

A SEMI-SMOOTH NEWTON METHOD FOR SOLVING ELLIPTIC EQUATIONS WITH GRADIENT CONSTRAINTS

ROLAND GRIESSE¹ AND KARL KUNISCH²

Abstract. Semi-smooth Newton methods for elliptic equations with gradient constraints are investigated. The one- and multi-dimensional cases are treated separately. Numerical examples illustrate the approach and as well as structural features of the solution.

Mathematics Subject Classification. 35J70, 49M15, 65K05, 90C33.

Received November 13, 2007.

Published online December 5, 2008.

1. INTRODUCTION

In this paper we investigate iterative methods for the numerical solution of an elliptic partial differential equation (PDE) with gradient constraints. In particular, we consider

$$\begin{aligned} \max\{-\Delta u - f, \quad |\nabla u| - g\} &= 0 \quad \text{a.e. in } \Omega, \\ u &= 0 \quad \text{on } \Gamma = \partial\Omega, \end{aligned} \tag{1.1}$$

where $\Omega \subset \mathbb{R}^d$ is a bounded domain with smooth boundary Γ .

This problem was originally studied in [5], where sufficient conditions for existence, uniqueness and regularity results of the type $u \in W^{1,\infty}(\Omega)$ and $u \in W_{\text{loc}}^{2,p}(\Omega)$ were obtained. These results were refined in [17], where sufficient conditions for $u \in C^{1,1}(\Omega)$ were given, and in [9] where among other aspects, more general first order constraints are admitted.

In the present work we aim for the efficient numerical treatment of (1.1). We shall analyze semi-smooth Newton methods for an appropriately defined family of approximating problems. It will be verified that this approximation is consistent in the sense that the solutions to this family of approximating problems converge to the solution of (1.1) and that the semi-smooth Newton method converges super-linearly for each member

Keywords and phrases. Gradient constraints, active set strategy, regularization, semi-smooth Newton method, primal-dual active set method.

¹ Chemnitz University of Technology, Faculty of Mathematics, 09107 Chemnitz, Germany.
roland.griesse@mathematik.tu-chemnitz.de; <http://www.tu-chemnitz.de/~griesse>

² Institute for Mathematics and Scientific Computing, Heinrichstraße 36, 8010 Graz, Austria. karl.kunisch@uni-graz.at;
<http://www.kfunigraz.ac.at/imawww/kunisch>

of the family. We recall here that semi-smooth Newton methods in a function space setting were successfully applied to diverse variational problems, including optimal control problems with constraints, for example. We refer to [8] and the references give there. Differently from previous applications of semi-smooth Newton methods, (1.1) does not directly arise from a variational setting. This will require us to use techniques different from those in [8] to analyze the proposed methods.

In special cases, however, (1.1) is equivalent to variational problems. In particular, if f and g are constant then (1.1) is equivalent to the elasto-plastic torsion problem, with u representing the stress function and Ω the cross section of an elasto-plastic beam. In [5] it is also pointed out that the solution u of (1.1) can be related to the minimal expected cost in stochastic optimal control problems.

Our motivation for investigating (1.1) is in part intrinsic and to the other part related to the fact that certain portfolio optimization problems have a structure related to (1.1). To explain the latter consider the system of stochastic differential equations

$$\begin{aligned} dX_0(t) &= rX_0(t) dt - (1 + \gamma) dL(t) + (1 - \gamma) dM(t) \\ dX_1(t) &= \mu X_1(t) dt + \sigma X_1(t) dW(t) + dL(t) - dM(t) \end{aligned}$$

where X_0, X_1 denote wealth processes for the bank account and stock respectively, where further r, μ, σ, γ stand for the interest rate, trend, volatility, trading costs and L, M are the cumulative processes describing purchases and sales of stock.

Maximizing the expected utility $U(x) = x^\alpha/\alpha$ for the terminal wealth after liquidating the position in the stocks requires us to consider the stochastic optimization problem

$$J(t, x_0, x_1) = \sup_{L, M} E\left[\frac{1}{\alpha}(X_0(T) + (1 - \gamma)X_1(T))^\alpha \mid X_0(t) = x_0, X_1(t) = x_1\right],$$

where E stands for expectation and $\alpha \in (0, 1)$.

Then, without entering into details, it can be shown [14] that J is a concave, continuous, viscosity solution of

$$\max \{ J_t + \mathcal{A}J, -(1 + \gamma)J_{x_0} + J_{x_1}, (1 - \gamma)J_{x_0} - J_{x_1} \} = 0 \quad (1.2)$$

on $[0, T) \times \mathcal{D}$ with $J(T, x_0, x_1) = \frac{1}{\alpha}(x_0 + (1 - \gamma)x_1)^\alpha$, where

$$\mathcal{A}h(x_0, x_1) = rx_0h_{x_0}(x_0, x_1) + \mu h_{x_1}(x_0, x_1) + \frac{1}{2}\sigma^2x_1^2h_{x_1, x_1}(x_0, x_1),$$

and \mathcal{D} is the solvency region. If the cost-functional is of infinite horizon type, *i.e.* if it extends over $[0, \infty)$, then the resulting Hamilton Jacobi Bellman equation corresponding to (1.2) is time independent. Using the so-called homotheticity property of (1.2), this equation can be reduced to a one-dimensional equation for the ‘spatial’ variable $\frac{x_0}{x_0+x_1}$, known as the risky fraction. The resulting one-dimensional equation has similar structure max-structure as (1.2). Clearly (1.2) is significantly more challenging to analyze than (1.1). But it is a worthwhile challenge to analyze (1.1) before (1.2). The numerical feasibility for solving one-dimensional portfolio problems was already established in earlier work [12].

Let us briefly comment on the following sections. In Section 2 we consider the one dimensional version of (1.1). It is distinctly different from the multi-dimensional one for two reasons. First in this case the nonlinear constraint $|\nabla u| \leq g$ is equivalent to the bilateral affine constraint $-g \leq u_x \leq g$, and secondly open sets in dimension one can be expressed as the countable union of open intervals. Both these structural properties have no immediate counterpart in the multi-dimensional case. After giving sufficient conditions for existence and uniqueness, a family of approximating problems is introduced and their asymptotic behavior is analyzed. Each member of the family can be solved by a semi-smooth Newton method or, equivalently a primal-dual active

set method, with super-linear convergence rate. In Section 3 we carry out essentially the same program for dimension $d \geq 2$, but now for the nonlinear constraint term $|\nabla u| \leq g$, which requires a significantly different treatment as the affine case. Concerning existence and uniqueness, we can quote [5]. Nevertheless we give an independent treatment here. In part this is done for the sake of completeness and to the other part, that we base the existence proof on the same family of *non-differentiable* approximating problems, as is used for the numerical treatment. Earlier work relies on smooth families of approximating problems. Section 4 is devoted to reporting on numerical experiments with the proposed algorithms. In a short final section we point out some open problems related to (1.1), of which there are, of course, many.

We recall here the weak maximum principle for the equation

$$Lu := -\Delta u + b \cdot \nabla u \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \Gamma$$

with $b \in (L^d(\Omega))^d$. If $u \in H_0^1(\Omega)$ as well as $Lu \leq 0$ in $H^{-1}(\Omega)$ and $u \leq 0$ on Γ holds, then $u \leq 0$. The proof in [16], Theorem 2.4, requires that $b \in (L^\infty(\Omega))^d$, but can be modified to the case $b \in (L^d(\Omega))^d$.

Notation. Throughout, c denotes a generic positive constant which can change its meaning in every occurrence.

2. ANALYSIS OF THE ONE-DIMENSIONAL CASE

In the one-dimensional case we choose $\Omega = (0, 1)$ and utilize the following assumption.

Assumption 2.1. We assume $f \in C(\Omega)$ and $f > 0$ as well as $g \in H^1(\Omega)$ and $g \geq 0$ in Ω .

Our approach is motivated by the following equivalent reformulation of (1.1) as a complementarity system:

$$\begin{aligned} -u'' + \lambda^+ + \lambda^- &= f && \text{a.e. in } \Omega \\ \lambda^+ \geq 0, \quad u' - g \leq 0, \quad \lambda^+(u' - g) &= 0 && \text{a.e. in } \Omega \\ \lambda^- \geq 0, \quad -u' - g \leq 0, \quad \lambda^-(-u' - g) &= 0 && \text{a.e. in } \Omega. \end{aligned} \tag{2.1}$$

2.1. Regularization

Instead of (2.1), we solve a sequence of regularized problems for an increasing sequence of parameters $\gamma \geq 0$:

$$-u'' + \gamma \max\{0, u' - g\} + \gamma \max\{0, -u' - g\} = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \Gamma. \tag{2.2}$$

The justification of calling (2.2) a regularization of (1.1) is due to the fact that diffusion is added on the active set $\{|u'| = g\}$.

Proposition 2.2. Problem (2.2) has a unique solution $u \in H^2(\Omega) \cap H_0^1(\Omega)$.

Proof. Existence. We define $T : H_0^1(\Omega) \rightarrow H_0^1(\Omega)$ by the relation $v = Tu$ and

$$-v'' + \gamma \max\{0, u' - g\} + \gamma \max\{0, -u' - g\} = f \quad \text{in } \Omega, \quad v = 0 \quad \text{on } \Gamma.$$

Let S be a bounded set in $H_0^1(\Omega)$. Then $\{\gamma \max\{0, u' - g\} + \gamma \max\{0, -u' - g\} : u \in S\}$ is a bounded set in $L^2(\Omega)$. Consequently, $T(S)$ is a bounded set in $H^2(\Omega) \cap H_0^1(\Omega)$ and thus $T(S)$ is precompact in $H_0^1(\Omega)$. This shows that T is compact. Suppose that $s \in [0, 1]$ and v solves $v = sT(v)$, i.e.,

$$-v'' + s\gamma \max\{0, v' - g\} + s\gamma \max\{0, -v' - g\} = sf \quad \text{in } \Omega, \quad v = 0 \quad \text{on } \Gamma.$$

By the maximum principle, $0 \leq v \leq (-\Delta)^{-1}f$. Testing the equation with v yields

$$\|\nabla v\|_{L^2(\Omega)}^2 \leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}.$$

Using Poincaré's and Young's inequalities, we derive the bound $\|v\|_{H^1(\Omega)} \leq c\|f\|_{L^2(\Omega)}$, independent of $s \in [0, 1]$. The existence of a solution $u \in H_0^1(\Omega)$ of (2.2), *i.e.*, of a fixed point of T , now follows from the Leray-Schauder fixed point theorem (see Appendix 5). By regularity of the data, $u \in H^2(\Omega)$ holds.

Uniqueness. Suppose that $u \neq v$ are solutions of (2.2). Using their continuity, one shows as in [5], p. 563, that there exists $\varepsilon > 0$ and $x_1 \in \Omega$ such that $u - v - \varepsilon u$ assumes a positive maximum in x_1 , hence

$$(1 - \varepsilon)u'(x_1) = v'(x_1). \quad (2.3)$$

Moreover, we have

$$-(1 - \varepsilon)u'' + v'' - C(x) = -\varepsilon f \quad \text{a.e. in } \Omega,$$

where

$$\begin{aligned} -C(x) &= (1 - \varepsilon)\gamma \max\{0, u' - g\} - \gamma \max\{0, v' - g\} \\ &\quad + (1 - \varepsilon)\gamma \max\{0, -u' - g\} - \gamma \max\{0, -v' - g\}. \end{aligned}$$

By Bony's maximum principle [1], see Appendix 5, we conclude that

$$0 \leq \lim \operatorname{ess\,sup}_{x \rightarrow x_1} (-(1 - \varepsilon)u'' + v'')$$

holds, and hence

$$0 \leq \lim \operatorname{ess\,sup}_{x \rightarrow x_1} (C(x) - \varepsilon f(x)) = C(x_1) - \varepsilon f(x_1), \quad (2.4)$$

due to continuity of C and f . We now distinguish cases and derive a contradiction to (2.4) in each case. All functions are evaluated at x_1 .

- (1) If $u' > g$ and $v' > g$, then (2.4) implies $\gamma(1 - \varepsilon)(u' - g) - \gamma(v' - g) \leq -\varepsilon f$, and (2.3) yields the contradiction $\gamma g \leq -f$.
- (2) If $u' > g$ and $|v'| \leq g$, then (2.4) implies the contradiction $\gamma(1 - \varepsilon)(u' - g) \leq -\varepsilon f$.
- (3) If $u' > g$ and $v' < -g$, then (2.3) implies a contradiction.
- (4) If $|u'| \leq g$ and $v' > g$, then (2.4) implies $-\gamma(v' - g) \leq -\varepsilon f$, and (2.3) yields $-\gamma(1 - \varepsilon)u' + \gamma g \leq -\varepsilon f$. In case $-g \leq u' \leq 0$, this implies the contradiction $0 \leq -\varepsilon f$. In case $0 \leq u' \leq g$, we infer the contradiction $\gamma \varepsilon u' \leq -\varepsilon f$.
- (5) If $|u'| \leq g$ and $|v'| \leq g$, then (2.4) implies the contradiction $0 \leq -\varepsilon f$.
- (6) If $|u'| \leq g$ and $v' < -g$, then (2.3) implies a contradiction.
- (7) If $u' < -g$ and $v' > g$, then (2.3) implies a contradiction.
- (8) If $u' < -g$ and $|v'| \leq g$, then (2.4) implies the contradiction $\gamma(1 - \varepsilon)(-u' - g) \leq -\varepsilon f$.
- (9) If $u' < -g$ and $v' < -g$, then (2.3) and (2.4) imply the contradiction $\varepsilon \gamma g \leq -\varepsilon f$.

