OPTIMAL CONTROL OF SWITCHING SURFACES IN HYBRID DYNAMIC SYSTEMS

Mauro Boccadoro [*] **Magnus Egerstedt** [**,1]
Yorai Wardi [**,1]

[*] *boccadoro@diei.unipg.it*
Dipartimento di Ingegneria Elettronica e dell'Informazione
Università di Perugia
06125, Perugia – Italy
[**] {*magnus,ywardi*}*@ece.gatech.edu*
Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332

Abstract: This paper studies the problem of optimal switching surface design for hybrid systems. In particular, a formula is derived for computing the gradient of a given integral performance cost with respect to the switching surface parameters. The novelty of the work lies both in the problem formulation itself, and in the introduction of a discrete event variational system for propagating the perturbations across switching surfaces.

Keywords: Optimal Control, Hybrid Systems, Switching Surfaces, Gradient Descent, Numerical Algorithms

1. INTRODUCTION

To view the switching surface parameters as free variables in an optimal control formulation for hybrid systems is an approach that previously has not been pursued. In this paper, we address this problem by computing the gradient of a given cost functional with respect to the switching surface parameters, which enables us to employ gradient-based numerical techniques for obtaining locally optimal solutions. The problem of optimal switching surface design draws its motivation from the many hybrid models where transitions between the continuous modes occur based on internal (state-based) events, as illustrated in Figure 1. Examples of such systems include autonomous navigation and obstacle avoidance for mobile robots (Arkin, 1998; Egerstedt, 2000), and production planning for manufacturing systems (Cassandras *et al.*, 2001; Boccadoro and Valigi, 2003). In this context, identifying optimal switching surfaces in the state space amounts to solving an optimal control problem in a closed loop fashion.

To control such hybrid systems is a problem that has received a significant amount of attention. In particular, the problem of finding optimal, continuous control laws, given a partition of the state space (Johansson and Rantzer, 1998) or a fixed set of switching surfaces (Shaikh and Caines, 2002; Sussmann, 1999; Xu and Antsaklis, 2002; Riedinger *et al.*, 1999) has been investigated. In (Giua *et al.*, 2001; Bemporad *et al.*, 2002) the authors addressed a timing optimization problem and derived homogeneous regions in the state space that determine the switchings. However, the

[1] The support from the Manufacturing Research Center at the Georgia Institute of Technology is gratefully acknowledged.

problem of optimal design of switching surfaces has not yet been addressed. In this paper, we assume that the switching surfaces are of the form $g : \Re^n \times \Re^m \to \Re$, where a transition occurs if $g(x, a) = 0$. Here, g is assumed to be continuously differentiable in both arguments, $x \in \Re^n$ is the state of the system, and $a \in \Re^m$ is a controllable parameter. *The main problem that will be considered in this paper is how to choose such parameters in order to minimize a suitable cost functional.*

This problem of optimal control of the switching surfaces will be approached from a numerical optimal control vantage point in the sense that the derivative of the cost with respect to the switching parameters will be computed. This derivative allows for gradient descent methods to be applied, and as a consequence, local optima can be obtained. The enabling result for this program to be successful is the computation of the derivative of the cost with respect to the switching parameters, and the program presents an extension to the work on optimal timing control (Egerstedt *et al.*, n.d.). The gradient will be shown to have a special structure, making it suitable as a basis for a numerical optimization algorithm. In fact, it will be based on the combination of a single, backwards costate computation (as established in (Egerstedt *et al.*, n.d.)) and a collection of forward integrations of the continuous dynamics as well as a discrete event system that describes, in an iterative fashion, how the perturbations propagate across the different switching surfaces.

This paper is organized as follows: In Section 2 we introduce the problem at hand, followed by a study of the single-switch case, in Section 3. In Section 4 a discrete event dynamics is introduced for the characterization of multiple switches trajectories. Section 5 presents the gradient computation, discusses its complexity, and includes an illustrative example.

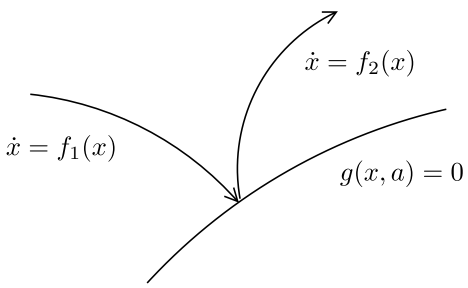

Fig. 1. The transition between modes occur as the state reaches a switching surface in the state space.

