ON CONSTRAINED NEWTON LINEARIZATION AND MULTIGRID FOR VARIATIONAL INEQUALITIES

RALF KORNHUBER

Abstract. We consider the fast solution of a class of large, piecewise smooth minimization problems. 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 polylogarithmic upper bounds for the asymptotic convergence rates. Efficiency is illustrated by numerical experiments.

1. Introduction

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

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

on a closed subspace $H \subset H^1(\Omega)$. For simplicity, we concentrate on $H = H_0^1(\Omega)$ and $d = 2$. 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 $a(\cdot,\cdot)$ and by a linear functional $\ell \in H'$. H is equipped with the energy norm $\|\cdot\| = a(\cdot,\cdot)^{1/2}$. The functional ϕ,

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

is generated by a convex function $\Phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ with the properties

$$\begin{aligned} &\Phi(z) = \infty \quad \forall z < 0, \qquad \Phi(z) < \infty \quad \forall z \ge 0 \\ &|\Phi(z) - \Phi(z')| \le G(|z| + |z'|)|z - z'| \qquad \forall z, z' \ge 0 \end{aligned} \tag{1.4}$$

where G is some scalar, affine function and

$$\Phi \in C^2(0,\infty), \quad \Phi'' \text{ is uniformly Lipschitz on compact subsets of } (0,\infty). \tag{1.5}$$

As a consequence of (1.4), ϕ is convex, lower semi–continuous and proper. Hence, (1.1) admits a unique solution $u \in H$ (cf. [12], pp. 28). This property and all results to be presented can be generalized to functions Φ with a finite number of singularities.

1991 *Mathematics Subject Classification.* Primary 65N55, 65K10. Secondary 49M20, 49M15.

Key words and phrases. finite elements, multigrid methods, variational inequalities.

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.

Let $\mathcal{T}_j$ be a 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

$$u_j \in \mathcal{S}_j : \quad \mathcal{J}(u_j) + \phi_j(u_j) \leq \mathcal{J}(v) + \phi_j(v) \quad \forall v \in \mathcal{S}_j. \tag{1.6}$$

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

$$\phi_j(v) = \sum_{p \in \mathcal{N}_j} \Phi(v(p)) h_p, \qquad h_p = \int_\Omega \lambda_p^{(j)}(x)\, dx \tag{1.7}$$

where $\lambda_p^{(j)}$, $p \in \mathcal{N}_j$, denote the nodal basis functions of $\mathcal{S}_j$. The discrete minimization problem (1.6) is uniquely solvable and can be reformulated as the variational inequality

$$u_j \in \mathcal{S}_j : \; a(u_j, v - u_j) + \phi_j(v) - \phi_j(u_j) \geq \ell(v - u_j) \;\; \forall v \in \mathcal{S}_j \tag{1.8}$$

or as the variational inclusion

$$u_j \in \mathcal{S}_j : \quad \ell(v) - a(u_j, v) \in \partial\phi_j(u_j)(v) \quad \forall v \in \mathcal{S}_j \tag{1.9}$$

where $\partial\phi_j$ is the set–valued subdifferential of ϕ_j. Problems of the form (1.1) and related discretizations arise in a wide range of applications and have been studied for quite a while. For further information we refer e.g. to [6, 7, 12] and the bibliography cited therein.

Here, we shall concentrate on the fast solution of the discrete minimization problem (1.6). It is clear that Newton–multigrid methods [1, 4] cannot be applied, because the functional ϕ_j is not differentiable. Nonlinear multigrid techniques [8] as well as nonlinear subspace corrections in the spirit of [5, 17] also rely on the smoothness of the nonlinearity. A common remedy is to use such methods after some suitable regularization of Φ. Unfortunately, reasonable convergence speed may then have to be paid by unacceptable discretization errors and vice versa. Similar problems arise in case of static piecewise quadratic approximation of Φ as applied in [12], pp. 138.

In this paper, we extend *monotone multigrid methods* [10, 11, 12] from piecewise quadratic functions Φ to the piecewise smooth case (1.5). To this end, monotone multigrid methods are regarded as two-stage iterations consisting of a globally convergent fine grid smoother $\mathcal{M}_j$ and a coarse grid correction $\mathcal{C}_j$ preserving global convergence by preserving monotonically decreasing energy.

Gauß–Seidel type relaxation is used for fine grid smoothing. As exact solutions of the occurring scalar problems might be unavailable, we present an inexact variant including a stopping criterion for simple bisection. The basic idea for constructing $\mathcal{C}_j$ is to *constrain* coarse grid corrections to a neighborhood of the actual smoothed iterate where *Newton linearization* can be controlled by pointwise Lipschitz constants. There is no coarse grid correction at the singularity. Hence, no regularization is involved. In contrast to piecewise quadratic Φ, suitable damping of coarse grid correction is required in order to preserve monotonicity. We propose *local damping* of each local correction associated with a fixed node on a fixed refinement level. This strategy is especially suited for heavily varying local Lipschitz constants.

Within this general framework, we derive standard and truncated versions of monotone multigrid methods. Similar algorithms were presented in [13] without proofs. Global convergence proofs carry over from [11, 12]. Detailed convergence analysis clarifies the interplay of fine grid smoothing, constrained Newton linearization and local damping and provides polylogarithmic bounds of asymptotic convergence rates. The practical relevance of our asymptotic analysis is supported by numerical experiments with a stationary porous medium flow. Despite of intrinsic singularities of the problem, we observed similar efficiency as in the linear self–adjoint case.

2. Inexact Gauss-Seidel Relaxation

Nonlinear Gauß–Seidel relaxation [7, 12] for the iterative solution of (1.6) is based on the successive minimization of the discrete energy functional $\mathcal{J} + \phi_j$ in the direction of the nodal basis functions $\lambda_{p_l}^{(j)}$, $l = 1, \ldots, n_j = \#\mathcal{N}_j$. For given $w \in \mathcal{S}_j$, the local correction $T_l w \in V_l = \operatorname{span}\{\lambda_{p_l}^{(j)}\}$ in the direction of $\lambda_{p_l}^{(j)}$ is the unique solution of

$$\begin{aligned} T_l w \in V_l: \quad & \mathcal{J}(w + T_l w) + \Phi(w(p_l) + T_l w(p_l)) h_{p_l} \\ & \leq \mathcal{J}(w + v) + \Phi(w(p_l) + v(p_l)) h_{p_l} \quad \forall v \in V_l. \end{aligned} \tag{2.1}$$

In general, the exact solution $T_l w$ of the scalar problem (2.1) is not available. For this reason, we consider *inexact Gauß–Seidel relaxation* defined as follows.

For given iterate u_j^ν, we introduce 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 approximations $v_l^\nu \in V_l$ of $T_l w_{l-1}^\nu$. For example, v_l^ν might be resulting from some steps of an iterative solver as applied to (2.1). 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*

$$\mathcal{M}_j u_j^0 \in \mathcal{K}_j = \{v \in \mathcal{S}_j \mid v(p) \geq 0 \ \forall p \in \mathcal{N}_j\} \qquad \forall u_j^0 \in \mathcal{S}_j$$

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 relaxation (2.3) is globally convergent.

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

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

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

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

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

Observe that the estimate

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

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

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

and (2.7), we obtain

$$\begin{aligned}
\bar{\mathcal{J}}(u_j^{\nu_k}) - \bar{\mathcal{J}}(u_j^{\nu_{k+1}}) &\geq \bar{\mathcal{J}}(u_j^{\nu_k}) - \bar{\mathcal{J}}(u_j^{\nu_k+1}) \\
&= \sum_{i=1}^{n_j} \left(\bar{J}(w_{i-1}^{\nu_k}) - \bar{J}(w_i^{\nu_k})\right) \\
&\geq \sum_{i=1}^{n_j} \left(\bar{J}(w_{i-1}^{\nu_k}) - \bar{J}(w_{i-1}^{\nu_k} + \omega_0 T_i w_{i-1}^{\nu_k})\right) \\
&= \sum_{i=1}^{n_j} \big(\omega_0(\ell(T_i w_{i-1}^{\nu_k}) - a(w_{i-1}^{\nu_k} + T_i w_{i-1}^{\nu_k}, T_i w_{i-1}^{\nu_k})) \\
&\qquad\quad + \phi_j(w_{i-1}^{\nu_k}) - \phi_j(w_{i-1}^{\nu_k} + \omega_0 T_i w_{i-1}^{\nu_k})\big) \\
&\quad + \omega_0(1 - \tfrac{\omega_0}{2}) \sum_{i=1}^{n_j} \|T_i w_{i-1}^{\nu_k}\|^2 \\
&\geq \omega_0(1 - \tfrac{\omega_0}{2}) \sum_{i=1}^{n_j} \|T_i w_{i-1}^{\nu_k}\|^2.
\end{aligned} \tag{2.8}$$

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

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

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

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

The monotonicity (2.5) yields

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

for each fixed $l = 1, \ldots, n_j$. 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 $l = 1, \ldots, n_j$ so that u_j^* must be a fixed point of the original nonlinear Gauß–Seidel relaxation 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

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

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

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 subdifferential of Φ. We shall 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 $(w + v_l)(p_l) \geq 0$, 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

Nonlinear or inexact Gauß-Seidel relaxation $\mathcal{M}_j$, as considered in the preceding section, typically suffer from rapidly deteriorating convergence rates when proceeding to more and more refined triangulations. As a possible remedy, we introduce 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 the *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*

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

Then the monotone iteration (3.1) is globally convergent.

