A Recursive Trust-Region Method for Non-Convex Constrained Minimization

Christian Groß[1] and Rolf Krause[1]

Institute for Numerical Simulation, University of Bonn.
{gross,krause}@ins.uni-bonn.de

1 Introduction

The mathematical modelling of mechanical or biomechanical problems involving large deformations or biological materials often leads to highly nonlinear and constrained minimization problems. For instance, the simulation of soft-tissues, as the deformation of skin, gives rise to a highly non-linear PDE with constraints, which constitutes the first order condition for a minimizer of the corresponding non-linear energy functional. Besides of the pure deformation of the tissue, bones and muscles have a restricting effect on the deformation of the considered material, leading to additional constraints. Although PDEs are usually formulated in the context of Sobolev spaces, their numerical solution is carried out using discretizations as, e.g., finite elements. Thus, in the present work we consider the following finite dimensional constrained minimization problem:

$$u \in \mathcal{B} : J(u) = \min! \tag{M}$$

where $\mathcal{B} = \{v \in \mathbb{R}^n \mid \underline{\varphi} \leq v \leq \overline{\varphi}\}$ and $\underline{\varphi} < \overline{\varphi} \in \mathbb{R}^n$ and the possibly nonconvex, but differentiable, objective function $J : \mathbb{R}^n \to \mathbb{R}$. Here, the occurring inequalities are to be understood pointwise. In the context of discretized PDEs, n corresponds to the dimension of the finite element space and may therefore be very large.

The design of robust and efficient solution methods for problems like (M) is a demanding task. Indeed, globalization strategies, such as trust-region methods (cf., [1, 10]) or line-search algorithms, succeed in computing local minimizers, but are based on the paradigm of Newton's method. This means that a sequence of iterates is computed by solving linear, but potentially large, systems of equations. The drawback is, that due to the utilization of a globalization strategy the computed corrections generally need to be damped. Hence, for two reasons the convergence of such an approach may slow down: often the linear systems of equations are ill-conditioned and therefore iterative linear solvers tend to converge slowly. Moreover, even if the linear system can be solved efficiently, for instance by employing a Krylov-space method in combination with a good preconditioner, globalization strategies tend to reduce the step-size depending on the non-linearity of the objective function.

Therefore, solution strategies which are just based on Newton's method can remain inefficient. In the context of quadratic minimization problems, linear multigrid methods have turned out to be highly efficient since these algorithms are able to resolve also the low frequency contributions of the solution. Similarly, nonlinear multigrids (cf., [4, 6, 8]) aim at a better resolution of the low-frequency contributions of non-linear problems.

Therefore, [3] introduced a class of non-linear multilevel algorithms, called RMTR_∞ (Recursive Multilevel Trust-Region method), to solve problems of the class (M). In the present work, we will introduce a V-cycle variant of the RMTR algorithm presented in [5, 6]. On each level of a given multilevel hierarchy, this algorithm employs a trust-region strategy to solve a constrained nonlinear minimization problem, which arises from level dependent representations of J, $\underline{\varphi}$, $\overline{\varphi}$. An important new feature of the RMTR_∞ algorithm is the L^2-projection of iterates to coarser levels to generate good initial iterates - in contrast to employing the restriction operator to transfer iterates to a coarse level. In fact, the new operator yields significantly better rates of convergence of the RMTR algorithm (for a complete discussion see [6]). To prove first-order convergence, we will state less restrictive assumptions on the smoothness of J than used by [3]. Moreover, we illustrate the efficiency of the RMTR_∞ - algorithm by means of an example from the field of non-linear elasticity in 3D.

2 The Multilevel Setting

The key concept of the RMTR_∞ algorithm, which we will present in Section 3, is to minimize on different levels arbitrary non-convex functions H_k approximating the fine level objective function J. The minimization is carried out employing a trust-region strategy which ensures convergence. Corrections computed on coarser levels will be summed up and interpolated which provide possible corrections on the fine level.

In particular, on each level, m_1 pre-smoothing and m_2 post-smooting trust-region steps are computed yielding trust-region corrections. In between, a recursion is called yielding a coarse level correction which is the interpolated difference between first and last iterate on the next coarser level.

