A MONOTONE MULTIGRID SOLVER FOR TWO BODY CONTACT PROBLEMS IN BIOMECHANICS

R. KORNHUBER, R. KRAUSE, O. SANDER, P. DEUFLHARD, S. ERTEL

ABSTRACT. The purpose of the paper is to apply monotone multigrid methods to static and dynamic biomechanical contact problems. In space, a finite element method involving a mortar discretization of the contact conditions is used. In time, a new contact–stabilized Newmark scheme is presented. Numerical experiments for a two body Hertzian contact problem and a biomechanical knee problem are reported.

1. INTRODUCTION

Mechanical loading in human joints plays a crucial role in medical treatment. Non-invasive approaches for in vivo measurements do not exist so that predictions of joint pressures are only accessible via numerical simulation. A detailed modelling of joints like the human knee leads to heterogeneous dynamic contact problems in three space dimensions with strongly varying behavior and complex 3D geometry of the interacting components such as bones, cartilage, tendons, ligaments, and soft tissue. The full complexity of the patient–specific problem is certainly intimidating. The present paper restricts its attention to a simplified core problem, the dynamic two body contact problem for two linear elastic, homogeneous bones with realistic 3D geometry.

A main feature of this problem is its intrinsic non–smooth nonlinearity as emerging even from linearized contact conditions. Popular approaches like multibody dynamics or mass–spring systems [4, 20, 36] have not yet reached a level of sophistication that permits reliable predictions for individual patients in clinical applications. In a finite element framework, the nonlinearity can be circumvented by explicit time discretization [29, 34]. However, the associated CFL condition might be quite restrictive so that extremely small time steps might be necessary in biomechanical applications. In contrast to that, implicit time discretizations are unconditionally stable, but require the solution of a static two body contact problem in each time step. As for the static problem, the typical approach in the engineering community is to apply Newton–Raphson solvers after regularization, see, e.g., [21, 29, 43]. Such a penalty approach aims at a problem–dependent compromise between regularization error and convergence speed. A systematic determination of suitable penalty parameters via dual methods like augmented Lagrangians [14] leads to indefinite saddle–point problems with additional unknowns.

Recently active set strategies [15, 19] attracted new interest, partly because of their reinterpretation as non–smooth Newton methods [28, 37]. Convergence is guaranteed, if the resulting local problems are solved exactly. However, sufficient

This work was supported by the DFG Research Center MATHEON.

accuracy criteria for inexact versions still seem to be an open problem. In addition, there are counterexamples showing that straightforward multigrid methods as applied to the linearized problems might fail to converge [25] and that special variants [24] adapted to current guesses of the free boundary should be used.

Such variants are similar to monotone multigrid methods (MMG) which can be regarded as multilevel versions of well–known projected block Gauß–Seidel relaxations [22, 23, 40]. Exploiting convexity rather than smoothness, monotone multigrid converges globally and does not involve any additional parameters. Asymptotic multigrid convergence rates were shown in the scalar case [22]. For simple model problems the convergence speed was observed to be comparable to usual linear multigrid as applied to related unconstrained problems [23, 40]. It is not all evident a–priori, whether this desirable feature carries over to the more complex biomechanical contact problems.

The purpose of the present paper is to shed some light into this question. In Section 2 we consider the static case. We use a finite element discretization in space involving a mortar discretization of the contact conditions [5, 11, 40]. Well–known instabilities and suboptimal convergence of straightforward node–to–surface representations are avoided in this way. In order to clarify the ideas behind MMG and its implementation, we present algorithms for the Laplacian, Signorini's problem and two body contact in the common framework of successive minimization and prove convergence. In our numerical computations, we consider the two body contact of the human tibia and femur. The highly resolved exact geometry is taken from the Visible Human Data Set [2]. It is clear that the individual geometry of the contact boundary has a strong influence on contact stresses. Hence, starting from a coarse approximation, we successively approximate the geometry by shifting new nodes to the exact boundary in the course of local refinement. A parametrization of the exact boundary is created automatically during the coarsening process [27]. We found that MMG performed like a linear multigrid algorithm. The simultaneous detection of the contact points is hardly visible in the convergence behavior. As for linear multigrid methods, the actual convergence speed was slightly decelerated by the shifting process and the reduced shape regularity of the mesh. Section 3 is devoted to time dependent problems. For time discretization, we present a new contact–stabilized variant from the Newmark family of variational integrators [12, 29, 33, 34]. Numerical comparisons with existing methods [34] of second order display significantly reduced oscillations at the contact boundary together with similar stability and conservation properties. A theoretical justification will be the subject of a forthcoming paper [9]. Using the implicit second–order version of our contact-stabilized time discretization, we finally compute the dynamic contact of tibia and femur illustrating the reliability and efficiency of our approach. Extensions to heterogeneous models involving cortical and trabecular bone [17] as well as tendons and ligaments are the subject of ongoing research.

2. Monotone multigrid methods

2.1. **Successive minimization and multigrid.** We consider the constrained minimization problem

$$u \in \mathcal{K}: \qquad \mathcal{J}(u) \leq \mathcal{J}(v) \qquad \forall v \in \mathcal{K} \tag{1}$$

where $\mathcal{K} \subset \mathcal{S}$ is a non-empty, closed, convex subset of the finite–dimensional linear space $\mathcal{S}$ and the quadratic functional

$$\mathcal{J}(v) = \tfrac{1}{2}a(v,v) - \ell(v) \tag{2}$$

is generated by a self-adjoint, positive definite bilinear form $a(\cdot,\cdot)$ and a linear functional ℓ on $\mathcal{S}$, respectively. With these assumptions, (1) admits a unique solution. Now let $\mathcal{K}$ allow for a decomposition

$$\mathcal{K} = \mathcal{K}_1 + \cdots + \mathcal{K}_n \ , \qquad \mathcal{K}_l \subset U_l \ , \tag{3}$$

with local linear subspaces $U_l \subset \mathcal{S}$ that generate an associated splitting of the global space $\mathcal{S}$,

$$\mathcal{S} = U_1 + \cdots + U_n \ . \tag{4}$$

Then the corresponding projected block Gauß–Seidel relaxation for the iterative solution of (1) reads as follows. For a given iterate $w_0 = u^\nu$, we subsequently solve the local minimization problems

$$\mathcal{J}(w_l) = \min_{\substack{v \in U_l \\ w_{l-1}+v \in \mathcal{K}}} \mathcal{J}(w_{l-1} + v) \ , \quad l = 1, \ldots, n, \tag{5}$$

and the next iterate is $u^{\nu+1} = w_n$. It is well–known [13, Chapter 5.4] that, for any initial iterate $u^0 \in \mathcal{K}$, the block Gauß–Seidel relaxation (5) converges to the solution u of (1).

In order to increase the convergence speed, one might select a nested sequence of subspaces

$$\mathcal{X}_0 \subset \mathcal{X}_1 \subset \cdots \subset \mathcal{X}_{J-1} \subset \mathcal{X}_J = \mathcal{S}$$

and realize a similar relaxation process on each subspace $\mathcal{X}_k$, $k = 0, \ldots, J-1$. To this end, we decompose $\mathcal{X}_k$ into subspaces

$$\mathcal{X}_k = U_1^{(k)} + \cdots + U_{n_k}^{(k)} \tag{6}$$

with $U_l^{(J)} = U_l$. Then an associated *block multilevel relaxation* is obtained as follows. From a given iterate u^ν, we first compute the so-called *smoothed iterate* $\overline{u}^\nu$ by block Gauß–Seidel relaxation (5) on $\mathcal{S}$. Then, starting with $w_{n_J}^{(J)} = \overline{u}^\nu$, we compute intermediate iterates $w_{n_k}^{(k)}$ for $k = J-1, \ldots, 0$ by successive restriction and constrained minimization

$$\mathcal{J}(w_l^{(k)}) = \min_{v \in \mathcal{D}_l^{(k)}} \mathcal{J}(w_{l-1}^{(k)} + v) \ , \quad l = 1, \ldots, n_k \ , \qquad w_0^{(k)} = w_{n_{k+1}}^{(k+1)} \ . \tag{7}$$

Finally, the next iterate is $u^{\nu+1} = w_{n_0}^{(0)}$. The condition

$$w_{l-1}^{(k)} + \mathcal{D}_l^{(k)} \subset \mathcal{K} \tag{8}$$

on the closed, convex subsets $\mathcal{D}_l^{(k)} \subset U_l^{(k)}$ preserves the convergence of the block Gauß–Seidel relaxation (5).

Theorem 1. *Assume that condition* (3) *on* $\mathcal{K}$ *and condition* (8) *are satisfied. Then, for any initial iterate* $u^0 \in \mathcal{K}$*, the block multilevel relaxation* (7) *converges to the solution* u *of* (1).

The proof is omitted here since it is essentially the same as for Theorem 3.1 in [30] or Theorem 2.1 in [22].

It turns out that a proper selection of the subspaces $\mathcal{X}_k$ is crucial for the *fast* convergence of the multilevel relaxation. This will be illustrated by two examples.

2.1.1. *Classical multigrid for the discretized Laplacian.* Let $\Omega \subset \mathbb{R}^d$ be a polyhedral domain and let $\Gamma_D \subset \partial\Omega$ have positive measure. Then

$$\mathcal{S} = \{v \in C(\overline{\Omega}) \mid v|_T \text{ is linear } \forall T \in \mathcal{T}\ , v|_{\Gamma_D} = 0\}$$

stands for the space of linear finite elements with Dirichlet boundary conditions with respect to some triangulation $\mathcal{T}$ of Ω. The standard nodal basis functions are denoted by λ_p, $p \in \mathcal{N}$, where $\mathcal{N}$ is the subset of the vertices of $\mathcal{T}$ lying in $\overline{\Omega} \setminus \Gamma_D$. We define $\mathcal{K} = \mathcal{S}$ and

$$a(v,w) = (\nabla v, \nabla w)_{L^2(\Omega)}\ , \quad \ell(v) = (f,v)_{L^2(\Omega)}\ , \qquad v,w \in \mathcal{S}\ ,$$

for some given $f \in L^2(\Omega)$. Then (1) becomes the discretization of a Poisson equation with homogeneous Dirichlet conditions on Γ_D and Neumann conditions on $\partial\Omega \setminus \Gamma_D$.

