A Stability Condition for the Numerical Simulation of Poroelastic Systems

Marco Favino[1], Alfio Grillo[2], and Rolf Krause[1]

[1] Institute of Computational Science,
University of Lugano, via Buffi 13, 6900, Lugano, Switzerland;
PH +41 58 666 4971; email: {marco.favino,rolf.krause}@usi.ch

[2] Dipartimento di Scienze Matematiche,
Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Torino, Italy.
email: alfio.grillo@polito.it

ABSTRACT

In the numerical simulation of the Biot model numerical oscillations may occur if the spatial and discretization parameters are not properly chosen. When the time-step is too small the grid has to be fine enough to resolve the boundary layers that appear in the early stages of the simulation. In this work, we extend a strategy that has been successfully employed in the one-dimensional case for the detection of a critical time-step below which instabilities appear. The idea is to study under which condition the Schur complement of the poroelastic system enjoys a discrete maximum principle property. Differently from the one-dimensional case, in the vectorial case the shear modulus plays a fundamental role for the computation of the Schur complement and the maximum principle argument can be applied only for small shear. We will also investigate the sign of the entries of the Schur complement when the shear and the bulk modulus are of the same magnitude.

INTRODUCTION

In the field of multiphase continuum mechanics, the Biot theory [Bio41] describes the motion of a deformable porous medium saturated with a fluid. It was successfully employed for the modeling of soils and reservoirs in the field of geomechanics. Recently, it has been studied also in case of growth [GFW12] and it has been applied in bio-mechanics, for example for the modeling of cartilage [FGLR+05], teeth [FGD+11], and heart walls.

For a poroelastic material, both momentum and mass balances of the mixture have to be solved and they form a time-dependent system of Partial Differential Equations (PDEs) whose unknowns are the solid displacement and the pore pressure.

In order to numerically solve the Biot system, it has to be discretized in time and space. Temporal discretization is usually addressed using θ-methods. At each time-step, a stationary saddle-point system similar to the Stokes problem has to be solved.

For its numerical solution, discretization techniques usually employed for the Stokes problem can be used, such as finite elements (FE) and finite differences.

Several authors observed that numerical simulations of Biot system may exhibit unphysical oscillations of pressure and displacement [AL92, GLV03, SCF$^+$10] if the spatial and temporal discretization step-size are not properly chosen. In particular, such instabilities appear if the time-step τ of the time integration scheme is smaller than a critical value τ_{crit} that depends on the physical parameters of the problem and on the spatial discretization step-size h.

The condition of Verruijt and Vermeer (V&V) is often employed to detect the source of unphysical oscillations [VV81]. The condition is obtained studying a consolidation problem in the one-dimensional case employing linear (P1) FE and states that the problem is stable if

$$\tau > \tau_{\text{crit}} = \frac{h^2}{6Pk}, \tag{1}$$

where P is the P-wave modulus and k the hydraulic conductivity. The proof is based on the fact that the eigenvalues of the iteration matrix of the time marching scheme must be smaller than one.

In [GLV03], a critical value for the time-step is found imposing that a maximum principle holds for the Schur complement for a one-dimensional poroelastic problem. This strategy provides a limit similar to the V&V condition:

$$\tau_{\text{crit}} = K\frac{h^2}{Bk}, \tag{2}$$

where K is a constant that depends on the choice of the FE spaces for displacements and pressure, and B is the bulk modulus in the one-dimensional case. The constant K is smaller when Taylor-Hood FE are employed allowing for smaller time-steps. In [AL92], the authors observed that the discretization with Taylor-Hood FE could remove the oscillations in some cases.

In our contribution, employing a strategy similar to the one presented in [GLV03], we will find the critical value τ_{crit} for the two-dimensional Biot problem employing Q2-Q1 and Q1-Q1 FE couples. Differently from the one-dimensional case, in the 2D case the maximum principle argument holds true only if $\mu \ll B$. Even if the maximum principle does not hold when shear modulus is comparable with the bulk, the critical time-step is still valid. In fact, assuming that τ_{crit} depends on $P > B$ as in (1), our limit results to be more restrictive and ensures that no oscillations appear in the numerical solution. Finally, we will present a numerical example that shows that the limits derived in the one-dimensional case are not strict enough to ensure that no oscillations appear.

