SUBSTRUCTURING PRECONDITIONER FOR NONCONFORMING FINITE ELEMENT APPROXIMATIONS OF SECOND ORDER ELLIPTIC PROBLEMS WITH ANISOTROPY

Y.A. KUZNETSOV AND S.Y. MALIASSOV*

Abstract. In this paper an algebraic substructuring preconditioner is considered for nonconforming finite element approximations of second order elliptic problems in 3D domains with diagonal anisotropic diffusion tensor. Using a block Gauss elimination and a substructuring idea part of the unknowns is eliminated and the Schur complement obtained is preconditioned by a spectrally equivalent matrix. When the domain considered is a parallelepiped and boundary conditions are uniform on entire faces this matrix is separable. Explicit estimates of condition numbers and implementation algorithms are established for the constructed preconditioner. It is shown that the condition number of the preconditioned matrix does not depend neither on meshsize parameter nor on the coefficients of the diffusion tensor. The numerical experiments show that proposed preconditioner is rather efficient and can be used for development of iteration solvers for general elliptic equations of second order on domains, topologically equivalent to a parallelepiped.

Key words. Second order elliptic problem, nonconforming finite element method, algebraic substructuring preconditioner, superelement analysis, separable matrix, condition number.

AMS(MOS) subject classifications. 65N30, 65N22, 65F10.

1. Introduction.

Let Ω be a convex bounded domain in $\mathbb{R}^3$ with boundary $\partial\Omega$. Consider an elliptic problem

$$
\begin{aligned}
-\nabla \cdot (K \cdot \nabla u) &= f && \text{in } \Omega, \\
u &= 0 && \text{on } \Gamma_0, \\
\frac{\partial u}{\partial n} &= 0 && \text{on } \Gamma_1,
\end{aligned}
\tag{1.1}
$$

where $K(\mathbf{x})$ is a uniformly positive definite bounded symmetric tensor, $f(\mathbf{x}) \in L^2(\Omega)$, $\overline{\Gamma_0 \cup \Gamma_1} = \partial\Omega$, $\Gamma_0 \cap \Gamma_1 = \emptyset$, and $\Gamma_0 \equiv \overline{\Gamma_0} \neq \emptyset$.

Note that an approach considered in this paper is valid also for the case of Neumann problem, i.e. $\Gamma_0 = \emptyset$, and it is not described here only for the sake of simplicity.

Let the bilinear form $a(\cdot,\cdot)$ is defined by

$$
a(u,v) = (K \cdot \nabla u, \nabla v), \qquad u, v \in V_0(\Omega) = \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_0\},
$$

where $(\cdot,\cdot)$ denotes the $L^2(\Omega)$ inner product. Then the usual weak form of (1.1) for the solution $u \in V_0(\Omega)$ is

$$
a(u,v) = (f,v), \qquad \forall v \in V_0(\Omega). \tag{1.2}
$$

Let $\mathcal{T}_T$ be a regular partitioning of Ω into simplexes T with a mesh size h and let $V_h(\Omega)$ be the P_1–nonconforming finite element space of functions $v \in L^2(\Omega)$ [1] such

* Yu.A. Kuznetsov, Institute of Numerical Mathematics, Russian Academy of Sciences, 32-a Leninski Prospect, 117334 Moscow, Russia. S.Yu. Maliassov, Institute for Scientific Computation and Department of Mathematics, Texas A&M University, 505 Blocker Bldg., College Station, TX 77843–3404, U.S.A.

that $v|_T$ are linear for all $T \in \mathcal{T}_T$, v are continuous at the barycenters of $T \in \mathcal{T}_T$ and vanish at the barycenters of the boundary faces on Γ_0. Note that the space $V_h(\Omega)$ is not a subspace of $H^1(\Omega)$.

Define the bilinear form on $V_h(\Omega)$ by

$$a_h(u,v) = \sum_{T\in\mathcal{T}_T} (K\cdot\nabla u, \nabla v)_T, \qquad \forall\ u,v \in V_h(\Omega), \tag{1.3}$$

where $(\cdot,\cdot)_T$ is the $L^2(T)$ inner product, $T \in \mathcal{T}_T$. Then the P_1–nonconforming finite element discretization of (1.1) is to find $u_h \in V_h$ such that

$$a_h(u_h, v) = (f, v), \qquad \forall v \in V_h(\Omega). \tag{1.4}$$

Once a basis $\{\varphi_i(\mathbf{x})\}_{i=1}^N$ for $V_h(\Omega)$ is chosen, (1.4) leads to a system of linear algebraic equations. Write $u(\mathbf{x}) = \sum_{i=1}^N\ u_i\varphi_i(\mathbf{x})$. Then (1.4) becomes

$$\sum_{i=1}^N u_i a_h(\varphi_i, \varphi_j) = (f, \varphi_j), \qquad j = 1, \ldots, N,$$

or in matrix representation

$$A\mathbf{u} = \mathbf{f}, \tag{1.5}$$

where $A_{ji} = a_h(\varphi_i, \varphi_j)$, $f_j = (f, \varphi_j)$, $i, j = 1, \ldots, N$.

Although the methods of solving (1.5) have been extensively studied in past few years (see, e.g., [1], [3], [5], [6], [8]), their efficiency depends on the coefficient matrix $K(\mathbf{x})$ and in the case of strong anisotropy in the coefficients the question of constructing effective solution techniques is still open.

In this paper we will describe and analyze a method of constructing the preconditioner for (1.5) using an idea of algebraic substructuring which can be described as follows [18].

Let us partition the domain Ω into subdomains Ω_s, $s = 1, \ldots, n$, in such a way that each Ω_s is a union of simplexes $T \in \mathcal{T}_T$,

$$\Omega = \bigcup_{s=1}^{n}\ \Omega_s, \qquad \Omega_s = \bigcup_{l=1}^{n_s} \{T_l \in \mathcal{T}_T : T_l \subset \Omega_s\}.$$

Below we call these subdomains Ω_s by superelements.

Let us introduce local stiffness matrices A_s on each superelement Ω_s by

$$(A_s\mathbf{u}_s, \mathbf{v}_s) = \sum_{T_l\subset\Omega_s} (K(\mathbf{x})\nabla u_h, \nabla v_h)_{T_l}, \qquad \forall u_h, v_h \in V_h(\Omega_s).$$

All these matrices are at least positive semidefinite and the global stiffness matrix is determined by assembling the local stiffness matrices over all superelements:

$$(A\mathbf{u}, \mathbf{v}) = \sum_{s=1}^n (A_s\mathbf{u}_s, \mathbf{v}_s), \qquad \forall \mathbf{u}, \mathbf{v} \in \mathrm{I\!R}^N.$$

We can symbolically write

$$A = \{A_s\}_{s=1}^n ,$$

where $\{\cdot\}_{s=1}^n$ means assembling with respect to the partitioning $\{\Omega_s\}_{s=1}^n$ of Ω.

In correspondence with this notation each superelement matrix A_s can be represented in terms of local stiffness matrices over simplexes T_l from Ω_s, i.e., $A_s = \{A_{sl}\}_{T_l \subset \Omega_s}$. Note that matrices A_{sl} also are at least positive semidefinite.

Following to [18], [16], let us introduce on each simplex $T \in \mathcal{T}_T$ another matrix $\hat{A}_{sl}$ which has the same kernel as A_{sl} (i.e., ker A_{sl} = ker $\hat{A}_{sl}$). Define on each superelement Ω_s matrix $\hat{A}_s$ by assembling $\hat{A}_{sl}$:

$$\hat{A}_s = \left\{\hat{A}_{sl}\right\}_{T_l \subset \Omega_s} .$$

Then it can be easily shown that ker A_s = ker $\hat{A}_s$, and matrices $\hat{A}_s$ are also at least positive semidefinite.

