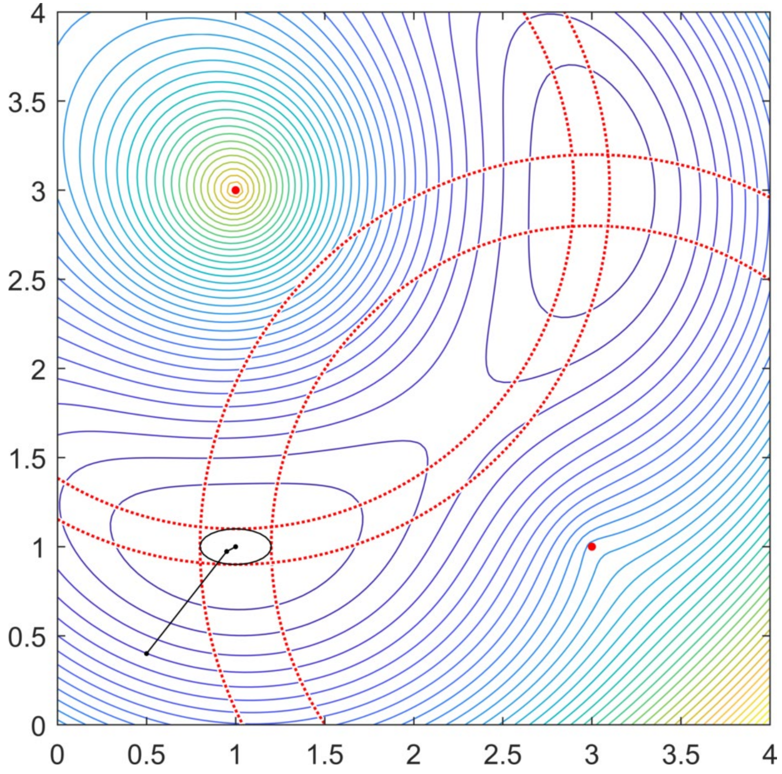
UiO : Department of Technology Systems
University of Oslo

Nonlinear MAP estimation



Trym Vegard Haavardsholm

Nonlinear MAP inference for state estimation

We want to solve **state estimation problems**
based on **measurements** and corresponding **measurement models**

Let X be the set of all unknown state variables,
and Z be the set of all measurements.

We are interested in estimating the unknown state variables X, given the measurements Z.
The **Maximum a Posteriori estimate** is given by:

$$X^{MAP} = \underset{X}{\operatorname{argmax}}\ p(X \mid Z)$$

Nonlinear MAP inference for state estimation

Measurement model:

$$\mathbf{z}_i = h_i(\underline{\mathcal{X}}_i) + \eta_i, \qquad \eta_i \sim N(\mathbf{0}, \boldsymbol{\Sigma}_i)$$

Measurement prediction function:

$$\hat{\mathbf{z}}_i = h_i(\underline{\mathcal{X}}_i)$$

Measurement error function:

$$e_i(\underline{\mathcal{X}}_i) = h_i(\underline{\mathcal{X}}_i) - \mathbf{z}_i$$

Objective function:

$$f(\underline{\mathcal{X}}) = \sum_{i=1}^{n} \left\| h_i(\underline{\mathcal{X}}_i) - \mathbf{z}_i \right\|_{\boldsymbol{\Sigma}_i}^2$$

where $\left\|\mathbf{e}\right\|_{\boldsymbol{\Sigma}}^2 = \mathbf{e}^T \boldsymbol{\Sigma}^{-1} \mathbf{e}$ is the squared Mahalanobis norm

Nonlinear MAP inference for state estimation

Measurement model:

$$\mathbf{z}_i = h_i(\underline{\mathcal{X}}_i) + \eta_i, \qquad \eta_i \sim N(\mathbf{0}, \mathbf{\Sigma}_i)$$

Measurement prediction function:

$$\hat{\mathbf{z}}_i = h_i(\underline{\mathcal{X}}_i)$$

Measurement error function:

$$e_i(\underline{\mathcal{X}}_i) = h_i(\underline{\mathcal{X}}_i) - \mathbf{z}_i$$

Objective function:

$$f(\underline{\mathcal{X}}) = \sum_{i=1}^{n} \left\| h_i(\underline{\mathcal{X}}_i) - \mathbf{z}_i \right\|_{\mathbf{\Sigma}_i}^2$$

This results in the nonlinear least squares problem:

$$\underline{\mathcal{X}}^* = \underset{\underline{\mathcal{X}}}{\operatorname{argmin}} \sum_{i=1}^{n} \left\| h_i(\underline{\mathcal{X}}_i) - \mathbf{z}_i \right\|_{\mathbf{\Sigma}_i}^2$$

It turns out that the nonlinear least squares solution to this problem is the MAP estimate!

Example:
Range-based localization

States: Our location

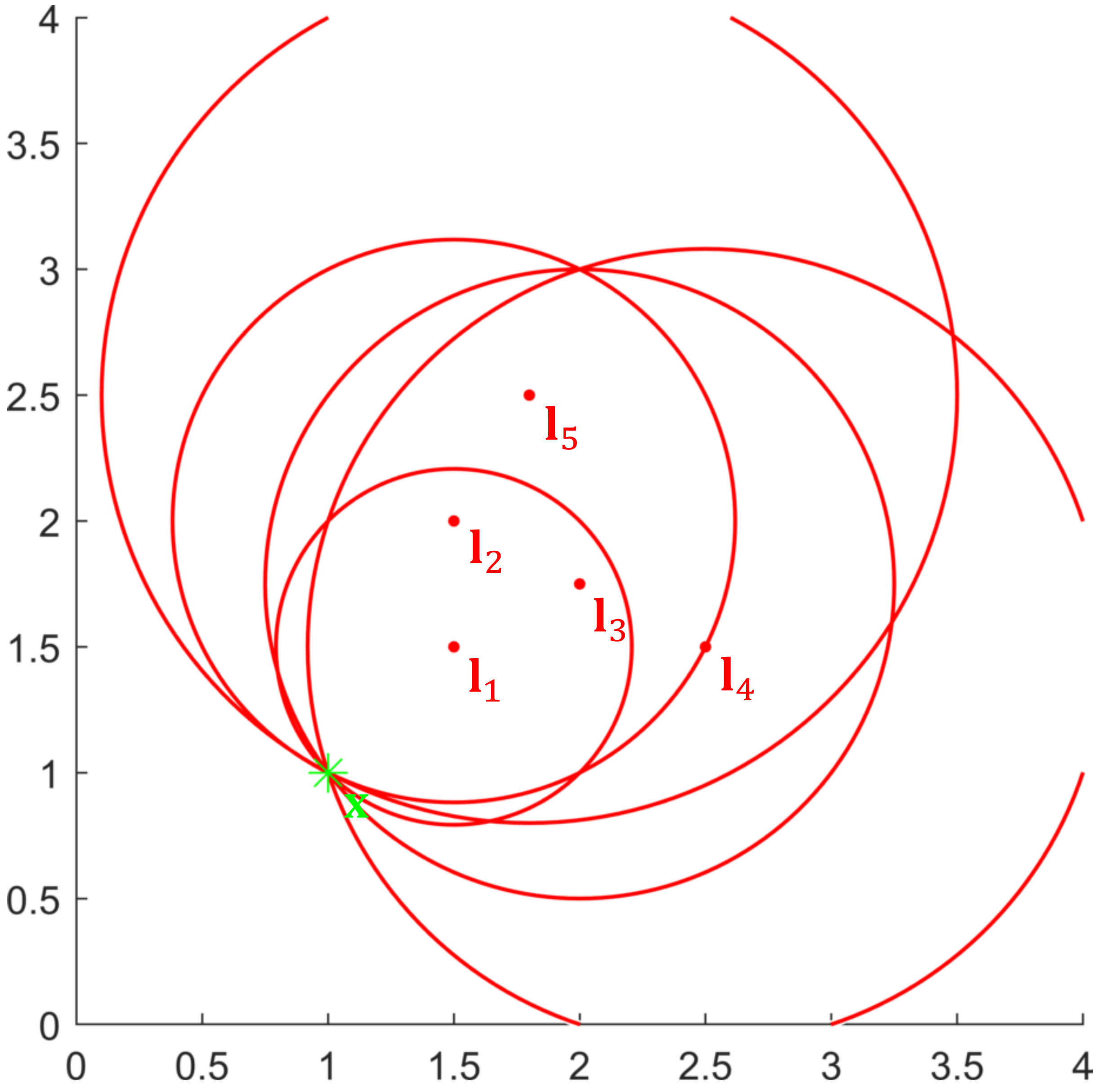
$$\underline{\mathcal{X}} = \mathbf{x}$$