Selecting the subspaces

$$U_p = \operatorname{span}\{\lambda_p\}\ , \qquad p \in \mathcal{N}\ ,$$

we recover the classical Gauß–Seidel relaxation. Now we assume that a hierarchy

$$\mathcal{S}_0 \subset \cdots \subset \mathcal{S}_J = \mathcal{S}$$

of finite element spaces $\mathcal{S}_k$ associated with a sequence of nested triangulations $\mathcal{T}_k$ with vertices $\mathcal{N}_k$ is available, e.g. from some adaptive refinement process. Let $\lambda_p^{(k)}$, $p \in \mathcal{N}_k$, denote the standard nodal basis of $\mathcal{S}_k$. Then the choice

$$\mathcal{X}_k = \mathcal{S}_k\ , \qquad U_p^{(k)} = \operatorname{span}\{\lambda_p^{(k)}\}\ , \quad p \in \mathcal{N}_k\ ,$$

provides a V–cycle of the classical multigrid method with one Gauß–Seidel presmoothing step. The weights $r_{pq} = \lambda_p^{(k)}(q)$ of the canonical Galerkin restriction and prolongation can be read from the nodal representation

$$\lambda_p^{(k)} = \sum_{q \in \mathcal{N}_{k+1}} \lambda_p^{(k)}(q) \lambda_q^{(k+1)}\ . \tag{9}$$

Additional smoothing steps, postsmoothing, or W–cycles can be formulated in a similar way (cf. [41, 42]).

2.1.2. *Truncated monotone multigrid for obstacle problems.* We consider the same situation as above with the only difference that $\mathcal{K}$ now takes the form

$$\mathcal{K} = \{v \in \mathcal{S} \mid v(p) \le \varphi(p),\ p \in \mathcal{N}\}$$

with some prescribed obstacle function $\varphi \in \mathcal{S}$.

From the previous example 2.1.1, we keep the choice $U_p = \operatorname{span}\{\lambda_p\}$ which now leads to the projected Gauß–Seidel relaxation on the fine grid. In order to construct suitable coarse grid spaces $\mathcal{X}_k$, we first introduce the coincidence set

$$\mathcal{N}^\bullet(\overline{u}^\nu) = \{p \in \mathcal{N} \mid \overline{u}^\nu(p) = \varphi(p)\}$$

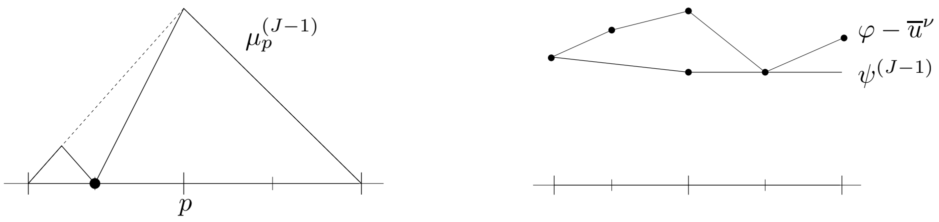

FIGURE 1. Truncated hat function $\mu_p^{(J-1)}$ and restriction r_{J-1}

of some smoothed iterate $\overline{u}^\nu$. As, in general, approximate or exact coincidence sets cannot be represented on coarse grids, the spaces $\mathcal{X}_k$ are constructed in such a way that no correction takes place at $p \in \mathcal{N}^\bullet(\overline{u}^\nu)$. Starting with

$$\mu_p^{(J)} = \begin{cases} 0\,, & \text{if } p \in \mathcal{N}^\bullet(\overline{u}^\nu) \\ \lambda_p\,, & \text{else} \end{cases}\,, \qquad p \in \mathcal{N}\,,$$

we recursively define *truncated* nodal basis functions $\mu_p^{(k)}$ by

$$\mu_p^{(k)} = \sum_{q \in \mathcal{N}_{k+1}} \lambda_p^{(k)}(q)\mu_q^{(k+1)}\,, \qquad p \in \mathcal{N}_k\,, \qquad k = J-1, \ldots, 0\,. \tag{10}$$

Note that (10) can be directly translated into a Galerkin restriction with the modified weights $r_{pq} = 0$ for $q \in \mathcal{N}_{k+1} \cap \mathcal{N}^\bullet(\overline{u}^\nu)$. We finally set

$$\mathcal{X}_k = \sum_{p \in \mathcal{N}_k} U_p^{(k)}\,, \qquad U_p^{(k)} = \operatorname{span}\{\mu_p^{(k)}\}\,,$$

to recover a truncated monotone multigrid method [22]. Note that $\mathcal{X}_k \not\subset \mathcal{S}_k$ in general and that $\mathcal{X}_k$ may vary in each iteration step. The local coarse grid constraints appearing in (7) take the form

$$\mathcal{D}_p^{(k)} = \{v = v_p \mu_p^{(k)} \mid v_p \le \psi^{(k)}(p)\} \subset U_p^{(k)} \tag{11}$$

where the defect obstacles $\psi^{(k)} \in \mathcal{S}_k$ are recursively defined by

$$\psi^{(k)} = r_k\left(\psi^{(k+1)} - v^{(k+1)}\right)\,, \qquad \psi^{(J)} = \varphi - u^\nu\,, \tag{12}$$

with $v^{(k)} = w_{n_k}^{(k)} - w_0^{(k)} \in \mathcal{X}_k$ denoting the current correction on level k. Condition (8) is then a consequence of the monotonicity

$$0 \le r_k v \le v\,, \qquad \forall v \in \mathcal{S}_{k+1}\,,\ \ 0 \le v\,.$$

For example, the restriction operators $r_k : \mathcal{S}_{k+1} \to \mathcal{S}_k$ defined by

$$r_k v(p) = \begin{cases} \min\{v(q) \mid q \in \mathcal{N}_{k+1} \cap \operatorname{int}\operatorname{supp}\mu_p^{(k)}\}\,, & \text{if } \mu_p^{(k)} \not\equiv 0 \\ 0\,, & \text{else} \end{cases}\,, \qquad p \in \mathcal{N}_k\,,$$

have this property. The resulting algorithm is globally convergent by Theorem 1 and can be implemented as a multigrid V–cycle with modified restriction/prolongation and projected Gauß–Seidel smoothing. Asymptotic multigrid convergence rates have been established for non–degenerate problems. We refer to [22] for details.

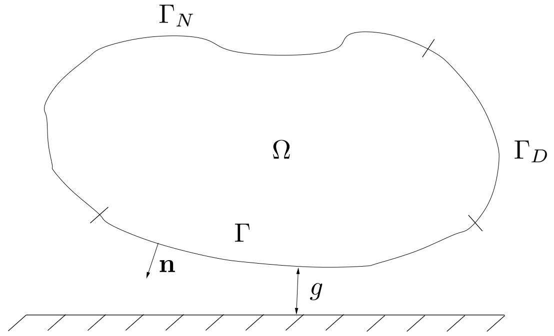

FIGURE 2. Signorini's problem

2.2. **Signorini's problem in linear elasticity.** Let $\Omega \subset \mathbb{R}^d$ be a polyhedral domain, representing a linear elastic body. The boundary $\partial\Omega$ is decomposed into three disjoint parts

$$\partial\Omega = \Gamma_D \cup \Gamma_N \cup \Gamma \ .$$

Assume that Ω is clamped at Γ_D and that deformation of Ω is caused by the volume force density $\boldsymbol{f}$ and the traction force density $\boldsymbol{t}$ on Γ_N. Signorini's problem amounts to the solution of the corresponding equilibrium conditions subject to the constraint that the normal displacements on Γ must not exceed the normal gap $g \geq 0$ between Ω and a rigid foundation (see Figure 2). More precisely, the set of admissible displacements is given by

$$\mathcal{K}^* = \{\boldsymbol{v} \in H \mid \boldsymbol{v} \cdot \boldsymbol{n} \leq g \text{ a.e. on } \Gamma\} \ ,$$

with $\boldsymbol{n}$ denoting the outward normal on $\partial\Omega$ and $H = \{\boldsymbol{v} \in H^1(\Omega)^d \mid \boldsymbol{v}|_{\Gamma_D} = 0\}$. Then the weak formulation of Signorini's problem can be written as the constrained minimization problem

$$\boldsymbol{u}^* \in \mathcal{K}^* : \quad \mathcal{J}(\boldsymbol{u}^*) \leq \mathcal{J}(\boldsymbol{v}) \quad \forall \boldsymbol{v} \in \mathcal{K}^* \ . \tag{13}$$

The quadratic energy $\mathcal{J}$ has the form (2) with $a(\cdot,\cdot)$ and ℓ defined by

$$\begin{aligned} a(\boldsymbol{v},\boldsymbol{w}) &= \int_\Omega \frac{E}{2(1+\nu)} \varepsilon(\boldsymbol{v}) : \varepsilon(\boldsymbol{w}) + \frac{E\nu}{2(1+\nu)(1-2\nu)} \mathrm{div}\boldsymbol{v}\ \mathrm{div}\boldsymbol{w}\ dx \ , \\ \ell(\boldsymbol{v}) &= \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}\ dx + \int_{\Gamma_N} \boldsymbol{t} \cdot \boldsymbol{v}\ d\sigma \ , \end{aligned} \tag{14}$$

involving Young's modulus E, Poisson ratio ν and the linearized strain tensor

$$\varepsilon_{ij}(\boldsymbol{v}) = \frac{1}{2}(\partial_i v_j + \partial_j v_i) \ , \qquad i,j = 1,\dots,d \ .$$

For existence and uniqueness results we refer to [21, p. 113].