Let us define now $N \times N$ matrix $\hat{A}$ by assembling $\hat{A}_s$ over all superelements

$$\hat{A} = \left\{\hat{A}_s\right\}_{s=1}^n .$$

To obtain an estimation of the condition number of $\hat{A}^{-1}A$ we use so called superelement analysis which we outline here. Suppose we have two sequences of nonnegative numbers $\{a_i\}_{i=1}^n$ and $\{b_i\}_{i=1}^n$ such that a_i and b_i, $i = 1, \ldots, n$, are simultaneously either positive or zeroes. And suppose we are looking for estimates the ratio $\sum_{i=1}^n a_i / \sum_{i=1}^n b_i$ from below and from above. There is well known solution for this problem [14]:

$$\min_{\substack{i \\ b_i \neq 0}} \frac{a_i}{b_i} \leq \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i} \leq \max_{\substack{i \\ b_i \neq 0}} \frac{a_i}{b_i} .$$

Then we can formulate the following lemma:

LEMMA 1.1.

$$\max_{(\hat{A}\mathbf{u},\mathbf{u}) \neq 0} \frac{(A\mathbf{u},\mathbf{u})}{(\hat{A}\mathbf{u},\mathbf{u})} = \max_{(\hat{A}\mathbf{u},\mathbf{u}) \neq 0} \frac{\sum_{s=1}^n (A_s\mathbf{u}_s,\mathbf{u}_s)}{\sum_{s=1}^n (\hat{A}_s\mathbf{u}_s,\mathbf{u}_s)} \leq \max_{\substack{s=1,\ldots,n \\ (\hat{A}_s\mathbf{u}_s,\mathbf{u}_s) \neq 0}} \frac{(A_s\mathbf{u}_s,\mathbf{u}_s)}{(\hat{A}_s\mathbf{u}_s,\mathbf{u}_s)}, \tag{1.6}$$

and

$$\min_{(\hat{A}\mathbf{u},\mathbf{u}) \neq 0} \frac{(A\mathbf{u},\mathbf{u})}{(\hat{A}\mathbf{u},\mathbf{u})} = \min_{(\hat{A}\mathbf{u},\mathbf{u}) \neq 0} \frac{\sum_{s=1}^n (A_s\mathbf{u}_s,\mathbf{u}_s)}{\sum_{s=1}^n (\hat{A}_s\mathbf{u}_s,\mathbf{u}_s)} \geq \min_{\substack{s=1,\ldots,n \\ (\hat{A}_s\mathbf{u}_s,\mathbf{u}_s) \neq 0}} \frac{(A_s\mathbf{u}_s,\mathbf{u}_s)}{(\hat{A}_s\mathbf{u}_s,\mathbf{u}_s)}. \tag{1.7}$$

From Lemma 1.1 it is easy to see that to estimate the extreme eigenvalues of $\hat{A}^{-1}A$ it is sufficient to consider the local problems (1.6) on all superelements Ω_s, $s = 1, \ldots, n$.

Because of this fact superelement analysis is very useful and rather simple tool to get the condition numbers of preconditioned matrices (see, e.g., [2], [16], and [12]). It can be shown that to estimate the extreme eigenvalues of $\hat{A}^{-1}A$ it is sufficient to consider the worst cases when superelements Ω_s have no common faces with Γ_0.

So, if superelement matrices A_s and $\hat{A}_s$ are spectrally equivalent with respect to $\ker A_s$, that is there exist constants $c_{0,s}$ and $c_{1,s}$ such that

$$c_{0,s}(\hat{A}_s\mathbf{u}_s,\mathbf{u}_s) \le (A_s\mathbf{u}_s,\mathbf{u}_s) \le c_{1,s}(\hat{A}_s\mathbf{u}_s,\mathbf{u}_s), \qquad \forall \mathbf{u}_s \in \mathbb{R}^{N_s}, \qquad N_s = \dim\, \Omega_s,$$

where constants $c_{0,s}$, $c_{1,s}$ do not depend on mesh size parameter h then matrices $\hat{A}$ and A are also spectrally equivalent, i.e.,

$$c_0(\hat{A}\mathbf{u},\mathbf{u}) \le (A\mathbf{u},\mathbf{u}) \le c_1(\hat{A}\mathbf{u},\mathbf{u}), \qquad \forall \mathbf{u} \in \mathbb{R}^N,$$

with $c_0 = \min\limits_s c_{0,s}$ and $c_1 = \max\limits_s c_{1,s}$.

Let us partition now all unknowns of (1.5) into two groups $\mathbf{u} = (\mathbf{u}_1^T, \mathbf{u}_2^T)^T$, $\dim\, \mathbf{u}_1 = N_1$, $\dim\, \mathbf{u}_2 = N - N_1$, in such a way that the matrix $\hat{A}$ represented in a block form

$$\hat{A} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix} \tag{1.8}$$

has the block $\hat{A}_{22}$ easily invertible. Then introducing Schur complement $S = \hat{A}_{11} - \hat{A}_{12}\hat{A}_{22}^{-1}\hat{A}_{21}$ the matrix $\hat{A}$ can be rewritten in the form

$$\hat{A} = \begin{bmatrix} S + \hat{A}_{12}\hat{A}_{22}^{-1}\hat{A}_{21} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix}. \tag{1.9}$$

Following ideas in [4], [16], [17], let us construct matrix $\tilde{S}$ spectrally equivalent to S with constants $0 < d_0 \le d_1$ independent of mesh size parameter h

$$d_0(\tilde{S}\mathbf{v},\mathbf{v}) \le (S\mathbf{v},\mathbf{v}) \le d_1(\tilde{S}\mathbf{v},\mathbf{v}), \qquad \forall \mathbf{v} \in \mathbb{R}^{N_1}.$$

Then the matrix

$$B = \begin{bmatrix} \tilde{S} + \hat{A}_{12}\hat{A}_{22}^{-1}\hat{A}_{21} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix} \tag{1.10}$$

is spectrally equivalent to the matrix A, i.e.

$$r_0(B\mathbf{u},\mathbf{u}) \le (A\mathbf{u},\mathbf{u}) \le r_1(B\mathbf{u},\mathbf{u}), \qquad \forall \mathbf{u} \in \mathbb{R}^N,$$

where $r_0 = c_0 \min\{1; d_0\}$, $r_1 = c_1 \max\{1; d_1\}$. To construct such a matrix $\tilde{S}$ we again can use the idea of the algebraic substructuring described above.

Concluding this overview the algebraic substructuring procedure consists of the following main steps:

1. reconstruction of the directed graph of the matrix A from (1.5) in such a way that resulting matrix $\hat{A}$ has the same kernel and is still positive definite (or positive semidefinite if the matrix A is singular);

2. representation of the matrix $\hat{A}$ in a 2×2 block form (1.8) in such a way that one of the blocks $\hat{A}_{11}$ or $\hat{A}_{22}$ is easily invertible;
3. change the Schur complement S in (1.9) by a spectrally equivalent matrix $\tilde{S}$; we can use steps 1 and 2 to construct such a matrix $\tilde{S}$.

Note that we can first represent the matrix A in a 2×2 block form (1.9) and use steps 1-3 to construct preconditioner for the Schur complement $S = A_{11} - A_{12}A_{22}^{-1}A_{21}$. Implementing a finite number of these steps we can get as a result the matrix B which is spectrally equivalent to the given matrix A.

Because of algebraic nature of such a procedure this approach strongly depends on the structure of the graph of matrix A and, consequently, on the type of nonconforming finite element space V_h. A detailed description of constructing algebraic substructuring preconditioners for one concrete case of the P_1–nonconforming space V_h was given in [10], [11], [7]. In all these papers authors defined partitioning $\mathcal{T}_h$ of the whole domain subdividing it into topological parallelepipeds and splitting each parallelepiped in turn into **six** tetrahedra. The case of splitting each topological parallelepiped into **five** tetrahedra when $K(\mathbf{x})$ in (1.1) is a scalar tensor $K(\mathbf{x}) = k(\mathbf{x})I$ with piece wise constant function $k(\mathbf{x})$ is studied in [19]. The present paper extends these results to the case of splitting each topological parallelepiped into **five** tetrahedra when $K(\mathbf{x})$ is a full-rank symmetric tensor. Briefly, the approach used here to construct preconditioners includes three main stages. First, from the given matrix A the matrix B is constructed using the substructuring algorithm. Then the unknowns corresponding to the faces of the topological parallelepipeds are eliminated and an obtained Schur complement is preconditioned by a spectrally equivalent matrix. For the case of Ω being parallelepiped this matrix is separable.

The explicit bounds of spectrum of the preconditioned matrix are obtained with the help of the superelement analysis [10], [17], [19].

The outline of the reminder of the paper is as follows. In Section 2 we consider a formulation of the model problem with diagonal constant tensor when Ω is a unit cube. Then, in Section 3 we develop algebraic substructuring preconditioner for the resulting linear system and give an implementation algorithm. In Section 4 we consider the case of full tensor function $K(\mathbf{x})$ and domain Ω being subdivided into topological parallelepipeds. Finally, the results of the numerical experiments and some conclusions are given in Section 5 to illustrate the theory being presented.

2. Problem Formulation.

To explain our approach we consider the model case when Ω is a unit cube in $\mathbb{R}^3$, Γ_0 is a union of entire faces of Ω, the boundary conditions are uniform, and $K(\mathbf{x})$ is a diagonal tensor with the constant coefficients provided some additional assumptions.

Assumption (**A1**): Suppose that the matrix coefficient of the equation (1.1) is a diagonal tensor $K(\mathbf{x}) = \text{diag}\{k_1, k_2, k_3\}$, where k_i, $i = 1, 2, 3$, are constants over the cube Ω such that $\kappa = \min\,\{k_3/k_1, k_3/k_2\} \geq 1$.

