Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

RALF KORNHUBER

Globally Convergent Multigrid Methods for Porous Medium Type Problems

Preprint SC 97-45 (January 1998)

GLOBALLY CONVERGENT MULTIGRID METHODS FOR POROUS MEDIUM TYPE PROBLEMS

RALF KORNHUBER*

Abstract. We consider the fast solution of large, piecewise smooth minimization problems as typically arising from the finite element discretization of porous media flow. For lack of smoothness, usual Newton multigrid methods cannot be applied. We propose a new approach based on a combination of convex minization with *constrained* Newton linearization. No regularization is involved. We show global convergence of the resulting monotone multigrid methods and give logarithmic upper bounds for the asymptotic convergence rates.

1. Introduction

Let Ω be a polyhedral domain in the Euclidean space $\mathbb{R}^d$. We consider the minimization problem

$$u \in H : \quad \mathcal{J}(u) + \phi(u) \leq \mathcal{J}(v) + \phi(v) \ \ \forall v \in H \tag{1.1}$$

on a closed subspace $H \subset H^1(\Omega)$. For simplicity, we concentrate on $H = H_0^1(\Omega)$ and $d = 2$. Other boundary conditions of Neumann or mixed type and the case of three space dimensions can be treated in a similar way [3, 4]. The quadratic functional $\mathcal{J}$,

$$\mathcal{J}(v) = \tfrac{1}{2}a(v,v) - \ell(v), \tag{1.2}$$

is induced by a continuous, symmetric and H–elliptic bilinear form $a(\cdot,\cdot)$ representing a differential operator of second order and by a linear functional $\ell \in H'$. H is equipped with the energy norm $\|\cdot\| = a(\cdot,\cdot)^{1/2}$. The convex functional ϕ of the form

$$\phi(v) = \int_\Omega \Phi(v(x))\, dx, \tag{1.3}$$

Date: January 30, 1998.

* Mathematisches Institut A, Universität Stuttgart Pfaffenwaldring 57, D-70569 Stuttgart.

The author gratefully acknowledges the hospitality of P. Deuflhard and his staff at the Konrad–Zuse–Center Berlin during the preparation of this manuscript. The work was supported by a Konrad–Zuse–Fellowship.

is generated by a scalar convex function $\Phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$. Assuming $\Phi(z) = \infty$ for all $z < 0$ and the local Lipschitz condition

$$|\Phi(z) - \Phi(z')| \leq C(|z| + |z'|))|z - z'| \quad \forall z, z' \geq 0 \tag{1.4}$$

with fixed $C > 0$, the functional ϕ is convex, lower semi–continuous and proper (i.e. $\phi(v) > -\infty$ and $\phi \not\equiv +\infty$). As a consequence, (1.1) admits a unique solution $u \in H$. Moreover, (1.1) can be rewritten as the variational inequality

$$u \in H : \ a(u, v - u) + \phi(v) - \phi(u) \geq \ell(v - u) \ \ \forall v \in H \tag{1.5}$$

or as the variational inclusion

$$u \in H : \quad a(u, v) - \ell(v) \in \partial\phi(u)(v) \quad \forall v \in H, \tag{1.6}$$

where $\partial\phi$ denotes the subdifferential of ϕ. See e.g. [7, 13] for further information.

Later on, we will additionally assume that the second derivative $\Phi''(z)$ exists and is locally Lipschitz for $z > 0$. We emphasize that Φ'' may have *no* continuous extension to $z = 0$.

As a typical example consider the porous medium equation

$$\rho_t = \Delta\rho^m - f(\rho), \quad \rho \geq 0, \tag{1.7}$$

with $m \geq 1$, monotonically increasing absorption f and suitable boundary conditions. After Kirchhoff type transformation $u = K(\rho) := \max\{0, \rho^m\}$ and implicit time discretization the weak formulation of the resulting spatial problems is given by (1.6). In this case, the scalar function Φ has to be chosen such that $\partial\Phi = (\mathrm{id} + \Delta t f)(K^{-1}(\cdot))$. Similar problems arise in a wide range of problems from mechanics, physical and biological science, metallurgy, etc.

Let $\mathcal{T}_j$ be a given partition of Ω in triangles $t \in \mathcal{T}_j$ with minimal diameter of order 2^{-j}. The set of interior nodes is called $\mathcal{N}_j$. Discretizing (1.1) by continuous, piecewise linear finite elements $\mathcal{S}_j \subset H$, we obtain the finite dimensional problem

$$u_j \in \mathcal{S}_j : \quad \mathcal{J}(u_j) + \phi_j(u_j) \leq \mathcal{J}(v) + \phi_j(v) \quad \forall v \in \mathcal{S}_j. \tag{1.8}$$

Observe that the functional ϕ is approximated by $\mathcal{S}_j$–interpolation of $\Phi(v)$, giving

$$\phi_j(v) = \sum_{p \in \mathcal{N}_j} \Phi(v(p))h_p, \qquad h_p = \int_\Omega \lambda_p^{(j)}(x)\, dx, \tag{1.9}$$

with $\Lambda_j = \{\lambda_p^{(j)},\ p \in \mathcal{N}_j\}$ denoting the nodal basis of $\mathcal{S}_j$. Observe that the discrete energy $\mathcal{J} + \phi_j$ is finite and continuous on the closed convex set $\mathcal{K}_j \subset \mathcal{S}_j$,

$$\mathcal{K}_j = \{v \in \mathcal{S}_j \mid v(p) \geq 0 \ \forall p \in \mathcal{N}_j\}.$$

Of course, (1.8) is uniquely solvable and can be reformulated as the variational inequality

$$u_j \in \mathcal{S}_j : \ a(u_j, v - u_j) + \phi_j(v) - \phi_j(u_j) \geq \ell(v - u_j) \ \ \forall v \in \mathcal{S}_j \tag{1.10}$$

or as the variational inclusion

$$u_j \in \mathcal{S}_j : \quad a(u_j, v) - \ell(v) \in \partial\phi_j(u_j)(v) \quad \forall v \in \mathcal{S}_j. \tag{1.11}$$

For convergence results we refer to [7, 13] and the bibliography cited therein.

In the sequel, we will concentrate on the fast solution of the discrete minimization problem (1.8). It is clear that usual Newton techniques cannot be applied, because the functional ϕ_j is not differentiable. If Φ is smooth on $z > 0$, then the Fréchèt derivative $\phi_j''(v)$ exists for positive $v \in \mathcal{S}_j$, but may *not* be uniformly Lipschitz on these functions. Hence, even for given coincidence set $\mathcal{N}_j^{\bullet} = \{p \in \mathcal{N}_j \mid u_j(p) = 0\}$, Newton–multigrid methods in the spirit of Bank and Rose [1] or Deuflhard and Weiser [5] as well as the nonlinear multigrid techniques of Hackbusch and Reusken [8] are not applicable.

A common remedy is to use Newton-type iterations after some suitable regularization of ϕ_j. Unfortunately, reasonable convergence speed then may have to be paid by unacceptable discretization errors and vice versa.

In this paper, we introduce a completely new approach, extending recent *monotone multigrid methods* [10, 11, 13] from piecewise quadratic functionals ϕ to the piecewise smooth case. Monotone multigrid methods can be regarded as two-stage iterations consisting of a globally convergent fine grid smoother $\mathcal{M}_j$ and a coarse grid correction $\mathcal{C}_j$ that has to provide monotonically decreasing energy in order to preserve global convergence. The basic idea for constructing $\mathcal{C}_j$ is first to choose a neighborhood of the actual smoothed iterate in which Newton linearization can be controlled by pointwise Lipschitz constants and then to constrain the coarse grid correction to this neighborhood. In this way, our approach avoids any regularization. As usual, we need a suitable damping of the Newton correction. Utilizing *local* damping parameters for the local corrections (each one associated with a fixed node on a fixed grid level), we get maximal effect of the coarse grid correction together with global convergence.

The paper is organized as follows. We first investigate inexact variants of the well-known nonlinear Gauß–Seidel smoother. Then, we provide a general framework for constructing monotone coarse grid corrections and state a general convergence result. On this background, we present standard and truncated monotone multigrid methods and prove logarithmic bounds for the asymptotic convergence rates. Numerical experiments, illustrating the efficiency and robustness of the method can be found in [12].

