

Recursive Triangulation Using Bearings-Only Sensors

Gustaf Hendeby, Rickard Karlsson, and

Fredrik Gustafsson

Div. Automatic Control

Linköpings universitet, Linköping, Sweden

{hendeby, rickard, fredrik}@isy.liu.se

Neil Gordon

Track & Sensor Fusion, DSTO

Edinburgh, Australia

Abstract— Recursive triangulation, using a bearings-only sensor, is investigated for a fly-by scenario. In a simulation study, several estimators are compared, fundamental estimation limits are calculated for different measurement noise assumptions. The quality of the estimated state distributions is evaluated.

I. INTRODUCTION

Passive ranging or bearings-only tracking is a common technique utilized in many applications. Many standard filtering algorithms have problems representing range or range rate uncertainty when the aim is to recursively triangulate objects using only bearings measurements. Especially for targets located close to the sensor or passing close and fast, where the sensor quadrant is changing rapidly. We discuss possible methods and associated problems using a benchmark example. Methods based on linearization as the *extended Kalman filter* (EKF, [1–3]) and also the *unscented Kalman filter* (UKF, [4]) have difficulties to solve this problem when the initial linearization error is large. A comparison with the *particle filter* (PF, [5–7]), and fundamental estimation bounds represented by the *Cramér-Rao lower bound* (CRLB) are given. Performance is discussed using information theoretical measures, such as *intrinsic accuracy* (IA, [8–10]), *relative accuracy* (RA), and *Kullback divergence*, [11, 12].

The problem of recursive triangulation occurs in classical bearings-only tracking using a passive radar sensor, but is likewise important in emerging applications that use vision systems for tracking. For instance, situational awareness in automotive collision avoidance systems is often based on *infrared* (IR) sensors, *laser radars* (LIDAR), and/or vision systems. In common for the sensors then used is accurate angle measurements, but poor or no range information. Thus, sensor fusion is often used to incorporate information from inertial measurements.

Applications and topics under investigation in this paper, theoretically and/or in simulation studies are:

- (i) A stationary sensor and an object passing fast and close (so that the bearing change is about π rad in a few measurement cycles), *e.g.*, a sonar buoy listening for passing maritime vessels.
- (ii) A multiple sensor scenario, *i.e.*, sensor networks, utilizing sensor fusion to gain range information.
- (iii) The impact on the estimation error of the distance from the sensor to the target.

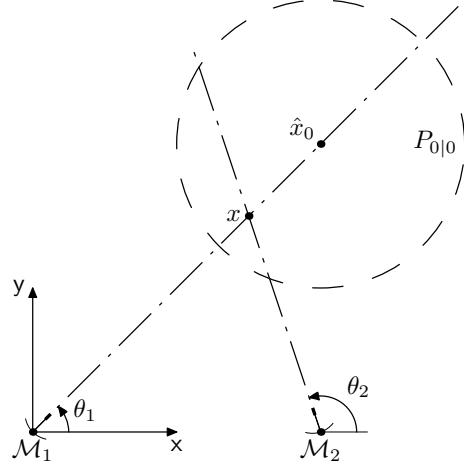


Fig. 1: Bearings-only problem, with two measurements M_1 and M_2 . The true target location x and initial estimate \hat{x}_0 , with covariance P_0 .

Here, these problems are dealt with using the simplified setup in Fig. 1, where the target is stationary and measurements are taken from different positions. This does not limit the analysis to stationary target, since target movements can be achieved moving the sensors. However, the dynamics of the system is neglected to simplify the analysis. This is motivated when the uncertainty between two measurements is small compared to other errors. Measuring a stationary target with a mobile sensor with a well working positioning system is one such case.

Usually, measurement noise is, implicitly or explicitly, considered to be Gaussian. This noise description is sometimes quite limiting, and this paper therefore analyzes the effects of non-Gaussianity for estimation using the generalized Gaussian distribution and Gaussian mixtures. Both these have the Gaussian distribution as a special case.

The paper is organized as follows: In Sec. II information theoretical bounds and definitions are introduced. Sec. III describe the bearings-only sensor model and the used measurement noise distributions. In Sec. IV Monte Carlo simulation studies are compared to fundamental limits. Finally, in Sec. V some concluding remarks are given.

