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.

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

- \( \mathcal{Q} = \{q_0, q_1, \ldots, q_J\} \) 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 \( q \in \mathcal{Q}, f_q : \mathcal{X}_q \to \mathbb{R}^n \)
- \( \mathcal{G} = \{G_e\}_{e \in \mathcal{E}} \) is a collection of guards. For all possible transitions \( e = (q_i, q_j) \in \mathcal{E}, G_e \subset \mathcal{X}_{q_i} \)
- \( \mathcal{R} = \{R_e\}_{e \in \mathcal{E}} \) is a collection of reset maps. For all \( e = (q_i, q_j) \in \mathcal{E}, R_e : \mathcal{G} \to \mathcal{X}_{q_j} \) where \( \mathcal{X}_{q_j} \) denotes the power set of \( \mathcal{X}_{q_i} \)

We assume that the vector fields \( f_q \) are smooth enough (see assumptions below), and that the sets \( \mathcal{G} \) and \( \mathcal{R} \) 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
\[
\dot{x}(t) = f_{q_0}(x(t)) \quad (1)
\]

The discrete state \( q(\cdot) = q_0 \) remains constant as long as the trajectory does not reach a guard \( G_{(q_0, q_1)} \). In our set
There exists a constant \( K \) for all \( \theta_0 \) in the set \( \{ \theta \} \) of initial states. For all \( x \) in the set \( \mathcal{X} \) of states, the functions \( f_q \) are continuously differentiable. Let \( L \) be a twice continuously differentiable function. There exists a constant \( M \) such that \( \| f_q(x) \| < M \) for all \( x \) in \( \mathcal{X} \) and \( q \) in \( \mathcal{Q} \). For \( i = 0, \ldots, K - 1 \), \( \inf_{y \in \theta(x(t_i), q_{i+1})} \| x(t_{i+1}) - y \| \geq \zeta > 0 \) 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

\[
H_{q_i}(x, \lambda) = L(x) + \lambda^i f_q(x, \lambda)
\]

where \( \lambda \) denotes the adjoint variables. We then have the following result:

**Proposition 1:** If \( \theta^* \) 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 \( \pi^*_i \) such that the following equations hold

\[
\lambda^*(t)^{\prime} = -\nabla_x H_{q_i}^*, \quad t \in [t_i, t_{i+1}]
\]

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

\[
\lambda^*[t^+_{i+1}] = \lambda^*[t_{i+1}] - \pi^*_i \nabla_x S_{q_i, q_{i+1}}[t_{i+1}]
\]

and

\[
\nabla \theta J[\theta^*] = 0
\]

**Proof:** The augmented Lagrangian can be written as

\[
L = \sum_{i=0}^{K} \int_{t_i}^{t_{i+1}} \left( H_{q_i}(x, \lambda) - \lambda^i \dot{x} \right) dt + \pi_i S_{q_i, q_{i+1}}[t_{i+1}]
\]

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 \( f_q \) are continuously differentiable

A3 \( L \) is a twice continuously differentiable function

A4 There exists a constant \( M \) such that \( \| f_q(x) \| < M \) for all \( x \) in \( \mathcal{X}_{q_i} \), and \( q_i \) in \( \mathcal{Q} \)

A5 For \( i = 0, \ldots, K - 1 \), \( \inf_{y \in \theta(x(t_i), q_{i+1})} \| x(t_{i+1}) - y \| \geq \zeta > 0 \)

A6 At the switching instances \( t_i \), the transversality condition \( \nabla_x S_{q_i, q_{i+1}}[t_{i+1}] = 0 \) holds.

Remark 2: Note that this is different from the standard hybrid automaton framework, where switches are enforced by the continuous state violating invariants.

