Unscented Kalman Filters for Multiple Target Tracking with Symmetric Measurement Equations

William F. Leven, *Student Member, IEEE,*

and Aaron D. Lanterman, *Member, IEEE*

Abstract

The symmetric measurement equation approach to multiple target tracking is revisited using unscented Kalman and particle filters. The characteristics and performance of these filters are compared to the original symmetric measurement equation implementation relying upon an extended Kalman filter. Counterintuitive results are presented and explained for two sets of symmetric measurement equations, including a previously unknown limitation of the unscented Kalman filter. The point is made that the performance of the SME approach is dependent on the interaction of the set of SME equations and the filter used. Furthermore, an SME/unscented Kalman filter pairing is shown to have improved performance versus previous approaches while possessing simpler implementation and equivalent computational complexity. Finally, Taylor series expansions are used to analyze the properties of the SME approach in conjunction with the Kalman filters.

Index Terms

nonlinear filtering, extended Kalman filter, particle filter, Taylor series

I. INTRODUCTION

In the early 1990's, Kamen and Sastry presented a novel approach to multiple target tracking based on symmetric measurement equations (SME) [1], [2]. The underlying idea was to create a "pseudomeasurement" vector $\tilde{y}$ consisting of symmetric functions of the original data y. For example, consider a simple case of tracking three targets in one dimension. Two possible sets of SME's are the *sum of products* and *sum of powers*[1] as follows:

$$\tilde{y}_{prod} = \begin{pmatrix} y_1 + y_2 + y_3 \\ y_1 y_2 + y_1 y_3 + y_2 y_3 \\ y_1 y_2 y_3 \end{pmatrix}, \tag{1}$$

$$\tilde{y}_{pow} = \begin{pmatrix} y_1 + y_2 + y_3 \\ y_1^2 + y_2^2 + y_3^2 \\ y_1^3 + y_2^3 + y_3^3 \end{pmatrix}. \tag{2}$$

Notice that the original y_i's may be rearranged without affecting $\tilde{y}$. It can be shown that the y_i's may be recovered uniquely (up to a permutation) from $\tilde{y}$, so there is no fundamental loss of information. This approach turns the data association problem, which has a discrete character, into an analytic nonlinearity. In this way, one difficult problem is traded for another difficult, but quite different, problem.

The first studies of the SME approach used extended Kalman filters to handle the nonlinearities. In practice, this turned out to have some difficulties. The EKF in this situation often exhibits instability, particularly when targets cross, and can be extremely sensitive to initial state and error covariance values as well as the chosen process noise covariance. The trouble lies not in the underlying SME idea, but in the fragility of the EKF. The EKF is based on first-order Taylor series expansions (linearization) around the state estimate, and the accuracy of these expansions breaks down if the estimated state is too far from the true state. The EKF seems well-suited to handle gentle nonlinearities, but not the sort which arise in the SME approach.

One problem inherent in pairing Kalman filters and the SME approach is that the nonlinear SME transformations produce measurements that cannot be perfectly modelled as signal plus additive Gaussian noise – violating a requirement for Kalman filters. One must find an additive Gaussian noise approximation to the true measurement probability density. The unscented

---

[1]In practice, we subtract some constants to ensure the pseudomeasurements are zero mean, as shown in (66). We suppress those constants here to avoid cluttering the exposition.

Kalman filter (UKF), introduced by Julier and Uhlmann [3] [4], conveniently finds this approximation as part of its basic structure, whereas the EKF requires a pre-computed analytical approximation. The UKF compares favorably to the EKF in two other aspects as well. The UKF, like the EKF, forces the posterior density to be Gaussian, but the posterior mean and covariance are accurate to a third-order Taylor series expansion compared to first-order accuracy for the EKF [4]. Finally, the UKF has the same order of computational complexity as the EKF [5]. With these credentials, the UKF was expected to consistently outperform the EKF.

This paper revisits the SME approach, replacing the extended Kalman filter with unscented Kalman and particle filters. The UKF experiment produced two unexpected results. First, the UKF appears to be fundamentally incompatible with the sum-of-products form of the SME. We show this to be a limitation of the UKF rather than a flaw in the SME approach. Second, we show that the sum-of-powers formulation, when paired with the UKF, outperforms the sum of products paired with the EKF. This contrasts with Kamen's early studies where the sum-of-products form outperformed the sum of powers form with the EKF; hence, all later work by him and his colleagues focused on the sum-of-products form.

The next step performed in this study was to apply particle filters to the SME approach. The particle filter was used to evaluate the effectiveness of using approximations to additive Gaussian noise with the SME formulation. The results from studying the EKF, UKF, and particle filter suggest the performance of the SME approach is extremely dependent on the pairing of an SME implementation and nonlinear filter, rather than dependent on either individually.

II. THE UNSCENTED KALMAN FILTER

A. *The Unscented Transform*

The unscented Kalman filter relies upon a mathematical technique referred to as the unscented transform [3]. The unscented transform captures the mean and variance of any random variable using a set of "sigma points." The sample mean and variance of these sigma points match the statistics of the original random variable.

Julier and Uhlmann show in [4] that finding these sigma points is a straightforward process based upon the square root of the random variable's covariance matrix. While any square root is acceptable, the Cholesky decomposition is generally chosen for its numerical stability.

For example, let X be an n-dimensional random variable with mean μ and covariance matrix P_{xx}. Let $\mathcal{X}$ be the set of sigma points for the random variable X. Julier and Uhlmann show in [3] that the $2n+1$ sigma points should be chosen as follows:

$$\mathcal{X}_0 = \mu, \tag{3}$$

$$\sigma = 2n \text{ columns of } \pm\sqrt{nP_{xx}}, \tag{4}$$

$$\mathcal{X}_i = \mu + \sigma_i, \tag{5}$$

where

$$nP_{xx} = \sqrt{nP_{xx}}^T \sqrt{nP_{xx}}. \tag{6}$$

Suppose we are interested in the random variable $Y = g(X)$ where $g(\cdot)$ is any function of an n-dimensional random variable. The estimated statistics of Y can be found by generating a new set of sigma points, $\mathcal{Y}$, where

$$\mathcal{Y}_i = g(\mathcal{X}_i) \qquad \text{for } i = 0,\ldots,2n. \tag{7}$$

The estimated statistics of Y are then the mean and variance of $\mathcal{Y}$. These estimates are accurate up to a third-order Taylor series expansion [5].

