

ICRA 2016 Tutorial on SLAM

Graph-Based SLAM and Sparsity

Cyrill Stachniss



Graph-Based SLAM ??

Graph-Based SLAM ??

SLAM = simultaneous localization and mapping

Graph-Based SLAM ??

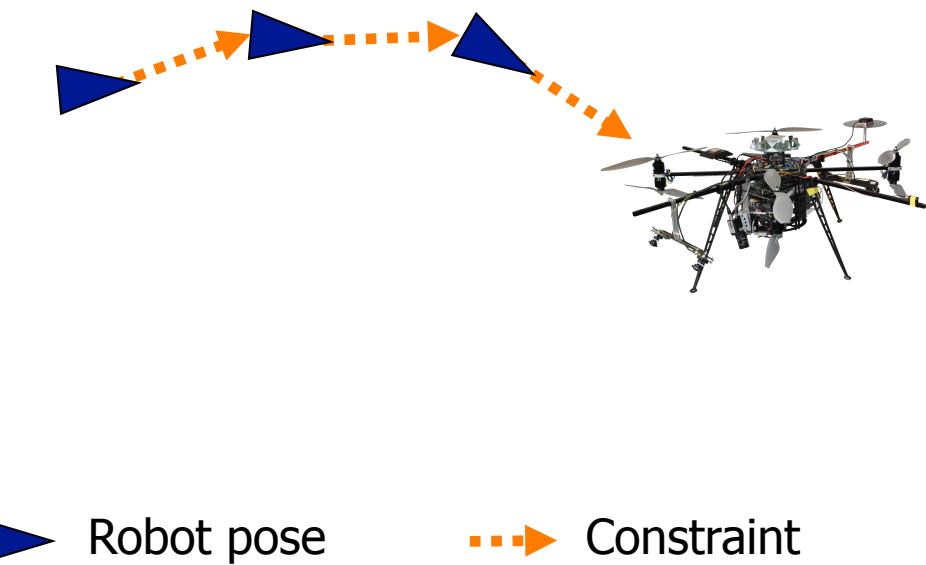
SLAM = simultaneous localization and mapping

graph = representation of a set of objects where pairs of objects are connected by links encoding relations between the objects

**What is my goal
for today?**

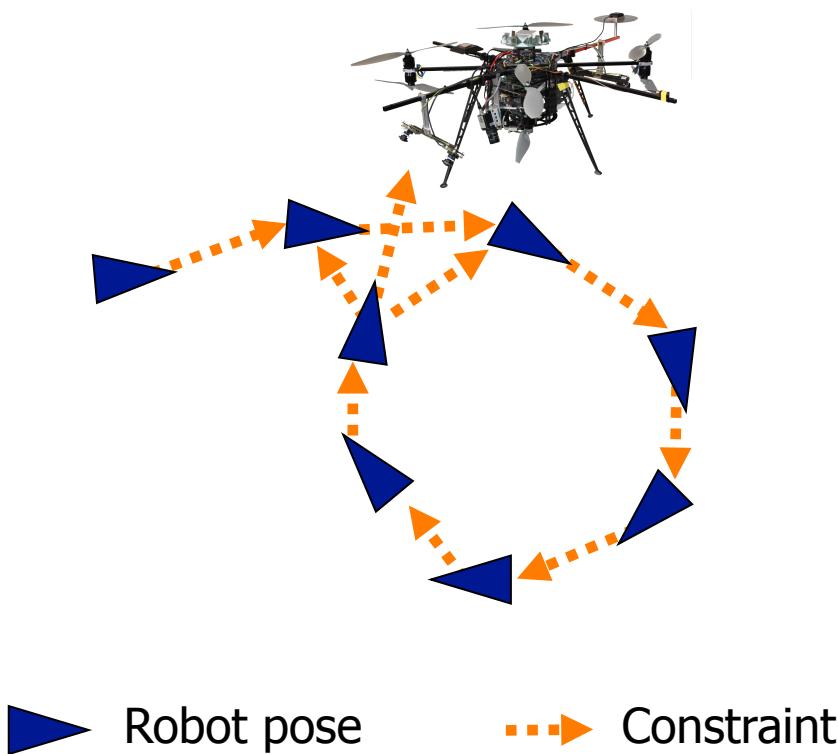
Graph-Based SLAM

- Nodes represent poses or locations
- Constraints connect the poses of the robot while it is moving
- Constraints are inherently uncertain



Graph-Based SLAM

- Observing previously seen areas generates constraints between non-successive poses



Idea of Graph-Based SLAM

- Use a **graph** to represent the problem
- Every **node** in the graph corresponds to a pose of the robot during mapping
- Every **edge** between two nodes corresponds to a spatial constraint between them
- **Graph-Based SLAM:** Build the graph and find a node configuration that minimize the error introduced by the constraints

Graph-SLAM and Least Squares

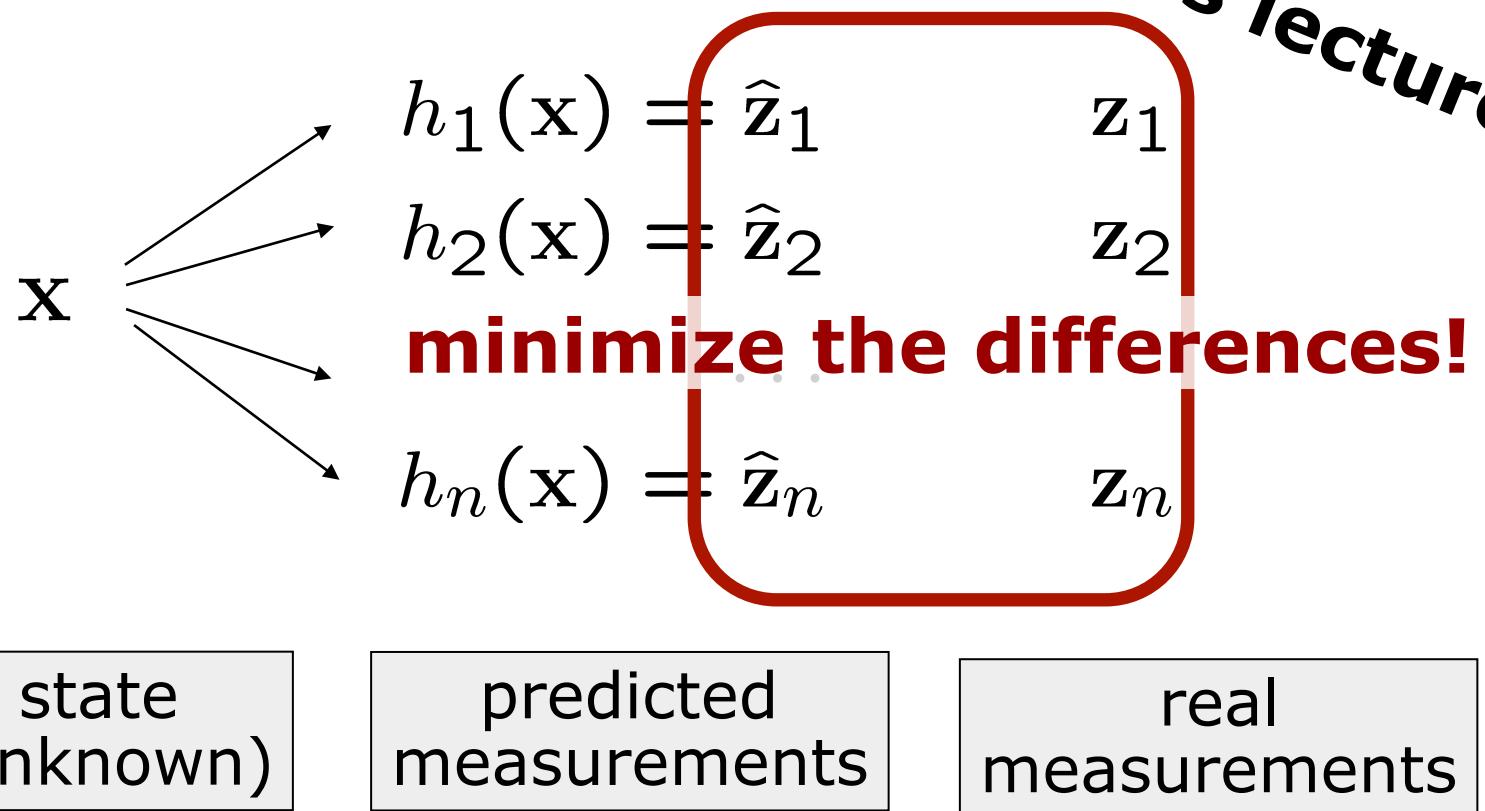
- The nodes represent the **state**
- Given a state, we can compute what we **expect** to perceive
- We have **real observations** relating the nodes with each other