2. PROBLEM FORMULATION

In this paper we consider a switched system described by (see (Branicky *et al.*, 1998)):

$$\dot{x}(t) = f_i(x(t)) \quad (1)$$

$$i^+(t) = s(x(t), i(t)) \quad (2)$$

In particular, Eq. (1) describes the continuous dynamics, Eq. (2) describes the discrete event dynamics of the system: the switch from the *mode* indexed by i to mode i^+ occurs as the trajectory of the system, in the continuous state space, intersects a guard/threshold set described by a function of the kind $g_i(x, a_i) = 0$, i.e.,

$$s(x, i) = \begin{cases} i & g_i(x, a_i) \neq 0 \\ i^+ & g_i(x, a_i) = 0 \end{cases}$$

Note that the threshold functions g_i do not partition the state space in different regions each characterized by a certain mode, rather they dictate the occurrence of the i^{th} switch and are "deactivated" thereafter. The functions g_i are assumed to be continuously differentiable in the x variable and the gradient w.r.t. the state variable $\partial g/\partial x$ will be denoted ∇g, a row vector. Also, the switching surfaces will be assumed to be always reachable. In formal terms:

Assumption 1.

$$\forall i \text{ and } \forall x \,|\, g_i(x, a_i) = 0, \; L_{f_i} g_i(x) \neq 0$$

where $L_{f_i} g_i(x) := \nabla g_i(x) f_i(x)$. The shape of the switching curves are determined by the form of the function g_i, given a priori, and by the value of the parameters a_i. These values constitute the control variables for the switched system. Consider a system evolution starting at $x(t_0) = x_0$, ruled by mode 1, which switches to mode 2 as the condition $g_1(x(t), a_1) = 0$ is satisfied, for some t. Similarly all subsequent switches from mode i to mode $i+1$ occur as the i^{th} guard condition is met, i.e. when $g_i(x(t), a_i) = 0$. In the following, the i^{th} switching instant will be denoted by t_i and the associated switching state as $x_i := x(t_i)$.

Here, the problem of tuning the values of parameters a_i, $i = 1, \ldots, N$ in order to minimize an integral cost function $J = \int_{t_0}^{T} L(x)$ is addressed. The final time T can be either a fixed final time or the time instant corresponding to the intersection with a terminal manifold $g_N(x, a_N) = 0$, i.e., in this case, system evolution terminates after $N-1$ switches at time $T = t_N$. The main contribution of this paper is to provide an expression for the derivatives dJ/da_i, which can be employed by a gradient descent algorithm for the (local) optimization of the switching surfaces. For simplicity, the case $m = 1$, will be considered, i.e., each switching surface is characterized by a single parameter, but it should be noted that similar arguments apply also in the general case $m > 1$.

3. PRELIMINARY RESULTS

Consider the trajectory of an autonomous system $\dot{x} = f(x)$, starting at time t_0 in x_0, hitting a surface $g(x, a) = 0$, $a \in \Re$, at time t_h, and incurring in a cost $J = \int_{t_0}^{t_h} L(x)$. Let $x_h = x(t_h)$, and let Φ denote the state transition matrix of the linearized system $\dot{z} = \frac{\partial f}{\partial x} x(t) z(t)$.

Lemma 1.

$$\frac{dt_h}{da} = -\frac{1}{L_f g(x_h)} \frac{\partial g}{\partial a}(x_h) \tag{3}$$

Proof. Let a and $\tilde{a} = a + \Delta a$ denote the nominal and perturbed value for the switching surface parameter, which correspond to switches at time t_h in x_h or at time $\tilde{t}_h$ in $\tilde{x}_h$, respectively. Under Assumption 1 the perturbations $\Delta t_h := \tilde{t}_h - t_h$ and $\Delta x_h := \tilde{x}_h - x_h$ can be made arbitrarily small by proper small values of Δa, thus the following first order approximations can be made:

$$g(\tilde{x}_h, \tilde{a}) = g(x, a) + \frac{\partial g}{\partial x}\Delta x_h + \frac{\partial g}{\partial a}\Delta a + o(\Delta x_h, \Delta a)$$
$$\Delta x_h = f(x_h)\Delta t_h + o(\Delta t_h).$$

Eq. (3) follows from the switching conditions $g(x, a) = 0$, $g(\tilde{x}_h, \tilde{a}) = 0$, and considering that by Assumption 1 $L_f g(x_h) \neq 0$. □