Therefore, we assume that a decomposition of the $\mathbb{R}^n$ into a sequence of nested subspaces is given, such as $\mathbb{R}^n = \mathbb{R}^{n_j} \supsetneq \mathbb{R}^{n_{j-1}} \supsetneq \cdots \supsetneq \mathbb{R}^{n_0}$. The spaces are connected to each other by full-rank linear interpolation, restriction and projection operators, i.e., $I_k : \mathbb{R}^{n_k} \to \mathbb{R}^{n_{k+1}}$, $R_k : \mathbb{R}^{n_k} \to \mathbb{R}^{n_{k-1}}$ and $P_k : \mathbb{R}^{n_k} \to \mathbb{R}^{n_{k-1}}$. Given these operators and a current fine level iterate, $u_{k+1} \in \mathbb{R}^{n_{k+1}}$, the nonlinear coarse level model (cf., [4, 6, 8]) is defined as

$$H_k(u_k) = J_k(u_k) + \langle \delta g_k, u_k - P_{k+1} u_{k+1} \rangle \tag{1}$$

where we assume that a fixed sequence of nonlinear functions $(J_k)_k$ is given, representing J on the coarser levels. Here, the residual $\delta g_k \in \mathbb{R}^{n_k}$ is given by

$$\delta g_k = \begin{cases} R_{k+1}\nabla H_{k+1}(u_{k+1}) - \nabla J_k(P_{k+1}u_{k+1}) & \text{if } j > k \geq 0 \\ 0 & \text{if } k = j \end{cases}$$

In the context of constrained minimization, the fine level obstacles $\underline{\varphi}, \overline{\varphi}$ have also to be represented on coarser levels. In our implementation we employ the approach introduced in [2]. Due to the definition of the coarse level obstacles, this ensures that the projection of a fine level iterate and each resulting coarse level correction is admissible.

3 Recursive Trust-Region Methods

The reliable minimization of nonconvex functions H_k depends crucially on the control of the "quality" of the iterates. Line-search algorithms, for instance, scale the length of Newton corrections in order to force convergence to first-order critical points. Similarly, in trust-region algorithms corrections are the solutions of constrained quadratic minimization problems. For a given iterate $u_{k,i} \in \mathbb{R}^{n_k}$, where i denotes the current iteration on level k, a correction $s_{k,i}$ is computed as an approximate solution of

$$\begin{aligned} s_{k,i} \in \mathbb{R}^{n_k} : \psi_{k,i}(s_{k,i}) = \min! \\ \text{w.r.t. } \|s_{k,i}\|_\infty \leq \Delta_{k,i} \text{ and } \underline{\varphi}_k \leq u_{k,i} + s_{k,i} \leq \overline{\varphi}_k \end{aligned} \tag{2}$$

Here, $\psi_{k,i}(s) = \langle \nabla H_k(u_{k,i}), s\rangle + \frac{1}{2}\langle B_{k,i}s, s\rangle$ denotes the trust-region model with $B_{k,i}$, a symmetric matrix, possibly approximating the Hessian $\nabla^2 H_k(u_{k,i})$ (if it exists) and $\Delta_{k,i}$ is the trust-region radius.

On the coarse level, the reduction of H_{k-1} starting at the initial iterate $u_{k-1,0} = P_k u_{k,m_1}$ yields a final coarse level iterate u_{k-1}. Therefore, the recursively computed correction is $s_{k,m_1} = I_{k-1}(u_{k-1} - P_k u_{k,m_1})$.

Trust-Region Algorithm, **Input**: $u_{k,0}, \Delta_{k,0}, \underline{\varphi}_k, \overline{\varphi}_k, H_k, m$

```
do m times {
    compute s_{k,i} as an approximate solution of (2)
    if (ρ_{k,i}(s_{k,i}) ≥ η_1)
        u_{k,i+1} = u_{k,i} + s_{k,i}
    otherwise
        u_{k,i+1} = u_{k,i}
    compute a new Δ_{k,i+1}
}
```

Algorithm 6: Trust-Region Algorithm

To ensure convergence, corrections are only added to the current iterate, if the *contraction rate* $\rho_{k,i}$ is sufficiently large. The contraction rate compares $H_k(u_{k,i}) - H_k(u_{k,i} + s_{k,i})$ to the reduction predicted by the underlying quadratic model $\psi_{k,i}$. The value of $\psi_{k,i}(s_{k,i})$ prognoses the reduction induced by corrections computed by means of (2). The underlying model for recursively computed corrections s_{k,m_1} is the coarse level objective function H_{k-1}. Thus, we define

$$\rho_{k,i}(s_{k,i}) = \begin{cases} \frac{H_k(u_{k,i}) - H_k(u_{k,i}+s_{k,i})}{-\psi_{k,i}(s_{k,i})} & \text{if } s_{k,i} \text{ computed by (2)} \\ \frac{H_k(u_{k,i}) - H_k(u_{k,i}+I_{k-1}s_{k-1})}{H_{k-1}(P_k u_{k,i}) - H_{k-1}(P_k u_{k,i}+s_{k-1})} & \text{otherwise} \end{cases}$$

