MATHEMATICS OF COMPUTATION
Volume 71, Number 240, Pages 1455–1472
S 0025-5718(01)01394-1
Article electronically published on December 5, 2001

CONJUGATE GRADIENT METHOD FOR DUAL-DUAL MIXED FORMULATIONS

GABRIEL N. GATICA AND NORBERT HEUER

ABSTRACT. We deal with the iterative solution of linear systems arising from so-called dual-dual mixed finite element formulations. The linear systems are of a two-fold saddle point structure; they are indefinite and ill-conditioned. We define a special inner product that makes matrices of the two-fold saddle point structure, after a specific transformation, symmetric and positive definite. Therefore, the conjugate gradient method with this special inner product can be used as iterative solver. For a model problem, we propose a preconditioner which leads to a bounded number of CG-iterations. Numerical experiments for our model problem confirming the theoretical results are also reported.

1. INTRODUCTION

As is well known, mixed finite element methods often lead to symmetric but indefinite linear systems. For standard variational formulations they represent a saddle point problem. Here we deal with linear systems of a two-fold saddle point structure. These systems arise, e.g., when using a combined dual-mixed finite element method with a Dirichlet-to-Neumann mapping to solve exterior transmission problems [15], or in elastostatics when coupling the primal mixed finite element method and the boundary element method [3]. In this paper we consider, as a model problem, the dual-dual mixed finite element method, which introduces the gradient as third explicit unknown, to solve linear second order elliptic equations in divergence form [13]. It is important to remark that this kind of formulation, with 3 (instead of 2) independent unknowns, is suitable for treating individual boundary conditions, for the case where the tensor κ defining the flux is large and does not need to be inverted, and more importantly when the constitutive equation describing the flux is nonlinear and cannot be inverted (see, e.g. [11], [15]). The technique of introducing further unknowns for the mixed formulation has also been applied in [9], [1], [7], and [8], where it was named expanded mixed finite element method. Further, this approach has been established in elasticity as the Hu-Washizu principle (see, e.g. [4]). However, the idea of writing the resulting variational formulation

Received by the editor September 22, 1999 and, in revised form, October 3, 2000.

2000 *Mathematics Subject Classification.* Primary 65N30, 65N22, 65F10.

Key words and phrases. Mixed finite elements, dual-dual variational formulation, conjugate gradient method.

This research was partially supported by CONICYT-Chile through Program A on Numerical Analysis of the FONDAP in Applied Mathematics and Fondecyt project No. 1980122, and by the Dirección de Investigación of the Universidad de Concepción through the Advanced Research Groups Program.

as a two-fold saddle point operator equation, so that an extension of the classical Babuška-Brezzi theory can be easily applied (see [10], [14]), has been utilized only by the present authors and some co-workers.

Now, there are several iterative methods for the solution of linear systems with single saddle point structure. For instance, preconditioned Krylov subspace methods can be efficiently applied to symmetric indefinite systems, see, e.g., [18, 19, 2, 21, 22, 17]. We also mention the Uzawa algorithm which is an iterative method for the corresponding positive definite Schur complement system. Finally, there is the method of Bramble and Pasciak [5] who solve an equivalent transformed system by using the conjugate gradient (CG) method with a special inner product.

Instead of investigating Krylov subspace methods for dual-dual type linear systems, for which we refer to [12], we here follow the ideas of Bramble and Pasciak. We define a specific transformation for linear systems of dual-dual type and present an inner product which makes the transformed system matrix symmetric and positive definite. Therefore, the CG method can be applied as iterative solver. Here, we have to deal with ill-conditioned transformed matrices and therefore, preconditioning strategies are in order. For our model problem we propose a preconditioner that leads to a bounded number of CG-iterations, i.e., that is independent of the underlying mesh size.

An outline of this paper is as follows. In Section 2 we consider in an abstract way linear systems which are of dual-dual structure. We introduce the specific transformation and define a bilinear form which is an inner product under certain assumptions. We show that the transformed system matrix is symmetric and positive definite with respect to the special inner product and we give a precise estimate for its extreme eigenvalues (Theorem 2.3). In Section 3 we apply this procedure to solve linear systems arising from our model problem. First the model problem and its discretization are briefly introduced. Then, in subsection 3.3, we consider in detail the condition number of the transformed linear system with and without additional preconditioner. We prove that the number of iterations of the preconditioned CG method, which are necessary to solve the system up to a given accuracy, is independent of the mesh size h (Theorem 3.3). Finally, in subsection 3.4, we present some numerical results which underline our theory.

2. Transformation of dual-dual type linear systems

The structure of the linear system we have to deal with looks like

$$\mathcal{A}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} := \begin{pmatrix} A_1 & B_1^* & 0 \\ B_1 & 0 & B^* \\ 0 & B & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix}. \tag{2.1}$$

Here, A_1 is symmetric positive definite and the matrices B_1 and B have full rank. The solution $(x_1, x_2, x_3)^T$ is sought within a vector space $S^1 \times S^2 \times S^3$ with

$$L := \dim S^1 \geq M := \dim S^2 \geq N := \dim S^3.$$

Certainly, by exchanging the second and third rows and columns of (2.1), we obtain a linear system with the usual saddle point structure, but with an upper diagonal block containing a null sub-block, which makes it only semi-definite. In principle the well-known Babuška-Brezzi theory could be applied in this case. However, the properties of the corresponding continuous operators yielding (2.1), and in particular the characterization of the null space involved, makes the discrete analysis too cumbersome. This fact can be verified in [7] where, instead of that theory, an alternative but too specific analysis was applied. These difficulties can be avoided by treating the system as in (2.1) and using the abstract theory from [10] and [14]. Indeed, the main advantage of using two-fold saddle point formulations, as compared with the approaches in [1] and [7], lies on the simplicity and generality of the analysis for both the continuous and discrete systems. We will go back to this point at the end of Section 3.1 when we consider a model boundary value problem.

The matrix $\mathcal{A}$ is symmetric and indefinite. It has $L+N$ positive and M negative eigenvalues; see [12]. We transform the linear system into an equivalent one which is symmetric and positive definite with respect to a certain bilinear form (Theorem 2.3). Therefore, due to that result, the conjugate gradient method can be used to solve the transformed linear system.

First, let us introduce step by step the transformation of the linear system. Then we define the special inner product and prove some technical results before we state the main result of this section (Theorem 2.3).

Let us denote a preconditioner for A_1 by A_0. Multiplying the first row of (2.1) by A_0^{-1}, subtracting the second row from B_1 times the new first row, and multiplying the last row by $-I$ we obtain

$$\begin{pmatrix} A_0^{-1}A_1 & A_0^{-1}B_1^* & 0 \\ B_1A_0^{-1}(A_1-A_0) & B_1A_0^{-1}B_1^* & -B^* \\ 0 & -B & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} A_0^{-1}f_1 \\ B_1A_0^{-1}f_1 - f_2 \\ -f_3 \end{pmatrix}.$$

This system can be written in the form

$$\begin{pmatrix} \mathcal{M}_1 & \mathcal{B}^* \\ \mathcal{B} & 0 \end{pmatrix} \begin{pmatrix} X \\ x_3 \end{pmatrix} = \begin{pmatrix} F \\ -f_3 \end{pmatrix} \tag{2.2}$$

where

$$\mathcal{M}_1 := \begin{pmatrix} A_0^{-1}A_1 & A_0^{-1}B_1^* \\ B_1A_0^{-1}(A_1-A_0) & B_1A_0^{-1}B_1^* \end{pmatrix}, \quad \mathcal{B} := \begin{pmatrix} 0 & -B \end{pmatrix} \tag{2.3}$$

and

$$X := \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad F := \begin{pmatrix} A_0^{-1}f_1 \\ B_1A_0^{-1}f_1 - f_2 \end{pmatrix}.$$

Let us repeat the analogous row manipulations to the system (2.2). In order to do so we need to take a preconditioner $\mathcal{M}_0$ for $\mathcal{M}_1$. As given by Lemma 2.1 below, $\mathcal{M}_1$ is spectrally equivalent (with respect to a special inner product) to the matrix

$$\tilde{\mathcal{M}}_1 := \begin{pmatrix} I & 0 \\ 0 & B_1A_1^{-1}B_1^* \end{pmatrix}. \tag{2.4}$$

It therefore suffices to consider a preconditioner for $\mathcal{M}_1$ which is of the special form

$$\mathcal{M}_0 := \begin{pmatrix} \rho I & 0 \\ 0 & M_0 \end{pmatrix}$$

where ρ is a positive number and M_0 is an appropriate symmetric positive definite matrix for the vector space S^2 with inner product $(\cdot,\cdot)$.

