OPTIMAL CONTROL OF SWITCHING TIMES IN HYBRID SYSTEMS

MAGNUS EGERSTEDT, YORAI WARDI, HENRIK AXELSSON

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

Abstract. This paper considers the problem of determining optimal switching times at which mode transitions should occur in multi-modal, hybrid systems. An expression for the gradient of the cost functional, defined with respect to the switching times, is presented in such a way that not only the switching times, but also the number of switches can be incorporated in the problem formulation. Numerical examples testify to the viability of the proposed approach.

Key Words. Switched dynamical systems, switching-time control, optimal control, hybrid systems, gradient-descent algorithms.

1. INTRODUCTION

Switched dynamical systems are often described by differential inclusions of the form

$$\dot{x}(t) \in \{f_\alpha(x(t), u(t))\}_{\alpha \in A}, \tag{1}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^k$, and $\{f_\alpha : \mathbb{R}^{n+k} \to \mathbb{R}^n\}_{\alpha \in A}$ is a collection of continuously differentiable functions, parameterized by α belonging to some given set A. The time t is confined to an interval $[0, T]$, where it is possible that $T = \infty$. Such systems arise in a variety of applications, including situations where a control module has to switch its attention among a number of subsystems [10, 12, 15], or collect data sequentially from a number of sensor sources [3, 5, 9]. A supervisory controller is engaged for dictating the switching law, i.e. the rule for switching among the functions f_α in the right-hand side of Eq. (1).

Recently, there has been a mounting interest in optimal switching time control of such hybrid systems, where the control variable consists of a proper switching law as well as the input function $u(t)$ (see [2, 4, 8, 13, 14, 16, 17]). A special class of these problems concerns autonomous systems, where the term $u(t)$ is absent, and the control variable consists solely of the switching times [9, 18].

Ref. [18] can be thought of as providing the starting point for the results presented in this paper, and we start by considering the problem in [18], where the sequence of switching functions as well as the number of switching times are fixed. We develop a formula, simpler than the one in [18], for the gradient of the cost functional, and combine it with an Armijo step-size gradient-descent algorithm. We moreover note that the gradient formula has a special structure that lends itself, in a natural way, to include the number of switching times and the sequence of switching functions as parameters to be optimized.

2. PROBLEM FORMULATION AND GRADIENT COMPUTATION

We consider the following problem, addressed in [6]: Let $\{f_i\}_{i=0}^N$ be a finite sequence of continuously differentiable functions from $\mathbb{R}^n$ to $\mathbb{R}^n$. Fix $T > 0$ and $x_0 \in \mathbb{R}^n$. Now, given a sequence of switching times τ_i, $i = 1, \ldots, N$, and let $\tau_0 = 0$ and $\tau_{N+1} = T$, the resulting dynamical system takes on the form

$$\dot{x}(t) = f_i(x(t)) \tag{2}$$

for all $t \in [\tau_i, \tau_{i+1})$ and for every $i \in \{0, \ldots, N\}$, with the given initial condition $x(0) = x_0$. Let furthermore $L : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function, and consider the cost functional J, defined by

$$J = \int_0^T L(x(t))dt. \tag{3}$$

We now consider the control parameter to consist of the switching times $\tau_1, \ldots, \tau_N$, and denote it by the N-dimensional variable $\bar{\tau} = (\tau_1, \ldots, \tau_N)^T$. Note that J is a function of $\bar{\tau}$ via Eq. (2), whose gradient is given in the following manner: The costate $p(t) \in \mathbb{R}^n$ associated with the dynamical system defined by Eq. (2) and the cost functional defined by Eq. (3), is given by the backward equation

$$\dot{p}(t) \;=\; -\Big(\frac{\partial f_i}{\partial x}(x(t))\Big)^T p(t) - \Big(\frac{\partial L}{\partial x}(x(t))\Big)^T, \quad (4)$$

where $t \in [\tau_i, \tau_{i+1}),\ i = N, \ldots, 1$, and $p(T) = 0$. It should be noted that $p(\tau_i)$ is defined by the regression (in time) of the above equation in the interval $[\tau_i, \tau_{i+1}]$, and the costate function $p(t)$ is assumed to be continuous at the point τ_i. Based on this costate equation, the following result was proved in [6]:

