Single-Integration Mode Scheduling for Linear Time-Varying Switched Systems

Anastasia Mavrommati, *Student Member, IEEE* and Todd D. Murphey, *Member, IEEE*

***Abstract*— This paper considers the problem of optimizing the schedule of modes in a linear time-varying switched system subject to a quadratic cost functional. The switched system optimization is formulated as an infinite-dimensional optimal control problem where a projection-based technique handles an integer constraint. In the proposed implementation, only a single set of differential equations needs to be solved off-line, with no additional simulation required during the optimization. Robustness to numerical errors is enhanced as these differential equations are as smooth as the system's vector fields, despite the optimization itself being non-smooth. An example demonstrates the optimization algorithm steps and verifies feasibility and convergence.**

I. INTRODUCTION

This paper considers the problem of scheduling the modes of a linear time-varying switched system so as to optimize a quadratic performance metric. By mode scheduling, we refer to the calculation of the optimal mode sequence and the corresponding switching times. The primary result of this paper is reformulating a projection-based mode scheduling algorithm so that no differential equation needs to be solved for during optimization. Therefore, we propose a modified iterative algorithm that makes full use of the system's linearity, hence requiring only a single integration of a set of smooth differential equations prior to optimization.

The problem is formulated as an infinite-dimensional optimal control problem where the variables to be optimized are a set of functions of time constrained to the integers. Several mode scheduling methods have been proposed to deal with this problem, including: mode injection methods [1], [2] which determine optimal switching times by computing when an injected mode will decrease the cost and embedding methods [3], [4], which relax, or embed, the integer constraint and find the optimal of the relaxed cost.

The iterative projection-based approach as introduced in Caldwell's work [5], [6], [7] form the basis for the work in this paper. For a projection-based method, the design variables live in an unconstrained space but the cost is computed on the projection of the design variables to the set of admissible switched system trajectories. In [5], an iterative optimization algorithm is synthesized that employs the Pontryagin Maximum Principle and a projection-based technique.

Anastasia Mavrommati and Todd D. Murphey are with the Department of Mechanical Engineering, Northwestern University, 2145 Sheridan Road Evanston, IL 60208, USA Email: stacymav@u.northwestern.edu ; t-murphey@northwestern.edu

This material is based upon work supported by the NIH Engineering for Neurologic Rehabilitation Center at the Rehabilitation Institute of Chicago and the National Science Foundation under award CMMI-1200321.

We reconsider this algorithm by taking advantage of the linearity of the dynamical system under concern. The specific case of linear switching control has been extensively investigated by others (see [8]). However, the approaches in [9], [10] solve for a differential equation at each step of the iterative algorithm. Previous attempts to avoid the on-line integration of differential equations are limited to switching time optimization problems where the mode sequence is fixed. More specifically, Xu and Antsaklis [10], [11] consider time-invariant switched systems and thus, employ the matrix exponential for simulating the state without solving for a differential equation. However, the linear relationship between state and co-state must still be simulated on-line at each iteration. On the other hand, Caldwell in [12], [13] considers linear time-varying systems and proves the existence of operators that, after they have been solved for off-line, can be used for the algebraic calculation of the state and co-state. In this paper, we extend this work to mode scheduling applications where the mode sequence is also optimized.

The objective of this paper is to present a projection-based mode scheduling algorithm that is not constrained by the system dynamics. In other words, only a single set of differential equations is solved off-line making it so that no additional simulation is required during the optimization routine. These differential equations are independent of the assumed mode sequence and switching times. Moreover, as in [12], no assumption about the time-variance of the modes is made. Under this formulation, the execution time is invariant to any choice of ODE solver, and only depends on the number of multiplications and inversions required for the calculation of the optimality condition. One of the strongest assets of the proposed algorithm is its high robustness to numerical errors arising from the fact that the off-line integrated differential equations are as smooth as each of the linear modes.

This paper is structured as follows: Section II reviews switched systems and their representations while stating our optimization problem. The single integration mode scheduling algorithm is proposed in Section III and an example is provided in Section IV. Section V discusses the algorithm complexity in terms of matrix multiplications per iteration and considers its robustness.

II. REVIEW

A. Switched Systems

Switched systems are a class of hybrid systems that evolve according to one of N modes $f_i : \mathbb{R}^n \to \mathbb{R}, i \in \{1, ..., N\}$ at any time over the finite time interval $[T_0, T_M]$, where

T_0 is the initial time and $T_M > 0$ is the final time. We consider two representations of the switched system, namely mode schedule and switching control. As a unique mapping exists between each representation, the two will be used interchangeably throughout the paper.