Lemma 2.

$$\frac{dt_h}{dx_0} = -\frac{\nabla g(x_h)\Phi(t_h, t_0)}{L_f g(x_h)}. \tag{4}$$

Proof. Consider a perturbed initial condition $\tilde{x}_0 = x_0 + \Delta x_0$, resulting in a perturbed trajectory $\tilde{x}(\cdot)$, and assume that the intersection with the surface $g(x, a) = 0$ takes place at time $\tilde{t}_h = t_h + \Delta t_h$. A first order approximation of the condition to be satisfied by the perturbed trajectory at the intersection with the surface $g(x, a) = 0$ is given by:

$$\begin{aligned} 0 = g(\tilde{x}(\tilde{t_h}), a) = g(x(t_h), a) + \\ \nabla g(x(t_h))\left(\Delta x(t_h) + f(x(t_h))\Delta t_h\right) + \\ o(\Delta x(t_h), \Delta t_h) \end{aligned} \tag{5}$$

where Δx denotes the variation in the nominal trajectory at a fixed time, hence:

$$\Delta x(t_h) = \Phi(t_h, t_0)\Delta x_0 + o(\Delta x_0) \tag{6}$$

Since $g(x(t_h), a) = 0$, substituting (6) in (5) yields the result. □

Lemma 3.

$$\frac{dJ}{dx_0} = p(t_0) \tag{7}$$

where the *costate* variable p satisfies the (backward) o.d.e.

$$\dot{p}^T = -p^T \frac{\partial f}{\partial x} - \frac{\partial L}{\partial x}, \tag{8}$$

$$p^T(t_h) = -\frac{L(x_h)\nabla g(x_h))}{L_f g(x_h)} \tag{9}$$

Also, if t_h is a *fixed final time* then $p^T(t_h) = 0$

Proof. This result is well known in optimal control theory (e.g. (Bryson and Ho, 1969)).

4. MULTIPLE SWITCHES

This section deals with the perturbations induced on the switching times and the corresponding states by small variations in the value of the threshold parameters. The manner in which these perturbations propagate across the switching surfaces will highlight the recursive nature of the solution, and how this is captured by introducing costate variables. The recursive propagation will take the form of a state-space-like discrete event system for the infinitesimal perturbations sampled in correspondence of the switches in the continuous dynamics. The variables $\tau_i(k)$ and $\chi_i(k)$ are introduced to represent the perturbations resulting at the k^{th} switching instant due to variations in the i^{th} threshold parameter a_i. In fact, by lifting the problem to a discrete event setting, we still capture all the information required to determine time and state perturbations at the next, $k + 1$ switching surface from these new variables. Hence they represent the *state* of the resulting discrete event system.

Since the switching surfaces are not time dependent, the hitting time at the k^{th} switching surface, t_k depends on the parameters a_k of such surface and on the hitting time and state at the previous switching surface t_{k-1} and x_{k-1}.

$$t_k = t_k(t_{k-1}, x_{k-1}, a_k) \tag{10}$$

The k^{th} switching state x_k, depends also on t_k, and is given by:

$$x_k = x_{k-1} + \int_{t_{k-1}}^{t_k} f_k(x(t))dt \tag{11}$$

Consider the infinitesimal approximation of the perturbations in the hitting times and states following a slight variation Δa_i on parameter a_i, assuming, for the moment, that this is independent by all the other parameters $a_j, j \neq i$. In this case, the variables $\tau_i \in \Re$ and $\chi_i \in \Re^n$ denote the following quantities:

$$\tau_i(k) = \frac{dt_k}{da_i}\Delta a_i + o(\Delta a_i) \tag{12}$$

$$\chi_i(k) = \frac{dx_k}{da_i}\Delta a_i + o(\Delta a_i); \tag{13}$$

According to this notation, differentiation of (10,11) give:

$$\tau_i(k) = \frac{\partial t_k}{\partial t_{k-1}}\tau_i(k-1) + \frac{\partial t_k}{\partial x_{k-1}}\chi_i(k-1) \quad (14)$$

$$\chi_i(k) = \frac{\partial x_k}{\partial t_k}\tau_i(k) + \frac{\partial x_k}{\partial t_{k-1}}\tau_i(k-1) + \frac{\partial x_k}{\partial x_{k-1}}\chi_i(k-1). \quad (15)$$

