Technical Notes and Correspondence_____

A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators

Simon Julier, Jeffrey Uhlmann, and Hugh F. Durrant-Whyte

Abstract—This paper describes a new approach for generalizing the Kalman filter to nonlinear systems. A set of samples are used to parameterize the mean and covariance of a (not necessarily Gaussian) probability distribution. The method yields a filter that is more accurate than an extended Kalman filter (EKF) and easier to implement than an EKF or a Gauss second-order filter. Its effectiveness is demonstrated using an example.

Index Terms—Covariance matrices, estimation, filtering, missile detection and tracking, mobile robots, nonlinear filters, prediction methods.

I. INTRODUCTION

The problem of generalizing the Kalman filter paradigm for nonlinear applications is considered. This work is motivated by the well-known limitations of the *extended Kalman filter* (EKF), which simply linearizes all nonlinear models so that the traditional linear Kalman filter can be applied. Although the EKF (in its many forms) is a widely used filtering strategy [1], [2], over 30 years of experience with it has led to a general consensus within the tracking and control community that it is difficult to implement, difficult to tune, and only reliable for systems that are almost linear on the time scale of the update intervals.

As is well known, the optimal solution to the nonlinear filtering problem is infinite dimensional [3] and a large number of suboptimal approaches have been developed [4], [5]. These methods can be broadly classed as numerical Monte Carlo [6] methods or analytical approximations [7]–[9]. However, the application of these methods to high-dimensioned systems is rarely practical, and it is a testament to the conceptual simplicity of the EKF that it is still widely used.

In this paper, a new linear estimator [10]–[12] is developed. It yields performance equivalent to the Kalman filter for linear systems, yet generalizes elegantly to nonlinear systems without the linearization steps required by the EKF. We show analytically that the expected performance of the new approach is superior to that of the EKF and, in fact, lies between those of the modified, truncated second-order filter [13] and the Gaussian second-order filter [14]. However, the algorithm is not restricted to Gaussian distributions. We demonstrate the performance benefits in an example application, and we argue that the ease of implementation and more accurate estimation features of the new filter recommend its use over the EKF in virtually all applications.

II. BACKGROUND

We seek the minimum-mean-squared error (MMSE) estimate of the state vector of the nonlinear discrete time system

$$\mathbf{x}(k+1) = \mathbf{f}[\mathbf{x}(k), \mathbf{u}(k), \mathbf{v}(k), k]$$

$$\mathbf{z}(k) = \mathbf{h}[\mathbf{x}(k), \mathbf{u}(k), k] + \mathbf{w}(k) \tag{1}$$

where $\mathbf{x}(k)$ is the state of the system at timestep k, $\mathbf{u}(k+1)$ is the input vector, $\mathbf{v}(k)$ is the noise process caused by disturbances and modelling errors, $\mathbf{z}(k)$ is the observation vector, and $\mathbf{w}(k)$ is additive measurement noise. It is assumed that the noise vectors $\mathbf{v}(k)$ and $\mathbf{w}(k)$ are zero mean and

$$E[\mathbf{v}(i)\mathbf{v}^T(j)] = \delta_{ij}\mathbf{Q}(i)$$
$$E[\mathbf{w}(i)\mathbf{w}^T(j)] = \delta_{ij}\mathbf{R}(i), \qquad \forall i, j$$
$$E[\mathbf{v}(i)\mathbf{w}^T(j)] = \mathbf{0}.$$

The MMSE estimate of $\mathbf{x}(k)$ is the *conditional mean*. Let $\hat{\mathbf{x}}(i|j)$ be the mean of $\mathbf{x}(i)$ conditioned on all of the observations up to and including time j

$$\hat{\mathbf{x}}(i|j) = E[\mathbf{x}(i)|\mathbf{Z}^j]$$

where $\mathbf{Z}^j = [\mathbf{z}(1), \cdots, \mathbf{z}(j)]$. The covariance of this estimate is denoted $\mathbf{P}(i|j)$.

The Kalman filter propagates the first two moments of the distribution of $\mathbf{x}(k)$ recursively and has a distinctive "predictor-corrector" structure. Given an estimate $\hat{\mathbf{x}}(k|k)$, the filter first predicts what the future state of the system will be using the process model. Ideally, the predicted quantities are

$$\hat{\mathbf{x}}(k+1|k) = E[\mathbf{f}[\mathbf{x}(k), \mathbf{u}(k), \mathbf{v}(k), k]|\mathbf{Z}^k]$$
$$\mathbf{P}(k+1|k) = E[\{\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k)\}$$
$$\times \{\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k)\}^T|\mathbf{Z}^k].$$

The expectations can be calculated only if the distribution of $\mathbf{x}(k)$ conditioned on $\mathbf{Z}^k$ is known. In general, the distribution cannot be described by a finite number of parameters and most practical systems employ an approximation of some kind. It is conventionally assumed that the distribution of $\mathbf{x}(k)$ is Gaussian at any time k. Two justifications are made. First, only the mean and covariance need to be maintained. Second, given just the first two moments the Gaussian distribution is the entropy maximizing or least informative distribution [15].