2. Inexact Gauss-Seidel Iteration

The well–known Gauß–Seidel method [7, 13] for the iterative solution of (1.8) is based on the successive minimization of the discrete energy functional $\mathcal{J} + \phi_j$ in the direction of the nodal basis functions $\lambda_{p_l}^{(j)}$, $l = 1, \ldots, n_j = \#\mathcal{N}_j$. The local correction $T_l w \in V_l = \operatorname{span}\{\lambda_{p_l}^{(j)}\}$ of some given $w \in \mathcal{K}_j$ is defined by

$$T_l w \in V_l : \quad \mathcal{J}(w + T_l w) + \phi_j(w + T_l w) \le \mathcal{J}(w + v) + \phi_j(w + v) \quad \forall v \in V_l \tag{2.1}$$

with straightforward modification for $w \notin \mathcal{K}_j$. In general, the solution $T_l w$ of the local problems (2.1) is not available in closed form. For this reason, we consider an *inexact Gauß–Seidel iteration* defined as follows.

For given iterate u_j^ν we compute a sequence of intermediate iterates w_l^ν according to

$$w_0^\nu = u_j^\nu, \qquad w_l^\nu = w_{l-1}^\nu + v_l^\nu, \quad l = 1, \ldots, n_j, \tag{2.2}$$

with suitable approximations $v_l^\nu \in V_l$ of $T_l w_{l-1}^\nu$. Finally, the new iterate is given by

$$u_j^{\nu+1} = \mathcal{M}_j u_j^\nu = w_{n_j}^\nu. \tag{2.3}$$

For notational convenience, the index ν will be frequently skipped in the sequel.

Theorem 2.1. *Assume that the corrections v_l in (2.2) are chosen in such a way that $\mathcal{M}_j u_j^0 \in \mathcal{K}_j$ holds for all $u_j^0 \in \mathcal{S}_j$ and*

$$v_l = \omega(w) T_l w, \quad \omega(w) \in [\omega_0, 1] \qquad \forall w \in \mathcal{K}_j \tag{2.4}$$

is valid with some fixed $\omega_0 \in (0,1]$. Then the inexact Gauß–Seidel iteration (2.3) is globally convergent.

Proof. Proof. We will use the abbreviation $\bar{\mathcal{J}} = \mathcal{J} + \phi_j$. Utilizing (2.4) and the convexity of $\bar{\mathcal{J}}$, we obtain the monotonicity

$$\bar{\mathcal{J}}(w_l) \le \bar{\mathcal{J}}(w_{l-1} + \omega_0 T_l w_{l-1}) \le \bar{\mathcal{J}}(w_{l-1}), \quad l = 1, \dots, n_j. \tag{2.5}$$

As a consequence, we get $\bar{\mathcal{J}}(u_j^{\nu+1}) \le \bar{\mathcal{J}}(u_j^{\nu}) \le \bar{\mathcal{J}}(u_j^1) < \infty$ for all $\nu \ge 1$. Since ϕ_j is convex, lower semicontinuous and proper, there exist c, $C \in \mathbb{R}$ such that

$$\phi_j(v) \ge c\|v\| + C \quad \forall v \in \mathcal{S}_j \tag{2.6}$$

(cf. e.g. [6]). From (2.6) and from the boundedness of $(\bar{\mathcal{J}}(u_j^{\nu}))_{\nu \ge 1}$ we conclude that the sequence $(u_j^{\nu})_{\nu \ge 0}$ must also be bounded. Let $(u_j^{\nu_k})_{k \ge 0} \subset \mathcal{K}_j$ be a convergent subsequence with limit $u_j^* \in \mathcal{K}_j$. We will show that $u_j^* = u_j$.

Observe that the estimate

$$\ell(T_l w) - a(w + T_l w, T_l w) + \phi_j(w) - \phi_j(w + T_l w) \ge 0 \tag{2.7}$$

is resulting from the variational formulation of (2.1). Utilizing the monotonicity (2.5), (2.7) and the convexity estimate

$$\phi_j(w) - \phi_j(w + \omega_0 T_l w) \ge \omega_0(\phi_j(w) - \phi_j(w + T_l w)),$$

we obtain

$$\bar{\mathcal{J}}(u_j^{\nu_k}) - \bar{\mathcal{J}}(u_j^{\nu_{k+1}}) \ge \omega_0(1 - \frac{\omega_0}{2}) \sum_{i=1}^{n_j} \|T_i w_{i-1}^{\nu_k}\|^2. \tag{2.8}$$

On the other hand, the triangle inequality, the Cauchy–Schwarz inequality and (2.4) lead to

$$\|u_j^{\nu_k} - w_{l-1}^{\nu_k}\|^2 \le n_j \sum_{i=1}^{n_j} \|T_i w_{i-1}^{\nu_k}\|^2, \qquad l = 1, \dots, n_j. \tag{2.9}$$

Since $\bar{\mathcal{J}}$ is continuous on $\mathcal{K}_j$, we conclude from (2.8) and (2.9) that

$$w_{l-1}^{\nu_k} \to u_j^*, \quad k \to \infty, \qquad l = 1, \dots, n_j.$$

The monotonicity (2.5) yields

$$\bar{\mathcal{J}}(u_j^{\nu_{k+1}}) \le \bar{\mathcal{J}}(u_j^{\nu_k + 1}) \le \bar{\mathcal{J}}(w_l^{\nu_k}) \le \bar{\mathcal{J}}(w_{l-1}^{\nu_k} + \omega_0 T_l w_{l-1}^{\nu_k}) \le \bar{\mathcal{J}}(u_j^{\nu_k}) \tag{2.10}$$

for each fixed $l = 1, \dots, n_j$. Since $\bar{\mathcal{J}}$ and T_l are continuous on $\mathcal{K}_j$, we can pass to the limit so that

$$\bar{\mathcal{J}}(u_j^*) = \bar{\mathcal{J}}(u_j^* + \omega_0 T_l u_j^*).$$

Moreover, the convexity of $\bar{\mathcal{J}}$ and (2.1) imply $\bar{\mathcal{J}}(u_j^*) = \bar{\mathcal{J}}(u_j^* + T_l u_j^*)$. As $T_l u_j^*$ is the unique solution of (2.1), we get $T_l u_j^* = 0$. The same holds true for all $l = 1, \dots, n_j$ so that u_j^* must be a fixed point of the exact Gauß–Seidel iteration which is well-known to have the unique fixed point u_j. This concludes the proof. ∎

Observe that condition (2.4) can be replaced by the energy reduction

$$\mathcal{J}(w + v_l) + \phi_j(w + v_l) \le \mathcal{J}(w + \omega_0 T_l w) + \phi_j(w + \omega_0 T_l w) \tag{2.11}$$

together with the additional assumption $\|v_l\| \le c\|T_l w\|$.

Theorem 2.1 can be used as a stopping criterion for the iterative solution of (2.1). To give an example, let us first reformulate (2.1) as the scalar inclusion

$$0 \in g(z_l) = \partial\Phi(w(p) + z_l)h_{p_l} + a_{ll}z_l - r_l \tag{2.12}$$

where

$$z_l\lambda_{p_l}^{(j)} = T_l w, \quad a_{ll} = a(\lambda_{p_l}^{(j)}, \lambda_{p_l}^{(j)}), \quad r_l = \ell(\lambda_{p_l}^{(j)}) - a(w, \lambda_{p_l}^{(j)})$$

and $\partial\Phi$ is the set-valued subdifferential of Φ [6]. We will now describe a simple bisection method for the approximate solution of (2.12). First, let $w_0 = \max\{0, -w(p)\}$. Now we have to distinguish three cases. Of course, $z_l = w_0$ is the exact solution, if $0 \in g(w_0)$. If $\overline{g} = \sup g(w_0) < 0$, then it is easily checked that $z_l \in [\underline{z}^0, \overline{z}^0]$ with $\underline{z}^0 = w_0$ and $\overline{z}^0 = -\overline{g}/a_{ll} > w_0$. Starting with $[\underline{z}^0, \overline{z}^0]$, we continue bisection until the new midpoint $z^i = (\underline{z}^i + \overline{z}^i)/2$ satisfies $0 \in g(z^i)$ or $\sup g(z^i) < 0$. Then $v_l = z^i\lambda_p^{(j)}$ has the property (2.4) with $\omega_0 = \frac{1}{2}$. In the remaining case $\inf g(w_0) > 0$ we first conclude $w_0 = 0$. Then we proceed in a symmetrical way starting with $\underline{z}^0 = -w(p) < 0$ and $\overline{z}^0 = 0$. Finally, it is clear that $(w + v_l)(p_l) \ge 0$, giving $\mathcal{M}_j u_j^0 \in \mathcal{K}_j$ for all $u_j^0 \in \mathcal{S}_j$.