Generally speaking, we need only the assumption that the coefficient k_* in some direction is not less then the coefficients in the other directions. For the sake of simplicity we assume that this is "z-direction".

Note that an extension of the method for the case of Ω being a union of parallelepipeds is straightforward.

Let $\mathcal{C}_h = \{C^{(i,j,k)}\}$ be a partition of Ω into uniform cubes with the length of the edge $h = 1/n$, where (x_i, y_j, z_k) is the right back upper corner of the cube $C^{(i,j,k)}$. Next, each cube $C^{(i,j,k)}$ is divided into 5 tetrahedra as shown in Figure 1 and denote this partitioning of Ω into tetrahedra by $\mathcal{T}_h$. Note that we have two types of the partitioning of the cubes $C^{(i,j,k)}$ into tetrahedra and the cube with one type of partitioning has all adjacent cubes of another type.

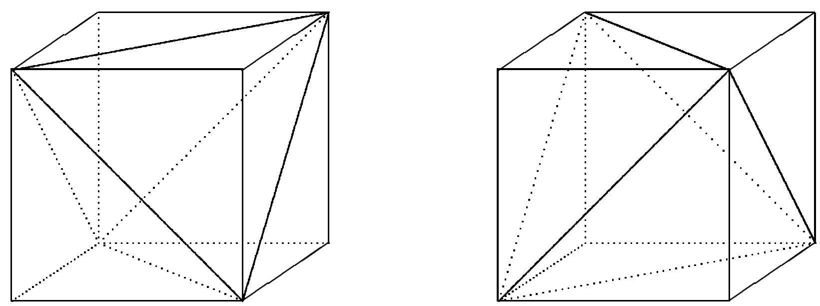


FIGURE 1. *Partition of cubes $C^{(i,j,k)}$ into tetrahedra.*

We introduce the set of barycenters of all faces of the tetrahedral partition of Ω, and the set Q_h of those barycenters that are not on Γ_0. The Crouzeix-Raviart P_1–nonconforming finite element space V_h is defined by

$$
(2.1)\quad V_h = \{v \in L^2(\Omega) : \ v|_T \in P_1(T), \ \forall T \in \mathcal{T}_h; \ v \text{ is continuous at the barycenters from } Q_h \text{ and vanishes at the barycenters of faces on } \Gamma_0\},
$$

and let its dimension be N. Note that $N \approx 10n^3$.

Now we define the bilinear form on V_h by

$$
(2.2)\qquad a_h(u,v) = \sum_{T\in\mathcal{T}_h} \int_T K(\mathbf{x})\,\nabla u \cdot \nabla v \; d\mathbf{x}, \qquad \forall\ u,v \in V_h.
$$

Thus the nonconforming discretization of the problem (1.1) is given by seeking $u_h \in V_h$ such that

$$
(2.3)\qquad a_h(u_h, v) = (f, v), \qquad \forall\ v \in V_h.
$$

For any function $v_h \in V_h$ we denote by $\mathbf{v} \in \mathbb{R}^N$ the corresponding vector of its degrees of freedom.

Let $(\mathbf{u},\mathbf{v})_N$ be a standard bilinear form defined on $\mathbb{R}^N$ by $(\mathbf{u},\mathbf{v})_N = \sum_{\mathbf{x}\in Q_h} u(\mathbf{x})v(\mathbf{x})$, $\forall u,v \in V_h$. Then the discretization operator $A : \mathbb{R}^N \to \mathbb{R}^N$, which is symmetric and positive definite is defined by

$$
(2.4)\qquad (A\mathbf{u},\mathbf{v})_N = a_h(u,v), \qquad u,v \in V_h.
$$

Similarly, we introduce the vector $\mathbf{f}$ as $(f, v) = (\mathbf{f}, \mathbf{v})_N$, $\forall\, v \in V_h$. Now, the problem (2.3) can be rewritten in a matrix form

$$A\mathbf{u} = \mathbf{f}. \tag{2.5}$$

For each cube $C = C^{(i,j,k)} \in \mathcal{C}_h$, denote by V_h^C the subspace of the restriction of the functions in V_h into C. For each $\mathbf{v} \in V_h^C$, we indicate by $\mathbf{v}_c$ the corresponding vector. The dimension of V_h^C is denoted by N_c. Obviously, for a cube without faces on Γ_0 we have $N_c = 16$.

The local stiffness matrix A^C on a cube $C \in \mathcal{C}_h$ is given by

$$(A^C\mathbf{u}_c, \mathbf{v}_c)_{N_c} = \sum_{T \subset C} (K(\mathbf{x})\nabla u_h, \nabla v_h)_T, \qquad \forall u_h, v_h \in V_h^C. \tag{2.6}$$

Note that matrices A^C are positive definite when $C \cap \Gamma_0 \neq 0$ and semidefinite otherwise. The global stiffness matrix is determined by assembling the local stiffness matrices:

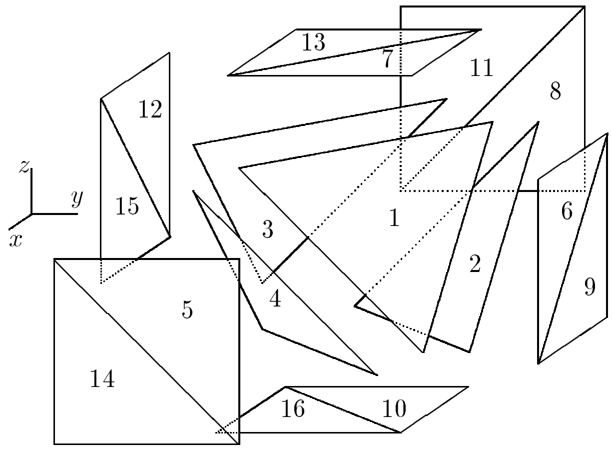
$$(A\mathbf{u}, \mathbf{v})_N = \sum_{C \in \mathcal{C}_h} (A^C\mathbf{u}_c, \mathbf{v}_c)_{N_c}, \qquad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^N. \tag{2.7}$$

3. Algebraic substructuring preconditioner over a cube. In this section we construct the algebraic substructuring preconditioner outlined in Introduction. Toward the end of the section, we divide all unknowns in the system into two groups:

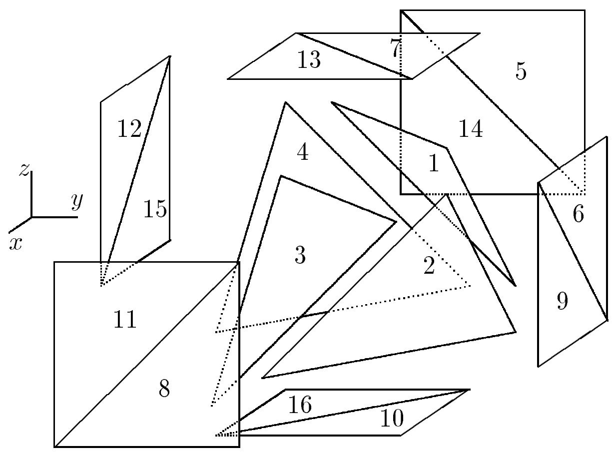
1. The first group consists of the unknowns corresponding to the faces of the tetrahedra that are internal for each cube (these are unknowns on faces 1, 2, 3 and 4 on Figure 2). We denote these unknowns by $VI_l^{(i,j,k)}$, $l = 1, 2, 3, 4$, $i, j, k = \overline{1, n}$.
2. The second group consists of all unknowns corresponding to faces of the cubes in the partition $\mathcal{C}_h$, without the faces on Γ_0 (Figure 2, faces $5, 6, \ldots, 16$).
   (a) First, we enumerate the unknowns on the faces perpendicular to x-axis (faces 5, 8, 11, 14). We denote these unknowns by $Vx_l^{(i,j,k)}$, $l = 1, 2$, $i = \overline{1, n-1}$, $j, k = \overline{1, n}$.
   (b) Second, we enumerate the unknowns on the faces perpendicular to y-axis (faces 6, 9, 12, 15). We denote these unknowns by $Vy_l^{(i,j,k)}$, $l = 1, 2$, $j = \overline{1, n-1}$, $i, k = \overline{1, n}$.
   (c) Finally, we enumerate the unknowns on the faces perpendicular to z-axis (faces 7, 10, 13, 16). We denote these unknowns by $Vz_l^{(i,j,k)}$, $l = 1, 2$, $k = \overline{1, n-1}$, $i, j = \overline{1, n}$.

