Necessary conditions for convergence rates of regularizations of optimal control problems

Daniel Wachsmuth and Gerd Wachsmuth

Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstraße 69, A–4040 Linz, Austria, Chemnitz University of Technology, Department of Mathematics, D-09107 Chemnitz, Germany

Abstract. We investigate the Tikhonov regularization of control constrained optimal control problems. We use a specialized source condition in combination with a condition on the active sets. In the case of high convergence rates, these conditions are necessary and sufficient.

Keywords: optimal control problem, inequality constraints, Tikhonov regularization, source condition

1 Introduction

In this article, we investigate regularization schemes for the following class of optimization problems:

$$\begin{aligned} &\text{Minimize} \quad \frac{1}{2}\|\mathcal{S}u - z\|_Y^2 + \beta\|u\|_{L^1(\Omega)} \\ &\text{such that} \quad u \in L^2(\Omega) \quad \text{and} \quad u_a \le u \le u_b \text{ a.e. on } \Omega. \end{aligned} \tag{P}$$

Here, Ω is a measurable subset of $\mathbb{R}^n$, $n \ge 1$, Y is a Hilbert space, $\mathcal{S} : L^2(\Omega) \to Y$ a bounded linear operator, and the function $z \in Y$ is given. The parameter β is assumed to be non-negative. The control constraints $u_a, u_b \in L^\infty(\Omega)$ satisfy $u_a \le 0 \le u_b$.

This model problem can be interpreted as an optimal control problem as well as an inverse problem. In the point of view of inverse problems, the unknown u has to be constructed in order to reproduce given measurements z. The inequality constraints on u reflect certain a-priori knowledge about the solution $u^\dagger$ of the linear ill-posed equation $\mathcal{S}u = z$. If the problem at hand is seen as an optimal control problem, then u is the control, $\mathcal{S}u$ the state of the system, which has to be close to a desired state z, the inequality constraints restrict the feasible set and may hinder the state $\mathcal{S}u$ to reach the target z. If the parameter β is positive, then the resulting optimal control will be sparse, that is, its support is a possibly small subset of Ω.

The resulting optimization problem (P) is nevertheless ill-posed if $\mathcal{S}$ is not continuously invertible. Due to the control constraints, problem (P) still possesses a solution, which is even unique if $\mathcal{S}$ is injective. However, the solution may be

unstable with respect to perturbations in the problem data, for instance in the given state z. Here small perturbations due to measurement errors may lead to large changes in the solution. Consequently, any numerical approximation of (P) is challenging to solve and numerical approximations of solutions may converge arbitrarily slow. Let us note, that a positive value of β does not make the problem well-posed. This is due to the fact, that $L^1(\Omega)$ is not a dual space and hence bounded sets in $L^1(\Omega)$ are not compact w.r.t. the weak(-star) topology, see also the discussions in [7,8].
In order to overcome this difficulty, we apply common ideas from inverse problem theory. We will study a regularization of the type

$$\begin{aligned} &\text{Minimize} \quad \frac{1}{2}\|\mathcal{S}u - z\|_Y^2 + \beta\|u\|_{L^1(\Omega)} + \frac{\alpha}{2}\|u\|_{L^2(\Omega)}^2 \\ &\text{such that} \quad u \in L^2(\Omega) \quad \text{and} \quad u_a \le u \le u_b \text{ a.e. on } \Omega, \end{aligned} \tag{P$_\alpha$}$$

where $\alpha > 0$ is given. Clearly, the problems (P$_\alpha$) are uniquely solvable for $\alpha > 0$. Now, the question arises, whether their solutions u_α converge (weakly or strongly) to a solution u_0 of (P) for $\alpha \to 0$. Moreover, in the case of convergence, one is interested in proving convergence rates of $\|u_\alpha - u_0\|_{L^2(\Omega)}$ and $\|\mathcal{S}u_\alpha - \mathcal{S}u_0\|_Y$ under suitable assumptions.
In this work, we will prove necessary conditions for convergence rates. In some parts, the necessary conditions are similar to sufficient conditions found in earlier works [7,8]. Moreover, the result of Theorem 3 leads to a weakened sufficient condition for convergence rates.