II. INFORMATION THEORETICAL BOUNDS

The *Fisher information* (IA) and the *Cramér-Rao lower bound* (CRLB), discussed in this section, offers a fundamental performance bound for unbiased estimators, *e.g.*, to be used for feasibility tests or to measure filter efficiency. The common Gaussian noise assumption yields the least favorable bound. Hence, it is important to compare the information for the given distribution against the Gaussian case. The difference is expressed in *intrinsic accuracy* and *relative accuracy*.

A. Cramér-Rao lower bound (CRLB)

For estimation of a parameter, x , from samples, y_i , of a distribution, in which x is a parameter, the CRLB limits the performance obtainable for any unbiased estimator:

$$\text{var}(\hat{x}) \succeq P, \quad (1)$$

where the P is the CRLB, under mild regularity conditions, given by

$$P^{-1} = -\mathbb{E}_y \Delta_x^x \log p(y|x) \quad (2)$$

evaluated for the true parameter value [13].

Assuming $y_i = h_i(x) + e_i$ and independent measurements $e_i \sim p_{e_i}$ yields for $\mathbb{Y}_n = \{y_1, \dots, y_n\}$ (see the Appendix for a derivation)

$$\begin{aligned} P_n^{-1} &= -\mathbb{E}_{\mathbb{Y}_n} \Delta_x^x \log p(\mathbb{Y}_n|x) \\ &= \sum_1^n \nabla_x h_i(x) (-\mathbb{E}_{e_i} \Delta_{e_i}^{e_i} \log p_{e_i}(e_i)) \nabla_x^T h_i(x). \end{aligned} \quad (3)$$

The terms in the sum are denoted *Fisher information* (FI), and describe how much information each measurement carries. Note that for independent noise new measurements just add to the total information available. Furthermore, the expectation term depends solely on the distribution of the measurement noise, and is hence a noise property. More information about the CRLB and its extension to dynamic systems can be found, *e.g.*, [6, 14, 15].

B. Intrinsic Accuracy

The expectation term in (3) is basically the *Fisher information* of the variable e_i , with respect to the expected value, $\mu_i = \mathbb{E} e_i$. In [8–10] this quantity is referred to as the *intrinsic accuracy* (IA) of the PDF for e_i . This name is inspired by the terminology used in [16] and IA is a measure of the information content in a noise. The IA is defined as

$$\mathcal{I}_e = -\mathbb{E}_e \Delta_e^e \log p_e(e|\mu), \quad (4)$$

evaluated for the true mean of the distribution.

The Gaussian distribution is the distribution with the lowest IA of all distributions with the same covariance [17]. Thus, in this sense, the worst case distribution.

The quantity *relative accuracy* (RA) is introduced to denote the relative difference in IA between a distribution and its second order Gaussian equivalent. If a scalar Ψ_e exists such that $\text{cov}(e) = \Psi_e \mathcal{I}_e^{-1}$, denote Ψ_e the RA for the distribution. It follows that when RA is defined, $\Psi_e \geq 1$,

with equality if and only if e is Gaussian. The RA thus is a measure of how much useful information there is in the distribution, compared to a Gaussian distribution with the same covariance.

C. Kullback Divergence

The *Kullback-Leibler information* [11, 12] quantifies the difference between two distributions. The Kullback-Leibler information is not symmetric in its arguments, and hence not a measure. If a measure is needed, the *Kullback divergence*, constructed as a symmetric sum of two Kullback-Leibler informations [12, 18], can be used as an alternative.

The Kullback-Leibler information is defined, for the two proper PDFs p and q , as

$$\mathcal{I}^{\text{KL}}(p, q) = \int p(x) \log \frac{p(x)}{q(x)} dx, \quad (5a)$$

when $p(x) \neq 0 \Leftrightarrow q(x) \neq 0$, otherwise $\mathcal{I}^{\text{KL}}(p, q) = +\infty$. The Kullback divergence is defined as

$$\mathcal{J}^K(p, q) = \mathcal{I}^{\text{KL}}(p, q) + \mathcal{I}^{\text{KL}}(q, p). \quad (5b)$$

A small value of $\mathcal{J}^K(p, q)$ indicates that p and q are similar, and a large value that they are easy to tell apart.