Now we consider a cube C that has no face on the boundary $\partial\Omega$ and enumerate the faces s_j, $j = 1, \ldots, 16$ of the tetrahedra in this cube in correspondence with the partitioning introduced above as it is shown in Figure 2. Then the local stiffness matrix of this cube has the following form:

$$A^C = \frac{3h}{2} \begin{bmatrix} A_{11,c} & A_{12,c} \\ A_{21,c} & A_{22,c} \end{bmatrix}, \tag{3.1}$$



(a) Cube of type I



(b) Cube of type II

FIGURE 2. *Local enumeration of faces in cubes.*

where

$$A_{11,c} = (k_1 + k_2 + k_3)\, I_c + \frac{1}{2}\left[k_1 T_1 + k_2 T_2 + k_3 T_3\right], \tag{3.2}$$

$$T_1 = \begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}, T_2 = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}, T_3 = \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix},$$

$$I_c = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}, \qquad A_{22,c} = \begin{bmatrix} D & & & \\ & D & & \\ & & D & \\ & & & D \end{bmatrix}, \qquad D = \begin{bmatrix} k_1 & 0 & 0 \\ 0 & k_2 & 0 \\ 0 & 0 & k_3 \end{bmatrix},$$

$$A_{12,c} = \begin{bmatrix} -k_1 & -k_2 & -k_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -k_1 & -k_2 & -k_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -k_1 & -k_2 & -k_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -k_1 & -k_2 & -k_3 \end{bmatrix}.$$

Along with the matrix A^C we introduce on each cube $C \in \mathcal{C}_h$ the matrix B^C as

$$B^C = A^C + \frac{3h}{2}\begin{bmatrix} \tilde{B}_{11,c} & 0 \\ 0 & 0 \end{bmatrix}, \tag{3.3}$$

where $\tilde{B}_{11,c} = (k_1 + k_2)T_3$. So, matrix B^C can be represented in the form

$$B^C = \frac{3h}{2}\begin{bmatrix} B_{11,c} & A_{12,c} \\ A_{21,c} & A_{22,c} \end{bmatrix},$$

where $B_{11,c} = (k_1 + k_2 + k_3)\, I_{11,c} + \frac{1}{2}\left[k_1(T_1 + T_3) + k_2(T_2 + T_3) + k_3 T_3\right]$.

Note that $\ker A^C = \ker B^C$.

We now define the $N \times N$ matrix B by the following equality:

$$(B\mathbf{u}, \mathbf{v})_N = \sum_{C \in \mathcal{C}_h} (B^C \mathbf{u}_c, \mathbf{v}_c)_{N^C}, \qquad \forall \mathbf{u}, \mathbf{v} \in \mathrm{I\!R}^N. \tag{3.4}$$

From Lemma 1.1 we see that to estimate the condition number of $B^{-1}A$, it is sufficient to consider the local eigenvalue problems for $\mu_c \neq 0$

$$A^C \mathbf{u}_c = \mu_c B^C \mathbf{u}_c, \qquad \mathbf{u}_c \neq \mathbf{0}, \qquad \mathbf{u}_c \in \mathrm{I\!R}^{N_c}.$$

From (3.1) and (3.3), a direct calculations show that the eigenvalues μ_c are within the interval $[1/3, 1]$ provided the Assumption A1.

Then the inequalities (1.7) yield:

PROPOSITION 3.1. *Suppose that the matrix coefficient of the equation* (1.1) *is a diagonal tensor* $K(\mathbf{x}) = diag\{k_1, k_2, k_3\}$, *where* k_i, $i = 1, 2, 3$, *are constants over the cube* Ω *such that* $\kappa = min\,\{k_3/k_1, k_3/k_2\} \geq 1$.

Then eigenvalues of the problem

$$A\mathbf{u} = \mu B\mathbf{u} \tag{3.5}$$

belong to the interval $[\kappa/(2+\kappa), 1]$ *and the condition number is thus estimated by*

$$\mathrm{cond}(B^{-1}A) \leq 1 + 2/\kappa \leq 3.$$

We stress that the condition number of the matrix $B^{-1}A$ is bounded by a constant independent of the mesh size h and the value of the coefficients k_i, $i = 1, 2, 3$, when $k_3 \geq \max\,\{k_1, k_2\}$.

The splitting of the space $\mathrm{I\!R}^N$ induces the presentation of the vectors: $\mathbf{v}^T = (\mathbf{v}_1^T, \mathbf{v}_2^T)$, where $\mathbf{v}_1 \in \mathrm{I\!R}^{N_1}$ and $\mathbf{v}_2 \in \mathrm{I\!R}^{N_2}$, where $\mathbf{v}_2$ corresponds to the unknowns of the 2-nd group. Obviously, $N_1 = 4n^3$ and $N_2 = N - 4n^3$. Then the matrices A and B can be presented in the following block form:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad B = \begin{bmatrix} B_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{3.6}$$

where $B_{11} : \mathbb{R}^{N_1} \to \mathbb{R}^{N_1}$.

Denote now by $\hat{B}_{11} = B_{11} - A_{12}A_{22}^{-1}A_{21}$ the Schur complement of B obtained by elimination of the vector $\mathbf{v}_2$. Then $B_{11} = \hat{B}_{11} + A_{12}A_{22}^{-1}A_{21}$, so the matrix B has the form

$$B = \begin{bmatrix} \hat{B}_{11} + A_{12}A_{22}^{-1}A_{21} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}. \tag{3.7}$$

Note that for each cube $C \in \mathcal{C}_h$ the unknowns of the 2nd group (unknowns on the faces $5, 6, \ldots, 16$, in the local enumeration, see Figure 2) are connected only with the unknowns of the 1st group and therefore the matrix A_{22} is diagonal and can be inverted locally (face by face). Thus, matrix $\hat{B}_{11}$ is easily computable. Important fact which can be established by direct computations is that the matrix $\hat{B}_{11}$ can be obtained by assembling over all cubes local matrices $\hat{B}_{11,c} = B_{11,c} - A_{12,c}A_{22,c}^{-1}A_{21,c}$:

$$(\hat{B}_{11}\mathbf{u}_1, \mathbf{v}_1) = \sum_{C \in \mathcal{C}_h} (\hat{B}_{11,c}\mathbf{u}_{1,c}, \mathbf{v}_{1,c}), \qquad \forall \mathbf{u}_1, \mathbf{v}_1 \in \mathbb{R}^{N_1}.$$

Here $\mathbf{u}_{1,c}$ is a restriction of $\mathbf{u}_1$ into the nodes of the first group of the cube $C \in \mathcal{C}_h$ and $\dim \mathbf{u}_{1,c} = 4$.

Although the dimension of the matrix $\hat{B}_{11}$ is approximately as 2.5 times smaller as the dimension of the matrix A, it is still hard to solve the system of the linear equations with this matrix. Below we show that using algebraic substructuring we can construct a sparse separable matrix $\tilde{B}_{11}$ spectrally equivalent to $\hat{B}_{11}$ so that the resulting matrix

$$\tilde{B} = \begin{bmatrix} \tilde{B}_{11} + A_{12}A_{22}^{-1}A_{21} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{3.8}$$

is spectrally equivalent to the initial matrix A. In this case we shall use method of separation of variables to solve the system of linear equations with matrix $\tilde{B}_{11}$.

First, consider the linear system

$$B\mathbf{v} = \mathbf{g}. \tag{3.9}$$

Let us write explicitly the elements of $B\mathbf{v}$ in the case of Dirichlet boundary conditions on all boundary $\partial\Omega$ in terms of the unknowns introduced earlier in this section, i.e. in terms of

$$gi_\ell^{(i,j,k)}, \quad VI_\ell^{(i,j,k)}, \quad \ell = 1,2,3,4 \quad i,j,k = \overline{1,n};$$

$$gx_\ell^{(i,j,k)}, \quad Vx_\ell^{(i,j,k)}, \quad \ell = 1,2, \qquad i = \overline{1,n-1}, \quad j,k = \overline{1,n};$$

$$gy_\ell^{(i,j,k)}, \quad Vy_\ell^{(i,j,k)}, \quad \ell = 1,2, \qquad j = \overline{1,n-1}, \quad i,k = \overline{1,n};$$

$$gz_\ell^{(i,j,k)}, \quad Vz_\ell^{(i,j,k)}, \quad \ell = 1,2, \qquad k = \overline{1,n-1}, \quad i,j = \overline{1,n}.$$

Note that the technique of constructing a separable matrix $\tilde{B}_{11}$ in the case of $\Gamma_1 \neq \emptyset$ when Γ_0 is the union of entire faces of the cube Ω is the same.

