Hybrid Function Approximation : A Variational Approach

Florent Delmotte, Magnus Egerstedt and Erik Verriest

{florent,magnus,erik.verriest}@ece.gatech.edu
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332, USA

***Abstract*— In this paper we study the problem of representing trajectories through a concatenation of trajectories obtained from solving different differential equations. These trajectories are moreover allowed to undergo discontinuous jumps in order to make the approximation better. Our solution is based on a variational approach, in which both jump heights as well as switch times are taken as optimization parameters. The main application for this method is data compression, where the original trajectory is represented through mode sequences, switch times, and jump heights.**

I. INTRODUCTION

Given a reference curve, a standard problem in signal processing, tracking and data compression is how to best represent/track this curve. In this paper we approach this problem using a hybrid approximation scheme. The main idea is to temporally concatenate a number of outputs from dynamical systems within a given class in order to minimize the tracking error. Moreover, as the resulting system switches between different modes, the state may experience, for a given cost, a discrete jump if that improves the result. In other words, switching times as well as jump heights constitute the free parameters in the optimization problem.

Similar problems can be found in two main areas, namely approximation theory and optimal control of hybrid systems. To our knowledge, no problems in approximation theory (e.g. [4], [5]) using discontinuous functions use a cost on the discontinuities. By adding costs on jumps, our problem formulation specifically targets hybrid systems realizations where state jumps often involve quick and resource intense tasks (e.g. remotely driving an autonomous robot from one place to another, switching the value of a potential in an electrical circuit, switching gears on a motorized vehicle, etc). Also, to our knowledge, no optimal control problems for autonomous hybrid systems with state jumps explicitly concern function approximation. One could argue that replacing the continuous state x of the hybrid system with $\tilde{x} = x - f(t)$ in such problems, where $f(t)$ is the function to be approximated, would allow us to extend the use of already available methods [1], [2], [6] to approximation applications. However, by doing so, we would modify the nature of the system and violate essential assumptions such as linearity or time-invariancy, on which such methods are based.

Note however that the idea to use hybrid systems to represent curves obtained from complex and possibly unknown systems is certainly not new. Gain scheduling represents such an effort, and the pioneering work in [7] triggered a flurry of activities in this area. In this paper we take a variational view of the problem and provide gradients of the cost with respect to the switching times and jump heights. As a result, locally optimal solutions can be obtained and numerical examples testify to the method's viability.

II. HYBRID FUNCTION APPROXIMATION

A. Problem Formulation

Given a continuous signal $x : [t_0, t_f] \to \mathbb{R}^n$, we propose to approximate x by a piecewise continuous function $\tilde{x}$ such that:

$$\dot{\tilde{x}}(t) = f_i(\tilde{x}(t), \tau_{i-1}), \quad t \in [\tau_{i-1}, \tau_i), \quad i = 1 : N+1 \tag{1}$$

$$\tilde{x}(\tau_i^+) = \tilde{x}(\tau_i^-) + G_i(x(\tau_i), \tilde{x}(\tau_i^-), u_i, \tau_i), \quad i = 1 : N \tag{2}$$

where the *derivative functions* $\{f_i\}_{i=1}^{N+1}$ and the *jump functions* $\{G_i\}_{i=1}^{N}$ are given smooth functions. The variables $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_N) \in [t_0, t_f]^N$ (note that $\tau_0 = t_0$ and $\tau_{N+1} = t_f$ are given) and $\mathbf{u} = (u_1, \ldots, u_N) \in \mathbb{R}^N$ are the control variables, and they are to be chosen so that the cost function:

$$J(\boldsymbol{\tau}, \mathbf{u}) = \sum_{i=1}^{N+1} \int_{\tau_{i-1}}^{\tau_i} L_i(x(t), \tilde{x}(t))dt + \sum_{i=1}^{N} K_i(x(\tau_i), \tilde{x}(\tau_i^-), u_i, \tau_i) \tag{3}$$

is minimized.

The arguments of the functions G_i allow for numerous applications, in particular the use of free jumps (for which $G_i = u_i$) or resets (for which $G_i = x(\tau_i) - \tilde{x}(\tau_i^-)$).
Note that we have not introduced any requirements on the initial condition for the approximation function. For the rest of the proof, we will assume that $\tilde{x}(t_0) = \tilde{x}_0$ is given. In particular, the examples provided in Section III use

$\tilde{x}(t_0) = x(t_0)$. We could also imagine a case where $\tilde{x}(t_0)$ is a free parameter. This can be done via the introduction of a non-movable jump at t_0 (hence new functions G_0 and a K_0 would be added to the problem). Similarly, a final cost at t_f could be added via the introduction of a function K_{N+1}.

