Electronic Transactions on Numerical Analysis.
Volume 40, pp. 249-267, 2013.

ISSN 1068-9613.

ETNA
Kent State University
http://etna.math.kent.edu

ADAPTIVE REGULARIZATION AND DISCRETIZATION OF BANG-BANG OPTIMAL CONTROL PROBLEMS*

DANIEL WACHSMUTH†

Abstract. In this article, Tikhonov regularization of control-constrained optimal control problems is investigated. Typically the solutions of such problems exhibit a so-called bang-bang structure. We develop a parameter choice rule that adaptively selects the Tikhonov regularization parameter depending on a posteriori computable quantities. We prove that this choice leads to optimal convergence rates with respect to the discretization parameter. The article is complemented by numerical results.

Key words. optimal control, bang-bang control, Tikhonov regularization, parameter choice rule

AMS subject classifications. 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

$$\frac{1}{2}\|Su - z\|_Y^2 \tag{1.1}$$

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

$$u_a \le u \le u_b \quad \text{a.e. on } D. \tag{1.2}$$

Here, 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 is a map from Y to $L^\infty(D)$.

Problems covered by this general framework include distributed or boundary control problems subject to elliptic and parabolic equations if one relaxes the requirements on S^* to map into $L^p(D)$, $p > 2$. Due to the appearance of the inequality constraints (1.2), a solution u_0 to the problem (1.1) often exhibits the so-called bang-bang structure, that is, $u_0(x) \in \{u_a(x), u_b(x)\}$ for almost all $x \in D$. Hence, the control constraints are active almost everywhere on D.

It is well-known that the problem (1.1) is solvable with a unique solution if S is injective. However, solutions of (1.1) are unstable with respect to perturbations in the problem data in general. Hence, a regularization of (1.1) is desirable that stabilizes the solutions of the problem while maintaining a certain accuracy of the discretization method. In this paper, we will study Tikhonov regularization methods, which are widely used in optimal control problems as well as in inverse problems; see, e.g., Engl, Hanke, Neubauer [7], Tröltzsch [22].

In order to solve (1.1) numerically, let us introduce a family of linear and continuous operators $\{S_h\}_{h>0}$ from $L^2(D)$ to Y with finite-dimensional ranges $Y_h \subset Y$, where h denotes the discretization parameter. The regularized and discretized problem now reads: minimize

$$\frac{1}{2}\|S_h u - z\|_Y^2 + \frac{\alpha}{2}\|u\|_{L^2(D)}^2 \tag{1.3}$$

*Received October 1, 2012. Accepted April 5, 2013. Published online on August 5, 2013. Recommended by R. Ramlau. This work was partially funded by the Austrian Science Fund (FWF) grant P23848.

†Institut für Mathematik, Universität Würzburg, 97074 Würzburg, Germany
(daniel.wachsmuth@mathematik.uni-wuerzburg.de).

ETNA
Kent State University
http://etna.math.kent.edu

subject to (1.2). This problem is solvable, where we will denote a solution by $u_{\alpha,h}$ in the sequel. In the case that S is a solution operator of an elliptic partial differential equation and Y_h is spanned by linear finite elements, then (1.3) is a variational discretization of (1.4) in the sense of Hinze [14]. Corresponding discretization error estimates can be found in [14]; the case $\alpha = 0$ is considered in Deckelnick, Hinze [6].

In order to develop an efficient approximation scheme to solve (1.1) by means of successively solving instances of (1.3), we are faced with two important questions:

1. Given a fixed discretization h, how should we choose $\alpha = \alpha(h)$?
2. Suppose we have solved (1.3) for fixed $(\alpha(h), h)$. How should we refine the discretization?

Of course, we want to have answers to both questions such that the resulting scheme is optimal. Here, we borrow the meaning of "optimal" from the results in linear inverse problems: choose the regularization parameter $\alpha(h)$ such that the regularization error is of the same size as the discretization error.

Let us briefly review the available literature. To the best of our knowledge, there are no results available concerning adaptive Tikhonov regularization and the discretization of optimal control problems with inequality constraints. Much more is known for linear inverse problems. Parameter choice rules and convergence results for a uniform discretization go back to Natterer [18] and Engl, Neubauer [8]. Parameter choice rules depending on a posteriori computable quantities can be found for instance in Groetsch, Neubauer [12] and King, Neubauer [16]. Adaptive wavelet based regularization was studied in Maaß, Pereverzev, Ramlau, and Solodky [17]. Adaptive discretization methods for parameter estimation problems based on residual error estimates can be found in Neubauer [19] and Ben Ameur, Chavent, Jaffré [3]. Adaptive refinement using goal-oriented estimators is investigated in Becker, Vexler [1, 2]. An adaptive regularization and discretization method is developed in Griesbaum, Kaltenbacher, Vexler [10] and Kaltenbacher, Kirchner, Vexler [15] for linear and non-linear inverse problems, respectively.

In this paper, we develop and analyze a parameter choice rule $\alpha = \alpha(h)$; see (2.2) below. There, the parameter $\alpha(h)$ can be determined solely by a posteriori available quantities. We will prove that the approximation error $\|u_0 - u_{\alpha(h),h}\|_{L^1(D)}$ is proportional to $\|u_0 - u_{0,h}\|_{L^1(D)}$, where u_0, $u_{0,h}$, $u_{\alpha(h),h}$ are the solutions of (1.1), (1.3) with $\alpha = 0$ and of (1.3) with $\alpha = \alpha(h)$, respectively. That is, the additional error introduced by the regularization is of the same size as the discretization error. The relevant estimates for the discretization error can be found in Proposition 1.9. The optimality of the parameter choice rule, which is the main result of this article, is shown in Theorem 2.7. Furthermore, we propose an adaptive procedure that involves both adaptive regularization and adaptive mesh-refinement; see Section 3.2.

In order to achieve these results, a certain regularity assumption, see Assumption 1.4, has to be fulfilled, which is an assumption on the solution of the continuous, undiscretized problem (1.1). This assumption guarantees that almost-active sets are small in some sense. Such a condition is used in Deckelnick, Hinze [6] to prove a priori error estimates for the discretized but unregularized problem. Moreover, this condition is used in Wachsmuth, Wachsmuth [24, 25] to prove a priori convergence of the regularization error $\|u_0 - u_\alpha\|_{L^1(D)}$ with respect to α; cf. Proposition 1.5.

1.1. Notation. In the sequel, we will use subscripts to indicate solutions to the optimization problems introduced above. That is, u_0, u_α, and $u_{\alpha,h}$ will denote solutions to (1.1), (1.4) (see below), and (1.3), respectively. Additionally, we will use $y_0 := Su_0$, $y_\alpha := Su_\alpha$, etc. We will work with generic constants $c > 0$, which may change from line to line, but which are independent of the relevant quantities such as α and h.

ETNA
Kent State University
http://etna.math.kent.edu

1.2. Assumptions and preliminary results. Let (D, Σ, μ) be a given 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 $L^\infty(D)$. By a result of Grothendieck [13], it follows that S has a closed range if and only if its range is finite-dimensional. This implies that the equation $Su = z$ is ill-posed, except for the trivial case that the range of S is finite-dimensional. Let us remark that the requirements on S^* could be relaxed to allow S^* mapping into $L^p(D)$, $p \in (2, \infty)$; see [25] and the comments after Proposition 1.5.

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

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

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 ranges. The adjoint operators S_h^* are assumed to map from the Hilbert space Y to $L^\infty(D)$. For later reference, let us introduce the regularized (undiscretized) problem: given $\alpha > 0$, minimize

$$\frac{1}{2}\|Su - z\|_Y^2 + \frac{\alpha}{2}\|u\|_{L^2(D)}^2 \tag{1.4}$$

subject to the inequality constraints (1.2). Regarding the existence of solutions of (1.1), (1.4), and (1.3), we have the following classical result.

PROPOSITION 1.1. *The problems* (1.1) *and* (1.3) *are solvable with convex and bounded sets of solutions. The problem* (1.4) *is uniquely solvable for $\alpha > 0$. The solutions of* (1.1) *are unique if S is injective. The solution of* (1.3) *is uniquely determined if $\alpha > 0$.*

While the optimal control of (1.1) and (1.3) may not be uniquely determined, the optimal states Su_0 and $Su_{0,h}$ are uniquely determined due to the strict convexity of the cost functional with respect to y.

Throughout the article, we assume the following properties of the discrete operators S_h and S_h^*.

ASSUMPTION 1.2. *There exist functions $\delta_2(h)$, $\delta_\infty(h) : \mathbb{R}^+ \to \mathbb{R}^+$, continuous and monotonically increasing with $\delta_2(0) = \delta_\infty(0) = 0$, such that it holds*

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

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

In this assumption, the convergence of the discretization depends on the approximation properties of *discrete solutions*. This form is especially useful in combination with a posteriori error estimators. We can replace this assumption with an assumption on the approximation properties of the solution of the continuous problem alone; see the remarks at the end of Section 1.5.

1.3. Necessary optimality conditions. Let us recall some standard results on the first-order necessary optimality conditions.

