UD Covariance Factorization for Unscented Kalman Filter using Sequential Measurements Update

H. Ghanbarpour Asl, and S. H. Pourtakdoust

Abstract—Extended Kalman Filter (EKF) is probably the most widely used estimation algorithm for nonlinear systems. However, not only it has difficulties arising from linearization but also many times it becomes numerically unstable because of computer round off errors that occur in the process of its implementation. To overcome linearization limitations, the unscented transformation (UT) was developed as a method to propagate mean and covariance information through nonlinear transformations. Kalman filter that uses UT for calculation of the first two statistical moments is called Unscented Kalman Filter (UKF). Square-root form of UKF (SR-UKF) developed by Rudolph van der Merwe and Eric Wan to achieve numerical stability and guarantee positive semi-definiteness of the Kalman filter covariances. This paper develops another implementation of SR-UKF for sequential update measurement equation, and also derives a new UD covariance factorization filter for the implementation of UKF. This filter is equivalent to UKF but is computationally more efficient.

Keywords—Unscented Kalman filter, Square-root unscented Kalman filter, UD covariance factorization, Target tracking.

I. INTRODUCTION

ONE of the most fundamental tasks in filtering and estimation is to calculate the statistics of a random variable which has undergone a transformation. Kalman filter, for example, uses two of such transformations. However when a transformation is nonlinear, no general closed form solutions exist [1], [2]. 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 proposed [2], [4]. These methods can be broadly classified as numerical Monte Carlo [5] methods or analytical approximations [6], [7]. 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 makes it widely utilized [8]. To handle this problem, many approximate methods have been proposed [9-12].

EKF simply applies the Taylor series expansion to the nonlinear system along with the observation equations, and takes terms only to the first order while the probability density function (PDF) is approximated by a Gaussian distribution. In

practice however, EKF has shown several limitations and easily exhibits divergent characteristics, [4], [10], and [14].

The development of the UT to approximate two of the first statistical moments was pioneered by Julier and Uhlmann [8], [13]. The UKF was developed based on UT with the underlying assumption that approximating a Gaussian distribution is easier than approximating a nonlinear transformation [8],[15]. The UKF uses deterministic sampling to approximate the state distribution as a Gaussian Random Variable (GRV). The sigma points are chosen to capture the true mean and covariance of state distribution and are propagated through the nonlinear system. The posterior mean and covariance are then calculated from the propagated sigma points. The UKF determines the mean and covariance accurately to the second order [13], while the EKF is only able to obtain first order accuracy [13]. Therefore the UKF provides better state estimates for nonlinear systems [15]. However UKF requires calculation the new set sigma points at each sample time which requires taking a matrix square-root of the state covariance matrix. While the square-root of covariance matrix is an integral part of the UKF, it is still the full covariance which is recursively updated. In the SR-UKF implementation, square-root matrix of state covariance will be propagated directly, avoiding the need to refactorize at each time step [16]. SRUKF uses QR factorization method, Cholesky of the Rank 1 update and efficient least square solution for linear systems [22], [23], [24].

In this study another implementation of SR-UKF like Potter's filter for vector measurements is proposed. This filter considers scalar measurement and sequentially updates the covariance matrix and state estimation [17], [18].

Also a UD (unit upper triangular matrix) covariance factorization is considered which makes our filter UD-UKF. This methodology produces comparable results to those of the SR-UKF, However with a computationally more efficient algorithm. Another advantage of our newly developed filter is that contrary to SR-UKF that uses rank 1 Cholesky update for square-root covariance matrix, no Cholesky updates is required for UD-UKF at all.

II. UNSCENTED KALMAN FILTER

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

Authors are with Department of Aerospace Engineering, Sharif University of Technology, Tehran, Iran (e-mail: habib_oxtay@yahoo.com).

system

$$x_k = f(x_{k-1}, u_{k-1}, w_{k-1})$$
$$y_k = h(x_k, u_k) + v_k \qquad (1)$$

Where $x_k \in \Re^{n_x}$ is the state of the system at time step k, $u_k \in \Re^{n_u}$ is the input vector, $w_k \in \Re^{n_w}$ is the noise process caused by disturbance and modeling errors, $y_k \in \Re^{n_y}$ is the observation vector and $v_k \in \Re^{n_y}$ is the additive measurement noise. It is assumed that the noise vector w_k, v_k are of zero mean and

$$E(v_i v_i^T) = \delta_{ij} R_i$$
$$E(w_i w_i^T) = \delta_{ij} Q_i \quad , \forall i, j \qquad (2)$$
$$E(v_i w_i^T) = 0$$

The Kalman filter propagates the first two moments of the distribution of x_k recursively and has a distinctive "predictive-corrector" structure. Let $\hat{x}_k$ be the estimate of x_k using the observation information up to and including time k. The covariance of this estimate is P_{x_k}. The recursive estimation for x_k can be expressed in the following from [19], [20], and [21]:

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - y_k^-) \qquad (3)$$
$$P_{x_k} = P_{x_k}^- - K_k P_{\tilde{y}_k} K_k^T \qquad (4)$$

Where $\hat{x}_k^-$ is the optimal prediction of the state at time k conditioned on all of the observed information up to and including time k-1, and $\hat{y}_k^-$ is the optimal prediction of the observation at time k. $P_{x_k}^-$ is the covariance of $\hat{x}_k^-$ and $P_{\tilde{y}_k}$ is the covariance of $r_k = y_k - \hat{y}_k^-$, termed the innovation process. The optimal parameters of this recursion are given by

$$\hat{x}_k^- = E[f(x_{k-1}, u_{k-1}, w_{k-1})] \qquad (5)$$
$$\hat{y}_k^- = E[h(x_k, u_k) + v_k] \qquad (6)$$
$$K_k = P_{x_k y_k} P_{\tilde{y}_k}^{-1} = E[(x_k - \hat{x}_k^-)(y_k - \hat{y}_k^-)^T] \, E[(y_k - \hat{y}_k^-)(y_k - \hat{y}_k^-)^T]^{-1} \qquad (7)$$

EKF calculates these quantities using linear functions, but UKF calculates these quantities from a set of weighted samples (sigma-points) that are deterministically calculated using the mean and square-root decomposition of the covariance matrix of x_{k-1}, w_{k-1} and v_k. When propagated through the nonlinear transformation, it captures the posterior covariance (3rd order accuracy is achieved if the prior random variable has a symmetric distribution, such as the exponential family of PDE) [13].