FINITE ELEMENT FORMULATION OF BIOT PROBLEM

Writing the elasticity tensor employing the deviatoric and volumetric terms, the Biot model has the following form:

$$\begin{cases} -\text{div}(2\hat{\mu}\text{dev}(\boldsymbol{\epsilon}(\boldsymbol{u})) + \text{tr}(\boldsymbol{\epsilon}(\boldsymbol{u}))\boldsymbol{I} - p\boldsymbol{I}) &=& \boldsymbol{0} \\ -\text{div}\dot{\boldsymbol{u}} &=& -\text{div}(\hat{k}\nabla p) \end{cases} \tag{3}$$

where $\boldsymbol{\epsilon}(\boldsymbol{u})$ is the linearized strain tensor of the displacement $\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{x}, t) \in \mathbb{R}^d$ and $p = p(\boldsymbol{x}, t) \in \mathbb{R}$ is the pore pressure scaled by the bulk modulus. The parameters $\hat{\mu} = \mu/B$ and $\hat{k} = kB$ are the scaled shear modulus (dimensionless) and hydraulic permeability (m^2s^{-1}). The functions $\boldsymbol{u}$ and p are defined over $\Omega \times I$, where $\Omega \in \mathbb{R}^d$ is the reference configuration, $I = (0, T_{Fin}]$ is a time interval, and d denotes the dimension of the problem.

In order for the equation system (3) to be well-posed, it has to be completed with an initial condition $\boldsymbol{u}(\boldsymbol{x}, 0) = \boldsymbol{u}(\boldsymbol{x})$ and with suitable boundary conditions:

$$\begin{cases} \boldsymbol{u}(\boldsymbol{x}, t) &=& \boldsymbol{0} & \text{on } \Gamma_D^u \\ \boldsymbol{\sigma} \cdot \nu &=& \boldsymbol{t} & \text{on } \Gamma_N^u \end{cases} \tag{4}$$

where

$$\boldsymbol{\sigma} = 2\hat{\mu}\text{dev}(\boldsymbol{\epsilon}(\boldsymbol{u})) + \text{tr}(\boldsymbol{\epsilon}(\boldsymbol{u}))\boldsymbol{I} - p\boldsymbol{I}$$

is the total stress, and Γ_D^u and Γ_N^u provide a partition of the boundary $\partial\Omega$ of the domain. In (4), ν is the outward unit normal vector to $\partial\Omega$ and $\boldsymbol{t}$ is a given function that represents an applied scaled boundary stress.

In principle on the pore pressure no initial condition is needed and the following boundary conditions are imposed

$$\begin{cases} p(\boldsymbol{x}, t) &=& 0 & \text{on } \Gamma_D^p \\ \hat{k}\nabla p \cdot \nu &=& 0 & \text{on } \Gamma_N^p. \end{cases} \tag{5}$$

In (5), Γ_D^p and Γ_N^p provide another partition of the boundary $\partial\Omega$.

The discretization in time of (3) with the implicit Euler method gives rise to a sequence of stationary problems that have to be solved at every time slice j:

$$\begin{cases} -\text{div}(\hat{\mu}\text{dev}(\boldsymbol{\epsilon}(\boldsymbol{u^j})) + \text{tr}(\boldsymbol{\epsilon}(\boldsymbol{u^j})) &-& p^j\boldsymbol{I}) &=& \boldsymbol{0} \\ -\text{div}\boldsymbol{u^j} &+& \tau\text{div}(k\nabla p^j) &=& -\text{div}\boldsymbol{u^{j-1}}. \end{cases} \tag{6}$$

