DFG Deutsche Forschungsgemeinschaft
Priority Programme 1962

A Proximal Gradient Method for Control Problems with Nonsmooth and Nonconvex Control Cost

Carolin Natemeyer, Daniel Wachsmuth

SPP 1962 | Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization

Preprint Number SPP1962-142

received on July 22, 2020

Edited by
SPP1962 at Weierstrass Institute for Applied Analysis and Stochastics (WIAS)
Leibniz Institute in the Forschungsverbund Berlin e.V.
Mohrenstraße 39, 10117 Berlin, Germany
E-Mail: spp1962@wias-berlin.de

World Wide Web: http://spp1962.wias-berlin.de/

A proximal gradient method for control problems with nonsmooth and nonconvex control cost

Carolin Natemeyer, Daniel Wachsmuth *

July 22, 2020

Abstract. We investigate the convergence of an application of a proximal gradient method to control problems with nonsmooth and nonconvex control cost. Here, we focus on control cost functionals that promote sparsity, which includes functionals of L^p-type for $p \in [0, 1)$. We prove stationarity properties of weak limit points of the method. These properties are weaker than those provided by Pontryagin's maximum principle and weaker than L-stationarity.

Keywords. Proximal gradient method, nonsmooth and nonconvex optimization, sparse control problems

1 Introduction

Let $\Omega \subset \mathbb{R}^n$ be Lebesgue measurable with finite measure. We consider a possibly non-smooth optimal control problem of type

$$\min_{u \in L^2(\Omega)} f(u) + \int_\Omega g(u(x))\,\mathrm{d}x. \tag{P}$$

Here, the function $g : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is nonconvex and nonsmooth. Examples include

$$g(u) = |u|^p, \quad p \in (0, 1),$$

and

$$g(u) = |u|_0 := \begin{cases} 1 & \text{if } u \neq 0 \\ 0 & \text{if } u = 0. \end{cases}.$$

The function $f : L^2(\Omega) \to \mathbb{R}$ is assumed to be smooth. Here, we have in mind to choose $f(u) := f(y(u))$ as the smooth part of an optimal control problem incorporating the state equation and possibly smooth cost functional. We will make the assumptions on the ingredients of the control problem precise below in Section 2.

Due to the properties of g, the optimization problem (P) is challenging in several ways. First of all, the resulting integral functional $u \mapsto \int_\Omega g(u(x))\,\mathrm{d}x$ is not weakly lower semicontinuous in $L^2(\Omega)$, so it is impossible to prove existence of solutions of (P) by the direct method. Second, it is challenging to solve numerically, i.e., to compute local minima or stationary points.

*Institut für Mathematik, Universität Würzburg, 97074 Würzburg, Germany, carolin.natemeyer@mathematik.uni-wuerzburg.de, daniel.wachsmuth@mathematik.uni-wuerzburg.de.
This research was partially supported by the German Research Foundation DFG under project grant Wa 3626/3-2.

In this paper, we address this second issue. Here, we propose to use the proximal gradient method (also called forward-backward algorithm [3]). The main idea of this method is as follows: Suppose the objective is to minimize a sum $f + j$ of two functions f and j on the Hilbert space H where f is smooth. Given an iterate u_k, the next iterate u_{k+1} is computed as

$$u_{k+1} = \operatorname*{arg\,min}_{u \in H} \left(f(u_k) + \nabla f(u_k)(u - u_k) + \frac{L}{2}\|u - u_k\|_H^2 + j(u) \right), \tag{1.1}$$

where $L > 0$ is a proximal parameter, and L^{-1} can be interpreted as a step-size. In our setting, the functional to be minimized in each step is an integral function, whose minima can be computed by minimizing the integrand pointwise. Using the so-called prox map, that is defined by

$$\operatorname{prox}_{\gamma j}(z) = \operatorname*{arg\,min}_{x \in H} \left(\frac{1}{2}\|x - z\|_H^2 + \gamma j(x) \right), \tag{1.2}$$

where $\gamma > 0$, the next iterate of the algorithm can be written as

$$u_{k+1} = \operatorname{prox}_{L^{-1} j} \left(u_k - \frac{1}{L} \nabla f(u_k) \right).$$

If $j \equiv 0$, the method reduces to the steepest descent method. If j is the indicator function of a convex set, then the method is a gradient projection method. If f and j are convex, then the convergence properties of the method are well-known: under mild assumptions the iterates (u_k) converge weakly to a global minimum of $f + j$, see, e.g., [3, Corollary 27.9]. If f is non-convex, then weak sequential limit points of (u_k) are stationary, that is, they satisfy $-\nabla f(u^*) \in \partial j(u^*)$. If in addition j is nonconvex, then much less can be proven. In finite-dimensional problems, limit points are fixed points of the iteration, and satisfy the so-called L-stationary type conditions, see [5] and [4, Chapter 10] for optimization problems with l^0-constraints. A feasible point u^* is called L-stationary if

$$u^* = \operatorname{prox}_{L^{-1} j} \left(u^* - \frac{1}{L} \nabla f(u^*) \right).$$

In a recent contribution [16], the method was analyzed when applied to control problems with L^0-control cost. There it was proven that weak sequential limit points of the iterates in $L^2(\Omega)$ satisfy the L-stationary type condition. An essential ingredient of the analysis in [16] was that the functional g is sparsity promoting: solutions of the proximal step are either zero or have a positive distance to zero. We will show how this property can be obtained under weak assumptions on the functional g in (P) near $u = 0$, see Section 3. Still this is not enough to conclude L-stationarity of limit points. We will show that weak limit points satisfy a weaker condition in general, see Theorem 4.18. Under stronger assumptions, L-stationarity can be obtained (Theorems 4.19, 4.20). Let us emphasize that, under weak assumptions, the sequence of iterates (u_k) contains weakly converging subsequences but is not weakly convergent in general. Pointwise a.e. and strong convergence is obtained in Theorem 4.25. We apply these results to $g(u) = |u|^p$, $p \in (0, 1)$ in Section 5.1.

Interestingly, the proximal gradient method sketched above is related to algorithms based on proximal minimization of the Hamiltonian in control problems. These algorithms are motivated by Pontryagin's maximum principle. First results for smooth problems can be found in [15]. There, stationarity of pointwise limits of (u_k) was proven. Under weaker conditions it was proved in [6] that the residual in the optimality conditions tends to zero. These results were transferred to control problems with parabolic partial differential equations in [7].

Notation. We will frequently use $\bar{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$.

2 Preliminary considerations

Throughout the paper, we will use the following assumption on the function f.

Assumption A. The functional $f : L^2(\Omega) \to \mathbb{R}$ is bounded from below and weakly lower semicontinuous. Moreover, f is Fréchet differentiable and $\nabla f : L^2(\Omega) \to L^2(\Omega)$ is Lipschitz continuous with constant L_f, i.e.,

$$\|\nabla f(u_1) - \nabla f(u_2)\|_{L^2(\Omega)} \leq L_f \|u_1 - u_2\|_{L^2(\Omega)}$$

holds for all $u_1, u_2 \in L^2(\Omega)$.

For the moment, let $g : \mathbb{R} \to \bar{\mathbb{R}}$ be lower semicontinuous and bounded from below. In Section 3 below, we will give the precise assumptions on g that allow sparse controls. Let $u \in L^2(\Omega)$ be given. Then $x \mapsto g(u(x))$ is a measurable function, and we define

$$j(u) := \int_\Omega g(u(x)) \, \mathrm{d}x.$$

Then $j : L^2(\Omega) \to \bar{\mathbb{R}}$ is well-defined and lower semicontinuous, but not weakly lower semicontinuous in general. Hence standard existence proofs cannot be applied. For a discussion, we refer to [11, 16]

Remark 2.1. The results are also valid for the general case that g depends on $x \in \Omega$, which results in the integral functional $j(u) = \int_\Omega g(x, u(x)) \, \mathrm{d}x$, provided $g : \Omega \times \mathbb{R} \to \bar{\mathbb{R}}$ is a normal integrand, for the definition we refer to [10, Definition VIII.1.1].

2.1 Necessary optimality conditions

The mapping $u \mapsto \int_\Omega g(u(x)) \, \mathrm{d}x$ is not directionally differentiable in general, and thus there is no first order optimality condition. In the following we are going to derive a necessary optimality condition for (P), known as Pontryagin maximum principle, where no derivatives of the functional are involved. We formulate the Pontryagin maximum principle (PMP) as in [16]. A control $\bar{u} \in L^2(\Omega)$ satisfies (PMP) if and only if for almost all $x \in \Omega$

$$\nabla f(\bar{u})(x)\bar{u}(x) + g(\bar{u}(x)) \leq \nabla f(\bar{u})(x) \cdot v + g(v) \tag{2.1}$$

holds true for all $v \in \mathbb{R}$. The following result is shown in [16, Thm. 2.5] for the special choice $g(u) := |u|_0$.

Theorem 2.2 (Pontryagin maximum principle)**.** *Let $\bar{u} \in L^2(\Omega) \cap L^\infty(\Omega)$ be locally optimal to* (P)*. Furthermore, assume f satisfies*

$$f(u) - f(\bar{u}) = \nabla f(\bar{u}) \cdot (u - \bar{u}) + o(\|u - \bar{u}\|_{L^1(\Omega)})$$

for all $u \in U_{ad}$. Then $\bar{u}$ satisfies the Pontryagin maximum principle (2.1)*.*

Proof. Let $\bar{u}$ be a local solution to (P). We will use needle perturbations of the optimal control. Let $E := \{(v_i, t_i),\ i \in \mathbb{N}\}$ be a countable dense subset of

$$\mathrm{epi}(g) = \{(v, t) \in \mathbb{R} \times \mathbb{R} : g(v) \leq t\}.$$

For arbitrary $x \in \Omega$ define $u_{r,i} \in L^2(\Omega)$ by

$$u_{r,i} := \begin{cases} v_i & \text{on } B_r(x), \\ \bar{u} & \text{otherwise} \end{cases}$$

for some $r > 0$ and $i \in \mathbb{N}$. Let $\chi_r := \chi_{B_r(x)}$, then we have $u_{r,i} = (1 - \chi_r)\bar{u} + \chi_r v_i$ and

$$\begin{aligned}\|u_{r,i} - \bar{u}\|_{L^1(\Omega)} &= \|\chi_r(v_i - \bar{u})\|_{L^1(\Omega)} \leq (|v_i| + \|\bar{u}\|_{L^\infty(\Omega)})\|\chi_r\|_{L^1(\Omega)} \\ &= (|v_i| + \|\bar{u}\|_{L^\infty(\Omega)})|B_r(x)|.\end{aligned}$$

With $j(u) := \int_\Omega g(u(x))\,\mathrm{d}x$ we get

$$\begin{aligned}0 &\leq f(u_{r,i}) + j(u_{r,i}) - f(\bar{u}) - j(\bar{u}) \\ &= \int_\Omega \nabla f(\bar{u})(u_{r,i} - \bar{u})\,\mathrm{d}t + o(\|u_{r,i} - \bar{u}\|_{L^1(\Omega)}) + \int_\Omega (g(u_{r,i}) - g(\bar{u}))\,\mathrm{d}t \\ &\leq \int_{B_r(x)} \nabla f(\bar{u})(v_i - \bar{u}) + (t_i - g(\bar{u}))\,\mathrm{d}t + o(\|u_{r,i} - \bar{u}\|_{L^1(\Omega)})\end{aligned}$$

After dividing above inequality by $|B_r(x)|$ and passing $r \searrow 0$, we obtain by Lebesgue's differentiation theorem

$$0 \leq \nabla f(\bar{u})(x) \cdot (v_i - \bar{u}(x)) + (t_i - g(\bar{u}(x))). \tag{2.2}$$

This holds for every Lebesgue point $x \in \Omega$ of the integrands, i.e., for all $x \in \Omega \setminus N_i$, where N_i is a set of zero Lebesgue measure, on which the above inequality is not satisfied. Since the union $\bigcup_i N_i$ is also of measure zero, (2.2) holds true for all $x \in \Omega \setminus \bigcup_i N_i$ for all i. Due to the density assumption, for $(v, t) \in \mathrm{epi}(g + I_{U_{ad}})$ we find a sequence $(v_k, t_k) \to (v, t)$ with $(v_k, t_k) \in E$, and hence for almost all $x \in \Omega$ it holds

$$0 \leq \nabla f(\bar{u})(x) \cdot (v - \bar{u}(x)) + (t - g(\bar{u}(x))).$$

for all $(v, t) \in \mathrm{epi}(g)$. Choosing $t = g(v)$ yields the claim. □

3 Sparsity promoting proximal operators

In this section, we will investigate the minimization problems that have to be solved in order to compute the proximal gradient step in (1.1). Let $g : \mathbb{R} \to \bar{\mathbb{R}}$ be proper and lower-semicontinuous. For $s > 0$ and $q \in \mathbb{R}$, we define the function

$$h_{q,s}(u) := -qu + \frac{1}{2}u^2 + sg(u).$$

Here, we have in mind to set $q := -\nabla f(u_k)(x)$. Let us investigate scalar-valued optimization problems of form

$$\min_{u \in \mathbb{R}} h_{q,s}(u). \tag{3.1}$$

The solution set is given by the proximal map $\mathrm{prox}_{sg} : \mathbb{R} \rightrightarrows \mathbb{R}$ of g,

$$\mathrm{prox}_{sg}(q) := \operatorname*{arg\,min}_{u \in \mathbb{R}} \left(\frac{1}{2}|u - q|^2 + sg(u)\right).$$

If g is convex then (3.1) is a convex problem, and the proximal map is single-valued. If g is bounded from below and lower semicontinuous, $\mathrm{prox}_{sg}(q)$ is nonempty for all q but may be multi-valued for some q.

