State Estimation with Sets of Densities Considering Stochastic and Systematic Errors

Benjamin Noack, Vesa Klumpp, and **Uwe D. Hanebeck**
Intelligent Sensor-Actuator-Systems Laboratory (ISAS),
Institute for Anthropomatics,
Universität Karlsruhe (TH), Germany.
{noack,klumpp}@ira.uka.de, uwe.hanebeck@ieee.org

Abstract – *In practical applications, state estimation requires the consideration of stochastic and systematic errors. If both error types are present, an exact probabilistic description of the state estimate is not possible, so that common Bayesian estimators have to be questioned. This paper introduces a theoretical concept, which allows for incorporating unknown but bounded errors into a Bayesian inference scheme by utilizing sets of densities. In order to derive a tractable estimator, the Kalman filter is applied to ellipsoidal sets of means, which are used to bound additive systematic errors. Also, an extension to nonlinear system and observation models with ellipsoidal error bounds is presented. The derived estimator is motivated by means of two example applications.*

Keywords: Bayesian estimation, Kalman filter, credal sets, ellipsoidal sets, systematic and stochastic errors.

1 Introduction

Obtaining insight to unkown quantities through indirect measurements appears to be a key problem in many technical applications. In general, the available information about the underlying system and the observations are corrupted by disturbances (e.g., noise), so that only an imprecise description of the inaccessible states is attainable. For random errors with known probability distributions, the Bayesian inference scheme embodies a theoretical solution for state estimation. In the case of linear systems and Gaussian noise terms, a well-known closed-form solution to the estimation problem is provided by the Kalman filter [1], where the state estimate is represented by conditional mean and covariance matrix. For nonlinear systems, the extended Kalman filter [2] has turned out to be a well-accepted approach. Bayesian inference models have in common that appropriate estimates are only guaranteed, when the probability distributions of the noise terms are known and the system is not additionally affected by other disturbances. Otherwise, the assumption, that the state is to be characterized by a unique probability density, cannot be maintained.

In order to cope with imprecisely known probabilities, a generic Bayesian estimator for sets of probability densities has been proposed in [3]. Here, the state estimate is given by a convex set of densities, the so-called *credal state*. Also, the requirement of *credal uniqueness* of the transition and likelihood densities has been relaxed to allow convex sets of transition and likelihood densities, respectively. While a purely stochastic description of a system implies a lack of precision, the incapability of specifying an appropriate probability distribution expresses a lack of knowledge. Due to incorrect assumptions, incomplete prior knowledge or unknown, possibly non-stochastic disturbances, uncertainty is overlaid by ignorance. Therefore, every distribution in the credal set is to be regarded as a seriously possible candidate for the true unknown probability distribution. A justification for demanding convexity of these sets is primarily given by Levi's epistemology [4, 5]: Since we are not capable of deciding, which distribution in the set is the true one, the probability judgement will remain the same, when adding all weighted averages to the set, i.e., taking the convex hull. Especially against the background of decision theory, this can easily be accepted [6]. When considering (linear) utility functions, an arbitrary set of distributions and its convex hull yield the same pattern of preferences. In this regard, a non-convex set of distributions is called *generator set*, whose convex hull yields the probability assessment of the state estimate then.

Besides Bayesian estimation, set-membership approaches [7, 8, 9] have been proposed, where no stochastic model of the system unknowns is used. With set-membership estimation, the unknown quantities are assumed to belong to bounded sets. The set of all possible states is computed through the system model and intersected with the set of states, which are consistent with the observations.

An approach for combining stochastic and set-theoretic estimation has been introduced by [10], where sets with stochastically uncertain bounds are used. In this paper, we suggest to use sets of densities instead.

In general, the effect of stochastic errors can be compensated by repeating measurements. In contrast, systematic errors cannot be described statistically and are often difficult to deal with, especially when they are not constant over time. One has to decide between a possible high technical effort to remove these errors or accepting them. In some situations, the consideration of systematic errors might even not be avoidable. In this paper, stochastic and unknown but bounded systematic errors are considered simultaneously. Due to the systematic errors, a unique probabilistic description of the state estimate is not achievable anymore. Therefore, sets of densities need to be processed and the concept proposed in [3] will be applied. The generic estimator is explained in Section 2. In view of practical applicability, Sections 3 and 4 present a generalization of the Kalman filtering scheme for linear and nonlinear system models, respectively. This set-valued Kalman filter is a further development of the filter proposed in [11], where only the prior knowledge is modeled as a set. Two promising applications are presented exemplary in Section 5. Section 6 concludes this paper with an outlook to prospective investigations.

2 Estimation Considering Stochastic and Systematic Errors

At first, we elucidate the basic concept of state estimation with stochastic and systematic errors and focus on a probabilistic discrete-time nonlinear dynamic system model

$$\underline{\boldsymbol{x}}_{k+1} = \underline{a}_k(\underline{\boldsymbol{x}}_k, \hat{\underline{u}}_k, \underline{\boldsymbol{w}}_k, \underline{d}_k) \ ,$$

where underlined letters denote vectors and random variables are written in boldface letters. The system function $\underline{a}_k$ maps the inaccessible n-dimensional state, which is characterized by the random vector $\underline{\boldsymbol{x}}_k$, to the state $\underline{\boldsymbol{x}}_{k+1}$ at time step $k+1$. Furthermore, the system function depends on the system noise $\underline{\boldsymbol{w}}_k$ with probabilitiy density f_k^w and on an unknown systematic error $\underline{d}_k \in \mathcal{D}_k \subset \mathbb{R}^n$. $\hat{\underline{u}}_k$ denotes the input value. The outcome of a measurement $\hat{\underline{y}}_k$ is given by a nonlinear observation model

$$\hat{\underline{y}}_k = \underline{h}_k(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{v}}_k, \underline{e}_k)$$

with respect to the state $\underline{\boldsymbol{x}}_k$, the measurement noise $\underline{\boldsymbol{v}}_k$, and an unknown error $\underline{e}_k \in \mathcal{E}_k \subset \mathbb{R}^m$. We require the noise variables $\underline{\boldsymbol{w}}_k$, $\underline{\boldsymbol{v}}_k$ and the initial state $\underline{\boldsymbol{x}}_0$ to be stochastically independent. A Bayesian estimator allows for recursively computing an estimate for the state $\underline{\boldsymbol{x}}_k$.

The time update or *prediction step* can be expressed in terms of probability density functions by the Chapman-Kolmogorov equation

$$f_{k+1}^{\mathrm{p}}(\underline{x}_{k+1}) = \int_{\mathbb{R}^n} f_k^{\mathrm{T}}(\underline{x}_{k+1}|\underline{x}_k, \hat{\underline{u}}_k, \underline{d}_k) f_k^{\mathrm{e}}(\underline{x}_k)\, d\underline{x}_k \ ,$$

where

$$f_k^{\mathrm{T}}(\underline{x}_{k+1}|\underline{x}_k, \hat{\underline{u}}_k, \underline{d}_k) = \int_{\mathbb{R}^n} \delta\big(\underline{x}_{k+1} - \underline{a}_k(\underline{x}_k, \hat{\underline{u}}_k, \underline{w}_k, \underline{d}_k)\big) f_k^w(\underline{w}_k) d\underline{w}_k$$

is the transition density. δ denotes the n-dimensional Dirac delta function. Since $\underline{d}_k$ is not known, we have to consider the entire set of possible transition densities

$$\boldsymbol{\mathcal{F}}_k^{\mathrm{T}} = \big\{ f_k^{\mathrm{T}}(\underline{x}_{k+1}|\underline{x}_k, \hat{\underline{u}}_k, \underline{d}_k) \ \big| \ \underline{d}_k \in \mathcal{D}_k \big\} \ ,$$

each of which is a seriously potential candidate for the true state transition density. Thus, we are obliged to propagate the entire set elementwise and we obtain a set $\boldsymbol{\mathcal{F}}_{k+1}^{\mathrm{p}}$ of predicted densities.