PROPOSITION 1.3. *For $\alpha \ge 0$, let u_α and $u_{\alpha,h}$ be solutions of* (1.4) *and* (1.3)*, respectively. Let us define $p_\alpha := S^*(y_\alpha - z)$ and $p_{\alpha,h} := S_h^*(y_{\alpha,h} - z)$. Then it holds that*

$$\begin{aligned} (\alpha u_\alpha + p_\alpha,\, u - u_\alpha) &\ge 0 \quad \forall u \in U_{ad}, \\ (\alpha u_{\alpha,h} + p_{\alpha,h},\, u - u_{\alpha,h}) &\ge 0 \quad \forall u \in U_{ad}. \end{aligned}$$

Using well-known arguments [22, Lemma 2.26], we have for almost all $x \in D$

$$u_\alpha(x) = \mathrm{proj}_{[u_a(x),u_b(x)]}\left(-\frac{1}{\alpha}p_\alpha(x)\right) \qquad \text{if } \alpha > 0$$

and

$$\begin{aligned} u_0(x) &= u_a(x) && \text{if } p_0(x) > 0, \\ u_0(x) &= u_b(x) && \text{if } p_0(x) < 0, \\ p_0(x) &= 0 && \text{if } u_a(x) < u_0(x) < u_b(x). \end{aligned} \tag{1.5}$$

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.

1.4. Regularization error estimate. We will now briefly investigate the convergence properties of u_α for $\alpha \to 0$. Since the problem (1.1) reduces to an ill-posed equation if the control constraints are not active at the solution, it is clear that the convergence $u_\alpha \to u_0$ cannot be achieved without any further conditions.

We will rely on the following assumption, which is an assumption about the measure of the set where the control constraints on u_0 are almost active.

ASSUMPTION 1.4. *Let us assume that there are $\kappa > 0$, $c > 0$ such that*

$$\mathrm{meas}\,\{x \in D : \, |p_0(x)| \le \epsilon\} \le c\,\epsilon^\kappa$$

for all $\epsilon > 0$.

This assumption implies that the set $\{x : \, p_0(x) = 0\}$ has measure zero, hence u_0 is a bang-bang control by the necessary optimality conditions (1.5). Since the adjoint state p_0 is uniquely determined, it is another consequence of this assumption that u_0 is the unique solution of (1.1).

Let us remark that such a condition is used in [6] to prove discretization error estimates for $\|u_0 - u_{0,h}\|_{L^1(D)}$ and in [9] to investigate stability properties of bang-bang solution to ODE control problems. Moreover, the assumption implies a local growth of the cost function with respect to the L^1-norm [24]. In connection with convergence rate estimates for interior point methods, this assumption was used as a strengthening of the strict complementarity conditions in [11, 23].

The main purpose of this assumption is to provide convergence rates for $\alpha \to 0$.

PROPOSITION 1.5. *Let Assumption 1.4 be satisfied. Let d be defined by*

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

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

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

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

For proofs, we refer to [24, 25]. Moreover, it was shown in [26] that Assumption 1.4 is necessary for the convergence rates $d > 1$, which corresponds to the case $\kappa > 1$.

ETNA
Kent State University
http://etna.math.kent.edu

Let us remark that $\kappa = 1$ is the maximal possible value for certain classes of problems. This is due to the fact that $\kappa > 1$ and $\kappa > 2$ imply $p_0 \notin C^1(\bar{D})$ and $p_0 \notin H^1(D)$, respectively. That said, if the range of the operator S^*S is a subspace of $C^1(\bar{D})$ or $H^1(D)$, then an upper bound on κ is provided by the smoothing properties of S and S^*.

REMARK 1.6. In the attainable case, which is the basis for many results in inverse problems, it holds that $p_0 = 0$. Thus, the assumption of bang-bang solutions is not valid in this case, and different techniques have to be used to prove a regularization error estimate; see [25].

REMARK 1.7. If S^* is a continuous operator mapping from Y to $L^p(D)$, $p \in (2, \infty)$, then the result of Proposition 1.9 still remains true with the modified convergence rate of $d = \min\left(\frac{1+\kappa}{2}, \frac{p+\kappa}{p(2-\kappa)+2\kappa}\right)$; see [25, Corollary 3.15].

1.5. Discretization error estimates. Let us first introduce the discretization of (1.1) by

$$\min_{u \in U_{ad}} \frac{1}{2}\|S_h u - z\|_Y^2. \tag{1.6}$$

We will now prove error estimates with respect to the discretization parameter h. To this end, let $u_{0,h}$ denote the solution of (1.6) with minimal L^2-norm.

PROPOSITION 1.8. *Let Assumption 1.2 be satisfied. Let $\alpha > 0$. Then there is a constant $c > 0$ independent of α, h such that*

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

holds for all $h > 0$.

Proof. The result is a consequence of the optimality condition in Proposition 1.3 as well as the assumptions on the discretization. Using the necessary optimality conditions we obtain

$$\begin{aligned}\alpha\|u_\alpha - u_{\alpha,h}\|_{L^2(D)}^2 &\le (p_{\alpha,h} - p_\alpha,\, u_\alpha - u_{\alpha,h})_{L^2(D)} \\ &= (S_h^*(y_{\alpha,h} - z) - S^*(y_\alpha - z),\, u_\alpha - u_{\alpha,h})_{L^2(D)} \\ &= \left((S_h^* - S^*)(y_{\alpha,h} - z) + S^*(y_{\alpha,h} - y_\alpha),\, u_\alpha - u_{\alpha,h}\right)_{L^2(D)}.\end{aligned} \tag{1.7}$$

We continue with

$$((S_h^* - S^*)(y_{\alpha,h} - z),\, u_\alpha - u_{\alpha,h})_{L^2(D)} \le \frac{\alpha}{2}\|u_\alpha - u_{\alpha,h}\|_{L^2(D)}^2 + \frac{1}{2\alpha}\delta_2(h)^2. \tag{1.8}$$

In addition, we estimate

$$\begin{aligned}&(S^*(y_{\alpha,h} - y_\alpha),\, u_\alpha - u_{\alpha,h})_{L^2(D)} \\ &\qquad = (y_{\alpha,h} - y_\alpha,\, y_\alpha - Su_{\alpha,h})_Y \\ &\qquad = -\|y_\alpha - y_{\alpha,h}\|_Y^2 + (y_{\alpha,h} - y_\alpha, (S_h - S)u_{\alpha,h})_Y \\ &\qquad \le -\frac{1}{2}\|y_\alpha - y_{\alpha,h}\|_Y^2 + \frac{1}{2}\,\delta_2(h)^2.\end{aligned} \tag{1.9}$$

The estimates (1.7)–(1.9) imply

$$\|y_\alpha - y_{\alpha,h}\|_Y^2 + \alpha\|u_\alpha - u_{\alpha,h}\|_{L^2(D)}^2 \le \left(1 + \alpha^{-1}\right)\delta_2(h)^2,$$

ETNA
Kent State University
http://etna.math.kent.edu

which proves the first claim. To obtain the second, observe that it holds that

$$
\begin{aligned}
\|p_\alpha - p_{\alpha,h}\|_{L^\infty(D)} &\le \|p_\alpha - S^*(y_{\alpha,h} - z) + S^*(y_{\alpha,h} - z) - p_{\alpha,h}\|_{L^\infty(D)} \\
&\le \|S^*(y_\alpha - y_{\alpha,h})\|_{L^\infty(D)} + \|(S^* - S_h^*)(y_{\alpha,h} - z)\|_{L^\infty(D)} \\
&\le c(\|y_\alpha - y_{\alpha,h}\|_Y + \delta_\infty(h)),
\end{aligned} \tag{1.10}
$$

and the estimate for the errors of the adjoint states is an immediate conclusion. ∎

As can be seen above, these error estimates are not robust with respect to $\alpha \to 0$. This is again due to the ill-posedness of the underlying equation $Su = z$. As demonstrated in [6], Assumption 1.4 is sufficient to prove convergence rates for $h \to 0$ in the case $\alpha = 0$.

PROPOSITION 1.9. *Let Assumptions 1.2 and 1.4 be satisfied. Let d be as in Proposition 1.5. Then for every $h_{max} > 0$, there is a constant $c > 0$ such that*

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

holds for all $h < h_{max}$.

Proof. The proof for the case $\kappa \le 1$ can be found in [6, Theorem 2.2]. We will briefly sketch the proof to indicate the necessary modifications for the case $\kappa > 1$. Using the estimates (1.7) and (1.9) with $\alpha = 0$ in the proof of Proposition 1.8, we obtain

$$
\frac{1}{2}\|y_0 - y_{0,h}\|_Y^2 \le c\,\delta_2(h)^2 + ((S_h^* - S^*)(y_{0,h} - z),\, u_0 - u_{0,h})_{L^2(D)}. \tag{1.11}
$$

The second term can be bounded as

$$
\begin{aligned}
((S_h^* - S^*)(y_{0,h} - z),\, u_0 - u_{0,h}) &\le \delta_\infty(h)\|u_0 - u_{0,h}\|_{L^1(D)} \\
&\le c\,\delta_\infty(h)\|p_0 - p_{0,h}\|_{L^\infty(D)}^\kappa,
\end{aligned} \tag{1.12}
$$

where the last estimate is a consequence of Assumption 1.4; see [6, (2.13)]. Using (1.10) to estimate $\|p_0 - p_{0,h}\|_{L^\infty(D)}^\kappa$, we find

$$
\|y_0 - y_{0,h}\|_Y^2 \le c\left(\delta_2(h)^2 + \delta_\infty(h)\|y_0 - y_{0,h}\|_Y^\kappa + \delta_\infty(h)^{\kappa+1}\right). \tag{1.13}
$$

If $\kappa < 2$, then the claim follows by an application of Young's inequality.

