Discrete Event Dyn Syst
DOI 10.1007/s10626-011-0109-8

A Proximal Point Based Approach to Optimal Control of Affine Switched Systems

Vadim Azhmyakov · Michael V. Basin · Jörg Raisch

Received: 30 September 2010 / Accepted: 18 May 2011

Abstract This paper focuses on the proximal point regularization technique for a class of optimal control processes governed by affine switched systems. We consider switched control systems described by nonlinear ordinary differential equations which are affine in the input. The affine structure of the dynamical models under consideration makes it possible to establish some continuity/approximability properties and to specify these models as convex control systems. We show that, for some classes of cost functionals, the associated optimal control problem (OCP) corresponds to a conventional convex optimization problem in a suitable Hilbert space. The latter can be reliably solved using standard first-order optimization algorithms and consistent regularization schemes. In particular, we propose a conceptual numerical approach based on the gradient-type method and classic proximal point techniques.

Keywords Switched control systems · Affine systems · Optimal control · Proximal point algorithm

V. Azhmyakov (✉)
Departamento de Control Automatico, CINVESTAV, A.P. 14-740,
Av. Instituto Politecnico Nacional No. 2508, 07360, Mexico D.F., Mexico
e-mail: vazhmyakov@ctrl.cinvestav.mx

M. V. Basin
Department of Physical and Mathematical Sciences, University of Nuevo Leon,
San Nicolas de los Garza, Mexico
e-mail: mbasin2007@gmail.com

J. Raisch
Fachgebiet Regelungssysteme, Technische Universität Berlin,
Einsteinufer 17, 10587 Berlin, Germany
e-mail: raisch@control.tu-berlin.de

J. Raisch
Systems and Control Theory Group, MPI for Dynamics of Complex Technical Systems,
Sandtorstr. 1, 39106 Magdeburg, Germany

Published online: 04 June 2011

1 Introduction

The study of hybrid and switched systems has gained lots of interest in the recent years (see Attia et al. 2005, 2010; Axelsson et al. 2008a, b; Azhmyakov et al. 2008, 2009, 2011; Cassandras et al. 2001; Ding et al. 2008, 2009; Egerstedt et al. 2006; Garavello and Piccoli 2005; Shaikh and Caines 2007; Sussmann 1999; Verriest et al. 2004; Xu and Antsaklis 2003a, b). Many applications from different disciplines have been studied in this modeling framework. The general switched systems constitute a class of control systems where two types of dynamics are present, continuous and discrete event dynamics (see e.g. Branicky et al. 1998; Lygeros 2003). In order to understand how these systems can be operated efficiently, both aspects of the dynamics have to be considered and taken into account during the optimal control design phase. An analytic approach that combines tools from discrete event systems and continuous systems is the most suitable but at the same time, probably the mostly complex. Evidently, each "component" of the dynamics contributes to the complexity of the resulting control design problem, namely, the combinatorial aspect of the discrete event part and the infinite dimensional problem related to the continuous part. Difficulties also arise from the mutual interaction/coupling of the two above-mentioned dynamic components. The stability analysis and the optimal design for switched systems are such intricate examples of the new system phenomena (see e.g., Branicky et al. 1998; Liberzon 2003). Many approaches have been recently proposed to tackle various OCPs with a particular emphasis on engineering inspired classes of switched systems. Hybrid systems with autonomous location transitions (switching triggered by the continuous part of the state) have been considered in Azhmyakov et al. (2008, 2009) and Xu and Antsaklis (2003a, b). The objective is to find the conventional optimal control and a sequence of switching instants such that a performance index is minimized. Optimization techniques for some delayed switched systems are studied in Verriest et al. (2004). Constructive approaches to optimal control of switched systems have been developed in Axelsson et al. (2008a, b), Ding et al. (2008, 2009), Egerstedt et al. (2006), Shaikh and Caines (2007) and the hybrid LQ optimization theory is discussed in Azhmyakov et al. (2009).

In this contribution we concentrate on the classes of switched systems which are affine in the input. Note that affine and polynomial-type control systems have become a modern application focus of practical control theory (Basin and Calderon-Alvarez 2009; Basin et al. 2011; Isidori 1989; Utkin 1992). The aim of our investigations is to study the analytical structure of the basic affine state equation with switchings and to establish some continuity properties of the corresponding trajectories. This result can also be interpreted as an approximability property of the dynamical models under consideration (see Azhmyakov et al. 2011). Next we characterize a class of affine switched systems as convex control systems (see e.g., Azhmyakov and Raisch 2008). This characterization allows to consider the optimal control processes governed by affine switched systems from the point of view of the classical convex programming and to apply the corresponding theoretical and computational optimization tools (see e.g., Azhmyakov and Schmidt 2003; Hiriart-Urruty and Lemarechal 1993; Ioffe and Tichomirov 1979; Polak 1997; Spellucci 1993). Moreover, we examine the above optimization methods in combination with a regularization scheme, namely, in combination with proximal point technique. This combination makes it possible to obtain a numerically stable general algorithm (see Azhmyakov and Noriega Morales 2010; Kaplan and Tichatschke 1994 for details).

Our paper is organized as follows. Section 2 contains the problem formulation and also an introduction to the conventional proximal point techniques. In Section 3 we study fundamental continuity properties of the classical affine systems. This general continuity result is next applied in Section 4 to our main dynamical model, namely, to switched systems with affine structure. Section 4 is also devoted to OCPs associated with convex affine switched systems. Section 5 discusses a regularized first-order numerical approach to the above OCPs. In this section we apply our main analytical results from the previous sections to optimal control processes governed by convex affine switched systems. Section 6 summarizes the paper.

2 Problem formulation and the classical proximal point method

Consider the following initial value problem

$$\begin{aligned} \dot{x}(t) &= \sum_{i=1}^{r} \chi_{[t_{i-1},t_i)}(t) \left[A_{q_i}(t, x(t)) + B_{q_i}(t)u(t) \right] \text{ a.e. on } [0, t_f], \\ x(0) &= x_0, \end{aligned} \tag{1}$$

where $x_0 \in \mathbb{R}^n$ is a fixed initial state, all functions $A_{q_i} : (0, t_f) \times \mathscr{R} \to \mathbb{R}^n$, where q_i are from an index set $\mathscr{Q}$, are measurable with respect to variable $t \in (0, t_f)$ and locally Lipschitz continuous in $x \in \mathscr{R} \subseteq \mathbb{R}^n$. We also assume that for every $t \in [0, t_f]$ the values $B_{q_i}(t) = (b_{i,j}(t))^n_m$ are $n \times m$ matrices with continuous components $b_{ij}(\cdot)$. We consider systems (1) with $r \in \mathbb{N}$ switching times $\{t_i\}$, $i = 1, ..., r$, where

$$0 = t_0 < t_1 < ... < t_{r-1} < t_r = t_f,$$

and denote by $\chi_{[t_{i-1},t_i)}(\cdot)$ the characteristic function of the disjunct time intervals of the type $[t_{i-1}, t_i)$, $i = 1, ..., r$. Motivated by numerous applications, let us examine Eq. 1 over a set $\mathscr{U}$ of bounded measurable control inputs. In this paper we assume that the set of admissible controls $\mathscr{U}$ has the following structure

$$\mathscr{U} := \left\{ v(\cdot) \in \mathbb{L}^2_m(0, t_f) \mid v(t) \in U \text{ a.e. on } [0, t_f] \right\},$$

where $U \subseteq \mathbb{R}^m$ is a compact and convex set. By $\mathbb{L}^2_m(0, t_f)$ we denote here the Hilbert space of square integrable functions. We assume that for every admissible control function $u(\cdot) \in \mathscr{U}$ the initial value problem (1) has an absolutely continuous solution $x^u(\cdot)$ (on the above time interval). However, due to the presence of a nonlinearity $A_{q_i}(t, x)$ in the above equation, the global existence of a solution to Eq. 1 is not guaranteed, unless there are some additional conditions imposed. We refer to Fattorini (1999), Hale (1969) and Poznyak (2008) for the corresponding details.