Measurements: Range to landmarks

$$Z = \{\rho_1, \ldots, \rho_n\}$$

Measurement model:

$$\rho_i = \|\mathbf{x} - \mathbf{l}_i\| + \eta_i, \qquad \eta_i \sim N(0, \sigma_i^2)$$



Example:
Range-based localization

States: Our location

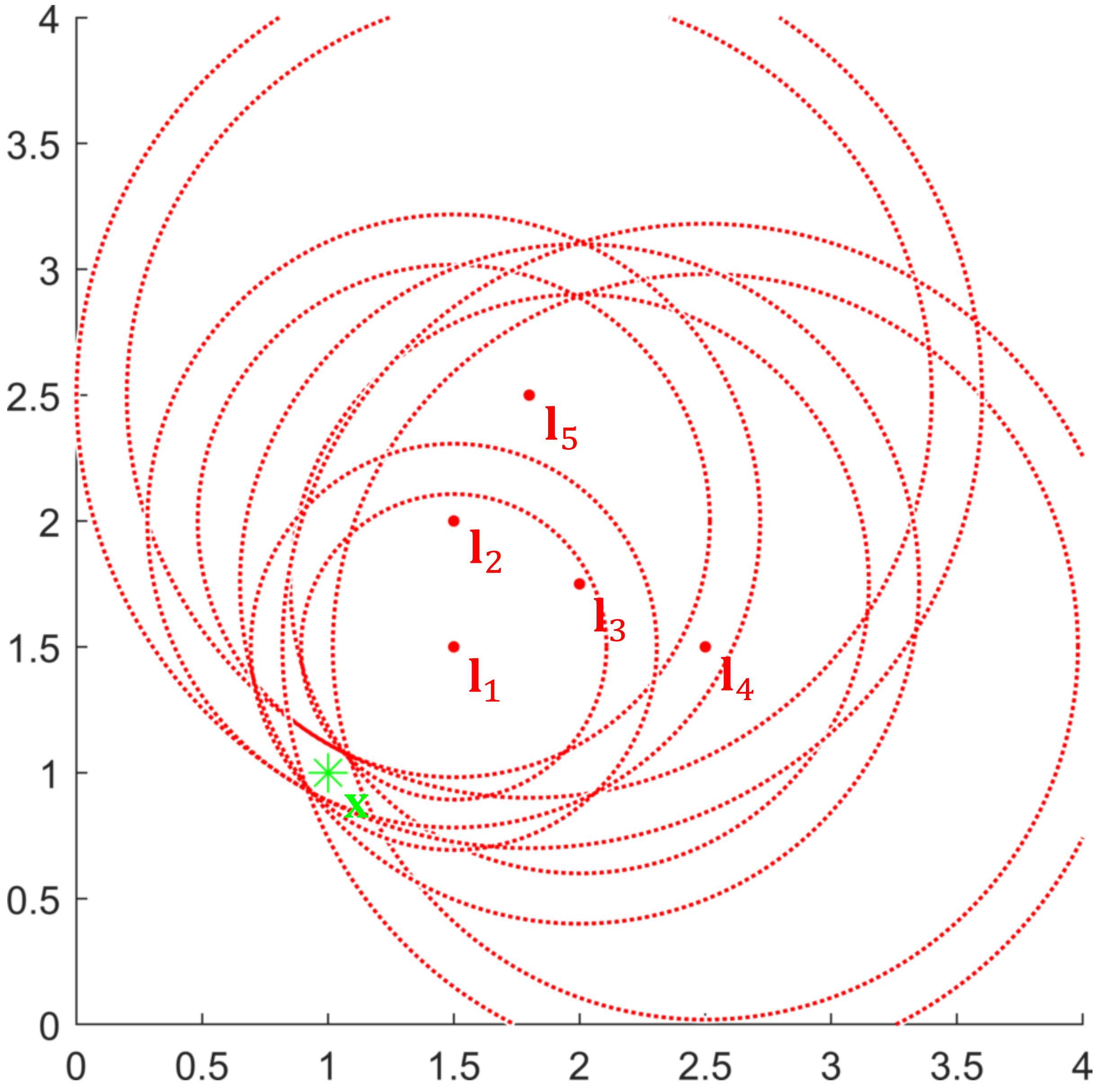
$$\underline{\mathcal{X}} = \mathbf{x}$$

Measurements: Range to landmarks

$$Z = \{\rho_1, \dots, \rho_n\}$$

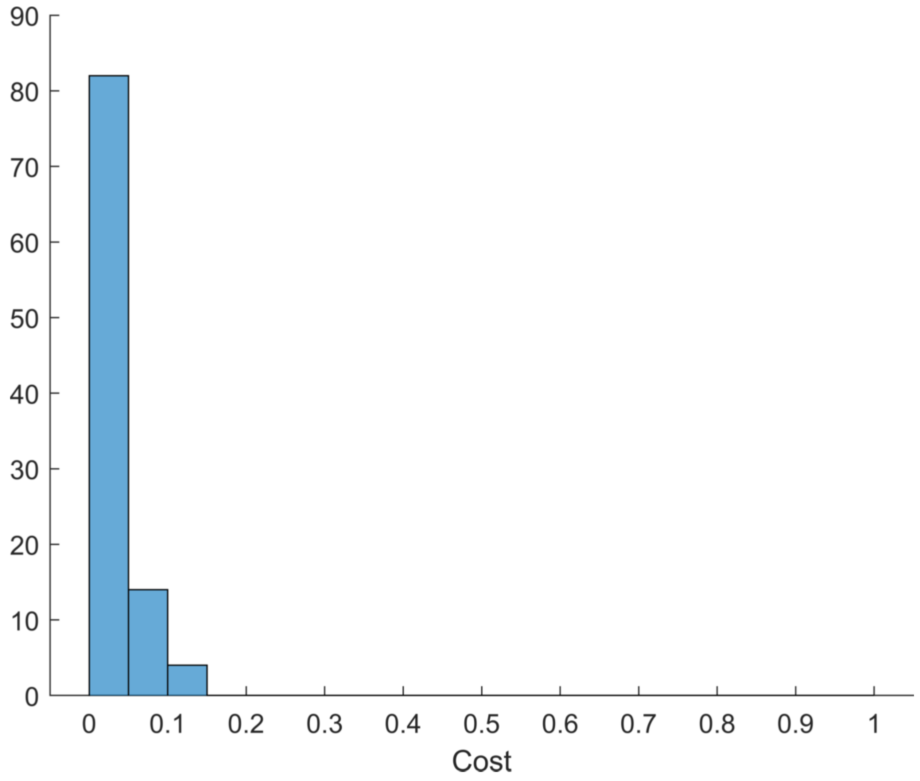
Measurement model:

