On Iterative Unscented Kalman Filter using Optimization

Martin Skoglund, Fredrik Gustafsson and Gustaf Hendeby

On Iterative Unscented Kalman Filter using Optimization

Martin A. Skoglund*†, Fredrik Gustafsson*, and Gustaf Hendeby*
*Division of Automatic Control, Linköping University, SE-581 83 Linköping, Sweden
†Eriksholm Research Centre, Rørtangvej 20, DK-3070 Snekkersten, Denmark
Email: {martin.skoglund,fredrik.gustafsson,gustaf.hendeby}@liu.se

***Abstract*—The *unscented Kalman filter* (UKF) is a very popular solution for estimation of the state in nonlinear systems. Similar to the *extended Kalman filter* (EKF) and contrary to the *Kalman filter* (KF) for linear systems, the UKF provides no guarantees that the filter updates will improve the filtered state estimate. In the past, the *iterated EKF* (IEKF) has been suggested as a way to online monitor the filter performance and try to improve it using optimization techniques. In this paper we do the same for the UKF, deriving six *iterated UKF* (IUKF) variations based on two cost functions and three optimization algorithms. The methods are evaluated and compared to IEKF versions and to two versions of the *iterative posterior linearization filter* (IPLF) in three benchmark simulation studies. The results show that IUKF algorithms can be used as a derivative free alternative to IEKF, and provide insights about the different design choices available in IUKF algorithms.**

I. Introduction

Since the 1990's the *unscented Kalman filter* (UKF, [1–4]) has become a very popular and derivative-free alternative to the *extended Kalman filter* (EKF, [5]) for nonlinear state-space problems. While the EKF works by linearizing the state-space model to be able to apply the *Kalman filter* (KF, [6, 7]), the UKF uses the *unscented transform* (UT) to estimate the distributions needed to apply the KF.

A central concept, often lacking in the text book presentations of the EKF, is online performance analysis of the measurement update step to ensure that the update in fact reduces the associated cost function. The cost function can both be used to adapt the step length of the update and to guide an iterative search to minimize the cost and hence to ensure that it is improved in each measurement update. Since the EKF measurement update is essentially a *Gauss-Newton* (GN) method [8] without iterations, efficient iterative optimization routines can be applied to the involved cost function as demonstrated in [9]. This class of filters can be referred to as *iterated extended Kalman filters* (IEKFs). In IEKF, the measurement Jacobian is re-linearized at the current estimate and Kalman-like updates are performed until convergence [8, 10, 11]. There are many other methods correcting for the linearization errors in the standard EKF; [12] includes higher-order moments; and [4] suggests second order derivative compensation. However, the key with IEKFs is not directly to compensate for linearization errors but rather to search for better linearization points using iterative search making it suitable for strong nonlinearities. Thus, IEKF is a natural extension of EKF as a *nonlinear least-squares* (NLS) solution using GN and as such, iterations are typically required for convergence to a local optimum. The IEKFs can result in a lower *root mean square error* (RMSE) and faster response to transients [13] than the EKF in cases when the *signal-to-noise ratio* (SNR) is high.

Given the beneficial properties of performing iterative updates in the EKF, a reasonable assumption is that the same applies for the UKF since the update equations are the same. Previous work on *iterated UKF* (IUKF) include [14] that develops an iterated sigma point filter in the context of long range stereo camera measurements. The presentation assumes a frequentist's or Fisherian point of view of estimation. That is, the objective is to obtain the best possible estimate of the unknown parameter, while minimizing in this case the squared error or variance. Other papers having a similar approach, but not based on GN, are *e.g.*, [15, 16] which update the covariance estimate in each iteration, guides the search minimizing the residual and uses a damping parameter, rather than a standard step-size parameter, to avoid divergence. Contrary to this, resorting to a Bayesian view point, a recent method based on statistical linear regression using sigma points KF was proposed [17, 18], where the objective is to derive an accurate distribution of a stochastic variable. Based on this, a Bayesian *minimum mean square estimator* (MMSE), denoted *iterative posterior linearization filter* (IPLF), is dervied. It can be seen as an iterative procedure minimizing a *Kullback-Leibler divergence* (KLD) to obtain the best posterior distribution approximation. A difference to the IEKFs is that the covariance approximation is iteratively updated and used in the update equations whereas the IEKFs only need to compute the covariance when the GN search has terminated. A further development is the damped version of IPLF [19] which applies GN optimisation to the mean estimate in a inner loop in the same fashion as damped IEKFs.

This paper extends on the work in [9, 14] and develops a number of iterative UKF filters based on different ways to optimize the underlying cost function. The algorithms are also compared to the standard IPLF [18] and a slight modification of IPLF using the state predicted measurements instead of the the sample mean. The paper is organized the following way. In Sec. II a short background to the EKF, UKF, and IEKF is given, providing a framework common to all the filters. Sec. III derives two alternative generic IUKF formulations, which are

Algorithm 1 Kalman Filter

Assume $\hat{x}_{0|0}$ and $P_{0|0}$ given.
for all t **do**
Time update:

$$\hat{x}_{t|t-1} = F\hat{x}_{t-1|t-1} \tag{2a}$$
$$P_{t|t-1} = FP_{t-1|t-1}F^T + GQ_tG^T \tag{2b}$$

Measurement update:

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(y_t - \hat{y}_t) \tag{3a}$$
$$P_{t|t} = (I - K_tH)P_{t|t-1}(I - K_tH)^T + K_tR_tK_t^T \tag{3b}$$
$$\hat{y}_t = H\hat{x}_{t|t-1} \tag{3c}$$
$$P^{xy}_{t|t-1} = P_{t|t-1}H^T \tag{3d}$$
$$P^{yy}_{t|t-1} = (HP_{t|t-1}H^T + R_t)^{-1} \tag{3e}$$
$$K_t = P^{xy}_{t|t-1}(P^{yy}_{t|t-1})^{-1}, \tag{3f}$$

end for

then further elaborated on in terms of different optimization algorithms. The derived IUKF filters are then evaluated using three benchmark examples in Sec. IV. Concluding remarks are given in Sec. V.