If $\kappa \ge 2$, we proceed as follows. As $\|u_0 - u_{0,h}\|_{L^1(D)}$ is uniformly bounded with respect to h by the control constraints, the inequality (1.11) implies the non-optimal error bound

$$
\frac{1}{2}\|y_0 - y_{0,h}\|_Y^2 \le c\,(\delta_2(h)^2 + \delta_\infty(h)).
$$

Substituting this in (1.13) yields

$$
\begin{aligned}
\|y_0 - y_{0,h}\|_Y^2 &\le c\left(\delta_2(h)^2 + \delta_\infty(h)\delta_2(h)^{2\kappa} + \delta_\infty(h)^{\kappa+1}\right) \\
&\le c\left(\delta_2(h)^2 + \delta_2(h)^{2(\kappa+1)} + \delta_\infty(h)^{\kappa+1}\right)
\end{aligned}
$$

by Young's inequality. The estimates of $\|p_0 - p_{0,h}\|_{L^\infty(D)}$ and $\|u_0 - u_{0,h}\|_{L^1(D)}$ can be derived as in (1.10) and (1.12), respectively. ∎

As can be seen from these results, the convergence rate of $\|p_0 - p_{0,h}\|_{L^\infty(D)}$ with respect to $\delta_\infty(h)$ saturates at 1, which is the rate given by Assumption 1.2.

ETNA
Kent State University
http://etna.math.kent.edu

Let us briefly sketch how to obtain similar results if Assumption 1.2 is replaced by the following, which is an assumption on the approximation properties of the solutions of (1.1) and (1.4).

ASSUMPTION 1.10. *There exist functions* $\delta_2'(h)$, $\delta_\infty'(h) : \mathbb{R}^+ \to \mathbb{R}^+$, *continuous and monotonically increasing with* $\delta_2'(0) = \delta_\infty'(0) = 0$, *such that it holds that*

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

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

COROLLARY 1.11. *Let Assumptions 1.4 and 1.10 be satisfied. Then the results of Propositions 1.8 and 1.9 are valid with* $\delta_2(h)$ *and* $\delta_\infty(h)$ *replaced by* $\delta_2'(h)$ *and* $\delta_\infty'(h)$.

Proof. Let us briefly outline the necessary modifications in the proofs of Propositions 1.8 and 1.9. The estimate (1.7) has to be replaced by

$$\begin{aligned} \alpha\|u_\alpha - u_{\alpha,h}\|^2_{L^2(D)} &\leq (S_h^*(y_{\alpha,h} - z) - S^*(y_\alpha - z),\, u_\alpha - u_{\alpha,h})_{L^2(D)} \\ &= ((S_h^* - S^*)(y_\alpha - z),\, u_\alpha - u_{\alpha,h})_{L^2(D)} - \|y_\alpha - y_{\alpha,h}\|_Y^2 \\ &\quad + (y_{\alpha,h} - y_\alpha,\, (S_h - S)u_\alpha)_Y. \end{aligned} \tag{1.14}$$

Now the estimates for $\|u_\alpha - u_{\alpha,h}\|_{L^2(D)}$ and $\|y_\alpha - y_{\alpha,h}\|_Y$ in the case $\alpha > 0$ are a straightforward consequence. The L^∞-error estimates for the adjoint states can be obtained from

$$\begin{aligned} \|p_\alpha - p_{\alpha,h}\|_{L^\infty(D)} &\leq \|p_\alpha - S_h^*(y_\alpha - z) + S_h^*(y_\alpha - z) - p_{\alpha,h}\|_{L^\infty(D)} \\ &\leq \|(S^* - S_h^*)(y_\alpha - z)\|_{L^\infty(D)} + \|S_h^*(y_\alpha - y_{\alpha,h})\|_{L^\infty(D)} \\ &\leq c(\delta_\infty'(h) + \|y_\alpha - y_{\alpha,h}\|_Y), \end{aligned} \tag{1.15}$$

which yields the claim for $\alpha > 0$. That is, the results analogous to Proposition 1.8 are proven.

Let us now prove the claimed estimate for $\alpha = 0$. Here, (1.14) with $\alpha = 0$ gives

$$\begin{aligned} \|y_0 - y_{0,h}\|_Y^2 &\leq c\left(\delta_\infty'(h)\|u_0 - u_{0,h}\|_{L^1(D)} + \delta_2'(h)^2\right) \\ &\leq c\left(\delta_\infty'(h)\|p_0 - p_{0,h}\|^\kappa_{L^\infty(D)} + \delta_2'(h)^2\right), \end{aligned}$$

where we applied Assumption 1.10 as well as $\|u_0 - u_{0,h}\|_{L^1(D)} \leq c\|p_0 - p_{0,h}\|^\kappa_{L^\infty(D)}$; see also the proof of Proposition 1.9. Using (1.15) we obtain

$$\|y_0 - y_{0,h}\|_Y^2 \leq c\left(\delta_\infty'(h)\|y_0 - y_{0,h}\|_Y^\kappa + \delta_\infty'(h)^{\kappa+1} + \delta_2'(h)^2\right).$$

Arguing as in the proof of Proposition 1.9, see the arguments following (1.13), completes the proof. □

2. Parameter choice rule depending on discretization errors. We describe our parameter choice rule $\alpha(h)$ that yields optimal convergence rates. Here, we mean by "optimal rate" that the error introduced by the regularization does not lead to convergence rates worse than the rates provided by Proposition 1.9. That is, for a given discretization parameter h, we want to choose $\alpha = \alpha(h)$ in such a way such that

$$\begin{aligned} \|y_0 - y_{\alpha(h),h}\|_Y &\approx \|y_0 - y_{0,h}\|_Y, \\ \|u_0 - u_{\alpha(h),h}\|_{L^1(D)} &\approx \|u_0 - u_{0,h}\|_{L^1(D)}. \end{aligned}$$

This is related to the classical motivation of the discrepancy principle in inverse problems: if only noisy data z^δ are available, then it suffices to solve the ill-posed equation $Sx = z^\delta$ up to an accuracy that corresponds to the noise level: choose $\alpha(\delta)$ such that $\|Sx_{\alpha(\delta)} - z^\delta\| \approx \delta$.

ETNA
Kent State University
http://etna.math.kent.edu

Similar results about a discrepancy principle for inequality constrained problems can be found in [25]. However, these results are not applicable here as they deal with perturbations of the data, $z^\delta \approx z$, while the problem under consideration involves a perturbation of the operator $S_h \approx S$.

Another challenge when devising the parameter choice rule is the fact that the parameter κ is not known a priori. This means that also the optimal convergence rate in Proposition 1.9 is unknown. Throughout this section, we assume that Assumptions 1.2 and 1.4 are satisfied. As we will show, the parameter choice rule will not rely on the actual value of κ.

Let us start with the observation that it is sufficient to choose $\alpha(h)$ such that

$$\begin{aligned} \|y_{0,h} - y_{\alpha(h),h}\|_Y &\approx \|y_0 - y_{0,h}\|_Y, \\ \|u_{0,h} - u_{\alpha(h),h}\|_{L^1(D)} &\approx \|u_0 - u_{0,h}\|_{L^1(D)}. \end{aligned} \tag{2.1}$$

Here, the term $\|y_{0,h} - y_{\alpha(h),h}\|_Y$ can be estimated numerically without knowing the solution $(y_{0,h}, u_{0,h})$ of (1.6).

LEMMA 2.1. *Let* $(u_{0,h}, y_{0,h}, p_{0,h})$ *and* $(u_{\alpha,h}, y_{\alpha,h}, p_{\alpha,h})$ *be solutions of the discretized unregularized problem* (1.6) *and regularized problem* (1.3) *for some* $\alpha > 0$*, respectively. Then it holds that*

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

with $I_{\alpha,h}$ *defined as*

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

Proof. The necessary optimality conditions for (1.6) imply

$$(p_{0,h} - p_{\alpha,h} + p_{\alpha,h},\, u_{\alpha,h} - u_{0,h})_{L^2(D)} \ge 0.$$

Since $p_{0,h} - p_{\alpha,h} = S_h(y_{0,h} - y_{\alpha,h})$, we obtain by transposition

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

which proves the first claim. Replacing $u_{0,h}$ by the suitable control bounds, the upper bound involving $I_{\alpha,h}$ is derived. ∎

Let us remark that $I_{\alpha,h}$ can be written as

$$I_{\alpha,h} = (u_{\alpha,h},\, p_{\alpha,h})_{L^2(D)} + \sup_{u \in U_{ad}} (u,\, -p_{\alpha,h})_{L^2(D)},$$

where the second addend is the support functional of U_{ad} evaluated at $-p_{\alpha,h}$.

