Switching Time Optimization in Discretized Hybrid Dynamical Systems

Kathrin Flaßkamp, Todd Murphey, and Sina Ober-Blöbaum

***Abstract*— Switching time optimization (STO) arises in systems that have a finite set of control modes, where a particular mode can be chosen to govern the system evolution at any given time. The STO problem has been extensively studied for switched systems that consists of time continuous ordinary differential equations with switching laws. However, it is rare that an STO problem can be solved analytically, leading to the use of numerical approximation using time discretized approximations of trajectories. Unlike the smooth optimal control problem, where differentiability of the discrete time control problem is inherited from the continuous time problem, in this contribution we show that the STO problem will in general be nondifferentiable in discrete time. Nevertheless, at times when it is differentiable the derivative can be computed using adjoint equations and when it is nondifferentiable the left and right derivatives can be computed using the same adjoint equation. We illustrate the results by a hybrid model of a double pendulum.**

I. INTRODUCTION

Physical processes as well as the dynamic behavior of technical systems are typically modeled by systems of continuous time differential equations. However, for an appropriate description of complex behavior and interactions, discrete effects have to be additionally accounted for. Hybrid systems provide a general framework to describe the interaction of continuous dynamics with discrete events. A great interest lies in the optimal control of hybrid systems since this includes not only the computation of optimal control trajectories for the continuous parts but also an optimization of the discrete variables. In this contribution, we focus on the switching time optimization (STO) of switched dynamical systems. That means, the time points, when the system instantaneously switches between different continuous subsystems during an evaluation are treated as design parameters that can be optimized w.r.t. a cost function. The STO problem has been extensively studied for switched systems that consists of time continuous ordinary differential equations with switching laws, e.g. in [1], [2], [3], [4], [5], [6], [7], [8] which we discuss in more detail in Section II.

a) Problem setting: For simplicity, we consider a switched dynamical system on state space $\mathcal{X} \subset \mathbb{R}^n$ that is described by only two different vector fields, f_1 and f_2 (autonomous and continuously differentiable in $x \in \mathcal{X}$). We aim to study the effects of one single switch at time τ, i.e.

$$\dot{x}(t) = \begin{cases} f_1(x(t)) & t < \tau \\ f_2(x(t)) & t \geq \tau \end{cases} \tag{1}$$

starting our observation at time $t = 0$ at the initial state $x(0) = x_{ini}$. (An extension to more than one switching times and several vector fields is straight forward though.) Then a switching time optimization problem is stated as follows

Problem 1.1: Let $\mathcal{X} \subset \mathbb{R}^n$ be a state space with $x_{ini} \in \mathcal{X}$. Let $T, \tau \in \mathbb{R}$ with $0 \leq \tau \leq T$, $f_1, f_2 \in C^1$ and $\ell \in C^1$ (continuously differentiable). We consider the following problem

$$\min_\tau J(\tau) = \int_0^T \ell(x(t), t)\, dt \tag{2}$$

w.r.t. $\dot{x}(t) = \begin{cases} f_1(x(t)) & t < \tau \\ f_2(x(t)) & t \geq \tau \end{cases} \quad x(0) = x_{ini}.$

A necessary condition for τ being the optimal switching time is $J'(\tau) = \frac{d}{d\tau} J(\tau) = 0$. Commonly, descent techniques are employed that are based on the derivative of J (cf. Section II for a formula to calculate $J'(\tau)$).

b) Discretization: In most cases, it is not possible to solve Problem 1.1 analytically. Therefore, numerical methods for integration and optimization have to be applied in order to approximate an optimal solution. They are based on a discretization of Problem 1.1. Euler integration is a common numerical integration scheme for differential equations, while an integral cost function can be approximated e.g. by the trapezoidal rule. To discretize (1), we choose a discrete time grid $\{t_k\}_{k=0}^N = \{t_0, t_1, \ldots, t_N\}$ (not necessarily equidistant) with $t_0 = 0$, $t_N = T$. A discretized version of Problem 1.1 is given by

Problem 1.2: Let $\{t_k\}_{k=0}^N = \{t_0, t_1, \ldots, t_N\}$ be a discrete time grid with $t_0 = 0$, $t_N = T$ and $\tau \in [t_i, t_{i+1}]$ for some $i \in \{0, \ldots, N\}$. Let $\mathcal{X} \subset \mathbb{R}^n$ be a state space with $x_{ini} \in \mathcal{X}$, $f_1, f_2 \in C^1$ and $\ell \in C^1$. Then we consider the following problem

$$\min_\tau J_d(\tau) = \sum_{k=0}^N \Psi(x_k) \approx \int_0^T \ell(x(t), t)\, dt \tag{3}$$

w.r.t. $\mathbf{K}\left(\{t_k\}_{k=0}^N, \tau, \{x_k\}_{k=0}^N\right) = 0$, with $\mathbf{K}$ being a system of algebraic equations resulting from the discretization of (1). The discretized trajectory $\{x_k\}_{k=0}^N$ is an approximation of the exact solution, i.e. $x_k \approx x(t_k)$.

For the minimization of $J_d(\tau)$ by descent directions, the derivative is required. If it exists, it is given by

$$J_d{}'(\tau) := \frac{d}{d\tau} J_d(\tau) = \sum_{k=0}^N D\Psi(x_k) \cdot \frac{d}{d\tau} x_k. \tag{4}$$

This contribution was partly developed and published in the course of the Collaborative Research Centre 614 Self-Optimizing Concepts and Structures in Mechanical Engineering funded by the German Research Foundation (DFG) under grant number SFB 614.

