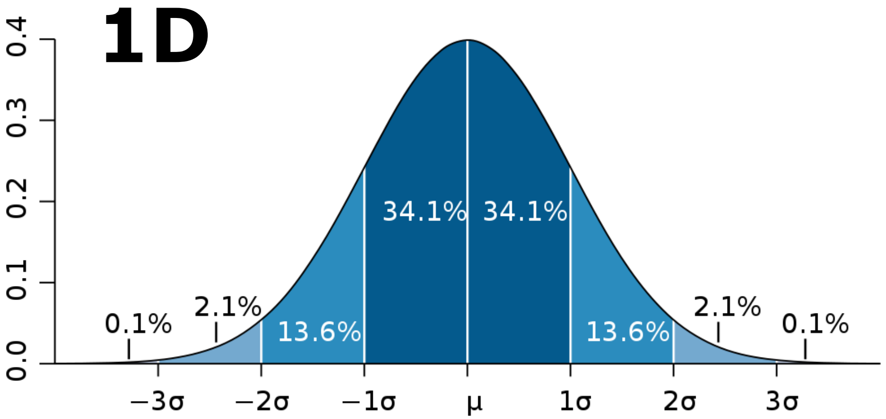
Sparse Extended Information Filter (SEIF) - SLAM

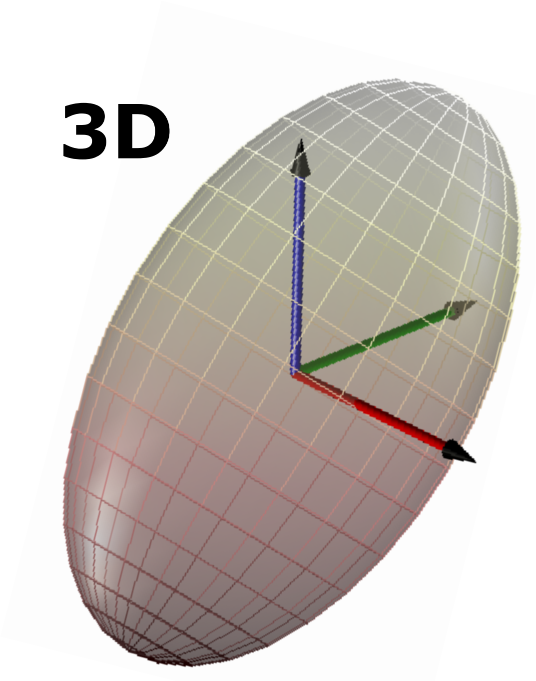
Dr. Gaurav Pandey

EE698G

Gaussian Distribution

$$p(x) = \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$$





Slide Credit: Cyrill Stachniss

Canonical Parametrization

- Alternate parametrization of Gaussian Distribution
- Parametrized by **Information Matrix** and **Information Vector**

$$\Omega = \Sigma^{-1} \qquad \xi = \Sigma^{-1}\mu$$

Information Matrix

Information Vector

$$p(x) = \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$$

Moments Parametrization

Moments vs Canonical Parametrization

$$
\begin{aligned}
&p(x)\\
&= \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\Big)\\
&= \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu - \frac{1}{2}\mu^T\Sigma^{-1}\mu\Big)\\
&= \boxed{\det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}\mu^T\Sigma^{-1}\mu\Big)}\\
&\qquad \exp\Big(-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu\Big)\\
&= \eta \exp\Big(-\frac{1}{2}x^T\underline{\Sigma^{-1}}x + x^T\underline{\Sigma^{-1}\mu}\Big)
\end{aligned}
$$

Dual Parametrization

- Moments Parametrization

$$p(x) = \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\Big)$$

- Canonical Parametrization

$$p(x) = \frac{\exp(-\frac{1}{2}\mu^T\xi)}{\det(2\pi\Omega^{-1})^{\frac{1}{2}}} \exp\Big(-\frac{1}{2}x^T\Omega x + x^T\xi\Big)$$

Information Filter

- Information Filter is Kalman filter with canonical parametrization i.e. beliefs are parametrized by **information matrix** and **information vector** instead of covariance matrix and mean vector.

$$\Omega = \Sigma^{-1}$$

$$\xi = \Sigma^{-1} \mu$$

Information Form

$$\Sigma = \Omega^{-1}$$

$$\mu = \Omega^{-1} \xi$$

Moments Form

Kalman Filter vs Information Filter

Algorithm Kalman_filter$(\mu_{t-1}, \Sigma_{t-1}, u_t, z_t)$**:**

$\bar{\mu}_t = A_t\ \mu_{t-1} + B_t\ u_t$
$\bar{\Sigma}_t = A_t\ \Sigma_{t-1}\ A_t^T + R_t$
$K_t = \bar{\Sigma}_t\ C_t^T (C_t\ \bar{\Sigma}_t\ C_t^T + Q_t)^{-1}$
$\mu_t = \bar{\mu}_t + K_t (z_t - C_t\ \bar{\mu}_t)$
$\Sigma_t = (I - K_t\ C_t)\ \bar{\Sigma}_t$
return μ_t, Σ_t

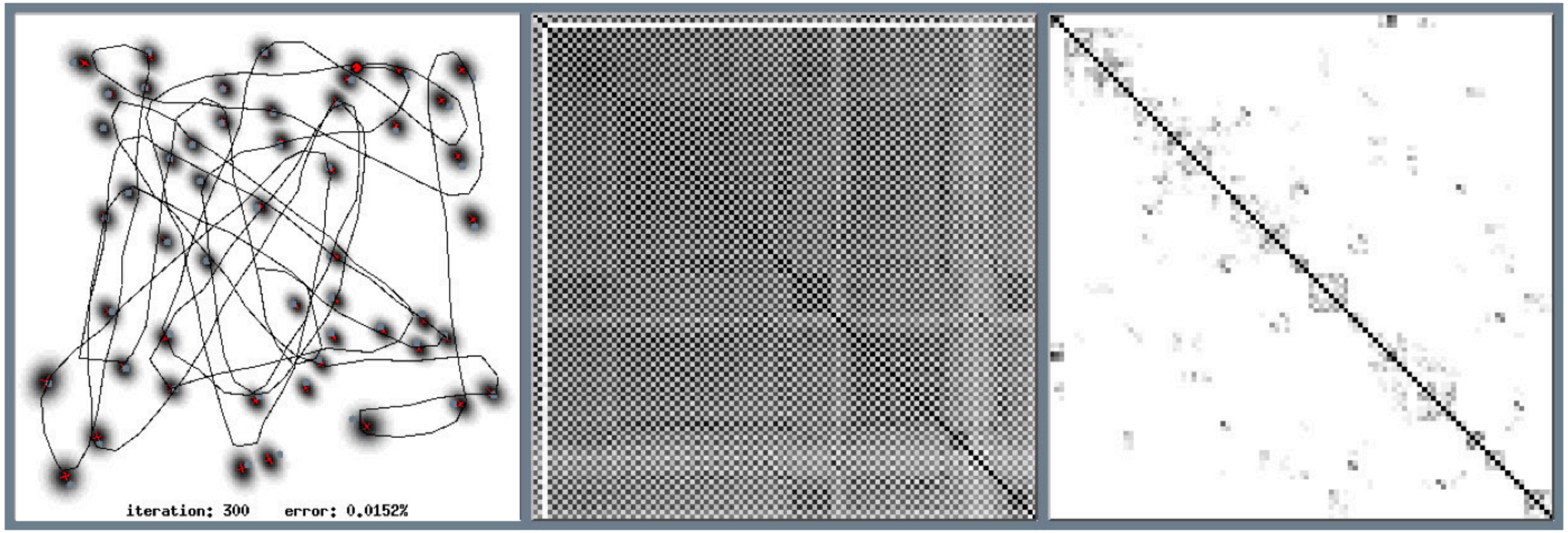
Algorithm Information_filter$(\xi_{t-1}, \Omega_{t-1}, u_t, z_t)$**:**

$\bar{\Omega}_t = (A_t\ \Omega_{t-1}^{-1}\ A_t^T + R_t)^{-1}$
$\bar{\xi}_t = \bar{\Omega}_t (A_t\ \Omega_{t-1}^{-1}\ \xi_{t-1} + B_t\ u_t)$
$\Omega_t = C_t^T\ Q_t^{-1}\ C_t + \bar{\Omega}_t$
$\xi_t = C_t^T\ Q_t^{-1}\ z_t + \bar{\xi}_t$
return ξ_t, Ω_t

Note: No need to find Kalman Gain in Information Form !!

Kalman vs Information Filter

- If you use information filter as it is then you will not gain anything.
- Note that information matrix is sparse or can be made sparse.
- We can utilize this fact to make the information filter computationally efficient.



Map   Covariance Matrix   Information Matrix

Image Credit: Cyrill Stachnis

Covariance vs Information Matrix

$$p(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu}_\alpha \\ \boldsymbol{\mu}_\beta \end{bmatrix}, \begin{bmatrix} \Sigma_{\alpha\alpha} & \Sigma_{\alpha\beta} \\ \Sigma_{\beta\alpha} & \Sigma_{\beta\beta} \end{bmatrix}\right) = \mathcal{N}^{-1}\left(\begin{bmatrix} \boldsymbol{\eta}_\alpha \\ \boldsymbol{\eta}_\beta \end{bmatrix}, \begin{bmatrix} \Lambda_{\alpha\alpha} & \Lambda_{\alpha\beta} \\ \Lambda_{\beta\alpha} & \Lambda_{\beta\beta} \end{bmatrix}\right)$$

	MARGINALIZATION $p(\boldsymbol{\alpha}) = \int p(\boldsymbol{\alpha}, \boldsymbol{\beta})\, d\boldsymbol{\beta}$	CONDITIONING $p(\boldsymbol{\alpha} \mid \boldsymbol{\beta}) = p(\boldsymbol{\alpha}, \boldsymbol{\beta}) / p(\boldsymbol{\beta})$
COVARIANCE FORM	$\boldsymbol{\mu} = \boldsymbol{\mu}_\alpha$ $\Sigma = \Sigma_{\alpha\alpha}$	$\boldsymbol{\mu}' = \boldsymbol{\mu}_\alpha + \Sigma_{\alpha\beta}\Sigma_{\beta\beta}^{-1}(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)$ $\Sigma' = \Sigma_{\alpha\alpha} - \Sigma_{\alpha\beta}\Sigma_{\beta\beta}^{-1}\Sigma_{\beta\alpha}$
INFORMATION FORM	$\boldsymbol{\eta} = \boldsymbol{\eta}_\alpha - \Lambda_{\alpha\beta}\Lambda_{\beta\beta}^{-1}\boldsymbol{\eta}_\beta$ $\Lambda = \Lambda_{\alpha\alpha} - \Lambda_{\alpha\beta}\Lambda_{\beta\beta}^{-1}\Lambda_{\beta\alpha}$	$\boldsymbol{\eta}' = \boldsymbol{\eta}_\alpha - \Lambda_{\alpha\beta}\boldsymbol{\beta}$ $\Lambda' = \Lambda_{\alpha\alpha}$

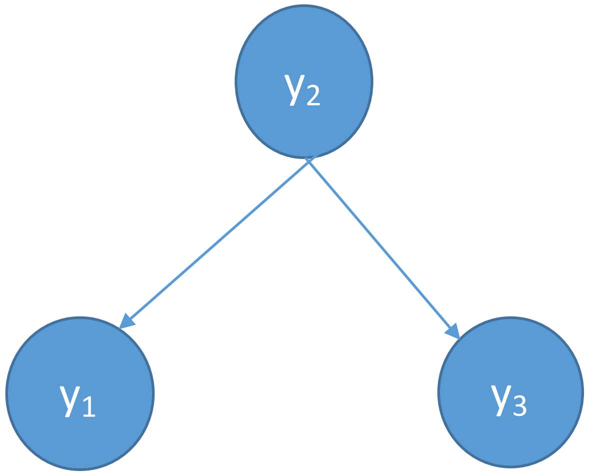
Slide Credit: R. Eustice

Covariance vs Information matrix

- Suppose you have 3 Gaussian random variables y_1, y_2, y_3 that are related as follows:

$$y_2 = n_2 \; ; \; y_1 = w_1 y_2 + n_1 \; ; \; y_3 = w_3 y_2 + n_3$$

n_1, n_2, n_3 are **independent** Gaussian random variables with zero mean and variance $[\sigma_1^2, \sigma_2^2, \sigma_3^2]$



Covariance Matrix

- What is the covariance of the joint random variable y = [y_1, y_2, y_3]

$$\Sigma = \begin{bmatrix} w_1^2\sigma_2^2 + \sigma_1^2 & w_1\sigma_2^2 & w_1 w_3 \sigma_2^2 \\ w_1\sigma_2^2 & \sigma_2^2 & w_3\sigma_2^2 \\ w_1 w_3 \sigma_2^2 & w_3\sigma_2^2 & w_3\sigma_2^2 + \sigma_3^2 \end{bmatrix}$$

- The covariance matrix is **Dense**

Information Matrix

- Consider the joint distribution

$$P(y_1, y_2, y_3) = P(y_1)P(y_1|y_2)P(y_3|y_2)$$

$$= \eta_2 \exp\left(-\frac{y_2^2}{2\sigma_2^2}\right) \eta_1 \exp\left(-\frac{(y_1 - w_1 y_2)^2}{2\sigma_1^2}\right) \eta_3 \exp\left(-\frac{(y_3 - w_3 y_2)^2}{2\sigma_3^2}\right)$$

$$= \eta \exp\left(-\frac{y_2^2}{2\sigma_2^2} - \frac{(y_1 - w_1 y_2)^2}{2\sigma_1^2} - \frac{(y_3 - w_3 y_2)^2}{2\sigma_3^2}\right)$$

Information Matrix

$$P(y_1, y_2, y_3) =$$

$$= \eta \exp\left(-\frac{1}{2} \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{w_1}{\sigma_1^2} & \mathbf{0} \\ -\frac{w_1}{\sigma_1^2} & \left(\frac{1}{\sigma_1^2} + \frac{w_1^2}{\sigma_1^2} + \frac{w_3^2}{\sigma_3^2} \right) & -\frac{w_3}{\sigma_3^2} \\ \mathbf{0} & -\frac{w_3}{\sigma_3^2} & -\frac{1}{\sigma_3^2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \right)$$

$$\Omega = \begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{w_1}{\sigma_1^2} & \mathbf{0} \\ -\frac{w_1}{\sigma_1^2} & \left(\frac{1}{\sigma_1^2} + \frac{w_1^2}{\sigma_1^2} + \frac{w_3^2}{\sigma_3^2} \right) & -\frac{w_3}{\sigma_3^2} \\ \mathbf{0} & -\frac{w_3}{\sigma_3^2} & -\frac{1}{\sigma_3^2} \end{bmatrix}$$

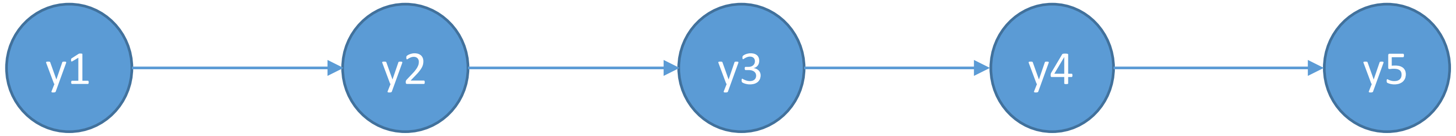
Information Matrix

$$\Omega = \begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{w_1}{\sigma_1^2} & \mathbf{0} \\ -\frac{w_1}{\sigma_1^2} & \left(\frac{1}{\sigma_1^2} + \frac{w_1^2}{\sigma_1^2} + \frac{w_3^2}{\sigma_3^2}\right) & -\frac{w_3}{\sigma_3^2} \\ \mathbf{0} & -\frac{w_3}{\sigma_3^2} & -\frac{1}{\sigma_3^2} \end{bmatrix}$$

- y_3 and y_1 are conditionally independent (i.e. they are independent when y_2 is known)
- If you marginalize out y_2 then y_1 and y_3 will no longer be independent.

