Cooperative Unscented Kalman Filter with Bank of Scaling Parameter Values

J. Duník, O. Straka
Deparment of Cybernetics
University of West Bohemia
Univerzitní 8, 306 14 Pilsen, Czech Republic
E-mails: {dunikj, straka30}@kky.zcu.cz

U. D. Hanebeck
Institute for Anthropomatics and Robotics
Karlsruhe Institute of Technology
Adenauerring 2, 76131 Karlsruhe, Germany
E-mail: uwe.hanebeck@kit.edu

***Abstract*—This paper is devoted to the Bayesian state estimation of the nonlinear stochastic dynamic systems. The stress is laid on Gaussian unscented Kalman filter (UKF) and, in particular, on a setting of its scaling parameter, which significantly affects the UKF estimation performance. Compared to the standard UKF design, where one scaling parameter per a time instant is selected, the proposed cooperative UKF combines estimates of the set of UKFs each designed with different value of the scaling parameter. The cooperative UKF reformulates the UKF scaling parameter selection task as the multiple model approach, which allows to extract more information from the measurement to provide estimates of better quality as indicated by the numerical simulations.**

Keywords: Nonlinear filtering; Gaussian estimators; Bayesian relations.

I. Introduction

State estimation of discrete-time stochastic dynamic systems from noisy or incomplete measurements has been a subject of considerable research interest for the last decades. The topic plays an important role in various fields such as navigation, speech and image processing, fault detection, and adaptive or optimal control.

Following the Bayesian approach, a general solution to the state estimation problem is given by the Bayesian recursive relations (BRRs) for computing the probability density functions (PDFs) of the state conditioned on the measurements. The conditional PDFs provide a full description of the immeasurable state. The relations are, however, analytically tractable for a limited set of models only for which the linearity is usually a common factor. This class of exact Bayesian estimators is represented, e.g., by the Kalman filter (KF). In other cases, an approximate solution to the BRRs has to be employed. These approximate filtering methods can be divided with respect to the validity of the estimates into global and local filters [1], [2].

Global filters, such as the particle or the point-mass filter, provide estimates in the form of conditional PDFs without any assumption on the conditional distribution family. These global filters are capable of estimating the state of a strongly nonlinear or non-Gaussian system but usually at the cost of higher computational demands. As opposed to global filters, *local*, or alternatively *Gaussian*, filters (GF) provide computationally efficient estimates predominantly in the form of the conditional mean and covariance matrix[1] with potentially limited performance due to inherent underlying Gaussian assumption[2]. This paper focuses on a class of GFs that rely on the specification of a *scaling parameter*. This class includes many modern widely-used derivative-free GFs, e.g., the unscented Kalman filter (UKF) with the scaling parameter κ affecting the σ-points spreading or divided-difference filter with the scaling parameter h affecting the interval, where the differences are computed [3]–[7].

The considered scaling parameter *significantly* affects the GF performance and its selection is a challenging task [6], [8]–[12]. In literature, various recommendations and algorithms for scaling parameter selection have been proposed. Unfortunately, they often result in contradictory scaling parameter setting, require some input from the filter user (or designer), and are designed for the GF filtering step only (i.e., for the measurement update). The scaling parameter can be selected according to

- *fixed* schemes and
- *adaptive* schemes.

Fixed schemes include "standard" scaling parameter settings based on the state dimension [3], [13], which is constant for all time instants, or the off-line tuned scaling parameter, which leads to a (time-varying) sequence of the scaling parameter setting found prior to the estimation experiment by a numerical analysis [8], [9], [11]. Although, such setting does not affect GF computational complexity, it cannot reflect the GF actual working conditions. On the other hand, adaptive schemes set scaling parameter online with respect to the current working conditions, but at the cost of higher computational complexity [6], [12]. Moreover, as adaptive schemes are based on a numerical solution, the user is required to chose a criterion and a scaling parameter range for which the optimization is performed. Often, the numerical optimization evaluates the criterion (and typically also a part of the GF filtering step) for multiple choices of the scaling parameter and

[1]The first two moments usually do not represent a full description of the immeasurable state.

[2]This assumption is not always realistic especially for strongly nonlinear systems.

then, select the most suitable one only. All intermediate results are forgotten. Thus, this concept of scaling parameter adaptation can be seen as *competitive* winner-takes-all scheme.

The goal of this paper is to propose a conceptually novel *cooperative* approach to on-line determination of the GF scaling parameter. Its basic idea lies in combination of multiple estimates provided by a set of the GFs configured for the same estimation task, but differing in the scaling parameter setting. The cooperative approach adopts the methodology of the generalized pseudo-Bayesian estimator design with the stress on a minimization of the user interaction.

The paper is organized as follows. In Section II, system description and a brief introduction to the Bayesian state estimation is presented. Section III focuses on introduction of the UKF, standard scaling parameter selection, and motivational example. Then, in Section IV, the concept of the cooperative scaling parameter selection is proposed and the cooperative UKF is designed. Sections V and VI provide numerical simulations and concluding remarks, respectively.

II. System Description and State Estimation

A discrete-time nonlinear stochastic dynamic system, described by the state-space model

$$\mathcal{M}: \quad \mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k) + \mathbf{w}_k \ , \tag{1}$$

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k \ , \tag{2}$$

is considered for $k = 0, 1, 2, \ldots, T$, where the vectors $\mathbf{x}_k \in \mathbb{R}^{n_x}$ and $\mathbf{z}_k \in \mathbb{R}^{n_z}$ represent the state of the system and the measurement at time instant k, respectively. The state and measurement functions $\mathbf{f}_k : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ and $\mathbf{h}_k : \mathbb{R}^{n_x} \to \mathbb{R}^{n_z}$ are supposed to be known. The state noise $\mathbf{w}_k \in \mathbb{R}^{n_x}$, the measurement noise $\mathbf{v}_k \in \mathbb{R}^{n_z}$, and the initial state $\mathbf{x}_0 \in \mathbb{R}^{n_x}$ are supposed to be independent of each other.

