On Performance Measures for Approximative Parameter Estimation

Gustaf Hendeby, Fredrik Gustafsson

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

Report no.: LiTH-ISY-R-2608
Submitted to Reglermöte 2004

Abstract

The Kalman filter computes the minimum variance state estimate as a linear function of measurements in the case of a linear model with Gaussian noise processes. There are plenty of examples of non-linear estimators that outperform the Kalman filter when the noise processes deviate from Gaussianity, for instance in target tracking with occasionally maneuvering targets. Here we present, in a preliminary study, a detailed analysis of the well-known parameter estimation problem. This time with Gaussian mixture measurement noise. We compute the discrepancy of the best linear unbiased estimator (BLUE) and the Cramér-Rao lower bound, and based on this conclude when computationally intensive Kalman filter banks or particle filters may be used to improve performance.

Keywords: Parameter estimation; Linear estimation; Maximum likelihood estimators; Model approximation; Performance analysis

ON PERFORMANCE MEASURES FOR APPROXIMATIVE PARAMETER ESTIMATION

Gustaf Hendeby * **Fredrik Gustafsson** *

* *Division of Automatic Control and Communication Systems*
Department of Electrical Engineering, Linköpings universitet
SE-581 83 LINKÖPING, SWEDEN
`{hendeby, fredrik}@isy.liu.se`

Abstract: The Kalman filter computes the minimum variance state estimate as a linear function of measurements in the case of a linear model with Gaussian noise processes. There are plenty of examples of non-linear estimators that outperform the Kalman filter when the noise processes deviate from Gaussianity, for instance in target tracking with occasionally maneuvering targets. Here we present, in a preliminary study, a detailed analysis of the well-known parameter estimation problem. This time with Gaussian mixture measurement noise. We compute the discrepancy of the best linear unbiased estimator (BLUE) and the Cramér-Rao lower bound, and based on this conclude when computationally intensive Kalman filter banks or particle filters may be used to improve performance.

Keywords: Parameter estimation; Linear estimation; Maximum likelihood estimators; Model approximation; Performance analysis

1. INTRODUCTION

For a linear state space model

$$\begin{aligned} x_{t+1} &= Ax_t + w_t \\ y_t &= Cx_t + e_t, \end{aligned}$$

the Kalman filter is the optimal state estimator in the following senses:

- It is optimal when the initial state, w_t and e_t are Gaussian processes.
- It is the optimal linear estimator independent of higher order noise moments.

However, the Bayesian framework yields the true posterior state distribution, where the conditional expectation is the minimum variance estimate. Note that generally, the estimator becomes a non-linear function of data.

The estimate can be approximated with several well-known techniques, ranging from point-mass filters (Bucy and Senne, 1971), spline filters (de Figueiredo and Jan, 1971), Kalman filter banks (Blom and Bar-Shalom, 1988) to the more recent and quite general particle filter (Doucet *et al.*, 2001). The central question is whether there is anything to gain, and if so, how much better estimates can be expected?

The case when the measurement noise is bi-Gaussian (a Gaussian mixture with two components) will be analyzed. This kind of noise has turned out to be a good model of outliers in many applications. One example is described in detail in (Bergman *et al.*, 1999) where the radar error for altitude measurements has a certain probability to be reflected in the tree tops. The distribution for different terrain types is demonstrated in (Dahlgren, 1998) to be well approximated with Gaussian mixtures. System performance gained a lot by including this empirical noise model in the filter (Bergman, 1999). Another example is the forward collision avoidance system for cars reported in (Jansson *et al.*, 2004), where reflections in wheel house and rear view mirror are argued to give Gaussian mixtures, in radar range errors.

The best linear unbiased estimator (BLUE) (Kay, 1993) will be compared, with the conditional expectation in terms of Cramér-Rao Lower Bound (CRLB) on the variance, for the parameter in bi-Gaussian noise problem. In this case the BLUE is the same as the Kalman filter estimate which makes this estimate easy to obtain.