Consequently, the solution of (2.2) is unique. □

For the subsequent discussion, we denote the unique solution of (2.2) by u_γ . We also introduce the abbreviations

$$\lambda_\gamma^+ := \gamma \max\{0, u' - g\}, \quad \lambda_\gamma^- := \gamma \max\{0, -u' - g\}, \quad \lambda_\gamma := \lambda_\gamma^+ + \lambda_\gamma^-. \quad (2.5)$$

Lemma 2.3 (*a priori estimates for u_γ*). *The unique solution of (2.2) satisfies the following a priori bounds:*

$$\|u_\gamma\|_{L^\infty(\Omega)} \leq c \|f\|_{L^2(\Omega)} \quad \text{and} \quad 0 \leq u_\gamma \leq (-\Delta)^{-1}f \quad (2.6)$$

$$\|u_\gamma\|_{H^1(\Omega)} \leq c \|f\|_{L^2(\Omega)} \quad (2.7)$$

$$\|\lambda_\gamma\|_{L^2(\Omega)} \leq \|f\|_{L^2(\Omega)} + \|g'\|_{L^2(\Omega)} \quad (2.8)$$

where all constants are independent of γ .

Proof. The pointwise inequality $0 \leq u_\gamma \leq (-\Delta)^{-1}f$ in (2.6) is a consequence of the maximum principle. The first inequality in (2.6) then follows from the embedding $H^2(\Omega) \hookrightarrow L^\infty(\Omega)$ and the *a priori* estimate $\|(-\Delta)^{-1}f\|_{H^2(\Omega)} \leq c \|f\|_{L^2(\Omega)}$. The bound (2.7) was already shown in the proof of Proposition 2.2.

We now define the disjoint sets $A^+ := \{x \in \Omega : \lambda_\gamma^+ > 0\}$ and $A^- := \{x \in \Omega : \lambda_\gamma^- > 0\}$ and discuss A^+ to get an estimate for λ_γ^+ . Owing to the continuity of u' and g , A^+ is an open set and thus it can be written as the countable union of pairwise disjoint open intervals $I_n = (l_n, r_n)$, i.e., $A^+ = \bigcup_{n=1}^\infty I_n$, see [15], p. 98. We write (2.2) as

$$-u'' + g' + \lambda_\gamma^+ + \lambda_\gamma^- = f + g' \quad \text{in } \Omega, \quad u(0) = u(1) = 0,$$

multiply with $u' - g$ and integrate over I_n to obtain

$$-\int_{I_n} (u'' - g')(u' - g) dx + \gamma \int_{I_n} |u' - g|^2 dx = \int_{I_n} (f + g')(u' - g) dx.$$

This implies

$$-\frac{1}{2}|u' - g|^2 \Big|_{l_n}^{r_n} + \gamma \|u' - g\|_{L^2(I_n)}^2 \leq \|f + g'\|_{L^2(I_n)} \|u' - g\|_{L^2(I_n)}.$$

For the first term, we distinguish cases:

- (1) $l_n = 0$. The contribution from the lower bound is $\frac{1}{2}|u' - g|^2(l_n) \geq 0$ and thus can be left out.
- (2) $l_n > 0$. By continuity of u' and g , $u'(l_n) - g(l_n) = 0$ holds and the contribution from the lower bound vanishes.
- (3) $r_n < 1$. As above, $u'(r_n) - g(r_n) = 0$ holds.
- (4) $r_n = 1$. This implies $u'(1) \geq g(1) \geq 0$. However, from the maximum principle we know that $u'(1) \leq 0$ holds. Hence $u'(r_n) = g(r_n) = 0$ and the contribution from the upper bound vanishes.

We thus obtain

$$\|\lambda_\gamma^+\|_{L^2(I_n)} = \gamma \|u' - g\|_{L^2(I_n)} \leq \|f + g'\|_{L^2(I_n)}$$

and summation over the intervals I_n yields the desired estimate

$$\|\lambda_\gamma^+\|_{L^2(\Omega)} \leq \|f + g'\|_{L^2(\Omega)}.$$

For the discussion of A^- and λ_γ^- , we use

$$-u'' - g' + \lambda_\gamma^+ + \lambda_\gamma^- = f - g' \quad \text{in } \Omega, \quad u(0) = u(1) = 0,$$

multiply by $-(u' + g)$ and proceed in the same way. We arrive at

$$\|\lambda_\gamma^-\|_{L^2(\Omega)} = \gamma \|u' + g\|_{L^2(\Omega)} \leq \|f - g'\|_{L^2(\Omega)}.$$

From here, (2.8) follows immediately since λ_γ^- and λ_γ^+ are orthogonal in $L^2(\Omega)$. \square

Theorem 2.4 (convergence as $\gamma \rightarrow \infty$). *The unique solutions u_γ of (2.2) converge strongly in $H^1(\Omega)$ to the unique solution u of (1.1).*

Proof. Let γ_n be a sequence tending to ∞ , and let u_n be the unique solution of (2.2) for γ_n . Moreover, let λ_n^+ and λ_n^- be defined as in (2.5), and recall that $-u_n'' + \lambda_n^+ + \lambda_n^- = f$ holds in Ω . Lemma 2.3 thus implies boundedness of $\{u_n\}$ in $H^2(\Omega)$, and hence there exists a subsequence (denoted by index n') such that

$$\begin{aligned} u_{n'} &\rightharpoonup \bar{u} && \text{in } H^2(\Omega), & u_{n'} &\rightarrow \bar{u} && \text{in } C^1(\Omega), \\ \lambda_n^+ &\rightharpoonup \bar{\lambda}^+ && \text{in } L^2(\Omega), & \lambda_n^- &\rightharpoonup \bar{\lambda}^- && \text{in } L^2(\Omega). \end{aligned}$$

We immediately have $-\bar{u}'' + \bar{\lambda}^+ + \bar{\lambda}^- = f$ and $\bar{\lambda}^+ \geq 0$ and $\bar{\lambda}^- \geq 0$. It remains to verify the complementarity conditions in (2.1). In view of the boundedness of $\lambda_{n'}^+$ in $L^2(\Omega)$, we obtain

$$\int_0^1 |\max\{0, u_{n'}' - g\}|^2 dx \leq \frac{c}{\gamma_{n'}^2} \xrightarrow{n' \rightarrow \infty} 0.$$

Moreover, the integrand converges uniformly to $|\max\{0, \bar{u}' - g\}|^2$ and hence $\bar{u}' - g \leq 0$ holds in Ω . Next we consider

$$\int_0^1 \lambda_{n'}^+(u_{n'}' - g) dx = \frac{1}{\gamma_{n'}} \int_0^1 |\lambda_{n'}^+|^2 dx \leq \frac{c}{\gamma_{n'}} \xrightarrow{n' \rightarrow \infty} 0.$$

Weak convergence of $\lambda_{n'}^+$ and strong convergence of $u_{n'}'$ in $L^2(\Omega)$ imply that the integral converges to $\int_0^1 \bar{\lambda}^+(\bar{u}' - g) dx$. The integrand is ≤ 0 a.e. and hence $\bar{\lambda}^+(\bar{u}' - g) = 0$ holds in Ω . The remaining complementarity condition which involves $\bar{\lambda}^-$ follows alike. □

2.2. Semi-smooth Newton method

The considerations above motivate the solution of (2.1) by solving a sequence of regularized problems (2.2) for increasing parameters γ . For every fixed $\gamma > 0$, we propose to solve (2.2) by a semi-smooth Newton method, formulated in terms of an active set strategy. The complete algorithm, together with an update of the regularization parameter γ as justified by Theorem 2.4, is provided below. In the numerical examples, we observed an increase of γ by a factor of 10 to be a suitable choice.

Algorithm 2.1 Semi-smooth Newton method in the one-dimensional case.

1: Choose initial u_0 and $\gamma \geq 0$ and set $n = 0$

2: **while** not converged **do**

3: **while** not converged **do**

4: Set

$$A_n^+ = \{x \in \Omega : u_n' - g > 0\} \quad \text{and} \quad A_n^- = \{x \in \Omega : -u_n' - g > 0\}$$

5: Solve for u_{n+1}

$$-u'' + \gamma \chi_{A_n^+}(u_{n+1}' - g) + \gamma \chi_{A_n^-}(-u_{n+1}' - g) = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \Gamma$$

6: Increase n

7: **end while**

8: Increase γ

9: **end while**

Lemma 2.5 (well-posedness of Step 5). *Let A^+, A^- be measurable subsets of Ω and $\gamma \geq 0$. Then the problem in Step 5 has a unique solution $u \in H^2(\Omega) \cap H_0^1(\Omega)$.*

Proof. The existence of a solution can be proved as in Proposition 2.2 where now $v = Tu$ is defined by

$$-v'' + \gamma \chi_{A^+}(u' - g) + \gamma \chi_{A^-}(-u' - g) = f \quad \text{in } \Omega, \quad v = 0 \quad \text{on } \Gamma.$$

Existence and uniqueness also follow directly from [16], Theorem 2.4 and Corollary. □

Proposition 2.6 (superlinear convergence of the active set loop for fixed γ). *For fixed γ , the inner loop (Steps 3–7 in Algorithm 2.1 converges locally superlinearly in $H^2(\Omega) \cap H_0^1(\Omega)$).*

Proof. We define the linear solution operator $Sf = (-\Delta)^{-1}f$, which maps $L^2(\Omega)$ to $H^2(\Omega) \cap H_0^1(\Omega)$, and

$$Gf = (Sf)',$$

which maps $L^2(\Omega)$ to $H^1(\Omega)$. Setting $\lambda := \lambda^+ + \lambda^-$ as in (2.5), we see that (2.2) is equivalent to $u = S(f - \lambda)$, and thus equivalent to

$$\lambda = \gamma \max\{0, G(f - \lambda) - g\} + \gamma \max\{0, -G(f - \lambda) - g\}. \tag{2.9}$$

Let $F : L^2(\Omega) \rightarrow L^2(\Omega)$ be defined by

$$F(\lambda) = \lambda - \gamma \max\{0, G(f - \lambda) - g\} - \gamma \max\{0, -G(f - \lambda) - g\}.$$

The $\max\{0, \cdot\}$ operator is well-known to be Newton (slantly) differentiable from $L^p(\Omega) \rightarrow L^2(\Omega)$ for any $p > 2$, see, e.g., [8], Proposition 4.1. The necessary smoothing is provided here by the operator G , which maps $L^2(\Omega)$ to $L^\infty(\Omega)$ in our one-dimensional setting. A generalized derivative (slanting function) is given by

$$(DF)(\lambda) = I + \gamma \chi_{A^+}G - \gamma \chi_{A^-}G,$$

where the active sets are

$$A^+ := \{x \in \Omega : G(f - \lambda) - g > 0\}, \quad A^- := \{x \in \Omega : -G(f - \lambda) - g > 0\}.$$

A semi-smooth Newton step towards $F(\lambda) = 0$ is given by $F(\lambda_n) + (DF)(\lambda_n) \delta\lambda = 0$ and $\lambda_{n+1} := \lambda_n + \delta\lambda$. This is equivalent to

$$\lambda_{n+1} - \gamma \chi_{A^+}(G(f - \lambda_{n+1}) - g) - \gamma \chi_{A^-}(-G(f - \lambda_{n+1}) - g) = 0 \tag{2.10}$$

with A^\pm determined from λ_n . Setting (for all n)

$$u_{n+1} = S(f - \lambda_{n+1}),$$

i.e., $-u_{n+1}'' + \lambda_{n+1} = f$, we obtain the following equivalent equation to (2.10):

$$-u_{n+1}'' + \gamma \chi_{A_n^+}(u_{n+1}' - g) + \gamma \chi_{A_n^-}(-u_{n+1}' - g) = f \quad \text{in } \Omega, \quad u_{n+1} = 0 \quad \text{on } \Gamma.$$

This is the iteration Step 5 of our Algorithm 2.1, and the active sets are determined as in Step 4. Hence analyzing the convergence for Algorithm 2.1 with $u \in H^2(\Omega) \cap H_0^1(\Omega)$ is equivalent to analyzing the semi-smooth Newton iteration applied to $F(\lambda) = 0$ in $L^2(\Omega)$.

Lemma 2.5 shows that Step 5, or equivalently (2.10), admits a unique solution at every iterate. Lemma 2.7, which is provided below, proves that $\|(DF)(\lambda)^{-1}\|_{\mathcal{L}(L^2(\Omega))}$ is uniformly bounded with respect to $\lambda \in L^2(\Omega)$.

To see this, note that $y = (DF)(\lambda)^{-1}z$ is equivalent to $y + \gamma \chi_{A^+} G y - \gamma \chi_{A^-} G y = z$. Letting $v = S y$, we see that this is in turn equivalent to

$$-v'' + \gamma \chi_{A^+} v' - \gamma \chi_{A^-} v' = z \quad \text{in } \Omega, \quad v = 0 \quad \text{on } \Gamma.$$

We obtain $\|y\|_{L^2(\Omega)} = \|\Delta v\|_{L^2(\Omega)} \leq \|v\|_{H^2(\Omega)} \leq c \|z\|_{L^2(\Omega)}$ by Lemma 2.7, with c independent of A^\pm . The local superlinear convergence is then a standard conclusion for generalized Newton methods, see [3], Remark 2.7, or [8], Theorem 1.1. \square

Lemma 2.7 (uniform boundedness). *The solution of Step 5 satisfies*

$$\|u\|_{H^2(\Omega)} \leq c (\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}),$$

where c does not depend on A^\pm .