The above technical assumptions associated with the initial value problem (1) are called the basic assumptions.

Control system (1) can be considered as a main formal dynamical model for various classes of switched and hybrid systems. We refer to Attia et al. (2005, 2010), Azhmyakov et al. (2008), Azhmyakov and Noriega Morales (2010), Branicky et al. (1998), Cassandras et al. (2001), Egerstedt et al. (2006), Garavello and Piccoli (2005), Liberzon (2003), Lygeros (2003), Moor and Raisch (1999) and Shaikh and Caines (2007), Sussmann (1999) for some concrete definitions and concepts of hybrid/switched control systems. Note that for some classes of the above systems the possible switching mechanisms are formalized by different types of the characteristic functions β. For example, some classes of hybrid systems with autonomous location transitions are considered in Azhmyakov et al. (2008), Garavello and Piccoli (2005) and Shaikh and Caines (2007). The general switched systems with controlled switchings are introduced in Liberzon (2003). Given an affine switched system (1) we now formulate an associated optimization problem, namely, the following main hybrid OCP

$$\begin{aligned} &\text{minimize } \phi(x(t_f)) \\ &\text{subject to (Eq. 1)}, \\ &q_i \in \mathcal{Q},\ u(\cdot) \in \mathcal{U}, \end{aligned} \tag{2}$$

where $\phi : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function. Note that the hybrid OCP (2) has a quite general nature. Consider some functions $\tilde{A}_i : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$, $i = 1, ..., r$ and assume that $\tilde{A}_i(\cdot, \cdot)$ are uniformly bounded on the set $(0, t_f) \times \mathcal{R}$, measurable with respect to $t \in (0, t_f)$ and uniformly Lipschitz continuous in $x \in \mathcal{R} \subseteq \mathbb{R}^n$. Moreover, let us also assume that $\tilde{B}_i : \mathbb{R}_+ \to \mathbb{R}^m$, $i = 1, ..., r$ are continuous functions. A hybrid OCP with the following performance criterion

$$\tilde{\phi}(x(t_f)) + \sum_{i=1}^{r} \int_{t_{i-1}}^{t_i} \left[\tilde{A}_{q_i}(t, x(t)) + \tilde{B}_{q_i}(t)u(t) \right] dt$$

where $q_i \in \mathcal{Q}$ and $\tilde{\phi} : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function, can be reduced to an OCP of the type (2) via simple state augmentation:

$$\dot{x}_{n+1}(t) := \sum_{i=1}^{r} \chi_{[t_{i-1}, t_i)}(t) \left[\tilde{A}_{q_i}(t, x(t)) + \tilde{B}_{q_i}(t)u(t) \right] \ \text{a.e. on } [0, t_f],$$

where x_{n+1} is considered as a new component of the extended $(n+1)$-dimensional state vector $\tilde{x} := (x^T, x_{n+1})^T$. The correspondingly transformed performance criterion can now be written as a terminal criterion

$$\phi\left(\tilde{x}\left(t_f\right)\right) := \tilde{\phi}\left(x\left(t_f\right)\right) + x_{n+1}(t_f).$$

Evidently, the above hybrid state space extension is similar to the usual state augmentation technique for the conventional OCPs (see e.g., Ioffe and Tichomirov 1979).

An OCP involving ordinary differential equations can be considered in various ways as an optimization problem in a suitable function space (Fattorini 1999). For instance, problem (2) can be expressed as an infinite-dimensional nonlinear program of the form:

$$\begin{aligned} &\text{minimize } J(u(\cdot)) \\ &\text{subject to } q_i \in \mathcal{Q},\ u(\cdot) \in \mathcal{U}, \end{aligned} \tag{3}$$

where $J : \mathbb{L}^2_m([0, t_f]) \to \mathbb{R}$ is a composite functional defined as

$$J(u(\cdot)) := \phi\left(x^u\left(t_f\right)\right)$$

(recall that $x^u(\cdot)$ is a solution to Eq. 1 generated by the input $u(\cdot)$).

We now shortly discuss the necessary technical facts related to the classical proximal point regularization method for convex optimization problems. We refer to Azhmyakov and Schmidt (2003), Kaplan and Tichatschke (1994), Martinet (1970), Rockafellar (1976) and Solodov and Svaiter (2000) for the corresponding proofs and some additional theoretical results. Let $\{Z, \ ||\cdot||_Z\}$ be a real Hilbert space. Consider the following convex minimization problem

$$\begin{aligned} &\text{minimize } f(z) \\ &\text{subject to } z \in S, \end{aligned} \tag{4}$$

where $S \subset Z$ is a bounded, convex, closed set, $f : Z \to \bar{\mathbb{R}} := \mathbb{R} \bigcup \{\infty\}$ is a lower semicontinuous and proper convex functional. Note that the existence of an optimal solution $z^{\text{opt}} \in S$ to Eq. 4 is guaranteed (see e.g., Hiriart-Urruty and Lemarechal 1993; Rockafellar 1970). Following Kaplan and Tichatschke (1994), Rockafellar (1976), Martinet (1970) and Azhmyakov and Schmidt (2003) we now introduce the proximal mapping

$$\begin{aligned} &\mathscr{P}_{f,S,K} : \alpha \to \text{Argmin}_{z\in S} \left[f(z) + \frac{K}{2} ||z - \alpha||_Z^2 \right], \\ &K > 0, \ \alpha \in Z \end{aligned}$$

and define the classical proximal point method

$$\begin{aligned} &z^{i+1} \approx \mathscr{P}_{f,S,K_j}\left(z^j\right), \\ &z^0 \in S, \ \ j \in \mathbb{N}, \end{aligned}$$

where $\{K_j\}$ is a given sequence with $0 < K_j \leq C < \infty$. Thus the original problem (4) is replaced by a sequence of the auxiliary (regularized) minimization problems

$$\begin{aligned} &f(z) + \frac{K_j}{2} ||z - z^j||_Z^2 \to \min, \\ &z \in S, \ \ j \in \mathbb{N} \end{aligned}$$

with strongly convex objective functionals. Recall that a sequence $\{z^l\}$, $l \in \mathbb{N}$ of elements from Z is a minimizing sequence for Eq. 4 if

$$\lim_{l\to\infty} f\left(z^l\right) = \min_{z\in S} f(z).$$

It is well known that minimizing sequences are of great importance in constructive optimization theory. A numerical method related with the original optimization problem (4) is called stable if the associated minimizing sequence $\{z^l\}$ converges (in a sense) to an element from the set of optimal solutions to Eq. 4. It is well-known that the proximal point algorithm introduced above is one of the most important stable methods for convex problems of the type (4) (see Azhmyakov and Schmidt 2003; Solodov and Svaiter 2000 and the references therein). A constructive solution procedure for every auxiliary problem from the above sequence of auxiliary problems can finally be obtained as a combination of the above proximal point regularization scheme and a suitable numerical optimization algorithm. We refer to Polak (1997), Poznyak (2008) and Spellucci (1993) for some implementable computational methods.

Suppose that the approximations $\{z^j\}, \ \ j \in S$ generated by the proximal point algorithm satisfy the estimate

$$||z^{j+1} - \mathscr{P}_{f,S,K_j}\left(z^j\right)||_Z \leq \varepsilon_j,$$
$$\sum_{j=0}^{\infty} \frac{\varepsilon_j}{K_j} < \infty,$$

for all $j \in \mathbb{N}$. It is well-known that in this case the sequence $\{z^j\}$ converges in the weak topology to a point from the solution set of problem (4) (Kaplan and Tichatschke 1994; Martinet 1970; Rockafellar 1976). Besides, $\{z^j\}, \ \ j \in \mathbb{N}$ is also a minimizing sequence for Eq. 4. Note that in specific cases one can improve the above-mentioned convergence property and obtain the strong convergence of a modified proximal-type minimizing sequences (see e.g., Azhmyakov and Schmidt 2003; Solodov and Svaiter 2000).

