OPTIMAL FILTERING OF NONLINEAR SYSTEMS BASED ON PSEUDO GAUSSIAN DENSITIES

Uwe D. Hanebeck

Institute of Computer Design and Fault Tolerance
Universität Karlsruhe
76128 Karlsruhe, Germany
Uwe.Hanebeck@ieee.org

Abstract: We consider the problem of estimating the state of a discrete–time dynamic system comprising a linear system equation and a nonlinear measurement equation based on measurements corrupted by non–Gaussian noise. The problem is solved by recursively calculating the complete posterior density of the state given the measurements. For representing the resulting non–Gaussian posterior, a new exponential type density, the so called pseudo Gaussian density, is introduced. By converting the original nonlinear system to an equivalent linear representation in a higher–dimensional space, the parameters of the pseudo Gaussian posterior are obtained by means of a linear estimator operating in the higher–dimensional space. The resulting filtering algorithms are easy to implement and always guarantee valid posterior densities.

Keywords: Estimators, Filtering Theory, Mathematical systems theory, Non–Gaussian processes, Nonlinear systems, Optimal filtering, Recursive estimation, Stochastic systems

1. INTRODUCTION

Filtering consists of estimating parameters of one process, the system state sequence, given uncertain information from another related process, the measurement sequence. When the measurements are related nonlinearly to the system state, this estimation problem is in general difficult to solve. Usually, a linearization is performed to permit application of filtering methods derived for linear systems (Anderson and Moore, 1979). Of course, this only works for certain type of nonlinearities. In addition, the presence of non–Gaussian measurement noise further limits the applicability of linear methods.

More advanced methods for providing state estimates in the nonlinear case have been developed by keeping nonlinear terms in a Taylor series expansion of the nonlinearity, see (Bohn and Unbehauen, 2000) for an elegant derivation. However, in this paper the focus is on calculating the *complete* posterior density of the unknown system state given all the measurements. A parametric closed–form density description is desired, which is defined by a finite number of parameters. In addition, the density representation should allow for recursive application and should not suffer from a permanently growing number of description parameters with an increasing number of available measurements.

A grid representation of densities for numerical nonlinear filtering based on quantization of the state space has been introduced in (Bucy and Senne, 1971), but has proven to be useful only for a limited state vector dimension (Bergman *et al.*, 1999). Monte Carlo techniques (Doucet *et al.*, 2000; Liu and Chen, 1998) use stochastic samples to represent density functions in order to numerically solve the filtering problem.

Closed–form representations of densities include the Edgeworth expansion, i.e., a Gaussian density times a sum of Hermite polynomials, which has been proposed in (Sorenson and Stubberud, 1968). A method for updating this type of density numerically is described in (Challa *et al.*, 2000). The approach has the disadvantage that truncated Edgeworth expansions are not themselves valid density functions and may give negative values (Jazwinski, 1970). A Gaussian mixture representation has been proposed in (Alspach and Sorenson, 1972), which always provides valid density functions. However, each term is individually updated based on linearization, which results in a bank of parallel extended Kalman filters.

The simplest form of the measurement update seems to be obtained when using exponential type densities (Kulhavý, 1992). In addition, these densities are always positive. However, depending on the exponent function, e.g. polynomials, numerical inaccuracies during the update recursion may lead to densities that are not integrable, i.e., the integral over the density does not give a finite value.

In this paper, a new type of exponential density, the so called pseudo Gaussian density, is proposed. It is defined by a standard Gaussian function in a hyperspace S^* related to the original state space S via a nonlinear transformation. Because of its special structure, pseudo Gaussians are always valid density functions even in the presence of numerical inaccuracies. In addition, it will be shown that under certain assumptions, this type of density can be updated by means of a linear filter operating in the hyperspace S^*.

The nonlinear filtering problem is formulated in Section 2. The concept of pseudo Gaussian densities is explained in detail in Section 3. Section 4 provides a generic conversion of a nonlinear system to an equivalent system, which is linear in a higher–dimensional space. The corresponding filtering algorithm is given in Section 5. An example in Section 6 illustrates the proposed approach.

2. PROBLEM FORMULATION

We consider estimating the state of a linear dynamic system, which may evolves in discrete time according to

$$\underline{x}_{k+1} = \mathbf{A}_k \, \underline{x}_k + \underline{\hat{u}}_k \ , \qquad (1)$$

where $\underline{\hat{u}}_k$ is a known input sequence. The system state $\underline{x}_k$ is not directly observable, but will instead be deduced from measurements of the system output. Measurements are assumed to be taken sequentially at discrete time steps $k = 1, 2, \ldots$ and are corrupted by white non–Gaussian noise.