Now we transform the system (2.2). We multiply the first row by $\mathcal{M}_0^{-1}$ and subtract the second row from $\mathcal{B}$ times the new first row. Then we obtain the system

(2.5)

$$\mathcal{M}\begin{pmatrix} X \\ x_3 \end{pmatrix} := \begin{pmatrix} \mathcal{M}_0^{-1}\mathcal{M}_1 & \mathcal{M}_0^{-1}\mathcal{B}^* \\ \mathcal{B}\mathcal{M}_0^{-1}(\mathcal{M}_1 - \mathcal{M}_0) & \mathcal{B}\mathcal{M}_0^{-1}\mathcal{B}^* \end{pmatrix}\begin{pmatrix} X \\ x_3 \end{pmatrix} = \begin{pmatrix} \mathcal{M}_0^{-1}F \\ \mathcal{B}\mathcal{M}_0^{-1}F + f_3 \end{pmatrix},$$

in short form

$$\mathcal{M}x = \mathcal{F},$$

where, in detail,

$$\text{(2.6)}\quad \mathcal{M} = \begin{pmatrix} \frac{1}{\rho}A_0^{-1}A_1 & \frac{1}{\rho}A_0^{-1}B_1^* & 0 \\ M_0^{-1}B_1A_0^{-1}(A_1 - A_0) & M_0^{-1}B_1A_0^{-1}B_1^* & -M_0^{-1}B^* \\ -BM_0^{-1}B_1A_0^{-1}(A_1 - A_0) & B - BM_0^{-1}B_1A_0^{-1}B_1^* & BM_0^{-1}B^* \end{pmatrix}$$

and

$$\mathcal{F} = \begin{pmatrix} \frac{1}{\rho}A_0^{-1}f_1 \\ M_0^{-1}(B_1A_0^{-1}f_1 - f_2) \\ BM_0^{-1}(f_2 - B_1A_0^{-1}f_1) + f_3 \end{pmatrix}.$$

Now we introduce a bilinear form on $S^1 \times S^2 \times S^3$ which becomes an inner product if some conditions are satisfied. This inner product then makes the system matrix $\mathcal{M}$ symmetric positive definite.

First, we define a bilinear form on $S^1 \times S^2$ by

$$\text{(2.7)}\qquad \left[\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\right]_1 := (A_1u_1, v_1) - (A_0u_1, v_1) + (u_2, v_2)$$

where $(\cdot, \cdot)$ is the inner product on S^1 or S^2, as appropriate. Under certain assumptions on A_0, which will be given below, $[\cdot, \cdot]_1$ is an inner product on $S^1 \times S^2$ and, moreover, $\mathcal{M}_1$ is symmetric positive definite with respect to $[\cdot, \cdot]_1$; see Lemma 2.1.

Using this bilinear form on $S^1 \times S^2$ we define the following bilinear form on $S^1 \times S^2 \times S^3$:

$$\text{(2.8)}\qquad \left[\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}\right] := \left[\mathcal{M}_1\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\right]_1 - \left[\mathcal{M}_0\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\right]_1 + (u_3, v_3).$$

Here, $(\cdot, \cdot)$ is also used as the inner product on S^3, like on S^1 and S^2 before.

Let us make the following assumptions.

(A1) There exist positive constants α_0, α_1 such that

$$\alpha_0(A_1u_1, u_1) \leq (A_0u_1, u_1) \leq \alpha_1(A_1u_1, u_1)$$

for all $u_1 \in S^1$.

(A2) $\alpha_1 < 1$.

We then conclude that

$$0 < (1 - \alpha_1)(A_1u_1, u_1) \leq ((A_1 - A_0)u_1, u_1) \leq \alpha(A_1u_1, u_1)$$

holds for all $u_1 \in S^1 \setminus \{0\}$ where $\alpha := 1 - \alpha_0$.

As shown by Bramble and Pasciak the following results hold.

Lemma 2.1 ([5]). *Let the assumptions* (A1) *and* (A2) *hold. Then*

(i) *the bilinear form* $[\cdot,\cdot]_1$ *is an inner product on* $S^1 \times S^2$.

(ii) $\mathcal{M}_1$ *is symmetric positive definite with respect to* $[\cdot,\cdot]_1$.

(iii) $\mathcal{M}_1$ *is spectrally equivalent to* $\tilde{\mathcal{M}}_1$; *see* (2.4). *More precisely, there holds*

$$\lambda_0 \left[\tilde{\mathcal{M}}_1 \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right]_1 \le \left[\mathcal{M}_1 \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right]_1 \le \lambda_1 \left[\tilde{\mathcal{M}}_1 \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right]_1$$

for any $u_1 \in S^1$ *and* $u_2 \in S^2$ *where*

$$\lambda_0 = \left(1 + \frac{\alpha}{2} + \sqrt{\alpha + \frac{\alpha^2}{4}}\right)^{-1} \quad \textit{and} \quad \lambda_1 = \frac{1+\sqrt{\alpha}}{1-\alpha}.$$

The assumptions (A1), (A2) are concerned with the preconditioner A_0 for A_1. They are used to handle the matrix $\mathcal{M}_1$ which is a transformation of the indefinite block $A = \begin{pmatrix} A_1 & B_1^* \\ B_1 & 0 \end{pmatrix}$ of $\mathcal{A}$. We repeat the procedure at the next level, i.e. we deal with the indefinite system (2.2). The transformed system matrix is $\mathcal{M}$, and we consider the preconditioner $\mathcal{M}_0$ for $\mathcal{M}_1$. However, since we already know that $\mathcal{M}_1$ is spectrally equivalent to $\tilde{\mathcal{M}}_1$, we do not specify the spectral equivalence of $\mathcal{M}_0$ and $\mathcal{M}_1$. Instead, it is more convenient to deal with the matrix $\tilde{\mathcal{M}}_1$ which is much simpler than $\mathcal{M}_1$.

The assumptions are the following.

(A3) There exist positive constants β_0, β_1 such that

$$\beta_0[\tilde{\mathcal{M}}_1 U, U]_1 \le [\mathcal{M}_0 U, U]_1 \le \beta_1[\tilde{\mathcal{M}}_1 U, U]_1$$

for all $U \in S^1 \times S^2$.

(A4) $\beta_1/\lambda_0 < 1$, where λ_0 is given in Lemma 2.1.

Note that Assumption (A3) makes sense because $\mathcal{M}_0 = \begin{pmatrix} \rho I & 0 \\ 0 & M_0 \end{pmatrix}$ is symmetric positive definite with respect to $[\cdot,\cdot]_1$ since M_0 is symmetric positive definite.

Then we obtain the following lemma.

Lemma 2.2. *Let the assumptions* (A1)–(A4) *be satisfied. Then* $\mathcal{M}_1 - \mathcal{M}_0$ *is positive definite with respect to* $[\cdot,\cdot]_1$. *More precisely there holds*

$$(1 - \frac{\beta_1}{\lambda_0})[\mathcal{M}_1 U, U]_1 \le [(\mathcal{M}_1 - \mathcal{M}_0)U, U]_1 \le (1 - \frac{\beta_0}{\lambda_1})[\mathcal{M}_1 U, U]_1$$

for all $U \in S^1 \setminus \{0\}$ *where* λ_0 *and* λ_1 *are given in Lemma* 2.1.

Proof. The proof is a straightforward combination of Assumption (A3) and Lemma 2.1. Assumption (A4) ensures the positive definiteness of $\mathcal{M}_1 - \mathcal{M}_0$ with respect to $[\cdot,\cdot]_1$. □

On the first level we dealt with the matrix $\mathcal{M}_1$ which is equivalent to the block-diagonal matrix $\tilde{\mathcal{M}}_1$ defined in (2.4); see Lemma 2.1. Now, for repeating the procedure at the second level, we consider the matrix $\mathcal{M}$ and the block-diagonal matrix $\tilde{\mathcal{M}}$ given by

$$\tilde{\mathcal{M}} := \begin{pmatrix} I & 0 \\ 0 & \mathcal{B}\mathcal{M}_1^{-1}\mathcal{B}^* \end{pmatrix}.$$

One finds that the symmetric part of $\tilde{\mathcal{M}}$ is positive definite with respect to $[\cdot,\cdot]$. The following theorem states that $\mathcal{M}$ and $\tilde{\mathcal{M}}$ are spectrally equivalent. Due to this result the conjugate gradient method can be used for the solution of (2.5) which is equivalent to solving (2.1).

