

Algebraic multigrid methods for saddle point systems arising from mortar contact formulations

T. A. Wiesner^{1,‡}, M. Mayr^{2,3,*}, A. Popp^{2,‡}, M. W. Gee⁴ and W. A. Wall⁵

¹ *Leica Geosystems AG, Heinrich-Wild-Strasse 201, 9435 Heerbrugg, Switzerland*

² *Institute for Mathematics and Computer-Based Simulation, University of the Bundeswehr Munich, Werner-Heisenberg-Weg 39, D-85577 Neubiberg, Germany*

³ *Data Science & Computing Lab, University of the Bundeswehr Munich, Werner-Heisenberg-Weg 39, D-85577 Neubiberg, Germany*

⁴ *Mechanics & High Performance Computing Group, Technical University of Munich, Parkring 35, D-85748 Garching, Germany*

⁵ *Institute for Computational Mechanics, Technical University of Munich, Boltzmannstr. 15, D-85748 Garching, Germany*

SUMMARY

In this paper, a fully aggregation-based algebraic multigrid strategy is developed for nonlinear contact problems of saddle point type using a mortar finite element approach. While the idea of extending multigrid methods to saddle point systems can already be found, e.g., in the context of Stokes and Oseen equations in literature, the main contributions of this work are (i) the development and open-source implementation of an interface aggregation strategy specifically suited for generating Lagrange multiplier aggregates that are required for coupling structural equilibrium equations with contact constraints and (ii) a review of saddle point smoothers in the context of constrained interface problems. The new interface aggregation strategy perfectly fits into an aggregation-based multigrid framework and can easily be combined with segregated transfer operators, which allow to preserve the saddle point structure on the coarse levels. Further analysis provides insight into saddle point smoothers applied to contact problems, while numerical experiments illustrate the robustness of the new method. We have implemented the proposed algorithm within the MUELU package of the open-source TRILINOS project. Numerical examples demonstrate the robustness of the proposed method in complex dynamic contact problems as well as its scalability up to 23.9 million unknowns on 480 MPI ranks.

Received ...

KEY WORDS: Algebraic multigrid methods, Contact mechanics, Mortar methods, Iterative linear solvers, Preconditioning

1. INTRODUCTION

Many engineering applications require the simulation of large-scale contact problems. Therefore, it is not surprising that recent years have seen significant progress in modelling and simulation of contact interaction and its associated phenomena, such as friction [28, 52, 66], wear [16, 24, 44, 48], adhesion [45, 65], or multi-scale contact phenomena [10, 73]. This is particularly true with regard to robust finite element based discretization techniques for finite deformations and efficient nonlinear solution algorithms. Above all, mortar finite element methods — originally introduced in the context of domain decomposition [6, 9] — are meanwhile well-established as a basis for state-of-the-art contact

*Correspondence to: M. Mayr, Institute for Mathematics and Computer-Based Simulation, University of the Bundeswehr Munich, Werner-Heisenberg-Weg 39, D-85577 Neubiberg, Germany, E-mail: matthias.mayr@unibw.de

‡This work was partially performed while these authors were affiliated with the Institute for Computational Mechanics, Technical University of Munich, Boltzmannstr. 15, D-85748 Garching, Germany.

formulations and widely accepted among researchers as being superior to more classical discretization techniques, such as the node-to-segment (NTS) method, the Gauss-point-to-segment (GPTS) method and other collocation based approaches [21, 59, 60, 89, 93].

Nowadays, constraint enforcement in the context of mortar methods is often based on a regularized Lagrange multiplier scheme or an augmented Lagrange method instead of a simple, yet often insufficient penalty approach. Independent from the actual details of the constraint enforcement implementation, the discrete Lagrange multipliers constitute an additional set of degrees of freedom in the mortar finite element contact formulation. When using a dual mortar approach [33, 36, 37, 54, 57, 87, 88], the discrete Lagrange multiplier basis is chosen based on a biorthogonality condition with the underlying finite element basis. This allows for the localization of the contact constraints and, thus, from a more algebraic point of view, for the trivial condensation of the additional Lagrange multiplier degrees of freedom from the final linearized systems of equations. If such a static condensation is not desired or not feasible (e.g. when choosing a standard basis rather than dual basis functions for the Lagrange multipliers, see e.g. [56, 88]), the linear system remains in its generalized saddle point format arising from the contact constraint equations. Both the standard and the dual mortar approach have become increasingly popular in recent years, with new contributions focusing for example on higher-order finite element interpolation [55, 90], isogeometric mortar methods [20, 67, 69, 70, 94], or improved robustness of the solution algorithms [25, 34, 58], to name only a few particularly active research directions.

It is striking, however, that almost all current research endeavors concerned with mortar finite element methods for contact mechanics focus exclusively on the modelling of various contact phenomena. Yet, for large-scale and industrial applications the appropriate modelling of contact problems is not sufficient. In fact, the demand for efficient solution strategies tailored to the specifics of contact simulations is eminent in order to achieve optimal over-all performance. Whereas one could use parallel direct solvers to solve the linear systems, they are not an option for very large problems. Iterative solvers for sparse systems (e.g. [31, 63]) combined with good preconditioning strategies are a far better choice with respect to computational resources. In particular, multigrid methods [32, 72] are known to be among the most efficient solution and preconditioning strategies, at least for certain classes of problems.

From the perspective of the linear solvers and multigrid-based preconditioners, the condensation of the Lagrange multipliers seems to be very attractive, since it allows to circumvent the more-sophisticated saddle point formulation. For contact problems though, we have experienced that the resulting linear systems after condensation suffer from some challenging matrix properties which cause severe convergence problems for standard preconditioning techniques. In particular, the matrices tend to be non-diagonally dominant due to different (local) coordinate systems that are typically used for the formulation of the structural equilibrium equations and the contact constraints. In our previous work [85], we have developed multilevel preconditioners that address such issues and are specifically tailored to contact problems using the dual mortar method in a condensed formulation.

On the other hand, multigrid methods already have been successfully applied to saddle point problems as they arise from different applications (e.g. Stokes flow [38] or incompressible Navier-Stokes problems [30, 46]) and even in the context of mortar finite element methods [83, 86]. The multigrid theory for this particular class of saddle point problems has evolved starting from special multigrid methods for mortar finite element methods (e.g. [12, 29, 41, 92]) to mortar finite element methods in saddle point formulation (e.g., [13, 14]). Based on these ideas, specific multigrid methods for contact problems in saddle point formulation have been developed in [82]. However, most of the literature available on multigrid for mortar finite element methods and contact problems in saddle point formulations is primarily on geometric multigrid methods with abundant work on saddle point smoothers (cf. [97]). A first algebraic multigrid preconditioner for mortar-based contact problems has been proposed by Adams [5], performing standard aggregation on the graph of an auxiliary matrix imitating the Lagrange multipliers. Alternatively, multigrid methods for contact problems not requiring an outer iteration loop or active set strategy have been developed in [40, 91].

In this paper, we address the case of mortar-based contact problems in saddle point formulation and show how to tailor iterative solvers with algebraic multigrid preconditioners to such problems. In contrast to geometric multigrid methods, algebraic multigrid methods (e.g. [68]) do not rely on geometric user-provided mesh information, but use only purely algebraic information from the fine level matrix. Since static condensation of Lagrange multiplier unknowns is not required, our approach is applicable to mortar methods using both standard or dual shape functions. The proposed multigrid method is based on the (smoothed) aggregation algebraic multigrid algorithms (cf. [64, 75, 76, 78, 79]) with special extensions for block matrices and some minor contact-specific adaptations. We propose a novel aggregation strategy for the discrete Lagrange multiplier unknowns along the contact interface, which we consider simpler to implement, computationally less expensive, and more intuitive for

contact problems compared to the ideas from [5]. Inspired by our prior work on fluid-structure interaction [27, 47], where we have investigated the beneficial effect of satisfying interface constraints within the preconditioner, we will then use segregated transfer operators suitable for block matrices to transfer and incorporate the contact constraints in all coarse levels. We analyze various Schur complement block smoothers and assess their suitability for satisfying the contact constraints. Finally, we demonstrate and assess the performance of the proposed preconditioner in several three-dimensional examples.

The remainder of this paper is organized as follows: Section 2 provides a brief introduction to mortar methods for finite deformation contact problems in saddle point formulation. After the basic notation is introduced, we specifically present the resulting linear system that is arising if the discrete Lagrange multipliers are explicitly included into the set of unknowns to be solved for. After a brief introduction to the general idea of multigrid methods, Section 3 describes our strategy to tailor a multigrid preconditioner to contact problems in saddle point formulation. It comprises the coarsening of the mortar contact constraints as detailed in Section 4 as well as suitable block smoothers as discussed in Section 5. Finally, Section 6 presents numerical examples that showcase the robustness, scalability, and performance of the proposed multigrid preconditioners, before we close with some final remarks.

2. MORTAR METHODS FOR FINITE DEFORMATION CONTACT

As this paper is concerned with preconditioning of the system of linear equations arising from contact problems, just a brief summary to the contact formulation and discretization is given here. For a detailed presentation, the reader is referred to our previous work [54].

2.1. Problem formulation and governing equations

We consider two solid bodies, which are represented by $\Omega_0^{(1)}, \Omega_0^{(2)} \subset \mathbb{R}^d$ with $d \in \{2, 3\}$ in the reference configuration. Their surfaces $\partial\Omega_0^{(i)}, i \in \{1, 2\}$ are decomposed into three disjoint subsets $\Gamma_D^{(i)}, \Gamma_N^{(i)}$ and $\Gamma_c^{(i)}$ denoting the Dirichlet boundary, the Neumann boundary, and the potential contact interface with unknown contact tractions $\mathbf{t}_c^{(i)}$, respectively. The solid bodies themselves are governed by nonlinear elasticity. Since we are only interested in the algebraic block structure of the final system of equations after discretization and linearization, it is sufficient to discuss a quasi-static contact problem with only two deformable bodies.

In order to describe the contact phenomenon, we state the Hertz-Signorini-Moreau conditions

$$g_n \geq 0 \quad \wedge \quad p_n \leq 0 \quad \wedge \quad g_n p_n = 0. \quad (1)$$

Therein, g_n defines a so-called gap function, which measures the distance of a point on the slave interface $\gamma_c^{(S)}$ to the projected corresponding point on the master side $\gamma_c^{(M)}$ of the contact interface in the current configuration. Furthermore, p_n denotes the normal contact traction. In the mathematical formulation, one introduces the negative slave side contact traction $\mathbf{t}_c^{(1)}$ as Lagrange multiplier, i.e., $\boldsymbol{\lambda} = -\mathbf{t}_c^{(1)}$. Using \mathbf{n} to denote the outward unit normal vector, the normal part of the contact stress can be denoted by $\lambda_n := \boldsymbol{\lambda}^T \mathbf{n}$ and the tangential part by $\boldsymbol{\lambda}_\tau := \boldsymbol{\lambda} - \lambda_n \mathbf{n}$.

We employ the usual function spaces $\mathcal{U}^{(i)}$ and $\mathcal{V}^{(i)}$ for the displacement field \mathbf{u} of the the solid body and its weighting function \mathbf{v} , respectively. Furthermore, a suitable function space \mathcal{M}_+ for the Lagrange multiplier field $\boldsymbol{\lambda}$ and its weighting function $\boldsymbol{\mu}$ is assumed. The weak form of the governing equations then reads: Find $(\mathbf{u}, \boldsymbol{\lambda}) \in \mathcal{U} \times \mathcal{M}_+$ such that

$$-\delta\mathcal{W}_{\text{int,ext}} + \int_{\gamma_c^{(S)}} \boldsymbol{\lambda} \left(\mathbf{v}^{(1)} - \mathbf{v}^{(2)} \right) dA = 0, \quad \forall \mathbf{v} \in \mathcal{V}, \quad (2a)$$

$$\int_{\gamma_c^{(S)}} (\mu_n - \lambda_n) g_n dA \geq 0, \quad \forall \boldsymbol{\mu} \in \mathcal{M}_+. \quad (2b)$$

Herein, the internal and external virtual work contributions $\delta\mathcal{W}_{\text{int,ext}}$ are defined as usual in nonlinear solid mechanics (cf. [96] for example) and, thus, further details are omitted. The second term in (2a) can be identified as contact virtual work $\delta\mathcal{W}_c$ and the expression in (2b) as variational inequality formulation of the contact constraints. An extension to frictional contact based on Coulomb's law is straightforward and can be found in our previous work [28] for example.

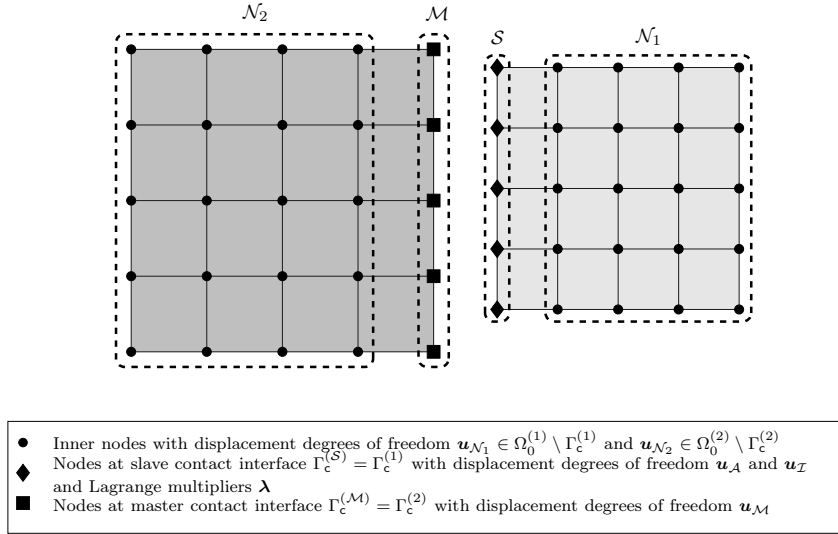


Figure 1. Schematic mesh illustrating interior, master and slave interface nodes.

2.2. Finite element discretization

For spatial discretization of the displacement field, either isoparametric finite elements with first-order and second-order Lagrange interpolation or isogeometric analysis (IGA) with NURBS-based shape functions are employed. After discretization, the discrete representation of displacement unknowns is given by the nodal degrees of freedom (DOFs)

$$\mathbf{u} = [\mathbf{u}_{\mathcal{N}_1}, \mathbf{u}_{\mathcal{S}}, \mathbf{u}_{\mathcal{M}}, \mathbf{u}_{\mathcal{N}_2}]^T.$$

Therein, the $\mathbf{u}_{\mathcal{N}_i}$, $i \in \{1, 2\}$, contain all degrees of freedom associated with the mesh nodes of the corresponding solid body without the nodes at the contact interface, where we use the convention that indices 1 and 2 denote the slave and master “body”, respectively. The degrees of freedom associated with the contact interface on the slave and master side are represented by $\mathbf{u}_{\mathcal{S}}$ and $\mathbf{u}_{\mathcal{M}}$, respectively.

Since (2) represents a mixed variational form, we also have to discretize the Lagrange multiplier field $\boldsymbol{\lambda}$. We choose to follow a mortar approach for its mathematical properties and its superiority to other schemes [21, 59, 60, 89, 93]. As usual for mortar methods, the Lagrange multiplier field is discretized on the slave side contact interface $\gamma_c^{(S)}$ in the current configuration. We either use standard ansatz functions, i.e. Lagrange polynomials with a trace space relation with the underlying volume element, or dual shape functions. The latter satisfy a biorthogonality property and, thus, allow for a computationally cheap condensation of the additional unknown Lagrange multipliers from the final system of equations. For details on dual basis functions in the context of mortar-based contact discretizations, we refer to [42, 55, 88]. Their interplay with preconditioners for iterative linear solvers has been discussed in our previous work [85]. The vector of discrete Lagrange multipliers is now referred to as $\boldsymbol{\lambda}$. A schematic mesh illustrating interior, slave interface, and master interface nodes is sketched in Figure 1.

The final spatially discretized formulation of the quasi-static frictionless problem (2) using the nodal vector representation now emerges as

$$\mathbf{f}_{\text{int}}(\mathbf{u}) + \mathbf{f}_{\text{co}}(\mathbf{u}, \boldsymbol{\lambda}) = \mathbf{f}_{\text{ext}}, \quad (3a)$$

$$(\tilde{g}_{n,h})_j \geq 0, \quad (\lambda_n)_j \geq 0, \quad (\tilde{g}_{n,h})_j (\lambda_n)_j = 0, \quad j = 1, \dots, n^{(S)}, \quad (3b)$$

$$(\boldsymbol{\lambda}_\tau)_j = 0, \quad j = 1, \dots, n^{(S)}. \quad (3c)$$

The internal forces $\mathbf{f}_{\text{int}}(\mathbf{u})$ and external forces \mathbf{f}_{ext} are common in nonlinear finite element methods and need no further explanation. The discrete vector of contact forces \mathbf{f}_{co} is computed based on two mortar matrices \mathbf{D} and \mathbf{M} , arising from the integral over the slave interface $\gamma_c^{(S)}$ in (2a), and the discrete Lagrange multiplier vector $\boldsymbol{\lambda}$. For details regarding the computation of \mathbf{D} and \mathbf{M} , see [54, 59]

for example. Using $(\tilde{g}_{n,h})_j$ to denote the discrete weighted gap function at node j , a closer look at the discrete contact constraints reveals that (3b) basically represents a discrete version of the Karush–Kuhn–Tucker (KKT) type conditions in (1) with an additional weighting based on the Lagrange multiplier shape functions ψ_j , while the nodal enforcement of frictionless sliding in (3c) is straightforward anyway.

