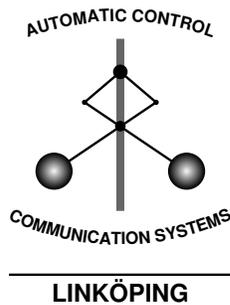# Performance Issues in Non-Gaussian Filtering Problems

Gustaf Hendeby, Rickard Karlsson, Fredrik Gustafsson, Neil Gordon

Division of Automatic Control
Department of Electrical Engineering
Linköpings universitet, SE-581 83 Linköping, Sweden
WWW: `http://www.control.isy.liu.se`
E-mail: `hendeby@isy.liu.se`, `rickard@isy.liu.se`,
`fredrik@isy.liu.se`, `-`

22nd June 2006

**AUTOMATIC CONTROL**

**COMMUNICATION SYSTEMS**

**LINKÖPING**

## Abstract

Performance for filtering problems is usually measured using the second-order moment. For non-Gaussian applications, this measure is not always sufficient. In this paper, the Kullback divergence is extensively used to compare estimated distributions. Several estimation techniques are compared, and methods with ability to express non-Gaussian posterior distributions are shown to give superior performance over classical second-order moment based estimators.

# 1  Introduction

Many estimation methods rely on linearization to handle nonlinear models, or consider just first- and second-order moment estimation of the underlying posterior *probability density function* (PDF). For instance, the *extended Kalman filter* (EKF, [1, 2, 3]) has difficulties when higher order moments, or the full PDF is needed. *Multiple model filters* (MMF, [2][4]) and more generally the *particle filter* (PF, [5, 6, 7]), are methods that can be used to overcome this problem.

A key question is how to evaluate how much there is to gain by estimating the full PDF. Traditionally, the *mean square error* (MSE) is used for this. However, if the problem is such that the second order moment can be sufficiently well approximated by a Gaussian density, but not the higher order moments, then another measure is needed. In this paper the *Kullback divergence* (KD) is introduced to compare the true PDF and the one obtained from an estimator.

Studying the full PDF or higher order moment estimation can be motivated by detection or hypothesis testing, where the posterior distribution is needed. Another situation is when the probability for an event is desired, this can be solved by integration of the PDF over the interesting region.

This technical report is an extended version of the paper [8] submitted to the Nonlinear Statistical Signal Processing Workshop in Cambridge, 2006.

# 2  Recursive Bayesian Estimation

Consider the discrete state-space model

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

with state variables $x_t \in \mathbb{R}^n$, input signal $u_t$ and measurements $\mathbb{Y}_t = \{y_i\}_{i=1}^t$, with known PDFs for the process noise, $p_w(w)$, and measurement noise, $p_e(e)$. The nonlinear prediction density $p(x_{t+1}|\mathbb{Y}_t)$ and filtering density $p(x_t|\mathbb{Y}_t)$ for the Bayesian inference, [1], are given by

$$p(x_{t+1}|\mathbb{Y}_t) = \int_{\mathbb{R}^n} p(x_{t+1}|x_t)p(x_t|\mathbb{Y}_t) \, dx_t, \tag{2a}$$
$$p(x_t|\mathbb{Y}_t) = \frac{p(y_t|x_t)p(x_t|\mathbb{Y}_{t-1})}{p(y_t|\mathbb{Y}_{t-1})}. \tag{2b}$$

For the important special case of linear-Gaussian dynamics and linear-Gaussian observations the Kalman filter (KF), [9], gives a finite dimensional recursive solution. For nonlinear and/or non-Gaussian systems, the PDF cannot in general be expressed with a finite number of parameters. Instead approximative methods must be used. Usually this is done in two ways; either by approximating the system or by approximating the posterior PDF. See for instance, [10, 11]. Here two different approaches of solving the Bayesian equations are considered; EKF and PF. The EKF will solve the problem using a linearization of the system and

assuming Gaussian noise. The PF on the other hand will approximately solve the Bayesian equations by stochastic integration. Hence, no linearization errors occur and non-Gaussian noise is not a problem.