Theorem 2.1[6] The following equation holds for every $i \in \{1, \ldots, N\}$,

$$\frac{dJ(\bar{\tau})}{d\tau_i} \;=\; p(\tau_i)^T \Big(f_{i-1}(x(\tau_i)) - f_i(x(\tau_i))\Big). \quad (5)$$

3. GRADIENT DESCENT

This section uses the gradient formula (5) in a descent algorithm. The particular algorithm that we apply is the steepest descent algorithm with Armijo stepsizes [1, 11]. Given an iteration point $\bar{\tau}(k)$ (where k is the iteration index), the next iteration point, $\bar{\tau}(k+1)$ is computed in the following way.

Steepest-descent algorithm with Armijo stepsizes.
Parameters: $\alpha \in (0,1)$ and $\beta \in (0,1)$.
Step 1: Compute $h(k) = \nabla J(\bar{\tau}(k))$.
Step 2: Compute $i(k)$, defined as the smallest non-negative i such that $J(\bar{\tau}(k) - \beta^i h(k)) - J(\bar{\tau}(k)) \leq -\alpha\beta^i ||h(k)||^2$.
Step 3: Set $\lambda(k) = \beta^{i(k)}$, and set $\bar{\tau}(k+1) = \bar{\tau}(k) - \lambda(k)h(k)$.

This algorithm is globally convergent to stationary points, and it does not jam at non-stationary points (see [11]). Practically, the parameters α and β often are set to $\alpha = \beta = 0.5$. The search for $i(k)$ in *Step 2* need not start at $i = 0$, rather, for improving the algorithm's efficiency, it can start at $i(k-1) - 1$ or $i(k-1) - 2$. (See [11].)

4. EXAMPLES

In order to verify the numerical feasibility of the proposed algorithm, we tested the gradient descent optimization on a collection of cases, different from those investigated in [6]. For example, let the Armijo parameters be given by $\alpha = \beta = 0.5$, and consider an experiment with three switches, initialized to $\tau_1 = 0.25,\ \tau_2 = 0.5,\ \tau_3 = 0.75$.

Let the system evolve over the time interval $[0, 1]$, and let the individual dynamical systems be given by the two-dimensional, linear systems

$$\begin{array}{l} \dot{x} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} x = A_1 x,\ t \in [0, \tau_1) \\ \dot{x} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} x = A_2 x,\ t \in [\tau_1, \tau_2) \\ \dot{x} = A_1 x,\ t \in [\tau_2, \tau_3) \\ \dot{x} = A_2 x,\ t \in [\tau_3, 1]. \end{array}$$

The cost function used in this example is

$$J = \frac{1}{2}\int_0^T ||x(t)||^2 dt. \quad (6)$$

Note that the dynamics switch between modes where either the first or the second component of x is unstable, which means that we can expect the optimal switching-time sequence to be such that both systems get enough “attention” in order to prevent any component of x from diverging. The result is shown in Figure 1, and it was found that the optimization algorithm terminated after 5 iterations. The optimal switching times were furthermore found to be $\tau_1 = 0.2195,\ \tau_2 = 0.5213,\ \tau_3 = 0.7405$.