B. Building a Kalman Filter

Here we present a brief overview of the UKF. See [3] and [5] for a detailed derivation. Building the unscented transform into a filter requires augmenting the state variable to include noise terms. Specifically, we now define our state at time instant k as:

$$X_k^a = \begin{pmatrix} X_k \\ U_k \\ V_k \end{pmatrix}, \tag{8}$$

where X_k is the original (desired) state information, U_k, is the observation noise and V_k is the process noise. The covariance, $P_{X_k^a X_k^a}$, is:

$$P_{X_k^a X_k^a} = \begin{pmatrix} P_{XX} & P_{XU} & P_{XV} \\ P_{XU} & P_{UU} & P_{UV} \\ P_{XV} & P_{UV} & P_{VV} \end{pmatrix}. \tag{9}$$

The unscented transform is then applied to the augmented state and covariance, X_k^a and $P_{X_k^a X_k^a}$, to generate a set of sigma points, $\mathcal{X}_k$. To find the predicted state, the state transition function,

$f(\cdot)$, is applied to the sigma points $\mathcal{X}_k$, generating a new set of sigma points, $\mathcal{X}_{k+1|k}$. The predicted state, X_{k+1}^-, is the sample mean of $\mathcal{X}_{k+1|k}$ and the predicted covariance, $P_{X_{k+1}^a X_{k+1}^a}^-$, is the sample variance.

Now suppose $g(\cdot)$ is the observation function. A third set of sigma points, $\mathcal{Y}_{k+1|k}$, are found to represent the predicted observation:

$$\mathcal{Y}_{k+1|k} = g(\mathcal{X}_{k+1|k}). \tag{10}$$

The predicted observation, Y_{k+1}^-, and the predicted observation covariance, $P_{Y_{k+1}Y_{k+1}}$, are the sample statistics of $\mathcal{Y}_{k+1|k}$:

$$Y_{k+1}^- = \sum_{i=0}^{2n} \mathcal{Y}_{i,k+1|k}, \tag{11}$$

$$P_{Y_{k+1}Y_{k+1}} = \sum_{i=0}^{2n} [\mathcal{Y}_{i,k+1|k} - Y_{k+1}^-][\mathcal{Y}_{i,k+1|k} - Y_{k+1}^-]^T. \tag{12}$$

The predicted cross correlation, $P_{X_{k+1}Y_{k+1}}$, is the sample cross correlation of $\mathcal{X}_{k+1|k}$ and $\mathcal{Y}_{k+1|k}$:

$$P_{X_{k+1}Y_{k+1}} = \sum_{i=0}^{2n} [\mathcal{X}_{i,k+1|k} - X_{k+1}^-][\mathcal{Y}_{i,k+1|k} - Y_{k+1}^-]^T. \tag{13}$$

We now have the familiar setup of a Kalman filter. The Kalman gain, K, is:

$$K = P_{X_{k+1}Y_{k+1}} P_{Y_{k+1}Y_{k+1}}^{-1}. \tag{14}$$

The filter estimate of the state is then:

$$X_{k+1}^a = X_{k+1}^- + K(Y_{k+1} - Y_{k+1}^-), \tag{15}$$

$$P_{X_{k+1}^a X_{k+1}^a} = P_{X_{k+1}^a X_{k+1}^a}^- - K P_{Y_{k+1}Y_{k+1}} K^T. \tag{16}$$

C. Breaking the Unscented Kalman Filter

This section uses a simple example, the sum-of-products SME, to illustrate how a specific observation function breaks the UKF. Suppose we are tracking two targets in one-dimensional space. Define the state to be:

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{17}$$

where x_1 and x_2 are the positions of targets one and two. Note, for simplification, we have not augmented the state vector with noise parameters, as is necessary for the actual UKF. This

simplifies analysis without changing the conclusion. Completing the same analysis with the augmented state vector is straightforward.

Following the unscented transform, we need $2n + 1 = 5$ sigma points. Let these sigma points be:

$$\mathcal{X} = \begin{bmatrix} x_1 & x_1 + p_{11} & x_1 - p_{11} & x_1 + p_{12} & x_1 - p_{12} \\ x_2 & x_2 + p_{21} & x_2 - p_{21} & x_2 + p_{22} & x_2 - p_{22} \end{bmatrix}, \tag{18}$$

where

$$\sqrt{nP_{XX}} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}. \tag{19}$$

The sum-of-products observation function is:

$$Y = g(X) = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1 x_2 \end{pmatrix}, \tag{20}$$

so the set of predicted sigma points becomes

$$\mathcal{Y}^T = \begin{bmatrix} y_1 & y_2 \\ y_1 + p_{11} + p_{21} & y_2 + x_1 p_{21} + x_2 p_{11} + p_{11} p_{21} \\ y_1 - p_{11} - p_{21} & y_2 - x_1 p_{21} - x_2 p_{11} + p_{11} p_{21} \\ y_1 + p_{12} + p_{22} & y_2 + x_1 p_{22} + x_2 p_{12} + p_{12} p_{22} \\ y_1 - p_{12} - p_{22} & y_2 - x_1 p_{22} - x_2 p_{12} + p_{12} p_{22} \end{bmatrix}. \tag{21}$$

The interesting case occurs when the targets cross paths and $x_1 = x_2$. The predicted sigma points become

$$\mathcal{Y}^T = \begin{bmatrix} y_1 & y_2 \\ y_1 + p_{11} + p_{21} & y_2 + x_1(p_{21} + p_{11}) + p_{11} p_{21} \\ y_1 - p_{11} - p_{21} & y_2 - x_1(p_{21} + p_{11}) + p_{11} p_{21} \\ y_1 + p_{12} + p_{22} & y_2 + x_1(p_{22} + p_{12}) + p_{12} p_{22} \\ y_1 - p_{12} - p_{22} & y_2 - x_1(p_{22} + p_{12}) + p_{12} p_{22} \end{bmatrix}. \tag{22}$$

