M a t h e m a t i c a
B a l k a n i c a

New Series Vol. 19, 2005, Fasc. 3-4

Robust multilevel preconditioners for second order elliptic equations and systems of equations

O. Axelsson[1], *L. Laayouni*[2] *and S. Margenov*[3]

Presented by

Preconditioners based on various multilevel extensions of two-level finite element methods lead to iterative methods which have an optimal order computational complexity with respect to the size (or discretization parameter) of the system. The methods can be on block matrix factorized form, recursively extended via certain matrix polynomial approximations of the arising Schur complement matrices or on additive, i.e. block diagonal form using stabilizations of the condition number at certain levels. The resulting spectral equivalence holds uniformly with respect to jumps in the coefficients of the differential operator and for arbitrary triangulations. Such methods were first presented by Axelsson and Vassilevski in the late 80s.

An important part of the algorithm is the treatment of the systems with the diagonal block matrix, which arises on each finer level in a recursive refinement method and corresponds to the added degrees of freedom on that level. This block is well-conditioned for model type problems but becomes increasingly ill-conditioned when the coefficient matrix becomes more anisotropic or, equivalently, when the mesh aspect ratio increases.

In the paper some methods are presented to approximate this matrix also leading to a preconditioner with spectral equivalence bounds which hold uniformly with respect to both the problem and discretization parameters. The same holds therefore also for the preconditioner to the global matrix. Such uniform bounds have not been achieved by other methods.

Key Words: multilevel preconditioners, partial differential equations and systems, hierarchical basis, optimal order preconditioners, uniform bounds.

1. Introduction

In many problems in mathematical modeling in natural sciences, engineering and in other areas as well where second order boundary value problems must be solved numerically, large scale linear systems arise which furthermore, frequently must be solved a number of times for each modeling case. Often, the arising systems are severely ill-conditioned due to some problem parameters taking near certain limit values. Examples of such parameters are

ratio of coefficient jumps, anisotropy, aspect ratio of the mesh and domain geometry, Poisson ratio for nearly incompressible materials etc. Furthermore, the condition number may increase rapidly when the discretization mesh is refined (due to both a smaller mesh parameter and possible irregularity of the mesh elements). As has been pointed out by many authors, see e.g. [10], the classical $V-$cycle multigrid method is inefficient in handling such problems where the usual regularity and approximations assumptions does not hold. In some special problems, a line smoother can be used to overcome this, but there is no multigrid theory available to handle more general problems. Therefore, instead, in finding a good solution method one should preferably search for efficient preconditioners for the, parameter free, conjugate gradient iterative solution method.

The method to be presented is a block matrix approximate factorization preconditioning of the algebraic multilevel iteration, AMLI type. It is based on two or multilevel finite element meshes and can handle arbitrary coefficient jumps on the coarsest mesh used and also ratio of anisotropy, using newly developed finite element based preconditioners for the block corresponding to the added nodes. The condition number is bounded for any ratio of coefficient jumps and anisotropy. The discontinuity of coefficients is assumed to occur across element edges of the coarsest mesh. It turns out that this can be chosen as a quite fine mesh itself, which permits modelling of problems with many different materials.

Algebraic multilevel preconditioners were first presented in [7, 8] and are multilevel extensions of the two-level methods in [9] and [4], see also [3]. Here block matrix approximate factorizations were considered and it was shown that by recursively extending the two level method using certain matrix polynomial approximations of the arising Schur complement matrices, one can derive a preconditioning with a condition number which is bounded independently on the number of levels and on jumps in the coefficients, assuming the coarsest mesh used had no jumps inside any element. As, in practice, one can use a coarsest mesh which is still quite fine, a significant number of jumps in coefficients, i.e. different materials in the physical model can be allowed. Similarly, preconditioners in additive form, i.e., using block diagonal preconditioners, but with stabilization at certain levels (see [2]) were developed with the same properties.

In the above methods the block matrix corresponding to the on each level added degrees of freedom gets increasingly ill-conditioned with increasing degree of anisotropy. Until recently, no efficient generally applicable method to handle this problem has been given. In [12], a preconditioner to this matrix in multiplicative form and in [6] an element by element preconditioner in additive form were suggested. The first method considered either x- or y-dominated anisotropy while the latter considered the general case with arbitrary coefficients in the differential operator. It was shown that the preconditioner is spectrally equivalent to the given matrix with bounds which hold uniformly in the number of levels

and in the coefficients of the operator. A preliminary idea of the present work has been considered in [5], in particular for a scalar case.
In the present paper we consider possible improvements of these methods. In particular it is shown that for the considered new element by element preconditioners, in multiplicative form or in block diagonal form, improvements in the condition number can be achieved. Furthermore, the results are here extended to elliptic systems of differential equations. The analysis of the computational complexity related to the constructed preconditioners is also considered for different model problems.
The remainder of the paper is organized as follows: In section 2 we survey shortly the main results for multiplicative and additive preconditioners in algebraic multilevel form. Section 3 deals with the construction of an additive (element by element) preconditioner for the block diagonal matrix corresponding to the added degrees of freedom on each level, while section 4 presents a preconditioner on multiplicative form for general systems of partial differential equations and a discussion on solutions of the algebraic systems arising when using those kind of preconditioners. Finally, in section 5 we present an extension of multiplicative preconditioners to three dimensional problems.
The following notation is used throughout the paper: $A \geq B$ means that $A - B$ is positive semidefinite.

2. General estimates of condition numbers for two-level methods

Consider the elliptic problem

$$\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) = f \qquad \text{in} \qquad \Omega,$$

where Ω is a polyhedral domain, with proper boundary conditions on $\partial\Omega$. For systems of PDEs of dimension d, u is a d-dimensional vector and a_{ij} are $d \times d$ matrices. Its variational formulation is : seek $u \in H_g^1(\Omega)$ such that

$$a(u, v) = \int_\Omega fv \qquad \text{for all } v \in H_0^1(\Omega),$$

where

$$a(u, v) = \int_\Omega \sum_{i,j} a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j}, \tag{1}$$

and where the function spaces $H_g^1(\Omega)$ and $H_0^1(\Omega)$ incorporate the Dirichlet portion of the boundary conditions. Further, the matrix $[a_{ij}]$ is assumed to be symmetric and positive definite. The domain of definition is partitioned in finite elements, such as triangles (d=2), tetrahedrons or prisms (d=3) and on each element we use piecewise linear finite element basis functions.

Remark 2.1 *For the analysis of certain uniform bounds for the finite element method such as bounds for the constant γ (see below), one can consider (1) for the reference triangle and arbitrary coefficients $[a_{ij}]$, or alternatively, the operator $-\Delta$ and an arbitrary triangle e. For details, see e.g. [2].*

Condition numbers for two-level finite element matrices
Each finite element is partitioned in 2^d elements of equal volume and the node set is partitioned in two sets, the old (coarse mesh) and the new (added) ones. The finite element matrix is partitioned correspondingly in 2×2 blocks

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{matrix} \text{(added node points)} \\ \text{(coarse mesh node points)}, \end{matrix}$$

which is the two level matrix, where A_{11} and A_{22} have orders $n_i \times n_i$, $i = 1, 2$. If we use basis functions for the arising small elements in both the old and new node points, A takes the nodal basis function form while if we keep the basis functions for the old node set corresponding to the whole (unrefined) elements, it takes the form of a hierarchical basis function matrix

$$\widehat{A} = \begin{bmatrix} \widehat{A}_{11} & \widehat{A}_{12} \\ \widehat{A}_{21} & \widehat{A}_{22} \end{bmatrix}.$$

We have the following relation

$$\widehat{A} = J^T A J,$$

where

$$J = \begin{bmatrix} I_1 & J_{12} \\ 0 & I_2 \end{bmatrix}$$

and J_{12} corresponds to the interpolation matrix. An elementary computation using this, shows the next relations between the corresponding matrix blocks,

$$\widehat{A}_{11} = A_{11}, \quad \widehat{A}_{12} = A_{12} + A_{11} J_{12}, \quad \widehat{A}_{21} = A_{21} + J_{12}^T A_{11},$$

$$\widehat{A}_{22} = A_{22} + J_{12}^T A_{12} + A_{21} J_{12} + J_{12}^T A_{11} J_{12}.$$

Further, $\widehat{A}_{22} = A_{2h}$, i.e., the nodal basis function matrix for the coarse (unrefined) mesh and $\widehat{S} = S$, where $\widehat{S} = \widehat{A}_{22} - \widehat{A}_{21} A_{11}^{-1} \widehat{A}_{12}$ and $S = A_{22} - A_{21} A_{11}^{-1} A_{12}$. Moreover, the following spectral relations hold (see [4, 7, 8]):

$$(1-\gamma) \begin{bmatrix} A_{11} & 0 \\ 0 & \widehat{A}_{22} \end{bmatrix} \le \begin{bmatrix} A_{11} & \widehat{A}_{12} \\ \widehat{A}_{21} & \widehat{A}_{22} \end{bmatrix} \le (1+\gamma) \begin{bmatrix} A_{11} & 0 \\ 0 & \widehat{A}_{22} \end{bmatrix},$$

$$(1-\gamma^2) \widehat{A}_{22} \le S \le \widehat{A}_{22},$$

where $\gamma = \left\{\rho\left(\widehat{A}_{22}^{-1/2} A_{21} A_{11}^{-1} A_{12} \widehat{A}_{22}^{-1/2}\right)\right\}^{1/2}$ and all inequalities are sharp. Further, it is known that the above block diagonal matrix is an optimal preconditioner, i.e. minimizes the spectral condition number, among all block-diagonal preconditioners.
Here γ, $0 \leq \gamma < 1$ is identical to the constant in the strengthened CBS-inequality

$$u^T A v \leq \gamma \{u^T A u v^T A v\}^{1/2},$$

which holds for all orthogonal vectors u, v, $u = [0, 0, 0, \alpha_1, \alpha_2, \alpha_3]$, $v = [\beta_1, \beta_2, \beta_3, 0, 0, 0]$. We let γ_1, γ_2 denote the constants for the h-version (i.e. for p=1) and the p-version (i.e. for p=2), of hierarchical basis functions, respectively. The following relation between γ_1 and γ_2 holds.