Definition 1: The *mode schedule* is defined as the pair $\{\Sigma, \mathcal{T}\}$ where $\Sigma = \{\sigma_1, ..., \sigma_M\}$ is the sequence of active modes $\sigma_i \in \{1, ..., N\}$ and $\mathcal{T} = \{T_1, ..., T_{M-1}\}$ is the set of the switching times $T_i \in [T_0, T_M]$. The total number of modes in the mode sequence is $M \in \mathbb{Z}^+$.

Definition 2: A *switching control* corresponds to a list of curves $u = [u_1, ..., u_N]^T$ composed of N piecewise constant functions of time, one for each different mode f_i. For each $t \in [T_0, T_M]$, $\sum_{i=1}^{N} u_i(t) = 1$, and for each $i \in \{1, ..., N\}, u_i(t) \in \{0, 1\}$. This dictates that the state evolves according to only one mode for all time. We represent the set of all admissible switching controls as Ω.

Throughout the paper, we will refer to the mode schedule corresponding to the switching control u as $\{\Sigma(u), \mathcal{T}(u)\}$.

B. Problem Statement

Our objective is the minimization of a quadratic cost function

$$J(x, u) = \int_{T_0}^{T_M} \frac{1}{2} x(\tau)^T Q(\tau) x(\tau) d\tau + \frac{1}{2} x(T_M)^T P_1 x(T_M) \quad (1)$$

subject to the pair (x, u) where x is the state and u the switching control. Here, Q and P_1 are the running and terminal cost respectively, and are both symmetric positive semi-definite. Note that this cost functional can also be adapted to include reference trajectory, in which case the objective would be to minimize the error between the state and the reference ([13]).

Any switched system (x, u) that optimizes the performance metric (1) is constrained by the state equations. For a system with n states $x = [x_1, ..., x_n]^T$ and N different modes, the state equations are given by

$$\dot{x}(t) = F(t, x(t), u(t)) := \sum_{i=1}^{N} u_i(t) f_i(x(t), t) \quad (2)$$

subject to the initial condition $x(T_0) = x_0$. For this paper, we restrict our focus to linear time-varying systems so that

$$F(t, x(t), u(t)) := \sum_{i=1}^{N} u_i(t) A_i(t) x(t). \quad (3)$$

Alternatively, we may express the system dynamics with respect to the current mode schedule as follows:

$$F(t, x(t), \Sigma, \mathcal{T}) := \overline{A}(t, \Sigma, \mathcal{T}) x(t) \quad (4)$$

where $\overline{A}(t, \Sigma, \mathcal{T}) = A_{\sigma_i}(t)$ for $T_{i-1} \leq t < T_i$.

From Definition 2 of an admissible switching control u, it follows that our optimization problem is subject to an integer constraint. Let $\mathcal{S}$ represent the set of all pairs of admissible state and switching control trajectories (x, u), i.e. all pairs that satisfy the constraint (2) and are consistent with Definition 2 so that $u \in \Omega$. In [6], authors propose a projection-based technique for handling these constraints set by $\mathcal{S}$. In particular, an equivalent problem is considered where the original design variables (α, μ) belong to an unconstrained set $(\mathcal{X}, \mathcal{U})$ and the cost J is evaluated on the projection of these variables to the set $\mathcal{S}$. Now, the problem is reformulated as

$$\arg\min_{(\alpha,\mu)} J(\mathcal{P}(\alpha, \mu)) \quad (5)$$

where $\mathcal{P}$ is a projection - i.e. $\mathcal{P}(\mathcal{P}(\alpha, \mu)) = (\mathcal{P}(\alpha, \mu))$ - that maps curves from the unconstrained set $(\mathcal{X}, \mathcal{U})$ to the set of admissible switched systems $\mathcal{S}$. As the cost is calculated on the admissible projected trajectories, this problem is equivalent to the original, as described in the beginning of this section.

The optimal mode scheduling algorithm developed in [5] utilizes the max-projection operator. The max-projection operator $\mathcal{P} : \mathcal{X} \times \mathcal{U} \to \mathcal{S}$ at time $t \in [T_0, T_M]$ is defined as

$$\mathcal{P}(\alpha, \mu) := \begin{cases} \dot{x}(t) = F(t, x(t), u(t)), & x(T_0) = x_0 \\ u(t) = \mathcal{Q}(\mu(t)) \end{cases} \quad (6)$$

where $\mathcal{Q}$ is a mapping from a list of N real-valued control trajectories, $\mu(t) = [\mu_1(t), ..., \mu_N(t)]^T \in \mathbb{R}^N$ to a list of N feasible switching controls, $u \in \Omega$. We may define $\mathcal{Q}$ as

$$\mathcal{Q}(\mu(t)) = \begin{bmatrix} \mathcal{Q}_1(\mu(t)) \\ \vdots \\ \mathcal{Q}_N(\mu(t)) \end{bmatrix} \text{ with } \mathcal{Q}_i(\mu(t)) := \prod_{j \neq i}^{N} 1(\mu_i(t) - \mu_j(t)) \quad (7)$$

where $1 : \mathbb{R} \to \{0, 1\}$ is the step function.