Information Matrix

- What will be the structure of information matrix for the following graph ?



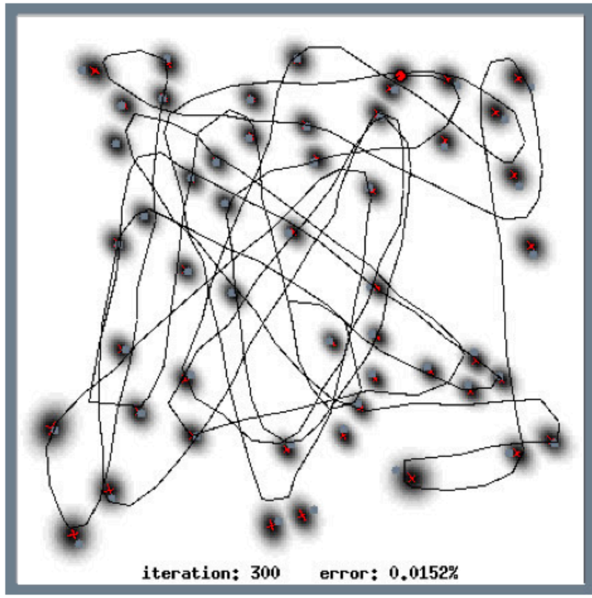
$$\Omega = \begin{bmatrix} c_{11} & c_{12} & 0 & 0 & 0 \\ c_{12} & c_{22} & c_{23} & 0 & 0 \\ 0 & c_{23} & c_{33} & c_{34} & 0 \\ 0 & 0 & c_{34} & c_{44} & c_{45} \\ 0 & 0 & 0 & c_{45} & c_{55} \end{bmatrix}$$

SEIF SLAM

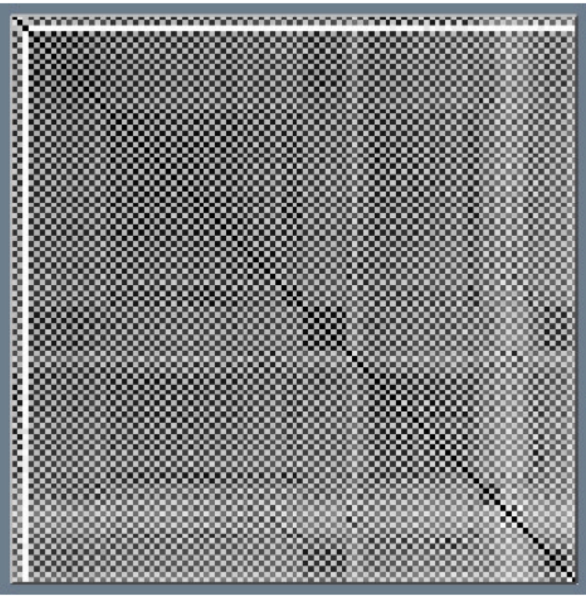
$\textbf{SEIF\_SLAM}(\xi_{t-1}, \Omega_{t-1}, \mu_{t-1}, u_t, z_t)$:

1: $\bar{\xi}_t, \bar{\Omega}_t, \bar{\mu}_t = \textbf{SEIF\_motion\_update}(\xi_{t-1}, \Omega_{t-1}, \mu_{t-1}, u_t)$
2: $\xi_t, \Omega_t = \textbf{SEIF\_measurement\_update}(\bar{\xi}_t, \bar{\Omega}_t, \bar{\mu}_t, z_t)$
3: $\mu_t = \textbf{SEIF\_update\_state\_estimate}(\xi_t, \Omega_t, \bar{\mu}_t)$
4: $\tilde{\xi}_t, \tilde{\Omega}_t = \textbf{SEIF\_sparsification}(\xi_t, \Omega_t, \mu_t)$
5: $\textit{return}\ \tilde{\xi}_t, \tilde{\Omega}_t, \mu_t$

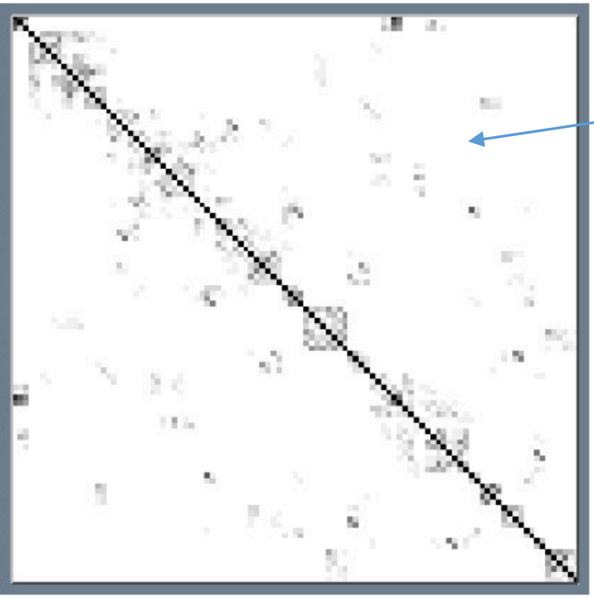
SLAM



Map



Covariance Matrix



Information Matrix

Small but
Non-zero

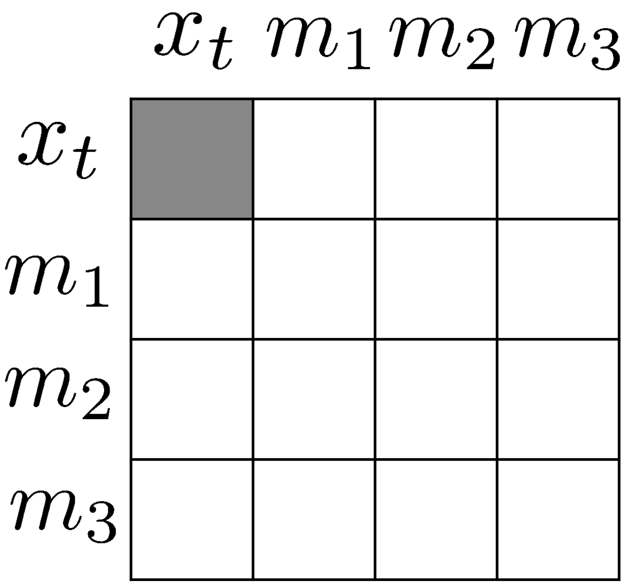
Slide Credit: Cyrill Stachniss

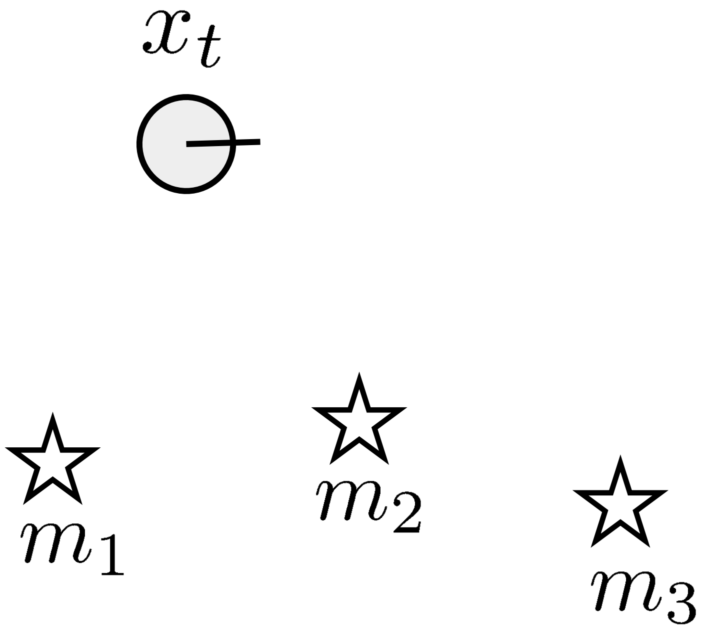
Sparseness of SEIF SLAM

- The key idea of SEIF SLAM is to maintain the sparsity of the information matrix and exploit it in computation
- Sparse matrix algebra allows efficient computation
- Sparse matrix means there are only finite (fixed) number of non-zero elements in the matrix, irrespective of its size.