We now consider a discretized version based on the finite element space

$$\mathcal{S} = \{\boldsymbol{v} \in H \mid v_i|_T \text{ is linear } \ \forall T \in \mathcal{T}, i = 1,\dots,d\} \tag{15}$$

associated with some given triangulation $\mathcal{T}$ of Ω. $\mathcal{S}$ is spanned by the nodal basis

$$\lambda_p \boldsymbol{E}^i, \quad i = 1,\dots,d, \ \ p \in \mathcal{N} \ ,$$

with the Cartesian unit vectors $\boldsymbol{E}^i \in \mathbb{R}^d$ and the subset $\mathcal{N}$ of the vertices of $\mathcal{T}$ lying in $\overline{\Omega} \setminus \Gamma_D$. The set $\mathcal{K} \subset \mathcal{S}$ of admissible approximate displacements is defined by

$$\mathcal{K} = \{\boldsymbol{v} \in \mathcal{S} \mid \boldsymbol{v}(p) \cdot \boldsymbol{n}(p) \leq g(p) \quad \forall p \in \Gamma \cap \mathcal{N}\} \ .$$

With these specifications our model problem (1) becomes a finite element discretization of Signorini's problem (13). Optimal error estimates are available for H^2–regular problems. We refer to [18, p. 109] or [21, p. 127] for details.

In order to derive a monotone multigrid method for the discretized Signorini problem, we have to specify local subspaces U_p and coarse grid spaces $\mathcal{X}_k$, $U_p^{(k)}$. The d–dimensional subspaces

$$U_p = \operatorname{span}\{\lambda_p \boldsymbol{E}^i \mid i = 1, \dots, d\} \ , \qquad p \in \mathcal{N} \ ,$$

generate a splitting (4) of $\mathcal{S}$. Note that $\mathcal{K}$ can be decomposed into the subsets

$$\mathcal{K}_p = \{\boldsymbol{v} \in U_p \mid \boldsymbol{v}(p) \cdot \boldsymbol{n}(p) \leq g(p)\}$$

according to (3). Hence, the associated block Gauß–Seidel relaxation is convergent.

The definition of $\mathcal{X}_k$ is based on an underlying hierarchy of triangulations $\mathcal{T}_0$, $\dots$, $\mathcal{T}_J = \mathcal{T}$ with nodes $\mathcal{N}_k$ and a corresponding nested sequence of scalar finite element spaces spanned by the nodal basis functions $\lambda_p^{(k)}$, $p \in \mathcal{N}_k$. Assuming that block Gauß–Seidel relaxation is applied to some iterate $\boldsymbol{u}^\nu$, the resulting smoothed iterate $\overline{\boldsymbol{u}}^{\boldsymbol{\nu}}$ gives rise to the current approximation approximate coincidence set

$$\mathcal{N}^\bullet(\overline{\boldsymbol{u}}^{\boldsymbol{\nu}}) = \left\{p \in \mathcal{N} \cap \Gamma \mid \overline{\boldsymbol{u}}^{\boldsymbol{\nu}}(p) \cdot \boldsymbol{n}(p) = g(p)\right\} \ .$$

of the exact coincidence set $\mathcal{N}^\bullet(\mathbf{u})$. As $\mathcal{N}^\bullet(\overline{\boldsymbol{u}}^{\boldsymbol{\nu}})$ has no representation on coarse grids the spaces $\mathcal{X}_k$ are constructed in such a way that no normal corrections at $p \in \mathcal{N}^\bullet(\overline{\boldsymbol{u}}^{\boldsymbol{\nu}})$ can occur. Following [23], we first choose a local orthonormal basis $\boldsymbol{e}^i(p)$ of $\mathbb{R}^d$ with the properties $\boldsymbol{e}^1(p) = \boldsymbol{n}(p)\ \forall p \in \mathcal{N} \cap \Gamma$ and $\boldsymbol{e}^i(p) = \boldsymbol{E}^i\ \forall p \in \mathcal{N} \setminus \Gamma$. Then, starting with

$$\left(\boldsymbol{\mu}_p^{(J)}\right)^i = \begin{cases} 0, & \text{if } i = 1 \text{ and } p \in \mathcal{N}^\bullet(\overline{\boldsymbol{u}}^{\boldsymbol{\nu}}) \\ \lambda_p \boldsymbol{e}^i(p), & \text{else} \end{cases} \ , \tag{16}$$

we recursively define truncated basis functions

$$\left(\boldsymbol{\mu}_p^{(k)}\right)^i = \sum_{q \in \mathcal{N}_{k+1}} \lambda_p^{(k)}(q) \sum_{j=1}^{d} \boldsymbol{e}^i(p) \cdot \boldsymbol{e}^j(q) \left(\boldsymbol{\mu}_q^{(k+1)}\right)^j \ . \tag{17}$$

We emphasize that (17) can be directly translated into restriction and prolongation operators. The weighting factors $\boldsymbol{e}^i(p) \cdot \boldsymbol{e}^j(q)$ are intended to avoid large energy of the functions $\left(\boldsymbol{\mu}_p^{(k)}\right)^i$ which in turn would lead to poor convergence speed of the corresponding iterative scheme. By construction, we have $\left(\boldsymbol{\mu}_p^{(k)}\right)^i(q) \cdot \boldsymbol{n}(q) = 0$, if $q \in \mathcal{N}^\bullet(\overline{\boldsymbol{u}}^{\boldsymbol{\nu}})$. Hence, $\mathcal{N}^\bullet(\overline{\boldsymbol{u}}^{\boldsymbol{\nu}})$ is not affected by corrections in the direction of $\left(\boldsymbol{\mu}_p^{(k)}\right)^i$. Otherwise, $\left(\boldsymbol{\mu}_p^{k}\right)^i$ behaves like a coarse grid nodal basis function, i.e.,

$$\left(\boldsymbol{\mu}_p^{(k)}\right)^i(q) \cdot \boldsymbol{e}^j(q) = \lambda_p^{(k)}(q)\ \boldsymbol{e}^i(p) \cdot \boldsymbol{e}^j(q)$$

holds for $j \neq 1$ or $q \notin \mathcal{N}^\bullet(\overline{\boldsymbol{u}}^{\boldsymbol{\nu}})$. Moreover, we have $\left(\boldsymbol{\mu}_p^{(k)}\right)^i = \lambda_p^{(k)}\ \boldsymbol{E}^i$ provided that p is sufficiently far away from Γ or, more precisely, int supp $\lambda_p^{(k)} \cap \mathcal{N}^\bullet(\overline{u}^\nu) = \emptyset$.

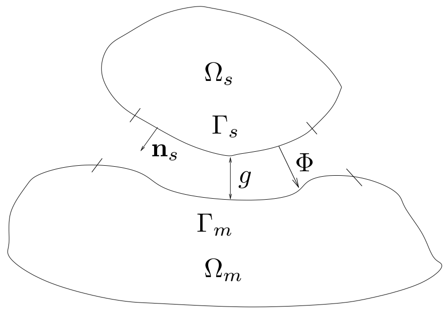

FIGURE 3. Two body contact problem

In the special case of constant normal directions, (17) reduces to the canonical restriction

$$\left(\boldsymbol{\mu}_p^{(k)}\right)^i = \sum_{q \in \mathcal{N}_{k+1}} \lambda_p^{(k)}(q) \, \left(\boldsymbol{\mu}_q^{(k+1)}\right)^i. \tag{18}$$

We finally set

$$\mathcal{X}_k = \sum_{p \in \mathcal{N}_k} U_p^{(k)} \,, \qquad U_p^{(k)} = \operatorname{span}\left\{ \left(\boldsymbol{\mu}_p^{(k)}\right)^1, \ldots, \left(\boldsymbol{\mu}_p^{(k)}\right)^d \right\} \,.$$

In analogy to (11), the coarse grid constraints $\mathcal{D}_p^{(k)}$ take the form

$$\mathcal{D}_p^{(k)} = \left\{ \boldsymbol{v} = \sum_{i=1}^d v_p^i \left(\boldsymbol{\mu}_p^{(k)}\right)^i \mid v_p^i \in \left[\left(\underline{\psi}_p^{(k)}\right)^i, \left(\overline{\psi}_p^{(k)}\right)^i \right], \; i = 1, \ldots, d \right\} \subset U_p^{(k)} \tag{19}$$

with obstacles $\left(\underline{\psi}_p^{(k)}\right)^i$, $\left(\overline{\psi}_p^{(k)}\right)^i$ providing the monotonicity condition (8). Similar to (12), such obstacles can be obtained by successive update and restriction. We refer to [23] for details.

The resulting monotone multigrid method is convergent by Theorem 1 and can be implemented as a multigrid V−cycle with projected block Gauß–Seidel smoothing.