The FE approximation of (6) is presented in [GGF$^+$12] and it is obtained writing the weak form of the problem (6), (4), and (5) and substituting the spaces of the unknowns with suitable interpolation spaces. Neglecting the temporal superscript j, the arising algebraic formulation is:

$$\begin{pmatrix} A & -B^T \\ -B & -\tau\hat{k}C \end{pmatrix} \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{p} \end{pmatrix} = \begin{pmatrix} \boldsymbol{f} \\ -\boldsymbol{g} \end{pmatrix}, \tag{7}$$

where A, B, and C are the elastic stiffness matrix, the divergence operator, and the discrete Laplacian matrix. The continuos problem (6) and its algebraic counterpart (7) are very similar to the well-known Stokes problem employed for the simulation of fluids. The main differences are that in Stokes system $K = 2\mu/d$ and the matrix C is vanishing. This latter property distinguishes the biphasic problem from the Stokes problem.

In the Stokes problem, the FE spaces must satisfy the Brezzi-Babuska-Ladyzhenskaya (LBB) condition [Bre91] and hence FE spaces of the same order for $\boldsymbol{u}$ and p are not stable. In order to employ FE spaces of the same order, in [BP84], an h-dependent diffusion matrix similar to C is added in the block (2,2) to remove spurious oscillations. In the biphasic problem, this stabilization effect is already given by the matrix C that is inherent to the physics of the problem.

At the algebraic level, the LBB condition implies that $\ker B^T = \boldsymbol{0}$ [Bre91]. In the poroelastic case, it is enough for the uniqueness of the solution that $\ker B^T \cap \ker C = \boldsymbol{0}$. This latter condition holds true even employing FE of the same degree for both displacement and pressure, as the $\ker B^T$ is given by the spurious modes while the $\ker C$ is constituted by the constant vectors.

The equation system can be written by static condensation of $\boldsymbol{u}$ in an equivalent problem whose matrix is the pressure Schur complement

$$S = \tau\hat{k}C + BA^{-1}B^T. \tag{8}$$

In the one-dimensional case, the deviatoric part of the elasticty tensor is null, the matrix A is a Laplacian and the matrix $M_S := BA^{-1}B^T$ is a mass matrix. In this case, the matrix S can be seen as the discretization of a diffusion-reaction problem. When the reactive term dominates the diffusion term, i.e. $\tau\hat{k} \ll 1$, the numerical solution may present unphysical oscillations which disappear for $h \to 0$. In [AL92], the authors state the hypothesis that the source of the oscillations is due to the loss of the discrete maximum principle for S. A critical value τ_{crit} can be found by studying for which τ the M-matrix property of S is lost.

SPECTRAL PROPERTY OF SCHUR COMPLEMENT

The pressure Schur complement is sum of two matrices: a discrete Laplacian $\tau\hat{k}C$ and the matrix M_S. The former has a condition number proportional to $1/h^2$, while the latter, in the case of Stokes problem, is known to be spectrally equivalent to the mass matrix M computed on the pressure FE space. Differently from the Stokes case, the matrix A, and hence the matrix M_S, depends on both the mechanical parameters μ and K. In order to study the effect of both these parameters, we will explicitly compute the matrix M_S for several μ while keeping K constant, and we will evaluate:
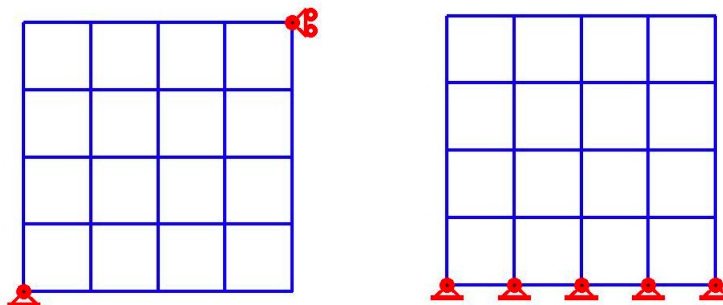
Figure 1. **Representation of the computational domain with** $N = 4$**: minimal boundary conditions (left) and rigid bottom side (right).**

- the norm $\|(M_S - M)/h^2\|_{\max}$

- the condition number of M_S

- the total sum of the entries of M_S.