3 Continuity property of the conventional and switched affine control systems

In this section we consider a fundamental continuity result for general conventional (non-switched) affine systems (Azhmyakov et al. 2011). The corresponding theorem will be then applied to our main dynamical model, namely, to affine switched systems of the type (1).

Theorem 1 *Consider the following initial value problem*

$$\begin{aligned} &\dot{x}(t) = a(t, x(t)) + b(t)u(t) \ \text{ a.e. on } [0, t_f], \\ &x(0) = x_0, \end{aligned} \tag{5}$$

where a function $a : (0, t_f) \times \mathscr{R} \to \mathbb{R}^n$ *is uniformly bounded on the open set* $(0, t_f) \times \mathscr{R}$*, measurable with respect to* $t \in (0, t_f)$ *and uniformly* Lipschitz *continuous in* $x \in \mathscr{R} \subseteq \mathbb{R}^n$*. Let* $b : \mathbb{R} \to \mathbb{R}^{n\times m}$ *be a continuous function. Let* $\{u^k(\cdot)\}, \ \ k \in \mathbb{N}$ *be an associated sequence of control functions* $u^k(\cdot) \in \mathbb{L}^2_m(0, t_f)$ *and* $\{x_0^k\}$ *be a sequence of vectors from* $\mathbb{R}^n$*. Assume that* $\{u^k(\cdot)\}$ *converges weakly to an element* $u(\cdot) \in \mathbb{L}^2_m(0, t_f)$ *and moreover,*

$$\lim_{k\to\infty} x_0^k = x_0.$$

Then, for all $k \in \mathbb{N}$ *the following problem*

$$\begin{aligned} &\dot{x}(t) = a(t, x(t)) + b(t)u^k(t) \ \text{a.e. on } [0, t_f], \\ &x(0) = x_0^k, \end{aligned} \tag{6}$$

has a unique (absolutely continuous) solution $x^k(\cdot)$ *on the time interval* $[0, t_f]$ *and*

$$\lim_{k\to\infty} ||x(\cdot) - x^k(\cdot)||_{\mathbb{C}_n(0,t_f)} = 0,$$

where $x(\cdot)$ *is the solution of Eq. 5 corresponding to* $u(\cdot)$*.*

The complete proof of Theorem 1 can be found in Azhmyakov et al. (2011). Recall that functions $a(\cdot,\cdot)$ from Theorem 1 are often called the Caratheodory functions (see e.g., Hale 1969; Poznyak 2008).

We now return to our main dynamical model, namely, to the affine switched system (1) and consider an associated weakly convergent sequence $\{u^k(\cdot)\}$, $k \in \mathbb{N}$ of control functions $u^k(\cdot)$ from $\mathbb{L}^2_m(0,t_f)$ and a convergent sequence of initial state vectors $\{x_0^k\} \subset \mathbb{R}^n$. The corresponding sequence of the associated initial value problems has the following form

$$\dot{x}(t) = \sum_{i=1}^{r} \chi_{[t_{i-1},t_i)}(t) \left[A_{q_i}(t, x(t)) + B_{q_i}(t)u^k(t) \right] \text{ a.e. on } [0,t_f],$$
$$x(0) = x_0^k. \tag{7}$$

Theorem 1 makes it possible to formulate the continuity result for the switched system (1).

Theorem 2 *For a sequence of locations $\{q_i\} \subset \mathcal{Q}$, where $i = 1, ...r$, consider the initial value problem (1) with the* Caratheodory *functions A_{q_i}, $q_i \in \mathcal{Q}$ and the associated problems (7). Assume that $\{u^k(\cdot)\}$ converges weakly to an element $u(\cdot) \in \mathbb{L}^2_m(0,t_f)$. Moreover, let*

$$\lim_{k\to\infty} x_0^k = x_0.$$

Then, for all $k \in \mathbb{N}$ the initial value problem (7) has a unique (absolutely continuous) solutions $x^k(\cdot)$ on the time interval $[0,t_f]$ and

$$\lim_{k\to\infty} ||x(\cdot) - x^k(\cdot)||_{\mathbb{C}_n(0,t_f)} = 0,$$

where $x(\cdot)$ is the solution of Eq. 1 corresponding to $u(\cdot)$.

Proof For a given sequence of locations $\{q_i\}$, $i = 1, ...r$ the structure of problem (7) is similar to the structure of an affine switched system (1).

Consider reductions $u_i^k(\cdot)$ and $u_i(\cdot)$ of $u^k(\cdot)$ and $u(\cdot)$ on the given time intervals $[t_{i-1}, t_i)$, $i = 1, ..., r$. The weak convergence of $\{u^k(\cdot)\}$ to $u(\cdot)$ implies the weak convergence of all sequences $\{u_i^k(\cdot)\}$ to the corresponding functions $u_i(\cdot)$. Note that $u_i^k(\cdot),\ u_i(\cdot) \in \mathbb{L}^2_m(t_{i-1},t_i)$. Therefore, from Theorem 1 we deduce that

$$\lim_{k\to\infty} ||x_i(\cdot) - x_i^k(\cdot)||_{\mathbb{C}_n(t_{i-1},t_i)} = 0$$

for all $i = 1, ..., r$, where $x_i(\cdot)$ are parts of the trajectory generated by Eq. 1 in location q_i corresponding to $u_i(\cdot)$ and $x_i^k(\cdot)$ are parts of the trajectory generated by Eq. 7 in location q_i corresponding to control $u_i^k(\cdot)$. Hence

$$\lim_{k\to\infty} ||x(\cdot) - x^k(\cdot)||_{\mathbb{C}_n(0,t_f)} = 0$$

and the proof is completed. □

Note that Theorem 2 can be interpreted as follows: for the affine switched system (1) the weak convergence of controls and the convergence of the initial conditions cause the uniform convergence of the corresponding state variables. Clearly, this result can be interpreted as a kind of "robustness" of systems (1) with respect to perturbations of controls and initial state variables. Evidently, the same interpretation can be given for our first Theorem 1 with respect to the general affine non-switched control system (5). We will next apply our last result in order to study the numerical consistence of a proposed combined proximal-based computational algorithm for hybrid OCP (2). Note that a consistence of an approximation or a numerical scheme is usually understood in computational control engineering as some well-established and stable convergence properties of the corresponding approximating procedures.

4 Optimal control of affine switched systems

An effective application of a numerical method can be usually realized under some necessary technical assumptions. The main condition for a possible implementation of the proximal point regularization scheme is the convexity of the minimization problem under consideration. To put it another way, the original hybrid OCP (2) needs to have a structure of the abstract convex optimization problem (4) in a suitable Hilbert space. Evidently, not every OCP of the type (2) is equivalent to the mentioned problem (4). Therefore, we now restrict the class of OCPs and introduce the following concept.

Definition 1 If the infinite-dimensional minimization problem (3) is equivalent to a convex optimization problem (4), then we call Eq. 2 a convex optimal control problem.

Our aim is to find some constructive characterizations of the elements of the hybrid problem (2) that make it possible to establish the convexity of a given OCP (1) in the sense of Definition 1. Our next definition characterize a sub-class of the affine switched systems of the type (1).

Definition 2 We call the affine switched control system (1) a convex affine system if every functional

$$V_l(u(\cdot)) := x_l^u(t),$$
$$u(\cdot) \in \mathscr{U},\ t \in [0, t_f],\ l = 1, ..., n$$

is convex.

Note that Definition 2 is similar to the concept of the conventional convex control systems introduced in Azhmyakov and Raisch (2008). We also recall a monotonicity concept for functionals in Euclidean spaces.

Definition 3 We say that a functional $T : \Gamma \subset \mathbb{R}^n \to \mathbb{R}$ is monotonically nondecreasing if $T(\xi) \geq T(\omega)$ for all $\xi, \omega \in \Gamma$ such that $\xi_l \geq \omega_l,\ l = 1, ..., n$.

