Impulse Optimization for Data Association

Matthew Travers[†], Todd Murphey[†], and Lucy Pao[‡]

Abstract— **This paper presents a new method that addresses measurement origin uncertainty. Measurement origin uncertainty occurs when the object a measurement originated from is not clear. The systems considered contain multiple bodies which are dynamically indistinguishable other than initial conditions. Each measurement originates from one of the bodies in the system. In the past, recursive data association methods have been used to address problems of this nature. A new technique is presented which treats the measurement association problem as a batch post-processing problem. Reformulating the problem as such, it is possible to transform the data association problem into a trajectory optimization problem. From this point of view it is then possible to solve the measurement association problem using first- and second-order optimization algorithms that rely on having first- and second-order derivatives for cost functions that depend on impulsive trajectories.**

I. Introduction

The problem addressed in this paper is that of associating measurements with the objects from which each measurement originated. The systems considered contain multiple objects from which each measurement could have originated. The dynamics for each of the objects from which a measurement is potentially received are indistinguishable other than initial conditions.

A variety of data association algorithms have been developed in the past to address problems of this nature. A majority of these techniques have focused on recursive formulations similar to a Kalman filter [6]. The key difference between the work presented in this paper and most other pre-existing data association methods [2] is that we treat the associations as a batch process. By assuming measurements are continuous (which is accomplished by interpolating the discrete set of measurements), it is possible to perform a "impulse optimization" that singles out the times at which incorrect measurements are received. By optimizing over the continuous measurements it is possible to incorporate every measurement in the same optimization procedure. Hence our algorithm is more analogous to Kalman smoothing than it is to Kalman filtering.

It is possible to treat the continuous measurements as a flow of a dynamical system with impulsive changes in the state. In past data association techniques, impulses have been addressed in terms of impulse response [3] and in terms of inteference [8]. We do not directly deal with filtering response and treat the impulses as inherent system characteristics upon which the data association is actually based (i.e., we do not treat the impulses as noise). Figure 1 illustrates this idea for a deterministic system. There are two objects with linear dynamics moving in close proximity to each other (in 1-D with unstable dynamics for purpose of illustration). The object that we are attempting to track, object 1, has a trajectory represented by the dotted line. A nearby object, object 2, with the same linear dynamics as object 1 but different initial condition has a trajectory represented by the dashed line. The solid line represents the measurements as seen by the sensor. Note that in this example we are assuming that the measurements are deterministic. In general this is not the case. The measurements are originally coming from

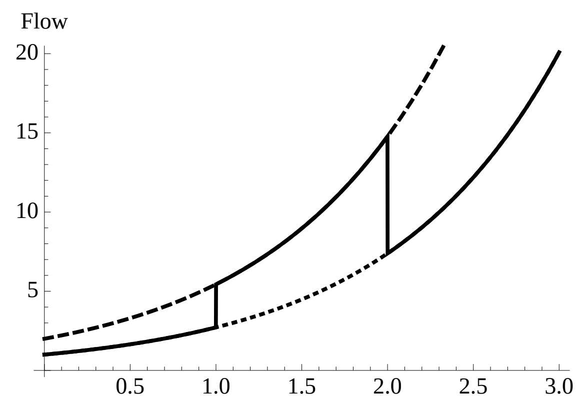

Fig. 1: Example of an impulsive flow generated by a combination of flows from two separate objects. The dotted line represents the flow of one object and the dashed line the flow of a second object. The solid line represents the impulsive measurement flow.

object 1, when at time $t = 1$ the measurements received originate from object 2. At time $t = 2$ there is another "switch" between trajectories and the measurements again originate from object 1. Although the measurements are actually coming from a combination of the two trajectories, we are able to treat the measurement trajectory (solid line) as a single trajectory with impulses. The impulses are a mechanism for switching between the trajectories of objects 1 and 2.