The noises and the initial state are assumed to be normally distributed, i.e.,

$$p(\mathbf{w}_k) = \mathcal{N}\{\mathbf{w}_k; \mathbf{0}_{n_x\times 1}, \mathbf{Q}_k\} \ , \tag{3}$$

$$p(\mathbf{v}_k) = \mathcal{N}\{\mathbf{v}_k; \mathbf{0}_{n_z\times 1}, \mathbf{R}_k\} \ , \tag{4}$$

$$p(\mathbf{x}_0) = \mathcal{N}\{\mathbf{x}_0; \bar{\mathbf{x}}_0, \mathbf{P}_0\} \ , \tag{5}$$

where $\mathbf{0}_{n_x\times 1}$ is the zero matrix of indicated dimension and the notation $\mathcal{N}\{\mathbf{x}; \bar{\mathbf{x}}, \mathbf{P}\}$ stands for the Gaussian PDF of a random variable $\mathbf{x}$ with the mean $\bar{\mathbf{x}}$ and covariance matrix $\mathbf{P}$. The first two moments of the random variables (3)–(5) are supposed to be known.

A. Bayesian Approach to State Estimation

The BRRs are given by [14]

$$p(\mathbf{x}_k|\mathbf{z}^{k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1}|\mathbf{z}^{k-1}) d\mathbf{x}_{k-1} \ , \tag{6}$$

$$p(\mathbf{x}_k|\mathbf{z}^{k}) = \frac{p(\mathbf{x}_k|\mathbf{z}^{k-1}) p(\mathbf{z}_k|\mathbf{x}_k)}{p(\mathbf{z}_k|\mathbf{z}^{k-1})} \ , \tag{7}$$

where $p(\mathbf{x}_k|\mathbf{z}^{k-1})$ is the one-step *predictive* PDF computed by the Chapman-Kolmogorov equation (6) and $p(\mathbf{x}_k|\mathbf{z}^{k})$ is the *filtering* PDF computed by the Bayes' rule (7). The PDFs $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ and $p(\mathbf{z}_k|\mathbf{x}_k)$ are the state transition PDF obtained from (1) and the measurement PDF obtained from (2), respectively. The PDF $p(\mathbf{z}_k|\mathbf{z}^{k-1}) = \int p(\mathbf{x}_k|\mathbf{z}^{k-1}) p(\mathbf{z}_k|\mathbf{x}_k) d\mathbf{x}_k$ is the one-step predictive PDF of the measurement. The symbol $\mathbf{z}^k$ represents the set of all measurements up to the time instant k, i.e., $\mathbf{z}^k = [\mathbf{z}_0, \mathbf{z}_1, \ldots \mathbf{z}_k]$. The estimate of the state is given by the filtering and the predictive PDFs. The recursion (6), (7) can be started from the initial PDF $p(\mathbf{x}_0|\mathbf{z}^0)$ stemming from $p(\mathbf{x}_0)$.

B. Gaussian Filter Design

Considering the system description (1)–(5), the BRRs are not exactly solvable. GFs assume the predictive conditional joint PDF [13], i.e.,

$$p(\mathbf{x}_{k+1}, \mathbf{z}_{k+1}|\mathbf{z}^k) \triangleq \mathcal{N}\{\left[\begin{smallmatrix}\mathbf{x}_{k+1}\\ \mathbf{z}_{k+1}\end{smallmatrix}\right]; \left[\begin{smallmatrix}\hat{\mathbf{x}}_{k+1|k}\\ \hat{\mathbf{z}}_{k+1|k}\end{smallmatrix}\right], \left[\begin{smallmatrix}\mathbf{P}^{xx}_{k+1|k} & \mathbf{P}^{xz}_{k+1|k}\\ \mathbf{P}^{zx}_{k+1|k} & \mathbf{P}^{zz}_{k+1|k}\end{smallmatrix}\right]\}, \tag{8}$$

to be Gaussian, which allows analytical (but inherently *approximate*) solution to the BRRs leading to the following recursive GF estimation algorithm[3]:

Step 1: Set the time instant $k = 0$ and define an initial condition $p(\mathbf{x}_0|\mathbf{z}^0) = \mathcal{N}\{\mathbf{x}_0; \hat{\mathbf{x}}_{0|0}, \mathbf{P}^{xx}_{0|0}\}$.

Step 2: The predictive mean and covariance matrix

$$\hat{\mathbf{x}}_{k+1|k} = \int \mathbf{f}_k(\mathbf{x}_k) \mathcal{N}\{\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}^{xx}_{k|k}\} d\mathbf{x}_k \ , \tag{9}$$

$$\mathbf{P}^{xx}_{k+1|k} = \int (\mathbf{f}_k(\mathbf{x}_k) - \hat{\mathbf{x}}_{k+1|k})(\mathbf{f}_k(\mathbf{x}_k) - \hat{\mathbf{x}}_{k+1|k})^T \times \mathcal{N}\{\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}^{xx}_{k|k}\} d\mathbf{x}_k + \mathbf{Q}_k \ , \tag{10}$$

are assumed to form the Gaussian PDF $p(\mathbf{x}_{k+1}|\mathbf{z}^k) \triangleq \mathcal{N}\{\mathbf{x}_k; \hat{\mathbf{x}}_{k+1|k}, \mathbf{P}^{xx}_{k+1|k}\}$.