Theorem 2.3. *Let the assumptions* (A1)–(A4) *be satisfied. Then* $[\cdot,\cdot]$ *is an inner product on* $S^1 \times S^2 \times S^3$. *Moreover,* $\mathcal{M}$ *is symmetric positive definite with respect to* $[\cdot,\cdot]$ *and there hold the following inequalities:*

$$\Lambda_0 \left[\tilde{\mathcal{M}}\begin{pmatrix} U \\ u_3 \end{pmatrix}, \begin{pmatrix} U \\ u_3 \end{pmatrix}\right] \le \left[\mathcal{M}\begin{pmatrix} U \\ u_3 \end{pmatrix}, \begin{pmatrix} U \\ u_3 \end{pmatrix}\right] \le \Lambda_1 \left[\tilde{\mathcal{M}}\begin{pmatrix} U \\ u_3 \end{pmatrix}, \begin{pmatrix} U \\ u_3 \end{pmatrix}\right] \tag{2.9}$$

for any $(U,u_3) \in S^1 \times S^2 \times S^3$ *where*

$$\Lambda_0 = \left(1 + \frac{1}{2}(1 - \frac{\beta_0}{\lambda_1}) + \sqrt{\frac{1}{4}(1 - \frac{\beta_0}{\lambda_1})^2 + 1 - \frac{\beta_0}{\lambda_1}}\right)^{-1}$$

and

$$\Lambda_1 = \frac{\lambda_1}{\beta_0}\left(1 + \sqrt{1 - \frac{\beta_0}{\lambda_1}}\right).$$

Proof. The proof of the theorem is similar to the one given in [5] (Theorem 1) for the standard saddle point problem. In principle one has to check the details given there within the framework used here, i.e., using the spaces S^1, S^2, S^3 and the inner products $[\cdot,\cdot]_1$ and $[\cdot,\cdot]$. However, we note that some relations make use of the special structure of the matrix $\mathcal{B} = \begin{pmatrix} 0 & -B \end{pmatrix}$ and the proofs of the intermediate results are not the same. For the convenience of the reader we briefly recall the main steps by Bramble and Pasciak, translated to the two-fold saddle point structure and without proving details.

The symmetry of $[\cdot,\cdot]$ is obvious and the positive definiteness is due to Lemma 2.2 and Assumption (A4). The symmetry of $\mathcal{M}$ with respect to $[\cdot,\cdot]$ can be directly checked and, once the lower bound in (2.9) is proved, the positive definiteness of $\mathcal{M}$ follows from the relation

$$\left[\tilde{\mathcal{M}}\begin{pmatrix} U \\ u_3 \end{pmatrix}, \begin{pmatrix} U \\ u_3 \end{pmatrix}\right] = \left[\begin{pmatrix} I & 0 \\ 0 & B(B_1A_1^{-1}B_1^*)^{-1}B^* \end{pmatrix}\begin{pmatrix} U \\ u_3 \end{pmatrix}, \begin{pmatrix} U \\ u_3 \end{pmatrix}\right] \tag{2.10}$$

for any $(U,u_3) \in S^1 \times S^2 \times S^3$. This relation can be derived in a straightforward way. It therefore remains to prove (2.9). For this we recall only the intermediate steps from the proof of Theorem 1 in [5].

Let $(U,u_3)^T = (u_1,u_2,u_3)^T \in S^1 \times S^2 \times S^3$ be given. We consider the representation

$$\begin{pmatrix} U \\ u_3 \end{pmatrix} = \begin{pmatrix} U_0 \\ 0 \end{pmatrix} + \begin{pmatrix} U_H \\ u_3 \end{pmatrix}$$

where U_H is defined by

$$\mathcal{M}_1 U_H + \mathcal{B}^* u_3 = 0.$$

As in [5, (i)–(iii)] one finds

$$[\mathcal{M}_1 U_H, U_H]_1 = [\mathcal{M}_1^{-1}\mathcal{B}^* u_3, \mathcal{B}^* u_3]_1, \tag{2.11}$$

$$\left[\mathcal{M}\begin{pmatrix} U_H \\ u_3 \end{pmatrix}, \begin{pmatrix} U_H \\ u_3 \end{pmatrix}\right] = (\mathcal{B}\mathcal{M}_1^{-1}\mathcal{B}^* u_3, u_3), \tag{2.12}$$

and

(2.13)

$$\left[\mathcal{M}\begin{pmatrix} U \\ u_3 \end{pmatrix}, \begin{pmatrix} U \\ u_3 \end{pmatrix}\right] = \left[\mathcal{M}\begin{pmatrix} U_0 \\ 0 \end{pmatrix}, \begin{pmatrix} U_0 \\ 0 \end{pmatrix}\right] + \left[\mathcal{M}\begin{pmatrix} U_H \\ u_3 \end{pmatrix}, \begin{pmatrix} U_H \\ u_3 \end{pmatrix}\right]$$

for any $(U, u_3) \in S^1 \times S^2 \times S^3$. One also obtains

$$\begin{aligned} &\left[\mathcal{M}\begin{pmatrix} U_0 \\ 0 \end{pmatrix}, \begin{pmatrix} U_0 \\ 0 \end{pmatrix}\right] \\ &\quad = [(\mathcal{M}_1 - \mathcal{M}_0)\mathcal{M}_0^{-1}(\mathcal{M}_1 - \mathcal{M}_0)U_0, U_0]_1 + [(\mathcal{M}_1 - \mathcal{M}_0)U_0, U_0]_1. \end{aligned} \tag{2.14}$$

An upper bound for $\tilde{\mathcal{M}}$ is as follows:

$$\begin{aligned} &\left[\tilde{\mathcal{M}}\begin{pmatrix} U \\ u_3 \end{pmatrix}, \begin{pmatrix} U \\ u_3 \end{pmatrix}\right] \\ &\quad \le (1+\gamma)[(\mathcal{M}_1 - \mathcal{M}_0)U_0, U_0]_1 \\ &\quad + (1 + \frac{1}{\gamma})[(\mathcal{M}_1 - \mathcal{M}_0)U_H, U_H]_1 + (\mathcal{B}\mathcal{M}_1^{-1}\mathcal{B}^* u_3, u_3) \end{aligned} \tag{2.15}$$

which holds for any positive number γ. A combination of Lemma 2.2 and (2.11) yields

$$[(\mathcal{M}_1 - \mathcal{M}_0)U_H, U_H]_1 \le (1 - \frac{\beta_0}{\lambda_1})[\mathcal{M}_1^{-1}\mathcal{B}^* u_3, \mathcal{B}^* u_3]_1 \tag{2.16}$$

and, therefore

(2.17)

$$\begin{aligned} &\left[\tilde{\mathcal{M}}\begin{pmatrix} U \\ u_3 \end{pmatrix}, \begin{pmatrix} U \\ u_3 \end{pmatrix}\right] \\ &\quad \le (1+\gamma)[(\mathcal{M}_1 - \mathcal{M}_0)U_0, U_0]_1 + \left(1 + (1 - \frac{\beta_0}{\lambda_1})(1 + \frac{1}{\gamma})\right)(\mathcal{B}\mathcal{M}_1^{-1}\mathcal{B}^* u_3, u_3). \end{aligned}$$

Now, using (2.13), (2.14), and (2.12), we obtain

$$\left[\mathcal{M}\begin{pmatrix} U \\ u_3 \end{pmatrix}, \begin{pmatrix} U \\ u_3 \end{pmatrix}\right] \ge [(\mathcal{M}_1 - \mathcal{M}_0)U_0, U_0]_1 + (\mathcal{B}\mathcal{M}_1^{-1}\mathcal{B}^* u_3, u_3). \tag{2.18}$$

Hence, the lower bound for $\mathcal{M}$ follows by combining (2.17) and (2.18) and by choosing

$$\gamma = \frac{1}{2}(1 - \frac{\beta_0}{\lambda_1}) + \sqrt{\frac{1}{4}(1 - \frac{\beta_0}{\lambda_1})^2 + 1 - \frac{\beta_0}{\lambda_1}}.$$

It remains to prove the upper bound for $\mathcal{M}$. First we estimate

$$\left[\mathcal{M}\begin{pmatrix} U_0 \\ 0 \end{pmatrix}, \begin{pmatrix} U_0 \\ 0 \end{pmatrix}\right].$$

Using Lemma 2.2, Lemma 2.1(iii), and Assumption (A3), we find that there holds

$$[(\mathcal{M}_1 - \mathcal{M}_0)^{-1}U, U]_1 \geq \frac{\beta_0}{\lambda_1 - \beta_0}[\mathcal{M}_0^{-1}U, U]_1.$$

Therefore, using this estimate, (2.14) and (2.16), we obtain

$$\left[\mathcal{M}\begin{pmatrix} U_0 \\ 0 \end{pmatrix}, \begin{pmatrix} U_0 \\ 0 \end{pmatrix}\right] \\ \leq \frac{\lambda_1}{\beta_0}(1+\gamma)[(\mathcal{M}_1 - \mathcal{M}_0)U, U]_1 + (1 - \frac{\beta_0}{\lambda_1})\frac{\lambda_1}{\beta_0}(1 + \frac{1}{\gamma})(\mathcal{B}\mathcal{M}_1^{-1}\mathcal{B}^* u_3, u_3)$$

for any positive number γ. Therefore, by (2.12) and (2.13), we eventually obtain the estimate

$$\begin{aligned}\left[\mathcal{M}\begin{pmatrix} U \\ u_3 \end{pmatrix}, \begin{pmatrix} U \\ u_3 \end{pmatrix}\right] \leq\ & \frac{\lambda_1}{\beta_0}(1+\gamma)[(\mathcal{M}_1 - \mathcal{M}_0)U, U]_1 \\ & + \left(1 + (\frac{\lambda_1}{\beta_0} - 1)(1 + \frac{1}{\gamma})\right)(\mathcal{B}\mathcal{M}_1^{-1}\mathcal{B}^* u_3, u_3).\end{aligned}$$

Comparing this bound with the relation (2.15), the upper bound for $\mathcal{M}$ then follows by choosing $\gamma = \sqrt{1 - \frac{\beta_0}{\lambda_1}}$. □

3. Application to a dual-dual mixed method

3.1. **The model problem.** We now present a model problem which leads to linear systems of the dual-dual form (2.1). In order to solve those systems we perform the transformation $\mathcal{A} \to \mathcal{M}$ as described in §2 and use the preconditioned conjugate gradient method with the special inner product $[\cdot,\cdot]$; see Section 3.3. This PCG-method requires only a bounded number of iterations and uses a preconditioner that is sparse and therefore cheap. Without preconditioner the transformation leads to a system matrix that is ill-conditioned and, thus, requires a large number of CG-iterations.

