Oberwolfach Preprints

OWP 2008 - 05

OWE AXELSSON AND JÁNOS KARÁTSON

Preconditioning of Block Tridiagonal Matrices

Mathematisches Forschungsinstitut Oberwolfach gGmbH
Oberwolfach Preprints (OWP) ISSN 1864-7596

Oberwolfach Preprints (OWP)

Starting in 2007, the MFO publishes a preprint series which mainly contains research results related to a longer stay in Oberwolfach. In particular, this concerns the Research in Pairs-Programme (RiP) and the Oberwolfach-Leibniz-Fellows (OWLF), but this can also include an Oberwolfach Lecture, for example.

A preprint can have a size from 1 - 200 pages, and the MFO will publish it on its website as well as by hard copy. Every RiP group or Oberwolfach-Leibniz-Fellow may receive on request 30 free hard copies (DIN A4, black and white copy) by surface mail.

Of course, the full copy right is left to the authors. The MFO only needs the right to publish it on its website *www.mfo.de* as a documentation of the research work done at the MFO, which you are accepting by sending us your file.

In case of interest, please send a **pdf file** of your preprint by email to *rip@mfo.de* or *owlf@mfo.de*, respectively. The file should be sent to the MFO within 12 months after your stay as RiP or OWLF at the MFO.

There are no requirements for the format of the preprint, except that the introduction should contain a short appreciation and that the paper size (respectively format) should be DIN A4, "letter" or "article".

On the front page of the hard copies, which contains the logo of the MFO, title and authors, we shall add a running number (20XX - XX).

We cordially invite the researchers within the RiP or OWLF programme to make use of this offer and would like to thank you in advance for your cooperation.

Imprint:

Mathematisches Forschungsinstitut Oberwolfach gGmbH (MFO)
Schwarzwaldstrasse 9-11
77709 Oberwolfach-Walke
Germany

Tel +49 7834 979 50
Fax +49 7834 979 55
Email admin@mfo.de
URL www.mfo.de

The Oberwolfach Preprints (OWP, ISSN 1864-7596) are published by the MFO.

Preconditioning of block tridiagonal matrices

by O. Axelsson[1], J. Karátson[2]

Abstract

Preconditioning methods via approximate block factorization for block tridiagonal matrices are studied. Bounds for the resulting condition numbers are given, and two methods for the recursive construction of the approximate Schur complements are presented. Illustrations for elliptic problems are also given, including a study of sensitivity to jumps in the coefficients and of a suitably modified Poincaré–Steklov operator on the continuous level.

Keywords: Preconditioning, Schur complement, domain decomposition, Poincaré–Steklov operator, approximate block factorization

1 Introduction

Block tridiagonal matrices arise in many applications. For instance, such a structure arises when decomposing the domain of definition of an elliptic operator using unidirectional stripes, or more generally, for a decomposition such that (in addition to a corresponding portion of the original boundary) each subdomain has a common boundary only with its previous and next neighbours in the sequence of subdomains. This subdivision can often be done according to different values of the coefficients in the differential operator, i.e. different materials in the underlying physical domain. Each diagonal block in the matrix corresponds to the restriction of the operator to one of the subdomains, and ordering the nodes in each domain in groups and then the domains consecutively, results in a block tridiagonal matrix.

Such problems are often split by ordering the interior domain nodes separately from the interface nodes and ordering all interface nodes last. This in turn results in a block diagonal submatrix with uncoupled blocks, which are only coupled to the interface nodes ordered last. The part of the system which corresponds to the different interior node sets can then be solved in parallel. The elimination of these nodes results, however, in a Schur complement matrix for the interface nodes, which in general is a full matrix. The corresponding system is often solved by a direct solution method which, however, can be costly for large-scale problems. Alternatively, one can construct various sparse preconditioners for the Schur complement matrix and solve that system by some iterative solution method. Such methods have been dealt with extensively in the literature, see e.g. [14, 15], based on certain Poincaré–Steklov operators corresponding to the interface. Also, various modifications for elliptic problems with coefficient jumps have been developed, see e.g. [9, 10, 12].

In this paper we keep instead the block tridiagonal structure

$$A = \begin{pmatrix} A_{11} & A_{12} & 0 & \dots & \dots & \dots & 0 \\ A_{21} & A_{22} & A_{23} & 0 & \dots & \dots & 0 \\ 0 & A_{32} & A_{33} & A_{34} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & 0 & A_{m,m-1} & A_{mm} \end{pmatrix} \tag{1.1}$$

[1]Department of Information Technology, Uppsala University, Sweden & Institute of Geonics AS CR, Ostrava, Czech Republic; owea@it.uu.se

[2]Department of Applied Analysis, ELTE University, H-1117 Budapest, Hungary; karatson@cs.elte.hu

and apply an appropriate (incomplete) block matrix factorization method. In this method we only need approximations of the local Schur complements which arise in eliminating one block matrix to form the next. We shall see that they can be constructed easily and the actions require little computational effort.

Bounds for the resulting condition number of the preconditioned matrix will be given. Thereby we show how the upper bound of the eigenvalues of the preconditioned matrix depends essentially only on the number of subdomains but does not depend on mesh size, that is, the bound is mesh independent. For the lower bound, assuming some additional properties of the matrix, one readily derives a fixed bound, frequently unity.

We present two methods for the construction of the approximate Schur complements. In both the action of the matrix on an arbitrary vector is readily computed. In the first method, called diagonal compensation, we correct the given diagonal block matrix with a diagonal matrix to preserve its action, compared with the corresponding exact matrix, on a particular positive vector. The second method involves a sum of actions of inverse matrices, and presents more accurate approximations.

Some illustrations for elliptic problems are included. We first show that the Schur complements for (1.1) are not sensitive to coefficient jumps, in contrast to the classical Schur complement for the interface. Then, turning to the continuous level for the corresponding operators, it is shown that the Schur complements for (1.1) correspond to dealing with the elliptic operator on the subdomains and certain modified Poincaré–Steklov operators for the interfaces. A continuous analogue of the first approximate factorization method is also presented. This illustrates properly some of the properties discussed on the discrete level.

The paper is organized as follows. Upper eigenvalue bounds for preconditioned matrices under approximate factorizations are given in Section 2. Two methods for the recursive construction of the approximate Schur complements, and condition number bounds for the corresponding preconditioned matrices, are presented in Section 3. In Section 4 we study the sensitivity of the Schur complements to jumps in the coefficients of the elliptic operator. Some analogues on the continuous level are given in Section 5, and we conclude in Section 6 with some additional remarks.

Except when it is otherwise stated, the inequalities

$$A \leq B, \quad A < B$$

between two symmetric matrices (of the same order) mean that $B - A$ is positive semidefinite or positive definite, respectively. The notation $\varrho(A)$ for a symmetric positive semidefinite matrix A stands for its maximal eigenvalue. The spectral condition number of A is defined by $\kappa(A) = \lambda_{max}(A)/\lambda_{min}(A)$.

2 Upper eigenvalue bounds for approximate factorizations

We now consider algebraic approaches to define preconditioners and estimate condition numbers for the corresponding preconditioned matrices.

Let A be a symmetric, positive definite (spd) matrix and partitioned in $m \times m$ block form. We split it as $A = D_A + L_A + L_A^T$ where D_A is the block diagonal part and L_A is the strictly lower block triangular part of A.

Let X be a spd block diagonal matrix and let L be strictly lower block triangular, both with a consistent partitioning to A. Let $K = A - L - L^T$ and $\tilde{L} = X^{-1/2} L X^{-1/2}$, $\tilde{K} = X^{-1/2} K X^{-1/2}$. The matrices $\tilde{L}$ and $\tilde{K}$ will only be used in the theoretical derivation, to follow. The actual

construction of X will be discussed later in this paper. We will use the following lemmata, see [6].

Lemma 2.1 *For any symmetric and positive semidefinite matrix*

$$\varrho(A) \le \sum_{i=1}^{m} \varrho(A_{ii})\,.$$

Lemma 2.2 *If* $(I+\tilde{L})^{-1} + (I+\tilde{L}^T)^{-1}$ *is positive semidefinite then*

$$(I+\tilde{L})^{-1} + (I+\tilde{L}^T)^{-1} \le 2m\,.$$

Let

$$C = (X+L)X^{-1}(X+L^T)\,.$$

The purpose of this paper is to derive estimates of the smallest and largest eigenvalues of the preconditioned matrix $C^{-1}A$. We shall thereby assume that X has been constructed to satisfy

$$\sigma := \varrho(X^{-1}K) < 2\,. \tag{2.1}$$

For the derivation of an upper bound, we consider the following similarity transformation of $C^{-1}A$:

$$M := X^{-1/2}(X+L^T)\,C^{-1}A\,(X+L^T)^{-1}X^{1/2}\,. \tag{2.2}$$

Then

$$M = X^{1/2}(X+L)^{-1}A(X+L^T)^{-1}X^{1/2}\,.$$

We shall derive two different types of bounds. For the derivation of the first bound, we note that

$$A = (K-2X) + (X+L) + (X+L^T)\,.$$

An elementary computation shows that

$$M = (I+\tilde{L})^{-1}(\tilde{K}-2I)(I+\tilde{L}^T)^{-1} + (I+\tilde{L})^{-1} + (I+\tilde{L}^T)^{-1}\,.$$

Then

$$\begin{aligned} M &\le (\sigma-2)(I+\tilde{L})^{-1}(I+\tilde{L}^T)^{-1} + (I+\tilde{L})^{-1} + (I+\tilde{L}^T)^{-1} \\ &= \frac{1}{2-\sigma}\,I - (2-\sigma)\Big((I+\tilde{L})^{-1} - \frac{1}{2-\sigma}\,I\Big)\Big((I+\tilde{L}^T)^{-1} - \frac{1}{2-\sigma}\,I\Big)\,. \end{aligned}$$

Hence

$$M \le \frac{1}{2-\sigma}\,I\,,$$

that is, since by the similarity transformation (2.2) the spectra of $C^{-1}A$ and M coincide,

$$\lambda_{max}(C^{-1}A) \le \frac{1}{2-\sigma}\,I\,, \tag{2.3}$$

which is the first estimate of the maximal eigenvalue.