In order to guarantee (2.1), we would like to choose α such that the discrepancy measure $I_{\alpha,h}$ is smaller than the optimal a priori rate given by Proposition 1.9. Since κ is not available, we will overestimate the convergence rate. Thus, we suggest to use

$$\alpha(h) := \sup\{\alpha > 0 :\, I_{\alpha,h} \le \delta_2(h)^2 + \delta_\infty(h)^2\}. \tag{2.2}$$

In the case $\kappa > 1$, the convergence of $\|y_{0,h} - y_0\|_Y$ is faster than $\delta_2(h) + \delta_\infty(h)$, which is exactly the convergence rate of $\|p_{0,h} - p_0\|_{L^\infty(D)}$. In this sense, (2.2) is an overestimation of the convergence rate of $\|p_{0,h} - p_0\|_{L^\infty(D)}$.

ETNA
Kent State University
http://etna.math.kent.edu

As it will turn out, this overestimation will not affect the convergence rate. In particular, $\alpha(h)$ will not be chosen too small. A lower bound of $\alpha(h)$ is provided in Corollary 2.5. First, let us prove that under weak assumptions the supremum exists, i.e., $\alpha(h) < +\infty$.

LEMMA 2.2. *Let us assume that $S^*z \neq 0$. Moreover, we assume that*

$$\|(S^* - S_h^*)z\|_{L^2(D)} \leq \delta_2(h).$$

Suppose furthermore that there is a number $\sigma > 0$ such that $u_b(x) > \sigma$ and $-\sigma > u_a(x)$ a.e. on D. Then for h sufficiently small, we have $\alpha(h) < +\infty$.

Proof. Since $0 \in U_{ad}$, we have for $\alpha \to \infty$ that $u_{\alpha,h} \to 0$ in $L^2(D)$. This implies $p_{\alpha,h} \to -S_h^*z$ for $\alpha \to \infty$. Arguing as in the proof of [25, Lemma 4.2], we obtain

$$\liminf_{\alpha\to\infty} I_{\alpha,h} \geq \sigma \|S_h^* z\|_{L^1(D)}.$$

Using the assumption in the lemma in order to estimate $\|(S^* - S_h^*)z\|_{L^2(D)}$, we obtain

$$\|S_h^* z\|_{L^1(D)} \geq c(\|S^*z\|_{L^1(D)} - \delta_2(h)).$$

Consequently, there must be a $h_0 > 0$ such that the set of all α with $I_{\alpha,h} \leq \delta_y(h)^2 + \delta_p(h)^2$ is bounded for all $h < h_0$, which implies $\alpha(h) < +\infty$ for $h < h_0$. □

Let us show that the assumption $S^*z \neq 0$ is not very restrictive. In fact, if this assumption is violated, then it follows that $u \equiv 0$ is a solution of (1.4) for all $\alpha \geq 0$.

COROLLARY 2.3. *Let $S^*z = 0$. Then $u_0 = 0$ is a solution of* (1.4) *for all $\alpha \geq 0$.*

Proof. If $u = 0$, then $y := Su = 0$ holds as well. The hypothesis $S^*z = 0$ then implies that $p := S^*(y - z) = 0$, which proves that $u = 0$ fulfills the necessary optimality condition of (1.4); cf. Proposition 1.3. □

In the next step, we will prove that $\alpha(h) > 0$ holds. That is, we prove that the discrepancy principle (2.2) yields a regularization of the discrete problem. To this end, let us first derive an upper bound on $I_{\alpha,h}$.

LEMMA 2.4. *Suppose that the operator S_h^* is uniformly bounded, i.e.,*

$$\|S_h^*\|_{\mathcal{L}(Y,L^\infty(D))} \leq M$$

with a constant $M > 0$ independent of h. Then there exists a constant c independent of α and h such that

$$I_{\alpha,h} \leq c\left(\alpha\left(\delta_2(h)^\kappa + \delta_\infty(h)^{\kappa\min(d,1)}\right) + \alpha^{\frac{2}{2-\kappa}} + \alpha^{\kappa+1}\right)$$

is satisfied for all $h > 0$, $\alpha > 0$ provided that $\kappa < 1$. In the case $\kappa \geq 1$, the estimate is valid except that the term $\alpha^{\frac{2}{2-\kappa}}$ has to be omitted.

Proof. Analogous to [25, Lemma 4.3] we obtain

$$I_{\alpha,h} \leq c\,\alpha(\|p_0 - p_{\alpha,h}\|^\kappa_{L^\infty(D)} + \alpha^\kappa). \tag{2.3}$$

We estimate $\|p_0 - p_{\alpha,h}\|_{L^\infty(D)}$ using Proposition 1.9, the uniform boundedness of S_h^*, and Lemma 2.1

$$\begin{aligned}\|p_0 - p_{\alpha,h}\|_{L^\infty(D)} &\leq \|p_0 - p_{0,h}\|_{L^\infty(D)} + \|p_{0,h} - p_{\alpha,h}\|_{L^\infty(D)} \\ &\leq c\left(\delta_2(h) + \delta_\infty(h)^{\min(d,1)} + \|y_{0,h} - y_{\alpha,h}\|_Y\right) \\ &\leq c\left(\delta_2(h) + \delta_\infty(h)^{\min(d,1)} + I_{\alpha,h}^{1/2}\right).\end{aligned}$$

ETNA
Kent State University
http://etna.math.kent.edu

In the case $\kappa < 1$, we obtain using Young's inequality

$$\begin{aligned} I_{\alpha,h} &\le c\,\alpha \left(\delta_2(h)^\kappa + \delta_\infty(h)^{\kappa \min(d,1)} + I_{\alpha,h}^{\kappa/2} + \alpha^\kappa\right) \\ &\le c\left(\alpha(\delta_2(h)^\kappa + \delta_\infty(h)^{\kappa\min(d,1)}) + \alpha^{\frac{2}{2-\kappa}} + \alpha^{\kappa+1}\right). \end{aligned}$$

Now let $\kappa > 1$. By Young's inequality we can estimate

$$c\,\alpha\, I_{\alpha,h}^{\kappa/2} = c\,\alpha\, I_{\alpha,h}^{\frac{\kappa-1}{2}}\; I_{\alpha,h}^{1/2} \le c\,\alpha^2 I_{\alpha,h}^{\kappa-1} + \frac{1}{2} I_{\alpha,h} \le c\,\alpha^{\kappa+1} + \frac{1}{2} I_{\alpha,h},$$

where in the last step we used that (2.3) implies $I_{\alpha,h} \le c\,\alpha$. Summarizing, we obtain the following inequality

$$I_{\alpha,h} \le c\left(\alpha(\delta_2(h)^\kappa + \delta_\infty(h)^{\kappa\min(d,1)}) + \alpha^{\kappa+1}\right),$$

which proves the claim. ∎

The upper bound of $I_{\alpha,h}$ provided by the previous lemma yields $\lim_{\alpha\to 0} I_{\alpha,h} = 0$ for fixed $h > 0$. Hence it is clear that the supremum in the parameter choice rule (2.2) is attained at a positive value of $\alpha(h)$. Moreover, with the help of this estimate we can prove lower bounds for $\alpha(h)$.

COROLLARY 2.5. *Let the assumptions of Lemma 2.4 be satisfied. Let $\alpha(h)$ be given by (2.2). Then it holds that $\alpha(h) > 0$.*

In the case $\kappa \le 1$, we have the following lower bound

$$\alpha(h) \ge \min\left(D^{\frac{4-3\kappa}{4-2\kappa}},\; D^{\frac{1}{\kappa+1}}\right)$$

with $D = c(\delta_2(h)^2 + \delta_\infty(h)^2)$, $c > 0$. If h is small enough, then

$$\alpha(h) \ge c\left(\delta_2(h)^2 + \delta_\infty(h)^2\right)^{\frac{4-3\kappa}{4-2\kappa}}.$$

If $\kappa > 1$, then it holds that

$$\alpha(h) \ge \min\left(D^{\frac{2-\kappa}{2}},\, D^{\frac{1}{\kappa+1}}\right),$$

hence if h is small enough, then

$$\alpha(h) \ge c\left(\delta_2(h)^2 + \delta_\infty(h)^2\right)^{\frac{1}{\kappa+1}}.$$

Proof. First, let us prove the explicit lower bound if $\kappa \le 1$. Since $\kappa \le 1$, the result of Proposition 1.3 implies $d = \frac{1}{2-\kappa} \le 1$. Let us fix $\alpha := 2\alpha(h) > 0$. Then the discrepancy principle (2.2) implies

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

Applying Young's inequality to the estimate of Lemma 2.4 gives

$$\begin{aligned} \delta_2(h)^2 + \delta_\infty(h)^2 &\le c\left(\alpha\left(\delta_2(h)^\kappa + \delta_\infty(h)^{\frac{\kappa}{2-\kappa}}\right) + \alpha^{\frac{2}{2-\kappa}} + \alpha^{\kappa+1}\right) \\ &\le c\left(\alpha^{\frac{2}{2-\kappa}} + \alpha^{\frac{4-2\kappa}{4-3\kappa}} + \alpha^{\frac{2}{2-\kappa}} + \alpha^{\kappa+1}\right) \\ &\le c\,\alpha\left(\alpha^{\frac{\kappa}{4-3\kappa}} + \alpha^{\frac{\kappa}{2-\kappa}} + \alpha^\kappa\right) \end{aligned}$$

ETNA
Kent State University
http://etna.math.kent.edu

with some constant $c > 0$ independent of α, h. As it can be seen, the smallest and largest exponent of α in the above estimate is $\frac{4-2\kappa}{4-3\kappa}$ and $\kappa + 1$, respectively. Using the convexity of $t \mapsto \alpha^t$, we can estimate

$$\delta_2(h)^2 + \delta_\infty(h)^2 \le c \left(\alpha^{\frac{4-2\kappa}{4-3\kappa}} + \alpha^{\kappa+1}\right),$$

which proves the claim.

