

# Optimal Control of Switched Dynamical Systems under Dwell Time Constraints

Usman Ali<sup>1</sup> and Magnus Egerstedt<sup>2</sup>

**Abstract**—This paper addresses the problem of optimally scheduling the mode sequence and mode duration for switched dynamical systems under dwell time constraints that describe how long a system has to stay in a mode before they can switch to another mode. The schedule should minimize a given cost functional defined on the state trajectory. The topology of the optimization space for switched dynamical systems with and without dwell time constraints is investigated and it is shown that the notion of local optimality must be replaced by stationarity with regards to a suitably chosen optimality function when dwell time constraints are present. Hence, an optimality function is proposed to characterize the solution to the dwell time problem as points that satisfy optimality condition defined in terms of optimality function. A conceptual algorithm is presented to solve the mode scheduling problem and its convergence to stationary points is proved. A numerical example is given to highlight the algorithm.

## I. INTRODUCTION

The problem of determining optimal switching law for systems that switch between different modes has been extensively investigated in recent years since such systems arise in a number of application areas and results of theoretical and computational significance have been derived; see e.g recent survey [1] and references therein. The switching law, which governs which mode is active at a given instant of time is defined using control variables that describe the sequence of modes and the duration of each mode and commonly referred to as mode schedule when considered together [2], [3].

While in theory, we can switch infinitely many times between different modes in a finite amount of time, most physical systems have to spend some minimum time in a mode before they can switch to another mode due to mechanical reasons, power constraints, information delays, stability considerations etc. This minimum time is known as the dwell time of the mode, a term previously used in the context of stability of switched linear systems [4], [5]. This important and practical constraint has not received attention in research on optimal control of switched systems.

In this paper, we consider the common optimal control problem of minimizing a cost functional defined on the state trajectory of nonlinear switched dynamical system under dwell time constraints where the control input consists of mode schedule only. The main contribution of this paper is the investigation of the topology of the optimization space under dwell time constraints and consequently, the need for

replacing the notion of local optimality by stationarity with regards to an optimality function. Secondly, an optimality function is proposed to characterize the optimal solution. Finally, a conceptual algorithm to solve the optimal mode scheduling problem based on this optimality function is presented and its convergence is proved.

Two fundamental problems investigated in the area of switched optimal control are the timing optimization problem and scheduling optimization problem. In first case, the mode sequence is kept fixed and optimization is performed over switching times between different modes while in the later case, the sequence of modes as well as the switching times between modes are the control variables. On the theoretical side, a general framework for optimal control problem in hybrid systems was formulated in [6]. Variations of the maximum principle have been derived in [7], [8], [9], [10].

For the timing optimization problem, the control variable is continuous which makes it amenable to nonlinear programming techniques and consequently algorithms were first developed to solve the timing optimization problem. In the case of scheduling optimization problem, the control variable has a continuous as well as discrete component which renders the problem significantly complex. A number of computational strategies have been developed to solve the scheduling problem. See [1] for a survey of these results as well as results for timing optimization problem. These optimization techniques do not impose any dwell time constraints.

While the gradient descent algorithms in [2], [3] and [11] are shown to be convergent, the question is what do they converge to, i.e. what is the local minima which leads to the question of what is the topology of the optimization space and whether these results can be extended to the case when the optimization space is restricted due to dwell time constraints. To answer this question, we explore in detail the topologies generated by the metrics used in [11] and [3] and investigate the impact of dwell time constraints on the structure of optimization space. We show that in the presence of dwell time constraints, the optimization space lacks structure to define a local minima. Hence, we propose to replace the notion of local optimality by stationarity and define the optimal solution for the scheduling problem as stationary points that satisfy optimality condition defined in terms of optimality function.

While the optimality functions in [2], [3] and [11] use insertion gradients derived in [12] using variational principles, the presence of dwell time constraint makes such techniques from calculus inapplicable since these methods insert modes at a point. To address this issue, we propose an optimality

<sup>1</sup>Usman Ali and <sup>2</sup>Magnus Egerstedt are with School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA. Email: usmanali@gatech.edu, magnus@ece.gatech.edu

\*Research supported in part by NSF through Grant CNS-1239225

function defined in terms of the cost differential resulting from mode insertion for duration equal to dwell time. In this paper, we assume that each mode has the same minimum dwell time.

Lastly, we present a conceptual algorithm to solve the dwell time problem that uses a two step strategy similar to the ones in [2] and [11]. At the lower level, we solve the timing optimization problem with dwell time constraints using gradient projection method. At the higher level, we do a single mode insertion for duration of dwell time at the point that gives the maximum decrease in the cost while ensuring that the dwell time constraints after mode insertion are still respected. The algorithm convergence is proved using the property of sufficient descent.

