An Algorithm for Discrete State Sequence and Trajectory Optimization for Hybrid Systems with Partitioned State Space

Benjamin Passenberg, Marion Sobotka, Olaf Stursberg, Martin Buss, and Peter E. Caines

***Abstract*— An algorithm for hybrid optimal control is proposed that varies the discrete state sequence based on gradient information during the search for an optimal trajectory. The algorithm is developed for hybrid systems with partitioned state space. It uses a version of the hybrid minimum principle that allows optimal trajectories to pass through intersections of switching manifolds, which enables the algorithm to vary the sequence. Consequently, the combinatorial complexity of former algorithms can be avoided, since not each possible sequence has to be investigated separately anymore. The convergence of the algorithm is proven and a numerical example demonstrates the efficiency of the algorithm.**

I. INTRODUCTION

Hybrid systems show a close interconnection of discrete and continuous dynamics [1]. The continuous dynamics is usually described by differential or difference equations including continuous variables. Logic variables are assigned to different discrete states of a hybrid system, where discrete states differ in their continuous dynamics. Furthermore, the discrete structure defines how a discrete state can be changed. The question how to control such systems optimally attracted remarkable attention in recent years. Answering the question turned up to be challenging since in general methods for optimizing the discrete or continuous part separately are not suitable for the optimization of their interconnection.

Different concepts emerged to tackle the problem like mixed-integer programming [2], value function approaches [3], and the hybrid maximum principle [4]. The task of finding an optimal discrete state sequence often leads to a combinatorial complexity in the number of discrete state transitions. Several approaches have been proposed on efficiently handling or avoiding the combinatorial complexity. For general hybrid systems, relaxations and a branch-and-bound method are applied in [5] to reduce the amount of discrete state sequences to be investigated.

Especially for the class of switched systems with only controlled switching, efficient methods exist that avoid exponential complexity. By the construction and pre-computation of optimality zones, it is possible to find the optimal discrete state schedule in a single run of the algorithm with an overall linear complexity in the amount of switchings [6]. In [7] and [8], the optimal sequence is found by relaxation techniques. In [9], needle variations of the discrete state are inserted in a given discrete state sequence and gradient methods are applied to find the optimal switching times. For switching systems with autonomous switching, approaches with low computational complexity are widely missing. For example in [10–12], the algorithms for autonomous switching have exponential complexity as every possible discrete state sequence is investigated separately. An approach for reducing the complexity in this system class is given in [13], where numerical techniques for solving continuous optimization problems are combined with symbolic techniques for solving constraint satisfaction problems of logic variables. Based on a value function approach, a method suitable for low dimensional systems is developed that shows a very low complexity increase with the number of controlled and autonomous switchings [14].

This motivates the introduction of an efficient algorithm developed for hybrid systems with partitioned state space and only autonomous switching is considered. The algorithm uses necessary optimality conditions provided by a novel version of the hybrid minimum principle (HMP) [15]. The introduced HMP can deal with non-smooth and intersecting switching manifolds, such that gradients of the cost function can be evaluated at these points. This allows us to develop an algorithm that varies the discrete state sequence during the optimization process, which reduces the computational complexity to find a (locally) optimal state trajectory and discrete state sequence significantly. The algorithm is initialized with a discrete state sequence and corresponding switching states and times. Next, an underlying optimization routine determines the optimal solution from switching point to switching point. By evaluating optimality conditions at the switching points, the latter can be shifted and consequently the discrete state sequence can be varied with a gradient descent method until an optimal solution is obtained.

Sec. II introduces hybrid systems and the non-smooth version of the HMP. In Sec. III the algorithm is presented, while its convergence is proved in Sec. IV. Sec. V discusses the effectiveness of the algorithm for a numerical example, and Sec. VI concludes the paper.

II. PROBLEM FORMULATION

Hybrid systems with partitioned state space and autonomous switching are introduced as follows.

B. Passenberg, M. Sobotka, and M. Buss are with the Institute of Automatic Control, Department of Electrical Engineering and Information Technology, Technische Universität München, München, Germany. {passenberg,marion.sobotka,mb}@tum.de

P.E. Caines is with the Department of Electrical and Computer Engineering and the Center for Intelligent Machines, McGill University, Montréal, Canada. peterc@cim.mcgill.ca

O. Stursberg is with the Institute of Control and System Theory, Department of Electrical Engineering and Computer Science, University of Kassel, Kassel, Germany. stursberg@uni-kassel.de

Definition 1: A hybrid system is a 7-tuple

$$\mathbb{H} := \{\mathcal{Q}, \mathcal{X}, \Gamma, U, \mathcal{F}, \mathcal{M}, \mathcal{I}\}. \tag{1}$$

□

Assumption 1:

- $\mathcal{Q} = \{1, 2, ..., N_q\}$: set of discrete states q with N_q being the number of discrete states.
- $\mathcal{X} = \bigcup_{1\leq q\leq N_q} \mathcal{X}_q$: union of all state spaces $\mathcal{X}_q = \partial\mathcal{X}_q \cup \hat{\mathcal{X}}_q \subset \mathbb{R}^{n_x}$ assigned to every discrete state q, where $\mathcal{X} = \mathbb{R}^{n_x}$ and $\dim(x) = n_x$. The interiors of all state spaces are disjoint: $\hat{\mathcal{X}}_j \cap \hat{\mathcal{X}}_k = \emptyset$, $j, k \in \mathcal{Q}$, $j \neq k$, while the boundaries of neighboring state spaces have a non-empty intersection $\partial\mathcal{X}_j \cap \partial\mathcal{X}_k \neq \emptyset$.
- $U = \{U_q\}_{q\in\mathcal{Q}}$: collection of compact sets $U_q \subset \mathbb{R}^{n_u}$ of admissible continuous control values with $\dim(u) = n_u$. The control functions $u_q(t)$, $q \in \mathcal{Q}$, $t \in [t_0, t_e]$, $t_e < \infty$ lie in the set of admissible control functions $\mathcal{U}_q$ denoting the set of all measurable and bounded functions on the interval $[t_0, t_e]$ taking values in the compact sets U_q. The collection of control functions is denoted by $\mathcal{U} = \{\mathcal{U}_q\}_{q\in\mathcal{Q}}$. $\chi : [t_0, t_e] \to \mathcal{X}$ denotes a state trajectory of the hybrid system with initial and final time t_0 and t_e.
- $\mathcal{F} = \{f_q\}_{q\in\mathcal{Q}}$: collection of vector fields $f_q : \mathbb{R}^{n_x} \times U_q \to \mathbb{R}^{n_x}$ defined for each $q \in \mathcal{Q}$. The vector fields are at least once continuously differentiable, i.e. $f_q \in C^1(\mathbb{R}^{n_x} \times U_q, \mathbb{R}^{n_x})$. They fulfill a uniform Lipschitz condition, i.e. $\exists L < \infty$ such that $||f_q(x_1, u_q) - f_q(x_2, u_q)|| \leq L||x_1 - x_2||$, $x_1, x_2 \in \mathbb{R}^{n_x}, u_q \in U_q$, and $||.|| := ||.||_2$.
- $\mathcal{M} = \{m_{j,k}\}_{j,k\in\mathcal{Q},j\neq k}$: collection of switching manifolds, where $m_{j,k} \in C^1(\mathcal{X}_j, \mathbb{R})$. An autonomous transition from discrete state j to k occurs at time t_i for $x(t_i)$ on the manifold $m_{j,k}(x(t_i)) = 0$ with codimension 1. The set of points on a switching manifold is denoted by $M_{j,k} = \{x|m_{j,k}(x) = 0\}$. In $\mathcal{X}$, the boundary between neighboring partitions is defined as the switching manifold $m_{j,k}(x) = 0 \ \forall x \in \partial\mathcal{X}_j \cap \partial\mathcal{X}_k \neq \emptyset$. This implies that $m_{j,k} = 0$, $j, k \in \mathcal{Q}, j \neq k$ may have a boundary $\partial m_{j,k}$ and that neighboring switching manifolds intersect at those boundaries, i.e. $\partial m_{j,k} \cap \partial m_{j,l} \neq \emptyset$ for some pairwise different $j, k, l \in \mathcal{Q}$. The intersection $\partial m_{j,k} \cap \partial m_{j,l}$ forms a piecewise C^1 submanifold of codimension 2, on which there exists a deterministic decision if it is switched to k or l. Switching manifolds are oriented: $m_{j,k}(x) < 0 \ \forall x \in \hat{\mathcal{X}}_j$.
- $\Gamma : \mathcal{Q} \times \mathcal{X} \to \mathcal{Q}$: discrete transition map.