Second, in the case $\kappa > 1$ we have $d = \frac{\kappa+1}{2} > 1$. Then the estimate of Lemma 2.4 reads

$$I_{\alpha,h} \le c\left(\alpha\left(\delta_2(h)^\kappa + \delta_\infty(h)^\kappa\right) + \alpha^{\kappa+1}\right).$$

Using $\delta_2(h)^2 + \delta_\infty(h)^2 < I_{\alpha,h}$, we conclude that either $\alpha \ge c\left(\delta_2(h)^2 + \delta_\infty(h)^2\right)^{\frac{2-\kappa}{2}}$ or $\alpha \ge c\left(\delta_2(h)^2 + \delta_\infty(h)^2\right)^{\frac{1}{\kappa+1}}$ holds true. $\square$

So far, we proved that $\alpha(h)$ is non-trivial (i.e., non-zero and finite). This enables us to prove optimal convergence rates for the errors in the states y and adjoint states p.

LEMMA 2.6. *Let $\alpha(h)$ be given by (2.2). Then for all $h_{max} > 0$, there is a constant $c > 0$ independent of h such that*

$$\|y_0 - y_{\alpha(h),h}\|_Y + \|p_0 - p_{\alpha(h),h}\|_{L^\infty(D)} \le c\left(\delta_2(h) + \delta_\infty(h)^{\min(d,1)}\right)$$

holds for all $h < h_{max}$.

Proof. Combining the result of Lemma 2.1 with the discrepancy principle (2.2) gives

$$\|y_{0,h} - y_{\alpha(h),h}\|_Y \le I^{1/2}_{\alpha(h),h} \le \delta_2(h) + \delta_\infty(h).$$

Using the a priori error estimate of Proposition 1.9 gives

$$\|y_0 - y_{0,h}\|_Y \le c\left(\delta_2(h) + \delta_\infty(h)^d\right),$$

and the claimed estimate for $\|y_0 - y_{\alpha(h),h}\|_Y$ follows. It remains to derive an upper bound for $\|p_0 - p_{\alpha(h),h}\|_{L^\infty(D)}$. Here, we estimate

$$\begin{aligned}\|p_0 - p_{\alpha(h),h}\|_{L^\infty(D)} &\le \|S^*(y_0 - y_{\alpha(h),h})\|_{L^\infty(D)} + \|(S^* - S_h^*)(y_{\alpha(h),h} - z)\|_{L^\infty(D)} \\ &\le c\left(\delta_2(h) + \delta_\infty(h)^{\min(d,1)} + \delta_\infty(h)\right),\end{aligned}$$

which completes the proof. $\square$

This result proves that the parameter choice rule (2.2) maintains the optimal convergence rates in the sense that $\|p_0 - p_{\alpha(h),h}\|_{L^\infty(D)} \sim \|p_0 - p_{0,h}\|_{L^\infty(D)}$. In the case $\kappa \le 1$, we moreover have that $\|y_0 - y_{\alpha(h),h}\|_Y \sim \|y_0 - y_{0,h}\|_Y$.

Now, let us prove the main result of this section, which states that the parameter choice rule (2.2) also leads to optimal convergence rates for $\|u_0 - u_{\alpha(h),h}\|_{L^1(D)}$ in the case $\kappa \le 1$.

THEOREM 2.7. *Let $\alpha(h)$ be given by (2.2). If $\kappa \le 1$, then for all $h_{max} > 0$ there is a constant $c > 0$ independent of h such that*

$$\|u_0 - u_{\alpha(h),h}\|_{L^1(D)} \le c\left(\delta_2(h)^\kappa + \delta_\infty(h)^{\frac{\kappa}{2-\kappa}}\right)$$

holds for all $h < h_{max}$. In the case $\kappa > 1$, we have the estimate

$$\|u_0 - u_{\alpha(h),h}\|_{L^1(D)} \le c\left(\delta_2(h)^{\frac{2\kappa}{\kappa+1}} + \delta_\infty(h)^{\frac{2\kappa}{\kappa+1}}\right).$$

ETNA
Kent State University
http://etna.math.kent.edu

Proof. Let us write for short $\alpha := \alpha(h)$ throughout the proof. Let us define the following subset of D

$$B := \{x \in D : \, p_{\alpha,h}(x) \neq 0, \ \operatorname{sign}(p_0(x)) = \operatorname{sign}(p_{\alpha,h}(x))\}.$$

The measure of its complement can be bounded using Assumption 1.4. Indeed, on $D \setminus B$ the signs of p_0 and $p_{\alpha,h}$ are different, which gives $|p_0| \le |p_0 - p_{\alpha,h}|$ a.e. on $D \setminus B$. Hence using Assumption 1.4 and Lemma 2.6, we obtain

$$\text{(2.4)} \qquad \operatorname{meas}(D \setminus B) \le c\, \|p_0 - p_{\alpha,h}\|^{\kappa}_{L^\infty(D)} \le c\, \left(\delta_2(h)^\kappa + \delta_\infty(h)^{\kappa \min(d,1)}\right).$$

Let us now investigate the L^1-norm of $u_0 - u_{\alpha,h}$ on B. For $\epsilon > 0$ we define the set

$$B_\epsilon := B \cap \{x \in D : \, |p_{\alpha,h}(x)| > \epsilon\}.$$

Since $|p_0| \le |p_{\alpha,h}| + |p_0 - p_{\alpha,h}| \le \epsilon + |p_0 - p_{\alpha,h}|$ a.e. on $B \setminus B_\epsilon$, we have by Assumption 1.4 and Lemma 2.6

$$\text{(2.5)} \quad \operatorname{meas}(B \setminus B_\epsilon) \le c(\|p_0 - p_{\alpha,h}\|^{\kappa}_{L^\infty(D)} + \epsilon^\kappa) \le c(\delta_2(h)^\kappa + \delta_\infty(h)^{\kappa \min(d,1)} + \epsilon^\kappa).$$

We recall that α satisfies the discrepancy estimate, cf. (2.2),

$$\int_{\{x:\, p_{\alpha,h} > 0\}} (u_{\alpha,h} - u_a) p_{\alpha,h} \, \mathrm{d}\mu + \int_{\{x:\, p_{\alpha,h} < 0\}} (u_{\alpha,h} - u_b) p_{\alpha,h} \, \mathrm{d}\mu \le \delta_2(h)^2 + \delta_\infty(h)^2.$$

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

$$\int_{\{x:\, p_{\alpha,h} > 0\} \cap B_\epsilon} (u_{\alpha,h} - u_a) p_{\alpha,h} \, \mathrm{d}\mu + \int_{\{x:\, p_{\alpha,h} < 0\} \cap B_\epsilon} (u_{\alpha,h} - u_b) p_{\alpha,h} \, \mathrm{d}\mu \\ \le \delta_2(h)^2 + \delta_\infty(h)^2.$$

Since $|p_{\alpha,h}| \ge \epsilon$ on B_ϵ, it holds that

$$\int_{\{x:\, p_{\alpha,h} > 0\} \cap B_\epsilon} \epsilon |u_{\alpha,h} - u_a| \, \mathrm{d}\mu + \int_{\{x:\, p_{\alpha,h} < 0\} \cap B_\epsilon} \epsilon |u_b - u_{\alpha,h}| \, \mathrm{d}\mu \\ \le \delta_2(h)^2 + \delta_\infty(h)^2.$$

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

$$\epsilon \int_{B_\epsilon} |u_0 - u_{\alpha,h}| \le \delta_2(h)^2 + \delta_\infty(h)^2,$$

which implies

$$\|u_0 - u_{\alpha,h}\|_{L^1(B_\epsilon)} \le \epsilon^{-1}(\delta_2(h)^2 + \delta_\infty(h)^2).$$

This together with (2.4), (2.5) gives the estimate

$$\|u_0 - u_{\alpha,h}\|_{L^1(D)} \le c(\delta_2(h)^\kappa + \delta_\infty(h)^{\kappa \min(d,1)} + \epsilon^\kappa + \epsilon^{-1}(\delta_2(h)^2 + \delta_\infty(h)^2)).$$

