PROCEEDINGS OF THE XXIV CONGRESS ON DIFFERENTIAL EQUATIONS AND APPLICATIONS
XIV CONGRESS ON APPLIED MATHEMATICS
Cádiz, June 8-12, 2015, pp. 1–6

Optimal Control of Partial Differential Equations with Nonsmooth Cost Functionals

Christian Clason* Karl Kunisch† Armin Rund ‡

Abstract— Over the last decade significant progress was made in the analysis and numerical treatment of optimal control problems with nonsmooth cost functionals. Such functionals are in the context of optimal control with sparsity constraints, for switching control and for multi-bang optimal control problems. The natural setting of such problems is given by non-reflexive Banach spaces, which leads to new analytical challenges. The lack of smoothness, on the other hand, demands novel numerical methods for practical solution of the resulting infinite dimensional optimization problems.

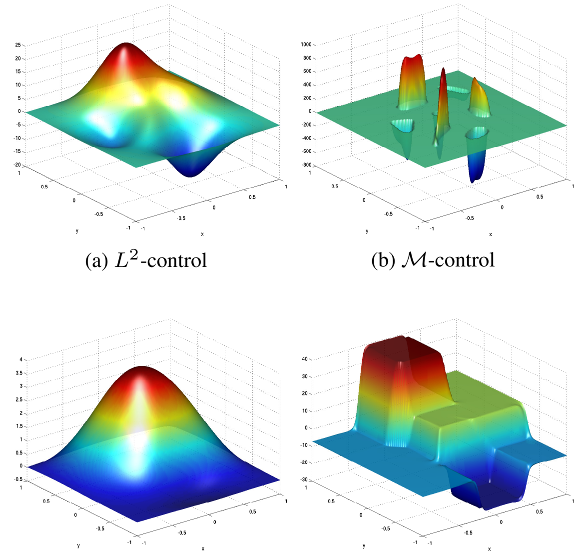
Keywords: Optimal Control with Sparsity Constraints, Multi-bang Control, Switching Control.

1 Introduction

We consider an optimal control problem of tracking type

$$(1) \qquad \begin{cases} \min_{u\in\mathcal{X}} \|y-y_{\mathrm{d}}\|^2_{L^2(\Omega)} + \alpha\mathcal{N}(u) \\ \text{s.t.} \quad Ay = u \end{cases}$$

where A is a differential operator, y is the state variable and y_{d} is a given desired state. The control u has to be determined in such a manner that the state is brought as close to z as possible while observing the control cost term $\alpha\mathcal{N}(u)$. The focus here lies on the choice of the functional $\mathcal{N}$ which has a significant effect on the optimal control as can be seen from Figure 1, which is obtained for the case where A is the Laplacian with homogenous Dirichlet boundary conditions on the unit square, and z is a scaled version of the `peaks` function from MATLAB. In the left column of Figure 1 the optimal controls for the choices $\mathcal{N}(u) = \|u\|^2_{L^2(\Omega)}$ and $\mathcal{N}(u) = \|u\|^2_{H^1(\Omega)}$ are depicted. We observe the global smoothing effects of these functionals when we compare with the choices $\mathcal{N}(u) = \mathcal{M}(u)$ and $\mathcal{N}(u) = BV(u)$, where $\mathcal{M}(u)$ stands for the Borel measure of u, and $BV(u)$ for the total bounded variation semi-norm. The choice $\mathcal{M}(u)$ promoted sparsity, i.e. $u = 0$ over large subsets of the domain Ω, while the BV-semi norm promotes sparsity of the derivative and results in piecewise constant values of the optimal controls. We note here that the choice $\mathcal{N}(u) = \mathcal{M}(u)$ as opposed to $\mathcal{N}(u) = \int_\Omega |u|\,dx$ results from difficulties when proving existence of optimal controls for (1) using the latter control cost.



