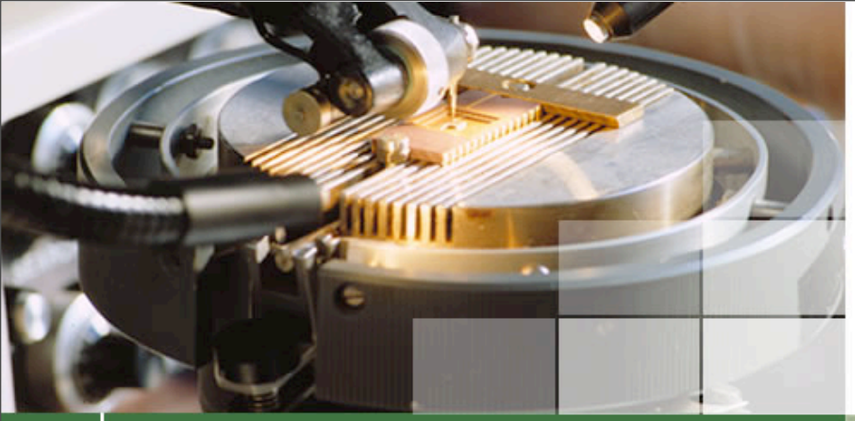


On the Use of Inverse Scaling in Monocular SLAM

Daniele Marzorati[1], Matteo Matteucci[2], Davide Migliore[2], *Domenico G. Sorrenti*[1]

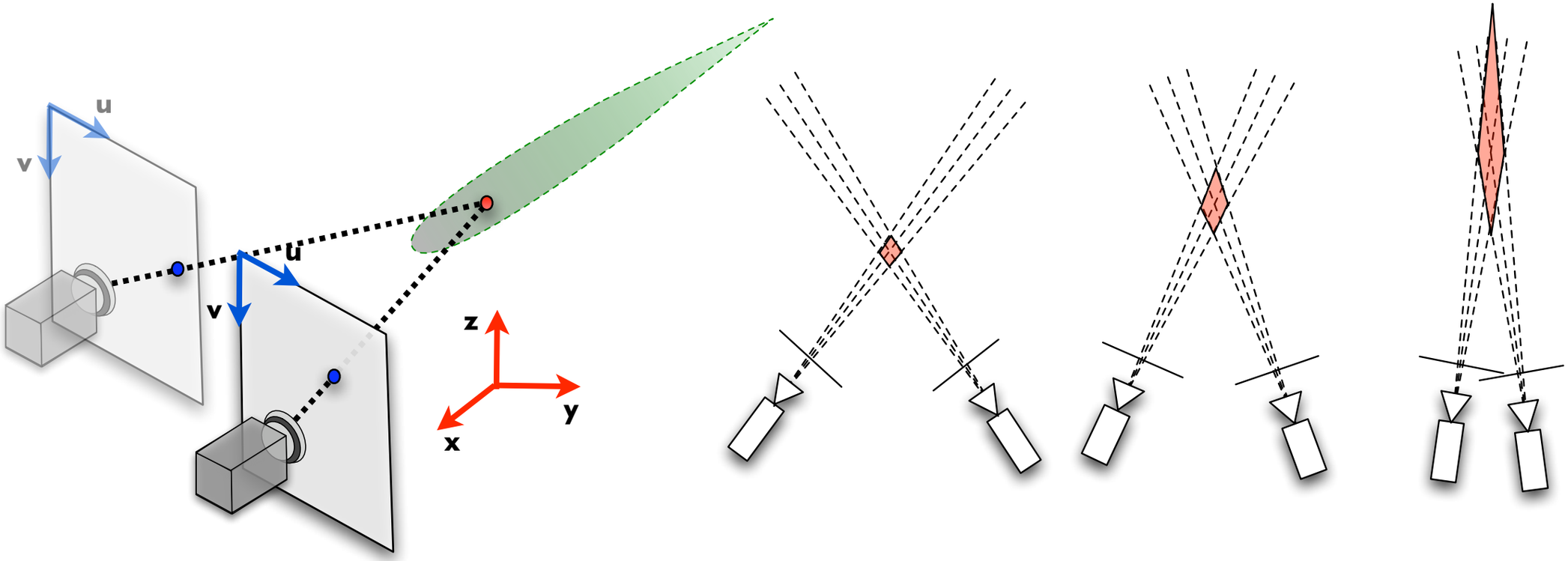
[1] Università degli Studi di Milano - Bicocca

[2] Politecnico di Milano

SLAM using Single Camera

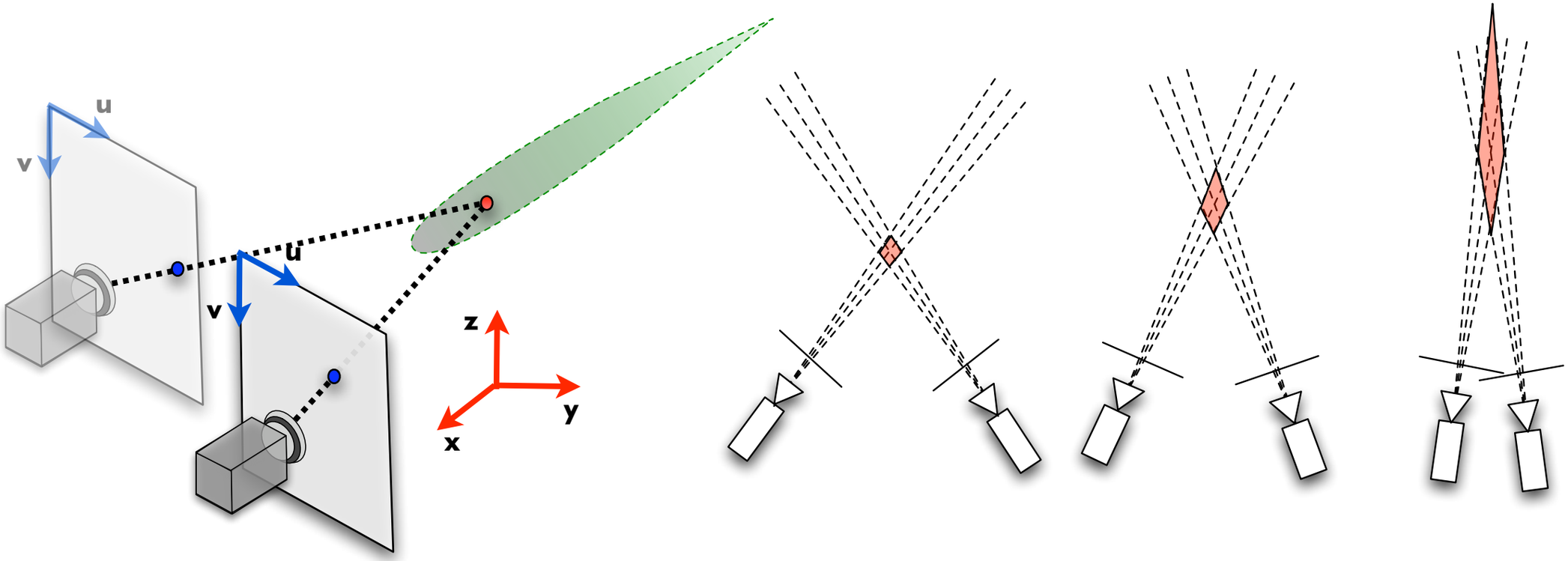
SLAM using Single Camera

- Why is this challenging?