More sophisticated algorithms based on secant approximations or Newton linearization can be constructed in a similar way.

3. Monotone Iterations

The (inexact) Gauß-Seidel iteration $\mathcal{M}_j$, as introduced in (2.3), typically suffers from rapidly deteriorating convergence rates when proceeding to more and more refined triangulations. As a remedy, we consider so-called *monotone iterations*

$$\begin{aligned} \bar{u}_j^\nu &= \mathcal{M}_j u_j^\nu \\ u_j^{\nu+1} &= \mathcal{C}_j \bar{u}_j^\nu \end{aligned} \tag{3.1}$$

where the additional substep $\mathcal{C}_j$ is intended to accelerate the convergence speed. Adopting multigrid terminology, $\mathcal{M}_j$ is called (fine grid) *smoother*, $\bar{u}_j^\nu$ is a *smoothed iterate* and $\mathcal{C}_j$ is called *coarse grid correction*.

Theorem 3.1. *Assume that the smoother $\mathcal{M}_j$ satisfies the conditions of Theorem 2.1 and that the coarse grid correction $\mathcal{C}_j$ has the monotonicity property*

$$\mathcal{J}(\mathcal{C}_j w) + \phi_j(\mathcal{C}_j w) \le \mathcal{J}(w) + \phi_j(w) \qquad \forall w \in \mathcal{K}_j. \tag{3.2}$$

Then the iteration (3.1) is globally convergent.

Proof. Proof. Exploiting (3.2), the proof is almost the same as for Theorem 2.1. For example, (2.10) now takes the form

$$\bar{\mathcal{J}}(u_j^{\nu_{k+1}}) \le \bar{\mathcal{J}}(\mathcal{C}_j \bar{u}_j^{\nu_k}) \le \bar{\mathcal{J}}(\bar{u}_j^{\nu_k}) \le \bar{\mathcal{J}}(w_{l-1}^{\nu_k} + \omega_0 T_l w_{l-1}^{\nu_k}) \le \bar{\mathcal{J}}(u_j^{\nu_k}).$$

□

As a by-product we also get the convergence of the smoothed iterates

$$\bar{u}_j^\nu \to u_j \qquad \nu \to \infty. \tag{3.3}$$

We emphasize that the coarse grid correction *alone* does not need to be convergent. This gives a lot of flexibility in constructing $\mathcal{C}_j$.

4. Monotone Coarse Grid Correction with Local Damping

Recall that classical Newton multigrid methods cannot be applied to (1.8) for lack of smoothness. In this section, we will derive *constrained* Newton multigrid methods to be used as coarse grid correction $\mathcal{C}_j$. Throughout the following, we assume that

$$\Phi \in C^2(0,\infty), \qquad \Phi'' \text{ is locally Lipschitz on } (0,\infty). \tag{4.1}$$

For given smoothed iterate $\bar{u}_j^\nu$, we introduce the subset of *regular nodes*

$$\mathcal{N}_j^\circ(\bar{u}_j^\nu) = \{p \in \mathcal{N}_j \mid \bar{u}_j^\nu(p) > 0\} \subset \mathcal{N}_j. \tag{4.2}$$

Consider some fixed $p \in \mathcal{N}_j^\circ(\bar{u}_j^\nu)$. Then, as a consequence of (4.1), there exists a neighborhood of $\bar{u}_j^\nu(p)$

$$0 < \underline{\varphi}_{\bar{u}_j^\nu}(p) < \bar{u}_j^\nu(p) < \overline{\varphi}_{\bar{u}_j^\nu}(p), \tag{4.3}$$

where the local Lipschitz condition

$$|\Phi''(z_1) - \Phi''(z_2)| \le L_p^\nu |z_1 - z_2| \quad \forall z_1, z_2 \in [\underline{\varphi}_{\bar{u}_j^\nu}(p), \overline{\varphi}_{\bar{u}_j^\nu}(p)] \tag{4.4}$$

holds with *pointwise Lipschitz constant* $L_p^\nu > 0$. At the remaining *critical nodes*

$$\mathcal{N}_j^\bullet(\bar{u}_j^\nu) = \mathcal{N}_j \setminus \mathcal{N}_j^\circ(\bar{u}_j^\nu)$$

we set

$$\underline{\varphi}_{\bar{u}_j^\nu}(p) = \overline{\varphi}_{\bar{u}_j^\nu}(p) = \bar{u}_j^\nu(p). \tag{4.5}$$

Collecting these intervals for all $p \in \mathcal{N}_j$, we introduce the neighborhood $\mathcal{K}_{\bar{u}_j^\nu}$ of $\bar{u}_j^\nu$,

$$\mathcal{K}_{\bar{u}_j^\nu} = \{w \in \mathcal{S}_j |\ \underline{\varphi}_{\bar{u}_j^\nu}(p) \le w(p) \le \overline{\varphi}_{\bar{u}_j^\nu}(p),\ p \in \mathcal{N}_j\} \subset \mathcal{S}_j.$$

From the above definitions, we obtain the local representation of ϕ_j

$$\phi_j(w) = \phi_{\bar{u}_j^\nu}(w) + \text{const.} \qquad \forall w \in \mathcal{K}_{\bar{u}_j^\nu} \tag{4.6}$$

by the smooth functional $\phi_{\bar{u}_j^\nu}$

$$\phi_{\bar{u}_j^\nu}(w) = \sum_{p \in \mathcal{N}_j^\circ(\bar{u}_j^\nu)} \Phi(w(p)) h_p, \qquad w \in \mathcal{K}_{\bar{u}_j^\nu}. \tag{4.7}$$

Let us consider the *constrained minimization* of the *smooth energy* $\mathcal{J} + \phi_{\bar{u}_j^\nu}$

$$u_j^* \in \mathcal{K}_{\bar{u}_j^\nu} : \quad \mathcal{J}(u_j^*) + \phi_{\bar{u}_j^\nu}(u_j^*) \le \mathcal{J}(v) + \phi_{\bar{u}_j^\nu}(v) \quad \forall v \in \mathcal{K}_{\bar{u}_j^\nu}. \tag{4.8}$$

We will see later on that, for non-degenerate problems (1.8), $u_j \in \mathcal{K}_{\bar{u}_j^\nu}$ holds after a finite number of iteration steps. In this case, our original non-smooth problem asymptotically reduces to the constrained smooth problem (4.8). Moreover, the convergence $\operatorname{dist}(u_j, \mathcal{K}_{\bar{u}_j^\nu}) \to 0$, $\nu \to \infty$, which is an immediate consequence of (3.3), suggests to improve the actual smoothed iterate $\bar{u}_j^\nu$ by the (approximate) solution of (4.8).