The Kullback-Leibler information is closely related to other statistical measures, *e.g.*, *Shannon's information* and *Akaike's information criterion* [18]. A connection to the FI can also be found [12]. Both the Kullback-Leibler information and the Kullback divergence are additive for independent variables.

III. SENSOR MODEL

In this section the measurement model is discussed in detail for the bearings-only application depicted in Fig. 1, where passive sensors measure the direction to the target. The sensor model describes the bearings measurement, θ_i , from \mathcal{M}_i , with additive measurement noise e_i , according to

$$y_i = h_i(x) + e_i = \theta_i + e_i = \arctan \left(\frac{y - y_i}{x - x_i} \right) + e_i, \quad (6)$$

where the target position is $x = (x, y)^T$ and $\mathcal{M}_i = (x_i, y_i)^T$. The measurement noise is considered to have the PDF p_e .

Based on the above sensor model, n measurements of the target and the initial information P_0^{-1} , yields according to (3), the CRLB

$$\begin{aligned} P_n^{-1} &= P_0^{-1} + \sum_1^n \nabla_x h_i(x) (-\mathbb{E}_{e_i} \Delta_{e_i}^{e_i} \log p(e_i)) \nabla_x^T h_i(x) \\ &= P_0^{-1} + \Psi_e (\text{var}(e))^{-1} \sum_1^n \begin{pmatrix} \delta_{y_i}^2 & \delta_{x_i} \delta_{y_i} \\ \delta_{x_i} \delta_{y_i} & \delta_{x_i}^2 \end{pmatrix}, \end{aligned} \quad (7)$$

where $(\delta_{x_i}, \delta_{y_i}) = (x - x_i, y - y_i)$. Note that since the measurements are scalar the RA of the measurement noise turns up as a factor in the CRLB when the initial information is disregarded.

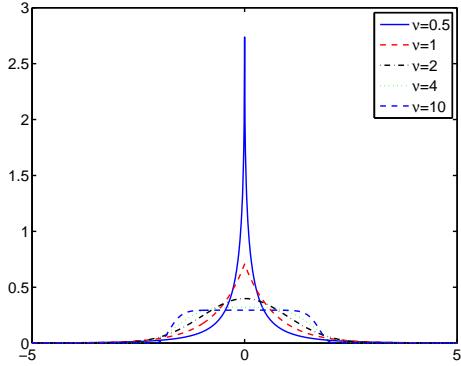


Fig. 2: PDF of the generalized Gaussian distribution as a function of the shape parameter ν .

A. Generalized Gaussian Distribution

The *generalized Gaussian* PDF is given by

$$p(e; \nu, \sigma) = \frac{\nu \eta(\nu, \sigma)}{2\Gamma(\nu^{-1})} e^{-\left(\eta(\nu, \sigma)|x|\right)^\nu}, \quad (8)$$

with

$$\eta(\nu, \sigma) = \sigma^{-1} \sqrt{\Gamma(3\nu^{-1})/\Gamma(\nu^{-1})},$$

and Γ the Gamma function [19]. The parameter ν acts as a shape parameter and σ^2 is the variance of the distribution. For $\nu \rightarrow 0^+$ the generalized Gaussian is a *Dirac distribution* and for $\nu \rightarrow +\infty$ a uniform distribution, both with variance σ^2 . This is illustrated in Fig. 2. Two important special cases are; $\nu = 1$ and $\nu = 2$, yielding the Laplacian and the Gaussian distribution, respectively.

The generalized Gaussian PDF provides means to study how small deviations from Gaussianity affect performance. Fig. 3 shows how the RA, Ψ , depends on γ .

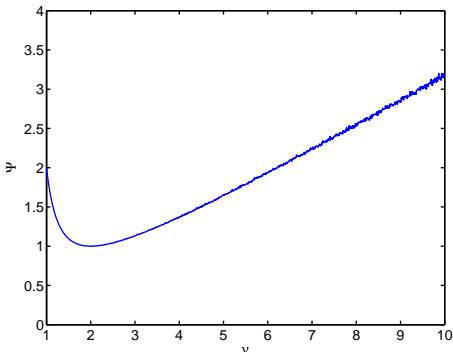


Fig. 3: RA, Ψ , of a generalized Gaussian distribution (8), parameterized in ν .