First let us describe the model problem. Let Ω be a polygonal domain in $\mathbb{R}^2$ with boundary $\Gamma := \partial\Omega$. Then, given $f \in L^2(\Omega)$, $g \in H^{1/2}(\Gamma)$ and a matrix valued continuous function $\boldsymbol{\kappa}$, we consider the non-homogeneous Dirichlet problem: *Find $u \in H^1(\Omega)$ such that*

$$\begin{aligned} -\operatorname{div}(\boldsymbol{\kappa}\nabla u) &= f \quad \textit{in}\ \Omega \\ u &= g \quad \textit{on}\ \Gamma. \end{aligned} \tag{3.1}$$

Here, we assume that $\boldsymbol{\kappa}$ is symmetric and that there exists $C > 0$ such that

$$C\|\xi\|^2 \leq \sum_{i,j=1}^{2} \kappa_{ij}(x)\xi_i\xi_j \quad \forall \xi := (\xi_1, \xi_2) \in \mathbb{R}^2, \quad \forall x \in \bar{\Omega}, \tag{3.2}$$

with κ_{ij} being the entries of $\boldsymbol{\kappa}$.

The standard mixed finite element method for (3.1) requires first the definition of the flux $\boldsymbol{\sigma} := \boldsymbol{\kappa}\nabla u$ as an auxiliary unknown. Then, using that $\boldsymbol{\kappa}$ is invertible (because of (3.2)), the integration by parts procedure is applied to the relation $\nabla u = \boldsymbol{\kappa}^{-1}\boldsymbol{\sigma}$. Following [13] we introduce the additional explicit unknown $\boldsymbol{\theta} := \nabla u$; see also [16, 7]. However, instead of proceeding as in [16, 7], we use the variational formulation as a dual-dual operator equation as in [13], see also [10].

Let $X_1 := [L^2(\Omega)]^2$, $M_1 := H(\mathrm{div};\Omega)$, $X := X_1 \times M_1$, $M := L^2(\Omega)$, and define the bounded linear operators $\mathbf{A}_1 : X_1 \to X_1'$, $\mathbf{B}_1 : X_1 \to M_1'$, $\mathbf{A} : X \to X'$ and $\mathbf{B} : M_1 \to M'$, and the functionals $\mathbf{F}_1 \in X_1'$, $\mathbf{G}_1 \in M_1'$ and $\mathbf{G} \in M'$, as follows:

$$(\mathbf{A}_1(\boldsymbol{\theta}), \boldsymbol{\zeta})' := \int_\Omega \kappa\boldsymbol{\theta} \cdot \boldsymbol{\zeta}\, dx, \quad (\mathbf{B}_1(\boldsymbol{\theta}), \boldsymbol{\tau})' := -\int_\Omega \boldsymbol{\theta} \cdot \boldsymbol{\tau}\, dx,$$

$$(\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\sigma}), (\boldsymbol{\zeta}, \boldsymbol{\tau}))' := (\mathbf{A}_1(\boldsymbol{\theta}), \boldsymbol{\zeta})' + (\mathbf{B}_1(\boldsymbol{\zeta}), \boldsymbol{\sigma})' + (\mathbf{B}_1(\boldsymbol{\theta}), \boldsymbol{\tau})',$$

$$(\mathbf{B}(\boldsymbol{\sigma}), v)' := -\int_\Omega v \mathrm{div} \boldsymbol{\sigma}\, dx, \quad (\mathbf{F}_1, \boldsymbol{\zeta})' := 0,$$

and

$$(\mathbf{G}_1, \boldsymbol{\tau})' := -\langle g, \boldsymbol{\tau} \cdot \nu \rangle_{L^2(\Gamma)}, \quad (\mathbf{G}, v)' := \int_\Omega f v\, dx$$

for all $(\boldsymbol{\theta}, \boldsymbol{\sigma})$, $(\boldsymbol{\zeta}, \boldsymbol{\tau}) \in X$ and for all $v \in M$, where $(\cdot, \cdot)'$ stands for the duality pairing induced by the operators appearing in each case. Further, let $\mathbf{B}_1^* : M_1 \to X_1'$ and $\mathbf{B}^* : M \to M_1'$ be the adjoints of $\mathbf{B}_1$ and $\mathbf{B}$, respectively, and let $\mathbf{O}$ denote the null operator. It follows that $\mathbf{A}$ can be equivalently defined as

$$\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\sigma}) := \begin{pmatrix} \mathbf{A}_1 & \mathbf{B}_1^* \\ \mathbf{B}_1 & \mathbf{O} \end{pmatrix} \begin{pmatrix} \boldsymbol{\theta} \\ \boldsymbol{\sigma} \end{pmatrix} \in X' := X_1' \times M_1'. \tag{3.3}$$

Then, the variational formulation of problem (3.1) can be stated as the following operator equation (see [13]): *Find* $(\boldsymbol{\theta}, \boldsymbol{\sigma}, u) \in X_1 \times M_1 \times M$ *such that*

$$\begin{pmatrix} \mathbf{A}_1 & \mathbf{B}_1^* & \mathbf{O} \\ \mathbf{B}_1 & \mathbf{O} & \mathbf{B}^* \\ \mathbf{O} & \mathbf{B} & \mathbf{O} \end{pmatrix} \begin{pmatrix} \boldsymbol{\theta} \\ \boldsymbol{\sigma} \\ u \end{pmatrix} = \begin{pmatrix} \mathbf{F}_1 \\ \mathbf{G}_1 \\ \mathbf{G} \end{pmatrix}, \tag{3.4}$$

or equivalently: *Find* $((\boldsymbol{\theta}, \boldsymbol{\sigma}), u) \in X \times M$ *such that*

$$\begin{pmatrix} \mathbf{A} & \tilde{\mathbf{B}}^* \\ \tilde{\mathbf{B}} & \mathbf{O} \end{pmatrix} \begin{pmatrix} (\boldsymbol{\theta}, \boldsymbol{\sigma}) \\ u \end{pmatrix} = \begin{pmatrix} \mathbf{F} \\ \mathbf{G} \end{pmatrix}, \tag{3.5}$$

where $\mathbf{F} := (\mathbf{F}_1, \mathbf{G}_1) \in X' := X_1' \times M_1'$ and $\tilde{\mathbf{B}} : X \to M'$ is given by $\tilde{\mathbf{B}} := (\mathbf{O} \quad \mathbf{B})$.

The equation (3.4) ((3.5)) constitutes the so-called *dual-dual mixed formulation* of our model problem (3.1) since the operator $\mathbf{A}$ itself has the dual-type structure given by (3.3). This problem is uniquely solvable:

Theorem 3.1 ([13, Theorem 4]). *There exists a unique solution* $((\boldsymbol{\theta}, \boldsymbol{\sigma}), u) \in X \times M$ *of the dual-dual mixed formulation* (3.5).

Before ending this subsection we remark that after swapping rows and columns 2 and 3 of the continuous dual-dual formulation (3.4), one obtains a standard saddle point problem that satisfies the hypotheses of the usual Babuška-Brezzi theory (see Lemmas 3.1 and 3.3 in [7]). In particular, the resulting operator $\tilde{\mathbf{B}}$ applies $X_1 \times M$ into M_1', and its null space is given by

$$\{ (\boldsymbol{\zeta}, v) \in X_1 \times M : \ v \in H_0^1(\Omega) \text{ and } \boldsymbol{\zeta} = \nabla v \}.$$

However, this characterization of the continuous kernel of $\tilde{\mathbf{B}}$ cannot be extended to the discrete one and hence the subsequent application of the standard Babuška-Brezzi theory to the corresponding Galerkin scheme, including the associated error analysis, becomes too complicated. This fact was confirmed by the alternative analysis developed in [7], which, however, has the drawback of being too particularized to the formulation and to the specific finite element subspaces utilized there. On the contrary, the approach based on the two-fold saddle point formulations (see [10], [14]) has been shown to be simpler, more general and hence of wider applicability.