The focus of this section is to investigate under which assumptions prox_{sg} is sparsity promoting: Here, we want to prove that there is $\sigma > 0$ such that

$$u \in \mathrm{prox}_{sg} \;\Rightarrow\; u = 0 \text{ or } |u| \geq \sigma.$$

In [13], this was also investigated for some special cases of non-convex functions. We will show that the following assumption is enough to guarantee the sparsity promoting property, it contains the result from [13] as a special case.

Assumption B.

(B1) $g : \mathbb{R} \to \bar{\mathbb{R}}$ is lower semicontinuous, symmetric with $g(0) = 0$.

(B2) There is $u \neq 0$ such that $g(u) \in \mathbb{R}$.

(B3) g satisfies one of the following properties:

(B3.a) g is twice differentiable on an interval $(0, \epsilon)$ for some $\epsilon > 0$ and $\limsup_{u \searrow 0} g''(u) \in (-\infty, 0)$,

(B3.b) g is twice differentiable on an interval $(0, \epsilon)$ for some $\epsilon > 0$ and $\lim_{u \searrow 0} g''(u) = -\infty$,

(B3.c) $0 < \liminf_{u \searrow 0} g(u)$.

(B4) $g(u) \geq 0$ for all $u \in \mathbb{R}$.

By assumption B, the function g is non-convex in a neighborhood of 0 and nonsmooth at 0. Some examples are given below.

Example 3.1. *Functions satisfying assumption B.*

(i) $g(u) := |u|_0 := \begin{cases} 1 & u \neq 0, \\ 0 & \text{else}, \end{cases}$

(ii) $g(u) := |u|^p, \quad p \in (0, 1)$,

(iii) $g(u) := \ln(1 + \alpha|u|)$, *with a given positive constant* α.

(iv) The indicator function of the integers $g(u) := \delta_{\mathbb{Z}}(u) = \begin{cases} 0 & \text{if } u \in \mathbb{Z}, \\ \infty & \text{otherwise.} \end{cases}$

We are interested in the characterization of global solutions to (3.1) in terms of q. It is well-known that for given $s > 0$ the proximal map $q \rightrightarrows \operatorname{prox}_{sg}(q)$ is monotone, i.e., the inequality

$$0 \leq (q_1 - q_2) \left(\operatorname{prox}_{sg}(q_1) - \operatorname{prox}_{sg}(q_2) \right)$$

is satisfied for all $q_1, q_2 \in \mathbb{R}$. In addition, the graph of prox_{sg} is a closed set. Moreover, the following results hold true.

Lemma 3.2. *Let $g : \mathbb{R} \to \bar{\mathbb{R}}$ satisfy Assumption (B1). Let $u \in \operatorname{prox}_{sg}(q)$. Then $u \geq 0$ if and only if $q \geq 0$.*

Proof. Due to (B1), we have $u \in \operatorname{prox}_{sg}(q)$ if and only if $-u \in \operatorname{prox}_{sg}(-q)$. The claim now follows from the monotonicity of the prox-mapping. □

Lemma 3.3. *Let $g : \mathbb{R} \to \bar{\mathbb{R}}$ satisfy Assumptions (B1), (B4). Then the growth condition*

$$|u| \leq 2|q| \quad \forall u \in \operatorname{prox}_{sg}(q)$$

is satisfied.

Proof. Let $u \in \operatorname{prox}_{sg}(q)$. By optimality, the following inequality

$$\frac{1}{2}u^2 - qu + g(u) \leq g(0) = 0.$$

is true. Since $g(u) \geq 0$, the claim follows. □

Lemma 3.4. *Let H be a Hilbert space. Let $f : H \to \bar{\mathbb{R}}$ be a function with $f(0) \in \mathbb{R}$. Then $0 \in \operatorname{prox}_f(q)$ for all $q \in H$ if and only if f is of the form $f(x) = f(0) + \delta_{\{0\}}(x)$. Here, $\delta_{\{0\}}$ is the indicator function of $\{0\}$ defined by $\delta_{\{0\}}(0) = 0$ and $\delta_{\{0\}}(x) = +\infty$ for all $x \neq 0$.*

Proof. If f is of the claimed form, then clearly $\operatorname{prox}_f(q) = \{0\}$ for all q. Now, let $0 \in \operatorname{prox}_f(q)$ for all $q \in H$. Then it holds

$$\frac{1}{2}\|u - q\|_H^2 + f(u) \geq \frac{1}{2}\|q\|_H^2 + f(0) \quad \forall u, q \in H.$$

This is equivalent to

$$f(u) + \frac{1}{2}\|u\|_H^2 \geq f(0) + (u, q)_H \quad \forall u, q \in H.$$

Setting $q := tu$ and letting $t \to +\infty$ shows $f(u) = +\infty$ for all $u \neq 0$. □

Lemma 3.5. *Let $g : \mathbb{R} \to \bar{\mathbb{R}}$ satisfy Assumption (B1). Let $s > 0$. Assume there is $q_0 \geq 0$ such that*

$$q_0|u| \leq \frac{1}{2}u^2 + sg(u) \quad \forall u \in \mathbb{R}. \tag{3.2}$$

Then $u = 0$ is a global solution to (3.1) if $|q| \leq q_0$. If $|q| < q_0$ then $u = 0$ is the unique global solution to (3.1). Moreover, if

$$q_0 := \sup\{q \geq 0 : q|u| \leq \frac{1}{2}u^2 + sg(u) \quad \forall u \in \mathbb{R}\},$$

then $|q| \leq q_0$ is also necessary for $u = 0$ being a global solution to (3.1).

Proof. Let $|q| \leq q_0$. Take $u \neq 0$, then we have

$$h_{q,s}(u) = \frac{1}{2}u^2 + sg(u) - uq \geq \frac{1}{2}u^2 + sg(u) - |u| \cdot |q| \geq \frac{1}{2}u^2 + sg(u) - q_0|u| \geq 0 = h_{q,s}(0).$$

Note that the second inequality is strict if $|q| < q_0$. For the second claim assume $u = 0$ is a global solution to (3.1). Assume $q > 0$. Then it holds

$$qu \leq \frac{1}{2}u^2 + sg(u) \quad \forall u \geq 0.$$

Since $g(u) = g(-u)$, this implies

$$q|u| \leq \frac{1}{2}u^2 + sg(u) \quad \forall u \in \mathbb{R}.$$

By the definition of q_0, the inequality $q \leq q_0$ follows. Similarly, one can prove $|q| \leq q_0$ for negative q. □

Together with Assumption B, these results allows us to show the following key observation concerning the characterization of solutions to (3.1). A similar statement to the following can be found in [13, Theorem 1.1].

Theorem 3.6. *Let $g : \mathbb{R} \to \bar{\mathbb{R}}$ satisfy Assumption B. Then there exists $s_0 \geq 0$ such that for every $s > s_0$ there is $u_0(s) > 0$ such that for all $q \in \mathbb{R}$ a global minimizer u of* (3.1) *satisfies*

$$u = 0 \text{ or } |u| \geq u_0(s).$$

In case g satisfies (B3.b) or (B3.c), s_0 can be chosen to be zero. Moreover, for all $s > 0$ there is $q_0 := q_0(s) > 0$ such that $u = 0$ is a global solution to (3.1) *if and only if $|q| \leq q_0$. If $|q| < q_0$ then $u = 0$ is the unique global solution to* (3.1).

Proof. Assume that the first claim does not hold. Then there are sequences (u_n) and (q_n) and $s > 0$ with $u_n \in \operatorname{prox}_{sg}(q_n)$ and $u_n \to 0$. W.l.o.g., (u_n) is a monotonically decreasing sequence of positive numbers, and hence (q_n) is monotonically decreasing and non-negative by Lemma 3.2. Let u and q denote the limits of both sequences. Since $u_n \neq 0$ is a global minimum of $h_{q_n,s}$, it follows $h_{q_n,s}(u_n) \leq h_{q_n,s}(0) = 0$. Passing to the limit in this inequality, we obtain $\liminf_{n\to\infty} h_{q_n,s}(u_n) \leq 0$, which implies

$$\liminf_{n\to\infty} g(u_n) \leq 0.$$

With $g(0) = 0$ by (B1), this contradicts (B3.c). Let now (B3.a) or (B3.b) be satisfied. Then for n sufficiently large the necessary second-order optimality condition $h''_{q_n,s}(u_n) \geq 0$ holds, and we obtain

$$\limsup_{n\to\infty} h''_{q_n,s}(u_n) \geq 0,$$

which implies

$$1 + s \limsup_{n\to\infty} g''(u_n) \geq 0.$$

This inequality is a contradiction to (B3.a) if $s > -1/\limsup_{u\searrow 0} g''(u) > 0$ and to (B3.b) for all s.

By (B1), it holds $\operatorname{prox}_{sg}(q) \neq \emptyset$ for all q. Due to (B2) and Lemma 3.4, there is $q \geq 0$ such that $0 \notin \operatorname{prox}_{sg}$. The claim concerning q_0 follows from Assumptions (B4), (B3) and Lemma 3.5. First, consider that case (B3.a) or (B3.b) is satisfied, i.e., there is $\epsilon_1 > 0$ such that g is strictly concave on $(0, \epsilon_1]$. By reducing ϵ_1 if necessary, we get $g(\epsilon_1) > 0$. Since $g(u) = 0$, it holds $g(u) \geq \frac{g(\epsilon_1)}{\epsilon_1}|u|$ for all $u \in (0, \epsilon_1)$ by concavity. Due to symmetry, this holds for all u with $|u| \leq \epsilon_1$. Since $g(u) \geq 0$ for all u by (B4), it holds $\frac{1}{2}u^2 + sg(u) \geq \frac{\epsilon_1}{2}|u|$ for all $|u| \geq \epsilon_1$. This proves $\frac{1}{2}u^2 + sg(u) \geq \min(\frac{\epsilon_1}{2}, \frac{sg(\epsilon_1)}{\epsilon_1})|u|$ for all u. Hence, the claim follows with $q_0 := \min(\frac{\epsilon_1}{2}, \frac{sg(\epsilon_1)}{\epsilon_1})$ by Lemma 3.5. Second, if (B3.c) is satisfied, then there are $\epsilon_2, \tau > 0$ such that $g(u) \geq \tau$ for all u with $|u| \in (0, \epsilon_2)$ as g is lower semicontinuous. Therefore, it holds $g(u) \geq \tau \geq \frac{\tau}{\epsilon_2}|u|$ if $|u| \in (0, \epsilon_2)$. The claim follows as above by Lemma 3.5. □

Remark 3.7.

1. In general, the constant u_0 in Theorem 3.6 depends on s and the structure of g.

2. We note the second claim concerning q_0 in Theorem 3.6 holds for all $s > 0$ and does not depend on the first claim due to Assumption (B4). One can replace $g(u) \geq 0$ by the pre-requisite of Lemma 3.5.

3. Assumption B also allows functions of form $g(u) = \tilde{q}(u) + \delta_D(u)$ with some $\tilde{g} : \mathbb{R} \to \bar{\mathbb{R}}$ and the indicator function δ_D of the set $D \subseteq \mathbb{R}$. This means the analysis includes constrained optimization problems, e.g., standard box constraints of form

$$\min_{u\in[a,b]} -qu + \frac{1}{2}u^2 + s\tilde{g}(u),$$

with $a, b \in \mathbb{R}, a < b$.

Example 3.8. *The proximal map of* (3.1) *with* $g(u) = |u|_0$ *is given by the hard-thresholding operator, defined by*

$$\operatorname{prox}_{sg}(q) = \begin{cases} 0 & \text{if } |q| \le \sqrt{2s}, \\ q & \text{else.} \end{cases}$$

With the above considerations in mind, let us discuss the minimization problem

$$\min_{u\in\mathbb{R}} g_k u + \frac{L}{2}(u-u_k)^2 + g(u). \tag{3.3}$$

This minimization corresponds to the pointwise minimization of the integrand in (1.1).