The estimate at time $k+1$ is given through updating the prediction by the linear update rule

$$\hat{\mathbf{x}}(k+1|k+1) = \hat{\mathbf{x}}(k+1|k) + \mathbf{W}(k+1)\boldsymbol{\nu}(k+1)$$
$$\mathbf{P}(k+1|k+1) = \mathbf{P}(k+1|k) - \mathbf{W}(k+1)\mathbf{P}_{\nu\nu}(k+1|k)$$
$$\cdot \mathbf{W}^T(k+1)$$
$$\boldsymbol{\nu}(k+1) = \mathbf{z}(k+1) - \hat{\mathbf{z}}(k+1|k)$$
$$\mathbf{W}(k+1) = \mathbf{P}_{x\nu}(k+1|k)\mathbf{P}_{\nu\nu}^{-1}(k+1|k).$$

The EKF exploits the fact that the error in the prediction, $\tilde{\mathbf{x}}(i|j) = \mathbf{x}(i) - \hat{\mathbf{x}}(i|j)$, can be attained by expanding (1) as a Taylor Series about the estimate $\hat{\mathbf{x}}(k|k)$. Truncating this series at the first order yields the approximate linear expression for the propagation of state errors as

$$\tilde{\mathbf{x}}(k+1|k) \approx \nabla\mathbf{f}_x\tilde{\mathbf{x}}(k|k) + \nabla\mathbf{f}_v\mathbf{v}(k)$$

where $\nabla\mathbf{f}_x$ is the Jacobian of $\mathbf{f}[\cdot]$ with respect to $\mathbf{x}(k)$, and $\nabla\mathbf{f}_v$ is that with respect to $\mathbf{v}(k)$. Using this approximation, the state prediction equations are

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{f}[\hat{\mathbf{x}}(k|k), \mathbf{u}(k), \mathbf{0}]$$

$$\mathbf{P}(k+1|k) = \nabla \mathbf{f}_x \mathbf{P}(k|k) \nabla^T \mathbf{f}_x + \nabla \mathbf{f}_v \mathbf{Q}(k+1) \nabla^T \mathbf{f}_v. \quad (2)$$

This approximation is valid if the contributions of the truncated higher order terms are negligible. However, in many practical situations (such as converting from polar to Cartesian coordinate systems [17]), linearization introduces significant biases or errors. The EKF can also be prohibitively difficult to implement because deriving Jacobians can be cumbersome and time consuming (especially when the system is complicated and of high order), and they must be reevaluated at every prediction step of the filter.

This paper focuses on the central problem of predicting the mean and covariance. We present an alternative algorithm that is more accurate, convenient, and uses approximately the same number of calculations as an EKF.

III. THE NEW FILTER

We use the intuition that *it is easier to approximate a probability distribution than it is to approximate an arbitrary nonlinear function or transformation* [18]. Following this intuition, we generate a set of points whose sample mean and sample covariance are $\hat{\mathbf{x}}(k|k)$ and $\mathbf{P}(k|k)$, respectively. The nonlinear function is applied to each of these points in turn to yield a transformed sample, and the predicted mean and covariance are calculated from the transformed sample.

Although this superficially resembles a Monte Carlo method, the samples are *not* drawn at random. Rather, the samples are deterministically chosen so that they capture specific information about the distribution. In general, this intuition can be applied to capture many kinds of information about many types of distribution [19], [20]. In this paper, we consider the special case of i) capturing the mean and covariance of an ii) assumed Gaussian distribution.

The n-dimensional random variable $\mathbf{x}(k)$ with mean $\hat{\mathbf{x}}(k|k)$ and covariance $\mathbf{P}(k|k)$ is approximated by $2n+1$ weighted samples or *sigma points* selected by the algorithm

$$\begin{aligned}
\mathcal{X}_0(k|k) &= \hat{\mathbf{x}}(k|k) \\
W_0 &= \kappa/(n+\kappa) \\
\mathcal{X}_i(k|k) &= \hat{\mathbf{x}}(k|k) + \left(\sqrt{(n+\kappa)\mathbf{P}(k|k)}\right)_i \\
W_i &= 1/\{2(n+\kappa)\} \\
\mathcal{X}_{i+n}(k|k) &= \hat{\mathbf{x}}(k|k) - \left(\sqrt{(n+\kappa)\mathbf{P}(k|k)}\right)_i \\
W_{i+n} &= 1/\{2(n+\kappa)\}.
\end{aligned} \quad (3)$$

$\kappa \in \Re$, $(\sqrt{(n+\kappa)\mathbf{P}(k|k)})_i$ is the ith row or column[1] of the matrix square root of $(n+\kappa)\mathbf{P}(k|k)$, and W_i is the weight that is associated with the ith point.

Theorem 1: The set of samples chosen by (3) have the same sample mean, covariance, and all higher odd-ordered central moments as the distribution of $\mathbf{x}(k)$. The matrix square root and κ affect the fourth and higher order sample moments of the sigma points.

Proof: The matching of the mean, covariance, and all odd-ordered moments can be readily demonstrated. Because the points are symmetrically distributed and chosen with equal weights about $\overline{\mathbf{x}}$, the sample mean is obviously $\overline{\mathbf{x}}$ and all odd-ordered moments are zero. The sample covariance $\mathbf{P}$ is

$$\begin{aligned}
\mathbf{P} &= \sum_{i=0}^{2n} W_i\{\mathcal{X}_i(k|k) - \hat{\mathbf{x}}(k|k)\}\{\mathcal{X}_i(k|k) - \hat{\mathbf{x}}(k|k)\}^T \\
&= \sum_{i=1}^{n} 2W_i(n+\kappa)\left(\sqrt{\mathbf{P}(k|k)}\right)_i\left(\sqrt{\mathbf{P}(k|k)}\right)_i^T
\end{aligned}$$

[1]If the matrix square root $\mathbf{A}$ of $\mathbf{P}$ is of the form $\mathbf{P} = \mathbf{A}^T\mathbf{A}$, then the sigma points are formed from the *rows* of $\mathbf{A}$. However, for a root of the form $\mathbf{P} = \mathbf{A}\mathbf{A}^T$, the *columns* of $\mathbf{A}$ are used.

$$\begin{aligned}
&= \sum_{i=1}^{n} \left(\sqrt{\mathbf{P}(k|k)}\right)_i \left(\sqrt{\mathbf{P}(k|k)}\right)_i^T \\
&= \mathbf{P}(k|k).
\end{aligned}$$

Appendix I sketches the principle of the argument for the higher moments, and a full proof can be found in [21]. ∎

Remark 1: The above properties hold for *any* choice of the matrix square root. Efficient and stable methods, such as the Cholesky decomposition, *should* be used.

