ON THE REGULARIZATION OF OPTIMIZATION PROBLEMS WITH INEQUALITY CONSTRAINTS

DANIEL WACHSMUTH* AND GERD WACHSMUTH†

Abstract. In this article we study the regularization of optimization problems by Tikhonov regularization. The optimization problems are subject to pointwise inequality constraints in $L^2(\Omega)$. We derive a-priori regularization error estimates if the regularization parameter as well as the noise level tend to zero. We rely on an assumption that is a combination of a source condition and of a structural assumption on the active sets. Moreover, we introduce a strategy to choose the regularization parameter in dependence of the noise level. We prove convergence of this parameter choice rule with optimal order.

Key words. source condition, discrepancy principle, non-smooth optimization, convex constraints, sparsity, regularization error estimates

AMS subject classifications. 49K20, 49N45, 65K10

1. Introduction. In this article we consider optimization problems that can be interpreted as optimal control problems or as inverse problems. We study the regularization of the minimization problem:

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\|\mathcal{S}u - z\|_Y^2 \\ \text{such that} \quad & u_a \le u \quad \text{a.e. on } \Omega_a \\ \text{and} \quad & u \quad \le u_b \text{ a.e. on } \Omega_b. \end{aligned}$$

Here, $\Omega \subset \mathbb{R}^n$, $n \geq 1$, is a bounded, measurable set, Y a Hilbert space, $z \in Y$ a given function. The operator $\mathcal{S} : L^2(\Omega) \to Y$ is linear and continuous. The inequality constraints are prescribed on measurable subsets $\Omega_a, \Omega_b \subset \Omega$. The functions u_a, u_b : $\Omega \to \mathbb{R} \cup \{-\infty, +\infty\}$ are given with $u_i \in L^\infty(\Omega_i)$, $i \in \{a, b\}$.
This simple model problem allows for two distinct interpretations. Viewed as an optimal control problem, the unknown u is the control, the inequality constraints are pointwise inequality constraints, the function z is the desired state. From the inverse problem point of view, the unknown u represents for example coefficients that have to be reconstructed from the (possible noisy) measurement z, the inequality constraints reflect a-priori information and restrict the solution space.
Although both interpretations sound very differently, the underlying problem is ill-posed, no matter which point of view one prefers. Ill-posedness may arise due to non-existence of unique solutions: If z is not in the range of $\mathcal{S}$ and inequality constraints are not prescribed on the whole domain Ω, i.e., $\Omega_a \neq \Omega$ or $\Omega_b \neq \Omega$, then a solution is not guaranteed to exist. Uniqueness of solutions can be proven only under additional assumptions, e.g. injectivity of $\mathcal{S}$. If solutions exist, they may be unstable with respect to perturbations, which is critical if only error-prone measurements $z_\delta \approx z$ of the exact data z are available. In addition, every discretization of the original problem introduces perturbations.
In order to overcome these difficulties, regularization methods were developed and investigated during the last decades. We will focus here on Tikhonov regularization

*Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstraße 69, A–4040 Linz, Austria, daniel.wachsmuth@oeaw.ac.at
†Chemnitz University of Technology, Faculty of Mathematics, D-09107 Chemnitz, Germany, gerd.wachsmuth@mathematik.tu-chemnitz.de

with some positive regularization parameter $\alpha > 0$. The regularized problem is given by:

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\|\mathcal{S}u - z_\delta\|_Y^2 + \frac{\alpha}{2}\|u\|_{L^2(\Omega)}^2 \\ \text{such that} \quad & u_a \leq u \quad \text{a.e. on } \Omega_a \\ \text{and} \quad & u \quad \leq u_b \text{ a.e. on } \Omega_b. \end{aligned}$$

Here, z_δ with $\|z - z_\delta\| \leq \delta$ is the perturbed state to the noise level $\delta > 0$. For given $\alpha > 0$ this problem has a unique solution, which is stable with respect to perturbations. The additional Tikhonov regularization term can be interpreted in the context of optimal control as control costs.

Once a regularized problem is solved, one is interested in the convergence for $(\alpha, \delta) \searrow 0$. Additionally, one wants to find conditions that guarantee (or explain) convergence rates with respect to α and δ. These questions are studied in the literature about inverse problems. Convergence results were developed for linear and nonlinear inverse problems, see e.g. [3]. One of the most famous sufficient conditions is the so-called source condition, which assumes that the solution of the original problem is in the range of the dual operator $\mathcal{S}^\star$.

A comprehensive study of inverse problems subject to convex constraints can be found in [3, Section 5.4]. There convergence of the regularization scheme given a source condition is proven. As mentioned in [8], a source condition is unlikely to hold in an optimal control setting if z is not attainable, i.e., there is no feasible u such that $z = \mathcal{S}u$. Then the optimal control u_0 might be bang-bang, i.e. it is a linear combination of characteristic functions, hence u_0 is in general discontinuous with $u_0 \notin H^1(\Omega)$. This contradicts a fulfillment of the source condition as in many examples the range of $\mathcal{S}^\star$ contains $H^1(\Omega)$ or $C(\bar{\Omega})$. In [8] a regularity assumption on the active sets is used as a suitable substitution of the source condition. The active set is the subset of Ω, where the inequality constraints for u_0 are active. Such a condition is also employed in [2, 4]. In [2] this condition was used to prove a-priori error estimates for the discretization error in the controls. In [4] the regularity condition was used to prove stability of bang-bang controls for problems in a non-autonomous ODE setting. However, the regularity assumption implies that the control constraints are active everywhere, i.e., $u_0 \in \{u_a, u_b\}$ a.e. on Ω. In particular, situations are not covered, where the control constraints are inactive on a large part of Ω or if only one-sided constraints are given.

In this paper, we will combine both approaches: we will use a source condition on the part of the domain, where the inequality constraints are inactive, and we will use a structural assumption on the active sets, see Section 3. Then we prove a-priori convergence rates if (α, δ) tends to zero, see Theorem 3.14. These rates allow for an a-priori choice of the regularization parameter α in dependence of δ.

However such an a-priori choice is not possible in practice, as it requires exact knowledge about the unknown solution u_0 in terms of parameters appearing in the structural assumption on the active sets. Here, one is interested to find a rule to determine α without any a-priori information on the unknown solution u_0. In the inverse problem context an important parameter choice rule is the so-called Morozov discrepancy principle [6]. There, α is determined as the parameter that brings the residual in the equation $\mathcal{S}u - z^\delta$ below a certain threshold. In Section 4, we extend this principle to account for the presence of control constraints and the non-attainability of the exact data z. Then we prove that the resulting regularization scheme gives under suitable assumptions the same convergence rate with respect to δ as the best a-priori choice,

see Theorem 4.7.
Simultaneously, we will study the regularization of the following minimization problem: given $\beta \geq 0$,

$$\text{Minimize} \quad \frac{1}{2}\|\mathcal{S}u - z\|_Y^2 + \beta\|u\|_{L^1(\Omega)}$$

subject to the inequality constraints. Here it is worth noting that the presence of the L^1-term does not make the problem well-posed. Indeed, the remarks about existence and stability of solutions above are still valid.
This is in contrast to the analysis of problems in sequence spaces. There one has $l^1 \hookrightarrow l^2$, and the regularization by l^1-norms is even stronger than the one by l^2-norms. Moreover, since $l^1 = (c_0)^*$ it is possible to prove the existence of solutions of optimization problems in l^1 by means of the weak-star topology. This is not the case for $L^1(\Omega)$: this space has no pre-dual, hence the notion of weak-star convergence does not make sense, and optimization problems may not have solutions in $L^1(\Omega)$.
In contrast to the existing literature on inverse problems, we do not assume that $\mathcal{S}u_0 = z$ holds, which corresponds to an optimal functional value of zero of problem (**P**). Instead we develop convergence results for $\mathcal{S}(u_0 - u_{\alpha,\delta})$. These are equivalent to estimates of $\mathcal{S}u_{\alpha,\delta} - z$ in the case $\mathcal{S}u_0 = z$.
The paper is organized as follows. In Section 2 we formulate the problem under consideration and derive some basic properties. Section 3 is devoted to the derivation of error estimates with respect to α and δ. There we use a combination of a (power-type) source condition and a structural assumption on the active set, see Assumption 3.2. Finally, in Section 4, we describe a parameter choice rule to determine the parameter α. We prove convergence rates for this method.

2. Problem setting and preliminary results. Let us recall the optimization problem that we want to regularize:

$$\left.\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\|\mathcal{S}u - z\|_Y^2 + \beta\|u\|_{L^1(\Omega)} \\ \text{such that} \quad & u_a \leq u \quad \text{a.e. on } \Omega_a \\ \text{and} \quad & u \quad \leq u_b \text{ a.e. on } \Omega_b. \end{aligned}\right\} \qquad (\mathbf{P})$$

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., $\mathcal{S}^\star \in \mathcal{L}(Y, L^\infty(\Omega))$. The parameter β is a non-negative number.
The set of feasible functions u is given by

$$U_{ad} := \{u \in L^2(\Omega) : \ u_a \leq u \text{ on } \Omega_a, \ u \leq u_b \text{ on } \Omega_b\}$$

with $u_i \in L^\infty(\Omega_i)$, $i \in \{a, b\}$, and $u_a \leq 0$ and $0 \leq u_b$ a.e. on Ω_a and Ω_b, respectively. For convenience we use sometimes $u_a(x) = -\infty$ if $x \notin \Omega_a$ and $u_b(x) = \infty$ if $x \notin \Omega_b$. In an inverse problem context U_{ad} represents given a-priori informations, whereas from an optimal control point of view, U_{ad} contains the admissible controls. We remark, that the assumption $u_a \leq 0 \leq u_b$ is not a restriction. If, e.g., $u_a > 0$ on a subset $\Omega_1 \subset \Omega$, we can decompose the L^1-norm as $\|u\|_{L^1(\Omega)} = \|u\|_{L^1(\Omega\setminus\Omega_1)} + \int_{\Omega_1} u$. Hence, on Ω_1 the L^1-norm in U_{ad} is in fact a linear functional, and thus the problem can be handled in an analogous way.

We will denote by $P_{U_{ad}}$ the L^2-projection onto the feasible set U_{ad}. This projection acts pointwise on functions $v \in L^2(\Omega)$ and can be written as

$$(P_{U_{ad}}(v))(x) = \min\big(\max\big(v(x),\, u_a(x)\big),\, u_b(x)\big)\,.$$

In the optimization problem ($\mathbf{P}$), the function $z \in Y$ corresponds to an exact measurement, i.e., it is obtained without measurement errors. In many applications only perturbed measurements $z_\delta \in Y$ are available for some error or noise level $\delta > 0$ with $\|z - z_\delta\|_Y \le \delta$.
In order to derive estimates w.r.t. the regularization parameter $\alpha \ge 0$ and noise level $\delta \ge 0$, we define a family of optimization problems

$$\left.\begin{aligned} &\text{Minimize} \quad J_{\alpha,\delta}(u) := \frac{1}{2}\|\mathcal{S}u - z_\delta\|_Y^2 + \frac{\alpha}{2}\,\|u\|_{L^2}^2 + \beta\|u\|_{L^1} \\ &\text{such that} \quad u_a \le u \quad \text{a.e. on } \Omega_a \\ &\qquad\quad \text{and} \quad u \;\; \le u_b \;\text{a.e. on } \Omega_b. \end{aligned}\right\} \qquad (\mathbf{P}_{\alpha,\delta})$$

We will use the conventions $z_0 := z$ and $J_\alpha(u) := J_{\alpha,0}(u)$. In the sequel, we use the following notation for the solutions, states and adjoint states for the problems ($\mathbf{P}_{\alpha,\delta}$) and ($\mathbf{P}_{\alpha,0}$):

$$\begin{aligned} u_\alpha \;\; &:= \operatorname{argmin}_{u\in U_{ad}} J_\alpha(u), & y_\alpha \;\; &:= \mathcal{S}u_\alpha, & p_\alpha \;\; &:= \mathcal{S}^\star(z_0 - y_\alpha),\\ u_{\alpha,\delta} &:= \operatorname{argmin}_{u\in U_{ad}} J_{\alpha,\delta}(u), & y_{\alpha,\delta} &:= \mathcal{S}u_{\alpha,\delta}, & p_{\alpha,\delta} &:= \mathcal{S}^\star(z_\delta - y_{\alpha,\delta}). \end{aligned}$$