We shall now derive an estimate which depends explicitly on the number of blocks of A. For this purpose, let Q be an orthonormal matrix that transforms $\tilde{K}$ to diagonal form,

$$\Lambda = Q\tilde{K}Q^T$$

where $\Lambda = diag(\mu_1,\ldots,\mu_m)$, i.e. μ_i are the eigenvalues of $X^{-1}K$. Then

$$N := QMQ^T = (I+\hat{L})^{-1}(\Lambda-2I)(I+\hat{L}^T)^{-1} + (I+\hat{L})^{-1} + (I+\hat{L}^T)^{-1}\,,$$

where $\hat{L} = Q^T\tilde{L}Q$. The following lemma will be used:

Lemma 2.3 *Let B be an arbitrary matrix of order n. Then*

$$(I+\hat{L})^{-1}B(I+\hat{L}^T)^{-1} = (I+\hat{L})^{-1}B + B(I+\hat{L}^T)^{-1} - B + (I+\hat{L})^{-1}\hat{L}B\hat{L}^T(I+\hat{L}^T)^{-1}\,. \quad (2.4)$$

PROOF. Write $\hat{L}$ and $\hat{L}^T$ in the last term of (2.4) as $\hat{L} = (I+\hat{L}) - I$, $\hat{L}^T = (I+\hat{L}^T) - I$. Then the equality in (2.4) follows by an elementary computation. ∎

Using the above lemma for $B = \Lambda - 2I$, it follows that

$$N = (I+\hat{L})^{-1}(\Lambda - I) + (\Lambda - I)(I+\hat{L}^T)^{-1} - (\Lambda - 2I) + (I+\hat{L})^{-1}\hat{L}(\Lambda - 2I)\hat{L}^T(I+\hat{L}^T)^{-1}\,. \quad (2.5)$$

Assume now that $\alpha C \leq A$ for some positive constant α. Then $N - \alpha I$ is positive semidefinite and it follows that

$$\lambda_{max}(C^{-1}A) = \lambda_{max}(N - \alpha I + \alpha I) \leq \ \lambda_{max}(N - \alpha I) + \alpha\,.$$

Further, since the diagonal blocks of $(I+\hat{L})^{-1}$ and $(I+\hat{L}^T)^{-1}$ are identity matrices, it follows from Lemmata 2.1-2.2 and of (2.5) that

$$\lambda_{max}(C^{-1}A) \leq \ \lambda_{max}(N - \alpha I) + \alpha \leq 2\sum_{i=1}^{m}(\mu_i - 1) + 2m - \sum_{i=1}^{m}\mu_i - \alpha m + \alpha = \sum_{i=1}^{m}\mu_i - \alpha(m-1). \quad (2.6)$$

Since, by assumption, $\mu_i \leq 2$, we obtain

$$\lambda_{max}(C^{-1}A) \leq (2-\alpha)m + \alpha. \quad (2.7)$$

A common case is $\alpha = 1$, see later. Then

$$\lambda_{max}(C^{-1}A) \leq m + 1.$$

However, (2.6) can be more accurate than the bound in (2.7) when several of the eigenvalues μ_i are (much) less than 2.

Together with (2.3), we have shown

Theorem 2.1 *Let A be a symmetric, positive definite (spd) matrix partitioned in $m \times m$ block form. Let $C = (X+L)X^{-1}(X+L^T)$ where X is spd block diagonal and let L is strictly lower block triangular, both with consistent partitioning to A. Let $\mu_i := \lambda_i(X^{-1}K)$ where $K = A - L - L^T$, let $\sigma := \max \mu_i$ and assume that $\sigma < 2$. Then*

$$\lambda_{max}(C^{-1}A) \leq min\Big\{\frac{1}{2-\sigma},\ \ \sum_{i=1}^{m}\mu_i - \alpha(m-1)\Big\}\,.$$

Remark 2.1 A common choice of L is $L = L_A$, which will be assumed in the remainder of the paper. Then $K = D_A$.

3 Recursive approximation of Schur complements

Let us consider a symmetric, positive definite (spd) matrix A with tridiagonal block structure as in (1.1):

$$A = \begin{pmatrix} A_{11} & A_{12} & 0 & \dots & \dots & \dots & 0 \\ A_{21} & A_{22} & A_{23} & 0 & \dots & \dots & 0 \\ 0 & A_{32} & A_{33} & A_{34} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & 0 & A_{m,m-1} & A_{mm} \end{pmatrix} . \tag{3.1}$$

Here $A_{ij} = A_{ji}^T$ for all i, j. The exact block factorization of A takes the form

$$A = (S + L_A)S^{-1}(S + L_A^T)$$

where $S = blockdiag(S_1, \dots, S_m)$ and the Schur complements S_i are determined recursively as

$$\begin{aligned} &S_1 := A_{11} \\ &S_2 := A_{22} - A_{21}S_1^{-1}A_{12} \\ &..... \\ &S_i := A_{ii} - A_{i,i-1}S_{i-1}^{-1}A_{i-1,i} \\ &..... \end{aligned} \tag{3.2}$$

for $i \leq m$.

The application of this factorization to solve a linear system involves the solution of the block triangular factors using a forward and a backward sweap. At each of them, systems with matrices S_i $(i = 1, \dots, m)$ appear that must be solved. In addition, matrix-vector multiplications with L_A and L_A^T, respectively, appear. In general, S_i are full matrices and their construction and the computation of actions of S_i^{-1} can be expensive.

Our goal now is to approximate S_i with some matrix X_i which is sparse, and such that the computation of X_i and actions of X_i^{-1} on vectors are cheap. At the same time, the approximation must be sufficiently accurate. For instance, it has been shown in [1] that the following lower bound holds for the condition number: $\kappa(C^{-1}A) \geq \min_i \kappa(X_i^{-1}S_i)$.

In the first method we let X_i be a correction to A_{ii} using a diagonal matrix, such that X_i takes the same action as the matrix that it approximates, on a certain positive vector. In the second method we approximate S_i^{-1} using a sum of matrices whose action is cheap.

3.1 Method 1: Diagonal compensation.

First let

$$X_i := A_{ii} - D_i \,, \tag{3.3}$$

where D_i is a diagonal matrix such that

$$D_i v_i = A_{i,i-1}X_{i-1}^{-1}A_{i-1,i}\, v_i \tag{3.4}$$

for some given positive vector v_i.

First, let v_i be the eigenvector to $A_{i,i-1}X_{i-1}^{-1}A_{i-1,i}$ corresponding to the smallest eigenvalue ξ_i of this matrix. Then

$$D_i v_i = \xi_i v_i \,,$$

i.e. $D_i = \xi_i I_i$ is a multiple of the identity matrix for the ith block. Since ξ_i is the smallest eigenvalue, it follows that $A_{i,i-1}X_{i-1}^{-1}A_{i-1,i} \geq D_i$ and hence $A_{ii} - A_{i,i-1}X_{i-1}^{-1}A_{i-1,i} \leq A_{ii} - D_i = X_i$. Here, using that $L = L_A$,

$$C - A = X + L_A X^{-1} L_A^T - D_A$$

and

$$(C - A)_{ii} = X_i + A_{i,i-1}X_{i-1}^{-1}A_{i-1,i} - A_{ii} \geq 0.$$

Hence $C \geq A$, which yields

$$\varrho(C^{-1}A) \leq 1.$$

In this method we must estimate the smallest eigenvalue of $C^{-1}A$, which we will not do here as the choice of X should rather be such that the smallest eigenvalue of $C^{-1}A$ is bounded by unity or some positive constant $\alpha \leq 1$.

Consider now the choice $v_i := \mathbf{e}_i = (1, \ldots, 1)$, i.e. $\mathbf{e}_i$ has all components equal to unity. Then X_i is obtained from

$$X_i := A_{ii} - D_i \tag{3.5}$$

where

$$D_i \mathbf{e}_i = A_{i,i-1}X_{i-1}^{-1}A_{i-1,i}\mathbf{e}_i\,. \tag{3.6}$$

Assume here that A is an M-matrix. Then we have componentwise

$$A_{ii}^{-1} \geq 0, \quad A_{i,i-1} \leq 0, \quad A_{i-1,i} \leq 0.$$

It follows by induction that $X_{i-1}^{-1} \geq 0$ componentwise, hence $A_{ii} - A_{i,i-1}X_{i-1}^{-1}A_{i-1,i}$ is a Z-matrix (i.e. all off-diagonal components are non-positive). Since $(A_{ii} - D_i)\mathbf{e}_i = (A_{ii} - A_{i,i-1}X_{i-1}^{-1}A_{i-1,i})\mathbf{e}_i$, it holds that if this vector is nonzero then $X_i = A_{ii} - D_i$ is positive definite, and also an M-matrix. Should the matrix lose positive definiteness (by having $(A_{ii} - D_i)\mathbf{e}_i = 0$), we must perturb the matrices A_{ii} with some (small) positive number. This will be discussed later.

Assuming that no perturbation is required, we have

$$X_i = A_{ii} - D_i \leq A_{ii} - A_{i,i-1}X_{i-1}^{-1}A_{i-1,i}$$

(here an inequality in a positive semidefinite sense). Therefore

$$(C - A)_{ii} = X_i + A_{i,i-1}X_{i-1}^{-1}A_{i-1,i} - A_{ii} \leq 0,$$

that is, $C \leq A$ and

$$\lambda_i(C^{-1}A) \geq 1.$$

Hence we have a lower bound. The upper bound follows from Theorem 2.1, so the condition number of $C^{-1}A$ is bounded as

$$\kappa(C^{-1}A) \leq min\Big\{\frac{1}{2-\sigma},\ \sum_{i=1}^{m}\mu_i - \alpha(m-1)\Big\},$$

where $\mu_i := \lambda_i(X^{-1}D_A)$ (using that $L = L_A$) and $\sigma := \max \mu_i$, further, it is assumed that $\sigma < 2$. In particular, if

$$D_i \leq \varrho A_{ii} \quad \text{for some} \quad \varrho < 1/2, \tag{3.7}$$

then by (3.5),

$$X_i \geq (1 - \varrho)A_{ii}$$

and hence

$$\sigma \leq \lambda_{max}(X_i^{-1}A_{ii}) \leq \frac{1}{1-\varrho} < 2. \tag{3.8}$$

Remark 3.1 The above two choices have somewhat opposite properties. In particular, if $\mathbf{e}_i$ is also an eigenvector of $A_{i,i-1}X_{i-1}^{-1}A_{i-1,i}$ for the smallest eigenvalue, then it can be seen that $A_{i,i-1}X_{i-1}^{-1}A_{i-1,i}$ is a multiple of the identity matrix. In the following we assume that this does not hold.

