AN UNSCENTED TRANSFORMATION FOR CONDITIONALLY LINEAR MODELS

Mark R. Morelande and Bill Moran

Melbourne Systems Laboratory
Department of Electrical and Electronic Engineering
The University of Melbourne, Australia
email: m.morelande@ee.unimelb.edu.au, b.moran@ee.unimelb.edu.au

ABSTRACT

A new method of applying the unscented transformation to conditionally linear transformations of Gaussian random variables is proposed. This method exploits the structure of the model to reduce the required number of sigma points. A common application of the unscented transformation is to nonlinear filtering where it used to approximate the moments required in the Kalman filter recursion. The proposed procedure is applied to a nonlinear filtering problem which involves tracking a falling object.

Keywords: Nonlinear filters, Kalman filtering.

1. INTRODUCTION

The unscented transformation (UT) is a method for approximating, in a computationally efficient manner, the moments of a random variable which has undergone a nonlinear transformation [7]. This is done by selecting a number of sigma points, passing them through the transformation and collecting the sample statistics of the transformed sigma points. The idea is similar to importance sampling except the UT selects samples deterministically rather than randomly.

The UT finds a useful application in nonlinear filtering as a means of approximating the Kalman filter (KF). Although the KF is the minimum mean square error estimator only for linear/Gaussian dynamic systems, it's use for nonlinear systems is desirable because it is computationally efficient and usually performs well [8]. In general the moments required in the KF recursion cannot be computed exactly. A popular method of approximating these moments is to linearise about a certain point and then use the usual KF recursion. This is known as the extended KF (EKF) [4]. The EKF performs adequately in many situations but has a tendency to underestimate the estimator covariance matrix. This can lead to divergence. The UKF, which uses UT moment approximations, generally avoids this problem [8]. Additional advantages of the UKF over the EKF are that it is easier to apply as it does not require derivation of the Jacobian of the transformation, and it can be more widely applied, e.g., to discontinuous transformations.

A basic requirement of the sigma points used in the UT is that the sample mean and covariance matrix of the sigma points should equal the mean and covariance matrix of the variable being transformed [7]. There are various ways this can be done. The basic implementation of the UT uses $2n+1$ sigma points for transformation of a n-dimensional random variable and includes a parameter for controlling the spread of the sigma points [7]. The scaled UT offers increased control of the sigma point distribution through the inclusion of additional parameters [5]. A method using $n+2$ sigma points, referred to as the simplex sigma points, has also been proposed in [6]. Alternatively, improved accuracy can be obtained through the use of an increased number of sigma points. This was shown in [13] for transformations of scalar variables. Another approach is to use numerical integration rules which are exact for monomials up to a certain order [9]. For instance, the basic scheme, using $2n+1$ points, is exact for monomials up to degree three. A scheme which is exact for monomials up to degree five requires $2n^2+1$ points.

A common element of the various methods of sigma point selection described above is that the number of sigma points is determined by the dimension of the random variable undergoing the transformation. In [10] it was shown that this requirement can be relaxed if the transformation is linear in some elements of the argument and nonlinear in others. For these partially linear transformations, the joint moments required in the KF can be accurately approximated using a number of sigma points depending on the number of elements which are nonlinearly transformed. A deficiency of the procedure proposed in [10] is that it does not apply to the important class of conditionally linear transformations, i.e., transformations which, with some of the elements of the argument held constant, are linear in the remaining elements [12]. This is a more general class of transformations which includes the partially linear transformations as a special case. Several applications which fit this class of models are given in [12]. In this paper a new method of performing a reduced point UT which is applicable when a conditionally linear, or more accurately affine, transformation is applied to a Gaussian random variable is proposed. The procedure has parallels with the process of Rao-Blackwellization in particle filtering [11] and applies to the same models. This link will be discussed in more detail in the body of the paper.