(a) L^2-control (b) $\mathcal{M}$-control

(c) H^1-control (d) BV-control

Figure 1: Optimal control for different functionals $\mathcal{N}$.

While the results of Figure 1 are given here for a simple test case problem it is by now well established that these features are generic for a wide class of optimal control problems related to different classes of partial differential equations. In the liter-

*Faculty of Mathematics, University Duisburg-Essen, 45117; Essen, Germany. Email: christian.clason@uni-due.de
†Institute for Mathematics and Scientific Computing, University of Graz, Heinrichstraße 36, A-8010 Graz, Austria, and Radon Institute, Austrian Academy of Sciences. Email: karl.kunisch@uni-graz.at
‡Institute for Mathematics and Scientific Computing, University of Graz, Heinrichstraße 36, A-8010 Graz, Austria. Email: armin.rund@uni-graz.at

ature $\mathcal{N}(u) = \|u\|^2_{L^2(\Omega)}$ is typically chosen as control cost for good reason: the simplicity of computing gradients and the use in the linear quadratic regulator theory are among them. In the case of nonlinear partial differential equations a specific type of nonlinearity may require to replace the L^2-norm by an L^p-norm, with $p > 2$ or, in the case of boundary controls, with the $H^{1/2}$-norm for example, to obtain well-posedness of the differential equation in variational form and existence to (1). The choices of $\mathcal{M}$ related to norm or semi-norms of non-reflexive Banach spaces are much more recent. In the references we give a list of some, but by no means all of the references which have become available in the recent past.

There are several reasons which make the choice $\mathcal{N}(u) = \mathcal{M}(u)$ an important one. First, as we shall see below, the control will be allowed to be shut off ($u = 0$) over subsets of the domain. Second, this norm expresses 'proportionality' which is not the case for the L^2-norm. Moreover, as already pointed out in [21] it provides an elegant solution to the problem of optimal actuator placement. For a nice application we refer to [2]. Also, as pointed out in [17], this formulation can be efficient to solve inverse source problems in the case that A in (1) represents a diffusion-convection operator.

In the following Section 2 we highlight some aspects related to sparse optimal control related to linear elliptic equations mostly following the results from [8]. Section 3 is devoted to multi-bang control and Section 4 to switching control.

2 Elliptic control problems with sparse solution

Here we consider

$$\min_{u\in\mathcal{M}(\Omega)} J(u) = \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \alpha\|u\|_{\mathcal{M}(\Omega)}, \tag{2}$$

where y is the unique solution to the Dirichlet problem

$$\left\{\begin{array}{rcll} -\Delta y + c_0 y &=& u & \text{in } \Omega,\\ y &=& 0 & \text{on } \Gamma,\end{array}\right. \tag{3}$$

with $c_0 \in L^\infty(\Omega)$ and $c_0 \geq 0$. We assume that $\alpha > 0$, $y_d \in L^2(\Omega)$ and Ω is a bounded domain in $\mathbb{R}^n$, $n = 2$ or 3, with a smooth boundary. The controls are taken in the space of regular Borel measures $\mathcal{M}(\Omega)$, which is identified with the dual space of $C_0(\Omega)$:

$$\|u\|_{\mathcal{M}(\Omega)} = \sup_{\|z\|_{C_0(\Omega)}\leq 1} \langle u, z\rangle = \sup_{\|z\|_{C_0(\Omega)}\leq 1} \int_\Omega z(x)\, du. \tag{4}$$

The results of this section can be extended to the case where the support of the controls is restricted to be in a subset $\omega \subset \Omega$.