Step 3: The moments of the filtering estimate $p(\mathbf{x}_{k+1}|\mathbf{z}^{k+1}) \triangleq \mathcal{N}\{\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{k+1|k+1}, \mathbf{P}^{xx}_{k+1|k+1}\}$ are

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1}(\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1|k}) \ , \tag{11}$$

$$\mathbf{P}^{xx}_{k+1|k+1} = \mathbf{P}^{xx}_{k+1|k} - \mathbf{K}_{k+1} \mathbf{P}^{zz}_{k+1|k} \mathbf{K}^T_{k+1} \ , \tag{12}$$

where $\mathbf{K}_{k+1} = \mathbf{P}^{xz}_{k+1|k} (\mathbf{P}^{zz}_{k+1|k})^{-1}$ is the filter gain,

$$\hat{\mathbf{z}}_{k+1|k} = \int \mathbf{h}_{k+1}(\mathbf{x}_{k+1}) \mathcal{N}\{\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{k+1|k}, \mathbf{P}^{xx}_{k+1|k}\} d\mathbf{x}_{k+1}, \tag{13}$$

$$\mathbf{P}^{zz}_{k+1|k} = \int (\mathbf{h}_{k+1}(\mathbf{x}_{k+1}) - \hat{\mathbf{z}}_{k+1|k})(\mathbf{h}_{k+1}(\mathbf{x}_{k+1}) - \hat{\mathbf{z}}_{k+1|k})^T \times \mathcal{N}\{\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{k+1|k}, \mathbf{P}^{xx}_{k+1|k}\} d\mathbf{x}_{k+1} + \mathbf{R}_{k+1}, \tag{14}$$

$$\mathbf{P}^{xz}_{k+1|k} = \int (\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k})(\mathbf{h}_{k+1}(\mathbf{x}_{k+1}) - \hat{\mathbf{z}}_{k+1|k})^T \times \mathcal{N}\{\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{k+1|k}, \mathbf{P}^{xx}_{k+1|k}\} d\mathbf{x}_{k+1} \ . \tag{15}$$

---

[3] All GFs are linear algorithms with respect to the actual measurement and have the same structure as the KF. However, contrary to the KF, the GF estimate is "only" assumed to be Gaussian.

Let $k = k + 1$. The algorithm then continues to **Step 2**.

III. Scaling Parameter in Unscented Kalman Filter and Goal of the Paper

Evaluation of the *Gaussian* PDF weighted integrals (9), (10) and (13)–(15) is based on various linearization techniques or numerical integration rules [3]–[7], [15]–[17]. Except for a few special cases (e.g., linear or polynomial functions $\mathbf{f}_k$ and $\mathbf{h}_k$ in (1), (2)) their evaluation is approximate and, typically, the integration result significantly depends on a scaling parameter selection.

The setting and impact of the scaling parameter is discussed and illustrated using the UKF, which solves the integrals using the *unscented transform* (UT) [3], [10]. Note, however, that analogous conclusions can be drawn for any GF of the considered class.

A. Unscented Transform

The UT computes the predictive moments as weighted sample means of nonlinearly transformed sets of deterministically chosen points (so called σ-points). In particular, the UT based *approximate* measurement predictive moments (13)–(15) are computed[4] as

$$\hat{\mathbf{z}}^{\mathrm{UKF}}_{k+1|k} = \sum_{i=0}^{2n_x} \mathcal{W}^i_{k+1|k} \mathcal{Z}^i_{k+1|k} \approx \hat{\mathbf{z}}_{k+1|k} \ , \tag{16}$$

$$\begin{aligned}\mathbf{P}^{zz,\mathrm{UKF}}_{k+1|k} &= \sum_{i=0}^{2n_x} \mathcal{W}^i_{k+1|k} (\mathcal{Z}^i_{k+1|k} - \hat{\mathbf{z}}_{k+1|k})(\cdot)^T + \mathbf{R}_{k+1} \\ &\approx \mathbf{P}^{zz}_{k+1|k} \ ,\end{aligned} \tag{17}$$

$$\begin{aligned}\mathbf{P}^{xz,\mathrm{UKF}}_{k+1|k} &= \sum_{i=0}^{2n_x} \mathcal{W}^i_{k+1|k} (\mathcal{X}^i_{k+1|k} - \hat{\mathbf{x}}_{k+1|k})(\mathcal{Z}^i_{k+1|k} - \hat{\mathbf{z}}_{k+1|k})^T \\ &\approx \mathbf{P}^{xz}_{k+1|k} \ ,\end{aligned} \tag{18}$$

where the transformed σ-points $\{\mathcal{Z}^i_{k+1|k}\}_{i=0}^{2n_x}$ are

$$\mathcal{Z}^i_{k+1|k} = \mathbf{h}_{k+1}(\mathcal{X}^i_{k+1|k}) \ , \forall i \ ,$$

and the σ-points $\{\mathcal{X}^i_{k+1|k}\}_{i=0}^{2n_x}$ and respective weights $\{\mathcal{W}^i_{k+1|k}\}_{i=0}^{2n_x}$ are

$$\mathcal{X}^{0:2n_x}_{k+1|k} = \hat{\mathbf{x}}_{k+1|k}\mathbf{1}_{1\times b} + c\left[\mathbf{0}_{n_x\times 1}, \mathbf{S}^{xx}_{k+1|k}, -\mathbf{S}^{xx}_{k+1|k}\right], \tag{19}$$

$$\mathcal{W}^{0:2n_x}_{k+1|k} = \tfrac{1}{n_x+\kappa}[\kappa, \tfrac{1}{2}, \ldots, \tfrac{1}{2}] \ . \tag{20}$$

In (16)–(20), $\mathbf{1}_{a\times b}$ and $\mathbf{0}_{a\times b}$ represent matrices of ones and zeros of indicated dimension, respectively. $\mathbf{S}^{xx}_{k+1|k}$ is a factor of the covariance matrix $\mathbf{P}^{xx}_{k+1|k}$ satisfying $\mathbf{P}^{xx}_{k+1|k} = \mathbf{S}^{xx}_{k+1|k}(\mathbf{S}^{xx}_{k+1|k})^T$, $b = 2n_x + 1$ is the number of σ-points, $c = \sqrt{n_x + \kappa}$, and $\kappa \in \mathbb{R}^+$ is the *scaling parameter* typically determined by the user.

[4]State predictive moments (9), (10) are computed analogously.

