PARAMETER-ROBUST DISCRETIZATION AND PRECONDITIONING OF BIOT'S CONSOLIDATION MODEL

JEONGHUN J. LEE, KENT-ANDRE MARDAL AND RAGNAR WINTHER

ABSTRACT. Biot's consolidation model in poroelasticity has a number of applications in science, medicine, and engineering. The model depends on various parameters, and in practical applications these parameters ranges over several orders of magnitude. A current challenge is to design discretization techniques and solution algorithms that are well behaved with respect to these variations. The purpose of this paper is to study finite element discretizations of this model and construct block diagonal preconditioners for the discrete Biot systems. The approach taken here is to consider the stability of the problem in non-standard or weighted Hilbert spaces and employ the operator preconditioning approach. We derive preconditioners that are robust with respect to both the variations of the parameters and the mesh refinement. The parameters of interest are small time-step sizes, large bulk and shear moduli, and small hydraulic conductivity.

1. INTRODUCTION

Biot's consolidation model describes the deformation of an elastic porous medium and the viscous fluid flow inside when the porous medium is saturated by the fluid. The unknowns are the displacement of the elastic medium, $\boldsymbol{u}$, and the fluid pressure, p_F. In homogeneous isotropic linear elastic porous media, the equations for the quasi-static Biot model are:

$$(1.1) \quad \begin{aligned} -\operatorname{div}(2\mu\underline{\boldsymbol{\epsilon}}(\boldsymbol{u}) + \lambda \operatorname{div}\boldsymbol{u}\underline{\boldsymbol{I}} - \alpha p_F\underline{\boldsymbol{I}}) &= \boldsymbol{f}, \\ s_0\dot{p}_F + \alpha \operatorname{div}\dot{\boldsymbol{u}} - \operatorname{div}(\kappa\nabla p_F) &= g, \end{aligned}$$

where the dots represent time derivatives, μ and λ are the Lamé coefficients of elastic medium, $\underline{\boldsymbol{\epsilon}}(\boldsymbol{u})$ is the symmetric gradient of $\boldsymbol{u}$, $\underline{\boldsymbol{I}}$ is the $n \times n$ identity matrix, $s_0 \geq 0$ is the constrained specific storage coefficient, $\kappa > 0$ is the hydraulic conductivity determined by the permeability of medium and the fluid viscosity, and $\alpha > 0$ is the Biot–Willis constant which is close to 1. The system (1.1) can be posed on a bounded domain in two and three space dimensions. The given functions $\boldsymbol{f}$ and g represent body force and source/sink of fluid, respectively. We will assume throughout the paper that the parameters μ, λ, and s_0 are scalar functions on the domain, while in general κ can be a symmetric positive definite matrix-valued function. To be a well-posed mathematical problem, the system (1.1) needs appropriate boundary and initial conditions. A discussion of general boundary conditions for

the Biot system will be given in Remark 3.3 below. Furthermore, a mathematical discussion of well-posedness of this model can be found in [1].

Due to importance of Biot's model in applications, ranging from geoscience to medicine, finite element methods for the model have been studied by many researchers. For example, various primal methods are studied in [2, 3, 4], mixed methods in [5, 6, 7], Galerkin least square methods in [8], discontinuous Galerkin methods in [9], and combinations of different methods in [10, 11, 12, 13], but this list is by no means complete.

It is important to construct numerical methods which are robust with respect to variation of model parameters since this variation in many practical problems is quite large. For example, relevant parameters in the soft tissue of the central nervous system are Young's modulus of $1 - 60$ kPa, Poisson ratio from 0.3 to almost 0.5 (0.499 in [14]), and the permeability is $10^{-14} - 10^{-16} m^2$ [15, 16]. In geophysics, Young's modulus is typically in the order of GPa, Poisson ratio $0.1 - 0.3$, while the permeability may vary from approximately 10^{-9} to $10^{-21} m^2$ [17, 18]. Relations of Young's modulus E, Poisson ratio ν and the two elastic moduli μ, λ are $\mu = E/2(1 + \nu)$ and $\lambda = E\nu/(1 + \nu)(1 - 2\nu)$. Consequently, μ and λ are in the ranges of $300 - 500$ MPa and $100 - 500$ MPa, respectively, in geoscience applications, whereas corresponding numbers are μ and λ in the ranges $300 - 2000$ Pa and $500 - 10^6$ Pa in neurological applications.

However, in the present paper we will not limit ourselves to the study of robustness with respect to model parameters of the finite element discretization of Biot's model. In fact, our main concern is to be able to construct preconditioners for the discrete systems which are well behaved both with respect to variations of the model parameters and the refinement of the discretization. When large discrete systems are solved by iterative methods, the convergence rate depends heavily on the construction of suitable preconditioners. Such preconditioners for finite element discretizations of Biot's model have been studied by many authors, cf. for example [19, 20, 21, 22, 23]. Recently, there is also an emerging interest for preconditioners which are robust with respect to model parameters [23, 24]. However, robustness with respect to all model parameters remains challenging. In particular, we will derive preconditioners that are robust as the medium approaches the incompressibility limit while the permeability is low. In our experience this represents the most difficult case, and it is also the case that occurs in many biomechanical applications.

The purpose of this paper is to develop a stable finite element method for Biot's model, and a corresponding preconditioner for the associated discrete systems, such that the preconditioned systems have condition numbers which are robust with respect to variations of model parameters. More precisely, we aim to have a preconditioned system which is robust for small κ, small time-steps, large λ, large μ, and mesh refinements. In order to obtain such a parameter-robust preconditioner we employ the operator preconditioning framework of [25]. It turns out that typical formulations of Biot's model are not appropriate to apply the framework, so we develop a new three-field formulation of Biot's model and propose a parameter-robust block diagonal preconditioner for it.

The present paper is organized as follows. In Section 2, we introduce some notation and conventions that will be used throughout the paper. Furthermore, we briefly discuss the preconditioning framework of [25] based on parameter-robust

stability of the continuous problems, and illustrate this with some numerical examples based on simplified models which can be seen as subsystems of the Biot system. In Section 3 we explain some difficulties related to more common formulations of the Biot system, and as a consequence we motivate a new three-field formulation. The discussion of finite element discretizations based on this three-field formulation is given in Section 4, and the stability results are used to motivate the construction of parameter robust preconditioners. The implementation of a special operator related to one of the blocks a block diagonal preconditioner is discussed in Section 5. Finally, in Section 6 we present some numerical experiments which illustrate our theoretical results.

2. Preliminaries

The system (1.1) can in principle be studied on rather general domains Ω in two or three dimensions. However, our main goal is to study finite element approximations of this system, and therefore we will assume throughout this paper that Ω is a bounded polyhedral domain in $\mathbb{R}^n$, with $n = 2$ or 3. We will use $H^k = H^k(\Omega)$ to denote the Sobolev space of functions on Ω with k derivatives in L^2 and the corresponding norm is denoted by $\| \cdot \|_k$. Further, let H_0^k be the closure of $C_0^\infty(\Omega)$ in H^k with dual space denoted by H^{-k} and $(\cdot, \cdot)$ denote the L^2 inner product of scalar, vector, and matrix valued functions as well as the duality paring between H_0^k and H^{-k}. The space L_0^2 is the space of L^2 functions with mean value zero. Boldface symbols are used to denote vector valued functions or spaces, and symbols of boldface with underline are used to denote matrix valued functions.

Throughout this paper we use $A \lesssim B$ to denote the inequality $A \leq CB$ with a generic constant $C > 0$ which is independent of the discretization parameters and the model parameters, and $A \sim B$ will stand for $A \lesssim B$ and $B \lesssim A$. If needed, we will use C to denote generic positive constants in inequalities. For a scalar valued function g, ∇g is a (column) vector valued function. For a matrix valued function $\underline{\boldsymbol{g}}$, div is understood as a row-wise divergence which results in the vector valued function div $\underline{\boldsymbol{g}}$. Adopting these conventions, the equations (1.1) are well-defined.

2.1. **Preconditioning of parameter-dependent systems.** Let us briefly review the abstract framework of parameter-robust preconditioning in [25]. Let X be a separable, real Hilbert space with inner product $\langle \cdot, \cdot \rangle_X$ and the associated norm $\| \cdot \|_X$. For two Hilbert spaces X and Y, $\mathcal{L}(X, Y)$ is the Hilbert space of bounded linear maps from X to Y. Let us denote the dual space of X by X^* and the duality pairing of X and X^* by $\langle \cdot, \cdot \rangle$. Suppose that $\mathcal{A} \in \mathcal{L}(X, X^*)$ is invertible and also symmetric in the sense that

$$\langle \mathcal{A}x, y \rangle = \langle x, \mathcal{A}y \rangle, \qquad x, y \in X.$$

For given $f \in X^*$ consider a problem finding $x \in X$ such that

$$(2.1) \qquad\qquad\qquad \mathcal{A}x = f.$$

Its preconditioned problem with a symmetric isomorphism $\mathcal{B} \in \mathcal{L}(X^*, X)$ is

$$\mathcal{B}\mathcal{A}x = \mathcal{B}f.$$

The convergence rate of a Krylov space method for this problem can be bounded by the condition number, $K(\mathcal{B}\mathcal{A})$, given by

$$K(\mathcal{B}\mathcal{A}) := \|\mathcal{B}\mathcal{A}\|_{\mathcal{L}(X,X)} \|(\mathcal{B}\mathcal{A})^{-1}\|_{\mathcal{L}(X,X)}.$$

Parameter-dependent problems are handled in this framework as follows. Let ε denote a collection of parameters, and $\mathcal{A}_\varepsilon$ the parameter-dependent coefficient operator. A systematic way to construct an ε-robust preconditioner $\mathcal{B}_\varepsilon$, as proposed in [25], is to consider the mapping property of $\mathcal{A}_\varepsilon$ in ε-dependent Hilbert spaces X_ε and X_ε^*. The key property is to choose the spaces X_ε and X_ε^* such that $\mathcal{A}_\varepsilon$ is a map from X_ε to X_ε^*, and with corresponding operator norms $\|\mathcal{A}_\varepsilon\|_{\mathcal{L}(X_\varepsilon, X_\varepsilon^*)}$ and $\|\mathcal{A}_\varepsilon^{-1}\|_{\mathcal{L}(X_\varepsilon^*, X_\varepsilon)}$ bounded independently of ε. In this case the preferred preconditioner, $\mathcal{B}_\varepsilon$, is a map from X_ε^* to X_ε with the property that $\|\mathcal{B}_\varepsilon\|_{\mathcal{L}(X_\varepsilon^*, X_\varepsilon)}$ and $\|\mathcal{B}_\varepsilon^{-1}\|_{\mathcal{L}(X_\varepsilon, X_\varepsilon^*)}$ are bounded independently of ε. If such an operator $\mathcal{B}_\varepsilon$ is identified, then the condition number $K(\mathcal{B}_\varepsilon \mathcal{A}_\varepsilon)$ is bounded independently of ε, since both $\mathcal{B}_\varepsilon \mathcal{A}_\varepsilon$ and $(\mathcal{B}_\varepsilon \mathcal{A}_\varepsilon)^{-1}$ are operators on $\mathcal{L}(X_\varepsilon, X_\varepsilon)$, with corresponding operator norms bounded independently of ε.

As we will illustrate below the discussion outlined above can often most easily be done in the continuous setting. On the other hand, in a computational setting we need preconditioners for the corresponding discrete problems. In fact, if we utilize a finite element discretization which is uniformly stable with respect to the parameters, then the structure of the preconditioners in the discrete case will be the natural discrete analogs of the preconditioners in the continuous case. However, the preconditioners derived by the procedure above will often require exact inverses of operators which cannot be inverted cheaply. Therefore, in order to obtain effective robust preconditioners in the discrete case, we also have to replace these exact inverses by related equivalent operators, often obtained by common procedures such as multilevel methods or domain decomposition methods. We refer to [25] and the examples below for more details.

A challenge of the Biot system is the dependency of several different and independent parameters like the Lamé elastic parameters as well as parameters related to porous flow such as permeability and the Biot-Willis constant. The aim is to achieve robustness with respect to all model parameters, as well as the resolution of the discretization. To motivate this discussion we start by considering two simplified examples related to the Biot equations. The first example illustrates the case where the permeability tends to zero, while the Lamé parameters are of unit scale. In the second example we consider the case where the elastic material tends towards the incompressible limit. Both examples are special cases of the Biot system.