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

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

□

As a by-product, we also get 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 considerable flexibility in constructing $\mathcal{C}_j$.

4. Monotone Coarse Grid Correction with Local Damping

Recall that classical Newton multigrid methods cannot be applied to (1.6) for lack of smoothness. In this section, we shall derive *constrained* Newton multigrid methods to be used as coarse grid correction $\mathcal{C}_j$.

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

$$\mathcal{N}_j^\circ(\bar{u}_j^\nu) = \{p \in \mathcal{N}_j \mid \bar{u}_j^\nu(p) > 0\} \subset \mathcal{N}_j. \tag{4.1}$$

Consider some fixed $p \in \mathcal{N}_j^\circ(\bar{u}_j^\nu)$. Then, as a consequence of (1.5), there exists a neighborhood of $\bar{u}_j^\nu(p)$,

$$0 < \underline{\varphi}_{\bar{u}_j^\nu}(p) < \bar{u}_j^\nu(p) < \overline{\varphi}_{\bar{u}_j^\nu}(p), \tag{4.2}$$

where uniform Lipschitz continuity

$$|\Phi''(z_1) - \Phi''(z_2)| \leq L_p^\nu |z_1 - z_2| \quad \forall z_1, z_2 \in [\underline{\varphi}_{\bar{u}_j^\nu}(p), \overline{\varphi}_{\bar{u}_j^\nu}(p)] \tag{4.3}$$

holds with *pointwise Lipschitz constant* $L_p^\nu > 0$. For instance, let us choose

$$\underline{\varphi}_{\bar{u}_j^\nu}(p) = \tfrac{1}{2}\bar{u}_j^\nu(p), \quad \overline{\varphi}_{\bar{u}_j^\nu}(p) = 2\bar{u}_j^\nu(p) \qquad \forall p \in \mathcal{N}_j^\circ(\bar{u}_j^\nu). \tag{4.4}$$

We define

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

at the remaining *critical nodes*

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

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

$$\mathcal{K}_{\bar{u}_j^\nu} = \{w \in \mathcal{S}_j \mid \underline{\varphi}_{\bar{u}_j^\nu}(p) \leq w(p) \leq \overline{\varphi}_{\bar{u}_j^\nu}(p),\ p \in \mathcal{N}_j\} \subset \mathcal{S}_j. \tag{4.6}$$

The above definitions were motivated by the following local representation of ϕ_j,

$$\phi_j(w) = \phi_{\bar{u}_j^\nu}(w) + \text{const.} \qquad \forall w \in \mathcal{K}_{\bar{u}_j^\nu}, \tag{4.7}$$

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

$$\phi_{\bar{u}_j^\nu}(w) = \sum_{p \in \mathcal{N}_j^\circ(\bar{u}_j^\nu)} \Phi(w(p)) h_p, \qquad w \in \mathcal{K}_{\bar{u}_j^\nu}. \tag{4.8}$$

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

$$u_{\bar{u}_j^\nu} \in \mathcal{K}_{\bar{u}_j^\nu}: \quad \mathcal{J}(u_{\bar{u}_j^\nu}) + \phi_{\bar{u}_j^\nu}(u_{\bar{u}_j^\nu}) \leq \mathcal{J}(v) + \phi_{\bar{u}_j^\nu}(v) \quad \forall v \in \mathcal{K}_{\bar{u}_j^\nu}. \tag{4.9}$$

As a consequence of (3.3), we have $\operatorname{dist}(u_j, \mathcal{K}_{\bar{u}_j^\nu}) \to 0$ as $\nu \to \infty$. Hence, the solutions of (4.9) tend to u_j. Moreover, we shall see later on that $u_j \in \mathcal{K}_{\bar{u}_j^\nu}$ holds for non–degenerate problems (1.6) after a finite number of iteration steps. In this case, we clearly have $u_{\bar{u}_j^\nu} = u_j$ or, equivalently, our original non–smooth problem (1.6) reduces to the constrained smooth problem (4.9). Hence, approximate solutions of (4.9) are good candidates for the next iterate $u_j^{\nu+1}$.

The main advantage of (4.9) 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

$$a_{\bar u_j^\nu}(w,w) = a(w,w) + \phi''_{\bar u_j^\nu}(\bar u_j^\nu)(w,w) \tag{4.10}$$

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 obtained by 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 resulting quadratic obstacle problem

$$w_{\bar u_j^\nu} \in \mathcal{K}_{\bar u_j^\nu} : \quad \mathcal{J}_{\bar u_j^\nu}(w_{\bar u_j^\nu}) \le \mathcal{J}_{\bar u_j^\nu}(v) \quad \forall v \in \mathcal{K}_{\bar u_j^\nu} \tag{4.11}$$

can be regarded as *constrained Newton linearization* reflecting that Φ is only piecewise differentiable.

We approximate (4.11) by one step of an *extended underrelaxation* as introduced in [10]. In contrast to [10], local damping parameters now have to be computed explicitly to enforce monotonicity (3.2) of the functional $\mathcal{J} + \phi_j$ which might be different from $\mathcal{J}_{\bar u_j^\nu}$. Hence, we briefly recall the basic algorithm for further reference and analysis. We choose scaled search directions μ_l^ν,

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

which may depend on the actual constraints $\mathcal{K}_{\bar u_j^\nu}$. It is convenient to start numeration at $n_j + 1$, because intermediate iterates w_l^ν, $l = 1, \dots, n_j$, are already given by (2.2). We now continue this sequence according to

$$w_{n_j}^\nu = \bar u_j^\nu, \qquad w_l^\nu = w_{l-1}^\nu + \omega_l^\nu v_l^\nu, \quad l = n_j + 1, \dots, m_j^\nu. \tag{4.12}$$

Each local correction v_l^ν is the solution of the local obstacle problem

$$v_l^\nu \in \mathcal{D}_l^\nu : \quad \mathcal{J}_{\bar u_j^\nu}(w_{l-1}^\nu + v_l^\nu) \le \mathcal{J}_{\bar u_j^\nu}(w_{l-1}^\nu + v) \quad \forall v \in \mathcal{D}_l^\nu \tag{4.13}$$

with constraints $\mathcal{D}_l^\nu \subset V_l^\nu := \operatorname{span}\{\mu_l^\nu\}$ satisfying

$$0 \in \mathcal{D}_l^\nu \subset \{v \in V_l^\nu \mid w_{l-1}^\nu + v \in \mathcal{K}_{\bar u_j^\nu}\}. \tag{4.14}$$

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

$$\mathcal{J}(w_l^\nu) + \phi_{\bar u_j^\nu}(w_l^\nu) \le \mathcal{J}(w_{l-1}^\nu) + \phi_{\bar u_j^\nu}(w_{l-1}^\nu). \tag{4.15}$$

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

$$\mathcal{C}_j \bar u_j^\nu = w_{m_j^\nu}^\nu = \bar u_j^\nu + \sum_{l=n_j+1}^{m_j^\nu} \omega_l^\nu v_l^\nu. \tag{4.16}$$

Using the monotonicity (4.15), general convergence results on extended underrelaxations carry over to the present case. For example, we get convergence of each infinite sequence of intermediate iterates (cf. [12], Corollary 2.3, p. 54)

$$w_l^\nu \to u_j \qquad \nu \to \infty. \tag{4.17}$$

We now derive a sufficient condition for the local monotonicity (4.15). Again, the index ν will be frequently suppressed. We shall use the notation $z_+ = \max\{0, z\}_+$.