In particular, we denote by u_0 a solution of the original problem ($\mathbf{P}$) if it exists. The solution set of ($\mathbf{P}$) is denoted with U_0. We will call the functions y and p states and adjoint states in accordance with the denotation in the optimal control literature.
Throughout the paper, c denotes a generic constant, which may change from line to line, but which does not depend on relevant quantities like α, δ.
REMARK 2.1. *All considerations can be transferred one-to-one to the case that (Ω, Σ, μ) is a given measure space with $\mu(\Omega) < +\infty$. Then one has to use $L^p(\Omega) := L^p(\Omega; \mu)$ with norm $\|u\|_{L^p} := (\int_\Omega |u|^p\, d\mu)^{1/p}$, $1 \le p < \infty$. This would allow to include boundary control problems and identification of initial values in time-dependent problems.*

2.1. Existence and optimality conditions. Let us first recall the results on existence of solutions of minimizers.
THEOREM 2.2. *Let $\alpha, \delta \ge 0$ be given. Assume further that $\alpha > 0$ or $\Omega_a = \Omega_b = \Omega$ holds. Then the problem ($\mathbf{P}_{\alpha,\delta}$) has a minimizer $u_{\alpha,\delta}$.*
If in addition $\alpha > 0$ holds or the operator $\mathcal{S}$ is injective then this solution is uniquely determined.
Please observe, that in the case $\alpha = 0$ one has to assume $\Omega_a = \Omega_b = \Omega$ to ensure existence of solutions of ($\mathbf{P}$) regardless of the value of β. Otherwise, minimizers will not exist in general. To obtain existence of minimizers in this case, one has to use a measure space setting, see [1].
If a solution of the minimization problem exists, then it can be characterized by first-order necessary optimality conditions.
THEOREM 2.3 ([8, Lemma 2.2]). *Let $\alpha, \delta \ge 0$ and let $u_{\alpha,\delta}$ be a solution of $J_{\alpha,\delta}$. Then, there exists a subgradient $\lambda_{\alpha,\delta} \in \partial\|u_{\alpha,\delta}\|_{L^1(\Omega)}$, such that with the adjoint state $p_{\alpha,\delta} = \mathcal{S}^\star(z_\delta - y_{\alpha,\delta})$ the variational inequality*

$$(\alpha\, u_{\alpha,\delta} - p_{\alpha,\delta} + \beta\,\lambda_{\alpha,\delta},\, u - u_{\alpha,\delta}) \ge 0 \quad \forall u \in U_{ad}, \tag{2.1}$$

is satisfied.

Since problem (**P**) is a convex optimization problem, the first order necessary optimality condition is also sufficient for optimality.

Standard arguments (see [7, Section 2.8]) lead to a pointwise a.e. interpretation of the variational inequality:

$$(\alpha\, u_{\alpha,\delta}(x) - p_{\alpha,\delta}(x) + \beta\, \lambda_{\alpha,\delta}(x),\, u - u_{\alpha,\delta}(x)) \geq 0 \quad \forall u \in \mathbb{R}:\; u_a(x) \leq u \leq u_b(x), \tag{2.2}$$

which in turn implies the following relation between $u_{\alpha,\delta}$ and $p_{\alpha,\delta}$ in the case $\alpha > 0$:

$$u_{\alpha,\delta}(x) = \begin{cases} u_a(x) & \text{if } p_{\alpha,\delta}(x) < \alpha\, u_a(x) - \beta \\ \frac{1}{\alpha}(p_{\alpha,\delta}(x) + \beta) & \text{if } \alpha\, u_a(x) - \beta \leq p_{\alpha,\delta}(x) \leq -\beta \\ 0 & \text{if } |p_{\alpha,\delta}(x)| < \beta \\ \frac{1}{\alpha}(p_{\alpha,\delta}(x) - \beta) & \text{if } \beta \leq p_{\alpha,\delta}(x) \leq \alpha\, u_b(x) + \beta \\ u_b(x) & \text{if } \alpha\, u_b(x) + \beta < p_{\alpha,\delta}(x) \end{cases} \quad \text{a.e. on } \Omega. \tag{2.3}$$

In the case $\alpha = 0$, (2.2) is equivalent to

$$u_{0,\delta}(x) \begin{cases} = u_a(x) & \text{if } p_{0,\delta}(x) < -\beta \\ \in [u_a(x), 0] & \text{if } p_{0,\delta}(x) = -\beta \\ = 0 & \text{if } |p_{0,\delta}(x)| < \beta \\ \in [0, u_b(x)] & \text{if } p_{0,\delta}(x) = \beta \\ = u_b(x) & \text{if } \beta < p_{0,\delta}(x) \end{cases} \quad \text{a.e. on } \Omega. \tag{2.4}$$

This implies that $u_0(x)$ is uniquely determined by $p_0(x)$ on the set, where $|p_0(x)| \neq \beta$ holds. Moreover, we conclude the bound $|p_{0,\delta}| \leq \beta$ on the parts of Ω where no inequality constraints are prescribed:

LEMMA 2.4. *Let $u_{0,\delta}$ be a solution of $J_{0,\delta}$ with associated adjoint state $p_{0,\delta}$. Then it holds*

$$\begin{aligned} p_{0,\delta}(x) &\geq -\beta \quad \text{a.e. on } \Omega \setminus \Omega_a, \\ p_{0,\delta}(x) &\leq +\beta \quad \text{a.e. on } \Omega \setminus \Omega_b. \end{aligned}$$

In particular, we have $|p_{0,\delta}(x)| \leq \beta$ a.e. on $\Omega \setminus (\Omega_a \cup \Omega_b)$.

Proof. Take $x \in \Omega \setminus \Omega_a$. The pointwise variational inequality (2.2) imply

$$(-p_{0,\delta}(x) + \beta\, \lambda_{0,\delta}(x),\, u - u_{0,\delta}(x)) \geq 0 \quad \forall u \in \mathbb{R}:\; u \leq u_b(x).$$

Since $u := u_{0,\delta} - 1 \leq u_b - 1 \leq u_b$, this implies $-p_{0,\delta}(x) + \beta\, \lambda_{0,\delta}(x) \leq 0$. Hence $p_{0,\delta}(x) \geq \beta\, \lambda_{0,\delta}(x) \geq -\beta$ almost everywhere on $x \in \Omega \setminus \Omega_a$. On $\Omega \setminus \Omega_b$ the inequality $p_{0,\delta}(x) \leq +\beta$ follows analogously. □

2.2. Structure of the solution set in case $\alpha = 0$. The aim of this section is to derive some basic properties of the solution set for the original problem (**P**). In general it is possible that there exists no solution, a unique solution or multiple solutions. For the remainder of this section we assume that the solution set U_0 is not empty.

Albeit U_0 is not necessarily single-valued, one can prove that the optimal state and the L^1-norm of the optimal control is uniquely determined.

LEMMA 2.5. *The set*

$$\{y \in Y : \exists u_0 \in U_0 \text{ with } y = \mathcal{S}u_0\}$$

is single-valued. If $\beta > 0$ the set

$$\{t \in \mathbb{R} : \exists u_0 \in U_0 \text{ with } t = \|u_0\|_{L^1}\}$$

is single-valued, too.

Proof. Let u_0, $\tilde{u}_0 \in U_0$ be given. Since $u \mapsto \|\mathcal{S}u - z_\delta\|_{L^2}^2$ and $u \mapsto \beta\|u\|_{L^1}$ are convex, both must be linear on the line segment $[u_0, \tilde{u}_0]$. This implies $\mathcal{S}u_0 = \mathcal{S}\tilde{u}_0$ and $\|u_0\|_{L^1} = \|\tilde{u}_0\|_{L^1}$ in case $\beta > 0$. □

As a consequence, there exists a unique solution if $\mathcal{S}$ is injective. However, even if $\mathcal{S}$ is not injective, the solution with minimal $L^2(\Omega)$-norm is unique.

LEMMA 2.6. *There exists a unique solution in U_0 with minimal L^2-norm.*

Proof. It is easy to see that the set U_0 is convex, non-empty, and closed in $L^2(\Omega)$. Due to the strict convexity of the $L^2(\Omega)$-norm, the problem

$$\min_{u \in U_0} \|u\|_{L^2}$$

has a unique solution. □

Later we shall see that if the sequence u_α converges, it converges to the solution of (**P**) with minimal $L^2(\Omega)$-norm.

2.3. Monotonicity and continuity. A direct consequence of the optimality is the monotonicity of the mapping $\alpha \mapsto \|u_\alpha\|_{L^2}$.

LEMMA 2.7. *The mapping $\alpha \mapsto \|u_\alpha\|_{L^2}$ is monotonically decreasing from $(0, +\infty)$ to $\mathbb{R}$. In addition, the mapping $\alpha \mapsto \frac{1}{2}\|y_\alpha - y_d\|_Y^2 + \beta\|u_\alpha\|_{L^1}$ is monotonically increasing from $(0, +\infty)$ to $\mathbb{R}$.*

If the problem (**P**) *has a solution then these mappings are monotone from $[0, +\infty)$ to $\mathbb{R}$, i.e.,*

$$\|u_\alpha\|_{L^2} \leq \|u_0\|_{L^2}$$

holds for all solutions u_0 of (**P**) *and all $\alpha > 0$.*

Proof. Let $\alpha, \alpha' \geq 0$ be given. Using the optimality of $(y_{\alpha'}, u_{\alpha'})$ and (y_α, u_α) for the functionals $J_{\alpha'}$ and J_α, respectively, we have

$$J_\alpha(y_\alpha, u_\alpha) \leq J_\alpha(y_{\alpha'}, u_{\alpha'}) \quad \text{and} \quad -J_{\alpha'}(y_\alpha, u_\alpha) \leq -J_{\alpha'}(y_{\alpha'}, u_{\alpha'}).$$

Adding both inequalities yields

$$(\alpha - \alpha')\|u_\alpha\|_{L^2}^2 \leq (\alpha - \alpha')\|u_{\alpha'}\|_{L^2}^2,$$

which is equivalent to $(\alpha - \alpha')(\|u_\alpha\|_{L^2}^2 - \|u_{\alpha'}\|_{L^2}^2) \leq 0$. Hence, the mapping $\alpha \mapsto \|u_\alpha\|_{L^2}$ is monotonically decreasing from $(0, +\infty)$ to $\mathbb{R}$.

Let us take $\alpha' > \alpha$. Then we have using the monotonicity of $\alpha \mapsto \|u_\alpha\|_{L^2}$

$$\begin{aligned} J_\alpha(y_\alpha, u_\alpha) \leq J_\alpha(y_{\alpha'}, u_{\alpha'}) &= \frac{1}{2}\|y_{\alpha'} - y_d\|_Y^2 + \beta\|u_{\alpha'}\|_{L^1} + \frac{\alpha'}{2}\|u_{\alpha'}\|_{L^2}^2 \\ &\leq \frac{1}{2}\|y_{\alpha'} - y_d\|_Y^2 + \beta\|u_{\alpha'}\|_{L^1} + \frac{\alpha'}{2}\|u_\alpha\|_{L^2}^2. \end{aligned}$$

Hence we have for $\alpha' > \alpha$

$$\frac{1}{2}\|y_\alpha - y_d\|_Y^2 + \beta\|u_\alpha\|_{L^1} \le \frac{1}{2}\|y_{\alpha'} - y_d\|_Y^2 + \beta\|u_{\alpha'}\|_{L^1}.$$

If a solution exists for $\alpha = 0$, then these arguments extend to the case $\alpha = 0$ with u_0 being any solution (**P**). □

For further references, we state the following obvious consequence of the previous result, which is boundedness of the sequence $\{u_\alpha\}_{\alpha>0}$ in $L^2(\Omega)$.

COROLLARY 2.8. *The set of solutions $\{u_\alpha\}_{\alpha>0}$ is bounded in $L^2(\Omega)$ if and only if (**P**) is solvable, i.e., $U_0 \neq \emptyset$.*

Proof. If $U_0 \neq \emptyset$, the assertion follows from Lemma 2.7.