K. Flaßkamp and S. Ober-Blöbaum are with the Department of Mathematics, University of Paderborn, Warburger Str. 100, 33098 Paderborn, Germany {kathrinf, sinaob}@math.upb.de

T. Murphey is with the Mechanical Engineering Department, Northwestern University, 2145 Sheridan Rd., Evanston, IL 60208, United States t-murphey@northwestern.edu

In the following, we will show that J_d is in general nondifferentiable. Nevertheless, at time points when $\frac{d}{d\tau}x_k$ does exist, we give explicit formulas for it. Unlike the smooth continuous case, nondifferentiable points occur when τ, which is still allowed to vary continuously, coincides with a discrete time point. Analogously to the continuous setting, ${J_d}'(\tau)$ can be evaluated by discrete adjoints, that originate from the necessary optimality conditions of an optimal control problem (cf. Section IV). We extend the discretization scheme for state-adjoint-equations, as e.g. given in [9] to the hybrid case. In our analysis and by illustrating numerical tests we show that the nondifferentiability is less severe for a refined time grid and vanishes when the step size goes to zero. However, our focus does not lie on time grid refinement but on dealing adequately with relatively large step sizes even for highly nonlinear problems typically arising in mechanics. In the research field of discrete mechanics, a great interest lies in structure preserving methods for simulation and optimization [10], [11]. Here, structure preservation and a good longtime energy preservation is not primarily achieved by reducing the time steps, but by finding a discretization that inherits the properties of the continuous time solution. Thus, optimization techniques that allow large step sizes are also of great interest for hybrid mechanical systems.While this is still future work, our contribution provides first ideas how to adress this challenge.

To emphasize the relevance of our research, we introduce the example of a double spherical pendulum that switches between a locked and an unlocked mode.

Example 1.1 (Hybrid locked double pendulum): Postponing all technical details to Section V, the optimization problem for the hybrid double pendulum is to approach a given final position. The numerical example is designed such that this requires one single switch from the locked to the unlocked mode with an optimal switching time $\tau^* = 0.33$. In Fig. 1 the cost function evaluation is given, which is obviously continuous w.r.t. τ. Approximating the derivative $DJ(\tau) = \frac{d}{d\tau}J(\tau)$ (cf. (4) and Sections III–IV for the corresponding formulas) reveals nondifferentiable points, i.e. jumps in the evaluation. They coincide with the grid points of the discretized time interval and do not depend on the specifically chosen discretization scheme for the trajectory, as we will see later. The nondifferentiable points generally lead to problems in gradient based STO techniques. One can observe in Fig. 1 that the discretized objective function in this example is still convex which leads to a uniquely defined τ^* though. However, for general applications, our analysis shows that nondifferentiability has to be accounted for, e.g. by providing subgradients at those points.

c) Outline: The remainder of this contribution is organized as follows: in Section II main results in STO for continuous time dynamics are recalled. In Section III, we study discretized STO problems and present general explicit and implicit integration schemes for hybrid dynamics. In Section IV corresponding discrete adjoint equations are introduced. We illustrate our results by the example of a locked double pendulum in Section V and conclude with an outlook to future work in Section VI.

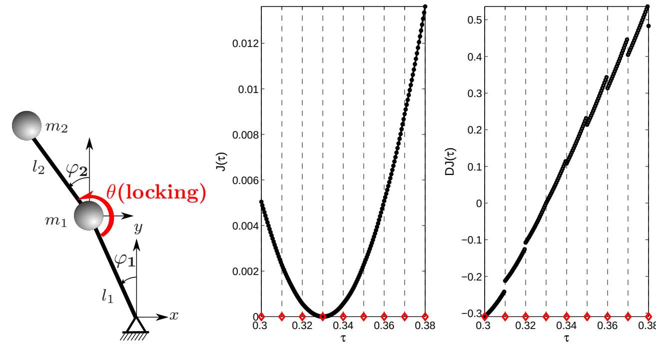

Fig. 1. *Left:* Sketch of the locked double pendulum: in mode 1, the outer pendulum is locked w.r.t. the inner pendulum with angle θ. In mode 2, the system is a normal planar double pendulum. *Right:* Cost function evaluations and its derivative for a switched trajectory of the pendulum: while J is continuous w.r.t. switching time τ, nondifferentiable points occur when τ coincides with a node of the discrete time grid (red dots). This is caused by the approx. trajectory, which is nondifferentiable w.r.t. τ at those points.

II. STO FOR CONTINUOUS PROBLEMS

Derivatives of the cost function w.r.t. switching times in a continuous setting have been studied in several works. We recall from [1]:

Lemma 2.1: Let f_1, f_2 and ℓ from Problem 1.1 be continuously differentiable. Define the costate by

$$\dot{\rho}(t) = -\left(\frac{\partial f_2}{\partial x}(x(t))\right)^T \rho(t) - \left(\frac{\partial \ell}{\partial x}(x(t))\right)^T \qquad (5)$$
$$\rho(T) = 0.$$

Then, $J'(\tau)$ has the following form,

$$J'(\tau) = \rho(\tau)^T[f_1(x(\tau)) - f_2(x(\tau))]. \qquad (6)$$

The result is extended to several switching times and vector fields in [1] and to second order derivatives in [7]. Later we will see that the case, when switching times coincide, are of special importance to us when dealing with discrete approximations of the system dynamics.