Assumption 2: The switching times t_i are well-defined, and $\forall x(t_i) \in M_{j,k}$ the vector fields $f_j(x, u_j)$ and $f_k(x, u_k)$ are non-vanishing at and transversal to $m_{j,k} = 0$ with $j, k \in \mathcal{Q}$ for the applied $u_j \in U_j$ and $u_k \in U_k$. This also implies that no accumulation points of switching (Zeno points) or sliding motions occur. □

Assumption 3: At time t_0 with given initial conditions $(q(t_0), x(t_0)) = (q_0, x_0) \in \mathcal{Q} \times \mathcal{X}$, for all switching manifolds it is assumed that $m_{q_0,j}(x_0) \neq 0 \ \forall j \in \mathcal{Q}$. □

Note that these assumptions are important for the existence of a unique execution of the hybrid system and for the existence of an optimal control.[1]

Assumption 4: Let a trajectory $\chi(t)$ go from discrete state j directly to $j+1$ through $x_s \in \partial m_{j,j+1} \cap \partial m_{j,k} \cap \partial m_{k,j+1}$, $j, j+1, k \in \mathcal{Q}$. Then it is assumed that there exists $u_k \in U_k$ such that a neighboring trajectory $\chi_k(t)$ passing the additional state k in-between j and $j+1$ can be found such that $\chi_k(t)$ is arbitrarily close to $\chi(t)$. □

Definition 2: An optimal sequence $q^*(t)$ and optimal control and state trajectories $u^*(t)$ and $\chi^*(t)$ minimize the cost functional

$$J = g(x_{q_e}(t_e)) + \sum_{i=0}^{N} \int_{t_i}^{t_{i+1}} \phi_{q_i}(x_{q_i}(t), u_{q_i}(t))\, dt \tag{2}$$

while satisfying initial conditions x_0 and q_0, final conditions $h(t_e, x(t_e), q(t_e)) = 0$, the discrete mapping Γ and the continuous dynamics

$$\dot{x}_{q_i} = f_{q_i}(x_{q_i}(t), u_{q_i}(t)) \tag{3}$$

for a.e. $t \in [t_i, t_{i+1})$ and for every $i \in \{0, ..., N\}$, where the number of switches N is an outcome of minimizing (2). The terminal cost function $g \in C^1(\mathbb{R}^{n_x}, \mathbb{R})$ and the running cost functions $\phi_q \in C^1(\mathbb{R}^{n_x} \times \mathbb{R}^{n_u}, \mathbb{R})$, $q \in \mathcal{Q}$ are at least once continuously differentiable. □

Definition 3: A switching point z_i is the vector consisting of the switching state-time pair $z_i = (x_i^T \ \ t_i)^T$ with $m_{q_{i-1},q_i}(x_i) = 0$. Here, i denotes the i-th switching. All switching points are summarized in the vector

$$z = (z_0 \ \ z_1 \ ... \ z_N \ z_{N+1})^T, \tag{4}$$

where z_0 is the initial state-time point $(x_0^T \ t_0)^T$. Usually, the terminal point $z_{N+1} = (x_e^T \ t_e)^T$ is only partially specified, e.g. x_e or t_e might be free. All switching times must satisfy $t_0 < t_1 < ... < t_N < t_e$. □

Definition 4: Suppose $\chi^*(t)$ to be an optimal trajectory and $x^*(t_i) \in M_{q_{i-1},q_i}$ to be an optimal switching state. Then $x^*(t_i)$ belongs to a closed region locally defining the maximally reachable region for the i-th switch by control trajectories $u(t) \in \mathcal{U}$. The time interval $I_i = [a_i, b_i]$ denotes the time span in which that region can be reached. All time intervals I_i are collected in $\mathcal{I} = \{I_i\}_{i\in\{1,...,N\}}$. □

Definition 5: Direct neighbors of discrete state $j \in \mathcal{Q}$ are discrete states $d_1, ..., d_{n_d} \in \mathcal{Q}$ whose state spaces $\mathcal{X}_{d_1}, ..., \mathcal{X}_{d_{n_d}}$ have a common boundary with $\mathcal{X}_j$. Switching manifolds $m_{j,d_1}, ..., m_{j,d_{n_d}}$ denoting the common boundaries are called *first-order neighboring switching manifolds*. Suppose a trajectory $\chi(t)$ passes discrete state j. Then those manifolds $m_{j,d_1}, ..., m_{j,d_{n_d}}$ that $\chi(t)$ crosses when leaving discrete state j are denoted by $\mathcal{N}_j^1 := \{m_{j,k}|m_{j,k}(x(t)) = 0, j, k \in \mathcal{Q}\}$. *Second-order neighboring switching manifolds* m_{d_k,d_l} with $d_k, d_l \in \{d_1, ..., d_{n_d}\}$ separate two neighboring discrete states d_k and d_l, which are reachable from state