Corollary 3.9. *Let* $g_k, u_k \in \mathbb{R}$, $L > 0$ *be given. Then the number* $u \in \mathbb{R}$ *is a solution to* (3.3) *if and only if*

$$u \in \operatorname{prox}_{L^{-1}g}\left(\frac{Lu_k - g_k}{L}\right).$$

If $\frac{1}{L} > s_0$, *see Theorem 3.6, then all global solutions* u *satisfy*

$$u = 0 \ \text{ or } \ |u| \ge u_0(L^{-1})$$

with some $u_0(L^{-1}) > 0$ *as in Theorem 3.6.*

Proof. Problem (3.3) is equivalent to

$$\min_{u\in\mathbb{R}} \frac{g_k - Lu_k}{L} u + \frac{1}{2}u^2 + \frac{1}{L}g(u)$$

and therefore of form (3.1). The claim follows by definition and from Theorem 3.6. □

4 Analysis of the proximal gradient algorithm

In this section, we will analyze the proximal gradient algorithm.

Algorithm 4.1 (Proximal gradient algorithm)**.** Choose $L > 0$ and $u_0 \in L^2(\Omega)$. Set $k = 0$.

1. Compute u_{k+1} as solution of

$$\min_{u\in L^2(\Omega)} f(u_k) + \nabla f(u_k)(u-u_k) + \frac{L}{2}\|u-u_k\|^2_{L^2(\Omega)} + j(u). \tag{4.1}$$

2. Set $k := k+1$, repeat.

The functional to be minimized in (4.1) can be written as an integral functional. In this representation the minimization can be carried out pointwise by using the previous results. The following statements are generalizations of [16, Lemmas 3.10, 3.11, Theorem 3.12], and the corresponding proofs can be carried over easily.

Lemma 4.2. *Let* $u_k \in U_{ad}$ *be given. Then*

$$\min_{u\in L^2(\Omega)} f(u_k) + \nabla f(u_k)(u-u_k) + \frac{L}{2}\|u-u_k\|^2_{L^2(\Omega)} + \int_\Omega g(u(x))\,\mathrm{d}x \tag{4.2}$$

is solvable, and $u_{k+1} \in L^2(\Omega)$ *is a global solution if and only if*

$$u_{k+1}(x) \in \operatorname{prox}_{L^{-1}g}\left(\frac{1}{L}(Lu_k(x) - \nabla f(u_k)(x))\right) \quad f.a.a.\ x \in \Omega. \tag{4.3}$$

Proof. Let us show, that we can choose a measurable function satisfying the inclusion (4.3). The set-valued mapping $\operatorname{prox}_{L^{-1}g}$ has closed graph and is thus outer semicontinuous. Then by [14, Corollary 14.14], the set-valued mapping $x \mapsto \operatorname{prox}_{L^{-1}g}\left(\frac{1}{L}(Lu_k(x) - \nabla f(u_k)(x))\right)$ is measurable. A well-known result [14, Corollary 14.6] implies the existence of a measurable function u such that $u(x) \in \operatorname{prox}_{L^{-1}g}\left(\frac{1}{L}(Lu_k(x) - \nabla f(u_k)(x))\right)$ for almost all $x \in \Omega$. Due to the growth condition of Lemma 3.3, we have $u \in L^2(\Omega)$, and hence u solves (4.2). If u_{k+1} solves (4.2) then (4.3) follows by a standard argument. □

We introduce the following notation. For a sequence $(u_k) \subset L^2(\Omega)$ define

$$I_k := \{x \in \Omega : \ u_k(x) \neq 0\}, \ \ \chi_k := \chi(u_k) = \chi_{I_k}.$$

Let us now investigate convergence properties of Algorithm 4.1. The following Lemma will be helpful for what follows.

Lemma 4.3. *Assume $\frac{1}{L} > s_0$ with s_0 from Theorem 3.6. Let $u_k, u_{k+1} \in L^2(\Omega)$ be consecutive iterates of Algorithm* (4.1)*. Then*

$$\|u_{k+1} - u_k\|^p_{L^p(\Omega)} \geq u_0^p \|\chi_k - \chi_{k+1}\|_{L^1(\Omega)}$$

holds for $p \in [1, \infty)$, where $u_0 := u_0(L^{-1})$ is as in Theorem 3.6.

Proof. Since $u_k(x) \neq 0$ and $u_{k+1}(x) = 0$ on $I_k \setminus I_{k+1}$, it holds $|u_{k+1}(x) - u_k(x)| \geq u_0$ for all $x \in I_k \setminus I_{k+1}$ by Corollary (3.9). Hence,

$$\begin{aligned}\|u_{k+1} - u_k\|^p_{L^p(\Omega)} &= \int_\Omega |u_{k+1}(x) - u_k(x)|^p \,\mathrm{d}x \\ &\geq \int_{(I_k \setminus I_{k+1}) \cup (I_{k+1} \setminus I_k)} |u_{k+1}(x) - u_k(x)|^p \,\mathrm{d}x \geq u_0^p \|\chi_{k+1} - \chi_k\|_{L^1(\Omega)}.\end{aligned}$$

□

Theorem 4.4. *For $L > L_f$ let (u_k) be a sequence of iterates generated by Algorithm 4.1. Then the following statements hold:*

(i) The sequence $(f(u_k) + j(u_k))$ is monotonically decreasing and converging.

(ii) The sequences (u_k) and $(\nabla f(u_k))$ are bounded in $L^2(\Omega)$ if $f + j$ is weakly coercive on $L^2(\Omega)$, i.e., $f(u) + j(u) \to \infty$ as $\|u_k\|_{L^2(\Omega)} \to \infty$.

(iii) $\|u_{k+1} - u_k\|_{L^2(\Omega)} \to 0$.

(iv) Let s_0 be as in Theorem 3.6. Assume $\frac{1}{L} > s_0$. Then the sequence of characteristic functions (χ_k) is converging in $L^1(\Omega)$ and pointwise a.e. to some characteristic function χ.

Proof. (i) Due to the Lipschitz continuity of ∇f it holds

$$f(u_{k+1}) \leq f(u_k) + \nabla f(u_k)(u_{k+1} - u_k) + \frac{L_f}{2}||u_{k+1} - u_k||^2_{L^2(\Omega)}.$$

Using the optimality of u_{k+1}, we find that the inequality

$$f(u_{k+1}) + j(u_{k+1}) \leq f(u_k) + j(u_k) - \frac{L - L_f}{2}\|u_{k+1} - u_k\|^2_{L^2(\Omega)} \tag{4.4}$$

holds.. Hence, $(f(u_k) + j(u_k))$ is decreasing. Convergence follows because f and j are bounded from below.

(ii) Weak coercivity of the functional implies that (u_k) is bounded. Furthermore, because of

$$\begin{aligned}\|\nabla f(u_k)\|_{L^2(\Omega)} &\leq \|\nabla f(u_k) - \nabla f(0)\|_{L^2(\Omega)} + \|\nabla f(0)\|_{L^2(\Omega)} \\ &\leq L_f \|u_k\|_{L^2(\Omega)} + \|\nabla f(0)\|_{L^2(\Omega)},\end{aligned}$$

boundedness of $(\nabla f(u_k))$ in $L^2(\Omega)$ follows.

(iii) Summation over $k = 1, \dots, n$ in (4.4) yields

$$\sum_{k=1}^{n} (f(u_{k+1}) + j(u_{k+1})) \leq \sum_{k=1}^{n} \left(f(u_k) + j(u_k) - \frac{L - L_f}{2} \|u_{k+1} - u_k\|^2_{L^2(\Omega)} \right)$$

and hence

$$f(u_{n+1}) + j(u_{n+1}) + \sum_{k=1}^{n} \frac{L - L_f}{2} \|u_{k+1} - u_k\|^2_{L^2(\Omega)} \leq f(u_1) + j(u_1) < \infty.$$

Letting $n \to \infty$ implies $\sum\limits_{k=1}^{\infty} \|u_{k+1} - u_k\|^2_{L^2(\Omega)} < \infty$ and therefore $\|u_{k+1} - u_k\|_{L^2(\Omega)} \to 0$.

(iv) By Lemma 4.3, we get

$$\frac{L - L_f}{2} u_0^2 \sum_{k=1}^{\infty} \|\chi_k - \chi_{k+1}\|_{L^1(\Omega)} \leq \frac{L - L_f}{2} \sum_{k=1}^{\infty} \|u_k - u_{k+1}\|_{L^2(\Omega)} < +\infty$$

Hence, (χ_k) is a Cauchy sequence in $L^1(\Omega)$, and therefore also converging in $L^1(\Omega)$, i.e., $\chi_k \to \chi$ for some characteristic function χ. Pointwise a.e. convergence of (χ_k) can be proven by Fatou's Lemma.

□

As a consequence, we get the following result.

Corollary 4.5. *Suppose $\frac{1}{L} > s_0$. Then for any weak sequential limit point $u^* \in L^2(\Omega)$ of iterates (u_k) of Algorithm 4.1 it holds*

$$(1 - \chi)u^* = 0$$

almost everywhere in Ω. Here, χ is as in Theorem 4.4.

Proof. See [16, Thm.3.15]. □

Corollary 4.6. *Let (u_k) be a sequence of iterates generated by Algorithm 4.1. Then $u_{k+1} - u_k \to 0$ pointwise almost everywhere on Ω.*

Proof. By the Lemma of Fatou, we have

$$\int_\Omega \liminf_{n\to\infty} \sum_{k=0}^{n} |u_{k+1}(x) - u_k(x)|^2 \, \mathrm{d}x \leq \liminf_{n\to\infty} \sum_{k=0}^{n} \|u_{k+1}(x) - u_k(x)\|^2_{L^2(\Omega)} < +\infty.$$

This implies $\sum_{k=0}^{n} |u_{k+1}(x) - u_k(x)|^2 < \infty$ for almost all $x \in \Omega$, and the claim follows. □

4.1 Stationarity conditions for weak limit points from inclusions

Under a weak coercivity assumption Theorem 4.4 implies that Algorithm 4.1 generates a sequence (u_k) with weak limit point $u^* \in L^2(\Omega)$. Due to the lack of weak lower semicontinuity in the term $u \mapsto \int_\Omega g(u)\,\mathrm{d}x$, however, we cannot conclude anything about the value of the objective functional in a weak limit point. Unfortunately, we are not able to show

$$f(u^*) + j(u^*) \leq \lim_{k\to\infty} f(u_k) + j(u_k),$$

as it was done in [16, Thm. 3.14] for the special choice $g(u) := |u|_0$. Nevertheless, by using results of set-valued analysis we will show that a weak limit point of a sequence (u_k) of iterates satisfies a certain inclusion in almost every point $x \in \Omega$, which can be interpreted as a pointwise stationary condition for weak limit points.

By definition, the iterates satisfy the inclusion

$$u_{k+1}(x) \in \operatorname{prox}_{L^{-1}g}\left(\frac{1}{L}(Lu_k(x) - \nabla f(u_k)(x))\right)$$

for almost all $x \in \Omega$, see e.g., (4.3). However, this inclusion seems to be useless for a convergence analysis as the function u_{k+1} to the left of the inclusion as well as the arguments $Lu_k - \nabla f(u_k)$ only have weakly converging subsequences at best. The idea is to construct a set-valued mapping $\mathcal{G} : \mathbb{R} \rightrightarrows \mathbb{R}$, such that a solution u_{k+1} of (4.2) satisfies the inclusion

$$u_{k+1}(x) \in \mathcal{G}(z_k(x)) \tag{4.5}$$

in almost every point $x \in \Omega$ for some $z_k \in L^2(\Omega)$, where (z_k) converges strongly or pointwise almost everywhere. Here, we will use

$$z_k := -(\nabla f(u_k) + L(u_{k+1} - u_k)).$$

By Theorem 4.4, we have $u_{k+1} - u_k \to 0$ in $L^2(\Omega)$ and pointwise almost everywhere. With the additional assumption that subsequences of $(\nabla f(u_k))$ are converging pointwise almost everywhere, the argument of the set-valued mapping is converging pointwise almost everywhere. In the context of optimal control problems, such an assumption is not a severe restriction. So there is a chance to pass to the limit in the inclusion (4.5).

Lemma 4.7. *Let u_{k+1} be a solution of (4.2). Then*

$$u_{k+1}(x) \in \mathcal{G}(z_k(x)) \;\; f.a.a.\;\; x \in \Omega,$$

where the set-valued mapping $\mathcal{G} : \mathbb{R} \rightrightarrows \mathbb{R}$ is given by

$$u \in \mathcal{G}(z) := \mathcal{G}_L(z) :\Longleftrightarrow u = \operatorname*{arg\,min}_{v\in\mathbb{R}} -zv + \frac{L}{2}(v-u)^2 + g(v).$$

Unfortunately, the set-valued map $\mathcal{G}$ is not monotone in general. If g would be convex, then the optimality condition of (4.2) is $z_k(x) \in \partial g(u_{k+1}(x))$ for almost all $x \in \Omega$, hence one could choose $\mathcal{G} = \operatorname{gph}(\partial(g^*))$, where g^* denotes the convex conjugate of g.

Remark 4.8. The definition of $\mathcal{G}$ is related to the concept of L-stationary points, introduced in [4, Definition 9.19] for l^0-optimization problems in $\mathbb{R}^n$.