Treating measurements as an impulsive trajectory, it is possible to reformulate the measurement association problem as an optimization procedure. The optimization procedure considers the entire measurement trajectory at one time and optimizes with respect to the times at which impulses occur. By determining the impulse times, it is possible to back out the times at which measurements are received from the object of interest (object 1 in Figure 1).

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

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

The organization of this paper is as follows. The problem is first formally defined in Section II. Adjoint equations for solving the first and second derivatives with respect to impulse times when there are N total impulses are then presented. The adjoint formulations are followed by a short discussion of optimization techniques. Simulated results and comparisons of the impulse optimization method are then presented. Finally, conclusions are made and future directions discussed.

II. Problem Definition

The systems considered in this work are in general nonlinear and have dynamics of the form

$$\dot{x}^i = f_i(x^i(t), t), \quad x^i(0) = x^i_0, \quad i = 1, 2, \dots, M, \tag{1}$$

where M is the total number of objects in the system. The given information is a stream of measurements that serves as a reference trajectory $x_d(t)$, which can be either deterministic or non-deterministic. For the purposes of optimizing, it will be necessary to treat the system as if there were only a single object with nonlinear dynamics that contain impulses

$$\dot{x} = f_i(x(t), t), \quad x(0) = x_0.$$

Recall that in the example from Section I we started by measuring the object of interest, at $t = 1$ measurements originating from object 2 started to be received, then at $t = 2$ measurements originating from object 1 were again received. The flow describing this behavior has the form

$$x(t) = \varphi^3_{t-\tau_2}(\varphi^2_{\tau_2-\tau_1}(\varphi^1_{\tau_1}(x_0) + \delta_1) - \delta_2), \tag{2}$$

where δ_1 and δ_2 are the magnitudes of the impulses. Note that δ_1 and δ_2 are simply the magnitudes of the impulses required to switch between the two trajectories of objects 1 and 2. The times at which the impulses occur are τ_1 and τ_2.

The measurement association problem can be abstracted to finding impulse times that minimize a total cost function:

$$J(\tau_1, \tau_2, \dots, \tau_N) = \int_0^{t_f} \ell(x(s), s)ds, \tag{3}$$

where N is the total number of switching times. The function $\ell(\cdot, \cdot)$ is arbitrary; one possible choice is $\ell(x(t), t) = (x_d(t) - x(t))^T(x_d(t) - x(t))$. Optimization of $J(\cdot)$ is accomplished through standard numerical techniques using a combination of gradient descent and Newton's method.

III. Derivatives of $J(\cdot)$

In this section we derive both the first and second derivatives of the cost with respect to multiple impulse times using adjoint formulations. It will be shown that by calculating adjoint equations it is possible to significantly reduce the computational complexity associated with finding the derivatives. In particular, it will be shown that by using the adjoint formulation, only a single integration is needed to calculate either the first or second derivatives at every impulse time.

While deriving the adjoint equation for the first derivative the following result will be useful. This result is analogous to the result that holds for impulse optimization. The proof for impulse optimization is nearly identical to the proof in [4] and [7] and is thus left out here.

Lemma 1: The first derivative of the trajectory with respect to the impulses times τ_i is

$$D_{\tau_i} x(t) \circ \partial\tau_i = \begin{cases} 0 & t < \tau_i \\ \Phi(t, \tau_i) \circ X^i & t \geq \tau_i \end{cases} \tag{4}$$

$$X^i = (f_{i-1}(x(\tau_i), \tau_i) - f_i(x(\tau_i), \tau_i))\partial\tau_i$$

where, in the general case

$$\dot{x} = f_i(x, t) \quad \tau_i \leq t < \tau_{i+1} \tag{5}$$

and $\Phi(t, \tau_i)$ is the state transition matrix for the system

$$\dot{z} = [D_1 f_i(x(t), t)]z = A(t)z(t).$$

Note that this is a general case. The systems of concern for this work deal with only a single set of dynamics, i.e., $f_i(x, t) = f(x, t)$. Note also that we are able to quote this result which comes from optimizing over switching times for dynamics as opposed to impulse times because the impulses show up linearly in the flow. Thus, the only contribution from the impulses comes from evaluating the dynamics.

