16.485: VNAV - Visual Navigation for Autonomous Vehicles

Luca Carlone

Lecture 23: SLAM I - Formulations and Sparsity

based on slides by Kasra Khosoussi

Today

Simultaneous Localization and Mapping

- "Holy grail of mobile robotics"
- Over 30 years of *robotic* research

Simultaneous Localisation and Mapping (SLAM): Part I The Essential Algorithms

Hugh Durrant-Whyte, *Fellow, IEEE*, and Tim Bailey

Abstract—**This tutorial provides an introduction to Simultaneous Localisation and Mapping (SLAM) and the extensive research on SLAM that has been undertaken over the past decade. SLAM is the process by which a mobile robot can build a map of an environment and at the same time use this map to compute it's own location. The past decade has seen rapid and exciting progress in solving the SLAM problem together with many compelling implementations of SLAM methods. Part I of this tutorial (this paper), describes the probabilistic form of the SLAM problem, essential solution methods and significant implementations. Part II of this tutorial will be concerned with recent advances in computational methods and new formulations of the SLAM**

this tutorial. Section V describes a number of important real-world implementations of SLAM and also highlights implementations where the sensor data and software are freely down-loadable for other researchers to study. Part II of this tutorial describes major issues in computation, convergence and data association in SLAM. These are subjects that have been the main focus of the SLAM research community over the past five years.

Past, Present, and Future of Simultaneous Localization And Mapping: Towards the Robust-Perception Age

Cesar Cadena, Luca Carlone, Henry Carrillo, Yasir Latif, Davide Scaramuzza, José Neira, Ian Reid, John J. Leonard

***Abstract*—Simultaneous Localization And Mapping (SLAM) consists in the concurrent construction of a model of the environment (the *map*), and the estimation of the state of the robot moving within it. The SLAM community has made astonishing progress over the last 30 years, enabling large-scale real-world applications, and witnessing a steady transition of this technology**

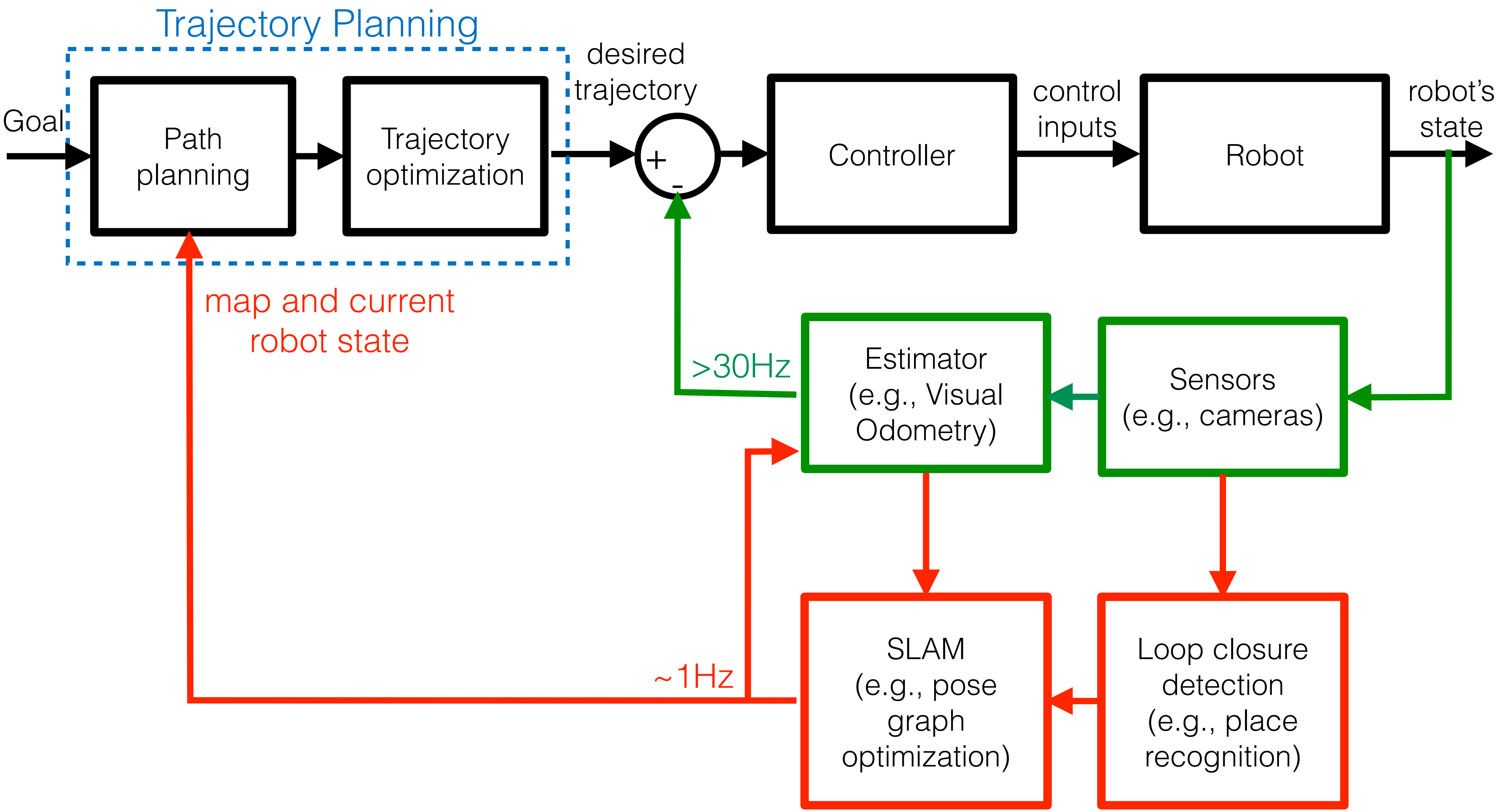
I. Introduction

SLAM comprises the simultaneous estimation of the state of a robot equipped with on-board sensors, and the construction of a model (the *map*) of the environment that the

> "The genesis of the probabilistic SLAM problem occurred at the 1986 IEEE Robotics and Automation Conference held in San Francisco. This was a time when probabilistic methods were only just beginning to be introduced into both robotics and AI. A number of researchers had been looking at applying estimation-theoretic methods to mapping and localisation problems; these included Peter Cheeseman, Jim Crowley, and Hugh Durrant-Whyte. *Over the course of the conference many paper table cloths and napkins were filled with long discussions about consistent mapping.* Along the way, Raja Chatila, Oliver Faugeras, Randal Smith and others also made useful contributions to the conversation."

MIT ASTRO

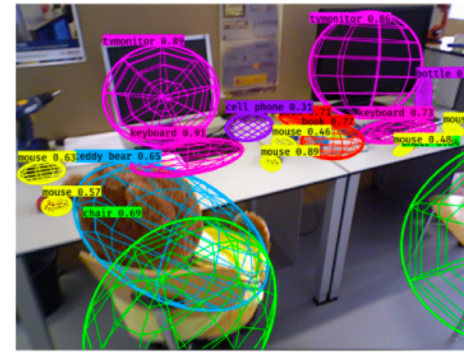
Big Picture


Trajectory Planning
Goal
Path planning
Trajectory optimization
desired trajectory
+
-
Controller
control inputs
Robot
robot's state
map and current robot state
>30Hz
Estimator (e.g., Visual Odometry)
Sensors (e.g., cameras)
~1Hz
SLAM (e.g., pose graph optimization)
Loop closure detection (e.g., place recognition)

"Map": Environment Representations

- **Sparse**:
  - Landmark-based
  - No explicit representation (pose graph)
  - Geometric primitives

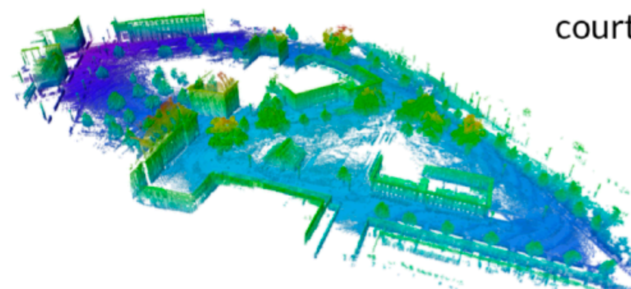
courtesy of Ranganathan et al.



courtesy of Nicholson et al.

