Systems & Control Letters 106 (2017) 1–8

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

Nonconvex penalization of switching control of partial differential equations

Christian Clason [a,*], Armin Rund [b], Karl Kunisch [c,d]

[a] Faculty of Mathematics, Universität Duisburg–Essen, 45117 Essen, Germany
[b] Institute for Mathematics and Scientific Computing, Karl-Franzens-Universität Graz, Heinrichstr. 36, 8010 Graz, Austria
[c] Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, 8010 Graz, Austria
[d] Radon Institute, Austrian Academy of Sciences, Linz, Austria

ARTICLE INFO

Article history:
Received 31 May 2016
Received in revised form 14 March 2017
Accepted 22 May 2017

Keywords:
Optimal control
Switching control
Partial differential equations
Nonsmooth optimization
Nonconvex analysis
Semi-smooth Newton method

ABSTRACT

This paper is concerned with optimal control problems for parabolic partial differential equations with pointwise in time switching constraints on the control. A standard approach to treat constraints in nonlinear optimization is penalization, in particular using L^1-type norms. Applying this approach to the switching constraint leads to a nonsmooth and nonconvex infinite-dimensional minimization problem which is challenging both analytically and numerically. Adding H^1 regularization or restricting to a finite-dimensional control space allows showing existence of optimal controls. First-order necessary optimality conditions are then derived using tools of nonsmooth analysis. Their solution can be computed using a combination of Moreau–Yosida regularization and a semismooth Newton method. Numerical examples illustrate the properties of this approach.

1. Introduction

Switching control refers to time-dependent optimal control problems with a vector-valued control of which at most one component should be active at any point in time. To partially set the stage, we consider for example optimal tracking control for a linear evolution equation $y_t + Au = Bu$ on $\Omega_T := (0,T)\times\Omega$ together with initial conditions $y(0) = y_0$ on Ω, where A is a linear second order elliptic operator defined on $\Omega \subset \mathbb{R}^n$ with homogeneous Neumann boundary conditions and the linear control operator $B : L^2(0,T;\mathbb{R}^N) \to L^2(\Omega_T)$ is given by

$$(Bu)(t,x) = \sum_{i=1}^{N} \chi_{\omega_i}(x)u_i(t), \tag{1.1}$$

where χ_{ω_i} are the characteristic functions of given control domains $\omega_i \subset \Omega$ of positive measure. Furthermore, let $\omega_{\text{obs}} \subset \Omega$ denote the observation domain and let $y^d \in L^2(0,T;L^2(\omega_{\text{obs}}))$ denote the target. Consider now the standard optimal control problem

$$\begin{cases} \min\limits_{u\in L^2(0,T;\mathbb{R}^N)} & \frac{1}{2}\|y-y^d\|^2_{L^2(0,T;L^2(\omega_{\text{obs}}))} + \frac{\alpha}{2}\int_0^T |u(t)|_2^2\,dt, \\ \text{s. t.} & y_t + Ay = Bu, \quad y(0) = y_0, \end{cases} \tag{1.2}$$

where $|v|_2^2 = \sum_{j=1}^N v_j^2$ denotes the squared ℓ^2-norm on $\mathbb{R}^N$. To promote the switching structure of optimal controls, we suggest adding the penalty term

$$\beta \int_0^T \sum_{\substack{i,j=1\\ i<j}}^{N} |u_i(t)u_j(t)|\,dt \tag{1.3}$$

with $\beta > 0$ to the objective, which can be interpreted as an L^1-penalization of the switching constraint $u_i(t)u_j(t) = 0$ for $i \neq j$ and $t \in [0,T]$. The combination of control cost and switching penalty is convex if and only if $\beta \leq \alpha$. The case $\beta = \alpha$ was investigated in [1]; the aim of this work is to treat the case $\beta > \alpha$, which allows choosing the switching penalty parameter independently of the control cost parameter. As can be verified for a simple scalar example, there exist sets of data for which the minimizer of the convex problem is not switching, while the nonconvex problem does admit (possibly multiple) minimizers that are switching.

