Combined Set-Theoretic and Stochastic Estimation: A Comparison of the SSI and the CS Filter

Vesa Klumpp, Benjamin Noack, Marcus Baum, and **Uwe D. Hanebeck**

Intelligent Sensor-Actuator-Systems Laboratory (ISAS),
Institute for Anthropomatics,
Karlsruhe Institute of Technology (KIT), Germany.
{vesa.klumpp,benjamin.noack,marcus.baum}@kit.edu, uwe.hanebeck@ieee.org

Abstract − *In estimation theory, mainly set-theoretic or stochastic uncertainty is considered. In some cases, especially when some statistics of a distribution are not known or additional stochastic information is used in a set-theoretic estimator, both types of uncertainty have to be considered. In this paper, two estimators that cope with combined stoachastic and set-theoretic uncertainty are compared, namely the Set-theoretic and Statistical Information filter, which represents the uncertainty by means of random sets, and the Credal State filter, in which the state information is given by sets of probability density functions. The different uncertainty assessment in both estimators leads to different estimation results, even when the prior information and the measurement and system models are equal. This paper explains these differences and states directions, when which estimator should be applied to a given estimation problem.*

Keywords: Filtering, state estimation, random sets, sets of probability densities.

1 Introduction

In state estimation problems, an internal system state is deduced from measurements by means of a measurement model, which relates these measurements to the system state. Various sources of uncertainty have to be considered, such as noise, which disturbs the measurements, and model errors, which lead to wrong state estimates.

Two common approaches to model uncertainty are stochastic, i.e., the uncertainty is described by means of a random variable, and set-theoretic, i.e., the uncertainty about a variable is given by a set of all possible values. For each single type of uncertainty models, optimal estimation approaches exist. In the case of stochastic uncertainty, a Bayesian estimator [1] can be applied. For set-theoretic uncertainty, set-theoretic estimators [2] lead to the desired result.

In a Bayesian state estimator, stochastic uncertainty is described by means of a random variable, for which the exact statistics must be known. Under certain conditions, the stochastic uncertainty vanishes by repeated measurements, until the true state is known. Set-theoretic uncertainty is represented by means of a set, which is often used to model a bounded error with an unknown distribution. It is often used to describe systematic uncertainty, like biases, which cannot be eliminated by repeated measurements.

Unfortunately, it is not always suitable to employ a model consisting of one single type of uncertainty only. Sometimes, certain parameters such as the exact statistics of the measurement noise or parameters of a system function are not known completely. This lack of knowledge, or ignorance, is modeled by a set of all possible values. When combining this lack of knowledge with other stochastic information, one has to deal with combined – set-theoretic and stochastic – uncertainties.

Aside from random sets [3, 4, 5, 6] and sets of densities [7, 8, 9], other approaches for dealing with combined uncertainty can be found in literature. The theory of imprecise probabilities [10, 11] deals with combined uncertainty by means of upper and lower expectations. In [12], an approach for the selection of an optimal Kalman gain under the consideration of stochastic and set-theoretic uncertainty is pursued and in [13] stochastic uncertainty is incorporated into set-theoretic estimation.

In this paper, we provide a unifying formalization and description of approaches for combined set-theoretic and stochastic uncertainty models in state estimation. The focus lies on comparing the SSI filter, which works with random sets, with the CS filter, which employs sets of densities. For this purpose, we state the basic estimation concepts of both filters in a general setting. Both filters are demonstrated by means of a one-dimensional fusion example, which directly allows to compare the estimation results. We interpret the estimation results of both filters in order to provide deeper insights. The different behavior of both filters is also discussed by means of several special cases. This discussion allows for deciding when to use which approach.

For the notation throughout this paper, vectors $\underline{x}$ are underlined, random variables $\boldsymbol{x}$ are bold face, and variables with combined set-theoretic and stochastic uncertainty are denoted by x.

2 Information Fusion for Single Type of Uncertainty

2.1 Set-Theoretic Estimation

For systems corrupted with unknown but bounded errors as in the system model

$$\underline{x}_{k+1} = \underline{a}_k\left(\underline{x}_k\right) + \underline{\eta}_k \qquad (1)$$

and measurement model

$$\underline{\hat{y}}_k = \underline{h}_k\left(\underline{x}_k\right) + \underline{\varepsilon}_k \ , \qquad (2)$$

depending on nonlinear system and measurement functions $\underline{a}_k$ and $\underline{h}_k$ and with bounded errors $\underline{\eta}_k \in \mathcal{M}_k^\eta$ and $\underline{\varepsilon}_k \in \mathcal{M}_k^\epsilon$, a set-theoretic estimator is suitable to estimate the unknown state $\underline{x}_k \in \mathbb{R}^N$ given specific measurements $\underline{\hat{y}}_k$.

In set-theoretic state estimation [14], the uncertainty about the true unknown state $\underline{x}_k$ at time step k is expressed by a solution set Δ_k^e such that $\underline{x}_k$ is known to lie in Δ_k^e, i.e., $\underline{x}_k \in \Delta_k^e$. For each measurement $\underline{\hat{y}}_k$, a new solution set

$$\Delta_{\underline{\hat{y}}_k} := \left\{ \underline{x}_k \mid \underline{\hat{y}}_k = h_k(\underline{x}_k) + \underline{\varepsilon}_k \text{ for } \underline{\varepsilon}_k \in \mathcal{M}_k^\epsilon \right\}$$

can be constructed.