B. Scaling Parameter Setting and Adaptation

In literature, various recommendations can be found. The recommendations result in either fixed time-invariant or adaptive time-variant scaling parameter κ. For fixed parameters, two recommendations can be found, i.e., $\kappa = 3 - n_x$ (if $n_x \leq 3$) minimizing the error of the fourth order term of the Taylor expansion of (16) [3] and $\kappa = 0$ leading to the cubature integration rule and the cubature KF [13]. Time-varying scaling parameters can be found either off-line [8], [9], [11] or on-line [6], [12]. The on-line identified parameter can be computed, for example, to maximize the likelihood function[5]

$$\kappa^{\mathrm{ML}}_k = \arg\max_{\kappa} \mathcal{N}\{\mathbf{z}_k; \hat{\mathbf{z}}^{\mathrm{UKF}}_{k|k-1}, \mathbf{P}^{zz,\mathrm{UKF}}_{k|k-1}\} \ , \tag{21}$$

which results in the scaling parameter value typically different from the fixed choices. With this short review, it is possible to see that there are multiple (and contradictory) choices of the scaling parameter. Each recommendation is, moreover, related to different criteria used for parameter selection. In addition, the criteria are only *superficially* related to the quantities, which are usually used for the filter performance evaluation (e.g., in terms of estimate accuracy and consistency).

C. Motivating Example

The performance of the UKF with three choices of the scaling parameter, namely two fixed values $\kappa = 2$, $\kappa = 0$, and one adaptive setting of κ_k according to (21), is illustrated using the univariate non-stationary model [18]

$$x_{k+1} = 0.5x_k + 25\tfrac{x_k}{1+x_k^2} + 20\sin(0.05k) + w_k \ , \tag{22}$$

$$z_k = \tfrac{x_k^2}{20} + v_k \ , \tag{23}$$

where $k = 0, 1, \ldots, T, T = 40$, $Q = 2$, $R = 0.1$, and $p(x_0) = \mathcal{N}\{x_0; -200, 1\}$. State estimation of the considered model state is a challenging task for any GF, including the UKF, especially if the true state is close to zero. In that region of the state space, the UKF may provide "inaccurate" estimates for a shorter or longer period of time, depending on the selected parameter κ.

In Figure 1, the filtering estimate error

$$\tilde{x}_{k|k} = x_k - \hat{x}_{k|k} \tag{24}$$

and the filtering standard deviation (STD)

$$S_{k|k} = \sqrt{P_{k|k}} \tag{25}$$

are plotted for *one* realization for three UKFs with two fixed scaling parameters and one adaptively set parameter. The figure illustrates enormous influence of the scaling parameter of the UKF performance and also a tendency of the UKF to provide *inconsistent* estimates in this case independently of the scaling parameter selection.

One of the possible *explanations* of the inconsistent behavior of the UKFs is that the considered state-of-the-art

[5]Other criteria can be found in [6].

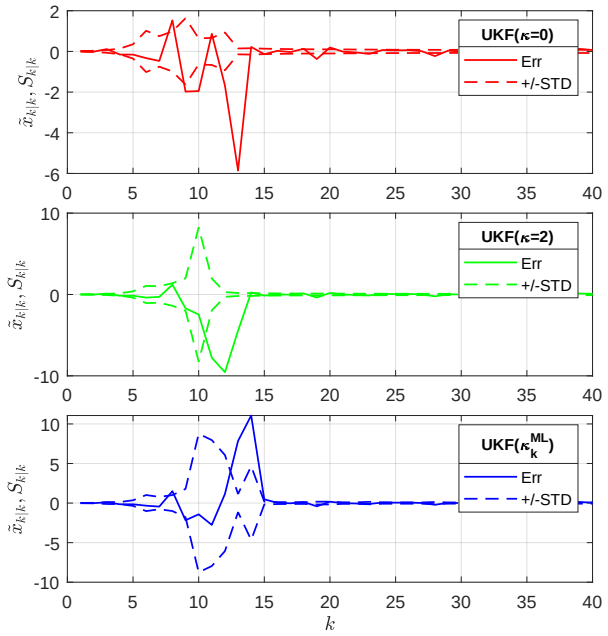

Figure 1. UKF performance depending on the scaling parameter setting.

scaling parameter selection techniques utilize *just* one (in some sense best) value of the scaling parameter at one time instant (either selected prior or adaptively during the the filter run). Thus, the filters are able to exploit a portion of the available information from the measurement. Moreover, as the adaptive strategies, such as in (21), are based on numerical optimization, there are many evaluations of the optimization criterion and the measurement prediction moments available, but only the one selected is used (and the remaining ones are simply disregarded).

D. Goal of the Paper

The goal of this paper is to propose a *cooperative* scaling parameter determination technique for the UKF that combines multiple UKF estimates differing by the scaling parameter choice. Such combination allows to exploit more information from the available measurement at each time instant. The cooperative determination is inspired by the recent techniques for UKF consistency monitoring [18], [19] and a theory of the generalized pseudo-Bayesian estimation [20], and, thus, benefits from a well-developed theoretical background.

IV. Cooperative Scaling Parameter Determination

The UKF with cooperative scaling parameter determination is based on the idea of combination of a set of UKF estimates, where each UKF is configured to perform the *same* estimation task, but uses a different scaling parameter. As a consequence, each filter provide an *unique*, but inherently *sub-optimal*, estimate that extracts just a portion of information from the conditioning data. Analysis and combination of such estimates result in such a scaling parameter selection, which may result in a higher quality state estimate in terms of accuracy and consistency.

A. Concept Illustration

The concept of the cooperative scaling parameter determination is introduced using the UKF (9)–(15), (16)–(20) with the initial condition

$$p(\mathbf{x}_0|\mathbf{z}^0) = \mathcal{N}\{\mathbf{x}_0; \hat{\mathbf{x}}_{0|0}, \mathbf{P}_{0|0}\} \tag{26}$$

and a set of allowed scaling parameters $\mathscr{K} = \{\kappa^{(i)}\}_{i=1}^{n_\kappa}$.