The paper is organised as follows. The UT is reviewed in Section 2. The proposed UT for conditionally linear transformations is described in Section 3 and applied to a filtering problem in Section 4. Conclusions are drawn in Section 5.

2. THE UNSCENTED TRANSFORMATION

Consider a random variable $\boldsymbol{x} \in \mathbb{R}^n$ subject to a nonlinear transformation $\boldsymbol{g} : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Letting $\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x})$, it is desired to find the moments $\boldsymbol{\mu}_Y \triangleq \mathsf{E}(\boldsymbol{y})$, $\boldsymbol{\Sigma}_Y \triangleq \mathsf{cov}(\boldsymbol{y},\boldsymbol{y})$ and $\boldsymbol{\Sigma}_{XY} \triangleq \mathsf{cov}(\boldsymbol{x},\boldsymbol{y})$. These are the moments which are required if the UT is used as part of the KF.

The expectation of $\boldsymbol{y}$ can be expanded as

$$\boldsymbol{\mu}_Y = \int \mathsf{E}(\boldsymbol{y}|\boldsymbol{x}) p(\boldsymbol{x})\, d\boldsymbol{x} = \int \boldsymbol{g}(\boldsymbol{x}) p(\boldsymbol{x})\, d\boldsymbol{x}. \qquad (1)$$

This work was supported by the Defense Advanced Research Projects Agency of the US Department of Defense and was monitored by the Office of Naval Research under Contract No. N00014-04-C-0437.

The notation p will be used to denote probability density functions (pdfs). The particular pdf under consideration will be clear from the arguments. In many cases of practical interest the integral (1) cannot be computed exactly so an approximation is required. The UT performs this approximation as follows. A collection of s sigma points $\mathcal{X}^i$, $i = 1, \ldots, s$ are selected and given a weighting with the weight for the ith sigma point denoted as w^i, $i = 1, \ldots, s$. The sigma points and weights are such that the sample mean and covariance matrix are equal to the mean and covariance matrix of $\boldsymbol{x}$. The sigma points are passed through the transformation $\boldsymbol{g}$ to obtain $\mathcal{Y}^i = \boldsymbol{g}(\mathcal{X}^i)$, $i = 1, \ldots, s$. The desired expectation (1) is then approximated by the sample mean

$$\hat{\boldsymbol{\mu}}_Y = \sum_{i=1}^{s} w^i \mathcal{Y}^i. \tag{2}$$

The remaining moments involve the evaluation of similarly intractable integrals. They can be approximated as

$$\hat{\boldsymbol{\Sigma}}_Y = \sum_{i=1}^{s} w^i \left(\mathcal{Y}^i - \hat{\boldsymbol{\mu}}_Y\right)\left(\mathcal{Y}^i - \hat{\boldsymbol{\mu}}_Y\right)', \tag{3}$$

$$\hat{\boldsymbol{\Sigma}}_{XY} = \sum_{i=1}^{s} w^i \left(\mathcal{X}^i - \boldsymbol{\mu}_X\right)\left(\mathcal{Y}^i - \hat{\boldsymbol{\mu}}_Y\right)', \tag{4}$$

where $'$ denotes the matrix transpose. As discussed in the Introduction, there are several ways of choosing the sigma points and weights [8].

3. THE UNSCENTED TRANSFORMATION FOR CONDITIONALLY LINEAR TRANSFORMATIONS

The transformation $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^m$ is applied to the random variable $\boldsymbol{x} \in \mathbb{R}^n$ with the following assumptions:

A1 $\boldsymbol{x}$ is Gaussian with mean $\boldsymbol{\mu}_X$ and covariance matrix $\boldsymbol{\Sigma}_X$.