Note that the presented monotonicity concept can be expressed by introducing the standard positive cone $\mathbb{R}^n_{\geq 0}$ (the positive orthant). For a similar monotonicity concept see, for example, Azhmyakov and Raisch (2008) and Göpffert et al. (2003). A general example of a monotonically nondecreasing functional can be deduced from the Mean Value Theorem (Clarke et al. 1998), namely, a differentiable functional

$$T : \mathbb{R}^n \to \mathbb{R},\ \ \partial T(\xi)/\partial \xi_l \geq 0\ \ \forall \xi \in \mathbb{R}^n,\ l = 1, ..., n,$$

where $\partial T(\xi)\partial\xi_l$ is the k-th component of the gradient of T, is monotonically nondecreasing.

Our next result gives a constructive characterization of a convex affine system from Definition 2.

Theorem 3 *Assume that the basic assumptions from Section 2 are satisfied. Let every component $A^l_{q_i}(t,\cdot),\ l=1,...,n$ of $A_{q_i},\ q\in\mathcal{Q}$ be a convex and monotonically nondecreasing functional for every $t\in[t_{i-1},t_i),\ i=1,...,r$. Then the control system (1) is a convex affine system.*

Proof Let

$$u_3(\cdot):=\lambda u_1(\cdot)+(1-\lambda)u_2(\cdot),$$

where $u_1(\cdot),u_2(\cdot)$ are elements of $\mathcal{U}$ and $0<\lambda<1$. Consider the corresponding restrictions

$$u_{3,i}(\cdot),u_{1,i}(\cdot),u_{2,i}(\cdot)$$

of the above control functions on the time intervals $[t_{i-1},t_i),\ i=1,...,r$. Let N be a (large) positive integer number and

$$G_i^N:=\left\{t_i^0=t_{i-1},t_i^1,...,t_i^N=t_i\right\}$$

be an equidistant grid associated with every interval $[t_{i-1},t_i],\ i=1,...,r$. Here we have $t_1^0=t_0=0$ and $t_r^N=t_r=t_f$. Consider the Euler discretization method applied to the affine differential equation from Eq. 1 on every time interval $[t_{i-1},t_i),\ i=1,...,r$ (for every location $q_i\in\mathcal{Q}$) with $u(\cdot)=u_3(\cdot)$

$$\begin{aligned}
&\tilde{x}_l^{u_{3,i}}\left(t_i^{s+1}\right)=\tilde{x}_l^{u_{3,i}}\left(t_i^s\right)+\Delta_i t\left[\left(A_{q_i}\left(t_i^s,\tilde{x}^{u_{3,i}}\left(t_i^s\right)\right)\right)_l+\left(B_{q_i}u_{3,i}\left(t_i^s\right)\right)_l\right],\\
&\tilde{x}_l^{u_{3,i}}\left(t_i^0\right)=\tilde{x}_l^{u_{3,i-1}}\left(t_{i-1}^N\right),\\
&\tilde{x}_l^{u_{3,1}}\left(t_1^0\right)=(x_0)_l\,,
\end{aligned}$$

where $l=1,...,n,s=0,...,N-1$ and

$$\Delta_i t:=(t_{i-1}-t_i)/(N+1).$$

By $\tilde{x}^{u_3}(t_i^s),\ t_i^s\in G_i^N$, we denote the corresponding Euler-discretization of the solution x^{u_3} on G_i^N. From the main convexity result of Azhmyakov and Raisch (2008) (Theorem 1, p. 994) it follows that the functional

$$u_q(\cdot)\to\tilde{x}_l^{u_q}\left(t_q^s\right),\ l=1,...,n,$$

where $q = 1$ (the first location), is convex. Using this fact and convexity of every matrix-function $A_{q_i}(t, \cdot),\ q_i \in \mathcal{Q}$, for $q = 2$ we obtain

$$\begin{aligned}\tilde{x}_l^{u_{3,q}}\left(t_q^1\right) &= \tilde{x}_l^{u_{3,q}}\left(t_q^0\right) + \Delta_q t\left[\left(A_q\left(t_q^0, \tilde{x}_l^{u_{3,q}}\left(t_q^0\right)\right)\right)_l + \left(B_q\left(t_q^0\right)u_{3,q}\left(t_q^0\right)\right)\right)_l\right] \\ &\leq \lambda\tilde{x}_l^{u_{1,q}}\left(t_q^0\right) + (1-\lambda)\,\tilde{x}_l^{u_{2,q}}\left(t_q^0\right) \\ &\quad + \lambda\Delta_q t\left[\left(A_q\left(t_q^0, \tilde{x}_l^{u_{1,q}}\left(t_q^0\right)\right)\right)_l + \left(B_q\left(t_q^0\right)u_1\left(t_q^0\right)\right)\right)_l\right] \\ &\quad + (1-\lambda)\,\Delta_q t\left[\left(A_q\left(t_q^0, \tilde{x}_l^{u_{2,q}}\left(t_q^0\right)\right)\right)_l + \left(B_q\left(t_q^0\right)u_2\left(t_q^0\right)\right)\right)_l\right] \\ &= \lambda\tilde{x}_l^{u_{1,q}}\left(t_q^1\right) + (1-\lambda)\,\tilde{x}_l^{u_{2,q}}\left(t_q^1\right).\end{aligned}$$

Thus, the functional

$$u_q(\cdot) \to \tilde{x}_l^{u_q}\left(t_q^1\right),\ l = 1, ..., n,$$

where $q = 2$ is also convex. Using the convexity of every $A_{q_i}(t, \omega)$ and monotonicity properties of $(A_{q_i}(t, x))_l$ with respect to x and the convex structure of the functional $u_q(\cdot) \to \tilde{x}_l^{u_q}(t_q^1)$, we deduce that functional

$$u_q(\cdot) \to \tilde{x}_l^{u_q}\left(t_q^2\right),\ q = 2$$

is also convex for all $l = 1, ..., n$ (see Azhmyakov and Raisch 2008 for details). Applying induction on s, we obtain convexity of the functional $u_q(\cdot) \to \tilde{x}_k^{u_q}(t_q^s),\ q = 2$ for all $s = 0, ..., N-1,\ N \in \mathbb{N}$.

Now we use induction on $q \in \mathcal{Q}$ and prove the convexity of every functional

$$\begin{aligned}&u_{q_i}(\cdot) \to \tilde{x}_l^{u_{q_i}}\left(t_i^s\right)\ \ \forall s = 0, ..., N-1, \\ &i = 1, ..., r,\ q_i \in \mathcal{Q}.\end{aligned}$$

From the last fact we can deduce the convexity of the general functionals (the global Euler discretization) $u(\cdot) \to \tilde{x}_l^u(t)$, where $t \in \bigcup_{i=1,\dots,r} G_i^N$ and $u(\cdot) \in \mathcal{U}$.

The convexity of all $A_{q_i},\ q \in \mathcal{Q}$ imply the global Lipschitz continuity of these functions on any closed subset of $\mathcal{R}$ that contains $x_l^{u_3}(t), t \in [t_{i-1}, t_i)$. From this fact we deduce the uniform convergence of the considered Euler approximations for $N \to \infty$. That means

$$\lim_{N\to\infty}\ \sup_{t\in[t_{i-1},t_i)} ||\tilde{x}_l^{u_3}(t) - x_l^{u_3}(t)|| = 0$$

on every interval $[t_{i-1}, t_i)$. Therefore, for every number $\varepsilon > 0$ there exists a number $N_\varepsilon \in \mathbb{N}$ such that for all $N > N_\varepsilon$ we have

$$||x_l^{u_3}(t) - \tilde{x}_l^{u_3}(t)|| \leq \varepsilon$$

where $t \in \bigcup_{i=1,\dots,r} G_i^N$. Therefore, we obtain the following inequality

$$x_l^{u_3}(t) \leq \tilde{x}_l^{u_3}(t) + \varepsilon$$

for all $t \in [0, t_f]$. From convexity of $\tilde{x}_k^{u_3}(t)$ we next deduce that

$$x_l^{u_3}(t) \leq \lambda \tilde{x}_l^{u_1}(t) + (1-\lambda)\tilde{x}_l^{u_2}(t) + \varepsilon, \ t \in [0, t_f].$$

That means

$$x_l^{u_3}(t) \leq \lim_{N\to\infty} \left(\lambda \tilde{x}_l^{u_1}(t) + (1-\lambda)\tilde{x}_l^{u_2}(t) + \varepsilon\right) = \lambda x_l^{u_1}(t) + (1-\lambda) x_l^{u_2}(t).$$

This shows that the functional $u(\cdot) \to x_l^u,\ u(\cdot) \in \mathscr{U}$ is convex for all $l = 1, ..., n$ and all $t \in [0, t_f]$ and the given control system (1) is a convex affine system. □

We now establish the convexity of an OCP governed by a convex affine switched system of the type (1).

