
On the Multiscale Solution of Constrained Minimization Problems

Rolf Krause

Institute for Numerical Simulation
Wegelerstraße 6
D-53115 Bonn
krause@ins.uni-bonn.de

1 Introduction

For the constrained minimization of convex or non-convex functionals on the basis of multilevel or domain decomposition methods, different strategies have been proposed within the last decades. These include nonlinear and monotone multigrid methods, see [5, 9, 12, 16, 20], multilevel optimization strategies and multilevel Trust-Region methods, see [8, 21], nonlinear domain decomposition methods [1, 6, 22, 23], multigrid methods as linear solvers in the framework of interior point based methods, see [4, 24] and multigrid methods applied in the framework of primal-dual active set strategies or semi-smooth Newton methods, see [11] for the latter. For a nonlinear multigrid method for smooth problems we refer to [10]. We remark that the references given here are far from exhaustive and refer the reader to the references cited therein.

From the multiscale point of view, two features might be employed in order to distinguish between the different methods. The first one is the way the constraints are incorporated into the multiscale hierarchy. The second one is the way the nonlinearity is intertwined with the multiscale structure.

On the one hand, in case interior point methods, active set strategies or semi-smooth Newton methods are used as solution method, domain decomposition or multilevel methods are often used as an inner linear solver within an outer smooth or non-smooth iteration process. Then, the outer iteration provides the convergence of the iterates to a minimizer whereas the inner solver is only applied to linear problems. In order to accelerate the overall iteration process, often the arising linear subproblems are solved inexactly. In this case, the choice of the linear solution method can also effect the convergence of the overall nonlinear scheme significantly, since the approximate correction given by the iterative linear solver might provide a completely different descent direction than the solution of the linear system itself. As a consequence, even if the original nonlinear constrained minimization problem is reduced to a se-

quence of linear subproblems, the nonlinearity shows also up within the linear subproblems.

On the other hand, following nonlinear domain decomposition and multi-level strategies, the nonlinear iteration process can in contrast be carried out within the subspaces provided by the considered splitting, see [1, 14, 23]. In case of a multilevel method, for example, the nonlinearity might be evaluated on all levels of the multilevel hierarchy. The resulting information gathered on the multilevel hierarchy can then be used to provide faster convergence of the nonlinear iteration process. A possible drawback of this approach is that spurious coarse grid corrections might spoil the convergence of the nonlinear method, cf. [17]. A remedy can be found in adapting the multilevel decomposition to the nonlinearities by, e.g. using solution dependent interpolation operators and bilinear forms. Although this requires at least partially reassembling of the coarse grid stiffness matrices, the additional effort is easily justified by the resulting gain in robustness and convergence speed of the multilevel method.

2 Constrained Minimization

Let H be a Hilbert space and $\emptyset \neq \mathcal{K} \subset H$ a closed and convex subset. We consider the constrained minimization Problem: find $u \in \mathcal{K}$

$$J(u) \leq J(v), \quad v \in \mathcal{K}, \quad (1)$$

where $J: H \rightarrow \mathbb{R}$ is a convex and l.s.c. functional. Under this assumptions, a minimizer exists, which is also unique if J is strictly convex, see, e.g. [7]. By introducing the characteristic functional

$$\chi_{\mathcal{K}}(v) = \begin{cases} 0, & \text{if } v \in \mathcal{K}, \\ \infty, & \text{else,} \end{cases}$$

the constraints can be translated into the non-smooth and nonlinear functional $\chi_{\mathcal{K}}$, leading to the unconstrained minimization problem: find $u \in H$

$$(J + \chi_{\mathcal{K}})(u) \leq (J + \chi_{\mathcal{K}})(v), \quad v \in H. \quad (2)$$

Since the resulting non-smooth energy $J + \chi_{\mathcal{K}}$ prevents the straight forward application of, e.g. a gradient method or Newton's method, often the functional $J + \chi_{\mathcal{K}}$ is replaced by a differentiable one, e.g. $J + \chi_{\mathcal{K}}^{\alpha}$, α a regularization parameter. This allows for applying Newton's method to the resulting first order conditions for a minimum

$$(J + \chi_{\mathcal{K}}^{\alpha})'(u^{\alpha})(v) = 0, \quad v \in H. \quad (3)$$

A different and non-smooth approach can be found by formulating the necessary conditions for a minimizer of J as variational inequality. In this case, the

energy functional J is generated by the H -elliptic bilinear form $a(\cdot, \cdot)$ and by the linear functional f on H as

$$J(u) = \frac{1}{2}a(u, u) - f(u), \quad (4)$$

the minimization problem (1) can equivalently be reformulated as the variational inequality: find $u \in \mathcal{K}$

$$a(u, v - u) \geq f(v - u), \quad v \in \mathcal{K}, \quad (5)$$

see [7]. The advantage of the latter formulation is that the non-smooth structure of the minimization problem (1) is preserved. Numerical methods based on (5) therefore can be expected to give results with higher accuracy.