The *prediction* (time-update) step of the UKF with cooperative scaling parameter determination (CUKF) is based on evaluation of n_κ UKF prediction steps, each with different scaling parameter $\kappa_{k+1|k}^{(i)} \in \mathscr{K}$ used in evaluation of (9), (10). This leads to a set of predictive state estimates, in the form of the filtering mean and covariance matrix defined as

$$\mathcal{P}_{k+1|k} \triangleq \left\{\mathcal{N}_{k+1|k}^{(i)}\right\}_{i=1}^{n_\kappa} , \tag{27}$$

where the notation $\mathcal{N}_{k+1|k}^{(i)} = \mathcal{N}\{\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{k+1|k}^{(i)}, \mathbf{P}_{k+1|k}^{xx,(i)}\}$ is used for convenience and the predictive moments are

$$\hat{\mathbf{x}}_{k+1|k}^{(i)} = \hat{\mathbf{x}}_{k+1|k}(\kappa_{k+1|k}^{(i)}) , \tag{28}$$

$$\mathbf{P}_{k+1|k}^{xx,(i)} = \mathbf{P}_{k+1|k}^{xx}(\kappa_{k+1|k}^{(i)}) . \tag{29}$$

The *filtering* (measurement update) step of the CUKF based on evaluation of n_κ^2 UKF filtering steps, each with different predictive PDF (27) and scaling parameter $\kappa_{k+1|k+1}^{(j)} \in \mathscr{K}$ used in (11)–(15), leads to the set of the filtering estimates

$$\mathcal{P}_{k+1|k+1} \triangleq \left\{\mathcal{N}_{k+1|k+1}^{(i,j)}\right\}_{i=1,j=1}^{n_\kappa,n_\kappa} , \tag{30}$$

where the filtering moments

$$\hat{\mathbf{x}}_{k+1|k+1}^{(i,j)} = \hat{\mathbf{x}}_{k+1|k+1}(\kappa_{k+1|k}^{(i)}, \kappa_{k+1|k+1}^{(j)}) , \tag{31}$$

$$\mathbf{P}_{k+1|k+1}^{xx,(i,j)} = \mathbf{P}_{k+1|k+1}^{xx}(\kappa_{k+1|k}^{(i)}, \kappa_{k+1|k+1}^{(j)}) \tag{32}$$

depends on *both* scaling parameters $\kappa_{k+1|k}^{(i)}$, $\kappa_{k+1|k+1}^{(j)}$.

Standard approaches to scaling parameter selection evaluate all the predictive and filtering state estimates (27), (30), and select just one estimate $\mathcal{N}_{k+1|k+1}^{(i,j)}$ according to some criterion, e.g., according to (21) [6]. It means, that all remaining estimates and the contained information is not exploited and a significant part of the calculations is wasted.

B. Formulation in Generalized Pseudo-Bayesian Estimator Design Framework

The proposed CUKF aims at utilization of *all* information the various UKFs estimates can capture due to different scaling parameter selection. The main question is how to fuse the estimates together to *exploit* maximum of available information being captured in (30)–(32). The chosen approach is based on a reformulation of the considered problem in the multiple model (MM) approach, namely using

the approximate *first order generalized pseudo-Bayesian* (GPB1) estimator design framework [20].

Because of the Bayesian approach to the GF design, the set of the UKFs can be, broadly speaking, thought of as a set of *KF* each designed for differently *linearized*[6] model $\mathcal{M}$ given by (1), (2), further denoted as

$$\mathcal{L} \triangleq \{\mathcal{L}^{(i,j)}(\kappa^{(i)}_{k+1|k}, \kappa^{(j)}_{k+1|k+1})\}^{n_\kappa,n_\kappa}_{i=1,j=1} \ , \qquad (33)$$

where the linearized model $\mathcal{L}^{(i,j)} = \mathcal{L}^{(i,j)}(\kappa^{(i)}_{k+1|k}, \kappa^{(j)}_{k+1|k+1})$ depends on the nonlinear model $\mathcal{M}$ and scaling parameters $\kappa^{(i)}_{k+1|k}$, $\kappa^{(j)}_{k+1|k+1}$ used by the UT in prediction and filtering step, respectively. Such interpretation of the set of the UKFs allows to create a set of n_κ^2 hypotheses $\mathcal{H}^{(i,j)}, \forall i,j$, that the considered linearized model $\mathcal{L}^{(i,j)}$ results in the approximate Gaussian posterior PDF $\mathcal{N}^{(i,j)}_{k+1|k+1}$, which is (in some sense) closest to the true (but unknown) posterior $p(\mathbf{x}_k|\mathbf{z}^k)$. The PDF $\mathcal{N}^{(i,j)}_{k+1|k+1}$ is, therefore, conditioned by the model $\mathcal{L}^{(i,j)}$, thus, by $\kappa^{(i)}_{k+1|k}$, $\kappa^{(j)}_{k+1|k+1}$.

Having the set of models, it is necessary to determine their weights. Following the GPB1, the predictive weights can be equal as we do not have any prior knowledge, i.e.,

$$\alpha^{(i)}_{k+1|k} = 1/n_\kappa \ . \qquad (34)$$

Then, the weight of the model $\mathcal{L}^{(i,j)}$ and of the filtering estimate $\mathcal{N}^{(i,j)}_{k+1|k+1}$, w.r.t. the last measurement $\mathbf{z}_{k+1}$, reads

$$\alpha^{(i,j)}_{k+1|k+1} = \alpha^{(i)}_{k+1|k}\mathcal{N}\{\mathbf{z}_{k+1}; \hat{\mathbf{z}}^{(i,j)}_{k+1|k}, \mathbf{P}^{zz,(i,j)}_{k+1|k}\}/c_k, \quad (35)$$

where $c_k = \sum_{i=1}^{n_\kappa}\sum_{j=1}^{n_\kappa} \alpha^{(i)}_{k|k-1}\mathcal{N}\{\mathbf{z}_k; \hat{\mathbf{z}}^{(i,j)}_{k|k-1}, \mathbf{P}^{zz,(i,j)}_{k|k-1}\}$ is the normalization and the measurement predictive moments are computed within the filtering step.

