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

Tine Lefebvre, Herman Bruyninckx, and Joris De Schutter

***Abstract*—The above paper generalizes the Kalman filter to nonlinear systems by transforming approximations of the probability distributions through the nonlinear process and measurement functions. This comment derives exactly the same estimator by linearizing the process and measurement functions by a statistical linear regression through some regression points (in contrast with the extended Kalman filter which uses an analytic linearization in one point). This insight allows: 1) to understand/predict the performance of the estimator for specific applications, and 2) to make adaptations to the estimator (i.e., the choice of the regression points and their weights) in those cases where the original formulation does not assure good results.**

***Index Terms*—Statistical linear regression, unscented Kalman filter.**

I. Introduction

The above paper[1] describes the so-called *unscented Kalman filter* (UKF) as a new minimum mean squared error (MMSE) state estimator for a nonlinear system

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

$$z(k+1) = \boldsymbol{h}[\boldsymbol{x}(k+1),\, \boldsymbol{u}(k+1),\, k+1] + \boldsymbol{w}(k+1) \tag{2}$$

where $\boldsymbol{x}(k)$ is the state of the system at time step k, $\boldsymbol{u}(k)$ is the input vector, $z(k)$ is the observation vector, and $\boldsymbol{w}(k)$ is additive measurement noise. $\boldsymbol{f}[.]$ and $\boldsymbol{h}[.]$ are the nonlinear system and measurement functions. The UKF formulas of the paper[1] assume that the state vector $\boldsymbol{x}(k)$ is augmented with the process noise vector $\boldsymbol{v}(k)$ (see Remark 4 of the aforementioned paper[1]). This explains why $\boldsymbol{v}(k)$ is not explicitly mentioned in (1). It is also assumed that the process noise $\boldsymbol{v}(k)$ and the measurement noise $\boldsymbol{w}(k)$ are zero mean and

$$E\left[\boldsymbol{v}(k)\boldsymbol{v}^T(j)\right] = \delta_{kj}\boldsymbol{Q}(k) \tag{3}$$

$$E\left[\boldsymbol{w}(k)\boldsymbol{w}^T(j)\right] = \delta_{kj}\boldsymbol{R}(k), \qquad \forall k,\, j \tag{4}$$

$$E\left[\boldsymbol{v}(k)\boldsymbol{w}^T(j)\right] = \boldsymbol{0}. \tag{5}$$

The Kalman filter update equations compute the n-dimensional state estimate $\hat{\boldsymbol{x}}(k|j)$ at time step k, given all observations up to and including time step j, and the covariance matrix $\boldsymbol{P}(k|j)$ of this estimate

$$\hat{\boldsymbol{x}}(k+1|k+1) = \hat{\boldsymbol{x}}(k+1|k) + \boldsymbol{W}(k+1)\boldsymbol{\nu}(k+1) \tag{6}$$

$$\boldsymbol{P}(k+1|k+1) = \boldsymbol{P}(k+1|k) - \boldsymbol{W}(k+1) \cdot \boldsymbol{P}_{\nu\nu}(k+1|k)\boldsymbol{W}^T(k+1) \tag{7}$$

Manuscript received September 19, 2001. Recommended by Associate Editor Q. Zhang. This work was supported by the Belgian Programme on Inter-University Attraction Poles initiated by the Belgian State—Prime Minister's Office—Science Policy Programme (IUAP), and by Katholieke Universiteit Leuven's Concerted Research Action GOA/99/04. T. Lefebvre and H. Bruyninckx are, respectively, Doctoral and Postdoctoral Fellows of the Fund for Scientific Research—Flander (F.W.O.), Belgium.

The authors are with the Department of Mechanical Engineering, Katholieke Universiteit Leuven, B-3001 Leuven, Belgium (e-mail: Tine.Lefebvre@mech.kuleuven.ac.be).

Publisher Item Identifier 10.1109/TAC.2002.800742.

