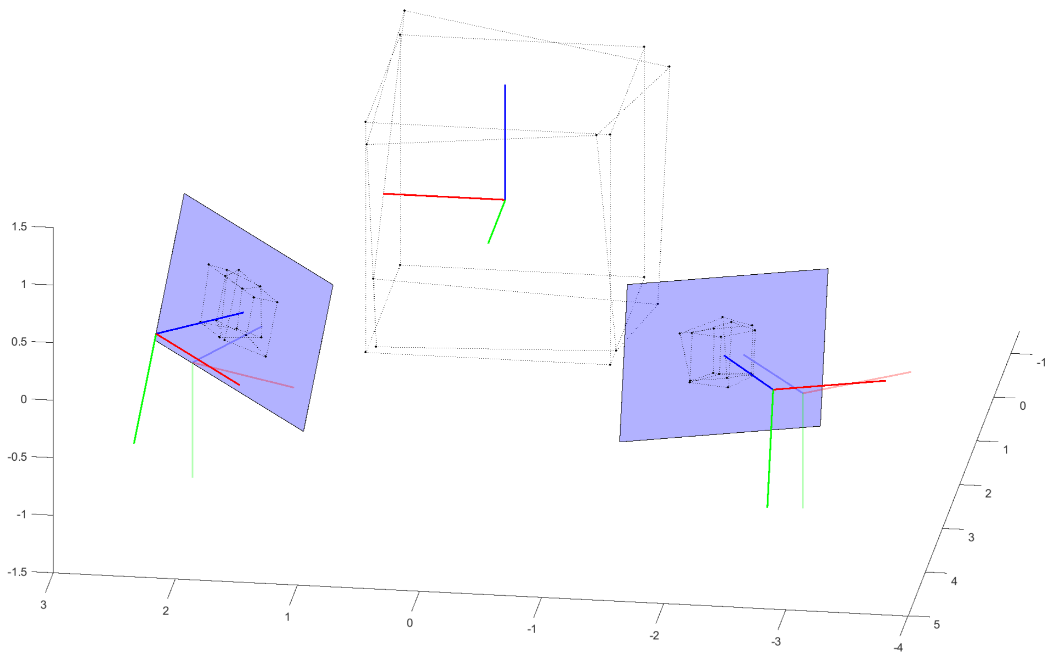
UiO : Department of Technology Systems
University of Oslo

Lecture 9.2
Full bundle adjustment

Trym Vegard Haavardsholm

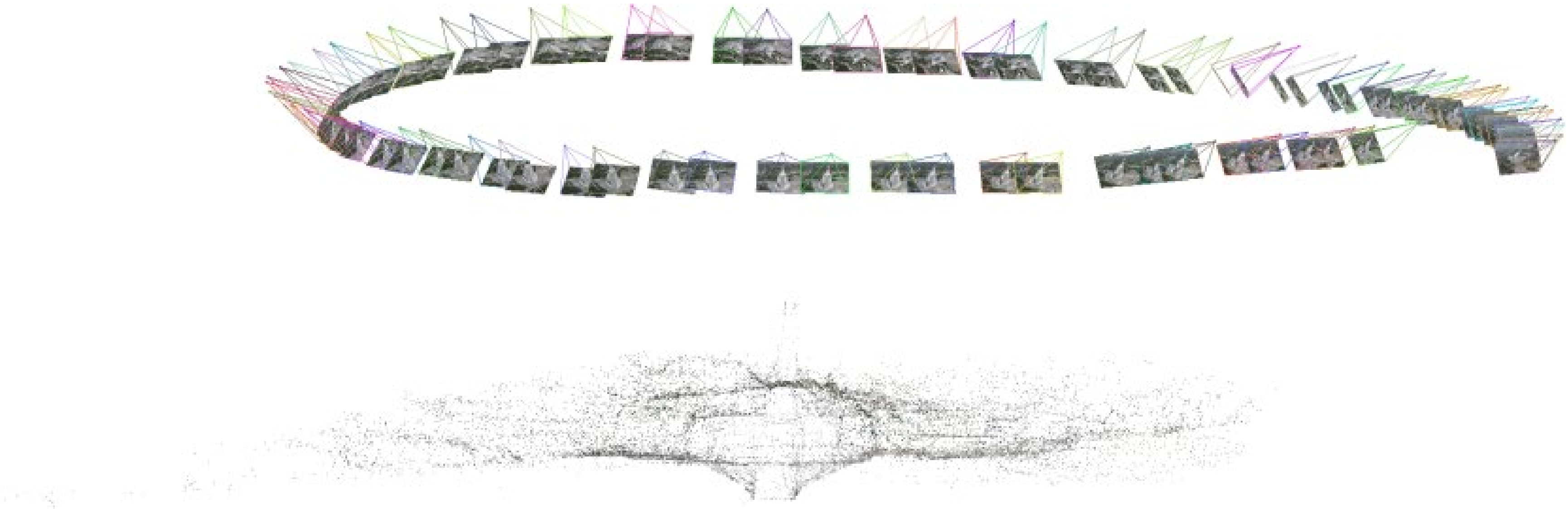


Bundle adjustment

Bundle Adjustment (BA)

Estimating the imaging geometry
based on minimizing reprojection error

- Motion-only BA
- Structure-only BA
- Full BA



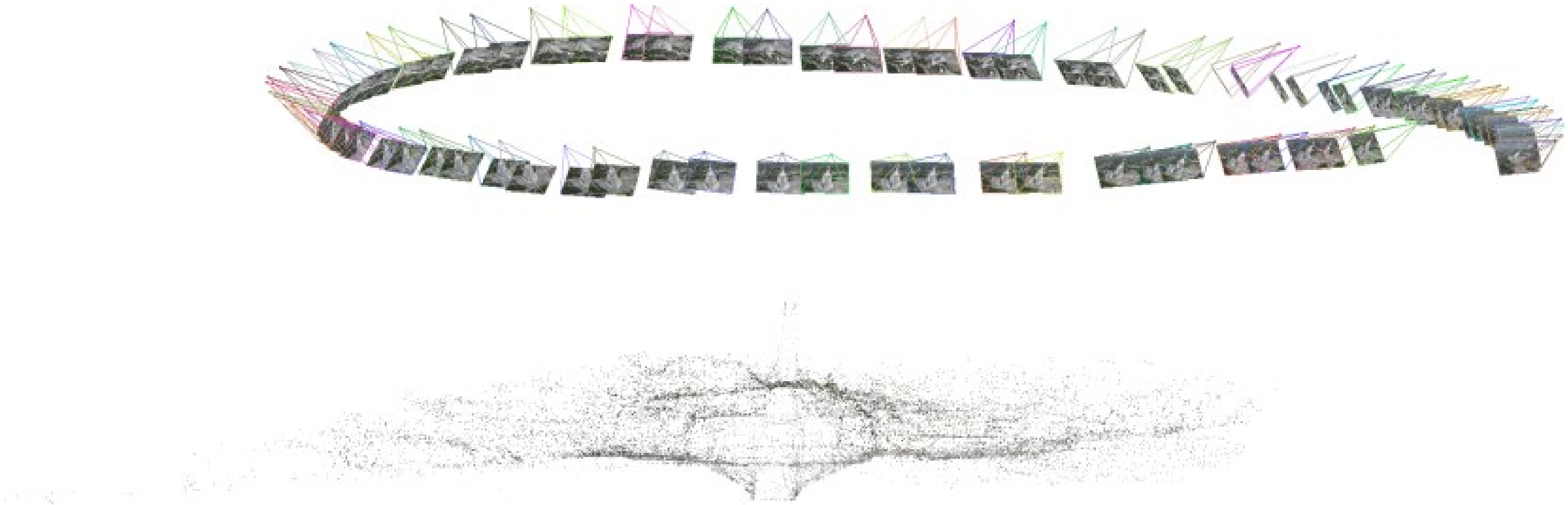
Bundle adjustment

$$\sum_i \sum_j \left\| \pi_i(\mathbf{T}_{wc_i}^{-1} \cdot \mathbf{x}_j^w) - \mathbf{u}_j^i \right\|^2$$

Bundle Adjustment (BA)

Estimating the imaging geometry based on minimizing reprojection error

- Motion-only BA
- Structure-only BA
- Full BA



Nonlinear MAP estimation

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

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

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

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

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

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

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

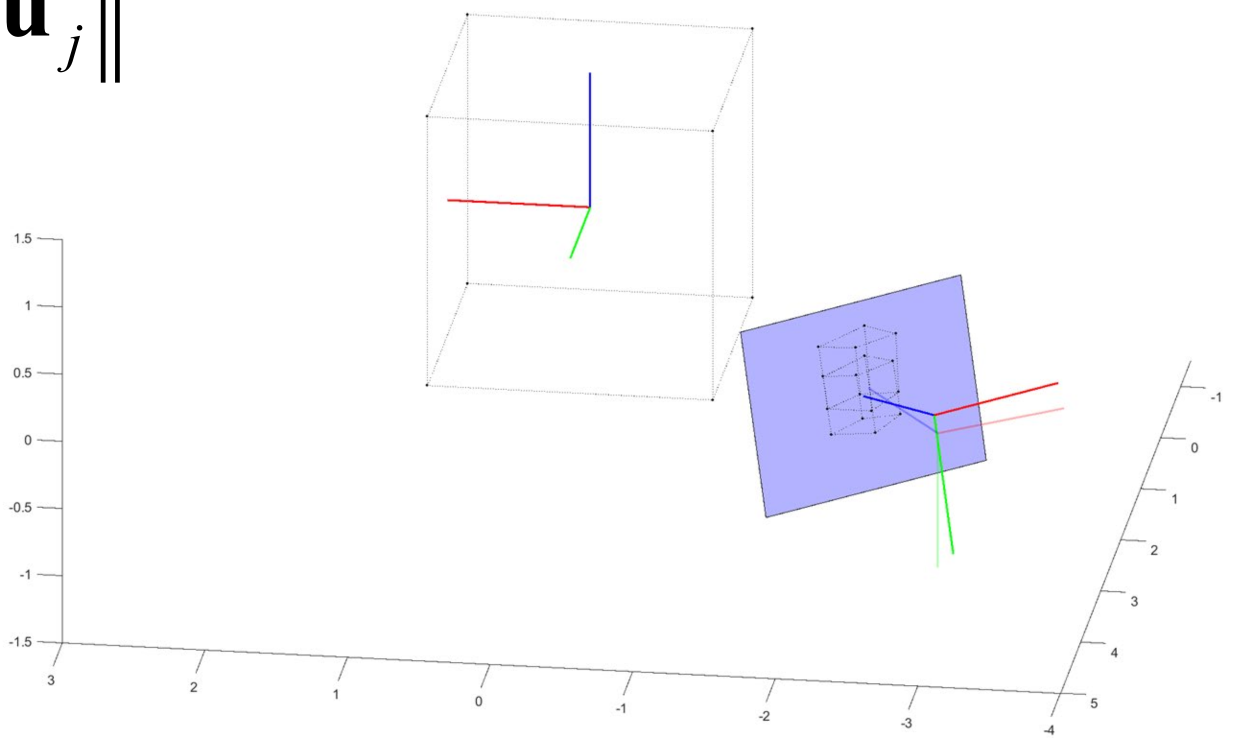
$$\hat{\underline{\mathcal{X}}}^{t+1} \leftarrow \hat{\underline{\mathcal{X}}}^{t} \oplus \underline{\boldsymbol{\tau}}^{*}$$

Pose estimation by minimizing reprojection error

Minimize **geometric error** over the **camera pose**

This is also sometimes called **Motion-Only Bundle Adjustment**

$$\mathbf{T}_{wc}^{*} = \underset{\mathbf{T}_{wc}}{\operatorname{argmin}} \sum_{j} \left\| \pi\left(\mathbf{T}_{wc}^{-1} \cdot \mathbf{x}_{j}^{w}\right) - \mathbf{u}_{j} \right\|^{2}$$

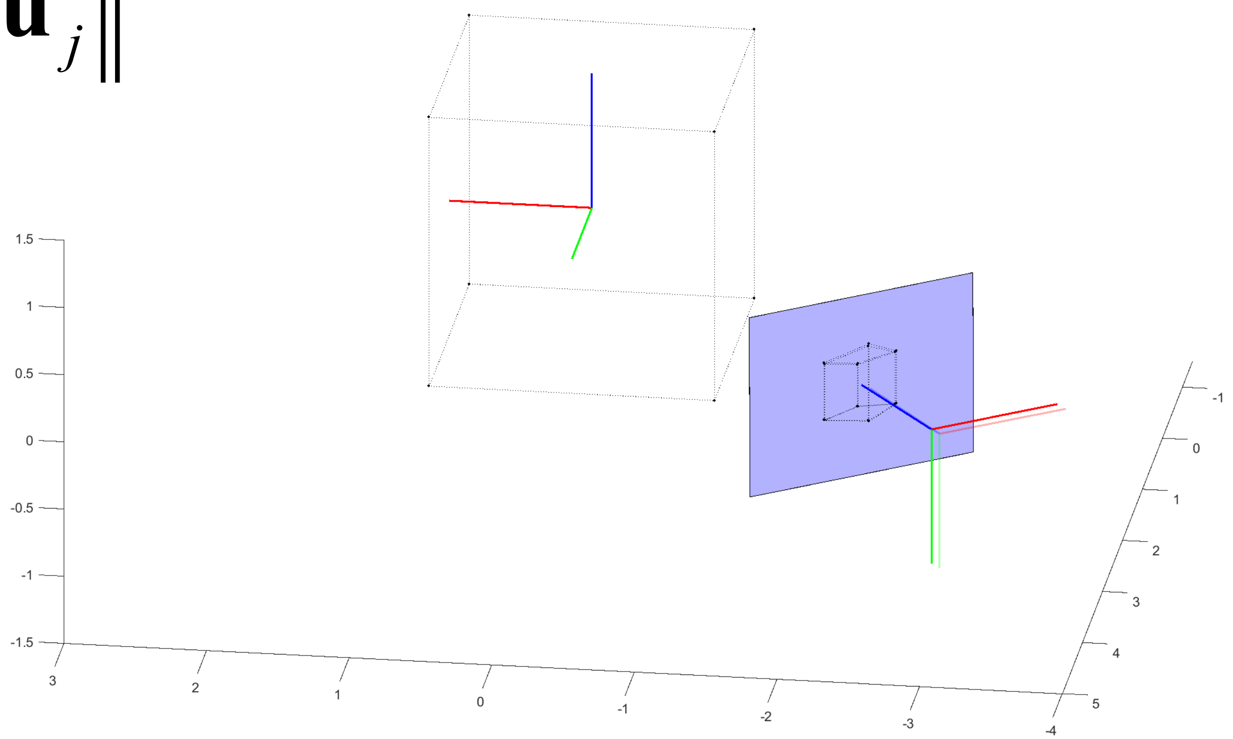


Pose estimation by minimizing reprojection error

Minimize **geometric error** over the **camera pose**

This is also sometimes called **Motion-Only Bundle Adjustment**

$$\mathbf{T}_{wc}^{*} = \underset{\mathbf{T}_{wc}}{\operatorname{argmin}} \sum_{j} \left\| \pi(\mathbf{T}_{wc}^{-1} \cdot \mathbf{x}_{j}^{w}) - \mathbf{u}_{j} \right\|^{2}$$