Proof. We abbreviate $w = f + \chi_{A^+} g + \chi_{A^-} g$. By Lemma 2.5, the unique solution u of Step 5 is of class C^1 . Define $z = u'$ and consider

$$-z' + \gamma \chi_{A^+} z - \gamma \chi_{A^-} z = w \quad \text{on } (0, 1). \quad (2.11)$$

By the intermediate value theorem, there exists $m \in (0, 1)$ such that $u'(m) = z(m) = 0$. Multiplication of (2.11) by z yields

$$\frac{1}{2} \frac{d}{dx}(z^2) = \gamma \chi_{A^+} z^2 - \gamma \chi_{A^-} z^2 - z w \leq \gamma z^2 + \frac{z^2}{2} + \frac{w^2}{2} \quad \text{on } (0, 1).$$

By integration over $[m, x]$ for $m \leq x \leq 1$, we obtain

$$z^2(x) \leq (2\gamma + 1) \int_m^x z^2(s) ds + \|w\|_{L^2}^2.$$

Gronwall's lemma implies that

$$z^2(x) \leq \|w\|_{L^2}^2 e^{(2\gamma+1)(x-m)} \quad \text{for all } x \in [m, 1].$$

Similarly, the application of Gronwall's lemma on $[x, m]$ yields

$$z^2(x) \leq \|w\|_{L^2}^2 e^{(2\gamma+1)(m-x)} \quad \text{for all } x \in [0, m].$$

The number m varies only in $(0, 1)$, so we finally get

$$z^2(x) \leq \|w\|_{L^2}^2 e^{(2\gamma+1)x+1} \quad \text{for all } x \in [0, 1],$$

which implies that

$$\|u'\|_{L^\infty} \leq c \|w\|_{L^2},$$

where c is independent of A^\pm . Integration over $[0, 1]$ yields the estimate

$$\|u\|_{W^{1,\infty}(\Omega)} \leq c (\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}).$$

Our claim follows by considering

$$-u'' = -\gamma \chi_{A^+} (u' - g) - \gamma \chi_{A^-} (-u' - g) + f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \Gamma$$

and applying a standard *a priori* estimate. \square

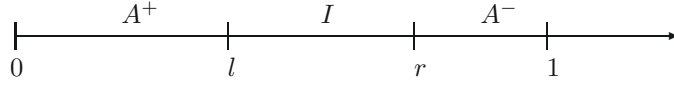


FIGURE 1. Assumed structure of the active and inactive sets for the primal-dual active set method (Algorithm 4.1).

2.3. An active set approach without regularization

Here we briefly discuss an alternative active set approach without regularization. We begin with the observation that (2.1) can be equivalently formulated in the following form:

$$\begin{aligned} -u'' + \lambda^+ + \lambda^- &= f && \text{a.e. in } \Omega, \\ \lambda^+ &= \max\{0, \lambda^+ + \bar{\gamma}(u' - g)\} && \text{a.e. in } \Omega, \\ \lambda^- &= \max\{0, \lambda^- + \bar{\gamma}(-u' - g)\} && \text{a.e. in } \Omega, \end{aligned} \tag{2.12}$$

where $\bar{\gamma}$ is an arbitrary positive constant. When the active sets at the solution

$$A^+ = \{x \in \Omega : u' - g > 0\} \quad \text{and} \quad A^- = \{x \in \Omega : -u' - g > 0\}$$

are known, problem (2.12) reduces to the solution of $-u'' = f$ on the inactive set with suitable boundary conditions. Here we assume the structure of the active and inactive sets depicted in Figure 1, which is typical for portfolio optimization problems. On the active sets, the solution is given by

$$u^+(x) = \tilde{G}(x) - \tilde{G}(0) \quad \text{in } A^+, \quad u^-(x) = -\tilde{G}(x) + \tilde{G}(1) \quad \text{in } A^-, \tag{2.13}$$

where \tilde{G} is any anti-derivative of g . In order to enforce continuity of u and u' at $x \in \{l, r\}$, we impose boundary conditions of Robin type there:

$$\begin{aligned} -\kappa u'(l) + u(l) &= -\kappa g(l) + \tilde{G}(l) - \tilde{G}(0), \\ \kappa u'(r) + u(r) &= -\kappa g(r) - \tilde{G}(r) + \tilde{G}(1), \end{aligned}$$

with some $\kappa > 0$. If the structure in Figure 1 is not assumed and one connected component of A lies strictly in the interior of Ω , the solution on this component involves unknown integration constants.

Numerical experience shows that the method based on regularization (Algorithm 2.1) is superior to the approach described here. Moreover, the regularization technique can be directly extended to the multi-dimensional case. In Section 4, we provide a comparison between the two approaches.

3. ANALYSIS OF THE MULTI-DIMENSIONAL CASE

Assumption 3.1. *Suppose that $\Omega \subset \mathbb{R}^d$, $d \geq 2$ is a bounded domain with a smooth boundary Γ . We assume that $f \in C^1(\bar{\Omega})$, $f > 0$ and $g \in C^2(\bar{\Omega})$, $g \geq 0$.*

3.1. Regularization

Similarly as before, we consider instead of (1.1) the regularized formulation

$$-\Delta u + \gamma \max\{0, |\nabla u|^2 - g^2\} = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \Gamma \tag{3.1}$$

for an increasing sequence of parameters $\gamma \geq 0$.

Proposition 3.2. *Problem (3.1) has a unique solution $u \in W^{3,q}(\Omega) \cap H_0^1(\Omega)$, $1 \leq q < \infty$.*

Proof. Existence. Let $q > 2d$. We use again Leray-Schauder theory and define $T : W_0^{1,4q}(\Omega) \rightarrow W_0^{1,4q}(\Omega)$ by the relation $v = Tu$ iff

$$-\Delta v + \gamma \max\{0, |\nabla u|^2 - g^2\} = f \quad \text{in } \Omega, \quad v = 0 \quad \text{on } \Gamma.$$

For $u \in W_0^{1,4q}(\Omega)$, we have $|\nabla u|^2 \in L^{2q}(\Omega)$, and the same regularity holds for $f - \gamma \max\{0, |\nabla u|^2 - g^2\}$. Thus $v \in W^{2,2q}(\Omega)$, which embeds compactly into $W^{1,4q}(\Omega)$, hence T is compact. Suppose now that $s \in [0, 1]$ and that u solves $u = sT(u)$, i.e.,

$$-\Delta u + s\gamma \max\{0, |\nabla u|^2 - g^2\} = sf \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \Gamma.$$

By the maximum principle, $0 \leq u \leq (-\Delta)^{-1}f$ holds, which implies

$$\|u\|_{L^\infty(\Omega)} \leq c \|f\|_{C^1(\Omega)},$$

independent of s . Testing the equation with u yields the *a priori* estimate $\|u\|_{H^1(\Omega)} \leq c \|f\|_{L^2(\Omega)}$ like in the one-dimensional case, but this is not strong enough for our purpose. Instead, we shall now derive a bound for $\|\nabla u\|_{L^\infty(\Omega)}$, independent of s .

We observe that $-\Delta u = sf - s\gamma \max\{0, |\nabla u|^2 - g^2\} \in L^{2q}(\Omega)$, hence $u \in W^{2,2q}(\Omega)$. This implies that the right hand side is even in $W^{1,q}(\Omega)$ and hence $u \in W^{3,q}(\Omega)$ and $u \in C^2(\Omega)$ hold. We define the auxiliary function

$$z = |\nabla u|^2 - 3u,$$

which is of class $W^{2,q/2}(\Omega)$ and thus $C^1(\Omega)$. Once a pointwise upper bound for z is established which depends only on the data, we obtain a bound for $\|\nabla u\|_{L^\infty(\Omega)}$. We therefore define \bar{x} to be the global maximum of z on $\bar{\Omega}$ and distinguish three cases:

- (1) If $\bar{x} \in I = \{x \in \Omega : |\nabla u| \leq g\}$, then

$$z(\bar{x}) = |\nabla u(\bar{x})|^2 - 3u(\bar{x}) \leq g(\bar{x})^2$$

and thus $\|\nabla u\|_{L^\infty(\Omega)}^2 \leq \|g\|_{L^\infty(\Omega)}^2 + 3\|u\|_{L^\infty(\Omega)}$ holds.

- (2) Suppose that $\bar{x} \in A = \{x \in \Omega : |\nabla u| > g\}$. Since A is open, we infer from the Bony maximum principle and the definition of z that

$$\begin{aligned} 0 &\leq \lim \operatorname{ess\,sup}_{x \rightarrow \bar{x}} (-\Delta z) \\ &= \lim \operatorname{ess\,sup}_{x \rightarrow \bar{x}} \left(-2 \sum_{i,k=1}^d ((u_{x_k x_i})^2 + u_{x_k} u_{x_k x_i x_i}) + 3 \Delta u \right) \\ &\leq \lim \operatorname{ess\,sup}_{x \rightarrow \bar{x}} \left(-2 \sum_{k=1}^d u_{x_k} \Delta(u_{x_k}) + 3 \Delta u \right) \end{aligned}$$

holds. On the open set A , we have the equation

$$-\Delta u + s\gamma (|\nabla u|^2 - g^2) = sf \quad \text{in } A.$$

Plugging in and using the continuity of the expression obtained, we get

$$0 \leq -2 \sum_{k=1}^d u_{x_k} [s\gamma (|\nabla u|^2 - g^2)_{x_k} - s f_{x_k}] + 3 [s\gamma (|\nabla u|^2 - g^2) - s f],$$

where everything is evaluated at \bar{x} , and we omit this argument throughout. We may cancel s since for $s = 0$, $u \equiv 0$ holds, and omit $-3\gamma g^2 - 3f \leq 0$ to obtain

$$0 \leq -2\gamma(\nabla u) \cdot [\nabla(|\nabla u|^2) - \nabla(g^2)] + 2(\nabla u) \cdot (\nabla f) + 3\gamma|\nabla u|^2.$$

Since z attains its maximum at \bar{x} , $\nabla z = 0$ holds, and thus

$$\nabla(|\nabla u|^2) - 3\nabla u = 0.$$

This yields

$$\begin{aligned} 0 &\leq -3\gamma|\nabla u|^2 + 2\gamma(\nabla u) \cdot \nabla(g^2) + 2(\nabla u) \cdot (\nabla f) \\ &\leq -3\gamma|\nabla u|^2 + \gamma|\nabla u|^2 + \gamma|\nabla(g^2)|^2 + |\nabla u|^2 + |\nabla f|^2. \end{aligned}$$

For $\gamma \geq 1$, this implies

$$|\nabla u(\bar{x})|^2 \leq \|g\|_{W^{1,\infty}(\Omega)}^4 + \|f\|_{W^{1,\infty}(\Omega)}^2.$$

Finally,

$$|\nabla u(x)|^2 - 3u(x) = z(x) \leq z(\bar{x}) = |\nabla u(\bar{x})|^2 - 3u(\bar{x})$$

yields the desired pointwise bound for $|\nabla u(x)|$ on Ω .

(3) If $\bar{x} \in \Gamma$, we consider

$$-\Delta w = f \quad \text{in } \Omega, \quad w = 0 \quad \text{on } \Gamma.$$

Since $-\Delta u \leq sf \leq f$ holds in Ω , the maximum principle implies that $0 \leq u \leq w$ in Ω . Therefore, we obtain

$$\frac{\partial w}{\partial n} \leq \frac{\partial u}{\partial n} \leq 0 \quad \text{on } \Gamma.$$

Choose $p > d$. The *a priori* estimate $\|w\|_{W^{2,p}(\Omega)} \leq c\|f\|_{L^\infty(\Omega)}$, the continuity of the normal trace operator from $W^{2,p}(\Omega)$ into $W^{1-1/p,p}(\Gamma)$ and the continuity of the embedding of the latter space into $L^\infty(\Gamma)$ imply $\|\frac{\partial u}{\partial n}\|_{L^\infty(\Gamma)} \leq \|\frac{\partial w}{\partial n}\|_{L^\infty(\Gamma)} \leq c\|f\|_{L^\infty(\Omega)}$. In tangential directions, we have $\frac{\partial u}{\partial \tau} = 0$ on Γ , thus $|\nabla u| \leq c\|f\|_{L^\infty(\Omega)}$ holds on Γ .

Altogether, we obtain a bound

$$\|u\|_{W^{1,\infty}(\Omega)} \leq c \left(\|f\|_{W^{1,\infty}(\Omega)} + \|f\|_{L^\infty(\Omega)}^{1/2} + \|g\|_{L^\infty(\Omega)} + \|g\|_{W^{1,\infty}(\Omega)}^2 \right),$$

where c does not depend on s nor γ . The Leray-Schauder theorem now implies the existence of a solution $u \in W_0^{1,4q}(\Omega)$ of (3.1). From the above bootstrapping argument, moreover $u \in W^{3,q}(\Omega)$ and $u \in C^2(\Omega)$ follow.

Uniqueness. We employ a similar argument as in the proof of Proposition 2.2. Suppose that $u \neq v$ are two solutions of (3.1). Without loss of generality, $u - v$ attains a positive maximum in Ω . Hence there exists $\varepsilon \in (0, 1)$ and $x_1 \in \Omega$ such that $u - v - \varepsilon u$ assumes a positive maximum in x_1 , which implies

$$(1 - \varepsilon)\nabla u(x_1) = \nabla v(x_1). \tag{3.2}$$

Moreover, the maximum principle implies

$$0 \leq -\Delta(u - v - \varepsilon u)(x_1) = C(x_1) - \varepsilon f(x_1), \tag{3.3}$$

where $-C(x) = (1 - \varepsilon)\gamma \max\{0, |\nabla u|^2 - g^2\} - \gamma \max\{0, |\nabla v|^2 - g^2\}$. We distinguish cases, where all functions are evaluated at x_1 .

- (1) If $|\nabla u| \leq g$, then by (3.2) also $|\nabla v| \leq g$ holds, and (3.3) yields the contradiction $0 \leq -\varepsilon f$.
- (2) If $|\nabla u| > g$ and $|\nabla v| \leq g$, then (3.3) yields the contradiction $0 \leq (1 - \varepsilon)\gamma(|\nabla u|^2 - g^2) \leq -\varepsilon f$.

(3) The case $|\nabla u| \leq g$ and $|\nabla v| > g$ violates (3.2).

(4) If $|\nabla u| > g$ and $|\nabla v| > g$, then (3.3) yields

$$(1 - \varepsilon)\gamma(|\nabla u|^2 - g^2) - \gamma(|\nabla v|^2 - g^2) \leq -\varepsilon f.$$

As $(1 - \varepsilon)^2 < 1 - \varepsilon$ holds, we infer from (3.2)

$$\gamma|\nabla v|^2 - (1 - \varepsilon)\gamma g^2 - \gamma|\nabla v|^2 + \gamma g^2 \leq -\varepsilon f,$$

which leads to the contradiction $\varepsilon\gamma g^2 \leq -\varepsilon f$.