- **Dense**:
  - Point clouds
  - 2D/3D occupancy grids
  - 3D meshes
  - ....

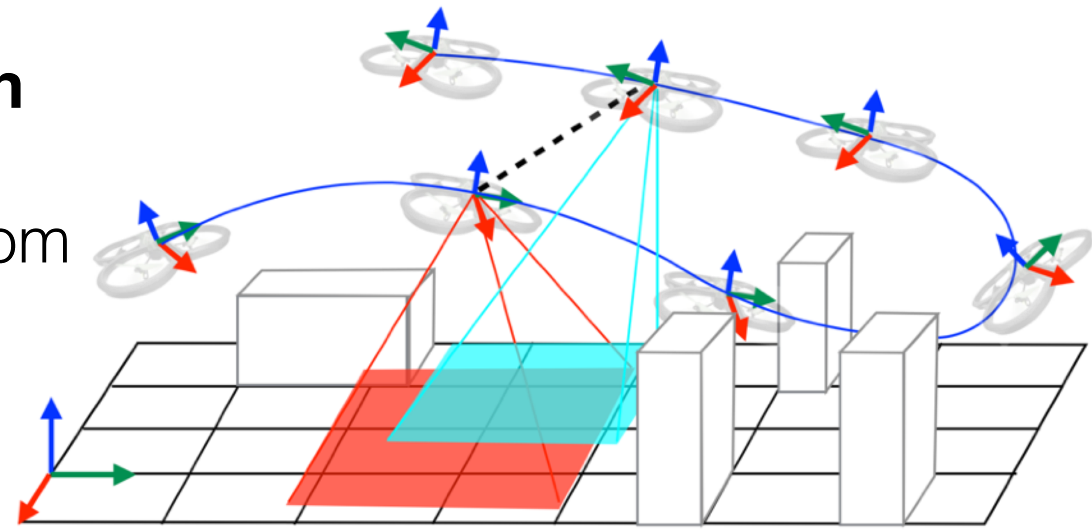
courtesy of Sameer Agarwal

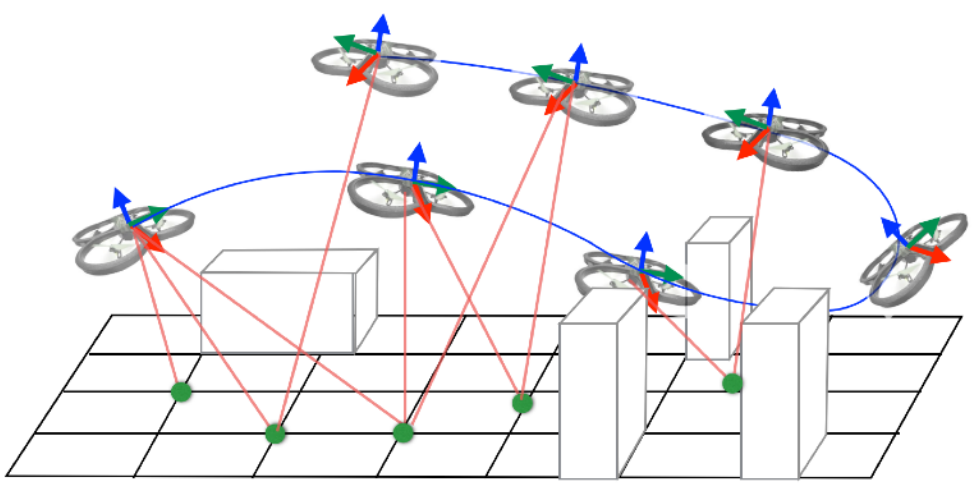


courtesy of octomap

Simultaneous Localization and Mapping

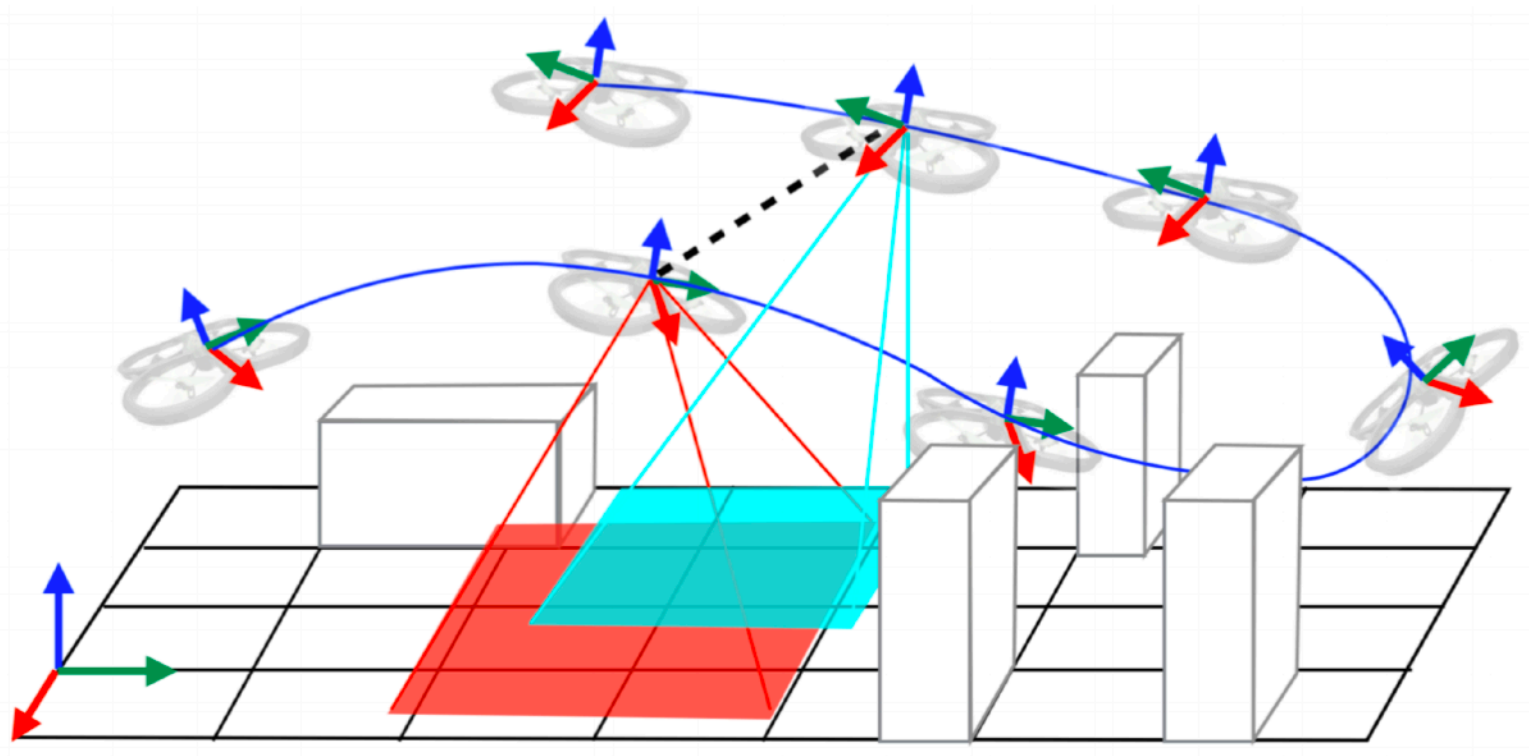
Pose Graph Optimization
(a.k.a. pose SLAM):
Estimate only <u>trajectory</u> from sensor data





