

# Impulsive Data Association with an Unknown Number of Targets

Matthew Travers  
Northwestern University  
Department of Mechanical  
Engineering  
Evanston, Illinois 60208  
botics11@gmail.com

Todd Murphey  
Northwestern University  
Department of Mechanical  
Engineering  
Evanston, Illinois 60208  
t-  
murphey@northwestern.edu

Lucy Pao  
University of Colorado  
Department of Electrical,  
Computer, and Energy  
Engineering  
Boulder, Colorado  
pao@colorado.edu

## ABSTRACT

First- and second-order solution methods for the multi-target data association problem with an unknown number of targets are presented. It is shown that by considering a single continuous measurement signal with impulsive switching between measuring the position of different objects, the data association problem can be recast as a continuous optimization over the impulse times and magnitudes. First- and second-order adjoint formulations are derived which reduce the calculation of either the first- or second-order derivative of the cost function to a single integration (over any number of impulse times and magnitudes). These adjoint formulations as well as a method for estimating the total number of impulses which occur are the main contributions of this work.

## Categories and Subject Descriptors

J [Computer Applications]: Miscellaneous; J.2 [Physical Sciences and Engineering]: Mathematics & Statistical Engineering

## General Terms

Measurements & Verification

## Keywords

Data Association, Optimization, & Filtering Theory

## 1. INTRODUCTION

Multi-target data association is a common interest across several different fields [2, 4, 10, 11, 16, 19, 26, 27, 29]. The systems of interest contain multiple targets from which measurements may originate. There is uncertainty in the positions of the targets due to an imperfect system model as well as noise contained in the measurements. The goal of

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

HSCC'11, April 12–14, 2011, Chicago, Illinois, USA.

Copyright 2011 ACM 978-1-4503-0629-4/11/04 ...\$10.00.

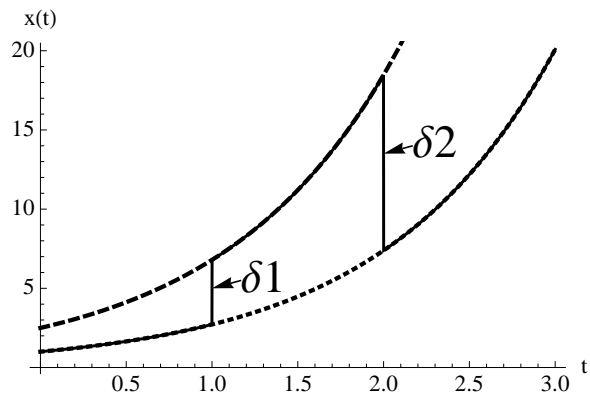


Figure 1: Trajectories of two dynamically identical objects with different initial conditions, represented by the dotted and dashed lines. The solid black line represents the portions of the two trajectories measured over the time horizon  $t = [0, 3]$ . The change between measuring different trajectories is an impulse in the measurement signal. In this example, the magnitudes of the impulses are  $\delta_1$  and  $\delta_2$ .

the data association problem is to assign the measurements to the correct target.

This paper presents a new method for nonlinear data association referred to as *impulsive data association* (IDA). Figure 1 shows a simple example. In Figure 1, the dotted line represents the trajectory of a pre-specified object of interest (object 1) and the dashed line some other object nearby (object 2). There is a single sensor which produces a single measurement signal that is assumed to be continuous. The solid black line in Figure 1 represents the portions of the two trajectories that the sensor measures over  $t = [0, 3]$ . The transitions between measuring objects 1 and 2 are impulses, which in this example have magnitudes  $\delta_1$  and  $\delta_2$ . Assuming the object of interest is originally being measured and that the times at which impulses occur in the “measurement trajectory” are known, it is possible to determine the origin of individual measurements.

The main contribution of this work is the derivation of the first- and second-order adjoint formulations for the derivative of a cost with respect to impulse times, magnitudes, and the second-order coupling between the impulse times and magnitudes. An interesting and computationally useful

result, presented in Section 3, is that the same adjoint operators which appear in the first- and second-order derivatives of the cost with respect to the impulse times also appear unmodified in the derivatives with respect to the impulse magnitudes. This result is computationally useful due to the fact that quadratic convergence with respect to the impulse times and magnitudes is achieved by calculating two integrals, regardless of the number of total impulses.

There are several techniques in the literature that address the multi-target data association problem with an unknown number of objects. The most popular approaches are referred to as *maximum a posterior (MAP)* techniques [3]. MAP techniques analytically define the posterior distribution and test various elements from the set of possible solutions, i.e., partitions of the discrete measurements, to determine which one maximizes the posterior. Pre-existing MAP techniques have been shown to perform well in many realistic situations. The main issue facing current MAP techniques is the way in which the solution space is searched. The main theoretical contribution of this work is a method that takes advantage of the maximum principle to search the space of all possible partitions of measurements for an optimum with convergence in quadratic time.

*Multiple-hypothesis tracking (MHT)* [6, 8, 28] is an example of a MAP technique. The power of MHT is that, given enough time, it will always find the correct answer. The issue with MHT is that it exhaustively tests every element of the solution space at each time step to determine the MAP solution. As the number of measurements increases the MHT algorithm experiences exponential growth in terms of running time. Heuristic methods for dealing with this exponential growth in complexity have been proposed [10, 11, 20, 28], but at the expense of optimality.

More recently, *Markov Chain Monte Carlo data association (MCMCDA)* techniques have been developed [5, 18, 22, 24, 25]. MCMCDA differs from MHT due to the fact that MCMCDA uses MCMC techniques to search the space of possible solutions for areas with nontrivial probability with respect to the posterior distribution. Although this is a more computationally efficient method than MHT, it still experiences run time problems as the number of measurements gets large.

The rest of this paper is organized as follows. The problem is formally defined in Section 2. After defining the problem, the first and second-order derivatives of the cost function (defined in Section 2) with respect to the impulse times and magnitudes are derived (as well as the cross terms for the second derivatives). Finally, a simulated implementation of the IDA algorithm is provided and results analyzed. In Section 5 conclusions and future directions of this work are provided.

## 2. PROBLEM DEFINITION

In general, impulsive data association is applied to nonlinear systems with dynamics of the form

$$\begin{aligned} \dot{x}^j &= f(x^j(t), t), \quad x^j(0) = x_0^j, \\ z(t) &= h(x^j(t), t) + \nu(t), \quad \text{where } j \in \{1, \dots, N\} \end{aligned} \quad (1)$$

where  $N$  is the total number of objects in the system, each with identical dynamics  $f(\cdot)$ , and  $z(t)$  is the *measurement signal* which contains the noise term  $\nu(\cdot)$ . In the work presented in this paper we do not assume that  $N$  is known.

An assumption that we are making is that we have a way of addressing the uncertainty in the system model, such as the *Kalman-Bucy Filter* [1]. In Section 3 we do not explicitly talk about the trajectory  $x(t)$  being the result of a filtered signal, but in practice this is the case. The derivatives of the trajectory that are defined in Section 3 depend only on evaluating the value of the trajectory at certain points. The trajectory  $x(t)$  can thus be a filtered signal without loss of generality.

We define a cost function of the form

$$J(\cdot) = \int_{t_0}^{t_f} \ell(x(s), s) ds \quad (2)$$

where, for example  $\ell(x(s), s) = (x_d(s) - x(s))^T (x_d(s) - x(s))$ ,  $x_d(\cdot)$  is the reference trajectory (which is the measurement signal in this work), and  $x(\cdot)$  is the model of the trajectory. Assume that  $\ell(\cdot)$  is  $C^2$ ,  $f(\cdot)$  is  $C^2$ , and  $x_d(\cdot)$  is  $C^1$ .

## 3. THE DERIVATIVES OF $J(\cdot)$

In this section, we analytically derive the first and second derivatives of the cost function (2) with respect to an arbitrary number of impulse times and magnitudes. The second-order cross terms between the impulse times and magnitudes are also derived. An interesting result is that the same adjoint operators that appear in the first- and second-order derivatives of the cost with respect to the impulse times appear in a similar way when taking derivatives with respect to an arbitrary impulse amplitude.