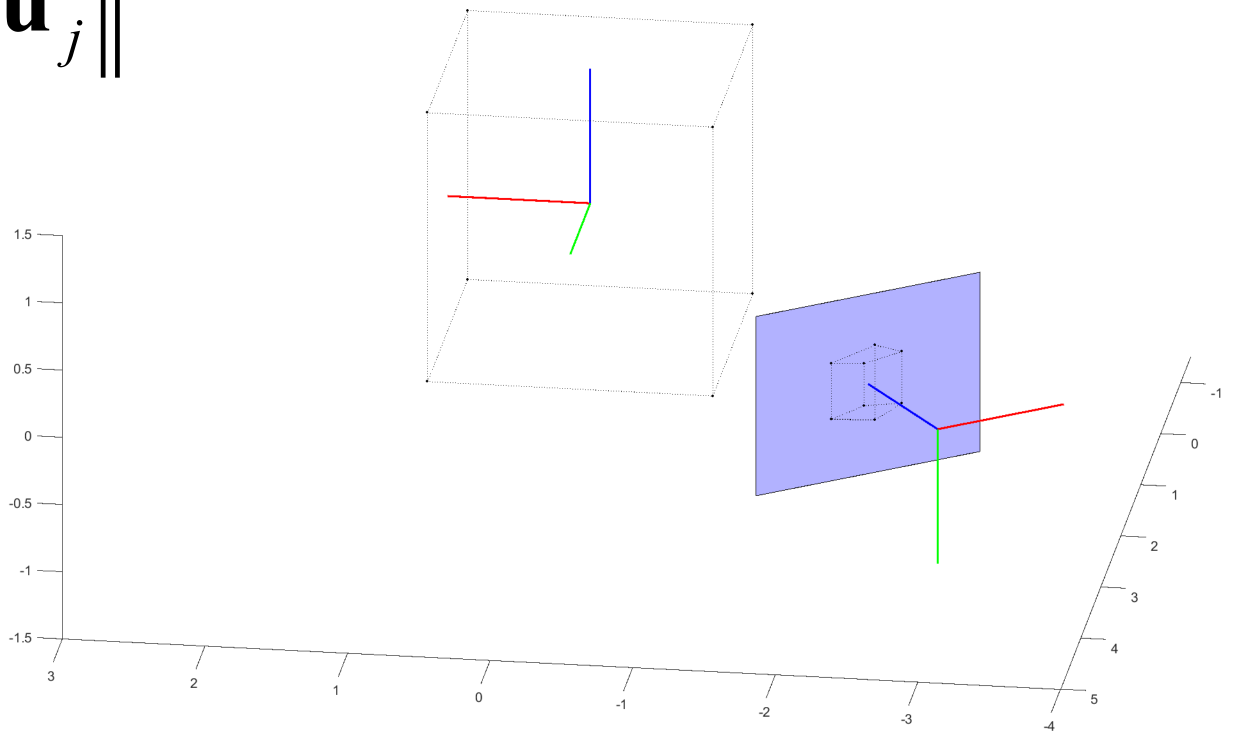


Pose estimation by minimizing reprojection error

Minimize **geometric error** over the **camera pose**

This is also sometimes called **Motion-Only Bundle Adjustment**

$$\mathbf{T}_{wc}^{*} = \underset{\mathbf{T}_{wc}}{\operatorname{argmin}} \sum_{j} \left\| \pi(\mathbf{T}_{wc}^{-1} \cdot \mathbf{x}_{j}^{w}) - \mathbf{u}_{j} \right\|^{2}$$

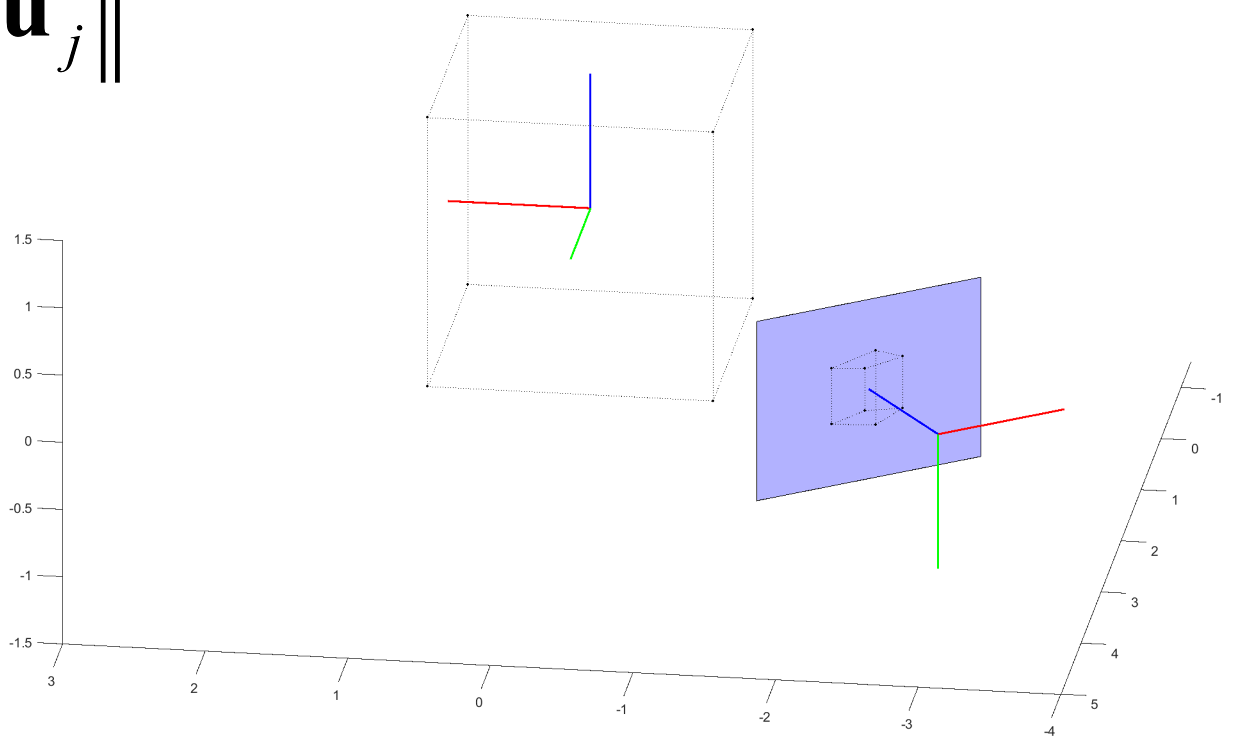


Pose estimation by minimizing reprojection error

Minimize **geometric error** over the **camera pose**

This is also sometimes called **Motion-Only Bundle Adjustment**

$$\mathbf{T}_{wc}^{*} = \underset{\mathbf{T}_{wc}}{\operatorname{argmin}} \sum_{j} \left\| \pi(\mathbf{T}_{wc}^{-1} \cdot \mathbf{x}_{j}^{w}) - \mathbf{u}_{j} \right\|^{2}$$

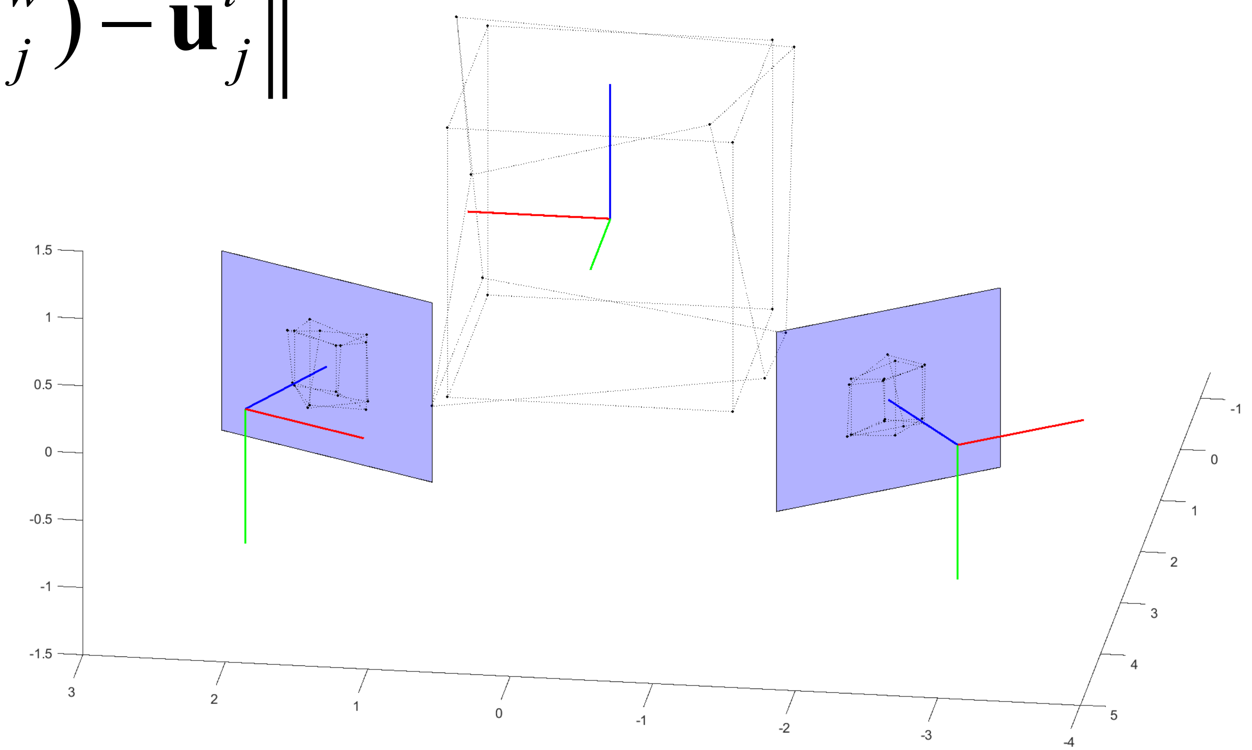


Triangulation by minimizing reprojection error

Minimize **geometric error** over the **world points**

This is also sometimes called **Structure-Only Bundle Adjustment**

$$\mathbf{x}_j^{w*} = \underset{\mathbf{x}_j^{w*}}{\operatorname{argmin}} \sum_i \sum_j \left\| \pi_i(\mathbf{T}_{wc_i}^{-1} \cdot \mathbf{x}_j^w) - \mathbf{u}_j^i \right\|^2$$

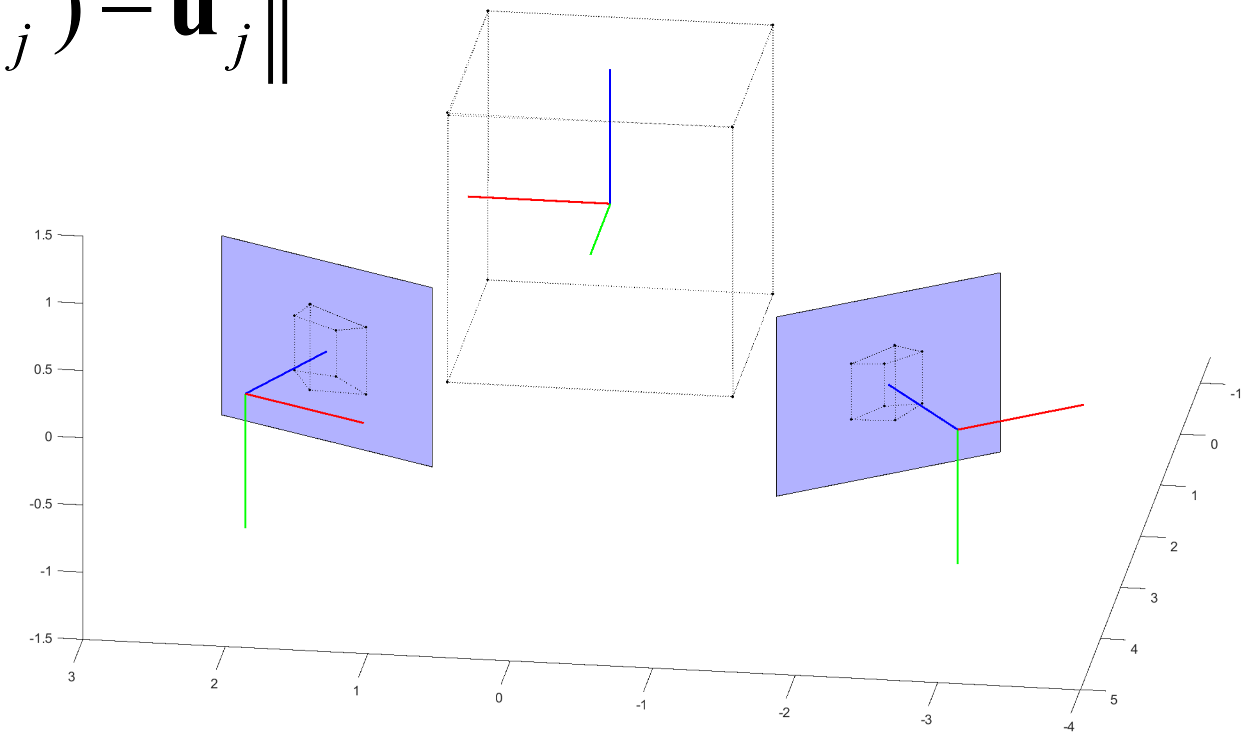


Triangulation by minimizing reprojection error

Minimize **geometric error** over the **world points**

This is also sometimes called **Structure-Only Bundle Adjustment**

$$\mathbf{x}_j^{w*} = \underset{\mathbf{x}_j^{w*}}{\operatorname{argmin}} \sum_i \sum_j \left\| \pi_i \left(\mathbf{T}_{wc_i}^{-1} \cdot \mathbf{x}_j^{w} \right) - \mathbf{u}_j^{i} \right\|^2$$

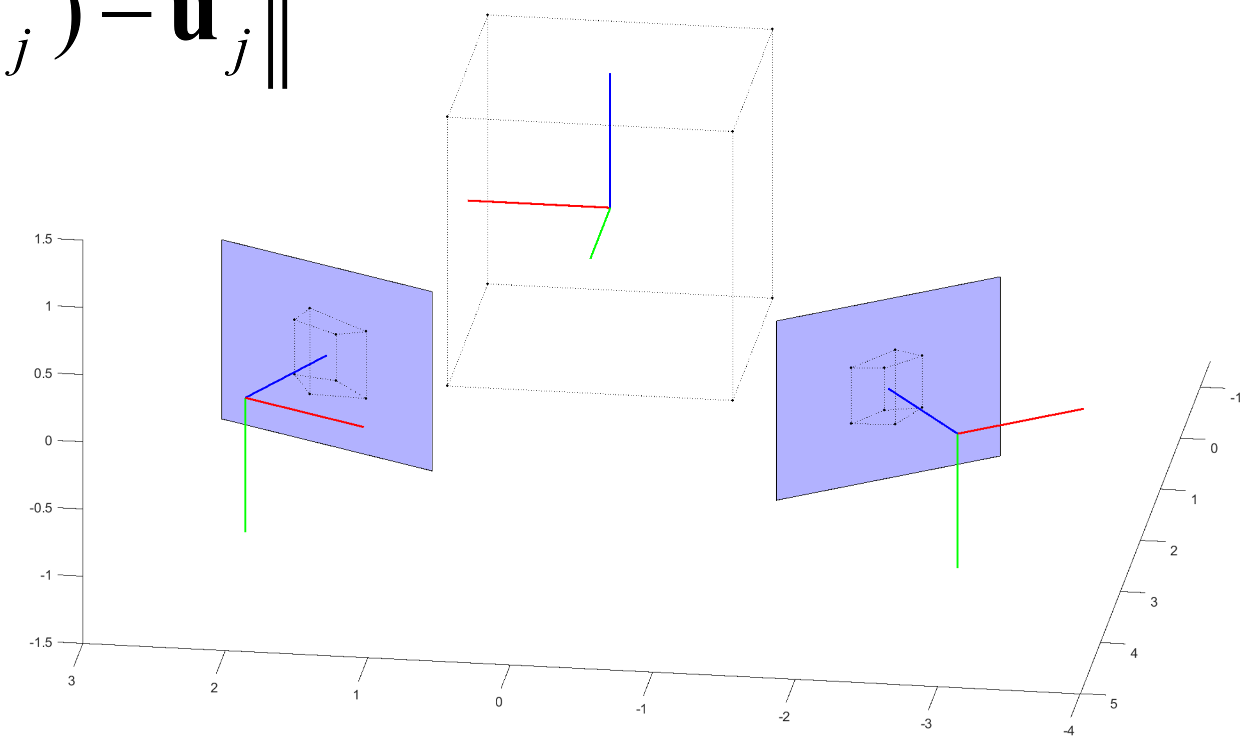


Triangulation by minimizing reprojection error

Minimize **geometric error** over the **world points**

This is also sometimes called **Structure-Only Bundle Adjustment**

$$\mathbf{x}_j^{w*} = \underset{\mathbf{x}_j^{w*}}{\operatorname{argmin}} \sum_i \sum_j \left\| \pi_i \left(\mathbf{T}_{wc_i}^{-1} \cdot \mathbf{x}_j^w \right) - \mathbf{u}_j^i \right\|^2$$