Remark 2: κ can be any number (positive or negative) providing that $(n+\kappa) \neq 0$.

Given the set of samples generated by (3), the prediction procedure is as follows.

1) Each sigma point is instantiated through the process model to yield a set of transformed samples

$$\mathcal{X}_i(k+1|k) = \mathbf{f}[\mathcal{X}_i(k|k), \mathbf{u}(k), k].$$

2) The predicted mean is computed as

$$\hat{\mathbf{x}}(k+1|k) = \sum_{i=0}^{2n} W_i \mathcal{X}_i(k+1|k). \quad (4)$$

3) The predicted covariance is computed as

$$\begin{aligned}
\mathbf{P}(k+1|k) = \sum_{i=0}^{2n} W_i &\{\mathcal{X}_i(k+1|k) - \hat{\mathbf{x}}(k+1|k)\} \\
&\times \{\mathcal{X}_i(k+1|k) - \hat{\mathbf{x}}(k+1|k)\}^T. \quad (5)
\end{aligned}$$

The mean and covariance are calculated using standard vector and matrix operations, which means that the algorithm is suitable for *any* choice of process model, and implementation is extremely convenient because it is not necessary to evaluate the Jacobians, which are needed in an EKF. The method has a further advantage: it yields more accurate predictions than those determined through linearization.

Theorem 2: The prediction algorithm introduces errors in estimating the mean and covariance at the fourth and higher orders in the Taylor series. These higher order terms are a function of κ and the matrix square root used.

Proof: Appendix II shows that evaluating the mean and covariance correctly up to the mth order requires approximating the moments of $\mathbf{x}(k)$ up to the mth order. From Theorem 1, the approximation is correct up to the third order, and so errors are introduced in the fourth and higher orders. These are a function of the matrix square root and κ. ∎

Remark 3: This is the same order of accuracy as the second-order Gaussian filter [14], but without the need to calculate Jacobians or Hessians.

κ provides an extra degree of freedom to "fine tune" the higher order moments of the approximation, and can be used to reduce the overall prediction errors. When $\mathbf{x}(k)$ is assumed Gaussian, Appendix I shows that a useful heuristic is to select $n+\kappa = 3$. If a different distribution is assumed for $\mathbf{x}(k)$, then a different choice of κ might be more appropriate.

Although κ can be positive or negative, Appendix III shows that when κ is negative it is possible to calculate a nonpositive, semidefinite $\mathbf{P}(k+1|k)$. This problem is not uncommon for methods that approximate higher order moments or probability density distributions (such as those described in [5], [4], and [8]). In this situation, it is possible to use a *modified* form of the prediction algorithm. The mean is still calculated using (4), but the "covariance" is evaluated about $\mathcal{X}_0(k+1|k)$. Appendix III shows that the modified form ensures positive semi-definiteness, and, in the limit $(n+\kappa) \to 0$, the prediction is the same as that of the modified, truncated second-order filter [4].

Remark 4: The method can be generalized to include the effects of process and observation noise by appending the noise vectors to the state vector [11].

IV. EXAMPLE APPLICATION

This section applies and compares the performance of the new filter against an EKF for a tracking problem that was presented in [14]. This example was chosen because it has significant nonlinearities in the process and observation models and has been analyzed extensively in the literature.

We wish to estimate the position $x_1(t)$, velocity $x_2(t)$, and constant ballistic coefficient $x_3(t)$ of a body as it reenters the atmosphere at a very high altitude at a very high velocity. Its motion is determined by altitude- and velocity-dependent drag terms, and it is constrained to fall vertically. The position of the body is measured at discrete points in time using a radar capable of measuring range corrupted by Gaussian measurement noise. The radar is at an altitude of H (100 000 ft), and the horizontal range between the body and the radar, M, is (100 000 ft).

The continuous time dynamics of this system are

$$\dot{x}_1(t) = -x_2(t) + w_1(t) \tag{6}$$
$$\dot{x}_2(t) = -e^{-\gamma x_1(t)} x_2(t)^2 x_3(t) + w_2(t) \tag{7}$$
$$\dot{x}_3(t) = w_3(t) \tag{8}$$

where $w_1(t)$, $w_2(t)$, and $w_3(t)$ are zero-mean, uncorrelated noises with covariances given by $\mathbf{Q}(t)$ and γ is a constant (5×10^{-5}) that relates the air density with altitude. The range at time t, $z(t)$, is

$$z(t) = \sqrt{(M^2 + [x_1(t) - H]^2)} + r(t), \tag{9}$$

where $r(t)$ is the uncorrelated observation noise with covariance $\mathbf{R}(t) = 10^4 \text{ ft}^2$. The measurements are made with a frequency of 1 Hz.

Tracking systems were implemented using the new filter and the EKF. The nonlinearities of the process model and the high velocities required the numerical integration of (6)–(8) to be carried out using extremely small time steps. In accordance with [14], a fourth-order Runge–Kutta scheme was employed with 64 steps between each observation. For the EKF, it was necessary to recalculate the Jacobian 64 times between each update. For the new filter, the trajectory of each sigma point was calculated using the small time steps, but it was only necessary to calculate the mean and covariance just before an observation was made. Because $n = 3$, κ was chosen to be zero in accordance with the heuristic $n + \kappa = 3$.

