A SQUARE–ROOT ALGORITHM FOR SET THEORETIC STATE ESTIMATION

U. D. Hanebeck

Institute of Automatic Control Engineering
Technische Universität München
80290 München, Germany
fax: +49-89-289-28340
e-mail: Uwe.Hanebeck@ieee.org

Keywords: Estimation, Bounded Uncertainty and Errors in Variables, Observers, Robust Filtering, Set–membership Estimation and Identification.

Abstract

This paper presents a modified set theoretic framework for estimating the state of a linear dynamic system based on uncertain measurements. The measurement errors are assumed to be unknown but bounded by ellipsoidal sets. Based on this assumption, a recursive state estimator is (re–)derived in a tutorial fashion. It comprises both the prediction step (time update), i.e., propagation of a set of feasible states by means of the system model and the filter step (measurement update), i.e., inclusion of a new measurement into the current estimate. The main contribution is an efficient square–root formulation of this estimator, which is well suited especially for practical applications.

1 Introduction

In many applications, it is necessary to estimate the state of a dynamic system based on a linear state–space model and erroneous observations of the system output. When (erroneous) measurements of the system input are available, they can be included in the estimation process. However, often only bounds for the measurement errors can be given, since a more detailed error description is not available or too complicated. In this case, set theoretic estimation methods have found to be useful and even superior to statistical estimation methods [10, 11, 12]. Set theoretic estimation provides the set of all feasible system states that are compatible with the system model, the observations, and the error bounds. This is in sharp contrast to statistical estimation concepts, where an "optimal" point estimate is supplied together with its associated distribution. Different descriptions of the feasible sets are found in literature: Rectangular sets [1], polytopes [22], and ellipsoids [25]. This paper focuses on ellipsoidal sets, since they can be manipulated by simple matrix operations and lead to computationally efficient estimation algorithms.

Set theoretic state estimation consists of 1. propagating a set of feasible states with the aid of the system model, 2. testing the propagated set of states and the set of states defined by a new observation for common points, and 3. fusing the two sets via set intersection in case they overlap. The exact description of the resulting feasible sets has been investigated in [30]. An approximation by parallelotopes is derived in [29]. Some early results on using ellipsoidal sets and approximating the intersection of two ellipsoids is found in [21]. The topic has then been extended in [28]. Parameter set estimation based on the approximation by bounding ellipsoids is discussed in [6, 7, 24, 8, 9, 3].

State estimation, i.e., a method for state propagation and several fusion schemes, is discussed in [2]. For the case of one–dimensional observations, a simple and efficient fusion method is given in [25], which is based on the results in [6, 7].

This paper first (re–)derives a solution framework for the above mentioned problems in the case of ellipsoids, i.e., 1. the propagation of an ellipsoid and 2. the fusion of an ellipsoid and a possibly degenerated ellipsoid based on the work in [28] and then recasts the resulting algorithms into a square–root form. The resulting algorithms are simple and compact, efficient, and easy to implement. Furthermore, they can be applied to arbitrary dimensional state and observation vectors.

The estimation problem is formulated in Sec. 3. Section 4 introduces the concept of time updating by propagating the last estimate through the system model. An efficient algorithm for performing measurement updates by fusion of an ellipsoid and a possibly degenerated ellipsoid is given in Sec. 5.

2 Preliminaries

In the following, we will represent symmetric, positive definite matrices by means of their square–roots, i.e,

$$\mathbf{F} = \mathbb{F}\mathbb{F}^T \quad ,$$

where $\mathbb{F}$ is a lower triangular matrix, e.g. the cholesky factor of $\mathbf{F}$. Furthermore, the following lemma is used extensively [26]:

Lemma 2.1 For two matrices $\mathbb{F}$, $\mathbb{G}$, there exists an orthogonal matrix Θ, $\Theta\Theta^T = \mathbf{I}$, such that

$$\mathbb{F}\Theta = \mathbb{G}$$

if, and only if, $\mathbb{F}\mathbb{F}^T = \mathbb{G}\mathbb{G}^T$ holds. $\mathbb{F}$ is called the pre–array, $\mathbb{G}$ is called the post–array [23].

3 Problem Formulation

Consider a discrete–time linear dynamic system given by

$$\begin{aligned} \underline{x}_{k+1} &= \mathbf{A}_k \, \underline{x}_k + \mathbf{B}_k \, \underline{u}_k && (1)\\ \underline{z}_k &= \mathbf{H}_k \, \underline{x}_k + \underline{e}_k \ , && (2) \end{aligned}$$

with N–dimensional state vector $\underline{x}_k$ and M–dimensional observation vector $\underline{z}_k$. The initial state $\underline{x}_0$ is assumed to be confined to an ellipsoidal set given by

$$\mathcal{X}_0 = \left\{ \underline{x}_0 : (\underline{x}_0 - \underline{\hat{x}}_0)^T \mathbf{X}_0^{-1} (\underline{x}_0 - \underline{\hat{x}}_0) \leq 1 \right\} \ .$$