The main advantage of the constrained problem (4.8) is that Newton linearization can be applied to the smooth energy $\mathcal{J} + \phi_{\bar{u}_j^\nu}$. More precisely, we approximate $\mathcal{J} + \phi_{\bar{u}_j^\nu}$ by the quadratic energy functional $\mathcal{J}_{\bar{u}_j^\nu}$,

$$\mathcal{J}_{\bar{u}_j^\nu}(w) = \tfrac{1}{2} a_{\bar{u}_j^\nu}(w,w) - \ell_{\bar{u}_j^\nu}(w) \approx \mathcal{J}(w) + \phi_{\bar{u}_j^\nu}(w) + \text{const.}, \quad w \in \mathcal{K}_{\bar{u}_j^\nu},$$

where the bilinear form

$$a_{\bar{u}_j^\nu}(w,w) = a(w,w) + \phi''_{\bar{u}_j^\nu}(\bar{u}_j^\nu)(w,w) \tag{4.9}$$

and the linear functional

$$\ell_{\bar{u}_j^\nu}(w) = \ell(w) - \phi'_{\bar{u}_j^\nu}(\bar{u}_j^\nu)(w) + \phi''_{\bar{u}_j^\nu}(\bar{u}_j^\nu)(\bar{u}_j^\nu, w)$$

are resulting from Taylor's expansion

$$\phi_{\bar{u}_j^\nu}(w) \approx \phi_{\bar{u}_j^\nu}(\bar{u}_j^\nu) + \phi'_{\bar{u}_j^\nu}(\bar{u}_j^\nu)(w - \bar{u}_j^\nu) + \tfrac{1}{2}\phi''_{\bar{u}_j^\nu}(\bar{u}_j^\nu)(w - \bar{u}_j^\nu, w - \bar{u}_j^\nu).$$

The solution of the resulting linearized problem

$$u_j^* \in \mathcal{K}_{\bar{u}_j^\nu} : \quad \mathcal{J}_{\bar{u}_j^\nu}(u_j^*) \le \mathcal{J}_{\bar{u}_j^\nu}(v) \quad \forall v \in \mathcal{K}_{\bar{u}_j^\nu} \tag{4.10}$$

is now approximated by one step of an *extended underrelaxation* as introduced in [10]. More precisely, we chose the search directions μ_l^ν,

$$\mu_l^\nu \in \mathcal{S}_j, \quad \max_{x \in \bar{\Omega}} \mu_l^\nu(x) = 1, \quad l = n_j + 1, \dots, m_j,$$

which may depend on the actual constraints $\mathcal{K}_{\bar{u}_j^\nu}$ and define the corresponding one-dimensional subspaces $V_l^\nu = \operatorname{span}\{\mu_l^\nu\}$. Then, we compute a sequence w_l^ν of intermediate iterates according to

$$w_{n_j}^\nu = \bar{u}_j^\nu, \qquad w_l^\nu = w_{l-1}^\nu + \omega_l^\nu v_l^\nu, \quad l = n_j + 1, \dots, m_j, \tag{4.11}$$

where each local correction v_l^ν is the solution of the obstacle problem

$$v_l^\nu \in \mathcal{D}_l^\nu : \quad \mathcal{J}_{\bar{u}_j^\nu}(w_{l-1}^\nu + v_l^\nu) \le \mathcal{J}_{\bar{u}_j^\nu}(w_{l-1}^\nu + v) \quad \forall v \in \mathcal{D}_l^\nu \tag{4.12}$$

with constraints $\mathcal{D}_l^\nu \subset V_l^\nu$ satisfying

$$0 \in \mathcal{D}_l^\nu \subset \{v \in V_l^\nu \mid w_{l-1}^\nu + v \in \mathcal{K}_{\bar{u}_j^\nu}\}. \tag{4.13}$$

In order to guarantee the monotonicity (3.2), the *local damping parameters* ω_l^ν are chosen such that

$$\mathcal{J}(w_l^\nu) + \phi_{\bar{u}_j^\nu}(w_l^\nu) \le \mathcal{J}(w_{l-1}^\nu) + \phi_{\bar{u}_j^\nu}(w_{l-1}^\nu). \tag{4.14}$$

Finally, our *monotone coarse grid correction with local damping* is given by

$$\mathcal{C}_j \bar{u}_j^\nu = w_{m_j^\nu}^\nu = \bar{u}_j^\nu + \sum_{l=1}^{m_j} \omega_l^\nu v_l^\nu. \tag{4.15}$$

Details on extended relaxations can be found in the textbook [13]. For instance, the convergence of the intermediate iterates

$$w_l^\nu \to u_j \qquad \nu \to \infty \tag{4.16}$$

can be shown in the same way as Corollary 2.3. We now derive a sufficient condition for the local monotonicity (4.14). As usual, the index ν will be frequently suppressed and we will use the abbreviation $z_+ = \max\{0, z\}_+$.

Proposition 4.1. *Let $v_l = z_l \mu_l$ be the solution of (4.12). Assume that $\omega_l \in [0,1]$ and*

$$\omega_l |z_l| \le 2 \left\{ \frac{|\ell_{\bar{u}_j^\nu}(\mu_l) - a_{\bar{u}_j^\nu}(w_{l-1}, \mu_l)| - L_l \|\bar{u}_j^\nu - w_{l-1}\|_{\infty,l}^2}{a_{\bar{u}_j^\nu}(\mu_l, \mu_l) + L_l \left(\|\bar{u}_j^\nu - w_{l-1}\|_{\infty,l} + \omega_l |z_l| \right)} \right\}_+ \tag{4.17}$$

with local Lipschitz constant

$$L_l = \sum_{p \in \mathcal{N}_j^\circ(\bar{u}_j^\nu)} L_p |\mu_l(p)|\, h_p \tag{4.18}$$

and local maximum norm

$$\|v\|_{\infty,l} = \max_{p \in \mathcal{N}_j \cap \text{ int supp } \mu_l} |v(p)|. \tag{4.19}$$

Then the damped correction $\omega_l v_l$ satisfies the local monotonicity condition (4.14).

Proof. Proof. The assertion is trivial for $z_l = 0$. Assuming $z_l \neq 0$, we introduce the scalar function

$$g(\omega) = \mathcal{J}(w_{l-1} + \omega v_l) + \phi_{\bar{u}_j^\nu}(w_{l-1} + \omega v_l).$$

Obviously, (4.14) is equivalent to $g(\omega_l) \leq g(0)$. As $g \in C^2[0,1]$, we can use Taylor's expansion to reformulate this condition as

$$0 \leq \omega_l \leq -2 \frac{g'(0)}{g''(\tau \omega_l)} \tag{4.20}$$

with suitable $\tau \in (0,1)$. To obtain a lower bound for $-g(0)$, we first state the estimate

$$\phi'_{\bar{u}_j^\nu}(w_{l-1})(v_l) \leq \phi'_{\bar{u}_j^\nu}(\bar{u}_j^\nu)(v_l) + \phi''_{\bar{u}_j^\nu}(\bar{u}_j^\nu)(w_{l-1} - \bar{u}_j^\nu, v_l) + L_l |z_l| \|w_{l-1} - \bar{u}_j^\nu\|_{\infty,l}^2$$

which is a consequence of Taylor's formula and the pointwise Lipschitz condition (4.4). Moreover, we have $\ell_{\bar{u}_j^\nu}(v_l) - a_{\bar{u}_j^\nu}(w_{l-1}, v_l) \geq 0$ because v_l is the solution of (4.12). Combining these estimates, we get the lower bound

$$\begin{aligned} -g'(0) &= -\mathcal{J}'(w_{l-1})(v_l) - \phi'_{\bar{u}_j^\nu}(w_{l-1})(v_l) \\ &\geq |\ell_{\bar{u}_j^\nu}(v_l) - a_{\bar{u}_j^\nu}(w_{l-1}, v_l)| - L_l |z_l| \|w_{l-1} - \bar{u}_j^\nu\|_{\infty,l}^2. \end{aligned} \tag{4.21}$$

Using

$$\begin{aligned} &\phi''_{\bar{u}_j^\nu}(w_{l-1} + \tau\omega v_l)(v_l, v_l) \\ &\qquad \leq \phi''_{\bar{u}_j^\nu}(\bar{u}_j^\nu)(v_l, v_l) + z_l^2 L_l \left(\|w_{l-1} - \bar{u}_j^\nu\|_{\infty,l} + \omega |z_l|\right) \end{aligned}$$

the upper bound

$$\begin{aligned} g''(\tau\omega) &= \mathcal{J}''(w_{l-1} + \tau\omega v_l)(v_l, v_l) + \phi''_{\bar{u}_j^\nu}(w_{l-1} + \tau\omega v_l)(v_l, v_l) \\ &\leq a_{\bar{u}_j^\nu}(v_l, v_l) + z_l^2 L_l \left(\|w_{l-1} - \bar{u}_j^\nu\|_{\infty,l} + \omega |z_l|\right) \end{aligned} \tag{4.22}$$

is obtained in a similar way. Inserting (4.21) and (4.22) in (4.20), it is clear that (4.17) implies (4.14) □

We emphasize that only *local properties* (i.e. properties on supp μ_l) enter the upper bound in (4.17).