The pseudo-code for UKF is given in Table I. In UKF state random variable (RV) is redefined as the concatenation of the original state plus the noise variables in an augmented state vector form $x_k^a = \begin{bmatrix} x_k^T & w_k^T & v_k^T \end{bmatrix}^T$. The sigma points selection scheme is applied to this new augmented state to calculate the corresponding sigma-point set, $\{\chi_{k,i}^a; \ i = 0, ..., 2L\}$, where $L = n_x + n_w + n_v$ and $\chi_{k,i}^a \in \Re^{2L+1}$. n_x, n_w and n_v are dimensions of state, noise process and noise of the measurements respectively. $\chi^a = \begin{bmatrix} (\chi^a)^T & (\chi^w)^T & (\chi^v)^T \end{bmatrix}^T$ is the augmented sigma points that are of dimension $L \times (2 \times L + 1)$. γ is scaling parameter that determines the spread of the sigma-points matrix around the prior mean. Q_k, R_k are covariances of the process and measurement noise processes respectively.

TABLE I
PSEUDO-CODE FOR UNSCENTED KALMAN FILTER

● Initialization

$$\hat{x}_0 = E[x_0] \quad , \quad P_{x_0} = E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T]$$

$$\hat{x}_0^a = E[x_0^a] = \begin{bmatrix} \hat{x}_0^T & \overline{w}_0^T & \overline{v}_0^T \end{bmatrix}^T$$

$$P_0^a = \begin{bmatrix} P_{x_k} & 0 & 0 \\ 0 & Q_k & 0 \\ 0 & 0 & R_k \end{bmatrix}$$

● For $k = 1, ..., \infty$

● Set $t = k - 1$

● Calculate sigma-points:

$$\chi_t^a = \begin{bmatrix} \hat{x}_t^a & \hat{x}_t^a + \gamma S_{x_t}^a & \hat{x}_t^a - \gamma S_{x_t}^a \end{bmatrix}$$

● time update equations:

$$\chi_{k|t}^x = f(\chi_t^x, u_t, \chi_t^w)$$

$$\hat{x}_k^- = \sum_{i=0}^{2L} w_i^m \chi_{i,k|t}^x$$

$$P_{x_k}^- = \sum_{i=0}^{2L} w_i^c (\chi_{i,k|t}^x - \hat{x}_k^-)(\chi_{i,k|t}^x - \hat{x}_k^-)^T$$

Measurement-update equations:

$$Y_{k|t} = h(\chi_t^x, u_k) + \chi_t^v$$

$$y_k^- = \sum_{i=0}^{2L} w_i^m Y_{i,k|t}$$

$$P_{\tilde{y}_k} = \sum_{i=0}^{2L} w_i^c (Y_{i,k|t} - y_k^-)(Y_{i,k|t} - y_k^-)^T$$

$$P_{x_k y_k} = \sum_{i=0}^{2L} w_i^c (\chi_{i,k|t}^x - \hat{x}_k^-)(Y_{i,k|t} - y_k^-)^T$$

$$K_k = P_{x_k y_k} P_{\tilde{y}_k}^{-1}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - y_k^-)$$

$$P_{x_k} = P_{x_k}^- - K_k P_{\tilde{y}_k}^{-1} K_k^T$$

End for

Where $\{w^i\}$ is a set of scalar weights, $w_0^m = \dfrac{\lambda}{L+\lambda}$,

$$w_0^c = \frac{\lambda}{L+\lambda} + (1 - \alpha^2 + \beta)$$

And

$$w_i^m = w_i^c = \frac{1}{2(L+\lambda)}, \quad i = 1,...,2L,$$

$\lambda = \alpha^2(L+\kappa) - L$ and $\gamma = \sqrt{L+\lambda}$. The constant α determines the spread of the sigma points around the prior mean. Typical range for α is $1e-3 < \alpha \le 1$. κ is a tertiary scaling factor and is usually set equal to 0. β is the secondary scaling factor used to emphasize the weighting on the zero's sigma-point for the posterior covariance calculation. β can be used to minimize certain higher-order error terms based on known moments of the prior random variable (RV). For Gaussian priors, $\beta = 2$ is optimal.

This algorithm requires factorizing the square-root form of $P_k^a = S_k^a (S_k^a)^T$ at each iteration; however this filter propagates the covariance of the states and is usually very sensitive to round off errors causing numerically instability.

III. SQUARE-ROOT UKF

As in the original UKF, the filter is initialized by calculating the matrix square-root of the state covariance once via a Cholesky decomposition method. However the propagated and updated Cholesky factor is then used in subsequent iterations to directly form the sigma points. For this reason in SR-UKF one should calculate and propagate the Square-root form of the state covariance matrix. Therefore, for generation of sigma-points, one should augment the square-root from of sate and noise covariance matrices.

The pseudo-code for UKF is given in Table II [21-24].

TABLE II
PSEUDO-CODE FOR SR-UKF

● Initialization

$$\hat{x}_0 = E[x_0] \quad , \quad S_{x_0} = \sqrt{E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T]}$$

$$\hat{x}_0^a = E[x_0^a] = \begin{bmatrix} \hat{x}_0^T & \overline{w}_0^T & \overline{v}_0^T \end{bmatrix}^T$$

$$S_0^a = \begin{bmatrix} S_{x_k} & 0 & 0 \\ 0 & S_{Q_k} & 0 \\ 0 & 0 & S_{R_k} \end{bmatrix}$$

● For $k = 1,...,\infty$

● Set $t = k-1$

● Calculate sigma-points:

$$\chi_t^a = \begin{bmatrix} \hat{x}_t^a & \hat{x}_t^a + \gamma S_{x_t}^a & \hat{x}_t^a - \gamma S_{x_t}^a \end{bmatrix}$$