In the nonconvex case, the approach followed in [1] is not applicable. The main difficulty stems from the fact that the integrand $g : \mathbb{R}^2 \to \mathbb{R}$, $(u_1,u_2) \mapsto |u_1u_2|$, is not convex, and hence the integral functional $G : L^2(0,T;\mathbb{R}^2) \to \mathbb{R}$, $u \mapsto \int_0^T |u_1(t)u_2(t)|\,dt$, is not weakly lower semicontinuous, which is an obstacle for proving existence. It is therefore necessary to enforce strong convergence of minimizing sequences, which is possible by either considering piecewise constant and hence finite-dimensional controls or by introducing an additional (small) $H^1(0,T;\mathbb{R}^2)$ penalty. Our analysis

* Corresponding author.
E-mail addresses: christian.clason@uni-due.de (C. Clason), armin.rund@uni-graz.at (A. Rund), karl.kunisch@uni-graz.at (K. Kunisch).

http://dx.doi.org/10.1016/j.sysconle.2017.05.006
0167-6911/

will cover both approaches. Besides the question of existence of optimal controls, their numerical computation is also challenging due to the nonconvexity of the problem. Here we proceed as follows: Using the calculus of Clarke's generalized derivative [2,3], we can derive first-order necessary optimality conditions. It then suffices to apply a Moreau–Yosida regularization only to the non-smooth but convex term in the optimality conditions in order to apply a semismooth Newton method.

This is a natural continuation of our previous works [1,4] on convex relaxation of the switching constraint. Let us briefly remark on further related literature. On switching control of ordinary and partial differential equations, there exists a large body of work; here we only mention [5–7] in the former context and [8–13] in the latter. A related topic is the control of switched systems, where we refer to, e.g., [14–16].

This paper is organized as follows. Section 2 is concerned with existence of optimal controls and their convergence as $\beta \to \infty$ to a "hard switching constrained" problem. Optimality conditions are then derived in Section 3, where the question of exact penalization is addressed as well. Section 4 discusses the numerical solution of the optimality conditions using a semismooth Newton method. Finally, Section 5 presents numerical examples illustrating the properties of the nonconvex penalty approach.

2. Existence

Here we describe the general framework that will be utilized and which will contain the example in the Introduction as a special case. Let W denote a Hilbert space of measurable functions on the space-time cylinder $\Omega_T = (0, T] \times \Omega$, where $\Omega \subset \mathbb{R}^n$ is a bounded domain with Lipschitz continuous boundary. This space will serve as the state space of the solutions of the control system which appears as a constraint in (1.2). It is assumed that $W \hookrightarrow L^2(0, T; L^2(\Omega))$, and that the embedding is continuous. Further let $\mathcal{U} \subset L^2(0, T; \mathbb{R}^N)$ denote the Hilbert space of controls. We assume that there exists an affine control-to-state mapping $u \mapsto S(u)$. Here, we suppress the dependence of S on y_0; for $y_0 = 0$, we denote the corresponding linear solution operator by S_0. Throughout it is assumed that S satisfies

(A1) $S : L^2(0, T; \mathbb{R}^N) \to W$ is a continuous mapping satisfying

$$\|S(u)\|_W \le C(\|u\|_{L^2(0,T;\mathbb{R}^N)} + \|y_0\|_{L^2(\Omega)})$$

for a constant C independent of u and y_0.

As mentioned in the Introduction, we need to restrict the set of feasible controls in order to obtain existence of an optimal control. We thus consider the following two cases for $\mathcal{U} \subset L^2(0, T; \mathbb{R}^N)$:

(i) $\mathcal{U} = H^1(0, T; \mathbb{R}^N)$;
(ii) $\mathcal{U}$ is finite-dimensional (e.g., consisting of piecewise constant controls).

For the sake of presentation, we further restrict ourselves in the following to the case of two control components; the results remain valid for $N > 2$ components (although it should be pointed out that, in contrast to the convex approach in [1], the number of terms in (1.3) grows as $\binom{N}{2}$). We hence consider for $\beta > \alpha > 0$ the problem

$$\min_{u \in \mathcal{U}} \frac{1}{2}\|Su - y^d\|^2_{L^2(0,T;L^2(\omega_{\text{obs}}))} + \frac{\alpha}{2}\|u\|^2_{L^2(0,T;\mathbb{R}^2)} + \frac{\varepsilon}{2}\|u_t\|^2_{L^2(0,T;\mathbb{R}^2)} + \beta \int_0^T |u_1(t)u_2(t)|\,dt \tag{2.1}$$

with $\omega_{\text{obs}} \subset \Omega$ and $y^d \in L^2(0, T; L^2(\omega_{\text{obs}}))$ as before. If $\mathcal{U}$ is finite-dimensional, it is understood that $\varepsilon = 0$; otherwise we require $\varepsilon > 0$. Keeping $\varepsilon \ge 0$ fixed, we will denote the cost functional in (2.1) by J_β.