1.1 Standing assumptions and notation

Let us fix the standing assumptions on the problem (P). We assume that $\mathcal{S} : L^2(\Omega) \to Y$ is linear and continuous. In many applications this operator $\mathcal{S}$ is compact. Furthermore, we assume that the Hilbert space adjoint operator $\mathcal{S}^\star$ maps into $L^\infty(\Omega)$, i.e., $\mathcal{S}^\star \in \mathcal{L}(Y, L^\infty(\Omega))$. These assumptions imply that the range of $\mathcal{S}$ is closed in Y if and only if the range of $\mathcal{S}$ is finite-dimensional, see [8, Prop. 2.1]. Hence, up to trivial cases, (P) is ill-posed. A typical example for $\mathcal{S}$ is the solution operator of the Poisson problem with homogeneous Dirichlet boundary conditions.
The set of feasible functions u is given by

$$U_{ad} := \{u \in L^2(\Omega) : \ u_a \le u \le u_b \text{ a.e. on } \Omega\}.$$

The problem (P$_\alpha$) is uniquely solvable for $\alpha > 0$. We will denote its solution by u_α, with the corresponding state $y_\alpha := \mathcal{S}u_\alpha$ and adjoint state $p_\alpha := \mathcal{S}^\star(z - y_\alpha)$. There is a unique solution of (P) with minimal $L^2(\Omega)$ norm, see [8, Thm. 2.3, Lem. 2.7]. This solution and the associated state and adjoint state will be denoted by u_0, y_0 and p_0, respectively. Note that the weak convergence $u_\alpha \rightharpoonup u^\star$ in $L^2(\Omega)$, where $u^\star$ is a solution of (P) already implies $u^\star = u_0$, see [8, Rem. 3.3].

1.2 Optimality conditions

As both problems (P) and (P_α) are convex, their solutions can be characterized by the following necessary and sufficient optimality conditions:

Theorem 1 ([7, Lemma 2.2]). *Let $\alpha \geq 0$ be given, and let u_α be a solution of (P_α) (or (P) in the case $\alpha = 0$).*
Then, there exists a subgradient $\lambda_\alpha \in \partial \|u_\alpha\|_{L^1(\Omega)}$, such that the variational inequality

$$(\alpha\, u_\alpha - p_\alpha + \beta\, \lambda_\alpha,\, u - u_\alpha) \geq 0 \quad \forall u \in U_{ad}, \tag{1}$$

is satisfied, where $p_\alpha = \mathcal{S}^\star(z - \mathcal{S}u_\alpha)$ is the associated adjoint state.

Here, $(\cdot,\cdot)$ refers to the scalar product in $L^2(\Omega)$.
Standard arguments (see [6, Section 2.8]) lead to a pointwise a.e. interpretation of the variational inequality, which in turn implies the following relation between u_α and p_α in the case $\alpha > 0$:

$$u_\alpha(x) = \begin{cases} u_a(x) & \text{if } p_\alpha(x) < \alpha\, u_a(x) - \beta \\ \frac{1}{\alpha}(p_\alpha(x) + \beta) & \text{if } \alpha\, u_a(x) - \beta \leq p_\alpha(x) \leq -\beta \\ 0 & \text{if } |p_\alpha(x)| < \beta \\ \frac{1}{\alpha}(p_\alpha(x) - \beta) & \text{if } \beta \leq p_\alpha(x) \leq \alpha\, u_b(x) + \beta \\ u_b(x) & \text{if } \alpha\, u_b(x) + \beta < p_\alpha(x) \end{cases} \quad \text{a.e. on } \Omega. \tag{2}$$

In the case $\alpha = 0$, we have

$$u_0(x) \begin{cases} = u_a(x) & \text{if } p_0(x) < -\beta \\ \in [u_a(x), 0] & \text{if } p_0(x) = -\beta \\ = 0 & \text{if } |p_0(x)| < \beta \\ \in [0, u_b(x)] & \text{if } p_0(x) = \beta \\ = u_b(x) & \text{if } \beta < p_0(x) \end{cases} \quad \text{a.e. on } \Omega. \tag{3}$$

Note that if $\beta = 0$, one obtains $u_0(x) \in [u_a(x), u_b(x)]$ where $p_0(x) = 0$ in (3). This implies that $u_0(x)$ is uniquely determined by $p_0(x)$ on the set, where it holds $|p_0(x)| \neq \beta$.

2 Sufficient conditions for convergence rates

Let us first recall the sufficient conditions for convergence rates as obtained in [8]. We will work with the following assumption. There we denote by $\mathrm{proj}_{[a,b]}(v)$ the projection of the real number v onto the interval $[a, b]$.