Now, a correction is then added to the current iterate if $\rho_{k,i}(s_{k,i}) \geq \eta_1$ where $\eta_1 > 0$. In this case, the next trust-region radius $\Delta_{k,i+1}$ will be chosen larger than the current one, i.e., $\gamma_3 \Delta_{k,i} \geq \Delta_{k,i+1} \geq \gamma_2 \Delta_{k,i} > \Delta_{k,i}$ with $\gamma_3 \geq \gamma_2 > 1$. Otherwise, if $\rho_{k,i}(s_{k,i}) < \eta_1$, the correction will be discarded and $\Delta_{k,i+1}$ chosen smaller than $\Delta_{k,i}$, i.e., $\Delta_{k,i} > \Delta_{k,i+1} \geq \gamma_1 \Delta_{k,i}$, with $0 < \gamma_1 < 1$.

These four steps, computing $s_{k,i}$ by means of (2), computing $\rho_{k,i}$, applying $s_{k,i}$ and the update of the trust-region radius are summarized in Algorithm 6.

Algorithm 7 on Page 141, introduces the RMTR_∞ algorithm, which is a V-cycle algorithm with an embedded trust-region solver.

3.1 Convergence to First-Order Critical Points

To show convergence of the RMTR_∞ algorithm to first-order critical points, we state the following assumptions on H_k, cf. [1].

(A1) For a given initial fine level iterate $u_{j,0} \in \mathbb{R}^{n_j}$, we assume that the level set $\mathcal{L}_j = \{u \in \mathbb{R}^{n_j} \mid \underline{\varphi}_j \leq u \leq \overline{\varphi}_j \text{ and } J_j(u) \leq J_j(u_{j,0})\}$ is compact. Moreover, for all initial coarse level iterates $u_{k,0} \in \mathbb{R}^{n_k}$, we assume that the level sets $\mathcal{L}_k = \{u \in \mathbb{R}^{n_k} \mid \underline{\varphi}_k \leq u \leq \overline{\varphi}_k \text{ and } H_k(u) \leq H_k(u_{k,0})\}$ are compact.

(A2) For all levels $k \in \{0, \dots, j\}$, we assume that H_k is continuously differentiable on $\mathcal{L}_k$. Moreover, we assume that there exists a constant $c_g > 0$ such that for all iterates $u_{k,i} \in \mathcal{L}_k$ holds $\|\nabla H_k(u_{k,i})\|_2 \leq c_g$.

(A3) For all levels $k \in \{0, \dots, j\}$, there exists a constant $c_B > 0$ such that for all iterates $u_{k,i} \in \mathcal{L}_k$ and for all symmetric matrices $B_{k,i}$ employed, the inequality $\|B_{k,i}\|_2 \leq c_B$ is satisfied.

Moreover, computations on level $k-1$ are carried out only if

$$\begin{aligned} \|R_k D_{k,m_1} g_{k,m_1}\|_2 &\geq \kappa_g \|D_{k,m_1} g_{k,m_1}\|_2 \\ \varepsilon_\varphi \geq \|D_{k,i}\|_\infty &\geq \kappa_\varphi > 0 \end{aligned} \tag{AC}$$

where $\varepsilon_\varphi, \kappa_\varphi, \kappa_g > 0$, $g_{k,i} = \nabla H_k(u_{k,i})$ and m_1 indexes the recursion. $D_{k,i}$ is a diagonal matrix given by

$$(D_{k,i})_{ll} = \begin{cases} (\overline{\varphi}_k - u_{k,i})_l & \text{if } (g_{k,i})_l < 0 \\ (u_{k,i} - \underline{\varphi}_k)_l & \text{if } (g_{k,i})_l \geq 0 \end{cases}$$

