Optimal Kalman Gains for Combined Stochastic and Set-Membership State Estimation

Benjamin Noack, Florian Pfaff, and Uwe D. Hanebeck

Abstract— **In state estimation theory, two directions are mainly followed in order to model disturbances and errors. Either uncertainties are modeled as stochastic quantities or they are characterized by their membership to a set. Both approaches have distinct advantages and disadvantages making each one inherently better suited to model different sources of estimation uncertainty. This paper is dedicated to the task of combining stochastic and set-membership estimation methods. A Kalman gain is derived that minimizes the mean squared error in the presence of both stochastic and additional unknown but bounded uncertainties, which are represented by Gaussian random variables and ellipsoidal sets, respectively. As a result, a generalization of the well-known Kalman filtering scheme is attained that reduces to the standard Kalman filter in the absence of set-membership uncertainty and that otherwise becomes the intersection of sets in case of vanishing stochastic uncertainty. The proposed concept also allows to prioritize either the minimization of the stochastic uncertainty or the minimization of the set-membership uncertainty.**

I. INTRODUCTION

In general, neither an exact model of the system evolution nor the precise properties of the measurement devices can be identified. Also, external influences cannot be taken into account in their entirety. A central challenge in state estimation theory therefore consists in defining appropriate models of uncertainty. Employing uncertainty models can significantly contribute to ensuring robustness and reliability in decision and control applications, but for this purpose, it is necessary to propagate and update uncertainty descriptions throughout the entire state estimation process. Essentially, two different directions have been pursued in estimation theory: Bayesian and set-membership uncertainty models.

Within the Bayesian framework, the Kalman filter [1] is the most well-known example as it embodies an optimal Bayesian solution to the state estimation problem when system dynamics and observation models are linear and perturbations are normally distributed. The estimate and the corresponding mean squared error (MSE) matrix are then directly related to the mean and covariance matrix of a Gaussian random variable. Both parameters can be calculated in closed form. The Kalman filtering scheme is also well-accepted for nonlinear estimation problems. Although nonlinear Bayesian estimation generally requires approximate solutions in order to propagate and update the non-Gaussian conditional probability densities, the Kalman filter still provides a minimum MSE solution if the first two moments of the state and error distributions are known. It is a best linear unbiased estimator (BLUE). In order to compute the first- and second-order statistics at least approximately, nonlinear system and observation models are usually linearized either by a Taylor series expansion, as it is done within the extended Kalman filter, or by a linear regression analysis, for which the unscented Kalman filter [2] is a candidate. Stochastic error characteristics are beneficial to model potentially unbounded disturbances, such as outliers.

Except for a bounding set, set-membership state estimators do not require the definition of a certain error characteristic or a distribution, i.e., no knowledge about the error behavior within the bounds is needed. A set-membership error description can be employed for uncertainties that are difficult to identify in terms of a probability distribution or that even do not reveal a probabilistic nature. Common implementations of set-membership estimation algorithms employ interval-based bounds [3] or ellipsoidal sets [4]–[7]. Both representations enable efficient computation of, for instance, linear transformations, Minkowski sums, and intersections. Estimation results constitute sets, which the actual system state is considered to certainly belong to. Filtering then involves the computation of the set of all possible values of the system state that are consistent with the sensor data, i.e., the computation of intersections. Set-membership estimators therefore often encounter difficulties in treating outliers that may even lead to non-intersecting sets.

A simultaneous consideration of stochastic and set-membership uncertainties may allow to flexibly model different sources of estimation uncertainty, to profit from the individual advantages, and to increase the reliability of the estimation results. Research towards combined stochastic and set-membership estimation has been conducted into different directions, which range from generalizations of classical probability theory that clearly distinguish between stochastic and set-membership uncertainties to concepts that subsume both types of uncertainty under an alternative description, such as constraints. Sec. II provides a brief overview of these concepts.

A linear estimator for a simultaneous treatment of both types of uncertainty is derived in this paper by minimizing the mean squared error of the filtered estimate. Set-membership uncertainties essentially affect the mean of the estimate and prevent it from being unbiased. Therefore, the mean squared error yields the sum of the variances and the squared bias. By employing ellipsoidal error bounds, the mean squared error can simply be expressed in terms of

Benjamin Noack, Florian Pfaff, and Uwe D. Hanebeck are with the Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics, Karlsruhe Institute of Technology (KIT), Germany. noack@kit.edu, florian.pfaff@student.kit.edu, uwe.hanebeck@ieee.org

the traces of the covariance matrix and the shape matrix of the ellipsoid. Sec. III briefly reviews the principles of ellipsoidal calculus. Sec. IV is dedicated to the derivation of an optimal Kalman gain in the presence of stochastic and unknown but bounded uncertainties. In Sec. V, an additional weighting parameter is introduced that allows to prioritize the minimization of the stochastic uncertainty or, alternatively, of the set-membership uncertainty. The presented approach is evaluated in Sec. VI and an outlook to prospective work is provided in Sec. VIII.

II. STATE OF THE ART

Several approaches have been proposed in order to make a simultaneous consideration of stochastic and set-membership uncertainties possible. For instance, the standard Kalman filter algorithm can be generalized to a set-valued filter that processes ellipsoids of estimated means. This concept has been introduced in [8], where the mean of the prior Gaussian density is unknown, but bounded by an ellipsoid. The ellipsoid is then predicted and updated by the standard Kalman filter formulas. In [9], this filter has been further developed in order to also incorporate set-membership disturbances affecting control inputs and measurements. Each uncertain quantity is characterized by a set of probability densities, which, more precisely, is a set of translated versions of a Gaussian density. The set-valued Kalman filter directly corresponds to the elementwise application of the Chapman-Kolmogorov integral and of Bayes' theorem to the sets of densities for prediction and filtering, respectively. Also, it is strongly related to Kalman filtering approaches based on coherent lower and upper previsions [10], which also correspond to convex sets of probability densities. As discussed in Sec. V, the concept proposed in this paper includes the set-valued Kalman filter as a special case.