## 2.1 The Extended Kalman Filter

For the special case of linear dynamics, linear measurements and additive Gaussian noise, the Bayesian recursions have a finite dimensional recursive solution, the Kalman filter. For many nonlinear problems the noise assumptions and the nonlinearity are such that a linearized solution will be a good approximation. This is the idea behind the EKF, [2, 3, 12], where the model is linearized around the previous estimate. Here the time update and measurement update for the EKF are presented,

$$
\begin{cases}
\hat{x}_{t+1|t} = f(\hat{x}_{t|t}, u_t, 0), \\
P_{t+1|t} = F_t P_{t|t} F_t^T + G_t Q_t G_t^T,
\end{cases} \tag{3a}
$$

$$
\begin{cases}
\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(y_t - h(\hat{x}_{t|t-1})), \\
P_{t|t} = P_{t|t-1} - K_t H_t P_{t|t-1}, \\
K_t = P_{t|t-1} H_t^T (H_t P_{t|t-1} H_t^T + R_t)^{-1},
\end{cases} \tag{3b}
$$

where the linearized matrices are given as

$$
F_t = \nabla_x f|_{x_t = \hat{x}_{t|t}}, \qquad G_t = \nabla_w f|_{x_t = \hat{x}_{t|t}}, \qquad H_t = \nabla_x h|_{x_t = \hat{x}_{t|t-1}}.
$$

The noise covariances are given by

$$
Q_t = \mathsf{cov}\, w_t \ \text{ and } \ R_t = \mathsf{cov}\, e_t.
$$

## 2.2 Grid-Based Approximation

The grid-based approximation for a general integral in $\mathbb{R}^n$ is

$$
\int_{\mathbb{R}^n} g(x_t)\, dx_t \approx \sum_{i=1}^{N} g(x_t^{(i)}) \Delta^n, \tag{4}
$$

using a regular grid, where $\Delta^n$ represents the volume and where $\{x_t^{(i)}\}_{i=1}^N$ represents the value a specific grid. The approximation error depends on the grid size, $\Delta$, hence in principle on the dimension of the state space. In [13], the Bayesian approach is investigated for a discrete-time nonlinear system using this approximation. The Bayesian time update and measurement update are solved using an approximative numerical method, where the density functions are piecewise constant on regular regions in the state space. Applying this grid-based integration to the general Bayesian estimation problem, using the model

with additive noise yields the following approximation

$$p(x_t^{(i)}|\mathbb{Y}_t) = \gamma_t^{-1} p_{e_t}(y_t|x_t^{(i)}) p(x_t^{(i)}|\mathbb{Y}_{t-1}), \tag{5a}$$

$$p(x_{t+1}^{(i)}|\mathbb{Y}_t) = \sum_{j=1}^{N} p_{w_t}(x_{t+1}^{(i)}|x_t^{(j)}) p(x_t^{(j)}|\mathbb{Y}_t) \Delta^n, \tag{5b}$$

where $\gamma_t$ is a normalization factor. The minimum variance estimate is approximated with

$$\hat{x}_{t|t} = \mathsf{E}\, x_t|\mathbb{Y}_t \approx \sum_{i=1}^{N} x_t^{(i)} p(x_t^{(i)}|\mathbb{Y}_t) \Delta^n. \tag{6}$$

In [14], the nonlinear and non-Gaussian problem is analyzed using grid-based integration as well as Monte Carlo integration, discussed next, for prediction and smoothing. Several nonlinear systems are compared using different techniques. Also [15], discusses grid-based or *point mass filter* (PMF) and PF based approaches.

## 2.3 The Particle Filter

In this section, the presentation of the particle filter theory is according to [5, 6, 7, 15]. The PF achieves an approximate solution to the discrete time Bayesian estimation problem formulated in (2) by updating an approximate description of the posterior filtering density. Let $x_t$ denote the state of the observed system and $\mathbb{Y}_t = \{y^{(i)}\}_{i=1}^t$ be the set of observed measurements until present time. The PF approximates the density $p(x_t|\mathbb{Y}_t)$ by a large set of $N$ samples (particles), $\{x_t^{(i)}\}_{i=1}^N$, where each particle has an assigned relative weight, $\gamma_t^{(i)}$, chosen such that all weights sum to unity. The location and weight of each particle reflect the value of the density in the region of the state space, The PF updates the particle location and the corresponding weights recursively with each new observed measurement. For the common special case of additive measurement noise the unnormalized weights are given by

$$\gamma_t^{(i)} = \gamma_{t-1}^{(i)} p_e(y_t - h(x_t^{(i)})), \quad i = 1, \ldots, N. \tag{7}$$

Using the samples (particles) and the corresponding weights the Bayesian equations can be approximately solved. To avoid divergence a resampling step is introduced, [5], which is referred to as *sampling importance resampling* (SIR). The PF method is given in Alg. 1.