Observations suffer from output noise $\underline{e}_k$, which is modeled via amplitude bounds according to $\underline{e}_k \in \mathcal{E}_k$, where $\mathcal{E}_k$ is an ellipsoid given by

$$\mathcal{E}_k = \left\{ \underline{e}_k : \underline{e}_k^T \mathbf{E}_k^{-1} \underline{e}_k \leq 1 \right\} \ . \tag{3}$$

In addition, the system input $\underline{u}_k$ is uncertain and also modeled by amplitude bounds $\underline{u}_k \in \mathcal{U}_k$ where $\mathcal{U}_k$ is an ellipsoid given by

$$\mathcal{U}_k = \left\{ \underline{u}_k : \underline{u}_k^T \mathbf{U}_k^{-1} \underline{u}_k \leq 1 \right\} \ . \tag{4}$$

When an uncertain measurement $\underline{\hat{u}}_k$ of the system input is available, the uncertainty set (4) will be centered around $\underline{\hat{u}}_k$ according to

$$\mathcal{U}_k = \left\{ \underline{u}_k : (\underline{u}_k - \underline{\hat{u}}_k)^T \mathbf{U}_k^{-1} (\underline{u}_k - \underline{\hat{u}}_k) \leq 1 \right\} \ . \tag{5}$$

The uncertainty models (3), (5) are more general than component–wise error bounds, since they also include correlation between variables. The sizes of the uncertainty ellipsoids are assumed to be a priori known for the purpose of this paper.

The state estimation or filtering task consists of calculating the set of feasible system states at time k denoted by $\mathcal{S}_k$ based on observations $\underline{z}_0, \underline{z}_1, \ldots, \underline{z}_k$. Rather than performing batch processing, an estimate should be made available at every time k. The following notation [1] is used:

[1]This notation is preferred to $\mathcal{X}_k^P$ $\mathcal{X}_k^M$, $\mathcal{X}_k^S$ to avoid excessive sub–/superscripting.

$\mathcal{P}_k$ denotes the set of *predicted* system states at time k with

$$\mathcal{P}_k = \left\{ \underline{x} : \left(\underline{x} - \underline{\hat{p}}_k\right)^T (\mathbf{P}_k)^{-1} \left(\underline{x} - \underline{\hat{p}}_k\right) \leq 1 \right\} \ .$$

$\mathcal{M}_k$ denotes the set of system states *defined by an observation* at time k, and $\mathcal{S}_k$ denotes the set of *estimated* system states at time k according to

$$\mathcal{S}_k = \left\{ \underline{x} : (\underline{x} - \underline{\hat{s}}_k)^T (\mathbf{S}_k)^{-1} (\underline{x} - \underline{\hat{s}}_k) \leq 1 \right\} \ .$$

The system equation (1) is used to predict all feasible system states $\mathcal{P}_k$ at time k based on an estimated system state $\mathcal{S}_{k-1}$ at time $k-1$ and the system input $\mathcal{U}_{k-1}$. A new observation $\underline{z}_k$ at time k defines a set $\mathcal{M}_k$ of states that could possibly cause the observation. Thus, at time k, two state estimates are available: The set of predicted states $\mathcal{P}_k$ and the set of states $\mathcal{M}_k$ compatible with the observation. Set theoretic estimation now calculates the set of feasible states via set intersection [5], i.e., the exact set of estimated system states $\mathcal{S}_k$ at time k is given by

$$\mathcal{S}_k = \mathcal{P}_k \cap \mathcal{M}_k \ .$$

Neither the set of predicted states $\mathcal{P}_k$ nor the fusion result $\mathcal{S}_k$ are in general ellipsoids. Hence, the remainder of this paper will be concerned with the efficient calculation of outer bounding ellipsoids for $\mathcal{P}_k$, $\mathcal{S}_k$ to arrive at a simple recursive estimation scheme similar to the Kalman filter.

4 Prediction Step (Time Update)

Given a set of estimated system states $\mathcal{S}_{k-1}$ at time $k-1$, the system equation (1) is used to predict all feasible system states $\mathcal{P}_k$ at time k (time update). Since ellipsoids are closed under affine transformations, the right hand side of (1) can be interpreted as the summation of two ellipsoids with centers $\mathbf{A}_{k-1} \, \underline{\hat{s}}_{k-1}$, $\mathbf{B}_{k-1} \, \underline{\hat{u}}_{k-1}$ and definition matrices $\mathbf{A}_{k-1}\mathbf{S}_{k-1}\mathbf{A}_{k-1}^T$, $\mathbf{B}_{k-1}\mathbf{U}_{k-1}\mathbf{B}_{k-1}^T$, respectively. Two ellipsoidal sets are added up by Minkowski summation, which is defined by adding every point contained in the first ellipsoid to all the points contained in the second ellipsoid. However, the Minkowski sum of two ellipsoids is not in general an ellipsoid. Nevertheless, a family of bounding ellipsoids can be found that contain the Minkowski sum [28]. The center is always given by the sum of centers

$$\underline{\hat{p}}_k = \mathbf{A}_{k-1}\underline{\hat{s}}_{k-1} + \mathbf{B}_{k-1}\underline{\hat{u}}_{k-1} \tag{6}$$

and the matrix $\mathbf{P}_k$ is given by

$$\begin{aligned} \mathbf{P}_k =& (0.5 - \kappa_k)^{-1} \mathbf{A}_{k-1}\mathbf{S}_{k-1}\mathbf{A}_{k-1}^T \\ &+ (0.5 + \kappa_k)^{-1} \mathbf{B}_{k-1}\mathbf{U}_{k-1}\mathbf{B}_{k-1}^T \end{aligned} \tag{7}$$

with parameter κ_k for $\kappa_k \in (-0.5, 0.5)$.