2. PARAMETER ESTIMATION

Consider the following simplification of the general dynamic system

$$y_i = x + e_i, \tag{1}$$

where x is to be determined from the noisy measurements y_i which are distorted by the uncorrelated noise e_i. This is in fact standard parameter estimation.

2.1 Posterior Distributions and Estimators

The complete description of what is known about the parameter, given some measurements, is gathered in the pdf $p(x|\mathbb{Y}_t)$. Here $\mathbb{Y}_k$ denote the k first measurements $\{y_i\}_{i=1}^k$. Furthermore, *Bayes' Rule* gives this expressed in measurement probabilities

$$p(x|\mathbb{Y}_k) = \frac{p(\mathbb{Y}_k|x)p(x)}{p(\mathbb{Y}_k)}, \tag{2}$$

where $p(\mathbb{Y}_k)$ is used to normalize the pdf.

From this, two important estimators of x can be constructed, the *minimum variance estimator*

$$\hat{x}_p^{\mathsf{MV}}(k) = \mathsf{E}_{p(x|\mathbb{Y}_k)}\, x \tag{3}$$

and the *maximum likelihood estimator*

$$\hat{x}_p^{\mathsf{ML}}(k) = \arg\min_x p(x|\mathbb{Y}_k). \tag{4}$$

Both of these will be further discussed below. However, first some special cases.

Gaussian Noise First assume that all e_i are identically distributed Gaussian noise, *viz.*

$$e = e_i \in \mathcal{N}(\mu, R). \tag{5}$$

Using this assumption calculate the probability for a certain measurement,

$$p(y|x) = \mathcal{N}(y; \mu + x, R), \tag{6}$$

and if generalized to k measurements on the same system

$$p(\mathbb{Y}_k|x) = \prod_{i=1}^k \mathcal{N}(y_i; \mu + x, R) = \\ = \bar{p}(\mathbb{Y}_k) \cdot \mathcal{N}(x; \bar{\mu}(\mathbb{Y}_k), \bar{R}) \tag{7}$$

where

$$\bar{R} = R/k$$
$$\bar{\mu}(\mathbb{Y}_k) = \sum_i (y_i - \mu)/k$$

and $\bar{p}(\mathbb{Y}_k)$ is a function independent of x.

The parameter posterior pdf is then

$$p(x|\mathbb{Y}_k) = \bar{p}(\mathbb{Y}_k)\mathcal{N}(x; \bar{\mu}(\mathbb{Y}_k), \bar{R})p(x)/p(\mathbb{Y}_k) = \\ = \mathcal{N}(x; \bar{\mu}(\mathbb{Y}_k), \bar{R}) \tag{8}$$

where the last equality follows by normalization if the non-informative (improper) prior $p(x) = 1$ is used.

The pdf is the well known Gaussian distribution, and the maximum likelihood estimate therefore coincides with the minimum variance estimate

$$\hat{x}_p^{\mathsf{ML}}(k) = \hat{x}_p^{\mathsf{MV}}(k) = \bar{\mu}(\mathbb{Y}_k) = \sum_i (y_i - \mu)/k. \tag{9}$$

Furthermore, it is easily shown that this is the Kalman filter solution and that it is unbiased with variance

$$\mathsf{Var}\,\hat{x}(k) = \mathsf{Var}(e)/k = R/k. \tag{10}$$

Gaussian Mixture Noise Now assume that the measurement noise is not Gaussian, as so often assumed, but instead a Gaussian mixture

$$e_i \in \sum_\delta p_\delta \cdot \mathcal{N}(\mu_\delta, R_\delta) \text{ with } \sum_\delta p_\delta = 1. \tag{11}$$

This is the same as to say that e_i is chosen from the Gaussian distribution $\mathcal{N}(\mu_\delta, R_\delta)$ with probability p_δ, where δ is referred to as the mode parameter. This kind of measurement noise is common *e.g.*, in some radar applications as indicated by (Dahlgren, 1998; Bergman *et al.*, 1999), and can be used to approximate any distribution p_e according to (Alspach and Sorenson, 1972; Anderson and Moore, 1979).

