Linear Time-Varying Impulse Optimization for Data Association

Matthew Travers[†], Todd Murphey[*], and Lucy Pao[⋆]

***Abstract*— This paper extends recent work in linear switching-time optimization to a class of impulsive systems. It is shown that for cases with and without reference the first derivative of a cost function over an arbitrary number of impulse times can be calculated by computing two or three total integrals respectively, which can be done offline. This makes it possible to numerically optimize a cost function online algebraically. A simulated example drawn from data association to which impulse optimization methods are applied is provided.**

I. Introduction

Switching-time optimization has been the subject of numerous recent works[1], [2], [4], [5], [9], [10], [11], [12], [13], [17]. The goal of switching-time optimization is to optimize a cost function with respect to the finite number of times at which the dynamics switch. The systems considered typically have a known sequence of switched dynamics ([1], [2], [4], [6], [13] focus on optimizing over the sequence order as well), where the resulting system trajectory is C^0.

Considering only first-order optimizations, the authors in [5], [10], [11], [12], [13] present an algorithm which reduces the calculation of the first-derivative of a cost function with respect to an arbitrary number of switching times to a single integration at each step of the numerical descent algorithm. The authors in [7] show that a further reduction in the number of total calculations is possible by considering only systems with linear dynamics.

This paper presents an extension of the previous work in [14], [15], which considers optimizing over an arbitrary number of impulse times for both linear and nonlinear systems. The main contribution of this paper is the development of the algorithm in Section III, which shows that by restricting the scope of the systems considered to those with linear dynamics, it is possible to optimize over an arbitrary number of impulse times by computing a single integral for the entire optimization. Previous work in [14], [15] requires a single integral be computed for each step of the optimization algorithm. This paper presents a direct extension of the work in [7] to a class of impulsive systems.

The rest of this paper is organized as follows: Section II analytically defines the type of impulsive systems considered in this work. Section III provides the derivation of the linear time-varying impulse optimization algorithm, considering both systems with and without reference signals. Section IV presents an example system for which the impulse optimization described in Section III solves a data association problem. Finally, section V gives conclusions and directions for future work.

[†]Department of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

[*]Department of Mechanical Engineering, Northwestern University, Evanston, Illinois 60208

[⋆]Department of Electrical, Computer, and Energy Engineering, University of Colorado, Boulder, Colorado 80309

II. Impulsive Systems

This paper considers linear-time varying (LTV) systems which experience a known number of impulses. In general, the assumption of a known number of impulses does not have to be made as has been shown in previous work [16]; it is possible to estimate the number of impulses based on a trajectory optimization procedure. In this work it is assumed that the dynamics of the system in between the times at which impulses occur are strictly linear, i.e., assuming that the times at which impulses occur belong to the set $\mathcal{T} = \{T_1, T_2, \ldots, T_{N-1}\}$

$$\dot{x}(t) = A(t)x(t), \quad x(T_i) = x_{T_i}, \quad T_i \leq t < T_{i+1}. \tag{1}$$

The solution to the linear dynamics in between impulses can be written

$$x(t) = \Phi(t, T_i)x(T_i), \quad T_i \leq t < T_{i+1}, \tag{2}$$

where $\Phi(t, T_i)$ is a state transition matrix defined to be the solution to $\dot{\Phi}(t, T_i) = A(t)\Phi(t, T_i)$, $\Phi(t, t) = I$. Equations (1) and (2) assume that the dynamics in between impulse times are always $A(t)$. This is a reasonable assumption for the example presented Section IV. Extending the results presented in this work to systems that experience switched dynamics at the times which impulses occur is a straightforward direction of future work, as discussed in Section V.