Theorem 4 *Let the control system from Eq. 1 be a convex affine system and let the function ϕ be convex and monotonically nondecreasing. Then the corresponding* OCP (2) *is convex.*

Proof Using the convexity and monotonicity of the function ϕ and the convexity of

$$V_l(u(\cdot)),\ u(\cdot) \in \mathscr{U},\ t \in [0, t_f],\ l = 1, ..., n,$$

we can prove that the functional J, where

$$J(u(\cdot)) = \phi\left(x^u\left(t_f\right)\right),$$

is convex. Since U is a convex subset of $\mathbb{R}^m$, the set $\mathscr{U}$ is also convex. Therefore, problem (2) is equivalent to a classical convex program of the type (4). □

Evidently, Theorem 4 provides a basis for a numerical consideration of convex OCPs from the point of view of the classical convex programming.

Let us now discuss shortly the applicability of the obtained convexity results. Consider the following simple initial value problem involving an one-dimensional Riccati differential equation

$$\dot{x}(t) = \sum_{i=1}^{r} \chi_{[t_{i-1}, t_i)}(t) \left(a_{q_i} t x^2(t) + b_{q_i} u(t)\right),$$
$$x(0) = x_0 > 0,$$

where $x(t) \in \mathbb{R}$, $u(t) \in \mathbb{R}$, $0 \leq u(t) \leq 1$ and the coefficients a_{q_i}, b_{q_i} are positive constants. Note that in this case the basic assumptions from Section 2 are fulfilled. Since

$$A_{q_i}(t, x) = a_{q_i} t x^2$$

is a locally Lipschitz function, the given initial value problem is locally solvable. Under the above positivity assumptions, $A_{q_i}(t, x)$ can be defined on the open set $\mathbb{R}_+ \times \mathbb{R}_+$. This nonlinear function is convex and monotonically nondecreasing. Since the conditions of Theorem 3 are satisfied, the control system under consideration is a convex affine system. Note that the same consideration can also be made for a more general controllable Riccati-type system (1) that admits positive solutions. Next we examine an example of a simplified controllable "explosion" model (see Zukas and Walters 1998)

$$\dot{x}(t) = \sum_{i=1}^{r} \chi_{[t_{i-1},t_i)}(t) \left(a^1_{q_i} e^{a^2_{q_i} x(t)} + a^3_{q_i} x(t) + b_{q_i} u(t) \right),$$
$$x(0) = x_0 > 0, \;\; 0 \leq u(t) \leq 1$$

with some positive scalar parameters $a^1_{q_i}$, $a^2_{q_i}$, $a^3_{q_i}$ and b_{q_i}, $q_i \in \mathcal{Q}$. This dynamic model also constitutes a convex affine system.

Finally note that our convexity results, namely, Theorem 3 and Theorem 4 primarily involve the nonlinearly affine systems (1). The convexity properties of a linear switched system and the associated OCP of the type (2) can be established under some weaker assumptions and are in fact simple consequences of the theory developed in Azhmyakov and Raisch (2008).

5 Some computational issues

It is well known that necessary optimality conditions are not only a main theoretical tool of the general optimization theory but also provide a basis for several numerical approaches in mathematical programming. The same is true in connection with various numerical solution schemes associated with OCPs. We refer to (see e.g., Azhmyakov and Raisch 2008; Azhmyakov and Noriega Morales 2010; Büskens and Maurer 2000; Cassandras et al. 2001; Maurer 1976; Polak 1997; Sakawa et al. 1981; Teo et al. 1991) for theoretical and some computational aspects. The necessary optimality conditions for convex minimization problems (3) and (4) are also sufficient. In this case a local solution of a convex minimization problem coincides with a global solution. Consider a convex OCP (2) and the associated minimization problem (3). In the case of a convex problem (2) one also can formulate necessary optimality conditions in the form of the conventional Pontryagin Maximum Principle (see e.g., Fattorini 1999; Ioffe and Tichomirov 1979). Note that in this case we need to use a special form of the Maximum Principle, namely, a hybrid version of this optimality criteria (see Azhmyakov et al. 2008; Shaikh and Caines 2007). The initial OCP (2) can be easily rewritten in the following equivalent form (see Ioffe and Tichomirov 1979)

$$\begin{aligned} &\text{minimize} \sum_{i=1}^{r} \int_{t_{i-1}}^{t_i} f^0_{q_i}(t, x(t))dt \\ &\text{subject to (Eq. 1)}, \; q_i \in \mathcal{Q}, \; u(\cdot) \in \mathcal{U}, \end{aligned} \tag{8}$$

where $f^0_q : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$, where $q \in \mathcal{Q}$, be continuously differentiable functions. We also assume that the control set has the simple box-structure:

$$U := \left\{ u \in \mathbb{R}^m \; : \; v^j_- \leq u_j \leq v^j_+, \;\; j = 1, ..., m \right\},$$

where v^j_-, v^j_+, $j = 1, ..., m$ are some constants. The application of the hybrid Maximum Principle to problem (8) involves the existence of the adjoint variable $\psi(\cdot)$ that is a solution of the corresponding boundary value problem (see Azhmyakov et al. 2008)

$$\dot{\psi}(t) = -\sum_{i=1}^{r} \chi_{[t_{i-1},t_i)}(t) \frac{\partial H_{q_i}(t, x^{\text{opt}}(t), u^{\text{opt}}(t), \psi(t))}{\partial x} \text{ a. e. on } [0, t_f],$$
$$\psi_r(t_f) = 0, \tag{9}$$

where

$$H_{q_i}(t, x, u, \psi) := \langle \psi, (A_{q_i}(t, x) + B_{q_i}(t)u \rangle - f^0_{q_i}(t, x)$$

is a "partial" Hamiltonian for the location $q_i \in \mathcal{Q}$, $x^{\text{opt}}(\cdot)$ is a solution to Eq. 1 associated with $u^{\text{opt}}(\cdot)$ and $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbb{R}^n$. Note that in difference to conventional optimal control problems the adjoint function $\psi(\cdot)$ determined in Eq. 9 is not an absolutely continuous function (has "jumps" at switching times t_i, $i = 1, ..., r$). We assume that problem (8) is regular (nonsingular) in the following sense: $\psi(t) \neq 0$ for all $t \in [0, t_f] \setminus \Upsilon$, where Υ is a subset of $[0, t_f]$ of measure zero. When solving sophisticated optimal control problems based on some necessary optimality conditions one can obtain a singular solution. Recall that there are two possible scenarios for a singularity: the irregularity of the Lagrange multiplier associated with the cost functional (Clarke et al. 1998; Ioffe and Tichomirov 1979) and the irregularity of the Hamiltonian. In the latter case the Hamiltonian is not an explicit function of the control function during a time interval. Various supplementary conditions (constraint qualifications) have been proposed under which it is possible to assert that the Lagrange Multiplier Rule (and the corresponding Maximum Principle) holds in "normal" form, i.e., that the first Lagrange multiplier is nonequal to zero. In this case the corresponding minimization problem is called regular.

