Systems & Control Letters 89 (2016) 66–73

Contents lists available at ScienceDirect

Systems & Control Letters

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

A convex penalty for switching control of partial differential equations

Christian Clason [a,*], Armin Rund [b], Karl Kunisch [b], Richard C. Barnard [c]

[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] Computational and Applied Mathematics Group, Computer Science and Mathematics Division, Oak Ridge National Laboratory, P.O. Box 2008 Oak Ridge, TN 37831, USA

ARTICLE INFO

Article history:
Received 26 March 2015
Received in revised form
10 November 2015
Accepted 21 December 2015
Available online 19 January 2016

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

ABSTRACT

A convex penalty for promoting switching controls for partial differential equations is introduced; such controls consist of an arbitrary number of components of which at most one should be simultaneously active. Using a Moreau–Yosida approximation, a family of approximating problems is obtained that is amenable to solution by a semismooth Newton method. The efficiency of this approach and the structure of the obtained controls are demonstrated by numerical examples.

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 every point in time. We focus here on optimal tracking control for a linear diffusion equation $Ly = Bu$ on $\Omega_T := (0, T] \times \Omega$, $y(0) = y_0$ on Ω, where $L = \partial_t - A$ for an elliptic operator A defined on $\Omega \subset \mathbb{R}^n$ carrying suitable boundary conditions. The control operator B is defined by

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

where χ_{ω_i} is the characteristic function of the given control domain $\omega_i \subset \Omega$ 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. We consider the standard optimal control problem

$$\begin{cases} \min\limits_{u\in L^2(0,T;\mathbb{R}^N)} \dfrac{1}{2}\|y - y^d\|^2_{L^2(0,T;L^2(\omega_{\text{obs}}))} + \dfrac{\alpha}{2}\displaystyle\int_0^T |u(t)|_2^2\,dt, \\ \text{s.t. } Ly = Bu, \qquad y(0) = y_0, \end{cases}$$

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 the optimal control $\bar{u} \in L^2(0, T; \mathbb{R}^N)$, we suggest adding an additional penalty term

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

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 $t \in [0, T]$. For the choice $\beta = \alpha$, the sum of the control cost and the penalty can be simplified to yield the problem

$$\begin{cases} \min\limits_{u\in L^2(0,T;\mathbb{R}^N)} \dfrac{1}{2}\|y - y^d\|^2_{L^2(0,T;L^2(\omega_{\text{obs}}))} + \dfrac{\alpha}{2}\displaystyle\int_0^T |u(t)|_1^2\,dt, \\ \text{s.t. } Ly = Bu, \qquad y(0) = y_0, \end{cases} \tag{P}$$

where $|v|_1 = \sum_{j=1}^N |v_j|$ denotes the ℓ^1-norm on $\mathbb{R}^N$. This is a convex optimization problem, for which we derive first-order optimality conditions in primal–dual form whose Moreau–Yosida regularization (to be introduced below) can be solved using a superlinearly convergent semismooth Newton method. The effect of other choices for β will be discussed in Section 3.3.

The approach which we follow here is related to both the switching control problem in [1] and the distributed parabolic sparse control problem in [2]. In [1] a nonconvex formulation in the case where $N = 2$ was considered; we compare its convex

* 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), barnardrc@ornl.gov (R.C. Barnard).

http://dx.doi.org/10.1016/j.sysconle.2015.12.013

relaxation to the present approach below. One advantage of the approach presented in this paper over that in [1] is given by the fact that there is no significant additional technical burden when considering switching between $N > 2$ controls. In [2] the L^2 norm in time of the measure norm in space was used to promote temporally varying sparsity in space. While the choice of the nonsmooth functional involving the controls in (P) is motivated by sparsity considerations, one can arrive at this functional also from controllability–observability considerations. In fact, it was shown in [3, Theorem 4.1] that – provided an appropriately defined controllability Gramian has full rank – exact null controls with perfect switching have minimal $L^2(0,T;\mathbb{R}^2)$ norm, where $\mathbb{R}^2$ is endowed with the ℓ^1 norm. In contrast, [4] follows a different approach where binary or integer decision variables are sought within a relaxation technique combined with a suitable rounding strategy.

Let us comment on further related work. While our work here aims at formulating optimal control problems with switching controls in a way that allows an efficient numerical treatment, the larger body of work focuses on the stabilization of switching systems. For ordinary differential equations we refer to e.g., [5–7]. For partial differential equations, this problem has received comparatively little attention. In both cases, one should distinguish switching control in the sense defined above from the control of switched systems. For the latter in the context of PDEs, we refer to [8,9]. In [10], converse Lyapunov theorems for abstract switched systems are developed. Lyapunov techniques are also used to study switches in hyperbolic systems in [11], and existence results for optimal control of switching systems modeling the use of bacteria for pollution removal are obtained in [12]. Exact null controls with switching structure for the heat and wave equation were treated in [3,13,14] and [15], respectively.

This work is organized as follows. Section 2 discusses the existence and first-order optimality conditions for solutions to (P) as well as its regularization within an abstract convex analysis framework. Explicit pointwise characterizations of the switching relations arising from the optimality system and for its regularization are given in Section 3. Here we also discuss the relation of the proposed switching functional in (P) to other possible choices of the penalty term. Section 4 is concerned with the numerical solution of the regularized optimality system by a semismooth Newton method. A numerical example for switching control of a two-dimensional linear heat equation is computed in Section 5.