A2 The random variable $\boldsymbol{x}$ partitions as $(\boldsymbol{a}', \boldsymbol{b}')'$ such that the transformation $\boldsymbol{g}$ can be written as:

$$\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{d}(\boldsymbol{a}) + \boldsymbol{G}(\boldsymbol{a})\boldsymbol{b}. \tag{5}$$

While assumption A1 may appear restrictive it is not unreasonable when the KF is applied to nonlinear filtering problems. In such cases it is customary to employ the maximum entropy principle and take the posterior density to be Gaussian [3]. Assumption A2 states that conditional on $\boldsymbol{a}$ the transformation $\boldsymbol{g}$ is linear (affine) in $\boldsymbol{b}$. In this section the structure implied by these assumptions will be exploited to reduce the number of sigma points required to approximate the joint moments of $\boldsymbol{y} = g(\boldsymbol{x})$ and $\boldsymbol{x}$.

The expectation of $\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x})$ can be expanded as

$$\boldsymbol{\mu}_Y = \int \mathsf{E}\left(\boldsymbol{y}|\boldsymbol{a}, \boldsymbol{b}\right) p(\boldsymbol{b}|\boldsymbol{a}) p(\boldsymbol{a})\, d\boldsymbol{b}\, d\boldsymbol{a}. \tag{6}$$

The mean and covariance matrix of $\boldsymbol{x}$ are partitioned as

$$\boldsymbol{\mu}_X = \begin{bmatrix} \mathsf{E}(\boldsymbol{a}) \\ \mathsf{E}(\boldsymbol{b}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_A \\ \boldsymbol{\mu}_B \end{bmatrix}, \tag{7}$$

$$\boldsymbol{\Sigma}_X = \begin{bmatrix} \mathsf{cov}(\boldsymbol{a}, \boldsymbol{a}) & \mathsf{cov}(\boldsymbol{a}, \boldsymbol{b}) \\ \mathsf{cov}(\boldsymbol{b}, \boldsymbol{a}) & \mathsf{cov}(\boldsymbol{b}, \boldsymbol{b}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_A & \boldsymbol{\Sigma}_{AB} \\ \boldsymbol{\Sigma}'_{AB} & \boldsymbol{\Sigma}_B \end{bmatrix}. \tag{8}$$

Then, assumption A1 and the matrix inversion lemma yield [1]

$$\boldsymbol{\nu}_B(\boldsymbol{a}) \triangleq \mathsf{E}(\boldsymbol{b}|\boldsymbol{a}) = \boldsymbol{\mu}_B + \boldsymbol{\Sigma}'_{AB}\boldsymbol{\Sigma}_A^{-1}(\boldsymbol{a} - \boldsymbol{\mu}_A), \tag{9}$$

$$\boldsymbol{\Gamma} \triangleq \mathsf{cov}(\boldsymbol{b}|\boldsymbol{a}) = \boldsymbol{\Sigma}_B - \boldsymbol{\Sigma}'_{AB}\boldsymbol{\Sigma}_A^{-1}\boldsymbol{\Sigma}_{AB}. \tag{10}$$

Substitution of (5) into (6) combined with (9) gives

$$\boldsymbol{\mu}_Y = \int \boldsymbol{\nu}_Y(\boldsymbol{a}) p(\boldsymbol{a})\, d\boldsymbol{a}. \tag{11}$$

where $\boldsymbol{\nu}_Y(\boldsymbol{a}) \triangleq \mathsf{E}(\boldsymbol{y}|\boldsymbol{a}) = \boldsymbol{d}(\boldsymbol{a}) + \boldsymbol{G}(\boldsymbol{a})\boldsymbol{\nu}_B(\boldsymbol{a})$. Note that (11) has the same form as (1) with the difference that A1 and A2 have been exploited to reduce the dimension of the numerical problem. The UT approximation to (11) proceeds by selecting s sigma points $\mathcal{A}^1, \ldots, \mathcal{A}^s$ along with weights $w^1, \ldots, w^s$. These sigma points can be found using any of the methods discussed previously. The transformed sigma points are calculated as

$$\mathcal{Y}^i = \boldsymbol{\nu}_Y(\mathcal{A}^i), \quad i = 1, \ldots, s, \tag{12}$$

and substituted in (2) to approximate the expectation $\boldsymbol{\mu}_Y$. Note that, for a given sigma point selection scheme, the number s of sigma points required to approximate $\boldsymbol{\mu}_Y$ via (11) will be smaller than that required by direct approximation via (1) due to the reduced dimension of the integration variable.