We shall refer to y as a solution if it satisfies the very weak solution concept, i.e.

$$\int_\Omega yAz\, dx = \int_\Omega z\, du \quad \text{for all } z \in H^2(\Omega) \cap H^1_0(\Omega), \tag{5}$$

where $A = -\Delta + c_0 I$. It is well known, see for instance [3], that there exists a unique solution to (3) in the sense of (5). Moreover, $y \in W^{1,p}_0(\Omega)$ for every $1 \leq p < \frac{n}{n-1}$ and

$$\|y\|_{W^{1,p}_0(\Omega)} \leq C_p\|u\|_{\mathcal{M}(\Omega)}.$$

Then, it can be obtained by the standard approach that (2) has a unique solution. Hereafter, this optimal solution will be denoted by $\bar u$ with an associated state $\bar y$. By using subdifferential calculus of convex functions and introducing the adjoint state we get the following first order necessary optimality condition, see [8].

THEOREM 1 *There exists a unique element $\bar p \in H^2(\Omega) \cap H^1_0(\Omega)$ satisfying*

$$\left\{\begin{array}{rcll} -\Delta\bar p + c_0\bar p &=& \bar y - y_d & \textit{in } \Omega,\\ \bar p &=& 0 & \textit{on } \Gamma,\end{array}\right. \tag{6}$$

such that

$$\begin{aligned} &\alpha\|\bar u\|_{\mathcal{M}(\Omega)} + \textstyle\int_\Omega \bar p\, d\bar u = 0,\\ &\|\bar p\|_{C_0(\Omega)} \left\{\begin{array}{l} = \alpha \textit{ if } \bar u \neq 0,\\ \leq \alpha \textit{ if } \bar u = 0.\end{array}\right.\end{aligned} \tag{7}$$

This optimality condition together with the Jordan decomposition of $\bar u = \bar u^+ - \bar u^-$, can be used to deduce from (7) that $\bar p \subset [-\alpha, \alpha]$ and

$$\left\{\begin{array}{l} \text{supp } (\bar u^+) \subset \{x \in \Omega : \bar p(x) = -\alpha\},\\ \text{supp } (\bar u^-) \subset \{x \in \Omega : \bar p(x) = +\alpha\}.\end{array}\right. \tag{8}$$

This implies that the optimal control is zero where $|\bar p(x)| \neq \alpha$ and implies the desired sparsity: unless the adjoint state, which is in some sense a sensitivity measure, is maximal, the optimal control is inactive.

In the case where we consider the observation of the state only in a subset $\omega_y \subset \Omega$, then we have the following property of the support of the optimal control.

PROPOSITION 2 *Let ω_y be an open subset of Ω such that $\Omega \setminus \omega_y$ is connected and consider the functional*

$$J_{\omega_y}(u) = \frac{1}{2}\|y - y_d\|^2_{L^2(\omega_y)} + \alpha\|u\|_{\mathcal{M}(\Omega)}. \tag{9}$$

Then the associated optimal control $\bar u$ satisfies supp $(\bar u) \subset \bar\omega_y$.

Further results on the support of the controls which are all based on the maximum principle are contained in [20].

To explain a numerical treatment for (2) we may commence by observing that the optimality conditions of Theorem 1 may be reformulated in primal-dual form as

$$\left\{\begin{array}{l} \bar p = A^{-*}(A^{-1}\bar u - y_{\mathrm{d}}),\\ 0 \geq \langle \bar u, \bar p - p\rangle \quad \text{for all } \|p\|_{C_0} \leq \alpha.\end{array}\right. \tag{10}$$

This system is still intractable in function spaces, and hence we consider its Moreau–Yosida regularization

$$
(11) \qquad \begin{cases} p_\gamma = A^{-*}(A^{-1}u_\gamma - y_\mathrm{d}), \\ u_\gamma = \dfrac{1}{\gamma}(\max(0, p_\gamma - \alpha) + \min(0, p_\gamma + \alpha)), \end{cases}
$$

for which convergence to a solution of (10) as $\gamma \to 0$ can be shown. Furthermore, this system can be solved efficiently by a semi-smooth Newton method [11, 8], which is given in Algorithm 1.