II. Filtering Preliminaries

Consider the problem to estimate the state in the following state-space model:

$$x_t = f(x_{t-1}, w_{t-1}) \tag{1a}$$
$$y_t = h(x_t) + e_t, \tag{1b}$$

where x_t is the state at time t, y_t a measurement, and w_t and e_t mutually independent and *Gaussian* white noise; f and h describe the state propagation and measurement function, respectively. It is assumed that $\text{cov}(w_t) = Q_t$ and $\text{cov}(e_t) = R_t$. Depending on the structure of the model, different solutions exist.

A. Kalman Filter

The famous Kalman filter is the *best linear unbiased estimator* (BLUE) when the model (1) is linear, *i.e.*, $f(x_t, w_t) = Fx_t + Gw_t$ and $h(x_t) = Hx_t$ and, furthermore, efficient if the noise is Gaussian. The KF is given in Algorithm 1, where Joseph's form is used for the covariance measurement update due to its better numerical properties and that it does not assume the filter gain is optimal. Given its favorable properties, the KF is often used as basis in algorithms to solve nonlinear problems.

B. Iterated Extended Kalman Filter

The KF is not directly applicable to nonlinear systems, however, the EKF provides a solution based on linearizing the model. The algorithm is identical to Algorithm 1 with $F = f'_x(\hat{x}_{t-1|t-1}, 0)$, $G = f'_w(\hat{x}_{t-1|t-1}, 0)$, and $H = h'_x(\hat{x}_{t|t-1})$ with the exceptions that

$$\hat{x}_{t|t-1} = f(\hat{x}_{t-1|t-1}, 0), \qquad \hat{y}_t = h(\hat{x}_{t|t-1}). \tag{4}$$

Algorithm 2 Iterated extended Kalman measurement update

The following measurement update is used instead of the measurement update in Algorithm 1.

Measurement update: (Let $x_0 = \hat{x} = \hat{x}_{t|t-1}$, $P = P_{t|t-1}$, $R = R_t$, and $y = y_t$.)
for $i = 0, \dots$ **do**

$$x_{i+1} = \hat{x} + K_i\big(y - h(x_i) - H_i(\hat{x} - x_i)\big) \tag{5a}$$
$$K_i = P^{xy}_i(P^{yy}_i)^{-1} \qquad H_i = h'_x(x_i) \tag{5b}$$
$$P^{xy}_i = PH_i^T \qquad P^{yy}_i = H_iPH_i^T + R \tag{5c}$$

end for
Output estimates:

$$\hat{x}_{t|t} = x_{i+1}, \tag{6a}$$
$$P_{t|t} = (I - K_iH_i)P(I - K_iH_i)^T + K_iRK_i^T \tag{6b}$$

In this case optimality, or even convergence, can no longer be guaranteed, and different improvements have been suggested. One alternative is to introduce an iterative measurement update step [5, 9, 20]. This can be motivated by the fact that the measurement update is a Gauss-Newton method, [11, 21], which in general relies on iterations. A generic measurement update step in the resulting *iterated extended Kalman filter* (IEKF) is given in Algorithm 2 [9].

C. Unscented Kalman Filter

Another alternative is to use the *unscented transform* (UT) to obtain the necessary quantities in Algorithm 1. This yields the unscented Kalman filter.

The UT approximates the distribution of a stochastic variable x after the mapping $y = f(x)$, assuming $\hat{x} = \mathbb{E}(x)$ and $P = P^{xx} = \text{cov}(x)$, using carefully selected and weighted samples, denoted *sigma points*. Usually the sigma points are selected according to

$$\mathcal{X}^{(0)} = \hat{x} \tag{7a}$$
$$\mathcal{X}^{(\pm k)} = \hat{x} \pm \Big[\sqrt{(n_x + \lambda)P}\Big]_{:k}, \quad k = 1, \dots, L, \tag{7b}$$

where n_x is the number of states and $\lambda = \alpha^2(n_x + \kappa) - n_x$, with the UT tuning parameters α, β, and κ. The $[A]_{:k}$ notation denotes selection of the k^{th} column of A, which are assigned the weights

$$W^{(0)} = \frac{\lambda}{n_x + \lambda} \qquad W^{(\pm k)} = \frac{1}{2(n_x + \lambda)}. \tag{7c}$$

The distribution of y is now approximated using weighted sample means,

$$\mathcal{Y}^{(k)} = f(\mathcal{X}^{(k)}) \tag{8a}$$
$$\hat{y} = \sum_k W^{(k)}\mathcal{Y}^{(k)} \tag{8b}$$
$$P^{yy} = \sum_k W^{(k)}(\mathcal{Y}^{(k)} - \hat{y})(\mathcal{Y}^{(k)} - \hat{y})^T + R + (1 - \alpha^2 + \beta)(\mathcal{Y}^{(0)} - \hat{y})(\mathcal{Y}^{(0)} - \hat{y})^T \tag{8c}$$

TABLE I
PARAMETER SETTINGS FOR SOME DIFFERENT UKF VERSIONS.

Parameter	UT1 [22]	UT2 [3]	CT [23]	DFT [24]
α	$\sqrt{3/n_x}$	10^{-3}	1	—
β	$3/n_x - 1$	2	0	—
κ	0	0	0	—
λ	$3 - n_x$	$10^{-6}n_x - n_x$	0	0
$\sqrt{n_x + \lambda}$	$\sqrt{3}$	$10^{-3}\sqrt{n_x}$	$\sqrt{n_x}$	$\frac{1}{a}\sqrt{n_x}$
$W^{(0)}$	$1 - n_x/3$	$1 - 10^6$	0	0

Algorithm 3 Unscented Kalman Filter

Assume $\hat{x}_{0|0}$ and $P_{0|0}$ given.
for all t **do**
Time update: Apply the UT to (1a), assuming the mean $\hat{x}_{t-1|t-1}$ and covariance $P_{t-1|t-1}$ for x_{t-1}, to obtain $\hat{x}_{t|t-1}$ and $P_{t|t-1}$.
Measurement update: Apply the UT to (1b), assuming the mean $\hat{x}_{t|t-1}$ and covariance $P_{t|t-1}$ for x_t, to obtain $\hat{y}_t$, $P^{yy}_{t|t-1}$, and $P^{xy}_{t|t-1}$.
Perform the regular KF update.
end for

$$\begin{aligned} P^{xy} = &\sum_k W^{(k)}(\mathcal{X}^{(k)} - \hat{x})(\mathcal{Y}^{(k)} - \hat{y})^T \\ &+ (1 - \alpha^2 + \beta)(\mathcal{X}^{(0)} - \hat{x})(\mathcal{Y}^{(0)} - \hat{y})^T \end{aligned} \tag{8d}$$

where $W_c^{(i)} = W^{(0)} + (1-\alpha^2+\beta)$ is often used as shorthand in the covariance expressions. The additional term in the covariance expressions try to ensure a positive definite result. Several different schemes for selecting the parameters α, β, and κ exist, see, *e.g.*, Table I and [4].