Evidently, the differential equation from Eq. 9 can be rewritten as follows

$$\dot{\psi}(t) = -\sum_{i=1}^{r} \chi_{[t_{i-1},t_i)}(t)$$
$$\times \left(\left[\frac{\partial A_{q_i}\left(t, x^{\text{opt}}(t)\right)}{\partial x} \right]^T + (u^{\text{opt}}(t))^T \left[\frac{\partial B_{q_i}(t)}{\partial x} \right]^T \psi(t) + \frac{\partial f^0_{q_i}(t, x^{\text{opt}}(t))}{x} \right) \tag{10}$$

The maximality condition from the above-mentioned hybrid Maximum Principle (see Azhmyakov et al. 2008) implies the "bang-bang" structure of the optimal control $u^{\text{opt}}(\cdot)$ given by components

$$u^{\text{opt}}_j(t) = \mathbf{1}\left(\left[\psi^T(t) B_{q_i}(t)\right]_j\right) v^j_+ + \left(1 - \mathbf{1}\left(\left[\psi^T(t) B_{q_i}(t)\right]_j\right)\right) v^j_-,$$
$$\forall t \in [t_{i-1}, t_i), \; j = 1, ..., m, \tag{11}$$

where $\mathbf{1}(z) \equiv 1$ if $z \geq 0$ and $\mathbf{1} \equiv 0$ if $z < 0$ for a scalar variable z. As we can see the optimal control is a function of $\psi(\cdot)$ and functions $B_{q_i}(\cdot)$, $q_i \in \mathcal{Q}$. A practical solution of the boundary value problem (9) is usually based on an iterative numerical scheme. We refer to Poznyak (2008) for some iterative computational approaches. Using (10), the $(l+1)$-iteration $\psi^{(l)}$ for Eq. 9 can be generally written as

$$\psi^{S,(l+1)}(t) = -L_S \left(\sum_{i=1}^{r} \chi_{[t_{i-1},t_i)}(t) \left(\left[\frac{\partial A_{q_i}\left(t, x^{\text{opt}}(t)\right)}{\partial x} \right]^T \right.\right.$$
$$\left.\left. + \left(u^{\text{opt}}(t)\right)^T \left[\frac{\partial B_{q_i}(t)}{\partial x} \right]^T \psi^{S,(l)}(t) + \frac{\partial f^0_{q_i}\left(t, x^{\text{opt}}(t)\right)}{x} \right) \right), \quad (12)$$

where components of $u^{\text{opt}}(\cdot)$ are given by Eq. 11 replacing $\psi(t)$ by $\psi^{S,(l)}(t),\ l \in \mathbb{N}$. Here $L_S(w(\cdot))$ is a sequence of Riemann sums

$$L_S(w(\cdot)) := \frac{t}{S} \sum_{s=1}^{S} w\left(\frac{t}{S}s\right),\ s \in \mathbb{N}$$

for the integral $L(w(\cdot)) := \int_0^t w(t)dt$ of a piecewise continuous function $w(\cdot)$. It is well known (see e.g., Atkinson and Han 2005) that a sequence $\{L_S\}(\cdot)$ converges pointwise (weakly) to L. To put it another way, we have a pointwise convergence of the approximations given by Eq. 12 to an exact solution $\psi(\cdot)$ of Eq. 9. Since a weak convergence in the space of piecewise continuous functions coincides with the weak convergence, we conclude that a sequence of controls $\{u^{S,(l)}(\cdot)\}$, where every j-component of $u^{S,(l)}(\cdot)$ is defined as follows

$$u_j^{S,(l)}(t) = \mathbf{1}\left(\left[\left(\psi^{S,(l)}(t)\right)^T B_{q_i}(t)\right]_j\right) v_+^j + \left(1 - \mathbf{1}\left(\left[\left(\psi^{S,(l)}(t)\right)^T B_{q_i}(t)\right]_j\right)\right) v_-^j,$$
$$\forall t \in [t_{i-1}, t_i),\ \ j = 1, ..., m,$$

converges weakly to $u^{\text{opt}}(\cdot)$. From our main continuity result for switched systems, namely, from Theorem 2 we now deduce the strong convergence of the sequence $x^{S,(l)}$ (the trajectories associated with $\{u^{S,(l)}(\cdot)\}$) to an optimal trajectory x^{opt}. That means a numerical consistence (in the sense of a strong convergence of trajectories) of the usual computational schemes applied to the boundary value problem (9) associated with the optimality conditions for the hybrid optimal control problem (8).

Alternatively, the convex structure of an infinite-dimensional minimization problem (3) allows us to write necessary and sufficient optimality conditions for a convex OCP (3) in the form of the Karush-Kuhn-Tucker Theorem or in the form of an easy variational inequality. We refer to (see Azhmyakov and Raisch 2008; Ioffe and Tichomirov 1979; Polak 1997; Spellucci 1993) for a concrete form of optimality conditions in the case of a convex OCP and also for additional theoretical details. Let (3) be a convex minimization problem with a (Fréchet) differentiable cost functional J. We now describe a gradient-based approach for this purpose (see e.g., Azhmyakov and Raisch 2008; Polak 1997; Teo et al. 1991). We refer to Azhmyakov and

Raisch (2008) and Fattorini (1999) for an explicit representation of the "reduced" gradient for the conventional OCPs. Using this representation, $\nabla J(u^{\text{opt}}(\cdot))$ can be expressed as

$$\begin{aligned}
&\nabla J(u^{\text{opt}}(\cdot))(t) = -H_u(t, x^{\text{opt}}(t), u(t), p^{\text{opt}}(t)),\\
&\dot{p}^{\text{opt}}(t) = -H_x(t, x^{\text{opt}}(t), u^{\text{opt}}(t), p^{\text{opt}}(t)),\\
&p^{\text{opt}}(t_f) = -\phi_x(x^{\text{opt}}(t_f)),\\
&\dot{x}^{\text{opt}}(t) = H_p(t, x^{\text{opt}}(t), u^{\text{opt}}(t), p^{\text{opt}}(t)),\\
&x(0) = x_0,
\end{aligned} \tag{13}$$

where

$$H(t, x, u, p) = \left\langle p, \sum_{i=1}^{r} \chi_{[t_{i-1}, t_i)}(t) \; \left[A_{q_i}(t, x(t)) + B_{q_i}(t)u(t)\right] \right\rangle_{\mathbb{R}^n}$$

is the joint Hamiltonian of the OCP problem (2). By H_u, H_p, H_x we denote here the partial derivatives of H with respect to u, p and x. Moreover, $x^{\text{opt}}(\cdot)$ and $p^{\text{opt}}(\cdot)$ are the optimal state and the optimal adjoint variable corresponding to the optimal control function $u^{\text{opt}}(\cdot) \in \mathscr{U}$. Note that the gradient ∇J in Eq. 13 is computed with help of the joint Hamiltonian H to the switched OCP (2). We do not use here a sequence of "partial" Hamiltonians associated with every location for a hybrid/switched system (see e.g., Azhmyakov et al. 2008, 2009; Shaikh and Caines 2007).

We may use the notation $\langle \cdot, \cdot \rangle_{\mathbb{L}^2_m([0,t_f])}$ to denote the inner product of the space $\mathbb{L}^2_m([0, t_f])$ and formulate the following optimality criterion.