B. Gaussian Sum

Another common way to approximate non-Gaussian distributions, for instance distributions with outliers, is to use a Gaussian sum,

$$e \sim \sum_i \alpha_i \mathcal{N}(\mu_i, \sigma_i^2), \quad (9)$$

where the weights $\alpha_i > 0$ sum to unity. The Gaussian sum is very general in that it can approximate any distribution arbitrarily well if the number of terms in the sum is allowed to grow [20].

IV. SIMULATIONS

The problem of a passing target object is here reformulated to coincide with a multiple sensor fusion problem. Hence, instead of a moving target the sensor is moved, as in the studied scenario in Fig. 4. The target dynamics is therefore neglected and only measurement noise affects the result. In order to avoid studying the filter initialization very carefully an initial PDF is used. In practice, this is done by considering several measurements, so the introduced simulation study starts after the first initial transient phase.

In the studies several different estimators are used: The *extended Kalman filter* (EKF, [2, 3]) linearizes the system around the current estimate in order to utilize the Kalman filter. The *iterated EKF* (IEKF, [1, 3]) utilizes the fact that in some cases, the linearization error can be reduced by iteratively re-linearize the system. The *unscented Kalman filter* (UKF, [4, 21]), uses a small set of *sigma points* in order to approximate the second order moment correctly when nonlinearities are present. The *particle filter* (PF, [5–7]) represents a Monte Carlo based filtering method known to work well in many situations.

In the Monte Carlo simulation studies, the *root mean square error* (RMSE) is compared to the parametric CRLB. Furthermore, the Kullback divergence, between the true state distribution and the distributions provided by the estimates, is compared to capture more of the differences not seen in the second moment.

A. Simulation I: Gaussian Measurement Noise

First the system is analyzed for Gaussian measurements,

$$e_t \sim \mathcal{N}(0, \sigma^2), \quad \sigma^2 = 10^{-4}, \quad (10)$$

for the measurement relation defined in (6), as the true target position, x , is moved along the y -axis. The five measurements, from \mathcal{M}_i , triangulate the target. The true state distribution for one realization, as obtained from a point-mass filter, is given in Fig. 5 and the estimate RMSE as a function of y is given in Fig. 6(a) for 1 000 Monte Carlo simulations.

Fig. 6(a) shows that with the target located relatively close to the sensor ($y \lesssim 2.5$), the difference in performance is substantial. It can also be seen that the nonlinearity in the measurement relation is not that severe for targets located far away ($y \gtrsim 3$). The performance of the studied methods varies with the true target position, the closer to the x -axis the greater difference in performance. This is most apparent for $\theta \approx \frac{\pi}{2}$ rad, i.e., measuring from \mathcal{M}_3 .

In Fig. 6(b), performance as a function of the number of measurements is evaluated for $y = 2$. The PF clearly outperforms the other methods.

To see if the degrading performance is an effect of a few outliers all the errors in distance are plotted in a MATLAB

