16th IEEE International Conference on Control Applications
Part of IEEE Multi-conference on Systems and Control
Singapore, 1-3 October 2007

State Jump Optimization for a Class of Hybrid Autonomous Systems

S.A. Attia, V. Azhmyakov and J. Raisch

Abstract— **In this contribution, optimization of state jumps for a class of hybrid systems is considered. Basically, the control variables to be determined are the amounts of jump in the continuous states such that a corresponding cost functional is minimized. Based on a variational approach, necessary conditions of optimality are first established. The problem is then cast as a parametric optimization problem where the gradient information is derived. Finally and under some assumptions, convergence to the optimal solution of a conceptual algorithm is established. A brief discussion on the main implementation issues is also included.**

Index Terms— **Hybrid systems, optimization, gradient methods, state jump, switching.**

I. INTRODUCTION

For a long time, discrete-continuous dynamical interactions have been recognized as a major challenge in the process control area. The emergence of a hybrid systems modelling framework is now providing a new perspective for some important problems. The ability to operate hybrid systems in an optimal way remains a challenging task. Indeed, for the general setting of hybrid systems, one has to deal not only with the infinite dimensional optimization problems related to the continuous dynamics, but also with a potential combinatorial explosion related to the discrete part. In this context and with focus on particular classes, many schemes have been proposed to tackle the problem. Some are based on a newly elaborated condition of optimality see e.g., [1], [2],[3], others are more related to semi-classical approaches see e.g., [4], [5], [6], [7]. In the last years, there has been a revival of gradient based methods see e.g., [8], [9], [10], [11], [12]. This fact is due to their intuitive interpretation, reliability and the existence of well established convergence results. It is the aim of this contribution to extend this approach to a particular class of hybrid systems with autonomous switching and controlled state jumps. This class arises most frequently in the area of process control and is a result of constraints satisfaction and material processing see e.g.,[13] for a typical benchmark problem.

Autonomous hybrid systems with uncontrolled state jumps have been considered in [14]. The objective is to find the sequence of jump instants such that a cost functional is minimized. For that purpose, the authors develop a second order scheme that is further specialized to linear systems in [15]. Recently, Verriest and coworkers [16] have considered another class of hybrid systems. This consists of autonomous systems with state delay where both the jump magnitudes and instants are unknowns. Based on variational arguments, necessary conditions of optimality are derived and used in a first order scheme, see [17] for an application. In this contribution, we consider a related problem where no delay is present on the state and where the switching is autonomous. The problem is motivated by a chemical process control application, namely a preferential crystallization process used to separate enantiomers see [18] for the physical aspects and [19] for more details on the control problems.

The paper organization is as follows : in Section 2 the problem is stated formally. Section 3 is devoted to the statement of the necessary conditions of optimality. In Section 4, the gradient formulas are derived and a conceptual algorithm together with some convergence properties are stated. Finally, some conclusions and suggestions for future work are given in Section 5.

S.A. Attia and J. Raisch are both with Fachgebiet Regelungssysteme, TU-Berlin Einsteinufer 17, 10587 Berlin and the Max-Planck-Institut fuer Dynamik komplexer technischer Systeme Sandtorstr.1, 39106 Magdeburg, Germany, attia@ieee.org, Raisch@control.tu-berlin.de

V. Azhmyakov is with Fachgebiet Regelungssysteme, TU-Berlin Einsteinufer 17, 10587 Berlin, Germany, Azhmyakov@control.tu-berlin.de

II. PROBLEM FORMULATION

We consider the following class of hybrid systems termed autonomous impulsive (see e.g., [20] for the general modelling framework).

Definition 1: An autonomous impulsive hybrid system is a collection

$$\mathcal{H} = (\mathcal{Q},\ \mathcal{E},\ \mathcal{X},\ \mathcal{F},\ \mathcal{G},\ \mathcal{R})$$

where

- $\mathcal{Q} = \{q_0, q_1 \dots, q_Q\}$ is a finite set of locations
- $\mathcal{E} \subseteq \mathcal{Q} \times \mathcal{Q}$ is a set of edges
- $\mathcal{X} = \{\mathcal{X}_q\}_{q\in\mathcal{Q}}$ is a collection of state spaces where for all $q \in \mathcal{Q}$, $\mathcal{X}_q$ is an open subset of $\mathbb{R}^n$, n is the dimension of the state space
- $\mathcal{F} = \{f_q\}_{q\in\mathcal{Q}}$ is a collection of vector fields. For all $q \in \mathcal{Q}$, $f_q : \mathcal{X}_q \to \mathbb{R}^n$
- $\mathcal{G} = \{\mathcal{G}_e\}_{e\in\mathcal{E}}$ is a collection of guards. For all possible transitions $e = (q_i,\ q_j) \in \mathcal{E}$, $\mathcal{G}_e \subset \mathcal{X}_{q_i}$
- $\mathcal{R} = \{\mathcal{R}_e\}_{e\in\mathcal{E}}$ is a collection of reset maps. For all $e = (q_i,\ q_j) \in \mathcal{E}$, $\mathcal{R}_e : \mathcal{G}_e \to 2^{\mathcal{X}_{q_j}}$ where $2^{\mathcal{X}_{q_j}}$ denotes the power set of $\mathcal{X}_{q_j}$