### 3.1 First derivatives of $J(\cdot)$

In finding the first derivative of  $J(\cdot)$  it is helpful to first derive the derivatives of the trajectory  $x(\cdot)$ . Note that in Lemmas 1 and 2, the first derivative of the trajectory with respect to the impulse times and magnitudes have the exact same linear form, differing only in initial conditions. Lemma 1 gives the derivative of  $x(\cdot)$  with respect to an arbitrary impulse magnitude.

**Lemma 1** The first derivative of the trajectory with respect to the impulse magnitudes  $\delta_i$  is

$$D_{\delta_i} x(t) \circ \partial \delta_i = \begin{cases} 0, & t < \tau_i \\ \Phi(t, \tau_i) \circ \Delta^i, & t \geq \tau_i \end{cases} \quad (3)$$

$$\Delta^i = D_{\delta_i} x(\tau_i) \circ \partial \delta_i$$

where  $\Phi(t, \tau_i)$  is the state transition [9] matrix for the system

$$\dot{q} = A(t)q$$

and  $A(t) = D_1 f(x(t), t)$  (where  $D_1 f$  means the derivative of  $f$  with respect to the first argument).

**Proof:** Using the fundamental theorem of calculus, the trajectory  $x(t)$  can be written as

$$x(0) = x_0, \quad x(t) = x(\tau_i) + \int_{\tau_i}^t f(x(s), s) ds, \quad (4)$$

where  $\tau_i$  can be any time such that  $\tau_i < t$ . Taking the

derivative of (4) with respect to  $\delta_i$

$$D_{\delta_i}x(\cdot) \circ \partial\delta_i = D_{\delta_i}x(\tau_i) \circ \partial\delta_i + \int_{\tau_i}^t D_1f(x(s), s) \circ D_{\delta_i}x(s) \circ \partial\delta_i ds \quad (5)$$

The fundamental theorem of calculus can be used to rewrite (5) in differential form

$$\frac{\partial}{\partial t} D_{\delta_i}x(t) \circ \partial\delta_i = D_1f(x(t), t) \circ D_{\delta_i}x(t) \circ \partial\delta_i.$$

with initial condition,

$$D_{\delta_i}x(\tau_i) \circ \partial\delta_i$$

This is a linear differential equation and can thus be represented as a state transition matrix operating on an initial condition as follows:

$$D_{\delta_i}x(t) \circ \partial\delta_i = \Phi(t, \tau_i) \circ D_{\delta_i}x(\tau_i) \circ \partial\delta_i.$$

This result proves the Lemma for  $t \geq \tau_i$ . For  $t < \tau_i$ , observe that (4), does not depend on  $\delta_i$  until time  $\tau_i$ . ■

The next lemma states the first derivative of the trajectory with respect to an arbitrary impulse time  $\tau_i$ . Again due to the application of the chain rule, the derivative of the trajectory appears in the derivative of the cost with respect to the impulse times.

**Lemma 2** The first derivative of the trajectory with respect to the impulse times  $\tau_i$  is

$$D_{\tau_i}x(t) \circ \partial\tau_i = \begin{cases} 0, & t < \tau_i \\ \Phi(t, \tau_i) \circ X^i, & t \geq \tau_i \end{cases} \quad (6)$$

$$X^i = (f(x(\tau_i^-), \tau_i^-) - f(x(\tau_i^+), \tau_i^+)) \partial\tau_i$$

where  $\Phi(t, \tau_i)$  is the same state transition matrix from Lemma 1,  $\tau_i^-$  refers to the time right before the  $i$ -th impulse, and  $\tau_i^+$  the time right after.

**Proof:** Taking the derivative of (4) with respect to  $\tau_i$

$$\begin{aligned} D_{\tau_i}x(t) \circ \partial\tau_i &= D_{\tau_i}x(\tau_i^-) \circ \partial\tau_i - f(x(\tau_i^+), \tau_i^+) \\ &\quad + \int_{\tau_i^+}^t D_1f(x(s), s) D_{\tau_i}x(s) \circ \partial\tau_i ds \\ &= f(x(\tau_i^-), \tau_i^-) - f(x(\tau_i^+), \tau_i^+) \\ &\quad + \int_{\tau_i^+}^t D_1f(x(s), s) D_{\tau_i}x(s) \circ \partial\tau_i ds. \end{aligned} \quad (7)$$

The first term in (7) is the result of taking the derivative of the initial condition in (4) with respect to its argument. The second term in (7) is the result of applying the Leibniz rule. The fundamental theorem of calculus can be used to rewrite (7) in differential form

$$\begin{aligned} D_{\tau_i}x(\tau_i) \circ \partial\tau_i &= f(x(\tau_i^-), \tau_i^-) - f(x(\tau_i^+), \tau_i^+) \\ \frac{\partial}{\partial t} D_{\tau_i}x(t) \circ \partial\tau_i &= D_1f(x(t), t) \circ D_{\tau_i}x(t) \circ \partial\tau_i. \end{aligned}$$

This is a linear differential equation and can thus be represented as a state transition matrix operating on an initial condition

$$D_{\tau_i}x(t) = \Phi(t, \tau_i) \circ D_{\tau_i}x(\tau_i) \circ \partial\tau_i.$$

This result is one part of the Lemma. To obtain the second part of the Lemma, go back to Equation (4) and take the derivative of  $x(t)$  with respect to an arbitrary  $\tau_k$  such that  $\tau_k < \tau_i$ . The initial condition  $x(\tau_i^-)$  is constant and thus does not depend on  $\tau_k$ . For the integral term,  $x(t)$  does not depend on  $\tau_k$  anywhere in the interval  $[\tau_i^+, t]$ . Thus  $D_{\tau_k}x(t) \circ \partial\tau_k = 0$  when  $t < \tau_i$ . ■

Having derived the first derivatives of the trajectory it is now straightforward to derive the first derivatives of the cost  $J(\cdot)$ .

**Theorem 1** The derivative of the cost function  $J(\cdot)$  with respect to the impulse magnitudes  $\delta_i$  is:

$$D_{\delta_i}J(\cdot) \circ \partial\delta_i = \psi(t_f, \tau_i) \circ \Delta^i. \quad (8)$$

The linear operator  $\psi(t_f, \tau_i) : \mathbb{R}^n \rightarrow \mathbb{R}$  is found by integrating

$$\begin{aligned} \psi(t, t) \circ U &= 0 \\ \frac{\partial}{\partial \tau} \psi(t, \tau) \circ U &= -D_1\ell(x(\tau), \tau) \circ U - \psi(t, \tau) \circ D_1f(x(\tau), \tau) \circ U \end{aligned}$$

backward along  $\tau$  from  $t_f$  to  $\tau_i$ .

**Proof:** Take derivative of (2) with respect to  $\delta_i$

$$D_{\delta_i}J(\cdot) \circ \partial\delta_i = \int_{\tau_i}^{t_f} D_1\ell(x(s), s) D_{\delta_i}x(s) \circ \partial\delta_i ds, \quad (10)$$

and substitute in (3) to obtain

$$D_{\delta_i}J(\cdot) \circ \partial\delta_i = \int_{\tau_i}^{t_f} D_1\ell(x(s), s) \circ \Phi(s, \tau_i) ds \circ \Delta^i \quad (11)$$

where  $\Delta^i$  has been taken out of the integral because it does not depend on  $s$ . Defining

$$\psi(t, \tau) \circ U = \left( \int_{\tau}^t D_1\ell(x(s), s) \circ \Phi(s, \tau) ds \right) \circ U, \quad (12)$$

(11) can be rewritten as

$$D_{\delta_i}J(\cdot) \circ \partial\delta_i = \psi(t_f, \tau_i) \circ \Delta^i$$

which is the first part of Theorem 1. To obtain the second part of Theorem 1, take the derivative of (12) with respect to  $\tau$  [12].

$$\begin{aligned} \frac{\partial}{\partial \tau} \psi(t, \tau) \circ U &= -D_1\ell(x(\tau), \tau) \circ U \\ &\quad - \int_{\tau}^t D_1\ell(x(s), s) \circ \Phi(s, \tau) \circ A(\tau) U ds \quad (13a) \\ &= -D_1\ell(x(\tau), \tau) \circ U \\ &\quad - \left( \int_{\tau}^t D_1\ell(x(s), s) \circ \Phi(s, \tau) ds \right) \circ A(\tau) \circ U \quad (13b) \\ &= -D_1\ell(x(\tau), \tau) \circ U - \psi(t, \tau) \circ D_1f(x(\tau), \tau) \circ U \quad (13c) \end{aligned}$$

Equation (13c) along with evaluating  $\psi$  in (12) at  $\tau = t$  yields the final two parts of Theorem 1. ■

The next theorem states the first derivative of the cost with respect to an arbitrary impulse time. The derivation of

Theorem 2 is similar to the derivation of Theorem 1, except that it relies on the application of Leibniz's rule.