Example 2.1. Consider a system of equations

(2.2)
$$-\Delta \boldsymbol{u} + \nabla p = \boldsymbol{f},$$
$$-\operatorname{div} \boldsymbol{u} + \operatorname{div}(\kappa \nabla p) = g,$$

with unknowns $\boldsymbol{u} : \Omega \to \mathbb{R}^n$ and $p : \Omega \to \mathbb{R}$. As boundary conditions we use homogeneous Dirichlet condition for $\boldsymbol{u}$, i.e., $\boldsymbol{u}|_{\partial \Omega} = 0$, while we use homogeneous Neumann condition for p. The parameter $\kappa > 0$ is taken to be a constant in this example. A variational formulation of this problem is to find $(\boldsymbol{u}, p) \in \boldsymbol{H}_0^1 \times H^1 \cap L_0^2$ satisfying

$$(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) - (p, \operatorname{div} \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}), \qquad \forall \boldsymbol{v} \in \boldsymbol{H}_0^1,$$
$$-(\operatorname{div} \boldsymbol{u}, q) - (\kappa \nabla p, \nabla q) = (g, q), \qquad \forall q \in H^1 \cap L_0^2.$$

This system has a form of (2.1) with $X = \boldsymbol{H}_0^1 \times H^1 \cap L_0^2$ and

$$\mathcal{A} = \begin{pmatrix} -\Delta & \nabla \\ -\operatorname{div} & \operatorname{div}(\kappa \nabla) \end{pmatrix}.$$

If we define X_κ as the Hilbert space $\boldsymbol{H}_0^1 \times H^1 \cap L_0^2$ with κ-dependent norm given by

$$\|(\boldsymbol{u}, p)\|_{X_\kappa}^2 = \|\boldsymbol{u}\|_1^2 + \|p\|_0^2 + \kappa \|\nabla p\|_0^2,$$

then one can check that $\mathcal{A} : X_\kappa \to X_\kappa^*$ is an isomorphism with corresponding operator norms of $\mathcal{A}$ and $\mathcal{A}^{-1}$ bounded independently of κ. Here the norm on $X_\kappa^* \supset \boldsymbol{L}^2 \times L^2$ is defined from the norm on X_κ by extending the L^2 inner product to a duality pairing. To define a robust preconditioner we need to identify an isomorphism $\mathcal{B} : X_\kappa^* \to X_\kappa$ with corresponding operator norms of $\mathcal{B}$ and $\mathcal{B}^{-1}$ bounded independently of κ. A natural choice is a block-diagonal operator of the form

$$(2.3) \qquad \mathcal{B} = \begin{pmatrix} -\Delta^{-1} & 0 \\ 0 & (I - \kappa\Delta)^{-1} \end{pmatrix}.$$

However, the operator $\mathcal{B}$ only indicates the structure of the desired preconditioner of the discrete system. We will discretize this problem on the unit square in $\mathbb{R}^2$ by the lowest order Taylor-Hood element with respect to a mesh of uniform squares. This method is uniformly stable with respect to κ in the proper norms. Furthermore, to obtain an effective preconditioner the exact inverses in the definition of $\mathcal{B}$ are replaced by corresponding algebraic multigrid preconditioners for the operators $-\Delta$ and $I - \kappa\Delta$, implemented in the software library Hypre [26] with default settings. The preconditioned system is implemented using cbc.block [27] and FEniCS [28]. The eigenvalue estimates are obtained by the conjugate gradient method of normal equation of the system with convergence criterion 10^{-16}, and we refer to [27] for more details. The same setup is used throughout the paper.

We present numerical results in Table 1. The numbers of iterations and condition numbers increase as the value of κ decreases but they are asymptotically stable and are still bounded in the limit case $\kappa = 0$, which is the Stokes equation. We remark that the zero eigenvalue of the system associated with $H^1 \cap L_0^2$ is ignored in the computation of condition numbers.

TABLE 1. Number of iterations of MinRes solver of system (2.2) with algebraic multigrid (AMG) preconditioner of the structure (2.3). Estimates of the condition numbers of the preconditioned system are given in parenthesis. ($\Omega = $ unit square, partitioned as bisections of $N \times N$ rectangles, convergence criterion[1] with relative residual of 10^{-6})

				N		
		16	32	64	128	256
	10^0	13 (1.3)	13 (1.2)	14 (1.2)	14 (1.2)	14 (1.2)
	10^{-1}	16 (1.7)	16 (1.8)	16 (1.8)	16 (1.8)	16 (1.8)
	10^{-2}	22 (3.1)	24 (3.4)	23 (3.6)	23 (3.8)	24 (3.9)
	10^{-3}	30 (4.6)	29 (4.9)	29 (5.4)	30 (5.8)	30 (6.1)
κ	10^{-4}	35 (6.1)	36 (6.2)	36 (6.4)	35 (6.6)	34 (7.0)
	10^{-5}	38 (7.0)	39 (7.8)	41 (7.9)	39 (7.7)	39 (7.7)
	10^{-6}	38 (7.1)	40 (8.3)	42 (9.2)	44 (9.4)	43 (9.0)
	0	38 (7.1)	42 (8.4)	44 (9.5)	47 (10.5)	48 (11.2)

Example 2.2. Consider a system related to the Lamé problem in linear elasticity:
Find $\boldsymbol{u} : \Omega \to \mathbb{R}^n$, $p : \Omega \to \mathbb{R}$ for

$$-\operatorname{div} \underline{\boldsymbol{\epsilon}}(\boldsymbol{u}) - \nabla p = \boldsymbol{f},$$

(2.4)
$$\operatorname{div} \boldsymbol{u} - \frac{1}{\lambda} p = g,$$

with $\boldsymbol{u}|_{\partial\Omega} = 0$, where $1 \le \lambda < +\infty$ is a positive constant and $\underline{\boldsymbol{\epsilon}}(\boldsymbol{u})$ is the symmetric
gradient of $\boldsymbol{u}$. When $g = 0$, this is the Lamé problem and $p = \lambda \operatorname{div} u$ is called
"solid pressure". Its variational form is to find $\boldsymbol{u} \in \boldsymbol{H}_0^1$ and $p \in L^2$ such that

$$(\underline{\boldsymbol{\epsilon}}(\boldsymbol{u}), \underline{\boldsymbol{\epsilon}}(\boldsymbol{v})) + (p, \operatorname{div} \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}), \qquad\qquad \forall \boldsymbol{v} \in \boldsymbol{H}_0^1,$$

$$(\operatorname{div} \boldsymbol{u}, q) - \frac{1}{\lambda}(p, q) = g, \qquad\qquad \forall q \in L^2.$$

This is a saddle point problem with a stabilizing term $-(1/\lambda)(p, q)$, and the sta-
bilization becomes weaker as λ becomes larger. Let $X = \boldsymbol{H}_0^1 \times L^2$ be the Hilbert
space with standard norm. In the limit when $\lambda = +\infty$ the system is not stable
in these norms. This is due to the fact that $\operatorname{div} \boldsymbol{H}_0^1 \subsetneq L^2$ and as a consequence
the Brezzi condition for stability of saddle point problem is not fulfilled [29]. More
precisely, $\operatorname{div} \boldsymbol{H}_0^1$ can control only the L^2 norm of the mean-value zero part of p,
and the stabilizing term is needed to control the mean-value part of p. Since the
stabilizing term is dependent on λ, we need a λ-dependent norm on $\boldsymbol{H}_0^1 \times L^2$ to
have λ-independent stability of the system.

Before we define an appropriate λ-dependent norm, we need some preliminaries.
Let P_m be the linear operator in L^2 such that

$$P_m \phi := \left(\frac{1}{|\Omega|} \int_\Omega \phi \, dx \right) \chi_\Omega, \qquad \forall \phi \in L^2,$$

where χ_Ω is the characteristic function on Ω and $|\Omega|$ is the Lebesgue measure of Ω.
Notice that $P_m \phi$ and $\phi - P_m \phi$ are the decomposition of ϕ into its *mean-value* part
and *mean-value zero* part. For $q \in L^2$ we denote its mean-value and mean-value
zero parts by

(2.5)
$$q_m := P_m q, \qquad q_0 = q - q_m.$$

We now define a Hilbert space X_λ by

$$\|(\boldsymbol{v}, q)\|_{X_\lambda}^2 = \|\boldsymbol{v}\|_1^2 + \left(\frac{1}{\lambda} \|q_m\|_0^2 + \|q_0\|_0^2 \right),$$

then the system is λ-independent stable in X_λ. We will not give a detailed proof of
stability here since it can be obtained by modifying the proof of Theorem 3.2 below.
But the rough explanation is that the mean value of p cannot be control by the
inf-sup condition, and therefore this part of the norm has to be weighted properly
in balance with the stabilizing term $\lambda^{-1} \|p_m\|_0^2$, while the rest of p is controlled by
the inf-sup condition.

We present numerical results for two different preconditioners. The Hilbert
space X and X_λ lead to preconditioners of the forms

(2.6)
$$\mathcal{B}_1 = \begin{pmatrix} -\Delta^{-1} & 0 \\ 0 & I^{-1} \end{pmatrix}, \qquad \mathcal{B}_2 = \begin{pmatrix} -\Delta^{-1} & 0 \\ 0 & \left(I_0 + \frac{1}{\lambda} I_m\right)^{-1} \end{pmatrix}.$$

---

[1]Convergence criterion is $(\mathcal{B}r_k, r_k)/(\mathcal{B}r_0, r_0) \le 10^{-6}$ where $\mathcal{B}$ is preconditioner and r_k is the
residual of k-th iteration.

Here the appearance of the operator I^{-1} calls for an explanation. In fact, the operator I should not be thought of as the identity operator on the Hilbert space L^2, but rather as the Riesz map between this space and its dual. In particular, in the corresponding discrete setting the operator I^{-1} is typically represented by the inverse of a mass matrix. In a similar manner, the operators I_0 and I_m should be interpreted as maps of L^2 into the duals of L_0^2 and its complement. We refer to [25, Section 6] for more details.

We employ the lowest order Taylor-Hood discretization and test the efficiency of the preconditioners on different refinements of the unit square. The preconditioner $\mathcal{B}_1$ is implemented by replacing the exact inverse of $-\Delta$ by an algebraic multigrid operator, and we use the Jacobi iteration to approximate I^{-1}. The construction of $\mathcal{B}_2$ is technical due to the second block, and we postpone the details of the construction to Section 5.

Numerical results for the preconditioners $\mathcal{B}_1$ and $\mathcal{B}_2$ are given in Tables 2 and 3. The results for the preconditioner with structure $\mathcal{B}_2$ is completely satisfactory. Both the number of iterations and the condition numbers appear to be uniformly bounded with respect to λ and mesh refinement. However, the results for the preconditioner $\mathcal{B}_1$ are different. In this case the condition numbers appear to grow linearly with λ, while the number of iterations still seems to be bounded, even if they appear to be slightly less robust in this case. So in the case of the preconditioner $\mathcal{B}_1$ the condition number will not lead to a sharp bound on the number of iterations. In fact, by comparing the operator $\mathcal{B}_1$ with the uniform preconditioner $\mathcal{B}_2$ one can argue that the operator $\mathcal{B}_1\mathcal{A}$ has one isolated eigenvalue that tends to 0 as λ increases, while the rest of the spectrum lie on an interval bounded independently of the mesh resolution and λ. In cases where only few eigenvalues are outside a bounded spectrum, it is well-known that the Conjugate Gradient and the Minimum Residual methods are very efficient [30, 31] and therefore $\mathcal{B}_1$ is about as efficient as $\mathcal{B}_2$, but slightly less robust.

TABLE 2. Number of iterations of MinRes solver of system (2.4) with preconditioner of the form $\mathcal{B}_1$ in (2.6). Estimates of the condition numbers of the preconditioned system are given in parenthesis. (Ω = unit square, partitioned as bisections of $N \times N$ rectangles, convergence criterion with relative residual of 10^{-6})

| | | N | | | |
		16	32	64	128	256
	10^0	29 (3.5)	29 (3.6)	29 (3.6)	29 (3.6)	29 (3.6)
	10^1	40 (10.8)	41 (10.8)	38 (10.9)	38 (10.9)	36 (10.9)
	10^2	53 (95.6)	59 (96.2)	54 (96.7)	53 (96.9)	52 (97.0)
λ	10^3	60 (958)	62 (964)	62 (969)	61 (972)	60 (973)
	10^4	66 (9589)	69 (9649)	44 (9697)	43 (9724)	42 (9734)
	10^5	44 (95892)	44 (96501)	44 (96981)	43 (97248)	42 (97344)
	10^6	44 (958942)	44 (964970)	44 (969940)	43 (972566)	41 (973587)

2.2. **Parameter rescaling of the Biot system.** The systems in the above two examples have only one parameter, so it is relatively easy to find function spaces

TABLE 3. Number of iterations of MinRes solver of system (2.4) with preconditioner of the form $\mathcal{B}_2$ in (2.6). Estimates of the condition numbers of the preconditioned system are given in parenthesis. (Ω = unit square, partitioned as bisections of $N \times N$ rectangles, convergence criterion with relative residual of 10^{-6})

		N				
		16	32	64	128	256
	10^0	29 (3.5)	29 (3.6)	29 (3.6)	29 (3.6)	29 (3.6)
	10^1	41 (11.6)	41 (11.6)	41 (11.6)	38 (11.6)	38 (11.6)
	10^2	46 (18.4)	45 (18.4)	44 (18.4)	44 (18.4)	44 (18.4)
λ	10^3	46 (19.6)	46 (19.6)	45 (19.6)	45 (19.6)	44 (19.6)
	10^4	45 (19.7)	46 (19.7)	45 (19.7)	45 (19.7)	44 (19.7)
	10^5	45 (19.7)	46 (19.7)	46 (19.7)	44 (19.7)	44 (19.7)
	10^6	46 (19.7)	46 (19.7)	43 (19.7)	44 (19.7)	44 (19.7)

and parameter-dependent norms such that the aforementioned preconditioner construction is applicable. However, the Biot system has several parameters of different ranges, so it is not easy to find appropriate function spaces and their parameter-dependent norms. In the rest of this section, we will rescale parameters of the Biot system and reduce it to a problem with three parameters. This procedure will not only simplify the problem but also clarify intrinsic parameters of the system. We emphasize that μ, λ, s_0, and κ are allowed to be functions on the domain, while α is assumed to be a constant.