After discretization of (1) by, e.g. finite elements, we obtain the finite dimensional minimization problem: find $u^L \in \mathcal{K}^L$

$$J(u^L) \leq J(v), \quad v \in \mathcal{K}^L, \quad (6)$$

where the closed and convex set $\emptyset \neq \mathcal{K}^L \subset \mathcal{S}^L$ approximates \mathcal{K} and \mathcal{S}^L is a finite dimensional subspace of H . Here, the index L serves as discretization parameter. We remark that instead of solving the nonlinear problem (6) in finite dimensions it is also possible to apply, e.g. an interior point method in the function space H directly, see [24]. The approximate computation of the resulting Newton corrections then gives rise to linear subproblems, which can be solved by linear multigrid methods. Here, we do not follow this approach but rather focus on the efficient computation of a solution to the finite dimensional constrained minimization problem (6). This solution can be obtained by either applying, e.g. a semi-smooth Newton method or a primal dual active set strategy to the necessary first order conditions, cf. (5), or by attacking the minimization problem (6) directly. Consequently, a multigrid method can either be used as a solver or preconditioner for the linearized problem, or it can serve as a nonlinear solver by itself.

3 Low Frequency Representation of Constraints

Here, as an example for (1), let us consider a contact problem in elasticity. Subject to volume and surface forces, an elastic body is pressed against a rigid foundation which cannot be penetrated, see, e.g., Figure 1. The actual zone of contact γ_C depends on the sought deformations and is unknown in advance. We identify the elastic body in its reference configuration with the (polyhedral) domain $\Omega \subset \mathbb{R}^3$ and set as solution space $H = (H^1(\Omega))^3$. The boundary $\partial\Omega$ is decomposed into three disjoint parts, Γ_D , the Dirichlet boundary with $\text{meas}_2(\Gamma_D) > 0$, Γ_N , the Neumann boundary and Γ_C , the possible contact boundary. We assume $\overline{\gamma_C} \Subset \Gamma_C$. At Γ_C , we enforce the linearized non-penetration condition $\mathbf{u} \cdot \mathbf{n} \leq g$, cf. [13], with respect to the outer

normal \mathbf{n} . Here, g is the distance in normal direction to the obstacle in the reference configuration. The normal and tangential displacements, respectively, are $u_n = \mathbf{u} \cdot \mathbf{n}$ and $\mathbf{u}_T = \mathbf{u} - u_n \cdot \mathbf{n}$. We use boldface symbols for tensor and vector quantities and the summation convention is enforced on indices running from $1, \dots, 3$. The stresses $\boldsymbol{\sigma}$ are given by Hooke's law $\sigma_{ij}(\mathbf{u}) = E_{ijml} u_{l,m}$, where Hooke's tensor $(E_{ijml})_{i,j,l,m=1}^3$, $E_{ijlm} \in L^\infty(\Omega)$, $1 \leq i, j, l, m \leq 3$ is assumed to be sufficiently smooth, symmetric and uniformly positive definite and $\boldsymbol{\epsilon}(\mathbf{u}) = \frac{1}{2}(\nabla \mathbf{u} + (\nabla \mathbf{u})^T)$ is the linearized strain tensor. The minimization problem (1) now constitutes the elastic contact problem without friction, if we define the set of admissible displacements by

$$\mathcal{K} = \{\mathbf{u} \in H \mid \mathbf{u} \cdot \mathbf{n} \leq g \text{ on } \Gamma_C\} \quad (7)$$

and choose J to be the quadratic elastic energy

$$J(\mathbf{v}) = \frac{1}{2}a(\mathbf{v}, \mathbf{v}) - f(\mathbf{v}) = \frac{1}{2} \int_{\Omega} \boldsymbol{\sigma}(\mathbf{v}) : \boldsymbol{\epsilon}(\mathbf{v}) \, dx - \int_{\Omega} \mathbf{f} \mathbf{v} \, dx, \quad (8)$$

see [13]. Here $\mathbf{f} \in \mathbf{L}^2(\Omega)$ accounts for the volume forces and surface tractions. The finite dimensional minimization problem (6) is now obtained by discretizing by finite elements. To this end, let $\mathcal{T} = (\mathcal{T}^\ell)_{\ell=0}^L$ denote a family of nested and shape regular meshes with discretization parameter h^ℓ . Here, $L > 0$ is the index of the finest level and h^ℓ is the mesh-size of \mathcal{T}^ℓ . The meshes may consist of tetrahedrons, hexahedrons, pyramids or prisms. We denote the set of all nodes of \mathcal{T}^ℓ by \mathcal{N}^ℓ and the nodes on the possible contact boundary Γ_C are $\mathcal{C}^\ell = \bar{\Gamma}_C \cap \mathcal{N}^\ell$. By $\mathcal{S}^\ell \subset \mathcal{S}^L$ we denote the spaces of first order Lagrangian finite elements on Level ℓ .

Multilevel methods for this type of problem have been considered by [3, 5, 12, 20] for scalar problems and by [16] for the system given above.

Construction of Subspaces and Coarse Level Energies

We first give the algorithmic formulation for a nonlinear and non-smooth multigrid method which has been implemented in the C++-toolbox ObsLib++, cf. [17].