These equations are different for different types of cubes. For any cube of the type I (see Fig. 2) we have

$$\left\{(k_1+k_2+k_3)I_c + \frac{1}{2}\left[k_1(T_1+T_3)+k_2(T_2+T_3)+k_3T_3\right]\right\}\mathbf{VI}^{(i,j,k)} - \tag{3.10}$$

$$-(1-\delta_{i1})k_1\begin{bmatrix}0&0\\1&0\\0&1\\0&0\end{bmatrix}\mathbf{Vx}^{(i-1,j,k)} - (1-\delta_{in})k_1\begin{bmatrix}1&0\\0&0\\0&0\\0&1\end{bmatrix}\mathbf{Vx}^{(i,j,k)} -$$

$$-(1-\delta_{j1})k_2\begin{bmatrix}0&0\\0&0\\1&0\\0&1\end{bmatrix}\mathbf{Vy}^{(i,j-1,k)} - (1-\delta_{jn})k_2\begin{bmatrix}1&0\\0&1\\0&0\\0&0\end{bmatrix}\mathbf{Vy}^{(i,j,k)} -$$

$$-(1-\delta_{k1})k_3\begin{bmatrix}0&0\\1&0\\0&0\\0&1\end{bmatrix}\mathbf{Vz}^{(i,j,k-1)} - (1-\delta_{kn})k_3\begin{bmatrix}1&0\\0&0\\0&1\\0&0\end{bmatrix}\mathbf{Vz}^{(i,j,k)} = \left(\tfrac{2}{3h}\right)\mathbf{gi}^{(i,j,k)},$$

$$2k_1\mathbf{Vx}^{(i,j,k)} - k_1\begin{bmatrix}1&0&0&0\\0&0&0&1\end{bmatrix}\mathbf{VI}^{(i,j,k)} - k_1\begin{bmatrix}1&0&0&0\\0&0&0&1\end{bmatrix}\mathbf{VI}^{(i+1,j,k)} = \left(\tfrac{2}{3h}\right)\mathbf{gx}^{(i,j,k)},$$

$$2k_2\mathbf{Vy}^{(i,j,k)} - k_2\begin{bmatrix}1&0&0&0\\0&1&0&0\end{bmatrix}\mathbf{VI}^{(i,j,k)} - k_2\begin{bmatrix}0&0&1&0\\0&0&0&1\end{bmatrix}\mathbf{VI}^{(i,j+1,k)} = \left(\tfrac{2}{3h}\right)\mathbf{gy}^{(i,j,k)},$$

$$2k_3\mathbf{Vz}^{(i,j,k)} - k_3\begin{bmatrix}1&0&0&0\\0&0&1&0\end{bmatrix}\mathbf{VI}^{(i,j,k)} - k_3\begin{bmatrix}0&1&0&0\\0&0&0&1\end{bmatrix}\mathbf{VI}^{(i,j,k+1)} = \left(\tfrac{2}{3h}\right)\mathbf{gz}^{(i,j,k)}. \tag{3.11}$$

For any cube of the type II the entries of the unknowns $\mathbf{Vx}^{(i,j,k)}$ are different from previous ones.

$$\left\{(k_1+k_2+k_3)I_c + \frac{1}{2}\left[k_1(T_1+T_3)+k_2(T_2+T_3)+k_3T_3\right]\right\}\mathbf{VI}^{(i,j,k)} - \tag{3.12}$$

$$-(1-\delta_{i1})k_1\begin{bmatrix}1&0\\0&0\\0&0\\0&1\end{bmatrix}\mathbf{Vx}^{(i-1,j,k)} - (1-\delta_{in})k_1\begin{bmatrix}0&0\\1&0\\0&1\\0&0\end{bmatrix}\mathbf{Vx}^{(i,j,k)} -$$

$$-(1-\delta_{j1})k_2\begin{bmatrix}0&0\\0&0\\1&0\\0&1\end{bmatrix}\mathbf{Vy}^{(i,j-1,k)}-(1-\delta_{jn})k_2\begin{bmatrix}1&0\\0&1\\0&0\\0&0\end{bmatrix}\mathbf{Vy}^{(i,j,k)}-$$

$$-(1-\delta_{k1})k_3\begin{bmatrix}0&0\\1&0\\0&0\\0&1\end{bmatrix}\mathbf{Vz}^{(i,j,k-1)}-(1-\delta_{kn})k_3\begin{bmatrix}1&0\\0&0\\0&1\\0&0\end{bmatrix}\mathbf{Vz}^{(i,j,k)}=\left(\frac{2}{3h}\right)\mathbf{gi}^{(i,j,k)},$$

$$2k_1\mathbf{Vx}^{(i,j,k)}-k_1\begin{bmatrix}0&1&0&0\\0&0&1&0\end{bmatrix}\mathbf{VI}^{(i,j,k)}-k_1\begin{bmatrix}0&1&0&0\\0&0&1&0\end{bmatrix}\mathbf{VI}^{(i+1,j,k)}=\left(\frac{2}{3h}\right)\mathbf{gx}^{(i,j,k)},$$

$$2k_2\mathbf{Vy}^{(i,j,k)}-k_2\begin{bmatrix}1&0&0&0\\0&1&0&0\end{bmatrix}\mathbf{VI}^{(i,j,k)}-k_2\begin{bmatrix}0&0&1&0\\0&0&0&1\end{bmatrix}\mathbf{VI}^{(i,j+1,k)}=\left(\frac{2}{3h}\right)\mathbf{gy}^{(i,j,k)},$$

$$2k_3\mathbf{Vz}^{(i,j,k)}-k_3\begin{bmatrix}1&0&0&0\\0&0&1&0\end{bmatrix}\mathbf{VI}^{(i,j,k)}-k_3\begin{bmatrix}0&1&0&0\\0&0&0&1\end{bmatrix}\mathbf{VI}^{(i,j,k+1)}=\left(\frac{2}{3h}\right)\mathbf{gz}^{(i,j,k)}.$$

(3.13)

Here the function $\delta_{ik}=\begin{cases}1, & i=k\\0, & i\neq k\end{cases}$ is introduced to take into account the Dirichlet boundary conditions, and any vector $\mathbf{vr}^{(i,j,k)}=\begin{bmatrix}vr_1^{(i,j,k)}\\vr_2^{(i,j,k)}\end{bmatrix}\in\mathbb{R}^2$.

Note that

$$\frac{1}{2}(T_1+T_3)=\begin{bmatrix}1&-1&0&0\\-1&1&0&0\\0&0&1&-1\\0&0&-1&1\end{bmatrix},\qquad \frac{1}{2}(T_2+T_3)=\begin{bmatrix}1&0&0&-1\\0&1&-1&0\\0&-1&1&0\\-1&0&0&1\end{bmatrix}.$$

After eliminating the unknowns $\mathbf{Vx}^{(i,j,k)}$, $\mathbf{Vy}^{(i,j,k)}$, $\mathbf{Vz}^{(i,j,k)}$ from equations (3.10), (3.12) we have a block "7-point" computational scheme with the 4×4-blocks for the unknowns $\mathbf{VI}^{(i,j,k)}$:

(3.14)
$$\left(\hat{B}_{11}\mathbf{VI}\right)^{(i,j,k)}\equiv$$
$$\equiv\left\{(k_1+k_2+k_3)I_c+\frac{1}{2}\left[k_1(T_1+T_3)+k_2(T_2+T_3)+k_3T_3\right]\right\}\mathbf{VI}^{(i,j,k)}-$$

$$-(1-\delta_{i1})\frac{k_1}{2}\begin{bmatrix}0&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&0\end{bmatrix}\left(\mathbf{VI}^{(i,j,k)}+\mathbf{VI}^{(i-1,j,k)}\right)$$

$$-(1-\delta_{in})\frac{k_1}{2}\begin{bmatrix}1&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&1\end{bmatrix}\left(\mathbf{VI}^{(i,j,k)}+\mathbf{VI}^{(i+1,j,k)}\right)$$

$$-(1-\delta_{j1})\frac{k_2}{2}\left(\begin{bmatrix} 0&0&0&0\\ 0&0&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{bmatrix}\mathbf{VI}^{(i,j,k)}+\begin{bmatrix} 0&0&0&0\\ 0&0&0&0\\ 1&0&0&0\\ 0&1&0&0 \end{bmatrix}\mathbf{VI}^{(i,j-1,k)}\right)$$

$$-(1-\delta_{jn})\frac{k_2}{2}\left(\begin{bmatrix} 1&0&0&0\\ 0&1&0&0\\ 0&0&0&0\\ 0&0&0&0 \end{bmatrix}\mathbf{VI}^{(i,j,k)}+\begin{bmatrix} 0&0&1&0\\ 0&0&0&1\\ 0&0&0&0\\ 0&0&0&0 \end{bmatrix}\mathbf{VI}^{(i,j+1,k)}\right)$$

$$-(1-\delta_{k1})\frac{k_3}{2}\left(\begin{bmatrix} 0&0&0&0\\ 0&1&0&0\\ 0&0&0&0\\ 0&0&0&1 \end{bmatrix}\mathbf{VI}^{(i,j,k)}+\begin{bmatrix} 0&0&0&0\\ 1&0&0&0\\ 0&0&0&0\\ 0&0&1&0 \end{bmatrix}\mathbf{VI}^{(i,j,k-1)}\right)$$