[1]S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, *IEEE Trans. Automat. Contr.*, vol. 45, pp. 477–482, Mar. 2000.

$$\boldsymbol{\nu}(k+1) = z(k+1) - \hat{z}(k+1|k); \tag{8}$$

$$\boldsymbol{W}(k+1) = \boldsymbol{P}_{xz}(k+1|k)\boldsymbol{P}_{\nu\nu}^{-1}(k+1|k) \tag{9}$$

$$\boldsymbol{P}_{\nu\nu}(k+1|k) = \boldsymbol{R}(k+1) + \boldsymbol{P}_{zz}(k+1|k). \tag{10}$$

Different Kalman filter variants for nonlinear systems propose other ways of computing $\hat{\boldsymbol{x}}(k+1|k)$, $\hat{z}(k+1|k)$, their covariance matrices $\boldsymbol{P}(k+1|k)$ and $\boldsymbol{P}_{zz}(k+1|k)$ and their cross covariance matrix $\boldsymbol{P}_{xz}(k+1|k)$.

The UKF formalism chooses $2n+1$ regression points $\boldsymbol{\mathcal{X}}_i$ in state space with weights W_i $(i = 1, \ldots, n)$

$$\boldsymbol{\mathcal{X}}_0(k|j) = \hat{\boldsymbol{x}}(k|j), \qquad W_0 = 2\kappa/\{2(n+\kappa)\} \tag{11}$$

$$\boldsymbol{\mathcal{X}}_i(k|j) = \hat{\boldsymbol{x}}(k|j) + \left(\sqrt{(n+\kappa)\boldsymbol{P}(k|j)}\right)_i \qquad W_i = 1/\{2(n+\kappa)\} \tag{12}$$

$$\boldsymbol{\mathcal{X}}_{i+n}(k|j) = \hat{\boldsymbol{x}}(k|j) - \left(\sqrt{(n+\kappa)\boldsymbol{P}(k|j)}\right)_i \qquad W_{i+n} = 1/\{2(n+\kappa)\} \tag{13}$$

where $(\sqrt{(n+\kappa)\boldsymbol{P}(k|j)})_i$ is the ith row or column of $\sqrt{(n+\kappa)\boldsymbol{P}(k|j)}$. κ is a degree of freedom in the choice of the regression points $\boldsymbol{\mathcal{X}}_i$ [1], [2], [3]. This sample set is chosen to have the same mean and covariance as the distribution of $\boldsymbol{x}(k)$

$$\hat{\boldsymbol{x}}(k|j) = \sum_{i=0}^{2n} W_i \boldsymbol{\mathcal{X}}_i(k|j) \tag{14}$$

$$\boldsymbol{P}(k|j) = \sum_{i=0}^{2n} W_i \{\boldsymbol{\mathcal{X}}_i(k|j) - \hat{\boldsymbol{x}}(k|j)\} \cdot \{\boldsymbol{\mathcal{X}}_i(k|j) - \hat{\boldsymbol{x}}(k|j)\}^T. \tag{15}$$

Starting from the regression points $\boldsymbol{\mathcal{X}}_i(k|k)$, obtained by (11)–(13), the UKF-specific process update equations compute the sample mean and covariance of the regression points passed through the nonlinear process function

$$\boldsymbol{\mathcal{X}}_i(k+1|k) = \boldsymbol{f}[\boldsymbol{\mathcal{X}}_i(k|k),\, \boldsymbol{u}(k),\, k] \tag{16}$$

$$\hat{\boldsymbol{x}}(k+1|k) = \sum_{i=0}^{2n} W_i \boldsymbol{\mathcal{X}}_i(k+1|k) \tag{17}$$

$$\boldsymbol{P}(k+1|k) = \sum_{i=0}^{2n} W_i \{\boldsymbol{\mathcal{X}}_i(k+1|k) - \hat{\boldsymbol{x}}(k+1|k)\} \cdot \{\boldsymbol{\mathcal{X}}_i(k+1|k) - \hat{\boldsymbol{x}}(k+1|k)\}^T. \tag{18}$$