SLAM using Single Camera

- Why is this challenging?

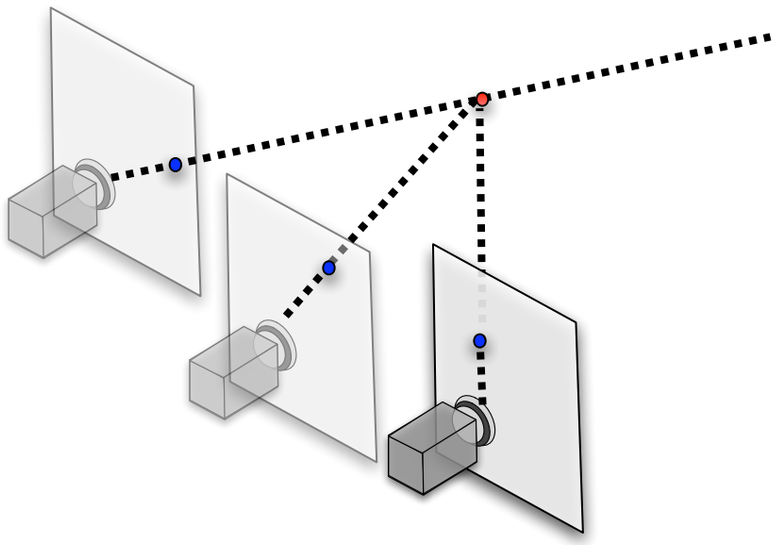


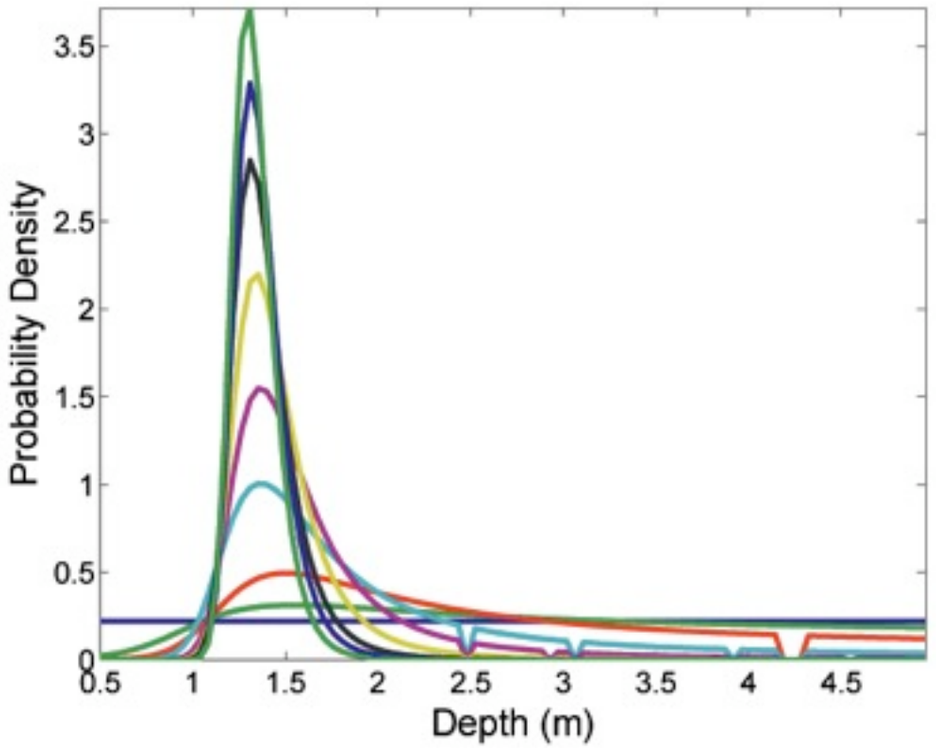
- Solutions:
  - (offline) - S. Soatto et al “Structure from motion casually integrated over time” - IEEE PAMI 2002
  - (online) - A. Davison “Real-time Simultaneous Localization And Mapping with a Single Camera” - ICCV 2003, PAMI 2007

SLAM using Single Camera

SLAM using Single Camera

- Particle Filter approach (Delayed)





- EKF based SLAM

$$\mathbf{x}_k = \begin{pmatrix} \mathbf{x}_c \\ \mathbf{m}_1 \\ \mathbf{m}_2 \\ \vdots \end{pmatrix}, \mathbf{P} = \begin{pmatrix} \mathbf{P}_{xx} & \mathbf{P}_{xm_1} & \mathbf{P}_{xm_2} & \cdots \\ \mathbf{P}_{m_1x} & \mathbf{P}_{m_1m_1} & \mathbf{P}_{m_1m_2} & \cdots \\ \mathbf{P}_{m_2x} & \mathbf{P}_{m_2m_1} & \mathbf{P}_{m_2m_2} & \cdots \\ \cdots & \cdots & \cdots & \ddots \end{pmatrix}.$$

$$\mathbf{x}_c = \begin{pmatrix} \mathbf{r}^{WC} \\ \mathbf{q}^{WC} \\ \mathbf{v}^W \\ \boldsymbol{\omega}^C \end{pmatrix} = \begin{pmatrix} \mathbf{r}^{WC} + (\mathbf{v}^W + \mathbf{V}^W)\Delta t \\ \mathbf{q}^{WC} \times \mathbf{q}((\boldsymbol{\omega}^C + \boldsymbol{\omega}^C))\Delta t \\ \mathbf{v}^W + \mathbf{V}^W \\ \boldsymbol{\omega}^C + \boldsymbol{\Omega}^C \end{pmatrix},$$