The paper organization is as follows. In section 2, we formulate the problem and define the optimization space. We discuss the topology of optimization space and define the optimality function in section 3. In section 4, we give a conceptual algorithm to solve the dwell time problem and prove its convergence. In section 5, a numerical example is given to show results of the algorithm. Section 6 concludes the paper.

## II. PROBLEM FORMULATION

In this paper, we consider nonlinear switched dynamical systems in which there is no external input. The dynamics of such systems can be mathematically described by differential equations of the form

$$\dot{x}(t) = f_{v(t)}(x(t)), \quad (1)$$

where  $\Phi = \{f_q : q \in \mathcal{N}\}$  is a family of modal functions from  $R^n$  to  $R^n$  parameterized by finite index set  $\mathcal{N} = \{1, 2, \dots, Q\}$  and  $v : [0, T] \rightarrow \mathcal{N}$  is a left continuous piecewise constant function of time called switching signal.

The switching signal  $v(t)$  corresponds to the mode schedule  $\xi = (\sigma, \tau)$  through a bijection as follows. Let  $\tau = (\tau_1, \tau_2, \dots, \tau_n)'$  be the vector corresponding to monotone increasing transition times in  $v(t)$  in interval  $(0, T)$ . Let  $\sigma = \{\sigma(1), \sigma(2), \dots, \sigma(n+1)\}$  denote the mode sequence, where  $\sigma(i) = v(\tau_i)$  is an element of  $\mathcal{N}$ , so  $\sigma \in \mathcal{N}^{n+1}$ . If we define  $\tau_0 = 0$  and  $\tau_{n+1} = T$ , then  $v(t) = \sigma(i)$  for all  $t \in [\tau_{i-1}, \tau_i]$  and  $i = 1, 2, \dots, n+1$  and  $\xi = (\sigma, \tau)$  is the corresponding mode schedule. We define the length of mode schedule  $\xi$  to be the same as the length of mode sequence  $\sigma$ . For schedule of length  $n+1$ , (1) can be written in the form

$$\dot{x}(t) = f_{\sigma(i)}(x(t)) \quad \forall \quad t \in [\tau_{i-1}, \tau_i] \quad (2)$$

and  $i = 1, \dots, n+1$ . While (1) is more compact, we will use representation in (2) in this paper since it explicitly shows the mode schedule.

Suppose the minimum dwell time for each mode is some  $\delta > 0$  and we are optimizing over finite time horizon. Then the maximum number of modes that can exist in the interval  $[0, T]$  is  $N = \lfloor \frac{T}{\delta} \rfloor$  where  $\lfloor \cdot \rfloor$  denotes the floor operator. A mode schedule acts as a feasible control input for our system

if it satisfies the dwell time constraints

$$\Delta\tau_i = \tau_i - \tau_{i-1} \geq \delta, \quad (3)$$

for all  $i = 1, \dots, n+1$ . The collection of all such feasible mode schedules constitute the optimization space  $\mathcal{X}$ . With this in mind, we explicitly define our optimization space.

For any positive integer  $n$ , the mode sequence  $\sigma$  is an element of  $\Sigma^n = \mathcal{N}^n$ . Then the mode sequence space

$$\Sigma = \bigcup_{n=1}^N \Sigma^n,$$

is collection of all such mode sequences. The mode sequence space is a finite space.

Also for any positive integer  $n$ , the dwell time constraints (3) define a polyhedron

$$\mathcal{S}^n = \{\tau \in R^n | a_j \tau \geq b_j, j = 1, 2, \dots, n\}, \quad (4)$$

where  $a_1 = (1, 0, 0, \dots, 0)$ ,  $b_1 = \delta$ ,  $a_n = (0, \dots, 0, -1)$ ,  $b_n = \delta - T$  and for the rest  $a_j = (0, \dots, -1, 1, \dots, 0)$ ,  $b_j = \delta$  with the non zero entries of  $a_j$  at the  $j$  and  $j+1$  positions. The transition time space is

$$\mathcal{S} = \bigcup_{n=1}^N \mathcal{S}^n.$$

The dwell time constraints automatically ensure that the transitions times  $\tau_i$  are monotone increasing, i.e  $\tau_{j-1} < \tau_j$  for all  $j = 1, \dots, n+1$ .

For any positive integer  $n$ , let  $\mathcal{X}^n = \Sigma^n \times \mathcal{S}^{n-1}$ . So the mode schedule  $\xi \in \mathcal{X}^n$  is the tuple  $\xi = (\sigma, \tau)$ . The optimization space, which is the collection of all such mode schedules is then

$$\mathcal{X} = \bigcup_{n=1}^N \mathcal{X}^n. \quad (5)$$

This optimization space is finite dimensional and is a subset of the infinite dimensional optimization space that would result if no dwell time constraints were considered.

