Kalman Filters for nonlinear systems: a comparison of performance.

Tine Lefebvre, Herman Bruyninckx and Joris De Schutter
Internal Report 01R033, October 2001,

Katholieke Universiteit Leuven,
Department of Mechanical Engineering,
Division of Production Engineering, Machine Design & Automation (PMA),
Celestijnenlaan 300B, B-3001 Leuven (Heverlee), Belgium.

Abstract

The Kalman Filter is a well-known recursive state estimator for linear systems. In practice the algorithm is often used for nonlinear systems by linearizing the system's process and measurement functions. Different Kalman Filter variants linearize the functions in different ways. This report explains how the best known Kalman Filter variants — i.e. the Extended Kalman Filter (EKF), Iterated Extended Kalman Filter (IEKF), the Central Difference Filter (CDF), the first order Divided Difference Filter (DD1) and the Unscented Kalman Filter (UKF)— (i) linearize the process and measurement functions; (ii) take the linearization errors into account; and (iii) how the quality of the state estimates depends on the previous two choices. This report uses two criteria to express this quality: i.e., the consistency and the information content of the estimates. Papers on nonlinear Kalman Filtering usually focus on the consistency of the estimates (the convergence of the filters) without considering their information content.

Besides some well known results, the report gives also insight in the following less known but important issues:

1. the CDF, DD1 and UKF are described in a unified way as Linear Regression Kalman Filters (LRKF). A LRKF (i) linearizes the process and measurement functions by a statistical linear regression of the functions through some regression points in the "uncertainty region" around the state estimate; and (ii) defines the uncertainty due to linearization errors on a linearized function as the covariance matrix of the deviations between the function values of the nonlinear and the linearized function in the regression points;

2. the LRKF measurement update often yields consistent, but non-informative state estimates;

3. the IEKF measurement update outperforms the EKF and LRKF updates for fully observed systems.

4. It can be interesting to use different filters for the process and measurement update. Bearing this in mind, the report compares the filter performances separately for the process and measurement update step, instead of analyzing their overall performance.

1 Problem formulation

During the last decades, many research areas (robotics, economic forecasting, etc.) looked into the matter of how a system can deal *on-line* with uncertainty in an "optimal" way. *Optimality* is hereby often translated into a minimum mean square error (MMSE) state estimation problem: given $\boldsymbol{x}_k$ (the real state of the system at time step k), the optimal state estimate $\hat{\boldsymbol{x}}_{k|i}$ minimizes the expectation

$$E[(\boldsymbol{x}_k - \boldsymbol{x}_{k|i})^T(\boldsymbol{x}_k - \boldsymbol{x}_{k|i})]. \tag{1}$$

The subscript i indicates that —besides prior information about $\boldsymbol{x}_k$— all measurement data up to time step i (i.e. $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_i\}$), are considered when making this best state estimate. The processing of these measurements relies on a stochastic *measurement function* $\boldsymbol{z}_k = \boldsymbol{h}_k(\boldsymbol{x}_k, .)$ linking the (noisy) measurements to the real state value. Eq. (1) is minimized by the expected value of the conditional probability distribution $p(\boldsymbol{x}_{k|i})$ of the state, given the measurements, [1]:

$$\hat{\boldsymbol{x}}_{k|i} = E[\boldsymbol{x}_{k|i}]. \tag{2}$$

In the remainder of the report, the "optimal" state estimate indicates the state estimate given by Eq. (2). Of course, calculating an "optimal" state estimate is not very useful if the system has no idea how much it can rely on it. Therefore, also the *uncertainty* of the state estimate $\hat{\boldsymbol{x}}_{k|i}$ is calculated. This uncertainty can be represented by a probability distribution $p(\boldsymbol{x}_{k|i})$ ("global estimation methods") or by some other representative value, as for instance the covariance matrix of the state estimate ("local estimation methods")

$$\boldsymbol{P}_{k|i} = E[(\boldsymbol{x}_{k|i} - \hat{\boldsymbol{x}}_{k|i})(\boldsymbol{x}_{k|i} - \hat{\boldsymbol{x}}_{k|i})^T]. \tag{3}$$

Note that this description is restricted to *stochastic* filters. Some deterministic filters also calculate the uncertainty on the state estimate, an example is interval filtering representing the estimate's uncertainty by an interval of feasible estimates.

Calculating the optimal state estimate (and its uncertainty) by a batch processing of all past measurements is a problem with increasing computational load. Hence, due to time constraints, *online* estimation requires a *recursive* formulation which needs at each time step only a fixed part of the past measurements. A widely considered class of systems are systems that can be modeled by *Markov models*. A Markov model is a model in which the state estimate and its uncertainty contain all necessary information to make future state estimates. In this case, a recursive implementation of the estimation problem only considers the new measurement, the last state estimate and their uncertainties. Recursive estimation algorithms can also easily be extended to the estimation of *dynamical* systems. In this case, the stochastic *process function* $\boldsymbol{x}_k = \boldsymbol{f}_{k-1}(\boldsymbol{x}_{k-1}, .)$, describing the dynamics of the system (subject to process uncertainty) is assumed to be known.

If the estimation problem has a *linear process function and a linear measurement function*; i.e., if the equations $\boldsymbol{x}_k = \boldsymbol{f}_{k-1}(\boldsymbol{x}_{k-1}, .)$ and $\boldsymbol{z}_k = \boldsymbol{h}_k(\boldsymbol{x}_k, .)$ reduce to $\boldsymbol{x}_k = \boldsymbol{F}_{k-1}\boldsymbol{x}_{k-1} + .$ and $\boldsymbol{z}_k = \boldsymbol{H}_k\boldsymbol{x}_k + .$; and if the process and measurement uncertainties are represented by their mean and covariance matrices, then the optimal state estimate and its covariance matrix are described by the recursive Kalman Filter algorithm [2, 1]. The optimal solution to a recursive estimation problem with *a nonlinear process function and/or a nonlinear measurement function* requires that the *complete* description of the probability distribution $p(\boldsymbol{x}_{k|i})$ is computed at each time step. For some systems this probability distribution can be described with a fixed number of parameters (see [3]

for continuous filtering), in general however, the number of necessary parameters grows boundless as a function of time. Hence, on-line implementations require *approximations*, e.g.:

1. filters that use the exact nonlinear functions and *approximate the probability distribution* by samples (e.g., Monte Carlo and grid-based Markov filters, [4, 5]) or by parameterization with a fixed number of parameters (e.g., the approximative Daum filters for continuous filtering, [3]);

2. filters that *approximate the nonlinear functions* and use known filter updates for these kind of approximations; e.g., the Kalman Filter update equations are used with *linearized* process and measurement functions to obtain state estimates and their covariance matrices [6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17];

3. filters that *approximate both the distribution and the functions*; e.g., higher order filters, [7, 8, 9] and Gaussian sum filters, [18, 19].

Kalman Filters are very popular because they are easy to use, have low computational cost and low memory requirements. Unfortunately, when applying them to systems with nonlinear process and/or measurement functions, their estimates $\hat{\boldsymbol{x}}_{k|i}$ and covariance matrices $\boldsymbol{P}_{k|i}$ can be far from the optimal values as given by equations (1)-(3). The quality of the estimates from the different Kalman Filter variants strongly depends on how these filters linearize the process and measurement functions and how they take linearization errors into account. This report evaluates the quality of an estimate with two contradictory criteria:

1. the *consistency* of the state estimate.
   A state estimate $\hat{\boldsymbol{x}}_{k|i}$ with covariance matrix $\boldsymbol{P}_{k|i}$ is called consistent if

   $$E[(\boldsymbol{x}_{k|i} - \hat{\boldsymbol{x}}_{k|i})(\boldsymbol{x}_{k|i} - \hat{\boldsymbol{x}}_{k|i})^T] \leq \boldsymbol{P}_{k|i}. \tag{4}$$

   The most encountered problem with the Kalman Filter variants is that the calculated state estimate $\hat{\boldsymbol{x}}_{k|i}$ and covariance matrix $\boldsymbol{P}_{k|i}$ become inconsistent ("divergence" of the filter). In this case, $\boldsymbol{P}_{k|i}$ is too small and does no longer represent a reliable measure for the uncertainty on $\hat{\boldsymbol{x}}_{k|i}$. Even more, once an inconsistent state estimate is encountered, the subsequent state estimates are also inconsistent. This is because the filter believes the inconsistent state estimate to be more accurate than it is in reality and hence, it attaches too much weight to this state estimate when processing new measurements.

