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

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.

0018-9286/02$17.00 © 2002 IEEE

The only difference between the extended Kalman filter (EKF), the UKF, and linear regression Kalman filter (LRKF) is how the prediction quantities are calculated. (We note in passing that other authors have done the same using quadrature and other methods [2]). They refer to their formulation of the estimation problem as being an LRKF. This analysis provides extremely useful insights into the properties of the UKF. For example, the matrix $\mathbf{A}$ in (28) of the paper[1] can be regarded as a meaningful approximation to the "Jacobian" of the observation. Such "Jacobians" are valuable in many applications including the fusion of measurements with known time delays [3]. Lefebvre *et al.* conclude that the UKF is a special case of a linear regression.

However, we believe that their conclusion is an unnecessarily narrow interpretation of this result. According to their analysis, *any* prediction algorithm which can be expressed in terms of (23) to (26) of the paper[1] is a type of linear regression. This not only includes the UKF, but many other Monte Carlo or particle-based filters, including the SIR filter described in [4]. Rather, we believe that a fairer interpretation is that any particle based filtering algorithm (including the UKF) implicitly performs a linear regression. Furthermore, unlike the LRKF, the UKF can be used to propagate higher order information about the distribution, if it is known, by extending the sigma point set. In [5] (which was cited in the aforementioned paper[1]), we demonstrated that the approach could be used to maintain and propagate information about the third-order moments.

We also note that the statement "This means that the linear regression is exact, i.e., the linearized function is hyperplane through the sampling points. Hence, $\mathbf{Q}^*$ and $\mathbf{R}^*$ are zero: the linearization errors are not taken into account" is *not* an inherent property of the reduced sigma point algorithm. It is a special case of the scaling algorithm in the limit as the scaling parameter tends to 0.

REFERENCES

[1] T. Lefebvre, H. Bruyninckx, and J. De Schutter, "Comment on'A new method for the nonlinear transformation of means and covariances in filters and estimators'," *IEEE Trans. Automat. Contr.*, vol. 47, pp. 1406–1408, Aug. 2002.

[2] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 910–927, May 2000.

[3] T. D. Larsen, N. A. Andersen, O. Ravn, and N. Kjolstad Poulson, "Incorporation of time delayed measurements in a discrete-time Kalman filter," in *Proc. 38th Conf. Decision Control*, vol. 4, 1998, pp. 3972–3977.

[4] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *Proc. Inst. Elect. Eng.*, pt. F, vol. 140, no. 2, pp. 107–113, Apr. 1993.

[5] S. J. Julier, "A skewed approach to filtering," in *Proc. AeroSense: 12th Int. Symp. Aerospace/Defense Sensing, Simulation, Controls*, vol. 3373, Orlando, FL, Apr. 1998, SPIE, Signal, and Data Processing of Small Targets, pp. 54–65.