Hence the solution u of (3.1) must be unique. \square

We may note that for the previous proposition, we only require $g \in W^{1,\infty}(\Omega)$. For the subsequent discussion, we denote the unique solution of (3.1) by u_γ . Similarly to the one-dimensional case, we introduce

$$\lambda_\gamma := \gamma \max\{0, |\nabla u_\gamma|^2 - g^2\}. \quad (3.4)$$

Lemma 3.3 (*a priori estimates for u_γ*). *The unique solution of (3.1) satisfies the following a priori bounds:*

$$\|u_\gamma\|_{W^{1,\infty}(\Omega)} \leq c \left(\|f\|_{C^1(\Omega)} + \|f\|_{C(\Omega)}^{1/2} + \|g\|_{C(\Omega)} + \|g\|_{C^1(\Omega)}^2 \right) \quad (3.5)$$

$$0 \leq u_\gamma \leq (-\Delta)^{-1}f \quad (3.6)$$

$$\|u_\gamma\|_{W^{2,p}(\Omega')} \leq C_1(p, \Omega') \quad \text{for all } 1 \leq p < \infty \text{ and all } \Omega' \subset\subset \Omega \quad (3.7)$$

$$\|\lambda_\gamma\|_{L^\infty(\Omega')} \leq C_2(\Omega'). \quad (3.8)$$

All constants are independent of γ , and C_1 and C_2 depend on $\|u_\gamma\|_{C^1(\Omega)}$, $\|f\|_{C^1(\Omega)}$ and $\|g\|_{C^2(\Omega)}$.

Proof. The bounds (3.5)–(3.6) were shown in the proof of Proposition 3.2. To show (3.7), we begin by proving a bound for $\|\Delta u\|_{L^\infty(\Omega)}$. (We omit the index γ throughout the proof.) We distinguish between

$$A = \{x \in \Omega : |\nabla u| > g\} \quad \text{and} \quad I = \{x \in \Omega : |\nabla u| \leq g\}.$$

On the inactive set, (3.1) immediately implies

$$\|\Delta u\|_{L^\infty(I)} \leq \|f\|_{L^\infty(I)}. \quad (3.9)$$

Now let $\Omega' \subset\subset \Omega$ and let ζ be a smooth function on $\overline{\Omega}$ satisfying $\zeta = 0$ on Γ , $\zeta = 1$ on Ω' . We consider the function

$$v = \zeta^2 \gamma (|\nabla u|^2 - g^2) \quad \text{in } \overline{A}.$$

Note that $v \geq 0$ and $v \in W^{2,q}(\Omega)$ for $1 \leq q < \infty$, thus $v \in C^1(\Omega)$ hold. Suppose that v attains its global maximum in $x_1 \in \overline{A}$. By construction, $v = 0$ on $\partial A \cup \Gamma$, and thus $x_1 \in A \setminus \Gamma$, and we infer

$$\nabla v(x_1) = 0 \quad \text{and} \quad \lim \operatorname{ess} \sup_{x \rightarrow x_1} \Delta v(x) \leq 0 \quad (3.10)$$

from Bony's maximum principle. (Here we use the $W^{2,\infty}$ regularity of g .) We also have

$$-\zeta^2 \Delta u + v = \zeta^2 f \quad \text{a.e. in } A.$$

Differentiation implies that

$$-\zeta^2 \Delta u_{\gamma, x_k} = (\zeta^2 f)_{x_k} + 2\zeta \zeta_{x_k} \Delta u - v_{x_k} \quad \text{a.e. in } A. \quad (3.11)$$

From the definition of v we infer

$$\begin{aligned} v_{x_i} &= \zeta^2 \gamma \left(2 \sum_k u_{x_k} u_{x_k x_i} - (g^2)_{x_i} \right) + 2 \zeta \zeta_{x_i} \gamma (|\nabla u|^2 - g^2) \\ v_{x_i x_i} &= \zeta^2 \gamma \left(2 \sum_k (u_{x_k x_i})^2 + 2 \sum_k u_{x_k} u_{x_k x_i x_i} - (g^2)_{x_i x_i} \right) \\ &\quad + 4 \zeta \zeta_{x_i} \gamma \left(2 \sum_k u_{x_k} u_{x_k x_i} - (g^2)_{x_i} \right) + 2 \gamma (\zeta_{x_i}^2 + \zeta \zeta_{x_i x_i}) (|\nabla u|^2 - g^2) \end{aligned}$$

a.e. in A . We calculate

$$\begin{aligned} \Delta v &= \zeta^2 \gamma \left(2 |D^2 u|^2 + 2 \sum_k u_{x_k} \Delta u_{x_k} - \Delta(g^2) \right) \\ &\quad + 4 \zeta \sum_i \zeta_{x_i} \gamma \left(2 \sum_k u_{x_k} u_{x_k x_i} - (g^2)_{x_i} \right) + 2 \gamma \left(\sum_i \zeta_{x_i}^2 + \zeta \Delta \zeta \right) (|\nabla u|^2 - g^2), \end{aligned}$$

where we used the abbreviation $|D^2 u|^2 := \sum_{i,k} (u_{x_i x_k})^2$. By (3.11), we find

$$\begin{aligned} \Delta v &= \gamma \left(2 \zeta^2 |D^2 u|^2 - 2 \sum_k u_{x_k} (\zeta^2 f)_{x_k} - 4 \zeta \sum_k \zeta_{x_k} u_{x_k} \Delta u + \sum_k v_{x_k} u_{x_k} - \zeta^2 \Delta(g^2) \right) \\ &\quad + 4 \zeta \sum_i \zeta_{x_i} \gamma \left(2 \sum_k u_{x_k} u_{x_k x_i} - (g^2)_{x_i} \right) + 2 \gamma \left(\sum_i \zeta_{x_i}^2 + \zeta \Delta \zeta \right) (|\nabla u|^2 - g^2). \end{aligned}$$

Realizing that all terms on the right hand side are continuous functions on A , we can evaluate (3.10) and obtain

$$\begin{aligned} 0 &\geq \Delta v(x_1) \\ &= \gamma \left(2 \zeta^2 |D^2 u|^2 - 2 \sum_k u_{x_k} (\zeta^2 f)_{x_k} - 4 \zeta \sum_k \zeta_{x_k} u_{x_k} \Delta u - \zeta^2 \Delta(g^2) \right) \\ &\quad + 4 \zeta \sum_i \zeta_{x_i} \gamma \left(2 \sum_k u_{x_k} u_{x_k x_i} - (g^2)_{x_i} \right) + 2 \gamma \left(\sum_i \zeta_{x_i}^2 + \zeta \Delta \zeta \right) (|\nabla u|^2 - g^2), \end{aligned}$$

where everything is evaluated at x_1 . Consequently, there exist C_1, C_2 , depending on $\|\zeta\|_{C^2(\Omega)}$, $\|f\|_{C^1(\Omega)}$, $\|g\|_{C^2(\Omega)}$, $\|u\|_{W^{1,\infty}(\Omega)}$, such that

$$\zeta^2 |D^2 u|^2 \leq c_1 + c_2 \zeta |D^2 u| \quad \text{in } x_1$$

and thus

$$\zeta |D^2 u| \leq c \quad \text{in } x_1.$$

Using the estimate $|\Delta u|^2 \leq d |D^2 u|^2$, we derive

$$\zeta |\Delta u| \leq c \sqrt{d} \quad \text{in } x_1.$$

Finally, we conclude

$$0 \leq v(x) = \zeta^2 (f + \Delta u)(x) \leq v_1(x) = \zeta^2 (f + \Delta u)(x_1) \quad \text{for a.a. } x \in A,$$

and thus

$$|\zeta \Delta u(x)| \leq \|f\|_{L^\infty(\Omega)} + \|\zeta\|_{C(\Omega)} c \sqrt{d} \quad \text{for a.a. } x \in A$$

and

$$\|\Delta u\|_{L^\infty(\Omega' \cap A)} \leq c.$$

Together with (3.9) this implies that

$$\|\Delta u\|_{L^\infty(\Omega')} \leq c.$$

By interior estimates for elliptic equations [13], Chapters III.7 and III.8, the final estimate (3.7) follows. (3.8) follows from $\lambda = f + \Delta u$. \square

Theorem 3.4 (convergence as $\gamma \rightarrow \infty$). *The unique solutions u_γ of (3.1) converge to the unique solution $u \in W^{1,\infty}(\Omega) \cap W_{\text{loc}}^{2,p}(\Omega)$ of (1.1) in the following sense:*

$$u_\gamma \rightarrow u \quad \text{in } C(\overline{\Omega}) \tag{3.12}$$

$$\nabla u_\gamma \rightarrow \nabla u \quad \text{in } W^{1,p}(\Omega') \text{ for all } 1 \leq p < \infty, \text{ and in } C(\Omega') \tag{3.13}$$

$$\Delta u_\gamma \rightharpoonup \Delta u \quad \text{in } L^p(\Omega') \text{ for all } 1 \leq p < \infty, \tag{3.14}$$

for every $\Omega' \subset\subset \Omega$.

Proof. Let γ_n be a sequence tending to ∞ , and let $u_n \in W^{1,\infty}(\Omega) \cap W_{\text{loc}}^{2,p}(\Omega)$ be the unique solution of (3.1) for γ_n . Moreover, let $\Omega' \subset\subset \Omega$ and choose $p > d$. By (3.5), there exists a subsequence satisfying $u_{n'} \rightharpoonup \bar{u}$ in $W^{1,p}(\Omega)$ for some $p > d$. By compactness of the embedding, $u_{n'} \rightarrow \bar{u}$ in $C(\overline{\Omega})$ follows. From (3.7) and the compact embedding of $W^{2,p}(\Omega')$ into $W^{1,p}(\Omega')$, we get $\nabla u_{n'} \rightarrow \nabla \bar{u}$ in $W^{1,p}(\Omega')$ and thus in $C(\Omega')$, and moreover $\nabla u_{n'} \rightharpoonup^* \nabla \bar{u}$ in $W^{1,\infty}(\Omega)$. (For convenience, we denote sub-subsequences by the same index n' .) Finally, (3.7) also implies $\Delta u_{n'} \rightharpoonup \Delta \bar{u}$ in $L^p(\Omega')$, thus we have (3.12)–(3.14) for subsequences.

We proceed to show that \bar{u} satisfies (1.1). From (3.1) we obtain

$$\int_{\Omega} (-\Delta u_{n'} - f) \varphi \, dx \leq 0 \quad \text{for all } \varphi \in L^2(\Omega), \varphi \geq 0.$$

Weak convergence implies that

$$\int_{\Omega} (-\Delta \bar{u} - f) \varphi \, dx \leq 0 \quad \text{for all } \varphi \in L^2(\Omega), \varphi \geq 0$$

and thus

$$-\Delta \bar{u} \leq f \quad \text{a.e. in } \Omega.$$

Moreover, by (3.8) and the definition $\lambda_\gamma = \gamma \max\{0, |\nabla u_\gamma|^2 - g^2\}$, we get

$$\max\{0, |\nabla u_{n'}|^2 - g^2\} \leq \frac{C_2(\Omega')}{\gamma_{n'}} \xrightarrow{n' \rightarrow \infty} 0 \quad \text{on } \Omega'.$$

Since $\nabla u_{n'} \rightarrow \nabla \bar{u}$ in $C(\Omega')$, the left hand side converges to $\max\{0, |\nabla \bar{u}|^2 - g^2\}$ and we obtain

$$\max\{0, |\nabla \bar{u}|^2 - g^2\} \leq 0.$$

This implies that

$$\max\{-\Delta \bar{u} - f, |\nabla \bar{u}| - g\} \leq 0 \quad \text{a.e. in } \Omega,$$

To prove that equality holds, assume that $|\nabla\bar{u}(x_0)| < g(x_0)$ for some $x_0 \in \text{int } \Omega'$ and indeed in a neighborhood $N(x_0) \subset \Omega'$. In view of $\nabla u_{n'} \rightarrow \nabla\bar{u}$ uniformly in Ω' , we have $|\nabla u_{n'}(x)| < g(x)$ for all $x \in N(x_0)$ and sufficiently large n' . Hence (3.1) implies that

$$-\Delta u_{n'} = f \quad \text{in } N(x_0)$$

holds for sufficiently large n' . Passing to the weak limit in $L^2(N(x_0))$, we get

$$-\Delta\bar{u} = f \quad \text{in } N(x_0).$$

To summarize, the limit $\bar{u} \in W^{1,\infty}(\Omega) \cap W_{\text{loc}}^{2,p}(\Omega)$ of the subsequence $u_{n'}$ solves (1.1). As the solution of (1.1) is unique [5], the convergence extends to the whole sequence. \square

3.2. Semi-smooth Newton method

This section is devoted to the analysis of the semi-smooth Newton method for solving (3.1). We start with a technical lemma which is used repeatedly throughout this section.

Lemma 3.5. *Consider*

$$-\Delta y + b \cdot \nabla y = f \quad \text{in } \Omega, \quad y = 0 \quad \text{on } \Gamma \tag{3.15}$$

in a bounded domain $\Omega \subset \mathbb{R}^d$ with sufficiently regular boundary Γ such that $(-\Delta)$ is a homeomorphism from $H^2(\Omega) \cap H_0^1(\Omega)$ to $L^2(\Omega)$. If $b \in [L^q(\Omega)]^d$ with $q > d$, then for every $f \in L^2(\Omega)$, there exists a unique solution y to (3.15). Moreover, there exists $K = K(b)$ such that

$$\|y\|_{H^2(\Omega)} \leq K(b) \|f\|_{L^2(\Omega)} \tag{3.16}$$

holds for all $f \in L^2(\Omega)$.