Since the discrete contact constraints summarized in (3b) are still formulated as inequalities, an active set strategy usually referred to as primal-dual active set strategy (PDASS) is needed in addition to the usual nonlinear solution procedure to identify the currently active and inactive contact regions \mathcal{A} and $\mathcal{I} = \mathcal{S} \setminus \mathcal{A}$, respectively. It has been demonstrated in [17, 35, 61] that the PDASS can equivalently be interpreted as a semi-smooth Newton method, thus allowing for an integrated treatment of all nonlinearities (including the search for the active set) within one single Newton–Raphson type iteration loop. Meanwhile, many successful applications to small and large deformation contact problems can be found in the literature [28, 36, 37, 54].

2.3. Algebraic formulation of linear systems

For efficient iterative solution strategies based on multigrid methods for nonlinear contact problems, one is primarily interested in the structure of the linear systems arising in each nonlinear iteration step of the underlying Newton–Raphson scheme. For the sake of brevity, details on the linearization process and on the Newton–Raphson procedure are omitted here and the interested reader is instead referred to [54, 57].

Consistent linearization of (3) and a subsequent update of the active set \mathcal{A} and inactive set \mathcal{I} yields the system

$$\begin{pmatrix} K_{\mathcal{N}_1\mathcal{N}_1} & K_{\mathcal{N}_1\mathcal{M}} & 0 & 0 & 0 & 0 & 0 \\ K_{\mathcal{M}\mathcal{N}_1} & K_{\mathcal{M}\mathcal{M}} & K_{\mathcal{M}\mathcal{I}} & K_{\mathcal{M}\mathcal{A}} & 0 & -M_{\mathcal{I}}^T & -M_{\mathcal{A}}^T \\ 0 & K_{\mathcal{I}\mathcal{M}} & K_{\mathcal{I}\mathcal{I}} & K_{\mathcal{I}\mathcal{A}} & K_{\mathcal{I}\mathcal{N}_2} & D_{\mathcal{I}\mathcal{I}}^T & D_{\mathcal{I}\mathcal{A}}^T \\ 0 & K_{\mathcal{A}\mathcal{M}} & K_{\mathcal{A}\mathcal{I}} & K_{\mathcal{A}\mathcal{A}} & K_{\mathcal{A}\mathcal{N}_2} & D_{\mathcal{A}\mathcal{I}}^T & D_{\mathcal{A}\mathcal{A}}^T \\ 0 & 0 & K_{\mathcal{N}_2\mathcal{I}} & K_{\mathcal{N}_2\mathcal{A}} & K_{\mathcal{N}_2\mathcal{N}_2} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & N_{\mathcal{M}} & N_{\mathcal{I}} & N_{\mathcal{A}} & 0 & 0 & 0 \\ 0 & 0 & F_{\mathcal{I}} & F_{\mathcal{A}} & 0 & 0 & T_{\mathcal{A}} \end{pmatrix} \begin{bmatrix} \Delta u_{\mathcal{N}_1} \\ \Delta u_{\mathcal{M}} \\ \Delta u_{\mathcal{I}} \\ \Delta u_{\mathcal{A}} \\ \Delta u_{\mathcal{N}_2} \\ \Delta \lambda_{\mathcal{I}} \\ \Delta \lambda_{\mathcal{A}} \end{bmatrix} = - \begin{bmatrix} r_{\mathcal{N}_1}^u \\ r_{\mathcal{M}}^u \\ r_{\mathcal{I}}^u \\ r_{\mathcal{A}}^u \\ r_{\mathcal{N}_2}^u \\ r_{\mathcal{I}}^{\lambda,n} \\ r_{\mathcal{A}}^{\lambda,\tau} \end{bmatrix} \quad (4)$$

to be solved in every nonlinear iteration. The 2×2 block matrix indicated by the dashed lines in (4) describes a linear system with a typical generalized saddle point structure. The upper left block consists of the entries of the tangential stiffness matrix (i.e. linearized internal forces) as well as linearizations of contact forces w.r.t. displacement degrees of freedom \mathbf{u} . The upper right block mirrors the discrete contact operator $\mathbf{C}(\mathbf{u})$, i.e. basically the two mortar matrices \mathbf{D} and \mathbf{M} , representing the linearizations of the contact forces w.r.t. the Lagrange multiplier unknowns λ . The kinematic constraints are incorporated in the bottom left block. The very simple sixth block row emerges from (3b) and (3c) for inactive nodes, while the last block row imposes frictionless sliding in the directions tangential to the contact interface.

The distinct pattern of zero entries in the upper left block reveals that the two solid bodies (indices \mathcal{N}_1 and \mathcal{N}_2) are indeed only coupled through the slave and master sides of the contact interface (indices \mathcal{S} and \mathcal{M}). Even though formulated for two solid bodies, the generalization to n solid bodies is straightforward and only a matter of notation.

In case of dual shape functions, the matrix \mathbf{D} reduces to a diagonal matrix and, thus, allows for a cheap condensation of the Lagrange multiplier unknowns. Algebraic multigrid preconditioners for this type of condensed system have been proposed in our earlier paper [85]. Furthermore, matrices $N_{\mathcal{M}}$, $N_{\mathcal{I}}$, and $N_{\mathcal{A}}$ denote the linearizations of the weighted gap function of (3b) at all active contact nodes. Finally, linearizations of the frictionless sliding condition (3c) are referred to by matrices $F_{\mathcal{I}}$, $F_{\mathcal{A}}$, and $T_{\mathcal{A}}$, respectively.

Note that the given matrix has 8 block rows but only 7 block columns in our notation in order to emphasize that the normal and tangential parts of the contact constraints for active nodes are considered separately, i.e. these two rows contain consistent linearizations of the active branch of (3b) and of (3c). Again, we point out that this separate notation is possible due to the fact that a local convective coordinate system is employed for evaluating the contact constraints / Lagrange multiplier weights μ , while the standard Cartesian frame is still applied for the discrete Lagrange multiplier values λ as well as the displacement unknowns. Yet, of course, the system matrix remains a square matrix with the total numbers of rows and columns being identical. The discrete vector $\mathbf{g}_{\mathcal{A}}$ contains all weighted gap values $(\tilde{g}_{n,h})_j$ associated with the active nodes at the contact interface.

MGV($\mathbf{A}^{(\ell)}$, x , b , ℓ):

if $\ell \neq \ell_{max}$

$x \leftarrow \mathcal{S}_\ell^{\nu_1}(\mathbf{A}^{(\ell)}, x, b)$

$r \leftarrow b - \mathbf{A}^{(\ell)}x$

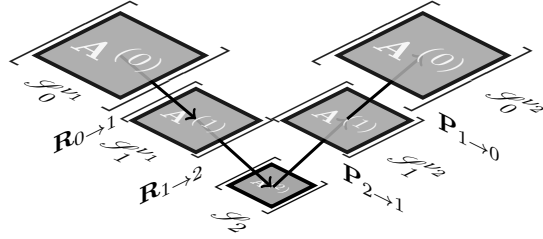
$c \leftarrow 0$

$c \leftarrow \text{MGV}(\mathbf{A}^{(\ell+1)}, x_{\ell+1}, \mathbf{R}_{\ell+1}r, \ell + 1)$

$x \leftarrow x + \mathbf{P}_{\ell+1}c$

$x \leftarrow \mathcal{S}_\ell^{\nu_2}(\mathbf{A}^{(\ell)}, x, b)$

else $x \leftarrow (\mathbf{A}^{(\ell)})^{-1}b$



(a) Recursive multigrid algorithm.

(b) Multigrid V-cycle in a 3 level setting.

Figure 2. Multigrid algorithm and V-cycle.

For ease of notation, the following short block notation is used in the remainder of the manuscript:

$$\begin{pmatrix} \mathbf{K} & \mathbf{C}_1^\top \\ \mathbf{C}_2 & -\mathbf{Z} \end{pmatrix} \begin{bmatrix} \Delta \mathbf{u} \\ \Delta \lambda \end{bmatrix} = - \begin{bmatrix} \mathbf{r}^u \\ \mathbf{r}^\lambda \end{bmatrix}. \quad (5)$$

3. MULTIGRID SCHEME FOR CONTACT PROBLEMS IN SADDLE POINT FORMULATION

Although multigrid methods can be used as standalone solvers for linear systems, they are usually incorporated into an iterative linear solver as a preconditioning method. Throughout this paper, we use a preconditioned generalized minimal residual (GMRES) solver [62] with one multigrid V-cycle sweep for preconditioning. A general introduction into the idea of preconditioning is beyond the scope of this paper. The reader is referred to the literature, e.g. [7].

3.1. Algebraic multigrid methods in a nutshell

Multigrid methods are based on the finding that many well-known and computationally cheap iterative methods (e.g. relaxation based iterative methods such as Jacobi or Gauss-Seidel methods) to solve linear systems $\mathbf{A}x = b$ effectively damp the high frequency part of an error vector but are less effective in damping out the low frequency error modes. Multigrid methods heavily make use of this smoothing property by applying such cheap smoothing methods on different coarsened representations of the original fine level problem.

3.1.1. Basic multigrid cycle and algorithm The multigrid algorithm given in Figure 2a is briefly described as follows: on each multigrid level ℓ , a level smoothing algorithm \mathcal{S}_ℓ performs ν_1 pre-smoothing sweeps before the residual vector r is transferred to the next coarser level $\ell + 1$ using the restriction operator \mathbf{R} . After the coarse level problem has been solved on the coarsest level, the correction c is then prolonged using the prolongation operator \mathbf{P} and the solution vector is smoothed using ν_2 post-smoothing sweeps. Figure 2b illustrates this basic multigrid V-cycle, exemplifying a three-level setting. As one can see from Figure 2b, applying a multigrid method basically means applying level smoothers on coarse representations \mathbf{A}_ℓ of the fine level problem \mathbf{A}_0 .

3.1.2. Algebraic multigrid methods There are different strategies for defining the transfer operators \mathbf{P} and \mathbf{R} which are necessary to generate coarse level matrices \mathbf{A}_ℓ ($\ell > 0$). For algebraic multigrid (AMG), the fine level operator \mathbf{A}_0 is sufficient to generate coarse level matrices \mathbf{A}_ℓ . An important class of AMG is given with the *smoothed aggregation AMG* which is based on so-called aggregates (see e.g. [75, 77]). The fine-level nodes are agglomerated and put into aggregates, which then represent ‘supernodes’ on the next coarser level. The aggregation information together with the null space information of \mathbf{A}_0 is used to construct the corresponding *tentative transfer operators* $\hat{\mathbf{R}}$ (for restriction) and $\hat{\mathbf{P}}$ (for prolongation). Transfer operators are used to restrict the fine level residual to the next coarser level and to interpolate the coarse level correction to the next finer level via prolongation. For an efficient multigrid method, the interaction of fine and coarse levels tackles those error modes, which appear as low-frequency modes on the fine level and cannot effectively be reduced by iterative smoothing methods on the fine level, but resemble high-frequency modes on a coarser level, such that iterative smoothing methods are effective again. On the coarsest level, a direct solver can take care of the

remaining error modes. For smoothed aggregation multigrid (cf. [26, 75]), the prolongation operator is found by applying one smoothing sweep with a damped Jacobi iteration using

$$\mathbf{P} = \widehat{\mathbf{P}} - \omega \mathbf{D}^{-1} \mathbf{A} \widehat{\mathbf{P}} \quad (6)$$

with \mathbf{D} being the diagonal part of \mathbf{A} and a damping parameter $\omega > 0$. Depending on the symmetry of the system matrix \mathbf{A} , the restriction operator is either chosen as $\mathbf{R} = \mathbf{P}^T$ or smoothed independently using, e.g. a Petrov-Galerkin approach (cf. [64]).

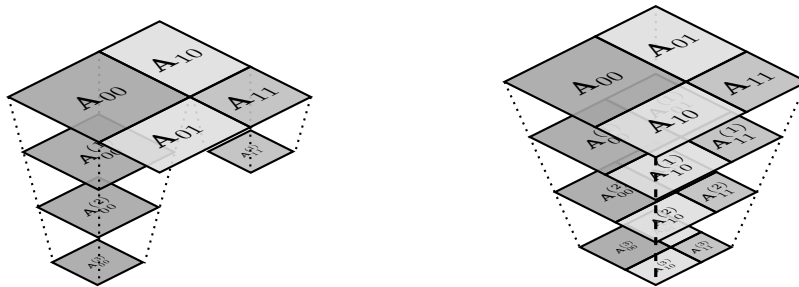
3.2. Algebraic multigrid methods for block matrices

Block matrices usually arise if multiple types of equations are coupled together. In the present context of contact problems in saddle point formulation, two types of equations, namely the balances of linear momentum of the solid bodies and the contact constraints, are connected via the off-diagonal blocks in (5). Similarly, multiphysics problems also yield block matrices where the coupling between different physical fields manifests itself in the off-diagonal blocks of the monolithic system matrix.

From a multigrid perspective, the most important question is where to consider the coupling between the different equations within the overall solver layout. In general, there are only two possible strategies to apply multigrid ideas to coupled block systems:

Nested multigrid approach: Multigrid methods can serve as local single field smoothers or solvers within well-known block preconditioners such as the SIMPLE method (cf. [50]) and variants for Schur complement based preconditioners or the block Gauss-Seidel (BGS) method. The coupling of the different fields or variables is only considered on the finest level in the outer (SIMPLE or BGS) iteration. This approach is well known in literature, e.g. for the Navier-Stokes equations [30, 68], fluid-structure interaction [27, 71] or general n -field problems [80]. The implementation is very easy and allows to use existing multigrid components in a standalone fashion within the solver. A graphic representation of this approach is shown in Figure 3a.

Fully coupled multigrid approach: Truly monolithic algebraic multigrid methods aim at coarsening the fully coupled fine level problem such that the block structure of the fine level matrix is preserved and the coupling information is present on all coarser levels, cf. Figure 3b. This is often achieved by using segregated transfer operators to preserve the characteristics of the sparsity pattern across all levels. Then, each level utilizes block smoothers to address the coupling. In [81], a coupled AMG method is developed and analyzed for a stabilized mixed finite element discretization of the Oseen equations. Fully coupled multigrid methods for multiphysics problems have been described in [27, 43, 80].



(a) Outer coupling iteration with nested multigrid methods. Example with four and nested coupling iteration on two multigrid levels for \mathbf{A}_{00} and \mathbf{A}_{11} . (b) Multiphysics multigrid approach with multigrid methods. Example with four and nested coupling iteration on all multigrid levels.

Figure 3. Multigrid for block matrices.

While the nested multigrid approach is easier to implement, it also allows for a high degree of modularity, since the multigrid hierarchies used to approximate the block inverses of the block smoother on the fine level can easily be swapped by any other method, either another type of multigrid algorithm, or another type of multi-level scheme (e.g. based on domain decomposition), or even any single-level approach if the block is of moderate size and scalability is not deteriorated. This flexibility is particularly useful if the coupled blocks differ a lot in size or if the user wants to apply a highly optimized solver for an individual block. The fully coupled multigrid method does

not offer such a degree of flexibility, yet it propagates the coupling conditions throughout the entire preconditioner. Thus, one expects a stronger and more robust preconditioning effect, since the coarse level corrections are aware of the coupling conditions. This expectation will later be confirmed in the numerical experiments, where the number of iterations for the fully coupled scheme is lower and more independent of the active contact nodes than for the nested multigrid approach.

3.3. Designing algebraic multigrid methods for contact problems

In the present context, we can interpret the mortar contact problem in saddle point formulation as the coupling of two types of equations: the structural equations and the contact equations which serve as constraint equations. Since the contact constraint equations are only defined along the contact interface, we can further classify the mortar contact problem as an interface-coupled problem (in contrast to volume-coupled problems). This information is important for the choice of coarsening strategy. The contact constraint equations are also responsible for the characteristic saddle point structure, which needs special attention when choosing an appropriate coupling algorithm between the structural equations and the contact constraints. Considering the class of fully coupled AMG schemes, the generalized saddle point problem (5) has to be preserved on all multigrid levels such that the contact constraints are considered on all levels. Due to the constraints, this will require Schur complement based level smoothers on all levels.

Alltogether, the key ingredients for designing an algebraic multigrid method for contact problems in saddle point formulation are the coarsening strategy as proposed in Section 4 and the level smoother and the coupling iteration as detailed in Section 5.

4. A COARSENING STRATEGY FOR MORTAR CONTACT CONSTRAINTS

4.1. Segregated transfer operators

To keep the characteristic saddle point block structure (5) on all multigrid levels, the common approach is to use *segregated* transfer operators

$$\mathbf{P}_{\ell+1} = \begin{pmatrix} \mathbf{P}^u & 0 \\ 0 & \hat{\mathbf{P}}^\lambda \end{pmatrix}_{\ell+1} \quad \text{and} \quad \mathbf{R}_{\ell+1} = \begin{pmatrix} \mathbf{R}^u & 0 \\ 0 & \hat{\mathbf{R}}^\lambda \end{pmatrix}_{\ell+1}, \quad (7)$$

as, e.g. introduced in [5, 11]. The segregated block transfer operators (7) are put together from the transfer operator blocks for the different physical and mathematical fields. Here, \mathbf{P}^u and \mathbf{R}^u describe the transfer operator blocks corresponding to the stiffness matrix block \mathbf{K} in (5). The transfer operators $\hat{\mathbf{P}}^\lambda$ and $\hat{\mathbf{R}}^\lambda$ define the level transfer for the Lagrange multipliers.