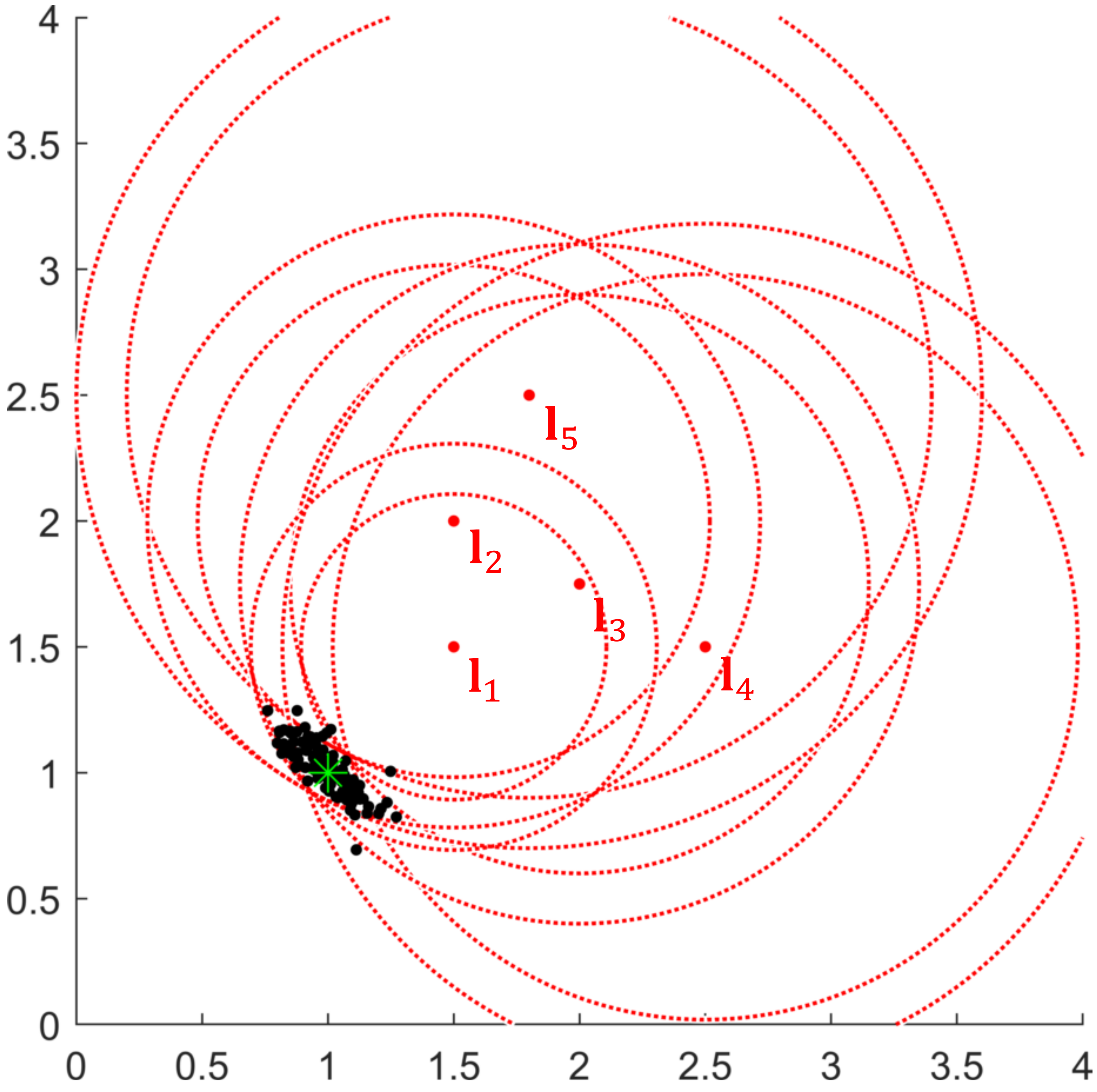
$$\rho_i = \|\mathbf{x} - \mathbf{l}_i\| + \eta_i, \qquad \eta_i \sim N(0, \sigma_i^2)$$



What happens when we ignore measurement noise?