Theorem 2.1 *[11],[2] For any finite element triangular mesh, where each element has been refined into congruent elements, it holds*

$$\gamma_2^2 = \frac{4}{3}\gamma_1^2, \tag{2}$$

where γ_1, γ_2 are the CBS constants for the piecewise linear and piecewise quadratic finite elements, respectively.

Corollary 1

$$\gamma_1^2 < \frac{3}{4}.$$

We recall now two preconditioners used to extend the two-level method to an arbitrary number of levels. For the hierarchical basis function matrix it is efficient to use a block diagonal preconditioner

$$D = \begin{bmatrix} B_{11} & 0 \\ 0 & \widetilde{A} \end{bmatrix}.$$

The matrices B_{11} and $\widetilde{A}$ are spectrally equivalent approximations to matrices A_{11} and $\widehat{A}_{22}$. The next result has been proven in [4].

Theorem 2.2 *Assume that*

$$b_1 v_1^T B_{11} v_1 \leq v_1^T A_{11} v_1 \leq b_0 v_1^T B_{11} v_1, \qquad \textit{for all} \quad v_1 \in \mathcal{R}^{n_1 - n_2}$$

and

$$a_1 v_2^T \widetilde{A} v_2 \leq v_2^T \widehat{A}_{22} v_2 \leq a_0 v_2^T \widetilde{A} v_2, \qquad \textit{for all} \quad v_2 \in \mathcal{R}^{n_2},$$

then

$$cond\left\{\begin{bmatrix} B_{11}^{-1} & 0 \\ 0 & \widetilde{A}^{-1} \end{bmatrix}\begin{bmatrix} \widehat{A}_{11} & \widehat{A}_{12} \\ \widehat{A}_{21} & \widehat{A}_{22} \end{bmatrix}\right\} \leq \frac{1+\gamma}{1-\gamma}\left(\frac{2}{1+\gamma}\right)^2 \frac{a_0+b_0}{2}\frac{1}{2}\left(\frac{1}{a_1}+\frac{1}{b_1}\right),$$

where

$$\gamma = \sup_{x,y} \frac{x^T \widehat{A}_{12} y}{\left\{x^T \widehat{A}_{11} x y^T \widehat{A}_{22} y\right\}^{1/2}}.$$

Furthermore if $a_0 \geq b_0$ *and* $a_1 \leq b_1$ *then it holds*

$$cond \leq \frac{1+\gamma}{1-\gamma}\frac{a_0}{a_1}.$$

This method can be extended recursively whereby each coarser matrix is approximated as above, except on certain levels, where one uses inner iterations to solve the arising coarse matrix system. Without such a stabilization the condition number would grow at least as $\left(\frac{1+\gamma}{1-\gamma}\right)^l$, with the level number distance l. For details, see [2].

Alternatively, one can use a multiplicative (=factorized) preconditioner. We use here the notations A_{2h} and A_h for the stiffness matrices corresponding to two consecutive levels. This preconditioner takes the form ([8])

$$M_h = \begin{bmatrix} B_{11} & 0 \\ \widetilde{A}_{21} & S_B \end{bmatrix}\begin{bmatrix} I_1 & B_{11}^{-1}\widetilde{A}_{12} \\ 0 & I_2 \end{bmatrix},$$

where

$$\begin{aligned} \widetilde{A}_{12} &= A_{12} + (A_{11} - B_{11})J_{12}, \\ \widetilde{A}_{21} &= A_{21} + J_{12}^T(A_{11} - B_{11}). \end{aligned} \tag{3}$$

The reason for perturbing the off-diagonal block matrices as done in (3) is that in this way

$$\widehat{M}_h = J^T M_h J, \tag{4}$$

and $\widehat{M}_h$ takes the form

$$\widehat{M}_h = \begin{bmatrix} B_{11} & \widehat{A}_{12} \\ \widehat{A}_{21} & S_B + \widehat{A}_{21}B_{11}^{-1}\widehat{A}_{12} \end{bmatrix},$$

which follows from an elementary computation. Here $\widehat{A}_{12} = A_{12} + A_{11}J_{12}$ is the off-diagonal block in the hierarchical basis function matrix

$$\widehat{A}_h = \begin{bmatrix} A_{11} & \widehat{A}_{12} \\ \widehat{A}_{21} & A_{2h} \end{bmatrix}.$$

Hence $\widehat{M}_h$ can be considered as a preconditioner to $\widehat{A}_h$ and the extreme eigenvalues of $M_h^{-1}A_h$ equal those of $\widehat{M}_h^{-1}\widehat{A}_h$, since

$$\sup_v \frac{v^T A_h v}{v^T M_h v} = \sup_{\widehat{v}} \frac{\widehat{v}^T \widehat{A}_h \widehat{v}}{\widehat{v}^T \widehat{M}_h \widehat{v}}, \qquad \inf_v \frac{v^T A_h v}{v^T M_h v} = \inf_{\widehat{v}} \frac{\widehat{v}^T \widehat{A}_h \widehat{v}}{\widehat{v}^T \widehat{M}_h \widehat{v}}.$$

Since the off-diagonal blocks in $\widehat{M}_h$ equal those in $\widehat{A}_h$ the estimate of the extreme eigenvalues of $\widehat{M}_h^{-1}\widehat{A}_h$ can be readily done. The following result has been proven in [8].

Theorem 2.3 *Assume that*

$$v_1^T A_{11} v_1 \le v_1^T B_{11} v_1 \le (1+b) v_1^T A_{11} v_1, \qquad \textit{for all} \quad v_1 \in \mathcal{R}^{n_2 - n_1}$$

and

$$v_2^T A_{2h} v_2 \le v_2^T S_B v_2 \le (1+d) v_2^T A_{2h} v_2, \qquad \textit{for all} \quad v_2 \in \mathcal{R}^{n_1}$$

then

$$cond\left(M_h^{-1}A_h\right) \le \frac{1+b+d}{1-\gamma^2}. \tag{5}$$

It follows from Corollary 1 that for piecewise linear functions on a triangular mesh it holds $\gamma^2 < 3/4$. For tetrahedrons it holds $\gamma^2 < 9/10$, see [1].

Remark 2.2 *The multiplicative method can be extended recursively replacing S_B with a matrix polynomial approximation*

$$S_B^{-1} = [I - P_\nu(M_{2h}^{-1}A_{2h})]A_{2h}^{-1},$$

where $P_\nu(0) = 1$ and P_ν is small on the interval of the eigenvalues of $M_h^{-1}A_h$, where M_{2h} is the preconditioner to A_{2h}. The best approximation is by a shifted and scaled Chebyshev polynomial, see [8]. In this way, the condition number can be stabilized, i.e. bounded by a number which does not depend on the number of levels. The construction is readily extended to multilevels. The polynomial degree doesn't have to be the same on each level.

As can be seen from the condition number estimates in Theorems 2.2 and 2.3, it is important to control the conditioning of A_{11}. The major task of this paper deals with some recent results in construction of preconditioners B_{11} to A_{11}, with condition number bounds which hold uniformly in both problem and mesh parameters.

3. A survey of previous results for scalar diffusion equations

On each level of the recursive multilevel extension of the additive or multiplicative method we must approximate the block matrix A_{11}. In this and next

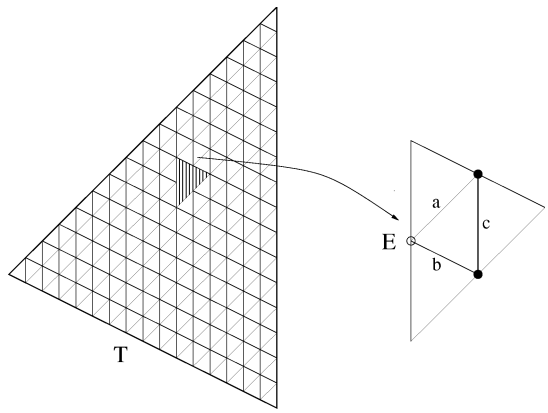


Figure 1: Four levels of uniform refinement of $T \in \mathcal{T}_0$ and macroelement $E \in \mathcal{T}_3$.

sections we describe two algorithms for construction of optimal order preconditioners B_{11} to the matrices A_{11} which are required within the AMLI methods under consideration. For each of the algorithms the condition numbers are bounded for all levels, i.e. independent on the level number, where the constant in the estimate is independent of the initial triangulation and the coefficients $a_{ij}(x)$ of the differential operator. The construction and the analysis of the preconditioners B_{11} are based on a macroelement-by-macroelement assembling procedure. The results in this section appeared earlier in a preliminary form in [5] but are included here for completeness.