The block diagonal structure in (7) guarantees that the primary displacement variables and the secondary Lagrange multipliers are not “mixed up” on the coarser levels. That is, the coarse level matrix still has the same block structure with a clear distinction of momentum and constraint equations as for the fine level problem since the columns and rows of the transfer operators $\mathbf{P}_{\ell+1}$ and $\mathbf{R}_{\ell+1}$ can be interpreted as some kind of basis functions.

For many volume-coupled problems, for example in thermo-structure-interaction problems [19], it is straightforward to generate $\hat{\mathbf{P}}^\lambda$ and $\hat{\mathbf{R}}^\lambda$ to be consistent with \mathbf{P}^u and \mathbf{R}^u . That is, in context of smoothed aggregation algebraic multigrid we just use the same aggregates for building \mathbf{P}^u and $\hat{\mathbf{P}}^\lambda$ (and the same for the restrictors, respectively).

For interface-coupled problems with interface constraints it is more difficult. Due to the saddle point structure of (5), the nonzero pattern of the \mathbf{Z} block is insufficient to generate valid aggregates for the Lagrange multipliers. Consequently, we need a special routine for finding aggregates for the Lagrange multipliers λ to be able to build the (non-smoothed) transfer operators $\hat{\mathbf{P}}^\lambda$ and $\hat{\mathbf{R}}^\lambda$. Nevertheless it seems natural to reflect the aggregation information of the structural equations along the contact interface in the choice of the aggregates for the corresponding Lagrange multipliers λ .

4.2. Aggregation strategy for displacement variables

In order to preserve the physics of the fine level contact problem, it is important to keep the two solid bodies separated in the matrix representation on all coarse levels. Therefore, we apply the standard aggregation strategy to a modified \mathbf{K} block from (5), where all off-diagonal entries representing connections between the two solid bodies are dropped — in particular the matrix

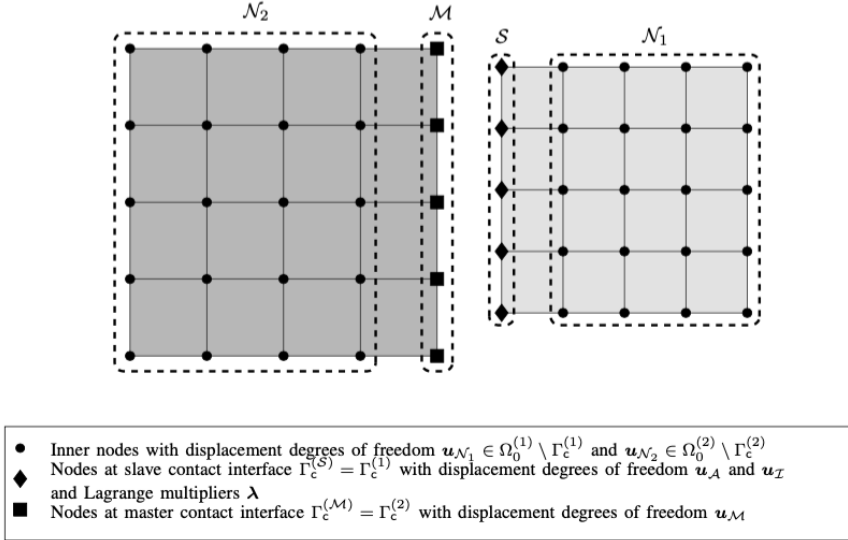


Figure 4. Aggregation for contact example in saddle point formulation.

blocks $\mathbf{K}_{\mathcal{I}\mathcal{M}}, \mathbf{K}_{\mathcal{A}\mathcal{M}}, \mathbf{K}_{\mathcal{M}\mathcal{I}}$ and $\mathbf{K}_{\mathcal{M}\mathcal{A}}$ — in order to make sure that the resulting displacement aggregates $\mathcal{A}_\ell^{\mathbf{u}}$ do not cross the contact interface (see Figure 4). Neglecting these blocks during aggregation guarantees that the two solid bodies are not melted together in the coarse matrix representation. We stress that the modified \mathbf{K} is never formed explicitly, but rather the off-diagonal entries are dropped on the fly during the aggregation process.

4.3. Aggregation strategy for Lagrange multipliers

In contrast to geometric multigrid methods, there is not so much literature on aggregation-based AMG methods for contact problems in saddle point formulation. The only publication, the authors are aware of covering all aspects of smoothed aggregation methods for structural contact problems in saddle point formulation, is [5], which also discusses a special aggregation strategy for the Lagrange multipliers. To find aggregates $\mathcal{A}_\ell^{\boldsymbol{\lambda}}$ for the Lagrange multipliers, Adams [5] proposes to apply the standard aggregation algorithm to the graph of a suitable matrix representing the Lagrange multipliers. However, this approach has some drawbacks: First, the graph used for the aggregation of the Lagrange multipliers $\boldsymbol{\lambda}$ has to be built explicitly to serve as input for the standard aggregation algorithm. Secondly, one has to run the aggregation algorithm sequentially both for the displacement degrees of freedom and for the Lagrange multipliers. For the second run of the aggregation method, one might have to use a different set of aggregation parameters to obtain optimal results, which further increases the complexity for the user. Algorithmically, the resulting aggregates $\mathcal{A}_\ell^{\boldsymbol{\lambda}}$ for the Lagrange multipliers are built independently from the displacement aggregates $\mathcal{A}_\ell^{\mathbf{u}}$.

In this work, we propose a different approach to build aggregates $\mathcal{A}_\ell^{\boldsymbol{\lambda}}$ for the Lagrange multipliers, which does not suffer from above drawbacks. Instead of explicitly building some helper matrix for the aggregation routine, interface aggregates $\mathcal{A}_\ell^{\boldsymbol{\lambda}}$ for the Lagrange multipliers are directly generated using the aggregation information of the displacement variables (see Figure 4). The resulting interface aggregates for the Lagrange multipliers are by construction aligned with the corresponding displacement aggregates.

The exact aggregation procedure is described in Algorithm 1. Assuming that the standard aggregates $\mathcal{A}_\ell^{\mathbf{u}}$ for the displacement degrees of freedom are available, new aggregates $\mathcal{A}_\ell^{\boldsymbol{\lambda}}$ are built by collecting the corresponding Lagrange multiplier degrees of freedom. Beside the displacement aggregates $\mathcal{A}_\ell^{\mathbf{u}}$, only the mortar matrix \mathbf{D} is needed to algebraically reconstruct the contact interface and find the associated Lagrange multipliers. The new aggregates $\mathcal{A}_\ell^{\boldsymbol{\lambda}}$ for the Lagrange multipliers can be interpreted as the natural extension of the displacement aggregates $\mathcal{A}_\ell^{\mathbf{u}}$ at the interface. This facilitates to keep the ratio of coarse level nodes at the slave contact interface and the coarse Lagrange multipliers constant, which also balances the ratio of contact constraints and inner structural displacement degrees of freedom over all multigrid levels.

Algorithm 1: Aggregation algorithm for Lagrange multipliers.

Procedure LagMultAggregation($\mathcal{A}_\ell^{\mathbf{u}}, \mathbf{D}$)

```

Initialize empty set and counter for aggregates  $\mathcal{A}_\ell^\lambda$ 
 $\mathcal{A}_\ell^\lambda \leftarrow \emptyset, l \leftarrow 0$ 

Initialize empty mapping of displacement aggregates to Lagrange multiplier aggregates
 $\mathbf{d}(k) \leftarrow \emptyset \quad \forall k = 1, \dots, m_{\mathcal{A}_\ell^{\mathbf{u}}}$ 

Loop over slave displacement DOFs (rows of  $\mathbf{D}$ )
for  $i \in \mathcal{D}_S$  do
  Find displacement node  $n^{\mathbf{u}}$  id corresponding to displacement DOF  $i$ 
   $n^{\mathbf{u}} \leftarrow \mathbf{n}(i)$ 

  Find aggregate index  $k$  that contains displacement node  $n^{\mathbf{u}}$ 
  Find  $k$  with  $\mathcal{A}_\ell^{(k)} \in \mathcal{A}_\ell^{\mathbf{u}}$  where  $n^{\mathbf{u}} \in \mathcal{A}_\ell^{(k)}$ 

  Loop over all Lagrange multipliers  $j$ 
  for  $j \in \mathcal{D}_\lambda$  do
    Check whether Lagrange multiplier  $j$  is coupled with row  $i$ 
    if  $\mathbf{D}_{i,j} \neq 0$  then
      Find pseudo node  $n^\lambda$  for Lagrange multiplier  $j$ 
       $n^\lambda \leftarrow \mathbf{n}(j)$ 

      Check whether to build a new Lagrange multiplier aggregate
      if  $\mathbf{d}(k) = \emptyset$  then
        Increment internal aggregation counter
         $l \leftarrow l + 1$ 

        Build a new aggregate and add Lagrange multiplier node  $n^\lambda$ 
         $\mathcal{A}_\ell^{(l)} \leftarrow \{n^\lambda\}$ 

        Associate displacement aggregate  $k$  with Lagrange multiplier aggregate  $l$ 
         $\mathbf{d}(k) \leftarrow \{l\}$ 

        Add new aggregate to set of Lagrange multiplier aggregates  $\mathcal{A}_\ell^\lambda$ 
         $\mathcal{A}_\ell^\lambda \leftarrow \mathcal{A}_\ell^\lambda \cup \mathcal{A}_\ell^{(l)}$ 
      else
        Extend aggregate  $0 \leq \mathbf{d}(k) \leq l$  with pseudo node
         $\mathcal{A}_\ell^{(\mathbf{d}(k))} \leftarrow \mathcal{A}_\ell^{(\mathbf{d}(k))} \cup \{n^\lambda\}$ 
      end
    end
  end
end
end

Return aggregates for Lagrange multipliers
return  $\mathcal{A}_\ell^\lambda$ 

```

The coarsening strategy outlined in Algorithm 1 as well as all other AMG components of the presented saddle-point preconditioner for contact problems have been implemented in MUELU [8], the next-generation multigrid package within the TRILINOS project [3]. For further details on the implementation, we refer to the MUELU User's Guide [8] and the MUELU website [2].

5. BLOCK SMOOTHING METHODS FOR MORTAR CONTACT PROBLEMS

Using saddle point preserving aggregation and segregated transfer operators as outlined in Section 4 to generate a fully coupled AMG hierarchy (see Section 3.3), the coupling of structural equilibrium equations and contact constraints on all levels of a fully coupled AMG hierarchy is now addressed by block smoothing methods on each level. Schur complement based coupling iterations present themselves as ideal candidates to deal with the saddle point structure resulting from the constraint-like contact equations.

Nevertheless, not all classical Schur complement based block smoothers for saddle point systems behave the same way when applied to contact problems. Specifically, it is important that the block smoothers account for the contact constraints in (5). While each iteration's intermediate solution itself is not of great interest, its satisfaction of the contact constraints impacts the overall number of required iterations to reach convergence, since only an intermediate solution that satisfies the contact constraints can then also be accepted as the final solution of the iterative process. While an exact satisfaction of the contact constraints by the block smoothers is desirable, practical computations require a compromise between the accuracy of the block smoothers and their computational effort in order to obtain computationally competitive preconditioners.

In Section 5.1, we first revisit some classical smoothing methods for saddle point systems. Therefore, we assume exact inverses for the predictor and corrector step of each smoother to focus on the systematic error resulting from the specific block structure of the smoother. The assumption of exact inverses allows to assess the behavior of each block smoother via an error matrix, such that the impact of the block smoother on the contact constraints can be characterized. Afterwards, we introduce computationally cheaper variants with inexact block inverses in the predictor and corrector step as a compromise between accuracy and performance in Section 5.2, however preventing the discourse on error matrices. While the approximations due to the block structure of a specific smoother originate from the definition of the smoothing method, the quality of the approximate block inverses can fully be controlled by the user. Finally, Section 5.3 discusses their application in the context of saddle point systems for contact formulations.

5.1. Block smoothers for saddle point problems

Now, we study the systematic errors introduced by specific block smoothers and their impact on the contact constraints. Therefore, we first assume exact mathematical block operations. To carefully distinguish between systematic errors and errors stemming from practical Schur complement approximations, we postpone the discussion of such approximations to Section 5.2.

The general block smoothing scheme can be written as

$$\begin{bmatrix} \Delta \mathbf{u}^{k+1} \\ \Delta \boldsymbol{\lambda}^{k+1} \end{bmatrix} = \begin{bmatrix} \Delta \mathbf{u}^k \\ \Delta \boldsymbol{\lambda}^k \end{bmatrix} + Q^{-1} \left(\begin{bmatrix} \mathbf{r}_u^k \\ \mathbf{r}_\lambda^k \end{bmatrix} - \begin{pmatrix} \mathbf{K} & \mathbf{C}_1^\top \\ \mathbf{C}_2 & -\mathbf{Z} \end{pmatrix} \begin{bmatrix} \Delta \mathbf{u}^k \\ \Delta \boldsymbol{\lambda}^k \end{bmatrix} \right) \quad (8)$$

where Q describes the 2×2 block preconditioning matrix approximating the 2×2 block operator in (5). Typical block smoothers consist of an outer coupling iteration with nested subsolvers to build the inverses of the diagonal blocks of Q in an algorithmic predictor-corrector scheme. In the following, a few classical block smoothers from literature (e.g. [49]) are introduced, stating that this list is by far not complete. All of them can be interpreted as block extensions of classical iterative smoothing methods following the general block scheme (8), but depending on the definition of Q with a different effect on the contact constraints by introducing certain systematic errors. In general, the better Q approximates the block operator from (5), the lower the number of linear iterations will be when using the block smoother within a multigrid preconditioner.

5.1.1. *Uzawa smoother* For the (inexact) Uzawa smoother, one chooses

$$Q_{\text{UZ}} := \frac{1}{\alpha} \begin{pmatrix} \mathbf{K} & \mathbf{0} \\ \mathbf{C}_2 & -\tilde{S} \end{pmatrix}. \quad (9)$$

The parameter $\alpha > 0$ is a damping parameter and \tilde{S} describes a cheap approximation of the Schur complement $S = \mathbf{Z} + \mathbf{C}_2 \mathbf{K}^{-1} \mathbf{C}_1^\top$. For a theoretical review of Uzawa like smoothers, the reader is referred to [15, 23, 98].

With the off-diagonal coupling block \mathbf{C}_2 in (9), the smoother performs a one-way coupling in the sense that the Lagrange multiplier increments now depend on the current increment of the displacement degrees of freedom. Algorithm 2 represents the practical implementation as a predictor-corrector method. In each smoothing iteration, one calculates a prediction for the displacement increments $\delta \mathbf{u}^{k+1}$, which are taken into account when solving for the corresponding Lagrange multiplier increments $\delta \boldsymbol{\lambda}^{k+1}$.

Assuming the smoothing iteration has converged, the error matrix for the Uzawa smoother is given as

$$\mathbf{E}_{\text{UZ}} := \mathbf{A} - Q_{\text{UZ}} = \begin{pmatrix} \left(1 - \frac{1}{\alpha}\right) \mathbf{K} & \mathbf{C}_1^\top \\ \left(1 - \frac{1}{\alpha}\right) \mathbf{C}_2 & -\mathbf{Z} + \frac{1}{\alpha} \tilde{S} \end{pmatrix}. \quad (10)$$

Algorithm 2: Uzawa smoother.

Procedure Uzawa(α, k_{\max})

Apply k_{\max} smoothing sweeps with the Uzawa algorithm
for $k \leftarrow 0$ **to** $k_{\max} - 1$ **do**
Prediction step: solve for $\delta \mathbf{u}^{k+1}$
 $\mathbf{K} \delta \mathbf{u}^{k+1} = \mathbf{r}_{\mathbf{u}}^k - \mathbf{K} \Delta \mathbf{u}^k - \mathbf{C}_1^\top \Delta \boldsymbol{\lambda}^k$
Correction step: solve for $\delta \boldsymbol{\lambda}^{k+1}$
 $-\tilde{\mathbf{S}} \delta \boldsymbol{\lambda}^{k+1} = \mathbf{r}_{\boldsymbol{\lambda}}^k - \mathbf{C}_2 \Delta \mathbf{u}^k + \mathbf{Z} \Delta \boldsymbol{\lambda}^k - \mathbf{C}_2 \delta \mathbf{u}^{k+1}$
Update step: update solution variables
 $\Delta \mathbf{u}^{k+1} \leftarrow \Delta \mathbf{u}^k + \alpha \delta \mathbf{u}^{k+1}$
 $\Delta \boldsymbol{\lambda}^{k+1} \leftarrow \Delta \boldsymbol{\lambda}^k + \alpha \delta \boldsymbol{\lambda}^{k+1}$
end
Return smooth solution vector
return ($\Delta \mathbf{u}^{k_{\max}}, \Delta \boldsymbol{\lambda}^{k_{\max}}$)

With $\alpha = 1$, $\tilde{\mathbf{S}} = \mathbf{Z} + \mathbf{C}_2 \tilde{\mathbf{K}}^{-1} \mathbf{C}_1^\top$ being an approximation to the Schur complement, and $\tilde{\mathbf{K}}$ denoting an easy-to-invert approximation of \mathbf{K} , e.g. the diagonal of \mathbf{K} as a cheap variant for the approximation $\tilde{\mathbf{K}}$, i.e. $\tilde{\mathbf{K}} = \text{diag}(\mathbf{K})$, it is easy to verify that the error matrix of the Uzawa smoother reduces to

$$\mathbf{E}_{\text{UZ}} = \begin{pmatrix} 0 & \mathbf{C}_1^\top \\ 0 & \mathbf{C}_2 \tilde{\mathbf{K}}^{-1} \mathbf{C}_1^\top \end{pmatrix}. \quad (11)$$

That is, since the second block row in the error matrices (10) or (11) does not vanish, the Uzawa smoother by definition cannot exactly fulfill the contact constraints, but adds a systematic error for the contact constraints. This might have a negative impact on the overall performance of the iterative linear solver.