The distribution in (11) has the following statistical properties:

$$\mu = \sum_\delta p_\delta \mu_\delta \tag{12a}$$

$$R = \sum_\delta p_\delta \big(R_\delta + \bar{\mu}_\delta^2\big) \tag{12b}$$

$$\gamma_1 = \sum_\delta p_\delta \bar{\mu}_\delta \big(3R_\delta + \bar{\mu_\delta}^2\big) \cdot R^{-\frac{3}{2}} \tag{12c}$$

$$\gamma_2 = \sum_\delta p_\delta \big(3R_\delta^2 + 6\bar{\mu}_\delta^2 R_\delta + \bar{\mu_\delta}^4\big) \cdot R^{-2} - 3, \tag{12d}$$

where $\bar{\mu}_\delta = \mu_\delta - \mu$. Above, μ is the regular mean, R is the variance, γ_1 is the skewness, and γ_2 the kurtosis. In the Gaussian case $\gamma_1 = \gamma_2 = 0$.

The probability for a given sequence of measurements, $\mathbb{Y}_k$, is given by

$$p(\mathbb{Y}_k|x, \delta^k) = \prod_{i=1}^k \mathcal{N}(y_i; \mu_{\delta_i^k} + x, R_{\delta_i^k}) = \\ = \bar{p}(\mathbb{Y}_k, \delta^k) \cdot \mathcal{N}\left(x; \bar{\mu}(\mathbb{Y}_k, \delta^k), \bar{R}(\delta^k)\right) \tag{13}$$

with $\delta^k := (\delta_1^k, \dots, \delta_k^k)$ an ordered k-tuple indicating the mode used in each measurement,

$$\bar{R}(\delta^k) = \left(\sum_{i=1}^k R_{\delta_i^k}^{-1}\right)^{-1}$$

$$\bar{\mu}(\mathbb{Y}_k, \delta^k) = \bar{R}(\delta^k) \sum_{i=1}^k R_{\delta_i^k}^{-1}(y_i - \mu_{\delta_i^k}),$$

and $\bar{p}(\mathbb{Y}_k, \delta_i^k)$ is a factor independent of x. Summing up all mode combinations results in the following likelihood for the measurements conditioned on x (*cf.* (7) for the Gaussian case),

$$p(\mathbb{Y}_k|x) = \sum_{\delta^k} p(\delta^k)\, p(\mathbb{Y}_k|x, \delta^k) \tag{14}$$

with the probability for a certain noise mode sequence

$$p(\delta^k) = \prod_{i=1}^{k} p_{\delta_i^k}.$$

Reversing the conditioning to get an expression for the parameter given the measurements yields, using *Bayes' rule* (*cf.* (8) for Gaussian measurements),

$$p(x|\mathbb{Y}_k) = p(\mathbb{Y}_k|x)p(x)/p(\mathbb{Y}_k) \tag{15}$$

where, if choosing the non-informative prior $p(x) = 1$,

$$p(\mathbb{Y}_k) = \sum_{\delta^k} \bar{p}(\mathbb{Y}_k, \delta^k)p(\delta^k). \tag{16}$$

Note that even though (15) provides complete statistical knowledge of the parameter, it is due to the exponential growth of mode combinations impossible to in practise make use of all this information. Nevertheless, is it of interest to see what the complete information has to offer.