Proof. The verification is quite standard and is included for the sake of completeness. Consider the operator $T : L^2(\Omega) \rightarrow L^2(\Omega)$, defined by

$$Tw = w + b \cdot \nabla(-\Delta)^{-1}w.$$

Since $w \mapsto \nabla(-\Delta)^{-1}w$ is compact from $H^1(\Omega)$ to $[L^p(\Omega)]^d$, for every $p \in [1, \frac{2d}{d-2})$, and since for $b \in [L^q(\Omega)]^d$, $q > d$, the mapping $v \mapsto bv$ is bounded from $[L^{\tilde{p}}(\Omega)]^d$ to $L^2(\Omega)$ for some $\tilde{p} \in [1, \frac{2d}{d-2})$, it follows that T is a compact perturbation of the identity from $L^2(\Omega)$ to itself. Moreover, T is injective. In fact, setting $v = (-\Delta)^{-1}w$, the equation $Tw = 0$ is equivalent to

$$-\Delta v + b \cdot \nabla v = 0 \quad \text{in } \Omega, \quad v = 0 \quad \text{on } \Gamma.$$

This implies that $v = 0$ by a corollary to the maximum principle [16], Theorem 2.4 and Corollary, and hence $w = 0$. By Fredholm's alternative, T is continuously invertible from $L^2(\Omega)$ to itself, and (3.16) follows. \square

Turning to (3.1), note that this equation is equivalent to

$$F(\lambda) = \lambda - \gamma \max\{0, |\nabla S(f - \lambda)|^2 - g^2\} = 0.$$

Here $S = (-\Delta)^{-1} : L^2(\Omega) \rightarrow H^2(\Omega) \cap H_0^1(\Omega)$, and F maps $L^2(\Omega)$ into itself. The relationship between λ and u is given by

$$u = S(f - \lambda).$$

The semi-smooth Newton step towards $F(\lambda) = 0$ reads

$$\delta\lambda + 2\gamma\chi_{A_n}\nabla S(f - \lambda_n)\nabla S\delta\lambda = -\lambda_n + \gamma \max\{0, |\nabla S(f - \lambda_n)|^2 - g^2\}, \tag{3.17}$$

or, in terms of the variable u ,

$$\begin{aligned} -\Delta\delta u + 2\gamma\chi_{A_n}\nabla u_n \cdot \nabla\delta u &= \Delta u_n - \gamma \max\{0, |\nabla u_n|^2 - g^2\} + f & \text{in } \Omega \\ \delta u &= 0 & \text{on } \Gamma. \end{aligned} \quad (3.18)$$

Let us henceforth assume $d \leq 3$. Then $H^1(\Omega) \hookrightarrow L^6(\Omega)$ and $\chi\nabla u_n \in [L^2(\Omega)]^d$ if $u_n \in H^2(\Omega) \cap H_0^1(\Omega)$. Lemma 3.5 implies the well-posedness of the Newton step and the H^2 regularity of the iterate $u_{n+1} = u_n + \delta u$. We are now prepared to define the semi-smooth Newton algorithm for the solution of (3.1) in the multi-dimensional case. As in the one-dimensional setting, we embed it into an update loop for the regularization parameter γ , see Algorithm 3.1.

Algorithm 3.1 Semi-smooth Newton method in the multi-dimensional case.

- 1: Choose initial u and $\gamma \geq 0$ and set $n = 0$
- 2: **while** not converged **do**
- 3: **while** not converged **do**
- 4: Set

$$A_n = \{x \in \Omega : |\nabla u_n| > g\}$$

- 5: Solve for $\delta u \in H^2(\Omega) \cap H_0^1(\Omega)$

$$\begin{aligned} -\Delta\delta u + 2\gamma\chi_{A_n}\nabla u_n \cdot \nabla\delta u &= \Delta u_n - \gamma \max\{0, |\nabla u_n|^2 - g^2\} + f & \text{in } \Omega \\ \delta u &= 0 & \text{on } \Gamma. \end{aligned}$$

- 6: Update $u_{n+1} = u_n + \delta u$ and increase n
 - 7: **end while**
 - 8: Increase γ
 - 9: **end while**
-

Before we turn to the convergence analysis of Algorithm 3.1, we need to introduce a technical assumption. We consider the following equation which corresponds to the formal linearization of (3.1) at u_γ :

$$-\Delta v + 2\gamma\chi_{A_\gamma}\nabla u_\gamma \cdot \nabla v = h \quad \text{in } \Omega, \quad v = 0 \quad \text{on } \Gamma.$$

Here, $A_\gamma = \{x \in \Omega : |\nabla u_\gamma| > g\}$. Since $u_\gamma \in H^2(\Omega)$, the coefficient $2\gamma\chi_{A_\gamma}\nabla u_\gamma \in L^{2d/(d-2)}(\Omega) \subset L^d(\Omega)$ for $d \leq 3$, and Lemma 3.5 implies the existence of a constant \overline{K} such that

$$\|v\|_{H^2(\Omega)} \leq \frac{\overline{K}}{2} \|h\|_{L^2(\Omega)}$$

holds for every $h \in L^2(\Omega)$. Moreover, there exist neighborhoods

$$N(u_\gamma) \subset H^2(\Omega), \quad N(\chi_{A_\gamma}) \subset L^6(\Omega)$$

of u_γ and χ_{A_γ} , such that the unique solution to

$$-\Delta v + 2\gamma\chi\nabla u \cdot \nabla v = h \quad \text{in } \Omega, \quad v = 0 \quad \text{on } \Gamma \quad (3.19)$$

satisfies

$$\|v\|_{H^2(\Omega)} \leq \overline{K} \|h\|_{L^2(\Omega)} \quad (3.20)$$

for every $h \in L^2(\Omega)$ and all $u \in N(u_\gamma)$, $\chi \in N(\chi_{A_\gamma})$. In fact, defining for $(\chi, u) \in L^6(\Omega) \times H^2(\Omega)$ the operator $\mathcal{B}(\chi, u) : H^2(\Omega) \cap H_0^1(\Omega) \rightarrow L^2(\Omega)$ by the left hand side of (3.19), we have

$$\|\mathcal{B}(\chi, u)^{-1}\|_{L(L^2(\Omega), H^2(\Omega))} \leq \bar{K}/2$$

and

$$\begin{aligned} \|\mathcal{B}(\chi_{A_\gamma}, u_\gamma)v - \mathcal{B}(\chi, u)v\|_{L^2(\Omega)} &\leq 2\gamma \left(\|\chi_{A_\gamma} - \chi_A\|_{L^6(\Omega)} \|\nabla u_\gamma\|_{L^6(\Omega)} + \|\chi\|_{L^6(\Omega)} \|\nabla u_\gamma - \nabla u\|_{L^6(\Omega)} \right) \|\nabla v\|_{L^6(\Omega)} \\ &\leq 2\gamma \kappa_6^2 \left(\|\chi_{A_\gamma} - \chi_A\|_{L^6(\Omega)} \|u_\gamma\|_{H^2(\Omega)} + \|\chi\|_{L^6(\Omega)} \|u_\gamma - u\|_{H^2(\Omega)} \right) \|v\|_{H^2(\Omega)}, \end{aligned}$$

where κ_6 is the embedding constant of $H^2(\Omega) \hookrightarrow W^{1,6}(\Omega)$. A perturbation argument now implies (3.20).

Theorem 3.6 (superlinear convergence of the active set loop for fixed γ). *Suppose that $d \leq 3$ holds and that the active sets A_n belonging to the iterates $\{u_n\}$ of the inner loop (Steps 3–7) of Algorithm 3.1 satisfy $\chi_{A_n} \in N(\chi_{A_\gamma})$ for all n . Then the inner loop converges locally superlinearly to u_γ .*

We shall discuss the assumption $\chi_{A_n} \in N(\chi_{A_\gamma})$ further below.

Proof. Step 5 of the algorithm and (3.1) imply the error equation

$$\begin{aligned} -\Delta(u_{n+1} - u_\gamma) + 2\gamma \chi_{A_n} \nabla u_n \cdot \nabla(u_{n+1} - u_\gamma) = \\ \gamma \max\{0, |\nabla u_\gamma|^2 - g^2\} - \gamma \max\{0, |\nabla u_n|^2 - g^2\} + 2\gamma \chi_{A_n} \nabla u_n \cdot \nabla(u_{n+1} - u_\gamma). \end{aligned}$$

In terms of λ , this reads

$$\begin{aligned} \lambda_{n+1} - \lambda_\gamma - 2\gamma \chi_{A_n} \nabla(S(f - \lambda_n)) \cdot \nabla(S(\lambda_{n+1} - \lambda_\gamma)) = \\ \gamma \max\{0, |\nabla S(f - \lambda_\gamma)|^2 - g^2\} - \gamma \max\{0, |\nabla S(f - \lambda_n)|^2 - g^2\} + 2\gamma \chi_{A_n} \nabla S(f - \lambda_n) \cdot \nabla S(\lambda_\gamma - \lambda_n). \end{aligned}$$

By assumption, $\chi_{A_n} \in N(\chi_{A_\gamma})$ holds for all n . We carry out an induction argument and assume that $u_n \in N(u_\gamma)$. Then by $u = S(f - \lambda)$ and (3.20),

$$\begin{aligned} \|\lambda_{n+1} - \lambda_\gamma\|_{L^2(\Omega)} \leq \bar{K} \gamma \left\| \max\{0, |\nabla S(f - \lambda_n)|^2 - g^2\} - \max\{0, |\nabla S(f - \lambda_\gamma)|^2 - g^2\} \right. \\ \left. + 2\chi_{A_n} \nabla S(f - \lambda_n) \cdot \nabla S(\lambda_\gamma - \lambda_n) \right\|_{L^2(\Omega)}. \quad (3.21) \end{aligned}$$

As already recalled in Section 2, the mapping $v \mapsto \max\{0, v\}$ is Newton differentiable from $L^3(\Omega)$ to $L^2(\Omega)$. Moreover, $\lambda \mapsto |\nabla S(f - \lambda)|^2$ is Fréchet differentiable with derivative $-2\nabla S(f - \lambda) \cdot \nabla S(\cdot)$ from $L^2(\Omega)$ to $L^3(\Omega)$. In fact, at a.e. $x \in \Omega$, we have for $\lambda, \bar{\lambda} \in L^2(\Omega)$

$$|\nabla S(f - \lambda)|_{\mathbb{R}^d}^2 - |\nabla S(f - \bar{\lambda})|_{\mathbb{R}^d}^2 + 2\nabla S(\lambda - \bar{\lambda}) \cdot \nabla S(f - \bar{\lambda}) = |\nabla S(\lambda - \bar{\lambda})|_{\mathbb{R}^d}^2$$

and hence

$$\left\| |\nabla S(f - \lambda)|_{\mathbb{R}^d}^2 - |\nabla S(f - \bar{\lambda})|_{\mathbb{R}^d}^2 + 2\nabla S(\lambda - \bar{\lambda}) \cdot \nabla S(f - \bar{\lambda}) \right\|_{L^3(\Omega)} \leq \kappa_3 \|\lambda - \bar{\lambda}\|_{L^2(\Omega)}^2,$$

where κ_3 is the embedding constant of $H^1(\Omega) \hookrightarrow L^3(\Omega)$. The chain rule for Newton differentiable functions [10], Lemma 7.2, implies that the composite function $\lambda \mapsto \max\{0, |\nabla S(f - \lambda)|^2 - g^2\}$ is Newton differentiable from $L^2(\Omega)$ into itself with Newton derivative given by $-2\chi_{\{|\nabla S(f - \lambda)| > g\}} \nabla S(f - \lambda) \cdot \nabla S(\cdot)$. It follows from (3.21) that

$$\|\lambda_{n+1} - \lambda_\gamma\|_{L^2(\Omega)} = o\left(\|\lambda_n - \lambda_\gamma\|_{L^2(\Omega)}\right).$$

If $\|u_0 - u_\gamma\|_{H^2(\Omega)}$, or equivalently $\|\lambda_0 - \lambda_\gamma\|_{L^2(\Omega)}$ is sufficiently small, then the next iterate is again in $N(u_\gamma)$, and superlinear convergence of $\{u_n\}$ to u_γ in $H^2(\Omega) \cap H_0^1(\Omega)$ follows. \square

Let us now turn to a discussion of the requirement that $\chi_{A_n} \in N(\chi_{A_\gamma})$ for all n . The difficulty is due to the fact that $u \mapsto \chi_{\{|\nabla u| > g\}} \in L^q(\Omega)$ for any $q \in [1, \infty]$ is not continuous, as can be seen by considering the sequence $u_n(x) = x/n$ and $g = 0$.

Assumption 3.7. *Suppose that $\text{meas}\{x \in \Omega : |\nabla u_\gamma| = g\} = 0$.*

Note that this is different from requiring that $\text{meas}\{x \in \Omega : |\nabla u| = g\} = 0$, where u is the solution to (1.1), which would exclude interesting cases. Rather, in view of the structure of (3.1), Assumption 3.7 is equivalent to

$$\text{meas}\{x \in \Omega : |\nabla u_\gamma| = g \text{ and } -\Delta u_\gamma = f\} = 0.$$

We can therefore interpret Assumption 3.7 as a strict complementarity condition for the regularized problem.

Lemma 3.8. *With Assumption 3.7 holding, there exists a neighborhood $\widehat{N}(u_\gamma) \subset W^{1,1}(\Omega)$ such that $u \in \widehat{N}(u_\gamma)$ implies that $\chi_A \in N(\chi_{A_\gamma})$, where $A = \{x \in \Omega : |\nabla u| > g\}$.*

Proof. In fact, if this is not the case, there exists a sequence $u_n \rightarrow u_\gamma$ in $W^{1,1}(\Omega)$ such that χ_{A_n} does not converge to χ_{A_γ} in $L^6(\Omega)$, where $A_n = \{x \in \Omega : |\nabla u_n| > g\}$. However, by Assumption 3.7, we have

$$\begin{aligned} \int_{\Omega} |\chi_{A_n} - \chi_{A_\gamma}|^6 dx &= \text{meas}\{x \in \Omega : |\nabla u_\gamma| > g, |\nabla u_n| \leq g\} \\ &\quad + \text{meas}\{x \in \Omega : |\nabla u_\gamma| < g, |\nabla u_n| > g\}. \end{aligned}$$

The expression on the right hand side converges to 0 as $n \rightarrow \infty$ by Lebesgue’s dominated convergence theorem. This is a contradiction. \square

Therefore, Assumption 3.7 implies the local superlinear convergence of the inner loop of Algorithm 3.1 in case $d \leq 3$.

We now develop an alternative to the requirement that $\chi_{A_n} \in N(\chi_{A_\gamma})$ for all n , which uses a uniform regularity assumption on the boundaries of A_n .