B. *Gradient Computation*

Here, we compute the directional derivative, also known as the Gateaux derivative, of the cost function defined in Equation (3). The derivative is computed at the point $(\boldsymbol{\tau}, \boldsymbol{u})$ and along the direction $(\boldsymbol{\theta}, \boldsymbol{\nu})$:

$$\nabla_{\boldsymbol{\theta},\boldsymbol{\nu}} J(\boldsymbol{\tau}, \mathbf{u}) = \lim_{\epsilon \to 0} \frac{J(\boldsymbol{\tau}+\epsilon\boldsymbol{\theta}, \mathbf{u}+\epsilon\boldsymbol{\nu}) - J(\boldsymbol{\tau}, \mathbf{u})}{\epsilon} \tag{4}$$

For ease of reading, we will use the alternate notation:

$$\nabla J = \lim_{\epsilon \to 0} \frac{J_\epsilon - J_0}{\epsilon} \tag{5}$$

where J_0 will be referred as the *unperturbed system*, and J_ϵ as the *perturbed system*. Note this is quite similar to the program in [8], [1].

First, we adjoin the dynamical constraints of Equations (1) and (2) into Equation (3) by defining Lagrange multipliers $\lambda_i : [\tau_{i-1}, \tau_i] \to \mathbb{R}^n$ and $\mu_i \in \mathbb{R}^n$ respectively. For the unperturbed system, we get:

$$\begin{aligned} J_0 = &\sum_{i=1}^{N+1} \int_{\tau_{i-1}}^{\tau_i} [L_i(x, \tilde{x}) + \lambda_i'(f_i(\tilde{x}, \tau_{i-1}) - \dot{\tilde{x}})] dt \\ &+ \sum_{i=1}^{N} [K_i(x(\tau_i), \tilde{x}(\tau_i^-), u_i, \tau_i) \\ &\qquad + \mu_i'(G_i(x(\tau_i), \tilde{x}(\tau_i^-), u_i, \tau_i) - \Delta\tilde{x}|_{\tau_i})] \\ = &\, J_0^{(1)} + J_0^{(2)}. \end{aligned} \tag{6}$$

By introducing the functions

$$M_i = K_i + \mu_i' G_i, \tag{7}$$

we get a simplified notation for $J_0^{(2)}$ in Equation (6):

$$J_0^{(2)} = \sum_{i=1}^{N} [M_i(x(\tau_i), \tilde{x}(\tau_i^-), u_i, \tau_i) - \mu_i' \Delta\tilde{x}|_{\tau_i})]. \tag{8}$$

Now consider the perturbed system where an ϵ-small variation of the control variables along $(\boldsymbol{\theta}, \boldsymbol{\nu})$ (i.e. variations $\tau_i \to \tau_i + \epsilon\theta_i$ and $u_i \to u_i + \epsilon\nu_i$ for $i = 1, \ldots, N$), is applied. Figure 1 illustrates the differences between the approximation functions of the unperturbed system $\tilde{x}$ and the perturbed one $\tilde{x}_\epsilon$. Because $\tilde{x}$ jumps at the instants $\{\tau_i\}$ while $\tilde{x}_\epsilon$ jumps at the instants $\{\tau_i + \epsilon\theta_i\}$, their respective derivatives on the intervals $(\tau_i, \tau_i + \epsilon\theta_i)$ are f_{i+1} and f_i. Also, $\tilde{x}_\epsilon$ can be expressed in terms of $\tilde{x}$:

$$\tilde{x}_\epsilon(t) = \tilde{x}(t) + \epsilon\eta(t) - \sum_{i=1}^{N} \Delta\tilde{x}|_{\tau_i} \xi_{[\tau_i, \tau_i+\epsilon\theta_i]}(t) \tag{9}$$

where $\eta(t)$ is a smooth function on the intervals $(\tau_i + \epsilon\theta_i, \tau_{i+1} + \epsilon\theta_{i+1})$ and ξ is the indicator function , i.e.

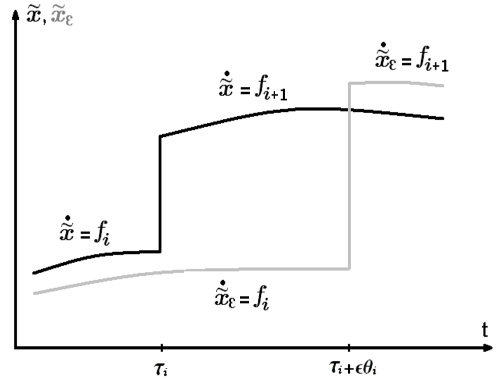

Fig. 1. Zoom on a jump at τ_i

$\xi_{[a,b]} = 1$ on $[a, b]$ and 0 elsewhere.