Proof: Given a direction vector $\underline{t}$ of unit length, i.e, $\|\underline{t}\| = 1$, and tangential planes for both the Minkowski sum and the bounding ellipsoid, for which $\underline{t}$ is the normal vector. Then, the normal distances from the center (6) to the planes are given by

$$\begin{aligned} p_{\text{Minkowski}} =& \sqrt{\underline{t}^T \mathbf{A}_{k-1}\mathbf{S}_{k-1}\mathbf{A}_{k-1}^T \underline{t}} \\ &+ \sqrt{\underline{t}^T \mathbf{B}_{k-1}\mathbf{U}_{k-1}\mathbf{B}_{k-1}^T \underline{t}} \end{aligned}$$

for the Minkowski sum and by

$$\begin{aligned} p_{\text{bound}} &= \sqrt{\underline{t}^T \mathbf{P}_k \underline{t}} \\ &= \sqrt{\begin{array}{l} (0.5-\kappa_k)^{-1}\underline{t}^T \mathbf{A}_{k-1}\mathbf{S}_{k-1}\mathbf{A}_{k-1}^T \underline{t} \\ +(0.5+\kappa_k)^{-1}\underline{t}^T \mathbf{B}_{k-1}\mathbf{U}_{k-1}\mathbf{B}_{k-1}^T \underline{t} \end{array}} \end{aligned}$$

for the bounding ellipsoid. $\mathbf{P}_k$ is in fact the definition matrix of a bounding ellipsoid, since

$$p_{\text{Minkowski}} \leq p_{\text{bound}} \tag{8}$$

for every $\underline{t}$, $\|\underline{t}\| = 1$. This is easily verified by squaring both sides of (8), which leads to

$$\begin{aligned} &\left(\sqrt{\underline{t}^T \mathbf{A}_{k-1}\mathbf{S}_{k-1}\mathbf{A}_{k-1}^T \underline{t}} + \sqrt{\underline{t}^T \mathbf{B}_{k-1}\mathbf{U}_{k-1}\mathbf{B}_{k-1}^T \underline{t}}\right)^2 \\ &\quad \leq (0.5-\kappa_k)^{-1}\underline{t}^T \mathbf{A}_{k-1}\mathbf{S}_{k-1}\mathbf{A}_{k-1}^T \underline{t} \\ &\quad +(0.5+\kappa_k)^{-1}\underline{t}^T \mathbf{B}_{k-1}\mathbf{U}_{k-1}\mathbf{B}_{k-1}^T \underline{t} \ . \end{aligned} \tag{9}$$

Hölder's inequality is given by

$$\sum_{i=1}^N |a_i b_i| \leq \left(\sum_{i=1}^N |a_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^N |b_i|^q\right)^{\frac{1}{q}}$$

for $p^{-1} + q^{-1} = 1$, $p > 1$, $q > 1$. Squaring both sides and setting $p = 2$, $q = 2$, Cauchy's inequality is obtained as

$$\left(\sum_{i=1}^N |a_i b_i|\right)^2 \leq \left(\sum_{i=1}^N |a_i|^2\right) \left(\sum_{i=1}^N |b_i|^2\right) \ .$$

The special case of $a_i = \alpha_i^{\frac{1}{2}}$, $b_i = \alpha_i^{-\frac{1}{2}}\beta_i$ gives

$$\left(\sum_{i=1}^N |\beta_i|\right)^2 \leq \left(\sum_{i=1}^N |\alpha_i|\right) \left(\sum_{i=1}^N |\alpha_i^{-1}|\beta_i^2\right) \ .$$

Setting $N = 2$, $\beta_1 = \sqrt{\underline{t}^T \mathbf{A}_{k-1}\mathbf{S}_{k-1}\mathbf{A}_{k-1}^T \underline{t}}$, $\beta_2 = \sqrt{\underline{t}^T \mathbf{B}_{k-1}\mathbf{U}_{k-1}\mathbf{B}_{k-1}^T \underline{t}}$, $\alpha_1 = 0.5 - \kappa_k$, $\alpha_2 = 0.5 + \kappa_k$ yields (9) and concludes the proof. □

Remark 4.1 κ_k may be selected in some optimum way, for example in such a way as to minimize the volume of $\mathbf{P}_k$.