Based on (15) the minimum variance estimator is calculated to be

$$\hat{x}_p^{\mathsf{MV}}(k) = \sum_{\delta^k} p(\delta^k)\bar{\mu}(\mathbb{Y}_k, \delta^k), \tag{17}$$

the weighted mean found in (13). The maximum likelihood estimate is not as easily derived in closed form, and it is often necessary to resort to numerical methods to find

$$\hat{x}_p^{\mathsf{ML}}(k) = \arg\min_x p(x|\mathbb{Y}_k). \tag{18}$$

3. CRAMÉR-RAO LOWER BOUND

The Cramér-Rao Lower Bound (CRLB) is a well known lower bound for the variance of any estimates achieved by an unbiased estimator, under weak regularity conditions (Kay, 1993, Theorem 3.1). These conditions are easily shown to hold both in the Gaussian and in the Gaussian mixture setting discussed in Sec. 2.1. The CRLB is an often used performance measure, and is therefore of interest here.

The CRLB for an estimate based on k measurements from system (1) is

$$P_{\text{CRLB}}(k) := \frac{-1}{\mathsf{E}_{p(\mathbb{Y}_k|x)} \frac{\partial^2}{\partial x^2} \log p(\mathbb{Y}_k|x)}, \tag{19}$$

evaluated for the true parameter value.

Given the uncorrelated property of the measurement noise and the invariance of the system, the denominator can be rewritten as

$$\mathsf{E}_{p(\mathbb{Y}_k|x)} \frac{\partial^2}{\partial x^2} \log p(\mathbb{Y}_k|x) = \\ = k\, \mathsf{E}_{p(y|x)} \frac{\partial^2}{\partial x^2} \log p(y|x), \tag{20}$$

still evaluated for the true parameter value. This shows that

$$P_{\text{CRLB}}(k) = \frac{P_{\text{CRLB}}(1)}{k}. \tag{21}$$

Note that the value decreases as $1/k$ due to the independence between the measurements and the the lack of parameter dynamics.

It is possible to show (see *e.g.*, (Kay, 1993)) that the maximum likelihood estimate, $\hat{x}_p^{\mathsf{ML}}(k)$, as k tends to infinity achieves the variance given by the CRLB. It should, however, be noted that the maximum likelihood estimate is in general not unbiased for finite number of observations.

3.1 Gaussian Noise

Performing the calculations in (20) for Gaussian measurements (5) yields,

$$P_{\text{CRLB}}(k) = \frac{R}{k}, \tag{22}$$

and the variance of an estimate of x in this setting is thus bounded from below by the variance of the measurements.

3.2 Gaussian Mixture Noise

Trying to calculating the CRLB for the case of Gaussian mixture noise (11) yields for the second order derivative in (20)

$$\frac{\partial^2}{\partial x^2} \log p(y|x) = \\ = \frac{\frac{\partial^2}{\partial x^2} p(y|x) \cdot p(y|x) - \left(\frac{\partial}{\partial x} p(y|x)\right)^2}{p^2(y|x)} = \\ = \frac{\sum_\delta p(\delta)p(y|x,\delta)\Big(\big(\frac{y-x-\mu_\delta}{R_\delta}\big)^2 - \frac{1}{R_\delta}\Big)}{p(y|x)} + \\ + \frac{-\Big(\sum_\delta p(\delta)p(y|x,\delta)\frac{y-x-\mu_\delta}{R_\delta}\Big)^2}{p^2(y|x)}. \tag{23}$$

Inserted into (19) this results in a much more complicated expression than the Gaussian counterpart in (22). It is impossible to calculate this expectancy value analytically in closed form. However, using Monte Carlo Integration (Robert and Casella, 1999) it is possible to numerically approximate the measure.

4. NON-LINEAR PERFORMANCE GAIN

The expression for the variance of an estimator based on Gaussian noise found in Sec. 2.1 together with expression for the CRLB for the Gaussian mixture noise found in Sec. 3.2 provide an opportunity to quantify the performance gained by using the a Gaussian mixture measurement noise instead of approximating it with a single Gaussian. Approximating the noise and then applying the Kalman filter yields the best linear unbiased estimator (BLUE). Knowing this, combining (10) and (21) yields the ratio

$$\frac{\operatorname{Var}\hat{x}^{\text{BLUE}}_{p^N}(k)}{P_{\text{CRLB}}(k)} = \frac{\operatorname{Var} e}{P_{\text{CRLB}}(1)} = \frac{R}{P_{\text{CRLB}}(1)} \tag{24}$$

which hence relate the CRLB to the BLUE variance with a constant scale factor. Note that this constant can be computed given just the measurement error pdf. It is therefore possible to compute tables with the performance loss for standard distributions.