### Dwell Time Problem

Let  $L : R^n \rightarrow R$  be a cost function defined on the state trajectory  $x(t)$ , and consider the cost functional  $J : \mathcal{X} \rightarrow R$ , defined by

$$J(\xi) = \int_0^T L(x(t)) dt. \quad (6)$$

Let  $x_0 \in R^n$  be the initial condition of the system (2). The dwell time problem is to solve

$$\min_{\xi \in \mathcal{X}} J(\xi), \quad (7)$$

subject to the dynamical constraints (2) and initial condition constraints. To ensure that a unique bounded solution exists for the differential equation, we make the following mild assumptions which are fairly standard in switched optimal control problems [2], [13].

*Assumptions:*

- (i) The modal functions  $f \in \Phi$  and cost function  $L$  are twice continuously differentiable.
- (ii) There exists a constant  $C > 0$  such that for every  $x \in R^n$  and every  $f \in \Phi$  we have  $\|f(x)\| \leq C(\|x\| + 1)$ .

### III. TOPOLOGY, LOCAL MINIMA AND OPTIMALITY FUNCTION

In this section, we begin with an explanation of the need for defining the minima for the dwell time problem using optimality functions and why defining local minima for such problems is not useful. The definition of local minima for the optimization space requires the concept of neighborhood which in turn depends on the topology defined on the space. While many topologies can be defined for the same optimization space, the choice of topology should be such that it resonates well with the underlying problem.

*Definition 1:* Given a topological space  $(X, T)$ , a subset  $N_x$  of  $X$  is a neighborhood of a point  $x \in X$  if  $N_x$  contains an open set  $U \subset T$  containing the point  $x$ .

To motivate what is to follow about the topology on our spaces, we consider few scenarios concerning the open neighborhoods and local minima. Suppose we want to find local minima of a continuous function  $f : R \rightarrow R$ . The decision variable in this case is continuous in the sense it takes values in  $R$  and we call a point  $x_0 \in R$  as its local minimum point if we can find an  $\epsilon > 0$  ball such that  $f(x_0) \leq f(x)$  for all  $x$  in this ball. The  $\epsilon$  ball defines a neighborhood of point  $x_0$  in this case. Suppose now the decision variable is discrete, for example, the set of integers and let say  $f : Z \rightarrow R$ . The  $\epsilon$  ball definition of local minima does not make sense in this case since for  $0 < \epsilon < 1$  the ball will contain only the point itself thereby making every point as a local minima. A more reasonable way of declaring a point  $n$  as a local minima would be if  $f(n)$  is less than  $f(n-1)$  and  $f(n+1)$ . The points  $\{n-1, n+1\}$  constitute the neighborhood of point  $n$ .

In case of the transition time space, the cost functional  $J$  is a continuous function of the decision variable  $\tau \in R^n$ . Hence the choice for topology is the usual topology of  $R^n$  obtained by the cartesian product of open intervals in  $R$ . However the mode sequence space is a discrete space and one would think that a reasonable definition of local minima for function defined on this space would be some natural extension of the local minima defined above for a discrete space. The problem is that while in the case of integers for example, it was easy to see that the points  $n-1$  and  $n+1$  are the neighbors of point  $n$ , we don't have such an obvious extension in the case of mode sequence space. To address this problem, one would like to define a metric on the space which can tell us which points are closer to a point than others. The question is does such a metric exist and is it useful. We first look at this question without dwell time constraints and then in the presence of dwell time constraints.

In [11], a metric is defined on the mode sequence space in a form inspired by the notion of Hamming distance between vectors and a metric on transition time space using  $\ell_1$  norm. The norm on  $\mathcal{X}$  is then defined as

$$d(\xi_x, \xi_y) = \{\sigma_x \neq \sigma_y\} \bar{1} + \|\tau_x - \tau_y\|_1, \quad (8)$$

and the algorithm convergence is proved using the property of sequential continuity in metric spaces. The topology

generated by this metric is such that the  $\epsilon$  neighborhood of a point  $\xi_0$  will contain only points of transition time space and no points from the mode sequence space. With dwell time constraints imposed on the optimization space, the situation remains the same.

If we use metric (8) and choose open balls of radius 2 to define neighborhoods so that it contains mode sequence points, we still run into problems. To see the neighborhoods generated using open balls of radius 2, consider the case where  $\mathcal{N} = \{1, 2, 3\}$  and  $N = 2$  due to dwell time constraints. Then the smallest non-trivial neighborhood of mode sequence point  $(1, 2)$  include the sequence points  $(1, 3)$  and  $(3, 2)$  both of which are distance 1 apart from the point under consideration. The point  $(1, 2)$  will be then locally optimal w.r.t. cost functional  $J$  if the cost associated with this mode sequence is less than its two neighboring points which constitute its neighborhood. The sequence point  $(3, 1)$  is distance 2 apart from point  $(1, 2)$  in this metric.