Graph-SLAM and Least Squares

- The nodes represent the **state**
- Given a state, we can compute what we **expect** to perceive
- We have **real observations** relating the nodes with each other
Giorgio's lecture

Find a configuration of the nodes so that the real and predicted observations are as similar as possible

Error Function



$$e_i(\mathbf{x}) = \mathbf{e}_i(\mathbf{x})^T \boldsymbol{\Omega}_i \mathbf{e}_i(\mathbf{x})$$

$$\mathbf{e}_i(\mathbf{x}) = \mathbf{z}_i - h_i(\mathbf{x})$$

Procedure in Brief

Iterate the following steps:

- Linearize around \mathbf{x} and compute for each measurement

$$\mathbf{e}_i(\mathbf{x} + \Delta\mathbf{x}) \simeq \mathbf{e}_i(\mathbf{x}) + \mathbf{J}_i \Delta\mathbf{x}$$

- Compute the terms for the linear system $\mathbf{b} = \sum_i \mathbf{J}_i^T \boldsymbol{\Omega}_i \mathbf{e}_i$ $\mathbf{H} = \sum_i \mathbf{J}_i^T \boldsymbol{\Omega}_i \mathbf{J}_i$
- Solve the linear system

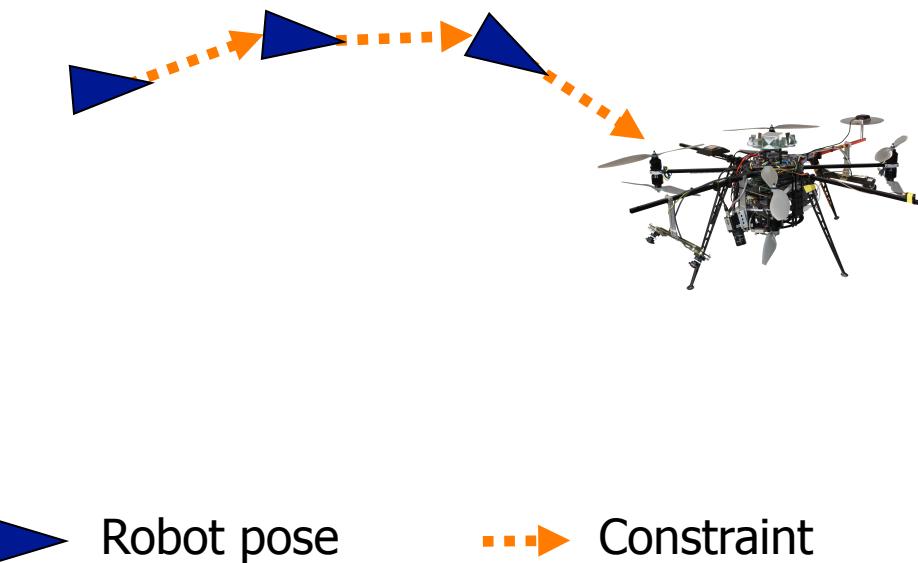
$$\Delta\mathbf{x}^* = -\mathbf{H}^{-1}\mathbf{b}$$

- Updating state $\mathbf{x} \leftarrow \mathbf{x} + \Delta\mathbf{x}^*$

Let's use that for SLAM

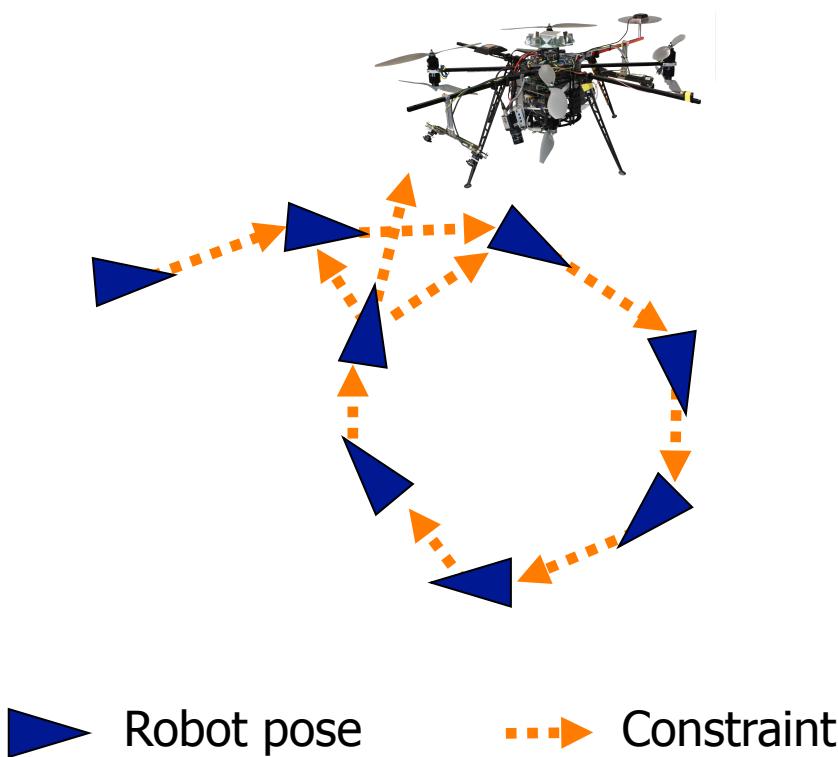
Pose-Graph-Based SLAM

- Nodes represent poses or locations
- Constraints connect the poses of the robot while it is moving
- Constraints are inherently uncertain



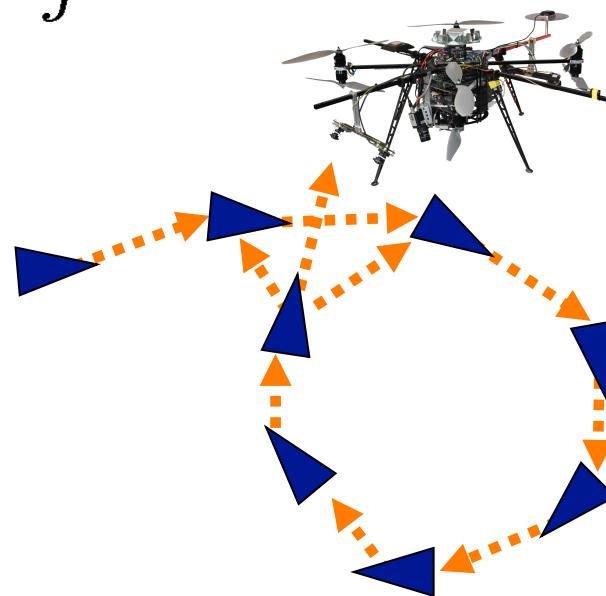
Pose-Graph-Based SLAM

- Observing previously seen areas generates constraints between non-successive poses



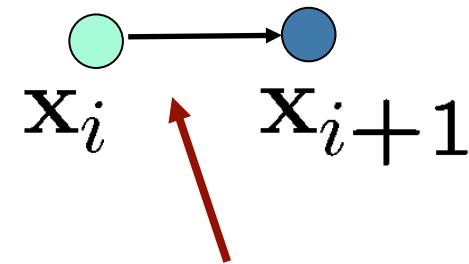
The Pose-Graph

- It consists of n nodes $\mathbf{x} = \mathbf{x}_{1:n}$
- Each \mathbf{x}_i is a 2D or 3D **pose** (position and orientation of the robot at time t_i)
- A constraint/edge exists between the nodes \mathbf{x}_i and \mathbf{x}_j if...