This section provides three such tables (actually contour plots) relating two sets of separated bi-Gaussian distributions to their Gaussian approximation and an example relating to outliers. Without loss of generality it is enough to consider only distributions that are approximated by $\mathcal{N}(0,1)$. If needed it is always possible to scale and/or add a mean to this.

4.1 *Bi-Gaussian Noise with Equal Weights*

The first evaluated mixture is

$$e \in \tfrac{1}{2}\mathcal{N}(\mu,R) + \tfrac{1}{2}\mathcal{N}(\mu',R'), \tag{25}$$

where μ and R are parameters, and μ' and R' are chosen to achieve zero mean and unit variance (*cf.* (11) with $p_\delta = \frac{1}{2}$). Use (12) to relate this distribution to statistical properties such as skewness and kurtosis.

The result of the comparison is shown in Fig. 1. As can be seen in the figure, with parameter values in the center of the plot, and a quite large region around it, the BLUE performance is close to the CRLB. The gain from more advanced estimators is hence negligible there.

Observe, that $P_{\text{CRLB}} = 0$ is the boundary of the feasible parameter region. Right on the boundary one, or both, Gaussian components have degenerated to point distributions. The result is quite intuitive, since with a point distribution the knowledge of the system is very high.

One special case of (25) is when the pdf is symmetric, *viz.*, $R = R'$. When this occurs is indicated in Fig. 1 with a dashed line and this part of the plot is then displayed in Fig. 2. Note in the figure that the modes must be well separated before this seriously affects the BLUE solution.

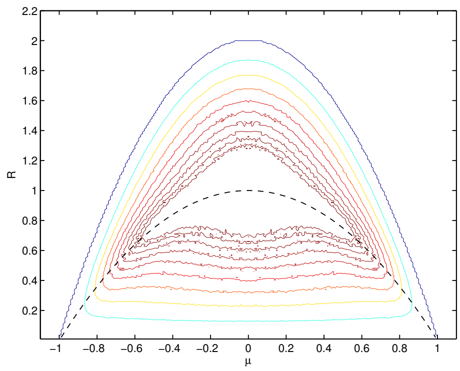

Fig. 1. The CRLB for the bi-Gaussian distribution (25) with equal weights. (Levels: [0.99, 0.98, 0.97, 0.95, 0.92, 0.87, 0.78, 0.64, 0.40 0], 0 being the outermost level.) As comparison using a Gaussian approximation will yield unit variance. Also included a dashed line indicating where the pdf is symmetric.

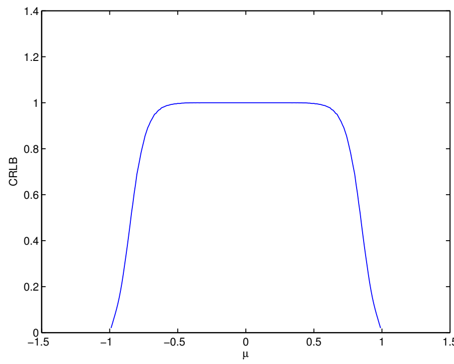

Fig. 2. Symmetric intersection of Fig. 1. The symmetric bi-Gaussian is $p_e(x) = \frac{1}{2}\mathcal{N}(x;\mu,1-\mu^2) + \frac{1}{2}\mathcal{N}(x;-\mu,1-\mu^2)$.