● Time update equations:

$$\chi_{k|t}^x = f(\chi_t^x, u_t, \chi_t^w)$$

$$\hat{x}_k^- = \sum_{i=0}^{2L} w_i^m \chi_{i,k|t}^x$$

$$[Q_{x_k}^-, S_{x_k}^-] = qr\left\{\sqrt{w_1^c}(\chi_{1:2L,k|t}^x - x_k^-)\right\}$$

$$S_{x_k}^x = cholupdate\{S_{x_k}^-, (\chi_{0,k|t}^x - x_k^-), w_0^c\}$$

$$Y_{k|t} = h(\chi_{k|t}^x, u_t) + \chi_t^v$$

$$y_k^- = \sum_{i=0}^{2L} w_i^m Y_{i,k|t}$$

● Measurement-update equations:

$$[Q_{\tilde{y}_k}, S_{\tilde{y}_k}] = qr\left\{\sqrt{w_1^c}(Y_{1:2L,k|t} - y_k^-)\right\}$$

$$S_{\tilde{y}_k} = cholupdate\{S_{\tilde{y}_k}, (Y_{1:2L,k|t} - y_k^-), w_0^c\}$$

$$P_{x_k y_k} = \sum_{i=0}^{2L} w_i^c (\chi_{i,k|t}^x - \hat{x}_k^-)(Y_{i,k|t} - y_k^-)^T$$

$$K_k = (P_{x_k y_k} / S_{\tilde{y}_k}^T) / S_{\tilde{y}_k}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - y_k^-)$$

$$U = K_k S_{\tilde{y}_k}$$

$$S_{x_k} = cholupdate\{S_{x_k}^-, U, -1\}$$

Now consider determination of $S_{x_k}^-$ from sigma points. From UKF,

$$P_{x_k}^- = S_{x_k}^-(S_{x_k}^-)^T = \sum_{i=0}^{2L} w_i^c (\chi_{i,k|t}^x - \hat{x}_k^-)(\chi_{i,k|t}^x - \hat{x}_k^-)^T$$

$$= S_{x_k}^- \theta \theta^T S_{x_k}^{-T} = AA^T \tag{8}$$

Where θ is a part of orthonormal matrix, $S_{x_k}^- \in R^{n_x \times n_x}$, $\theta \in R^{n_x \times (2L+1)}$ and $S_{x_k}^-$ is the triangular part of QR decomposition of A, and A is [Appendix 2] :

$$A = \begin{bmatrix} \sqrt{w_0^c}(\chi_{0,k|t}^x - \hat{x}_k^-) & \sqrt{w_1^c}(\chi_{1:2L,k|t}^x - \hat{x}_k^-) \end{bmatrix} \tag{9}$$

Equation (9) assumes $w_0^c \geq 0$ which is not necessarily the case, especially for small α (note that $w_1^c \geq 0$ is always true). The correct way to handle $w_0^c < 0$ is with a Cholesky down date algorithm (appendix 1). Therefore UKF uses a QR decomposition of compound matrix containing the weighted propagated sigma-points for calculating Cholesky factor of $P_{x_k}^-$. The subsequent Cholesky update (or down date) in SR-UKF is necessary since the zero's weight w_0^c, may be negative. The same two-step approach is applied to the calculation of Cholesky factor $S_{\widetilde{y}_k}$ of the observation prediction error covariance. The Kalman gain can be calculated directly from the original UKF, but one can use two nested inverse solution for Kalman gain that have simpler algorithm [18]. For posterior measurement update of Cholesky factor of the state covariance, Cholesky down dates of $S_{x_k}^-$ are sequentially applied. The down date vectors are the columns of $U = K_k S_{\widetilde{y}_k}$, that is the square root of $U = K_k P_{\widetilde{y}_k} K_k^T$. This algorithm is more efficient compared with the original UKF and while guarantees the semi-positive form of the state covariance matrix, has less computational load compared to that of the original UKF.

IV. SEQUENTIAL SQUARE-ROOT UKF

In this proposed new filter, another approach similar to Potter's filter [17] [18] with scalar measurement is taken for the implementation of SR-UKF, but for multi-dimensional measurement systems with sequential updates. This algorithm does not need Cholesky update to achieve S_{x_k} or the inversion of two time $n_y \times n_y$ upper triangular matrix. To start consider the Square-root matrices of R_k and Q_k which are $\sqrt{R_k}$ and $\sqrt{Q_k}$, respectively. By multiplying the measurement equation with inverse of $\sqrt{R_k}$, a new measurement equation results:

$$\widetilde{y}_k = \left(\sqrt{R_k}\right)^{-1}\left[h(x_k,u_k) + \widetilde{v}_k\right] = \widetilde{h}(x_k,u_k) + \widetilde{v}_k \quad (10)$$

where $\widetilde{v}_k$ can be considered as another noise process:

$$E(\widetilde{v}_k) = \overline{v}_k = 0 \text{ And } \widetilde{R}_k = E(\widetilde{v}_k \widetilde{v}_k^T) = I_{n_y \times n_y} \quad (11)$$

With the property of: $E(\widetilde{v}_k(i)\widetilde{v}_k(j)) = 0$, $\forall i \neq j$

Now consider a discrete system model with state, z_j defined as follows:

$$z_j = z_{j-1} \quad (12)$$

$$\widetilde{y}_{k,j} = \widetilde{h}_j(z_j,u_k) + \widetilde{v}_{k,j} \quad (13)$$

Where $\widetilde{v}_{k,j}$ represents the j-th element of $\widetilde{v}_k$ and $\widetilde{h}_j$ represents the j-the row of h.

Further, suppose that $z_0 = x_k$ then

$$z_j = x_k \text{ For all } j = 1,...,n_y \quad (14)$$

And $\widetilde{y}_{k,j}$ is j-th component of the new measurement equation described in Eq. (10). Now consider the discrete time Kalman filtering problem for the system given by Eq. (12, 13). Since $\widetilde{R}_k$ is a diagonal matrix, then $\{\widetilde{v}_{k,j}, j = 1:,...,n_y\}$ is a white-noise process with unit covariance. Further, since x_k and $\widetilde{v}_k$ are uncorrelated, then x_k is uncorrelated to each element of $\{\widetilde{v}_{k,j}, j = 1:,...,n_y\}$ and one can apply the discrete-time Kalman filtering problem with one-dimensional measurement and no process noise. Let the initial estimate for z_0 be:

$$\hat{z}_0 = \hat{x}_k^- \quad (15)$$

Where $\hat{x}_k^-$ is least mean square error (LMS) estimate of the system state x_k biased on $y_1,...,y_{k-1}$. Then $\hat{z}_0$ is a linear minimum variance (LMV) estimate of state z_0, biased on the measurements $y_1,...,y_{k-1}$ and has an error square root covariance, denoted by $S_{z,0}$ and given by:

$$S_{z,0} = S_{x_k}^- \quad (16)$$

Then, following Kalman filtering theory, we can sequentially Apply a scalar measurement to system describe by Eq. (12, 13) with initial conditions specified by Eq. (14,15) for n_y (number of measurements) iterations. This way a LMV estimate of state z_m based on $\widetilde{y}_1,...,\widetilde{y}_{k-1},\widetilde{y}_{k,1},...,\widetilde{y}_{k,n_y}$ (denoted by $\hat{z}_m$) or equivalently, x_k biased on $y_1,...,y_k$ (we denoted d by $\hat{x}_k$) is obtained. Further

$$\hat{x}_k = \hat{z}_m , S_{x_k} = S_{z,n_y} \quad (17)$$

Where $S_{\widetilde{y}_k,m}$, is a matrix square root of the error covariance matrix for $\hat{z}_{n_y}$ formed at the n_y-th iteration of the scalar measurement.

In this algorithm it is very important to update $S_{z,j}$ from $S_{z,j-1}$. Following UKF we have:

$$S_{z,j}S_{z,j}^T = S_{z,j-1}S_{z,j-1}^T - S_{z,j-1}Q_{x_k}^- Q_{z,j}^T S_{z,j}^T P_{\widetilde{y}_{k,j}} S_{z,j}Q_{z,j}Q_{x_k}^{-T}S_{z,j-1}^T$$
$$= S_{z,j-1}\left(I - b_j a_j a_j^T\right)S_{z,j-1}^T \quad (18)$$

where:

$$b_j = \frac{1}{P_{\widetilde{y}_{k,j}}} , a_j = S_{z,j-1}Q_{x_k}^- Q_{z,j}^T \quad (19)$$

Note that b_j is a scalar and a_j is a vector of dimension n_x.

If we calculate the Cholesky factor for the term in the parenthesis of Eq. (18), we get:

$$\left(I - b_j a_j a_j^T\right) = \left(I - c_j b_j a_j a_j^T\right)\left(I - c_j b_j a_j a_j^T\right)^T \quad (20)$$

Where c_j is [17], [18] :

$$c_j = \frac{1}{1 + \sqrt{1 - b_j a_j^T a_j}} \quad (21)$$

The pseudo-code for sequential SR-UKF is given in Table III.

<div align="center">

TABLE III
PSEUDO-CODE FOR SSR-UKF
</div>

- Initialization

$$\hat{x}_0 = E[x_0] \quad , \quad S_{x_0} = \sqrt{E\left[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T\right]}$$

$$\hat{x}_0^a = E[x_0^a] = \begin{bmatrix} \hat{x}_0^T & \overline{w}_0^T & \overline{v}_0^T \end{bmatrix}^T$$

$$S_0^a = \begin{bmatrix} S_{x_k} & 0 & 0 \\ 0 & S_{Q_k} & 0 \\ 0 & 0 & I \end{bmatrix}$$

- For $k = 1, \ldots, \infty$
- Set $t = k - 1$
- Calculate sigma-points:

$$\chi_t^a = \begin{bmatrix} \hat{x}_t^a & \hat{x}_t^a + \gamma S_{x_t}^a & \hat{x}_t^a - \gamma S_{x_t}^a \end{bmatrix}$$

- Time update equations:

$$\chi_{k|t}^x = f(\chi_t^x, u_t, \chi_t^w)$$

$$\hat{x}_k^- = \sum_{i=0}^{2L} w_i^m \chi_{i,k|t}^x$$

$$\left[Q_{x_k}^-, S_{x_k}^-\right] = qr\left\{\sqrt{w_1^c}\left(\chi_{1:2L,k|t}^x - x_k^-\right)\right\}$$

$$S_{x_k}^- = cholupdate\{S_{x_k}^-, (\chi_{0,k|t}^x - x_k^-), w_0^c\}$$

$$\widetilde{y}_k = \sqrt{R_k}\, y_k$$

- Measurement-update equations:

Set $\hat{z}_0 = \hat{x}_k^-, S_{z,0} = S_{x_k}^-$

For $j = 1, \ldots, n_y$

$$\widetilde{Y}_{k|t,j} = \widetilde{h}_j(Z_{k|t}, u_t) + \chi_t^{\widetilde{v}}$$

$$\widetilde{y}_{k,j}^- = \sum_{i=0}^{2L} w_i^m \widetilde{Y}_{i,k|t,j}$$

$$\left[Q_{\widetilde{y}_k,j}, S_{\widetilde{y}_k,j}\right] = qr\left\{\sqrt{w_1^c}\left(\widetilde{Y}_{1:2L,k|t,j} - \widetilde{y}_{k,j}^-\right)\right\}$$

$$S_{\widetilde{y}_k,j} = cholupdate\{S_{\widetilde{y}_k,j}, (\widetilde{Y}_{0,k|t,j} - \widetilde{y}_{k,j}^-), w_0^c\}$$

$$P_{\widetilde{y}_{k,j}} = S_{\widetilde{y}_k,j} Q_{\widetilde{y}_k,j} Q_{\widetilde{y}_k,j}^T S_{\widetilde{y}_k,j}^T$$

$$a_j = Q_{x_k}^- Q_{\widetilde{y}_k,j,k}^T S_{\widetilde{y}_k,j}^T$$

$$b_j = \frac{1}{P_{\widetilde{y}_{k,j}}}$$

$$c_j = \frac{1}{1 + \sqrt{1 - b_j a_j^T a_j}}$$

$$K_j = b_j S_{\widetilde{y}_k,j} a_j$$

$$\hat{z}_j = \hat{z}_{j-1} + K_j(\widetilde{y}_{k,j} - \widetilde{y}_{k,j}^-)$$

$$S_{\widetilde{y}_k,j} = S_{\widetilde{y}_k,j-1}\left(I - c_j b_j a_j a_j^T\right)$$