Hence, the updated solution set becomes $\Delta_k^e = \Delta_k^p \cap \Delta_{\underline{\hat{y}}_k}$, where Δ_k^p denotes the prediction to time step k. The predicted solution set Δ_{k+1}^p is the result of propagating Δ_k^e element-wise through the system model (1), i.e.,

$$
\begin{aligned}
\Delta_{k+1}^p \quad := \quad & a_k(\Delta_k^e) + \underline{\eta}_k \\
= \quad & \left\{ \underline{x}_{k+1} \mid \underline{x}_{k+1} = a_k(\underline{x}_k) + \underline{\eta}_k \ , \right. \\
& \left. \text{with } \underline{x}_k \in \Delta_k^e \text{ and } \underline{\eta}_k \in \mathcal{M}_k^\eta \right\} \ .
\end{aligned}
$$

The exact computation of the sets Δ_k^p and Δ_k^e can be difficult and time consuming. Hence, in order to allow a fast recursive computation, the sets Δ_k^p and Δ_k^e are usually outer-bounded with supersets that can be described with a (fixed) number of parameters like ellipsoids [15] and polytopes [16].

2.2 Bayesian Estimation

In Bayesian state estimation, the uncertainty about an unknown quantity, i.e, the system state, is expressed by a probability density function. The relationship between the state and the measurement may be specified by a stochastic measurement model

$$\underline{\hat{y}}_k = \underline{h}_k\left(\boldsymbol{x}_k\right) + \underline{\boldsymbol{w}}_k \ , \qquad (3)$$

where $\underline{\boldsymbol{x}}_k$ is the state variable at time step k, $\underline{\hat{y}}_k$ is a specific measurement, and $\underline{\boldsymbol{w}}_k$ is a random variable modeling stochastic measurement noise. The temporal evolution of the state estimate may be specified by means of a stochastic dynamic model

$$\underline{\boldsymbol{x}}_{k+1} = \underline{a}_k\left(\underline{\boldsymbol{x}}_k\right) + \underline{\boldsymbol{w}}_k \ , \qquad (4)$$

where $\underline{\boldsymbol{w}}_k$ is a random variable modeling system noise.[1]

Given the probability density function $f^e(\underline{x}_k)$ for time step k, the predicted probability density $f^p(\underline{x}_{k+1})$ for time step $k+1$ can be computed according to the Chapman-Kolomogorov equation

$$f^p(\underline{x}_{k+1}) = \int_{\mathbb{R}^N} f^T(\underline{x}_{k+1}|\underline{x}_k) \cdot f^e(\underline{x}_k) \, \mathrm{d}\underline{x}_k \ , \qquad (5)$$

with a transition density f^T resulting from the system model (4).

The measurement $\underline{\hat{y}}_k$ is fused with $f^p(\underline{x}_k)$ following Bayes' theorem

$$f^e(\underline{x}_k) = \frac{f^p(\underline{x}_k) \cdot f^L(\underline{\hat{y}}_k|\underline{x}_k)}{\int_{\mathbb{R}^N} f^p(\underline{\xi}) \cdot f^L(\underline{\hat{y}}_k|\underline{\xi}) \, \mathrm{d}\underline{\xi}} \qquad (6)$$

with the likelihood f^L resulting from the measurement model (3).

In general, the resulting probability densities cannot be computed in closed form. However, for linear systems and white Gaussian noise, the well-known Kalman filter provides the optimal solution [1]. For nonlinear systems, approximation techniques like the EKF, UKF, or particle filters [1] have to be employed.

3 Models Considering Combined Uncertainty

In several applications, the uncertainties of a dynamic system are composed of both stochastic noise with known distribution and bounded noise with unknown distribution. In the case, when the uncertainties can be modeled by means of additive terms, such a system can be described by a special measurement model

$$\underline{\hat{y}}_k = h_k(\mathrm{x}_k) + \mathrm{v}_k, \text{ with } \mathrm{v}_k = \underline{\varepsilon}_k + \underline{\boldsymbol{v}}_k \ , \qquad (7)$$

where $\underline{\boldsymbol{v}}_k$ is a random variable modeling stochastic measurement noise and $\underline{\varepsilon}_k$ is an unknown but bounded error with $\underline{\varepsilon}_k \in \mathcal{M}_k^\epsilon$. A system model that models both types of uncertainty is given by

$$\mathrm{x}_{k+1} = a_k(\mathrm{x}_k) + \mathrm{w}_k, \text{ with } \mathrm{w}_k = \underline{\eta}_k + \underline{\boldsymbol{w}}_k \ , \qquad (8)$$

where $\underline{\boldsymbol{w}}_k$ is a random variable modeling stochastic system noise and $\underline{\eta}_k$ is an unknown but bounded error with $\underline{\eta}_k \in \mathcal{M}_k^\eta$. We assume that the errors $\underline{\boldsymbol{v}}_k$, $\underline{\varepsilon}_k$, $\underline{\boldsymbol{w}}_k$, and $\underline{\eta}_k$ are mutually independent.

---

[1]For the sake of simplicity, a system input $\underline{\hat{u}}_k$ has been omitted.

In order to derive a state estimator for the above system, the question arises how to interpret the sum of a set and a random variable given by

$$\varepsilon_k + \underline{\boldsymbol{v}}_k \ (\text{or } \underline{\eta}_k + \underline{\boldsymbol{w}}_k) \ .$$

Actually, two possibilities seem obvious: First, it can be considered as a set of translated random variables, i.e., the set of densities $\{f_v(\underline{x}_k - \varepsilon_k) \,|\, \varepsilon_k \in \mathcal{M}_k^\epsilon\}$. Second, it can be interpreted as a set that is translated by a random variable $\mathcal{M}_k^\epsilon + \underline{\boldsymbol{v}}_k$ with $\underline{\boldsymbol{v}}_k \sim f_v(\underline{x}_k)$.