Instead of employing sets of densities, an uncertain quantity can also be characterized by a random set, as it is done in [11]. Two random set observations become deterministic sets when they are conditioned specific outcomes of the underlying random parameter. This implies that the fusion of two random sets yields a random set that encloses all random intersections that are possible outcomes. The challenge is to compute and parameterize the resulting random sets, which may often require rough approximations. Bayesian and set-membership estimators are special cases of this approach.

In [12], a linear estimator is derived that simultaneously minimizes the stochastic and set-membership uncertainty. The filter gain is therefore optimized by solving a linear matrix inequality problem.

Other approaches utilize alternative models of uncertainties, such as constraint-based descriptions [13], [14]. Irrespective of their actual nature, uncertainties can then be subsumed under the considered model. In some cases, such as quantized measurements, it is possible to stay in the Bayesian framework [15]. The state estimate is then conditioned on the entire set of possible measurements, which results in a nonlinear estimation problem even if all models are linear and the stochastic disturbances are Gaussian.

In this work, we derive a linear estimator that minimizes the mean squared error in the presence of both stochastic and set-membership disturbances. The optimization of the filter gain is analogous to the derivation of the standard Kalman gain.

III. ELLIPSOIDAL CALCULUS IN A NUTSHELL

In order to represent and bound set-membership uncertainties, ellipsoidal sets

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

are employed, which are each parameterized by a midpoint $\underline{\hat{c}} \in \mathbb{R}^n$ and a symmetric nonnegative definite shape matrix $\mathbf{X} \in \mathbb{R}^{n \times n}$. Affine transformations can easily be computed by means of the corresponding transformations of the parameters, i.e.,

$$\mathbf{A}\mathcal{E}(\underline{\hat{c}}, \mathbf{X}) + \underline{b} = \mathcal{E}(\mathbf{A}\,\underline{\hat{c}} + \underline{b}, \mathbf{A}\mathbf{X}\mathbf{A}^{\mathrm{T}}) \ . \tag{2}$$

The elementwise sum of two ellipsoids, i.e., the Minkowski sum, does not yield an ellipsoid anymore, but an outer approximation

$$\mathcal{E}(\underline{\hat{c}}_1, \mathbf{X}_1) \oplus \mathcal{E}(\underline{\hat{c}}_2, \mathbf{X}_2) \subseteq \mathcal{E}(\underline{\hat{c}}_1 + \underline{\hat{c}}_2, \mathbf{X}(p)) \ , \tag{3}$$

with

$$\mathbf{X}(p) = (1 + p^{-1})\mathbf{X}_1 + (1 + p)\mathbf{X}_2 \tag{4}$$

can be computed. The inclusion (3) holds for every $p > 0$. The trace of $\mathbf{X}(p)$ corresponds to the sum of squares of the semixaxes and can be minimized by choosing

$$p = \mathrm{trace}(\mathbf{X}_1)^{\frac{1}{2}} \cdot \mathrm{trace}(\mathbf{X}_2)^{-\frac{1}{2}} \ . \tag{5}$$

In general, outer approximations require p to be determined numerically. For any element $\underline{e}$ of (1), the Euclidean distance to the midpoint is related to the trace of $\mathbf{X}$ via

$$\|\underline{\hat{c}} - \underline{e}\|^2 = \mathrm{trace}\left((\underline{\hat{c}} - \underline{e})^2\right) \leq \mathrm{trace}(\mathbf{X}) \ , \tag{6}$$

where we employ the notation $(\underline{x})^2 = (\underline{x}) \cdot (\underline{x})^{\mathrm{T}} \in \mathbb{R}^{n \times n}$ for $\underline{x} \in \mathbb{R}^n$ throughout this paper.

IV. KALMAN GAINS FOR STOCHASTIC AND SET-MEMBERSHIP UNCERTAINTIES

For a simultaneous consideration and treatment of stochastic and set-membership uncertainty, a linear discrete-time estimator is derived in this section that minimizes the mean squared error (MSE) of the estimate. Both types of uncertainty additively affect control inputs and measurements and are assumed to be independent of each other. $\underline{\hat{x}}_k^{\mathrm{e}}$ denotes the estimate of the state conditioned on all measurements up to the current discrete time instant k. The standard Kalman filter provides an estimate that is related to the true state $\underline{\boldsymbol{x}}_k$ by

$$\underline{\hat{x}}_k^{\mathrm{e}} = \underline{\boldsymbol{x}}_k + \boldsymbol{\Delta}_k^{\mathrm{stoc}} \ ,$$

where $\boldsymbol{\Delta}_k^{\mathrm{stoc}} \sim \mathcal{N}(\underline{0}, \mathbf{C}_k^{\mathrm{e}})$ denotes the stochastic error on the state estimate and $\underline{\boldsymbol{x}}_k$ is the true state. In this case, $\mathrm{E}[\underline{\hat{x}}_k^{\mathrm{e}}] = \mathrm{E}[\underline{\boldsymbol{x}}_k]$ holds. But, the additional presence of an additive unknown but bounded disturbance causes the estimator to possibly be biased, i.e., $\mathrm{E}[\hat{\underline{x}}^{\mathrm{e}}] \neq \mathrm{E}[\underline{\boldsymbol{x}}]$ and

$$\hat{\underline{x}}_k^{\mathrm{e}} = \underline{\boldsymbol{x}}_k + \boldsymbol{\Delta}_k^{\mathrm{stoc}} + \Delta_k^{\mathrm{set}} \ .$$