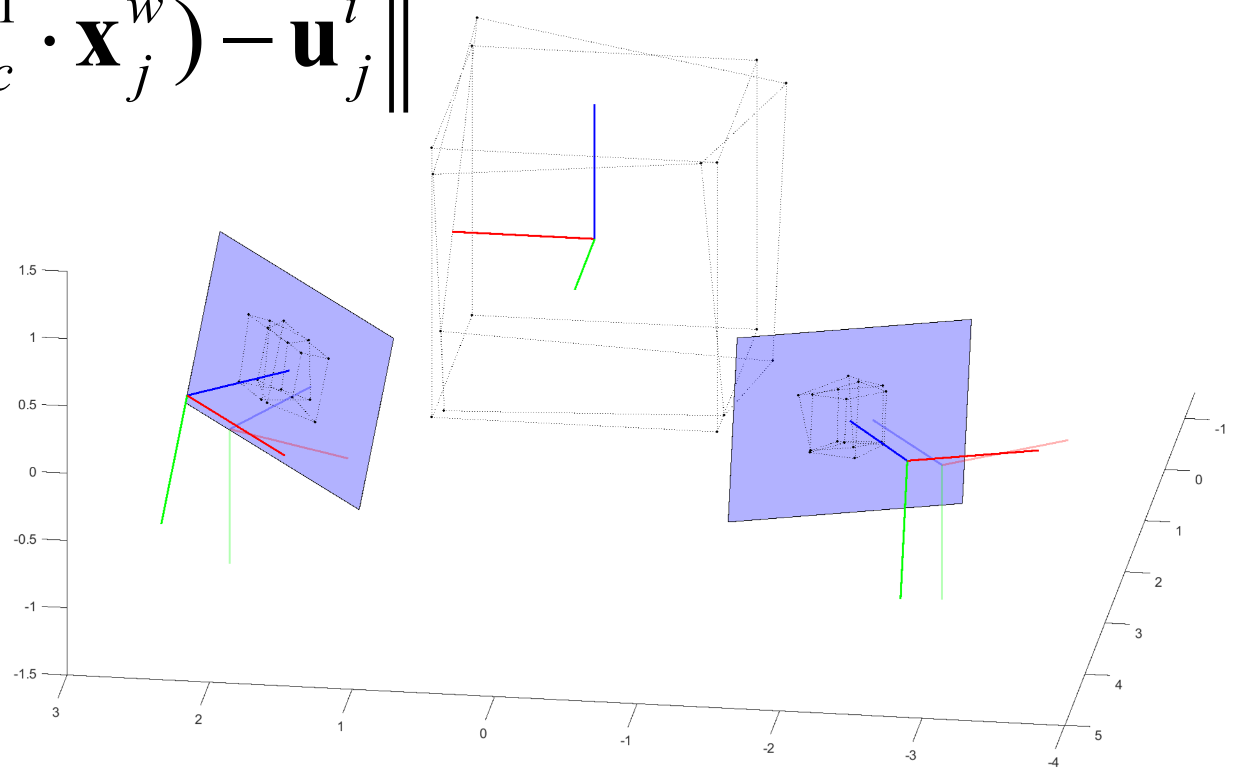


Pose *and* structure estimation by minimizing reprojection error

Minimize **geometric error** over the **camera poses** and **world points**

This is also sometimes called **Full Bundle Adjustment**

$$\left\{ \mathbf{T}_{wc_i}^{*}, \mathbf{x}_j^{w*} \right\} = \underset{\mathbf{T}_{wc_i}, \mathbf{x}_j^w}{\operatorname{argmin}} \sum_i \sum_j \left\| \pi_i \left(\mathbf{T}_{wc}^{-1} \cdot \mathbf{x}_j^w \right) - \mathbf{u}_j^i \right\|^2$$



Pose *and* structure estimation by minimizing reprojection error

Given:

–

Measurements:

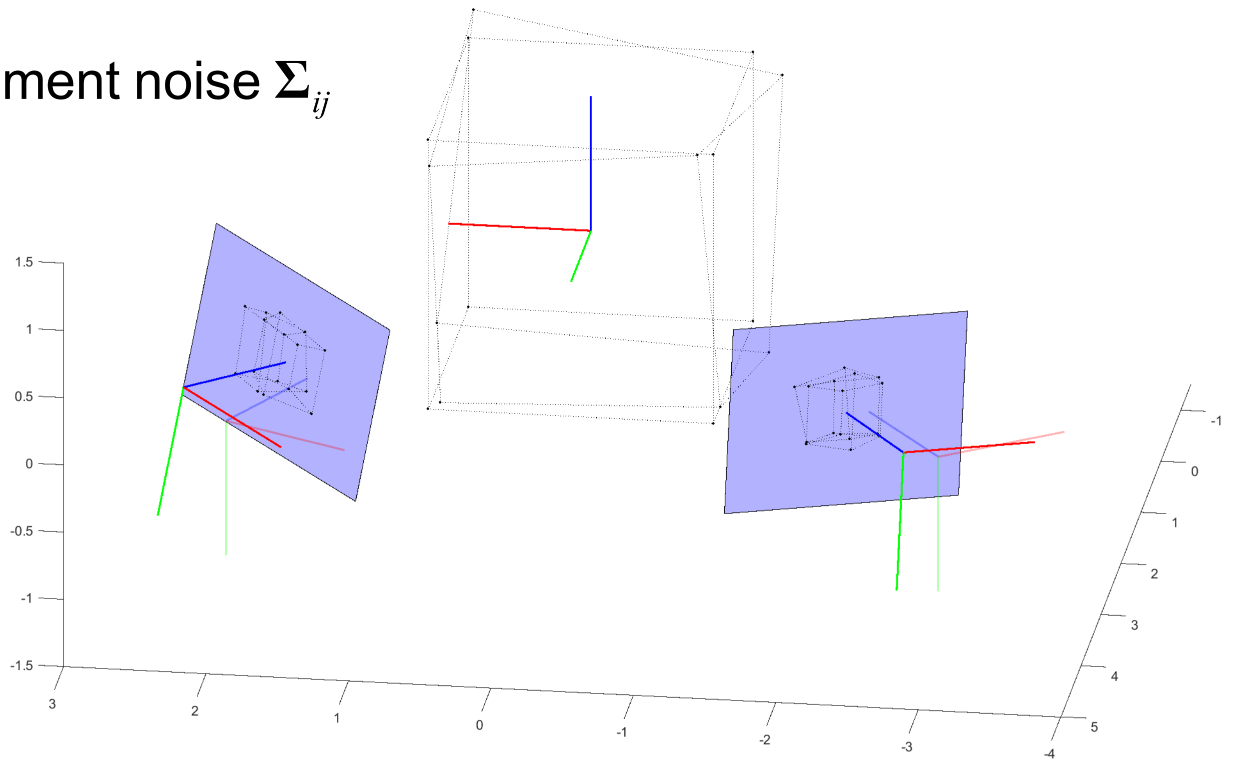
- Correspondences $\mathbf{u}_j^i \leftrightarrow \mathbf{x}_j^w$ with measurement noise $\mathbf{\Sigma}_{ij}$

States we wish to estimate:

- Camera poses $\mathbf{T}_{wc_i}$ *and* world points $\mathbf{x}_j^w$

Initial estimates:

- Pairwise two-view constraints (from the essential matrix)
- Triangulated points



Applying the MAP framework

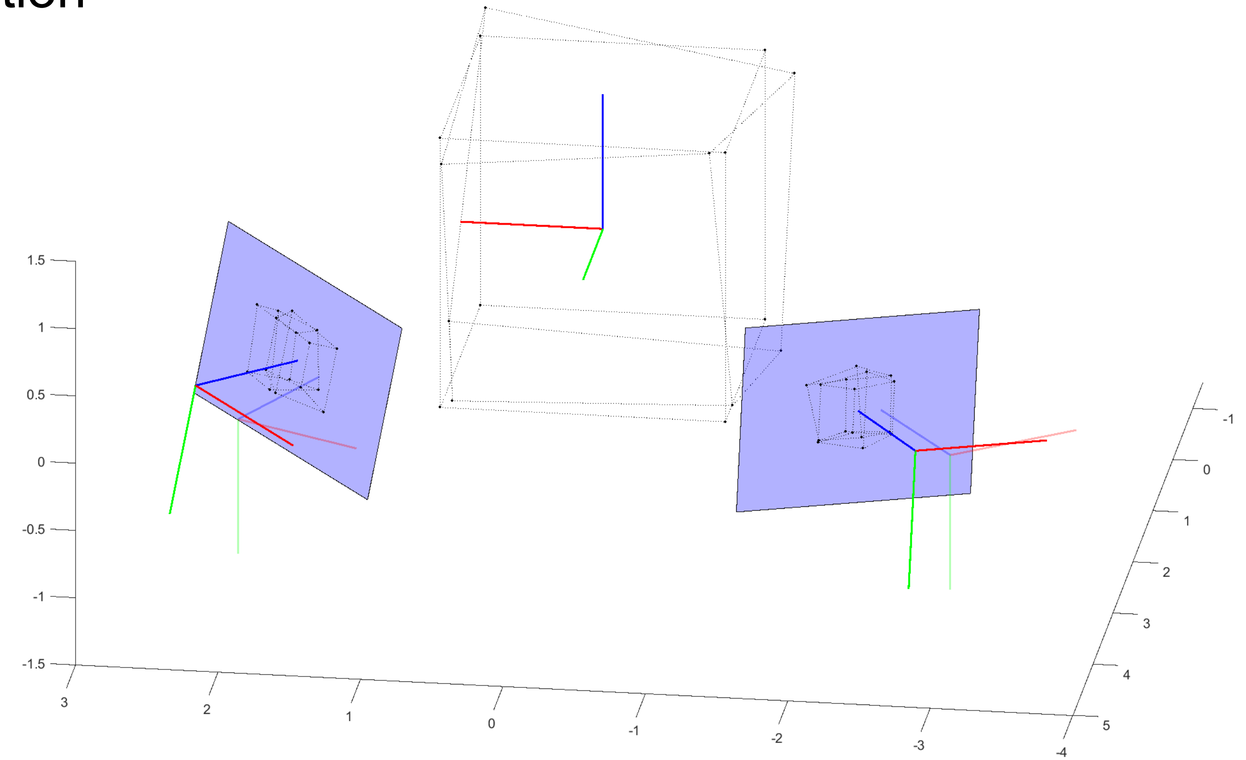
For simplicity,
we pre-calibrate to normalized image coordinates (and propagate the noise)

This gives us the measurement prediction function

$$h_{ij}(\mathbf{T}_{wc_i}, \mathbf{x}_j^w) = \pi_n(\mathbf{T}_{wc_i}^{-1} \cdot \mathbf{x}_j^w)$$

and measurement error function

$$e_{ij}(\mathbf{T}_{wc_i}, \mathbf{x}_j^w) = \pi_n(\mathbf{T}_{wc_i}^{-1} \cdot \mathbf{x}_j^w) - \mathbf{x}_{n\,j}^{i}$$



Applying the MAP framework

Since the measurement prediction function is a function of two variables, we linearize it at the current state estimates as

$$\begin{aligned} h_{ij}(\mathbf{T}_{wc_i}, \mathbf{x}_j^w) &= h_{ij}(\hat{\mathbf{T}}_{wc_i} \oplus \boldsymbol{\xi}_i, \hat{\mathbf{x}}_j^w + \delta \mathbf{x}_j) \\ &\approx h_{ij}(\hat{\mathbf{T}}_{wc_i}, \hat{\mathbf{x}}_j^w) + \mathbf{J}_{\hat{\mathbf{T}}_{wc_i}}^{h_{ij}} \boldsymbol{\xi}_i + \mathbf{J}_{\hat{\mathbf{x}}_j^w}^{h_{ij}} \delta \mathbf{x}_j \end{aligned}$$

These measurement Jacobians are given in earlier lectures on motion-only BA and structure-only BA.

Applying the MAP framework

This results in the linearized weighted least squares problem

$$\begin{aligned}\underline{\boldsymbol{\tau}}^* &= \arg\min_{\underline{\boldsymbol{\tau}}} \sum_{i=1}^{k}\sum_{j=1}^{n} \left\| \mathbf{P}_{ij}\boldsymbol{\xi}_i + \mathbf{S}_{ij}\delta\mathbf{x}_j - \mathbf{b}_{ij} \right\|^2 \\ &= \arg\min_{\underline{\boldsymbol{\tau}}} \left\| \mathbf{A}\underline{\boldsymbol{\tau}} - \mathbf{b} \right\|^2,\end{aligned}$$

where

$$\begin{aligned}\mathbf{P}_{ij} &= \boldsymbol{\Sigma}_{n\,ij}^{-1/2}\mathbf{J}_{\mathbf{T}_{wc_i}}^{h_{ij}} \\ \mathbf{S}_{ij} &= \boldsymbol{\Sigma}_{n\,ij}^{-1/2}\mathbf{J}_{\mathbf{x}_j^w}^{h_{ij}} \\ \mathbf{b}_{ij} &= \boldsymbol{\Sigma}_{n\,ij}^{-1/2}(\mathbf{x}_{n\,j}^{i} - h_{ij}(\mathbf{T}_{wc_i}, \mathbf{x}_j^w)),\end{aligned}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{P}_{11} & & & \mathbf{S}_{11} & & \\ \vdots & & & & \ddots & \\ \mathbf{P}_{1n} & & & & & \mathbf{S}_{1n} \\ & \ddots & & & \vdots & \\ & & \mathbf{P}_{k1} & \mathbf{S}_{k1} & & \\ & & \vdots & & \ddots & \\ & & \mathbf{P}_{kn} & & & \mathbf{S}_{kn} \end{bmatrix} \quad \underline{\boldsymbol{\tau}} = \begin{bmatrix} \boldsymbol{\xi}_1 \\ \vdots \\ \boldsymbol{\xi}_k \\ \delta\mathbf{x}_1 \\ \vdots \\ \delta\mathbf{x}_n \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} \mathbf{b}_{11} \\ \vdots \\ \mathbf{b}_{1n} \\ \vdots \\ \mathbf{b}_{k1} \\ \vdots \\ \mathbf{b}_{kn} \end{bmatrix}.$$

Linear least-squares

The **measurement Jacobian** $\mathbf{A}$ is a **block sparse matrix**.