3.1 Some basic relations

Let us consider two consecutive levels of uniform refinement ($l1$) and ($l2$). They correspond to the triangulations $\mathcal{T}_{2h}$ and $\mathcal{T}_h$ where each element of $\mathcal{T}_{2h}$ is divided into four congruent triangles of $\mathcal{T}_h$. We call the union of these four triangles *macroelement* $E \in \mathcal{T}_h$ (see figure 1).

Following the standard FEM assembling procedure we can write A_{11} in the form

$$A_{11} = \sum_{E \in \mathcal{T}_h} L_E^T A_{11:E} L_E, \tag{6}$$

where L_E stands for the restriction mapping of the global vector of unknowns to the local one corresponding to the macroelement E. Accounting for the general form of the element stiffness matrix corresponding to $T \in \mathcal{T}_0$ we get the following

simple representation of $A_{11:E}$, see e.g. [3],

$$(7)\quad A_{11:E} = 2\, r_T \begin{bmatrix} a_T + b_T + c_T & -c_T & -b_T \\ -c_T & a_T + b_T + c_T & -a_T \\ -b_T & -a_T & a_T + b_T + c_T \end{bmatrix},$$

where r_T depends on the shape of $T \in \mathcal{T}_0$ and on the related coefficients of the differential operator and a_T, b_T, c_T equal the cotan of the angels in T.
In what follows we will simplify the notations omitting the subscript T. This will not lead to any confusion as all constructions we will introduce are local, that is, they are within one and the same element of the initial triangulation $T \in \mathcal{T}_0$. Now without loss of generality we can assume that $|a| \le b \le c$. This concludes from the following relations.

Lemma 3.1 *Let $\theta_1, \theta_2, \theta_3$ be the angles in an arbitrary triangle. Then with $a = \cot\theta_1$, $b = \cot\theta_2$, $c = \cot\theta_3$ it holds*

(i) $a = (1 - bc)/(b + c)$

(ii) *If* $\theta_1 \ge \theta_2 \ge \theta_3$ *then* $|a| \le b \le c$

(iii) $a + b > 0$.

*Proof.*see [5]. ∎

Then

$$(8)\qquad A_{11:E} = 2\, r\, c \begin{bmatrix} \alpha+\beta+1 & -1 & -\beta \\ -1 & \alpha+\beta+1 & -\alpha \\ -\beta & -\alpha & \alpha+\beta+1 \end{bmatrix},$$

where $\alpha = a/c$, $\beta = b/c$. Taking into account that $a = \cot\theta_T^{(1)}$, $b = \cot\theta_T^{(2)}$, and $c = \cot\theta_T^{(3)}$ where $\theta_T^{(1)} + \theta_T^{(2)} + \theta_T^{(3)} = \pi$ are the angles of some auxiliary triangle depending on $T \in \mathcal{T}_0$ and on the corresponding coefficients $a_{ij}(T)$ of the differential operator (see e.g. [3]), we get that $(\alpha, \beta) \in D$ where $\alpha > -1/(1 + c/b) > -1/2$ and

$$(9)\quad D = \{(\alpha,\beta) \in \mathcal{R}^2 : -\frac{1}{2} < \alpha \le 1,\ 0 < \beta \le 1,\ \alpha + \beta > 0,\ \textit{and}\ |\alpha| \le \beta\}.$$

The next pure algebraic inequality will also be used in the following considerations.

Lemma 3.2 *For all $(\alpha, \beta) \in D$ holds the inequality*

$$(10)\qquad \frac{\alpha\beta + \alpha + \beta + 1}{(\alpha+\beta+1)(\alpha+\beta+2)} > \frac{4}{15}.$$

*Proof.*see [5]. ■

The approach used to construct the preconditioner discussed in the next section can be summarized as *preserving the links between the mesh nodes along the dominating anisotropy.*

3.2 An additive optimal order preconditioner of A_{11}

The additive preconditioner is defined as follows

$$B_{11} = \sum_{E \in \mathcal{T}_h} {L_E}^T B_{11:E} L_E. \tag{11}$$

The local matrix $B_{11:E}$ is obtained by preserving only the *strongest* off-diagonal entries, i.e., we have

$$B_{11:E} = 2\ r\ c \begin{bmatrix} \alpha+\beta+1 & -1 & 0 \\ -1 & \alpha+\beta+1 & 0 \\ 0 & 0 & \alpha+\beta+1 \end{bmatrix}. \tag{12}$$

To estimate the condition number of the preconditioner (11) to A_{11} we consider the local generalized eigenvalue problem

$$A_{11:E} v_E = \lambda_E B_{11:E} v_E. \tag{13}$$

The characteristic equation for λ_E, $\det(A_{11:E} - \lambda_E B_{11:E}) = 0$ can be written in the form

$$\begin{vmatrix} (\alpha+\beta+1)\mu_E & -\mu_E & -\beta \\ -\mu_E & (\alpha+\beta+1)\mu_E & -\alpha \\ -\beta & -\alpha & (\alpha+\beta+1)\mu_E \end{vmatrix} = 0, \tag{14}$$

where $\mu_E = 1 - \lambda_E$. For the solutions of (14) we get

$$\mu_E^{(1)} = 0, \quad \textit{and} \quad \left(\mu_E^{(2,3)}\right)^2 = \frac{(\alpha+\beta+1)(\alpha^2+\beta^2)+2\alpha\beta}{(\alpha+\beta+1)[(\alpha+\beta+1)^2-1]},$$

or, after simplification,

$$\left(\mu_E^{(2,3)}\right)^2 = \frac{\alpha^2+\beta^2+\alpha+\beta}{(\alpha+\beta+1)(\alpha+\beta+2)} = 1 - 2\frac{\alpha+\beta+1+\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta+2)}.$$

Hence, applying the inequality (10), it follows that $\left(\mu_E^{(2,3)}\right)^2 < 7/15$, and the local eigenvalue estimate

$$1 - \sqrt{7/15} < \lambda_E < 1 + \sqrt{7/15} \tag{15}$$

holds.

Theorem 3.1 *The additive preconditioner (A) of A_{11} has an optimal order convergence rate with a relative condition number uniformly bounded by*

$$\kappa\left({B_{11}}^{-1}A_{11}\right) < \frac{1}{4}(11+\sqrt{105}) \approx 5.31. \tag{16}$$

This condition number holds independent on shape and size of each element and on the coefficients in the differential operator.

*Proof.*see [5]. ∎

Next we will consider a multiplicative preconditioner which is applicable to more general systems of partial differential equations.

4. Multiplicative preconditioner

4.1 A multiplicative preconditioner of A_{11}, analysed for systems of partial differential equations

We consider now the construction of a multiplicative preconditioner. We partition then the nodes corresponding to the block A_{11} into two groups where the first one contains the centers of parallelogram superelements Q (see figure 2) which are weakly connected in a sense to be defined below. It is important to note that the parallelograms $Q \subset T \in \mathcal{T}_0$, i.e. it is not allowed to be composed of triangles of neighbour elements from the coarsest triangulation $\mathcal{T}_0$. With respect to this partitioning, A_{11} admits the following two-by-two block-factored form

$$A_{11} = \begin{bmatrix} D_{11} & F_{11} \\ {F_{11}}^T & E_{11} \end{bmatrix} = \begin{bmatrix} D_{11} & 0 \\ {F_{11}}^T & S_{11} \end{bmatrix} \begin{bmatrix} I & {D_{11}}^{-1}F_{11} \\ 0 & I \end{bmatrix}, \tag{17}$$

where S_{11} stands for the related Schur complement.

For simplicity, we consider here only a plane domain problem (d=2). The preconditioner to be presented will be applicable to systems of the differential equations. The order of the system equals the dimension of the space domain. We consider then the matrix corresponding to the mid-edge points in a $2D$ triangulation of the given domain where each macroelement contains four congruent triangles, as previously.

Here each element matrix is a 3×3 block matrix where each block has order 2×2.

It is readily seen that the element matrix takes the structure

$$J = \begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \end{bmatrix} \quad \text{, where} \quad J_{11} = J_{22} = J_{33} \quad \text{and} \quad J_{ij} = J_{ji}^T.$$

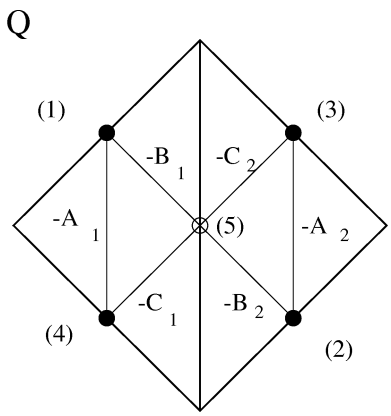


Figure 2: Weakest couplings in the union of two adjacent elements.

J is identical to the finite element matrix for the above element where we have Dirichlet boundary conditions at the vertex nodes and Neumann boundary conditions at the mid-edge node points. In particular J is positive definite.