$$-(1-\delta_{kn})\frac{k_3}{2}\left(\begin{bmatrix} 1&0&0&0\\ 0&0&0&0\\ 0&0&1&0\\ 0&0&0&0 \end{bmatrix}\mathbf{VI}^{(i,j,k)}+\begin{bmatrix} 0&1&0&0\\ 0&0&0&0\\ 0&0&0&1\\ 0&0&0&0 \end{bmatrix}\mathbf{VI}^{(i,j,k+1)}\right),$$

$$i,j,k=1,\dots,n.$$

Along with the Schur matrix $\hat{B}_{11}$ we define the matrix $\tilde{B}_{11}$ in the form

$$\begin{aligned}
&\left(\tilde{B}_{11}\mathbf{VI}\right)^{(i,j,k)}\equiv \\
&\equiv\left\{(k_1+k_2+\tfrac{1}{2}k_3)I_c+\tfrac{1}{2}\left[k_1(T_1+T_3)+k_2(T_2+T_3)+k_3T_3\right]\right\}\mathbf{VI}^{(i,j,k)}-
\end{aligned}\tag{3.15}$$

$$-(1-\delta_{i1})\frac{k_1}{2}\mathbf{VI}^{(i-1,j,k)}-(1-\delta_{in})\frac{k_1}{2}\mathbf{VI}^{(i+1,j,k)}-$$

$$-(1-\delta_{j1})\frac{k_2}{2}\begin{bmatrix} 0&0&1&0\\ 0&0&0&1\\ 1&0&0&0\\ 0&1&0&0 \end{bmatrix}\mathbf{VI}^{(i,j-1,k)}-(1-\delta_{jn})\frac{k_2}{2}\begin{bmatrix} 0&0&1&0\\ 0&0&0&1\\ 1&0&0&0\\ 0&1&0&0 \end{bmatrix}\mathbf{VI}^{(i,j+1,k)}-$$

$$-(1-\delta_{k1})\frac{k_3}{2}\begin{bmatrix} 0&0&0&0\\ 1&0&0&0\\ 0&0&0&0\\ 0&0&1&0 \end{bmatrix}\mathbf{VI}^{(i,j,k-1)}-(1-\delta_{kn})\frac{k_3}{2}\begin{bmatrix} 0&1&0&0\\ 0&0&0&0\\ 0&0&0&1\\ 0&0&0&0 \end{bmatrix}\mathbf{VI}^{(i,j,k+1)},$$

$$i,j,k=1,\dots,n.$$

Let us consider an eigenvalue problem

$$\hat{B}_{11}\mathbf{u} = \lambda \tilde{B}_{11}\mathbf{u}, \qquad \mathbf{u} \in \mathbb{R}^{N_1}. \tag{3.16}$$

PROPOSITION 3.2. *The eigenvalues of the problem* (3.16) *belong to the interval* [1/6, 1].

Proof. Note first that the matrices $\hat{B}_{11}$ and $\tilde{B}_{11}$ may be represented in the form

$$\begin{aligned} \hat{B}_{11} &= k_1\hat{B}^{(1)} + k_2\hat{B}^{(2)} + k_3\hat{B}^{(3)}, \\ \tilde{B}_{11} &= k_1\tilde{B}^{(1)} + k_2\tilde{B}^{(2)} + k_3\hat{B}^{(3)}, \end{aligned} \tag{3.17}$$

where the matrices $\hat{B}^{(i)}$, $i = 1, 2, 3$, and $\tilde{B}^{(j)}$, $j = 1, 2$, do not depend on the coefficients of the problem k_1, k_2, k_3.

Since all components in the right hand sides are nonnegative we can estimate eigenvalues λ of the problem (3.16) by inequalities

$$\min_{i=1,2}\left\{\mu^{(i)}_{min}; 1\right\} \le \lambda \le \max_{i=1,2}\left\{\mu^{(i)}_{max}; 1\right\}, \tag{3.18}$$

where $\mu^{(i)}_*$ are extremal eigenvalues of auxiliary problems

$$\hat{B}^{(i)}\mathbf{u} = \mu^{(i)}\tilde{B}^{(i)}\mathbf{u}, \qquad i = 1, 2.$$

Direct calculations show that $\mu^{(i)} \in [1/6, 1]$. Taking into account the inequalities (3.18) we get the proposition. □

Using Propositions 3.1 and 3.2, and Lemma 1.1 we have the following theorem.

THEOREM 3.3. *Suppose that the matrix coefficient of the equation* (1.1) *is a diagonal tensor* $K(\mathbf{x}) = diag\{k_1, k_2, k_3\}$, *where* k_i, $i = 1, 2, 3$, *are constants over the cube* Ω *such that* $\kappa = min\,\{k_3/k_1, k_3/k_2\} \ge 1$.

Then the matrix $\tilde{B}$ *defined in* (3.8) *with the block* $\tilde{B}_{11}$ *defined in* (3.15) *is spectrally equivalent to the matrix* A. *Moreover,*

$$\mu_*\tilde{B} \le A \le \mu^*\tilde{B},$$

where $\mu_* = \kappa/6(2+\kappa)$ *and* $\mu^* = 1$, *hence*

$$\mathrm{cond}(\tilde{B}^{-1}A) \le \overline{\mu} \equiv \mu^*/\mu_* \le 6(1 + 2/\kappa) \le 18. \tag{3.19}$$

Instead of the matrix B in the form (3.7) we take the matrix $\tilde{B}$ from (3.8) with the block $\tilde{B}_{11}$ in the form of (3.15) as a preconditioner for the matrix A. As we noted earlier, the matrix A_{22} is block-diagonal and can be inverted locally face-by-face.

Again, we note that the condition number does not depend neither on mesh size h nor the value of the coefficients, when $k_3 \ge \max\,\{k_1, k_2\}$. Because the condition number of the matrix $\tilde{B}^{-1}A$ depends on the value of the parameter κ it is very important to choose "z-direction" in a right way. If, for example, we have the problem where the coefficient k_1 is not less than the coefficients k_2 and k_3 we can change variables in such a way that a new "z" variable coincides with old "x" variable. It means that we simply redirect axis of the coordinate system.

From the representations (3.15), (3.17) it is easy to see that matrix $\tilde{B}_{11}$ is separable. It is also separable in the case of $\Gamma_1 \neq 0$ when Γ_0 is a union of entire faces of the cube Ω. To solve the system of linear equations with matrix $\tilde{B}_{11}$ from (3.15) we use the method of separation of variables which is described in the next subsection.

3.1. Implementation of the method of separation of variables. Since the matrix $\tilde{B}_{11}$ is separable we can use method of separation of variables to solve the problem

$$\tilde{B}_{11}\mathbf{w} = \mathbf{g}, \qquad \mathbf{w}, \mathbf{g} \in \mathbb{R}^{N_1}. \tag{3.20}$$

The matrix $\tilde{B}_{11}$ can be represented in the form

$$\tilde{B}_{11} = k_1 B_x + k_2 B_y + k_3 B_z, \tag{3.21}$$

$$B_x = I_z \otimes I_y \otimes (I_x \otimes D_1 + K_x \otimes I_0),$$

$$B_y = I_z \otimes (I_y \otimes I_x \otimes D_2 + K_y \otimes I_x \otimes D_0),$$

$$B_z = I_z \otimes I_y \otimes I_x \otimes D_3 + K_{lz} \otimes I_y \otimes I_x \otimes D_{3l} + K_{uz} \otimes I_y \otimes I_x \otimes D_{3u},$$

where I_0, D_* are 4×4-matrices, I_x, I_y, I_z, K_* are $n \times n$-matrices,

$$D_1 = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 2 & 0 & -1 & -1 \\ 0 & 2 & -1 & -1 \\ -1 & -1 & 2 & 0 \\ -1 & -1 & 0 & 2 \end{bmatrix}, \quad D_0 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

$$D_3 = \frac{1}{2}\begin{bmatrix} 2 & -1 & 1 & -1 \\ -1 & 2 & -1 & 1 \\ 1 & -1 & 2 & -1 \\ -1 & 1 & -1 & 2 \end{bmatrix}, \quad D_{3l} = \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad D_{3u} = \frac{1}{2}\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$K_x = K_y = \frac{1}{2}\begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix}, \qquad K_{lz} = K_{uz}^T = \begin{bmatrix} 0 & & & \\ -1 & 0 & & \\ & \ddots & \ddots & \\ & & -1 & 0 \end{bmatrix}.$$

Let us represent matrices $K_x, K_y, D_1, D_2, D_0, D_3$ in the form

$$K_\alpha = Q_\alpha \Lambda_\alpha Q_\alpha^T, \qquad \alpha = x, y$$

(3.22)

$$K_\beta = Q_0 \Lambda_\beta Q_0^T, \qquad \beta = 0, 1, 2, 3,$$

where

$$Q_x = Q_y = \{q_{ij}\}_{i,j=1}^n, \qquad q_{ij} = \sqrt{\frac{2}{n+1}}\ \sin\left(\frac{\pi}{n+1} \cdot i \cdot j\right),$$

$$Q_0 = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix},$$

and Λ_x, Λ_y are $(n \times n)$-diagonal matrices, and $\Lambda_0, \Lambda_1, \Lambda_2, \Lambda_3$ are 4×4-diagonal matrices.