The equivalent of Equation (6) for the perturbed system is:

$$\begin{aligned} J_\epsilon = &\sum_{i=1}^{N+1} \int_{\tau_{i-1}+\epsilon\theta_{i-1}}^{\tau_i+\epsilon\theta_i} [L_i(x, \tilde{x}_\epsilon) + \lambda_{i(t)}'(f_i(\tilde{x}_\epsilon, \tau_{i-1} + \epsilon\theta_{i-1}) - \dot{\tilde{x}}_\epsilon)] dt \\ &+ \sum_{i=1}^{N} [M_i(x(\tau_i + \epsilon\theta_i), \tilde{x}_\epsilon(\tau_i + \epsilon\theta_i^-), u_i + \epsilon\nu_i, \tau_i + \epsilon\theta_i) \\ &\qquad - \mu_i' \Delta\tilde{x}_\epsilon|_{\tau_i+\epsilon\theta_i})] \\ = &\, J_\epsilon^{(1)} + J_\epsilon^{(2)} \end{aligned} \tag{10}$$

Recall that the Lagrange multipliers λ_i are defined on the intervals $[\tau_{i-1}, \tau_i]$. Because the integrals in the previous equation are now calculated on the intervals $[\tau_{i-1} + \epsilon\theta_{i-1}, \tau_i + \epsilon\theta_i]$, we have indexed λ with a function $i(t)$ equal to i on $[\tau_{i-1}, \tau_i]$. If we brake the integrals in the expression for $J_\epsilon^{(1)}$, we get:

$$\begin{aligned} J_\epsilon^{(1)} = &\sum_{i=1}^{N+1} \int_{\tau_{i-1}+\epsilon\theta_{i-1}}^{\tau_i} [L_i(x, \tilde{x}_\epsilon) + \lambda_i'(f_i(\tilde{x}_\epsilon, \tau_{i-1} + \epsilon\theta_{i-1}) - \dot{\tilde{x}}_\epsilon)] dt \\ &+ \sum_{i=1}^{N+1} \int_{\tau_i}^{\tau_i+\epsilon\theta_i} [L_i(x, \tilde{x}_\epsilon) + \lambda_{i+1}'(f_i(\tilde{x}_\epsilon, \tau_{i-1} + \epsilon\theta_{i-1}) - \dot{\tilde{x}}_\epsilon)] dt \end{aligned} \tag{11}$$

Also, given that :

$$\begin{aligned} \tilde{x}_\epsilon(t) &= \tilde{x}(t) + \epsilon\eta(t) + o(\epsilon) && \text{on } [\tau_{i-1} + \epsilon\theta_{i-1}, \tau_i] \\ \tilde{x}_\epsilon(t) &= \tilde{x}(\tau_i^-) + O(\epsilon) && \text{on } [\tau_i, \tau_i + \epsilon\theta_i], \end{aligned} \tag{12}$$

Equation (11) becomes:

$$\begin{aligned} J_\epsilon^{(1)} = &\sum_{i=1}^{N+1} \int_{\tau_{i-1}+\epsilon\theta_{i-1}}^{\tau_i} [L_i(x, \tilde{x} + \epsilon\eta) + \lambda_i'(f_i(\tilde{x} + \epsilon\eta, \tau_{i-1} + \epsilon\theta_{i-1}) - \dot{\tilde{x}} - \epsilon\dot{\eta})] dt \\ &+ \epsilon \sum_{i=1}^{N+1} [L_i(x, \tilde{x}(\tau_i^-)) + \lambda_{i+1}'(f_i(\tilde{x}(\tau_i^-), \tau_{i-1}) - \dot{\tilde{x}}_\epsilon)]\Big|_{\tau_i^+} \theta_i + o(\epsilon). \end{aligned} \tag{13}$$

Apply a Taylor expansion to the terms in the integral, and compute an integration by part of the terms $\lambda_i' \dot{\eta}$:

$$J_\epsilon^{(1)} = \sum_{i=1}^{N+1} \int_{\tau_{i-1}+\epsilon\theta_{i-1}}^{\tau_i} [L_i(x, \tilde{x}) + \lambda_i'(f_i(\tilde{x}, \tau_{i-1}) - \dot{\tilde{x}})] dt$$

$$+\epsilon\sum_{i=1}^{N+1}\int_{\tau_{i-1}+\epsilon\theta_{i-1}}^{\tau_i}[\frac{\partial L_i}{\partial\tilde{x}}\eta+\lambda_i'(\frac{\partial f_i}{\partial\tilde{x}}\eta+\frac{\partial f_i}{\partial\tau_{i-1}}\theta_{i-1})+\dot{\lambda}_i'\eta]dt$$
$$+\epsilon\sum_{i=1}^{N+1}[L_i(x,\tilde{x}(\tau_i^-))+\lambda_{i+1}'(f_i(\tilde{x}(\tau_i^-),\tau_{i-1})-\dot{\tilde{x}}_\epsilon)]\Big|_{\tau_i^+}\theta_i$$
$$-\epsilon\sum_{i=1}^{N+1}[\lambda_i'\eta]_{\tau_{i-1}+\epsilon\theta_{i-1}}^{\tau_i}\quad+o(\epsilon). \tag{14}$$