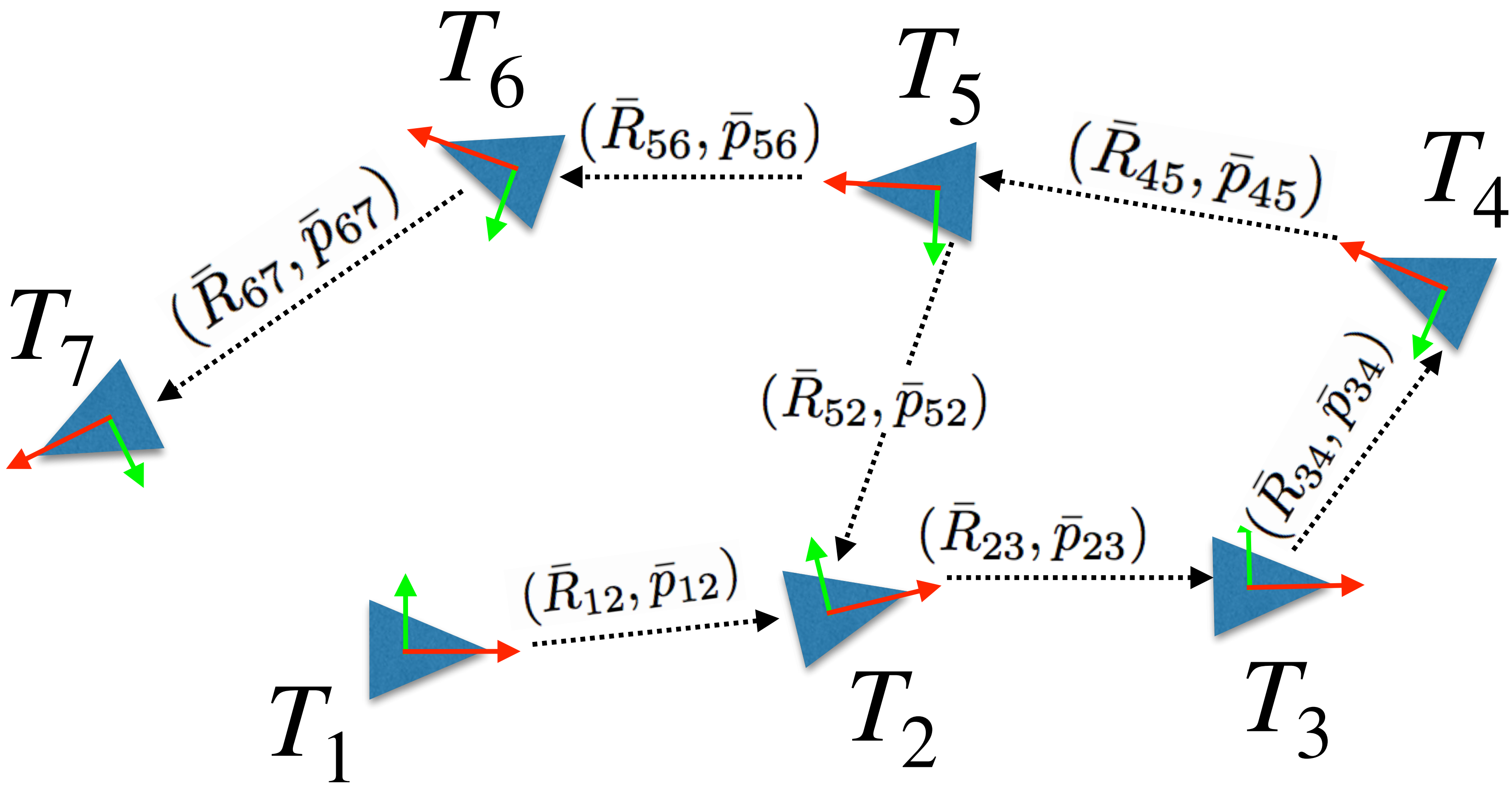
Landmark-based SLAM:
Estimate <u>trajectory</u> of robot and position of external <u>landmarks</u> from sensor data

Pose Graph Optimization



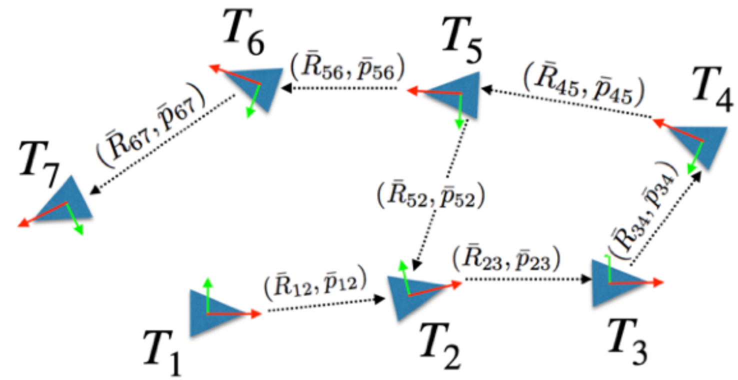
- **Measurements**: odometry + loop closures (i.e., relative pose measurements between non-consecutive poses obtained via place recognition & 2-view geometry, or similar)
- **Variables**: robot poses