The covariance matrix $\boldsymbol{\Sigma}_Y$ of $\boldsymbol{y}$ can be expanded as

$$\begin{aligned}
\boldsymbol{\Sigma}_Y &= \int \left(\boldsymbol{y} - \boldsymbol{\mu}_Y\right)\left(\boldsymbol{y} - \boldsymbol{\mu}_Y\right)' p(\boldsymbol{y})\, d\boldsymbol{y} \\
&= \int \left(\boldsymbol{y} - \boldsymbol{\mu}_Y\right)\left(\boldsymbol{y} - \boldsymbol{\mu}_Y\right)' p(\boldsymbol{y}|\boldsymbol{x}) p(\boldsymbol{x})\, d\boldsymbol{y}\, d\boldsymbol{x} \\
&= \int \left[\boldsymbol{d}(\boldsymbol{a}) + \boldsymbol{G}(\boldsymbol{a})\boldsymbol{b} - \boldsymbol{\mu}_Y\right]\left[\boldsymbol{d}(\boldsymbol{a}) + \boldsymbol{G}(\boldsymbol{a})\boldsymbol{b} - \boldsymbol{\mu}_Y\right]' \\
&\qquad \times p(\boldsymbol{b}, \boldsymbol{a})\, d\boldsymbol{b}\, d\boldsymbol{a} \\
&= \int \left\{\left[\boldsymbol{\nu}_Y(\boldsymbol{a}) - \boldsymbol{\mu}_Y\right]\left[\boldsymbol{\nu}_Y(\boldsymbol{a}) - \boldsymbol{\mu}_Y\right]' + \boldsymbol{G}(\boldsymbol{a})\boldsymbol{\Gamma}\boldsymbol{G}(\boldsymbol{a})'\right\} \\
&\qquad \times p(\boldsymbol{a})\, d\boldsymbol{a}.
\end{aligned} \tag{13}$$

The UT approximation to (13) can then be calculated using the sigma points $\mathcal{A}^1, \ldots, \mathcal{A}^s$ and the transformed sigma points of (12):

$$\hat{\boldsymbol{\Sigma}}_Y = \sum_{i=1}^{s} w^i \left\{(\mathcal{Y}^i - \hat{\boldsymbol{\mu}}_Y)(\mathcal{Y}^i - \hat{\boldsymbol{\mu}}_Y)' + \boldsymbol{G}(\mathcal{A}^i)\boldsymbol{\Gamma}\boldsymbol{G}(\mathcal{A}^i)'\right\}. \tag{14}$$

The cross-covariance matrix $\boldsymbol{\Sigma}_{XY}$ between $\boldsymbol{x}$ and $\boldsymbol{y}$ can be expanded as

$$\begin{aligned}
\boldsymbol{\Sigma}_{XY} &= \int (\boldsymbol{x} - \boldsymbol{\mu}_X)(\boldsymbol{y} - \boldsymbol{\mu}_Y)' p(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{y}\, d\boldsymbol{x} \\
&= \int \left(\begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix} - \boldsymbol{\mu}_X\right)\left[\boldsymbol{d}(\boldsymbol{a}) + \boldsymbol{G}(\boldsymbol{a})\boldsymbol{b} - \boldsymbol{\mu}_Y\right]' \\
&\qquad \times p(\boldsymbol{b}, \boldsymbol{a})\, d\boldsymbol{b}\, d\boldsymbol{a} \\
&= \int \left\{\left(\begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{\nu}_B(\boldsymbol{a}) \end{bmatrix} - \boldsymbol{\mu}_X\right)\left[\boldsymbol{\nu}_Y(\boldsymbol{a}) - \boldsymbol{\mu}_Y\right]' \right. \\
&\qquad \left. + \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{\Gamma}\boldsymbol{G}(\boldsymbol{a})' \end{bmatrix}\right\} p(\boldsymbol{a})\, d\boldsymbol{a}.
\end{aligned} \tag{15}$$