Define a matrix Q as

$$Q = I_z \otimes Q_y \otimes Q_x \otimes Q_0. \tag{3.23}$$

Remark. Q is an $(4n^3 \times 4n^3)$-orthogonal matrix.

Then the matrix $\tilde{B}_{11}$ can be represented in the form

$$\tilde{B}_{11} = Q \Lambda Q^T, \tag{3.24}$$

where

$$\Lambda = Q^T \tilde{B}_{11} Q =$$

$$k_1 I_z \otimes I_y \otimes (I_x \otimes \Lambda_1 + \Lambda_x \otimes I_0) + k_2 I_z \otimes (I_y \otimes I_x \otimes \Lambda_2 + \Lambda_y \otimes I_x \otimes \Lambda_0) +$$

$$k_3 \Big(I_z \otimes I_y \otimes I_x \otimes \Lambda_3 + K_{lz} \otimes I_y \otimes I_x \otimes (Q_0^T D_{3l} Q_0) + K_{uz} \otimes I_y \otimes I_x \otimes (Q_0^T D_{3u} Q_0)\Big),$$

(3.25)

$$Q_0^T D_{3l} Q_0 = \frac{1}{4}\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \end{bmatrix}, \qquad Q_0^T D_{3u} Q_0 = \frac{1}{4}\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}.$$

Then we can use the following method to solve the system (3.20):

(3.26)

$$(1) \quad \tilde{\mathbf{g}} = Q^T \mathbf{g},$$

$$(2) \quad \Lambda \tilde{\mathbf{w}} = \tilde{\mathbf{g}},$$

$$(3) \quad \mathbf{v} = Q \tilde{\mathbf{w}}.$$

We note that due to the form of the matrix Λ (3.25) the solution procedure of the stage (2) is equivalent to solving $2n^2$ independent tridiagonal linear systems of the order $2n \times 2n$.

If we use fast fourier transform to implement the procedure (3.26) then the number of operations to solve the system (3.20) is estimated by $cN_1 \ln(N_1)$ or $cN \ln(0.4N)$, where the constant c does not depend on the number of unknowns N and on the coefficients k_1, k_2, and k_3.

3.2. Preconditioned conjugate gradient method. The underlying method to solve (2.5) is a preconditioned iterative method. The choice of a particular iterative method within a certain class is not essential, but for the purpose of this exposition we may think of the well-known preconditioned conjugate gradient method [13], [20] which is often used in practice.

PROPOSITION 3.4. *The number of operations for solving the system* (2.5) *by preconditioned conjugate gradient method with the preconditioner* $\tilde{B}$ *defined in* (3.8) *and with accuracy* ε *in the sense*

$$\|\mathbf{u}^{k_\varepsilon+1} - \mathbf{u}^*\|_A \le \varepsilon \|\mathbf{u}^0 - \mathbf{u}^*\|_A, \qquad 0 < \varepsilon \ll 1, \tag{3.27}$$

is estimated by $cN \ln(0.4N) \ln\left(\frac{2}{\varepsilon}\right)$, *where* $\mathbf{u}^* = A^{-1}\mathbf{f}$, $\mathbf{u}^0 \in \mathbb{R}^N$ *is any initial vector, and the constant* $c > 0$ *does not depend on* N *and the coefficients of the matrix function* $K(\mathbf{x})$.

4. Preconditioners for a general case. In this section we consider the case where the coefficient matrix $K(\mathbf{x})$ is a full tensor and the domain Ω satisfies the assumption that there is an orientation-preserving smooth mapping $\mathcal{L}$ from the unit cube $\hat{\Omega}$ onto Ω and there are positive constants d and C (see [9]) such that

$$d^{-1}\|J(x)\| \le C, \qquad \forall x \in \hat{\Omega}, \tag{4.1}$$

and

$$d\,\|J^{-1}(x)\| \le C, \qquad \forall x \in \Omega, \tag{4.2}$$

where $J(x)$ is the Jacobian matrix of $\mathcal{L}$ at x and $\|\cdot\|$ denotes a matrix norm. Note that the domain Ω is roughly of size d.

Next, we consider the definition of the nonconforming finite element space. Let $\mathcal{C}_{\hat{h}}$ and $\mathcal{T}_{\hat{h}}$ be the partitions of $\hat{\Omega}$ into cubes and tetrahedra, respectively, associated with the mesh size $\hat{h} = 1/n$, as defined in Section 2, and let $V_{\hat{h}}$ be the P_1 nonconforming space associated with $\mathcal{T}_{\hat{h}}$, as given in (2.1). Set $h = d\hat{h}$ and define

$$V_h(\Omega) = \left\{\varphi = \psi \circ \mathcal{L}^{-1} \,:\, \psi \in V_{\hat{h}}(\hat{\Omega})\right\}.$$

Also, we introduce the mapping $\mathcal{I} : V_h(\Omega) \to V_{\hat{h}}(\hat{\Omega})$ defined by $\mathcal{I}v = v \circ \mathcal{L}$.

Now we define the stiffness matrix A on the domain Ω by

$$(A\mathbf{u}, \mathbf{v})_N = a_h(u, v), \qquad \forall u, v \in V_h, \tag{4.3}$$

where

$$
\begin{aligned}
a_h(u,v) &= \sum_{T\in\mathcal{T}_h} \int_T K(\mathbf{x})\,\nabla u\cdot\nabla v\ d\mathbf{x} =, \\
&= \sum_{\hat T\in\mathcal{T}_{\hat h}} \int_{\hat T} \frac{1}{|\det(\mathcal{J})|}\mathcal{J}^T K(\mathbf{x})\mathcal{J}\ \nabla\mathcal{I}u\cdot\nabla\mathcal{I}v\ d\mathbf{x},
\end{aligned}
\tag{4.4}
$$

where $|\det(\mathcal{J})|$ is the Jacobian of the mapping.

Note that owing to (4.4) we can consider the bilinear form (4.3) as a form generated by some elliptic positive definite operator with piece-wise smooth 3×3 symmetric matrix-valued function $K(\mathbf{x})$ on the cube $\hat\Omega$. This function satisfies the uniform positive definiteness condition. For this reason below without loss of generality we suppose that $\Omega\equiv\hat\Omega$ is parallelepiped with partitioning into cubes $\mathcal{C}_h$ and into tetrahedra $\mathcal{T}_h$ as described in Section 2.

For each cube $C\in\mathcal{C}_{\hat h}$, we introduce the diagonal matrix $\mathcal{K}_C=\mathrm{diag}\{k_{1,C},k_{2,C},k_{3,C}\}$ with some as yet unspecified constants $k_{i,C}$, $i=1,2,3$. Then we define on the reference parallelepiped $\hat\Omega$ a bilinear form

$$
b_h(u,v)=\sum_{C\in\mathcal{C}_h}\delta_C\left(\sum_{T\in C}\int_T \mathcal{K}_C\nabla u\cdot\nabla v\ d\mathbf{x}\right),\qquad \forall u,v\in V_{\hat h},
\tag{4.5}
$$

where the constants δ_C are scaling factors. One reasonable choice is to take $\delta_C=(\lambda_{1,C}+\lambda_{0,C})/2$, where $\lambda_{1,C}$ and $\lambda_{0,C}$ are the largest and smallest eigenvalues of the eigenvalue problem

$$
\hat K(\mathbf{x}_0)\mathbf{v}=\lambda_C\mathcal{K}\mathbf{v},\qquad \mathbf{v}\in\mathbb{R}^3,
\tag{4.6}
$$

where $\hat K(\mathbf{x})=\frac{1}{|\det(\mathcal{J})|}\mathcal{J}^T K(\mathbf{x})\mathcal{J}$ and $\mathbf{x}_0\in\mathcal{L}(C)\subset\Omega$ is some point.

