Assignment 1

Computational Intelligence, SS2020

Team Members		
Last name	First name	Matriculation Number
Blöcher	Christian	01573246
Bürgener	Max	01531577

1 Maximum Likelihood Estimation of Model Parameters

1.1 Which measurement in sc.2 is exponentially distributed?

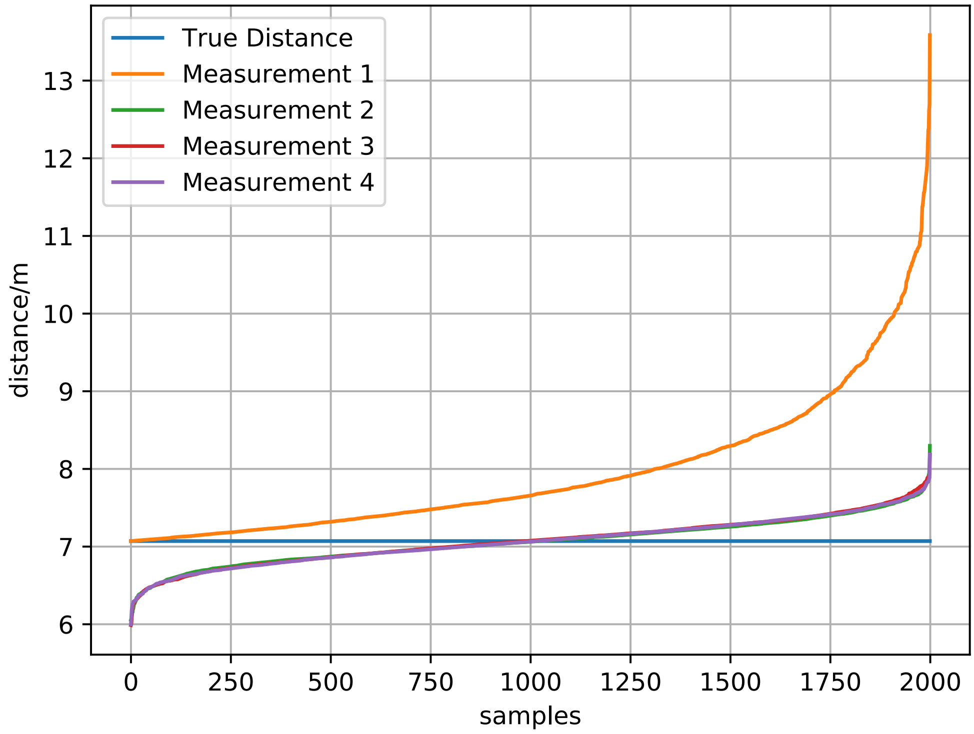


Figure 1: Scenario 2 with mixed measurement models for the anchors

- In figure 1 we can see that measurement 1 in scenario 2 is exponentially distributed. It is the only distribution which is not negative over all samples and its slope is rising exponentially for rising x-values.

1.2 Derivation for the maximum likelihood solution

- Analytical derivation for the Gaussian distribution:

$$p(\tilde{d}_n(a_i, \mathbf{p}) \mid \mathbf{p}) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{[\tilde{d}_n(a_i,\mathbf{p}) - d(a_i,\mathbf{p})]^2}{2\sigma^2}}$$

The data is independent identically distributed (iid), therefore the likelihood function is the product of all individual likelihoods

$$P(\tilde{d}_n(a_i, \mathbf{p}) \mid \mathbf{p}) = \prod_{n=0}^{N-1} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{[\tilde{d}_n(a_i,\mathbf{p}) - d(a_i,\mathbf{p})]^2}{2\sigma^2}}$$

To convert the product to a sum we apply the natural logarithm.

$$L(\tilde{d}_n(a_i,\mathbf{p}) \mid \mathbf{p}) \quad = \quad ln\left[\prod_{n=0}^{N-1} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{[\tilde{d}_n(a_i,\mathbf{p})-d(a_i,\mathbf{p})]^2}{2\sigma^2}}\right]$$

$$L(\tilde{d}_n(a_i,\mathbf{p}) \mid \mathbf{p}) \quad = \quad \sum_{n=0}^{N-1}\left[ln(\frac{1}{\sqrt{2\pi\sigma^2}}) - \frac{[\tilde{d}_n(a_i,\mathbf{p}) - d(a_i,\mathbf{p})]^2}{2\sigma^2}\right]$$