In the *filtering step*, an observation $\hat{\underline{y}}_k$ is incorporated into a prior density f_k^{p} in order to obtain an improved estimated density f_k^{e}, i.e.,

$$f_k^{\mathrm{e}}(\underline{x}_k) = \frac{f_k(\hat{\underline{y}}_k|\underline{x}_k, \underline{e}_k) f_k^{\mathrm{p}}(\underline{x}_k)}{\int_{\mathbb{R}^n} f_k(\hat{\underline{y}}_k|\underline{x}_k, \underline{e}_k) f_k^{\mathrm{p}}(\underline{x}_k)\, d\underline{x}_k} \ ,$$

where $f_k(\hat{\underline{y}}_k|\underline{x}_k, \underline{e}_k)$ denotes the likelihood

$$f(\hat{\underline{y}}_k\,|\underline{x}_k, \underline{e}_k) = \int_{\mathbb{R}^n} \delta\big(\hat{\underline{y}}_k - \underline{h}_k(\underline{x}_k, \underline{v}_k, \underline{e}_k)\big) f_k^v(\underline{v}_k)\, d\underline{v}_k \ .$$

Again, this function is not unique, since we only know $\underline{e}_k \in \mathcal{E}_k$. Therefore, the filtering step is performed for every $f(\hat{\underline{y}}_k\,|\cdot, \underline{e}_k), \underline{e}_k \in \mathcal{E}_k$ and, if f_k^{p} is not unique either, for every $f_k^{\mathrm{p}} \in \boldsymbol{\mathcal{F}}_k^{\mathrm{p}}$. This elementwise filtering yields a set $\boldsymbol{\mathcal{F}}_k^{\mathrm{e}}$ of estimated densities, which can then processed elementwise with the set of transition densities or likelihoods in the following prediction or filtering step, respectively.

Commonly, additive perturbation terms are considered. The system and observation model are then of the form

$$\underline{\boldsymbol{x}}_{k+1} = \underline{a}_k(\underline{\boldsymbol{x}}_k, \hat{\underline{u}}_k) + \underline{\boldsymbol{w}}_k + \underline{d}_k$$

and

$$\hat{\underline{y}}_k = \underline{h}_k(\underline{x}_k) + \underline{\boldsymbol{v}}_k + \underline{e}_k \ .$$

In this case, the state transition simplifies to

$$f_k^{\mathrm{T}}(\underline{x}_{k+1}|\underline{x}_k, \hat{\underline{u}}_k, \underline{d}_k) = f_k^w(\underline{x}_{k+1} - \underline{a}_k(\underline{x}_k, \hat{\underline{u}}_k) - \underline{d}_k) \ ,$$

which leads to the set

$$\boldsymbol{\mathcal{F}}_k^{\mathrm{T}} = \big\{ f_k^w(\underline{x}_{k+1} - \underline{a}_k(\underline{x}_k, \hat{\underline{u}}_k) - \underline{d}_k) \ \big| \ \underline{d}_k \in \mathcal{D}_k \big\}$$

of transition densities. $\boldsymbol{\mathcal{F}}_k^{\mathrm{T}}$ then only contains translations of f_k^w. Analogously, we obtain the set

$$\boldsymbol{\mathcal{F}}_k^{\mathrm{L}} = \Big\{ f_k^v(\hat{\underline{y}}_k - \underline{h}_k(\underline{x}_k) - \underline{e}_k) \ \Big| \ \underline{e}_k \in \mathcal{E}_k \Big\}$$

of translated likelihoods in the filtering step. Such a set can be interpreted as a density with imprecisely known mean and can therefore be parameterized by the set of possible means.

For linear equations

$$\underline{x}_{k+1} = \mathbf{A}_k\,\underline{x}_k + \mathbf{B}_k\,\underline{u}_k = \mathbf{A}_k\,\underline{x}_k + \mathbf{B}_k\,(\underline{\hat{u}}_k + \underline{\boldsymbol{w}}_k)\ ,$$
$$\underline{\hat{y}}_k = \underline{y}_k + \underline{\boldsymbol{v}}_k = \mathbf{H}_k\,\underline{x}_k + \underline{\boldsymbol{v}}_k\ ,$$

where the known input and measurement values $\underline{\hat{u}}_k$ and $\underline{\hat{y}}_k$ are affected by (zero-mean) Gaussian white noise terms $\underline{\boldsymbol{w}}_k$, $\underline{\boldsymbol{v}}_k$, the Kalman filter is usually utilized. Then, only the first two moments need to be propagated. The predicted mean and covariance matrix are obtained by

$$\underline{\hat{x}}^{\mathrm{p}}_{k+1} = \mathbf{A}_k\,\underline{\hat{x}}^{\mathrm{e}}_k + \mathbf{B}_k\,\underline{\hat{u}}_k \tag{1}$$

and

$$\mathbf{C}^{\mathrm{p}}_{k+1} = \mathbf{A}_k\mathbf{C}^{\mathrm{e}}_k\mathbf{A}^{\mathrm{T}}_k + \mathbf{B}_k\mathbf{C}^{w}_k\mathbf{B}^{\mathrm{T}}_k\ ,$$

which yield the predicted Gaussian density $f^{\mathrm{p}}_{k+1} = \mathcal{N}(\underline{\hat{x}}^{\mathrm{p}}_{k+1}, \mathbf{C}^{\mathrm{p}}_{k+1})$ of the state. The filtering step gives the estimated conditional mean

$$\underline{\hat{x}}^{\mathrm{e}}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\underline{\hat{x}}^{\mathrm{p}}_k + \mathbf{K}_k\,\underline{\hat{y}}_k \tag{2}$$

and the covariance

$$\mathbf{C}^{\mathrm{e}}_k = \mathbf{C}^{\mathrm{p}}_k - \mathbf{K}_k\mathbf{H}_k\mathbf{C}^{\mathrm{p}}_k$$

with Kalman gain

$$\mathbf{K}_k = \mathbf{C}^{\mathrm{p}}_k\mathbf{H}^{\mathrm{T}}_k(\mathbf{C}^{v}_k + \mathbf{H}_k\mathbf{C}^{\mathrm{p}}_k\mathbf{H}^{\mathrm{T}}_k)^{-1}\ .$$

