New Estimators for Mixed Stochastic and Set Theoretic Uncertainty Models: The Scalar Measurement Case

Uwe D. Hanebeck
Institute of Automatic Control Engineering
Technische Universität München
80290 München, Germany
e-mail: Uwe.Hanebeck@ei.tum.de

Joachim Horn
Siemens AG, Corporate Technology
Information and Communications
81730 München, Germany
e-mail: Joachim.Horn@mchp.siemens.de

Abstract

New filters are derived for estimating the state of a linear dynamic system based on uncertain observations, which suffer from two types of uncertainties simultaneously. The first uncertainty is a stochastic process with given distribution. The second uncertainty is only known to be bounded, the exact underlying distribution is unknown. The new estimators combine set theoretic and stochastic estimation in a rigorous manner and provide a continuous transition between the two classical estimation concepts. They converge to a set theoretic estimator, when the stochastic error goes to zero, and to a Kalman filter, when the bounded error vanishes. In the mixed noise case, solution sets are provided that are uncertain in a stochastic sense.

1 Introduction

Estimating the state of a dynamic system based on uncertain observations is a topic of extraordinary importance. In a wide variety of applications, where an appropriate system model together with a stochastic noise model is given, the Kalman filter [1] and its many variations have proven to be useful.

In many cases, however, uncertainties arise, for example, from unmodeled dynamics or unmodeled nonlinearities, which cannot satisfactorily be described as stochastic signals with known distribution. In addition, correlated noise terms or systematic errors may be present but neglected for the sake of simplicity. In that case, Kalman filter estimates tend to be overoptimistic [12], i.e., the covariance is underestimated. Several heuristics have been suggested for coping with this problem, which of course do not provide optimal estimators.

In some situations, however, bounds for these uncertainties can be provided. In that case, set theoretic estimation can be applied [14], which often leads to good results [4]. However, when additional uncorrelated noise is present, the error bounds become unnecessarily conservative.

In [5, 8], a concept for estimation in the presence of both bounded and stochastic uncertainties has been introduced. The proposed algorithm for the case of a scalar state is exact, but computationally complex. In [6, 7], an approximate solution for the case of a scalar state has been derived, that is computationally attractive. Furthermore, a generalization towards arbitrary dimensional states and observations of the same dimension has been proposed in [9].

This paper is concerned with updating the estimate of an arbitrary dimensional state based on scalar observations. For this very relevant case, a new, approximate solution is derived, that is computationally attractive. Nevertheless, it combines both set theoretic and stochastic estimation in a rigorous manner. It bridges the gap between both estimation schemes, because a set theoretic estimator is obtained, when the stochastic error goes to zero. A Kalman filter is obtained, when the bounded error vanishes. When both types of uncertainty are present, the new estimator provides solution sets that are uncertain in a stochastic sense. The propagation of estimates suffering from both uncertainties through a dynamic system is discussed in [10].

In Section 2, the use of a mixed stochastic and set theoretic uncertainty model is motivated, and a rigorous problem formulation is given. In Section 3, the estimation concept is presented. In Section 4, the estimation problem is solved on the basis of a sum approximation. An exact second-order description is derived in Section 5. In Section 6, the results are summarized and border cases of the new estimator are discussed. In Section 7, the proposed framework is applied to estimating the state of a nonlinear system by first converting it to a linear system with mixed stochastic and bounded uncertainties.
2 Problem Formulation

The key point of this paper is the use of a generalized uncertainty model unifying stochastic and set theoretic modeling [2, 3]. This allows the treatment of systems corrupted by both bounded and stochastic uncertainties simultaneously. Hence, the model is well-suited for, but not limited to, the combination of deterministic / systematic errors and random noise.

To be specific, we consider a linear measurement equation given by

\[ y = H^T \bar{x} + e_y + c_y \]

with scalar observation \( y \), state vector \( \bar{x} \), and additive uncertainties \( e_y, c_y \). Furthermore, there exists a prior estimate \( \bar{x}_p \) of the state vector. \( \bar{x}_p \) also suffers from additive uncertainties \( \bar{c}_p \), \( c_p \) according to

\[ \bar{x}_p = \bar{x} + \bar{c}_p + c_p. \]

The corresponding additive uncertainties are of different type:

1) \( \bar{c}_p, e_y \) are uncertainties, where the only prior knowledge is their boundedness, which is expressed by

\[ \bar{c}_p^T E_p^{-1} \bar{c}_p \leq 1, \quad e_y^2 \leq E_y. \]

2) \( \bar{c}_p, c_y \) are Gaussian random variables

\[ \bar{c}_p \sim N(0, C_p), \quad c_y \sim N(0, \sigma_y), \]

which are assumed to be uncorrelated.