Lemma 1 The state $x(\cdot)$ at time t, where $T_i \leq t < T_{i+1}$ can be written

$$\begin{array}{ll} x(t) = \Phi(t, T_0)x_0 + \sum_{i=1}^{j-1} \Phi(t, T_i)\delta_i, & T_1 \leq t < T_j \\ x(t) = \Phi(t, T_0)x_0, & t < T_1. \end{array} \tag{3}$$

Pf: Using the properties of the state transition matrix and assuming that the impulse magnitudes are finite, it is possible to write

$$\begin{aligned} x(t) = \Phi(t, T_{j-1}) \left[\Phi(T_{j-1}, T_{j-2}) \left[\ldots \left[\Phi(T_1, T_0)x_0 \right.\right.\right. \\ \left.\left.\left. +\delta_1\right] \ldots\right] + \delta_{j-1}\right], \quad T_{j-1} \leq t < T_j. \end{aligned} \tag{4}$$

Equation (3) is easily obtained by expanding (4) using the properties of the state transition matrix. The second part of (3) is obtained directly from the definition of the state transition matrix. ■

III. Optimality Conditions for Impulsive Systems with Quadratic Cost

A. No Reference Signal

The cost functional is defined to be

$$J(\cdot) = \int_0^{t_f} \ell(s, x(s))ds + M(x(T_n)), \tag{5}$$

where $\ell(s, x(s))$ is quadratic in the state $x(s)$. In this subsection it is assumed that $\ell(s, x(s)) = \frac{1}{2}x^T(s)Q(s)x(s)$ and $M(x(T_N)) = \frac{1}{2}x^T(T_N)P_1x(T_N)$. In the rest of this paper it is assumed, for simplicity, that the impulse magnitudes are known. A method for simultaneously optimizing over the impulse times and magnitudes was presented in prior work [15]. In order to optimize over the times at which impulses occur in the system, the following optimization problem is defined

$$\min_{\mathcal{T}} J(\cdot) \quad \text{s.t. } \dot{x}(t) = A(t)x(t). \tag{6}$$

It was shown in previous work [14] that the first-order optimality condition of the cost functional $J(\cdot)$ with respect to the impulse times can be written in terms of the co-state $\rho(t)$

$$D_{T_i}J(\cdot) \circ \partial T_i = \rho(t, T_i) \circ X^i + \ell(x(T_i^-), T_i^-) - \ell(x(T_i^+), T_i^+), \tag{7}$$

where $X^i = (A(T_i^-)x(T_i^-) - A(T_i^+)x(T_i^+))$ is a term which comes from taking the time derivative of the trajectory $x(t)$. The second and third terms on the right hand side of (7) come from applying Leibniz's rule at the impulse time T_i. The co-state can be shown [14] to be equal to

$$\dot{\rho}(t) = -A^T(t)\rho(t) - Q(t)x(t), \quad \rho(T_N) = P_1x(T_N). \tag{8}$$

The following theorem states the relationship between the state and costate.

Theorem 1 For LTV impulsive systems, the state $x(t)$ and co-state $\rho(t)$ are related by the equation

$$\rho(t) = P(t)x(t) + \sum_{i=j}^{N-1} \Phi^T(T_i, t)P(T_i)\delta_i, \ T_{j-1} \le t < T_j,$$
$$\forall j \in \{2, 3, \ldots, N-1\}, \text{ and}$$
$$\rho(t) = P(t)x(t), \text{ for } T_{N-1} \le t$$

where $P(t)$ is the solution to the following matrix valued equation

$$\dot{P}(t) = -A^T(t)P(t) - P(t)A(t) - Q(t), \quad P(T_N) = P_1. \tag{9}$$

Pf: Using the fundamental theorem of calculus, the solution to (8) can be written

$$\rho(t) = \Phi^T(T_N, t)P_1x(T_N) + \int_t^{T_N} \Phi^T(s, t)Q(s)x(s)ds. \tag{10}$$