C. Mode Insertion Gradient

The *mode insertion gradient* appears in previous studies [1], [2], [14]. Here, it is defined as the list of functions $d = [d_1(t), ..., d_N(t)] \in \mathbb{R}^N$ that calculate the change to the cost J from inserting one of the N modes at some time t for an infinitesimal interval. Each function element of d is given by:

$$d_\alpha(t) := \rho(t)^T (f_\alpha(x(t), t) - f_{\sigma(t)}(x(t), t)) \quad (8)$$

where $x \in \mathbb{R}^n$ is the solution to the state equations (2) and $\rho \in \mathbb{R}^n$ is the co-state and solution to the adjoint equation

$$\dot{\rho}(t) = -\mathcal{D}F(t, x(t), u(t))^T \rho(t) - Q(t) x(t), \quad (9)$$

subject to $\rho(T_M) = P_1 x(T_M)$. In (8), $\sigma(t) : [T_0, T_M] \to \{1, ..., N\}$ is the function that returns the active mode at any time t.

It has been shown in [12], that when a quadratic cost is optimized subject to a linear time-varying switched system, a linear mapping between state x and co-state ρ exists. Thus, we may express the co-state as

$$\rho(t) = P(t) x(t) \quad (10)$$

where $P(t) \in \mathbb{R}^{n \times n}$ is the Riccati relation and is calculated by the following differential equation:

$$\dot{P}(t) = -\overline{A}(t, \Sigma, \mathcal{T})^T P(t) - P(t) \overline{A}(t, \Sigma, \mathcal{T}) - Q(t) \quad (11)$$

subject to $P(T_N) = P_1$. Note that this is the linear switched system analog to the Riccati Equation from the LQR problem in classical control theory. Using (3) and (10), the mode insertion gradient element can be written as

$$d_\alpha(t) := x(t)^T P(t)[A_\alpha(t) - A_{\sigma(t)}(t)]x(t). \tag{12}$$

III. Single Integration Optimal Mode Scheduling Algorithm

Caldwell in [5] considers the problem of optimizing an arbitrary cost functional $J(x,u)$ subject to the switching control $u(t)$ and switching system state $x(t)$, using projection-based techniques. Here, we further this work by restricting our focus to linear time-varying systems with quadratic performance metric. In particular, we seek to reformulate this problem so that no differential equations are solved during the iterative optimization routine.

Consider the switched system (x,u) and the corresponding optimization problem constrained by the system dynamics, as described in the previous section. The dynamic constraint dictates that a system simulation should be performed at each iteration as soon as the next switching control has been calculated. In particular, the calculation of the mode insertion gradient (8) involves the solution of the state and adjoint equations, (4) and (9), while the max-projection operator also includes the state equation. In this paper, we consider a modified version of the projection-based iterative optimization algorithm in [5] that depends on a set of differential equations that are only solved off-line once, hence alleviating the need for additional integration during the optimization process.

A similar approach to [12] will be followed. Building on the existence of a linear relationship between the state and co-state as described in the previous section, we will utilize operators to formulate algebraic expressions for the calculation of the state $x(t)$ and the Riccati relation $P(t)$ at any time $t \in [T_0, T_M]$. The operators should be available prior to optimization through an off-line solution to a differential equation. Moreover, they should be switching control invariant.

It should be noted that for an infinite-dimensional optimal control problem, the definition of the full state and co-state trajectories is required at each iteration, unlike the switching time optimization case [12], where only a finite number of state and co-state evaluations are involved. Remember that in the latter case, the mode sequence is constant and the problem is finite-dimensional. Therefore, in order for the proposed algorithm to be feasible, an explicit mapping from time t to x and P is defined at each iteration, depending on the current mode schedule $\{\Sigma, \mathcal{T}\}$. The proposed mapping only includes algebraic expressions. The exact number of multiplications executed in each iteration depends on the number of times the state and co-state must be evaluated. Later on, we will see that this comes down to the minimization of a single function, namely the mode insertion gradient.

For the rest of the paper, a variable with the superscript k implies that the variable depends directly on u^k i.e. the switching control at the k^{th} algorithm iteration.

A. *Defining $x(t)$*

The operators for defining $x(t)$ are the state-transition matrices (STM) of the N different modes. Let $\Phi^j(\cdot,\cdot) \in \mathbb{R}^{n\times n}$ denote the STM for the linear mode $j \in \{1,...,N\}$ with $A_j(t)$. The STM are the solutions to the N differential equations

$$\frac{d}{dt}\Phi^j(t,T_0) = A_j(t)\cdot \Phi^j(t,T_0), \quad j = 1,...,N \tag{13}$$

subject to the initial condition $\Phi^j(T_0,T_0) = I_n$.