Further, the discrete analog of $\frac{dx(t)}{d\tau}$ is required for the derivative of the cost function (cf. (4)). In the continuous setting of Lemma 2.1, it is given for $t \in (\tau, T)$ by (cf. [1])

$$\frac{dx(t)}{d\tau} = \Phi(t,\tau)(f_1(x(\tau)) - f_2(x(\tau))), \qquad (7)$$

with $\Phi(t,\tau)$ being the state transition matrix of the autonomous linear system $\dot{z} = \frac{\partial f_2(x(t))}{\partial x} z$ (cf. [7]).

III. DISCRETIZED STO PROBLEMS

In the following, we consider discretized problems such as Problem 1.2 for explicit and implicit one-step integration schemes.

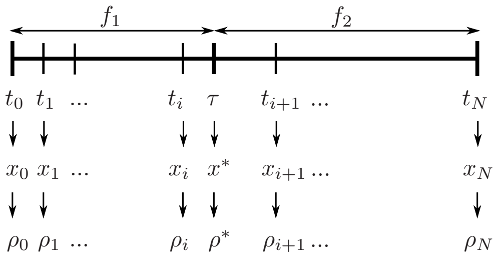

Fig. 2. Notation for discretization as used in the explicit and implicit integration schemes and for the definition of discrete adjoints.

A. Explicit One-step Integration Schemes

Problem 3.1 (STO with explicit scheme): Based on the setting in Problem 1.2, we consider the problem

$$\min_{\tau} J_d = \sum_{k=0}^{N} \Psi(x_k) \quad \text{w.r.t.}$$

$$\mathbf{F} = \begin{cases} x_{k+1} - F_1(x_k, t_k, t_{k+1}) = 0 & k = 0, \dots, i-1, \\ x^* - F_1(x_i, t_i, \tau) = 0 & \text{and} \\ x_{i+1} - F_2(x^*, \tau, t_{i+1}) = 0 & \text{for } k = i, \\ x_{k+1} - F_2(x_k, t_k, t_{k+1}) = 0 & k = i+1, \dots, N-1. \end{cases} \tag{8}$$

F_1 and F_2 denote the schemes for the different vector fields f_1 and f_2 that switch at τ (cf. Fig. 2). Thus, it holds $\tau \in [t_i, t_{i+1}]$ for some $i \in \{0, \dots, N\}$. It can be seen that $\{x_k\}_{k=0}^N$ is continuous w.r.t. τ. For the derivative of $\{x_k\}_{k=0}^N$ w.r.t. τ, the following holds

$$\frac{d}{d\tau} x_{k+1} = \begin{cases} 0 & \text{for } k = 0, \dots, i-1, \\ D_1F_2(x^*, \tau, t_{i+1}) \cdot D_3F_1(x_i, & \\ t_i, \tau) + D_2F_2(x^*, \tau, t_{i+1}) & \text{for } k = i, \\ D_1F_2(x_k, t_k, t_{k+1}) \cdot \frac{d}{d\tau} x_k & \text{for } k = i+1, \dots, N-1. \end{cases} \tag{9}$$

Here, we use the slot derivative notation, i.e. $D_1F_2(\cdot,\cdot,\cdot)$ is the partial derivative of F_2 w.r.t. its first argument, $D_2F_2(\cdot,\cdot,\cdot)$ is the derivative w.r.t. the argument in the second slot and so forth. For $\tau \in (t_i, t_{i+1})$, $\frac{d}{d\tau} x_{k+1}$ for $k = 0, \dots, N-1$ is continuous, if F_1 and F_2 are continuously differentiable, which is a reasonable requirement on an explicit integration scheme. Now the case when τ coincides with a grid point, say t_{i+1} is studied. Therefore, we look at the left and right limits: while $\lim_{\substack{\tau \to t_{i+1} \\ \tau > t_{i+1}}} \frac{d}{d\tau} x_{i+1} = 0$, because switching happens afterwards, in general,

$$\begin{aligned} \lim_{\substack{\tau \to t_{i+1} \\ \tau < t_{i+1}}} \frac{d}{d\tau} x_{i+1} = \lim_{\substack{\tau \to t_{i+1} \\ \tau < t_{i+1}}} & D_1F_2(x^*, \tau, t_{i+1}) \cdot D_3F_1(x_i, t_i, \tau) \\ & + D_2F_2(x^*, \tau, t_{i+1}) \neq 0 \end{aligned} \tag{10}$$

and thus, $\frac{d}{d\tau} x_{i+1}$ is nondifferentiable for $\tau = t_{i+1}$. Although (10) has to be checked for each integration scheme and each system individually, most likely the nondifferentiability of x_k, $(k = i+1, \dots, N)$ at time points is existent for a system with arbitrary switching vector fields. As we saw in (4), $\frac{d}{d\tau} x_k$ is part of the discrete cost function derivative and thus, nondifferentiability of the discrete trajectory generally leads to nondifferentiability of J_d.