While the definitions are consistent, this does not resonate well with the problem we are trying to solve. To see this, even for fixed switching times, the state trajectory  $x(t)$  resulting from mode sequence  $(3, 1)$  might be closer to the optimal state trajectory resulting from  $(1, 2)$  than the mode sequences  $(1, 3)$  and  $(3, 2)$  in  $L^p$  norm which were in its neighborhood and thus does not appeal to our intuitive understanding of the local minima for the actual problem. This problem however persists even if no dwell time constraints were considered.

In [3], the mode schedules are represented using switching signals  $v(t)$  and the topology is induced on the optimization space  $\mathcal{V}$  using the  $L_1$  norm

$$d(v_x, v_y) = \|v_x - v_y\|_{L_1}. \quad (9)$$

In the topology generated by this metric, the  $\epsilon$  neighborhood of a point  $\xi_0$  contains points of mode schedule space as well as transition time space when considered in the corresponding mode schedule representation form. When two switching signals are close to each other in the  $L_1$  norm, the corresponding state and co-state trajectories are close to each other in the  $L_\infty$  norm. The cost function  $J$  is thus a continuous function of  $v(t)$  and a useful definition of local minima is obtained that resonates well with the problem.

The scenario changes however when dwell time constraints are imposed on the optimization space. The  $\epsilon$  neighborhoods generated by metric (9) under dwell time constraints do not contain mode sequence points but just transition time points when considered in the corresponding mode schedule representation form. Here, the problem primarily lies not in the metric used but in the inherent lack of structure in the optimization space when dwell time constraints are present. There exists no meaningful similarities between different elements of the optimization space. In such cases the feasible topology is the trivial topology in which all points are neighbors to any point [14].

In summary, to use the topological definition of local minima would mean to use trivial topology for the mode sequence space and in which case we have to be globally optimal with respect to the mode sequence space. This

requires that we solve the timing optimization problem for every mode sequence which is not feasible. The local minima in such a case results from being locally optimal with respect to time only.

Thus to define the minima for our dwell time problem, we resort to the use of optimality functions which are semicontinuous functions of the form  $\theta : X \rightarrow R^-$ . Points in  $X$  that satisfy the optimality condition  $\theta_X(x) = 0$  are referred to as stationary points and constitute the optimal solutions for the dwell time problem. The magnitude of optimality function  $|\theta_X(\cdot)|$  can be seen as a measure by which point  $\xi \in \mathcal{X}$  fails to satisfy the optimality condition.

The optimality functions used in [2], [3] and [11] rely on insertion gradient which tests the sensitivity of cost due to each mode over an arbitrarily small interval. Due to the presence of dwell time constraints, such calculus based gradient insertion at a point is not possible. We propose an optimality function based on the cost differential resulting from a feasible mode insertion for duration of dwell time  $\delta$ .

*Definition 2:* A mode insertion is feasible if the resulting mode sequence does not violate the dwell time constraints.

Let  $U \subset [0, T]$  be set of all points where the mode insertion is feasible. At given iteration  $k$ , let  $\sigma_0$  be time optimized mode sequence to which we want to insert a new mode. Let  $J_{k0}$  be the cost associated with this mode sequence. Let  $J(t, f_j)$  denote the cost associated with mode sequence obtained by replacing the mode at time  $t \in U$  by mode  $f_j$  for  $\delta$  seconds. Define

$$J_\Delta(t, f_j) = J(t, f_j) - J_{k0}. \quad (10)$$

We then define our optimality function as

$$\theta_{\mathcal{X}} = \min \{J_\Delta(t, f) : f \in \Phi, t \in U\}. \quad (11)$$

Now  $\theta_{\mathcal{X}} \leq 0$  since we can always insert the same mode. The condition  $\theta_{\mathcal{X}} = 0$  is then the necessary condition for optimality which we refer to as the optimality condition. We define the optimal solution for the dwell time problem as:

*Definition 3:* A point  $\xi_0 = (\sigma_0, \tau_0) \in \mathcal{X}$  is optimal solution for the dwell time problem (7) if  $\tau_0$  is Karush Kuhn Tucker (KKT) point for the timing optimization problem and  $\theta_{\mathcal{X}}(\xi_0) = 0$ .

