

# Robust error estimates for regularization and discretization of bang-bang control problems <sup>\*</sup>

Daniel Wachsmuth<sup>†</sup>

September 2, 2015

## Abstract

We investigate the simultaneous regularization and discretization of an optimal control problem with pointwise control constraints. Typically such problems exhibit bang-bang solutions: the optimal control almost everywhere takes values at the control bounds. We derive discretization error estimates that are robust with respect to the regularization parameter. These estimates can be used to make an optimal choice of the regularization parameter with respect to discretization error estimates.

**Keywords.** Optimal control, bang-bang control, Tikhonov regularization, parameter-choice rule.

**AMS classification.** 49K20, 49N45, 65K15

## 1 Introduction

In this article we investigate the regularization and discretization of bang-bang control problems. The class of problems that we consider can be described as the minimization of

$$\frac{1}{2} \|Su - z\|_Y^2 \tag{P}$$

over all  $u \in L^2(D)$  satisfying the constraint

$$u_a \leq u \leq u_b \text{ a.e. on } D. \tag{1.1}$$

In line with the usual nomenclature in optimal control, the variable  $u$  will be called control, the variable  $y := Su$  will be called state. In the problem above  $D$  is a bounded subset of  $\mathbb{R}^n$ . The operator  $S$  is assumed to be linear and continuous from  $L^2(D)$  to  $Y$  with  $Y$  being a Hilbert space. The target state  $z \in Y$  is a given desired state. Moreover, we assume that the Hilbert space adjoint operator  $S^*$  of  $S$  maps from  $Y$  to  $L^\infty(D)$ . Here, we have in mind optimal control problems for linear partial differential equations.

In order to numerically solve (P), let us introduce a family of linear and continuous operators  $\{S_h\}_{h>0}$  from  $L^2(D)$  to  $Y$  with finite-dimensional range  $Y_h \subset Y$ ,

---

<sup>\*</sup>This work was partially funded by Austrian Science Fund (FWF) grant P23848.

<sup>†</sup>Institut für Mathematik, Universität Würzburg, 97074 Würzburg, Germany, [daniel.wachsmuth@mathematik.uni-wuerzburg.de](mailto:daniel.wachsmuth@mathematik.uni-wuerzburg.de)

where  $h$  denotes the discretization parameter. We assume  $S_h \rightarrow S$  for  $h \searrow 0$  in a sense clarified below, see Section 1.4. Furthermore, we consider the Tikhonov regularization with parameter  $\alpha > 0$ . The regularized and discretized problem now reads: Minimize

$$\frac{1}{2} \|S_h u - z\|_Y^2 + \frac{\alpha}{2} \|u\|_{L^2(D)}^2 \quad (\text{P}_{\alpha,h})$$

subject to (1.1). Let us note that the control space is not discretized, which is the variational discretization concept introduced in [5].

Here, one is interested in convergence results with respect to  $(\alpha, h) \searrow 0$ . Moreover, the choice of the parameter  $\alpha$  depending on discretization parameters is important for actual numerical computations. Let us briefly review existing literature on this subject. Most of the existing results assume a bang-bang structure of the optimal control  $u_0$ :  $u_0(x) \in \{u_a(x), u_b(x)\}$  a.e. on  $D$ . Convergence rate estimates with respect to  $\alpha \searrow 0$  for the undiscretized problem can be found in [12, 13]. There, an assumption on the measure of the almost-inactive set is used. Such an assumption was applied in different situation in the literature as well, we mention only [4, 10]. Convergence rate estimates of the regularization error are also available without the assumption of bang-bang structure. There one has to resort to source conditions, see e.g. [7], and combinations of source condition and active set conditions [13]. Discretization error estimates are well known in the literature. A-priori discretization error estimates concerning the discretization of  $(\text{P}_\alpha)$  for fixed  $\alpha > 0$  can be found for instance in [5, 6]. In the case  $\alpha = 0$  the analysis is much more delicate, a-priori error discretization estimates for this case can be found in [3]. Let us mention that the error estimates for the regularized problem  $\alpha > 0$  are not robust with respect to  $\alpha$ . Consequently, the results in the case  $\alpha = 0$  cannot be obtained by passing to the limit  $\alpha \searrow 0$ . The coupling of discretization and regularization was discussed in [11]. There the regularization parameter  $\alpha$  is chosen depending on a-priori or a-posteriori discretization error estimates.

Let us emphasize, that almost all of the above cited literature is concerned with convex problems. The extension to the non-convex case is not straight-forward, we refer to [1] for results on second-order sufficient optimality conditions for bang-bang control problems. In a recent work, multi-bang control problems were investigated [2].

In this article, we discuss robust discretization error estimates for  $\alpha > 0$ . In these estimates, the regularization parameter can tend to zero, and the resulting estimate coincides with that of [3] in the case  $\alpha = 0$ . Of course, the robust estimates are not optimal in the discretization parameters. A second result is the arising choice of the regularization parameter depending on discretization quantities. It turns out that it is optimal to choose  $\alpha$  proportional to the  $L^\infty$ -discretization error. The surprising fact is that this choice is *independent* of some unknown constants appearing in the assumption on the almost-inactive set.

## Notation

In the sequel, we will frequently use generic constants  $c > 0$  that may change from line to line, but which are independent of relevant quantities such as  $\alpha$  and  $h$ .

## 1.1 Assumptions and preliminary results

Let us specify the standing assumptions that we will use throughout the paper. Let  $(D, \Sigma, \mu)$  be a given finite measure space. The operator  $S$  is assumed to be linear and continuous from  $L^2(D)$  to the Hilbert space  $Y$ . Moreover, we assume that the Hilbert space adjoint operator  $S^*$  of  $S$  maps from  $Y$  to  $L^\infty(D)$ . Let us remark that the requirements on  $S^*$  could be relaxed to allow  $S^*$  mapping into  $L^p(D)$ ,  $p \in (2, \infty)$ , see [13].

The control constraints are given with  $u_a, u_b \in L^\infty(D)$  and  $u_a \leq u_b$  a.e. on  $D$ . The set of admissible controls is defined by

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

As already introduced, we will work with a family of operators  $\{S_h\}_{h>0}$ ,  $S_h \in \mathcal{L}(L^2(D), Y)$  with finite-dimensional range. The adjoint operators  $S_h^*$  are assumed to map from  $Y$  to  $L^\infty(D)$ .

For completeness, let us define the regularized version of the undiscretized problem: Given  $\alpha > 0$ , minimize

$$\frac{1}{2} \|Su - z\|_Y^2 + \frac{\alpha}{2} \|u\|_{L^2(D)}^2 \quad (\text{P}_\alpha)$$

subject to the inequality constraints (1.1).

Due to classical results, the problems (P),  $(\text{P}_\alpha)$ ,  $(\text{P}_{\alpha,h})$  admit solutions.

**Proposition 1.1.** *The problems (P) and  $(\text{P}_{\alpha,h})$ , are solvable with convex and bounded sets of solutions. The problems  $(\text{P}_\alpha)$  and  $(\text{P}_{\alpha,h})$  are uniquely solvable for  $\alpha > 0$ .*

*The set of optimal states of (P)  $\{Su : u \text{ solves (P)}\}$  is a singleton. Moreover, (P) is uniquely solvable if  $S$  is injective.*

*Proof.* Due to the control constraints, the admissible set  $U_{ad}$  is bounded in  $L^2(D)$ . Moreover, it is convex and closed in  $L^2(D)$  hence weakly-closed. By assumption, the functionals to be minimized are convex and continuous from  $L^2(D)$  to  $\mathbb{R}$ , which implies weakly lower semicontinuity. The existence of solutions of the problems (P),  $(\text{P}_\alpha)$ , and  $(\text{P}_{\alpha,h})$  follows from the Weierstraß theorem, where  $L^2(D)$  is supplied with the weak topology. Since the objective functional and the admissible set are convex, the set of solutions of these problems is convex.