The following two STM properties will be useful for defining the state $x(t)$ given a mode schedule $\{\Sigma, \mathcal{T}\}$. For an arbitrary STM, Φ, characterized by $A(t)$, we have:

1) $x(t) = \Phi(t,\tau)x(\tau)$
2) $\Phi(t1,t3) = \Phi(t1,t2)\Phi(t2,t3) = \Phi(t1,t2)\Phi(t3,t2)^{-1}$

Then, the state x at the i^{th} switching time is

$$x(T_i) = \overline{\Phi}(T_i,T_0)x_0 = [\prod_{j=i}^{1} \Phi^{\sigma_j}(T_j,T_{j-1})]x_0 \tag{14}$$

where $\overline{\Phi}(T_i,T_0)$ is the state-transition matrix corresponding to $\overline{A}(t,\Sigma,\mathcal{T})$ as defined above.

But for an infinite-dimensional optimal control algorithm, the value of the state should be available for all $t \in [T_0, T_M]$, as needed throughout the algorithm steps. Hence, the state evolution is defined as a piecewise function of time, each piece corresponding to a time interval between consecutive switching times $\{T_i, T_{i+1}\}$:

$$x(t) = \begin{cases} \Phi^{\sigma_1}(t,T_0)x(T_0), & T_0 \le t < T_1 \\ \Phi^{\sigma_2}(t,T_1)\Phi^{\sigma_1}(T_1,T_0)x(T_0), & T_1 \le t < T_2 \\ \vdots & \vdots \\ \Phi^{\sigma_M}(t,T_{M-1})[\prod_{j=M-1}^{1} \Phi^{\sigma_j}(T_j,T_{j-1})]x(T_0) & T_{M-1} \le t \le T_M \end{cases} \tag{15}$$

A more compact representation of the state is provided here, by employing unit step functions and (14) . In particular, the value of the state $x(t)$ at all $t \in [T_0, T_M]$ depends on

---

Algorithm 1

Off-line:

- Solve for the STM $\Phi^j(t,T_0)$ and ATM $\Psi^j(t,T_M)$ $\forall j \in \{1,...,N\}$ and $t \in [T_0,T_M]$ and store in memory.
- Choose initial $u^0 \to \{\Sigma(u^0), \mathcal{T}(u^0)\}$.
- Set $x(T_0) = x_0$ and $P(T_M) = P_1$.

On-line iterative process:
Set $k = 0$, $u^k = u^0$.

1) Define $x^k(t) := \chi(t, \Sigma(u^k), \mathcal{T}(u^k))$ as in Eq. (17).
2) Define $P^k(t) := \varrho(t, \Sigma(u^k), \mathcal{T}(u^k))$ as in Eq. (24).
3) Define the descent direction $-d^k(t)$ as in Eq. (27).
4) Calculate step size γ^k by backtracking.
5) Update: $u^{k+1}(t) = Q(u^k(t) - \gamma^k d^k(t))$.
6) If u^{k+1} satisfies a terminating condition, then exit, else, increment k and repeat from step 1.

---

the current mode schedule $\{\Sigma, \mathcal{T}\}$ and is given by

$$x(t) := \chi(t, \Sigma, \mathcal{T}) \tag{16}$$

where

$$\chi(t, \Sigma, \mathcal{T}) = \sum_{i=1}^{M} \Big\{ [1(t - T_{i-1}) - 1(t - T_i)] \Phi^{\sigma_i}(t, T_{i-1}) x(T_{i-1}) \Big\} \tag{17}$$

where, from STM property 2,

$$\Phi^{\sigma_i}(t, T_{i-1}) = \Phi^{\sigma_i}(t, T_0) \Phi^{\sigma_i}(T_{i-1}, T_0)^{-1}. \tag{18}$$

Prior to the iterative optimization, the operators (STM) $\Phi^j(t, T_0)$ are solved off-line for $t \in [T_0, T_M]$ and for all different modes $j = 1, ..., N$ and stored in memory. Thus, given a mode schedule, the calculation of state $x(t)$ via (17) requires no additional integrations.