Algorithm 1 Semismooth Newton method for (11)

1: Set $k = 0$, Choose $u^0 \in L^2(\Omega)$
2: **repeat**
3: Solve for y^k in $Ay = u^k$
4: Solve for p^k in $A^*p = y^k - y_\mathrm{d}$
5: Set
$$\mathcal{A}_k^+ = \{x \in \Omega : p^k(x) > \alpha\},$$
$$\mathcal{A}_k^- = \{x \in \Omega : p^k(x) < -\alpha\},$$
6: Set $F(u^k) = u^k - \frac{1}{\gamma}(\chi_{\mathcal{A}_k^+}(p^k - \alpha) + \chi_{\mathcal{A}_k^-}(p^k + \alpha))$
7: Solve for $\delta u \in L^2(\Omega)$
$$\delta u - \frac{1}{\gamma}(\chi_{\mathcal{A}_k^+} + \chi_{\mathcal{A}_k^-})A^{-*}A^{-1}\delta u = -F(u^k)$$
using a matrix-free Krylov method
8: Set $u^{k+1} = u^k + \delta u$ and $k = k + 1$
9: **until** $(\mathcal{A}_{k+1}^+ = \mathcal{A}_k^+)$ and $(\mathcal{A}_{k+1}^- = \mathcal{A}_k^-)$

This is still an infinite dimensional problem. In [8] a framework based on approximation of the measure valued controls by the linear combination of Dirac deltas was proposed and analyzed which allows taking the limit $\gamma \to 0$ computationally. The analysis was later improved in [20].

The necessity to utilize measure spaces for the controls can be avoided and replaced by $L^1(\Omega)$ if additional constraints on the norms are utilized, see e.g. [21, 4, 15]. Directional sparsity was analyzed first in [15]. Results on sparse controls have been extended to nonlinear elliptic [5], to parabolic [17, 6], and to wave equations [18], where the citations only point at part of the literature.

3 Multi-bang control

Multi-bang control refers to optimal control problems for partial differential equations where a distributed control should only take on values from a discrete set of values u_i. This property can be promoted by a combination of L^2 and L^0-type control costs. The resulting functional, however, is non-convex and lacks weak lower-semicontinuity. More specifically we consider the problem

(12)
$$
\begin{cases} \min\limits_{u,y} \dfrac{1}{2}\|y - y_\mathrm{d}\|_{L^2}^2 + \dfrac{\alpha}{2}\|u\|_{L^2}^2 + \beta \displaystyle\int_\Omega \prod_{i=1}^d |u(x) - u_i|_0 \, dx \\ \text{s.t. } Ay = u, \quad u_1 \le u(x) \le u_d \text{ for a.e. } x \in \Omega \end{cases}
$$

where $\alpha > 0, \beta > 0$, $d \in \mathbb{N}$, and the binary term is given by

$$
|t|_0 := \begin{cases} 0 \text{ if } t = 0, \\ 1 \text{ if } t \neq 0. \end{cases}
$$

Below we summarize some results from [10] which rely in an essential manner on a convexification process applied to (12). For this purpose we define

$$\mathcal{F} : L^2(\Omega) \to \mathbb{R}, u \mapsto \tfrac{1}{2}\|A^{-1}u - y_\mathrm{d}\|_{L^2}^2,$$

$$\mathcal{G}_0 : L^2(\Omega) \to \overline{\mathbb{R}},$$

$$u \mapsto \int_\Omega \left(\frac{\alpha}{2}|u(x)|^2 + \beta \prod_{i=1}^d |u(x) - u_i|_0 \right) dx + \delta_U(u),$$

where δ_U is the indicator function of the admissible set

$$U := \left\{ u \in L^2(\Omega) : u_1 \le u(x) \le u_d \text{ for a.e. } x \in \Omega \right\}.$$