**Theorem 2** The derivative of the cost function  $J(\cdot)$  with respect to each of the impulse times  $\tau_i$  is

$$D_{\tau_i} J(\cdot) \circ \partial \tau_i = \Psi(t_f, \tau_i) \quad (14)$$

where  $\Psi(t_f, \tau_i) : \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$\Psi(t_f, \tau_i) = \psi(t_f, \tau_i) \circ X^i + \ell(x(\tau_i^-), \tau_i^-) - \ell(x(\tau_i^+), \tau_i^+).$$

**Proof:** Take the derivative of the cost function as it is written in Equation (2) with respect to  $\tau_i$ . The derivative is the sum of three parts, the derivative of the integrand itself along with two terms that come from applying Leibniz's rule. Recall that in Equation (6),  $D_{t_i} x(t) = 0$  up until  $t = \tau_i$ . Therefore the derivative of the integrand only needs to be integrated from  $\tau_i$  up to  $t_f$ . Thus,

$$D_{\tau_i} J(\cdot) \circ \partial \tau_i = \int_{\tau_i}^{t_f} D_1 \ell(x(s), s) \circ D_{\tau_i} x(s) \circ \partial \tau_i ds + \ell(x(\tau_i^-), \tau_i^-) - \ell(x(\tau_i^+), \tau_i^+). \quad (15)$$

Substituting in (6) and observing that  $X^i$  is independent of  $s$ , we can write

$$D_{\tau_i} J(\cdot) \circ \partial \tau_i = \psi(t, \tau_i) \circ X^i + \ell(x(\tau_i^-), \tau_i^-) - \ell(x(\tau_i^+), \tau_i^+)$$

where  $\psi(\cdot)$  is found by integrating (13c) backwards in time. ■

Note that the adjoint operator  $\psi(\cdot)$  appears in both (8) and (14), and thus the calculation of the first derivatives of  $J(\cdot)$  with respect to each impulse time as well as impulse magnitude requires only a single integration. This result is independent of the total number of impulses.

### 3.2 Second Derivatives of $J(\cdot)$

In deriving the second derivatives of the cost  $J(\cdot)$ , we will proceed in a similar manner to the derivations of the first derivatives, i.e., we will first derive the second derivatives of the trajectory.

The following two lemmas are provided in order to define the initial conditions which appear in the second derivatives of the trajectory and thus in the second derivatives of the cost due to the chain rule.

**Lemma 3** For  $i \geq j$  and  $t \geq \tau_i$ , the derivative of  $D_{\delta_i} x(t) \circ \partial \delta_i$  with respect to  $\delta_j$  satisfies the differential equation (with initial condition  $\Delta^{i,j}$ )

$$\begin{aligned} \frac{d}{dt} D_{\delta_j} D_{\delta_i} x(t) \circ (\partial \delta_j, \partial \delta_i) = & D_1 f(x(t), t) \circ D_{\delta_j} D_{\delta_i} x(t) \circ (\partial \delta_j, \partial \delta_i) \\ & + D_1^2 f(x(t), t) \circ (D_{\delta_j} x(t) \circ \partial \delta_j, D_{\delta_i} x(t) \circ \partial \delta_i) \end{aligned} \quad (16a)$$

$$\Delta^{i,j} = D_{\delta_j} D_{\delta_i} x(\tau_i) \circ (\partial \delta_j, \partial \delta_i) \quad (16b)$$

**Proof:** Differentiate (5) and apply the fundamental theorem of calculus. This is straightforward so it is not included here. ■

The next lemma provides an initial condition that appears in the second derivatives of the cross terms between the impulse times and magnitudes.

**Lemma 4** For  $i \geq j$  and  $t \geq \tau_i$ , the derivative of  $D_{\tau_i} x(t) \circ \partial \tau_i$  with respect to  $\delta_j$  satisfies the differential equation (with initial condition  $\Delta X^{i,j}$ )

$$\begin{aligned} \frac{d}{dt} D_{\delta_j} D_{\tau_i} x(t) \circ (\partial \delta_j, \partial \tau_i) = & D_1 f(x(t), t) \circ D_{\delta_j} D_{\tau_i} x(t) \circ (\partial \delta_j, \partial \tau_i) \\ & + D_1^2 f(x(t), t) \circ (D_{\delta_j} x(t) \circ \partial \delta_j, D_{\tau_i} x(t) \circ \partial \tau_i) \end{aligned} \quad (17a)$$

$$\begin{aligned} \Delta X^{i,j} = D_{\delta_j} D_{\tau_i} x(\tau_i) \circ (\partial \delta_j, \partial \tau_i) = & D_1 f(x(\tau_i^-), \tau_i^-) \circ D_{\delta_j} x(\tau_i^-) \circ \partial \delta_j \\ & - D_1 f(x(\tau_i^+), \tau_i^+) \circ D_{\delta_j} x(\tau_i^+) \circ \partial \delta_j \end{aligned} \quad (17b)$$

**Proof:** As before, differentiate  $D_{\tau_i} x(t) \circ \partial \tau_i$  and apply the fundamental theorem of calculus. ■

The following lemma provides the second derivative of the trajectory with respect to two impulse times.

**Lemma 5** For  $i \geq j$  and  $t \geq \tau_i$ , the second derivative of  $x(t)$  satisfies the differential equation (with initial condition  $X^{i,j}$ )

$$\begin{aligned} \frac{d}{dt} D_{\tau_j} D_{\tau_i} x(t) \circ (\partial \tau_j, \partial \tau_i) = & D_1 f(x(t), t) \circ D_{\tau_j} D_{\tau_i} x(t) \circ (\partial \tau_j, \partial \tau_i) \\ & + D_1^2 f(x(t), t) \circ (D_{\tau_j} x(t) \circ \partial \tau_j, D_{\tau_i} x(t) \circ \partial \tau_i) \end{aligned} \quad (18a)$$

$$\begin{aligned} X^{i,j} = D_{\tau_j} D_{\tau_i} x(\tau_i) \circ (\partial \tau_j, \partial \tau_i) = & \begin{cases} D_1 f(x(\tau_i^+), \tau_i^+) \circ f(x(\tau_i^+), \tau_i^+) \partial \tau_j \partial \tau_i \\ \quad + D_1 f(x(\tau_i^-), \tau_i^-) \circ f(x(\tau_i^-), \tau_i^-) \partial \tau_j \partial \tau_i \\ \quad - 2D_1 f(x(\tau_i^+), \tau_i^+) \circ f(x(\tau_i^-), \tau_i^-) \partial \tau_j \partial \tau_i \\ \quad + D_2 f(x(\tau_i^-), \tau_i^-) \circ \partial \tau_j \partial \tau_i \\ \quad \quad - D_2 f(x(\tau_i^+), \tau_i^+) \circ \partial \tau_j \partial \tau_i, & i = j \\ (D_1 f(x(\tau_i^-), \tau_i^-) \\ \quad - D_1 f(x(\tau_i^+), \tau_i^+)) \circ \Phi(\tau_i, \tau_j) \circ X^j \partial \tau_i, & i > j. \end{cases} \end{aligned} \quad (18b)$$

**Proof:** Differentiate (7) and apply the fundamental theorem of calculus. ■

Looking at Lemmas 3, 4, and 5, observe that the ODE's for the second derivatives of the trajectory are not linear as in the first derivatives, they are *affine*. From the form of the solution to a linear affine system, the next few lemmas complete our derivations of the second-order derivatives of the trajectories with respect to the impulse times and magnitudes.