Algorithm 1 (Non-smooth Multigrid Method)

- (1) Initialize \mathbf{u}_0^L . For $k = 0, \dots, k_{\max}$ do:
- (2) Compute an approximate solution \mathbf{c}^L of the problem: find $\mathbf{w}^L \in \mathcal{S}^L$, such that

$$(J + \chi_{\mathcal{K}^L})(\mathbf{u}_k^L + \mathbf{w}^L) \leq (J + \chi_{\mathcal{K}^L})(\mathbf{u}_k^L + \mathbf{w}), \quad \mathbf{v} \in \mathcal{S}^L.$$
 Set $\bar{\mathbf{u}}^L = \mathbf{u}_k^L + \mathbf{c}^L$.
- (3) For $\ell < L$ do:
 - Choose subspace $\mathcal{X}_{\bar{\mathbf{u}}^L}^\ell$, convex set $\mathcal{D}_{\bar{\mathbf{u}}^L}^\ell$, $\bar{\mathbf{u}}^l \in \mathcal{D}_{\bar{\mathbf{u}}^L}^\ell$ and functional $\mathcal{Q}_{\bar{\mathbf{u}}^L}^\ell$
 - Coarse grid correction: find $\mathbf{c}^\ell \in \mathcal{D}_{\bar{\mathbf{u}}^L}^\ell$, such that

$$\mathcal{Q}_{\bar{\mathbf{u}}^L}^\ell(\bar{\mathbf{u}}^l + \mathbf{c}^\ell) \leq \mathcal{Q}_{\bar{\mathbf{u}}^L}^\ell(\bar{\mathbf{u}}^l + \mathbf{v}), \quad \mathbf{v} \in \mathcal{D}_{\bar{\mathbf{u}}^L}^\ell.$$
- (5) Set $\mathbf{u}_k^L = P^L(\bar{\mathbf{u}}^L + \sum_{\ell < L} \mathbf{c}^\ell)$

Here, in order to allow for an adaptation of the coarse grid basis to the actual iterate, we have replaced the multilevel decomposition induced by the spaces \mathcal{S}^ℓ by the subspaces $\{\mathcal{X}_{\bar{\mathbf{u}}^L}^\ell\}_{\ell < L}$ which may depend on the smoothed iterate $\bar{\mathbf{u}}^L$ obtained after the leading fine grid smoothing (2) in Algorithm 1. The convex sets $\mathcal{D}_{\bar{\mathbf{u}}^L}^\ell$ provide a multilevel decomposition of \mathcal{K}^L , see [1, 16, 23]. By means of the mapping $P^L: \mathcal{S}^L \rightarrow \mathcal{S}^L$, the feasibility of the iterates is ensured. Examples are global damping of the coarse grid corrections or line search. In case, the coarse grid corrections are feasible by construction of $\mathcal{D}_{\bar{\mathbf{u}}^L}^\ell$, P^L will be the identity. In case of a parallel multiscale method, P^L may also serve as a non-linear synchronization which is necessary for synchronizing the different iterates obtained on the different processors. For an example, we refer to Figure 2 and Table 3. Finally, on each Level $0 \leq \ell < L$, the correction in $\mathcal{X}_{\bar{\mathbf{u}}^L}^\ell$ is computed with respect to the possibly level dependent convex functional $\mathcal{Q}_{\bar{\mathbf{u}}^L}^\ell$. This step requires either the restriction of the linear or nonlinear defect or the projection of the smoothed iterate $\bar{\mathbf{u}}^L$ onto $\mathcal{X}_{\bar{\mathbf{u}}^L}^\ell$ in order to obtain the start iterate $\bar{\mathbf{u}}^\ell$. Concerning the construction of the coarse grid models, the straight forward approach would be to set $\mathcal{Q}_{\bar{\mathbf{u}}^L}^\ell = J + \chi_{\mathcal{K}^\ell}$. However, the characteristic functional $\chi_{\mathcal{K}^\ell}$ in general cannot be represented on the coarser grids $\ell < L$. As a consequence, coarse grid corrections originating from $\mathcal{Q}_{\bar{\mathbf{u}}^L}^\ell = J + \chi_{\mathcal{K}^\ell}$ might interfere in an undesirable way with $J + \chi_{\mathcal{K}^L}$, thus spoiling the convergence or efficiency of the multilevel method, see [17]. Therefore, a suitable multiscale representation of the non-smooth nonlinearities has to be constructed, which guarantees the nonlinear convergence of our multiscale method as well as their efficiency and robustness.

For the contact problem, the leading minimization step (2) in Algorithm 1 can be realized by applying a nonlinear Gauß-Seidel method. By means of the resulting smoothed iterate $\bar{\mathbf{u}}^L$, we can define the set

$$\mathcal{A}_{\bar{\mathbf{u}}^L}^L = \{p \in \mathcal{C}^L \mid \bar{\mathbf{u}}^L(p) \cdot \mathbf{n}(p) = g(p)\} \quad (9)$$

of active nodes on level L . In order to ensure the feasibility of the coarse grid corrections, they must at least vanish at all active nodes $p \in \mathcal{A}_{\bar{\mathbf{u}}^L}^L$ in normal direction. In general, using the standard nodal multilevel basis this is not possible. We now show how suitable subspaces $\mathcal{X}_{\bar{\mathbf{u}}^L}^L$ can be obtained. Let λ_p^ℓ be the standard nodal hat function for $p \in \mathcal{N}^\ell$ and let $\{\mathbf{E}_i\}_{1 \leq i \leq 3}$ denote the Cartesian basis vectors of \mathbb{R}^d . We replace the standard nodal basis functions $\boldsymbol{\lambda}_p^\ell = (\lambda_p^\ell \cdot \mathbf{E}_1, \dots, \lambda_p^\ell \cdot \mathbf{E}_d)^T$ of \mathcal{S}^ℓ for $p \in \mathcal{C}^\ell$ by