Having the filtering PDFs (30) and the weights (35), the combined state estimate, which take into account UKFs with all admissible combinations of $\kappa^{(i)}_{k+1|k}$, $\kappa^{(j)}_{k+1|k+1}$, is [20]

$$p^{\mathrm{C}}(\mathbf{x}_{k+1}|\mathbf{z}^{k+1}) = \sum_{i=1}^{n_\kappa}\sum_{j=1}^{n_\kappa} \alpha^{(i,j)}_{k+1|k+1}\mathcal{N}^{(i,j)}_{k+1|k+1} \ , \qquad (36)$$

with mean

$$\hat{\mathbf{x}}^{\mathrm{C}}_{k+1|k+1} = \sum_{i=1}^{n_\kappa}\sum_{j=1}^{n_\kappa} \alpha^{(i,j)}_{k+1|k+1}\hat{\mathbf{x}}^{(i,j)}_{k+1|k+1} \qquad (37)$$

and covariance matrix

$$\mathbf{P}^{xx,\mathrm{C}}_{k+1|k+1} = \sum_{i=1}^{n_\kappa}\sum_{j=1}^{n_\kappa} \alpha^{(i,j)}_{k+1|k+1} \qquad (38)$$
$$\times \left(\mathbf{P}^{xx,(i,j)}_{k+1|k+1} + (\hat{\mathbf{x}}^{(i,j)}_{k+1|k+1} - \hat{\mathbf{x}}^{\mathrm{C}}_{k+1|k+1})(\cdot)^T\right).$$

To keep the computational complexity constant $\forall k$, it is necessary to approximate the combined PDF (36) with the Gaussian PDF, such as

$$p^{\mathrm{C},\mathcal{N}}(\mathbf{x}_{k+1}|\mathbf{z}^{k+1}) = \mathcal{N}\{\mathbf{x}_{k+1}; \hat{\mathbf{x}}^{\mathrm{C}}_{k+1|k+1}, \mathbf{P}^{xx,\mathrm{C}}_{k+1|k+1}\} \ , \qquad (39)$$

which is then used as an initial PDF for the CUKF prediction (27). Note that the Gaussian approximation (39) is optimal according to the Kullback-Leibler distance.

C. CUKF Algorithm Illustration

The CUKF algorithm is illustrated by the diagram in Figure 2.

D. Notes and Extensions

1) User Defined Parameter and Computational Complexity: The only input required from the user, is related to the specification of the allowed scaling parameter set $\mathscr{K}$. The cardinality of the set is driven by the available computational power. In particular, the CUKF evaluates n_κ prediction and n_κ^2 filtering steps of the UKF plus n_κ^2 likelihood functions.

Regarding the particular values of the set and assuming an ordered set $\mathscr{K}$, i.e., $\kappa^{(1)} < \kappa^{(2)} < \ldots < \kappa^{(n_\kappa)}$, the lower bound can be $\kappa_0 = 0$, as for this value the covariance matrix of the transformed random variable is guaranteed to be positive definite. The upper bound $\kappa^{(n_\kappa)}$ can be determined so that the random variable $\mathbf{x}_k$ lies with a required probability P^* within a hyper-ellipsoid determined by the σ-points. The upper bound, them, satisfies the condition [6], [21]

$$P^* = \frac{2^{1-n_x/2}}{\Gamma(n_x/2)} \int_0^{\sqrt{n_x+\kappa^{(n_\kappa)}}} e^{-r^2/2} r^{n_x-1} dr \ , \qquad (40)$$

where $\Gamma(\cdot)$ is the Gamma function. The closed-form solution to (40) for $\kappa^{(n_\kappa)}$ exists for some state dimensions n_x only, for other cases a numerical solution can be used instead.

2) Comparison with Standard Competitive Adaptations and Criteria: The standard on-line "competitive" scaling parameter adaptation techniques for the UKF (or derivative-free filters in general) [6], [8], [11] select the scaling parameter to minimize/maximize some criterion, e.g., (21), which is only *loosely* (or indirectly) related to the state estimate accuracy or consistency. This means, that the selected scaling parameter need not necessarily result in the most accurate estimate. On the other hand, the multi-model based interpretation used in the CUKF design allows to interpret the resulting estimates in the Bayesian context [20]. The resulting estimate is, thus, a weighted mixture of all partial estimates based on all allowed scaling parameters.

3) Information Extraction: Each UKF can be viewed as a KF designed for a linearized model, i.e., as a sub-optimal filter with some (unique) gain parametrized by the scaling parameter. As such, each filter is able to extract a portion of information about the desired state from the measurement. By the combination of the sub-optimal estimates, thus, it could be possible to extract much more information about the true state. An analysis and comparison of in literature available combinations of the two GF estimates based on different linearization can be found in [22].

[6]The UT can be interpreted as a statistical (linear) regression [2], [16].

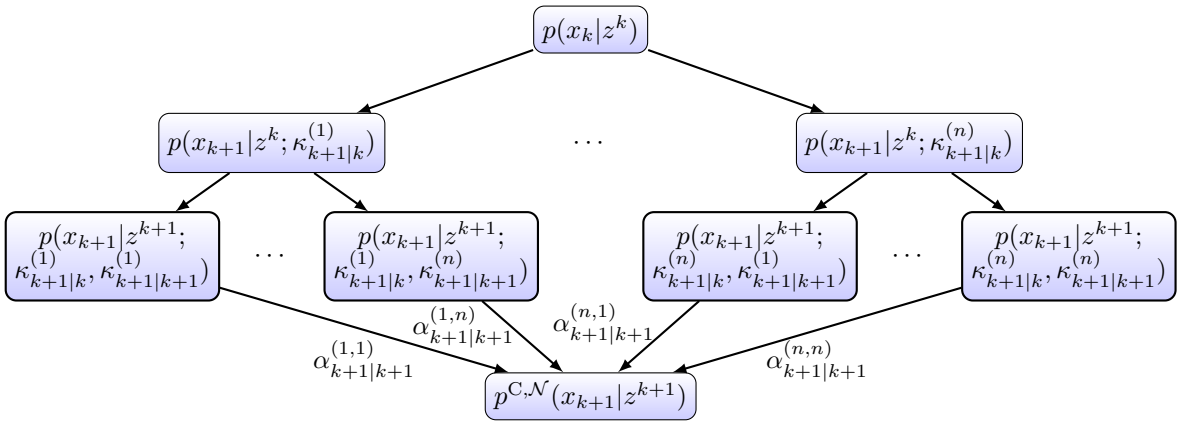

Figure 2. CUKF information flow.