Starting from the regression points $\boldsymbol{\mathcal{X}}_i(k+1|k)$, obtained by (11)–(13), the UKF-specific measurement update equations compute the sample mean and covariance of the regression points passed through the nonlinear measurement function[2]:

$$\boldsymbol{\mathcal{Z}}_i(k+1|k) = \boldsymbol{h}[\boldsymbol{\mathcal{X}}_i(k+1|k),\, \boldsymbol{u}(k+1),\, k+1] \tag{19}$$

$$\hat{z}(k+1|k) = \sum_{i=0}^{2n} W_i \boldsymbol{\mathcal{Z}}_i(k+1|k) \tag{20}$$

$$\boldsymbol{P}_{zz}(k+1|k) = \sum_{i=0}^{2n} W_i \{\boldsymbol{\mathcal{Z}}_i(k+1|k) - \hat{z}(k+1|k)\} \cdot \{\boldsymbol{\mathcal{Z}}_i(k+1|k) - \hat{z}(k+1|k)\}^T \tag{21}$$

$$\boldsymbol{P}_{xz}(k+1|k) = \sum_{l=0}^{2n} W_i \{\boldsymbol{\mathcal{X}}_i(k+1|k) - \hat{\boldsymbol{x}}(k+1|k)\} \cdot \{\boldsymbol{\mathcal{Z}}_i(k+1|k) - \hat{z}(k+1|k)\}^T. \tag{22}$$

[2]The paper[1] does not write down these formulas explicitly, but they can be found in the companion publications [1]–[3].

The motivation behind the aforementioned paper[1] is to avoid the linearization of the process and measurement functions $\boldsymbol{f}$ and $\boldsymbol{h}$, however, (16)–(22) are the same as for an estimator which *linearizes* these functions by statistical linear regression and which we will refer to as the linear regression Kalman filter (LRKF). The LRKF has the following properties:

1) it linearizes the process and measurement functions by statistical linear regression of the functions through some regression points in state space;
2) it defines the uncertainty due to linearization errors on the linearized process or measurement function as the covariance matrix of the deviations between the function values of the nonlinear and the linearized function in the regression points.

II. Equivalence of the UKF and the LRKF

Section II-A describes the linear regression formulas. Sections II-B and II-C define the LRKF and show that the UKF is a LRKF for both the process and the measurement update.

A. Statistical Linear Regression of a Nonlinear Function

Consider a nonlinear function $\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x})$ evaluated in r points ($\boldsymbol{\mathcal{X}}_i$, $\boldsymbol{\mathcal{Y}}_i$) where $\boldsymbol{\mathcal{Y}}_i = \boldsymbol{g}(\boldsymbol{\mathcal{X}}_i)$. Define

$$\overline{\boldsymbol{x}} = \frac{1}{r}\sum_{i=1}^{r}\boldsymbol{\mathcal{X}}_i \qquad \overline{\boldsymbol{y}} = \frac{1}{r}\sum_{i=1}^{r}\boldsymbol{\mathcal{Y}}_i \tag{23}$$

$$\boldsymbol{P}_{xx} = \frac{1}{r}\sum_{i=1}^{r}(\boldsymbol{\mathcal{X}}_i - \overline{\boldsymbol{x}})(\boldsymbol{\mathcal{X}}_i - \overline{\boldsymbol{x}})^T \tag{24}$$

$$\boldsymbol{P}_{xy} = \frac{1}{r}\sum_{i=1}^{r}(\boldsymbol{\mathcal{X}}_i - \overline{\boldsymbol{x}})(\boldsymbol{\mathcal{Y}}_i - \overline{\boldsymbol{y}})^T \tag{25}$$

$$\boldsymbol{P}_{yy} = \frac{1}{r}\sum_{i=1}^{r}(\boldsymbol{\mathcal{Y}}_i - \overline{\boldsymbol{y}})(\boldsymbol{\mathcal{Y}}_i - \overline{\boldsymbol{y}})^T. \tag{26}$$

The *y-regression* is the *linear regression* $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$ that minimizes the sum of the squared errors

$$\min_{\boldsymbol{A},\boldsymbol{b}} \sum_{i=1}^{r}\left\{\boldsymbol{e}_i^T\boldsymbol{e}_i\right\} \tag{27}$$

with $\boldsymbol{e}_i = \boldsymbol{\mathcal{Y}}_i - (\boldsymbol{A}\boldsymbol{\mathcal{X}}_i + \boldsymbol{b})$, i.e., the deviations between the function values of the nonlinear and linearized functions in the regression points.