Graphical representation of pose graph optimization

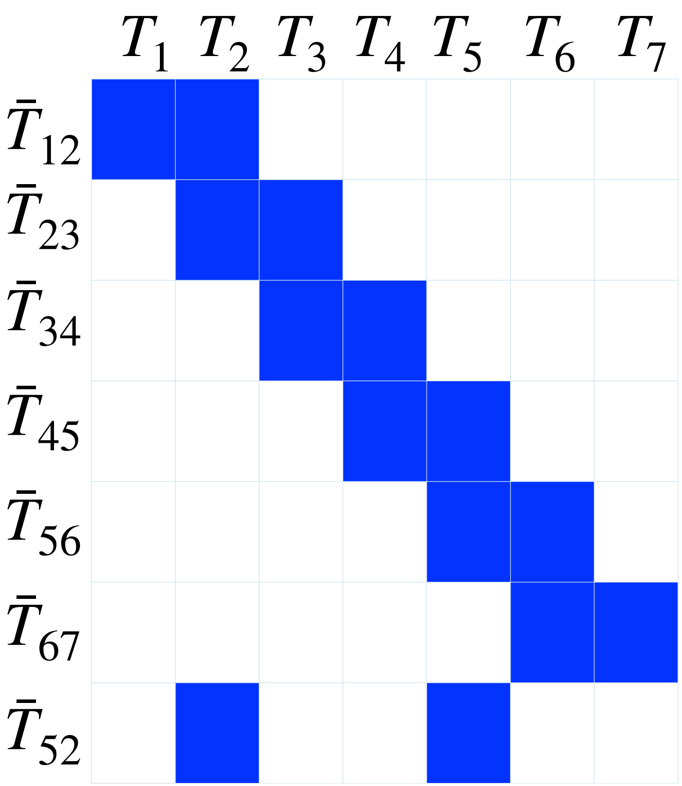


pose graph

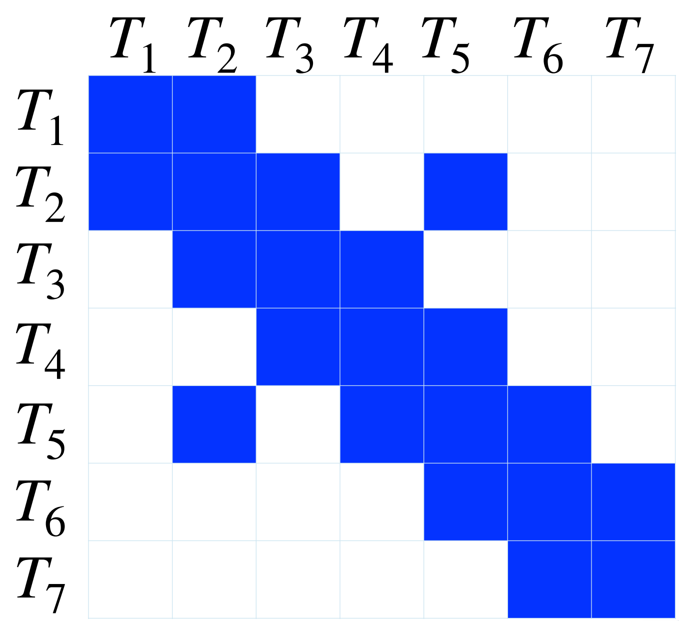
Pose Graph Optimization: Sparsity



Jacobian **J**

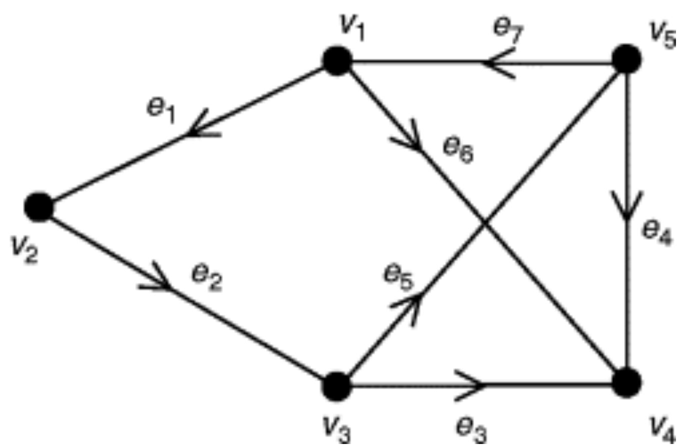


Hessian $\mathbf{J}^T \mathbf{J}$

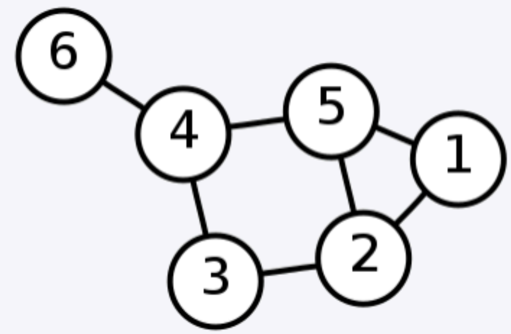


a.k.a. Information Matrix of the estimate

Graphical representation of pose graph optimization



Labelled graph



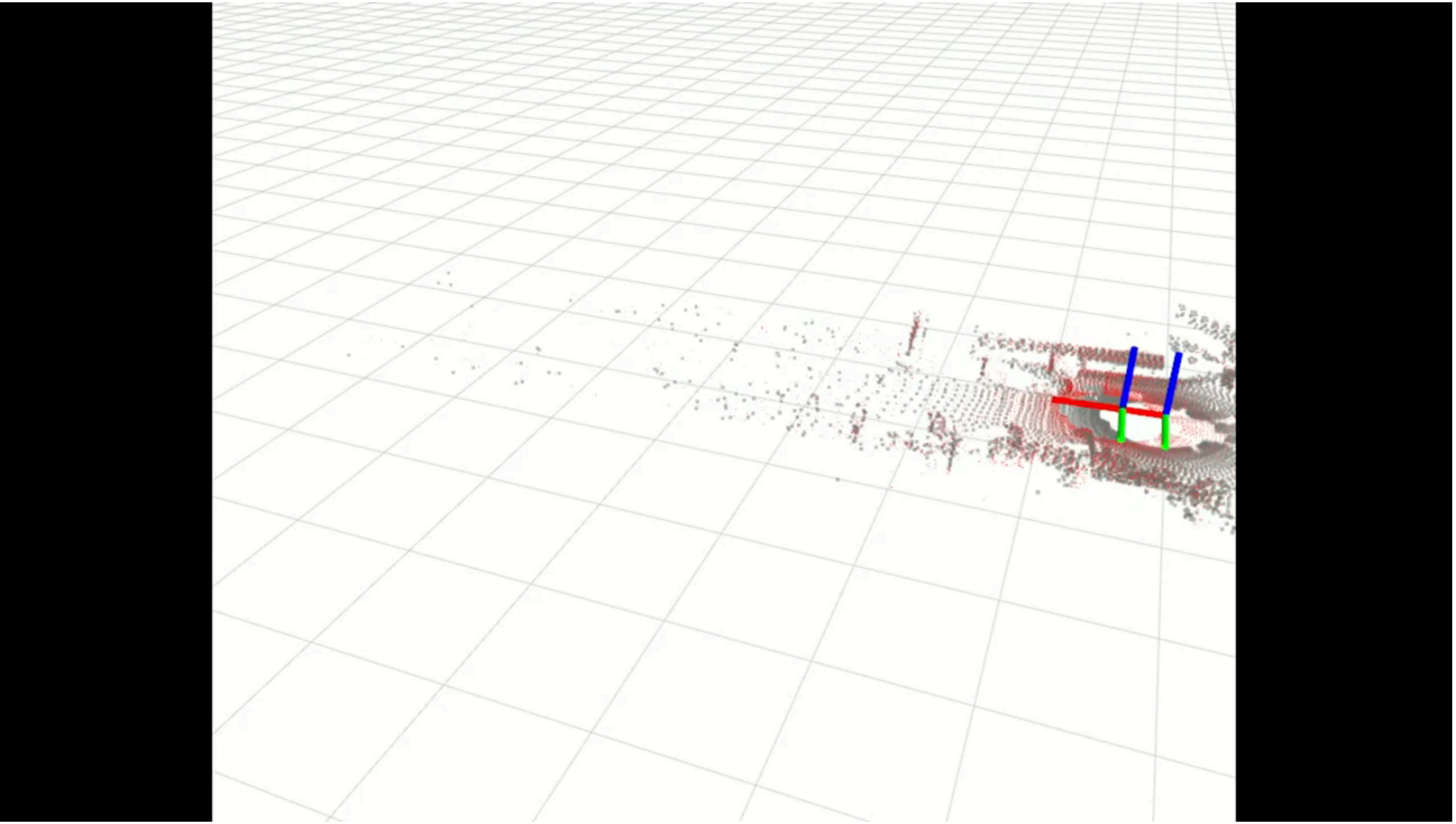
Incidence Matrix

	e_1	e_2	e_3	e_4	e_5	e_6	e_7
v_1	1	0	0	0	0	1	–1
v_2	–1	1	0	0	0	0	0
v_3	0	–1	1	0	1	0	0
v_4	0	0	–1	–1	0	–1	0
v_5	0	0	0	1	–1	0	1

Laplacian matrix

$$\begin{pmatrix} 2 & -1 & 0 & 0 & -1 & 0 \\ -1 & 3 & -1 & 0 & -1 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -1 & -1 \\ -1 & -1 & 0 & -1 & 3 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{pmatrix}$$