While at this point, these two interpretations can easily be converted into each other, we will see that they lead to two different fusion rules. These two different approaches, namely the SSI filter and CS filter, for dealing with combined uncertainty models are introduced and discussed in the rest of this paper.

4   The SSI Filter

One possible method to deal with models consisting of set-theoretic and stochastic uncertainties is the so-called *Set-theoretic and Statistical Information filter (SSI filter)*, which was introduced in [3] and extended in [17, 18]. The SSI filter can be seen as a combination of set-theoretic and stochastic estimation and provides a continuous transition between these two concepts. The uncertainty about the unknown state is captured with a solution *set* that is uncertain in a stochastic sense. In the following, we present a general formulation of the SSI filter, which is independent of a particular implementation.

Consider a measurement $\hat{y}_k$ related to the unknown state $\underline{x}_k$ according to (7). The basic idea is to construct a solution set, which consists of all $\underline{x}_k$ that are compatible with $\hat{y}_k$ in the following manner:

$$\boldsymbol{\Delta}_k^y := \left\{ \underline{x}_k \,|\, \hat{\underline{y}}_k - \underline{\boldsymbol{v}}_k = h_k(\underline{x}_k) + \underline{\varepsilon}_k \text{ with } \underline{\varepsilon}_k \in \mathcal{M}_k^\epsilon \right\} \ .$$

Since $\underline{\boldsymbol{v}}_k$ is a random variable, the set $\boldsymbol{\Delta}_k^y$ is random as well. It can be interpreted as a random set in the classical sense [5, 6], which is a random experiment whose outcomes are sets. Formally, a random set is a set-valued random variable. The set $\boldsymbol{\Delta}_k^y$ is also called *measured random solution set*, since it results from a particular measurement. It can be seen as an analogon of the likelihood function in Bayesian state estimation. The *measured random solution set* is fused at each time step with a *predicted random solution set* $\boldsymbol{\Delta}_k^p$. Since the true value of $\underline{x}_k$ is known to be consistent with both $\boldsymbol{\Delta}_k^p$ and $\boldsymbol{\Delta}_k^y$ these random sets are fused by intersection

$$\boldsymbol{\Delta}_k^e := \boldsymbol{\Delta}_k^y \cap \boldsymbol{\Delta}_k^p \ .$$

This intersection has to be conditioned on non-emptiness, since the true value of $\underline{x}_k$ is known to be an element of both sets.

The prediction step of the SSI filter propagates the *random solution set* through the system model (8) in the following manner:

$$\begin{aligned}
\boldsymbol{\Delta}_k^p \ &:= \ a_k(\boldsymbol{\Delta}_{k-1}^e) + \underline{\eta}_k + \underline{\boldsymbol{w}}_k \\
&= \ \{\underline{x}_{k+1} \,|\, \underline{x}_{k+1} = a_k(\underline{x}_k) + \underline{\eta}_k + \underline{\boldsymbol{w}}_k \\
&\qquad \text{and } \underline{x}_k \in \boldsymbol{\Delta}_{k-1}^e \text{ and } \underline{\eta}_k \in \mathcal{M}_k^\eta\} \ .
\end{aligned}$$

It is worth mentioning that the SSI filter turns into a pure set-theoretic filter if the stochastic noise vanishes. Furthermore, if the bounded noise error is zero, a stochastic filter is obtained. This results from the fact that a random variable can be considered as a singleton random set. From this point of view, the Bayesian fusion of two probability density functions coincides with random set intersection conditioned on non-emptiness.

In practicable applications, a parametric representation of random solution sets is chosen. The SSI filter has been formulated for the scalar case with random intervals [3] and for the multi-dimensional case with random ellipsoids [17, 18]. Since the intersection of two ellipsoids is not an ellipsoid anymore, proper outer-bounding techniques have to be employed in [17, 18].

4.1   Special Case: Intervals

For the purpose of illustration, we describe a particular SSI filter for simplified linear one-dimensional models given by the measurement equation

$$\hat{y}_k = \mathbb{x}_k + \epsilon_k + \boldsymbol{v}_k \ , \tag{9}$$

and system equation

$$\mathbb{x}_{k+1} = \mathbb{x}_k + \eta_k + \boldsymbol{w}_k \ , \tag{10}$$

where the error bounds are $\eta_k \in [-b_\eta, b_\eta]$ and $\epsilon_k \in [-b_\epsilon, b_\epsilon]$. Similar to [3], the particular SSI filter for the above models employs random intervals in order to describe the random solution sets. An interval is represented by its center m and width b, i.e.,

$$[m - b, m + b] \ .$$

The random solution set $\boldsymbol{\Delta}_k^e$ at time step k is specified by a one random variable for the center $\boldsymbol{m}_k^e$ and one random variable $\boldsymbol{b}_k^e$ for the width, i.e.,

$$\boldsymbol{\Delta}_k^e := [\boldsymbol{m}_k^e - \boldsymbol{b}_k^e, \boldsymbol{m}_k^e + \boldsymbol{b}_k^e] \ .$$

In the same manner, the *predicted random solution set* turns out to be

$$\boldsymbol{\Delta}_k^p := [\boldsymbol{m}_k^p - \boldsymbol{b}_k^p, \boldsymbol{m}_k^p + \boldsymbol{b}_k^p] \ .$$

The random measurement solution set resulting from measurement $\hat{y}_k$ is given by

$$\boldsymbol{\Delta}_k^y := [\boldsymbol{m}_k^y - b_\epsilon, \boldsymbol{m}_k^y + b_\epsilon] \ ,$$

where $\boldsymbol{m}_k^y = \hat{y}_k - \underline{\boldsymbol{v}}_k$. Note that this random interval has a deterministic width.