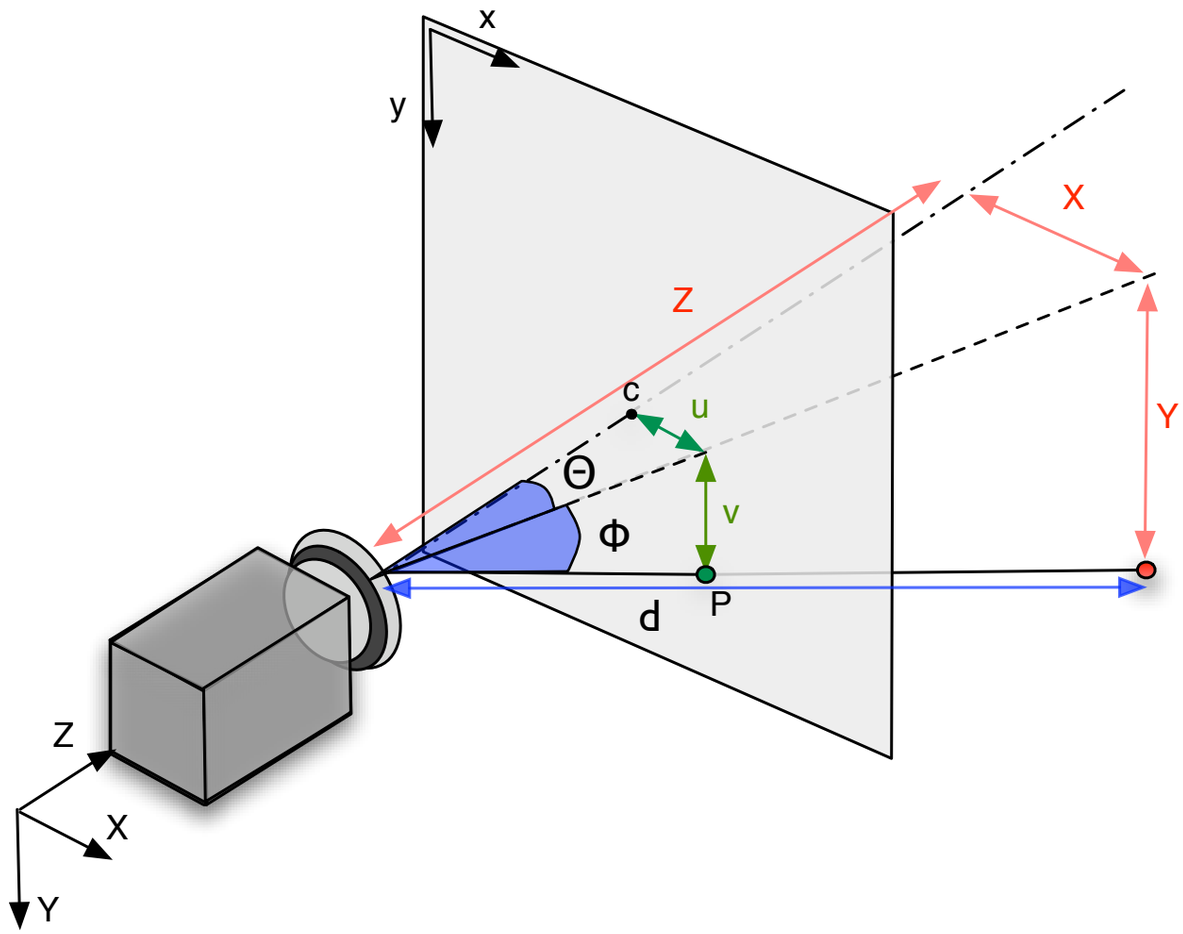
$$\mathbf{h}_n^C = \mathbf{R}^{CW}(\mathbf{m}_n^W - \mathbf{r}^{WC}),$$

- **3** Parameters for Each Point
- A. Davison et al. "*MonoSLAM: Real-Time Single Camera SLAM*" - PAMI 2007

SLAM using Single Camera

SLAM using Single Camera

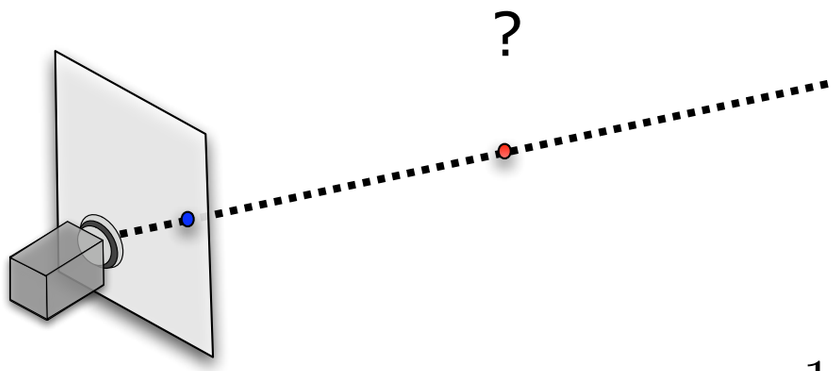
- Unified Inverse Depth Parametrization (Undelayed)

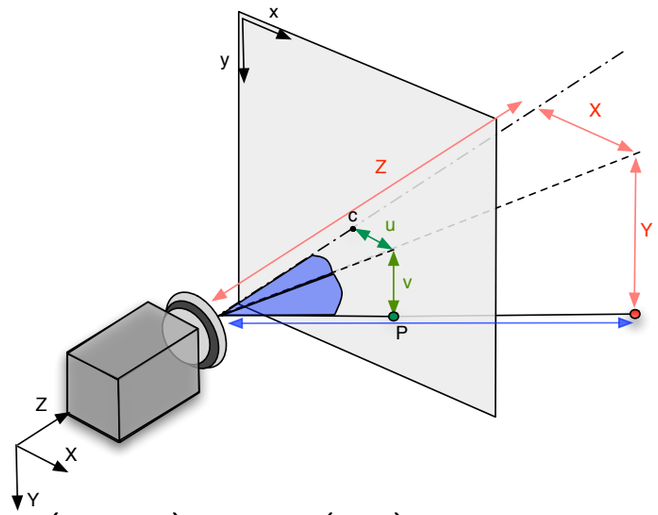


- J. Civera et al. "*Inverse Depth Parametrization for Monocular SLAM*" - TRO 2008