[1] It is possible to consider hybrid system executions beyond Zeno points [16], but as Zeno points are not in the scope of this paper, they are excluded by assumption.

j, from each other. The switching manifolds of second-order, whose boundaries $\partial m_{d_k,d_l} \subset \partial\mathcal{X}_j$ are hit by $\chi(t)$, are collected in the set $\mathcal{N}_j^2$. □

Theorem 1 ([15]): Let *Assum.* 1, 2, 3, and 4 hold, then all controls u^* (locally) minimizing the cost functional

$$J^*(u^*) = \inf_{u \in \mathcal{U}} J(u) \tag{5}$$

lead to the fulfillment of the following conditions:

1) The differential equation (3) is satisfied.
2) There exists an optimal, absolutely continuous adjoint process λ^* such that:

$$\dot{\lambda}^* = -\nabla_x H_{q_i^*} \tag{6}$$

for a.e. $t \in [t_i, t_{i+1})$ and every $i \in \{0, 1, ..., N\}$, $q_i^* \in q^*(t)$ with the Hamiltonian $H_{q_i}(x(t), \lambda(t), u_{q_i}(t)) = \phi_{q_i}(x(t), u_{q_i}(t)) + \lambda(t)^T f_{q_i}(x(t), u_{q_i}(t))$. The following boundary conditions hold for λ^*:

a) Terminal condition if $x(t_e)$ is not specified:

$$\lambda^*(t_e) = \nabla_x g(x(t_e)) \tag{7}$$

b) If an autonomous transition with $m_{q_{i-1}^*,j} \in \mathcal{N}_{q_{i-1}^*}^1$ and $m_{j,k} \in \mathcal{N}_{q_{i-1}^*}^2$ with $q_{i-1}^*, j, k \in \mathcal{Q}$ is triggered at time t_i, $i \in \{1, ..., N\}$, then:

$$\begin{aligned} \lambda^*(t_i-) &= \lambda^*(t_i) \\ &+ \sum_{m_{q_{i-1}^*,j} \in \mathcal{N}_{q_{i-1}}^1} \pi_{i,(q_{i-1},j)}^* \nabla_x m_{q_{i-1}^*,j}(x(t_i)) \\ &+ \sum_{m_{j,k} \in \mathcal{N}_{q_{i-1}}^2} \pi_{i,(j,k)}^* \nabla_x m_{j,k}(x(t_i)) \end{aligned} \tag{8}$$

with constant and optimal Lagrange multipliers $\pi_{i,(q_{i-1}^*,j)}^*$, $\pi_{i,(j,k)}^* \in \mathbb{R}$, $q_{i-1}^*, j, k \in \mathcal{Q}$.

3) The Hamiltonian has to fulfill the following conditions:

a) If at time t_i, $i \in \{0, 1, ..., N\}$ the system switches autonomously from q_{i-1}^* to q_i^*, $q_{i-1}^*, q_i^* \in \mathcal{Q}$, then:

$$\begin{aligned} H_{q_{i-1}^*}(t_i-) &= H_{q_i^*}(t_i) \quad \text{if} \quad t_i \in (a_i, b_i) \\ H_{q_{i-1}^*}(t_i-) &\geq H_{q_i^*}(t_i) \quad \text{if} \quad t_i = a_i \\ H_{q_{i-1}^*}(t_i-) &\leq H_{q_i^*}(t_i) \quad \text{if} \quad t_i = b_i \ . \end{aligned} \tag{9}$$

b) The minimization condition with respect to $u_{q_i^*}^*$, $q_i^* \in Q$ is:

$$H_{q_i^*}(x^*, \lambda^*, u_{q^*}^*) \leq H_{q_i^*}(x^*, \lambda^*, v) \tag{10}$$

for a.e. $t \in [t_i, t_{i+1})$, $i \in \{0, 1, ..., N\}$, and for every $v \in U_{q_i^*}$. □

III. ALGORITHM

This section introduces an algorithm that finds a locally optimal trajectory $\chi^*(t)$ based on the optimality conditions of *Theorem* 1. The optimality conditions are utilized such that the discrete state sequence may be varied during the search. This implies that the algorithm can be initialized with an almost arbitrary sequence, which will be changed until a local minimum of the cost is reached. By varying the switching sequence during a single run of the algorithm, the exponential complexity of earlier algorithms in a neighborhood around a local minimum is avoided. Earlier algorithms that use indirect shooting methods can only solve the HOCP for a pre-specified sequence of discrete states q, i.e. finding a locally optimal solution required to search over every possible sequence in a certain neighborhood, leading to a combinatorial complexity.

A. Concept

The algorithm contains two layers. The optimal control problem $J(u)$ (2) is transferred to a problem $J(z)$ depending on the switching point vector z. This is achieved by considering between two consecutive switching points z_i and z_{i+1} only state and control trajectories, that satisfy the optimality conditions (3), (6), (7), and (10). Finding the trajectories between two switching points forms the lower layer. On a higher layer of the algorithm, the switching points are varied until the optimality conditions (8) and (9) are satisfied.

In every iteration l, the algorithm produces a feasible sequence of discrete states $q^l = (q_0^l, ..., q_{N(l)}^l)$ and switching points $z_i^l = (x_i^{lT} \ t_i^l)^T$, $i = 1, ..., N(l)$. Here, feasibility means that an optimal control $u(t) \in \mathcal{U}$ exists, which transfers system (3) from switching point z_i^l to z_{i+1}^l, $i \in \{0, 1, ..., N(l)\}$, without leaving the state space $\mathcal{X}_{q_i}$.

Algorithm 1:

Step 0: Find a feasible initial sequence $q^0 = (q_0^0, ..., q_{N(0)}^0)$ with switching points z_i^0. Set $l = 0$ and $j = 1$.

Step 1: Solve the optimal control subproblems

$$\min_{u(t)} J_{q_i} = \min_{u(t)} \int_{t_i^l}^{t_{i+1}^l} \phi_{q_i}(x(t), u(t))\, dt \tag{11}$$

for every discrete state $q_i \in q^l$, $i \in \{0, 1, ..., N(l)\}$. In the last discrete state $q_{N(l)}$, the terminal costs $g(x(t_e))$ are added to subproblem (11). The optimization uses z_i^l as initial value and z_{i+1}^l as bounding condition, and it takes the dynamics (3) and feasible control trajectories $\mathcal{U}_{q_i}$ into account. Indirect shooting on first-order optimality conditions is used [8, 17].

Step 2: Determine the gradient of the cost $\nabla_{z_i^l}^m J(z^l)$ projected onto the corresponding switching manifolds $m_{q_{i-1}^l, q_i^l}$ for all z_i^l. If $\nabla_{z_i^l}^m J(z^l) = 0 \ \forall z_i^l$, then stop.