End for

$$\hat{x}_k = \hat{z}_{n_y}, S_{x_k} = S_{z,n_y}$$

V. UD COVARIANCE FACTORIZATION OF UKF WITH SEQUENTIAL UPDATE MEASUREMENT EQUATIONS

The UD covariance factorization of the unscented Kalman filter (UD-UKF) is an error covariance factorization filter of the system state x_k. based on the measurements $y_1, \ldots, y_k$, which is mathematically equivalent to the UKF. Although UD-UKF provides performance comparable to that of the SR-UKF and the original UKF, It is computationally more efficient algorithm. Any symmetric semi-definite matrix can be written in the form of UD factorization [25], [26]. Suppose that:

$$P_{x_k} = U_k D_k U_k^T \quad (22)$$

$$P_{x_{k-1}}^- = U_{k-1}^- D_{k-1}^- U_{k-1}^{-T} \quad (23)$$

$$Q_{k-1} = U_{Q,k-1} D_{Q,k-1} U_{Q,k-1}^T \quad (24)$$

$$R_k = U_{R,k} D_{R,k} U_{R,k}^T \quad (25)$$

where U is an upper triangular matrix with 1's on the diagonal (a unit upper triangular matrix) and D is diagonal matrix with positive elements. In this filter we want to update and propagate U_k and $\sqrt{D_K}$ (Square-root matrix of D_k) for covariance error matrix. Equations (22) and (23) can be written as:

$$P_{x_k} = U_k\sqrt{D_k}\sqrt{D_k}U_k^T = S_{x_k}S_{x_k}^T \quad (26)$$

$$P_{x_k}^- = U_k^-\sqrt{D_k^-}\sqrt{D_k^-}U_k^{-T} = S_{x_k}^- S_{x_k}^{-T} \quad (27)$$

If we propagate and update $\hat{x}_k, U_k$ and D_k, we can drive the sigma points from $\hat{x}_{k-1}$ and $S_{x_{k-1}} = U_{k-1}\sqrt{D_{k-1}}$. By substitution of these sigma-points in the nonlinear transformation (1), we get $\chi_{k|t}^x$, that is a matrix of dimension $n_x \times (2L+1)$. Calculation of $\hat{x}_k^-$ is straightforward and for

covariance of prediction we have:

$$P_{x_k}^- = \left(\chi_{k|t}^x - \hat{\chi}_k^-\right) W^c \left(\chi_{k|t}^x - \hat{\chi}_k^-\right)^T \tag{28}$$

Where W^c is a diagonal matrix of w^c, and $\hat{\chi}_k^- \in \Re^{n_x \times (2L+1)}$ is a matrix with columns $\hat{x}_k^-$. Substituting Eq. (27) in Eq. (28) and factorizing results in: $\left(\chi_{k|t}^x - \hat{\chi}_k^-\right)$ we have:

$$P_{x_k}^- = U_k^- D_k^- U_k^{-T} = R_{x_k}^- Q_{x_k}^- W^- Q_{x_k}^{-T} R_{x_k}^{-T} \tag{29}$$

where $R_{x_k}^-$ is a triangular matrix and $Q_{x_k}^-$ is an orthonormal matrix (Appendix 1). For any triangular matrix we can have (Appendix 3):

$$R_{x_k}^- = U_k^{\sim} D_k^{\sim} \tag{30}$$

Here $U_k^{\sim}$ is triangular with ones on the diagonal, and $D_k^{\sim}$ has square roots of the diagonal elements of $R_{x_k}^-$.

Substituting equation (30) in (29) yields:

$$U_k^- = U_k^{\sim} U_k^{\approx} \quad , \quad D_k^- = D_k^{\approx} \tag{31}$$

where, $U_k^{\approx}, D_k^{\approx}$ are UD factors of the matrix $D_k^{\sim} Q_{x_k}^- W^c Q_{x_k}^{-T} R_{x_k}^{-T} D_k^{\sim T}$.

If we suppose that measurement is scalar, we can have for the posterior covariance:

$$P_{x_k} = P_{x_k}^- - K_k P_{\widetilde{y}_k}^{-1} K_k^T \tag{32}$$

Substituting from equation (29) and (30) gives:

$$U_k D_k U_k^T = U_k^{\sim} D_k^{\sim} Q_k^- W^c \left(U_k^{\sim} D_k^{\sim} Q_k^-\right)^T \tag{33}$$
$$- U_k^{\sim} D_k^{\sim} Q_k^- W^c \left(\widetilde{Y}_{k|t,j} - y_{k,j}^-\right) P_{\widetilde{y}}^{-1} \left(\widetilde{Y}_{k|t,j} - y_{k,j}^-\right) \left(U_k^{\sim} D_k^{\sim} Q_k^- W^c\right)^T$$

From Eq. (33) we have:

$$U_k = U_k^{\sim} \overline{U}_k, \quad D_k = \overline{D}_k \tag{34}$$

Where $\overline{U}_k, \overline{D}_k$ are factors of matrix $G_{k,j}$ that is:

$$G_{k,j} = \begin{bmatrix} D_k^{\sim} Q_k^- W^c \left(U_k^{\sim} D_k^{\sim} Q_k^-\right)^T \\ - U_k^{\sim} D_k^{\sim} Q_k^- W^c \left(\widetilde{Y}_{k|t,j} - y_{k,j}^-\right) P_{\widetilde{y}}^{-1} \left(\widetilde{Y}_{k|t,j} - y_{k,j}^-\right) \left(D_k^{\sim} Q_k^- W^c\right)^T \end{bmatrix} \tag{35}$$

Like SR-UKF for sequential updates, we can have uncorrelated measurements and for this reason we can change the measurement equation to:

$$\widetilde{y}_k = \left(U_{R,k}\right)^{-1} h(x_k, u_k) + \widetilde{v}_k = \widetilde{h}(x_k, u_k) + \widetilde{v}_k \tag{36}$$

Where $E[\widetilde{v}_k] = D_{R,k}$, it is a diagonal matrix.