4.2 *Bi-Gaussian Noise with 9:1 Weights*

In the previous section a Gaussian mixture with two equally important Gaussians were analyzed, in this section a bi-Gaussian mixture where 90% of the probability is contained in one Gaussian is studied. This is the result if

$$e \in \tfrac{9}{10}\mathcal{N}(\mu,R) + \tfrac{1}{10}\mathcal{N}(\mu',R'). \tag{26}$$

Once again μ and R are parameters whereas μ' and R' are used to achieve zero mean and unit variance of e (*cf.* (11) and (25)).

The CRLB with the noise in (26) is presented in Fig. 3. Note the similarity between between Fig. 1 and 3, but also that a larger region in the parameter space would favor from using the true p_e. That is, the effect of a smaller disturbance to the measurements is in some sense more serious and gain more from proper modeling.

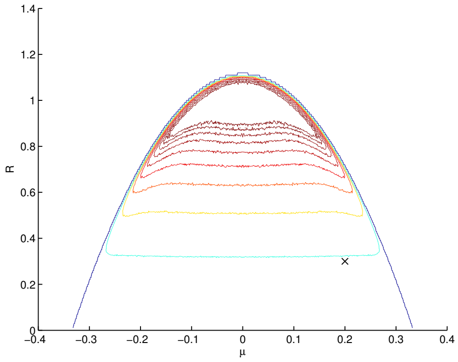

Fig. 3. The CRLB for the bi-Gaussian mixture distribution in (26). (Levels: [0.99, 0.98, 0.97, 0.95, 0.92, 0.87, 0.78, 0.64, 0.40 0], 0 being the outermost level.) Note that a Gaussian approximation yields unit variance. ($\times$ denotes the distribution simulated in Sec. 5.)

4.3 *Uni-Modal Bi-Gaussian Noise Modelling Outliers*

In this section the attention is pointed towards the case with one mode, but heavier tails than is to be expected from a Gaussian distribution. One way to model this is with

$$e \in (1-p)\mathcal{N}(0,R') + p\mathcal{N}(0,kR'), \qquad (27)$$

where p and k is parameters and R' is used to obtain unit variance. One interpretation of this is that the first part is the expected measurement and the second part describes outliers. The risk of an outlier is then p and the outliers are distributed with k times larger variance than the proper measurements. For this distribution the skewness is $\gamma_1 = 0$ due to symmetry and the kurtosis is

$$\gamma_2 = \frac{3(k-1)^2 p(1-p)}{(1+(k-1)p)^2} \geq 0 \text{ for } p \in [0,1]. \qquad (28)$$

Note that this distribution by construction is sub-Gaussian, which shows in the non-negative kurtosis.

Fig. 4 shows how the CRLB is affected by outliers. As seen in the figure, relatively few outliers degrade the performance of the approximative filter substantially, especially if the outliers have high variance.

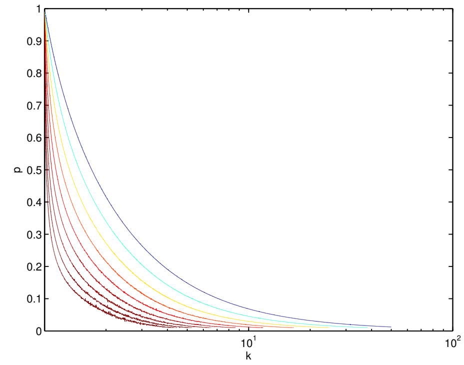

Fig. 4. The CRLB for the outlier situation. (Levels: [0.99, 0.98, 0.97, 0.95, 0.92, 0.87, 0.78, 0.64, 0.40 0], 0 being the uppermost level.) As a comparison, the BLUE yields unit variance for the estimator.

The measurement noise used for the approximation, e_i^N, then becomes,

$$e_i^N \in \mathcal{N}(\mathrm{E}\, e_i, \mathrm{Var}\, e_i) = \mathcal{N}(0,1), \qquad (30)$$

this corresponds to the analysis in Sec. 4.2. Furthermore, from here on the superscript ^{N} will denote a quantity originating in this approximation.