Step 3: Execute an Armijo-like step [18] for the update of the switching point z_j^l to reduce the cost based on the gradient information. The Armijo-like step is taken along the switching manifold $m_{q_{j-1}^l, q_j^l} = 0$ in direction $-\nabla_{z_j^l}^m J(z^l)$, such that $x_j^{l+1} \in M_{q_{j-1}^l, q_j^l}$. If the step proceeds beyond the boundary $\partial m_{q_{j-1}^l, q_j^l}$, then z_j^{l+1} is projected onto the neighboring switching manifold $m_{q_{j-1}^{l+1}, q_j^{l+1}}$ with $q_{j-1}^l = q_{j-1}^{l+1}$ and $q_j^l \neq q_j^{l+1}$ and possibly further manifolds. The Armijo step is taken for each switching point z_j^l separately, as in simulations, this turned out to be more effective than updating all z_j^l at once. The optimization problems (11) are solved for all discrete states in the sequence resulting from the Armijo step. If the Armijo step is successful, then increase l and j and goto Step 4, else repeat Step 3 with a decreased step-size.

Step 4: If $j > N(l)$, which means that the algorithm has just updated the last switching point $z^l_{N(l)}$, then set $j = 1$ to start updating z^{l+1}_1 and goto Step 1. □

After giving a short summary of the algorithm, the most important steps are discussed in detail.

B. Gradient of Costs on Switching Manifold

In order to update a switching point z^l_i, a direction $d_{z^l_i}$ is chosen, which leads to a lower overall cost $J(z^{l+1})$. Here, the projected gradient $-\nabla^m_{z^l_i} J(z^l)$ of the cost $J(z^l)$ is used as descent direction. Projected means that only the part of the gradient $\nabla_{z^l_i} J(z^l)$ is considered, which is tangential to the switching manifold $m_{q^l_{i-1},q^l_i}$. The analytic expression for

$$\nabla^m_{z^l_i} J(z^l) = \left(\nabla^m_{x^l_i} J^T(z^l)\ \nabla^m_{t^l_i} J(z^l)\right)^T \tag{12}$$

is basically given by the optimality conditions (8) and (9):

$$\begin{aligned}\nabla^m_{x^l_i} J(z^l) = &\lambda(t^l_i) - \lambda(t^l_i-)\\ &+ \sum_{m_{q_{i-1},j} \in \mathcal{N}^1_{q_{i-1}}(0)} \pi_{i,(q_{i-1},j)} \nabla_x m_{q_{i-1},j}(x(t^l_i))\\ &+ \sum_{m_{j,k} \in \mathcal{N}^2_{q_{i-1}}(0)} \pi_{i,(j,k)} \nabla_x m_{j,k}(x(t^l_i))\end{aligned} \tag{13}$$

$$\nabla^m_{t^l_i} J(z^l) = H_{q_{i-1}}(t^l_i-) - H_{q_i}(t^l_i)\ , \tag{14}$$

where $q_{i-1}, j, k \in \mathcal{Q}$. This result follows directly from the proof of (8) and (9), e.g. in [15], or from results and proofs of the standard calculus of variations [19]. To obtain $\nabla^m_{x^l_i} J(z^l)$, it is necessary to know all π values in (13). Unique solutions are obtained, if only the largest set of linearly independent vectors $\nabla_x m_{q_{i-1},j}(x(t^l_i))$ and $\nabla_x m_{j,k}(x(t^l_i))$ from (13) is considered. Combining the selected vectors in the matrix Ψ and defining $\theta := \lambda(t^l_i-) - \lambda(t^l_i)$, then the vector $\pi = (\pi_{i,(q_{i-1},q_i)}, \ldots)^T$ can be found as the minimizer of $\nabla^m_{x^l_i} J(z^l)$ by a least squares regression, see [10]:

$$\pi = (\Psi^T \Psi)^{-1} \Psi^T \theta\ . \tag{15}$$

The descent direction for the Armijo step at z^l_i is:

$$d_{z^l_i} = -\nabla^m_{z^l_i} J(z^l)\ . \tag{16}$$

C. Projection of Switching Points

In general, the projection of switching points on nonlinear switching manifolds and for non-convex state partitions $\mathcal{X}_q$ can be handled numerically. To enable a simplified implementation of the algorithm, the switching manifolds are assumed to be affine and all partitions $\mathcal{X}_q$ to be convex. For affine switching manifolds, non-convex $\mathcal{X}_q$ can be convexified by introducing additional switching manifolds. For convex $\mathcal{X}_q$, also all switching manifolds are convex.

To update the switching point z^l_i, a candidate $\bar{z}^{l+1}_i$ is found by proceeding from z^l_i in the descent direction with step size α:

$$\bar{z}^{l+1}_i = z^l_i + \alpha d_{z^l_i}\ . \tag{17}$$

If $\bar{z}^{l+1}_i \in M_{q^l_{i-1},q^l_i}$ the update of the switching point is valid: $z^{l+1}_i = \bar{z}^{l+1}_i$. If $\bar{z}^{l+1}_i \notin M_{q^l_{i-1},q^l_i}$, it is on the unbounded extension of $M_{q^l_{i-1},q^l_i}$. Then, $\bar{z}^{l+1}_i$ has crossed the boundary $\partial m_{q^l_{i-1},q^l_i}$, which implies that the update step $\alpha d_{z^l_i}$ enforces a new sequence q^{l+1} of discrete states. Note that for reasons of convergence only sequences q^{l+1} are considered that are neighbors of the old sequence q^l, i.e. there must exist trajectories $\chi^{l+1}(t)$ with q^{l+1} and $\chi^l(t)$ with q^l that are arbitrarily close to each other. The old sequence $q^l = (*, \Delta, *)$ differs from the new one $q^{l+1} = (*, \Box, *)$ in the subsequences $\Delta = (q^l_i, \ldots, q^l_{i+\rho-1})$ and $\Box = (q^{l+1}_i, \ldots, q^{l+1}_{i+\varpi-1})$ with $\rho, \varpi \in \mathbb{N}_0$. Here, ρ and ϖ denote how many discrete states disappear from the old sequence q^l and how many are added to q^{l+1}, respectively. For example, $\rho = 0$ and $\varpi = 1$ imply that $\Delta = \emptyset$ and $\Box = \{q^{l+1}_i\}$. For simplicity and without loss of generality, in the following only $\rho = 0$ and $\varpi = 1$ is considered.