Consider $T_{j-1} \le t < T_j$, where $j < N-1$. Plugging in (3),

$$\begin{aligned}
\rho(t) &= \Phi^T(T_N, t)P_1\left(\Phi(T_N, T_0)x_0 + \sum_{i=1}^{N-1}\Phi(T_N, T_i)\delta_i\right) \\
&+ \int_{T_{N-1}^+}^{T_N} \Phi^T(s, t)Q(s)\left(\Phi(s, T_0)x_0 + \sum_{i=1}^{N-1}\Phi(s, T_i)\delta_i\right)ds \\
&+ \int_{T_{N-2}^+}^{T_{N-1}^-} \Phi^T(s, t)Q(s)\left(\Phi(s, T_0)x_0 + \sum_{i=1}^{N-2}\Phi(s, T_i)\delta_i\right)ds \\
&+ \ldots \\
&+ \int_t^{T_j^-} \Phi^T(s, t)Q(s)\left(\Phi(s, T_0)x_0 + \sum_{i=1}^{j-1}\Phi(s, T_i)\delta_i\right)ds.
\end{aligned}$$

Using only the properties of state transition matrices, this can be rewritten as

$$\begin{aligned}
\rho(t) &= \Phi^T(T_N, t)P_1\Phi(T_N, t) \\
&\quad \circ \left(\Phi(t, T_0)x_0 + \sum_{i=1}^{N-1}\Phi(t, T_i)\delta_i\right) \\
&+ \int_{T_{N-1}^+}^{T_N} \Phi^T(s, t)Q(s)\Phi(s, t) \\
&\quad \circ \left(\Phi(t, T_0)x_0 + \sum_{i=1}^{N-1}\Phi(t, T_i)\delta_i\right)ds \\
&+ \int_{T_{N-2}^+}^{T_{N-1}^-} \Phi^T(s, t)Q(s)\Phi(s, t) \\
&\quad \circ \left(\Phi(t, T_0)x_0 + \sum_{i=1}^{N-2}\Phi(t, T_i)\delta_i\right)ds + \ldots \\
&+ \int_t^{T_j^-} \Phi^T(s, t)Q(s)\Phi(s, t)\Phi(t, T_0)x_0ds.
\end{aligned} \tag{11}$$

Regrouping and again using Equation (3), (11) can be written as

$$\begin{aligned}
\rho(t) &= \Phi^T(T_N, t)P_1\Phi(T_N, t) \cdot x(t) \\
&\quad + \int_t^{T_N} \Phi^T(s, t)Q(s)\Phi(s, t)ds \cdot x(t) \\
&\quad + \sum_{i=j}^{N-1}\Big[\Phi^T(T_N, t)P_1\Phi(T_N, T_i) \cdot \delta_i \\
&\qquad + \int_{T_j}^{T_N} \Phi^T(s, t)Q(s)\Phi(s, T_i)ds \cdot \delta_i\Big]
\end{aligned}$$

Again, using the properties of the state transition matrix

$$\rho(t) = \Bigg[\Phi^T(T_N,t)P_1\Phi(T_N,t) + \int_t^{T_N}\Phi^T(s,t)Q(s)\Phi(s,t)ds\Bigg]\cdot x(t) + \sum_{i=j}^{N-1}\Bigg[\Phi^T(T_i,t)\Bigg[\Phi^T(T_N,T_i)P_1\Phi(T_N,T_i) + \int_{T_i}^{T_N}\Phi^T(s,T_i)Q(s)\Phi(s,T_i)ds\Bigg]\cdot\delta_i\Bigg] \quad (12)$$

Define $P(t)$ to be

$$P(t) = \Phi^T(T_N,t)P_1\Phi(T_N,t) + \int_t^{T_N}\Phi^T(s,t)Q(s)\Phi(s,t)ds. \quad (13)$$

Taking the derivative of (13) with respect to t gives the second part of the theorem. To obtain the first part of the theorem, plug (13) into (12) to obtain

$$\rho(t) = P(t)x(t) + \sum_{i=j}^{N-1}\Phi^T(T_i,t)P(T_i)\delta_i, \quad T_1 \le t < T_j, \quad \forall j \in \{2,3,\ldots,N-1\}. \quad (14)$$

The second part of the theorem, i.e., the solution for $\rho(t)$ when $t \ge \tau_{N-1}$ is easy to see. When j in (14) is larger than $N-1$, the sum is equal to zero, thus $\rho(t) = P(t)x(t)$. ■