Assumption 2 *Let u_0 be a solution of (P). Let us assume that there exist a measurable set $I \subset \Omega$, a function $w \in Y$, and positive constants κ, c such that it holds:*

1. **(source condition)** *$I \supset \{x \in \Omega : \ |p_0(x)| = \beta\}$, and for almost all $x \in I$*

$$u_0(x) = \begin{cases} \operatorname{proj}_{[u_a(x),0]}\big((\mathcal{S}^\star w)(x)\big) & \text{if } \beta > 0,\ p_0(x) \le -\frac{\beta}{2}, \\ \operatorname{proj}_{[0,u_b(x)]}\big((\mathcal{S}^\star w)(x)\big) & \text{if } \beta > 0,\ p_0(x) \ge \frac{\beta}{2}, \\ \operatorname{proj}_{[u_a(x),u_b(x)]}\big((\mathcal{S}^\star w)(x)\big) & \text{if } \beta = 0. \end{cases} \tag{4}$$

2. **(structure of active set)** *$A = \Omega \setminus I$ and for all $\epsilon > 0$*

$$\begin{aligned} \operatorname{meas}\big(\{x \in A : \ 0 < \big||p_0(x)| - \beta\big| < \epsilon\}\big) &\le c\,\epsilon^\kappa \quad \text{if } w \ne 0, \\ \operatorname{meas}\big(\{x \in A : \ 0 < \ |p_0(x)| - \beta \ < \epsilon\}\big) &\le c\,\epsilon^\kappa \quad \text{if } w = 0. \end{aligned} \tag{5}$$

Some remarks are in order. The first part of the assumption is analogous to source conditions in inverse problems: we assume that on the set $I \subset \Omega$ the solution u_0 is the restriction to I of a certain pointwise projection of an element in the range of $\mathcal{S}^\star$. This part of the condition is different from other conditions in the literature: in our earlier work [8] we used the assumption $u_0(x) = \operatorname{proj}_{[u_a(x),u_b(x)]}\big((\mathcal{S}^\star w)(x)\big)$ on I. However, in the light of the derivation of necessary conditions it turns out that such a condition can be weakened without losing anything with respect to convergence rates. In works on inverse problems [3,5], the source condition $u_0 = \operatorname{proj}_{U_{ad}}(\mathcal{S}^\star w)$ is used, which is retained as the special case $I = \Omega$ in Assumption 2.
The assumption (5) (without the second alternative) on the active sets was already employed to obtain regularization error estimates [7,8], error estimates for finite-element discretizations of (P) [2], as well as stability results of bang-bang controls [4]. Note that in the case $\beta = 0$, both conditions in (5) are equivalent. However, if $\beta > 0$ and $w = 0$ (in particular, if I has measure zero), the second alternative provides a weaker condition than the first one. Hence, condition (5) is weaker than the condition used in our earlier work [8].

Theorem 3. *Let Assumption 2 be satisfied.*
Let d be defined as

$$d = \begin{cases} \frac{1}{2-\kappa} & \text{if } \kappa \le 1, \\ 1 & \text{if } \kappa > 1 \text{ and } w \ne 0, \\ \frac{\kappa+1}{2} & \text{if } \kappa > 1 \text{ and } w = 0. \end{cases}$$

Then there is $\alpha_{max} > 0$ and a constant $c > 0$, such that

$$\begin{aligned} \|y_0 - y_\alpha\|_Y &\le c\,\alpha^d \\ \|p_0 - p_\alpha\|_{L^\infty(\Omega)} &\le c\,\alpha^d \\ \|u_0 - u_\alpha\|_{L^2(\Omega)} &\le c\,\alpha^{d-1/2} \end{aligned}$$

holds for all $\alpha \in (0, \alpha_{max}]$.

Under the assumptions of the theorem, one can prove also convergence rates for $\|u_\alpha - u_0\|_{L^1(A)}$ [8].