Proposition 4.1. *Let $v_l = z_l \mu_l$ be the solution of (4.13). Assume that $\omega_l \in [0,1]$ satisfies*

$$\omega_l |z_l| \leq 2 \left\{ \frac{|\ell_{\bar{u}_j^\nu}(\mu_l) - a_{\bar{u}_j^\nu}(w_{l-1}, \mu_l)| - L_l \|\bar{u}_j^\nu - w_{l-1}\|_{\infty,l}^2}{a_{\bar{u}_j^\nu}(\mu_l, \mu_l) + L_l \left(\|\bar{u}_j^\nu - w_{l-1}\|_{\infty,l} + \omega_l |z_l|\right)} \right\}_+ \tag{4.18}$$

with local Lipschitz constant

$$L_l = \sum_{p \in \mathcal{N}_j^\circ(\bar{u}_j^\nu)} L_p |\mu_l(p)|\, h_p \tag{4.19}$$

and local maximum norm

$$\|v\|_{\infty,l} = \max_{p \in \mathcal{N}_j \cap \text{ int supp } \mu_l} |v(p)|. \tag{4.20}$$

Then the damped correction $\omega_l v_l$ fulfills the local monotonicity condition (4.15).

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, (4.15) 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

$$0 \leq \omega_l \leq -2 \frac{g'(0)}{g''(\tau \omega_l)} \tag{4.21}$$

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 (4.3). 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 (4.13). Combining these estimates, we get the lower bound

$$\begin{aligned} -g'(0) &= -\mathcal{J}'(w_{l-1})(v_l) - \phi'_{\bar{u}_j^\nu}(w_{l-1})(v_l) \\ &\geq |\ell_{\bar{u}_j^\nu}(v_l) - a_{\bar{u}_j^\nu}(w_{l-1}, v_l)| - L_l |z_l| \|w_{l-1} - \bar{u}_j^\nu\|_{\infty,l}^2. \end{aligned} \tag{4.22}$$

Using

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

the upper bound

$$\begin{aligned} g''(\tau\omega_l) &= \mathcal{J}''(w_{l-1} + \tau\omega_l v_l)(v_l, v_l) + \phi''_{\bar{u}_j^\nu}(w_{l-1} + \tau\omega_l v_l)(v_l, v_l) \\ &\leq a_{\bar{u}_j^\nu}(v_l, v_l) + z_l^2 L_l \left(\|w_{l-1} - \bar{u}_j^\nu\|_{\infty,l} + \omega_l |z_l|\right) \end{aligned} \tag{4.23}$$

is obtained in a similar way. Inserting (4.22) and (4.23) in (4.21), it is clear that (4.18) implies (4.15) □

We emphasize that only *local properties* (i.e. properties on supp μ_l) enter the upper bound in (4.18).

As an alternative to local damping, one might always set $\omega_l = 1$ in (4.12) and enforce monotonicity (3.2) by *global* damping

$$u_j^{\nu+1} = \bar{u}_j^\nu + \bar{\omega} \sum_{l=n_j+1}^{m_j} v_l \tag{4.24}$$

with suitable $\bar{\omega} \in [0,1]$. This would simplify convergence analysis, because e.g. the results from [10] could be applied directly. However, upper bounds for $\bar{\omega}$ (cf. e.g. [1, 4]) typically deteriorate for increasing *global* Lipschitz constant

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

Hence, for heavily varying L_p as considered here, global damping (4.24) is likely to provide very little progress in comparison with the local strategy (4.12).

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 $\mathcal{T}_0, \dots, \mathcal{T}_j$ 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 nonuniform grids, we assume for convenience that the triangulations are uniformly refined. More precisely, each triangle $t \in \mathcal{T}_k$ is subdivided into 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 $\Lambda_{\mathcal{S}}$,

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

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

$$\lambda_l = \lambda_{p_l}^{(k_l)}, \qquad l = n_j + 1, \dots, m_j = n_j + m_{\mathcal{S}},$$

are ordered from fine to coarse.

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

$$\mu_l^\nu = \lambda_l, \qquad l = n_j + 1, \dots, m_j, \qquad \forall \nu \geq 0.$$

For each $\bar{u}_j^\nu$ the admissible set $\mathcal{K}_{\bar{u}_j^\nu}$ is chosen according to (4.6) with $\underline{\varphi}_{\bar{u}_j^\nu}$, $\overline{\varphi}_{\bar{u}_j^\nu}$ taken from (4.4). The constraints $\mathcal{D}_l^\nu$, appearing in the local problems (4.13), take the form

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

where local obstacles $\underline{\psi}_l^\nu, \overline{\psi}_l^\nu \in V_l$ are intended to approximate the fine grid constraints $\underline{\varphi}_{\bar{u}_j^\nu} - w_{l-1}^\nu$, $\overline{\varphi}_{\bar{u}_j^\nu} - w_{l-1}^\nu$, respectively. The property $\underline{\psi}_l^\nu, \overline{\psi}_l^\nu \in V_l$ allows to check the constraints directly on the coarse grid. In order to guarantee (4.15), we impose the condition

$$\underline{\varphi}_{\bar{u}_j^\nu}(p) - w_{l-1}^\nu(p) \leq \underline{\psi}_l^\nu(p) \leq 0 \leq \overline{\psi}_l^\nu(p) \leq \underline{\varphi}_{\bar{u}_j^\nu}(p) - w_{l-1}^\nu(p) \quad \forall p \in \mathcal{N}_j. \tag{5.3}$$

Finally, we assume that

$$\underline{\psi}_l^\nu = \underline{\psi}_l(\underline{\varphi}_{\bar{u}_j^\nu}, w_{n_j}^\nu, \dots, w_{l-1}^\nu), \qquad \overline{\psi}_l^\nu = \overline{\psi}_l(\overline{\varphi}_{\bar{u}_j^\nu}, w_{n_j}^\nu, \dots, w_{l-1}^\nu)$$

are continuous functions of $\underline{\varphi}_{\bar u_j^\nu}, \overline{\varphi}_{\bar u_j^\nu}, w_{n_j}^\nu, \dots, w_{l-1}^\nu$, satisfying

$$(5.4) \qquad \underline{\psi}_l(\underline{\varphi}_{u_j}, u_j, \dots, u_j)(p) < 0 < \overline{\psi}_l(\overline{\varphi}_{u_j}, u_j, \dots, u_j)(p) \quad \forall p \in \text{ int supp } \lambda_l,$$

if int supp $\lambda_l \subset \mathcal{N}_j^\circ(u_j)$. Local obstacles $\underline{\psi}_l^\nu, \overline{\psi}_l^\nu \in V_l$ with the properties (5.3) and (5.4) can be obtained inductively by *quasioptimal monotone restriction.* We refer to [10] or [12], pp. 74, for details. As usual, the index ν is mostly skipped in the sequel.

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

$$(5.5) \qquad \omega_l = \min\left\{1, \left\{\frac{2(|\ell_{\bar u_j^\nu}(\lambda_l) - a_{\bar u_j^\nu}(w_{l-1}, \lambda_l)| - L_l B_l^2)}{|z_l|(a_{\bar u_j^\nu}(\lambda_l, \lambda_l) + L_l(B_l + |z_l|))}\right\}_+\right\}$$

for all non–zero local corrections $v_l = z_l \lambda_l$ obtained from (4.13). Denoting

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

the upper bounds

$$(5.6) \qquad B_l = \sum_{k=n_j+1}^{l-1} \omega_k \|v_k\|_\infty \geq \|\bar u_j^\nu - w_{l-1}\|_{\infty, l}$$

make ω_l computable without visiting the fine grid (cf. [13]).

As a consequence of the above considerations, the resulting coarse grid correction $\mathcal{C}_j^{\text{std}}$ can be implemented as a classical V–cycle with optimal numerical complexity. For further reference, the monotone iteration

$$(5.7) \qquad \begin{aligned} \bar u_j^\nu &= \mathcal{M}_j u_j^\nu \\ u_j^{\nu+1} &= \mathcal{C}_j^{\text{std}} \bar u_j^\nu \end{aligned}$$

is called *standard monotone multigrid method.*

It is clear from Theorem 3.1 and Proposition 4.1 that (5.7) is globally convergent, if the smoother $\mathcal{M}_j$ satisfies the conditions of Theorem 2.1. We shall now derive upper bounds for the asymptotic convergence rates with respect to the local energy norm

$$(5.8) \qquad \|v\|_{u_j} = a_{u_j}(v, v)^{1/2}.$$

The symmetric, positive definite bilinear form $a_{u_j}(v, v)$ is defined according to (4.10). We first state that the discrete free boundary is detected after a finite number of steps.