Since we want to find the parameter σ^2, which maximizes the probability of the distance, we derive $L(\tilde{d}_n(a_i,\mathbf{p}) \mid \mathbf{p})$ and set it to zero.

$$L(\tilde{d}_n(a_i,\mathbf{p}) \mid \mathbf{p}) \quad = \quad \sum_{n=0}^{N-1}\left[ln(1) - \frac{1}{2}ln(2\pi\sigma^2) - \frac{[\tilde{d}_n(a_i,\mathbf{p}) - d(a_i,\mathbf{p})]^2}{2\sigma^2}\right] \quad \mid \frac{\partial}{\partial\sigma^2}$$

$$\frac{\partial}{\partial\sigma^2}L(\tilde{d}_n(a_i,\mathbf{p}) \mid \mathbf{p}) \quad = \quad \sum_{n=0}^{N-1}\left[-\frac{1}{\sigma^2} + \frac{[\tilde{d}_n(a_i,\mathbf{p}) - d(a_i,\mathbf{p})]^2}{\sigma^4}\right] \quad \stackrel{!}{=} 0$$

$$0 \quad = \quad \sum_{n=0}^{N-1}\left[-\frac{1}{\sigma^2} + \frac{[\tilde{d}_n(a_i,\mathbf{p}) - d(a_i,\mathbf{p})]^2}{\sigma^4}\right]$$

$$\frac{N}{\sigma^2} \quad = \quad \sum_{n=0}^{N-1}\frac{[\tilde{d}_n(a_i,\mathbf{p}) - d(a_i,\mathbf{p})]^2}{\sigma^4}$$

$$\sigma^2 \quad = \quad \frac{\sum_{n=0}^{N-1}[\tilde{d}_n(a_i,\mathbf{p}) - d(a_i,\mathbf{p})]^2}{N}$$

- Analytical derivation for the Exponential distribution:

$$p(\tilde{d}_n(a_i,\mathbf{p}) \mid \mathbf{p}) = \begin{cases} \lambda_i e^{-\lambda_i[\tilde{d}_n(a_i,\mathbf{p})-d(a_i,p)]} & , \tilde{d}_n(a_i,\mathbf{p}) \geq d(a_i,\mathbf{p}) \\ 0 & , \text{else} \end{cases}$$

$$L(\tilde{d}_n(a_i,\mathbf{p}) \mid \mathbf{p}) \quad = \quad \sum_{n=0}^{N-1} ln(\lambda_i) - \lambda_i[\tilde{d}_n(a_i,\mathbf{p}) - d(a_i,\mathbf{p})] \quad \mid \frac{\partial}{\partial\lambda_i}$$

$$\frac{\partial}{\partial\lambda_i}L(\tilde{d}_n(a_i,\mathbf{p}) \mid \mathbf{p}) \quad = \quad \frac{N}{\lambda_i} - \sum_{n=0}^{N-1}[\tilde{d}_n(a_i,\mathbf{p}) - d(a_i,\mathbf{p})] \quad \stackrel{!}{=} 0$$

$$\frac{N}{\lambda_i} \quad = \quad \sum_{n=0}^{N-1}[\tilde{d}_n(a_i,\mathbf{p}) - d(a_i,\mathbf{p})]$$

$$\lambda_i \quad = \quad \frac{N}{\sum_{n=0}^{N-1}[\tilde{d}_n(a_i,\mathbf{p}) - d(a_i,\mathbf{p})]} \quad , \tilde{d}_n(a_i,\mathbf{p}) \geq d(a_i,\mathbf{p})$$

2 Estimation of the Position

2.1 Least-Squares Estimation of the Position

- Analytical conversion of the ML estimation equation:

$$
\begin{aligned}
\hat{\mathbf{p}}_{ML}(n) &= \underset{\mathbf{p}}{\operatorname{argmax}} \prod_{i=0}^{N_A-1} p(\tilde{d}_n(a_i,\mathbf{p}) \mid \mathbf{p}) \\
\hat{\mathbf{p}}_{ML}(n) &= \underset{\mathbf{p}}{\operatorname{argmax}}\, ln \left[\prod_{i=0}^{N_A-1} p(\tilde{d}_n(a_i,\mathbf{p}) \mid \mathbf{p})\right] \\
\hat{\mathbf{p}}_{ML}(n) &= \underset{\mathbf{p}}{\operatorname{argmax}} \sum_{i=0}^{N_A-1} ln\left(\frac{1}{\sqrt{2\pi\sigma_i^2}}\right) - \frac{[\tilde{d}_n(a_i,\mathbf{p}) - d(a_i,\mathbf{p})]^2}{2\sigma_i^2}
\end{aligned}
$$