2.3. **Static two body contact.** We consider the polyhedral domain $\Omega = \Omega_m \cup \Omega_s$ representing two elastic bodies Ω_m, $\Omega_s \subset \mathbb{R}^d$, where the subscripts m, s stand for master and slave in the mortar discretization below. We assume that Ω_m, Ω_s are sufficiently close in the reference configuration and have empty intersection. Each of the boundaries is divided into three disjoint subsets

$$\partial\Omega_i = \Gamma_{D,i} + \Gamma_{N,i} + \Gamma_i \qquad i = m, s \,.$$

The bodies are clamped at $\Gamma_{D,i}$. Deformation is caused by volume forces and traction forces at $\Gamma_{N,i}$. No penetration must occur at the possible contact boundary $\Gamma = \Gamma_m \cup \Gamma_s$. For a more precise formulation of this constraint, we assume that Γ_m and Γ_s can be identified by a smooth, bijective mapping

$$\Phi : \Gamma_s \to \Gamma_m \,.$$

In the numerical experiments to be reported later, we simply use normal projection. Introducing the jump

$$[\boldsymbol{v}] = \boldsymbol{v}|_{\Gamma_s} - \boldsymbol{v}|_{\Gamma_m} \circ \Phi$$

of functions $v : \Gamma_m \cup \Gamma_s \to \mathbb{R}^d$, the non–penetration condition can be written as

$$[\boldsymbol{v}] \cdot \boldsymbol{n}_s \leq g \; . \tag{20}$$

Here, $\boldsymbol{n}_s$ denotes the outward normal on Γ_s, and $g : \Gamma_s \to \mathbb{R}$, $g \geq 0$, is the initial normal gap between Ω_m and Ω_s (see Figure 3). Condition (20) leads to the closed convex subset of admissible displacements

$$\mathcal{K}^* = \{\boldsymbol{v} \in H \mid [\boldsymbol{v}] \cdot \boldsymbol{n}_s \leq g \;\text{ a.e. on } \Gamma_s\}$$

of the solution space H,

$$H = H_m \times H_s \; , \qquad H_i = \{\boldsymbol{v} \in \left(H^1(\Omega)\right)^d \bigm| \boldsymbol{v}|_{\Gamma_{D,i}} = 0\} \; .$$

We consider the two body contact problem

$$\boldsymbol{u}^* \in \mathcal{K}^* : \quad \mathcal{J}(\boldsymbol{u}^*) \leq \mathcal{J}(\boldsymbol{v}) \quad \forall \boldsymbol{v} \in \mathcal{K}^* \tag{21}$$

with a quadratic energy $\mathcal{J}$ of the form (2). Denoting $\boldsymbol{v} = (\boldsymbol{v}_m, \boldsymbol{v}_s)$, $\boldsymbol{w} = (\boldsymbol{w}_m, \boldsymbol{w}_s) \in H$ the bilinear form $a(\cdot,\cdot)$ and the linear functional ℓ are given by

$$a(\boldsymbol{v}, \boldsymbol{w}) = a_m(\boldsymbol{v}_m, \boldsymbol{w}_m) + a_s(\boldsymbol{v}_s, \boldsymbol{w}_s) \; , \qquad \ell(\boldsymbol{v}) = \ell_m(\boldsymbol{v}_m) + \ell_s(\boldsymbol{v}_s) \; ,$$

where $a_i(\cdot,\cdot)$, ℓ_i are defined according to (14). For existence, uniqueness and regularity results, we refer, e.g., to [7, 16].

We now describe a finite element discretization of (21). In analogy to (15), the underlying finite element space $\mathcal{S}$, based on a triangulation $\mathcal{T} = \mathcal{T}_m \cup \mathcal{T}_s$ of $\Omega = \Omega_m \cup \Omega_s$, is spanned by the canonical nodal basis functions $\lambda_p \boldsymbol{E}^i$, $p \in \mathcal{N} = \mathcal{N}_m \cup \mathcal{N}_s$. We assume that Γ_s is resolved by $\mathcal{T}_s$ in the sense that Γ_s can be represented as the union of faces $F \in \mathcal{F}_{\Gamma_s}$ of elements $T \in \mathcal{T}_s$. Following [40], the constraints are then incorporated in a weak sense. More precisely, the set $\mathcal{K}^*$ is approximated by

$$\mathcal{K} = \{\boldsymbol{v} \in \mathcal{S} \mid ([\boldsymbol{v}] \cdot \boldsymbol{n}_s, \psi_q)_{\Gamma_s} \leq (g, \psi_q)_{\Gamma_s} \; \forall q \in \mathcal{N}_{\Gamma_s}\}$$

denoting $(v, w)_{\Gamma_s} = \int_{\Gamma_s} v(x) w(x) \; d\sigma$ and $\mathcal{N}_{\Gamma_s} = \mathcal{N}_s \cap \Gamma_s$. Dual basis functions ψ_q with the bi–orthogonality property

$$(\lambda_p|_{\Gamma_s}, \psi_q)_{\Gamma_s} = \delta_{p,q} \quad \forall p, q \in \mathcal{N}_{\Gamma_s}$$

have been introduced and analyzed in [39].

We now concentrate on monotone multigrid methods for the resulting discrete two body problem of the form (1). The crucial step towards a nodal splitting (4) is the following hierarchical decomposition

$$\mathcal{S} = \mathcal{V} \oplus \mathcal{W} \tag{22}$$

into the subspaces

$$\mathcal{V} = \{\boldsymbol{v} \in \mathcal{S} \mid ([\boldsymbol{v} \cdot \boldsymbol{E}^i])_{\Gamma_s} = 0 \; \forall i = 1, \ldots, d, \; p \in \mathcal{N}_{\Gamma_s}\} \; ,$$

$$\mathcal{W} = \{\boldsymbol{v} \in \mathcal{S} \mid \boldsymbol{v}(p) = 0 \; \forall p \in \mathcal{N} \setminus \mathcal{N}_{\Gamma_s}\} \; ,$$

as proposed in [40]. The space $\mathcal{V}$ is spanned by the nodal basis

$$\mu_p \boldsymbol{E}^i \; , \qquad i = 1, \ldots, d \; , \;\; p \in \mathcal{N} \setminus \mathcal{N}_{\Gamma_s} \; ,$$

with projected hat functions μ_p,

$$\mu_p = \Pi\lambda_p \ , \qquad \Pi v = v - \sum_{q\in\mathcal{N}_{\Gamma_s}} ([v],\psi_q)_{\Gamma_s}\lambda_q \ .$$

Obviously, we have $\mu_p = \lambda_p$ for $p \in \mathcal{N}\setminus\Gamma$. This choice of basis gives rise to the decomposition of $\mathcal{V}$ into the local subspaces

$$U_p = \operatorname{span}\{\mu_p\boldsymbol{E}^i \mid i = 1,\dots,d\} \ , \qquad p \in \mathcal{N}\setminus\mathcal{N}_{\Gamma_s} \ .$$

Together with (22) and the canonical splitting of $\mathcal{W}$ into the spaces

$$U_p = \operatorname{span}\{\lambda_p\boldsymbol{E}^i \mid i = 1,\dots,d\} \ , \qquad p \in \mathcal{N}_{\Gamma_s} \ ,$$

we get the decomposition (4) of the whole space $\mathcal{S}$. Observe that, by construction, $\mathcal{K}$ can be written in the form (3) with

$$\mathcal{K}_p = \{\boldsymbol{v} \in U_p \mid \boldsymbol{v}\cdot\boldsymbol{n}_s(p) \le (g,\psi_p)_{\Gamma_s}\} \qquad p \in \mathcal{N}_{\Gamma_s} \ ,$$

and $\mathcal{K}_p = U_p$, $p \in \mathcal{N}\setminus\mathcal{N}_{\Gamma_s}$, see [40].

In order to construct suitable coarse grid spaces $\mathcal{X}_k$, we assume that a hierarchy of triangulations $\mathcal{T}_0,\dots,\mathcal{T}_J = \mathcal{T}$ and a corresponding nested sequence of scalar finite element spaces, spanned by the nodal basis functions $\lambda_p^{(k)}$, $p \in \mathcal{N}_k = \mathcal{N}_{k,m}\cup\mathcal{N}_{k,s}$, is available. It is understood that Γ_s is the union of faces $F \in \mathcal{F}_{k,\Gamma_s}$. Setting

$$\mathcal{X}_k = \mathcal{V}_k \cup \mathcal{W}_k \ ,$$

we still have to define the spaces $\mathcal{V}_k$, $W_k \subset \mathcal{S}$. Let $\mathcal{N}_{k,\Gamma_s} = \mathcal{N}_{k,s}\cap\Gamma_s$. We introduce the projection operators

$$\Pi_k v = v - \sum_{q\in\mathcal{N}_{k,\Gamma_s}} ([v],\psi_q^{(k)})_{\Gamma_s}\lambda_q^{(k)} \ , \quad k = 0,\dots,J \ , \tag{23}$$

using the coarse grid dual basis functions $\psi_q^{(k)}$. Obviously,

$$\Pi_k\lambda_p^{(k)} = \lambda_p^{(k)} \qquad \forall p \in \mathcal{N}_k\setminus\Gamma \ . \tag{24}$$

Direct coarse grid analogues of $\mathcal{V}$ are generated by the basis functions $\Pi_k\lambda_p^{(k)}\boldsymbol{E}^i$. Unfortunately, such spaces are not nested in the case of non–matching triangulations. On the other hand, the functions $\Pi_k\lambda_p^{(k)}$, $p \in \mathcal{N}_{k,m}\cap\Gamma_m$, have large support in Ω_s, as desired according to common multigrid philosophy. This motivates the choice

$$\mathcal{V}_k = \sum_{p\in\mathcal{N}_k\setminus\mathcal{N}_{k,\Gamma_s}} U_p^{(k)} \ , \qquad U_p^{(k)} = \operatorname{span}\{\mu_p^{(k)}\boldsymbol{E}^i \mid i = 1,\dots,d\} \ ,$$

with functions $\mu_p^{(k)}$ defined by successive projection

$$\mu_p^{(k)} = \Pi_J\circ\Pi_{J-1}\circ\cdots\circ\Pi_k\lambda_p^{(k)} \ ,$$

see [40]. By construction, we have $\mathcal{V}_0 \subset \dots\mathcal{V}_{J-1} \subset \mathcal{V}_J = \mathcal{V}$. Moreover, the projected hat functions $\mu_p^{(k)}$ are characterized by the recursion formula

$$\mu_p^{(k)} = \sum_{q\in\mathcal{N}_{k+1}\setminus\mathcal{N}_{k+1,\Gamma_s}} \big(\Pi_k\lambda_p^{(k)}\big)(q)\mu_q^{(k+1)} \ . \tag{25}$$