The solution to (27) is [4]

$$\boldsymbol{A} = \boldsymbol{P}_{xy}^T\,\boldsymbol{P}_{xx}^{-1} \qquad \boldsymbol{b} = \overline{\boldsymbol{y}} - \boldsymbol{A}\overline{\boldsymbol{x}}. \tag{28}$$

The covariance matrix of the deviations $\boldsymbol{e}_i$ is

$$\begin{aligned}\boldsymbol{P}_{ee} &= \frac{1}{r}\sum_{i=1}^{r}\boldsymbol{e}_i\boldsymbol{e}_i^T \\ &= \frac{1}{r}\sum_{i=1}^{r}\{(\boldsymbol{\mathcal{Y}}_i - \overline{\boldsymbol{y}}) - \boldsymbol{A}(\boldsymbol{\mathcal{X}}_i - \overline{\boldsymbol{x}})\} \\ &\quad\cdot\{(\boldsymbol{\mathcal{Y}}_i - \overline{\boldsymbol{y}}) - \boldsymbol{A}(\boldsymbol{\mathcal{X}}_i - \overline{\boldsymbol{x}})\}^T \\ &= \boldsymbol{P}_{yy} - \boldsymbol{A}\,\boldsymbol{P}_{xy} - \boldsymbol{P}_{yx}\,\boldsymbol{A}^T + \boldsymbol{A}\,\boldsymbol{P}_{xx}\,\boldsymbol{A}^T \\ &= \boldsymbol{P}_{yy} - \boldsymbol{A}\,\boldsymbol{P}_{xx}\,\boldsymbol{A}^T.\end{aligned} \tag{29}$$

B. The UKF Is a LRKF: Process Update

The LRKF evaluates the nonlinear process function in r points $\boldsymbol{\mathcal{X}}_i(k|k)$. These points have mean $\hat{\boldsymbol{x}}(k|k)$ and covariance $\boldsymbol{P}(k|k)$. The updated values of the regression points are $\boldsymbol{\mathcal{X}}_i(k+1|k) = \boldsymbol{f}[\boldsymbol{\mathcal{X}}_i(k|k),\ \boldsymbol{u}(k),\ k]$. The LRKF algorithm uses a linearized process function obtained by statistical linear regression through $(\boldsymbol{\mathcal{X}}_i(k|k),\ \boldsymbol{\mathcal{X}}_i(k+1|k))$, $i = 1, \ldots, r$

$$\boldsymbol{x}(k+1) = \boldsymbol{F}_d(k)\boldsymbol{x}(k) + \boldsymbol{F}_c(k) + \boldsymbol{v}^*(k). \tag{30}$$

The analogy with Section II-A becomes clear by replacing the symbols according to Table I.

TABLE I
Equivalence of the Symbols in Sections II-A, II-B, and II-C

Section II-A	Section II-B	Section II-C
$\boldsymbol{x}$	$\boldsymbol{x}(k)$	$\boldsymbol{x}(k+1)$
$\boldsymbol{y}$	$\boldsymbol{x}(k+1)$	$\boldsymbol{z}(k+1)$
$\boldsymbol{g}$	$\boldsymbol{f}$	$\boldsymbol{h}$
$\boldsymbol{A}$ (Eq. (28))	$\boldsymbol{F}_d(k)$	$\boldsymbol{H}_d(k+1)$
$\boldsymbol{b}$ (Eq. (28))	$\boldsymbol{F}_c(k)$	$\boldsymbol{H}_c(k+1)$
$\boldsymbol{\mathcal{X}}_i$	$\boldsymbol{\mathcal{X}}_i(k\|k)$	$\boldsymbol{\mathcal{X}}_i(k+1\|k)$
$\boldsymbol{\mathcal{Y}}_i$	$\boldsymbol{\mathcal{X}}_i(k+1\|k)$	$\boldsymbol{\mathcal{Z}}_i(k+1\|k)$

$\boldsymbol{v}^*(k)$ is the process uncertainty that accounts for the linearization errors. Its covariance matrix $\boldsymbol{Q}^*(k)$ is defined as the covariance matrix of the deviations between the function values of the nonlinear and the linearized function in the regression points, i.e., $\boldsymbol{Q}^*(k) = \boldsymbol{P}_{ee}$ of (29).