1. Note that this is different from the standard hybrid automaton framework, where switches are enforced by the continuous state violating invariants.
where for simplicity time dependence is dropped. The increment of $L$ with respect to $x$ can be written

$$
\Delta x L = L(x + h) - L(x)
$$

(17)

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

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

(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

$$
\Delta x L = \sum_{i=0}^{K} \left[ \int_{t_i}^{t_{i+1}} \left( \nabla \Delta \theta_i \cdot H_{q_i} + \lambda' \dot{h} \right) d\tau - (H_{q_i}(x, \lambda) - \lambda' \dot{x}) \right]_{t_i}^{t_{i+1}} dt_i + \left( H_{q_i}(x, \lambda) - \lambda' \dot{x} \right)_{t_i}^{t_{i+1}} dt_i
$$

(19)

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

$$
\Delta x L = \sum_{i=0}^{K} \left[ \int_{t_i}^{t_{i+1}} \left( \nabla \Delta \theta_i \cdot H_{q_i} + \lambda' \dot{h} \right) d\tau - (H_{q_i}(x, \lambda) - \lambda' \dot{x}) \right]_{t_i}^{t_{i+1}} dt_i + \left( H_{q_i}(x, \lambda) - \lambda' \dot{x} \right)_{t_i}^{t_{i+1}} dt_i
$$

(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

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

(21)

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

$$
\delta x L = \sum_{i=0}^{K} \left[ \int_{t_i}^{t_{i+1}} \left( \nabla \Delta \theta_i \cdot H_{q_i} + \lambda' \dot{h} \right) d\tau - (H_{q_i}(x, \lambda) - \lambda' \dot{x}) \right]_{t_i}^{t_{i+1}} dt_i + \left( H_{q_i}(x, \lambda) - \lambda' \dot{x} \right)_{t_i}^{t_{i+1}} dt_i
$$

(22)

Along the optimal pair $(x^*, \theta^*)$ the following is satisfied

$$
\delta x L = 0
$$

(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.

**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 $\psi$ 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 $\Delta^\theta_i x(\cdot)$ of the state trajectory $x(\cdot)$, corresponding to the dynamics (5)-(9) under assumptions $A1 - A6$, w.r.t. the vectors $\theta_i$ can be computed as a solution to the following variational equation

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

(24)

under the following initial conditions

$\Delta^\theta_i x(t_i) = 0$. 

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 trajecory, iteratively by solving a set of initial value problems.
Choose parameters

\[ \Delta \theta^0 x(t_0) = I_{n \times n} \]
\[ \Delta \theta_{i+1} x(t_{i+1}) = \Delta \theta_i x(t_{i+1}) + I_{n \times n} + \left( f_{q_i}(x(t_{i+1})) - f_{q_{i+1}}(x(t_{i+1})) \right) \Delta \theta_{i+1} t_{i+1} \]

with

\[ \Delta \theta_{i+1} t_{i+1} = - \frac{\nabla_x S_{(q_i, q_{i+1})}}{\nabla_x S_{(q_i, q_{i+1})} \mid t_{i+1}} \Delta \theta_i x(t_{i+1}) \]

**Algorithm 1:**

**Step 0** Choose parameters $\beta$, $\mu$ as positive real numbers from the set $(0, 1)$, a small positive real number $\epsilon$ 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 $\dot{\theta}^{(k)}$ be the resulting sequence of switching instants $\{t_i\}$.

**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 $\dot{\theta}^{(k)}$ compute the multipliers $\pi_i^{(k)}$ in (12) such that equation (13) is enforced. Update the adjoint at the switching instants $\dot{\theta}^{(k)}$ using Equation (12) and the so far computed multipliers $\pi_i^{(k)}$.

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

\[ \theta^{(k+1)} = P_{\Theta} \left( \theta^{(k)} - \gamma^{(k)} \nabla_{\theta} J \right) \] with $\gamma^{(k)} = \mu^j k$ and $j_k$ as the smallest nonnegative integer $j$ satisfying the following inequality

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

A good choice of the algorithm parameters $\beta$, $\mu$ and $\epsilon$ 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 applications of Algorithm 1 applied to a first-order system.
cost $J[\theta(\epsilon)]$ can be decomposed and written as follows

$$
\frac{d}{d\epsilon} J[\theta(\epsilon)] = 0
$$

Computing the derivative of (33) with respect to $\epsilon$, one obtains the following expression

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

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

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

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

\[
\int_{t_i}^{t_{i+1}} \frac{d}{d\tau} [\lambda'(\tau)\Delta q_{q_i}\lambda(\tau)] d\tau = \lambda'[t_{i+1}]\Delta q_{q_i}x(t_{i+1}) - \lambda'[t_i]\Delta q_{q_i}x(t_i) 
\]