One final observation is required to understand when the UKF breaks down. When the targets behave independently, the cross-correlation terms of nP_{XX} head towards zero as the filter locks on to the targets' behavior. In equation form:

$$nP_{XX} \rightarrow \begin{pmatrix} p_{11}^2 & 0 \\ 0 & p_{22}^2 \end{pmatrix}. \tag{23}$$

Recall (19), where we defined $\sqrt{nP_{XX}}$ such that

$$nP_{XX} = \begin{pmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \tag{24}$$

$$= \begin{pmatrix} p_{11}^2 + p_{21}^2 & p_{11} p_{12} + p_{21} p_{22} \\ p_{11} p_{12} + p_{21} p_{22} & p_{22}^2 + p_{12}^2 \end{pmatrix}. \tag{25}$$

Comparing (23) and (25), p_{12} and p_{21} head towards zero when our assumption that the targets move independently is met. This leaves our predicted sigma points in the following form:

$$\mathcal{Y}^T = \begin{bmatrix} y_1 & y_2 \\ y_1+p_{11} & y_2+x_1 p_{11} \\ y_1-p_{11} & y_2-x_1 p_{11} \\ y_1+p_{22} & y_2+x_1 p_{22} \\ y_1-p_{22} & y_2-x_1 p_{22} \end{bmatrix}. \tag{26}$$

The covariance matrix is calculated according to the following:

$$P_{YY} = \sum_{i=0}^{2n} [\mathcal{Y}_i - Y][\mathcal{Y}_i - Y]^T, \tag{27}$$

where the $\mathcal{Y}_i$ are the individual sigma points of $\mathcal{Y}$. Applying equation (27) to the set of sigma points (26) leads to:

$$P_{YY} = \begin{bmatrix} 2(p_{11}^2 + p_{22}^2) & 2x_1(p_{11}^2 + p_{22}^2) \\ 2x_1(p_{11}^2 + p_{22}^2) & 2x_1^2(p_{11}^2 + p_{22}^2) \end{bmatrix}. \tag{28}$$

P_{YY} is a singular matrix with the second column being the first scaled by x_1. This means that the previously two-dimensional space has degenerated into a one-dimensional space – all the sigma points lie along a single line as seen in the top, right portion of Fig. 1. The reduction in dimensionality means the covariance matrix P_{YY} is ill-conditioned and the inverse, if it exists at all, will have elements with extremely large values. The Kalman gain, $K = P_{XY} P_{YY}^{-1}$, will then multiply the observed data by enormous values and produce inaccurate target tracks.

The need to artificially increase the covariance for numerical stability has been addressed in [6], [7]. A simple solution would be to hard limit the minimum value of the covariance matrix. Experimental results show that this approach reduces, but does not remove, the degeneration of P_{yy} and may lead to problems maintaining its positive-definiteness. This is, however, simply applying hacks to the UKF, which defeats the purpose of using it in the first place.

D. When the Unscented Kalman Filter Works

In Section II-C, we showed how the UKF breaks down when paired with the sum-of-products form of SME. This section follows a similar derivation for the sum-of-powers SME implementation and illustrates how the problems of the previous section are avoided by the sum of powers.

The setup is the same as in Section II-C with

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{29}$$

and

$$\mathcal{X} = \begin{bmatrix} x_1 & x_1 + p_{11} & x_1 - p_{11} & x_1 + p_{12} & x_1 - p_{12} \\ x_2 & x_2 + p_{21} & x_2 - p_{21} & x_2 + p_{22} & x_2 - p_{22} \end{bmatrix}. \tag{30}$$

Instead of (20), the observation function is:

$$g(X) = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1^2 + x_2^2 \end{pmatrix}, \tag{31}$$

which leads to the following set of predicted sigma points:

$$\mathcal{Y}^T = \begin{bmatrix} y_1 & y_2 \\ y_1 + p_{11} + p_{21} & y_2 + 2x_1 p_{11} + 2x_2 p_{21} + p_{11}^2 + p_{21}^2 \\ y_1 - p_{11} - p_{21} & y_2 - 2x_1 p_{11} - 2x_2 p_{21} + p_{11}^2 + p_{21}^2 \\ y_1 + p_{12} + p_{22} & y_2 + 2x_1 p_{12} + 2x_2 p_{22} + p_{12}^2 + p_{22}^2 \\ y_1 - p_{12} - p_{22} & y_2 - 2x_1 p_{12} - 2x_2 p_{22} + p_{12}^2 + p_{22}^2 \end{bmatrix}. \tag{32}$$

Again, examining the case of targets crossing and assuming that p_{12} and p_{21} are nearly zero, the predicted sigma points become:

$$\mathcal{Y}^T = \begin{bmatrix} y_1 & y_2 \\ y_1 + p_{11} & y_2 + 2x_1 p_{11} + p_{11}^2 \\ y_1 - p_{11} & y_2 - 2x_1 p_{11} + p_{11}^2 \\ y_1 + p_{22} & y_2 + 2x_1 p_{22} + p_{22}^2 \\ y_1 - p_{22} & y_2 - 2x_1 p_{22} + p_{22}^2 \end{bmatrix}. \tag{33}$$

Applying equation (27) to the set of sigma points (33) leads to:

$$P_{YY} = \begin{bmatrix} 2(p_{11}^2 + p_{22}^2) & 4x_1(p_{11}^2 + p_{22}^2) \\ 4x_1(p_{11}^2 + p_{22}^2) & 8x_1^2(p_{11}^2 + p_{22}^2) + 2(p_{11}^4 + p_{22}^4) \end{bmatrix}. \tag{34}$$

