Online Performance Assessment of Multi-Sensor Kalman Filters Based on Subjective Logic

Thomas Griebel, Jonas Heinzler, Michael Buchholz, and Klaus Dietmayer

This paper has been accepted for presentation and publication at the 2023 26th International Conference on Information Fusion (FUSION), June 27 - 30, 2023, Charleston, SC, USA. This is the accepted version of the paper, which has not been fully edited and the layout may differ from the original publication.

Citation information of the original publication:

T. Griebel, J. Heinzler, M. Buchholz and K. Dietmayer, "Online Performance Assessment of Multi-Sensor Kalman Filters Based on Subjective Logic," 2023 26th International Conference on Information Fusion (FUSION), Charleston, SC, USA, 2023, pp. 1-8, doi: 10.23919/FUSION52260.2023.10224188.

Online Performance Assessment of Multi-Sensor Kalman Filters Based on Subjective Logic

Thomas Griebel, Jonas Heinzler, Michael Buchholz, and Klaus Dietmayer
Institute of Measurement, Control and Microtechnology, *Ulm University*, Germany
{firstname}.{lastname}@uni-ulm.de

***Abstract*—Operation monitoring for automation systems requires self-assessment of all data processing modules. In this work, we extend our new self-assessment method for linear Kalman filters based on subjective logic to nonlinear Kalman filtering. Furthermore, we propose novel approaches within this subjective logic-based framework to assess the overall filter performance in multi-sensor systems online, i.e., in real-time without ground truth data. The results of the proposed self-assessment method for nonlinear Kalman filtering are demonstrated through simulation studies, showing advantages compared to classical consistency measures, like the normalized innovation squared. In addition, the results of the proposed online overall filtering assessment for multi-sensor systems can even compete with consistency measures based on ground truth data, which cannot be applied in online applications.**

I. Introduction

Dynamic state estimation is a widely recognized and frequently employed technique for filtering and tracking, and the Kalman filter is one of its most well-known methods. Historically, simple measures such as the *normalized innovation squared* (NIS) and the *normalized estimation error squared* (NEES) have been employed in Kalman filtering to check for consistency [1], [2]. Like all other available methods in literature so far, both only check single criteria or perform specific statistical tests, as the NIS applied to nonlinear Kalman filtering in [3] or additional checks to test some particular aspects as, e.g., for nonlinearity in [4]. Hence, these methods only provide limited information about the reliability of the estimation results and the algorithms themselves.

However, in the automotive industry, e.g., recent developments put a growing focus on ensuring reliability beyond functional safety, particularly in light of the ISO 21448 *safety of the intended functionality* (SOTIF) [5] standard, which requires self-assessment (SA) of the functional modules. Hence, what is missing until now is a unified, holistic concept and framework for SA in filtering and tracking algorithms.

In this paper, we present a novel overall online assessment concept and framework for linear and nonlinear Kalman filtering that is based on subjective logic (SL) [6]; see Fig. 1. Our contribution is twofold: First, we extend our SA method for linear Kalman filtering from [7] to nonlinear Kalman filtering.

Parts of this research have been conducted as part of the EVENTS project, which is funded by the European Union, under grant agreement No 101069614. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Commission. Neither the European Union nor the granting authority can be held responsible for them.

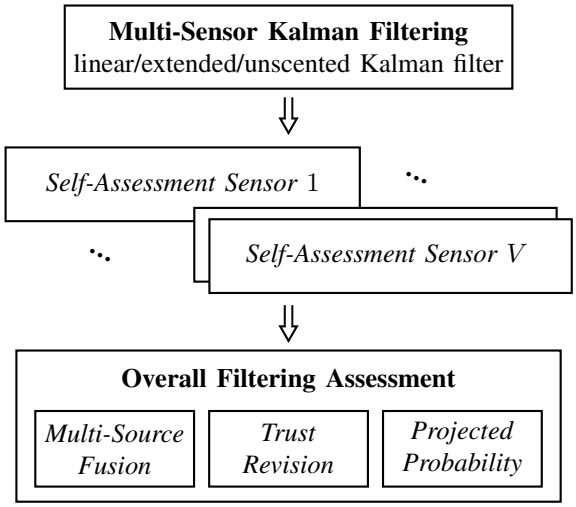

Fig. 1. Our proposed self-assessment framework for linear and nonlinear multi-sensor Kalman filtering. The self-assessment framework proposes three different overall filtering assessment approaches based on the multi-source fusion, trust revision, and projected probability concepts from subjective logic.

Second, based on this extension, we propose a novel overall SA framework for multi-sensor systems based on it. This framework provides a more flexible and adaptable approach to SA, allowing for real-time monitoring of the system's performance and identification of potential issues. Our simulation experiments demonstrate that our proposed overall online assessment can even achieve comparable outcomes to the NEES, however, without ground truth (GT) data. The presented approach addresses the need for reliability beyond functional safety and offers a powerful and comprehensive framework for SA in multi-sensor Kalman filtering.

Our main contributions are:

- an extension of the SA approach based on SL towards nonlinear Kalman filtering in Section IV,
- a novel overall online assessment concept and framework based on SL for multi-sensor systems in Section V,
- an extensive evaluation of our SA methods in scenarios with challenging disturbance modes of the filter and comparison to classical consistency measures in Section VI.

II. Related Work

In [2], Bar-Shalom et al. present the classical consistency measures in Kalman filtering. These are the NIS and the

NEES, where the NEES requires GT data, in contrast to the NIS. Based on these, further consistency assessments, like their transfer to nonlinear filters, have been introduced in recent years. For example, Scalzo et al. [3] select an approach that adaptively switches between the *extended Kalman filter* (EKF), the *unscented Kalman filter* (UKF), and the *particle filter* (PF) after evaluating the respective NIS values. Niu et al. in [4] use two measures to quantitatively assess the nonlinearity of nonlinear filtering problems using the EKF, UKF, and PF. Specifically, curvature measures, as well as the NIS, are used to measure nonlinearity. In [8], Ivanov et al. introduce the *normalized deviation squared* (NDS) and the p-equivalence tests, which also can be used when the system is nonlinear, and a suboptimal filter is applied. In addition, Duník et al. [9] apply a non-Gaussian measure that relies on the conditional third moment of the state to identify a potential decrease in estimation quality. In [10], Mahler presents a broader approach that relies on the NIS and terms it a divergence detector. This approach incorporates the idea of the *generalized NIS* (GNIS), which is further developed into the *multi-target generalized NIS* (MGNIS).

Regarding the application of SL for SA in filtering, our earlier work [7] introduced a novel SA technique for linear Kalman filtering that includes explicit modeling of statistical uncertainty. Based on this, we proposed a SA framework for single-object tracking in clutter using linear Kalman filtering and nearest neighbor association in [11]. This approach considers various aspects of reliability assessment, such as the correctness of the assumed measurement noise, the detection probability, and the clutter rate.