Lemma 5.1. *Assume that the discrete minimization problem (1.6) satisfies the non–degeneracy condition*

$$(5.9) \qquad \ell(\lambda_p^{(j)}) - a(u_j, \lambda_p^{(j)}) \in \text{int } \partial\phi_j(u_j)(\lambda_p^{(j)}) \quad \forall p \in \mathcal{N}_j^\bullet(u_j)$$

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

$$(5.10) \qquad \mathcal{N}_j^\circ(u_j^\nu) = \mathcal{N}_j^\circ(\bar u_j^\nu) = \mathcal{N}_j^\circ(u_j) \qquad \forall \nu \geq \nu_0.$$

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 (5.10). Recall that $\bar{u}_j^\nu \to u_j$ (cf. (3.3)). Hence, we have $\bar{u}_j^\nu(p) > 0$, if $u_j(p) > 0$ and ν is large enough. This implies

$$\mathcal{N}_j^\circ(u_j) \subset \mathcal{N}_j^\circ(\bar{u}_j^\nu) \tag{5.11}$$

for sufficiently large ν. It remains to show

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

Let $p_l \in \mathcal{N}_j^\bullet(u_j)$ or, equivalently, $u_j(p) = 0$. Rewriting (2.1) as a variational inclusion, we get

$$\ell(v) - a(w_l^\nu, v) \in \partial\Phi(w_l^\nu(p_l))v(p)\, h_{p_l} \qquad \forall v \in V_l = \operatorname{span}\{\lambda_{p_l}^{(j)}\}. \tag{5.12}$$

As $w_l^\nu(p_l) = \bar{u}_j^\nu(p_l)$ for all $l = 1, \ldots, n_j$, this leads to

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

Recall that $w_l^\nu \to u_j$ (cf. (4.17)). Hence, (5.9) yields

$$\ell(\lambda_{p_l}^{(j)}) - a(w_l^\nu, \lambda_{p_l}^{(j)}) \in \partial\Phi(0)\, h_{p_l} \tag{5.13}$$

for sufficiently large ν. As $\partial\Phi$ is maximal monotone, these two inclusions imply $\bar{u}_j^\nu(p_l) = 0$. We finally chose ν_0 such that (5.11) and (5.13) are satisfied. ☐

Note that Lemma 5.1 can not be extended to inexact Gauß–Seidel smoothing.

We continue with an asymptotic error estimate for nonlinear Gauß–Seidel relaxation. Note that known results on nonlinear subspace correction methods (cf. [5, 17]) cannot be applied, because Φ is not uniformly Lipschitz.

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

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

Proof. Choose $l = 1, \ldots, n_j$ and arbitrary $p_l \in \mathcal{N}_j^\circ(u_j)$. Using (5.10), the minimization problem (2.1) for the correction $v_l^\nu = T_l w_{l-1}^\nu$ can be equivalently rewritten as

$$\ell(v) - a(w_{l-1}^\nu + v_l^\nu, v) = \Phi'(w_{l-1}^\nu(p_l) + v_l^\nu(p_l))v(p_l)\, h_{p_l} \qquad \forall v \in V_l.$$

Observe that $w_l^\nu(p_l) = w_{l-1}^\nu(p_l) + v_l^\nu(p_l) = \bar{u}_j^\nu(p_l)$, $l = 1, \ldots, n_j$. Inserting $v = v_l^\nu$ we get

$$a(w_{l-1}^\nu, v_l^\nu) = \ell(v_l^\nu) - a(v_l^\nu, v_l^\nu) - \Phi'(\bar{u}_j^\nu(p_l))v_l^\nu(p_l)\, h_{p_l}.$$

On the other hand, use (1.9) with $v = v_l^\nu$ to obtain

$$a(u_j, v_l^\nu) = \ell(v_l^\nu) - \Phi'(u_j(p_l))v_l^\nu(p_l)\, h_{p_l}.$$

Now the mean-value theorem gives

$$a(w_{l-1}^\nu - u_j, v_l^\nu) = -a(v_l^\nu, v_l^\nu) - \Phi''(\tilde{w}(p_l))(\bar{u}_j^\nu(p_l) - u_j(p_l))v_l^\nu(p)\, h_{p_l} \tag{5.15}$$

denoting $\tilde w(p_l) = u_j(p_l) + \tau(\bar u_j^\nu(p_l) - u_j(p_l))$ with suitable $\tau \in (0,1)$. Using (5.15) and again $w_{l-1}^\nu(p_l) + v_l^\nu(p_l) = \bar u_j^\nu(p_l)$, we compute

$$
\begin{aligned}
\|w_l^\nu - u_j\|_{u_j}^2 &= \|w_{l-1}^\nu + v_l^\nu - u_j\|_{u_j}^2 \\
&= \|w_{l-1}^\nu - u_j\|_{u_j}^2 + \|v_l^\nu\|_{u_j}^2 \\
&\qquad +2(a(w_{l-1}^\nu - u_j, v_l^\nu) + \phi''_{u_j}(u_j)(w_{l-1}^\nu - u_j, v_l^\nu)) \\
&= \|w_{l-1}^\nu - u_j\|_{u_j}^2 + \|v_l^\nu\|_{u_j}^2 \\
&\qquad +2\left(-a(v_l^\nu, v_l^\nu) - \Phi''(u_j(p_l)) v_l^\nu(p_l)^2\, h_{p_l}\right. \\
&\qquad\quad +\Phi''(u_j(p_l))(v_l^\nu(p_l)^2 + v_l^\nu(p_l)(w_{l-1}^\nu(p_l) - u_j(p_l)))\, h_{p_l} \\
&\qquad\qquad \left. -\, \Phi''(\tilde w(p_l)) v_l^\nu(p_l)(\bar u_j^\nu(p_l) - u_j(p_l))\, h_{p_l}\right) \\
&= \|w_{l-1}^\nu - u_j\|_{u_j}^2 - \|v_l^\nu\|_{u_j}^2 \\
&\qquad +2\left(\Phi''(u_j(p_l)) - \Phi''(\tilde w(p_l))\right)(\bar u_j^\nu(p_l) - u_j(p_l)) v_l^\nu(p_l)\, h_{p_l}.
\end{aligned}
$$

We now derive an upper bound for $|\Phi''(u_j(p_l)) - \Phi''(\tilde w(p_l))|$. Convergence (3.3) provides

$$
\tfrac14 u_j(p) - |\bar u_j^\nu(p) - u_j(p)| \ge 0 \qquad \forall p \in \mathcal{N}_j^\circ(u_j) \tag{5.16}
$$

for sufficiently large ν. Choosing

$$
\underline{\varphi} = \tfrac12 \min_{p\in\mathcal{N}_j^\circ(u_j)} u_j(p), \qquad \overline{\varphi} = 2 \max_{p\in\mathcal{N}_j^\circ(u_j)} u_j(p), \tag{5.17}
$$

(5.16) implies

$$
\bar u_j^\nu(p) \in [\underline{\varphi}, \overline{\varphi}] \subset (0,\infty) \qquad \forall p \in \mathcal{N}_j^\circ(u_j). \tag{5.18}
$$

Hence, assumption (1.5) yields

$$
|\Phi''(u_j(p)) - \Phi''(\tilde w(p))| \le L^* |\bar u_j^\nu(p) - u_j(p)| \qquad \forall p \in \mathcal{N}_j^\circ(u_j)
$$

with suitable L^* independent of $l = 1, \dots, n_j$ and ν. Inserting this estimate in the above representation of $\|w_l^\nu - u_j\|_{u_j}^2$, we get

$$
\|w_l^\nu - u_j\|_{u_j}^2 \le \|w_{l-1}^\nu - u_j\|_{u_j}^2 + 2L^* |v_l^\nu(p_l)| (\bar u_j^\nu(p_l) - u_j(p_l))^2 h_{p_l}.
$$

Successive application gives

$$
\|\bar u_j^\nu - u_j\|_{u_j}^2 \le \|u_j^\nu - u_j\|_{u_j}^2 + 2L^* \|\bar u_j^\nu - u_j^\nu\|_\infty \sum_{l=1}^{n_j} (\bar u_j^\nu(p_l) - u_j(p_l))^2 h_{p_l}.
$$

It is well–known that shape regularity of $\mathcal{T}_j$, Poincaré's inequality, ellipticity of $a(\cdot,\cdot)$ and convexity of Φ provide

$$
\sum_{l=1}^{n_j} v(p)^2 h_{p_l} \le c \int_\Omega v(x)^2\, dx \le C\|v\|^2 \le C\|v\|_{u_j}^2 \qquad \forall v \in \mathcal{S}_j \tag{5.19}
$$

with suitable $c, C \in \mathbb{R}$ independent of ν and j. The last two estimates imply

$$
(1 - 2CL^* \|\bar u_j^\nu - u_j^\nu\|_\infty)\, \|\bar u_j^\nu - u_j\|_{u_j}^2 \le \|u_j^\nu - u_j\|_{u_j}^2.
$$

We finally chose ν_ε such that (5.10), (5.18) and

$$
1 - 2CL^* \|\bar u_j^\nu - u_j^\nu\|_\infty \ge (1+\varepsilon)^{-1} \tag{5.20}
$$

are valid for all $\nu \ge \nu_\varepsilon$. □

We now prove that the coarse grid correction $\mathcal{C}_j$ is asymptotically based on a *smooth* nonlinear problem.