Remark 3.2 A similar method has been used in [11]. Let

$$X_i = A_{ii} - A_{i,i-1}Y_{i-1}A_{i-1,i} + D_i', \qquad i = 2, \ldots, m,$$

where Y_{i-1} is a bandmatrix, possibly diagonal, approximation of X_{i-1}^{-1}, such that the off-diagonal entries of Y_{i-1} are not larger than those of D_{i-1}^{-1}, and D_i' is a diagonal matrix determined such that

$$D_i'\mathbf{e}_i = A_{i,i-1}(Y_{i-1} + D_{i-1}^{-1})A_{i-1,i}\mathbf{e}_i\,.$$

Here we have $D_1\mathbf{e}_1 = A_{11}\mathbf{e}_1$ and

$$D_i\mathbf{e}_i = (A_{ii} - A_{i,i-1}D_{i-1}^{-1}A_{i-1,i})\mathbf{e}_i, \qquad i = 2, \ldots, m.$$

It follows that $A - C$ is a Z-matrix and $(A - C)\mathbf{e} = 0$, so $\lambda_{min}(C^{-1}A) \geq 1$.

Remark 3.3 The above method has been applied in [11] to elliptic problems

$$-\frac{\partial}{\partial x}\left(a_1\,\frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial y}\left(a_2\,\frac{\partial u}{\partial y}\right) = f$$

on a rectangular domain $\{(x, y) :\ 0 < x < a,\ 0 < y < b\}$ with Dirichlet boundary condition except possibly at $x = a$, where a Neumann boundary condition may hold. It has been proved that if we use a columnwise ordering starting at the line $x = 0$, and if $x \mapsto a_1(x, y)$ is nonincreasing, then $\mu_i \leq 2$ for $i = 2, \ldots, m$, $\mu_1 = 1$ and Theorem 2.1 implies

$$\lambda_{max}(C^{-1}A) \leq \sum_{i=1}^{m}(\mu_i - 1) \leq m\,.$$

The upper bound $\varrho(X_i^{-1}A_{ii}) < 2$ may, however, not always hold.

To achieve a sequence X_i for which it holds, the following linear combination of the two diagonal compensation methods may be used. Let $0 < \alpha_i \leq 1$ and

$$X_i = A_{ii} - \alpha_i D_i - (1 - \alpha_i)\xi_i I_i\,,$$

where

$$D_i\mathbf{e}_i = A_{i,i-1}X_{i-1}^{-1}A_{i-1,i}\mathbf{e}_i\,,$$

$$A_{i,i-1}X_{i-1}^{-1}A_{i-1,i}\,v_i = \xi_i v_i$$

and ξ_i is the smallest eigenvalue of $A_{i,i-1}X_{i-1}^{-1}A_{i-1,i}$. We require that

$$X_i \geq \frac{1}{2}A_{ii}\,. \tag{3.9}$$

This condition, $2X_i \geq A_{ii}$ takes the form

$$A_{ii} - 2\alpha_i D_i - 2(1 - \alpha_i)\xi_i I_i \geq 0.$$

This matrix is a Z-matrix and is positive semidefinite if $\zeta_i := A_{ii}\mathbf{e}_i - 2\alpha_i D_i \mathbf{e}_i - 2(1-\alpha_i)\xi_i \mathbf{e}_i \geq 0$. Assume that the smallest component of ζ_i is taken for its jth component ($j = j_i$). Then, letting $\hat{e}_j$ be the jth unit vector, it holds that

$$\hat{e}_j^T \left(A_{ii}\mathbf{e}_i - 2\alpha_i A_{i,i-1} X_{i-1}^{-1} A_{i-1,i}\mathbf{e}_i - 2(1-\alpha_i)\xi_i \mathbf{e}_i \right) \geq 0$$

or

$$2(1-\alpha_i)\hat{e}_j^T (A_{i,i-1} X_{i-1}^{-1} A_{i-1,i}\mathbf{e}_i - \xi_i \mathbf{e}_i) \geq \mathbf{e}_j^T \left(2A_{i,i-1} X_{i-1}^{-1} A_{i-1,i} - A_{ii}\right)\mathbf{e}_i \,.$$

Hence (3.9) holds if

$$1 - \alpha_i = \max \left\{ 0, \; \frac{\hat{e}_j^T \left(A_{i,i-1} X_{i-1}^{-1} A_{i-1,i} - \frac{1}{2} A_{ii}\right)\mathbf{e}_i}{\mathbf{e}_j^T \left(A_{i,i-1} X_{i-1}^{-1} A_{i-1,i} - \xi_i\right)\mathbf{e}_i} \right\} .$$

It is seen that $\alpha_i = 1$ as long as

$$\hat{e}_j^T \, A_{i,i-1} X_{i-1}^{-1} A_{i-1,i}\mathbf{e}_i \leq \frac{1}{2} e_j^T \, A_{ii}\mathbf{e}_i \,,$$

i.e. in particular if

$$\frac{1}{2} A_{ii} \geq A_{i,i-1} X_{i-1}^{-1} A_{i-1,i} \,.$$

The actual choice of $\{X_i\}$ is a balance between a good upper bound and a good lower bound. Accurate estimates of lower bounds can be difficult. A technique used in [1] is based on perturbations of the diagonal of the given matrix. More generally, one can perturb with a positive semidefinite matrix. For instance, when the given matrix is not an M-matrix, one can perturb it with semidefinite matrices of the form

$$\beta \begin{pmatrix} \ddots & & & & \\ & 1 & \dots & -1 & \\ & & \ddots & & \\ & -1 & \dots & 1 & \\ & & & & \ddots \end{pmatrix}$$

where $\beta > 0$ is chosen to make the matrix an M-matrix. This technique is identical to moving the positive off-diagonal entries to the diagonal in the same row. It has been called diagonal compensation, see [5], where resulting eigenvalue bounds can be found.

If the matrix is already an M-matrix, then one can add (small) positive entries to its diagonal and construct the preconditioner for the perturbed matrix. In this way the preconditioner becomes more 'stable' (positive definite) than if it had been applied to the unperturbed matrix.

Let A be the given matrix, δ the diagonal perturbation matrix, $\tilde{A} = A + \delta$ and let C be the preconditioner based on $\tilde{A}$. The aim of the perturbation is to limit the upper eigenvalue bound while still not decrease the lower eigenvalue bound too much. Often the preconditioner satisfies $\lambda_{min}(C^{-1}\tilde{A}) \geq 1$. It holds then $\lambda_{min}(C^{-1}A) \geq \lambda_{min}(C^{-1}\tilde{A})\lambda_{min}(\tilde{A}^{-1}A) \geq \lambda_{min}(\tilde{A}^{-1}A) = \lambda_{min}((A+\delta)^{-1}A)$. Estimates of such bounds can be based on the length of certain directed paths in the matrix graph for A from the node point where a perturbation takes place to a nearby Dirichlet boundary node, see [1] and the references therein. In general, the amount of perturbations should be limited by the order of the mesh width. Another application of the perturbation method can be found in [3].

3.2 Method 2: Approximation with sums of inverse matrices

We now consider a more accurate approximation of the Schur complements. We note first that at each stage of the factorization method, we must compute a Schur complement for the matrix

$$\begin{pmatrix} X_{i-1} & A_{i-1,i} \\ A_{i,i-1} & A_{ii} \end{pmatrix} \tag{3.10}$$

$(i = 2, \ldots, m)$, which is positive definite.

The chosen form of the approximation arises from the relation between the inverses of the Schur complements of a two-by-two block matrix, which relation is based on the Sherman-Morrison formula, see e.g. [1]. For the matrix in (3.10) it takes the form

$$S_i^{-1} = A_{ii}^{-1} + A_{ii}^{-1} A_{i,i-1} \tilde{S}_{i-1}^{-1} A_{i-1,i} A_{ii}^{-1} \,, \tag{3.11}$$

$$\tilde{S}_{i-1}^{-1} = X_{i-1}^{-1} + X_{i-1}^{-1} A_{i-1,i} S_i^{-1} A_{i,i-1} A_{i-1,i-1}$$

where for the standard forms it holds

$$S_i = A_{ii} - A_{i,i-1} X_{i-1}^{-1} A_{i-1,i} \,,$$

$$\tilde{S}_{i-1} = X_{i-1} - A_{i-1,i} S_i^{-1} A_{i,i-1} \,.$$

For the application of the preconditioner C, we note that we need only actions of the inverses X_i^{-1}, i.e. we do not need actions of X_i themselves. Therefore the expression in (3.11) is viable if we can approximate S_{i-1}^{-1} in a proper way. The actions of the other matrices in (3.11) are cheap and the approximate action of S_{i-1}^{-1} must also be cheap. The approximation can also not involve X_{i-1}^{-1} as this would lead to a recursive computation, involving matrices on all previous levels. The approximation we choose is to replace $\tilde{S}_{i-1}^{-1}$ by $A_{i-1,i-1}^{-1} + D_{i-1}$, where D_{i-1} is a diagonal matrix. Then X^{-1} takes the form

$$X^{-1} = D_A^{-1} + D_A^{-1} L (D_A^{-1} + D) L^T D_A^{-1}$$

where we want to determine $D \geq 0$ such that

$$A - C = D_A - X - L X^{-1} L^T \tag{3.12}$$

becomes close to zero. However, at the same time we want to limit that choice to satisfy

$$\sigma \leq \| X^{-1} D_A \|_\infty = \max_i |X^{-1} D_A \mathbf{e}|_i \leq 2 \,, \tag{3.13}$$

where we have used the fact that $X^{-1} > 0$ when A is an M-matrix. Since

$$X^{-1} D_A = I + D_A^{-1} L (D_A^{-1} + D) L^T \,, \tag{3.14}$$

we require that

$$D_A^{-1} L (D_A^{-1} + D) L^T \mathbf{e} \leq \mathbf{e}$$

or

$$D_A^{-1} L D L^T \mathbf{e} \leq (I - D_A^{-1} L D_A^{-1} L^T) \mathbf{e} \,.$$

Making $A - C = 0$ in (3.12) gives the relation

$$X^{-1} D_A - I = X^{-1} L X^{-1} L^T \,.$$

Relating this to (3.14) gives

$$D_A^{-1}L(D_A^{-1}+D)L^T = X^{-1}LX^{-1}L^T\,.$$

Therefore we wish to choose D such that

$$D_A^{-1}LDL^T = (X^{-1}LX^{-1} - D_A^{-1}LD_A^{-1})L^T\,.$$

For practical reasons we can only satisfy this relation in a small subspace. We determine here a diagonal matrix D such that

$$D_A^{-1}LDL^T\mathbf{e} = (X^{-1}LX^{-1} - D_A^{-1}LD_A^{-1})L^T\mathbf{e}\,,$$

at the same time limiting that choice to satisfy (3.13). Whereas the computation of actions of X^{-1} is now more involved, the increased computational effort is expected to be outweighed by an increased accuracy of X^{-1}. The computation of $D = block(D_1,\ldots,D_m)$ takes place recursively for D_i $(i = 2,\ldots,m)$ starting with $D_1 = 0$. The approximation X^{-1} can be seen as a truncated Neumann series, $X_i^{-1} = A_{ii}^{-1} + G_i + G_iA_{ii}G_i + \ldots$ where $G_i = A_{ii}^{-1}A_{i,i-1}A_{i-1,i-1}^{-1}A_{i-1,i}A_{ii}^{-1}$ and the matrix term D_{i-1} tries to compensate for the truncated terms.