Let $\mathcal{X}^i = [\mathcal{A}^{i\prime}, \boldsymbol{\nu}_B(\mathcal{A}^i)']'$ for $i = 1, \ldots, s$. Then the cross-covariance matrix $\boldsymbol{\Sigma}_{XY}$ between $\boldsymbol{x}$ and $\boldsymbol{y}$ can be approximated as

$$\hat{\boldsymbol{\Sigma}}_{XY} = \sum_{i=1}^{s} w^i \left[\left(\mathcal{X}^i - \boldsymbol{\mu}_X\right)\left(\mathcal{Y}^i - \hat{\boldsymbol{\mu}}_Y\right)' + \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{\Gamma}\boldsymbol{G}(\mathcal{A}^i)' \end{bmatrix}\right]. \tag{16}$$

This completes the derivation of the UT moment approximations for a conditionally linear transformation of a Gaussian random variable. The procedure for computing the moment approximations is summarised in Table 1.

Table 1: UT moment approximation for a conditionally linear transformation.

1. Select sigma points $\mathcal{A}^1, \ldots, \mathcal{A}^s$ and weights $w^1, \ldots, w^s$.
2. Compute $\boldsymbol{\nu}_B(\mathcal{A}^1), \ldots, \boldsymbol{\nu}_B(\mathcal{A}^s)$ and $\boldsymbol{\Gamma}$ using (9) and (10).
3. Compute the transformed sigma points $\mathcal{Y}^1, \ldots, \mathcal{Y}^s$ using (12).
4. Approximate $\boldsymbol{\mu}_Y$, $\boldsymbol{\Sigma}_Y$ and $\boldsymbol{\Sigma}_{XY}$ using (2), (14) and (16), respectively.

4. APPLICATION TO NONLINEAR FILTERING

Consider a stochastic dynamic system with a state at time t_k, $k = 0, 1, \ldots$ denoted as $\boldsymbol{x}_k \in \mathbb{R}^{n_x}$. The state evolves according to

$$\boldsymbol{x}_k = \boldsymbol{f}_k(\boldsymbol{x}_{k-1}) + \boldsymbol{v}_k, \quad k = 1, 2, \ldots, \tag{17}$$

where the process noise $\boldsymbol{v}_k$ is zero-mean and $\mathsf{cov}(\boldsymbol{v}_k, \boldsymbol{v}_k) = \mathbf{Q}_k \delta_{k-l}$. The state at time t_0 has initial distribtuion π_0. The state is observed through the measurement equation

$$\boldsymbol{y}_k = \boldsymbol{h}_k(\boldsymbol{x}_k) + \boldsymbol{w}_k, \qquad k = 1, 2, \ldots, \tag{18}$$

where the measurement noise $\boldsymbol{w}_k$ is zero-mean, $\mathsf{cov}(\boldsymbol{w}_k, \boldsymbol{w}_l) = \boldsymbol{R}_k \delta_{k-l}$ and $\mathsf{cov}(\boldsymbol{w}_k, \boldsymbol{v}_l) = \mathbf{0}$. The goal of the filtering problem is to estimate the state $\boldsymbol{x}_k$ given the observations $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$ for $k = 1, 2, \ldots$. The Kalman filter (KF) is a well-known solution which is optimal in the minimum mean square error sense when $\boldsymbol{f}_k$ and $\boldsymbol{h}_k$ are linear and the process and measurement noises are Gaussian. More generally, for a nonlinear/non-Gaussian stochastic dynamic system the KF provides a computationally efficient estimator which is quite accurate in many cases.