We assume that the vector fields f_q are smooth enough (see assumptions below), and that the sets $\mathcal{G}_e$ and $\mathcal{R}_e$ are nonempty for all $e \in \mathcal{E}$. An execution is as follows: starting from an initial condition (x_0, q_0) the continuous state evolves according to the autonomous differential equations

$$\dot{x}(t) = f_{q_0}(x(t)) \tag{1}$$

The discrete state $q(\cdot) = q_0$ remains constant as long as the trajectory does not reach a guard $\mathcal{G}_{(q_0,q_1)}$. In our set

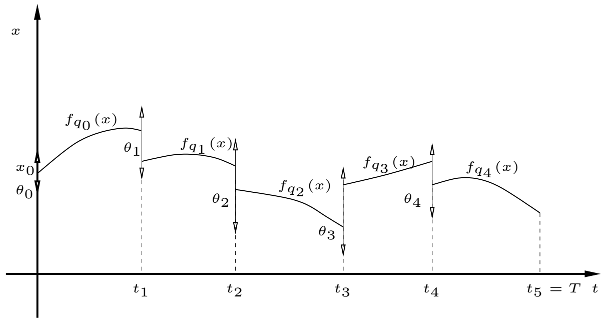

Fig. 1. An example of execution with $K = 4$ switches.

up, once this guard is reached, $q(\cdot)$ will switch from q_0 to q_1,[1] at the same time the continuous state gets reset to some value according to the map $\mathcal{R}_{(q_0,q_1)}$ and the whole process is repeated. Next, we suppose that the guards can be described by smooth $(n-1)$ dimensional surfaces in $\mathbb{R}^n$

$$\mathcal{G}_e = \{x \mid S_e(x) = 0\}, \qquad \text{for all } e \in \mathcal{E} \tag{2}$$

and that the reset maps are linear maps of the form, for all $e \in \mathcal{E}$,

$$\mathcal{R}_{(q_{i-1},q_i)}(x) = x + \theta_i \tag{3}$$

with θ_i belonging to Θ_i a compact subset of $\mathbb{R}^n$.

Remark 1: The variables θ_i may represent an amount of material to be added at each end of processing cycle.

The dynamic optimization problem of interest can now be formulated.

Problem 1: Under a fixed switching sequence of locations $\{q_i\}_i$, solve the following optimization problem

$$\min_\theta J[\theta] := \sum_{i=0}^{K-1} \int_{t_i}^{t_{i+1}} L(x(\tau))d\tau + \int_{t_K}^{t_{K+1}} L(x(\tau))d\tau \tag{4}$$

such that

$$\dot{x}(t) = f_{q(t)}(x(t)) \tag{5}$$
$$q(t) = q_i, \qquad t \in [t_i, t_{i+1}) \tag{6}$$
$$x(t_0^+) = x(t_0) + \theta_0 \tag{7}$$
$$x(t_{i+1}^+) = x(t_{i+1}) + \theta_{i+1}, \ S_{(q_i,q_{i+1})}(x(t_{i+1})) = 0 \tag{8}$$
$$\dot{x}(t) = f_{q_K}(x(t)), \qquad t \in [t_K, t_{K+1}] \tag{9}$$

with $x(t_0) = x_0$, $K \in \mathbb{Z}^+$ is the total number of switches, the vector θ denotes the $n \times (K+1)$ dimensional vector $(\theta_0' \dots \theta_K')'$, $L : \mathbb{R}^n \to \mathbb{R}^+$ is a supplied cost function, t_0 and $t_{K+1} = T$ are both finite and given.

Remark 2: Note that although the sequence of locations is fixed, the switching instants are not defined à priori. This follows from the degree of freedom allowed on the state jump i.e., variation of these quantities induces variations of the switching instants. Note also the absence of endpoint constraints. This allows the operation of the system within a prespecified amount of time, see Figure 1 for an example of execution.

[1] Note that this is different from the standard hybrid automaton framework, where switches are enforced by the continuous state violating invariants.

The following assumptions are made.

Assumptions

- A1 For all q_i in $\mathcal{Q}$, $\mathcal{X}_{q_i} = \mathbb{R}^n$
- A2 For all q_i in $\mathcal{Q}$, the functions f_{q_i} are continuously differentiable
- A3 L is a twice continuously differentiable function
- A4 There exists a constant M such that $\|f_{q_i}(x)\| < M$ for all $x \in \mathcal{X}_{q_i}$ and $q_i \in \mathcal{Q}$
- A5 For $i = 0, \dots K-1$, $\inf_{y \in \mathcal{G}_{(q_i,q_{i+1})}} \|x(t_i^+)) - y\| \geq \zeta > 0$
- A6 At the switching instants t_{i+1}, the transversality condition $\nabla_x S_{(q_i,q_{i+1})} f_{q_i} \neq 0$ holds

Remark 3: Assumptions $A5$-$A6$ are introduced for well posedness of any execution. Indeed, they avoid Zeno phenomena and sliding behaviors from occurring.

Problem 1 could be seen as a collection of initial value optimization subproblems. Using classical approaches, the subproblems could be solved separately but nothing guarantees that the trajectory obtained by concatenation of the different solutions; although optimal; is optimal this follows from the principle of optimality. The link between the different subproblems is provided by a set of necessary conditions stated in the next section.