The intersection of the two random solution sets $\boldsymbol{\Delta}_k^e$ and $\boldsymbol{\Delta}_k^p$ can be expressed in terms of the parameters specifying the random intervals. Let $[m_k^e, b_k^e]^T := \mathcal{G}(m_k^p, b_k^p, m_k^m, b_k)$ denote the function which maps the parameters of the predicted solution set and the measurement solution set to the parameters of their intersection interval, i.e.,

$$\mathcal{G}(m_k^p, b_k^p, m_k^m, b_k) := \begin{bmatrix} u_k + 0.5(o_k - u_k) \\ (o_k - u_k) \end{bmatrix} \ ,$$

where $u_k := \max(m_k^p - b_k^p, m_k^e - b_k^e)$ and $o_k := \max(m_k^p + b_k^p, m_k^y + b_\epsilon)$.

The probability density for the parameters of the updated interval is given by

$$f(m_k^e, b_k^e \mid b_k^e > 0) \ , \tag{11}$$

where the random vector $[\boldsymbol{m}_k^e, \boldsymbol{b}_k^e]^T$ is defined as

$$[\boldsymbol{m}_k^e, \boldsymbol{b}_k^e]^T = \mathcal{G}(\boldsymbol{m}_k^p, \boldsymbol{b}_k^p, \boldsymbol{m}_k^m, \boldsymbol{b}_k^m) \ ,$$

which is a stochastic forward mapping that computes the interval intersection. The condition $b_k^e > 0$ ensures non-emptiness of the interval intersection. Note that the resulting probability densities are in general not Gaussian, even if the stochastic noise is assumed to be Gaussian. However, it is possible to approximate the densities with a Gaussian distribution by means of analytic moment matching [3].

The result of the above SSI filter is a probability distribution over the bounds of an interval. It states how probable it is that the true value is an element of a particular interval. A random interval, i.e., the random solution set, can be considered as a random constraint on the true value.

5 The CS Filter

The second method for dealing with models consisting of set-theoretic and stochastic uncertainties is given by the processing of sets of probability density functions in the so called *credal state filter* (CS filter) [7, 8]. Within this approach, the uncertain state information is given by means of a set of densities, i.e, the set-theoretic uncertainty is exploited to address a certain density within the set.

An important property of this representation form is that no assessment between the densities of the set are made. Every density can be considered as an own independent state estimate with stochastic uncertainty only, but no decision for a specific density is made.

The primary uncertainty in this approach is stochastic. Every density in the set is processed independently in a stochastic estimator leading to multiple independent resulting densities, which form a new set.

In the prediction step, the set of estimated densities $\boldsymbol{\mathcal{F}}_k^e$ is propagated according to the system model (8),

which depends on the set-valued parameter $\underline{\eta}_k$. This value parameterizes the set of transition densities

$$\boldsymbol{\mathcal{F}}_k^T = \left\{ f_k^t(\underline{x}_{k+1} | \underline{x}_k - \underline{\eta}_k) \, | \, \underline{\eta}_k \in \mathcal{M}_k^\eta \right\} \ .$$

Element-wise processing of the set of estimated densities with the set of transition densities according to the Chapman-Kolmogorov equation (5) leads to the set of predicted densities

$$\boldsymbol{\mathcal{F}}_{k+1}^p = \left\{ f^p(\underline{x}_{k+1}) = \int_{\mathbb{R}^N} f^T(\underline{x}_{k+1} | \underline{x}_k - \underline{\eta}_k) \cdot f^e(\underline{x}) \, \mathrm{d}\underline{x}_k, \right. \\ \left. f^T \in \boldsymbol{\mathcal{F}}_k^T, f^e \in \boldsymbol{\mathcal{F}}_k^e \right\} \ .$$

One important property of this element-wise processing is that the resulting set is now parameterized by the parameters of both sets.

In the filter step, a given measurement $\underline{\hat{y}}_k$ is incorporated into the set of predicted densities $\boldsymbol{\mathcal{F}}_k^p$ according to the measurement equation (7). A set of likelihoods

$$\boldsymbol{\mathcal{F}}_k^L = \left\{ f_k^L(\underline{\hat{y}}_k | \underline{x}_k - \underline{\varepsilon}_k) \, | \, \underline{\varepsilon}_k \in \mathcal{M}_k^\epsilon \right\} \ ,$$

which is parameterized by the set-theoretic parameter $\underline{\eta}_k$, is used to express the information from this measurement. Element-wise processing of the elements of both sets with the Bayes equation (6) leads to the set of posterior densities

$$\boldsymbol{\mathcal{F}}_k^e = \left\{ f^e(\underline{x}_k) = \frac{f^p(\underline{x}_k) \cdot f^L(\underline{\hat{y}}_k | \underline{x}_k - \underline{\varepsilon}_k)}{\int_{\mathbb{R}^N} f^p(\underline{\xi}) \cdot f^L(\underline{\hat{y}}_k | \underline{\xi} - \underline{\varepsilon}_k) \, \mathrm{d}\underline{\xi}}, \right. \\ \left. f^p \in \boldsymbol{\mathcal{F}}_k^p, f^L \in \boldsymbol{\mathcal{F}}_k^L \right\} \ .$$

Again, the number of parameters for describing the posterior set of densities has increased.

5.1 Special Case: Intervals

For comparison with the SSI filter, the special case of estimation with sets of densities for one-dimensional linear systems is depicted. It is a special case of the Kalman filter for ellipsoidal sets [8], where sets of densities are parameterized by a displacement term, which lies within an ellipsoidal set.

For the one-dimensional case, only the interval of means and the variance is needed to represent the estimated state.