Because in scenario 1 we only use Gaussian models for all anchors that were calibrated with the same distance to the reference position, we can assume that $\sigma_i^2 = \sigma^2\ \forall i$. That means the ln-term can be neglected since it only shifts the value of the maximum by a constant but does not affect its position. Similarly $\frac{1}{2\sigma^2}$ can be omitted, as it is also just a scaling factor. Furthermore $\underset{\mathbf{p}}{\operatorname{argmax}}\,(-\dots)$ is equivalent to $\underset{\mathbf{p}}{\operatorname{argmin}}\,(\dots)$. Thus:

$$
\hat{\mathbf{p}}_{ML}(n) = \underset{\mathbf{p}}{\operatorname{argmin}} \sum_{i=0}^{N_A-1} [\tilde{d}_n(a_i,\mathbf{p}) - d(a_i,\mathbf{p})]^2 = \hat{\mathbf{p}}_{LS}(n)
$$

2.2 Gauss-Newton Algorithm for Position Estimation

- Analytical solution for the Jacobian matrix.

$$
\begin{aligned}
[J(p)]_{i,1} &= \tfrac{\partial}{\partial x}\left[\tilde{d}_n(a_i,\mathbf{p}) - \sqrt{(x_i-x)^2+(y_i-y)^2}\right] \\
[J(p)]_{i,1} &= -\tfrac{[(-)2(x_i-x)]}{2\cdot\sqrt{(x_i-x)^2+(y_i-y)^2}} \\
[J(p)]_{i,1} &= \tfrac{(x_i-x)}{\sqrt{(x_i-x)^2+(y_i-y)^2}} \\
[J(p)]_{i,2} &= \tfrac{(y_i-y)}{\sqrt{(x_i-x)^2+(y_i-y)^2}}
\end{aligned}
$$

- Mean and variance

	Scenario 1	Scenario 2		Scenario 3
		with exponential anchor	without exponential anchor	
Error mean μ_e	0.278	0.640	0.399	1.265
Error variance σ_e^2	0.022	0.275	0.054	0.939

Table 1: .

Table 1 shows that the Gaus-Newton Algorithm is optimized for Gauss distributions. Mean μ_e and variance σ_e^2 is much higher in the mixed scenario 2 and scenario 3 which consists of 4 exponential distributions.

- Multivariat Gaussian distributions

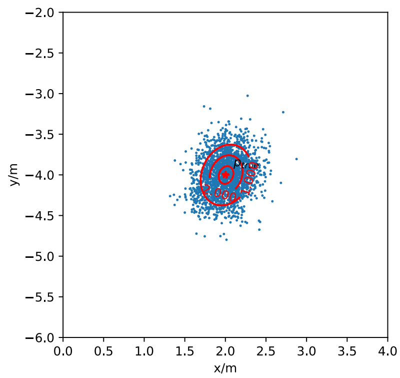


Figure 2: Gaussian distribution of scenario 1

	X	Y
X	0,04603	0,00869
Y	0,00869	0,05823

Table 2: Covariance matrix

The Gaussian distribution is computed by the Covariance matrix which is next to Figure 2. The distribution for scenario 1 is almost *spherically shaped* due to the small difference in the main diagonal and the very low values in the secondary diagonal.

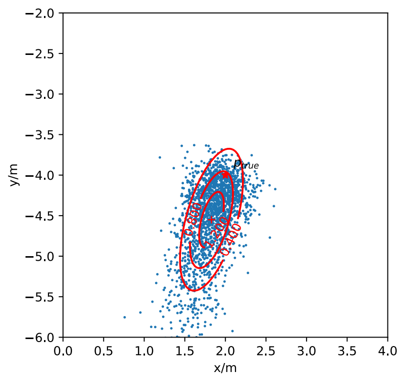


Figure 3: Gaussian distribution of scenario 2

	X	Y
X	0,05783	0,07271
Y	0,07271	0,29491

Table 3: Covariance matrix

With rising covariance values the variance around the mean value is rising. Therefore we can see a general Gaussian distribution in Figure 3. Furthermore the position estimation is not very precise in scenario 2.