For such a biased estimator, the MSE matrix then yields

$$\mathrm{E}\left[(\hat{\underline{x}}_k^{\mathrm{e}} - \underline{\boldsymbol{x}}_k)^2\right] = \mathrm{E}\left[(\boldsymbol{\Delta}_k^{\mathrm{stoc}})^2\right] + \left(\Delta_k^{\mathrm{set}}\right)^2 \ , \tag{7}$$

where $\mathbf{C}_k^{\mathrm{e}} = \mathrm{E}\left[(\boldsymbol{\Delta}_k^{\mathrm{stoc}})^2\right]$ is the covariance matrix and the unknown but bounded uncertainty Δ_k^{set} is characterized by its membership to the ellipsoid $\mathcal{E}(\underline{0}, \mathbf{X}_k^{\mathrm{e}})$. The MSE is related to the traces of the matrices by

$$\mathrm{E}\left[\|\hat{\underline{x}}_k^{\mathrm{e}} - \underline{\boldsymbol{x}}_k\|^2\right] = \underbrace{\mathrm{E}\left[\|\boldsymbol{\Delta}_k^{\mathrm{stoc}}\|^2\right]}_{=\,\mathrm{trace}(\mathbf{C}^{\mathrm{e}})} + \underbrace{\left\|\Delta_k^{\mathrm{set}}\right\|^2}_{\leq\,\mathrm{trace}(\mathbf{X}^{\mathrm{e}})} \ , \tag{8}$$

which is the quantity to be minimized by the estimator. The inequality $\left\|\Delta_k^{\mathrm{set}}\right\|^2 \leq \mathrm{trace}(\mathbf{X}^{\mathrm{e}})$ justifies the use of ellipsoidal error bounds, since the Euclidean length of the error is directly bounded by the sum of the squared lengths of the semiaxes, i.e., the trace of $\mathbf{X}_k^{\mathrm{e}}$. Hence, the trace of the covariance matrix characterizes the mean squared error of the stochastic term and the trace of the shape matrix bounds the maximum squared error of the set-membership term.

Of course, in several cases, non-stochastic and systematic errors affecting the state estimate can be estimated and canceled out, e.g., by following the direction of [16] and subsequent approaches. However, by contenting ourselves with error bounds, we bypass the need to specify a certain error characteristic, such as constancy. For instance, the unknown but bounded error term can even be a second stochastic disturbance with a compactly supported probability density and therefore, may not behave systematically.

Before we derive the Kalman gains for the considered situation, we extend the notion of unbiasedness to set-membership uncertainties in the following subsection.

A. *On the Unbiased Condition*

Set-membership perturbations may directly affect the mean of a linear estimator, according to

$$\begin{aligned}
&\mathrm{E}\left[\mathbf{K}'\hat{\underline{x}}_1^{\mathrm{e}} + \mathbf{K}\hat{\underline{x}}_2^{\mathrm{e}}\right] = \mathbf{K}'(\mathrm{E}[\underline{\boldsymbol{x}}] + \Delta_1^{\mathrm{set}}) + \mathbf{K}(\mathrm{E}[\underline{\boldsymbol{x}}] + \Delta_2^{\mathrm{set}}) \\
&= (\mathbf{K}' + \mathbf{K})\,\mathrm{E}[\underline{\boldsymbol{x}}] + \mathbf{K}'\Delta_1^{\mathrm{set}} + \mathbf{K}\Delta_2^{\mathrm{set}} \ ,
\end{aligned}$$

where $\mathbf{K}'$ and $\mathbf{K}$ are the gains that are used to fuse the two estimates $\hat{\underline{x}}_1^{\mathrm{e}}$ and $\hat{\underline{x}}_2^{\mathrm{e}}$. For known deviations Δ_1^{set} and Δ_2^{set}, the gains can be determined to eliminate the bias and to minimize (8), i.e., $(\mathbf{K}' + \mathbf{K})\,\mathrm{E}[\underline{\boldsymbol{x}}] = \mathrm{E}[\underline{\boldsymbol{x}}] - \mathbf{K}'\Delta_1^{\mathrm{set}} - \mathbf{K}\Delta_2^{\mathrm{set}}$. Since the deviations are unknown and are bounded by ellipsoids centered at the origin, even $-\Delta_1^{\mathrm{set}}$ and $-\Delta_2^{\mathrm{set}}$ are possible. The mean then yields

$$\mathrm{E}\left[\mathbf{K}'\hat{\underline{x}}_1^{\mathrm{e}} + \mathbf{K}\hat{\underline{x}}_2^{\mathrm{e}}\right] = (\mathbf{K}' + \mathbf{K})\,\mathrm{E}[\underline{\boldsymbol{x}}] - 2\mathbf{K}'\Delta_1^{\mathrm{set}} - 2\mathbf{K}\Delta_2^{\mathrm{set}} \ ,$$

which could significantly increase the actual MSE. Therefore, $\mathbf{K}' = \mathbf{I} - \mathbf{K}$ minimizes the risk of a high error. A second argument is that the set-membership also includes zero-mean stochastic perturbations, for which we expect an unbiased estimation result.