Building on our previous research, this paper proposes the first extensive SA framework in the literature for multi-sensor systems that employ both linear and nonlinear Kalman filtering. This framework allows for real-time evaluation of the performance of individual sensors, as well as the overall filtering performance.

III. Fundamentals

This section summarizes the foundations of nonlinear Kalman filtering based on [2] and the mathematical concepts of SL based on [6].

A. Nonlinear Kalman Filtering

The goal of nonlinear filtering approaches is to estimate the state of a discrete-time nonlinear dynamic-system

$$\boldsymbol{x}_{k+1} = \boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{v}_k), \tag{1}$$

$$\boldsymbol{z}_k = \boldsymbol{h}(\boldsymbol{x}_k, \boldsymbol{w}_k) \tag{2}$$

in each time step $k \in \mathbb{N}$, where $\boldsymbol{x}_k \in \mathbb{R}^n$ is the unobserved state vector of the system with state dimension $n \in \mathbb{N}$ and $\boldsymbol{z}_k \in \mathbb{R}^m$ is the observed measurement vector with measurements dimension $m \in \mathbb{N}$. The vector-valued motion and measurement function $\boldsymbol{f}$ and $\boldsymbol{h}$ are generally time-varying and nonlinear. Moreover, $\boldsymbol{v}_k \in \mathbb{R}^n$ and $\boldsymbol{w}_k \in \mathbb{R}^m$ are the process and measurement noise.

For linear systems with Gaussian distributed noise, the Kalman filter [12] is a Bayes-optimal estimator [13] and allows a closed-form implementation of the Bayes filter for recursive state estimation. For nonlinear systems with Gaussian distributed noise, the EKF [14], [15] is a nonlinear implementation of the Kalman filter, which uses linearization in terms of Taylor approximation typically of first order. Moreover, the UKF [16], [17] is another nonlinear implementation of the Kalman filter, which uses stochastic linearization in terms of the so-called unscented transform [17].

For online consistency checking of linear and nonlinear Kalman filters, the time-average NIS [2], [3]

$$\bar{\varepsilon}_{\boldsymbol{\gamma}} = \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{\gamma}_k^T \boldsymbol{S}_k^{-1} \boldsymbol{\gamma}_k =: \frac{1}{K}\sum_{k=1}^{K} \varepsilon_{\boldsymbol{\gamma}_k} \tag{3}$$

can be used as a time-average value of the classical NIS $\varepsilon_{\boldsymbol{\gamma}_k}$ over a window size $K \in \mathbb{N}$. The NIS is calculated using the innovation $\boldsymbol{\gamma}_k := \boldsymbol{z}_k - \hat{\boldsymbol{z}}_{k|k-1}$ based on the measurement prediction with mean $\hat{\boldsymbol{z}}_{k|k-1} \in \mathbb{R}^m$ and its innovation covariance $\boldsymbol{S}_k \in \mathbb{R}^{m\times m}$. If all Kalman filter assumptions are met in the linear case, $K\bar{\varepsilon}_{\boldsymbol{\gamma}}$ is χ^2 distributed with Km degrees of freedom. For nonlinear Kalman filtering, this can only be approximately true.

For offline consistency monitoring with GT data, the time-average NEES [2]

$$\bar{\varepsilon}_{\boldsymbol{x}} = \frac{1}{K}\sum_{k=1}^{K} \tilde{\boldsymbol{x}}_k^T \boldsymbol{P}_{k|k}^{-1} \tilde{\boldsymbol{x}}_k =: \frac{1}{K}\sum_{k=1}^{K} \varepsilon_{\boldsymbol{x}_k} \tag{4}$$

can be used similarly to the time-average NIS. The NEES $\varepsilon_{\boldsymbol{x}_k}$ is calculated using the state residual $\tilde{\boldsymbol{x}}_k := \boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k}$ based on the Kalman filter estimates, i.e., the state vector $\hat{\boldsymbol{x}}_{k|k} \in \mathbb{R}^n$ and its covariance $\boldsymbol{P}_{k|k} \in \mathbb{R}^{n\times n}$. If all Kalman filter assumptions are met, $K\bar{\varepsilon}_{\boldsymbol{x}}$ is χ^2 distributed with Kn degrees of freedom. However, this only applies in the linear case. For the nonlinear filters, this applies only approximately. When checking consistency using the NIS and NEES, the values are expected to be within a certain confidence interval of the χ^2 distribution.

In Kalman filtering for multi-sensor systems, there are two main approaches for combining measurements from multiple sensors: the sequential update and the group update [18]. The sequential update processes the sensor measurements one by one for real-time processing. In contrast, the group update combines all sensor measurements into a single batch for increased accuracy but may be computationally more expensive. The NIS can be used in both sequential and group update to evaluate consistency. Applying the NIS for the sequential update involves calculating it for each sensor measurement. Applying the NIS for the group update requires calculating the entire batch of sensor measurements to assess the consistency. In this paper, we assume synchronous sensor measurements and use the multi-sensor Kalman filter approach with the sequential update as presented in [18] because only this setup allows for separate consistency statements for each sensor.

B. Subjective Logic

This section outlines the main principles of SL theory as needed for this work, mainly based on the work by Jøsang et al. [6]. SL is s a mathematical framework that explicitly represents statistical uncertainty [6], comparable to the Dempster-Shafer theory [19], [20]. The key aspect of SL is the representation of opinions. Typically, a multinomial opinion embodies information about a discrete random variable X for each event x in the sample space $\mathbb{X}$, in terms of belief, uncertainty, and base rate.

Definition 1 (Multinomial Opinion)**.** *Consider a random variable X in the finite domain $\mathbb{X}$ with cardinality $W = |\mathbb{X}| \geq 2$. A multinomial opinion can be defined as an ordered triple $\omega_X = (\boldsymbol{b}_X, u_X, \boldsymbol{a}_X)$ with*

$$\boldsymbol{a}_X(x) : \mathbb{X} \mapsto [0,1], \qquad 1 = \sum_{x \in \mathbb{X}} \boldsymbol{a}_X(x)\,, \tag{5a}$$

$$\boldsymbol{b}_X(x) : \mathbb{X} \mapsto [0,1], \qquad 1 = u_X + \sum_{x \in \mathbb{X}} \boldsymbol{b}_X(x)\,. \tag{5b}$$

The belief mass distribution $\boldsymbol{b}_X$ over $\mathbb{X}$ reflects the belief in each event, the uncertainty mass $u_X \in [0,1]$ signifies the lack of evidence, and the base rate distribution $\boldsymbol{a}_X$ over $\mathbb{X}$ represents the prior probability for each event.

To project a multinomial opinion into a classical probability distribution, we use the projected probability

$$\boldsymbol{P}_X(x) = \boldsymbol{b}_X(x) + \boldsymbol{a}_X(x) u_X, \quad \forall x \in \mathbb{X}. \tag{6}$$

This projected probability represents the expected outcome of the opinion in a classical probability space.