TABLE IV
PSEUDO-CODE FOR UD-UKF

- Initialization

$$\hat{x}_0 = E[x_0] \quad , \quad [U_{x_0}, D_{x_0}) = UDU\{E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T]\}$$

$$\hat{x}_0^a = E[x_0^a] = \begin{bmatrix} \hat{x}_0^T & \overline{w}_0^T & \overline{v}_0^T \end{bmatrix}^T$$

$$U_0^a = \begin{bmatrix} S_{x_k} & 0 & 0 \\ 0 & S_{Q_k} & 0 \\ 0 & 0 & U_{R,k} \end{bmatrix}, \sqrt{D_0^a} = \begin{bmatrix} \sqrt{D_{x_0}} & 0 & 0 \\ 0 & \sqrt{D_{Q,0}} & 0 \\ 0 & 0 & \sqrt{D_{R,0}} \end{bmatrix}, \quad S_0^a = U_0^a \sqrt{D_0^a}$$

- For $k = 1, \ldots, \infty$
- Set $t = k - 1$
- Calculate sigma-points:

$$\chi_t^a = \begin{bmatrix} \hat{x}_t^a & \hat{x}_t^a + \gamma S_{x_t}^a & \hat{x}_t^a - \gamma S_{x_t}^a \end{bmatrix}$$

- Time update equations:

$$\chi_{k|t}^x = f(\chi_t^x, u_t, \chi_t^w)$$

$$\hat{x}_k^- = \sum_{i=0}^{2L} w_i^m \chi_{i,k|t}^x$$

$$[U_k^{\sim}, D_k^{\sim}, Q_k^-] = QR\_UD\{\chi_{k|t}^x - \hat{x}_k^-\}$$

$$[U_k^{\approx}, D_k^{\approx}] = UDU\{D_k^{\sim} Q_k^- W^c Q_k^{-T} D_k^{\sim T}\}$$

$$U_k^- = U_k^{\sim} U_k^{\approx}, D_k^- = D_k^{\approx}$$

$$\widetilde{y}_k = U_{R,k} y_k$$

- Measurement-update equations:

Set $\hat{z}_0 = \hat{x}_k^-, U_{z,0} = U_k^-, D_{z,0} = D_k^-$

For $j = 1, \ldots, n_y$

$$\widetilde{Y}_{k|t,j} = \widetilde{h}_j(Z_{k|t}, u_t) + \chi_t^{\widetilde{v}}$$

$$\widetilde{y}_{k,j}^- = \sum_{i=0}^{2L} w_i^m \widetilde{Y}_{i,k|t,j}$$

$$P_{\widetilde{y}_k} = \left(\widetilde{Y}_{k|t,j} - y_{k,j}^-\right) W^c \left(\widetilde{Y}_{k|t,j} - y_{k,j}^-\right)^T$$

$$[\overline{U}_{k,j}, \overline{D}_{k,j}] = UDU\{G_{k,j}\}$$

$$U_{z,j} = U_k^{\sim} \overline{U}_{k,j}, D_{z,j} = \overline{D}_{k,j}$$

$$K_j = \left(\chi_{k|t}^x - \hat{x}_k^-\right) W^c \left(\widetilde{Y}_{k|t,j} - y_{k,j}^-\right)^T$$

$$\hat{z}_j = \hat{z}_{j-1} + K_j (\widetilde{y}_{k,j} - \widetilde{y}_{k,j}^-)$$

End for

$$\hat{x}_k = \hat{z}_{n_y}, U_k = U_{z,n_y}, D_k = D_{z,n_y}$$

VI. EXAMPLE APPLICATION

In this section we consider the problem of tracking a vehicle that enters the atmosphere at high altitude with a very high speed. The position of the body is to be tracked by radar which accurately measures range and bearing. This type of problem has been identified by some authors [13] as being practically stressful for filters and trackers because of the strong nonlinearities exhibited by the forces which act on the

vehicle. There are three types of forces which act on a reentry vehicle. The most dominant one being the aerodynamic drag, which is a function of the vehicle speed and has substantial nonlinear variation with altitude. The second type is gravity which accelerates the vehicle towards the center of earth. Finally, there exists a random type buffeting force. The combined effect of these forces generates the trajectory shown in Fig. 1. Initially the trajectory is almost ballistic but as density of the atmosphere increases, drag effects become important and the vehicle rapidly decelerates until its motion is almost vertical. The tracking problem is made more difficult by the fact that the drag properties of the vehicle might be only crudely known.

In summery we can formulate this problem in state space form [9], [13]:

$$\dot{x}_1 = x_3$$
$$\dot{x}_2 = x_4$$
$$\dot{x}_3 = Dx_3 + Gx_1 + w_1 \qquad (28)$$
$$\dot{x}_4 = Dx_4 + Gx_2 + w_2$$
$$\dot{x}_5 = w_3$$

Where x_1 and x_2 represent position in two dimension. x_3 and x_4 represent velocity in two dimensions, and x_5 is a parameter of aerodynamic properties. D Is the drag-related force term, G is the Gravity-related force term and w is process noise. Defining $R = \sqrt{x_1^2 + x_2^2}$ as distance from the center of the Earth and $V = \sqrt{x_3^2 + x_4^2}$ as the absolute vehicle speed, then the drag and gravitational terms will be:

$$D = -\beta \exp\left[\frac{[R_0 - R]}{H_0}\right]V \;,\; G = \frac{\mu}{r^3} \qquad (29)$$

$$\beta = \beta_0 \exp(x_5) \qquad (30)$$

for this example the parameter values are $\beta_0 = -0.59783, H_0 = 13.406, Gm_0 = 3.9860 \times 10^5 \frac{km^3}{\sec^2}$ and $R_0 = 6374 \; km$, which are reflective of typical environmental and vehicle characteristics. The parameterization of the ballistic coefficient, β_k, reflects the uncertainty in vehicle characteristics. β_0 is the ballistic coefficient of a typical vehicle and it is scaled bye $\exp(x_5)$ to ensure that its value is always positive. This is vital for filter stability.
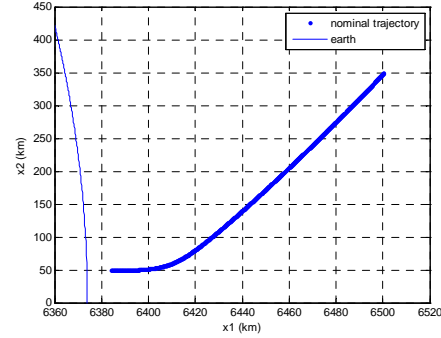

Fig. 1 Nominal trajectory

The motion of the vehicle is measured by radar that is located at (x_r, y_r). The radar is able to measure range, r and bearing, θ at a frequency of 20 Hz, where

$$r_r(k) = \sqrt{(x_1(k) - x_r)^2 + (x_2(k) - y_r)^2} + v_1(k)$$

$$\theta(k) = \tan^{-1}\left(\frac{(x_2(k) - y_r)}{(x_1(k) - x_r)}\right) + v_2(k)$$

$v_1(k), v_2(k)$ Are zero-mean uncorrelated noise processes with variances of $1 \, m$ and $17 \, mrd$, respectively [35].The high update rate and extreme accuracy of the sensor means that a large quantity of extremely high quality data is available for the filter. The bearing uncertainty is sufficiently small that the EKF is able to predict the sensor readings accurately with very little bias.