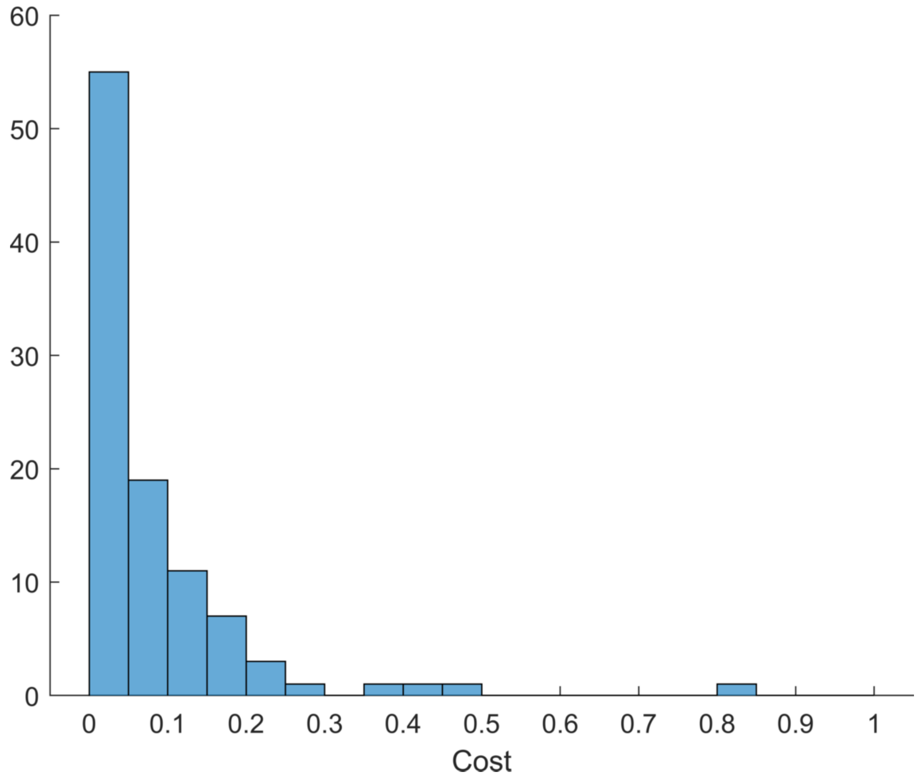
100 runs, $\sigma_i = 0.1$

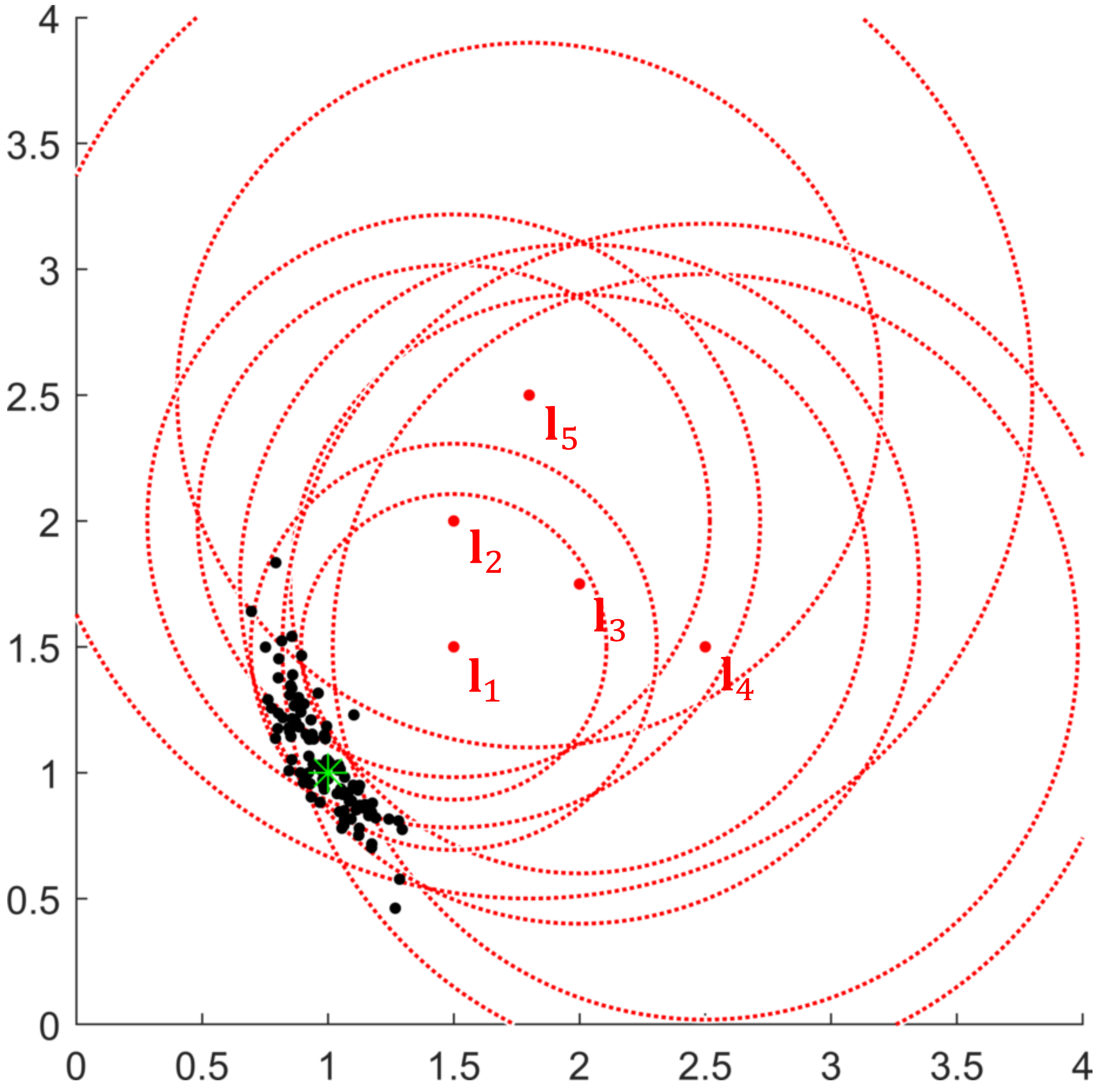




What happens when we ignore measurement noise?

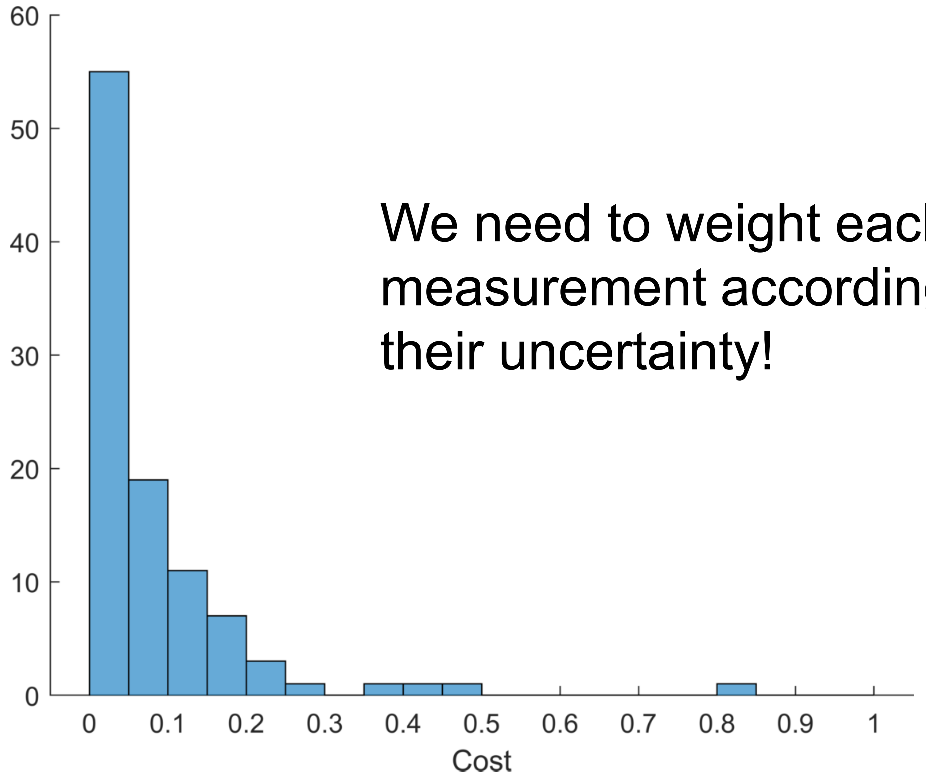
100 runs, $\sigma_1, \dots, \sigma_4 = 0.1, \sigma_5 = 0.3$