Create an Edge If... (1)

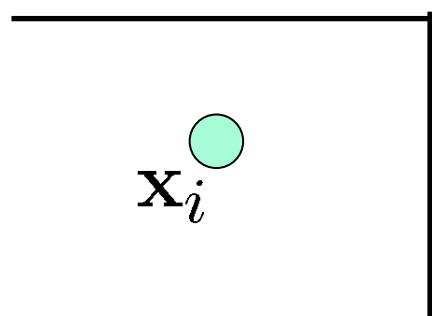
- ...the robot moves from \mathbf{x}_i to \mathbf{x}_{i+1}
- Edge corresponds to odometry



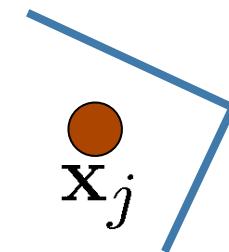
The edge represents the
odometry measurement

Create an Edge If... (2)

- ...the robot observes the same part of the environment from \mathbf{x}_i and from \mathbf{x}_j



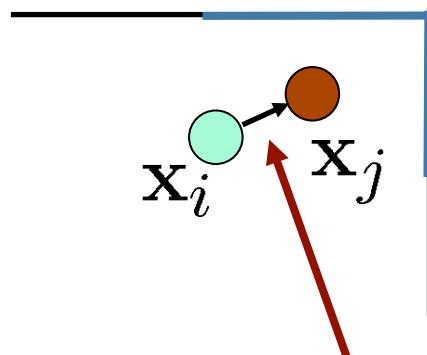
Measurement from \mathbf{x}_i



Measurement from \mathbf{x}_j

Create an Edge If... (2)

- ...the robot observes the same part of the environment from x_i and from x_j
- Construct a **virtual measurement** about the position of x_j seen from x_i



Edge represents the position of x_j seen from x_i based on the **observation**

Transformations

- **How to express x_j relative to x_i ?**
- Express this through transformations
- Let X_i be transformation of the origin into x_i
- Let X_i^{-1} be the inverse transformation
- We can express relative transformation $X_i^{-1}X_j$

Transformations

- **How to express x_j relative to x_i ?**
- Express this through transformations
- Let X_i be transformation of the origin into x_i
- Let X_i^{-1} be the inverse transformation
- We can express relative transformation $X_i^{-1}X_j$
- Transformations can be expressed using **homogenous coordinates**

Transformations

- Transformations can be expressed using **homogenous coordinates**
- Odometry-Based edge

$$(\mathbf{X}_i^{-1} \mathbf{X}_{i+1})$$

- Observation-Based edge

$$(\mathbf{X}_i^{-1} \mathbf{X}_j)$$

describes “how node i sees node j”

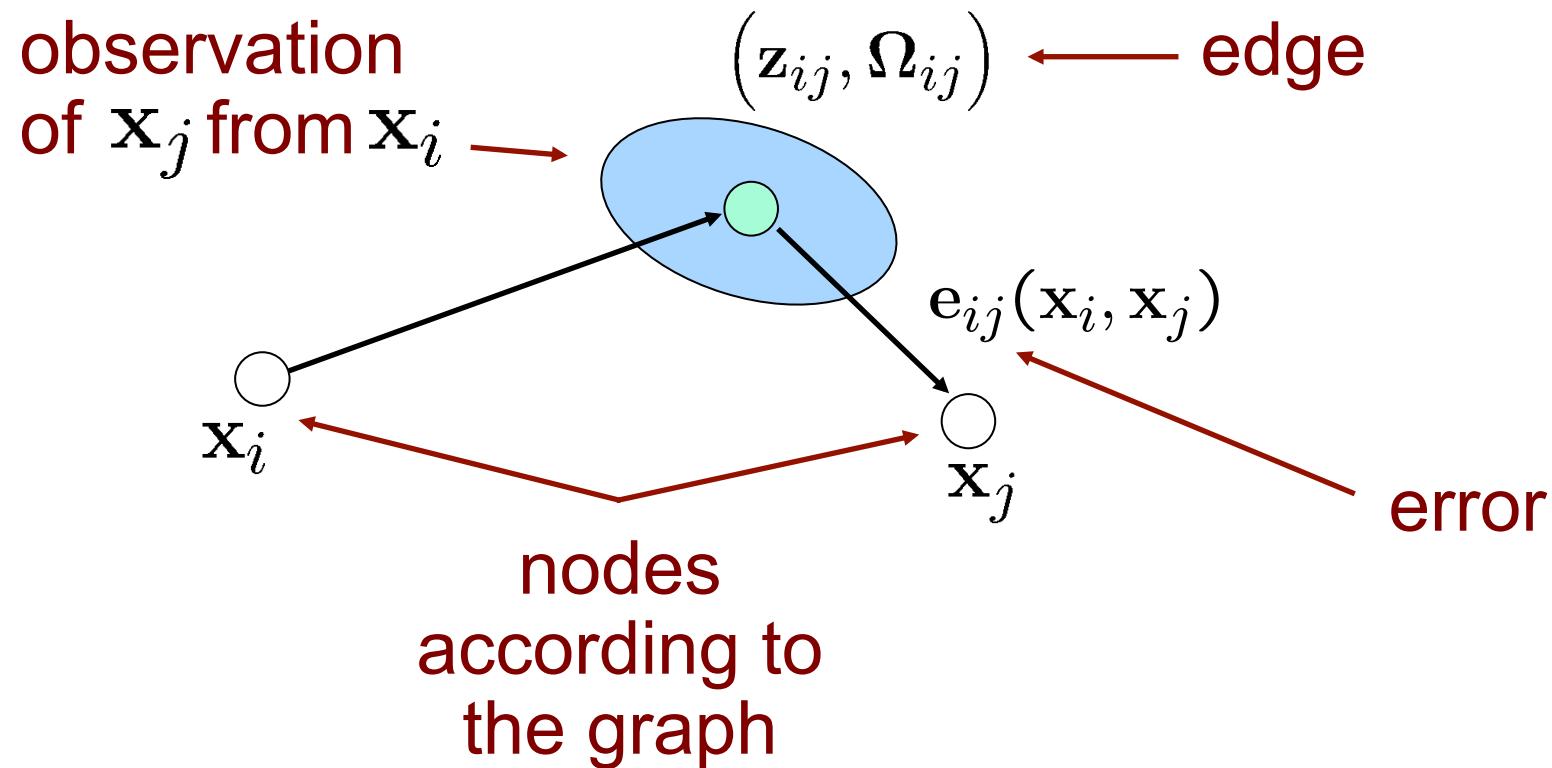
The Edge Information Matrices

- Observations are affected by noise
- Information matrix Ω_{ij} for each edge to encode its uncertainty
- The “bigger” Ω_{ij} , the more the edge “matters” in the optimization