5.1.2. Braess–Sarazin smoother Originally introduced for the Stokes problem in [11], the Braess–Sarazin smoother belongs to the class of block approximate smoothers and is based on the choice

$$\mathbf{Q}_{\text{BS}} := \begin{pmatrix} \alpha \tilde{\mathbf{K}} & \mathbf{C}_1^\top \\ \mathbf{C}_2 & -\mathbf{Z} \end{pmatrix} \quad (12)$$

for the block preconditioning matrix \mathbf{Q} in (8). Again, the parameter $\alpha > 0$ denotes a scaling parameter and $\tilde{\mathbf{K}}$ refers to an easy-to-invert approximation of \mathbf{K} .

Algorithm 3: Braess–Sarazin smoother.

Procedure BraessSarazin(α, k_{\max})

Apply k_{\max} smoothing sweeps with Braess–Sarazin algorithm
for $k \leftarrow 0$ **to** $k_{\max} - 1$ **do**
Prediction step: determine prediction $\Delta \mathbf{u}^{k+\frac{1}{2}}$ by calculating
 $\Delta \mathbf{u}^{k+\frac{1}{2}} = \Delta \mathbf{u}^k + \frac{1}{\alpha} \tilde{\mathbf{K}}^{-1} (\mathbf{r}_{\mathbf{u}}^k - \mathbf{K} \Delta \mathbf{u}^k - \mathbf{C}_1^\top \Delta \boldsymbol{\lambda}^k)$
Correction step: solve for $\delta \boldsymbol{\lambda}^{k+\frac{1}{2}}$
 $-(\mathbf{Z} + \frac{1}{\alpha} \mathbf{C}_2 \tilde{\mathbf{K}}^{-1} \mathbf{C}_1^\top) \delta \boldsymbol{\lambda}^{k+\frac{1}{2}} = \mathbf{r}_{\boldsymbol{\lambda}}^k + \mathbf{Z} \Delta \boldsymbol{\lambda}^k - \mathbf{C}_2 \Delta \mathbf{u}^{k+\frac{1}{2}}$
Update step: update solution variables
 $\Delta \boldsymbol{\lambda}^{k+1} \leftarrow \Delta \boldsymbol{\lambda}^k + \delta \boldsymbol{\lambda}^{k+\frac{1}{2}}$
 $\Delta \mathbf{u}^{k+1} \leftarrow \Delta \mathbf{u}^{k+\frac{1}{2}} - \frac{1}{\alpha} \tilde{\mathbf{K}}^{-1} \mathbf{C}_1^\top \delta \boldsymbol{\lambda}^{k+\frac{1}{2}}$
end
Return smooth solution vector
return ($\Delta \mathbf{u}^{k_{\max}}, \Delta \boldsymbol{\lambda}^{k_{\max}}$)

Assuming convergence of the smoother, the error matrix for the Braess–Sarazin smoother is given as

$$\mathbf{E}_{\text{BS}} := A - Q_{\text{BS}} = \begin{pmatrix} \mathbf{K} - \alpha \tilde{\mathbf{K}} & 0 \\ 0 & 0 \end{pmatrix}. \quad (13)$$

With the second block row in the blocked operator (5) being retained in (12), the Braess–Sarazin smoother seems to be a reasonable choice for dealing with contact constraints. The error matrix in (13) reveals that the quality of the block smoother only depends on the choice of $\tilde{\mathbf{K}}$.

The implementation as a predictor-corrector method is based on the splitting of (12) into

$$\begin{pmatrix} \alpha \tilde{\mathbf{K}} & \mathbf{C}_1^T \\ \mathbf{C}_2 & -\mathbf{Z} \end{pmatrix} = \begin{pmatrix} \alpha \tilde{\mathbf{K}} & 0 \\ \mathbf{C}_2 & -\mathbf{Z} - \frac{1}{\alpha} \mathbf{C}_2 \tilde{\mathbf{K}}^{-1} \mathbf{C}_1^T \end{pmatrix} \begin{pmatrix} I & \frac{1}{\alpha} \tilde{\mathbf{K}}^{-1} \mathbf{C}_1^T \\ 0 & I \end{pmatrix}. \quad (14)$$

As one can easily see from Algorithm 3, the prediction step can be understood as one hard-coded sweep with a (damped) Jacobi iteration. In other words, the quality of the prediction for the displacement degrees of freedom must be considered rather poor. Exactly fulfilling contact constraints with respect to a rather poor prediction of the displacement variables might not be optimal for the overall performance of the preconditioner.

5.1.3. SIMPLE variants Originally introduced in [50, 51], the SIMPLE method is based on the approximate block factorization

$$Q_{\text{SIMPLE}} := \begin{pmatrix} \mathbf{K} & 0 \\ \mathbf{C}_2 & -\tilde{S} \end{pmatrix} \begin{pmatrix} I & \tilde{\mathbf{K}}^{-1} \mathbf{C}_1^T \\ 0 & \frac{1}{\alpha} I \end{pmatrix} = \begin{pmatrix} \mathbf{K} & \mathbf{K} \tilde{\mathbf{K}}^{-1} \mathbf{C}_1^T \\ \mathbf{C}_2 & \left(1 - \frac{1}{\alpha}\right) \mathbf{C}_2 \tilde{\mathbf{K}}^{-1} \mathbf{C}_1^T - \frac{1}{\alpha} \mathbf{Z} \end{pmatrix} \quad (15)$$

for the iterative method in (8). In (15), \tilde{S} denotes an approximation of the Schur complement $S := \mathbf{Z} + \mathbf{C}_2 \mathbf{K}^{-1} \mathbf{C}_1^T$ with a cheap and easy-to-invert approximation $\tilde{\mathbf{K}}$ of the block \mathbf{K} . Algorithm 4 shows the implementation of the SIMPLE method using the predictor-corrector scheme (cf. [22]).

Algorithm 4: SIMPLE smoother.

Procedure SIMPLE(α, k_{max})

Apply k_{max} smoothing sweeps with SIMPLE algorithm

for $k \leftarrow 0$ **to** $k_{\text{max}} - 1$ **do**

 Prediction step: solve for $\Delta \mathbf{u}^{k+\frac{1}{2}}$

$$\mathbf{K} \Delta \mathbf{u}^{k+\frac{1}{2}} = \mathbf{r}_{\mathbf{u}}^k - \mathbf{C}_1^T \Delta \boldsymbol{\lambda}^k$$

 Correction step: solve for $\delta \boldsymbol{\lambda}^{k+\frac{1}{2}}$

$$-\tilde{S} \delta \boldsymbol{\lambda}^{k+\frac{1}{2}} = \mathbf{r}_{\boldsymbol{\lambda}}^k + \mathbf{Z} \Delta \boldsymbol{\lambda}^k - \mathbf{C}_2 \Delta \mathbf{u}^{k+\frac{1}{2}}$$

 Update step: update solution variables

$$\Delta \boldsymbol{\lambda}^{k+1} \leftarrow \Delta \boldsymbol{\lambda}^k + \alpha \delta \boldsymbol{\lambda}^{k+\frac{1}{2}}$$

$$\Delta \mathbf{u}^{k+1} \leftarrow \Delta \mathbf{u}^{k+\frac{1}{2}} - \alpha \tilde{\mathbf{K}}^{-1} \mathbf{C}_1^T \delta \boldsymbol{\lambda}^{k+\frac{1}{2}}$$

end

Return smooth solution vector

return $(\Delta \mathbf{u}^{k_{\text{max}}}, \Delta \boldsymbol{\lambda}^{k_{\text{max}}})$

For our applications, we found the diagonal matrix containing the row sums of $|\mathbf{K}| = (|a_{ij}|)_{i,j=1,\dots,n_{\mathbf{K}}}$ to be a good approximation for block \mathbf{K} . This corresponds to the SIMPLEC method as introduced in [74]. That is, $\tilde{\mathbf{K}}$ is defined as the diagonal lumping of $|\mathbf{K}|$ with

$$\tilde{\mathbf{K}} = \text{diag} \left(\sum_{j=1}^{n_{\mathbf{K}}} |a_{ij}| \right), \quad i = 1, \dots, n_{\mathbf{K}}. \quad (16)$$

The default choice for \tilde{S} is consequently $\tilde{S} = \alpha \mathbf{Z} + \alpha \mathbf{C}_2 \tilde{\mathbf{K}}^{-1} \mathbf{C}_1^T$ with $\tilde{\mathbf{K}}$ as defined in (16). A more theoretical discussion on the mathematical consequences of approximations for the Schur complement S can be found in [97].

For the SIMPLE preconditioner, the error matrix is calculated by

$$E_{\text{SIMPLE}} := A - Q_{\text{SIMPLE}} = \begin{pmatrix} 0 & C_1^T - K\tilde{K}^{-1}C_1^T \\ 0 & -Z - C_2\tilde{K}^{-1}C_1^T + \frac{1}{\alpha}\tilde{S} \end{pmatrix}. \quad (17)$$

As one can see from (17), SIMPLE perturbs the Lagrange multipliers, but it does not affect the terms that operate on the primary displacement variables since the first block column in (17) is zero. Choosing $\tilde{S} = \alpha Z + \alpha C_2\tilde{K}^{-1}C_1^T$, the error matrix reduces to

$$E_{\text{SIMPLE}} = \begin{pmatrix} 0 & C_1^T - K\tilde{K}^{-1}C_1^T \\ 0 & 0 \end{pmatrix}. \quad (18)$$

That is, an appropriate approximation \tilde{S} of the Schur complement S allows to exactly satisfy the contact constraints within one smoothing sweep. Depending on the choice for the approximation \tilde{K} , the SIMPLE method admits an error in the coupling between displacements and contact constraints. However, compared to the Braess-Sarazin method from Section 5.1.2, we put more focus on a good prediction for the displacements with a consistent update for fulfilling the contact constraints.

5.2. Cheap variants of block smoothers

As one can easily see from the Algorithms 2 to 4, all block smoothing methods internally require inverses of the matrix blocks on the matrix diagonal and of the Schur complement operator \tilde{S} . Specifically, there is one linear system to be solved in the prediction step and one during the correction step. To keep the computational costs low in practical computations, one does not solve for the block inverses exactly, but apply a cheap approximation, e.g. by using a fixed number of smoothing sweeps with a relaxation-based smoothing method such as symmetric Gauss–Seidel or an ILU sweep. This allows for more flexibility for finding a good compromise between quality and performance, as the user can decide how much effort should be put on finding a good prediction and fulfilling the Schur complement equation in the corrector step. While the systematic errors introduced by the choice of the block smoother from Section 5.1 are fixed, the practitioner has full control over the quality of the block inverses.

The numerical examples in Section 6 show that such an approximation leads to efficient and computationally reasonable block smoothing methods. As a naming convention, the prefix ‘‘Cheap’’ is added to the name of the block smoothing method to indicate the usage of a cheap approximation for finding the inverse of the diagonal blocks in addition to the systematic approximations of building the Schur complement operator as discussed in Section 5.1.

The block smoothing methods and their cheap variants are available in the XPETRA package of the TRILINOS project [3]. For further details on the implementation, we refer to the MUELU User’s Guide [8] and the MUELU and XPETRA websites [2, 4], respectively.

5.3. Comparison of saddle point smoothing methods for contact problems

Considering the block scheme of (5), there are two main challenges for contact problems. First, we have the two distinct sets of equations as described in Section 2.3: the structural equations formulated in cartesian coordinates and the set of contact constraints formulated in normal-tangential coordinates relative to the contact surface. Second, the coupling of those two distinct sets of structural equations and contact constraint equations.

Algorithmically, the coupling of structural equations at the contact interface via the off-diagonal blocks in (5) is only considered within the block smoothers from Section 5.1. Therefore, constraint smoothers (cf. [39]) are the natural choice for contact problems, since the contact problem is implicitly governed by the contact constraint equations. Only solutions that are in alignment with the contact constraints are of interest.

Under certain preconditions as discussed in Section 5.1 and assuming sufficient smoother iterations to reach convergence of the smoother, both the Braess–Sarazin method and the SIMPLE-based methods would exactly fulfill the contact constraints as one can see from equations (13) or (18). Therefore, with a systematic error in the lower-right block of (11), the Uzawa method seems to be less promising for contact problems. In the Braess–Sarazin method, the approximation $\tilde{K} = \text{diag}(K)$ is hard-coded with some scaling parameter $\alpha > 0$ and consistently used within the approximate Schur complement operator, which is defined by $\tilde{S} = Z + \frac{1}{\alpha}C_2\tilde{K}^{-1}C_1^T$.

In contrary to the Braess–Sarazin method, the SIMPLE based methods keep the full K block whenever possible in the block factorization and use \tilde{K} only where its inverse is required. Consequently,

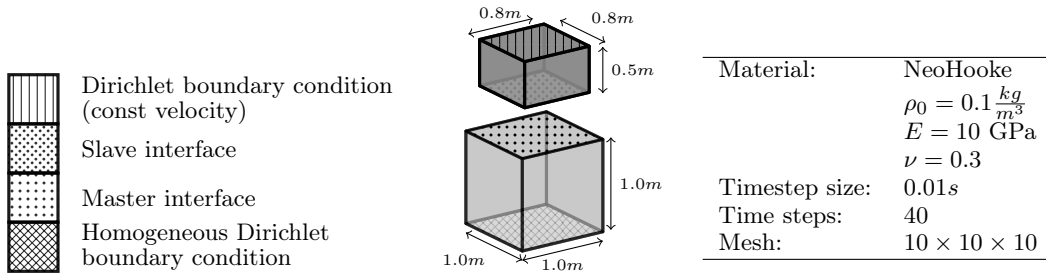


Figure 5. Two solid bodies example — Geometric configuration and parameters.

in “cheap” variants of the SIMPLE method, more elaborate smoothing strategies can be used for the K block instead of a hard-coded Jacobi sweep. Therefore, one can think of the SIMPLE methods to allow for a more balanced quality of approximations for the displacement degrees of freedom and Lagrange multipliers for the contact constraints, whereas the Braess–Sarazin method exhibits an imbalance in the sense that the computational effort spent for approximating the constraints is much higher than for dealing with the displacement variables.

6. NUMERICAL EXAMPLES

For the numerical examples, we use our in-house code BACI [1] that internally uses various capabilities from the TRILINOS project [3]. The implementation of the multigrid algorithms is based on Trilinos’ MueLu package [8, 84]. In particular, all block smoothers from Section 5 as well as the contact specific aggregation strategy for the Lagrange multiplier unknowns as described in Section 4 are readily available in MueLu.

6.1. Two solid bodies example

With the first example, we want to study the effect of the block smoothers from Section 5 for contact problems. Here, we not only compare different block smoothers, but also highlight the effect of varying the quality of the sub-smoothing steps within the block smoother versus increasing the number of outer coupling iterations.

Motivated by findings in our previous work [85], this example briefly revisits a detail that has been problematic in the context of contact problems in condensed contact formulations. While the discrete global unknowns $(\mathbf{u}, \boldsymbol{\lambda})$ are – as usual – formulated with respect to the global Cartesian frame, the discrete Lagrange multiplier weights $\boldsymbol{\mu}$ and therefore the contact constraint equations in (3b) and (3c) are formulated with respect to a local convective coordinate system. This local system is defined at each slave node j by a surface normal vector and two tangent vectors, i.e. by a triad of orthonormal basis vectors $(\mathbf{n})_j$, $(\boldsymbol{\tau}_\xi)_j$ and $(\boldsymbol{\tau}_\eta)_j$. Although it represents a quite intuitive and natural choice in contact mechanics, this local constraint formulation may lead to non-diagonally dominant system matrices and therefore poses a serious challenge to the development of iterative linear solvers as has been elaborated in [85]. In the following, we will also investigate the susceptibility of the proposed saddle point preconditioners to this phenomenon.

6.1.1. Geometrical setup Since we are interested in the solver behavior, by intention we choose a simple 3D contact example as shown in Figure 5. There are two solid bodies with the same material parameters using a Neo-Hookean material (density $\rho_0 = 0.1 \frac{kg}{m^3}$, Young’s modulus $E = 10 \text{ GPa}$, Poisson’s ratio $\nu = 0.3$). The initial gap between the two solid bodies is 0.02 meters. The upper solid body (size: $0.8m \times 0.8m \times 0.5m$) is moving down with constant velocity along the normal to the contact interface towards the lower fixed solid body (size: $1.0m \times 1.0m \times 1.0m$).