The norm $\| \cdot \|_{\max}$ is the so-called uniform norm that corresponds to the maximum of the absolute values of its entries:

$$\|A\|_{\max} = \max_{i,j} |A_{ij}|.$$

The sum of the entries in a mass matrix corresponds to the global mass of the system which equals the domain size. This further property is used to check on which parameters the total mass of the system depends on.

Minimal Boundary Condition Experiment

In the first numerical experiment, the domain is the unit square $\Omega = (0, 1)^2$ (see Fig. 1(left)), and it is discretized into N^2 elements of side length $h = 1/N$. We employ Q2-Q1 FE for displacement and pressure. In this experiment, the minimum number of boundary conditions to make A invertible is imposed. In the two-dimensional case, A has three null eigenvalues, corresponding to the three rigid motions of the body and constraints have to be imposed to make it solvable. In Fig. 1(left), the applied boundary conditions are reported: we impose Dirichlet conditions in both directions at point $(0, 0)$ and in one of the two directions at point $(1, 1)$. In Tab. 1, the absolute error of $N^2(M - M_S)$ is reported. In particular, we observe that the error does not depend on the discretization step-size h, and it is almost constant for $\mu > 1$. For every considered N, M_S converges linearly to a mass matrix when μ tends to zero.

In Tab. 1, also the condition number of the matrix M_S is reported. It is increasing when μ or N tends to infinity, while it converges to the condition number of a mass matrix when $\mu \to 0$. In all the numerical tests with *minimal* boundary conditions the global mass of the system was 1, corresponding to the domain area.

μ	$N = 8$	$N = 16$	$N = 32$	$N = 64$
10^4	$4.2{\cdot}10^{-1}$ $(0.8{\cdot}10^5)$	$4.3{\cdot}10^{-1}$ $(1.1{\cdot}10^5)$	$4.4{\cdot}10^{-1}$ $(1.5{\cdot}10^5)$	$4.4{\cdot}10^{-1}$ $(2.2{\cdot}10^5)$
10^2	$4.1{\cdot}10^{-1}$ $(0.8{\cdot}10^3)$	$4.3{\cdot}10^{-1}$ $(1.1{\cdot}10^3)$	$4.3{\cdot}10^{-1}$ $(1.7{\cdot}10^3)$	$4.3{\cdot}10^{-1}$ $(2.3{\cdot}10^3)$
1	$1.43{\cdot}10^{-1}$ (26.54)	$1.47{\cdot}10^{-1}$ (31.77)	$1.48{\cdot}10^{-1}$ (37.74)	$1.49{\cdot}10^{-1}$ (44.14)
10^{-2}	$2.18{\cdot}10^{-3}$ (22.45)	$2.25{\cdot}10^{-3}$ (22.52)	$2.27{\cdot}10^{-3}$ (22.61)	$2.27{\cdot}10^{-3}$ (22.70)
10^{-4}	$2.21{\cdot}10^{-5}$ (22.39)	$2.27{\cdot}10^{-5}$ (22.39)	$2.29{\cdot}10^{-5}$ (22.39)	$2.30{\cdot}10^{-5}$ (22.39)
10^{-6}	$2.21{\cdot}10^{-7}$ (22.39)	$2.27{\cdot}10^{-7}$ (22.39)	$2.29{\cdot}10^{-7}$ (22.39)	$2.30{\cdot}10^{-7}$ (22.39)

Table 1. Absolute norm of $N^2(M - M_S)$ and, in parenthesis, the condition number of M_S.

Rigid Bottom Side Experiment

In Tabs. 2, numerical results with a rigid bottom side of the domain are reported (see Fig. 1(right)), i.e. $\boldsymbol{u} = \boldsymbol{0}$ on $x = 0$. We observe that also in this case M_S converges to M for μ tending to zero and a behavior similar to the previous case has been observed for the condition number, but it is not reported here.