The initial true state of the system is $\mathbf{x}(0) = [3 \times 10^5 \ 2 \times 10^4 \ 10^{-3}]$, and the initial estimates and covariances of these states are

$$\hat{\mathbf{x}}(0|0) = [3 \times 10^5 \ 2 \times 10^4 \ 3 \times 10^{-5}]$$
$$\mathbf{P}(0|0) = \begin{bmatrix} 10^6 & 0 & 0 \\ 0 & 4 \times 10^6 & 0 \\ 0 & 0 & 10 \end{bmatrix}.$$

Although the initial estimates of altitude and velocity are correct, $\hat{\mathbf{x}}_3(0|0)$ is very bad. The body is assumed to be "heavy," whereas in reality it is "light." The behavior of the two filters differs if the second and higher order terms are significant. Because process noise can be used to mask linearization errors, we adopt the practice from [14] and do not introduce any process noise into the simulation—$\mathbf{Q}(k) = \mathbf{0}$ for both filters.

In Fig. 1, we show the average magnitude of the state errors committed by each filter across a Monte Carlo simulation consisting of 50
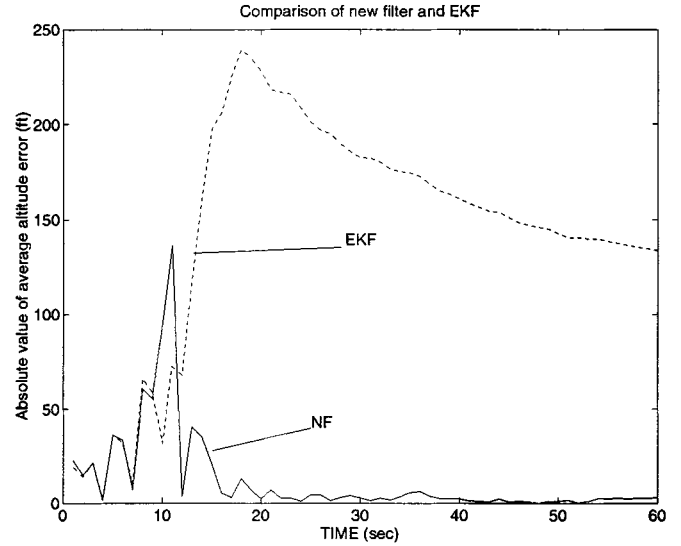

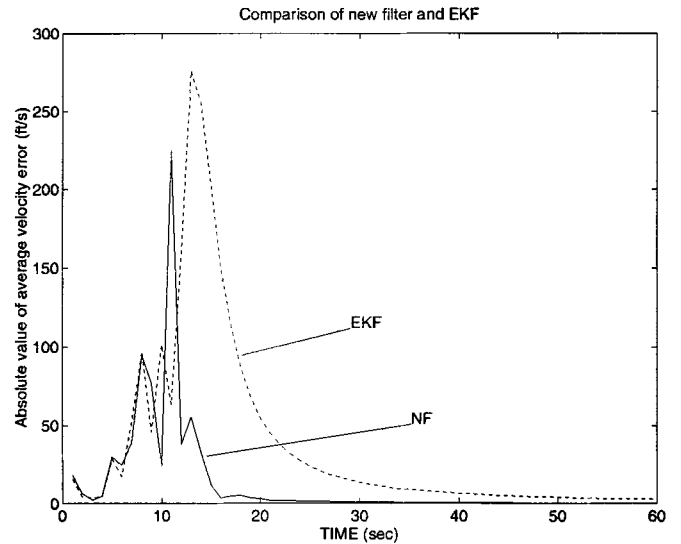Fig. 1. Absolute mean error position error.

Fig. 2. Absolute mean velocity error.

runs. At high altitude, the drag effects are minimal, and the body falls approximately linearly. However, after about 10 s, drag becomes significant, motion becomes noticeably nonlinear, and the two filters differ significantly. The velocity estimates are shown in Fig. 2 and indicate large error spikes in both filters. These occur when the altitude of the body is the same as that of the radar and range information provides less data about body motion. The new filter recovers quickly, but the EKF has a larger error spike and only slowly converges. Fig. 3 shows the errors in estimating $x_3(\cdot)$. The error in the EKF estimate is biased and is an order of magnitude larger than that for the new filter.