Intuition behind SEIF-SLAM

- Before any measurement

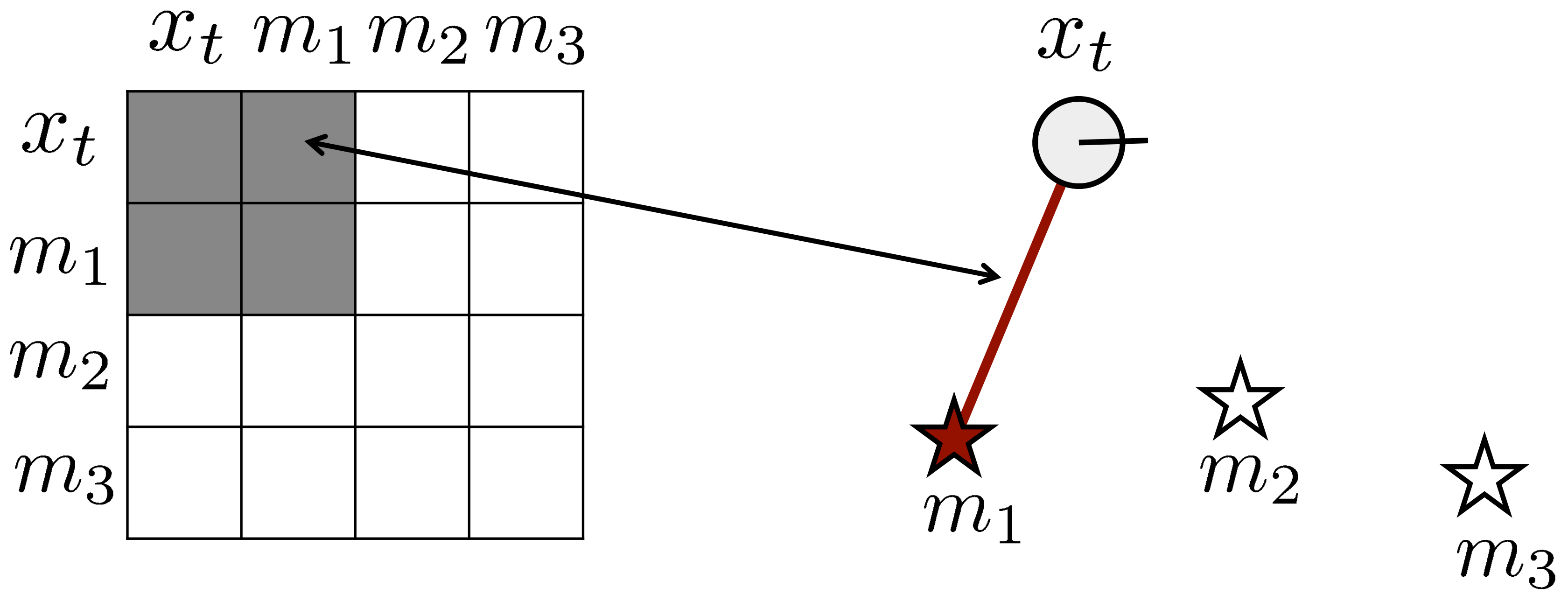




Slide Credit: Cyrill Stachniss

Intuition behind SEIF-SLAM

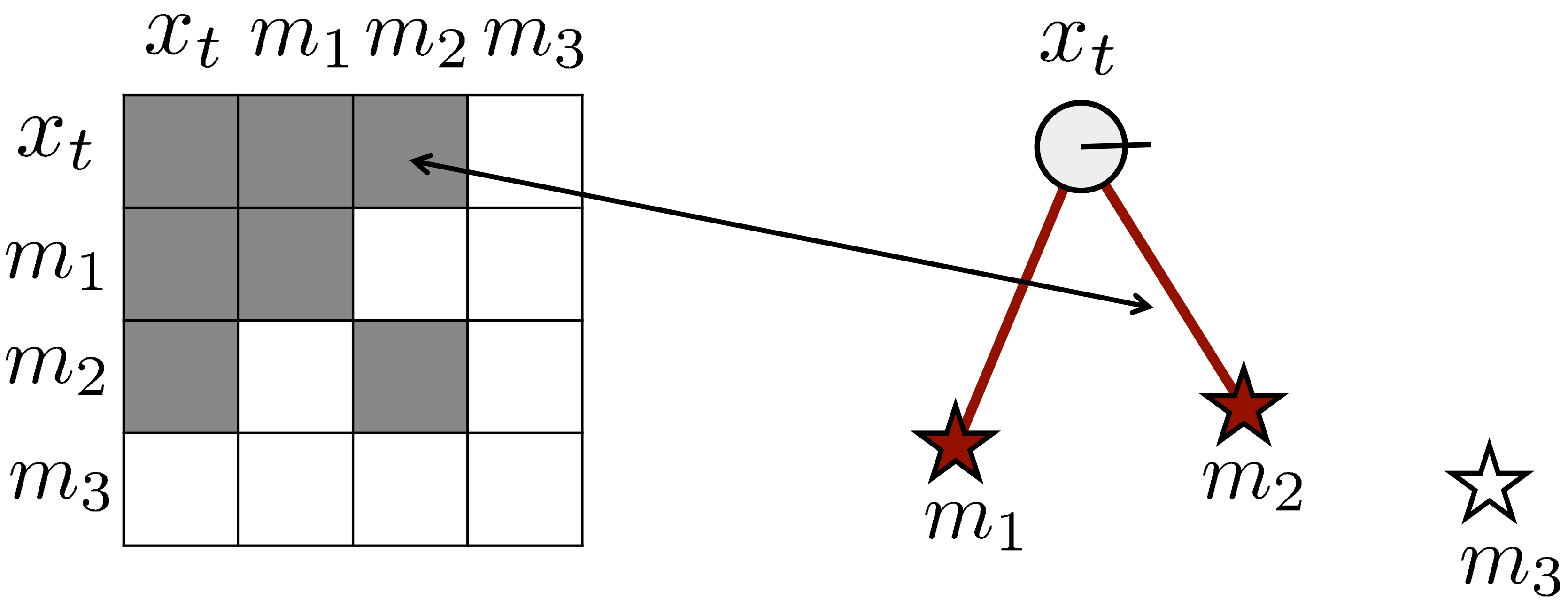
- Robot observes a landmark



Slide Credit: Cyrill Stachniss

Intuition behind SEIF-SLAM

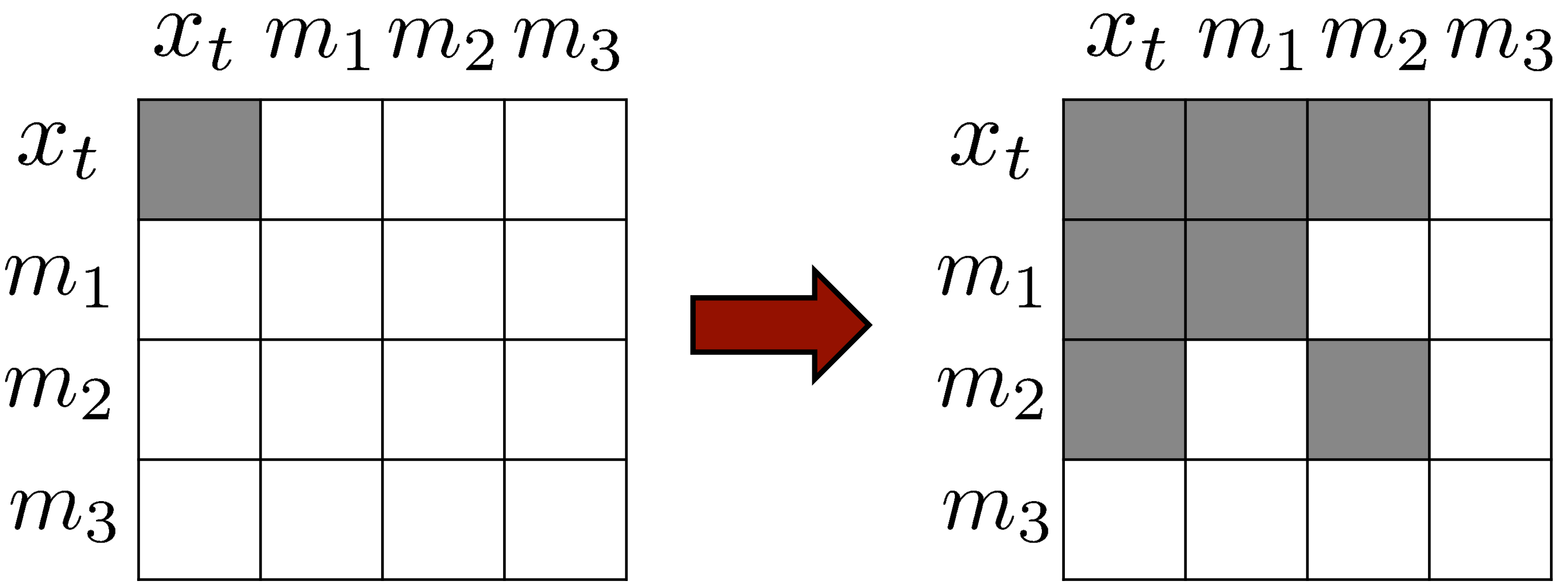
- Robot observes another landmark



Slide Credit: Cyrill Stachniss

Intuition behind SEIF-SLAM

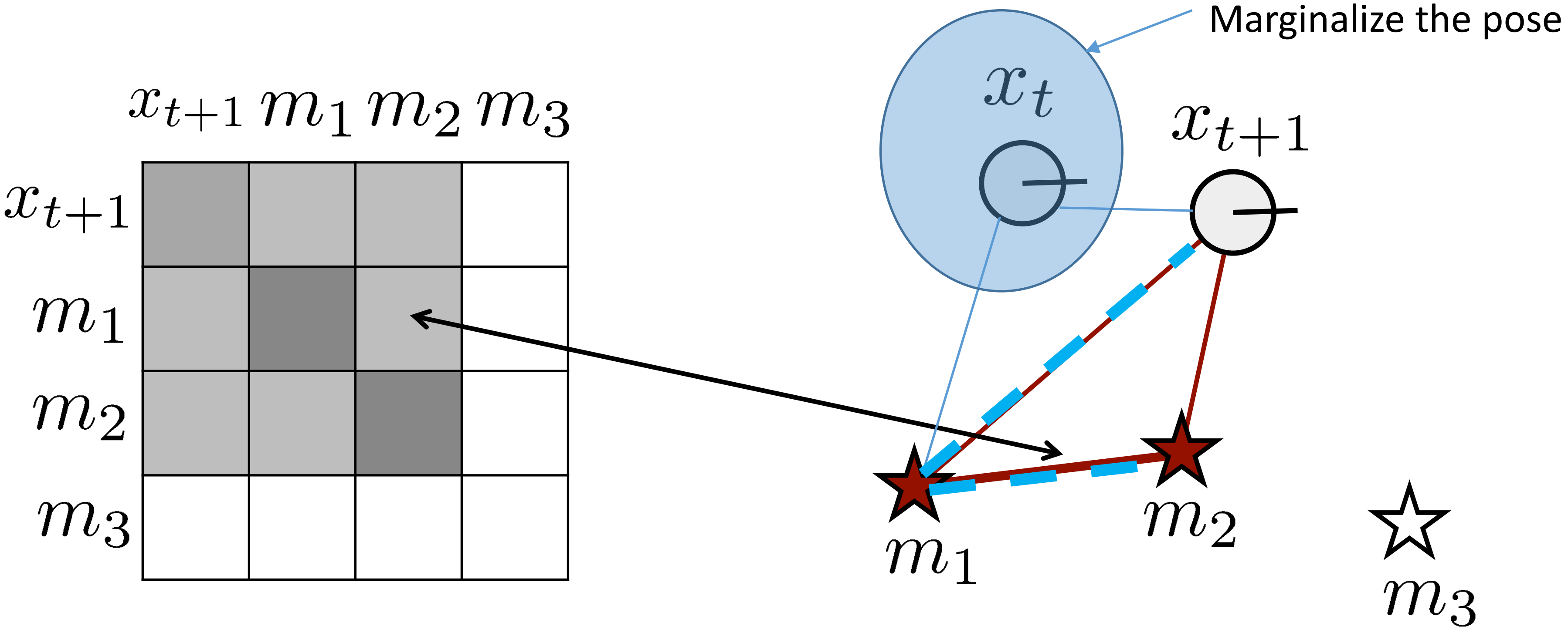
- Effect of measurement update



Slide Credit: Cyrill Stachniss

Intuition behind SEIF-SLAM

- Now the robot moves again, observes landmark 2 but do not observe landmark 1 this time



Slide Credit: Cyrill Stachniss

Intuition behind SEIF-SLAM

- Motion Update: Since we track only the latest pose, this is equivalent to marginalising out the past pose. This introduces new links between the features.

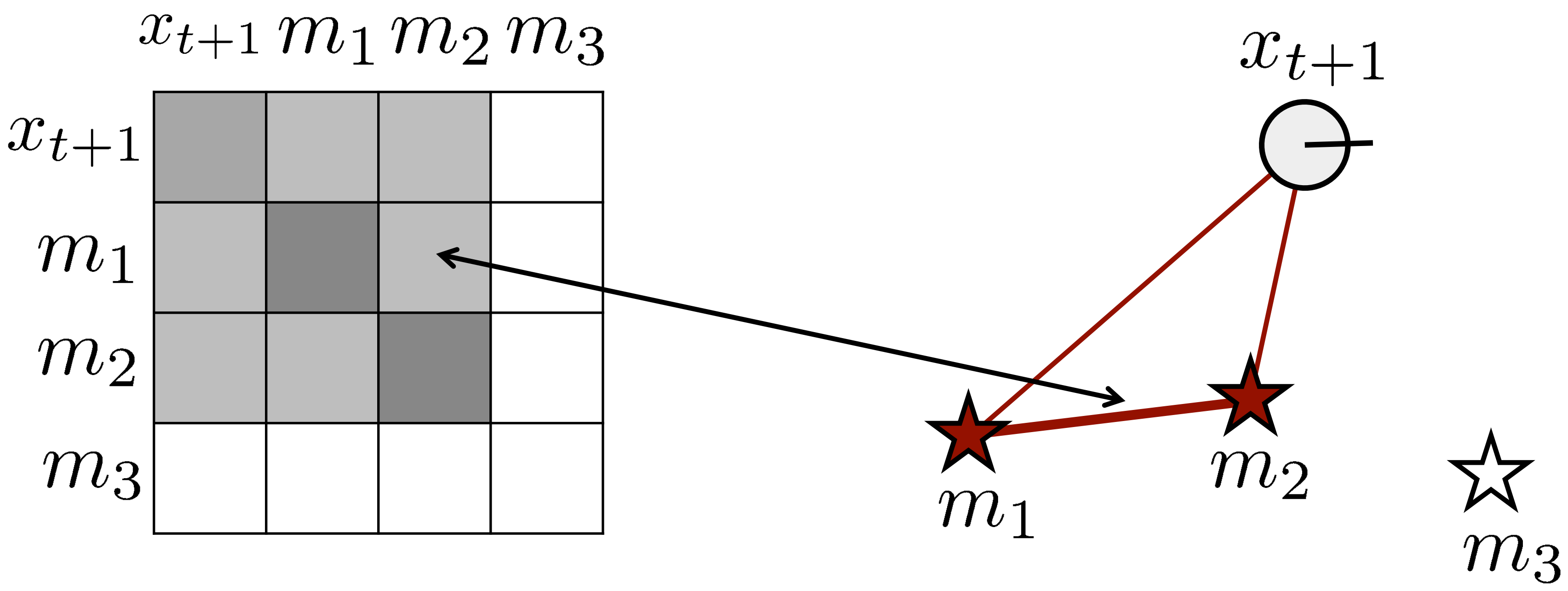


Image Credit: Cyrill Stachniss

Intuition behind SEIF-SLAM

- Motion update: Marginalize out previous state
- The information matrix becomes dense
- The links between the robot pose and features m1 and m2 become weak because of motion uncertainty.
- A new link between features is created. Equivalent to distribution of information.

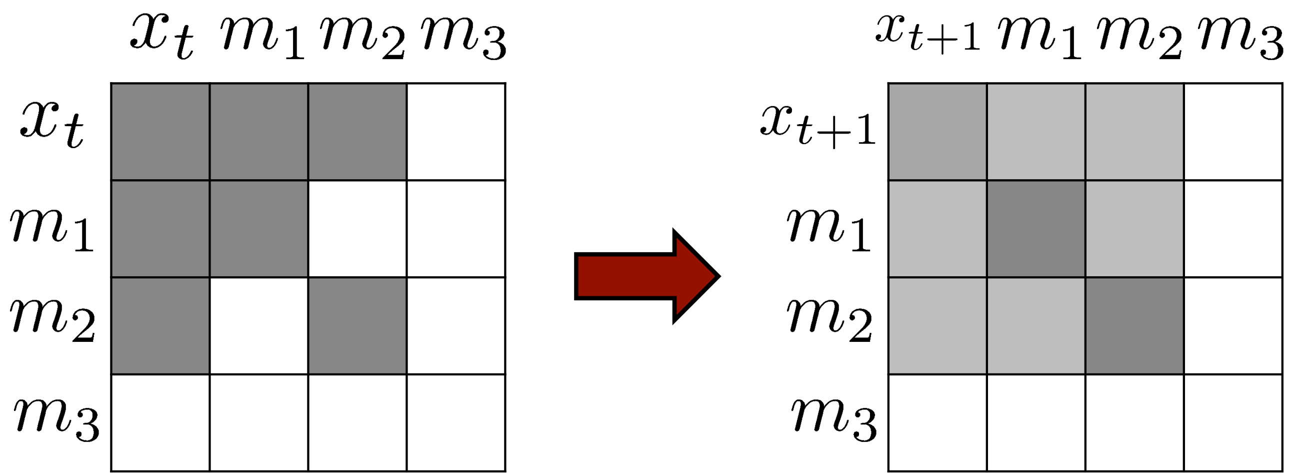
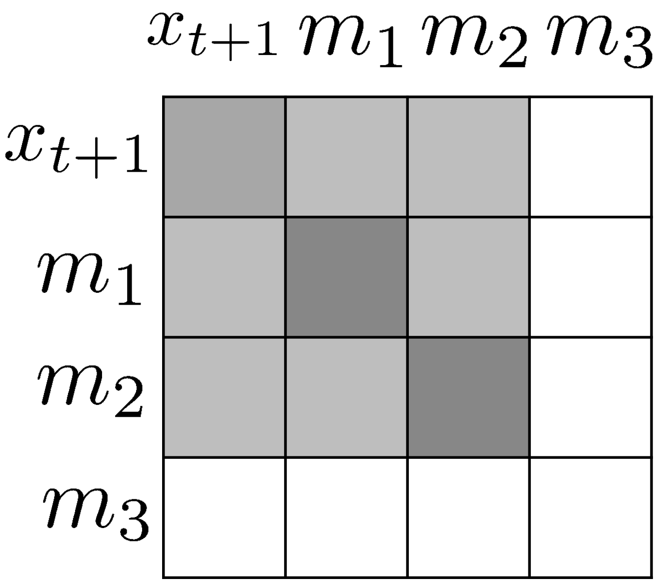
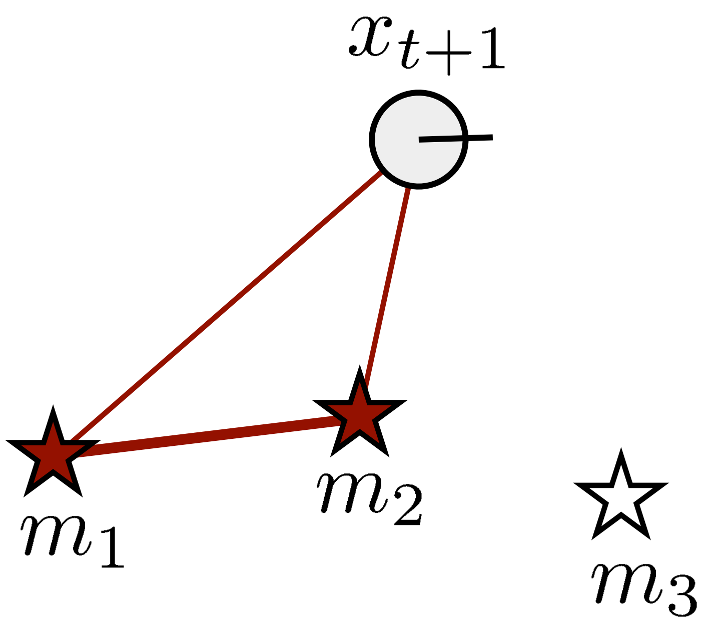