3 The Estimation Concept

For deriving an appropriate state estimator, we define

\[ \bar{y} = y - e_y. \]

Since there is no prior information about the remaining uncertainties \( \bar{c}_p, e_y \) besides their boundedness, we make the worst case assumption that \( \bar{c}_p, e_y \) are fully correlated. Hence, a set theoretic estimator is appropriate for fusing \( \bar{y} \) and \( \bar{x}_p \). The fusion result is then given by the set

\[ \bar{X}_s = \{ \bar{x}_s : (\bar{x}_s - \bar{x}_p)^T E_p^{-1} (\bar{x}_s - \bar{x}_p) \leq 1 \} \]

with

\[ \bar{x}_s = \bar{x}_p + \lambda \frac{E_p H}{E_y + \lambda H^T E_p H} \eta, \quad \eta = \bar{y} - H^T \bar{x}_p \quad (1) \]

and

\[ E_s = d P_s, \quad (2) \]

where \( d \) is given by

\[ d = 1 + \lambda \frac{\eta^2}{E_y + \lambda H^T E_p H} \]

and \( P_s \) is given by

\[ P_s = E_p - \lambda \frac{E_p H H^T E_p}{E_y + \lambda H^T E_p H} \]

The appropriate selection of the parameter \( \lambda \in [0, \infty] \) will be discussed later. (1) can be rewritten as

\[ \bar{x}_s = W_x \bar{x}_p + W_y \bar{y}, \quad (3) \]

with

\[ W_x = I - \lambda \frac{E_p H H^T}{E_y + \lambda H^T E_p H}, \quad W_y = \frac{\lambda E_p H}{E_y + \lambda H^T E_p H} \]

and

\[ W_x + W_y H^T = I. \]

To simplify the following derivations, we note the following: The set theoretic uncertainty \( E_s \) given by (2) depends on \( \bar{y} \) and \( \bar{x}_p \). Setting \( \eta = 0 \) leads to \( d = 1 + \lambda \) and is equivalent to bounding \( E_s \) from above. The resulting \( E_s \) is then given by

\[ E_s = (1 + \lambda) E_p - (1 + \lambda) \lambda \frac{E_p H H^T E_p}{E_y + \lambda H^T E_p H}. \quad (4) \]

Since the simplified \( E_s \) in (4) does not depend on the actual values, it is not a random variable. On the other hand, \( \bar{x}_s \) according to

\[ \bar{x}_s = W_x \bar{x}_p + W_y \bar{y}, \]

is a random variable. We obtain \( \bar{x}_s = \bar{x}_s + \bar{c}_s \), with \( \bar{c}_s = W_x \bar{c}_p + W_y c_y \). In this paper, we provide two different solutions for estimating the unknown state vector \( \bar{x}_s \):

- An approximation of the density of \( \bar{x}_s \), by a sum (Sec. 4), which in contrast to the exact density can be used for recursive estimation.
- An exact second-order description, i.e., mean and covariance (Sec. 5).

4 Calculating the Density

Since \( \bar{x}_s \) is defined only for

\[ |y - H^T \bar{x}_p| \leq K \]

with \( K = \sqrt{E_y} + \sqrt{H^T E_p H} \),
the density of $\bar{x}$, is given by [13]

$$f_{\bar{x}}(\bar{x}_s) = \frac{1}{|W_x|} \int_{-\infty}^{\infty} f_{\bar{x},y}(W_x^{-1}(\bar{x}_s - W_y y), y) \, dy$$  \tag{5}$$

with

$$f_{\bar{x},y}(\bar{x}_p, y) = \begin{cases} f_{\bar{x}}(\bar{x}_p) f(y) & \text{for } |y - H^T \bar{x}_p| \leq K \\ 0 & \text{elsewhere} \end{cases},$$

where $W_x$ is assumed to be regular. Calculating the exact density $f_{\bar{x}}(\bar{x}_s)$ directly from (5) gives a rather complicated and not very useful expression. Hence, a series expansion will be calculated instead. For that purpose, we use

$$\text{rect}(x) = \begin{cases} 1 & \text{for } |x| \leq 1 \\ 0 & \text{elsewhere} \end{cases}$$

to interpret the constraint $|y - H^T \bar{x}_p| \leq K$ as

$$\text{rect}(\frac{y - H^T \bar{x}_p}{K}) \approx \sum_{i=-L}^{L} \frac{1}{\sqrt{2\pi c}} \exp \left\{ -\frac{1}{2} \left( \frac{y - H^T \bar{x}_p - m_i^t}{C_g} \right)^2 \right\}$$

with $m_i^t = \frac{K}{L}, C_g = \frac{K}{L}$. Note that the integral from $-\infty$ to $\infty$ over the sum yields $2K$ independent of $L$. The free parameter $c \in (0, \infty)$ may be obtained by a one-dimensional search to give the best approximation of the rect–function according to a given norm.