For $m_{q^l_{i-1},q^l_i}(x^l_i) = 0$, assume that the trajectory of iteration l jumps from discrete state q^l_{i-1} to q^l_i. Then the sequence $(q^l_0, \ldots, q^l_{i-1})$ remains the same for iteration $l+1$, since only the switching point z^l_i is shifted. From the switching point z^l_i lying on the manifold $m_{q^l_{i-1},q^l_i}(x^l_i) = 0$ and with the scaled descent direction $\alpha d_{z^l_i}$, the switching point z^{l+1}_i has to be found on a different switching manifold $m_{q^{l+1}_{i-1},q^{l+1}_i}(x^{l+1}_i) = 0$ with $q^l_{i-1} = q^{l+1}_{i-1}$ and $q^l_i \neq q^{l+1}_i$. To find z^{l+1}_i, a projection $P(z^l_i, \alpha d_{z^l_i})$ that maintains the length $||\alpha d_{z^l_i}||$ is applied. The projection works as follows: Basically a circle around z^l_i is considered with radius $||\alpha d_{z^l_i}||$ and tangent $\nabla_x m_{q^l_{i-1},q^l_i}(x^l_i)$ at $\bar{z}^{l+1}_i$. Each intersection of the circle with switching manifolds delivers a candidate for a new switching point. Out of the set of candidate points, a subset is picked, which contains the switching points z^{l+1}_i and z^{l+1}_{i+1} leading from $q^l_{i-1} = q^{l+1}_{i-1}$ over q^{l+1}_i to $q^{l+1}_{i+1} = q^l_i$ (since $\rho = 0$ and $\varpi = 1$ in this example). From $q^{l+1}_{i+1} = q^l_i$ on, the rest of the old sequence $(q^l_{i+1}, \ldots, q^l_{N(l)+1})$ with the old switching points completes the new sequence q^{l+1}. If no intersection of the circle with a switching manifold restricting q^l_{i-1} exists, then the projection is declared to be infeasible and a new trial is started with a reduced step size $||\alpha d_{z^l_i}||$.

The property of the projection $P(z^l_i, \alpha d_{z^l_i})$ to keep the length of the update step equal for all candidate points becomes important in the convergence analysis.

D. Update Step

The new sequence q^{l+1} and the new switching points z^{l+1} found in Part III-C are checked with an Armijo-like condition [18] for suitability. The update is accepted if:

$$J(z^{l+1}) - J(z^l) \leq \begin{cases} -\tau_0 \alpha d^T_{z^l_i} d_{z^l_i} & \text{if } z^{l+1}_i \in M_{q^l_{i-1},q^l_i} \\ -\tilde{\tau}_0 \alpha d^T_{min} d_{min} & \text{if } z^{l+1}_i \notin M_{q^l_{i-1},q^l_i} \end{cases}\ , \tag{18}$$

with

$$||d_{min}|| = \min(||d_{z^l_i}||, ||d_{z^{l+1}_i}||, ||d_{z^{l+1}_{i+1}}||) \tag{19}$$

for $\varpi = 1$, with $\tilde{\tau}_0 = \tau_0\beta$, and the design parameters $\tau_0, \beta \in (0,1)$. Introducing d_{min} in (18) extends the usual Armijo criterion [18]. The extension turned out to be necessary for the convergence analysis and guarantees that the discrete state sequence changes if this is required to find a locally optimal sequence. If the updated switching point vector satisfies (18), then z^{l+1} and q^{l+1} are valid and the algorithm updates z_{i+1}^{l+1} in the next iteration, otherwise α is decreased.

Here, $\alpha := \{\max(c\beta^p) | p \in \mathbb{N}_0, p \text{ s.t. } c\beta^p \text{ fulfills } (18)\}$ with $c \in \mathbb{R}^+$ is chosen as the largest value $c\beta^p$ fulfilling (18) [20]. Armijo's condition assures sufficient decrease of the cost J for a sufficiently large distance between z_i^l and the corresponding points z_i^{l+1} and z_{i+1}^{l+1} in the sequence q^{l+1}. In the following, the scaling factor c is omitted for shorter notation.

IV. CONVERGENCE

Here, the classical concept of convergence of an algorithm [21] cannot be applied, since the dimension $N(l)$ of each switching point vector z^l (4) in the sequence of iterations $\{z^l\}_{l=1}^{\infty}$ may change at every iteration l. Furthermore, the gradient of the cost projected onto the switching manifolds $\nabla_z^m J(z^l) = (\nabla_{z_1^l}^m J^T(z^l), ..., \nabla_{z_{N(l)}^l}^m J^T(z^l))^T$ is only piecewise continuous. However, the cost J is continuous on the whole state space according to *Lemma* 1 in [15]. For these reasons, a different concept of *convergent algorithm* similar to the one used in [22] is introduced.

Definition 6: An algorithm is convergent if every sequence $\{z^l\}_{l=1}^{\infty}$ fulfills the elementwise condition $\lim_{l\to\infty} ||z_i^l - z_j^*|| = 0$ for $i \in \{1, ..., N(l)\}$ and its corresponding $j \in \{1, ..., N^*\}$, where $\nabla_z^m J(z^*) = 0$. □

Remark 1: For simplicity, it is assumed that an optimal switching point vector z^* satisfies $\nabla_z^m J(z^*) = 0$. But note that there exist cases, where $\nabla_{t_j^l}^m J(z^*) \neq 0$ for some $j \in \{1, ..., N^*\}$ according to (9). □

In the sequel, it is shown step-by-step that *Algorithm* 1 is convergent, compare also to the proof of the algorithm for systems with controlled switching in [22]. As derived later, the algorithm converges, if there is sufficient descent for a bounded cost J.

Lemma 1: The cost J is bounded for any feasible trajectory $\chi(t)$ with initial condition x_0 at t_0, switching points z_i, and final condition x_e or t_e. □

Proof: From *Assum.* 1, $f_q(x(t), u_q(t))$ is Lipschitz continuous and bounded with respect to its arguments. For all $q \in \mathcal{Q}$, the trajectory $\chi(t)$ is in a compact set. Here, it is assumed that, if x_e is specified, it is reached in finite time. As the cost J is continuous and $g(x(t_e))$ and $\phi_q(x(t), u_q(t))$ are bounded with respect to their arguments, $|J(z)|$ is bounded from above. ■

Definition 7: *Algorithm* 1 has *sufficient descent* if for every iteration l and every $\kappa > 0$, there exists $\eta > 0$ such that, if elementwise $||z_i^l - z_j^*|| > \kappa$ for $i \in \{1, ..., N(l)\}$ and its corresponding $j \in \{1, .., N^*\}$, then

$$J(z^{l+1}) - J(z^l) \leq -\eta \,. \tag{20}$$

□

Proposition 1: If *Algorithm* 1 has sufficient descent and bounded cost, then it is convergent. □

Proof: If the algorithm has sufficient descent, then at every switching point z^l in the infinite sequence $\{z^l\}_{l=1}^{\infty}$ the cost J will decrease by a non-zero amount. As J is bounded, this must lead to convergence in the sense of *Def.* 6. ■

Now, it is shown that the applied Armijo-like criterion ensures that the algorithm has sufficient descent. At first, it is proven that an Armijo-like step is always possible if elementwise $||z_i^l - z_j^*|| > \kappa$ for any $\kappa > 0$, for $i \in \{1, ..., N(l)\}$ and corresponding $j \in \{1, .., N^*\}$. Here, the piecewise differentiability of J projected onto switching manifolds is applied.