4) Adaptation in Smoothing: As the CUKF is based on the GPB1 framework, the cooperative scaling parameter adaptation concept can be straightforwardly extended to smoothing.

5) Cooperative Approach in GF Performance Monitoring: In [18], [19], consistency (or integrity) of the filter output was monitored by statistical testing of estimate properties of the set of filters configured for the same estimation task, but differing in the scaling parameter. In these tests, the particular local filters were interpreted also as the set of KFs designed for a set of differently linearized models and, roughly saying, the filter output was considered to be consistent if all KFs provided "similar" results. If not, the user was notified and such estimates were treated with care.

6) Extensions: The CUKF should be still understood as an initial concept with some open questions remaining. For example, utilization of other estimate fusion techniques such as the second-order generalized pseudo-Bayesian estimator design should be treated. Also, the focus should be laid on the selection of the initial (possibly non-equal) weights (34).

V. Numerical Illustration

Let the system definition from the motivational example (22), (23) be recalled and the following evaluation criteria for the $M = 10^3$ Monte-Carlo simulations be defined

- Root mean square error (RMSE) of the filtering estimate

$$\mathrm{RMSE}_k = \sqrt{\tfrac{1}{M}\sum_{m=1}^{M}(\tilde{x}_{k|k,m})^2}\ , \tag{41}$$

- Averaged standard deviation (ASTD) of the filtering estimate

$$\mathrm{ASTD}_k = \sqrt{\tfrac{1}{M}\sum_{m=1}^{M}P_{k|k,m}}\ , \tag{42}$$

where $\tilde{x}_{k|k,m} = x_{k,m} - \hat{x}_{k|k,m}$ is the estimate error at m-th MC simulation with $x_{k,m}$ being the true state, $\hat{x}_{k|k,m}$ its estimate, and $P_{k|k,m}$ the respective variance at m-th MC simulation.

Table I
RMSE of UKFs estimate.

	UKFfix(0)	UKFfix(2)	UKFML	**CUKF**
$\mathrm{RMSE}(\mathscr{K}_{0:4})$	12.4	6.2	5.2	2.8
$\mathrm{RMSE}(\mathscr{K}_{0:20})$			3.4	1.9

Four implementations of the UKF are considered, namely

- UKF with fixed $\kappa = 0$, denoted as *UKFfix(0)*, (algorithm also known as the cubature KF [13]),
- UKF with fixed $\kappa = 2$, denoted as *UKFfix(2)*, (according to [3]),
- UKF with adaptive κ_k computed according to (21), denoted as *UKFML*, [6],
- *proposed* UKF with the cooperative scaling parameter adaptation, denoted as *CUKF*.

The UKFs with adaptation are designed for two sets $\mathscr{K}_{0:4} = \{0, 1, 2, 3, 4\}$ and $\mathscr{K}_{0:20} = \{0, 1, \ldots, 20\}$.

The results in the form of the averaged RMSE over all time instants

$$\mathrm{RMSE} = \tfrac{1}{T+1}\sum_{k=0}^{T}\mathrm{RMSE}_k\ , \tag{43}$$

are given in Table I, where the UKFs with adaptive selection of κ are evaluated for both sets $\mathscr{K}_{0:4}$ and $\mathscr{K}_{0:20}$. It can be seen, that the impact of the scaling parameter selection is significant and the proposed cooperative CUKF *outperforms* not only the UKFs with fixed parameters but also the standard competitive UKFML. Also, the UKFs with scaling parameter adaptation improve their accuracy with increasing cardinality of $\mathscr{K}$.

The RMSE assesses the quality of the estimated mean. The GF, however, provides also the conditional variance (or the standard deviation), which should be consistent with the state estimate error. To compare consistency of the state estimates, the criteria (41), (42) are plotted in Figure 3 (the UKFs with adaptive settings are plotted for $\mathscr{K}_{0:20}$). It can

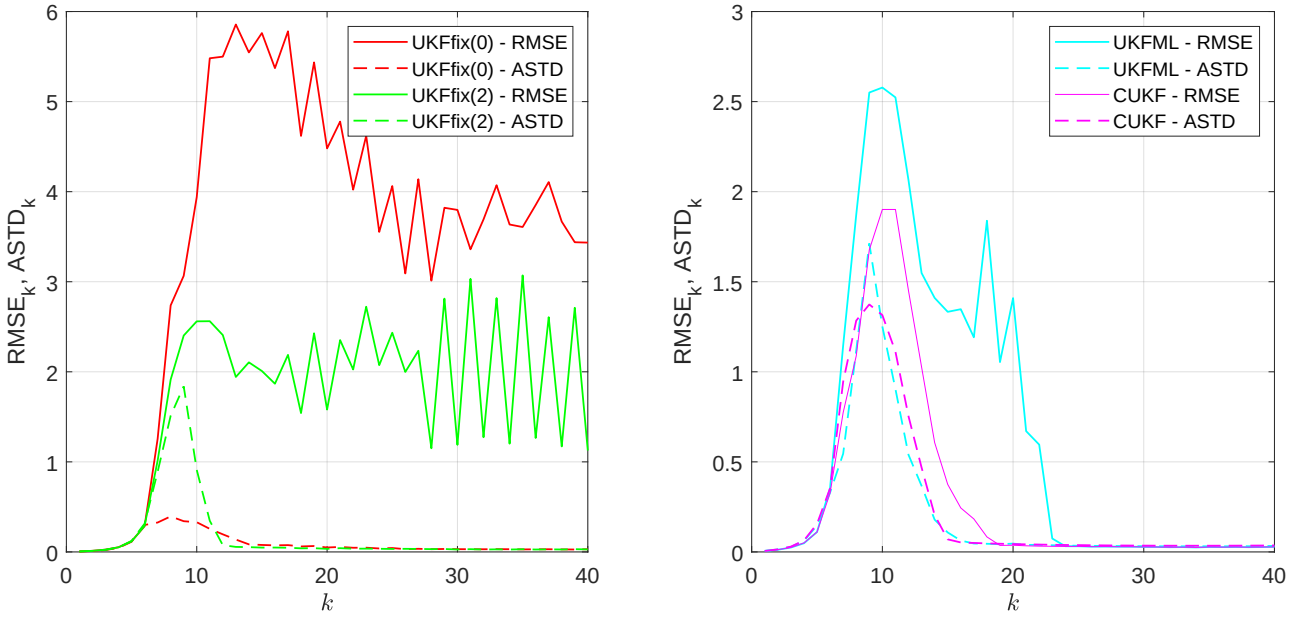

Figure 3. RMSE and ASTD time behavior of UKFs estimates.