SLAM using Single Camera

- Unified Inverse Depth Parametrization (Undelayed)





$$d_n = \frac{1}{\rho_n}$$

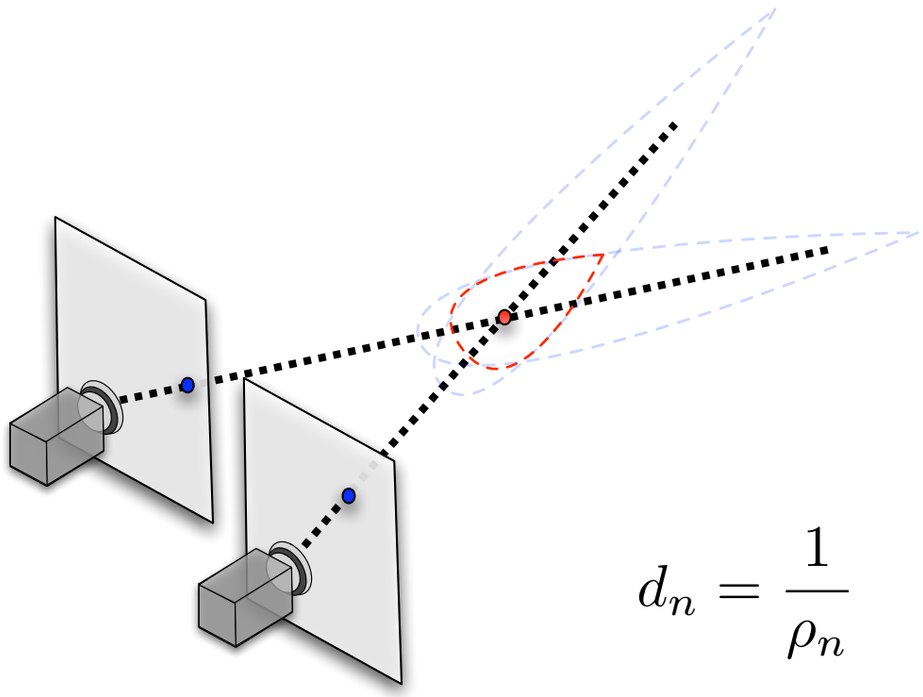
$$\mathbf{m}_n = \begin{pmatrix} m_{x,n} \\ m_{y,n} \\ m_{z,n} \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} + \frac{1}{\rho_n}\mathbf{\Gamma}(\theta_n, \phi_n),$$

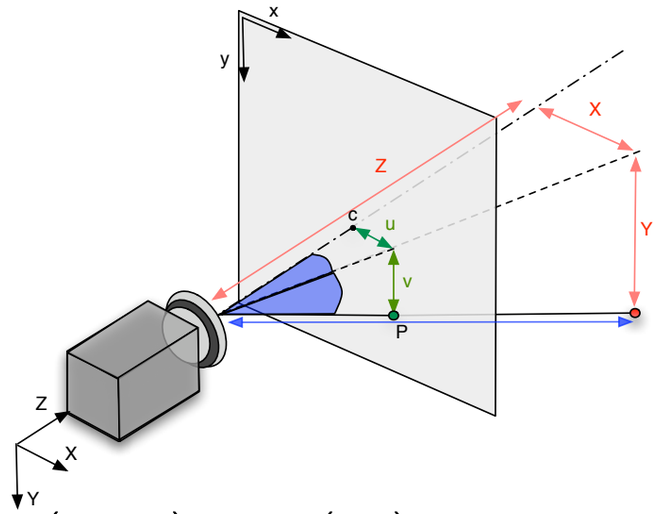
$$\mathbf{\Gamma}(\theta_n, \phi_n) = \left(cos\phi_n sin\theta_n, \quad -sin\phi_n, \quad cos\phi_n cos\theta_n\right)^T.$$

- J. Civera et al. "*Inverse Depth Parametrization for Monocular SLAM*" - TRO 2008

SLAM using Single Camera

- Unified Inverse Depth Parametrization (Undelayed)





$$d_n = \frac{1}{\rho_n}$$

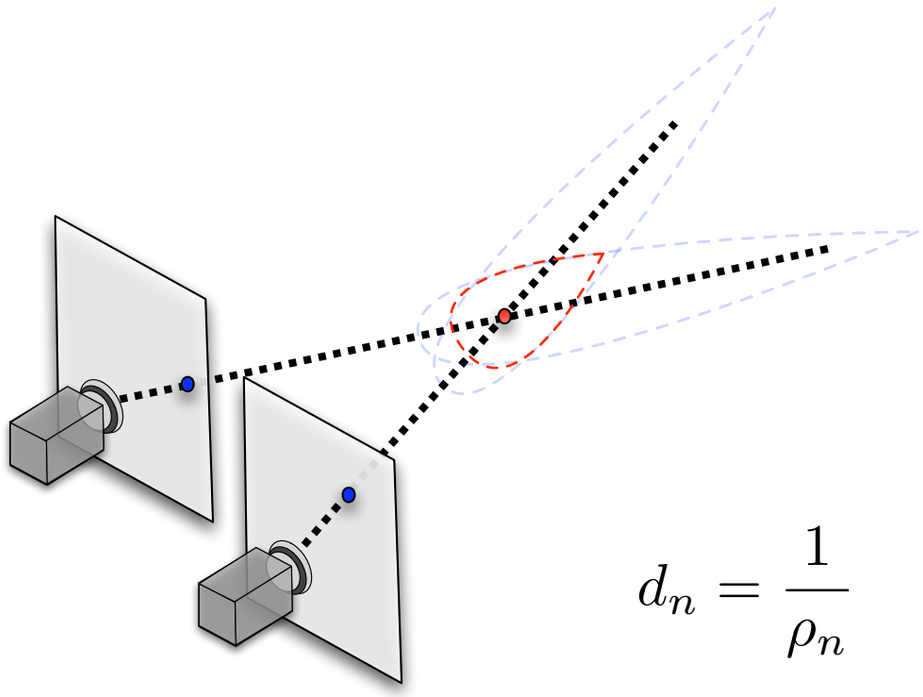
$$\mathbf{m}_n = \begin{pmatrix} m_{x,n} \\ m_{y,n} \\ m_{z,n} \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} + \frac{1}{\rho_n}\mathbf{\Gamma}(\theta_n, \phi_n),$$

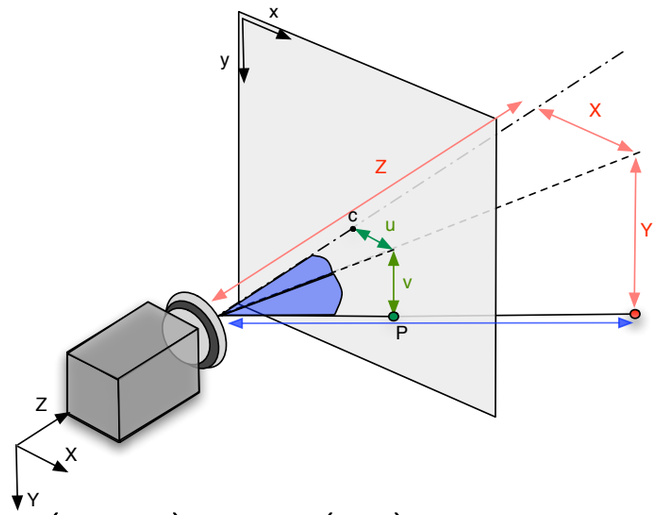
$$\mathbf{\Gamma}(\theta_n, \phi_n) = \begin{pmatrix} cos\phi_n sin\theta_n, & -sin\phi_n, & cos\phi_n cos\theta_n \end{pmatrix}^T.$$