Recall that when we solve a time-dependent problem numerically, we discretize the problem in time and solve a static problem at each time step, so preconditioning of time-dependent problem is reduced to preconditioning of static problem at each time step. If we consider an implicit time discretization (e.g., the backward Euler method) applied to (1.1) with time-step size δ^2 ($0 < \delta \leq 1$), and multiply the second equation with $-\delta^2$, then we obtain a static problem

$$- \operatorname{div}(2\mu \underline{\boldsymbol{\epsilon}}(\boldsymbol{u}) + \lambda \operatorname{div} \boldsymbol{u} \underline{\boldsymbol{I}} - \alpha p_F \underline{\boldsymbol{I}}) = \boldsymbol{f},$$
$$-s_0 p_F - \alpha \operatorname{div} \boldsymbol{u} + \delta^2 \operatorname{div}(\kappa \nabla p_F) = \tilde{g}$$

with some right hand side $\tilde{g}$.

To reduce this system further, we recall the physical derivation of s_0. The storage coefficient s_0 is the increase of the amount of fluid for the unit increase of fluid pressure, when volumetric strain is kept as constant. More precisely, $s_0 = \phi c_F + (1 - \phi)c_S$ where ϕ is the porosity of solid, and $c_F \geq 0$, $c_S \geq 0$ are compressibilities of the fluid and solid. For derivation of these equations from physical modeling, we refer to standard porous media references, for instance, [32]. The parameter c_S and other parameters μ and λ are related so that $c_S \sim 1/(2\mu/n + \lambda)$ holds with n, the spatial dimension of Ω. In many practical applications, $\mu \lesssim \lambda$ and $c_F \approx 0$ hold, so we have $s_0 \sim 1/\lambda$. Furthermore, α is a constant close to 1, so we will assume that s_0 scales like α^2/λ. Therefore, to reduce the number of parameters in our system we will simply let $s_0 = \alpha^2/\lambda$ for the rest of the discussion in this paper. We emphasize that the equality is not essential. Our analysis below can easily be adopted to the situation where s_0 scales like α^2/λ. However, this rather artificial expression of s_0 is useful when we normalize the system later. In addition, $\delta^2 \kappa$ can

be regarded as one small parameter because the hydraulic conductivity κ is small in general. As a consequence, introducing $\tilde{\kappa} = \delta^2 \kappa$, we have a simplified system

$$- \operatorname{div}(2\mu \underline{\epsilon}(\boldsymbol{u}) + \lambda \operatorname{div} \boldsymbol{u} \underline{\boldsymbol{I}} - \alpha p_F \underline{\boldsymbol{I}}) = \boldsymbol{f},$$

$$-\frac{\alpha^2}{\lambda} p_F - \alpha \operatorname{div} \boldsymbol{u} + \operatorname{div}(\tilde{\kappa} \nabla p_F) = \tilde{g},$$

with $1 \le \lambda < +\infty$, small $\tilde{\kappa}$, and $\alpha \sim 1$. However, in practical applications, μ can be much larger than α, so we rescale the above equations to include this factor. To do so we assume μ is a spatial function with a uniform scale, i.e., there is a constant $\bar{\mu}$ such that $\mu/\bar{\mu} \sim 1$. Let

$$\lambda' = \frac{\lambda}{2\bar{\mu}}, \quad \mu' = \frac{\mu}{\bar{\mu}}, \quad \alpha' = \frac{\alpha}{2\bar{\mu}}, \quad \kappa' = \frac{\tilde{\kappa}}{2\bar{\mu}}, \quad \boldsymbol{f}' = \frac{\boldsymbol{f}}{2\bar{\mu}}, \quad g' = \frac{\tilde{g}}{2\bar{\mu}}.$$

Dividing the above two equations by $2\bar{\mu}$, the final simplified equations are

$$- \operatorname{div}(\mu' \underline{\epsilon}(\boldsymbol{u}) - \lambda' \operatorname{div} \boldsymbol{u} \underline{\boldsymbol{I}} - \alpha' p_F \underline{\boldsymbol{I}}) = \boldsymbol{f}',$$

(2.7)

$$-\frac{(\alpha')^2}{\lambda'} p_F - \alpha' \operatorname{div} \boldsymbol{u} + \operatorname{div}(\kappa' \nabla p_F) = g',$$

where α' is a small positive constant, λ' and μ' are positive scalar functions such that λ' is bounded from below and $\mu' \sim 1$, while κ' is a positive definite matrix valued function with eigenvalues bounded from above.

In summary, we have reduced the original Biot system to a system with three intrinsic parameters λ', α', and κ' which all may be unbounded. More precisely, the positive constant α' may be arbitrarily small, the scalar function λ' may be arbitrarily large, while the positive eigenvalues of κ' may be close to zero. Since μ' is bounded from above and below this parameter has no essential effect on the properties of the system. Therefore, to reduce the number of parameters of the system, we take $\mu' = 1$ in the discussion below.

For the biomedical applications discussed in the introduction in units of Pascal, gram, milimeter and second, μ was in the order of $1 - 60$ kPa which makes λ', α', κ' in the ranges of $0.25 - 500$, $10^{-3} - 10^{-5}$, $10^{-8} - 10^{-12}$, respectively. In geoscience, a representative μ is 10 GPa and units for pressure, viscosity and permeability are pounds per square inch (psi), centi Poise (cP), mili Darcy (mD). Using these units, λ', α', κ' are in the ranges of $0.25 - 3.5$, 10^{-10}, $10^{-9} - 10^{-13}$, respectively.

3. Parameter-robust stability of the continuous problems

For the rest of this paper we will discuss a system of the form (2.7), where the parameter $\mu' = 1$. By omitting the primes on the parameters we obtain a system of the form

$$- \operatorname{div}(\underline{\epsilon}(\boldsymbol{u}) - \lambda \operatorname{div} \boldsymbol{u} \underline{\boldsymbol{I}} - \alpha p_F \underline{\boldsymbol{I}}) = \boldsymbol{f},$$

(3.1)

$$-\frac{\alpha^2}{\lambda} p_F - \alpha \operatorname{div} \boldsymbol{u} + \operatorname{div}(\kappa \nabla p_F) = g.$$

Throughout the rest of the paper we will assume that the parameters λ, α, and κ satisfy

(3.2) $$1 \le \lambda < +\infty, \qquad 0 < \alpha \le 1, \qquad 0 < \kappa \le 1,$$

where the assumption on the matrix valued function κ has the interpretation that κ is uniformly positive definite, and with all eigenvalues bounded above by one.

3.1. **Difficulties in typical formulations.** In the discussion of the two examples in Section 2 we saw that simplified versions of the Biot system were efficiently handled with straightforward extensions of standard preconditioners for Stokes problem. In particular, in Example 2.1, the presence of a small permeability was handled by extending a standard Stokes preconditioner in the canonical way with an operator of the form $\nabla \cdot (\kappa \nabla)$ applied to the pressure. Furthermore, in Example 2.2, the use of "solid pressure" gave a stable formulation and an efficient preconditioner even in the limit of an incompressible material. We will now demonstrate that extending these two approaches to the Biot system is not straightforward.

For simplicity we will first consider the system with homogeneous Dirichlet boundary conditions of the form $\boldsymbol{u} = 0$ and $p_F = 0$ on $\partial\Omega$. A discussion of other possible boundary conditions is given in Remark 3.3 below. Throughout this subsection we will make the simplifying assumption that λ is a constant, and we will illustrate that parameter-robust preconditioning is difficult even in that case. For a variational formulation of (3.1) with Dirichlet boundary conditions, we will use the function spaces $\boldsymbol{H}_0^1$ and H_0^1 for the unknowns $\boldsymbol{u}$ and p_F, respectively, and obtain

$$(\underline{\boldsymbol{\epsilon}}(\boldsymbol{u}), \underline{\boldsymbol{\epsilon}}(\boldsymbol{v})) + \lambda(\operatorname{div}\boldsymbol{u}, \operatorname{div}\boldsymbol{v}) - \alpha(p_F, \operatorname{div}\boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}), \qquad \forall \boldsymbol{v} \in \boldsymbol{H}_0^1,$$

$$-\alpha(\operatorname{div}\boldsymbol{u}, q_F) - \frac{\alpha^2}{\lambda}(p_F, q_F) - (\kappa\nabla p_F, \nabla q_F) = (g, q_F), \qquad \forall q_F \in H_0^1.$$

In matrix form, the system is

$$(3.3) \qquad \mathcal{A}\begin{pmatrix} \boldsymbol{u} \\ p_F \end{pmatrix} := \begin{pmatrix} -\operatorname{div}(\underline{\boldsymbol{\epsilon}} + \lambda \operatorname{div}\underline{\boldsymbol{I}}) & \alpha\nabla \\ -\alpha\operatorname{div} & -\frac{\alpha^2}{\lambda} + \operatorname{div}(\kappa\nabla) \end{pmatrix} \begin{pmatrix} \boldsymbol{u} \\ p_F \end{pmatrix} = \begin{pmatrix} \boldsymbol{f} \\ g \end{pmatrix}.$$

This system has a perturbed saddle point problem structure. Following the preconditioner construction framework in the previous section, we define Hilbert spaces $\boldsymbol{V}_1$ and $Q_{F,1}$ with parameter-dependent norms on $\boldsymbol{H}_0^1$ and H_0^1,

$$\|\boldsymbol{v}\|_{\boldsymbol{V}_1}^2 := \|\underline{\boldsymbol{\epsilon}}(\boldsymbol{v})\|_0^2 + \lambda\|\operatorname{div}\boldsymbol{v}\|_0^2, \qquad \boldsymbol{v} \in \boldsymbol{H}_0^1,$$

$$\|q_F\|_{Q_{F,1}}^2 := \frac{\alpha^2}{\lambda}\|q_F\|_0^2 + (\kappa\nabla q_F, \nabla q_F), \qquad q_F \in H_0^1.$$

Then we are able to show that $\mathcal{A} : \boldsymbol{V}_1 \times Q_{F,1} \to \boldsymbol{V}_1^* \times Q_{F,1}^*$ in (3.3) is an isomorphism with upper and lower bounds uniform in λ, κ, and α. However, this formulation has a nontrivial difficulty to achieve parameter-robust preconditioner for its discrete counterpart. For example, if we consider a block-diagonal preconditioner as in examples in the previous section, we need a good preconditioner of $-\operatorname{div}\underline{\boldsymbol{\epsilon}} - \lambda\operatorname{grad}\operatorname{div}$ in first block of the preconditioner. However, usually accepted preconditioners (e.g., algebraic multigrid preconditioner) do not perform well when λ is large. This is observed in [23] for the simplified McKenzie equations, and can be explained by the fact that it is hard to construct discretizations which are uniformly stable with respect to λ.

This difficulty is similar to volumetric locking problem in linear elasticity [33], which arises when λ is very large. Thus, we expect that resolutions of the locking problem in elasticity are useful to circumvent this preconditioning problem. There are two mathematically equivalent ways to avoid the locking problem: one is reduced integration technique [34] and the other is the mixed method (see, e.g., [35, 36, 37, 38]). However, both of them have technical difficulties in parameter-robust preconditioning. Here we will discuss the difficulty with the mixed approach.