The estimated density is then given by $f^{\mathrm{e}}_k = \mathcal{N}(\underline{\hat{x}}^{\mathrm{e}}_k, \mathbf{C}^{\mathrm{e}}_k)$. If input and measurement are affected by both stochastic and systematic perturbations

$$\underline{u}_k = \underline{\hat{u}}_k + \underline{\boldsymbol{w}}_k + \underline{d}_k\ ,$$
$$\underline{\hat{y}}_k = \underline{y}_k + \underline{\boldsymbol{v}}_k + \underline{e}_k\ ,$$

then obviously only the means are biased by systematic errors. Since these errors are unknown, the means are not unique anymore. With

$$\mathcal{U}_k = \{\underline{\hat{u}}_k + \underline{d}_k \mid \underline{d}_k \in \mathcal{D}_k\}\ ,$$
$$\mathcal{Y}_k = \left\{\underline{\hat{y}}_k - \underline{e}_k \mid \underline{e}_k \in \mathcal{E}_k\right\}\ ,$$

equations (1) and (2) become

$$\mathcal{X}^{\mathrm{p}}_{k+1} = \mathbf{A}_k\,\mathcal{X}^{\mathrm{e}}_k \oplus \mathbf{B}_k\,\mathcal{U}_k \tag{3}$$

and

$$\mathcal{X}^{\mathrm{e}}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathcal{X}^{\mathrm{p}}_k \oplus \mathbf{K}_k\,\mathcal{Y}_k\ , \tag{4}$$

respectively, where $\mathcal{X}_k$ is the set of conditional means. The operator $\oplus$ denotes the Minkowski sum, which is the elementwise summation. In the following sections, we derive a tractable estimator by parameterizing these sets properly.

3 Kalman Filter for Ellipsoidal Sets

In the case of linear dynamical systems perturbed by additive white Gaussian noise and unknown but bounded disturbances, we will employ a generalization of the Kalman filtering scheme. The unknown systematic errors will each be characterized by an enclosing ellipsoidal set

$$\mathcal{E}(\underline{\hat{c}}, \mathbf{X}) = \left\{\underline{x} \in \mathbb{R}^n \mid (\underline{x} - \underline{\hat{c}})^{\mathrm{T}}\mathbf{X}^{-1}(\underline{x} - \underline{\hat{c}}) \leq 1\right\}$$

with midpoint $\underline{\hat{c}} \in \mathbb{R}^n$ and symmetric positive definite matrix $\mathbf{X} \in \mathbb{R}^{n\times n}$. As revealed below, the latter requirement may be weakened by allowing for singular matrices, so that the ellipsoids may become degenerate. Some preparatory work will be required, before establishing the ellipsoidal estimation concept.

A more general definition of an ellipsoidal set is obtained by utilizing support functions $\rho(\,\cdot\,, \mathcal{C}) : \mathbb{R}^n \to \mathbb{R}$ with

$$\rho(\underline{l}, \mathcal{C}) = \sup_{\underline{x}\in\mathcal{C}} \langle \underline{l}, \underline{x}\rangle\ ,$$

where $\mathcal{C} \subseteq \mathbb{R}^n$ denotes a closed convex set and $\langle\,\cdot\,,\,\cdot\,\rangle$ is the scalar product. The principal property of support functions is stated in the next lemma [8, 9].

Lemma 1 *For a closed convex set $\mathcal{C} \subseteq \mathbb{R}^n$, the inclusion $\underline{x} \in \mathcal{C}$ holds, if and only if the inequality*

$$\langle \underline{l}, \underline{x}\rangle \leq \rho(\underline{l}, \mathcal{C}) \quad \forall \underline{l} \in \mathbb{R}^n$$

is satisfied.

An ellipsoidal set $\mathcal{E}(\underline{\hat{c}}, \mathbf{X})$ is a closed convex set, whose support function is given by

$$\rho\big(\underline{l}, \mathcal{E}(\underline{\hat{c}}, \mathbf{X})\big) = \langle \underline{l}, \underline{\hat{c}}\rangle + \langle \underline{l}, \mathbf{X}\,\underline{l}\rangle^{\frac{1}{2}}\ .$$

According to its support function, an ellipsoid can now be defined by

$$\mathcal{E}(\underline{\hat{x}}, \mathbf{X}) = \left\{\underline{x} \in \mathbb{R}^n \middle|\, \langle \underline{l}, \underline{x}\rangle \leq \langle \underline{l}, \underline{\hat{c}}\rangle + \langle \underline{l}, \mathbf{X}\,\underline{l}\rangle^{\frac{1}{2}}\,, \forall \underline{l} \in \mathbb{R}^n\right\}$$

with midpoint $\underline{\hat{c}}$ and symmetric *nonnegative* definite matrix $\mathbf{X}$. If the shape matrix $\mathbf{X}$ is singular, some semiaxes of the ellipsoid are zero and thus the volume is equal to zero, too. For instance, such a degenerate ellipsoid is given by a disc in three-dimensional space. The example presented in Section 5.2 will confirm the necessity of considering degenerate ellipsoids. Herein, scalar distance measurements are deployed to confront a two-dimensional localization problem. Therefore, the error bound of a single measurement is a degenerate ellipsoid, a line, in the two-dimensional space.

The next lemma summarizes further important properties of support functions.

Lemma 2 *Support functions fulfill the following two assertions.*

- *Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a closed convex set and let $\underline{y} = \mathbf{A}\,\underline{x} + \underline{b}$ be an affine transformation. For $\underline{x} \in \mathcal{C}$, we obtain the inequality*

$$\big\langle \underline{l}, \underline{y}\big\rangle = \langle \underline{l}, \mathbf{A}\,\underline{x} + \underline{b}\rangle \leq \rho(\underline{l}, \mathbf{A}\mathcal{C} + \underline{b}) \quad \forall \underline{l} \in \mathbb{R}^n\ .$$

- *The support function of the Minkowski sum of two closed convex sets $\mathcal{C}_1, \mathcal{C}_2 \subseteq \mathbb{R}^n$ is given by*

$$\rho(\underline{l}, \mathcal{C}_1 \oplus \mathcal{C}_2) = \rho(\underline{l}, \mathcal{C}_1) + \rho(\underline{l}, \mathcal{C}_2)\ .$$

Remark 1 *An affine transformation $\underline{y} = \mathbf{A}\,\underline{x} + \underline{b}$ applied to every element*

$$\underline{x} \in \mathcal{E}(\hat{\underline{c}}, \mathbf{X})$$

yields the ellipsoid

$$\mathbf{A}\mathcal{E}(\hat{\underline{c}}, \mathbf{X}) + \underline{b} = \mathcal{E}(\mathbf{A}\,\hat{\underline{c}} + \underline{b}, \mathbf{A}\mathbf{X}\mathbf{A}^{\mathrm{T}}) \,.$$

Assuming that the set $\mathcal{X}_k^{\mathrm{e}}$ of estimated means and the error bound $\mathcal{U}_k$ of input values are closed and convex, we are able to express equation (3), the prediction step, in terms of support functions, i.e.,

$$\begin{aligned} \rho\left(\underline{l}, \mathcal{X}_{k+1}^{\mathrm{p}}\right) &= \rho\left(\underline{l}, \mathbf{A}_k\,\mathcal{X}_k^{\mathrm{e}} \oplus \mathbf{B}_k\,\mathcal{U}_k\right) \\ &= \rho\left(\underline{l}, \mathbf{A}_k\,\mathcal{X}_k^{\mathrm{e}}\right) + \rho\left(\underline{l}, \mathbf{B}_k\,\mathcal{U}_k\right) \,. \end{aligned} \tag{5}$$