Lemma 5.3. *Assume that the conditions in Lemma 5.1 are satisfied. Then there is a ν_0 such that for $\nu \geq \nu_0$ the constrained smooth problem (4.9) has the solution u_j and can be equivalently rewritten as*

$$u_j \in U_j^\circ : \qquad a(u_j, v) + \phi'_{u_j}(u_j)(v) = \ell(v) \qquad \forall v \in \mathcal{S}_j^\circ \tag{5.21}$$

with ϕ_{u_j} defined according to (4.8),

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

and $U_j^\circ = \{v \in \mathcal{S}_j^\circ \mid v(p) > 0 \ \forall p \in \mathcal{N}_j^\circ(u_j)\} \subset \mathcal{S}_j^\circ$.

Proof. We first show that there is ν_0 such that

$$u_j \subset \mathcal{K}_{\bar{u}_j^\nu} \qquad \forall \nu \geq \nu_0. \tag{5.22}$$

Using (5.10), we only have to prove that

$$\underline{\varphi}_{\bar{u}_j^\nu}(p) = \tfrac{1}{2}\bar{u}_j^\nu(p) \leq u_j(p) \leq 2\bar{u}_j^\nu(p) = \overline{\varphi}_{\bar{u}_j^\nu}(p) \qquad \forall p \in \mathcal{N}_j^\circ(u_j)$$

which is an immediate consequence of (5.16). Exploiting Lemma 5.1 and the convergence of $\bar{u}_j^\nu$ (cf. (3.3)), we finally choose ν_0 such that (5.10) and (5.16) are satisfied for $\nu \geq \nu_0$.

By definition, u_j is minimizing $\mathcal{J} + \phi_j$ on the whole space $\mathcal{S}_j$. As $u_j \in \mathcal{K}_{\bar{u}_j^\nu}$ and $\phi_{\bar{u}_j^\nu} = \phi_j \ \forall \nu \geq \nu_0$, u_j must be the solution of (4.9) for all $\nu \geq \nu_0$. The formulation (5.21) follows immediately from (1.9). □

As a consequence of Lemma 5.3, *exact* solution of the constrained smooth problem (4.9) in each iteration step would provide an *asymptotically exact* method.

Lemma 5.4. *Assume that the conditions in Lemma 5.1 are satisfied. Then there is a ν_0 such that constrained Newton linearization (4.11) is equivalent to classical Newton linearization of (5.21) at $\bar{u}_j^\nu$*

$$w_{\bar{u}_j^\nu} \in \mathcal{S}_j^\circ : \quad a_{\bar{u}_j^\nu}(w_{\bar{u}_j^\nu}, v) = \ell_{\bar{u}_j^\nu}(v) \quad \forall v \in \mathcal{S}_j^\circ. \tag{5.23}$$

Moreover, for each $\varepsilon > 0$ there is a $\nu_\varepsilon \geq \nu_0$ such that

$$\|w_{\bar{u}_j^\nu} - u_j\|_{u_j} \leq \varepsilon \|\bar{u}_j^\nu - u_j\|_{u_j} \qquad \forall \nu \geq \nu_\varepsilon. \tag{5.24}$$

Proof. Note that (5.10) yields $\bar{u}_j^\nu \in U_j^\circ$ so that (5.23) is well-defined for sufficiently large ν. For the moment, let w^ν denote the solution of (5.23). We first prove that for given $\varepsilon > 0$ there is a ν_ε such that

$$\|w^\nu - u_j\|_{u_j} \leq \varepsilon \|\bar{u}_j^\nu - u_j\|_{u_j} \qquad \forall \nu \geq \nu_\varepsilon. \tag{5.25}$$

We subtract (5.21) from (5.23), use the mean-value theorem and insert $v = w^\nu - u_j$ to get the equality

$$\|w^\nu - u_j\|_{u_j}^2 + \phi''_{u_j}(\tilde{w})(\bar{u}_j^\nu - u_j, w^\nu - u_j) - \phi''_{u_j}(\bar{u}_j^\nu)(\bar{u}_j^\nu - u_j, w^\nu - u_j) = 0 \tag{5.26}$$

where $\tilde{w} \in U_j^\circ \subset \mathcal{S}_j^\circ$ is given by the nodal values $\tilde{w}(p) = \bar{u}_j^\nu(p) + \tau_p(u_j(p) - \bar{u}_j^\nu(p))$ with suitable $\tau_p \in (0,1)$. Using (5.16), we get

$$|\Phi''(\tilde{w}(p)) - \Phi''(\bar{u}_j^\nu(p))| \leq L^* |\bar{u}_j^\nu(p) - u_j(p)| \qquad \forall p \in \mathcal{N}_j^\circ(u_j)$$

in the same way as in the proof of Lemma 5.2. Together with the Cauchy-Schwarz inequality and (5.19) this leads to

$$
\begin{aligned}
&|\phi''_{u_j}(\tilde{w})(\bar{u}_j^\nu - u_j, w^\nu - u_j) - \phi''_{u_j}(\bar{u}_j^\nu)(\bar{u}_j^\nu - u_j, w^\nu - u_j)| \\
&\leq L^* \sum_{p \in \mathcal{N}_j^\circ(u_j)} (\bar{u}_j^\nu(p) - u_j(p))^2 |w^\nu(p) - u_j(p)|\ h_p \qquad (5.27)\\
&\leq CL^* \|\bar{u}_j^\nu - u_j\|_\infty \|\bar{u}_j^\nu - u_j\|_{\bar{u}_j^\nu} \|w^\nu - u_j\|_{\bar{u}_j^\nu}
\end{aligned}
$$

Inserting this estimate in (5.26), we get

$$
\|w^\nu - u_j\|_{\bar{u}_j^\nu} \leq CL^* \|\bar{u}_j^\nu - u_j\|_\infty \|\bar{u}_j^\nu - u_j\|_{\bar{u}_j^\nu}. \qquad (5.28)
$$

We finally choose ν_ε such that (5.10) (5.16) and $CL^*\|\bar{u}_j^\nu - u_j\|_\infty \leq \varepsilon$ are satisfied for $\nu \geq \nu_\varepsilon$.

We still have to show that $w^\nu = w_{\bar{u}_j^\nu}$ holds for sufficiently large ν. First note that (5.10) yields

$$
w^\nu(p) = w_{\bar{u}_j^\nu}(p) \qquad \forall p \in \mathcal{N}_j^\bullet(u_j).
$$

Convergence (3.3) in combination with (5.28) provides

$$
\tfrac{1}{2}u_j(p) - |w^\nu(p) - u_j(p)| - |\bar{u}_j^\nu(p) - u_j(p)| \geq 0 \qquad \forall p \in \mathcal{N}_j^\circ(u_j) \qquad (5.29)
$$

for sufficiently large ν, giving

$$
\underline{\varphi}_{\bar{u}_j^\nu}(p) = \tfrac{1}{2}\bar{u}_j^\nu(p) \leq w^\nu(p) \leq 2\bar{u}_j^\nu(p) = \overline{\varphi}_{\bar{u}_j^\nu}(p) \qquad \forall p \in \mathcal{N}_j^\circ(u_j).
$$

Hence, $w^\nu \in \mathcal{K}_{\bar{u}_j^\nu}$ so that (4.11) and (5.23) must have the same solution $w_{\bar{u}_j^\nu} = w^\nu$. We finally choose ν_0 such that (5.10) and (5.29) are valid. □

Now we shall see that $\mathcal{C}_j^{\mathrm{std}}$ is asymptotically becomes a linear subspace correction method for the reduced linear problem (5.23).