2. its *information content*.
   The covariance matrix $\boldsymbol{P}_{k|i}$ indicates how uncertain the state estimate $\hat{\boldsymbol{x}}_{k|i}$ is: a large covariance matrix indicates an inaccurate (and little useful) state estimate; the smaller the covariance matrix, the larger the information content of the state estimate.

There is a trade-off between consistent and informative state estimates: inconsistency can be avoided by making $\boldsymbol{P}_{k|i}$ artificially larger. Making $\boldsymbol{P}_{k|i}$ too large (larger than necessary for consistency) corresponds to loosing information about the real accuracy of the state estimate.

This report focuses on the following variants of the Kalman Filter: the Extended Kalman Filter (EKF, [6, 7, 8, 9]), the Iterated Extended Kalman Filter (IEKF, [6, 7, 8, 9]), the Central Difference Filter (CDF, [10]), the first order Divided Difference Filter

(DD1, [11, 12]) and the Unscented Kalman Filter (UKF, [13, 14, 15, 16, 17]). The CDF, DD1 and UKF are described in a unified way as Linear Regression Kalman Filters (LRKF). The following sections explain (i) how the Kalman Filter variants linearize the process and measurement functions; (ii) how they take the linearization errors into account; and (iii) how the quality of the state estimates depends on these two choices[1]. Besides well known issues, such as the inconsistency of the EKF state estimates and consistency of the UKF state estimates, this report clarifies the following less known but important facts:

1. the CDF, DD1 and UKF are LRKFs (Sections 3.2, 4.3), meaning that (i) they linearize process and measurement functions by statistical linear regression of the functions through some regression points in the "uncertainty region" around the state estimate; and (ii) they define the uncertainty due to linearization errors on a linearized process or measurement function as the covariance matrix of the deviations between the function values of the nonlinear and the linearized function in the regression points.

   Originally, the CDF, DD1 and UKF were not derived this way, but the insight is important to understand their drawbacks and possible choices of the regression points in those cases where the original formulation does not assure good results (e.g., when dealing with discontinuous functions);

2. for large uncertainties on the state estimate, the LRKF measurement update yields consistent but non-informative state estimates (Section 4.3);

3. the IEKF measurement update returns very good (consistent and informative) state estimates when the state —or at least the part of it that causes the nonlinearity in the measurement function— is instantaneously fully observable (Section 4.2);

4. a good performance of a filter in the process or measurement update, does not guarantee a good performance in the other update. Therefore, it can be interesting to use different filters for both updates. Hence, this report compares the filter performances separately for process updates (Section 3) and measurement updates (Section 4) instead of their overall performance when they are both combined.

First of all, Section 2 describes the (linear) Kalman Filter in order to introduce notation.

2 The Kalman Filter

Consider a linear dynamic system whose dynamics are described by

$$\boldsymbol{x}_k = \boldsymbol{F}_{k-1}\boldsymbol{x}_{k-1} + \boldsymbol{b}_{k-1} + \boldsymbol{C}_{k-1}\boldsymbol{w}_{k-1}. \tag{5}$$

$\boldsymbol{F}_{k-1}$, $\boldsymbol{b}_{k-1}$ and $\boldsymbol{C}_{k-1}$ are supposed to be exactly known; $\boldsymbol{b}_{k-1}$ can be a function of the input $\boldsymbol{u}_{k-1}$ to the system; $\boldsymbol{w}_{k-1}$ is the process uncertainty and is a random vector sequence with zero mean and a known covariance matrix $\boldsymbol{Q}_{k-1}$.

The measurement $\boldsymbol{z}_k$ has a linear relationship to the state:

$$\boldsymbol{z}_k = \boldsymbol{H}_k\boldsymbol{x}_k + \boldsymbol{d}_k + \boldsymbol{E}_k\boldsymbol{v}_k; \tag{6}$$

[1] The more nonlinear the behavior of the process or measurement function in the "uncertainty region" around the state estimate, the more pronounced the difference in quality performance (consistency and information content of the state estimates) between the Kalman Filter variants.

$\boldsymbol{H}_k$, $\boldsymbol{d}_k$ and $\boldsymbol{E}_k$ are supposed to be known; $\boldsymbol{d}_k$ can be a function of the input $\boldsymbol{u}_{k-1}$ to the system; $\boldsymbol{v}_k$ is the measurement uncertainty and is a random vector sequence with zero mean and a known covariance matrix $\boldsymbol{R}_k$. Furthermore, $\boldsymbol{v}$ and $\boldsymbol{w}$ are mutually uncorrelated and uncorrelated between sampling times (white noise sequences).[2]

The state estimate at time step k, based on the measurements up to time step i, is denoted as $\hat{\boldsymbol{x}}_{k|i}$; its covariance matrix is $\boldsymbol{P}_{k|i}$. $\hat{\boldsymbol{x}}_{k|k-1}$ is called the *predicted* state estimate and $\hat{\boldsymbol{x}}_{k|k}$ the *updated* state estimate. The initial state estimate $\hat{\boldsymbol{x}}_{1|0}$ and its covariance matrix $\boldsymbol{P}_{1|0}$ represent the prior knowledge. For the system described by Eq. (5) and (6), the MMSE state estimates and their covariance matrices are obtained by the Kalman Filter algorithm [6]:

$$\hat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{F}_{k-1}\hat{\boldsymbol{x}}_{k-1|k-1} + \boldsymbol{b}_{k-1}; \tag{7}$$

$$\boldsymbol{P}_{k|k-1} = \boldsymbol{F}_{k-1}\boldsymbol{P}_{k-1|k-1}\boldsymbol{F}_{k-1}^T + \boldsymbol{C}_{k-1}\boldsymbol{Q}_{k-1}\boldsymbol{C}_{k-1}^T; \tag{8}$$

$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k(\boldsymbol{z}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{d}_k); \tag{9}$$

$$\boldsymbol{P}_{k|k} = \boldsymbol{P}_{k|k-1} - \boldsymbol{K}_k\boldsymbol{S}_k\boldsymbol{K}_k^T; \tag{10}$$

where

$$\boldsymbol{K}_k = \boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T\boldsymbol{S}_k^{-1}; \tag{11}$$

$$\boldsymbol{S}_k = \boldsymbol{E}_k\boldsymbol{R}_k\boldsymbol{E}_k^T + \boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T. \tag{12}$$

The process update equations (7)–(8) propagate the updated state estimate $\hat{\boldsymbol{x}}_{k-1|k-1}$ through the dynamics of the system described by Eq. (5) and adapt the covariance matrix accordingly. This results in a predicted state estimate $\hat{\boldsymbol{x}}_{k|k-1}$ and its covariance matrix. The measurement update equations (9)–(12) describe how the updated state estimate $\hat{\boldsymbol{x}}_{k|k}$ and its covariance matrix are calculated from this predicted state estimate, the measurement $\boldsymbol{z}_k$ and their covariance matrices.

3 Nonlinear process functions.

This section compares the *process* updates of the different Kalman Filter variants when dealing with a nonlinear process function[3]

$$\boldsymbol{x}_k = \boldsymbol{f}_{k-1}(\boldsymbol{x}_{k-1}, \boldsymbol{u}_{k-1}) + \boldsymbol{C}_{k-1}\boldsymbol{w}_{k-1}, \tag{13}$$

with linearization

$$\boldsymbol{x}_k = \boldsymbol{F}_{k-1}\boldsymbol{x}_{k-1} + \boldsymbol{b}_{k-1} + \boldsymbol{w}_{k-1}^* + \boldsymbol{C}_{k-1}\boldsymbol{w}_{k-1}. \tag{14}$$

$\boldsymbol{w}_{k-1}^*$ represents the linearization error. The extra uncertainty due to this linearization error is represented by the covariance matrix $\boldsymbol{Q}_{k-1}^*$[4]. Kalman Filter variants differ by

[2]Correlated uncertainties can be dealt with by augmenting the state vector, this is the original formulation of the Kalman Filter [2]. In this new state vector, process and measurement functions are of the form (5) and (6) where the uncertainties are uncorrelated.

[3]It is possible to consider process uncertainty that enters the process function in a nonlinear way. To simplify notation in this report, the state vector is assumed to be augmented with these uncertainty terms.

[4]$\boldsymbol{w}_{k-1}^*$ is correlated to the state estimate and is correlated over time. The Kalman Filter is able to cope with correlated uncertainties by augmenting the state vector, *if* the correlation is linear and

their choice of $\boldsymbol{F}_{k-1}$, $\boldsymbol{b}_{k-1}$ and $\boldsymbol{Q}^*_{k-1}$; they all use process update equations (7) and (8) to update the state estimate and its uncertainty[5].