Analogously, equation (4) can be rewritten as

$$\begin{aligned} \rho\left(\underline{l}, \mathcal{X}_k^{\mathrm{e}}\right) &= \rho\left(\underline{l}, (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathcal{X}_k^{\mathrm{p}} \oplus \mathbf{K}_k\,\mathcal{Y}_k\right) \\ &= \rho\left(\underline{l}, (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathcal{X}_k^{\mathrm{p}}\right) + \rho\left(\underline{l}, \mathbf{K}_k\,\mathcal{Y}_k\right) \,, \end{aligned} \tag{6}$$

where $\mathcal{X}_k^{\mathrm{p}}$ and $\mathcal{Y}_k$ are closed convex sets of predicted states and measurements, respectively. Equations (5) and (6) state that the resulting sets of conditional means $\mathcal{X}_{k+1}^{\mathrm{p}}$ and $\mathcal{X}_k^{\mathrm{e}}$ are convex as well, so linear prediction and filtering preserve convexity. The justification for demanding the sets to be convex is constituted by the following idea: The sets of means are intended to bound maximal possible systematic errors, so if $\underline{d}_1$ and $\underline{d}_2$ are two possible biases, then every $\underline{d}_\alpha = \alpha\underline{d}_1 + (1-\alpha)\underline{d}_2, \alpha \in [0,1]$ is obviously bounded by $\underline{d}_1$ and $\underline{d}_2$. Hence, $\underline{d}_\alpha$ should be also regarded as possible bias.

In general, equations (5) and (6) do *not* imply that a specific representation of sets will be preserved. For instance, the Minkowski sum of ellipsoids does not usually yield an ellipsoid. However, especially for ellipsoidal sets, there exists a simple approach to obtain an enclosing ellipsoid of the Minkowski sum, as stated in the following theorem.

Theorem 1
Let $\mathcal{E}(\hat{\underline{c}}_1, \mathbf{X}_1)$ and $\mathcal{E}(\hat{\underline{c}}_2, \mathbf{X}_2)$ be two ellipsoids. The inclusion

$$\mathcal{E}(\hat{\underline{c}}_1, \mathbf{X}_1) \oplus \mathcal{E}(\hat{\underline{c}}_2, \mathbf{X}_2) \subseteq \mathcal{E}(\hat{\underline{c}}_1 + \hat{\underline{c}}_2, \mathbf{X}(p))$$

holds for every matrix

$$\mathbf{X}(p) = (1 + p^{-1})\mathbf{X}_1 + (1 + p)\mathbf{X}_2$$

with $p > 0$. $\mathcal{E}(\hat{\underline{c}}_1 + \hat{\underline{c}}_2, \mathbf{X}(p))$ is a properly defined ellipsoid and an outer approximation of the Minkowski sum.

PROOF. This assertion can be proven by means of support functions. For a thorough proof refer to [9]. □

Theorem 1 provides a parametric family of outer approximations, for which the equality

$$\mathcal{E}(\hat{\underline{c}}_1, \mathbf{X}_1) \oplus \mathcal{E}(\hat{\underline{c}}_2, \mathbf{X}_2) = \bigcap \left\{ \mathcal{E}(\hat{\underline{c}}_1 + \hat{\underline{c}}_2, \mathbf{X}(p)) \;\middle|\; p > 0 \right\}$$

holds. An appropriate enclosing ellipsoid can be chosen in view of different optimality criteria. Typically, the parameter p is determined such that the volume of $\mathcal{E}(\hat{\underline{c}}_1 + \hat{\underline{c}}_2, \mathbf{X}(p))$ becomes minimal or the trace of $\mathbf{X}(p)$, i.e., the sum of squares of the semiaxes, is minimized.

Theorem 2
The parametric family of outer approximations given by Theorem 1 contains the following optimal ellipsoids.

1. *The minimum volume ellipsoid enclosing $\mathcal{E}(\hat{\underline{c}}_1, \mathbf{X}_1) \oplus \mathcal{E}(\hat{\underline{c}}_2, \mathbf{X}_2)$ is constituted by $\mathcal{E}(\hat{\underline{c}}_1 + \hat{\underline{c}}_2, \mathbf{X}(p^*))$, where $p^* > 0$ denotes the unique solution of the equation*

$$\sum_{i=1}^{n} \frac{1}{\lambda_i + p} = \frac{n}{p(p+1)} \tag{7}$$

and $0 \le \lambda_1 \le \ldots \le \lambda_n < \infty$ are the roots of

$$\det(\mathbf{X}_1 - \lambda\mathbf{X}_2) = 0 \,.$$

2. *The trace of $\mathbf{X}(p)$ is minimal for*

$$p^+ = \frac{\mathrm{trace}(\mathbf{X}_1)^{\frac{1}{2}}}{\mathrm{trace}(\mathbf{X}_2)^{\frac{1}{2}}} \,. \tag{8}$$

Then, the enclosing ellipsoid $\mathcal{E}(\hat{\underline{c}}_1 + \hat{\underline{c}}_2, \mathbf{X}(p^+))$ features minimal sum of squares of semiaxes.

Calculating the minimal volume enclosing ellipsoid requires to solve the algebraic equation (7), which can turn out to be a laborious task. In contrast, considering the trace as optimality criterion provides an easy way to compute an outer approximation, since the required parameter is directly given by equation (8). Furthermore, the trace appears to be an adequate criterion, because the maximal length of the semiaxes is a bound of the maximal presumed systematic error.

With this spadework, we are now in a position to depict the generalized Kalman filtering scheme for ellipsoidal sets allowing for simultaneous treatment of stochastic and systematic disturbances.

3.1 Prediction Step

Consider a linear system

$$\underline{x}_{k+1}^{\mathrm{p}} = \mathbf{A}_k\,\underline{x}_k^{\mathrm{e}} + \mathbf{B}_k(\hat{\underline{u}}_k + \underline{\boldsymbol{w}}_k + \underline{d}_k) \,,$$

where the state $\underline{x}_k^{\mathrm{e}}$ is characterized by an ellipsoidal set of means $\mathcal{X}_k^{\mathrm{e}} = \mathcal{E}(\hat{\underline{c}}_k^{\mathrm{e}}, \mathbf{X}_k^{\mathrm{e}})$ and covariance matrix $\mathbf{C}_k^{\mathrm{e}}$. The input value $\hat{\underline{u}}_k$ is corrupted by zero-mean white Gaussian noise $\underline{\boldsymbol{w}}_k$ with covariance $\mathbf{C}_k^w$ and an unknown systematic error $\underline{d}_k$, which is bounded by the ellipsoid $\mathcal{E}(\underline{0}, \mathbf{U}_k)$. Both perturbations can be combined to a set of translated Gaussian densities, which is characterized by the set of means $\mathcal{U}_k = \mathcal{E}(\hat{\underline{u}}_k, \mathbf{U}_k)$ and covariance matrix $\mathbf{C}_k^w$. The Kalman prediction step consists of calculating the enclosing ellipsoid