**Lemma 6** The second derivative  $D_{\delta_j} D_{\delta_i} x(t) \circ (\partial \delta_j, \partial \delta_i)$  is

$$\begin{aligned} D_{\delta_j} D_{\delta_i} x(t) \circ (\partial \delta_j, \partial \delta_i) = & \Phi(t, \tau_i) \circ \Delta^{i,j} + \phi(t, \tau_i) (\Phi(\tau_i, \tau_j) \circ \Delta^j, \Delta^i) \end{aligned} \quad (19)$$

where  $\Phi(t, \tau)$  is the state transition matrix from Lemma 1 and the bilinear operator  $\phi(t, \tau) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  is defined as

$$\begin{aligned} \phi(t, \tau) \circ (U, V) &= \\ \int_{\tau}^t \Phi(t, s) \circ D_1^2 f(x(s), s) \circ (\Phi(s, \tau) \circ U, \Phi(s, \tau) \circ V) ds \end{aligned} \quad (20)$$

with  $\Delta^{i,j}$  as the initial condition from (16b).

**Proof:** Notice that (16a) is in the affine form  $\dot{x} = A(t)x + B(t)$ . Recalling that the solution to an affine system is  $x(t) = \Phi(t, t_0) \circ x_0 + \int_{\tau}^t \Phi(t, s) \circ B(s) ds$  [9], as well as using (3) and (16b), we find that

$$\begin{aligned} D_{\delta_j} D_{\delta_i} x(t) \circ (\partial \delta_j, \partial \delta_i) &= \\ &= \Phi(t, \tau_i) \circ \Delta^{i,j} + \int_{\tau_i}^t \Phi(t, s) \circ D_1^2 f(x(s), s) \\ &\quad \circ (D_{\delta_j} x(s) \circ \partial \delta_j, D_{\delta_i} x(s) \circ \partial \delta_i) ds \\ &= \Phi(t, \tau_i) \circ \Delta^{i,j} + \int_{\tau_i}^t \Phi(t, s) \circ D_1^2 f(x(s), s) \\ &\quad \circ (\Phi(s, \tau_j) \circ \Delta^j, \Phi(s, \tau_i) \circ \Delta^i) ds \\ &= \Phi(t, \tau_i) \circ \Delta^{i,j} + \int_{\tau_i}^t \Phi(t, s) \circ D_1^2 f(x(s), s) \\ &\quad \circ (\Phi(s, \tau_i) \circ \Phi(\tau_i, \tau_j) \circ \Delta^j, \Phi(s, \tau_i) \circ \Delta^i) ds \\ &= \Phi(t, \tau_i) \circ \Delta^{i,j} + \phi(t, \tau_i) \circ (\Phi(\tau_i, \tau_j) \circ \Delta^j, \Delta^i) \end{aligned}$$

where  $\Delta^i$  and  $\Delta^j$  have been pulled out of the integral because they do not depend on  $s$ . ■

By direct inspection of (20) we can write

$$\begin{aligned} \phi(t, t) \circ (U, V) &= 0 \\ \frac{\partial}{\partial \tau} \phi(t, \tau) \circ (U, V) &= -\Phi(t, \tau) \circ D_1^2 f(x(\tau), \tau) \circ (U, V) \\ &\quad - \phi(t, \tau) \circ (D_1 f(x(\tau), \tau) \circ U, V) \\ &\quad - \phi(t, \tau) \circ (U, D_1 f(x(\tau), \tau) \circ V). \end{aligned} \quad (21)$$

In a similar way to the calculation of  $\psi(\cdot)$ ,  $\phi(\cdot)$  can be found by integrating (21) backwards in time.

The next lemma derives the second derivative of the trajectory for the cross terms between impulse times and magnitudes.

**Lemma 7** The second derivative  $D_{\delta_j} D_{\tau_i} x(t) \circ (\partial \delta_j, \partial \tau_i)$  is

$$\begin{aligned} D_{\delta_j} D_{\tau_i} x(t) \circ (\partial \delta_j, \partial \tau_i) &= \\ \Phi(t, \tau_i) \circ \Delta X^{i,j} + \phi(t, \tau_i) \circ (\Phi(\tau_i, \tau_j) \circ \Delta^j, X^i) \end{aligned} \quad (22)$$

**Proof:** Using (17a) and (17b) and plugging in (3) and (6)

$$\begin{aligned} D_{\delta_j} D_{\tau_i} x(t) \circ (\partial \delta_j, \partial \tau_i) &= \\ &= \Phi(t, \tau_i) \circ \Delta X^{i,j} + \int_{\tau_i}^t \Phi(t, s) \circ D_1^2 f(x(s), s) \\ &\quad \circ (D_{\delta_j} x(s) \circ \partial \delta_j, D_{\tau_i} x(s) \circ \partial \tau_i) ds \\ &= \Phi(t, \tau_i) \circ \Delta X^{i,j} + \int_{\tau_i}^t \Phi(t, s) \circ D_1^2 f(x(s), s) \\ &\quad \circ (\Phi(s, \tau_j) \circ \Delta^j, \Phi(s, \tau_i) \circ X^i) ds \\ &= \Phi(t, \tau_i) \circ \Delta X^{i,j} + \int_{\tau_i}^t \Phi(t, s) \circ D_1^2 f(x(s), s) \\ &\quad \circ (\Phi(s, \tau_i) \circ \Phi(\tau_i, \tau_j) \circ \Delta^j, \Phi(s, \tau_i) \circ X^i) ds \\ &= \Phi(t, \tau_i) \circ \Delta X^{i,j} + \phi(t, \tau_i) \circ (\Phi(\tau_i, \tau_j) \circ \Delta^j, X^i) \end{aligned}$$

where  $\Delta^i$  and  $\Delta^j$  have been taken out of the integral because they do not depend on  $s$ . ■

The following lemma provides the second-order derivative of the trajectory with respect to the impulse times  $\tau_i$  and  $\tau_j$ .

**Lemma 8** The second derivative  $D_{\tau_j} D_{\tau_i} x(t) \circ (\partial \tau_j, \partial \tau_i)$  is

$$\begin{aligned} D_{\tau_j} D_{\tau_i} x(t) \circ (\partial \tau_j, \partial \tau_i) &= \Phi(t, \tau_i) \circ X^{i,j} \\ &\quad + \phi(t, \tau_i) \circ (\Phi(\tau_i, \tau_j) \circ X^j, X^i) \end{aligned} \quad (23)$$

where  $\Phi(t, \tau)$  is the state transition matrix from Lemma 1 and  $X^{i,j}$  is the initial condition from (18b).

**Proof:** Using (18a) and (18b)

$$\begin{aligned} D_{\tau_j} D_{\tau_i} x(t) \circ (\partial \tau_j, \partial \tau_i) &= \\ &= \Phi(t, \tau_i) \circ X^{i,j} + \int_{\tau_i}^t \Phi(t, s) \circ D_1^2 f(x(s), s) \\ &\quad \circ (D_{\tau_j} x(s) \circ \partial \tau_j, D_{\tau_i} x(s) \circ \partial \tau_i) ds \\ &= \Phi(t, \tau_i) \circ X^{i,j} + \int_{\tau_i}^t \Phi(t, s) \circ D_1^2 f(x(s), s) \\ &\quad \circ (\Phi(s, \tau_j) \circ X^j, \Phi(s, \tau_i) \circ X^i) ds \\ &= \Phi(t, \tau_i) \circ X^{i,j} + \int_{\tau_i}^t \Phi(t, s) \circ D_1^2 f(x(s), s) \\ &\quad \circ (\Phi(s, \tau_i) \circ \Phi(\tau_i, \tau_j) \circ X^j, \Phi(s, \tau_i) \circ X^i) ds \\ &= \Phi(t, \tau_i) \circ X^{i,j} + \phi(t, \tau_i) \circ (\Phi(\tau_i, \tau_j) \circ X^j, X^i) \end{aligned}$$

where  $X^i$  and  $X^j$  have been pulled out of the integral because they do not depend on  $s$ . ■

Having derived the second derivatives of the trajectory with respect to the impulse times, magnitudes, and cross terms between the impulse times and magnitudes, it is now possible to derive the second derivatives of the cost.