RMTR$_\infty$ Algorithm, **Input**: $u_{k,0}, \Delta_{k,0}, \underline{\varphi}_k, \overline{\varphi}_k, H_k$
Pre-smoothing call Algorithm 6 with $u_{k,0}, \Delta_{k,0}, \underline{\varphi}_k, \overline{\varphi}_k, H_k, m_1$ receive u_{k,m_1}, Δ_{k,m_1} *Recursion* compute $\underline{\varphi}_{k-1}, \overline{\varphi}_{k-1}, H_{k-1}$ call RMTR$_\infty$ with $u_{k,m_1}, \Delta_{k,m_1}, \underline{\varphi}_{k-1}, \overline{\varphi}_{k-1}, H_{k-1}$, receive s_{k-1} and compute $s_{k,m_1} = I_{k-1}s_{k-1}$ **if** $(\rho_{k,m_1}(s_{k,m_1}) \geq \eta_1)$ $u_{k,m_1+1} = u_{k,m_1} + s_{k,m_1}$ **otherwise** $u_{k,m_1+1} = u_{k,m_1}$ compute a new Δ_{k,m_1+1} *Post-smoothing* call Algorithm 6 with $u_{k,m_1+1}, \Delta_{k,m_1+1}, \underline{\varphi}_k, \overline{\varphi}_k, H_k, m_2$ receive u_{k,m_2}, Δ_{k,m_2} **if** $(k == j)$ **goto** *Pre-smoothing* **else return** $u_{k,m_2} - u_{k,0}$

Algorithm 7: RMTR$_\infty$

In the remainder, we abbreviate $\widehat{g}_{k,i} = D_{k,i}g_{k,i}$. Finally, we follow [1] and assume that corrections computed in Algorithm 6 satisfy

$$\psi_{k,i}(s_{k,i}) < \beta_1 \psi_{k,i}(s^C_{k,i}) \tag{CC}$$

where $\beta_1 > 0$ and $s^C_{k,i} \in \mathbb{R}^{n_k}$ solves

$$\psi_{k,i}(s^C_{k,i}) = \min_{t\geq 0:\ s=-tD^2_{k,i}g_{k,i}} \{\psi_{k,i}(s) : \|s\|_\infty \leq \Delta_{k,i} \text{ and } \underline{\varphi}_k \leq u_{k,i} + s \leq \overline{\varphi}_k\} \tag{3}$$

Now, we can now cite Lemma 3.1 from [1].

Lemma 1. *Let* (A1)–(A3) *and* (AC) *hold. Then if* $s_{k,i}$ *in Algorithm* 6 *satisfies* (CC)*, we obtain*

$$-\psi_{k,i}(s_{k,i}) \geq c\|\widehat{g}_{k,i}\|_2 \min\{\Delta_{k,i}, \|\widehat{g}_{k,i}\|_2\} \tag{4}$$

To obtain the next results, the number of applied V-cycles will be indexed by ν, so that $u^\nu_{k,i}$, denotes the i-th iterate on Level k in Cycle ν.

Theorem 1. *Assume that* (A1)–(A3) *and* (AC) *hold. Moreover, assume that in Algorithm* 7 *at least* $m_1 > 0$ *or* $m_2 > 0$ *holds. We also assume that for each* $s_{k,i}$ *computed* *in Algorithm* 6 (CC) *holds. Then for each sequence of iterates* $(u^{\nu}_{j,i})_{\nu,i}$*, we obtain* $\liminf_{\nu\to\infty}\|\widehat{g}^{\nu}_{j,i}\|_2 = 0$.

Proof. We will prove the result by contradiction, i.e., we assume that $\exists\varepsilon > 0$ and a sequence $(u^{\nu}_{j,i})_{\nu,i}$ such that $\liminf_{\nu\to\infty}\|\widehat{g}^{\nu}_{j,i}\|_2 \geq \varepsilon$.

In this case, one can show that $\Delta^{\nu}_{j,i} \to 0$. Namely, if $\rho^{\nu}_{j,i} \geq \eta_1$ holds only finitely often, we obtain that $\Delta^{\nu}_{j,i}$ is increased finitely often but decreased infinitely often. On the other hand, for infinitely many iterations with $\rho^{\nu}_{j,i} \geq \eta_1$ we obtain

$$H^{\nu}_k(u^{\nu}_{k,i}) - H^{\nu}_k(u^{\nu}_{k,i} + s^{\nu}_{k,i}) \geq -\eta_1 \psi^{\nu}_{k,i}(s^{\nu}_{k,i})$$

We now exploit Lemma 1, (A1), and $\|\widehat{g}^{\nu}_{j,i}\|_2 \geq \varepsilon$ and obtain for sufficiently large ν that

$$H^{\nu}_k(u^{\nu}_{k,i}) - H^{\nu}_k(u^{\nu}_{k,i} + s^{\nu}_{k,i}) \geq c\Delta^{\nu}_{k,i} \geq c\|s^{\nu}_{k,i}\|_\infty \to 0$$