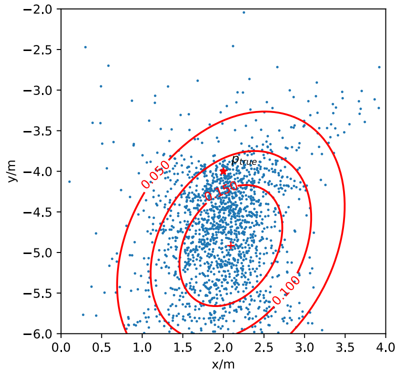


Figure 4: Gaussian distribution of scenario 3

	X	Y
X	0,73497	0,25884
Y	0,25884	0,98659

Table 4: Covariance matrix

The densitiy function of scenario 3 is also computed by a general covariance matrix with higher values in all dimensions. Therefore the point estimation not precise and the variance around the mean value is very high.

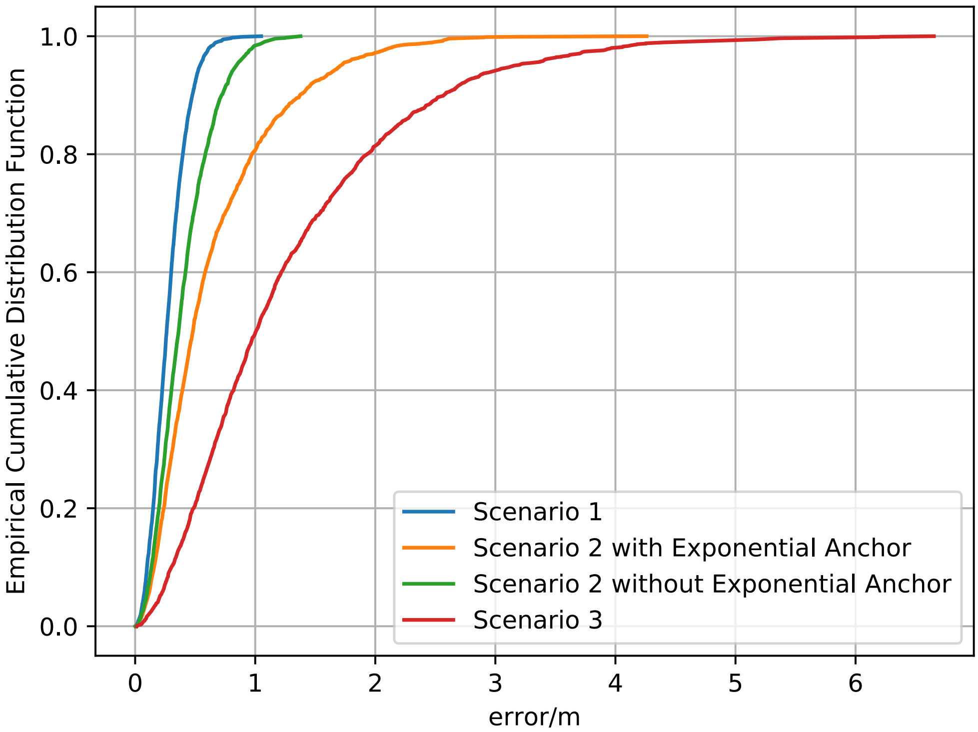


Figure 5: Gaussian distribution of scenario 3

The cumulative distribution functions of the position errors emphasize our finding that the Gauss-Newton Algorithm is optimized for Gauss distributed data sets. For every exchanged gaussian distributed anchor with an exponential distributed anchor, the results for our position estimation are getting worse in variance and error distance.

2.3 Numerical Maximum Likelihood Estimation of the Position

2.3.1 Single Measurement

- The numerical maximum likelihood estimate is computed by finding the maximum of the joint likelihood of all anchors evaluated within a 2D-grid enclosed by the anchors. Because of the i.i.d.-assumption the joint likelihood can be calculated as the product of all individual exponential likelihoods:

$$p(\tilde{\mathbf{d}}_n(\mathbf{p})|\mathbf{p}) = \begin{cases} \prod_{i=0}^{N_A-1} p(\tilde{d}_n(a_i,\mathbf{p})|\mathbf{p}), & \text{if } \tilde{d}_n(a_i,\mathbf{p}) \geq d(a_i,\mathbf{p})\ \forall i \\ 0 & \text{else} \end{cases}$$