In contrast to Section II-C, the columns of P_{YY} are clearly not linearly dependent. Consequently, P_{YY} remains full rank and its inverse is well-defined. The lower right portion of Fig. 1 shows the graphical perspective. Again, all the sigma points fall onto the same curve, but this time the curve is quadratic, maintaining two dimensions rather than collapsing to one, as with the sum of products.
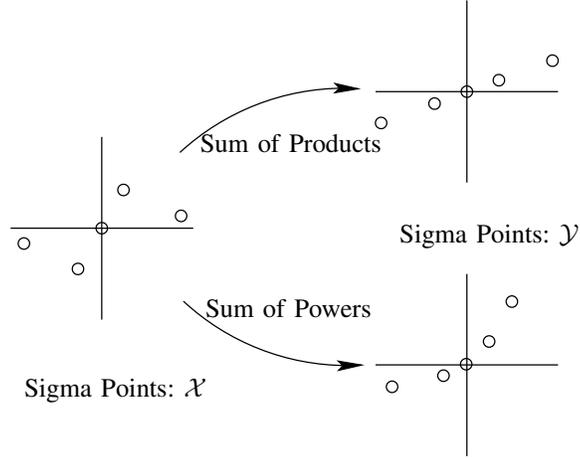
Fig. 1.   Sigma Point Transformation

III. THE PARTICLE FILTER

The particle filter implementation was chosen to provide a bound on the performance of the EKF and UKF. Since the particle filter allows the posterior to be non-Gaussian, we expected it to have the best performance [8]. In practice, however, finding the exact likelihood function for the particle filter is difficult because the SME transformations produce highly nonlinear and difficult to calculate measurement noise densities. Rather than calculate the noise density exactly, the particle filter was implemented with two different approximations. The first is an analytically derived additive Gaussian noise approximation from the EKF [9], and the second is also an additive Gaussian approximation, but based upon the unscented transform discussed in Section II-A. Thus, the particle filters may still provide an upper limit on performance, and also allow analysis of the Gaussian noise approximations.

A. Noise in the SME Formulation

Suppose our system is tracking N targets, and let the actual received measurement, y_i, be modelled as the truth, x_i, plus an additive Gaussian noise, u_i, so that

$$y_i = x_i + u_i \qquad \text{for } i = 1, \ldots, N. \tag{35}$$

Applying an SME observation function $g(\cdot)$ to the actual measurements yields:

$$\tilde{y}_i = g(y_1, y_2, \cdots, y_N) \quad \text{for } i = 1, \ldots, N. \tag{36}$$

The restrictions on the choice of SME functions given in [1] require that equation (36) can be written as:

$$\tilde{y}_i = g(x_1, x_2, \cdots, x_N) + v_i \quad \text{for } i = 1, \ldots, N, \tag{37}$$

where v_i is a zero mean white noise term. Note that v_i is not Gaussian, because the SME functions are nonlinear.

B. Analytical Gaussian Approximation

Despite the statement in Section III-A above that the noise terms v_i are not Gaussian, they can be approximated as such. Kamen presents a detailed derivation of an analytical Gaussian approximation in [9] for the sum of products. Unfortunately, the derivation for the sum-of-powers case is not nearly as neat, but the result has been included in the Appendix. Furthermore, deriving the approximations becomes increasingly tedious as the number of targets increases. It is this approximation, however, that Kamen *et al.* used with the SME-EKF pairing. The same approximation can be used with a particle filter. The results using this approximation with the particle filter are given in the *Particle Filter - Analytical* section of Tables I, II, and III.

C. Unscented Transform based Gaussian Approximation

Another approach to generating a Gaussian approximation is to use the unscented transform discussed in II-A. The advantage of using the unscented transform is that it tries to approximate the non-Gaussian white noise. Consequently, it should produce a better-performing particle filter.

To implement this approach, the unscented transform is applied to each particle of the filter at each time step. This is expensive computationally; however, for our analytical purposes, we do not need the particle filter to run in real-time. Unfortunately, the unscented transform does not work for the sum-of-products formulation for the same reasons that it broke down in the UKF, as discussed in Section II-C. The results in Section IV, however, show improved performance in the sum-of-powers case over the analytical Gaussian approximation. Note that this is not the "unscented particle filter" as presented in [5]. The unscented particle filter in [5] uses the UKF to estimate the transition prior density; whereas, we are investigating using the unscented transform to provide a better estimate of the likelihood density.

IV. RESULTS

A. Simulation Setup

The simulation consisted of three targets moving in one dimension. Each target moved independently with nearly constant velocity and a known process noise variance. Initial positions and velocities were chosen so that the targets were likely to cross paths. The observation data was generated by adding Gaussian noise with zero mean and a known variance. Software developed by Wan and van der Merwe provided the basis for these simulations [10].

To ease discussion, the filters and SME pairing will be referred to by a pair of abbreviations, e.g. UKF-Products for the unscented Kalman filter paired with the sum of products.

B. Data

TABLE I

PERCENTAGE OF TIME FILTERS MAINTAINED CORRECT TARGET ASSOCIATIONS

	Sum of Products	Sum of Powers
Extended Kalman	73.75	37.25
Unscented Kalman	40.50	82.25
PF - Analytical	86.00	82.75
PF - Unscented	13.50	87.50
Global NN	41.00	
Associated KF	98.00	

Table I contains the percentage of time that each filter maintained the correct target associations for the three targets moving in one-dimensional space. Table II contains an associated mean squared error (MSE) measurement. The term "associated" indicates that the error is only included in the average when the filter maintains the correct target association. Table III contains the "set estimation" MSE. The set estimation MSE is calculated by finding the track-estimate/track-truth association with the smallest MSE at each time step and then averaging over all time steps. Consequently, target associations are not considered.

Several relationships evident in the tables are worthy of mention. First, the EKF clearly performs better with the sum-of-products implementation (73.75%) than with the sum of powers

TABLE II

MEAN SQUARED ASSOCIATED ERROR

	Sum of Products	Sum of Powers
Extended Kalman	5.52	5.09
Unscented Kalman	12.86	5.06
PF - Analytical	5.75	4.57
PF - Unscented	5.90	4.56
Global NN	3.08	
Associated KF	3.36	

TABLE III

SET ESTIMATION MEAN SQUARED ERROR

	Sum of Products	Sum of Powers
Extended Kalman	3.26	3.21
Unscented Kalman	7.75	2.91
PF - Analytical	3.33	2.84
PF - Unscented	$- - --$	2.73
Global NN	2.01	
Associated KF	2.51	

(37.25%). This result was expected since the original investigators, using only the EKF, quickly abandoned the sum of powers.