$$\mathbf{A}_k\mathcal{X}_k^{\mathrm{e}} \oplus \mathbf{B}_k\mathcal{U}_k = \mathcal{X}_{k+1}^{\mathrm{p}} \subseteq \mathcal{E}(\hat{\underline{c}}_{k+1}^{\mathrm{p}}, \mathbf{X}_{k+1}^{\mathrm{p}})$$

of predicted conditional means and of determining the covariance matrix $\mathbf{C}_{k+1}^{\mathrm{p}}$. According to Remark 1 and Theorem 1, the midpoint of the outer approximation is

$$\hat{\underline{c}}_{k+1}^{\mathrm{p}} = \mathbf{A}_k\,\hat{\underline{c}}_k^{\mathrm{e}} + \mathbf{B}_k\,\hat{\underline{u}}_k$$

and the matrix $\mathbf{X}_{k+1}^{\mathrm{p}}$ is given by

$$\mathbf{X}_{k+1}^{\mathrm{p}} = (1+p^{-1})\mathbf{A}_k\mathbf{X}_k^{\mathrm{e}}\mathbf{A}_k^{\mathrm{T}} + (1+p)\mathbf{B}_k\mathbf{U}_k\mathbf{B}_k^{\mathrm{T}} , \quad (9)$$

where p may be chosen by means of Theorem 2. The covariance matrix is obtained in a known manner by

$$\mathbf{C}_{k+1}^{\mathrm{p}} = \mathbf{A}_k\mathbf{C}_k^{\mathrm{e}}\mathbf{A}_k^{\mathrm{T}} + \mathbf{B}_k\mathbf{C}_k^{w}\mathbf{B}_k^{\mathrm{T}} . \quad (10)$$

Finally, $\mathcal{E}(\hat{\underline{c}}_{k+1}^{\mathrm{p}}, \mathbf{X}_{k+1}^{\mathrm{p}})$ and $\mathbf{C}_{k+1}^{\mathrm{p}}$ are the result of the prediction step.

3.2 Filtering Step

The obervation model of the form

$$\hat{\underline{y}}_k = \mathbf{H}_k\,\underline{x}_k + \underline{\boldsymbol{v}}_k + \underline{e}_k$$

is affected by additive Gaussian white noise $\underline{\boldsymbol{v}}_k$ and a systematic error $\underline{e}_k$. The random variable $\underline{\boldsymbol{v}}_k$ is zero-mean with covariance matrix $\mathbf{C}_k^v$. Furthermore, the error $\underline{e}_k$ is contained in the ellipsoid $\mathcal{E}(\underline{0}, \mathbf{Y}_k)$. Therefore, $(\hat{\underline{y}}_k - \underline{e}_k)$ lies in the ellipsoid $\mathcal{Y}_k = \mathcal{E}(\hat{\underline{y}}_k, \mathbf{Y}_k)$. This set of corrupted measurements is now fused elementwise with the ellipsoid $\mathcal{X}_k^{\mathrm{p}} = \mathcal{E}(\underline{x}_k^{\mathrm{p}}, \mathbf{X}_k^{\mathrm{p}})$ of predicted states, i.e.,

$$\mathcal{X}_k^{\mathrm{e}} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathcal{X}_k^{\mathrm{p}} \oplus \mathbf{K}_k\mathcal{Y}_k$$

with Kalman gain

$$\mathbf{K}_k = \mathbf{C}_k^{\mathrm{p}}\mathbf{H}_k^{\mathrm{T}}(\mathbf{C}_k^v + \mathbf{H}_k\mathbf{C}_k^{\mathrm{p}}\mathbf{H}_k^{\mathrm{T}})^{-1} . \quad (11)$$

In analogy to the prediction step, the outer approximation $\mathcal{E}(\hat{\underline{c}}_k^{\mathrm{e}}, \mathbf{X}_k^{\mathrm{e}}) \supseteq \mathcal{X}_k^{\mathrm{e}}$ has the midpoint

$$\hat{\underline{c}}_k^{\mathrm{e}} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\hat{\underline{c}}_k^{\mathrm{p}} + \mathbf{K}_k\hat{\underline{y}}_k$$

and matrix

$$\begin{aligned}\mathbf{X}_k^{\mathrm{e}} = {} & (1+p^{-1})(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{X}_k^{\mathrm{p}}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^{\mathrm{T}} \\ & + (1+p)\mathbf{K}_k\mathbf{Y}_k\mathbf{K}_k^{\mathrm{T}} ,\end{aligned} \quad (12)$$

where p is chosen appropriately. Since the covariance matrix is independent of the means, we have

$$\mathbf{C}_k^{\mathrm{e}} = \mathbf{C}_k^{\mathrm{p}} - \mathbf{K}_k\mathbf{H}_k\mathbf{C}_k^{\mathrm{p}} . \quad (13)$$

The result of the filtering step is given by the ellipsoid of means $\mathcal{E}(\hat{\underline{c}}_k^{\mathrm{e}}, \mathbf{X}_k^{\mathrm{e}})$ and covariace matrix $\mathbf{C}_k^{\mathrm{e}}$.

3.3 Discussion

For linear dynamic systems with linear observation models, we have proposed a concept for incorporating stochastic and systematic errors simultaneously. The influence of unknown systematic disturbances is expressed in terms of sets of means. In comparison with the stochastic Kalman filter, only one additional parameter has to be considered, when utilizing ellipsoidal bounds for the systematic error. Whereas the midpoints and the covariance matrices are computed by the well-known Kalman equations for mean and covariance, respectively, the shape matrices of the ellipsoids are obtained by equations (9) and (12). Prediction and filtering step involve the computation of Minkowski sums, which yield the set of conditional means, i.e., the credal state of the system. Since the Minkowski sum is not ellipsoidal anymore, an enclosing ellipsoid is determined. When considering the trace as an optimality criterion for the ellipsoidal approximation of the Minkowski sum, the additional computational demand compared to the standard Kalman filter is low.

It is worth recalling that the ellipsoids of means represent sets of translated Gaussian densities with the same covariance matrix. Such a set of densities itself is not convex and therefore to be regarded as a generator set. The presented estimator is a further development of the set-valued filter proposed in [11], where only the state is modeled by a set, but likelihood and transition densities are supposed to be unique.

4 Extended Kalman Filter for Ellipsoidal Sets

A common approach to cope with nonlinear state transition and observation models is the extended Kalman filter (EKF). Here, the standard Kalman filter is applied to linear approximations of the nonlinear functions, where the linearizations are obtained by first-order Taylor series expansions about the current mean. Since we have relaxed the requirement of credal uniqueness, i.e., we have a credal set of conditional means, a linearization of the system model cannot be achieved in exact the same manner. Thus, we aspire to find a linear approximation over the entire set of means. For this purpose, we will deploy the approach proposed in [12].

The linear mapping is calculated to minimize the sum of squared errors between the function values of the nonlinear function and those of the linearization at appropriately chosen points. In [12], it is suggested to select $4n+1$ approximation points for a n-dimensional ellipsoid, which lie equidistantly on the principal axes of the ellipsoid, as depicted in Figure 1. Subsequently, prediction and filtering step will be discussed in detail.