Image Credit: Probabilistic Robotics

Sparsification

- Before Sparsification

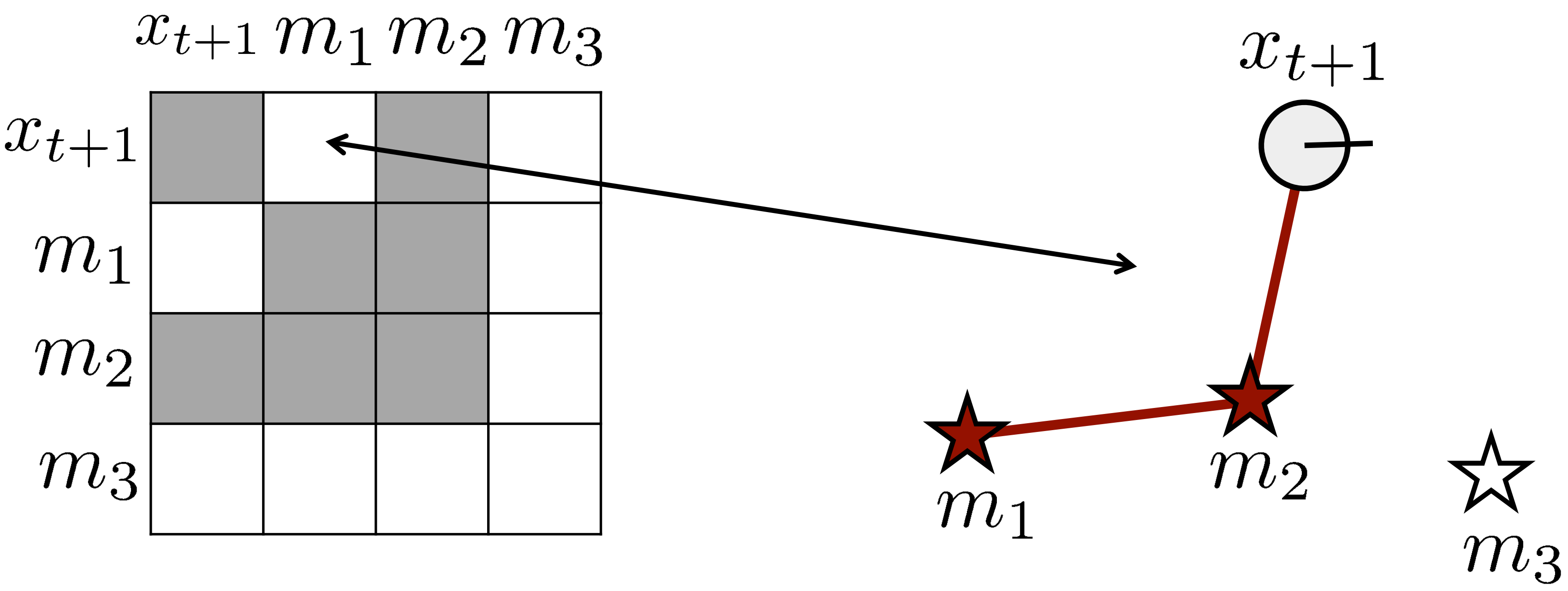




Slide Credit: Cyrill Stachniss

Sparsification

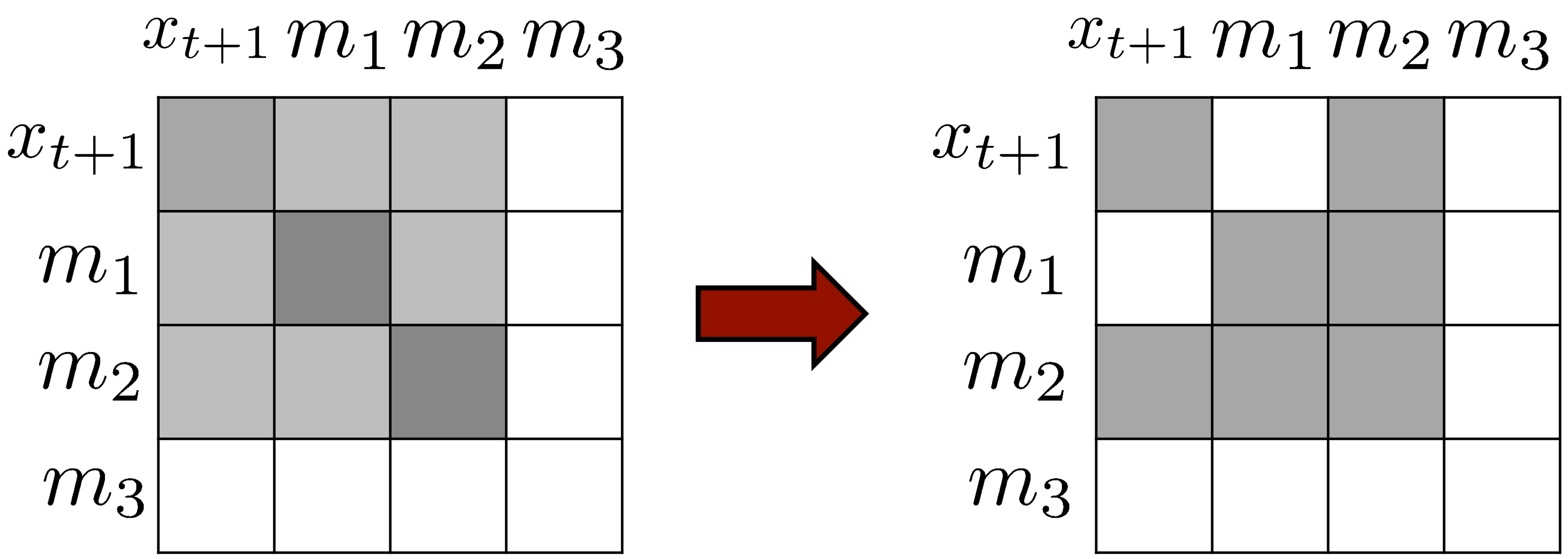
- After Sparsification: This is an approximation, we remove the link between robot state and un-seen landmarks.



Slide Credit: Cyrill Stachniss

Sparsification

- Removing links between robot state and some features / landmarks

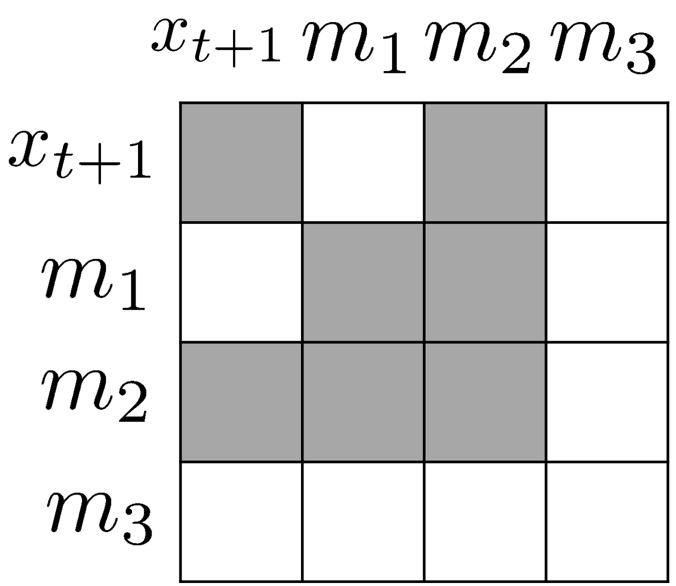


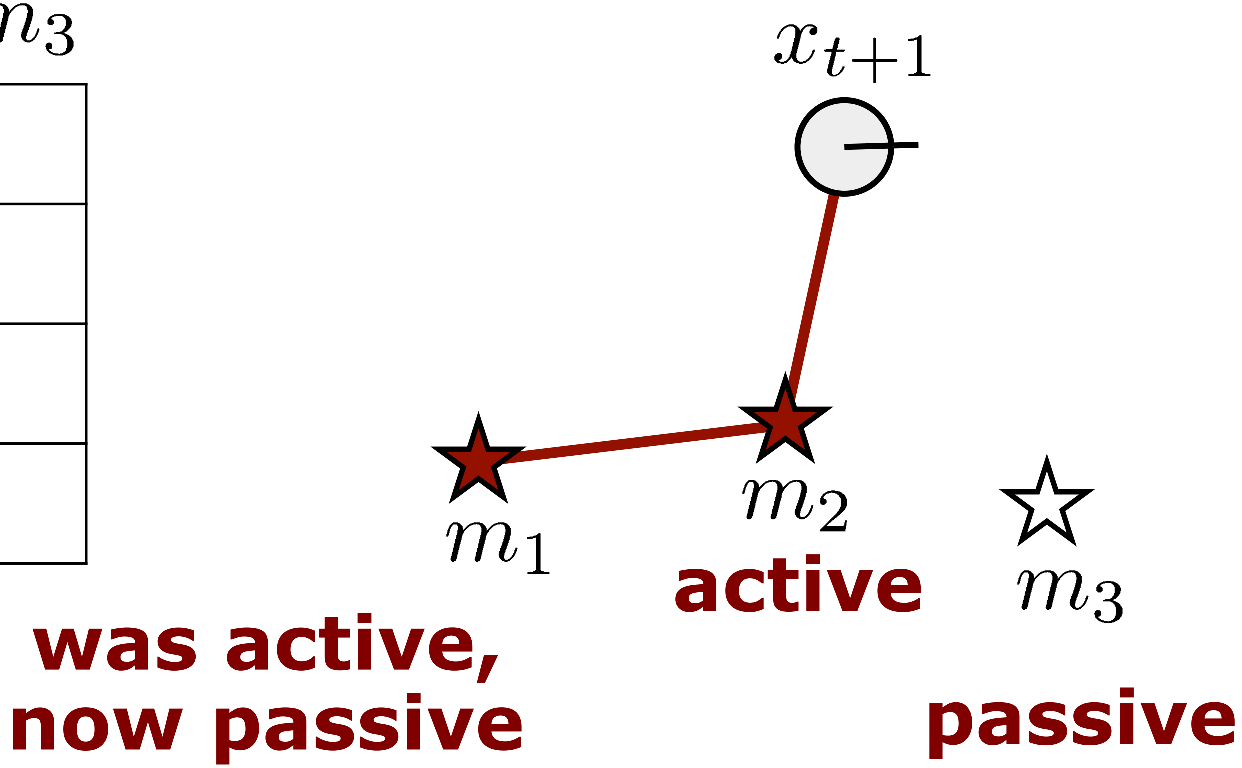
Slide Credit: Cyrill Stachniss

SEIF-SLAM: Active vs Passive landmarks

- Sparsification is the key to the success of SEIF-SLAM.
- In order to maintain sparsity of the information matrix we classify the landmarks into two categories:
- Active Landmarks:
  - These are all the landmarks that are currently observed by the robot
- Passive Landmarks:
  - All other landmarks that are not currently observed

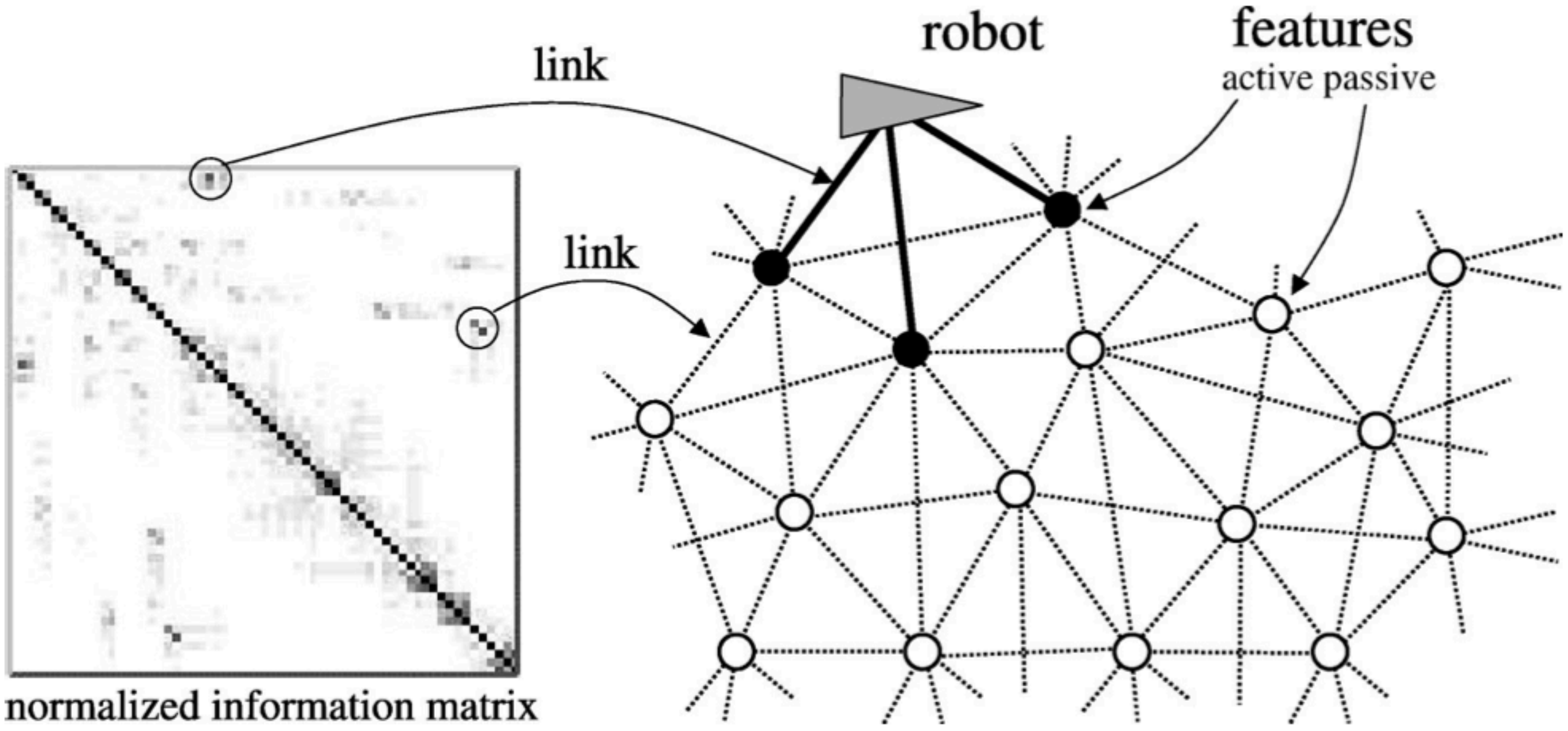
Active vs Passive Landmarks





Slide Credit: Cyrill Stachniss

SEIF - SLAM



Slide Credit: Cyrill Stachniss

SEIF-SLAM Sparsification

- SEIF maintains sparsity of Information Matrix by keeping the links between robot state and active landmarks only !
- Some measurement may add small information between robot state and passive landmarks, those are forced to be zero.
- SEIF performs the sparsification (approximation) step after each iteration.
- Maintaining sparsity allows constant time updates of the information matrix and information vector even for very large matrices (large number of landmarks).

SEIF SLAM

$$\textbf{SEIF\_SLAM}(\xi_{t-1}, \Omega_{t-1}, \mu_{t-1}, u_t, z_t):$$

1: $\bar{\xi}_t, \bar{\Omega}_t, \bar{\mu}_t = \textbf{SEIF\_motion\_update}(\xi_{t-1}, \Omega_{t-1}, \mu_{t-1}, u_t)$

2: $\xi_t, \Omega_t = \textbf{SEIF\_measurement\_update}(\bar{\xi}_t, \bar{\Omega}_t, \bar{\mu}_t, z_t)$

3: $\mu_t = \textbf{SEIF\_update\_state\_estimate}(\xi_t, \Omega_t, \bar{\mu}_t)$

4: $\tilde{\xi}_t, \tilde{\Omega}_t = \textbf{SEIF\_sparsification}(\xi_t, \Omega_t, \mu_t)$

5: $return\ \tilde{\xi}_t, \tilde{\Omega}_t, \mu_t$

Slide Credit: Probabilistic Robotics

SEIF-SLAM example

- Robot moves in 2D plane
- Number of landmarks are known “n”
- Dimension of state space (3+2n)
- Velocity motion model
- Range and bearing sensor
- Robot observes point landmarks
- Data association is known

Motion update

$$\begin{pmatrix} x' \\ y' \\ \theta' \end{pmatrix} = \underbrace{\begin{pmatrix} x \\ y \\ \theta \end{pmatrix} + \begin{pmatrix} -\frac{v_t}{\omega_t}\sin\theta + \frac{v_t}{\omega_t}\sin(\theta+\omega_t\Delta t) \\ \frac{v_t}{\omega_t}\cos\theta - \frac{v_t}{\omega_t}\cos(\theta+\omega_t\Delta t) \\ \omega_t\,\Delta t \end{pmatrix}}_{g_{x,y,\theta}(u_t,(x,y,\theta)^T)}$$

$$\begin{pmatrix} x' \\ y' \\ \theta' \\ \vdots \end{pmatrix} = \underbrace{\begin{pmatrix} x \\ y \\ \theta \\ \vdots \end{pmatrix} + \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0\ldots 0 \\ 0 & 1 & 0 & 0\ldots 0 \\ 0 & 0 & 1 & \underbrace{0\ldots 0}_{2N cols} \end{pmatrix}^T}_{F_x^T} \begin{pmatrix} -\frac{v_t}{\omega_t}\sin\theta + \frac{v_t}{\omega_t}\sin(\theta+\omega_t\Delta t) \\ \frac{v_t}{\omega_t}\cos\theta - \frac{v_t}{\omega_t}\cos(\theta+\omega_t\Delta t) \\ \omega_t\,\Delta t \end{pmatrix}}_{g(u_t,x_t)}$$