$$\{\lambda_p^\ell \cdot \mathbf{e}_1(p), \dots, \lambda_p^\ell \cdot \mathbf{e}_d(p)\}, \quad (10)$$

where $\{\mathbf{e}_i\}_{1 \leq i \leq 3}$ is an orthonormal basis associated with $p \in \mathcal{C}^L$ and with $\mathbf{e}_1(p) = \mathbf{n}(p)$. As a consequence of (10), the first component of the displacements at the nodes $p \in \mathcal{C}^\ell$ is always the displacement in normal direction. Let now $I_\ell^{\ell+1}: \mathcal{S}^\ell \rightarrow \mathcal{S}^{\ell+1}$ denote the interpolation operator with respect to the local transformation (10). The algebraic representation $\mathbf{I}_\ell^{\ell+1} = (\mathbf{i}_{pq})_{p \in \mathcal{N}^{\ell+1}, q \in \mathcal{N}^\ell}$ of $I_\ell^{\ell+1}$ is a rectangular matrix with the 3×3 blocks $\mathbf{i}_{pq} \in \mathbb{R}^{3 \times 3}$. We note

that due to (10) for $p \in \mathcal{C}^L$ the blocks \mathbf{i}_{pq} in general are not diagonal, if the normals differ along Γ_C . Now, we introduce the sets $A^\ell \subset \mathcal{C}^\ell \times \{1, \dots, d\}$ of active degrees of freedom for each Level $\ell \leq L$. On Level L , we set $A^L = \{(p, 1) \mid p \in \mathcal{A}_{\bar{\mathbf{u}}^L}^L\}$. On the coarser levels, the spaces $\mathcal{X}_{\bar{\mathbf{u}}^L}^\ell$ can be defined by removing the degrees of freedom in A^ℓ from the nodal basis of \mathcal{S}^ℓ . Possible choices for the multiscale representation of the set A^L include now ($1 \leq i, j \leq 3$)

1. $A^\ell = \{(q, 1) \mid q \in \mathcal{C}^\ell \text{ and } \bar{\mathbf{u}}^L(q) \cdot \mathbf{n}(q) = g(q)\}$.
2. Set recursively for $\ell < L$: $A^\ell = \{(q, 1) \mid \exists (p, 1) \in A^{\ell+1} : (\mathbf{i}_\ell^{\ell+1})_{pq}^{11} \neq 0\}$.
3. Set recursively for $\ell < L$: $A^{\ell-1} = \{(q, j) \mid \exists (p, i) \in A^{\ell+1} : (\mathbf{i}_\ell^{\ell+1})_{pq}^{ij} \neq 0\}$.

In addition to 1–3, we employ truncated basis functions $\{\boldsymbol{\mu}_q^\ell\}_{q \in \mathcal{N}^\ell}$, see [16]. For $\ell < L$, they can be defined by

$$\boldsymbol{\mu}_q^\ell = \boldsymbol{\lambda}_q^\ell - \sum_{p \in \text{int supp } \boldsymbol{\lambda}_q^\ell \cap \mathcal{A}_{\bar{\mathbf{u}}^L}^L} \omega_{qp} \boldsymbol{\lambda}_p^L \cdot \mathbf{n}(p),$$

where the weights ω_{qp} are such that for all active nodes $p \in \mathcal{A}_{\bar{\mathbf{u}}^L}^L$ it holds for $\ell \leq L$ that $\boldsymbol{\mu}_q^\ell(p) \cdot \mathbf{n}(p) = 0$. Thus, the resulting multilevel basis provides a multiscale representation of the active constraints $\mathcal{A}_{\bar{\mathbf{u}}^L}^L$ on all coarser levels $\ell < L$. We remark that the search directions $\boldsymbol{\mu}_q^\ell$ are never explicitly computed, since the corresponding stiffness matrix can be obtained recursively by modifying the interpolation operator and using local reassembling.

Global Convergence

Despite the coarse grid spaces, we also have to choose the coarse grid energies $\mathcal{Q}_{\bar{\mathbf{u}}^L}^\ell$ and the convex sets of feasible corrections $\mathcal{D}_{\bar{\mathbf{u}}^L}^\ell$. Using the multigrid method as nonlinear solver by itself, following the idea of monotone multigrid methods, see [14], global convergence is achieved by guaranteeing that during the multigrid iteration process the convex functional $J + \chi_{\mathcal{K}^L}$ always decreases. The minimizer of (6) is sought by successive minimization in direction of all basis functions of the subspaces $\mathcal{X}_{\bar{\mathbf{u}}^L}^\ell$ originating from the truncated basis. In order to ensure the feasibility of the coarse grid corrections, inner approximations $\mathcal{D}_{\bar{\mathbf{u}}^L}^\ell$ of the set \mathcal{K}^L are constructed for $\ell < L$. Choosing $\mathcal{Q}_{\bar{\mathbf{u}}^L}^\ell = J$, then the monotonicity of the iteration guarantees the global convergence of the resulting monotone multigrid method for contact problems, see [16]. In the following, we denote this method by M-MG.