For the rest of this section, we will always suppose that g satisfies Assumption B. As a first direct consequence from the definition of $\mathcal{G}$ we get

Corollary 4.9. *Assume* $\frac{1}{L} > s_0$. *Let* $u, z \in \mathbb{R}$ *such that* $u \in \mathcal{G}(z)$. *Then we have: If* $u > 0$ *then* $u \geq \max\left(u_0, \frac{Lq_0 - z}{L}\right)$, *and if* $u < 0$ *then* $u \leq \min\left(-u_0, -\frac{Lq_0 + z}{L}\right)$. *In case* $u = 0$ *it holds* $|z| \leq Lq_0$. *Here,* $u_0 := u_0(L^{-1})$ *and* $q_0 := q_0(L^{-1})$ *are the positive constants from Theorem 3.6.*

Proof. By construction of $\mathcal{G}$, we have

$$u \in \mathcal{G}(z) \iff u = \operatorname{prox}_{L^{-1}g}\left(\frac{Lu + z}{L}\right).$$

If $u \neq 0$ then by Lemma 3.2 and Theorem 3.6, it follows that $u \geq u_0(L^{-1})$ if and only if $\frac{Lu+z}{L} \geq q_0(L^{-1})$ and likewise $u < -u_0(L^{-1})$ iff $\frac{Lu+z}{L} \leq -q_0(L^{-1})$. The claim follows for $u > 0$ and $u < 0$, respectively. On the other hand $u = 0$ is a solution if and only if $|\frac{z}{L}| \leq q_0$, which implies the claim for $u = 0$. □

4.2 A convergence result for inclusions

Let us recall a few helpful notions and results from set-valued analysis that can be found in the literature, see e.g., [2, 14].

Definition 4.10. For a sequence of sets $A_n \subset \mathbb{R}^n$ we define the *outer limit* by

$$\limsup_{n\to\infty} A_n := \{x : \ \exists (x_{n_k}), x_{n_k} \to x, \ \ x_{n_k} \in A_{n_k}\}.$$

Definition 4.11. Let $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ be a set-valued map.

1. The domain and graph of S are defined by
$$\operatorname{dom} S := \{x : S(x) \neq \emptyset\}, \ \ \operatorname{gph} S := \{(x, y) : y \in S(x)\}.$$
2. S is called *outer semicontinuous* in $\bar{x}$ if
$$\limsup_{x\to\bar{x}} S(x) \subseteq S(\bar{x}).$$
3. S is called *locally bounded* at $x \in \mathbb{R}^m$ if there is a neighborhood U of x such that $S(U)$ is bounded.

A set-valued mapping S is outer semicontinuous if and only if it has a closed graph.

The following convergence analysis relies on [2, Thm. 7.2.1]. We want to extend this result to set-valued maps into $\mathbb{R}^n$ that are not locally bounded. Let us define the following set-valued map that serves as a generalization of $x \to \operatorname{conv}(F(x))$ for the locally unbounded situation.

Definition 4.12. Let $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ be a set-valued map.

Define the set-valued map $\operatorname{conv}^\infty F : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ by

$$(\operatorname{conv}^\infty F)(x) := \limsup_{k\to\infty} \operatorname{conv}\left(F\left(x + B_{1/k}(0)\right)\right).$$

By definition, it holds $\operatorname{gph} F \subset \operatorname{gph} \operatorname{conv}^\infty F$. In addition, we have $\overline{\operatorname{conv}}(F(x)) \subset (\operatorname{conv}^\infty F)(x)$. If F is locally bounded in x, then $(\operatorname{conv}^\infty)F(x) = \overline{\operatorname{conv}}(F(x))$, which can be proven using Carathéodory's theorem. In general, $\operatorname{dom} \operatorname{conv}^\infty F$ is strictly larger than $\operatorname{dom} F$.

Example 4.13. *Define* $F : \mathbb{R} \rightrightarrows \mathbb{R}$ *by*

$$\operatorname{gph} F = \{(x, y) : \ yx = 1\}.$$

Then F *is not locally bounded near* $x = 0$. *Here it holds* $\operatorname{gph}(\operatorname{conv}^\infty F) = \operatorname{gph} F \cup (\{0\} \times \mathbb{R})$, *so that* $\operatorname{dom}(\operatorname{conv}^\infty F) = \mathbb{R} \neq \operatorname{dom} F$.

Theorem 4.14. *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ be a set-valued map. Let sequences of measurable functions $(x_n), (y_n)$ be given such that*

1. *x_n converges almost everywhere to some function $x : \Omega \to \mathbb{R}^m$,*
2. *y_n converges weakly to a function y in $L^1(\mu, \mathbb{R}^n)$,*
3. *$y_n(t) \in F(x_n(t))$ for almost all $t \in \Omega$.*

Then for almost all $t \in \Omega$ it holds:

$$y(t) \in (\operatorname{conv}^\infty F)(x(t)).$$

Proof. Arguing as in the proof of [2, Thm. 7.2.1], we find

$$y(t) \in \bigcap_{k \in \mathbb{N}} \overline{\operatorname{conv}} \left(F(x(t) + B_{1/k}(0)) \right)$$

for almost all $t \in \Omega$. Take $t \in \Omega$ such that the above inclusion is satisfied. Then there is a sequence (u_k) such that $u_k \to y(t)$, $u_k \in \operatorname{conv}(F(x(t) + B_{1/k}(0)))$. This implies $y(t) \in \limsup_{k\to\infty} \operatorname{conv} \left(F(x(t) + B_{1/k}(0)) \right)$, or equivalently $y(t) \in (\operatorname{conv}^\infty F)(x(t))$. □

4.3 Stationarity conditions for weak limit points

Recall, for iterates (u_k) of Algorithm 4.1 and the corresponding sequence z_k we have by construction

$$u_{k+1}(x) \in \mathcal{G}(z_k(x)) \quad \text{f.a.a.} \quad x \in \Omega.$$

Then by Theorem 4.14, we could expect the inclusion $u^*(t) \in (\operatorname{conv}^\infty \mathcal{G})(-\nabla f(u^*)(x))$ pointwise almost everywhere to hold in the subsequential limit. However, the convexification of $\mathcal{G}$ results in a set-valued map that is very large. In order to obtain a smaller inclusion in the limit, we will employ the result of Corollary 4.9: the graph of $\mathcal{G}$ can be split into three clearly separated components. In the sequel, we will show that we can pass to the limit with each component separately, which leads to a smaller set-valued map in the limit. This observation motivates the following splitting of the map $\mathcal{G}$.

Definition 4.15. For $L > 0$ we define the following set-valued mappings.

1. $\mathcal{G}^+ : \mathbb{R} \rightrightarrows \mathbb{R}$ with $u \in \mathcal{G}^+(z) :\Longleftrightarrow u \in \mathcal{G}(z)$ and $u > 0$,
2. $\mathcal{G}^- : \mathbb{R} \rightrightarrows \mathbb{R}$ with $u \in \mathcal{G}^-(z) :\Longleftrightarrow u \in \mathcal{G}(z)$ and $u < 0$,
3. $\mathcal{G}^0 : \mathbb{R} \rightrightarrows \mathbb{R}$ with $u \in \mathcal{G}^0(z) :\Longleftrightarrow u \in \mathcal{G}(z)$ and $u = 0$.

The mappings $\mathcal{G}^+, \mathcal{G}^-$ and $\mathcal{G}^0$ are depicted in Figure 2 for the special choice $g(u) := \frac{\alpha}{2}|u|^2 + |u|^p + \delta_{[-b,b]}(u)$, $p \in (0,1), b \in (0,\infty)$.

Obviously we have by construction

$$u_{k+1}(x) \in (\mathcal{G}^+ \cup \mathcal{G}^- \cup \mathcal{G}^0)(z_k(x)) \quad \text{f.a.a. } x \in \Omega. \tag{4.6}$$

Corollary 4.16. *The mappings $\mathcal{G}, \mathcal{G}^0$ are outer semicontinuous. If $L^{-1} > s_0$ the same holds for $\mathcal{G}^+$ and $\mathcal{G}^-$.*

Proof. $\mathcal{G}$ being outer semicontinuous is equivalent to the closedness of its graph. Let $(u_n), (q_n)$ be sequences such that $u_n \to u, q_n \to q$ and $u_n \in \mathcal{G}(q_n)$. By definition it holds

$$0 \leq -q_n(v - u_n) + (g(v) - g(u_n)) + \frac{L}{2}(v - u_n)^2$$

for all $v \in \mathbb{R}$. Passing to the limit in above inequality yields

$$0 \leq -q(v - u) + (g(v) - g(u)) + \frac{L}{2}(v - u)^2$$

due to the lower semicontinuity of g. Hence,

$$u = \arg\min_{v \in \mathbb{R}} -qv + \frac{L}{2}(v - u)^2 + g(v),$$

i.e., $u \in \mathcal{G}(q)$, which is the claim for $\mathcal{G}$. For $\mathcal{G}^+, \mathcal{G}^-, \mathcal{G}^0$ the claim follows as their graphs are intersections of closed sets with $\operatorname{gph} \mathcal{G}$, which follows from Corollary 4.9 (for suitable chosen L in case of $\mathcal{G}^+, \mathcal{G}^-$). □

In the sequel we want to apply Theorem 4.14 to each of the set-valued maps in (4.6) separately. Let us first show the next helpful result.

Lemma 4.17. *Let (u_k) be a sequence of iterates generated by Algorithm 4.1. Let $b > a$ be given. Define*

$$A_k^+ := \{x \in \Omega : \ u_k(x) \geq b\},$$
$$A_k^- := \{x \in \Omega : \ u_k(x) \leq a\},$$

and $\chi_k^+ := \chi_{A_k^+}$, $\chi_k^- := \chi_{A_k^-}$. Then it holds

$$\sum_{k=1}^{\infty} \|\chi_{k+1}^+ \chi_k^- + \chi_{k+1}^- \chi_k^+\|_{L^1(\Omega)} < +\infty.$$

If $\chi_k^+ + \chi_k^- = 1$ for all k almost everywhere, then there are characteristic functions χ^+, χ^- such that $\chi^+ + \chi^- = 1$ almost everywhere, $\chi_k^+ \to \chi^+$ and $\chi_k^- \to \chi^-$ strongly in $L^1(\Omega)$ and pointwise almost everywhere.

Proof. Let $x \in \Omega$. If $\chi_{k+1}^+(x)\chi_k^-(x) = 1$, then $u_{k+1}(x) - u_k(x) \geq b - a$. This proves $\|\chi_{k+1}^+ \chi_k^-\|_{L^1(\Omega)} \leq (b-a)^{-2}\|u_{k+1} - u_k\|_{L^2(\Omega)}^2$. Similarly, we obtain $\|\chi_{k+1}^- \chi_k^+\|_{L^1(\Omega)} \leq (b-a)^{-2}\|u_{k+1} - u_k\|_{L^2(\Omega)}^2$. Since $\sum_{k=1}^{\infty} \|u_{k+1} - u_k\|_{L^2(\Omega)}^2 < +\infty$, the claim follows. Suppose $\chi_k^+ + \chi_k^- = 1$ almost everywhere. Then we have

$$\chi_{k+1}^+ \chi_k^- + \chi_{k+1}^- \chi_k^+ = \chi_{k+1}^+(1 - \chi_k^+) + (1 - \chi_{k+1}^+)\chi_k^+ = |\chi_{k+1}^+ - \chi_k^+|,$$

which implies the second claim. □

Theorem 4.18. *Let s_0 be as in Theorem 3.6. Assume $\frac{1}{L} > s_0$. Let (u_k) be a sequence of iterates generated by Algorithm 4.1 with weak limit point $u^* \in L^2(\Omega)$, i.e., $u_{k_n} \rightharpoonup u^*$. Assume $\nabla f(u_{k_n})(x) \to \nabla f(u^*)(x)$ for almost every $x \in \Omega$. Let $\mathcal{G}^0, \mathcal{G}^+, \mathcal{G}^- : \mathbb{R} \rightrightarrows \mathbb{R}$ be as in Definition 4.15. Then*

$$u^*(x) \in \left(\mathcal{G}_0 \cup \operatorname{conv}^\infty \mathcal{G}^+ \cup \operatorname{conv}^\infty \mathcal{G}^-\right)(-\nabla f(u^*)(x))$$

holds for almost all $x \in \Omega$.

Proof. By Theorem 4.4 and Corollary 4.6, we have $u_{k_n+1} \rightharpoonup u^*$ in $L^2(\Omega)$ and

$$z_{k_n} := -\left(\nabla f(u_{k_n}) + L(u_{k_n+1} - u_{k_n})\right) \to -\nabla f(u^*) := z$$

pointwise almost everywhere on Ω. Let us define $I_k^+ := \{x \in \Omega : u_k(x) > 0\}$ and $I_k^- := \{x \in \Omega : u_k(x) < 0\}$ with associated characteristic functions χ_k^+, χ_k^-. Then by Lemma 4.17 with $a = 0$ and $b = u_0$ with u_0 from Theorem 3.6, we obtain $\chi_k^+ \to \chi^+$ in $L^1(\Omega)$ and pointwise almost everywhere. Similarly, $\chi_k^- \to \chi^-$ in $L^1(\Omega)$ and pointwise almost everywhere.