Proposition 2: Let J be continuously differentiable in a neighborhood $B(z^l, r)$, $r > 0$, $||z^{l+1} - z^l|| \leq r$ and let $\tau_0 \in (0,1)$, $\beta \in (0,1)$ be specified. Set $z^{l+1} = z^l + \beta^p d_{z^l}$ and $d_{z^l} = -\nabla_z^m J(z^l)$, so that $\nabla_z^m J^T(z^l)\, d_{z^l} < 0$, then there exists a finite $p \in \mathbb{N}_0$, such that

$$J(z^l + \beta^p d_{z^l}) - J(z^l) \leq -\tau_0 \beta^p d_{z^l}^T\, d_{z^l} \,, \tag{21}$$

i.e. the Armijo-like criterion is well-defined for updates z^{l+1} that leave the sequence $q^l = q^{l+1}$ unchanged. □

Proof: See [20]. ■

Condition (21) can easily be shown to hold also for variations at single switching points z_i^l with the gradient of the cost $\nabla_{z_i^l}^m J(z^l)$. Next, it is shown that the criterion also holds in a similar form for updates leading to a change in the discrete state sequence. Therefore, define the switching point $z_{s,i}^l = z_i^l + \beta^s d_{z_i^l}$ with $s \in \mathbb{R}$ on the intersection of two neighboring switching surfaces, that means $z_{s,i}^l \in \partial m_{q_{i-1}^l, q_i^l} \cap \partial m_{q_{i-1}^{l+1}, q_i^{l+1}}$. Furthermore, a switching point z^l, that remains unchanged except for a change of $\beta^p d_{z_i^l}$ in the i-th subvector, is denoted by $z^l(\beta^p d_{z_i^l})$.

Proposition 3: For iteration k, let $z_j^* \notin M_{q_{i-1}^k, q_i^k}$ $\forall j \in \{1, ..., N^*\}$, $||z_j^* - z|| > \delta$ $\forall z \in M_{q_{i-1}^k, q_i^k}$, $\delta > 0$, and $z_i^k \in M_{q_{i-1}^k, q_i^k}$. Then there exists an iteration $l \in \mathbb{N}$, $l \geq k$, and a $p \in \mathbb{N}_0$ such that

$$J(z^{l+1}) - J(z^l) \leq -\tilde{\tau}_0 \beta^p d_{min}^T d_{min} \,, \tag{22}$$

where $||d_{min}|| = \min(||d_{z_i^l}||, ||d_{z_i^{l+1}}||, ||d_{z_{i+1}^{l+1}}||)$ and the updated sequence q^{l+1} is different from the old one q^l. □

The idea of the proof is the following: An update of a switching point leading to a new discrete state sequence is separated into a part that remains on the old switching manifold and a part on one of the new switching manifolds. Formulating the difference between the costs of the new and the old discrete state sequence in terms of Armijo steps in both parts of the update step and combining the results, then it can be shown that (22) can always be fulfilled.

Proof: Assume first, that $z_i^k \in \hat{M}_{q_{i-1}^k, q_i^k}$, where $\hat{M}_{q_{i-1}^k, q_i^k}$ is the interior of $M_{q_{i-1}^k, q_i^k}$. By *Prop.* 2 and $z_i^l \in \hat{M}_{q_{i-1}^k, q_i^k}$ being the result of several Armijo steps (21) starting from z_i^k with $l \geq k$, $||z_i^l - z_{s,i}^l||$ can be made arbitrarily small with increasing l. In particular for some l, there exists κ^l,

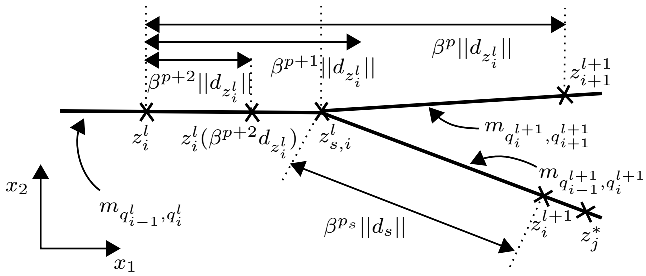

Fig. 1. Switching points used in the proof.

s.t. $||z_i^l - z_j^*|| > \kappa^l > \frac{1}{\beta^2}||z_i^l - z_{s,i}^l||$. Now, let $l, p \in \mathbb{N}$, such that

$$z_i^l + \beta^{p+2} d_{z_i^l} \in M_{q_{i-1}^l, q_i^l} \tag{23}$$

$$z_i^l + \beta^{p+1} d_{z_i^l} \notin M_{q_{i-1}^l, q_i^l} \tag{24}$$

$$J(z^l(\beta^{p+y} d_{z_i^l})) - J(z^l) \leq -\tau_0 \beta^{p+y} d_{z_i^l}^T d_{z_i^l} \tag{25}$$

for every $y \in \{0, 1, 2\}$ and $d_{z_i^l} = -\nabla^m_{z_i^l} J(z^l)$, see Fig. 1. This implies

$$\beta^{p+2}||d_{z_i^l}|| \leq ||z_i^l - z_{s,i}^l|| \leq \beta^{p+1}||d_{z_i^l}||\,, \tag{26}$$

which leads to $\beta^p||d_{z_i^l}|| \leq \frac{1}{\beta^2}||z_i^l - z_{s,i}^l|| < \kappa^l$. The projection $P(z_i^l, \beta^p||d_{z_i^l}||)$ delivers the new switching points $z_v \in \{z_i^{l+1}, z_{i+1}^{l+1}\}$, which are used if the Armijo-like test (22) is successful. As the projection maintains the step length,

$$||z_v - z_i^l|| = \beta^p||d_{z_i^l}|| < \kappa^l \tag{27}$$

is true $\forall z_v$. The following inequalities hold $\forall z_v$:

$$\begin{aligned} ||z_v - z_i^l|| &= ||z_v - z_{s,i}^l + z_{s,i}^l - z_i^l|| \\ &\leq ||z_v - z_{s,i}^l|| + ||z_{s,i}^l - z_i^l|| \\ \Rightarrow ||z_v - z_{s,i}^l|| &\geq ||z_v - z_i^l|| - ||z_{s,i}^l - z_i^l|| \\ &\overset{(27),(26)}{\geq} \beta^p(1-\beta)||d_{z_i^l}||\ . \end{aligned} \tag{28}$$