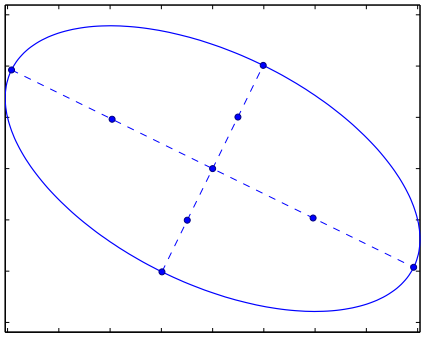


Figure 1: Approximation points of an ellipsoidal set.

4.1 Prediction Step

The system model will be approximated by an affine mapping, i.e.,

$$\underline{x}_{k+1} = \underline{a}_k(\underline{x}_k, \underline{u}_k) \approx \mathbf{A}_k\,\underline{x}_k + \mathbf{B}_k\,\underline{u}_k + \underline{a}_k^0 . \quad (14)$$

Let $\underline{x}_k^{\mathrm{e},(1)}, \ldots, \underline{x}_k^{\mathrm{e},(N)}$ denote the approximation points of the ellipsoid $\mathcal{E}(\hat{\underline{c}}_k^{\mathrm{e}}, \mathbf{X}_k^{\mathrm{e}})$ of filtered states. The input corrupted by an unknown systematic error lies in the ellipsoidal set $\mathcal{E}(\hat{\underline{u}}_k, \mathbf{U}_k)$, whose approximation points are given by $\underline{u}_k^{(1)}, \ldots, \underline{u}_k^{(M)}$. Then,

$$\underline{d}_k^{(i,j)} := \underline{a}_k\left(\underline{x}_k^{\mathrm{e},(i)}, \underline{u}_k^{(j)}\right) - \mathbf{A}_k\,\underline{x}_k^{\mathrm{e},(i)} - \mathbf{B}_k\,\underline{u}_k^{(j)} - \underline{a}_k^0\,,$$

for $i = 1, \ldots, N$, $j = 1, \ldots, M$, are the errors between the nonlinear mapping and its linear approximation at this points. The mappings $\mathbf{A}_k$, $\mathbf{B}_k$ and the base point $\underline{a}_k^0$ are computed to satisfy

$$\left[\mathbf{A}_k, \mathbf{B}_k, \underline{a}_k^0\right] = \arg\min_{\mathbf{A}_k, \mathbf{B}_k, \underline{a}_k^0} \sum_{i,j=1}^{N,M} \omega_{i,j}^{-1} \left[\underline{d}_k^{(i,j)}\right]^{\mathrm{T}} \left[\underline{d}_k^{(i,j)}\right],$$

with weighting factors $\omega_{i,j}^{-1}$. The solution of this weighted least squares problem [13] is given by

$$\begin{bmatrix} \mathbf{A}_k^{\mathrm{T}} \\ \mathbf{B}_k^{\mathrm{T}} \\ (\underline{a}_k^0)^{\mathrm{T}} \end{bmatrix} = \left(\mathbf{F}_k \mathbf{Q}_k^{-1} \mathbf{F}_k^{\mathrm{T}}\right)^{-1} \mathbf{F}_k \mathbf{Q}_k^{-1} \mathbf{f}_k\,,$$

where $\mathbf{F}_k$, $\mathbf{f}_k$ and $\mathbf{Q}$ are defined by

$$\mathbf{F}_k = \begin{bmatrix} \underline{x}_k^{\mathrm{e},(1)} & \ldots & \underline{x}_k^{\mathrm{e},(1)} & \underline{x}_k^{\mathrm{e},(2)} & \ldots & \underline{x}_k^{\mathrm{e},(N)} \\ \underline{u}_k^{(1)} & \ldots & \underline{u}_k^{(M)} & \underline{u}_k^{(1)} & \ldots & \underline{u}_k^{(M)} \\ 1 & \ldots & 1 & 1 & \ldots & 1 \end{bmatrix}, \quad (15)$$

$$\mathbf{f}_k = \begin{bmatrix} \underline{a}_k\left(\underline{x}_k^{\mathrm{e},(1)}, \underline{u}_k^{(1)}\right) & \cdots & \underline{a}_k\left(\underline{x}_k^{\mathrm{e},(N)}, \underline{u}_k^{(M)}\right) \end{bmatrix}^{\mathrm{T}}$$

and $\mathbf{Q}_k = \mathrm{diag}(\omega_{1,1}, \ldots, \omega_{N,M})$.

The midpoint of the ellipsoid $\mathcal{E}(\hat{\underline{c}}_{k+1}^{\mathrm{p}}, \mathbf{X}_{k+1}^{\mathrm{p}})$ of predicted means is obtained by applying the nonlinear function to the midpoints, i.e.,

$$\hat{\underline{c}}_{k+1}^{\mathrm{p}} = \underline{a}_k(\hat{\underline{c}}_k^{\mathrm{e}}, \hat{\underline{u}}_k)\,.$$

With the previously determined matrices $\mathbf{A}_k$ and $\mathbf{B}_k$, the shape matrix $\mathbf{X}_{k+1}^p$ and the covariance $\mathbf{C}_{k+1}^p$ are computed by means of equation (9) and (10), respectively.

4.2 Filtering Step

In order to compute a linear approximation for the nonlinear observation model

$$\underline{y}_k = h_k(\underline{x}_k) \approx \mathbf{H}_k\,\underline{x}_k + h_k^0\,,$$

we only need to consider the ellpsoidal set $\mathcal{E}(\hat{\underline{c}}_k^{\mathrm{p}}, \mathbf{X}_k^{\mathrm{p}})$ of predicted states. Let

$$\underline{e}_k^{(i)} = \underline{h}_k\left(\underline{x}_k^{\mathrm{p},(i)}\right) - \mathbf{H}_k\,\underline{x}_k^{\mathrm{p},(i)} - \underline{h}_k^0$$

denote the error between the nonlinearity and its linearization at the approximation points $\underline{x}_k^{\mathrm{p},(1)}, \ldots, \underline{x}_k^{\mathrm{p},(N)}$ of $\mathcal{E}(\hat{\underline{c}}_k^{\mathrm{p}}, \mathbf{X}_k^{\mathrm{p}})$. Again, the matrix $\mathbf{H}_k$ and the vector $\underline{h}_k^0$ are chosen to be the solution of the weighted least squares problem

$$\left[\mathbf{H}_k, \underline{h}_k^0\right] = \arg\min_{\mathbf{H}_k, \underline{h}_k^0} \sum_{i=1}^{N} \omega_i^{-1} \left[\underline{e}_k^{(i)}\right]^{\mathrm{T}} \left[\underline{e}_k^{(i)}\right].$$

For

$$\mathbf{G}_k = \begin{bmatrix} \underline{x}_k^{\mathrm{p},(1)} & \ldots & \underline{x}_k^{\mathrm{p},(N)} \\ 1 & \ldots & 1 \end{bmatrix}$$

and

$$\mathbf{g}_k = \begin{bmatrix} \underline{h}_k\left(\underline{x}_k^{\mathrm{p},(1)}\right) & \cdots & \underline{h}_k\left(\underline{x}_k^{\mathrm{p},(N)}\right) \end{bmatrix}^{\mathrm{T}},$$

$\mathbf{H}_k$ and $\underline{h}_k^0$ are given by

$$\begin{bmatrix} \mathbf{H}_k^{\mathrm{T}} \\ (\underline{h}_k^0)^{\mathrm{T}} \end{bmatrix} = \left(\mathbf{G}_k \mathbf{Q}_k^{-1} \mathbf{G}_k^{\mathrm{T}}\right)^{-1} \mathbf{G}_k \mathbf{Q}_k^{-1} \mathbf{g}_k\,,$$

where $\mathbf{Q}_k = \mathrm{diag}(\omega_1, \ldots, \omega_N)$ is a weighting matrix.