For an example with *two cameras* and *three points* we have

$$\mathbf{A} = \begin{bmatrix} \mathbf{P}_{11} & & \mathbf{S}_{11} & & \\ \mathbf{P}_{12} & & & \mathbf{S}_{12} & \\ \mathbf{P}_{13} & & & & \mathbf{S}_{13} \\ & \mathbf{P}_{21} & \mathbf{S}_{21} & & \\ & \mathbf{P}_{22} & & \mathbf{S}_{22} & \\ & \mathbf{P}_{23} & & & \mathbf{S}_{23} \end{bmatrix} \qquad \underline{\boldsymbol{\tau}} = \begin{bmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \\ \delta\mathbf{x}_1 \\ \delta\mathbf{x}_2 \\ \delta\mathbf{x}_3 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} \mathbf{b}_{11} \\ \mathbf{b}_{12} \\ \mathbf{b}_{13} \\ \mathbf{b}_{21} \\ \mathbf{b}_{22} \\ \mathbf{b}_{23} \end{bmatrix}$$

Applying the MAP framework

The solution can be found by solving the normal equations

$$\left(\mathbf{A}^T\mathbf{A}\right)\underline{\boldsymbol{\tau}}^* = \mathbf{A}^T\mathbf{b}$$

Since $\mathbf{A}$ is sparse,
a sparse solver should be used.

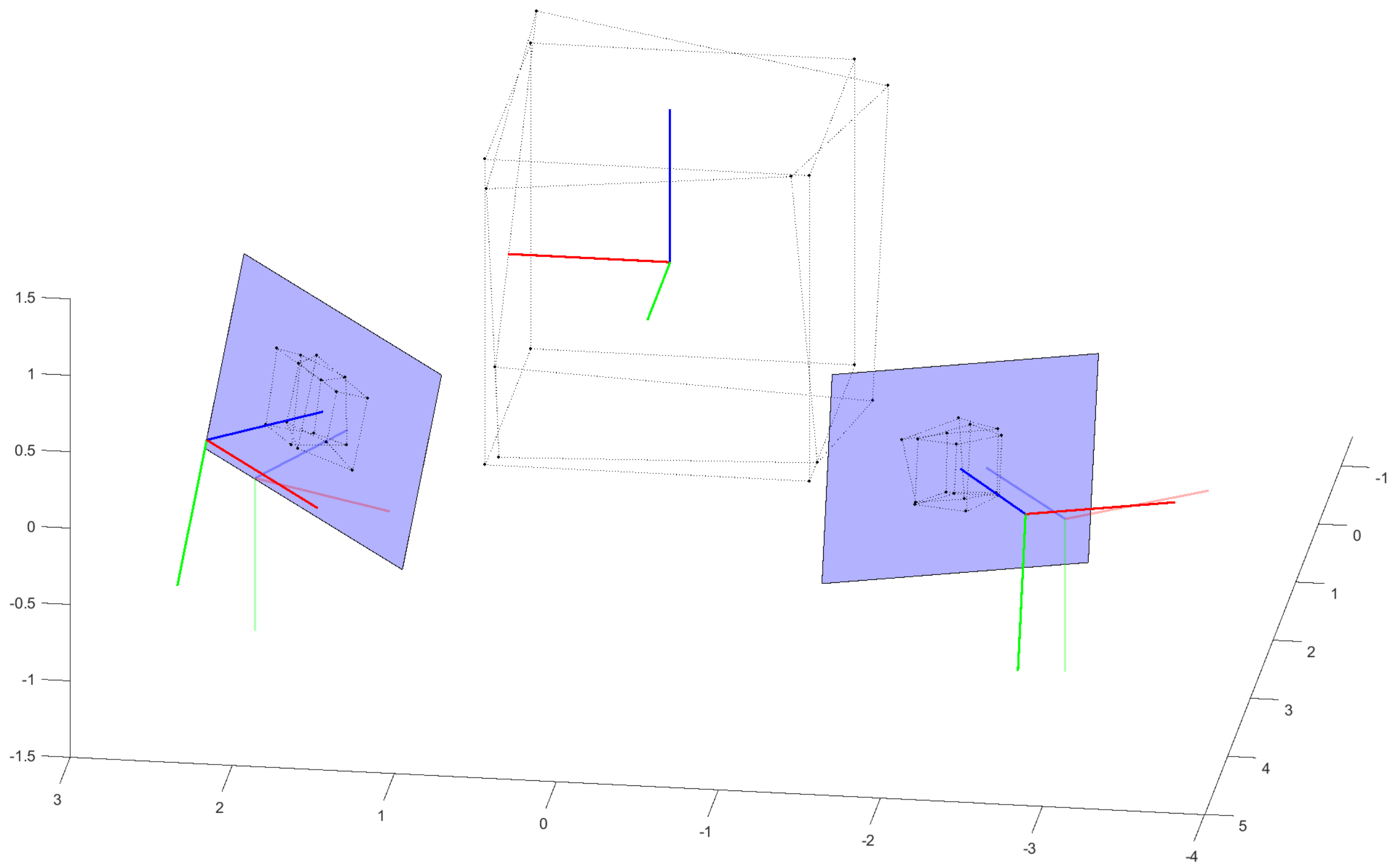
Choose a suitable inital estimate $\hat{\underline{\mathcal{X}}}^0$

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

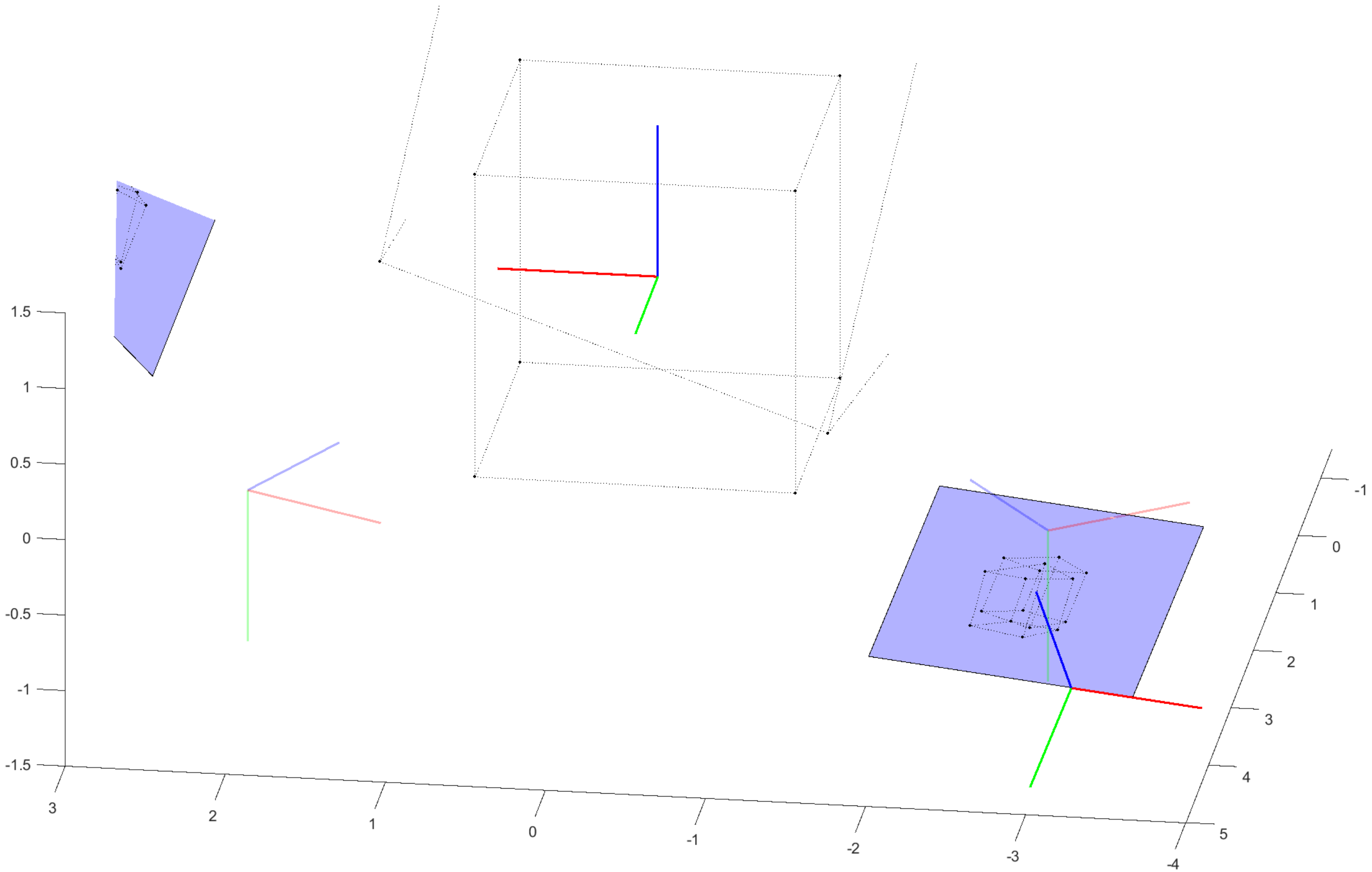
$$\underline{\boldsymbol{\tau}}^* \leftarrow \textit{Solve } \underset{\underline{\boldsymbol{\tau}}}{\operatorname{argmin}} \left\|\mathbf{A}\underline{\boldsymbol{\tau}} - \mathbf{b}\right\|^2$$