In Figs. 4 and 5, we show the errors in the position estimates made by the EKF and the new filter and the associated estimates of the two standard deviation bounds. These bounds are given by twice the square root of the diagonals of the covariance matrix, and, if the filter is consistent, the state errors should lie within these bounds 95% of the time. However, the error in the EKF drifts outside of these bounds after 30 s, showing that it does not yield consistent estimates. However, the errors in the new filter always lie well within the two standard deviations, implying that the new filter is consistent.
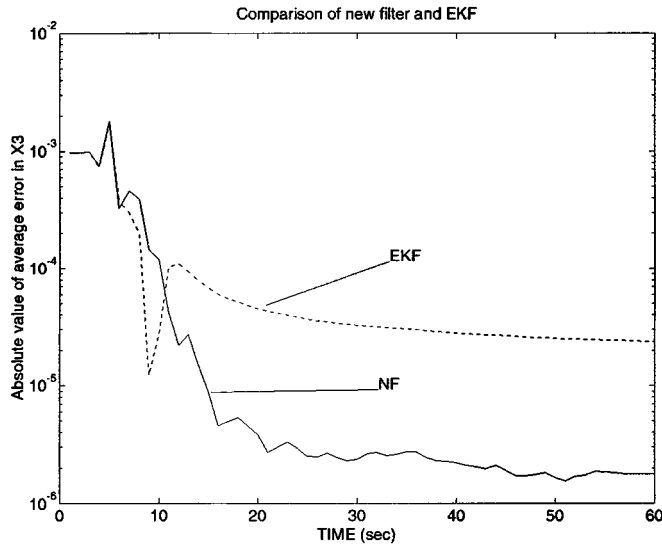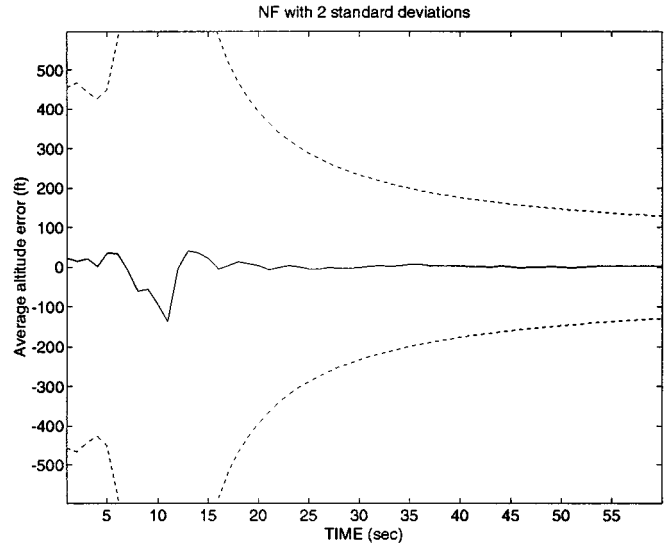
Fig. 3. Absolute mean error in x_3.
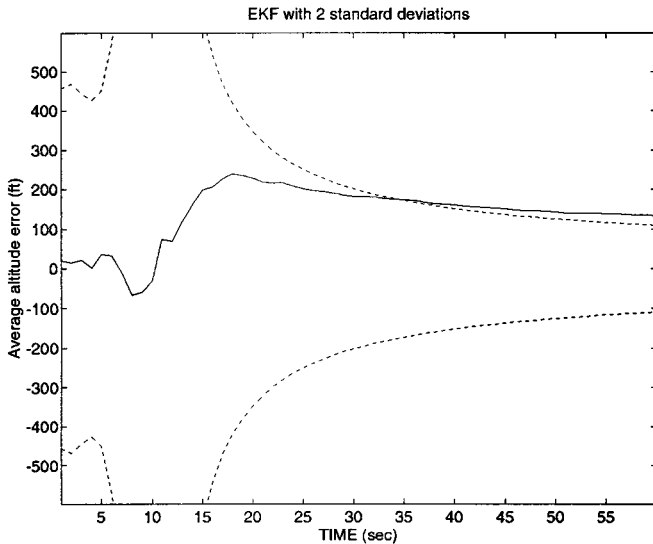

Fig. 5. New filter mean position error.

Fig. 4. EKF mean position error.

Therefore, we conclude that in this example the new filter has substantial advantages over the EKF both in implementation and performance.

V. CONCLUSIONS

Motivated by the deficiencies of the EKF, we have examined a completely new approach for applying linear estimation theory to nonlinear systems. Rather than approximate the Taylor series to an arbitrary order, we approximate the first three moments of the prior distribution accurately using a set of samples. The algorithm predicts the mean and covariance accurately up to the third order and, because the higher order terms in the series are not truncated, it is possible to reduce the errors in the higher order terms as well. We have provided empirical evidence that supports the theoretical conclusion and have demonstrated that the new filter is far easier to implement because it does not involve any linearization steps, eliminating the derivation and evaluation of Jacobian matrices.

This algorithm has found a number of applications in high-order, nonlinear coupled systems, including navigation systems for high-speed road [21], [16] vehicles, public transportation systems [22], and underwater vehicles [23]. Square root filters can be formulated by discrete sets of points (as demonstrated in [24]), and iterated filters can be constructed using the predictions [25]. The algorithm has been extended to capture the first four moments of a Gaussian distribution [12] and the first three moments of an arbitrary distribution [20].

Given its performance and implementation advantages, we conclude that the new filter should be preferred over the EKF in virtually all nonlinear estimation applications.

APPENDIX I
HIGHER MOMENTS OF THE APPROXIMATION

This appendix analyzes the properties of the higher moments of the sigma point selection scheme of (3). The selection process consists of three stages. First, a set of sigma points are drawn to approximate an n-dimensional standard Gaussian with mean $\mathbf{0}$ and covariance $\mathbf{I}$. A linear transformation [a matrix square root $\mathbf{P}(k|k)$] is applied to each point so that the transformed sample still has mean $\mathbf{0}$, but the covariance is $\mathbf{P}(k|k)$. Finally, $\hat{\mathbf{x}}(k|k)$ is added to each transformed sigma point to ensure the correct sample mean. κ affects the first stage of the approximation, and the matrix square root affects the second.

The sigma points, which are assumed approximate the standard Gaussian, lie on the coordinate axes.[2] The orthogonality and the symmetry of these points means that the only nonzero sample moments are those that are an even order power of a single coefficient. The $2k$th-order moment for any coefficient is $(n + \kappa)^{(k-1)}$. Therefore, the higher order moments scale geometrically by a factor determined by κ. However, the moments for the standard Gaussian are different from those of the sigma points, which approximate it. Because the covariance matrix is $\mathbf{I}$, the different components are independent of one another and

$$E[x_1^{\alpha_1} \times \cdots \times x_n^{\alpha_n}] = \prod_{i=1}^{n} E[x_i^{\alpha_i}]$$

[2]In effect, these points are drawn from the *orthogonal* matrix square root of $\mathbf{I}$.

where the ith component in the moment is raised to the power α_i. It can be shown that [13]

$$E[x_i^{\alpha_i}] = \begin{cases} 0, & \alpha_i \text{ is odd.} \\ 1 \times 3 \times \cdots \times (\alpha_i - 1), & \alpha_i \text{ is even.} \end{cases}$$

Two differences exist in the approximation by the sigma points. First, the "scaling" of nonzero moments are different. Second, a number of moment terms are "missed out." These differences become more marked as the order increases, but, as explained in Appendix II, they only affect the higher order terms of the Taylor Series of the process model/nonlinear function. Assuming that the contribution of terms diminishes as the order increases, the emphasis is to minimize the errors in the lowest order terms. Because $E[x_i^4] = 3$ for the Gaussian distribution, choosing $(n + \kappa) = 3$ minimizes the difference between the moments of the standard Gaussian and the sigma points up to the fourth order.