The KF represents the state estimate at time $k-1$ by a mean $\boldsymbol{x}_{k-1|k-1}$ and covariance matrix $\boldsymbol{P}_{k-1|k-1}$ which, for general nonlinear/non-Gaussian models, can be interpreted as approximations of the true posterior mean and covariance matrix. The first step in the KF recursion is to compute the predicted mean and covariance matrix,

$$\hat{\boldsymbol{x}}_{k|k-1} = \mathsf{E}(\boldsymbol{x}_k | \boldsymbol{y}_{1:k-1}), \tag{19}$$

$$\boldsymbol{P}_{k|k-1} = \mathsf{cov}(\boldsymbol{x}_k, \boldsymbol{x}_k | \boldsymbol{y}_{1:k-1}). \tag{20}$$

The predicted mean and covariance matrix are then corrected using the current measurement to give

$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{\Psi}_k \boldsymbol{S}_k^{-1} (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}), \tag{21}$$

$$\boldsymbol{P}_{k|k} = \boldsymbol{P}_{k|k-1} - \boldsymbol{\Psi}_k \boldsymbol{S}_k^{-1} \boldsymbol{\Psi}_k', \tag{22}$$

where

$$\hat{\boldsymbol{y}}_{k|k-1} = \mathsf{E}(\boldsymbol{y}_k | \boldsymbol{y}_{1:k-1}), \tag{23}$$

$$\boldsymbol{S}_k = \mathsf{cov}(\boldsymbol{y}_k, \boldsymbol{y}_k | \boldsymbol{y}_{1:k-1}), \tag{24}$$

$$\boldsymbol{\Psi}_k = \mathsf{cov}(\boldsymbol{x}_k, \boldsymbol{y}_k | \boldsymbol{y}_{1:k-1}). \tag{25}$$

In accordance with the maximum entropy principle, all expectations are taken with respect to Gaussian distributions.

The conditional expectations (19), (20) and (23)-(25) cannot in general be computed exactly. The unscented KF (UKF) approximates each of these moments using the UT. If the functions $\boldsymbol{f}_k$ and $\boldsymbol{h}_k$ have the structure given in (5) then the procedure of Table 1 can be used to reduce the required number of sigma points. This procedure has the same aim as Rao-Blackwellisation in particle filtering, i.e., to reduce the required amount of numerical approximation by evaluating integrals analytically wherever possible. The filter which results from applying the transformation of Table 1 to the KF recursion will thus be referred to as the Rao-Blackwellised UKF (RB-UKF). However, a complete correspondence between the RB-UKF and RB particle filters is not possible because the UKF and the particle filter are fundamentally different approaches to the nonlinear filtering problem. The UKF seeks to approximate the KF recursion and uses samples generated by the UT to approximate the required moments. On the other hand, the samples generated in a particle filter are intended to be propagated through time as an approximation to the posterior density. Thus although both filters compute the conditional moments (9) for each sample point, these conditional moments are used in different ways.

4.1. Filtering example

The RB-UKF will be demonstrated for the problem of tracking a vertically falling body, previously considered in [2, 7]. The quantities to be estimated are the height, velocity and ballistic coefficient of the body. The state vector is $\boldsymbol{x}_k = [x_k, \dot{x}_k, \alpha_k]'$ where x_k is the height of the object, $\dot{x}_k$ is the downward speed and α_k is a ballistic coefficient. The evolution of this state vector cannot be written in the form (17) because the differential equation which governs the motion of the body cannot be discretised exactly. However by splitting the sampling period into short intervals and using a first-order approximation across each interval the evolution of the state can be accurately approximated. Let $\boldsymbol{x}(t)$ denote the state at time t so that $\boldsymbol{x}(t_k) = \boldsymbol{x}_k$ and let $T_k = t_k - t_{k-1}$ denote the sampling period. The interval $[t_{k-1}, t_k]$ is split into $M \geq 1$ sub-intervals. Then, for $t = t_{k-1}, t_{k-1} + T_k/M, \ldots, t_k - T_k/M$,