An M–dimensional measurement $\underline{\hat{y}}_k$ [1] at time step k is related to the N–dimensional system state $\underline{x}_k$ via the *nonlinear time–variant* measurement equation

$$\underline{\hat{y}}_k = \underline{h}_k(\underline{x}_k) + \underline{v}_k \qquad (2)$$

and is corrupted by *additive* white noise $\underline{v}_k$ from a possibly non–Gaussian noise density $p_k^v(\underline{v}_k)$.

Instead of providing point estimates of the unknown state $\underline{x}_k$, an estimator is used to construct the complete conditional density of the state

$$p_k^e(\underline{x}_k) = p(\underline{x}_k | \underline{\hat{y}}_k, \underline{\hat{y}}_{k-1}, \ldots, \underline{\hat{y}}_1)$$

given all observations up to time step k. A recursive estimation procedure is preferred, which calculates a state estimate based on the estimate at the previous time step and hence, does not require to store all measurements. A suitable time update procedure produces a predicted density

$$p_k^p(\underline{x}_k) = p(\underline{x}_k | \underline{\hat{y}}_{k-1}, \ldots, \underline{\hat{y}}_1)$$

by propagating the previous estimate $p_{k-1}^e(\underline{x}_{k-1})$ through the system model. Although not strictly required, an initial density $p_0^e(\underline{x}_0)$ is assumed to be given.

Arbitrary characteristic values of the estimate such as mean, covariance matrix, mode, or median can be derived once the estimated density is available.

3. PSEUDO GAUSSIANS

The key idea is to represent complicated probability density functions in the N–dimensional original state space S_x by simpler densities in a higher–dimensional space S_x^*. Points $\underline{x}_k$ in S_x are related to points $\underline{x}_k^*$ in S_x^* via a nonlinear transformation $\underline{T}_x(.)$ according to

$$\underline{x}_k^* = \underline{T}_x(\underline{x}_k) = [T_1(\underline{x}_k), \ldots, T_{L_x}(\underline{x}_k)]^T \ ,$$

where L_x denotes the dimension of space S_x^*. Hence, the original space S_x is transformed by $\underline{T}_x(.)$ to an N–dimensional manifold U_x^* in the L_x–dimensional space S_x^*.

---

[1] We use a hat to indicate that the given measurement at time step k is a non–random quantity and is used as an estimate of the true measurement.

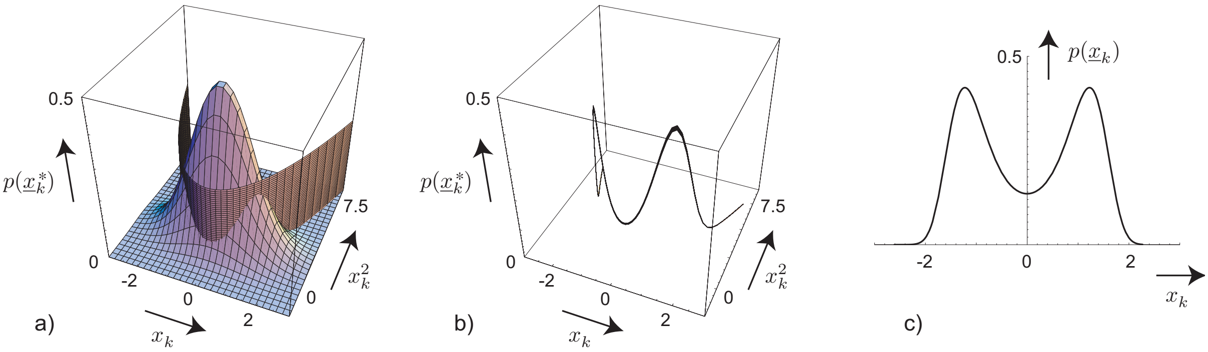

Fig. 1. Example for demonstrating the concept of pseudo Gaussians with scalar state x_k and twodimensional hyperspace S_x^*. a) Pseudo Gaussian in hyperspace S^* with mean and covariance matrix according to example 3.1. b) Parts of the pseudo Gaussian density lying on the manifold U_x^*. c) Corresponding density in the original space S_x.