Motivated by Example 2.2 and the mixed finite element technique to avoid the locking problem in linear elasticity, it is tempting to employ the "solid pressure" $p_S := -\lambda \operatorname{div} \boldsymbol{u}$. Recall that our purpose here is to show that this formulation is *not* appropriate for parameter-robust preconditioner construction, so we show the lack of stability and bad numerical results only for one specific case $\alpha = 1$. Introducing $p_S = -\lambda \operatorname{div} \boldsymbol{u}$ in the first equation of (3.1), with some algebraic manipulation, we have a three-field formulation

$$
\begin{aligned}
-\operatorname{div} \underline{\boldsymbol{\epsilon}}(\boldsymbol{u}) + \nabla p_S + \nabla p_F &= \boldsymbol{f}, \\
-\operatorname{div} \boldsymbol{u} - \frac{1}{\lambda} p_S &= 0, \\
-\operatorname{div} \boldsymbol{u} - \frac{1}{\lambda} p_F + \operatorname{div}(\kappa \nabla p_F) &= g.
\end{aligned}
$$
(3.4)

Since $\boldsymbol{u} \in \boldsymbol{H}_0^1$, $\operatorname{div} \boldsymbol{u}$ is mean-value zero, and therefore $p_S = -\lambda \operatorname{div} \boldsymbol{u}$ is mean-value zero as well. This means that, L_0^2 is an appropriate function space for p_S in variational formulation. Thus, a variational form of (3.4) is to find $(\boldsymbol{u}, p_S, p_F) \in \boldsymbol{H}_0^1 \times L_0^2 \times H_0^1$ such that

$$
\begin{aligned}
(\underline{\boldsymbol{\epsilon}}(\boldsymbol{u}), \underline{\boldsymbol{\epsilon}}(\boldsymbol{v})) - (p_S + p_F, \operatorname{div} \boldsymbol{v}) &= (\boldsymbol{f}, \boldsymbol{v}), && \forall \boldsymbol{v} \in \boldsymbol{H}_0^1, \\
-(\operatorname{div} \boldsymbol{u}, q_S) - \frac{1}{\lambda}(p_S, q_S) &= 0, && \forall q_S \in L_0^2, \\
-(\operatorname{div} \boldsymbol{u}, q_F) - \frac{1}{\lambda}(p_F, q_F) - (\kappa \nabla p_F, \nabla q_F) &= (g, q_F), && \forall q_F \in H_0^1.
\end{aligned}
$$

In order to have parameter-robust stability, we need to find parameter-dependent norms of $\boldsymbol{H}_0^1 \times L_0^2 \times H_0^1$ such that the above system gives a linear isomorphism from the Hilbert space to its dual space, and norms of the linear isomorphism and its inverse are independent of the parameters. The bilinear forms in the system,

$$(3.5) \qquad (\underline{\boldsymbol{\epsilon}}(\boldsymbol{u}), \underline{\boldsymbol{\epsilon}}(\boldsymbol{v})), \qquad\qquad\qquad \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{H}_0^1,$$

$$(3.6) \qquad (\operatorname{div} \boldsymbol{v}, p_S), \qquad\qquad\qquad \boldsymbol{v} \in \boldsymbol{H}_0^1, p_S \in L_0^2,$$

$$(3.7) \qquad (\operatorname{div} \boldsymbol{v}, p_F), \qquad\qquad\qquad \boldsymbol{v} \in \boldsymbol{H}_0^1, p_F \in H_0^1,$$

$$\qquad\qquad \frac{1}{\lambda}(p_S, q_S), \qquad\qquad\qquad p_S, q_S \in L_0^2,$$

$$(3.8) \qquad \frac{1}{\lambda}(p_F, q_F) - (\kappa \nabla p_F, \nabla q_F), \qquad p_F, q_F \in H_0^1,$$

have to be bounded for the parameter-dependent norms, so some necessary conditions of the norms will be given. The bilinear form (3.5) suggests H^1-norm for $\boldsymbol{H}_0^1$. To make (3.6) and (3.7) bounded, the chosen norms of L_0^2 and H_0^1 have to bound the standard L^2-norms of p_S and p_F, respectively. Finally, (3.8) enforces the norm of H_0^1 to be an upper bound of $(\kappa \nabla p_F, \nabla p_F)^{1/2}$. Thus the smallest possible norms for $\boldsymbol{H}_0^1$, L_0^2, H_0^1 from these observations are

$$(3.9) \qquad \|\underline{\boldsymbol{\epsilon}}(\boldsymbol{u})\|_0, \qquad \|p_S\|_0, \qquad \left(\|p_F\|_0^2 + (\kappa \nabla p_F, \nabla p_F)\right)^{\frac{1}{2}}$$

for $\boldsymbol{u} \in \boldsymbol{H}_0^1$, $p_S \in L_0^2$, $p_F \in H_0^1$.

We use $\boldsymbol{V}_2$, Q_S, $Q_{F,2}$ to denote the Hilbert spaces on $\boldsymbol{H}_0^1$, L_0^2, H_0^1 with the above norms. It is easy to check that all bilinear forms are bounded with these norms uniformly in λ and κ. In other word, the linear map from $\boldsymbol{V}_2 \times Q_S \times Q_{F,2}$ to $\boldsymbol{V}_2^* \times Q_S^* \times Q_{F,2}^*$, given by the above three-field formulation has a uniform bound

independent of the parameters. Unfortunately, it does not seem to be the case for the inverse of the linear map. Although the system is a stabilized saddle point problem, the stabilization terms $-\lambda^{-1}(p_S, q_S)$ and $-\lambda^{-1}(p_F, q_F)$ are not enough to control the L^2 norms of p_S and p_F when λ is very large. Therefore, we need to control the norms by $\boldsymbol{v} \in \boldsymbol{V}_2$ with inf-sup condition. In other words, we need

$$(3.10) \qquad \inf_{(q_S, q_F) \in Q_S \times Q_{F,2}} \sup_{\boldsymbol{v} \in \boldsymbol{V}_2} \frac{(\operatorname{div} \boldsymbol{v}, q_S + q_F)}{\|\boldsymbol{v}\|_1 (\|q_S\|_0 + \|q_F\|_0)} \geq \beta > 0,$$

with a constant β which is independent of the parameters. However, both q_S and q_F interact with $\operatorname{div} \boldsymbol{v}$ in the bilinear form $(\operatorname{div} \boldsymbol{v}, q_S + q_F)$, and it is difficult to control two independent terms with one object, $\operatorname{div} \boldsymbol{v}$. When κ is not small the stabilization term $(\kappa \nabla p_F, \nabla p_F)$ can be used to control the L^2-norm of p_F. However, as we have seen in the model reduction in the previous section, the smallness of κ is given not only by small hydraulic conductivity in the model, but also by a small time-step. Thus it is inevitable to assume that κ is small when we solve static problems at each time step.

Example 3.1. We present a computational result which provides numerical evidence for the above discussion. Suppose that Ω is the unit square in $\mathbb{R}^2$ and

$$\Gamma_d = \{(x, y) \in \mathbb{R}^2 \ : \ (x, y) \in \partial\Omega, x < 1\}.$$

For simplicity of implementation, we assume that $\boldsymbol{u}$ is vanishing on Γ_d, not on $\partial\Omega$, and therefore the appropriate function space for p_S is L^2, since $\lambda \operatorname{div} \boldsymbol{u} \notin L_0^2$. This is a reasonable assumption since robust preconditioners should cover problems with general boundary conditions. For discretization we use elements inspired by

TABLE 4. Number of iterations for different λ and κ with the preconditioner of the form in (3.11). (Ω = unit square, partitioned as bisections of $N \times N$ rectangles, convergence criterion with relative residual of 10^{-6})

N	λ	k			($\kappa = 10^k$)			
		0	−1	−2	−3	−4	−5	−6
32	10^0	21	20	21	22	24	25	25
	10^2	45	46	49	88	115	111	84
	10^4	46	46	50	102	247	359	198
	10^6	46	46	53	102	251	451	238
64	10^0	20	20	20	21	23	23	24
	10^2	44	44	49	85	107	107	92
	10^4	44	44	48	98	242	315	207
	10^6	44	44	50	98	244	370	231
128	10^0	19	19	19	21	22	23	23
	10^2	40	40	44	82	97	99	93
	10^4	42	42	47	94	220	241	181
	10^6	40	41	46	94	228	262	189

the lowest order Taylor–Hood element, i.e., $(\boldsymbol{\mathcal{P}}_2, \mathcal{P}_1, \mathcal{P}_1)$ Lagrange finite elements for $(\boldsymbol{u}, p_S, p_F)$. In the experiment, block-diagonal preconditioner $\mathcal{B}$ based on the

norms in (3.9), i.e.,

$$(3.11) \qquad \mathcal{B} = \begin{pmatrix} -\Delta^{-1} & 0 & 0 \\ 0 & I^{-1} & 0 \\ 0 & 0 & (I - \kappa\Delta)^{-1} \end{pmatrix}$$

is implemented using FEniCS with Hypre. As above, the exact inverses are replaced by suitable algebraic multigrid operators. A heuristic way to validate this choice of preconditioner is to consider a special case of (3.4) with $\lambda = 1$. Noting that this special case has the systems in Example 2.1 with $\lambda = 1$ and in Example 2.2 as subsystems, the operator $\mathcal{B}$ in (3.11) can be viewed as a combination of the robust preconditioners in those examples.

Numbers of iterations for different constant parameter values are given in Table 4. When λ is not too large the numbers of iterations are more or less robust with respect to variation of κ. However, the number of iterations clearly increases with increasing λ, and the increment is quite large when κ is small. The growth in iterations as λ increases is milder when κ is larger, which supports our heuristic analysis of "partial remedy", that $(\kappa\nabla p_F, \nabla p_F)$ may play a role of a stabilization term for the L^2 norm of p_F.

3.2. A new three-field formulation.

The discussion in the previous subsection suggests that we need a different formulation to obtain parameter-robust preconditioners. We will present such a new formulation of the system (3.1) here, where parameters λ, α, κ are assumed to be as specified in the beginning of this section. In particular, λ and κ are allowed to be functions of the spatial domain.

The main obstacle in the discussion above was that the inf-sup condition (3.10) is not fulfilled. To circumvent this, we use a different system with unknowns $(\boldsymbol{u}, p_T, p_F)$, where $p_T := -\lambda \operatorname{div} \boldsymbol{u} + \alpha p_F$. With this new unknown p_T, which will be called total pressure, we can rewrite the equations (3.1) as

$$-\operatorname{div} \underline{\boldsymbol{\epsilon}}(\boldsymbol{u}) + \nabla p_T = \boldsymbol{f},$$
$$-\operatorname{div} \boldsymbol{u} - \lambda^{-1}(p_T - \alpha p_F) = 0,$$
$$\lambda^{-1}(\alpha p_T - 2\alpha^2 p_F) + \operatorname{div}(\kappa\nabla p_F) = g.$$

The matrix form of this system is

$$(3.12) \quad \mathcal{A} \begin{pmatrix} \boldsymbol{u} \\ p_T \\ p_F \end{pmatrix} := \begin{pmatrix} -\operatorname{div}\underline{\boldsymbol{\epsilon}} & \nabla & 0 \\ -\operatorname{div} & -\lambda^{-1} & \alpha\lambda^{-1} \\ 0 & \alpha\lambda^{-1} & -2\alpha^2\lambda^{-1} + \operatorname{div}(\kappa\nabla) \end{pmatrix} \begin{pmatrix} \boldsymbol{u} \\ p_T \\ p_F \end{pmatrix} = \begin{pmatrix} \boldsymbol{f} \\ 0 \\ -g \end{pmatrix}.$$

With function spaces $\boldsymbol{H}_0^1$, L^2, H_0^1 for $\boldsymbol{u}$, p_T, p_F, respectively, corresponding to Dirichlet boundary conditions for $\boldsymbol{u}$ and p_F, we obtain a variational form

$$\begin{aligned} (\underline{\boldsymbol{\epsilon}}(\boldsymbol{u}), \underline{\boldsymbol{\epsilon}}(\boldsymbol{v})) - (\operatorname{div}\boldsymbol{v}, p_T) &= (\boldsymbol{f}, \boldsymbol{v}), & \boldsymbol{v} &\in \boldsymbol{H}_0^1, \\ (3.13) \qquad -(\operatorname{div}\boldsymbol{u}, q_T) - (\lambda^{-1}p_T, q_T) + (\alpha\lambda^{-1}p_F, q_T) &= 0, & q_T &\in L^2, \\ (\alpha\lambda^{-1}p_T, q_F) - 2(\alpha^2\lambda^{-1}p_F, q_F) - (\kappa\nabla p_F, \nabla q_F) &= (g, q_F), & q_F &\in H_0^1. \end{aligned}$$

Recall that we used the decomposition $p = p_m + p_0$ and the stabilization term in order to obtain the stability of the system in Example 2.2. We need a similar argument to show the stability of (3.13) due to the same reason, $\operatorname{div}\boldsymbol{H}_0^1 \subsetneq L^2$. Denoting the mean-value zero part of q_T by $q_{T,0}$ as in (2.5), we define norms by

$$(3.14) \quad \|\underline{\boldsymbol{\epsilon}}(\boldsymbol{u})\|_0, \quad \left((\lambda^{-1}p_T, p_T) + \|p_{T,0}\|_0^2\right)^{\frac{1}{2}}, \quad \left((\alpha^2\lambda^{-1}p_F, p_F) + (\kappa\nabla p_F, \nabla p_F)\right)^{\frac{1}{2}}$$

for $\boldsymbol{u} \in \boldsymbol{H}_0^1$, $p_T \in L^2$, $p_F \in H_0^1$. Let us denote these spaces with the norms in (3.14) by $\boldsymbol{V}$, Q_T, Q_F, and let $\mathcal{X} = \boldsymbol{V} \times Q_T \times Q_F$. Then it can be shown that all the bilinear forms in (3.13) are uniformly bounded in λ, α, κ. In other words, the operator $\mathcal{A}$ appearing in (3.12) is a bounded linear map from $\mathcal{X}$ to $\mathcal{X}^*$ and its norm is independent of the three parameters. Here the norm on the space $\mathcal{X}^*$ is derived from the norm on $\mathcal{X}$ exactly as we explained in Example 2.1.