B. *Filtering*

In the filtering step, a prior or predicted state estimate $\hat{\underline{x}}_k^{\mathrm{p}}$ with the error covariance matrix $\mathbf{C}_k^{\mathrm{p}}$ and the ellipsoidal error matrix $\mathbf{X}_k^{\mathrm{p}}$ is fused with observation data $\hat{\underline{z}}_k$ that are linearly related to the system state via

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

where $\underline{\boldsymbol{v}}_k \sim \mathcal{N}(\underline{0}, \mathbf{C}_k^z)$ is an additive zero-mean white noise with covariance matrix $\mathbf{C}_k^z$ and $\underline{e}_k$ denotes an unknown but bounded error enclosed by the ellipsoid $\mathcal{E}(\underline{0}, \mathbf{X}_k^z)$. W.l.o.g., the set-membership error is assumed to be centered at $\underline{0}$. Otherwise, we consider a shifted version $\hat{\underline{z}}_k' = \hat{\underline{z}}_k - \hat{\underline{c}}_k$ of the measurement, when $\underline{e}_k \in \mathcal{E}(\hat{\underline{c}}_k, \mathbf{X}_k^z)$. Furthermore, the stochastic perturbation $\underline{\boldsymbol{v}}_k$ is independent of the outcome of $\underline{e}_k$.

Based upon the measurement data and system dynamics up to a time instance k, we are looking for the Kalman gain $\mathbf{K}_k$ that combines the prior state estimate with the measurement information according to

$$\hat{\underline{x}}_k^{\mathrm{e}} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\hat{\underline{x}}_k^{\mathrm{p}} + \mathbf{K}_k\hat{\underline{z}}_k = \hat{\underline{x}}_k^{\mathrm{p}} + \mathbf{K}_k\left(\hat{\underline{z}}_k - \mathbf{H}_k\,\hat{\underline{x}}_k^{\mathrm{p}}\right) \ ,$$

and concurrently minimizes the trace of (7), i.e. eq. (8). As discussed in Subsec. IV-A, we require the estimator to be unbiased. With $\Delta_k^{\mathrm{p}} = \boldsymbol{\Delta}_k^{\mathrm{stoc}} + \Delta_k^{\mathrm{set}}$ denoting the errors within the prior estimate $\hat{\underline{x}}_k^{\mathrm{p}}$, the MSE matrix yields

$$\begin{aligned}
&\mathrm{E}\left[(\hat{\underline{x}}_k^{\mathrm{e}} - \underline{\boldsymbol{x}}_k)^2\right] = \mathrm{E}\left[\left(\hat{\underline{x}}_k^{\mathrm{p}} + \mathbf{K}_k(\hat{\underline{z}}_k - \mathbf{H}_k\hat{\underline{x}}_k^{\mathrm{p}}) - \underline{\boldsymbol{x}}_k\right)^2\right] \\
&= \mathrm{E}\left[\left(\Delta_k^{\mathrm{p}} + \mathbf{K}_k(\mathbf{H}_k\underline{\boldsymbol{x}}_k + \underline{\boldsymbol{v}}_k + \underline{e}_k - \mathbf{H}_k\hat{\underline{x}}_k^{\mathrm{p}})\right)^2\right] \\
&= \mathrm{E}\left[\left((\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\Delta_k^{\mathrm{p}} + \mathbf{K}_k(\underline{\boldsymbol{v}}_k + \underline{e}_k)\right)^2\right] \\
&= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\,\mathrm{E}[(\boldsymbol{\Delta}_k^{\mathrm{stoc}})^2](\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^{\mathrm{T}} + \mathbf{K}_k\,\mathrm{E}[(\underline{\boldsymbol{v}}_k)^2]\mathbf{K}_k^{\mathrm{T}} \\
&\quad + \mathrm{E}\left[\left((\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\Delta_k^{\mathrm{set}} + \mathbf{K}_k\underline{e}_k\right)^2\right] \\
&= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{C}_k^{\mathrm{p}}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^{\mathrm{T}} + \mathbf{K}_k\mathbf{C}_k^z\mathbf{K}_k^{\mathrm{T}} \\
&\quad + \left((\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\Delta_k^{\mathrm{set}} + \mathbf{K}_k\underline{e}_k\right)^2 \ .
\end{aligned} \tag{9}$$

Due to the set-membership of Δ_k^{set} and $\underline{e}_k$, the latter sum can be considered as a Minkowki sum, i.e.,

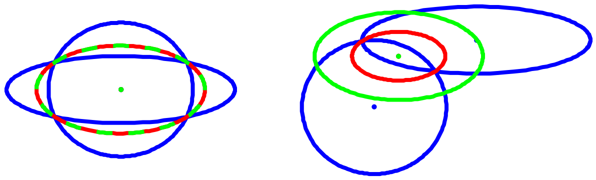
$$\begin{aligned}
&(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\Delta_k^{\mathrm{set}} + \mathbf{K}_k\underline{e}_k \\
&\qquad \in (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathcal{E}(\underline{0}, \mathbf{X}_k^{\mathrm{p}}) \oplus \mathbf{K}_k\mathcal{E}(\underline{0}, \mathbf{X}_k^z) \\
&\qquad \overset{(2)}{=} \mathcal{E}(\underline{0}, (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{X}_k^{\mathrm{p}}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^{\mathrm{T}}) \\
&\qquad\quad \oplus \mathcal{E}(\underline{0}, \mathbf{K}_k\mathbf{X}_k^z\mathbf{K}_k^{\mathrm{T}}) \\
&\qquad \overset{(3)}{\subset} \mathcal{E}(\underline{0}, \mathbf{X}_k^{\mathrm{e}}(p)) \ ,
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{X}_k^{\mathrm{e}}(p) = {} & (1 + p^{-1})(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{X}_k^{\mathrm{p}}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^{\mathrm{T}} \\
& + (1 + p)\mathbf{K}_k\mathbf{X}_k^z\mathbf{K}_k^{\mathrm{T}}
\end{aligned} \tag{10}$$

denotes the shape matrix of the bounding ellipsoid, according to (4). For this matrix, we have

$$\mathrm{trace}\left(\left((\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\Delta_k^{\mathrm{set}} + \mathbf{K}_k\underline{e}_k\right)^2\right) \leq \mathrm{trace}\left(\mathbf{X}_k^{\mathrm{e}}(p)\right)$$



(a) Ellipsoids with the same midpoint. The optimal and the centered approximation of the intersection are identical.