Various fusion operators are available for merging opinions from multiple sources about a common domain of interest. The idea is to bring together a set of sources to reach a joint conclusion about a specific task through the use of a fusion operator. It is important to note that the suitability of a fusion operator varies depending on the task at hand. In our work, we use the *cumulative belief fusion* (CBF) and the *averaging belief fusion* (ABF) [21], [22].

Definition 2 (Cumulative Belief Fusion)**.** *Let $\mathbb{A}$ be a set of $N \in \mathbb{N}$ sources represented by $A_1, \ldots, A_N$, each having a different opinion $\omega_X^{A_1}, \ldots, \omega_X^{A_N}$ on the same variable X. The fused opinion of all sources in the set $\mathbb{A}$ is referred to $\diamond(\mathbb{A})$ such that $\omega_X^{\diamond(\mathbb{A})} = \left(\boldsymbol{b}_X^{\diamond(\mathbb{A})}, u_X^{\diamond(\mathbb{A})}, \boldsymbol{a}_X^{\diamond(\mathbb{A})}\right)$. This is calculated for $\exists A \in \mathbb{A} : u_X^A \neq 0$ and $\exists A \in \mathbb{A} : u_X^A \neq 1$ by*

$$\boldsymbol{b}_X^{\diamond(\mathbb{A})}(x) = \frac{\sum\limits_{A \in \mathbb{A}} \left(\boldsymbol{b}_X^A(x) \prod\limits_{A^\star \neq A} u_X^{A^\star} \right)}{\sum\limits_{A \in \mathbb{A}} \left(\prod\limits_{A^\star \neq A} u_X^{A^\star} \right) - (N-1) \prod\limits_{A \in \mathbb{A}} u_X^A}, \tag{7a}$$

$$u_X^{\diamond(\mathbb{A})} = \frac{\prod\limits_{A \in \mathbb{A}} u_X^A}{\sum\limits_{A \in \mathbb{A}} \left(\prod\limits_{A^\star \neq A} u_X^{A^\star} \right) - (N-1) \prod\limits_{A \in \mathbb{A}} u_X^A}, \tag{7b}$$

$$\boldsymbol{a}_X^{\diamond(\mathbb{A})}(x) = \frac{\sum\limits_{A \in \mathbb{A}} \left(\boldsymbol{a}_X^A(x) \prod\limits_{A^\star \neq A} u_X^{A^\star} \right) - \sum\limits_{A \in \mathbb{A}} \boldsymbol{a}_X^A(x) \prod\limits_{A \in \mathbb{A}} u_X^A}{\sum\limits_{A \in \mathbb{A}} \left(\prod\limits_{A^\star \neq A} u_X^{A^\star} \right) - N \left(\prod\limits_{A \in \mathbb{A}} u_X^A \right)}. \tag{7c}$$

The general notation of CBF is often abbreviated as $\omega_X^{\diamond(\mathbb{A})} = \bigoplus\limits_{A \in \mathbb{A}} \left(\omega_X^A\right) = \omega_X^{A_1} \oplus \ldots \oplus \omega_X^{A_N}$. Here, associativity, commutativity, and non-idempotent can be verified.

CBF assumes that incorporating additional, independent sources of evidence will accumulate and strengthen the overall belief. ABF [21], [22], on the other hand, takes into account the interdependence between sources and assumes that adding more sources does not necessarily lead to a stronger conclusion with lower uncertainty.

Definition 3 (Averaging Belief Fusion)**.** *Let $\mathbb{A}$ be a set of $N \in \mathbb{N}$ sources represented by $A_1, \ldots, A_N$, each having a different opinion $\omega_X^{A_1}, \ldots, \omega_X^{A_N}$ on the same variable X. The fused opinion of all sources in the set $\mathbb{A}$ is referred to $\underline{\diamond}(\mathbb{A})$ such that $\omega_X^{\underline{\diamond}(\mathbb{A})} = \left(\boldsymbol{b}_X^{\underline{\diamond}(\mathbb{A})}, u_X^{\underline{\diamond}(\mathbb{A})}, \boldsymbol{a}_X^{\underline{\diamond}(\mathbb{A})}\right)$. Again, this is calculated for $\exists A \in \mathbb{A} : u_X^A \neq 0$ and $\exists A \in \mathbb{A} : u_X^A \neq 1$ by*

$$\boldsymbol{b}_X^{\underline{\diamond}(\mathbb{C})}(x) = \frac{\sum\limits_{A \in \mathbb{A}} \left(\boldsymbol{b}_X^A(x) \prod\limits_{A^\star \neq C} u_X^{A^\star} \right)}{\sum\limits_{A \in \mathbb{A}} \left(\prod\limits_{A^\star \neq A} u_X^{A^\star} \right)}, \tag{8a}$$

$$u_X^{\underline{\diamond}(\mathbb{A})} = \frac{N \prod\limits_{A \in \mathbb{A}} u_X^A}{\sum\limits_{A \in \mathbb{A}} \left(\prod\limits_{A^\star \neq A} u_X^{A^\star} \right)}, \tag{8b}$$

$$\boldsymbol{a}_X^{\underline{\diamond}(\mathbb{C})}(x) = \frac{\sum\limits_{A \in \mathbb{A}} \left(\boldsymbol{a}_X^A(x) \prod\limits_{A^\star \neq A} u_X^{A^\star} \right) - \sum\limits_{A \in \mathbb{A}} \boldsymbol{a}_X^A(x) \prod\limits_{A \in \mathbb{A}} u_X^A}{\sum\limits_{A \in \mathbb{A}} \left(\prod\limits_{A^\star \neq A} u_X^{A^\star} \right) - N \left(\prod\limits_{A \in \mathbb{A}} u_X^A \right)}. \tag{8c}$$

The general notation of ABF is often abbreviated as $\omega^{\underline{\diamond}(\mathbb{C})} = \underline{\bigoplus\limits_{C \in \mathbb{C}}} \left(\omega_X^C\right) = \omega_X^{C_1} \underline{\oplus} \ldots \underline{\oplus} \omega_X^{C_N}$. Here, commutativity, idempotent and non-associativity can be verified.

Further, we evaluate the difference between two opinions about the same variable X using the *degree of conflict* (DC).

Definition 4 (Degree of Conflict)**.** *Let ω_X^A and ω_X^B be multinomial opinions of source A and B, respectively, over the same variable $X \in \mathbb{X}$. The degree of conflict between these two opinions is defined as $\mathrm{DC}\left(\omega_X^A, \omega_X^B\right) \in [0,1]$ with*

$$\mathrm{DC}\left(\omega_X^A, \omega_X^B\right) = \mathrm{PD}\left(\omega_X^A, \omega_X^B\right) \cdot \mathrm{CC}\left(\omega_X^A, \omega_X^B\right), \tag{9}$$

using the projected distance $\mathrm{PD}\left(\omega_X^A, \omega_X^B\right) = \frac{1}{2} \sum_{x \in \mathbb{X}} |\boldsymbol{P}_X^A(x) - \boldsymbol{P}_X^B(x)| \in [0,1]$ and the conjunctive certainty $\mathrm{CC}\left(\omega_X^A, \omega_X^B\right) = \left(1 - u_X^A\right)\left(1 - u_X^B\right) \in [0,1]$.