As an alternative to local damping (4.11), one may formally set $\omega_l \equiv 1$ and enforce the monotonicity (3.2) by *global* damping

$$u_j^{\nu+1} = \bar{u}_j^\nu + \bar{\omega} \sum_{l=1}^{m_j} v_l, \qquad \bar{\omega} \in [0,1]. \tag{4.23}$$

However, upper bounds for $\bar{\omega}$ (cf. e.g. [1, 5]) typically deteriorate for increasing global Lipschitz constant

$$\bar{L} = \max_{p \in \bigcup_{l=1}^m \text{ int supp } \mu_l} L_p.$$

Hence, for heavily varying L_p, global damping (4.23) is likely to provide very little progress in comparison with the local strategy (4.11).

In order to avoid overflow in numerical computations, one may select the regular nodes $\mathcal{N}_j^\circ(\bar{u}_j^\nu)$ according to the more restrictive condition

$$\mathcal{N}_j^\circ(\bar{u}_j^\nu) = \{p \in \mathcal{N}_j \mid \bar{u}_j^\nu(p) > 0 \text{ and } L_p < L_{\max}\} \tag{4.24}$$

with some given threshold $L_{\max} > 0$. This modification does not affect the above considerations.

5. Standard Monotone Multigrid Methods

Assume that $\mathcal{T}_j$ is resulting from j refinements of an intentionally coarse triangulation $\mathcal{T}_0$. In this way, we obtain a sequence of triangulations $\mathcal{T}_0, \ldots, \mathcal{T}_j$ and corresponding nested finite element spaces $\mathcal{S}_0 \subset \cdots \subset \mathcal{S}_j$. Though the algorithms and convergence results to be presented can be easily generalized to the nonuniform grids, we assume for convenience that the triangulations are uniformly refined. More precisely, each triangle $t \in \mathcal{T}_k$ is subdivided in four congruent subtriangles in order to produce the next triangulation $\mathcal{T}_{k+1}$. Collecting all nodal basis functions from all refinement levels, we obtain the multilevel nodal basis $\Lambda_{\mathcal{S}}$,

$$\Lambda_{\mathcal{S}} = \left(\lambda_{p_1}^{(j)}, \lambda_{p_2}^{(j)} \ldots, \lambda_{p_{n_j}}^{(j)}, \ldots, \lambda_{p_1}^{(0)}, \ldots, \lambda_{p_{n_0}}^{(0)}\right), \tag{5.1}$$

with $m_{\mathcal{S}} = n_j + \cdots + n_0$ elements.

Using the abstract framework of the preceding section, we now specify the coarse grid correction $\mathcal{C}_j^{\text{std}}$ by selecting constant search directions

$$\mu_l^\nu = \lambda_l = \lambda_{p_l}^{(k_l)}, \qquad l = n_j + 1, \ldots, m_j = n_j + m_{\mathcal{S}}, \quad \nu \geq 0.$$

As usual, the ordering is taken from fine to coarse. The constraints $\mathcal{D}_l$, appearing in the local problems (4.12), are given by

$$\mathcal{D}_l = \{v \in V_l \mid \underline{\psi}_l \leq v \leq \overline{\psi}_l\}, \tag{5.2}$$

where $\underline{\psi}_l, \overline{\psi}_l \in V_l$ are obtained by *quasioptimal monotone restriction* as introduced in [10]. In this way, we end up with a *standard monotone multigrid method* (cf. [10, 13]) for the approximate solution of the linearized problem (4.10).

In the light of Proposition 4.1, we finally choose local damping parameters

$$\omega_l = \min\left\{1, \left\{\frac{2(|\ell_{\bar{u}_j^\nu}(\lambda_l) - a_{\bar{u}_j^\nu}(w_{l-1}, \lambda_l)| - L_l B_l^2)}{|z_l|(a_{\bar{u}_j^\nu}(\lambda_l, \lambda_l) + L_l(B_l + |z_l|))}\right\}_+\right\} \tag{5.3}$$

for non-zero local corrections $v_l = z_l \lambda_l$ obtained from (4.12). Denoting

$$\|v_k\|_\infty = \max_{x \in \Omega} |v_k(x)|,$$

the upper bounds

$$B_l = \sum_{k=1}^{l-1} \omega_k \|v_k\|_\infty \geq \|\bar{u}_j^\nu - w_{l-1}\|_{\infty,l} \tag{5.4}$$

are used to make ω_l computable without visiting the fine grid. As a consequence, $\mathcal{C}_j^{\text{std}}$ can be implemented as a classical V–cycle with optimal numerical complexity. Algorithmic details are reported in [12].

Monotone iterations of the form

$$\begin{aligned} \bar{u}_j^\nu &= \mathcal{M}_j u_j^\nu \\ u_j^{\nu+1} &= \mathcal{C}_j^{\mathrm{std}} \bar{u}_j^\nu \end{aligned} \tag{5.5}$$

are called *standard monotone multigrid methods with local damping.* It is clear from Theorem 3.1 and Proposition 4.1 that (5.5) is globally convergent, if the smoother $\mathcal{M}_j$ satisfies the conditions of Theorem 2.1. We will now derive upper bounds for the asymptotic convergence rates with respect to the local energy norm

$$\|v\|_{u_j} = a_{u_j}(v,v)^{1/2}$$

with $a_{u_j}(v,v)$ defined according to (4.9).

The following lemma will serve as a basis for the rest of the exposition.

Lemma 5.1. *Assume that the discrete minimization problem (1.8) satisfies the non–degeneracy condition*

$$\ell(\lambda_p^{(j)}) - a(u_j, \lambda_p^{(j)}) \in \operatorname{int} \partial\phi_j(u_j)(\lambda_p^{(j)}) \quad \forall p \in \mathcal{N}_j^\bullet(u_j) \tag{5.6}$$

and that exact Gauß–Seidel iteration (2.1) is used as smoother $\mathcal{M}_j$.
Then there is a $\nu_0 \geq 0$ such that

$$\mathcal{N}_j^\circ(u_j^\nu) = \mathcal{N}_j^\circ(\bar{u}_j^\nu) = \mathcal{N}_j^\circ(u_j) \qquad \forall \nu \geq \nu_0. \tag{5.7}$$

Proof. Proof. Note that

$$\mathcal{N}_j^\circ(u_j^\nu) = \mathcal{N}_j^\circ(\bar{u}_j^{\nu-1})$$

follows directly from (4.5). Hence, it is sufficient to show the second equality in (5.7). It is clear that

$$\mathcal{N}_j^\circ(u_j) \subset \mathcal{N}_j^\circ(\bar{u}_j^\nu)$$

holds for sufficiently large ν, because the convergence $\bar{u}_j^\nu \to u_j$ (see (3.3)) clearly implies $\bar{u}_j^\nu(p) > 0$ if $u_j(p) > 0$ and ν is large enough. It remains to show that asymptotically

$$\mathcal{N}_j^\bullet(u_j) \subset \mathcal{N}_j^\bullet(\bar{u}_j^\nu).$$

Let $p_l \in \mathcal{N}_j^\bullet(u_j)$ and compute $w_l^\nu = w_{l-1}^\nu + T_l w_{l-1}^\nu$ according to (2.2). Exploiting (5.6) and the convergence $w_l^\nu \to u_j$, we obtain

$$\ell(\lambda_{p_l}^{(j)}) - a(w_l^\nu, \lambda_{p_l}^{(j)}) \in \ \operatorname{int} \partial\Phi(0)\, h_{p_l}$$

for sufficiently large ν. On the other hand, we get from (2.1) that

$$\ell(\lambda_{p_l}^{(j)}) - a(w_l^\nu, \lambda_{p_l}^{(j)}) \in \partial\Phi(\bar{u}_j^\nu(p_l))\, h_{p_l}.$$

As $\partial\Phi$ is maximal monotone, these two inclusions imply $\bar{u}_j^\nu(p_l) = 0$. □

Note that (5.7) may be wrong, if inexact Gauß–Seidel smoothing is used.