With this notation (12) can be expressed as

$$(13) \qquad \min_u \mathcal{F}(u) + \mathcal{G}_0(u).$$

Since $\mathcal{G}_0$ is not convex standard convex analysis techniques are not applicable. We therefore pass to the convexification of (12) and consider

$$(14) \qquad \min_u \mathcal{F}(u) + \mathcal{G}(u),$$

where $\mathcal{G} = \mathcal{G}_0^{**}$, which is the biconjugate of $\mathcal{G}$. The necessary optimality condition of (13) can be expressed as: there exists a $\bar{p} = -\mathcal{F}'(\bar{u})$ such that $\bar{p} \in \partial\mathcal{G}(\bar{u})$, which holds if and only if $\bar{u} \in \partial\mathcal{G}^*(\bar{p})$. Here, $\mathcal{G}^*$ denotes the Fenchel conjugate of the convex functional $\mathcal{G}$, and $\partial\mathcal{G}^*$ denotes its convex subdifferential. We thus obtain the primal-dual optimality system

$$
(15) \qquad \begin{cases} -\bar{p} = \mathcal{F}'(\bar{u}) = A^{-*}(A^{-1}\bar{u} - y_\mathrm{d}), \\ \bar{u} \in \partial\mathcal{G}^*(\bar{p}). \end{cases}
$$

We have the following result concerning existence and structure of the solution.

THEOREM 3 *There exists a solution* $(\bar{u}, \bar{p}) \in L^2(\Omega) \times H_0^1(\Omega)$ *of* (15)*. Moreover, if*

$$(16) \qquad \sqrt{2\beta/\alpha} \ge \frac{1}{2}(u_{i+1} - u_i) \quad \text{for all } 1 \le i < d,$$

then

$$\Omega = \bigcup_{i=1}^d \{x \in \Omega : \bar{u}(x) = u_i\} \cup \{x \in \Omega : \bar{y}(x) = y_\mathrm{d}(x)\}$$

where $\bar{y}$ *is that the state corresponding to* $\bar{u}$*.*

To quantify the effect of the convexification process, i.e. the passage from $\mathcal{G}_0$ to $\mathcal{G}$, we introduce the critical set

$$\mathcal{C} = \{x \in \Omega : \bar{p}(x) = \tfrac{1}{2}(u_i + u_{i+1}) \text{ for some } i \in \{1, \dots, d-1\}, \text{ and } \bar{u}(x) \notin \{u_i, u_{i+1}\}\}. \tag{17}$$

Then we have

Theorem 4 *If $(\bar{u}, \bar{p}) \in L^2(\Omega) \times H_0^1(\Omega)$ is a solution of (15) and (16) holds, then for every $u \in L^2(\Omega)$*

$$J(\bar{u}) \le J(u) + \beta|\mathcal{C}|, \tag{18}$$

where J denotes the cost functional in (12), and $|\mathcal{C}|$ stands for the measure of $\mathcal{C}$.

However, the optimality conditions (15) are not directly amenable to numerical solution by Newton-type techniques. For this reason we consider a regularized optimality system

$$\begin{cases} -p_\gamma = \mathcal{F}'(u_\gamma) = A^{-*}(A^{-1}u_\gamma - y_\mathrm{d}), \\ \ \ u_\gamma = (\partial\mathcal{G}^*)_\gamma(p_\gamma), \end{cases} \tag{19}$$

where $(\partial\mathcal{G}^*)_\gamma$ is the Moreau–Yosida approximation of the subdifferential of the Fenchel conjugate $\mathcal{G}^*$. Thus for the numerical realization, only $(\partial\mathcal{G}^*)_\gamma$ is needed which can be computed without explicit knowledge of $\mathcal{G}$, since $\mathcal{G}^* = \mathcal{G}_0^*$. For system (19) semi-smooth Newton methods are applicable. We close this section with a numerical example. Again, z is a scaled version of the MATLAB `peaks` function, and we choose the desired control values $\{u_1, \dots, u_5\} = \{-2, \dots, 2\}$.

