A New Estimator for Mixed Stochastic and Set Theoretic Uncertainty Models Applied to Mobile Robot Localization

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
This work presents new results for state estimation based on noisy observations suffering from two different types of uncertainties: The first uncertainty is a stochastic process with given statistics. The second uncertainty is only known to be bounded, the exact underlying statistics are unknown. State estimation tasks of this kind typically arise in target localization, navigation, and sensor data fusion. A new estimator has been developed, that combines set theoretic and stochastic estimation in a rigorous manner. The estimator is efficient and, hence, well-suited for practical applications. It provides a continuous transition between the two classical estimation concepts, because it converges 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, the new estimator provides solution sets that are uncertain in a statistical sense.

1 Introduction
In many applications, it is important to deduce a system’s state on the basis of uncertain observations of the system’s output. In addition, uncertain results of different estimators must be combined. Applications include vehicle or missile localization, target tracking, navigation, and sensor data fusion. The goal of these estimation procedures is, of course, to reduce the uncertainty about the system’s state as much as possible.

When an appropriate system model together with noise statistics is given, the Kalman filter and its descendants [1] have been successfully applied for more than 30 years. However, in the applications cited above, a detailed statistical noise model is often either not available or impractical. Special caution is in order when neglecting strongly correlated noise or systematic errors. In that case, Kalman filter estimates tend to be overoptimistic [8], i.e., the covariance estimate is unrealistically small. Several heuristics have been suggested for coping with this problem, ranging from artificially increasing the covariance from time to time to employing nonlinear pre-filters. Of course, these techniques do not provide optimal estimators.

In some situations, although a statistical noise description cannot be given, bounds for the noise can be provided. This may be the case for unmodeled dynamics, unmodeled nonlinearities, correlated noise, and systematic errors. In that case, set theoretic estimation can be applied [10], which often leads to good results [4]. However, when additional uncorrelated noise is present, the error bounds become unnecessarily conservative.

In [5], [6], a concept for estimation in the presence of both set theoretic and stochastic uncertainties has been introduced. The proposed algorithm is exact, but computationally complex. This paper presents a new, approximate solution, 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 attained, when the stochastic error goes to zero, and a Kalman filter, when the bounded error vanishes. When both types of uncertainty are present, the new estimator provides solution sets that are uncertain in a statistical sense.

We restrict attention to the scalar case for the sake of simplicity.

2 Estimation Concept
Consider two uncertain measurements of an unknown state \( z \) given by

\[
\begin{align*}
    x &= z + e_x + c_x, \quad e_x^2 \leq E_x, \quad c_x \sim N(0, \sigma_x), \\
    y &= z + e_y + c_y, \quad e_y^2 \leq E_y, \quad c_y \sim N(0, \sigma_y). 
\end{align*}
\]


\[ x \text{ and } y \text{ suffer from two types of additive noise } [2, [3]: \\
1 \text{ Uncertainties } e_x, \ e_y, \text{ where the only prior knowledge is their boundedness and } 2 \text{ zero mean Gaussian noise processes } e_x, \ e_y, \text{ which are assumed to be uncorrelated.}
\]

First, assume that \( x, y \) can be observed without stochastic uncertainty. Then, since there is no prior information about \( e_x, \ e_y \) besides their boundedness, we make the worst case assumption that \( e_x, \ e_y \) are fully correlated. In that case, a set theoretic estimator is appropriate for fusing the information sources. An efficient form of a set theoretic estimator, which is an approximation based on the convex combination of the original sets, is given by the set \([10]\)

\[
\mathcal{S} = \{ z : (z - \hat{z})^2 \leq E_z \}, \tag{2}
\]

which is an interval in the scalar case. The interval midpoint is given by

\[
\hat{z} = w_x x + w_y y \tag{3}
\]

with weighting factors

\[
w_x = (0.5 - \lambda) \frac{P_z}{E_x}, \\
w_y = (0.5 + \lambda) \frac{P_z}{E_y} \tag{4}
\]

with

\[
P_z = \left( \frac{0.5 - \lambda}{E_x} + \frac{0.5 + \lambda}{E_y} \right)^{-1} \tag{5}
\]

where \( w_x + w_y = 1 \). The appropriate selection of the parameter \( \lambda \in [-0.5, 0.5] \) will be discussed later. The set theoretic uncertainty is given by

\[
E_z = d \cdot P_z \tag{6}
\]

with

\[
d = 1 - (0.25 - \lambda^2)(x - y)^2 \frac{P_z}{E_x E_y} \tag{7}
\]