Theorem 5 *Consider a regular convex* OCP (2) *and assume that the basic assumptions from Section 2 are satisfied. Then $u^{\text{opt}}(\cdot) \in \mathscr{U}$ is an optimal solution of Eq. 2 if and only if*

$$\left\langle \nabla J(u^{\text{opt}}(\cdot)), u(\cdot) - u^{\text{opt}}(\cdot) \right\rangle_{\mathbb{L}^2_m([0,t_f])} \geq 0 \;\; \forall u(\cdot) \in \mathscr{U}. \tag{14}$$

Proof Under the basic assumptions introduced in Section 2 the following mapping

$$x^u(t_f) : \mathbb{L}^2_m([0, t_f]) \to \mathbb{R}^n$$

is Fréchet differentiable (Ioffe and Tichomirov 1979). Since ϕ is a differentiable function, the composition

$$J(u(\cdot)) = \phi(x^u(t_f))$$

is also (Fréchet) differentiable. The set of admissible controls $\mathscr{U}$ is convex, closed and bounded. Hence, we obtain the necessary and sufficient optimality condition for OCP (2) in the form (14) (see Ioffe and Tichomirov 1979). The proof is completed. □

Generally, the reduced gradient $\nabla J(u(\cdot))$ at $u(\cdot) \in \mathbb{L}^2_m([0, t_f])$ can be computed as follows

$$
\begin{aligned}
&\nabla J(u(\cdot))(t) = -H_u(t, x^u(t), u(t), p^u(t)),\\
&\dot{p}^u(t) = -H_x(t, x^u(t), u(t), p^u(t)),\\
&p^u(t_f) = -\phi_x(x^u(t_f)),\\
&\dot{x}^u(t) = H_p(t, x^u(t), u(t), p^u(t)),\\
&x(0) = x_0,
\end{aligned}
\tag{15}
$$

where $x^u(\cdot)$ and $p^u(\cdot)$ are the state and the adjoint variable corresponding to $u(\cdot)$. The presented formalism provides a basis for a variety of useful gradient-type algorithms. In the case of a convex OCP (2), these algorithms are particularly attractive in the context of an effective numerical treatment of a regularized problem (3). We now introduce the sequence of iterative prox-regularized auxiliary problems for the initial minimization problem (3) (see also Section 2)

$$
\begin{aligned}
&\text{minimize } J^l(u(\cdot)) := J(u(\cdot)) + \frac{\chi_l}{2}||u(\cdot) - u^l(\cdot)||^2_{\mathbb{L}^2_m([0,t_f])}\\
&\text{subject to } u(\cdot) \in \tilde{Q},
\end{aligned}
\tag{16}
$$

where $l = 0, 1...$ and $u^0(\cdot) \in \mathscr{U}$. Here $\{\chi_l\}$ is a given sequence with $0 < \chi_l \leq C < \infty$. Clearly,

$$
\nabla J^l(u(\cdot)) = \nabla J(u(\cdot)) + \chi_l||u(\cdot) - u^l(\cdot)||_{\mathbb{L}^2_m([0,t_f])}.
$$

where $\nabla \tilde{J}$ can be computed by using the formulas (15).

Since $J^l(u(\cdot))$ is a strongly convex functional, the minimization problem (16) has a unique solution. Under some mild assumptions, the gradient-type method converges to this solution of Eq. 12. Under the basic assumptions presented in Section 2 a sequence $\{u^l(\cdot)\}$ generated by the proximal point method (16) converges weakly to an optimal solution of the initial problem (3). Moreover, this sequence is also a minimizing sequence in the sense of Eq. 3. Combining this property of the proximal-sequence and Theorem 4 we obtain our consistence result.

Theorem 6 *Let Eq. 2 be a convex* OCP *and the corresponding initial value problem* (1) *involves the* Caratheodory *functions* A_{q_i}, $q_i \in \mathscr{Q}$. *Let* $\{u^l(\cdot)\}$ *be a sequence generated by the proximal point regularization* (16). *Assume that* $\{x^l(\cdot)\}$ *is a sequence of solutions to Eq. 1 associated with* $\{u^l(\cdot)\}$. *Then this sequence of trajectories* $\{x^l(\cdot)\}$ *converges strongly to an optimal trajectory* $x^{\text{opt}}(\cdot)$ *of Eq. 2, namely,*

$$
\lim_{l\to\infty} ||x^{\text{opt}}(\cdot) - x^l(\cdot)||_{\mathbb{C}_n(0,t_f)} = 0.
$$

Proof Since Eq. 2 is assumed to be convex and the sequence $\{u^l(\cdot)\}$ is generated by the proximal point method, we obtain a weak convergence of this sequence to an optimal control $u^{\text{opt}}(\cdot)$ for Eq. 2. Using this fact and Theorem 4, we deduce the above strong convergence of the associated trajectories $\{x^l(\cdot)\}$ to an optimal trajectory of Eq. 2. □

Finally, let us note that the application of a gradient-based approach to the prox-regularized initial OCP (2) provides a stable numerical procedure for this problem. The numerical "stability" is considered here with respect to a strong convergence of the corresponding approximative trajectories to an optimal trajectory of the given affine switched systems.

6 Conclusion

This paper studies some practically motivated switched control systems, namely, the switched systems with affine structure. We study the approximability property associated with the affine switched systems and also consider a subclass of the above dynamical models that can be characterized as a class of convex affine (switched) systems. Roughly speaking, convex control systems are those for which all components of every admissible trajectory are convex functionals with respect to the control inputs. Under some additional assumptions, OCPs associated with the above dynamics correspond to convex optimization problems in a real Hilbert space. We examine a specific type of convex affine systems with switchings and the related convex OCPs. For constructively solving convex optimal control problems we propose a conceptual computational method. This method is a combination of a conventional first-order algorithms and the classical proximal point regularization scheme. The approximability property of the general affine switched systems makes it possible to establish numerical stability of the above computational procedures for the switched OCPs under consideration.

Acknowledgements The authors thank anonymous referees for valuable remarks and suggestions from which the final version of the paper greatly benefited.

References

Atkinson K, Han W (2005) Theoretical numerical analysis. Springer, New York

Attia SA, Alamir M, Canudas de Wit C (2005) Suboptimal control of switched nonlinear systems under location and switching constraints. In: Proceedings of the 16th IFAC world congress, Prague, Czech Republick

Attia SA, Azhmyakov V, Raisch J (2010) On an optimization problem for a class of impulsive hybrid systems. Discret Event Dyn Syst 20:215–231

Axelsson H, Boccardo H, Egerstedt M, Valigi M, Wardi Y (2008a) Optimal mode-switching for hybrid systems with varying initial states. Nonlinear Anal: Hybrid Syst 2:765–772

Axelsson H, Wardi Y, Egerstedt M, Verriest EI (2008b) Gradient descent approach to optimal mode scheduling in hybrid dynamical systems. J Optim Theory Appl 136:167–186

Azhmyakov V, Raisch J (2008) Convex control systems and convex optimal control problems with constraints. IEEE Trans Automat Contr 53:993–998

Azhmyakov V, Schmidt WH (2003) Strong convergence of a proximal-based method for convex optimization. Math Methods Oper Res 57:393–407

Azhmyakov V, Attia SA, Raisch J (2008) On the maximum principle for the impulsive hybrid systems. Lecture notes in computer science, vol 4981. Springer, Berlin, pp 30–42

Azhmyakov V, Galvan-Guerra R, Egerstedt M (2009) Hybrid LQ-optimization using dynamic programming. In: Proceedings of the 2009 American control conference, St. Louis, USA, pp 3617–3623

Azhmyakov V, Noriega Morales S (2010) Proximal point method for optimal control processes governed by ordinary differential equations. Asian J Control 12:15–25

Azhmyakov V, Egerstedt M, Fridman L, Poznyak A (2011) Approximability of nonlinear affine control systems. Nonlinear Analysis: Hybrid Systems 5:275–288

Basin M, Calderon-Alvarez D (2009) Optimal controller for uncertain stochastic polynomial systems with deterministic disturbances. Int J Control 82:1435–1447

Basin M, Loukianov A, Hernandez-Fabian R (2011) A sliding mode-like optimal regulator for nonlinear polynomial systems. Int J Syst Sci. doi:10.1080/00207721.2010.545491

Branicky MS, Borkar VS, Mitter SK (1998) A unifed framework for hybrid control: model and optimal control theory. IEEE Trans Automat Contr 43:31–45

Büskens C, Maurer H (2000) SQP-methods for solving optimal control problems with control and state constraints: adjoint variables, sensitivity analysis and real-time control. J Comput Appl Math 120:85–108

Cassandras C, Pepyne DL, Wardi Y (2001) Optimal control of a class of hybrid systems. IEEE Trans Automat Contr 46:398–415

Clarke FH, Yu Ledyaev S, Stern RJ, Wolenski PR (1998) Nonsmooth analysis and control theory. Springer, New York

Ding XC, Wardi Y, Taylor D, Egerstedt M (2008) Optimization of switched-mode systems with switching costs. In: Proceedings of the 2008 American control conference, Seattle, USA, pp 3965–3970