A smaller DC means the opinions are more similar, while a larger DC indicates more significant differences. Conflicting opinions from multiple sources about X may suggest some statements about trust, which is why we later turn to the concept of trust revision in SL. In Section V-B, when proposing our multi-sensor overall assessment method, we then go into trust revision in greater detail. Further insight into trust revision can be found in [6], [23].

IV. Self-Assessment For Nonlinear Kalman Filtering

The presented SL SA for Kalman filtering is mainly based on [7], [11]. A detailed algorithm and its comprehensive explanation of the SA method for the linear case can be found in [7]. Some extensions, like a threshold derivation for the SL comparison measure DC, are presented in [11]. In the following, we concisely explain the significant points of the SA procedure based on SL with a primary focus on its extension for nonlinear filtering.

At the beginning of each time step, we obtain an SL multinomial opinion $\omega_X^{\boldsymbol{z}}$ based on the incoming measurements $\boldsymbol{z} \in \mathbb{R}^m$. This multinomial opinion $\omega_X^{\tilde{\boldsymbol{z}}}$ models the correctness of the Kalman filter's assumptions with respect to the incoming measurements. Hence, we need to calculate the measurement prediction consisting of mean $\hat{\boldsymbol{z}}$ and its innovation covariance $\boldsymbol{S}$. For nonlinear Kalman filtering, the measurement prediction is approximately calculated as Gaussian; see the EKF and UKF described in Section III-A. Thus, a linearization error can be injected and assessed here. In contrast, this is obtained in closed form for linear filtering.

Using the linearized measurement prediction of nonlinear Kalman filtering, the transformed measurement

$$\tilde{\boldsymbol{z}} = \boldsymbol{S}^{-1/2} \left(\boldsymbol{z} - \hat{\boldsymbol{z}} \right) \tag{10}$$

can be obtained, where $\tilde{\boldsymbol{z}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$ with $\tilde{\boldsymbol{z}} \in \mathbb{R}^m$ and $\boldsymbol{0} = [0, \ldots, 0]^T \in \mathbb{R}^m$ denotes a zero vector and $\boldsymbol{I} \in \mathbb{R}^{m \times m}$ the identity matrix. For calculating $\boldsymbol{S}^{-1/2}$, the Cholesky factorization can be used [2]. To obtain the SL opinion about the incoming measurement $\omega_X^{\tilde{\boldsymbol{z}}}$, the evidence vector $\boldsymbol{r}_X(x)$ for $x \in \mathbb{X}$ in the underlying evidence space of SL [6] is updated according to

$$\tilde{\boldsymbol{z}} \to x \;\Rightarrow\; \boldsymbol{r}_X(x) = \boldsymbol{r}_X(x) + 1, \quad x \in \mathbb{X}. \tag{11}$$

Thus, we can obtain a multinomial opinion

$$\omega_X^{\tilde{\boldsymbol{z}}} = (\boldsymbol{b}_X^{\tilde{\boldsymbol{z}}}, u_X^{\tilde{\boldsymbol{z}}}, \boldsymbol{a}_X) \tag{12}$$

using the bijective mapping theorem in SL between Dirichlet probability density functions with evidence vector notation and multinomial opinions (see e.g. [6]).

Besides the opinion of the incoming measurements, two additional opinions, ω_X^{st} and ω_X^{lt}, are calculated to represent a short-term and long-term memory, respectively. Every time step, the newly created opinion $\omega_X^{\tilde{\boldsymbol{z}}}$ is fused with ω_X^{st} using the CBF operator (see Definition 2) to update the short-term memory. After a certain number of measurement opinions $n_{\text{st}} \in \mathbb{N}$ has been accumulated and fused into ω_X^{st}, a long-term short-term comparison is performed using the DC (9) with ω_X^{st} and ω_X^{lt}. If the DC exceeds a derived threshold using [11], indicating a conflict between short-term and long-term opinion, we overwrite our long-term opinion ω_X^{lt} with the short-term opinion ω_X^{st} and reset the latter. Conversely, if the DC is below a threshold, we use CBF to fuse the long-term and short-term opinion, resulting in our final opinion

$$\omega_X = (\boldsymbol{b}_X, u_X, \boldsymbol{a}_X) \tag{13}$$

over a time interval of consistency.

To obtain our SA scores, we use the resulting opinion ω_X and the related reference opinion ω_X^{ref}, which represents the initial Kalman filter assumptions. For performing the comparison, we again use the DC as follows:

$$\delta = \text{DC}\left(\omega_X, \omega_X^{\text{ref}}\right). \tag{14}$$

With this, we obtain the online SA measures

$$(\delta, u_X, \eta), \tag{15}$$

which are the assessment score $\delta \in [0, 1]$ itself, the associated uncertainty $u_X \in [0, 1]$, and, additionally, a corresponding confidence-level threshold $\eta \in [0, 1]$. The threshold η is directly derived from the opinion ω_X and a certain confidence level $\alpha \in [0, 1]$; see [11]. It specifies a limit for the SA score δ with respect to the given confidence level α. In our case, exceeding the threshold η indicates that the assumptions made in the filter algorithm are violated.

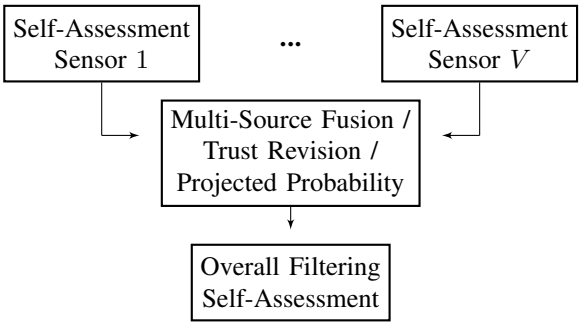

Fig. 2. Our three proposed overall assessment methods for multi-sensor systems directly utilize the single-sensor self-assessments and use the SL multi-source fusion, trust revision, and projected probability concept.

V. Overall Assessment For Multi-Sensor Systems

This section proposes an overall online assessment framework to monitor the general performance of a multi-sensor filtering system. Specifically, we propose three different concepts for implementing an overall assessment mechanism using concepts of SL; see Fig. 2. This enables a closed-form solution in SL which directly utilizes our single-sensor SA modules presented in Section IV for the multi-sensor assessment. First, we use classical multi-source fusion of SL to obtain a fundamental overall online performance measure. Second, we include the trust revision concept of SL into

the approach based on multi-source fusion in order to handle opinion conflicts better. Here, source trust refers to the reliability of the sources, where the sources, in our case, are the single-sensor SAs. Third, we take advantage of the projected probability concept of SL to calculate an overall online performance measure. A comparison overall assessment measure for the filter performance is the NEES. However, the NEES requires GT data that are not available in our considered online applications.