Lemma 2: The derivative of the cost function $J(\cdot)$ with respect to each of the impulse times τ_i is

$$D_{\tau_i} J(\cdot) \circ \partial\tau_i = \Psi(t_f, \tau_i) \tag{6}$$

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

$$\Psi(t_f, \tau_i) = \psi(t_f, \tau_i) \circ X^i + \ell(x(\tau_i^-), \tau_i^-) - \ell(x(\tau_i^+), \tau_i^+), \tag{7}$$

and $\psi(t_f, \tau_i) : \mathbb{R}^n \to \mathbb{R}$ is found by integrating

$$\psi(t, t) \circ U = 0 \tag{8a}$$

$$\frac{\partial}{\partial\tau}\psi(t, \tau) \circ U = -D_1\ell(x(\tau), \tau) \circ U - \psi(t, \tau) \circ D_1 f(x(\tau), \tau) \circ U \tag{8b}$$

backward along τ from t_f to τ_i.

Proof: Take the derivative of (3) with respect to τ_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 (4), $D_{\tau_i} x(t) \circ \delta\tau_i = 0$ up until $t = \tau_i$. The result of this fact is that the derivative of the integrand only needs to be integrated from τ_i up to t_f. Thus,

$$D_{\tau_i} J(\cdot) \circ \partial\tau_i = \int_{\tau_i}^{t_f} D_1\ell(x(s), s) \circ D_{\tau_i} x(s) \circ \partial\tau_i ds + \ell(x(\tau_i^-), \tau_i^-) - \ell(x(\tau_i^+), \tau_i^+). \tag{9}$$

Substituting in (4) and noting that X^i is independent of s, we can define the linear operator $\psi(t, \tau)$ such that

$$D_{\tau_i} J(\cdot) \circ \partial\tau_i = \psi(t, \tau_i) \circ X^i + \ell(x(\tau_i^-), \tau_i^-) - \ell(x(\tau_i^+), \tau_i^+) \tag{10}$$

where

$$\psi(t,\tau)\circ U = \left(\int_\tau^t D_1\ell(x(s),s)\circ\Phi(s,\tau)ds\right)\circ U. \quad (11)$$

Equation (10) gives the first part of Lemma 2. This result provides a method for calculating the first derivative of the cost function $J(\cdot)$ with respect to each of the impulse times $\tau_1, \tau_2, \ldots, \tau_N$, but each of these derivatives requires recalculating the value of $\psi(t,\tau)$, i.e., computing the integral in (9). In previous work [5] it has been noted that taking the derivative of (11) with respect to τ yields the equation:

$$\begin{aligned}\frac{\partial}{\partial\tau}\psi(t,\tau)\circ U &= -D_1\ell(x(\tau),\tau)\circ U \\ &\quad - \int_\tau^t D_1\ell(x(s),s)\circ\Phi(s,\tau)\circ A(\tau)U ds && (12a)\\ &= -D_1\ell(x(\tau),\tau)\circ U \\ &\quad - \left(\int_\tau^t D_1\ell(x(s),s)\circ\Phi(s,\tau)ds\right)\circ A(\tau)\circ U && (12b)\\ &= -D_1\ell(x(\tau),\tau)\circ U - \psi(t,\tau)\circ D_1 f(x(\tau),\tau)\circ U && (12c)\end{aligned}$$

Equation (12c) along with evaluating ψ in (11) at $\tau = t$ yield the final two parts of the Lemma. ■

The adjoint equation, (8b), and its terminal condition, (8a), are helpful for several reasons. The first is that the adjoint equation does not include any flow notation. The second and most important is that a terminal condition exists and the adjoint equation can be integrated backward along τ. In particular, ψ can be integrated backward from t_f to τ_1. Thus, in a single integration it is possible to calculate the value of Ψ at each τ_i.