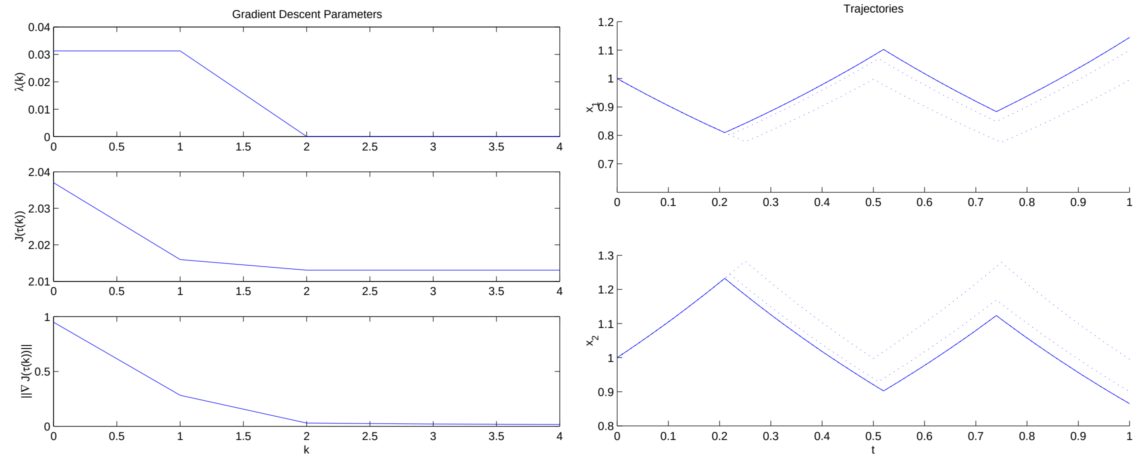
Now, it is conceivable that the algorithm may produce switching times where some switches are eliminated (either by having $\tau_1 = 0,\ \tau_N = T$ or $\tau_i = \tau_{i+1}$). This is the case in the following modified version of the previous example:

$$\begin{array}{l} \tau_1 = 0.25,\ \tau_2 = 0.5,\ \tau_3 = 0.75 \\ A_1 = \begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix} \\ A_2 = \begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix} \\ \dot{x} = A_1 x,\ t \in [0, \tau_1) \\ \dot{x} = A_2 x,\ t \in [\tau_1, \tau_2) \\ \dot{x} = A_1 x,\ t \in [\tau_2, \tau_3) \\ \dot{x} = A_2 x,\ t \in [\tau_3, 1]. \end{array}$$

Using the same Armijo parameters as in the previous example, the optimal switching times were found to be $\tau_1 = 0,\ \tau_2 = 0.6978,\ \tau_3 = 0.6978$. What this means is that the system never evolves according to $\dot{x} = A_1 x$ since the intervals on which this dynamics is defined are empty (on $[0, \tau_1)$ and $[\tau_2, \tau_3)$). The results are depicted in Figure 2.

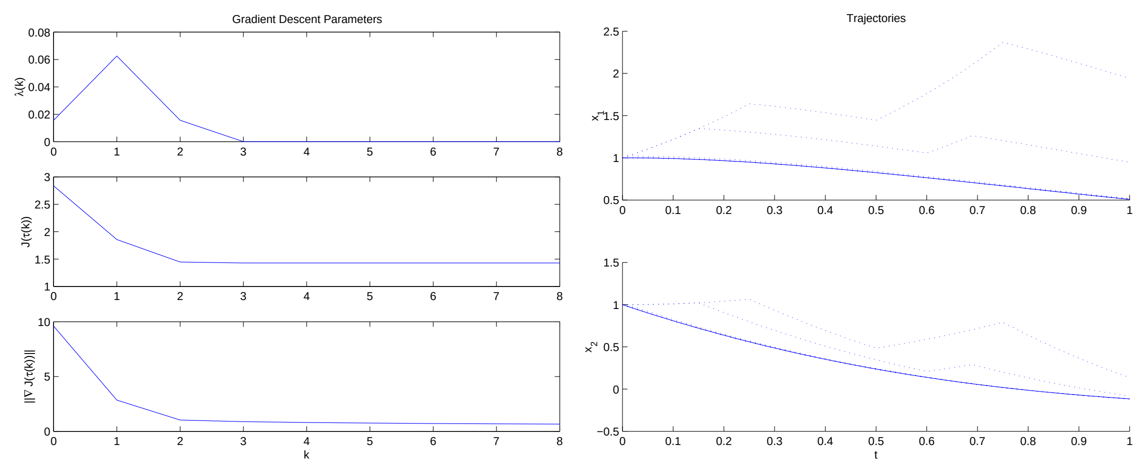
5. OPTIMIZATION OF THE NUMBER OF SWITCHING TIMES