(b) Ellipsoids with different midpoints. An optimal approximation of the intersection is much smaller than the centered approximation.

Fig. 1: Centered (green) and uncentered (red) ellipsoidal approximations of intersections. Fusing the blue ellipsoids with the gain (13) yields the green ellipsoids.

for all $p > 0$, due to the inequality (6). In view of (8) and (9), the actual MSE can be bounded from above by

$$\begin{aligned}
&\mathrm{E}\left[\|\hat{\underline{x}}_k^{\mathrm{e}} - \underline{\boldsymbol{x}}_k\|^2\right] = \mathrm{trace}\big(\,\mathrm{E}\left[(\hat{\underline{x}}_k^{\mathrm{e}} - \underline{\boldsymbol{x}}_k)^2\right]\big) \\
&= \mathrm{trace}\big(\mathbf{C}_k^{\mathrm{e}}\big) + \mathrm{trace}\Big(\big((\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\Delta_k^{\mathrm{set}} + \mathbf{K}_k\underline{e}_k\big)^2\Big) \\
&\leq \mathrm{trace}\big(\mathbf{C}_k^{\mathrm{e}}\big) + \mathrm{trace}\big(\mathbf{X}_k^{\mathrm{e}}(p)\big) \\
&= \mathrm{trace}\big((\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)\mathbf{C}_k^{\mathrm{p}}(\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)^{\mathrm{T}}\big) + \mathrm{trace}\big(\mathbf{K}_k\mathbf{C}_k^{z}\mathbf{K}_k^{\mathrm{T}}\big) \\
&\quad + (1+p^{-1})\,\mathrm{trace}\big((\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{X}_k^{\mathrm{p}}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^{\mathrm{T}}\big) \\
&\quad + (1+p)\,\mathrm{trace}\big(\mathbf{K}_k\mathbf{X}_k^{z}\mathbf{K}_k^{\mathrm{T}}\big) \ .
\end{aligned} \tag{11}$$

Analogously to the derivation of the standard Kalman gain and by utilizing the derivative rules for the trace, i.e.,

$$\begin{aligned}
&\frac{\partial}{\partial \mathbf{K}_k}\mathrm{trace}\big(\mathbf{A}\mathbf{K}_k\big) = \frac{\partial}{\partial \mathbf{K}_k}\mathrm{trace}\big(\mathbf{K}_k\mathbf{A}\big) = \mathbf{A}^{\mathrm{T}} \text{ and} \\
&\frac{\partial}{\partial \mathbf{K}_k}\mathrm{trace}\big(\mathbf{K}_k\mathbf{A}\mathbf{K}_k\big) = \mathbf{K}_k(\mathbf{A} + \mathbf{A}^{\mathrm{T}}) \ ,
\end{aligned}$$

the optimal gain yields

$$\begin{aligned}
&\mathbf{K}_k(p) = \Big((1+p^{-1})\mathbf{X}_k^{\mathrm{p}}\mathbf{H}_k^{\mathrm{T}} + \mathbf{C}_k^{\mathrm{p}}\mathbf{H}_k^{\mathrm{T}}\Big)\cdot \\
&\Big((1+p^{-1})\mathbf{H}_k\mathbf{X}_k^{\mathrm{p}}\mathbf{H}_k^{\mathrm{T}} + (1+p)\mathbf{X}_k^{z} + \mathbf{H}_k\mathbf{C}_k^{\mathrm{p}}\mathbf{H}^{\mathrm{T}} + \mathbf{C}_k^{z}\Big)^{-1}
\end{aligned} \tag{12}$$

for an arbitrary but fixed $p > 0$. Unfortunately, a convex optimization is needed to find that value p which minimizes (11). The need for such an optimization is a usual issue of ellipsoidal approximations of, inter alia, Minkowski sums or intersections. With $\mathbf{K}_k = \mathbf{K}_k(p^{\mathrm{opt}})$, the updated covariance matrix is computed by

$$\mathbf{C}_k^{\mathrm{e}} = (\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)\mathbf{C}_k^{\mathrm{p}}(\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)^{\mathrm{T}} + \mathbf{K}_k\mathbf{C}_k^{z}\mathbf{K}_k^{\mathrm{T}}$$

and the shape matrix $\mathbf{X}_k^{\mathrm{e}}(p^{\mathrm{opt}})$ of the bounding ellipsoid by means of (10). The following subsection considers two important special cases of this estimator.