The estimate for each time, $t$, is often chosen as the *minimum mean square estimate*, i.e.,

$$\hat{x}_t = \mathsf{E}\, x_t = \int_{\mathbb{R}^n} x_t p(x_t|\mathbb{Y}_t)\, dx_t \approx \sum_{i=1}^{N} \gamma_t^{(i)} x_t^{(i)}. \tag{8}$$

The PF approximates the posterior PDF, $p(x_t|\mathbb{Y}_t)$, by a finite number of particles. There exist theoretical limits [6], that show that the approximated PDF converges to the true as the number of particles tends to infinity.

3

---

**Alg. 1** Sampling Importance Resampling (SIR)

---

1: Generate $N$ samples $\{x_0^{(i)}\}_{i=1}^N$ from $p(x_0)$.
2: Compute $\gamma_t^{(i)} = \gamma_{t-1}^{(i)} p_e\big(y_t - h(x_t^{(i)})\big)$ and normalize, *i.e.*, $\bar{\gamma}_t^{(i)} = \gamma_t^{(i)} / \sum_{j=1}^N \gamma_t^{(j)}$, $i = 1, \ldots, N$.
3: Generate a new set $\{x_t^{(i\star)}\}_{i=1}^N$ by resampling with replacement $N$ times from $\{x_t^{(i)}\}_{i=1}^N$, with probability $\bar{\gamma}_t^{(i)} = \Pr\{x_t^{(i\star)} = x_t^{(i)}\}$. Set $q_t^{(i)} = 1/N$.
4: $x_{t+1}^{(i)} = f(x_t^{(i\star)}, u_t, w_t^{(i)})$, $i = 1, \ldots, N$ using different noise realizations, $w_t^{(i)}$.
5: Increase $t$ and continue to step 2.

---

## 2.4 Multiple Model Filtering

By combining *multiple models* (MM) in a filter it is possible to obtain a better approximation of the underlying PDF than with a single linear Gaussian model. The general multiple model idea is often based on the *Gaussian sum* (GS) approximation, described in [2, 4]. The GS method approximates the PDF with a sum of Gaussian densities,

$$p(x_t|\mathbb{Y}_t) = \sum_{i=1}^N \gamma_t^{(i)} \mathcal{N}(\hat{x}_t^{(i)}, P_t^{(i)}), \tag{9}$$

where $x_t^{(i)}$ and $P_t^{(i)}$ represent mean and covariance of one hypothesis, and $\gamma_t^{(i)}$ represents the trust in the hypothesis.

The different hypotheses are then handles separately, usually using a KF to update $x_t^{(i)}$ and $P_t^{(i)}$. The probability of the hypotheses $\gamma_t^{(i)}$, that always should add to one, is updated based both on the likelihood of the system switching between the different modes (a model feature) and the likelihood of the obtained measurements,

$$\gamma_{t+1}^{(i)} \propto p(y_t|x_{t+1}^{(i)}) \sum_j \Pr(\text{migrating to hypothesis } j \text{ from } i) \cdot \gamma_t^{(j)}. \tag{10}$$

The minimum variance estimate of the state can then be obtained as

$$\hat{x}_t = \sum_{i=1}^N \gamma_t^{(i)} \hat{x}_t^{(i)} \tag{11a}$$

$$P_t = \sum_{i=1}^N \gamma_t^{(i)} \big(P_t^{(i)} + (\hat{x}_t^{(i)} - \hat{x}_t)(\hat{x}_t^{(i)} - \hat{x}_t)^T\big). \tag{11b}$$

In order to avoid exponential growth of hypotheses, pruning or merging techniques needs to be applied. Two common merging base methods are *generalized pseudo Bayesian* (GPB) method [12, 16] and *interacting multiple model* (IMM) method [16, 17, 18]. Both methods are filter algorithms for linear discrete-time filters with Markovian switching coefficients, and differs only in when merging

is performed. In this paper a pruning algorithm is used where the least probable branches in the hypotheses tree are cut off to keep the number of parallel hypotheses at the desired level [12, 19].

# 3   Statistical Properties

The Cramér-Rao lower bound (bounding the MSE) and the Kullback divergence will be used to evaluate the estimated posterior distributions from filters.