In S_x^*, L_x–dimensional Gaussian probability density functions are defined according to

$$p(\underline{x}_k^*) = c_k^x \exp\left\{ -\frac{1}{2}(\underline{x}_k^* - \hat{\underline{x}}_k^*)^T (\mathbf{C}_x^*)^{-1} (\underline{x}_k^* - \hat{\underline{x}}_k^*) \right\}$$

with mean $\hat{\underline{x}}_k^*$, symmetric positive definite covariance matrix $\mathbf{C}_k^{x,*}$, and normalizing constant c_k^x. Densities of this type will be called pseudo Gaussian in the following, because the components of $\underline{x}_k^*$ are not independent.

The intersection of a pseudo Gaussian $p(\underline{x}_k^*)$ with the manifold U_x^* defines a non–Gaussian, e.g. multimodal, probability density function in the original space S_x.

Remark 3.1. A non–Gaussian density in the original space S_x is defined by *both* the transformation $\underline{T}_x(.)$ *and* the mean $\hat{\underline{x}}_k^*$ and covariance matrix $\mathbf{C}_k^{x,*}$ of the pseudo Gaussian $p(\underline{x}_k^*)$.

Example 3.1. A scalar state x_k is considered, which is related to a two–dimensional state $\underline{x}_k^*$ via

$$\underline{x}_k^* = \underline{T}_x(\underline{x}_k) = [x_k, x_k^2]^T \quad .$$

An example of a pseudo Gaussian density defined in the space S_x^* with mean

$$\hat{\underline{x}}_k^* = \begin{bmatrix} 0 & 2 \end{bmatrix}^T$$

and covariance matrix

$$\mathbf{C}_k^{x,*} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

is shown in Fig. 1 a) together with the manifold U_x^*. Fig. 1 b) then shows that part of the pseudo Gaussian density lying on the manifold U_x^*, which defines the density in the original space shown in Fig. 1 c).

The selection of the functions $T_i(\underline{x}_k)$, $i = 1, \ldots, L_x$ depends on the type of nonlinearity considered. However, multidimensional Bernstein polynomials appear to be a good choice in many cases, e.g. polynomial nonlinearities. They are defined on the basis of one–dimensional Bernstein polynomials, which on the interval $[l, r]$ are given by

$$H_i^n(x) = \binom{n}{i} \left(\frac{l-x}{l-r}\right)^i \left(\frac{r-x}{r-l}\right)^{n-i}$$

for $i = 0, \ldots, n$. With

$$\underline{x}_k = \begin{bmatrix} x_k^1 & x_k^2 & \ldots & x_k^N \end{bmatrix}^T \quad ,$$

the above transformation is defined by

$$T_i(\underline{x}_k) = \prod_{j=1}^{N} H_{i_j}^{L_j - 1}(x_k^j) \quad ,$$

for $i_j = 0, \ldots, L_j{-}1$, $j = 1, \ldots, N$, $L_x = \prod_{j=1}^N L_j$, and $i = \sum_{j=1}^N i_j$. For example, in two dimensions this gives

$$T_i(\underline{x}_k) = H_{i_1}^{L_1-1}(x_k^1)\, H_{i_2}^{L_2-1}(x_k^2) \quad ,$$

for $i_1 = 0, \ldots, L_1 - 1$, $i_2 = 0, \ldots, L_2 - 1$, $L_x = L_1 L_2$, and $i = i_1 + i_2$.

4. SYSTEM TRANSFORMATION

The original nonlinear system given by the linear system equation (1) and the nonlinear measurement equation (2) will now be converted to an equivalent linear representation in a higher–dimensional space.

For that purpose, the nonlinear measurement equation is transformed according to

$$\underline{T}_v(\hat{\underline{y}}_k - \underline{v}_k) = \underline{T}_v(\underline{h}_k(\underline{x}_k)) \quad . \qquad (3)$$

The left hand side is then converted into an affine function of $\underline{v}_k^*$

$$\underline{T}_v(\hat{\underline{y}}_k - \underline{v}_k) = -\mathbf{G}_k^* \underline{v}_k^* + \hat{\underline{y}}_k^* \quad ,$$

where the term $\hat{\underline{y}}_k^*$ does not depend on elements of $\underline{v}_k^*$. Of course, $\mathbf{G}_k^*$ and $\hat{\underline{y}}_k^*$ are polynomial functions