In the following, we assume that the conditions in Lemma 5.1 are fulfilled and that ν is large enough. Denote

$$\mathcal{K}_j^* = \{v \in \mathcal{S}_j |\, \mathcal{N}_j^\circ(v) = \mathcal{N}_j^\circ(u_j),\ |v(p) - u_j(p)| < \tfrac{\varepsilon^*}{2},\ p \in \mathcal{N}_j^\circ(u_j)\},$$

with sufficiently small ε^*, say $\varepsilon^* = \frac{1}{2}\min_{p\in\mathcal{N}_j^\circ(u_j)} u_j(p)$. Then, utilizing (5.7) and the convergence of the intermediate iterates $w_l^\nu \to u_j$ (cf. (4.16)), we get

$$w_l^\nu \in \mathcal{K}_j^*, \qquad l = 1, \ldots, m_j. \tag{5.8}$$

In the light of (5.7) and $\bar{u}_j^\nu \to u_j$, we can choose pointwise obstacles $\underline{\varphi}_{\bar{u}_j^\nu}$, $\overline{\varphi}_{\bar{u}_j^\nu}$ in (4.3) such that, for all $p \in \mathcal{N}_j^\circ(\bar{u}_j^\nu)$, the estimate

$$0 < \underline{\varphi}^* \leq \underline{\varphi}_{\bar{u}_j^\nu}(p) \leq \bar{u}_j^\nu(p) - \varepsilon^* < \bar{u}_j^\nu(p) + \varepsilon^* \leq \overline{\varphi}_{\bar{u}_j^\nu}(p) \leq \overline{\varphi}^* \tag{5.9}$$

holds with fixed $\underline{\varphi}^*$, $\overline{\varphi}^* \in \mathbb{R}$. Hence, we can select pointwise Lipschitz constants L_p^ν such that

$$L_p^\nu \leq L^* \qquad \forall p \in \mathcal{N}_j^\circ(\bar{u}_j^\nu) \tag{5.10}$$

holds with L^* independent of ν. Together with (5.7) and $\bar{u}_j^\nu \to u_j$ (5.9) leads to

$$\mathcal{K}_j^* \subset \mathcal{K}_{\bar{u}_j^\nu}. \tag{5.11}$$

In particular, we have $u_j \in \mathcal{K}_{\bar{u}_j^\nu}$. Combining (5.8) with (5.11), we finally get

$$w_l^\nu \in \mathcal{K}_j^* \subset \mathcal{K}_{\bar{u}_j^\nu} \qquad l = 1, \dots, m_j. \tag{5.12}$$

In order to illustrate first consequences of this observation, note that the original discrete problem (1.8) can be rewritten as the *reduced smooth problem*

$$u_j \in \mathcal{S}_j^\circ : \qquad a(u_j, v) + \phi'_{u_j}(u_j)(v) = \ell(v) \qquad \forall v \in \mathcal{S}_j^\circ \tag{5.13}$$

with ϕ_{u_j} defined according to (4.7) and $\mathcal{S}_j^\circ$ given by

$$\mathcal{S}_j^\circ = \{v \in \mathcal{S}_j \mid v(p) = 0,\ p \in \mathcal{N}_j^\bullet(u_j)\} \subset \mathcal{S}_j.$$

It follows from (5.12) that the standard monotone multigrid method (5.5) asymptotically reduces to an iterative method for the reduced smooth problem (5.13). Moreover, the constrained minimization problem (4.8) turns out to be asymptotically equivalent to (5.13) so that, for large ν, the coarse grid correction $\mathcal{C}_j^{\mathrm{std}}$ becomes an inexact Newton method.

Let us now proceed with two further auxiliary results.

Lemma 5.2. *Assume the conditions in Lemma 5.1 are satisfied. Then, for each $\varepsilon > 0$ there is a $\nu_\varepsilon \geq 0$ such that*

$$\|u_j - \bar{u}_j^\nu\|_{u_j} \leq (1 + \varepsilon)\|u_j - u_j^\nu\|_{u_j} \qquad \forall \nu \geq \nu_\varepsilon. \tag{5.14}$$

Proof. Proof. Let ν be large enough to ensure (5.8). Choose arbitrary $p = p_l \in \mathcal{N}_j^\circ(u_j)$. Rewriting (2.1) in variational form, we obtain

$$a(w_l^\nu, v_l^\nu) + \Phi'(w_l^\nu(p))v_l^\nu(p)h_p = \ell(v_l^\nu)$$

with $v_l = T_l w_{l-1}^\nu$. Inserting $v = v_l^\nu$ in (1.11), we also get

$$a(u_j, v_l^\nu) + \Phi'(u_j(p))v_l^\nu(p)h_p = \ell(v_l^\nu).$$

Now the mean-value theorem gives

$$a(w_{l-1}^\nu - u_j, v_l^\nu) = -a(v_l^\nu, v_l^\nu) - \Phi''(\tilde{w}(p))(w_l^\nu(p) - u_j(p))v_l^\nu(p)h_p \tag{5.15}$$

denoting $\tilde{w}(p) = u_j(p) + \tau(w_l^\nu(p) - u_j(p))$ with suitable $\tau \in (0,1)$. Using (5.15), straightforward computation leads to

$$\|w_l^\nu - u_j\|_{u_j}^2 = a(w_l^\nu - u_j, w_l^\nu - u_j) + \phi''_{u_j}(u_j)(w_l^\nu - u_j, w_l^\nu - u_j) =$$
$$\|w_{l-1}^\nu - u_j\|_{u_j}^2 - \|v_l^\nu\|_{u_j}^2 + 2(\Phi''(u_j(p))) - \Phi''(\tilde{w}(p)))v_l^\nu(p)(w_l^\nu(p) - u_j(p))h_p.$$

As Φ'' is locally Lipschitz, we get

$$|\Phi''(u_j(p)) - \Phi''(\tilde{w}(p))| \leq L|w_l^\nu(p) - u_j(p)|$$

with suitable L independent of ν. We have shown

$$\|w_l^\nu - u_j\|_{u_j}^2 \le \|w_{l-1}^\nu - u_j\|_{u_j}^2 + 2L|v_l^\nu(p)|(w_l^\nu(p) - u_j(p))^2 h_p.$$

Now the assertion follows from the equivalence of norms on finite dimensional spaces. $\square$

Lemma 5.3. *Assume that the conditions in Lemma 5.1 are satisfied and that the pointwise obstacles $\underline{\varphi}_{\bar u_j^\nu}$, $\overline{\varphi}_{\bar u_j^\nu}$ and pointwise Lipschitz constants L_p are chosen such that (5.9) and (5.10) hold, respectively. Assume further that non-zero corrections $v_l^\nu = z_l^\nu \lambda_l$, $\lambda_l = \lambda_{p_l}^{(k_l)} \in \Lambda_{\mathcal S} \cap \mathcal S_j^\circ$, obtained from (4.12) have the property*

$$\|v_k^\nu\|_\infty^2 = o(\|v_l^\nu\|_\infty), \quad \nu \to \infty, \qquad k = 1, \dots l-1. \tag{5.16}$$

Then there is a ν_0 such that the damping parameters ω_l^ν defined in (5.3) reduce to

$$\omega_l^\nu = 1 \qquad \forall \nu \ge \nu_0.$$

Proof. Proof. Let ν_0 be large enough to guarantee (5.7), (5.9) and (5.10) for $\nu \ge \nu_0$. First note that L_l^ν then is uniformly bounded, say $L_l^\nu \le L_l^*$, for $\nu \ge \nu_0$, because $\lambda_l \in \mathcal S_j^\circ$.

Recall that the local constraints $\underline{\psi}_l^\nu$, $\overline{\psi}_l^\nu$ appearing in (5.2) are generated by successive quasioptimal restriction. Hence, utilizing (5.9) we can find $\psi_l^* \in V_l$, independent of ν, such that

$$\underline{\psi}_l^\nu(p_l) \le -\psi_l^*(p_l) < 0 < \psi_l^*(p_l) \le \overline{\psi}_l^\nu(p_l) \quad \forall \nu \ge \nu_0$$

holds for sufficiently large ν_0 (see [10]). As a consequence, the local problems (4.12) are reducing to variational equalities, if $|v_l^\nu(p_l)| < \psi_l^*(p_l)$, i.e. if ν is large enough. In this case, the solution $v_l^\nu = z_l^\nu \lambda_l$ of (4.12) is given by

$$z_l^\nu = \frac{\ell_{\bar u_j^\nu}(\lambda_l) - a_{\bar u_j^\nu}(w_{l-1}, \lambda_l)}{a_{\bar u_j^\nu}(\lambda_l, \lambda_l)}. \tag{5.17}$$

Let $z_l^\nu \ne 0$. Then the lower bound

$$\omega_l^\nu \ge \min\left\{1, \frac{2a_{u_j}(\lambda_l, \lambda_l)}{a_{u_j}(\lambda_l, \lambda_l) + L_l^* \sum_{k=1}^{l} \|v_k^\nu\|_\infty} \left(1 - \frac{(l-1)L_l^* \sum_{k=1}^{l-1} \|v_k^\nu\|_\infty^2}{|\ell_{u_j}(\lambda_l) - a_{u_j}(w_{l-1}, \lambda_l)|}\right)\right\}$$

follows directly from (5.7) and (5.17). Using again (5.17) and the equivalence of norms on finite dimensional spaces, we get

$$\frac{\sum_{k=1}^{l-1} \|v_k^\nu\|_\infty^2}{|\ell_{u_j}(\lambda_l) - a_{u_j}(w_{l-1}, \lambda_l)|} = \sum_{k=1}^{l-1} \frac{\|v_k^\nu\|_\infty^2}{\|v_l^\nu\|_{u_j}^2} |z_l^\nu| \le c \max_{k=1,\dots,l-1} \frac{\|v_k^\nu\|_\infty^2}{\|v_l^\nu\|_\infty}$$

with constant c independent of ν. Now the assertion follows from (5.16) and the convergence $v_k^\nu \to 0$, $\nu \to \infty$. $\square$

Now we are ready to state the main result of this section.