## 3.1   Cramér-Rao Lower Bound (CRLB)

The *Cramér-Rao lower bound* (CRLB), [20, 21, 22], offers a fundamental performance bound for unbiased estimators. For instance, the CRLB can be used for feasibility tests or to measure filter efficiency in combination with MSE.

The CRLB along a given state trajectory for a system with Gaussian noise can in principle be found as

$$\mathsf{cov}(x_t - \hat{x}_{t|t}) \succeq P_{t|t}, \tag{12}$$

where $P_{t|t}$ is the CRLB, given by the EKF, around the true state, $x_t$. If process/measurement noise is non-Gaussian slight modifications to this approach are needed. More information about the CRLB and its extension to dynamic systems for the posterior CRLB can be found in [6, 15, 22].

## 3.2   Kullback Divergence (KD)

The *Kullback-Leibler information* (KLI) [23, 24] quantifies the difference between two distributions. The KLI is not symmetric in its arguments, and hence not a measure. If symmetry is needed the *Kullback divergence* (KD), constructed as a symmetric sum of two KLI [24, 25], can be used as an alternative.

The KLI is defined, for two proper PDFs $p$ and $q$, as

$$\mathcal{I}^{\mathrm{KL}}(p,q) = \int p(x) \log \frac{p(x)}{q(x)} \, dx, = \mathsf{E}_p \log \frac{p(x)}{q(x)} \tag{13a}$$

when $p(x) \neq 0 \Rightarrow q(x) \neq 0$, otherwise $\mathcal{I}^{\mathrm{KL}}(p,q) = +\infty$. It can be shown that $\mathcal{I}^{\mathrm{KL}}(p,q) \geq 0$ for all proper PDFs $p$ and $q$ and that $\mathcal{I}^{\mathrm{KL}}(p,q) = 0 \Leftrightarrow p = q$. A small $\mathcal{I}^{\mathrm{KL}}(p,q)$ indicates that the distribution $p$ is similar to $q$. For $p(x_1, x_2) = p_1(x_1)p_2(x_2)$ and $q(x_1, x_2) = q_1(x_1)q_2(x_2)$,

$$\mathcal{I}^{\mathrm{KL}}(p,q) = \mathcal{I}^{\mathrm{KL}}(p_1, q_1) + \mathcal{I}^{\mathrm{KL}}(p_2, q_2)$$

That is, new independent observations just add to the total information available, [24].

The symmetric KD is defined as

$$\mathcal{J}^{\mathrm{K}}(p,q) = \mathcal{I}^{\mathrm{KL}}(p,q) + \mathcal{I}^{\mathrm{KL}}(q,p), \tag{13b}$$

and the KLI properties above hold.

The KLI is closely related to other statistical measures, *e.g.*, *Shannon's information* and *Akaike's information criterion* [25]. A connection to *Fisher information* can also be found [24], and KLI can be used to derive the EM-algorithm [26]. In [27], the use of information bounds for dynamic systems discussed in quite general terms.

The KD evaluates any PDF against for instance the true posterior PDF. The true PDF can in simulations be provided by a fine gridded PMF. This way a measure of the quality of an estimator can be obtained.

**Example (comparing two Gaussian distributions):**

Assume $p_i = \mathcal{N}(\mu_i, \Sigma_i)$, $i = 1, 2$ ($\mu_i$ and $\Sigma_i$ scalar):

$$\mathcal{I}^{\mathrm{KL}}(p_1, p_2) = \mathsf{E}_{p_1} \log \frac{p_1}{p_2} = \frac{1}{2} \log \frac{\Sigma_2}{\Sigma_1} - \frac{1}{2} + \frac{\Sigma_1 + (\mu_1 - \mu_2)^2}{2\Sigma_2}.$$

For a different mean only ($\Sigma_1 = \Sigma_2 =: \Sigma$) this yields

$$\mathcal{I}^{\mathrm{KL}}(p_1, p_2) = 0 - \frac{1}{2} + \frac{\Sigma + (\mu_1 - \mu_2)^2}{2\Sigma} = \frac{(\mu_1 - \mu_2)^2}{2\Sigma},$$

where the main component is the normalized squared difference in mean. For equal mean $\mu_1 = \mu_2$, but different variance

$$\mathcal{I}^{\mathrm{KL}}(p_1, p_2) = \frac{1}{2} \log \frac{\Sigma_2}{\Sigma_1} - \frac{1}{2} + \frac{\Sigma_1}{2\Sigma_2} = \frac{1}{2} \left( \frac{\Sigma_1}{\Sigma_2} - 1 - \log \frac{\Sigma_1}{\Sigma_2} \right).$$