Definition 3.9. A bounded set $D \subset \mathbb{R}^d$ with nonempty boundary satisfies the *uniform cone property* if there exist $\vartheta, h > 0$ and $r > 0$ with the following property: For every $x \in \partial D$, there exists a cone $C(\xi_x, \vartheta, h) = \{x \in \mathbb{R}^d : x^\top \xi_x > |x| \cos \vartheta, |x| > h\}$ with $|\xi_x| = 1$, such that $y + C(\xi_x, \vartheta, h) \subset \text{int } D$ for all $y \in B_r(x) \cap \overline{D}$, where $B_r(x)$ is the open ball of radius r centered at x . Further we set

$$L(\Omega, \vartheta, h, r) = \{D \in \overline{\Omega} : D \text{ satisfies the uniform cone property with } (\vartheta, h, r)\}.$$

Theorem 3.10 (superlinear convergence of the active set loop for fixed γ). *Suppose that $d \leq 3$ holds and that the active sets A_n belonging to the iterates $\{u_n\}$ of the inner loop (Steps 3–7) of Algorithm 3.1 satisfy $A_n \in L(\Omega, \vartheta, h, r)$, independent of n . Then the inner loop converges locally superlinearly to u_γ .*

Proof. In view of the proof of Theorem 3.6, it suffices to establish the *a priori* bound

$$\|\mathcal{B}(\chi_{A_n}, u_n)^{-1}\|_{L(L^2(\Omega), H^2(\Omega))} \leq \widehat{K} \tag{3.22}$$

for a constant \widehat{K} independent of n . Let $N(u_\gamma)$ be an open neighborhood of u_γ in $H^2(\Omega) \cap H_0^1(\Omega)$, and set

$$\mathcal{S} = \{(\chi_A, u) : A \in L(\Omega, \vartheta, h, r), u \in N(u_\gamma)\}.$$

By Lemma 3.5, the operator $\mathcal{B}(\chi_A, u)$ is continuously invertible from $H^2(\Omega) \cap H_0^1(\Omega)$ to $L^2(\Omega)$ for every $(\chi_A, u) \in \mathcal{S}$. Moreover, $(\chi, u) \mapsto \mathcal{B}(\chi, u)$ is continuous from $L^{12}(\Omega) \times W^{1,4}(\Omega)$ to $L(H^2(\Omega) \cap H_0^1(\Omega), L^2(\Omega))$. Hence there exist neighborhoods $U(\chi_A) \times U(u) \subset L^{12}(\Omega) \times W^{1,4}(\Omega)$ such that $\mathcal{B}(\psi, v)$ is uniformly continuously invertible for all $(\psi, v) \in U(\chi_A) \times U(u)$. The set

$$\mathcal{C} = \{U(\chi_A) \times U(u) : (\chi_A, u) \in \mathcal{S}\}$$

is an open covering of \mathcal{S} in $L^{12}(\Omega) \times W^{1,4}(\Omega)$. Moreover, \mathcal{S} is a precompact subset of $L^{12}(\Omega) \times W^{1,4}(\Omega)$, see [4], p. 253. Hence \mathcal{C} contains a finite open subcovering and thus there exists \widehat{K} such that

$$\|\mathcal{B}(\chi_A, u)^{-1}\|_{L(L^2(\Omega), H^2(\Omega))} \leq \widehat{K}$$

holds for all $(\chi, u) \in \mathcal{S}$. One now proceeds with an induction argument as in the proof of Theorem 3.6. At every iteration level, $u_n \in U(u_\gamma)$ and $A_n \in L(\Omega, \vartheta, h, r)$ by assumption, so that (3.22) is applicable. \square

For numerical implementation, we note that the Newton step (3.18) is equivalent to solving the convection-diffusion equation

$$\begin{aligned} -\Delta u + 2\gamma \chi_{A_n} \nabla u_n \cdot \nabla u &= f + \gamma \chi_{A_n} (|\nabla u_n|^2 + g^2) && \text{in } \Omega \\ u &= 0 && \text{on } \Gamma \end{aligned} \tag{3.23}$$

for the next iterate u_{n+1} .

Remark 3.11. In Algorithm 3.1, both u and the active set are updated simultaneously in the inner loop. In the remainder of this section, we briefly comment on an alternative approach (Algorithm 3.2), where u is updated repeatedly until convergence, before an update of the active set is performed.

Algorithm 3.2 Partial semi-smooth Newton method in the multi-dimensional case.

- 1: Choose initial u and $\gamma \geq 0$ and set $n = 0$
- 2: **while** not converged **do**
- 3: **while** not converged **do**
- 4: Set
- $A_n = \{x \in \Omega : |\nabla u_n| > g\}$
- 5: Solve by Newton’s method for u_{n+1} (initial iterate u_n)

$$-\Delta u + \gamma \chi_{A_n} (|\nabla u|^2 - g^2) = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \Gamma$$

- 6: Increase n
 - 7: **end while**
 - 8: Increase γ
 - 9: **end while**
-

Note that the Newton iteration in Step 5 is given by an iteration over (3.23), but with fixed active set:

$$\begin{aligned} -\Delta u^{k+1} + 2\gamma \chi_{A_n} \nabla u^k \cdot \nabla u^{k+1} &= f + \gamma \chi_{A_n} (|\nabla u^k|^2 + g^2) && \text{in } \Omega \\ u^{k+1} &= 0 && \text{on } \Gamma \end{aligned}$$

and initial iterate $u^0 = u_n$. We recover Algorithm 3.1 when only one Newton step is carried out.

We recall from Proposition 3.2 that the problem in Step 5 has a unique solution $u_{n+1} \in W^{3,q}(\Omega) \cap H_0^1(\Omega)$ for all $1 \leq q < \infty$. Concerning the Newton method, the Jacobian $J(u) : H^2(\Omega) \cap H_0^1(\Omega) \rightarrow L^2(\Omega)$ is given by

$$J(u)v = -\Delta v + 2\gamma \chi_{A_n} \nabla u_\gamma \cdot \nabla v.$$

Evaluating J at u_{n+1} , we find that $J(u_{n+1})$ is a continuously invertible operator from $H^2(\Omega) \cap H_0^1(\Omega)$ to $L^2(\Omega)$. Moreover, $u \mapsto J(u)$ is continuous from $H^2(\Omega) \cap H_0^1(\Omega)$ to $L(H^2(\Omega) \cap H_0^1(\Omega), L^2(\Omega))$, and hence $J(u)^{-1}$ is uniformly bounded in a neighborhood of u_{n+1} . Consequently, the Newton loop in Step 5 converges locally quadratically.

We turn to a short discussion of the convergence of the partial semi-smooth Newton loop (Steps 3–7 in Algorithm 3.2). In this case the error equation is given by

$$-\Delta(u_{n+1} - u_\gamma) + \gamma \chi_{A_n} \nabla(u_{n+1} + u_\gamma) \cdot \nabla(u_{n+1} - u_\gamma) = \gamma \max\{0, |\nabla u_\gamma|^2 - g^2\} - \gamma \max\{0, |\nabla u_n|^2 - g^2\} - \gamma \chi_{A_n} (|\nabla u_\gamma|^2 - |\nabla u_n|^2).$$

One can proceed as in the proof of Theorem 3.6 to argue local superlinear convergence of $u_n \rightarrow u_\gamma$ in $H^2(\Omega) \cap H_0^1(\Omega)$, provided that $\chi_{A_n} \in N(\chi_{A_\gamma})$ and that $u_{n+1} \in N(u_\gamma)$ hold for all n . The assumption that $u_{n+1} \in N(u_\gamma)$ involving the new iterate of the state variable is an additional requirement compared to the assumptions of Theorem 3.6. Again the condition $\chi_{A_n} \in N(\chi_{A_\gamma})$ for all n can be replaced by the strict complementarity type assumption (Assumption 3.7). We may conclude that if Assumption 3.7 holds, and if $u_n \rightarrow u_\gamma$, then it converges locally superlinearly.

4. NUMERICAL RESULTS

In this section we present numerical results for the proposed methods. We distinguish between the one- and multi-dimensional settings.

4.1. Semi-smooth Newton method in 1D

We applied Algorithm 2.1 to problem (1.1) and provide two examples. We discretized each example using the finite difference method on a uniform grid with mesh size $h = 1/2000$. The Laplacian was discretized by the standard stencil $[1 \ -2 \ 1]/h^2$. The convective terms in Step 5 of Algorithm 2.1 need to be stabilized, and we used upwind differences for this purpose. That is, u' on the set A^+ was discretized using backward differences, while u' on the set A^- was approximated by forward differences. The same upwind differences were used to discretize u' in Step 4, where the active indices are determined.

Example 4.1. We set the problem data to

$$f(x) = 1 + 15x, \quad g \equiv 2$$

on the domain $\Omega = (0, 1)$.

The unconstrained solution of (1.1), corresponding to $g = \infty$, is given by $u_{\text{unc}}(x) = -\frac{5}{2}x^3 - \frac{1}{2}x^2 + 3x$. Since $u'_{\text{unc}}(0) = 3$ and $u'_{\text{unc}}(1) = -\frac{11}{2}$, we expect active sets at both ends of the interval.

We started Algorithm 2.1 with $\gamma = 1$ and initial guess $u = 0$. The inner *while* loop (active set loop) was terminated upon coincidence of the active sets in two consecutive iterations. Then the regularization parameter γ was increased by a factor of 10. The final iteration was carried out for $\gamma = 10^6$. Table 1 and Figure 2 summarize the convergence behavior of Algorithm 2.1 for this example. It should be noted that the method also converges when the regularization parameter is fixed at $\gamma = 10^6$. However, it then takes 200 iterations to arrive at the same solution and hence we conclude that homotopy significantly improves the convergence.

The final iterate is shown in Figure 3. It has a residual of

$$\|\max\{-\Delta u - f, |\nabla u| - g\}\|_{L^\infty(\Omega)} = 1.60 \times 10^{-5}, \quad (4.1)$$

TABLE 1. Convergence history of Algorithm 2.1 for Example 4.1. $|A^\pm|$ and $|I|$ denote the cardinalities of the active and inactive grid points, and *residual* is the residual norm in (4.1).

Iter	γ	$ A^+ $	$ A^- $	$ I $	Residual
1	1.00 E+00	0	0	2001	3.50 E+00
2		609	495	897	3.12 E+00
3		604	497	900	3.12 E+00
4	1.00 E+01	604	497	900	1.37 E+00
5		561	518	922	1.37 E+00
6		556	518	927	1.37 E+00
7	1.00 E+02	556	518	927	2.66 E+01
8		518	565	918	1.73 E+01
9		486	565	950	9.56 E+00
10		462	564	975	4.19 E+00
11		448	564	989	1.22 E+00
12		442	564	995	1.58 E-01
13	1.00 E+03	442	564	995	3.78 E+01
14		436	573	992	2.92 E+01
15		430	573	998	2.06 E+01
16		425	573	1003	1.35 E+01
17		421	573	1007	7.89 E+00
18		418	573	1010	3.73 E+00
19		416	573	1012	9.75 E-01
20		415	573	1013	1.60 E-02
21	1.00 E+04	415	573	1013	9.14 E+00
22		414	574	1013	5.72 E+00
23		413	574	1014	2.30 E+00
24		412	574	1015	1.60 E-03
25	1.00 E+05	412	574	1015	1.60 E-04
26	1.00 E+06	412	574	1015	1.44 E-02
27		411	574	1016	1.60 E-05

where the maximum is evaluated at all grid points. The active sets found are

$$A^+ = [0, 0.2055], \quad A^- = [0.7130, 1],$$

see Figure 3.

Example 4.2. In our second example we set

$$f(x) = 1 + 15x, \quad g(x) = \begin{cases} 1 & \text{on } [0, 1/8] \\ 0.1 & \text{on } [3/8, 5/8] \\ 2 & \text{elsewhere} \end{cases}$$

and again $\Omega = (0, 1)$.

This example differs from the first one in that the active set has a component which is strictly contained in the interior of Ω . We ran the algorithm up to and including $\gamma = 10^8$. Table 2 and Figure 4 show the convergence history, and the final solution is displayed in Figure 5. Superlinear convergence can be observed

for $\gamma = 10^2, \dots, 10^4$. The final residual is

$$\max\{-\Delta u - f, |\nabla u| - g\} = 1.97 \times 10^{-5},$$

and the active sets are

$$A_1^+ = [0, 0.1250], \quad A_2^+ = [0.3750, 0.6250], \quad A^- = [0.8765, 1].$$

4.2. Active set approach without regularization in 1D

For comparison, we also applied Algorithm 4.1, which is based on the method outlined in Section 2.3, to Example 4.1. Step 4 of this algorithm takes care that the assumed structure of the active sets at the solution (Fig. 1) is preserved throughout the iteration. Differently from the regularized semi-smooth Newton Algorithm 2.1, in this algorithm typically only very few indices migrate between the active and inactive sets from one iteration to the next. This was observed for numerous parameter pairs $(\bar{\gamma}, \kappa) \in [10^{-2}, 10^2] \times [10^{-6}, 10^2]$. After more than 400 iterations, the same solution as before was obtained, see Figure 3. The choice $\kappa = 0$ leads to Dirichlet conditions, which causes Algorithm 4.1 to produce a (false) solution with kinks.

Algorithm 4.1 Active set method without regularization in 1D.