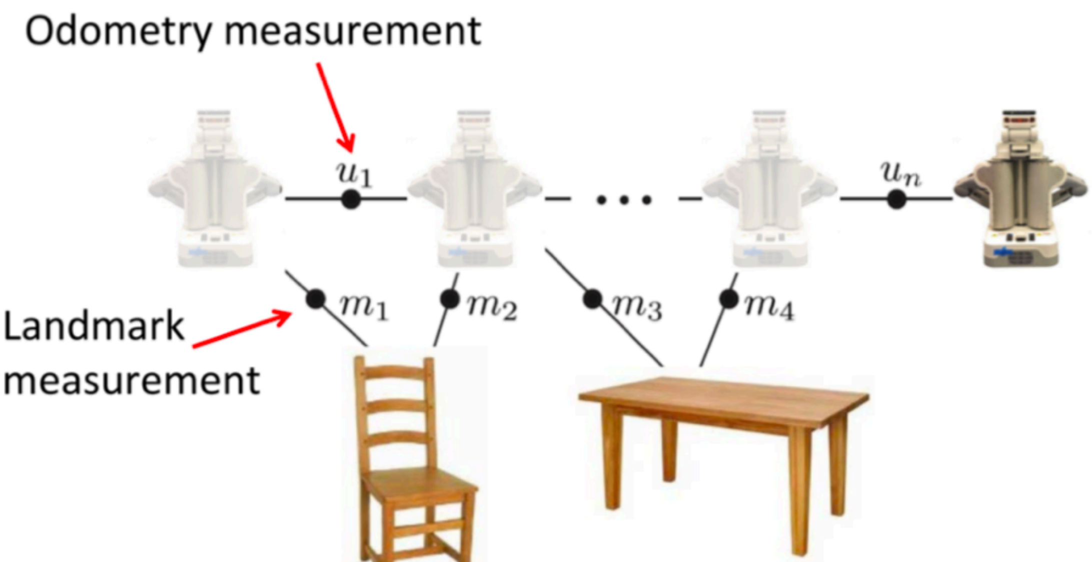
Pose Graph Optimization: Example



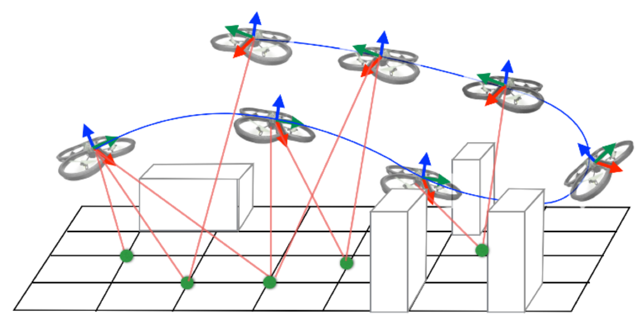
https://www.youtube.com/watch?v=KYvOqUB_odg

Landmark-based SLAM

- Sequence of robot (camera) poses $\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_t \in \mathrm{SE}(d)$
- Robot measures the relative pose between $\mathbf{T}_i$ and $\mathbf{T}_{i+1}$ (odometry)
- Robot measures the environment (e.g., point landmarks $\mathbf{p}_i \in \mathbb{R}^d$)



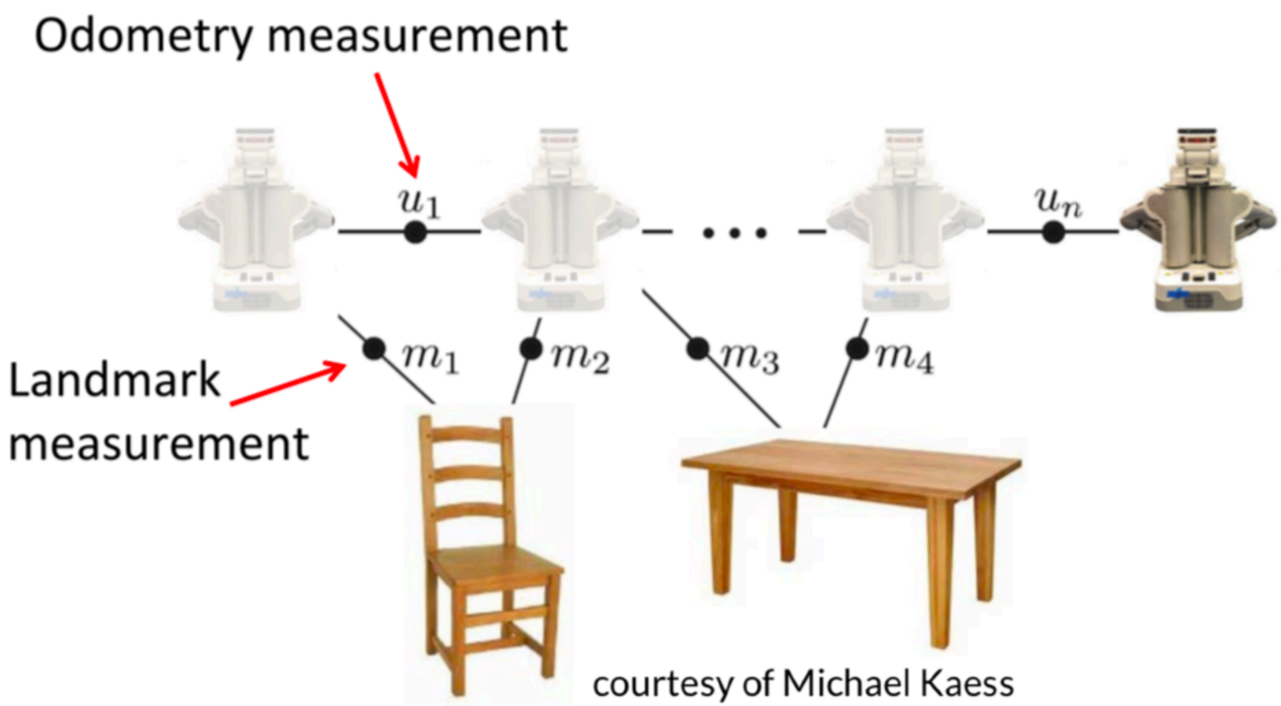
courtesy of Michael Kaess



- **Measurements**: odometry + measurements of (projection, range, position, or others) of external landmarks
- **Variables**: robot poses and landmark positions

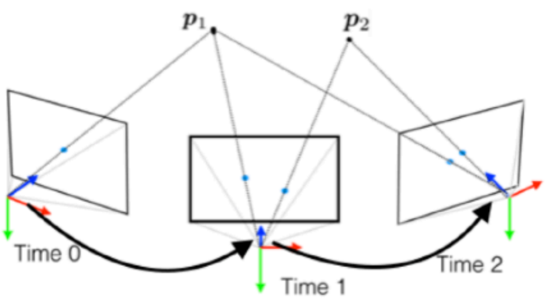
MIT AEROASTRO

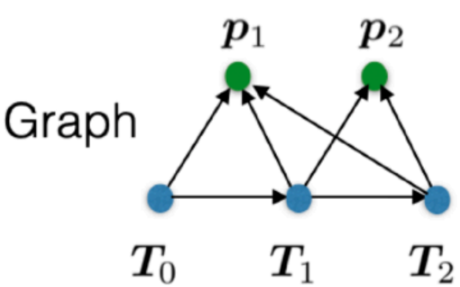
Graphical representation of landmark-based SLAM