be seen that the proposed CUKF provides not only the most accurate but also almost consistent estimates. On the other hand, the UKFs with fixed scaling parameters provide very optimistic estimates. This may be a problem especially in safety-critical scenarios [18].

VI. Concluding Remarks

The paper dealt with the state estimation of the nonlinear stochastic dynamic systems by the Gaussian unscented Kalman filter. The stress was laid on the setting of the scaling parameter, which is typically left on the user although it significantly affects the estimation performance. Contrary to the standard scaling parameter selection strategies choosing one particular parameter value according to a criterion, the proposed *cooperative unscented Kalman filter* fuses a set of estimates provided by the set of the unscented Kalman filters, which are all configured for the same estimation task, but with *different* scaling parameter value. The estimate fusion is based in the well-developed multiple-model approach, which requires minimal user interaction and offers interpretation of the results in the Bayesian context. The improved performance of the cooperative unscented Kalman filter in terms of accuracy and consistency was illustrated in a numerical example.

Acknowledgment

This work was supported by the Czech Science Foundation (GACR) under grant GA 20-06054J (J. Duník, O. Straka) and by the German Research Foundation (DFG) under grant HA 3789/20-1 (U. D. Hanebeck).

References

[1] H. W. Sorenson, "On the development of practical nonlinear filters," *Information Sciences*, vol. 7, pp. 230–270, 1974.

[2] M. Šimandl and J. Duník, "Derivative-free estimation methods: New results and performance analysis," *Automatica*, vol. 45, no. 7, pp. 1749–1757, 2009.

[3] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *IEEE Proceedings*, vol. 92, no. 3, pp. 401–421, 2004.

[4] M. Nørgaard, N. K. Poulsen, and O. Ravn, "New developments in state estimation for nonlinear systems," *Automatica*, vol. 36, no. 11, pp. 1627–1638, 2000.

[5] J. Steinbring and U. Hanebeck, "S2KF: The smart sampling Kalman filter," in *Proceedings of the 16th International Conference on Inference Fusion*, Istanbul, Turkey, July 2013.

[6] O. Straka, J. Duník, and M. Šimandl, "Unscented Kalman filter with advanced adaptation of scaling parameter," *Automatica*, vol. 50, no. 10, pp. 2657–2664, 2014.

[7] J. Steinbring and U. D. Hanebeck, "LRKF revisited: The smart sampling Kalman filter (S^2KF)," *Journal of Advances in Information Fusion*, vol. 9, no. 2, pp. 106–123, 2014.

[8] A. Sakai and Y. Kuroda, "Discriminatively trained unscented Kalman filter for mobile robot localization," *Journal of Advanced Research in Mechanical Engineering*, vol. 1, no. 3, pp. 153–161, 2010.

[9] R. Turner and C. E. Rasmussen, "Model based learning of sigma points in unscented Kalman filter," *Neurocomputing*, vol. 80, pp. 47–53, 2012.

[10] J. Duník, M. Šimandl, and O. Straka, "Unscented Kalman filter: Aspects and adaptive setting of scaling parameter," *IEEE Transactions on Automatic Control*, vol. 57, no. 9, pp. 2411–2416, 2012.

[11] S. L.A. and da Cruz J.J., "Adaptively tuning the scaling parameter of the unscented Kalman filter," in *Proceedings of the 11th Portuguese Conference on Automatic Control*, Porto, Portugal, Jul. 2014.

[12] Y. Nie and T. Zhang, "Scaling parameters selection principle for the scaled unscented Kalman filter," *Journal of Systems Engineering and Electronics*, vol. 29, no. 3, pp. 601–610, 2018.

[13] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," *IEEE Trans. on Automatic Control*, vol. 54, no. 6, pp. 1254–1269, 2009.

[14] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House, 2004.

[15] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Prentice Hall, New Jersey, 1979.

[16] T. Lefebvre, H. Bruyninckx, and J. De Schutter, "Comment on "A new method for the nonlinear transformation of means and covariances in filters and estimators"," *IEEE Transactions on Automatic Control*, vol. 47, no. 8, pp. 1406–1409, 2002.

[17] J. Duník, O. Straka, and M. Šimandl, "Stochastic integration filter," *IEEE Transactions on Automatic Control*, vol. 58, no. 6, pp. 1561–1566, 2013.
[18] J. Duník and O. Straka, "State estimate consistency monitoring in Gaussian filtering framework," *Signal Processing*, no. 148, pp. 145–156, 2018.
[19] J. Duník, O. Straka, and E. Blasch, "Solution separation unscented kalman filter," in *Proceedings of the 22nd International Conference on Information Fusion*, Ottawa, Canada, Jul. 2019.
[20] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. John Wiley & Sons, 2001.
[21] O. Straka, J. Dunik, and M. Simandl, "Unscented Kalman filter with controlled adaptation," in *Proceedings of 16th IFAC Symposium on System Identification (SYSID)*, vol. 16, no. 1, 2012, pp. 906–911.
[22] J. Duník, O. Straka, J. Ajgl, and E. Blasch, "From competitive to cooperative filter design," in *Proceedings of the 20th International Conference on Information Fusion*, Xi'an, China, Jul. 2017.