- 1: Choose initial $u_0, \lambda_0^+, \lambda_0^-$ and parameters $\bar{\gamma} > 0$ and $\kappa > 0$ and set $n = 0$
- 2: **while** not converged **do**
- 3: Set the preliminary active and inactive sets

$$\begin{aligned} \tilde{A}_n^+ &= \{x \in \Omega : \lambda_n^+ + \bar{\gamma}(u_n' - g) > 0\}, & \tilde{I} &= \Omega \setminus (\tilde{A}_n^+ \cup \tilde{A}_n^-), \\ \tilde{A}_n^- &= \{x \in \Omega : \lambda_n^- + \bar{\gamma}(-u_n' - g) > 0\} \end{aligned}$$

- 4: Set $l = \inf \tilde{I}_n, r = \sup \tilde{I}_n$ and set the final active and inactive sets

$$I = (l, r), \quad A_n^+ = (0, l], \quad A_n^- = [r, 1)$$

- 5: Solve for u_{n+1}

$$\begin{aligned} -u'' &= f \quad \text{in } I_n, & -\kappa u'(l) + u(l) &= -\kappa g(l) + G(l) - G(0), \\ & & \kappa u'(r) + u(r) &= -\kappa g(r) - G(r) + G(1) \end{aligned}$$

- 6: Set

$$u_{n+1}(x) = G(x) - G(0) \quad \text{in } A_n^+, \quad u_{n+1}(x) = -G(x) + G(1) \quad \text{in } A_n^-$$

- 7: Set

$$\begin{aligned} \lambda_{n+1}^+ &= f + u_{n+1}'' \quad \text{in } A_n^+, & \lambda_{n+1}^+ &= 0 \quad \text{elsewhere} \\ \lambda_{n+1}^- &= f + u_{n+1}'' \quad \text{in } A_n^-, & \lambda_{n+1}^- &= 0 \quad \text{elsewhere} \end{aligned}$$

- 8: Increase n
 - 9: **end while**
-

4.3. Regularized active set method in 2D

In this section, we provide results for Algorithm 3.1, applied to two-dimensional problems. The discretization was carried out by piecewise linear finite elements on triangular meshes.

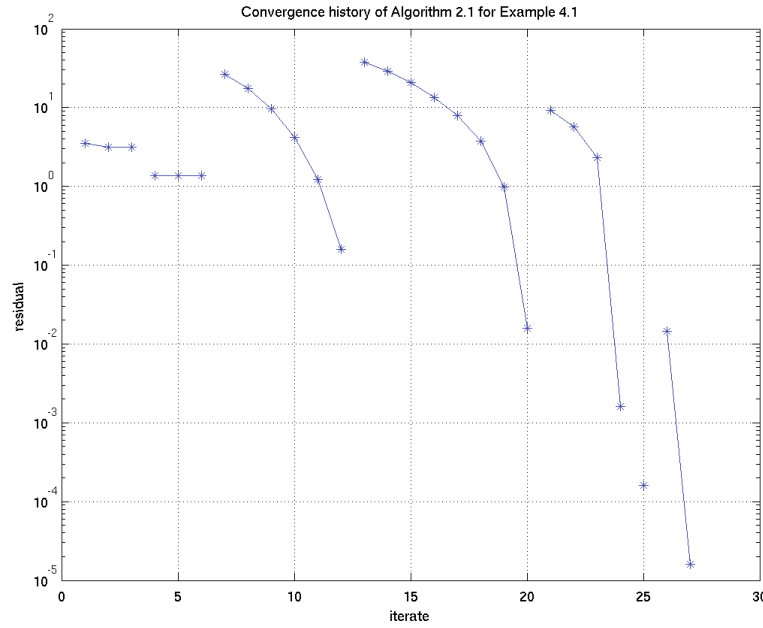


FIGURE 2. Convergence history (residual *vs.* iterate) of Algorithm 2.1 for Example 4.1. Values belonging to one and the same parameter γ are shown connected.

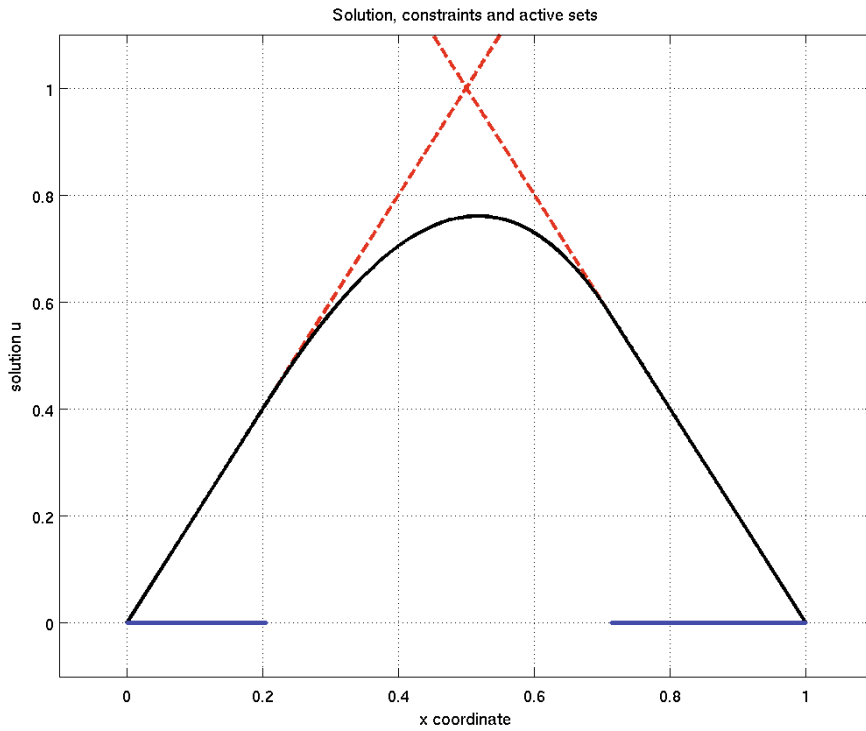


FIGURE 3. The figure shows the final iterate u_γ for Example 4.1. The active sets A^+ (left) and A^- (right) are depicted in blue on the x -axis. The dashed red lines represent the constraint envelope.

TABLE 2. Convergence history of Algorithm 2.1 for Example 4.2. See Table 1 for a legend.

Iter	γ	$ A^+ $	$ A^- $	$ I $	Residual
1	1.00 E+00	0	0	2001	3.50 E+00
2		977	587	437	3.07 E+00
3		956	566	479	3.07 E+00
4	1.00 E+01	956	566	479	6.66 E+00
5		850	446	705	2.51 E+00
6		766	410	825	1.30 E+00
7		733	406	862	1.30 E+00
8		733	405	863	1.30 E+00
9	1.00 E+02	733	405	863	1.00 E+02
10		751	360	890	6.72 E+01
11		751	323	927	3.91 E+01
12		751	295	955	1.91 E+01
13		751	277	973	6.79 E+00
14		751	268	982	8.09 E-01
15		751	266	984	8.09 E-01
16	1.00 E+03	751	266	984	7.85 E+01
17		751	261	989	5.45 E+01
18		751	257	993	3.53 E+01
19		751	253	997	1.63 E+01
20		751	251	999	6.84 E+00
21		751	250	1000	2.12 E+00
22		751	249	1001	6.48 E-01
23	1.00 E+04	751	249	1001	1.40 E+01
24		751	248	1002	2.20 E+00
25		751	247	1003	1.64 E-01
26	1.00 E+05	751	247	1003	1.93 E-02
27	1.00 E+06	751	247	1003	1.97 E-03
28	1.00 E+07	751	247	1003	1.97 E-04
29	1.00 E+08	751	247	1003	1.97 E-05

Each Newton step (3.23) requires the solution of a convection-diffusion equation which may be convection dominated. Moreover, the convection coefficient is discontinuous, and the jump height increases with γ . Naturally, we found the stabilization of the Newton step to be necessary for most examples. In our context, artificial diffusion stabilization turned out to be inadequate and led to cyclic behavior of the active sets, even for small γ . We thus employed a streamline upwind Petrov-Galerkin (SUPG) stabilization [2,11] in every Newton step.

For fixed γ , the residual norm of

$$-\Delta u + \gamma \max\{0, |\nabla u|^2 - g^2\} - f = 0 \quad (4.2)$$

serves as a stopping criterion for the inner *while* loop of Algorithm 3.1. In our examples, we used a tolerance of 10^{-6} for the $L^2(\Omega)$ norm of the weak form of the residual.

We recall that our original problem (1.1) is to solve

$$\max\{-\Delta u - f, |\nabla u| - g\} = 0. \quad (4.3)$$

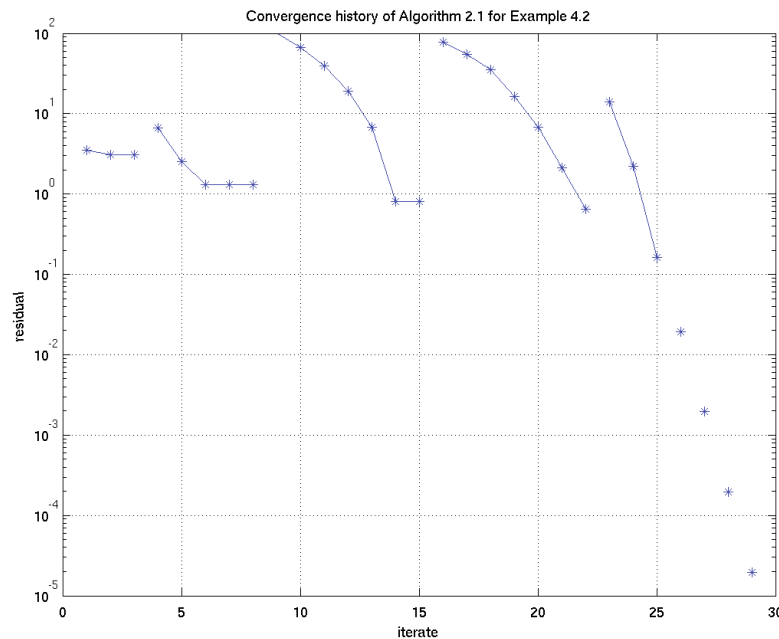


FIGURE 4. Convergence history (residual *vs.* iterate) of Algorithm 2.1 for Example 4.2. Values belonging to one and the same parameter γ are shown connected.

At the end of Step 7 of Algorithm 3.1, the iterates satisfy (4.2) for the current value of γ and we obtain

$$-\Delta u_\gamma - f = -\gamma \max\{0, |\nabla u_\gamma|^2 - g^2\} \leq 0 \quad \text{in } \Omega.$$

On the inactive set $I = \{x \in \Omega : |\nabla u_\gamma| \leq g\}$, this implies $-\Delta u_\gamma - f = 0$ and (4.3) holds. The only violation of (4.3) can occur on $A = \{x \in \Omega : |\nabla u_\gamma| > g\}$. Hence the constraint violation $\|\max\{0, |\nabla u_\gamma| - g\}\|_{L^\infty(\Omega)}$ can serve as a stopping criterion for the outer *while* loop.

Example 4.3. We set Ω to the unit disk in \mathbb{R}^2 and

$$f \equiv 1, \quad g \equiv 0.3.$$

The unconstrained solution is $u_{\text{unc}}(r) = \frac{1}{4}(1 - r^2)$ in terms of polar coordinates. Hence $|\nabla u(r)| = r/2$ holds, and we expect the active set to be an annulus near the boundary of Ω .

The behavior of Algorithm 3.1 for this problem is reported in Figure 7. In Step 8, the regularization parameter γ is increased by a factor of $\sqrt{10}$. Superlinear convergence can be observed for $\gamma > 10^3$ excluding possibly the last iteration for fixed γ . The discretization was carried out on a regular mesh with 817 vertices and 1536 triangles.

Indeed, it is easy to compute the exact solution of (1.1) as

$$u(r) = \begin{cases} \frac{1}{4}(0.84 - r^2), & r \in [0, 0.6] \\ 0.3(1 - r), & r \in [0.6, 1]. \end{cases}$$

The boundary of the active set of the numerical solution numerical solution (Fig. 6) is within sub-grid accuracy to $r = 0.6$ despite the smooth transit of u between the active and inactive sets.

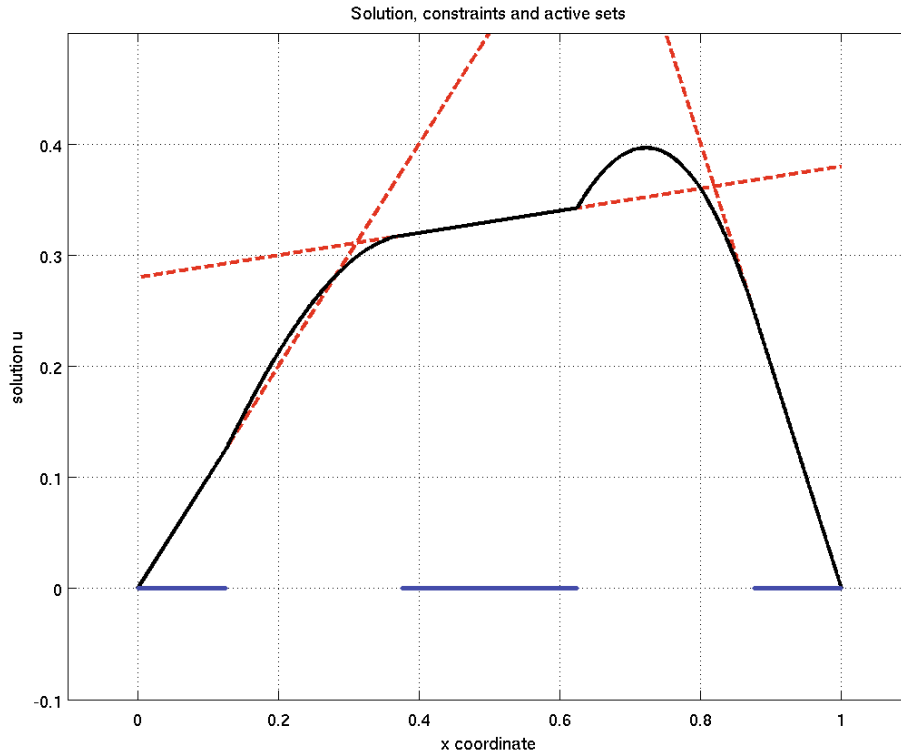


FIGURE 5. The figure shows the final iterate u_γ for Example 4.2. The active sets are again depicted in blue on the x -axis, and A^+ (left and middle) is disconnected. The dashed red lines represent the constraint envelope.

We observed that beyond the value of $\gamma \approx 10^6$, the determination of the active sets became unreliable so that no convergence of the inner *while* loop could be achieved for these values of γ . This is due to numerical error which biases the sign of $|\nabla u| - g$.

We also tested the partial semi-smooth Newton method (Algorithm 3.2) and observed very similar convergence behavior for Examples 4.3 and 4.4.