For unconstrained convex minimization problems, [21] has shown the convergence of a multilevel optimization method if the coarse grid problems are solved “accurately enough”. Then, it can be guaranteed that a descent direction is provided by the coarse grid corrections. In [8], Trust-Region strategies are intertwined with multilevel optimization methods. In all cases, the sufficient decrease of the functional J is used to ensure the convergence of the multilevel method. Let us remark that the convergence proof in [10] for a smooth nonlinear multigrid method is also based on a minimization property.

Influence of the Multilevel Splitting

As an alternative to using the multigrid method as nonlinear solution method, it can also be applied as linear solver or as a preconditioner within a nonlinear strategy as, e.g. a primal dual active set strategy. We therefore consider the influence of the multilevel decompositions $\mathcal{X}_{\bar{\mathbf{u}}^L}^\ell$ given above in the context of a monotone multigrid method as well as in the context of a primal dual active set strategy. Let us define the active set of \mathbf{u}_k^L by $\mathcal{A}_k^L = \{p \in \mathcal{C}^L \mid s_n(p) + \rho(\mathbf{u}_k^L(p) \cdot \mathbf{n}(p) - g(p)) > 0\}$, where $s_n(p) = a(\mathbf{u}_k^L, \lambda_p^L \cdot \mathbf{n}(p)) - f(\lambda_p^L \cdot \mathbf{n}(p))$ are the discrete normal stresses and $\rho > 0$ is an algorithmic parameter. An inexact multigrid based primal-dual active set strategy can be obtained from Algorithm 1 by replacing $\chi_{\mathcal{K}^L}$ in step (2) by the characteristic functional of the set $\mathcal{X}_{\bar{\mathbf{u}}_{k-1}^L}^L = \{\mathbf{v} \in \mathcal{S}^L \mid \mathbf{v}(p) \cdot \mathbf{n}(p) = 0, p \in \mathcal{A}_{k-1}^L\}$ and by using the linear coarse grid corrections induced by setting $\mathcal{Q}_{\bar{\mathbf{u}}^L}^\ell = J$ and $\mathcal{X}_{\bar{\mathbf{u}}^L}^\ell = \mathcal{D}_{\bar{\mathbf{u}}^L}^\ell$. In each step k , the steps (2) and (3) in Algorithm 1 amount to the inexact solution of a linear sub-problem of the form: find $\mathbf{c} \in \mathcal{X}_{\bar{\mathbf{u}}_{k-1}^L}^L$, such that

$$a(\mathbf{c}, \mathbf{v}) = (\mathbf{f}, \mathbf{v}) - a(\mathbf{w}, \mathbf{v}), \quad \mathbf{v} \in \mathcal{X}_{\bar{\mathbf{u}}_{k-1}^L}^L, \quad (11)$$

where $\mathbf{w} \in \mathcal{S}^L$ and $\mathbf{w}(p) \cdot \mathbf{n}(p) = g(p)$ for $p \in \mathcal{A}_{k-1}^L$.

Primal-dual active set strategies are known to converge superlinearly, if the initial iterate \mathbf{u}_0^L is sufficiently close to the solution. If, in addition the stiffness matrix is an M -matrix and the linear systems (11) are solved exactly, also global convergence can be shown, see [11]. Global convergence can also be obtained using Trust-Region Strategies. As a matter of fact, in case the linear sub-problems (11) are solved only inexactly, the choice of the employed multilevel decomposition strongly influences the convergence of the overall nonlinear strategy. We illustrate this for a Hertzian contact problem in 3d. Here, a sphere is pressed in z -direction against the rigid plane $\{z = 0\}$. The material parameters are $E = 10^5$ and $\nu = 0.3$ and we have $L = 5$ levels of adaptive refinement and 659.409 degrees of freedom on Level 5. In Table 1, the resulting numbers of $\mathcal{W}(3, 3)$ -cycles are shown for this multigrid based active set strategy. We use the coarse grid spaces induced by the active sets \mathcal{A}^ℓ given on the previous page and the truncated basis as well as the globally convergent monotone multigrid method M-MG with the truncated basis. The iteration is stopped, if $\|u_{k+1}^L - u_k^L\|_a / \|u_k^L - u_{k-1}^L\|_a \leq 10^{-12}$, $\|u\|_a = a(u, u)^{1/2}$. The initial iterate \mathbf{u}_0^L is given by random values in the interval $[-0.2, -0.1]$. For the definition of the set \mathcal{K}^L we consider two different cases. Firstly, the case of constant normal direction at Γ_C , i.e. we take as normal direction a $\mathbf{n}(p) = (0, 0, -1)^T$ for all $p \in \mathcal{C}^L$ (“equal normals”), and secondly, $\mathbf{n}(p)$ the outer normal at $p \in \mathcal{C}^L$ (“outer normals”). As can be seen from Table 1, for the case of the outer normals, the constraints at the interface are locally not decoupled and spurious corrections from the coarser grids can spoil the convergence. We emphasize that the truncated basis functions showed to provide the best nonlinear search directions with respect to both, efficiency and

robustness. As a by product, they can also be used for the multilevel representation of Dirichlet values. The slightly higher iteration numbers for the monotone multigrid method show the influence of the multilevel decomposition of the set \mathcal{K}^L , since for M-MG the feasibility of the coarse grid corrections is enforced by the construction of the sets $\mathcal{D}_{\bar{\mathbf{u}}^L}^\ell$.

Summing up, regardless of using a multigrid method as inexact solver within a non-smooth solution method or as nonlinear solver by itself, the multilevel decomposition has to be adapted to the active constraints in order to provide a fast and robust method.