B. *Defining $P(t)$*

Caldwell has proved in [12] that an analogous operator to the STM exists for the definition of the Riccati relation $P(t)$ appearing in (10). As in [12], we will refer to the operator as the adjoint-transition matrix (ATM) and use $\Psi^j(\cdot\,,\cdot) \in \mathbb{R}^{n \times n}$ to denote the ATM corresponding to each mode $j \in \{1, ..., N\}$. The ATM are defined to be the solutions to the following N differential equations:

$$\frac{d}{dt} \Psi^j(t, T_M) = -A_j(t)^T \Psi^j(t, T_M) - \Psi^j(t, T_M) A_j(t) - Q(t) \tag{19}$$

subject to the initial condition $\Psi^j(T_M, T_M) = 0_{n \times n}$. Note that the state and the adjoint are solved in opposing directions.

The following two ATM properties will be useful for defining $P(t)$ given a mode schedule $\{\Sigma, \mathcal{T}\}$. For an arbitrary ATM, Ψ, characterized by $A(t)$ and associated STM Φ, and cost function defined by $Q(t)$, we have:

1) $P(t) = \Psi(t, \tau) \circ P(\tau) = \Psi(t, \tau) + \Phi(\tau, t)^T P(\tau) \Phi(\tau, t)$
2) $\Psi(t1, t3) = \Psi(t1, t2) \circ \Psi(t2, t3) = \Psi(t1, t2) + \Phi(t2, t1)^T \Psi(t2, t3) \Phi(t2, t1)$

Each of the above properties has been proved in [12]. Then, the $P(t)$ at the i^{th} switching time is

$$\begin{aligned} P(T_i) &= \overline{\Psi}(T_i, T_M) \circ P(T_M) \\ &= \overline{\Psi}(T_i, T_M) + \overline{\Phi}(T_M, T_i)^T P(T_M) \overline{\Phi}(T_M, T_i) \end{aligned} \tag{20}$$

where $\overline{\Psi}(T_i, T_M)$ is the adjoint-transition matrix corresponding to $\overline{A}(t, \Sigma, \mathcal{T})$ as defined above. From ATM property 2, this is equal to

$$\begin{aligned} \overline{\Psi}(T_i, T_M) &= \Psi^{\sigma_{i+1}}(T_i, T_{i+1}) \circ \cdots \circ \Psi^{\sigma_M}(T_{M-1}, T_M) \\ &= \sum_{k=i+1}^{M} \overline{\Phi}(T_{k-1}, T_i)^T \Psi^{\sigma_k}(T_{k-1}, T_k) \overline{\Phi}(T_{k-1}, T_i) \end{aligned} \tag{21}$$

As in the previous case, we aim to derive an expression for the evaluation of $P(t)$ at random time instances, as needed. Again, we will represent the Riccati relation as a piecewise function of time:

$$P(t) = \begin{cases} \Psi^{\sigma_M}(t, T_M) \circ P(T_M), & T_{M-1} \le t < T_M \\ \Psi^{\sigma_{M-1}}(t, T_{M-1}) \circ P(T_{M-1}), & T_{M-2} \le t < T_{M-1} \\ \quad\vdots & \quad\vdots \\ \Psi^{\sigma_1}(t, T_1) \circ P(T_1), & T_0 \le t < T_1 \end{cases} \tag{22}$$

In particular, the value of the Riccati relation $P(t)$ at all $t \in [T_0, T_M]$ depends on the current mode schedule $\{\Sigma, \mathcal{T}\}$ and is given in a more compact form by the following

$$P(t) := \varrho(t, \Sigma, \mathcal{T}) \tag{23}$$

with

$$\varrho(t, \Sigma, \mathcal{T}) = \sum_{i=1}^{M} \Big\{ [1(t - T_{i-1}) - 1(t - T_i)][\Psi^{\sigma_i}(t, T_i) \circ P(T_i)] \Big\} \tag{24}$$

where, from ATM property 2,

$$\Psi^{\sigma_i}(t, T_i) = \Psi^{\sigma_i}(t, T_M) + \Phi^{\sigma_i}(T_i, t)^T \Psi^{\sigma_i}(T_i, T_M) \Phi^{\sigma_i}(T_i, t). \tag{25}$$

Combining ATM property 1 with (24) and (25), we end up with the expression

$$\begin{aligned} \varrho(t, \Sigma, \mathcal{T}) = \sum_{i=1}^{M} \Big\{ &[1(t - T_{i-1}) - 1(t - T_i)] \cdot \\ &[\Psi^{\sigma_i}(t, T_M) + \Phi^{\sigma_i}(T_i, t)^T [P(T_i) - \Psi^{\sigma_i}(T_i, T_M)] \Phi^{\sigma_i}(T_i, t)] \Big\}. \end{aligned} \tag{26}$$

Prior to the iterative optimization, the operators (ATM) $\Psi^j(t, T_M)$ are solved off-line for all $t \in [T_0, T_M]$ and for all different modes $j = 1, ..., N$ and stored in memory. Thus, given a mode schedule, the calculation of $P(t)$ via (26) requires no additional integrations.