The gradient $\nabla J(\bar{\tau})$, based on the formula for the partial derivatives, given in Eq. (5), has a special structure that makes it possible to extend and enhance gradient-descent algorithms in a number of directions. This special structure stems from the fact that the same costate $p(t)$ is used, in (5), for all the partial derivatives $\frac{dJ(\bar{\tau})}{d\tau_i}$, $i \in \{1, \ldots, N\}$. Thus, the computation of these partial derivatives involves two stages:



(a) (b)

Figure 1: Three Switches: In the left figure, the Armijo stepsize $\lambda(k)$ is shown (top) together with the cost J (middle) and $\|\frac{\partial J}{\partial \bar{\tau}}\|$ (bottom). In the right figure the two x-components are displayed as functions of t. The dotted lines depict the state trajectories associated with the variable $\bar{\tau}$ computed in odd-numbered iterations (iteration 1,3, etc.), and the solid line corresponds to the final trajectory.



(a) (b)

Figure 2: Three Switches: The dynamics in this particular example is such that no switches do in fact occur, which is correctly detected by the proposed algorithm.

(i) Computing $p(t)$ by first solving (numerically) Eq. (2) forwards and then solving Eq. (4) backwards; and

(ii) Evaluating the terms $f_{i-1}(x(\bar{\tau}_i)) - f_i(x(\bar{\tau}_i))$ and multiplying them by $p(\tau_i)^T$ to obtain the partial derivatives (5).

This special structure allows extensions of the optimization technique discussed here in a number of directions. In the following paragraph we discuss one such extension, to the case where the number of switching times (N) is a variable parameter. To explain this point, consider the following simple case, where the right-hand side of Eq. (2) alternates between two given functions, $g_1(x)$ and $g_2(x)$, in a round-robin fashion. Thus, assuming that g_1 and g_2 are continuously differentiable functions from $\mathbb{R}^n$ to $\mathbb{R}^n$, suppose that $f_i = g_1$ for all odd i, and $f_i = g_2$ for all even i.

Now, consider an interval $[\tau_i, \tau_{i+1}]$ for some odd i, so that $\dot{x} = g_1(x)$ throughout that interval. Fix $t \in (\tau_i, \tau_{i+1})$, and let $\lambda > 0$ be such that the interval $(t-\lambda/2, t+\lambda/2)$ is contained in the interval (τ_i, τ_{i+1}). We now insert two switching times, one at $t - \lambda/2$ and one at $t + \lambda/2$, and let the switching function be g_2 in the interval $[t - \lambda/2, t + \lambda/2)$. Considering J as a function of $\lambda \geq 0$, it has the following one-sided derivative at $\lambda = 0$ (see [7]):

$$\frac{dJ(0)}{d\lambda^+} = p(t)^T \Big(g_2(t) - g_1(t) \Big). \qquad (7)$$

With this equation, it is possible to extend the gradient-descent algorithm discussed in Section 3 to include the number and location of the switching points.

6. CONCLUSIONS

This paper presents an algorithm for computing optimal switching times for mode transition systems. A key aspect of the algorithm is the formula for the gradient, consisting of two elements: a common costate, and various time-dependent function evaluations. The computation of the costate generally imposes a larger computational burden than the function evaluations. However, once computed, the same costate can be used for all the partial derivatives of the cost functional. The obtained gradient formula is applied in conjunction with a steepest-descent optimization algorithm.

The special structure of the gradient can be exploited by extending the scope of the gradient-based algorithm to situations where the number of the switching points is part of the variable parameter.