C. Special Cases

The proposed combined estimator reduces to well-known estimation principles in special cases, namely, in the situation of vanishing set-membership or, respectively, vanishing stochastic uncertainty. The

Kalman Filter appears in its standard formulation, if $\mathbf{X}_k^{\mathrm{p}} = \mathbf{X}_k^{z} = \mathbf{0}$. The gain simply becomes

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

This result is expected, since we have strictly followed and generalized the derivation of the standard Kalman gain. More surprising is the special case of a

Centered Intersection, if $\mathbf{C}_k^{\mathrm{p}} = \mathbf{C}_k^{z} = \mathbf{0}$. The gain (12) then reduces to

$$\begin{aligned}
\mathbf{K}_k = (1+p^{-1})&\mathbf{X}_k^{\mathrm{p}}\mathbf{H}_k^{\mathrm{T}}\cdot \\
&\Big((1+p^{-1})\mathbf{H}_k\mathbf{X}_k^{\mathrm{p}}\mathbf{H}_k^{\mathrm{T}} + (1+p)\mathbf{X}_k^{z}\Big)^{-1} \ .
\end{aligned} \tag{13}$$

With this gain, the formula (10) for the shape matrix of the bounding ellipsoid can be simplified to

$$\mathbf{X}_k^{\mathrm{e}}(\omega) = \big((1-\omega)(\mathbf{X}_k^{\mathrm{p}})^{-1} + \omega\mathbf{H}_k^{\mathrm{T}}(\mathbf{X}_k^{z})^{-1}\mathbf{H}_k\big)^{-1}$$

with $\omega = \frac{1}{1+p} \in [0,1]$. By letting $\mathbf{H}_k = \mathbf{I}$ in order to simplify matters, this matrix characterizes the centered intersection

$$\mathcal{E}(\underline{0}, \mathbf{X}_k^{\mathrm{p}}) \cap \mathcal{E}(\underline{0}, \mathbf{X}_k^{z}) \subset \mathcal{E}(\underline{0}, \mathbf{X}_k^{\mathrm{e}}(\omega)) \ ,$$

which is analogous to the results of covariance intersection [17]. More precisely, the fused bounding ellipsoid bounds the maximally possible intersection, as illustrated in Fig. 1 .

D. Prediction

Approaches such as [9] and [11] come to the same conclusion with regard to the prediction. We consider linear system dynamics

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

where $\underline{\boldsymbol{w}}_k \sim \mathcal{N}(\underline{0}, \mathbf{C}_k^{u})$ and $\underline{d}_k \in \mathcal{E}(\underline{0}, \mathbf{X}_k^{u})$. Predicting the estimate by means of this model gives

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

For this predicted estimate, the error matrices can be calculated by means of (7). The predicted covariance matrix then yields

$$\begin{aligned}
\mathbf{C}_{k+1}^{\mathrm{p}} &= \mathrm{E}\left[(\mathbf{A}_k\boldsymbol{\Delta}_k^{\mathrm{stoc}} + \mathbf{B}_k\underline{\boldsymbol{w}}_k)^2\right] \\
&= \mathbf{A}_k\mathbf{C}_k^{\mathrm{e}}\mathbf{A}_k^{\mathrm{T}} + \mathbf{B}_k\mathbf{C}_k^{u}\mathbf{B}_k^{\mathrm{T}} \ .
\end{aligned}$$

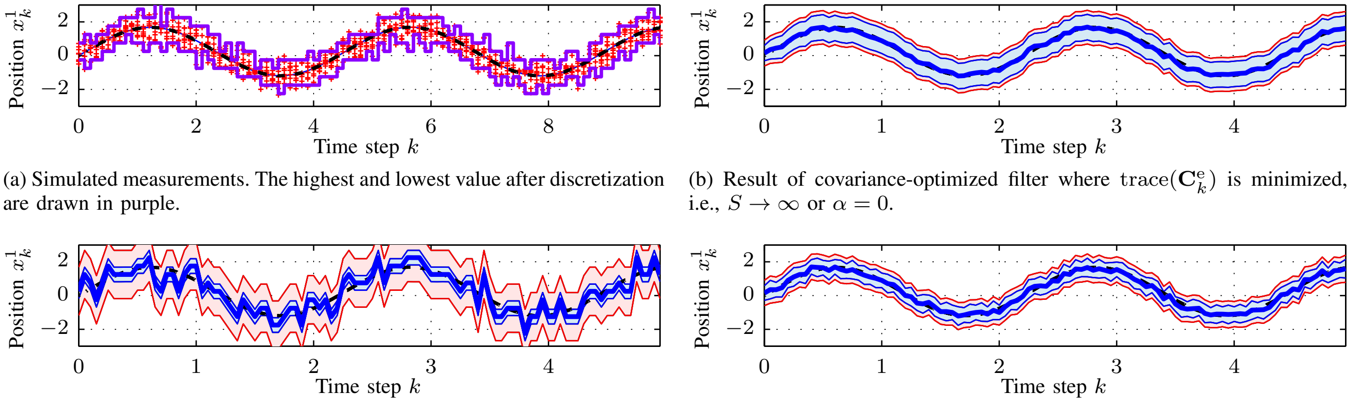
The predicted set-membership error is bounded by the Minkowski sum

$$\begin{aligned}
\mathbf{A}_k\Delta_k^{\mathrm{set}} + \mathbf{B}_k\underline{d}_k &\in \mathbf{A}_k\mathcal{E}(\underline{0}, \mathbf{X}_k^{\mathrm{e}}) \oplus \mathbf{B}_k\mathcal{E}(\underline{0}, \mathbf{X}_k^{u}) \\
&\subset \mathcal{E}\big(\underline{0}, \mathbf{X}_{k+1}^{\mathrm{p}}(p)\big)
\end{aligned}$$

with

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

and $p > 0$. In this case, p can analytically be determined to minimize the trace of the error matrix $\mathbf{X}_{k+1}^{\mathrm{p}}(p)$ by means of (5).



(a) Simulated measurements. The highest and lowest value after discretization are drawn in purple.

(b) Result of covariance-optimized filter where $\mathrm{trace}(\mathbf{C}_k^{\mathrm{e}})$ is minimized, i.e., $S \to \infty$ or $\alpha = 0$.

(c) Result of set-optimized filter where $\mathrm{trace}(\mathbf{X}_k^{\mathrm{e}})$ is minimized, i.e., $S = 0$ or $\alpha = 1$.