The time update step of the UKF is a direct application of the UT to (1a). For the measurement update, direct application of the UT to (1b) is used to obtain $\hat{y}_t$, $P^{yy}_{t|t-1}$, and $P^{xy}_{t|t-1}$. The UKF algorithm is summarized in Algorithm 3.

The method suffers from many of the same problems as the EKF do, and the remainder of this paper is dedicated to deriving and evaluating different iterated UKF filters.

III. ITERATED UNSCENTED KALMAN FILTER

Deriving an IUKF within the Fisher estimation framework, similar to the IEKF presented in [9], amounts to minimizing the following cost function in the measurement update step of the filter:

$$\hat{x}_{t|t} = \arg\min_x V(x) = \arg\min_x \tfrac{1}{2} r^T(x) r(x)$$

where

$$r(x) = \begin{bmatrix} R_t^{-1/2}\big(y_t - h(x)\big) \\ P_{t|t-1}^{-1/2}(\hat{x}_{t|t-1} - x) \end{bmatrix}. \tag{9}$$

Depending on the optimization method used and the interpretation of the cost function $V(x)$, different filters will be derived.

A. Basic Iterated Unscented Kalman Filter

First a UKF version of Algorithm 2 is derived. As the IEKF version, it assumes that the measurement function is affine in a neighborhood of x and x_i and hence that $h'_x(x) = h'_x(x_i) = H_i$. The Jacobian H_i is not explicitly computed in the UKF, but a stochastic linearization can be derived from the fact that $P^{xy} = PH^T$ in the linear case. Hence, a reasonable approximation of H_i in the IUKF is

$$H_i = (P_i^{xy})^T P^{-1}, \tag{10}$$

where symmetry in P has been utilized. As also suggested in [8], P is assumed constant during the iterations and hence the spread sigma points are constant as in [14]. This is different than the approach in [15, 16] where the covariance is updated in each iteration. Note, however, that as P and P^{xy} implicitly include second order transformation effects [4], the expression is only an approximation of the Jacobian for nonlinear models.

Using the above stochastic linearization argument, the state iteration in IEKF (5a) can be used to obtain

$$\begin{aligned} x_{i+1} &= \hat{x} + K_i\big(y - \hat{y}_i - (P_i^{xy})^T P^{-1}(\hat{x} - x_i)\big) \\ &= \hat{x} + K_i\big(y - \hat{y}_i - H_i(\hat{x} - x_i)\big), \end{aligned} \tag{11}$$

where

$$K_i = P_i^{xy}(P_i^{yy})^{-1}, \tag{12}$$

which can be used as basis in the IUKF. Note that $\hat{x} = \hat{x}_{t|t-1}$ is kept constant throughout the iterations (5). That is, the exact same procedure as in the IEKF (Algorithm 2) can be used in the IUKF by deriving P_i^{xy} and P_i^{yy} from the UT instead of a linearization as in the IEKF. It remains to determine the predicted measurement $\hat{y}_i$. There are two different natural choices:

$$\hat{y}_i = \sum_k W^{(k)} \mathcal{Y}_i^{(k)}, \tag{13a}$$

i.e., as the UT prediction of the measurement, or as

$$\hat{y}_i^* = \mathcal{Y}_i^{(0)}, \tag{13b}$$

i.e., the transformed center sigma point which is here denoted by superscript *. The two choices result in two slightly different interpretations of the cost function

$$\begin{aligned} V(x) = &\big(y_t - \mathbb{E}[h(x)]\big)^T R_t^{-1}\big(y_t - \mathbb{E}[h(x)]\big) \\ &+ (\hat{x}_{t|t-1} - x)^T P_{t|t-1}^{-1}(\hat{x}_{t|t-1} - x) \end{aligned} \tag{14a}$$

$$\begin{aligned} V^*(x) = &\big(y_t - h(x)\big)^T R_t^{-1}\big(y_t - h(x)\big) \\ &+ (\hat{x}_{t|t-1} - x)^T P_{t|t-1}^{-1}(\hat{x}_{t|t-1} - x), \end{aligned} \tag{14b}$$

both constituting different approximations of the cost function in (9).

Two things are worth noticing at this point.

- The former expression defines the cost in terms of the expected value of the measurement, not the measurement given by the current state estimate as the latter expression does. Hence, the latter gives a more direct connection to the current state.

Algorithm 4 Iterated Unscented Kalman Measurement Update

The following measurement update is used instead of the measurement update in Algorithm 3.

Measurement update: (Let $x_0 = \hat{x}_{t|t-1}$, $\hat{x} = \hat{x}_{t|t-1}$, $R = R_t$, and $y = y_t$.)

for $i = 0, \dots$ **do** until convergence

$$x_{i+1} = \hat{x} + K_i\big(y - \hat{y}_i - (P_i^{xy})^T P^{-1}(\hat{x} - x_i)\big), \quad (15)$$

with

$$K_i = P_i^{xy}(P_i^{yy})^{-1}, \quad (16)$$

where as in the UKF, P_i^{yy} and P_i^{xy} are all derived using the UT using new sigma points around the current estimate. $\hat{y}_i$ is either the predicted measurement from the UT or the $\mathcal{Y}_i^{(0)}$ sigma point.

end for

Output estimates:

$$\hat{x}_{t|t} = x_{i+1}, \quad (17a)$$

$$P_{t|t} = (I - K_i H_i)P(I - K_i H_i)^T + K_i R K_i^T. \quad (17b)$$

- For the EKF the two expressions are identical due to the underlying approximation. At first, it is not obvious which solution would be the best; therefore, both will be considered when evaluating the IUKF.

The measurement update in the generic IUKF is summarized in Algorithm 4.