Section 3.1 evaluates the linearization of the process function by the EKF and IEKF; Section 3.2 by the LRKF. Section 3.3 illustrates. Section 3.4 summarizes.

3.1 (Iterated) Extended Kalman Filter

The Extended Kalman Filter (EKF) and the Iterated Extended Kalman Filter (IEKF) linearize the process function in the same way, i.e. by a first order Taylor series around the *updated state estimate* $\hat{\boldsymbol{x}}_{k-1|k-1}$:

$$\boldsymbol{F}_{k-1} = \left.\frac{\partial \boldsymbol{f}_{k-1}}{\partial \boldsymbol{x}}\right|_{\hat{\boldsymbol{x}}_{k-1|k-1}}; \tag{15}$$

$$\boldsymbol{b}_{k-1} = \boldsymbol{f}_{k-1}(\hat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_{k-1}) - \boldsymbol{F}_{k-1}\hat{\boldsymbol{x}}_{k-1|k-1}. \tag{16}$$

The basic (I)EKF algorithms *do not take the linearization errors into account*:

$$\boldsymbol{Q}^*_{k-1} \equiv \boldsymbol{0}. \tag{17}$$

This leads to *inconsistent* state estimates when these errors can not be neglected.

Some implementations of the (I)EKF *try to take the linearization errors into account* by considering a covariance matrix $\boldsymbol{Q}^*_{k-1} \neq \boldsymbol{0}$ (or by multiplying the calculated covariance matrix $\boldsymbol{P}_{k|k-1}$ by a fading factor larger than 1 [6]). However, the linearization errors depend on the nonlinear behavior of the function $\boldsymbol{f}_{k-1}$ in the *"uncertainty region" around* the state estimate $\hat{\boldsymbol{x}}_{k-1|k-1}$ and only the evaluation of the function and its Jacobian *in* $\hat{\boldsymbol{x}}_{k-1|k-1}$ are available. Only *off-line tuning or on-line parameter learning* (adaptive filtering, [20]) can give good values for $\boldsymbol{Q}^*_{k-1}$; i.e., values that result in consistent and informative state estimates (for a particular problem !). In almost all cases, the used added $\boldsymbol{Q}^*_{k-1}$ terms are chosen too large (e.g. constant in time, compensating for decreasing linearization errors); this leads to less informative estimates.

3.2 Linear Regression Kalman Filter

The Linear Regression Kalman Filter (LRKF) uses the function values of *some regression points* $\boldsymbol{\mathcal{X}}^i_{k-1|k-1}$ *in state space* to model the behavior of the process function in the "uncertainty region" around the updated state estimate $\hat{\boldsymbol{x}}_{k-1|k-1}$. The regression points are chosen such that their mean and covariance matrix equal the state estimate $\hat{\boldsymbol{x}}_{k-1|k-1}$ and its covariance matrix $\boldsymbol{P}_{k-1|k-1}$:

$$\hat{\boldsymbol{x}}_{k-1|k-1} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{\mathcal{X}}^i_{k-1|k-1}; \tag{18}$$

$$\boldsymbol{P}_{k-1|k-1} = \frac{1}{n}\sum_{i=1}^{n} \left\{(\boldsymbol{\mathcal{X}}^i_{k-1|k-1} - \hat{\boldsymbol{x}}_{k-1|k-1})(\boldsymbol{\mathcal{X}}^i_{k-1|k-1} - \hat{\boldsymbol{x}}_{k-1|k-1})^T\right\}. \tag{19}$$

---

known. For this new state vector, a process function in the form of Eq. (14) can be written. A remaining unmodeled correlation between $\boldsymbol{w}^*_{k-1}$ and the state estimate results in a $\boldsymbol{Q}^*_{k-1}$ that is larger than $E[\boldsymbol{w}^*_{k-1}\boldsymbol{w}^{*T}_{k-1}]$; a remaining correlation over time results in *biased* state estimates; i.e. the mean of the state estimates does not correspond to the real value. This is not a problem, as long as the state estimates are *consistent* (i.e. $\boldsymbol{Q}^*_{k-1}$ also includes the uncertainty due to this bias).

[5]Note that $\boldsymbol{C}_{k-1}\boldsymbol{w}_{k-1}$ from Eq. (5) corresponds to $\boldsymbol{w}^*_{k-1} + \boldsymbol{C}_{k-1}\boldsymbol{w}_{k-1}$ of Eq. (14); hence instead of $\boldsymbol{C}_{k-1}\boldsymbol{Q}_{k-1}\boldsymbol{C}^T_{k-1}$ in Eq. (8), $\boldsymbol{Q}^*_{k-1} + \boldsymbol{C}_{k-1}\boldsymbol{Q}_{k-1}\boldsymbol{C}^T_{k-1}$ is used.

The function values of these points are $\boldsymbol{\mathcal{X}}^i_{k|k-1} = \boldsymbol{f}_{k-1}(\boldsymbol{\mathcal{X}}^i_{k-1|k-1}, \boldsymbol{u}_{k-1})$.

The LRKF linearizes the process function by statistical linear regression through the $(\boldsymbol{\mathcal{X}}^i_{k-1|k-1}, \boldsymbol{\mathcal{X}}^i_{k|k-1})$ points; i.e., the deviations $\boldsymbol{e}_i$ between the function values in $\boldsymbol{\mathcal{X}}^i_{k-1|k-1}$ for the nonlinear and the linearized function are minimized in least squares sense:

$$\min_{\boldsymbol{F}_{k-1}, \boldsymbol{b}_{k-1}} \sum_i \left\{ \boldsymbol{e}_i^T \boldsymbol{e}_i \right\}; \tag{20}$$

with

$$\boldsymbol{e}_i = \boldsymbol{\mathcal{X}}^i_{k|k-1} - (\boldsymbol{F}_{k-1} \boldsymbol{\mathcal{X}}^i_{k-1|k-1} + \boldsymbol{b}_{k-1}). \tag{21}$$

The covariance of the deviations $\boldsymbol{e}_i$

$$\boldsymbol{Q}^*_{k-1} = E[\boldsymbol{e}_i \boldsymbol{e}_i^T] \tag{22}$$

gives an idea of the magnitude of the linearization errors in the "uncertainty region" around $\hat{\boldsymbol{x}}_{k-1|k-1}$. [21] describes the analytical update equations and shows that they are equivalent to the UKF equations, using the Unscented Transform to transform $\hat{\boldsymbol{x}}_{k-1|k-1}$ and $\boldsymbol{P}_{k-1|k-1}$ into $\hat{\boldsymbol{x}}_{k|k-1}$ and $\boldsymbol{P}_{k|k-1}$ [13, 14, 15, 16, 17].

Of course, the state estimates of the LRKF can only be consistent and informative if the $(\boldsymbol{\mathcal{X}}^i_{k-1|k-1}, \boldsymbol{\mathcal{X}}^i_{k|k-1})$ points are representative for the behavior of the process function in the "uncertainty region" around $\hat{\boldsymbol{x}}_{k-1|k-1}$. A discussion on different choices of the $\boldsymbol{\mathcal{X}}^i_{k-1|k-1}$ points is out of the scope of this report. [10, 11, 12, 13, 14, 15, 16, 17] propose choices of the regression points which assure that —for *continuous* process functions— the calculated covariance matrix $\boldsymbol{P}_{k|k-1}$ is at least as accurate as the one calculated by the (I)EKF:

- the original UKF formalism [13, 14, 15, 16, 17] chooses $2n + 2\kappa$ regression points, where n is the dimension of the state space and κ is a degree of freedom in the choice of the regression points.
- the CDF and DD1 filters [10, 11, 12] choose $2n$ regression points.
- an adaptation of the original UKF algorithm (called the Reduced Sigma Point Filters, [22, 15]) chooses only $n + 1$ regression points to reduce the computational costs. This means that the linear regression through these points is exact and hence $\boldsymbol{Q}^*_{k-1} \equiv \boldsymbol{0}$: this filter does not take the linearization errors into account. Hence, as was the case for the (I)EKF, consistent estimates can only be obtained by adding a $\boldsymbol{Q}^*_{k-1}$ obtained by off-line tuning or on-line parameter learning.

In general —when enough regression points are taken— the state estimates of the LRKF process update are *consistent* and *informative*. Consistent because $\boldsymbol{Q}^*_{k-1}$ gives a well founded approximation of the linearization errors (Eq. (22)); informative because the linearized function is a good approximation of the process function in the "uncertainty region" around $\hat{\boldsymbol{x}}_{k-1|k-1}$ (Eqs. (20)–(21)).