3.2. **The Galerkin scheme.** We now recall the Galerkin procedure from [13] for the approximate solution of (3.4) (or (3.5)). Let $\mathcal{T}_h$ be a regular triangulation of Ω made up of triangles T of diameter h_T such that $h := \sup_{T\in\mathcal{T}_h} h_T$ and $\bar{\Omega} = \cup\{T : T \in \mathcal{T}_h\}$. Next, we consider the canonical triangle with vertices $\hat{P}_1 = (0,0)^T$, $\hat{P}_2 = (1,0)^T$ and $\hat{P}_3 = (0,1)^T$ as a reference triangle $\hat{T}$, and introduce the family of bijective affine mappings $\{F_T\}_{T\in\mathcal{T}_h}$, such that $F_T(\hat{T}) = T$. It is well known that $F_T(\hat{x}) = B_T\hat{x} + b_T$ for all $\hat{x} \in \hat{T}$, where the square matrix B_T of order 2 and $b_T \in \mathbb{R}^2$ depend only on the vertices of T.

We consider the lowest order Raviart-Thomas spaces. For each triangle $T \in \mathcal{T}_h$ let

$$\mathcal{RT}_0(T) := \{\boldsymbol{\tau} : \boldsymbol{\tau} = |\det(B_T)|^{-1} B_T\, \hat{\boldsymbol{\tau}} \circ F_T^{-1},\ \hat{\boldsymbol{\tau}} \in \mathcal{RT}_0(\hat{T})\},$$

where

$$\mathcal{RT}_0(\hat{T}) := \text{span}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix}\right\}.$$

Then, we define the finite element subspaces for the unknowns $\boldsymbol{\theta}$ and $\boldsymbol{\sigma}$, respectively, as follows:

$$X_{1,h} := \{\boldsymbol{\zeta} \in [L^2(\Omega)]^2 : \boldsymbol{\zeta}|_T \in \mathcal{RT}_0(T) \quad \forall T \in \mathcal{T}_h\} \tag{3.6}$$

and

$$M_{1,h} := \{\boldsymbol{\tau} \in H(\text{div};\Omega) : \boldsymbol{\tau}|_T \in \mathcal{RT}_0(T) \quad \forall T \in \mathcal{T}_h\}. \tag{3.7}$$

Next, we put

$$X_h := X_{1,h} \times M_{1,h},$$

and consider the piecewise constant functions as the finite element subspace for the unknown u, i.e.

$$M_h := \{v \in L^2(\Omega) : v|_T \text{ is constant } \forall T \in \mathcal{T}_h\}. \tag{3.8}$$

Then, the Galerkin scheme for the continuous problem (3.4) ((3.5)) reads:
Find $((\boldsymbol{\theta}_h, \boldsymbol{\sigma}_h), u_h) \in X_h \times M_h$ such that

$$\begin{array}{llllll} (\mathbf{A}_1(\boldsymbol{\theta}_h),\boldsymbol{\zeta})' & + & (\mathbf{B}_1(\boldsymbol{\zeta}),\boldsymbol{\sigma}_h)' & & = & 0 \\ (\mathbf{B}_1(\boldsymbol{\theta}_h),\boldsymbol{\tau})' & & & + \ (\mathbf{B}(\boldsymbol{\tau}),u_h)' & = & (\mathbf{G}_1,\boldsymbol{\tau})' \\ & & (\mathbf{B}(\boldsymbol{\sigma}_h),v)' & & = & (\mathbf{G},v)' \end{array} \tag{3.9}$$

for all $((\boldsymbol{\zeta}, \boldsymbol{\tau}), v) \in X_h \times M_h$.

The convergence of the Galerkin scheme is as follows.

Theorem 3.2 ([13, Theorems 5, 6]). *There exists a unique solution $((\boldsymbol{\theta}_h, \boldsymbol{\sigma}_h), u_h) \in X_h \times M_h$ of the Galerkin system* (3.9). *Let $((\boldsymbol{\theta}, \boldsymbol{\sigma}), u)$ be the unique solution of*

(3.5). *Assume that* $\boldsymbol{\theta}|_T \in [H^1(T)]^2$ $\forall T \in \mathcal{T}_h$, $\boldsymbol{\sigma} \in [H^1(\Omega)]^2$, $\operatorname{div}\boldsymbol{\sigma} \in H^1(\Omega)$ *and* $u \in H^1(\Omega)$. *Then, there exists* $C > 0$ *which is independent of* h *such that*

$$\|((\boldsymbol{\theta},\boldsymbol{\sigma}),u) - ((\boldsymbol{\theta}_h,\boldsymbol{\sigma}_h),u_h)\|$$
$$\leq C\,h \left\{ \sum_{T\in\mathcal{T}_h} \|\boldsymbol{\theta}\|^2_{[H^1(T)]^2} + \|\boldsymbol{\sigma}\|^2_{[H^1(\Omega)]^2} + \|\operatorname{div}\boldsymbol{\sigma}\|^2_{H^1(\Omega)} + \|u\|^2_{H^1(\Omega)} \right\}^{1/2}.$$

3.3. **Conjugate gradient method.** Now let us study the preconditioned conjugate gradient method for the solution of the linear system (3.9). This system is exactly of the form (2.1) where the matrices A_1, B_1, and B correspond to discretizations of the operators $\mathbf{A}_1$, $\mathbf{B}_1$, and $\mathbf{B}$, respectively. In the following we use plain symbols for matrices which correspond to operators with the analogous bold symbol. For a coefficient vector which belongs to a certain function we use the same symbol as for the function.

Given preconditioners A_0 for A_1 and $\mathcal{M}_0$ for $\mathcal{M}_1$ (as defined in (2.3)) we transform the linear system (3.9) into the form

(3.10)
$$\mathcal{M}\begin{pmatrix} X \\ x_3 \end{pmatrix} := \begin{pmatrix} \mathcal{M}_0^{-1}\mathcal{M}_1 & \mathcal{M}_0^{-1}\mathcal{B}^* \\ \mathcal{B}\mathcal{M}_0^{-1}(\mathcal{M}_1 - \mathcal{M}_0) & \mathcal{B}\mathcal{M}_0^{-1}\mathcal{B}^* \end{pmatrix}\begin{pmatrix} X \\ x_3 \end{pmatrix} = \begin{pmatrix} \mathcal{M}_0^{-1}F \\ \mathcal{B}\mathcal{M}_0^{-1}F + f_3 \end{pmatrix}$$

given by (2.5) and use the preconditioned conjugate gradient method (CG) with the special inner product $[\cdot,\cdot]$ (2.8) as iterative solver. We note that the preconditioners A_0 and $\mathcal{M}_0$ are used only for the transformation above and both matrices are simple scalings for our model problem.

In view of Theorem 2.3 it is already clear what an efficient preconditioner for the transformed matrix $\mathcal{M}$ looks like. Since $\mathcal{M}$ is uniformly spectrally equivalent to $\tilde{\mathcal{M}}$ (if (A1)–(A4) are uniformly satisfied) one only needs to find a preconditioner for $B(B_1A_1^{-1}B_1^*)^{-1}B^*$, cf. (2.10). But since $B_1A_1^{-1}B_1^*$ is uniformly well conditioned we can simply take the sparse matrix BB^* as a preconditioner for this block. We then obtain an efficient preconditioned iterative method for the solution of (3.10) (and of (3.9)). Since all the matrices which need to be inverted in this procedure, i.e. A_0, M_0, and P, are either just scalings or sparse and since we obtain a bounded number of iterations, the complexity of this method is comparable to that of the preconditioned minimum residual method as proposed in [12] (for a different model problem).