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

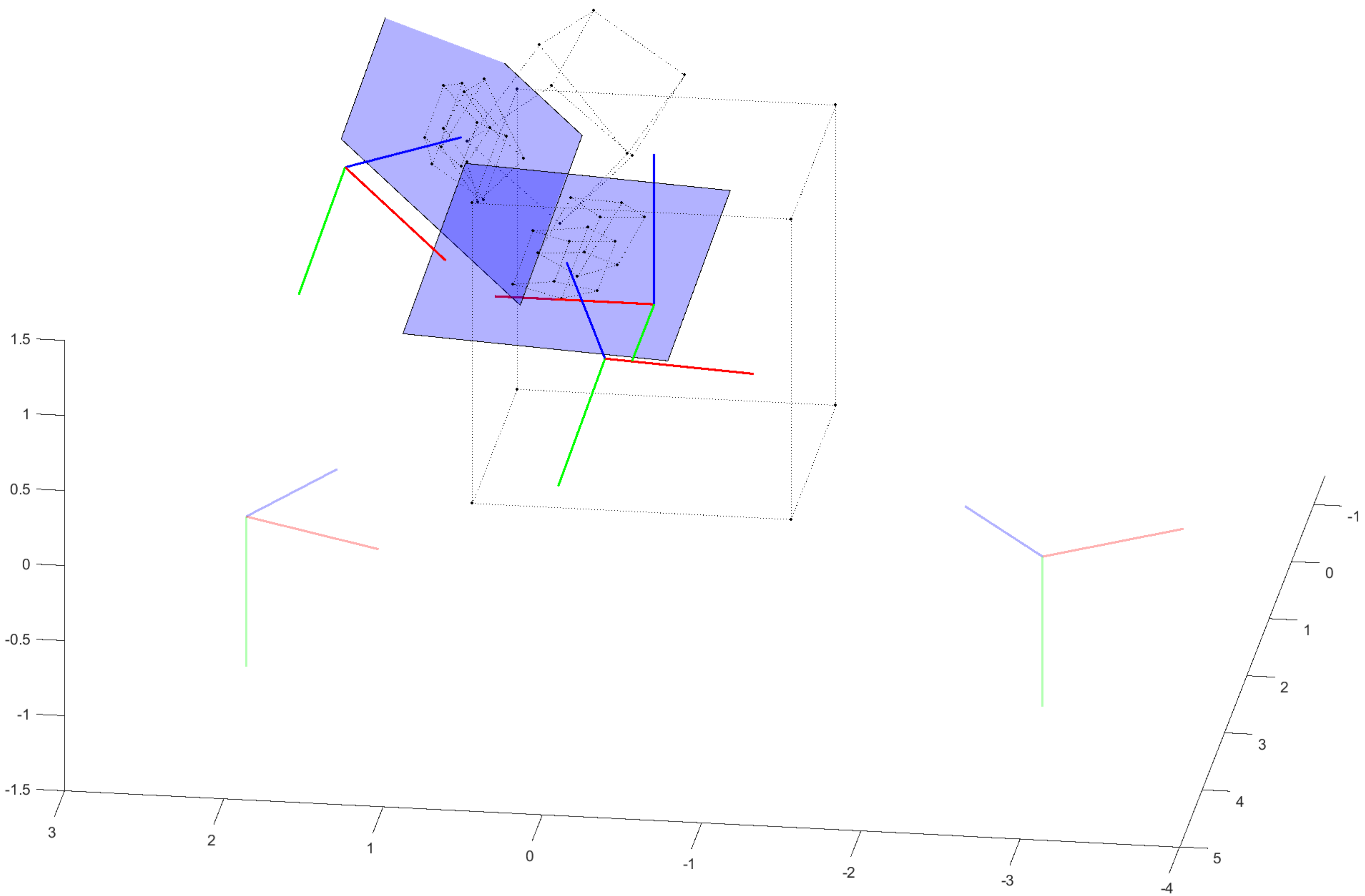
Example



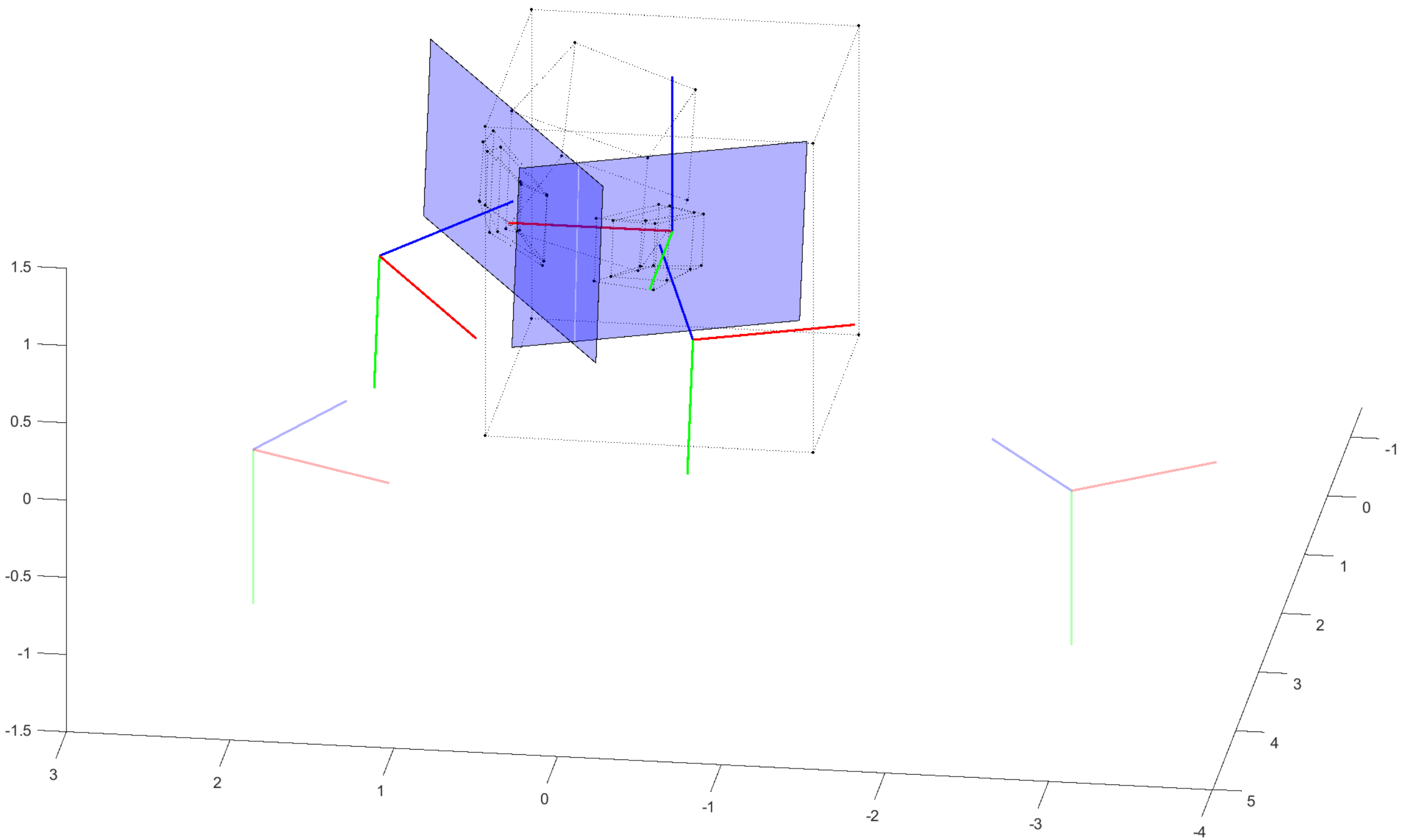
Example



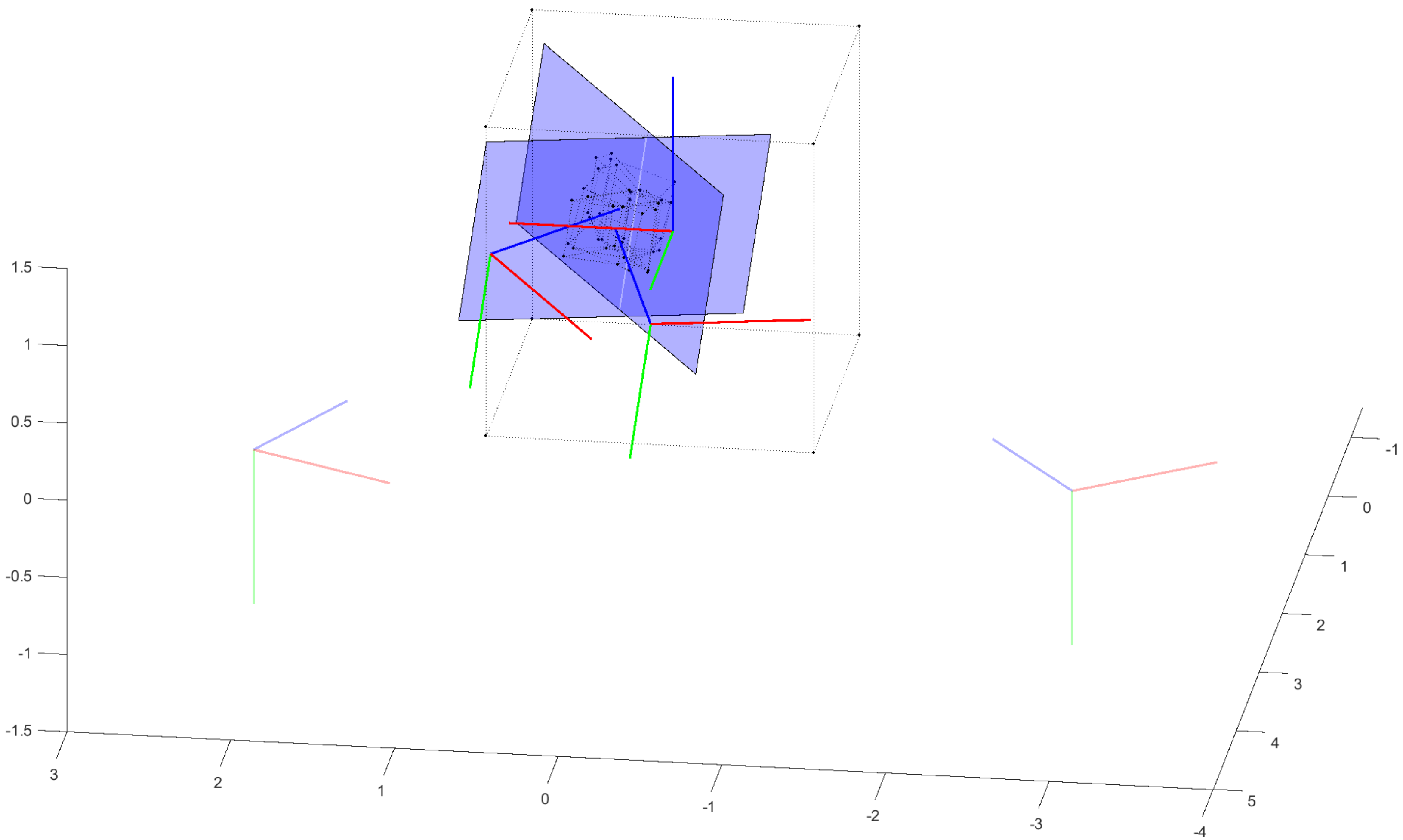
Example



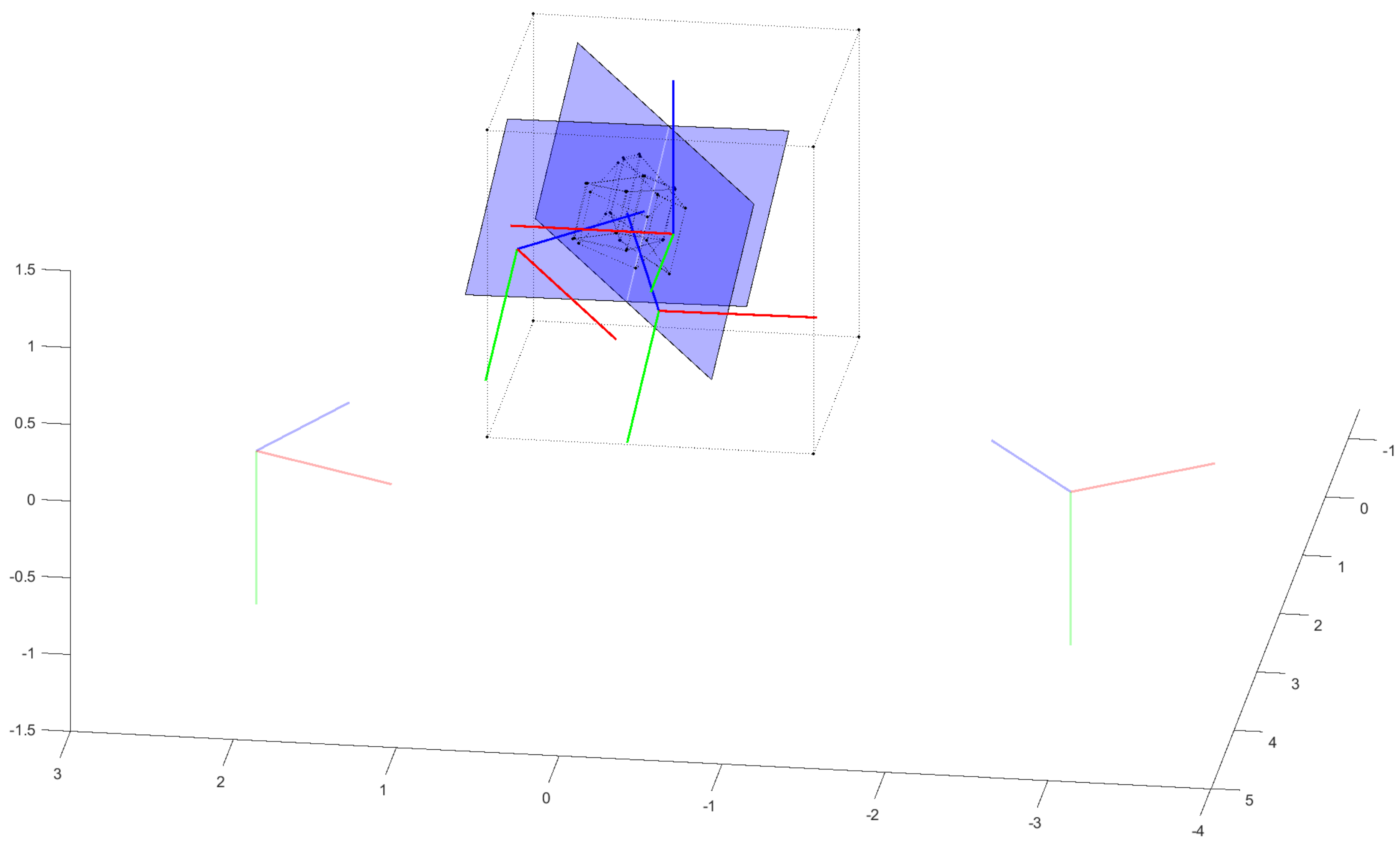
Example



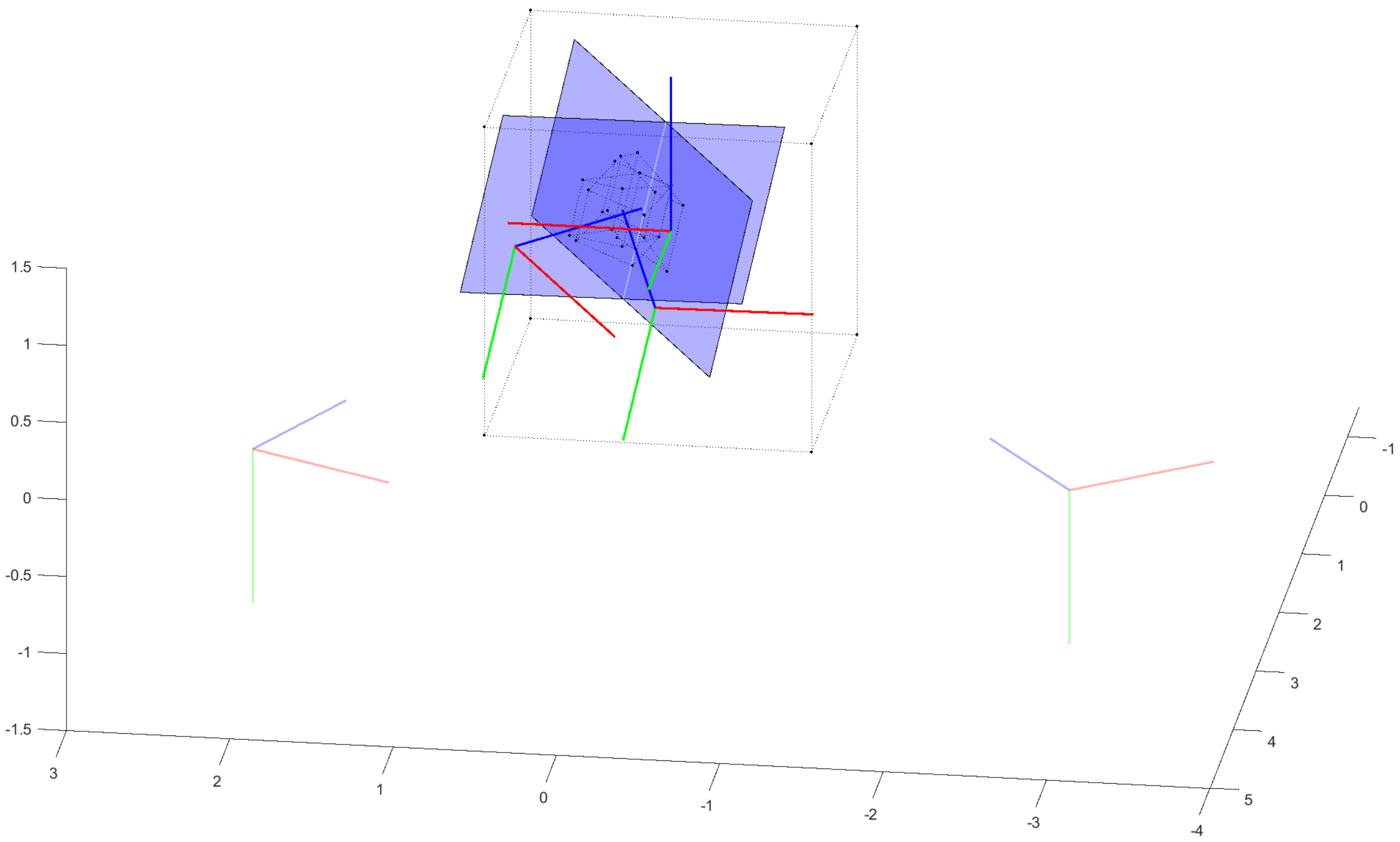
Example



Example

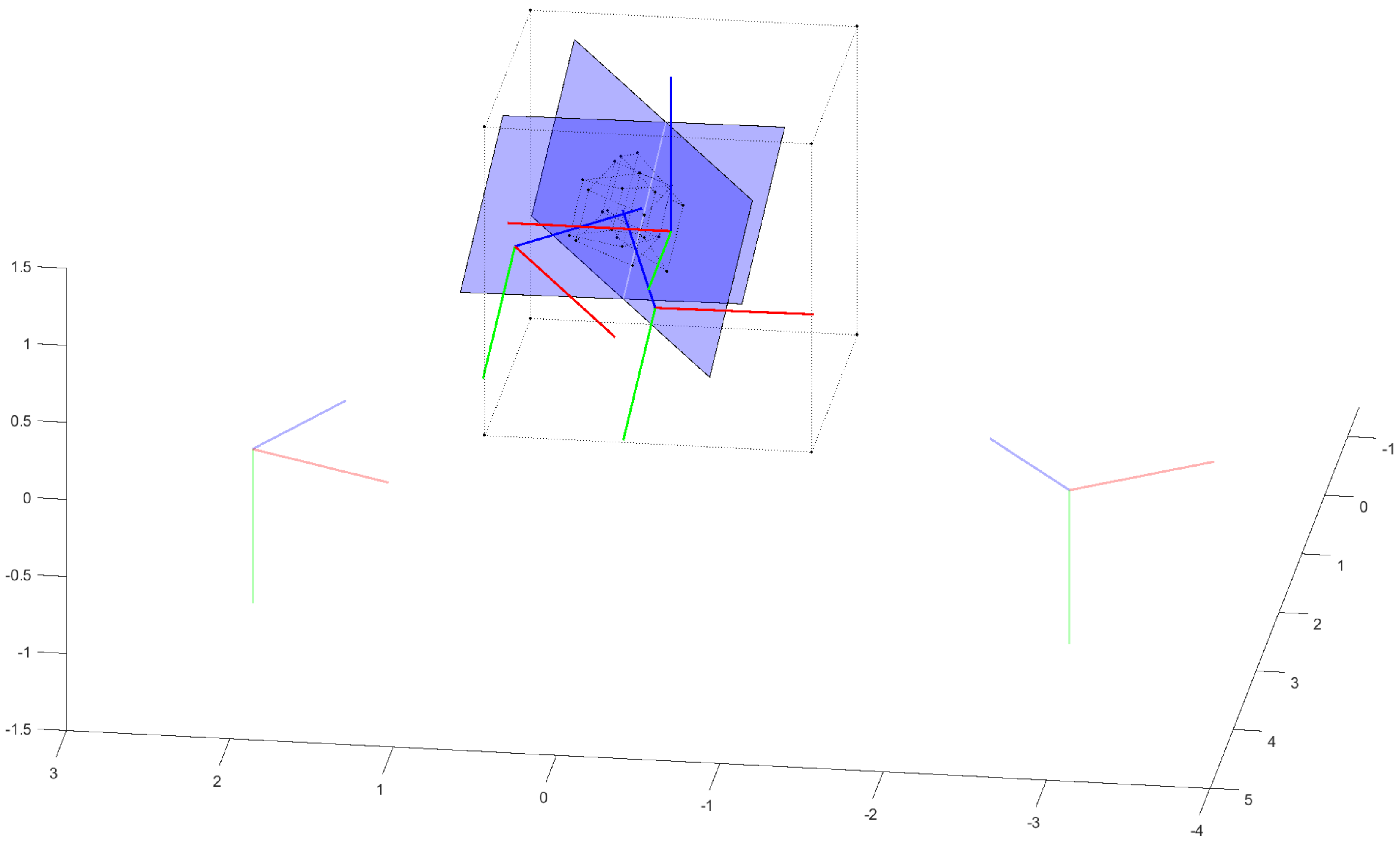


Example



Example

Why does this fail?