3.3 Illustrations

The different process updates are illustrated by a simple 2D nonlinear process function ($\boldsymbol{x}(i)$ denotes the ith element of $\boldsymbol{x}$):

$$\begin{cases} \boldsymbol{x}_k(1) &= (\boldsymbol{x}_{k-1}(1))^2; \\ \boldsymbol{x}_k(2) &= \boldsymbol{x}_{k-1}(1) + 3\, \boldsymbol{x}_{k-1}(2); \end{cases} \tag{23}$$

with no process uncertainty: $\boldsymbol{w}_{k-1} \equiv \mathbf{0}$. The updated state estimate and its uncertainty at time step $k-1$ are:

$$\hat{\boldsymbol{x}}_{k-1|k-1} = \begin{bmatrix} 10 \\ 15 \end{bmatrix}, \qquad \boldsymbol{P}_{k-1|k-1} = \begin{bmatrix} 36 & 0 \\ 0 & 3600 \end{bmatrix}. \tag{24}$$

The (in)consistency and information content of the state estimates is illustrated graphically by comparing their uncertainty ellipses[6] with the uncertainty ellipse obtained by Monte Carlo simulations. When the former encloses the latter, the state estimate is *consistent*. A good filter calculates consistent *and* informative state estimates, i.e., the uncertainty ellipse based on the state estimate and covariance matrix encloses the Monte Carlo uncertainty ellipse and is not much larger.

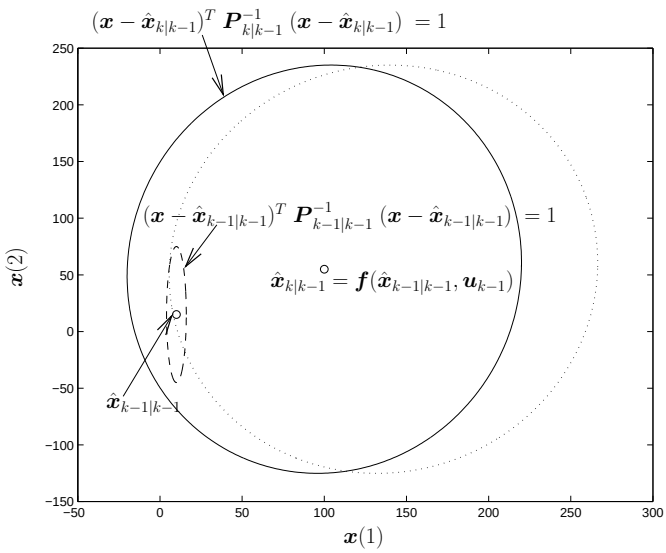

Figure 1: Nonlinear process model. Uncertainty ellipses for the updated state estimate at $k-1$ (dashed line), for the (I)EKF predicted state estimate (full line) and the Monte Carlo uncertainty ellipse (dotted line). The predicted state estimate is inconsistent (its uncertainty ellipse does not enclose the dotted line) due to the neglected linearization errors.

(I)EKF Figure 1 shows the updated and (I)EKF predicted state estimates and their uncertainty ellipses. The dotted line is the uncertainty ellipse of the distribution obtained by Monte Carlo simulation. The IEKF's state prediction and its covariance matrix are:

$$\hat{\boldsymbol{x}}_{k|k-1} = \begin{bmatrix} 100 \\ 55 \end{bmatrix}, \qquad \boldsymbol{P}_{k|k-1} = \begin{bmatrix} 14400 & 720 \\ 720 & 32436 \end{bmatrix}. \tag{26}$$

Due to the neglected linearization errors, the state estimate is *inconsistent*: the uncertainty ellipse of the predicted state estimate does not enclose the Monte Carlo uncertainty ellipse.

[6]The uncertainty ellipsoid

$$(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|i})^T \boldsymbol{P}_{k|i}^{-1} (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|i}) = 1 \tag{25}$$

is a graphical representation of the uncertainty on the state estimate $\hat{\boldsymbol{x}}_{k|i}$. Starting from the point $\hat{\boldsymbol{x}}_{k|i}$, the distance to the ellipse in each direction is a measure for the uncertainty on $\hat{\boldsymbol{x}}_{k|i}$ in that direction.

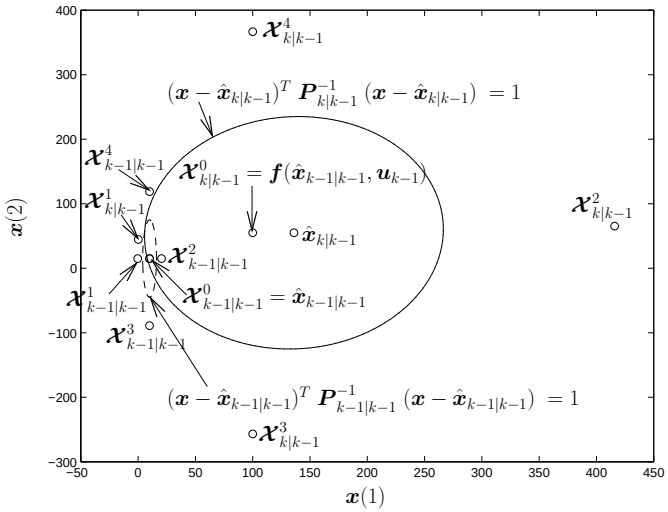

Figure 2: Nonlinear process model. Uncertainty ellipses for the updated state estimate at $k-1$ (dashed line), for the UKF predicted state estimate (full line), and Monte Carlo uncertainty ellipse (dotted line, coincides with the full line). The predicted state estimate is consistent (its uncertainty ellipse encloses the dotted line) and informative (its uncertainty ellipse is close to the dotted line).

LRKF Figure 2 shows the $\boldsymbol{\mathcal{X}}^i_{k-1|k-1}$ and $\boldsymbol{\mathcal{X}}^i_{k|k-1}$ points, the updated and predicted state estimate and their uncertainty ellipses for the UKF. The $\boldsymbol{\mathcal{X}}^i_{k-1|k-1}$ points are chosen with the algorithm of [13, 14, 15, 16, 17] where $\kappa = 1$. The uncertainty ellipse obtained by Monte Carlo simulation coincides with the final uncertainty ellipse on the scale of the figure. The predicted state estimate and its covariance matrix are

$$\hat{\boldsymbol{x}}_{k|k-1} = \begin{bmatrix} 136 \\ 55 \end{bmatrix}, \qquad \boldsymbol{P}_{k|k-1} = \begin{bmatrix} 16992 & 720 \\ 720 & 32436 \end{bmatrix}. \tag{27}$$

The uncertainty ellipse of the predicted state estimate (i) encloses the Monte Carlo uncertainty ellipse, indicating *consistency*; and (ii) is close to the Monte Carlo uncertainty ellipse, indicating an *informative* state estimate.

3.4 Conclusion: the process update

The LRKF performs better than the (I)EKF when dealing with nonlinear process functions:

1. the LRKF linearizes the function based on its behavior in the "uncertainty region" *around* the updated state estimate. The (I)EKF on the other hand only uses the function evaluation and its Jacobian *in* this state estimate.

2. the LRKF deals with linearization errors in a theoretically founded way (provided that enough regression points are chosen). The (I)EKF on the other hand needs trial and error for each particular example to obtain good values for the covariance matrix which models the linearization errors.

3. unlike the (I)EKF, the LRKF does not need the function Jacobian. This is an advantage where this Jacobian is difficult to obtain or non-existing (e.g. for a discontinuous function).

4 Nonlinear measurement functions.

The previous section compared the (I)EKF and LRKF *process* updates; this section compares their *measurement* updates for a nonlinear measurement function[7]

$$\boldsymbol{z}_k = \boldsymbol{h}_k(\boldsymbol{x}_k, \boldsymbol{u}_k) + \boldsymbol{E}_k \boldsymbol{v}_k; \tag{28}$$

with linearization

$$\boldsymbol{z}_k = \boldsymbol{H}_k \boldsymbol{x}_k + \boldsymbol{d}_k + \boldsymbol{v}_k^* + \boldsymbol{E}_k \boldsymbol{v}_k. \tag{29}$$

$\boldsymbol{v}_k^*$ represents the linearization errors. The additional uncertainty on the linearized measurement function due to these linearization errors is $\boldsymbol{R}_k^*$.[8] The EKF, IEKF and LRKF choose $\boldsymbol{H}_k$, $\boldsymbol{d}_k$ and $\boldsymbol{R}_k^*$ in a different way; afterwards, they use the Kalman Filter update equations[9] (9)–(12).