Define z_s^l to be the vector of switching points that only differs from z^l in replacing z_i^l by $z_{s,i}^l$. By *Prop.* 2 it can be concluded that there exists a finite $l \in \mathbb{N}$ leading to a sufficiently small $\beta^p||d_{z_i^l}||$, such that $\beta^{p_s}||d_s||$ exists to satisfy:

$$J(z^{l+1}) - J(z_s^l) \leq -\tau_0 \beta^{p_s} d_s^T d_s\,, \tag{29}$$

where d_s and β^{p_s} such that $||d_s|| = \min(||d_v||)$, $v \in \{i, i+1\}$, $d_v = -\nabla^m_{z_v} J(z_s^l)$, and the corresponding

$$z_v = z_{s,i}^l + \beta^{p_s} d_s\,. \tag{30}$$

By the continuity of J and $||z_v - z_i^l|| < ||z_i^l - z_j^*||$, it follows for sufficiently small $\beta^p||d_{z_i^l}||$ that $d_s^T d_v > 0$. Now, it remains to show that (22) holds:

$$\begin{aligned} J(z^{l+1}) - J(z^l) &= J(z^{l+1}) - J(z_s^l) + J(z_s^l) - J(z^l) \\ &\leq J(z^{l+1}) - J(z_s^l) + J(z^l(\beta^{p+2}d_{z_i^l})) - J(z^l) \\ &\overset{(29),(25)}{\leq} -\tau_0\beta^{p_s} d_s^T d_s - \tau_0 \beta^{p+2} d_{z_i^l}^T d_{z_i^l} \\ &\overset{(30),(28)}{\leq} -\tau_0\beta^p||d_{z_i^l}|| \left[(1-\beta)||d_s|| + \beta^2||d_{z_i^l}||\right] \\ &\leq -\tau_0\beta^p||d_{z_i^l}|| \left[\beta(1-\beta)||d_s|| + \beta^2||d_{z_i^l}||\right] \\ &\overset{(19)}{\leq} -\tau_0\beta\beta^p||d_{z_i^l}||\,||d_{min}|| \\ &\leq -\tilde{\tau}_0\beta^p d_{min}^T d_{min}\,. \end{aligned} \tag{31}$$

The same result can be shown in a similar way, if $z_i^l \in \partial m_{q_{i-1}^l, q_i^l} \cap \partial m_{q_{i-1}^{l+1}, q_i^{l+1}}$ with $q_i^l \neq q_i^{l+1}$. ■

Remark 2: For simplified notation, *Prop.* 3 is given for $\rho = 0$ and $\varpi = 1$. The proof can be extended to handle all cases of possible combinations of $\rho, \varpi \in \mathbb{N}_0$. □

Combining the results above, the sufficient descent property of the algorithm follows.

Proposition 4: For every $l \in \mathbb{N}$ and $\kappa^l > 0$ there exists $\eta > 0$, such that, if $||z_i^l - z_j^*|| > \kappa^l$ for $i \in \{1, ..., N(l)\}$ and corresponding $j \in \{1, .., N^*\}$, then

$$J(z^{l+1}) - J(z^l) \leq -\eta\,. \tag{32}$$

□

Proof: If $||z_i^l - z_j^*|| > \kappa^l$ for $i \in \{1, ..., N(l)\}$ and corresponding $j \in \{1, .., N^*\}$, then there exists $\epsilon > 0$ such that $||\nabla^m_{z_i^l} J(z^l)|| = ||d_{z_i^l}|| > \epsilon$. *Prop.* 2 and 3 imply that there always exists a finite $p \in \mathbb{N}_0$, such that a new switching point z^{l+1} can be determined with $||z_v - z_i^l|| = \beta^p||\nabla^m_{z_i^l} J(z^l)||$ $\forall z_v \in \{z_i^{l+1}, ..., z_{i+\varpi}^{l+1}\}$. Here, z_v is any updated switching point and may also be on $M_{q_{i-1}^l, q_i^l}$. If the algorithm does not converge in a finite amount of iterations, there exists again a $\kappa^{l+1} > 0$ such that $||z_v - z_j^*|| > \kappa^{l+1}$. This again implies that $||\nabla^m_{z_v} J(z^{l+1})|| = ||d_v|| > \epsilon\ \forall z_v$, where ϵ is set to fulfill $||d_{min}|| > \epsilon > 0$. Since at each iteration l either (21) or (22) are satisfied, where $\tilde{\tau}_0 < \tau_0$, $||d_{z_i^l}|| > \epsilon$ as well as $||d_{min}|| > \epsilon$, the sufficient descent condition holds:

$$J(z^{l+1}) - J(z^l) \leq -\tilde{\tau}_0 \beta^p \epsilon^2 = -\eta\,. \tag{33}$$

■

Theorem 2: A sequence $\{z^l\}_{l=1}^{\infty}$ of switching points, which is a solution of *Algorithm* 1, satisfies the elementwise equality:

$$\lim_{l \to \infty} ||z_i^l - z_j^*|| = 0 \tag{34}$$

for $i \in \{1, ..., N(l)\}$ and corresponding $j \in \{1, .., N^*\}$ with $\nabla^m_z J(z^*) = 0$. □

Proof: The result is a direct consequence of *Prop.* 4 and 1. ■

Remark 3: If it is not possible to shift z_i^l in direction of $d_{z_i^l}$ because the control $u \notin \mathcal{U}$ is not admissible, then it might be necessary to shift some other switching points z_j^l, $j \in \{1, ..., N(l)\}$, $j \neq i$ before. If no switching point z_i^l can be shifted due to restrictions on the control, then $\nabla^m_z J(z^l)$ is zero (consider *Remark* 1) according to *Theorem* 1. □

Remark 4: In the current formulation of the algorithm and the convergence analysis, it is assumed that the optimal trajectory $\chi(t)$ which starts in z_i^l and reaches z_{i+1}^l stays in $\mathcal{X}_i$ for $t \in [t_i^l, t_{i+1}^l]$. If the assumption cannot be met for a specific problem, then it is necessary to add and remove switching points in-between two switching points. □

V. NUMERICAL EXAMPLES

The example from [15] is considered again, with the modification that $\mathcal{X}_4$ is convexified. The partitioned state space contains 11 discrete states, where each discrete state has a linear dynamics $\dot{x} = A_q x + B_q u$ with $x \in \mathbb{R}^2$, $u \in U_q \subset \mathbb{R}^2$ and a running cost function $\phi_q = 0.5x^T S_q x + 0.5u^T R_q u$. There is no terminal cost function and no terminal state and discrete state specified, the initial states are $x_0 = (-8\ -8)^T$ and $q_0 = 1$, and initial and final time are $t_0 = 0$ and $t_e = 2$. The system is designed such that there exists a unique optimum. The algorithm changes the initial sequence $q = (1, 3, 4, 6, 7, 10)$ to $q = (1, 3, 4, 7, 10)$, $q = (1, 3, 4, 7, 8, 10)$, and finally to the optimal sequence $q^* = (1, 3, 4, 5, 7, 8, 10)$ with cost $J = 14.6966$ in 11.4 min computational time. The computations were performed on a 3.0 GHz processor using Matlab 2009a. The results are illustrated in Fig. 2.