Lemma 5.5. *Assume that the conditions in Lemma 5.1 are satisfied. Then there is a ν_0 such that the local obstacle problems (4.13) can be equivalently rewritten as*

$$
v_l^\nu \in V_l: \quad a_{\bar{u}_j^\nu}(v_l^\nu, v) = \ell_{\bar{u}_j^\nu}(v) - a_{\bar{u}_j^\nu}(w_{l-1}^\nu, v) \quad \forall v \in V_l, \qquad (5.30)
$$

if int supp $\lambda_l \subset \mathcal{N}_j^\circ(u_j)$ *and we have*

$$
v_l^\nu = 0 \qquad \textit{if} \ \ \text{int supp } \lambda_l \not\subset \mathcal{N}_j^\circ(u_j). \qquad (5.31)
$$

Assume further that non–zero corrections v_l^ν have the property

$$
\|v_k^\nu\|_\infty^2 = o(\|v_l^\nu\|_\infty), \quad \nu \to \infty, \qquad k = n_j + 1, \dots, l-1. \qquad (5.32)
$$

Then $\nu_1 \geq \nu_0$ can be chosen such that the damping parameters ω_l^ν defined in (5.5) satisfy

$$
\omega_l^\nu = 1 \qquad \forall \nu \geq \nu_1. \qquad (5.33)
$$

Proof. Let $p \in$ int supp $\lambda_l \cap \mathcal{N}_j^\bullet(u_j)$. Then (5.10) provides

$$
\underline{\varphi}_{\bar{u}_j^\nu}(p) = w_{l-1}^\nu(p) = \overline{\varphi}_{\bar{u}_j^\nu}(p) = 0.
$$

Hence, $\underline{\psi}_l^\nu(p) = \overline{\psi}_l^\nu(p) = 0$, due to (5.3). As $\underline{\psi}_l^\nu, \overline{\psi}_l^\nu \in V_l = \text{span}\{\lambda_l\}$, this leads to $\underline{\psi}_l^\nu(p) \equiv \overline{\psi}_l^\nu(p) \equiv 0$ and (5.31) follows. Now let int supp $\lambda_l \subset \mathcal{N}_j^\circ(u_j)$. By assumption, $\underline{\psi}_l^\nu, \overline{\psi}_l^\nu$, depend continuously on $\underline{\varphi}_{\bar{u}_j^\nu}, \overline{\varphi}_{\bar{u}_j^\nu}, w_{n_j}^\nu, \dots, w_{l-1}^\nu$. Hence, convergence

(4.17) of the intermediate iterates w_l^ν and condition (5.4) imply that we can find $\varepsilon > 0$ independent of l and ν such that

$$\underline{\psi}_l^\nu(p) < -\varepsilon < 0 < \varepsilon < \overline{\psi}_l^\nu(p) \qquad \forall p \in \text{int supp } \lambda_l$$

holds for sufficiently large ν. Convergence (4.17) yields $v_l^\nu \to 0$ so that we can finally choose ν_0 in such a way that (5.10) and

$$\underline{\psi}_l^\nu(p) < v_l^\nu(p) < \overline{\psi}_l^\nu(p) \qquad \forall p \in \text{int supp } \lambda_l \tag{5.34}$$

are valid for $\nu \geq \nu_0$. This proves (5.30) and (5.31).

We still have to show (5.33). The solution $v_l^\nu = z_l^\nu \lambda_l$ of (5.30) is given by

$$z_l^\nu = \frac{\ell_{\bar{u}_j^\nu}(\lambda_l) - a_{\bar{u}_j^\nu}(w_{l-1}, \lambda_l)}{a_{\bar{u}_j^\nu}(\lambda_l, \lambda_l)}. \tag{5.35}$$

Let $z_l^\nu \neq 0$. Inserting (5.35) in (5.5), we get

$$\omega_l^\nu \geq 2 \frac{a_{\bar{u}_j^\nu}(\lambda_l, \lambda_l)|z_l^\nu| - L_l B_l^2}{|z_l^\nu|(a_{\bar{u}_j^\nu}(\lambda_l, \lambda_l) + L_l(B_l + |z_l^\nu|))}.$$

In order to estimate L_l, note that (5.16) provides

$$\underline{\varphi} < \underline{\varphi}_{\bar{u}_j^\nu}(p) < \overline{\varphi}_{\bar{u}_j^\nu}(p) < \overline{\varphi} \qquad \forall p \in \mathcal{N}_j^\circ(u_j)$$

with $\underline{\varphi}$, $\overline{\varphi}$ defined in (5.17). Hence, exploiting property (1.5) of Φ, we get

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

with L^* independent of ν. Using (5.10) this leads to

$$L_l = \sum_{p \in \mathcal{N}_j^\circ(u_j)} L_p^\nu |\lambda_l(p)| \; h_p \leq |\Omega| L^*$$

with $|\Omega|$ denoting the area of Ω. In the light of $a_{\bar{u}_j^\nu}(\lambda_l, \lambda_l) \geq a(\lambda_l, \lambda_l)$ we now obtain

$$\omega_l^\nu \geq \frac{2}{1 + \frac{L^*|\Omega|}{a(\lambda_l,\lambda_l)} \sum_{k=n_j+1}^{l} \|v_k^\nu\|_\infty} - \frac{2L^*|\Omega|}{a(\lambda_l, \lambda_l)} \left(\sum_{k=n_j+1}^{l-1} \left(\frac{\|v_k^\nu\|_\infty^2}{\|v_l^\nu\|_\infty} \right)^{1/2} \right)^2 .$$

Exploiting $\|v_k^\nu\|_\infty \to 0$ and (5.32) we can finally choose $\nu_1 \geq \nu_0$ such that (5.16) and (5.33) are satisfied. □

The technical assumption (5.32) is the price that we have to pay for evaluating derivatives of $\phi_{\bar{u}_j^\nu}$ at $\bar{u}_j^\nu \neq w_{l-1}^\nu$. However, evaluating derivatives at w_{l-1}^ν would require additional interpolations that lead to suboptimal complexity ranging from $\mathcal{O}(n \log n)$ (uniform refinement) to $\mathcal{O}(n^2)$ (highly nonuniform case).

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

Theorem 5.6. *Assume that the conditions of Lemma 5.5 are satisfied. Let*

$$\|v\|_{u_j} \leq \gamma_j \|v\| \qquad \forall v \in \mathcal{S}_j^\circ. \tag{5.36}$$

Then there is a $\nu_j \geq 0$ such that the iterates produced by the standard monotone multigrid method (5.7) fulfill the error estimate

$$\|u_j - u_j^{\nu+1}\|_{u_j} \leq (1 - c\gamma_j^{-1}(j+1)^{-4}) \|u_j - u_j^\nu\|_{u_j} \qquad \forall \nu \geq \nu_j. \tag{5.37}$$

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

Proof. We have seen in Lemma 5.5 that coarse grid correction $\mathcal{C}_j^{\mathrm{std}}$ becomes a linear subspace correction for the linear reduced problem (5.21) if $\nu \geq \nu_1$. The corresponding 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 these subspaces, the bilinear form $a_{\bar{u}_j^\nu}(\cdot,\cdot)$ 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. More precisely, we want to show

$$\|w_{\bar{u}_j^\nu} - \mathcal{C}_j^{\mathrm{std}} \bar{u}_j^\nu\|_{\bar{u}_j^\nu} \leq q_j \|w_{\bar{u}_j^\nu} - \bar{u}_j^\nu\|_{\bar{u}_j^\nu} \quad \forall \nu \geq \nu_1 \tag{5.38}$$

with $q_j = (1 - \tilde{c}\gamma_j^{-1}(j+1)^{-4})$ and a constant $\tilde{c}$ depending only on the shape regularity of $\mathcal{T}_0$ and on the ellipticity of $a(\cdot,\cdot)$.

In order to prove (5.38), we shall apply Theorem 2.5 from a recent paper of Neuss [15]. To this end, we have to check the conditions (V0)–(V2) stated there. Let $b_k^{\mathrm{s}}(\cdot,\cdot)$ denote the symmetric bilinear form as induced by the symmetric Gauß-Seidel relaxation on $\mathcal{W}_k$. For some $v \in \mathcal{S}_j^{\circ}$, we 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 modified interpolation operators I_k defined by

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

The smoothing property (V0)

$$a_{\bar{u}_j^\nu}(v,v) \leq \omega b_k(v,v) \quad \forall v \in \mathcal{W}_k \tag{5.39}$$

with suitable $\omega \in (0,2)$ depending only on $\mathcal{T}_0$ is explicitly stated in Theorem 3.2 in [15].

Stability condition (V1) with $K_1 = C\gamma_j(j+1)^2$ takes the form

$$\sum_{k=0}^{j} b_k^{\mathrm{s}}(v_k, v_k) \leq C\gamma_j(j+1)^2 \|v\|_{\bar{u}_j^\nu}^2. \tag{5.40}$$

In order to prove (5.40) with a constant C depending only on $\mathcal{T}_0$ and on the ellipticity of $a(\cdot,\cdot)$, we use the estimate

$$b_k^{\mathrm{s}}(v_k, v_k) \leq c \sum_{i=1}^{n_k} a_{\bar{u}_j^\nu}(\lambda_{p_i}^{(k)}, \lambda_{p_i}^{(k)}) v_k(p_i)^2$$

which holds for all $v_k \in \mathcal{W}_k$ with $c > 0$ depending only on $\mathcal{T}_0$ (see e.g. (29) in [15]), condition (5.36) and recent results on modified hierarchical splittings as contained in section 5 of [14].