It is important to note that the set theoretic uncertainty \( E_z \) depends on the actual observations \( x, y \). From now on, we will bound \( E_z \) from above by setting \( d = 1 \), which gives

\[
E_z = \left( \frac{0.5 - \lambda}{E_x} + \frac{0.5 + \lambda}{E_y} \right)^{-1} \tag{8}
\]

The so obtained interval always contains the true interval. Most importantly, \( E_z \) does not depend on the actual observations, which simplifies some of the following derivations.

Since \( x, y \) cannot be observed directly, but are corrupted by Gaussian noise, \( \hat{z} \) is a random variable with statistics that can be obtained from (3). Note that \( E_z \) in (8) is not a random variable, because it does not depend on the actual observation. We provide three solutions:

\[ \diamond \text{ The exact density of } \hat{z} \text{ (Sec. 3),} \]
\[ \diamond \text{ an approximation of the density by a sum (Sec. 4),} \]
\[ \diamond \text{ and an exact second-order description, i.e., mean and variance (Sec. 5).} \]

The first solution is mainly useful for validation purposes, the other two solutions are useful for data-recursive estimation.

### 3 The Exact Density

For \( w_x \neq 0 \), the density \( f_z(z) \) of \( \hat{z} \) is given by

\[
f_z(z) = \frac{1}{w_x} \int_{-\infty}^{\infty} f_{xy} \left( \frac{z - w_y y}{w_x}, y \right) dy \tag{9}
\]

From (1) we deduce the constraint

\[
|x - y| \leq B, \quad B = \sqrt{E_x} + \sqrt{E_y}, \tag{10}
\]

which leads to

\[
f_{xy}(x, y) = \begin{cases} 
  f_x(x) \cdot f_y(y) & \text{for } |x - y| \leq B \\
  0 & \text{elsewhere}
\end{cases} \tag{11}
\]

The constraint can be rewritten as

\[
-B \leq \frac{z - w_y y}{w_x} \leq B \iff \frac{z - w_x B}{w_x + w_y} \leq y \leq \frac{z + w_y B}{w_x + w_y}.
\]

Hence, (9) gives

\[
f_z(z) = \frac{1}{w_x} \int_{-\infty}^{\infty} f_x \left( \frac{z - w_y y}{w_x} \right) f_y(y) dy \tag{12}
\]

and after some manipulations

\[
f_z(z) = \frac{1}{w_x 2 \pi \sigma_x \sigma_y} \exp \left\{ -\frac{1}{2} \left( \frac{z - (w_x m_x + w_y m_y)}{w_x \sigma_x + w_y \sigma_y} \right)^2 \right\} \hat{I}(z) \tag{13}
\]

with

\[
\hat{I}(z) = \int_{\frac{-w_x B}{w_x + w_y}}^{\frac{w_x B}{w_x + w_y}} \exp \left\{ \frac{w_x^2 S_x + w_y^2 S_y}{w_x^2 S_x + w_y^2 S_y} \right\} \left( y - \frac{w_x^2 m_x S_x - w_x w_y m_x S_y + w_y S_y z}{w_x^2 S_x + w_y^2 S_y} \right)^2 dy.
\]
\[ I(z) = \frac{\sqrt{2\pi w_x \sigma_x \sigma_y}}{\sqrt{w_x^2 S_x + w_y^2 S_y}} \left\{ \text{erf} \left( \frac{\sigma_x \sigma_y (w_x + w_y) \sqrt{w_x^2 S_x + w_y^2 S_y}}{2} \right) \right\} + \frac{\text{erf} \left( \frac{\sigma_x \sigma_y (w_x + w_y) \sqrt{w_x^2 S_x + w_y^2 S_y}}{2} \right)}{\sqrt{2\pi \sigma}} \]

where \( \text{erf}(x) \) is defined as

\[ \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp \left( -\frac{y^2}{2} \right) dy \] (14)

according to [9].