Next, we employ the mean value theorem, i.e., $\langle s^{\nu}_{k,i}, \overline{g}^{\nu}_{k,i}\rangle = H_k(u^{\nu}_{k,i} + s^{\nu}_{k,i}) - H_k(u^{\nu}_{k,i})$, as well as (A2) and (A3) and obtain for each trust-region correction

$$\begin{aligned}|\mathrm{pred}^{\nu}_{k,i}(s^{\nu}_{k,i})|\,|\rho^{\nu}_{k,i} - 1| &= \left|H^{\nu}_k(u^{\nu}_{k,i} + s^{\nu}_{k,i}) - H^{\nu}_k(u^{\nu}_{k,i}) + \langle s^{\nu}_{k,i}, g^{\nu}_{k,i}\rangle + \frac{1}{2}\langle s^{\nu}_{k,i}, B^{\nu}_{k,i}s^{\nu}_{k,i}\rangle\right| \\ &\leq \frac{1}{2}|\langle s^{\nu}_{k,i}, B^{\nu}_{k,i}s^{\nu}_{k,i}\rangle| + |\langle s^{\nu}_{k,i}, \overline{g}^{\nu}_{k,i} - g^{\nu}_k(u^{\nu}_{k,i})\rangle| \\ &\leq \frac{1}{2}c_B(\Delta^{\nu}_{k,i})^2 + \|\overline{g}^{\nu}_{k,i} - g^{\nu}_k(u^{\nu}_{k,i})\|_2\Delta^{\nu}_{k,i}\end{aligned}$$

Due to the convergence of $\Delta^{\nu}_{k,i}$, and, hence, of $(u^{\nu}_{k,i})_{\nu,i}$, and the continuity of $g^{\nu}_{k,i}$, we obtain $\rho^{\nu}_{k,i} \to 1$ for $i \neq m_1$. Hence, on each level, for sufficiently small $\Delta^{\nu}_{j,i}$, trust-region corrections are successful and applied.

One can also show, that for sufficiently small $\Delta^{\nu}_{j,i}$ recursively computed corrections will be computed and applied: we find that there exists a $c > 0$ such that $\Delta^{\nu}_{k,i} \geq c\Delta^{\nu}_{j,m_1}$ (cf., [6]). In turn, (AC) provides that there exists another constant such that

$$H^{\nu}_k(u^{\nu}_{k,m_1}) - H^{\nu}_k(u^{\nu}_{k,m_1} + s^{\nu}_{k,m_1}) \geq c\|\widehat{g}^{\nu}_{k,m_1}\|_2 \min\{c\Delta^{\nu}_{j,m_1}, \|\widehat{g}^{\nu}_{k,m_1}\|_2\}$$

(cf., the proof of Lemma 4.4, [6]). Now, one can show, cf. Theorem 4.6, [6], that the contraction rates for recursively computed corrections also tend to one, i.e., $\rho^{\nu}_{j,m_1} \to 1$.

Since $\Delta^{\nu}_{j,i} \to 0$, we obtain $(\rho^{\nu}_{j,i})_{\nu,i} \to 1$. But this contradicts $\Delta^{\nu}_{j,i} \to 0$ and $\liminf_{\nu\to\infty}\|\widehat{g}^{\nu}_{j,i}\|_2 \to 0$ must hold. □

Using exactly the same argumentation as used in Theorem 6.6 of [6], we obtain convergence to first-order critical points, i.e., $\lim_{\nu\to\infty}\|\widehat{g}^{\nu}_{j,i}\|_2 = 0$.

Theorem 2. *Let assumptions* (A1)–(A3), (AC) *hold. Moreover, assume that al least one pre- or post-smoothing step in Algorithm* 7 *will be performed and that* (CC) *holds for each correction computed in Algorithm* 6*. Then, for each sequence of iterates* $(u^{\nu}_{j,i})_{\nu,i}$ *we obtain* $\lim_{\nu\to\infty}\|\widehat{g}^{\nu}_{j,i}\|_2 = 0$.

4 Numerical Example

In this section, we present an example from the field of non-linear elasticity computed with the RMTR$_\infty$ algorithm which is implemented in OBSLIB++, cf. [7].