Our starting point for developing the three overall assessment methods is the multi-sensor Kalman filter described in Section III-A. During the sequential update, we apply the single-sensor SA mechanism described in Section IV to each sensor $s \in \mathbb{S} = \{1, \dots, V\}$ with $V \in \mathbb{N}$ using the common prediction for all sensors. This yields the evidence-based opinions $\omega_X^{(s)}$ from (13) and the corresponding single-sensor SA measures $\left(\delta^{(s)}, u_X^{(s)}, \eta^{(s)}\right)$ from (15) for each sensor $s \in \mathbb{S}$.

A. Multi-Source Fusion

Based on all sensor-specific SAs, a comprehensive fusion is enabled that combines and unifies all collected evidence from all sensors as a whole. Note that, in this context, past evidence from each sensor s is implicitly taken into account and represented by $\omega_X^{(s)}$. In SL, there are multiple possibilities to fuse opinions [6]. Within this work, two fundamental fusion operators are considered, presented in Section III-B. These are the CBF operator, which accumulates evidence, and the ABF operator, which averages evidence. Note that performing the SA mechanism for each sensor using partially the same information, such as the measurement prediction, can lead to dependencies. Furthermore, adding more single-sensor SA opinions does not necessarily mean more evidence supports the conclusion. Therefore, for the multi-source SA fusion approach, the opinions are fused using the ABF. Then, an overall assessment opinion ω_X^{all} can be obtained by

$$\omega_X^{\text{all}} = \omega_X^{(\diamondsuit s)} = \omega_X^{(1)} \underline{\oplus} \omega_X^{(2)} \underline{\oplus} \dots \underline{\oplus} \omega_X^{(V)} \tag{16}$$

using ABF according to Definition 3. The overall assessment measure is then calculated by

$$\delta^{\text{all}} = \text{DC}\left(\omega_X^{\text{all}}, \omega_X^{\text{ref}}\right). \tag{17}$$

where ω_X^{ref} is the corresponding reference opinion, which can be obtained similarly to (16) using the single-sensor reference opinions. With this, we get the SA measures for the overall filtering performance

$$\left(\delta^{\text{all}}, u_X^{\text{all}}, \eta^{\text{all}}\right), \tag{18}$$

consisting of the overall assessment score $\delta^{\text{all}} \in [0, 1]$, the corresponding overall uncertainty $u_X^{\text{all}} \in [0, 1]$, and the new corresponding threshold $\eta^{\text{all}} \in [0, 1]$, which are all calculated and automatically obtained by the SA mechanism. When the overall assessment score δ^{all} exceeds the threshold η^{all}, the SA module indicates a violation of the filter assumptions to a certain confidence level. It is worth mentioning that with the proposed SL SA framework, it is possible to easily fuse and combine each sensor's SA outputs. In contrast, combining classical statistical tests for each sensor, such as the NIS, into an overall assessment is challenging and not yet available in the literature.

B. Trust Revision

The second approach to obtain an overall performance assessment is again based on all sensor-specific SAs but additionally includes the SL concept trust revision before continuing with the first approach using SL fusion. So far, beyond the usual modeling of the measurement uncertainties in the filtering process, the initial trust in the single-sensor SA opinions for the fusion process is equal from the SL perspective. The concept of trust revision [6], [23] enables the additional consideration of trust and weighting for all single-sensor SA statements. The initial step is to establish the reliability-based weighting binomial opinion $\omega_{\text{weight}}^{(s)}$ for each sensor s, which is initialized as a dogmatic opinion with total certainty and complete reliability. In order to update $\omega_{\text{weight}}^{(s^\star)}$ of a specific sensor $s^\star \in \mathbb{S}$, the belief conflict (BC) of $s^\star$ regarding a fusion reference is calculated first. The BC of sensor $s^\star$ represents the conflict existing to the majority belief. For the majority belief, we determine the fusion reference using ABF

$$\omega_X^{\text{FR}} = \omega_X^{(1)} \underline{\oplus} \omega_X^{(2)} \underline{\oplus} \dots \underline{\oplus} \omega_X^{(V)}. \tag{19}$$

With this, the BC is calculated for sensor $s^\star$ using

$$\text{BC}\left(\omega_X^{(s^\star)}\right) = \text{DC}\left(\omega_X^{(s^\star)}, \omega_X^{\text{FR}}\right). \tag{20}$$

Additionally, the maximum conflict (MC)

$$\text{MC}\left(\omega_X^{(\mathbb{S})}\right) = \max_{s \in \mathbb{S}} \text{BC}\left(\omega_X^{(s)}\right) \tag{21}$$

and the average conflict (AC)

$$\text{AC}\left(\omega_X^{(\mathbb{S})}\right) = \frac{1}{V} \sum_{s \in \mathbb{S}} \text{BC}\left(\omega_X^{(s)}\right) \tag{22}$$

over the sensor set $\mathbb{S}$ are generated, which are used to calculate the revision weight (RW). Thus, the RW is calculated according to

$$\text{RW}\left(\omega_X^{(s^\star)}\right) = \begin{cases} \frac{\text{MC}\left(\omega_X^{(\mathbb{S})}\right)\left(\text{BC}\left(\omega_X^{(s^\star)}\right) - \text{AC}\left(\omega_X^{(\mathbb{S})}\right)\right)}{\text{MC}\left(\omega_X^{(\mathbb{S})}\right) - \text{AC}\left(\omega_X^{(\mathbb{S})}\right)} & , \ \tilde{d} > 0 \\ 0 & , \ \text{else} \end{cases} \tag{23}$$

with $\tilde{d} = \text{BC}\left(\omega_X^{(s^\star)}\right) - \text{AC}\left(\omega_X^{(\mathbb{S})}\right)$. The revised weighting opinion $\check{\omega}_{\text{weight}}^{(s^\star)} = \left(\check{b}_{\text{weight}}^{(s^\star)}, \check{d}_{\text{weight}}^{(s^\star)}, \check{u}_{\text{weight}}^{(s^\star)}, \check{a}_{\text{weight}}^{(s^\star)}\right)$ consequently consists of

$$\check{b}_{\text{weight}}^{(s^\star)} = b_{\text{weight}}^{(s^\star)} - b_{\text{weight}}^{(s^\star)} \cdot \text{RW}\left(\omega_X^{(s^\star)}\right), \tag{24a}$$

$$\check{d}_{\text{weight}}^{(s^\star)} = d_{\text{weight}}^{(s^\star)} + \left(1 - d_{\text{weight}}^{(s^\star)}\right) \cdot \text{RW}\left(\omega_X^{(s^\star)}\right), \tag{24b}$$

$$\check{u}_{\text{weight}}^{(s^\star)} = u_{\text{weight}}^{(s^\star)} - u_{\text{weight}}^{(s^\star)} \cdot \text{RW}\left(\omega_X^{(s^\star)}\right), \tag{24c}$$

$$\check{a}_{\text{weight}}^{(s^\star)} = a_{\text{weight}}^{(s^\star)}. \tag{24d}$$