C. *Defining the descent direction*

An iterative optimization method computes a new estimate of the optimum by taking a step in a search direction from the current estimate of the optimum so that a sufficient decrease in cost is achieved. For projection-based infinite-dimensional optimization, the cost does not have a natural gradient. However, the mode insertion gradient $d(t)$ defined in the previous section, has a similar role in the mode scheduling optimization as the gradient does for finite-dimensional optimization. It was proved in [5] that $-d^k(t)$ is a descent direction.

After the definition for the state and co-state, deriving an equivalent expression for the mode insertion gradient is straightforward from (17),(23) and (8). Thus, an element of $d^k(t)$ is defined as

$$\begin{aligned} d^k_\alpha(t) := \chi(t, \Sigma(u_k), \mathcal{T}(u_k))^T \varrho(t, \Sigma(u_k), \mathcal{T}(u_k)) \\ [A_\alpha(t) - A_{\sigma^k(t)}(t)] \chi(t, \Sigma(u_k), \mathcal{T}(u_k)) \end{aligned} \tag{27}$$

where $\alpha = \{1, ..., N\}$.

D. *Update rule*

A new estimate of the optimal switching control u^{k+1} is obtained by varying from the current optimum u^k in the descent direction and projecting the result to the set of admissible switching control trajectories. For this purpose, the max-projection operator (6) is employed. Here, the projection operator is adjusted to reflect the fact that the problem is no longer constrained by the system dynamics. We refer to the new projection operator as $\mathcal{P}_n$ and define it as follows:

$$\mathcal{P}_n(\alpha,\mu) := \begin{cases} x(t) = \chi(t, \Sigma(u), \mathcal{T}(u)) \\ u(t) = Q(\mu(t)) \end{cases} \quad (28)$$

where Q is again given by (7). Consequently, the update rule is $u^{k+1}(t) = Q(u^k(t) - \gamma^k d^k(t))$. For choosing a sufficient step size γ^k, we may utilize a projection-based backtracking process as described in [7].

As the focus of this paper is very specific i.e. the reformulation of the optimization problem to a problem unconstrained by the system dynamics, the reader is referred to [5], [6] for a more detailed description of these algorithm steps, along with the associated proofs for convergence.

E. Calculating the optimality condition

As no natural gradient exists for this infinite-dimensional optimal control problem, the optimality function $\theta^k \in \mathbb{R}$ has been defined in [5] as

$$\theta^k := d^k_{a_0}(t_0) \quad (29)$$

where

$$(a_0, t_0) = \arg \min_{a \in \{1,\dots,N\}, t \in [T_0, T_M]} d_a(t). \quad (30)$$

This choice is fairly intuitive, if we consider that as the minimum value of the mode insertion gradient goes to zero, the possibility of cost reduction with a mode insertion drops. This rationale follows directly from the definition of the mode insertion gradient. As far as convergence is concerned, the limit of the sequence of optimality functions is proved to go to zero in [5]. This allows us to utilize θ^k also as a terminating condition for the iterative algorithm.

Building on the definition of the optimality condition in (29), the amount of memory calls and matrix algebra is tied to the selection of a specialized global optimization algorithm for expensive-to-evaluate functions (see [15]).

IV. Example

We demonstrate an example to showcase the algorithm feasibility and convergence. Consider a linear time-invariant switched system of the form in (4) with $N = 2$ possible modes

$$A_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix} \quad (31)$$

and initial configuration $x_0 = [0.5, 0]^T$ at time $T_0 = 0$. We wish to find the switching controls that minimize the quadratic cost functional (1) with $Q = I_n$ and $P_1 = 0_n$ over the time interval $[0, 6.5]$.

Before initiating the on-line iterative process, we calculate the operators $\Phi^j(t, 0)$ and $\Psi^j(t, 6.5)$ for all $t \in [0, 6.5]$ and $j = \{1, 2\}$ by numerically solving (13) and (19) respectively. One way to do that is to store in memory discrete data points of the trajectories, so that any single value of $\Phi^j(t, 0)$ and $\Psi^j(t, 6.5)$ at time t can be given by a polynomial interpolation of the data points surrounding this time instance.