Consider the factorization (17) where the partitioning of the nodes corresponds to superelements Q consisting of two adjacent triangles chosen such that the coupling along the edges parallel to the interface edge is the weakest coupling, see figure 2. The union of such superelements is denoted by $\mathcal{T}_h^Q$. The preconditioner B_{11} is defined by block approximation of the related Schur complement, i.e.,

$$B_{11} = \begin{bmatrix} D_{11} & 0 \\ {F_{11}}^T & \widehat{S}_{11} \end{bmatrix} \begin{bmatrix} I & {D_{11}}^{-1} F_{11} \\ 0 & I \end{bmatrix}, \tag{18}$$

where

$$S_{11} = S_{11:I} + \sum_{Q \in \mathcal{T}_h} S_{11:Q}, \qquad \widehat{S}_{11} = S_{11:I} + \sum_{Q \in \mathcal{T}_h} \widehat{S}_{11:Q}, \tag{19}$$

and

$$S_{11:Q} = \begin{bmatrix} \overline{S}_{11} & \overline{S}_{12} \\ \overline{S}_{21} & \overline{S}_{22} \end{bmatrix}, \qquad \widehat{S}_{11:Q} = \begin{bmatrix} \overline{S}_{11} & 0 \\ 0 & \overline{S}_{22} \end{bmatrix}.$$

Here, the blocks $\overline{S}_{ij}$, $i, j = 1, 2$ are 4×4 matrices where the partitioning follows the local numbering of the nodes from figure 2. The first additive term in (19) can be written in the form

$$S_{11:I} = \sum_{E \in \{\mathcal{T}_h - \mathcal{T}_h^Q\}} A_{11:E}, \qquad \mathcal{T}_h^Q = \bigcup_{Q \in \mathcal{T}_h} Q, \tag{20}$$

i.e., $S_{11:I}$ is the part of A_{11} corresponding to the node on the interface between the triangles from $\mathcal{T}_0$, and therefore is unchanged by a static condensation.
For the local analysis we need the superelement matrices $A_{11:Q}$ and $S_{11:Q}$. The corresponding matrix, ordered as in figure 2 takes the following structure:

$$A_{11:Q} = \begin{bmatrix} K_{11} & 0 & 0 & K_{14} & K_{15} \\ 0 & K_{22} & K_{23} & 0 & K_{25} \\ 0 & K_{32} & K_{33} & 0 & K_{35} \\ K_{41} & 0 & 0 & K_{44} & K_{45} \\ K_{51} & K_{52} & K_{53} & K_{54} & K_{55} \end{bmatrix},$$

where $K_{ij} = K_{ji}^T$, $K_{23} = K_{14} \equiv A$, $K_{25} = K_{15} \equiv B$, $K_{45} = K_{35} \equiv C$, $K_{11} = K_{22} = K_{33} = K_{44} \equiv D$ and $K_{55} = 2D$.
Here the matrix $A_{11:Q}$ and, in particular, D is symmetric and positive definite. Hence the assembled matrix has the structure:

$$A_{11:Q} = \begin{bmatrix} D & 0 & 0 & A & B \\ 0 & D & A & 0 & B \\ 0 & A^T & D & 0 & C \\ A^T & 0 & 0 & D & C \\ B^T & B^T & C^T & C^T & 2D \end{bmatrix}.$$

Following the above scheme here we will eliminate by static condensation the interior node-point to form a 4×4 block matrix and use its block diagonal part as preconditioner.
Since the bilinear form of the problem (1) is symmetric it follows that all matrices A, B, C are symmetric. Further, it holds

$$-A - B - C = D. \tag{21}$$

For systems of differential equations we must give a criteria how to define which couplings are weakest.
"Weakest" coupling
From Theorem 4.1 it follows that the smallest condition number will be achieved when we define the weakest couplings in the following way.
Let $A_1 = D^{-1/2} A D^{-1/2}$, $A_2 = D^{-1/2} B D^{-1/2}$, $A_3 = D^{-1/2} C D^{-1/2}$. Then for $i = 1, 2, 3$, compute $\rho_i = \rho(A_i A_i^T)$ and let

$$\tau_i = \begin{cases} \dfrac{1}{\sqrt{\rho_i}} - 1, & \text{if} \quad \rho_i < 1 \\ 0, & \text{if} \quad \rho_i \geq 1. \end{cases}$$

Further, let τ_i' be the biggest value of τ for which

$$-\tau(A_i + A_j) \leq -(A_j + A_k), \quad j \neq i, \quad k \neq j, \quad k \neq j, \quad i = 1, 2, 3,$$

holds. Then let

$$\tau = \max_{i=1,2,3} \min\{\tau_i, \tau_i'\}, \tag{22}$$

and let A_i correspond to the weakest coupling, being an index i where equality in (22) is taken.

Since $A_1 + A_2 + A_3 = -I$, it follows that $-(A_j + A_k) > 0$.

For each element pair, in this way one selects the weakest coupling. As all arising systems have small order, the computations required are viable. As it turns out, the method and results we shall present are equally applicable for the case of a pair of triangles having jumps in the coefficients but otherwise with the same matrices, i.e., where the relations

$$K_{ij}^{(2)} = \nu K_{ij}^{(1)} \quad , \quad i,j = 1,2,3, \quad \text{where} \quad 0 < \nu \leq 1, \tag{23}$$

hold. For this case the corresponding stiffness matrix takes the form

$$A_{11:Q} = \begin{bmatrix} D_1 & 0 & 0 & A_1 & B_1 \\ 0 & D_2 & A_2 & 0 & B_2 \\ 0 & A_2^T & D_2 & 0 & C_2 \\ A_1^T & 0 & 0 & D_1 & C_1 \\ B_1^T & B_2^T & C_2^T & C_1^T & D_1 + D_2 \end{bmatrix},$$

where $D_2 = \nu D_1$, $A_2 = \nu A_1$, $B_2 = \nu B_1$, $C_2 = \nu C_1$. Let $D_1 = D$, $A_1 = A$, $B_1 = B$, $C_1 = C$.

Next we eliminate by static condensation the couplings to the interior node point to form a 4×4 block matrix:

$$S_{11:Q} = \begin{bmatrix} D - \alpha E & -\beta E & -\beta F & A - \alpha F \\ -\beta E & \nu(D - \beta E) & \nu(A - \beta F) & -\beta F \\ -\beta F^T & \nu(A - \beta F)^T & \nu(D - \beta G) & -\beta G \\ (A - \alpha F)^T & -\beta F^T & -\beta G & D - \alpha G \end{bmatrix}, \tag{24}$$

where $E = BD^{-1}B^T$, $F = BD^{-1}C^T$, $G = CD^{-1}C^T$ and $\alpha = \dfrac{1}{1+\nu}$, $\beta = \dfrac{\nu}{1+\nu}$.

Note that $\alpha + \beta = 1$. We precondition $S_{11:Q}$ with $\widehat{S}_{11:Q}$.

Again, we recall that (apart from a scalar factor) this is the optimal preconditioner among all block diagonal preconditioners and to find the condition number $\kappa\left(\widehat{S}_{11:Q}^{-1} S_{11:Q}\right)$ we can compute $\gamma^2 = \|\overline{S}_{11}^{-1}\overline{S}_{12}\overline{S}_{22}^{-1}\overline{S}_{21}\|$ and use the fact that $\kappa\left(\widehat{S}_{11:Q}^{-1} S_{11:Q}\right) = \dfrac{1+\gamma}{1-\gamma}$.

Alternatively we can compute the eigenvalues of the generalized eigenvalue problem (where it is known that $\underline{y} = \underline{x}$ or $\underline{y} = -\underline{x}$),

$$S_{11:Q} \begin{pmatrix} \underline{x} \\ \underline{y} \end{pmatrix} = \lambda \widehat{S}_{11:Q} \begin{pmatrix} \underline{x} \\ \underline{y} \end{pmatrix}. \tag{25}$$

Analysis:
For the analysis it is convenient to first symmetrically scale the matrices with $D^{-1/2}$ to get the matrix (keeping for simplicity the same notations for $S_{11:Q}$ and $\widehat{S}_{11:Q}$)

$$\widehat{S}_{11:Q} = \begin{bmatrix} S_{11} & 0 \\ 0 & S_{22} \end{bmatrix},$$

where now

$$S_{11:Q} = \begin{bmatrix} I-\alpha\widetilde{E} & -\beta\widetilde{E} & -\beta\widetilde{F} & \widetilde{A}-\alpha\widetilde{F} \\ -\beta\widetilde{E} & \nu(I-\beta\widetilde{E}) & \nu(\widetilde{A}-\beta\widetilde{F}) & -\beta\widetilde{F} \\ -\beta\widetilde{F}^T & \nu(\widetilde{A}-\beta\widetilde{F})^T & \nu(I-\beta\widetilde{G}) & -\beta\widetilde{G} \\ (\widetilde{A}-\alpha\widetilde{F})^T & -\beta\widetilde{F}^T & -\beta\widetilde{G} & I-\alpha\widetilde{G} \end{bmatrix},$$

where $\widetilde{E} = \widetilde{B}\widetilde{B}^T$, $\widetilde{F} = \widetilde{B}\widetilde{C}^T$, $\widetilde{G} = \widetilde{C}\widetilde{C}^T$.
We see that there are three matrices, $\widetilde{A} := D^{-1/2}AD^{-1/2}$, $\widetilde{B} := D^{-1/2}BD^{-1/2}$ and $\widetilde{C} := D^{-1/2}CD^{-1/2}$ involved while the other matrices depend on those. Further, we recall that $-\widetilde{A}-\widetilde{B}-\widetilde{C} = I$. The reduced eigenvalue problem for (25) takes the form:

$$(1-\lambda)^2\underline{X} = \overline{S}_{11}^{-1}\overline{S}_{12}\overline{S}_{22}^{-1}\overline{S}_{21}\underline{X}, \qquad \text{or}$$

$$\begin{aligned}(1-\lambda)^2\underline{X} \;=\; & \begin{bmatrix} I-\alpha\widetilde{E} & -\beta\widetilde{E} \\ -\beta\widetilde{E} & \nu(I-\beta\widetilde{E}) \end{bmatrix}^{-1} \begin{bmatrix} \beta\widetilde{F} & \alpha\widetilde{F}-\widetilde{A} \\ \nu(\beta\widetilde{F}-\widetilde{A}) & \beta\widetilde{F} \end{bmatrix} \\ & \times \begin{bmatrix} \nu(I-\beta\widetilde{G}) & -\beta\widetilde{G} \\ -\beta\widetilde{G} & I-\alpha\widetilde{G} \end{bmatrix}^{-1} \begin{bmatrix} \beta\widetilde{F}^T & \nu(\beta\widetilde{F}-\widetilde{A})^T \\ (\alpha\widetilde{F}-\widetilde{A})^T & \beta\widetilde{F}^T \end{bmatrix} \underline{X}.\end{aligned}$$