Let us fix $(u', q') \in \operatorname{gph} \mathcal{G}^+$. Then the following inclusion

$$\chi_{k+1}^+ u_{k+1} + (1 - \chi_{k+1}^+)u' \in \mathcal{G}^+(\chi_{k+1}^+ z_k + (1 - \chi_{k+1}^+)q')$$

is satisfied almost everywhere on Ω. By Theorem 4.14, we obtain

$$\chi^+ u^* + (1 - \chi^+)u' \in \operatorname{conv}^\infty \mathcal{G}^+(\chi^+ z + (1 - \chi^+)q')$$

almost everywhere on Ω. Similarly, we obtain for $(u'', q'') \in \operatorname{gph} G^-$

$$\chi^- u^* + (1 - \chi^-)u'' \in \operatorname{conv}^\infty \mathcal{G}^-(\chi^- z + (1 - \chi^-)q'')$$

and

$$(1 - \chi)u^* \in \mathcal{G}^0((1 - \chi)z)$$

almost everywhere, where χ_k and χ are as in Theorem 4.4. Note that $\operatorname{conv}^\infty \mathcal{G}^0 = \mathcal{G}^0$. By construction, $\chi_k^+ + \chi_k^- = \chi_k$, which implies $\chi^+ + \chi^- = \chi$. Then we can combine all the inclusions above into one, which is

$$u^+(x) \in \left(\mathcal{G}_0 \cup \operatorname{conv}^\infty \mathcal{G}^+ \cup \operatorname{conv}^\infty \mathcal{G}^-\right)(-\nabla f(u^*)(x))$$

for almost all $x \in \Omega$. □

Let us remark that the assumption of pointwise convergence of $(\nabla f(u_k))$ is not a severe restriction. If $\nabla f : L^2(\Omega) \to L^2(\Omega)$ is completely continuous, then this assumption is fulfilled. For many control problems, this property of ∇f is guaranteed to hold.

Interestingly, we can get rid of the convexification operator $\operatorname{conv}^\infty$ if we assume that the whole sequence $(\nabla f(u_k))$ converges pointwise almost everywhere.

Theorem 4.19. *Let (u_k) be a sequence of iterates generated by Algorithm 4.1 with weak limit point $u^* \in L^2(\Omega)$. Assume $\nabla f(u_k) \to \nabla f(u^*)$ pointwise almost everywhere. Then*

$$u^*(x) \in \mathcal{G}(-\nabla f(u^*)(x))$$

holds for almost all $x \in \Omega$.

Proof. Denote $z(x) := -\nabla f(u^*)(x)$. Then $z_k(x) \to z(x)$ almost everywhere.

Let $(\tilde{z}, \tilde{u}) \notin \operatorname{gph} \mathcal{G}$. Since $\operatorname{gph} \mathcal{G}$ is closed, there is $\epsilon > 0$ such that

$$(B_\epsilon(\tilde{z}) \times B_\epsilon(\tilde{u})) \cap \operatorname{gph} \mathcal{G} = \emptyset.$$

Let $\epsilon' \in (0, \epsilon)$. Set

$$I := \{x : \ |\tilde{z} - z(x)| < \epsilon'\},$$

and

$$I_K := \{x \in I : \ |\tilde{z} - z_k(x)| < \epsilon \quad \forall k > K\}.$$

The sequence (I_K) is monotonically increasing. Since $z_k(x) \to z(x)$ for almost all $x \in \Omega$, we have $\cup_{K\in\mathbb{N}} I_K = I$.

Define

$$A_k^+ := \{x \in \Omega : \ u_k(x) \geq \tilde{u} + \epsilon\},$$
$$A_k^- := \{x \in \Omega : \ u_k(x) \leq \tilde{u} - \epsilon\},$$

and $\chi_k^+ := \chi_{A_k^+}$, $\chi_k^- := \chi_{A_k^-}$. By Lemma 4.17 above, we have $\sum_{k=1}^{\infty} \|\chi_{k+1}^+\chi_k^- + \chi_{k+1}^-\chi_k^+\|_{L^1(\Omega)} < +\infty$, $\chi_{k+1}^+\chi_k^- + \chi_{k+1}^-\chi_k^+ \to 0$ in $L^1(\Omega)$ and pointwise almost everywhere.

Let $x \in I$. Then there is K such that $x \in I_K$. This implies $u_k(x) \notin B_\epsilon(\tilde{u})$ for all $k > K$. Here, the pointwise convergence of the whole sequence (z_k) is needed. The sum $\sum_{k=K+1}^{\infty}(\chi_{k+1}^+\chi_k^- + \chi_{k+1}^-\chi_k^+)(x)$ counts the number of switches between values larger than $\tilde{u} + \epsilon$ and smaller than $\tilde{u} - \epsilon$ from $u_k(x)$ to $u_{k+1}(x)$. Since this sum is finite for almost all $x \in \Omega$, there is only a finite number of such switches. Then there is $K' > K$ such that either $u_k(x) \geq \tilde{u} + \epsilon$ for all $k > K'$ or $u_k(x) \leq \tilde{u} - \epsilon$ for all $k > K'$. Set

$$S_K^+ := \{x \in I : \ u_k(t) \geq \tilde{u} + \epsilon \quad \forall k > K\},$$

$$S_K^- := \{x \in I : \ u_k(t) \leq \tilde{u} - \epsilon \quad \forall k > K\}.$$

The sequences (S_K^+) and (S_K^-) are increasing, and $\cup_{K\in\mathbb{N}}(S_K^+ \cup S_K^-) = I$.

Since $u_{k_n} \rightharpoonup u^*$, this implies $u^* \geq \tilde{u}+\epsilon$ on S_K^+ and $u^* \leq \tilde{u}-\epsilon$ on S_K^-. Since $\cup_{K\in\mathbb{N}}(S_K^+\cup S_K^-) = I$, this implies

$$u^*(x) \notin B_\epsilon(\tilde{u})$$

for almost all $x \in I$, which implies

$$((z(x), u^*(x)) \notin B_{\epsilon'}(\tilde{z}) \times B_\epsilon(\tilde{u})$$

for almost all $x \in \Omega$. Since we can cover the complement of $\operatorname{gph}\mathcal{G}$ by countably many such sets, the claim follows. □

For convex functions g, the result above is equivalent to

$$-\nabla f(u^*) \in \partial g(u^*),$$

see, e.g., [3, Cor. 27.9].

4.4 Pointwise convergence of iterates

So far we were able to show that weak limit points of iterates (u_k) satisfy a certain inclusion in a pointwise sense. However, the resulting set in the limit might still be large or even unbounded in general. Assuming that $\mathcal{G}$ is (locally) single-valued on its components $\mathcal{G}^+, \mathcal{G}^-, \mathcal{G}^0$, we can show local pointwise convergence of a subsequence of iterates (u_{k_n}) to a weak limit point $u^* \in L^2(\Omega)$. In the next result this is illustrated for the map $\mathcal{G}^+$, however it can be shown for the components $\mathcal{G}^-, \mathcal{G}^0$ similarly. To this end, we set in the following $\chi_k^+ := \chi_{\{x\in\Omega:\, u_k(x)>0\}}$ with $\chi_k^+ \to \chi^+$ in $L^1(\Omega)$ and pointwise almost everywhere by Lemma 4.17.

Theorem 4.20. *Let $\bar{z} \in \operatorname{dom}(\mathcal{G}^+)$. Assume that $\mathcal{G}^+ : \mathbb{R} \to \mathbb{R}$ is single-valued and locally bounded on $B_\epsilon(\bar{z}) \cap \operatorname{dom}(\mathcal{G}^+)$ for some $\epsilon > 0$. Let $u_{k_n} \rightharpoonup u^*$ in $L^2(\Omega)$ and assume $\nabla f(u_{k_n})(x) \to \nabla f(u^*)(x)$ pointwise almost everywhere. For $\epsilon' \in (0, \epsilon]$ define the set*

$$I_{\epsilon'} := \left\{x \in \operatorname{supp}(\chi^+) : -\nabla f(u^*)(x) \in B_{\epsilon'}(z) \cap \operatorname{dom}(\mathcal{G}^+)\right\}.$$

Then

$$u_{k_n}(x) \to u^*(x)$$

holds for almost all $x \in I$. Furthermore, we have

$$u^*(x) \in \operatorname{prox}_{L^{-1}g}\left(\frac{1}{L}(Lu^*(x) - \nabla f(u^*)(x))\right) \text{ f.a.a. } x \in I_\epsilon.$$

Proof. Let $u_{k_n+1} \rightharpoonup u^*$ in $L^2(\Omega)$. By the assumption and Corollary 4.9 it holds $z_{k_n}(x) \to z(x) := -\nabla f(u^*)(x)$ pointwise almost everywhere. In addition, $u_{k_n+1} \rightharpoonup u^*$ in $L^2(\Omega)$ holds. Let $\epsilon' \in (0, \epsilon)$ be given. Take $x \in I_{\epsilon'}$ such that $z_{k_n}(x) \to z(x)$. Then there is $K > 0$ such that $|z_{k_n}(x) - \bar{z}| < \epsilon$ for all $k_n > K$. Since $x \in \operatorname{supp}(\chi^+)$ and $\chi_k^+ \to \chi^+$ in $L^1(\Omega)$ and pointwise almost everywhere there is $K' > 0$ such that $x \in \operatorname{supp}(\chi_k^+)$ for all $k > K'$. Hence, for k_n sufficiently large we have

$$z_{k_n}(x) \in B_\epsilon(\bar{z}) \cap \operatorname{dom}(\mathcal{G}^+).$$

Since $\mathcal{G}^+$ is single-valued, locally bounded and outer semicontinuous in $B_\epsilon(\bar{z}) \cap \operatorname{dom}(\mathcal{G}^+)$, it is continuous, see also [14, Cor. 5.20]. This implies

$$\lim_{n\to\infty} u_{k_n+1}(x) = \lim_{n\to\infty} \mathcal{G}^+(z_{k_n}(x)) = \mathcal{G}^+(\lim_{n\to\infty} z_{k_n}(x)) = \mathcal{G}^+(z(x)).$$

The continuity property mentioned above implies $\operatorname{conv}^\infty \mathcal{G}^+(z(x)) = \mathcal{G}^+(z(x))$. Then by Theorem 4.18, $\mathcal{G}^+(z(x)) = \{u^*(x)\}$, and the convergence $u_{k_n}(x) \to u^*(x)$ follows. The fixed-point property is a consequence of the closedness of the graph of the proximal operator. As $x \in I_{\epsilon'}$ was chosen arbitrary, and $I_\epsilon = \cup_{\epsilon' \in (0,\epsilon)} I_{\epsilon'}$, the claim is proven. □

The above result requires local boundedness of the set-valued map $\mathcal{G}$, which is not satisfied in general. For some interesting choices of g, e.g. $g(u) := |u|^p$, it can be proven, see Section 5. Let us give an example of a locally unbounded map $\mathcal{G}$ below.

Example 4.21. *Let $L > 0$ and define $g(u) := \delta_{\mathbb{Z}}(u) := \begin{cases} 0 & \text{if } u \in \mathbb{Z} \\ +\infty & \text{else.} \end{cases}$ with the associated map $\mathcal{G}_L$. Set $U := [-\frac{L}{2}, \frac{L}{2}]$. Then it holds that $\mathcal{G}(z) = \mathbb{Z}$ for all $z \in U$, i.e., $\mathcal{G}$ is clearly not locally bounded in the origin.*

4.5 Strong convergence of iterates

Many optimal control problems of type (P) include a smooth cost functional of form $u \to \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)}$, $\alpha > 0$. For the rest of the sequel, we will treat this term explicitly in the convergence analysis to obtain an almost everywhere and strong convergence of a subsequence. Therefore let $\tilde{g} : \mathbb{R} \to \mathbb{R}$ satisfy Assumption B and consider a sequence of iterates computed by

$$u_{k+1} := \operatorname*{arg\,min}_{u \in L^2(\Omega)} f(u_k) + \nabla f(u_k)(u - u_k) + \frac{L}{2}\|u - u_k\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} + \int_\Omega \tilde{g}(u(x))\,\mathrm{d}x. \quad (4.7)$$

The solution to (4.7) is now given by

$$u_{k+1}(x) \in \operatorname{prox}_{\frac{1}{L+\alpha}\tilde{g}}\left(\frac{1}{L+\alpha}(Lu_k(x) - \nabla f(u_k)(x))\right)$$

for almost every $x \in \Omega$. It follows that all the analysis that was done in this section still applies in this case and all results can be transferred except for a possible change of notation. Furthermore, we adapt the set-valued map $\mathcal{G} : \mathbb{R} \to \mathbb{R}$ from Lemma 4.7 which is then defined by

$$u \in \mathcal{G}(z) :\Longleftrightarrow u = \operatorname*{arg\,min}_{v \in \mathbb{R}} -zv + \frac{L}{2}(v - u)^2 + \frac{\alpha}{2}v^2 + \tilde{g}(v).$$

For simplicity we assume $\operatorname{dom}(\tilde{g}) = [-b, b]$ with $b \in (0, \infty]$, i.e., the subproblem (4.7) is equivalent to a box constrained optimization problem of form

$$u_{k+1} := \operatorname*{arg\,min}_{u \in L^2(\Omega)} f(u_k) + \nabla f(u_k)(u - u_k) + \frac{L}{2}\|u - u_k\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} + \int_\Omega \tilde{g}(u(x))\,\mathrm{d}x.$$

subject to $|u(x)| \leq b$ for almost every $x \in \Omega$. To obtain strong convergence of iterates in $L^1(\Omega)$ and an L-stationary condition almost everywhere, we need to put stronger and more restricting assumptions on $\tilde{g}$, as the next theorem shows. To this end, let us introduce the following extension of Assumption B.