Based on this approximation of the rect–function, the exact density $f_{\bar{x}}$ in (5) can be approximated by a sum of simple densities. For that purpose, we first consider one term of the sum (6) which gives

$$f_{\bar{x}_s}(\bar{x}_s) = \frac{1}{|W_x|} \int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{2} \left[ (x_p - \hat{x}_p)^T C_p^{-1} (x_p - \hat{x}_p) \right. \right.$$  

$$+ \frac{1}{C_y} (y - \hat{y})^2 + \frac{1}{C_g} (y - H^T \bar{x}_p - m_i^t)^2 \left. \right\} dy$$

for $i = -L, \ldots, L$ and $\bar{x}_p = W_x^{-1}[\bar{x}_s - W_y y]$. A tedious calculation reveals that this approximation can be simplified to

$$f_{\bar{x}_s}(\bar{x}_s) = g_i \exp \left\{ -\frac{1}{2} (x_p - \hat{x}_i)^T C_i^{-1} (x_p - \hat{x}_i) \right\}$$

with weighting factors

$$g_i = \exp \left\{ -\frac{1}{2} \left( \frac{y - H^T \hat{\bar{x}}_p - m_i^t}{2 C_i^y} \right)^2 \right\}$$

and individual means

$$\hat{x}_i = W_x \hat{x}_p + W_y \hat{y}$$

$$- \frac{W_x C_p H}{H^T C_p H + C_y + C_g} \left( \hat{y} - H^T \hat{\bar{x}}_p - m_i^t \right).$$

for $i = -L, \ldots, L$. The covariance matrices are the same for each term in the sum and given by

$$C_s = W_x C_p W_x^T + W_y W_y^T C_y - \frac{W_x C_p H}{H^T C_p H + C_y + C_g} (W_x C_p H - W_y C_y)^T H^T C_p H + C_y + C_g.$$ 

The approximate solution for the density $f_{\bar{x}}$ is then given by

$$f_{\bar{x}}(\bar{x}_s) \approx \sum_{i=-L}^{L} f_{\bar{x}_i}(\bar{x}_s),$$

which is a weighted sum of Gaussian densities, where the weighting factors $g_i$ are themselves values of a Gaussian function.

**Note:** It can be proven that this approximation converges to the exact density for $L \to \infty$.

## 5 Exact Analytic Solutions for Mean and Covariance

In the following, an exact second-order description for $\bar{x}$, i.e., mean and covariance, will be derived.

### 5.1 Exact analytic solution for the mean

An approximate expression for the mean or expected value $\hat{x}_s = E[\bar{x}_s]$ of $\bar{x}_s$ is given by

$$\hat{x}_s \approx \frac{\sum_{i=-L}^{L} g_i \hat{x}_i}{\sum_{i=-L}^{L} g_i}.$$ 

For $L \to \infty$, this expression gives the exact mean of $\hat{x}_s$; $L \to \infty$ also implies $C_g \to 0$, and the summation can be replaced by integration. A lengthy calculation gives

$$\hat{x}_s = W_x \hat{x}_p + W_y \hat{y}$$


$$- (W_x C_p H - W_y C_y) F_1 (\hat{y} - H^T \hat{\bar{x}}_p),$$

with $F_1 (\hat{y} - H^T \hat{\bar{x}}_p)$ according to the appendix.
5.2 Exact Solution for the Covariance

For obtaining the exact covariance of \( \mathbf{x} \), the first step is to calculate the covariance of the (approximate) Gaussian sum density \( f_\mathbf{x} (\mathbf{x}) \) in (7) based on the relation

\[
C_s = \frac{1}{L} \sum_{i=-L}^{L} g_i \left\{ \mathbf{C}_s + \hat{\mathbf{x}}_i (\hat{\mathbf{x}}_i^T) \right\} - \hat{\mathbf{x}}_i \hat{\mathbf{x}}_i^T,
\]

which gives the exact \( C_s \) for \( L \to \infty \) as

\[
C_s = W_x C_p W_x^T + W_x W_y C_y - F_2 (\hat{\mathbf{y}} - H^T \hat{\mathbf{x}}_p) \left( W_x C_p H - W_y C_y \right)^T F_2^T,
\]

with \( F_2 (\hat{\mathbf{y}} - H^T \hat{\mathbf{x}}_p) \) according to the appendix.