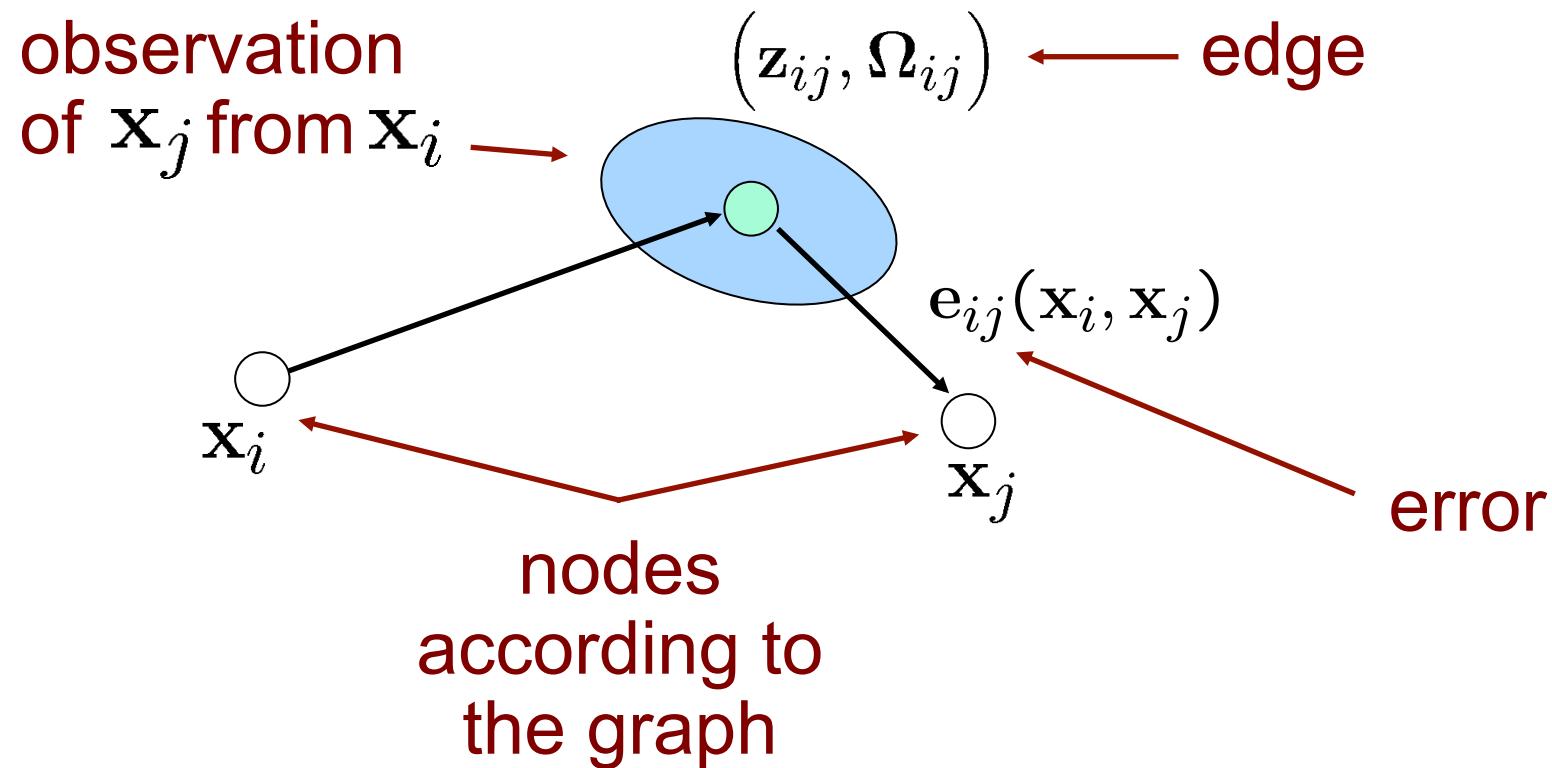
Question

- What should these matrices look like when moving in a long, featureless corridor?

Pose-Graph



Pose-Graph



Goal: $\mathbf{x}^* = \operatorname{argmin}_{\mathbf{x}} \sum_{ij} \mathbf{e}_{ij}^T \Omega_{ij} \mathbf{e}_{ij}$

The Error Function

- Error function for a single constraint

$$e_{ij}(\mathbf{x}_i, \mathbf{x}_j) = t2v(\underline{\mathbf{Z}_{ij}^{-1}} \underline{(\mathbf{X}_i^{-1} \mathbf{X}_j)})$$

↑
measurement

↑
 \mathbf{x}_j seen from \mathbf{x}_i

- Error as a function of the whole state vector

$$e_{ij}(\mathbf{x}) = t2v(\mathbf{Z}_{ij}^{-1}(\mathbf{X}_i^{-1} \mathbf{X}_j))$$

- Error takes a value of zero if

$$\mathbf{Z}_{ij} = (\mathbf{X}_i^{-1} \mathbf{X}_j)$$

Error Minimization Procedure

- Define the error function
- Linearize the error function
- Compute its derivative
- Set the derivative to zero
- Solve the linear system
- Iterate this procedure until convergence

Linearizing the Error Function

- We can approximate the error functions around an initial guess \mathbf{x} via Taylor expansion

$$e_{ij}(\mathbf{x} + \Delta\mathbf{x}) \simeq e_{ij}(\mathbf{x}) + J_{ij}\Delta\mathbf{x}$$

with $J_{ij} = \frac{\partial e_{ij}(\mathbf{x})}{\partial \mathbf{x}}$

Derivative of the Error Function

- Does one error term $e_{ij}(x)$ depend on all state variables?

Derivative of the Error Function

- Does one error term $e_{ij}(x)$ depend on all state variables?
 - ➡ No, only on x_i and x_j

Derivative of the Error Function

- Does one error term $e_{ij}(x)$ depend on all state variables?
 - No, only on x_i and x_j
- Is there any consequence on the **structure** of the Jacobian?

Derivative of the Error Function

- Does one error term $e_{ij}(\mathbf{x})$ depend on all state variables?
 - ➡ No, only on \mathbf{x}_i and \mathbf{x}_j
- Is there any consequence on the **structure** of the Jacobian?
 - ➡ Yes, it will be non-zero only in the rows corresponding to \mathbf{x}_i and \mathbf{x}_j

$$\begin{aligned}\frac{\partial e_{ij}(\mathbf{x})}{\partial \mathbf{x}} &= \left(0 \dots \frac{\partial e_{ij}(\mathbf{x}_i)}{\partial \mathbf{x}_i} \dots \frac{\partial e_{ij}(\mathbf{x}_j)}{\partial \mathbf{x}_j} \dots 0 \right) \\ J_{ij} &= \left(0 \dots A_{ij} \dots B_{ij} \dots 0 \right)\end{aligned}$$

Jacobians and Sparsity

- Error $e_{ij}(x)$ depends only on the two parameter blocks x_i and x_j

$$e_{ij}(x) = e_{ij}(x_i, x_j)$$

- The Jacobian will be zero everywhere except in the columns of x_i and x_j

$$\mathbf{J}_{ij} = \begin{pmatrix} 0 \cdots 0 & \underbrace{\frac{\partial e(x_i)}{\partial x_i}}_{A_{ij}} & 0 \cdots 0 & \underbrace{\frac{\partial e(x_j)}{\partial x_j}}_{B_{ij}} & 0 \cdots 0 \end{pmatrix}$$

Consequences of the Sparsity

- We need to compute the coefficient vector b and matrix H :

$$b = \sum_{ij} b_{ij} = \sum_{ij} J_{ij}^T \Omega_{ij} e_{ij}$$

$$H = \sum_{ij} H_{ij} = \sum_{ij} J_{ij}^T \Omega_{ij} J_{ij}$$

- The sparse structure of J_{ij} will result in a sparse structure of H
- This structure reflects the adjacency matrix of the graph

Illustration of the Structure

$$\mathbf{b}_{ij} = \mathbf{J}_{ij}^T \boldsymbol{\Omega}_{ij} \mathbf{e}_{ij}$$

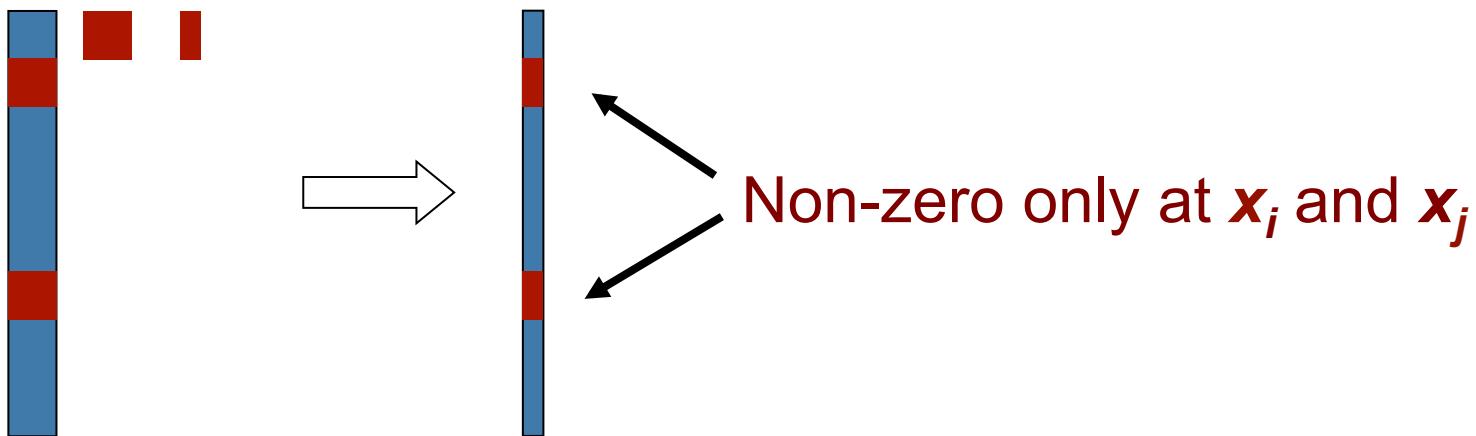
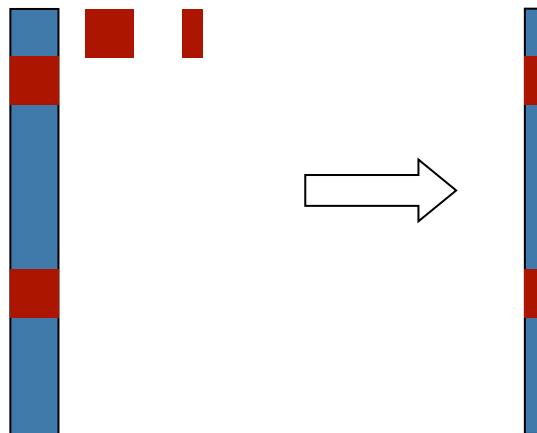