Theorem 5.4. *Assume that the conditions of Lemma 5.3 are satisfied. Let*

$$\|v\|_{u_j} \le \gamma_j \|v\| \qquad \forall v \in \mathcal S_j^\circ. \tag{5.18}$$

Then there is a $\nu_0 = \nu_0(j) \ge 0$ such that the iterates produced by the standard monotone multigrid method (5.5) fulfill the error estimate

$$\|u_j - u_j^{\nu+1}\|_{u_j} \le (1 - c\gamma_j^{-1}(j+1)^{-4})\|u_j - u_j^\nu\|_{u_j} \qquad \forall \nu \ge \nu_0. \tag{5.19}$$

with a constant $c < 1$ depending only on the ellipticity of $a(\cdot,\cdot)$ and on the initial triangulation $\mathcal{T}_0$.

Proof. Proof. For the moment choose ν_0 such that (5.7) is valid and the constrained smooth problem (4.8) is equivalent to the reduced smooth problem (5.13) for $\nu \geq \nu_0$. Consider the reduced linearized problem

$$u_j^* \in \mathcal{S}_j^\circ : \quad a_{u_j}(u_j^*, v) = \ell_{u_j}(v) \quad \forall v \in \mathcal{S}_j^\circ. \tag{5.20}$$

We subtract (5.13) from (5.20), use the mean-value theorem and insert $v = u_j^* - u_j$ to get the equality

$$\|u_j^* - u_j\|_{u_j}^2 + \phi''_{u_j}(w)(\bar{u}_j^\nu - u_j, u_j^* - u_j) - \phi''_{u_j}(\bar{u}_j^\nu)(\bar{u}_j^\nu - u_j, u_j^* - u_j) = 0$$

where $w \in \mathcal{S}_j^\circ$ is defined by $w(p) = \bar{u}_j^\nu(p) + \omega_p(u_j^*(p) - \bar{u}_j^\nu(p))$ with $\omega_p \in (0,1)$ for all $p \in \mathcal{N}_j$. Now the quadratic estimate

$$\|u_j^* - u_j\|_{u_j} \leq c\|\bar{u}_j^\nu - u_j\|_{u_j}^2 \tag{5.21}$$

with c independent of ν follows from (5.10) and from the equivalence of norms on finite dimensional spaces.

A first consequence of (5.21) is the asymptotic equivalence of the constrained linearized problem (4.12) with (5.20). Moreover, it follows from Lemma 5.3 and its proof, i.e. from (5.17), that $\mathcal{C}^{\mathrm{std}}$ asymptotically becomes a linear subspace correction method (cf. [14, 17, 18]) for the linear reduced problem (5.20). The subspaces $\mathcal{W}_k$, $k = 0, \ldots, j$, are given by

$$\mathcal{W}_k = \operatorname{span}\{\lambda_p^{(k)} \in \Lambda_{\mathcal{S}} \cap \mathcal{S}_j^\circ, p \in \mathcal{N}_k\}.$$

On all these subspaces the bilinear form $a_{u_j}(\cdot,\cdot)$ is approximated by the non-symmetric bilinear form $b_k(\cdot,\cdot)$ representing the standard Gauß-Seidel smoother. We now give an upper bound for the convergence rate of this linear iteration using a general result of Neuß [15]. To this end, we have to check three properties.

As

$$a_{u_j}(v,v) \leq c\sum_{i=1}^{n_k} a_{u_j}(\lambda_{p_i}^{(k)}, \lambda_{p_i}^{(k)}) v(p_i)^2 \quad \forall v \in \mathcal{W}_k$$

holds with a constant c depending only on the initial triangulation $\mathcal{T}_0$, we get the smoothing property

$$a_{u_j}(v,v) \leq \omega b_k(v,v) \quad \forall v \in \mathcal{W}_k \tag{5.22}$$

with some $\omega \in (0,2)$ depending only on $\mathcal{T}_0$.

Let $v \in \mathcal{S}_j^\circ$. Consider the splitting

$$v = \sum_{i=0}^{j} v_k, \qquad v_0 = I_0 v, \quad v_k = I_k v - I_{k-1} v,$$

induced by the modified interpolation operators I_k, defined by

$$(I_k v)(p) = \begin{cases} v(p) & \text{if } \lambda_p^{(k)} \in \mathcal{S}_j^\circ \\ 0 & \text{else} \end{cases}.$$

We want to show that

$$\sum_{k=0}^{j} b_k^{\mathrm{s}}(v_k, v_k) \leq C\gamma_j (j+1)^2 \|v\|_{u_j}^2 \tag{5.23}$$

holds with a constant C depending only on $\mathcal{T}_0$ and on the ellipticity of $a(\cdot,\cdot)$. The symmetric bilinear form $b_k^{\mathrm{s}}(\cdot,\cdot)$ stands for the symmetric Gauß-Seidel iteration on $\mathcal{W}_k$. Indeed, (5.23) is a consequence of the estimate

$$b_k^{\mathrm{s}}(v_k, v_k) \le c \sum_{i=1}^{n_k} a_{u_j}(\lambda_{p_i}^{(k)}, \lambda_{p_i}^{(k)}) v_k(p_i)^2$$

which holds for all $v_k \in \mathcal{W}_k$, of condition (5.18) and the results in section 5 of [14].

Finally,

$$a_{u_j}(v_l, v_k) \le \omega^{\frac{1}{2}} b_l(v_l, v_l)^{\frac{1}{2}} b_k^{\mathrm{s}}(v_k, v_k)^{\frac{1}{2}} \qquad \forall v_l \in \mathcal{W}_l, v_k \in \mathcal{W}_k \tag{5.24}$$

follows directly from the Cauchy-Schwarz inequality, (5.22) and the well-known smoothing property of $b_k^{\mathrm{s}}(\cdot,\cdot)$. Utilizing (5.22), (5.23) and (5.24), we now can apply Satz 2.3.16 in [15] in order to get the asymptotic error estimate

$$\|u_j^* - \mathcal{C}^{\mathrm{std}} \bar{u}_j^\nu\|_{u_j} \le (1 - c\gamma_j^{-1}(j+1)^{-4}) \|u_j^* - \bar{u}_j^\nu\|_{u_j} \quad \forall \nu \ge \nu_0 \tag{5.25}$$

with sufficiently large ν_0.

To conclude the proof, we combine the estimates (5.14), (5.21) and (5.25) by the triangle inequality in order to get the asymptotic error estimate (5.19). □

We emphasize that (5.19) describes the worst case and can be easily improved on suitable regularity assumptions. For example, let

$$\sup_{j=0,1,\dots} \max_{p \in \mathcal{N}_j^\circ(u_j)} \Phi''(u_j(p)) \le \text{const.} < \infty$$

and assume that the bilinear form $a(\cdot,\cdot)$ takes the form

$$a(v, w) = \int_\Omega \sum_{l,k=1}^{2} a_{lk}\ \partial_l v\ \partial_k w\ dx, \tag{5.26}$$

with coefficients $a_{lk} \in C^1(\bar{\Omega})$. Then, exploiting a sharpened Cauchy-Schwarz inequality instead of (5.24), we get the usual $\mathcal{O}(j^{-2})$-estimate for hierarchical bases. Further improvements can be made by using L^2-like projections instead of the modified interpolations I_k. We refer to [14, 16] for further information.