To find the second derivative of the cost function (3), we will proceed in nearly the same way as in finding the first derivative. In particular, the second derivative of the trajectory (5) will be calculated first directly. This calculation will be used to show that the second derivative of the trajectory can be expressed as an affine linear system (note that the first derivative of the trajectory could be expressed as a purely linear system). The second derivative of the trajectory will then be substituted into the second derivative of the cost to yield a general equation for the derivative. Then, similar to the first derivative, an adjoint equation will be found such that when integrated backward along τ reduces the calculation of the second derivative of the cost with respect to each impulse time down to a single integration.

The following result, which has been derived previously [4], [7], will be used to express the second derivative of the trajectory. We assume $i > j$ without loss of generality.

Proposition 3:

$$\begin{aligned}\frac{d}{dt}D_{\tau_j}D_{\tau_i}x(t)\circ(\partial\tau_j,\partial\tau_i) &= \\ D_1 f(x(t),t)\circ D_{\tau_j}D_{\tau_i}x(t)&\circ(\partial\tau_j,\partial\tau_i) \\ + D_1^2 f(x(t),t)\circ(D_{\tau_j}x(t)\circ\partial\tau_j,&\, D_{\tau_i}x(t)\circ\partial\tau_i) \quad (13a)\end{aligned}$$

$$D_{\tau_j}D_{\tau_i}x(\tau_i)\circ(\partial\tau_j,\partial\tau_i) = \begin{cases} D_1 f_i(x(\tau_i),\tau_i)\circ f_i(x(\tau_i),\tau_i)\partial\tau_j\partial\tau_i \\ \quad +D_1 f_{i-1}(x(\tau_i),\tau_i)\circ f_{i-1}(x(\tau_i),\tau_i)\partial\tau_j\partial\tau_i \\ \quad -2D_1 f_i(x(\tau_i),\tau_i)\circ f_{i-1}(x(\tau_i),\tau_i)\partial\tau_j\partial\tau_i \\ \quad +D_2 f_{i-1}(x(\tau_i),\tau_i)\circ\partial\tau_j\partial\tau_i \\ -D_2 f_i(x(\tau_i),\tau_i)\circ\partial\tau_j\partial\tau_i, & i=j \\ (D_1 f_{i-1}(x(\tau_i),\tau_i)\partial\tau_j \\ -D_1 f_i(x(\tau_i),\tau_i))\circ\Phi(\tau_i,\tau_j)\circ X^j\partial\tau_i & i>j\end{cases} \quad (13b)$$

Note that the proof of this result is again analogous to the proof for impulse optimization. Combining the result in Proposition 3 with the fundamental theorem of calculus leads to the following.

Lemma 4: The second derivative $D_{\tau_j}D_{\tau_i}x(t)\circ(\partial\tau_j,\partial\tau_i)$ is

$$\begin{aligned}D_{\tau_j}D_{\tau_i}x(t)&\circ(\partial\tau_j,\partial\tau_i) = \\ &\Phi(t,\tau_i)\circ X^{i,j} + \phi(t,\tau_i)(\Phi(\tau_i,\tau_j)\circ X^j, X^i) \quad (14)\end{aligned}$$

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

$$\begin{aligned}&\phi(t,\tau)\circ(U,V) = \\ &\int_\tau^t \Phi(t,s)\circ D_1^2 f(x(s),s)\circ(\Phi(s,\tau)\circ U, \Phi(s,\tau)\circ V)ds \quad (15)\end{aligned}$$

and $X^{i,j}$ is the intial condition from (13b). This result has also previously been derived [4], [7].

The results from Proposition 3 and Lemma 4 are both used in the proof of the following theorem which calculates the second derivative of the cost. In the proof of the theorem, the second derivative of the cost is first derived explicitly. An adjoint equation is then found which eliminates the need to explicitly calculate $\phi(t,\tau)$. This is very similar to the adjoint equation for the first derivative which eliminates the need to explicitly calculate $\Phi(t,\tau)$.