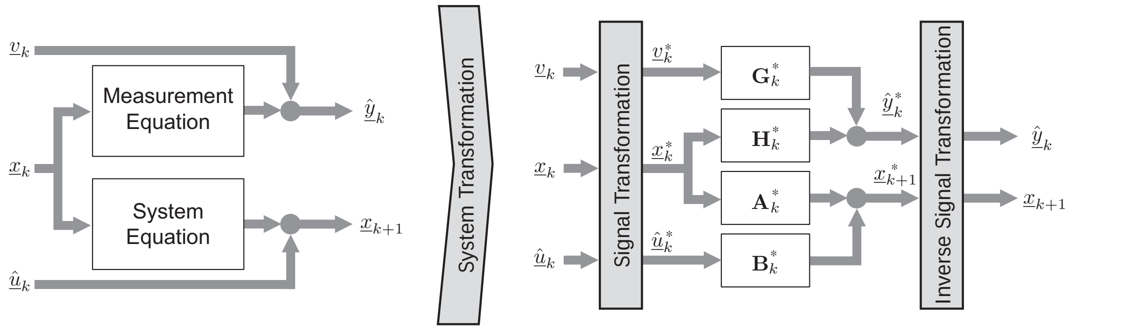

Fig. 2. Block diagram of the system transformation: The original system comprising a linear system equation and a nonlinear measurement equation is converted to a higher–dimensional representation, which is linear in a higher–dimensional space.

of the measurements $\hat{\underline{y}}_k$. The right hand side of (3) is expanded into a linear function of $\underline{x}_k^*$

$$\underline{T}_v(\underline{h}_k(\underline{x})) = \mathbf{H}_k^* \underline{x}_k^*$$

with

$$\underline{x}_k^* = \underline{T}_x(\underline{x}_k)$$

and $L_x \geq \max(N, L_v)$. This expansion is exact for a polynomial measurement nonlinearity $\underline{h}_k(.)$. Finally, we obtain a *linear* measurement equation

$$\hat{\underline{y}}_k^* = \mathbf{H}_k^* \underline{x}_k^* + \mathbf{G}_k^* \underline{v}_k^*$$

in S_x^* with $\hat{\underline{y}}_k^* \in \mathbb{R}^{L_v}$, $\underline{x}_k^* \in \mathbb{R}^{L_x}$, $\underline{v}_k^* \in \mathbb{R}^{L_v}$.

In addition, the system equation given by (1) is transformed according to

$$\underline{T}_x(\underline{x}_{k+1}) = \underline{T}_x(\mathbf{A}_k \, \underline{x}_k + \hat{\underline{u}}_k) \ ,$$

which is rewritten as

$$\underline{x}_{k+1}^* = \mathbf{A}_k^* \, \underline{x}_k^* + \mathbf{B}_k^* \, \hat{\underline{u}}_k^* \ .$$

5. FILTERING

Given the linear representation from Section 4, the desired densities are obtained by a linear filter operating in a higher–dimensional space S_x^* with state dimension L_x, provided the noise density $p_k^v(\underline{v}_k)$ is given as a pseudo Gaussian

$$p_k^v(\underline{v}_k^*) = c_k^v \exp\left\{ -\frac{1}{2} (\underline{v}_k^* - \hat{\underline{v}}_k^*)^T (\mathbf{C}_k^{v,*})^{-1} (\underline{v}_k^* - \hat{\underline{v}}_k^*) \right\}$$

in a space S_v^* with dimension L_v and the initial state is characterized by a pseudo Gaussian density $p_0^e(\underline{x}_0)$ defined by $\hat{\underline{x}}_0^{e,*}$ and $\mathbf{C}_0^{e,*}$. The prediction step is given by

$$\begin{aligned} \hat{\underline{x}}_{k+1}^{p,*} &= \mathbf{A}_k^* \, \hat{\underline{x}}_k^{e,*} + \mathbf{B}_k^* \, \hat{\underline{u}}_k^* \ , \\ \mathbf{C}_{k+1}^{p,*} &= \mathbf{A}_k^* \, \mathbf{C}_k^{p,*} \, (\mathbf{A}_k^*)^T \ , \end{aligned} \tag{4}$$

the filter step is given by

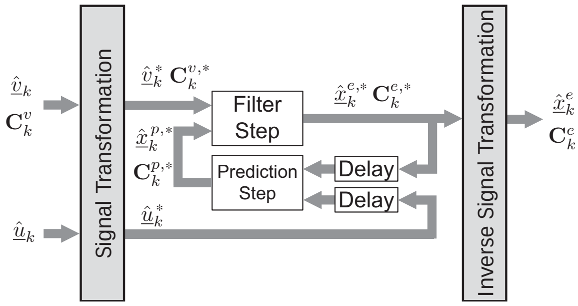

Fig. 3. Block diagram of the proposed new estimator: Estimation is performed by a linear estimator, e.g. a Kalman filter, in the higher–dimensional space. Please note that the recursion is completely performed in the higher–dimensional space. The inverse transformation is done outside of the recursion loop.