Assumption B$^+$.

(B5) $\tilde{g}$ is C^1 on $(0, b)$ with $g'(b) := \lim_{u \nearrow b} g'(u)$.

(B6) For $s > 0$ there is $u_I := u_I(s) > 0$ such that $u \mapsto \frac{1}{2}u^2 + s\tilde{g}(u)$ is strictly convex on $[u_I, b]$.

First, we have the following necessary optimality condition for (4.7) due to Assumption (B5).

Corollary 4.22. *Let u_{k+1} be a solution to (4.7) and $\tilde{g}$ satisfy in addition (B5). Then the pointwise inequality in $\mathbb{R}$*

$$\begin{aligned}(\nabla f(u_k)(x) + L(u_{k+1}(x) - u_k(x)) + \alpha u_{k+1}(x)\\ + \tilde{g}'(u_{k+1}(x)))(v - u_{k+1}(x)) \geq 0\end{aligned}$$

for all $v \in [-b, b]$ holds for almost all $x \in I_{k+1}$.

Proof. Since $\operatorname{dom}(g) = [-b, b]$, minimizing the integrand in

$$\min_{u \in L^2(\Omega)} \int_\Omega \nabla f(u_k)(x)u(x) + \frac{L}{2}(u(x) - u_k(x))^2 + \frac{\alpha}{2}|u(x)|^2 + \tilde{g}(u(x))\,\mathrm{d}x. \tag{4.8}$$

pointwise is equivalent to solve the constrained problem

$$\min_{u:|u| \leq b} f(u_k)(x)u + \frac{L}{2}(u - u_k(x))^2 + \frac{\alpha}{2}|u|^2 + \tilde{g}(u)$$

in every Lebesgue point x. For $x \in I_{k+1}$ it holds $u_{k+1}(x) \neq 0$, and therefore above problem is differentiable. The claimed inequality is the corresponding necessary optimality condition. □

Let us for the rest of the sequel assume that $\tilde{g}$ satisfies (B5) and (B6) in addition to Assumption B. This enables us to give more information about the set-valued map $\mathcal{G}$ as the next result shows. That is, elements in $\mathcal{G}$ are (possibly unique) solutions of an associated variational inequality.

Lemma 4.23. *Let $u_0(\frac{1}{L+\alpha}), q_0(\frac{1}{L+\alpha})$ be constants as in Theorem 3.6 and $|u| \geq u_0(\frac{1}{L+\alpha})$. Then $u \in \mathcal{G}(z)$ satisfies the variational inequality*

$$(-z + \alpha u + \tilde{g}'(u))(v - u) \geq 0 \tag{4.9}$$

for all $v \in [-b, b]$. If in addition $|\frac{z+Lu}{L+\alpha}| \geq q_0\left(\frac{1}{L+\alpha}\right)$ and $u_0 \geq u_I$ with $u_I := u_I(\frac{1}{L+\alpha})$ as in (B6), then we have $u \in \mathcal{G}(z)$ if and only u satisfies (4.9).

Proof. Let us discuss the case $u \geq u_0$ only. If $u \in \mathcal{G}(z)$ for some $z \in \mathbb{R}$, then by definition

$$\begin{aligned} u &= \operatorname*{arg\,min}_{v\in\mathbb{R}} -zv + \frac{L}{2}(v-u)^2 + \frac{\alpha}{2}v^2 + \tilde{g}(v) \\ &= \operatorname*{arg\,min}_{|v|\leq b} -zv + \frac{L}{2}(v-u)^2 + \frac{\alpha}{2}v^2 + \tilde{g}(v) \\ &= \operatorname*{arg\,min}_{|v|\leq b} -(z+Lu)v + \frac{L+\alpha}{2}v^2 + \tilde{g}(v) \end{aligned}$$

Hence, by first order necessary optimality condition it holds

$$\begin{aligned} 0 &\leq \left(-(z+Lu) + (L+\alpha)u + \tilde{g}'(u)\right)(v-u) \\ &= (-z+\alpha u + \tilde{g}'(u))(v-u) \end{aligned}$$

for all $v \in [-b, b]$, which is the claim.

Assume $u_I \leq u_0$ holds, and let $u > 0$ satisfy (4.9), then u satisfies in particular

$$0 \leq (-z+\alpha u + \tilde{g}'(u))(v-u) = (-z - Lu + (\alpha+L)u + \tilde{g}'(u))(v-u)$$

for all $v \in [u_I, b]$, i.e., it is stationary to

$$\min_{v\in[u_I,b]} -zv + \frac{\alpha}{2}v^2 + \tilde{g}(v) \tag{4.10}$$

and also to

$$\min_{v\in[u_I,b]} -(z+Lu)v + \frac{L+\alpha}{2}v^2 + \tilde{g}(v).$$

By convexity u is the unique solution of the latter and since by assumption $\frac{z+Lu}{L+\alpha} \geq q_0\left(\frac{1}{L+\alpha}\right)$, it follows from Theorem 3.6 that there is a global solution larger than u_0 to the unconstrained problem which together implies $u \in \mathcal{G}(z)$.

□

Lemma 4.24. *Let $\alpha > 0$. Assume u_{k+1} is a global solution to (4.7) with $|u_{k+1}(x)| \geq u_0 \geq u_I(\frac{1}{\alpha})$ for almost all $x \in I_{k+1}$, where $u_I(\frac{1}{\alpha})$ is as in (B6). Then there is a continuous mapping $G : L^2(\Omega) \to L^2(\Omega)$ such that*

$$u_{k+1} = \chi_{k+1} G\left(\frac{z_k}{\alpha}\right).$$

Proof. We set $s := \frac{1}{\alpha}$ and $u_I := u_I(s)$ as in (B6). Note that by assumptions the following holds for $\alpha > 0$ and $|u| \geq u_0 \geq u_I$:

$$u \in \mathcal{G}(z) \Longleftrightarrow u \in \operatorname{prox}_{(L+\alpha)^{-1}\tilde{g}}\left(\frac{z+Lu}{L+\alpha}\right) \Longrightarrow u \in \operatorname{prox}_{s\tilde{g}}^{u_I}\left(\frac{z}{\alpha}\right),$$

where we define, corresponding to (4.10),

$$u \in \operatorname{prox}_{s\tilde{g}}^{u_I}(z) :\Longleftrightarrow u = \operatorname*{arg\,min}_{|v|\in[u_I,b]} -zv + \frac{1}{2}v^2 + s\tilde{g}(v).$$

Due to assumption (B6) and Lemma 4.23, $u_{k+1}(x)$ is the only element in $\mathcal{G}(z_k(x)) \setminus \{0\}$ for almost all $x \in I_{k+1}$ and it holds $u_{k+1}(x) = \operatorname{prox}_{s\tilde{g}}^{u_I}\left(\frac{z_k(x)}{\alpha}\right)$. Set

$$z_I := \sup\{q > 0 : u_I = \operatorname{prox}_{s\tilde{g}}^{u_I}(q)\}.$$

It is easy to see that $\operatorname{prox}_{s\tilde{g}}^{u_I}$ is single-valued for $|z| > 0$. Since it is in addition outer semicontinuous and locally bounded for $|z| \geq z_I$, it is also continuous on $\{z : |z| \geq z_I\}$, see also [14, Corollary 5.20]. Let $u \in \operatorname{prox}_{s\tilde{g}}^{u_I}(z)$. By optimality of u we have

$$-zu + \frac{1}{2}u^2 + s\tilde{g}(u) \leq -z \cdot \operatorname{sign}(u)u_I + \frac{1}{2}u_I^2 + s\tilde{g}(u_I).$$

Dividing by $|u| > 0$, we get

$$\frac{1}{2}|u| \leq \left(\frac{u - \operatorname{sign}(u)u_I}{|u|}\right) z + \frac{u_I^2}{|u|} + s\frac{\tilde{g}(u_I) - \tilde{g}(u)}{|u|}.$$

Having in mind that $\frac{u_I}{|u|} \leq 1$, the growth estimate $|\operatorname{prox}_{s\tilde{g}}^{u_I}(z)| \leq 2|z| + c$ for all $|z| \geq z_I$ with some $c > 0$ independent of z follows.

Let $l : \mathbb{R} \to \mathbb{R}$ denote a continuous function defined by

$$l(z) := \begin{cases} \operatorname{prox}_{sg}^{u_I}(z) & \text{if } |z| \geq z_I, \\ \frac{u_I}{z_I} z & \text{if } |z| \leq z_I. \end{cases}$$

Define

$$G : L^2(\Omega) \to L^2(\Omega),\ G(z)(x) = l(z(x))$$

for $z : \Omega \to \mathbb{R}$. Then by a well-known result, see e.g. [1, Theorem 3.1], the superposition operator G is continuous from $L^2(\Omega) \to L^2(\Omega)$ and the claim follows. □

Now, we are able to prove strong convergence of a subsequence of (u_k) similar to [16, Thm. 3.17].

Theorem 4.25. *Suppose complete continuity of ∇f and let $(u_k) \subset L^2(\Omega)$ be a sequence generated by Algorithm 4.7 with weak limit point u^*. Under the same assumptions as in Lemma 4.24 u^* is a strong sequential limit point of (u_k) in $L^1(\Omega)$.*

Proof. By Lemma 4.24 there exists a continuous mapping $G : L^2(\Omega) \to L^2(\Omega)$ such that $u_{k+1} = \chi_{k+1}\left(G(\frac{z_k}{\alpha})\right)$. Let $u_{k_n} \rightharpoonup u^*$ in $L^2(\Omega)$. Again, by Theorem 4.4 and complete continuity of ∇f, we obtain strong convergence of the sequence

$$z_{k_n} := -\left(\nabla f(u_{k_n}) + L(u_{k_n+1} - u_k)\right) \to -\nabla f(u^*) =: z^*$$

in $L^2(\Omega)$ as well as $\chi_k \to \chi$ in $L^p(\Omega)$ for all $p < \infty$ and $u_{k_n+1} \rightharpoonup u^*$. Then the convergence

$$u_{k_n+1} = \chi_{k_n+1} G\left(\frac{1}{\alpha} z_{k_n}\right) \to \chi G\left(\frac{1}{\alpha} z^*\right)$$

in $L^1(\Omega)$ follows by Hölder's inequality. Since strong and weak limit points coincide, it follows $u_{k_n} \to u^*$ in $L^1(\Omega)$ and

$$u^* = \chi G\left(-\frac{1}{\alpha}\nabla f(u^*)\right).$$

□

With the assumptions in Theorem 4.25 we can find an almost everywhere converging subsequence of iterates, i.e., $u_{k_n}(x) \to u^*(x)$ for almost every $x \in \Omega$. By the closedness of the mapping $\operatorname{prox}_{s\tilde{g}}$, we get

$$u^*(x) \in \operatorname{prox}_{\frac{1}{L+\alpha}\tilde{g}}\left(\frac{1}{L+\alpha}(Lu^*(x) - \nabla f(u^*)(x))\right) \text{ f.a.a } x \in \Omega, \tag{4.11}$$

i.e., u^* is L-stationary to the problem in almost every point. If $L = 0$ in (4.11), then we obtain by Lemma 4.2

$$u^*(x) = \arg\min_{u\in\mathbb{R}} f(u_k)(x)u(x) + \frac{\alpha}{2}|u(x)|^2 + \tilde{g}(u(x)) \quad \text{f.a.a. } x \in \Omega.$$

Hence, in this case u^* satisfies the Pontryagin maximum principle.

4.6 The proximal gradient method with variable stepsize

The convergence results of this section require the knowledge of the Lipschitz modulus L_f of ∇f. This can be overcome by line-search with respect to the parameter L subject to a suitable decrease condition, which is a widely applied technique.

Algorithm 4.26 (Proximal gradient with variable step-size)**.** Choose $\eta > 0$ and $u_0 \in U_{ad}$. Set $k = 0$.

1. Determine $L_k \geq 0$ and u_{k+1} as global solution of

$$\min_{u\in L^2(\Omega)} f(u_k) + \nabla f(u_k)(u - u_k) + \frac{L_k}{2}\|u - u_k\|^2_{L^2(\Omega)} + j(u)$$

   such that

$$\eta\|u_{k+1} - u_k\|^2_{L^2(\Omega)} \leq (f(u_k) + j(u_k)) - (f(u_{k+1} + j(u_{k+1})) \tag{4.12}$$

   is satisfied.