We define lower triangular square–root factors according to

$$\mathbf{S}_{k-1} = \mathbb{S}_{k-1}\left(\mathbb{S}_{k-1}\right)^T$$

and

$$\mathbf{P}_k = \mathbb{P}_k\left(\mathbb{P}_k\right)^T \ .$$

Theorem 4.1 A square–root algorithm for the prediction step (6), (7) is given by

$$\begin{aligned} &\begin{bmatrix} (0.5-\kappa_k)^{-\frac{1}{2}}\mathbf{A}_{k-1}\mathbb{S}_{k-1} & (0.5+\kappa_k)^{-\frac{1}{2}}\mathbf{B}_{k-1}\mathbb{U}_{k-1} \\ (0.5-\kappa_k)^{\frac{1}{2}}\underline{\hat{s}}_{k-1}^T\left(\mathbb{S}_{k-1}^{-1}\right)^T & (0.5+\kappa_k)^{\frac{1}{2}}\underline{\hat{u}}_{k-1}^T\left(\mathbb{U}_{k-1}^{-1}\right)^T \end{bmatrix}\Theta_1 \\ &\quad = \begin{bmatrix} \mathbb{P}_k & \mathbf{0} \\ \underline{\hat{p}}_k^T\left(\mathbb{P}_k^{-1}\right)^T & * \end{bmatrix} \ , \end{aligned} \tag{10}$$

where Θ_1 is an arbitrary orthogonal rotation matrix with $\Theta_1\Theta_1^T = \mathbf{I}$, that performs the desired triangularization. $*$ denotes elements, that are not of interest for the prediction.

Proof: According to Lemma 2.1, "squaring" the left–hand–side of (10) leads to

$$\begin{aligned} &\begin{bmatrix} (0.5-\kappa_k)^{-\frac{1}{2}}\mathbf{A}_{k-1}\mathbb{S}_{k-1} & (0.5+\kappa_k)^{-\frac{1}{2}}\mathbf{B}_{k-1}\mathbb{U}_{k-1} \\ (0.5-\kappa_k)^{\frac{1}{2}}\underline{\hat{s}}_{k-1}^T\left(\mathbb{S}_{k-1}^{-1}\right)^T & (0.5+\kappa_k)^{\frac{1}{2}}\underline{\hat{u}}_{k-1}^T\left(\mathbb{U}_{k-1}^{-1}\right)^T \end{bmatrix}\Theta_1 \\ &\Theta_1^T\begin{bmatrix} (0.5-\kappa_k)^{-\frac{1}{2}}\mathbb{S}_{k-1}^T\mathbf{A}_{k-1}^T & (0.5-\kappa_k)^{\frac{1}{2}}\mathbb{S}_{k-1}^{-1}\underline{\hat{s}}_{k-1} \\ (0.5+\kappa_k)^{-\frac{1}{2}}\mathbb{U}_{k-1}^T\mathbf{B}_{k-1}^T & (0.5+\kappa_k)^{\frac{1}{2}}\mathbb{U}_{k-1}^{-1}\underline{\hat{u}}_{k-1} \end{bmatrix} \\ &\quad = \begin{bmatrix} K_{k-1}^{11} & K_{k-1}^{12} \\ K_{k-1}^{21} & K_{k-1}^{22} \end{bmatrix} \end{aligned}$$

with

$$\begin{aligned} K_{k-1}^{11} =& (0.5-\kappa_k)^{-1}\mathbf{A}_{k-1}\mathbf{S}_{k-1}\mathbf{A}_{k-1}^T \\ &+(0.5+\kappa_k)^{-1}\mathbf{B}_{k-1}\mathbf{U}_{k-1}\mathbf{B}_{k-1}^T \\ K_{k-1}^{12} =& \mathbf{A}_{k-1}\underline{\hat{s}}_{k-1} + \mathbf{B}_{k-1}\underline{\hat{u}}_{k-1} \\ K_{k-1}^{21} =& \underline{\hat{s}}_{k-1}^T\mathbf{A}_{k-1}^T + \underline{\hat{u}}_{k-1}^T\mathbf{B}_{k-1}^T \\ K_{k-1}^{22} =& (0.5-\kappa_k)\underline{\hat{s}}_{k-1}^T\mathbf{S}_{k-1}^{-1}\underline{\hat{s}}_{k-1} \\ &+(0.5+\kappa_k)\underline{\hat{u}}_{k-1}^T\mathbf{U}_{k-1}^{-1}\underline{\hat{u}}_{k-1} \ . \end{aligned}$$

"Squaring" the right–hand–side of (10) gives

$$\begin{bmatrix} \mathbb{P}_k & \mathbf{0} \\ \underline{\hat{p}}_k^T\left(\mathbb{P}_k^{-1}\right)^T & * \end{bmatrix}\begin{bmatrix} \mathbb{P}_k^T & \mathbb{P}_k^{-1}\underline{\hat{p}}_k \\ \mathbf{0} & * \end{bmatrix} = \begin{bmatrix} \mathbf{P}_k & \underline{\hat{p}}_k \\ \underline{\hat{p}}_k^T & \underline{\hat{p}}_k^T\mathbf{P}_k^{-1}\underline{\hat{p}}_k + * \end{bmatrix} \ .$$