4 Low Frequency Representation of Nonlinearities

Approximation by Quadratic Functionals

We now consider the case where the functional to be minimized is non-quadratic and non-differentiable. As example we choose elastic contact with Tresca friction. The corresponding minimization problem, after discretization, is given by: find $u \in \mathcal{K}^L$

$$(J + j_{s_n}^L)(\mathbf{u}^L) \leq (J + j_{s_n}^L)(\mathbf{v}), \quad \mathbf{v} \in \mathcal{K}^L. \quad (12)$$

Here, J is the elastic energy (8) and the friction functional $j_{s_n}^L$ is given by

$$j_{s_n}^L(\mathbf{v}) = \sum_{p \in \mathcal{C}^L} \mathcal{F} |s_{p,n}| |\mathbf{v}_{p,T}|, \quad (13)$$

$|\cdot|$ the Euclidean norm in \mathbb{R}^2 and $s_n = (s_{p,n})_{p \in \mathcal{C}^L}$ are the prescribed scaled boundary stresses, and $\mathcal{F} > 0$ is the coefficient of friction. We write $u_{p,n} = u_n(p)$ and $\mathbf{u}_{p,T} = \mathbf{u}_T(p)$. Tresca's friction law induces a non-smooth relationship between tangential displacements and tangential stresses, cf. [13]. Taking into account the efficiency and robustness of SQP-methods, the construction of the coarse level functionals $\mathcal{Q}_{\bar{\mathbf{u}}^L}^\ell$ might be based on a quadratic approximation of $J + j_{s_n}^L$. However, since $j_{s_n}^L$ is non-differentiable, this turns out to be a non-trivial task. We therefore proceed as follows, see [15, 17]. After $\bar{\mathbf{u}}^L$ in Algorithm 1 has been obtained, the subsequent coarse grid corrections are restricted to a neighborhood $\mathcal{D}_{\bar{\mathbf{u}}^L}^L$ of $\bar{\mathbf{u}}^L$ where $\bar{\mathbf{u}}_T^L \neq 0$ and therefore the energy $J + j_{s_n}^L$ is smooth. In contrast to the Trust-Region techniques given in [8], here the neighborhood $\mathcal{D}_{\bar{\mathbf{u}}^L}^L$ is locally defined by box constraints. This allows us to construct the coarse grid models $\mathcal{Q}_{\bar{\mathbf{u}}^L}^\ell$ on the basis of the quadratic approximation

$$\begin{aligned} \mathcal{Q}_{\bar{\mathbf{u}}^L}(\mathbf{v}) = & \frac{1}{2} (a(\mathbf{v}, \mathbf{v}) + j_{\bar{\mathbf{u}}^L}''(\bar{\mathbf{u}}^L)(\mathbf{v}, \mathbf{v})) \\ & - (f(\mathbf{v}) - j_{\bar{\mathbf{u}}^L}'(\bar{\mathbf{u}}^L)(\mathbf{v}) + j_{\bar{\mathbf{u}}^L}''(\bar{\mathbf{u}}^L)(\bar{\mathbf{u}}^L, \mathbf{v})) \end{aligned} \quad (14)$$

of $J + j_{s_n}^L$ on $\mathcal{D}_{\bar{\mathbf{u}}^L}^L$, where we have set

$$j_{\bar{\mathbf{u}}^L}(\mathbf{v}) = \sum_{p \in \mathcal{B}^L} \mathcal{F}|s_{p,n}||\mathbf{v}_{p,T}|, \quad \mathbf{v} \in \mathcal{S}^L. \quad (15)$$

Here, $\mathcal{B}^L \subset \mathcal{C}^L$ denotes the set of all sliding nodes w.r.t. $\bar{\mathbf{u}}^L$, i.e. all nodes $p \in \mathcal{C}^L$ with $\bar{\mathbf{u}}_{p,T}^L \neq 0$. As coarse grid model we use $\mathcal{Q}_{\bar{\mathbf{u}}^L}^\ell = \mathcal{Q}_{\bar{\mathbf{u}}^L}$ from (14). For the construction of the subspaces $\mathcal{X}_{\bar{\mathbf{u}}^L}^\ell$ we again use the truncated basis functions. At all sticky nodes, i.e. $p \in \mathcal{C}^L$ with $\bar{\mathbf{u}}_{p,T} = 0$, truncation is employed in all directions, such that $\boldsymbol{\mu}_q^\ell \cdot \mathbf{e}_i(p) = 0$ for $1 \leq i \leq 3$. Now setting $\mathcal{D}_{\bar{\mathbf{u}}^L}^\ell = \{\mathbf{v} \in \mathcal{X}_{\bar{\mathbf{u}}^L}^\ell \mid (\bar{\mathbf{u}}_{p,T}^L, \mathbf{v}_{p,T}) \geq 0, \quad p \in \mathcal{C}^L\}$ the global convergence of the resulting multigrid method can be shown, see [18]. Again, the convergence proof relies on the successive minimization of the frictional energy, but now the coarse grid functionals $\mathcal{Q}_{\bar{\mathbf{u}}^L}^\ell$ are different from the fine grid functional.