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

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

where δ_i^j is the Kronecker delta and $\Omega(t,\tau)\circ(U,V):\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ is the bilinear operator found by integrating

$$\begin{aligned}&\Omega(t,t)\circ(U,V) = 0_{n\times n} && (16a)\\ &\frac{\partial}{\partial\tau}\Omega(t,\tau)\circ(U,V) = -D_1^2\ell(x(\tau),\tau)\circ(U,V) \\ &- \psi(t,\tau)\circ D_1^2 f(x(\tau),\tau)\circ(U,V) - \\ &\Omega(t,\tau)\circ(D_1 f(x(\tau),\tau)\circ U, V) \\ &\quad - \Omega(t,\tau)\circ(U, D_1 f(x(\tau),\tau)\circ V) && (16b)\end{aligned}$$

backwards over τ from t_f to τ_i.

Proof: Take the derivative of (9) with respect to τ_j:

$$
\begin{aligned}
&D_{\tau_j} D_{\tau_i} J(\cdot) \circ (\partial\tau_j, \partial\tau_i) = \\
&\quad \frac{\partial}{\partial \tau_j}\Big(\int_{\tau_i^+}^{t_f} D_1\ell(x(s), s) \circ D_{\tau_i} x(s) \circ \partial\tau_i ds \\
&\qquad + \ell(x(\tau_i^-), \tau_i^-) - \ell(x(\tau_i^+), \tau_i^+)\Big) \\
&= 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(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) \\
&\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 \Phi(s, \tau_i) \circ X^{i,j} \\
&\quad + D_1\ell(x(s), s) \circ \phi(s, \tau_i) \circ (\Phi(\tau_i, \tau_j) \circ X^j, X^i)) ds \\
&\quad + \int_{\tau_i^+}^{t_f} D_1^2\ell(x(s), s) \circ (\Phi(s, \tau_i) \circ \Phi(\tau_i, \tau_j) \\
&\qquad \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 \\
&\quad - f(x(\tau_i^-), \tau_i^-)) - D_1\ell(x(\tau_i^+), \tau_i^+) \\
&\quad \circ (D_{\tau_j} x_d(\tau_i^+) \circ \partial\tau_j \delta_j^i - f(x(\tau_i^+), \tau_i^+)) \\
&\quad - D_1\ell(x(\tau_i^+), \tau_i^+) \circ X^i \partial\tau_j \delta_i^j + \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} \tag{17}
$$

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

$$
\begin{aligned}
\Omega(t, \tau) \circ (U, V) = &\\
&\int_{\tau}^{t} D_1\ell(x(s), s) \circ \phi(s, \tau) \circ (U, V) \\
&+ D_1^2\ell(x(s), s) \circ (\Phi(s, \tau) \circ U, \Phi(s, \tau) \circ V) ds.
\end{aligned} \tag{18}
$$

The calculation in (17) provides a method for calculating the second derivative of the cost function $J(\cdot)$ using forward integration. Similar to the first derivative of the cost, this method produces the correct calculation but is computationally costly due to the fact that a separate forward integration is required for each derivative.

The computational burden of calculating the second derivative of the cost function using forward integration can be avoided by finding an adjoint equation for the bilinear operator $\Omega(t, \tau)$. This is accomplished by taking the derivative of (18) with respect to τ, as in (16b). Equation (16b) can then be integrated backwards from t_f to τ_1 using the initial/terminal condition (16a). The value of $\Omega(\cdot, \cdot)$ and thus each component of the second derivative of the cost with respect to each impulse time can be calculated in a single backward integration.