III. NECESSARY CONDITIONS OF OPTIMALITY

In this section, necessary conditions of optimality of a solution to Problem 1 are derived. The arguments used are of variational type (see e.g.,[21],[22] for a basic material on the Euler-Lagrange theory). Let us first define the Hamiltonian associated to location q_i as

$$H_{q_i}(x, \lambda) = L(x) + \lambda' f_{q_i}(x) \tag{10}$$

where λ denotes the adjoint variables. We then have the following result

Proposition 1: If θ^* is an interior optimal solution to Problem 1 under assumptions $A1 - A6$ and $x^*(t)$ its corresponding state trajectory for $t \in [t_0, T]$, then there exists a nontrivial adjoint $\lambda^*(t)$ and multipliers π_i^* such that the following equations hold

$$\dot{\lambda}^*(t)' = -\nabla_x H_{q_i}^*, \qquad t \in [t_i, t_{i+1}) \tag{11}$$

At the switching instants t_{i+1}, the following jump conditions are satisfied

$$\lambda^*|_{t_{i+1}^+} = \lambda^*|_{t_{i+1}} - \pi_i^* \nabla_x S'_{(q_i,q_{i+1})}\Big|_{t_{i+1}} \tag{12}$$
$$H_{q_{i+1}}^*|_{t_{i+1}^+} = H_{q_i}^*|_{t_{i+1}} \tag{13}$$

for $i = 0, \dots K-1$, with

$$\lambda^*|_{t_{K+1}} = \mathbf{0} \tag{14}$$

and

$$\nabla_\theta J[\theta^*] = \mathbf{0} \tag{15}$$

Proof: The augmented Lagrangian can be written as

$$\mathcal{L} = \sum_{i=0}^{K} \left[\int_{t_i}^{t_{i+1}} \left(H_{q_i}(x, \lambda) - \lambda' \dot{x} \right) d\tau + \pi_i S_{(q_i,q_{i+1})}|_{t_{i+1}} \right] \tag{16}$$

where for simplicity time dependence is dropped. The increment of $\mathcal{L}$ with respect to x can be written

$$\Delta_x \mathcal{L} = \mathcal{L}(x+h) - \mathcal{L}(x) \quad (17)$$

where h is a continuously differentiable function of time. Development of the these terms lead to the following

$$\begin{aligned}\Delta_x \mathcal{L} \;=\; & \sum_{i=0}^{K} \Big[\int_{t_i+dt_i}^{t_{i+1}+dt_{i+1}} \Big(H_{q_i}(x+h,\lambda) - \lambda'(\dot{x}+\dot{h})\Big) d\tau \\ & - \int_{t_i}^{t_{i+1}} \left(H_{q_i}(x,\lambda) - \lambda'\dot{x}\right) d\tau + \pi_i \Delta_x S_{(q_i,q_{i+1})}|_{t_{i+1}})\Big]\end{aligned} \quad (18)$$

with dt_i a small time increment (the existence of which follows from the smoothness assumptions). After integration by part of the second term under the integral sign and rearrangement, Equation (18) can be written as

$$\begin{aligned}\Delta_x \mathcal{L} = & \sum_{i=0}^{K} \Big[\int_{t_i}^{t_{i+1}} \Big(H_{q_i}(x+h,\lambda) - H_{q_i}(x,\lambda) + \dot{\lambda}'h\Big) d\tau \\ & - \left(H_{q_i}(x,\lambda) - \lambda'\dot{x}\right)|_{t_i^+} dt_i + \left(H_{q_i}(x,\lambda) - \lambda'\dot{x}\right)|_{t_{i+1}^+} dt_{i+1} \\ & -\lambda' h|_{t_i^+}^{t_{i+1}} + \pi_i \Delta_x S_{(q_i,q_{i+1})}|_{t_{i+1}}\Big]\end{aligned} \quad (19)$$

Now, using Taylor's theorem, we obtain (up to first order) the following expression

$$\begin{aligned}\Delta_x \mathcal{L} \;=\; & \sum_{i=0}^{K} \Big[\int_{t_i}^{t_{i+1}} \Big(\nabla_x H_{q_i} + \dot{\lambda}'\Big) h d\tau \\ & - \left(H_{q_i}(x,\lambda) - \lambda'\dot{x}\right)|_{t_i^+} dt_i \\ & + \left(H_{q_i}(x,\lambda) - \lambda'\dot{x}\right)|_{t_{i+1}} dt_{i+1} \\ & -\lambda' h|_{t_i^+}^{t_{i+1}} + \pi_i \frac{\partial S_{(q_i,q_{i+1})}}{\partial x}\Big|_{t_{i+1}} dx(t_{i+1})\Big]\end{aligned} \quad (20)$$

where $dx(t_{i+1})$ is the exact state variation at the instant t_{i+1}. Following simple geometrical arguments, it can be approximated to the first order by the following expression

$$dx(t_{i+1}) = h(t_{i+1}) + \dot{x}(t_{i+1}) dt_{i+1} \quad (21)$$

Using (21), the first order variation of the augmented Lagrangian $\mathcal{L}$ can be written as

$$\begin{aligned}\delta_x \mathcal{L} = & \sum_{i=0}^{K} \Big[\int_{t_i}^{t_{i+1}} \Big(\nabla_x H_{q_i} + \dot{\lambda}'\Big) h d\tau - \\ & \Big(H_{q_i}|_{t_{i+1}} dt_{i+1} - H_{q_i}|_{t_i^+} dt_i\Big) \\ & +\lambda|_{t_i^+} dx(t_i) + \Big(\pi_i \nabla_x S_{(q_i,q_{i+1})}\Big|_{t_{i+1}} - \lambda|_{t_{i+1}}\Big) dx(t_{i+1})\Big]\end{aligned} \quad (22)$$