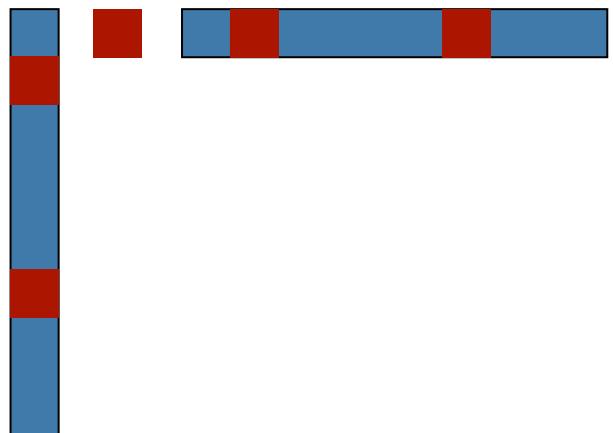
Illustration of the Structure

$$\mathbf{b}_{ij} = \mathbf{J}_{ij}^T \boldsymbol{\Omega}_{ij} \mathbf{e}_{ij}$$



Non-zero only at \mathbf{x}_i and \mathbf{x}_j

$$\mathbf{H}_{ij} = \mathbf{J}_{ij}^T \boldsymbol{\Omega}_{ij} \mathbf{J}_{ij}$$



Non-zero on the main
diagonal at \mathbf{x}_i and \mathbf{x}_j

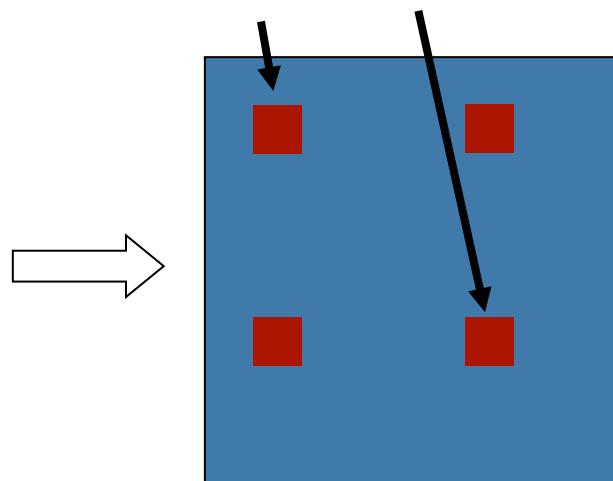
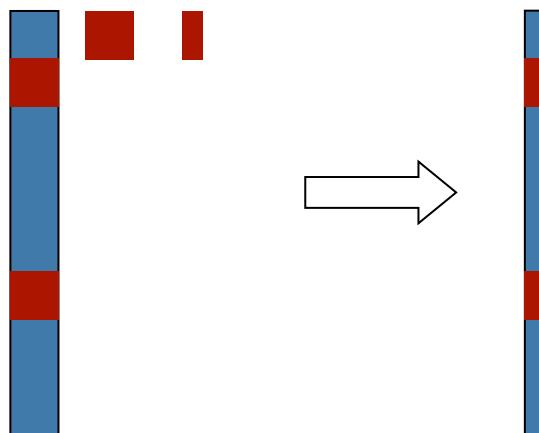


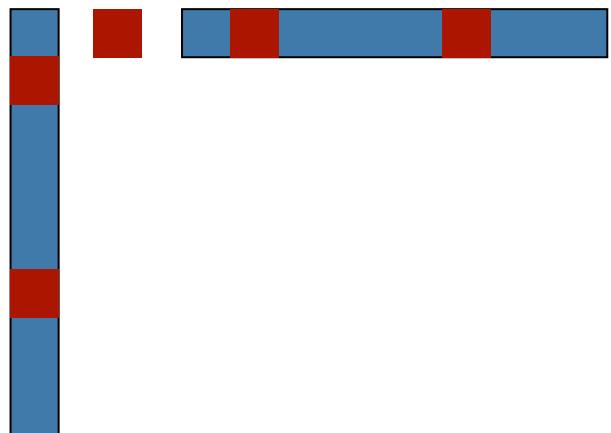
Illustration of the Structure

$$\mathbf{b}_{ij} = \mathbf{J}_{ij}^T \boldsymbol{\Omega}_{ij} \mathbf{e}_{ij}$$

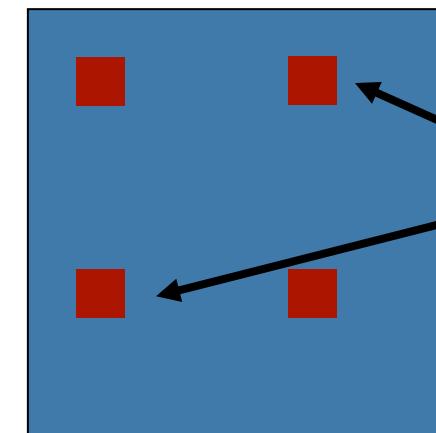


Non-zero only at \mathbf{x}_i and \mathbf{x}_j

$$\mathbf{H}_{ij} = \mathbf{J}_{ij}^T \boldsymbol{\Omega}_{ij} \mathbf{J}_{ij}$$



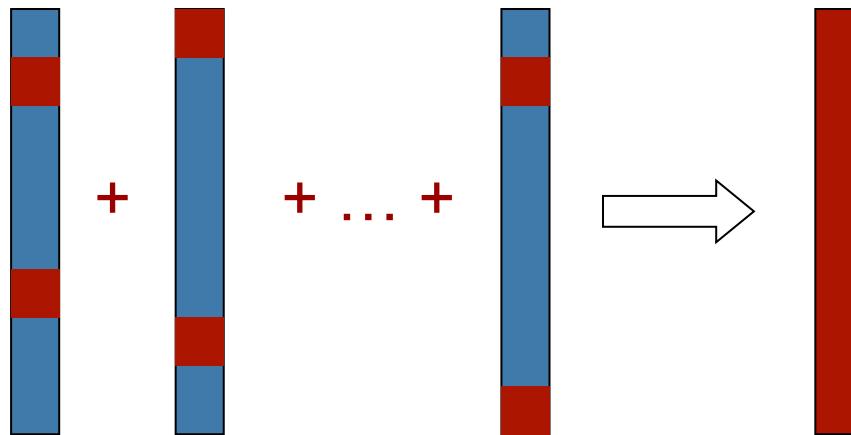
Non-zero on the main
diagonal at \mathbf{x}_i and \mathbf{x}_j



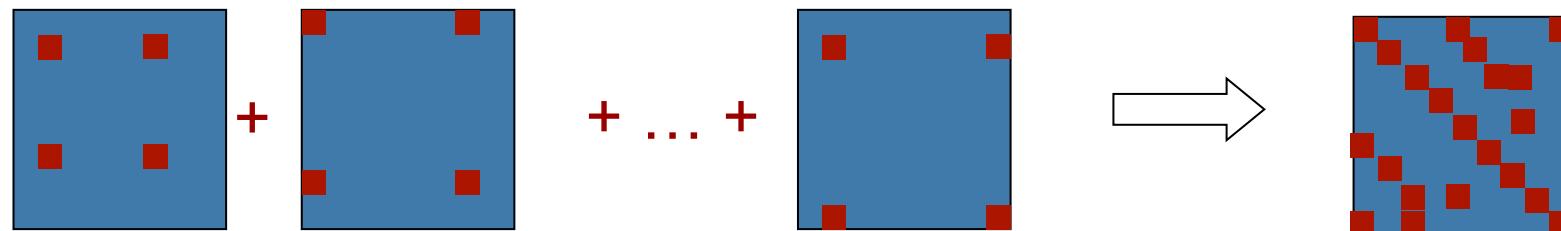
... and at
the blocks
 ij, ji

Illustration of the Structure

$$\mathbf{b} = \sum_{ij} \mathbf{b}_{ij}$$



$$\mathbf{H} = \sum_{ij} \mathbf{H}_{ij}$$



Sparsity Effect on \mathbf{b}

- An edge contributes to the linear system via b_{ij} and H_{ij}
- The coefficient vector is:

$$\begin{aligned} \mathbf{b}_{ij}^T &= \mathbf{e}_{ij}^T \Omega_{ij} \mathbf{J}_{ij} \\ &= \mathbf{e}_{ij}^T \Omega_{ij} \left(\begin{array}{cccc} 0 & \cdots & \mathbf{A}_{ij} & \cdots & \mathbf{B}_{ij} & \cdots & 0 \end{array} \right) \\ &= \left(\begin{array}{cccc} 0 & \cdots & \mathbf{e}_{ij}^T \Omega_{ij} \mathbf{A}_{ij} & \cdots & \mathbf{e}_{ij}^T \Omega_{ij} \mathbf{B}_{ij} & \cdots & 0 \end{array} \right) \end{aligned}$$