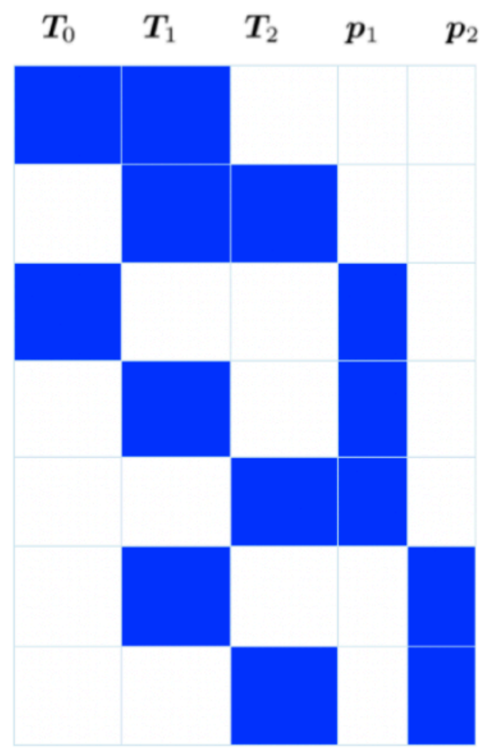
- Each variable (robot pose, landmark position/pose) is a node in the graph
- Each (usually) pairwise measurement denotes an edge between the corresponding two variables (nodes)

Graphical representation of landmark-based SLAM

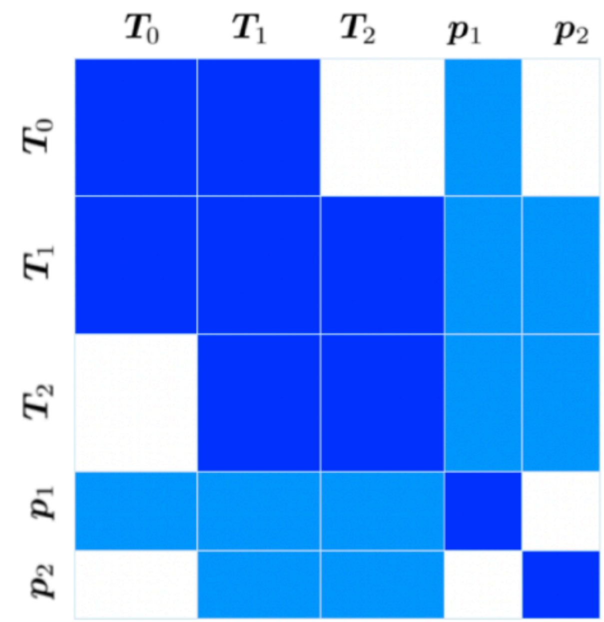




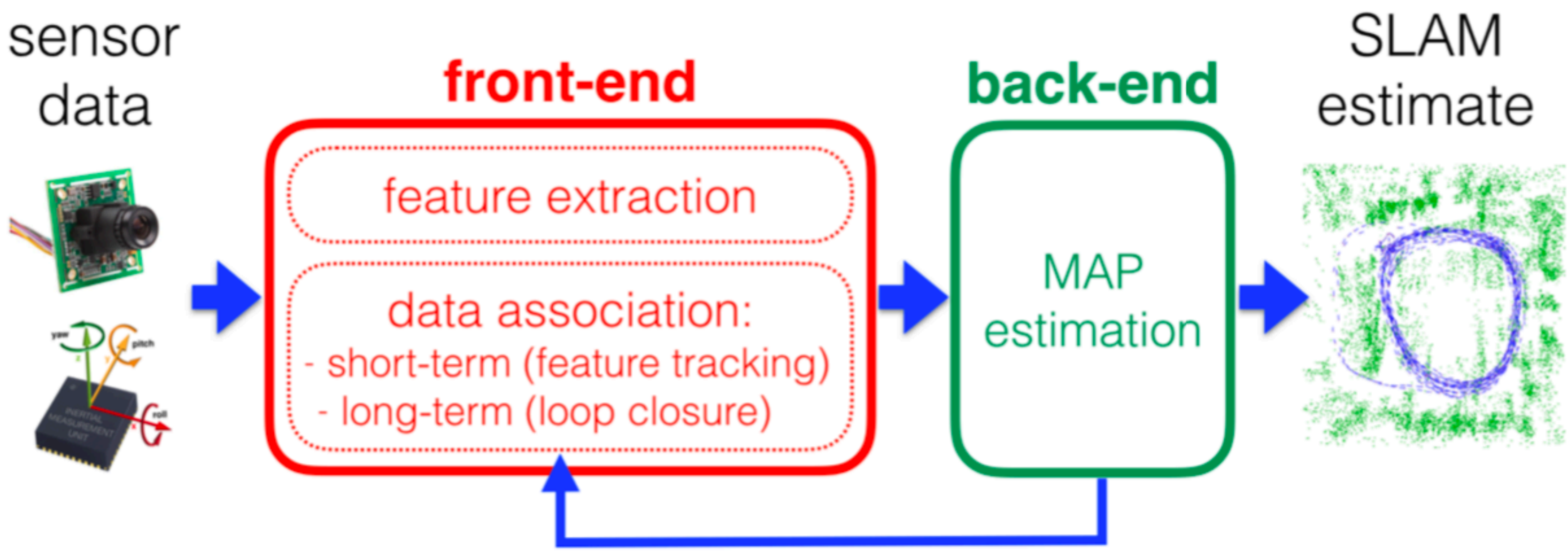
Jacobian **J**



Hessian **J**



Some terminology

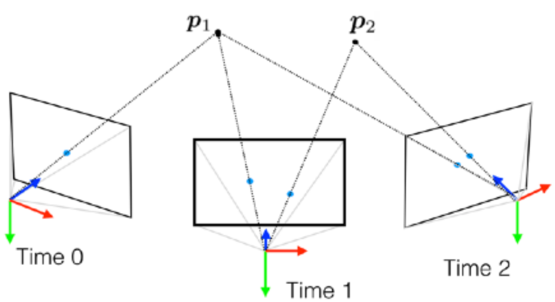


MAP is maximum *a posteriori* estimation
(MLE if no prior is available ["uninformative" prior])

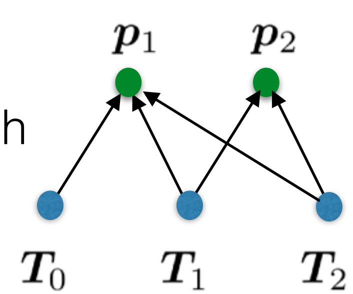
courtesy of Cadena et al.