If  $\alpha > 0$  then the objective functionals of  $(\text{P}_\alpha)$  and  $(\text{P}_{\alpha,h})$  are strictly convex, which implies that the solutions are unique. Moreover, the mapping  $Su \mapsto \frac{1}{2} \|Su - z\|_Y^2$  is strictly convex, hence the sets of optimal states for (P) and  $(\text{P}_{\alpha,h})$  for  $\alpha = 0$  are singletons. If in addition  $S$  is injective, then the uniqueness of solutions of (P) follows.  $\square$

Let us note that the solutions of (P) and  $(\text{P}_{\alpha,h})$  may not be uniquely determined if  $\alpha = 0$ . However, the optimal states of (P) and  $(\text{P}_{0,h})$  are uniquely determined due to the strict convexity of the cost functional w.r.t.  $Su$ .

## 1.2 Necessary optimality conditions

The solutions of the considered problems can be characterized by means of first-order necessary optimality conditions, which are sufficient as well due to the convexity of the problems.

**Proposition 1.2.** For  $\alpha \geq 0$  let  $u_\alpha$  and  $u_{\alpha,h}$  be solutions of  $(P_\alpha)$  and  $(P_{\alpha,h})$ , respectively. Let us define  $y_\alpha := Su_\alpha$ ,  $y_{\alpha,h} := S_h u_{\alpha,h}$ ,  $p_\alpha := S^*(y_\alpha - z)$ , and  $p_{\alpha,h} := S_h^*(y_{\alpha,h} - z)$ . Then it holds

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

and

$$(\alpha u_{\alpha,h} + p_{\alpha,h}, u - u_{\alpha,h}) \geq 0 \quad \forall u \in U_{ad}.$$

These variational inequalities imply the following pointwise variational inequality [9, Lemma 2.26]

$$(\alpha u_\alpha(x) + p_\alpha(x))(u - u_\alpha(x)) \geq 0 \quad \forall u \in [u_a(x), u_b(x)] \quad \text{f.a.a. } x \in D. \quad (1.2)$$

In the case  $\alpha > 0$  this yields pointwise a.e. representations for the optimal control

$$u_\alpha(x) = \text{proj}_{[u_a(x), u_b(x)]} \left( -\frac{1}{\alpha} p_\alpha(x) \right) \quad \text{f.a.a. } x \in D.$$

If  $\alpha = 0$  we have

$$u_0(x) \begin{cases} = u_a(x) & \text{if } p_0(x) > 0 \\ \in [u_a(x), u_b(x)] & \text{if } p_0(x) = 0 \\ = u_b(x) & \text{if } p_0(x) < 0 \end{cases} \quad \text{f.a.a. } x \in D. \quad (1.3)$$

Similar relations hold for  $u_{\alpha,h}$  and  $u_{0,h}$ . For  $\alpha = 0$ , the controls  $u_0$  and  $u_{0,h}$  are bang-bang if  $p_0 \neq 0$  and  $p_{0,h} \neq 0$  a.e. on  $D$ , respectively. Moreover, if  $p_0 = 0$  and  $p_{0,h} = 0$  on sets of positive measure then the values of  $u_0$  and  $u_{0,h}$  cannot be determined by the respective variational inequalities (1.2).

### 1.3 Regularization error estimate

Let us now recall the assumption on the almost-inactive sets. It is widely used in the literature, as it can be viewed as a strengthened complementarity condition.

**Assumption 1.** Let us assume that there are  $\kappa > 0$ ,  $c > 0$  such that

$$\text{meas} \{x \in D : |p_0(x)| \leq \epsilon\} \leq c \epsilon^\kappa$$

holds for all  $\epsilon > 0$ .

As discussed above, the optimal state  $y_0 = Su_0$  and consequently the optimal adjoint state  $p_0 = S^*(y_0 - z)$  are uniquely determined. Under the assumption above, the optimal control  $u_0$  is uniquely determined as well and has bang-bang type.

**Corollary 1.3.** Let Assumption 1 be satisfied. Then the problem (P) is uniquely solvable, and its solution  $u_0$  is bang-bang.

*Proof.* By proposition 1.1 the optimal state  $y_0 := Su_0$  with  $u_0$  being a solution of (P) is uniquely determined. Hence, also the adjoint  $p_0 = S^*(y_0 - z)$  is uniquely determined. Assumption 1 implies that the set  $\{x \in D : p_0(x) = 0\}$  has measure zero. Then the pointwise representation (1.3) of solutions of (P) completely determines the values of a solution  $u_0$  on  $D$  up to sets of zero measure. This implies that (P) is uniquely solvable. Moreover, it follows  $u_0(x) \in \{u_a(x), u_b(x)\}$  for almost all  $x \in D$  by (1.3). Hence,  $u_0$  is of bang-bang type.  $\square$

Assumption 1 not only guarantees uniqueness of solutions of (P), but it is also sufficient to prove convergence rates with respect to  $\alpha$  for  $\alpha \searrow 0$ .

**Proposition 1.4.** *Let Assumption 1 be satisfied. Let  $d$  be defined by*

$$d = \begin{cases} \frac{1}{2-\kappa} & \text{if } \kappa \leq 1, \\ \frac{\kappa+1}{2} & \text{if } \kappa > 1. \end{cases}$$

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

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

*holds for all  $\alpha \in (0, \alpha_{\max}]$ . The constant  $c$  depends on  $\alpha_{\max}$ .*

*Proof.* For the proof we refer to [13, Theorem 3.2].  $\square$

Let us present a small result to ease the work with the convergence rates stated above.

**Lemma 1.5.** *Let  $\kappa > 0$  be given. Let  $d$  satisfy*

$$d = \begin{cases} \frac{1}{2-\kappa} & \text{if } \kappa \leq 1, \\ \frac{\kappa+1}{2} & \text{if } \kappa > 1. \end{cases}$$

*Then it holds*

$$\kappa \min(1, d) = 2d - 1.$$

*Proof.* In the case  $\kappa \leq 1$  we have  $\kappa \min(1, d) + 1 = \frac{\kappa}{2-\kappa} + 1 = \frac{2}{2-\kappa} = 2d$ , whereas in the case  $\kappa > 1$  we obtain  $\kappa \min(1, d) + 1 = \kappa + 1 = 2d$ .  $\square$

**Remark 1.6.** *With this small lemma at hand, we can simplify the convergence rate of Proposition 1.4 to*

$$\|u_0 - u_\alpha\|_{L^1(D)} \leq c \alpha^{2d-1},$$

*since it holds  $\frac{\kappa}{2} \min(1, d) = d - \frac{1}{2}$  by Lemma 1.5 above.*

## 1.4 Discretization error estimates

Let us now turn to the discretization of the considered optimal control problems. As already mentioned in the introduction, we consider approximations  $S_h$  of the operator  $S$ . In order to control the discretization error we make the following assumption.