6.1.2. Experimental setup To investigate a potential impact of the contact formulation in different coordinate systems (cartesian coordinates for the structural degrees of freedom and normal-tangential coordinates for the Lagrange multipliers) on the linear solver, we perform a similar experiment as introduced in [85] and rotate the example setup around α_y and α_z as shown in Figure 6. We expect

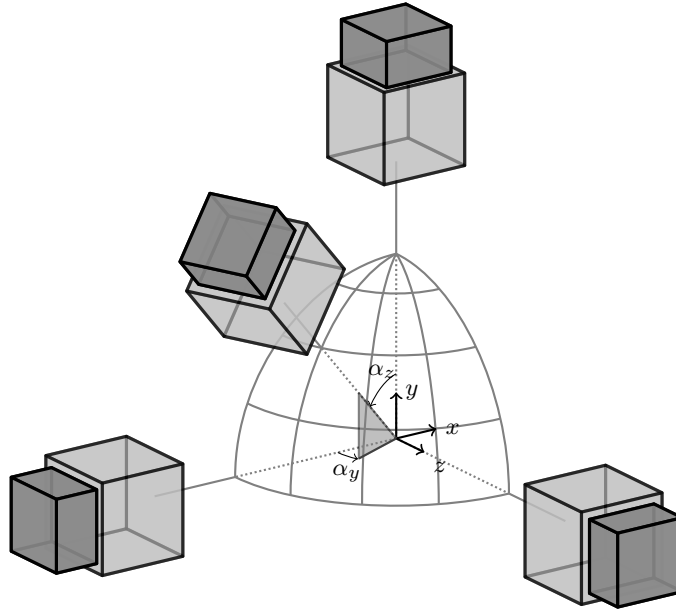


Figure 6. Two solid bodies example — Experimental setup to study independence of spatial orientation.

the number of linear iterations to be independent of the rotation angles α_y and α_z , since the underlying physics do not change. Any dependency of the linear solver on α_y and α_z would be a result of purely numerical effects and would turn out highly problematic for the iterative solution of large and complex contact problems. For reasons of symmetry, it is sufficient to vary α_y and α_z within $0 \leq \alpha_y, \alpha_z \leq \frac{\pi}{2}$.

6.1.3. Discretization The spatial discretization is based on a $10 \times 10 \times 10$ mesh for each solid block with altogether 6000 displacement degrees of freedom and 300 Lagrange multipliers modeling the contact coupling constraints for the 10×10 slave nodes at the contact interface (see Figure 5). The simulation runs for 40 time steps with a time step size of 0.01s on 4 processors. After 6 time steps ($t = 0.06s$) both bodies come into contact and are deformed. We assume frictionless contact here. With this example we reduce the contact-specific effects (such as the contact search based on an active set strategy) to a minimum, such that one can focus on the linear solvers. That is, the contact zone is not changing once the two solid bodies are in contact.

6.1.4. Stopping criteria The nonlinear iteration inside each time step stops if either $\|\Delta \mathbf{u}\|_e < 10^{-8}$ holds for the Newton increment of the displacement degrees of freedom, or alternatively, if the conditions

$$\|\mathbf{r}_i^{\mathbf{u}}\|_e < 10^{-6} \wedge \|\mathbf{r}_i^{\boldsymbol{\lambda}}\|_e < 10^{-4} \quad (19)$$

hold for the nonlinear residuals $\mathbf{r}_i^{\mathbf{u}}$ and $\mathbf{r}_i^{\boldsymbol{\lambda}}$ in (4) after applying i Newton iterations. Thereby, $\|\bullet\|_e$ denotes the Euclidian vector norm. Those stopping criteria for the nonlinear solver are chosen to result in the same number of nonlinear iterations in each time step for an easier comparison of the linear solver behavior.

Within each Newton iteration, the saddle point system (4) is solved iteratively using a preconditioned GMRES method with a 3-level AMG preconditioner as described in Section 3. The iterative process for the linear system is considered to be converged, if it is

$$\left\| \frac{\mathbf{r}^k}{\mathbf{r}^0} \right\|_e < 10^{-8} \quad (20)$$

for the full residual vector $\mathbf{r}^k = \begin{bmatrix} \mathbf{r}^{\mathbf{u}} \\ \mathbf{r}^{\boldsymbol{\lambda}} \end{bmatrix}$ in the linear iteration step k . Here, the subscript i for the nonlinear Newton iteration is dropped.

Table I. Two solid bodies example — Average (maximum) number of linear GMRES iterations per nonlinear iteration (over all 40 time steps) for different combinations of rotation angles α_y and α_z . As preconditioner, a 3 level AMG method (PA-AMG + PA-AMG, minimum aggregate size: 6 nodes) is used with different level smoothers. Within the block smoothers on all multigrid levels, symmetric Gauss–Seidel is used for the structural degrees of freedoms and ILU for the Lagrange multipliers.

(a) **Level smoother:** 3 CheapUzawa(0.7) with 1 SGS (0.7) + ILU(0)

		α_y				
		0	$\frac{1}{8}\pi$	$\frac{1}{4}\pi$	$\frac{3}{8}\pi$	$\frac{1}{2}\pi$
α_z	0	26.2 (30)	26.0 (30)	24.6 (29)	24.6 (30)	26.0 (29)
	$\frac{1}{8}\pi$	26.8 (30)	25.8 (37)	25.3 (30)	24.7 (30)	27.3 (31)
	$\frac{1}{4}\pi$	25.9 (31)	25.1 (32)	25.3 (30)	28.7 (40)	26.0 (31)
	$\frac{3}{8}\pi$	25.9 (31)	25.0 (30)	24.9 (30)	25.5 (33)	25.3 (29)
	$\frac{1}{2}\pi$	26.1 (30)	25.4 (32)	25.5 (30)	26.8 (31)	26.0 (30)

(b) **Level smoother:** 3 CheapBraessSarazin(1.9) with ILU(0)

		α_y				
		0	$\frac{1}{8}\pi$	$\frac{1}{4}\pi$	$\frac{3}{8}\pi$	$\frac{1}{2}\pi$
α_z	0	29.9 (37)	29.6 (36)	27.9 (33)	30.7 (38)	29.7 (37)
	$\frac{1}{8}\pi$	41.2 (59)	43.1 (68)	42.0 (58)	43.6 (63)	41.4 (59)
	$\frac{1}{4}\pi$	56.8 (82)	64.8 (86)	68.6 (95)	64.3 (86)	54.9 (75)
	$\frac{3}{8}\pi$	40.9 (60)	55.5 (74)	73.1 (102)	111.1 (141)	41.3 (61)
	$\frac{1}{2}\pi$	29.9 (37)	41.0 (55)	56.8 (79)	41.3 (61)	29.7 (36)

(c) **Level smoother:** 3 CheapSIMPLEC(0.7) with 1 SGS (0.7) + ILU(0)

		α_y				
		0	$\frac{1}{8}\pi$	$\frac{1}{4}\pi$	$\frac{3}{8}\pi$	$\frac{1}{2}\pi$
α_z	0	20.5 (26)	19.3 (21)	18.8 (23)	19.3 (26)	19.3 (20)
	$\frac{1}{8}\pi$	20.1 (22)	19.7 (27)	20.0 (24)	19.8 (21)	21.2 (25)
	$\frac{1}{4}\pi$	20.1 (25)	19.9 (23)	20.1 (23)	22.0 (26)	20.8 (30)
	$\frac{3}{8}\pi$	19.7 (22)	19.8 (22)	19.8 (23)	20.0 (26)	19.7 (23)
	$\frac{1}{2}\pi$	19.3 (20)	19.6 (22)	20.4 (27)	20.7 (25)	20.5 (28)

In this work, we focus on the behavior of the linear solver. Therefore, a fixed stopping criterion for all tested variants is chosen in (20). This allows the comparison of different preconditioning techniques including their effect on the linear solution strategy. For real world problems, and especially for coupled multiphysics problems, the task of choosing appropriate stopping criteria for both the nonlinear and linear solver turns out to be quite challenging. Usually, one would choose a combination of different (length-scaled) norms for the partial vectors \mathbf{r}^u and \mathbf{r}^λ . In order to reduce the solver time in the inner linear solver, it is recommended to adapt the linear (relative) solver tolerance according to the residual norms of the outer nonlinear solver.

6.1.5. Results First, the effect of the different saddle point smoothers on the number of linear iterations is explored. Table I summarizes the average number of linear iterations per time step for different combinations of the rotation angles α_y and α_z . The numbers in brackets denote the maximum

Table II. Two solid bodies example — Average (maximum) number of linear GMRES iterations per nonlinear iteration (over all 40 time steps) for different rotation angles α_y and α_z . As preconditioner, a 3 level AMG method (minimum aggregate size: 6 nodes) is used with different variants of CheapSIMPLEC. Within the CheapSIMPLEC method on all multigrid levels, symmetric Gauss-Seidel sweeps are used for the structural degrees of freedoms and ILU for the Lagrange multipliers.

(a) **Level smoother:** 1 CheapSIMPLEC(0.7) with 1 SGS (0.7) + ILU(0)

		α_y				
		0	$\frac{1}{8}\pi$	$\frac{1}{4}\pi$	$\frac{3}{8}\pi$	$\frac{1}{2}\pi$
α_z	0	37.5 (57)	35.6 (43)	35.2 (46)	36.1 (44)	35.7 (45)
	$\frac{1}{8}\pi$	36.4 (45)	37.6 (44)	38.5 (47)	36.3 (43)	38.1 (53)
	$\frac{1}{4}\pi$	39.1 (50)	37.8 (47)	37.0 (47)	39.2 (48)	37.0 (44)
	$\frac{3}{8}\pi$	36.3 (43)	37.0 (48)	37.8 (46)	37.2 (51)	36.4 (43)
	$\frac{1}{2}\pi$	36.2 (47)	38.0 (51)	39.8 (50)	36.9 (50)	36.6 (53)

(b) **Level smoother:** 1 CheapSIMPLEC(0.7) with 3 SGS (0.7) + ILU(0)

		α_y				
		0	$\frac{1}{8}\pi$	$\frac{1}{4}\pi$	$\frac{3}{8}\pi$	$\frac{1}{2}\pi$
α_z	0	29.5 (41)	27.5 (34)	28.4 (36)	28.5 (37)	27.1 (36)
	$\frac{1}{8}\pi$	29.3 (37)	30.1 (40)	32.7 (40)	28.8 (35)	28.4 (38)
	$\frac{1}{4}\pi$	30.0 (35)	29.6 (37)	28.5 (37)	31.1 (38)	29.2 (36)
	$\frac{3}{8}\pi$	28.8 (36)	27.8 (34)	28.1 (34)	29.1 (39)	28.1 (34)
	$\frac{1}{2}\pi$	28.5 (38)	28.4 (34)	29.7 (40)	27.9 (35)	27.1 (35)

(c) **Level smoother:** 3 CheapSIMPLEC(0.7) with 3 SGS (0.7) + ILU(0)

		α_y				
		0	$\frac{1}{8}\pi$	$\frac{1}{4}\pi$	$\frac{3}{8}\pi$	$\frac{1}{2}\pi$
α_z	0	15.5 (16)	17.0 (19)	16.9 (20)	15.3 (16)	15.2 (16)
	$\frac{1}{8}\pi$	16.2 (17)	16.1 (17)	16.3 (17)	15.9 (17)	16.1 (17)
	$\frac{1}{4}\pi$	16.0 (17)	16.0 (17)	16.0 (17)	15.6 (18)	15.6 (16)
	$\frac{3}{8}\pi$	15.7 (17)	15.9 (17)	15.9 (17)	15.5 (19)	15.6 (16)
	$\frac{1}{2}\pi$	15.6 (16)	15.8 (16)	15.7 (16)	15.4 (16)	15.1 (16)

number of linear iterations needed for solving one linear system during the full simulation, roughly indicating the variation of the number of linear iterations within the simulation. For the CheapUzawa smoother, the number of iterations does not show a dependence on the rotation angles α_y and α_z . Comparing the numbers from Table Ia with the results for the CheapBraessSarazin smoother in Table Ib, the CheapBraessSarazin smoother heavily suffers from the worse approximation of the displacement degrees of freedom using one internal hard-coded Jacobi sweep (cf. Section 5.1.2). The resulting iteration numbers show an obvious dependency on the rotation angles. With a CheapSIMPLEC block smoother, the number of iterations is lower than for the CheapUzawa smoother and independent from α_y and α_z when compared with the CheapBraessSarazin smoother (see Table Ic). So, the linear solver has some benefit from the two-way coupling of displacements and Lagrange multipliers within the AMG preconditioner. Compared to the Uzawa smoother, the additional computational costs for the CheapSIMPLEC method are very low with only one additional matrix-vector product by $\tilde{\mathbf{K}}^{-1}\mathbf{C}_1^T$

per iteration. Therefore, CheapSIMPLEC is the preferred level smoother for our further experiments with some cheap approximations for the internal single fields using some sweeps with a (symmetric) Gauss–Seidel (SGS) method for the structural block or incomplete LU factorization (ILU) for the Lagrange multipliers.

Table II illustrates how the number of CheapSIMPLEC coupling iterations and the quality of the single field smoothing methods within the CheapSIMPLEC smoother affect the number of linear iterations. Improving the quality of the Schur complement approximations within CheapSIMPLEC (see Tables IIa vs. IIb) as well as increasing the number of CheapSIMPLEC coupling iterations (see Table IIc) unsurprisingly reduces the number of linear solver iterations. Aside from the concrete parameter choices for the level smoother, one can even further reduce the number of linear iterations with a reasonable transfer operator smoothing strategy for the displacement block, e.g. as indicated in (6).

By intention, we do not report solver timings, since this example is too small to perform reasonable time measurements, especially when using 4 processors for altogether only 6300 degrees of freedom.

The intention of this example is to compare typical saddle point smoothers within a fully coupled AMG preconditioner. One can observe the expected behavior that increasing the number of smoothing sweeps reduces the number of linear GMRES iterations. However, in practice, the variant with a smaller number of GMRES iterations may not always be the fastest method. This example shows that the proper choice of block level smoothing is essential for the overall performance of a saddle point multigrid method. The particular choice of the block smoothing method gives the user full control over the quality of the coupling with field-specific parameters and allows for fine-grained adaptations and problem-specific optimizations.

With the experience from this example one can choose efficient level smoothers which provide results independent from the exact geometric configuration.

6.2. Weak scaling behavior

To assess the behavior of the proposed multilevel preconditioner applied to large-scale examples, we now report a weak scaling study. To exclude side effects (such as changes in the contact active set) and to fully focus on the behavior of the preconditioner and iterative linear solver, we study a simplified and linear contact problem.

6.2.1. Setup We consider a small block (dimensions $0.8m \times 0.8m \times 0.4m$) and a slightly bigger block (dimensions $1.0m \times 1.0m \times 0.5m$), where contact will occur between the large faces of both blocks. To reduce the complexity of the contact problem and to exclude nonlinearities due to changes in the contact active set, the faces opposite to the contact interface are fixed with Dirichlet boundary conditions, while the blocks initially penetrate each other at the contact interface by 0.001. The smaller block acts as the slave side and its entire contact area is initialized as “active”. Application of the contact algorithms will then result in a slight compression of both blocks, such that the initial penetration vanishes. This problem setup allows to distill the performance of the AMG preconditioner under uniform mesh refinement and weak scaling conditions.

Both blocks use a Neo–Hooke material with Young’s modulus $E = 10$ MPa and Poisson’s ratio $\nu = 0.3$. Denoting the mesh refinement factor with κ , both blocks are discretized with 2κ linear hexahedral elements along their longer edges and κ elements along the shorter edges. The Lagrange multiplier field is discretized with standard shape functions, i.e., linear Lagrange polynomials.

6.2.2. Solver and preconditioner settings As linear solver, a preconditioned GMRES method is used with a fully coupled multigrid approach as described in Section 3.2. The convergence criterion for the linear GMRES solver is set to

$$\left\| \frac{\mathbf{r}^k}{\mathbf{r}^0} \right\|_e < 10^{-8} \quad (21)$$

for the full residual vector $\mathbf{r}^k = \begin{bmatrix} \mathbf{r}^u \\ \mathbf{r}^\lambda \end{bmatrix}$ in the linear iteration step k . For the segregated block transfer operator as introduced in (7), we combine smoothed aggregation (SA-AMG) with a prolongator smoothing factor $\omega = 4/3$ for the displacement aggregates and plain non-smoothed (PA-AMG) transfer operators for the Lagrange multiplier aggregates. The restriction operators are built as the transposed of the prolongation operators. Coarsening stops when the total number of rows in the saddle-point system drops below 5000.