Comparing the elements completes the proof. $\square$

The elements of the resulting post–array (10) are then used for calculating the desired prediction ellipsoid $\mathcal{P}_k$ given by the center $\hat{\underline{p}}_k$ and the definition matrix $\mathbf{P}_k$. For that purpose, we use boxes to denote the block matrices

$$\boxed{\mathbb{P}_k} \text{ and } \boxed{\hat{\underline{p}}_k^T \left(\mathbb{P}_k^{-1}\right)^T}$$

that are directly obtained from the post–array in (10). The center and the definition matrix are then calculated according to

$$\begin{aligned}\hat{\underline{p}}_k &= \boxed{\mathbb{P}_k}\boxed{\hat{\underline{p}}_k^T \left(\mathbb{P}_k^{-1}\right)^T}^T \quad ,\\ \mathbf{P}_k &= \boxed{\mathbb{P}_k}\boxed{\mathbb{P}_k}^T \quad .\end{aligned}$$

5 Fusion Step (Measurement Update)

An observation $\underline{z}_k$ at time k defines a set $\mathcal{M}_k$ of states given by

$$\mathcal{M}_k = \{\underline{m}_k : \underline{z}_k - \mathbf{H}_k\underline{m}_k \in \mathcal{E}_k\}$$

according to (2) or equivalently by

$$\mathcal{M}_k = \left\{\underline{m}_k : \left(\underline{z}_k - \mathbf{H}_k\underline{m}_k\right)^T \mathbf{E}_k^{-1} \left(\underline{z}_k - \mathbf{H}_k\underline{m}_k\right) \leq 1\right\} \quad .$$

Calculation of the minimum–volume bounding ellipsoid for the intersection of $\mathcal{P}_k$ and $\mathcal{M}_k$ requires numerical optimization. Here, a suboptimal method is pursued, that 1. provides a whole family of bounding ellipsoids and 2. yields recursion equations similar to the Kalman filter. This method has first been reported in [28] and consists of bounding the intersection set by the convex combination of $\mathcal{P}_k$ and $\mathcal{M}_k$ according to

$$\begin{aligned}\mathcal{S}_k = \Big\{\underline{s}_k : &(0.5 - \lambda_k^*)\left(\underline{z}_k - \mathbf{H}_k\underline{s}_k\right)^T \mathbf{E}_k^{-1}\left(\underline{z}_k - \mathbf{H}_k\underline{s}_k\right)\\ &+(0.5 + \lambda_k^*)(\underline{s}_k - \hat{\underline{p}}_k)^T\mathbf{P}_k^{-1}(\underline{s}_k - \hat{\underline{p}}_k)\Big\}\end{aligned} \tag{11}$$

After some manipulations and bilinear–transformation $\lambda_k = \dfrac{0.5+\lambda_k^*}{0.5-\lambda_k^*}$, where $\lambda_k \in [0,\infty]$ and $\lambda_k^* \in [-0.5, 0.5]$, $\mathcal{S}_k$ can be written in the following feedback form.

Lemma 5.1 A bounding ellipsoid for the intersection of $\mathcal{P}_k$ and $\mathcal{M}_k$ is given by

$$\begin{aligned}\hat{\underline{s}}_k &= \hat{\underline{p}}_k + \lambda_k\mathbf{P}_k\mathbf{H}_k^T\mathbf{R}_k^{-1}\hat{\underline{\epsilon}}_k\\ \mathbf{S}_k &= d_k\mathbf{C}_k\end{aligned}$$

with

$$\begin{aligned}d_k &= 1 + \lambda_k - \lambda_k\hat{\underline{\epsilon}}_k^T\mathbf{R}_k^{-1}\hat{\underline{\epsilon}}_k\\ \mathbf{C}_k &= \mathbf{P}_k - \lambda_k\mathbf{P}_k\mathbf{H}_k^T\mathbf{R}_k^{-1}\mathbf{H}_k\mathbf{P}_k\\ \mathbf{R}_k &= \mathbf{E}_k + \lambda_k\mathbf{H}_k\mathbf{P}_k\mathbf{H}_k^T\\ \hat{\underline{\epsilon}}_k &= \underline{z}_k - \mathbf{H}_k\hat{\underline{p}}_k\end{aligned}$$

for $\lambda_k \in [0,\infty]$.

Remark 5.1 Obviously, we have

$$(\mathcal{P}_k \cap \mathcal{M}_k) \subset \mathcal{S}_k \subset (\mathcal{P}_k \cup \mathcal{M}_k) \quad ,$$

i.e., the bounding ellipsoid $\mathcal{S}_k$ not only contains the intersection of $\mathcal{P}_k$ and $\mathcal{M}_k$, but is also itself contained in their union.

Theorem 5.1 A square–root algorithm for the fusion step (11) is given by:

$$\begin{aligned}&\begin{bmatrix}\mathbb{E}_k & \lambda_k^{\frac{1}{2}}\mathbf{H}_k\mathbb{P}_k\\ \mathbf{0} & \mathbb{P}_k\\ -\lambda_k^{\frac{1}{2}}\underline{z}_k^T\left(\mathbb{E}_k^{-1}\right)^T & \hat{\underline{p}}_k^T\left(\mathbb{P}_k^{-1}\right)^T\end{bmatrix}\Theta_2\\ &= \begin{bmatrix}\mathbb{R}_k & \mathbf{0}\\ \lambda_k^{\frac{1}{2}}\mathbf{P}_k\mathbf{H}_k^T\left(\mathbb{R}_k^{-1}\right)^T & \mathbb{C}_k\\ -\lambda_k^{\frac{1}{2}}\hat{\underline{\epsilon}}_k^T\left(\mathbb{R}_k^{-1}\right)^T & \hat{\underline{s}}_k^T\left(\mathbb{C}_k^{-1}\right)^T\end{bmatrix}\end{aligned} \tag{12}$$

with $\mathbf{R}_k = \mathbb{R}_k\,\mathbb{R}_k^T$, $\mathbf{C}_k = \mathbb{C}_k\,\mathbb{C}_k^T$, ans $\mathbf{E}_k = \mathbb{E}_k\,\mathbb{E}_k^T$.