Notice that because $P(t)$ and $\Phi(T,t)$ only depend on the matrices $A(t)$ and $Q(t)$, both can be solved for offline, i.e., before beginning the optimization. This fact implies that once $P(t)$ and $\Phi(T,t)$ are calculated, the costate equation, (14), is completely determined by function evaluations and matrix multiplications. Noting that the terms in (7) are simply function evaluations, this makes it is possible to optimize over multiple impulse times and magnitudes as well as over multiple iterations of the descent calculating only two integrals. Notice also that when each of the impulse magnitudes $\delta_i = 0$, the standard relationship between the state and costate, $\rho(t) = P(t)x(t)$, is recovered.

B. Adding a Reference Signal

It is possible to incorporate a reference signal $x_d(t)$ into the cost $J(\cdot)$ by modifying the objective function $\ell(t,x(t))$ and terminal condition $M(x(T_N))$ such that

$$\ell(t,x(t)) = (x(t)-x_d(t))^T Q(t)(x(t)-x_d(t))$$
$$M(x(T_N)) = (x(T_N)-x_d(T_N))^T Q(t)(x(T_N)-x_d(T_N))$$

In order to determine the first derivative of a cost function which incorporates a reference signal, the costate equation $\rho(t)$ also needs to be modified,

$$\dot\rho(t) = -A(t)^T\rho(t) - Q(t)(x(t)-x_d(t)),$$
$$\rho(T_N) = P_1(x(T_N)-x_d(T_N)). \quad (15)$$

Theorem 2 For LTV systems that include a reference signal, the relationship between the state and co-state is

$$\rho(t) = P(t)x(t) + \sum_{i=j}^{N-1}\Phi^T(T_i,t)P(T_i)\delta_i - r(t), \quad T_1 \le t < T_j, \forall j \in \{2,3,\ldots,N-1\}, \quad (16)$$

such that $P(t)$ is the solution to the Riccati equation (9), and $r(t)$ is a linear mapping of the reference signal

$$r(t) = P_r(t)x_d(t),$$

where

$$\dot P_r(t) = -A^T(t)P_r(t) - P_r(t)A_r(t) - Q(t), \quad P_r(T_N) = P_1$$

and $A_r(t)$ satisfies

$$\dot x_d(t) = A_r(t)x_d(t).$$

Pf: Starting with equation (15) and applying the fundamental theorem of calculus

$$\rho(t) = \Phi(T_N,t)^T P_1(x(T_n)-x_d(T_N)) + \int_t^{T_n}\Phi(s,t)^T Q(s)(x(s)-x_d(s))ds. \quad (17)$$

Equation (17) can be written to produce

$$\rho(t) = \Phi(T_N,t)^T P_1 x(T_n) + \int_t^{T_n}\Phi(s,t)^T Q(s)x(s)ds - \Phi(T_N,t)^T P_1 x_d(T_N) - \int_t^{T_n}\Phi(s,t)^T Q(s)x_d(s)ds. \quad (18)$$

The first two terms on the right-hand side of (18) are the same as the right-hand side of (10), and are thus determined by Theorem 1. In Theorem 1, the relationship between the state and costate was found to be an algebraic relationship, once the operator $P(t)$ was determined. It is desirable to find a similar algebraic relationship between the state and costate when a reference is added. In order to accomplish this, the second two terms on the right-hand side of (18) need to be rewritten in a way such that they can be computed online algebraically.

Define the residual, $r(t)$, to be

$$r(t) = \Phi(T_N,t)^T P_1 x_d(T_N) + \int_t^{T_n}\Phi(s,t)^T Q(s)x_d(s)ds. \quad (19)$$