Since the sliding directions $\bar{\mathbf{u}}_{p,T}^L$ differ along Γ_C , we again equilibrate the constraints by applying a basis transformation as in (10), but now only in the tangential space $\text{span}\{\mathbf{e}_2(p), \mathbf{e}_3(p)\}$. This allows for a better representation of the sets $\mathcal{D}_{\bar{\mathbf{u}}^L}^\ell$ in $\mathcal{X}_{\bar{\mathbf{u}}^L}^\ell$.

As an example, we consider an elastic block pressed onto a rigid plane. A coarse triangulation of a cube with eight hexahedrons is refined adaptively until 190,888 elements are obtained on Level $L = 6$. In Figure 1, the resulting number of iterates of M-MG on Level 6 are shown if this additional basis transformation is applied (lower line) or not (upper line), again for the stopping criterion given above. As can be seen, adapting the spaces $\mathcal{X}_{\bar{\mathbf{u}}^L}^\ell$ to the nonlinearity j_{s_n} improves the robustness and efficiency of the method. For details, we refer to [18]. As an additional example, Figure 1 shows a torus in contact with a rigid foundation and the tangential stresses for $\mathcal{F} = 0.3$ at the contact interface. As can be seen, the sharp interface between sliding and

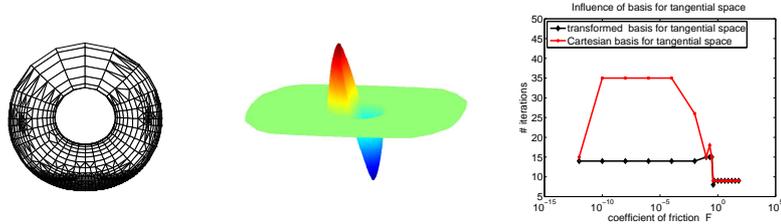


Fig. 1. Torus in contact with a rigid foundation *Left*: Deformed geometry. *Middle*: First component of the tangential stresses. *Right*: Block on plane: Robustness w.r.t to the coefficient of friction. Influence of coarse grid spaces $\mathcal{X}_{\bar{\mathbf{u}}^L}^\ell$

sticky nodes in the tangential stresses is perfectly resolved by our non-smooth minimization approach.

Semi-smooth Approach

The regularized problem (3) motivates another possibility to ensure that the coarse grid corrections provide a descent direction for the minimization problem (6). To enforce the pointwise given constraints $u_n \leq g$ for our contact problem, one could use the classical logarithmic barrier function to obtain the smooth energy functional

$$(J + \chi_{\mathcal{K}_\mu^L})(\mathbf{u}) = J(\mathbf{u}) - \mu \sum_{p \in \mathcal{C}^L} \ln(g(p) + \varepsilon - u_n(p)), \quad (16)$$

$\mu > 0, \varepsilon \geq 0$ parameters. The disadvantage of this formulation is that ill-conditioning of the resulting Hessian may occur. Moreover, due to the regularization, the solution of the minimization problem (6) is only obtained in the limit $\mu \rightarrow 0$ and therefore some accuracy is lost. However, in the context of our nonlinear multigrid method, this approach can be used to construct the coarse grid energies $Q_{\mathbf{u}_L}^\ell$ for $\ell < L$ on the basis of the formulation (16). To this end, on Level L , the leading minimization step (2) in Algorithm 1 is done by means of a non-smooth method as, e.g. a nonlinear Gauß-Seidel method. Then, the spaces $\mathcal{X}_{\mathbf{u}_L}^\ell$ are constructed using the truncated basis functions w.r.t (9). As coarse grid energies, we use the quadratic approximation (14) for the smooth energy (16). By means of this semi-smooth method, the coarse

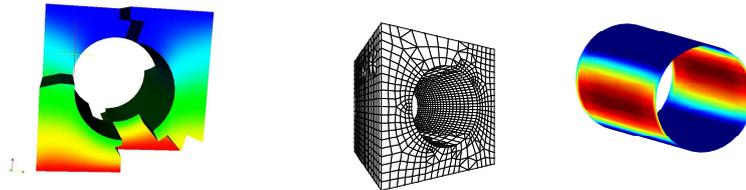


Fig. 2. Cube with hole in contact with a rigid cylinder inside *Left:* Parallel decomposition with 16 subdomains. *Middle:* Grid on Level $\ell = 1$. *Right:* Normal stresses.

grid corrections are encouraged to stay within the feasible set, which gives rise to a “better” descent direction. In addition, the regularization does not influence the accuracy of the results, since it is only applied on the coarser grids. In Table 2, the resulting number of iterates for the contact problem from Figure 2 for $\mu = 10^{-4}$ and $\varepsilon = 10^{-7}$ are shown. Here, we compare the monotone multigrid method using the truncated basis functions with (M-MG) and without (NL-MG) enforcing the feasibility of the coarse grid corrections with the combined approach (C-MG). In order to stress the nonlinear iteration process, the components of the initial iterate \mathbf{u}_0^L are chosen randomly in $[-100, 100]$. For the coarse grid problems, the respective method as algebraic multigrid method is used. As can be seen, the inner approximation of

Table 1. Iteration numbers illustrating the resolution of constraints for the different multilevel splittings given in Section 3

Level 5	Splitting 1	Splitting 2	Splitting 3	trc. Basis	M-MG
equal normals	15	34	34	15	17
outer normals	no conv.	no conv.	> 100	15	25

Table 2. Non-smooth and combined non-smooth and regularization approach for randomly chosen initial iterate

Level	# it. M-MG	# it. NL-MG	# it C-MG	# dof	# contacts
1	34	34	34	5,016	228
2	44	24	23	22,326	854
3	70	140	35	142,146	3,226

Table 3. Iteration numbers illustrating the scalability for the parallelized non-smooth multigrid method M-MG. Nested iteration, example from Figure 2.