μ	$N = 8$	$N = 16$	$N = 32$	$N = 64$
10^4	4.4×10^{-1} (0.30)	4.4×10^{-1} (0.36)	4.4×10^{-1} (0.38)	4.4×10^{-1} (0.39)
10^2	4.3×10^{-1} (0.62)	4.3×10^{-1} (0.63)	4.3×10^{-1} (0.63)	4.3×10^{-1} (0.63)
1	1.6×10^{-1} (0.95)	1.7×10^{-1} (0.95)	1.7×10^{-1} (0.95)	1.7×10^{-1} (0.95)
10^{-2}	2.9×10^{-3} (0.99)	3.1×10^{-3} (0.99)	3.1×10^{-3} (0.99)	3.1×10^{-3} (0.99)
10^{-4}	3.0×10^{-5} (1.00)	3.1×10^{-5} (1.00)	3.2×10^{-5} (1.00)	3.2×10^{-5} (1.00)
10^{-6}	3.0×10^{-7} (1.00)	3.1×10^{-7} (1.00)	3.2×10^{-7} (1.00)	3.2×10^{-7} (1.00)

Table 2. Absolute norm of $N^2(M - M_S)$ and, in parenthesis, the sum of the entries of M_S.

The main difference concerns the mass preservation. With these boundary conditions we observe that the total mass is decreasing when $\hat{\mu} \to \infty$.

Different Boundary Conditions and Finite Element Couples

A similar behavior has been observed in numerical experiments performed on different boundary conditions, namely rigid bottom and top side and confined compression.

Employing Q1-Q1 FE spaces, we obtain that the mass is preserved with minimal boundary conditions and it is a function of $\hat{\mu}$ when more complicated boundary conditions are applied. The condition number of M_S is infinity since it is singular employing FE spaces of the same order. When $\hat{\mu} \to 0$, the matrix M_S is converging to a mass matrix whose local matrix built on the reference element is

$$M_\ell = \frac{1}{48} \begin{vmatrix} 5 & 3 & 3 & 1 \\ 3 & 5 & 1 & 3 \\ 3 & 1 & 5 & 1 \\ 1 & 3 & 3 & 5 \end{vmatrix}. \tag{9}$$

MAXIMUM PRINCIPLE FOR THE SCHUR COMPLEMENT

For $\hat{\mu} \to 0$, the matrix M_S tends to a mass matrix and hence S is the discretization of a two-dimensional diffusion-reaction problem and the argument presented in [GLV03] can be applied. In order to ensure that S enjoys the M-matrix property, the following three conditions are sufficient:

1) $S_{ii} > 0$

2) $S_{ij} \leq 0 \quad i \neq j$

3) $S_{ii} \geq \sum_{j=1}^{i-1} |S_{ij}| + \sum_{j=i+1}^{N+1} |S_{ij}| \quad \forall i$

For a matrix derived from a FE discretization, a sufficient condition for being an M-matrix is that each local matrix enjoys the same property. The local matrix S_ℓ is the sum of a local Laplacian $\tau \hat{k} C_\ell$ and a local mass matrix $h^2 M_\ell$, where

$$C_\ell = \frac{1}{6} \begin{vmatrix} 4 & -1 & -1 & -2 \\ -1 & 4 & -2 & -1 \\ -1 & -2 & 4 & -1 \\ -2 & -1 & -1 & 4 \end{vmatrix} \tag{10}$$

and

$$M_\ell = \frac{1}{36} \begin{vmatrix} 4 & 2 & 2 & 1 \\ 2 & 4 & 1 & 2 \\ 2 & 1 & 4 & 2 \\ 1 & 2 & 2 & 4 \end{vmatrix}. \tag{11}$$