Here, only the relative difference in variance, $\Sigma_1/\Sigma_2$, has any significance. ◇

# 4 Probability Calculations and Hypothesis Testing

Often, just getting a point estimate of the state of a system is not enough. Utilizing the full information in the estimated posterior state distribution, it is possible to determine how likely an event, $\mathcal{A}(x_t)$, is by integrating the PDF:

$$\Pr\left(\mathcal{A}(x_t)\right) = \int_{\left\{x_t : \mathcal{A}(x_t)\right\}} p(x_t | \mathbb{Y}_t) \, dx_t. \tag{14}$$

This is straightforward to compute from the PMF and the PF, whereas the EKF often calls for numerical methods.

Another application is to decide between different possible hypotheses, $\mathcal{H}_0$ and $\mathcal{H}_1$. For this, one method is to use *Bayes factor* [28, 29],

$$B_{01}^{\pi}(\mathbb{Y}_t) = \frac{\Pr(\mathcal{H}_0 | \mathbb{Y}_t)}{\Pr(\mathcal{H}_1 | \mathbb{Y}_t)} \left/ \frac{\pi(\mathcal{H}_0)}{\pi(\mathcal{H}_1)} \mathop{\gtrless}_{\mathcal{H}_0}^{\mathcal{H}_1} k, \right. \tag{15}$$

where $\pi$ is the prior for the hypothesis. The threshold $k$ should be chosen to obtain an acceptable compromise between low risk of false alarms and good detection rate. Usually $k = 1$ is a good choice.

To perform any of the tasks above a good knowledge of the posterior state distribution is needed, and the result depends on the level of precision in the approximations made.

# 5  Simulations

In the simulation study two different examples are studied to highlight the performance gain obtained from estimating the full PDF instead of just first and second order moments. Also, the estimated PDFs are used to determine the probability for the state to be in a given region. Several estimators are used: EKF, MMF (only Example I), PF, and PMF (representing the truth). The MMF uses a pruning algorithm, where the least probable branches in the hypotheses tree are cut off, after the introduction of the measurement information, to keep the number of parallel hypotheses at the desired level [12, 19].

In the Monte Carlo simulation studies, the MSE is compared to the parametric CRLB. Furthermore, the KD, between the true state distribution (from the fine gridded PMF) and the distributions provided by the filters, are compared to capture more of the differences not seen in the second-order moment.

## 5.1  Example I — Multi-Modal Posterior

In this example, the process noise is given by a bimodal Gaussian mix, comprising possible target maneuvers. The measurement noise is given by a quite uninformative Gaussian, *i.e.* with a large variance. The posterior PDF tends to a multimodal Gaussian mix, with approximately four distinct peaks, for the selected system:

$$x_{t+1} = 0.4x_t + w_t \tag{16a}$$

$$y_t = x_t + e_t \tag{16b}$$

with $w_t \sim \frac{1}{2}\mathcal{N}(-1, 0.1^2) + \frac{1}{2}\mathcal{N}(1, 0.1^2)$, $e_t \sim \mathcal{N}(0, 2^2)$, and $x_0 \sim \mathcal{N}(0, 0.1)$. The system has a clearly non-Gaussian posterior state distribution. See Fig. 1 for a typical example, where the PDF is estimated from several different filters.

The four different filters, EKF, MMF ($2^3 = 8$ hypothesis, *i.e.*, 3 correct time steps), PF (SIR with 1000 particles), and PMF (representing the true PDF), have been used to track the state of the system. With 100 Monte Carlo simulations the results in Fig. 2(a) were obtained. As can be seen, the more advanced methods do not yield any noticeable improvement in terms of the MSE, which here matches the *best linear unbiased estimate* (BLUE) performance for all filters even though the CRLB is much lower. However, instead studying the estimated PDFs using KD shows a clear difference in performance. This is shown in Fig. 2(b), where the KD is used to compare the estimates provided by the different filtering methods to the true $p(x_t|\mathbb{Y}_t)$. Note how the PF and MMF, both allowing for non-Gaussian posterior distributions, are better than the KF. In this case the MMF approximates the true PDF very well with 8 Gauss modes, as seen in Fig. 1, and the performance is therefore better than for the PF.