Theorem 3.3. *Let the basis functions of all the three spaces* $X_{1,h}$, $M_{1,h}$ *and* M_h *be scaled such that their* L^∞*-norms are* $O(h^{-1})$. *Further, let us take a preconditioner*

$$\mathcal{P} := \begin{pmatrix} I & 0 \\ 0 & P \end{pmatrix}$$

with $I \in \mathbb{R}^{(L+M)\times(L+M)}$ *and* $P \in \mathbb{R}^{N\times N}$ *such that* P *is uniformly spectrally equivalent to* BB^*. *Then there exist constants* μ, ρ, ω *such that with*

$$A_0 := \mu I \in \mathbb{R}^{L\times L}, \quad M_0 := \omega I \in \mathbb{R}^{M\times M}, \quad \mathcal{M}_0 := \begin{pmatrix} \rho I & 0 \\ 0 & M_0 \end{pmatrix} \in \mathbb{R}^{(L+M)\times(L+M)}$$

there holds

$$\Lambda_0\left[\mathcal{P}\begin{pmatrix} \mathrm{U} \\ u \end{pmatrix}, \begin{pmatrix} \mathrm{U} \\ u \end{pmatrix}\right] \leq \left[\mathcal{M}\begin{pmatrix} \mathrm{U} \\ u \end{pmatrix}, \begin{pmatrix} \mathrm{U} \\ u \end{pmatrix}\right] \leq \Lambda_1\left[\mathcal{P}\begin{pmatrix} \mathrm{U} \\ u \end{pmatrix}, \begin{pmatrix} \mathrm{U} \\ u \end{pmatrix}\right]$$

for any $(\mathbf{U}, u) \in X_h \times M_h$. Λ_0 *and* Λ_1 *are positive constants being independent of* h. *Moreover, choosing* A_0, $\mathcal{M}_0$ *and the preconditioner* $\mathcal{P}$ *as above, the number of iterations of the preconditioned conjugate gradient method with inner product* $[\cdot,\cdot]$ *for the solution of* (3.10) *is bounded.*

In order to prove Theorem 3.3 we need to collect some estimates for the eigenvalues and singular values of A_1, B_1, and B. Let us denote the eigenvalues of A_1 by

$$0 < \lambda_1 \leq \lambda_1 \leq \cdots \leq \lambda_L.$$

We also need the singular values of B_1,

$$0 < \sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_M$$

and those of B,

$$0 < \eta_1 \leq \eta_2 \leq \cdots \leq \eta_N.$$

These values depend on the scaling of the basis functions in use.

Lemma 3.4. *Let the basis functions of all the three ansatz spaces* $X_{1,h}$, $M_{1,h}$ *and* M_h *be scaled to* L^∞*-norm being* $O(h^{-1})$. *Then there exist generic positive constants* c_0 *and* c_1 *being independent of* h *such that there holds*

$$c_0 \leq \lambda_1 \leq \lambda_L \leq c_1, \tag{3.11}$$

$$c_0 \leq \sigma_1 \leq \sigma_M \leq c_1, \tag{3.12}$$

and

$$c_0 \leq \eta_1 \leq \eta_N \leq c_1 h^{-1}. \tag{3.13}$$

Proof. Since A_1 and B_1 are simple Gram matrices for the spaces $X_{1,h} \times X_{1,h}$ and $X_{1,h} \times M_{1,h}$, respectively, the estimates (3.11) and (3.12) are obvious. The proof of (3.13) is also straightforward, cf. the general result [12, Lemma 1]. For the convenience of the reader we recall the short proof. First we note that $\mathbf{B}$ is bounded and that it satisfies an inf-sup condition (see [20], [6]): There exists $\beta > 0$ which is independent of h such that

$$\sup_{\boldsymbol{\tau} \in M_{1,h}\setminus\{0\}} \frac{(\mathbf{B}(\boldsymbol{\tau}), v)'}{\|\boldsymbol{\tau}\|_{M_1}} \geq \beta \|v\|_M \quad \text{for any } v \in M_h.$$

Now, taking the scalings within $M_{1,h}$, M_h and the inf-sup condition for $\mathbf{B}$ into account, we obtain

$$\begin{aligned}
\eta_1 &= \min_{v \in M_h\setminus\{0\}} \max_{\boldsymbol{\tau} \in M_{1,h}\setminus\{0\}} \frac{v^T B \boldsymbol{\tau}}{\sqrt{v^T v}\sqrt{\boldsymbol{\tau}^T \boldsymbol{\tau}}} \\
&= \min_{v \in M_h\setminus\{0\}} \max_{\boldsymbol{\tau} \in M_{1,h}\setminus\{0\}} \frac{(\mathbf{B}(\boldsymbol{\tau}), v)'}{\|v\|_M \|\boldsymbol{\tau}\|_{M_1}} \frac{\|v\|_M \|\boldsymbol{\tau}\|_{M_1}}{\sqrt{v^T v}\sqrt{\boldsymbol{\tau}^T \boldsymbol{\tau}}} \\
&\geq \min_{v \in M_h\setminus\{0\}} \max_{\boldsymbol{\tau} \in M_{1,h}\setminus\{0\}} \frac{(\mathbf{B}(\boldsymbol{\tau}), v)'}{\|v\|_M \|\boldsymbol{\tau}\|_{M_1}} \frac{\|v\|_{L^2(\Omega)} \|\boldsymbol{\tau}\|_{L^2(\Omega)\times L^2(\Omega)}}{\sqrt{v^T v}\sqrt{\boldsymbol{\tau}^T \boldsymbol{\tau}}} \geq \beta c.
\end{aligned}$$

For an upper bound of η_N we write

$$\eta_N = \max_{v \in M_h\setminus\{0\}} \max_{\boldsymbol{\tau} \in M_{1,h}\setminus\{0\}} \frac{(\mathbf{B}(\boldsymbol{\tau}), v)'}{\|v\|_M \|\boldsymbol{\tau}\|_{M_1}} \frac{\|v\|_M \|\boldsymbol{\tau}\|_{M_1}}{\sqrt{v^T v}\sqrt{\boldsymbol{\tau}^T \boldsymbol{\tau}}}.$$

By the inverse property of the basis functions within $M_{1,h}$ there holds with a constant $c > 0$

$$\|\boldsymbol{\tau}\|^2_{M_1} = \|\boldsymbol{\tau}\|^2_{L^2(\Omega)\times L^2(\Omega)} + \|\operatorname{div}\boldsymbol{\tau}\|^2_{L^2(\Omega)} \leq c(1+h^{-2})\|\boldsymbol{\tau}\|^2_{L^2(\Omega)\times L^2(\Omega)}$$

and thus, using the boundedness of $\mathbf{B}$,

$$\eta_N \leq c(1+h^{-2})^{1/2}\|\mathbf{B}\| \max_{v\in M_h\setminus\{0\}} \max_{\boldsymbol{\tau}\in M_{1,h}\setminus\{0\}} \frac{\|v\|_{L^2(\Omega)}\|\boldsymbol{\tau}\|_{L^2(\Omega)\times L^2(\Omega)}}{\sqrt{v^Tv}\sqrt{\boldsymbol{\tau}^T\boldsymbol{\tau}}} \leq ch^{-1}.$$

Therefore, the proof of the lemma is finished. □

We now prove the main theorem of this section.

Proof of Theorem 3.3. By Lemma 3.4 the mass matrix A_1 is uniformly spectrally equivalent to the identity matrix $I \in \mathbb{R}^{L\times L}$. Therefore, there exists a constant $\mu > 0$ such that Assumptions (A1), (A2) hold for $A_0 := \mu I$ and constants α_0, α_1 in (A1) being independent of h. Thus, in particular, Lemma 2.1(iii) holds with λ_0 and λ_1 being independent of h. Now, again by Lemma 3.4, $B_1A_1^{-1}B_1^*$ is uniformly spectrally equivalent to $I \in \mathbb{R}^{M\times M}$. We conclude, that there exist constants ρ, $\omega > 0$ such that Assumptions (A3), (A4) hold for $\mathcal{M}_0 = \begin{pmatrix} \rho I & 0 \\ 0 & \omega I \end{pmatrix}$ and constants β_0, β_1 being independent of h. Eventually, we know that Theorem 2.3 holds with constants Λ_0 and Λ_1 that are also independent of h. Using Theorem 2.3 we conclude that $\mathcal{M}$ is uniformly spectrally equivalent to $\tilde{\mathcal{M}}$. For the latter matrix there holds

$$\left[\tilde{\mathcal{M}}\begin{pmatrix} \mathbf{U} \\ u \end{pmatrix}, \begin{pmatrix} \mathbf{U} \\ u \end{pmatrix}\right] = \left[\begin{pmatrix} I & 0 \\ 0 & B(B_1A_1^{-1}B_1^*)^{-1}B^* \end{pmatrix}\begin{pmatrix} \mathbf{U} \\ u \end{pmatrix}, \begin{pmatrix} \mathbf{U} \\ u \end{pmatrix}\right]$$

$$\simeq \left[\begin{pmatrix} I & 0 \\ 0 & BB^* \end{pmatrix}\begin{pmatrix} \mathbf{U} \\ u \end{pmatrix}, \begin{pmatrix} \mathbf{U} \\ u \end{pmatrix}\right] \simeq \left[\mathcal{P}\begin{pmatrix} \mathbf{U} \\ u \end{pmatrix}, \begin{pmatrix} \mathbf{U} \\ u \end{pmatrix}\right]$$

for all $\mathbf{U} \in X_{1,h} \times M_{1,h}$ and $u \in M_h$ by (2.10), Lemma 3.4 and since $BB^* \simeq P$ by assumption. Here, $\simeq$ means spectral equivalence of the terms involved.