B. Iterative Posterior Linearization Filter IPLF

The IPLF has similar update equations as the GN based IEKF/IUKF but it differs in three ways:

- IPLF approximates the measurement by moment matching using statistical linear regression.
- It introduces an extra matrix Ω_i, corresponding to the estimation error covariance as a result of the linearization of $h(x)$.
- The iterated covariance P_i is used to compute the Kalman gain. This is also done in the Levenberg-Marquardt versions of IEKF/IUKF but only to compute the search directions, and the final covariance estimates are here computed using (17b).

The IPLF derived in [17, 18] results in the following update equations, here stated in the notation used for the IUKF in this paper

$$x_{i+1} = \hat{x} + K_i\big(y - \hat{y}_i - H_i(\hat{x} - x_i)\big), \quad (18a)$$

$$P_{i+1} = P - K_i H_i P, \quad (18b)$$

$$K_i = P H_i^T (H_i P H_i^T + \Omega_i + R)^{-1}, \quad (18c)$$

$$\Omega_i = P_i^{yy} - H_i P_i H_i^T - R, \quad (18d)$$

where $\hat{y}_i, H_i$ and P_i^{yy} are computed using the current state iterate x_i and the UT with sigma points chosen using the iterated covariance P_i in (7b). As with the IUKF the predicted iterative measurement in the IPLF update can be chosen using (13a) (as used in [17, 18]) or (13b).

C. IUKF with Line Search

With a cost function at hand it is sensible to check whether the state correction results in a cost decrease in the first place, that is, $V(x_{i+1}) < V(x_i)$. If this is not the case a step size parameter α can be introduced

$$x_{i+1} = x_i + \alpha_i\Big(\hat{x} - x_i + K_i\big(y - \hat{y}_i - H_i(\hat{x} - x_i)\big)\Big). \quad (19)$$

The step size can be calculated using line search, *e.g.*, by checking for cost decrease, see [25] for detailed options. For GN, $\alpha_i = 1$, should be tried first. A line search can be added to both the IEKF and IUKF, and following our earlier convention the IUKF with lines search will be referred to as IUKF-L, in accordance with IEKF-L which was introduced in [9]. When the iterations have finished the state and covariance are updated using (17).

It is also straightforward to introduce a step size parameter to the IPLF (18a) by comparing it to (5a) but there is no obvious criterion to minimize. One option is given in [19], where (18a) is used as an inner GN search with the covariance fixed, in which the cost function $V(x)$ (14a) is used.

D. Quasi-Newton IUKF

Needing only to compute first order derivatives, as compared to Newton's method, makes the Gauss-Newton method an attractive option. However, second order terms may be of great importance for finding a good search direction when the starting point is far from the optimum or when the Jacobian is rank-deficient. As in [9], a model of the second order terms matrix, T_i, can be added to the IUKF update as

$$x_{i+1} = \hat{x} + K_i^q\big(y - \hat{y}_i - H_i\tilde{x}_i\big) - S_i^q T_i \tilde{x}_i, \quad (20a)$$

$$S_i^q = (H_i^T R^{-1} H_i + P^{-1} + T_i)^{-1}, \quad (20b)$$

$$K_i^q = S_i^q H_i^T R^{-1}, \quad (20c)$$

where $\tilde{x}_i = \hat{x} - x_i$ has been introduced. Adding a step size parameter to (20) results in

$$x_{i+1} = x_i + \alpha_i\big(\tilde{x}_i + K_i^q(y - \hat{y}_i - H_i\tilde{x}_i) - S_i^q T_i \tilde{x}_i\big), \quad (21)$$

which we shall refer to as *Quasi-Newton-IUKF* (QN-IUKF). The updated covariance is again computed using (17b).

The matrix T_i can be computed numerically at run-time, see, *e.g.*, [26] or [27, Eq. (10.143)], using the scheme in Algorithm 5.

E. Levenberg-Marquardt IUKF

The third IUKF is based on the trust-region method called Levenberg-Marquard (LM), after its inventors [28, 29]. In LM a single parameter, μ, is used to control the search strategy by interpolating between steepest-descent and GN. The iterations for the *Levenberg-Marquardt-IUKF* (LM-IUKF) are

$$x_{i+1} = \hat{x} + K_i(y - \hat{y}_i - H_i\tilde{x}_i) - \mu_i(I - K_i H_i)\tilde{P}_i B_i \tilde{x}_i, \quad (23a)$$

$$\tilde{P}_i = \big(I - P(P + \mu_i^{-1} B_i^{-1})^{-1}\big)P, \quad (23b)$$

$$K_i = \tilde{P}_i H_i^T (H_i \tilde{P}_i H_i^T + R)^{-1}, \quad (23c)$$

$$B_i = \mathrm{diag}(H_i^T R^{-1} H_i + P^{-1}), \quad (23d)$$

Algorithm 5 Hessian Model Calculation

Require: $H_i, H_{i-1}, \hat{y}_i, \hat{y}_{i-1}$ using UT and initialize $T_0 = 0$.

for $i = 1, \ldots$ **do** until convergence of (21)

$$v_i = -H_i^T R^{-1}(y - \hat{y}_i) + H_{i-1}^T R^{-1}(y - \hat{y}_{i-1}),$$

$$z_i = (H_{i-1} - H_i)^T R^{-1}(y - \hat{y}_i), \tag{22a}$$

$$s_i = x_i - x_{i-1}, \tag{22b}$$

$$\tau_i = \min\left(1, \frac{|s_i^T z_i|}{|s_i^T T_i s_i|}\right), \tag{22c}$$

$$T_{i-1} = \tau_i T_{i-1}, \tag{22d}$$

$$w_i = z_i - T_{i-1} s_i, \tag{22e}$$

$$T_i = T_{i-1} + \frac{w_i v_i^T + v_i w_i^T}{v_i^T s_i} - \frac{w_i^T s_i}{(v_i^T s_i)^2} v_i v_i^T, \tag{22f}$$

end for

as shown in [9], where again $\hat{y}_i$ and H_i are computed using UT. As with IUKF-L and QN-IUKF the final covariance is computed using (17b). Note that (23b) is used in the sigma point transformation (7b) similar to the IPLF update (18). The damping factor μ is, in our examples, initially set to 0.01 and then multiplied by 10 if no cost decrease is obtained and divided by 10 until no further cost decrease can be found.

IV. Results

In this section all iterative and non-iterative EKF and UKF versions are studied using three examples with simulated data. The IUKFs/IPLF with measurement predictions using (13b) and cost function (14b) are denoted with superscript * and the other IUKFs/IPLF use the standard UT prediction (13a) and the cost function (14a). All results are obtained using sigma points computed using UT2 with the standard setting, see Table I.