The iterative relation of the derivatives at neighboring trajectory points gives rise to a transition operator

$$\mathbf{\Phi}(k+1, k) := D_1F_2(x_k, t_k, t_{k+1}) \tag{11}$$

for $k \in \{i+1, \dots, N-1\}$. Further, we define $\mathbf{\Phi}(k,k) := 1$ and for $l > k+1$ $\mathbf{\Phi}(l,k) := \mathbf{\Phi}(l, l-1) \cdot \ldots \cdot \mathbf{\Phi}(k+2, k+1) \cdot \mathbf{\Phi}(k+1, k)$. Thus, for $k \in \{i+1, \dots, N-1\}$ one receives the propagation scheme

$$\frac{d}{d\tau} x_{k+1} = \mathbf{\Phi}(k+1, i+1) \cdot \frac{d}{d\tau} x_{i+1}.$$

Example 3.1 (Explicit Euler): Now we investigate one specific integration scheme, i.e. an explicit Euler approximation to illustrate the general result of the nondifferentiability of x_k w.r.t. τ as stated above. The explicit Euler scheme for a switched system is defined for $k \in \{0, \dots, N-1\}$ as $F_j(x_k, t_k, t_{k+1}) = x_k + (t_{k+1} - t_k) \cdot f_j(x_k)$, $j = \{1, 2\}$ and on the switching interval with x^* and τ in the appropriate arguments. Thus, by using the partial derivatives $D_3F_1(x_k, t_k, t_{k+1}) = f_1(x_k)$, $D_1F_2(x_k, t_k, t_{k+1}) = 1 + \frac{\partial}{\partial x} f_2(x_k)(t_{k+1} - t_k)$ and $D_2F_2(x_k, t_k, t_{k+1}) = -f_2(x_k)$, we verifiy that

$$\frac{d}{d\tau} x_{i+1} = f_1(x_i) + \frac{\partial}{\partial x} f_2(x^*) \cdot (t_{i+1} - \tau) \cdot f_1(x_i) - f_2(x^*)$$

with $x^* = x_i + f_1(x_i) \cdot (\tau - t_i)$. The left hand side limit for $\tau = t_{i+1}$ is

$$\lim_{\substack{\tau \to t_{i+1} \\ \tau < t_{i+1}}} \frac{d}{d\tau} x_{i+1} = f_1(x_i) - f_2(x_{i+1}).$$

In general, $f_1(x_i)$ and $f_2(x_{i+1})$ will not coincide. Then $\{x_k\}_{k=0}^N$ is nondifferentiable at $\tau = t_{i+1}$. For $dt = t_{i+1} - t_i \to 0$, it holds $x^* = x_i = x_{i+1}$ and thus, the result matches the continuous case (cf. (7)).

For the next node, x_{i+2} we get the following derivative and limits $\frac{d}{d\tau} x_{i+2} = \mathbf{\Phi}(i+2, i+1) \frac{d}{d\tau} x_{i+1} = (1 + (t_{i+2} - t_{i+1}) \frac{\partial}{\partial x} f_2(x_{i+1})) \frac{d}{d\tau} x_{i+1}$, and hence $\lim_{\substack{\tau \to t_{i+1} \\ \tau < t_{i+1}}} \frac{d}{d\tau} x_{i+2} = \left(1 + (t_{i+2} - t_{i+1}) \frac{\partial}{\partial x} f_2(x_{i+1})\right) \cdot (f_1(x_i) - f_2(x_{i+1}))$, but $\lim_{\substack{\tau \to t_{i+1} \\ \tau > t_{i+1}}} \frac{d}{d\tau} x_{i+2} = \left(1 + (t_{i+2} - t_{i+1}) \frac{\partial}{\partial x} f_2(x_{i+1})\right) \cdot f_1(x_{i+1}) - f_2(x_{i+2})$, where the second limit is received by the second case of (8) for an index shifted by one. Again, only for vanishing time steps these limits will coincide.

Remark 3.1: To sum up, $\{x_k\}_{k=0}^N$ generated by an arbitrary one-step explicit integration scheme is not guaranteed to be differentiable for $\tau \in \{t_0, \dots, t_N\}$, but everywhere else. This is consistent with the continuous setting described in Section II, because we can also interpret the approximated trajectory, e.g. from an explicit Euler scheme as a piecewise linear function given by

$$x(t) = \begin{cases} x_k + f_1(x_k)(t - t_k) & \text{if } k < i, \text{ and } t_k \le t \le t_{k+1}, \\ x_i + f_1(x_i)(t - t_i) & \text{if } k = i \text{ and } t_i \le t \le \tau, \\ x^* + f_2(x^*)(t - \tau) & \text{if } k = i \text{ and } \tau \le t \le t_{i+1}, \\ x_k + f_2(x_k)(t - t_k) & \text{if } k > i \text{ and } t_k \le t \le t_{k+1}. \end{cases}$$

This can be seen as the hybrid trajectory of a switched linear system with switching points $t_0, \dots, t_N$ and τ. For disjoint switching points, the theory presented in Section II can be applied. However, if two switching points coincide, i.e. $\tau = t_{i+1}$ for some i as studied above, x is not differentiable there. This is in correspondence with [3], in which the nonexistence of a gradient in case of coinciding switching points is shown.

B. Implicit One-step Integration Schemes

When using an implicit integration scheme instead of an explicit, (8) of Problem 3.1 is replaced by

$$\mathbf{G} = \begin{cases} G_1(x_k, x_{k+1}, t_k, t_{k+1}) = 0 & \text{for } k = 0, \dots, i-1, \\ G_1(x_k, x^*, t_k, \tau) = 0 & \text{and} \\ G_2(x^*, x_{k+1}, \tau, t_{k+1}) = 0 & \text{for } k = i, \\ G_2(x_k, x_{k+1}, t_k, t_{k+1}) = 0 & \text{for } k = i+1, \dots, N-1. \end{cases} \tag{12}$$