4 Schur complements for elliptic problems with jumps in their coefficients

Let us consider a domain decomposition method for an elliptic problem discretized with FEM, such that (in addition to a corresponding portion of the outer boundary) each subdomain has a common boundary only with its previous and next neighbours in the sequence of subdomains. Let the elliptic operator have constant diffusion coefficients in each subdomain, the value of which can vary between subdomains. Such problems often arise in the context of various domain decomposition procedures [2, 8, 9, 10, 12]. Our goal in this section is to study the condition numbers of the arising Schur complements.

In the classical domain decomposition (DD) approach, the interior domain nodes are ordered separately from the interface nodes and all interface nodes are ordered last. It has been observed that the condition number of the corresponding Schur complements deteriorate as the magnitude of jumps increases. Like in multigrid methods, to avoid this, an efficient method has proved to be to introduce one or more proper auxiliary coarse spaces that have a global balancing effect, see e.g. the BDD method [12] and the approach of so-called exotic coarse spaces [9] in a Schwarz method framework. The definition of these coarse spaces involves the solution of local Dirichlet subproblems, but it turns out that they can be avoided by using suitable approximate harmonic extensions, which lead to an optimal solution method (except possibly the solution of the coarse problem).

An alternative to the above approach is to take the interface nodes into account together with the previous subdomain in the mentioned sequence of subdomains. This approach, considered in the present paper, leads to a tridiagonal block structure as in (1.1). In this section it is verified for a model problem that the condition numbers of the Schur complements are sensitive to the jump in the first approach (namely, proportional to the magnitude of the jump) but are not in the second approach. That is, one can have independence of jumps without introducing auxiliary problems.

For simplicity, in this section we will study a decomposition of the domain Ω in three subdomains Ω_1, Ω_2 and Ω_3. According to the above, we have common boundaries $\Gamma_1 := \overline{\Omega}_1 \cap \overline{\Omega}_2$ and $\Gamma_2 := \overline{\Omega}_2 \cap \overline{\Omega}_3$, but Ω_1 and Ω_3 have no common boundary.

The discussion below is built up as follows. We will first formulate the block forms of the stiffness matrix under the two mentioned approaches for an isotropic Poisson equation. Then we will consider a different diffusion coefficient in each Ω_i, rewrite the stiffness matrices, and study the variation of the corresponding condition numbers.

4.1 Basic block forms for the isotropic Poisson equation

Let us consider the Poisson equation where, in weak form, one seeks $u \in V_h \subset H_0^1(\Omega)$ such that

$$\int_\Omega \nabla u \cdot \nabla v = \int_\Omega f v \qquad (v \in V_h). \tag{4.1}$$

The FEM subspace is chosen with piecewise linear basis functions, assumed either to have node points on one of Γ_i or to have its support entirely in one of Ω_i.

In the classical DD approach, the stiffness matrix is written in the block form

$$A = \begin{pmatrix} A_{11} & 0 & 0 & A_{1,\Gamma_1} & 0 \\ 0 & A_{22} & 0 & A_{2,\Gamma_1} & A_{2,\Gamma_2} \\ 0 & 0 & A_{33} & 0 & A_{3,\Gamma_2} \\ A_{\Gamma_1,1} & A_{\Gamma_1,2} & 0 & A_{\Gamma_1,\Gamma_1} & 0 \\ 0 & A_{\Gamma_2,2} & A_{\Gamma_2,3} & 0 & A_{\Gamma_2,\Gamma_2} \end{pmatrix}. \tag{4.2}$$

Here $A_{i,\Gamma_j} = A_{j,\Gamma_i}^T$ for all i, j. Then one lets

$$A_{\Gamma 1} := A_{1\Gamma}^T := \begin{pmatrix} A_{\Gamma_1,1} \\ 0 \end{pmatrix}, \qquad A_{\Gamma 2} := A_{2\Gamma}^T := \begin{pmatrix} A_{\Gamma_1,2} \\ A_{\Gamma_2,2} \end{pmatrix}, \qquad A_{\Gamma 3} := A_{3\Gamma}^T := \begin{pmatrix} 0 \\ A_{\Gamma_2,3} \end{pmatrix}, \tag{4.3}$$

$$A_{\Gamma\Gamma} := \begin{pmatrix} A_{\Gamma_1,\Gamma_1} & 0 \\ 0 & A_{\Gamma_2,\Gamma_2} \end{pmatrix} \tag{4.4}$$

and thus obtains the more concise form

$$A = \begin{pmatrix} A_{11} & 0 & 0 & A_{1\Gamma} \\ 0 & A_{22} & 0 & A_{2\Gamma} \\ 0 & 0 & A_{33} & A_{3\Gamma} \\ A_{\Gamma 1} & A_{\Gamma 2} & A_{\Gamma 3} & A_{\Gamma\Gamma} \end{pmatrix}. \tag{4.5}$$

The solution of the corresponding linear system can be reduced to solving systems with $\Sigma_i := A_{ii}$ $(i = 1, 2, 3)$ and an additional system with the Schur complement matrix

$$\Sigma := A_{\Gamma\Gamma} - A_{\Gamma 1} A_{11}^{-1} A_{1\Gamma} - A_{\Gamma 2} A_{22}^{-1} A_{2\Gamma} - A_{\Gamma 3} A_{33}^{-1} A_{3\Gamma}\,. \tag{4.6}$$

In the other approach, the interface nodes are taken into account together with the previous subdomain. Under this reordering, the stiffness matrix in (4.2) can be rewritten as

$$\tilde{A} = \begin{pmatrix} A_{11} & A_{1,\Gamma_1} & 0 & 0 & 0 \\ A_{\Gamma_1,1} & A_{\Gamma_1,\Gamma_1} & A_{2,\Gamma_1} & 0 & 0 \\ 0 & A_{\Gamma_1,2} & A_{22} & A_{2,\Gamma_2} & 0 \\ 0 & 0 & A_{\Gamma_2,2} & A_{\Gamma_2,\Gamma_2} & A_{3,\Gamma_2} \\ 0 & 0 & 0 & A_{\Gamma_2,3} & A_{33} \end{pmatrix}, \tag{4.7}$$

where we introduce the notations

$$\tilde{A}_{11} := \begin{pmatrix} A_{11} & A_{1,\Gamma_1} \\ A_{\Gamma_1,1} & A_{\Gamma_1,\Gamma_1} \end{pmatrix}, \qquad \tilde{A}_{12} := \begin{pmatrix} 0 & 0 \\ A_{2,\Gamma_1} & 0 \end{pmatrix}, \tag{4.8}$$

$$\tilde{A}_{21} := \begin{pmatrix} 0 & A_{\Gamma_1,2} \\ 0 & 0 \end{pmatrix}, \qquad \tilde{A}_{22} := \begin{pmatrix} A_{22} & A_{2,\Gamma_2} \\ A_{\Gamma_2,2} & A_{\Gamma_2,\Gamma_2} \end{pmatrix}, \tag{4.9}$$

$$\tilde{A}_{23} := \begin{pmatrix} 0 \\ A_{3,\Gamma_2} \end{pmatrix}, \qquad \tilde{A}_{32} := \begin{pmatrix} 0 \\ A_{\Gamma_2,3} \end{pmatrix} \tag{4.10}$$

to obtain the concise form

$$\tilde{A} = \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} & 0 \\ \tilde{A}_{21} & \tilde{A}_{22} & \tilde{A}_{23} \\ 0 & \tilde{A}_{32} & A_{33} \end{pmatrix}. \tag{4.11}$$

In the Schur complement approach, here only the first block remains unchanged: $S_1 := \tilde{A}_{11}$, and the solution of the original system can now be reduced to solving two additional systems corresponding to Schur complements, determined recursively as

$$S_2 := \tilde{A}_{22} - \tilde{A}_{21} S_1^{-1} \tilde{A}_{12}, \qquad S_3 := A_{33} - \tilde{A}_{32} S_2^{-1} \tilde{A}_{23}\,. \tag{4.12}$$

Using notations (4.8)–(4.10) and letting

$$S_{\Gamma_1} := A_{\Gamma_1,\Gamma_1} - A_{\Gamma_1,1} A_{11}^{-1} A_{1,\Gamma_1}, \tag{4.13}$$

we obtain

$$S_2 = \begin{pmatrix} A_{22} - A_{\Gamma_1,2} S_{\Gamma_1}^{-1} A_{2,\Gamma_1} & A_{2,\Gamma_2} \\ A_{\Gamma_2,2} & A_{\Gamma_2,\Gamma_2} \end{pmatrix}. \tag{4.14}$$

(The similar formula for S_3 will not be needed here.)

4.2 Conditioning properties for problems with jumps in their coefficients

Now we can turn to the case of our interest. Instead of the above Poisson equation, we consider the FEM solution of an elliptic problem with a different constant diffusion coefficient in each Ω_i. That is, in weak form, one seeks $u \in V_h \subset H_0^1(\Omega)$ such that

$$\int_\Omega w\, \nabla u \cdot \nabla v = \int_\Omega f v \qquad (v \in V_h), \tag{4.15}$$

where w is a weight function on Ω such that

$$w_{|\Omega_i} \equiv w_i \qquad (i = 1, 2, 3).$$

In our model problem we assume

$$w_1 \geq w_2 \geq w_3 \tag{4.16}$$

and are interested in the case

$$w_1 >> w_2\,. \tag{4.17}$$

When varying these coefficients, in order to avoid the loss of ellipticity in the limit, we also assume that there exists a constant $\alpha > 0$ such that

$$w_3 \geq \alpha w_2\,. \tag{4.18}$$

Below, we will find that if we vary the ratio $\frac{w_1}{w_2}$ unboundedly, then the condition numbers also grow to infinity for the Schur complement in (4.6) but remain bounded for the Schur complements in (4.12).