With the aid of this linearization, the predicted states can now be fused with the set of measurements $\mathcal{E}(\hat{\underline{y}}_k, \mathbf{Y}_k)$. The ellipsoid $\mathcal{E}(\hat{\underline{c}}_k^{\mathrm{e}}, \mathbf{X}_k^{\mathrm{e}})$ of filtered means has the midpoint

$$\hat{\underline{c}}_k^{\mathrm{e}} = \underline{\hat{c}}_k^{\mathrm{p}} + \mathbf{K}_k\left(\underline{\hat{y}}_k - \underline{h}_k\left(\underline{\hat{c}}_k^{\mathrm{p}}\right)\right),$$

where $\mathbf{K}_k$ is the Kalman gain (11). The shape matrix $\mathbf{X}_k^{\mathrm{e}}$ is computed from equation (12) and the covariance matrix $\mathbf{C}_k^{\mathrm{e}}$ is given by (13).

4.3 Discussion

The presented approach allows for considering both systematic and stochastic errors in a nonlinear estimation scenario. In compliance with Section 3, the unknown systematic error can be incorporated as a bias in the mean, which is enclosed by an ellipsoidal set. The difficulty of nonlinear estimation is the necessity of propagating the entire densities. A proposal to overcome this issue is to linearize the transition and observation model, such that only the first two moments need to be propagated. The linearization in this section is obtained by the least squares fitting method, in contrast to calculate Jacobians with the conventional EKF.

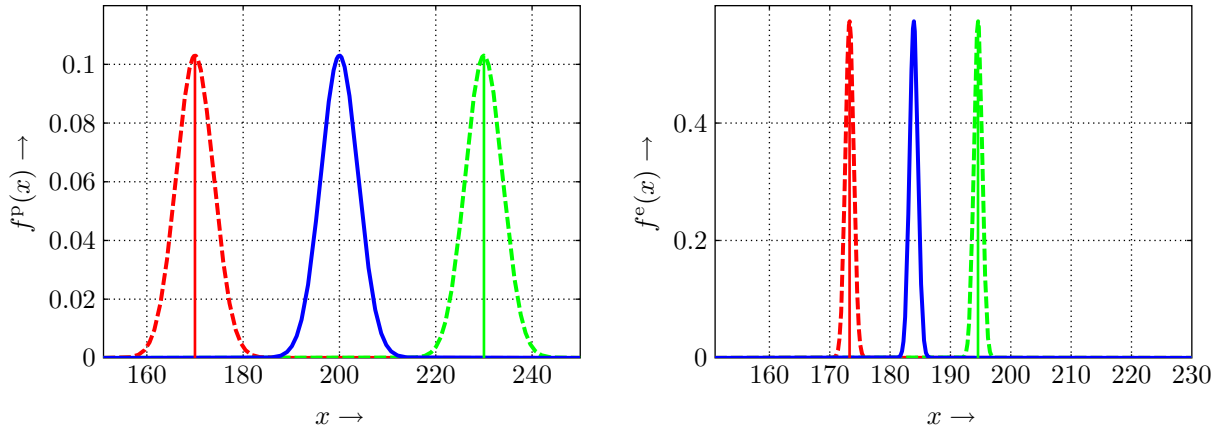
The approximation accuracy strongly depends on the state and input dimensions and on the number of approximation points. For high-dimensional systems, especially matrix (15) is large and the least squares calculation may become cumbersome. In that case it might be favorable to use less approximation points at the cost of lower accuracy.

5 Examples

In this section, two example situations are presented, where set-valued estimation seems promising. In the first example, sets of densities are employed in order to reduce the model complexity. In the second example, the ellipsoidal Kalman filter is used to cope with unknown systematic errors.

5.1 Radar Altimeter

In [14], terrain aided positioning for improving the position estimate of air vehicles is used. For this, precise estimates of ground clearance, i.e., the distance between aircraft and ground, which is measured by radar, has to be known. By flying over a tree-covered area, echos from the ground or echos from trees occur, which can be misinterpreted as



(a) Set of prior state estimates. The set of means is given by $[170\ \mathrm{m}, 230\ \mathrm{m}]$.

(b) Set of state estimates after 20 measurements. The set of means is given by $[173.3\ \mathrm{m}, 194.6\ \mathrm{m}]$.

Figure 2: Fusion result of example in Section 5.1. Figures show the set of estimated densities for the ground clearance. The minimum (red, dashed), the center (blue), and the maximum (green, dashed) elements of the set are shown. The true value is $180\ \mathrm{m}$.

ground clearance. The approach in [14] is to model the likelihood as a Gaussian mixture

$$f^{\mathrm{L}}(x) = \pi \cdot \mathcal{N}(x - \mu_1, \sigma_1^2) + (1 - \pi) \cdot \mathcal{N}(x - \mu_2, \sigma_2^2)\ ,$$

with different means μ_1, μ_2 and variances σ_1^2, σ_2^2. When using this model, it is necessary to give a proper probability π of receiving an echo from ground. This probability is difficult to determine a priori, especially over different terrains.

Instead of a mixture, a set of translated densities is used in the proposed approach. This set of Gaussian densities is given by

$$\boldsymbol{\mathcal{F}}^{v} = \left\{ \mathcal{N}(x_k - d, \sigma_{\boldsymbol{v}}^2) \mid d \in \mathcal{D} \right\}\ ,$$

with a translation within the set $\mathcal{D}$. Measurements $\hat{y}_k$ at discrete time steps are related to the system state by

$$\hat{y}_k = x_k + \boldsymbol{v}_k + d\ ,$$

with $\boldsymbol{v}_k$ zero-mean Gaussian noise with variance $\sigma_v^2 = 10\ \mathrm{m}^2$. The parameter d describes the inability to decide from where the echo came from, here we assume, $d \in \mathcal{D} = [-10\ \mathrm{m}, 10\ \mathrm{m}]$. The true, but unknown, ground clearance is $180\ \mathrm{m}$ and the prior set of densities is set to

$$\boldsymbol{\mathcal{F}}^{\mathrm{p}} = \left\{ \mathcal{N}(x_0 - 200\ \mathrm{m} - d^{\mathrm{p}}, \sigma_{\boldsymbol{x}}^2) \mid d^{\mathrm{p}} \in [-30\ \mathrm{m}, 30\ \mathrm{m}] \right\}\ .$$

In Fig. 2, simulation results are visualized. Fig. 2(a) shows a set of prior densities. Note, that only the densities with minimum, maximum, and central mean are displayed. After several measurements, the width of the interval of means, which is a one-dimensional ellipsoid, converges against the width of $\mathcal{D}$ and the stochastic uncertainty is reduced. The estimated set after 20 measurement steps is shown in Fig. 2(b). As can be seen clearly, the true ground clearance of $180\ \mathrm{m}$ is within the set.