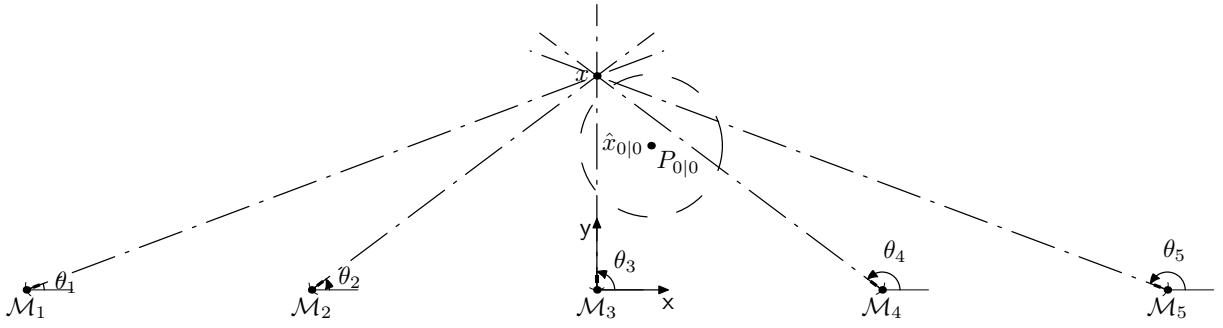


Fig. 4: The scenario, where the true target position is denoted $x = (x, y)$. The sensors, \mathcal{M}_i , are the positions where the measurements are taken and i indicates the order. $\mathcal{M}_i = (-30 + 10i, 0)^T$, for $i = 1, \dots, 5$.

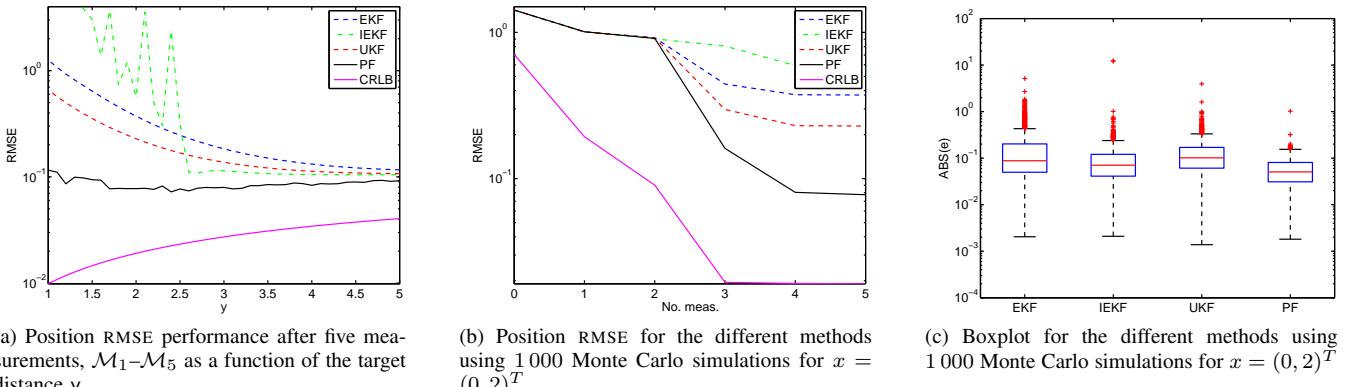


Fig. 6: Simulation I, different performance measures.

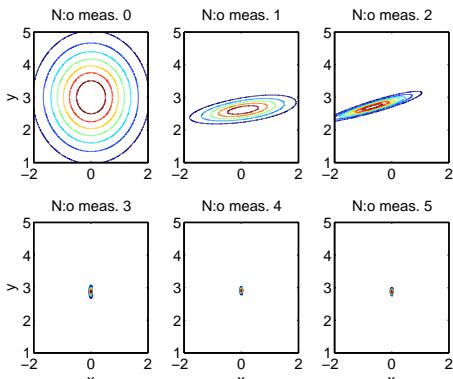


Fig. 5: True inferred position distribution for the scenario, utilizing one realization of the measurement noise to illustrate the reduction of uncertainty.

box plot. The box plot yields the median, quartile values and outliers of the errors, as seen in Fig. 6(c). See MATLAB (`help boxplot`) for details about the plot. Note how the PF is more precise than the other methods and have fewer and less severe outliers.

B. Simulation II: Outliers

Measurement outliers are a common problem in applications, *i.e.*, the sensor performs well on the average but in a few cases performance degrades considerably. Here,

a Gaussian sum is used to describe this behavior; one mode describe nominal measurements whereas another mode models the outliers. In this section the simulations in Sec. IV-A are repeated with outliers in the measurements,