Because the fourth and higher order moments are not precisely matched, the choice of matrix square root affects the errors in the higher order terms by adjusting the way in which the errors are distributed among the different states. However, in general, this information cannot be exploited because it would require knowledge of the higher order derivatives of the process model equation. In the absence of this information, the choice of matrix square root is governed by other issues, such as numerical stability or computational cost.

APPENDIX II
MOMENT APPROXIMATION AND PERFORMANCE

In this appendix we provide a justification for approximating a probability distribution by a set of samples that match its moments. We show that an approximation is only correct to the mth order if its moments up to that order are correctly approximated.

Consider a Gaussian-distributred random variable $\mathbf{x}$ with mean $\overline{\mathbf{x}}$ and covariance $\mathbf{P}_{xx}$. We wish to calculate the mean $\overline{\mathbf{y}}$ and covariance $\mathbf{P}_{yy}$ of the random variable $\mathbf{y}$, which is related to $\mathbf{x}$ through the nonlinear (analytic) function $\mathbf{y} = \mathbf{f}[\mathbf{x}]$.

Noting that $\mathbf{x}$ can be written as $\mathbf{x} = \overline{\mathbf{x}} + \tilde{\mathbf{x}}$, where $\tilde{\mathbf{x}}$ is a zero-mean Gaussian random variable with covariance $\mathbf{P}_{xx}$, the nonlinear transformation can be expanded as a Taylor Series about $\overline{\mathbf{x}}$

$$\mathbf{y} = \mathbf{f}[\overline{\mathbf{x}} + \tilde{\mathbf{x}}]$$
$$= \mathbf{f}[\overline{\mathbf{x}}] + \mathbf{D}_{\Delta x}\mathbf{f} + \frac{\mathbf{D}_{\Delta x}^2\mathbf{f}}{2!} + \frac{\mathbf{D}_{\Delta x}^3\mathbf{f}}{3!} + \frac{\mathbf{D}_{\Delta x}^4\mathbf{f}}{4!} + \cdots \quad (10)$$

where the $\mathbf{D}_{\Delta x}\mathbf{f}$ operator evaluates the total differential of $\mathbf{f}[\cdot]$ when perturbed around a nominal value $\overline{\mathbf{x}}$ by $\tilde{\mathbf{x}}$. The ith term in the Taylor series for $\mathbf{f}[\cdot]$ is given by

$$\frac{\mathbf{D}_{\Delta x}^i\mathbf{f}}{i!} = \frac{1}{i!}\left(\sum_{j=1}^n \tilde{x}_j \frac{\partial}{\partial x_j}\right)^i \mathbf{f}[\mathbf{x}]\Bigg|_{\mathbf{x}=\overline{\mathbf{x}}} \quad (11)$$

where $\tilde{x}_j$ is the jth component of $\tilde{\mathbf{x}}$. Therefore, the ith term in the series is an ith-order polynomial in the coefficients of $\tilde{\mathbf{x}}$, whose coefficients are given by derivatives of $\mathbf{f}[\cdot]$.

$\overline{\mathbf{y}}$ is the expected value of (10)

$$\overline{\mathbf{y}} = E[\mathbf{f}[\overline{\mathbf{x}} + \tilde{\mathbf{x}}]]$$
$$= \mathbf{f}[\overline{\mathbf{x}}] + E\left[\mathbf{D}_{\Delta x}\mathbf{f} + \frac{\mathbf{D}_{\Delta x}^2\mathbf{f}}{2!} + \frac{\mathbf{D}_{\Delta x}^3\mathbf{f}}{3!} + \frac{\mathbf{D}_{\Delta x}^4\mathbf{f}}{4!} + \cdots\right] \quad (12)$$

where the ith term in the series is given by

$$E\left[\frac{\mathbf{D}_{\Delta x}^i\mathbf{f}}{i!}\right] = \frac{1}{i!} E\left[\left(\sum_{j=1}^n \tilde{\mathbf{x}}_j \frac{\partial}{\partial x_j}\right)^i\right] \mathbf{f}[\overline{\mathbf{x}}]$$

$$= \frac{1}{i!}\left(m_{1\cdots11}\frac{\partial^i\mathbf{f}}{\partial x_1^i} + m_{1\cdots12}\frac{\partial^i\mathbf{f}}{\partial x_1^{i-1}\partial x_2} + \cdots\right) \quad (13)$$

and $m_{c_1 c_2 \cdots c_i}$ is the ith-order moment $m_{c_1 c_2 \cdots c_i} = E[\tilde{x}_{c_1} \tilde{x}_{c_2} \cdots \tilde{x}_{c_i}]$.

The ith term is a function of the ith central moments of the distribution of $\tilde{\mathbf{x}}$. Therefore, for an approximation to be accurate up to the ith order, it must be able to approximate the moments correctly up to the ith order. The EKF truncates this series after the first term, and so its error in predicting the mean is in the second and higher orders. The new filter matches the mean and covariance correctly, and so it is correct up to the second order. Because the new filter does not truncate the filter at any order, it can be expected that the errors in the fourth and higher order terms are smaller than those committed by the EKF.