2. Set $k := k + 1$, repeat.

The convergence results as in Theorem 4.4 can be carried over. Then theorem 4.4 holds without the assumption $L > L_f$. The assumptions $1/L > s_0$ has to be replaced by $(\limsup L_k)^{-1} > s_0$. This is satisfied if $s_0 = 0$, which is true by Theorem 3.6 if one of (B3.b), (B3.c) is valid.

5 Applications of the proximal gradient method

5.1 Optimal control with L^p control cost, $p \in (0, 1)$

In [16], the discussed proximal method was analyzed and applied to optimal control problems with L^0 control cost, i.e., $g(u) := \frac{\alpha}{2}u^2 + |u|_0$. In this section, we discuss the problem with $g(u) := \frac{\alpha}{2}u^2 + \beta|u|^p + \delta_{[-b,b]}$, where $p \in (0, 1)$ and $b \in (0, \infty]$ and consider

$$\min_{u\in L^2(\Omega)} f(u) + \frac{\alpha}{2}\|u\|_{L^2(\Omega)} + \beta\int_\Omega |u(x)|^p \,\mathrm{d}x \tag{5.1}$$

s.t.

$$u \in U_{ad} := \{u \in L^2(\Omega) : \ |u(x)| \leq b \ \text{ a.e. in } \Omega\}$$

with $\alpha \geq 0$, $\beta > 0$.

To find a solution to (5.1)with Algorithm 4.1, the subproblem, interpreted in terms of (4.7) with $\tilde{g} := |u|^p + \delta_{[-b,b]}$,

$$\min_{u\in U_{ad}} f(u_k) + \nabla f(u_k)(u - u_k) + \frac{L}{2}\|u - u_k\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|_{L^2(\Omega)} + \beta\int_\Omega |u(x)|^p \,\mathrm{d}x$$

has to be solved in every iteration. According to Theorem 4.2, u_{k+1} is a solution to (5.1) if and only if

$$u_{k+1}(x) \in \operatorname{prox}_{\frac{\beta}{L+\alpha}\tilde{g}}\left(\frac{1}{L+\alpha}(Lu_k(x) - \nabla f(u_k)(x))\right) \quad f.a.a.\ x \in \Omega.$$

Due to Theorem 3.6 it holds $u_{k+1}(x) = 0$ or $|u_{k+1}(x)| \geq u_0$ for all k. The particular choice of g allows to compute the constant u_0 explicitly by solving $\min_{u \neq 0} \frac{u}{2} + s\frac{g(u)}{2}$ and is given by

$$u_0\left(\frac{\beta}{\alpha+L}\right) = \min\left(b, \left(\frac{\alpha+L}{2\beta(1-p)}\right)^{\frac{1}{p-2}}\right)$$

as a consequence of Lemma 3.5.

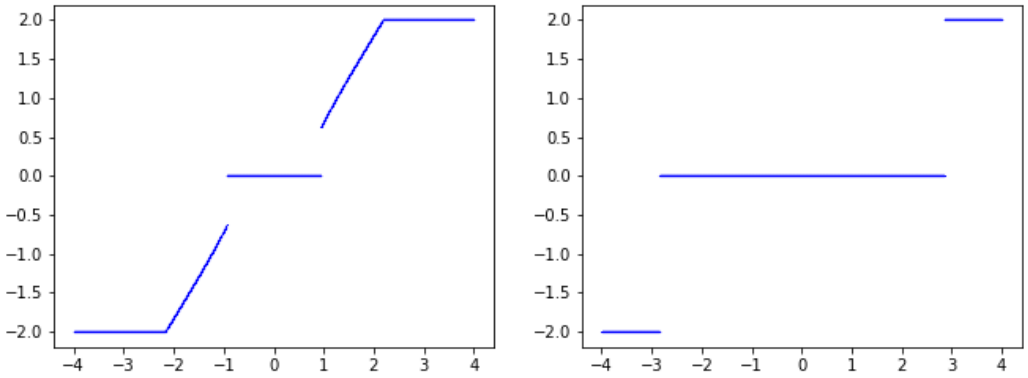

Figure 1: The mapping $\operatorname{prox}_{sg}(q)$ for parameters $(s, b, p) = (0.5, 2, 0.5)$ (left) and $(s, b, p) = (3, 2, 0.3)$ (right) with $\tilde{g}(u) := |u|^p + \delta_{[-b,b]}$.

We recall the definition of the set-valued map $\mathcal{G} : \mathbb{R} \to \mathbb{R}$, which reads in this case

$$u \in \mathcal{G}(z) := \mathcal{G}_{L,\alpha,s} :\Longleftrightarrow u = \operatorname*{arg\,min}_{|v| \leq b} -zv + \frac{L}{2}(u-v)^2 + \frac{\alpha}{2}v^2 + s|v|^p.$$

Note that g satisfies assumptions (B5) and (B6) due to its structure. This allows to give an equivalent but more precise characterization of $\mathcal{G}$ as Lemma 4.23 applies to $u_{k+1}(x)$ on I_{k+1}.

Corollary 5.1. *Let $u \geq u_0(\frac{\beta}{L+\alpha})$. Then $u \in \mathcal{G}(z_k(x)) \Longleftrightarrow u$ is a stationary point of*

$$\min_{u:|u| \leq b} -z_k(x)u + \frac{\alpha}{2}u^2 + \beta|u|^p$$

for almost all $x \in I_{k+1}$.

A visualization of $\mathcal{G}$ is given in Figure 2 below.

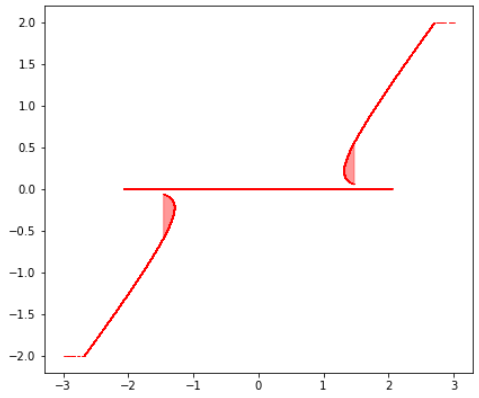

Figure 2: The union $(\mathcal{G}_0 \cup \mathcal{G}^+ \cup \mathcal{G}^-)(q)$ and the convexified map $(\mathcal{G}_0 \cup \overline{\text{conv}}\mathcal{G}^+ \cup \overline{\text{conv}}\mathcal{G}^-)(q)$ (filled area) (right) for parameters $(L, \alpha, s, b) = (0.1, 0.01, 0.01, 2)$ and $\tilde{g}(u) := |u|^{0.8} + \delta_{[-b,b]}$.

With a suitable choice of parameters, we can apply Theorem 4.25 to the L^p problem to obtain a strong convergent subsequence.

Corollary 5.2. *Let $\alpha > 0$ and (u_k) a sequence of iterates. Furthermore, assume $L \leq (\frac{2}{p} - 1)\alpha$. Then the assumptions of Theorem 4.25 are satisfied. If in addition ∇f is completely continuous from $L^2(\Omega)$ to $L^2(\Omega)$, then every weak sequential limit point $u^* \in L^2(\Omega)$ is a strong sequential limit point in $L^1(\Omega)$.*

Proof. Let $k \in \mathbb{N}$. It holds $|u_{k+1}(x)| \geq u_0$ with $u_0 := \min\left(b, \left(\frac{\alpha+L}{2\beta(1-p)}\right)^{\frac{1}{p-2}}\right)$ on I_{k+1}. A short calculation yields that the assumptions on the parameters imply

$$\left(\frac{\alpha + L}{2\beta(1-p)}\right)^{\frac{1}{p-2}} \geq \left(\frac{\alpha}{\beta p(1-p)}\right)^{\frac{1}{p-2}} =: u_I.$$

Here, u_I is the positive point of inflection of (5.1) and it holds that

$$h_{q,\frac{\beta}{\alpha}}(u) = -qu + \frac{1}{2}u^2 + \frac{\beta}{\alpha}|u|^p$$

is convex for all $q \in \mathbb{R}$ on $[u_I, \infty)$ and $(-\infty, u_I)$, respectively, which corresponds to Assumption (B6). The claim now follows by Lemma 4.24 and Theorem 4.25. □

5.2 Optimal control with discrete-valued controls

Let us investigate the optimization problem with optimal control taking discrete values. That is, we choose $g(u)$ as the indicator function of integers, i.e.,

$$g(u) := \delta_{\mathbb{Z}}(u) := \begin{cases} 0 & \text{if } u \in \mathbb{Z}, \\ \infty & \text{else} \end{cases}.$$

The problem now reads

$$\min_{u \in L^2(\Omega)} f(u) + \int_\Omega \delta_{\mathbb{Z}}(u(x))\,\mathrm{d}x. \tag{5.2}$$

Note, this choice satisfies Assumption (B3.c). Applying Algorithm 4.1, the subproblem to solve is given by

$$\min_{u\in L^2(\Omega)} f(u_k) + \nabla f(u_k)(u-u_k) + \frac{L}{2}\|u-u_k\|^2_{L^2(\Omega)} + \int_\Omega \delta_{\mathbb{Z}}(u(x))\,\mathrm{d}x \tag{5.3}$$

and can be solved pointwise and explicitly. The analysis carried out in Chapter 4 is applicable, however, the special choice of g comes along with the following desirable result.

Lemma 5.3. *Let $u_k, u_{k+1} \in U_{ad}$ be consecutive iterates of Algorithm 4.1. Then*

$$\|u_{k+1}-u_k\|^p_{L^p(\Omega)} \geq \|u_{k+1}-u_k\|_{L^1(\Omega)}$$

holds for all $p \in [1,\infty)$.

Proof. The claim follows directly, since either $|u_{k+1}(x)-u_k(x)| = 0$ or $|u_{k+1}(x)-u_k(x)| \geq 1$ as the iterates are integer-valued in almost every point. □

Lemma 5.3 implies strong convergence of iterates (u_k) in $L^1(\Omega)$.

Theorem 5.4. *Let (u_k) be a sequence generated by Algorithm 4.1 with weak limit point u^*. Then $u_k \to u^*$ in $L^1(\Omega)$.*

Proof. As in the proof of Theorem 4.4, we get

$$\sum_{k=1}^{\infty} \|u_{k+1}-u_k\|^2_{L^2(\Omega)} < \infty$$

and therefore by Lemma 5.3

$$\sum_{k=1}^{\infty} \|u_{k+1}-u_k\|_{L^1(\Omega)} \leq \sum_{k=1}^{\infty} \|u_{k+1}-u_k\|^2_{L^2(\Omega)} < \infty$$

Thus, (u_k) is a Cauchy sequence in $L^1(\Omega)$ and therefore convergent in $L^1(\Omega)$ and it holds $u_k \to u^*$. □

6 Numerical experiments

In this section we finally apply the proximal gradient method to optimal control problems of type (P) and carry out numerical experiments for cost functionals with different g.

Let in the following denote f_l the reduced tracking-type functional

$$f_l(u) := \|S_l u - y_d\|^2_{L^2(\Omega)},$$

where S_l is the weak solution operator of the linear Poisson equation

$$-\Delta y = u \quad \text{in } \Omega,\ \ y = 0 \text{ on } \partial\Omega. \tag{6.1}$$

Further we define the nonlinear solution operator S_{sl} of the semilinear equation

$$-\Delta y + d(y) = u \quad \text{in } \Omega,\ \ y = 0 \text{ on } \partial\Omega \tag{6.2}$$

where $d(x,y) : \Omega \times \mathbb{R} \to \mathbb{R}$ is a C^2 Carathéodory Function with respect to y with $d(\cdot,0)$ in $L^p(\Omega)$, $n < p$, satisfying

1. $\dfrac{\partial d}{\partial y}(x,y) \geq 0$ for almost all $x \in \Omega$,

2. $\forall\, M > 0\, \exists C_M > 0$ s.t. $\left|\dfrac{\partial d(x,y)}{\partial y}\right| + \left|\dfrac{\partial^2 d(x,y)}{\partial y^2}\right| \leq C_M$ for almost all $x \in \Omega$ and $|y| \leq M$.