Second, we see the effect of the UKF breaking down under the sum of products (40.50%) as discussed in Section II-C. An example of the UKF breaking down is given in Fig. 2. Notice how the estimated target tracks may jump when two estimates cross.

The most exciting result is that UKF-Powers outperforms (82.25%) both EKF implementations and approaches the performance of the particle filters. Not only is this result new and unexpected, but the UKF is significantly easier to implement and runs with similar computational complexity to the EKF. Also, by both MSE metrics, the UKF-Powers has the best performance.

Judging by the percentage of correctly maintained tracks, the PFUT-Powers showed the

PSfrag replacements

Sigma Points: $\mathcal{X}$

Sigma Points: $\mathcal{Y}$

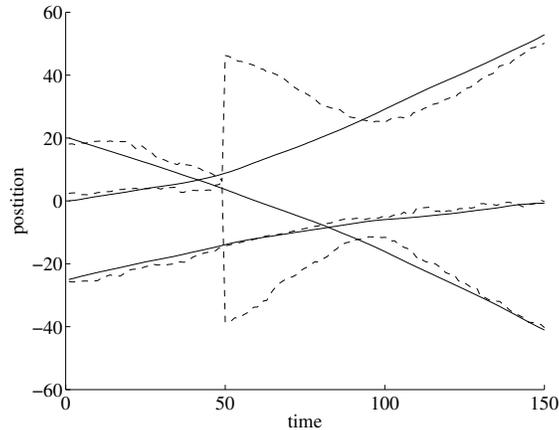Sum of Products

Sum of Powers

Fig. 2.   UKF breaking under the sum of products.

best performance (87.50%), followed by the PFAN-Products (86.00%) and the PFAN-Powers (82.75%). However, by both MSE metrics, the PFAN-Products suffers noticeably worse error performance.

Finally, a note on the global nearest neighbor (GNN) association and associated KF is necessary. GNN represents a simple solution to the multiple target tracking problem. Since the simulation scenario does not include false alarms and missed detections, each observation is used to update exactly one track. GNN associates the observations with tracks such that the sum of the distances from the observations to the predicted positions is minimized [11]. This approach may also be thought of as a 2-D assignment algorithm. Separate standard Kalman filters are run for each track. The associated KF, on the other hand, has knowledge of the correct target associations and runs a standard Kalman filter for each track. In Table I it may appear incorrect that the associated KF maintains the correct track 98% of the time instead of 100%, since it knows the correct associations. The discrepancy appears when the tracks for two targets are very close and run nearly parallel. The associated KF track for target 1 is actually closer to the truth for target 2 and vice versa. This scenario causes the track-switching algorithm to detect a loss of track.

V. ANALYSIS WITH TAYLOR SERIES EXPANSIONS

Several authors have presented Taylor series expansions about a random variable's mean as a method for comparing the accuracies of the EKF and UKF [3], [4], [5]. This section gives a

brief introduction to this analysis tool and uses it to examine the filters' performances with the SME nonlinearities.

In this paper, the system has linear dynamics, but the SME formulation generates a nonlinear measurement equation. Under these conditions, the EKF uses the measurement covariance update

$$P_{YY}(k+1|k) = \mathcal{J}_g P_{XX}(k+1|k)\mathcal{J}_g^T + R(k+1), \tag{38}$$

and the UKF uses

$$\mathcal{Y}_{k+1|k} = g(\mathcal{X}_{k+1|k}), \tag{39}$$

$$Y_{k+1}^- = \sum_{i=0}^{2n} \mathcal{Y}_{i,k+1|k}, \tag{40}$$

$$P_{Y_{k+1}Y_{k+1}} = \sum_{i=0}^{2n} [\mathcal{Y}_{i,k+1|k} - Y_{k+1}^-][\mathcal{Y}_{i,k+1|k} - Y_{k+1}^-]^T, \tag{41}$$

where $g(\cdot)$ is the nonlinear SME function, $\mathcal{J}_g$ is the Jacobian matrix of $g(\cdot)$, and R is the measurement noise covariance matrix. The $\mathcal{J}_g P_{XX}(k+1|k)\mathcal{J}_g^T$ term in (38) is a direct linearization of the covariance from the system space to the measurement space. Although the mechanics of the UKF are not as clear, the Taylor series expansion will provide some insight.

Following the derivation in [5], we represent the prior variable, $\mathbf{x}$, as a zero-mean disturbance $\delta\mathbf{x}$ about the mean $\bar{\mathbf{x}}$. Then the Taylor series expansion of $g(\mathbf{x})$ becomes

$$g(\mathbf{x}) = g(\bar{\mathbf{x}} + \delta\mathbf{x}) = \sum_{n=0}^{\infty} \left[\frac{(\delta\mathbf{x} \cdot \nabla_{\mathbf{x}})^n g(\mathbf{x})}{n!} \right]_{\mathbf{x}=\bar{\mathbf{x}}}. \tag{42}$$

To simplify notation, we define

$$\mathbf{D}_{\delta\mathbf{x}}^n g \triangleq [(\delta\mathbf{x} \cdot \nabla_{\mathbf{x}})^n g(\mathbf{x})]_{\mathbf{x}=\bar{\mathbf{x}}} \tag{43}$$

so that (42) can be written

$$\mathbf{y} = g(\mathbf{x}) = g(\bar{\mathbf{x}}) + \mathbf{D}_{\delta\mathbf{x}}g + \frac{1}{2!}\mathbf{D}_{\delta\mathbf{x}}^2 g + \frac{1}{3!}\mathbf{D}_{\delta\mathbf{x}}^3 g + \dots \tag{44}$$

