle,$$

where by (12), $\langle Q_{\mathcal{N}_i}F_{\mathcal{N}_i}\mathbf{row}_i^T(A),\, F_{\mathcal{N}_i}\mathbf{b}\rangle = \langle F_{\mathcal{N}_i}\mathbf{row}_i^T(A),\, Q_{\mathcal{N}_i}F_{\mathcal{N}_i}\mathbf{b}\rangle = 0$, completing the proof.

By construction, (see Alg. 4) columns of the *composite tentative prolongator* P_j^l are orthonormal vectors formed by (scaled) zero-energy modes restricted to the disjoint *composite aggregates* corresponding to the disaggregation from the level $j > l$ to the finer level l. Example of such a composite aggregate in case of semicoarsening is given by Fig. 3. Here, the composite aggregate $\mathcal{A}$ is formed by nodes connected in strong connection adjacency in such a way that for all nodes i except the end-points A and B, their "strong" neighbourhoods $\mathcal{N}_i$ are also contained in the aggregate, i.e.

$$\mathcal{N}_i \subset \mathcal{A} \quad \text{for all } i \neq A, B. \tag{17}$$

Denoting the $i-$th canonical basis vector of $\Re^{n_j}$ by $\mathbf{e}_i$, it follows by well-known arguments that

$$\|D_l P_j^l\| \leq C \max_{i \in 1,\ldots,n_j} \|D_l P_j^l \mathbf{e}_i\|, \tag{18}$$

where C depends on overlaps of columns of $D_l P_j^l$ (vectors $D_l P_j^l \mathbf{e}_i$.) Since the matrices D_l and A_l have the same nonzero structure and columns of P_j^l "do not overlap", it is realistic to assume that C is small. Trivially,

$$\|D_l P_j^l \mathbf{e}_k\|^2 = (D_l P_j^l \mathbf{e}_k)_A^2 + (D_l P_j^l \mathbf{e}_k)_B^2 + \sum_{i \in \text{inter}} (D_l P_j^l \mathbf{e}_k)_i^2.$$

We make a (realistic) heuristical assumption that for a large aggregate, the contribution from the interior points dominates and we can neglect the boundary nodes A and B. In other words, we assume that

$$\|D_l P_j^l \mathbf{e}_k\|^2 \leq C \sum_{i \in \text{inter}} (D_l P_j^l \mathbf{e}_k)_i^2. \tag{19}$$

Then we estimate using the fact the $P_j^l \mathbf{e}_k$ is a zero-energy mode of $\mathcal{A}$, inclusion (17), Lemma 7, condition (4) and bounded overlaps of $dist = 1$ neighborhoods $\mathcal{N}_i$ as follows:

$$\sum_{i \in \text{inter}} (D_l P_j^l \mathbf{e}_k)_i^2 = \sum_{i \in \text{inter}} \langle \mathbf{row}_i^T(D_l),\, P_j^l \mathbf{e}_k\rangle^2 = \sum_{i \in \text{inter}} \langle \mathbf{row}_i^T(A_l),\, F_{\mathcal{N}_i} P_j^l \mathbf{e}_k\rangle$$

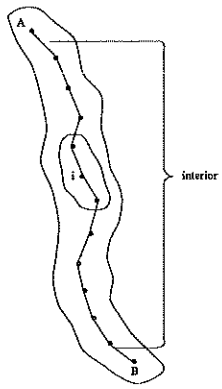

Fig. 3. Composite aggregate in case of semicoarsening

$$= \sum_{i \in \text{inter}} E^2(i, \mathcal{N}_i) \|F_{\mathcal{N}_i} P_j^l \mathbf{e}_k\|^2 \leq \alpha^2 \bar{\lambda}_l^2 \sum_{i \in \text{inter}} \|F_{\mathcal{N}_i} P_j^l \mathbf{e}_k\|^2$$
$$\leq C\alpha^2 \bar{\lambda}_l^2 \|P_j^l \mathbf{e}_k\|^2.$$

Since $(P_j^l)^T P_j^l$ is an identity matrix, we have $\|P_j^l \mathbf{e}_k\| = 1$, and therefore

$$\|D_l P_j^l \mathbf{e}_k\|^2 \leq \alpha^2 \bar{\lambda}_l^2 + (D_l P_j^l \mathbf{e}_k)_A^2 + (D_l P_j^l \mathbf{e}_k)_B^2.$$

The last inequality together with (19) and (18) gives (16).

Thus we see that the strong coupling detection algorithm implicitly enforces ((14)), which is one of the two key conditions in the convergence theory for the smoothed aggregation multigrid method.