(d) Result of combined filter from Sec. IV, i.e., $S = 1$ or $\alpha = 0.5$.

Fig. 2: Simulated measurements and filtering results.

V. GENERALIZATION

This section provides a generalization of the proposed estimation concept in order to enable the user to decide whether the stochastic or the set-membership uncertainty shall primarily be minimized. The idea behind it rests upon the fact that for the standard Kalman gain, the same result is attained when instead of

$$\mathrm{trace}\big(\mathbf{C}_k^{\mathrm{e}}\big)=\mathrm{trace}\big((\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)\mathbf{C}_k^{\mathrm{p}}(\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)^{\mathrm{T}}+\mathbf{K}_k\mathbf{C}_k^{z}\mathbf{K}_k^{\mathrm{T}}\big)$$

a scaled version $\mathrm{trace}\left(S \cdot \mathbf{C}_k^{\mathrm{e}}\right)$ with $S \geq 0$ is considered, which corresponds to a scaled covariance ellipsoid around the current state estimate $\hat{\underline{x}}_k^{\mathrm{e}}$. By striving for

$$\mathbf{K}_k(p) = \arg\min \left\{ \mathrm{trace}\big(S \cdot \mathbf{C}_k^{\mathrm{e}}\big) + \mathrm{trace}\big(\mathbf{X}_k^{\mathrm{e}}(p)\big) \right\} \quad (14)$$

instead of minimizing (11) in Sec. IV-B, the gain

$$\mathbf{K}_k(p) = \Big(\big(1+\frac{1}{p}\big)\mathbf{X}_k^{\mathrm{p}}\mathbf{H}_k^{\mathrm{T}} + S\mathbf{C}_k^{\mathrm{p}}\mathbf{H}_k^{\mathrm{T}}\Big)\cdot$$
$$\Big(\big(1+\frac{1}{p}\big)\mathbf{H}_k\mathbf{X}_k^{\mathrm{p}}\mathbf{H}_k^{\mathrm{T}} + (1+p)\mathbf{X}_k^{z} + S\mathbf{H}_k\mathbf{C}_k^{\mathrm{p}}\mathbf{H}^{\mathrm{T}} + S\mathbf{C}_k^{z}\Big)^{-1} \quad (15)$$

is obtained. Alternatively to (14), the problem can also be formulated as a convex combination

$$\mathbf{K}_k(p) = \arg\min \left\{ (1-\alpha)\,\mathrm{trace}\big(\mathbf{C}_k^{\mathrm{e}}\big) + \alpha\,\mathrm{trace}\big(\mathbf{X}_k^{\mathrm{e}}(p)\big) \right\}$$

with $\alpha \in [0, 1]$.

For $S = 0$ or, respectively, $\alpha = 1$, we obtain the formulas of the centralized intersection. For $S \to \infty$ or, respectively, $\alpha = 0$, the gain (15) becomes the standard Kalman gain. This case corresponds to the set-valued Kalman filter derived in [8] and [9]. It only considers the covariance matrices when computing the gains.

VI. SIMULATIONS

For evaluation purposes, we simulate a dynamic system that suffers from both stochastic and set-membership errors. The system model has been generated from time-discretization and linearization of a differential equation. It features a three-dimensional state vector and the following transition matrix depending on the time-step $h = \Delta t$:

$$\mathbf{A} = \begin{bmatrix} 1 & h & h \\ -h & 1 & h \\ -0.51h & -0.51h & 1 \end{bmatrix} .$$

The differential equation is approximated better the smaller h is chosen. The ground truth is generated by setting h as low as $0.0001s$, for the system model the value is chosen identical to the time-interval between two subsequent batches of measurements: every $h = 0.1s$ ten fully-dimensional samples $\hat{\underline{z}}_k = \underline{x}_k + \underline{\boldsymbol{v}}$ are generated, altered by the random noise $\underline{\boldsymbol{v}} \sim \mathcal{N}(\underline{0}, \mathbf{C}^z)$. A set-membership error is added by discretization according to $f(z) = 2\left\lfloor \frac{z}{2} \right\rfloor + 0.25$, dividing the measurement space into $[0.5)^3$-cubes and mapping each value to its respective centroid. The trace-minimal ellipsoids containing such cubes are centered in the centroid of the cube and their shape is described by $\mathbf{X}^z = \frac{3}{16}\mathbf{I}$.

For the stochastic and set-membership system noise matrices $\mathbf{C}^u$ and $\mathbf{X}^u$, reasonable matrices are chosen that provide a consistent approximation for the true system noise for each time-step. The covariance matrices are

$$\mathbf{C}^u = \mathrm{diag}([0.2, 0.15, 0.1]), \quad \mathbf{C}^z = \mathrm{diag}([0.25, 0.5, 0.75])$$

and the shape matrices of the bounding ellipsoids are

$$\mathbf{X}^u = 0.1\mathbf{I} \;, \qquad \mathbf{X}^z = \frac{3}{16}\mathbf{I} \;.$$

The system has been simulated for 10 seconds with the initial estimate $\begin{bmatrix} 0 & 1 & 1 \end{bmatrix}^{\mathrm{T}}$ and the estimation results of the first component are shown in Fig. 2. The dashed black line represents the ground truth in all plots. Measurements altered by stochastic noise are depicted in Fig. 2a as tiny red crosses and - to give an impression of the applied discretization - the highest and the lowest discretized values of each time-step are drawn in purple. While the state estimate is drawn in dark blue, the whole cyan area represents the set-membership uncertainty at each point of time. The red area is the 2-sigma boundary added upon the outermost elements of the set.