A. Analyzing the Mean

From [5], the mean after the nonlinear transformation is:

$$\bar{\mathbf{y}}_{\mathbf{T}} = \mathbb{E}[\mathbf{y}] = \mathbb{E}[g(\mathbf{x})], \tag{45}$$

$$= \mathbb{E}\left[g(\bar{\mathbf{x}}) + \mathbf{D}_{\delta\mathbf{x}}g + \frac{1}{2!}\mathbf{D}_{\delta\mathbf{x}}^2 g + \frac{1}{3!}\mathbf{D}_{\delta\mathbf{x}}^3 g + \dots \right], \tag{46}$$

where the subscript T indicates that the expression represents the true mean. Equation (46) can be simplified by assuming that $\mathbf{x}$ is a symmetrically distributed random variable, which causes all odd-moments to equal zero. Two identities will also help simplify (46). First, $\mathbb{E}[\delta\mathbf{x}\delta\mathbf{x}^T] = \mathbf{P_x}$. Second, $\mathbb{E}[\mathbf{D}_{\delta\mathbf{x}}g] = \frac{1}{2}[(\nabla^T\mathbf{P_x}\nabla)g(\mathbf{x})]_{\mathbf{x}=\bar{\mathbf{x}}}$. This leads to

$$\bar{\mathbf{y}}_{\mathbf{T}} = g(\bar{\mathbf{x}}) + \frac{1}{2}[(\nabla^T\mathbf{P_x}\nabla)g(\mathbf{x})]_{\mathbf{x}=\bar{\mathbf{x}}} + \mathbb{E}\left[\frac{1}{4!}\mathbf{D}_{\delta\mathbf{x}}^4 g + \frac{1}{6!}\mathbf{D}_{\delta\mathbf{x}}^6 g + \ldots\right]. \tag{47}$$

In [5], the authors show that the unscented transform calculates the mean as:

$$\bar{\mathbf{y}}_{\mathbf{UT}} = g(\bar{\mathbf{x}}) + \frac{1}{2}[(\nabla^T\mathbf{P_x}\nabla)g(\mathbf{x})]_{\mathbf{x}=\bar{\mathbf{x}}} + \frac{1}{2(L+\lambda)}\sum_{i=1}^{2L}\left(\frac{1}{4!}\mathbf{D}_{\boldsymbol{\sigma}_i}^4 g + \frac{1}{6!}\mathbf{D}_{\boldsymbol{\sigma}_i}^6 g + \ldots\right), \tag{48}$$

where $\boldsymbol{\sigma_i}$ denotes the ith column of the matrix square root of $\sqrt{(L+\lambda)}\mathbf{P_x}$. The linear approximation used in the EKF calculates the mean as

$$\bar{\mathbf{y}}_{\mathbf{LIN}} = g(\bar{\mathbf{x}}). \tag{49}$$

Comparing equations (47), (48), and (49), the unscented transform matches the true mean in the first and second order terms, whereas the linearization only matches the true mean for the first term. Julier and Uhlman show in [3] that the errors in the higher order terms of $\bar{\mathbf{y}}_{\mathbf{UT}}$ are smaller than the error in assuming these terms are all zero as the linear approximation does.

B. Mean Approximation for the Sum-of-Products SME

For the three-target sum-of-products case,

$$\nabla^T = \left[\frac{\partial}{\partial x_1} \; \frac{\partial}{\partial x_2} \; \frac{\partial}{\partial x_3}\right], \tag{50}$$

and describing $\mathbf{P_x}$ as

$$\mathbf{P_x} = \begin{bmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ \rho_{21} & \rho_{22} & \rho_{23} \\ \rho_{31} & \rho_{32} & \rho_{33} \end{bmatrix}, \tag{51}$$

leads to

$$[(\nabla^T\mathbf{P_x}\nabla)g_{prod}(\mathbf{x})]_{\mathbf{x}=\bar{\mathbf{x}}} = \begin{bmatrix} 0 \\ 2(\rho_{12} + \rho_{23} + \rho_{13}) \\ 2(\rho_{12}x_3 + \rho_{23}x_1 + \rho_{13}x_2) \end{bmatrix}. \tag{52}$$

Throughout this paper, targets have been assumed to move independently, which makes $\mathbf{P_x}$ a diagonal matrix. Since all the elements of the second order term (52) depend only on the cross-covariances, the entire second order term is zero. Furthermore, the derivatives in the fourth and

higher order terms of (48) reduce to zero for the three target case. This leaves $\bar{\mathbf{y}}_{\mathbf{UT}} = g(\bar{\mathbf{x}})$, which is exactly the same as $\bar{\mathbf{y}}_{\mathbf{LIN}}$. With the three-target sum of products SME and a diagonal $\mathbf{P_x}$, the UKF and EKF employ the same approximation for the measurement noise mean after the SME transformation.

C. Mean Approximation for the Sum of Powers SME

Following the derivation from the previous section, the second order term in the Taylor expansion for the sum-of-powers SME is

$$\left[(\nabla^T \mathbf{P_x} \nabla) g_{pow}(\mathbf{x})\right]_{\mathbf{x}=\bar{\mathbf{x}}} = \begin{bmatrix} 0 \\ 2(\rho_{11} + \rho_{22} + \rho_{33}) \\ 6(\rho_{11}x_1 + \rho_{22}x_2 + \rho_{33}x_3) \end{bmatrix}. \tag{53}$$

The second order term (53) depends only on the variance terms of $\mathbf{P_x}$. Consequently, the second order term is not zero, although the derivatives in the fourth and higher order terms reduce to zero for a three target case. Thus, with one more non-zero term than the EKF, the UKF will provide a better estimate of the mean than the linearization used by the EKF.