Let us first consider the classical DD approach again. The stiffness matrix (4.2) is then modified as follows. The entries corresponding to basis functions with support in Ω_i are multiplied by the weight w_i. For simplicity, assume that for the node points on one of Γ_i, the support of the basis function is symmetric w.r.t the node point, and thus its parts intersecting with the two domains have equal measure. (An opposite case will be mentioned in Remark 4.1.) Then the entries corresponding to such basis functions are multiplied by $(w_i + w_j)/2$. Therefore, the stiffness matrix has the form

$$A = \begin{pmatrix} w_1 A_{11} & 0 & 0 & w_1 A_{1,\Gamma_1} & 0 \\ 0 & w_2 A_{22} & 0 & w_2 A_{2,\Gamma_1} & w_2 A_{2,\Gamma_2} \\ 0 & 0 & w_3 A_{33} & 0 & w_3 A_{3,\Gamma_2} \\ w_1 A_{\Gamma_1,1} & w_2 A_{\Gamma_1,2} & 0 & \frac{w_1+w_2}{2} A_{\Gamma_1,\Gamma_1} & 0 \\ 0 & w_2 A_{\Gamma_2,2} & w_3 A_{\Gamma_2,3} & 0 & \frac{w_2+w_3}{2} A_{\Gamma_2,\Gamma_2} \end{pmatrix} . \tag{4.19}$$

With these modifications, one readily sees that the Schur complement (4.6) becomes

$$\Sigma(w) := W A_{\Gamma\Gamma} - w_1 A_{\Gamma 1} A_{11}^{-1} A_{1\Gamma} - w_2 A_{\Gamma 2} A_{22}^{-1} A_{2\Gamma} - w_3 A_{\Gamma 3} A_{33}^{-1} A_{3\Gamma} \tag{4.20}$$

where W is the diagonal matrix

$$W := \begin{pmatrix} \frac{w_1+w_2}{2} & 0 \\ 0 & \frac{w_2+w_3}{2} \end{pmatrix} .$$

Proposition 4.1 *There exist constants $c_1, c_2 > 0$ independent of w such that*

$$\kappa(\Sigma(w)) \geq c_1 \frac{w_1}{w_2} + c_2 . \tag{4.21}$$

Proof. Using (4.3)-(4.4), a simple calculation yields

$$\tilde{\Sigma}(w) := \frac{1}{w_2} \Sigma(w) = \begin{pmatrix} \frac{w_1}{w_2} \Sigma_1 + \frac{1}{2} A_{\Gamma_1,\Gamma_1} & 0 \\ 0 & \frac{1}{2}(1 + \frac{w_3}{w_2}) A_{\Gamma_2,\Gamma_2} \end{pmatrix} - A_{\Gamma 2} A_{22}^{-1} A_{2\Gamma} - \frac{w_3}{w_2} A_{\Gamma 3} A_{33}^{-1} A_{3\Gamma} \tag{4.22}$$

where

$$\Sigma_1 := \frac{1}{2} A_{\Gamma_1,\Gamma_1} - A_{\Gamma_1,1} A_{11}^{-1} A_{1,\Gamma_1} .$$

Here $\Sigma_1 \geq 0$ (i.e. it is positive semidefinite) and is not a zero matrix since it is a Schur complement, corresponding to the positive definite matrix $\tilde{A}_{11}$ modified by setting a zero diffusion coefficient outside Ω_1. Further, $A_{\Gamma_i,\Gamma_i} > 0$ $(i = 1, 2)$ and $\frac{1}{2}(1 + \frac{w_3}{w_2}) \leq 1$ owing to (4.16). Hence, the matrix

$$G(w) := \begin{pmatrix} \frac{w_1}{w_2} \Sigma_1 + \frac{1}{2} A_{\Gamma_1,\Gamma_1} & 0 \\ 0 & \frac{1}{2}(1 + \frac{w_3}{w_2}) A_{\Gamma_2,\Gamma_2} \end{pmatrix}$$

satisfies

$$\lambda_{max}(G(w)) \geq \frac{w_1}{w_2} \lambda_{max}(\Sigma_1), \qquad \lambda_{min}(G(w)) \leq \lambda_{min}(A_{\Gamma_2,\Gamma_2}),$$

which yields for the condition number of $G(w)$ that

$$\kappa(G(w)) \geq \frac{w_1}{w_2} \frac{\lambda_{max}(\Sigma_1)}{\lambda_{min}(A_{\Gamma_2,\Gamma_2})} .$$

The condition numbers of the other two terms in (4.22) are bounded. Since $\kappa(\Sigma(w)) = \kappa(\tilde{\Sigma}(w))$, we obtain (4.21). ∎

Corollary 4.1 *If we vary* $\frac{w_1}{w_2}$ *unboundedly, then*

$$\kappa(\Sigma(w)) = O\big(\frac{w_1}{w_2}\big) \to \infty \quad as\ \frac{w_1}{w_2} \to \infty.$$

Remark 4.1 The above sensitivity to $\frac{w_1}{w_2}$ may be reduced if the supports of the basis functions on Γ_1 are not assumed to be symmetric w.r.t. the node point, but their parts intersecting with Ω_2 have small measure. However, this would in turn lead to inpractically small element widths and very large gradients of the basis functions near Γ_1.

Let us now consider the second approach. We study the Schur complements (4.12) modified w.r.t. the diffusion coefficient w. The corresponding modification of the matrix $\tilde{A}$ in (4.7) comes by first replacing the considered blocks of (4.2) by the corresponding blocks of (4.19), and then using the same reassembling as for (4.7). Then the Schur complement S_2 in (4.14) becomes modified as

$$S_2(w) := \begin{pmatrix} w_2 A_{22} - \ w_2^2\, A_{\Gamma_1,2}\, S_{\Gamma_1}(w)^{-1} A_{2,\Gamma_1} & w_2 A_{2,\Gamma_2} \\ w_2 A_{\Gamma_2,2} & \frac{1}{2}(w_2 + w_3) A_{\Gamma_2,\Gamma_2} \end{pmatrix}, \tag{4.23}$$

where S_{Γ_1} in (4.13) has been replaced by

$$S_{\Gamma_1}(w) := \frac{w_1 + w_2}{2}\, A_{\Gamma_1,\Gamma_1} - w_1 A_{\Gamma_1,1} A_{11}^{-1} A_{1,\Gamma_1}\,. \tag{4.24}$$

Introducing the notation

$$S_2^{11}(w) := A_{22} - \ w_2\, A_{\Gamma_1,2} S_{\Gamma_1}(w)^{-1} A_{2,\Gamma_1}\,, \tag{4.25}$$

we have

$$S_2(w) := \begin{pmatrix} w_2 S_2^{11}(w) & w_2 A_{2,\Gamma_2} \\ w_2 A_{\Gamma_2,2} & \frac{1}{2}(w_2 + w_3) A_{\Gamma_2,\Gamma_2} \end{pmatrix} \tag{4.26}$$

Lemma 4.1 *There holds* $S_{\Gamma_1}(w) \geq w_2\, S_{\Gamma_1}$.

Proof. We have

$$\begin{aligned} S_{\Gamma_1}(w) &= w_2 \Big[\frac{1}{2}\big(\frac{w_1}{w_2} + 1\big) A_{\Gamma_1,\Gamma_1} - \frac{w_1}{w_2} A_{\Gamma_1,1} A_{11}^{-1} A_{1,\Gamma_1}\Big] \\ &= w_2 \Big[\frac{w_1}{w_2}\big(\frac{1}{2} A_{\Gamma_1,\Gamma_1} - A_{\Gamma_1,1} A_{11}^{-1} A_{1,\Gamma_1}\big) + \frac{1}{2} A_{\Gamma_1,\Gamma_1}\Big]\,. \end{aligned}$$

Since, by assumption, $w_1 \geq w_2$, we obtain

$$S_{\Gamma_1}(w) \geq w_2 \Big[\big(\frac{1}{2} A_{\Gamma_1,\Gamma_1} - A_{\Gamma_1,1} A_{11}^{-1} A_{1,\Gamma_1}\big) + \frac{1}{2} A_{\Gamma_1,\Gamma_1}\Big] \ = w_2\, S_{\Gamma_1}.$$ ∎

Similarly to (4.25), let us denote the top left block of (4.14) by

$$S_2^{11} := A_{22} - A_{\Gamma_1,2} S_{\Gamma_1}^{-1} A_{2,\Gamma_1}\,, \tag{4.27}$$

and then let

$$\tilde{S}_2 := \begin{pmatrix} S_2^{11} & A_{2,\Gamma_2} \\ A_{\Gamma_2,2} & \frac{1}{2}(1 + \alpha) A_{\Gamma_2,\Gamma_2} \end{pmatrix} \tag{4.28}$$

with α from (4.18). Now we can prove the required boundedness:

Proposition 4.2 *The condition number of $S_2(w)$ satisfies*

$$\kappa(S_2(w)) \le \frac{\lambda_{max}(\tilde{A}_{22})}{\lambda_{min}(\tilde{S}_2)}\,,$$

hence it is bounded independently of w.

PROOF. Clearly $S_2^{11}(w) \le A_{22}$, and $\frac{1}{2}(w_2 + w_3) \le w_2$ owing to (4.16), hence

$$S_2(w) \le w_2 \begin{pmatrix} A_{22} & A_{2,\Gamma_2} \\ A_{\Gamma_2,2} & A_{\Gamma_2,\Gamma_2} \end{pmatrix} = w_2\, \tilde{A}_{22}\,. \tag{4.29}$$

To find a lower bound for $S_2(w)$, note that Lemma 4.1 and the definitions (4.25) and (4.27) yield

$$S_2^{11}(w) \ge S_2^{11}\,. \tag{4.30}$$

Substituting (4.30) into (4.26), and using (4.18) and (4.28), respectively, we then obtain

$$S_2(w) \ge \begin{pmatrix} w_2 S_2^{11} & w_2 A_{2,\Gamma_2} \\ w_2 A_{\Gamma_2,2} & \frac{1}{2}(1+\alpha) w_2\, A_{\Gamma_2,\Gamma_2} \end{pmatrix} = w_2\, \tilde{S}_2\,.$$

Here $\tilde{S}_2 > 0$, since by the above, $w_2\tilde{S}_2$ is the Schur complement $S_2(w)$ in the case $w_3 = \alpha w_2$. Together with (4.29), we obtain the required statement. ∎

Finally, we consider the second Schur complement S_3 from (4.12). When replacing its considered blocks from (4.2) by the corresponding blocks of (4.19), we observe that each of the blocks A_{33}, $\tilde{A}_{32}$ and $\tilde{A}_{23}$ is multiplied by w_3. Hence the matrix S_3 becomes modified as

$$S_3(w) := w_3\, A_{33} - w_3^2\, \tilde{A}_{32}\, S_2(w)^{-1} \tilde{A}_{23}\,. \tag{4.31}$$

We can easily prove again the required boundedness:

Proposition 4.3 *The condition number of $S_3(w)$ satisfies*

$$\kappa(S_3(w)) \le \frac{\lambda_{max}(A_{33})}{\lambda_{min}(S_3)}\,,$$

hence it is bounded independently of w.