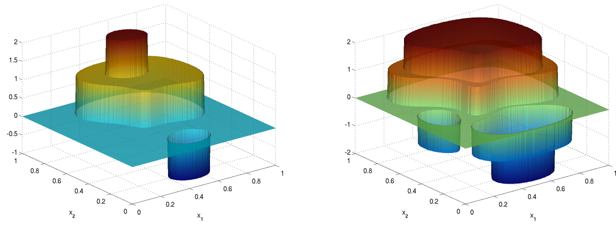


Figure 2: Effect of α, β on the structure of the control u, left: $\alpha = 5.10^{-3}, \beta = 10^{-3}$, right: $\alpha = 10^{-3}, \beta = 10^{-3}$.

4 Switching control

Here we briefly describe two choices of cost functionals which are amenable for computing multiswitching controls for differential equations. Such controls consist of an arbitrary number of components of which at most one should be simultaneously active. Switching can refer to alternating between differential spatial control-subdomains or between different components of vector valued controls, when we refer to time dependent control systems. Switching control is quite well-studied for controlled ordinary differential equations, see e.g. [19] but much less is known for partial differential equations, see, however, [13, 22]. In either case, little attention has been paid to the efficient numerical solution of switching control problems. In this respect we refer to [14] where a relaxation technique combined with rounding strategies is proposed to solve mixed-integer programming problems arising in optimal control of partial differential equations. Here we follow a quite different route which is based on the choice of specially tuned cost functionals and convex analysis techniques. Below we draw on results from [9, 1].

Here we consider the parabolic controlled partial differential equation $Ly = Bu$ on $\Omega_T := [0,T] \times \Omega$, where $L = \partial_t - A$ for an elliptic operator A defined on $\Omega \subset \mathbb{R}^n$, with homogenous boundary conditions, and B is defined by $(Bu)(t,x) = \chi_{\omega_1}(x)u_1(t) + \chi_{\omega_2}(x)u_2(t)$ for given control domains $\omega_1, \omega_2 \subset \overline{\Omega}$.

To promote a switching structure between the temporal control functions u_1 and u_2, we first propose to use the regularized binary function for

$$g(v) = \frac{\alpha}{2}(v_1^2 + v_2^2) + \beta|v_1 v_2|_0,$$

for $v = (v_1, v_2) \in \mathbb{R}^2$. This term combines in a single functional both switching enhancement and a quadratic cost for the active control. For some $\omega_T \subset \Omega_T$ we then consider the problem

$$\begin{cases} \displaystyle\min_{u \in L^2(0,T;\mathbb{R}^2)} \frac{1}{2}\|y - y_\mathrm{d}\|^2_{L^2(\omega_T)} + \int_0^T g(u(t))\,dt, \\ \qquad \text{s.t.} \quad Ly = Bu \text{ on } \Omega_T, y(0) = y_0, \end{cases} \tag{20}$$

for given $y_0 \in L^2(\Omega)$, and $y_\mathrm{d} \in \omega_T$. Using the solution operator $S = L^{-1}B : u \mapsto y$, problem (20) can be expressed in reduced form as

$$\min_u \mathcal{F}(u) + \mathcal{G}_0(u), \tag{21}$$

where $\mathcal{F}$ is smooth and convex, and $\mathcal{G}_0$ is neither smooth nor convex nor, in fact, weakly lower semicontinuous. We therefore consider the relaxed problem

$$\min_u \mathcal{F}(u) + \mathcal{G}(u), \tag{22}$$

where, as in the previous section we set $\mathcal{G} = \mathcal{G}_0^{**}$. Existence and optimality conditions for the relaxed problem can readily be obtained. We again consider the regularized system (19) for S in place of A^{-1}, for which semi-smooth Newton methods are applicable. In [9] the asymptotic behavior for $\gamma \to 0^+$ and sufficient conditions are given for the limit problem, which guarantee that both controls cannot be simultaneously nontrivial except for a singular set, for which $|u_1(t)| = |u_2(t)| \le \sqrt{\frac{2\beta}{\alpha}}$. The practical realization of the approach requires to characterize $(\partial\mathcal{G}^*)_\gamma$ which is quite involved, and a generalization to more than two controls is not straightforward.