We almost recognize $J_0^{(1)}$ in the first term of Equation(14). So we propose to break the integrals of $J_0^{(1)}$ as we did for $J_\epsilon^{(1)}$:

$$J_0^{(1)}=\sum_{i=1}^{N+1}\int_{\tau_{i-1}}^{\tau_i}[L_i(x,\tilde{x})+\lambda_i'(f_i(\tilde{x},\tau_{i-1})-\dot{\tilde{x}})]dt$$
$$=\sum_{i=1}^{N+1}\int_{\tau_{i-1}+\epsilon\theta_{i-1}}^{\tau_i}[L_i(x,\tilde{x})+\lambda_i'(f_i(\tilde{x},\tau_{i-1})-\dot{\tilde{x}})]dt$$
$$+\sum_{i=1}^{N+1}\int_{\tau_{i-1}}^{\tau_{i-1}+\epsilon\theta_{i-1}}[L_i(x,\tilde{x})+\lambda_i'(f_i(\tilde{x},\tau_{i-1})-\dot{\tilde{x}})]dt$$
$$=\sum_{i=1}^{N+1}\int_{\tau_{i-1}+\epsilon\theta_{i-1}}^{\tau_i}[L_i(x,\tilde{x})+\lambda_i'(f_i(\tilde{x},\tau_{i-1})-\dot{\tilde{x}})]dt$$
$$+\epsilon\sum_{i=1}^{N+1}[L_i(x,\tilde{x})+\lambda_i'(f_i(\tilde{x},\tau_{i-1})-\dot{\tilde{x}})]\Big|_{\tau_{i-1}^+}\theta_{i-1}\;+o(\epsilon) \tag{15}$$

After a simple change of variable in the previous line, and regrouping with Equation (14), we get the first part of the directional derivative of J:

$$\nabla J^{(1)}=\lim_{\epsilon\to 0}\frac{J_\epsilon^{(1)}-J_0^{(1)}}{\epsilon}$$
$$=\sum_{i=1}^{N+1}\int_{\tau_{i-1}}[\frac{\partial L_i}{\partial\tilde{x}}\eta+\lambda_i'(\frac{\partial f_i}{\partial\tilde{x}}\eta+\frac{\partial f_i}{\partial\tau_{i-1}}\theta_{i-1})+\dot{\lambda}_i'\eta]dt$$
$$+\sum_{i=1}^{N}\Big[L_i(x(\tau_i),\tilde{x}(\tau_i^-))-L_{i+1}(x(\tau_i),\tilde{x}(\tau_i^+))$$
$$+\lambda_{i+1}(\tau_i^+)'(f_i(\tilde{x}(\tau_i^-),\tau_{i-1})-f_{i+1}(\tilde{x}(\tau_i^+),\tau_i))\Big]\theta_i$$
$$-\sum_{i=1}^{N+1}[\lambda_i'\eta]_{\tau_{i-1}^+}^{\tau_i^-} \tag{16}$$

Note that we used $\theta_0=0$ and $\theta_{N+1}=0$ as $\tau_0=t_0$ and $\tau_{N+1}=t_f$ are fixed.

Now we can focus our attention on the second part of the variation of J.
From Figure 1, we get the following approximations:

$$\begin{aligned} x(\tau_i+\epsilon\theta_i) &= x(\tau_i)+\epsilon\theta_i\dot{x}(\tau_i)+o(\epsilon)\\ \tilde{x}_\epsilon((\tau_i+\epsilon\theta_i)^-) &= x(\tau_i^-)+\epsilon\eta(\tau_i^-)+\epsilon\theta_i f_i(x(\tau_i^-))+o(\epsilon)\\ \Delta\tilde{x}_\epsilon|_{\tau_i+\epsilon\theta_i} &= \Delta\tilde{x}|_{\tau_i}+\epsilon\Delta\eta|_{\tau_i+\epsilon\theta_i}+o(\epsilon)\end{aligned} \tag{17}$$