■

Note that it is somewhat helpful to write (16b) in matrix form to facilitate making numerical calculations

$$
\begin{aligned}
&\frac{\partial}{\partial \tau}[\Omega(t, \tau)] = \\
&\quad - [D_1^2\ell(x(\tau), \tau)] - [\psi(t, \tau) \circ D_1^2 f(x(\tau), \tau)] \\
&\quad - [D_1 f(x(\tau), \tau)]^T [\Omega(t, \tau)] - [\Omega(t, \tau)][D_1 f(x(\tau), \tau)].
\end{aligned}
$$

Note also that the impulses do not show up directly in the calculation of either $\psi(t, \tau)$ or $\Omega(t, \tau)$. This is again due to the fact that the impulses enter the flow linearly. Thus the only contribution due to the impulses comes from evaluating the derivatives $\ell(x(t), t)$ and $f(x(t), t)$ along $x(t)$.

IV. Optimization Algorithms

The derivatives calculated in Section III are for use in optimizing (3). First and second-order iterative methods [9] are relied on to accomplish the optimization. Both the first and second-order methods have iterates that take the form

$$x_{k+1} = x_k + \alpha_k z_k, \tag{19}$$

where x_k is the position of the current step, x_{k+1} the position of the next step, z_k contains the descent information, and α_k is a step size parameter. The parameter z_k has the general form $z = -[H]^{-1}[D_{\tau_i} J(\cdot)]^T$ where H is a positive definite matrix. In first-order methods, $H = I$, where I is the identity matrix. This choice of H results in a steepest descent algorithm. Choosing $H = D_{\tau_j} D_{\tau_i} J(\cdot)$ results in Newton's method which is a second-order optimization algorithm.

In many of the systems of interest for this problem $D_{\tau_j} D_{\tau_i} J(\cdot)$ will not be positive definite. In this case it is necessary to implement a *quasi-Newton's method*. There are a variety of choices of *quasi-Newton's methods*, the one chosen in this work is as follows [9], [10]: the Hessian is decomposed into matrices containing the eigenvectors and a matrix containing the eigenvalues. The eigenvalues that are either close to zero or negative are replaced with a value of unity. The Hessian is then recomposed using the original eigenvectors with the modified matrix of "eigenvalues." Note that by replacing an eigenvalue of the original Hessian with a unity value, we are essentially performing steepest descent in the associated direction.

It was mentioned above that α_k in (19) is a step size parameter. The value of α_k is chosen using the *Armijo Line Search* algorithm [1]. This algorithm uses a line search technique to ensure a sufficient decrease. The sufficient decrease condition guarantees that the optimization will eventually converge (assuming no issue with numerical precision).

V. Example

In this section we consider several different examples that all deal with the same system. The example system chosen has one-dimensional dynamics

$$\dot{x}(t) = \sin(x(t))$$
$$y(t) = h(x(t)) \tag{20}$$

where $y(t)$ are the measurements. In the first example the function $h(\cdot)$ in (20) is deterministic. In the second two examples the function $h(\cdot)$ contains a noise term sampled from the normal distribution $N(0, \sigma^2)$. Note that it is possible to apply the impulse optimization results to either deterministic or non-deterministic systems due to the fact that the analytical results derived in this paper depend only on the local convexity of the cost function.

The overall system contains two dynamically identical objects except for initial conditions. Measurements initially originate from the object of interest, object 1. At time $t = 1$ measurements begin to originate from object 2. Measurements originate from object 2 up until $t = 2$, at which point they again originate from object 1. A cost function of the form (3) is defined where $\ell(x(t), t) = (x_d(t) - x(t))^2$ and $x_d(t)$ represents the measurements $y(t)$.

Figure 2 shows a convergence plot for a second-order method for a single-impulse deterministic system. The vertical axis in Figure 2 is the cost on a logarithmic scale. The horizontal axis is iteration number. Figure 2 shows quadratic

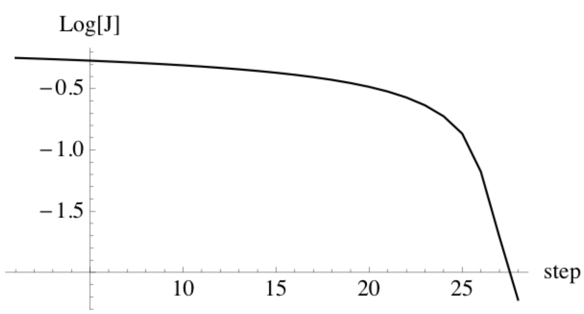

Fig. 2: Quadratic convergence using a second-order method for a single-impulse deterministic system.

convergence for this case where the measurements were assumed to be deterministic.