The covariance is given by $\mathbf{P}_{yy} = E[\{\mathbf{y} - \overline{\mathbf{y}}\}\{\mathbf{y} - \overline{\mathbf{y}}\}^T]$. Now

$$\mathbf{y} - \overline{\mathbf{y}} = \mathbf{f}[\overline{\mathbf{x}} + \tilde{\mathbf{x}}] - E[\mathbf{f}[\overline{\mathbf{x}} + \tilde{\mathbf{x}}]]$$
$$= \mathbf{D}_{\Delta x}\mathbf{f} + \frac{\mathbf{D}_{\Delta x}^2\mathbf{f}}{2!} + \frac{\mathbf{D}_{\Delta x}^3\mathbf{f}}{3!} + \frac{\mathbf{D}_{\Delta x}^4\mathbf{f}}{4!}$$
$$- E\left[\frac{\mathbf{D}_{\Delta x}^2\mathbf{f}}{2!} + \frac{\mathbf{D}_{\Delta x}^4\mathbf{f}}{4!} + \cdots\right] \quad (14)$$

with substitutions from (10) and (12). The true covariance is found by postmultiplying the state error by the transpose of itself and taking expectations. Exploiting the fact that $\tilde{\mathbf{x}}$ is symmetric, the odd terms all evaluate to zero and the covariance is given by

$$\mathbf{P}_{yy} = \nabla\mathbf{f}_x\mathbf{P}_{xx}\nabla\mathbf{f}_x^T$$
$$+ E\left[\frac{\mathbf{D}_{\Delta x}\mathbf{f}(\mathbf{D}_{\Delta x}^3\mathbf{f})^T}{3!} + \frac{\mathbf{D}_{\Delta x}^2\mathbf{f}(\mathbf{D}_{\Delta x}^2\mathbf{f})^T}{2 \times 2!}\right.$$
$$+ \left.\frac{\mathbf{D}_{\Delta x}^3\mathbf{f}(\mathbf{D}_{\Delta x}\mathbf{f})^T}{3!}\right] - E\left[\frac{\mathbf{D}_{\Delta x}^2\mathbf{f}}{2}\right] E\left[\frac{\mathbf{D}_{\Delta x}^2\mathbf{f}}{2}\right]^T + \cdots$$
$$(15)$$

where $E[\mathbf{D}_{\Delta x}\mathbf{f}(\mathbf{D}_{\Delta x}\mathbf{f})^T] = \nabla\mathbf{f}_x\mathbf{P}_{xx}\nabla\mathbf{f}_x^T$ has been used. Comparing this with (2), it can be seen that the EKF truncates this series after the first term. The new filter does not truncate the series at any arbitrary order, and, applying the analysis from above, it is correct up to the second order with errors in the fourth and higher order terms. Therefore, both the EKF and the new filter predict the covariance correctly up to the third order.

APPENDIX III
THE MODIFIED FORM OF THE ALGORITHM

When κ is negative, it is possible that the predicted covariance will not be positive semidefinite. This can be demonstrated by taking the limit of (15) as $(n + \kappa) \to 0$. As shown in Appendix I, the fourth and higher order moments tend to zero and

$$\lim_{(n+\kappa)\to0} \mathbf{P}_{yy} = \nabla\mathbf{f}_x\mathbf{P}_{xx}\nabla\mathbf{f}_x^T - E\left[\frac{\mathbf{D}_{\Delta x}^2\mathbf{f}}{2}\right] E\left[\frac{\mathbf{D}_{\Delta x}^2\mathbf{f}}{2}\right]^T.$$

It can be seen that a fourth-order, positive-semidefinite matrix is subtracted. However, this term originates from the outer product of two second-order expected terms and does not scale with κ.

The modified form of the algorithm evaluates the covariance about the projected mean. Taking the sigma points for $\mathbf{x}$ are $\mathcal{X}_i$ and those for $\mathbf{y}$ are $\mathcal{Y}_i$, the modified form for calculating $\mathbf{P}_{yy}$ is[3]

$$\mathbf{P}_{yy}^{\text{MOD}} = \sum_{i=1}^{2n} W_i[\mathcal{Y}_i - \mathcal{Y}_0][\mathcal{Y}_i - \mathcal{Y}_0]^T.$$

[3]This is equivalent to calculating the mean and covariance using (4) and (5) and adding a term $\{\overline{\mathbf{y}} = \mathcal{Y}_0\}\{\overline{\mathbf{y}} - \mathcal{Y}_0\}^T$.

Positive semidefiniteness is guaranteed by the fact that the covariance matrix is evaluated as the sum of outer products of vectors. It can be shown that the calculated covariance is

$$\mathbf{P}_{yy} = \nabla \mathbf{f}_x \mathbf{P}_{xx} \nabla \mathbf{f}_x^T$$
$$+ E\left[\frac{\mathbf{D}_{\Delta x}\mathbf{f}(\mathbf{D}_{\Delta x}^3\mathbf{f})^T}{3!} + \frac{\mathbf{D}_{\Delta x}^2\mathbf{f}(\mathbf{D}_{\Delta x}^2\mathbf{f})^T}{2 \times 2!} \right.$$
$$\left. + \frac{\mathbf{D}_{\Delta x}^3\mathbf{f}(\mathbf{D}_{\Delta x}\mathbf{f})^T}{3!} \right] + \cdots$$

which introduces errors in the fourth and higher orders.