Recall EKF-SLAM Prediction step

EKF_SLAM_Prediction$(\mu_{t-1}, \Sigma_{t-1}, u_t, z_t, c_t, R_t)$**:**

2: $$F_x = \begin{pmatrix} 1 & 0 & 0 & 0 \cdots 0 \\ 0 & 1 & 0 & 0 \cdots 0 \\ 0 & 0 & 1 & 0 \cdots 0 \end{pmatrix}$$

3: $$\bar{\mu}_t = \mu_{t-1} + F_x^T \begin{pmatrix} -\frac{v_t}{\omega_t} \sin \mu_{t-1,\theta} + \frac{v_t}{\omega_t} \sin(\mu_{t-1,\theta} + \omega_t \Delta t) \\ \frac{v_t}{\omega_t} \cos \mu_{t-1,\theta} - \frac{v_t}{\omega_t} \cos(\mu_{t-1,\theta} + \omega_t \Delta t) \\ \omega_t \Delta t \end{pmatrix}$$

4: $$G_t = I + F_x^T \begin{pmatrix} 0 & 0 & -\frac{v_t}{\omega_t} \cos \mu_{t-1,\theta} + \frac{v_t}{\omega_t} \cos(\mu_{t-1,\theta} + \omega_t \Delta t) \\ 0 & 0 & -\frac{v_t}{\omega_t} \sin \mu_{t-1,\theta} + \frac{v_t}{\omega_t} \sin(\mu_{t-1,\theta} + \omega_t \Delta t) \\ 0 & 0 & 0 \end{pmatrix} F_x$$

5: $$\bar{\Sigma}_t = G_t \, \Sigma_{t-1} \, G_t^T + \underbrace{F_x^T \, R_t^x \, F_x}_{R_t}$$

Lets compute information matrix first

SEIF-SLAM Motion Update

Algorithm SEIF_motion_update$(\xi_{t-1}, \Omega_{t-1}, \mu_{t-1}, u_t)$**:**

2: $$F_x = \begin{pmatrix} 1 & 0 & 0 & 0 \cdots 0 \\ 0 & 1 & 0 & 0 \cdots 0 \\ 0 & 0 & 1 & \underbrace{0 \cdots 0}_{2N} \end{pmatrix}$$

3: $$\delta = \begin{pmatrix} -\frac{v_t}{\omega_t} \sin \mu_{t-1,\theta} + \frac{v_t}{\omega_t} \sin(\mu_{t-1,\theta} + \omega_t \Delta t) \\ \frac{v_t}{\omega_t} \cos \mu_{t-1,\theta} - \frac{v_t}{\omega_t} \cos(\mu_{t-1,\theta} + \omega_t \Delta t) \\ \omega_t \Delta t \end{pmatrix}$$

4: $$\Delta = \begin{pmatrix} 0 & 0 & \frac{v_t}{\omega_t} \cos \mu_{t-1,\theta} - \frac{v_t}{\omega_t} \cos(\mu_{t-1,\theta} + \omega_t \Delta t) \\ 0 & 0 & \frac{v_t}{\omega_t} \sin \mu_{t-1,\theta} - \frac{v_t}{\omega_t} \sin(\mu_{t-1,\theta} + \omega_t \Delta t) \\ 0 & 0 & 0 \end{pmatrix}$$

Prediction Step

- We need to show that information matrix can be computed efficiently if the previous information is sparse:

$$
\begin{aligned}
\bar{\Omega}_t &= \bar{\Sigma}_t^{-1} \\
&= \left[G_t\,\Omega_{t-1}^{-1}\,G_t^T + R_t\right]^{-1} \\
&= \left[\Phi_t^{-1} + R_t\right]^{-1}
\end{aligned}
$$

where

$$
\begin{aligned}
\Phi_t &= \left[G_t\,\Omega_{t-1}^{-1}\,G_t^T\right]^{-1} \\
&= [G_t^T]^{-1}\,\Omega_{t-1}\,G_t^{-1}
\end{aligned}
$$

Prediction Step

$$\begin{aligned}\bar{\Omega}_t &= \left[\Phi_t^{-1} + R_t\right]^{-1} \\ &= \left[\Phi_t^{-1} + F_x^T\ R_t^x\ F_x\right]^{-1}\end{aligned}$$

- Apply Matrix inversion lemma

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & 0\ldots0 \\ 0 & 1 & 0 & 0\ldots0 \\ 0 & 0 & 1 & \underbrace{0\ldots0}_{2N cols} \end{pmatrix}^T}_{F_x^T}$$

$$\begin{aligned}\bar{\Omega}_t &= \left[\Phi_t^{-1} + R_t\right]^{-1} \\ &= \left[\Phi_t^{-1} + F_x^T\ R_t^x\ F_x\right]^{-1} \\ &= \Phi_t - \Phi_t\ F_x^T \underbrace{(R_t^{x-1} + F_x\ \Phi_t\ F_x^T)^{-1}}_{\text{3X3 matrix}}\ F_x\ \Phi_t\end{aligned}$$

Prediction Step

$$\begin{aligned}\bar{\Omega}_t &= \left[\Phi_t^{-1} + R_t\right]^{-1} \\ &= \left[\Phi_t^{-1} + F_x^T\ R_t^x\ F_x\right]^{-1} \\ &= \Phi_t - \Phi_t\ \underset{\substack{\uparrow \\ \text{Zero except} \\ \text{3x3 block}}}{F_x^T} \underbrace{(R_t^{x-1} + F_x\ \Phi_t\ F_x^T)^{-1}}_{\text{3x3 matrix}} \underset{\substack{\uparrow \\ \text{Zero except} \\ \text{3x3 block}}}{F_x}\ \Phi_t\end{aligned}$$

- **Constant complexity if Φ_t is sparse!**

Slide Credit: Cyrill Stachniss

Prediction Step

- Computing Φ_t

$$\Phi_t = [G_t^T]^{-1} \, \Omega_{t-1} \, G_t^{-1}$$

- Lets assume that Ω_{t-1} is sparse.

$$\begin{aligned} G_t^{-1} &= (I + F_x^T \, \Delta \, F_x)^{-1} \\ &= \begin{pmatrix} \Delta + I_3 & 0 \\ 0 & I_{2N} \end{pmatrix}^{-1} \end{aligned}$$

3X3 2NX2N

Prediction Step

- Therefore

$$\begin{aligned} G_t^{-1} &= (I + F_x^T \, \Delta \, F_x)^{-1} \\ &= \begin{pmatrix} \Delta + I_3 & 0 \\ 0 & I_{2N} \end{pmatrix}^{-1} \\ &= \begin{pmatrix} (\Delta + I_3)^{-1} & 0 \\ 0 & I_{2N} \end{pmatrix} \end{aligned}$$

- True for all block diagonal matrices for which off-diagonal elements are zeros.

Prediction Step

$$
\begin{aligned}
G_t^{-1} &= (I + F_x^T\ \Delta\ F_x)^{-1} \\
&= \begin{pmatrix} \Delta + I_3 & 0 \\ 0 & I_{2N} \end{pmatrix}^{-1} \\
&= \begin{pmatrix} (\Delta + I_3)^{-1} & 0 \\ 0 & I_{2N} \end{pmatrix} \\
&= I_{3+2N} + \begin{pmatrix} (\Delta + I_3)^{-1} - I_3 & 0 \\ 0 & 0 \end{pmatrix} \\
&= I + \underbrace{F_x^T\ [(I + \Delta)^{-1} - I]\ F_x}_{\Psi_t} \\
&= I + \Psi_t
\end{aligned}
$$

Computing $\Phi_t = [G_t^T]^{-1} \, \Omega_{t-1} \, G_t^{-1}$

- We have

$$G_t^{-1} = I + \Psi_t \qquad\qquad [G_t^T]^{-1} = I + \Psi_t^T$$

- with

$$\Psi_t = F_x^T \, \underbrace{\left[(I + \Delta)^{-1} - I\right]}_{\textbf{3x3 matrix}} \, F_x$$

- Ψ_t is zero except of a 3x3 block
- G_t^{-1} is an identity except of a 3x3 block

Slide Credit: Cyrill Stachniss

Prediction Step

- Assuming sparse information matrix Ω_{t-1}

$$
\begin{aligned}
\Phi_t &= [G_t^T]^{-1}\,\Omega_{t-1}\,G_t^{-1} \\
&= (I + \Psi_t^T)\,\Omega_{t-1}\,(I + \Psi_t) \\
&= \Omega_{t-1} + \underbrace{\Psi_t^T\,\Omega_{t-1} + \Omega_{t-1}\,\Psi_t + \Psi_t^T\,\Omega_{t-1}\,\Psi_t}_{\lambda_t} \\
&= \Omega_{t-1} + \underline{\lambda_t}
\end{aligned}
$$

all elements zero except a constant number of entries

Slide Credit: Cyrill Stachniss

Prediction Step

SEIF_motion_update($\xi_{t-1}, \Omega_{t-1}, \mu_{t-1}, u_t$)**:**

2: $F_x = \cdots$

3: $\delta = \cdots$

4: $\Delta = \cdots$

5: $\Psi_t = F_x^T \, [(I + \Delta)^{-1} - I] \, F_x$

6: $\lambda_t = \Psi_t^T \, \Omega_{t-1} + \Omega_{t-1} \, \Psi_t + \Psi_t^T \, \Omega_{t-1} \, \Psi_t$

7: $\Phi_t = \Omega_{t-1} + \lambda_t$

8: $\kappa_t = \Phi_t \, F_x^T (R_t^{-1} + F_x \, \Phi_t \, F_x^T)^{-1} \, F_x \, \Phi_t$

9: $\bar{\Omega}_t = \Phi_t - \kappa_t$

Predicted Mean

- Same as EKF

$$\bar{\mu}_t \quad = \quad \mu_{t-1} + F_x^T \; \delta$$

$$F_x = \begin{pmatrix} 1 & 0 & 0 & 0 \cdots 0 \\ 0 & 1 & 0 & 0 \cdots 0 \\ 0 & 0 & 1 & \underbrace{0 \cdots 0}_{2N} \end{pmatrix}$$

$$\delta = \begin{pmatrix} -\frac{v_t}{\omega_t} \sin \mu_{t-1,\theta} + \frac{v_t}{\omega_t} \; \sin(\mu_{t-1,\theta} + \omega_t \Delta t) \\ \frac{v_t}{\omega_t} \cos \mu_{t-1,\theta} - \frac{v_t}{\omega_t} \; \cos(\mu_{t-1,\theta} + \omega_t \Delta t) \\ \omega_t \Delta t \end{pmatrix}$$

Predicted Information Vector

$$
\begin{aligned}
&\bar{\xi}_t \\
&= \bar{\Omega}_t \, (\mu_{t-1} + F_x^T \, \delta_t) \\
&= \bar{\Omega}_t \, (\Omega_{t-1}^{-1} \, \xi_{t-1} + F_x^T \, \delta_t) \\
&= \bar{\Omega}_t \, \Omega_{t-1}^{-1} \, \xi_{t-1} + \bar{\Omega}_t \, F_x^T \, \delta_t \\
&= (\bar{\Omega}_t \underbrace{- \Phi_t + \Phi_t}_{=0} \underbrace{- \Omega_{t-1} + \Omega_{t-1}}_{=0}) \, \Omega_{t-1}^{-1} \, \xi_{t-1} + \bar{\Omega}_t \, F_x^T \, \delta_t \\
&= (\underbrace{\bar{\Omega}_t - \Phi_t}_{= -\kappa_t} + \underbrace{\Phi_t - \Omega_{t-1}}_{= \lambda_t}) \, \underbrace{\Omega_{t-1}^{-1} \, \xi_{t-1}}_{= \mu_{t-1}} + \underbrace{\Omega_{t-1} \, \Omega_{t-1}^{-1}}_{= I} \, \xi_{t-1} + \bar{\Omega}_t \, F_x^T \, \delta_t \\
&= \xi_{t-1} + (\lambda_t - \kappa_t) \, \mu_{t-1} + \bar{\Omega}_t \, F_x^T \, \delta_t
\end{aligned}
$$

SEIF Prediction Step

SEIF_motion_update($\xi_{t-1}, \Omega_{t-1}, \mu_{t-1}, u_t$)**:**

$F_x = \cdots$

$\delta = \cdots$

$\Delta = \cdots$

$\Psi_t = F_x^T \left[(I + \Delta)^{-1} - I \right] F_x$

$\lambda_t = \Psi_t^T \, \Omega_{t-1} + \Omega_{t-1} \, \Psi_t + \Psi_t^T \, \Omega_{t-1} \, \Psi_t$

$\Phi_t = \Omega_{t-1} + \lambda_t$

$\kappa_t = \Phi_t \, F_x^T (R_t^{-1} + F_x \, \Phi_t \, F_x^T)^{-1} \, F_x \, \Phi_t$

$\bar{\Omega}_t = \Phi_t - \kappa_t$

$\bar{\xi}_t = \xi_{t-1} + (\lambda_t - \kappa_t) \, \mu_{t-1} + \bar{\Omega}_t \, F_x^T \, \delta_t$

$\bar{\mu}_t = \mu_{t-1} + F_x^T \, \delta$

return $\bar{\xi}_t, \bar{\Omega}_t, \bar{\mu}_t$

SEIF SLAM

$\textbf{SEIF\_SLAM}(\xi_{t-1}, \Omega_{t-1}, \mu_{t-1}, u_t, z_t)$:

1: $\bar{\xi}_t, \bar{\Omega}_t, \bar{\mu}_t = \textbf{SEIF\_motion\_update}(\xi_{t-1}, \Omega_{t-1}, \mu_{t-1}, u_t)$

2: $\boxed{\xi_t, \Omega_t = \textbf{SEIF\_measurement\_update}(\bar{\xi}_t, \bar{\Omega}_t, \bar{\mu}_t, z_t)}$

3: $\mu_t = \textbf{SEIF\_update\_state\_estimate}(\xi_t, \Omega_t, \bar{\mu}_t)$

4: $\tilde{\xi}_t, \tilde{\Omega}_t = \textbf{SEIF\_sparsification}(\xi_t, \Omega_t, \mu_t)$

5: $\textit{return}\ \tilde{\xi}_t, \tilde{\Omega}_t, \mu_t$

Measurement Update

- Initialization of landmarks: Similar to EKF

$$\textbf{SEIF\_measurement\_update}(\bar{\xi}_t, \bar{\Omega}_t, \mu_t, z_t)$$

$$Q_t = \begin{pmatrix} {\sigma_r}^2 & 0 \\ 0 & {\sigma_\phi}^2 \end{pmatrix}$$

$$\textit{for all observed features } z_t^i = (r_t^i, \phi_t^i)^T \textit{ do}$$

$$\quad j = c_t^i \qquad \text{(data association)}$$

$$\quad \textit{if landmark } j \textit{ never seen before}$$

$$\qquad \begin{pmatrix} \bar{\mu}_{j,x} \\ \bar{\mu}_{j,y} \end{pmatrix} = \begin{pmatrix} \bar{\mu}_{t,x} \\ \bar{\mu}_{t,y} \end{pmatrix} + \begin{pmatrix} r_t^i \ \cos(\phi_t^i + \bar{\mu}_{t,\theta}) \\ r_t^i \ \sin(\phi_t^i + \bar{\mu}_{t,\theta}) \end{pmatrix}$$

$$\quad \textit{endif}$$

$$\quad \delta = \begin{pmatrix} \delta_x \\ \delta_y \end{pmatrix} = \begin{pmatrix} \bar{\mu}_{j,x} - \bar{\mu}_{t,x} \\ \bar{\mu}_{j,y} - \bar{\mu}_{t,y} \end{pmatrix}$$

$$\quad q = \delta^T \delta$$

$$\quad \hat{z}_t^i = \begin{pmatrix} \sqrt{q} \\ \operatorname{atan2}(\delta_y, \delta_x) - \bar{\mu}_{t,\theta} \end{pmatrix}$$

Slide Credit: Cyrill Stachniss

Measurement Update: Same as EIF

$$H_t^i = \frac{1}{q}\begin{pmatrix} -\sqrt{q}\delta_x & -\sqrt{q}\delta_y & 0 & 0\ldots0 & +\sqrt{q}\delta_x & \sqrt{q}\delta_y & 0\ldots0 \\ \delta_y & -\delta_x & -q & \underbrace{0\ldots0}_{2j-2} & -\delta_y & +\delta_x & \underbrace{0\ldots0}_{2N-2j} \end{pmatrix}$$

endfor

$\xi_t = \bar{\xi}_t + \sum_i H_t^{iT}\; Q_t^{-1}\; [z_t^i - \hat{z}_t^i + H_t^i\; \mu_t]$

$\Omega_t = \bar{\Omega}_t + \sum_i H_t^{iT}\; Q_t^{-1}\; H_t^i$

return ξ_t, Ω_t

$$\xi_t \quad = \quad \bar{\xi}_t + \sum_i H_t^{iT}\; Q_t^{-1}\; [z_t^i - \hat{z}_t^i + H_t^i\; \mu_t]$$

$$\Omega_t \quad = \quad \bar{\Omega}_t + \sum H_t^{iT}\; Q_t^{-1}\; H_t^i$$

H_t is mostly zero and Q_t is a 2X2 matrix ➔ Update step is constant time !