#### IV. ALGORITHM

In this section we present a conceptual algorithm that employs a basic two step strategy similar to the one used in [2] and [11]. At the higher level, the mode sequence is optimized by inserting a single mode for  $\delta$  seconds and at the lower level, for a fixed mode sequence, the switching times between different modes are optimized subject to the dwell time constraints. We begin with the switch time optimization problem first.

##### A. Switch Time Optimization

For a fixed mode sequence  $\sigma_0$  of length  $n + 1$ , the cost functional in (6) is only function of  $\tau \in \mathcal{S}^n$  and the switch time optimization problem becomes

$$\min_{\tau \in \mathcal{S}^n} J(\tau) \quad (12)$$

subject to (2) and initial condition constraints. In [12], the derivative of the cost with respect to switching times for this problem without dwell time constraints is shown to be

$$\frac{dJ}{d\tau_i} = p(\tau_i) (f_{\sigma_0(i)}(x(\tau_i)) - f_{\sigma_0(i+1)}(x(\tau_i))) \quad (13)$$

where  $x(t)$  is the solution of the system (2) and  $p(t)$  is solution of the costate equation

$$\dot{p}(t) = -p(t) \frac{\partial f_{\sigma_0}(x(t))}{\partial x} - \frac{\partial L(x(t))}{\partial x}, \quad p(T) = 0. \quad (14)$$

For sets that are convex and compact, we can find the feasible descent direction by taking the projection of vector found using steepest descent method onto the constraint set, which in general requires solving a quadratic optimization problem. When the constraint set has the structure of polyhedron, the direction finding problem can be greatly simplified using manifold suboptimization methods based on [15] which are a type of gradient projection methods. In this method, instead of projection onto the entire constraint set, the gradient is projected on a linear manifold of active constraints which makes computation of the projection quite easy [16]. Since our constraint set  $\mathcal{S}^n$  has the structure of a polyhedron, we use manifold suboptimization method to solve our timing optimization problem.

Let  $\mathcal{I}(\tau_k) = \{j | a_j \tau_k = b_j, j = 1, 2, \dots, r\}$  denote the index set corresponding to active constraints at feasible point  $\tau_k$ . We assume without loss of generality that vectors  $\{a_j, j \in \mathcal{I}(\tau_k)\}$  are linearly independent at every  $\tau_k$ . The feasible descent direction is obtained from the subspace  $M(\tau_k) = \{h | a_j h_k = b_j, j \in \mathcal{I}(\tau_k)\}$ . The projection of the gradient onto  $M(\tau_k)$  can be obtained by solving the following quadratic optimization problem.

$$\min_{h_k \in M(\tau_k)} \nabla J(\tau_k)' h_k + \frac{1}{2} h_k' h_k.$$

The unique optimal solution of this problem can be easily computed from Karush Kuhn Tucker (KKT) conditions to be

$$h_k = -P_M \nabla J(\tau_k), \quad (15)$$

where  $P_M = I - L_k' (L_k L_k')^{-1} L_k$  is the projection matrix and  $L_k$  is the matrix that has as rows the vectors  $a_j, j \in \mathcal{I}(\tau_k)$ . The KKT multiplier  $\mu_k$  for this problem is given by

$$\mu_k = -(L_k L_k')^{-1} L_k' \nabla J(\tau_k). \quad (16)$$

To ensure that the algorithm converges, the step size in the descent direction  $h_k$  is obtained by using Armijo step size rule [17] over the set  $\Lambda = \{\lambda > 0 | a_j'(\tau^k + \lambda h_k) \geq b_j, j \notin \mathcal{I}(\tau_k)\}$ . From this we can compute the upper limit on  $\lambda_k$  as

$$\lambda = \min_{j \notin \mathcal{I}(\tau_k)} \frac{a_j \tau_k - b_j}{a_j h_k}, \quad (17)$$

such that  $a_j h_k < 0$ , otherwise there is no upper limit due to the  $j^{\text{th}}$  inactive constraint on  $\lambda$ . In case there is no upper limit on  $\lambda$ , we set  $\lambda = \bar{\lambda}$  for some fixed  $\bar{\lambda}$ . So we have  $\lambda_k = c_k \lambda$  where  $c_k \in (0, 1]$ . Let  $h_k' = \lambda h_k$  be the scaled descent direction then the next iteration can be written as

$$\tau_{k+1} = \tau_k + c_k h_k'. \quad (18)$$

Let  $\alpha \in (0, 1]$ ,  $\beta \in (0, 1)$  and let  $m$  be the smallest positive integer for which

$$J(\tau + \beta^m h'_k) - J(\tau_k) \leq \beta^m \alpha \langle \nabla J(\tau_k), h'_k \rangle. \quad (19)$$

Then  $c_k = \beta^m$  is our step size to be used for updating  $\tau_{k+1}$  via (18). The switch time optimization algorithm can be summarized as follows.