ETNA
Kent State University
http://etna.math.kent.edu

With $\epsilon := \left(\delta_2(h)^2 + \delta_\infty(h)^2\right)^{\frac{1}{\kappa+1}}$ we obtain

$$\|u_0 - u_{\alpha,h}\|_{L^1(D)} \le c(\delta_2(h)^\kappa + \delta_\infty(h)^{\kappa \min(d,1)} + \delta_2(h)^{\frac{2\kappa}{\kappa+1}} + \delta_\infty(h)^{\frac{2\kappa}{\kappa+1}}).$$

For $\kappa \le 1$ we have $d = \frac{1}{2-\kappa} \le 1$, hence it holds that

$$\|u_0 - u_{\alpha,h}\|_{L^1(D)} \le c\left(\left(1 + \delta_2(h)^{\frac{\kappa(1-\kappa)}{1+\kappa}}\right)\delta_2(h)^\kappa + \left(1 + \delta_\infty(h)^{\frac{3\kappa(1-\kappa)}{(2-\kappa)(1+\kappa)}}\right)\delta_\infty(h)^{\frac{\kappa}{2-\kappa}}\right).$$

For $\kappa > 1$, we have $d = \frac{\kappa+1}{2} > 1$, hence we obtain in this case

$$\|u_0 - u_{\alpha,h}\|_{L^1(D)} \le c\left(\left(1 + \delta_2(h)^{\frac{\kappa(\kappa-1)}{1+\kappa}}\right)\delta_2(h)^{\frac{2\kappa}{\kappa+1}} + \left(1 + \delta_\infty(h)^{\frac{\kappa(\kappa-1)}{1+\kappa}}\right)\delta_\infty(h)^{\frac{2\kappa}{\kappa+1}}\right),$$

which completes the proof. ∎

REMARK 2.8. Analogous results can be obtained when using the modified Assumption 1.10 instead of Assumption 1.2, except that $\delta_2(h)$ and $\delta_\infty(h)$ have to be replaced by $\delta_2'(h)$ and $\delta_\infty'(h)$ in the above estimates.

3. Application to the optimal control of an elliptic equation. Let us report on the application of the adaptive parameter choice rule to an optimal control problem subject to an elliptic equation. The optimal control problem is stated as follows: minimize $J(y, u)$, which is given by

$$J(y, u) = \frac{1}{2}\|y - z\|^2_{L^2(\Omega)} \tag{3.1}$$

subject to the elliptic equation

$$\begin{aligned} -\Delta y &= u + f, \\ y|_\Gamma &= 0 \end{aligned} \tag{3.2}$$

and to the control constraints

$$u_a(x) \le u(x) \le u_b(x) \quad \text{a.e. on } \Omega.$$

Here, $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, is a bounded domain with a polygonal boundary Γ. Moreover, we set $D = \Omega$ endowed with the Lebesgue measure. In addition, $z, f \in L^2(\Omega)$, $u_a, u_b \in L^\infty(\Omega)$ are given.

Let us show that this problem fits into the framework developed in the above sections. Clearly, the partial differential equation (3.2) admits for each $u \in L^2(\Omega)$ a unique solution $Su := y \in H^1_0(\Omega)$. With the choice $Y = L^2(\Omega)$, we have $S = S^*$. Due to Stampacchia's classical result [21], the operator S^* is continuous from $L^2(\Omega)$ to $L^\infty(\Omega)$. Then the problem (3.1) is equivalent to the minimization of $\frac{1}{2}\|S(u + f) - z\|^2_{L^2(\Omega)}$, and the theory of the previous sections applies if one sets $z := z - Sf$.

Let us fix the assumptions on the discretization of (3.2) by finite elements. We will work with a family of regular meshes $\mathcal{F} = \{\mathcal{T}_h\}_{h>0}$, where $\mathcal{T}$ is a regular mesh consisting of closed cells T. These meshes are indexed by their mesh sizes, i.e., $h(\mathcal{T}_h) = h$ with the setting $h(\mathcal{T}) := \max_{T\in\mathcal{T}} h_T$, $h_T := \operatorname{diam} T$. We assume in addition that there exists a positive constant R such that $\frac{h_T}{R_T} \le R$ hold for all cells $T \in \mathcal{T}_h$ and all $h > 0$, where R_T is the diameter of the largest ball contained in T.

With each mesh $\mathcal{T}_h$, we associate a finite-dimensional subspace $V_h \subset H^1_0(\Omega)$ consisting of functions whose restriction to a cell $T \in \mathcal{T}_h$ is a linear polynomial. The operator S_h is then defined by the notion of weak solutions: we set $S_h u := y_h$, where $y_h \in V_h$ solves

$$\int_\Omega \nabla y_h \cdot \nabla v_h \,\mathrm{d}\mu = \int_\Omega (u + f) \cdot v_h \,\mathrm{d}\mu \qquad \forall v_h \in V_h.$$

Please note, that also in the discrete case we have $S_h = S_h^*$.

3.1. A priori error estimates. Let us first discuss the a priori error estimates. In addition, we assume now that Ω is convex and that $\max_{T\in\mathcal{T}_h} h_T \le C_M \min_{T\in\mathcal{T}_h} h_T$ is satisfied for all $h > 0$ with a given constant $C_M > 1$. Then we have the following classical estimates; see [5, Theorem 5.7.6], [4, p. 87]: there is an $h_0 > 0$ such that

$$\|(S - S_h)f\|_{L^2(\Omega)} \le c\,h^2\,\|f\|_{L^2(\Omega)},$$
$$\|(S^* - S_h^*)f\|_{L^\infty(\Omega)} \le c\,h^{2-n/2}\,\|f\|_{L^2(\Omega)},$$

holds for all $h \in (0, h_0)$ and $f \in L^2(\Omega)$ with a constant $c > 0$ independent of h and f. Hence, Assumption 1.2 is fulfilled with $\delta_2(h) = c\,h^2$, $\delta_\infty(h) = c\,h^{2-n/2}$. Of course, these rates will not lead to optimal discretization error estimates since they are limited by the regularity of the finite element solutions, where in general $y_h \notin H^2(\Omega)$. In order to obtain optimal a priori rates, we will have to resort to the modified Assumption 1.10. Here it turns out that under additional regularity assumptions, see [6], we get $\delta_\infty(h)' = h^2|\log h|^{r(n)}$, with $r(2) = 2$ and $r(3) = 11/4$.

3.2. A posteriori error estimates. In addition to relying on a priori rates, we apply the following reliable and efficient error estimator from [20]. For simplicity, we use L^∞-error estimators for both the state and the adjoint equation. Define $\eta_{y_{\alpha,h},\infty} := \max_{T\in\mathcal{T}_h} \eta_{T,y_{\alpha,h},\infty}$ 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 that, cf. [20],

$$c_0\|(S - S_h)u_{\alpha,h}\|_{L^\infty(\Omega)} \le \eta_{y_{\alpha,h},\infty} \le c_1\|(S - S_h)u_{\alpha,h}\|_{L^\infty(\Omega)}$$

with constants $c_0, c_1 > 0$ depending on Ω, the polynomial degree l, and the shape regularity of the triangulation. Note that convexity of Ω is not required. Analogously, we define $\eta_{p_{\alpha,h},\infty} := \max_{T\in\mathcal{T}_h} \eta_{T,p_{\alpha,h},\infty}$ and

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

with the error bound

$$c_0\|(S^* - S_h^*)(y_{\alpha,h} - z)\|_{L^\infty(\Omega)} \le \eta_{p_{\alpha,h},\infty} \le c_1\|(S^* - S_h^*)(y_{\alpha,h} - z)\|_{L^\infty(\Omega)}.$$

With the help of these error estimators, we used the following parameter choice rule:

$$\alpha(h) := \sup\left\{\alpha:\ \alpha = 2^{-j}\alpha_0,\ j \in \mathbb{N},\ I_{\alpha,h} \le \tau(\eta^2_{y_{\alpha,h},\infty} + \eta^2_{p_{\alpha,h},\infty})\right\}. \tag{3.3}$$

Here, $\alpha_0 > 0$ is an a priori chosen initial regularization parameter. It is not obvious whether the supremum in (3.3) exists and is non-zero since the term $\eta^2_{y_{\alpha,h},\infty} + \eta^2_{p_{\alpha,h},\infty}$ does depend on α.

LEMMA 3.1. *Let $\mathcal{T}_h$ be a fixed mesh. Let Assumption 1.4 be satisfied. If $z \ne 0$, then the supremum in (3.3) exists and $\alpha(h) > 0$.*

Proof. Let us assume that the supremum in (3.3) does not exist. Then it holds that

$$I_{\alpha_j,h} > \tau(\eta^2_{y_{\alpha_j,h},\infty} + \eta^2_{p_{\alpha_j,h},\infty})$$

ETNA
Kent State University
http://etna.math.kent.edu

for all $\alpha_j := 2^{-j}\alpha_0$, $j \in \mathbb{N}$. Let us abbreviate $u_j := u_{\alpha_j,h}$, $y_j := y_{\alpha_j,h}$, $p_j := p_{\alpha_j,h}$. Due to the control constraints, the sequence $\{u_j\}$ is uniformly bounded in $L^\infty(\Omega)$. By ellipticity, the sequences $\{y_j\}$ and $\{p_j\}$ are uniformly bounded in $H^1_0(\Omega)$. Since these sequences belong to the finite-dimensional space V_h, we get by the equivalence of norms in finite-dimensional spaces that $\{y_j\}$ and $\{p_j\}$ are bounded in $L^\infty(\Omega)$.