Now, let us assume that $\{u_\alpha\}_{\alpha>0}$ is bounded in $L^2(\Omega)$. Due to the reflexivity of $L^2(\Omega)$, there is a sequence α_n and $u \in L^2(\Omega)$, such that $\alpha_n \searrow 0$ and $u_{\alpha_n} \rightharpoonup u$ in $L^2(\Omega)$ as $n \to \infty$. Since U_{ad} is weakly closed, we obtain $u \in U_{ad}$. Let $\tilde{u} \in U_{ad}$ be arbitrary. We obtain

$$\begin{aligned}
J_0(\mathcal{S}u, u) &\le \liminf_{n\to\infty} J_0(\mathcal{S}u_{\alpha_n}, u_{\alpha_n}) && \text{(since } \mathcal{S} \text{ is weakly continuous)}\\
&\le \liminf_{n\to\infty} J_{\alpha_n}(\mathcal{S}u_{\alpha_n}, u_{\alpha_n}) && \text{(by definition)}\\
&\le \liminf_{n\to\infty} J_{\alpha_n}(\mathcal{S}\tilde{u}, \tilde{u}) && \text{(by optimality of } u_{\alpha_n})\\
&= J_0(\mathcal{S}\tilde{u}, \tilde{u}),
\end{aligned}$$

which implies $u \in U_0$, and in particular $U_0 \neq \emptyset$. □

Before we study the behavior of solutions for $\alpha \to 0$, let us state the following result, which will be one key to prove convergence rates.

PROPOSITION 2.9. *Let $\alpha, \alpha' > 0$ and $\delta, \delta' \ge 0$ be given. Then it holds*

$$\begin{aligned}
&\|y_{\alpha',\delta'} - y_{\alpha,\delta}\|_Y^2 + \alpha\,\|u_{\alpha',\delta'} - u_{\alpha,\delta}\|_{L^2}^2\\
&\qquad\qquad \le (\alpha' - \alpha)\,(u_{\alpha',\delta'},\, u_{\alpha,\delta} - u_{\alpha',\delta'}) + (z_\delta - z_{\delta'},\, y_{\alpha',\delta'} - y_{\alpha,\delta})_Y.
\end{aligned}$$

If $(\mathbf{P}_{\alpha,\delta})$ is solvable for $\alpha = 0$ and noise levels δ, δ', the estimate holds true for $\alpha, \alpha' \ge 0$.

Proof. For $\delta = \delta'$ this result can be found in [8, Lemma 3.1].

We start with the variational inequalities (2.1) for (α, δ) and (α', δ'). Testing with $u_{\alpha',\delta'}$ and $u_{\alpha,\delta}$, respectively, leads to

$$\begin{aligned}
(\alpha\, u_{\alpha,\delta} - p_{\alpha,\delta} + \beta\,\lambda_{\alpha,\delta},\, u_{\alpha',\delta'} - u_{\alpha,\delta}) &\ge 0,\\
(\alpha'\, u_{\alpha',\delta'} - p_{\alpha',\delta'} + \beta\,\lambda_{\alpha',\delta'},\, u_{\alpha,\delta} - u_{\alpha',\delta'}) &\ge 0.
\end{aligned}$$

Adding both inequalities yields

$$\begin{aligned}
&-\alpha\,\|u_{\alpha',\delta'} - u_{\alpha,\delta}\|_{L^2}^2 - (\alpha' - \alpha)\,(u_{\alpha',\delta'},\, u_{\alpha',\delta'} - u_{\alpha,\delta})\\
&\qquad + \beta\,(\lambda_{\alpha,\delta} - \lambda_{\alpha',\delta'},\, u_{\alpha',\delta'} - u_{\alpha,\delta}) - (p_{\alpha,\delta} - p_{\alpha',\delta'},\, u_{\alpha',\delta'} - u_{\alpha,\delta}) \ge 0.
\end{aligned}$$

Due to the monotonicity of the subdifferential, we have $(\lambda_{\alpha',\delta'} - \lambda_{\alpha,\delta},\, u_{\alpha',\delta'} - u_{\alpha,\delta}) \ge 0$. Inserting the definition of the adjoint state, we get immediately

$$(p_{\alpha,\delta} - p_{\alpha',\delta'},\, u_{\alpha',\delta'} - u_{\alpha,\delta}) = (z_\delta - z_{\delta'}, y_{\alpha',\delta'} - y_{\alpha,\delta}) + \|y_{\alpha',\delta'} - y_{\alpha,\delta}\|_Y^2,$$

which implies

$$- (\alpha' - \alpha)\,(u_{\alpha',\delta'},\, u_{\alpha',\delta'} - u_{\alpha,\delta}) - (z_\delta - z_{\delta'}, y_{\alpha',\delta'} - y_{\alpha,\delta}) \\ \geq \alpha\, \|u_{\alpha',\delta'} - u_{\alpha,\delta}\|^2_{L^2} + \|y_{\alpha',\delta'} - y_{\alpha,\delta}\|^2_Y.$$

□

A first consequence of this result is the local Lipschitz continuity of the map $\alpha \mapsto u_{\alpha,\delta}$ from $(0, +\infty)$ to $L^2(\Omega)$ for fixed δ.

COROLLARY 2.10. *Let us fix $\delta \geq 0$. Then the mapping $\alpha \mapsto u_{\alpha,\delta}$ is locally Lipschitz continuous from $(0, +\infty)$ to $L^2(\Omega)$.*

Setting $\alpha' = \delta = \delta' = 0$ in Proposition 2.9, yields

LEMMA 2.11. *Let $\alpha \geq 0$ be given. Let u_0 be a solution of $(\mathbf{P})$. Then it holds*

$$\|y_0 - y_\alpha\|^2_Y + \alpha\, \|u_0 - u_\alpha\|^2_{L^2} \leq \alpha\, (u_0,\, u_0 - u_\alpha). \tag{2.5}$$

3. Error estimates. In this section, we will develop the a-priori convergence result. Let some measurement z_δ be given with measurement error $\delta > 0$, $\|z - z_\delta\|_Y \leq \delta$, where z corresponds to exact measurement. Here, z is unknown to us, only z_δ is accessible. In order to approximate a solution u_0 of $(\mathbf{P})$, we solve the regularized problem $(\mathbf{P}_{\alpha,\delta})$ for $(\alpha, \delta) \to 0$.

Throughout this section we assume that a solution u_0 of $(\mathbf{P})$ exists. We will estimate the error as

$$\|u_{\alpha,\delta} - u_0\|_{L^2} \leq \|u_{\alpha,\delta} - u_\alpha\|_{L^2} + \|u_\alpha - u_0\|_{L^2},$$

i.e., we will separate the noise error and the regularization error.

3.1. Estimate of the noise error. At first, let us estimate the error $\|u_\alpha - u_{\alpha,\delta}\|_{L^2}$, arising due to the measurement error or noise level δ.

THEOREM 3.1. *Let $\alpha > 0$ be given. Then for the solution with noise level $\delta > 0$ we obtain the estimates*

$$\|u_\alpha - u_{\alpha,\delta}\|_{L^2} \leq \frac{\delta}{2\sqrt{\alpha}}, \quad \textit{and} \quad \|y_\alpha - y_{\alpha,\delta}\|_Y \leq \delta.$$

Proof. Proposition 2.9 with $\delta' = 0$ and $\alpha = \alpha'$ gives

$$\|y_\alpha - y_{\alpha,\delta}\|^2_Y + \alpha\, \|u_\alpha - u_{\alpha,\delta}\|^2_{L^2} \leq (z_\delta - z,\, y_\alpha - y_{\alpha,\delta})_Y \\ \leq \delta\, \|y_\alpha - y_{\alpha,\delta}\|_Y$$

The assertion of the theorem follows immediately by Young's inequality. □

3.2. Regularity assumption. Let us now state our regularity assumption, which will us allow later to prove convergence rates for $\alpha \to 0$.

ASSUMPTION 3.2. *Let u_0 be a solution of $(\mathbf{P})$. Let us assume that there exist a 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*

$$\chi_I\, u_0 = \chi_I\, P_{U_{ad}}(\mathcal{S}^\star w)\,.$$

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

$$\text{meas}\left(\{x \in A: \ 0 < \big||p_0(x)| - \beta\big| < \epsilon\}\right) \le c\,\epsilon^\kappa, \qquad (3.1)$$

with the convention that $\kappa := +\infty$ *if the left-hand side of* (3.1) *is* 0 *for some* $\epsilon > 0$.

The set I contains the set $\{x \in \Omega: \ |p_0(x)| = \beta\}$, which is the set of points, where $u_0(x)$ cannot be uniquely determined from $p_0(x)$, compare (2.4). On this set, we assume that u_0 fulfills a local source condition. The set A contains the points, where the inequality constraints are active, since it holds by construction that $|p_0(x)| \neq \beta$ on A, which implies $u_0(x) \in \{u_a(x), 0, u_b(x)\}$, cf. (2.4) and Lemma 3.4 below.

Let us comment on the relation of Assumption 3.2 to other conditions in the literature. The classical source condition for linear inverse problems reads $u_0 = \mathcal{S}^\star w$. In [3], this was slightly adapted to inequality constrained problems. There the condition $u_0 = P_{U_{ad}}(\mathcal{S}^\star w)$ was employed. For the choice $I = \Omega$, $A = \emptyset$, this condition is a special case of Assumption 3.2, see also Corollary 3.11 below. The authors used in [8] an approach different to source conditions. The condition investigated there corresponds to the case $I = \emptyset$, $A = \Omega$, $\kappa = 1$ of Assumption 3.2. This condition was also employed in [2, 4] in a different context.

REMARK 3.3. *We will show in Theorem 3.14, that if some* $u_0 \in U_0$ *(together with* $p_0 = \mathcal{S}^\star(z_\delta - \mathcal{S}u_0)$*) fulfills Assumption 3.2, the sequence of regularized solutions* u_α *will converge towards* u_0*. This implies, that at most one* $u_0 \in U_0$ *can satisfy Assumption 3.2. In view of Lemmata 2.6 and 2.7 this has to be the solution with the minimal* L^2*-norm in* U_0.

Under Assumption 3.2, we will derive a boundedness result for u_0 on the active set A.

LEMMA 3.4. *Let Assumption 3.2 be satisfied. Then it holds* $|p_0(x)| \neq \beta$ *on* A *and*

$$\begin{aligned} u_a(x) > -\infty \ \text{a.e. on } \{x \in \Omega: \ p_0(x) < -\beta\} \\ u_b(x) < +\infty \ \text{a.e. on } \{x \in \Omega: \ p_0(x) > \beta\} \end{aligned}$$

Moreover, for almost all $x \in A$ *we have* $u_0(x) \in \{u_a(x), 0, u_b(x)\}$*, hence* $u_0|_A \in L^\infty(A)$.

Proof. By definition of A, we have that $|p_0(x)| \neq \beta$. Hence, the characterization of u_0 in (2.4) gives $u_0(x) \in \{u_a(x), 0, u_b(x)\}$. Moreover, this implies that u_a is finite on the set $\{x: \ p_0(x) < -\beta\}$, i.e., u_a has a representative that is bounded from below on this set. The same argumentation applies to the set $\{x: \ p_0(x) > \beta\}$, which proves $u_0|_A \in L^\infty(A)$. □

REMARK 3.5. *Following [2, Lemma 3.2], one can prove that for* $p_0 \in C^1(\bar\Omega)$ *satisfying*

$$\nabla p_0(x) \neq 0 \ \text{for all } x \in \bar\Omega \ \text{with } |p_0(x)| = \beta$$

Assumption 3.2 is fulfilled with $A = \Omega$ *and* $\kappa = 1$*. Under these conditions, a-priori discretization error estimates for the variational discretization of a distributed optimal control for an elliptic equation were proven in [2]. This is remarkable since classical error estimates contain negative powers of* α *and thus are not applicable to the case* $\alpha = 0$.

REMARK 3.6. *Let us consider an one-dimensional example with* $\Omega = (0, 1)$*. First, let the function* p_0 *be given by* $p_0(x) = \beta + x^s$*, with some power* $s > 0$*. Then the measure of the set* $\{|p_0| = \beta\}$ *is zero, and Assumption 3.2 is fulfilled with* $A = \Omega$ *and* $\kappa = 1/s$.