The true initial conditions for the vehicle are

$$x(0) = \begin{bmatrix} 6500.4 \\ 349.14 \\ -1.8093 \\ -6.7967 \\ 0.6932 \end{bmatrix} \quad, \quad P(0) = \begin{bmatrix} 10^{-6} & 0 & 0 & 0 \\ 0 & 10^{-6} & 0 & 0 \\ 0 & 0 & 10^{-6} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

In other words, the vehicle's ballistic coefficient is twice the nominal coefficient.

The vehicle is buffeted by random accelerations,

$$Q = \begin{bmatrix} 2.4064 \times 10^{-5} & 0 & 0 \\ 0 & 2.4064 \times 10^{-5} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The initial conditions assumed by the filter are:

$$x(0|0) = \begin{bmatrix} 6500.4 \\ 349.14 \\ -1.8093 \\ -6.7967 \\ 0 \end{bmatrix} \quad, \quad P(0) = \begin{bmatrix} 10^{-6} & 0 & 0 & 0 \\ 0 & 10^{-6} & 0 & 0 \\ 0 & 0 & 10^{-6} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The filter uses the nominal initial conditions and, in order to offset the uncertainty, the variance on this initial estimate is

taken as one.

Here we consider the three square root forms of UKF, namely the Square-Root Unscented Kalman filter (SR-UKF), Sequential update of Square-Root Unscented Kalman filter (SSR-UKF) and UD Unscented Kalman filter (UD-UKF) respectively. These filters have the same initial conditions. Three filters were implemented in discrete time and observations were taken at a frequency $20Hz$. For discrete estimation of x_4, a simple Euler method with $\Delta t = \dfrac{1}{20}\sec$ is utilized. Here estimation of $x(5)$ is hard, for this reason the three above mentioned algorithms are compared for 100 time execution of the Monte Carlo simulation. Fig. 2 shows the mean square estimation of $x(5)$ with time. For all filters we have:
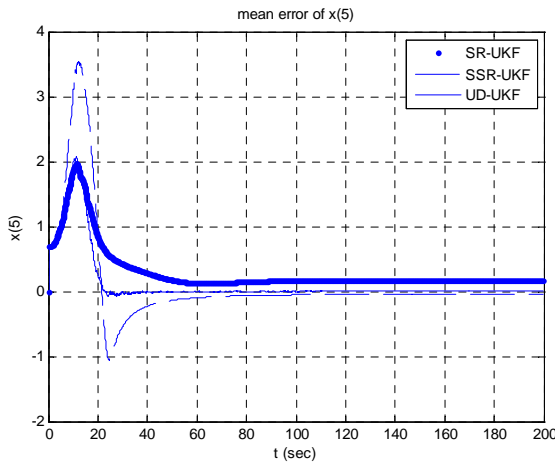
$$\kappa = 0 \ , \ \alpha = 0.55 \ , \ \beta = 2$$

Fig. 2 Mean error of x(5) in time

This figure shows that the UD-UKF is far superior once compared with the other filters. Note that, $\alpha = 0.55$ was used for simulation because we need to have a positive value for w_0^c.

Table V shows the mean square error for the three filters used:

TABLE V
MEAN SQUARE ERROR FOR THREE VERSIONS UKF

Mean square error for 100 Monte Carlo simulation	X1	X2	X3	X4	X4
SR-UKF	29.57	327.54	0.0018	0.097	0.1965
SSR-UKF	0.0453	0.1667	0.0011	0.015	0.2961
UD-UKF	0.0043	0.1017	5e-5	0.001	0.1697

VII. CONCLUSION

In this paper we have presented and developed new additions to the UKF. In real applications, the original UKF may fail due to loss of positive-definiteness property of the state covariance matrix. In this condition not only the filter will be unstable, but also the Cholesky factorization algorithm would not work. Van Der Merwe [22] has developed a square root formulation of UKF (SR-UKF) which propagates the mean and square root form of the covariance matrix, rather than the covariance matrix itself. His filter has good numerical properties compared with those of the original UKF. However, SR-UKF uses Cholesky downdate algorithm sequentially for the calculation of the posterior square root matrix which reduces the accuracy of the filter. In our newly developed sequential update of SR-UKF filter (SSR-UKF), a sequential updating on all of the measurements equations is performed that relieves the need for downdate calculations for a given posterior square root matrix. SSR-UKF uses cholupdate algorithm in time update as well as measurement update equations, while w_0^c is negative. In addition UD-UKF eliminates the need for cholupdate or downdate and uses UD factorization scheme. Simulation results in this paper show that UD-UKF has higher accuracy compared with SR-UKF and SSR-UKF.

APPENDIX 1: CHOLESKY UPDATE/DOWNDATE [27]

Consider $A = RR^T$ where $R = \text{chol}(A)$ the original Cholesky factorization of A, returns the upper triangular Cholesky factor of A. In order to calculate the Cholesky factor of $R' = chol(A + w_0 xx^T)$, where $(A + w_0 xx^T) = R'R'^T$, one can use rank 1 update of Cholesky factor of R. This is denoted by:

$$R' = cholupdate(R, x, w_0)$$

for downdate of R, we have $R' = cholupdate(R, x, -w_0)$.

APPENDIX 2: ORTHOGONAL-TRIANGULAR DECOMPOSITION [27]

The QR function performs the orthogonal-triangular decomposition of a matrix. This factorization is useful for both square and rectangular matrices. It expresses the matrix as the product of a real orthonormal or complex unitary matrix and an upper triangular matrix.

Consider x in a rectangular form with $x \in R^{n \times L}$, $L \geq n$. One can factorize this matrix in the form of:

$$x = \begin{bmatrix} S & 0 \end{bmatrix} Q' \text{ and } S \in R^{n \times n} \text{ and } Q' \in R^{L \times L}, \text{ where}$$

Q' is an orthonormal matrix . One can write:

$$P = xx^T = \begin{bmatrix} S & 0 \end{bmatrix} Q'Q'^T \begin{bmatrix} S & 0 \end{bmatrix}^T$$
$$= SQQ^T S^T$$

Here $Q \in R^{n \times L}$ is a part of Q', and every row of Q is a unit vector which is orthogonal to others.