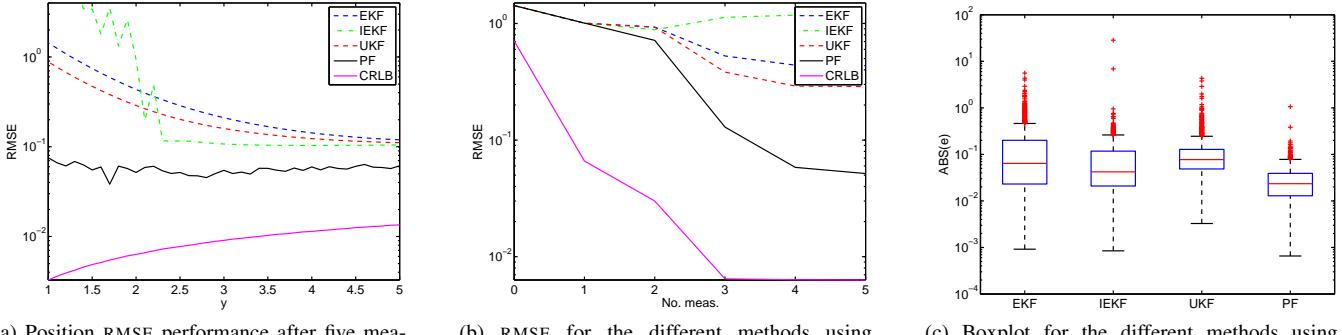
$$e \sim 0.9\mathcal{N}(0, \sigma^2) + 0.1\mathcal{N}(0, 100\sigma^2), \quad (11)$$

with Σ such that variance $\text{var}(e) = 10^{-4}$, *i.e.*, the same variance as in the previous simulations. The RA for e is $\Psi_e = 9.0$.

The results obtained using this measurement noise are found in Fig. 7. As seen in Fig. 7(a), the performance between the different estimators is rather similar, as in Simulation I. However, the PF stands out as better, hence utilizing the non-Gaussian noise better than the other methods. As before, it is rather difficult to come close the asymptotic CRLB.

C. Simulation III: Generalized Gaussian Noise

In Fig. 8, the relative CRLB is depicted for different values of ν , using the generalized Gaussian distribution. The CRLB is computed using (3) under the assumption of 10 measurements in each of the sensors according to the configuration in Fig. 4. To see the effects of how non-Gaussian noise affects the CRLB, the results are normalized to be 1 for Gaussian noise ($\nu = 2$). For the relative CRLB achieved this way, values less than 1 indicate, at least theoretical, a gain in performance for this scenario.



(a) Position RMSE performance after five measurements, $\mathcal{M}_1\text{--}\mathcal{M}_5$, as a function of the target position $x = (0, y)^T$.

(b) RMSE for the different methods using 1 000 Monte Carlo simulations for $x = (0, 2)^T$

(c) Boxplot for the different methods using 1 000 Monte Carlo simulations for $x = (0, 2)^T$

Fig. 7: Simulation II, different performance measures for measurements with outliers.

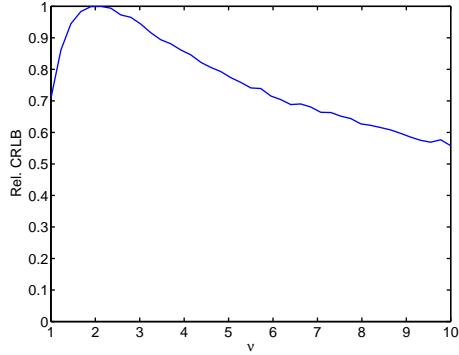


Fig. 8: Relative CRLB for the system with generalized Gaussian measurement noise, parameterized in ν , with 10 measurements from each sensor position.

It is interesting to see what performance the different estimators give, using Monte Carlo simulations. In Fig. 9, the RMSE is plotted as a function of the generalized Gaussian parameter ν . As seen, the PF and the UKF give virtually the same performance. The linearized filters, EKF and IEKF, all give much worse performance. Also note that it is hard to reach the CRLB, and the information gain indicated in Fig. 8 is not achieved. This difficulty to reach the CRLB is most likely a result of too few measurements being available to get an almost asymptotic behavior.