$$\boldsymbol{x}(t + T_k/M) = \boldsymbol{f}(\boldsymbol{x}(t)) + \boldsymbol{v}(t) \tag{26}$$

where $\mathsf{cov}(\boldsymbol{v}(t), \boldsymbol{v}(\tau)) = \mathbf{Q}\delta_{t-\tau}$ and

$$\boldsymbol{f}(x(t)) = \begin{bmatrix} x(t) - T_k \dot{x}(t)/M \\ \dot{x}(t) - T_k e^{-\gamma x(t)} \dot{x}(t)^2 \alpha(t)/M \\ \alpha(t) \end{bmatrix}. \tag{27}$$

The constant $\gamma = 5 \times 10^{-5}$ relates air density to altitude.

A sensor located at a height h above the ground and a horizontal distance l from the vertically falling body provides noisy measurements of the range of the body. Thus, (18) applies with

$$h_k(\boldsymbol{x}_k) = \sqrt{l^2 + (x_k - h)^2}. \tag{28}$$

The dynamic and measurement equations are conditionally linear for this example. Consider first the dynamic equation (27). For the transition between t and $t + T_k/M$, this is equivalent to (5) with $\boldsymbol{a} = [x(t), \dot{x}(t)]'$, $b = \alpha(t)$ and

$$\boldsymbol{d}(\boldsymbol{a}) = \begin{bmatrix} x(t) - T_k \dot{x}(t)/M & \dot{x}(t) & 0 \end{bmatrix}', \tag{29}$$

$$\boldsymbol{G}(\boldsymbol{a}) = \begin{bmatrix} 0 & T_k e^{-\gamma x(t)} \dot{x}(t)^2/M & 1 \end{bmatrix}'. \tag{30}$$

Thus the procedure of Table 1 can be used to find the moments (19) and (20) at each time $t = t_{k-1} + T_k/M, \ldots, t_{k-1} + T_k = t_k$. If the basic sigma point selection scheme of [7] is used, five sigma points are required rather than the seven which would be required using (2)-(4).

The measurement equation is equivalent to (5) with $a = x_k$, $\boldsymbol{b} = [\dot{x}_k, \alpha_k]'$, $d(a) = \sqrt{l^2 + (x_k - h)^2}$ and $\boldsymbol{G}(a) = [0, 0]$.

With the basic sigma point selection scheme of [7], the moments (23)-(25) can be approximated using three sigma points rather than the seven which would be needed using (2)-(4).

The parameters used here are the same as those used in [7]. The height of the sensor is $h = 100\,000$ft and its horizontal distance from the object is $l = 100\,000$ft. The measurement noise is Gaussian with variance 10^4ft^2. The sampling period is $T_k = 1$s for all k. The object parameters at time t_0 are $x_0 = 300\,000$ft, $\dot{x}_0 = 20\,000\text{fts}^{-1}$ and $\alpha_0 = 10^{-3}\text{ft}^{-1}$. The initial state is Gaussian with mean $[300\,000, 20\,000, 3 \times 10^{-5}]'$ and covariance matrix $\text{diag}(10^6, 4 \times 10^6, 10^{-4})$. The initial estimates of the height and velocity are correct but there is a large error in the initial estimate of the ballistic coefficient. No process noise is used in the dynamic equation. Results presented in [7] show that the EKF performs poorly in this scenario so it will not be considered here.