Proof. The proof is analogous to the proof of [8, Thm. 3.14]. We have to take into account the modification of the source condition (4) in the case $\beta > 0$ and the modification of (5) in the case $w = 0$. By [8, Lemma 2.12], we have

$$\|y_0 - y_\alpha\|_Y^2 + \alpha\,\|u_0 - u_\alpha\|_{L^2(\Omega)}^2 \le \alpha\,(u_0,\, u_0 - u_\alpha). \tag{6}$$

Since U_{ad} is bounded, we obtain $\|p_0 - p_\alpha\|_{L^\infty(\Omega)} \le c\,\alpha^{1/2}$ for some $c > 0$ independent of α.
Let now α be small enough such that $\|p_0 - p_\alpha\|_{L^\infty(\Omega)} < \beta/2$. This implies that p_0 and p_α have the same sign on the set $\{x \in I : \ |p_0(x)| \ge \beta/2\}$. Consequently, u_0 and u_α have the same sign on this set, too. Moreover, on the set $\{x \in I : |p_0(x)| < \beta/2\}$ it holds $|p_\alpha| < \beta$, and hence $u_\alpha = 0 = u_0$ on this set. This yields

$$(\chi_I u_0,\, u_0 - u_\alpha) \le (\chi_I \mathcal{S}^\star w,\, u_0 - u_\alpha)$$

for $\alpha > 0$ small enough. Note that in case of $w = 0$, the right-hand side in the previous estimate vanishes and it remains to estimate $(\chi_A\, u_0,\, u_0 - u_\alpha)$. Taking into account that $u_0(x) = 0$ whenever $|p_0(x)| < \beta$, the weakend estimate (5) is sufficient in this case. Arguing as in the proof of [8, Thm. 3.14] proves the claim. □

3 Necessary conditions for convergence rates

3.1 Necessity of the source condition (4)

Theorem 4. *Let us suppose that $\|y_\alpha - y_0\|_Y = O(\alpha)$ with $y_0 = \mathcal{S}u_0$. Then there exists $w \in Y$ such that*

$$u_0(x) = \begin{cases} \operatorname{proj}_{[u_a(x),0]}\big((\mathcal{S}^\star w)(x)\big) & \text{if } \beta > 0,\ p_0(x) = -\beta, \\ \operatorname{proj}_{[0,u_b(x)]}\big((\mathcal{S}^\star w)(x)\big) & \text{if } \beta > 0,\ p_0(x) = +\beta, \\ \operatorname{proj}_{[u_a(x),u_b(x)]}\big((\mathcal{S}^\star w)(x)\big) & \text{if } \beta = 0,\ p_0(x) = 0. \end{cases}$$

If moreover $\|y_\alpha - y_0\|_Y = o(\alpha)$, then $u_0 = 0$ on $\{x \in \Omega : \ |p_0(x)| = \beta\}$, i.e. $w = 0$.

This result shows that the source condition (4) is necessary on the set $\{x \in \Omega : |p_0(x)| = \beta\}$.

Proof. Let us prove the claim in the case $\beta > 0$. The result in the case $\beta = 0$ can be proved with obvious modifications. Let us take a test function $u \in U_{ad}$ defined as

$$u(x) \begin{cases} = u_a(x) & \text{if } p_0(x) < -\beta, \\ \in [u_a(x), 0] & \text{if } p_0(x) = -\beta, \\ = 0 & \text{if } |p_0(x)| < \beta, \\ \in [0, u_b(x)] & \text{if } p_0(x) = \beta, \\ = u_b(x) & \text{if } p_0(x) > \beta. \end{cases}$$

Due to the relation

$$\lambda_0 = \mathrm{proj}_{[-1,1]}\left(\frac{1}{\beta}p_0\right), \tag{7}$$

which is a consequence of the necessary optimality condition, see [1] for a proof, we obtain $\lambda_0 = \pm 1$ where $p_0 = \pm\beta$. Hence it holds

$$(-p_0,\ u - u_0) = (-\beta\,\lambda_0,\ u - u_0) = \beta\,\|u_0\|_{L^1(\Omega)} - \beta\,\|u\|_{L^1(\Omega)} \tag{8}$$

for u as above.
Since $\lambda_0 \in \partial\|u_0\|_{L^1(\Omega)}$, we obtain

$$(\lambda_0,\ u_\alpha - u_0) \le \|u_\alpha\|_{L^1(\Omega)} - \|u_0\|_{L^1(\Omega)}. \tag{9}$$