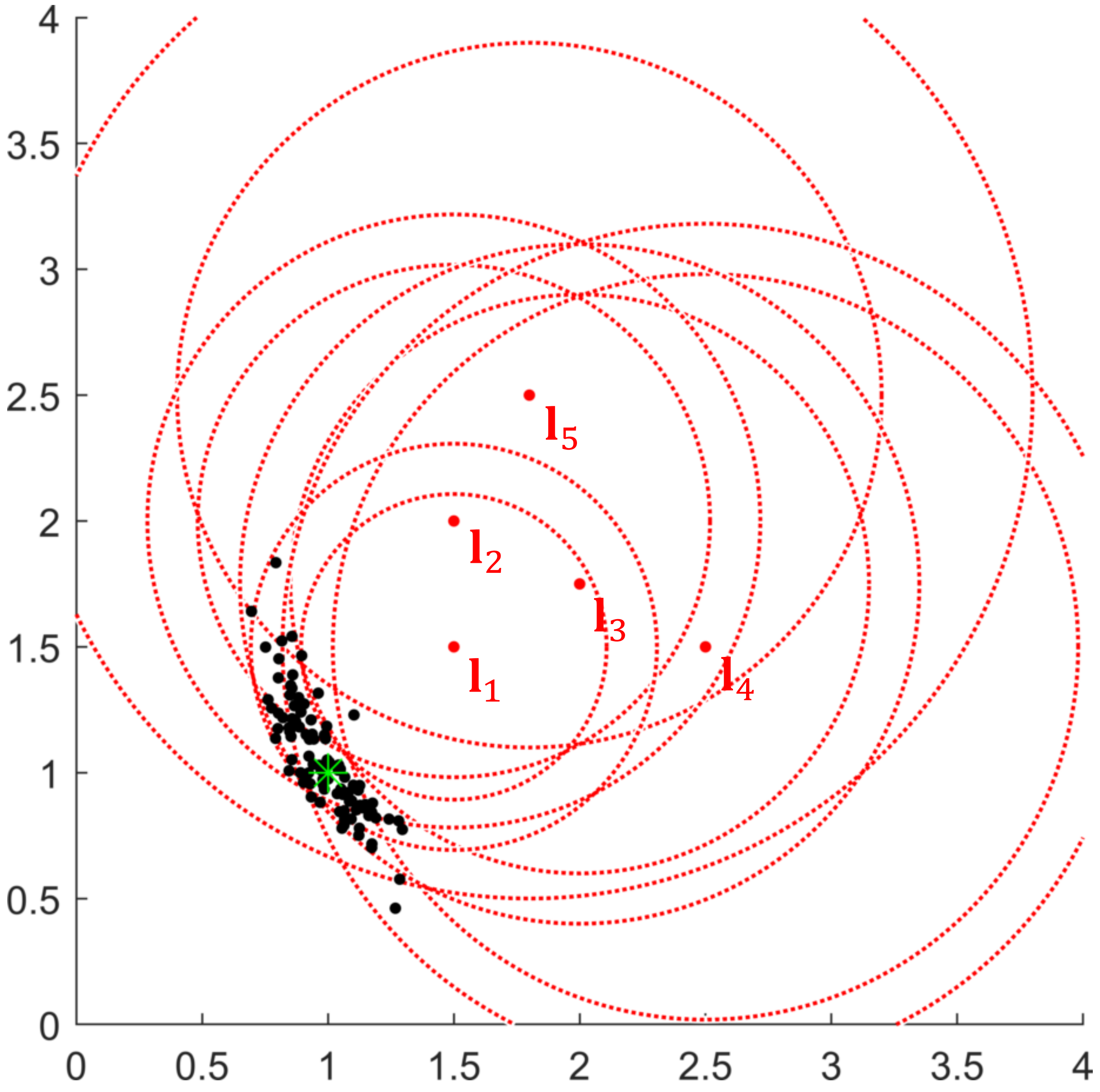


What happens when we ignore measurement noise?

100 runs, $\sigma_1, \dots, \sigma_4 = 0.1, \sigma_5 = 0.3$



We need to weight each measurement according to their uncertainty!



Weighted nonlinear least squares

We can rewrite the Mahalanobis norms as

$$\|\mathbf{e}\|_{\boldsymbol{\Sigma}}^2 \triangleq \mathbf{e}^T \boldsymbol{\Sigma}^{-1} \mathbf{e} = \left(\boldsymbol{\Sigma}^{-1/2}\mathbf{e}\right)^T \left(\boldsymbol{\Sigma}^{-1/2}\mathbf{e}\right) = \left\|\boldsymbol{\Sigma}^{-1/2}\mathbf{e}\right\|^2$$

Hence, we can eliminate the covariances
by weighting the Jacobian and the prediction error:

$$\mathbf{A}_i = \boldsymbol{\Sigma}_i^{-1/2} \mathbf{J}_{\hat{\underline{\mathcal{X}}}_i}^{h_i}$$

$$\mathbf{b}_i = \boldsymbol{\Sigma}_i^{-1/2} \left(\mathbf{z}_i - h_i(\hat{\underline{\mathcal{X}}}_i)\right)$$

This is a form of whitening,
which eliminates the units of the measurements

Weighted nonlinear least squares

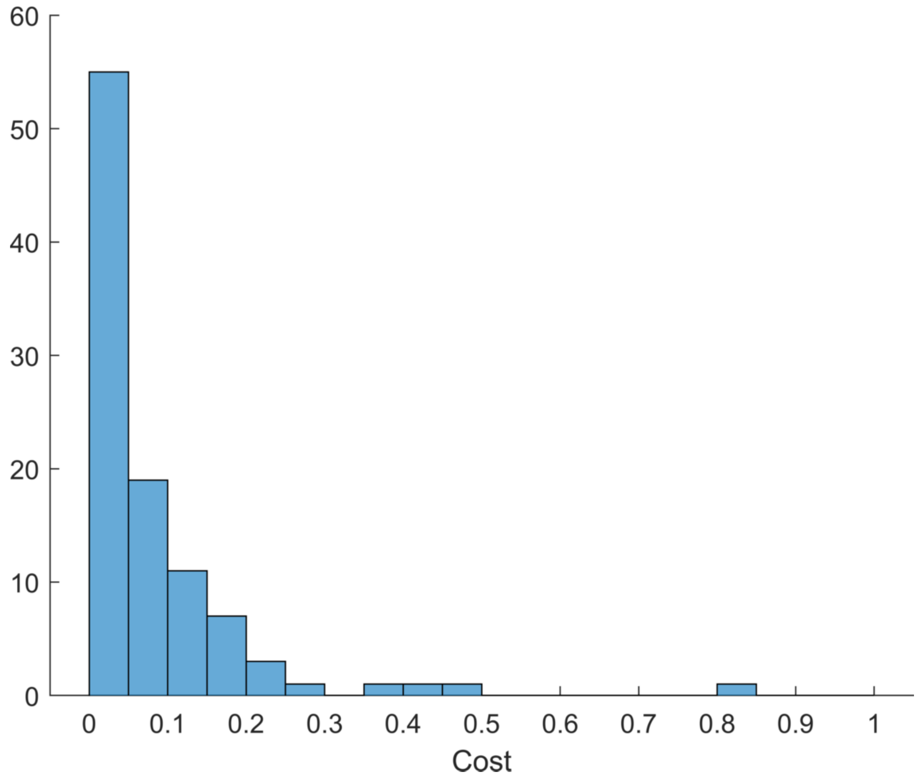
The objective function in the weighted least squares problem is now given by