Slide Credit: Cyrill Stachniss

SEIF SLAM

$\textbf{SEIF\_SLAM}(\xi_{t-1}, \Omega_{t-1}, \mu_{t-1}, u_t, z_t)$:

1: $\bar{\xi}_t, \bar{\Omega}_t, \bar{\mu}_t = \textbf{SEIF\_motion\_update}(\xi_{t-1}, \Omega_{t-1}, \mu_{t-1}, u_t)$

2: $\xi_t, \Omega_t = \textbf{SEIF\_measurement\_update}(\bar{\xi}_t, \bar{\Omega}_t, \bar{\mu}_t, z_t)$

3: $\mu_t = \textbf{SEIF\_update\_state\_estimate}(\xi_t, \Omega_t, \bar{\mu}_t)$

4: $\tilde{\xi}_t, \tilde{\Omega}_t = \textbf{SEIF\_sparsification}(\xi_t, \Omega_t, \mu_t)$

5: *return* $\tilde{\xi}_t, \tilde{\Omega}_t, \mu_t$

Update Mean

- We need to keep track of mean because of following reasons:

1. Mean is used for linearization of motion model
2. It is also used for linearization of measurement model
3. It is also used in sparsification step

- However, we never need the full mean vector. We only need an estimate of the robot pose and all the active landmarks. Exact recovery of mean from information vector is computationally expensive

$$\mu = \Omega^{-1}\xi$$

Approximate Mean Recovery

- The key insight is that the information matrix is sparse, therefore we can formulate the mean recovery as an optimization problem.

$$\hat{\mu} \quad = \quad \operatorname{argmax} p(\mu)$$

$$p(\mu) \quad = \quad \eta \ \exp\left\{-\tfrac{1}{2}\ \mu^T\ \Omega\ \mu + \xi^T\ \mu\right\}$$

$$\frac{\partial p(\mu)}{\partial \mu} \quad = \quad \eta\ (-\Omega\ \mu + \xi)\ \exp\left\{-\tfrac{1}{2}\ \mu^T\ \Omega\ \mu + \xi^T\ \mu\right\} \quad = \quad 0$$

$$\mu = \Omega^{-1}\xi$$

Approximate Mean Recovery

- Solve the following using gradient descent:

$$\begin{aligned}\hat{\mu} &= \underset{\mu}{\operatorname{argmax}}\ \exp\left\{-\tfrac{1}{2}\,\mu^T\,\Omega\,\mu + \xi^T\,\mu\right\} \\ &= \underset{\mu}{\operatorname{argmin}}\ \tfrac{1}{2}\,\mu^T\,\Omega\,\mu - \xi^T\,\mu\end{aligned}$$

- Since we only need the estimate of robot pose and active landmarks, the dimensionality of the optimization problem is small ! For details see the probabilistic robotics book.

SEIF SLAM

$\textbf{SEIF\_SLAM}(\xi_{t-1}, \Omega_{t-1}, \mu_{t-1}, u_t, z_t)$:

1: $\bar{\xi}_t, \bar{\Omega}_t, \bar{\mu}_t = \textbf{SEIF\_motion\_update}(\xi_{t-1}, \Omega_{t-1}, \mu_{t-1}, u_t)$

2: $\xi_t, \Omega_t = \textbf{SEIF\_measurement\_update}(\bar{\xi}_t, \bar{\Omega}_t, \bar{\mu}_t, z_t)$

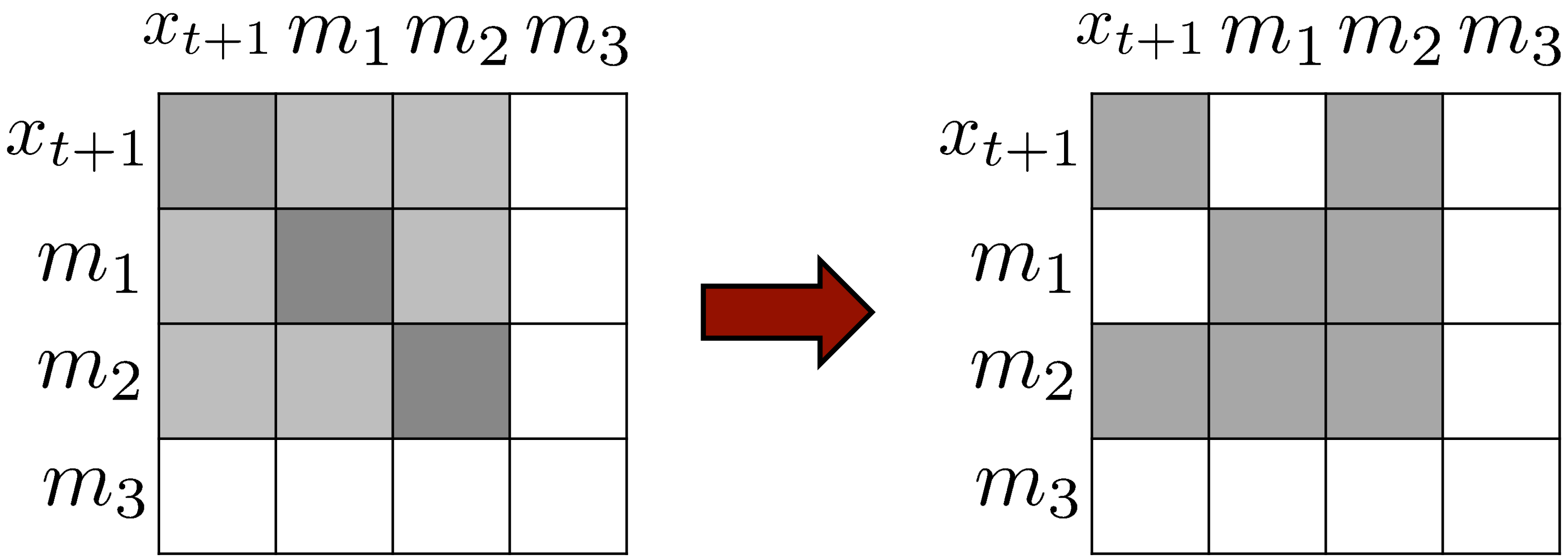
3: $\mu_t = \textbf{SEIF\_update\_state\_estimate}(\xi_t, \Omega_t, \bar{\mu}_t)$

4: $\tilde{\xi}_t, \tilde{\Omega}_t = \textbf{SEIF\_sparsification}(\xi_t, \Omega_t, \mu_t)$

5: $return\ \tilde{\xi}_t, \tilde{\Omega}_t, \mu_t$

Sparsification

- Ignoring some links in the information matrix.
- Assuming “conditional independence” of some random variables



Slide Credit: Cyrill Stachniss

Sparsification: General Idea

- Suppose we are given a joint distribution between three random variables:

$$p(a, b, c)$$

- To sparsify this distribution means, to approximate ‘p’ by a distribution $\tilde{p}$ such that ‘a’ and ‘b’ are conditionally independent given ‘c’.

- In multivariate Gaussian distribution this conditional independence is equivalent to the absence of a direct link between ’a’ and ‘b’.

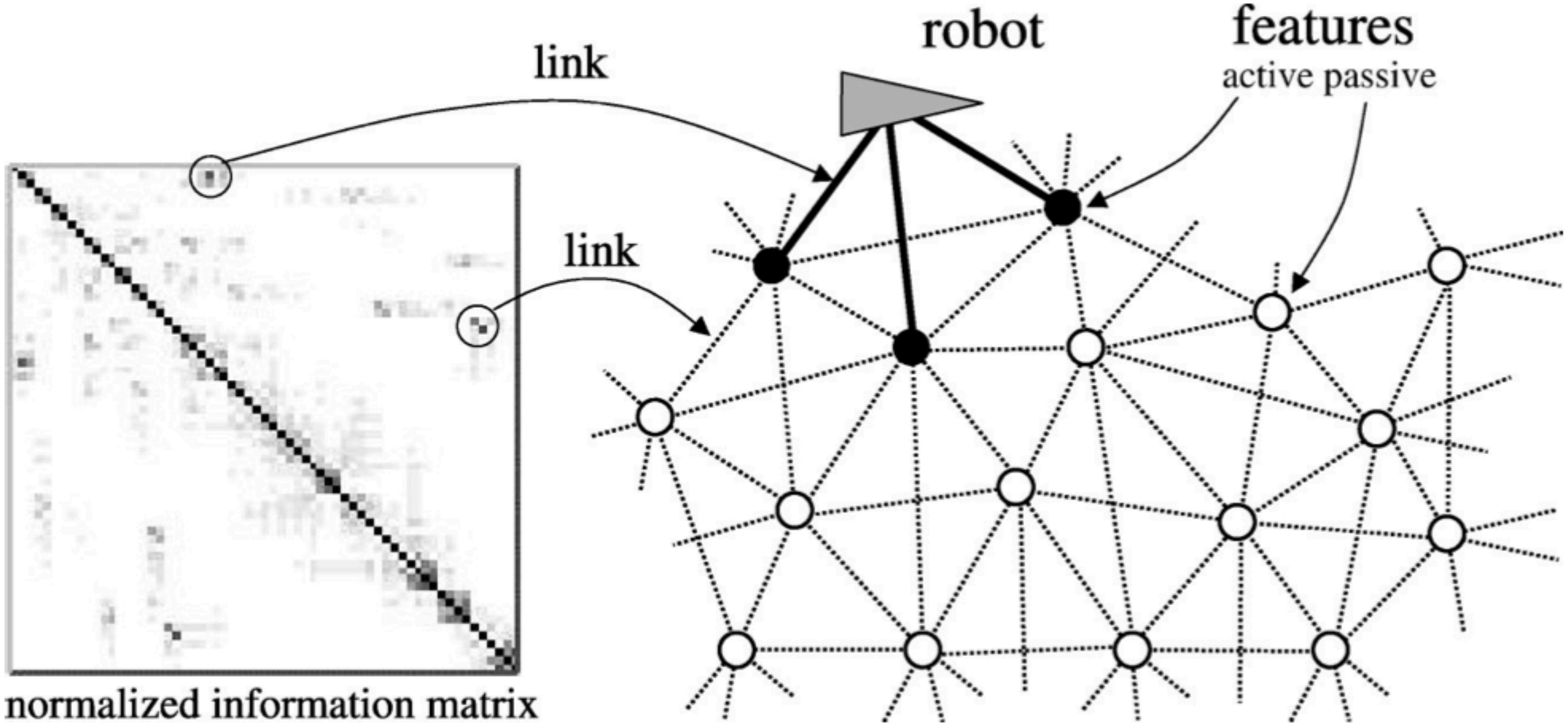
Sparsification: General Idea

$$
\begin{aligned}
p(a,b,c) &= p(a \mid b,c)\, p(b \mid c)\, p(c) \\
&\simeq p(a \mid c)\, p(b \mid c)\, p(c) \\
&= p(a \mid c)\, \frac{p(c)}{p(c)}\, p(b \mid c)\, p(c) \\
&= \frac{p(a,c)\, p(b,c)}{p(c)}
\end{aligned}
$$

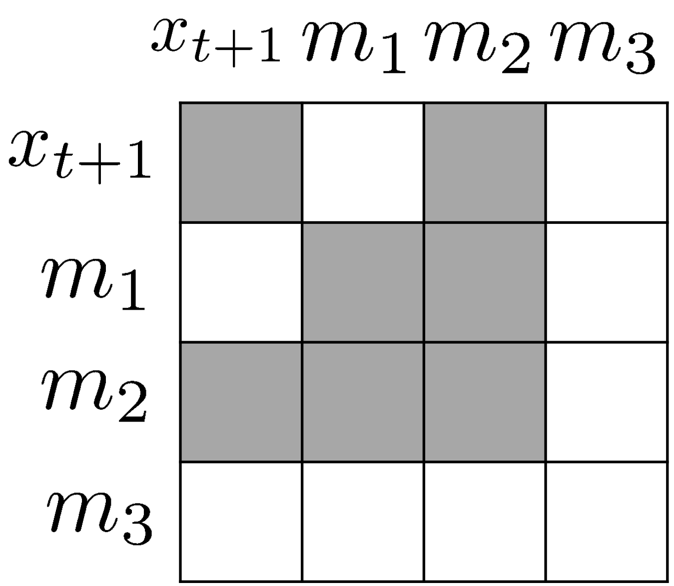
approximation

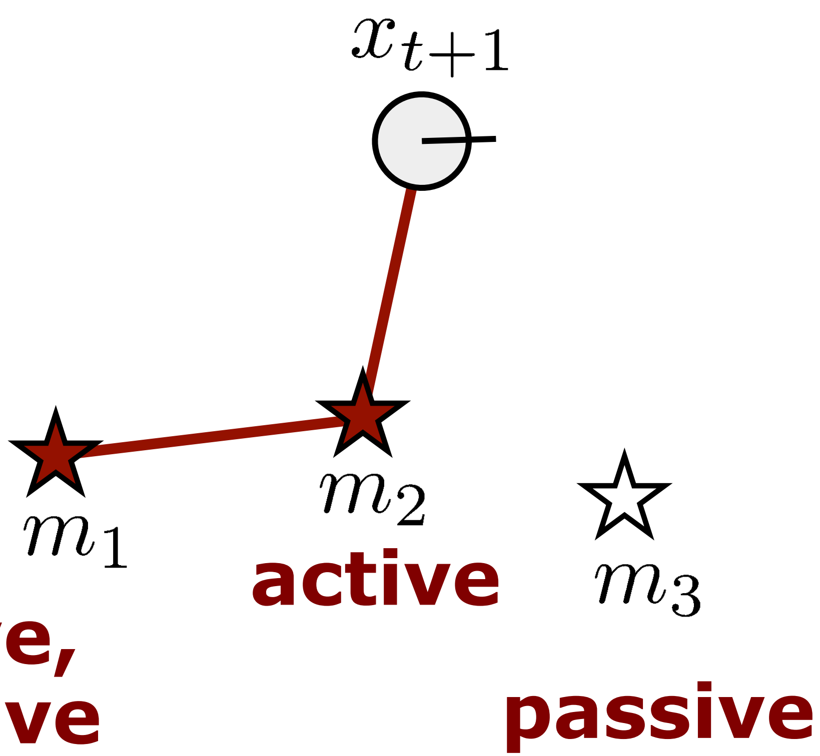
- No direct dependence between ‘a’ and ‘b’ !