Along the optimal pair (x^*, θ^*) the following is satisfied

$$\delta_x \mathcal{L} = 0 \quad (23)$$

After rearrangement and using the fact that the optimal problem is without terminal constraints but with fixed initial and final time, the desired result (11)-(14) follows and ends the proof of the first part. Using similar ideas, the condition (15) follows, this ends the proof. ■

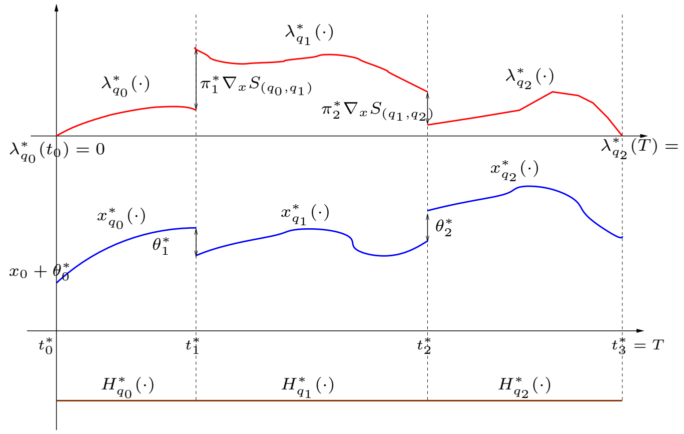

Fig. 2. A figure showing how the optimal solution to problem 1, under assumptions $A1 - A6$, looks like. The implementable Algorithm 1 in the Section IV shows how to find the optimal parameters, and thus the optimal trajectroy, iteratively by solving a set of initial value problems.

Remark 4: For the general case i.e., a reset map described by $x(t_i^+) = \psi(x(t_i))$ and provided that some smoothness requirements hold the gradient of ψ would appear in the necessary conditions.

Remark 5: Stronger necessary conditions of optimality have been reported in the literature in the sense that the result (12)-(14) in Proposition 1 follows from e.g., [23], [24]. However, the arguments used here are of the variational type which make them valid for, smooth in location, hybrid systems under the aforementioned assumptions.

Remark 6: The conditions stated above characterize a local minimum in the sense that the optimal trajectory is compared only to trajectories that hits the switching surfaces and have the same switching sequence of locations.

IV. A GRADIENT BASED APPROACH

The system is an impulsive system with a free initial condition. The basic idea is to cast Problem 1 into a parameter optimization framework and then to compute, as required by the necessary conditions of optimality, the gradient of the cost functional. Gradient descent techniques can then be used to compute the optimal parameter values. Before the gradient formula is stated in a proposition, a lemma concerning the sensitivity of the states is given.

Lemma 1: The sensitivity $\Delta^{\theta_i}_{q_i} x(\cdot)$ of the state trajectory $x(\cdot)$, corresponding to the dynamics (5)-(9) under assumptions $A1 - A6$, w.r.t. the vectors θ_i can be computed as a solution to the following variational equation

$$\Delta^{\theta_i}_{q_i} \dot{x}(t) = \frac{\partial f_{q_i}}{\partial x} \Delta^{\theta_i}_{q_i} x(t), \qquad t \in [t_i, t_{i+1}) \quad (24)$$

under the following initial conditions

$$\begin{aligned}\Delta_{q_0}^{\theta_0} x(t_0) &= I_{n\times n} &(25)\\ \Delta_{q_{i+1}}^{\theta_{i+1}} x(t_{i+1}^+) &= \Delta_{q_i}^{\theta_i} x(t_{i+1}) + I_{n\times n} + \\ &\left(f_{q_i}(x(t_{i+1})) - f_{q_{i+1}}(x(t_{i+1}^+))\right)\Delta^{\theta_{i+1}} t_{i+1} &(26)\end{aligned}$$

with

$$\Delta^{\theta_{i+1}} t_{i+1} = -\frac{\nabla_x S_{(q_i,q_{i+1})}\Big|_{t_{i+1}}}{\nabla_x S_{(q_i,q_{i+1})}\Big|_{t_{i+1}} f_{q_i}(x(t_{i+1}))}\Delta_{q_i}^{\theta_i} x(t_{i+1}) \quad (27)$$

Proof: See [25]. ■

Equations (24)-(27) allow for the computation of the sensitivity trajectories. This result is used to establish the following proposition.