We can then apply a Taylor expansion to $J_\epsilon^{(2)}$:

$$J_\epsilon^{(2)}=\sum_{i=1}^{N}[M_i(x(\tau_i+\epsilon\theta_i),\tilde{x}_\epsilon(\tau_i+\epsilon\theta_i^-),u_i+\epsilon\nu_i,\tau_i+\epsilon\theta_i)-\mu_i'\Delta\tilde{x}_\epsilon|_{\tau_i+\epsilon\theta_i})]$$
$$=\sum_{i=1}^{N}[M_i(x(\tau_i),\tilde{x}(\tau_i^-),u_i,\tau_i)-\mu_i'\Delta\tilde{x}|_{\tau_i})]$$
$$+\epsilon\sum_{i=1}^{N}[\frac{\partial M_i}{\partial x}\dot{x}(\tau_i)\theta_i+\frac{\partial M_i}{\partial\tilde{x}}(\eta(\tau_i^-)+f_i(x(\tau_i^-))\theta_i)$$
$$+\frac{\partial M_i}{\partial u_i}\nu_i+\frac{\partial M_i}{\partial\tau_i}\theta_i-\mu_i'\Delta\eta|_{\tau_i+\epsilon\theta_i}]\quad+o(\epsilon) \tag{18}$$

where the first term is nothing less than $J_0^{(2)}$. We thus get the second part of the directional derivative of J:

$$\nabla J^{(2)}=\lim_{\epsilon\to 0}\frac{J_\epsilon^{(2)}-J_0^{(2)}}{\epsilon}$$
$$=\sum_{i=1}^{N}[\frac{\partial M_i}{\partial x}\dot{x}(\tau_i)\theta_i+\frac{\partial M_i}{\partial\tilde{x}}(\eta(\tau_i^-)+f_i(x(\tau_i^-))\theta_i)$$
$$+\frac{\partial M_i}{\partial u_i}\nu_i+\frac{\partial M_i}{\partial\tau_i}\theta_i-\mu_i'\Delta\eta|_{\tau_i}] \tag{19}$$

Now, regrouping Equations (16) and (19) and rearranging terms, we get the total cost variation:

$$\nabla J=\sum_{i=1}^{N+1}\int_{\tau_{i-1}}^{\tau_i}[(\frac{\partial L_i}{\partial\tilde{x}}+\lambda_i'\frac{\partial f_i}{\partial\tilde{x}}+\dot{\lambda}_i')\eta+\lambda_i'\frac{\partial f_i}{\partial\tau_{i-1}}\theta_{i-1})]dt$$
$$+\sum_{i=1}^{N}\Big[L_i(x(\tau_i),\tilde{x}(\tau_i^-))-L_{i+1}(x(\tau_i),\tilde{x}(\tau_i^+))$$
$$+\lambda_{i+1}(\tau_i^+)'(f_i(\tilde{x}(\tau_i^-),\tau_{i-1})-f_{i+1}(\tilde{x}(\tau_i^+),\tau_i))\Big]\theta_i$$
$$+\lambda_{N+1}(t_f)'\eta(t_f)+\sum_{i=1}^{N}(\lambda_{i+1}(\tau_i^+)'-\mu_i)\eta(\tau_i^+)$$
$$+\sum_{i=1}^{N}(\frac{\partial M_i}{\partial\tilde{x}}+\mu_i-\lambda_i(\tau_i^-)')\eta(\tau_i^-)$$
$$+\sum_{i=1}^{N}[\frac{\partial M_i}{\partial x}\dot{x}(\tau_i)\theta_i+\frac{\partial M_i}{\partial\tilde{x}}f_i(x(\tau_i^-),\tau_{i-1})\theta_i$$
$$+\frac{\partial M_i}{\partial u_i}\nu_i+\frac{\partial M_i}{\partial\tau_i}\theta_i] \tag{20}$$

Now, in order to cancel all the terms containing η, we choose the Lagrange multipliers or "costates" λ_i and μ_i such that:

$$\lambda_{N+1}(t_f)'=0 \tag{21}$$
$$\dot{\lambda}_i'(t)=-\frac{\partial L_i}{\partial\tilde{x}}(x(t),\tilde{x}(t))-\lambda_i'(t)\frac{\partial f_i}{\partial\tilde{x}}(\tilde{x}(t),\tau_{i-1}) \tag{22}$$
$$\lambda_i(\tau_i^-)'=\lambda_{i+1}(\tau_i^+)'+\frac{\partial M_i}{\partial\tilde{x}}(x(\tau_i),\tilde{x}(\tau_i^-),u_i,\tau_i) \tag{23}$$
$$\mu_i=\lambda_{i+1}(\tau_i^+)' \tag{24}$$

The final expression for ∇J is:

$$\nabla J=\sum_{i=1}^{N}\Big\{\int_{\tau_i}^{\tau_{i+1}}[\lambda_{i+1}'\frac{\partial f_{i+1}}{\partial\tau_i}]dt+\frac{\partial M_i}{\partial x}\dot{x}(\tau_i)+\frac{\partial M_i}{\partial\tau_i}$$

$$+H_i(\tau_i^-) - H_{i+1}(\tau_i^+)\Big\}\theta_i + \sum_{i=1}^{N} \frac{\partial M_i}{\partial u_i}\nu_i \quad (25)$$

where $H_i = L_i + \lambda_i' f_i$.
Finally, note that this last expression for the directional derivative gives access to the partial derivatives $\frac{\partial J}{\partial \tau_i}$ and $\frac{\partial J}{\partial u_i}$ as we know that:

$$\nabla_{\boldsymbol{\theta},\boldsymbol{\nu}} J(\boldsymbol{\tau}, \mathbf{u}) = \sum_{i=1}^{N} \frac{\partial J}{\partial \tau_i}\theta_i + \sum_{i=1}^{N} \frac{\partial J}{\partial u_i}\nu_i. \quad (26)$$