To define a second choice for a switching functional we consider control operators of the form

$$(Bu)(t,x) = \sum_{i=1}^{N} \chi_{\omega_i}(x)u_i(t), \tag{23}$$

where the χ_{ω_i} are the characteristic function of given control domains $\omega_i \subset \Omega$ of positive measure, and $(u_1,\dots,u_N)$ is the time-dependent control vector, of which only one component should be nontrivial at any instance in time. For this purpose we consider the optimal control problem

$$\begin{cases} \displaystyle\min_{u\in L^2(0,T;\mathbb{R}^N)} \frac{1}{2}\|y-y_{\mathrm{d}}\|^2_{L^2(\omega_T)} + \frac{\alpha}{2}\int_0^T |u(t)|_1^2\,dt, \\ \text{s.t.} \quad Ly = Bu, \quad y(0)=y_0, \end{cases} \tag{24}$$

where $|\cdot|_1$ stands for the ℓ^1-norm on $\mathbb{R}^N$. This is a convex optimization problem, for which the same steps can be carried out as for (20), but without the need for the convexification step. In the context of exact controllability problems with switching controls (24) was utilized in [22]. In [1], it was shown that (24) coincides with (20) for special choices of β sufficiently large. We close with a numerical example from [1], where

$$y_{\mathrm{d}} = \sum_{i=1}^{N} \cos(i+t)\sin^2\left(2\pi\frac{t}{T}\right)|x-x_i|^2,$$

and the observation and the seven control domains (of which not all need to be active at any time) are depicted in Figure 3.

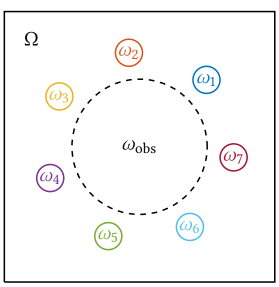


Figure 3: Problem setting for $N = 7$ control components.

Figure 4 depicts the various control branches for two different choices of the control weight α.

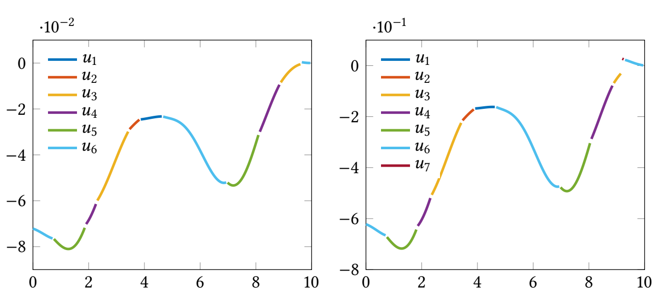


Figure 4: Dependence of the controls on α for $N = 7$.

Acknowledgements

The concepts and results presented in this note originate from collaborations with E. Casas, K. Ito, K. Pieper, P. Trautmann, and B. Vexler. Their contributions are highly appreciated.

Support by Austrian Science Fund (FWF) under grant SFB F32 (SFB “Mathematical Optimization and Applications in Biomedical Sciences”) is gratefully acknowledged.

References

[1] R. Barnard, C. Clason, K. Kunisch, and A. Rund. *A convex analysis approach to multiswitching control of partial differential equations.* Submitted, 2015.

[2] P. Brunner, C. Clason, M. Freiberger, and H. Scharfetter. *A deterministic approach to the adapted optode placement for illumination of highly scattering tissue*, *Biomedical Optics Express*, 3(7):1732–1743, 2012. doi:10.1364/BOE.3.001732.

[3] E. Casas. *Control of an elliptic problem with pointwise state constraints*, *SIAM J. Control Optim.*, 24:1309–1318, 1986. doi:10.1137/0324078.

[4] E. Casas, R. Herzog, and G. Wachsmuth. *Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional.* *SIAM J. Optim.*, 22(3):795–820, 2012. doi:10.1137/110834366.