Proposition 2: The gradient of the cost functional J corresponding to Problem 1, under Assumptions $A1 - A6$, can be computed as follows for $i = 0, \dots K-1$

$$\begin{aligned}\nabla_{\theta_0} J &= \lambda'|_{t_0} &(28)\\ \nabla_{\theta_{i+1}} J &= \lambda'|_{t_{i+1}} - \pi_{i+1}^* \nabla_x S_{(q_i,q_{i+1})} \times \\ &\left(\Delta_{q_{i+1}}^{\theta_{i+1}} x(t_{i+1}^+) + f_{q_{i+1}}(x(t_{i+1}^+))\Delta^{\theta_{i+1}} t_{i+1}\right) &(29)\end{aligned}$$

Proof: See Appendix 1. ■

Note that a formula similar to Equation (28) is already known in the literature and is used as a measure of the influence that the initial conditions can have on conventional optimal control problems. As will be described in more details in the forthcoming paragraphs, the computational complexity in this type of problems is non negligible. Indeed, the difficulty consists in solving a boundary value problem of a special type. We will now state a conceptual algorithm and show that, under some additional assumptions, it converges to the infinimum of the optimization problem 1 under assumptions $A1 - A6$.

Conceptual Algorithm 1:

1) Choose an admissible parameter vector $\theta^{(0)}$ and set $k = 0$
2) Compute the trajectory $x^{(k)}(\cdot)$ and the corresponding adjoint $\lambda^{(k)}(\cdot)$ such that the conditions (12)-(14) are satisfied,
3) Update the parameter vector $\theta^{(k)}$ using the gradient information (28)-(29) in a gradient projection algorithm, set $k := k+1$ and go to step 2

This is summarized in the following results.

Proposition 3: If θ^* is an accumulation point of the sequence $\{\theta^{(k)}\}_k$ generated by the Conceptual Algorithm 1 then it is a stationary point ($\nabla J[\theta^*] = 0$).

Proof: The proof of Proposition 3 follows using descent properties see e.g.,[26],[27]. ■

In the following paragraph, some important implementation issues are discussed.

Denote by P_Θ the projection operator on the collection of sets Θ_i. An implementable version of the preceding Conceptual Algorithm can be stated as follows

Algorithm 1:

Step 0 Choose parameters β, μ as positive real numbers from the set $(0,1)$, a small positive real number ϵ and an admissible parameter vector $\theta^{(0)}$. Set $k = 0$

Step 1 Compute the trajectory $x^{(k)}(\cdot)$ by forward integrating the state Equations (5)-(9) under the specified initial condition. Let $t^{(k)}$ be the resulting sequence of switching instants $\{t_i\}_i$

Step 2 Compute the sensitivity trajectories using (24) and the corresponding initial conditions (25)-(26) and Equation (27)

Step 3 Backward integrate the adjoint $\lambda^{(k)}(\cdot)$ using equation (11) with the terminal condition (14). At the switching instants $t^{(k)}$ compute the multipliers $\pi_i^{(k)}$ in (12) such that equation (13) is enforced. Update the adjoint at the switching instants $t^{(k)}$ using Equation (12) and the so far computed multipliers $\pi^{(k)}$.

Step 4 Compute the gradient using (28)-(29). Update the parameter vector $\theta^{(k+1)}$

$$\theta^{(k+1)} = P_\Theta(\theta^{(k)} - \gamma^{(k)}\nabla_\theta J) \quad (30)$$

with $\gamma^{(k)} = \mu^{j_k}$ and j_k as the smallest nonnegative integer j satisfying the following inequality

$$J[\theta^{(k)} - \mu^j \nabla_\theta J] - J[\theta^{(k)}] \le -\beta\mu^j \|\nabla_\theta J\|^2 \quad (31)$$

Step 5 *If* $J[\theta^{(k)}] - J[\theta^{(k+1)}] \le \epsilon$ *Then* STOP *Else* set $k := k+1$ and *go to* Step 1

A good choice of the algorithm parameters β, μ and ϵ depends on the problem at hand. Numerical experience has shown the universality of some values see e.g.[28] for indications. The computations in Step 1 involves the solution of $(K+1)$ Initial Value Problems (IVP). A particular attention should be paid to the detection of the switching instants. This can be done using the event location capabilities of Matlab IVP solvers [29]. Step 2 involves the solution of a linear time varying system that should pose no major difficulties. In Step 3, the multipliers are computed such that the Hamiltonian continuity condition is enforced. Recall here that a closed form solution of the multipliers can be found by combining Equations (12) and (13). Step 4 is the costliest part of the algorithm. Indeed, evaluation of the cost functional in the right-hand side of inequality (31) makes internal calls to Step1-Step3. However, the number of such calls is finite. In figure 3 and figure 4 are schematized two successive iterations of Algorithm 1 applied to a first-order system.

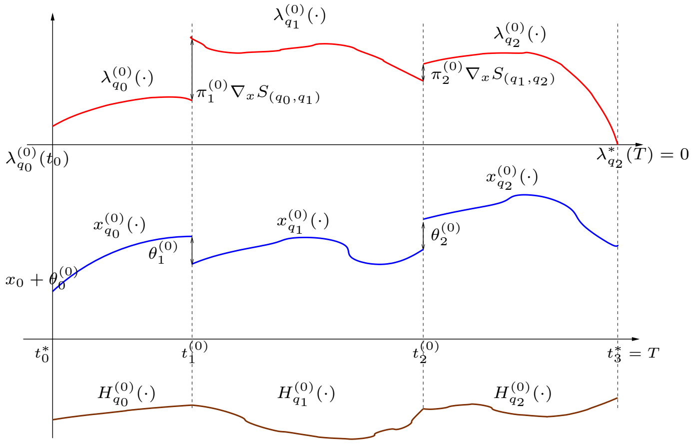

Fig. 3. Iteration 1 of the algorithm under $K = 2$ switches for a generic example.