Sparsification in SEIF



Active vs Passive Landmarks





Slide Credit: Cyrill Stachniss

Sparsification in SEIF

- For sparsification we partition the set of all features/landmarks into three disjoint sets

$$m \quad = \quad \underbrace{m^{+}}_{\text{active}} \; + \; \underbrace{m^{0}}_{\text{active to passive}} \; + \; \underbrace{m^{-}}_{\text{passive}}$$

- We remove the links between the **robot pose** and the landmarks that transition from **active → passive**

Sparsification in SEIF

- The posterior can be factored as

$$
\begin{aligned}
& p(y_t \mid z_{1:t}, u_{1:t}, c_{1:t}) \\
= \quad & p(x_t, m^0, m^+, m^- \mid z_{1:t}, u_{1:t}, c_{1:t}) \\
= \quad & p(x_t \mid m^0, m^+, m^-, z_{1:t}, u_{1:t}, c_{1:t})\, p(m^0, m^+, m^- \mid z_{1:t}, u_{1:t}, c_{1:t}) \\
= \quad & p(x_t \mid m^0, m^+, m^- = 0, z_{1:t}, u_{1:t}, c_{1:t})\, p(m^0, m^+, m^- \mid z_{1:t}, u_{1:t}, c_{1:t})
\end{aligned}
$$

- In the last step we have used the fact that if we know all the active features (including those that were active and now became passive) then the robot pose does not depend on the passive features and it can be set to any arbitrary value (chosen to be 0 here).

Sparsification

- Following the general sparsification idea we can rewrite the approximate posterior as:

$$\tilde{p}(x_t, m \mid z_{1:t}, u_{1:t}, c_{1:t})$$

$$= \quad \frac{p(x_t, m^+ \mid m^- = 0, z_{1:t}, u_{1:t}, c_{1:t})}{p(m^+ \mid m^- = 0, z_{1:t}, u_{1:t}, c_{1:t})}\, p(m^0, m^+, m^- \mid z_{1:t}, u_{1:t}, c_{1:t})$$

- Now we need to find the information matrix for the given approximated posterior. Since all the distributions are Gaussians

$$\tilde{\Omega}_t \quad = \quad \Omega_t^1 - \Omega_t^2 + \Omega_t^3$$

Sparsification

- Consider the information matrix of $p(x_t, m^0, m^+ \mid m^- = 0)$

$$\Omega_t^0 \quad = \quad F_{x,m^+,m^0}\, F^T_{x,m^+,m^0}\, \Omega_t\, F_{x,m^+,m^0}\, F^T_{x,m^+,m^0}$$

- This is obtained by extracting the sub-matrix of all state variables except the passive landmarks. Recall

	Marginalisation	Conditioning
INFORMATION FORM	$\boldsymbol{\eta} = \boldsymbol{\eta}_\alpha - \Lambda_{\alpha\beta}\Lambda_{\beta\beta}^{-1}\boldsymbol{\eta}_\beta$	$\boldsymbol{\eta}' = \boldsymbol{\eta}_\alpha - \Lambda_{\alpha\beta}\boldsymbol{\beta}$
	$\Lambda = \Lambda_{\alpha\alpha} - \Lambda_{\alpha\beta}\Lambda_{\beta\beta}^{-1}\Lambda_{\beta\alpha}$	$\Lambda' = \Lambda_{\alpha\alpha}$

Sparsification

- Now marginalising $p(x_t, m^0, m^+ \mid m^- = 0)$ gives

$$p(x_t, m^+ \mid m^- = 0, z_{1:t}, u_{1:t}, c_{1:t})$$

$$\Omega_t^1 = \Omega_t^0 - \Omega_t^0\, F_{m_0}\, (F_{m_0}^T\, \Omega_t^0\, F_{m_0})^{-1}\, F_{m_0}^T\, \Omega_t^0$$

$$p(m^+ \mid m^- = 0, z_{1:t}, u_{1:t}, c_{1:t})$$

$$\Omega_t^2 = \Omega_t^0 - \Omega_t^0\, F_{x,m_0}\, (F_{x,m_0}^T\, \Omega_t^0\, F_{x,m_0})^{-1}\, F_{x,m_0}^T\, \Omega_t^0$$

$$p(m^0, m^+, m^- \mid z_{1:t}, u_{1:t}, c_{1:t})$$

$$\Omega_t^3 = \Omega_t - \Omega_t F_x (F_x^T \Omega_t F_x)^{-1} F_x^T \Omega_t$$

Sparsification

- Approximated information vector

$$
\begin{aligned}
\tilde{\xi}_t &= \tilde{\Omega}_t \, \mu_t \\
&= (\Omega_t - \Omega_t + \tilde{\Omega}_t) \, \mu_t \\
&= \Omega_t \, \mu_t + (\tilde{\Omega}_t - \Omega_t) \, \mu_t \\
&= \xi_t + (\tilde{\Omega}_t - \Omega_t) \, \mu_t
\end{aligned}
$$

SEIF: Motion Update

Algorithm SEIF_motion_update($\xi_{t-1}, \Omega_{t-1}, \mu_{t-1}, u_t$)**:**

$$F_x = \begin{pmatrix} 1 & 0 & 0 & 0 \cdots 0 \\ 0 & 1 & 0 & 0 \cdots 0 \\ 0 & 0 & 1 & \underbrace{0 \cdots 0}_{2N} \end{pmatrix}$$

$$\delta = \begin{pmatrix} -\frac{v_t}{\omega_t} \sin \mu_{t-1,\theta} + \frac{v_t}{\omega_t} \sin(\mu_{t-1,\theta} + \omega_t \Delta t) \\ \frac{v_t}{\omega_t} \cos \mu_{t-1,\theta} - \frac{v_t}{\omega_t} \cos(\mu_{t-1,\theta} + \omega_t \Delta t) \\ \omega_t \Delta t \end{pmatrix}$$

$$\Delta = \begin{pmatrix} 0 & 0 & \frac{v_t}{\omega_t} \cos \mu_{t-1,\theta} - \frac{v_t}{\omega_t} \cos(\mu_{t-1,\theta} + \omega_t \Delta t) \\ 0 & 0 & \frac{v_t}{\omega_t} \sin \mu_{t-1,\theta} - \frac{v_t}{\omega_t} \sin(\mu_{t-1,\theta} + \omega_t \Delta t) \\ 0 & 0 & 0 \end{pmatrix}$$

$\Psi_t = F_x^T \; [(I + \Delta)^{-1} - I] \; F_x$

$\lambda_t = \Psi_t^T \; \Omega_{t-1} + \Omega_{t-1} \; \Psi_t + \Psi_t^T \; \Omega_{t-1} \; \Psi_t$

$\Phi_t = \Omega_{t-1} + \lambda_t$

$\kappa_t = \Phi_t \; F_x^T (R_t^{-1} + F_x \; \Phi_t \; F_x^T)^{-1} \; F_x \; \Phi_t$

$\bar{\Omega}_t = \Phi_t - \kappa_t$

$\bar{\xi}_t = \xi_{t-1} + (\lambda_t - \kappa_t) \; \mu_{t-1} + \bar{\Omega}_t \; F_x^T \; \delta_t$

$\bar{\mu}_t = \mu_{t-1} + F_x^T \; \delta$

return $\bar{\xi}_t, \bar{\Omega}_t, \bar{\mu}_t$

SEIF Measurement Update

Algorithm SEIF_measurement_update($\bar{\xi}_t, \bar{\Omega}_t, \mu_t, z_t, c_t$)**:**

$$Q_t = \begin{pmatrix} \sigma_r & 0 & 0 \\ 0 & \sigma_\phi & 0 \\ 0 & 0 & \sigma_s \end{pmatrix}$$

for all observed features $z_t^i = (r_t^i \;\; \phi_t^i \;\; s_t^i)^T$ *do*

$j = c_t^i$

if landmark j *never seen before*

$$\begin{pmatrix} \mu_{j,x} \\ \mu_{j,y} \\ \mu_{j,s} \end{pmatrix} = \begin{pmatrix} \mu_{t,x} \\ \mu_{t,y} \\ s_t^i \end{pmatrix} + r_t^i \begin{pmatrix} \cos(\phi_t^i + \mu_{t,\theta}) \\ \sin(\phi_t^i + \mu_{t,\theta}) \\ 0 \end{pmatrix}$$

endif

$$\delta = \begin{pmatrix} \delta_x \\ \delta_y \end{pmatrix} = \begin{pmatrix} \mu_{j,x} - \mu_{t,x} \\ \mu_{j,y} - \mu_{t,y} \end{pmatrix}$$

$q = \delta^T \delta$

$$\hat{z}_t^i = \begin{pmatrix} \sqrt{q} \\ \operatorname{atan2}(\delta_y, \delta_x) - \mu_{t,\theta} \\ \mu_{j,s} \end{pmatrix}$$

$$H_t^i = \frac{1}{q} \begin{pmatrix} \sqrt{q}\delta_x & -\sqrt{q}\delta_y & 0 & 0\cdots0 & -\sqrt{q}\delta_x & \sqrt{q}\delta_y & 0 & 0\cdots0 \\ \delta_y & \delta_x & -1 & 0\cdots0 & -\delta_y & -\delta_x & 0 & 0\cdots0 \\ 0 & 0 & 0 & \underbrace{0\cdots0}_{3j-3} & 0 & 0 & 1 & \underbrace{0\cdots0}_{3j} \end{pmatrix}$$

endfor

$\xi_t = \bar{\xi}_t + \sum_i H_t^{iT} \, Q_t^{-1} \, [z_t^i - \hat{z}_t^i - H_t^i \, \mu_t]$

$\Omega_t = \bar{\Omega}_t + \sum_i H_t^{iT} \, Q_t^{-1} \, H_t^i$

return ξ_t, Ω_t

SEIF Mean Recovery

Algorithm SEIF_update_state_estimate($\bar{\xi}_t, \bar{\Omega}_t, \bar{\mu}_t$):

for a small set of map features m_i *do*

$$F_i = \begin{pmatrix} 0 \cdots 0 & 1 & 0 & 0 \cdots 0 \\ \underbrace{0 \cdots 0}_{2(N-i)} & 0 & 1 & \underbrace{0 \cdots 0}_{2(i-1)x} \end{pmatrix}$$

$$\mu_{i,t} = (F_i \; \Omega_t \; F_i^T)^{-1} \; F_i \; [\xi_t - \Omega_t \; \bar{\mu}_t + \Omega_t \; F_i^T \; F_i \; \bar{\mu}_t]$$

endfor

for all other map features m_i *do*

$$\mu_{i,t} = \bar{\mu}_{i,t}$$

endfor

$$F_x = \begin{pmatrix} 1 & 0 & 0 & 0 \cdots 0 \\ 0 & 1 & 0 & 0 \cdots 0 \\ 0 & 0 & 1 & \underbrace{0 \cdots 0}_{2N} \end{pmatrix}$$

$$\mu_{x,t} = (F_x \; \Omega_t \; F_x^T)^{-1} \; F_x \; [\xi_t - \Omega_t \; \bar{\mu}_t + \Omega_t \; F_x^T \; F_x \; \bar{\mu}_t]$$

return μ_t

SEIF Sparsification

Algorithm SEIF_sparsification(ξ_t, Ω_t):

define F_{m_0}, F_{x,m_0}, F_x *as projection matrices from* y_t *to* m_0, $\{x, m_0\}$, *and* x, *respectively*

$$\begin{aligned}\tilde{\Omega}_t \;=\;& \Omega_t - \Omega_t^0\, F_{m_0}\, (F_{m_0}^T\, \Omega_t^0\, F_{m_0})^{-1}\, F_{m_0}^T\, \Omega_t^0 \\ & + \Omega_t^0\, F_{x,m_0}\, (F_{x,m_0}^T\, \Omega_t^0\, F_{x,m_0})^{-1}\, F_{x,m_0}^T\, \Omega_t^0 \\ & - \Omega_t\, F_x\, (F_x^T\, \Omega_t F_x)^{-1}\, F_x^T\, \Omega_t\end{aligned}$$

$$\tilde{\xi}_t \;=\; \xi_t \;+\; \mu_t\, (\tilde{\Omega}_t \;-\; \Omega_t)$$

return $\tilde{\xi}_t, \tilde{\Omega}_t$

SEIF SLAM vs EKF SLAM

- Roughly **constant time** complexity vs. quadratic complexity of the EKF
- **Linear memory** complexity vs. quadratic complexity of the EKF
- SEIF SLAM is **less accurate** than EKF SLAM
- Sparsification and mean recovery are approximations.