Properties 1) and 3) are trivially satisfied for matrix S_ℓ. In order for the property 2) to hold, the requirement is that

$$-\tau \hat{k} \frac{1}{6} + h^2 \frac{1}{18} < 0,$$

resulting in a lower bound for the discretization step-size

$$\tau_{\mathrm{crit}} = \frac{h^2}{3\hat{k}} = \frac{h^2}{3kB} \tag{12}$$

The limit above holds for quadrilateral elements employing Q2-Q1 finite elements. This formula is the extension of (1) and (2) to the two-dimensional case. As in (2), the
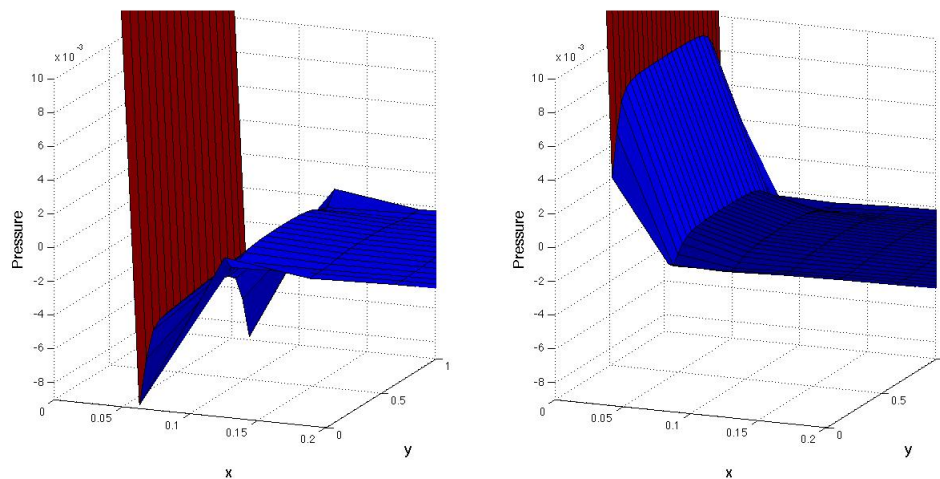
Figure 2. Pressure distribution after the first time step: $N = 16$ **left and** $N = 24$ **right.**

limit depends on the bulk modulus B. The limit case for $\hat{\mu} \to 0$ also corresponds to the limit of V&V but with a larger numerical constant.

In the Q1-Q1 case, the maximum principle argument gives to the following limit:

$$\tau > \frac{3h^2}{8\hat{k}} = \frac{3h^2}{8kB}. \tag{13}$$

As observed in [AL92], the use of FE spaces of the same order gives rise to a stricter limit on the time-step.

NUMERICAL TEST

In this section we compare the limit presented in the previous section with the V&V limit. The domain is the unit square and the physical parameters [GGF$^+$12] are:

- $\mu = 0.3137$ Pa m^{-2}

- $B = 0.6703$ Pa m^{-2}

- $k = 3.6454$ m^4 Pa^{-1} s^{-1}

In this numerical experiment, we consider the boundary conditions shown in Fig. 1(right) for the displacement, while for the pressure we set $p = 0$ at the bottom side and $p = 0.5$ Pa at the top side. Employing a time step $\tau = 2.5 \times 10^{-4}$ s, we have two different limits N_V and N_F for the V&V condition and our approach, respectively. Inverting (1) and (13) we find $N_V = 16$ and $N_F = 24$.

In Fig. 2 the numerical results with the two limits are reported. In the left plot, we can observe that the resolution given by the V&V limit is not enough to detect the boundary layer while with the limit presented in this paper no instabilities occur.
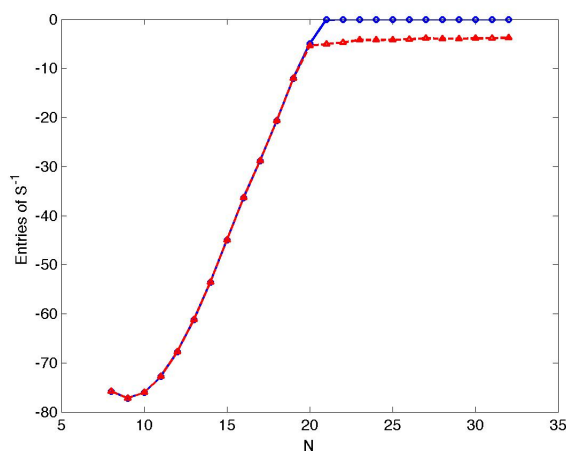
Figure 3. Minimum magnitude of the entries in the matrix (triangles) and minimum magnitude of the entries connected to the boundary (circle).