Then the equation is uniquely solvable, we refer to e.g., [8,9] In addition, we define

$$f_{sl} := \|S_{sl}(u) - y_d\|^2_{L^2(\Omega)}.$$

Furthermore, we choose $\Omega := (0,1)^2$ to be the underlying domain in all following examples. To solve the partial differential equation, the domain is divided into a regular triangular mesh and the PDE (6.1),(6.2) is discretized with piecewise linear finite elements. The controls are discretized with piecewise constant functions on the triangles. The finite-element matrices were created with FEnicCS [12]. If not mentioned otherwise, the meshsize is approximately $h = \sqrt{2}/160 \approx 0.00884$. In each iteration a suitable constant $L_k > 0$ needs to be determined, that satisfies the decrease condition

$$\eta\|u_{k+1} - u_k\|^2_{L^2(\Omega)} \leq (f(u_k) + j(u_k)) - (f(u_{k+1} + j(u_{k+1})), \tag{6.3}$$

see (4.12). Note, L_k^{-1} can be seen as a stepsize. In [16] several stepsize selection strategies are proposed. In our tests, we use a simple Armijo-like backtracking line search method (**BT**). That is, having an initial $L^0 > 0$ and a widening factor $\theta \in (0,1)$, determine L_k as the smallest accepted number of form $L^0\theta^{-i},\ \ i = 0, 1, ...$ This method ensures a decrease in the objective values along the iterates, but it turns out to be very slow for large L_0, as the corresponding stepsize L_k^{-1} gets smaller. For all our tests we choose

$$\eta = 10^{-4}, \quad \theta = 0.5.$$

The stopping criterion is as follows:

If $|f(u_{k+1}) + g(u_{k+1}) - (f(u_k) + g(u_k)| \leq 10^{-12}$:
STOP.

First, we consider control problems with L^p control cost, which were investigated in chapter 5.1, i.e., $g(u) := |u|^p + \delta_{[-b,b]}$ with $p \in (0,1)$.

Example 1 Let $g(u) := |u|^p + \delta_{[-b,b]}$ for $p \in (0,1)$ and find

$$\min_{u \in L^2(\Omega)} f_l(u) + \|u\|^2_{L^2(\Omega)} + \beta \int_\Omega g(u(x))\,\mathrm{d}x.$$

Setting $U_{ad} := \{L^2(\Omega) : |u(x)| \leq b \text{ a.e. on } \Omega\}$ the problem is equivalent to

$$\min_{u \in U_{ad}} f_l(u) + \|u\|^2_{L^2(\Omega)} + \beta \int_\Omega |u(x)|^p\,\mathrm{d}x.$$

The first example is taken from [16], where the proximal gradient algorithm was investigated for (sparse) optimal control problems with $L^0(\Omega)$ control cost. Since $\int_\Omega |u|^p dx \to \int_\Omega |u|^0 dx$ as $p \searrow 0$, we expect similar solutions. We choose the same problem data as in [11,16]. That is, if not mentioned otherwise,

$$y_d(x,y) = 10x\sin(5x)\cos(7y)$$

and $\alpha = 0.01,\ \ \beta = 0.01,\ \ b = 4.$

A computed solution for $p = 0.8$ is shown in Figure 3.

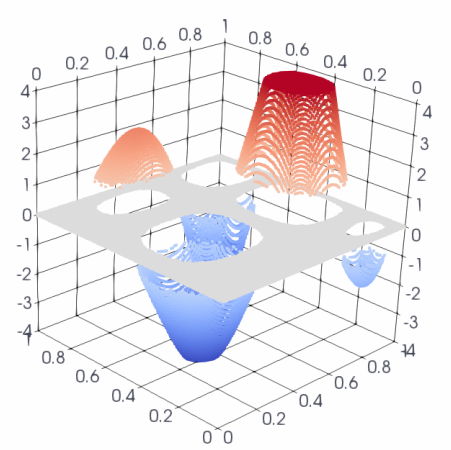

Figure 3: Solution u

Convergence for decreasing $p-$values. In the following we consider solutions for different values of p. We use the same data and discretization as above. We set $L_0 = 0.0001$. In Table

p	$J(u^*)$	$N_p(u^*)$	no. pde
0.5	5.3831	0.6711	15
0.3	5.3819	0.5725	15
0.1	5.3808	0.4841	15
0.01	5.3804	0.4482	15
0.001	5.3804	0.4448	15
0	5.38034	0.4445	15

Table 1: Decreasing values of p

1 it can be seen that $J(u^*)$ and $\int_\Omega |u^*|^p dx$ converge for decreasing values of p. The last row in Table 1 shows the result of applying the iterative hard-thresholding algorithm IHT-LS from [16] to the problem with $p = 0$, which is in agreement with our expectation. In the implementation we used a meshsize of $h = \sqrt{2}/500 \approx 0.0028$.

Discretization. Next, we solved the problem on different levels of discretization to investigate the influence. As can be seen in Table 2 the algorithm stays robust across different mesh sizes.

h	$J(u^*)$	$N_p(u^*)$	no. pde
0.071	5.2239	0.6371	13
0.035	5.3429	0.6581	15
0.0177	5.3732	0.6686	15
0.00884	5.3808	0.6704	15
0.00442	5.3827	0.6710	15
0.00221	5.3832	0.6711	15

Table 2: influence of meshsize

Convergence in the case $L > (2/p - 1)\alpha$. So far, in every experiment the assumption on the parameters was naturally satisfied, such that strong convergence of iterates can be proven according to Theorem 5.2. The numerical results confirmed the theory. We will now investigate the case where the assumption is not satisfied, i.e., we choose parameters such that $L > (2/p - 1)\alpha$. In the following we present the result for the problem parameters

$$\alpha = 0.001, \quad p = 0.9, \quad L_0 = 0.005.$$

Furthermore, we set $b = 6$. In our computations the algorithm needed very long to reach the stopping criteria $|J(u_{k+1}) - J(u_k)| \leq 10^{-12}$ as can be seen in Table 3. This might be due to the parameter choice and the step-size strategy. For smaller mesh-sizes more iterations are needed.

h	$J(u^*)$	$N_p(u^*)$	no. pde
0.00884	5.3567	1.1246	395
0.00442	5.3567	1.1247	601
0.00221	5.3567	1.1253	821

Table 3: performance for bad choice of parameters across different mesh-sizes

Recall, the problem in the analysis that comes with this choice of parameters is that the map $\mathcal{G}$ in Lemma 4.7 is not necessarily single-valued anymore on the set of points where an iterate is not vanishing, see also Figure 2. Let $u_I := u_I(\beta/\alpha) > 0$ denote the constant from Assumption (B6) and define the set

$$\Omega_{m,k} := \{x \in \Omega : 0 < |u_k(x)| < u_I\}.$$

Then $\Omega_{m,k}$ is the set of points for which the crucial assumption in Lemma 4.24 that implies single-valuedness of $\mathcal{G} \setminus \{0\}$ is not satisfied. In our numerical experiments, however, we made the observation that the measure of the set $\Omega_{m,k}$ is decreasing as $k \to \infty$, see Figure 4. Across different mesh-sizes h, the measure decreases and tends to zero along the iterations.

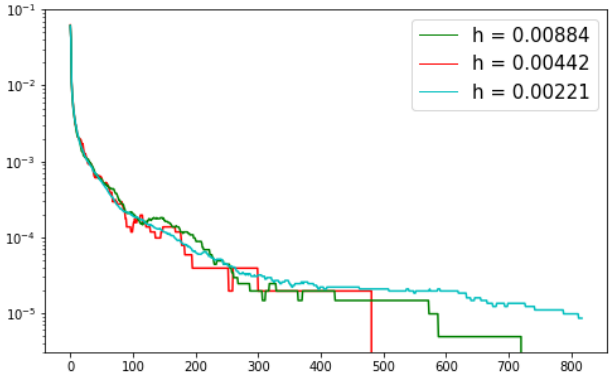

Figure 4: Measure of $\Omega_{m,k}$ at iteration k for different discretization levels

Unfortunately, we were not able to prove such a behavior in the analysis and have no theoretical evidence whether this can be expected in general. But assuming

$$|\Omega_{m,k}| \to 0$$

based on our numerical result, strong convergence of the sequence (u_k) can be concluded similar to Theorem 4.25.

Example 2 Let us now consider the semilinear problem

$$\min_{u\in U_{ad}} f_{sl}(u) + \|u\|^2_{L^2(\Omega)} + \beta \int_\Omega g(u(x))\,\mathrm{d}x$$

with $g(u) = |u|^p$, $p \in (0,1)$. This example can be found in [9] for semilinear control problems with L^1-cost. Here, f_{sl} is given by the standard tracking type functional $u \mapsto \|y_u - y_d\|^2_{L^2(\Omega)}$, where y_u is the solution of the semilinear elliptic state equation

$$-\Delta y + y^3 = u \quad \text{in } \Omega, \quad y = 0 \quad \text{on } \partial\Omega.$$

The data is given by $\alpha = 0.002$, $\beta = 0.03$, $b = 12$ and $y_d = 4\sin(2\pi x_1)\sin(\pi x_2)e^{x_1}$. We use the parameter $L_0 = 0.001$.

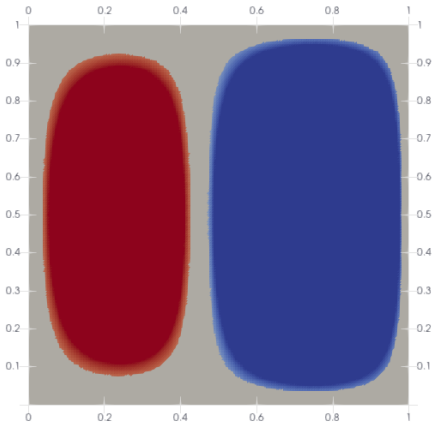

Figure 5: solution u of the semilinear optimal control problem with $g(u) := |u|^{0.5}$.

We made similar observations as in the linear case concerning the influence of discretization and different values of p. Also the behavior of the algorithm in case of a bad choice of parameters is as before (see Example 1).

Example 3 In this last test, we consider an optimal control problem with discrete-valued controls. That is, we choose

$$g(u) := \delta_{\mathbb{Z}}(u),$$

where δ_M denotes the indicator function of a set M, i.e., $\delta_M(u) := \begin{cases} 0 & \text{if } u \in M, \\ \infty & \text{else} \end{cases}$. Here, the subproblem in Algorithm 4.1 can be solved pointwise and explicitly. We adapt again the setting from Example 1. In Figure 6, a solution plot of the optimal control is displayed. We used exactly the same problem data as before in Example 1, but set $b = 2$ and $L_0 = 0.001$. Again, we find the algorithm is robust with respect to the discretization.

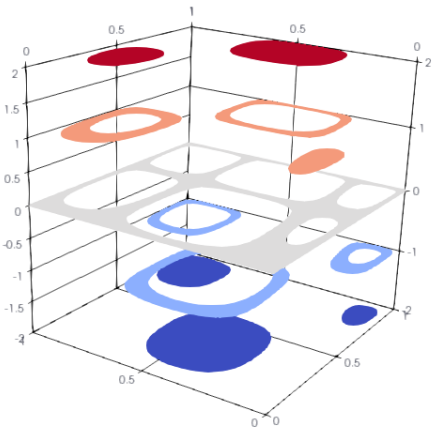

Figure 6: optimal control with discrete values

References

[1] J. Appell and P. P. Zabrejko. *Nonlinear superposition operators*, volume 95 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1990.

[2] J.-P. Aubin and H. Frankowska. *Set-valued analysis*, volume 2 of *Systems & Control: Foundations & Applications*. Birkhäuser Boston, Inc., Boston, MA, 1990.

[3] H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, New York, 2011.

[4] A. Beck. *Introduction to nonlinear optimization*, volume 19 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2014. Theory, algorithms, and applications with MATLAB.

[5] A. Beck and Y. C. Eldar. Sparsity constrained nonlinear optimization: optimality conditions and algorithms. *SIAM J. Optim.*, 23(3):1480–1509, 2013.

[6] J. F. Bonnans. On an algorithm for optimal control using Pontryagin's maximum principle. *SIAM J. Control Optim.*, 24(3):579–588, 1986.

[7] T. Breitenbach and A. Borzì. A sequential quadratic Hamiltonian method for solving parabolic optimal control problems with discontinuous cost functionals. *J. Dyn. Control Syst.*, 25(3):403–435, 2019.

[8] E. Casas. Boundary control of semilinear elliptic equations with pointwise state constraints. *SIAM J. Control Optim.*, 31(4):993–1006, 1993.

[9] E. Casas, R. Herzog, and G. Wachsmuth. Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional. *SIAM J. Optim.*, 22(3):795–820, 2012.

[10] I. Ekeland and R. Témam. *Convex analysis and variational problems*, volume 28 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, 1999. Translated from the French.

[11] K. Ito and K. Kunisch. Optimal control with $L^p(\Omega)$, $p \in [0, 1)$, control cost. *SIAM J. Control Optim.*, 52(2):1251–1275, 2014.

[12] H. P. Langtangen and A. Logg. *Solving PDEs in Python*, volume 3 of *Simula SpringerBriefs on Computing*. Springer, Cham, 2016. The FEniCS tutorial I.

[13] M. Nikolova, M. K. Ng, S. Zhang, and W.-K. Ching. Efficient reconstruction of piecewise constant images using nonsmooth nonconvex minimization. *SIAM J. Imaging Sci.*, 1(1):2–25, 2008.

[14] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.

[15] Y. Sakawa and Y. Shindo. On global convergence of an algorithm for optimal control. *IEEE Trans. Automat. Control*, 25(6):1149–1153, 1980.

[16] D. Wachsmuth. Iterative hard-thresholding applied to optimal control problems with $L^0(\Omega)$ control cost. *SIAM J. Control Optim.*, 57(2):854–879, 2019.