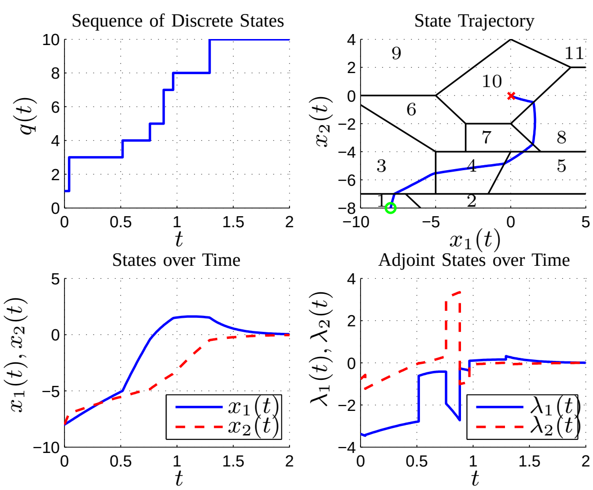

Fig. 2. Optimization results: optimal discrete state sequence, optimal system trajectory, corresponding state and adjoint states.

The optimal solution and the computations are compared to results of a brute force implementation based on multiple shooting. For a global approach, the brute force algorithm investigates every possible discrete state sequence in the partitioned state space with at most 6 switchings beginning from q_0. The brute force solver finds the same optimum except for higher accuracy with $q^* = (1, 3, 4, 5, 7, 8, 10)$ and $J = 14.6949$. The brute force solver looks at 3976 different sequences separately and needs $18,508.4$ min for the computation on a 1.6 GHz processor using Matlab 2008b.

VI. CONCLUSIONS

An algorithm is introduced that finds a locally optimal discrete state sequence and locally optimal state and control trajectories of hybrid systems defined on a partitioned state space. Here, the algorithm avoids the problem of former algorithms based on the HMP that the computational complexity increases exponentially with the amount of switchings. In convex hybrid optimal control problems, the global optimal solution can be found in a single run of the algorithm. The basis of the algorithm is a version of the HMP that can deal with piecewise differentiable and intersecting manifolds. The convergence of the algorithm is proven and a numerical example shows the efficiency of the approach.

REFERENCES

[1] M. Branicky, V. Borkar, and S. Mitter, "A unified framework for hybrid control: Model and optimal control theory," *IEEE Trans. on Aut. Control*, vol. 43, no. 1, pp. 31–45, 1998.

[2] A. Bemporad and M. Morari, "Control of systems integrating logic, dynamics, and constraints," *Automatica*, vol. 35, pp. 407–427, 1999.

[3] S. Hedlund and A. Rantzer, "Optimal control of hybrid systems," in *38th Conf. on Decision and Control*, 1999, pp. 3972–3977.

[4] H. Sussmann, "A maximum principle for hybrid optimal control problems," in *38th Conf. on Decision and Control*, 1999, pp. 425–430.

[5] B. Passenberg and O. Stursberg, "Graph search for optimizing the discrete location sequence in hybrid optimal control," in *3rd IFAC Conf. on Analysis and Design of Hybrid Systems*, 2009, pp. 304–309.

[6] P. Caines and M. Shaikh, "Optimality zone algorithms for hybrid systems: Efficient algorithms for optimal location and control computation," in *Hybrid Systems: Computation and Control*, ser. LNCS 3927, 2006, pp. 123–137.

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

[8] P. Riedinger, J. Daafouz, and C. Iung, "About solving hybrid optimal control problems," in *IMACS05*, 2005.

[9] M. Egerstedt, Y. Wardi, and H. Axelsson, "Transition-time optimization for switched-mode dynamical systems," *IEEE Trans. on Aut. Control*, vol. 51, no. 1, pp. 110–115, 2006.

[10] M. Shaikh and P. Caines, "On the hybrid optimal control problem: Theory and algorithms," *IEEE Trans. on Aut. Control*, vol. 52, no. 9, pp. 1587–1603, Sept. 2007.

[11] ——, "Correction to "on the hybrid optimal control problem: Theory and algorithms"," *IEEE Trans. on Aut. Control*, vol. 54, no. 6, p. 1440, 2009.

[12] X. Xu and P. Antsaklis, "Optimal control of switched systems based on parameterization of the switching instants," *IEEE Trans. on Aut. Control*, vol. 49, no. 1, pp. 2–16, 2004.

[13] A. Bemporad and N. Giorgetti, "Logic-based solution methods for optimal control of hybrid systems," *IEEE Trans. on Aut. Control*, vol. 51, no. 6, pp. 963–976, 2006.

[14] M. Rungger and O. Stursberg, "Optimal control for deterministic hybrid systems using dynamic programming," in *3rd IFAC Conf. on Analysis and Design of Hybrid Systems*, 2009, pp. 316–321.

[15] B. Passenberg, P. Caines, M. Sobotka, O. Stursberg, and M. Buss, "The minimum principle for hybrid systems with partitioned state space and unspecified discrete state sequence," in *Proc. of the 49th IEEE Conf. on Decision and Control*, 2010.

[16] K. Johansson, M. Egerstedt, J. Lygeros, and S. Sastry, "On the regularization of zeno hybrid automata," *Sys. & Ctrl. Letters*, vol. 38, pp. 141–150, 1999.

[17] H. Bock, "Numerical solution of nonlinear multipoint boundary value problems with applications to optimal control," *ZAMM*, vol. 58, pp. 407–409, 1978.

[18] L. Armijo, "Minimum of functions having lipschitz-continuous first partial derivatives," *Pacific J. of Mathematics*, vol. 16, pp. 1–3, 1966.

[19] A. Bryson and Y. Ho, *Applied Optimal Control: Optimization, Estimation and Control*. Hemisphere Publishing Corporation, 1975.

[20] C. Geiger and C. Kanzow, *Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben*. Springer, 1999.

[21] E. Polak, *Computational Methods in Optimization*, ser. Math. in Science and Eng., R. Bellman, Ed. Academic Press, 1971, vol. 77.

[22] H. Axelsson, Y. Wardi, M. Egerstedt, and E. Verriest, "A provably convergent algorithm for transition-time optimization in switched systems," in *44th IEEE Conf. on Decision and Control*, 2005, pp. 1397–1402.