Since the equation $\dot{x} = f_k(x)$ is time-invariant in the interval $[t_{k-1}, t_k)$, it follows that $\partial t_k/\partial t_{k-1} = 1$. The term $\partial t_k/\partial x_{k-1}$ is the variation in the k^{th} hitting time w.r.t. variations on x_{k-1} considering a fixed threshold surface, i.e., by Proposition 2:

$$\frac{\partial t_k}{\partial x_{k-1}} = -\frac{\nabla g_k(x_k)\Phi_k(t_k, t_{k-1})}{L_{f_k} g_k(x_k)} \quad (16)$$

Moreover, $dx_k/dt_k = f_k(x_k)$, since it is the equation $\dot{x} = f_k(x)$ in the interval $[t_{k-1}, t_k)$ that determines x_k (see Eq. 1); and $\partial x_k/\partial x_{k-1} = \Phi_k(t_k, t_{k-1})$. The term $\partial x_k/\partial t_{k-1}$ refers to a variation considering a fixed initial state x_{k-1} and fixed final time t_k; being the system time invariant it results: $\partial x_k/\partial t_{k-1} = -f_k(x_k)$. Finally it results: $\partial x_k/\partial x_{k-1} = \Phi_k(t_k, t_{k-1})$. Having evaluated all the coefficients of Eqs. (14, 15), by substituting the former for the term $\tau_i(k)$ in the latter, it is possible to give a state space like expression for the iterative evolution of the variables τ_i and χ_i. In order to present this result with the least possible notational burden, the obvious arguments of functions f, g, and Φ and their (repeated) indexes will be dropped, adopting the following conventions: $f_k := f_k(x_k)$; $\nabla g_k := \nabla g_k(x_k)$; $L_k := L_{f_k} g_k(x_k)$ $\Phi_{k,k-1} = \Phi_k(t_k, t_{k-1})$.

$$\tau_i(k) = \tau_i(k-1) - \frac{\nabla g_k \Phi_{k,k-1}}{L_k}\chi_i(k-1) \quad (17)$$

$$\chi_i(k) = \left(I - \frac{f_k \nabla g_k}{L_k}\right)\Phi_{k,k-1}\chi_i(k-1) \quad (18)$$

4.1 Not independent threshold parameters

Assuming the a_i not independent (e.g. two distinct switches are determined by the same surface) adds forcing terms proportional to $\alpha_i(k) := \frac{da_k}{da_i}\Delta a_i + o(\Delta a_i)$ to the "autonomous" dynamics (17,18). In fact, (14) modifies as follows:

$$\tau_i(k) = \frac{\partial t_k}{\partial t_{k-1}}\tau_i(k-1) + \frac{\partial t_k}{\partial x_{k-1}}\chi_i(k-1) + \frac{\partial t_k}{\partial a_k}\alpha_i(k)$$

which yields:

$$\begin{pmatrix}\tau_i(k)\\ \chi_i(k)\end{pmatrix} = A(k)\begin{pmatrix}\tau_i(k-1)\\ \chi_i(k-1)\end{pmatrix} - \begin{bmatrix}\frac{\frac{\partial g_k}{\partial a}}{L_k}\\ \frac{\frac{\partial g_k}{\partial a}}{L_k} f_k\end{bmatrix}\alpha_i(k) \quad (19)$$

where $A(k)$ is the matrix whose members derive from Eqs. (17,18). The above dynamics describe how the perturbation in the hitting times due to a slight modification in the parameter a_i propagates from the i^{th} hit. As initial conditions set $\tau_i(0) = 0$ and $\chi_i(0) = 0$; observe that the system describing the propagation of the perturbations remains at such a "state" zero until the first non null forcing term, i.e. $\alpha_i(i) = 1$, causes it to drift ($\alpha_i(k) = 0$ for $k < i$).