Gauge freedom

The solution is not uniquely determined!

- The Hessian is singular!
- We can apply any 7DOF similarity transform to the cameras without affecting the objective function

Gauge freedom

The solution is not uniquely determined!

- The Hessian is singular!
- We can apply any 7DOF similarity transform to the cameras without affecting the objective function

Possible solutions:

- ~~Use **Levenberg-Marquardt** optimization~~
- Add priors on poses and points
- Fuse with other information, such as GPS and IMU

Adding priors

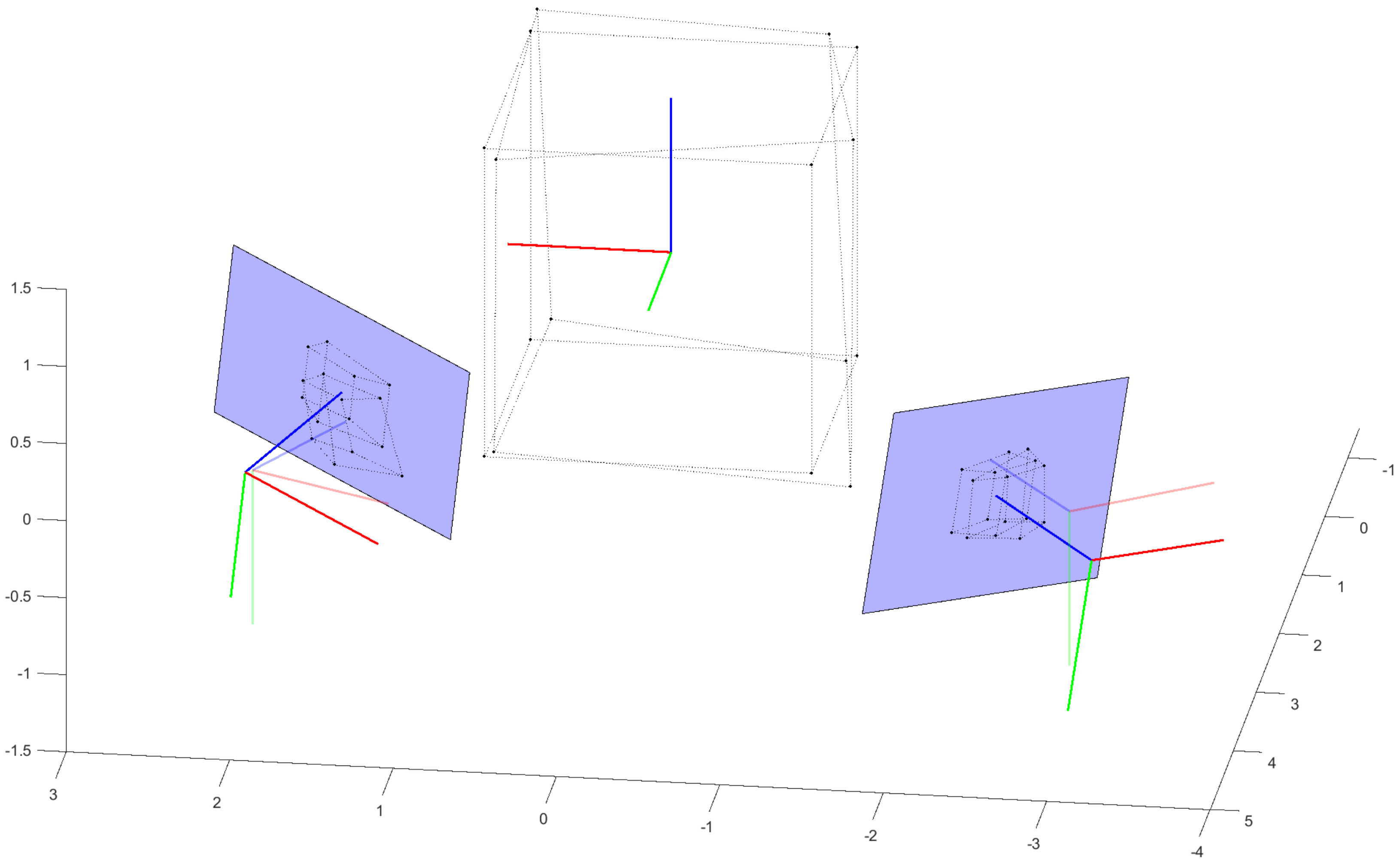
Prior on first pose and first point

$$
\mathbf{A} = \begin{bmatrix}
\mathbf{P}_{11} & & \mathbf{S}_{11} & & \\
\mathbf{P}_{12} & & & \mathbf{S}_{12} & \\
\mathbf{P}_{13} & & & & \mathbf{S}_{13} \\
& \mathbf{P}_{21} & \mathbf{S}_{21} & & \\
& \mathbf{P}_{22} & & \mathbf{S}_{22} & \\
& \mathbf{P}_{23} & & & \mathbf{S}_{23} \\
\mathbf{I}_{6\times 6} & & & & \\
& & \mathbf{I}_{3\times 3} & &
\end{bmatrix}
\qquad
\underline{\boldsymbol{\tau}} = \begin{bmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \\ \delta\mathbf{x}_1 \\ \delta\mathbf{x}_2 \\ \delta\mathbf{x}_3 \end{bmatrix}
\qquad
\mathbf{b} = \begin{bmatrix} \mathbf{b}_{11} \\ \mathbf{b}_{12} \\ \mathbf{b}_{13} \\ \mathbf{b}_{21} \\ \mathbf{b}_{22} \\ \mathbf{b}_{23} \\ \mathbf{b}_{\boldsymbol{\xi}_1^{prior}} \\ \mathbf{b}_{\delta\mathbf{x}_1^{prior}} \end{bmatrix}
$$

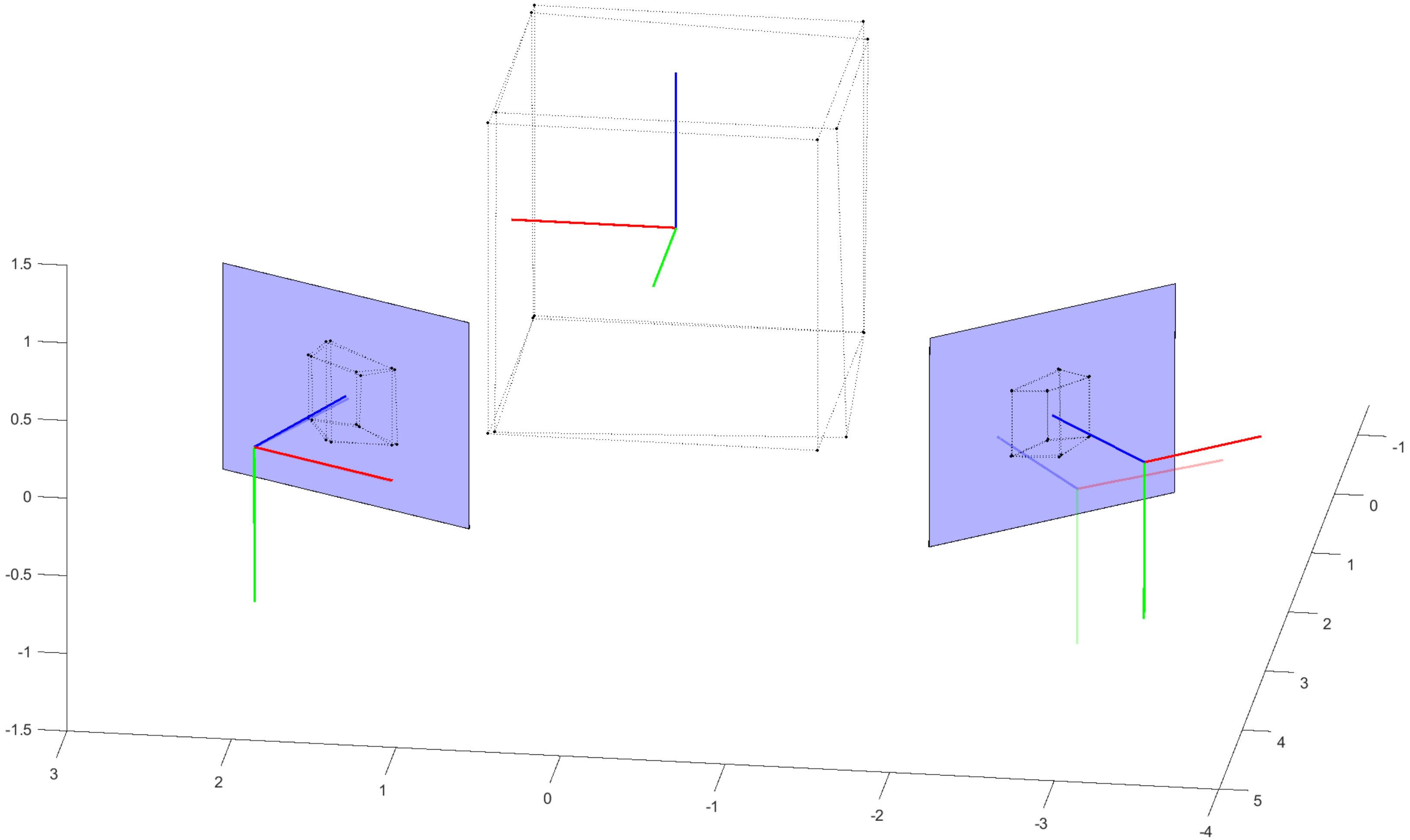
$$
\mathbf{b}_{\boldsymbol{\xi}_1^{prior}} = -(\mathbf{T}_{wc_1} \ominus \mathbf{T}_{wc_1}^{prior})
$$

$$
\mathbf{b}_{\delta\mathbf{x}_1^{prior}} = -(\mathbf{x}_1^{w} - \mathbf{x}_1^{w,prior})
$$