Therefore, the transformed matrix $\mathcal{M}$ and the preconditioner $\mathcal{P}$ are uniformly spectrally equivalent and the number of iterations of the preconditioned conjugate gradient method with inner product $[\cdot,\cdot]$ (and preconditioner $\mathcal{P}$) for the solution of (3.10) is bounded. □

As a conclusion of Theorem 3.3 we have the following result when no preconditioner is used.

Corollary 3.5. *Let the basis functions of all the three spaces $X_{1,h}$, $M_{1,h}$ and M_h be scaled such that their L^∞-norms are $O(h^{-1})$. Then there exist constants μ, ρ, ω such that with*

$$A_0 := \mu I \in \mathbb{R}^{L\times L}, \quad M_0 := \omega I \in \mathbb{R}^{M\times M}, \quad \mathcal{M}_0 := \begin{pmatrix} \rho I & 0 \\ 0 & M_0 \end{pmatrix} \in \mathbb{R}^{(L+M)\times(L+M)}$$

there holds

$$\Lambda_0\left[\begin{pmatrix} \mathbf{U} \\ u \end{pmatrix}, \begin{pmatrix} \mathbf{U} \\ u \end{pmatrix}\right] \leq \left[\mathcal{M}\begin{pmatrix} \mathbf{U} \\ u \end{pmatrix}, \begin{pmatrix} \mathbf{U} \\ u \end{pmatrix}\right] \leq \Lambda_1 h^{-2}\left[\begin{pmatrix} \mathbf{U} \\ u \end{pmatrix}, \begin{pmatrix} \mathbf{U} \\ u \end{pmatrix}\right]$$

for any $(\mathbf{U}, u) \in X_h \times M_h$. Here, Λ_0, Λ_1 are positive constants being independent of h, I denotes the identity matrix of generic size and A_0, $\mathcal{M}_0$ (together with A_1 and $\mathcal{M}_1$) define the inner product $[\cdot,\cdot]$, cf. (2.8), (2.7). Moreover, choosing A_0 and

$\mathcal{M}_0$ as above, the number of iterations of the conjugate gradient method with inner product $[\cdot,\cdot]$ for the solution of (3.10) *is bounded by $O(h^{-1})$.*

Proof. From the proof of Theorem 3.3 we know that $\mathcal{M}$ is uniformly spectrally equivalent to $\begin{pmatrix} I & 0 \\ 0 & B(B_1A_1^{-1}B_1^*)^{-1}B^* \end{pmatrix}$. Due to Lemma 3.4 the minimum eigenvalue of $B(B_1A_1^{-1}B_1^*)^{-1}B^*$ is bounded from below by a positive constant which is independent of h and the maximum eigenvalue of $B(B_1A_1^{-1}B_1^*)^{-1}B^*$ may grow like $O(h^{-2})$. Therefore, the spectral condition number of $\mathcal{M}$ behaves like $O(h^{-2})$ and the number of iterations of the CG method is bounded by $O(h^{-1})$. □

3.4. **Numerical results.** For the computational implementation of (3.9) we choose the finite element subspaces according to (3.6), (3.7) and (3.8). Let M and N be the number of edges and the number of triangles, respectively, of the triangulation $\mathcal{T}_h$, and let $L = 3N$. Then, we let $\{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, ..., \boldsymbol{\theta}_L\}$, $\{\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, ..., \boldsymbol{\sigma}_M\}$ and $\{u_1, u_2, ..., u_N\}$ be bases of $X_{1,h}$, $M_{1,h}$ and M_h, respectively. In particular, if $\{e_1, e_2, ..., e_M\}$ denote the edges of $\mathcal{T}_h$, the functions $\boldsymbol{\sigma}_j$ can be characterized by the relation

$$\boldsymbol{\sigma}_j \in M_{1,h} \quad \text{and} \quad \boldsymbol{\sigma}_j|_{e_i} \cdot \nu_i = c_j \delta_{ij} \quad \forall i,j \in \{1,2,...,M\},$$

where the c_j are scaling constants and ν_i denotes the unit normal on the edge e_i (in a previously chosen direction). In addition, if $\{T_1, T_2, ..., T_N\}$ denote the triangles of $\mathcal{T}_h$, we can take u_i such that $u_i|_{T_j} = \hat{c}_i\,\delta_{ij}$ for all $i,j \in \{1,2,...,N\}$, where the $\hat{c}_i$ are also scaling constants. We scale all the basis functions to $O(h^{-1})$.

We find that

$$A_1 = (a_{ij})_{L\times L} \quad \text{with} \quad a_{ij} := \int_\Omega \boldsymbol{\kappa}\boldsymbol{\theta}_i \cdot \boldsymbol{\theta}_j \, dx,$$

$$B_1 = (b_{ij}^{(1)})_{M\times L} \quad \text{with} \quad b_{ij}^{(1)} := -\int_\Omega \boldsymbol{\sigma}_i \cdot \boldsymbol{\theta}_j \, dx,$$

$$B = (b_{ij})_{N\times M} \quad \text{with} \quad b_{ij} := -\hat{c}_i \int_{T_i} \operatorname{div} \boldsymbol{\sigma}_j \, dx = \begin{cases} -\hat{c}_i \text{div}(\boldsymbol{\sigma}_j)|T_i| & \text{if } e_j \subset \bar{T}_i, \\ 0 & \text{otherwise,} \end{cases}$$

$$G_1 = (g_i^{(1)})_{M\times 1} \quad \text{with} \quad g_i^{(1)} := -\int_\Gamma g\boldsymbol{\sigma}_i \cdot \nu \, ds = \begin{cases} -(\boldsymbol{\sigma}_i \cdot \nu_i) \int_{e_i} g \, ds & \text{if } e_i \subset \Gamma, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$G = (g_i)_{N\times 1} \quad \text{with} \quad g_i := \hat{c}_i \int_{T_i} f \, dx.$$

For our numerical example we choose $\Omega = (0,1)\times(0,1)$ and take the right-hand side functions f and g in (3.1) such that $u(x_1,x_2) = 1/(x_1+x_2+1)$ and $\boldsymbol{\kappa} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$. Of course, the choice of the right-hand side does not influence the properties of the stiffness matrix. Moreover, we consider uniform triangular meshes.

For our computations we have to specify the preconditioners A_0 and $\mathcal{M}_0 = \begin{pmatrix} \rho I & 0 \\ 0 & M_0 \end{pmatrix}$. To confirm Theorem 3.3 and Corollary 3.5 we have to take $A_0 = \mu I$ and $M_0 = \omega I$. The parameters μ, ρ and ω have to be chosen sufficiently small such that (A1)–(A4) are satisfied. In practice, one can perform a power method to estimate the needed parameters. This can be done for a rather coarse mesh size

TABLE 1. The extreme eigenvalues of $\mathcal{M}$: (1) pure scalings for A_0, M_0, (2) pure scalings for A_0, M_0 plus preconditioner BB^*. dim is the number of unknowns $L + M + N$.

		(1)		(2)	
dim	$1/h$	$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\min}$	$\lambda_{\max}$
48	2	0.9912	481	0.8935	32.28
184	4	0.9878	2056	0.8742	38.68
408	6	0.9869	4712	0.8624	41.28
720	8	0.9864	8440	0.8549	42.57
1120	10	0.9861	13236	0.8500	43.36
1608	12	0.9859	19100	0.8466	43.90
2184	14	0.9857	26032	0.8442	44.28
2848	16	0.9856	34030	0.8424	44.56
3600	18	0.9855	43096	0.8410	44.78
4440	20	0.9855	53228	0.8399	44.95
5368	22	0.9854	64427	0.8390	45.09

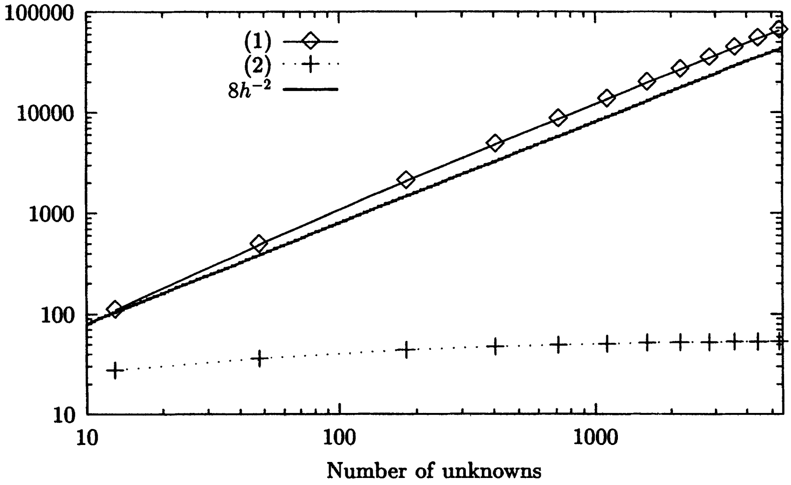

FIGURE 1. The condition numbers of $\mathcal{M}$: (1) pure scalings for A_0, M_0, (2) pure scalings for A_0, M_0 plus preconditioner BB^*.

since the behavior of the spectrum of the appearing matrices is quite stable. In our actual experiments we just tested some parameters and checked the positive definiteness of the inner product $[\cdot,\cdot]$ during the computations. The choice $\mu = 0.3$, $\rho = 0.7$ and $\omega = 0.06$ performed well.