6 The New Estimators

In Section 4, it has been shown, that the uncertainty of the fusion result is given by a bounded uncertainty and a sum of Gaussian densities. When the number of terms included in the Gaussian sum tends towards infinity, the exact density is approached. In addition, mean and covariance of the exact density have been given in closed form. These important results can now be applied to derive two different estimators for solving practical estimation problems.

The first estimator is obtained by keeping a finite number of, say \( M \), terms in the Gaussian sum. Of course, when using this estimator recursively, there will be \( M^2 \) terms after the first recursion step. Hence, for recursive application, the number of terms must be kept fixed by selecting the \( M \) most important terms after each step.

The second estimator only keeps second order information on both the set theoretic and the stochastic uncertainty. Here, the stochastic uncertainty is given by the exact mean and covariance derived in Sec. 5.

Both estimators unify Kalman filtering and set theoretic estimation: A Kalman filter is approached, when the bounded error vanishes. On the other hand, a set theoretic estimator is attained, when the stochastic error goes to zero. When both types of uncertainty are present simultaneously, the new estimator provides solution sets that are uncertain in a stochastic sense.

7 Simulative Example

Consider two nonlinear measurement equations according to

\[
\begin{align*}
\mathbf{y}_1^k &= x_1 + 2 \sin(x_1) + c_1^k, \\
\mathbf{y}_2^k &= 2x_1 + x_2 + 3 \cos(x_2) + c_2^k,
\end{align*}
\]

with \( c_1^k, c_2^k \) samples from independent, zero mean white Gaussian random processes with standard deviation \( \sigma_1 = 3, \sigma_2 = 3 \), respectively. Prior knowledge is given by a \( \mathbf{x}_0 = [20, 20]^T \) with covariance matrix \( \mathbf{C}_0 = \text{diag}(100^2, 100^2) \).

A standard approach for estimating the state \( \mathbf{z} = [x_1, x_2]^T \) is the extended Kalman filter, which applies the Kalman filter to the measurement equations linearized about the best available estimate. The evolution of the resulting confidence set over time is depicted in Figure 1. The confidence set has been calculated based on 9 times the covariance matrix centered at \( \hat{\mathbf{x}}_0^k \). The true state \( \mathbf{z} = [17, 13]^T \) is marked by a dot. Note: The confidence set for \( k \to \infty \) does not contain the true state.

For estimating the state \( \mathbf{z} \) by using the above framework, the original nonlinear measurement equations are written as linear equations

\[
\begin{align*}
\mathbf{y}_1^k &= \mathbf{H}_1^k \mathbf{z} + e_1^k + e_1^k, \\
\mathbf{y}_2^k &= \mathbf{H}_2^k \mathbf{z} + e_2^k + e_2^k,
\end{align*}
\]

with \( \mathbf{H}_1^k = [1, 0]^T, \mathbf{H}_2^k = [2, 1]^T \). The neglected nonlinearities are captured by additional bounded uncertainties \( e_1^k, e_2^k \). The bounds of \( e_1^k, e_2^k \) are given by

\[
\begin{align*}
E_1 &= \max_{x \in \mathbb{R}^2} \left[ 2 \sin(x_1) \right]^2 = 2^2, \\
E_2 &= \max_{x \in \mathbb{R}^2} \left[ 3 \cos(x_2) \right]^2 = 3^2.
\end{align*}
\]

The proposed estimator is evaluated by recursively updating the state estimate using the equation for \( \mathbf{z}^k \) in (8), \( \mathbf{E}^k \) in (4), and \( \mathbf{C}^k \) in (9) for both measurement equations at time \( k \). The parameter \( \lambda_k \) is chosen such that \( \text{det}(\mathbf{E}^k) + 9 \text{det}(\mathbf{C}^k) \) is minimized. The initial set theoretic uncertainty is set to \( \mathbf{E}_0^k = \text{diag}(\epsilon, \epsilon) \), with a small \( \epsilon > 0 \). Figure 2 depicts how the resulting estimate evolves over time. Here, the confidence set is given by the Minkowski sum of \( \mathbf{E}_k^k \) and \( 9 \mathbf{C}_k^k \) centered at \( \hat{\mathbf{z}}_0^k \). The optimal estimate for an infinite number of measurements would be given by the set resulting from intersecting the two strips representing the neglected nonlinearities. Note: The confidence set for \( k \rightarrow \infty \) bounds the exact set from above, and hence contains the true state.