## 5.2  Example II — Range-Only Measurement

In a range-only measurement application, two range sensors are used. The measurements are illustrated in Fig. 3(a). They provide the relative range to
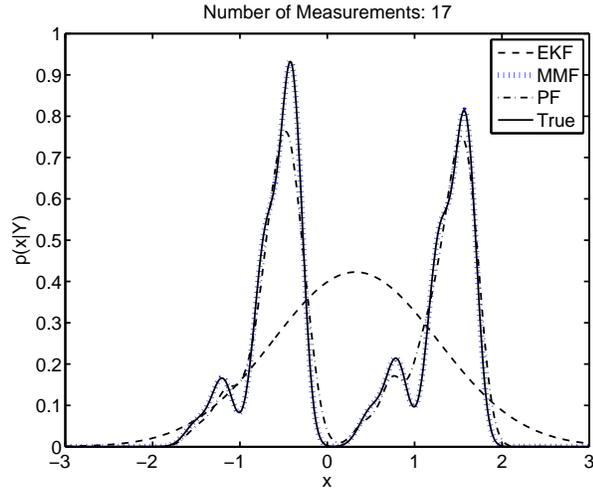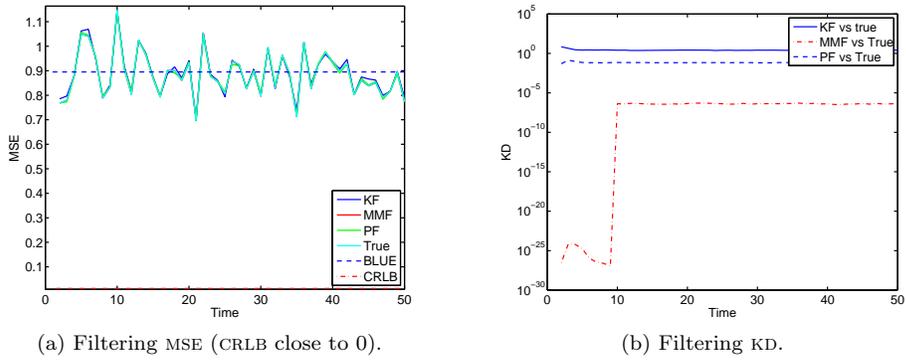
Fig. 1: Example I. Typical posterior state distribution for (16). (The MMF coincides with the true PDF, given by a PMF.)



(a) Filtering MSE (CRLB close to 0).



(b) Filtering KD.

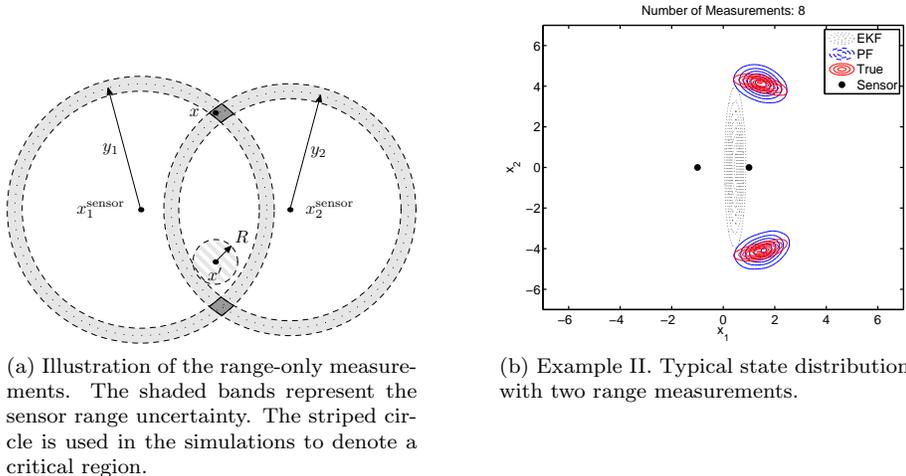Fig. 2: Example I. Simulated MSE and KD for (16).

(a) Illustration of the range-only measurements. The shaded bands represent the sensor range uncertainty. The striped circle is used in the simulations to denote a critical region.



(b) Example II. Typical state distribution with two range measurements.

Fig. 3: Example II. Scenario and typical PDF.

an unknown target, with measurement uncertainty. Hence, producing a natural bimodal posterior distribution. The model used is:

$$x_{t+1} = x_t + w_t \tag{17a}$$

$$y_t = \begin{pmatrix} \|x_t - x_1^{\text{sensor}}\| \\ \|x_t - x_2^{\text{sensor}}\| \end{pmatrix} + e_t, \tag{17b}$$

with $w_t \sim \mathcal{N}(0, 0.1I_2)$, $e_t \sim \mathcal{N}(0, 0.1I_2)$, and initial knowledge $x_0 \sim \mathcal{N}(0, 3I_2)$.