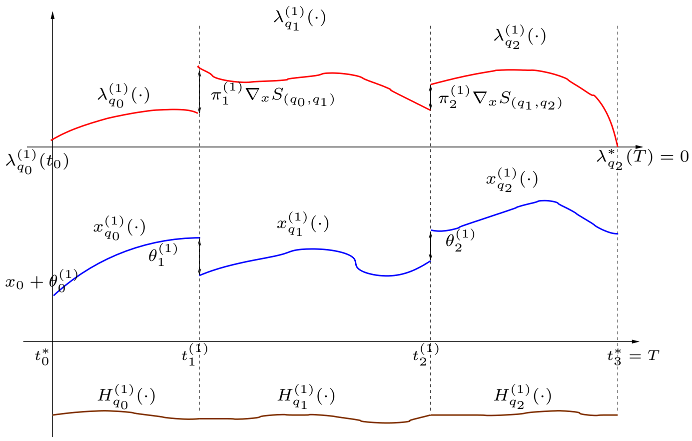

Fig. 4. Iteration 2 of the algorithm under $K = 2$ switches for a generic example. Notice here, the change of the switching instants with respect to those in Figure 3 although these were not explictly considered as optimization variables. See Remark 2 for a thorough discussion. Notice also that the initial value of the adjoint is converging to the point $\mathbf{0}$, this is a direct consequence of the necessary conditions of optimality see equation (15) and equation (28).

V. CONCLUSIONS

This paper addresses optimization problems for a class of hybrid systems. Using variational principles, necessary conditions of optimality and a gradient formulas are derived. A conceptual algorithm is then presented together with convergence analysis and implementation issues.

Future work will concentrate on numerical experimentations with this algorithm and on the study of the cyclic operation i.e., periodic switching sequence of locations. This is actually the way the preferential crystallization process is operated.

***Appendix 1**: Proof of Proposition 2:* *Proof:* Let ρ_i be an arbitrary but a fixed vector in $\mathbb{R}^n$ with $i \in \mathcal{K} = \{0,\ldots,K\}$. Define

$$\theta_i(\epsilon) = \theta_i + \epsilon\rho_i \tag{32}$$

Where ϵ is an arbitrary small real number. The pertubed cost $J[\theta(\epsilon)]$ can be decomposed and written as follows

$$J[\theta(\epsilon)] = \sum_{i=0}^{K}\left[\int_{t_i(\epsilon)}^{t_{i+1}(\epsilon)} \left[H_{q_i}(x(\tau;\epsilon),\lambda(\tau)) - \lambda'(\tau) f_{q_i}(x(\tau;\epsilon))\right] d\tau\right]$$

Computing the derivative of (33) with respect to ϵ, one obtains the following expression

$$\begin{aligned}
&\frac{d}{d\epsilon}J[\theta(\epsilon)] = \\
&\sum_{i=0}^{K}\Bigg[\left[H_{q_i}(x(t_{i+1};\epsilon),\lambda(t_{i+1})) - \lambda'|_{t_{i+1}} f_{q_i}(x(t_{i+1};\epsilon))\right] \times \\
&\frac{d}{d\epsilon}t_{i+1}(\epsilon) \\
&-\left[H_{q_i}(x(t_i^+;\epsilon),\lambda(t_i^+)) - \lambda'|_{t_i^+} f_{q_i}(x(t_i^+;\epsilon))\right]\frac{d}{d\epsilon}t_i(\epsilon) \\
&+\int_{t_i(\epsilon)}^{t_{i+1}(\epsilon)} \frac{\partial}{\partial x}\left[H_{q_i}(x(\tau;\epsilon),\lambda(\tau)) - \lambda'(\tau) f_{q_i}(x(\tau;\epsilon))\right] \times \\
&\frac{\partial x(\tau;\epsilon)}{\partial \epsilon}d\tau\Bigg]
\end{aligned} \tag{33}$$

Evaluating expression (33) at $\epsilon = 0$ and using proposition 1 (equation (11)) we obtain

$$\begin{aligned}
&\left.\frac{d}{d\epsilon}J[\theta(\epsilon)]\right|_{\epsilon=0} = \\
&\sum_{i=0}^{K}\Bigg[\left[H_{q_i}|_{t_{i+1}} - \lambda'|_{t_{i+1}} f_{q_i}(x(t_{i+1}))\right]\Delta^\epsilon t_{i+1} \\
&-\left[H_{q_i}|_{t_i^+} - \lambda'|_{t_i^+} f_{q_i}(x(t_i^+))\right]\Delta^\epsilon t_i \\
&-\int_{t_i}^{t_{i+1}}\left[\dot{\lambda}'(\tau)\Delta^\epsilon_{q_i}x(\tau) + \lambda'(\tau)\Delta_{q_i}\dot{x}(\tau)\right]d\tau\Bigg]
\end{aligned} \tag{34}$$