This opinion is a binomial SL opinion, a special case of multinomial opinions, with a binary domain meaning that the belief can be explicitly expressed by belief and disbelief as denoted. Note that in the following, all revised elements are marked by ' ˇ '. After the revised weighting opinions $\check{\omega}_{\text{weight}}^{(s^\star)}$ have been calculated for all sensors $s^\star \in \mathbb{S}$, SL trust-discounting [6] is applied to $\omega_X^{(s^\star)}$ using $\check{\omega}_{\text{weight}}^{(s^\star)}$. To do this, the projected probability $\check{P}_{\text{weight}}^{(s^\star)}$ of $\check{\omega}_{\text{weight}}^{(s^\star)}$ is calculated using (6), which is then applied to $\omega_X^{(s^\star)}$ such that

$$\check{\boldsymbol{b}}_X^{(s^\star)}(x) = \check{P}_{\text{weight}}^{(s^\star)} \boldsymbol{b}_X^{(s^\star)}(x), \tag{25a}$$

$$\check{u}_X^{(s^\star)} = 1 - \check{P}_{\text{weight}}^{(s^\star)} \sum\nolimits_{x \in \mathbb{X}} \boldsymbol{b}_X^{(s^\star)}(x), \tag{25b}$$

$$\check{\boldsymbol{a}}_X^{(s^\star)}(x) = \boldsymbol{a}_X^{(s^\star)}(x). \tag{25c}$$

With the revised single-sensor SA opinions $\check{\omega}_X^{(s^\star)} = \left(\check{\boldsymbol{b}}_X^{(s^\star)}, \check{u}_X^{(s^\star)}, \check{\boldsymbol{a}}_X^{(s^\star)}\right)$, the SL multi-source fusion approach from Section V-A can be used from now on. For the fusion process in (16), the CBF operator is selected in this case because the revised single-sensor SA opinions make an accumulation of evidence suitable. By applying trust revision, conflicts are already reduced before the actual fusion process. The outputs of the trust revision approach are, again, overall SA measures $\left(\delta^{\text{all}}, u_X^{\text{all}}, \eta^{\text{all}}\right)$.

C. Projected Probability

The third approach for the overall assessment of the multi-sensor system can be developed by using the projected probability operator in (6). Instead of the first two approaches utilizing the single-sensor SA opinions $\omega_X^{(s)}$, the projected probability approach directly utilizes the outputs of the single-sensor SAs, i.e., $\left(\delta^{(s)}, u_X^{(s)}, \eta^{(s)}\right), \forall s \in \mathbb{S}$. Based on these outputs, binary information can be extracted in each time step. This information is generated by simply checking whether the considered assessment measures $\delta^{(s)}$ exceed the corresponding threshold $\eta^{(s)}$ or not. To this end, the binary evidence vector $\boldsymbol{r} = [r_{\text{true}}, r_{\text{false}}]$ is generated using

$$r_{\text{true}} = \frac{1}{V} \sum_{s \in \mathbb{S}} \mathbb{1}\left\{\delta^{(s)} < \eta^{(s)}\right\} \tag{26}$$

with $\mathbb{1}\{\cdot\}$ as the indicator function. Note that $r_{\text{true}} + r_{\text{false}} = 1$ must be fulfilled in each time step. Using the fundamental SL principle of bijective mapping [6], this evidence vector $\boldsymbol{r}$ can be mapped to a binomial opinion $\omega_{\boldsymbol{r}}$ in each time step. These opinions are then collected using a short-term sliding window of n_{st} time steps by fusing $\omega_{\boldsymbol{r}}$ into $\omega_{\boldsymbol{r}}^{\text{st}}$ based on CBF, i.e., $\omega_{\boldsymbol{r}}^{\text{st}} = \omega_{\boldsymbol{r}}^{\text{st}} \oplus \omega_{\boldsymbol{r}}$. Finally, the projected probability $\boldsymbol{P}_{\boldsymbol{r}}$ of $\omega_{\boldsymbol{r}}^{\text{st}}$ using (6) provides a probability measure in $[0, 1]$ of whether the overall system is currently operating as expected. A high projected probability indicates an overall assessment that says everything runs as expected and all assumptions are fulfilled.

VI. Experiments

This section evaluates our proposed methods in challenging simulation scenarios. In these scenarios, we consider sensors measuring a single object's position in two dimensions $(x, y) \in \mathbb{R}^2$. All sensors are initially modeled identically for tracking purposes. Moreover, as a transition model, a constant turn and velocity model [2] with state vector $\boldsymbol{x} = [x, y, v_x, v_y, \phi]^T$ is used, which leads to using a nonlinear Kalman filter. For the purpose of clarity in the plots, we only show the results using the UKF because the EKF results are basically similar. We have determined the parameters for the single-sensor SL SA method as follows: n_{st} has been set to 35, which ensures sufficient evidence for the formation of a dependable short-term opinion; for opinion building, we use a domain with $W = |\mathbb{X}| = 7$; the threshold is calculated based on a 0.1 confidence level. If not stated otherwise, we consider the time-average versions for NIS and NEES with a window size of 35 and 99% confidence intervals.

A. Single-Sensor Nonlinear Kalman Filtering

In the first simulation scenario, we consider a single-sensor single-object simulation setup to evaluate our proposed method in Section IV. In time steps 300 to 600, the true measurement noise is abruptly increased by a factor of 2. Then, from time step 900 to 1400, the true measurement noise is first increased by a drift to the level 2.5 times the original noise level and afterward again decreased by a drift to the original noise level of $\sigma_w = 1.0\,\text{m}$. Furthermore, from time step 1700 to 2000, measurement outliers with an additional noise factor of 3 to the original noise level occur every 30 time step. Finally, the GT movement of the object to be tracked shows a higher nonlinearity from time step 2300 to 2600. All these disturbances are not known a priori to the filter method, which uses its default parameterization for the entire simulation.

Fig. 3 visualizes the simulation results, which are averaged values of 200 Monte Carlo runs. In general, our SL-based approach and the time-average NIS show comparable results during the disturbances. Additionally, our approach tends to be slightly more conservative in reacting to changes caused by disturbances, which can provide additional safety margins in certain scenarios. However, it is important to note that our approach is designed to react only when statistically significant changes are detected, which ensures that the response is reliable and not affected by false alarms. This can especially be seen for the disturbance by outliers, where our method does not report any violation, whereas the time-average NIS exceeds its confidence interval. Although the NEES also exceeds its confidence interval, it is questionable whether an SA system should continuously report a violation of the filter assumption when outliers occur only once every 30 time steps.