The following theorem implies that $\mathcal{A}$ is invertible and that the inverse is a map from $\mathcal{X}^*$ to $\mathcal{X}$ with operator norm independent of the three parameters.

Theorem 3.2. *For the system* (3.13) *there exists a constant* $\beta > 0$, *independent of* λ, α, κ *satisfying* (3.2), *such that the following inf-sup condition holds:*

$$\inf_{(\boldsymbol{u},p_T,p_F)\in\mathcal{X}} \sup_{(\boldsymbol{v},q_T,q_F)\in\mathcal{X}} \frac{(\mathcal{A}(\boldsymbol{u},p_T,p_F),(\boldsymbol{v},q_T,q_F))_{(\mathcal{X}^*,\mathcal{X})}}{\|(\boldsymbol{u},p_T,p_F)\|_{\mathcal{X}}\|(\boldsymbol{v},q_T,q_F)\|_{\mathcal{X}}} \geq \beta.$$

Proof. To prove the inf-sup condition, we will use a standard technique: For given $(0,0,0) \neq (\boldsymbol{u},p_T,p_F) \in \mathcal{X}$, we will find $(\boldsymbol{v},q_T,q_F) \in \mathcal{X}$ such that

$$(3.15) \qquad\qquad \|(\boldsymbol{v},q_T,q_F)\|_{\mathcal{X}} \leq C_1 \|(\boldsymbol{u},p_T,p_F)\|_{\mathcal{X}},$$

$$(3.16) \qquad (\mathcal{A}(\boldsymbol{u},p_T,p_F),(\boldsymbol{v},q_T,q_F))_{(\mathcal{X}^*,\mathcal{X})} \geq C_2 \|(\boldsymbol{u},p_T,p_F)\|_{\mathcal{X}}^2,$$

with positive constants C_1, C_2 which are independent of λ, κ, and α. From these two inequalities we obtain that the desired inf-sup condition holds with $\beta = C_2/C_1$.

Suppose that $(0,0,0) \neq (\boldsymbol{u},p_T,p_F) \in \mathcal{X}$ is given. Recall that $p_{T,0}$ is the mean-value zero part of p_T. It is well-known from the theory of Stokes equation, cf. [39, Theorem 5.1] that there exists a constant $\beta_0 > 0$, depending only on the domain Ω, and $\boldsymbol{w} \in \boldsymbol{V}$, such that

$$(3.17) \qquad (\operatorname{div} \boldsymbol{w}, p_T) = \|p_{T,0}\|_0^2, \qquad (\underline{\boldsymbol{\epsilon}}(\boldsymbol{w}),\underline{\boldsymbol{\epsilon}}(\boldsymbol{w})) \leq \beta_0^2 \|p_{T,0}\|_0^2.$$

We set $\boldsymbol{v} = \boldsymbol{u} - \delta_0 \boldsymbol{w}$, $q_T = -p_T$, $q_F = -p_F$ with a constant δ_0 which will be determined later. One can check that

$$\|(\boldsymbol{v},q_T,q_F)\|_{\mathcal{X}} \leq \sqrt{2(1+\delta_0^2\beta_0^2)} \|(\boldsymbol{u},p_T,p_F)\|_{\mathcal{X}},$$

and (3.15) follows, if δ_0 is independent of the parameters of our interest.

To establish (3.16) and determine δ_0, we use the chosen $\boldsymbol{v}$, q_T, q_F, and (3.17) to obtain

$$(\mathcal{A}(\boldsymbol{u},p_T,p_F),(\boldsymbol{v},q_T,q_F))_{(\mathcal{X}^*,\mathcal{X})}$$
$$= \|\boldsymbol{u}\|_{\boldsymbol{V}}^2 - \delta_0(\underline{\boldsymbol{\epsilon}}(\boldsymbol{u}),\underline{\boldsymbol{\epsilon}}(\boldsymbol{w})) + \delta_0(\operatorname{div}\boldsymbol{w},p_T)$$
$$(3.18) \qquad + ((\lambda^{-1}p_T,p_T) + 2(\alpha^2\lambda^{-1}p_F,p_F) - 2(\alpha\lambda^{-1}p_T,p_F)) + (\kappa\nabla p_F,\nabla p_F)$$
$$= \|\boldsymbol{u}\|_{\boldsymbol{V}}^2 - \delta_0(\underline{\boldsymbol{\epsilon}}(\boldsymbol{u}),\underline{\boldsymbol{\epsilon}}(\boldsymbol{w})) + \delta_0\|p_{T,0}\|_0^2$$
$$+ ((\lambda^{-1}p_T,p_T) + 2(\alpha^2\lambda^{-1}p_F,p_F) - 2(\alpha\lambda^{-1}p_T,p_F)) + (\kappa\nabla p_F,\nabla p_F).$$

By Young's inequality and (3.17), we also have

$$(\underline{\boldsymbol{\epsilon}}(\boldsymbol{u}),\underline{\boldsymbol{\epsilon}}(\boldsymbol{w})) \leq \frac{1}{2\theta_0}\|\boldsymbol{u}\|_{\boldsymbol{V}}^2 + \frac{\theta_0}{2}\|\boldsymbol{w}\|_{\boldsymbol{V}}^2 \leq \frac{1}{2\theta_0}\|\boldsymbol{u}\|_{\boldsymbol{V}}^2 + \frac{\theta_0\beta_0^2}{2}\|p_{T,0}\|_0^2, \quad \forall\theta_0 > 0.$$

Using the above inequality with the choice $\theta_0 = \delta_0 = \beta_0^{-2}$, we derive

$$(3.19) \qquad \|\boldsymbol{u}\|_{\boldsymbol{V}}^2 - \delta_0(\underline{\boldsymbol{\epsilon}}(\boldsymbol{u}),\underline{\boldsymbol{\epsilon}}(\boldsymbol{w})) + \delta_0\|p_{T,0}\|_0^2 \geq \frac{1}{2}\|\boldsymbol{u}\|_{\boldsymbol{V}}^2 + \frac{\delta_0}{2}\|p_{T,0}\|_0^2.$$

Again by Young's inequality, for any $\theta_1 > 0$,

$$2(\alpha\lambda^{-1}p_T, p_F) = 2(\lambda^{-1/2}p_T, \alpha\lambda^{-1/2}p_F) \leq 2\theta_1(\lambda^{-1}p_T, p_T) + \frac{1}{2\theta_1}(\alpha^2\lambda^{-1}p_F, p_F).$$

If we take $\theta_1 = 3/8$, then we get

$$(3.20) \quad ((\lambda^{-1}p_T, p_T) + 2(\alpha^2\lambda^{-1}p_F, p_F) - 2(\alpha\lambda^{-1}p_T, p_F))$$
$$\geq \frac{1}{4}(\lambda^{-1}p_T, p_T) + \frac{2}{3}(\alpha^2\lambda^{-1}p_F, p_F).$$

The inequality (3.16) is obtained by combining (3.18), (3.19), and (3.20). Finally, we remark that the constants C_1 and C_2 in (3.15)–(3.16) depend only on δ_0, which is β_0^{-2} in the argument, so they are independent of λ, α, κ. $\qquad\square$

Remark 3.3. The set up in Theorem 3.2 is suitable for homogeneous Dirichlet boundary conditions for the displacement $\boldsymbol{u}$ and the fluid pressure p_F. However, similar result can be obtained for more general boundary conditions. For this, we first review possible boundary conditions for Biot's model. Suppose that there are two partitions of $\partial\Omega$,

$$(3.21) \qquad\qquad \partial\Omega = \Gamma_p \cup \Gamma_f, \qquad \partial\Omega = \Gamma_d \cup \Gamma_t,$$

with $|\Gamma_p|, |\Gamma_d| > 0$, i.e., the Lebesgue measures of Γ_p and Γ_d are positive. General homogeneous boundary conditions of (1.1) are given by

$$p_F(t) = 0 \quad \text{on } \Gamma_p, \quad -\kappa\nabla p_F(t) \cdot \boldsymbol{n} = 0 \quad \text{on } \Gamma_f,$$
$$\boldsymbol{u}(t) = 0 \quad \text{on } \Gamma_d, \quad \underline{\boldsymbol{\sigma}}(t)\boldsymbol{n} = 0 \quad \text{on } \Gamma_t,$$

for time variable $t \in [0, T]$, $T > 0$, in which $\boldsymbol{n}$ is the outward unit normal vector field on $\partial\Omega$ and $\underline{\boldsymbol{\sigma}}(t) := 2\mu\underline{\boldsymbol{\epsilon}}(\boldsymbol{u}(t)) + (\lambda\operatorname{div}\boldsymbol{u}(t) - \alpha p_F(t))\underline{\boldsymbol{I}}$. The conditions for p_F is a combination of pressure-flux boundary condition as in Darcy flow and the conditions for $\boldsymbol{u}$ is a combination of displacement-traction boundary conditions as in elasticity problems. The proper function spaces for this variational formulation are

$$(3.22) \quad \boldsymbol{V} := \{\boldsymbol{v} \in \boldsymbol{H}^1 : \boldsymbol{v}|_{\Gamma_d} = 0\}, \quad Q_T = L^2, \quad Q_F := \{q_F \in H^1 : q_F|_{\Gamma_p} = 0\},$$

for $\boldsymbol{u}$, p_T, and p_F. When $\Gamma_d \neq \partial\Omega$, we choose parameter-dependent norms by

$$(3.23) \qquad \|\underline{\boldsymbol{\epsilon}}(\boldsymbol{u})\|_0, \quad \|p_T\|_0, \quad \left((\alpha^2\lambda^{-1}p_F, p_F) + (\kappa\nabla p_F, \nabla p_F)\right)^{\frac{1}{2}}.$$

We can prove a stability result similar to Theorem 3.2 with these norms. In fact, the proof is easier in this case since the inf-sup condition

$$(3.24) \qquad\qquad \inf_{p_T \in L^2} \sup_{\boldsymbol{v} \in \boldsymbol{H}^1_{\Gamma_d}} \frac{(\operatorname{div}\boldsymbol{v}, p_T)}{\|\boldsymbol{v}\|_1\|p_T\|_0} \geq \beta_0$$

holds, and therefore a decomposition of p_T, into its mean-value and mean-value zero components, is not necessary. We omit the details since the proof is completely analogous to the proof of Theorem 3.2.

4. Discretization and construction of preconditioners

In this section we propose finite element discretizations of the three field formulation introduced above, and show that parameter-robust preconditioners for the discrete problems can be found. In contrast to the discussion above, we will first consider problems with general boundary conditions, i.e., boundary conditions with $\Gamma_d \neq \partial\Omega$, cf. (3.21). The reason for this reversed order is that the construction of preconditioners in the case when $\Gamma_d = \partial\Omega$, resulting in the choice $\boldsymbol{V} = \boldsymbol{H}_0^1$, requires a nontrivial technical discussion.

We have shown above that (3.13) is a linear system with parameter-robust stability for the function spaces $\boldsymbol{V}$, Q_T, Q_F with parameter-dependent norms given by (3.14). If we discretize the system (3.13) with the finite element spaces $\boldsymbol{V}_h \subset \boldsymbol{V}$, $Q_{T,h} \subset Q_T$, $Q_{F,h} \subset Q_F$, then the discrete counterpart of (3.13) is to find $(\boldsymbol{u}_h, p_{T,h}, p_{F,h}) \in \boldsymbol{V}_h \times Q_{T,h} \times Q_{F,h}$ such that

$$(\underline{\epsilon}(\boldsymbol{u}_h), \underline{\epsilon}(\boldsymbol{v})) - (\operatorname{div}\boldsymbol{v}, p_{T,h}) = (\boldsymbol{f}, \boldsymbol{v}), \qquad \forall \boldsymbol{v} \in \boldsymbol{V}_h,$$

$$(4.1) \quad -(\operatorname{div}\boldsymbol{u}_h, q_T) - (\lambda^{-1} p_{T,h}, q_T) + (\alpha\lambda^{-1} p_{F,h}, q_T) = 0, \qquad \forall q_T \in Q_{T,h},$$

$$(\alpha\lambda^{-1} p_{T,h}, q_F) - 2(\alpha^2\lambda^{-1} p_{F,h}, q_F) - (\kappa\nabla p_{F,h}, \nabla q_F) = (g, q_F), \quad \forall q_F \in Q_{F,h}.$$

A basic stability assumption for this discretization is that the pair $\boldsymbol{V}_h \times Q_{T,h}$ satisfies a discrete version of (3.24), i.e.,

$$(4.2) \qquad \inf_{p_T \in Q_{T,h}} \sup_{\boldsymbol{v} \in \boldsymbol{V}_h} \frac{(\operatorname{div}\boldsymbol{v}, p_T)}{\|\boldsymbol{v}\|_1 \|p_T\|_0} \geq \beta_0 > 0,$$

where β_0 is independent of h. In other words, $\boldsymbol{V}_h \times Q_{T,h}$ is a stable Stokes pair.