#### Gradient Projection Algorithm

Given a fixed  $\sigma_0$  and  $\tau_k$ , identify the set  $\mathcal{I}(\tau_k)$ , choose  $\bar{\lambda}$ , set  $k = 0$  and do the following.

- 1) Compute  $\nabla J(\tau_k)$  via (13) and  $h_k$  using (15).
- 2) If  $h_k = 0$ , compute  $\mu_k$  using (16). If all entries of  $\mu_k$  are non-negative **stop**. If not, eliminate the active constraint associated with the most negative  $\mu_k$  and go to step 1.
- 3) Compute  $\lambda$  using (17) and  $c_k$  using Armijo Rule (19).
- 4) Update the  $\tau$  using (18) and update the set of active constraints  $\mathcal{I}(\tau_k)$ . Set  $k = k + 1$  and go to step 1.

#### B. Mode Sequence Optimization

In this part we explain how to modify the existing mode sequence by inserting a new mode at each iteration of the algorithm. Due to dwell time constraints, the insertion of new mode is quite complex in comparison to mode insertion at a point as in [12]. The new mode inserted must exist for at least  $\delta$  seconds. If we insert a mode at a single point and do switch time optimization as in [12] by introducing two switching times initially  $\epsilon$  apart, they may not be apart by greater than or equal to  $\delta$  at the end of this optimization. However, still introducing the new mode forcibly for  $\delta$  seconds might result in a lower cost than the original.

To work around this problem, we take the route of starting feasible and staying feasible. By this, we mean that instead of doing mode insertion at a single point and then doing time optimization, we introduce a new mode for  $\delta$  seconds so that we start with the dwell time constraints satisfied and then the switch time optimization ensures that we remain feasible.

There is one more issue that needs to be addressed to start feasible. Since the mode inserted has to exist at least for  $\delta$  seconds, we cannot come very close to the existing switching times, otherwise their dwell time constraint will be violated. This problem is addressed by searching in the areas at and between switching intervals shown bold in the Fig. 1, assuming that the difference  $\tau_{k+1} - \tau_k \geq 3\delta$ . In this case, we check for insertion at points  $P_{k+1} = \{\tau_k, \tau_{k+1} - \delta\}$  and the interval  $I_{k+1} = [\tau_k + \delta, \tau_{k+1} - 2\delta]$ . For the case where  $2\delta \leq \tau_{k+1} - \tau_k < 3\delta$ , the only points to be checked for insertion are  $P_{k+1} = \{\tau_k, \tau_{k+1} - \delta\}$ . If  $\delta < \tau_{k+1} - \tau_k < 2\delta$ , we do not insert any mode. If  $\tau_{k+1} - \tau_k = \delta$ , then  $P_{k+1} = \{\tau_k\}$  is the only feasible insertion point. For mode schedule  $\xi$  of length  $n$ , the feasible insertion region is then  $U = \bigcup_{k=1}^n (P_k \cup I_k)$ .

The mode to be inserted and its location is given by argmin of the optimality function  $\theta_{\mathcal{X}}$  defined in (11) and can be determined as follows. For every  $k \in \mathcal{N}$  compute

$$t_k^* = \arg \min_{t \in U} J_{\Delta}(t, f_j), \quad (20)$$

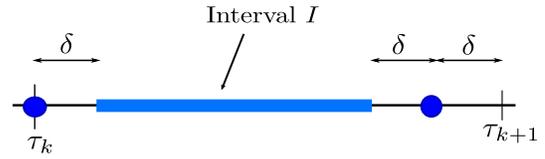


Fig. 1. Scanning area for mode insertion. The interval I and the points marked as circles form region of feasible mode insertion

where  $J_{\Delta}(t, f_j)$  is defined in (10). Then the insertion point and the mode to be inserted at a given iteration are

$$(t^*, f^*) = \arg \min_{k \in \mathcal{N}} J_{\Delta}(t_k^*, f_k), \quad (21)$$

so long as  $J_{\Delta}(t_k^*, f_k) < 0$ .

Solving (20) to find the optimal insertion point for every mode is computationally expensive since it seeks minimum over the entire set  $U \subset [0, T]$ . As opposed to the optimal insertion point in [12] which utilizes the one time computed state and co-state trajectory, this requires solving the state and co-state trajectory for each point to be checked for insertion. However, as the number of modes in a schedule increase, the region to be searched for mode insertion decreases. Accordingly the computational complexity for the mode insertion part of the algorithm decreases as the algorithm converges towards the stationary point.