D. Analyzing the Covariance

Now using the Taylor series approximation to find an expression for the true covariance [5],

$$\mathbf{P_y} = \mathcal{J}_{\mathbf{x}} \mathbf{P_x} \mathcal{J}_{\mathbf{x}}^T \tag{54}$$

$$-\frac{1}{4}\left\{\left[(\nabla^T \mathbf{P_x} \nabla) g(\mathbf{x})\right]\left[(\nabla^T \mathbf{P_x} \nabla) g(\mathbf{x})\right]^T\right\}_{\mathbf{x}=\bar{\mathbf{x}}}$$

$$+\mathbb{E}\left[\underbrace{\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\frac{1}{i!j!}\mathbf{D}_{\delta\mathbf{x}}^i g \ (\mathbf{D}_{\delta\mathbf{x}}^j g)^T}_{i \neq j = 1}\right]$$

$$-\left[\underbrace{\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\frac{1}{(2i)!(2j)!}\mathbb{E}\left[\mathbf{D}_{\delta\mathbf{x}}^{2i} g\right]\mathbb{E}\left[\mathbf{D}_{\delta\mathbf{x}}^{2j} g\right]^T}_{i \neq j = 1}\right],$$

where $\mathcal{J}_{\mathbf{x}}$ is the Jacobian matrix of $g(\mathbf{x})$ evaluated at $\mathbf{x} = \bar{\mathbf{x}}$.

Using a similar approach as with the mean [3], it can be shown that the unscented transform calculates the posterior covariance as

$$
\begin{aligned}
(\mathbf{P_y})_{UT} = \;& \mathcal{J}_\mathbf{x} \mathbf{P_x} \mathcal{J}_\mathbf{x}^T \\
& - \frac{1}{4} \left\{ \left[(\nabla^T \mathbf{P_x} \nabla) g(\mathbf{x}) \right] \left[(\nabla^T \mathbf{P_x} \nabla) g(\mathbf{x}) \right]^T \right\}_{\mathbf{x}=\bar{\mathbf{x}}} \\
& + \frac{1}{2(L+\lambda)} \sum_{k=1}^{2L} \left[\underbrace{ \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{1}{i!j!} \mathbf{D}_{\sigma_\mathbf{k}}^i g \, (\mathbf{D}_{\sigma_\mathbf{k}}^j g)^T }_{i \neq j = 1} \right] \\
& - \left[\underbrace{ \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{1}{(2i)!(2j)!4(L+\lambda)^2} }_{i \neq j = 1} \sum_{k=1}^{2L} \sum_{m=1}^{2L} \left[\mathbf{D}_{\sigma_\mathbf{k}}^{2i} g \right] \left(\mathbf{D}_{\sigma_\mathbf{m}}^{2j} g \right)^T \right].
\end{aligned}
\tag{55}
$$

For the covariance, the unscented transform correctly calculates the first two terms. The linearization, on the other hand, calculates the posterior covariance as

$$
(\mathbf{P_y})_{LIN} = \mathcal{J}_\mathbf{x} \mathbf{P_x} \mathcal{J}_\mathbf{x}^T,
\tag{56}
$$

which is only accurate for the first term. Again, the authors of [3] show that the error for unscented transform's approximation is less that the assumption that all higher order terms are zero.

E. Covariance Approximation for the Sum of Products

Section V-B showed that the second term of (55) is zero for the sum of products when the prior covariance matrix is diagonal. The third and fourth terms are also zero for the sum of products when the prior covariance matrix is diagonal. Any element of these terms that is non-zero after taking all the appropriate derivatives is multiplied by a cross-covariance that is zero. The unscented transform's mean and covariance produce an approximation similar to that of the linearization used in the EKF for the sum-of-products SME. Unfortunately, a direct comparison of the EKF and UKF is not practical because the UKF is incompatible with the sum of products.

F. Covariance Approximation for the Sum of Powers

Section V-C it was shown that the second term of (55) depends on the variance terms of the prior covariance matrix and will be nonzero. Thus, the unscented transform will generate a more

accurate approximation of the measurement posterior than the linearization used in the EKF. The effect of this improved approximation can be seen in the results. The UKF has significantly better performance than the EKF when paired with the sum-of-powers SME.

G. Other Observations from Taylor Series Analysis

1) Data Dependence: Equations 52 and 53 reveal one drawback to the SME method. Since the data appears in the third row of the expression, the covariance estimate is a function of the data. This is equivalent to saying that the covariance depends upon the location of the mean. While this problem can be alleviated by using a centered coordinate system, ideally the covariance estimate would be independent of the raw data's choice of Cartesian origin.

2) Increasing the Number of Targets: The multiple target tracking problem increases in complexity as the number of targets increases. In the SME approach, this is manifested by an increase in the number of equations needed, the magnitude of the numbers used in these equations, and the difficulty posed by increasing the number of variables maintained by the filter. From the Taylor series expansion analysis presented above, it becomes evident that increasing the number of targets reduces the accuracy of the EKF and UKF's approximations. In terms of the Taylor series, both can be thought of as truncating the Taylor series after some number of terms. The derivatives in the higher order terms generally cause those terms to be zero for the SME case. Hence, the error can be thought of as all the nonzero terms not included by the EKF or UKF's approximations. As the number of targets increases, the number of nonzero terms increases, creating a corresponding increase in the error.

VI. CONCLUSIONS

Several interesting phenomena were observed in this study. In Kamen's early studies, he found the sum-of-products form of the SME to work better than the sum of powers with the EKF; hence, all later work by him and his colleagues focused on the sum-of-products form. We have found that an UKF implementation of the sum of powers – the form originally abandoned by Kamen - actually performs better than the EKF implementation of either form. Also, we have discovered that the sum-of-products nonlinearity is inherently incompatible with the UKF, inadvertently uncovering an aspect of UKFs that does not seem to have been previously discussed in the literature. These results suggest the performance of the SME approach is dependent on the

pairing of a specific SME implementation and nonlinear filter rather than dependent on either individually.

VII. FUTURE WORK

This paper focused on exploring and comparing different implementations of the SME. A clear avenue for future work would be to compare our various SME implementations with other association algorithms [11] such as JPDA (Joint Probabilistic Data Association), N-D multiassignment for $N > 2$ (i.e., going several scans into the past) and Kastella's EAMLE (Event-Averaged Maximum-Likelihood Estimation).