By using the fact, that the term under the integral sign can be written for any i in $\mathcal{K}$ as

$$\begin{aligned}
&\int_{t_i}^{t_{i+1}} \frac{d}{d\tau}\left[\lambda'(\tau)\Delta^\epsilon_{q_i}x(\tau)\right]d\tau = \\
&\lambda'|_{t_{i+1}}\Delta^\epsilon_{q_i}x(t_{i+1}) - \lambda'|_{t_i^+}\Delta^\epsilon_{q_i}x(t_i^+)
\end{aligned} \tag{35}$$

We then obtain the simplified expression below

$$\begin{aligned}
&\left.\frac{d}{d\epsilon}J[\theta(\epsilon)]\right|_{\epsilon=0} = \\
&\sum_{i=0}^{K-1}\Bigg[\left[H_{q_i}|_{t_{i+1}} - \lambda'|_{t_{i+1}} f_{q_i}(x(t_{i+1}))\right]\Delta^\epsilon t_{i+1} \\
&-\left[H_{q_i}|_{t_i^+} - \lambda'|_{t_i^+} f_{q_i}(x(t_i^+))\right]\Delta^\epsilon t_i \\
&-\lambda'|_{t_{i+1}}\Delta^\epsilon_{q_i}x(t_{i+1}) + \lambda'|_{t_i^+}\Delta^\epsilon_{q_i}x(t_i^+)\Bigg]
\end{aligned} \tag{36}$$

After rearrangement and using that $\lambda|_T = \mathbf{0}$ (see equation (14) from Proposition 1), t_0 and T are fixed, the last expression (36) may be written as

$$\frac{d}{d\epsilon}J[\theta(\epsilon)]\Big|_{\epsilon=0} = \lambda'|_{t_0^+}\Delta^{\epsilon}_{q_0}x(t_0^+) + \sum_{i=0}^{K-1}\Bigg[\Big(H_{q_i}|_{t_{i+1}} - H_{q_{i+1}}|_{t_{i+1}^+} + \lambda'|_{t_{i+1}^+} f_{q_{i+1}}(x(t_{i+1}^+)) - \lambda'|_{t_{i+1}} f_{q_i}(x(t_{i+1}))\Big)\Delta^{\epsilon}t_{i+1} + \lambda'|_{t_{i+1}^+}\Delta^{\epsilon}_{q_{i+1}}x(t_{i+1}^+) - \lambda'|_{t_{i+1}}\Delta^{\epsilon}_{q_i}x(t_{i+1})\Bigg] \quad (37)$$

Simplification of the equation (37) can be now carried out using equation (**??**) from Lemma 1 that is written as follows

$$\Delta^{\epsilon}_{q_{i+1}}x(t_{i+1}^+) = \Delta^{\epsilon}_{q_i}x(t_{i+1}) + \rho_{i+1} - (f_{q_{i+1}}(x(t_{i+1}^+)) - f_{q_i}(x(t_{i+1})))\Delta^{\epsilon}t_{i+1} \quad (38)$$

Extracting the term $\Delta^{\epsilon}_{q_i}x(t_{i+1})$ from the preceding equation (38) and using it in equation (37). This gives after simplification

$$\frac{d}{d\epsilon}J[\theta(\epsilon)]\Big|_{\epsilon=0} = \lambda'|_{t_0^+}\Delta^{\epsilon}_{q_0}x(t_0^+) + \sum_{i=0}^{K-1} H_{q_i}|_{t_{i+1}} - H_{q_{i+1}}|_{t_{i+1}^+} + \Big(\lambda'|_{t_{i+1}+} - \lambda'|_{t_{i+1}}\Big)\Big(\Delta^{\epsilon}_{q_{i+1}}x(t_{i+1}^+) + f_{q_{i+1}}(x(t_{i+1}^+))\Delta^{\epsilon}t_{i+1}\Big) + \lambda'|_{t_{i+1}}\rho_{i+1} \quad (39)$$

Now using Proposition 1 the terms involving the Hamiltonians can be eliminated and along the optimal trajectory the gradient can be written as

$$\frac{d}{d\epsilon}J[\theta(\epsilon)]\Big|_{\epsilon=0} = \nabla_{\theta}J[\theta]\rho = \lambda'(t_0^+)\Delta^{\theta_0}_{q_0}x(t_0^+)\rho_0 - \sum_{i=0}^{K-1} \pi_i^*\nabla_x S_{(q_i,q_{i+1})}\Big(\Delta^{\theta_{i+1}}_{q_{i+1}}x(t_{i+1}^+) + f_{q_{i+1}}(x(t_{i+1}^+))\Delta^{\theta_{i+1}}t_{i+1}\Big)\rho_{i+1} + \lambda'|_{t_{i+1}}\rho_{i+1}$$

With $\rho = (\rho_0' \;\; \rho_1' \dots \rho_K')'$. Direct decomposition and simplification since the arbitrariness of ρ gives the desired result and end the proof. ■

REFERENCES

[1] M. Shahid Shaikh and P. E. Caines. On the optimal control of hybrid systems : Optimization of trajectories, switching times, and location schedules. In *HSCC 2003*, 2003.

[2] P. E. Caines and S. Shahid Shaikh. New results in optimality zone hybrid optimal control algorithms: Halting and geometry. In *Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems*, pages 619–624, Kyoto, Japan, Jul 2006.

[3] A. Rantzer. On relaxed dynamic programming in switching systems. *IEE Proceedings on Control Theory and Applications*, 153(5):567–574, Sep 2006.