D. Simulation IV: Kullback Divergence

The Kullback divergence between the true state distribution, $p(x|\mathbb{Y})$, and the estimated state distribution, $\hat{p}(x|\mathbb{Y})$, from the different filters are found in Tables I and II, for the Gaussian and the outliers noise, respectively. The results presented are from one realization (depicted in Figure 4) where all the estimators work reasonably well. A reason for just using one realization is that the computations are expensive. The divergence has been computed numerically using a grid approach, where each particle has been approximated with a Gaussian with low variance compared to the total variance to get a smooth PDF. Another reason for using just one realization is that it is not obvious how to combine the

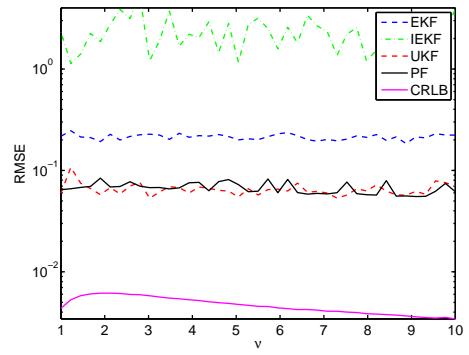


Fig. 9: RMSE for different estimators using a generalized Gaussian measurement noise.

Table I: Kullback divergence between true state distribution and estimated state distribution with Gaussian measurement noise.

Filter	Number of measurements					
	0	1	2	3	4	5
EKF	3.16	9.19	8.88	9.31	9.32	9.15
IEKF	3.16	9.25	9.06	11.68	11.62	11.61
UKF	3.16	9.19	8.87	9.31	9.33	9.15
PF	3.32	9.38	9.12	9.33	9.18	9.59

results of several simulations in a fair way. The Kullback divergence compares two distributions and is not an average performance measure.

Without measurements, the estimated state distribution is Gaussian. The Kullback divergence is non-zero due to numerical errors in the integration. The error in the PF is slightly larger because it uses a non-parametric representation of the Gaussian, where as the other filters represent it analytically.

Once the measurements arrive, the PF estimate improves in relation to the other estimates, indicating that it better captures the true nature of the state distribution. The improvement is greater for the case with outliers, indicating that the PF better handles the non-Gaussianity.

To capture the complete state distribution, not just the variance moment, is important in for instance detection

Table II: Kullback divergence between true state distribution and estimated state distribution with measurement noise affected by outliers.

Filter	Number of measurements					
	0	1	2	3	4	5
EKF	3.16	10.15	10.64	11.53	10.81	11.23
IEKF	3.16	10.12	10.40	11.55	11.14	11.61
UKF	3.16	10.15	10.62	11.53	11.14	11.63
PF	3.32	9.17	8.99	8.87	9.87	9.98

applications to make appropriate decisions.

Questions that arise from this initial Kullback divergence analysis are: How many particles are needed to approximate higher order moments in the particle filter ($N = 50\,000$ here, and simulations, not shown here, indicate further improvements if N is increased)? What is the significance of the proposal distribution?

V. CONCLUSIONS

The bearings-only problem is analyzed for a target passing close and fast to the observing sensor. Various estimators are evaluated in Monte Carlo simulation studies, and compared to fundamental theoretical estimation limits. Many of the methods based on linearization suffer severe performance losses compared to the particle filter for some of the discussed simulations. The outlier problem is discussed in detail, and the sensor modeling error is discussed in terms of the generalized Gaussian distribution. The Kullback divergence is also computed in a study and the PF is shown to better capture the inferred state distribution.

APPENDIX

The derivation of the CRLB expression (3) follows below for $y_i = h_i(x) + e_i$, where all e_i are mutually independent distributed with the PDF p_{e_i} :

$$\begin{aligned} P_n^{-1} &= -\mathbb{E}_{\mathbb{Y}_n} \Delta_x^x \log p(\mathbb{Y}_n|x) = -\mathbb{E}_{\mathbb{Y}_n} \Delta_x^x \log \prod_1^n p(y_i|x) \\ &= -\mathbb{E}_{\mathbb{Y}_n} \Delta_x^x \sum_1^n \log p(y_i|x) = -\sum_1^n \mathbb{E}_{y_i} \Delta_x^x \log p(y_i|x), \end{aligned}$$

where each term in the sum can be treated independently, due to the additivity of information. Now, look at $\Delta_x^x \log p(y_i|x)$,