Proof: Again, according to Lemma 2.1, the left–hand–side of (12) is "squared"

$$\begin{bmatrix}\mathbb{E}_k & \lambda_k^{\frac{1}{2}}\mathbf{H}_k\mathbb{P}_k\\ \mathbf{0} & \mathbb{P}_k\\ -\lambda_k^{\frac{1}{2}}\underline{z}_k^T\left(\mathbb{E}_k^{-1}\right)^T & \hat{\underline{p}}_k^T\left(\mathbb{P}_k^{-1}\right)^T\end{bmatrix}\Theta_2$$

$$\Theta_2^T\begin{bmatrix}\mathbb{E}_k^T & \mathbf{0} & -\lambda_k^{\frac{1}{2}}\mathbb{E}_k^{-1}\underline{z}_k\\ \lambda_k^{\frac{1}{2}}\mathbb{P}_k^T\mathbf{H}_k^T & \mathbb{P}_k^T & \mathbb{P}_k^{-1}\hat{\underline{p}}_k\end{bmatrix} =$$

$$\begin{bmatrix}\mathbf{E}_k + \lambda_k\mathbf{H}_k\mathbf{P}_k\mathbf{H}_k^T & \lambda_k^{\frac{1}{2}}\mathbf{H}_k\mathbf{P}_k & -\lambda_k^{\frac{1}{2}}\left(\underline{z}_k - \mathbf{H}_k\hat{\underline{p}}_k\right)\\ \lambda_k^{\frac{1}{2}}\mathbf{P}_k\mathbf{H}_k^T & \mathbf{P}_k & \hat{\underline{p}}_k\\ -\lambda_k^{\frac{1}{2}}\left(\underline{z}_k^T - \hat{\underline{p}}_k^T\mathbf{H}_k^T\right) & \hat{\underline{p}}_k^T & \hat{\underline{p}}_k^T\mathbf{P}_k^{-1}\hat{\underline{p}}_k - \lambda_k\underline{z}_k^T\mathbf{E}_k^{-1}\underline{z}_k\end{bmatrix}$$

and compared with the "squared" right–hand–side of (12)

$$
\begin{bmatrix}
\mathbb{R}_k & \mathbf{0} \\
\lambda_k^{\frac{1}{2}} \mathbf{P}_k \mathbf{H}_k^T \left(\mathbb{R}_k^{-1}\right)^T & \mathbb{C}_k \\
-\lambda_k^{\frac{1}{2}} \underline{\hat{\epsilon}}_k^T \left(\mathbb{R}_k^{-1}\right)^T & \underline{\hat{s}}_k^T \left(\mathbb{C}_k^{-1}\right)^T
\end{bmatrix}
\begin{bmatrix}
\mathbb{R}_k^T & \lambda_k^{\frac{1}{2}} \mathbb{R}_k^{-1} \mathbf{H}_k \mathbf{P}_k & -\lambda_k^{\frac{1}{2}} \mathbb{R}_k^{-1} \underline{\hat{\epsilon}}_k \\
\mathbf{0} & \mathbb{C}_k^T & \mathbb{C}_k^{-1} \underline{\hat{s}}_k
\end{bmatrix} =
$$

$$
\begin{bmatrix}
\mathbf{R}_k & \lambda_k^{\frac{1}{2}} \mathbf{H}_k \mathbf{P}_k & -\lambda_k^{\frac{1}{2}} \underline{\hat{\epsilon}}_k \\
\lambda_k^{\frac{1}{2}} \mathbf{P}_k \mathbf{H}_k^T & \mathbf{C}_k + \lambda_k \mathbf{P}_k \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{P}_k & \underline{\hat{s}}_k - \lambda_k \mathbf{P}_k \mathbf{H}_k^T \mathbf{R}_k^{-1} \underline{\hat{\epsilon}}_k \\
-\lambda_k^{\frac{1}{2}} \underline{\hat{\epsilon}}_k^T & \underline{\hat{s}}_k^T - \lambda_k \underline{\hat{\epsilon}}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{P}_k & \underline{\hat{s}}_k^T \mathbf{C}_k^{-1} \underline{\hat{s}}_k + \lambda_k \underline{\hat{\epsilon}}_k^T \mathbf{R}_k^{-1} \underline{\hat{\epsilon}}_k
\end{bmatrix}
$$

to prove (12). □

Again, we use boxes to denote the block matrices

$$
\boxed{\mathbb{C}_k}\,, \quad \boxed{\underline{\hat{s}}_k^T \left(\mathbb{C}_k^{-1}\right)^T}\,, \quad \text{and} \quad \boxed{-\lambda_k^{\frac{1}{2}} \underline{\hat{\epsilon}}_k^T \left(\mathbb{R}_k^{-1}\right)^T}
$$