Although the filters use similar linearization techniques for the linearization of process and measurement functions, there can be a substantial difference in their performance for both: this is due to the fact that in the *process* update the state estimate and its uncertainty are the only available information to linearize the nonlinear function; but in the *measurement* update also the new measurement is available. The IEKF (Section 4.2) takes this measurement into account; on the other hand the EKF (Section 4.1) and LRKF (Section 4.3) linearize the measurement function based only on the predicted state estimate and its uncertainty. For the latter filters, the linearization errors ($\boldsymbol{R}_k^*$) are larger, especially when the measurement function is quite nonlinear in the uncertainty region around the predicted state estimate. The Kalman Filter considers $\boldsymbol{R}_k^* + \boldsymbol{E}_k \boldsymbol{R}_k \boldsymbol{E}_k^T$ to be the uncertainty on the measurement, hence, a large uncertainty on the *state* estimate results in throwing away the greater part of the information of the (possibly very accurate) *measurement*.

Section 4.4 illustrates the different measurement updates. Section 4.5 summarizes.

4.1 Extended Kalman Filter

The Extended Kalman Filter (EKF) linearizes the measurement function around the predicted state estimate $\hat{\boldsymbol{x}}_{k|k-1}$:

$$\boldsymbol{H}_k = \left.\frac{\partial \boldsymbol{h}_k}{\partial \boldsymbol{x}}\right|_{\hat{\boldsymbol{x}}_{k|k-1}}; \tag{30}$$

$$\boldsymbol{d}_k = \boldsymbol{h}_k(\hat{\boldsymbol{x}}_{k|k-1}) - \boldsymbol{H}_k \hat{\boldsymbol{x}}_{k|k-1}. \tag{31}$$

The *basic* EKF algorithm does not take the linearization errors into account:

$$\boldsymbol{R}_k^* \equiv \boldsymbol{0}. \tag{32}$$

If the function is nonlinear in the "uncertainty region" around the predicted state estimate, the linearization errors are not negligible. This results in a true state value "far from"[10] the linearized measurement function and leads to an *inconsistent* updated state estimate.

[7] It is possible to consider measurement uncertainty that enters the measurement function in a nonlinear way. To simplify notation in this report, the state vector is assumed to be augmented with these uncertainty terms.

[8] The same remarks as for $\boldsymbol{w}_{k-1}^*$ and $\boldsymbol{Q}_{k-1}^*$ in Section 3 (footnote 4) can be made here for $\boldsymbol{v}_k^*$ and $\boldsymbol{R}_k^*$.

[9] Note that $\boldsymbol{E}_k \boldsymbol{v}_k$ from Eq. (6) corresponds to $\boldsymbol{v}_k^* + \boldsymbol{E}_k \boldsymbol{v}_k$ of Eq. (29); hence instead of $\boldsymbol{E}_k \boldsymbol{R}_k \boldsymbol{E}_k^T$ in Eq. (12), $\boldsymbol{R}_k^* + \boldsymbol{E}_k \boldsymbol{R}_k \boldsymbol{E}_k^T$ is used here.

[10] "Far from" (and "close to") must be understood as: the deviation of the true state with respect to the linearized measurement function is not justified (is justified) by the measurement uncertainty $\boldsymbol{R}_k^* + \boldsymbol{E}_k \boldsymbol{R}_k \boldsymbol{E}_k^T$.

In order to make the state estimates consistent, many EKF implementations take $\boldsymbol{R}_k^* \neq \mathbf{0}$. As was the case for the linearization of process functions by the (I)EKF (Section 3.1), no information is available about the magnitude of the linearization errors. Only *off-line tuning or on-line parameter learning* can lead to a good value for $\boldsymbol{R}_k^*$ for a particular problem. In almost all cases, the used added $\boldsymbol{R}_k^*$ terms are chosen too large (e.g. constant in time, compensating for decreasing linearization errors); this leads to less informative estimates.

4.2 Iterated Extended Kalman Filter

The EKF of the previous section, linearizes the measurement function around the *predicted* state estimate. The Iterated Extended Kalman Filter (IEKF) tries to do better by linearizing the measurement function around the *updated* state estimate:

$$\boldsymbol{H}_k = \left.\frac{\partial \boldsymbol{h}_k}{\partial \boldsymbol{x}}\right|_{\hat{\boldsymbol{x}}_{k|k}}; \tag{33}$$

$$\boldsymbol{d}_k = \boldsymbol{h}_k(\hat{\boldsymbol{x}}_{k|k}) - \boldsymbol{H}_k \hat{\boldsymbol{x}}_{k|k}. \tag{34}$$

This is achieved by iteration: the filter first linearizes the function around the predicted state estimate $\hat{\boldsymbol{x}}_{k|k-1}$ and calculates the updated state estimate (like the EKF). Then the filter linearizes the function around this new state estimate $\hat{\boldsymbol{x}}_{k|k}^1$ and calculates (starting from $\hat{\boldsymbol{x}}_{k|k-1}$, $\boldsymbol{P}_{k|k-1}$ and the new linearized measurement function) a new updated state estimate. This process is iterated till a state estimate $\hat{\boldsymbol{x}}_{k|k}^i$ is found for which $\hat{\boldsymbol{x}}_{k|k}^i$ is very close to $\hat{\boldsymbol{x}}_{k|k}^{i-1}$ (or till a maximum number of iterations i has been executed). The state estimate $\hat{\boldsymbol{x}}_{k|k}$ and uncertainty $\boldsymbol{P}_{k|k}$ are calculated starting from the state estimate $\hat{\boldsymbol{x}}_{k|k-1}$ with its uncertainty $\boldsymbol{P}_{k|k-1}$ and the measurement function linearized around $\hat{\boldsymbol{x}}_{k|k}^i$. Like the EKF algorithm, the *basic* IEKF algorithm does not take the linearization errors into account:

$$\boldsymbol{R}_k^* \equiv \mathbf{0}. \tag{35}$$

If the function is nonlinear in the "uncertainty region" around the updated state estimate $\hat{\boldsymbol{x}}_{k|k}$, state estimates will be *inconsistent* (analogous to Section 4.1). In case of a measurement model that instantaneously *fully observes* the state (or at least the part of the state that causes the nonlinearities in the measurement function), the linearization errors will be small in the "uncertainty region" around $\hat{\boldsymbol{x}}_{k|k}$. The true state estimate is then "close to"[10] the linearized measurement function and the updated state estimate is *consistent*.

In order to obtain consistent state estimates for systems that do not fully observe the state, a *non-zero* $\boldsymbol{R}_k^*$ can be used. Again, this value has to be tuned off-line or learned on-line and is often chosen too large.

4.3 Linear Regression Kalman Filter

The LRKF evaluates the measurement function in some regression points $\boldsymbol{\mathcal{X}}_{k|k-1}^i$ in the "uncertainty region" around the predicted state estimate $\hat{\boldsymbol{x}}_{k|k-1}$. The $\boldsymbol{\mathcal{X}}_{k|k-1}^i$ are chosen such that their mean and covariance matrix are equal to the predicted state estimate

and its covariance:

$$\hat{\boldsymbol{x}}_{k|k-1} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{\mathcal{X}}^i_{k|k-1}; \tag{36}$$

$$\boldsymbol{P}_{k|k-1} = \frac{1}{n}\sum_{i=1}^{n} \left\{(\boldsymbol{\mathcal{X}}^i_{k|k-1} - \hat{\boldsymbol{x}}_{k|k-1})(\boldsymbol{\mathcal{X}}^i_{k|k-1} - \hat{\boldsymbol{x}}_{k|k-1})^T\right\}. \tag{37}$$

The function values of the regression points through the nonlinear function are $\boldsymbol{\mathcal{Z}}^i_k = \boldsymbol{h}_k(\boldsymbol{\mathcal{X}}^i_{k|k-1}, \boldsymbol{u}_k)$.

The LRKF algorithm uses a linearized measurement function obtained by statistical linear regression through the ($\boldsymbol{\mathcal{X}}^i_{k|k-1}$, $\boldsymbol{\mathcal{Z}}^i_k$) points. The statistical linear regression is such that the deviations $\boldsymbol{e}_i$ between the nonlinear and the linearized function in the regression points are minimized in least squares sense:

$$\boldsymbol{e}_i = \boldsymbol{\mathcal{Z}}^i_k - (\boldsymbol{H}_k\boldsymbol{\mathcal{X}}^i_{k|k-1} + \boldsymbol{d}_k); \tag{38}$$

$$\min_{\boldsymbol{H}_k,\boldsymbol{d}_k} \sum_i \left\{\boldsymbol{e}_i^T\boldsymbol{e}_i\right\}. \tag{39}$$

The covariance matrix of the deviations $\boldsymbol{e}_i$ give an idea of the magnitude of the linearization errors:

$$\boldsymbol{R}^*_k = E[\boldsymbol{e}_i\boldsymbol{e}_i^T]. \tag{40}$$

[21] describes the analytical measurement update equations and shows that they are equivalent to the UKF equations, using an Unscented Transform to calculate the predicted measurement $\hat{\boldsymbol{z}}_k = \boldsymbol{h}_k(\hat{\boldsymbol{x}}_{k|k-1}, \boldsymbol{u}_k)$ and its covariance matrix based on the predicted state estimate $\hat{\boldsymbol{x}}_{k|k-1}$ with covariance matrix $\boldsymbol{P}_{k|k-1}$ [13, 14, 15, 16, 17].