Theorem 4.1. *Suppose that $\boldsymbol{V}$, Q_T, Q_F are as in (3.22) with $\Gamma_d \neq \partial\Omega$ and Γ_p as in (3.21), and that $\boldsymbol{V}_h \subset \boldsymbol{V}$, $Q_{T,h} \subset Q_T$, $Q_{F,h} \subset Q_F$ are corresponding finite element spaces. Furthermore, assume that the pair $\boldsymbol{V}_h \times Q_{T,h}$ satisfies the inf-sup condition (4.2). Let $\mathcal{X}_h = \boldsymbol{V}_h \times Q_{T,h} \times Q_{F,h}$ be the Hilbert space with norm given in (3.23), and $\mathcal{A}_h : \mathcal{X}_h \to \mathcal{X}_h^*$ the operator given by (4.1). Then there exists a constant $\beta > 0$, such that*

$$(4.3) \qquad \inf_{(\boldsymbol{u}, p_T, p_F) \in \mathcal{X}_h} \sup_{(\boldsymbol{v}, q_T, q_F) \in \mathcal{X}_h} \frac{(\mathcal{A}_h(\boldsymbol{u}, p_T, p_F), (\boldsymbol{v}, q_T, q_F))_{(\mathcal{X}_h^*, \mathcal{X}_h)}}{\|(\boldsymbol{u}, p_T, p_F)\|_{\mathcal{X}_h} \|(\boldsymbol{v}, q_T, q_F)\|_{\mathcal{X}_h}} \geq \beta,$$

for all parameters λ, α, and κ satisfying (3.2).

We do not prove this result here since the proof is completely analogous to the proof of Theorem 3.2. We observe that the norms given in (3.14) shows that a preconditioner of the form

$$(4.4) \qquad \mathcal{B} = \begin{pmatrix} -\Delta^{-1} & 0 & 0 \\ 0 & I^{-1} & 0 \\ 0 & 0 & \left(\alpha^2\lambda^{-1}I - \operatorname{div}(\kappa\nabla)\right)^{-1} \end{pmatrix}$$

will be a parameter-robust preconditioner.

We now turn to the case with $\Gamma_d = \partial\Omega$ such that $\boldsymbol{V} = \boldsymbol{H}_0^1$. We recall that L_0^2 is the space of L^2 functions with mean value zero. The proper discrete inf-sup condition in this case takes the form

$$(4.5) \qquad \inf_{p_T \in Q_{T,h} \cap L_0^2} \sup_{\boldsymbol{v} \in \boldsymbol{V}_h} \frac{(\operatorname{div}\boldsymbol{v}, p_T)}{\|\boldsymbol{v}\|_1 \|p_T\|_0} \geq \beta_0 > 0,$$

where again β_0 is independent of h.

The following is a discrete analogue of Theorem 3.2, and its proof is completely analogous to the proof of that theorem.

Theorem 4.2. *Suppose that V, Q_T, Q_F are as in (3.22) with $\Gamma_d = \partial\Omega$ and Γ_p as in (3.21), and that $V_h \subset V$, $Q_{T,h} \subset Q_T$, $Q_{F,h} \subset Q_F$ are corresponding finite element spaces. Furthermore, assume that the pair $V_h \times Q_{T,h}$ satisfy the inf-sup condition (4.5). Let $\mathcal{X}_h = V_h \times Q_{T,h} \times Q_{F,h}$ be the Hilbert space with norm given in (3.14), and $\mathcal{A}_h : \mathcal{X}_h \to \mathcal{X}_h^*$ the operator given by (4.1). Then there is a constant $\beta > 0$ such that (4.3) holds for all parameters λ, α, and κ satisfying (3.2).*

There exist a number of choices of stable Stokes pairs $V_h \times Q_{T,h}$, and in Section 6 below we will present numerical results for two examples, the lowest order Taylor-Hood element and the MINI element. For more examples of stable Stokes pairs we refer to [29, 39]. The parameter-dependent norms in (3.14) suggest a block diagonal preconditioner of the form

$$(4.6) \qquad \mathcal{B} = \begin{pmatrix} -\Delta^{-1} & 0 & 0 \\ 0 & \left(\lambda^{-1}I + I_0\right)^{-1} & 0 \\ 0 & 0 & \left(\alpha^2\lambda^{-1}I - \operatorname{div}(\kappa\nabla)\right)^{-1} \end{pmatrix},$$

for the continuous system. We recall that I is the Riesz map of Q_T into its dual Q_T^*, and I_0 is the corresponding map into the dual of $Q_T \cap L_0^2$. The first and third blocks of this block diagonal operator are inverses of standard second-order elliptic operators, and corresponding preconditioners to replace the exact inverses in the discrete case are well-studied. In contrast, the operator in the second block is less standard, and, as far as we know, a construction of an effective preconditioner to replace it has not been proposed. We will discuss such a construction below.

5. A preconditioner for the operator $\lambda^{-1}I + I_0$

Throughout this section the parameter λ is assumed to be a constant. We recall from the discussion above that in order to construct an effective block diagonal preconditioner of the form (4.6) we need to replace the inverse of the operator $\lambda^{-1}I + I_0$ by a spectrally equivalent operator which can be cheaply evaluated. In fact, when $\lambda \geq 1$ the operators $\lambda^{-1}I + I_0$ and $\lambda^{-1}I_m + I_0$ are spectrally equivalent, so it is enough to approximate the inverse of the latter.

Let N be the dimension of $Q_{T,h}$ and $\{\phi^i\}_{i=1}^N$ be the standard nodal basis of $Q_{T,h}$. Let 1_Ω be the constant function on Ω with value $1/\sqrt{|\Omega|}$ where $|\Omega|$ is the volume of Ω. Denoting the mean-value zero and mean-value parts of ϕ^i by ϕ_0^i and ϕ_m^i as before, we can observe that

$$(5.1) \qquad \phi_m^i = m_i 1_\Omega, \quad m_i := (\phi^i, 1_\Omega), \qquad \phi^i = \phi_0^i + \phi_m^i, \qquad \forall 1 \leq i \leq N.$$

If we let $\mathbb{M}$, $\mathbb{M}_0$, $\mathbb{M}_m$ be mass matrices corresponding to the operators I, I_0, and I_m, then their (i, j)-entries are

$$\mathbb{M}(i,j) := (\phi^i, \phi^j), \quad \mathbb{M}_0(i,j) := (\phi_0^i, \phi_0^j), \quad \mathbb{M}_m(i,j) := (\phi_m^i, \phi_m^j), \quad \forall 1 \leq i, j \leq N,$$

and the matrix corresponding to $\lambda^{-1}I_m + I_0$ is $\lambda^{-1}\mathbb{M}_m + \mathbb{M}_0$. Since $(\phi^i, \phi^j) = (\phi_0^i + \phi_m^i, \phi_0^j + \phi_m^j) = (\phi_0^i, \phi_0^j) + (\phi_m^i, \phi_m^j)$, one can see $\mathbb{M} = \mathbb{M}_0 + \mathbb{M}_m$, and therefore

$\lambda^{-1}\mathbb{M}_m + \mathbb{M}_0 = \mathbb{M} + (\lambda^{-1} - 1)\mathbb{M}_m$. In addition, observe that $\mathbb{M}_m(i,j) = m_i m_j$ by (5.1), so

$$(5.2) \qquad \lambda^{-1}\mathbb{M}_m + \mathbb{M}_0 = \mathbb{M} + (\lambda^{-1} - 1)\mathbf{m}\mathbf{m}^T,$$

with

$$(5.3) \qquad \mathbf{m} = \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_N \end{pmatrix}.$$

To construct a preconditioner we need to find an approximate inverse of $\lambda^{-1}\mathbb{M}_m + \mathbb{M}_0$. Since $\mathbb{M}$ is positive definite, we can rewrite the right-hand side of (5.2) as

$$(5.4) \qquad \mathbb{M}_\lambda := \mathbb{M} + (\lambda^{-1} - 1)\mathbf{m}\mathbf{m}^T = (\mathbb{I} + (\lambda^{-1} - 1)\mathbf{m}\mathbf{m}^T\mathbb{M}^{-1})\mathbb{M},$$

where $\mathbb{I}$ is the $N \times N$ identity matrix. Recall the Sherman–Morrison–Woodbury formula,

$$(5.5) \qquad (\mathbb{I} + \mathbf{u}\mathbf{v}^T)^{-1} = \mathbb{I} - \frac{\mathbf{u}\mathbf{v}^T}{1 + \mathbf{u}^T\mathbf{v}}, \qquad \mathbf{u}, \mathbf{v} \in \mathbb{R}^N \text{ with } \mathbf{u}^T\mathbf{v} \neq -1.$$

We will use it to find the inverse of $\mathbb{I} + (c^{-1} - 1)\mathbf{m}\mathbf{m}^T\mathbb{M}^{-1}$ for a constant $c \neq 0$.

Lemma 5.1. *Let $\mathbf{w} = (1 \cdots 1)^T \in \mathbb{R}^N$. For the mass matrix $\mathbb{M}$ and $\mathbf{m}$ in (5.3), the following two identities holds:*

$$(5.6) \qquad \mathbb{M}\mathbf{w} = \sqrt{|\Omega|}\mathbf{m}, \qquad \mathbf{m}^T\mathbf{w} = \sqrt{|\Omega|}.$$

Proof. Note that

$$(5.7) \qquad \frac{1}{\sqrt{|\Omega|}} \sum_{j=1}^N \phi^j = 1_\Omega,$$

because $\{\phi^j\}_{1 \leq j \leq N}$ is the standard nodal basis and no boundary condition is imposed on $Q_{T,h}$.

If we consider the i-th row of the left-hand side of the first identity in (5.6), then the definition of $\mathbb{M}$, (5.7), and (5.1) give

$$\sum_{j=1}^N \mathbb{M}(i,j) = \sum_{j=1}^N (\phi^i, \phi^j) = \left(\phi^i, \sum_{j=1}^N \phi^j\right) = \sqrt{|\Omega|}m_i.$$

This proves the first identity in (5.6). The second identity follows by

$$\sum_{i=1}^N m_i = \sum_{i=1}^N (\phi^i, 1_\Omega) = \sqrt{|\Omega|}(1_\Omega, 1_\Omega) = \sqrt{|\Omega|}.$$

$\square$

Corollary 5.2. *For the mass matrix $\mathbb{M}$, $\mathbf{m}$ in (5.3), $\mathbf{w}$ in Lemma 5.1 and any constant $c \neq 0$, the following holds:*

$$(5.8) \qquad (\mathbb{I} + (c^{-1} - 1)\mathbf{m}\mathbf{m}^T\mathbb{M}^{-1})^{-1} = \mathbb{I} + (c - 1)(\sqrt{|\Omega|})^{-1}\mathbf{m}\mathbf{w}^T.$$

Proof. Since $\mathbb{M}$ is symmetric, the first identity in (5.6) gives $\mathbf{m}^T\mathbb{M}^{-1} = (\sqrt{|\Omega|})^{-1}\mathbf{w}^T$. If we set $\mathbf{u} = (c^{-1} - 1)\mathbf{m}$ and $\mathbf{v}^T = \mathbf{m}^T\mathbb{M}^{-1}$, then the second identity in (5.6) gives $\mathbf{u}^T\mathbf{v} = c^{-1} - 1 \neq -1$. The assertion follows from (5.5). $\square$

Theorem 5.3. *Let* $\mathbb{V}_\lambda = (\mathbb{I} + a\mathbf{m}\mathbf{w}^T)^{-1}$ *with* $a = (-1 + \sqrt{\lambda})/\sqrt{|\Omega|}$*. Then, for* $\mathbb{M}_\lambda$ *in* (5.4), $\mathbb{M}_\lambda = \mathbb{V}_\lambda \mathbb{M} \mathbb{V}_\lambda^T$*. Thus, if* $\mathbb{D}$ *is a preconditioner of* $\mathbb{M}$ *with condition number* K*, then* $\mathbb{V}_\lambda^{-T} \mathbb{D} \mathbb{V}_\lambda^{-1}$ *is a preconditioner of* $\mathbb{M}_\lambda$ *with same condition number.*