Level ℓ	#it. 1 Processor	#it. 2 Proc.	#it. 4 Proc.	#it. 8 Proc.	#it. 16 Proc.
2	16	17	17	17	16
3	17	18	18	18	18

the feasible set in M-MG requires additional iterations to identify the contact boundary. Setting $\mathcal{D}_{\mathbf{u}^L}^\ell = \mathcal{X}_{\mathbf{u}^L}^\ell$, as in NL-MG, in contrast to the previous section, here does not improve the convergence speed. However, the combined approach C-MG provides a good multilevel search strategy for bad initial iterates. We remark that in case of a better start iterate, all three strategies show similar iteration numbers. Our numerical experiments have been carried out in the framework of the finite element toolbox [2] and the C++-toolbox ObsLib++, see [17]. The hexahedral grids have been created using the Cubit grid generator, see [19].

Acknowledgement. We thank C. Groß for his help in implementing the UG/Cubit Interface and the anonymous referees for their remarks.

References

- [1] L. Badea and J. Wang. An additive Schwarz method for variational inequalities. *Math. Comp.*, 69(232):1341–1354, 1999.
- [2] P. Bastian, K. Birken, K. Johannsen, S. Lang, N. Neuß, H. Rentz-Reichert, and C. Wieners. UG – a flexible software toolbox for solving partial differential equations. *Comput. Vis. Sci.*, 1:27–40, 1997.
- [3] V. Belsky. A multi-grid method for variational inequalities in contact problems. *Computing*, 51:293–311, 1993.
- [4] M. Benzi, E. Haber, and L. R. Hansson. Multilevel algorithms for large scale interior point methods in bound constraint optimization. Technical report, Emory University, Atlanta, 2005.

- [5] A. Brandt and C.W. Cryer. Multigrid algorithms for the solution of linear complementarity problems arising from free boundary problems. *SIAM J. Sci. Stat. Comput.*, 4:655–684, 1983.
- [6] Z. Dostál, F.A.M. Gomes Neto, and S.A. Santos. Duality based domain decomposition with natural coarse space for variational inequalities. *J. Comput. Appl. Math.*, 126(1-2):397–415, 2000.
- [7] R. Glowinski. *Numerical Methods for Nonlinear Variational Problems*. Series in Computational Physics. Springer, New York, 1984.
- [8] S. Gratton, A. Sartenaer, and P. Toint. Recursive trust-region methods for multiscale nonlinear optimization. *CERFACS*, 06(32), 05 2006. 06.
- [9] W. Hackbusch and H.D. Mittelmann. On multi-grid methods for variational inequalities. *Numer. Math.*, 42:65–76, 1983.
- [10] W. Hackbusch and A. Reusken. Analysis of a damped nonlinear multi-level method. *Numer. Math.*, 55:225–246, 1989.
- [11] M. Hintermüller, K. Ito, and K. Kunisch. The primal dual active set strategy as a semismooth Newton method. *SIAM J. Optim.*, 13(3):865–888, 2003.
- [12] R.H.W. Hoppe. Multigrid algorithms for variational inequalities. *SIAM J. Numer. Anal.*, 24:1046–1065, 1987.
- [13] N. Kikuchi and J.T. Oden. *Contact Problems in elasticity*. SIAM, 1988.
- [14] R. Kornhuber. *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems*. Teubner-Verlag, Stuttgart, 1997.
- [15] R. Kornhuber. On constrained Newton linearization and multigrid for variational inequalities. *Numer. Math.*, 91:699–721, 2002.
- [16] R. Kornhuber and R. Krause. Adaptive multigrid methods for Signorini’s problem in linear elasticity. *Comput. Vis. Sci.*, 4:699–721, 2001.
- [17] R. Krause. From inexact active set strategies to nonlinear multigrid methods. In P. Wriggers and U. Nackenhorst, editors, *Analysis and Simulation of Contact Problems*, volume 27. Springer, 2006.
- [18] R. Krause. A non-smooth multiscale method for solving frictional two-body contact problems in 2d and 3d with multigrid efficiency. Technical Report INS-Preprint No. 604, INS, University of Bonn, 2006.
- [19] Sandia National Laboratories. Cubit 9.1 mesh generation toolkit, 2004.
- [20] J. Mandel. A multi-level iterative method for symmetric, positive definite linear complementarity problems. *Appl. Math. Optim.*, 11:77–95, 1984.
- [21] S. G. Nash. A multigrid approach to discretized optimization problems. *Optim. Methods Softw.*, 14:99–116, 2000.
- [22] X. C. Tai. Rate of convergence for some constraint decomposition methods for nonlinear variational inequalities. *Numer. Math.*, 93:755–786, 2003.
- [23] X.-C. Tai and J. Xu. Global and uniform convergence of subspace correction methods for some convex optim. problems. *Math. Comp.*, 71:105–124, 2002.
- [24] M. Weiser. Interior point methods in function space. Technical report, Zuse Institute Berlin (ZIB), Berlin, 2003.