The state estimates can only be consistent if the ($\boldsymbol{\mathcal{X}}^i_{k|k-1}$, $\boldsymbol{\mathcal{Z}}^i_k$) points are representative for the behavior of the measurement function in the "uncertainty region" around $\hat{\boldsymbol{x}}_{k|k-1}$. A discussion on possible choices of regression points is out of the scope of this report. [10, 11, 12, 13, 14, 15, 16, 17] analyze choices of the $\boldsymbol{\mathcal{X}}^i_{k|k-1}$ points for which —in case of a *continuous* measurement function— the calculated $\boldsymbol{P}_{\boldsymbol{\mathcal{Z}}_k\boldsymbol{\mathcal{Z}}_k}$ (the covariance matrix of $\boldsymbol{\mathcal{Z}}^i_k$) is at least as accurate as the one calculated by the EKF ($\boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T$):

- the original UKF formalism [13, 14, 15, 16, 17] chooses $2n + 2\kappa$ regression points, where n is the dimension of the state space and κ is a degree of freedom in the choice of the regression points.
- the CDF and DD1 filters [10, 11, 12] choose $2n$ regression points.
- an adaptation of the original UKF algorithm (called the Reduced Sigma Point Filters, [22, 15]) reduces the computational cost by choosing only $n + 1$ regression points. This means that the linear regression through these points is exact and hence $\boldsymbol{R}^*_k \equiv \boldsymbol{0}$: this filter does not take the linearization errors into account and inconsistent estimates are obtained. As was also the case for the (I)EKF, consistent estimates can only be obtained by adding a $\boldsymbol{R}^*_{k-1}$ obtained by trial and error.

In general —when enough regression points are taken— the state estimates of the LRKF measurement update are *consistent* because $\boldsymbol{R}^*_k$ gives a well founded approximation of the linearization errors (Eq. (40)). If the measurement function is highly nonlinear in the "uncertainty region" around $\hat{\boldsymbol{x}}_{k|k-1}$, then the ($\boldsymbol{\mathcal{X}}^i_{k|k-1}$, $\boldsymbol{\mathcal{Z}}^i_k$) points deviate substantially from a hyperplane. This results in a large $\boldsymbol{R}^*_k$ and *non-informative* state estimates (see also the illustration in the next section).

4.4 Illustrations

The comparison between the different measurement updates is illustrated with the measurement function $z_k = h_1(\boldsymbol{x}_k) + v_1$;

$$h_1(\boldsymbol{x}_k) = (\boldsymbol{x}_k(1))^2 + (\boldsymbol{x}_k(2))^2 \,. \tag{41}$$

$\boldsymbol{x}_k = [15\ 20]^T$ is the real value; $\hat{\boldsymbol{x}}_{k|k-1} = [10\ 15]^T$ is the predicted state estimate with covariance matrix $\boldsymbol{P}_{k|k-1} = \begin{bmatrix} 36 & 0 \\ 0 & 3600 \end{bmatrix}$. The processed measurement is $z_k = 630$ and its covariance is $R_k = 400$.

In all figures, the real state value $\boldsymbol{x}_k$ is plotted; if this value is far outside the uncertainty ellipse, the corresponding state estimate is inconsistent. As the measurement is accurate and the initial state estimate is not, the uncertainty on the state estimate should drop considerably when the measurement is processed. If this is not the case, then the updated state estimate is not informative.

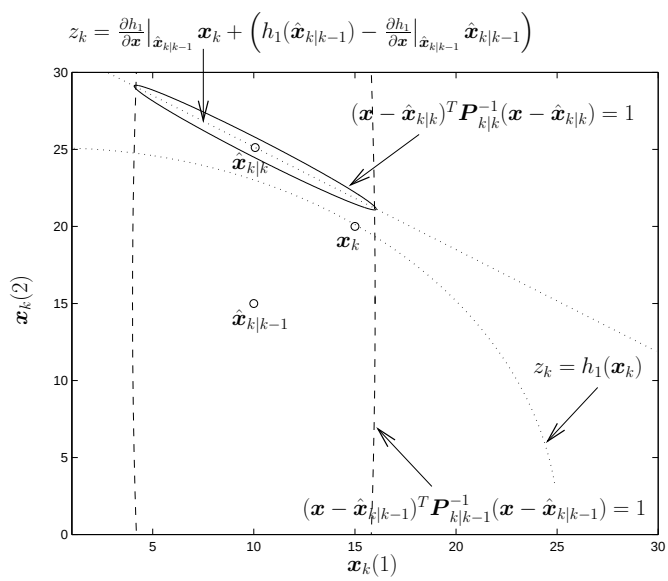

Figure 3: Nonlinear measurement model $z = h_1(\boldsymbol{x}_k)$ and EKF linearization around $\hat{\boldsymbol{x}}_{k|k-1}$ (dotted lines). The real state $\boldsymbol{x}_k$ is "far from" this linearization and the obtained state estimate $\hat{\boldsymbol{x}}_{k|k}$ (uncertainty ellipse in full line) is inconsistent.

EKF Figure 3 shows the state estimates, uncertainty ellipses and measurement functions for the EKF. The real measurement function is nonlinear. $\boldsymbol{x}_k$ is the true value of the state, and is "close to"[10] this function. The linearization around the uncertain predicted state estimate is not a good approximation of the function around the true state value: the true state value is "far from"[10] the linearized measurement function. The updated state estimate

$$\hat{\boldsymbol{x}}_{k|k} = \begin{bmatrix} 10.0675 \\ 25.1204 \end{bmatrix}; \boldsymbol{P}_{k|k} = \begin{bmatrix} 35.8407 & -23.8909 \\ -23.8909 & 16.3697 \end{bmatrix} \tag{42}$$

is *inconsistent*.

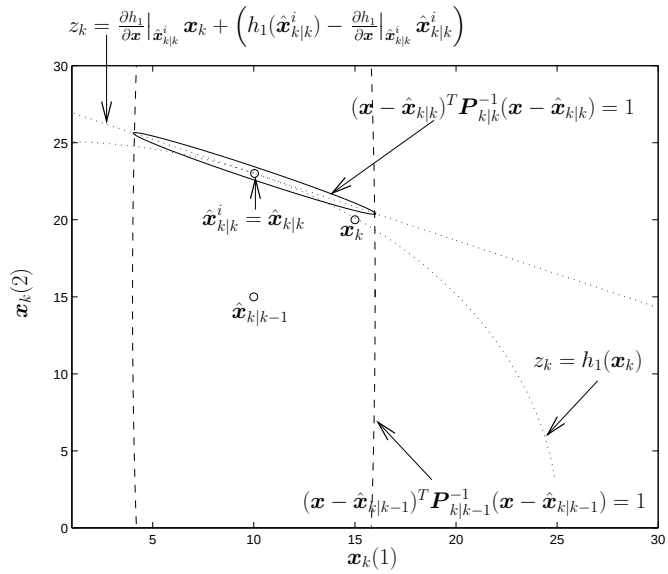

Figure 4: Nonlinear measurement model $z = h_1(\boldsymbol{x}_k)$ that does *not* observe the full state, and its IEKF linearization around $\hat{\boldsymbol{x}}_{k|k}$ (dotted lines). The real state $\boldsymbol{x}_k$ is "far from" this linearization, leading to an inconsistent state estimate $\hat{\boldsymbol{x}}_{k|k}$ (uncertainty ellipse in full line).