The M-matrix property ensures that the entries of the inverse Schur complement are non-negative. The loss of this discrete maximum principle results in negative entries, especially in the elements near the boundary. A similar effect can be observed also if μ is of the same order of magnitude of B. In Fig. 3, we plot the minimum entry of the S^{-1} and the minimum entries of the nodes that are connected to the boundary. We observe that both entries become constant when $N \geq 18$. This shows that our limit is strict enough for ensuring that no change of sign of the entries near the boundary occurs.

CONCLUSIONS

In this work, we presented an approach to study the source of instabilities in the numerical simulation of the Biot model. Employing a maximum principle argument for the Schur complement of the system, we obtained a lower bound for τ_{crit} that ensures that no oscillations occur and that the grid is fine enough to resolve the boundary layer of the problem. The limit has been found by studying under which condition the Schur complement of the system is an M-matrix. In the two-dimensional case, this argument is valid only for $\mu \ll B$, but the presented limit seems to provide a stricter condition that ensures that the grid is fine enough. Numerical experiments have been presented in the two-dimensional case but our methodology is readily applicable for the computation of the constant in two- and three-dimensional cases, with different FE couples, and on different element types.

REFERENCES

[AL92]   Murad M. A. and A. F. D. Loula. Improved accuracy in finite element analysis of Biot's consolidation problem. *Computer Methods in Applied Mechanics and Engineering*, 95(3):359 – 382, 1992.

[Bio41]   M. A. Biot. General theory of three dimensional consolidation. *J. Appl. Physics*, 1941.

[BP84]   F. Brezzi and J. Pitkäranta. On the stabilization of finite element approximations of the Stokes equations. In *Efficient Solutions of Elliptic Systems*, pages 11–19. Vieweg, 1984.

[Bre91]   F. Brezzi. *Mixed and Hybrid Finite Element Method*. Springer-Verlag, New York, 1991.

[FGD$^+$11]   M. Favino, C. Gross, M. Drolshagen, L. Keilig, J. Deschner, C. Bourauel, and R. Krause. Validation of a heterogeneous elastic-biphasic model for the numerical simulation of the PDL. *Computer Methods in Biomechanics and Biomedical Engineering*, 2011. to appear.

[FGLR$^+$05]   S. Federico, A. Grillo, G. La Rosa, G. Giaquinta, and W. Herzog. A transversely isotropic, transversely homogeneous microstructural-statistical model of articular cartilage. *Journal of Biomechanics*, 38(10):2008 – 2018, 2005.

[GFW12]   Alfio Grillo, Salvatore Federico, and Gabriel Wittum. Growth, mass transfer, and remodeling in fiber-reinforced, multi-constituent materials. *International Journal of Non-Linear Mechanics*, 2012.

[GGF$^+$12]   A. Grillo, C. Giverso, M. Favino, R. Krause, M. Lampe, and G. Wittum. Mass transport in porous media with variable mass. In J.M.P.Q. Delgado, Antonio Gilson Barbosa Lima, and Marta Vazquez Silva, editors, *Numerical Analysis of Heat and Mass Transfer in Porous Media*, volume 27 of *Advanced Structured Materials*, pages 27–61. Springer Berlin Heidelberg, 2012.

[GLV03]   F.J. Gaspar, F.J. Lisbona, and P.N. Vabishchevich. A finite difference analysis of biot's consolidation model. *Applied Numerical Mathematics*, 44(4):487 – 506, 2003.

[SCF$^+$10]   I. A. Stokes, S. Chegini, S. J. Ferguson, M. G. Gardner-Morse, J. C. Iatridis, and J. P. Laible. Limitation of finite element analysis of poroelastic behavior of biological tissues undergoing rapid loading. *Computer Methods in Applied Mechanics and Engineering*, 2010.

[VV81]   P. A. Vermeer and A. Verruijt. An accuracy condition for consolidation by finite elements. *International Journal for Numerical and Analytical Methods in Geomechanics*, 5(1):1–14, 1981.