Using the optimality of u_α and the relation $-p_\alpha = -p_0 + \mathcal{S}^\star\mathcal{S}(u_\alpha - u_0)$ we get

$$(-p_0 + \mathcal{S}^\star\mathcal{S}(u_\alpha - u_0) + \alpha\,u_\alpha,\ u - u_\alpha) + \beta\,\|u\|_{L^1(\Omega)} - \beta\,\|u_\alpha\|_{L^1(\Omega)} \ge 0.$$

Adding $(-p_0 + \beta\,\lambda_0,\ u_\alpha - u_0) \ge 0$ to the left-hand side yields

$$\begin{aligned}(\mathcal{S}^\star\mathcal{S}(u_\alpha - u_0) + \alpha\,u_\alpha,\ u - u_\alpha) + (-p_0,\ u - u_0) + (\beta\,\lambda_0,\ u_\alpha - u_0)&\\ +\beta\,\|u\|_{L^1(\Omega)} - \beta\,\|u_\alpha\|_{L^1(\Omega)} &\ge 0.\end{aligned}$$

Using (8) and (9) we obtain

$$(\mathcal{S}^\star\mathcal{S}(u_\alpha - u_0) + \alpha\,u_\alpha,\ u - u_\alpha) \ge 0.$$

Due to the assumptions of the theorem, the functions $\frac{1}{\alpha}(\mathcal{S}(u_\alpha - u_0)) = \frac{1}{\alpha}(y_\alpha - y_0)$ are uniformly bounded for $\alpha \searrow 0$. As a consequence, $\alpha \searrow 0$ implies

$$(\mathcal{S}^\star\dot{y}_0 + u_0,\ u - u_0) \ge 0$$

for any weak subsequential limit $\dot{y}_0$ of $\frac{1}{\alpha}(y_\alpha - y_0)$. Due to the construction of the test function u, we obtain

$$u_0 = \begin{cases}\mathrm{proj}_{[u_a,0]}(\mathcal{S}^\star\dot{y}_0) & \text{where } p_0 = -\beta,\\ \mathrm{proj}_{[0,u_b]}(\mathcal{S}^\star\dot{y}_0) & \text{where } p_0 = +\beta.\end{cases}$$

If $\|y_\alpha - y_0\|_Y = o(\alpha)$ then $\frac{1}{\alpha}(y_\alpha - y_0) \to 0$ strongly in Y for $\alpha \to 0$, hence $\dot{y}_0 = 0$, and $u_0 = 0$ on the set $\{|p_0| = \beta\}$. □

As can be seen from the proof, the element that realizes the source condition can be interpreted as the (weak) directional derivate of $\alpha \mapsto y_\alpha$ at $\alpha = 0$.
The result of the theorem resembles known results of necessity of the source condition in linear inverse problems, see e.g. [3,5].

3.2 Necessity of the condition (5) on the active set

In this section, we want to prove the necessity of (5) in the case of high convergence rates $d > 1$. In this case, we have $w = 0$, see Theorem 4. It remains to show that the second condition in (5) is necessary to obtain convergence rates $d > 1$. Hence, we derive a bound on

$$\mu(\epsilon) := |\{x \in \Omega :\ 0 < |p_0(x)| - \beta < \epsilon\}|\,,$$

which is the measure of a subset of

$$A := \{x \in \Omega :\ \beta < |p_0(x)|\}.$$

For $\alpha > 0$ let $\tilde{u}_\alpha$ denote the unique solution of

$$\min_{u \in U_{ad}} -(u, p_0) + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} + \beta\|u\|_{L^1(\Omega)}. \qquad (\mathrm{P}^{\mathrm{aux}}_\alpha)$$

Analogous to (2), we have the representation

$$\tilde{u}_\alpha(x) = \begin{cases} u_a(x) & \text{if } p_0(x) < \alpha\, u_a(x) - \beta \\ \frac{1}{\alpha}(p_0(x) + \beta) & \text{if } \alpha\, u_a(x) - \beta \le p_0(x) \le -\beta \\ 0 & \text{if } |p_0(x)| < \beta \\ \frac{1}{\alpha}(p_0(x) - \beta) & \text{if } \beta \le p_0(x) \le \alpha\, u_b(x) + \beta \\ u_b(x) & \text{if } \alpha\, u_b(x) + \beta < p_0(x) \end{cases} \quad \text{a.e. on } \Omega. \qquad (10)$$

Let us first prove a relation between the convergence rates of $\|u_0 - \tilde{u}_\alpha\|_{L^2(A)}$ and $\mu(\epsilon)$ for $\alpha \to 0$ and $\epsilon \to 0$, respectively.