We emphasize that this representation can be directly translated into restriction and prolongation operators. The spaces $\mathcal{W}_k$ are defined by

$$\mathcal{W}_k = \sum_{\mathcal{N}_{k,\Gamma_s}} U_p^{(k)} , \qquad U_p^{(k)} = \operatorname{span}\left\{ \big(\boldsymbol{\mu}_p^{(k)}\big)^i \mid i = 1, \dots, d \right\} ,$$

where the basis functions $\big(\boldsymbol{\mu}_p^{(k)}\big)^i$ are obtained recursively according to (16), (17). Observe that in general $\mathcal{W}_k \notin \mathcal{W}_{k+1}$, because the support of functions in $\mathcal{W}_k$ typically becomes larger with decreasing level k. However, utilizing (24), it is easily checked that each $\boldsymbol{v} \in \mathcal{W}_k$ can be represented as the sum $\boldsymbol{v} = \boldsymbol{v}_{k+1} + \boldsymbol{w}_{k+1}$ with $\boldsymbol{v}_{k+1} \in \mathcal{V}_{k+1}$ and $\boldsymbol{w}_{k+1} \in \mathcal{W}_{k+1}$. Hence, we get

$$\mathcal{X}_0 \subset \cdots \subset \mathcal{X}_{J-1} \subset \mathcal{X}_J = \mathcal{S} ,$$

as required.

Coarse grid constraints $\mathcal{D}_p^{(k)}$, $p \in \mathcal{N}_{k,\Gamma_s}$, of the form (19) can be derived by successive update and restriction, exactly in the same way as for Signorini's problem [23].

As in the previous examples, the resulting monotone multigrid method is convergent by Theorem 1 and can be reformulated as a multigrid V–cycle with projected block Gauß–Seidel smoother. Let us briefly comment on the implementation of the restriction and prolongation operators. As usual, functions $\boldsymbol{v} \in \mathcal{X}_k$ are represented algebraically in terms of the coefficients of the basis functions $\mu_p^{(k)} \boldsymbol{E}^i$, of $\mathcal{V}_k$ and $\big(\boldsymbol{\mu}_p^{(k)}\big)^i$ of $\mathcal{W}_k$, respectively. Coefficients related to $\mathcal{X}_k \subset \mathcal{X}_{k+1}$ are expressed by coefficients related to $\mathcal{X}_{k+1}$ by multiplication with the restriction matrix

$$\boldsymbol{R}_k = (r_{pq})_{p \in \mathcal{N}_k,\, q \in \mathcal{N}_{k+1}} , \qquad r_{pq} \in \mathbb{R}^{d \times d} .$$

The prolongation matrix is $\boldsymbol{P}_k = \big(\boldsymbol{R}_k\big)^T$. Using (25), we immediately get

$$r_{pq} = \big(\Pi_k \lambda_p^{(k)}\big)(q)\; I_d , \qquad p \in \mathcal{N}_k \setminus \mathcal{N}_{k,\Gamma_s} ,$$

with I_d denoting the unit matrix in $\mathbb{R}^d$. Let $\mathcal{N}_{k,\Gamma_m} = \mathcal{N}_{k,m} \cap \Gamma_m$. For $p \notin \mathcal{N}_{k,\Gamma_m}$, the identity (24) yields the canonical weights

$$\big(\Pi_k \lambda_p^{(k)}\big)(q) = \lambda_p^{(k)}(q) .$$

In the case $p \in \mathcal{N}_{k,\Gamma_m}$, we essentially have to evaluate the entries

$$([\lambda_p^{(k)}], \psi_q^{(k)})_{\Gamma_s} = -\int_{\Gamma_s} \big(\lambda_p^{(k)} \circ \Phi\big)(x) \psi_q^{(k)}(x)\, d\sigma , \qquad q \in \mathcal{N}_{k,\Gamma_s} . \tag{26}$$

For mortar elements on curvilinear boundaries we also refer to [11] and the references therein. We start out with the fine grid $k = J$. For an approximation by suitable quadrature rules, we need an algorithm to evaluate $\big(\lambda_p \circ \Phi\big)(x)$ for arbitrary $x \in \Gamma_s$. This is far from trivial, in particular for $d = 3$. We refer to [26] where normal projection is treated in some detail. In order to compute the coarse–grid entries by successive restriction, we use the embedding

$$\mathcal{M}_k = \operatorname{span}\{\psi_q^{(k)} \mid q \in \mathcal{N}_{k,\Gamma_s}\} \subset \widetilde{\mathcal{M}}_k = \{v : \Gamma_s \to \mathbb{R} \mid v|_F \text{ linear } \ \forall F \in \mathcal{F}_{k,\Gamma_s}\} , \tag{27}$$

because, in contrast to the larger spaces $\widetilde{\mathcal{M}}_k$, the spaces $\mathcal{M}_k$ are not nested. Selecting suitable basis functions $\widetilde{\psi}_q^{(k)}$ of $\widetilde{\mathcal{M}}_k$, the auxiliary coefficients $([\lambda_p^{(k)}], \widetilde{\psi}_q^{(k)})_{\Gamma_s}$ can be computed from $([\lambda_p], \widetilde{\psi}_q)_{\Gamma_s}$ by successive restriction. Finally, an additional restriction from $\widetilde{\mathcal{M}}_k$ to $\mathcal{M}_k$ according to the embedding (27) provides the desired

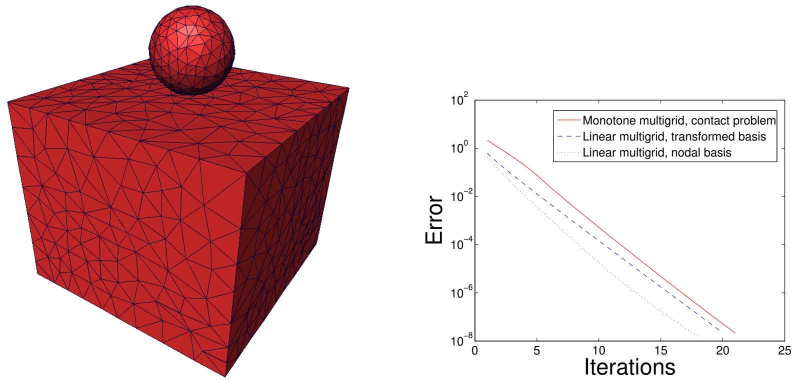

FIGURE 4. Initial triangulation $\mathcal{T}_0$ and iteration history of MMG on the final triangulation $\mathcal{T}_3$ in comparison with two multigrid methods for a corresponding linear problem

entries (26) on each level k. For the remaining Signorini–type nodes $p \in \mathcal{N}_{k,\Gamma_s}$, the evaluation of the corresponding entries r_{pq} is explained in [23].

2.4. **Numerical experiments.** In what follows we present two examples. First a rather simple model problem is used as a test for the algorithm as such. Second, in view of future medical applications, an example from biomechanics with complex 3D geometry from the Visible Human Data Set.

2.4.1. *A static two body Hertzian contact problem.* We consider the Hertzian contact of a unit ball Ω_s and a cuboid Ω_m with side lengths $6 \times 6 \times 5$. We prescribe a downward displacement of 0.1 at an upper polar cap $\Gamma_{D,s}$ of Ω_s and zero displacements at the vertical faces $\Gamma_{D,m}$ of Ω_m. After selecting a lower polar cap Γ_s of Ω_s the corresponding part Γ_m of $\partial\Omega_m$ is determined by normal projection of Γ_s onto $\partial\Omega_m$. The remaining part of the boundary is kept traction–free. The material parameters are $E = 2.5 \cdot 10^5$ and $\nu = 0.3$. Starting from the coarse grid $\mathcal{T}_0$ as depicted in the left picture of Figure 4 with 1278 vertices, we perform $J = 3$ local refinement steps to obtain the fine triangulation $\mathcal{T}_J$ with 76 765 vertices. To concentrate refinement on the contact boundary Γ, all tetrahedra which can be connected to Γ by at most 15 edges are refined in the usual way [6, 8]. The coarse grid problems on $\mathcal{T}_0$ are solved by the open–source interior point solver IPOPT [38] which converges much faster than usual projected block Gauß–Seidel relaxation. Implementation is realized in the framework of the DUNE library [1] using the grid manager of UG [3].

The right picture of Figure 4 illustrates the convergence behavior of the monotone multigrid method (MMG) as described in the preceding section. The solid curve shows the reduction of the approximate algebraic error of a $V(3,3)$ cycle in the course of iteration. The algebraic error is measured in the energy norm. The initial iterate is obtained by nested iteration. Once the contact points, i.e., the vertices $p \in \Gamma_s$ where contact actually occurs, are detected after some steps, MMG becomes a linear iteration for the resulting linear Dirichlet–type problem with vanishing normal displacements at the contact points. For a comparison, we consider the related Neumann–type problem where the resulting normal contact pressure is

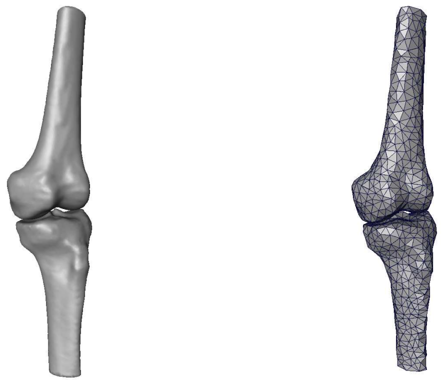

FIGURE 5. The computational domain Ω and its coarse grid approximation $\mathcal{T}_0$

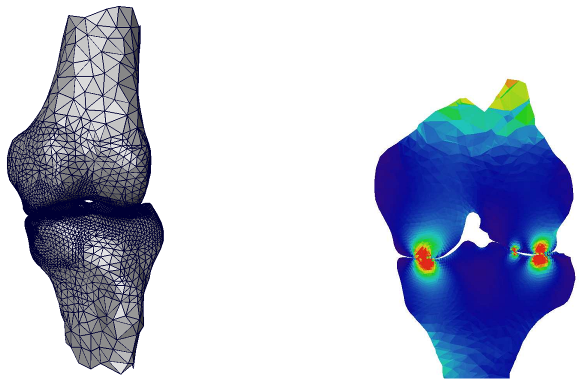

FIGURE 6. Section of the final triangulation $\mathcal{T}_3$ and von Mises stress along a cut through the contact boundary

prescribed at the contact boundary Γ. Observe that both the classical multigrid method based on the nodal basis and the hierarchical variant based on the splitting (22) exhibit very similar convergence behavior to MMG. The asymptotic convergence rates range from 0.45 (linear multigrid, nodal basis) to 0.41 (linear multigrid, hierarchical splitting) and 0.41 (MMG).