**Theorem 3** The second derivative of the cost function  $J(\cdot)$  with respect to the impulse magnitude  $\delta_j$  where  $\tau_i \geq \tau_j$  is

$$\begin{aligned} D_{\delta_j} D_{\delta_i} J(\cdot) \circ (\partial \delta_j, \partial \delta_i) &= \\ \psi(t_f, \tau_i) \circ \Delta^{i,j} + \Omega(t_f, \tau_i) \circ (\Phi(\tau_i, \tau_j) \circ \Delta^j, \Delta^i) \end{aligned}$$

where  $\Omega(t, \tau) \circ (U, V) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  is the bilinear operator

found by integrating

$$\Omega(t, t) \circ (U, V) = 0_{n \times n} \quad (24a)$$

$$\begin{aligned} \frac{\partial}{\partial \tau} \Omega(t, \tau) \circ (U, V) &= -D_1^2 \ell(x(\tau), \tau) \circ (U, V) \\ &\quad - \psi(t, \tau) \circ D_1^2 f(x(\tau), \tau) \circ (U, V) \\ &\quad - \Omega(t, \tau) \circ (D_1 f(x(\tau), \tau) \circ U, V) \\ &\quad - \Omega(t, \tau) \circ (U, D_1 f(x(\tau), \tau) \circ V) \end{aligned} \quad (24b)$$

backwards over  $\tau$  from  $t_f$  to  $\tau_i$ .

**Proof:** Take the derivative of (10) with respect to  $\delta_j$  and plug in (19)

$$\begin{aligned} D_{\delta_j} D_{\delta_i} J(\cdot) \circ (\partial \delta_j, \partial \delta_i) &= \\ &\int_{\tau_i}^{t_f} D_1 \ell(x(s), s) \circ D_{\delta_j} D_{\delta_i} x(s) \circ (\partial \delta_j, \partial \delta_i) \\ &\quad + D_1^2 \ell(x(s), s) \circ (D_{\delta_j} x(s) \circ \partial \delta_j, D_{\delta_i} x(s) \circ \partial \delta_i) ds \\ &= \int_{\tau_i}^{t_f} D_1 \ell(x(s), s) \circ \Phi(s, \tau_i) \circ \Delta^{i,j} \\ &\quad + D_1 \ell(x(s), s) \circ \phi(s, \tau_i) \circ (\Phi(\tau_i, \tau_j) \circ \Delta^j, \Delta^i) \\ &\quad + D_1^2 \ell(x(s), s) \circ (\Phi(s, \tau_i) \circ \Phi(\tau_i, \tau_j) \circ \Delta^j, \Phi(s, \tau_i) \circ \Delta^j) \\ &= \psi(t_f, \tau_i) \circ \Delta^{i,j} + \Omega(t_f, \tau_i) \circ (\Phi(\tau_i, \tau_j) \circ \Delta^j, \Delta^i) \end{aligned}$$

where  $\Omega(t, \tau)$  is defined as

$$\begin{aligned} \Omega(t, \tau) \circ (U, V) &= \int_{\tau}^t D_1 \ell(x(s), s) \circ \Phi(s, \tau) \circ (U, V) \\ &\quad + D_1^2 \ell(x(s), s) \circ (\Phi(s, \tau) \circ U, \Phi(s, \tau) \circ V) ds. \end{aligned} \quad (25)$$

This provides the first part of the proof. To obtain the second parts, take the derivative of (25) with respect to  $\tau$  (this calculation is straight forward and is thus omitted here). ■

The following theorem provides the second-order cross derivatives of the cost with respect to the impulse magnitude  $\delta_j$  and impulse time  $\tau_i$ .

**Theorem 4** The second order cross derivative of the cost function  $J(\cdot)$  with respect to the impulse time  $\tau_i$  and magnitude  $\delta_j$  where  $\tau_i \geq \tau_j$  is

$$\begin{aligned} D_{\delta_j} D_{\tau_i} J(\cdot) \circ (\partial \delta_j, \partial \tau_i) &= \\ &\quad D_1 \ell(x(\tau_i^-, \tau_i^-) \circ D_{\delta_j} x(\tau_i^-) \circ \partial \delta_j \\ &\quad - D_1 \ell(x(\tau_i^+, \tau_i^+) \circ D_{\delta_j} x(\tau_i^+) \circ \partial \delta_j \\ &\quad + \psi(t_f, \tau_i) \circ \Delta X^{i,j} + \Omega(t_f, \tau_i) \circ (\Phi(\tau_i, \tau_j) \circ \Delta^j, X^i) \end{aligned}$$

**Proof:** Take the derivative  $D_{\tau_i} J(\cdot) \circ \partial \tau_i$  with respect to  $\delta_j$  and plug in (22)

$$\begin{aligned} D_{\delta_j} D_{\tau_i} J(\cdot) \circ (\partial \delta_j, \partial \tau_i) &= \\ &\quad D_1 \ell(x(\tau_i^-, \tau_i^-)) \circ D_{\delta_j} x(\tau_i^-) \circ \partial \delta_j \\ &\quad - D_1 \ell(x(\tau_i^+, \tau_i^+)) \circ D_{\delta_j} x(\tau_i^+) \circ \partial \delta_j \\ &\quad + \int_{\tau_i}^{t_f} D_1 \ell(x(s), s) \circ D_{\delta_j} D_{\tau_i} x(s) \circ (\partial \delta_j, \partial \tau_i) \\ &\quad + D_1^2 \ell(x(s), s) \circ (D_{\delta_j} x(s) \circ \partial \delta_j, D_{\delta_i} x(s) \circ \partial \delta_i) ds \end{aligned}$$

$$\begin{aligned} &= D_1 \ell(x(\tau_i^-, \tau_i^-)) \circ D_{\delta_j} x(\tau_i^-) \circ \partial \delta_j \\ &\quad - D_1 \ell(x(\tau_i^+, \tau_i^+)) \circ D_{\delta_j} x(\tau_i^+) \circ \partial \delta_j \\ &\quad + \int_{\tau_i}^{t_f} D_1 \ell(x(s), s) \circ \Phi(s, \tau_i) \circ \Delta X^{i,j} \\ &\quad + D_1 \ell(x(s), s) \circ \phi(s, \tau_i) \circ (\Phi(\tau_i, \tau_j) \circ \Delta^j, X^i) \\ &\quad + D_1^2 \ell(x(s), s) \circ (\Phi(s, \tau_i) \circ \Phi(\tau_i, \tau_j) \circ \Delta^j, \Phi(s, \tau_i) \circ X^i) \\ &= D_1 \ell(x(\tau_i^-, \tau_i^-)) \circ D_{\delta_j} x(\tau_i^-) \circ \partial \delta_j \\ &\quad - D_1 \ell(x(\tau_i^+, \tau_i^+)) \circ D_{\delta_j} x(\tau_i^+) \circ \partial \delta_j \\ &\quad + \psi(t_f, \tau_i) \circ \Delta X^{i,j} + \Omega(t_f, \tau_i) \circ (\Phi(\tau_i, \tau_j) \circ \Delta^j, X^i) \blacksquare \end{aligned}$$

The following theorem provides the second-order derivatives of the cost with respect to the impulse times  $\tau_i$  and  $\tau_j$ .