By extracting subsequences if necessary, we have $u_j \rightharpoonup \tilde u$ in $L^2(\Omega)$ and $u_j \rightharpoonup^* \tilde u$ in $L^\infty(\Omega)$. Moreover, it holds that $y_j \to \tilde y_h = S_h\tilde u \in V_h$ and $p_j \to \tilde p_h = S_h^*(\tilde y_h - z) \in V_h$ with respect to any norm as V_h is finite-dimensional. By the lower semicontinuity of norms, we obtain

$$\eta^2_{\tilde y_h,\infty} + \eta^2_{\tilde p_h,\infty} \le \liminf_{j\to\infty}(\eta^2_{y_j,\infty} + \eta^2_{p_j,\infty}),$$

which implies

$$\eta^2_{\tilde y_h,\infty} + \eta^2_{\tilde p_h,\infty} \le \tau^{-1}\liminf_{j\to\infty} I_{\alpha_j,h}.$$

By the inequality (2.3), we obtain $I_{\alpha_j,h} \to 0$ as $j \to \infty$ because of the convergence $\alpha_j \to 0$ and the uniform boundedness of $\{p_j\}$ in $L^\infty(\Omega)$. Hence, we find $\eta_{\tilde y_h,\infty} = \eta_{\tilde p_h,\infty} = 0$. This implies that $\tilde y_h$ and $\tilde p_h$ are affine-linear on Ω. As these functions satisfy homogeneous Dirichlet-boundary conditions, it holds that $\tilde y_h = \tilde p_h = 0$. Now, $\eta_{\tilde y_h,\infty} = 0$ implies $\tilde u + f = 0$, which in turn gives $\tilde y_h = 0$. In addition, $\eta_{\tilde p_h,\infty} = 0$ implies $\tilde y_h - z = 0$, hence $z = 0$, which is a contradiction to the assumptions. ∎

The following corollary follows directly from the parameter choice rule (3.3) and Lemma 2.1.

COROLLARY 3.2. *Let Assumption 1.4 be satisfied with $\kappa \le 1$. Let $\alpha(h)$ be given by (3.3). Then for all $h_{max} > 0$, there is a constant $c > 0$ independent of h such that*

$$\|y_{0,h} - y_{\alpha(h),h}\|_Y + \|p_{0,h} - p_{\alpha(h),h}\|_{L^\infty(D)} \le c\left(\eta_{y_{\alpha(h),h},\infty} + \eta_{p_{\alpha(h),h},\infty}\right)$$

holds for all $h < h_{max}$.

3.3. Example 1. Let us now show some results using uniform meshes. Let us take the following data

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

It is easy to check that (1.1) admits the following unique solution:

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

In addition, it turns out that the regularity assumption is satisfied for all $\kappa < 1$. This implies by Proposition 1.5 that $d = 1$ and $\|u_0 - u_\alpha\|_{L^1(D)} \le c\,\alpha$.

Hence, if $\alpha(h)$ is chosen as

$$\alpha(h) := \sup\{\alpha > 0:\ I_{\alpha,h} \le \delta_2'(h)^2 + \delta_\infty'(h)^2\}, \tag{3.4}$$

we expect to observe the convergence rates

$$\begin{aligned}&\|y_0 - y_{\alpha(h),h}\|_Y + \|p_0 - p_{\alpha(h),h}\|_{L^\infty(\Omega)} + \|u_0 - u_{\alpha(h),h}\|_{L^1(\Omega)}\\ &\qquad\le c\,(\delta_2'(h) + \delta_\infty'(h)) = c\,h^2(1 + |\log h|^{r(n)}),\end{aligned}$$

cf. Lemma 2.6 and Theorem 2.7, taken into account Remark 2.8.

ETNA
Kent State University
http://etna.math.kent.edu

TABLE 3.1
Example 1 using a priori rates.

h	$\|u_0-u_{\alpha(h),h}\|_{L^1(D)}$	$\|y_0-y_{\alpha(h),h}\|_{L^2(D)}$	$\|p_0-p_{\alpha(h),h}\|_{L^\infty(D)}$	$\alpha(h)$
$3.5355 \cdot 10^{-1}$	$3.7125 \cdot 10^{-1}$	$5.9522 \cdot 10^{-2}$	$1.0123 \cdot 10^{-2}$	$1.9531 \cdot 10^{-3}$
$1.7678 \cdot 10^{-1}$	$1.0430 \cdot 10^{-1}$	$1.4209 \cdot 10^{-2}$	$3.4092 \cdot 10^{-3}$	$4.8828 \cdot 10^{-4}$
$8.8388 \cdot 10^{-2}$	$2.8239 \cdot 10^{-2}$	$3.2977 \cdot 10^{-3}$	$8.7735 \cdot 10^{-4}$	$1.2207 \cdot 10^{-4}$
$4.4194 \cdot 10^{-2}$	$7.5361 \cdot 10^{-3}$	$7.7452 \cdot 10^{-4}$	$2.1348 \cdot 10^{-4}$	$3.0518 \cdot 10^{-5}$
$2.2097 \cdot 10^{-2}$	$1.7609 \cdot 10^{-3}$	$1.8484 \cdot 10^{-4}$	$5.1677 \cdot 10^{-5}$	$3.8147 \cdot 10^{-6}$
$1.1049 \cdot 10^{-2}$	$4.3913 \cdot 10^{-4}$	$4.4649 \cdot 10^{-5}$	$1.2606 \cdot 10^{-5}$	$9.5367 \cdot 10^{-7}$
	$\sim h^2$	$\sim h^2$	$\sim h^2$	$\sim h^2$

TABLE 3.2
Example 1 using a posteriori estimates.

h	$\|u_0-u_{\alpha(h),h}\|_{L^1(D)}$	$\|y_0-y_{\alpha(h),h}\|_{L^2(D)}$	$\|p_0-p_{\alpha(h),h}\|_{L^\infty(D)}$	$\alpha(h)$
$3.5355 \cdot 10^{-1}$	$3.7125 \cdot 10^{-1}$	$5.9522 \cdot 10^{-2}$	$1.0123 \cdot 10^{-2}$	$1.9531 \cdot 10^{-3}$
$1.7678 \cdot 10^{-1}$	$1.0430 \cdot 10^{-1}$	$1.4209 \cdot 10^{-2}$	$3.4092 \cdot 10^{-3}$	$4.8828 \cdot 10^{-4}$
$8.8388 \cdot 10^{-2}$	$2.8239 \cdot 10^{-2}$	$3.2977 \cdot 10^{-3}$	$8.7735 \cdot 10^{-4}$	$1.2207 \cdot 10^{-4}$
$4.4194 \cdot 10^{-2}$	$6.9357 \cdot 10^{-3}$	$7.7398 \cdot 10^{-4}$	$2.1340 \cdot 10^{-4}$	$1.5259 \cdot 10^{-5}$
$2.2097 \cdot 10^{-2}$	$1.7609 \cdot 10^{-3}$	$1.8484 \cdot 10^{-4}$	$5.1677 \cdot 10^{-5}$	$3.8147 \cdot 10^{-6}$
$1.1049 \cdot 10^{-2}$	$4.3913 \cdot 10^{-4}$	$4.4649 \cdot 10^{-5}$	$1.2606 \cdot 10^{-5}$	$9.5367 \cdot 10^{-7}$
	$\sim h^2$	$\sim h^2$	$\sim h^2$	$\sim h^2$

3.3.1. Results using a priori rates and uniform refinement. In our computations, we did not calculate the supremum in (2.2) or (3.4) accurately. Rather, we choose some initial value $\alpha_0 > 0$, and then computed $\alpha(h)$ as

$$\alpha(h) := \sup\{\alpha : \ \alpha = 2^{-j}\alpha_0, \ j \in \mathbb{N}, \ I_{\alpha,h} \leq \tau h^4\} \tag{3.5}$$

ignoring the logarithmic term in $\delta'_\infty(h)$. Here, $\tau > 0$ is an additional factor. After the parameter $\alpha(h)$ has been found according to (3.5), the mesh was refined uniformly.

Let us report about the results using the parameter choice rule (3.5), which is based on a priori convergence rates. We chose $\alpha_0 = 1$ and $\tau = 0.01$. The domain Ω was first divided into 32 triangles with mesh size $h = 0.3536$. The results are displayed in Table 3.1.

As expected, we observed that the errors $\|y_0 - y_{\alpha(h),h}\|_{L^2(D)}$, $\|p_0 - p_{\alpha(h),h}\|_{L^\infty(\Omega)}$, and $\|u_0 - u_{\alpha(h),h}\|_{L^1(\Omega)}$ converge at the order of h^2. As can be found in Table 3.1, the parameter choice rule selects $\alpha(h) \sim h^2$.