Second, let the function p_0 be defined as

$$p_0(x) = \beta + \begin{cases} 0 & \text{if } x \le 1/2 \\ (x-1/2)^s & \text{if } x > 1/2 \end{cases}$$

with $s > 0$. If s is integer then p_0 belongs to $C^s(\bar{\Omega})$. In order that Assumption 3.2 can be fulfilled, the sets I and A must be chosen such that $I \supset (0, 1/2)$ and $A \subset (1/2, 1)$. If $A = (1/2, 1)$ is chosen then it follows that $\kappa = 1/s$.

REMARK 3.7. *Let us comment on situations, where the special case $\kappa = +\infty$ in Assumption 3.2.2 occurs. If Ω is connected, p_0 is not continuous and bounded away from zero then Assumption 3.2 is fulfilled with $\kappa = +\infty$, e.g. $\Omega = (-1, 1)$, $p_0(x) = \operatorname{sign}(x)$. If p_0 is a continuous function but Ω is not connected, then one can construct an analogous example allowing to set $\kappa = +\infty$.*

3.3. Estimate of the regularization error. Now let us start with the convergence analysis of the regularization error $u_\alpha - u_0$. Invoking the source condition part of Assumption 3.2, we have

LEMMA 3.8. *Let Assumption 3.2.1 (source condition) be satisfied. Then there is a constant $c > 0$ independent of α such that*

$$\|y_0 - y_\alpha\|_Y^2 + \alpha \|u_0 - u_\alpha\|_{L^2}^2 + \|p_0 - p_\alpha\|_{L^\infty}^2 \le c\,\alpha \{\alpha + \operatorname{meas}(A_\alpha)\},$$

where $\operatorname{meas}(A_\alpha)$ is the Lebesgue-measure of the set $A_\alpha \subset A$, where A_α is given by

$$A_\alpha = \{x \in A : \; u_\alpha(x) \ne u_0(x)\}. \tag{3.2}$$

Proof. Due to the properties of the projection, we have from the source condition in Assumption 3.2

$$(\chi_I(u_0 - \mathcal{S}^\star w),\, u - u_0) \ge 0 \quad \forall u \in U_{ad},$$

which implies

$$(\chi_I u_0,\, u_0 - u_\alpha) \le (\chi_I S^* w,\, u_0 - u_\alpha).$$

Using this in inequality (2.5), we can estimate

$$\begin{aligned} \|y_0 - y_\alpha\|_Y^2 + \alpha \|u_0 - u_\alpha\|_{L^2}^2 &\le \alpha\,(u_0, u_0 - u_\alpha) \\ &\le \alpha \{(\chi_I \mathcal{S}^\star w, u_0 - u_\alpha) + (\chi_A u_0, u_0 - u_\alpha)\}. \end{aligned}$$

Writing $\mathcal{S}\chi_I(u_0 - u_\alpha) = \mathcal{S}((1-\chi_A)(u_0 - u_\alpha)) = y_0 - y_\alpha - \mathcal{S}\chi_A(u_0 - u_\alpha)$ we find

$$\begin{aligned} \|y_0 - y_\alpha\|_Y^2 + \alpha \|u_0 - u_\alpha\|_{L^2}^2 &\le \alpha \{(\mathcal{S}^\star w, \chi_I(u_0 - u_\alpha)) + (\chi_A u_0, u_0 - u_\alpha)\} \\ &= \alpha \{(w, y_0 - y_\alpha) - (\mathcal{S}^\star w, \chi_A(u_0 - u_\alpha)) + (u_0, \chi_A(u_0 - u_\alpha))\}. \end{aligned}$$

On the set A, we have $|p_0| \ne \beta$, which by Lemma 3.4 implies that $u_0 \in L^\infty(A)$. This enables us to estimate

$$\|y_0 - y_\alpha\|_Y^2 + \alpha \|u_0 - u_\alpha\|_{L^2}^2 \le c\,\alpha \left\{\|y_0 - y_\alpha\|_Y + \|u_0 - u_\alpha\|_{L^1(A)}\right\}$$

with a constant $c > 0$ depending on $\|w\|_Y$, $\|\mathcal{S}^\star w\|_{L^\infty(A)}$, and $\|u_0\|_{L^\infty(A)}$ but independent of α. Since $\text{meas}(A_\alpha)$ is finite, we have

$$\|u_0 - u_\alpha\|_{L^1(A)} \leq \text{meas}(A_\alpha)^{1/2} \|u_0 - u_\alpha\|_{L^2}, \tag{3.3}$$

which gives

$$\|y_0 - y_\alpha\|_Y^2 + \alpha \, \|u_0 - u_\alpha\|_{L^2}^2 \leq c \, \alpha \left\{ \|y_0 - y_\alpha\|_Y + \text{meas}(A_\alpha)^{1/2} \|u_0 - u_\alpha\|_{L^2} \right\}.$$

Applying Young's inequality and continuity of $\mathcal{S}^\star : Y \mapsto L^\infty(\Omega)$ finishes the proof for $I \neq \emptyset$. ☐

If the set I, on which the source condition should hold, is empty, we can strengthen the result of the Lemma. This will allow us later to prove improved convergence rates in this case.

Corollary 3.9. *Let Assumption 3.2 be satisfied with $A = \Omega$. Then there is a constant $c > 0$ independent of α, such that*

$$\|y_0 - y_\alpha\|_Y^2 + \alpha \, \|u_0 - u_\alpha\|_{L^2}^2 + \|p_0 - p_\alpha\|_{L^\infty}^2 \leq c \, \alpha \; \text{meas}(A_\alpha)$$

with A_α as in (3.2).

Proof. Using (2.5), the boundedness of $\|u_0\|_{L^\infty}$ and Young's inequality, we obtain

$$\begin{aligned} \|y_0 - y_\alpha\|_Y^2 + \alpha \, \|u_0 - u_\alpha\|_{L^2}^2 &\leq \alpha \, (u_0, \, u_0 - u_\alpha) \\ &\leq c \, \alpha \; \text{meas}(A_\alpha)^{1/2} \|u_0 - u_\alpha\|_{L^2} \\ &\leq c \, \alpha \; \text{meas}(A_\alpha) + \frac{\alpha}{2} \|u_0 - u_\alpha\|_{L^2}^2. \end{aligned}$$

☐

With these results we can prove a first basic estimate of the regularization error in states and adjoints.

Corollary 3.10. *Let Assumption 3.2.1 (source condition) be satisfied. Then there is a constant $c > 0$ independent of α, such that*

$$\|y_0 - y_\alpha\|_Y + \|p_0 - p_\alpha\|_{L^\infty} \leq c \, \alpha^{1/2}.$$

Proof. The claim follows directly from Lemma 3.8, since by Corollary 2.8, the $L^1(\Omega)$-norms of u_α are uniformly bounded with respect to $\alpha \geq 0$. ☐

If we assume a source condition on the whole domain Ω, i.e., $I = \Omega$, as for example in [3], one can prove rates as a consequence of Lemma 3.8, since $A_\alpha \subset A = \emptyset$.

Corollary 3.11. *Let Assumption 3.2 with $I = \Omega$ be satisfied, that is, we assume that there is $w \in Y$, such that $u_0 = P_{U_{ad}}(\mathcal{S}^\star w)$. Then*

$$\begin{aligned} \|y_0 - y_\alpha\|_Y &\leq \alpha \, \|w\|_Y, \\ \|u_0 - u_\alpha\|_{L^2} &\leq \frac{\sqrt{\alpha}}{2} \, \|w\|_Y \end{aligned}$$

for all $\alpha \geq 0$.

Now we will make use of the structural assumption on the active set in Assumption 3.2.

Lemma 3.12. *Let Assumption 3.2.2 (structure of the active set) be satisfied with $\kappa < \infty$ and $A = \Omega \setminus I$ having positive measure.*

Then there exists a constant $c > 0$, such that

$$\text{meas}(A_\alpha) \le c\,(\alpha^\kappa + \|p_0 - p_\alpha\|_{L^\infty}^\kappa)$$

for all $\alpha > 0$.

Proof. Let us divide A in disjoint sets depending on the values of p_0 and p_α, see also Table 3.1 below,

$$\begin{aligned} A_1 &:= \{x \in A : \beta \text{ or } -\beta \text{ lies between } p_0 \text{ and } p_\alpha\} \\ A_2 &:= \{x \in A : p_0, p_\alpha \le -\beta \text{ and } p_\alpha \ge -\beta + \alpha\, u_a\} \\ A_3 &:= \{x \in A : p_0, p_\alpha \ge +\beta \text{ and } p_\alpha \le \beta + \alpha\, u_b\}. \end{aligned} \tag{3.4}$$

	$p_0 < -\beta$	$\lvert p_0\rvert < \beta$	$p_0 > \beta$
$p_\alpha \le -\beta + \alpha\, u_a$	$u_0 = u_\alpha = u_a$	A_1	A_1
$p_\alpha \in (-\beta + \alpha\, u_a, -\beta]$	A_2	A_1	A_1
$\lvert p_\alpha\rvert < \beta$	A_1	$u_0 = u_\alpha = 0$	A_1
$p_\alpha \in [\beta, \beta + \alpha\, u_b)$	A_1	A_1	A_3
$p_\alpha \ge \beta + \alpha\, u_b$	A_1	A_1	$u_0 = u_\alpha = u_b$

TABLE 3.1
Partition of A, used in Proof of Lemma 3.12

Let us recall the definition of $A_\alpha = \{x \in \Omega :\ u_\alpha(x) \ne u_0(x)\}$ as given in (2.5). Then it follows that it holds $A_\alpha = A_1 \cup A_2 \cup A_3$: In fact, on $A \setminus (A_1 \cup A_2 \cup A_3)$ we have $u_0 = u_\alpha$ due to the necessary optimality condition Theorem 2.3, confer Table 3.1.

Let us now derive bounds of the measures of the sets A_1, A_2 and A_3. Here, we will develop upper bounds of $\big||p_0| - \beta\big|$ to apply Assumption 3.2. On A_1 we find $\big||p_0| - \beta\big| \le |p_0 - p_\alpha|$.

On A_2 we have that $p_0 < -\beta$, and hence by Lemma 3.4 we obtain $u_a > -\infty$, i.e., $A_2 \subset \Omega_a$. Additionally, it holds $\alpha\, u_a < p_\alpha + \beta \le 0$ on A_2. Hence, we can estimate on A_2

$$\big||p_0| - \beta\big| = |p_0 + \beta| \le |p_0 - p_\alpha| + |p_\alpha + \beta| \le |p_0 - p_\alpha| + \alpha |u_a|.$$

Analogously we get that $\big||p_0| - \beta\big| \le |p_0 - p_\alpha| + \alpha\, u_b$ holds on A_3. Consequently, it holds

$$0 < \big||p_0| - \beta\big| \le \max(\|u_a\|_{L^\infty(\Omega_a)}, \|u_b\|_{L^\infty(\Omega_b)})\,\alpha + |p_0 - p_\alpha| \le c\,\alpha + \|p_0 - p_\alpha\|_{L^\infty}$$

a.e. on A_α. Applying Assumption 3.2 we can bound the measure of A_α and obtain $\text{meas}(A_\alpha) \le c\,(\alpha + \|p_0 - p_\alpha\|_{L^\infty})^\kappa$. □

Let us prove the corresponding result for the special case $\kappa = +\infty$.

COROLLARY 3.13. *Let Assumption 3.2 be satisfied with $\kappa = +\infty$ and $A = \Omega \setminus I$ having positive measure. Then there exists a number α_∞, such that*

$$\text{meas}(A_\alpha) = 0$$

for all $\alpha < \alpha_\infty$.

Proof. As in the proof of the previous Lemma we obtain

$$\text{meas}(A_\alpha) \le c\,(\alpha + \|p_0 - p_\alpha\|_{L^\infty})^{\kappa'}$$

for all $\kappa' > 0$. By Corollary 3.10, there exists α_∞, such that the term $\alpha + \|p_0 - p_\alpha\|_{L^\infty}$ is smaller than one for all $\alpha \in (0, \alpha_\infty)$. Since κ' can be chosen arbitrarily large, this proves that $\text{meas}(A_\alpha) = 0$ for all $\alpha \in (0, \alpha_\infty)$. ☐

With these results we can prove our convergence result.