APPENDIX 3: ORTHOGONAL-TRIANGULAR DECOMPOSITION OF A VECTOR

Suppose the r is a vector and $r \in R^{n \times 1}$. We can factorize this vector as:

$$r^T = S_r Q_r$$

Here $S_r = \|r\|$ is a scalar and $Q_r \in R^{1 \times n}$, $Q_r = \frac{1}{\|r\|} r^T$.

It is easy to factorize a vector.

APPENDIX 4: PSEUDO-CODE FOR QR_UD

This pseudo-code uses QR factorization to factorize a matrix $x \in R^{n \times L}, L \geq n$;

$$x = UDQ$$

$$[U, D, Q] = QR\_UD\{x\}$$

Here $U \in R^{n \times n}$ is an upper triangular matrix with ones on the diagonal, $D \in R^{n \times n}$ is a diagonal positive matrix and $Q \in R^{n \times L}$ is a part of orthogonal matrix. For this factorization, QR method is utilized. We have: $x = RQ$.

Here $R \in R^{n \times n}$ is an upper triangular matrix, which has non-zero diagonal elements. Also we can write: $R = UD$.

REFERENCES

[1] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte. A new approach for filtering nonlinear systems. In Proc. of the 1995 American Control Conference, pages 1628–1632, Seattle, Washington, June 1995.

[2] P. S. Maybeck, Stochastic Models, Estimation, and Control, P. S. Maybeck, Ed. New York: Academic, 1982, vol. 1 and 2.

[3] H.J. Kushner, "Dynamical equatins for optimum no-linear filtering", J. Differential Equations, vol. 3,pp. 179-190, 1967.

[4] A. H. Jazwinski. Stochastic Processes and Filtering Theory. Academic Press, New York, 1970.

[5] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to Nonlinear/Non-Gaussian Bayesian State Estimation", In IEE Proceedings on Radar and Signal Processing, volume 140, no. 2, pp. 107-113,1993.

[6] H. J. Kushner, "Approximations to optimal nonlinear filters", IEEE Trans. Automat. Contr., vol. AC-12, pp. 546-556, Oct. 1967.

[7] 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.

[8] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte," A new Approach for the Nonlinear Transformation of Means and Covariance in linear Filters", IEEE Transactions on Automatic Control, Vol. 5, No. 3, pp. 477-482, March 2000.

[9] S. J. Julier and J. K. Uhlmann, "A New Extension of the Kalman Filter to Nonlinear Systems," in Proc. of eroSense: The 11th Int. Symp. On Aerospace/Defense Sensing, Simulation and Controls, 1997.

[10] S. J. Julier and J. K. Uhlmann. "A general method for approximating nonlinear transformations of probability distributions. Technical report, Robotics Research Group, Department of Engineering Science, University of Oxford, 1994. (Internet publication:http://www.robots.ox.ac.uk/ fi siju/index.html).

[11] Rudolph van der Merwe and Eric A. Wan, "Efficient Derivative-Free Kalman Filters for Online Learning", ESANN'2001 proceedings - European Symposium on Artificial Neural Networks Bruges (Belgium), 25-27 April 2001, D-Facto public., ISBN 2-930307-01-3, pp. 205-210.

[12] Kazufumi Ito and Kaiqi Xiong, "Gaussian Filters for Nonlinear Filtering Problems," IEEE Transactions on Automatic Control, vol. 45, no. 5, pp. 910–927, may 2000.

[13] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," Proc. IEEE, vol. 92, no. 3, pp. 401–422, Mar. 2004.

[14] S. Haykin, "Adaptive filter Theory", Prentice Hall, Inc, fourth edition, 2002.

[15] E. A. Wan and R. v. d. Merwe, "The unscented Kalman filter," In Kalman Filtering and Neural Networks, S. Haykin, Ed. New York: Wiley, 2001, ch. 7, pp. 221–280.

[16] R. v. d. Merwe, "Sigma-point Kalman filters for probabilistic inference in dynamic state-space models," electrical and computer engineering Ph.D. dissertation, Oregon Health Sciences Univ., Portland, OR, 2004.

[17] G. Minkler, J. Minkler, "Theory and Application of Kalman Filtering", Magellan Book Company, 1993.

[18] Grewal, M.S., Andrews, A.P., "Kalman Filtering Theory and Practice using Matlab," John Wiley & Sons, INC., 2001.

[19] R. van der Merwe, E. Wan, and S. J. Julier. Sigma-Point Kalman Filters for Nonlinear Estimation and Sensor-Fusion: Publications to Integrated Navigation. In Proceedings of the AIAA Guidance, Navigation & Control Conference, Providence, RI, Aug 2004.

[20] R. van der Merwe. Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models. PhD thesis, OGI School of Science & Engineering at Oregon Health & Science University, Portland, OR, April 2004.

[21] Rudolph van der Merwe and Eric A. Wan, "Sigma-Point Kalman Filters for Integrated Navigation", in Proceedings of the 60th Annual Meeting of The Institute of Navigation (ION), Dayton, OH, Jun, 2004.

[22] R. van der Merwe and E. A. Wan, "The Square-Root Unscented Kalman Filter for State and Parameter- Estimation", in International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, Utah, Vol. 6, May, 2001, pp. 3461-3464.

[23] R. van der Merwe and E. A. Wan, "Efficient Derivative-Free Kalman Filters for Online Learning", in European Symposium on Artificial Neural Networks (ESANN), Bruges, Belgium, Apr, 2001.

[24] Eric A. Wan and Rudolph van der Merwe and Alex T. Nelson, "Dual Estimation and the Unscented Transformation", in Advances in Neural Information Processing Systems 12, pp. 666-672, MIT Press, Eds. S.A. Solla and T. K. Leen and K.-R. Muller, Nov, 2000.

[25] G. J. Bierman, "Factorization METHODS FOR Discrete Sequential Estimation", Academic, New York, 1977.

[26] C.L. Thornton, "Triangular Covariance Factorizations for Kalman Filtering", PHD. thesis, University of California at Los Angeles, School of Engineering, 1976.

[27] http://www.mathworld.com