- It is non-zero only at the indices corresponding to x_i and x_j

Sparsity Effect on H

- The coefficient matrix of an edge is:

$$\begin{aligned} H_{ij} &= \mathbf{J}_{ij}^T \boldsymbol{\Omega}_{ij} \mathbf{J}_{ij} \\ &= \begin{pmatrix} \vdots \\ \mathbf{A}_{ij}^T \\ \vdots \\ \mathbf{B}_{ij}^T \\ \vdots \end{pmatrix} \boldsymbol{\Omega}_{ij} \left(\cdots \mathbf{A}_{ij} \cdots \mathbf{B}_{ij} \cdots \right) \\ &= \begin{pmatrix} \mathbf{A}_{ij}^T \boldsymbol{\Omega}_{ij} \mathbf{A}_{ij} & \mathbf{A}_{ij}^T \boldsymbol{\Omega}_{ij} \mathbf{B}_{ij} \\ \mathbf{B}_{ij}^T \boldsymbol{\Omega}_{ij} \mathbf{A}_{ij} & \mathbf{B}_{ij}^T \boldsymbol{\Omega}_{ij} \mathbf{B}_{ij} \end{pmatrix} \end{aligned}$$

- Non-zero only in the blocks relating i,j

Sparsity Summary

- An edge ij contributes only to the
 - i^{th} and the j^{th} block of b_{ij}
 - to the blocks ii , jj , ij and ji of H_{ij}
- Resulting system is sparse
- System can be computed by summing up the contribution of each edge
- Efficient solvers can be used
 - Sparse Cholesky decomposition
 - Conjugate gradients
 - ... many others

All We Need...

- Vector of the states increments:

$$\Delta \mathbf{x}^T = (\Delta \mathbf{x}_1^T \ \Delta \mathbf{x}_2^T \ \dots \ \Delta \mathbf{x}_n^T)$$

- Coefficient vector:

$$\mathbf{b}^T = (\bar{\mathbf{b}}_1^T \ \bar{\mathbf{b}}_2^T \ \dots \ \bar{\mathbf{b}}_n^T)$$

- System matrix:

$$\mathbf{H} = \begin{pmatrix} \bar{\mathbf{H}}^{11} & \bar{\mathbf{H}}^{12} & \dots & \bar{\mathbf{H}}^{1n} \\ \bar{\mathbf{H}}^{21} & \bar{\mathbf{H}}^{22} & \dots & \bar{\mathbf{H}}^{2n} \\ \vdots & \ddots & & \vdots \\ \bar{\mathbf{H}}^{n1} & \bar{\mathbf{H}}^{n2} & \dots & \bar{\mathbf{H}}^{nn} \end{pmatrix}$$

small blocks
(or vectors)
corresponding
to the individual
constraints

... for the Linear System

For each constraint:

- Compute error $e_{ij} = t2v(Z_{ij}^{-1}(X_i^{-1}X_j))$
- Compute the blocks of the Jacobian:

$$A_{ij} = \frac{\partial e(x_i, x_j)}{\partial x_i} \quad B_{ij} = \frac{\partial e(x_i, x_j)}{\partial x_j}$$

- Update the coefficient vector:

$$\bar{b}_i^T + = e_{ij}^T \Omega_{ij} A_{ij} \quad \bar{b}_j^T + = e_{ij}^T \Omega_{ij} B_{ij}$$

- Update the system matrix:

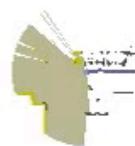
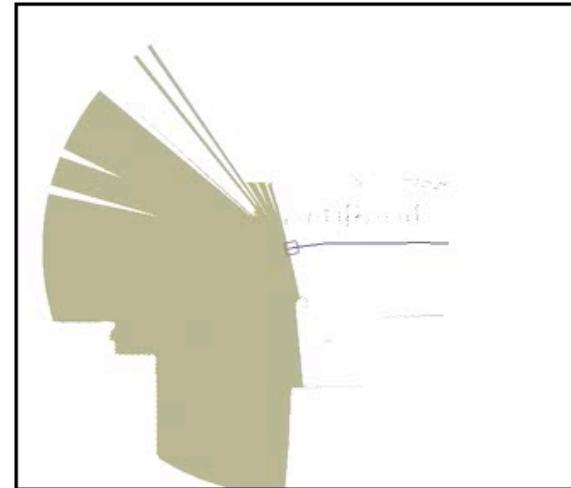
$$\bar{H}^{ii} + = A_{ij}^T \Omega_{ij} A_{ij} \quad \bar{H}^{ij} + = A_{ij}^T \Omega_{ij} B_{ij}$$

$$\bar{H}^{ji} + = B_{ij}^T \Omega_{ij} A_{ij} \quad \bar{H}^{jj} + = B_{ij}^T \Omega_{ij} B_{ij}$$

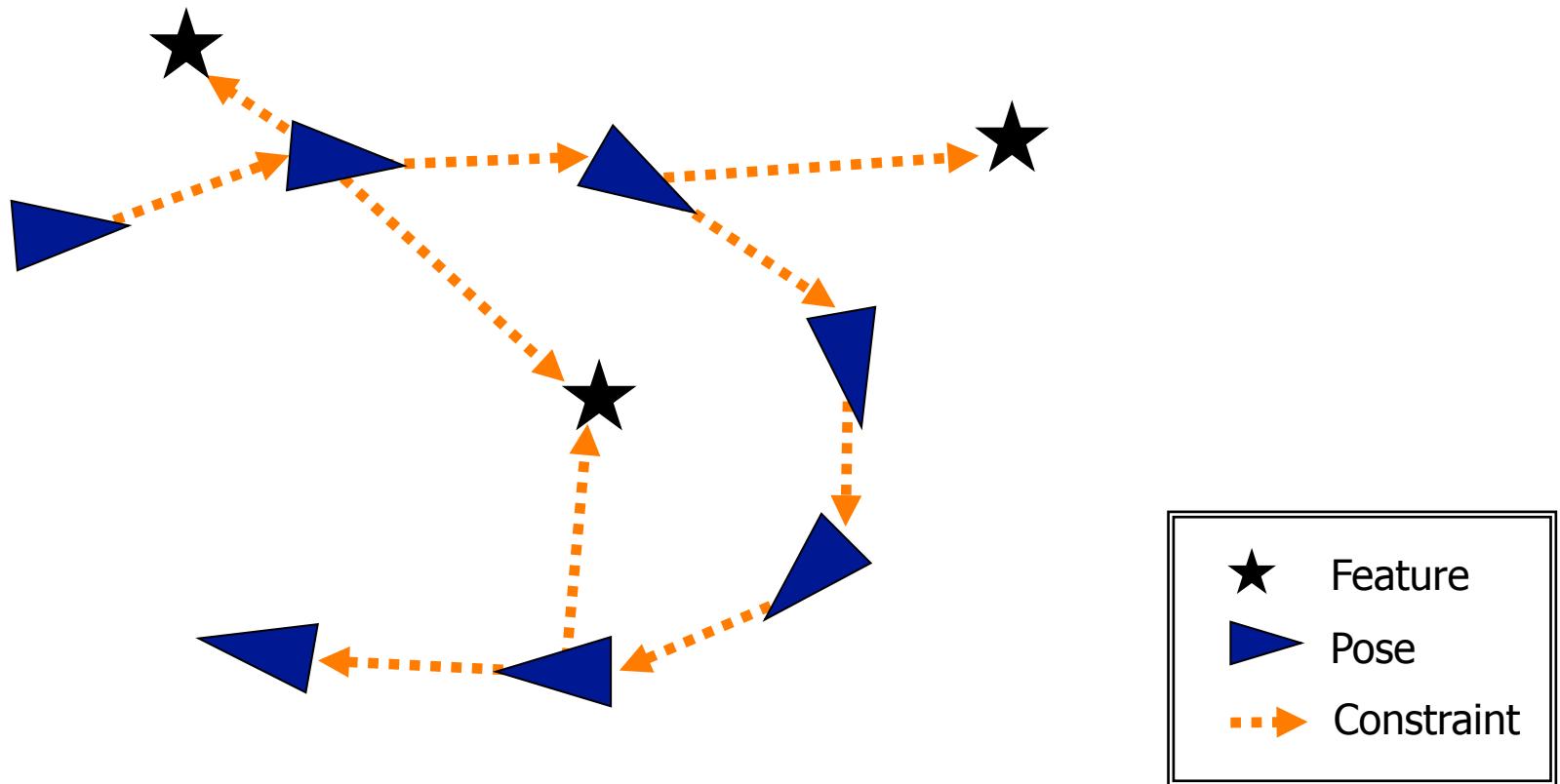
Algorithm

```
1: optimize(x):  
2:     while (!converged)  
3:         ( $\mathbf{H}$ ,  $\mathbf{b}$ ) = buildLinearSystem( $\mathbf{x}$ )  
4:          $\Delta\mathbf{x}$  = solveSparse( $\mathbf{H}\Delta\mathbf{x} = -\mathbf{b}$ )  
5:          $\mathbf{x} = \mathbf{x} + \Delta\mathbf{x}$   
6:     end  
7:     return  $\mathbf{x}$ 
```