Lemma 5. *Let us assume that there is $\sigma > 0$ such that $u_a \le -\sigma < 0 < \sigma \le u_b$ a.e. on Ω. Then it holds: If $\|u_0 - \tilde{u}_\alpha\|_{L^2(A)} = O(\alpha^d)$, $d > 0$, for $\alpha \to 0$, then $\mu(\epsilon) = O(\epsilon^{2d})$ for $\epsilon \to 0$.*

Proof. Due to the pointwise representations of $\tilde{u}_\alpha$ and u_0 in (10) and (3), respectively, it holds

$$\|u_0 - \tilde{u}_\alpha\|^2_{L^2(A)} = \int_{\{\beta < p_0 < \alpha u_b + \beta\}} (u_b - \alpha^{-1}(p_0 - \beta))^2 + \int_{\{\alpha u_a - \beta < p_0 < -\beta\}} (u_a - \alpha^{-1}(p_0 + \beta))^2.$$

Due to the assumption on the control constraints we have

$$\begin{aligned} \int_{\{\beta < p_0 < \alpha u_b + \beta\}} (u_b - \alpha^{-1}(p_0 - \beta))^2 &\ge \int_{\{\beta < p_0 < \alpha\sigma/2 + \beta\}} (u_b - \alpha^{-1}(p_0 - \beta))^2 \\ &\ge \int_{\{\beta < p_0 < \alpha\sigma/2 + \beta\}} (\sigma/2)^2 \\ &\ge (\sigma/2)^2 \big|\{x \in \Omega : 0 < p_0(x) - \beta < \alpha\sigma/2\}\big|. \end{aligned}$$

Similarly, we obtain

$$\int_{\{\alpha u_a-\beta<p_0<-\beta\}}(u_a-\alpha^{-1}(p_0+\beta))^2 \geq (\sigma/2)^2\big|\{x\in\Omega : 0<-p_0(x)-\beta<\alpha\sigma/2\}\big|.$$

This implies

$$\|u_0-\tilde u_\alpha\|^2_{L^2(A)} \geq (\sigma/2)^2\mu(\alpha\sigma/2).$$

Hence if $\|u_0-\tilde u_\alpha\|_{L^2(A)} = O(\alpha^d)$ holds, then

$$\mu(\alpha\sigma/2) \leq O(\alpha^{2d}),$$

for $\alpha\to 0$, which proves the claim. □

Using the same arguments, we can prove the following result.

Corollary 6. *Let the requirements of Lemma 5 be satisfied. Let $p\in[1,\infty)$ be given. Then it holds*

$$\left(\frac{\sigma}{2}\right)^p \mu\left(\frac{\sigma}{2}\alpha\right) \leq \|u_0-\tilde u_\alpha\|^p_{L^p(A)} \leq M^p\,\mu(M\alpha)$$

with $M=\max(\|u_a\|_{L^\infty(\Omega)},\|u_b\|_{L^\infty(\Omega)})$.

Lemma 7. *Let $\tilde u_\alpha$ be defined as above. Then it holds*

$$\alpha\|\tilde u_\alpha-u_\alpha\|^2_{L^2(\Omega)}+\|y_0-y_\alpha\|^2_Y \leq (p_0-p_\alpha,\,\tilde u_\alpha-u_0).$$

Proof. Since u_α and $\tilde u_\alpha$ solve (P_α) and $(\mathrm{P}^{\mathrm{aux}}_\alpha)$, respectively, we have

$$\begin{aligned}(\alpha u_\alpha-p_\alpha+\beta\lambda_\alpha,\,\tilde u_\alpha-u_\alpha)&\geq 0,\\ (\alpha\tilde u_\alpha-p_0+\beta\tilde\lambda_\alpha,\,u_\alpha-\tilde u_\alpha)&\geq 0,\end{aligned}$$

with some $\tilde\lambda_\alpha\in\partial\|\tilde u_\alpha\|_{L^1(\Omega)}$. Due to the monotonicity of the subdifferential we have $(\lambda_\alpha-\tilde\lambda_\alpha,\,u_\alpha-\tilde u_\alpha)\geq 0$. This gives

$$\alpha\|\tilde u_\alpha-u_\alpha\|^2_{L^2(\Omega)} \leq (p_0-p_\alpha,\,\tilde u_\alpha-u_\alpha).$$