**Assumption 2.** *There exist continuous and monotonically increasing functions  $\delta_2(h), \delta_\infty(h) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  with  $\delta_2(0) = \delta_\infty(0) = 0$  such that it holds*

$$\begin{aligned} \|(S - S_h)u_{\alpha,h}\|_Y + \|(S^* - S_h^*)(y_{\alpha,h} - z)\|_{L^2(D)} &\leq \delta_2(h), \\ \|(S^* - S_h^*)(y_{\alpha,h} - z)\|_{L^\infty(D)} &\leq \delta_\infty(h) \end{aligned} \tag{1.4}$$

*for all  $h > 0$  and  $\alpha \geq 0$ .*

Note that this assumption contains the unknown solutions of discretized optimal control problems. The functions  $\delta_2(h)$  and  $\delta_\infty(h)$  can be realized by means of a-posteriori error estimators, see e.g. [11]. In the analysis, we can as well allow for a-priori discretization error estimates, see the comments below in Remark 1.9.

Under this assumption one can prove discretization error estimates for  $P_\alpha$ .

**Proposition 1.7.** *Let Assumption 2 be satisfied. Let  $\alpha > 0$ . Then there is a constant  $c > 0$  independent of  $\alpha, h$  such that*

$$\begin{aligned} \|y_\alpha - y_{\alpha,h}\|_Y + \alpha^{\frac{1}{2}} \|u_\alpha - u_{\alpha,h}\|_{L^2(D)} &\leq c \left(1 + \alpha^{-\frac{1}{2}}\right) \delta_2(h), \\ \|p_\alpha - p_{\alpha,h}\|_{L^\infty(D)} &\leq c \left(\delta_\infty(h) + \left(1 + \alpha^{-\frac{1}{2}}\right) \delta_2(h)\right). \end{aligned}$$

holds for all  $h > 0$ .

*Proof.* For the proof, we refer to [5, Theorem 2.4] and [11, Proposition 1.8].  $\square$

Obviously these error estimates are not robust with respect to  $\alpha \searrow 0$ . In the case  $\alpha = 0$  we have an independent discretization error estimate of [3], which relies on Assumption 1.

**Proposition 1.8.** *Let Assumptions 1 and 2 be satisfied. Let  $d$  be as in Proposition 1.4. Then for every  $h_{\max} > 0$  there is a constant  $c > 0$  such that*

$$\begin{aligned} \|y_0 - y_{0,h}\|_Y &\leq c \left(\delta_2(h) + \delta_\infty(h)^d\right), \\ \|p_0 - p_{0,h}\|_{L^\infty(D)} &\leq c \left(\delta_2(h) + \delta_\infty(h)^{\min(d,1)}\right), \\ \|u_0 - u_{0,h}\|_{L^1(D)} &\leq c \left(\delta_2(h)^\kappa + \delta_\infty(h)^{\kappa \min(d,1)}\right) \end{aligned}$$

holds for all  $h < h_{\max}$ . The constant  $c$  depends on  $h_{\max}$ .

*Proof.* For the proof, we refer to [3, Theorem 2.2] and [11, Proposition 1.9].  $\square$

**Remark 1.9.** *As already mentioned, the estimates above are also valid if Assumption 2 on the discretization error is replaced by an a-priori variant. To this end, let us assume that there exist continuous and monotonically increasing functions  $\delta'_2(h), \delta'_\infty(h) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  with  $\delta'_2(0) = \delta'_\infty(0) = 0$  such that it holds*

$$\begin{aligned} \|(S - S_h)u_\alpha\|_Y + \|(S^* - S_h^*)(y_\alpha - z)\|_{L^2(D)} &\leq \delta'_2(h), \\ \|(S^* - S_h^*)(y_\alpha - z)\|_{L^\infty(D)} &\leq \delta'_\infty(h) \end{aligned} \tag{1.5}$$

for all  $h > 0$  and  $\alpha \geq 0$ . Here, the error estimate depends still on unknown solutions of  $(P_\alpha)$ . However, it is standard to prove uniform norm bounds on  $y_\alpha$  and  $u_\alpha$ , which then can be used to apply standard (e.g. finite element) a-priori error estimates to deduce the existence of  $\delta'_2$  and  $\delta'_\infty$ , see [3, 5] for the related analysis with  $S$  being the solution operator of an elliptic equation.

According to the discussion in [11], the results of Propositions 1.7 and 1.8 are valid if  $\delta_2$  and  $\delta_\infty$  are replaced by  $\delta'_2$  and  $\delta'_\infty$ , respectively.

## 2 Robust discretization error estimates

As one can see from the statement of Proposition 1.7, the error estimate is not robust with respect to  $\alpha \searrow 0$ , as the constant on the right-hand side behaves like  $\alpha^{-1/2}$  for  $\alpha \searrow 0$ . In particular, this means that the estimates of Proposition 1.8 cannot be obtained from those of Proposition 1.7. The purpose of this section is to prove an a-priori error estimate for  $\alpha > 0$  that is robust for  $\alpha \searrow 0$ .

**Lemma 2.1.** *Let Assumption 2 be satisfied. Let  $\alpha > 0$ . Then it holds*

$$\frac{1}{2}\|y_\alpha - y_{\alpha,h}\|_Y^2 + \alpha\|u_\alpha - u_{\alpha,h}\|_{L^2(D)}^2 \leq \frac{1}{2}\delta_2(h)^2 + \delta_\infty(h)\|u_{\alpha,h} - u_\alpha\|_{L^1(D)},$$

for all  $h > 0$  and  $\alpha > 0$ .

*Proof.* Since  $u_{\alpha,h}$  and  $u_\alpha$  are feasible for  $P_\alpha$  and  $P_{\alpha,h}$ , respectively, we can use both functions in the variational inequalities of Proposition 1.2. Adding the obtained inequalities yields

$$(\alpha u_\alpha + p_\alpha, u_{\alpha,h} - u_\alpha) + (\alpha u_{\alpha,h} + p_{\alpha,h}, u_\alpha - u_{\alpha,h}) \geq 0$$

This implies

$$\alpha\|u_\alpha - u_{\alpha,h}\|_{L^2(D)}^2 \leq (p_\alpha - p_{\alpha,h}, u_{\alpha,h} - u_\alpha). \quad (2.1)$$

Using the definitions of  $p_\alpha$  and  $p_{\alpha,h}$ , we obtain

$$\begin{aligned} (p_\alpha - p_{\alpha,h}, u_{\alpha,h} - u_\alpha) &= (S^*(y_\alpha - z) - S_h^*(y_{\alpha,h} - z), u_{\alpha,h} - u_\alpha) \\ &= (S^*(y_\alpha - y_{\alpha,h}) + (S - S_h^*)(y_{\alpha,h} - z), u_{\alpha,h} - u_\alpha). \end{aligned}$$