2. Convex analysis approach

We recall the convex analysis approach for (nonconvex) switching controls for partial differential equations from [1], which is also applicable to the convex penalty considered here. For this purpose, we consider Problem (P) in the reduced form

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

with $\mathcal{F} : L^2(0,T;\mathbb{R}^N) \to \mathbb{R}$ and $\mathcal{G} : L^2(0,T;\mathbb{R}^N) \to \mathbb{R}$ given by

$$\mathcal{F}(u) = \frac{1}{2}\|Su - y^d\|^2_{L^2(0,T;L^2(\omega_{\text{obs}}))},$$

$$\mathcal{G}(u) = \frac{\alpha}{2}\int_0^T |u(t)|_1^2\,dt,$$

where the continuous affine solution operator $S : u \mapsto y$ assigns to any control $u \in L^2(0,T;\mathbb{R}^N)$ the unique state $y \in L^2(\Omega_T)$ satisfying the state equation $Ly = Bu$ with initial condition $y(0) = y_0$ subject to appropriate boundary conditions. Here we assume that the coefficients of A, the boundary and initial conditions as well as the domain Ω are sufficiently regular that the range of S is contained in

$$\begin{aligned} W(0,T) &:= L^2(0,T;H_0^1(\Omega)) \cap W^{1,2}(0,T;H^{-1}(\Omega)) \\ &\hookrightarrow C([0,T];L^2(\Omega)). \end{aligned} \tag{2.1}$$

Since S is affine, $\mathcal{F}$ is proper, convex and lower semicontinuous. Furthermore, since the squared norm $|\cdot|_1^2$ is convex, $\mathcal{G}$ is proper, convex, lower semicontinuous, and, in addition, radially unbounded. Existence of a solution thus follows from standard arguments, e.g., Tonelli's direct method.

Proposition 2.1. *There exists a minimizer $\bar{u}$ for Problem (P).*

We next derive first-order optimality conditions in primal–dual form. Throughout, for any proper convex function $\mathcal{H}$, we denote by $\mathcal{H}^*$ its Fenchel conjugate and by $\partial\mathcal{H}$ its subdifferential; see, e.g., [16,17] for their definitions. The following proposition is a direct consequence of the sum rule and inversion formula for convex subdifferentials (see, e.g., [16, Corollary 16.24] for the latter) as well as the Fréchet-differentiability of $\mathcal{F}$.

Proposition 2.2. *The control $\bar{u} \in L^2(0,T;\mathbb{R}^N)$ is a minimizer for (P) if and only if there exists a $\bar{p} \in L^2(0,T;\mathbb{R}^N)$ such that*

$$\begin{cases} -\bar{p} = \mathcal{F}'(\bar{u}), \\ \;\;\bar{u} \in \partial\mathcal{G}^*(\bar{p}), \end{cases} \tag{OS}$$

holds.

Since $\mathcal{F}$ is a standard quadratic tracking term, the first relation in (OS) can be expressed in a straightforward manner in terms of the solution operator $S = L^{-1}B$ and its adjoint $S^* = B^*L^{-*}$ (with homogeneous boundary and initial conditions), i.e., $\bar{p} = -S^*(S\bar{u} - y^d)$. For later use, we point out that due to (2.1) and the specific choice of S there holds $\bar{p} \in V := B^*(W(0,T)) \hookrightarrow L^r(0,T;\mathbb{R}^N)$ for any $r > 2$.

The second relation is responsible for the switching structure of the optimal control $\bar{u}$, and we will give a pointwise characterization in Proposition 3.1.

Our numerical approach is based on the Moreau–Yosida regularization of (OS). Specifically, we replace $\partial\mathcal{G}^*$ for $\gamma > 0$ by

$$\partial\mathcal{G}^*_\gamma(p) := (\partial\mathcal{G}^*)_\gamma(p) := \frac{1}{\gamma}\left(p - \operatorname{prox}_{\gamma\mathcal{G}^*}(p)\right),$$

where

$$\begin{aligned} \operatorname{prox}_{\gamma\mathcal{G}^*}(v) &:= \arg\min_{w\in L^2(0,T;\mathbb{R}^N)} \frac{1}{2\gamma}\|w - v\|^2_{L^2(0,T;\mathbb{R}^N)} + \mathcal{G}^*(w) \\ &= \left(\mathrm{Id} + \gamma\partial\mathcal{G}^*\right)^{-1}(v) \end{aligned}$$

is the proximal mapping of $\mathcal{G}^*$, which in Hilbert spaces coincides with the resolvent of $\partial\mathcal{G}^*$. Note that the proximal mapping and thus the Moreau–Yosida regularization of a proper and convex functional is always single-valued and Lipschitz continuous; see, e.g., [16, Corollary 23.10].

We then consider the regularized system

$$\begin{cases} -p_\gamma &= \mathcal{F}'(u_\gamma), \\ u_\gamma &= \partial\mathcal{G}^*_\gamma(p_\gamma). \end{cases} \tag{OS$_\gamma$}$$

Again, we will give an explicit formulation of (OS$_\gamma$) in the next section.

Proposition 2.3. *For each $\gamma > 0$ system (OS$_\gamma$) admits a unique solution (u_γ, p_γ).*

Proof. Using convex analysis techniques (see, e.g., [16, Chapter 12]), we obtain that (OS$_\gamma$) is the necessary optimality condition for

$$\min_u \mathcal{F}(u) + \left(\mathcal{G}^*_\gamma\right)^*(u). \tag{P$_\gamma$}$$

In fact, since $\mathcal{F}$ is globally defined and continuous, we have the necessary optimality condition

$$0 \in \partial\mathcal{F}(u_\gamma) + \partial\left(\mathcal{G}^*_\gamma\right)^*(u_\gamma).$$