For the prediction step according to the linear model (10), the set of transition densities is given by

$$\boldsymbol{\mathcal{F}}_k^T = \{ f_w(x_{k+1} - x_k - \eta_k), \eta_k \in \mathcal{M}_k^\eta \} \ .$$

The set of predicted densities can be characterized by its set of means $\mathcal{X}_{k+1}^p$ and the variance $\left(\sigma_{k+1}^p\right)^2$ with

$$\mathcal{X}_{k+1}^p = \mathcal{X}_k^e \oplus \mathcal{M}_k^\eta \ ,$$

where the element-wise addition of both intervals is represented by the Minkowski sum $\oplus$. The variance is independent of the set-valued uncertainty η_k and thus

$$\left(\sigma_{k+1}^p\right)^2 = \left(\sigma_k^e\right)^2 + \left(\sigma_k^w\right)^2 \ ,$$

where $\left(\sigma_k^w\right)^2$ denotes the variance of $\boldsymbol{w}_k$. In contrast to the processing of general models, the parameterization of this linear model can be reduced to an interval for the mean and a single variance. This is due to the lucky fact that the addition of two intervals according to the Minkowski sum is an interval again.

For the filter step, the set of likelihoods derived from the measurement model (9) is given by

$$\boldsymbol{\mathcal{F}}_k^L = \{f_v(\hat{y}_k - x_k - \epsilon_k)\,|\,\epsilon_k \in \mathcal{M}_k^\epsilon\} \ .$$

The set of estimated means is calculated employing the Kalman filter equations for the set of estimated means

$$\mathcal{X}_k^e = \left(\left(\sigma_k^p\right)^2 \mathcal{X}_k^p \oplus \left(\sigma_k^v\right)^2 \left(\hat{y}_k + \mathcal{M}_k^\epsilon\right)\right) \big/ \left(\left(\sigma_k^p\right)^2 + \left(\sigma_k^v\right)^2\right)$$

and the estimated variance

$$\left(\sigma_k^e\right)^2 = \left(\sigma_k^v\right)^2 \left(\sigma_k^p\right)^2 \big/ \left(\left(\sigma_k^v\right)^2 + \left(\sigma_k^p\right)^2\right) \ ,$$

with the variance $\left(\sigma_k^v\right)^2$ of the noise term $\boldsymbol{v}_k$. Again, the mean of estimated densities can be represented by means of a single interval.

6 Comparison

When comparing the results of the SSI and the CS filter, the problem of comparability of random sets and sets of densities arises. Different solutions have been proposed in literature.

One solution is to consider every measurable selection [5, 19, 20] from a random set, i.e., the set of densities, which is compatible to the given random set. A measurable selection is a random variable that is an element of the random set with probability one.

A simplified approach for linear one-dimensional models is the selection of extremal lower and upper elements in the set. This results in a lower extremal density and an upper extremal density. The lower extremal density is obtained by the selection of the lowest element in the corresponding set ϵ_k, which is characterized by the random variable $\boldsymbol{m}_k - \boldsymbol{b}_k$. The higher extremal random variable is given by $\boldsymbol{m}_k + \boldsymbol{b}_k$. For comparison, a center density can be stated, which is distributed according to $\boldsymbol{m}_k$.

Another approach is the comparison of confidence sets, which can be stated for a given probability, i.e., confidence level. For random sets [17, 18], the confidence set is the union of the most probable sets. For sets of densities, the confidence set is given by the union of the confidence sets for every density.

When considering the processing of combined uncertainty in the SSI filter and in the CS filter alone, both approaches are rational. Unfortunately, the fusion results can differ, even when the original stochastic and set-theoretic uncertainties are equal.

Example: In order to visualize the different results of the fusion of two random sets and the fusion of two sets of densities, a simple example is stated here. The measurement equation is given by

$$\hat{y}_k = \mathrm{x}_k + \epsilon_k + \boldsymbol{v}_k \ ,$$

with set-theoretic uncertainty $\eta_k \in [-0.5, 0.5]$ and stochastic uncertainty $\boldsymbol{v}_k \sim \mathcal{N}(x - 0, 1)$. For two time steps $k = 1, 2$, the measurements $\hat{y}_1 = 0$, $\hat{y}_2 = 2.5$ were received. Because no information about the prior state estimate at time $k = 0$ is available, the estimation results after the first measurement are set to the corresponding likelihoods of both filters. For the SSI filter, the measured random solution sets are described by the density functions $f^y(m_1, b_1)$ and $f^y(m_2, b_2)$ for both measurements, which are displayed in Fig. 1. The set of likelihoods $\mathcal{F}_1^L$ and $\mathcal{F}_2^L$, as used in the CS filter, are visualized in Fig. 2, whereas only the likelihoods for $\underline{\varepsilon}_k \in \{-0.5, 0, 0.5\}$ are shown. The result of both filters is visualized in Fig. 3. The posterior density $f^e(m_2, b_2)$ of the parameters of the random set (top) is converted to lower extremal, upper extremal, and center densities (middle). The resulting set of densities of the CS filter is displayed at the bottom. It can be seen that the densities resulting from the SSI filter are different from the CS filter. For the SSI filter, the means are closer to each other and the variances differ. For the CS filter, the set of means is wider, contains the means obtained from the SSI filter, and has equal variances for every density.

Explanation: For both approaches, the assessment of uncertainty differs. The set-theoretic uncertainty is treated in different ways. This can be seen when fusing a single set with a probability density. In the SSI filter, the set describes a constraint to the density, since the true element lies in this set. This differs from the CS filter, where the set-theoretic uncertainty is treated as ignorance. Each single element of the set describes a density function (i.e., a Dirac impulse).