2.4.2. *A static two body contact problem from biomechanics.* The computational domain $\Omega = \Omega_m \cup \Omega_s$ is shown in the left picture of Figure 5. It represents the left proximal tibia Ω_m and distal femur Ω_s from the Visible Human Data Set [2]. The lower third $\Gamma_{D,m}$ of the boundary of the tibia Ω_m is clamped, whereas a downward displacement of 4 mm is prescribed at the upper third $\Gamma_{D,s}$ of the boundary of the femur Ω_s. The possible contact boundaries Γ_m and Γ_s at the upper part of the tibia and the lower part of the femur are determined in such a way that both can be identified by normal projection. Here, we use a straightforward algorithm described in [26]. The remaining parts of the boundaries are kept traction–free. The bone is assumed to be homogeneous, isotropic and linear elastic with material parameters $E = 17$GPa and $\nu = 0.3$ which are realistic choices for real life.

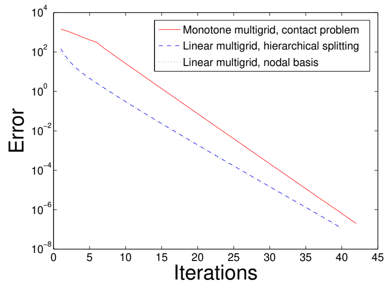

FIGURE 7. Iteration history of MMG on the final triangulation $\mathcal{T}_3$ in comparison with two multigrid methods for a corresponding linear problem

In order to construct a coarse triangulation $\mathcal{T}_0$, the given, highly resolved representation of the boundary surface $\partial\Omega$ is coarsened by successive point removals. The resulting coarse surface is shown in the right picture of Figure 5. It is used to generate $\mathcal{T}_0$ by an advancing front tetrahedral mesh generator [35]. The resulting triangulation $\mathcal{T}_0$ has 1662 vertices. To focus on the particular difficulties of the problem, local refinement again concentrates on the possible contact boundary Γ by selecting all tetrahedra which can be connected to Γ by at most 15 edges. However, in order to improve the approximation of $\partial\Omega$ in the course of refinement, the new vertices located on the approximate boundary are now shifted to the exact boundary $\partial\Omega$. A parametrization of $\partial\Omega$ is created simultaneously to the coarsening process [27]. To avoid a deterioration of mesh quality, the coarse triangulation $\mathcal{T}_0$ should be sufficiently fine, in particular in concave regions. If the quality of a new tetrahedron would still miss certain thresholds, then the corresponding vertex is not moved. Due to the displaced vertices the resulting sequence of triangulations $\mathcal{T}_0, \dots, \mathcal{T}_J$ is no longer nested. Nevertheless, we still use the weights introduced in the preceding section. For a theoretical justification in the linear self–adjoint case, we refer to [32]. Using IPOPT [38] as a coarse–grid solver, we can deal with the 4986 unknowns on the coarse grid $\mathcal{T}_0$ in a reasonable way. Implementation is realized in the framework of the DUNE library [1].

After $J = 3$ local refinement steps, we obtain the final triangulation $\mathcal{T}_J$ with 88 334 vertices. This leads to a discrete two body contact problem with 265 002 unknowns. The left picture in Figure 6 shows the distribution of the approximate von Mises stress field along a cut through the actual coincidence set. As in the preceding experiment, the right picture illustrates the convergence behavior of MMG. We always consider $V(3,3)$ cycles. First observe that the detection of the exact coincidence set still plays a minor role in the iteration history of MMG (solid line). As in the previous experiment, MMG almost immediately reduces to a linear multigrid method for the asymptotic linear Dirichlet–type problem with prescribed normal displacements at the contact points. The asymptotic convergence rate is 0.56. The curves for the two linear multigrid methods (nodal basis and hierarchical splitting) as applied to the related Neumann problem with prescribed contact pressure at the contact boundary Γ almost coincide. In both cases we observe the asymptotic convergence rate 0.62. Thus, the two body contact problem is still solved with linear convergence speed. The slowdown in comparison with the model problem above is due to the shifting process and the reduced shape regularity of the mesh.

Summarizing, even though this example is much more challenging than scalar obstacle problems or simple model problems in elasticity, the behavior of monotone multigrid is very much the same. This means that the method is robust even in highly complex situations.

3. Dynamic contact problems

We now turn to the time dependent case. As already mentioned above, implicit time discretization requires the solution of a static contact problem in each time step.

3.1. **A stabilized Newmark scheme.** Using the notation of Section 2.3 a solution of the dynamic two body contact problem in the time interval $[0, T_0]$ can be regarded as a stationary point of the action integral (unit density normalization)

$$\int_0^{T_0} \mathcal{L}(\mathbf{v}(t), \dot{\mathbf{v}}(t))\, dt\ , \qquad \mathcal{L}(\mathbf{v}, \dot{\mathbf{v}}) = \tfrac{1}{2}|\dot{\mathbf{v}}|^2 - \mathcal{J}(\mathbf{v}) - \chi(\mathbf{v})\ ,$$

with suitable initial and boundary conditions [31, 34]. The constraints $\mathbf{v}(t) \in \mathcal{K}^*$ for almost all $t \in [0, T_0]$ are enforced by the characteristic functional $\chi(\mathbf{v})$,

$$\chi(\mathbf{v}) = \begin{cases} 0\ , & \text{if } \mathbf{v} \in \mathcal{K}^* \\ \infty\ , & \text{else} \end{cases}\ , \qquad \mathbf{v} \in H\ .$$

The total energy of a state $\mathbf{v}$ is given by

$$\mathcal{E}(\mathbf{v}) = \tfrac{1}{2}|\dot{\mathbf{v}}|^2 + \mathcal{J}(\mathbf{v}) + \chi(\mathbf{v})\ .$$

From the unconstrained case, we expect the energy $\mathcal{E}(\mathbf{u}^*)$ of the solution $\mathbf{u}^*$ to be preserved throughout the evolution. The solution $\mathbf{u}^*$ satisfies the hyperbolic variational inequality [29, p. 81]

$$\mathbf{u}^*(\cdot, t) \in \mathcal{K}^*: \quad (\ddot{\mathbf{u}}^*, \mathbf{v} - \mathbf{u}^*) + a(\mathbf{u}^*, \mathbf{v} - \mathbf{u}^*) \geq \ell(\mathbf{v} - \mathbf{u}^*)\ , \quad \forall \mathbf{v} \in \mathcal{K}^*\ . \tag{28}$$

for each $t \in [0, T_0]$. Here and in the following we use the same notation $(\cdot, \cdot)$ for the scalar product in $\left(L^2(\Omega)\right)^d$ and for the pairing of H and its dual. It is convenient to represent the internal and external forces according to

$$(\mathbf{F}(\mathbf{w}), \mathbf{v}) = a(\mathbf{w}, \mathbf{v}) - \ell(\mathbf{v})\ , \qquad \mathbf{w}, \mathbf{v} \in H\ .$$

Then (28) can be reformulated as the inclusion

$$0 \in \ddot{\mathbf{u}}^* + \mathbf{F}(\mathbf{u}^*) + \boldsymbol{\partial\chi}(\mathbf{u}^*) \tag{29}$$

utilizing the subdifferential $\boldsymbol{\partial\chi}$ of χ (see, e.g., [10]). In the unconstrained case, $\mathcal{K}^* = H$, this inclusion reduces to Newton's equations of motion

$$\ddot{\mathbf{u}}^* = -\mathbf{F}(\mathbf{u}^*)\ . \tag{30}$$

We start by considering the time discretization of this differential equation in the dual of H. The classical Newmark scheme [29, p. 50] for Newton's equations (30) is based on Taylor expansions of the displacement $\mathbf{u}^*$ and of the velocity $\dot{\mathbf{u}}^*$ with respect to a given time step τ. Neglecting terms of higher order and introducing the traditional parameters $\gamma, 2\beta \in [0,1]$, this leads to

$$\begin{aligned} \mathbf{u}^{n+1} &= \mathbf{u}^n + \tau\dot{\mathbf{u}}^n + \tfrac{\tau^2}{2}\big((1-2\beta)\ddot{\mathbf{u}}^n + 2\beta\ddot{\mathbf{u}}^{n+1}\big) , \\ \dot{\mathbf{u}}^{n+1} &= \dot{\mathbf{u}}^n + \tau\big((1-\gamma)\ddot{\mathbf{u}}^n + \gamma\ddot{\mathbf{u}}^{n+1}\big) \end{aligned} \tag{31}$$

with $\mathbf{u}^n$ approximating $\mathbf{u}^*(\cdot, t_n)$, $t_n = n\tau$. Upon inserting (30) we get

$$\mathbf{u}^{n+1}_{\text{pred}} = \mathbf{u}^n + \tau\dot{\mathbf{u}}^n , \tag{32a}$$

$$\mathbf{u}^{n+1} = \mathbf{u}^{n+1}_{\text{pred}} - \tfrac{\tau^2}{2}\big((1-2\beta)\mathbf{F}(\mathbf{u}^n) + 2\beta\mathbf{F}(\mathbf{u}^{n+1})\big) , \tag{32b}$$

$$\dot{\mathbf{u}}^{n+1} = \dot{\mathbf{u}}^n - \tau\big((1-\gamma)\mathbf{F}(\mathbf{u}^n) + \gamma\mathbf{F}(\mathbf{u}^{n+1})\big) . \tag{32c}$$

For the symmetric case $\gamma = 2\beta = \frac{1}{2}$, the Newmark scheme (32) reduces to the trapezoidal rule, which is second order, stable and energy conserving [12, Chapter 10]. For $2\beta > 0$, a linear elasticity problem (32b) has to be solved in each time step. For $2\beta = 0$, the scheme is explicit and unstable. It can only be applied in a method of lines approach, i.e., after discretization in space first – observing a CFL–condition.