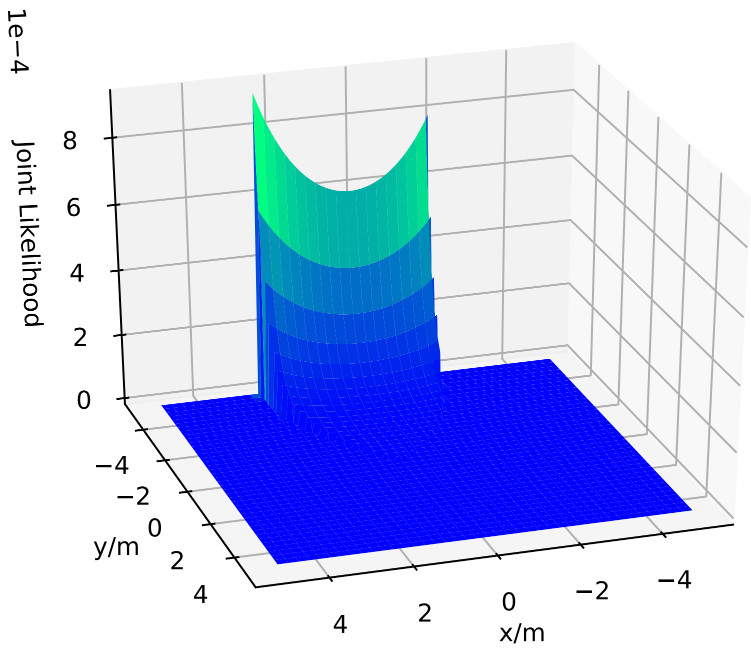


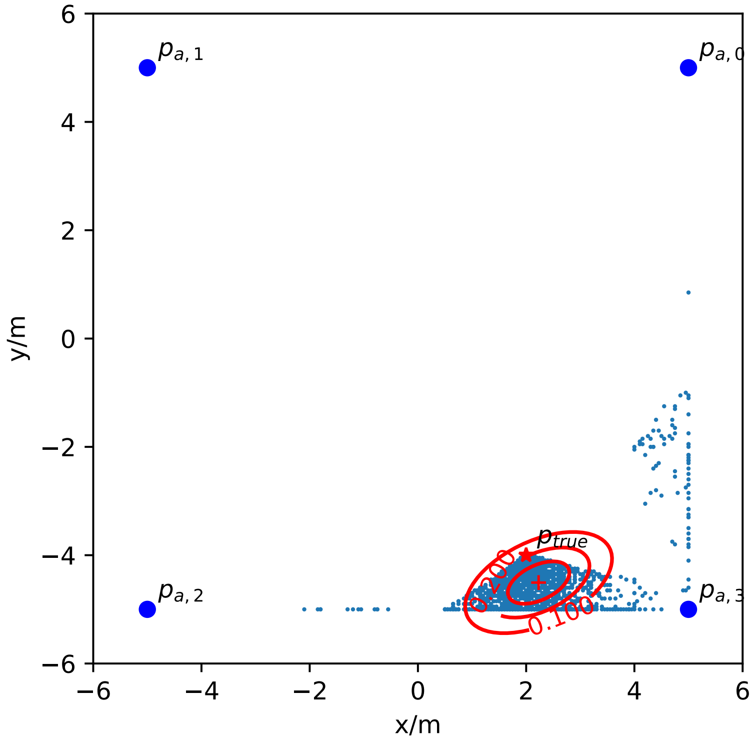
Figure 6: $p(\tilde{\mathbf{d}}_0(\mathbf{p})|\mathbf{p})$ evaluated within a 2D-grid enclosed by the anchors.

- Because of its nonlinearity the joint likelihood function of the first sample $n = 0$ has two local maxima (s. figure 6). If we used a gradient ascent algorithm with a random starting position, it would stop once it reaches any of them, possibly leading to a false estimation of the position, if the found maximum is not global.