Here the second addend can be estimated as

$$|(S - S_h^*)(y_{\alpha,h} - z), u_{\alpha,h} - u_\alpha| \leq \delta_\infty(h)\|u_{\alpha,h} - u_\alpha\|_{L^1(D)}. \quad (2.2)$$

where we used Assumption 2. (If we would have estimated  $|(S - S_h^*)(y_{\alpha,h} - z), u_{\alpha,h} - u_\alpha| \leq \|(S - S_h^*)(y_{\alpha,h} - z)\|_{L^2(D)}\|u_{\alpha,h} - u_\alpha\|_{L^2(D)}$  instead, we would obtain the non-robust estimate of Proposition 1.7.) We continue with investigating the first addend in the above estimate

$$\begin{aligned} (S^*(y_\alpha - y_{\alpha,h}), u_{\alpha,h} - u_\alpha) &= (y_\alpha - y_{\alpha,h}, Su_{\alpha,h} - Su_\alpha)_Y \\ &= (y_\alpha - y_{\alpha,h}, (S - S_h)u_{\alpha,h} + S_h u_{\alpha,h} - Su_\alpha)_Y \\ &= -\|y_\alpha - y_{\alpha,h}\|_Y^2 + (y_\alpha - y_{\alpha,h}, (S - S_h)u_{\alpha,h})_Y \\ &\leq -\frac{1}{2}\|y_\alpha - y_{\alpha,h}\|_Y^2 + \frac{1}{2}\delta_2(h)^2, \end{aligned} \quad (2.3)$$

where we used Assumption 2 in the last step. Combining the estimates (2.1), (2.2), and (2.3) yields

$$\frac{1}{2}\|y_\alpha - y_{\alpha,h}\|_Y^2 + \alpha\|u_\alpha - u_{\alpha,h}\|_{L^2(D)}^2 \leq \frac{1}{2}\delta_2(h)^2 + \delta_\infty(h)\|u_{\alpha,h} - u_\alpha\|_{L^1(D)},$$

which is the claim.  $\square$

The  $L^1$ -error in the previous estimate can be bounded if the regularity assumption, i.e. Assumption 1, is fulfilled.

**Lemma 2.2.** *Let Assumption 1 be satisfied. Let  $\alpha > 0$ . Let  $v_\alpha, q_\alpha \in L^\infty(D)$  be given satisfying the projection formula*

$$v_\alpha(x) = \text{proj}_{[u_a(x), u_b(x)]} \left( -\frac{1}{\alpha} q_\alpha(x) \right) \quad \text{f.a.a. } x \in D.$$

*Then there is a constant  $c > 0$  independent of  $\alpha$  and  $(v_\alpha, q_\alpha)$  such that*

$$\begin{aligned} \|u_0 - v_\alpha\|_{L^1(D)} &\leq c(\alpha^\kappa + \|p_0 - q_\alpha\|_{L^\infty(D)}^\kappa) \\ \|u_0 - v_\alpha\|_{L^2(D)} &\leq c(\alpha^{\kappa/2} + \|p_0 - q_\alpha\|_{L^\infty(D)}^{\kappa/2}) \end{aligned}$$

*holds for all  $\alpha > 0$ .*

*Proof.* This follows from [13, Lemma 3.3]. □

Now we have everything at hand to derive the robust error estimate.

**Theorem 2.3.** *Let Assumptions 1 and 2 be satisfied. Then for every  $h_{\max} > 0$  and  $\alpha_{\max} > 0$  there is a constant  $c > 0$  such that*

$$\begin{aligned} \|y_\alpha - y_{\alpha,h}\|_Y &\leq c \left( \delta_2(h) + \delta_\infty(h)^d + \alpha^{d-1/2} \delta_\infty(h)^{1/2} \right), \\ \|p_\alpha - p_{\alpha,h}\|_{L^\infty(D)} &\leq c \left( \delta_2(h) + \delta_\infty(h)^{\min(d,1)} + \alpha^{d-1/2} \delta_\infty(h)^{1/2} \right), \\ \|u_\alpha - u_{\alpha,h}\|_{L^1(D)} &\leq c \left( \delta_2(h)^\kappa + \delta_\infty(h)^{\kappa \min(d,1)} + \alpha^{\kappa(d-1/2)} \delta_\infty(h)^{\kappa/2} + \alpha^{\kappa \min(1,d)} \right) \end{aligned}$$

*holds for all  $h < h_{\max}$  and  $\alpha \in (0, \alpha_{\max}]$ . Here,  $d$  is given by Proposition 1.4. The constant  $c$  depends on  $\alpha_{\max}$  and  $h_{\max}$ .*

*Proof.* Let  $h_{\max} > 0$  and  $\alpha_{\max} > 0$  be given, and let  $h < h_{\max}$  and  $\alpha \in (0, \alpha_{\max}]$  be arbitrary.

In order to invoke Lemma 2.1, we start with an estimate of  $\|u_{\alpha,h} - u_\alpha\|_{L^1(D)}$ . With the result of Lemma 2.2 we obtain

$$\begin{aligned} \|u_{\alpha,h} - u_\alpha\|_{L^1(D)} &\leq \|u_{\alpha,h} - u_0\|_{L^1(D)} + \|u_0 - u_\alpha\|_{L^1(D)} \\ &\leq c(\alpha^\kappa + \|p_0 - p_{\alpha,h}\|_{L^\infty(D)}^\kappa + \|p_0 - p_\alpha\|_{L^\infty(D)}^\kappa) \\ &\leq c(\alpha^\kappa + \|p_0 - p_\alpha\|_{L^\infty(D)}^\kappa + \|p_\alpha - p_{\alpha,h}\|_{L^\infty(D)}^\kappa) \end{aligned} \quad (2.4)$$

with constants  $c > 0$  independent of  $\alpha$  and  $h$ . Due to the regularization error estimate of Proposition 1.4, we have

$$\|p_0 - p_\alpha\|_{L^\infty(D)}^\kappa \leq c \alpha^{\kappa d}.$$

Here,  $c > 0$  depends on  $\alpha_{\max}$ . The discretization error  $\|p_\alpha - p_{\alpha,h}\|_{L^\infty(D)}^\kappa$  can be estimated by

$$\begin{aligned} \|p_\alpha - p_{\alpha,h}\|_{L^\infty(D)} &\leq \|p_\alpha - S^*(y_{\alpha,h} - z) + S^*(y_{\alpha,h} - z) - p_{\alpha,h}\|_{L^\infty(D)} \\ &\leq \|S^*(y_\alpha - y_{\alpha,h})\|_{L^\infty(D)} + \|(S^* - S_h^*)(y_{\alpha,h} - z)\|_{L^\infty(D)} \\ &\leq c(\|y_\alpha - y_{\alpha,h}\|_Y + \delta_\infty(h)). \end{aligned} \quad (2.5)$$

This proves

$$\begin{aligned} \|u_{\alpha,h} - u_\alpha\|_{L^1(D)} &\leq c(\alpha^{\kappa \min(1,d)} + \|y_\alpha - y_{\alpha,h}\|_Y^\kappa + \delta_\infty(h)^\kappa) \\ &= c(\alpha^{2d-1} + \|y_\alpha - y_{\alpha,h}\|_Y^\kappa + \delta_\infty(h)^\kappa) \end{aligned} \quad (2.6)$$

where in the last step we used Lemma 1.5.