$$\begin{aligned} \hat{\underline{x}}_k^{e,*} &= \hat{\underline{x}}_k^{p,*} + \mathbf{C}_k^{p,*} (\mathbf{H}_k^*)^T \left\{ \mathbf{G}_k^* \mathbf{C}_k^{v,*} (\mathbf{G}_k^*)^T \right. \\ &\quad \left. + \mathbf{H}_k^* \mathbf{C}_k^{p,*} (\mathbf{H}_k^*)^T \right\}^{-1} (\hat{\underline{y}}_k^* - \mathbf{G}_k^* \hat{\underline{v}}_k^* - \mathbf{H}_k^* \hat{\underline{x}}_k^{p,*}) \ , \\ \mathbf{C}_k^{e,*} &= \mathbf{C}_k^{p,*} - \mathbf{C}_k^{p,*} (\mathbf{H}_k^*)^T \left\{ \mathbf{G}_k^* \mathbf{C}_k^{v,*} (\mathbf{G}_k^*)^T \right. \\ &\quad \left. + \mathbf{H}_k^* \mathbf{C}_k^{p,*} (\mathbf{H}_k^*)^T \right\}^{-1} \mathbf{H}_k^* \mathbf{C}_k^{p,*} \ , \end{aligned} \tag{5}$$

where nonzero mean measurement noise has been considered. However, to ensure symmetry and positive definiteness of the covariance matrix $\mathbf{C}_k^{e,*}$, square–root forms of the Kalman filter (Park and Kailath, 1995; Sayed and Kailath, 1994) are a better choice.

The recursion is completely performed in the higher–dimensional space, only the calculation of characteristic values of the estimate is done outside of the recursion loop. Typically, the mean $\hat{\underline{x}}_k^e$ and the covariance matrix $\mathbf{C}_k^e$ are provided, which in general requires numerical computation. However, an efficient algorithm for calculating moments of general exponential densities including pseudo–Gaussian densities by means of differential algebraic equations is given in (Rauh and Hanebeck, 2003).

6. EXAMPLE

To illustrate the proposed filtering algorithm, the following dynamic system with scalar state x_k is considered, which evolves in discrete time according to the linear system equation

$$x_{k+1} = a\, x_k + \hat{u}_k \ , \tag{6}$$

with a known scalar input sequence $\hat{u}_k$. Measurements $\hat{y}_k$ of the system output are related to the system state x_k via the nonlinear measurement equation

$$\hat{y}_k = x_k^2 + v_k \ . \tag{7}$$

The noise distribution is given by a two–dimensional pseudo Gaussian $p_k^v(\underline{v}_k^*)$, i.e., $L_v = 2$, with mean and covariance matrix

$$\hat{\underline{v}}_k^* = \begin{bmatrix} 0.5 \\ 2 \end{bmatrix} \ , \ \mathbf{C}_k^{v,*} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 2 \end{bmatrix} \ ,$$

where $\underline{v}_k^*$ is selected as

$$\underline{v}_k^* = \begin{bmatrix} v_k \\ v_k^2 \end{bmatrix}$$

for illustration purposes. $p_k^v(v_k)$ is visualized in Fig. 4.

Transformation of the original nonlinear measurement equation (7) and the system equation (6) according to Section 4 yields a linear measurement equation

$$\underbrace{\begin{bmatrix} \hat{y}_k \\ \hat{y}_k^2 \end{bmatrix}}_{\hat{\underline{y}}_k^*} \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{H}_k^*} \underbrace{\begin{bmatrix} \hat{x}_k \\ \hat{x}_k^2 \\ \hat{x}_k^3 \\ \hat{x}_k^4 \end{bmatrix}}_{\underline{x}_k^*} + \underbrace{\begin{bmatrix} 1 & 0 \\ 2\,\hat{y}_k & -1 \end{bmatrix}}_{\mathbf{G}_k^*} \underbrace{\begin{bmatrix} v_k \\ v_k^2 \end{bmatrix}}_{\underline{v}_k^*}$$