Ding XC, Wardi Y, Egerstedt M (2009) On-line optimization of switched-mode dynamical systems. IEEE Trans Automat Contr 54:2266–2271

Egerstedt M, Wardi Y, Axelsson H (2006) Transition-time optimization for switched-mode dynamical systems. IEEE Trans Automat Contr 51:110–115

Fattorini HO (1999) Infinite dimensional optimization and control theory. Cambridge University Press, Cambridge

Garavello M, Piccoli B (2005) Hybrid necessary priniple. SIAM J Control Optim 43:1867–1887

Göpffert A, Riahi H, Tammer C, Zălinescu C (2003) Variational methods in partially ordered spaces. Springer, New York

Hale J (1969) Ordinary differential equations. Wiley, New York

Hiriart-Urruty JB, Lemarechal C (1993) Convex analysis and minimization algorithms. Springer, Berlin

Ioffe AD, Tichomirov VM (1979) Theory of extremal problems. North Holland, Amsterdam

Isidori A (1989) Nonlinear control systems. Springer, New York

Kaplan A, Tichatschke R (1994) Stable methods for ill-posed variational problems. Akademie Verlag, Berlin

Liberzon D (2003) Switching in systems and control. Birkhäuser, Boston

Lygeros J (2003) Lecture notes on hyrid systems. Cambridge University Press, Cambridge

Martinet B (1970) Regularisation d'inequations variationelles par approximations successives, Revue Francaise Informat. Recherche Operationnelle 4:154–159

Maurer H (1976) Numerical solution of singular control problems using multiple shooting techniques. J Optim Theory Appl 18:235–257

Moor T, Raisch J (1999) Supervisory control of hybrid systems within a behavioural framework. Syst Control Lett 38:157–166

Polak E (1997) Optimization. Springer, New York

Poznyak A (2008) Advanced mathematical tools for automatic control engineers. Elsevier, Amsterdam

Rockafellar RT (1970) Convex analysis. Princeton University Press, Princeton

Rockafellar RT (1976) Monotone operators and the proximal point algorithm. SIAM J Control Optim 14:877–898

Sakawa Y, Shindo Y, Hashimoto Y (1981) Optimal control of a rotary crane. J Optim Theory Appl 35:535–557

Shaikh MS, Caines PE (2007) On the hybrid optimal control problem: theory and algorithms. IEEE Trans Automat Contr 52:1587–1603

Solodov MV, Svaiter BF (2000) Forcing strong convergence of proximal point iterations in a Hilbert space. Math Program Ser A 87:189–202

Spellucci P (1993) Numerische Verfahren der Nichtlinearen Optimierung. Birkhäuser, Basel

Sussmann HJ (1999) A maximum principle for hybrid optimization. In: Proceedings of the 38th IEEE conference on decision and control, Phoenix, USA, pp 425–430

Teo KL, Goh CJ, Wong KH (1991) A unifed computational approach to optimal control problems. Wiley, New York
Utkin V (1992) Sliding modes in control and optimization. Springer, Berlin
Verriest E, Delmotte F, Egerstedt M (2004) Optimal impulsive control of point delay systems with refractory period. In: Proceedings of the 5th IFAC workshop on time delay systems, Leuven, Belgium
Xu X, Antsaklis P (2003a) Optimal control of hybrid autonomous systems with state jumps. In: Proceedings of the American control conference, Denver, USA, pp 5191–5196
Xu X, Antsaklis P (2003b) Quadratic optimal control problems for hybrid linear autonomous systems with state jumps. In: Proceedings of the American control conference, Denver, USA, pp 3393–3398
Zukas JA, Walters WP (1998) Explosive effects and applications. Springer, New York

Vadim Azhmyakov received his M.S. degree in Aerospace Control Technology from the Faculty for Aerospace Engineering the Moscow Technical University “N. E. Bauman” (Moscow, Russia) in 1989. His Ph.D. degree in Control Engineering has been completed in 1994 at the Institute of Control Sciences (Moscow, Russia). In addition, Dr. V. Azhmyakov received his Habilitation in Applied Mathematics from the University of Greifswald (Greifswald, Germany) in 2006.

From 1989 until 1990 he was a software engineer in the Construction Office for Missiles Technology (Reutov, Russia) and from 1991 until 1994 he was worked as a researcher at the Institute of Control Sciences (Moscow, Russia). From October 1996 to April 1999, he was a Lecturer at the Department of Mathematical Statistics, University of Bayreuth (Bayreuth, Germany). From 1999 until 2005 Dr. Vadim Azhmyakov was an Assistant Professor at the Faculty for Numerical Analysis, University of Greifswald (Greifswald, Germany). From 2005 until 2006 he was a Researcher at the Department of Electrical Engineering, University of Magdeburg (Magdeburg, Germany) and from 2006 until 2007 a Researcher at the Department of Electrical Engineering of the Technical University of Berlin (Berlin, Germany). Since 2007 Dr. Vadim Azhmyakov is with the Department of Control Automation the Center for Advanced Studies and Research in Mexico City, Mexico. His research interests include hybrid/switched systems theory, computational methods in optimization and control of complex dynamic systems, nonlinear and variational analysis.

Michael V. Basin received his Ph.D. degree in Physical and Mathematical Sciences with major in Automatic Control and System Analysis from the Moscow Aviation Institute in 1992. His work experience includes Senior Scientist position in the Institute of Control Science (Russian Academy of Sciences) in 1992–96, Visiting Professor position in the University of Nevada at Reno in 1996–97, and Full Professor position in the Autonomous University of Nuevo Leon, Mexico, from 1998.

Starting from 1992, Dr. Basin published more than 80 research papers in international referred journals and more than 130 papers in Proceedings of the leading IEEE and IFAC conferences and symposiums. He is the author of the monograph "New Trends in Optimal Filtering and Control for Polynomial and Time-Delay Systems," published by Springer. His works are cited more than 650 times. Dr. Basin has supervised 7 doctoral and 4 master's theses (all of them are successfully defended) and is currently supervising 6 doctoral and 2 master's theses. He serves as the Editor-in-Chief of Journal of The Franklin Institute, an Associate Editor of the IEEE Control System Society Conference Editorial Board, Automatica, International Journal of Innovative Computing, Information and Control, International Journal of Systems Science, Dynamics of Discrete, Continuous, and Impulsive Systems, Series B, and other journals, as a session chairman at a series of the leading IEEE and IFAC conferences, and as a reviewer for a number of leading international journals and conferences in the area of automatic control. Dr. Basin serves as a member of the IEEE Control System Society Technical Committee on Intelligent Control, Program Committee member of IEEE Conference on Decision and Control 2008, IEEE Conference on Control Applications 2009, IEEE 11th Workshop on Variable Structure Systems 2010, IEEE International Conferences on Innovative Computing, Information and Control 2006–2010, Mexican Automatic Control Conferences 2006, 2008. Dr. Basin was awarded a title of Highly Cited Researcher by Thomson Reuters (International Science Institute), the publisher of Science Citation Index, in 2009; he is a Senior Member of the IEEE Control Systems Society and a regular member of the Mexican Academy of Sciences. His research interests include optimal filtering and control problems, stochastic systems, time-delay systems, identification, sliding mode control and variable structure systems.

Jörg Raisch is a professor at Technische Universitaet Berlin (Germany), where he heads the Control Systems Group within the Department of Electrical Engineering and Computer Science. He is also head of the Systems and Control Theory Group at the Max Planck Institute for Dynamics of Complex Technical Systems in Magdeburg (Germany). He studied Engineering Cybernetics and Control Systems at Stuttgart University and UMIST, Manchester (UK). He got his Ph.D. and "Habilitation", both from Stuttgart University, in 1991 and 1998, respectively. From 1991–1993 he was a postdoc in the Systems Control Group at the University of Toronto (Canada). From 2000–2006 he was a professor at the Otto-von-Guericke University Magdeburg, where he headed the automatic control lab. His research interests are in hybrid systems and hierarchical control and include biomedical control and chemical process control applications.