The identity

$$\begin{aligned}(p_0-p_\alpha,\,\tilde u_\alpha-u_\alpha)&=(p_0-p_\alpha,\,\tilde u_\alpha-u_0+u_0-u_\alpha)\\ &=(p_0-p_\alpha,\,\tilde u_\alpha-u_0)-\|y_0-y_\alpha\|^2_Y\end{aligned}$$

finishes the proof. □

Theorem 8. *Let us assume that there is $\sigma>0$ such that $u_a\leq-\sigma<0<\sigma\leq u_b$ a.e. on Ω. The we have the following implication: If*

$$\|u_0-u_\alpha\|_{L^2(\Omega)}=O(\alpha^{d-1/2}),\qquad \|y_0-y_\alpha\|_Y=O(\alpha^d)\ \text{for}\ \alpha\to 0$$

holds with $d>1$, then it follows

$$\mu(\epsilon)\leq O(\epsilon^{2d-1})\ \text{for}\ \epsilon\to 0.$$

Proof. Let us begin with

$$\begin{aligned}\|u_0 - \tilde{u}_\alpha\|^2_{L^2(\Omega)} &\le 2(\|u_0 - u_\alpha\|^2_{L^2(\Omega)} + \|u_\alpha - \tilde{u}_\alpha\|^2_{L^2(\Omega)}) \\ &\le O(\alpha^{2d-1}) + \alpha^{-1}(p_0 - p_\alpha,\, \tilde{u}_\alpha - u_0) \\ &\le O(\alpha^{2d-1}) + O(\alpha^{d-1})\|u_0 - \tilde{u}_\alpha\|_{L^2(\Omega)},\end{aligned}$$

which gives $\|u_0 - \tilde{u}_\alpha\|_{L^2(\Omega)} = O(\alpha^{d-1})$. Hence by Lemma 5, we obtain $\mu(\epsilon) = O(\epsilon^{2d-2})$. Let us note that the convergence rates imply $u_0(x) = 0$ if $|p_0(x)| = \beta$ by Theorem 4. Moreover, we have $u_0 = \tilde{u}_\alpha = 0$ on $\{x \in \Omega : |p_0(x)| \le \beta\}$ by (3) and (10). This implies $u_0 = \tilde{u}_\alpha = 0$ on the set $\{x \in \Omega :\ |p_0(x)| \le \beta\} = \Omega \setminus A$, cf. (10). Using the convergence rate $\|p_0 - p_\alpha\|_{L^\infty(\Omega)} = O(\alpha^d)$ and Corollary 6, we find

$$\begin{aligned}\alpha^{-1}|(p_0 - p_\alpha,\, \tilde{u}_\alpha - u_0)| &= O(\alpha^{d-1})\|\tilde{u}_\alpha - u_0\|_{L^1(\Omega)} \\ &= O(\alpha^{d-1})\|\tilde{u}_\alpha - u_0\|_{L^1(A)} \\ &\le O(\alpha^{d-1})\mu(M\,\alpha).\end{aligned}$$

Since by the above considerations we already got $\mu(\epsilon) = O(\epsilon^{2d-2})$ this gives

$$\|u_0 - \tilde{u}_\alpha\|^2_{L^2(\Omega)} = O(\alpha^{2d-1}) + O(\alpha^{3(d-1)}).$$

Repeating this process k times until $k(d-1) \ge 2d-1$ yields

$$\|u_0 - \tilde{u}_\alpha\|^2_{L^2(\Omega)} = O(\alpha^{2d-1}),$$

which finishes the proof. □

Together with Theorem 4, this result shows that the requirements of Theorem 3 for convergence rates $d > 1$ are sharp. It is an open question, whether the requirement (5) on the active set is also necessary for convergence rates $d \le 1$. In our opinion, this condition is too strong and has to be relaxed in order to obtain a characterization for convergence rates $d \le 1$.

3.3 Necessary conditions for exact reconstruction with $\alpha > 0$

Let us now investigate the case of exact reconstruction. That is, the solutions of the regularized problem u_α coincide with the (minimal L^2-norm) solution u_0 of the original problem.