and a linear system equation

$$\underbrace{\begin{bmatrix} x_{k+1} \\ x_{k+1}^2 \\ x_{k+1}^3 \\ x_{k+1}^4 \end{bmatrix}}_{\underline{x}_{k+1}^*} = \underbrace{\begin{bmatrix} a & 0 & 0 & 0 \\ 2a\hat{u}_k & a^2 & 0 & 0 \\ 3a\hat{u}_k^2 & 3a^2\hat{u}_k & a^3 & 0 \\ 4a\hat{u}_k^3 & 6a^2\hat{u}_k^2 & 4a^3\hat{u}_k & a^4 \end{bmatrix}}_{\mathbf{A}_k^*} \underbrace{\begin{bmatrix} x_k \\ x_k^2 \\ x_k^3 \\ x_k^4 \end{bmatrix}}_{\underline{x}_k^*}$$

$$+ \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{B}_k^*} \underbrace{\begin{bmatrix} \hat{u}_k \\ \hat{u}_k^2 \\ \hat{u}_k^3 \\ \hat{u}_k^4 \end{bmatrix}}_{\hat{\underline{u}}_k^*}$$

in a higher–dimensional space with $L_x = 4$ and $L_u = 4$.

The state of the system given by (7) and (6) can now be estimated by means of a Kalman filter in the L_x–dimensional space according to Fig. 3 with the prediction step given by (4) and the filter step given by (5).

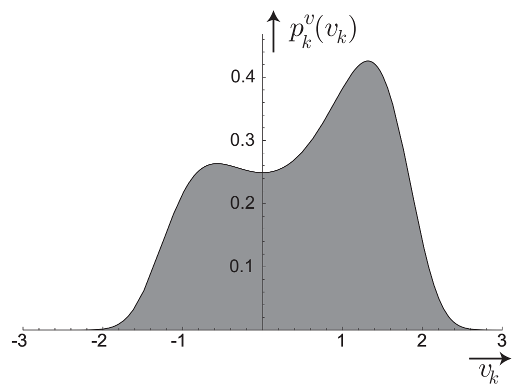

Fig. 4. Pseudo Gaussian noise density $p_k^v(v_k)$ discussed in the example.

Numerical simulations of a scalar system comprising a cubic measurement equation and a linear system equation are given in (Hanebeck, 2001*a*).

7. CONCLUSIONS

Estimating the state of a nonlinear dynamic system from measurements corrupted by non–Gaussian measurement noise is reformulated as a linear filtering problem in a higher–dimensional space. This is similar to the concept of support vector machines (Schölkopf, 1998), which perform nonlinear classification by means of linear hyperplane classifiers in a higher–dimensional space nonlinearly related to the input or problem space.

For that purpose, a specific type of exponential probability density function, the so–called pseudo Gaussian density, is introduced for representing the resulting non–Gaussian posterior densities. Furthermore, the measurement nonlinearity is converted into a linear measurement equation in a higher–dimensional space. Hence, proven linear filtering techniques can be employed for solving the nonlinear estimation problem.

A similar approach has been proposed for the case of nonlinear set–theoretic estimation in the case of unknown–but–bounded noise descriptions (Hanebeck, 2001*b*) and evaluated extensively in applications (Horn *et al.*, 2002; Briechle and Hanebeck, 2003).

8. EXTENSIONS

So far, an exact expansion of the measurement nonlinearity was assumed to exist. In that case a sufficient statistic is provided by the mean vectors and the covariance matrices of the pseudo

Gaussians used to represent the posterior densities. However, in many practical applications it is not possible to use an exact expansion of the nonlinearity. In addition, an approximation may be desirable to keep the dimensions of the hyperspace low even when an exact expansion is known. The resulting pseudo Gaussian densities are then approximations of the true posterior densities and are described by a nonsufficient or reduced statistic. However, an approximate expansion can be selected in such a way that a certain distance, e.g. the Kullback–Leibler distance, between the approximate and the exact posterior is minimized.

The proposed technique can also be applied to more complex additive noise descriptions, for example colored noise or noise with partially known statistics. These problems can be solved analogously by applying the appropriate linear filter in the higher–dimensional space, e.g. (Hanebeck *et al.*, 1999; Hanebeck and Horn, 2000; Julier and Uhlmann, 1997).