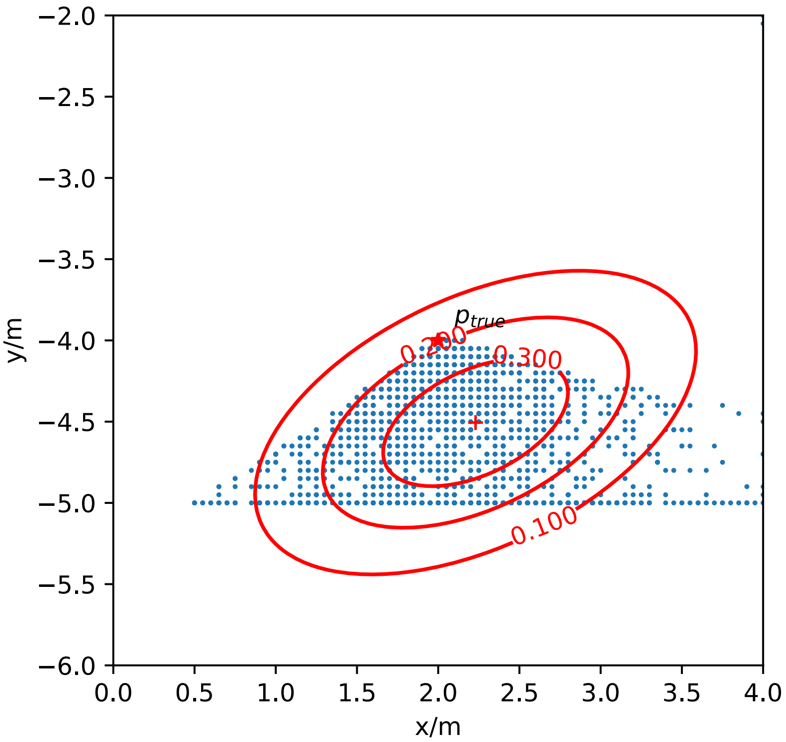
- The found maximum at $\mathbf{p}_{0,NML} = \begin{bmatrix} 2.5 \\ -5 \end{bmatrix}$ is not at the true position. **Of Course That's Because...???**

2.3.2 Multiple Measurements

-



(a) Full view of scatter plot.



(b) Detailed view centered on $\mathbf{p_{true}}$.

Figure 7: Scatter plot of NML-estimated positions with Gaussian-contour.

	Least-Squares	Numerical Maximum Likelihood	Bayes
Error mean μ_e	1.265	0.915	0.680
Error variance σ_e^2	0.939	0.489	0.095

Table 5: Error-mean and -variance of Least-Squares-, Numerical Maximum Likelihood- and Bayes-Estimation.