A natural extension of the deterministic case above is to systems that contain noise. Figure 3 shows a plot of the types of signals being considered (note that the deterministic flow is also plotted). Figure 4 shows three different cost functions associated with varying degrees of noise (the dashed line contains the most noise and the solid line the least). Note that the general trend for these costs functions is the more noise in the measurement signal the higher the cost values.

Figure 5 shows three convergence plots that correspond to the three cost functions in Figure 4, all three starting at $\tau_2 = 1.2$. The scaling on the vertical axes of Figures 4 and 5 have the same linear scaling, not logarithmic scaling. The horizontal axis in Figure 5 is iteration number. Note

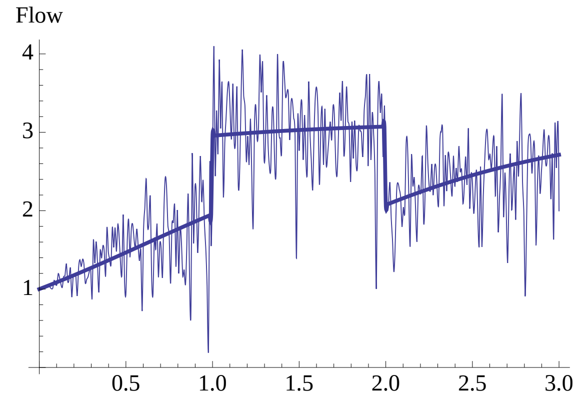

Fig. 3: Deterministic and stochastic flows for a nonlinear system with two impulses.

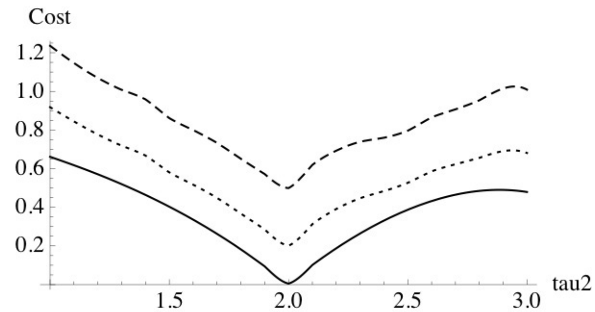

Fig. 4: One dimensional cross-sections of several cost functions for a stochastic linear system with two impulses and varying degrees of noise, where $\tau_1 = 1$. The dashed line represents the case with the highest amount of noise and the solid line the least.

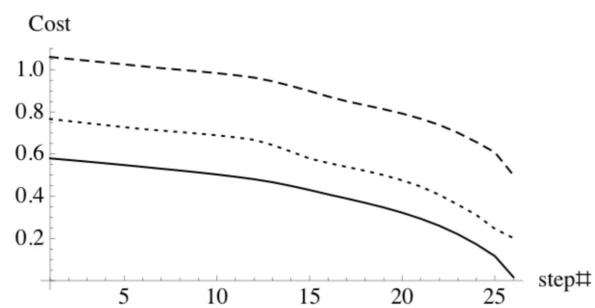

Fig. 5: Convergence plots for the three cost functions shown in Figure 4 using a seond-order method.

that although the final cost values are different, based on the amount of noise present, the value of τ_2 at which the minimum is attained as well as the number of steps (with the same initial guess) it takes to get there remains unchanged by varying the amount of noise present. Note also that although the convergence plots in Figure 5 were not plotted on a logarithmic scale, quadratic convergence is achieved in these cases as well.