Perhaps the most exciting aspect of this work is the potential for using the SME as a framework for deriving multitarget Cramér-Rao bounds, as suggested by Fred Daum [12]. While developing an EKF-based SME filter for a large number of targets is a challenging task, the same system is easily implemented using an UKF. Thus the UKF/SME pair may allow easy implementation and efficient computation of a CR bound for multiple target tracking systems. Furthermore, the SME framework allows for straightforward analysis of the correlation between errors in tracking separate targets, particularly when the targets cross paths.

APPENDIX

A. Sum of Products

Recall that for the sum of products, the measurements are found by taking the received data, y, and calculating a pseudomeasurement, $\tilde{y}_{prod}$, as follows:

$$\tilde{y}_{prod} = \begin{pmatrix} y_1 + y_2 + y_3 \\ y_1 y_2 + y_1 y_3 + y_2 y_3 \\ y_1 y_2 y_3 \end{pmatrix}. \tag{57}$$

Also, recall that y_i can be written as $x_i + u_i$ where x_i is the truth and u_i is an additive Gaussian noise term with variance σ_{u_i}. Restrictions on the SME approach require that y_i can be written as

$$y_i = g(x_1, x_2, \cdots, x_N) + v_i \quad \text{for } i = 1 \cdots N. \tag{58}$$

Following the derivation in [9], let R be the covariance matrix of the new measurement noise, v. For the three-target case, R can be approximated as

$$R \approx \sigma_1^2 V_1 V_1^T + \sigma_2^2 V_2 V_2^T + \sigma_3^2 V_3 V_3^T, \tag{59}$$

where

$$\sigma_1^2 = \sigma_{u_1}^2 + \sigma_{u_2}^2 + \sigma_{u_3}^2, \tag{60}$$

$$\sigma_2^2 = \sigma_{u_1}^2 \sigma_{u_2}^2 + \sigma_{u_1}^2 \sigma_{u_3}^2 + \sigma_{u_2}^2 \sigma_{u_3}^2, \tag{61}$$

$$\sigma_3^2 = \sigma_{u_1}^2 \sigma_{u_2}^2 \sigma_{u_3}^2, \tag{62}$$

and

$$V_1 = \begin{pmatrix} 1 & 1 & 1 \\ x_2 + x_3 & x_1 + x_3 & x_1 + x_2 \\ x_2 x_3 & x_1 x_3 & x_1 x_2 \end{pmatrix}, \tag{63}$$

$$V_2 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ x_3 & x_2 & x_1 \end{pmatrix}, \tag{64}$$

$$V_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{65}$$

B. Sum of Powers

To further simplify calculations in the sum-of-powers case, each measurement was assumed to have the same noise variance such that $\sigma_{u_1}^2 = \sigma_{u_2}^2 = \sigma_{u_3}^2 = \sigma_u^2$. For the sum of powers,

$$y_{pow} = \begin{pmatrix} y_1 + y_2 + y_3 \\ y_1^2 + y_2^2 + y_3^2 - 3\sigma_u^2 \\ y_1^3 + y_2^3 + y_3^3 - 3\sigma_u^2(y_1 + y_2 + y_3) \end{pmatrix}, \tag{66}$$

where the σ_u^2 terms are included to ensure that the measurement noise after the sum-of-powers transformation is zero mean. The calculation required to approximate the new measurement noise covariance matrix, R, for the sum of powers is long and tedious, so the details have been omitted here. However, the result is the following:

$$R \approx \begin{pmatrix} 3\sigma_u^2 & 2\sigma_u^2 z_1 & 3\sigma_u^2 z_2 + 9\sigma_u^4 \\ 2\sigma_u^2 z_1 & 4\sigma_u^2 z_2 + 6\sigma_u^4 & 6\sigma_u^2 z_3 + 12\sigma_u^4 z_1 \\ 3\sigma_u^2 z_2 + 9\sigma_u^4 & 6\sigma_u^2 z_3 + 12\sigma_u^4 z_1 & 9\sigma_u^2 z_4 + 36\sigma_u^4 z_2 + 45\sigma_u^6 \end{pmatrix}, \tag{67}$$

where

$$z_1 = x_1 + x_2 + x_3, \tag{68}$$

$$z_2 = x_1^2 + x_2^2 + x_3^2, \tag{69}$$

$$z_3 = x_1^3 + x_2^3 + x_3^3, \tag{70}$$

$$z_4 = x_1^4 + x_2^4 + x_3^4. \tag{71}$$

REFERENCES

[1] E. Kamen, "Multiple target tracking based on symmetric measurement equations," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 371–374, April 1992.

[2] E. Kamen and C. Sastry, "Multiple target tracking using products of position measurements," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 2, pp. 476–493, 1993.

[3] S. J. Julier and J. K. Uhlmann, "New extension of the Kalman filter to nonlinear systems," in *Proceedings of the SPIE*. Orlando, FL: SPIE, Apr. 1997, pp. 182–193.

[4] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Trans. Automat. Contr.*, vol. 45, no. 3, pp. 477–482, Mar. 2000.

[5] E. Wan and R. van der Merwe, *Kalman Filtering and Neural Networks*. Wiley, 2001, ch. 7.

[6] S. J. Julier, "The scaled unscented transformation," in *Proceedings of the 2002 American Control Conference*, vol. 6, Anchorage, AK, May 2002, pp. 4555 – 4559.

[7] T. Lefebvre, H. Bruyninckx, and J. D. Schuller, "Comment on 'A new method for the nonlinear transformation of means and covariances in filters and estimators' [and authors' reply]," *IEEE Trans. Automat. Contr.*, vol. 47, pp. 1406– 1409, Aug. 2002.

[8] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, 2001.

[9] E. Kamen and J. Su, *Introduction to Optimal Estimation*. Springer-Verlag, 1999.

[10] E. Wan and R. van der Merwe, *Recursive Bayesian Estimation Library*. Beaverton, OR: OGI School of Science and Engineering, 2003.

[11] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Artech House, 1999.

[12] F. Daum, "A Cramér-Rao bound for multiple target tracking," in *Signal and Data Processing of Small Targets*, O. Drummond, Ed., vol. SPIE Proc. 1481, Orlando, FL, Apr. 1991, pp. 341–344.