With the result of Lemma 2.1 we obtain

$$\begin{aligned} \frac{1}{2} \|y_\alpha - y_{\alpha,h}\|_Y^2 &\leq \frac{1}{2} \delta_2(h)^2 + \delta_\infty(h) \|u_{\alpha,h} - u_\alpha\|_{L^1(D)} \\ &\leq \frac{1}{2} \delta_2(h)^2 + c \delta_\infty(h) (\alpha^{2d-1} + \|y_\alpha - y_{\alpha,h}\|_Y^\kappa + \delta_\infty(h)^\kappa). \end{aligned} \quad (2.7)$$

Let us first consider the case  $\kappa < 2$ , in which the term  $\|y_\alpha - y_{\alpha,h}\|_Y^\kappa$  can be absorbed by the left-hand side. By Young's inequality we find

$$c \delta_\infty(h) \|y_\alpha - y_{\alpha,h}\|_Y^\kappa \leq \frac{1}{4} \|y_\alpha - y_{\alpha,h}\|_Y^2 + c' \delta_\infty(h)^{\frac{2}{2-\kappa}}.$$

This implies

$$\begin{aligned} \frac{1}{4} \|y_\alpha - y_{\alpha,h}\|_Y^2 &\leq c(\delta_2(h)^2 + \alpha^{2d-1} \delta_\infty(h) + \delta_\infty(h)^{\frac{2}{2-\kappa}} + \delta_\infty(h)^{\kappa+1}) \\ &\leq c(\delta_2(h)^2 + \alpha^{2d-1} \delta_\infty(h) + \delta_\infty(h)^{2d}) \end{aligned}$$

with  $c > 0$  depending additionally on  $h_{\max}$ . Using (2.5) and (2.6) to estimate  $\|p_\alpha - p_{\alpha,h}\|_{L^\infty(D)}$  and  $\|u_{\alpha,h} - u_\alpha\|_{L^1(D)}$ , respectively, yields the claim in the case  $\kappa < 2$ .

Let us now prove the claim for the case  $\kappa \geq 2$ . Here, it is not possible to absorb  $\|y_\alpha - y_{\alpha,h}\|_Y^\kappa$ . We will derive a rough estimate of this quantity first. Due to the control constraints, the term  $\|u_0 - u_\alpha\|_{L^1(D)}$  is uniformly bounded. Hence, we obtain by Lemma 2.1

$$\|y_\alpha - y_{\alpha,h}\|_Y^2 \leq c(\delta_2(h)^2 + \delta_\infty(h)).$$

Using this upper bound of  $\|y_\alpha - y_{\alpha,h}\|_Y$  in (2.7) yields

$$\frac{1}{2} \|y_\alpha - y_{\alpha,h}\|_Y^2 \leq \frac{1}{2} \delta_2(h)^2 + c \delta_\infty(h) (\alpha^{2d-1} + \delta_2(h)^{2\kappa} + \delta_\infty(h)^\kappa).$$

By Young's inequality we obtain

$$\delta_\infty(h) \delta_2(h)^{2\kappa} \leq c(\delta_\infty(h)^{\kappa+1} + \delta_2(h)^{2\kappa \cdot \frac{\kappa+1}{\kappa}}) = c(\delta_\infty(h)^{\kappa+1} + \delta_2(h)^{2(\kappa+1)}).$$

Since  $\kappa + 1 = 2d > 1$ , this proves

$$\|y_\alpha - y_{\alpha,h}\|_Y^2 \leq c(\delta_2(h)^2 + \delta_\infty(h) \alpha^{2d-1} + \delta_\infty(h)^{\kappa+1}),$$

with  $c > 0$  depending additionally on  $h_{\max}$ . The estimates of  $\|p_\alpha - p_{\alpha,h}\|_{L^\infty(D)}$  and  $\|u_{\alpha,h} - u_\alpha\|_{L^1(D)}$  follow now from (2.5) and (2.6), respectively.  $\square$

Let us compare the results of this theorem to the results of Proposition 1.7 and 1.8. Clearly, the convergence rates of Theorem 2.3 with respect to the discretization quantities are smaller than the rates given by Proposition 1.7 in

the case  $\alpha > 0$ . But the estimates of Theorem 2.3 are not only robust with respect to  $\alpha \searrow 0$  but also optimal in the case  $\alpha = 0$  as they coincide with the rates given by Proposition 1.8. Here, it would be interesting to search for results that combine both advantages: namely, provide convergence rates that are similar to Proposition 1.7 in the case  $\alpha > 0$  and that are on the same time robust with respect to  $\alpha \searrow 0$  and yield the convergence rate of Proposition 1.8 in the case  $\alpha = 0$ .

### 3 A-priori regularization parameter choice

The results of the previous section give rise to a-priori parameter choice rules  $\alpha(h)$ , where  $\alpha$  is chosen depending on  $\delta_\infty(h)$ . Here it is important to ensure that the additional error introduced by the regularization is of the same order as the discretization error. It is favorable to obtain error estimates that are of the same order as the ones that are available for  $\alpha = 0$ , see Proposition 1.8 above.

**Theorem 3.1.** *Let Assumptions 1 and 2 be satisfied. Let  $\alpha$  be chosen such that*

$$\alpha(h) = \delta_\infty(h).$$

*Then for every  $h_{max} > 0$  there is  $c > 0$  such that*

$$\begin{aligned} \|y_0 - y_{\alpha(h),h}\|_Y &\leq c (\delta_2(h) + \delta_\infty(h)^d), \\ \|p_0 - p_{\alpha(h),h}\|_{L^\infty(D)} &\leq c (\delta_2(h) + \delta_\infty(h)^{\min(d,1)}), \\ \|u_0 - u_{\alpha(h),h}\|_{L^1(D)} &\leq c (\delta_2(h)^\kappa + \delta_\infty(h)^\kappa \min(d,1)) \end{aligned}$$

*holds for all  $h < h_{max}$ , where  $c$  depends on  $h_{max}$  but is independent of  $h$ .*

*Proof.* Let us first investigate the error  $\|y_0 - y_{\alpha(h),h}\|_Y$ . We have

$$\|y_0 - y_{\alpha(h),h}\|_Y \leq \|y_0 - y_{\alpha(h)}\|_Y + \|y_{\alpha(h)} - y_{\alpha(h),h}\|_Y.$$

The first addend can be estimated by Proposition 1.4 using the assumption

$$\|y_0 - y_{\alpha(h)}\|_Y \leq c \alpha(h)^d \leq c \delta_\infty(h)^d.$$

Applying Theorem 2.3 we can bound the second term from above as

$$\begin{aligned} \|y_{\alpha(h)} - y_{\alpha(h),h}\|_Y &\leq c (\delta_2(h) + \delta_\infty(h)^d + \alpha^{d-1/2} \delta_\infty(h)^{1/2}) \\ &\leq c (\delta_2(h) + \delta_\infty(h)^d). \end{aligned}$$

This implies the claimed estimate

$$\|y_0 - y_{\alpha(h),h}\|_Y \leq c (\delta_2(h) + \delta_\infty(h)^d).$$

Similarly, we can prove

$$\|p_0 - p_{\alpha(h),h}\|_{L^\infty(D)} \leq c (\delta_2(h) + \delta_\infty(h)^{\min(1,d)})$$

and the estimate of  $\|u_\alpha - u_{\alpha,h}\|_{L^1(D)}$ . □

This result proves convergence rates with respect to the discretization while still allowing for some regularization. The obtained convergence rate with respect to  $\delta_2$  and  $\delta_\infty$  is optimal in the following sense: We prove the same convergence rates as in the case  $\alpha = 0$ .

The surprising fact about this result is that the optimal parameter choice  $\alpha(h) = \delta_\infty(h)$  is independent of the unknown parameter  $\kappa$  in Assumption 1, which played a key role in all the analysis above. Hence, this result is perfectly suited to be used in adaptive computations that both are adaptive in the discretization as well as in the regularization.