[4] S. A. Attia, M. Alamir, and C. Canudas de Wit. Suboptimal control of switched nonlinear systems under location and switching constraints. In *16th IFAC World Congress*, 2005.

[5] M. Alamir and S. A. Attia. An efficient algorithm to solve optimal control problems for nonlinear switched hybrid systems. In *IFAC NOLCOS*, 2004.

[6] S. Hedlund and A. Rantzer. Convex dynamic programming for hybrid systems. *IEEE Transactions on Automatic Control*, 47(9):1536–1540, Sep 2002.

[7] X. Xu and P. J. Antsaklis. Results and perspectives on computational methods for optimal control of switched systems. In O. Maler and A. Pnueli, editors, *HSCC 2003*, Lecture Notes in Computer Science, pages 540–555. Springer Verlag, 2003.

[8] J. Lu, L. Z. Liao, A. Nerode, and J. H. Taylor. Optimal control of systems with continuous and discrete states. In *Proceedings of the IEEE Conference on Decision and Control*, pages 2292–2297, Texas, USA, 1993.

[9] C. G. Cassandras, D. L. Pepyne, and Y. Wardi. Generalized gradient algorithms for hybrid system models of manufacturing systems. In *Proceedings of the IEEE Conference on Decision and Control*, pages 2627–2632, Tampa, USA, Dec 1998.

[10] V. Azhmyakov and J. Raisch. A gradient-based approach to a class of hybrid optimal control problems. In *Proceedings of the Conference on Analysis and Design of Hybrid Systems ADHS*, pages 89–94, Alghero, Italy, Jun 2006.

[11] M. Egerstedt, Y. Wardi, and H. Axelsson. Transition-time optimization for switched-mode dynamical systems. *IEEE Transactions on Automatic Control*, 51(1):110–115, Jan 2006.

[12] H. Axelsson, Y. Wardi, and M. Egerstedt. Convergence of gradient-descent algorithms for mode-scheduling problems in hybrid systems. In *Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems*, pages 625–627, Kyoto, Japan, Jul 2006.

[13] M. Alamir. A benchmark for optimal control solvers for hybrid nonlinear systems. *Automatica*, 42:1593–1598, Sep 2006.

[14] X. Xu and P. Antsaklis. Optimal control of hybrid autonomous systems with state jumps. In *Proceedings of the American Control Conference*, pages 5191–5196, Denver, USA, Jun 2003.

[15] X. Xu and P. Antsaklis. Quadratic optimal control problems for hybrid linear autonomous systems with state jumps. In *Proceedings of the American Control Conference*, pages 3393–3398, Denver, USA, Jun 2003.

[16] E. Verriest, F. Delmotte, and M. Egerstedt. Optimal impulsive control of point delay systems with refractory period. In *Proceedings of the 5th IFAC workshop on Time Delay Systems*, Leuven, Belgium, Sep 2004.

[17] E. Verriest, F. Delmotte, and M. Egerstedt. Control of epidemics by vaccination. In *Proceedings of the American Control Conference*, pages 985–990, Portland, USA, Jun 2005.

[18] M. P. Elsner, D. F. Mendez, and A. E. Muslera A. Seidel-Morgenstern. Experimental study and simplified mathematical description of preferential crystallisation. *Chirality*, (17):183–195, 2005.

[19] J. Raisch, U. Vollmer, and I. Angelov. Control problems in batch crystallization of enantiomers. In *Computer Methods and Systems, CMS*, Krakow, Poland, Nov 2005.

[20] S. N. Simic, K. H. Johansson, S. Sastry, and J. Lygeros. Towards a geometric theory of hybrid systems. In N. Lynch and B. Krogh, editors, *HSCC 2000*, Lecture Notes in Computer Science, pages 421–436. Springer Verlag, 2000.

[21] I. M. Gelfand and S. V. Fomin. *Calculus of Variations*. Prentice-Hall, 1963.

[22] A. E. Bryson and Y-C. Ho. *Applied optimal control: optimization, estimation and control*. Hemisphere Publishing Corp., 1975.

[23] H. J. Sussmann. A nonsmooth hybrid maximum principle. In D. Aeyels, F. Lamnabhi-Lagarrigue, and A. J. van der Schaft, editors, *Stability and Stabilization of Nonlinear Systems*, number 246 in Lecture Notes in Control and Information Sciences, pages 325–354. Springer Verlag, 1999.

[24] V. Azhmyakov, S.A. Attia, D. Gromov, and J. Raisch. Necessary optimality conditions for a class of hybrid optimal control problems. In *Hybrid Systems: Computation and Control*, Pisa, Italy, Avril 3-5 2007.

[25] S. A. Attia, V. Azhmyakov, and J. Raisch. On some aspects of hybrid optimal control. Technical report, MPI and TU-Berlin, 2007.

[26] D. P. Bertsekas. *Nonlinear programming*. Athena Scientific, 1995.

[27] B. N. Pshenichny and Y. M. Danilin. *Numerical Methods In Extremal Problems*. Mir Publishers, 1982.

[28] E. Polak. *Computational Methods in Optimization: A Unified Approach*. Academic press, 1971.

[29] L.F. Shampine and S. Thompson. Event location for ordinary differential equations. *Comp. Math with Appl.*, (5-6):43–54, 2000.