5. COMPUTATION OF THE GRADIENT OF THE COST FUNCTION

Consider a nominal trajectory $x(\cdot)$ characterized by N switches, occurring at time instants t_k and states x_k, and a perturbed trajectory $\tilde{x}(\cdot)$, due to a small variation in parameter a_i, with switches at time instants $\tilde{t}_k$ at states $\tilde{x}_k$. Let $\Theta_k := \max\{t_k, \tilde{t}_k\}$, $\theta_k := \min\{t_k, \tilde{t}_k\}$, and consider the piece of nominal and perturbed trajectories delimited by the time instant Θ_k and by their respective intersections with the next, $(k+1)^{st}$ switching surface. Notice that both nominal and perturbed trajectories evolve according to the same mode $k+1$. The corresponding variation in the cost function J, i.e.,

$$\int_{\Theta_k}^{\tilde{t}_{k+1}} L(\tilde{x}(s))ds - \int_{\Theta_k}^{t_{k+1}} L(x(s))ds, \quad (20)$$

can be referred to Lemma 3, where the perturbation on the initial conditions, according to (12,13) are given by:

$$\Delta x(\Theta_k) = \tilde{x}(\Theta_k) - x(\Theta_k) = \chi_i(k) - f_{k+1}(x_k)\tau_i(k), \quad (21)$$

independently by the relative ordering of t_k and $\tilde{t}_k$ (see Figure 2). Therefore, by (7,21) a first order approximation of Eq. (20) is:

$$p_{k+1}^T(t_k)(\chi_i(k) - f_{k+1}(x_k)\tau_i(k)) \quad (22)$$

where the costate variable evolves backwards according to:

$$\dot{p}_k^T = -p_k^T \frac{\partial f_k}{\partial x} - \frac{\partial L}{\partial x}, \quad (23)$$

with boundary conditions (see Eq. 9) :

$$p_k^T(t_k) = -\frac{L(x_k)\nabla g(x_k)}{L_f g(x_k)}, \quad (24)$$

A unique costate variable p can be defined by means of the switched dynamics (23), with reset conditions (24), for $k = 1, \ldots N$, so that: $p^T(t_k^+) = p_{k+1}^T(t_k)$. Notice that the choice of time t_k instead of time Θ_k in (22) is justified since

$$p(\Theta_k) = p(t_k) + \dot{p}(t_k)\tau_i(k) + o(\Delta a_i),$$

so that Eq. (22) still constitutes a proper first order approximation to (20).

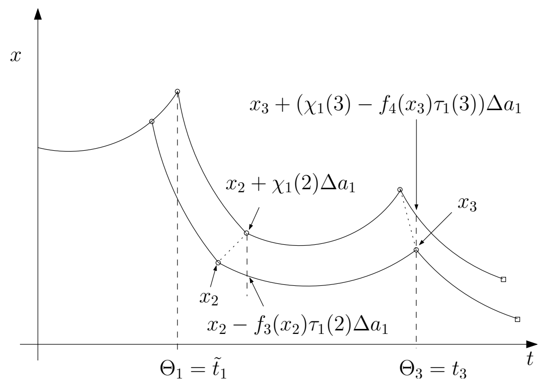

Fig. 2. An example of nominal (below) vs. perturbed (above) trajectories for one component of the state variables, denoted by x. A variation in parameter a_1 generates perturbations in the hitting times and states for all switches

For completing the derivation of dJ/da_i it is convenient to introduce the following function:

$$I_i(k) = \begin{cases} -\int_{t_k}^{\tilde{t}_k} L(x(s))ds, \ t_k \leq \tilde{t}_k \\ \int_{\tilde{t}_k}^{t_k} L(\tilde{x}(s))ds, \quad \tilde{t}_k < t_k \end{cases} \tag{25}$$

It is easy to see that independently by the relative ordering of t_k and $\tilde{t}_k$, it results:

$$I_i(k) = -L(x_k)\tau_i(k) \tag{26}$$

The variation in the cost function J due to a variation in a_i, is (see Figure 2):

$$\Delta J_i = \int_{\theta_i}^{\tilde{t}_N} L(\tilde{x}) - \int_{\theta_i}^{t_N} L(x) =$$
$$\int_{\theta_i}^{\Theta_i} L(\tilde{x}(s)) - L(x(s))ds + \sum_{k=i+1}^{N-1} I_i(k) +$$
$$\sum_{k=i}^{N-1} \left(\int_{\Theta_k}^{\tilde{t}_{k+1}} L(\tilde{x}(s))ds - \int_{\Theta_k}^{t_{k+1}} L(x(s))ds \right). \tag{27}$$