Noticing that equation (19) looks similar to the equation for $\rho(t)$ in Theorem 1, we would like to write $x_d(t)$ in terms of a state transition matrix operating on an initial condition. To define the state transition matrix that governs the time evolution of the reference signal, the assumption that $A_r(t)$ is known must be made, where

$$\dot x_d(t) = A_r(t)x_d(t), \quad x_d(t_0) = x_{0,d}. \quad (20)$$

Equation (20) is an LTV equation; its solution is known to have the form [8]

$$x_d(t) = \Phi_d(t, t_0)x_{0,d},$$

which can be used to rewrite (19) (using the properties of the state transition matrix)

$$\begin{aligned} r(t) &= \Phi(T_N, t)^T P_1 \Phi_d(T_n, t) x_d(t) \\ &\quad + \int_t^{T_n} \Phi(s,t)^T Q(s) \Phi_d(s,t) ds \cdot x_d(t) \\ &= \Bigg(\Phi(T_N, t)^T P_1 \Phi_d(T_n, t) \\ &\quad + \int_t^{T_n} \Phi(s,t)^T Q(s) \Phi_d(s,t) ds \Bigg) x_d(t) \\ &= P_r(t) x_d(t), \end{aligned} \tag{21}$$

where $P_r(t)$ is defined to be the solution to

$$\dot{P}_r(t) = -A^T(t)P_r(t) - P_r(t)A_r(t) - Q(t), \quad P_r(T_N) = P_1.$$

Thus, (18) can be rewritten as

$$\begin{aligned} \rho(t) &= P(t)x(t) + \sum_{i=j}^{N-1} \Phi^T(T_i, t) P(T_i) \delta_i - P_r(t) x_d(t), \\ &\quad T_1 \leq t < T_j, \quad \forall j \in \{2, 3, \ldots, N-1\} \\ &= P(t)x(t) + \sum_{i=j}^{N-1} \Phi^T(T_i, t) P(T_i) \delta_i - r(t), \\ &\quad T_1 \leq t < T_j, \quad \forall j \in \{2, 3, \ldots, N-1\}. \end{aligned}$$ ■

Although we will not typically know the exact form of $A_r(t)$, we can make the assumption that

$$\begin{aligned} A_r(t) = \text{diag} \Bigg(& \dot{x}_{\{d,1\}}(t)/x_{\{d,1\}}(t), \ldots, \\ & \dot{x}_{\{d,N\}}(t)/x_{\{d,N\}}(t) \Bigg), \end{aligned} \tag{22}$$

where N is the dimension of $x_d(t)$ and the notation diag() is meant to represent a diagonal matrix. It is possible to make this assumption if we choose a coordinate parameterization such that $x_{\{d,i\}}(t) \neq 0 \quad \forall i = 1, 2, \ldots N$.

Theorem 2 illustrates that it is possible to algebraically optimize online over an arbitrary number of impulse times when the cost function (5) contains a reference signal.

IV. Examples

The results derived in Section III are general in the sense that they are applicable to any non-hybrid impulsive system. In order to provide a direct comparison of the results presented in this work to those derived in previous work ([14], [15], [16]), an example addressing a data association problem is selected.

A. Example System

Consider a system which contains two airplanes both with dynamics described by

$$\dot{q}(t) = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \vartheta(t)\cos(\theta(t)) \\ \vartheta(t)\sin(\theta(t)) \\ \psi(t) \end{bmatrix} \tag{23}$$

where $\vartheta(t)$ and $\psi(t)$ are inputs. We assume that these dynamics can be linearized, and that they will in general not be known explicitly, i.e.,

$$\dot{q} = A(t)q(t) + w(t),$$

where $w(t)$ is a noise term sampled from a zero-mean normal distribution with covariance $Q(t)$. It is assumed that a single sensor makes continuous measurements of the form

$$y(t) = C(t)q(t) + v(t),$$

where $v(t)$ is a noise term sampled from a zero-mean normal distribution with covariance $R(t)$. The problem to be solved is: how can clear associations on the origin of measurements at certain times be made when the level of noise contained in the measurement is relatively high with respect to the distance between the two airplanes?

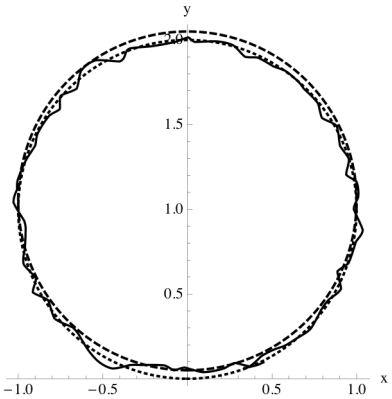

Fig. 1: The dotted and dashed lines again represent circular trajectories of two separate airplanes. The solid black line represents a nondeterministic measurement signal.