- J. Civera et al. "*Inverse Depth Parametrization for Monocular SLAM*" - TRO 2008

SLAM using Single Camera

- Unified Inverse Depth Parametrization (Undelayed)





$$d_n = \frac{1}{\rho_n}$$

$$\mathbf{m}_n = \begin{pmatrix} m_{x,n} \\ m_{y,n} \\ m_{z,n} \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} + \frac{1}{\rho_n} \mathbf{\Gamma}(\theta_n, \phi_n),$$

$$\mathbf{\Gamma}(\theta_n, \phi_n) = \left(cos\phi_n sin\theta_n, \quad -sin\phi_n, \quad cos\phi_n cos\theta_n \right)^T .$$

- EKF based SLAM

$$\mathbf{h}_n^C = \mathbf{R}^{CW} \left(\rho_n \left(\begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} - \mathbf{r}^{WC} \right) + \mathbf{\Gamma}(\theta_n, \phi_n) \right) .$$

- **6** Parameters for Each Point
- J. Civera et al. "*Inverse Depth Parametrization for Monocular SLAM*" - TRO 2008

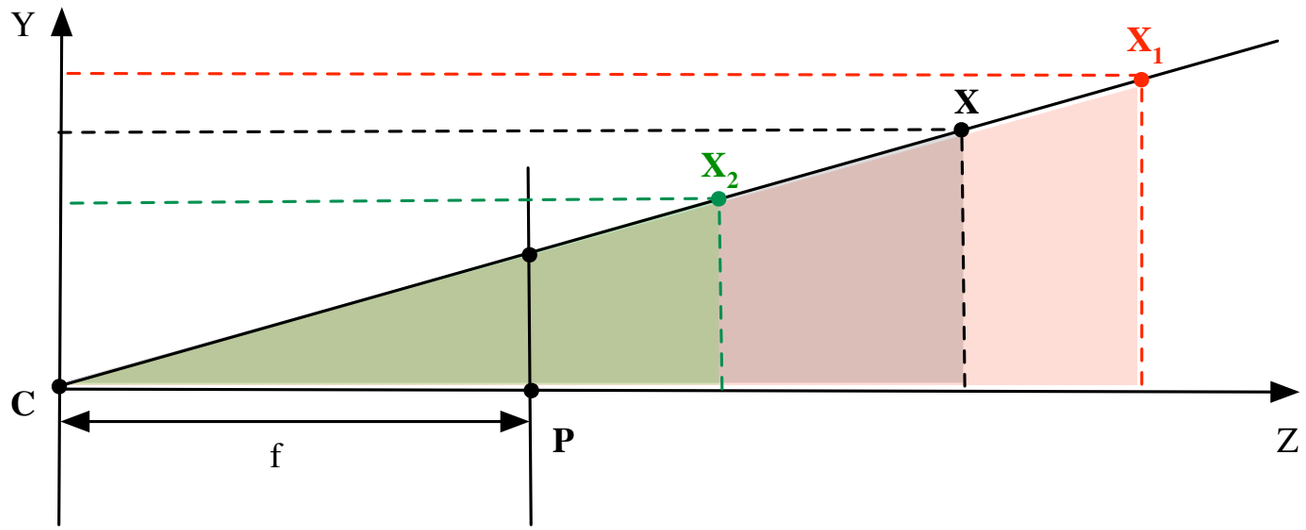
UID Improvements

- J. Civera et al. - "*Inverse depth to depth conversion for Monocular SLAM*" - ICRA 2007
- M. Pupilli et al. - "*Real Time Visual SLAM with resilience to Erratic Motion*" - CVPR 2006
- A. P. Gee et al. - "*Discovering Planes and Collapsing the State Space in Visual SLAM*" - BMVC 2007
- E. Eade et al. - "*Monocular SLAM as a Graph of Coalesced Observations*" - ICCV 2007
- G. Klein et al. - "*Parallel Tracking And Mapping for Small AR workspaces*" - ISMAR 2007
- T. Pietzsch - "*Efficient Feature Parameterisation for visual SLAM using Inverse Depth Bundles*" - BMVC 2008
- E. Imre et al. - "*Improved Inverse Depth Parameterization for Monocular SLAM*" - ICRA 2009

Inverse Scaling Parametrization

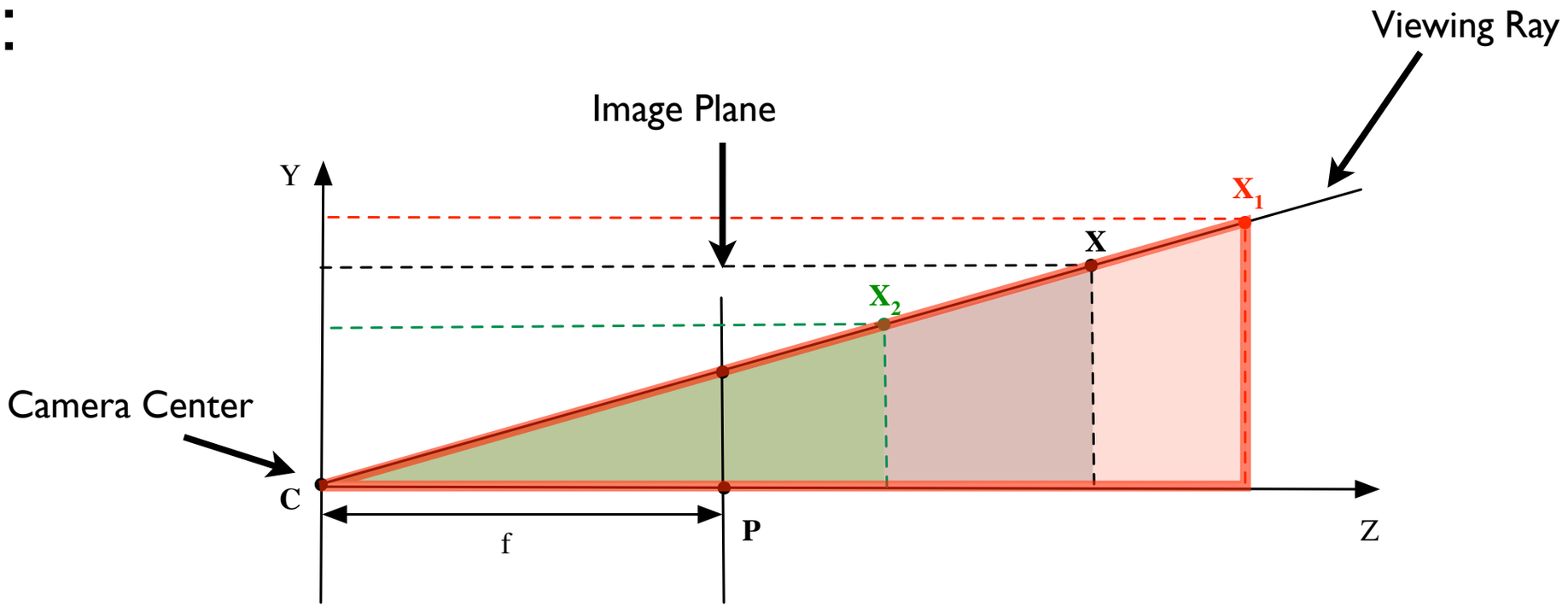
Inverse Scaling Parametrization

- Idea:



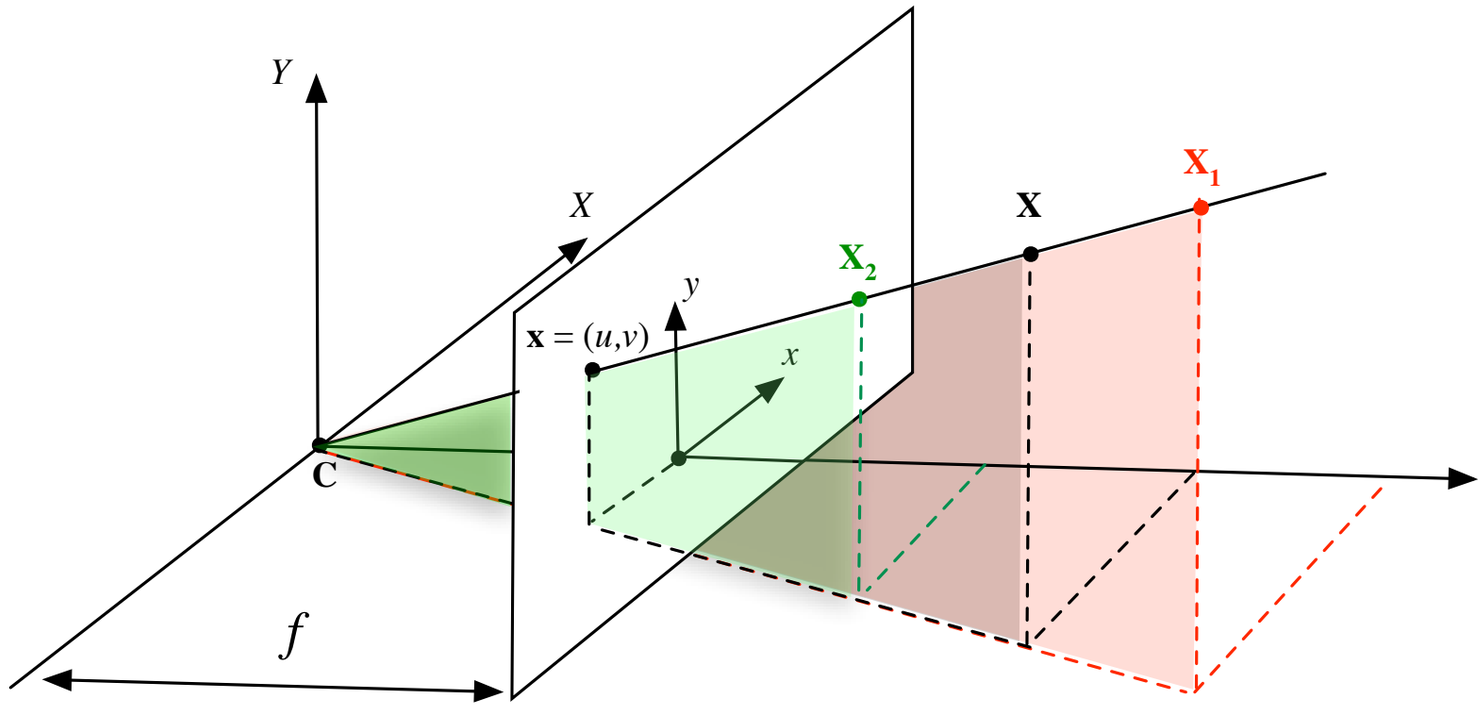
Inverse Scaling Parametrization

- Idea:



Inverse Scaling Parametrization

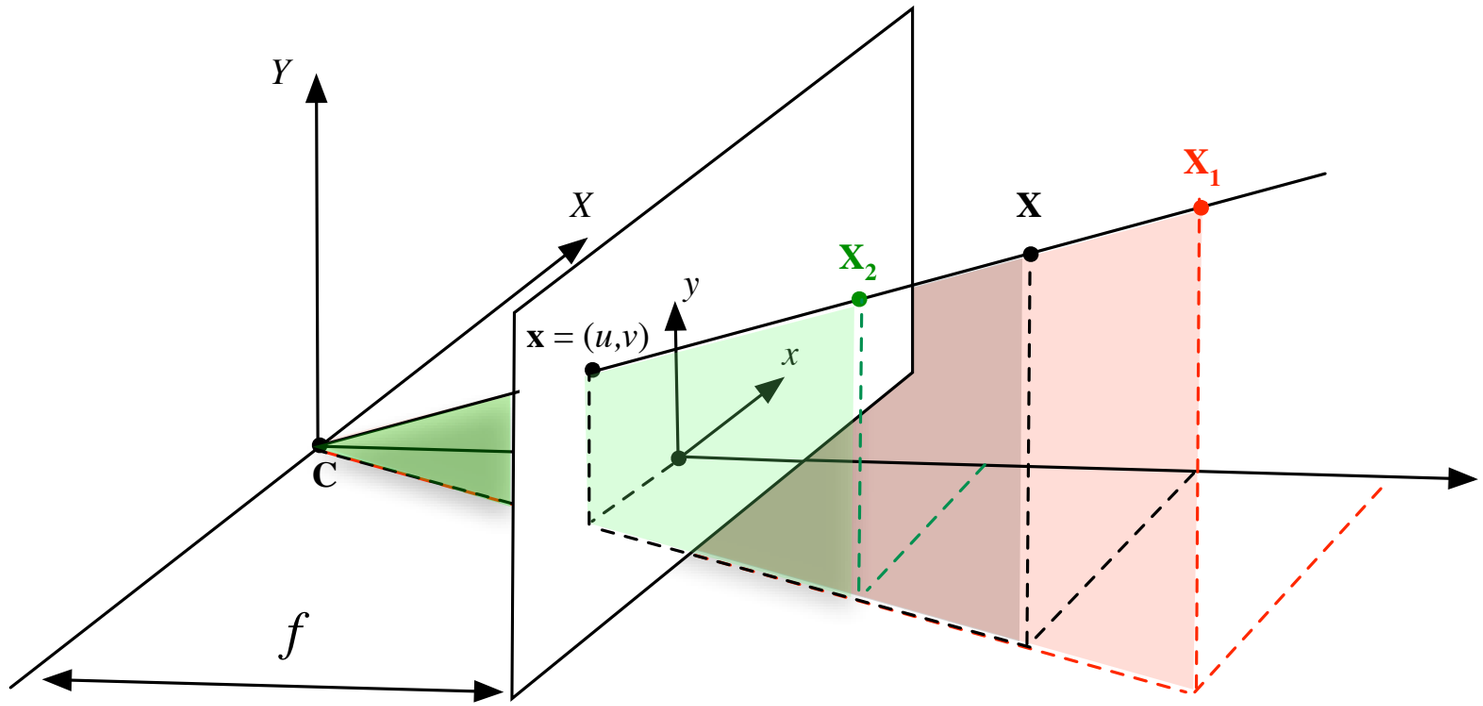
- Idea:



$$\mathbf{X} = \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \qquad \mathbf{X}_i = \begin{pmatrix} \alpha_i X \\ \alpha_i Y \\ \alpha_i Z \\ 1 \end{pmatrix} \quad \rightarrow \quad \mathbf{X}_i \equiv \frac{1}{\alpha_1}\mathbf{X}_i = \begin{pmatrix} X \\ Y \\ Z \\ 1/\alpha_i \end{pmatrix}$$

Inverse Scaling Parametrization

- Idea:

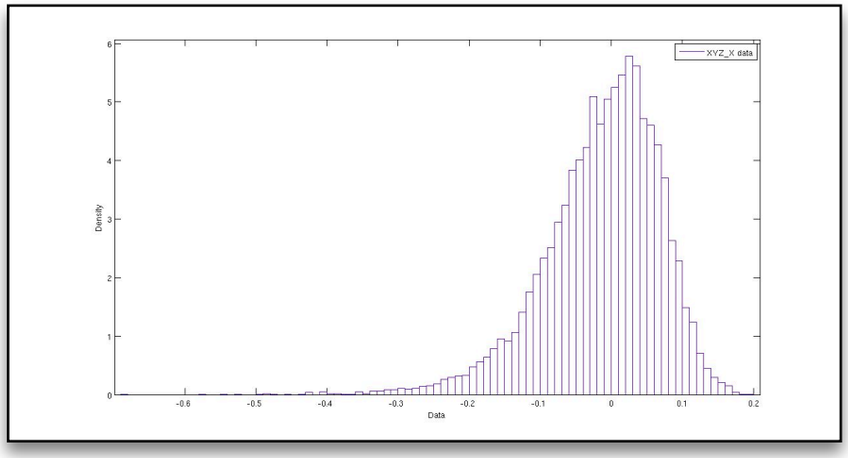


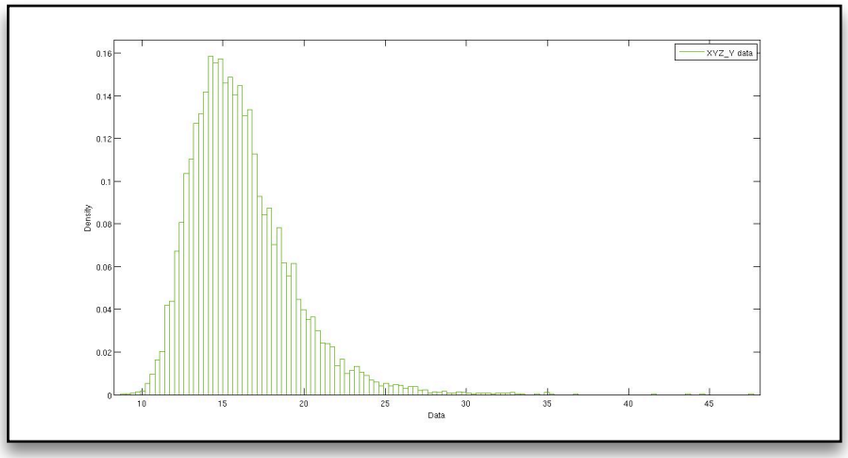
$$\mathbf{X} = \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \qquad \mathbf{X}_i = \begin{pmatrix} \alpha_i X \\ \alpha_i Y \\ \alpha_i Z \\ 1 \end{pmatrix} \longrightarrow \qquad \mathbf{X}_i \equiv \frac{1}{\alpha_1}\mathbf{X}_i = \begin{pmatrix} X \\ Y \\ Z \\ 1/\alpha_i \end{pmatrix}$$

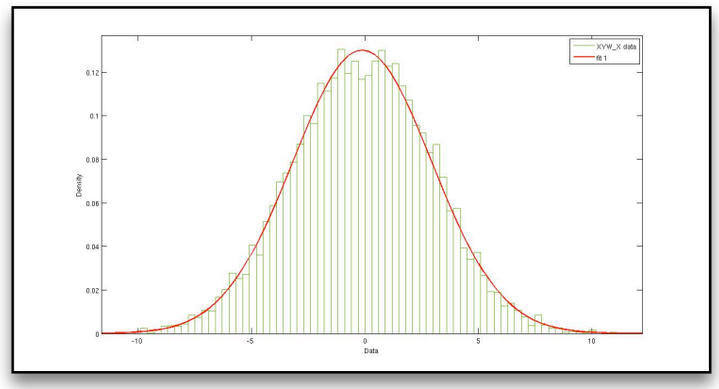
$$\mathbf{X} = \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \equiv \begin{pmatrix} u \\ v \\ f \\ 1/\alpha \end{pmatrix} \equiv \begin{pmatrix} u \\ v \\ f \\ \omega \end{pmatrix},$$

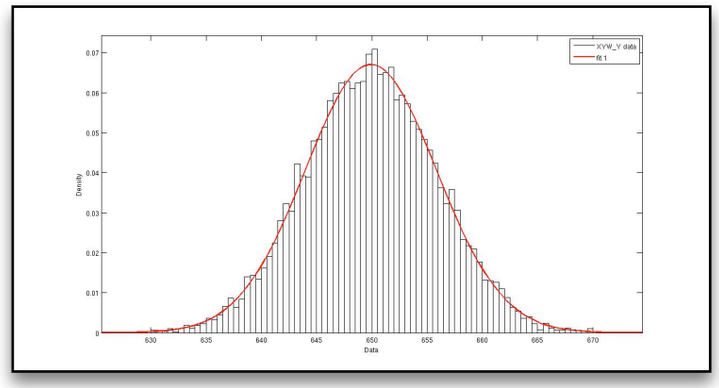
Inverse Scaling Parametrization

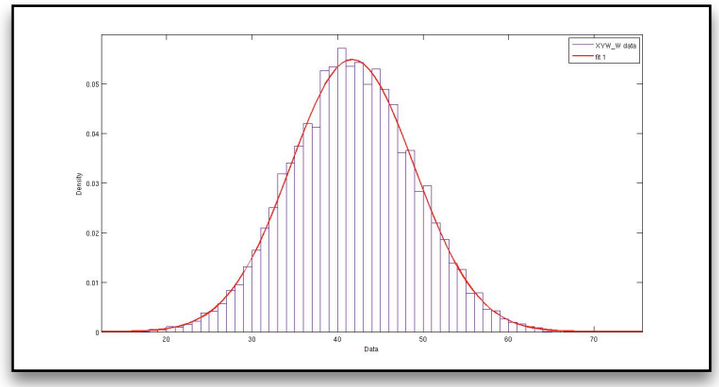
- Uncertainty Modeling - Gaussian approximation