5 Numerical experiments

We now give two examples illustrating the behavior of the proposed coupling detection algorithm in connection with smoothed aggregation multigrid V-cycle method. In both experiments, multigrid is used as a preconditioner for conjugate gradients.

The first example is a model Poisson problem discretized using bilinear quadrilateral elements on the rectangular grid with elements stretched to a $1 : 10$ aspect ratio. The discretization gives matrix A with stencil (5). Difficulties with coupling detection for (5) are discussed in Sect. 2.

The second experiment has been done with an industrial structural mechanics problem on an unstructured mesh discretized using QUAD4 (shell), HEXA (solid) and BEAM elements. The maximal aspect ratio exceeded 100.

Results are summed up in the following table.

	Poisson problem	Industrial problem
Dofs:	160000	612863
Total memory used:	23MB	665.6MB
Memory used/data size	241%	331.6%
threshold α	0.01	0.01
Achieved accuracy:	5.4812×10^{-7}	1.5394×10^{-6}
Iterations done:	6	23
conv rate:	0.07	0.558
setup time:	$20s$	$339s$
iteration time:	$10s$	$424s$
Architecture:	2×R8000, 95MHz	2×R8000, 95MHz

The convergence rate of 0.07 for the model problem is quite satisfactory. Even though the convergence rate for the industrial problem is quite a bit larger, it is in fact quite good considering the difficulty of the problem.

References

1. 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.
2. A. Brandt, *Algebraic multigrid theory: The symmetric case*, Appl. Math. Comput., 19 (1986), pp. 23–56.
3. T. F. Chan, J, Xu, and L. Zikatanov, *An agglomeration multigrid for unstructured meshes.*, In: Domain Decomposition Methods 10, (Proceedings of the tenth international conference on domain decomposition methods) Mandel, Farhat, Cai Eds., AMS 1998.
4. T. F. Chan, S. Go, and L. Zikatanov, *Lecture Notes on Multilevel Methods for Elliptic Problems on Unstructured Grids*, UCLA CAM Report 97-11, March 1997. Lectures notes for the lecture series "Computational Fluid Dynamics", von Karman Inst., Belgium, March 3-7, 1997. An abridged version has been published as CAM Report 97-36, August, 1997 and appeared in "Computational Fluid Dynamics Review 1997", Hafez and Oshima (eds.), Wiley.
5. ——, *Algebraic multigrid (AMG)*, in Multigrid Methods, S. F. McCormick, ed., vol. 3 of Frontiers in Applied Mathematics, SIAM, Philadelphia, PA, 1987, pp. 73–130.
6. P. Vaněk, J. Mandel, and M. Brezina, *Algebraic multigrid based on smoothed aggregation for second and fourth order problems*, Computing, 56 (1996), pp. 179–196.
7. P. Vaněk, M. Brezina, and J. Mandel, *Convergence of Algebraic Multigrid Based on Smoothed Aggregation*, Submitted to Num. Math.
8. P. Vaněk, A. Janka, and H. Guillard, *Convergence of Petrov-Galerkin Smoothed Aggregation Method*, To appear.
9. P. Vaněk, *Fast multigrid solver.* Applications of Mathematics, to appear.
10. *Acceleration of convergence of a two-level algorithm by smoothing trarsfer operator*, Applications of Mathematics, 37 (1992), pp. 265–274.

11. P. VANĚK, M. BREZINA, AND R. TEZAUR, *Two-Grid Method for Linear Elasticity on Unstructured Meshes*, To appear in SIAM J. Sci. Comp.
12. J. MANDEL, M. BREZINA, AND P. VANĚK, *Energy Optimization of Algebraic Multigrid Bases*, To appear in Computing
13. W. L. WAN, T. F. CHAN, AND B. SMITH, *An energy-minimizing interpolation for robust multigrid methods*, UCLA CAM Report 98-6, Department of Mathematics, UCLA, February 1998. To appear in SIAM J. Sci. Comp.
14. P. VANĚK, J. MANDEL, T.F. CHAN, *Smoothed aggregation method for anisotropic problems*, in preparation
15. M. BREZINA, A. J. CLEARY, R. D. FALGOUT, V. E. HENSON, J. E. JONES, T. A. MANTEUFFEL, S. F. MCCORMIC, J. W. RUGE, *Algebraic multigrid based on element interpolation* Tech. report UCRL - JC - 131752, LLNL, 1998
16. A. BRANDT, *General highly accurate algebraic coarsening schemes*, GMC tech. report 13