In Figure 1 the dotted and dashed lines represent an example of trajectories for the two airplane example system with constant inputs. The solid black line represents a nondeterministic "measurement signal." Figure 1 visually displays the typical measurement noise level to airplane separation distance considered in this work.

It is possible to associate measurements in the two airplane example by first assuming that from the perspective of the sensor, only a single object with a single trajectory exists. This is a reasonable assumption due to the fact that it is assumed that the sensor only makes a single measurement at each time. The single trajectory experiences impulses with respect to its spatial position at a number of different times. In the actual system containing two airplanes, the times at which impulses occur correspond to the sensor

switching between measuring the different airplanes. Using the measurement signal as a reference, the results of Section III-B can be applied to find the optimal set of times at which impulses occur in the single model trajectory. With a priori knowledge of which airplane is initially being measured, determining the times at which the sensor switches between measuring the trajectories of the two airplanes is equivalent to associating measurements to the airplane from which they originate.

B. Convergence Results

This section provides results of applying the LTV impulsive data association algorithm to the two-airplane example described in Section IV-A. Both deterministic as well as nondeterministic results are given.

Figure 2 shows first-order convergence results for the deterministic two-airplane system, optimizing over six impulse times. The horizontal axis is the iteration number, and the vertical axis the first derivative of the cost, plotted on a $\log$ scale. Figure 2 shows near perfect linear convergence.

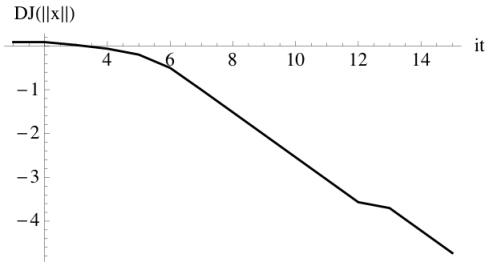

Fig. 2: Linear convergence results for the deterministic two-airplane example, optimizing over six impulse times.

The purpose of providing Figure 2 is to provide a basis of comparison for the nondeterministic results shown in Figures 3 and 4. Figure 3(a) shows the y-component of the measurement signal for a particular noise level in the two-airplane example (due to the initial conditions the impulses occur in the y-direction, all process noise is sampled from a zero mean normal distribution with standard deviation 0.05, and all measurement noise is sampled from a zero mean normal distribution whose standard deviation is 30% of the average separation distance of the two airplanes). The gray line represents the non-deterministic measurement signal, and the black line denotes the portions of the two separate trajectories that are measured. In this example, the sensor switches between measuring the two airplanes,, i.e., impulses occur, at times $\mathcal{T} = \{4, 6, 8, 10, 12, 14\}$. Figure 3(b) shows the first-order convergence results derived using the reference signal shown in Figure 3 (a). Figure 4 shows results similar to those shown in Figure 3, except for a higher noise level (due to the initial conditions the impulses occur in the y-direction, all process noise is sampled from a zero mean normal distribution with standard deviation 0.05, and all measurement noise is sampled from a zero mean normal distribution whose standard deviation is 50% of the average separation distance of the two airplanes). Figure 4 (a) shows