By computations similar to those in Section III-A, we derive for $\tau \in (t_i, t_{i+1})$

$$\frac{d}{d\tau} x_{i+1} = - D_2 G_2(x^*, x_{i+1}, \tau, t_{i+1})^{-1} \cdot (D_1 G_2(x^*, x_{i+1}, \tau, t_{i+1}) \cdot \frac{d}{d\tau} x^* + D_3 G_2(x^*, x_{i+1}, \tau, t_{i+1}))$$

with $\frac{d}{d\tau} x^* = -D_2 G_1(x_i, x^*, t_i, \tau)^{-1} D_4 G_1(x_i, x^*, t_i, \tau)$. Defining the discrete transition operator as $\mathbf{\Phi}(k+1, k) := -D_2 G_2(x_k, x_{k+1}, t_k, t_{k+1})^{-1} \cdot D_1 G_2(x_k, x_{k+1}, t_k, t_{k+1})$ for $k \in \{i+1, \dots, N-1\}$, the propagation rule can be again written as $\frac{d}{d\tau} x_{k+1} = \mathbf{\Phi}(k+1, i+1) \cdot \frac{d}{d\tau} x_{i+1}$ for $k = i+1, \dots, N-1$.

In this section, we showed that the STO problem in discrete time is in general not differentiable everywhere. The points at which the objective function is generally nonsmooth are the time grid points. However, in between neighboring time points, the discrete objective function inherits the smoothness of the corresponding continuous problem. Therefore, the derivative of the cost function can be computed using discrete adjoint equations analogously to the continuous case. At nondifferentiable points the left and right derivatives can be computed using the same adjoint equation, as we will show in the following section.

IV. DISCRETE ADJOINTS

Optimal solutions of continuous optimal control problems (under appropriate regularity assumptions) satisfy first-order optimality conditions, the well known minimum principle (see e.g. [9]). A discretization leads to a nonlinear constraint optimization problem, where the constraint on (t_i, t_{i+1}) is dependent on τ. Here, necessary optimality conditions are called Kuhn-Tucker equations and give rise to discrete adjoint multipliers (cf. [9]). The optimality conditions can be either formulated in terms of a Hamiltonian or a Lagrangian function, we choose the latter for our problem settings. The aim is to compute $J_d{}'(\tau)$ in terms of the discrete adjoints (cf. (6) for the time continuous case).

A. Discrete Adjoints for Explicit Schemes

Definition 4.1 (Discrete Lagrangian): The discrete Lagrangian of Problem 3.1 is given by

$$\begin{aligned} &\mathcal{L}_d(\{x_k\}_{k=0}^N, \{\rho\}_{k=0}^N, \tau, x^*, \rho^*) \\ &= \sum_{k=0}^{N} \Psi_k(x_k) - \sum_{k=0}^{i-1} \rho_{k+1}(x_{k+1} - F_1(x_k, t_k, t_{k+1})) \\ &\quad - \rho_0 \cdot (x_0 - x_{ini}) - \rho^* \cdot (x^* - F_1(x_i, t_i, \tau)) \\ &\quad - \rho_{i+1} \cdot (x_{i+1} - F_2(x^*, \tau, t_{i+1})) \\ &\quad - \sum_{k=i+1}^{N-1} \rho_{k+1} \cdot (x_{k+1} - F_2(x_k, t_k, t_{k+1})) \end{aligned}$$

with the discrete adjoints $\{\rho\}_{k=0}^N$ and ρ^* (cf. Fig. 2). Note that the discrete adjoints are treated as row vectors in contrast to the continuous formulation in Section II.

Theorem 4.1 (Discrete adjoints for explicit schemes): The backwards difference equations defining the discrete adjoints for an explicit integration scheme analogously to the continuous case (cf. Section II) are given by

$$\begin{aligned} \rho_N &= D\Psi_N(x_N) \\ \rho_k &= D\Psi_k(x_k) + \rho_{k+1} \cdot D_1 F_2(x_k, t_k, t_{k+1}) \\ &\quad \text{for } k = N-1, \dots, i+2, \\ \rho_{i+1} &= D\Psi_{i+1}(x_{i+1}) + \rho_{i+2} D_1 F_2(x_{i+1}, t_{i+1}, t_{i+2}) \\ \rho^* &= \rho_{i+1} \cdot D_1 F_2(x^*, \tau, t_{i+1}) \\ \rho_i &= D\Psi_i(x_i) + \rho^* D_1 F_1(x_i, t_i, \tau) \\ \rho_k &= D\Psi_k(x_k) + \rho_{k+1} \cdot D_1 F_1(x_k, t_k, t_{k+1}) \\ &\quad \text{for } k = i-1, \dots, 0. \end{aligned}$$

Proof: Taking variations w.r.t. x_k, ρ_k, x^*, ρ^* and τ leads to the necessary optimality conditions, i.e. the discrete equations of motions, the boundary condition $x_0 = x(0) = x_{ini}$ and also the discrete adjoint equations as given above. It further holds $\rho^* \cdot D_3 F_1(x_i, t_i, \tau) + \rho_{i+1} \cdot D_2 F_2(x^*, \tau, t_{i+1}) = 0$ which defines τ. ■

These adoint equations are consistent with the system given in [9]. Using the operator $\mathbf{\Phi}(k+1, k)$ from (11), the difference equation can be written as

$$\rho_k = D\Psi_k(x_k) + \rho_{k+1} \cdot \mathbf{\Phi}(k+1, k)$$

for $k = N, \dots, i+1$, with boundary value $\rho_N = D\Psi_N(x_N)$, or alternatively, $\rho_k = \sum_{j=k}^N D\Psi_j(x_j) \cdot \Phi(j, k)$, where ρ_{i+1} is the last adjoint before switching (looking backwards in time). Thus, the adjoints are continuous w.r.t. τ, if the $D\Psi_k$ and the transition operator are continuous, which is reasonable to assume.