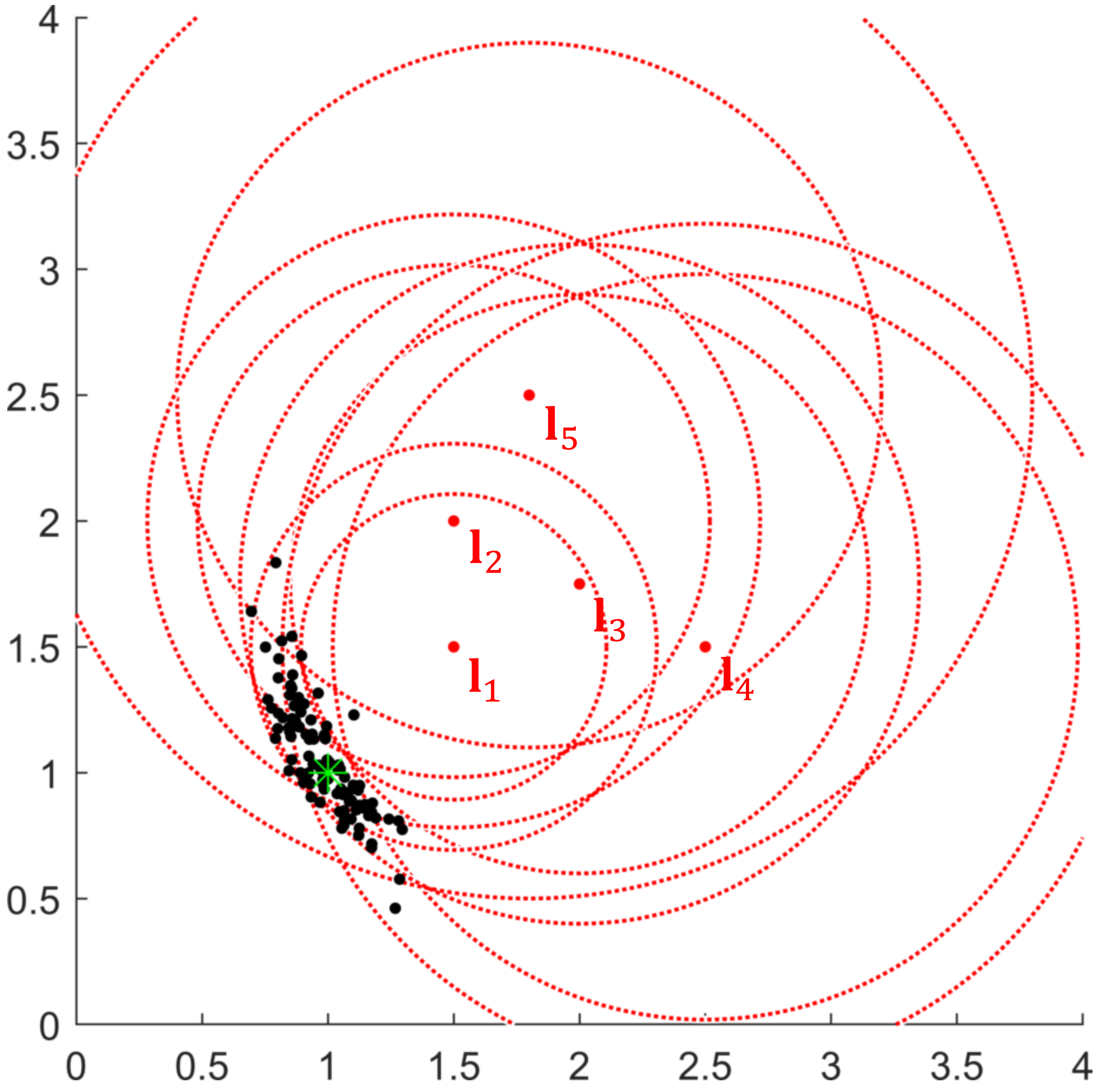
$$
\begin{aligned}
f(\underline{\mathcal{X}}) = f(\underline{\hat{\mathcal{X}}} \oplus \underline{\boldsymbol{\tau}}) &= \sum_{i=1}^{n} \left\| e_i(\underline{\hat{\mathcal{X}}}_i \oplus \underline{\boldsymbol{\tau}}_i) \right\|_{\boldsymbol{\Sigma}_i}^2 \\
&\approx \sum_{i=1}^{n} \left\| h_i(\underline{\hat{\mathcal{X}}}_i) + \mathbf{J}_{\underline{\hat{\mathcal{X}}}_i}^{h_i} \underline{\boldsymbol{\tau}}_i - \mathbf{z}_i \right\|_{\boldsymbol{\Sigma}_i}^2 \\
&= \sum_{i=1}^{n} \left\| \mathbf{J}_{\underline{\hat{\mathcal{X}}}_i}^{h_i} \underline{\boldsymbol{\tau}}_i - \left(\mathbf{z}_i - h_i(\underline{\hat{\mathcal{X}}}_i) \right) \right\|_{\boldsymbol{\Sigma}_i}^2 \\
&= \sum_{i=1}^{n} \left\| \boldsymbol{\Sigma}_i^{-1/2} \mathbf{J}_{\underline{\hat{\mathcal{X}}}_i}^{h_i} \underline{\boldsymbol{\tau}}_i - \boldsymbol{\Sigma}_i^{-1/2} \left(\mathbf{z}_i - h_i(\underline{\hat{\mathcal{X}}}_i) \right) \right\|^2 \\
&= \sum_{i=1}^{n} \left\| \mathbf{A}_i \underline{\boldsymbol{\tau}}_i - \mathbf{b}_i \right\|^2 \\
&= \left\| \mathbf{A} \underline{\boldsymbol{\tau}} - \mathbf{b} \right\|^2
\end{aligned}
$$