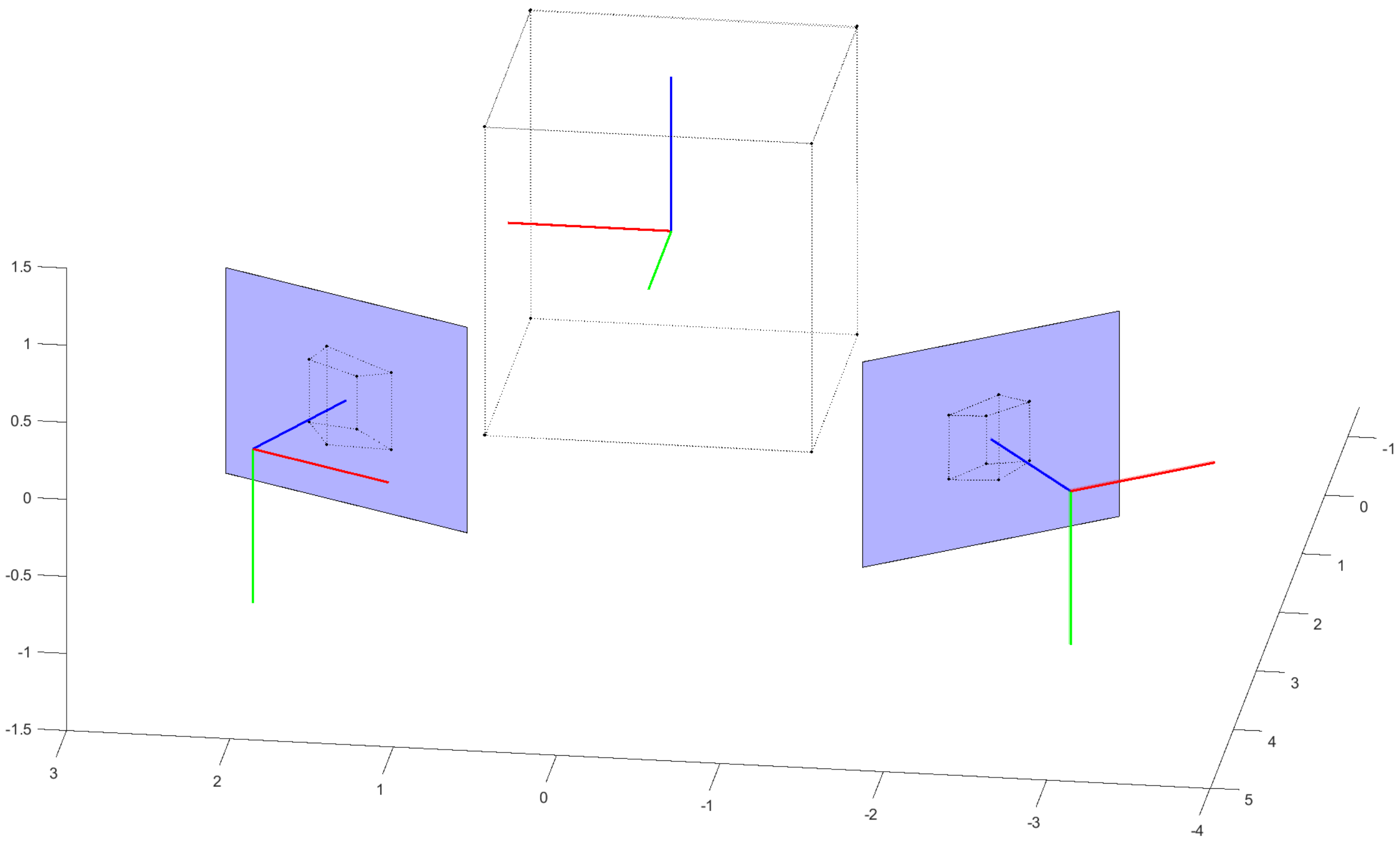
Example



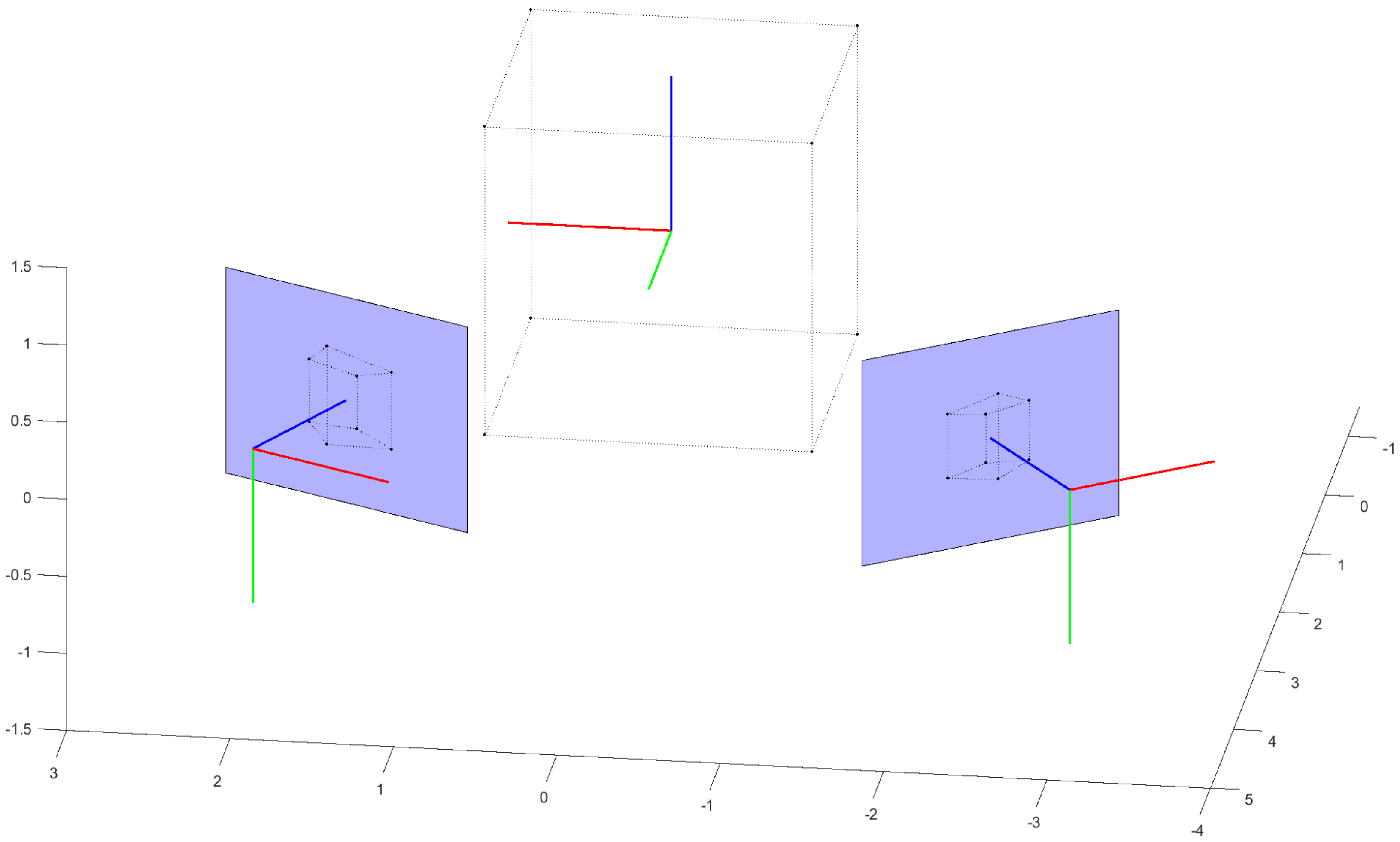
Example



Example



Example



Adding priors

Prior on first pose and distance between first two points

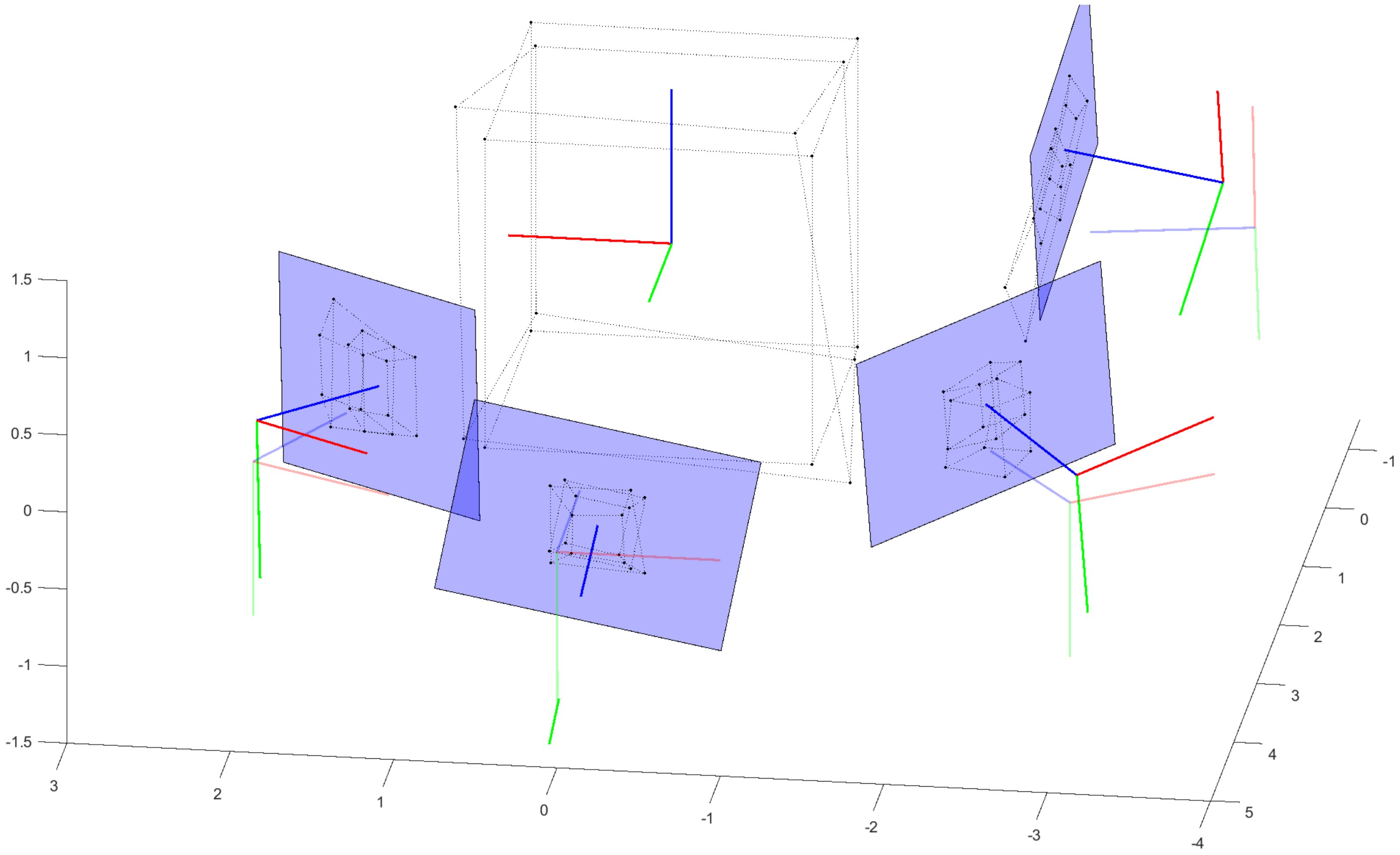
$$\mathbf{A} = \begin{bmatrix} \mathbf{P}_{11} & & \mathbf{S}_{11} & & \\ \mathbf{P}_{12} & & & \mathbf{S}_{12} & \\ \mathbf{P}_{13} & & & & \mathbf{S}_{13} \\ & \mathbf{P}_{21} & \mathbf{S}_{21} & & \\ & \mathbf{P}_{22} & & \mathbf{S}_{22} & \\ & \mathbf{P}_{23} & & & \mathbf{S}_{23} \\ \mathbf{I}_{6\times 6} & & & & \\ & & \mathbf{D}_{12} & -\mathbf{D}_{12} & \end{bmatrix} \qquad \underline{\boldsymbol{\tau}} = \begin{bmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \\ \delta\mathbf{x}_1 \\ \delta\mathbf{x}_2 \\ \delta\mathbf{x}_3 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} \mathbf{b}_{11} \\ \mathbf{b}_{12} \\ \mathbf{b}_{13} \\ \mathbf{b}_{21} \\ \mathbf{b}_{22} \\ \mathbf{b}_{23} \\ \mathbf{b}_{\xi_1^{prior}} \\ \mathbf{b}_{d_{12}^{prior}} \end{bmatrix}$$

$$\mathbf{b}_{\xi_1^{prior}} = -(\mathbf{T}_{wc_1} \ominus \mathbf{T}_{wc_1}^{prior})$$

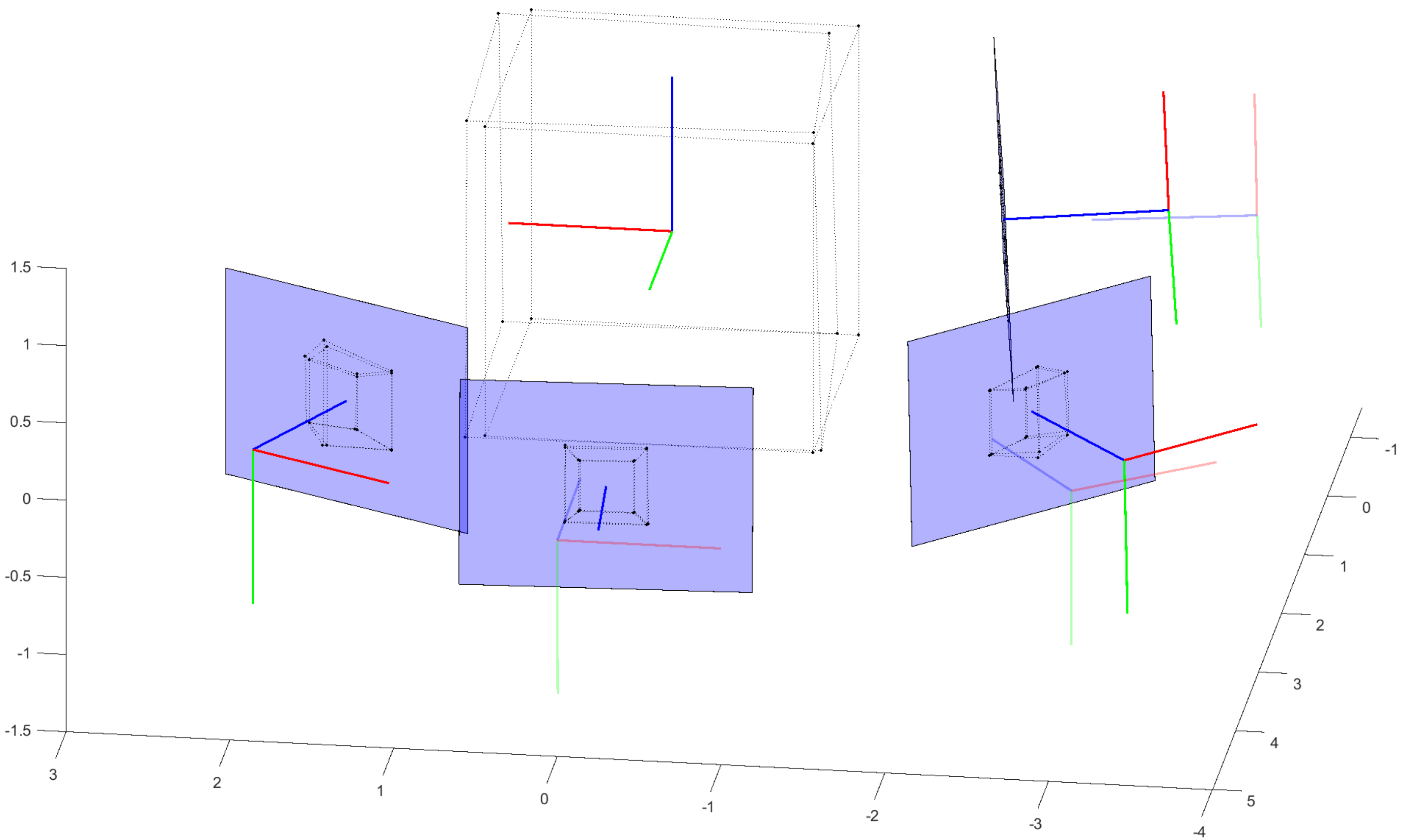
$$\mathbf{b}_{d_{12}^{prior}} = -\left(\left\|\mathbf{x}_2^w - \mathbf{x}_1^w\right\| - d_{12}^{prior}\right)$$

Try to compute the Jacobian $\mathbf{D}$!