The only term of (27) which has not been characterized yet is: $\int_{\theta_i}^{\Theta_i}(L(\tilde{x}(s)) - L(x(s))ds$ which results to be a $o(\Delta a_i)$. According to this consideration, evaluating (27) at the light of (22,26), letting $\Delta a_i \to 0$, and recalling the definitions (12,13) yields:

$$\frac{dJ}{da_i} = \sum_{k=i}^{N-1} p^T(t_k^+)(\frac{dx_k}{da_i} - f_{k+1}(x_k)\frac{dt_k}{da_i})$$
$$- \sum_{k=i+1}^{N-1} L(x_k)\frac{dt_k}{da_i}. \tag{28}$$

5.1 Computational Issues

Eq. (28) describes the generic term of the gradient dJ/da, $a := (a_1, \ldots, a_N)^T$. The computation of each term dJ/da_i needs: i) the set of switching time instants t_k and switching states x_k, for $k = i, \ldots, N$; ii) the perturbation terms $\tau_i(k)$ $\chi_i(k)$, approximating $\frac{dt_k}{da_i}$ and $\frac{dx_k}{da_i}$ in (28), for $k = i, \ldots, N-1$; iii) the value of the costate variable p at times t_k^+, for $k = i, \ldots, N-1$. The computational cost for each of the above can be evaluated as follows:

I) A *single* forward integration of the state equations provides t_k, x_k, for $k = 1, \ldots, N$. Since those are nominal values they are independent by which parameter a_i is considered varying.

II) The characterization of the perturbation terms is provided via (19). Associated to the variations on each parameter a_i are the relative set of time and state perturbations $\tau_i(k)$ and $\chi_i(k)$, $k = i, \ldots, N$, whose computation needs *independent* forward integrations, one for each varying threshold parameter. Although the computational burden may seem large, it is worth noting that in some cases the number of switching surfaces is far less than the associated switchings (e.g. stability).

III) A *single* backwards integration for the costate dynamics (23) provides the needed values for the costate: in fact, such dynamics as well as the corner conditions (24) are independent of the varying parameter a_i.

5.2 Example

We present an example of switching surfaces optimization, using the gradient formulae derived (28) in a gradient descent algorithm. We consider the system, (after (Branicky, 1998)), $\dot{x} = A_i x$, characterized by two modes:

$$A_1 = \begin{pmatrix} -1 & -100 \\ 10 & -1 \end{pmatrix}; A_2 = \begin{pmatrix} 1 & 10 \\ -100 & 1 \end{pmatrix};$$

System switches from mode 1 to mode 2 upon hitting the line $x_1 + a_1^2 x_2 = 0$ in the fourth quadrant; and from mode 2 to mode 1 upon hitting $a_2^2 x_1 + x_2 = 0$ in the second quadrant. The switching surfaces optimization is carried out by a gradient descent algorithm for the switching parameters. Following the procedure described in the preceding Section, the gradient dJ/da, $a = \{a_1, a_2\}$ is computed for the current, nominal instance of the switching parameters, and their value is iteratively optimized, by gradient descent, until some terminating condition for the algorithm is satisfied. As a cost function let:

$$J = \frac{1}{2} \int_0^T ||x^2(t)||dt.$$

System evolution starts at $x_0 = (-1, 0)^T$ and terminates at a fixed final time $T = 0.3$. The optimization is achieved by a gradient descent algorithm with Armijo stepsize (Polak, 1997), and terminating condition $||dJ/da|| < 10^{-6}$. Figure

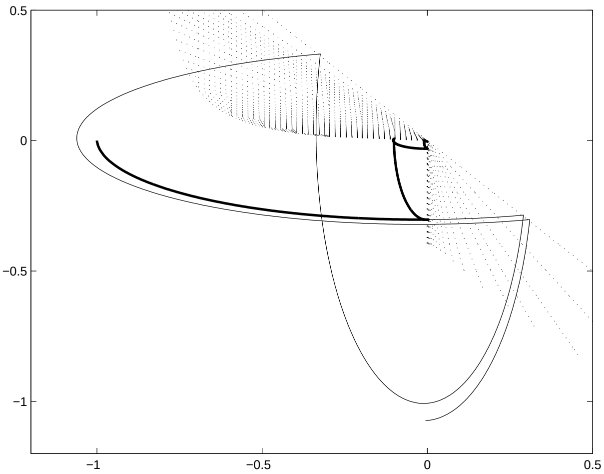

Fig. 3. Evolution of switching surfaces (dashed lines). System trajectories for nominal and terminal values of the switching parameters represented by thin and thick lines, respectively