THEOREM 3.14. *Let Assumption 3.2 be satisfied.*
Let d be defined as

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

Then for every $\alpha_{max} > 0$ there exists a constant $c > 0$, such that

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

holds for all $\alpha \in (0, \alpha_{max}]$. If $\kappa < \infty$ then we have in addition

$$\|u_0 - u_\alpha\|_{L^1(A)} \le c\,\alpha^{d-1/2+\kappa\, d/2}$$

for all $\alpha \in (0, \alpha_{max}]$.
If $\kappa = \infty$ there is a constant $\alpha_\infty > 0$, such that

$$u_0 = u_\alpha \text{ a.e. on } A$$

holds for all $\alpha \in (0, \alpha_\infty)$.

Proof. The case $I = \Omega$, $A = \emptyset$ with the convention $\kappa = +\infty$ is proven in Corollary 3.11, which yields the claimed estimates for $d = 1$.

Let us assume now that A has positive measure. By Lemma 3.8, we have

$$\|y_0 - y_\alpha\|_Y^2 + \alpha\,\|u_0 - u_\alpha\|_{L^2}^2 + \|p_0 - p_\alpha\|_{L^\infty}^2 \le c\,\alpha\,\{\alpha + \text{meas}(A_\alpha)\}.$$

If on one hand $\kappa = \infty$, the claim follows from Corollary 3.13 and (3.3). If on the other hand κ is finite, then according to Lemma 3.12, we can bound the measure of A_α and obtain

$$\|y_0 - y_\alpha\|_Y^2 + \alpha\,\|u_0 - u_\alpha\|_{L^2}^2 + \|p_0 - p_\alpha\|_{L^\infty}^2 \le c\,\alpha\,\{\alpha + \alpha^\kappa + \|p_0 - p_\alpha\|_{L^\infty}^\kappa\}.$$

Let us consider the case $\kappa < 2$. Then by Young's inequality

$$\|y_0 - y_\alpha\|_Y^2 + \alpha\,\|u_0 - u_\alpha\|_{L^2}^2 + \|p_0 - p_\alpha\|_{L^\infty}^2 \le c\,\left\{\alpha^2 + \alpha^{\kappa+1} + \alpha^{\frac{2}{2-\kappa}}\right\}.$$

Since

$$\min\{2,\, 1+\kappa,\, 2/(2-\kappa)\} = \begin{cases} 2/(2-\kappa) & \text{if } 0 \le \kappa \le 1 \\ 2 & \text{if } 1 \le \kappa \le 2. \end{cases}$$

this proves the claim for $\kappa < 2$.

If $\kappa \ge 2$, then we find using Corollary 3.10

$$\|y_0 - y_\alpha\|_Y^2 + \alpha\,\|u_0 - u_\alpha\|_{L^2}^2 + \|p_0 - p_\alpha\|_{L^\infty}^2 \le c\,\alpha\left\{\alpha + \alpha^\kappa + \alpha^{\kappa/2}\right\},$$

which proves the claimed convergence rates of $\|y_0 - y_\alpha\|_Y$, $\|u_0 - u_\alpha\|_{L^2}$, and $\|p_0 - p_\alpha\|_{L^\infty}$. The convergence result for $\|u_0-u_\alpha\|_{L^1(A)}$ follows now from (3.3), Lemma 3.12, and Corollary 3.13.
If $A = \Omega$ and $\kappa \geq 1$ hold, we have from Corollary 3.9 and Lemma 3.12

$$\begin{aligned}\|y_0 - y_\alpha\|_Y^2 + \alpha\,\|u_0 - u_\alpha\|_{L^2}^2 + \|p_0 - p_\alpha\|_{L^\infty}^2 &\leq c\,\alpha\ \mathrm{meas}(A_\alpha)\\ &\leq c\,\alpha\,(\alpha^\kappa + \|p_0 - p_\alpha\|_{L^\infty}^\kappa).\end{aligned}$$

We already proved $\|p_0 - p_\alpha\|_{L^\infty} \leq c\,\alpha$. Using this in the above inequality gives the claim with $d = \frac{\kappa+1}{2}$. □

Example 3.15. *Let us discuss Assumption 3.2 and the resulting convergence rates for the simple settings discussed in Remark 3.6 for $\Omega = (0, 1)$.*
If p_0 is given by $p_0(x) = x^s$, $s > 0$, then Assumption 3.2 is fulfilled with $A = \Omega$ and $\kappa = 1/s$. Then with σ given by

$$\sigma = \begin{cases} \frac{1}{2s} & \text{if } s < 1.\\ \frac{1}{2(2s-1)} & \text{if } s \geq 1,\end{cases}$$

Theorem 3.14 yields the convergence rate $\|u_0 - u_\alpha\|_{L^2} \leq c\,\alpha^\sigma$. Here, we see that with increasing smoothness of p_0 the convergence rate tends to zero.
For the second example, we take the function p_0 be defined as

$$p_0(x) = \beta + \begin{cases} 0 & \text{if } x \leq 1/2\\ (x-1/2)^s & \text{if } x > 1/2\end{cases}$$

with $s > 0$. Let us set $I = (0, 1/2)$, $A = [1/2, 1)$. Suppose that the source condition Assumption 3.2.1 is fulfilled. Then we can expect the convergence rates

$$\|u_0 - u_\alpha\|_{L^2} \leq c\,\alpha^{\min\left(\frac{1}{2},\,\frac{1}{2(2s-1)}\right)}.$$

Again, the convergence rate degenerates with increasing smoothness of p_0.

3.4. Power-type source condition. The aim of this section is to choose a weaker source condition. We replace Assumption 3.2.1 with
Assumption 3.16. *Let u_0 be a solution of* (**P**). *Let us assume that there exists a function $\tilde{w} \in L^2(\Omega)$, and a positive constant $\nu \in (0, 1)$, such that*

$$\chi_I\,u_0 = \chi_I\,P_{U_{ad}}\left((\mathcal{S}^\star\mathcal{S})^{\nu/2}\tilde{w}\right)$$

holds.
Let us prove an analogous result to Theorem 3.14. We will outline the main steps. First, we use [5, Theorem 1] to turn the power-type source condition into an approximate source condition. This implies the existence of $K > 0$, such that for all $R > 0$ there exists $w \in Y$, $v \in L^2(\Omega)$ with

$$\|w\|_Y \leq R, \qquad \|v\|_{L^2} \leq K\,R^{\frac{\nu}{\nu-1}}, \qquad \chi_I\,u_0 = \chi_I\,P_{U_{ad}}\left(\mathcal{S}^\star w + v\right).$$

Proceeding as in Lemma 3.8, we obtain

$$\|y_0-y_\alpha\|_Y^2+\alpha\,\|u_0-u_\alpha\|_{L^2}^2+\|p_0-p_\alpha\|_{L^\infty}^2 \leq c\left(\alpha^2 R^2+\alpha\,R^{\frac{2\nu}{\nu-1}}+\alpha\,(R+c)^2\ \mathrm{meas}(A_\alpha)\right)$$

for all $R > 0$. By Lemma 3.12, we conclude

$$\|p_0 - p_\alpha\|_{L^\infty}^2 \le c\left(\alpha^2 R^2 + \alpha R^{\frac{2\nu}{\nu-1}} + \alpha (R+c)^2 (\alpha^\kappa + \|p_0 - p_\alpha\|_{L^\infty}^\kappa)\right).$$

Now we use the approach $R = \alpha^\gamma$, with $\gamma < 0$. This yields the rate

$$\|p_0 - p_\alpha\|_{L^\infty}^2 \le c\left(\alpha^{2+2\gamma} + \alpha^{1+\gamma\frac{2\nu}{\nu-1}} + \alpha^{1+2\gamma} \|p_0 - p_\alpha\|_{L^\infty}^\kappa\right). \tag{3.5}$$

In case $\kappa \ge 2$, (3.5) with $\gamma = \frac{\nu-1}{2}$ and Corollary 3.10 implies

$$\|p_0 - p_\alpha\|_{L^\infty} \le c\,\alpha^{\frac{1+\nu}{2}}.$$

In case $\kappa < 2$ we use Young's inequality in (3.5) and obtain

$$\|p_0 - p_\alpha\|_{L^\infty}^2 \le c\left(\alpha^{2+2\gamma} + \alpha^{1+\gamma\frac{2\nu}{\nu-1}} + \alpha^{\frac{2(1+2\gamma)}{2-\kappa}}\right).$$

In case $\kappa(1+\nu) \le 2$, using $\gamma = \frac{\kappa(\nu-1)}{2(2-\nu\kappa)}$, this yields

$$\|p_0 - p_\alpha\|_{L^\infty} \le c\,\alpha^{\frac{1}{2-\nu\kappa}},$$

whereas in case $\kappa(1+\nu) > 2$, using $\gamma = \frac{\nu-1}{2}$, this yields

$$\|p_0 - p_\alpha\|_{L^\infty} \le c\,\alpha^{\frac{1+\nu}{2}}.$$

THEOREM 3.17. *Let Assumptions 3.2.2 and 3.16 be satisfied. Let d be defined as*

$$d = \begin{cases} \frac{1}{2-\kappa\nu} & \text{if } \kappa(1+\nu) \le 2 \\ (1+\nu)/2 & \text{if } \kappa(1+\nu) > 2. \end{cases}$$

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

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

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

We briefly comment on the case $I = \Omega$ (in particular $\kappa = +\infty$ in Assumption 3.2.2) and $\nu < 1$, i.e., a power-type source condition on Ω is satisfied. According to the arguments given above, our technique resembles the standard rate $\|u_0 - u_\alpha\|_{L^2} \le c\,\alpha^{\nu/2}$, which is known from the literature for linear-quadratic objectives.

3.5. A-priori parameter choice. We will now combine the error estimates with respect to noise level and regularization. This will give an a-priori choice $\alpha = \alpha(\delta)$ with best possible convergence order.

THEOREM 3.18. *Let Assumption 3.2 be satisfied.*
Let us choose $\alpha := \alpha(\delta) = \delta^{1/d}$ with d as in Theorem 3.14. Then for every $\delta_{max} > 0$ there is a positive constant $c = c(\delta_{max})$ independent of δ, such that

$$\|y_\alpha^\delta - y_0\|_Y \le c\,\delta, \quad \text{and} \quad \|u_\alpha^\delta - u_0\|_{L^2} \le c\,\delta^s$$

holds for all $\delta \in (0, \delta_{max})$ with s defined by

$$s = 1 - \frac{1}{2\,d} = \begin{cases} \frac{\kappa}{2} & \text{if } \kappa \le 1, \\ \frac{1}{2} & \text{if } \kappa > 1 \text{ and } A \neq \Omega, \\ \frac{\kappa}{\kappa+1} & \text{if } \kappa > 1 \text{ and } A = \Omega. \end{cases}$$

A similar result can be derived for the power-type source condition Assumption 3.16 as employed in the previous section.
Let us remark that such an a-priori choice of α is barely possible in practice, as the constants κ and ν appearing in Assumptions 3.2 and 3.16 are not known a-priori, as they depend heavily on the unknown solution of the unregularized problem and the possibly unaccessible noise-less data. Nevertheless, such an a-priori convergence result can be used as benchmark to compare the convergence order of a-posteriori parameter choices.