Real World Examples

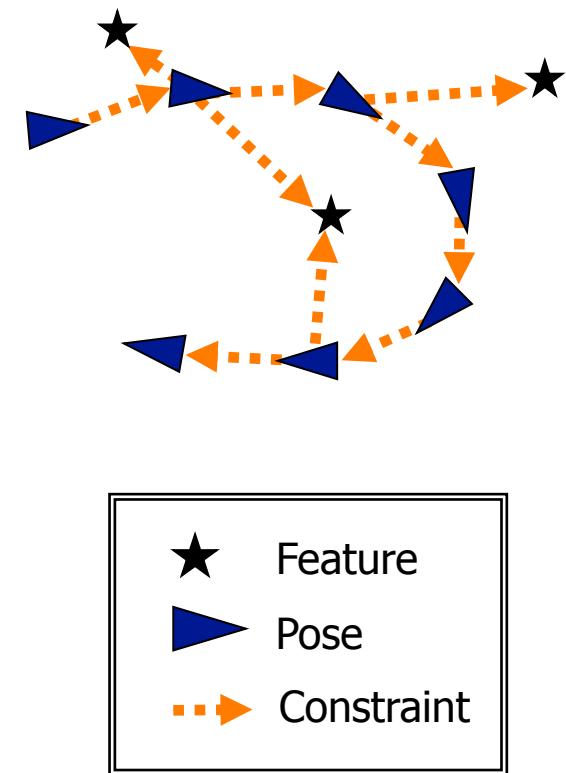


The Graph with Landmarks



The Graph with Landmarks

- **Nodes** can represent:
 - Robot poses
 - Landmark locations
- **Edges** can represent:
 - Landmark observations
 - Odometry measurements
- The minimization optimizes the **landmark locations and robot poses**



Landmarks Observation

- Expected observation (x-y sensor)

$$\hat{z}_{il}(x_i, x_l) = X_i^{-1} \begin{pmatrix} x_l \\ 1 \end{pmatrix}$$


robot landmark

Landmarks Observation

- Expected observation (x-y sensor)

$$\hat{z}_{il}(x_i, x_l) = X_i^{-1} \begin{pmatrix} x_l \\ 1 \end{pmatrix}$$

↑ ↑
robot landmark

- Error function (in Euclidian space)

$$e_{il}(x_i, x_l) = \hat{z}_{il} - z_{il}$$

Bearing Only Observations

- A landmark is still a 2D point
- The robot observe only the bearing towards the landmark
- 1D Observation function

$$\hat{z}_{il}(\mathbf{x}_i, \mathbf{x}_l) = \text{atan} \frac{(\mathbf{x}_l - \mathbf{t}_i).y}{(\mathbf{x}_l - \mathbf{t}_i).x} - \theta_i$$

$\uparrow \quad \uparrow \quad \uparrow \quad \uparrow$
robot landmark robot-landmark robot
 angle orientation

Bearing Only Observations

- ## ■ Observation function

$$\hat{z}_{il}(\mathbf{x}_i, \mathbf{x}_l) = \text{atan} \frac{(\mathbf{x}_l - \mathbf{t}_i) \cdot \mathbf{y}}{(\mathbf{x}_l - \mathbf{t}_i) \cdot \mathbf{x}} - \theta_i$$

↑ ↑ ↑
 robot landmark robot-landmark
 angle robot
 orientation

- ## ■ Error function

$$e_{il}(x_i, x_l) = \text{atan} \frac{(x_l - t_i) \cdot y}{(x_l - t_i) \cdot x} - \theta_i - z_{il}$$

The Rank of the Matrix \mathbf{H}

- What is the rank of \mathbf{H}_{ij} for a 2D landmark-pose constraint?

The Rank of the Matrix \mathbf{H}

- What is the rank of \mathbf{H}_{ij} for a 2D landmark-pose constraint?
 - The blocks of \mathbf{J}_{ij} are a 2×3 matrices
 - \mathbf{H}_{ij} cannot have more than rank 2
 $\text{rank}(\mathbf{A}^T \mathbf{A}) = \text{rank}(\mathbf{A}^T) = \text{rank}(\mathbf{A})$

The Rank of the Matrix \mathbf{H}

- What is the rank of \mathbf{H}_{ij} for a 2D landmark-pose constraint?
 - The blocks of \mathbf{J}_{ij} are a 2×3 matrices
 - \mathbf{H}_{ij} cannot have more than rank 2
 $\text{rank}(\mathbf{A}^T \mathbf{A}) = \text{rank}(\mathbf{A}^T) = \text{rank}(\mathbf{A})$
- What is the rank of \mathbf{H}_{ij} for a bearing-only constraint?

The Rank of the Matrix \mathbf{H}

- What is the rank of \mathbf{H}_{ij} for a 2D landmark-pose constraint?
 - The blocks of \mathbf{J}_{ij} are a 2×3 matrices
 - \mathbf{H}_{ij} cannot have more than rank 2
 $\text{rank}(\mathbf{A}^T \mathbf{A}) = \text{rank}(\mathbf{A}^T) = \text{rank}(\mathbf{A})$
- What is the rank of \mathbf{H}_{ij} for a bearing-only constraint?
 - The blocks of \mathbf{J}_{ij} are a 1×3 matrices
 - \mathbf{H}_{ij} has rank 1

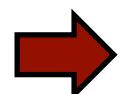
Rank

- In landmark-based SLAM, the system can be under-determined
- The rank of H is **less or equal** to the sum of the ranks of the constraints
- To determine a **unique solution**, the system must have **full rank**