A. Bearings Only Tracking

A bearings only example with a stationary target is evaluated. This example was also studied in [9, 30, 31] and we will therefore only briefly describe the model. The target with $2D$ state $x = [\mathsf{X}, \mathsf{Y}]^T$ is stationary at the true position $x^* = [1.5, 1.5]^T$. The bearing measurement function from the j-th sensor S^j at time t is

$$y_t^j = h_j(x_t) + e_t = \arctan2(\mathsf{Y}_t - S_{\mathsf{Y}}^j, \mathsf{X}_t - S_{\mathsf{X}}^j) + e_t, \tag{24}$$

where $\arctan2()$ is the two argument arctangent function, S_{Y} and S_{X} denotes the Y and the X coordinates of the sensors, respectively. With the two sensors having positions $S^1 = [0, 0]^T$ and $S^2 = [1.5, 0]^T$. The filters are initialised with $\hat{x}_{0|0} = [0.5, 0.1]^T$, $P_{0|0} = 0.1 I_2$ and $R = \pi^2 10^{-5} I_2$. After 10 iterations IPLF and IPLF* (*i.e.*, using (13b)) has converged to an error of 11.7 and 11.8, respectively. All GN based iterative filters have converged to an error of 0.02 while both the EKF and the UKF have an error of 3.19. Also the covariances for all GN iterated filters cover the true state at 2σ, while the UKF, the EKF and the IPLFs' covariances are too small which is also indicated in Table II by comparing the square root of the trace of covariance matrix, $\sqrt{\mathrm{tr}(P)}$, to the error. Hence, for this application all the GN iterated methods work well.

B. Simple Quadratic

A simple quadratic measurement function is

$$h(x) = -x^2 \tag{25}$$

with $x \in \mathbb{R}$ for which the UKF is known to produce strange results [4], while the IEKF handles this well [9]. It is therefore interesting to further analyze why the UKF has difficulties to produce small estimation errors and whether this could be resolved using any of the IUKF versions.

Here, 1000 measurements are generated, distributed according to $y \sim \mathcal{N}(h(x^*), R)$, with $R = 10^{-3}$, and the true state is distributed $x^* \sim \mathcal{N}(1, 10^{-1})$. The filters are initialized with $x_0 = 0.1$, $P_{0|0} = 1$ and 10 iterations are used in all iterative filters. In Fig. 1 the distribution of the estimation errors from all filters are shown. It can be seen that the IEKF and IUKF* (using (14b)) filters yield slightly better RMSEs than the EKF and UKF, while all the IUKF (using (14a)) filters all have a few large errors. The exception is the QN-based ones that all have large estimation errors and do not seem to move from the initial point. The IPLF have a few large errors but most of them are small while the IPLF* is almost on par with the best IEKFs. Using a smaller spread of the sigma points, as in [18], then IPLF and IUKF-L* performs marginally better.

To understand why the IUKFs using the loss function (14a) work considerably worse than the ones using (14b), lets study the special case with a noise-free measurement $y = -1$ and the prediction $x = 0.1$ and $P = 1$. Here, $-h(x) \sim \chi_1^2(.1^2)$[1] and hence $\mathbb{E}[h(x)] = 1.01$, and $y - \mathbb{E}[h(x)] = 0.01$. The result is that the error due to the measurement is minor, whereas $y - h(x) = -0.99$ in the second case. Here, it would seem it is not favorable to use the expected measurement instead of the transformed estimate only. This explains the difference observed, but provides no definitive answer to what is generally the best thing to use.

C. Convex Search

A difficult, but convex, nonlinearity is given by the two dimensional problem

$$h(x) = e^{\mathsf{X}+3\mathsf{Y}-0.1} + e^{\mathsf{X}-3\mathsf{Y}-0.1} + e^{\mathsf{X}-0.1} \tag{26}$$

with $x = [\mathsf{X}, \mathsf{Y}]^T \in \mathbb{R}^2$ which is used in many examples in [32]. All filters are initialized with $x_0 = [-0.9,\ 0.8]^T$ and initial covariance $P_0 = 0.1 \cdot I_2$ while the true state is located in $x^* = [-0.5,\ 0]^T$. Ten iterations are used in all iterative filters and $R = 10^{-2}$, while the measurement is ideal. The covariance estimate for UKF and the IPLFs resulted in a negative eigenvalue and hence there is no covariance to show for these two in Fig. 2. The EKF fails completely as the covariance approximation does not capture the true state. The IUKFs does not improve the estimation error as much as the

[1] $\chi_n^2(\lambda)$ is the non-central chi-2 distribution of degree n and with offset λ.

TABLE II
FINAL ESTIMATION ERROR FOR THE BEARINGS ONLY EXAMPLE.

Method	UKF	EKF	IUKF-L	LM-IUKF	QN-IUKF	IUKF-L*	LM-IUKF*	QN-IUKF*	IEKF-L	LM-IEKF	QN-IEKF	IPLF	IPLF*
$\|x - x*\|$	3.19	3.19	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	11.70	11.77
$\sqrt{\mathrm{tr}(P)}$	0.03	0.03	0.03	0.03	0.03	0.03	0.44	0.03	0.03	0.03	0.03	0.30	0.30