The exact density (13) can easily be calculated, but it is not useful for recursive applications, because the derivation has been performed for the case that the primary stochastic uncertainties are Gaussian. Hence, the first of the two approximation approaches of the exact density \( f_z(z) \) proposed in this paper is given by a sum of Gaussian densities.

\section{The Approximate Density}

The key idea to finding an approximate solution for the probability density function is to look at the constraint (10) as rect \( \frac{x-y}{B} \) and to approximate this function by a weighted sum of Gaussian densities

\[ \text{rect} \left( \frac{x-y}{B} \right) \approx \sum_{i=-L}^{L} \frac{1}{\sqrt{2\pi c}} \exp \left\{ -\frac{1}{2} \left( \frac{x-y-m_{g,i}}{\sigma_g} \right)^2 \right\} \]

with \( m_{g,i} = i \frac{B}{2}, \sigma_g = c \frac{B}{2} \), so that the integral from \(-\infty \) to \( \infty \) over the sum gives \( 2B \). The free parameter \( c \in (0, \infty) \) may be obtained by a one-dimensional search to yield the best function approximation according to a given norm. A lot of manipulation gives

\[ f_z(z) \approx \sum_{i=-L}^{L} \frac{g_i \exp \left\{ -\frac{1}{2} \left( \frac{z-m_{g,i}}{\eta} \right)^2 \right\}}{\sqrt{2\pi \sigma_g} g_i} \]

with

\[ g_i = \exp \left\{ -\frac{1}{2} \left( \frac{m_x - m_y - m_{g,i}}{S_g + S_x + S_y} \right)^2 \right\} \] (15)

\[ z_i = (S_g [w_x m_x + w_y m_y] + S_x [w_x (m_{g,i} + m_y) + w_y m_y] + S_y [w_y (-m_{g,i} + m_x) + w_x m_x]) / (S_g + S_x + S_y) \] (16)

\[ \eta^2 = \left( S_g (w_x^2 S_x + w_y^2 S_y) + (w_x + w_y)^2 S_x S_y \right) / (S_g + S_x + S_y) \] (17)

where the shorthand notation \( S_g = \sigma_g^2 \) has been used. Using this approximate pdf, we can easily calculate the approximate mean and variance of \( z \). Furthermore, the exact mean and variance can be derived for \( L \to \infty \).

\section{Exact Mean and Variance}

Given (15), the approximate mean is calculated as

\[ m_z = \frac{\sum_{i=-L}^{L} g_i z_i}{\sum_{i=-L}^{L} g_i} \] (18)

The approximation of the variance is given by \( \sigma_z^2 = \text{E} \{ z^2 \} - m_z^2 \) with

\[ \text{E} \{ z^2 \} \approx \frac{\sum_{i=-L}^{L} g_i (\eta^2 + z_i^2)}{\sum_{i=-L}^{L} g_i} \] (19)

Analytic expressions for the exact mean and variance are calculated from (18), (19) for \( L \to \infty \), which implies \( S_g \to 0 \). The mean is then given by

\[ m_z = \frac{\int_{-B}^{B} \exp \left\{ -\frac{1}{2} \left( \frac{m_x - m_y - v}{S_x + S_y} \right)^2 \right\} \frac{1}{S_x + S_y} \left( S_x \left[ w_x (v + m_y) + w_y m_y \right] + S_y \left[ w_y (-v + m_x) + w_x m_x \right] \right) dv}{\int_{-B}^{B} \exp \left\{ -\frac{1}{2} \left( \frac{m_x - m_y - v}{S_x + S_y} \right)^2 \right\} dv} \]

which results in

\[ m_z = w_x m_x + w_y m_y - (w_x S_x - w_y S_y) G \] (20)

with

\[ g_i = \exp \left\{ -\frac{1}{2} \left( \frac{m_x - m_y - m_{g,i}}{S_g + S_x + S_y} \right)^2 \right\} \] (15)

\[ z_i = (S_g [w_x m_x + w_y m_y] + S_x [w_x (m_{g,i} + m_y) + w_y m_y] + S_y [w_y (-m_{g,i} + m_x) + w_x m_x]) / (S_g + S_x + S_y) \] (16)

\[ \eta^2 = \left( S_g (w_x^2 S_x + w_y^2 S_y) + (w_x + w_y)^2 S_x S_y \right) / (S_g + S_x + S_y) \] (17)
with

\[ G = \left( \exp \left\{ -\frac{1}{2} \frac{(B - m_x + m_y)^2}{S_x + S_y} \right\} ight) \\
- \exp \left\{ -\frac{1}{2} \frac{(B + m_x - m_y)^2}{S_x + S_y} \right\} \\
/ \left( 2\pi \right) \frac{1}{\sqrt{S_x + S_y}} \left( \text{erf} \left( B - m_x + m_y \right) \right) \\
+ \text{erf} \left( B + m_x - m_y \right) \right) \, . \quad (21) \]