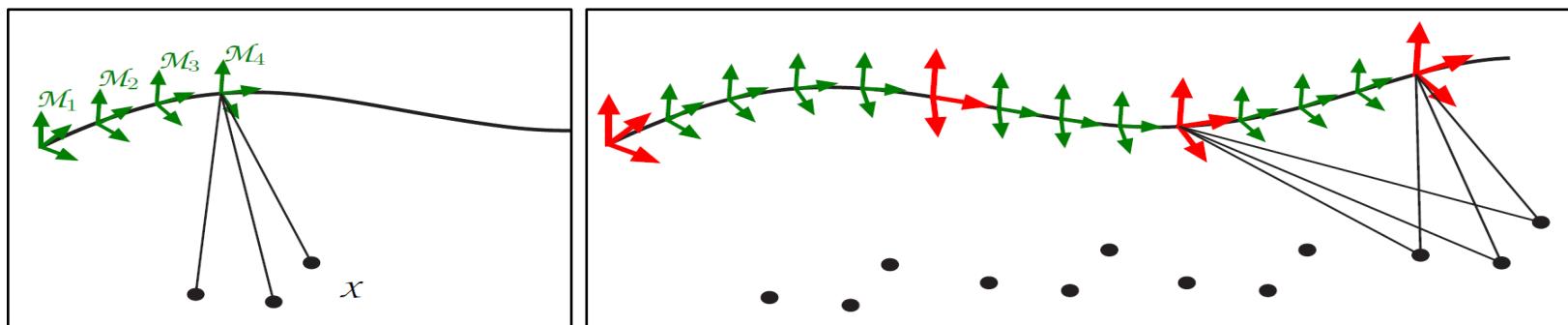
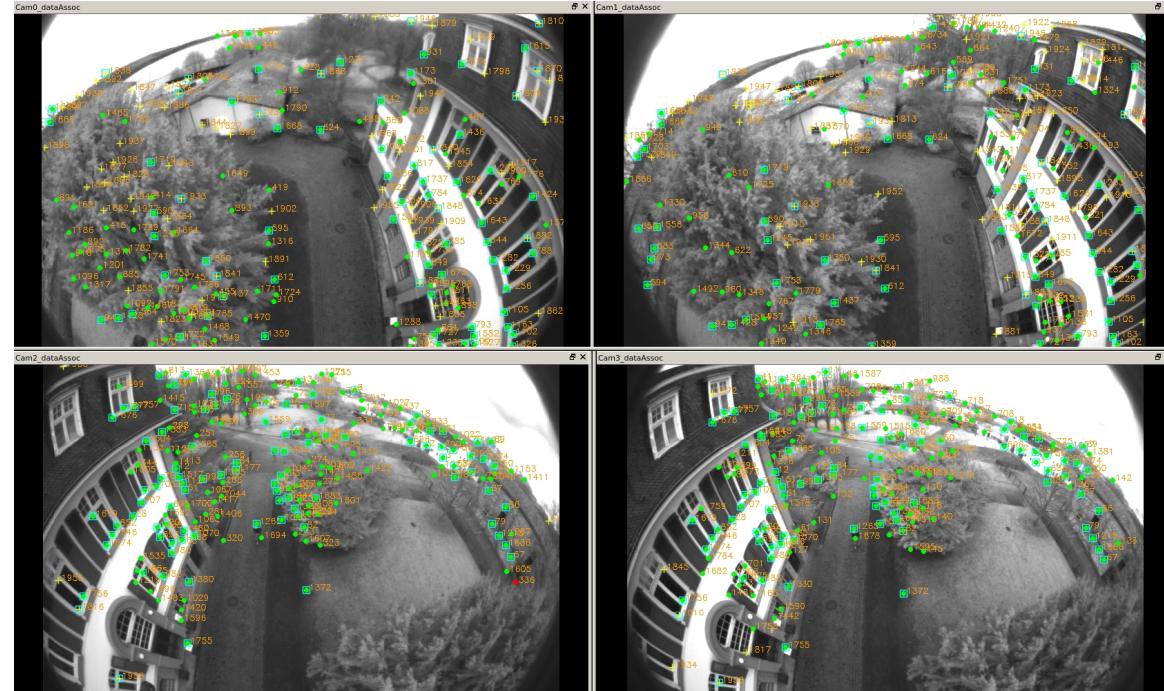
Under-Determined Systems

- No guarantee for a full rank system
 - Landmarks may be observed only once
 - Robot might have no odometry
- We can still deal with these situations by adding a “damping” factor to H
- Instead of solving $H\Delta x = -b$, we solve

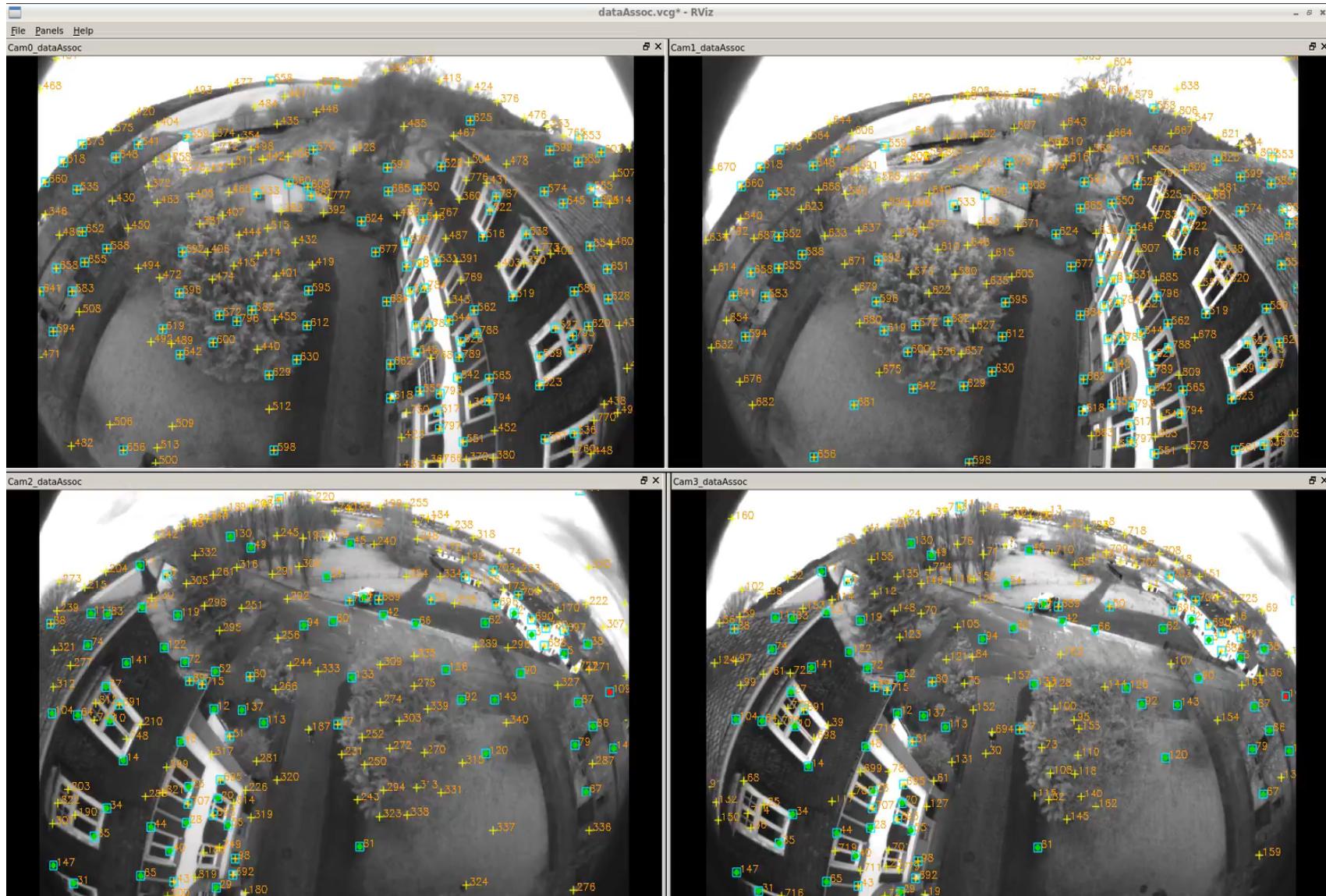
$$(H + \lambda I)\Delta x = -b$$

 **Levenberg Marquardt**

UAV Example



UAV Example



Summary

- The back-end part of the SLAM problem can be solved with GN or LM
- The H matrix is typically sparse
- This sparsity allows for efficiently solving the linear system
- There are several extensions (online, robust methods wrt outliers or initialization, hierarchical approaches, exploiting sparsity, multiple sensors)

YouTube Lectures

Cyrill Stachniss Videos Playlists Channels Discussion About

 SLAM Course - WS13/14

• Cyrill Stachniss • 22 videos • 52,746 views • Last updated on Jul 1, 2014

Lecture Recordings from my winter 2013/14 course on SLAM taught in Freiburg.

Lecture material can be found here:
<http://ais.informatik.uni-freiburg.de/teaching/ws13/mapping/>

▶ Play all  Share  Playlist settings Add videos

1		SLAM-Course - 00 - Course Introduction (2013/14; Cyrill Stachniss)	18:06
2		SLAM-Course - 01 - Introduction to Robot Mapping (2013/14; Cyrill Stachniss)	1:16:35
3		SLAM-Course - 02 - Homogeneous Coordinates (2013/14; Cyrill Stachniss)	28:35
4		SLAM-Course - 03 - Bayes Filter (2013/14; Cyrill Stachniss)	53:17

https://www.youtube.com/playlist?list=PLgnQpQtFTOGQrZ4O5QzbIHgI3b1JHimN_

Thank you for your attention!

Slide Information

- These slides have been created by Cyrill Stachniss, Giorgio Grisetti and Wolfram Burgard evolving from different courses and tutorials that we taught over the years between 2010 and 2016.
- I tried to acknowledge all people that contributed image or video material. In case I missed something, please let me know. If you adapt this course material, please make sure you keep the acknowledgements.
- Feel free to use and change the slides. If you use them, I would appreciate an acknowledgement as well. To satisfy my own curiosity, I appreciate a short email notice in case you use the material in your course.
- My video recordings of my lectures on robot mapping are available through YouTube:
http://www.youtube.com/playlist?list=PLgnQpQtFTOGQrZ4O5QzbIHgI3b1JHimN_&feature=g-list

Cyrill Stachniss, cyrill.stachniss@igg.uni-bonn.de