We summarize our results in the following theorem:

Theorem:
Given a smooth function $x(t) \in \mathbb{R}^n$, and its approximation function $\tilde{x}(t) \in \mathbb{R}^n$ given by:

$$\begin{cases} \tilde{x}(t_0) &= \tilde{x}_0 \\ \dot{\tilde{x}}(t) &= f_i(\tilde{x}(t), \tau_{i-1}) \quad \text{for } t \in [\tau_{i-1}, \tau_i) \\ \tilde{x}(\tau_i^+) &= \tilde{x}(\tau_i^-) + G_i(x(\tau_i), \tilde{x}(\tau_i^-), u_i, \tau_i) \end{cases} \quad (27)$$

an extremum to the cost function:

$$J(\boldsymbol{\tau}, \mathbf{u}) = \sum_{i=1}^{N+1} \int_{\tau_{i-1}}^{\tau_i} L_i(x(t), \tilde{x}(t))dt + \sum_{i=1}^{N} K_i(x(\tau_i), \tilde{x}(\tau_i^-), u_i, \tau_i) \quad (28)$$

is reached when the control variables $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_N) \in [t_0, t_f]^N$ and $\mathbf{u} = (u_1, \ldots, u_N) \in \mathbb{R}^N$ are chosen such that: Define:

$$M_i = K_i + \mu_i G_i \quad (29)$$
$$H_i = L_i + \lambda_i f_i \quad (30)$$

Euler Lagrange-Equations:

$$\dot{\lambda}_i'(t) = -\frac{\partial L_i}{\partial \tilde{x}}(x(t), \tilde{x}(t)) - \lambda_i'(t)\frac{\partial f_i}{\partial \tilde{x}}(\tilde{x}(t), \tau_{i-1}) \quad (31)$$

Boundary Conditions:

$$\lambda_{N+1}(t_f)' = 0 \quad (32)$$

$$\lambda_i(\tau_i^-)' = \lambda_{i+1}(\tau_i^+)' + \frac{\partial M_i}{\partial \tilde{x}}(x(\tau_i), \tilde{x}(\tau_i^-), u_i, \tau_i) \quad (33)$$

Multipliers

$$\mu_i = \lambda_{i+1}(\tau_i^+)' \quad (34)$$

Optimality Conditions

$$\frac{\partial J}{\partial \tau_i} = \int_{\tau_i}^{\tau_{i+1}} [\lambda_{i+1}' \frac{\partial f_{i+1}}{\partial \tau_i}]dt + \frac{\partial M_i}{\partial x}\dot{x}(\tau_i) + \frac{\partial M_i}{\partial \tau_i} + H_i(\tau_i^-) - H_{i+1}(\tau_i^+) = 0 \quad (35)$$

$$\frac{\partial J}{\partial u_i} = \frac{\partial M_i}{\partial u_i} = 0. \quad (36)$$

C. Search for a Local Minimum

The motivation behind getting an expression for the gradient is that we can now perform a gradient descent method. The following numerical algorithm is proposed :

At each iteration k, where $(\boldsymbol{\tau}^{(k)}, \boldsymbol{u}^{(k)}) = (\tau_1, \ldots, \tau_N, u_1, \ldots, u_N)$ is the current vector of control variables, we follow these steps :

1) Compute the approximation function $\tilde{x}(t)$ forward in time from t_0 to t_f using Equation (27).

2) Compute the co-state variables λ_i backward in time from t_f to t_0 using Equations (31)-(33).

3) Compute the gradient $\nabla J\big((\boldsymbol{\tau}^{(k)}, \boldsymbol{u}^{(k)}\big)\big)$ using Equations (35) and (36).

4) Update the control variables as follow :

$$(\boldsymbol{\tau}^{(k+1)}, \boldsymbol{u}^{(k+1)}) = (\boldsymbol{\tau}^{(k)}, \boldsymbol{u}^{(k)}) - \gamma^{(k)}\nabla J\big((\boldsymbol{\tau}^{(k)}, \boldsymbol{u}^{(k)}\big)\big) \quad (37)$$

Note that the choice of the stepsize $\gamma^{(k)}$ can be critical for the method to converge. An efficient method among others is the use of Armijo's algorithm presented in [3].
Because of the non-convex nature of the cost function J, this gradient descent algorithm can only converge to a local minimum. The obtention of a "good" local minimum can thus be quite dependent on the choice of a "good" initial guess for the control variables. However, the method presented here still offers significant reductions in the cost function.