Finally,

$$a_{\bar{u}_j^\nu}(v_l, v_k) \leq \omega^{\frac{1}{2}} b_l^{\mathrm{s}}(v_l, v_l)^{\frac{1}{2}} b_k(v_k, v_k)^{\frac{1}{2}} \qquad \forall v_l \in \mathcal{W}_l, v_k \in \mathcal{W}_k \tag{5.41}$$

follows directly from the Cauchy-Schwarz inequality, (5.39) and

$$a_{\bar{u}_j^\nu}(v_l, v_l) \leq b_k^{\mathrm{s}}(v_l, v_l)$$

which is the well-known smoothing property of the symmetric Gauß-Seidel relaxation. As a consequence of (5.41), (V2) holds with $K_2 = \sqrt{2}(j+1)$. Now, we can apply Theorem 2.5 in [15] in order to get the desired estimate (5.38).

Note that for given $\delta > 0$ (5.10) and convergence (3.3) provide the norm equivalence

$$(1-\delta)\|v\|_{u_j} \le \|v\|_{\bar{u}_j^\nu} \le (1+\delta)\|v\|_{u_j}$$

uniformly on bounded subsets of $\mathcal{S}_j^\circ$ for sufficiently large ν. Hence, for given $\varepsilon > 0$ we can find ν_ε such that

$$\|w_{\bar{u}_j^\nu} - \mathcal{C}_j^{\mathrm{std}} \bar{u}_j^\nu\|_{u_j} \le (1+\varepsilon) q_j \|w_{\bar{u}_j^\nu} - \bar{u}_j^\nu\|_{u_j} \quad \forall \nu \ge \nu_\varepsilon. \tag{5.42}$$

To conclude the proof, we combine the estimates (5.14), (5.24) and (5.42) by the triangle inequality in order to get

$$\begin{aligned}
\|u_j - u_j^{\nu+1}\|_{u_j} &= \|u_j - \mathcal{C}_j^{\mathrm{std}} \bar{u}_j^\nu\|_{u_j} \\
&\le \|w_{\bar{u}_j^\nu} - u_j\|_{u_j} + \|w_{\bar{u}_j^\nu} - \mathcal{C}_j^{\mathrm{std}} \bar{u}_j^\nu\|_{u_j} \\
&\le \varepsilon \|\bar{u}_j^\nu - u_j\|_{u_j} + (1+\varepsilon) q_j \|w_{\bar{u}_j^\nu} - \bar{u}_j^\nu\|_{u_j} \\
&\le \varepsilon(1+\varepsilon)\|u_j^\nu - u_j\|_{u_j} + (1+\varepsilon) q_j (\|w_{\bar{u}_j^\nu} - u_j\|_{u_j} + \|\bar{u}_j^\nu - u_j\|_{u_j}) \\
&\le \big(\varepsilon(1+\varepsilon) + (\varepsilon(1+\varepsilon)^2 + (1+\varepsilon)^2) q_j\big) \|u_j - u_j^\nu\|_{u_j}.
\end{aligned}$$

We finally choose $\nu_\varepsilon \ge \nu_1$ such that (5.14), (5.24) and (5.42) hold with

$$\varepsilon \le \frac{1}{18} \frac{\tilde{c}}{\gamma_j (j+1)^4}. \tag{5.43}$$

Then, the desired estimate (5.37) follows with $c = \frac{\tilde{c}}{2}$ and $\nu_j = \nu_\varepsilon$. □

We emphasize that (5.37) describes the worst case and can be easily improved on suitable regularity assumptions. For example, let

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

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

$$a(v,w) = \int_\Omega \sum_{l,k=1}^{2} a_{lk}\, \partial_l v\, \partial_k w \, dx, \tag{5.44}$$

with coefficients $a_{lk} \in C^1(\bar{\Omega})$. Then, exploiting a sharpened Cauchy-Schwarz inequality instead of (5.41), 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 numerical computations [13], we also observed mesh-independent convergence rates with respect to the usual energy norm induced by $a(\cdot,\cdot)$. A theoretical justification will be subject of future research.

The preceding convergence analysis clarifies the basic idea behind monotone iterations (3.1). Fine grid smoother $\mathcal{M}_j$ provides global convergence exploiting convexity of the underlying minimization problem. Additional coarse grid correction $\mathcal{C}_j^{\mathrm{std}}$ asymptotically becomes a Newton multigrid method with polylogarithmic convergence rates exploiting local smoothness of Φ. The accuracy of iterates u_j^ν required to enter the asymptotic regime depends on stability of critical nodes $\mathcal{N}_j^\bullet(u_j)$ and on Lipschitz continuity of Φ'' at $u_j(p)$, $p \in \mathcal{N}_j^\circ(u_j)$. Numerical experiments indicate that initial iterates u_j^0 as resulting from nested iteration are frequently good enough to provide multigrid convergence rates immediately.

6. Truncated Monotone Multigrid Methods

Monotone iterations (3.1) are constructed in such a way that coarse grid correction $\mathcal{C}_j$ does 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 $\mathcal{C}_j^{\text{std}}$. It is well-known (cf. eg. [14]) that this may lead to poor representation of low frequency parts of the error. In order to improve the convergence rates by improved coarse grid transport, we shall now modify all $\lambda_l \in \Lambda_\mathcal{S}$ with the property (6.1) according to $\mathcal{N}_j^\bullet(\bar{u}_j^\nu)$.

Following [10, 14], we define 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 truncation operators $T_{j,k}^\nu$, $k = 0, \dots, j$,

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

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

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

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

The resulting *truncated multilevel nodal basis* $\tilde{\Lambda}_\mathcal{S}^\nu$,

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

clearly depends on the set $\mathcal{N}_j^\bullet(\bar{u}_j^\nu)$ which may change in each iteration step. We now derive a *truncated* coarse grid correction $\mathcal{C}_j^{\text{trc}}$ by the same reasoning as described in the previous section. More precisely, introducing some ordering from fine to coarse

$$\tilde{\lambda}_l = \tilde{\lambda}_{p_l}^{(k_l)}, \qquad l = n_j + 1, \dots, m_j^\nu = n_j + \tilde{m}_\mathcal{S}^\nu$$

of the $\tilde{m}_\mathcal{S}^\nu$ non–zero elements of $\tilde{\Lambda}_\mathcal{S}^\nu$, we now use the search directions

$$\mu_l^\nu = \tilde{\lambda}_l, \qquad l = n_j + 1, \dots, m_j^\nu, \quad \nu \geq 0.$$

Local constraints $\mathcal{D}_l$, as appearing in (4.13), are obtained from slightly modified monotone restrictions (see [10, 13]) and local damping parameters ω_l are obtained by replacing λ_l by $\tilde{\lambda}_l$ in (5.5).

The resulting iterative scheme

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

is called *truncated monotone multigrid method.* Global convergence of (6.5) follows from Theorem 3.1 and Proposition 4.1.

Theorem 6.1. *Assume that the conditions of Lemma 5.1 and (5.36) hold. Assume further that all non–zero corrections* $v_l^\nu = z_l^\nu \tilde\lambda_l$, $\tilde\lambda_l \in \tilde\Lambda_{\mathcal{S}}^\nu$, *as resulting from (4.13) have property (5.32).*

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

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

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

Proof. The proof is essentially the same as for Theorem 5.6. We only have to establish an analogue of Lemma 5.5 involving $\tilde\lambda_l = \tilde\lambda_{p_l}^{(k_l)}$ instead of $\lambda_l = \lambda_{p_l}^{(k_l)}$ and an error estimate of the form (5.42) for the reduced linear iteration. Note that (5.39) and (5.41) still hold if $\mathcal{W}_k$ is replaced by the larger space $\tilde{\mathcal{W}}_k$,

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

□

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.

Consider some $p \in$ int supp $\tilde\lambda_l$ with $L_p^\nu >> 1$. Then our local damping strategy clearly gives $\omega_l \approx 0$ so that there is almost no contribution from $\tilde\lambda_l$. Hence, such p play a similar role as critical nodes $\mathcal{N}_j^\bullet(\bar u_j^\nu)$ in (6.1) and it seams reasonable to treat them similarly in the truncation process. This can be done by replacing definition (4.1) of regular nodes by

$$\mathcal{N}_j^\circ(\bar u_j^\nu) = \{p \in \mathcal{N}_j \mid \bar u_j^\nu(p) > 0 \text{ and } L_p^\nu < L_{\max}\} \tag{6.7}$$

with some given threshold $L_{\max} > 0$. Of course, this modification preserves global convergence. If $L_{\max}$ is sufficiently large, then there is a ν_0 such that (4.1) and (6.7) define the same sets for $\nu \geq \nu_0$. Hence, we still have asymptotic bounds of the convergence rates in this case. For numerical results, we refer to [13] and the experiments to be reported below.