Before we turn to address existence for (2.1), we describe three typical cases of interest for which assumption (A1) is satisfied. Throughout the following, A will denote a linear second-order uniformly elliptic operator with smooth coefficients.

Distributed control. We return to the case considered in the Introduction, i.e., we consider the equation in (1.2) with A together with homogeneous Dirichlet, Neumann, or Robin boundary conditions and the control operator $B \in \mathcal{L}(L^2(0, T; \mathbb{R}^N), L^2(\Omega_T))$ as in (1.1). It is then well-known, see, e.g., [17, Chap. 4], that (A1) is satisfied with $W = W(0, T) := H^1(0, T; V^*) \cap L^2(0, T; V)$, where $V = H^1_0(\Omega)$ in the case of homogeneous Dirichlet boundary conditions and $V = H^1(\Omega)$ for homogeneous Neumann or Robin conditions.

Neumann boundary control. Here we consider the case of Neumann boundary control. Thus the control system is given by

$$\begin{cases} y_t + Ay = 0 \text{ in } Q_T, \\ \dfrac{\partial y}{\partial n} = Bu \text{ on } \Sigma_T, \\ y(0) = y_0 \text{ in } \Omega, \end{cases}$$

where $\Sigma_T := (0, T] \times \partial\Omega$, and analogous to (1.1) we now take B to be of the form

$$(Bu)(t, s) = \sum_{i=1}^{2} \chi_{\omega_i}(s)u_i(t),$$

with χ_{ω_i} the characteristic functions of given control domains $\omega_i \subset \partial\Omega$ of positive measure relative to $\partial\Omega$. Again, A1 is satisfied, this time with $W = W(0, T)$ and $V = H^1(\Omega)$. For a reference, see, e.g., [18, Chap. 3.3] and the references given there.

Dirichlet boundary control. Finally we consider the case of Dirichlet boundary control given by

$$\begin{cases} y_t + Ay = 0 \text{ in } Q_T, \\ y = Bu \text{ on } \Sigma_T, \\ y(0) = y_0 \text{ in } \Omega, \end{cases}$$

where B is defined as in the case of Neumann control just above. In this case, A1 can be verified by the method of transposition, and one arrives at the state space

$$W = L^2(0, T; L^2(\Omega)) \cap H^1(0, T; (H^1_0(\Omega) \cap H^2(\Omega))^*) \cap C([0, T]; H^{-1}(\Omega)).$$

This was carried out in, e.g., [19, Thm. 2.1] with leading term in A taken as the Laplacian for simplicity.

Theorem 2.1. *There exists a minimizer $\bar{u} \in \mathcal{U}$ to (2.1).*

Proof. We first consider the case of $\mathcal{U} = H^1(0, T; \mathbb{R}^2)$. Since J_β is bounded from below, there exists a minimizing sequence $\{u_n\}_{n\in\mathbb{N}}$ that is bounded in $H^1(0, T; \mathbb{R}^2)$. Hence, by coercivity of J_β, there exists a subsequence, still denoted by $\{u_n\}_{n\in\mathbb{N}}$, with $u_n \rightharpoonup \bar{u}$ in $H^1(0, T; \mathbb{R}^2)$ and $u_n \to \bar{u}$ pointwise in $(0, T)$. This implies pointwise convergence of $|u_{n,1}(t)u_{n,2}(t)| \to |\bar{u}_1(t)\bar{u}_2(t)|$. Together with the continuity of S and the weak lower semicontinuity of norms, this implies

$$J_\beta(\bar{u}) \le \liminf_{n\to\infty} J_\beta(u_n) = \inf_{u\in\mathcal{U}} J_\beta(u),$$

i.e., $\bar{u}$ is a minimizer.

The case of $\mathcal{U}$ finite dimensional follows similarly, since boundedness in $L^2(0, T; \mathbb{R}^N)$ then directly implies strong and hence pointwise convergence. □