The SSI filter computes a probability distribution over intervals. It gives the probability that the true state is an element of a particular interval. The set-theoretic uncertainty directly restricts the possible values. Hence, the SSI filter calculates a probability density for the points that both random sets have in common, since it is based on the rule of intersection.

In the CS filter, sets are used to describe additional uncertainty by means of several possible stochastic state estimates. In the general case, the set-theoretic uncertainty may be given by an arbitrary parameterization of the densities, which is not directly related to the state space. When both information imply combined uncertainty, the CS filter results in a worst-case estimate, because of the element-wise processing of all single densities within both sets. For sets of densities, the translation term $\underline{\eta}_k$ in the measurement model (7)
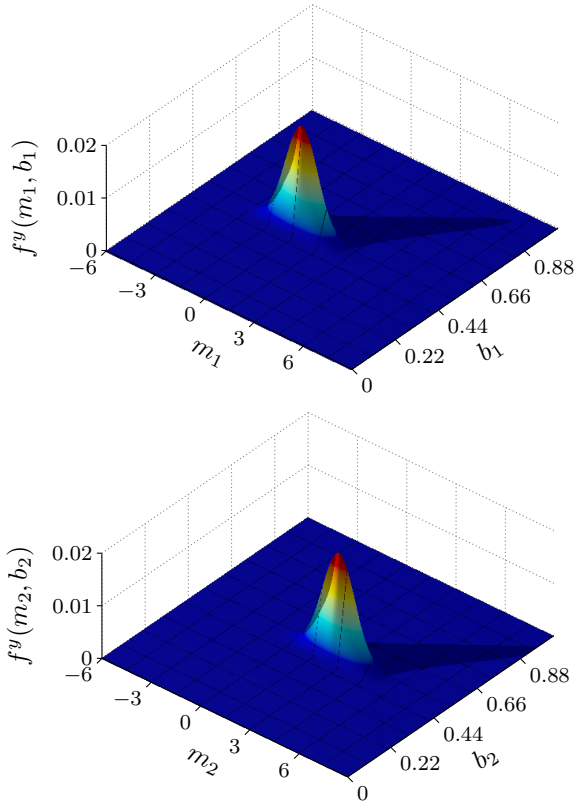
Figure 1: SSI filter: Densities $f^y(m_1, b_1)$ and $f^y(m_2, b_2)$ for the measured random sets at time step $k = 1, 2$.
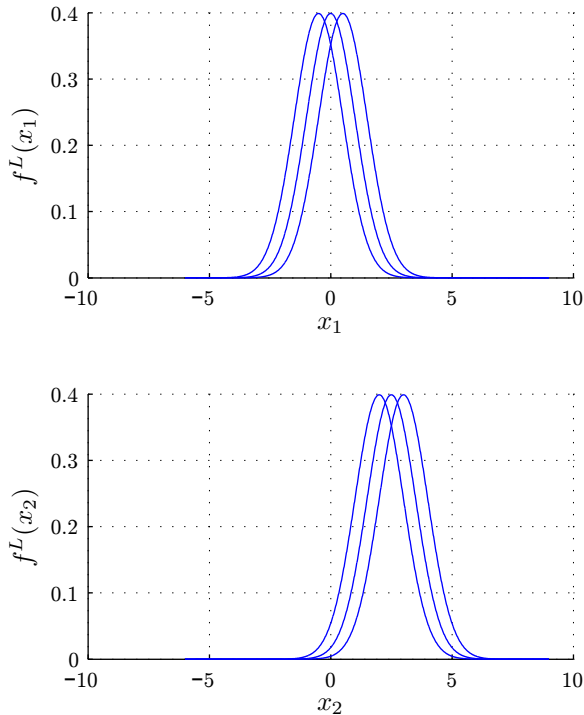

Figure 2: CS filter: Set of likelihoods $\boldsymbol{\mathcal{F}}_1^L$ and $\boldsymbol{\mathcal{F}}_2^L$ for the measurements at time steps $k = 1, 2$.
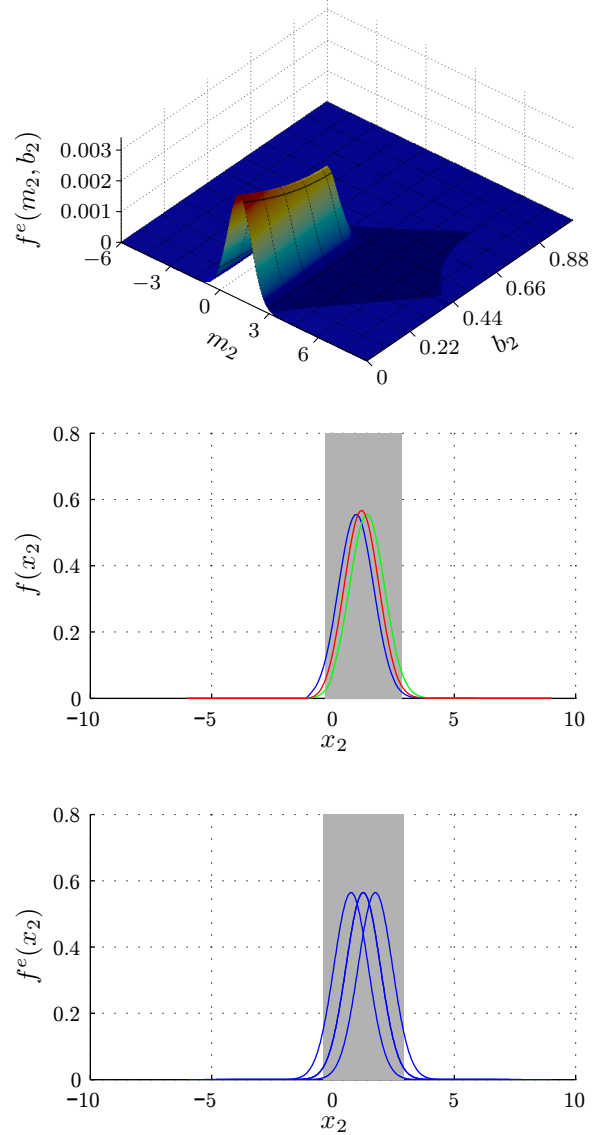
Figure 3: Fusion results for the SSI filter and the CS filter. Top: Density $f^e(m_1, b_1)$ over the parameters of the resulting random set. Middle: Extremal densities of the resulting random set. Bottom: Resulting set of estimated densities $\boldsymbol{\mathcal{F}}_2^e$. The 0.9 confidence sets are visualized by the shaded regions.