8 Conclusions

Many estimation problems can be converted to the problem of estimating the state of a linear system from uncertain observations, where the uncertainties are additively composed of both 1) noise with known distribution and 2) noise with known bounds. The new estimators then provide a rigorous framework for solving these problems efficiently.
This paper focused on the measurement update, i.e., on updating the estimate of an arbitrary dimensional state based on given scalar observations. The time update, i.e., propagating the state estimate through a system model, is discussed in [10].

As an example, a nonlinear system has been converted to a linear system by representing the neglected nonlinearities as bounded noise terms. The new estimator then provided state estimates that account for both uncertainties due to measurement noise and neglected nonlinearities.

**Figure 1:** Results of applying the extended Kalman filter.

**Figure 2:** Results of applying the new estimator.

**References**


The nonlinear functions $F_1 (\hat{y} - H^T \hat{x}_p)$, $F_2 (\hat{y} - H^T \hat{x}_p)$ of the innovation $\hat{y} - H^T \hat{x}_p$ are given by

$$F_1 (\hat{y} - H^T \hat{x}_p) = G_0 \left( \hat{y} - H^T \hat{x}_p, \sqrt{E_y} + \sqrt{H^T E_p H} C_y + \sqrt{H^T C_p H} \right),$$

$$F_2 (\hat{y} - H^T \hat{x}_p) = \left[ G_0 \left( \hat{y} - H^T \hat{x}_p, \sqrt{E_y} + \sqrt{H^T E_p H} C_y + \sqrt{H^T C_p H} \right) \right]^2$$

$$+ \frac{G_1 \left( \hat{y} - H^T \hat{x}_p, \sqrt{E_y} + \sqrt{H^T E_p H} C_y + \sqrt{H^T C_p H} \right)}{C_y + H^T C_p H}.$$

with functions $G_0$ and $G_1$

$$G_0(x, B, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} \exp \left\{ -\frac{1}{2} \frac{(x-B)^2}{\sigma^2} \right\} - \exp \left\{ -\frac{1}{2} \frac{(x+B)^2}{\sigma^2} \right\}$$

$$- \text{erf}\left\{ \frac{x-B}{\sigma} \right\} - \text{erf}\left\{ \frac{x+B}{\sigma} \right\},$$

$$G_1(x, B, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} (x-B) \exp \left\{ -\frac{1}{2} \frac{(x-B)^2}{\sigma^2} \right\} - (x+B) \exp \left\{ -\frac{1}{2} \frac{(x+B)^2}{\sigma^2} \right\}$$

$$- \text{erf}\left\{ \frac{x-B}{\sigma} \right\} - \text{erf}\left\{ \frac{x+B}{\sigma} \right\}.$$

Appendix

The nonlinear functions $F_1 (\hat{y} - H^T \hat{x}_p)$, $F_2 (\hat{y} - H^T \hat{x}_p)$ of the innovation $\hat{y} - H^T \hat{x}_p$ are given by

$$F_1 (\hat{y} - H^T \hat{x}_p) = G_0 \left( \hat{y} - H^T \hat{x}_p, \sqrt{E_y} + \sqrt{H^T E_p H} C_y + \sqrt{H^T C_p H} \right),$$

$$F_2 (\hat{y} - H^T \hat{x}_p) = \left[ G_0 \left( \hat{y} - H^T \hat{x}_p, \sqrt{E_y} + \sqrt{H^T E_p H} C_y + \sqrt{H^T C_p H} \right) \right]^2$$

$$+ \frac{G_1 \left( \hat{y} - H^T \hat{x}_p, \sqrt{E_y} + \sqrt{H^T E_p H} C_y + \sqrt{H^T C_p H} \right)}{C_y + H^T C_p H}.$$

with functions $G_0$ and $G_1$

$$G_0(x, B, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} \exp \left\{ -\frac{1}{2} \frac{(x-B)^2}{\sigma^2} \right\} - \exp \left\{ -\frac{1}{2} \frac{(x+B)^2}{\sigma^2} \right\}$$

$$- \text{erf}\left\{ \frac{x-B}{\sigma} \right\} - \text{erf}\left\{ \frac{x+B}{\sigma} \right\},$$

$$G_1(x, B, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} (x-B) \exp \left\{ -\frac{1}{2} \frac{(x-B)^2}{\sigma^2} \right\} - (x+B) \exp \left\{ -\frac{1}{2} \frac{(x+B)^2}{\sigma^2} \right\}$$

$$- \text{erf}\left\{ \frac{x-B}{\sigma} \right\} - \text{erf}\left\{ \frac{x+B}{\sigma} \right\}.$$