3.6. Additional results for the special case $A = \Omega$. If Assumption 3.2 holds with $A = \Omega$, one can obtain additional results regarding stability of solutions. At first, we show a superlinear growth rate of the functional J_0 with respect to the $L^1(\Omega)$-norm. A related result can be found in [4, Theorem 3.4] corresponding to the case $\kappa = 1$.
THEOREM 3.19. *Let us suppose that Assumption 3.2 is fulfilled with $A = \Omega$. Then there exists a constant $c > 0$ such that for all $u \in U_{ad} \cap L^\infty(\Omega)$ with $y := \mathcal{S}u$ it holds*

$$J_0(y,u) - J_0(y_0,u_0) \ge \frac{1}{2}\|y - y_0\|_Y^2 + c\,\frac{\|u-u_0\|_{L^1}^{1+\frac{1}{\kappa}}}{\|u-u_0\|_{L^\infty}^{\frac{1}{\kappa}}}.$$

Proof. Using the relation $p_0 = \mathcal{S}^\star(z - y_0)$ we can rewrite

$$J_0(y,u) - J_0(y_0,u_0) = \frac{1}{2}\|y-y_0\|_Y^2 + (-p_0, u-u_0) + \beta\,(\|u\|_{L^1} - \|u_0\|_{L^1}). \tag{3.6}$$

Let us define the set B by

$$B := \{x \in \Omega :\ u_0(x) = 0\}.$$

Let $\lambda_0 \in \partial\|u_0\|_{L^1}$ be given by the optimality condition (2.1). Then it holds

$$\beta\,(\|u\|_{L^1} - \|u_0\|_{L^1}) \ge \int_{\Omega\setminus B} \beta\,\lambda_0\,(u-u_0) + \beta\,\|u-u_0\|_{L^1(B)}$$

since $u_0 = 0$ on B.
Let $\epsilon > 0$ be given. Let us define $B_\epsilon := \{x \in \Omega \setminus B :\ |p_0(x)| \ge \beta + \epsilon\}$. Then it holds

$$\begin{aligned} \int_{\Omega\setminus B} (\beta\,\lambda_0 - p_0)\,(u-u_0) &= \int_{B_\epsilon} (\beta\,\lambda_0 - p_0)\,(u-u_0) + \int_{\Omega\setminus(B\cup B_\epsilon)} (\beta\,\lambda_0 - p_0)\,(u-u_0) \\ &\ge \epsilon\,\|u-u_0\|_{L^1(B_\epsilon)} - \epsilon\,\|u-u_0\|_{L^1(\Omega\setminus(B\cup B_\epsilon))}, \end{aligned}$$

where we used that $\lambda_0 = \operatorname{sign}(u_0)$ on $\Omega \setminus B$, and $(\beta\,\lambda_0 - p_0)\,(u-u_0) \ge 0$ by (2.1). Using Assumption 3.2 to estimate the measure of the set $\Omega\setminus(B\cup B_\epsilon)$ we proceed with

$$\begin{aligned} &\epsilon\,\|u-u_0\|_{L^1(B_\epsilon)} - \epsilon\,\|u-u_0\|_{L^1(\Omega\setminus(B\cup B_\epsilon)} \\ &\quad\ge \epsilon\,\|u-u_0\|_{L^1(\Omega\setminus B)} - 2\,\epsilon\,\|u-u_0\|_{L^1(\Omega\setminus(B\cup B_\epsilon))} \\ &\quad\ge \epsilon\,\|u-u_0\|_{L^1(\Omega\setminus B)} - 2\,\epsilon\,\|u-u_0\|_{L^\infty}\ \operatorname{meas}(\Omega\setminus(B\cup B_\epsilon)) \\ &\quad\ge \epsilon\,\|u-u_0\|_{L^1(\Omega\setminus B)} - c\,\epsilon^{\kappa+1}\|u-u_0\|_{L^\infty}, \end{aligned}$$

where c is a constant $c \geq 1$. Setting $\epsilon := c^{-2/\kappa} \|u - u_0\|_{L^1}^{1/\kappa} \|u - u_0\|_{L^\infty}^{-1/\kappa}$ yields

$$\int_{\Omega \setminus B} (\beta \lambda_0 - p_0)(u - u_0) \geq c \frac{\|u - u_0\|_{L^1(\Omega \setminus B)}^{1+\frac{1}{\kappa}}}{\|u - u_0\|_{L^\infty}^{\frac{1}{\kappa}}}. \tag{3.7}$$

It remains to estimate $\int_B -p_0 (u - u_0) + \beta \|u - u_0\|_{L^1(B)}$. Since $u_0 = 0$ on B, we have

$$\int_B -p_0 (u - u_0) + \beta \|u - u_0\|_{L^1(B)} = \int_B (\beta \operatorname{sign}(u) - p_0) u.$$

Defining $\tilde{B}_\epsilon := \{x \in B : \ |p_0(x)| \leq \beta - \epsilon\}$, we can estimate

$$\int_B (\beta \operatorname{sign}(u) - p_0) u \geq \int_{\tilde{B}_\epsilon} \epsilon |u| \geq \epsilon \|u\|_{L^1(B)} - \epsilon \|u\|_{L^1(B \setminus \tilde{B}_\epsilon)}.$$

Again the measure of $B \setminus \tilde{B}_\epsilon$ can be bounded proportionally to ϵ^κ, which gives with similar arguments as above

$$\int_B -p_0(u - u_0) + \beta \|u - u_0\|_{L^1(B)} \geq c \frac{\|u - u_0\|_{L^1(B)}^{1+\frac{1}{\kappa}}}{\|u - u_0\|_{L^\infty}^{\frac{1}{\kappa}}}. \tag{3.8}$$

Combining (3.6), (3.7), and (3.8) gives the claim. ☐

As one can see in the proof, this result cannot be strengthened if one assumes that β is positive. There is also a connection to sufficient optimality conditions of second-order that take strongly active constraints into account. There, the strongly active sets are used to obtain a certain growth of the objective with respect to the L^1-norm.

Using the previous result on the growth of the functional, one can also prove a stability result for the controls. From Theorem 3.1 we can deduce $|J_0(y_0^\delta, u_0^\delta) - J_0(y_0, u_0)| \leq c\,\delta$ in the case $\beta = 0$. This would give together with the previous theorem the estimate $\|u_0^\delta - u_0\|_{L^1} \leq c\,\delta^{\frac{\kappa}{\kappa+1}}$. We will however derive a slightly stronger result by a direct proof.

THEOREM 3.20. *Let control constraints be prescribed everywhere on Ω, i.e. $\Omega_a = \Omega_b = \Omega$. Let us suppose that Assumption 3.2 is fulfilled with $A = \Omega$. Then there exists a constant $c > 0$ independent of δ such that*

$$\|u_0 - u_0^\delta\|_{L^1} \leq c\,\delta^\kappa,$$
$$\|u_0 - u_0^\delta\|_{L^2} \leq c\,\delta^{\frac{\kappa}{2}}.$$

Proof. Let us define the following subset of Ω

$$A_0^\delta := \{x \in \Omega : \ u_0 \neq u_0^\delta\}.$$

By Theorem 3.1, we have $\|p_0 - p_0^\delta\|_{L^\infty} \leq c\,\delta$. On the set $\{x \in \Omega : \ \big| |p_0(x)| - \beta \big| > \|p_0 - p_0^\delta\|_{L^\infty}\}$ we have that $p_0 - \beta$ and $p_0^\delta - \beta$ as well as $p_0 + \beta$ and $p_0^\delta + \beta$ have the same signs, which implies that on this set $u_0 = u_0^\delta$ holds. Hence, we obtain the inclusion

$$A_0^\delta = \{x \in \Omega : \ u_0 \neq u_0^\delta\} \subset \left\{x \in \Omega : \ \big| |p_0(x)| - \beta \big| \leq \|p_0 - p_0^\delta\|_{L^\infty} \right\},$$

and the measure of this set can be bounded by $\|p_0 - p_0^\delta\|_{L^\infty}^\kappa$ due to Assumption 3.2. This implies

$$\begin{aligned} \|u_0 - u_0^\delta\|_{L^2} &\leq \|u_0 - u_0^\delta\|_{L^\infty} \operatorname{meas}(A_0^\delta) \\ &\leq c\,\|p_0 - p_0^\delta\|_{L^\infty}^\kappa \\ &\leq c\,\delta^\kappa. \end{aligned}$$

The estimate $\|u_0 - u_0^\delta\|_{L^2} \le \|u_0 - u_0^\delta\|_{L^\infty}^{1/2} \|u_0 - u_0^\delta\|_{L^\infty}^{1/2}$ together with the fact that u_0, u_0^δ are bounded uniformly in $L^\infty(\Omega)$ due to the control constraints yields the claim. ∎

4. A parameter choice rule. An important issue in regularization methods is the choice of the regularization parameter in dependence of the noise level or discretization. Several principles are known in the context of inverse problems, see [3], with the Morozov discrepancy principle [6] being one of the most important ones. There the parameter α is defined as a function of δ as

$$\alpha(\delta) := \sup\{\alpha > 0 : \; \|\mathcal{S}u_\alpha^\delta - z^\delta\|_Y \le \tau\delta\}$$

with a given constant $\tau > 1$. That is, the parameter is chosen such that the residuum in the ill-posed equation $\mathcal{S}u = z$ is below a certain threshold that is proportional to δ. A direct application of this principle to the regularization of our optimization problem is not possible since the residual $\mathcal{S}u_0 - z$ is in general non-zero.
Using the inequality $\|\mathcal{S}u_\alpha^\delta - z^\delta\|_Y \le \|\mathcal{S}(u_\alpha - u_0)\|_Y + \delta$, we can replace the residual $\mathcal{S}u_\alpha^\delta - z^\delta$ by the error in the states $\mathcal{S}(u_\alpha - u_0)$: choosing α according to

$$\hat\alpha(\delta) = \sup\{\alpha > 0 : \; \|\mathcal{S}(u_\alpha^\delta - u_0)\|_Y \le (\tau - 1)\delta\}$$

gives $\hat\alpha(\delta) \le \alpha(\delta)$. While there is no sensible upper bound available of $\|\mathcal{S}u_\alpha^\delta - z^\delta\|_Y$ in the context of our optimization problem, we can derive an upper bound of $\|\mathcal{S}(u_\alpha^\delta - u_0)\|_Y$ that can be computed explicitly without the knowledge of u_0.
Throughout this section we require that control constraints are prescribed everywhere on Ω, i.e., $\Omega_a = \Omega_b = \Omega$. In order to simplify the exposition of the results, we set the parameter $\beta = 0$ in the sequel.
LEMMA 4.1. *It holds*

$$\frac{1}{4}\|y_\alpha^\delta - y_0\|_Y^2 \le \int_{\{p_\alpha^\delta > 0\}} p_\alpha^\delta \,(u_b - u_\alpha^\delta) + \int_{\{p_\alpha^\delta < 0\}} p_\alpha^\delta \,(u_a - u_\alpha^\delta) + \|z - z_\delta\|_Y^2.$$

Proof. By optimality of (y_0, u_0) we have

$$\begin{aligned}
0 &\le J_0(y_\alpha^\delta, u_\alpha^\delta) - J_0(y_0, u_0)\\
&= \frac{1}{2}\|y_\alpha^\delta - z\|_Y^2 - \frac{1}{2}\|y_0 - z\|_Y^2 + (\mathcal{S}(u_0 - u_\alpha^\delta) - (y_0 - y_\alpha^\delta),\, z_\delta - y_\alpha^\delta)_Y\\
&= -(y_\alpha^\delta - z, y_0 - y_\alpha^\delta) - \frac{1}{2}\|y_0 - y_\alpha^\delta\|_Y^2 - (y_0 - y_\alpha^\delta,\, z_\delta - y_\alpha^\delta)_Y + (u_0 - u_\alpha^\delta,\, p_\alpha^\delta)_{L^2}\\
&= -\frac{1}{2}\|y_0 - y_\alpha^\delta\|_Y^2 + (u_0 - u_\alpha^\delta,\, p_\alpha^\delta)_{L^2} - (z_\delta - z,\, y_0 - y_\alpha^\delta)_Y,
\end{aligned}$$

where we did a Taylor expansion of J_0 at y_α^δ. It remains to derive an upper bound for $(u_0 - u_\alpha^\delta,\, p_\alpha^\delta)_{L^2}$, in order to eliminate the unknown u_0 from the final estimate. We can bound the integral by bounding u_0 by the control bounds in dependence of the sign of p_α^δ:

$$(u_0 - u_\alpha^\delta,\, p_\alpha^\delta)_{L^2} \le \int_{\{p_\alpha^\delta > 0\}} p_\alpha^\delta \,(u_b - u_\alpha^\delta) + \int_{\{p_\alpha^\delta < 0\}} p_\alpha^\delta \,(u_a - u_\alpha^\delta).$$

Applying Young's inequality to the term $-(z_\delta - z,\, y_0 - y_\alpha^\delta)_Y$ gives the claim. ∎

This result motivates the following definition of the regularization parameter in dependence of δ:

$$\alpha(\delta) := \sup \left\{ \alpha > 0 : \int_{\{p_\alpha^\delta > 0\}} p_\alpha^\delta \, (u_b - u_\alpha^\delta) + \int_{\{p_\alpha^\delta < 0\}} p_\alpha^\delta \, (u_a - u_\alpha^\delta) \le \frac{1}{2} \tau^2 \delta^2 \right\} \quad (4.1)$$

Here, the constant $\tau > 0$ can be chosen arbitrary. By convention, we set $\alpha(\delta) = 0$ if the set on the right hand side in (4.1) is empty.