The RB-UKF and the conventional UKF are applied to the tracking problem with the target dynamics evaluated using $M = 32$ sub-intervals per sampling period. The sigma points for both methods are selected using the basic sigma point selection scheme of [7]. The number of sigma points then required for each filter has been discussed above. Figures 1 and 2 show the RMS errors in height and the ballistic coefficent plotted against time. The errors were obtained by averaging over 100 realisations. The two filters perform similarly with the RB-UKF performing better at the start of the observation interval and worse towards the end. The most significant difference occurs in the estimation of the ballistic coefficient at the beginning of the observation interval. Here the RB-UKF provides significantly more accurate estimates.

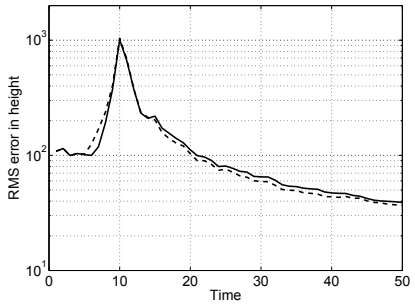

Figure 1: RMS error in height plotted against time for the RB-UKF (solid) and the UKF (dashed).

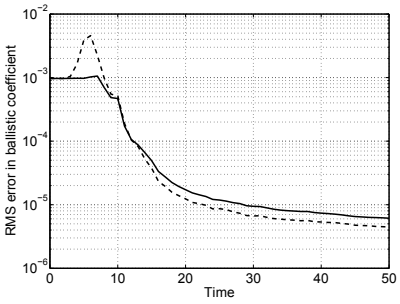

Figure 2: RMS error in ballistic coefficient plotted against time for the RB-UKF (solid) and the UKF (dashed).

5. CONCLUSIONS

A new unscented transformation for conditionally linear functions of Gaussian random variables was proposed. For such functions, the proposed method permits a reduction in the required number of sigma points compared to the conventional approach. The aim of the proposed method is not to improve upon the accuracy of the conventional unscented transformation. Rather, the aim is to provide similar accuracy with a reduced number of sigma points. The example given here demonstrates that the proposed approach is capable of achieving this aim. However, a reduced number of sigma points does not necessarily mean a reduced computational expense since the number of computations per sigma point is higher for the proposed method. Therefore it is anticipated that the proposed approach will find most application in cases where a large proportion of the elements of the argument are linear conditional on the remaining elements or where the expense of computing and transforming sigma points is large.

6. REFERENCES

[1] B. Anderson and J. Moore, *Optimal Filtering*. Prentice Hall, 1979.

[2] M. Athans, R. Wishner, and A. Bertolini, "Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements," *IEEE Transactions on Automatic Control*, vol. 13, no. 5, pp. 504–514, 1968.

[3] D. Catlin, *Estimation, Control and the Discrete Kalman Filter*. New York: Springer-Verlag, 1989.

[4] A. Jazwinski, *Stochastic Processes and Filtering Theory*. Academic Press, 1970.

[5] S. Julier, "The scaled unscented transformation," in *Proceedings of the American Control Conference*, Anchorage, USA, 2002, pp. 4555–4559.

[6] S. Julier and J. Uhlmann, "Reduced sigma point filters for the propagation of means and covariances through nonlinear transformations," in *Proceedings of the American Control Conference*, Anchorage, USA, 2002, pp. 887–892.

[7] S. Julier, J. Uhlmann, and H. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 477–482, 2000.

[8] S. Julier and J. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[9] U. Lerner, "Hybrid Bayesian Networks for Reasoning about Complex Systems," Ph.D. dissertation, Stanford University, 2002.

[10] M. Morelande and B. Ristic, "Reduced sigma point filtering for partially linear models," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, (to appear), 2006.

[11] T. Schon, F. Gustafsson, and P.-J. Nordlund, "Marginalized particle filters for mixed linear/nonlinear state-space models," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2279–2289, 2005.

[12] N. Shephard, "Partial non-Gaussian state space," *Biometrika*, vol. 81, no. 1, pp. 115–131, 1994.

[13] D. Tenne and T. Singh, "The higher order unscented filter," in *Proceedings of the American Control Conference*, Denver, USA, 2003.