IEKF Figure 4 shows the measurement function, the linearized measurement function around the point $\hat{\boldsymbol{x}}^i_{k|k}$, the true state value $\boldsymbol{x}_k$ and the state estimates for the IEKF. The measurement model does not fully observe the state; this results in an uncertain updated state estimate $\hat{\boldsymbol{x}}^i_{k|k}$ around which the filter linearizes the measurement function. As was the case for the EKF, the linearization errors are not negligible and the true value is "far from"[10] the linearized measurement function. The updated state estimate

$$\hat{\boldsymbol{x}}_{k|k} = \begin{bmatrix} 10.0349 \\ 23.0061 \end{bmatrix}; \boldsymbol{P}_{k|k} = \begin{bmatrix} 35.9316 & -15.6720 \\ -15.6720 & 7.0245 \end{bmatrix} \tag{43}$$

is *inconsistent.*

If however the measurement model fully observes the state, the IEKF updated state estimate is accurately known; hence, the linearization errors are small and the true state value is "close to"[10] the linearized measurement function. In this case, the updated state estimate is *consistent.* To illustrate this, a second example is given. The measurement function

$$\boldsymbol{z}_k = \boldsymbol{h}(\boldsymbol{x}_k) + \boldsymbol{v} = \begin{bmatrix} h_1(\boldsymbol{x}_k) + v_1 \\ h_2(\boldsymbol{x}_k) + v_2 \end{bmatrix}; \tag{44}$$

with

$$\begin{cases} h_1(\boldsymbol{x}_k) &= (\boldsymbol{x}_k(1))^2 + (\boldsymbol{x}_k(2))^2; \\ h_2(\boldsymbol{x}_k) &= 3\,(\boldsymbol{x}_k(2))^2 / \boldsymbol{x}_k(1); \end{cases} \tag{45}$$

observes the state completely. The processed measurement and its covariance matrix are:

$$\boldsymbol{z}_k = \begin{bmatrix} 630 \\ 85 \end{bmatrix}; \qquad \boldsymbol{R}_k = \begin{bmatrix} 400 & 0 \\ 0 & 400 \end{bmatrix}. \tag{46}$$

Figure 5 shows the measurement function, the linearized measurement function, the true state value $\boldsymbol{x}_k$, the state estimates and the uncertainty ellipses. The updated state

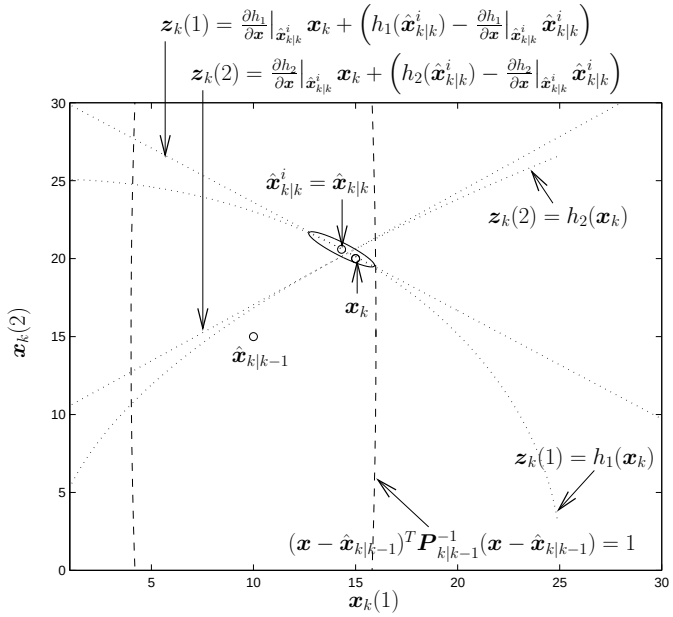

Figure 5: Nonlinear measurement model $\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{x}_k)$ that fully observes the state, and its IEKF linearization around $\boldsymbol{x}^i_{k|k}$ (dotted lines). The real state $\boldsymbol{x}_k$ is "close to" this linearization, leading to a consistent state estimate (uncertainty ellipse in full line).

estimate and covariance matrix

$$\hat{\boldsymbol{x}}_{k|k} = \begin{bmatrix} 14.3164 \\ 20.5963 \end{bmatrix}; \boldsymbol{P}_{k|k} = \begin{bmatrix} 2.5983 & -1.6513 \\ -1.6513 & 1.2753 \end{bmatrix} \tag{47}$$

are *consistent and informative* due to the small, ignored, linearization errors.

LRKF An UKF is run on the original example (Eq. (41)). The $\boldsymbol{\mathcal{X}}^i_{k|k-1}$ points are chosen as in [13, 14, 15, 16] with $\kappa = 1$. Figures 6 and 7 give a view of the nonlinear measurement function, the $\boldsymbol{\mathcal{X}}^i_{k|k-1}$-points and the linearization. The predicted state estimate is uncertain, hence the $\boldsymbol{\mathcal{X}}^i_{k|k-1}$-points are widespread. Due to the large deviations between the ($\boldsymbol{\mathcal{X}}^i_{k|k-1}$, $\boldsymbol{\mathcal{Z}}^i_k$) points and the linearized measurement function (See Figure 7), R^*_k is large ($R^*_k = 2.5664\ 10^7$). The updated state estimate and its covariance matrix are

$$\hat{\boldsymbol{x}}_{k|k} = \begin{bmatrix} 9.9171 \\ 2.5598 \end{bmatrix}; \qquad \boldsymbol{P}_{k|k} = \begin{bmatrix} 36.000 & 0.0000 \\ 0.0000 & 3600.0 \end{bmatrix}. \tag{48}$$

Figure 8 shows the $\boldsymbol{\mathcal{X}}^i_{k|k-1}$ points, the measurement function, the UKF linearized measurement function, the true state value $\boldsymbol{x}_k$, the state estimates and the uncertainty ellipses. The updated state estimate is consistent, but it can hardly be called an improvement over the previous state estimate ($\boldsymbol{P}_{k|k} \approx \boldsymbol{P}_{k|k-1}$)! *The information in the measurement is neglected* due to the very high "measurement" uncertainty $\boldsymbol{R}^*_k + \boldsymbol{E}_k \boldsymbol{R}_k \boldsymbol{E}^T_k$ on the linearized function.

Note that some kind of *Iterated* LRKF (analogous to the Iterated EKF) would not solve this problem: the updated state estimate $\hat{\boldsymbol{x}}_{k|k}$ and its covariance matrix $\boldsymbol{P}_{k|k}$ are more or less the same as the predicted state estimate $\hat{\boldsymbol{x}}_{k|k-1}$ and its covariance matrix $\boldsymbol{P}_{k|k-1}$. Hence, the regression points and the linearization would approximately be the same after iteration.

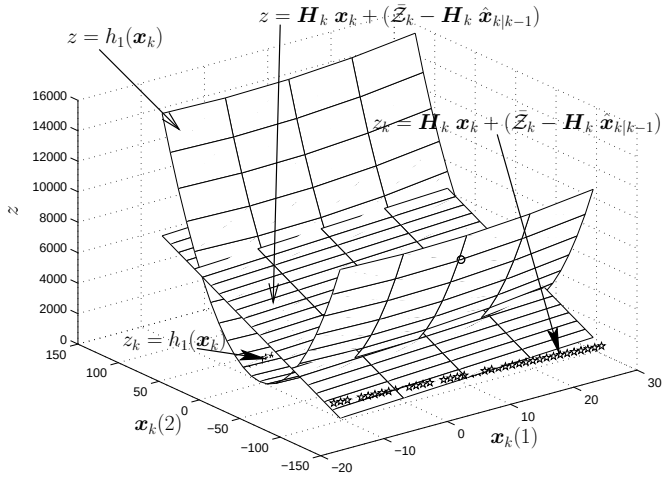

Figure 6: Nonlinear measurement model $z = h_1(\boldsymbol{x})$ and UKF linearization. The linearization errors are large.

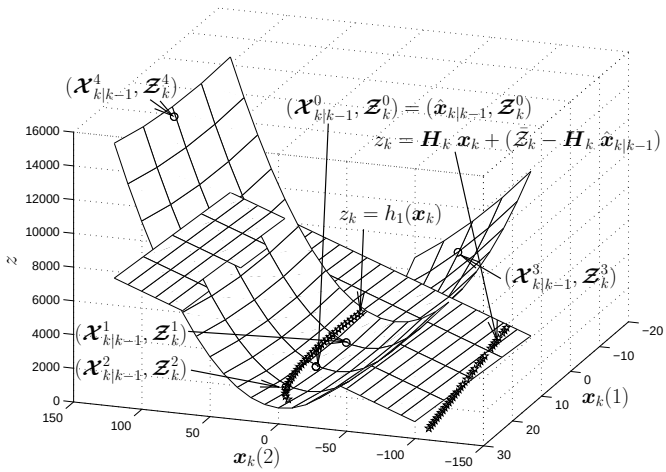

Figure 7: Figure 6 seen from another viewpoint.