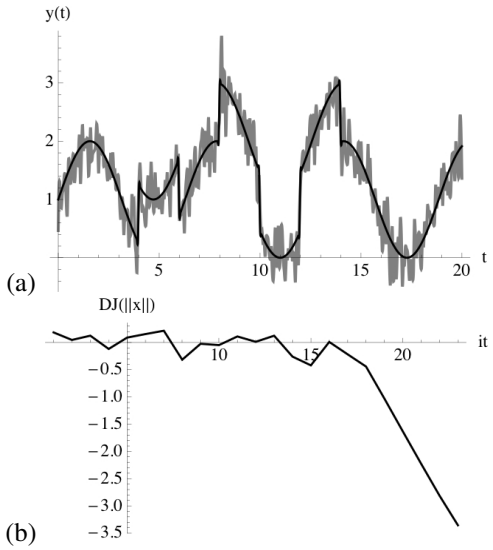

Fig. 3: (a) The y-component of the measurement signal in the two-airplane example when the signal to noise ratio is relatively high (gray line), and deterministic portions of individual trajectories actually measured (black line). The continuous measurement signal for this example was produced by interpolating discrete measurements, where $dt = 0.05$. (b) First-order convergence results with respect to the reference signal shown in (a).

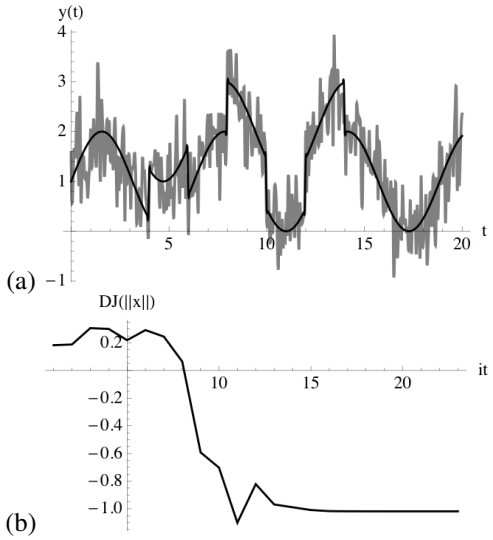

Fig. 4: (a) The y-component of the measurement signal in the two-airplane example when the signal to noise ratio is relatively low (gray line), and deterministic portions of individual trajectories actually measured (black line). The continuous measurement signal for this example was produced by interpolating discrete measurements, where $dt = 0.05$. (b) First-order convergence results with respect to the reference signal shown in (a).

the y-component of the measurement signal used to produce the first-order convergence results shown in Figure 4 (b). The convergence results in Figure 4 (b) no longer exhibit the same kind of linear behavior as those shown in Figures 2 and 3 (b),

but the algorithm somewhat surprisingly does still converge with this low signal to noise ratio. In this example, the sensor again switches between measuring the two airplanes,, i.e., impulses occur, at times $\mathcal{T} = \{4, 6, 8, 10, 12, 14\}$. The initial guess for the optimization which produced the results in Figures 2, 3, and 4 was $\mathcal{T}_0 = \{3.3, 6.2, 7, 9.4, 11, 16.5\}$.

C. Timing Results

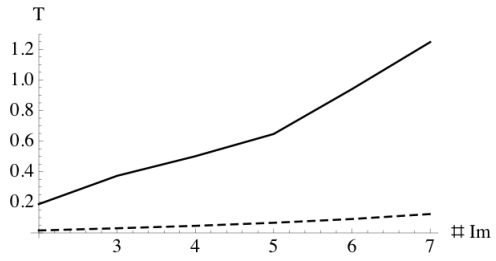

Fig. 5: Timing plot comparing adjoint formulations (solid line) and LTV formulations (dashed line) for the first derivative of the cost vs. the number of impulses.

Assuming that a nonlinear system may be linearized about a nominal system trajectory, the algorithm derived in Section III can be compared to the results of the algorithm described in [14], which optimize over impulse times for either linear or nonlinear systems, but require a single integration for each step in the descent algorithm, e.g., steepest descent [3].

Figure 5 shows timing results that compare the adjoint calculations of [14] to the LTV calculations presented in this work (applied to the system of Section IV-B). The horizontal axis represents the total number of impulses optimized over, and the vertical axis the time in seconds it takes to calculate the first derivative of the cost. The dashed line represents the results using the LTV formulation (online), and the solid line using the adjoint formulation. As we would expect, the LTV calculations are much faster than the adjoint calculations; the LTV calculations rely on a single integration over all steps of the descent algorithm and the adjoint calculations rely on a single integration for each step of the descent algorithm.