Proof. From the definition of $\mathbb{V}_\lambda$ and the second identity in (5.6), we can see

$$\mathbb{V}_\lambda^{-2} = (\mathbb{I} + a\mathbf{m}\mathbf{w}^T)^2 = \mathbb{I} + (2a + a^2\sqrt{|\Omega|})\mathbf{m}\mathbf{w}^T = \mathbb{I} + (\lambda - 1)(\sqrt{|\Omega|})^{-1}\mathbf{m}\mathbf{w}^T.$$

By the identity (5.8) with $c = \lambda$ and (5.4), we have $\mathbb{M}_\lambda = \mathbb{V}_\lambda^2 \mathbb{M}$. If we use (5.8) for $\mathbb{V}_\lambda^{-1}$, then one can verify that $\mathbb{V}_\lambda = \mathbb{I} + \bar{a}\mathbf{m}\mathbf{m}^T\mathbb{M}^{-1}$ with $\bar{a} = 1 - (\sqrt{\lambda})^{-1}$. From this expression of $\mathbb{V}_\lambda$, it is easy to check that $\mathbb{V}_\lambda\mathbb{M} = \mathbb{M}\mathbb{V}_\lambda^T$, so $\mathbb{M}_\lambda = \mathbb{V}_\lambda^2 \mathbb{M} = \mathbb{V}_\lambda\mathbb{M}\mathbb{V}_\lambda^T$. The assertion for preconditioner $\mathbb{V}_\lambda^{-T} \mathbb{D} \mathbb{V}_\lambda^{-1}$ follows from the identity $\mathbb{M}_\lambda \mathbb{V}_\lambda^{-T} \mathbb{D} \mathbb{V}_\lambda^{-1} \mathbb{M}_\lambda = \mathbb{V}_\lambda\mathbb{M}\mathbb{D}\mathbb{M}\mathbb{V}_\lambda^T$. □

For the preconditioner $\mathbb{D}$ for $\mathbb{M}$, it is known that the Jacobi preconditioner, i.e., the inverse of diagonal of mass matrix as a preconditioner has explicit condition number bounds [40]. If $Q_{T,h}$ is the piecewise linear continuous finite element, then the Jacobi preconditioner $\mathbb{D}$ for $\mathbb{M}$ is a constant multiple of the diagonal matrix $\mathrm{diag}(m_1^{-1}, m_2^{-1}, ..., m_N^{-1})$. As a consequence, $\mathbf{w}\mathbf{m}^T\mathbb{D} = \mathbb{D}\mathbf{m}\mathbf{w}^T$, so $\mathbb{V}_\lambda^{-T} \mathbb{D} \mathbb{V}_\lambda^{-1}$ can be reduced to $\mathbb{D}\mathbb{V}_\lambda^{-2} = \mathbb{D}(\mathbb{I} + (\lambda - 1)(\sqrt{|\Omega|})^{-1}\mathbf{m}\mathbf{w}^T)$.

There is one caution in the implementation of the preconditioner $\mathbb{V}_\lambda^{-T} \mathbb{D} \mathbb{V}_\lambda^{-1}$ because $\mathbf{m}\mathbf{w}^T$ and $\mathbf{m}\mathbf{m}^T$ in $\mathbb{V}_\lambda^{-1}$ and $\mathbb{M}_\lambda$ are dense matrices in general. We remark that the minimum residual method requires only matrix-vector multiplication operations. Therefore, to avoid computation with these dense matrices, we use the structure of the matrix $\mathbf{m}\mathbf{w}^T$. More precisely, $\mathbf{m}\mathbf{w}^T\mathbf{v}$ for an $\mathbb{R}^N$-vector $\mathbf{v}$ can be computed with two operations, the inner product $\mathbf{w}^T\mathbf{v}$ and constant-vector multiplication $(\mathbf{w}^T\mathbf{v})\mathbf{m}$. Similarly, we can avoid generating $\mathbf{m}\mathbf{m}^T$ in $\mathbb{M}_\lambda = \mathbb{M} + (\lambda^{-1} - 1)\mathbf{m}\mathbf{m}^T$. Finally, we remark that the preconditioner $\mathbb{V}_\lambda^{-T} \mathbb{D} \mathbb{V}_\lambda^{-1}$ is useful when a piecewise discontinuous finite element is used for $Q_{T,h}$ because $\mathbf{m}\mathbf{w}^T$ and $\mathbb{M}_\lambda$ are not sparse.

TABLE 5. Boundary conditions (BC), preconditioners (PC), and finite elements of test cases. The first three cases use the lowest order Taylor–Hood element and the last case uses the MINI element ($\boldsymbol{B}$ = vector-valued bubble function).

	BC	PC	finite elements	system size for N		
				32	64	128
Case 1	$\Gamma_d \neq \partial\Omega,\ \Gamma_p = \partial\Omega$	(4.4)	$\boldsymbol{\mathcal{P}}_2$-$\mathcal{P}_1$-$\mathcal{P}_1$	10628	41732	165380
Case 2	$\Gamma_d = \partial\Omega,\ \Gamma_p = \partial\Omega$	(4.6)	$\boldsymbol{\mathcal{P}}_2$-$\mathcal{P}_1$-$\mathcal{P}_1$	10628	41732	165380
Case 3	$\Gamma_d = \partial\Omega,\ \Gamma_p = \partial\Omega$	(4.4)	$\boldsymbol{\mathcal{P}}_2$-$\mathcal{P}_1$-$\mathcal{P}_1$	10628	41732	165380
Case 4	$\Gamma_d \neq \partial\Omega,\ \Gamma_p = \partial\Omega$	(4.4)	$(\boldsymbol{\mathcal{P}}_1 + \boldsymbol{B})$-$\mathcal{P}_1$-$\mathcal{P}_1$	8452	33284	132100

6. Numerical results

In this section we present some numerical results which illustrate our theoretical results for the proposed preconditioners (4.4) and (4.6). As before, all numerical experiments are carried out using FEniCS with Hypre algebraic multigrid operators as replacements for the exact inverses appearing in the first and third block of (4.4) and (4.6). Furthermore, in the preconditioner of the form (4.6), the second block is

TABLE 6. Numbers of iteration and condition numbers of Case 1 (cf. Table 5). ($\Omega =$ unit square, partitioned as bisections of $N \times N$ rectangles, convergence criterion with relative residual of 10^{-6})

N	α	λ	κ			
			10^0	10^{-4}	10^{-8}	10^{-12}
32	10^0	10^0	33 (3.8)	43 (6.3)	47 (7.6)	47 (7.6)
		10^4	52 (21.7)	52 (21.7)	65 (21.7)	63 (21.7)
		10^8	52 (21.7)	54 (21.7)	52 (21.7)	62 (21.7)
	10^{-2}	10^0	33 (3.8)	33 (3.9)	43 (6.3)	47 (7.6)
		10^4	52 (21.7)	52 (21.7)	52 (21.7)	63 (21.7)
		10^8	52 (21.7)	52 (21.7)	52 (21.7)	52 (21.7)
	10^{-4}	10^0	33 (3.8)	33 (3.8)	33 (3.8)	43 (6.3)
		10^4	50 (21.7)	52 (21.7)	50 (21.7)	52 (21.7)
		10^8	54 (21.7)	52 (21.7)	52 (21.7)	52 (21.7)
64	10^0	10^0	33 (3.9)	40 (5.6)	47 (7.6)	47 (7.6)
		10^4	52 (21.7)	52 (21.7)	63 (21.7)	63 (21.7)
		10^8	46 (21.7)	52 (21.7)	52 (21.7)	62 (21.7)
	10^{-2}	10^0	33 (3.8)	33 (3.8)	40 (5.6)	47 (7.6)
		10^4	52 (21.7)	52 (21.7)	52 (21.7)	58 (21.7)
		10^8	52 (21.7)	46 (21.7)	52 (21.7)	48 (21.7)
	10^{-4}	10^0	33 (3.9)	33 (3.8)	33 (3.8)	40 (5.6)
		10^4	50 (21.7)	46 (21.7)	50 (21.7)	52 (21.7)
		10^8	52 (21.7)	52 (21.7)	50 (21.7)	52 (21.7)
128	10^0	10^0	32 (3.8)	39 (5.4)	46 (7.7)	45 (7.7)
		10^4	51 (21.7)	52 (21.7)	61 (21.7)	59 (21.7)
		10^8	51 (21.7)	50 (21.7)	50 (21.7)	54 (21.7)
	10^{-2}	10^0	33 (3.8)	33 (3.8)	39 (5.2)	46 (7.7)
		10^4	48 (21.7)	45 (21.7)	52 (21.7)	58 (21.7)
		10^8	50 (21.7)	52 (21.7)	50 (21.7)	52 (21.7)
	10^{-4}	10^0	33 (3.8)	32 (3.8)	33 (3.8)	39 (5.2)
		10^4	44 (21.7)	52 (21.7)	52 (21.7)	52 (21.7)
		10^8	50 (21.7)	48 (21.7)	48 (21.7)	50 (21.7)

constructed by using the technique outlined in Section 5, while the standard Jacobi preconditioner is used in the second block of (4.4).

In all the experiments the domain Ω is unit square. We show numerical results for four different combinations of boundary conditions, finite element spaces, and preconditioners. The different combinations are presented as Case 1-4 in Table 5. In Case 1 and 4 the statement $\Gamma_d \neq \partial\Omega$ means that Γ_d is taken as in Example 3.1, while problems with $\Gamma_d = \partial\Omega$ are consider in Case 2 and Case 3. We compare numerical results obtained by the two preconditioners with structure of the form (4.6) and (4.4). The result of Theorem 4.2 suggests that preconditioners of the form (4.6) are more robust than the ones of the form (4.4) in the case of Dirichlet boundary conditions. In other words, we expect that the results of Case 2 are more robust than the ones of Case 3. However, the system preconditioned with a preconditioner of the form (4.4) has only one bad eigenvalue. Therefore, as in Example 2.2, we can expect small differences in the number of iterations. Finally,

TABLE 7. Numbers of iteration of test cases in Table 5. ($\Omega =$ unit square, partitioned into bisections of $N \times N$ rectangles, convergence criterion with relative residual of 10^{-6})

			Case 1			Case 2			Case 3			Case 4		
			N			N			N			N		
κ	α	λ	32	64	128	32	64	128	32	64	128	32	64	128
10^0	10^0	10^0	33	33	32	29	29	29	29	29	29	34	34	34
		10^4	52	52	51	46	46	46	66	45	44	60	61	60
		10^8	52	46	51	46	46	45	44	44	43	60	61	60
	10^{-4}	10^0	33	33	33	29	29	29	29	29	29	34	34	34
		10^4	50	50	44	46	46	44	66	44	44	60	60	60
		10^8	54	52	50	46	46	45	45	44	43	60	60	60
10^{-4}	10^0	10^0	43	40	39	39	38	36	39	38	36	46	43	40
		10^4	52	52	52	46	46	45	68	44	44	60	61	60
		10^8	54	52	50	46	46	45	66	44	43	60	60	60
	10^{-4}	10^0	33	33	32	29	29	29	29	29	29	34	34	34
		10^4	52	46	52	46	46	45	66	44	42	60	60	60
		10^8	52	52	48	46	46	45	44	44	43	60	61	60
10^{-8}	10^0	10^0	47	47	46	42	42	42	42	42	42	52	52	52
		10^4	65	63	61	61	59	58	60	58	57	73	73	72
		10^8	52	52	50	46	46	45	44	44	43	60	60	60
	10^{-4}	10^0	47	33	33	29	29	29	29	29	29	34	34	34
		10^4	65	50	52	46	45	45	66	44	43	60	60	60
		10^8	52	50	48	46	46	44	44	44	40	60	60	60
10^{-12}	10^0	10^0	47	47	45	42	42	42	42	42	42	52	52	52
		10^4	63	63	59	58	58	57	57	56	56	72	72	72
		10^8	62	62	54	58	58	50	57	56	54	72	72	71
	10^{-4}	10^0	43	40	39	39	38	36	37	38	36	47	43	40
		10^4	52	52	52	46	46	44	66	44	44	60	61	60
		10^8	52	52	50	46	45	44	44	43	43	60	61	60

in Case 4 we use the MINI element instead of the Taylor–Hood element in order to show that our results are robust with respect to the choice of finite element spaces, as long as they fulfill the assumptions of the theory. In most of the examples the parameters λ, α and κ are taken to be constants. However, in the last experiment, presented in Table 8, κ varies with the spatial variable.