Moreover, the theorem above yields the optimal convergence rate for all  $\kappa > 0$ , whereas the discrepancy-principle-based parameter choice rule of the previous work [11] only yields optimal rates for  $\kappa \leq 1$ . Additionally, the result gives a theoretical explanation for the numerical results in [11]. There, the discrepancy principle selects  $\alpha(h) \sim h^2$ , which is (up to logarithmic terms) the underlying convergence rate of the  $L^\infty$ -error for the problem considered there.

Let us mention that an a-priori choice of the regularization parameter based on the non-robust estimate of Proposition 1.7 does *not* yield optimal convergence rates. With the estimates of Proposition 1.4 and 1.7 we have

$$\begin{aligned} \|y_0 - y_{\alpha,h}\|_Y &\leq \|y_0 - y_\alpha\|_Y + \|y_\alpha - y_{\alpha,h}\|_Y \\ &\leq c \left( \alpha^d + \alpha^{-1/2} \delta_2(h) + \delta_2(h) \right). \end{aligned}$$

In order to balance the regularization and discretization error, one would have to choose  $\hat{\alpha}(h) \sim \delta_2(h)^{\frac{2}{2d+1}}$ , which would lead to

$$\|y_0 - y_{\hat{\alpha}(h),h}\|_Y \leq c \left( \delta_2(h)^{\frac{2d}{2d+1}} + \delta_2(h) \right).$$

In the case  $\kappa = 1$ , we would get  $\|y_0 - y_{\hat{\alpha}(h),h}\|_Y \leq c \delta_2(h)^{2/3}$ . The resulting convergence rate thus is much lower than the optimal one obtained for the choice  $\alpha \sim \delta_\infty(h)$ .

**Remark 3.2.** *The same results can be obtained if we want to use a-priori type discretization error estimates as in Remark 1.9 above. The results of Theorems 2.3 and 3.1 remain valid if  $\delta_2$  and  $\delta_\infty$  are replaced by  $\delta_2'$  and  $\delta_\infty'$ , respectively.*

## 4 Relation to parameter choice rules

In the previous work [11] the following parameter choice rule was studied. There  $\alpha$  was chosen by

$$\tilde{\alpha}(h) := \sup\{\alpha > 0 : I_{\alpha,h} \leq \delta_2(h)^2 + \delta_\infty(h)^2\}. \quad (4.1)$$

Here, the discrepancy measure  $I_{\alpha,h}$  was defined by

$$I_{\alpha,h} := \int_{\{x: p_{\alpha,h} > 0\}} (u_{\alpha,h} - u_a) p_{\alpha,h} \, d\mu + \int_{\{x: p_{\alpha,h} < 0\}} (u_{\alpha,h} - u_b) p_{\alpha,h} \, d\mu.$$

This choice was motivated by the estimate

$$\|y_{0,h} - y_{\alpha,h}\|_Y^2 \leq (u_{\alpha,h} - u_{0,h}, p_{\alpha,h})_{L^2(D)} \leq I_{\alpha,h},$$

see [11, Lemma 2.1]. Let us now study, whether the choice  $\alpha(h) = \delta_\infty(h)$  satisfies the relation

$$I_{\alpha(h),h} \leq \delta_2(h)^2 + \delta_\infty(h)^2,$$

which would imply  $\tilde{\alpha}(h) \geq \alpha(h)$ . In order to show this, let us cite the following estimate of  $I_{\alpha,h}$

$$I_{\alpha,h} \leq c \alpha(\|p_0 - p_{\alpha,h}\|_{L^\infty(D)}^\kappa + \alpha^\kappa), \quad (4.2)$$

which is taken from [11, (2.3)].

**Theorem 4.1.** *Let Assumptions 1 and 2 be satisfied. Let  $\alpha$  be chosen such that*

$$\alpha(h) = \delta_\infty(h).$$

*Then for every  $h_{max} > 0$  there is  $c > 0$  such that*

$$I_{\alpha(h),h} \leq c(\delta_\infty(h)^{2d} + \delta_2(h)^{2d})$$

*holds for all  $h < h_{max}$ . Hence, if  $\kappa > 1$  then there is  $h_0 > 0$  such that*

$$I_{\alpha(h),h} \leq c(\delta_2(h)^2 + \delta_\infty(h)^2)$$

*holds for all  $h < h_0$ . The constant  $c$  depends on  $h_{max}$  but is independent of  $h$ .*

*Proof.* Using estimate (4.2) and Theorem 3.1 we find

$$\begin{aligned} I_{\alpha(h),h} &\leq c \alpha(h)(\|p_0 - p_{\alpha(h),h}\|_{L^\infty(D)}^\kappa + \alpha^\kappa) \\ &\leq c \delta_\infty(h)(\delta_2(h)^\kappa + \delta_\infty(h)^{\kappa \min(1,d)} + \delta_\infty(h)^\kappa). \end{aligned}$$

By Young's inequality we obtain using  $2d \leq \kappa + 1$

$$\delta_\infty(h)\delta_2(h)^\kappa \leq c(\delta_\infty(h)^{\kappa+1} + \delta_2(h)^{\kappa \frac{\kappa+1}{\kappa}}) \leq c(\delta_\infty(h)^{2d} + \delta_2(h)^{2d}).$$

With the help of Lemma 1.5 we find

$$I_{\alpha(h),h} \leq c(\delta_\infty(h)^{2d} + \delta_2(h)^{2d}),$$

which proves the claim.  $\square$

This shows that the a-priori choice of the regularization parameter given by Theorem 2.3 satisfies the discrepancy principle (4.1) above in the case  $\kappa > 1$ . In the case  $\kappa = 1$  (and hence  $d = 1$ ) one can prove a similar result, if one replaces (4.1) with

$$\tilde{\alpha}(h) := \sup\{\alpha > 0 : I_{\alpha,h} \leq \tau(\delta_2(h)^2 + \delta_\infty(h)^2)\},$$

where  $\tau > 0$  has to be sufficiently small.

The theorem above shows that for sufficiently small  $h$  the inequality  $\alpha(h) \leq \tilde{\alpha}(h)$  is satisfied. Here one would like to prove the reverse estimate. Such an estimate seems not to be available, as it would require to work with estimates of the discrepancy measure  $I_{\alpha,h}$  from below.

## 5 Application to an elliptic optimal control problem

Let us now report about numerical experiments to solve an elliptic optimal control problem. This problem is given by: Minimize

$$J(y, u) := \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}$$

subject to the elliptic equation

$$\begin{aligned} -\Delta y &= u + f && \text{in } \Omega \\ y &= 0 && \text{on } \partial\Omega \end{aligned}$$

and the control constraints

$$u_a \leq u \leq u_b \quad \text{a.e. on } \Omega.$$

Here,  $\Omega \subset \mathbb{R}^n$  is a given bounded, polygonal, and convex domain with boundary  $\partial\Omega$ . Moreover, we have  $y_d, f, u_a, u_b \in L^\infty(\Omega)$  with  $u_a \leq u_b$  a.e. on  $\Omega$ . The solution operator  $S$ ,  $S : u \mapsto (-\Delta)^{-1}u$ , satisfies the assumptions made at the beginning of the paper. We set  $z := y_d - Sf$ . Then the control problem can be rephrased equivalently in the form (P).

Let us now describe the discretization by finite elements. We use a family  $\mathcal{F} = \{\mathcal{T}_h\}_{h>0}$  of regular meshes  $\mathcal{T}_h$ , consisting of closed cells  $T \in \mathcal{T}_h$ . For  $T \in \mathcal{T}_h$  let us define  $h_T := \text{diam } T$ . We assume that there is a constant  $R$ , such that  $h_T \leq R R_T$  for all  $h > 0$ ,  $T \in \mathcal{T}_h$ , where  $R_T$  is the diameter of the largest ball contained in  $T$ .