Although Figure 5 provides a clear advantage to using the LTV method to optimize over impulse times, it is not always possible to simply rely on the less computationally burdensome method. The LTV formulations depend on the quality of the linearization whereas the adjoint formulations do not.

V. Conclusions and Future Work

This paper presents an extension of the work in [7], where an optimization algorithm capable of optimizing over an arbitrary number of switching times (for switched dynamic systems) by numerically calculating two integrals is derived. The work in this paper extends the concepts taken from switched dynamical systems to impulsive systems. The algorithm presented in Section III represents this extension. Results of applying this algorithm to a data association example are presented in Section IV. These results showed that convergence over multiple impulse times was possible in cases where the signal to noise ratio was relatively low. The results also demonstrated a significant improvement in terms of the time needed to complete the optimization between the method presented in this paper and the previous state of the art.

An extension of the work presented in this paper is to include hybrid dynamics in the impulse optimization algorithm for LTV systems. This extension is straight forward, as it is a combination of the work presented in this paper and that found in [7].

References

[1] M. Alamir and S. A. Attia. Discussion on the paper "an optimal control approach for hybrid systems" by P. Riefinger, C. Iung, and F. Kratz. *European Journal of Control*, pages 459–460, 2003.

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

[3] L. Armijo. Minimization of functions having Lipschitz continuous first-partial derivatives. *Pacific Journal of Mathematics*, 16:1–3, 1966.

[4] S. A. Attia, M. Alamir, and C. Canudas de Wit. Sub-optimal control of switched nonlinear systems under location and switching constraints. In *Proceedings of the 16th IFAC World Congress*, pages 558–563, 2005.

[5] H. Axelsson, Y. Wardi, M. Egerstedt, and E. I. Verriest. Gradient descent approach to optimal mode scheduling in hybrid dynamical systems. *Journal of Optimization Theory and Applications*, 136:167–186, 2008.

[6] T. Caldwell and T. D. Murphey. Switching mode generation and optimal estimation with application to skid-steering. *Automatica*, 2010. In Press.

[7] T. M. Caldwell and T. D. Murphey. Single integration optimization of linear time-varying switched systems. In *American Control Conference*, San Francisco, CA, 2011.

[8] C. T. Chen. *Linear Systems Theory and Design, 3rd ed.* Oxford University Press, 1999.

[9] X. C. Ding, Y. Wardi, and F. Delmotte. On-line optimization of switched-mode dynamical systems. *IEEE Transactions on Automatic Control*, 54:2266–2271, 2009.

[10] M. Egerstedt, S. Azuma, and H. Axelsson. Transition-time optimization for switched-mode dynamical systems. *IEEE Transactions on Automatic Control*, 51:110–115, 2006.

[11] M. Egerstedt, S. Azuma, and Y. Wardi. Optimal timing control of switched linear systems based on partial information. *Nonlinear Analysis: Theory, Methods, and Applications*, 65:1736–1750, 2006.

[12] M. Egerstedt, Y. Wardi, and F. Delmotte. Optimal control of switching times in switched dynamical systems. In *IEEE Conference of Decision and Control*, pages 133–139, 2003.

[13] S. Hedlund and A. Rantzer. Convex dynamic programming for hybrid systems. *IEEE Transactions on Automatic Control*, 47:1536–1540, 2002.

[14] M. Travers, T. D. Murphey, and L. Pao. Impulse optimization for data association. In *IEEE Conference on Decision and Control*, pages 2204–2209, 2010.

[15] M. Travers, T. D. Murphey, and L. Pao. Impulsive data association with an unknown number of targets. In *Hybrid Systems: Computation and Control*, 2011.

[16] M. Travers, T. D. Murphey, and L. Pao. Trajectory optimization estimator for impulsive data association. In *American Control Conference*, pages 4514–4519, San Francisco, CA, 2011.

[17] X. Xu and P. J. Antsaklis. Optimal control of switched systems via non-linear optimization based on direct differentiations of value functions. *International Journal of Control*, 75:1406–1426, 2002.