Taking measurement noise into account

100 runs, $\sigma_1, \dots, \sigma_4 = 0.1, \sigma_5 = 0.3$

Unweighted

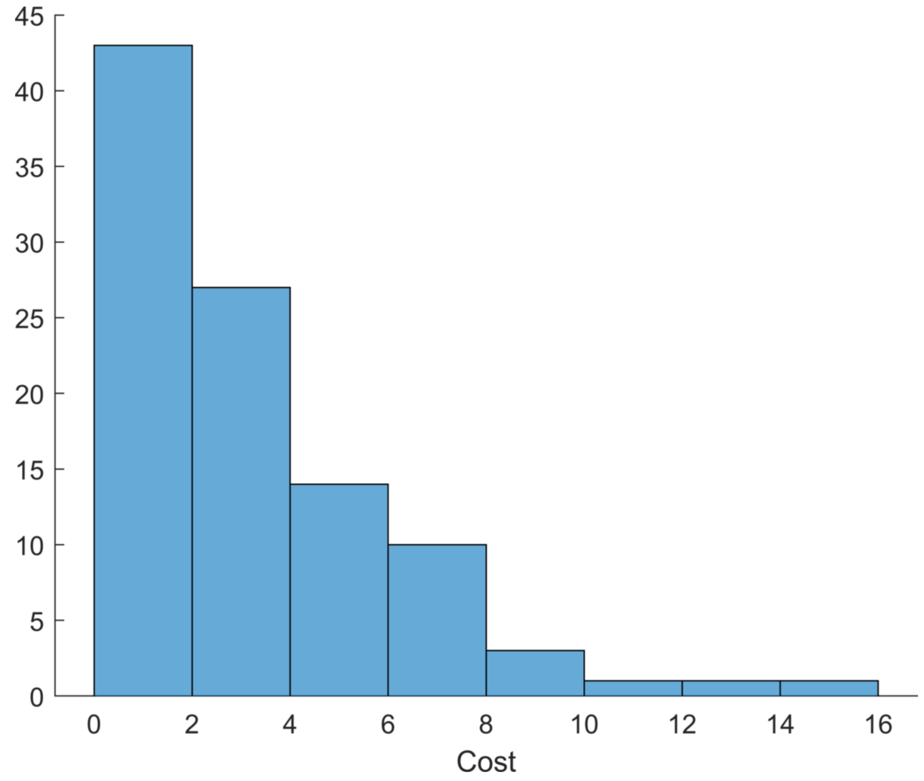


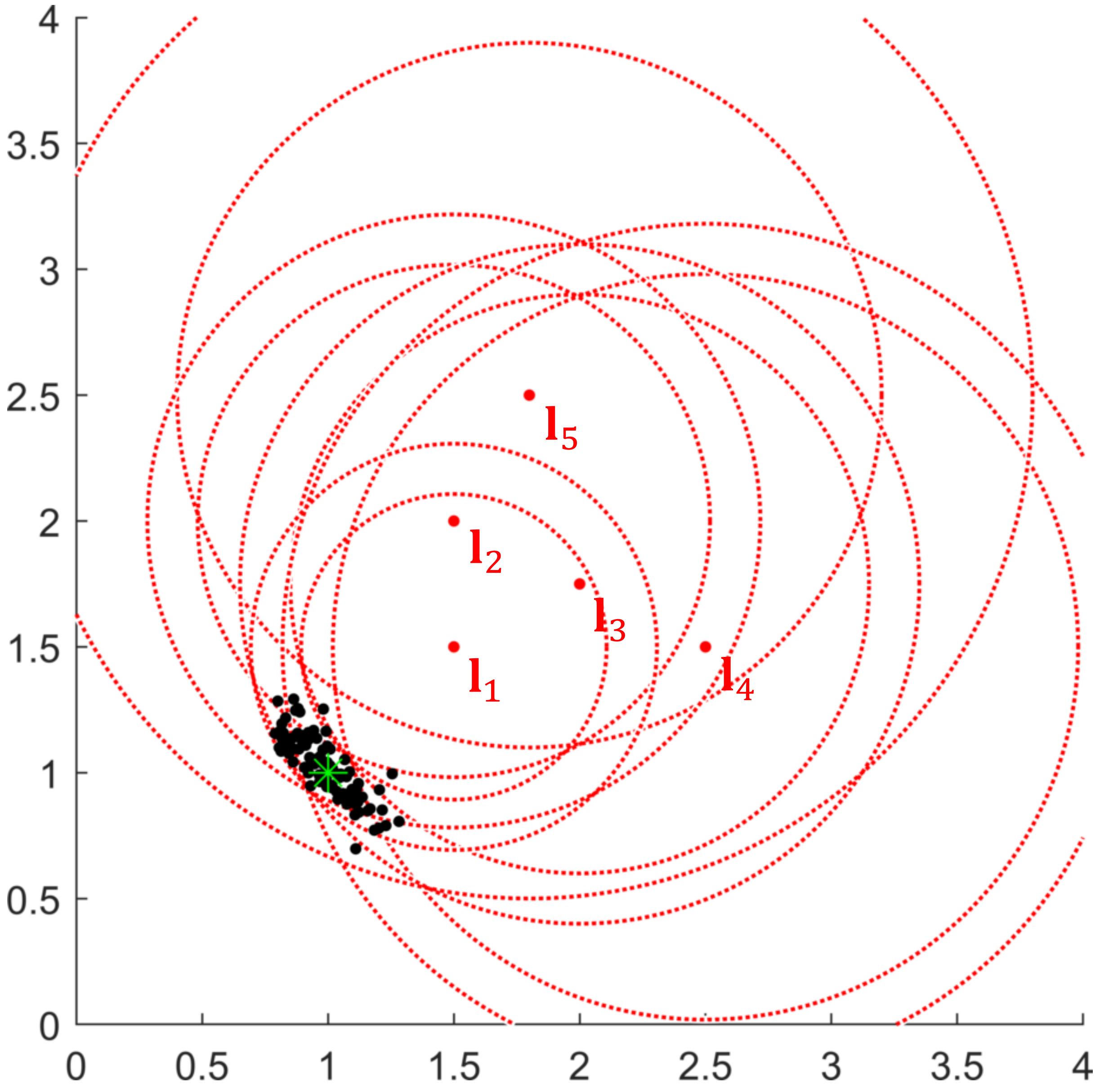


Taking measurement noise into account

100 runs, $\sigma_1, \ldots, \sigma_4 = 0.1, \sigma_5 = 0.3$

Covariance weighted (whitened)





Estimating uncertainty in the MAP estimate

The Hessian at the solution for the weighted problem
is the inverse of the covariance matrix (the information matrix)!

$$\frac{\partial^2 f(\underline{\hat{\mathcal{X}}}^*)}{\partial \underline{\hat{\mathcal{X}}}^* \partial \underline{\hat{\mathcal{X}}}^{*T}} = \Lambda = \Sigma^{-1}$$

Using our approximated Hessian,
we obtain a first order approximation of the true covariance for all states

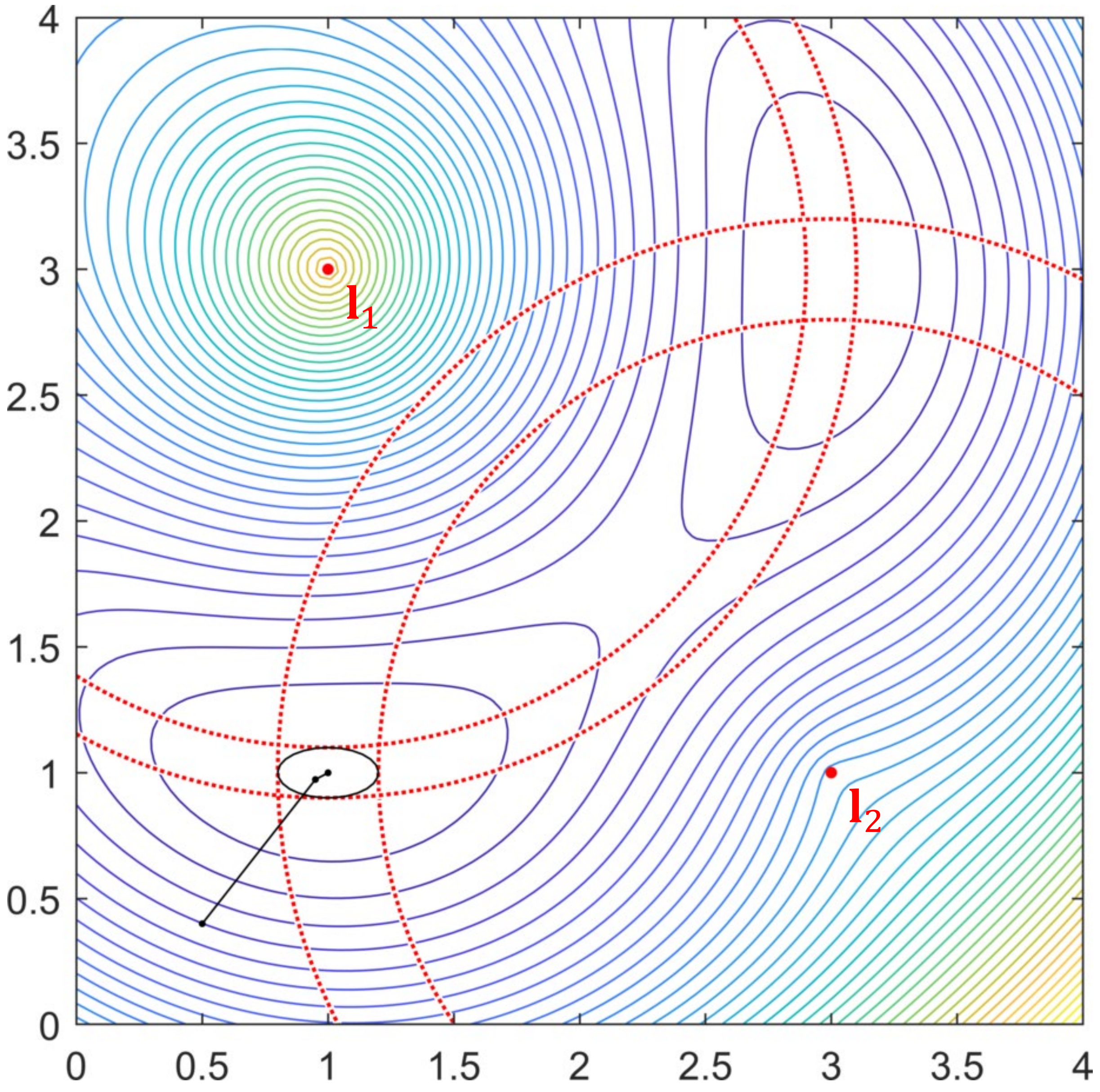
$$\Sigma_{\underline{\hat{\mathcal{X}}}^*} \approx (\mathbf{A}_{\underline{\hat{\mathcal{X}}}^*}^T \mathbf{A}_{\underline{\hat{\mathcal{X}}}^*})^{-1}$$