Following the guidance from Section 5.3 and the findings from Section 6.1.5, we apply 3 sweeps of CheapSIMPLE(0.8) as a level smoother, where both the predictor and corrector step are approximated

Table III. Weak scaling study — mesh refinement and hierarchy details

n^{proc}	κ	n_{DOF}^u	n_{DOF}^λ	$n_{\text{DOF}}^{\text{total}}$	$n_{\text{DOF}}^u/\text{proc}$	n^ℓ	$n_{\text{DOF}}^{\text{total}(n^\ell)}$	\mathcal{C}_A
4	20	211806	5043	216849	52951.5	3	948	1.16
8	25	405756	7803	413559	50719.5	3	1275	1.18
16	32	836550	12675	849225	52284.4	3	2661	1.23
24	36	1183038	15987	1199025	49293.2	3	3426	1.24
48	46	2439018	25947	2464965	50812.9	4	393	1.26
72	52	3505950	33075	3539025	48693.8	4	570	1.27
96	58	4845906	41067	4886973	50478.2	4	762	1.28
144	66	7110978	53067	7164045	49381.8	4	1071	1.28
192	73	9594396	64827	9659223	49970.8	4	1401	1.29
240	79	12134880	75843	12210723	50562.0	4	1713	1.29
288	84	14566110	85683	14651793	50576.8	4	2103	1.30
336	88	16729686	93987	16823673	49790.7	4	2421	1.30
384	92	19097550	102675	19200225	49733.2	4	2709	1.30
432	96	21678918	111747	21790665	50182.7	4	3090	1.30
480	99	23760600	118803	23879403	49501.2	4	3884	1.30

by 1 sweep of SGS each. We use the same level smoother layout on all multigrid levels except of the coarsest level. For the coarsest level, we compare two variants: a direct solver (marked as “LU”, requiring an expensive merging of the block matrix into a regular sparse matrix format and a subsequent LU factorization) vs. the CheapSIMPLE level smoother as on all other multigrid levels (marked as “CheapSIMPLE”). We will study the impact of the coarse solver on scaling behavior, iteration counts, as well as AMG setup and V-cycle timings.

6.2.3. Results For the weak scaling study, we target the load per rank to be 50k displacement DOFs. A uniform mesh refinement as outlined in Table III is performed with the finest mesh consisting of roughly 23.9 million unknowns in the saddle-point system. Therein, n^{proc} denotes the number of MPI ranks, κ the mesh refinement factor introduced in Section 6.2.1, n_{DOF}^u , n_{DOF}^λ , and $n_{\text{DOF}}^{\text{total}}$ the overall number of displacement, Lagrange multiplier, and total number of unknowns, $n_{\text{DOF}}^u/\text{proc}$ the average load per MPI rank. The mesh refinement factor κ is chosen such that the targeted average number of displacement unknowns per MPI rank is met. The number of levels n^ℓ , the actual size of the coarse level system $n_{\text{DOF}}^{\text{total}(n^\ell)}$ as well as the operator complexity \mathcal{C}_A (defined as the ratio of non-zero entries of all multigrid level matrices A_ℓ with $\ell = 0, \dots, n^\ell - 1$ and the number of non-zeros on the finest level) are reported in the last three columns of Table III. The operator complexity \mathcal{C}_A slightly increases with larger problem sizes, but does not grow beyond $\mathcal{C}_A \approx 1.30$.

The scaling study is run on our in-house cluster (20 nodes with 2x Intel Xeon Gold 5118 (Skylake-SP) 12 core CPUs, 480 cores in total, Mellanox Infiniband Interconnect). Solver performance in terms of the number of GMRES iterations as well as wall clock time spent in the AMG preconditioner construction and its application (i.e. sweeping through the V-cycle once per GMRES iteration) are summarized in Figure 7. With respect to iteration counts, both the direct solver as well as the block smoother yield the GMRES iteration count to be fairly independent of the overall mesh size, especially when taking the slight deviations of the load per core as detailed in Table III into account.

Using a direct solver on the coarsest multigrid level perfectly deals with all left-over error modes that are not properly handled by the block smoothers on the finer multigrid levels. Therefore, the “LU” variant results in a lower iteration count and is less sensitive to the actual size of the coarse level system. In contrast, when using the CheapSIMPLE block smoother as coarse level solver, we observe a somewhat higher iteration count together with an obvious impact of the number of multigrid levels on the iteration count (e.g., when moving from a 3-level hierarchy on 24 MPI ranks to a 4-level hierarchy on 48 MPI ranks). Increasing the number of multigrid levels helps to compensate the lack of coupling of the CheapSIMPLE block smoother compared to a direct solve on the coarsest level.

We show the CheapSIMPLE variants for the coarse solve since this variant in practice might be superior with respect to the overall performance: Regarding AMG setup time, the cost for the direct coarse solver becomes more expensive with a growing coarse system size, since (i) more matrix entries need to be moved from a block sparse matrix layout to a regular sparse matrix layout and (ii) a larger matrix needs to be factorized. For the CheapSIMPLE block smoother as coarse level solver, no additional communication is required on the coarsest level. Hence, the mild slope in the graph for preconditioner setup time for the CheapSIMPLE coarse level solver originates only from the Galerkin

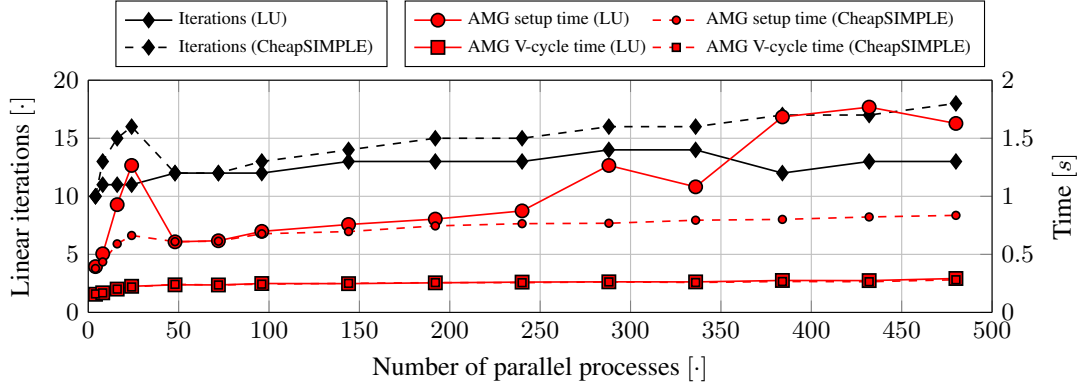


Figure 7. Weak scaling study — Linear solver iterations and AMG setup and V-cycle timings demonstrate proper weak scaling behavior. The type of coarse solver is indicated in parentheses.

product during hierarchy construction. In terms of preconditioner application, i.e. time spent in the AMG V-cycle, both methods are almost identical and scale very well.

While proper weak scalability up to 23.9 million unknowns on 480 MPI ranks has been demonstrated, the actual choice of the coarse solver in practical applications also depends on other factors such as the frequency of rebuilding the AMG hierarchy or the balance of setup time, V-cycle time, and cost of the additional GMRES iterations, which all together impact the overall time to solution. In the next example, we will put attention on effects for the linear solver caused by changes in the active set of contact nodes for larger problems.

6.3. 1000 deformable rings

Even though only a 2D example of moderate size, the 1000 rings example comes with frequent changes in the contact active set and, thus, tests the robustness of the proposed multigrid methods. Particularly, we are interested in the comparison of the fully coupled multigrid approach and the nested multigrid approach as described in Section 3.2.

6.3.1. Setup This example consists of 1000 deformable rings (Neo-Hookean material with $E = 210$ GPa, $\nu = 0.3$ and $\rho_0 = 7.83 \cdot 10^{-6} \frac{kg}{m^3}$) arranged in a rectangle (see Figure 8). A gravitational force is inducing an acceleration towards a rigid wall. The simulated time extends to 2.0s with a time step size of $\Delta t = 0.0005s$, yielding 4000 time steps in total. The full mesh consists 110000 nodes with 110 nodes for each deformable ring.

6.3.2. Stopping criteria In each time step, the nonlinear system is handled by a semi-smooth Newton method. As convergence criteria, we have chosen

$$\|\Delta \mathbf{u}\|_e < 10^{-8} \wedge \left(\|\mathbf{r}_i^{\mathbf{u}}\|_e < 10^{-8} \wedge \|\mathbf{r}_i^{\lambda}\|_e < 10^{-6} \right). \quad (22)$$

Here, $\mathbf{r}_i^{\mathbf{u}}$ and \mathbf{r}_i^{λ} denote the nonlinear residual for the displacement and Lagrange multiplier variables after i Newton iterations, respectively. Similarly, $\Delta \mathbf{u}$ denotes the solution increment for the displacement variables in the i -th Newton iteration.

For solving the linear systems arising during the simulation a GMRES solver is applied with different variants of AMG preconditioners listed in Table IV. The relative tolerance of convergence for the GMRES solver is set to $\left\| \frac{\mathbf{r}^k}{\mathbf{r}^0} \right\|_e < 10^{-8}$ with $\mathbf{r}^k = \begin{bmatrix} \mathbf{r}^{\mathbf{u}} \\ \mathbf{r}^{\lambda} \end{bmatrix}$ being the full residual vector in the linear iteration step k . Here, the subscript i for the nonlinear Newton step is dropped.

The stopping criteria (22) for the nonlinear solver are carefully chosen in such a way that the simulations with all the tested preconditioner variants shown in Table IV always result in the same number of nonlinear iterations. This way we can directly compare the number of linear iterations of all tested solver variants which allow to draw some conclusions on the multigrid preconditioners.

6.3.3. Results Table IV provides an overview of the chosen preconditioner parameters for the level smoothers. For both the nested multigrid schemes and the fully coupled multigrid schemes we apply

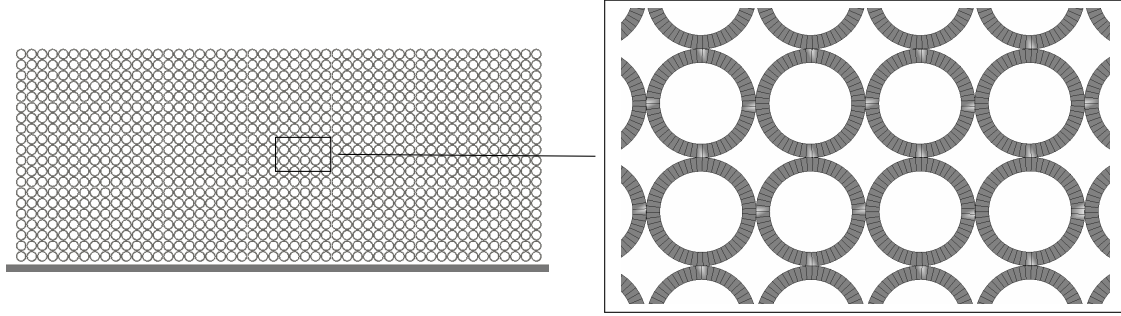
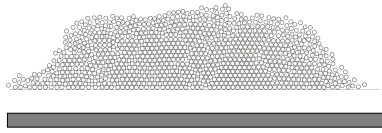
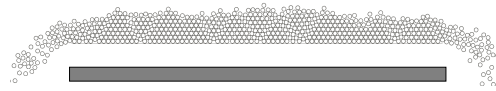
(a) Initial configuration with close-up view of the meshed deformable rings at $t = 0.0s$ (b) Configuration at $t = 0.5s$ (c) Configuration at $t = 2.0s$

Figure 8. 1000 deformable rings — Characteristic stages at different times.

Table IV. 1000 deformable rings — Different AMG variants.

Preconditioner type	
Full multigrid based methods	Nested multigrid based methods
PA-AMG (CheapSIMPLE) Transfer operators: PA-AMG Level smoother: 1 CheapSIMPLEC(0.8) – Pred. smoother: 3 SGS (0.8) – Corr. smoother: ILU (0)	CheapSIMPLE (PA-AMG) Block prec.: 1 CheapSIMPLEC(0.8) – Pred. smoother: AMG – Transfer op.: PA-AMG – Level sm.: 1 SGS (0.8) – Corr. smoother: KLU
Emin (CheapSIMPLE) Transfer operators: Emin Level smoother: 1 CheapSIMPLEC(0.8) – Pred. smoother: 3 SGS (0.8) – Corr. smoother: ILU (0)	CheapSIMPLE (SA-AMG) Block prec.: 1 CheapSIMPLEC(0.8) – Pred. smoother: AMG – Transfer op.: SA-AMG (0.8) – Level sm.: 1 SGS (0.8) – Corr. smoother: KLU

the level smoother on all multigrid levels including the coarsest level. For each class of multigrid preconditioners, only those variants are presented, that give the best overall timings and are able to accomplish the all 4000 time steps of the full simulation.

The multigrid parameters are chosen to be the same for all preconditioner variants: the minimum size of the aggregates is set to 6 nodes for the two-dimensional problem and the maximum coarse level size is set to 1000 degrees of freedom, yielding a 3 level multigrid method.

For the fully coupled AMG variants, different transfer operator strategies are compared, namely the non-smoothed (PA-AMG) transfer operators and the energy minimization approach with local damping parameters for transfer operator smoothing denoted by Emin, cf. [64]. For the nested AMG variants, we compare transfer operators based on PA-AMG and SA-AMG. All the simulations have been run on 16 cores (spread over 2 Intel Xeon E5-2670 Octocore CPUs).

Figure 9a shows for each time step the overall number of linear iterations for all nonlinear iterations stacked onto each other. Compared to the CheapSIMPLE based methods, the fully coupled AMG

Table V. 1000 deformable rings — Exemplary timings in [s] of the different preconditioning variants from Table IV for the full simulation (4000 time steps).

Method	Setup costs	Solver time	Overall solver time
PA-AMG (CheapSIMPLE)	11870	10013	21883
Emin (CheapSIMPLE)	12820	8679	21499
CheapSIMPLE (PA-AMG)	11730	11103	22833
CheapSIMPLE (SA-AMG)	12300	10763	23063

variants need a significantly lower number of linear iterations to solve the problem in each time step. This can be explained by the better approximation of the displacement degrees of freedom using 3 instead of 1 damped Gauss–Seidel sweeps. One can also see how the number of linear iterations correlates with the number of active contact nodes. Obviously, the fully coupled AMG variants are less sensitive to the changes in the number of active contact nodes than the CheapSIMPLE based methods. In Figure 9b, the overall linear solver time is given for each time step. Again, the fully coupled AMG variants perform better, although the difference is less pronounced due to the higher computational effort of fully coupled AMG implementations. The number of nonlinear iterations per time step is 2 for the initial phase without contact and then varies between 4 and 6, when contact is active. It is the same for all preconditioner variants due to the particular choice of stopping criteria of the nonlinear solver (see Section 6.3.2). Moreover, the number of nonlinear iterations is independent of the size of the active set.

Finally, overall timings for the linear solver including preconditioner setup are reported in Table V. The overall solver time only varies slightly for all given variants. The variants Emin (CheapSIMPLE) and CheapSIMPLE (SA-AMG) need more setup time due to the additional effort of prolongator smoothing. Emin (CheapSIMPLE) can fully amortize the additional setup cost during the solve phase due to coarse grid corrections, that respect the contact constraints, while CheapSIMPLE (SA-AMG) does not have this benefit and, thus, requires the largest overall solver time. To show a fair comparison, the preconditioner has been rebuilt in every nonlinear iteration step. Of course, re-use strategies might increase amortization of setup costs and reduce overall solver timings when solving actual problems.

6.4. Two tori impact example

Inspired by a similar analysis in [95], with the two tori impact example we test the proposed multigrid variants from Section 3.2 on a complex 3D contact example with more than 1 million unknowns. Please refer to [53] for the detailed problem setup of the two tori impact example with geometry and load conditions.

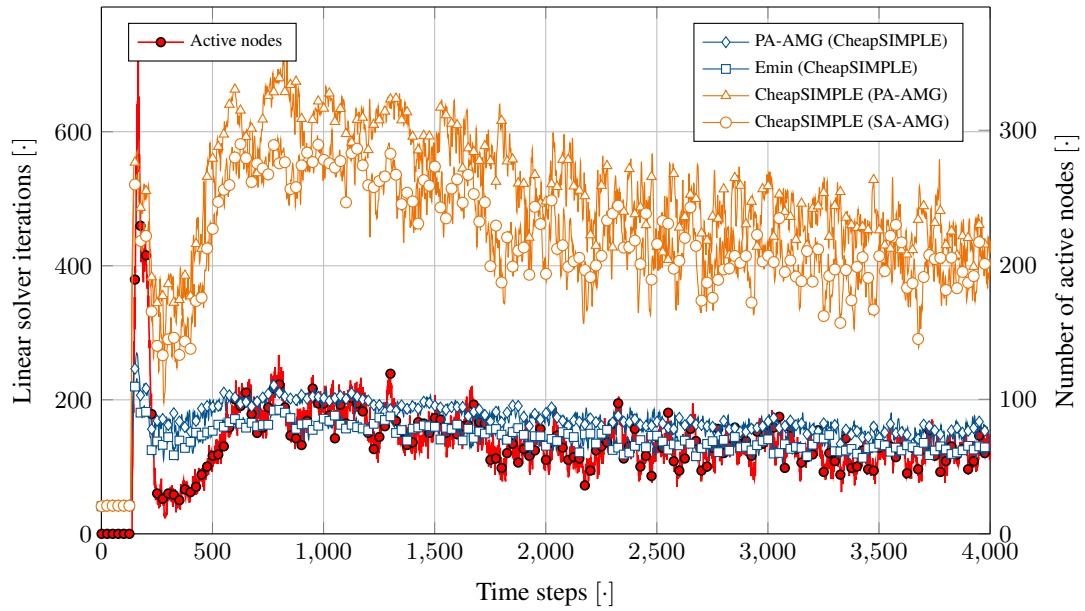
6.4.1. Setup The example consists of two thin-walled tori with a Neo-Hookean material model ($E = 2250$ GPa, $\nu = 0.3$, $\rho_0 = 0.1 \frac{kg}{m^3}$) with a major and minor radius of $76m$ and $24m$ and a wall thickness of $4.5m$. The left torus in Figure 10 lies in the xy -plane with resting initial conditions. The right torus has been rotated around the y -axis by 45 degrees and has an initial velocity of $1.0 \frac{m}{s}$ directed towards the left torus. The simulated time is $10s$ divided into 200 time steps with a time step size of $0.05s$ using a generalized- α time integration scheme [18]. The mesh consists of 284672 first-order hexahedral elements with 350208 nodes.