This provides an elegant way to write the discrete cost function derivative (cf. (4))

$$\begin{aligned} J_d{}'(\tau) &= \sum_{k=0}^{N} D\Psi_k(x_k) \frac{d}{d\tau} x_k \\ &= \sum_{k=i+1}^{N} D\Psi_k(x_k) \cdot \Phi(k, i+1) \cdot \frac{d}{d\tau} x_{i+1} = \rho_{i+1} \cdot \frac{d}{d\tau} x_{i+1}. \end{aligned}$$

So it can be nicely seen that although the adjoint itself is continuous, its argument, i.e. $\frac{d}{d\tau}x_{i+1}$ leads to nondifferentiability of J_d. In fact, if $\tau = t_i$ for $i = 0, \ldots, N$, $\frac{d}{d\tau}x_{i+1}$ and therefore ${J_d}'$ can only be defined by either the left or the right limit as defined in (10).

Example 4.1 (Adjoints for explicit Euler): For a specific integration scheme, the formula for the discrete adjoints can be explicitly computed and analyzed. Here, we consider again the explicit Euler as an example. Recall that in the explicit Euler scheme (cf. Example 3.1), it holds $\mathbf{\Phi}(k+1,k) = D_1F_2(x_k, t_k, t_{k+1}) = 1 + (t_{k+1} - t_k)\frac{\partial}{\partial x}f_2(x_k)$ for $k = i+1, \ldots, N-1$ with f_2 the active vector field after switch. If we plug this in the adjoint equation, we get

$$\begin{aligned}\rho_k &= D\Psi_k(x_k) + \rho_{k+1}\left(1 + (t_{k+1} - t_k)\frac{\partial}{\partial x}f_2(x_k)\right)\\ &= \rho_{k+1} + \left(\frac{D\Psi_k(x_k)}{t_{k+1} - t_k} + \rho_{k+1}\frac{\partial}{\partial x}f_2(x_k)\right)\cdot(t_{k+1} - t_k).\end{aligned} \tag{13}$$

For choosing $\Psi(x_k) = (t_{k+1} - t_k)\cdot\ell(x_k)$, (13) is a direct discretization of the continuous formulation in (5). The resulting adjoint scheme itself is explicit, since we are going backwards in time. However, because the computation of ρ_k explicitly depends on x_k, the discrete scheme for the system of equations consisting of (1) and (5) is a symplectic or semi implicit Euler scheme (cf. [9], where general Runge-Kutta, but non-hybrid schemes are studied).

Example 4.2 (Switched linear system): We compare the analytic solutions of the commonly used continuous setting to the results we received for the discrete time setting. Therefore, consider the following simple one-dimensional linear switched system

$$\dot{x} = \begin{cases} x & t \leq \tau \\ 2x & t > \tau \end{cases}$$

with linear vector fields $f_1(x) = x$, $f_2 = 2x$, $x(0) = 10$ and switching time τ. The corresponding flow, i.e. the solution of the switched differential equation is hence given by

$$x(t,\tau) = \begin{cases} x_0\exp(t) & t \leq \tau \\ x(\tau)\exp(2(t-\tau)) & t > \tau. \end{cases}$$

The cost function to be minimized is chosen to be $J(\tau) = \int_0^T x^2(t,\tau)\,dt$ and depends on τ through the hybrid trajectory $x(t,\tau)$. The derivative of $x(t,\tau)$ w.r.t. τ equals $\frac{d}{d\tau}x(t) = -x_0\cdot\exp(2t-\tau)$ for $t \geq \tau$ with $f_1(x(\tau)) - f_2(x(\tau)) = -x_0\exp(\tau)$ and $\Phi(t,\tau) = \exp(2(t-\tau))$ (cf. Section II). Further, the analytic solution of the adjoint equation is given by $\rho(t) = -\frac{1}{2}x(\tau)\exp(2t-2\tau) + \frac{1}{2}x(\tau)\exp(4T-2\tau-2t)$. Thus, $J'(\tau)$ can be exactly determined by equation (6). In general applications, analytical solutions cannot be found and one therefore has to approximate $x(t,\tau)$ as well as $\rho(t)$ and also the evaluation of the cost function integral. For this example, we approximate $x(t,\tau)$ by an explicit Euler scheme with $x_0 = x(0)$ and for $k = 1, \ldots, N$

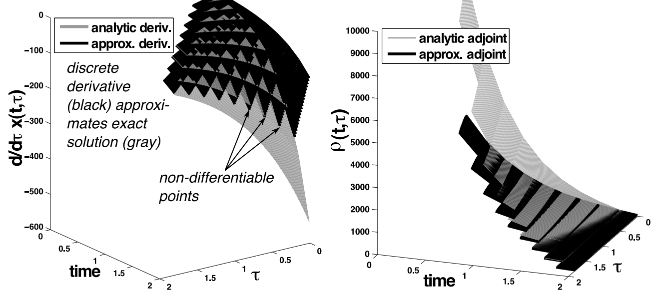

Fig. 3. *Left:* The derivative of the approximated trajectory $\frac{d}{d\tau}x_k$ (black) on the discrete time grid $\{t_k\}_{k=0}^N$ with $t_k = k\cdot 0.2$ and $\tau \in (0,2)$ approximates the analytically computed exact solution (gray). Non-differentiable points occur when $\tau \in \{t_0, \ldots, t_N\}$. *Right:* The corresponding discrete adjoints (black) are continuous w.r.t. τ and they approximate the adjoints (gray) from the analytic solution of the continuous problem.