(26)
We consider the first inverse matrix

$$\left[\begin{bmatrix} 1 & 0 \\ 0 & \nu \end{bmatrix} I - \begin{bmatrix} 1 & \nu \\ \nu & \nu^2 \end{bmatrix} \frac{\widetilde{E}}{1+\nu}\right]^{-1} = D_1 \left[\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} I - \begin{bmatrix} 1 & \sqrt{\nu} \\ \sqrt{\nu} & \nu \end{bmatrix} \frac{\widetilde{E}}{1+\nu}\right]^{-1} D_1,$$

where $D_1 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1/\nu} \end{bmatrix}$. A computation shows that

$$\left[\begin{bmatrix} 1 & 0 \\ 0 & \nu \end{bmatrix} I - \begin{bmatrix} 1 & \nu \\ \nu & \nu^2 \end{bmatrix} \frac{\widetilde{E}}{1+\nu}\right]^{-1}$$

$$= D_1 \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} I + \begin{bmatrix} 1 & \sqrt{\nu} \\ \sqrt{\nu} & \nu \end{bmatrix} \frac{\widetilde{E}(I-\widetilde{E})^{-1}}{1+\nu} \right\} D_1.$$

A further computation shows that

$$\left[\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} I + \begin{bmatrix} 1 & \sqrt{\nu} \\ \sqrt{\nu} & \nu \end{bmatrix} \frac{\widetilde{E}(I-\widetilde{E})^{-1}}{1+\nu}\right]^{1/2}$$

$$= \alpha \left\{ \begin{bmatrix} \nu & -\sqrt{\nu} \\ -\sqrt{\nu} & 1 \end{bmatrix} I + \begin{bmatrix} 1 & \sqrt{\nu} \\ \sqrt{\nu} & \nu \end{bmatrix} (I-\widetilde{E})^{-1/2} \right\}. \tag{27}$$

Similarly,

$$\left[\begin{bmatrix} \nu & 0 \\ 0 & 1 \end{bmatrix} I - \begin{bmatrix} \nu^2 & \nu \\ \nu & 1 \end{bmatrix} \frac{\widetilde{G}}{1+\nu}\right]^{-1}$$

$$= D_2 \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} I + \begin{bmatrix} \nu & \sqrt{\nu} \\ \sqrt{\nu} & 1 \end{bmatrix} \frac{\widetilde{G}(I-\widetilde{G})^{-1}}{1+\nu} \right\} D_2,$$

where $D_2 = \begin{bmatrix} \sqrt{1/\nu} & 0 \\ 0 & 1 \end{bmatrix}$, and

$$\left[\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} I + \begin{bmatrix} \nu & \sqrt{\nu} \\ \sqrt{\nu} & 1 \end{bmatrix} \frac{\widetilde{G}(I-\widetilde{G})^{-1}}{1+\nu}\right]^{1/2}$$

$$= \alpha \left\{ \begin{bmatrix} 1 & -\sqrt{\nu} \\ -\sqrt{\nu} & \nu \end{bmatrix} I + \begin{bmatrix} \nu & \sqrt{\nu} \\ \sqrt{\nu} & 1 \end{bmatrix} (I-\widetilde{G})^{-1/2} \right\}. \tag{28}$$

Hence, after a similarity transformation, the matrix in (26) can be written in the form

$$H = LL^T,$$

where

$$\begin{aligned} L \;=\; & \left[\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} I + \begin{bmatrix} 1 & \sqrt{\nu} \\ \sqrt{\nu} & \nu \end{bmatrix} \frac{\widetilde{E}(I-\widetilde{E})^{-1}}{1+\nu}\right]^{1/2} \\ & \times D_1 \left[\begin{bmatrix} \nu & 1 \\ \nu^2 & \nu \end{bmatrix} \frac{\widetilde{F}}{1+\nu} - \begin{bmatrix} 0 & 1 \\ \nu & 0 \end{bmatrix} \widetilde{A} \right] \\ & \times D_2 \left[\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} I + \begin{bmatrix} \nu & \sqrt{\nu} \\ \sqrt{\nu} & 1 \end{bmatrix} \frac{\widetilde{G}(I-\widetilde{G})^{-1}}{1+\nu}\right]^{1/2}. \end{aligned}$$

Using (27) and (28) we obtain

$$\begin{aligned} L &= \alpha^2 \left\{ \begin{bmatrix} \nu & -\sqrt{\nu} \\ -\sqrt{\nu} & 1 \end{bmatrix} I + \begin{bmatrix} 1 & \sqrt{\nu} \\ \sqrt{\nu} & \nu \end{bmatrix} (I - \widetilde{E})^{-1/2} \right\} \\ &\quad \times \left\{ \begin{bmatrix} \sqrt{\nu} & 1 \\ \nu & \sqrt{\nu} \end{bmatrix} \frac{\widetilde{F}}{1+\nu} - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \widetilde{A} \right\} \\ &\quad \times \left\{ \begin{bmatrix} 1 & -\sqrt{\nu} \\ -\sqrt{\nu} & \nu \end{bmatrix} I + \begin{bmatrix} \nu & \sqrt{\nu} \\ \sqrt{\nu} & 1 \end{bmatrix} (I - \widetilde{G})^{-1/2} \right\}, \end{aligned}$$

which, after simplifications, using the orthogonality of the matrices $\begin{bmatrix} \nu & -\sqrt{\nu} \\ -\sqrt{\nu} & 1 \end{bmatrix}$ and $\begin{bmatrix} \sqrt{\nu} & 1 \\ \nu & \sqrt{\nu} \end{bmatrix}$, takes the form

$$\begin{aligned} L &= \alpha^2 \left\{ \begin{bmatrix} \sqrt{\nu} & -\nu \\ -1 & \sqrt{\nu} \end{bmatrix} \widetilde{A} + \begin{bmatrix} \sqrt{\nu} & 1 \\ \nu & \sqrt{\nu} \end{bmatrix} (I - \widetilde{E})^{-1/2} (\widetilde{F} - \widetilde{A}) \right\} \\ &\quad \times \left\{ \begin{bmatrix} 1 & -\sqrt{\nu} \\ -\sqrt{\nu} & \nu \end{bmatrix} I + \begin{bmatrix} \nu & \sqrt{\nu} \\ \sqrt{\nu} & 1 \end{bmatrix} (I - \widetilde{G})^{-1/2} \right\} \\ &= \alpha \left\{ \begin{bmatrix} \sqrt{\nu} & -\nu \\ -1 & \sqrt{\nu} \end{bmatrix} \widetilde{A} + \begin{bmatrix} \sqrt{\nu} & 1 \\ \nu & \sqrt{\nu} \end{bmatrix} (I - \widetilde{E})^{-1/2} (\widetilde{F} - \widetilde{A}) (I - \widetilde{G})^{-1/2} \right\}. \end{aligned}$$

Hence

$$H = LL^T = \alpha \begin{bmatrix} \nu & -\sqrt{\nu} \\ -\sqrt{\nu} & 1 \end{bmatrix} \widetilde{A}\widetilde{A}^T$$

$$+ \alpha \begin{bmatrix} 1 & \sqrt{\nu} \\ \sqrt{\nu} & \nu \end{bmatrix} (I - \widetilde{E})^{-1/2} (\widetilde{F} - \widetilde{A}) (I - \widetilde{G})^{-1} (\widetilde{F} - \widetilde{A})^T (I - \widetilde{E})^{-1/2}.$$

It is immediately seen that the eigenvectors of H are the two orthogonal block vectors, $\begin{pmatrix} \underline{X} \\ \sqrt{\nu}\underline{X} \end{pmatrix}$ and $\begin{pmatrix} \sqrt{\nu}\underline{Y} \\ -\underline{Y} \end{pmatrix}$, where $\underline{X}$ is the eigenvector of

$$(I - \widetilde{E})^{-1/2} (\widetilde{F} - \widetilde{A}) (I - \widetilde{G})^{-1} (\widetilde{F} - \widetilde{A})^T (I - \widetilde{E})^{-1/2} \underline{X} = \lambda_1 \underline{X},$$

and $\underline{Y}$ is the eigenvector of

$$\widetilde{A}\widetilde{A}^T \underline{Y} = \lambda_2 \underline{Y}.$$

To analyze the eigenvalues λ_1 of the first matrix, we see that

$$\widetilde{F} - \widetilde{A} = \widetilde{B}\widetilde{C} + \widetilde{B} + \widetilde{C} + I = (I + \widetilde{B})(I + \widetilde{C})$$

and

$$I - \widetilde{E} = (I + \widetilde{B})(I - \widetilde{B}), \quad I - \widetilde{G} = (I + \widetilde{C})(I - \widetilde{C}).$$