$$\begin{aligned} \Delta_x^x \log p(y_i|x) &= \Delta_x^x \log p(y_i|x) = \nabla_x \nabla_x \log p(y_i|x) \\ &= \nabla_x (\nabla_{y_i-h_i(x)} \log p(y_i|x) \nabla_x (y_i - h_i(x))) \\ &= \nabla_x (\nabla_{y_i-h_i(x)} \log p(y_i|x) \nabla_x h_i(x)) \\ &= \nabla_x (\nabla_{y_i-h_i(x)} \log p(y_i|x)) \nabla_x^T h_i(x) \\ &\quad + \nabla_{y_i-h_i(x)} \log p(y_i|x) (\nabla_x \nabla_x h_i(x))^T \\ &= \nabla_x h_i(x) (\Delta_{y_i-h_i(x)}^x \log p(y_i|x)) \nabla_x^T h_i(x) \\ &\quad + \nabla_{y_i-h_i(x)} \log p(y_i|x) (\Delta_x^x h_i(x))^T \\ &= \nabla_x h_i(x) (\Delta_{e_i}^x \log p_{e_i}(e_i)) \nabla_x^T h_i(x), \end{aligned}$$

where the last equality follows from the regularity conditions, see [9]. Now, since x is given the expected value with respect to y_i is equivalent one with respect to e_i , yielding

$$P_n^{-1} = \sum_1^n \nabla_x h(x) (-\mathbb{E}_{e_i} \Delta_{e_i}^{e_i} \log p_{e_i}(e_i)) \nabla_x^T h(x),$$

for the true value of x .

REFERENCES

- [1] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, ser. Mathematics in Science and Engineering. Academic Press, Inc, 1970, vol. 64.
- [2] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, Inc, 1979.
- [3] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Prentice-Hall, Inc, 2000.
- [4] S. J. Julier and J. K. Uhlmann, “Unscented filtering and nonlinear estimation,” *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.
- [5] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, “Novel approach to nonlinear/non-Gaussian Bayesian state estimation,” *IEE Proc.-F*, vol. 140, no. 2, pp. 107–113, Apr. 1993.
- [6] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*, ser. Statistics for Engineering and Information Science. New York: Springer-Verlag, 2001.
- [7] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House, Inc, 2004.
- [8] S. M. Kay and D. Sengupta, “Optimal detection in colored non-Gaussian noise with unknown parameter,” in *Proc. IEEE Int. Conf. on Acoust., Speech, Signal Processing*, vol. 12, Dallas, TX, USA, Apr. 1987, pp. 1087–1089.
- [9] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice-Hall, Inc, 1998, vol. 2.
- [10] D. R. Cox and D. V. Hinkley, *Theoretical Statistics*. New York: Chapman and Hall, 1974.
- [11] S. Kullback, J. C. Keegel, and J. H. Kullback, *Topics in Statistical Information Theory*, ser. Lecture Notes in Statistics. Springer-Verlag, 1987, vol. 42.
- [12] S. Kullback and R. A. Leibler, “On information and sufficiency,” *Ann. Math. Statist.*, vol. 22, no. 1, pp. 79–86, Mar. 1951.
- [13] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc, 1993, vol. 1.
- [14] E. L. Lehmann, *Theory of Point Estimation*, ser. Probability and Mathematical Statistics. John Wiley & Sons, Ltd, 1983.
- [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] R. A. Fisher, “Theory of statistical estimation,” in *Proceedings of the Cambridge Philosophical Society*, vol. 22, 1925, pp. 700–725.
- [17] D. Sengupta and S. M. Kay, “Efficient estimation for non-Gaussian autoregressive processes,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 6, pp. 785–794, June 1989.
- [18] C. Arndt, *Information Measures*. Springer-Verlag, 2001.
- [19] G. E. P. Box and G. C. Tao, *Bayesian Inference in Statistical Analysis*. Addison-Wesley, 1973.
- [20] H. W. Sorenson and D. L. Alspach, “Recursive Bayesian estimation using Gaussian sums,” *Automatica*, vol. 7, no. 4, pp. 465–479, July 1971.
- [21] S. J. Julier, “The scaled unscented transformation,” in *Proc. American Contr. Conf.*, vol. 6, Anchorage, AK, USA, May 2002, pp. 4555–4559.