In our numerical computations as reported [12], we observed optimal convergence rates with respect to the usual energy norm induced by $a(\cdot,\cdot)$. A theoretical justification is subject of future research.

6. Truncated Monotone Multigrid Methods

The standard multigrid method relies on the condition that the coarse grid correction must not change the values of the smoothed iterate $\bar{u}_j^\nu$ at the critical nodes $p \in \mathcal{N}_j^\bullet(\bar{u}_j^\nu)$. Hence, only functions $\lambda_l \in \Lambda_{\mathcal{S}}$ with the property

$$\text{int supp } \lambda_l \cap \mathcal{N}_j^\bullet(\bar{u}_j^\nu) = \emptyset \tag{6.1}$$

actually contribute to the coarse grid correction. In comparison with the linear selfadjoint case, this leads to a poor representation of the low frequency parts of the error. In order to improve the convergence rates by improved coarse grid transport, we will now modify all $\lambda_l \in \Lambda_{\mathcal{S}}$ with the property (6.1) according to the actual guess of the free boundary.

Following [10, 11, 12, 13], we define the modified basis functions

$$\tilde{\lambda}_p^{(k)} = T_{j,k}^\nu \lambda_p^{(k)}, \quad p \in \mathcal{N}_k, \tag{6.2}$$

by using the truncation operators $T^{\nu}_{j,k}$, $k = 0, \ldots, j$,

$$T^{\nu}_{j,k} = I_{\mathcal{S}^{\nu}_j} \circ \cdots \circ I_{\mathcal{S}^{\nu}_k}. \tag{6.3}$$

Here $I_{\mathcal{S}^{\nu}_k} : \mathcal{S}_j \to \mathcal{S}^{\nu}_k$ denotes the $\mathcal{S}^{\nu}_k$–interpolation, and the spaces $\mathcal{S}^{\nu}_k \subset \mathcal{S}_k$,

$$\mathcal{S}^{\nu}_k = \{v \in \mathcal{S}_k \mid v(p) = 0,\ p \in \mathcal{N}^{\nu}_k\} \subset \mathcal{S}_k, \tag{6.4}$$

are the reduced subspaces with respect to $\mathcal{N}^{\nu}_k = \mathcal{N}_k \cap \mathcal{N}^{\bullet}_j(\bar{u}^{\nu}_j)$, $k = 0, \ldots, j$. Similar subspaces of $\mathcal{S}_j$ have been considered recently by other authors [2, 9, 14] in connection with the coarsening of a given mesh.

Replacing the multilevel nodal basis $\Lambda_{\mathcal{S}}$ by the actual truncation $\tilde{\Lambda}^{\nu}_{\mathcal{S}}$,

$$\tilde{\Lambda}^{\nu}_{\mathcal{S}} = \left(\lambda^{(j)}_{p_1}, \ldots, \lambda^{(j)}_{p_{n_j}}, \tilde{\lambda}^{(j-1)}_{p_1}, \ldots, \tilde{\lambda}^{(j-1)}_{p_{n_{j-1}}}, \ldots, \tilde{\lambda}^{(0)}_{p_1}, \ldots, \tilde{\lambda}^{(0)}_{p_{n_0}}\right), \quad \nu \geq 0,$$

we can now derive a *truncated* coarse grid correction $\mathcal{C}^{\text{trc}}$ by the same reasoning as described in the previous section. More precisely, all non-zero elements of $\tilde{\Lambda}^{\nu}_{\mathcal{S}}$ are now used as search directions

$$\mu^{\nu}_l = \tilde{\lambda}_l = \tilde{\lambda}^{(k_l)}_{p_l}, \qquad l = n_j + 1, \ldots, m_j = n_j + \tilde{m}_{\mathcal{S}}, \quad \nu \geq 0.$$

Local constraints $\mathcal{D}_l$, as appearing in (4.12), are obtained from slightly modified monotone restrictions (see [10, 12, 13]) and local damping parameters ω_l are given by a straightforward analogue of (5.3).

Monotone iterations of the form

$$\begin{aligned} \bar{u}^{\nu}_j &= \mathcal{M}_j u^{\nu}_j \\ u^{\nu+1}_j &= \mathcal{C}^{\text{trc}}_j \bar{u}^{\nu}_j \end{aligned} \tag{6.5}$$

are called *truncated monotone multigrid methods with local damping*. Again, the global convergence follows from Theorem 3.1 and Proposition 4.1.

Theorem 6.1. *Assume that the conditions of Lemma 5.1 are satisfied and that (5.18) holds. Assume further that non-zero corrections* $v^{\nu}_l = z^{\nu}_l \tilde{\lambda}_l$, $\tilde{\lambda}_l = \tilde{\lambda}^{(k_l)}_{p_l} \in \tilde{\Lambda}^{\nu}_{\mathcal{S}}$, *obtained from (4.12) have property (5.16).*

Then there is a $\nu_0 = \nu_0(j) \geq 0$ *such that the iterates produced by the truncated monotone multigrid method (6.5) fulfill the error estimate*

$$\|u_j - u^{\nu+1}_j\|_{u_j} \leq (1 - c\gamma_j^{-1}(j+1)^{-4})\|u_j - u^{\nu}_j\|_{u_j} \qquad \forall \nu \geq \nu_0. \tag{6.6}$$

with a constant $c < 1$ *depending only on the ellipticity of* $a(\cdot,\cdot)$ *and on the initial triangulation* $\mathcal{T}_0$.

Proof. Proof. The proof is essentially the same as for Theorem 5.4. We only have to establish a straightforward analogue of Lemma 5.3 and an error estimate of the form (5.25) for the reduced linear iteration. Note that (5.22) and (5.24) still hold if $\mathcal{W}_k$ is replaced by the larger space $\tilde{\mathcal{W}}_k$,

$$\tilde{\mathcal{W}}_k = \text{span}\{\tilde{\lambda}^{(k)}_p \in \tilde{\Lambda}^{\nu}_{\mathcal{S}}, p \in \mathcal{N}_k\}.$$

□

As functions $v \in \tilde{\mathcal{W}}_k$ in general do not satisfy a strengthened Cauchy-Schwarz inequality, further improvements of (6.6) are more difficult than in the standard case. Nevertheless, we found much faster convergence for the truncated version in our numerical experiments. We refer to [12, 13] for further information.

References

[1] R.E. Bank and D.J. Rose. Analysis of a multilevel iterative method for nonlinear finite element equations. *Math. Comp.*, 39:453–665, 1982.

[2] R.E. Bank and J. Xu. An algorithm for coarsening unstructured meshes. *Numer. Math.*, 73:1–36, 1996.

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

[4] F.A. Bornemann, B. Erdmann, and R. Kornhuber. A posteriori error estimates for elliptic problems in two and three space dimensions. *SIAM J. Numer. Anal.*, 33:1188–1204, 1996.

[5] P. Deuflhard and M. Weiser. Local inexact Newton multilevel FEM for nonlinear elliptic problems. Preprint SC 96-30, Konrad–Zuse–Zentrum (ZIB), Berlin, 1996.

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

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

[8] W. Hackbusch and A. Reusken. Analysis of a damped nonlinear multilevel method. *Numer. Math.*, 55:225–246, 1989.

[9] W. Hackbusch and S.A. Sauter. Composite finite elements for the approximation of PDEs on domains with complicated micro–structures. *Numer. Math.*, 75:447–472, 1997.

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

[11] R. Kornhuber. Monotone multigrid methods for elliptic variational inequalities II. *Numer. Math.*, 72:481 – 499, 1996.

[12] R. Kornhuber. On robust multigrid methods for non–smooth variational problems. Preprint SC 96-38, Konrad–Zuse–Zentrum (ZIB), Berlin, 1996.

[13] R. Kornhuber. *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems*. Teubner, Stuttgart, 1997.

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

[15] N. Neuß. Homogenisierung und Mehrgitter. Dissertation, Bericht N96/7, ICA Stuttgart, 1996.

[16] P. Oswald. Stable subspace splittings for Sobolev spaces and their applications. Forschungsergebnisse der Friedrich-Schiller-Universität Jena Math/93/7, Jena, 1993.

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

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