Hence

$$\begin{aligned} H_1 &:= (I - \widetilde{E})^{-1/2}(\widetilde{F} - \widetilde{A})(I - \widetilde{G})^{-1}(\widetilde{F} - \widetilde{A})^T(I - \widetilde{E})^{-1/2} \\ &= (I - \widetilde{B})^{-1/2}(I + \widetilde{B})^{-1/2}(I + \widetilde{B})(I + \widetilde{C})(I + \widetilde{C})^{-1}(I - \widetilde{C})^{-1}(I + \widetilde{C}) \\ &\quad \times (I + \widetilde{B})(I + \widetilde{B})^{-1/2}(I - \widetilde{B})^{-1/2} \\ &= (I - \widetilde{B})^{-1/2}(I + \widetilde{B})^{1/2}(I - \widetilde{C})^{-1}(I + \widetilde{C})(I + \widetilde{B})^{1/2}(I - \widetilde{B})^{-1/2}. \end{aligned}$$

This matrix is similarly equivalent to

$$\begin{aligned} &(I - \widetilde{B})^{-1}(I + \widetilde{B})^{1/2}(I - \widetilde{C})^{-1}(I + \widetilde{C})(I + \widetilde{B})^{1/2} \\ &= (I + \widetilde{B})^{1/2}(I - \widetilde{B})^{-1}(I + \widetilde{C})(I - \widetilde{C})^{-1}(I + \widetilde{B})^{1/2}, \end{aligned}$$

which in its turn is similarly equivalent to

$$(I - \widetilde{B})^{-1}(I + \widetilde{C})(I - \widetilde{C})^{-1}(I + \widetilde{B}).$$

Since $-\widetilde{A} - \widetilde{B} - \widetilde{C} = I$, it holds

$$\begin{aligned} (I - \widetilde{B})^{-1}(I + \widetilde{C}) &= (I - \widetilde{B})^{-1}(-\widetilde{A} - \widetilde{B}) \\ &= (-\widetilde{A} - \widetilde{B} - \widetilde{B} - \widetilde{C})^{-1}(-\widetilde{A} - \widetilde{B}). \end{aligned}$$

Since by assumptions made,

$$(-\widetilde{A} - \widetilde{B}) + (-\widetilde{B} - \widetilde{C}) \geq (1 + \tau)(-\widetilde{A} - \widetilde{B}),$$

it follows that

$$\|(I - \widetilde{B})^{-1}(I + \widetilde{C})\| \leq \frac{1}{1 + \tau}.$$

Similarly,

$$(I - \widetilde{C})^{-1}(I + \widetilde{B}) = (-\widetilde{A} - \widetilde{C} - \widetilde{B} - \widetilde{C})^{-1}(-\widetilde{A} - \widetilde{C})$$

and

$$\|(I - \widetilde{C})^{-1}(I + \widetilde{B})\| \leq \frac{1}{1 + \tau}.$$

It follows that the eigenvalues of H_1 satisfy

$$0 \leq \lambda_1 \leq \frac{1}{(1 + \tau)^2}.$$

As before (section 3), the proof of the next theorem follows directly from the derived local estimate.

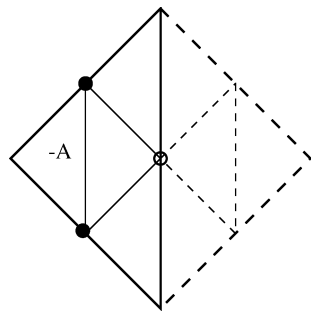


Figure 3: Static condensation in a single element.

Theorem 4.1 *The multiplicative preconditioner of A_{11} has an optimal order convergence rate with a relative condition number uniformly bounded by*

$$\kappa\left({B_{11}}^{-1}A_{11}\right) \leq \frac{1+\rho_0}{1-\rho_0}, \tag{29}$$

where $\rho_0 = max(\frac{1}{1+\tau}, \rho^{1/2}(\widetilde{A}^T\widetilde{A}))$, where $\widetilde{A}$ corresponds to the weakest coupling and τ is defined in (22). Furthermore, the result holds uniformly in problem parameters and shape of triangles.

Remark 4.1 *The bound in Theorem 4.1 does not depend on the jump ratio ν. Hence it holds also for $\nu \to 0$ showing the same bound for a single triangle, where the node opposite to the weakest coupling has been eliminated, see figure 3. This property can be of importance for the fictitious domain method, among others.*

Remark 4.2 *In the case of elasticity problem on a uniform mesh of isosceles triangles it can be shown that the parameter τ depends on the Poisson ratio $\tilde{\nu}$, $(\tilde{\nu} < 1/2)$ as follows*

$$\tau \leq 2\frac{\sqrt{8(4\tilde{\nu}^2 - 6\tilde{\nu} + 2) + 1} - 1}{\sqrt{8(4\tilde{\nu}^2 - 6\tilde{\nu} + 2) + 1} + 1}.$$

In this case the behavior of the condition number with respect to the Poisson ratio $(\tilde{\nu})$ is shown in Figure 4.

Remark 4.3 *Using the same derivations as before, for the scalar partial differential equations it is readily seen that the multiplicative preconditioner gives a condition number which is bounded above by 3 where the coefficients $-A = a$, $-B = b$ and $-C = c$ are weakly connected in the sense of the relation, $|a| \leq b \leq c$.*

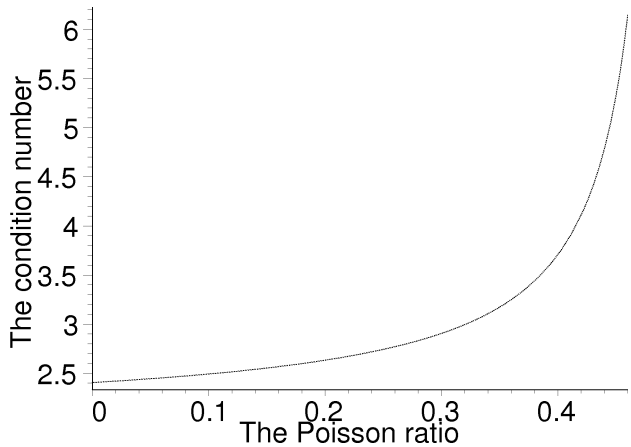


Figure 4: The condition number of elasticity on uniform mesh.

4.2 Solution of linear systems with the preconditioners B_{11}

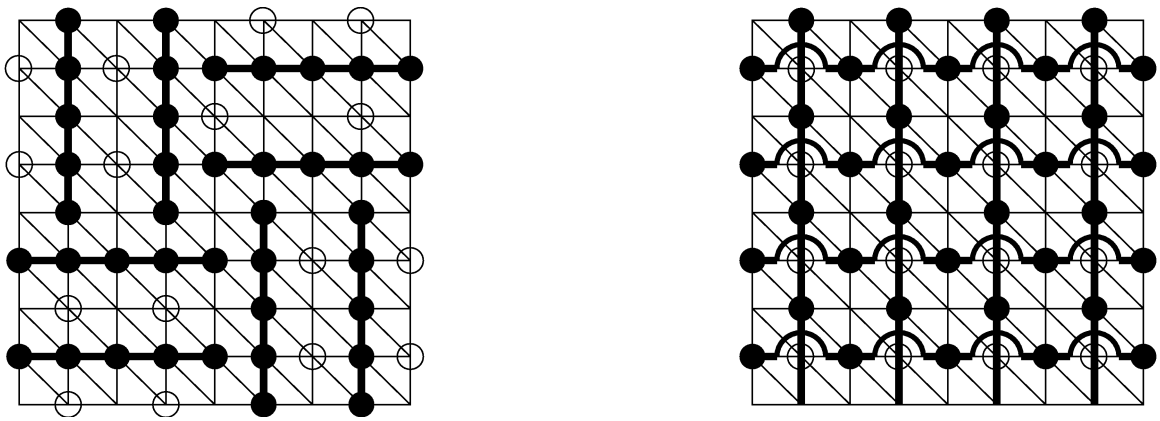
We consider briefly the particular case when the coefficient matrix of the differential problem is diagonal, i.e., $[a_{ij}(x)] = diag\,[a_1(x), a_2(x)]$, and the initial triangulation $\mathcal{T}_0$ consists of right triangles with legs parallel to the coordinate axes. The goal of this consideration is better to illustrate the behaviour of the related condition numbers. This model problem was studied during the years by various authors, applying different preconditioning techniques, and the results we present here will allow better to recognize the advantages of the here reported results. Here $\cot \theta_T^{(1)} = 0$ and consequently $a_T = \alpha_T = 0$ for the problem under consideration. The parameter $\beta_T \in (0, 1]$ is referred to as a *ratio of anisotropy*. Then the estimates (16) and (29) of the additive, multiplicative and multiplicative preconditioner take the following explicit forms,

$$\kappa^{(A)}\left({B_{11}}^{-1}A_{11}\right) \leq \max_{T\in\mathcal{T}_0}\left\{1+\beta_T+\sqrt{\beta_T(\beta_T+2)}\right\} < 2+\sqrt{3} \approx 3.73 \tag{30}$$

$$\kappa^{(M)}\left({B_{11}}^{-1}A_{11}\right) \leq \frac{1+1/3}{1-1/3} = 2, \tag{31}$$

respectively. It is important to stress here that the model problem considered in this subsection includes the interesting case when the direction of dominating anisotropy varies in different $T \in \mathcal{T}_0$.