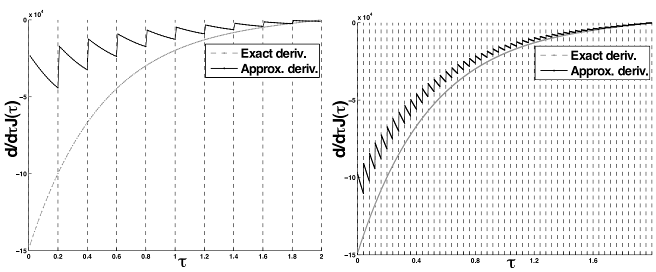

Fig. 4. *Left:* The discrete cost function derivative (black) shows non-differentiability at discrete time points. The dashed lines illustrate the discrete time grid (same as in Fig. 3, $dt = 0.2$). *Right:* Reducing the grid width to $t_k = k\cdot 0.04$, the jumps at the nondifferentiable points of the discrete derivative ${J_d}'(\tau) = \frac{d}{d\tau}J_d$ get smaller and the discrete derivative approaches the continuous solution.

$$x_{k+1} = \begin{cases} x_k + f_1(x_k)(t_{k+1} - t_k) & \text{if } k+1 < i \\ x_i + f_1(x_i)(\tau - t_i) & \\ \;+(t_{i+1} - \tau)\cdot f_2(x^*) & \text{if } k = i, \\ x_k + f_2(x_k)(t_{k+1} - t_k) & \text{if } k+1 > i, \end{cases}$$

with the approximated switching state $x^* = x_i + f_1(x_i)(\tau - t_i)$ and the switching interval $[t_i, t_{i+1}]$ as depicted in Fig. 2. The trapezoidal rule for a quadrature of the cost function is chosen, i.e. $J(\tau) \approx \sum_{k=0}^N \Psi(x_k) = \sum_{k=1}^{N-1}\ell(x_k)\cdot\frac{t_{k+1}-t_{k-1}}{2} + \ell(x_0)\cdot\frac{t_1 - t_0}{2} + \ell(x_N)\cdot\frac{t_N - t_{N-1}}{2}$ (cf. (3)).

In Fig. 3 (left) we compare $\frac{d}{d\tau}\{x_k\}_{k=0}^N$ to the exact values of $\frac{d}{d\tau}x(t)$ evaluated on $\{t_k\}_{k=0}^N$. It can be observed that the derivative of the approximation is not well defined if $\tau = t_k$ for $k \in \{1, \ldots, N-1\}$ (time grid marked as dashed lines) since left and right hand side limits are not equal (cf. Section III-A). The discrete adjoints (see Fig. 3 (right), cf. Section IV for their definition) are continuous w.r.t. τ. Thus, since ${J_d}'(\tau) = \sum_{k=0}^N D\Psi(x_k)\cdot\frac{d}{d\tau}x_k$ (cf. (4)), $J_d(\tau)$ is nondifferentiable for $\tau \in \{t_k\}_{k=0}^N$ (cf. Fig. 4) as the results of Section III state. However, this nondifferentiability vanishes when the grid width tends to zero, as Fig. 4 (right) illustrates. Here, the step size is reduced from $dt = 0.2$ to $dt = 0.04$.

B. Discrete Adjoints for Implicit Schemes

For an implicit integration scheme as in (12), a Lagrangian can be defined and adjoints can be derived analogously to the explicit scheme (details have to be postponed to a future publication). Although the adjoints are continuous under normal smoothness conditions, $\frac{d}{d\tau} x_{i+1}$ (cf. Section III-B) may generally be not well defined on time grid points, as in explicit schemes. Thus, for the cost function derivative the same problem of nondifferentiability occurs.

V. NUMERICAL EXAMPLE: THE HYBRID LOCKED DOUBLE PENDULUM

As an illustrating example for nondifferentiable points of a cost function for discretized switched systems, we consider the double pendulum. The model consists of two mass points m_1, m_2 on massless rods of length l_1, l_2. The motion of the pendula are described by two angles, φ_1 and φ_2 (cf. Fig. 1). The standard double pendulum is turned into a hybrid system by introducing two different modes: **M1**: The outer pendulum is locked w.r.t. the inner pendulum with angle θ, i.e. the system behaves like a single pendulum with a special inertia tensor. **M2**: Both pendula can move freely as in the standard case.

In M1, the following energy terms are valid

$$K_1(\varphi_1, \dot{\varphi}_1) = \frac{1}{2}(m_1 l_1^2 + m_2 r^2) \cdot \dot{\varphi}_1^2$$
$$V_1(\varphi_1) = (m_1 + m_2) g l_1 \cos(\varphi_1) + m_2 g l_2 \cos(\varphi_1 + \theta - \pi)$$

with distance of outer mass to origin $r^2 = l_1^2 + l_2^2 - 2 l_1 l_2 \cos(\theta)$. The position of the outer mass can be updated according to $\varphi_2 = \varphi_1 + \theta - \pi$ and it naturally follows that $\dot{\varphi}_1 = \dot{\varphi}_2$. In M2, the system is defined by

$$K_2(\varphi_1, \varphi_2, \dot{\varphi}_1, \dot{\varphi}_2) = \frac{1}{2} \begin{pmatrix} \dot{\varphi}_1 \\ \dot{\varphi}_2 \end{pmatrix}^T \cdot$$
$$\begin{pmatrix} (m_1 + m_2) l_1^2 & m_2 l_1 l_2 \cos(\varphi_1 - \varphi_2) \\ m_2 l_1 l_2 \cos(\varphi_1 - \varphi_2) & m_2 l_2^2 \end{pmatrix} \cdot \begin{pmatrix} \dot{\varphi}_1 \\ \dot{\varphi}_2 \end{pmatrix}$$
$$V_2(\varphi_1, \varphi_2) = m_1 g l_1 \cos(\varphi_1) + m_2 g (l_1 \cos(\varphi_1) + l_2 \cos(\varphi_2)).$$