III. EXAMPLES

A. Constant approximations with resets and jumps

We first try our algorithm on a simple example where the function $x(t)$ is approximated by piecewise constant functions i.e. by functions $\tilde{x}(t)$ such that $\dot{\tilde{x}}(t) = f_i(\tilde{x}(t), \tau_{i-1}) = 0$ on $[\tau_{i-1}, \tau_i)$, $i = 1, \ldots, N+1$. We find a solution and compare the results for two kinds of approximating functions:

1) functions where the error between x and $\tilde{x}$ is reset to zero at the instants τ_i, i.e. $G_i(x(\tau_i), \tilde{x}(\tau_i^-), u_i, \tau_i) = x(\tau_i) - \tilde{x}(\tau_i^-)$. The only control variables in this case are the instants $\{\tau_i\}_{i=1}^N$.
2) functions where the jumps at the instants τ_i are free variables, i.e. $G_i(x(\tau_i), \tilde{x}(\tau_i^-), u_i, \tau_i) = u_i$. The control variables in this case are both $\{\tau_i\}_{i=1}^N$ and $\{u_i\}_{i=1}^N$.

We choose the same cost function for both problems: $J = \int_{t_0}^{t_f}(x(t) - \tilde{x}(t))^2 dt$. Such a cost function only penalizes the error between $x(t)$ and $\tilde{x}(t)$. No costs associated with the jumps at τ_i are added.
Figure 2 shows the optimal approximating functions of the same function $x(t) = 0.5(0.5 - t)\sin^2(3\pi t) + 0.3t$ defined on $t \in [0, 1]$ and for $N = 8$. Our initial guess for the approximation using resets was such that the number of resets was more important in regions where the curve's

variation is important, i.e. where the error is likely to increase fast. For the approximation using free jumps, we initiated the algorithm with the values τ_i and $u_i = \Delta\tilde{x}|_{\tau_i}$ of the optimal solution of the first method. Figure 3 shows how fast the two algorithms converge to a solution.

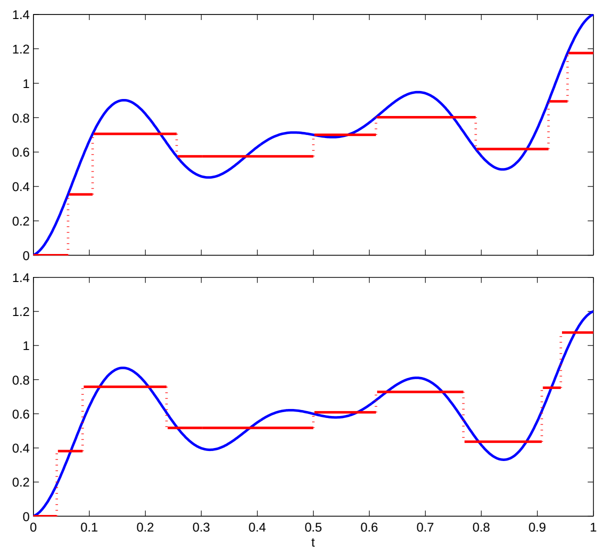

Fig. 2. Optimal Approximation using resets, jumps

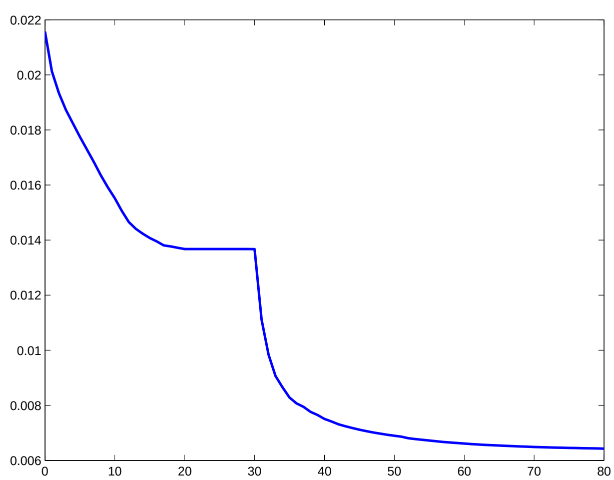

Fig. 3. Cost function decrease (for a constant stepsize γ)