The ability for efficient solution of systems with the previously introduced preconditioning matrices B_{11} is determined by their connectivity pattern, assuming that rapid solution methods are used at this step of the algorithm.



(a) (b)

Figure 5: Connectivity pattern of B_{11} for the model problem of varying ratio of anisotropy in $(0,1)^2$: (a) Additive preconditioner (A); (b) Multiplicative preconditioner (M).

Let us first consider the model problem in $\Omega = (0,1) \times (0,1)$ where the mesh is rectangular and uniform and the bilinear functional is as follows:

$$a(u,v) = \int_\Omega a_1 \frac{\partial u}{\partial x_1}\frac{\partial v}{\partial x_1} + a_2 \frac{\partial u}{\partial x_2}\frac{\partial v}{\partial x_2},$$

where the coefficients (a_1, a_2) are piecewise constants in subdomains Ω_i, $i \in \{1,2,3,4\}$ varying the anisotropy ratio as follows: in $\Omega_1 = (0,1/2)\times(0,1/2)$ $a_1 > a_2$; in $\Omega_2 = (1/2,1) \times (0,1/2)$ $a_1 < a_2$; in $\Omega_3 = (1/2,1) \times (1/2,1)$ $a_1 > a_2$; and in $\Omega_4 = (0,1/2) \times (1/2,1)$ $a_1 < a_2$. Figure 5 illustrates the connectivity pattern of B_{11} where the dense circles and bold lines show the remaining links after the local modification in the additive algorithm, and after the static condensation in the multiplicative variants. What one can see for this example is that the solution of systems with the preconditioning matrices B_{11} is split in a number of tridiagonal systems. Our final goal is to generalize these observations.

4.2.1 Additive algorithm (A)

Consider now a more irregular mesh, as shown in figure 6. It is readily seen, that in this case, the matrix B_{11} has a generalized tridiagonal structure (see [6] and also [13]), that is, the solution of linear systems with B_{11} has a computational cost which is proportional to the related problem size. In some more details, due to the form of the corresponding element matrices $B_{11:E}$, the related connectivity

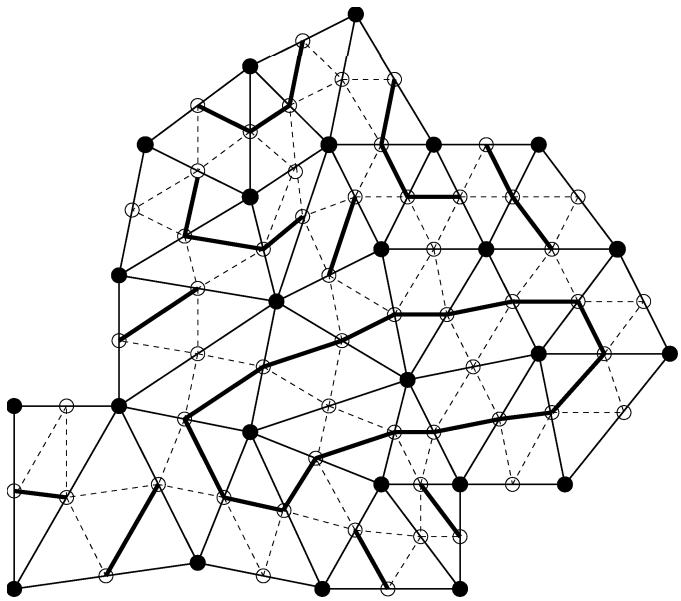


Figure 6: An example of the connectivity pattern for the additive preconditioner (A).

pattern of the preconditioner B_{11} is such as shown in figure 6; i.e., each node is coupled to none, one or at most two neighbors. This means that the coupled nodes form either a single point, a polyline or a polygon. Therefore, there are no cross-points. If we order the nodes along the connectivity lines, we get a block-diagonal form of the matrix B_{11}, where each block matrix is tridiagonal and corresponds to such a group of coupled nodes. Clearly, each of the blocks can be solved by a direct method with an arithmetic cost proportional to its dimension. Furthermore, an algorithm for ordering the unknowns can also be implemented with such an optimal order of complexity. Finally, we summarize the major result of this subsection in the next theorem.

Theorem 4.2 *The additive preconditioner of A_{11} has an optimal computational complexity with respect to both problem and discretization parameters.*

4.2.2 Multiplicative algorithm

The graph of connectivity of the multiplicative preconditioner is not planar. The design of special separators for this particular class of graph is out of the scope of this paper. We will note here, that due to the regular structure of the preconditioning matrix, algorithm (M) seems to be the best candidate, if the union of

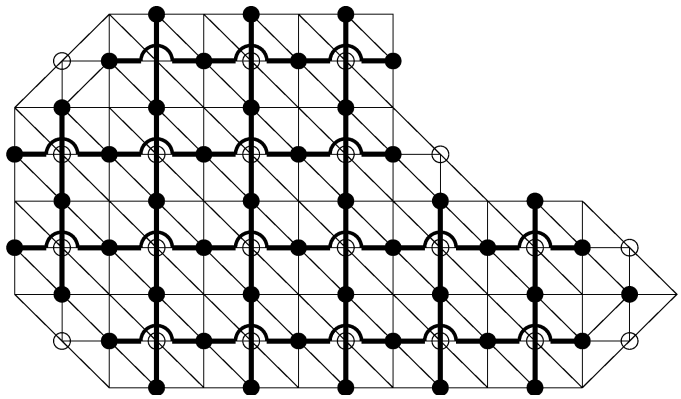


Figure 7: First example of the connectivity pattern for the multiplicative preconditioner (M).

the parallelograms used in the construction (see figure 5) is equal to Ω. Then, as for the model problem, the matrix B_{11} is completely decoupled to a number of tridiagonal blocks. This assumption obviously holds for discretizations on rectangular meshes of problems with a diagonal coefficient matrix. Note, that in this case, the estimates of the condition number have to be applied with a coefficient jumps parameter $\nu = 0$, (see (23), when the hypotenuse of a triangle from $\mathcal{T}_0$ is a part of the boundary of Ω.

Remark 4.4 *Let us assume that $\mathcal{T}_h^Q = \Omega$ (see (20)). This is a sufficient condition for the multiplicative algorithm (M) to have also an optimal computational complexity.*

We present here two examples to illustrate the potential of this remark. The first example problem is defined on a uniform polygonal domain discretized using a rectangular mesh, see figure 7, where the coefficients matrix is diagonal as in the model problem. Here we have also only two uncoupled directions of connectivity. Using a proper ordering of the nodes along connectivity pattern we get a block diagonal structure for the matrix B_{11}, which means that the multiplicative preconditioner has an optimal computational complexity for the considered problem. The second example is obtained from the first one by locally perturbing the shape of each element. It is readily seen that the algorithm is fully applicable if the alterations are slightly done.
We consider now the general case of connectivity pattern illustrated in Fig. 8. This means that the coarse mesh is composed by quadilaterals where the diagonals correspond to the weakest coupling for both adjacent triangles. We allow in this modification of the multiplicative algorithm to have superelements

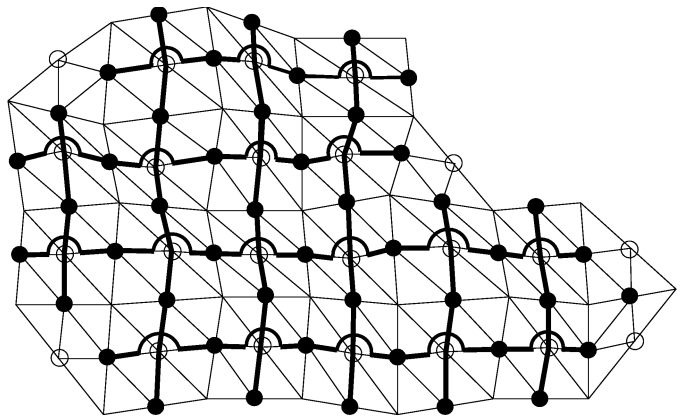


Figure 8: Second example of the connectivity pattern for the multiplicative preconditioner (M).