that are directly obtained from the post–array in (12). The center and the definition matrix of the fusion ellipsoid $\mathcal{S}_k$ are then calculated on the basis of these block matrices according to

$$
\begin{aligned}
\underline{\hat{s}}_k &= \boxed{\mathbb{C}_k} \boxed{\underline{\hat{s}}_k^T \left(\mathbb{C}_k^{-1}\right)^T}^T \quad , \\
\mathbf{S}_k &= d_k \boxed{\mathbb{C}_k} \boxed{\mathbb{C}_k}^T \quad , \\
d_k &= 1 + \lambda_k - \boxed{-\lambda_k^{\frac{1}{2}} \underline{\hat{\epsilon}}_k^T \left(\mathbb{R}_k^{-1}\right)^T} \boxed{-\lambda_k^{\frac{1}{2}} \underline{\hat{\epsilon}}_k^T \left(\mathbb{R}_k^{-1}\right)^T}^T \quad .
\end{aligned}
$$

6 Recursive State Estimation

At every time k, the set of feasible states is predicted using (10). In addition, when a new observation $\underline{z}_k$ is available, a fusion step according to (12) is subsequently performed. The resulting recursion equations are similar in appearance to the Kalman filter equations. However, the resulting state estimates are different because of the different uncertainty model.

At first glance, four inversions of square–root factors $\mathbb{U}_k^{-1}$, $\mathbb{E}_k^{-1}$, $\mathbb{P}_k^{-1}$, $\mathbb{S}_k^{-1}$ must be performed for setting up the left–hand–side block matrices for prediction and fusion, respectively. However, in fact only $\mathbb{U}_k^{-1}$ and $\mathbb{E}_k^{-1}$ must be calculated, since $\boxed{\underline{\hat{p}}_k^T \left(\mathbb{P}_k^{-1}\right)^T}$ is available as a by–product from the prediction post–array. Furthermore, the lower left block matrix of the prediction pre–array is calculated on the basis of the elements of the fusion post–array according to

$$
\begin{aligned}
&\boxed{(0.5 - \kappa_k)^{\frac{1}{2}} \underline{\hat{s}}_{k-1}^T \left(\mathbb{S}_{k-1}^{-1}\right)^T} \\
&= (0.5 - \kappa_k)^{\frac{1}{2}} \frac{1}{d_k} \boxed{\underline{\hat{s}}_{k-1}^T \left(\mathbb{C}_{k-1}^{-1}\right)^T} \quad .
\end{aligned}
$$

Some technical details like the determination of overlap of $\mathcal{P}_k$ and $\mathcal{M}_k$ and the calculation of κ_k, λ_k for obtaining minimum volume bounding ellipsoids at every time step have been omitted for the sake of brevity.

It is important to note, that the calculation of a minimum volume ellipsoid at each time step allows for recursive computation. Hence, it is useful from a practical point of view. However, doing so does not mean that the final ellipsoid recursively obtained after several time steps is of minimum volume. For calculating a minimum volume ellipsoid after several time steps, all previous measurements have to be reconsidered.

7 Conclusions

Efficient approximate set theoretic state estimators have been derived in square–root form by generalizing the results for Kalman filtering reported in [23]. The estimators apply to linear state–space models for the case, that both system and observation uncertainties are modeled as unknown but bounded by ellipsoidal sets. Besides having numerical advantages, the square–root form allows a parallel or even decentralized implementation of set theoretic estimators similar to the treatment of Kalman filtering in [4].

The proposed square–root algorithms are also useful in the context of the unification of stochastic and set theoretic estimators proposed in [13] – [20].

References

[1] S. Atiya and G. D. Hager, "Real–Time Vision–Based Localization", *IEEE Transactions on Robotics and Automation*, Vol. 9, No. 6, pp. 785–800, 1993.

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

[3] M.-F. Cheung, S. Yurkovich, and K. M. Passino, "An Optimal Volume Ellipsoid Algorithm for Parameter Set Identification", *IEEE Transactions on Automatic Control*, Vol.38, No. 8, pp. 1292–1296, 1993.

[4] T. M. Chin, W. C. Karl, and A. S. Willsky, "A Distributed and Iterative Method for Square Root Filtering in Space–time Estimation", *automatica*, Vol. 31, No. 1, pp. 67–82, 1995.

[5] P. L. Combettes, "The Foundations of Set Theoretic Estimation", *Proceedings of the IEEE*, Vol. 81, No. 2, pp. 182–208, 1993.

[6] J. R. Deller, "Set Membership Identification in Digital Signal Processing", *IEEE ASSP Magazine*, Vol. 6, pp. 4–20, 1989.

[7] J. R. Deller and T. C. Luk, "Linear Prediction Analysis of Speech Based on Set–Membership Theory",

Computer Speech and Language, Vol. 3, pp. 301–327, 1989.

[8] J. R. Deller, M. Nayeri, and S. F. Odeh, "Least–Squares Identification with Error Bounds for Real–Time Signal Processing and Control", *Proceedings of the IEEE*, Vol. 81, No. 6, pp. 815–849, 1993.