9. REFERENCES

Alspach, D. L. and H. W. Sorenson (1972). Nonlinear Bayesian Estimation Using Gaussian Sum Approximation. *IEEE Transactions on Automatic Control* **AC–17**(4), 439–448.

Anderson, B. D. O. and J. B. Moore (1979). *Optimal Filtering.* Prentice–Hall.

Bergman, N., L. Ljung and F. Gustafsson (1999). Terrain Navigation Using Bayesian Statistics. *IEEE Control Systems Magazine* **19**(3), 33–40.

Bohn, C. and H. Unbehauen (2000). The Application of Matrix Differential Calculus for the Derivation of Simplified Expressions in Approximate Non–Linear Filtering Algorithms. *Automatica* **36**(10), 1553–1560.

Briechle, K. and U. D. Hanebeck (2003). Localization of Mobile Robots Using Relative Bearing Measurements. *to appear in: IEEE Transactions on Robotics and Automation.*

Bucy, R. S. and K. D. Senne (1971). Digital Synthesis of Non–linear Filters. *Automatica* **7**, 287–298.

Challa, S., Y. Bar-Shalom and V. Krishnamurthy (2000). Nonlinear Filtering via Generalized Edgeworth Series and Gauss–Hermite Quadrature. *IEEE Transactions on Signal Processing* **48**(6), 1816–1820.

Doucet, A., S. Godsill and C. Andrieu (2000). On Sequential Monte Carlo Sampling Methods for Bayesian Filtering. *Statistics and Computing* **10**(3), 197–208.

Hanebeck, U. D. (2001*a*). Optimal Filtering for Polynomial Measurement Nonlinearities with Additive Non-Gaussian Noise. In: *Proceedings of the American Control Conference (ACC'2001), Arlington, Virginia.*

Hanebeck, U. D. (2001*b*). Recursive Nonlinear Set–Theoretic Estimation Based on Pseudo–Ellipsoids. In: *Proceedings of the IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI'2001), Baden–Baden.* pp. 159–164.

Hanebeck, U. D. and J. Horn (2000). Fusing Information Simultaneously Corrupted by Uncertainties with Known Bounds and Random Noise with Known Distribution. *Information Fusion, Elsevier Science* **1**(1), 55–63.

Hanebeck, U. D., J. Horn and G. Schmidt (1999). On Combining Statistical and Set Theoretic Estimation. *Automatica* **35**(6), 1101–1109.

Horn, J., U. D. Hanebeck, K. Riegel, K. Heesche and W. Hauptmann (2002). Nonlinear Set–Theoretic Position Estimation of Cellular Phones. *Ortung und Navigation, Deutsche Gesellschaft für Ortung und Navigation e.V. (DGON)* (1), 93–99.

Jazwinski, A. H. (1970). *Stochastic Processes and Filtering Theory.* Academic Press.

Julier, S. and J. K. Uhlmann (1997). A Nondivergent Estimation Algorithm in the Presence of Unknown Correlations. In: *Proceedings of the American Control Conference (ACC'1997).*

Kulhavý, R. (1992). Recursive Nonlinear Estimation: Geometry of a Space of Posterior Densities. *Automatica* **28**(2), 313–323.

Liu, J. S. and R. Chen (1998). Sequential Monte Carlo Methods for Dynamic Systems. *Journal of American Statistical Association* **93**, 1032–1043.

Park, P. and T. Kailath (1995). New Square–Root Algorithms for Kalman Filtering. *IEEE Transactions on Automatic Control* **40**(5), 895–899.

Rauh, A. and U. D. Hanebeck (2003). Calculating Moments of Exponential Densities Using Differential Algebraic Equations. *to appear in: IEEE Signal Processing Letters.*

Sayed, A. H. and T. Kailath (1994). A State–Space Approach to Adaptive RLS Filtering. *IEEE Signal Processing Magazine* **11**(3), 18–70.

Schölkopf, B. (1998). Support Vector Machines – A Practical Consequence of Learning Theory. In: M. A. hearst, B. Schölkopf, S. Dumais, E. Osuna, J. Platt. Trends and Controversies – Support Vector Machines. *IEEE Intelligent Systems* **13**(4), 18–28.

Sorenson, H. W. and A. R. Stubberud (1968). Non–Linear Filtering by Approximation of the A Posteriori Density. *International Journal of Control* **8**(1), 33–51.