Q consisting of two adjacent triangles which belong to different elements from the initial mesh. The results of the numerically performed local analysis of the condition number $\kappa(B_{11}^{-1}A_{11})$ are presented in Table 2. The local analysis is performed under the assumption that the angles

$$\theta_1 \geq \theta_2 \geq \theta_3, \qquad \theta_1 + \theta_2 + \theta_3 = \pi,$$

$$\bar{\theta}_1 \geq \bar{\theta}_2 \geq \bar{\theta}_3, \qquad \bar{\theta}_1 + \bar{\theta}_2 + \bar{\theta}_3 = \pi,$$

of the triangles are varied with a stepsize $\tau = \pi/m$. The table has two rows corresponding to the cases $m = 100$ and $m = 200$. Then, the coefficients of the element stiffness matrices are computed as

$$a = \cot\theta_1, \quad b = \cot\theta_2, \quad c = \cot\theta_3, \tag{32}$$

and

$$\bar{a} = \nu\cot\bar{\theta}_1, \quad \bar{b} = \nu\cot\bar{\theta}_2, \quad \bar{c} = \nu\cot\bar{\theta}_3. \tag{33}$$

Here $\nu \in \{1, 10, 100, 1000\}$ represents the influence of the possible coefficient jumps trough the interface between the elements of the initial coarse grid.
What we observe from the presented numerical data is the stable behaviour of the condition number. This shows that the multiplicative preconditioner (M) is applicable to problems in the general setting (32) and (33) of this consideration, see figure 9.
Finally, it should be noted here that when the assumption of remark 4.4 is satisfied then the multiplicative preconditioner (M) is the best candidate for

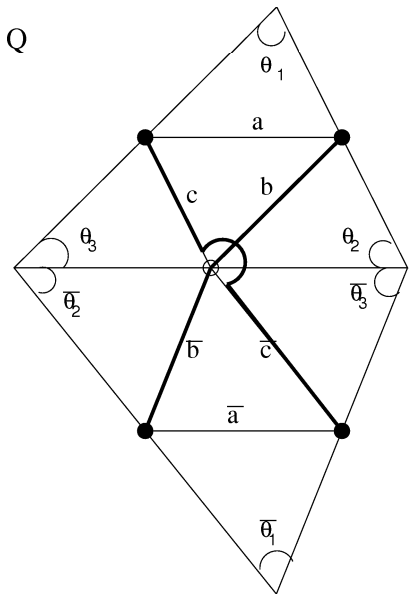


Figure 9: Multiplicative preconditioner. Connectivity pattern in the union of two adjacent elements.

m	$\nu = 1$	$\nu = 10$	$\nu = 100$	$\nu = 1000$
100	4.49	4.83	4.37	2.94
200	4.78	5.05	4.98	3.73

Table 1: Multiplicative preconditioner. Numerically computed estimates of $\kappa\left(B_{11:Q}^{-1}A_{11:Q}\right)$.

parallel computations. This is due to the decoupled structure of the algorithm, which does not require any special reordering of the unknowns.

5. Extension to three dimensional problems

The previous analysis can be extended to a three-dimensional case, where the mesh is constructed as a tensor product of an arbitrary one-dimensional grid T_z and an arbitrary two-dimensional triangulation T_{xy},

$$T = T_{xy} \otimes T_z.$$

The meshes T consist of prisms aligned along the z-direction. Similarly to the 2D case, we eliminate by static condensation the nodes (a) and (a') such that

the couplings (bc) and $(b^{'}c^{'})$ are the weakest couplings in the corresponding face, see figure 4.2.2.
The element stiffness matrix corresponding to A_{11} takes the following structure

$$K = \begin{bmatrix} K_{aa} & K_{aa'} & K_{ab} & K_{ab'} & K_{ac} & K_{ac'} \\ K_{a'a} & K_{a'a'} & K_{a'b} & K_{a'b'} & K_{a'c} & K_{a'c'} \\ K_{ba} & K_{ba'} & K_{bb} & K_{bb'} & K_{bc} & K_{bc'} \\ K_{b'a} & K_{b'a'} & K_{b'b} & K_{b'b'} & K_{b'c} & K_{b'c'} \\ K_{ca} & K_{ca'} & K_{cb} & K_{cb'} & K_{cc} & K_{cc'} \\ K_{c'a} & K_{c'a'} & K_{c'b} & K_{c'b'} & K_{c'c} & K_{c'c'} \end{bmatrix},$$

where each matrix K_{ij} has 3×3 size (order of the PDE'S). Note that K can be written in the following form

$$K = \begin{bmatrix} K_{11} & K_{12} & K_{13} \\ K_{21} & K_{22} & K_{23} \\ K_{31} & K_{32} & K_{33} \end{bmatrix},$$

where now the blocks K_{ij} has the size 6×6. Furthermore $K_{11} = K_{22} = K_{33}$ and $K_{ij} = K_{ji}^T$, $i, j = 1, \cdots, 3$ and K is symmetric positive definite matrix. The same analysis as in 2D can be done readily where we consider a polyhedral element instead of a polygonal element. The corresponding polyhedral element consist of (a pair of) two adjacent macroelements chosen such that the weakest coupling, in the same sense as in 2D, occurs along an edge parallel to the interface, see figure 4.2.2 for an illustration.

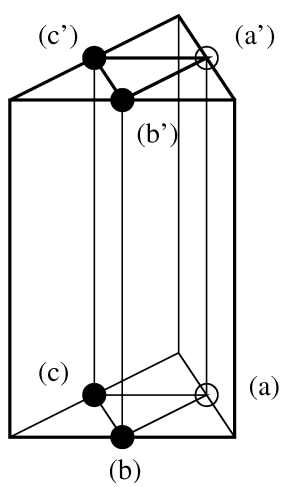


Fig.10: A prism macroelement.

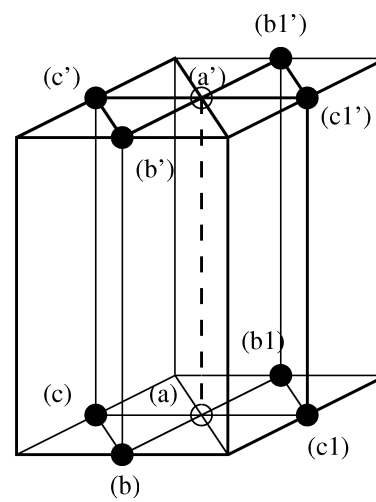


Fig.11: Union of two prism macroelements.

Remark 5.1 *By the static condensation all the nodes that are connected to the nodes* (a) *and* $(a^{'})$ *and which form a straight line parallel to the z-direction will be eliminated.*

6. Concluding remarks

The approximation accuracy for the discretization used may in some cases be bad for problems with bad aspect ratio or for nearly incompressible materials. This can be improved by proper mesh refinements or for nearly incompressible materials by introducing the pressure as an additional variable (see e.g., [6]). Even if the accuracy is bad it is of interest to solve the arising linear systems. For instance, based on the numerical solution on two mesh-levels, one can estimate the discretization error. As a general conclusion we summarize here some of the benefits from using the aforementioned preconditioning methods. The additive preconditioner (A) has a block diagonal structure which makes the implementation of the algorithm more efficient and easy. Furthermore it has an optimal computational complexity with respect to the size of the considered problem when dealing with scalar equations. To achieve such a result for more general PDE's systems one can use the multiplicative preconditioner (M) which turns out to be also a good candidate for parallelization computations, in particular when the assumption in remark (4.4) is fulfilled.

References

[1] B. Achchab, O. Axelsson, L. Laayouni and A. Souissi , *Strengthened Cauchy-Bunyakowski-Schwarz inequality for a three dimensional elasticity system*, Numerical Linear Algebra with Applications, Vol. 8(3), 191-205 (2001).

[2] O. Axelsson, *Stabilization of algebraic multilevel iteration methods; additive methods*, Numerical Algorithms, 21 (1999), 23–47.

[3] O. Axelsson and V. A. Barker, *Finite Element Solution of Boundary Value Problems: Theory and Computations*, Academic Press, Orlando, (1984)

[4] O. Axelsson and I. Gustafsson, *Preconditioning and two-level multigrid methods of arbitrary degree of approximations*, Math. Comp., 40(1983), 21-9-242.

[5] O. Axelsson and S. Margenov, *On multilevel preconditioners which are optimal with respect to both problem and discretization parameters*, Computational Methods in Applied Mathematics, Vol. 3 (1)(2003), 6–22.

[6] O. Axelsson and A. Padiy, *On the additive version of the algebraic multilevel iteration method for anisotropic elliptic problems*, SIAM J. Sci. Comput., 20 (1999), 1807–1830.

[7] O. Axelsson and P. S. Vassilevski, *Algebraic multilevel preconditioning methods, I*, Numer. Math., 56 (1989), 157 -177.

[8] O. Axelsson and P. S. Vassilevski, *Algebraic multilevel preconditioning methods, II*, SIAM J. Numer. Anal., 27 (1990), 1569–1590.

[9] R. Bank and T. Dupont, *An optimal order process for solving finite element equations*, Math. Comp., 36(1981), 35-51

[10] J. H. Bramble and X. Zhang, *Uniform convergence of the multigrid V-cycle for an anisotropic problem.* Math. Comp. 70 (2001), no. 234, 453–470.

[11] J. F. Maitre and F. Musy, *The contraction number of a class of two-level methods; an exact evaluation for some finite element subspaces and model problems*, Lect. Notes Math., 960 (1982), 535–544.

[12] S. Margenov and P.S. Vassilevski, *Algebraic multilevel preconditioning of anisotropic elliptic problems*, SIAM J. Sci. Comp., 15(5) (1994), 1026–1037.

[13] Y. Notay, *A multilevel block incomplete factorization preconditioning*, Applied Numerical Mathematics, 31 (1999), 209–225.

Received 19.03.2004

[1]*Department of Mathematics,*
University of Nijmegen,
The Netherlands

[2]*Department of Mathematics and Statistics,*
McGill University,
Montreal, Qc, Canada

[3]*Institute for Parallel Processing,*
Bulgarian Academy of Sciences,
Sofia, Bulgaria