Backup

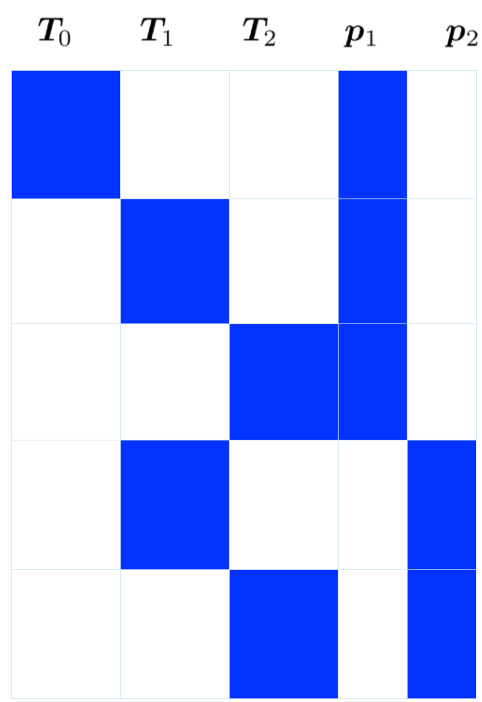
Windowed Bundle Adjustment



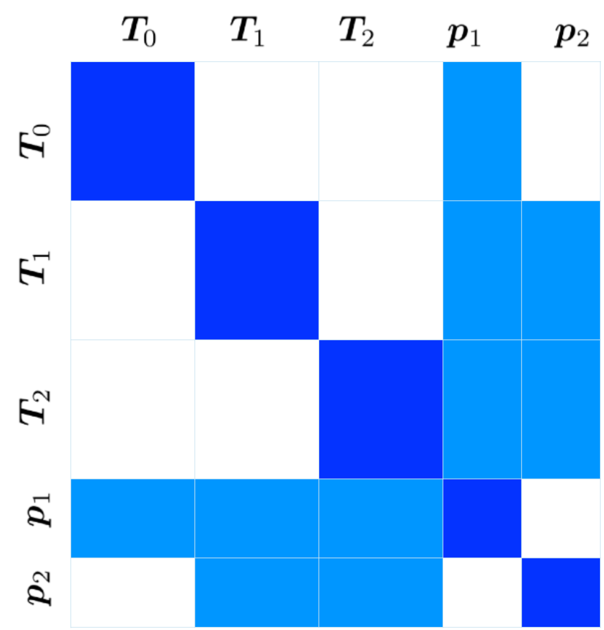
Graph



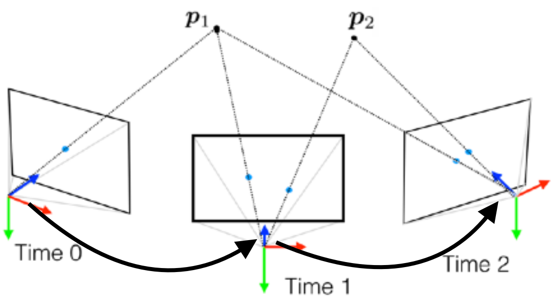
Jacobian **J**



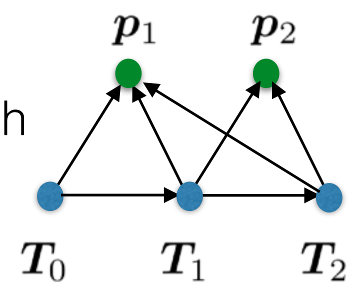
Hessian **J^TJ**



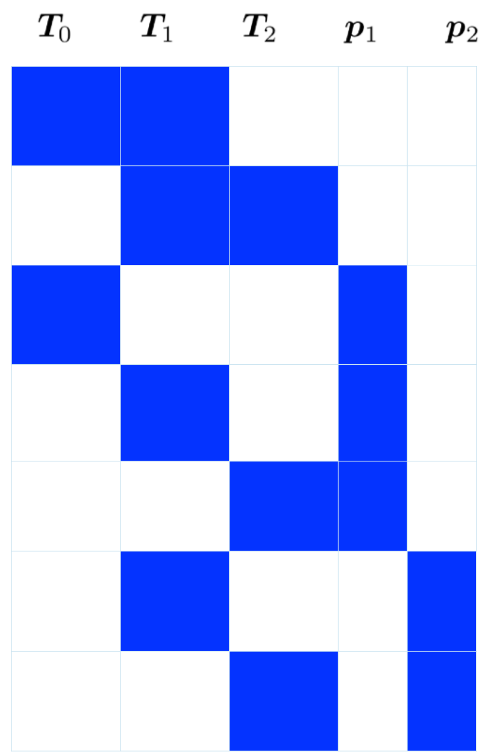
Windowed Bundle Adjustment



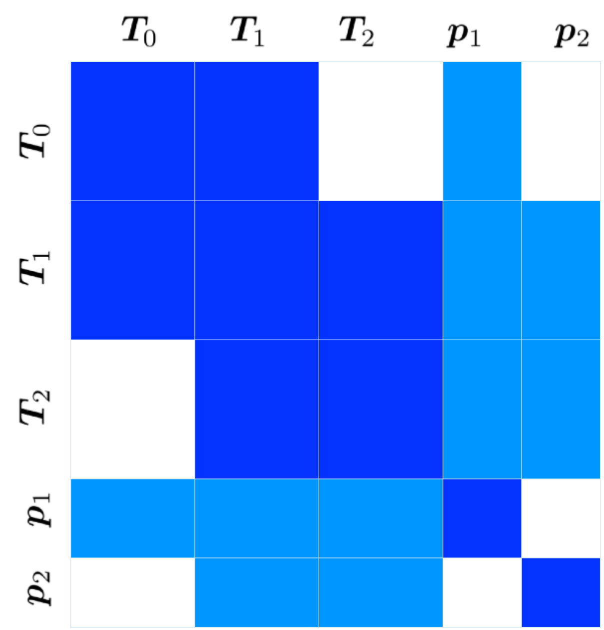
Graph



Jacobian **J**

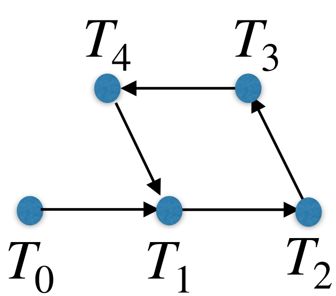


Hessian **J^TJ**



Windowed Bundle Adjustment

Graph



Transpose Incidence Matrix

	T_0	T_1	T_2	T_3	T_4
$\bar{T}_{01}$	-1	1			
$\bar{T}_{12}$		-1	1		
$\bar{T}_{23}$			-1	1	
$\bar{T}_{34}$				-1	1
$\bar{T}_{41}$		1			-1

Laplacian Matrix

	T_0	T_1	T_2	T_3	T_4
T_0	1	-1			
T_1	-1	3	-1		-1
T_2		-1	2	-1	
T_3			-1	2	-1
T_4		-1		-1	2

A

Cont'd

- $\widetilde{\mathbf{T}}_{ij} \to$ measured pose between $\mathbf{T}_i$ and $\mathbf{T}_j$

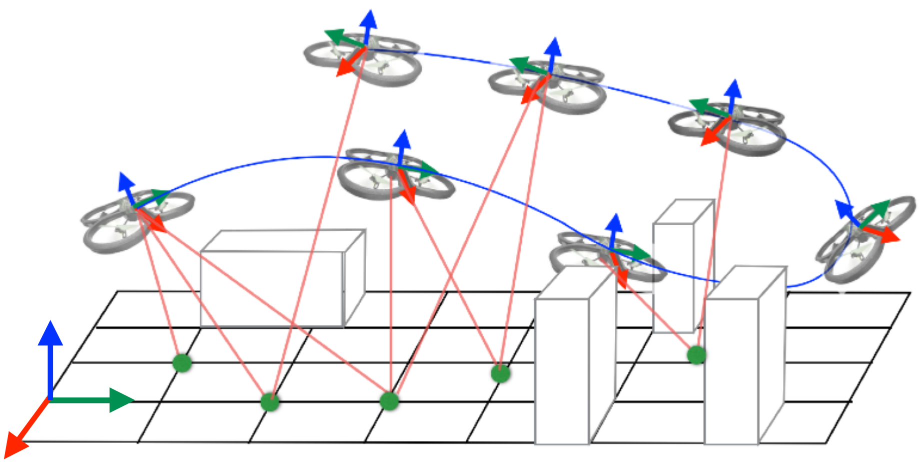
- $\widetilde{\mathbf{T}}_{ij} = \mathbf{T}_i^{-1}\mathbf{T}_j \exp(\widehat{\boldsymbol{\epsilon}_{ij}})$ where $\boldsymbol{\epsilon}_{ij} \sim \mathcal{N}(\mathbf{0}, \Sigma_{ij})$
- $\|\mathbf{r}_{ij}\|^2_{\Sigma_{ij}^{-1}}$ where $\widehat{\mathbf{r}_{ij}} = \log_{\mathsf{SE}(3)}(\mathbf{T}_j^{-1}\mathbf{T}_i\widetilde{\mathbf{T}}_{ij}) = -\log_{\mathsf{SE}(3)}(\widetilde{\mathbf{T}}_{ij}^{-1}\mathbf{T}_i^{-1}\mathbf{T}_j)$

- Other noise models (and thus residual/MLE formulations) also exist and are commonly used — we'll see one next week