Results are as to be expected: usage of the covariance-minimal estimator results in relatively large set-membership uncertainty, while the set-optimal filter yields high stochastic uncertainty and provides a noisy estimated signal. Whereas not outperforming them in the individual criteria, the combined filter surpasses them at having the smallest interval of points that lie within any possible distributions 2-sigma boundary. The cyan and red area together contain the actual state with $> 95\%$ confidence. In this regard, the plot for the combined filter shows the best result.

VII. COMPARISONS

As stated in Sec. II, the presented Kalman filter is not the only concept enabling a simultaneous consideration of stochastic and set-membership errors.

In [18], set measurements can be incorporated by computing the likelihood for the entire set, which implies that a uniform distribution of the set-membership error is assumed. Without approximations, this procedure turns a linear estimation problem into a nonlinear one. Set-membership uncertainties affecting the state transition model are not taken into account.

In [11], stochastic and set-membership uncertainties are characterized by means of random sets. A comparison of the proposed approach for $S \to \infty$ or $\alpha = 0$ is carried out in [19] and unveils that a different interpretation of set-membership uncertainty is also reasonable.

[12] proposes a similar approach that minimizes a cost function under linear matrix inequality (LMI) conditions. In the special case that the set-membership is weighted with 0, i.e., $S \to \infty$ or $\alpha = 0$, the approach of [12] and the approach in this paper coincide with [8], [9]. Prospective work will compare both approaches in terms of performance and estimation quality.

VIII. CONCLUSION AND OUTLOOK

The discrete-time Kalman filter has been generalized in order to allow for a simultaneous treatment of stochastic and set-membership uncertainties. In this work, an optimal Kalman gain has been derived that additionally takes possible set-membership uncertainties into account and minimizes the maximum possible MSE. Set-membership uncertainties have been circumscribed by enclosing ellipsoids. With regard to the MSE, we consider ellipsoidal sets to be most appropriate and the corresponding shape matrices imply a strong analogy to covariance matrices. As an important feature, the proposed concept not only includes the standard (purely stochastic) Kalman filter as a special case, but also the outer ellipsoidal approximation of intersecting sets in the absence of stochastic uncertainty. We have further generalized the estimator, such that the user can prioritize the minimization of the stochastic uncertainty or the minimization of the set-membership uncertainty.

Prospective research will focus on further extensions of the proposed concept. By following the direction of [9], it can be applied in nonlinear estimation problems where the ellipsoidal bounds can be utilized to account for linearization errors. It also appears promising to use two estimators in parallel, for instance, the covariance-optimal ($S \to \infty$) and the set-optimal ($S = 0$) estimator. The prioritization of one type of uncertainty can then be adapted afterward when the corresponding estimation results are fused. Also, the requirement of unbiasedness in IV-A can possibly be relaxed, which might lead to further insights with respect to the set-membership uncertainties.

ACKNOWLEDGMENT

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*, pp. 35–45, 1960.

[2] S. J. Julier and J. K. Uhlmann, "A New Extension of the Kalman Filter to Nonlinear Systems," in *Proceedings the 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls (AeroSense 1997)*, Orlando, Florida, USA, 1997.

[3] L. Jaulin, M. Kieffer, O. Didrit, and É. Walter, *Applied Interval Analysis: With Examples in Parameter and State Estimation, Robust Control and Robotics*. London: Springer, 2001.

[4] D. P. Bertsekas and I. B. Rhodes, "Recursive State Estimation for a Set-membership Description of Uncertainty," *IEEE Transactions on Automatic Control*, vol. 16, pp. 117–128, 1971.

[5] F. C. Schweppe, *Uncertain Dynamic Systems*. Prentice-Hall, 1973.

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

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

[8] 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.

[9] 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.

[10] A. Benavoli, M. Zaffalon, and E. Miranda, "Robust Filtering Through Coherent Lower Previsions," *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1567–1581, Jul. 2011.

[11] 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.

[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)*, Cancún, Mexico, Dec. 2008.

[13] I. R. Petersen, M. R. James, and P. Dupuis, "Minimax Optimal Control of Stochastic Uncertain Systems with Relative Entropy Constraints," *IEEE Transactions on Automatic Control*, vol. 45, pp. 398–412, 2000.

[14] A. Kallapur, I. R. Petersen, and S. Anavatti, "A Discrete-time Robust Extended Kalman Filter," in *Proceedings of the 2009 American Control Conference (ACC 2009)*, St. Louis, Missouri, USA, Jun. 2009.

[15] R. P. S. Mahler, "General Bayes Filtering of Quantized Measurements," in *Proceedings of the 14th International Conference on Information Fusion (Fusion 2011)*, Chicago, Illinois, USA, Jul. 2011.

[16] B. Friedland, "Treatment of Bias in Recursive Filtering," *IEEE Transactions on Automatic Control*, vol. 14, no. 4, pp. 359–367, Aug. 1969.

[17] S. J. Julier and J. K. Uhlmann, *Handbook of Multisensor Data Fusion: Theory and Practice*, 2nd ed. CRC Press, 2009, ch. General Decentralized Data Fusion with Covariance Intersection, pp. 319–343.

[18] Z. Duan, X.-R. Li, and V. P. Jilkov, "State Estimation with Point and Set Measurements," in *Proceedings of the 13th International Conference on Information Fusion (Fusion 2010)*, Edinburgh, United Kingdom, Jul. 2010.

[19] V. Klumpp, B. Noack, M. Baum, and U. D. Hanebeck, "Combined Set-Theoretic and Stochastic Estimation: A Comparison of the SSI and the CS Filter," in *Proceedings of the 13th International Conference on Information Fusion (Fusion 2010)*, Edinburgh, United Kingdom, Jul. 2010.