Lemma 9. *Let us assume that $u_{\alpha^*} = u_0$ a.e. on Ω for some $\alpha^* > 0$. Then $u_\alpha = u_0$ a.e. on Ω for all $\alpha \in (0, \alpha^*)$.*

Proof. The claim follows from known monotonicity results: The mapping $\alpha \mapsto \|u_\alpha\|_{L^2}$ is monotonically decreasing, while $\alpha \mapsto \frac{1}{2}\|y_\alpha - y_d\|^2_Y + \beta\|u_\alpha\|_{L^1}$ is monotonically increasing from $(0, +\infty)$ to $\mathbb{R}$, see e.g. [8, Lemma 2.8]. □

Theorem 10. *Let us assume that there is $\sigma > 0$ such that $u_a \le -\sigma < 0 < \sigma \le u_b$ a.e. on Ω. Then the exact recovery $u_{\alpha^*} = u_0$ a.e. on Ω for some $\alpha^* > 0$ is equivalent to*

$$\left.\begin{aligned} u_0 = 0 \quad &\text{on } \{x \in \Omega : \ |p_0(x)| = \beta\} \qquad \text{and} \\ \mu(\epsilon) = \big|\{x \in \Omega :\ 0 < &|p_0(x)| - \beta < \epsilon\}\big| = 0 \end{aligned}\right\} \tag{11}$$

for some $\epsilon > 0$.

Proof. Let us assume $u_{\alpha^*} = u_0$ for some $\alpha^* > 0$. Lemma 9 and Theorem 4 imply $u_0(x) = 0$ for $x \in \{x \in \Omega : \ |p_0(x)| = \beta\}$. Moreover, due to $p_0 = p_{\alpha^*}$ we infer $u_0 = u_{\alpha^*} = \tilde{u}_{\alpha^*}$ from Lemma 7, where $\tilde{u}_{\alpha^*}$ is defined by (10). Hence, Corollary 6 implies $\mu(\sigma\,\alpha^*/2) = 0$.
To prove the converse, let (11) be satisfied for some $\epsilon > 0$. Using (6) we obtain

$$\alpha\,\|u_0 - u_\alpha\|^2_{L^2(\Omega)} \le \alpha\,(u_0, u_0 - u_\alpha) = \alpha\,(\chi_A\,u_0, u_0 - u_\alpha) \le C\,\alpha\,|A_\alpha|,$$

where $A = \{x \in \Omega : \ |p_0(x)| > \beta\}$ and $A_\alpha = \{x \in A : \ u_0(x) \ne u_\alpha(x)\}$. Arguing similarly as in [8, Corollary 3.13], we have $|A_\alpha| = 0$, and hence $\|u_0 - u_\alpha\|_{L^2(A)} = 0$ holds for $\alpha > 0$ small enough. □

In many applications, the adjoint state p_0 belongs to $C(\Omega)$. In this case, the result of Theorem 10 shows that an exact reconstruction is only possible if $|p_0(x)| \ne \beta$ for all $x \in \Omega$. This in turn implies either $u_0 \equiv u_a$ or $u_0 \equiv 0$ or $u_0 \equiv u_b$ on every connected component of Ω.

References

1. Casas, E., Herzog, R., Wachsmuth, G.: Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional. SIAM J. Optim. (2012), to appear
2. Deckelnick, K., Hinze, M.: A note on the approximation of elliptic control problems with bang-bang controls. Comput. Optim. Appl. 51(2), 931–939 (2012)
3. Engl, H.W., Hanke, M., Neubauer, A.: Regularization of inverse problems, Mathematics and its Applications, vol. 375. Kluwer Academic Publishers Group, Dordrecht (1996)
4. Felgenhauer, U.: On stability of bang-bang type controls. SIAM J. Control Optim. 41(6), 1843–1867 (2003)
5. Neubauer, A.: Tikhonov-regularization of ill-posed linear operator equations on closed convex sets. J. Approx. Theory 53(3), 304–320 (1988)
6. Tröltzsch, F.: Optimal Control of Partial Differential Equations, Graduate Studies in Mathematics, vol. 112. American Mathematical Society, Providence (2010), theory, methods and applications, Translated from the 2005 German original by J. Sprekels
7. Wachsmuth, D., Wachsmuth, G.: Convergence and regularization results for optimal control problems with sparsity functional. ESAIM Control Optim. Calc. Var. 17(3), 858–886 (2011)
8. Wachsmuth, D., Wachsmuth, G.: On the regularization of optimization problems with inequality constraints. Control and Cybernetics (2011), to appear