In Table 6, we present numbers of iteration of Case 1. The results are fairly robust with respect to parameter changes and mesh refinements. To compare robustness of preconditioners of all the cases, we present numbers of iteration for all the different four cases in Table 7. As expected, the results of Case 2 are slightly better than the ones of Case 3, in particular for $N = 32$. Although there are no remarkable differences in the presented results, in the full numerical results which are not included here, the results for Case 3 shows that this method sometimes need about $30 - 45\%$ more iterations than those of Case 2. In Case 4 we use the MINI element instead of the Taylor-Hood element. Although the numbers of iteration are larger than those for the Taylor–Hood element, the results are still quite robust with respect to changes of parameters. As the final experiment, a model problem

TABLE 8. Numbers of iteration and condition numbers of Case 1 (cf. Table 5) with nonconstant κ. (Ω = unit square, partitioned as bisections of $N \times N$ rectangles, $\Omega_1 = \{(x, y) : 0 \le x \le 1, 1/4 \le y \le 3/4\}$, $\kappa = 1$ on $\Omega \setminus \Omega_1$, convergence criterion with relative residual of 10^{-6})

N	α	λ	κ on Ω_1				
			10^{-2}	10^{-4}	10^{-6}	10^{-8}	10^{-10}
32	10^0	10^0	34 (4.5)	42 (6.1)	46 (7.4)	46 (7.5)	46 (7.4)
		10^4	54 (21.7)	52 (21.7)	56 (21.7)	65 (21.7)	61 (21.7)
		10^8	54 (21.7)	54 (21.7)	46 (21.7)	52 (21.7)	55 (21.7)
	10^{-2}	10^0	33 (3.8)	33 (3.8)	34 (4.4)	42 (6.1)	46 (7.4)
		10^4	52 (21.7)	48 (21.7)	50 (21.7)	53 (21.7)	56 (21.7)
		10^8	52 (21.7)	52 (21.7)	50 (21.7)	52 (21.7)	46 (21.7)
	10^{-4}	10^0	33 (3.8)	33 (3.8)	33 (3.8)	33 (3.8)	34 (4.5)
		10^4	52 (21.7)	52 (21.7)	52 (21.7)	52 (21.7)	52 (21.7)
		10^8	52 (21.7)	52 (21.7)	50 (21.7)	53 (21.7)	50 (21.7)
64	10^0	10^0	34 (4.5)	39 (5.4)	44 (7.4)	46 (7.5)	46 (7.5)
		10^4	52 (21.7)	52 (21.7)	55 (21.7)	63 (21.7)	60 (21.7)
		10^8	52 (21.7)	46 (21.7)	46 (21.7)	50 (21.7)	51 (21.7)
	10^{-2}	10^0	33 (3.8)	33 (3.8)	34 (4.4)	39 (5.4)	45 (7.4)
		10^4	46 (21.7)	52 (21.7)	52 (21.7)	52 (21.7)	51 (21.7)
		10^8	48 (21.7)	52 (21.7)	52 (21.7)	52 (21.7)	52 (21.7)
	10^{-4}	10^0	33 (3.8)	33 (3.8)	33 (3.8)	33 (3.8)	34 (4.4)
		10^4	52 (21.7)	52 (21.7)	52 (21.7)	52 (21.7)	52 (21.7)
		10^8	52 (21.7)	52 (21.7)	52 (21.7)	52 (21.7)	46 (21.7)
128	10^0	10^0	34 (4.4)	37 (5.0)	44 (7.2)	46 (7.5)	44 (7.5)
		10^4	50 (21.7)	48 (21.7)	49 (21.7)	56 (21.7)	61 (21.7)
		10^8	50 (21.7)	50 (21.7)	52 (21.7)	50 (21.7)	47 (21.7)
	10^{-2}	10^0	33 (3.8)	32 (3.9)	34 (4.4)	37 (5.0)	44 (7.1)
		10^4	50 (21.7)	46 (21.7)	50 (21.7)	52 (21.7)	49 (21.7)
		10^8	52 (21.7)	52 (21.7)	50 (21.7)	52 (21.7)	50 (21.7)
	10^{-4}	10^0	32 (3.8)	33 (3.9)	33 (3.8)	33 (3.9)	34 (4.4)
		10^4	52 (21.7)	50 (21.7)	52 (21.7)	49 (21.7)	46 (21.7)
		10^8	52 (21.7)	50 (21.7)	48 (21.7)	52 (21.7)	50 (21.7)

with nonconstant κ is considered. We assume that κ is small on

$$\Omega_1 = \{(x, y) : 0 \le x \le 1, 1/4 \le y \le 3/4\} \subset \Omega,$$

and $\kappa = 1$ on $\Omega \setminus \Omega_1$. The numerical results in Table 8 are fairly robust for mesh refinements and changes of parameters, including high contrasts of κ.

7. CONCLUSION

We have studied parameter-robust discretizations and construction of preconditioners for Biot's consolidation model. To apply the framework of [25] we have proposed a new three-field formulation of the Biot system. We have showed that preconditioners based on mapping properties and parameter-dependent norms are robust with respect to variations of the model parameters, choice of finite element spaces satisfying the proper stability condition, and the discretization parameters.

In particular, the variations of parameters in our consideration cover large shear and bulk elastic moduli, small hydraulic conductivity, small time-step, including the ranges of interest in geophysics and computational biomechanics applications. Furthermore, our theoretical results are confirmed by a number of numerical experiments.

References

[1] R. E. Showalter. Diffusion in poro-elastic media. *J. Math. Anal. Appl.*, 251(1):310–340, 2000.

[2] M. B. Reed. An investigation of numerical errors in the analysis of consolidation by finite elements. *Internat. J. Numer. Analyt. Methods Geomech.*, 8(3):243–257, 1984.

[3] P. A. Vermeer and A. Verruijt. An accuracy condition for consolidation by finite elements. *Internat. J. Numer. Analyt. Methods Geomech.*, 5(1):1–14, 1981.

[4] O. C. Zienkiewicz and T. Shiomi. Dynamic behaviour of saturated porous media; the generalized Biot formulation and its numerical solution. *Internat. J. Numer. Analyt. Methods Geomech.*, 8(1):71–96, 1984.

[5] M. A. Murad and A. F. D. Loula. Improved accuracy in finite element analysis of Biot's consolidation problem. *Comput. Methods Appl. Mech. Engrg.*, 95(3):359–382, 1992.

[6] M. A. Murad and A. F. D. Loula. On stability and convergence of finite element approximations of Biot's consolidation problem. *Internat. J. Numer. Methods Engrg.*, 37(4):645–667, 1994.

[7] M. A. Murad, V. Thomée, and A. F. D. Loula. Asymptotic behavior of semidiscrete finite-element approximations of Biot's consolidation problem. *SIAM J. Numer. Anal.*, 33(3):1065–1083, 1996.

[8] J. Korsawe and G. Starke. A least-squares mixed finite element method for Biot's consolidation problem in porous media. *SIAM J. Numer. Anal.*, 43(1):318–339, 2005.

[9] Y. Chen, Y. Luo, and M. Feng. Analysis of a discontinuous Galerkin method for the Biot's consolidation problem. *Appl. Math. Comput.*, 219(17):9043–9056, 2013.

[10] P. J. Phillips and M. F. Wheeler. A coupling of mixed and continuous Galerkin finite element methods for poroelasticity. I. The continuous in time case. *Comput. Geosci.*, 11(2):131–144, 2007.

[11] P. J. Phillips and M. F. Wheeler. A coupling of mixed and continuous Galerkin finite element methods for poroelasticity. II. The discrete-in-time case. *Comput. Geosci.*, 11(2):145–158, 2007.

[12] P. J. Phillips and M. F. Wheeler. A coupling of mixed and discontinuous Galerkin finite-element methods for poroelasticity. *Comput. Geosci.*, 12(4):417–435, 2008.

[13] S.-Y. Yi. A coupling of nonconforming and mixed finite element methods for Biot's consolidation model. *Numer. Methods Partial Differ. Equ.*, 29(5):1749–1777, 2013.

[14] F. Ben-Hatira, K. Saidane, and A. Mrabet. A finite element modeling of the human lumbar unit including the spinal cord. *J. Biomed. Sci. Eng.*, 5:146–152, 2012.

[15] J. H. Smith and J. A. Humphrey. Interstitial transport and transvascular fluid exchange during infusion into brain and tumor tissue. *Microvasc. Res.*, 73(1):58–73, 2007.

[16] K. Støverud, M. Alnæs, H. P. Langtangen, V. Haughton, and K.-A. Mardal. Poroelastic modeling of syringomyelia – a systematic study of the effects of pia mater, central canal, median fissure, white and grey matter on pressure wave propagation and fluid movement within the cervical spinal cord. *Comput. Methods Biomech. Biomed. Engin.* Accepted 2015.

[17] O. Coussy. *Poromechanics*. John Wiley & Sons, 2004.

[18] H. F. Wang. Theory of linear poroelasticity. *Princeton Series in Geophysics, Princeton University Press, Princeton, NJ*, 2000.

[19] O. Axelsson, R. Blaheta, and P. Byczanski. Stable discretization of poroelasticity problems and efficient preconditioners for arising saddle point type matrices. *Comput. Vis. Sci.*, 15(4):191–207, 2012.

[20] S.-H. Chan, K.-K. Phoon, and F. H. Lee. A modified Jacobi preconditioner for solving ill-conditioned Biot's consolidation equations using symmetric quasi-minimal residual method. *Internat. J. Numer. Analyt. Methods Geomech.*, 25(10):1001–1025, 2001.

[21] J. B. Haga, H. Osnes, and H. P. Langtangen. A parallel block preconditioner for large-scale poroelasticity with highly heterogeneous material parameters. *Comput. Geosci.*, 16(3):723–734, 2012.

[22] K. K. Phoon, K. C. Toh, S. H. Chan, and F. H. Lee. An efficient diagonal preconditioner for finite element solution of Biot's consolidation equations. *Int. J. Numer. Methods Eng.*, 55(4):377–400, 2002.

[23] S. Rhebergen, G. N. Wells, R. F. Katz, and A. J. Wathen. Analysis of block preconditioners for models of coupled magma/mantle dynamics. *SIAM J. Sci. Comput.*, 36(4):A1960–A1977, 2014.

[24] S. Rhebergen, G. N. Wells, A. J. Wathen, and R. F. Katz. Three-field block preconditioners for models of coupled magma/mantle dynamics. *SIAM J. Sci. Comput.*, 37(5):A2270–A2294, 2015.

[25] K.-A. Mardal and R. Winther. Preconditioning discretizations of systems of partial differential equations. *Numer. Linear Algebra Appl.*, 18(1):1–40, 2011.

[26] R. D. Falgout and U. M. Yang. *Computational Science — ICCS 2002: International Conference Amsterdam, The Netherlands, April 21–24, 2002 Proceedings, Part III*, chapter hypre: A Library of High Performance Preconditioners, pages 632–641. Springer Berlin Heidelberg, Berlin, Heidelberg, 2002.

[27] K.-A. Mardal and J. B. Haga. Block preconditioning of systems of PDEs. In *Automated Solution of Differential Equations by the Finite Element Method*, pages 643–655. Springer Berlin Heidelberg, 2012.

[28] A. Logg, K.-A. Mardal, and G. N. Wells, editors. *Automated solution of differential equations by the finite element method*, volume 84 of *Lecture Notes in Computational Science and Engineering*. Springer, Heidelberg, 2012.

[29] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*, volume 15 of *Springer Series in computational Mathematics*. Springer, 1992.

[30] O. Axelsson and G. Lindskog. On the rate of convergence of the preconditioned conjugate gradient method. *Numer. Math.*, 48(5):499–523, 1986.

[31] B. F. Nielsen and K.-A. Mardal. Analysis of the minimal residual method applied to ill posed optimality systems. *SIAM J. Sci. Comput.*, 35(2):A785–A814, 2013.

[32] A. Anandarajah. *Computational Methods in Elasticity and Plasticity: Solids and Porous Media*. Springer New York, 2010.

[33] I. Babuška and M. Suri. Locking effects in the finite element approximation of elasticity problems. *Numer. Math.*, 62(4):439–463, 1992.

[34] D. S. Malkus and T. J. Hughes. Mixed finite element methods reduced and selective integration techniques: A unification of concepts. *Comput. Methods Appl. Mech. Engrg.*, 15(1):63 – 81, 1978.

[35] D. Boffi, F. Brezzi, and M. Fortin. Finite elements for the Stokes problem. In D. Boffi and L. Gastaldi, editors, *Mixed Finite Elements: Compatibility Conditions*, volume 1939 of *Lecture Notes in Mathematics*. Springer, 2008.

[36] S. C. Brenner and L.-Y. Sung. Linear finite element methods for planar linear elasticity. *Math. Comp.*, 59(200):321–338, 1992.

[37] M. Crouzeix and P.-A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge*, 7(R-3):33–75, 1973.

[38] K.-A. Mardal, X.-C. Tai, and R. Winther. A robust finite element method for Darcy-Stokes flow. *SIAM J. Numer. Anal.*, 40(5):1605–1631, 2002.

[39] V. Girault and P.-A. Raviart. *Finite element methods for Navier-Stokes equations*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1986.

[40] A. J. Wathen. Realistic eigenvalue bounds for the Galerkin mass matrix. *IMA J. Numer. Anal.*, 7(4):449–457, 1987.