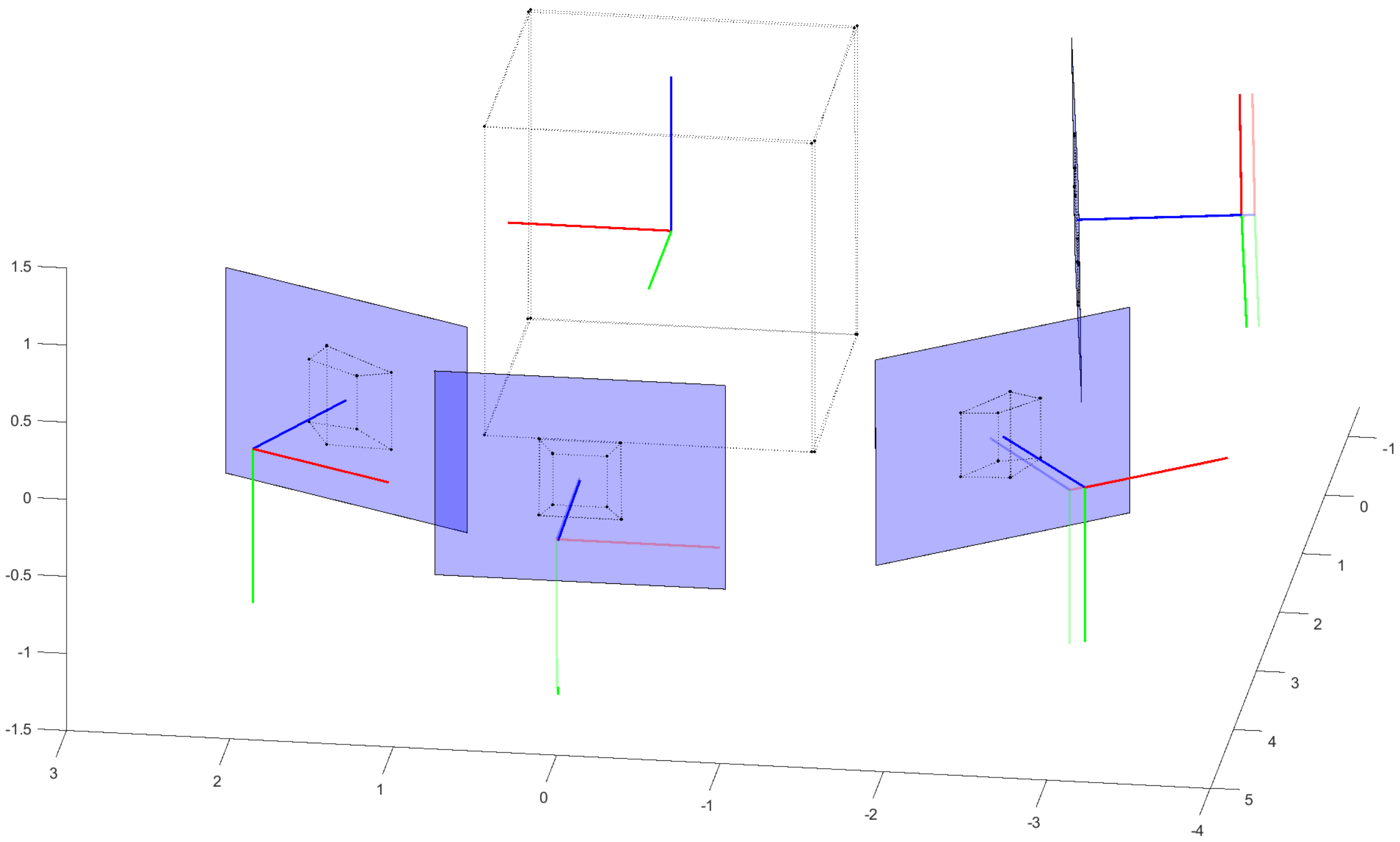
Example



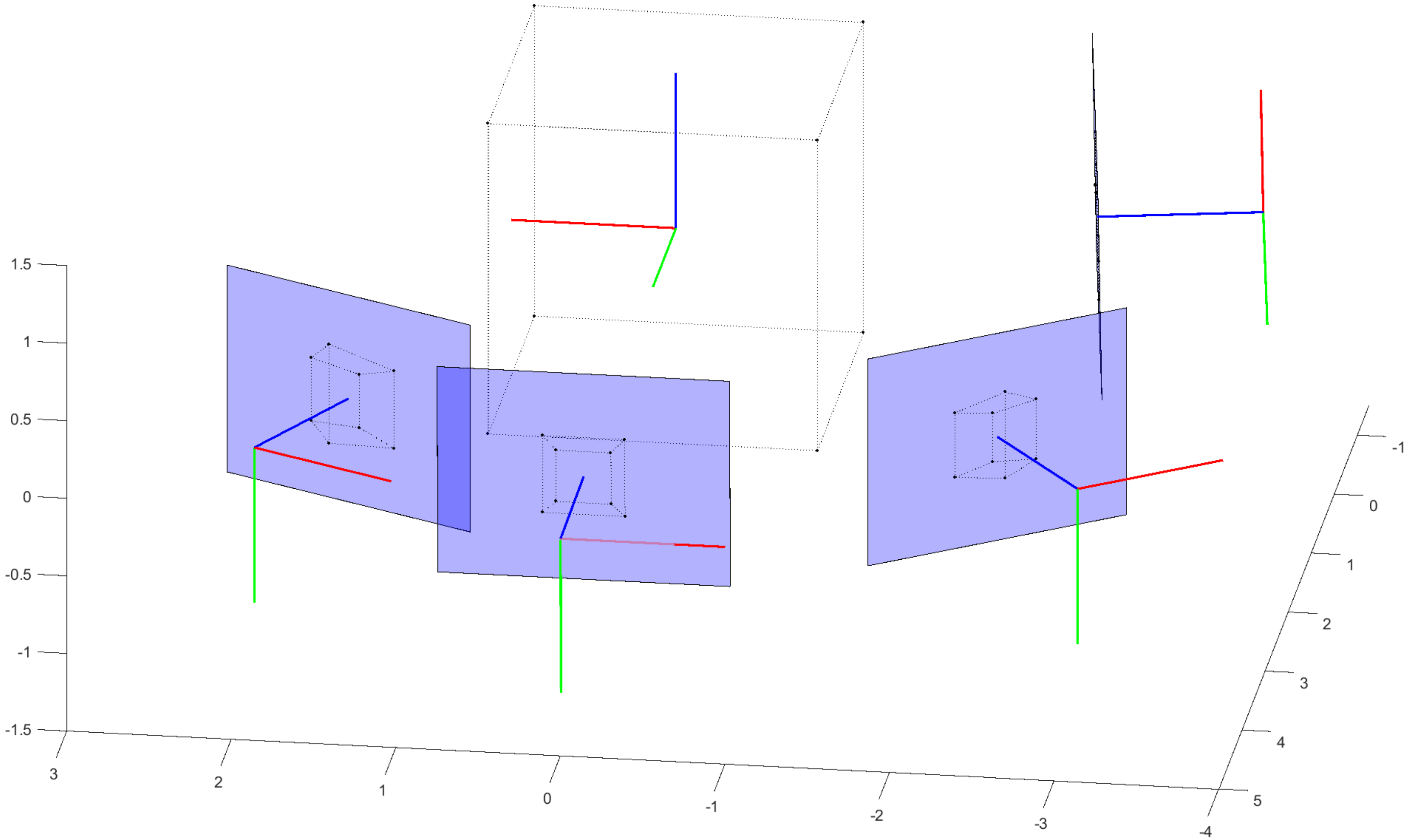
Example



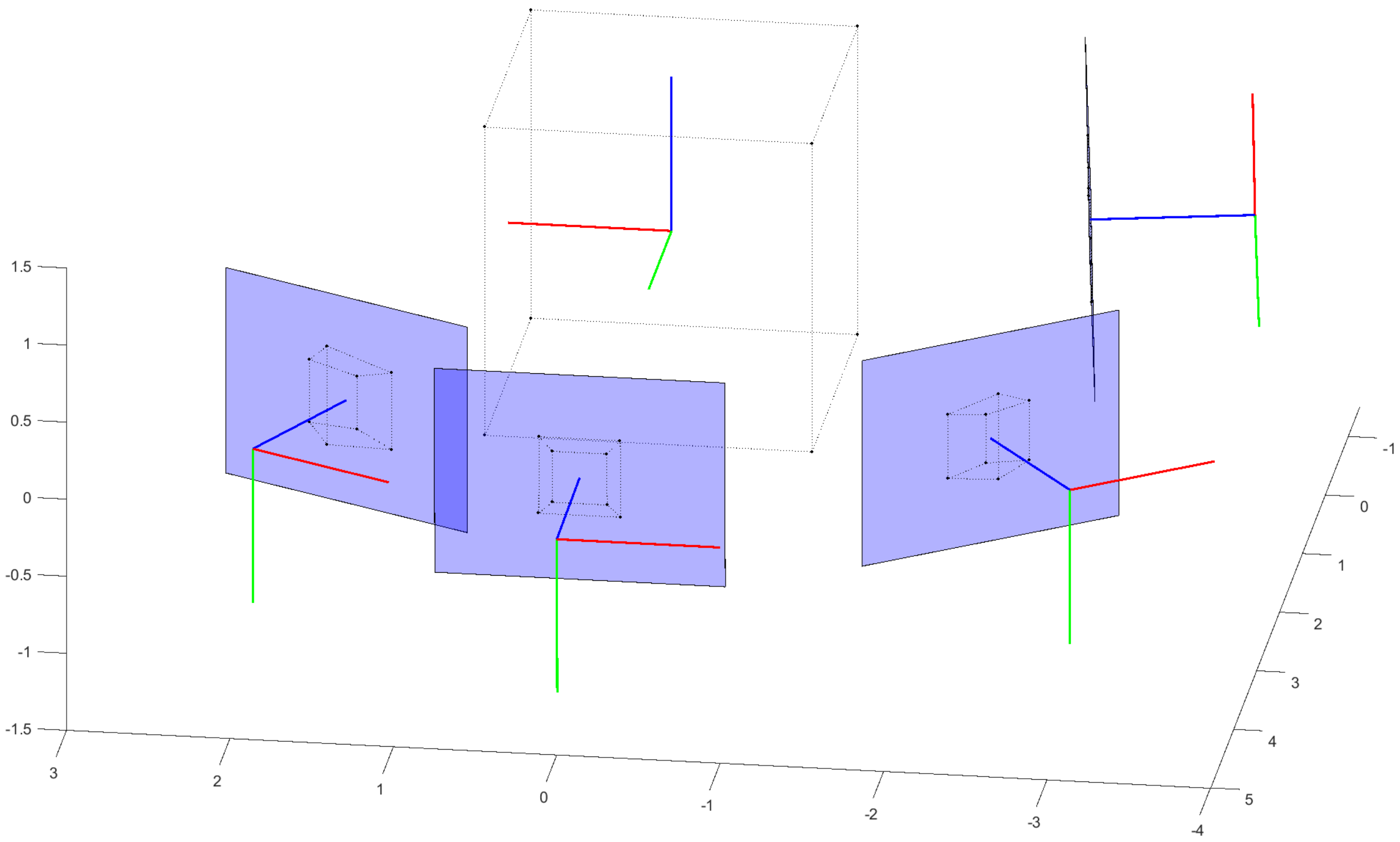
Example



Example



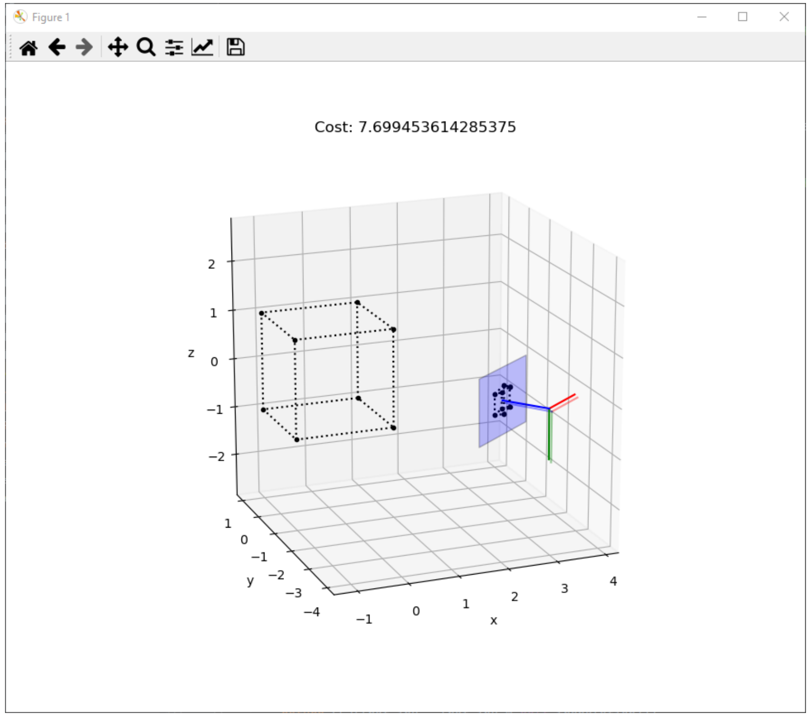
Example



Supplementary material

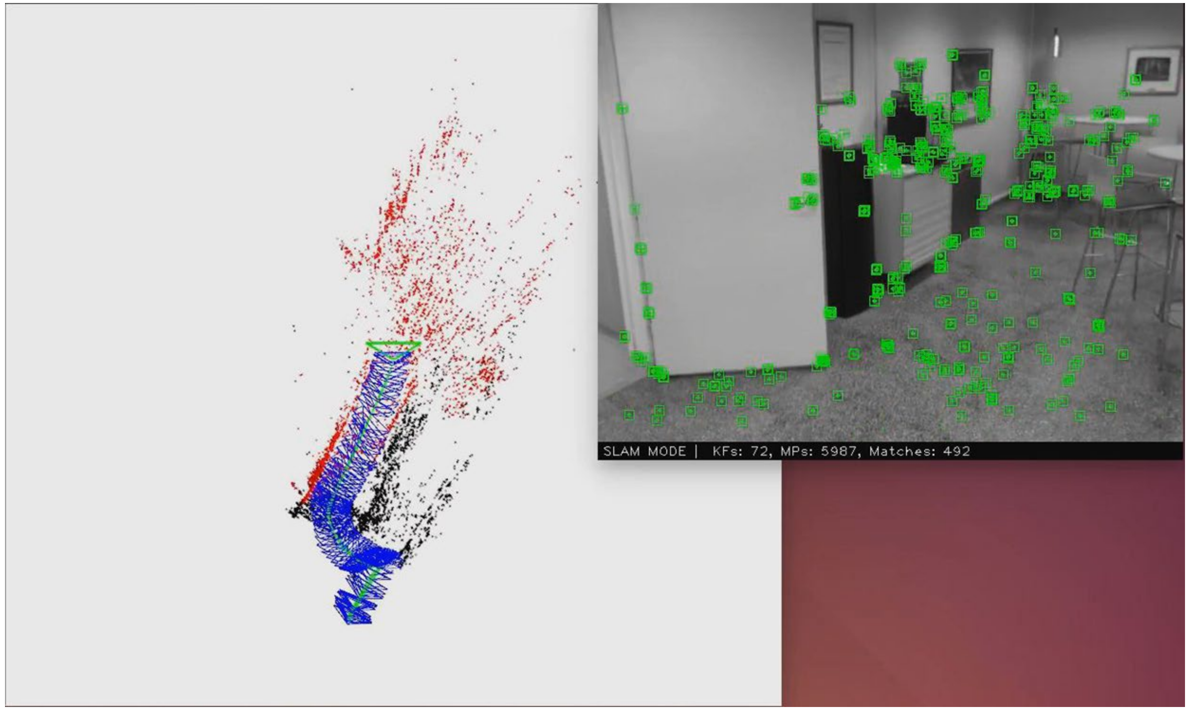
- The compendium!
  - Let me know if you would like to go through the derivations in greater detail!

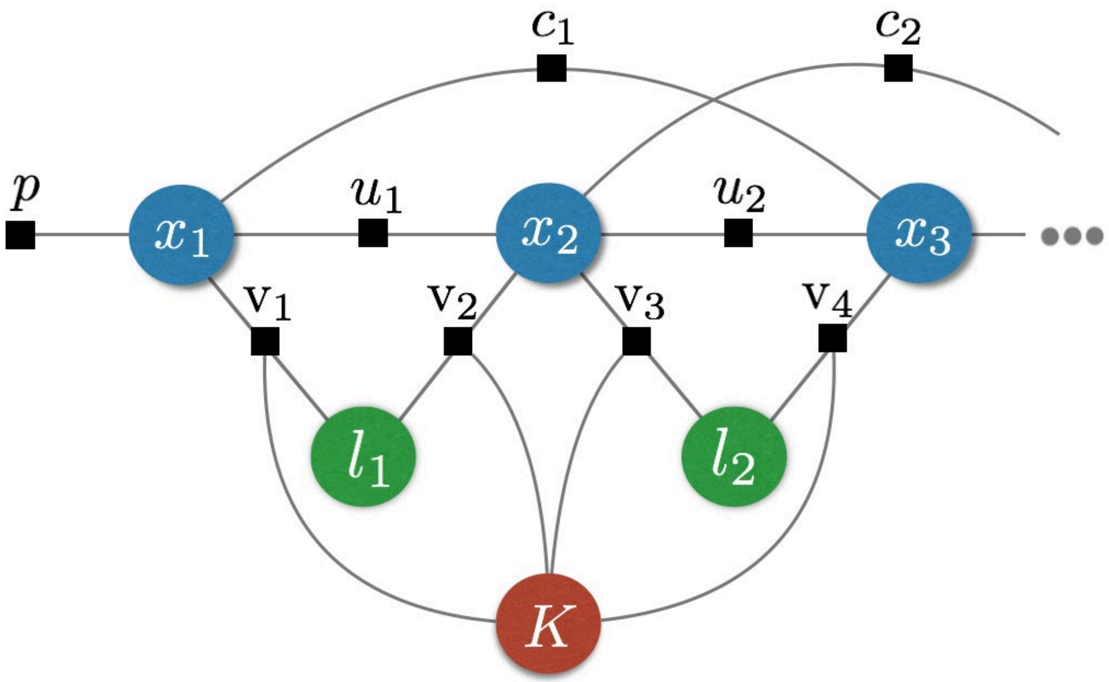
- Python implementation of the bundle adjustment examples:
  - https://github.com/ttk21/lab_05



Next lecture: Multiple-view stereo (for 3D reconstruction)

Next week: Visual SLAM





Cadena, C., et al. (2016). Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age. *IEEE Transactions on Robotics*, *32*(6), 1309–1332