#### Dwell Time Algorithm

Given A mode sequence  $\sigma$  having  $n$  modes and vector  $\tau$  satisfying the dwell time constraints and time optimized, do the following

- 1) Compute the optimality function  $\theta_{\mathcal{X}}$  as defined in (11).
- 2) If  $\theta_{\mathcal{X}} = 0$  then **stop**. Otherwise insert mode  $f^*$  in the interval  $[t^*, t^* + \delta]$ , by appending two new switching instants to vector  $\tau$  at times  $t^*$  and  $t^* + \delta$ . We get new mode sequence  $\sigma$  and switching time vector  $\tau$ .
- 3) For new  $\sigma$ , solve for optimal  $\tau$  using gradient projection algorithm of subsection (A) and go to step 1.

#### C. Convergence

To prove the convergence of our algorithm, we follow the approach used in [18] that relies on the property of sufficient descent. Let  $\{\xi_j\}$  be the sequence of points computed by the algorithm, then the property of sufficient descent for an algorithm is defined as follows.

*Definition 4:* An algorithm  $a : \mathcal{X} \rightarrow \mathcal{X}$  has the property of sufficient descent with respect to an optimality function  $\theta$  if for every  $\delta' > 0$ , there exists  $\eta > 0$  such that when  $\theta(\xi_j) < -\delta'$  then  $J(\xi_{j+1}) - J(\xi_j) < -\eta$ .

Since  $J(\xi) \geq 0$  for all  $\xi \in \mathcal{X}$ , the algorithm having the property of sufficient descent converges i.e for every infinite sequence  $\{\xi_j\}$  computed by algorithm,  $\lim_{j \rightarrow \infty} \theta(\xi_j) = 0$ . Otherwise  $\lim_{j \rightarrow \infty} J(\xi_j) = -\infty$  which contradicts the fact that  $J$  is bounded from below. The convergence is obvious if  $\{\xi_j\}$  is a finite sequence.

*Lemma 1:* The dwell time algorithm has the property of sufficient descent.

*Proof:* Consider our optimality function (11). Since  $\theta_{\mathcal{X}}(\xi_j) = J(\xi_{j+1}) - J(\xi_j)$  and  $\theta_{\mathcal{X}}(\xi_j) < -\delta'$  with  $\delta' > 0$

at all non-stationary points, we have  $J(\xi_{j+1}) - J(\xi_j) < -\eta$  with  $\eta = \delta'$ . Hence proved. ■

Again, since our optimization space is finite dimensional, we do not run into problems associated with infinite dimensional parameter spaces [2] to prove the sufficient descent property of our algorithm. We now state the main theorem about algorithm's convergence.

*Theorem 1:* The dwell time algorithm converges to stationary points of the optimality function  $\theta_{\mathcal{X}}$ .

*Proof:* The proof follows immediately from the property of sufficient descent proved for our algorithm and the discussion following the definition of sufficient descent regarding algorithm convergence. ■

## V. EXAMPLE

We consider the linear system in [12] that can switch between two modes described by the matrices

$$A_1 = \begin{pmatrix} -1 & 0 \\ 1 & 2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}.$$

We define a quadratic cost functional on the state trajectory.

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

The switched system is initialized to  $x_0 = (1, 0)'$  and optimized over time interval  $[0, 2]$ . The Armijo parameters are set to  $\alpha = 0.5$  and  $\beta = 0.5$ . The time step for simulation was set to  $dt = 1e^{-3}$  s. For intervals in the set  $U \subset [0, T]$  for feasible mode insertion, the time step used was  $ds = 1e^{-3}$  s. The mode sequence is initialized to  $\sigma_0 = \{f_1\}$  for which the cost is  $J = 40.75$ . The algorithm is terminated when  $|\theta_{\mathcal{X}}(\cdot)|$  falls below  $\epsilon = 0.1e^{-3}$ . The simulation is performed for dwell time constraint of 0.01s for both modes and results are shown in Fig. 2. To emphasize the variation of cost with each iteration, the initial cost  $J = 40.75$  which is relatively large is not shown for the cost trajectory and the graph starts with iteration 1. It takes about 8 iterations before the optimality conditions are met and the final cost is  $J = 2.39$ .

## VI. CONCLUSION

In this paper, we investigated the structure of our optimization space and made a case for defining the optimal solution as stationary points of optimality function. Subsequently we defined an optimality function and presented a conceptual algorithm. The algorithm convergence was proved and a numerical example was provided to highlight the algorithm.

The dwell time constraint makes the mode insertion a complex task and as a result this step is computationally expensive. A future work in this direction would be to come up with strategies to reduce the computational complexity associated with mode insertion.