The LRKF-specific process update equations then follow from the well-known linear Kalman filter equations

$$\begin{aligned}\hat{\boldsymbol{x}}(k+1|k) &= \boldsymbol{F}_d(k)\hat{\boldsymbol{x}}(k|k) + \boldsymbol{F}_c(k) \\ &= \frac{1}{r}\sum_{i=1}^{r}\boldsymbol{\mathcal{X}}_i(k+1|k)\end{aligned} \tag{31}$$

$$\begin{aligned}\boldsymbol{P}(k+1|k) &= \boldsymbol{F}_d(k)\boldsymbol{P}(k|k)\boldsymbol{F}_d(k)^T + \boldsymbol{Q}^*(k) \\ &= \frac{1}{r}\sum_{i=1}^{r}\{\boldsymbol{\mathcal{X}}_i(k+1|k) - \hat{\boldsymbol{x}}(k+1|k)\} \\ &\quad\cdot\{\boldsymbol{\mathcal{X}}_i(k+1|k) - \hat{\boldsymbol{x}}(k+1|k)\}^T.\end{aligned} \tag{32}$$

These update equations correspond to the UKF equations (16)–(18), knowing that the UKF defines $r = 2(n+\kappa)$ regression points: n points $\boldsymbol{\mathcal{X}}_i(k|k)$, n points $\boldsymbol{\mathcal{X}}_{i+n}(k|k)$, and 2κ points $\boldsymbol{\mathcal{X}}_o(k|k)$, (11)–(13).

C. The UKF Is a LRKF: Measurement Update

The LRKF evaluates the nonlinear measurement function in r regression points $\boldsymbol{\mathcal{X}}_i(k+1|k)$. The points are chosen such that their mean and their covariance matrix equal $\hat{\boldsymbol{x}}(k+1|k)$ and $\boldsymbol{P}(k+1|k)$. The function values of the regression points through the measurement function are $\boldsymbol{\mathcal{Z}}_i(k+1|k) = \boldsymbol{h}[\boldsymbol{\mathcal{X}}_i(k+1|k),\ \boldsymbol{u}(k+1),\ k+1]$. The LRKF algorithm uses a linearized measurement function obtained by statistical linear regression through $(\boldsymbol{\mathcal{X}}_i(k+1|k),\ \boldsymbol{\mathcal{Z}}_i(k+1|k))$, $i = 1, \ldots, r$

$$\begin{aligned}\boldsymbol{z}(k+1) &= \boldsymbol{H}_d(k+1)\boldsymbol{x}(k+1) + \boldsymbol{H}_c(k+1) \\ &\quad + \boldsymbol{w}(k+1) + \boldsymbol{w}^*(k+1).\end{aligned} \tag{33}$$

The analogy with Section II-A becomes clear by replacing the symbols according to Table I.

$\boldsymbol{w}(k+1)$ is the measurement noise on the nonlinear measurement function (2), $\boldsymbol{w}^*(k+1)$ is the extra measurement uncertainty on the linearized function due to linearization errors. The covariance matrix of the latter is $\boldsymbol{R}^*(k+1)$ and is defined as the covariance matrix of the deviations between the function values of the nonlinear and the linearized function in the regression points, i.e., $\boldsymbol{R}^*(k+1) = \boldsymbol{P}_{ee}$ of (29).