With the rather complex geometry and contact configuration, that heavily and frequently changes over time, this example can be considered as a representative test for the robustness and efficiency of the tested numerical methods.

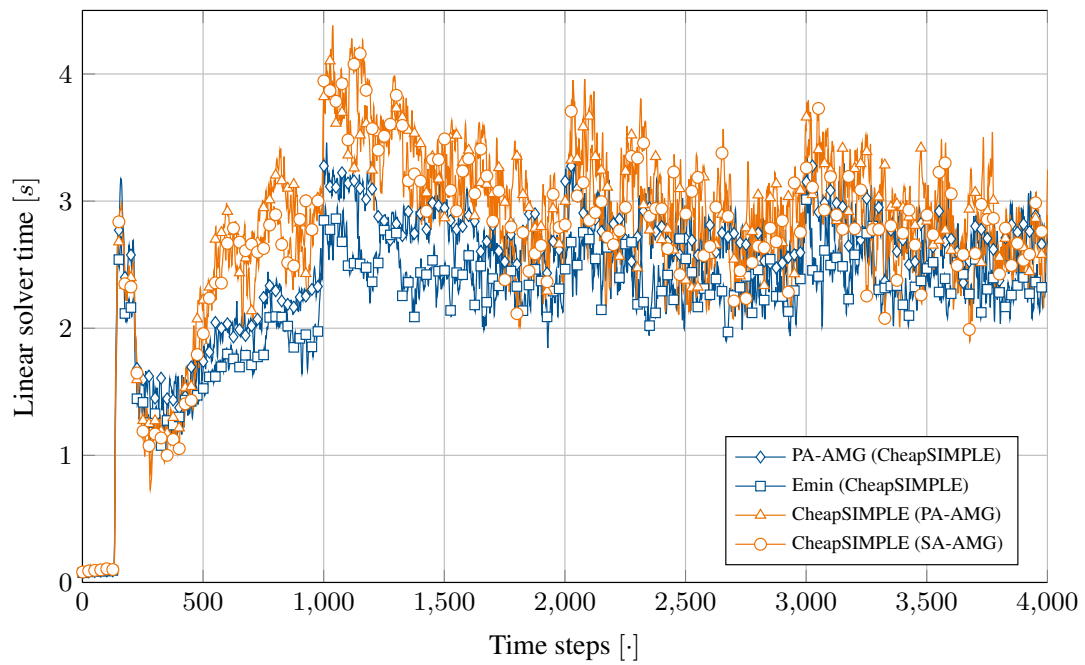
6.4.2. Stopping criteria The stopping criteria for the semi-smooth Newton method are chosen as

$$\|\Delta \mathbf{u}\|_e < 10^{-7} \wedge \left(\left\| \begin{array}{c} \mathbf{r}_i^u \\ \mathbf{r}_0^u \end{array} \right\|_e < 10^{-8} \wedge \left\| \mathbf{r}_i^\lambda \right\|_e < 10^{-4} \right). \quad (23)$$

Here, \mathbf{r}_i^u and \mathbf{r}_i^λ denote the (nonlinear) residual for the displacement variables and Lagrange multipliers in the i -th Newton iteration, respectively. The quantity $\Delta \mathbf{u}$ describes the solution increment for the displacement variables only. Again, the stopping criteria in (23) for the nonlinear solver are specifically chosen to produce the same number of nonlinear iterations for all tested solver variants. This allows for an easy comparison of the number of linear iterations during the simulation.



(a) Number of linear GMRES iterations for all nonlinear iterations over time steps for different preconditioner variants. The red curve shows the number of active nodes for each time step.



(b) Overall linear solver time over time steps.

Figure 9. 1000 collapsing rings example — Results for different AMG preconditioner variants.

As linear solver a preconditioned GMRES method is used with the convergence criterion

$$\left\| \frac{\mathbf{r}^k}{\mathbf{r}^0} \right\|_e < 10^{-8} \quad (24)$$

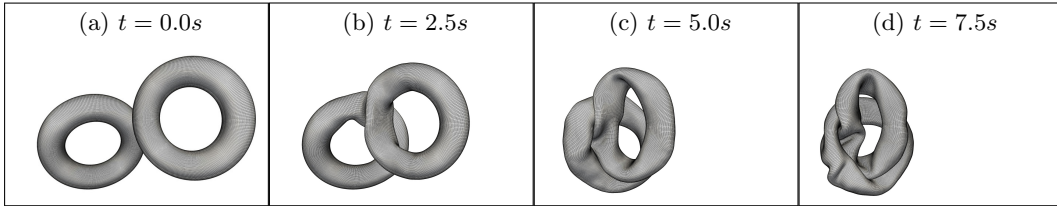


Figure 10. Two tori impact example — Characteristic stages of deformation.

Table VI. Two tori impact example — Different AMG variants.

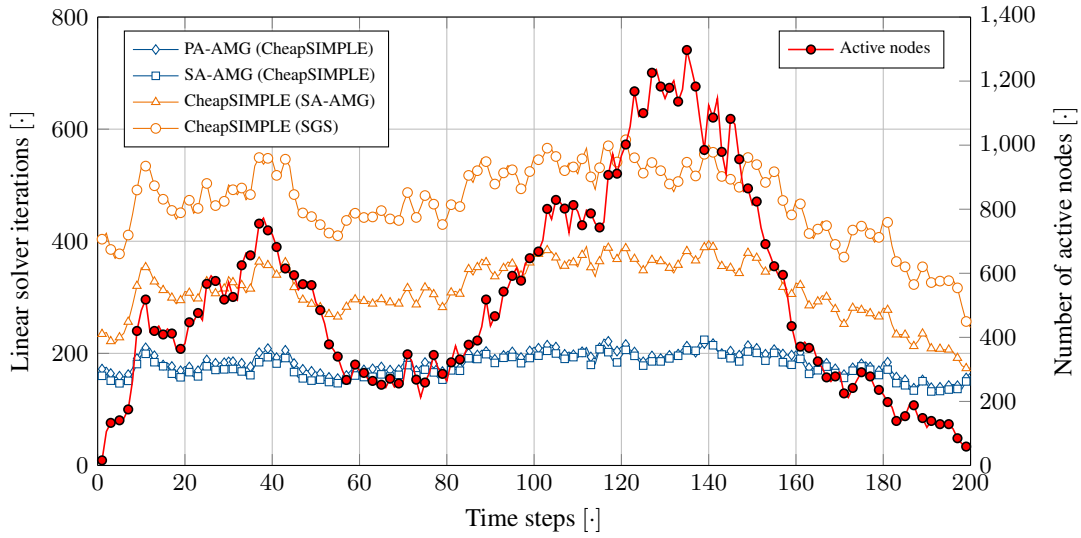
Full multigrid based methods	SIMPLE based methods
PA-AMG (CheapSIMPLE) Transfer operators: PA-AMG Level smoother: 1 CheapSIMPLEC(0.8) – Pred. smoother: 1 SGS (0.8) – Corr. smoother: ILU (0)	CheapSIMPLE (SGS) Transfer operators: – Block prec.: 2 CheapSIMPLEC(0.8) – Pred. smoother: 3 SGS (0.8) – Corr. smoother: ILU (0)
SA-AMG (CheapSIMPLE) Transfer operators: SA-AMG (0.4) Level smoother: 1 CheapSIMPLEC(0.8) – Pred. smoother: 1 SGS (0.8) – Corr. smoother: ILU (0)	CheapSIMPLE (SA-AMG) Block prec.: 2 CheapSIMPLEC(0.8) – Pred. smoother: AMG – Transfer op.: SA-AMG (0.4) – Level sm.: 2 SGS (0.8) – Corr. smoother: ILU (0)

for the full residual vector $\mathbf{r}^k = \begin{bmatrix} \mathbf{r}^u \\ \mathbf{r}^\lambda \end{bmatrix}$ in the linear iteration step k .

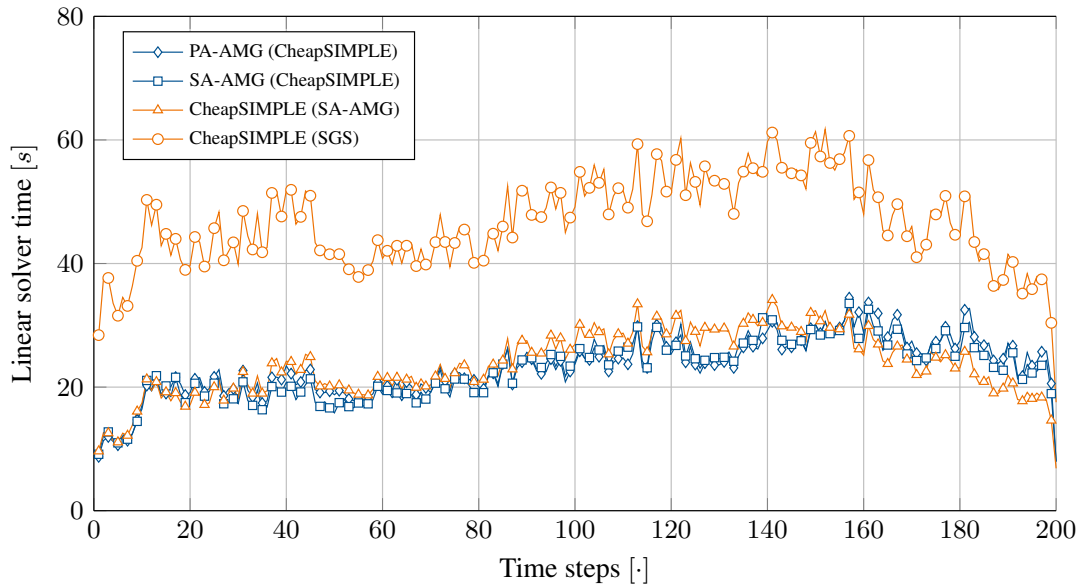
6.4.3. Results An overview of the different tested preconditioner variants is given in Table VI. We study variants with the fully coupled multigrid approach, the nested multigrid approach, and a SIMPLE based variant without multigrid at all. For the fully coupled AMG variants, the transfer operators for the displacement blocks are varied. Particularly, non-smoothed transfer operators (PA-AMG) are compared to smoothed aggregation transfer operators (SA-AMG). For the multigrid schemes there is no special treatment of the coarsest level. For the fully coupled multigrid schemes as well as for the nested multigrid method we apply the same level smoother on all multigrid levels including the coarsest level. All the simulations have been run on 16 cores (spread over 2 Intel Xeon E5-2670 Octocore CPUs).

Figure 11a shows the overall number of linear iterations performed to solve all nonlinear iterations for each time step. All preconditioner variants require exactly the same number of nonlinear iterations per time step due to the particular choice of stopping criteria in (23), ranging between 6 and 10 nonlinear iterations per time step, while also the number of nonlinear iterations is independent of the size of the active set. Obviously, the SIMPLE based methods need more linear iterations than the AMG based methods. In this example, there is nearly no difference between the non-smoothed transfer operator variant PA-AMG (CheapSIMPLE) and the smoothed transfer operator variant SA-AMG (CheapSIMPLE). Furthermore, there is no clear and obvious correlation between the number of linear iterations and the number of active nodes. This indicates that the fully coupled multigrid method is robust and efficient with regard to the increasing complexity of the contact configuration over time, which is not the case for cheaper methods such as the SIMPLE based methods. Evidently, one can see a significant drop in the linear iterations for the last 20 time steps of the simulation, which may correspond to the small number of nodes in contact.

When looking at the corresponding solver timings over the time steps in Figure 11b, one finds the CheapSIMPLE (SA-AMG) method to be very close to the AMG based methods PA-AMG (CheapSIMPLE) and SA-AMG (CheapSIMPLE). For the AMG based methods, one sweep with a CheapSIMPLEC method is applied on each level, which internally uses one sweep with a symmetric Gauss-Seidel iteration for the primary variable and one ILU sweep for the constraint equation. That is, quite a lot of time is invested in the coupling on all levels with the comparably expensive ILU method.



(a) Number of linear GMRES iterations for all nonlinear iterations over time steps for different preconditioner variants. The red curve shows the number of active nodes for each time step.



(b) Overall linear solver time over time steps.

Figure 11. Two tori impact example — Results for different saddle point preconditioner variants.

In contrast to the AMG based method, the CheapSIMPLE (SA-AMG) method uses 2 sweeps with a CheapSIMPLE preconditioner for the coupling (on the finest level only). Internally, a 3 level AMG multigrid is used with 2 symmetric Gauss–Seidel sweeps for the level smoother and an ILU sweep for the constraint correction equation. These parameters have been found to result in a reasonably low number of linear iterations. For this example the experiment shows that the CheapSIMPLE (SA-AMG) method needs twice as many iterations as the SA-AMG (CheapSIMPLE) method, but the costs per iteration are only half of the costs of the SA-AMG (CheapSIMPLE). Nevertheless, the AMG based methods seem to have a small advantage, when the number of nodes in contact increases.

Last but not least, the overall timings for the linear solver are reported in Table VII. Except for the CheapSIMPLE (SGS) variant, which is far away from the others, there is no clear winner. The setup

Table VII. Two tori impact example — Exemplary timings in [s] of the different preconditioning variants from Table VI for the full simulation over 200 time steps.

Method	Setup costs	Solver time	Overall solver time
PA-AMG (CheapSIMPLE)	11630	4658	16288
SA-AMG (CheapSIMPLE)	12250	4564	16814
CheapSIMPLE (SA-AMG)	12130	4731	16861
CheapSIMPLE (SGS)	10270	9320	19590

costs are quite close, since the same transfer operators have to be built for all methods with only a small difference for smoothed versus non-smoothed transfer operators. To show a fair comparison, the preconditioner has been rebuilt in every nonlinear iteration step. Of course, re-use strategies might increase amortization of setup costs and reduce overall solver timings when solving actual problems.

7. CONCLUSION

We have presented algebraic multigrid schemes designed for saddle point problems arising from contact problems using mortar finite element methods. The new fully coupled multigrid scheme has the advantage that the contact constraints are considered on all multigrid levels, which significantly reduces the number of iterations. It gives the user full control over the coupling process by appropriately choosing the solver parameters. Additionally, we have proposed a novel aggregation method for the Lagrange multipliers, which reuses existing aggregation information at the contact interface. We have demonstrated the robustness and efficiency of the overall multigrid method for large examples with increasingly complex contact configurations over time as well as its weak scalability up to 23.9 million unknowns on 480 MPI ranks.

REFERENCES

- [1] BACI: A Comprehensive Multi-Physics Simulation Framework. <https://baci.pages.gitlab.lrz.de/website>.
- [2] <https://trilinos.github.io/muelu.html>.
- [3] <https://trilinos.github.io>.
- [4] <https://trilinos.github.io/xpetra.html>.
- [5] M. F. Adams. Algebraic multigrid methods for constrained linear systems with applications to contact problems in solid mechanics. *Numerical Linear Algebra with Applications*, 11(2-3):141–153, 2004.
- [6] F. B. Belgacem. The mortar finite element method with Lagrange multipliers. *Numerische Mathematik*, 84(2):173–197, 1999.
- [7] M. Benzi. Preconditioning techniques for large linear systems: a survey. *Journal of Computational Physics*, 182(2):418–477, 2002.
- [8] L. Berger-Vergiat, C. A. Glusa, J. J. Hu, M. Mayr, A. Prokopenko, C. M. Siefert, R. S. Tuminaro, and T. A. Wiesner. MueLu User’s Guide 2.0. Technical Report SAND2019-0537, Sandia National Laboratories, Albuquerque, NM (USA) 87185, 2019.
- [9] C. Bernardi, Y. Maday, and A. T. Patera. A new nonconforming approach to domain decomposition: The mortar element method. In H. Brezis and J. Lions, editors, *Nonlinear partial differential equations and their applications*, pages 13–51. Pitman/Wiley: London/New York, 1994.
- [10] J. Bonari, M. R. Marulli, N. Hagemeyer, M. Mayr, A. Popp, and M. Paggi. A multi-scale FEM-BEM formulation for contact mechanics between rough surfaces. *Computational Mechanics*, 65:731–749, 2020.
- [11] D. Braess and R. Sarazin. An efficient smoother for the stokes problem. *Applied Numerical Mathematics*, 23:3–19, 1997.
- [12] D. Braess, M. Dryja, and W. Hackbush. A multigrid method for nonconforming FE-discretisations with application to non-matching grids. *Computing*, 63(1):1–25, 1999.
- [13] D. Braess and W. Dahmen. Stability estimates of the mortar finite element method for 3-dimensional problems. *East-West Journal of Numerical Mathematics*, 6:249–264, 1998.
- [14] D. Braess, W. Dahmen, and C. Wieners. A multigrid algorithm for the mortar finite element method. *SIAM Journal on Numerical Analysis*, 37(1):pp. 48–69, 2000.
- [15] J. H. Bramble, J. E. Pasciak, and A. T. Vassilev. Analysis of the inexact uzawa algorithm for saddle point problems. *SIAM Journal on Numerical Analysis*, 34(3):pp. 1072–1092, 1997.
- [16] F. J. Cavalieri and A. Cardona. Three-dimensional numerical solution for wear prediction using a mortar contact algorithm. *International Journal for Numerical Methods in Engineering*, 96(8):467–486, 2013.
- [17] P. W. Christensen, A. Klarbring, J. S. Pang, and N. Strömberg. Formulation and comparison of algorithms for frictional contact problems. *International Journal for Numerical Methods in Engineering*, 42(1):145–173, 1998.