PROOF. Obviously $S_3(w) \le w_3\, A_{33}$. Further, in (4.26) we can estimate each w_2 below by w_3 and $S_2^{11}(w)$ below by S_2^{11} using (4.30), such that we obtain

$$S_2(w) \ge w_3 \begin{pmatrix} S_2^{11} & A_{2,\Gamma_2} \\ A_{\Gamma_2,2} & A_{\Gamma_2,\Gamma_2} \end{pmatrix} = w_3\, S_2\,,$$

and substituting into (4.31) yields

$$S_3(w) \ge w_3\, A_{33} - w_3\, \tilde{A}_{32}\, S_2^{-1} \tilde{A}_{23} \;= w_3\, S_3\,.$$

The two bounds imply the desired estimate. ∎

5 Some model analysis on the continuous level

A continuous analogue of the preceding method is presented on some model problems, including the introduction of a certain modified Poincaré–Steklov operator for the interfaces. This study on the continuous level can help the understanding of the properties of the studied factorization approach.

5.1 Preliminaries: the Poincaré–Steklov operator

As pointed out in Section 4, the analysis of standard domain decomposition methods relies strongly on the Poincaré–Steklov operator, see e.g. [14, 15]. In this subsection we give a brief description, following [15].

Let us consider a boundary value problem

$$\left\{ \begin{array}{ll} -\Delta u = f & \text{in } \Omega \\ u_{|\partial\Omega} = 0 & \end{array} \right. \tag{5.1}$$

on a bounded domain Ω with Lipschitz boundary and some $f \in L^2(\Omega)$. The domain Ω is decomposed in two nonoverlapping domains Ω_1 and Ω_2, whose common boundary is denoted by Γ, further, we let $\Gamma_1 := \partial\Omega_1 \setminus \Gamma$ and $\Gamma_2 := \partial\Omega_2 \setminus \Gamma$.

The Poincaré–Steklov operator is then defined in the following way. Let us choose an arbitrary function $\gamma \in H_{00}^{1/2}(\Gamma)$. (For the definition of $H_{00}^{1/2}(\Gamma)$ and other related Sobolev spaces, see also [14].) Let $H_1\gamma$ and $H_2\gamma$ denote the harmonic extensions of γ in Ω_1 and Ω_2, respectively, with zero boundary condition on $\partial\Omega$. That is, $H_i\gamma$ is the solution of the problem

$$\left. \begin{array}{ll} -\Delta\, H_i\gamma = 0 & \text{in } \Omega_i \\ H_i\gamma\,_{|\Gamma_i} = 0 & \\ H_i\gamma\,_{|\Gamma} = \gamma & \end{array} \right\} \qquad (i = 1, 2). \tag{5.2}$$

Then the Poincaré–Steklov operator is $R : H_{00}^{1/2}(\Gamma) \to H_{00}^{-1/2}(\Gamma)$ that assigns to γ the jump of the normal derivatives of its harmonic extensions on Γ, i.e.

$$R\gamma := \frac{\partial}{\partial n} H_1\gamma + \frac{\partial}{\partial n} H_2\gamma \qquad \text{on } \Gamma. \tag{5.3}$$

(The plus sign expresses the jump with the convention that the outward normal vector n w.r.t. Ω_1 is opposite to n w.r.t. Ω_2 on Γ, which will be understood throughout this paper. That is, for a smooth function on Ω, the two normal derivatives are the opposite of each other and hence the jump on Γ equals zero.)

Remark 5.1 Problem (5.1) can then be reduced to equation

$$R\gamma = \psi \tag{5.4}$$

with ψ defined as follows. Let $T_1 f$ and $T_2 f$, respectively, denote the solutions of the problems

$$\left. \begin{array}{ll} -\Delta\, T_i f = f & \text{in } \Omega_i \\ T_i f\,_{|\partial\Omega_i} = 0 & \end{array} \right\} \qquad (i = 1, 2), \tag{5.5}$$

and let

$$\psi := -\frac{\partial}{\partial n} T_2 f - \frac{\partial}{\partial n} T_1 f \qquad \text{on } \Gamma$$

(which is the negative jump of the corresponding normal derivatives). Then $u := H_i\gamma + T_i f$ on Ω_i $(i = 1, 2)$ satisfies $-\Delta u = f$ on both Ω_1 and Ω_2 and is continuous on Ω. Hence u solves (5.1) if and only if its normal derivative has zero jump on Γ, which is equivalent to (5.4).

Remark 5.2 Green's formula implies that the bilinear form of the Poincaré–Steklov operator R is

$$\langle R\gamma, \mu\rangle = \int_{\Omega_1} \nabla H_1\gamma \cdot \nabla H_1\mu + \int_{\Omega_2} \nabla H_2\gamma \cdot \nabla H_2\mu \qquad (\gamma, \mu \in H_{00}^{1/2}(\Gamma)), \tag{5.6}$$

whence R is a symmetric and strictly positive operator.

On the discrete level, let us now consider a FEM discretization of problem (5.1) and let us decompose the stiffness matrix as

$$A = \begin{pmatrix} A_{11} & 0 & A_{1\Gamma} \\ 0 & A_{22} & A_{2\Gamma} \\ A_{\Gamma 1} & A_{\Gamma 2} & A_{\Gamma\Gamma} \end{pmatrix}, \tag{5.7}$$

corresponding to the node points in Ω_1, in Ω_2 and on Γ, respectively. The linear system can be reduced to the Schur complement

$$\Sigma := A_{\Gamma\Gamma} - A_{\Gamma 1}A_{11}^{-1}A_{1\Gamma} - A_{\Gamma 2}A_{22}^{-1}A_{2\Gamma}, \tag{5.8}$$

i.e. Σ is the Schur complement for Γ w.r.t. both Ω_1 and Ω_2. Then, as pointed out in [15], Σ is the discrete analogue of the Poincaré–Steklov operator (5.3). Essentially, the term $A_{\Gamma\Gamma}$ is responsible for the boundary values of the considered function and the two other terms represent the procedures involving the two harmonic extensions.

Remark 5.3 The generalization of the above notions to the case of more (say, k) subdomains is straightforward. Then the Poincaré–Steklov operator involves harmonic extensions from the union of interfaces to all subdomains, and its bilinear formulation will contain a sum of k terms, e.g. for $k = 3$ the form (5.6) is replaced by

$$\langle R\gamma, \mu\rangle = \int_{\Omega_1} \nabla H_1\gamma \cdot \nabla H_1\mu + \int_{\Omega_2} \nabla H_2\gamma \cdot \nabla H_2\mu + \int_{\Omega_3} \nabla H_3\gamma \cdot \nabla H_3\mu\,. \tag{5.9}$$

Similarly, the stiffness matrix (5.7) and the corresponding Schur complement (5.8) will include k interior blocks A_{ii}: e.g., for the above example $k = 3$, we have

$$\Sigma := A_{\Gamma\Gamma} - A_{\Gamma 1}A_{11}^{-1}A_{1\Gamma} - A_{\Gamma 2}A_{22}^{-1}A_{2\Gamma} - A_{\Gamma 3}A_{33}^{-1}A_{3\Gamma} \tag{5.10}$$

as in (4.6).

5.2 The modified Poincaré–Steklov operator

Let us consider again a FEM discretization of problem (5.1). We decompose the domain Ω in subdomains $\Omega_1, \ldots, \Omega_m$ such that, in addition to a corresponding portion of the original boundary $\partial\Omega$, each Ω_i has a common boundary only with its neighbours Ω_{i-1} and Ω_{i+1}. Denoting here these common boundaries by $\Gamma_{i-1,i}$ and $\Gamma_{i,i+1}$, respectively, we decompose the stiffness

matrix as in (3.1), corresponding to the subdomains $\Omega_1, \dots, \Omega_m$ such that the node points on $\Gamma_{i,i+1}$ are taken into account in A_{ii} (i.e. together with Ω_i). Our goal is to study the factorization (3.2). Since, in contrast to the idea of (5.7), the boundary node points are not considered here separately, the Schur complements in (3.2) are understood recursively as complements for Ω_i w.r.t. Ω_{i-1}. This is an important difference as compared to (5.8), and therefore the continuous analogues of the Schur complements in (3.2) will also be appropriate modifications of the Poincaré–Steklov operator (5.3). In fact, the proper operator takes into account the previous subdomain Ω_{i-1} only.

First, for simplicity, let us consider the case of two subdomains Ω_1 and Ω_2, where one can follow more clearly how the operator in subsection 5.1 is modified. Similarly as therein, the common boundary of Ω_1 and Ω_2 is denoted by Γ, further, we let $\Gamma_1 := \partial\Omega_1 \setminus \Gamma$ and $\Gamma_2 := \partial\Omega_2 \setminus \Gamma$. We wish to define the continuous analogue of the Schur complement $S_2 := A_{22} - A_{21}A_{11}^{-1}A_{12}$.

Let us take a function u_2 on Ω_2 such that $u_{2\,|\Gamma_2} = 0$. Applying the operator $-\Delta_{|\Omega_2}$ to u_2 (which corrresponds to the term A_{22} in S_2), we want it to equal f. Let us further consider the restriction $u_{2\,|\Gamma}$, and calculate its harmonic extension to Ω_1, i.e., let H_1u_2 be the solution of the problem

$$\left.\begin{aligned} -\Delta\, H_1u_2 &= 0 \quad \text{in } \Omega_1 \\ H_1u_{2\,|\Gamma_1} &= 0 \\ H_1u_{2\,|\Gamma} &= u_2\,. \end{aligned}\right\} \tag{5.11}$$

(That is, we solve the analogue of (5.2) only on Ω_1.) Accordingly, the modified Poincaré–Steklov operator P assigns to u_2 the jump of the normal derivative of its harmonic extension and of itself, i.e.

$$Pu_2 := \frac{\partial}{\partial n}\, H_1u_2 + \frac{\partial}{\partial n}\, u_2 \qquad \text{on } \Gamma. \tag{5.12}$$

Remark 5.4 Similarly as in Remark 5.1, problem (5.1) can now be reduced to the equation

$$Pu_2 = \chi \tag{5.13}$$

where $\chi := -\frac{\partial}{\partial n}\, T_1f$ with T_1f defined in (5.5). Letting $u = u_1 := H_1u_2 + T_1f$ on Ω_1 and $u = u_2$ on Ω_2, it is readily seen that u solves (5.1) if and only if $Lu_2 = f$ in Ω_2 and (5.13) holds on Γ.