## REFERENCES

- [1] F. Zhu and P. J. Antsaklis, "Optimal control of switched hybrid systems: a brief survey," *ISIS*, vol. 3, 2011.
- [2] H. Axelsson, Y. Wardi, M. Egerstedt, and E. Verriest, "Gradient descent approach to optimal mode scheduling in hybrid dynamical systems," *Journal of Optimization Theory and Applications*, vol. 136, no. 2, pp. 167–186, 2008.

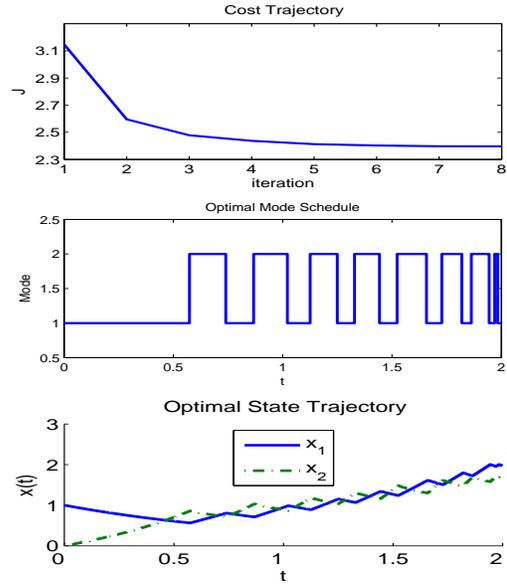


Fig. 2. The cost trajectory, optimal mode sequence and state trajectories are shown respectively for dwell time constraint of 0.01s

- [3] Y. Wardi, M. Egerstedt, and P. Twu, "A controlled-precision algorithm for mode-switching optimization," in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*. IEEE, 2012, pp. 713–718.
- [4] A. S. Morse, "Supervisory control of families of linear set-point controllers part i. exact matching," *Automatic Control, IEEE Transactions on*, vol. 41, no. 10, pp. 1413–1431, 1996.
- [5] J. P. Hespanha and A. S. Morse, "Stability of switched systems with average dwell-time," in *Decision and Control, 1999. Proceedings of the 38th IEEE Conference on*, vol. 3. IEEE, 1999, pp. 2655–2660.
- [6] M. S. Branicky, V. S. Borkar, and S. K. Mitter, "A unified framework for hybrid control: Model and optimal control theory," *Automatic Control, IEEE Transactions on*, vol. 43, no. 1, pp. 31–45, 1998.
- [7] B. Piccoli, "Hybrid systems and optimal control," in *Decision and Control, 1998. Proceedings of the 37th IEEE Conference on*, vol. 1. IEEE, 1998, pp. 13–18.
- [8] H. J. Sussmann, "A maximum principle for hybrid optimal control problems," in *Decision and Control, 1999. Proceedings of the 38th IEEE Conference on*, vol. 1. IEEE, 1999, pp. 425–430.
- [9] M. Shahid Shaikh and P. E. Caines, "On trajectory optimization for hybrid systems: theory and algorithms for fixed schedules," in *Decision and Control, 2002. Proceedings of the 41st IEEE Conference on*, vol. 2. IEEE, 2002, pp. 1997–1998.
- [10] M. S. Shaikh and P. E. Caines, "On the hybrid optimal control problem: theory and algorithms," *Automatic Control, IEEE Transactions on*, vol. 52, no. 9, pp. 1587–1603, 2007.
- [11] H. Gonzalez, R. Vasudevan, M. Kamgarpour, S. S. Sastry, R. Bajcsy, and C. Tomlin, "A numerical method for the optimal control of switched systems," in *Decision and Control (CDC), 2010 49th IEEE Conference on*. IEEE, 2010, pp. 7519–7526.
- [12] M. Egerstedt, Y. Wardi, and F. Delmotte, "Optimal control of switching times in switched dynamical systems," in *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, vol. 3. IEEE, 2003, pp. 2138–2143.
- [13] M. Egerstedt, Y. Wardi, and H. Axelsson, "Transition-time optimization for switched-mode dynamical systems," *Automatic Control, IEEE Transactions on*, vol. 51, no. 1, pp. 110–115, 2006.
- [14] F. Rothlauf, *Design of modern heuristics: principles and application*. Springer, 2011.
- [15] J. B. Rosen, "The gradient projection method for nonlinear programming. part i. linear constraints," *Journal of the Society for Industrial & Applied Mathematics*, vol. 8, no. 1, pp. 181–217, 1960.
- [16] D. P. Bertsekas, "Nonlinear programming," 1999.
- [17] L. Armijo, "Minimization of functions having lipschitz continuous first partial derivatives," *Pacific Journal of mathematics*, vol. 16, no. 1, pp. 1–3, 1966.
- [18] E. Polak, *Optimization: algorithms and consistent approximations*. Springer-Verlag New York, Inc., 1997.