[9] J. R. Deller and S. F. Odeh, "Adaptive Set–Membership Identification in O(m) Time for Linear–in–Parameters Models", *IEEE Transactions on Signal Processing*, Vol. 41, No. 5, pp. 1906–1924, 1993.

[10] G. D. Hager, S. P. Engelson, and S. Atiya, "On Comparing Statistical and Set–Based Methods in Sensor Data Fusion", *Proceedings of the 1993 IEEE International Conference on Robotics and Automation, Atlanta, GA*, pp. 352–358, 1993.

[11] U. D. Hanebeck and G. Schmidt, "Set–theoretic Localization of Fast Mobile Robots Using an Angle Measurement Technique", *Proceedings of the 1996 IEEE International Conference on Robotics and Automation, Minneapolis, Minnesota*, pp. 1387–1394, 1996.

[12] U. D. Hanebeck and G. Schmidt, "Mobile Robot Localization Based on Efficient Processing of Sensor Data and Set–theoretic State Estimation", in A. Casals and A. T. de Almeida, eds., *Experimental Robotics V, The Fifth International Symposium, Barcelona, Catalonia, 1997, Lecture Notes in Control and Information Sciences 223*, pp. 385–396, Springer, 1998.

[13] U. D. Hanebeck and J. Horn, "A New State Estimator for a Mixed Stochastic and Set Theoretic Uncertainty Model", *Proceedings of SPIE Vol. 3720*, AeroSense Symposium, Orlando, Florida, USA, 1999.

[14] U. D. Hanebeck and J. Horn, "A New Estimator for Mixed Stochastic and Set Theoretic Uncertainty Models Applied to Mobile Robot Localization", *Proceedings of the 1999 IEEE International Conference on Robotics and Automation (ICRA'99)*, Detroit, Michigan, USA, 1999.

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

[16] U. D. Hanebeck and J. Horn, "New Estimators for Mixed Stochastic and Set Theoretic Uncertainty Models: The Vector Case", *Proceedings of the Fifth European Control Conference (ECC'99)*, Karlsruhe, Germany, 1999.

[17] U. D. Hanebeck and J. Horn, "New Results for State Estimation in the Presence of Mixed Stochastic and Set Theoretic Uncertainties", *Proceedings of the 1999 IEEE Systems, Man, and Cybernetics Conference (SMC'99)*, Tokyo, Japan, 1999.

[18] U. D. Hanebeck, J. Horn, "A New Concept for State Estimation in the Presence of Stochastic and Set Theoretic Uncertainties, *Proceedings of the 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'99)*, Kjongju, Korea, 1999.

[19] U. D. Hanebeck, J. Horn, "New Estimators for Mixed Stochastic and Set Theoretic Uncertainty Models: The Scalar Measurement Case", *Proceedings of the 1999 IEEE Conference on Decision and Control (CDC'99)*, Phoenix, Arizona, 1999.

[20] U. D. Hanebeck, 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.

[21] W. Kahan, "Circumscribing An Ellipsoid About the Intersection of Two Ellipsoids", *Canad. Math. Bull.*, Vol. 11, No. 3, pp. 437–441, 1968.

[22] M. Milanese and G. Belaforte, "Estimation Theory and Uncertainty Intervals Evaluation in the Presence of Unknown but Bounded Errors: Linear Families of Models and Estimates", *IEEE Transactions on Automatic Control*, Vol. 27, pp. 408–414, 1982.

[23] P. Park and T. Kailath, "New Square–Root Algorithms for Kalman Filtering", *IEEE Transactions on Automatic Control*, Vol. 40, No. 5, pp. 895–899, 1995.

[24] A. K. Rao, Y.-F. Huang, and S. Dasgupta, "ARMA Parameter Estimation Using a Novel Recursive Estimation Algorithm with Selective Updating", *IEEE Transactions on Signal Processing*, Vol. 38, No. 3, pp. 447–457, 1990.

[25] A. Sabater and F. Thomas, "Set Membership Approach to the Propagation of Uncertain Geometric Information" *Proceedings of the 1991 IEEE International Conference on Robotics and Automation, Sacramento, CA*, pp. 2718–2723, 1991.

[26] A. H. Sayed and T. Kailath, "A State–Space Approach to Adaptive RLS Filtering", *IEEE Signal Processing Magazine*, Vol. 11, No. 3, pp. 18–70, 1994.

[27] F. C. Schweppe, "Recursive State Estimation: Unknown but Bounded Errors and System Inputs", *IEEE Transactions on Automatic Control*, Vol. 13, pp. 22–28, 1968.

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

[29] A. Vicino and G. Zappa, "Sequential Approximation of Feasible Parameter Sets for Identification with Set Membership Uncertainty", *IEEE Transactions on Automatic Control*, Vol. 41, No. 6, pp. 774–785, 1996.

[30] E. Walter and H. Piet-Lahanier, "Exact Recursive Polyhedral Description of the Feasible Parameter Set for Bounded Error Models", *IEEE Transactions on Automatic Control*, Vol. 34, No. 8, pp. 911–915, 1989.