The modified form has the advantage that, irrespective of the choice of the matrix square root

$$\lim_{(n+\kappa)\to 0} \overline{\mathbf{y}} = \mathbf{f}[\overline{\mathbf{x}}] + E\left[\frac{\mathbf{D}_{\Delta x}^2\mathbf{f}}{2!} \right]$$
$$\lim_{(n+\kappa)\to 0} \mathbf{P}_{yy} = \nabla \mathbf{f}_x \mathbf{P}_{xx} \nabla \mathbf{f}_x^T.$$

These are the values calculated by the modified, truncated second-order filter [4], but without the need to evaluate Jacobians or Hessians.[4]

ACKNOWLEDGMENT

REFERENCES

[1] J. K. Uhlmann, "Algorithms for multiple target tracking," *Am. Sci.*, vol. 80, no. 2, pp. 128–141, 1992.
[2] H. W. Sorenson, *Kalman Filtering: Theory and Application*. New York: IEEE Press, 1985.
[3] H. J. Kushner, "Dynamical equations for optimum non-linear filtering," *J. Differential Equations*, vol. 3, pp. 179–190, 1967.
[4] P. S. Maybeck, *Stochastic Models, Estimation, and Control*, P. S. Maybeck, Ed. New York: Academic, 1982, vol. 2.
[5] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic, 1970.
[6] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *Proc. Inst. Elect. Eng. F*, vol. 140, no. 2, pp. 107–113, Apr. 1993.
[7] H. J. Kushner, "Approximations to optimal nonlinear filters," *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 546–556, Oct. 1967.
[8] H. W. Sorenson and A. R. Stubberud, "Non-linear filtering by approximation of the *a posteriori* density," *Int. J. Contr.*, vol. 8, no. 1, pp. 33–51, 1968.
[9] F. E. Daum, "New exact nonlinear filters," in *Bayesian Analysis of Time Series and Dynamic Models*, J. C. Spall, Ed. New York: Marcel Dekker, 1988, pp. 199–226.
[10] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *Proc. Am. Contr. Conf.*, Seattle, WA, 1995, pp. 1628–1632.
[11] S. J. Julier and J. K. Uhlmann. (1994, Aug.) A general method for approximating nonlinear transformations of probability distributions. [Online]. Available: http://www.robots.ox.ac.uk/~siju
[12] ——, "A new extension of the Kalman filter to nonlinear systems," in *Proc. AeroSense: 11th Int. Symp. Aerosp./Defense Sensing, Simulat. Contr.*, Orlando, FL, 1997.
[13] P. S. Maybeck, *Stochastic Models, Estimation, and Control*. New York: Acdemic, 1979, vol. 1.
[14] M. Athans, R. P. Wishner, and A. Bertolini, "Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements," *IEEE Trans. Automat. Contr.*, vol. 13, pp. 504–518, Oct. 1968.
[15] D. E. Catlin, "Estimation, control and the discrete Kalman filter," in *Applied Mathematical Sciences 71*. New York: Springer-Verlag, 1989, p. 84.
[16] S. Clark, "Autonomous land vehicle navigation using millimetre wave radar," Ph.D. dissertation, University of Sydney, Australia, 1999.
[17] D. Lerro and Y. K. Bar-Shalom, "Tracking with debiased consistent converted measurements vs. EKF," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, pp. 1015–1022, July 1993.
[18] J. K. Uhlmann, "Simultaneous map building and localization for real time applications," Univ. Oxford, U.K., Tech. Rep., 1994.
[19] S. J. Julier and J. K. Uhlmann, "A consistent, debiased method for converting between polar and Cartesian coordinate systems," in *Proc. AeroSense: 11th Int. Symp. Aerosp./Defense Sensing, Simulat. Contr.*, Orlando, FL, 1997.
[20] S. J. Julier, "A skewed approach to filtering," in *Proc. AeroSense: 12th Int. Symp. Aerosp./Defense Sensing, Simulat. Contr.*, Orlando, FL, 1998.
[21] ——, "Comprehensive process models for high-speed navigation," Ph.D. dissertation, Univ. Oxford, U.K., 1997.
[22] A. Montobbio, "Sperimentazione ed affinamento di un localizzatore," B.S. thesis, Politecnico di Torino, Italy, 1998.
[23] R. Smith, "Navigation of an underwater remote operated vehicle," Univ. Oxford, Tech. Rep., 1995.
[24] T. S. Schei, "A finite difference approach to linearization in nonlinear estimation algorithms," in *Proc. Am. Contr. Conf.*, vol. 1, Seattle, WA, 1995, pp. 114–118.
[25] R. L. Bellaire, E. W. Kamen, and S. M. Zabin, "A new nonlinear iterated filter with applications to target tracking," in *Proc. AeroSense: 8th Int. Symp. Aerosp./Defense Sensing, Simulat. Contr.*, vol. 2561, Orlando, FL, 1995, pp. 240–251.
[26] B. M. Quine, J. K. Uhlmann, and H. F. Durrant-Whyte, "Implicit Jacobians for linearized state estimation in nonlinear systems," in *Proc. Amer. Contr. Conf.*, Seattle, WA, 1996, pp. 1645–1646.

An $O(T^2)$ Boundary Layer in Sliding Mode for Sampled-Data Systems

Wu-Chung Su, Sergey V. Drakunov, and Ümit Özgüner

Abstract—The use of a discontinuous control law (typically, sign functions) in a sampled-data system will bring about chattering phenomenon in the vicinity of the sliding manifold, leading to a boundary layer with thickness $O(T)$, where T is the sampling period. However, by proper consideration of the sampling phenomenon in the discrete-time sliding mode control design, the thickness of the boundary layer can be reduced to $O(T^2)$. In contrast to discontinuous control for continuous-time VSS, the discrete-time sliding mode control need not be of switching type.

Index Terms—Boundary layer, discrete-time sliding mode, sampled-data systems.