The variance is given by

\[ S_z = \int_{v=-B}^{B} \exp \left\{ -\frac{1}{2} \frac{(m_x - m_y - v)^2}{S_x + S_y} \right\} \frac{1}{(S_x + S_y)^2} \]

\[ \frac{(w_x + w_y)^2 S_x S_y}{S_x + S_y} + \left( S_x [w_x (v + m_y) + w_y m_y] + S_y [w_y (-v + m_x) + w_x m_x] \right)^2 \]

\[ \int_{v=-B}^{B} \exp \left\{ -\frac{1}{2} \frac{(m_x - m_y - v)^2}{S_x + S_y} \right\} dv - m_z^2 \, , \]

which results in

\[ S_z = w_x^2 S_x + w_y^2 S_y - (w_x S_x - w_y S_y)^2 G^2 - \frac{(w_x S_x - w_y S_y)^2}{2\pi (S_x + S_y)^{3/2}} \]

\[ \left( B - m_x + m_y \right) \exp \left\{ -\frac{1}{2} \frac{(B - m_x + m_y)^2}{S_x + S_y} \right\} \]

\[ + \left( B + m_x - m_y \right) \exp \left\{ -\frac{1}{2} \frac{(B + m_x - m_y)^2}{S_x + S_y} \right\} \]

\[ \int \left( \text{erf} \left( B - m_x + m_y \right) + \text{erf} \left( B + m_x - m_y \right) \right) \, . \quad (22) \]

6 The New Estimator

Given two uncertain information sources according to (1), it has been shown in Sec. 4, that the fusion result is given as the sum of 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 derived in Sec. 3 is approached. This important result 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.

An even simpler estimator is obtained, when using the additional results derived in Sec. 5, i.e., the exact second-order description of the density of the interval midpoint. The estimator is then given by the interval

\[ \mathcal{Z} = \{ z : (z - \hat{z})^2 \leq E_z \} \, , \]

with midpoint \( \hat{z} \), that is uncertain in a stochastic sense. Mean \( m_z \) and variance \( S_z \) of \( \hat{z} \) are given by (20) and (22). The set theoretic uncertainty \( E_z \) is not a random variable and given by

\[ E_z = \left( \frac{0.5 - \lambda}{E_x} + \frac{0.5 + \lambda}{E_y} \right)^{-1} \]

for \( \lambda \in [-0.5, 0.5] \).

The new estimator contains the classical estimation concepts, i.e., the set theoretic estimator and the Kalman filter, as border cases. We first investigate the case of a vanishing set theoretic uncertainty.

6.1 Border Case: Stochastic Uncertainty Only

This case is equivalent to \( B = 0 \). Hence, we have to examine the expression for \( m_z \) from (20) and \( S_z \) from (22) for the limit \( B \to 0 \). By applying the rule of l'Hospital, we obtain for the mean

\[ m_z(B \to 0) = (w_x + w_y) \frac{S_x m_y + S_y m_x}{S_x + S_y} \, , \]

which can be simplified using \( w_x + w_y = 1 \) to yield

\[ m_z(B \to 0) = \frac{m_x}{S_x} + \frac{m_x}{S_y} \, , \]

which is exactly the mean of a Kalman filter estimate. When examining (22), we obtain

\[ S_z(B \to 0) = (w_x + w_y)^2 \frac{S_x S_y}{S_x + S_y} \, . \]

This is simplified by using \( w_x + w_y = 1 \) again to yield

\[ S_z(B \to 0) = \left( \frac{1}{S_x} + \frac{1}{S_y} \right)^{-1} \, , \]

which is the variance of a Kalman filter estimate.
6.2 Border Case: Set Theoretic Uncertainty Only
The second border case is equivalent to \( S_x = 0, S_y = 0 \). Now we have to examine the expression for \( m_z \) from (20) and \( S_z \) from (22) for the limit \( S_x \to 0, S_y \to 0 \). For the mean, we obtain
\[
m_z(S_x \to 0, S_y \to 0) = w_x m_x + w_y m_y.
\]
The limiting variance of the new estimator vanishes, i.e.,
\[
S_z(S_x \to 0, S_y \to 0) = 0
\]
as expected. Hence, for the border case of vanishing stochastic uncertainties, the new estimator converges to a set theoretic estimator according to (3) and (8).