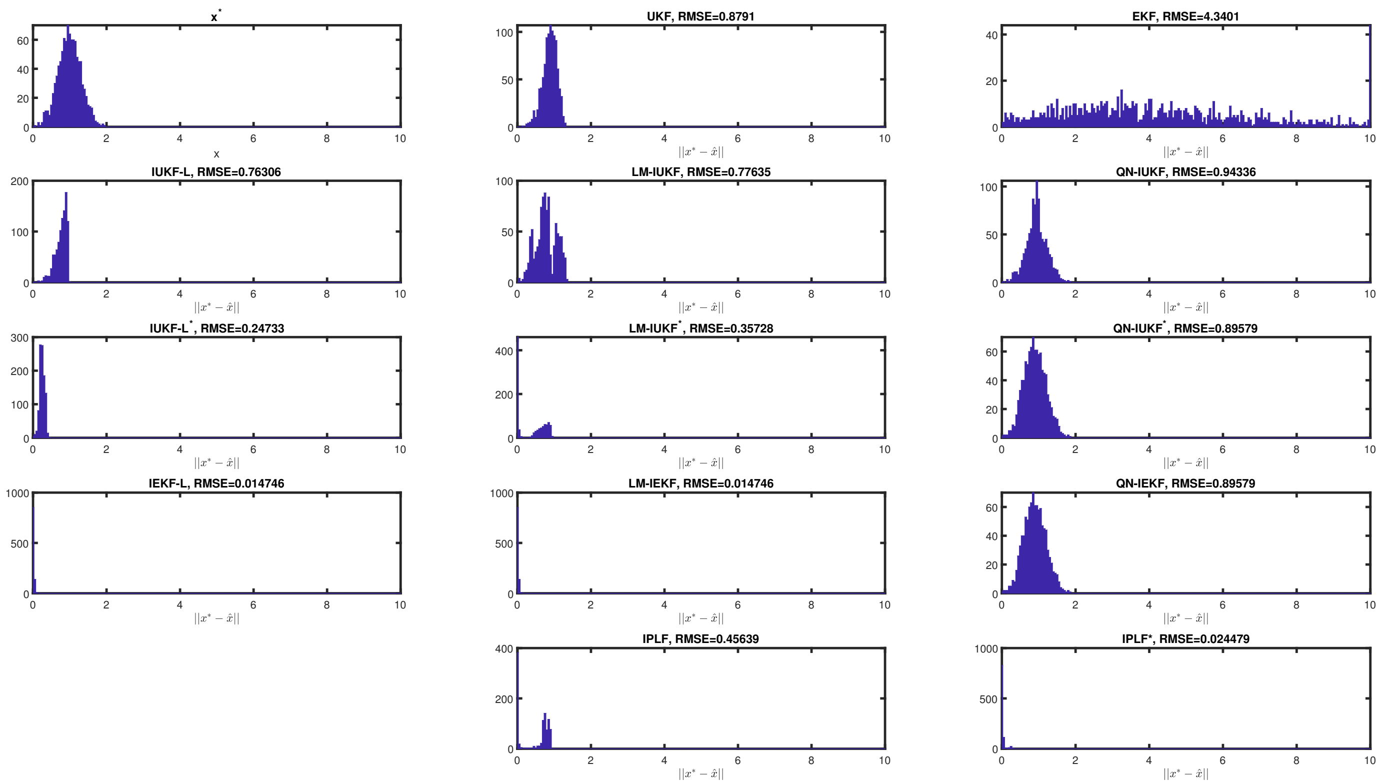

Fig. 1. Distribution of the estimation errors on the measurement function (25). The first row shows the distribution of the true state, the UKF and the EKF estimation errors. The second row shows the IUKFs (using (14a)). The third row show the IUKF* (using (14b)). The forth row shows the IEKFs. The last row shows the IPLFs.

IUKF*s, again suggesting that it may be better to use (13b) when there are strong nonlinearities. The IEKFs all produce small estimation errors with trustworthy covariance estimates. The IPLF* does improve a little over the starting point while the IPLF fails completely.

In a second evaluation of (26), 1000 starting points are distributed $x_0 \sim \mathcal{N}\left(x^*, 10^{-1} I_2\right)$ with $x^* = [-0.5,\ 0]^T$ and the measurements are distributed $y \sim \mathcal{N}\left((h(x^*), 10^{-2}\right)$. The filters are initialized with $x_0 = [-0.9,\ 0.8]^T$ and covariance $P_0 = 0.1 \cdot I_2$ in all 1000 estimates. In all iterative filters 10 iterations were used. In Fig. 3 the distribution of the absolute values of the estimation errors are shown. As expected, given the results from the first evaluation of this function, the IEKFs yields small estimation errors. Both the UKF, IUKF-L IUKF-L* and IEKF-L clearly result in some really bad estimates. Notably all QN and LM-based filters have rather small estimation errors with a few exceptions in LM-IUKF and LM-IUKF*. Again the IPLF* has smaller RMSE than the IPLF.

V. CONCLUSIONS

In this paper six versions of iterated unscented Kalman filters have been presented and compared with three iterated extended Kalman filters, the recent IPFL, and the non-iterated UKF and EKF. The filters have been evaluated using simulations in three benchmark examples. Three methods were used to compute the measurement prediction corresponding to two different cost functions. The first cost function uses the UT and the other uses only the UT center point as in the EKF, and an interpretation is given to the difference between the two.

For mild nonlinearities, as in the bearings only example Sec. IV-A, all iterated IUKFs and IEKFs seem to give nearly identical results while the IPLFs' covariances are inconsistent with too small covariances. The UKF can give poor results on quadratic functions as noted in [4]. A quadratic function is studied in the second example in Sec. IV-B highlighting why the IUKF*s can here yield smaller estimation errors which also seem to apply to the IPLF*. The last example is a difficult,