Simple example:
Two landmarks

(No noise added to measurements)

1σ covariance contours

$$\Sigma_{\underline{\hat{x}}^*} \approx (\mathbf{A}_{\underline{\hat{x}}^*}^T \mathbf{A}_{\underline{\hat{x}}^*})^{-1}$$

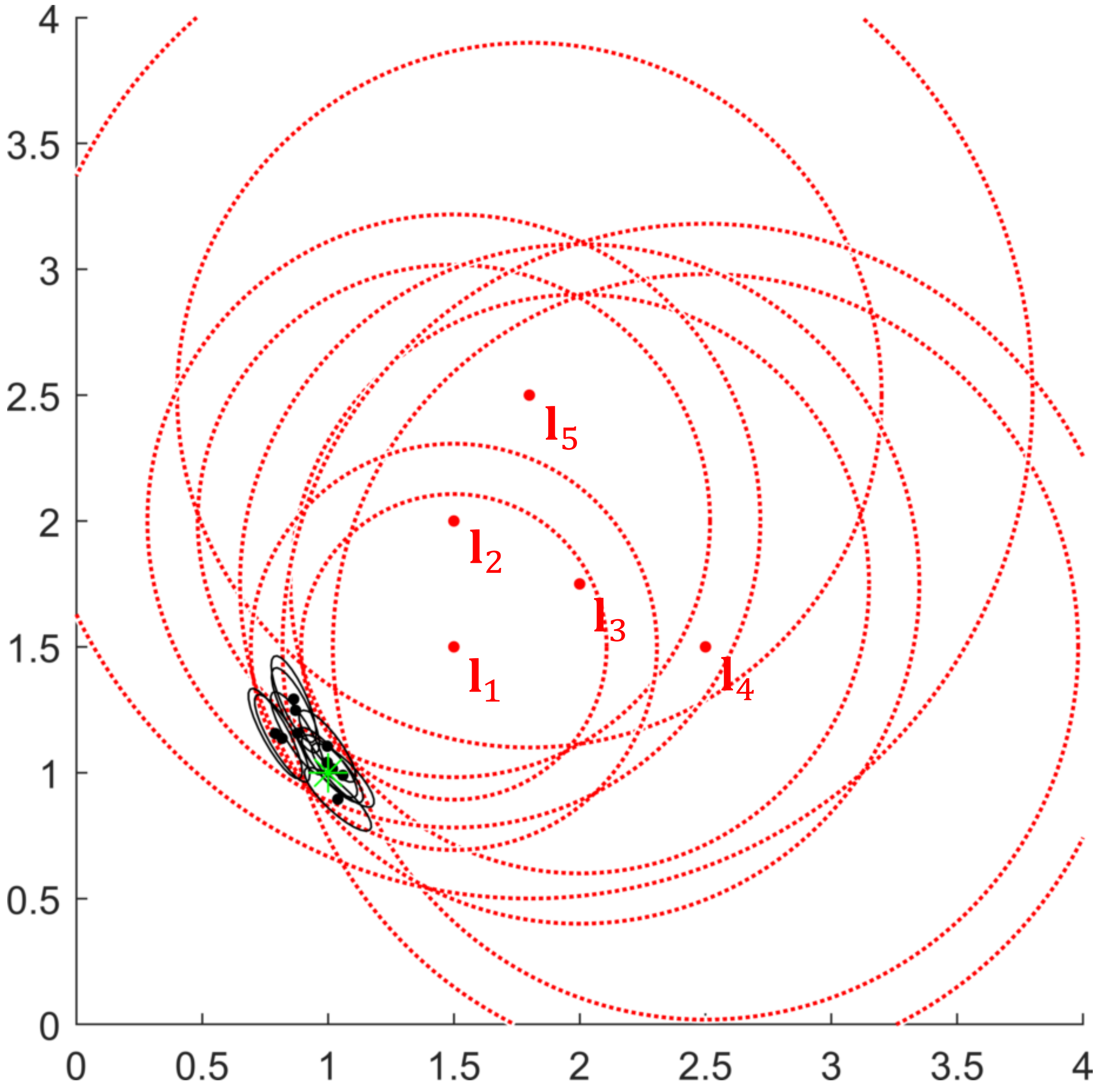


Example:
Range-based localization

10 runs, $\sigma_1, \ldots, \sigma_4 = 0.1, \sigma_5 = 0.3$

1σ covariance contours

$$\Sigma_{\underline{\hat{\mathcal{X}}}^*} \approx (\mathbf{A}_{\underline{\hat{\mathcal{X}}}^*}^T \mathbf{A}_{\underline{\hat{\mathcal{X}}}^*})^{-1}$$



Summary

We have seen how we can find the MAP estimate of our unknown states given measurements

$$X^{MAP} = \underset{X}{\operatorname{argmax}}\; p(X \mid Z)$$

by representing it as
a nonlinear least squares problem

$$\underline{\mathcal{X}}^{*} = \underset{\underline{\mathcal{X}}}{\operatorname{argmin}} \sum_{i=1}^{n} \left\| h_i(\underline{\mathcal{X}}_i) - \mathbf{z}_i \right\|_{\Sigma_i}^{2}$$

The resulting estimate is the (joint) probability distribution

$$\hat{\underline{\mathcal{X}}} \sim N(\hat{\bar{\underline{\mathcal{X}}}}, \hat{\Sigma}_{\hat{\underline{\mathcal{X}}}})$$

$$\hat{\bar{\underline{\mathcal{X}}}} = \hat{\underline{\mathcal{X}}}^{*}$$

$$\hat{\Sigma}_{\hat{\underline{\mathcal{X}}}} = (\mathbf{A}_{\hat{\underline{\mathcal{X}}}^{*}}^{T} \mathbf{A}_{\hat{\underline{\mathcal{X}}}^{*}})^{-1}$$

Choose a suitable inital estimate $\hat{\underline{\mathcal{X}}}^{0}$

⇩

$\mathbf{A}, \mathbf{b} \leftarrow$ Linearize at $\hat{\underline{\mathcal{X}}}^{t}$

⇩

$\underline{\boldsymbol{\tau}}^{*} \leftarrow$ Solve $\underset{\underline{\boldsymbol{\tau}}}{\operatorname{argmin}} \left\| \mathbf{A}\underline{\boldsymbol{\tau}} - \mathbf{b} \right\|^{2}$

⇩

$\hat{\underline{\mathcal{X}}}^{t+1} \leftarrow \hat{\underline{\mathcal{X}}}^{t} \oplus \underline{\boldsymbol{\tau}}^{*}$