Note that the assumptions (4.1) and (4.2) imply that there are two constants c_0 and c_1 independent of d and $\hat h$ such that

$$
c_0 a_h(u,u)\le d\cdot b_h(\mathcal{I}u,\mathcal{I}u)\le c_1 a_h(u,u),\qquad \forall u\in V_h.
\tag{4.7}
$$

We choose the matrices $\mathcal{K}_C$ in the form $\mathcal{K}_C=\mathrm{diag}\{\hat K(\mathbf{x}_0)\}$, $\forall C\in\mathcal{C}_h$, i.e., the matrix $\mathcal{K}_C$ is the diagonal part of $\hat K(x_0)$ at some point $x_0\in\mathcal{L}(C)$. In this case the constants c_0 and c_1 in (4.7) depend only on the local variation of the coefficients $\left\{\left(\hat K\right)_{kl}\right\}$. Hence the problem of defining a preconditioner for $a_h(\cdot,\cdot)$ is reduced to the problem of finding a preconditioner for $d\cdot b_h(\cdot,\cdot)$, which has a diagonal coefficient tensor and is defined on the unit cube $\hat\Omega$. Namely, all the analysis in Section 3 can be carried out here.

5. Results of the numerical experiments. In this section the method of preconditioning being presented is tested on the model problem

$$
-\sum_{i=1}^{3}\frac{\partial}{\partial x_i}\left(k_i\frac{\partial u}{\partial x_i}\right)=f\quad\text{in }\Omega,\qquad u=0\quad\text{on }\partial\Omega,
$$

where the domain Ω is the unit cube: $\Omega=[0,1]^3$. The domain is divided into n^3 cubes (n in each direction) and each cube is partitioned into 5 tetrahedra. The dimension of the original algebraic system is $N = 10n^3 - 6n^2$. The right hand side is generated randomly, and the accuracy parameter is taken as $\varepsilon = 10^{-6}$. The condition number of the matrix $\tilde{B}^{-1}A$ is calculated by the relation between the conjugate gradient and Lanczos algorithm [13]. The coefficients k_i, $i = 1, 2, 3$, are constants on each cube. The results are summarized in Table 1, where n_{iter} and $Cond$ denote the iteration number and condition number, respectively. All experiments are carried out on Sun Workstation.

TABLE 1.

			$16 \times 16 \times 16$ $N = 39424$		$20 \times 20 \times 20$ $N = 77600$		$30 \times 30 \times 30$ $N = 264600$	
k_1	k_2	k_3	n_{iter}	$Cond$	n_{iter}	$Cond$	n_{iter}	$Cond$
1	1	1	14	4.87	14	4.93	14	5.03
1	1	10	12	3.72	12	3.94	12	4.28
1	1	100	9	2.28	10	2.55	10	3.00
1	1	1000	8	1.55	8	1.58	8	1.73
1	1	10000	8	1.48	8	1.49	8	1.51
1	1	0.1	31	19.4	31	19.6	31	19.8
1	1	0.01	62	133.	71	149.	82	168.
10	1	1	24	12.0	25	12.1	25	12.1
1	10	1	24	12.1	24	12.1	24	12.0
100	1	1	58	99.3	63	100.	62	100.
1	100	1	62	100.	60	100.	60	99.5
1	10	10	14	4.72	14	4.81	14	4.94
1	10	100	12	3.62	12	3.85	12	4.25
1	10	1000	9	2.14	10	2.42	10	2.92
1	100	10000	9	2.20	10	2.42	10	2.92

From Table 1 we see that the condition number depends on the maximal ratio $\kappa = \max\limits_{C \in \mathcal{C}_h} \left\{ \frac{k_1}{k_3}, \ \frac{k_2}{k_3} \right\}$. The numerical results are in full agreement with the theoretical estimates. One can see that the studied preconditioner is optimal if $\kappa \leq 1$. In the case of $\kappa < 1$ the method has a better convergence than in the case of the Poisson equation (i.e., $k_1 = k_2 = k_3 = 1$). If $\kappa > 1$, the preconditioner looses its optimal order and the corresponding relative condition numbers increased strongly with κ. It is rather predictable result since we defined local preconditioning matrices B^C in (3.3) on each cube taking some "additional positiveness" from the direction with the dominated anisotropy ($z-$direction) to other directions. Experiments show that this procedure has "a good behavior" if coefficient in $z-$direction (k_3) is greater than the coefficients k_1 and k_2. And method looses its effectiveness if we choose wrong direction, i.e. coefficient k_3 is small in comparison with the coefficients k_1 and k_2.

Remind that actually in the method described above we need only an assumption that the coefficient k_* in some direction is not less then the coefficients in the other directions. So, if, for example, we have the problem where the coefficient k_1 is not less than the coefficients k_2 and k_3 we can change variables in such a way that a new z variable coincides with old x variable. It means that we can simply redirect axis of the coordinate system. The results will be the same.

Acknowledgment. The authors are grateful to Professors Richard Ewing and Raytcho Lazarov for fruitful discussions and support given through the Institute for Scientific Computation, Texas A&M University, College Station, Texas.

REFERENCES

[1] D. Arnold and F. Brezzi, *Mixed and nonconforming finite element methods: Implementation, postprocessing and error estimates*, RAIRO Model. Math. Anal. Numer., 19 (1985), pp. 7–32.

[2] O. Axelsson, *On multigrid methods of the two-level type*, Multigrid Methods, 1981, W. Hackbush and U. Trottenberg, eds., no. 960 in LNM, Köln-Porz, 1981, Springer, 1982, pp. 352–367.

[3] D. Braess and R. Verfürth, *Multigrid methods for nonconforming finite element methods*, SIAM J. Numer. Anal., 27 (1990), pp. 979–986.

[4] J. Bramble, J. Pasciak, and A. Schatz, *The construction of preconditioners for elliptic problems by substructuring, I*, Math. Comp., 47 (1986), 103–134.

[5] J. Bramble, J. Pasciak, and J. Xu, *The analysis of multigrid algorithms with non-nested spaces or non-inherited quadratic forms*, Math. Comp., 56 (1991), pp. 1–34.

[6] S. Brenner, *An optimal-order multigrid method for P1 nonconforming finite elements*, Math. Comp., 52 (1989), pp. 1–16.

[7] Z. Chen, R. Ewing, Y. Kuznetsov, R. Lazarov, and S. Maliassov, *Multilevel preconditioners for mixed methods for second order elliptic problems*, IMA Preprint #1269 (1994). (Submitted to SIAM Numer. Anal.).

[8] Z. Chen and D. Kwak, *The analysis of multigrid algorithms for nonconforming and mixed methods for second order elliptic problems*, IMA Preprint #1277 (1994).

[9] E. D'yakonov, *On triangulations in the finite element method and efficient iterative methods*, in Topics in Numerical Analysis. III, I. Miller, ed., Academic Press, London, 1977, pp. 103–124.

[10] R. Ewing, Y. Kuznetsov, R. Lazarov, S. Maliassov, *Substructuring preconditioning for finite element approximations of second order elliptic problems. I. Nonconforming linear elements for the Poisson equation in parallelepiped*, IMA Preprint #1280 (1994).

[11] ———, *Preconditioning of nonconforming finite element approximations of second order elliptic problems*, in The Third Int. Conf. on Advances in Numerical Methods and Applications, I. Dimov, B. Sendov, and P. Vassilevski, eds., Bulgaria, 1994, World Scientific, pp. 101–110.

[12] R. Ewing, R. Lazarov, and P. Vassilevski, *Local refinement techniques for elliptic problems on cell-centered grids, I: Error Analysis*, Math. Comput., 56 (1991), 437–462.

[13] G. Golub and C. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, 1989.

[14] G. Hardy, J. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, 1952.

[15] Y. Kuznetsov, *Conjugate gradient method, its generalizations and applications*, in Computational System and Processes, Nauka, Moscow, Russia, 1983, pp. 264–300 (in Russian).

[16] ———, *Multigrid domain decomposition methods for elliptic problems*, in Proc. of 8th Int. Conf. on Comput. Methods in Applied Science and Engineering, Paris, 1987. In: Comput. Meth. Appl. Mech. and Eng., 75 (1989), pp. 185–193.

[17] ———, *Multigrid domain decomposition methods*, in Proc. of 3rd International Symposium on Domain Decomposition Methods, T. Chan. R. Glowinski, J. Periaux, and O. Widlund, eds., Philadelphia, 1990, 1989, SIAM, pp. 290–313.

[18] ———, *Multilevel substructuring preconditioners*. Invited presentation at the 7th Int. Symp. on Domain Decomp. Methods for PDEs., October 1993. Penn State University.

[19] S. Maliassov, *Substructuring preconditioning for finite element approximations of second order elliptic problems. II. Mixed method for an elliptic operator with scalar tensor*, Institute for Scientific Computation, Technical Report ISC–94–19–MATH, Texas A&M University, 1994.

[20] G. Marchuk and Y. Kuznetsov, *On optimal iteration processes*, Soviet Math. Dokl., 9 (1968), pp. 1041–1045.