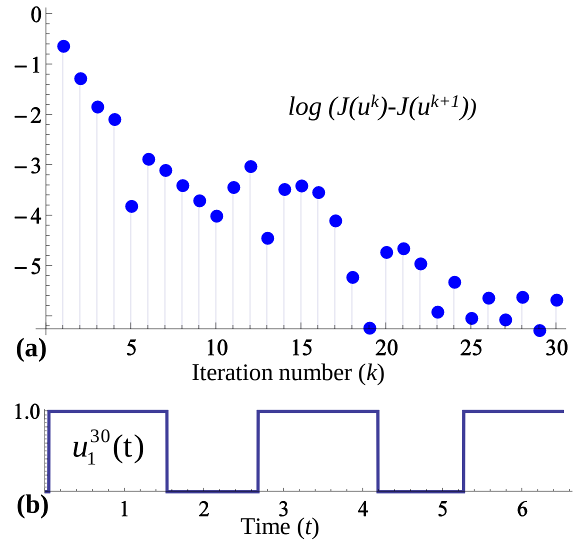

Fig. 1. (a) Plot of $\log(J(u^k) - J(u^{k+1}))$ for 30 iterations. The difference between consecutive costs decreases with iteration. (b) The optimal switching control after 30 iterations. Note that this is the scheduling of mode 1, u_1^{30}, and u_2^{30} is its supplementary i.e. where u_1 is 1, u_2 is 0 and vice versa.

Following, we execute 30 iterations of Algorithm 1 starting with initial switching control $u^0 = [1, 0]^T(t)$ or equivalently $\Sigma(u^0) = \{1\}$ and $\mathcal{T}(u^0) = \{\}$. The procedure for the first iteration is described next. The state $x^0(t)$ and Riccati relation $P^0(t)$ are defined as in (17) and (26) respectively, with respect to the current mode schedule. Now, the negative mode insertion gradient $-d^0(t)$ is given by (27). The next optimum u^1 is computed by taking a sufficient step in the descent direction and utilizing the max-projection operator so as to project the unconstrained trajectories to the feasible switching control trajectories constrained to the integers.

After the 30th iteration, the cost is reduced from 0.812499 to 0.483374 and the optimality condition θ^k trends toward 0, i.e. $\theta^{30} = -0.00754861$. The reduction of the cost at each iteration is illustrated in Fig. 1(a). The optimal mode schedule is $\Sigma(u^{30}) = \{2, 1, 2, 1, 2, 1\}$ and $\mathcal{T}(u^{30}) = \{0.0447234, 1.53061, 2.67752, 4.18317, 5.25856\}$. The switching control after 30 iterations $u^{30}(t)$ was evaluated at all $t \in [0, 6.5]$ and the first element corresponding to mode 1 is plotted against time in Fig. 1(b).

The executional efficiency of the algorithm is not addressed in this example, since a conventional function minimizer (i.e. built-in Mathematica function) was used for the calculation of the optimality condition.

V. Complexity and Robustness

It is generally the case that the complexity of an optimal control algorithm is discussed in terms of the forward and backward system simulations involved in each iteration. However, in the previous section, we showed that all the state and co-state information we need, is encoded in the STM, $\Phi^j(t, T_0)$, and ATM, $\Psi^j(t, T_M)$, $\forall j \in \{1, \dots, N\}$ which are solved for and saved in memory for all $t \in [T_0, T_M]$ prior to the optimization routine. Therefore, the calculation of $x^k(t)$ and $P^k(t)$ and consequently the optimality condition θ^k relies simply on memory calls and matrix algebra. No additional differential equations need to be solved for during optimization.

The algorithm complexity can be discussed instead in terms of the number of matrix multiplications involved in each iteration. Recall that at each iteration, the state is defined as in (17) and the Riccati relation as in (26), but the total number of function evaluations depends on the selection of the global optimization algorithm for minimizing the mode insertion gradient. Taking this into consideration, we will instead look at the algebraic calculations required for the evaluation of the state and co-state at a single time instance t.

First, for executional efficiency, one may calculate all the state and co-state values at the switching times, $x(T_i)$ and $P(T_i)$, given the current mode schedule $(\Sigma(u^k), \mathcal{T}(u^k))$ at the beginning of each iteration. To compute the state, begin with $x(T_0) = x_0$ and then recursively calculate

$$x(T_i) = \Phi^{\sigma_i}(T_i, T_{i-1})x(T_{i-1})\forall i \in \{1, ..., M-1\}. \quad (32)$$

Using STM property 2 and following a similar approach as in the derivation of (17) , this computation comes down to $2(M-1)$ multiplications, assuming that all $\Phi^j(t, T_0)^{-1}$ for all $j \in \{1, ..., N\}$ have also been stored in memory. Similarly for the Riccati relation, begin with $P(T_M) = P_1$ and then recursively calculate

$$\begin{aligned} P(T_i) =& \Psi^{\sigma_{i+1}}(T_i, T_M) + \Phi^{\sigma_{i+1}}(T_{i+1}, T_i)^T[P(T_{i+1}) \\ &- \Psi^{\sigma_{i+1}}(T_{i+1}, T_M)]\Phi^{\sigma_{i+1}}(T_{i+1}, T_i) \end{aligned} \quad (33)$$

for all $i \in \{1, ..., M-1\}$. Note that the derivation of the above expression is identical to the derivation of (26). Knowing that all $\Phi^{\sigma_{i+1}}(T_{i+1}, T_i)$ have already been calculated in (32), another $2(M-1)$ multiplications are required for the calculation of $P(t)$. To summarize, the standard computational cost of the algorithm comes down to a total of $4(M-1)$ multiplications per iteration.