To check Theorem 3.3 we need to take a preconditioner P that is equivalent to BB^*; we simply take BB^* itself. Actually, an application of this preconditioner requires only the solution of a sparse linear system. Figure 1 shows the condition numbers of the transformed stiffness matrices (plus preconditioner in the second case) in a double logarithmic scale. As given by Corollary 3.5 and Theorem 3.3 the spectral condition numbers increase like $O(h^{-2})$ without preconditioner and are bounded with preconditioner. Table 1 gives the extreme eigenvalues for both cases. As proved by Corollary 3.5 and Theorem 3.3 the minimum eigenvalues are

TABLE 2. The numbers of iterations of the CG method to reduce the initial residual by 10^{-6}: (1) pure scalings for A_0, M_0, (2) pure scalings for A_0, M_0 plus preconditioner BB^*. dim is the number of unknowns $L + M + N$.

dim	$1/h$	(1)	(2)
48	2	24	15
184	4	73	33
408	6	131	37
720	8	177	40
1120	10	223	40
1608	12	270	41
2184	14	316	42
2848	16	364	42
3600	18	410	41
4440	20	458	41
5368	22	504	41
6384	24	552	41
7488	26	598	41

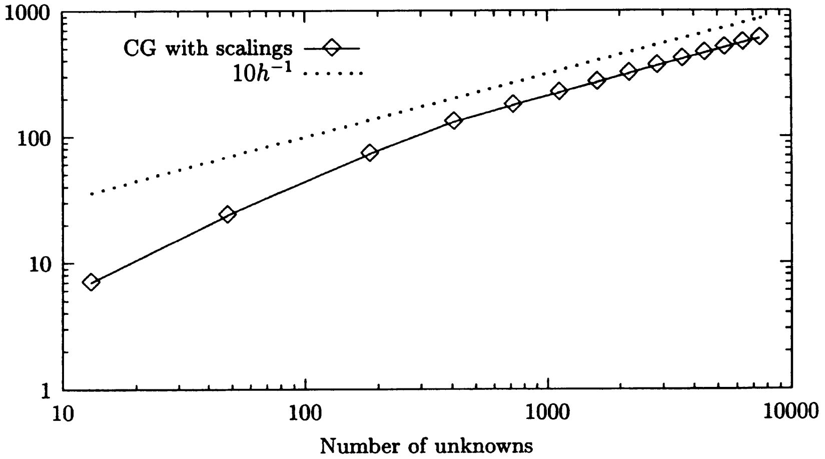

FIGURE 2. The number of iterations of the CG method to reduce the initial residual by 10^{-6} (no preconditioner, scalings for A_0 and M_0).

bounded from below for both methods. Indeed, this is already clear from the general Theorem 2.3 (by setting $\beta_0 = 0$). The magnitude of the condition number hinges only on the size of the maximum eigenvalue.

Table 2 presents the iteration numbers of the CG method which are required to reduce the initial residual in discrete l^2-norm by the factor 10^{-6}. They are quite large for the first method and appear to be bounded for the preconditioned method. The number of iterations without preconditioner increases asymptotically like $O(h^{-1})$ as is confirmed by Figure 2 where a double logarithmic scale is used. This is exactly what is expected due to Corollary 3.5.

Acknowledgment

This research was done while N. Heuer was visiting the Departamento de Ingeniería Matemática of the Universidad de Concepción, Chile, during the period April–September, 1999. This author expresses his gratitude to his colleagues in that department for their great hospitality.

References

[1] T. Arbogast, M.F. Wheeler and I. Yotov, *Mixed finite elements for elliptic problems with tensor coefficients as cell-centered finite differences*, SIAM J. Numer. Anal., 34 (1997), pp. 828–852. MR **98g**:65105

[2] S. F. Ashby, T. A. Manteuffel, and P. E. Saylor, *A taxonomy for conjugate gradient methods*, SIAM J. Numer. Anal., 27 (1990), pp. 1542–1568. MR **91i**:65062

[3] G. R. Barrenechea, G. N. Gatica, and J.-M. Thomas, *Primal mixed formulations for the coupling of FEM and BEM. part I: Linear problems*, Numer. Funct. Anal. Optim., 19 (1998), pp. 7–32. MR **99d**:65310

[4] D. Braess, Finite Elements. Theory, Fast Solvers, and Applications in Solid Mechanics, Cambridge University Press, 1997. MR **98f**:65002

[5] J. H. Bramble and J. E. Pasciak, *A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems*, Math. Comp., 50 (1988), pp. 1–17. MR **89m**:65097a

[6] F. Brezzi and M. Fortin, Mixed and Hybrid Finite Element Methods, Springer Verlag, 1991. MR **92d**:65187

[7] Z. Chen, *Expanded mixed finite element methods for linear second-order elliptic problems*, RAIRO Math. Model. Numer. Anal., 32 (1998), pp. 479–499. MR **99j**:65201

[8] Z. Chen, *Expanded mixed finite element methods for quasilinear second-order elliptic problems*, RAIRO Math. Model. Numer. Anal., 32 (1998), pp. 501–520. MR **99j**:65202

[9] L. Franca and A. Loula, *A new mixed finite element method for the Timoshenko beam problem*, RAIRO Modélisation Mathématique et Analyse Numérique, 25 (1991), pp. 561–578. MR **92e**:65150

[10] G. N. Gatica, *Solvability and Galerkin approximations of a class of nonlinear operator equations.* Technical Report 99-03, Departamento de Ingeniería Matemática, Universidad de Concepción, Chile. Submitted for publication. `http://www.ing-mat.udec.cl/inf-loc-dim.html`

[11] G. N. Gatica and N. Heuer, *A dual-dual mixed formulation for the coupling of mixed-FEM and BEM in hyperelasticity*, SIAM J. Numer. Anal., 38 (2000), pp. 380–400. MR **2001e**:65193

[12] G. N. Gatica and N. Heuer, *Minimum residual iteration for a dual-dual mixed formulation of exterior transmission problems*, Numer. Linear Algebra Appl., 8 (2001), pp. 147–164. CMP 2001:09

[13] G. N. Gatica and N. Heuer, *An expanded mixed finite element method via a dual-dual formulation and the minimum residual method*, J. Comput. Appl. Math., 132 (2001), pp. 371–385.

[14] G. N. Gatica, N. Heuer and S. Meddahi, *Solvability and fully discrete Galerkin schemes of nonlinear two-fold saddle point problems.* Technical Report 00-03, Departamento de Ingeniería Matemática, Universidad de Concepción, Chile. `http://www.ing-mat.udec.cl/inf-loc-dim.html`

[15] G. N. Gatica and S. Meddahi, *A dual-dual mixed formulation for nonlinear exterior transmission problems*, Math. Comp., 70 (2001), pp. 1461–1480.

[16] G. N. Gatica and W. L. Wendland, *Coupling of mixed finite elements and boundary elements for linear and nonlinear elliptic problems*, Appl. Anal., 63 (1996), pp. 39–75. MR **99a**:65167

[17] N. Heuer, M. Maischak, and E. P. Stephan, *Preconditioned minimum residual iteration for the h-p version of the coupled FEM/BEM with quasi-uniform meshes*, Numer. Linear Algebra Appl., 6 (1999), pp. 435–456. MR **2000j**:65112

[18] V. I. Lebedev, *Iterative methods for solving operator equations with a spectrum contained in several intervals*, USSR Comput. Math. and Math. Phys., 9 (1969), pp. 17–24. MR **42**:7052

[19] C. C. Paige and M. A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629. MR **52:**4595

[20] J. E. Roberts and J.-M. Thomas, Mixed and Hybrid Methods. In: *Handbook of Numerical Analysis*, edited by P.G. Ciarlet and J.L. Lions, vol. II, *Finite Element Methods* (Part 1), North-Holland, Amsterdam, 1991. CMP 91:14

[21] T. Rusten and R. Winther, *A preconditioned iterative method for saddlepoint problems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 887–904. MR **93a:**65043

[22] A. J. Wathen, B. Fischer, and D. J. Silvester, *The convergence of iterative solution methods for symmetric and indefinite linear systems*, in Numerical Analysis 1997, D. F. Griffiths and G. A. Watson, eds., Pitman Research Notes in Mathematics, Harlow, England, 1997, pp. 230–243. MR **99e:**65058

GI^{2}MA, Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile.

E-mail address: ggatica@ing-mat.udec.cl

GI^{2}MA, Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile.

E-mail address: norbert@ing-mat.udec.cl