5.2 Localization of Mobile Robot

In this example, a vehicle moving on a 2D-plane is modeled. The motion model is given by

$$\begin{bmatrix} \boldsymbol{x}_{k+1}^1 \\ \boldsymbol{x}_{k+1}^2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_k^1 \\ \boldsymbol{x}_k^2 \end{bmatrix} + (v_k + \boldsymbol{w}_{v_k}) \cdot T \cdot \begin{bmatrix} \cos(\varphi_k + \boldsymbol{w}_{\varphi_k}) \\ \sin(\varphi_k + \boldsymbol{w}_{\varphi_k}) \end{bmatrix}\ ,$$

with input vector $\underline{\hat{u}}_k = [v_k, \varphi_k]^{\mathrm{T}}$ and zero-mean Gaussian noise terms $\boldsymbol{w}_{v_k}$ and $\boldsymbol{w}_{\varphi_k}$. The robot takes distance measurements to landmarks P_i for localization, which is descibed by the measurement model

$$\hat{y}_k = \sqrt{(\boldsymbol{x}_k^1 - P_i^1)^2 + (\boldsymbol{x}_k^2 - P_i^2)^2} + \boldsymbol{v}_k\ .$$

In this example, the standard Extended Kalman Filter is compared to the Extended Kalman Filter for ellipsoidal sets. Besides the random noise, the system is affected by systematic, deterministic errors in the measurement $\hat{y}_k$ and in the system input $\underline{\hat{u}}_k$. Both errors are additive, where the measurements are each influenced by an offset of 0.1 and every system input is biased by the vector $[0.05, (\pi/360)^{1/2}]$. For simulation, they are assumed to lie within the ellipsoids $\mathcal{E}(\hat{y}_k, 0.025)$ and $\mathcal{E}(\underline{\hat{u}}_k, \mathrm{diag}(0.01, \pi/360))$, respectively. For the standard EKF, they are ignored. The unknown prior state is modeled as a credal set with the ellipsoid $\mathcal{E}(\underline{\hat{c}}_0^{\mathrm{p}}, \mathbf{X}_0^{\mathrm{p}})$ with $\underline{\hat{c}}_0^{\mathrm{p}} = [-0.4, 0.3]^{\mathrm{T}}$ and $\mathbf{X}_0^{\mathrm{p}} = \mathbf{I}$. The standard EKF is initialized with the true position $\underline{\hat{x}}_0^{\mathrm{p}}$.

Simulation results are depicted in Fig. 3. The true trajectory (blue), the estimation result of the standard Extended Kalman Filter (red, dashed), and the fusion result of the Extended Kalman Filter for ellipsoidal sets (green, dash-dotted) are shown. It can be seen that the ellipsoids of the means contain the fusion result of the standard EKF.

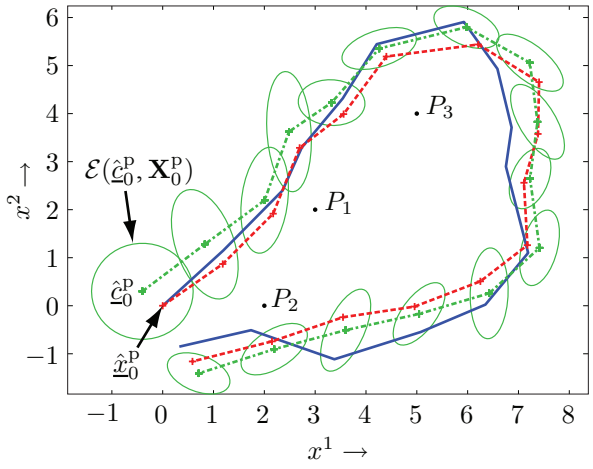


Figure 3: Localization example of Section 5.2. The Extended Kalman Filter for ellipsoidal Sets (green, dash-dotted) and the standard Extended Kalman Filter (red, dashed) are compared to the true trajectory (blue).

6 Conclusions and Future Work

Using a Bayesian estimator will not yield reliable results, if besides random noise further unknown disturbances are present. The dynamics and observations are then influenced by unknown biases, which can vary over time. This paper introduced a theoretical framework for modelling stochastic and systematic errors simultaneously. This is achieved by characterizing a random variable by a set of probability densities instead of a single density. For additive stochastic and systematic errors, these sets contain translations of one density and can be parameterized by sets of means.

In the case of linear transition and observation models, the Kalman filtering scheme was utilized and, by using ellipsoidal error bounds, an efficient estimator was obtained. Sets of Gaussian densities allows the following interpretation [11]: The covariance characterizes the possible dispersion *about* a single estimated state, whereas the set of means accounts for a possible deviation *of* the estimated state.

For nonlinear systems, a method for linearization was proposed in order to apply an extended Kalman filter. Especially for high-dimensional systems, the herein used approximation points may represent the nonlinearities insufficiently. Therefore, nonlinear set-valued state estimation will be the focus of further research. In addition, it seems promising to enclose linearization errors by sets when applying an extended Kalman filter.

Of course, incorporating systematic error bounds does not yield better estimation results, but allows insight into the sensitivity towards non-stochastic disturbances.

7 Acknowledgments

This work was partially supported by the German Research Foundation (DFG) within the Research Training Group GRK 1194 “Self-organizing Sensor-Actuator-Networks”.

References

[1] R. E. Kalman, “A new Approach to Linear Filtering and Prediction Problems,” *Transactions of the ASME, Journal of Basic Engineering*, vol. 82, pp. 35–45, 1960.

[2] H. W. Sorenson, “Least-Squares Estimation: From Gauss to Kalman,” *IEEE Spectrum*, vol. 7, pp. 63–68, 1970.

[3] 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.

[4] I. Levi, *The Enterprise of Knowledge*. MIT Press, 1983.

[5] ——, “Imprecision and Indeterminacy in Probability Judgment,” *Philosophy of Science*, vol. 52, no. 3, pp. 390–409, September 1985.

[6] F. G. Cozman, “A Brief Introduction to the Theory of Sets of Probability Measures,” Robotics Institute, Carnegie Mellon University, Tech. Rep., 1999. [Online]. Available: http://www.cs.cmu.edu/~fgcozman/qBayes.html

[7] F. C. Schweppe, *Uncertain Dynamic Systems*. Prentice–Hall, 1973.

[8] F. L. Chernousko, *State Estimation for Dynamic Systems*. CRC Press, 1994.

[9] A. Kurzhanski and I. Vályi, *Ellipsoidal Calculus for Estimation and Control*. Birkhäuser, 1997.

[10] U. D. Hanebeck and J. Horn, “Fusing Information Simultaneously Corrupted by Uncertainties with Known Bounds and Random Noise with Known Distribution,” *Information Fusion, Elsevier Science*, vol. 1, no. 1, pp. 55–63, 2000.

[11] 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, Jan. 1991.

[12] ——, “An Extended Set-Valued Kalman Filter,” in *ISIPTA*, 2003, pp. 395–405.

[13] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*. SIAM Classics in Applied Mathematics, 1995.

[14] T. Schön, F. Gustafsson, and P.-J. Nordlund, “Marginalized Particle Filters for Nonlinear State-space Models,” Division of Communication Systems, Department of Electrical Engineering, Linköpings University, Tech. Rep., Oct. 2003.