The construction of $\alpha(\delta)$ is tailored to the case that u_0 is at the bounds everywhere in Ω. This can be seen in the proof of Lemma 4.1 above, where we replaced u_0 by u_a and u_b. Hence, the results that follow rely on the fulfillment of Assumption 3.2 with $A = \Omega$. That is, we assume that u_0 is everywhere at the bounds. It is an open problem, to extend these considerations to the general case $I \neq \emptyset, A \neq \Omega$.

As it will turn out later, we have $\alpha(\delta) > 0$ under Assumption 3.2. At first let us prove that $\alpha(\delta) < +\infty$ for sufficiently small δ.

Lemma 4.2. *Let us assume that $\mathcal{S}^\star z \neq 0$. Suppose further that there is a number $\sigma > 0$ such that $u_b(x) > \sigma$ and $-\sigma > u_a(x)$ a.e. on Ω.*

Then for δ sufficiently small, we have $\alpha(\delta) < +\infty$.

Proof. Since $\hat{u} \equiv 0$ is admissible, we have for $\alpha \to \infty$ that $u_\alpha^\delta \to 0$ in $L^2(\Omega)$, and hence $y_\alpha^\delta \to 0$ in Y, and $p_\alpha^\delta \to \mathcal{S}^\star z^\delta$ in $L^\infty(\Omega)$. This implies

$$-\int_{\{p_\alpha^\delta > 0\}} p_\alpha^\delta \, u_\alpha^\delta - \int_{\{p_\alpha^\delta < 0\}} p_\alpha^\delta \, u_\alpha^\delta = -\int_\Omega p_\alpha^\delta \, u_\alpha^\delta \to 0$$

Due to the assumption it holds

$$\int_{\{p_\alpha^\delta > 0\}} p_\alpha^\delta \, u_b + \int_{\{p_\alpha^\delta < 0\}} p_\alpha^\delta \, u_a \ge \sigma \|p_\alpha^\delta\|_{L^1}.$$

Hence we have

$$\begin{aligned} \lim_{\alpha \to \infty} \int_{\{p_\alpha^\delta > 0\}} p_\alpha^\delta \, (u_b - u_\alpha^\delta) + \int_{\{p_\alpha^\delta < 0\}} p_\alpha^\delta \, (u_a - u_\alpha^\delta) &\ge \sigma \|\mathcal{S}^\star z_\delta\|_{L^1} \\ &\ge \sigma \left(\|\mathcal{S}^\star z\|_{L^1} - \text{meas}(\Omega) \|\mathcal{S}^\star\|_{\mathcal{L}(Y, L^\infty)} \, \delta \right). \end{aligned}$$

If δ satisfies

$$\delta \ge \min \left(\frac{\|\mathcal{S}^\star z\|_{L^1}}{2 \, \text{meas}(\Omega) \|\mathcal{S}^\star\|_{\mathcal{L}(Y, L^\infty)}}, \frac{(\sigma \|\mathcal{S}^\star z\|_{L^1})^{1/2}}{\tau} \right)$$

the term $\int_{\{p_\alpha^\delta > 0\}} p_\alpha^\delta \, (u_b - u_\alpha^\delta) + \int_{\{p_\alpha^\delta < 0\}} p_\alpha^\delta \, (u_a - u_\alpha^\delta)$ is larger than $\frac{1}{2}\tau^2\delta^2$ for large α. Consequently $\alpha(\delta) < +\infty$ for these small δ. □

Second, we prove $\alpha(\delta) > 0$ under our regularity assumption. In order to get this lower bound on $\alpha(\delta)$ we prove an upper bound of the quantity $\int_{\{p_\alpha^\delta > 0\}} p_\alpha^\delta \, (u_b - u_\alpha^\delta) + \int_{\{p_\alpha^\delta < 0\}} p_\alpha^\delta \, (u_a - u_\alpha^\delta)$ first.

Lemma 4.3. *Let us suppose Assumption 3.2 is satisfied with $A = \Omega$ and $\kappa < \infty$. Then it holds*

$$0 \le \int_{\{p_\alpha^\delta > 0\}} p_\alpha^\delta \, (u_b - u_\alpha^\delta) + \int_{\{p_\alpha^\delta < 0\}} p_\alpha^\delta \, (u_a - u_\alpha^\delta) \le c \, \alpha \left(\|p_\alpha^\delta - p_0\|_{L^\infty}^\kappa + \alpha^\kappa \right) \quad (4.2)$$

with a constant $c > 0$ independent of α, δ.

Proof. The non-negativity of the integrals follows from $u_\alpha^\delta \in U_{ad}$, i.e., $u_a \le u_\alpha^\delta \le u_b$ a.e. in Ω.

By the optimality conditions for the regularized problem, cf. Theorem 2.3 and inequality (2.2), it holds

$$p_\alpha^\delta(x)\big(u(x) - u_\alpha^\delta(x)\big) \le \alpha\, u_\alpha^\delta(x)\big(u(x) - u_\alpha^\delta(x)\big) \quad \text{f.a.a. } x \in \Omega \text{ and for all } u \in U_{ad}.$$

Let us first consider the integral over the set $\{p_\alpha^\delta > 0\}$. Here, we distinguish subsets according to the sign of p_0. By Assumption 3.2, we have that the set $\{p_0 = 0\}$ has zero measure. On the set $\{p_0 > 0\}$ we have $u_0 = u_b$, whereas it holds $u_0 = u_a$ on $\{p_0 < 0\}$. Hence we have

$$\int_{\{p_\alpha^\delta > 0,\, p_0 > 0\}} p_\alpha^\delta\,(u_b - u_\alpha^\delta) \le \int_{\{p_\alpha^\delta > 0,\, p_0 > 0,\, u_0 \ne u_\alpha^\delta\}} \alpha\, u_\alpha^\delta\,(u_0 - u_\alpha^\delta).$$

Analogously to Lemma 3.12, we obtain

$$\operatorname{meas}(\{u_0 \ne u_\alpha^\delta\}) \le c(\|p_\alpha^\delta - p_0\|_{L^\infty}^\kappa + \alpha^\kappa).$$

This implies the estimate

$$\begin{aligned}\int_{\{p_\alpha^\delta > 0,\, p_0 > 0\}} p_\alpha^\delta\,(u_b - u_\alpha^\delta) &\le \alpha\ \operatorname{meas}(\{u_0 \ne u_\alpha^\delta\})\, \|u_\alpha^\delta\|_{L^\infty}\|\, \|u_\alpha^\delta - u_0\|_{L^\infty} \\ &\le c\,\alpha\,(\|p_\alpha^\delta - p_0\|_{L^\infty}^\kappa + \alpha^\kappa).\end{aligned}$$

In addition we have with $\epsilon > 0$

$$\begin{aligned}\int_{\{p_\alpha^\delta > 0,\, p_0 < 0\}} p_\alpha^\delta\,(u_b - u_\alpha^\delta) &\le \int_{\{p_\alpha^\delta > 0,\, p_0 < 0\}} \alpha\, u_\alpha^\delta\,(u_b - u_\alpha^\delta) \\ &\le \int_{\{p_\alpha^\delta > 0,\, p_0 \le -\epsilon\}} \alpha\, u_\alpha^\delta\,(u_b - u_\alpha^\delta) + \int_{\{p_\alpha^\delta > 0,\, -\epsilon < p_0 < 0\}} \alpha\, u_\alpha^\delta\,(u_b - u_\alpha^\delta).\end{aligned}$$

Let us estimate the measure of both integration regions. By Chebyshev's inequality it holds for all $q \ge 1$

$$\operatorname{meas}(\{p_\alpha^\delta > 0,\, p_0 \le -\epsilon\}) \le \operatorname{meas}(\{|p_\alpha^\delta - p_0| > \epsilon\}) \le \frac{\|p_\alpha^\delta - p_0\|_{L^q}^q}{\epsilon^q}.$$

By Assumption 3.2.2 we have

$$\operatorname{meas}(\{-\epsilon < p_0 < 0\}) \le c\,\epsilon^\kappa.$$

With $\epsilon := \|p_\alpha^\delta - p_0\|_{L^q}^{\frac{q}{\kappa+q}}$ we obtain

$$\left|\int_{\{p_\alpha^\delta > 0,\, p_0 < 0\}} p_\alpha^\delta\,(u_b - u_\alpha^\delta)\right| \le c\,\alpha\,\|p_\alpha^\delta - p_0\|_{L^q}^{\frac{\kappa q}{\kappa+q}},$$

where the constant c is in particular independent of q. Hence we obtain for $q \to \infty$

$$\left|\int_{\{p_\alpha^\delta > 0,\, p_0 < 0\}} p_\alpha^\delta\,(u_b - u_\alpha^\delta)\right| \le c\,\alpha\,\|p_\alpha^\delta - p_0\|_{L^\infty}^\kappa.$$

With similar arguments we find the upper bound for the second integral in (4.2). A close inspection of the proof yields that the constant c in (4.2) depends only on $\|u_a\|_{L^\infty}, \|u_b\|_{L^\infty}$, and the constant of Assumption 3.2.2 appearing in (3.1). ☐

Let us remark that the result of this Lemma is not true if Assumption 3.2 is fulfilled with $A \neq \Omega$, that is if a non-trivial subset $I \subset \Omega$ exists, where p_0 is zero, and u_0 is not at the bounds.

With the help of the previous Lemma 4.3, we can prove that $\alpha(\delta)$ is positive.

COROLLARY 4.4. *Let us suppose Assumption 3.2 is satisfied with $A = \Omega$ and $\kappa < \infty$. Then $\alpha(\delta) > 0$.*

Proof. Since control constraints are given everywhere on Ω, u_α^δ and hence p_α^δ are uniformly bounded in $L^\infty(\Omega)$. Then the right-hand side of (4.2) tends to zero for $\alpha \to 0$, which implies that for sufficiently small α, this quantity is smaller than $\frac{1}{2}\tau^2\delta^2$. Therefore, the supremum in (4.1) is positive. ☐

This proves that $\alpha(\delta)$ is well-defined and not trivial under certain conditions. Now let us turn to the convergence analysis of the regularization scheme for $\delta \to 0$. Here, one has to ensure that the convergence $\alpha(\delta) \to 0$ is not too fast, which would result in a non-optimal convergence order.

Let us first prove the optimal convergence order for the states and adjoints.

LEMMA 4.5. *Let $\delta > 0$ and $\alpha(\delta)$ be given from* (4.1). *Then there is $c > 0$ independent of δ such that*

$$\|y_{\alpha(\delta)} - y_0\|_Y + \|y^\delta_{\alpha(\delta)} - y_0\|_Y \le c\,\delta.$$

Proof. Let $\alpha := \alpha(\delta)$. Then by Lemma 4.1 and (4.1)

$$\frac{1}{4}\|y_\alpha^\delta - y_0\|_Y^2 \le \frac{1}{2}\tau^2\delta^2 + \delta^2,$$

which gives $\|y_\alpha^\delta - y_0\|_Y \le 2\sqrt{1+\tau^2/2}\;\delta$. By Proposition 2.9 we obtain

$$\|y_\alpha - y_0\|_Y \le \|y_\alpha^\delta - y_\alpha\|_Y + \|y_\alpha^\delta - y_0\|_Y \le (1 + 2\sqrt{1+\tau^2/2})\;\delta.$$

☐

The next step is to establish a lower bound on $\alpha(\delta)$.