B. Multi-Sensor Overall Filtering Assessment

In the second simulation scenario, we consider a multi-sensor single-object simulation setup to evaluate our three proposed approaches for an overall assessment of multi-sensor Kalman filters from Section V. Here, 5 sensors deliver measurements to track the object. In time steps 300 to 600, the true measurement noise of all 5 sensors is abruptly disturbed with a jump by a factor of 2. Then, from time step 900 to 1200, the true measurement noise of only 4 sensors is disturbed with

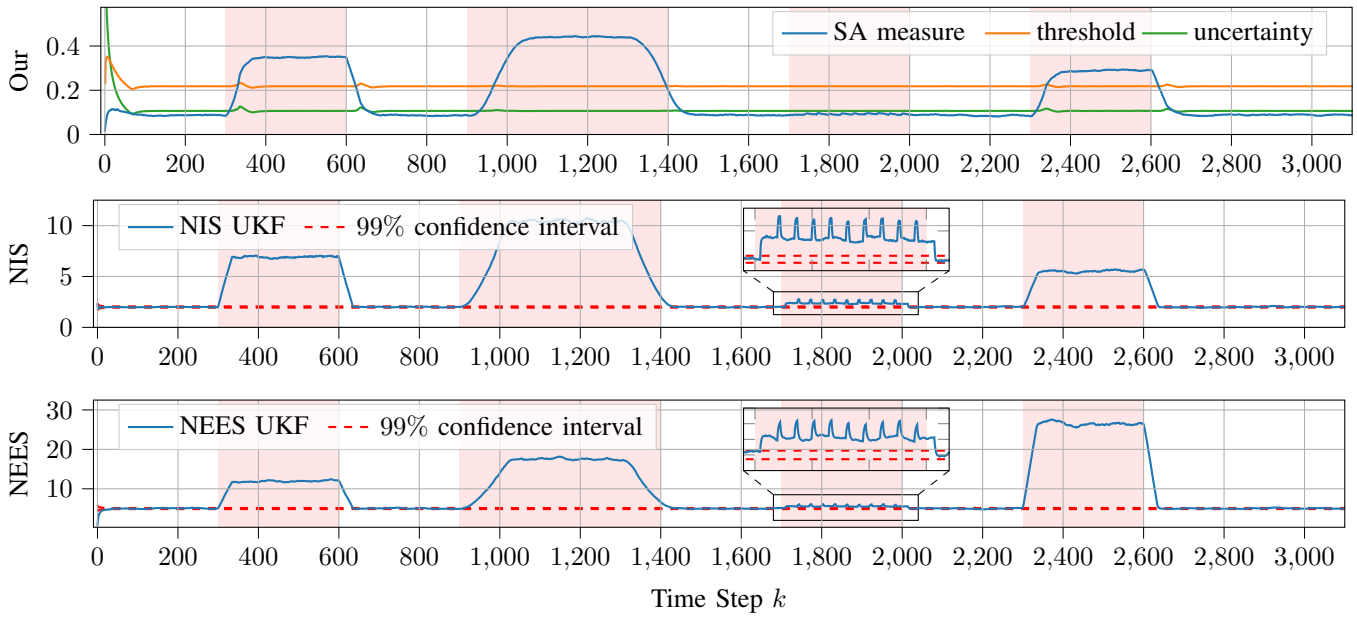

Fig. 3. Results of our SL SA in a challenging nonlinear simulation scenario with a single sensor. First, the measurement noise is disturbed by a jump, a drift, and multiple outliers, and then a higher nonlinearity is consecutively added as a disturbance. These disturbances are indicated with red shadow areas. Our SL-based SA method, consisting of an assessment measure, threshold, and uncertainty, is compared to the time-average NIS and referenced with the time-average NEES. The results are averaged values from 200 Monte Carlo runs.

a jump by a factor of 2, while one sensor functions normally. Subsequently, the number of disturbed sensors by a jump in the true measurement noise decreases gradually until only one sensor is disturbed from time step 2700 to 3000. Again, all these disturbances are not known a priori to the filter method.

The simulation results from 200 Monte Carlo runs are illustrated in Fig. 4. The multi-source fusion and trust revision approaches have the same SA output types as the single-sensors SA modules. Here, the SL-based approach with trust revision shows comparable results when referenced to the time-averaged NEES. Both methods report violations by exceeding the threshold or confidence interval at the same time. However, the SL-based approach does not report a violation if only one sensor is disturbed. We want to emphasize that our proposed method is able to perform these results online without using GT data, whereas the time-average NEES requires GT data and thus cannot be applied for online assessment. Our approach based on the projected probability provides a reliability score in $[0, 1]$, where a score near one means that the estimation results are reliable with respect to the filter assumptions. This also shows that the filter estimation is strongly affected when all sensors are disturbed, which is correctly detected in the overall assessment score based on the projected probability. On the other hand, if only one sensor is disturbed, the projected probability does not indicate a loss of reliability of the estimation results to the same extent. For the applied multi-sensor Kalman filter setup (see Section III-A), using the NIS as a comparison measure for this application is not appropriate because the sequential update concept leads to NIS values for each individual sensor and does not provide an overall measure.

VII. Conclusion

In this work, we first extended our SL-based SA approach to nonlinear Kalman filtering. Compared to the NIS, our approach shows similar results in a simulation scenario with multiple assumption violations, while being slightly more robust in the SA when dealing with occasional outliers. Moreover, we investigated an overall assessment for multi-sensor systems in the Kalman filter framework based on our developed SL SA approach. Thus, we proposed three approaches for an overall assessment of the filter performance, which directly utilize our single-sensor SAs and fundamental concepts of SL. In simulation scenarios, we showed that with our proposed approach, we are even able to reach similar results as the NEES but without requiring GT data. Thus, unlike the NEES, our proposed method is suitable for online applications such as autonomous driving that require an online SA.

In future work, we want to use our SA results to adapt the filter assumptions online and, thus, improve the filter performance through this automatic adaptation process. Furthermore, we want to incorporate our proposed approaches in the works of [11] to enhance the overall SA framework for single-object tracking in clutter.

References

[1] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*. Academic Press, New York, 1988.

[2] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorthims and Software: Theory Algorithms and Software*. John Wiley & Sons, 2001.

[3] M. Scalzo, G. Horvath, E. Jones, A. Bubalo, M. Alford, R. Niu, and P. K. Varshney, "Adaptive filtering for single target tracking," in *Signal Processing, Sensor Fusion, and Target Recognition XVIII*, vol. 7336. SPIE, 2009, pp. 172–181.