The discrete space  $V_h$  is defined as

$$V_h := \{v \in H_0^1(\Omega) : v|_T \in P_1(T)\}.$$

The discrete operator  $S_h$  is defined as the solution mapping of the discrete weak formulation of the elliptic equation. That is,  $S_h u := y_h \in V_h$ , where  $y_h$  solves

$$\int_{\Omega} \nabla y_h \cdot \nabla v_h \, d\mu = \int_{\Omega} u v_h \, d\mu \quad \forall v_h \in V_h.$$

Please note, that  $S$  and  $S_h$  are self-adjoint,  $S = S^*$  and  $S_h = S_h^*$ .

We will report about experiments using a-priori or a-posteriori estimates to calculate  $\delta_\infty(h)$ .

### 5.1 A-priori error estimates

Let us assume in addition that the meshes  $\mathcal{T}_h$  are quasi-uniform. That is, there is a constant  $M > 1$  such that  $\max_{T \in \mathcal{T}_h} h_T \leq M \min_{T \in \mathcal{T}_h} h_T$  for all  $h > 0$ . In this case, we have the following optimal a-priori rates, which hold under additional regularity assumptions, see [3],

$$\|(S - S_h)f\|_{L^\infty(\Omega)} \leq c h^2 |\log h|^{r(n)}$$

with  $r(2) = 2$ ,  $r(3) = 11/4$ .

## 5.2 A-posteriori error estimates

For comparison, we also used a-posteriori error estimates to determine the discretization error. We used the reliable and efficient error estimator from [8], which is defined as follows. Let us set  $\eta_{y_{\alpha,h},\infty} := \max_{T \in \mathcal{T}_h} \eta_{T,y_{\alpha,h},\infty}$  and

$$\eta_{T,y_{\alpha,h},\infty} := |\log h_{\min}|^2 \left( h_T^2 \|\Delta y_{\alpha,h} + u_{\alpha,h} + f\|_{L^\infty(T)} + h_T \left\| \left[ \frac{\partial y_{\alpha,h}}{\partial n} \right] \right\|_{L^\infty(\partial T \setminus \Gamma)} \right),$$

where  $h_{\min} := \min_{T \in \mathcal{T}_h} h_T$ , and  $[v]_E$  denotes the jump of the quantity  $v$  across an edge  $E$ . Then it holds

$$\|(S - S_h)u_{\alpha,h}\|_{L^\infty(\Omega)} \leq c_1 \eta_{y_{\alpha,h},\infty}$$

with  $c > 0$  independent of  $h$  and  $\alpha$ . Analogously, the error estimator  $\eta_{p_{\alpha,h},\infty}$  for the error in the adjoint equation is defined.

## 5.3 Numerical experiments

Let us now report about the outcome of numerical experiments. The following data was used:

$$\begin{aligned} \Omega &= (0, 1)^2, \quad u_a = -1, \quad u_b = +1, \\ z(x_1, x_2) &= \sin(\pi x_1) \sin(\pi x_2) + \sin(2\pi x_1) \sin(2\pi x_2) \\ f(x_1, x_2) &= -\text{sign}(\sin(2\pi x_1) \sin(2\pi x_2)) + 2\pi^2 \sin(\pi x_1) \sin(\pi x_2). \end{aligned}$$

It is easy to check that (P) admits the following unique solution:

$$\begin{aligned} u_0(x_1, x_2) &= \text{sign}(\sin(2\pi x_1) \sin(2\pi x_2)) \\ y_0(x_1, x_2) &= \sin(\pi x_1) \sin(\pi x_2) \\ p_0(x_1, x_2) &= -\frac{1}{8\pi^2} \sin(2\pi x_1) \sin(2\pi x_2). \end{aligned}$$

In addition, it turns out that the regularity assumption is satisfied for all  $\kappa < 1$ . This implies by Proposition 1.4  $d = 1$ ,  $\|u_0 - u_\alpha\|_{L^1(D)} \leq c \alpha$ . Moreover, with the choice  $\alpha(h) = \delta_\infty(h)$ , we have by Theorem 3.1

$$\|u_0 - u_{\alpha(h),h}\|_{L^1(D)} \leq c(\delta_2(h) + \delta_\infty(h)).$$

Since  $\Omega$  is bounded by assumption, we can estimate the  $L^2$ -error against the  $L^\infty$ -error to obtain

$$\|u_0 - u_{\alpha(h),h}\|_{L^1(D)} \leq c \delta_\infty(h).$$

The initial mesh was obtained by dividing the domain  $\Omega$  into 32 triangles with mesh size  $h = 0.3536$ . This mesh is then successively refined using uniform refinement.

Let us first describe the different tests with regularization parameter choices. We emphasize that both choices of the regularization parameter do *not* need information about the value of  $\kappa$ . We comment on the results below in Section 5.4.

$h$	$\ u_0 - u_{\alpha(h),h}\ _{L^1(D)}$	$\ p_0 - p_{\alpha(h),h}\ _{L^\infty(D)}$	$\alpha(h)$
$3.5355 \cdot 10^{-1}$	$3.2840 \cdot 10^{-1}$	$1.0098 \cdot 10^{-2}$	$9.7656 \cdot 10^{-4}$
$1.7678 \cdot 10^{-1}$	$1.0145 \cdot 10^{-1}$	$3.4007 \cdot 10^{-3}$	$2.4414 \cdot 10^{-4}$
$8.8388 \cdot 10^{-2}$	$3.4515 \cdot 10^{-2}$	$8.7735 \cdot 10^{-4}$	$1.2207 \cdot 10^{-4}$
$4.4194 \cdot 10^{-2}$	$7.7750 \cdot 10^{-3}$	$2.1340 \cdot 10^{-4}$	$1.5259 \cdot 10^{-5}$
$2.2097 \cdot 10^{-2}$	$1.8164 \cdot 10^{-3}$	$5.1677 \cdot 10^{-5}$	$3.8147 \cdot 10^{-6}$
$1.1049 \cdot 10^{-2}$	$4.4220 \cdot 10^{-4}$	$1.2606 \cdot 10^{-5}$	$9.5367 \cdot 10^{-7}$
	$\sim h^2$	$\sim h^2$	$\sim h^2$

Table 1: Choice by discrepancy principle, a-posteriori error estimate

$h$	$\ u_0 - u_{\alpha(h),h}\ _{L^1(D)}$	$\ p_0 - p_{\alpha(h),h}\ _{L^\infty(D)}$	$\alpha(h)$
$3.5355 \cdot 10^{-1}$	$5.8409 \cdot 10^{-1}$	$1.0281 \cdot 10^{-2}$	$6.2500 \cdot 10^{-3}$
$1.7678 \cdot 10^{-1}$	$1.7616 \cdot 10^{-1}$	$3.4236 \cdot 10^{-3}$	$1.5625 \cdot 10^{-3}$
$8.8388 \cdot 10^{-2}$	$5.2228 \cdot 10^{-2}$	$8.8031 \cdot 10^{-4}$	$3.9063 \cdot 10^{-4}$
$4.4194 \cdot 10^{-2}$	$1.4560 \cdot 10^{-2}$	$2.1445 \cdot 10^{-4}$	$9.7656 \cdot 10^{-5}$
$2.2097 \cdot 10^{-2}$	$3.4108 \cdot 10^{-3}$	$5.2005 \cdot 10^{-5}$	$2.4414 \cdot 10^{-5}$
$1.1049 \cdot 10^{-2}$	$8.0186 \cdot 10^{-4}$	$1.2699 \cdot 10^{-5}$	$6.1035 \cdot 10^{-6}$
	$\sim h^2$	$\sim h^2$	$\sim h^2$