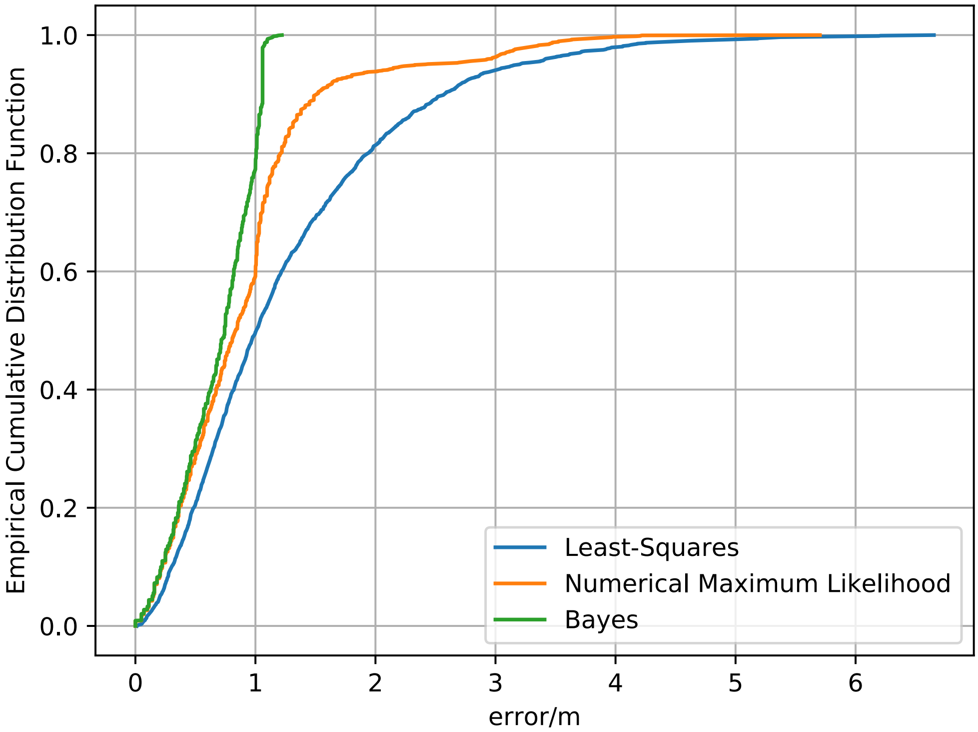
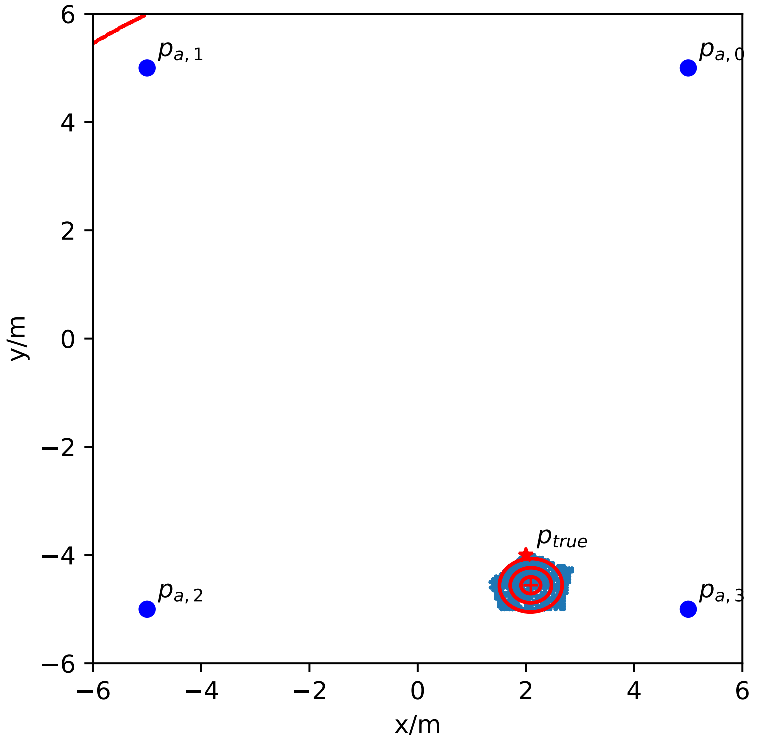
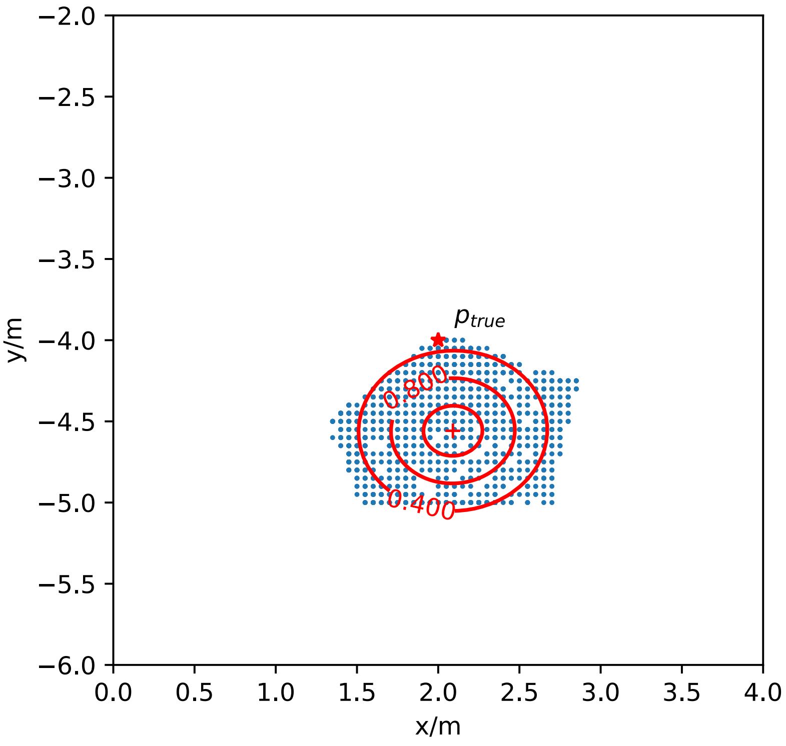


Figure 8: ECDF of Least-Squares-, Numerical Maximum Likelihood- and Bayes-Estimation.

-
-



(a) Full view of scatter plot.



(b) Detailed view centered on $\mathbf{p_{true}}$.

Figure 9: Scatter plot of Bayes-estimated positions with Gaussian-contour.