Given the system above, either look up the CRLB in Fig. 3 (at the $\times$) or compute the CRLB to be $P_{\text{CRLB}}(1) \approx 0.37$, and $\mathrm{Var}\, \hat{x}_{p^N}^{\text{BLUE}}(1) = 1$. Hence, the Kalman filter variance correlates to the CRLB as $\mathrm{Var}\, \hat{x}_{p^N}^{\text{BLUE}}(k) \approx 2.7 P_{\text{CRLB}}(k)$ independently on the number of measurements used. Thus, using the approximative measurement noise increases the variance of the estimate 2.7 times compared to if an optimal filter is used. Moreover, we know that $\mathrm{Var}\, \hat{x}_p^{\text{ML}}$ eventually reaches the CRLB given enough measurements. The latter estimate therefore outperforms the former given enough measurements.

The system above has been evaluated using 1000 Monte Carlo simulations. Each simulation involves twelve measurements, and the different parameter are calculated for one through twelve measurements. The exponential complexity of the problem makes it impossible to simulate many more measurements. The results are studied in the sequel.

5. SIMULATIONS

In this section the results from Sec. 4 will be verified using simulations. To do this measurements according to (1) is generated with Gaussian mixture noise (11). From these measurements the parameter is estimated both using the true noise and a Gaussian approximation (5). The latter estimate coinciding with the BLUE.

Let the true measurement noise be defined by

$$e_i \in \underbrace{0.9 \cdot \mathcal{N}(0.2, 0.3)}_{p_0 \cdot \mathcal{N}(\mu_0, R_0)} + \underbrace{0.1 \cdot \mathcal{N}(-1.8, 3.7)}_{p_1 \cdot \mathcal{N}(\mu_1, R_1)}. \qquad (29)$$

5.1 *Posterior pdfs*

In Fig. 5, the pdfs $p(x|\mathbb{Y}_k)$ and $p^N(x|\mathbb{Y}_k)$ are plotted for one realization of measurements. Note that the true pdf differs in the behavior compared to approximated. As could be expected, the extra weight in one tail makes the true pdf less tempted to be affected by an outlier. For example see the changes from $k=2$ to $k=3$ and from $k=10$ to $k=11$ in Fig. 5. In this particular realization the difference between the pdfs is notable. This difference may be especially important if the result is to be used for statistical decision making.

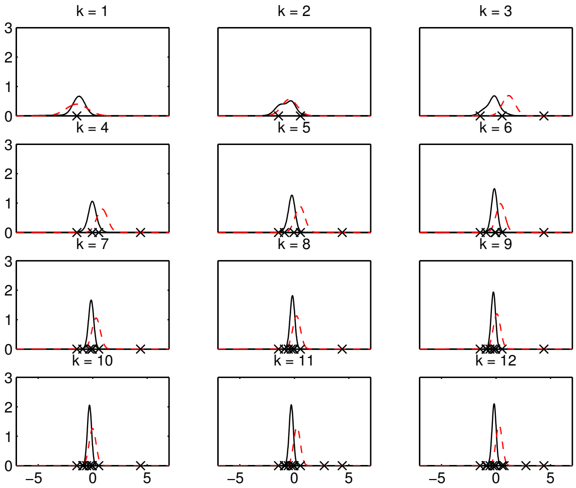

Fig. 5. Comparison of the parameter pdfs achieved using true and approximated measurement noise for k measurements. (Solid line true pdf, dashed line approximated pdf, and $\times$ denoting measurements.)

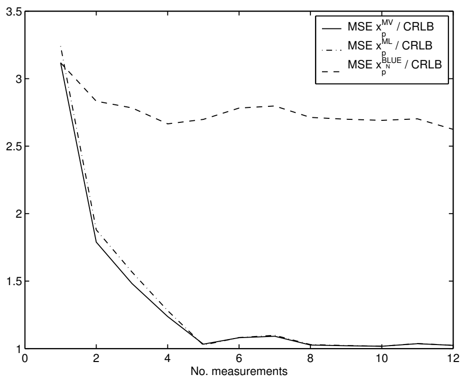

Fig. 6. Comparison of the ratios between MSE and CRLB for the different estimation schemes. (1000 Monte Carlo simulations.)