The LRKF-specific measurement equations then follow from the well-known linear Kalman filter equations

$$\begin{aligned}\hat{z}(k+1|k) &= \boldsymbol{H}_d(k+1)\hat{\boldsymbol{x}}(k+1|k) + \boldsymbol{H}_c(k+1) \\ &= \frac{1}{r}\sum_{i=1}^{r}\boldsymbol{\mathcal{Z}}_i(k+1|k)\end{aligned} \tag{34}$$

$$\begin{aligned}\boldsymbol{P}_{\nu\nu}(k+1|k) &= \boldsymbol{R}(k+1) + \boldsymbol{R}^*(k+1) \\ &\quad + \boldsymbol{H}_d(k+1)\boldsymbol{P}(k+1|k)\boldsymbol{H}_d^T(k+1) \\ &= \boldsymbol{R}(k+1) \\ &\quad + \frac{1}{r}\sum_{i=1}^{r}\{\boldsymbol{\mathcal{Z}}_i(k+1|k) - \hat{\boldsymbol{z}}(k+1|k)\} \\ &\quad \cdot \{\boldsymbol{\mathcal{Z}}_i(k+1|k) - \hat{\boldsymbol{z}}(k+1|k)\}^T \qquad (35) \\ \boldsymbol{P}_{xz}(k+1|k) &= \boldsymbol{P}(k+1|k)\boldsymbol{H}_d^T(k+1); \\ &= \frac{1}{r}\sum_{i=1}^{r}\{\boldsymbol{\mathcal{X}}_i(k+1|k) - \hat{\boldsymbol{x}}(k+1|k)\} \\ &\quad \cdot \{\boldsymbol{\mathcal{Z}}_i(k+1|k) - \hat{\boldsymbol{z}}(k+1|k)\}^T. \qquad (36)\end{aligned}$$

These update equations are identical to the UKF process update equations (19)–(22), knowing that the UKF defines $r = 2(n+\kappa)$ regression points: n points $\boldsymbol{\mathcal{X}}_i(k+1|k)$, n points $\boldsymbol{\mathcal{X}}_{i+n}(k+1|k)$, and 2κ points $\boldsymbol{\mathcal{X}}_o(k+1|k)$ (11)–(13).

III. Adaptations to the Original UKF

Several adaptations to the original UKF framework have been proposed. These adaptations use other criteria to choose the regression points and/or their weights; sometimes the calculated covariance matrix is increased artificially.

- The weights can be chosen as real numbers, the UKF then performs a weighted linear regression.
- Sometimes, the UKF is used with a $\kappa < 0$ (resulting in negative weights) (see the aforementioned paper[1] and [1]). In this case, the calculated covariance matrices can be nonpositive, semidefinite. To overcome this problem, the covariances are artificially increased

$$\begin{aligned}\boldsymbol{P}^{\mathrm{mod}}(k+1|k) &= \boldsymbol{P}(k+1|k) \\ &\quad + \{\boldsymbol{\mathcal{X}}_0(k+1|k) - \hat{\boldsymbol{x}}(k+1|k)\} \\ &\quad \cdot \{\boldsymbol{\mathcal{X}}_0(k+1|k) - \hat{\boldsymbol{x}}(k+1|k)\}^T \qquad (37) \\ \boldsymbol{P}_{zz}^{\mathrm{mod}}(k+1|k) &= \boldsymbol{P}_{zz}(k+1|k) \\ &\quad + \{\boldsymbol{\mathcal{Z}}_0(k+1|k) - \hat{\boldsymbol{z}}(k+1|k)\} \\ &\quad \cdot \{\boldsymbol{\mathcal{Z}}_0(k+1|k) - \hat{\boldsymbol{z}}(k+1|k)\}^T. \qquad (38)\end{aligned}$$

- The scaled unscented transformation [5] introduces one more degree of freedom in the choice of the regression points and their weights (parameter α). In this case, the calculated covariances lie between those of the original formulation and the previously described modified form. The covariances can be increased even more by introducing another parameter (β).
- The reduced sigma point filters [2], [6] minimize the number of regression points to $n+1$ (the so-called simplex sigma points) for an n-dimensional state space. This means that the linear regression is exact, i.e., the linearized function is a hyperplane through the regression points. Hence, $\boldsymbol{Q}^*$ and $\boldsymbol{R}^*$ are zero: the linearization errors are not taken into account. In this case, the calculated covariances are too small and need to be increased artificially.
- Also some other filters are linear regression Kalman filters, e.g., the central difference filter [7] and the first-order divided difference filter [8], [9] which choose $2n$ regression points.