- [18] J. Chung and G. Hulbert. A Time Integration Algorithm for Structural Dynamics With Improved Numerical Dissipation: The Generalized- α Method. *Journal of Applied Mechanics*, 60(2):371–375, 1993.
- [19] C. Danowski, V. Gravemeier, L. Yoshihara, and W. A. Wall. A monolithic computational approach to thermo-structure interaction. *International Journal for Numerical Methods in Engineering*, 95(13):1053–1078, 2013.
- [20] L. De Lorenzis, I. Temizer, P. Wriggers, and G. Zavarise. A large deformation frictional contact formulation using NURBS-based isogeometric analysis. *International Journal for Numerical Methods in Engineering*, 87(13):1278–1300, 2011.
- [21] L. De Lorenzis, P. Wriggers, and T. J. Hughes. Isogeometric contact: a review. *GAMM-Mitteilungen*, 37(1):85–123, 2014.
- [22] H. Elman, V. E. Howle, J. Shadid, R. Shuttleworth, and R. Tuminaro. A taxonomy and comparison of parallel block multi-level preconditioners for the incompressible Navier-Stokes equations. *Journal of Computational Physics*, 227(3):1790–1808, January 2008.
- [23] H. C. Elman and G. H. Golub. Inexact and preconditioned Uzawa algorithms for saddle point problems. *SIAM Journal on Numerical Analysis*, 31(6):pp. 1645–1661, 1994.
- [24] P. Farah, W. A. Wall, and A. Popp. An implicit finite wear contact formulation based on dual mortar methods. *International Journal for Numerical Methods in Engineering*, 111(4):325–353, 2016.
- [25] M. Franke, C. Hesch, and P. Betsch. An augmentation technique for large deformation frictional contact problems. *International Journal for Numerical Methods in Engineering*, pages –, 2013.
- [26] M. W. Gee, J. J. Hu, and R. S. Tuminaro. A new smoothed aggregation multigrid method for anisotropic problems. *Numerical Linear Algebra with Applications*, 16:19–37, 2009.
- [27] M. W. Gee, U. Küttler, and W. A. Wall. Truly monolithic algebraic multigrid for fluid–structure interaction. *International Journal for Numerical Methods in Engineering*, 85(8):987–1016, 2011.
- [28] M. Gitterle, A. Popp, M. W. Gee, and W. A. Wall. Finite deformation frictional mortar contact using a semi-smooth Newton method with consistent linearization. *International Journal for Numerical Methods in Engineering*, 84(5):543–571, 2010.
- [29] J. Gopalakrishnan and J. E. Pasciak. Multigrid for the mortar finite element method. *SIAM Journal on Numerical Analysis*, 37(3):1029–1052, 2000.
- [30] M. Griebel, T. Neunhoffer, and H. Regler. Algebraic multigrid methods for the solution of the Navier–Stokes equations in complicated geometries. *International Journal for Numerical Methods in Fluids*, 26(3):281–301, 1998.
- [31] W. Hackbusch. *Iterative Solution of Large Sparse Systems of Equations*, volume 95 of *Applied Mathematical Sciences*. Springer, 1994.
- [32] W. Hackbusch. *Multi-Grid Methods and Applications*, volume 4 of *Series in Computational Mathematics*. Springer, 1st edition edition, 1985.
- [33] S. Hartmann, S. Brunssen, E. Ramm, and B. I. Wohlmuth. Unilateral non-linear dynamic contact of thin-walled structures using a primal-dual active set strategy. *International Journal for Numerical Methods in Engineering*, 70(8):883–912, 2007.
- [34] M. Hiermeier, W. A. Wall, and A. Popp. A truly variationally consistent and symmetric mortar-based contact formulation for finite deformation solid mechanics. *Computer Methods in Applied Mechanics and Engineering*, 342:532–560, 2018.
- [35] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semi-smooth Newton method. *SIAM Journal on Optimization*, 13(3):865–888, 2002.
- [36] S. Hieber and B. I. Wohlmuth. A primal-dual active set strategy for non-linear multibody contact problems. *Computer Methods in Applied Mechanics and Engineering*, 194(27-29):3147–3166, 2005.
- [37] S. Hieber, G. Stadler, and B. I. Wohlmuth. A primal-dual active set algorithm for three-dimensional contact problems with Coulomb friction. *SIAM Journal on Scientific Computing*, 30(2):572–596, 2008.
- [38] A. Janka. Smoothed aggregation multigrid for a Stokes problem. *Computing and Visualization in Science*, 11(3):169–180, 2008.
- [39] C. Keller, N. I. M. Gould, and A. J. Wathen. Constrained preconditioning for indefinite linear systems. *SIAM Journal on Matrix Analysis & Applications*, 21(4):1300 – 1317, 2000.
- [40] R. Kornhuber and R. Krause. Adaptive multigrid methods for signorini’s problem in linear elasticity. *Computing and Visualization in Science*, 4:9–20, 2001.
- [41] R. H. Krause and B. I. Wohlmuth. Multigrid methods for mortar finite elements. In *Multigrid methods, VI (Gent, 1999)*, volume 14 of *Lecture Notes in Computational Science and Engineering*, pages 136–142. Springer, Berlin, 2000.
- [42] B. P. Lamichhane, R. P. Stevenson, and B. I. Wohlmuth. Higher order mortar finite element methods in 3D with dual Lagrange multiplier bases. *Numerische Mathematik*, 102(1):93–121, 2005.
- [43] U. Langer and H. Yang. Numerical Simulation of Fluid–Structure Interaction Problems with Hyperelastic Models: A Monolithic Approach. *Mathematics and Computers in Simulation*, 145:186–208, 2018.
- [44] J. Lengiewicz and S. Stupkiewicz. Continuum framework for finite element modelling of finite wear. *Computer Methods in Applied Mechanics and Engineering*, 205-208:178–188, 2012.
- [45] Q. Li and V. L. Popov. Adhesive contact between a rigid body of arbitrary shape and a thin elastic coating. *Acta Mechanica*, 230(7):2447–2453, 2019.
- [46] R. Lonsdale. *An algebraic multigrid scheme for solving the Navier-Stokes equations on unstructured meshes*, volume 7 of *Numerical Methods in Laminar and Turbulent Flow*, pages 1432–1442. Pineridge press, Swansea, U.K., 1991.
- [47] M. Mayr, M. Noll, and M. W. Gee. A hybrid interface preconditioner for monolithic fluid-structure interaction solvers. *Advanced Modeling and Simulation in Engineering Sciences*, 7:15, 2020.
- [48] E. Milanese, T. Brink, R. Aghababaei, and J.-F. Molinari. Emergence of self-affine surfaces during adhesive wear. *Nature Communications*, 10(1116), 2019.

- [49] Y. Notay. A new analysis of block preconditioners for saddle point problems. *SIAM Journal on Matrix Analysis and Applications*, 35(1):143–173, 2014.
- [50] S. V. Patankar and D. B. Spalding. A calculation procedure for heat, mass and momentum transfer in three-dimensional parabolic flows. *International Journal of Heat and Mass Transfer*, 15:1787–1972, 1972.
- [51] S. V. Patankar. *Numerical Heat Transfer and Fluid Flow*. Hemisphere Publishing Corporation, New York, 1980.
- [52] V. L. Popov. *Contact Mechanics and Friction: Physical Principles and Applications*. Springer, Berlin, 2nd edition, 2017.
- [53] A. Popp. *Mortar Methods for Computational Contact Mechanics and General interface Problems*. PhD thesis, Technische Universität München, 2012.
- [54] A. Popp, M. Gitterle, M. W. Gee, and W. A. Wall. A dual mortar approach for 3D finite deformation contact with consistent linearization. *International Journal for Numerical Methods in Engineering*, 83(11):1428–1465, 2010.
- [55] A. Popp, B. I. Wohlmuth, M. W. Gee, and W. A. Wall. Dual quadratic mortar finite element methods for 3D finite deformation contact. *SIAM Journal on Scientific Computing*, 34:B421–B446, 2012.
- [56] A. Popp and W. A. Wall. Dual mortar methods for computational contact mechanics – overview and recent developments. *GAMM-Mitteilungen*, 37(1):66–84, 2014.
- [57] A. Popp, M. W. Gee, and W. A. Wall. A finite deformation mortar contact formulation using a primal–dual active set strategy. *International Journal for Numerical Methods in Engineering*, 79(11):1354–1391, 2009.
- [58] A. Popp, A. Seitz, M. W. Gee, and W. A. Wall. Improved robustness and consistency of 3D contact algorithms based on a dual mortar approach. *Computer Methods in Applied Mechanics and Engineering*, 264:67–80, 2013.
- [59] M. A. Puso and T. A. Laursen. A mortar segment-to-segment contact method for large deformation solid mechanics. *Computer Methods in Applied Mechanics and Engineering*, 193(6-8):601–629, 2004.
- [60] M. A. Puso and T. A. Laursen. A mortar segment-to-segment frictional contact method for large deformations. *Computer Methods in Applied Mechanics and Engineering*, 193(45-47):4891–4913, 2004.
- [61] L. Qi and J. Sun. A nonsmooth version of Newton’s method. *Mathematical Programming*, 58(1):353–367, 1993.
- [62] Y. Saad and M. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 7(3):856–869, 1986.
- [63] Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, second edition edition, 2003.
- [64] M. Sala and R. S. Tuminaro. A new Petrov-Galerkin smoothed aggregation preconditioner for nonsymmetric linear systems. *SIAM Journal on Scientific Computing*, 31(1):143–166, 2008.
- [65] R. A. Sauer. Enriched contact finite elements for stable peeling computations. *International Journal for Numerical Methods in Engineering*, 87(6):593–616, 2011.
- [66] A. Seitz, A. Popp, and W. A. Wall. A semi-smooth newton method for orthotropic plasticity and frictional contact at finite strains. *Computer Methods in Applied Mechanics and Engineering*, 285(0):228–254, 2015.
- [67] A. Seitz, P. Farah, J. Kremheller, B. I. Wohlmuth, W. A. Wall, and A. Popp. Isogeometric dual mortar methods for computational contact mechanics. *Computer Methods in Applied Mechanics and Engineering*, 301:259–280, 2016.
- [68] K. Stüben. *Multigrid*, Appendix A: An introduction to algebraic multigrid, pages 413–533. U. Trottenberg and C.W. Oosterlee and A. Schüller, 2001.
- [69] I. Temizer, P. Wriggers, and T. J. R. Hughes. Contact treatment in isogeometric analysis with NURBS. *Computer Methods in Applied Mechanics and Engineering*, 200(9, 12):1100–1112, 2011.
- [70] I. Temizer, P. Wriggers, and T. J. R. Hughes. Three-dimensional mortar-based frictional contact treatment in isogeometric analysis with NURBS. *Computer Methods in Applied Mechanics and Engineering*, 209–212(0):115–128, 2012.
- [71] T. E. Tezduyar, S. Sathe, and K. Stein. Solution techniques for the fully discretized equations in computation of fluid–structure interactions with the space–time formulations. *Computer Methods in Applied Mechanics and Engineering*, 195(41–43):5743–5753, 2006.
- [72] U. Trottenberg, C. Oosterlee, and A. Schüller. *Multigrid*. Academic Press, 2001. ISBN 9780127010700.
- [73] A. I. Vakis, V. A. Yastrebov, J. Scheibert, L. Nicola, D. Dini, C. Minfray, A. Almqvist, M. Paggi, S. Lee, G. Limbert, J.-F. Molinari, G. Ancaix, R. Aghababaei, S. Echeverri Restrepo, A. Papangelo, A. Cammarata, P. Nicolini, C. Putignano, G. Carbone, S. Stupkiewicz, J. Lengiewicz, G. Costagliola, F. Bosia, R. Guarino, N. M. Pugno, M. H. Müser, and M. Ciavarella. Modeling and simulation in tribology across scales: An overview. *Tribology International*, 125:169–199, 2018.
- [74] J. Van Doormaal and G. Raithby. Enhancements of the simple method for predicting incompressible fluid flows. *Numerical heat transfer*, 7(2):147–163, 1984.
- [75] P. Vaněk, J. Mandel, and M. Brezina. Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems. *Computing*, 56:179–196, 1996.
- [76] P. Vaněk. Acceleration of convergence of a two-level algorithm by smoothing transfer operators. *Applications of Mathematics*, 37:265–274, 1992.
- [77] P. Vaněk, J. Mandel, and M. Brezina. Algebraic multigrid on unstructured meshes. Technical report, University of Colorado at Denver, 1994.
- [78] P. Vaněk, A. Janka, and H. Guillard. Convergence of algebraic multigrid based on smoothed aggregation ii: Extension to a Petrov-Galerkin method. *INRIA Technical Report 3683*, 1999.
- [79] P. Vaněk, M. Brezina, and J. Mandel. Convergence of algebraic multigrid based on smoothed aggregation. *Numerische Mathematik*, 88:559–579, 2001.
- [80] F. Verdugo and W. A. Wall. Unified computational framework for the efficient solution of n -field coupled problems with monolithic schemes. *Computer Methods in Applied Mechanics and Engineering*, 310:335–366, 2016.

- [81] M. Wabro. Coupled algebraic multigrid methods for the Oseen problem. *Computing and Visualization in Science*, 7:141–151, 2004.
- [82] C. Wieners and B. Wohlmuth. Duality estimates and multigrid analysis for saddle point problems arising from mortar discretizations. *SIAM Journal on Scientific Computing*, 24:2163–2184, 2003.
- [83] C. Wieners and B. Wohlmuth. A general framework for multigrid methods for mortar finite elements. Technical Report 415, Universität Augsburg, Institut für Mathematik, 1999.
- [84] T. A. Wiesner, M. W. Gee, A. Prokopenko, and J. Hu. The MueLu Tutorial. Technical report, Sandia National Laboratories, 2014.
- [85] T. A. Wiesner, A. Popp, M. W. Gee, and W. A. Wall. Algebraic multigrid methods for dual mortar finite element formulations in contact mechanics. *International Journal for Numerical Methods in Engineering*, 144(4):399–430, 2018.
- [86] B. Wohlmuth. A multigrid methods for saddlepoint problems arising from mortar finite element discretizations. *ETNA. Electronic Transactions on Numerical Analysis*, 11:43–54, 2000.
- [87] B. Wohlmuth. *Discretization Techniques and Iterative Solvers Based on Domain Decomposition*, volume 17. Springer, Heidelberg, 2001.
- [88] B. I. Wohlmuth. A mortar finite element method using dual spaces for the Lagrange multiplier. *SIAM Journal on Numerical Analysis*, 38(3):989–1012, 2000.
- [89] B. I. Wohlmuth. Variationally consistent discretization schemes and numerical algorithms for contact problems. *Acta Numerica*, 20:569–734, 2011.
- [90] B. I. Wohlmuth, A. Popp, M. W. Gee, and W. A. Wall. An abstract framework for a priori estimates for contact problems in 3D with quadratic finite elements. *Computational Mechanics*, 49:735–747, 2012.
- [91] B. I. Wohlmuth and R. Krause. Monotone Multigrid Methods on Nonmatching Grids for Nonlinear Multibody Contact Problems. *SIAM Journal on Scientific Computing*, 25(1):324–347, 2003.
- [92] B. Wohlmuth and R. Krause. Multigrid methods based on the unconstrained product space for mortar finite element discretizations. *SIAM Journal on Numerical Analysis*, 39:192–213, 2001.
- [93] P. Wriggers. *Computational contact mechanics*. Springer-Verlag Berlin Heidelberg, 2006.
- [94] L. Wunderlich, A. Seitz, M. D. Alaydin, B. I. Wohlmuth, and A. Popp. Biorthogonal splines for optimal weak patch-coupling in isogeometric analysis with applications to finite deformation elasticity. *Computer Methods in Applied Mechanics and Engineering*, 346:197–215, 2019.
- [95] B. Yang and T. Laursen. A contact searching algorithm including bounding volume trees applied to finite sliding mortar formulations. *Computational Mechanics*, 41(2):189–205, 2008.
- [96] O. C. Zienkiewicz, R. L. Taylor, and D. D. Fox. *The Finite Element Method for Solid and Structural Mechanics*. Butterworth-Heinemann, Oxford, 7th edition, 2014.
- [97] W. Zulehner. A class of smoothers for saddle point problems. *Computing*, 65:227–246, 2000.
- [98] W. Zulehner. Analysis of iterative methods for saddle point problems: A unified approach. *Mathematics of Computation*, 71:479–505, 2002.