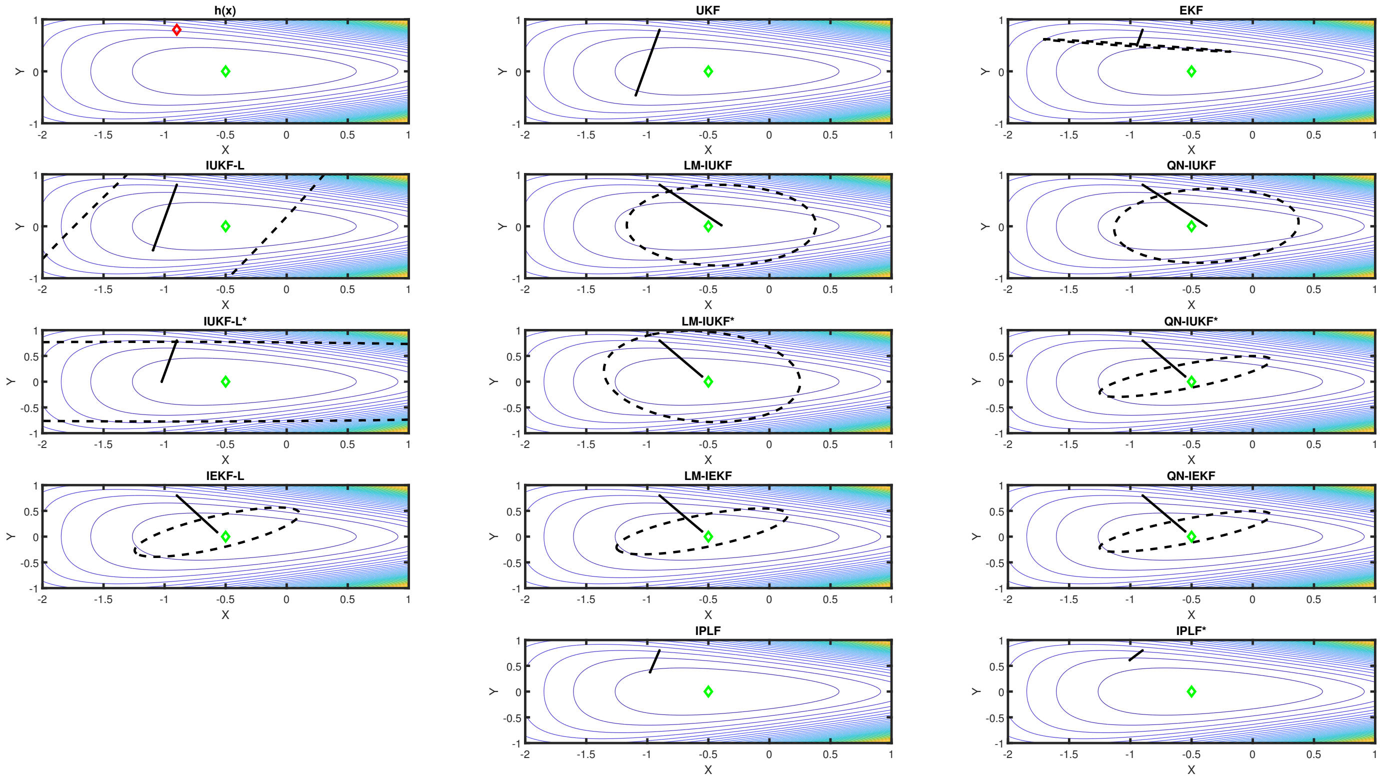

Fig. 2. The difficult, but convex, measurement function (26) is evaluated using 10 iterations. All IEKFs have small estimation errors while the EKF have a large error and a poor covariance estimate. The best performing the QN and LM-based IUKFs are on par with the IEKFs while IUKF-L and IUKF-L* are not that good. The IPLFs covariances (not shown) are degenerate since they have a negative eigenvalue.

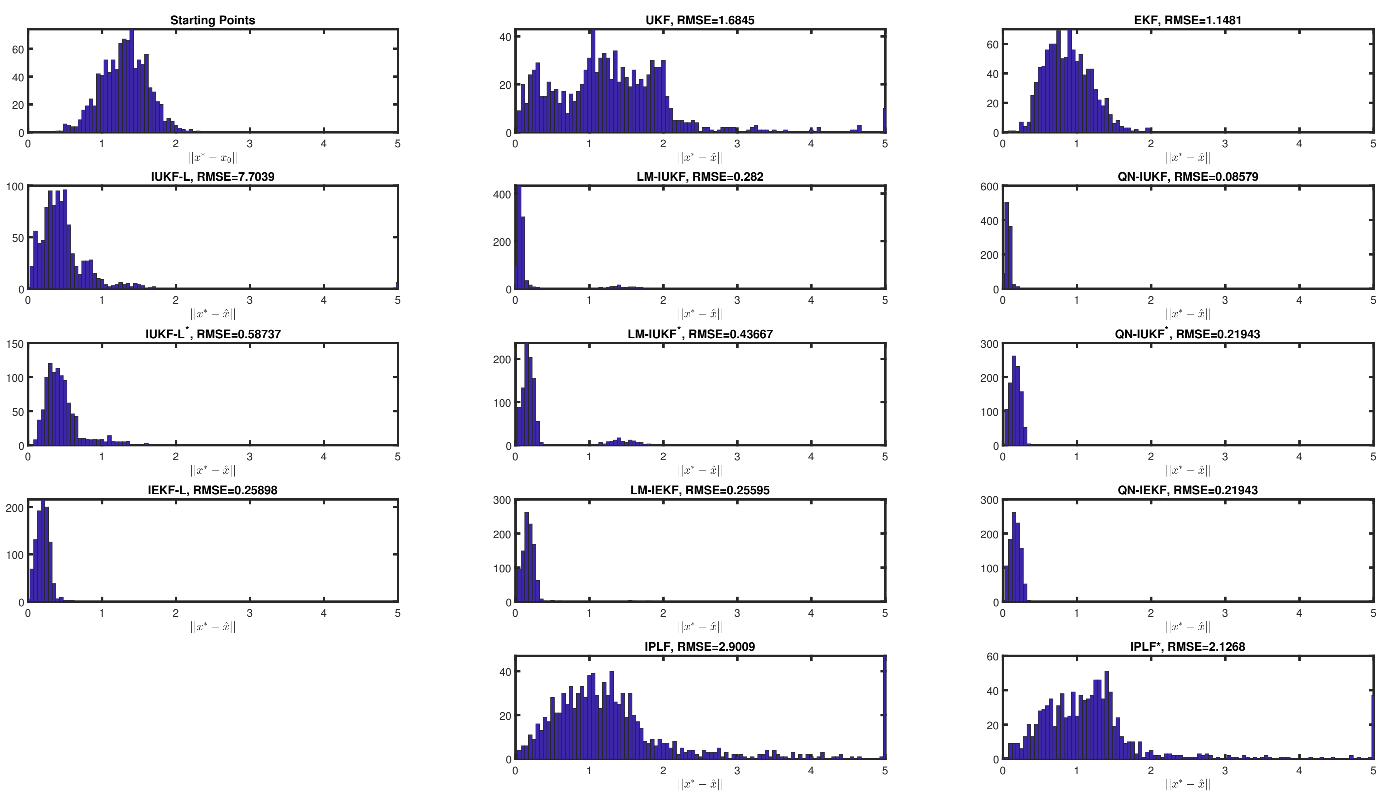

Fig. 3. Distribution of the absolute values of the estimation errors using the measurement function (26). The upper left plot shows the distribution of the absolute values of the starting coordinates. The second row shows that the IUKFs sometimes can have some rather poor estimates. The third row shows the IUKF*. The forth row shows the IEKFs and the fifth show the IPLFs.

but convex, nonlinearity for which line search based IEKF and IUKFs can have difficulties finding a good search direction. For this strong nonlinearity LM and QN-based filters are the most promising candidates, especially QN-IUKF which has the smallest RMSE. The IPLFs does not promise much here and would probably benefit from a guided search using a cost function.