[5] E. Casas and K. Kunisch. *Optimal control of semilinear elliptic equations in measure spaces.* *SIAM Journal on Control and Optimization*, 52:339–364, 2014. doi:10.1137/13092188X.

[6] E. Casas and K. Kunisch. *Parabolic control problems in space-time measure spaces.* *ESAIM: Control, Optimisation and Calculus of Variations*, 2015. doi:10.1051/cocv/2015008.

[7] E. Casas, C. Clason, and K. Kunisch. *Parabolic control problems in measure spaces with sparse solutions.* *SIAM Journal on Control and Optimization*, 51(1):28–63, 2013. doi:10.1137/120872395.

[8] E. Casas, C. Clason, and K. Kunisch. *Approximation of elliptic control problems in measure spaces with sparse solutions.* *SIAM Journal on Control and Optimization*, 50(4):1735–1752, 2012. doi:10.1137/110843216.

[9] C. Clason, K. Ito, and K. Kunisch. *A convex analysis approach to optimal controls with switching structure for partial differential equations.* *ESAIM: Control, Optimisation and Calculus of Variations*, 2015. doi:10.1051/cocv/2015017.

[10] C. Clason and K. Kunisch. *Multi-bang control of elliptic systems.* *Annales de l'Institut Henri Poincaré (C) Analyse Non Linéaire*, 31:1109–1130, 2014. doi:10.1016/j.anihpc.2013.08.005.

[11] C. Clason and K. Kunisch. *A measure space approach to optimal source placement.* *Computational Optimization and Applications*, 53(1):155–171, 2012. doi:10.1007/s10589-011-9444-9.

[12] C. Clason and K. Kunisch. *A duality-based approach to elliptic control problems in non-reflexive Banach spaces.* *ESAIM: Control, Optimisation and Calculus of Variations*, 17(1):243–266, 2011. doi:10.1051/cocv/2010003.

[13] F. M. Hante, G. Leugering, and T. I. Seidman. *Modeling and analysis of modal switching in networked transport systems.* *Appl. Math. Optim.*, 59(2):275–292, 2009. doi:10.1007/s00245-008-9057-6.

[14] F. M. Hante and S. Sager. *Relaxation methods for mixed-integer optimal control of partial differential equations.* *Comput. Optim. Appl.*, 55(1):197–225, 2013. doi:10.1007/s10589-012-9518-3.

[15] R. Herzog, G. Stadler, and G. Wachsmuth. *Directional sparsity in optimal control of partial differential equations.* *SIAM J. Control Optim.*, 50(2):943–963, 2012. doi:10.1137/100815037.

[16] K. Ito and K. Kunisch. *Optimal control with $l^p(\omega)$, $p \in [0,1)$, control cost.* *SIAM Journal on Control and Optimization*, 52(2):1251–1275, 2014. doi:10.1137/120896529

[17] K. Kunisch, K. Pieper, and B. Vexler. *Measure valued directional sparsity for parabolic optimal control problems.* *SIAM Journal on Control and Optimization*, 52(5):3078–3108, 2014. doi:10.1137/140959055.

[18] K. Kunisch, P. Trautmann, and B. Vexler. *Optimal control of the undamped linear wave equation with measure valued controls.* Submitted, 2014.

[19] D. Liberzon. Switching in systems and control. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 2003. doi:10.1007/978-1-4612-0017-8.

[20] K. Pieper and B. Vexler. *A priori error analysis for discretization of sparse elliptic optimal control problems in measure space.* *SIAM J. Control Optim.*, 51(4):2788–2808, 2013. doi:10.1137/120889137.

[21] G. Stadler. *Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices.* *Comput. Optim. Appl.*, 44(2):159–181, 2009. doi:10.1007/s10589-007-9150-9.

[22] E. Zuazua. Switching control. *J Eur Math Soc*, 13(1):85–117, 2011.