Now, to evaluate equation (17) and (26) at any random time t during the optimization process, we only need four additional multiplications, two for the state and two for the Riccati relation. Therefore, to evaluate the mode insertion gradient at any random time, 7 multiplications are required in total, including the algebra involved in (8).

On the subject of the algorithm robustness to numerical errors, the determinative factor here is the numerical solver used to solve for the operators $\Phi^j(t, T_0)$ and $\Psi^j(t, T_M)$. Since the numerical integration of these operators is as smooth as each of the modes, the proposed implementation has additional numerical robustness compared to the original algorithm in [5] where simulation of non-smooth switched dynamical systems is involved.

VI. Conclusions and Future Work

In this paper, we proposed an approach for scheduling the modes of linear switched autonomous systems to optimize a quadratic cost functional. In particular, we revised a projection-based optimal control algorithm aiming to take full advantage of the system's linearity. With the proposed implementation, no differential equations are solved during the optimization routine. Only a single set of differential equations, which are as smooth as the system's vector fields, need to be simulated off-line, independent of the assumed mode sequence and switching times. Now, the computational cost does not depend on the choice of the ODE solver but on the number of matrix multiplications per iteration.

Future work will focus on assessing the computational efficiency of the algorithm, by employing global optimization techniques.

References

[1] 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.

[2] H. Gonzalez, R. Vasudevan, M. Kamgarpour, S. S. Sastry, R. Bajcsy, and C. J. Tomlin, "A descent algorithm for the optimal control of constrained nonlinear switched dynamical systems," in *Proceedings of the 13th ACM international conference on Hybrid systems: computation and control*. ACM, 2010, pp. 51–60.

[3] S. C. Bengea and R. A. DeCarlo, "Optimal control of switching systems," *automatica*, vol. 41, no. 1, pp. 11–27, 2005.

[4] S. Wei, K. Uthaichana, M. Žefran, R. A. DeCarlo, and S. Bengea, "Applications of numerical optimal control to nonlinear hybrid systems," *Nonlinear Analysis: Hybrid Systems*, vol. 1, no. 2, pp. 264–279, 2007.

[5] T. M. Caldwell and T. D. Murphey, "Projection-based optimal mode scheduling," in *IEEE Conference on Decision and Control*, 2013, (Accepted).

[6] T. M. Caldwell and T. D. Murphey, "Projection-based switched system optimization," in *American Control Conference (ACC), 2012*. IEEE, 2012, pp. 4552–4557.

[7] T. M. Caldwell and T. D. Murphey, "Projection-based switched system optimization: Absolute continuity of the line search," in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*. IEEE, 2012, pp. 699–706.

[8] Z. Sun and S. S. Ge, "Analysis and synthesis of switched linear control systems," *Automatica*, vol. 41, no. 2, pp. 181–195, 2005.

[9] A. Giua, C. Seatzu, and C. Van Der Mee, "Optimal control of switched autonomous linear systems," in *Decision and Control, 2001. Proceedings of the 40th IEEE Conference on*, vol. 3. IEEE, 2001, pp. 2472–2477.

[10] X. Xu and P. J. Antsaklis, "An approach for solving general switched linear quadratic optimal control problems," in *Decision and Control, 2001. Proceedings of the 40th IEEE Conference on*, vol. 3. IEEE, 2001, pp. 2478–2483.

[11] X. Xu and P. J. Antsaklis, "Optimal control of switched autonomous systems," in *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, vol. 4. IEEE, 2002, pp. 4401–4406.

[12] T. M. Caldwell and T. D. Murphey, "Single integration optimization of linear time-varying switched systems," in *American Control Conference (ACC), 2011*. IEEE, 2011, pp. 2024–2030.

[13] T. M. Caldwell and T. D. Murphey, "Single integration optimization of linear time-varying switched systems," *IEEE transactions on automatic control*, vol. 57, no. 6, pp. 1592–1597, 2012.

[14] Y. Wardi and M. Egerstedt, "Algorithm for optimal mode scheduling in switched systems," in *American Control Conference (ACC), 2012*. IEEE, 2012, pp. 4546–4551.

[15] J. Villemonteix, E. Vazquez, and E. Walter, "An informational approach to the global optimization of expensive-to-evaluate functions," *Journal of Global Optimization*, vol. 44, no. 4, pp. 509–534, 2009.