Table 2: A-priori choice of regularization parameter, a-priori error estimates

### 5.3.1 Choice by discrepancy principle, a-posteriori error estimate

First, let us report about the outcome of the computations if the regularization parameter is chosen by the discrepancy principle (4.1). The costly computation of the supremum in (4.1) was replaced by the following principle

$$\tilde{\alpha}(h) := \sup\{\alpha : \alpha = 2^{-j}\alpha_0, j \in \mathbb{N}, I_{\alpha,h} \leq \tau(\eta_{y_{\alpha,h},\infty} + \eta_{p_{\alpha,h},\infty})\}$$

with  $\tau = 2 \cdot 10^{-4}$  and  $\alpha_0 = 10^{-4}$ . If  $z \neq 0$ , then the supremum exists, see [11, Lemma 3.1]. The results of the computation can be found in Table 1.

### 5.3.2 A-priori choice of regularization parameter, a-priori error estimates

Second, we applied the following strategy. According to Theorem 3.1, we choose  $\alpha(h) \sim \delta_\infty(h)$ . Here, we used  $\delta_\infty(h) = h^2$ , ignoring the logarithmic term, cf. Section 5.1. The results of the computation can be found in Table 2.

### 5.3.3 A-priori choice of regularization parameter, a-posteriori error estimates

The third strategy we investigated was the following. Again  $\alpha$  was chosen proportionally to the  $L^\infty$ -error. Here, we used the a-posteriori estimates of Section 5.2. Then the regularization parameter  $\alpha(h)$  was chosen iteratively: For

$h$	$\ u_0 - u_{\alpha(h),h}\ _{L^1(D)}$	$\ p_0 - p_{\alpha(h),h}\ _{L^\infty(D)}$	$\alpha(h)$
$3.5355 \cdot 10^{-1}$	$4.8075 \cdot 10^{-1}$	$1.0177 \cdot 10^{-2}$	$3.6464 \cdot 10^{-3}$
$1.7678 \cdot 10^{-1}$	$1.6251 \cdot 10^{-1}$	$3.4200 \cdot 10^{-3}$	$1.2837 \cdot 10^{-3}$
$8.8388 \cdot 10^{-2}$	$4.8981 \cdot 10^{-2}$	$8.7957 \cdot 10^{-4}$	$3.3204 \cdot 10^{-4}$
$4.4194 \cdot 10^{-2}$	$1.3059 \cdot 10^{-2}$	$2.1395 \cdot 10^{-4}$	$8.3717 \cdot 10^{-5}$
$2.2097 \cdot 10^{-2}$	$3.1173 \cdot 10^{-3}$	$5.1785 \cdot 10^{-5}$	$2.0964 \cdot 10^{-5}$
$1.1049 \cdot 10^{-2}$	$7.6073 \cdot 10^{-4}$	$1.2645 \cdot 10^{-5}$	$5.2427 \cdot 10^{-6}$
	$\sim h^2$	$\sim h^2$	$\sim h^2$

Table 3: A-priori choice of regularization parameter, a-posteriori error estimates

a fixed mesh, we solved the regularized discretized problem. If  $\alpha \leq \eta_{y_{\alpha,h},\infty} + \eta_{p_{\alpha,h},\infty}$  was satisfied, we set  $\alpha(h) := \alpha$  and accepted the discrete solution. If  $\alpha$  was larger than  $\eta_{y_{\alpha,h},\infty} + \eta_{p_{\alpha,h},\infty}$ , we set  $\alpha := \frac{1}{2} (\eta_{y_{\alpha,h},\infty} + \eta_{p_{\alpha,h},\infty})$  and solved the discrete problem again. The results of the computation can be found in Table 3.

## 5.4 Comments on the computational results

All three tests showed similar performances. Although the choices of the regularization parameter were conceptually different, the outcome is very similar. All the methods choose

$$\alpha(h) \sim h^2.$$

The resulting errors in control, state, and adjoint behaved like

$$\|u_0 - u_{\alpha(h),h}\|_{L^1(D)}, \|y_0 - y_{\alpha(h),h}\|_{L^2(D)}, \|p_0 - p_{\alpha(h),h}\|_{L^\infty(D)} \sim h^2.$$

Moreover, the results using the optimistic a-priori rate  $\delta_\infty(h) = h^2$  and the results using the a-posteriori estimator agree to large extend. Hence, the choice of the a-priori rate is justified.

These results show that both parameter choices - a-priori as well as a-posteriori - lead to comparable results.

## 6 Conclusion

In this paper, we presented a robust error estimate for the optimization problem under consideration. This robust estimate lends itself to an a-priori regularization parameter choice. The numerical results confirm these findings, results in optimal convergence rates with respect to the discretization parameter. Moreover, they show that this a-priori choice leads to the same accuracy as the a-posteriori choice based on a discrepancy principle.

## References

- [1] E. Casas. Second order analysis for bang-bang control problems of PDEs. *SIAM J. Control Optim.*, 50(4):2355–2372, 2012.

- [2] C. Clason and K. Kunisch. Multi-bang control of elliptic systems. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 2013.
- [3] K. Deckelnick and M. Hinze. A note on the approximation of elliptic control problems with bang-bang controls. *Comput. Optim. Appl.*, 51(2):931–939, 2012.
- [4] U. Felgenhauer. On stability of bang-bang type controls. *SIAM J. Control Optim.*, 41(6):1843–1867, 2003.
- [5] M. Hinze. A variational discretization concept in control constrained optimization: the linear-quadratic case. *J. Computational Optimization and Applications*, 30:45–63, 2005.
- [6] C. Meyer and A. Rösch. Superconvergence properties of optimal control problems. *SIAM J. Control Optim.*, 43(3):970–985, 2004.
- [7] A. Neubauer. Tikhonov-regularization of ill-posed linear operator equations on closed convex sets. *J. Approx. Theory*, 53(3):304–320, 1988.
- [8] R. H. Nochetto, A. Schmidt, K. G. Siebert, and A. Veerer. Pointwise a posteriori error estimates for monotone semi-linear equations. *Numer. Math.*, 104(4):515–538, 2006.
- [9] F. Tröltzsch. *Optimal control of partial differential equations*, volume 112 of *Graduate Studies in Mathematics*. AMS, Providence, RI, 2010.
- [10] M. Ulbrich and S. Ulbrich. Superlinear convergence of affine-scaling interior-point Newton methods for infinite-dimensional nonlinear problems with pointwise bounds. *SIAM J. Control Optim.*, 38(6):1938–1984, 2000.
- [11] D. Wachsmuth. Adaptive regularization and discretization of bang-bang optimal control problems. *ETNA*, 40:249–267, 2013.
- [12] G. Wachsmuth and D. Wachsmuth. Convergence and regularization results for optimal control problems with sparsity functional. *ESAIM Control Optim. Calc. Var.*, 17(3):858–886, 2011.
- [13] G. Wachsmuth and D. Wachsmuth. On the regularization of optimization problems with inequality constraints. *Control and Cybernetics*, 4:1125–1154, 2011.