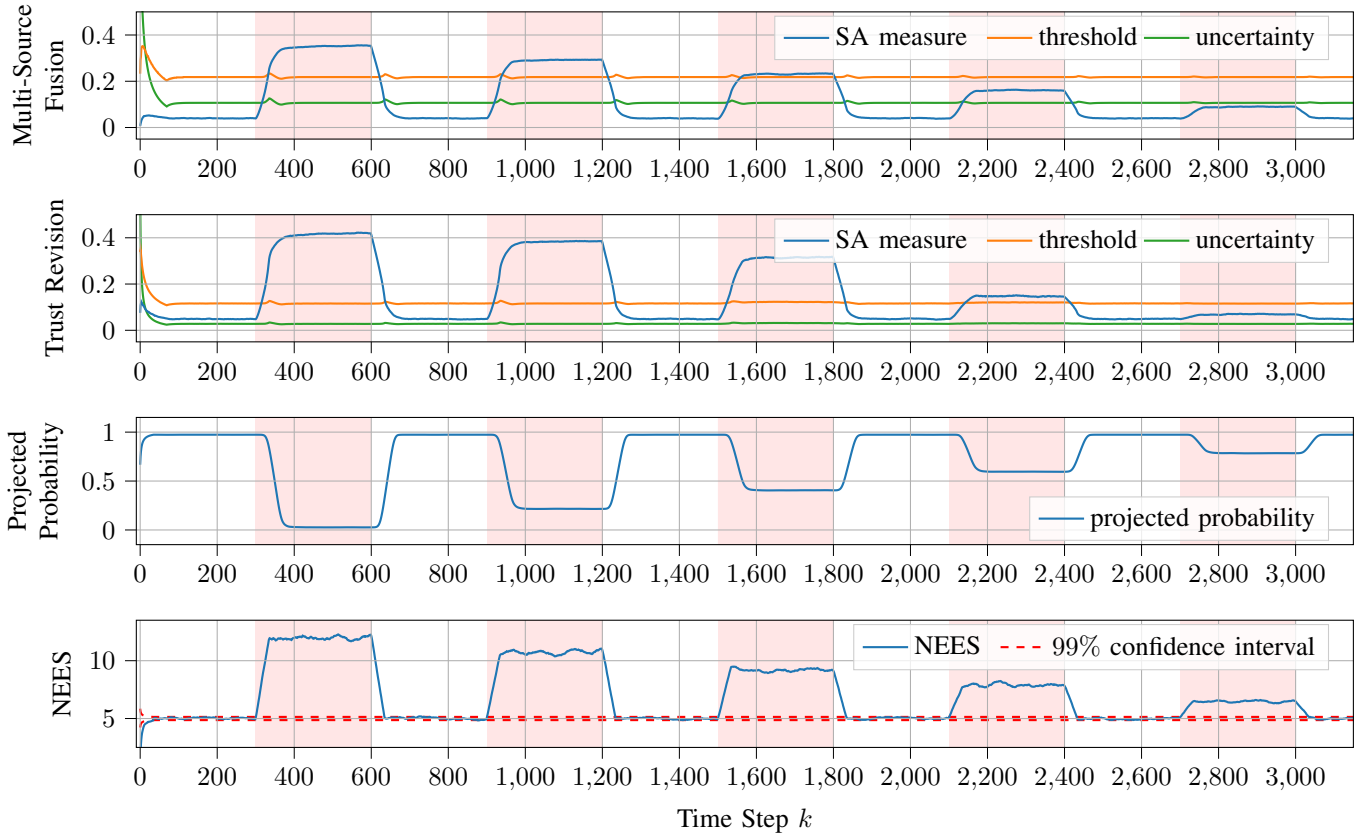

Fig. 4. Results of our proposed overall assessment of multi-sensor Kalman filters in a challenging nonlinear simulation scenario with 5 sensors. First, the measurement noises of all 5 sensors are disturbed simultaneously by jumps. The amount of disturbed sensors decreases gradually until only one sensor is disturbed. These disturbances are indicated with red shadow areas. Our three proposed approaches using the SL concepts of multi-source fusion, trust revision, and projected probability are compared to the GT data required time-average NEES. The results are averaged values from 200 Monte Carlo runs.

[4] R. Niu, P. K. Varshney, M. Alford, A. Bubalo, E. Jones, and M. Scalzo, "Curvature nonlinearity measure and filter divergence detector for nonlinear tracking problems," in *2008 11th International Conference on Information Fusion*, 2008, pp. 1–8.

[5] International Organization for Standardization, "ISO/PAS 21448: Road Vehicles —– Safety of the Intended Functionality," *ISO, Publicly Available Specification*, 2019.

[6] A. Jøsang, *Subjective Logic: A Formalism for Reasoning Under Uncertainty*. Springer International Publishing, 2016.

[7] T. Griebel, J. Müller, M. Buchholz, and K. Dietmayer, "Kalman filter meets subjective logic: A self-assessing Kalman filter using subjective logic," in *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*. IEEE, 2020, pp. 1–8.

[8] P. Ivanov, S. Ali-Löytty, and R. Piché, "Evaluating the consistency of estimation," in *International Conference on Localization and GNSS 2014 (ICL-GNSS 2014)*, 2014, pp. 1–5.

[9] J. Duník, O. Straka, and M. Šimandl, "Self-assessment of local filters by non-Gaussianity measures," in *2014 American Control Conference*, 2014, pp. 1723–1728.

[10] R. Mahler, "Divergence detectors for multitarget tracking algorithms," in *Signal Processing, Sensor Fusion, and Target Recognition XXII*, vol. 8745. International Society for Optics and Photonics, 2013, p. 87450F.

[11] T. Griebel, J. Müller, P. Geisler, C. Hermann, M. Herrmann, M. Buchholz, and K. Dietmayer, "Self-assessment for single-object tracking in clutter using subjective logic," in *2022 25th International Conference on Information Fusion (FUSION)*. IEEE, 2022, pp. 1–8.

[12] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME - Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[13] Y. Ho and R. Lee, "A Bayesian approach to problems in stochastic estimation and control," *IEEE transactions on automatic control*, vol. 9, no. 4, pp. 333–339, 1964.

[14] A. H. Jazwinski, *Stochastic processes and filtering theory*. Courier Corporation, 2007.

[15] H. W. Sorenson, *Kalman filtering: theory and application*. IEEE, 1985.

[16] S. J. Julier and J. K. Uhlmann, "New extension of the Kalman filter to nonlinear systems," in *Signal processing, sensor fusion, and target recognition VI*, vol. 3068. Spie, 1997, pp. 182–193.

[17] ——, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[18] D. Willner, C. B. Chang, and K. P. Dunn, "Kalman filter algorithms for a multi-sensor system," in *1976 IEEE Conference on Decision and Control including the 15th Symposium on Adaptive Processes*, 1976, pp. 570–574.

[19] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *The Annals of Mathematical Statistics*, vol. 38, no. 2, pp. 325–339, 1967.

[20] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976, vol. 42.

[21] A. Jøsang, D. Wang, and J. Zhang, "Multi-source fusion in subjective logic," in *2017 20th International Conference on Information Fusion (FUSION)*. IEEE, 2017, pp. 1–8.

[22] R. W. Van Der Heijden, H. Kopp, and F. Kargl, "Multi-source fusion operations in subjective logic," in *2018 21st International Conference on Information Fusion (FUSION)*. IEEE, 2018, pp. 1990–1997.

[23] A. Jøsang, J. Zhang, and D. Wang, "Multi-source trust revision," in *2017 20th International Conference on Information Fusion (FUSION)*. IEEE, 2017, pp. 1–8.