Example 4.4. The second example features a disconnected active set and a radially unsymmetric configuration, again on the unit disk Ω in \mathbb{R}^2 . We set

$$f(x) = 0.9 + x_1, \quad g(x) = \begin{cases} 0.1 & \text{for } \|x\| \leq 0.3 \\ 0.4 & \text{for } \|x\| > 0.3. \end{cases}$$

We refer to Figure 8 for the solution and to Figure 10 for the convergence behavior of Algorithm 3.1. In this case the iteration was stopped before $\gamma = 10^5$, where again the determination of the active set starts to become unstable.

Finally, we illustrate Assumption 3.7 numerically. For this purpose, we plot the cross section of $|\nabla u_\gamma| - g$ for $\gamma = 10^2$, see Figure 9 (right). This figure was obtained using P_2 finite elements.

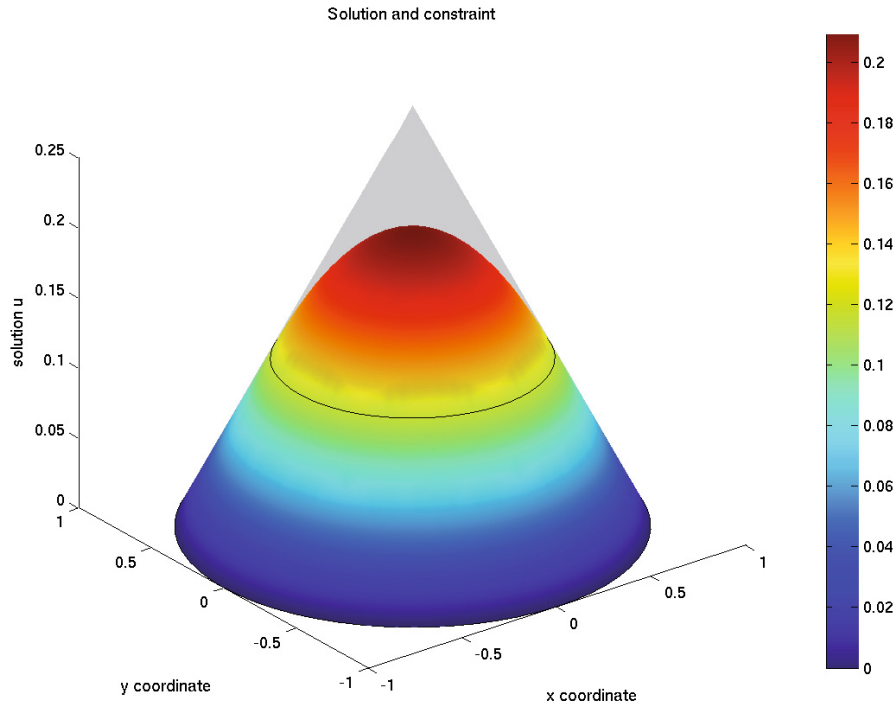


FIGURE 6. The figure shows the final iterate u_γ for Example 4.3. The constraint cone is superimposed on the solution. The contour line marks the boundary of the active set.

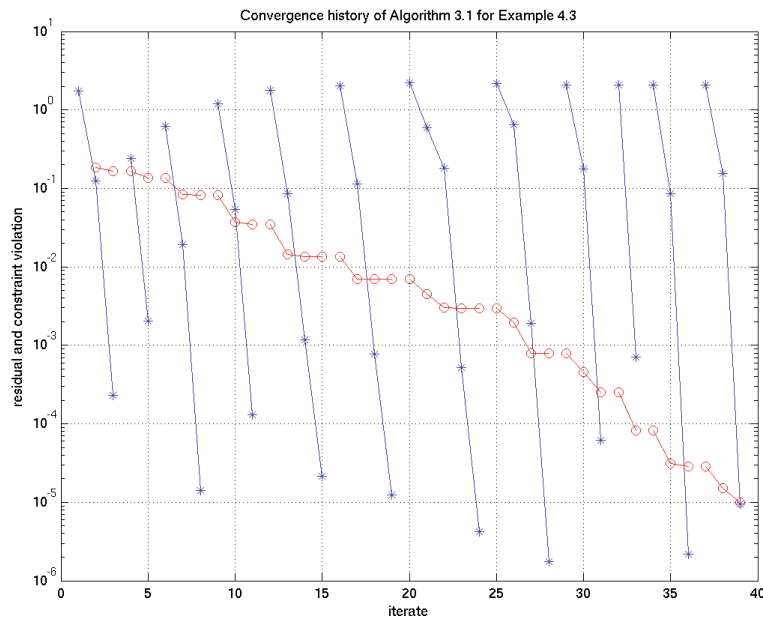


FIGURE 7. Convergence history of Algorithm 3.1 for Example 4.3. The residual (shown with stars) refers to the $L^2(\Omega)$ norm of the residual of equation (4.2). The circles display the the $L^\infty(\Omega)$ norm of the constraint violation $\max\{0, |\nabla u| - g\}$. Values belonging to one and the same parameter γ are shown connected. The values of γ range from 1 to 3.16×10^5 .

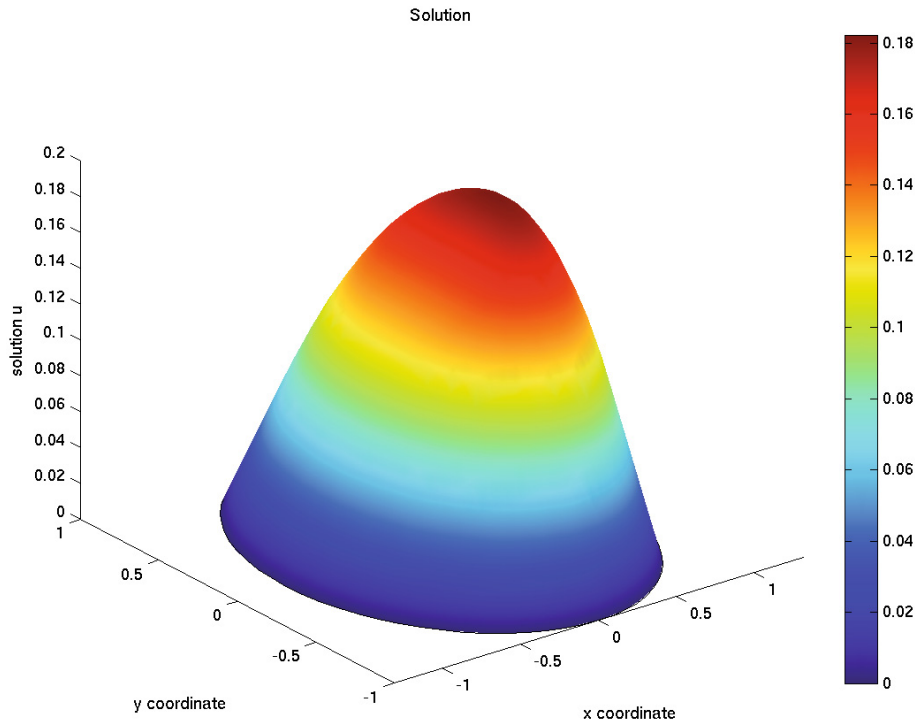


FIGURE 8. The figure shows the final iterate u_γ for Example 4.4.

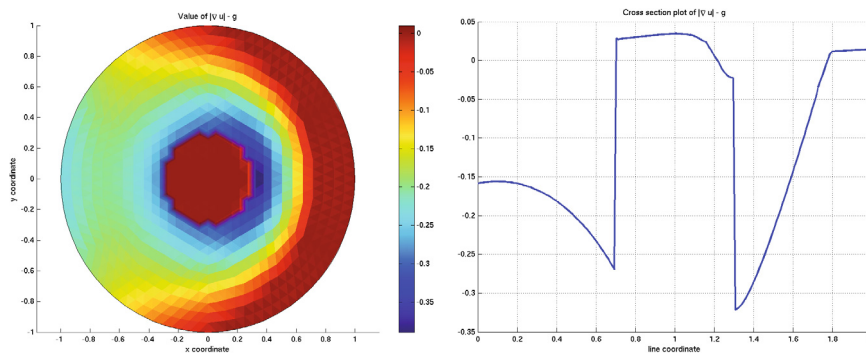


FIGURE 9. The figure shows the value of $|\nabla u_\gamma| - g$ for the final iterate of Example 4.4 on the domain (left), and its cross section along the first median for $\gamma = 10^2$ (right).

5. CONCLUSION AND OUTLOOK

This paper proposes the solution of elliptic equations with gradient constraints by semi-smooth Newton methods. Many extensions and further investigations are possible. These may concern the efficient numerical treatment of the convection dominated problems with discontinuous coefficients which arise in (3.23), or the systematic increase of the parameter γ based, for example, on path following techniques as developed in [7] for a class of variational problems with constraints. Further the extensions of the analysis of this paper to Hamilton Jacobi Bellman equations which arise in portfolio optimization as pointed out in the introduction is of particular interest.

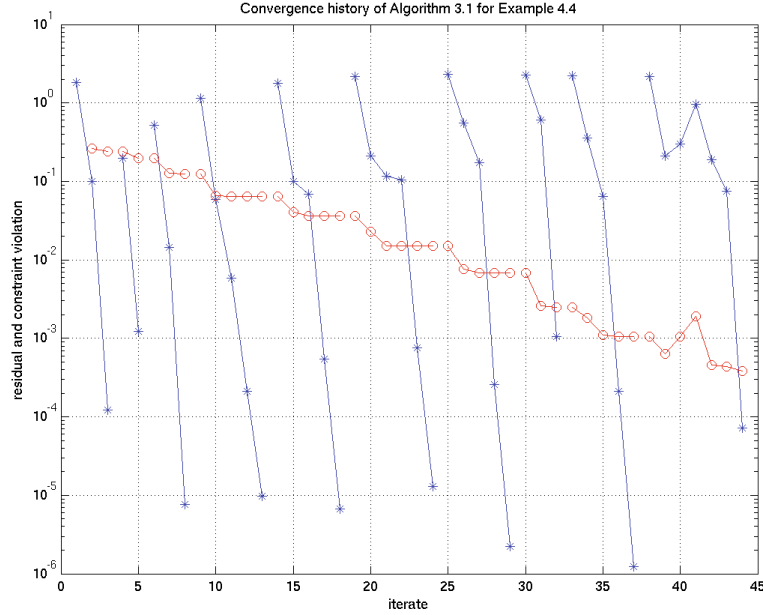


FIGURE 10. Convergence history of Algorithm 3.1 for Example 4.4. See Figure 7 for a legend. The values of γ range from 1 to 3.16×10^4 .

APPENDIX

A – Leray-Schauder fixed point theorem

Theorem 5.1 (Leray-Schauder [6]). *Let T be a compact operator of a Banach space B into itself. Suppose that for all $s \in [0, 1]$, there exists a constant $M > 0$, independent of s , such that $v = sT(v)$ implies that $\|v\|_B \leq M$. Then T has a fixed-point.*

B – Bony maximum principle

Theorem 5.2 (Bony [1], Thm. 1). *Suppose that the differential operator*

$$Lu = \sum_{i,j=1}^d a_{ij}u_{ij} + \sum_{i=1}^d a_i u_i + au$$

has coefficients in $L^\infty(\Omega)$ satisfying $\sum_{i,j} a_{ij}(x) \xi_i \xi_j \geq 0$ on Ω and $a \leq 0$ almost everywhere. If $u \in W^{2,p}(\Omega)$ with $p > d$ attains a positive maximum in $x_0 \in \Omega$, then

$$\lim \text{ess inf}_{x \rightarrow x_0} Lu(x) \leq 0$$

holds.

We apply the maximum principle to $L = -\Delta$ with a change of signs.

REFERENCES

- [1] J.-M. Bony, Principe du maximum dans les espaces de Sobolev. *C. R. Acad. Sci. Paris Sér. A-B* **265** (1967) 333–336.
- [2] A. Brooks and T. Hughes, Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Comput. Methods Appl. Mech. Eng.* **32** (1982) 99–259.
- [3] X. Chen, Superlinear convergence and smoothing quasi-Newton methods for nonsmooth equations. *J. Comput. Appl. Math.* **80** (1997) 105–126.
- [4] M. Delfour and J.-P. Zolésio, *Shapes and Geometries. Analysis, Differential Calculus, and Optimization*. Philadelphia (2001).
- [5] L.C. Evans, A second order elliptic equation with gradient constraint. *Comm. Partial Differ. Equ.* **4** (1979) 555–572.
- [6] D. Gilbarg and N.S. Trudinger, *Elliptic Differential Equations of Second Order*. Springer, New York (1977).
- [7] M. Hintermüller and K. Kunisch, Stationary optimal control problems with pointwise state constraints. *SIAM J. Optim.* (to appear).
- [8] M. Hintermüller, K. Ito and K. Kunisch, The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.* **13** (2002) 865–888.
- [9] H. Ishii and S. Koike, Boundary regularity and uniqueness for an elliptic equation with gradient constraint. *Comm. Partial Differ. Equ.* **8** (1983) 317–346.
- [10] K. Ito and K. Kunisch, The primal-dual active set method for nonlinear optimal control problems with bilateral constraints. *SIAM J. Contr. Opt.* **43** (2004) 357–376.
- [11] C. Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method*. Cambridge University Press, Cambridge (1987).
- [12] K. Kunisch and J. Sass, Trading regions under proportional transaction costs, in *Operations Research Proceedings*, U.M. Stocker and K.-H. Waldmann Eds., Springer, New York (2007) 563–568.
- [13] O.A. Ladyzhenskaya and N.N. Ural'tseva, *Linear and Quasilinear Elliptic Equations*. Academic Press, New York (1968).
- [14] S. Shreve and H.M. Soner, Optimal investment and consumption with transaction costs. *Ann. Appl. Probab.* **4** (1994) 609–692.
- [15] K. Stromberg, *Introduction to Classical Real Analysis*. Wadsworth International, Belmont, California (1981).
- [16] G. Troianiello, *Elliptic Differential Equations and Obstacle Problems*. Plenum Press, New York (1987).
- [17] M. Wiegner, The $C^{1,1}$ -character of solutions of second order elliptic equations with gradient constraint. *Comm. Partial Differ. Equ.* **6** (1981) 361–371.