**Theorem 5** The second derivative of the cost function  $J(\cdot)$  with respect to the impulse times  $\tau_j$  and  $\tau_i$  where  $\tau_i \geq \tau_j$  is

$$\begin{aligned} D_{\tau_j} D_{\tau_i} J(\cdot) \circ (\partial \tau_j, \partial \tau_i) &= \\ &\quad D_1 \ell(x(\tau_i^-, \tau_i^-) \circ (D_{\tau_j} x_d(\tau_i^-) \circ \frac{\partial \tau_i}{\partial \tau_j} - D_{\tau_j} x(\tau_i^-) \circ \partial \tau_j) \\ &\quad - D_1 \ell(x(\tau_i^+, \tau_i^+) \circ (D_{\tau_j} x_d(\tau_i^+) \circ \frac{\partial \tau_i}{\partial \tau_j} - D_{\tau_j} x(\tau_i^+) \circ \partial \tau_j) \\ &\quad - D_1 \ell(x(\tau_i), \tau_i) \circ X^i \partial \tau_i \delta_j^i + \psi(t_f, \tau_i) \circ X^{i,j} \\ &\quad + \Omega(t_f, \tau_i) \circ (\Phi(\tau_i, \tau_j) \circ X^j, X^i) \end{aligned}$$

**Proof:** Take the derivative of (15) with respect to  $\tau_j$  and plugging in (23)

$$\begin{aligned} D_{\tau_j} D_{\tau_i} J(\cdot) \circ (\partial \tau_j, \partial \tau_i) &= \\ &\quad \frac{\partial}{\partial \tau_j} \left( \int_{\tau_i^+}^{t_f} D_1 \ell(x(s), s) \circ D_{\tau_i} x(s) \circ \partial \tau_i ds \right. \\ &\quad \left. + \ell(x(\tau_i^-, \tau_i^-) - \ell(x(\tau_i^+, \tau_i^+)) \right) \\ &= D_1 \ell(x(\tau_i^-, \tau_i^-) \circ (D_{\tau_j} x_d(\tau_i^-) \circ \frac{\partial \tau_i}{\partial \tau_j} \\ &\quad - D_{\tau_j} x(\tau_i^-) \circ \partial \tau_j) - D_1 \ell(x(\tau_i^+, \tau_i^+) \\ &\quad \circ (D_{\tau_j} x_d(\tau_i^+) \circ \frac{\partial \tau_i}{\partial \tau_j} - D_{\tau_j} x(\tau_i^+) \circ \partial \tau_j) \\ &\quad - D_1 \ell(x(\tau_i^+, \tau_i^+) \circ D_{\tau_i} x(\tau_i^+) \circ \partial \tau_i \frac{\partial \tau_i}{\partial \tau_j} \\ &\quad + \int_{\tau_i^+}^{t_f} (D_1 \ell(x(s), s) \circ D_{\tau_j} D_{\tau_i} x(s) \circ (\partial \tau_j, \partial \tau_i) \\ &\quad + D_1^2 \ell(x(s), s) \circ (D_{\tau_j} x(s) \circ \partial \tau_j, D_{\tau_i} x(s) \circ \partial \tau_i)) ds \\ &= D_1 \ell(x(\tau_i^-, \tau_i^-) \circ (D_{\tau_j} x_d(\tau_i^-) \circ \frac{\partial \tau_i}{\partial \tau_j} \\ &\quad - D_{\tau_j} x(\tau_i^-) \circ \partial \tau_j) - D_1 \ell(x(\tau_i^+, \tau_i^+) \\ &\quad \circ (D_{\tau_j} x_d(\tau_i^+) \circ \frac{\partial \tau_i}{\partial \tau_j} - D_{\tau_j} x(\tau_i^+) \circ \partial \tau_j) \end{aligned}$$

$$\begin{aligned}
& -D_1\ell(x(\tau_i^+), \tau_i^+) \circ D_{\tau_i}x(\tau_i^+) \circ \partial\tau_i \frac{\partial\tau_i}{\partial\tau_j} \\
& + \int_{\tau_i^+}^{t_f} (D_1\ell(x(s), s) \circ \Phi(s, \tau_i) \circ X^{i,j}) \\
& + D_1\ell(x(s), s) \circ \phi(s, \tau_i) \circ (\Phi(\tau_i, \tau_j) \circ X^j, X^i)) ds \\
& + \int_{\tau_i^+}^{t_f} D_1^2\ell(x(s), s) \circ (\Phi(s, \tau_i) \circ \Phi(\tau_i, \tau_j) \\
& \quad \circ X^j, \Phi(s, \tau_i) \circ X^i) ds \\
= & D_1\ell(x(\tau_i^-), \tau_i^-) \circ (D_{\tau_j}x_d(\tau_i^-) \circ \partial\tau_j\delta_j^i \\
& - D_{\tau_j}x(\tau_i^-) \circ \partial\tau_j) - D_1\ell(x(\tau_i^+), \tau_i^+) \\
& \quad \circ (D_{\tau_j}x_d(\tau_i^+) \circ \partial\tau_j\delta_j^i - D_{\tau_j}x(\tau_i^+) \circ \partial\tau_j) \\
& - D_1\ell(x(\tau_i^+), \tau_i^+) \circ X^i \partial\tau_j\delta_j^i + \psi(t_f, \tau_i) \circ X^{i,j} \\
& + \Omega(t_f, \tau_i) \circ (\Phi(\tau_i, \tau_j) \circ X^j, X^i)
\end{aligned}$$

where  $\delta_j^i$  is the Kronecker delta and

$$\frac{\partial\tau_i}{\partial\tau_j} = \begin{cases} \partial\tau_j & i = j \\ 0 & i \neq j. \end{cases} \quad (26)$$

The first two terms in (26) are the derivatives of the two Leibniz terms in (15),  $\ell(x(\tau_i^-), \tau_i^-)$  and  $\ell(x(\tau_i^+), \tau_i^+)$ . It is important to note that the derivative of the reference  $x_d(\cdot)$  appears in the second derivative of  $J(\cdot)$  with respect to the impulse times  $\tau_i$  and  $\tau_j$  for impulsive systems.

To explain the appearance of the derivative of the reference in the second derivatives, return to the Leibniz terms in (15) (recalling that  $\ell(x(t), t) = (x_d(t) - x(t))^T(x_d(t) - x(t))$ ). The Leibniz terms in the first derivatives of  $J(\cdot)$  (with respect to the impulse time  $\tau_i$ ) are evaluated at  $\tau_i^-$  and  $\tau_i^+$ , respectively, and thus both the reference and the trajectory depend explicitly on  $\tau_i$ . Thus, when taking the derivatives of these Leibniz terms with respect to  $\tau_i$  a second time, the derivatives of both the reference and model must be present. Taking the derivatives of the reference is a nontrivial operation. It turns out that evaluating the derivatives of the model that result from the Leibniz terms is also nontrivial. When  $i \neq j$  the evaluation of the  $D_{\tau_j}x(\tau_i) \circ \partial\tau_j$  is straight forward and can be calculated using (6). When  $i = j$  care must be taken due to the fact that the derivative is now being taken with respect to the argument of  $x(\cdot)$ . In this case ( $i = j$ ),  $D_{\tau_j}x(\tau_i) \circ \partial\tau_j = f(x(\tau_i), \tau_i)$ . ■

## 4. IMPLEMENTATION

### 4.1 Trajectory Optimization

It was mentioned earlier that in order to properly implement the impulse optimization, we first need a way in which to estimate the total number of impulses that occur. Nonlinear trajectory optimization is the method proposed as a pre-step used to estimate the number of impulses that occur.

The exact form of the nonlinear trajectory optimization problem being solved is

$$\dot{x} = U(t)f(x, t) = F(x, U) \quad (27)$$

where  $U(t)$  is a diagonal matrix function unless otherwise noted and  $f(x, t)$  are the same dynamics from Equation (1). Note that for the system in (27) the trajectory optimization will always be nonlinear, even when the dynamics  $f(\cdot)$  are linear.