Remark 5.5 The analogue of Remark 5.2 holds if, according to our setting, we handle the operators $-\Delta_{|\Omega_2}$ and P together. Using Green's formula, the pair $\tilde{P}$ of these operators satisfies

$$\begin{aligned} \Big\langle \tilde{P}\,(u_2,\, u_{2\,|\Gamma}),\, (\varphi,\, \varphi_{|\Gamma}) \Big\rangle \equiv \Big\langle \begin{pmatrix} -\Delta \\ P \end{pmatrix} (u_2,\, u_{2\,|\Gamma}),\, (\varphi,\, \varphi_{|\Gamma}) \Big\rangle &= \int_{\Omega_2} (-\Delta u_2)\varphi + \int_\Gamma (Pu_2)\varphi \\ &= \int_{\Omega_1} \nabla H_1u_2 \cdot \nabla H_1\varphi + \int_{\Omega_2} \nabla u_2 \cdot \nabla\varphi \end{aligned} \tag{5.14}$$

(for all $\varphi \in H^1_D(\Omega_2) := \{\varphi \in H^1(\Omega_2) :\ \varphi_{|\Gamma_2} = 0\}$), whence it is a symmetric and strictly positive operator.

Remark 5.6 For more subdomains, one can define P_i in just an analogous way. Namely, for simplicity, let Γ_{i-1} denote the common boundary of Ω_{i-1} and Ω_i. Letting u_i be defined on Ω_i such that $u_{i\,|\partial\Omega_i \setminus \Gamma_{i-1}} = 0$, we consider $u_{i\,|\Gamma_{i-1}} = 0$ and solve the Dirichlet problem on $\Omega_1 \cup \dots \cup \Omega_{i-1}$ with this boundary condition (which can be reduced to previous subproblems in a recursive way, just as is the Schur complement reduced to previous Schur complements), and finally calculate

the jump of the corresponding normal derivatives on Γ. Here the bilinear form that replaces (5.14) will thus include a term on $\Omega_1 \cup \cdots \cup \Omega_{i-1}$ and a term on Ω_i: for instance, in the case of three subdomains, we have

$$\left\langle \tilde{P}_3 \, (u_3, \, u_{3\,|\Gamma}), \, (\varphi, \, \varphi_{|\Gamma}) \right\rangle = \int_{\Omega_1 \cup \Omega_2} \nabla H_{12} u_3 \cdot \nabla H_{12} \, \varphi + \int_{\Omega_3} \nabla u_3 \cdot \nabla \varphi \tag{5.15}$$

(for all $\varphi \in H^1_D(\Omega_3) := \{\varphi \in H^1(\Omega_3) : \ \varphi_{|\partial\Omega_3 \setminus \Gamma_2} = 0\}$) where $H_{12} u_3$ denotes the harmonic extension of $u_{3\,|\Gamma_2}$ to $\Omega_1 \cup \Omega_2$.

Remark 5.7 For problems with jumps in the diffusion coefficients, the conditioning properties observed in Section 4 are in accordance with their analogues on the continuous level. This will be outlined here. Namely, we have observed in Section 4 that the condition numbers of the Schur complements are sensitive to jumps in the first approach but not in the second approach. Accordingly, one can indicate for the same example that the standard Poincaré–Steklov operator is sensitive to the jumps whereas the modified Poincaré–Steklov operator is not.

Let us therefore consider the model problem of Section 4. The domain Ω is decomposed in three subdomains Ω_1, Ω_2 and Ω_3, such that there are common boundaries $\Gamma_1 := \overline{\Omega}_1 \cap \overline{\Omega}_2$ and $\Gamma_2 := \overline{\Omega}_2 \cap \overline{\Omega}_3$, but Ω_1 and Ω_3 have no common boundary. We consider an elliptic problem, formally as $-\mathrm{div}\,(w \, \nabla u) = f$ with $u_{|\partial\Omega} = 0$, with weak form (4.15), where w is a weight function on Ω such that $w_{|\Omega_i} \equiv w_i$ $(i = 1, 2, 3)$. We assume $w_1 \geq w_2 \geq w_3$ and, varying the coefficients, we are interested in the case $w_1/w_2 \to \infty$.

The standard Poincaré–Steklov operator can be extended directly to such piecewise constant coefficient problems, such that one considers weighted normal derivatives on the interfaces with weights w_i. Considering the bilinear form for our model problem with three subdomains, the form (5.9) is replaced by

$$\langle R(w)\gamma, \mu \rangle = w_1 \int_{\Omega_1} \nabla H_1 \gamma \cdot \nabla H_1 \mu + w_2 \int_{\Omega_2} \nabla H_2 \gamma \cdot \nabla H_2 \mu + w_3 \int_{\Omega_3} \nabla H_3 \gamma \cdot \nabla H_3 \mu \, . \tag{5.16}$$

Factoring out w_2, we see that $R(w)$ is the constant multiple of an operator where the first term is proportional to w_1/w_2 and the other two terms are bounded as $w_1/w_2 \to \infty$, i.e. $R(w)$ behaves similarly as $\Sigma(w)$ in Corollary 4.1.

The modified Poincaré–Steklov operator can be extended similarly to piecewise constant coefficient problems, using the same weighted normal derivatives as above. The bilinear form for our model problem with three subdomains is the proper modification of (5.15):

$$\left\langle \tilde{P}_3(w) \, (u_3, \, u_{3\,|\Gamma}), \, (\varphi, \, \varphi_{|\Gamma}) \right\rangle = \int_{\Omega_1 \cup \Omega_2} w \, \nabla H_{12} u_3 \cdot \nabla H_{12} \, \varphi + w_3 \int_{\Omega_3} \nabla u_3 \cdot \nabla \varphi \tag{5.17}$$

for all $\varphi \in H^1_D(\Omega_3) := \{\varphi \in H^1(\Omega_3) : \ \varphi_{|\partial\Omega_3 \setminus \Gamma_2} = 0\}$ where $H_{12} u_3$ denotes the "w-harmonic" extension of $u_{3\,|\Gamma_2}$ to $\Omega_1 \cup \Omega_2$, that is, $H_{12} u_3 = v$ if and only if $v_{|\Gamma_2} = u_3$ and $v_{|\partial(\Omega_1 \cup \Omega_2) \setminus \Gamma_2} = 0$, and further

$$\int_{\Omega_1 \cup \Omega_2} w \, \nabla v \cdot \nabla \phi \equiv w_1 \int_{\Omega_1} \nabla v \cdot \nabla \phi + w_2 \int_{\Omega_2} \nabla v \cdot \nabla \phi = 0 \qquad \forall \phi \in H^1_0(\Omega_1 \cup \Omega_2). \tag{5.18}$$

Let us now consider an arbitrary test function $\varphi \in H^1_D(\Omega_3)$ as required for (5.17), and denote by $\tilde{\varphi}$ an extension of φ to Ω such that $\tilde{\varphi}_{|\Omega_1} \equiv 0$ and $\tilde{\varphi}_{|\partial\Omega} \equiv 0$. Then $\tilde{\varphi}$ coincides with the w-harmonic extension $H_{12}\varphi$ on Γ_2, and also on $\partial(\Omega_1 \cup \Omega_2) \setminus \Gamma_2$ since both vanish on the latter.

Hence $H_{12}\varphi - \tilde{\varphi}$ equals zero on the entire $\partial(\Omega_1 \cup \Omega_2)$, i.e. $H_{12}\varphi - \tilde{\varphi} \in H_0^1(\Omega_1 \cup \Omega_2)$. Setting $\phi := H_{12}\varphi - \tilde{\varphi}$ in (5.18) and using $\tilde{\varphi}_{|\Omega_1} \equiv 0$, we obtain

$$\int_{\Omega_1 \cup \Omega_2} w \nabla v \cdot \nabla H_{12}\varphi = \int_{\Omega_1 \cup \Omega_2} w \nabla v \cdot \nabla \tilde{\varphi} = w_2 \int_{\Omega_2} \nabla v \cdot \nabla \tilde{\varphi}\,.$$

Since by definition $H_{12}u_3 = v$, we have just obtained an equality for the first term of (5.17). Substituting this into the whole expression in (5.17), we obtain a form for $\tilde{P}_3(w)$ that contains integrals only on Ω_2 and Ω_3 with respective weights w_2 and w_3:

$$\left\langle \tilde{P}_3(w)\,(u_3,\, u_{3\,|\Gamma}),\, (\varphi,\, \varphi_{|\Gamma}) \right\rangle = w_2 \int_{\Omega_2} \nabla v \cdot \nabla \tilde{\varphi} + w_3 \int_{\Omega_3} \nabla u_3 \cdot \nabla \varphi\,. \tag{5.19}$$

To sum up, the behaviour of the Schur complements under jumps in Section 4 follows that of their continuous analogues.

5.3 Approximate modified Poincaré–Steklov operator on a model problem

In this subsection we consider a continuous analogue of the procedure (3.5)-(3.6), and show on a model problem that it can be carried out in a similar way as on the discrete level. This gives an alternate illustration for the fact that the condition numbers in Theorem 2.1 are mesh independent.