A typical state distribution is given in Fig. 3(b). Note the distinct bimodal characteristics of the distribution, as well as how poorly the EKF approximation describes the situation.

The MSE and KD from 100 Monte Carlo simulations are given in Fig. 4 for an EKF, a PF (SIR with 20 000 particles[1]), and a PMF (regarded as the truth). Here, the MSE performance of the PF is slightly better than for the EKF, but not as good as the PMF. (Two poor PF estimates has a large impact on the MSE.) However, the KD gives a clear indication that the PDF estimated with the PF is better than the one the EKF, as was to be expected due to the true bimodal PDF. This difference may be important, as will be shown next.

Using the estimated PDFs, it is also possible to detect if the tracked target is in the neighborhood of the point $x'$ (not affecting the measurements), as illustrated in Fig. 3(a). Assume that it is of interest to detect if $\|x - x'\|_2 < R$, where $x' = (0, 1)$ and $R = 0.5$. The probability for the target in the critical region, given the estimates from the different filters is given in Fig. 5. Note that the EKF throughout the whole simulation indicates a much higher probability

---

[1]The number of particles has been chosen excessively large to get an accurate PDF estimate, with further tuning could be reduced significantly.
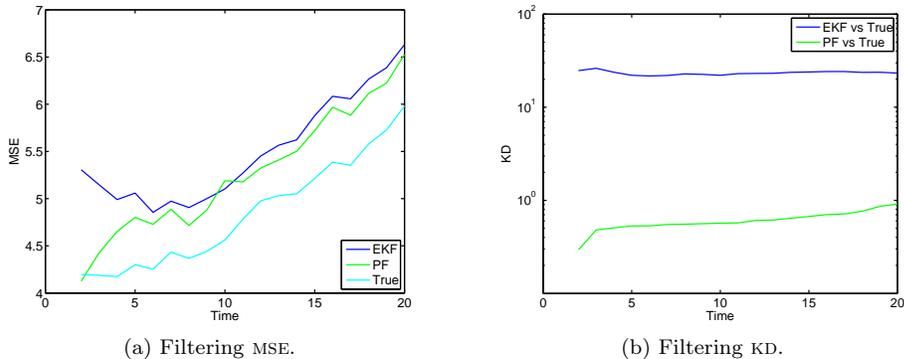
| (a) Filtering MSE. | (b) Filtering KD. |

Fig. 4: Example II. Simulated MSE and KD for the range-only system.

than the true situation. Furthermore, the PF reflects the actual situation well. The lack of descriptive power of the EKF results in an unnecessary high degree of detections, which could be costly.

# 6    Conclusions

In extensive simulation studies the Kullback divergence is shown to indicate the performance gain for estimators such as the MMF and the PF with more accurate PDF estimates compared to more traditional KF/EKF. It is shown that just the second-order moment can be quite misleading as a performance measure. In for instance hypothesis testing accurate estimation of the entire PDF is important.

# References

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

[2] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice-Hall, Inc, Englewood Cliffs, NJ, 1979.

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

[4] H. W. Sorenson and D. L. Alspach, "Recursive Bayesian estimation using Gaussian sums," *Automatica*, vol. 7, no. 4, pp. 465–479, July 1971.

[5] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to non-linear/non-Gausian Bayesian state estimation," *IEE Proc.-F*, vol. 140, no. 2, pp. 107–113, Apr. 1993.
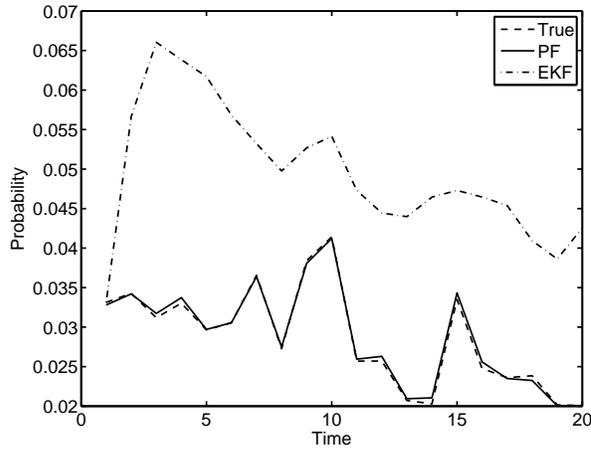
Fig. 5: Probability of the target belonging to the critical region given by $\|x - (0,1)\|_2 < 0.5$, see Fig. 3(a).