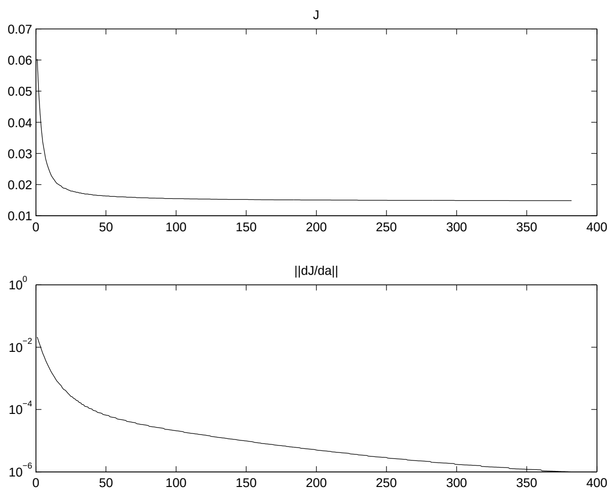

Fig. 4. Performance criteria of the gradient descent algorithm

3 illustrates the modifications incurred by the switching surfaces from the initial choice for the switching parameters $a_1 = a_2 = 1$, and how the system trajectory changes accordingly. Notice that, by the cost function chosen, the originally unstable system becomes stable through optimization of the switching surfaces.

REFERENCES

Arkin, R.C. (1998). *Behavior Based Robotics*. The MIT Press. Cambridge, MA.

Bemporad, A., A. Giua and C. Seatzu (2002). A master-slave algorithm for the optimal control of continuous-time switched affine systems. In: 41^{th} *IEEE Conf. on Decision and Control (CDC 2002)*. Las Vegas, USA. pp. 1976 –1981.

Boccadoro, M. and P. Valigi (2003). A modelling approach for the dynamic scheduling problem of manufacturing systems with non negligible setup times and finite buffers. In: *42nd IEEE Conference on Decision and Control (CDC '03)*. Maui, Hawaii, USA.

Branicky, M. S. (1998). Multiple lyapunov functions and other analysis tools for switched and hybrid systems. *IEEE Trans. on Automatic Control* **43**(4), 475–482.

Branicky, M., V. Borkar and S. Mitter (1998). A unified framework for hybrid control: model and optimal control theory. *IEEE Trans. on Automatic Control* **43**(1), 31–45.

Bryson, A. E. and Y. C. Ho (1969). *Applied Optimal Control*. Ginn and Co.

Cassandras, C. G., D. L. Pepyne and Y. Wardi (2001). Optimal control of a class of hybrid systems. *IEEE Trans. on Automatic Control* **46**(3), 398–415.

Egerstedt, M. (2000). Behavior based robotics using hybrid automata. In: *Lecture Notes in Computer Science: Hybrid Systems III: Computation and Control*. Pittsburgh, PA. pp. 103–116.

Egerstedt, M., Y. Wardi and H. Axelsson (n.d.). Transition time optimization for switched systems. *IEEE Trans. on Automatic Control, submitted*.

Giua, A., C. Seatzu and C. Van Der Mee (2001). Optimal control of switched autonomous linear systems. In: 40^{th} *IEEE Conf. on Decision and Control (CDC 2001)*. Orlando, FL, USA. pp. 2472–2477.

Johansson, M. and A. Rantzer (1998). Computation of piecewise quadratic lyapunov functions for hybrid systems. *IEEE Trans. on Automatic Control* **43**(4), 555–559.

Polak, E. (1997). *Optimization: Algorithms and Consistent Approximations*. Springer.

Riedinger, P., F. Kratz, C. Iung and C. Zanne (1999). Linear quadratic optimization for hybrid systems. In: *Proceedings of the 38th IEEE Conference on Decision and Control*. Phoenix, AZ. pp. 3059–3064.

Shaikh, M. S. and P.E. Caines (2002). On trajectory optimization for hybrid systems: theory and algorithms for fixed schedules. In: 41^{th} *IEEE Conf. on Decision and Control (CDC 2002)*. Las Vegas, USA. pp. 1997–1998.

Sussmann, H.J. (1999). A maximum principle for hybrid optimal control problems. In: *Proceedings of the 38th IEEE Conference on Decision and Control*. Phoenix, AZ. pp. 425–430.

Xu, X. and P.J. Antsaklis (2002). An approach to optimal control of switched systems with internally forced switchings. In: *Proceedings of the American Control Conference*. Anchorage, USA. pp. 148 –153.