We are now ready to tackle the time discretization of our contact problem (29). In this case, the forces $\mathbf{F}(\mathbf{u}^*)$ have to be augmented with unknown contact forces $\mathbf{F}_{\text{con}} \in \boldsymbol{\partial\chi}(\mathbf{u}^*)$. Persistently implicit treatment of the contact forces, as proposed in [34], yields the following Newmark method

$$\mathbf{u}^{n+1}_{\text{pred}} = \mathbf{u}^n + \tau\dot{\mathbf{u}}^n , \tag{33a}$$

$$0 \in \mathbf{u}^{n+1} - \mathbf{u}^{n+1}_{\text{pred}} + \tfrac{\tau^2}{2}\big((1-2\beta)\mathbf{F}(\mathbf{u}^n) + 2\beta\mathbf{F}(\mathbf{u}^{n+1}) + \boldsymbol{\partial\chi}(\mathbf{u}^{n+1})\big) , \tag{33b}$$

$$\mathbf{F}^{n+1}_{\text{con}} = \tfrac{2}{\tau^2}(\mathbf{u}^{n+1}_{\text{pred}} - \mathbf{u}^{n+1}) - \big((1-2\beta)\mathbf{F}(\mathbf{u}^n) + 2\beta\mathbf{F}(\mathbf{u}^{n+1})\big) , \tag{33c}$$

$$\dot{\mathbf{u}}^{n+1} = \dot{\mathbf{u}}^n - \tau\big((1-\gamma)\mathbf{F}(\mathbf{u}^n) + \gamma\mathbf{F}(\mathbf{u}^{n+1}) + \mathbf{F}^{n+1}_{\text{con}}\big) . \tag{33d}$$

The inclusion (33b) can be rewritten as a two body contact problem of the form (21). The corresponding residual provides the current approximation of the contact forces $\mathbf{F}^{n+1}_{\text{con}} \in \boldsymbol{\partial\chi}(\mathbf{u}^{n+1})$ according to (33c). For what follows we fix the parameters to the symmetric case $\gamma = 2\beta = \frac{1}{2}$ which leads to a stable and dissipative scheme [9]. Energy is conserved under the condition that the number of contact points decreases in the course of evolution [9]. However, as confirmed by the numerical computations below, the above discretization tends to produce oscillations at the contact boundary: A short calculation shows that $\mathbf{u}^{n+1} = \mathbf{u}^n$ implies reversion of the velocity $\dot{\mathbf{u}}^{n+1} = -\dot{\mathbf{u}}^n$ so that new contact points are likely to be detached in the next but one time step.

In order to avoid such a kind of oscillations, we apply an additional L^2–projection in the prediction step. Thus we arrive at a *contact–stabilized Newmark method*, wherein (33a) is replaced by

$$0 \in \mathbf{u}^{n+1}_{\text{pred}} - (\mathbf{u}^n + \tau\dot{\mathbf{u}}^n) + \boldsymbol{\partial\chi}(\mathbf{u}^{n+1}_{\text{pred}}) , \tag{34}$$

while the remaining steps (33b) – (33d) are left unchanged. Under the assumption that $\mathbf{u}^{n+1} = \mathbf{u}^{n+1}_{\text{pred}} = \mathbf{u}^n$, we here compute $\dot{\mathbf{u}}^{n+1} = \dot{\mathbf{u}}^n$ and non–zero velocities are swallowed by the L^2–projection (34). Apart from that, (33) and the modified scheme have quite similar properties. We refer to [9] for a detailed investigation.

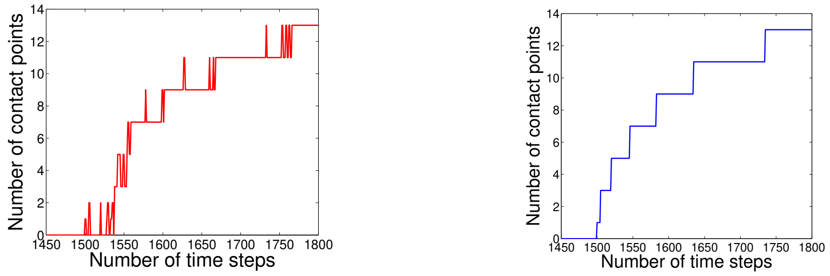

FIGURE 8. Number of contact points. *Left:* classical predictor (33a). *Right:* projected predictor (34).

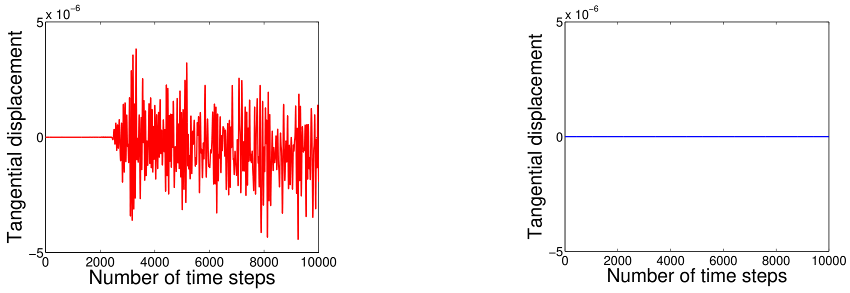

FIGURE 9. Tangential displacements at the south pole of Ω_s. *Left:* classical predictor (33a). *Right:* projected predictor (34).

3.2. **Numerical experiments.** As in the stationary case above (cf. Section 2.4) we present numerical results for a simpler 2D model problem and a biomechanical application.

3.2.1. *A dynamic two body Hertzian contact problem.* In order to compare the numerical properties of the symmetric Newmark scheme (33) with the contact–stabilized variant, we consider a simple model problem in two space dimensions. At initial time $t = 0$, two circles Ω_s, Ω_m with radii $r = 8$, midpoints on the y–axis, and distance 1.5 are moving with vertical speed $v_{0,s} = -1$, $v_{0,m} = +1$, respectively. The contact boundaries consist of quarter circles located at the bottom of Ω_s and at the top of Ω_m, respectively. The remaining part of the boundary is traction–free and no volume forces occur. We choose the material parameters $E = 5 \cdot 10^2$ and $\nu = 0.3$. The computations are carried out until $T_0 = 5$.

We select uniform time steps $\tau = 5 \cdot e^{-4}$. As explained in Sections 2.2 and 2.3, the spatial discretization is performed by finite elements utilizing a lumped version of the L^2–scalar product. The underlying triangulation $\mathcal{T}_J$ with 1035 vertices is resulting from $J = 6$ refinement steps of a coarse triangulation $\mathcal{T}_0$ with 8 vertices. The discretized analogues of the spatial two body contact problems (33b) are solved by monotone multigrid (cf. Section 2.3). Figure 8 shows the number of contact points, i.e., the number of vertices $p \in \Gamma_s$ where contact actually occurs, over a typical section of time steps. On the left, the oscillating behavior of the Newmark scheme (33) is reflected by the comparative number of contact points. The right picture clearly

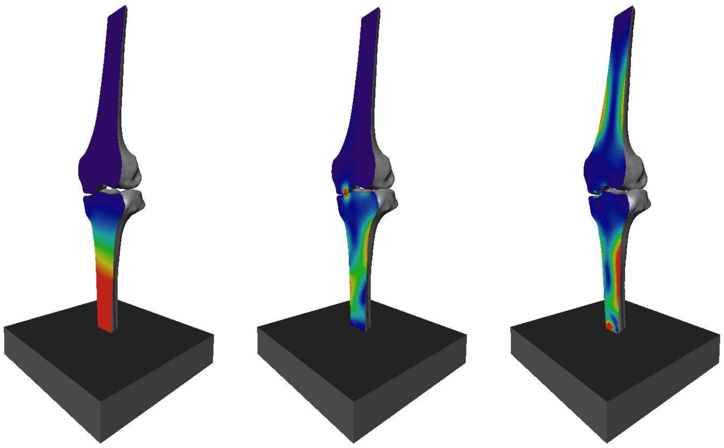

FIGURE 10. Von Mises stresses along a plane cut. *Left:* at contact with the rigid plane ($n = 6$). *Middle:* first contact of tibia and femur ($n = 21$). *Right:* after detachment of tibia and femur ($n = 50$).

displays the contact–stabilizing effect of our suggested prediction step (34). This observation is confirmed by Figure 9 showing the tangential displacements of the point at the south pole of Ω_m. While the Newmark scheme (33) is generating spurious oscillations, the contact–stabilized version (34) reproduces the behavior of the exact solution. Both schemes generated a slight energy loss of less than 0.2%. Moreover, there is numerical evidence that both schemes are dissipative and that the energy loss depends on the variation of the contact points and the time step. A detailed analysis of the contact–stabilized Newmark scheme is left to a forthcoming paper [9].