SEIF SLAM experiment

Image Credit: S Thrun

SEIF vs EKF Computation Time

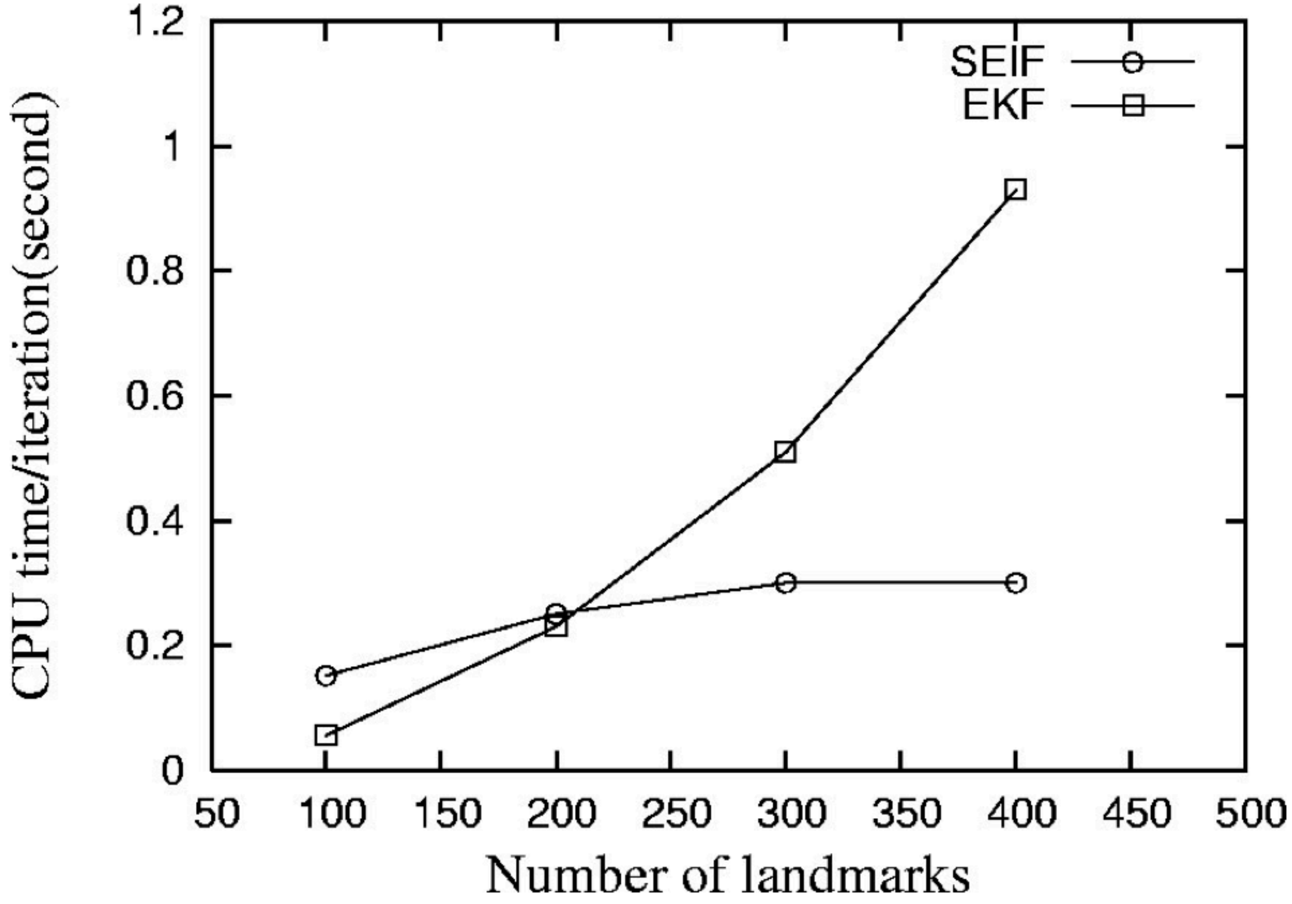


Image Credit: S Thrun

SEIF vs EKF memory requirement

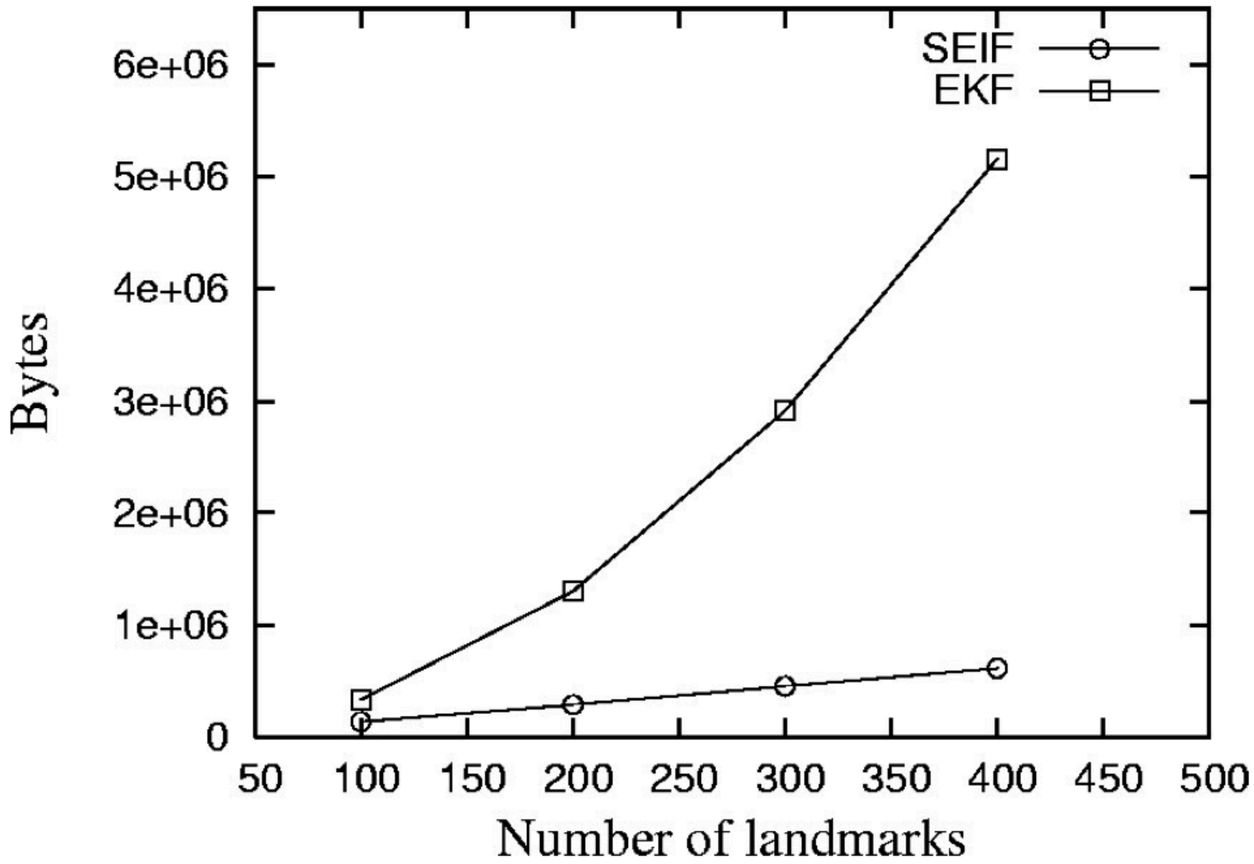


Image Credit: S Thrun

SEIF vs EKF

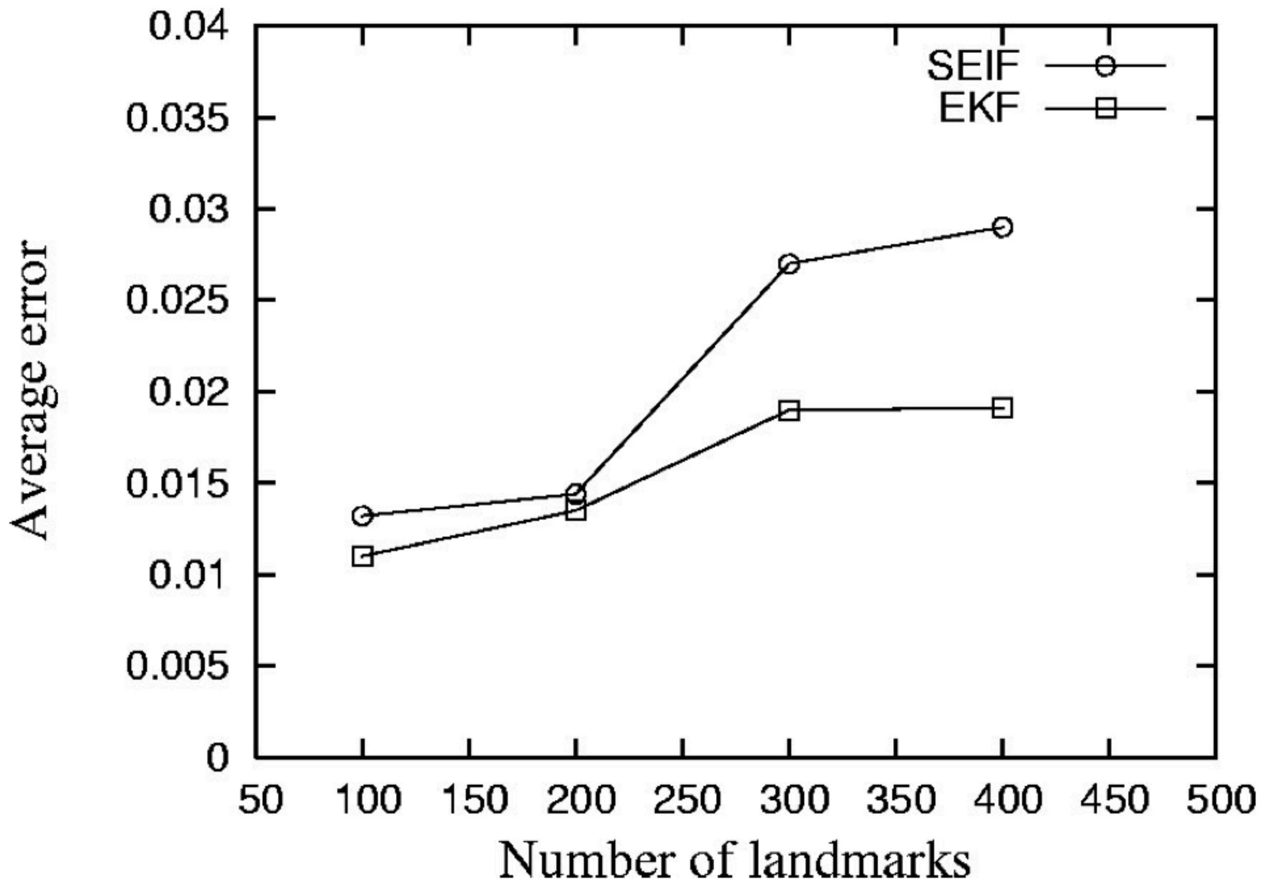


Image Credit: S Thrun

Influence of Active Features on Computation Time

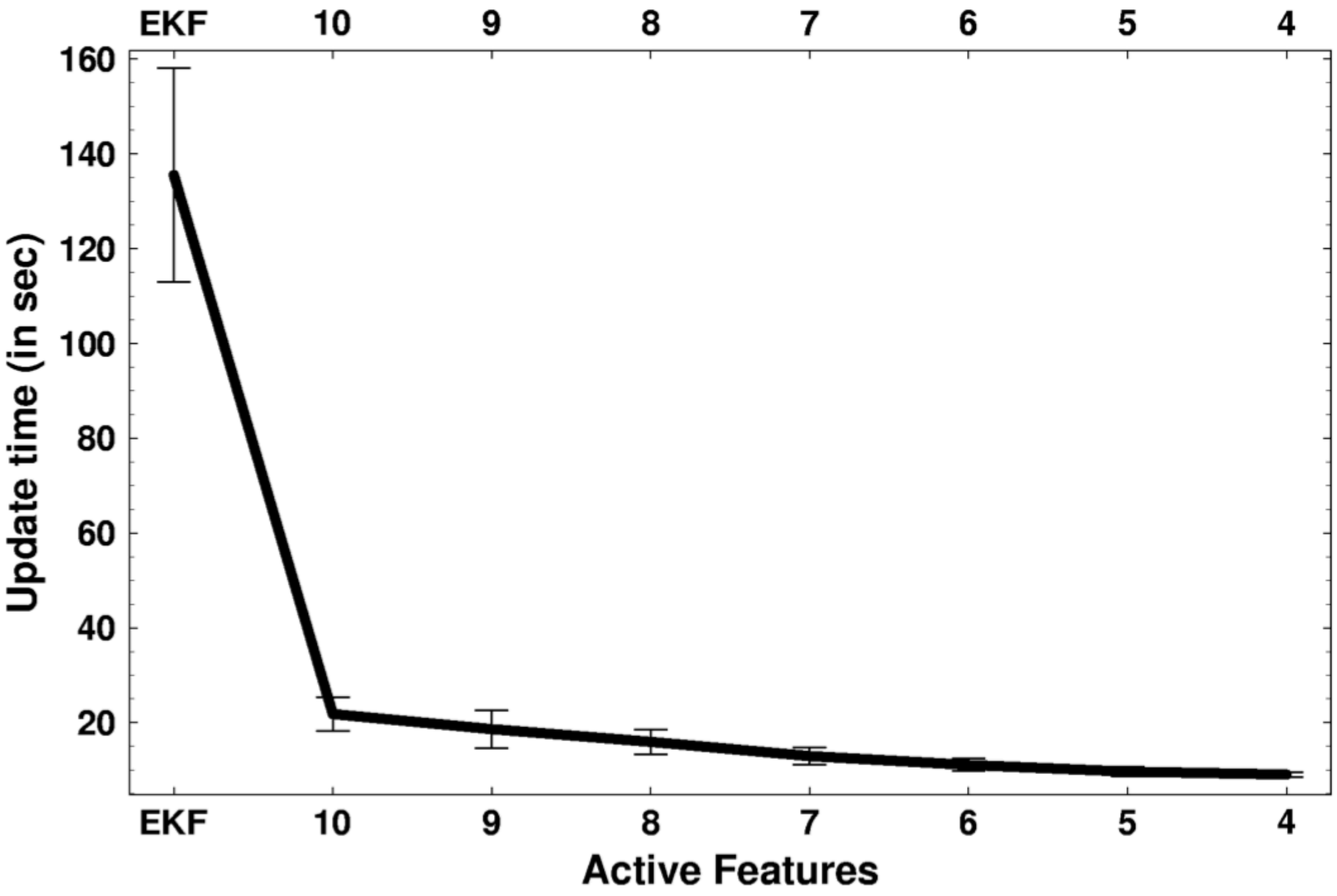


Image Credit: S Thrun

Influence of Active Features on Error

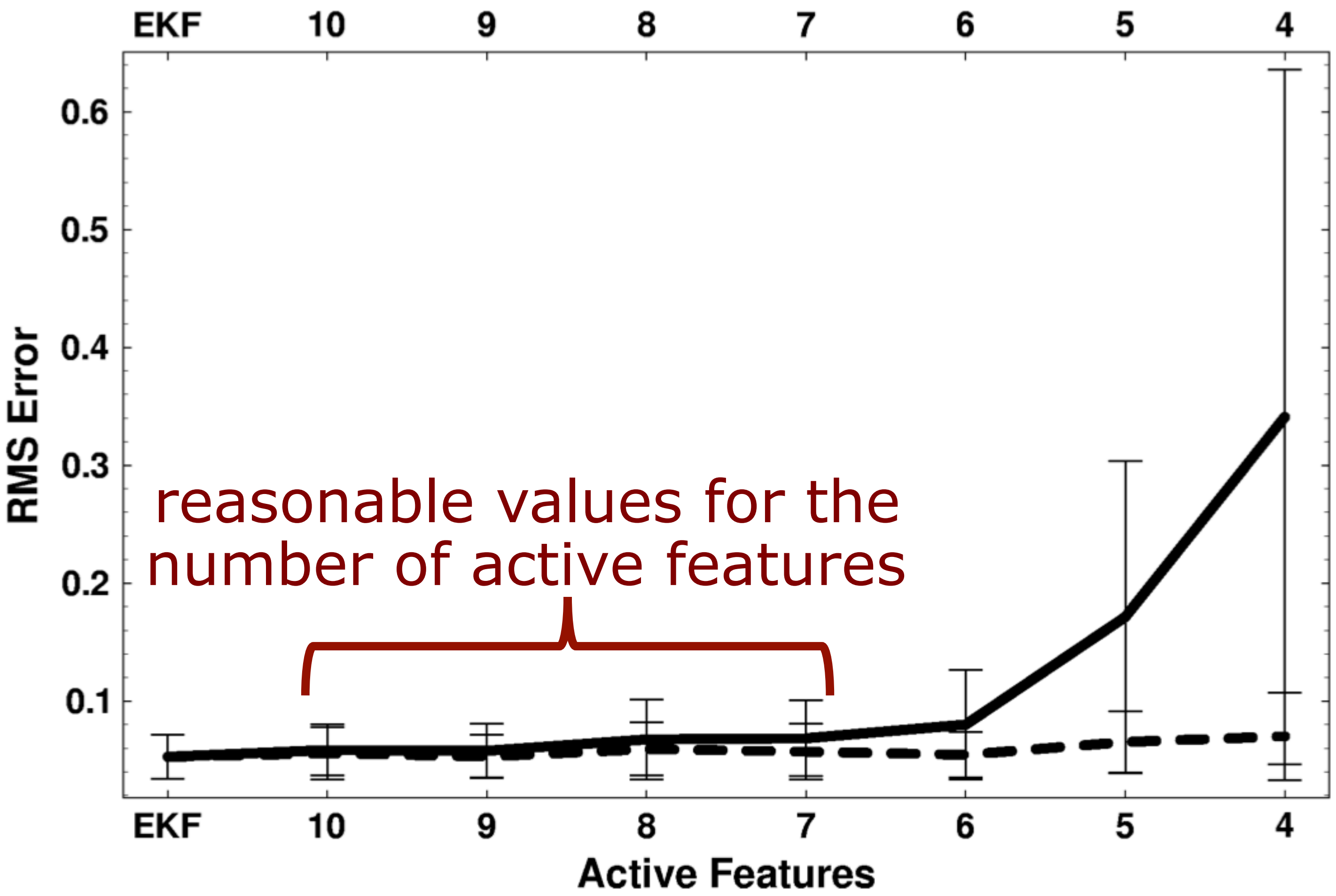


Image Credit: S Thrun