[6] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*, Statistics for Engineering and Information Science. Springer-Verlag, New York, 2001.

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

[8] G. Hendeby, R. Karlsson, F. Gustafsson, and N. Gordon, "Performance issues in non-Gaussian filtering problems," in *Proc. Nonlinear Statistical Signal Processing Workshop*, Cambridge, UK, Sept. 2006.

[9] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME*, vol. 82, no. Series D, pp. 35–45, Mar. 1960.

[10] H. W. Sorenson, "Recursive estimation for nonlinear dynamic systems," in *Bayesian Analysis of Time Series and Dynamic Models*, J. C. Spall, Ed., pp. 126–165. Dekker, 1988.

[11] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[12] F. Gustafsson, *Adaptive Filtering and Change Detection*, John Wiley & Sons, Ltd, Chichester, West Sussex, England, 2000.

[13] S. C. Kramer and H. W. Sorenson, "Bayesian parameter estimation," *IEEE Trans. Automat. Contr.*, vol. 33, no. 2, pp. 217–222, Feb. 1988.

[14] H. Tanizaki, "Nonlinear and nonnormal filters using Monte Carlo methods," *Computational Statistics and Data Analysis*, vol. 25, pp. 417–439, 1997.

[15] N. Bergman, *Recursive Bayesian Estimation: Navigation and Tracking Applications*, Dissertations no 579, Linköping Studies in Science and Technology, SE-581 83 Linköping, Sweden, May 1999.

[16] H. A. P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. Automat. Contr.*, vol. 33, no. 8, pp. 780–783, Aug. 1988.

[17] H. A. P. Blom, "An efficient filter for abruptly changing systems," in *Proc. 23rd IEEE Conf. Decis. and Contr*, Las Vegas, NV, USA, 1984, pp. 656–658.

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

[19] G. Hendeby, *Fundamental Estimation and Detection Limits in Linear Non-Gaussian Systems*, Lic. thesis no 1199, Dept. Electr. Eng, Linköpings universitet, Sweden, Nov. 2005.

[20] H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, Princeton, NJ, 1946.

[21] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, vol. 1, Prentice-Hall, Inc, 1993.

[22] E. L. Lehmann, *Theory of Point Estimation*, Probability and Mathematical Statistics. John Wiley & Sons, Ltd, 1983.

[23] S. Kullback, J. C. Keegel, and J. H. Kullback, *Topics in Statistical Information Theory*, vol. 42 of *Lecture Notes in Statistics*, Springer-Verlag, 1987.

[24] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Statist.*, vol. 22, no. 1, pp. 79–86, Mar. 1951.

[25] C. Arndt, *Information Measures*, Springer-Verlag, 2001.

[26] S. Gibson and B. Ninness, "Robust maximum-likelihood estimation of multivariable dynamic systems," *Automatica*, vol. 41, pp. 1667–1682, 2005.

[27] J. Bröcker, "On comparing nonlinear filtering algorithms," in *Proc. 2005 Int. Symp. Nonlin. Theory App.*, Bruges, Belgium, Oct. 2005.

[28] C. P. Robert, *The Bayesian Choise: From Decision-Theoretic Foundation to Computational Implementation*, Springer texts in Statistics. Springer-Verlag, 2 edition, 2001.

[29] I. J. Good, "Significance test in parallel and in series," *JASA*, vol. 53, no. 284, pp. 799–813, Dec. 1958.

**Titel**
Title

Performance Issues in Non-Gaussian Filtering Problems

**Författare**
Author

Gustaf Hendeby, Rickard Karlsson, Fredrik Gustafsson, Neil Gordon

**Sammanfattning**
Abstract

Performance for filtering problems is usually measured using the second-order moment. For non-Gaussian applications, this measure is not always sufficient. In this paper, the Kullback divergence is extensively used to compare estimated distributions. Several estimation techniques are compared, and methods with ability to express non-Gaussian posterior distributions are shown to give superior performance over classical second-order moment based estimators.

**Nyckelord**
Keywords      Kullback divergence, Particle filtering