Let us consider the 3D model problem

$$\begin{cases} -\Delta u = f \quad \text{in } B \\ u_{|\partial B} = 0 \end{cases} \tag{5.20}$$

where $B \subset \mathbf{R}^3$ is the unit ball. Let us fix a positive integer k and numbers $0 = R_0 < R_1 < \cdots < R_{k-1} < R_k = 1$. Using notation $r := |x|$ for the Euclidean norm of vectors $x \in \mathbf{R}^3$, we define annular subdomains

$$\Omega_j := \{x \in B :\ R_{k-j} < r < R_{k-j+1}\} \qquad (i = 1, \ldots, k). \tag{5.21}$$

First, for simplicity, let $k = 2$ and $R_1 = 1/2$. Then (3.6) becomes $D_2\mathbf{e} = A_{2,1}A_{11}^{-1}A_{1,2}\mathbf{e}$ for the constant vector $\mathbf{e} = (1, \ldots, 1)$. Its continuous analogue, with the notations of subsection 5.2, is to find an operator $\hat{D}_2$ such that

$$\hat{D}_2 e = Pe \qquad \text{on } \Gamma \tag{5.22}$$

for the constant function $e \equiv 1$. Here $\Gamma := \{x \in \mathbf{R}^3 :\ r = 1/2\}$, and P is defined in (5.12) and the procedure before that. We have

$$\Omega_1 = \{x \in B :\ 1/2 < r < 1\} \quad \text{and} \quad \Omega_2 = \{x \in B :\ 0 < r < 1/2\}, \tag{5.23}$$

further, $\Gamma_1 := \partial\Omega_1 \setminus \Gamma = \partial B$ and $\Gamma_2 := \partial\Omega_2 \setminus \Gamma = \emptyset$. Then Pe can be calculated explicitly. First, the harmonic extension of e to Ω_1 is $H_1 e =: v$, where v is the solution of

$$\begin{cases} -\Delta\, v = 0 \quad \text{in } \Omega_1 \\ v_{|\partial B} = 0 \\ v_{|\Gamma} = 1\,. \end{cases} \tag{5.24}$$

Here we use the form of the Laplace operator in 3D spherical coordinates, which reduces to $\Delta v = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2\frac{\partial v}{\partial r})$ for radially symmetric functions. Then an elementary calculation yields

$$v(r) = \frac{1}{r} - 1$$

hence by (5.12) and using that now $\frac{\partial}{\partial n} = -\frac{\partial}{\partial r}$ on Γ, we obtain $Pe = -\big(\frac{\partial v}{\partial r} + \frac{\partial e}{\partial r}\big)_{|r=1/2} = 4$. That is, Pe is constant on Γ, i.e. we can write $Pe = 4e$ on Γ, which means that the operator required in (5.22) can be defined as

$$\hat{D}_2 := 4I \tag{5.25}$$

where I is the identity operator on Γ.

Our goal now is to verify a continuous analogue of condition (3.7). According to the above, the operator $4I$ corresponds to D_2, further, as seen before, the analogue of A_{22} is the operator $-\Delta$ such that homogeneous Dirichlet boundary conditions are considered on $\partial\Omega_2 = \partial B$. Hence the required analogue of (3.7) reads as

$$4I \le -\varrho\Delta \quad \text{for some} \quad \varrho < 1/2. \tag{5.26}$$

Denoting by λ_1 the smallest eigenvalue of $-\Delta$ with the given boundary conditions, and taking into account the condition $\varrho < 1/2$, inequality (5.26) is equivalent to $8 < \lambda_1$. Here the eigenfunctions of $-\Delta$ are the restrictions to Ω_2 of the eigenfunctions on B with homogeneous Dirichlet boundary conditions on ∂B. The first eigenfunction is the first three-dimensional Bessel function $w(r) := \frac{\sin \pi r}{\pi r}$ with eigenvalue $\lambda_1 = \pi^2 > 8$, therefore (5.26) is satisfied.

Now let us consider more subdomains. Here by (5.21),

$$\Omega_k := \{x \in B : \ 0 < r < R_1\}. \tag{5.27}$$

In order to determine the operator P_k, problem (5.24) has now to be solved with Γ replaced by $\Gamma_{k-1} = \{x \in \mathbf{R}^3 : \ r = R_1\}$. The solution is

$$v(r) = \frac{\frac{1}{r} - 1}{\frac{1}{R_1} - 1},$$

hence the constant 4 in (5.25) is replaced by $\frac{1}{R_1(1-R_1)}$, and accordingly, the above property $8 < \pi^2$ is replaced by condition

$$2 < \pi^2 R_1(1 - R_1). \tag{5.28}$$

If this holds then the operator $\hat{D}_k := \frac{1}{R_1(1-R_1)}\, I$ satisfies the required analogue of (3.7), i.e. $\hat{D}_k \le -\varrho\Delta$ for some $\varrho < 1/2$. Analogous calculations can be carried out to find $\hat{D}_1, \dots, \hat{D}_{k-1}$.

Inequality (5.28) is satisfied if, up to four digits, $0.2824 < R_1 < 0.7176$. Concerning the case of several subdomains, one may define $R_j := (\frac{j}{k})^{1/3}$ in (5.21) to have equal volume of the subdomains for technical convenience. Then the condition $0.2824 < R_1 = (\frac{1}{k})^{1/3}$ is satisfied up to $k = 44$, i.e. (5.28) is satisfied for any reasonable number of subdomains.

6 Concluding remarks; multilevel methods and parallelism

Using pure algebraic tools, condition number bounds have been established for approximate block factorizations for matrices in block tridiagonal form. The condition number of the corresponding preconditioned matrix depends linearly on the number of blocks but does not depend

on the order of the system. An important application is for the solution of elliptic problems where the domains have been partitioned in substructures using unidirectional stripes. For discontinuous coefficient problems, the partitioning can often be done so that the diffusion coefficient is constant on each subdomain. The condition number then depends at most linearly on the number of subdomains. The method is applicable for both 2D and 3D problems. Each subdomain problem can be solved readily by a direct solution method or possibly by some, inner, iterative method. It has been shown in various publications, see e.g. [7], that this influences little the rate of convergence. The performance of the standard domain decomposition method using iterative methods for the resulting global Schur complement matrix depends on the ratios of diffusion coefficients, while the present method, using only local Schur complements, does not depend on those.

Although this has not been dealt with in the paper, the proposed method can be coupled with a multilevel approach. For the standard global Schur complement domain decomposition, this has been considered in a number of publications, see e.g. [9, 12]. Here the local subdomain problems are coupled with a coarse problem that, as in multigrid methods, is used to propagate the information globally. In this way the bounds become independent of coefficient jumps between subdomains and may grow only with the square of the number of subdivisions.

The present method can also be combined with a multilevel approach. Assuming for simplicity just two levels of meshes, we order the meshpoints that are not in the coarse set first and the matrix takes the form

$$\begin{pmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{22} & \hat{A}_{22} \end{pmatrix}.$$

This matrix can be factorized approximately and the arising Schur complement matrix $\hat{A}_{22} - \hat{A}_{21}\hat{A}_{11}^{-1}\hat{A}_{12}$ can be preconditioned by the coarse mesh FE matrix. For the solution of the matrix $\hat{A}_{11}$, which corresponds to all meshpoints that are not in the coarse mesh, we can use the incomplete factorization method presented in this paper. It suffices normally to make a subdivision with few blocks. Each subdomain problem can be solved by a direct solution method or some inner, iterative method based on another incomplete factorization method.

One advantage with domain decomposition methods is that it can have a high degree of parallelism. As the present method makes use of a recursive computation, possible parallelism is available only on the local subdomain level, if a proper version of incomplete factorization, such as based on a red-black reordering of meshpoints, is used there. An additional parallelism is achieved if the substructures are ordered from both ends of the domain, preferably both with Dirichlet boundary conditions, to meet in the middle of the domain. This enables a two-fold parallel implementation. For certain elasticity problems in 2D or 3D, one can use a preconditioning method based on separate displacements, enabling the solution of the 2 or 3 separate displacement problems, arising in the preconditioner, in parallel. See [4] for further details on separate displacement preconditions.

Acknowledgements. The authors acknowledge the generous provisions at the *Mathematisches Forschungsinstitut Oberwolfach* which enabled writing this paper during their 'Research in Pairs' stay at the Institute from 24 February–8 March 2008. The second author was also supported by the Hungarian Research Grant OTKA No.K 67819.

References

[1] Axelsson, O., *Iterative Solution Methods,* Cambridge University Press, 1994.

[2] Axelsson, O., Faragó I., Karátson J., Sobolev space preconditioning for Newton's method using domain decomposition, *Numer. Lin. Alg. Appl.*, 9 (2002), 585-598.

[3] Axelsson, O., Hakopian, Yu. R., Kuznetsov, Yu. A., Multilevel preconditioning for perturbed finite element matrices, *IMA J. Numer. Anal.* 17 (1997), no. 1, 125–149.

[4] Axelsson, O., Karátson J., Conditioning analysis of separate displacement preconditioners for some nonlinear elasticity systems, *Math. Comput. Simul.* 64 (2004), No.6, pp. 649-668.

[5] Axelsson, O., Kolotilina, L., Diagonally compensated reduction and related preconditioning methods, *Numer. Linear Algebra Appl.* 1 (1994), no. 2, 155–177.

[6] Axelsson, O., Lu, H., A survey of some estimates of eigenvalues and condition numbers for certain preconditioned matrices, *J. Comput. Appl. Math.* 80 (1997), no. 2, 241–264.

[7] Axelsson, O., Vassilevski, P. S., Variable-step multilevel preconditioning methods. I. Selfadjoint and positive definite elliptic problems, *Numer. Linear Algebra Appl.* 1 (1994), no. 1, 75–101.

[8] Dryja, M., An iterative substructuring method for elliptic mortar finite element problems with discontinuous coefficients, in: *Domain decomposition methods*, pp. 94-103; Contemp. Math. 218, Amer. Math. Soc., Providence, RI, 1998.

[9] Dryja, M., Sarkis, M. V., Widlund, O. B., Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions, *Numer. Math.* 72 (1996), no. 3, 313–348.

[10] Langer, U., Steinbach, O., Coupled finite and boundary element domain decomposition methods, in: *Boundary Element Analysis*, pp. 61–95; Lect. Notes Appl. Comput. Mech. 29, Springer, Berlin, 2007;

[11] Lu, H., Axelsson, O., Conditioning analysis of block incomplete factorizations and its application to elliptic equations, *Numer. Math.* 78 (1997), no. 2, 189–209.

[12] Mandel, J., Brezina, M., Balancing domain decomposition for problems with large jumps in coefficients, *Math. Comp.* 65 (1996), no. 216, 1387–1401.

[13] Mandel, J., Dohrmann, C. R., Convergence of a balancing domain decomposition by constraints and energy minimization, *Numer. Linear Algebra Appl.* 10 (2003), no. 7, 639–659.

[14] Toselli, A., Widlund, O., *Domain Decomposition Methods – Algorithms and Theory,* Springer Series in Computational Mathematics 34, Springer-Verlag, Berlin, 2005.

[15] Quarteroni, A., Valli, A., *Domain Decomposition Methods for Partial Differential Equations,* Oxford University Press, New York, 1999.