In order to perform the trajectory optimization, we define a cost function  $S(\cdot)$  separate from the cost  $J(\cdot)$  in Equation (2) such that

$$S(\eta) = \int_0^T \ell(s, x(s), U(s)) ds + m(x(T)) \quad (28)$$

where  $\eta \in \mathcal{T}$ , and  $\mathcal{T}$  is the trajectory manifold associated with the dynamics defined in (1). The problem can thus be stated as the constrained optimization

$$\min_{\eta \in \mathcal{T}} S(\eta). \quad (29)$$

Through the use of the projection operator ([7, 13, 14, 15], the constrained optimization in (29) can be rewritten as the unconstrained optimization

$$\min_{\xi \in \mathcal{L}} S(P(\xi))$$

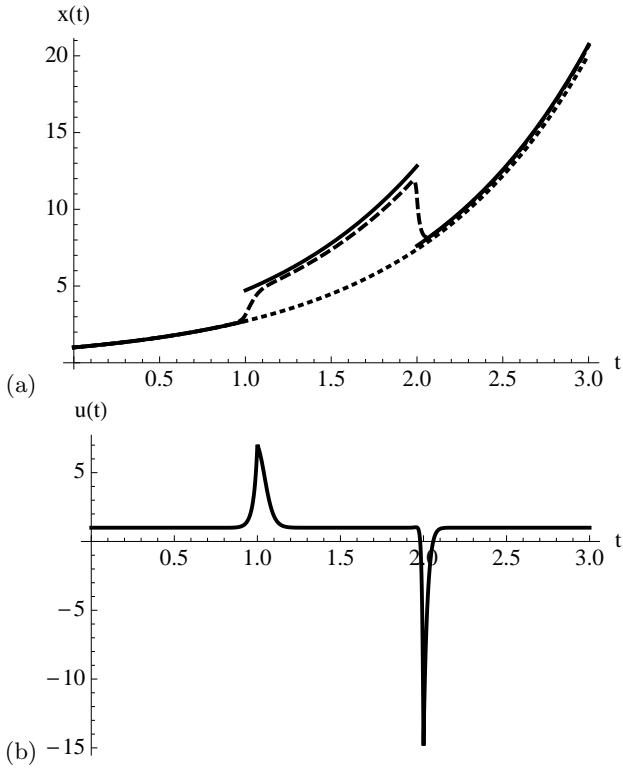
where  $\xi$  is an element of the infinite dimensional function space  $\mathcal{L}$ . This optimization problem is solved using a first-order descent method with a line search. Most of the technical details of the trajectory optimization will be left to several references ([7, 13, 14, 15]) and are thus not provided here.

In standard implementations of trajectory optimization the goal is to find an optimal trajectory with respect to a cost of the form (28). For the purposes of estimating the total number of impulses as well as approximating the times at which the impulses occur, the goal of the trajectory optimization is slightly modified. Using trajectory optimization as a pre-step to impulse optimization, the goal is to find deviations in the control  $U(\cdot)$ . It is explained further below, but we wish the nominal value of  $U(\cdot)$  to be equal to the identity. That is, we want—when possible—to have  $\dot{x} = f(x, t)$ . The goal of the trajectory optimization is to find deviations away from  $1(t)$ .

To explain why we would like to constrain  $U = I$ , consider the 1-D system containing two separate bodies each with dynamics  $\dot{x} = x$ , but with different initial conditions. This situation is shown in Figure 1, where a single sensor measures the position of one object over the intervals  $t = (0, 1)$  and  $t = (2, 3)$ , and a second object over the interval  $t = (1, 2)$ . Note that this example is deterministic and intentionally oversimplified for the purposes of clarity.

Figure 2 shows results of applying trajectory optimization to this 1-D system. In Figure 2(a) the solid line represents the desired signal, the dotted line the initial guess in the trajectory optimization, and the dashed line the current guess in the optimization after several iterations.

Figure 2(b) shows the control signal associated with the same step in the trajectory optimization that produced the dashed line in Figure 2(a). Through inspection of Figure 2 we can see that when the reference signal (solid line) is very close to the current trajectory in the optimization procedure (dashed line), the control signal value is close to one. When the difference between the current trajectory and the reference trajectory changes suddenly, the difference is reflected in the control signal as spikes that generate the delta function in the state (as can be seen in Figure 2(b)). The spikes in Figure 2(b) are the deviations from  $1(t)$  in the control that are desired. In this example, we would use thresholding [21, 23] to determine that there are two impulse times that we need to optimize over and our initial guess for the impulse optimization would be  $(\tau_1, \tau_2) \approx (1, 2)$ .



**Figure 2:** (a) Measurement signal (solid line), initial guess (dotted line) for trajectory optimization, and result (dashed line) of applying trajectory optimization to the system with dynamics  $\dot{x} = ux$ . (b) Control signal obtained as a result of applying trajectory optimization in (a).

## 4.2 Descent Methods

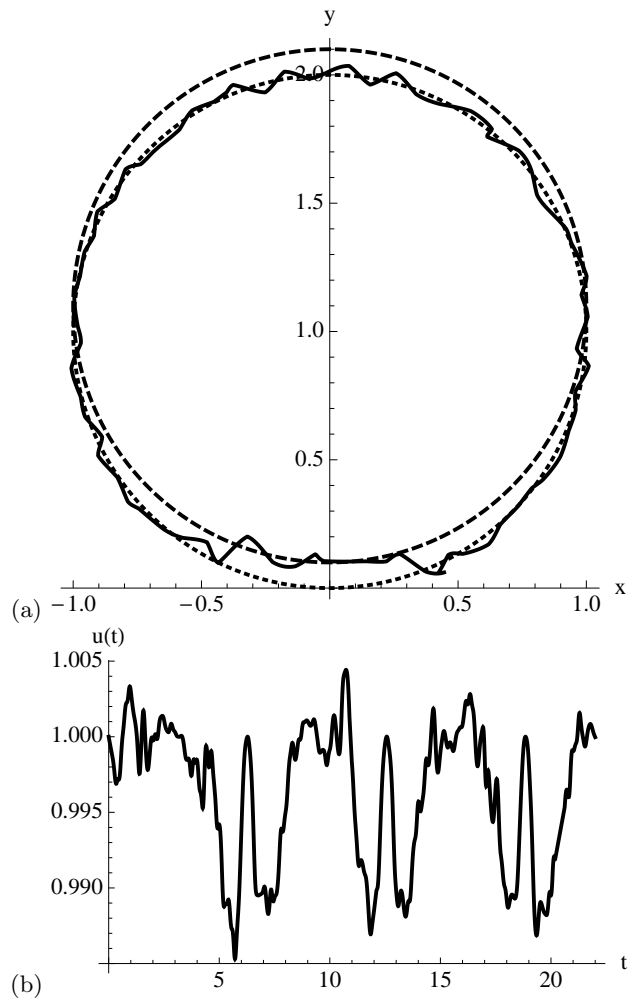
Having found the derivatives in Section 3 and a method for estimating the total number of impulses in Section 4.1, it is possible to implement both first- and second-order optimizations on (2), such as steepest descent with a line search and Newton’s method [17]. Note that in the example in Section 4.3, a combination of a quasi-Newton’s method and standard Newton’s method are used to produce the convergence results shown in Figure 4. The quasi-Newton’s method checks the eigenvalues of the Hessian and replaces any negative eigenvalues with 1, thus performing steepest descent in that subspace.

## 4.3 Example

The example selected to demonstrate the IDA algorithm has dynamics

$$\begin{aligned}\dot{x} &= v(t) \cos(\theta(t)) \\ \dot{y} &= v(t) \sin(\theta(t)) \\ \dot{\theta} &= \omega(t).\end{aligned}$$

where  $v(t)$  and  $\omega(t)$  are some inputs. Two objects are present in the system. Figure 3 shows an example of the two objects’ trajectories for  $v(t) = 1$  and  $\omega(t) = 1$ . In Figure 3(a), the dotted line represents the trajectory of a pre-specified object of interest, the dashed line represents the trajectory of the second nearby object, and the solid line represents the



**Figure 3:** (a) Measurements for the two-plane example (solid line). The dotted line represents the trajectory of the object of interest and the dashed line the trajectory of some plane nearby. (b) Result of applying the trajectory optimization of Section 4.1 to the two-plane system when six total impulses occur. The six spikes correspond to the six impulses.

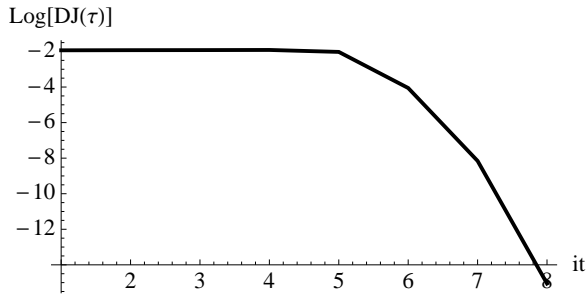
continuous measurement signal, found by interpolating the set of discrete noisy measurements.