7. Numerical Results

We consider the stationary porous medium equation

$$-\Delta\theta^2 - f(\theta) = 0 \qquad \theta \geq 0 \tag{7.1}$$

with absorption term

$$f(\theta) = \begin{cases} \theta, & \text{if } \theta \in [0,1) \\ [1,2], & \text{if } \theta = 1 \\ 2, & \text{if } \theta \geq 1 \end{cases}$$

and constant Dirichlet boundary conditions $\theta \equiv 2$ on $\partial\mathcal{O}$, $\mathcal{O} = (-10,10)^2$. After Kirchhoff–type transformation $u = \theta^2$ the weak formulation of (7.1) takes the form (1.1) with $a(v,w) = (\nabla v, \nabla w)$,

$$\partial\Phi(u) = \begin{cases} (-\infty, 0], & \text{if } u = 0 \\ f(\sqrt{u}), & \text{if } u > 0 \end{cases},$$

$\Omega = (0, 10)^2$ and appropriate boundary conditions. Observe that our model problem combines an obstacle condition, a jump and unbounded Lipschitz constants for $u \in (0, 1)$.

The initial triangulation $\mathcal{T}_0$ is obtained by subdividing Ω in 4 congruent triangles. Triangulation $\mathcal{T}_{j+1}$ is obtained from $\mathcal{T}_j$ by an ad hoc local refinement strategy: A triangle $t \in \mathcal{T}_j$ is marked for refinement if $\tilde{u}_j$ does not vanish on t. The approximate solution $\tilde{u}_j$ on $\mathcal{T}_j$ is computed up to 0.05% accuracy. More precisely, $\tilde{u}_j = u_j^{\nu^*}$ is accepted as soon as the stopping criterion

$$\|u_j^{\nu^*} - u_j^{\nu^*-1}\|_{u_j^{\nu^*}} \leq 5 \cdot 10^{-4} \|u_j^{\nu^*}\|_{u_j^{\nu^*}} \tag{7.2}$$

is fulfilled. Note that $\|\cdot\|_{u_j^{\nu^*}}$ is intended to approximate the local energy norm $\|\cdot\|_{u_j}$ as defined in (5.8). For iterative solution of the discrete problems (1.6) on each refinement level j we use the standard monotone multigrid method (STDKH) and the truncated variant (TRCKH) as described in Sections 5 and 6, respectively. The second singularity at $u = 1$ is incorporated as described in [12], pp. 65. Truncation is based on the modification (6.7) with $L_{\max} = 10^{12}$. In the light of Theorems 5.6 and 6.1 nonlinear Gauß-Seidel smoothing with exact evaluation of (2.1) is applied. Using the initial iterate $u_j^0 = \tilde{u}_{j-1}$, $j = 1, \ldots, 8$, (nested iteration) at most 7 (STDKH) or 6 (TRCKH) iteration steps were needed in order to meet the accuracy requirement (7.2). We found similar results for the inexact variant as described in Section 2. Implementation was carried out in the framework of the finite element toolbox KASKADE [3].

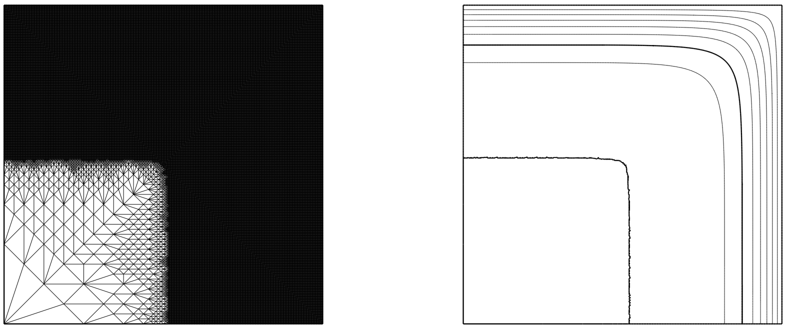

FIGURE 1. Final grid $\mathcal{T}_8$ and level curves of final approximation $\tilde{u}_8$

Figure 1 shows the final triangulation $\mathcal{T}_8$ together with the level curves of the final approximation $\tilde{u}_8$. Bold lines are used for the free boundaries $\tilde{u}_8 \equiv 0$ and $\tilde{u}_8 \equiv 1$. Observe that in large parts of the computational domain $\tilde{u}_8$ is close to the singularity zero where local Lipschitz constants tend to infinity.

We take a closer look at the convergence behavior of our monotone multigrid methods on the final level $j = 8$ with 97 285 unknowns. The left picture of Figure 2 shows the algebraic error $\|u_8 - u_8^\nu\|_{u_8}$ over the number ν of iteration steps. The initial iterate is $u_8^0 = \tilde{u}_7$ (nested iteration). The exact solution u_8 is precomputed up to machine precision. For both methods, we observe a fast reduction of the high frequency contributions to the error in the first iteration step. Then, asymptotic linear convergence dominates the whole iteration history. This supports the practical relevance of our asymptotic convergence analysis. In the leading iteration

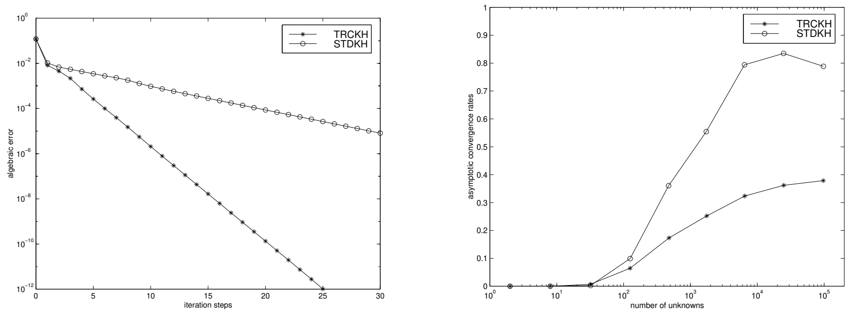

FIGURE 2. Iteration history and asymptotic convergence rates

steps, the algorithms provide damping of at most 126 (STDKH) or 153 (TRCKH) local corrections. Minor effects on the convergence rates illustrate the benefits of local damping. No damping occurs after 28 (STDKH) or 24 (TRCKH) iteration steps confirming our theoretical analysis (cf. Lemma 5.5). In comparison with the standard method the truncated variant exhibits a considerable improvement of convergence speed. This justifies our heuristic reasoning in Section 6. Of course, dominance of asymptotic convergence rates is a consequence of sufficiently accurate initial iterates as obtained by nested iteration. Starting from $u_8^0 \equiv 0$, i.e. directly from the singularity, TRCKH required 180 iteration steps to enter the asymptotic regime.

The right picture in Figure 2 shows approximate asymptotic convergence rates

$$\rho_j = \frac{\|u_j - u_j^{\nu^*}\|_{u_j}}{\|u_j - u_j^{\nu^*-1}\|_{u_j}}, \qquad j = 0, \dots, 8.$$

Here, ν^* is chosen such that $\|u_j - u_j^{\nu^*}\|_{u_j} < 10^{-10}$ and again u_j is precomputed up to machine precision. The asymptotic convergence rates seem to saturate with increasing refinement level j confirming the convergence results as stated in Theorems 5.6 and 6.1, respectively.

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] R. Beck, B. Erdmann, and R. Roitzsch. KASKADE Manual, Version 3.0. Technical Report TR95-4, Konrad–Zuse–Zentrum (ZIB), Berlin, 1995.

[4] P. Deuflhard and M. Weiser. Global inexact Newton multilevel FEM for nonlinear elliptic problems. In W. Hackbusch and G. Wittum, editors, *Multigrid methods V*, Lecture Notes in Computational Science and Engineering, Berlin, Heidelberg, 1998. Springer.

[5] M. Dryja and W. Hackbusch. On the nonlinear domain decomposition method. *BIT*, 32:296–311, 1997.

[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. *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems*. Teubner, Stuttgart, 1997.
[13] R. Kornhuber. On robust multigrid methods for non–smooth variational problems. In W. Hackbusch and G. Wittum, editors, *Multigrid Methods V*, pages 173–188, Berlin, 1998. Springer.
[14] R. Kornhuber and H. Yserentant. Multilevel methods for elliptic problems on domains not resolved by the coarse grid. *Contemporary Mathematics*, 180:49–60, 1994.
[15] N. Neuß. V-cycle convergence with unsymmetric smoothers and application to an anisotropic model problem. *SIAM J.Numer.Anal.*, 35:1201–1212, 1998.
[16] P. Oswald. Stable subspace splittings for Sobolev spaces and their applications. Forschungsergebnisse der Friedrich-Schiller-Universität Jena Math/93/7, Jena, 1993.
[17] X.C. Tai and J. Xu. Global and uniform convergence of subspace correction methods for some convex optimization methods. Preprint, Pennsylvania State University, 1999.

FU Berlin, Math. Institut II, Arnimallee 2-6, D-14195 Berlin, Germany
E-mail address: `kornhuber@math.fu-berlin.de`