parameterizes the densities. By the generation of the set of densities, the semantics about this displacement are ignored, because the uncertainty leading to this set of densities is regarded as lack of knowledge about the true parameter. When fusing a set of densities with a set of likelihoods, the element-wise processing leads to a consistent, but conservative solution, where the true density is assumed to lie within the resulting set.

Special Cases

In order to demonstrate the differences of the two approaches for dealing with combined set-theoretic and

stochastic uncertainties, several special cases are discussed in the following. For this purpose, two pieces of information $\mathbb{x}_1$ and $\mathbb{x}_2$ about the same quantity are fused according to the SSI filter and CS filter update step. The result of the fusion is denoted with $\mathbb{x}^e$. The considered special cases differ in the type of uncertainty (set-theoretic or stochastic) with which the information is described.

Case 1: Stochastic/stochastic uncertainty If both pieces of information are described by means of a classical probability densities and no set-theoretic uncertainties are present, the fusion rule of both filters turns into standard Bayesian fusion of two probability densities.

Case 2: Set-theoretic/set-theoretic uncertainty If both information are described by means of a set, i.e., Δ_1 and Δ_2, the SSI filter becomes a classic set-theoretic estimator. Hence, $\mathbb{x}^e$ is the set intersection $\Delta_1 \cap \Delta_2$.

The CS filter interprets Δ_1 (and Δ_2) as a set of Dirac densities $\boldsymbol{\mathcal{D}}_1 = \{\delta(x - x_1) \,|\, x_1 \in \Delta_1\}$ (and $\boldsymbol{\mathcal{D}}_2 = \{\delta(x - x_2) \,|\, x_2 \in \Delta_2\}$). Each Dirac density in $\boldsymbol{\mathcal{D}}_1$ is fused with each Dirac density in $\boldsymbol{\mathcal{D}}_2$ according to Bayes' rule. Hence, the fusion result is again a set of densities, which consists of all densities with support $\Delta_1 \cap \Delta_2$.

Case 3: Set-theoretic/stochastic uncertainty If $\mathbb{x}_1$ is a classical set Δ_1 (without stochastic uncertainties) and $\mathbb{x}_2$ is a classic probability density $f_2(x)$, the following results for $\mathbb{x}^e$ are obtained:

- SSI filter
  The set Δ_1 serves as a constraint on the probability density $f_2(x)$. Hence, the fusion result is again a probability density. This density results from truncating $f_2(x)$ for $x \notin \Delta_1$.

- CS filter
  The fusion result is a set of densities. The CS filter interprets Δ_1 as a set of Dirac densities $\boldsymbol{\mathcal{D}}_1 = \{\delta(x - x_1) \,|\, x_1 \in \Delta_1\}$. Each of these Dirac densities is fused with $f_2(x)$ in order to obtain the set of densities $\mathbb{x}^e$. This set of densities consists of all densities that have the support Δ_1.

Case 4: Combined/stochastic uncertainty Now, the first piece of information $\mathbb{x}_1$ is described with both set-theoretic and stochastic uncertainty, and $\mathbb{x}_2$ is a probability density $f_2(x)$.

- SSI filter
  In this case, $\mathbb{x}_1$ is a a random set $\boldsymbol{\Delta}_1$ that serves as a soft constraint (in a stochastic sense) on the probability density $f_2(x)$. Hence, the fusion result is again a probability density.

- CS filter
  If $\mathbb{x}_1$ is interpreted as a set of densities $\boldsymbol{\mathcal{F}}_1$, the fusion result is in general a a set of densities again, as each probability density in $\boldsymbol{\mathcal{F}}_1$ is fused with $f_2(x)$.

Case 5: Combined/set-theoretic uncertainty In this case, one quantity is corrupted with both set-theoretic and stochastic uncertainty, and the second quantity is described by means of a set Δ_2.

- SSI filter
  If $\mathbb{x}_1$ is described with a random set $\boldsymbol{\Delta}_1$, the fusion result is again a random set, namely $\boldsymbol{\Delta}_1 \cap \Delta_2$.

- CS filter
  Again, Δ_2 is interpreted as a set of Dirac densities $\boldsymbol{\mathcal{D}}_2 = \{\delta(x - x_2) \,|\, x_2 \in \Delta_2\}$ and the uncertainty of $\mathbb{x}_1$ is described with a set of densities $\boldsymbol{\mathcal{F}}_1$. Then, each Dirac density in $\boldsymbol{\mathcal{D}}_2$ is fused with each density in $\boldsymbol{\mathcal{F}}_1$. Hence, the fusion result is again a set of densities, which consists of all densities that have the support Δ_2.