B. Piecewise linear systems with resets

In this more elaborate example, we approximate the original function $x(t) = 3t^2\sin(3\pi t) - t$ on $t \in [0,1]$ by piecewise linear systems.
Assuming $\dot{x}(t) = g(x(t))$ and writing $x(t) = x(\tau_i) + e(t)$ on $t \in [\tau_i, \tau_{1+1})$ where $e(t)$ is the error function such that $e(\tau_i) = 0$, a first order Taylor expansion gives:

$$\begin{aligned}\dot{x}(t) &= g(x(\tau_i) + e(t)) \\ &\approx g(x(\tau_i)) + \frac{\partial g}{\partial x}(x_{\tau_i})e(t) \\ &\approx \dot{x}(\tau_i) + \frac{\ddot{x}(\tau_i)}{\dot{x}(\tau_i)}(x(t) - x(\tau_i))\end{aligned}$$

We can then approximate $x(t)$ by the piecewise linear systems such that:

$$\left\{\begin{aligned}\tilde{x}(\tau_i) &= x(\tau_i), \qquad i = 1, \ldots, N \\ \dot{\tilde{x}}(t) &= \dot{x}(\tau_i) + \tfrac{\ddot{x}(\tau_i)}{\dot{x}(\tau_i)}(\tilde{x}(t) - x(\tau_i)) \\ &= A(\tau_i)\tilde{x}(t) + B(\tau_i) \qquad t \in [\tau_i, \tau_{i+1})\end{aligned}\right.$$

Hence we will use our algorithm with

$$\left\{\begin{array}{l} f_i(\tilde{x}, \tau_{i-1}) = \dot{x}(\tau_{i-1}) + \frac{\ddot{x}(\tau_{i-1})}{\dot{x}(\tau_{i-1})}(\tilde{x}(t) - x(\tau_{i-1})) \\ G_i(x(\tau_i), \tilde{x}(\tau_i^-), u_i, \tau_i) = x(\tau_i) - \tilde{x}(\tau_i^-) \end{array}\right.$$

Figure 4 shows the optimal solution where the cost function to minimize was $\int_0^1 (x(t) - \tilde{x}(t))^2 dt$.

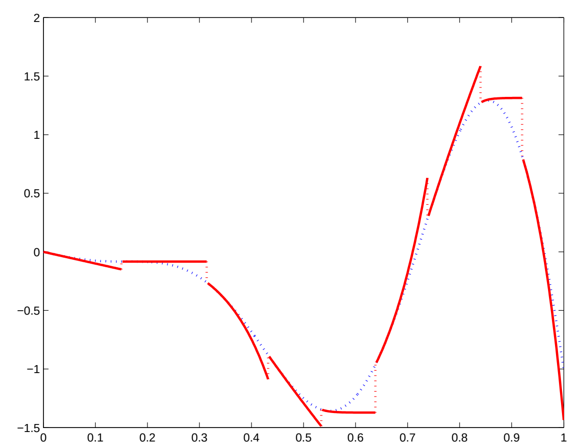

Fig. 4. Optimal approximation using Linear Systems

IV. CONCLUSIONS

In this paper we used a variational approach to derive necessary conditions for the stationarity of the cost function of a hybrid function approximation system. Such equations have been successfully integrated in a numerical gradient descent algorithm. The association of such a local method with heuristic strategies in order to find a global minimum has not been investigated here but could be considered.
The main application envisioned for the proposed method involves data compression, where continuous trajectories are stored as mode sequences, switch times, and jump heights. This is an endeavor that will be pursued in the future.

REFERENCES

[1] X. Xu, P.J. Antsaklis. "Optimal Control of Hybrid Autonomous Systems with State Jumps". Proceedings of the American Control Conference 2003, Volume: 6, Pages: 5191 - 5196, 2003.

[2] X. Xu, P.J. Antsaklis. "Quadratic Optimal Control Problems for Hybrid Linear Autonomous Systems with State Jumps". Proceedings of the American Control Conference, 2003.

[3] L. Armijo. ”Minimization of Functions Having Lipschitz Continuous First-Partial Derivatives”. Pacific Journal of Mathematics, Vol. 16, ppm. 1-3, 1966.
[4] E. W. Cheney. ”Introduction to approximation theory”. Publisher: New York, Mcgraw-hill Book Co. 1966.
[5] O. Christensen and K.L. Christensen. ”Approximation theory : from Taylor polyonomials to wavelets”. Publisher: Boston : Birkhuser, 2005.
[6] A. Giua, C. Seatzu and C. Van Der Mee. ”Optimal Control of Switched Autonomous Linear Systems”. Proceedings of the 40th IEEE CDC, Orlando, FL, Dec. 2001.
[7] E.D. Sontag. ”Nonlinear Regulation: the Piecewise Linear Approach”. IEEE Trans. Autom. Control, Volume: 26 , Issue: 2 , Pages: 346 - 358, 1981.
[8] E. Verriest, F. Delmotte and M. Egerstedt. ”Optimal Impulsive Control for Point Delay Systems with Refractory Period”. IFAC Workshop on Time-Delay Systems, Leuven, Belgium, Sept. 2004.