7 Simulative Verification
Consider a vehicle equipped with a range sensor that measures the distance to a number of boxes, Fig. 1. The box positions are known within a given geometric tolerance, i.e.,
\[
i x_B = i \hat{x}_B + i \Delta x_B, \quad \text{with } |i \Delta x_B| \leq \frac{b_i}{2},
\]
where \( i \hat{x}_B \) denotes the unknown true value and \( i \Delta x_B \) is the unknown but bounded deviation of the nominal value \( i x_B \). The range sensor is corrupted by additive white Gaussian noise with zero mean and a variance \( \sigma_i^2 \) which depends on the surface characteristic of the box \( i \). The measurement equation is thus given by
\[
i x_B + D^k_i = x + i \Delta x_B + c^k_i, \quad \text{with } x = 200 \text{ is assumed.}
\]
The recursive estimation scheme is initialized to
\[
m(1) = i x_B + D^1_i, \quad E(1) = b_i^2, \quad S(1) = \sigma_i^2.
\]
This estimate is recursively updated using (4), (8), (20), and (22) with
\[
m_x = m(k - 1), \quad E_x = E(k - 1), \quad S_x = S(k - 1)
\]
and
\[
m_y = i x_B + D^k_i, \quad E_y = b_i^2, \quad S_y = \sigma_i^2.
\]
Table 1: Parameters of localization experiment.

<table>
<thead>
<tr>
<th>Box</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>True value ( i \hat{x}_B )</td>
<td>140</td>
<td>95</td>
<td>50</td>
<td>5</td>
</tr>
<tr>
<td>Nominal value ( i x_B )</td>
<td>120</td>
<td>80</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td>Bound ( b_i )</td>
<td>40</td>
<td>30</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>Standard deviation ( \sigma_i )</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

Thus, the estimate
\[
m(k) = m_z, \quad E(k) = E_z, \quad S(k) = S_z
\]
incorporates all measurements available up to time \( k \). The parameter \( \lambda \) in (4), (8) is chosen such that
\[
\sqrt{E(k)} + 3 \sqrt{S(k)} \text{ is minimized.}
\]
Fig. 2 depicts the expected value and the confidence interval using the proposed estimator. Sampling a specific box reduces the stochastic uncertainty. The initial deterministic uncertainty given by the interval \([180, 260] \) is reduced each time the box is changed. When changing from box 1 to box 2, the resulting intersection set without stochastic noise would be \([185, 245] \), which is eventually approached when sampling box 2 for an infinite number of times. Traversing from box 2 to box 3 gives the underlying set \([190, 230] \), from box 3 to box 4 it is given by \([195, 215] \). The realistic quantification may be exploited when attempting to navigate a mobile robot through narrow openings.

It is common practice to approximate systematic errors by stochastic noise. To demonstrate the negative effect of doing so, the deterministic error of the box positions will now be modeled as additional uncorrelated noise with the standard deviation set to \( b_i \). For this noise description, estimates can be obtained by Kalman filtering. Fig. 3 depicts the corresponding expected value and the confidence interval. The resulting estimate is obviously overoptimistic, because the
confidence interval does not always include the true vehicle position.

8 Conclusions
A vast class of estimation problems can be attacked as a mixed noise problem, i.e., the arising uncertainties can be modeled as being additively composed of both 1) noise with known bounds and 2) noise with known statistics. For this class, a new estimator has been derived, which combines set theoretic and stochastic estimation in a rigorous manner. Hence, it provides solution sets that are uncertain in a statistical sense.

The proposed estimator is efficient and, hence, well-suited for practical applications. Our research also includes the vector case. However, this presentation has been restricted to the scalar case for ease of explanation.

References