3.2.2. *A dynamic two body contact problem from biomechanics.* At the initial time $t = 0$, the left proximal tibia Ω_m and distal femur Ω_s are located according to the left picture in Figure 5 with a common downward velocity $v_0 = -1\text{m/sec}$ and a distance of 0.5mm of the tibia to a rigid plane below. We select realistic material parameters $E = 17\text{GPa}$, $\nu = 0.3$ (cf. Section 2.4.2) and density $\rho = 2\text{g/cm}^3$. The computations are carried out until $T_0 = 0.01\text{sec}$.

The resulting dynamic two body contact problem is discretized in time by the second–order contact–stabilized Newmark scheme. We select uniform time steps $\tau = 10^{-4}$. In passing, we mention that the explicit Newmark scheme ($2\beta = 0$) would require time steps of the order of 10^{-7} for stability reasons. The spatial problems (33b) are discretized by finite elements as explained in 2.4.2 utilizing a lumped version of the L^2–scalar product. We now perform $J = 2$ uniform refinement steps to the initial triangulation $\mathcal{T}_0$. Figure 10 illustrates the evolution of the displacements and the internal stresses on a fixed cutting plane. The discretized static two body problems are solved by the monotone multigrid method described in Section 2.3 using a $V(3,3)$ cycle. In each time step the multigrid iteration is terminated as soon as the relative algebraic error in the energy norm is less than 10^{-7} to avoid that accumulated errors affect the accuracy of the discretization. Figure 11 shows the required number of iteration steps over the time steps. Once

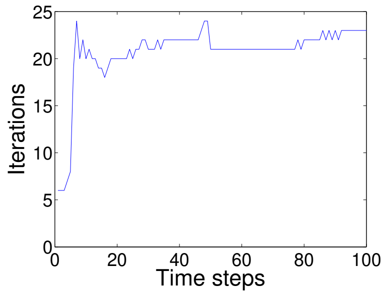

FIGURE 11. Number of required multigrid iterations over time

a first contact has occurred, the number of iterations mostly ranges from 20 to 25 steps which translates into averaged convergence rates of about 0.5.

References

[1] DUNE – Distributed and Unified Numerics Environment. `http://hal.iwr.uni-heidelberg.de/dune/`.

[2] The Visible Human Project. `http://www.nlm.nih.gov/research/visible/visible_human.html`.

[3] P. Bastian, K. Birken, K. Johannsen, S. Lang, N. Neuß, H. Rentz–Reichert, and C. Wieners. UG – a flexible software toolbox for solving partial differential equations. *Computing and Visualization in Science*, 1:27–40, 1997.

[4] Y. Bei and B.J. Fregly. Multibody dynamic simulation of knee contact mechanics. *Med. Engrg. Phys.*, 26:777–789, 2004.

[5] C. Bernardi, Y. Maday, and A. Patera. A new nonconforming approach to domain decomposition: The mortar element method. In H. B. et al., editors, *Nonlinear Partial Differential Equations and their Applications*, pages 269–286, Dordrecht, 1993. Reidel.

[6] J. Bey. *Finite-Volumen- und Mehrgitterverfahren für elliptische Randwertprobleme*. Teubner, Stuttgart, 1998.

[7] P. Boieri, F. Gastaldi, and D. Kinderlehrer. Existence, uniqueness, and regularity results for the two–body contact problem. *Appl. Math. Opt.*, 15:251–277, 1987.

[8] F.A. Bornemann, B. Erdmann, and R. Kornhuber. Adaptive multilevel methods in three space dimensions. *Int. J. Numer. Meth. Engrg.*, 36:3187–3203, 1993.

[9] P. Deuflhard, R. Krause, and S. Ertel. A contact–stabilized Newmark method for dynamical contact problems. In preparation.

[10] I. Ekeland and R. Temam. *Convex Analysis and Variational Problems*. North–Holland, Amsterdam, 1976.

[11] B. Flemisch, M.A. Puso, and B.I. Wohlmuth. A new dual mortar method for curved interfaces: 2D elasticity. *Int. J. Numer. Meth. Engrg.*, 63(6):813–832, 2005.

[12] M. Geradin and A. Cardona. *Flexible Multibody Dynamics: A Finite Element Approach*. John Wiley & Sons, 2001.

[13] R. Glowinski. *Numerical Methods for Nonlinear Variational Problems.* Springer, New York, 1984.

[14] R. Glowinski and P. Le Tallec. *Augmented Lagrangian and Operator–Splitting Methods in Nonlinear Mechanics.* SIAM, Philadelphia, 1989.

[15] W. Hackbusch and H.D. Mittelmann. On multigrid methods for variational inequalities. *Numer. Math.*, 42:65–76, 1983.

[16] J. Haslinger and I. Hlaváček. Contact between elastic bodies I. Continuous problems. *Apl. Mat.*, 25:324–327, 1980.

[17] C. Hellmich, F.-J. Ulm, and L. Dormieux. Can the diverse elastic properties of trabecular and cortical bone be attributed to only a few tissue-independent phase properties and their interactions? - Arguments from a multiscale approach. *J. Biomech.*, 2:219–238, 2004.

[18] I. Hlaváček, J. Haslinger, J. Nečas, and J. Lovíšek. *Solution of Variational Inequalities in Mechanics.* Springer, Berlin, 1988.

[19] R.H.W. Hoppe. Multigrid algorithms for variational inequalities. *SIAM J. Numer. Anal.*, 24:1046–1065, 1987.

[20] E. Keeve and R. Kikinis. Biomechanic based simulation of knee dynamics. In *Proceedings of the BMES/EMBS Conference 1999*, volume 1, page 558. 1999.

[21] N. Kikuchi and J.T. Oden. *Contact Problems in Elasticity.* SIAM, Philadelphia, 1988.

[22] R. Kornhuber. Monotone multigrid methods for elliptic variational inequalities I. *Numer. Math*, 69:167–184, 1994.

[23] R. Kornhuber and R. Krause. Adaptive multigrid methods for Signorini's problem in linear elasticity. *Comp. Visual. Sci.*, 4:9–20, 2001.

[24] R. Kornhuber and H. Yserentant. Multilevel methods for elliptic problems on domains not resolved by the coarse grid. *Cont. Math.*, 180:49–60, 1994.

[25] R. Krause. From inexact active set strategies to nonlinear multigrid methods. Uni Bonn, 2006. INS–Preprint 0602.

[26] R. Krause and O. Sander. Fast solving of contact problems on complicated geometries. In R. Kornhuber et al., editors, *Domain Decomposition Methods in Science and Engineering*, pages 495–502. Springer Verlag, 2005.

[27] R. Krause and O. Sander. Automatic construction of boundary parametrizations for geometric multigrid solvers. *Comp. Visual. Sci.*, 9:11–22, 2006.

[28] K. Kunisch and G. Stadler. Generalized Newton methods for the 2D Signorini contact problem with friction in function space. M^2AN, 39:827–854, 2005.

[29] T.A. Laursen. *Computational Contact and Impact Mechanics.* Springer, Berlin, Heidelberg, New York, 2003.

[30] J. Mandel. A multilevel iterative method for symmetric, positive definite linear complementarity problems. *Appl. Math. Optimization*, 11:77–95, 1984.

[31] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numerica*, pages 357–514, 2001.

[32] N. Neuss and Ch. Wieners. Criteria for the approximation property for multigrid methods in nonnested spaces. *Math. Comp.*, 73:1583–1600, 2004.

[33] N.M. Newmark. A method of computation for structural dynamics. *J. Eng. Div. Proc., ASCE*, 85:67–94, 1959.

[34] A. Pandolfi, C. Kane, J. E. Marsden, and M. Ortiz. Time–discretized variational formulation of non-smooth frictional contact. *Int. J. Numer. Meth. Engrg.*, 53:1801–1829, 2002.

[35] D. Stalling, M. Westerhoff, and H.-C. Hege. Amira: A highly interactive system for visual data analysis. In C.D. Hansen and C.R. Johnson, editors, *The Visualization Handbook*, chapter 38, pages 749–767. Elsevier, 2005.

[36] W.R. Taylor, R. Ehrig, M.O. Heller, H. Schell, P. Seebeck, and G.N. Duda. Tibio–femoral joint contact forces in sheep. *J. Biomech.*, to appear.

[37] M. Ulbrich. Nonsmooth Newton-like methods for variational inequalities and constrained optimization problems in function spaces. Fakultät für Mathematik, Technische Universität München, 2002. Habilitationsschrift.

[38] A. Wächter and L. T. Biegler. On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming. *Math. Progr.*, 106(1):25–57, 2006.

[39] B. Wohlmuth. A mortar finite element method using dual spaces for the Lagrange multiplier. *SINUM*, 38:989–1014, 2000.

[40] B. Wohlmuth and R. Krause. Monotone methods on nonmatching grids for nonlinear contact problems. *SISC*, 25(1):324–347, 2003.

[41] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Review*, 34:581–613, 1992.

[42] H. Yserentant. Old and new convergence proofs for multigrid methods. *Acta Numerica*, pages 285–326, 1993.

[43] O. C. Zienkiewicz and R. L. Taylor. *The Finite Element Method*, volume 2. McGraw–Hill Book Company, London, 4. edition, 1991.

Fachbereich Mathematik und Informatik, Freie Universität Berlin, Arnimallee 2–6, 14195 Berlin, Germany

Institute for Numerical Simulation — Institute for Applied Mathematics, University of Bonn, Wegeler Str. 6, 53115 Bonn, Germany

Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB), Division Scientific Computing, Department Numerical Analysis and Modelling, Takustr. 7, 14195 Berlin, Germany