IV. Conclusion

This comment has shown that the UKF is a special case of the LRKF, i.e., 1) it linearizes the process and measurement functions by statistical linear regression of the functions through some regression points; and 2) it represents the extra uncertainty on a linearized function due to linearization errors by the covariance of the deviations between the nonlinear and the linearized function in the regression points. Looking at the UKF in this way: 1) allows a better understanding of the performance of the estimator for specific applications; and 2) allows to understand/develop adaptations to the estimator which guarantee better performance in applications where the original estimator does not assure good results (e.g., when dealing with discontinuous functions).

References

[1] S. J. Julier and J. K. Uhlmann. (1996) A general method for approximating nonlinear transformations of probability distributions. [Online]. Available: http://www.robots.ox.ac.uk/~siju/work/work.html
[2] J. Uhlmann, "Dynamic map building and localization: New theoretical foundations," Ph.D. dissertation, Univ. Oxford, Oxford, U.K., 1995.
[3] S. Julier and J. K. Uhlmann, "Data fusion in nonlinear systems," in *Handbook of Multisensor Data Fusion*, D. L. Hall and J. Llinas, Eds. Boca Raton, FL: CRC Press, 2001.
[4] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. New York: Wiley, 1984.
[5] S. Julier. (1999, Jan.) The scaled unscented transformation. [Online]. Available: http://citeseer.nj.nec.com/julier99scaled.html
[6] ——, (1998, Dec.) Reduced sigma point filters for the propagation of means and covariances through nonlinear transformations. [Online]. Available: http://citeseer.nj.nec.com/julier98reduced.html
[7] T. S. Schei, "A finite-difference method for linearization in nonlinear estimation algorithms," *Automatica*, vol. 33, no. 11, pp. 2053–2058, Nov. 1997.
[8] O. Ravn, M. Nørgaard, and N. K. Poulsen, "New developments in state estimations for nonlinear systems," *Automatica*, vol. 36, no. 11, pp. 1627–1638, Nov. 2000.
[9] ——, "Advances in derivative-free state estimation for nonlinear systems," Technical Univ. Denmark, Denmark, Tech. Rep. IMM-REP-1998-15 (revised ed.), Apr. 2000.

Authors' Reply

Simon Julier and Jeffrey Uhlmann

In [1], Lefebvre *et al.* demonstrate the validity of the so-called unscented Kalman filter (UKF) approach[1] from a least-squares regression perspective. We use the definition that the Kalman filter is the minimum least squares update algorithm. From first principles and using notation from the aforementioned paper,[1] it can be proved that this is

$$\begin{aligned}\hat{\mathbf{x}}(k+1|k+1) &= \hat{\mathbf{x}}(k+1|k) + \mathbf{W}(k+1)\nu(k+1) \\ \mathbf{P}(k+1|k+1) &= \mathbf{P}(k+1|k) \\ &\quad - \mathbf{W}(k+1)\mathbf{P}_{\nu\nu}(k+1|k)\mathbf{W}^T(k+1) \\ \nu(k+1) &= \mathbf{z}(k+1) - \hat{\mathbf{z}}(k+1|k) \\ \mathbf{W}(k+1) &= \mathbf{P}_{x\nu}(k+1|k)\mathbf{P}_{\nu\nu}^{-1}(k+1|k).\end{aligned}$$

Manuscript received May 7, 2002.

S. Julier is with IDAK Industries, Jefferson City, MO 65109 USA (e-mail: sjulier@idak.com).

J. Uhlmann was with Robotics Research Group, Oxford University, Oxford 0X1 3PJ, U.K. He is now with the Department of Computer Engineering and Computer Science, University of Missouri, Columbia, MO 65211-2060 USA (e-mail: uhlmannj@missouri.edu).

Publisher Item Identifier 10.1109/TAC.2002.800741.