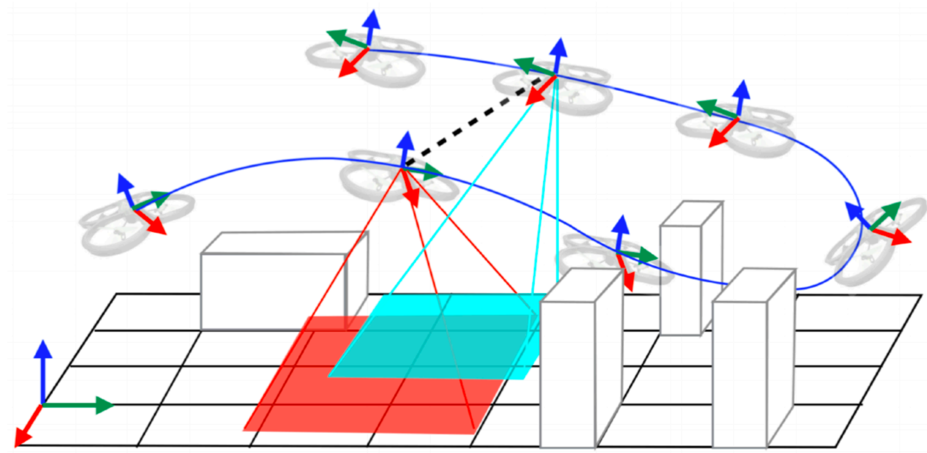
- e.g., Langevin noise for rotational measurements and additive Gaussian noise for translational measurements (i.e., similar to odometry measurement residuals in **Lab 9** individual deliverable)

- e.g., another commonly used model uses wrapped Gaussian on rotational measurements ($\mathrm{SO}(3)$) and additive Gaussian on translational measurements

A

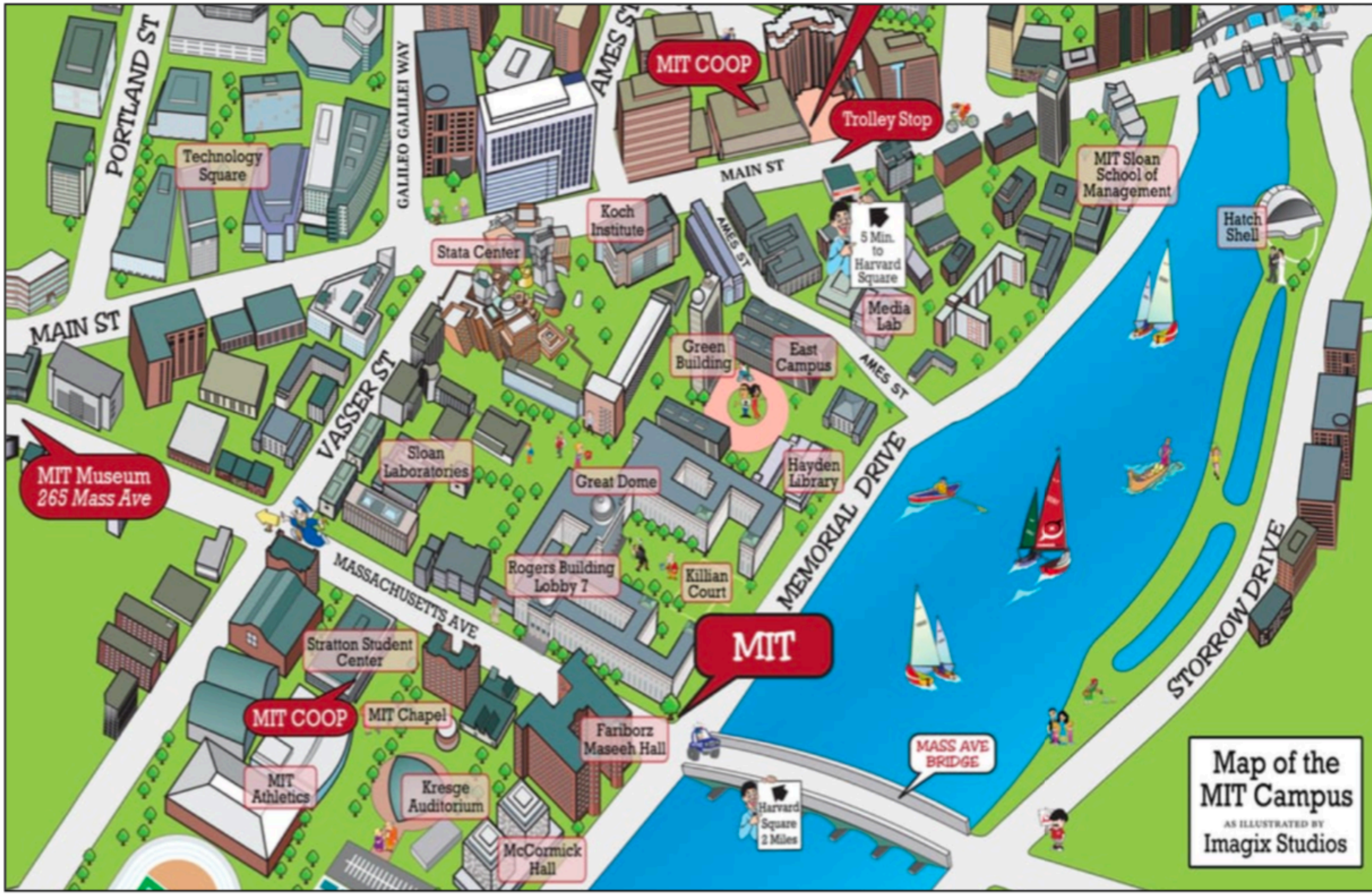


(a)



(b)

Loop Closure



On board?

Typical Back-End (MLE)

$$f(\mathbf{x}) = \sum_{(i,j)\in E} \|\mathbf{z}_{ij} \boxminus \mathbf{h}_{ij}(\mathbf{x}_i,\mathbf{x}_j)\|^2_{\Sigma_{ij}^{-1}}$$

- $\mathbf{x}_1, \ldots, \mathbf{x}_n \rightarrow$ robot poses, landmark positions, …
- $\mathbf{z}_{ij} \rightarrow$ actual measurement
- $\mathbf{h}_{ij} \rightarrow$ measurement model
- $\Sigma_{ij} \rightarrow$ noise covariance matrix
- $\mathbf{z}_{ij} \boxminus \mathbf{h}_{ij}(\mathbf{x}_i,\mathbf{x}_j) \rightarrow$ residual (error)
- $\boxminus \rightarrow$ "generalized $-$"
  would be defined differently based on the specific measurement model

MIT AEROASTRO

A

Typical Back-End (MAP)

$$f(\mathbf{x}) = \sum_{(i,j)\in E} \|\mathbf{z}_{ij} \boxminus \mathbf{h}_{ij}(\mathbf{x}_i,\mathbf{x}_j)\|^2_{\Sigma_{ij}^{-1}} + \sum_{i} \underbrace{\|\mathbf{x}_i \boxminus \mathbf{s}_i\|^2_{\Sigma_i^{-1}}}_{\text{potential priors for } \mathbf{x}_i \text{ at } \mathbf{s}_i}$$

- $\mathbf{x}_1,\ldots,\mathbf{x}_n \rightarrow$ robot poses, landmark positions, …
- $\mathbf{z}_{ij} \rightarrow$ actual measurement
- $\mathbf{h}_{ij} \rightarrow$ measurement model
- $\Sigma_{ij} \rightarrow$ noise covariance matrix
- $\mathbf{z}_{ij} \boxminus \mathbf{h}_{ij}(\mathbf{x}_i,\mathbf{x}_j) \rightarrow$ residual (error)
- $\boxminus \rightarrow$ "generalized $-$"
  would be defined differently based on the specific measurement model