LEMMA 4.6. *Let $\delta > 0$ be given. Let $\alpha(\delta)$ satisfy $0 < \alpha(\delta) < \infty$.*
Let us suppose that Assumption 3.2 is satisfied with $A = \Omega$.
If $\kappa \le 1$ and $\alpha(\delta) \le 1$ then there exists a constant c independent of δ such that

$$\frac{\delta^2}{\alpha(\delta)} \le c\,\delta^\kappa.$$

If $1 \le \kappa < \infty$ then for each $\delta_{max} > 0$ there exists a constant $c = c(\delta_{max})$ independent of δ such that if $\delta < \delta_{max}$

$$\frac{\delta^2}{\alpha(\delta)} \le c\,\delta^{\frac{2\kappa}{\kappa+1}}.$$

Proof. Let us consider the regularized problem with regularization parameter $\alpha := 2\,\alpha(\delta) > \alpha(\delta)$. Since $\alpha > \alpha(\delta)$ it holds

$$\frac{1}{2}\tau^2\delta^2 \le \int_{\{p_\alpha^\delta > 0\}} p_\alpha^\delta\,(u_b - u_\alpha^\delta) + \int_{\{p_\alpha^\delta < 0\}} p_\alpha^\delta\,(u_a - u_\alpha^\delta).$$

By Lemma 4.3, we have

$$\delta^2 \le c\,\alpha\left(\|p_\alpha^\delta - p_0\|_{L^\infty}^\kappa + \alpha^\kappa\right).$$

Let us first investigate the case $\kappa \le 1$. Using the results of Theorems 3.1 and 3.14 on a-priori regularization error estimates we obtain

$$\delta^2 \le c\,\alpha\left(\delta^\kappa + \alpha^{\frac{\kappa}{2-\kappa}} + \alpha^\kappa\right) \le \frac{1}{2}\delta^2 + c\left(\alpha^{\frac{2}{2-\kappa}} + \alpha^{\kappa+1}\right). \tag{4.3}$$

Here we used again Young's inequality. Hence, we get

$$\delta^2 \le c\left(\alpha^{\frac{2}{2-\kappa}} + \alpha^{\kappa+1}\right) = c\,\alpha^{\frac{2}{2-\kappa}}\ \left(1 + \alpha^{\frac{\kappa(1-\kappa)}{2-\kappa}}\right).$$

Since $\alpha(\delta) \le 1$ implies $\alpha \le 2$, the term in brackets on the right-hand side is bounded uniformly with respect to α, and we obtain

$$\delta^{2-\kappa} \le c\,\alpha,$$

which is equivalent to

$$\frac{\delta^2}{\alpha} \le c\,\delta^\kappa,$$

where the constant c is independent of δ.
In the case $\kappa > 1$ we have to replace (4.3) according to Theorem 3.14 by $\delta^2 \le c\,\alpha\,(\delta^\kappa + \alpha^\kappa)$. Let $\delta_{\max} > 0$ be given. Applying Young's inequality we get with $c = c(\delta_{\max})$

$$\delta^2 \le \frac{1}{2\delta_{\max}^{\kappa-1}}\delta^{\kappa+1} + c\,\alpha^{\kappa+1},$$

which is equivalent to

$$\delta^2\left(1 - \frac{1}{2}\left(\frac{\delta}{\delta_{\max}}\right)^{\kappa-1}\right) \le c\,\alpha^{\kappa+1}.$$

This implies for $\delta < \delta_{\max}$ that $\frac{\delta^2}{\alpha(\delta)} \le c\,\delta^{\frac{2\kappa}{\kappa+1}}$ with c depending on $\delta_{\max}$ but not on δ.
□
Let us remark that in the case $\alpha(\delta) > 1$ the inequality $\frac{\delta^2}{\alpha} \le \delta^2$ holds. That means, we do not need an upper bound on $\alpha(\delta)$ in the subsequent convergence analysis, only the existence of the supremum is needed, i.e., $\alpha(\delta) < \infty$.
Now, we have everything at hand to prove the main result of this section: convergence of the regularization scheme with the same order with respect to δ as given by best a-priori parameter choice, see Section 3.5.
THEOREM 4.7. *Let control constraints be given everywhere, i.e.,* $\Omega_a = \Omega_b = \Omega$. *Moreover, let us suppose Assumption 3.2 is satisfied with* $A = \Omega$.
Then for every $\delta_{max} > 0$ *there is a positive constant* $c = c(\delta_{max})$ *such that*

$$\|u_\alpha^\delta - u_0\|_Y \le c\,\delta^s$$

holds for all $\delta \in (0, \delta_{max})$ *with* s *defined by*

$$s = \begin{cases} \frac{\kappa}{2} & \text{if } \kappa \le 1, \\ \frac{\kappa}{\kappa+1} & \text{if } 1 < \kappa < \infty. \end{cases}$$

Proof. Let us set $\alpha := \alpha(\delta)$. Let us first prove the claim for $\kappa \le 1$. We have by Proposition 2.9

$$\alpha \, \|u_0 - u_\alpha^\delta\|_{L^2}^2 \le \alpha \, (u_0, u_0 - u_\alpha^\delta) + (z^\delta - z, \, y_0 - y_\alpha^\delta).$$

Since control constraints are prescribed everywhere, we have $u_0 \in L^\infty(\Omega)$. Using Lemma 4.5 we can then estimate

$$\alpha \, \|u_0 - u_\alpha^\delta\|_{L^2}^2 \le c \, (\alpha \, \|u_0 - u_\alpha^\delta\|_{L^1} + \delta^2).$$

Applying Lemma 4.6 gives in the case that $\alpha \le 1$

$$\|u_0 - u_\alpha^\delta\|_{L^2}^2 \le c(\|u_0 - u_\alpha^\delta\|_{L^1} + \delta^\kappa). \tag{4.4}$$

If α is greater than 1, then it trivially holds

$$\frac{\delta^2}{\alpha} \le \delta^2,$$

and (4.4) is valid in this case, too, with the constant c depending on $\delta_{\max}$.
It remains to bound $\|u_0 - u_\alpha^\delta\|_{L^1}$ in terms of δ. To this end, let us define the following subset of Ω:

$$B := \{p_\alpha^\delta \ne 0, \ \operatorname{sign}(p_0) = \operatorname{sign}(p_\alpha^\delta)\}.$$

The measure of its complement can be bound using Assumption 3.2. Indeed, on $\Omega \setminus B$ the signs of p_0 and p_α^δ are different, which gives $|p_0| \le |p_0 - p_\alpha^\delta|$ on $\Omega \setminus B$. Hence using Assumption 3.2 and Lemma 4.5 we obtain

$$\operatorname{meas}(\Omega \setminus B) \le c \, \|p_0 - p_\alpha^\delta\|_{L^\infty}^\kappa \le c \, \delta^\kappa. \tag{4.5}$$

Let us investigate now the L^1-norm of $u_0 - u_\alpha^\delta$ on B. For $\epsilon > 0$ let us define the set

$$B_\epsilon := B \cap \{|p_\alpha^\delta| > \epsilon\}.$$

Since $|p_0| \le |p_\alpha^\delta| + |p_0 - p_\alpha^\delta| \le \epsilon + |p_0 - p_\alpha^\delta|$ on $B \setminus B_\epsilon$, we have with Assumption 3.2 and Lemma 4.5

$$\operatorname{meas}(B \setminus B_\epsilon) \le c(\|p_0 - p_\alpha^\delta\|_{L^\infty}^\kappa + \epsilon^\kappa) \le c(\delta^\kappa + \epsilon^\kappa). \tag{4.6}$$

Let us recall that α satisfies the discrepancy estimate, cf. (4.1),

$$\int_{\{p_\alpha^\delta > 0\}} p_\alpha^\delta \, (u_b - u_\alpha^\delta) + \int_{\{p_\alpha^\delta < 0\}} p_\alpha^\delta \, (u_a - u_\alpha^\delta) \le \frac{1}{2}\tau^2\delta^2.$$

Here, the integrands in both integrals are positive functions, which allows us to restrict the integration regions

$$\int_{\{p_\alpha^\delta > 0\} \cap B_\epsilon} p_\alpha^\delta \, (u_b - u_\alpha^\delta) + \int_{\{p_\alpha^\delta < 0\} \cap B_\epsilon} p_\alpha^\delta \, (u_a - u_\alpha^\delta) \le \frac{1}{2}\tau^2\delta^2.$$

Since $|p_\alpha^\delta| \ge \epsilon$ on B_ϵ, it holds

$$\int_{\{p_\alpha^\delta > 0\} \cap B_\epsilon} \epsilon |u_b - u_\alpha^\delta| + \int_{\{p_\alpha^\delta < 0\} \cap B_\epsilon} \epsilon |u_\alpha^\delta - u_a| \le \frac{1}{2}\tau^2\delta^2$$

Since p_0 and p_α^δ have equal signs on B_ϵ, and $\{p_\alpha^\delta \neq 0\} \supset B_\epsilon$, we have

$$\epsilon \int_{B_\epsilon} |u_0 - u_\alpha^\delta| \leq \frac{1}{2}\tau^2\delta^2,$$

which implies

$$\|u_0 - u_\alpha^\delta\|_{L^1(B_\epsilon)} \leq \frac{1}{2}\tau^2\delta^2\epsilon^{-1}.$$

This implies together with (4.5), (4.6) that

$$\|u_0 - u_\alpha^\delta\|_{L^1} \leq c(\delta^\kappa + \epsilon^\kappa + \delta^2\epsilon^{-1}).$$

With $\epsilon := \delta^{\frac{2}{\kappa+1}}$ we obtain for $\kappa \leq 1$ and $\delta \leq \delta_{\max}$

$$\|u_0 - u_\alpha^\delta\|_{L^1} \leq c(\delta^\kappa + \delta^{\frac{2\kappa}{\kappa+1}}) = c\,(1 + \delta^{\frac{\kappa(1-\kappa)}{\kappa+1}})\delta^\kappa \leq c\,\delta^\kappa \tag{4.7}$$

which proves with (4.4)

$$\|u_0 - u_\alpha^\delta\|_{L^2} \leq c\,\delta^{\kappa/2}$$

which is the optimal rate. Here the constant c depends on $\delta_{\max}$.
Let us now sketch the proof for the case $\kappa > 1$. Here, we have to replace (4.4) according to Lemma 4.6 by

$$\|u_0 - u_\alpha^\delta\|_{L^2}^2 \leq c(\|u_0 - u_\alpha^\delta\|_{L^1} + \delta^{\frac{2\kappa}{\kappa+1}}). \tag{4.8}$$

Since we are in the situation $\kappa > 1$, we have to modify estimate (4.7) to

$$\|u_0 - u_\alpha^\delta\|_{L^1} \leq c(\delta^\kappa + \delta^{\frac{2\kappa}{\kappa+1}}) = c(\delta^{\frac{\kappa(\kappa-1)}{\kappa+1}} + 1)\delta^{\frac{2\kappa}{\kappa+1}} \leq c\,\delta^{\frac{2\kappa}{\kappa+1}},$$

where c depends on $\delta_{\max}$. This finishes the proof. □
As mentioned in this section, it is an open problem, how these ideas can be transferred to the the case that the control constraints are not active everywhere.
Moreover, it will be interesting to see how these results can be used to choose the regularization parameter α in dependence of discretization parameters. In particular, an adaptive algorithm that combines adaptive discretization schemes and adaptive regularization parameter choices has to be developed.

REFERENCES

[1] C. Clason and K. Kunisch. A duality-based approach to elliptic control problems in non-reflexive Banach spaces. *ESAIM: Control, Optimisation and Calculus of Variations*, 2009. To appear.

[2] K. Deckelnick and M. Hinze. A note on the approximation of elliptic control problems with bang-bang controls. *Comput. Optim. Appl.*, 2010. To appear.

[3] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.

[4] U. Felgenhauer. On stability of bang-bang type controls. *SIAM J. Control Optim.*, 41(6):1843–1867, 2003.

[5] B. Hofmann, D. Düvelmeyer, and K. Krumbiegel. Approximative source conditions in Tikhonov regularization — new analytical results and some numerical studies. *Mathematical Modeling and Analysis*, 11(1):41–56, 2006.

[6] V. A. Morozov. *Regularization methods for ill-posed problems*. CRC Press, Boca Raton, FL, 1993. Translated from the 1987 Russian original.

[7] F. Tröltzsch. *Optimal Control of Partial Differential Equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2010. Theory, methods and applications, Translated from the 2005 German original by J. Sprekels.

[8] D. Wachsmuth and G. Wachsmuth. Convergence and Regularisation Results for Optimal Control Problems with Sparsity Functional. *ESAIM: Control, Optimisation and Calculus of Variations*, 2009. To appear.