In both cases, the equations of motion are derived by the Euler-Lagrange equations $\frac{d}{dt}\frac{\partial L_i}{\partial \dot{q}} - \frac{\partial L_i}{\partial q} = 0$ for $L_i(q, \dot{q}) = K_i(q, \dot{q}) - V_i(q, \dot{q})$ $(i = 1, 2)$ with $q = (\varphi_1, \varphi_2)$ being the configurations and $\dot{q} = (\dot{\varphi}_1, \dot{\varphi}_2)$ the corresponding velocities. We focus on the scenario, when the system switches a single time from M1 to M2. One can check that the energies of M1 and M2 coincides in a switching point $x_\tau = (\varphi_1, \varphi_1 + \theta - \pi, \dot{\varphi}_1, \dot{\varphi}_1)$ and thus we will have energy conservation along the entire hybrid trajectory. We assume that the velocities directly before and after the switch are the same, i.e. $\dot{\varphi}_1^- = \dot{\varphi}_1^+ = \dot{\varphi}_2^+$. As a cost function we choose $J(\tau) = \Psi(x(T)) = \left\| \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} - q_{\text{final}} \right\|^2$ to minimize the distance to a given final point. This is an algebraic cost function as considered in Problem 1.2. The final point $q_{\text{final}} = (-1.5487, -1.9733)$ is chosen such that the optimal value is $\tau^* = 0.33$. We approximate the switching time derivative $J_d{}'(\tau)$ by evaluating the corresponding formula for $\frac{d}{d\tau} x_{i+1}$ and the appropriate discrete adjoints. In Fig. 1 (right) the nondifferentiable points of $J_d{}'(\tau)$, i.e. points in which the left hand right hand side derivatives do not coincide, can be clearly seen.

VI. CONCLUSION AND FUTURE WORK

In this contribution, we show that in contrast to time continuous STO, in discretized problems the differentiability of a cost function w.r.t. the switching time is not guaranteed if the switching time matches grid points of the time grid. Consequently, smaller time steps actually make the neighborhood in which a smooth optimality condition may be used smaller. This indicates the need for application of optimality conditions appropriate for nonsmooth systems.

So far, we restrict our numerical test to implicit and explicit Euler methods. Thus it is straight forward to extend to other integration methods, e.g. multi step methods, higher order Runge Kutta schemes or geometric integrators. The latter are of special importance for structure preserving integration, e.g. energy or momentum preservation in mechanical systems (cf. [10]). Therefore, our future work will focus on STO of discrete hybrid mechanics (cf. [12] for a discrete variational modeling of hybrid mechanical systems). From a numerical point of view, the information of nondifferentiable points should be used to improve STO algorithms by providing e.g. subgradients in those cases. Applications to more complex examples will then be of great interest.

REFERENCES

[1] M. Egerstedt, Y. Wardi, and F. Delmotte, "Optimal control of switching times in switched dynamical systems," in *IEEE Conference on Decision and Control*, 2003, pp. 2138–2143.

[2] M. Alamir and S. A. Attia, "On solving optimal control problems for switched hybrid nonlinear systems by strong variations algorithms," in *Proceedings of 6th IFAC Symposium on Nonlinear Control Systems*, 2004, pp. 558–563.

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

[4] A. Schild, X. C. Ding, M. Egerstedt, and J. Lunze, "Design of optimal switching surfaces for switched autonomous systems," in *IEEE Conference on Decision and Control*, 2009, pp. 5293–5298.

[5] X. Xu and P. Antsaklis, "Optimal control of switched autonomous systems," in *IEEE Conference on Decision and Control*, 2002, pp. 4401–4406.

[6] ——, "Optimal control of switched systems based on parameterization of the switching instants," *IEEE Transactions on Automatic Control*, vol. 49, no. 1, pp. 2–16, 2004.

[7] T. Caldwell and T. Murphey, "Switching mode generation and optimal estimation with application to skid-steering," *Automatica*, vol. 47, no. 1, pp. 50–64, 2011.

[8] E. Johnson and T. Murphey, "Second-order switching time optimization for nonlinear time-varying dynamic systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 8, pp. 1953–1957, 2011.

[9] W. W. Hager, "Runge-Kutta methods in optimal control and the transformed adjoint system," *Numerische Mathematik*, vol. 87, pp. 247–282, 1999.

[10] J. E. Marsden and M. West, "Discrete mechanics and variational integrators," *Acta Numerica*, vol. 10, pp. 357–514, 2001.

[11] S. Ober-Blöbaum, O. Junge, and J. E. Marsden, "Discrete mechanics and optimal control: an analysis," *Control, Optimisation and Calculus of Variations*, vol. 17, no. 2, pp. 322–352, 2011.

[12] K. Flaßkamp and S. Ober-Blöbaum, "Variational formulation and optimal control of hybrid Lagrangian systems," in *Proceedings of the 14th International Conference on Hybrid Systems: Computation and Control*, 2011, pp. 241–250.