Acknowledgments

The work of the Magnus Egerstedt was sponsored in part by ECS NSF-CAREER award (grant # 0237971). The work by Yorai Wardi and Henrik Axelsson was sponsored through the Georgia Institute of Technology Manufacturing Research Center.

REFERENCES

1. L. Armijo. Minimization of Functions Having Lipschitz Continuous First-Partial Derivatives. *Pacific Journal of Mathematics*, Vol. 16, ppm. 1-3, 1966.

2. M.S. Branicky, V.S. Borkar, and S.K. Mitter. A Unified Framework for Hybrid Control: Model and Optimal Control Theory. *IEEE Transactions on Automatic Control*, Vol. 43, pp. 31-45, 1998.

3. R. Brockett. Stabilization of Motor Networks. *IEEE Conference on Decision and Control*, pp. 1484–1488, 1995.

4. J. Chudoung and C. Beck. The Minimum Principle for Deterministic Impulsive Control Systems. *IEEE Conference on Decision and Control*, Vol. 4, pp. 3569 - 3574, Dec. 2001.

5. M. Egerstedt and Y. Wardi. Multi-Process Control Using Queuing Theory. *IEEE Conference on Decision and Control*, Las Vegas, NV, Dec. 2002.

6. M. Egerstedt, Y. Wardi, and F. Delmotte. Optimal Control of Switching Times in Switched Dynamical Systems. Submitted to *IEEE Conference on Decision and Control*, Maui, Hawaii, Dec. 2003.

7. M. Egerstedt, Y. Wardi, and H. Axelsson. Simultaneous Scheduling and Switching Time Optimization for Hybrid Systems. *Technical Memorandum, School of Electrical and Computer Engineering, Georgia Institute of Technology*, 2003.

8. S. Hedlund and A. Rantzer. Optimal Control of Hybrid Systems. *Proceedings of the 38th IEEE Conference on Decision and Control*, pp. 3972-3977, 1999.

9. D. Hristu-Varsakelis. Feedback Control Systems as Users of Shared Network: Communication Sequences that Guarantee Stability. *IEEE Conference on Decision and Control*, pp. 3631–3631, Orlando, FL, 2001.

10. B. Lincoln and A. Rantzer. Optimizing Linear Systems Switching. *IEEE Conference on Decision and Control*, pp. 2063–2068, Orlando, FL, 2001.

11. E. Polak. *Optimization Algorithms and Consistent Approximations*. Springer-Verlag, New York, New York, 1997.

12. H. Rehbinder and M. Sanfirdson. Scheduling of a Limited Communication Channel for Optimal Control. *IEEE Conference on Decision and Control*, Sidney, Australia, Dec. 2000.

13. M.S. Shaikh and P. Caines. On Trajectory Optimization for Hybrid Systems: Theory and Algorithms for Fixed Schedules. *IEEE Conference on Decision and Control*, Las Vegas, NV, Dec. 2002.

14. H.J. Sussmann. Set-Valued Differentials and the Hybrid Maximum Principle. *IEEE Conference on Decision and Control*, Vol. 1, pp. 558 -563, Dec. 2000.

15. G. Walsh, H. Ye, and L. Bushnell. Stability Analysis of Networked Control Systems. *American Control Conference*, pp. 2876–2880, 1999.

16. L.Y. Wang, A. Beydoun, J. Cook, J. Sun, and I. Kolmanovsky. Optimal Hybrid Control with Applicatiuons to Automotive Powertrain Systems. In *Control Using Logic-Based Swirtching*, Vol. 222 of LNCIS, pp. 190-200, Springer-Verlag, 1997.

17. X. Xu and P.J. Antsaklis. An Approach for Solving General Switched Linear Quadratic Optimal Control Problems. In *Proceedings of the 40th CDC*, pp. 2478-2483, 2001.

18. X. Xu and P. Antsaklis. Optimal Control of Switched Autonomous Systems. *IEEE Conference on Decision and Control*, Las Vegas, NV, Dec. 2002.