In conclusion, in the evaluated examples in this paper, the IEKFs and IUKFs gives similar results in most examples highlighting that the GN interpretation is indeed applicable to IUKF. The choice of method could therefore be motivated based on other aspects such as, *e.g.*, simpler implementation without derivatives using IUKFs or fewer computations using IEKFs.

For future work the different properties of the IUKF*s, IUKFs, IPLFs and IPLF* should be studied. The effect on estimator performance using different sigma point methods and damped IPLFs [19] should be analyzed further. Work in the direction of iterated extended Kalman smoothers, [33], could perhaps be utilized to yield the unscented counterpart.

References

[1] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.

[2] ——, "A new extension of the Kalman filter to nonlinear systems," in *Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*, ser. SPIE proceedings series, Orlando, FL, USA, April 1997, pp. 182–193.

[3] E. A. Wan and R. van der Merwe, "The unscented Kalman filter for nonlinear estimation," in *Proceedings of IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium*, Lake Louise, AB, Canada, Oct. 2000, pp. 153–158.

[4] F. Gustafsson and G. Hendeby, "Some relations between extended and unscented Kalman filters," *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 545–555, Feb. 2012.

[5] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, ser. Mathematics in Science and Engineering. Academic Press, Inc, 1970, vol. 64.

[6] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the American Society of Mechanical Engineering — Journal Basic Engieering, Series D*, vol. 82, pp. 35–45, Mar. 1960.

[7] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Prentice-Hall, Inc, 2000.

[8] B. M. Bell and F. W. Cathey, "The iterated Kalman filter update as a Gauss-Newton method," *IEEE Transactions on Automatic Control*, vol. 38, no. 2, pp. 294–297, Feb. 1993.

[9] M. A. Skoglund, G. Hendeby, and D. Axehill, "Extended Kalman filter modifications based on an optimization view point," in *Proceedings of 18th IEEE International Conference on Information Fusion*, Washington, DC, USA, Jul. 6–9 2015.

[10] Y. Bar-Shalom and X.-R. Li, *Estimation and tracking: principles, techniques and software*. Boston: Artech House, 1993.

[11] P. J. G. Teunissen, "On the local convergence of the iterated extended Kalman filter," in *Proceeding of the XX General Assembly of the IUGG, IAG Section IV*. Vienna: IAG, Aug. 1991, pp. 177–184.

[12] N. U. Ahmed and S. M. Radaideh, "Modified extended Kalman filtering," *IEEE Transactions on Automatic Control*, vol. 39, no. 6, pp. 1322–1326, Jun. 1994.

[13] J. Havlík and O. Straka, "Performance evaluation of iterated extended Kalman filter with variable step-length," *Journal of Physics: Conference Series*, vol. 659, no. 1, 2015. [Online]. Available: http://stacks.iop.org/1742-6596/659/i=1/a=012022

[14] G. Sibley, G. S. Sukhatme, and L. Matthies, "The iterated sigma point Kalman filter with applications to long range stereo." *Robotics: Science and Systems*, vol. 8, no. 1, pp. 235–244, 2006.

[15] R. Zhan and J. Wan, "Iterated unscented kalman filter for passive target tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 3, pp. 1155–1163, July 2007.

[16] Z. Xie and J. Feng, "Real-time nonlinear structural system identification via iterated unscented kalman filter," *Mechanical Systems and Signal Processing*, vol. 28, pp. 309 – 322, 2012, interdisciplinary and Integration Aspects in Structural Health Monitoring. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0888327011000690

[17] Á. F. García-Fernández, L. Svensson, and M. R. Morelande, "Iterated statistical linear regression for Bayesian updates," in *17th International Conference on Information Fusion (FUSION)*, Jul. 2014.

[18] Á. F. García-Fernández, L. Svensson, M. R. Morelande, and S. Särkkä, "Posterior linearization filter: Principles and implementation using sigma points," *IEEE Transactions on Signal Processing*, vol. 63, no. 20, pp. 5561–5573, Oct. 2015.

[19] M. Raitoharju, L. Svensson, Á. F. García-Fernández, and R. Piché, "Damped posterior linearization filter," *IEEE Signal Processing Letters*, vol. 25, no. 4, pp. 536–540, Apr. 2018.

[20] R. L. Bellaire, E. W. Kamen, and S. M. Zabin, "New nonlinear iterated filter with applications to target tracking," in *SPIE Signal and Data Processing of Small Targets*, vol. 2561, San Diego, CA, USA, Sep. 1 1995, pp. 240–251.

[21] D. Bertsekas, "Incremental least squares methods and the extended Kalman filter," in *Proceedings of the 33rd IEEE Conference on Decision and Control*, vol. 2, Lake Buena Vista, Florida, USA, Dec. 1994, pp. 1211–1214.

[22] S. J. Julier, J. K. 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, Mar. 2000.

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

[24] B. M. Quine, "A derivative-free implementation of the extended Kalman filter," *Automatica*, vol. 42, no. 11, pp. 1927–1934, Nov. 2006.

[25] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York: Springer, 2006.

[26] J. E. Dennis, D. M. Gay, and R. E. Welsh, "Algorithm 573 - NL2SOL, An adaptive nonlinear least-squares algorithm," *ACM Transactions on Mathematical Software*, vol. 7, pp. 348–368, 1981.

[27] J. Nocedal and S. Wright, *Numerical Optimization (Springer Series in Operations Research and Financial Engineering)*, 2nd ed. Springer, Jul. 2006.

[28] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *Quarterly Journal of Applied Mathmatics*, vol. II, no. 2, pp. 164–168, 1944.

[29] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *SIAM Journal on Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[30] G. Hendeby, "Performance and implementation aspects of nonlinear filtering," Dissertations No 1161, Linköping Studies in Science and Technology, Mar. 2008.

[31] ——, "Fundamental estimation and detection limits in linear non-Gaussian systems," Licentiate Thesis No 1199, Department of Electrical Engineering, Linköpings universitet, Sweden, Nov. 2005.

[32] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[33] B. M. Bell, "The iterated Kalman smoother as a Gauss-Newton method," *SIAM Journal on Optimization*, vol. 4, no. 3, 1994.