Figure 6 shows a convergence plot for the optimization over both of the impulses in the system discussed above (the true minimum is at $(\tau_1, \tau_2) = (1, 2)$). The reference signal for this example is the noisy signal in Figure 3. The starting point for this optimization was $(\tau_1, \tau_2) = (0.7, 2.4)$. This plot shows that when optimizing over both impulses, quadratic convergence is achieved when using a second order method.

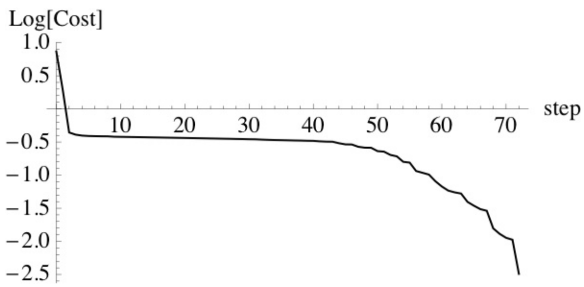

Fig. 6: Convergence plot for second-order optimization over two impulses where the reference signal is the noisy signal in Figure 3.

VI. Conclusions and Future Works

This paper presents a new method for measurement association in systems with measurement origin uncertainty. The method considers impulsive flows as a model for continuous measurements. The measurement associations are analogous to optimizing a cost function with respect to the times at which impulse times occur. By determining the impulse times it is possible to determine the time periods over which incorrect measurements are being received.

The optimization of the cost function is accomplished using both first- and second-order optimization algorithms. These optimization algorithms relies on analytical first- and second-order derivatives. Results for calculating these derivatives have been provided using both flow and adjoint formulations.

Simulated results of applying impulse optimization to a nonlinear system have been provided in Section V. Several different cases with varying degrees of noise have presented. It has been shown explicitly that in the case of deterministic measurements quadratic convergence is achieved. It has also been shown that when the measurement signal incorporated noise, the location and ability to reach the minimum of the noisy cost functions does not change by varying the noise within a reasonable amount.

Several different future directions of this work are currently being considered. The first is to incorporate the impulse magnitude into the optimization procedure. In the work presented in this paper the impulse magnitude has thus far been treated as a constant (the exact amount needed to switch between the trajectories of various objects). By incorporating the impulse magnitudes, fewer assumptions need to be made and a wider variety of systems can be considered. A second direction of future works is to find an analytical bound on the amount of noise that can be present such that convergence of the optimization is still achieved and to create the same relationship between the stochastic analysis and optimization that the Kalman-Bucy filter creates for smooth systems.

References

[1] L. Armijo. Minimization of Functions having Lipschitz Continuous First-Partial Derivatives. *Pacific Journal of Mathematics*, 16:1–3, 1966.

[2] Y. Bar-Shalom and T.E. Fortmann. *Tracking and Data Association*. Academic Press, Inc., 1988.

[3] S. Beheshti and M. A. Dahleh. Noisy data and impulse response estimation. *Trans. Sig. Proc.*, 58(2):510–521, 2010.

[4] T. Caldwell and T.D. Murphey. Switching mode generation and optimal estimation with application to skid-steering. *Automatica*, 2010. Accepted for Publication.

[5] M. Egerstedt, Y. Wardi, and F. Delmotte. Optimal Control of Switching Times in Switched Dynamical Systems. *IEEE Conference on Decision and Control*, 2003.

[6] A. Gelb. *Applied Optimal Estimation*. MIT Press, 1974.

[7] E. Johnson and T.D. Murphey. Second order switching time optimization for time-varying nonlinear systems. In *Conference on Decision and Control*, pages 5281 – 5286, 2009.

[8] C. Kabakchiev, I. Garvanov, L. Doukovska, V. Kyovtorov, and H. Rohling. Data Association Algorithm in Multi-Radar Systems. In *IEEE Radar Conference*, Rome, Italy, 2008.

[9] C.T. Kelly. *Iterative Methods for Optimization*. SIAM, 1999.

[10] J. Nocedal and S. Wright. *Numerical Optimization*. Springer, 2000.