4.5 Conclusion: the measurement update

Nonlinear measurement models that fully observe the part of the state that makes the measurement function nonlinear, are best processed by the IEKF. In the other cases, none of the presented filters gives the "best" solution. The filter should be chosen for each specific application: the LRKF makes an estimate of its linearization errors ($\boldsymbol{R}^*_k$), the EKF and IEKF on the other hand require off-line tuning or on-line parameter learning of $\boldsymbol{R}^*_k$ to yield consistent state estimates. Because the IEKF also takes the measurement into account when linearizing the measurement function, its linearization errors are smaller than those of the EKF and LRKF. This means that once a well-tuned IEKF is available, the state estimates it returns can be far more informative than those of the LRKF or (a well-tuned) EKF.

The LRKF does not use the Jacobian of the measurement function, this makes it possible to process discontinuous measurement functions.

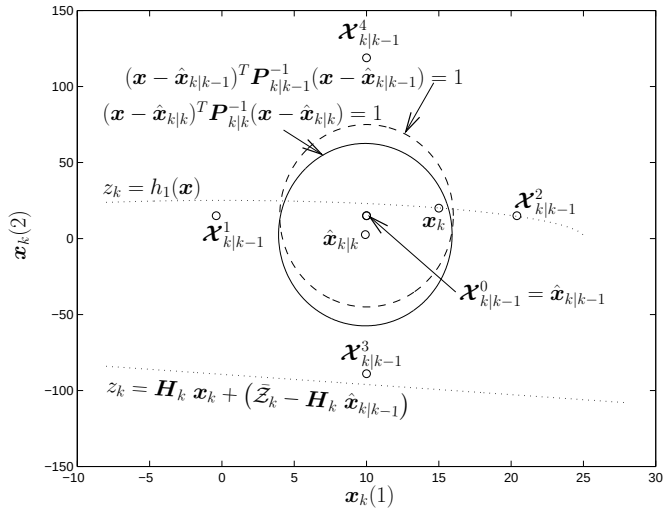

Figure 8: Nonlinear measurement model $z = h_1(\boldsymbol{x})$ and UKF linearization (dotted lines). The large linearization errors make the measurement uncertainty $\boldsymbol{R}^*_k + \boldsymbol{E}_k\boldsymbol{R}_k\boldsymbol{E}^T_k$ on the linearized measurement function very large. The updated state estimate (uncertainty ellipse in full line) is consistent but non-informative.

5 Conclusions

This report evaluates how the common Kalman Filter variants linearize the process and measurement functions and how they deal with linearization errors.

The performance (consistency and information content of the state estimates) of the different filters is compared for the process and measurement updates separately, because a good performance for one of these updates does not necessarily mean a good performance for the other update. This makes it interesting in some cases to use different filters for both updates. The difference in performance for the process and measurement updates is due to the fact that for the linearization of the process function, the state estimate and its uncertainty are the only available information; while for the linearization of the measurement function, also the measurement itself and its uncertainty can be used. The IEKF uses this extra information in contrast to the EKF and LRKF which linearize the measurement function based on only the state estimate. For the latter filters, the linearization errors are larger, especially when the measurement function is quite nonlinear in the uncertainty region around the state estimate. The uncertainty on the state estimate then results in throwing away the greater part of the information of the (possibly very accurate) measurement.

The report describes the CDF, DD1 and UKF in a unified way: the Linear Regression Kalman Filter (LRKF). These filters (i) linearize the process and measurement function by statistical linear regression of the function through some regression points in the "uncertainty region" around the state estimate, and (ii) approximate the linearization errors by the covariance matrix of the deviations between the function values of the nonlinear and the linearized function in these points. A nice advantage of the LRKF is that it does not need the Jacobian of the nonlinear function.

For *process* updates the LRKF performs better than the other mentioned Kalman Filter variants because (i) the LRKF linearizes the process function based on its behavior in the "uncertainty region" *around* the updated state estimate. The (I)EKF on the other hand only uses the function evaluation and its Jacobian *in* this state estimate; and (ii)

the LRKF deals with linearization errors in a theoretically founded way (provided that enough regression points are chosen). The (I)EKF on the other hand needs trial and error for each particular example to obtain good values for the covariance matrix which models the linearization errors.

The IEKF is the best way to process nonlinear *measurement* models that fully observe the part of the state that makes the measurement model nonlinear. In the other cases, none of the presented filters gives the "best" measurement update: the LRKF makes an estimation of the linearization errors, the EKF and IEKF on the other hand require extensive off-line tuning or on-line parameter learning in order to yield consistent state estimates. But, as the IEKF also uses the measurement in order to linearize the measurement function (in contrast to the EKF and LRKF), its linearization errors are smaller and, once a well-tuned IEKF is available, the state estimates it returns can be far more informative than those of the LRKF or (a well-tuned) EKF.

Acknowledgments

T. Lefebvre and H. Bruyninckx are, respectively, Doctoral and Postdoctoral Fellows of the Fund for Scientific Research–Flanders (F.W.O.) in Belgium. Financial support by the Belgian Programme on Inter-University Attraction Poles initiated by the Belgian State—Prime Minister's Office—Science Policy Programme (IUAP), and by K.U.Leuven's Concerted Research Action GOA/99/04 are gratefully acknowledged.

References

[1] Harold W. Sorenson, *Kalman filtering: theory and application*, IEEE Press, New York, NY, 1985.

[2] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME, Journal of Basic Engineering*, vol. 82, pp. 34–45, March 1960.

[3] F. Daum, "New exact nonlinear filters," *J. C. Spall, editor, Bayesian Analysis of Time Series and Dynamic Models*, pp. 199–226, 1988.

[4] D. Fox, W. Burgard, and S. Thrun, "Markov localization for mobile robots in dynamic environments," *Journal of Artificial Intelligence Research*, vol. 11, pp. 391–427, 1999.

[5] D. J.C. MacKay, *Information Theory, Inference and Learning Algorithms*, http://wol.ra.phy.cam.ac.uk/mackay/itprnn/book.html, 2001.

[6] Bar-Shalom and X.R. Li, *Estimation and Tracking: Principles, Techniques and Software*, Artech House, 1993.

[7] H. Tanizaki, *Nonlinear Filters: Estimation and Applications - Second, Revised and Enlarged Edition*, Springer-Verlag, Berlin-Heidelberg, 1996.

[8] A. Gelb, J.F. Kasper, R.A. Nash, C.F. Price, and A.A. Sutherland, *Applied Optimal Estimation*, MIT Press, Cambridge, MA, 1974.

[9] P.S. Maybeck, *Stochastic Models, Estimation, and Control, volume 2*, Academic Press, 1982.

[10] T.S. Schei, "A finite-difference method for linearization in nonlinear estimation algorithms," *Automatica*, vol. 33, no. 11, pp. 2053–2058, November 1997.

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

[12] M. Nørgaard, N.K. Poulsen, and O. Ravn, "Advances in derivative-free state estimation for nonlinear systems," Technical Report IMM-REP-1998-15 (revised edition), Technical University of Denmark, Denmark, April 2000.

[13] S. J. Julier and J. K. Uhlmann, "A general method for approximating nonlinear transformations of probability distributions," *http://www.robots.ox.ac.uk/~siju/work/work.html*, 1996.

[14] S. Julier, J. Uhlmann, and H.F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 477–482, March 2000.

[15] J. Uhlmann, *Dynamic Map Building and Localization: New Theoretical Foundations*, Ph.D. thesis, University of Oxford, UK, 1995.

[16] S. Julier and J.K. Uhlmann, "Data fusion in non-linear systems," in *Handbook of Multisensor Data Fusion*, D.L. Hall and J. Llinas, Eds. CRC Press, 2001.

[17] E.A. Wan and R. van der Merwe, *Kalman Filtering and Neural Networks, chapter 7*, Wiley Publishing, Edited by S. Haykin, (in press) 2001.

[18] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Transactions on Automatic Control*, vol. 45, no. 5, pp. 910–927, May 2000.

[19] D.L. Alspach and H.W. Sorenson, "Nonlinear bayesian estimation using gaussian sum approximations," *IEEE Transactions on Automatic Control*, vol. 17, no. 4, pp. 439–448, August 1972.

[20] R.K. Mehra, "Approaches to adaptive filtering," *IEEE Transactions on Automatic Control*, vol. 17, no. 5, pp. 693–698, October 1972.

[21] 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"," *Accepted for publication in the IEEE Transactions on Automatic Control*, Scheduled for June 2002.

[22] S. Julier, "Reduced sigma point filters for the propagation of means and covariances through nonlinear transformations," *http://citeseer.nj.nec.com/julier98reduced.html*, December 1998.