5.2 Variance Behavior

The ratios between the obtained MSE (approximating the estimator variance) and the CRLB for the estimators discussed are visualized in Fig. 6. The result coincides with the theoretically predicted results.

The ratio between CRLB and MSE $\hat{x}_{pN}^{\text{BLUE}}$ (the Kalman filter variance estimate) is somewhat unsteady around 2.7 as can be seen in Fig. 6. This well coincides with the value theoretically derived earlier. Statistical fluctuations should account for this unsteadiness.

From Fig. 6 it further seems to be true that $\hat{x}_p^{\text{ML}}$ tends to the CRLB quite quickly. The minimum variance estimate based on the true distribution follows the same pattern, and seems to approach the CRLB at approximately the same rate as the maximum likelihood estimate. This shows that there is precision to gain from using the true distributions as compared to a crude Gaussian approximation in this case. How much is indicated by the ratio $\text{Var}\, e/P_{\text{CRLB}}(1)$, in this case 2.7. If this ratio is close to one, the approximation is probably valid.

6. CONCLUSIONS

In this paper, the effect of approximating a Gaussian mixture noise (11) with Gaussian noise (5), in a parameter in noise setting (1), is investigated; both using theoretically analysis and simulated data. A theoretical performance loss, in terms of how much larger the best unbiased linear estimator (BLUE) variance is than the CRLB, is derived. It turns out that, using precompiled tables it is possible to in advance decide whether a certain approximation is good enough. If there is a large difference between CRLB and approximative estimate variance this is an indication that using more elaborate, and computationally intensive, methods pay off. This result is also exemplified and verified using simulated data.

ACKNOWLEDGEMENTS

This work is supported by the competence center ISIS (Information Systems for Industrial Control and Supervision) at Linköpings universitet.

REFERENCES

Alspach, D. L. and H. W. Sorenson (1972). Recursive Bayesian estimation using Gaussian sum. *IEEE Trans. Automat. Contr.* **17**(4), 439–448.

Anderson, B. D. O. and J. B. Moore (1979). *Optimal Filtering*. Prentice Hall. Englewood Cliffs, NJ.

Bergman, N. (1999). Recursive Bayesian Estimation: Navigation and Tracking Applications. Ph.D. thesis no 579. Linköping Studies in Science and Technology. SE-581 83 Linköping, Sweden.

Bergman, N., L. Ljung and F. Gustafsson (1999). Terrain navigation using Bayesian statistics. *IEEE Control Syst. Mag.* **19**(3), 33–40.

Blom, H. A. P. and Y. Bar-Shalom (1988). The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Trans. Automat. Contr.* **33**(8), 780–783.

Bucy, R. S. and K. D. Senne (1971). Digital synthesis of non-linear filters. *Automatica* **7**(3), 287–298.

Dahlgren, C. (1998). Nonlinear black box modelling of JAS 39 Gripen's radar altimeter. Master's thesis n:o LiTH-ISY-EX-1958. Dept. E.E., Linköpings universitet, Sweden.

de Figueiredo, R. J. P. and Y. G. Jan (1971). Spline filters. In: *Proc. 2nd symp. on nonlinear estimation theory and its applications*. pp. 127–141.

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

Jansson, J., R. Karlsson and F. Gustafsson (2004). Model-based statistical tracking and decision making for collision avoidance application. In: *Accepted for: American Control Conference*. Boston.

Kay, S. M. (1993). *Fundamentals of Statistical Signal Processing: Estimation Theory*. Vol. 1. Prentice Hall, Inc.

Robert, C. P. and G. Casella (1999). *Monte Carlo Statistical Methods*. Springer Texts in Statistics. Springer-Verlag. New York.