3.3.2. Results using a posteriori estimates and uniform refinement. Let us now report about the results when instead of the a priori error estimates, an a posteriori error estimator is used. We choose $\tau = 2 \cdot 10^{-4}$ in (3.3) to obtain results which are comparable with those in Section 3.3.1. Using the same data and setup as in Section 3.3.1, we obtained the results as depicted in Table 3.2. The errors almost coincide with those of Table 3.1. This is due to the fact that the only difference in the computations is in the parameter choice rules (3.5) and (3.3). As both select the same values of $\alpha(h)$ (except at one discretization level), this is the expected outcome.

ETNA
Kent State University
http://etna.math.kent.edu

TABLE 3.3
Example 2 using a posteriori estimates and uniform refinement.

N	$\alpha(h)$	$I_{\alpha(h),h}$
71	$1.5625 \cdot 10^{-2}$	$8.6100 \cdot 10^{-4}$
253	$7.8125 \cdot 10^{-3}$	$2.4128 \cdot 10^{-4}$
953	$3.9062 \cdot 10^{-3}$	$6.9901 \cdot 10^{-5}$
3697	$1.9531 \cdot 10^{-3}$	$1.9965 \cdot 10^{-5}$
14561	$9.7656 \cdot 10^{-4}$	$5.6060 \cdot 10^{-6}$
57793	$2.4414 \cdot 10^{-4}$	$4.8650 \cdot 10^{-7}$

TABLE 3.4
Example 2 using a posteriori estimates and adaptive refinement.

N	$\alpha(h)$	$I_{\alpha(h),h}$
71	$1.5625 \cdot 10^{-2}$	$8.6100 \cdot 10^{-4}$
99	$1.5625 \cdot 10^{-2}$	$8.5669 \cdot 10^{-4}$
165	$7.8125 \cdot 10^{-3}$	$2.2686 \cdot 10^{-4}$
377	$3.9062 \cdot 10^{-3}$	$6.7847 \cdot 10^{-5}$
721	$3.9062 \cdot 10^{-3}$	$6.5876 \cdot 10^{-5}$
1410	$1.9531 \cdot 10^{-3}$	$1.8306 \cdot 10^{-5}$
2874	$9.7656 \cdot 10^{-4}$	$5.2227 \cdot 10^{-6}$
5557	$9.7656 \cdot 10^{-4}$	$5.1821 \cdot 10^{-6}$
11671	$4.8828 \cdot 10^{-4}$	$1.5013 \cdot 10^{-6}$
22015	$2.4414 \cdot 10^{-4}$	$4.0858 \cdot 10^{-7}$
46428	$1.2207 \cdot 10^{-4}$	$1.0967 \cdot 10^{-7}$

3.4. Example 2. Our second example is posed on a convex polygonal domain Ω that is the interior of the convex hull of the points $(0,0)$, $(2,0)$, $(2,2)$, $(1,2)$, and $(0,1)$. The following data are given

$$z(x_1,x_2) = 10\frac{x_1^2 - x_2^2}{x_2^2+1},\ f=0,\ u_a=-1,\ u_b=+1.$$

3.4.1. Results using a posteriori estimates and uniform refinement. The discretization error was estimated as in Section 3.3.2, and we used the parameter choice rule (3.3). The coarsest mesh was obtained by dividing the domain Ω into 112 triangles. The results for the computations on a hierarchy of uniformly refined meshes can be observed in Table 3.3. There, N corresponds to the number of nodes of the mesh, which is the number of degrees of freedom of the state and adjoint variables.

3.4.2. Results using a posteriori estimates and adaptive refinement. Using the same data as in the previous section, we also performed computations with an adaptive mesh refinement. Let us describe the marking strategy, which we used to mark triangles for refinement. After $\alpha(h)$ has been found according to (3.3), we refined using the error estimator $\eta_{y,\infty} + \eta_{p,\infty}$. Here, an element T was marked for refinement if

$$\eta_{T,y,\infty} + \eta_{T,p,\infty} \geq \theta(\eta_{y,\infty} + \eta_{p,\infty})$$

with $\theta = 0.5$. After refinement, the parameter $\alpha(h)$ is determined on the new mesh according to the parameter choice rule (3.3).

When comparing the results of the computations with uniform and adaptive refinement, cf. Tables 3.3 and 3.4, the adaptive method clearly gives better results: for the same number of degrees of freedom, the discrepancy $I_{\alpha(h),h}$ is roughly four times smaller as in the case with uniform refinement. Due to Lemma 2.1, this implies that the error $\|y_{0,h} - y_{\alpha(h),h}\|_Y$ decays faster with adaptive refinement than with uniform refinement. Hence, our computational experience suggests to use the fully adaptive algorithm, where the discretization and regularization is chosen adaptively.

REFERENCES

[1] R. BECKER AND B. VEXLER, *A posteriori error estimation for finite element discretization of parameter identification problems*, Numer. Math., 96 (2004), pp. 435–459.
[2] ———, *Mesh refinement and numerical sensitivity analysis for parameter calibration of partial differential equations*, J. Comput. Phys., 206 (2005), pp. 95–110.
[3] H. BEN AMEUR, G. CHAVENT, AND J. JAFFRÉ, *Refinement and coarsening indicators for adaptive parametrization: application to the estimation of hydraulic transmissivities*, Inverse Problems, 18 (2002), pp. 775–794.
[4] D. BRAESS, *Finite Elemente – Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie*, Springer, Berlin, 1997.
[5] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, Springer, New York, 1994.
[6] K. DECKELNICK AND M. HINZE, *A note on the approximation of elliptic control problems with bang-bang controls*, Comput. Optim. Appl., 51 (2012), pp. 931–939.
[7] H. W. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.
[8] H. W. ENGL AND A. NEUBAUER, *Convergence rates for Tikhonov regularization in finite-dimensional subspaces of Hilbert scales*, Proc. Amer. Math. Soc., 102 (1988), pp. 587–592.
[9] U. FELGENHAUER, *On stability of bang-bang type controls*, SIAM J. Control Optim., 41 (2003), pp. 1843–1867.
[10] A. GRIESBAUM, B. KALTENBACHER, AND B. VEXLER, *Efficient computation of the Tikhonov regularization parameter by goal-oriented adaptive discretization*, Inverse Problems, 24 (2008), 025025 (20 pages).
[11] R. GRIESSE AND M. WEISER, *On the interplay between interior point approximation and parametric sensitivities in optimal control*, J. Math. Anal. Appl., 337 (2008), pp. 771–793.
[12] C. W. GROETSCH AND A. NEUBAUER, *Regularization of ill-posed problems: optimal parameter choice in finite dimensions*, J. Approx. Theory, 58 (1989), pp. 184–200.
[13] A. GROTHENDIECK, *Sur certains sous-espaces vectoriels de L^p*, Canadian J. Math., 6 (1954), pp. 158–160.
[14] M. HINZE, *A variational discretization concept in control constrained optimization: the linear-quadratic case*, Comput. Optim. Appl., 30 (2005), pp. 45–61.
[15] B. KALTENBACHER, A. KIRCHNER, AND B. VEXLER, *Adaptive discretizations for the choice of a Tikhonov regularization parameter in nonlinear inverse problems*, Inverse Problems, 27 (2011), 125008 (28 pages).
[16] J. T. KING AND A. NEUBAUER, *A variant of finite-dimensional Tikhonov regularization with a posteriori parameter choice*, Computing, 40 (1988), pp. 91–109.
[17] P. MAASS, S. V. PEREVERZEV, R. RAMLAU, AND S. G. SOLODKY, *An adaptive discretization for Tikhonov-Phillips regularization with a posteriori parameter selection*, Numer. Math., 87 (2001), pp. 485–502.
[18] F. NATTERER, *The finite element method for ill-posed problems*, RAIRO Anal. Numér., 11 (1977), pp. 271–278.
[19] A. NEUBAUER, *Solution of ill-posed problems via adaptive grid regularization: convergence analysis*, Numer. Funct. Anal. Optim., 28 (2007), pp. 405–423.
[20] R. H. NOCHETTO, A. SCHMIDT, K. G. SIEBERT, AND A. VEESER, *Pointwise a posteriori error estimates for monotone semi-linear equations*, Numer. Math., 104 (2006), pp. 515–538.
[21] G. STAMPACCHIA, *Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus*, Ann. Inst. Fourier (Grenoble), 15 (1965), pp. 189–258.
[22] F. TRÖLTZSCH, *Optimal Control of Partial Differential Equations*, AMS, Providence, 2010.
[23] 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 (2000), pp. 1938–1984.

ETNA
Kent State University
http://etna.math.kent.edu

[24] G. WACHSMUTH AND D. WACHSMUTH, *Convergence and regularization results for optimal control problems with sparsity functional*, ESAIM Control Optim. Calc. Var., 17 (2011), pp. 858–886.

[25] ———, *Regularization error estimates and discrepancy principle for optimal control problems with inequality constraints*, Control and Cybernet., 40 (2011), pp. 1125–1158.

[26] ———, *Necessary conditions for convergence rates of regularizations of optimal control problems*, in System Modelling and Optimization, D. Hömberg and F. Tröltzsch, eds., IFIP Advances in Information and Communication Technology vol. 391, Springer, Heidelberg, 2013, pp. 145–154.