R. W. Ogden has introduced a material law for rubber-like materials (cf., [9]). The associated stored energy function is highly non-linear due to a penalty term which prevents the inversion of element volumes:

$$W(\nabla\varphi) = a \cdot \mathrm{tr}E + b \cdot (\mathrm{tr}E)^2 + c \cdot \mathrm{tr}(E^2) + d \cdot \Gamma(\det(\nabla\varphi)) \tag{5}$$

where $\varphi = id + u$. This function is a polyconvex stored energy function depending on the Green - St. Venant strain tensor $E(u)$, i.e., $E(u) = \frac{1}{2}(\nabla u^T + \nabla u + \nabla u^T \nabla u)$, a penalty function $\Gamma(x) = -\ln(x)$ for $x \in \mathbb{R}^+$, and $a = -d, b = \lambda - d, c = \mu + d, d > 0$.

Fig. 1 shows the results of the computation of a contact problem with parameters $\lambda = 34$, $\mu = 136$, and $d = 100$. In particular, a skewed pressure is applied on the top side of the cube, which results in that the cube is pressed towards an obstacle. Problem (M) with $J(u) = \int_\Omega W(\nabla\varphi)dx$ was solved in a finite element framework using both a fine level trust-region strategy and our RMTR$_\infty$ algorithm with $m_1 = m_2 = 2$. Equation (2) was (approximately) solved using 10 projected cg-steps.

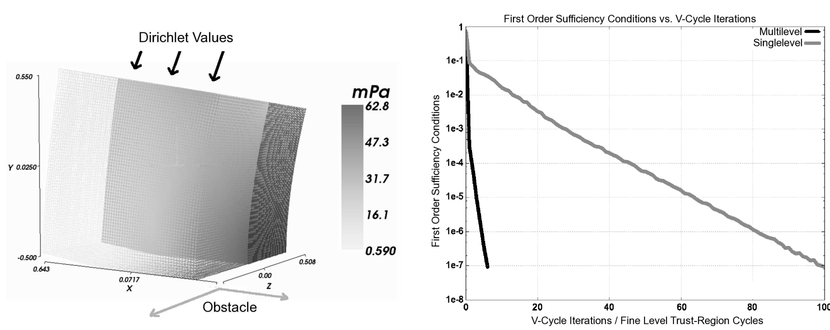

Fig. 1. Nonlinear Elastic Contact Problem 823,875 degrees of freedom. *Left image:* Deformed mesh and von-Mises stress distribution. *Right image:* Comparison of $(\|\widehat{g}^{\nu}_{j,0}\|_2)_\nu$ computed by our RMTR$_\infty$ algorithm (black line) and by a trust-region strategy, performed only on the finest grid (grey line).

References

[1] Coleman, T.F., Li, Y.: An interior trust region approach for nonlinear minimization subject to bounds. *SIAM J. Optim.*, 6(2):418–445, 1996.

[2] Gelman, E., Mandel, J.: On multilevel iterative methods for optimization problems. *Math. Program.*, 48(1):1–17, 1990.

[3] Gratton, S., Mouffe, M., Toint, P.L., Weber-Mendonca, M.: A recursive trust-region method in infinity norm for bound-constrained nonlinear optimization. *IMA J. Numer. Anal.*, 28(4):827–861, 2008.

[4] Gratton, S., Sartenaer, A., Toint, P.L.: Recursive trust-region methods for multiscale nonlinear optimization. *SIAM J. Optim.*, 19(1):414–444, 2008.

[5] Groß, C.: *A Multiscale Approach for Convex and Non-convex Constrained Minimization*. Diploma thesis, University Bonn, April 2007.

[6] Groß, C., Krause, R.: On the convergence of recursive trust-region methods for multiscale non-linear optimization and applications to non-linear mechanics. Technical Report 710, Institute for Numerical Simulation, University of Bonn, March 2008. Submitted.

[7] Krause, R.: A parallel decomposition approach to non-smooth minimization problems—concepts and implementation. Technical Report 709, Institute for Numerical Simulation, University of Bonn, Germany, 2007.

[8] Nash, S.G.: A multigrid approach to discretized optimization problems. *Optim. Methods Softw.*, 14(1-2):99-116, 2000. ISSN 1055-6788. International Conference on Nonlinear Programming and Variational Inequalities (Hong Kong, 1998).

[9] Ogden, R.W.: *Nonlinear elastic deformations*. Ellis Horwood Series: Mathematics and its Applications. Ellis Horwood, Chichester, 1984. ISBN 0-85312-273-3.

[10] Ulbrich, M., Ulbrich, S., Heinkenschloss, M.: Global convergence of trust-region interior-point algorithms for infinite-dimensional nonconvex minimization subject to pointwise bounds. *SIAM J. Control Optim.*, 37(3):731–764 (electronic), 1999. ISSN 0363-0129.