It was mentioned earlier that in order to determine the times at which impulses occur in the measurement trajectory, the total number of impulses must first be estimated. Figure 3(b) shows the results of applying the trajectory optimization algorithm of Section 4.1 to the two airplane system over three periods of the circular trajectories shown in Figure 3(a). Note that there are six distinct peaks in the control signal shown in Figure 3(b) that correspond to the six impulses that occur in the measurement trajectory. Note also that we are showing a single component of the matrix  $U(\cdot)$  in Figure 3(b) due to the fact that for this example

$$U(t) = \begin{pmatrix} u_1 & 0 & 0 \\ 0 & u_2 & 0 \\ 0 & 0 & u_3 \end{pmatrix}, \quad (30)$$



where  $u_1 = u_3 = 1$  because the impulse always occurs in the  $y$ -direction.



**Figure 4: Second-order convergence results of applying impulse optimization over impulse times and magnitudes to the two plane, six impulse system.**

Figure 4 presents the second-order convergence that results from optimizing (2) over six impulse times as well as magnitudes. Figure 4 shows that we do in fact achieve quadratic convergence and that we reach a solution, with magnitude of the gradient less than  $10^{-14}$  on a log based 10 scale, within about 8 iterations.

## 5. CONCLUSIONS AND FUTURE WORK

This paper presented a new method for solving the multi-target data association problem for systems with an unknown number of targets, called impulsive data association. The algorithm is “impulsive” in that the sensor switching between measuring the trajectories of different targets is modeled as an impulse in a continuous measurement signal.

The main contribution of the work presented in this paper is the derivation of first- and second-order adjoint equations for the first- and second-order derivatives of the cost function (2) with respect to impulse times and magnitudes. An interesting result that arises in deriving the derivatives of the cost is that the same adjoint operator appears in the derivatives with respect to both the impulse times and magnitudes. This result means that it is possible to compute the entire second derivative of the cost with respect to an unknown number of impulse times and magnitudes by integrating two separate equations. This helps minimize the computational complexity.

In Section 2, we mentioned the fact that we were ignoring the presence of process noise in the work presented in this paper. In an actual implementation of IDA the uncertainty in the initial distribution as well as the uncertainty associated with the process noise needs to be explicitly addressed. One particular choice of methods to address these two sources of additional uncertainty in the IDA framework is the Kalman-Bucy filter. The resulting method would simultaneously run the filter and IDA (with a moving window in time), where  $x(\cdot)$  becomes the estimate calculated by the filter. The implementation of an example that includes this combination of methods is a current direction of future work.

Another direction of future work is addressing the possibility of impulsively switching between the trajectories of two objects that are not the object of interest. In this paper we have made the assumption that the only impulses that occur are those between the object of interest and another

object nearby. A method similar to gating [3] is currently being developed to handle this possibility.

## 6. ACKNOWLEDGMENTS

The authors would like to thank each of the reviewers as well as Benjamin Tovar for their help with this paper.

## 7. REFERENCES

- [1] B. Anderson and J. B. Moore. *Optimal Control: Linear Quadratic Methods*. Dover Publications, Inc., 1990.
- [2] T. Bailey and H. Durrant-Whyte. Simultaneous Localization and Mapping (SLAM): Part II. *Robotics Automation Magazine, IEEE*, 13(3):108–117, Sep. 2006.
- [3] Y. Bar-Shalom and T. E. Fortmann. *Tracking and Data Association*. Academic Press, Inc., 1988.
- [4] Y. Bar-Shalom, F. Daum, and J. Huang. The Probabilistic Data Association Filter. *IEEE Control Systems Magazine*, 29(6):82–100, 2009.
- [5] N. Bergman and A. Doucet. Markov Chain Monte Carlo Data Association Target Tracking Problems. In *IEEE Conference on Decision and Control*, pages 735–742, 2004.
- [6] S. S. Blackman. Multiple Hypothesis Tracking for Multiple Target Tracking. *IEEE Aerospace and Electronic Systems Magazine*, pages 5–18, 2004.
- [7] T. Caldwell and T. Murphey. Switching Mode Generation and Optimal Estimation with Application to Skid-Steering. *Automatica*, 47(1):50–64, 2011.
- [8] K. Chang, S. Mori, and C. Chong. Evaluating a Multiple-Hypothesis Multi-Target Tracking Algorithm. *IEEE Transactions on Aerospace and Electronic Systems*, 30(2):578–590, 1994.
- [9] C. T. Chen. *Linear Systems Theory and Design, 3rd ed.* Oxford University Press, 1999.
- [10] I. Cox. A Review of Statistical Data Association Techniques for Motion Correspondence. *International Journal of Computer Vision*, 10(1):53–66, 1993.
- [11] I. Cox and S. Hingorani. An Efficient Implementation of Reid’s Multiple-Hypothesis Tracking Algorithm and its Evaluation for the Purpose of Visual Tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(2):138–150, 1996.
- [12] M. Egerstedt, Y. Wardi, and F. Delmotte. Optimal Control of Switching Times in Switched Dynamical Systems. In *IEEE Conference on Decision and Control*, pages 2138–2143, 2003.
- [13] J. Hauser. A Projection Operator Approach to the Optimization of Trajectory Functionals. *IFAC World Congress*, 2002.
- [14] J. Hauser and D. G. Meyer. The Trajectory Manifold of a Nonlinear Control System. In *IEEE Conference on Decision and Control*, pages 1034–1039, 1998.
- [15] J. Hauser and A. Saccon. A Barrier Function Method for the Optimization of Trajectory Functionals with Constraints. In *IEEE Conference on Decision and Control*, pages 864–869, 2006.
- [16] M. Kalandros, L. Trailović, L. Y. Pao and Y. Bar Shalom. Tutorial on Multisensor Management and

- Fusion Algorithms for Target Tracking. In *American Control Conference (ACC)*, 2004.
- [17] C. T. Kelley. *Iterative Methods for Optimization*. Society for Industrial and Applied Mathematics, 1999.
- [18] Z. Khan, T. Balch, and F. Dellaert. *Lecture Notes in Computer Science*, chapter : An MCMC-Based Particle Filter for Tracking Multiple Interacting Targets. Springer, Berlin / Heidelberg, Germany, 2004.
- [19] Z. Khan, T. Balch, and F. Dellaert. MCMC Data Association and Sparse Factorization Updating for Real Time Multi-target Tracking with Merged and Multiple Measurements. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(12):1960–72, 2006.
- [20] T. Kurien. *Multi-Target Multi-Sensor Tracking: Advanced Applications*, chapter : Issues in the Design of Practical Multi-Target Tracking Algorithms. Y. Bar-Shalom, Ed. Artech House, Norwood, MA, 1990.
- [21] M. Mariscotti. A Method for Automatic Identification of Peaks in the Presence of Background and its Application to Spectrum Analysis. *Nucl. Instrum. Methods*, 50:309–320, 1967.
- [22] B. Milch, B. Marthi, S. Russell, D. Sontag, D. L. Ong, and A. Kolobov. *BLOG: Probabilistic Models with Unknown Objects*. MIT Press, 1999.
- [23] N. Mtetwa and L. S. Smith. Smoothing and Thresholding in Neuronal Spike Detection. *Neurocomput.*, 69(10-12):1366–1370, 2006.
- [24] S. Oh, S. Russell, and S. Sastry. Markov Chain Monte Carlo Data Association for General Multiple-Target Tracking Problems. In *IEEE Conference on Decision and Control*, 2004.
- [25] S. Oh, S. Russell, and S. Sastry. Markov Chain Monte Carlo Data Association for Multi-Target Tracking. *IEEE Transactions on Automatic Control*, 54(3):481–497, 2009.
- [26] R. Powers and L. Pao. Power and Robustness of a Track-Loss Detector Based on Kolmogorov-Smirnov Tests. In *American Control Conference (ACC)*, pages 3757 – 3764, 2006.
- [27] K. M. Reichard, E. C. Crow, and D. C. Swanson. Automated Situational Awareness Sensing for Homeland Defense. *SPIE System Diagnosis and Prognosis: Security and Condition Monitoring Issues Conference*, pages 64 – 71, 2003.
- [28] D. Reid. An Algorithm for Tracking Multiple Targets. *IEEE Transactions on Automatic Control*, 24.6:843–854, 1979.
- [29] B. Schumitsch, S. Thrun, G. Bradski, and K. Olukotun. The Information Form Data Association Filter. In *Conference on Neural Information Processing Systems*, 2005.