We then obtain the simplified expression below

\[
\frac{d}{d\epsilon} J[\theta(\epsilon)] \bigg|_{\epsilon=0} = \sum_{i=0}^{K-1} \left[H_{q_i}[t_{i+1} - \lambda'[t_{i+1}]f_{q_i}(x(t_{i+1})]] \Delta t_{i+1} - \left[H_{q_i}[t_i] - \lambda'[t_i]f_{q_i}(x(t_i)) \right] \Delta t_i - \lambda'[t_{i+1}]\Delta q_{q_i}x(t_{i+1}) + \lambda'[t_i]\Delta q_{q_i}x(t_i) 
\]

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

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 experiments 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 $\rho_i$ be an arbitrary but a fixed vector in $\mathbb{R}^n$ with $i \in K = \{0, \ldots, K\}$. Define

\[
\theta_i(\epsilon) = \theta_i + \epsilon \rho_i 
\]

Where $\epsilon$ is an arbitrary small real number. The perturbed
\[
\frac{d}{d\epsilon} J(\theta(\epsilon)) \Big|_{\epsilon=0} = \lambda'_{i+1} \Delta_{q_0}^{\epsilon} x(t^{+}_{i+1}) + \sum_{i=0}^{K-1} \left( H_{q_i}|t_{i+1} - H_{q_{i+1}} |t_{i+1}^{+} + \lambda'_{i+1} f_{q_{i+1}}(x(t^{+}_{i+1})) - \lambda'_{i+1} f_{q_i}(x(t^{+}_{i+1})) \right) \Delta_{t_{i+1}}^{\epsilon} + \\
\lambda'_{i+1} \Delta_{q_{i+1}}^{\epsilon} x(t^{+}_{i+1}) - \lambda'_{i+1} \Delta_{q_i}^{\epsilon} x(t^{+}_{i+1}) \right] \tag{37}
\]

Simplification of the equation (37) can now be carried out using equation (22) from Lemma 1 that is written as follows

\[
\Delta_{q_{i+1}}^{\epsilon} x(t^{+}_{i+1}) = \Delta_{q_i}^{\epsilon} x(t^{+}_{i}) + \rho_{i+1} - (f_{q_{i+1}}(x(t^{+}_{i+1})) - f_{q_i}(x(t^{+}_{i+1}))) \Delta_{t_{i+1}}^{\epsilon} 
\]

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

\[
\frac{d}{d\epsilon} J(\theta(\epsilon)) \Big|_{\epsilon=0} = \lambda'_{i+1} \Delta_{q_0}^{\epsilon} x(t^{+}_{i+1}) + \\
\sum_{i=0}^{K-1} H_{q_i}|t_{i+1} - H_{q_{i+1}} |t_{i+1}^{+} + \\
\left( \lambda'_{i+1} \Delta_{t_{i+1}}^{\epsilon} x(t^{+}_{i+1}) + f_{q_{i+1}}(x(t^{+}_{i+1})) \Delta_{t_{i+1}}^{\epsilon} \right) - \lambda'_{i+1} f_{q_i}(x(t^{+}_{i+1})) \Delta_{t_{i+1}}^{\epsilon} + \\
\lambda'_{i+1} \rho_{i+1} \tag{39}
\]

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

\[
\frac{d}{d\epsilon} J(\theta(\epsilon)) \Big|_{\epsilon=0} = \nabla_{\theta} J(\theta) = \lambda'_{i+1} \Delta_{q_0}^{\epsilon} x(t^{+}_{i+1}) \rho_0 - \\
\sum_{i=0}^{K-1} \left( \pi_i^{+} \nabla_{x} S_{Q_i}(t^{+}_{i+1}) \Delta_{q_{i+1}}^{\epsilon} x(t^{+}_{i+1}) + f_{q_{i+1}}(x(t^{+}_{i+1})) \Delta_{t_{i+1}}^{\epsilon} \right) \rho_{i+1} + \\
\lambda'_{i+1} \rho_{i+1} \tag{39}
\]

With \( \rho = (\rho_0, \rho_1, \ldots, \rho_K)' \). Direct decomposition and simplification since the arbitrariness of \( \rho \) gives the desired result and end the proof.

REFERENCES