Hands-On Guide for Filter Selection

The assumptions made within the SSI and CS filters have to be considered for the choice of a proper estimator for a given estimation problem, especially because the results differ. The SSI filter returns the most probable state estimate that fits best to the received measurements. As Special Case 3 turned out, the set-theoretic part also serves as a constraint in state space and can therefore provide additional information. The CS filter gives the most conservative result, by assuming ignorance about the set-theoretic uncertainty. This estimator can be used to derive conservative bounds for worst-case estimates, that are important in the task of avoiding certain states, such as collisions.

7   Conclusions

In this contribution, an overview of combined set-theoretic and stochastic estimation concepts has been given. In particular, two special approaches, the SSI filter and the CS filter, have been described and discussed in detail. The representation of combined uncertainty differs for both approaches, as the SSI filter employs random sets and the CS filter uses sets of densities. In direct comparison of both estimators, it becomes clear that both filters make different assumptions on the set-theoretic uncertainty, which leads to differences in the estimation results. The estimation results of both filters have been compared by means of examples and special cases, where the CS filter has produced more conservative estimates compared to the SSI filter results. It was shown that depending on requirements of the considered problem, either the SSI filter or the CS filter may be suitable for state estimation in the presence of combined set-theoretic and stochastic uncertainties. Possible application scenarios for the SSI filter are localization and extended object tracking, where set-theoretic

uncertainties frequently occur. The CS Filter is suitable for scenarios in which conservative bounds are desired, for instance in control applications.

References

[1] D. Simon, *Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches.* John Wiley & Sons, 2006.

[2] P. L. Combettes, "Foundations of Set Theoretic Estimation," *Proceedings of IEEE*, vol. 81, no. 2, pp. 182–208, 1993.

[3] U. D. Hanebeck, J. Horn, and G. Schmidt, "On Combining Statistical and Set Theoretic Estimation," *Automatica*, vol. 35, no. 6, pp. 1101–1109, Jun. 1999.

[4] M. Baum and U. D. Hanebeck, "Extended Object Tracking based on Combined Set-Theoretic and Stochastic Fusion," in *Proceedings of the 12th International Conference on Information Fusion (Fusion 2009)*, Seattle, Washington, Jul. 2009.

[5] H. T. Nguyen, *An Introduction to Random Sets.* Chapman & Hall/CRC, 2006.

[6] I. Couso, L. Sánchez, and P. Gil, "Imprecise Distribution Function Associated to a Random Set," *Information Sciences*, vol. 159, no. 1-2, pp. 109–123, 2004.

[7] B. Noack, V. Klumpp, D. Brunn, and U. D. Hanebeck, "Nonlinear Bayesian Estimation with Convex Sets of Probability Densities," in *Proceedings of the 11th International Conference on Information Fusion (Fusion 2008)*, Cologne, Germany, Jul. 2008, pp. 1–8.

[8] B. Noack, V. Klumpp, and U. D. Hanebeck, "State Estimation with Sets of Densities considering Stochastic and Systematic Errors," in *Proceedings of the 12th International Conference on Information Fusion (Fusion 2009)*, Seattle, Washington, Jul. 2009.

[9] D. R. Morrell and W. C. Stirling, "Set-Valued Filtering and Smoothing," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 21, no. 1, pp. 184–193, January/February 1991.

[10] P. Walley, *Statistical Reasoning with Imprecise Probabilities.* London: Chapman and Hall, 1991, vol. Monographs on Statistics and Applied Probability 42.

[11] A. Benavoli and A. Antonucci, "Aggregating Imprecise Probabilistic Knowledge," in *Proceedings of 6th International Symposium on Imprecise Probability: Theories and Applications*, Durham, United Kingdom, 2009.

[12] T. Henningsson, "Recursive State Estimation for Linear Systems with Mixed Stochastic and Set-Bounded Disturbances," in *Proceedings of the 47th IEEE Conference on Decision and Control (CDC 2008)*, Cancun, Mexico, 2008.

[13] P. L. Combettes, M. Benidir, and B. Picinbono, "A General Framework for the Incorporation of Uncertainty in Set Theoretic Estimation," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 92)*, San Francisco, California, 1992, pp. 349–352.

[14] P. Combettes and M. Civanlar, "The Foundations of Set Theoretic Estimation," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 91)*, Toronto, Canada, 1991, pp. 2921–2924 vol. 4.

[15] S. Gollamudi, S. Nagaraj, S. Kapoor, and Y. F. Huang, "Set Membership State Estimation with Optimal Bounding Ellipsoids," in *Proceedings of the 1996 International Symposium on Information Theory and its Applications (ISITA 1996)*, 1996, pp. 262–265.

[16] J. Shamma and K.-Y. Tu, "Approximate Set-Valued Observers for Nonlinear Systems," *IEEE Transactions on Automatic Control*, vol. 42, no. 5, pp. 648–658, May 1997.

[17] U. D. Hanebeck and J. Horn, "New Estimators for Mixed Stochastic and Set Theoretic Uncertainty Models: The Vector Case," in *Proceedings of the 5th European Control Conference (ECC 1999)*, Karlsruhe, Germany, Sep. 1999.

[18] ——, "New Estimators for Mixed Stochastic and Set Theoretic Uncertainty Models: The General Case," in *Proceedings of the 2001 American Control Conference (ACC 2001)*, Arlington, Virginia, 2001.

[19] A. P. Dempster, "Upper and Lower Probabilities Generated by a Random Closed Interval," *The Annals of Mathematical Statistics*, vol. 39, no. 3, pp. 957–966, 1968.

[20] E. Miranda, I. Couso, and P. Gil, "Random Sets as Imprecise Random Variables," *Journal of Mathematical Analysis and Applications*, vol. 307, no. 1, pp. 32 – 47, 2005.