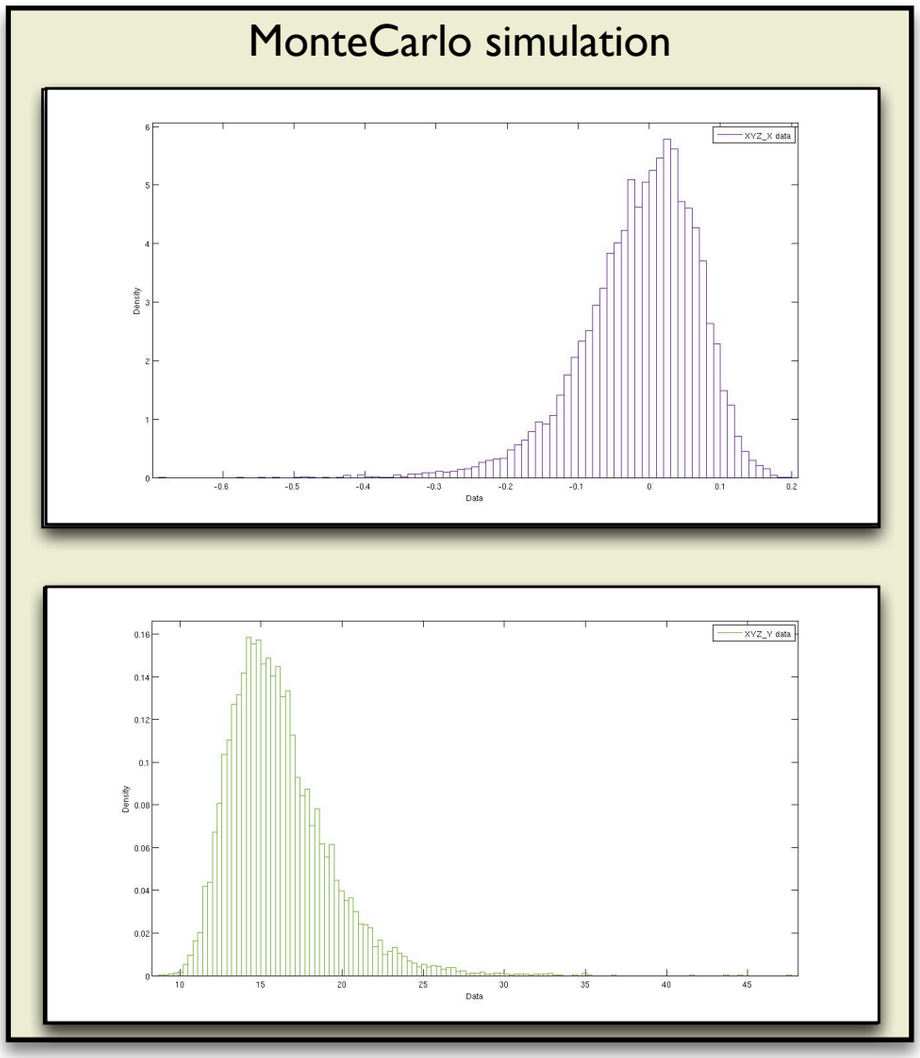


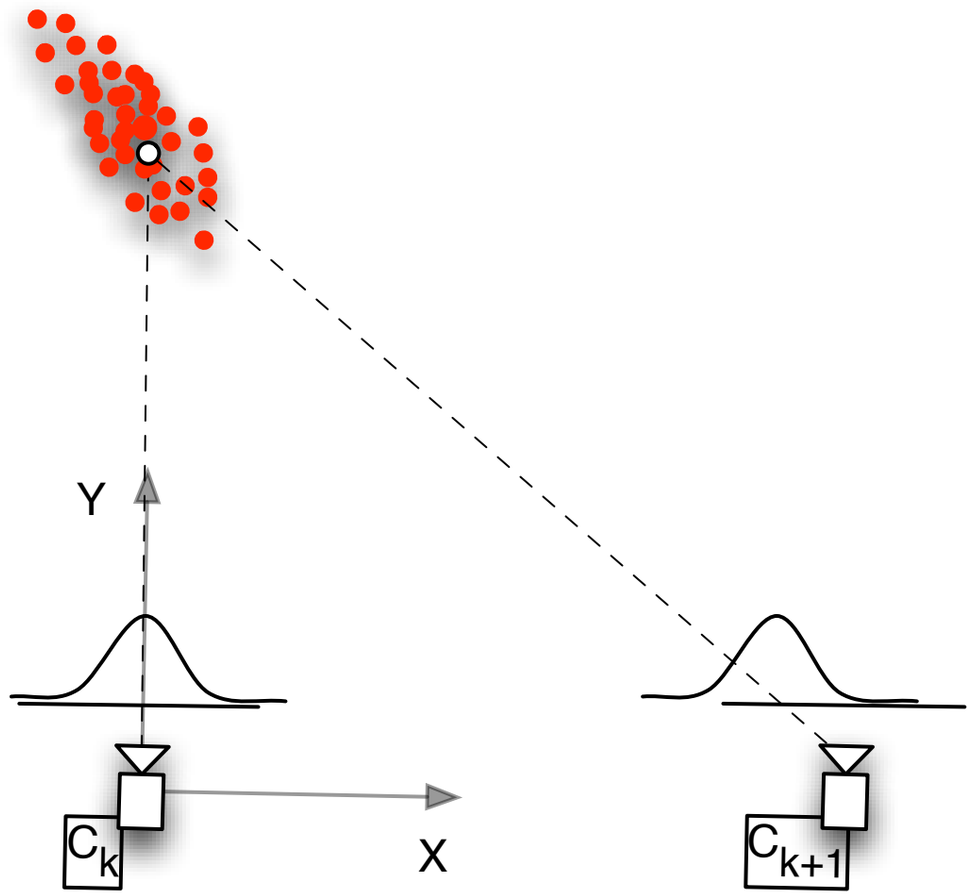




Inverse Scaling Parametrization

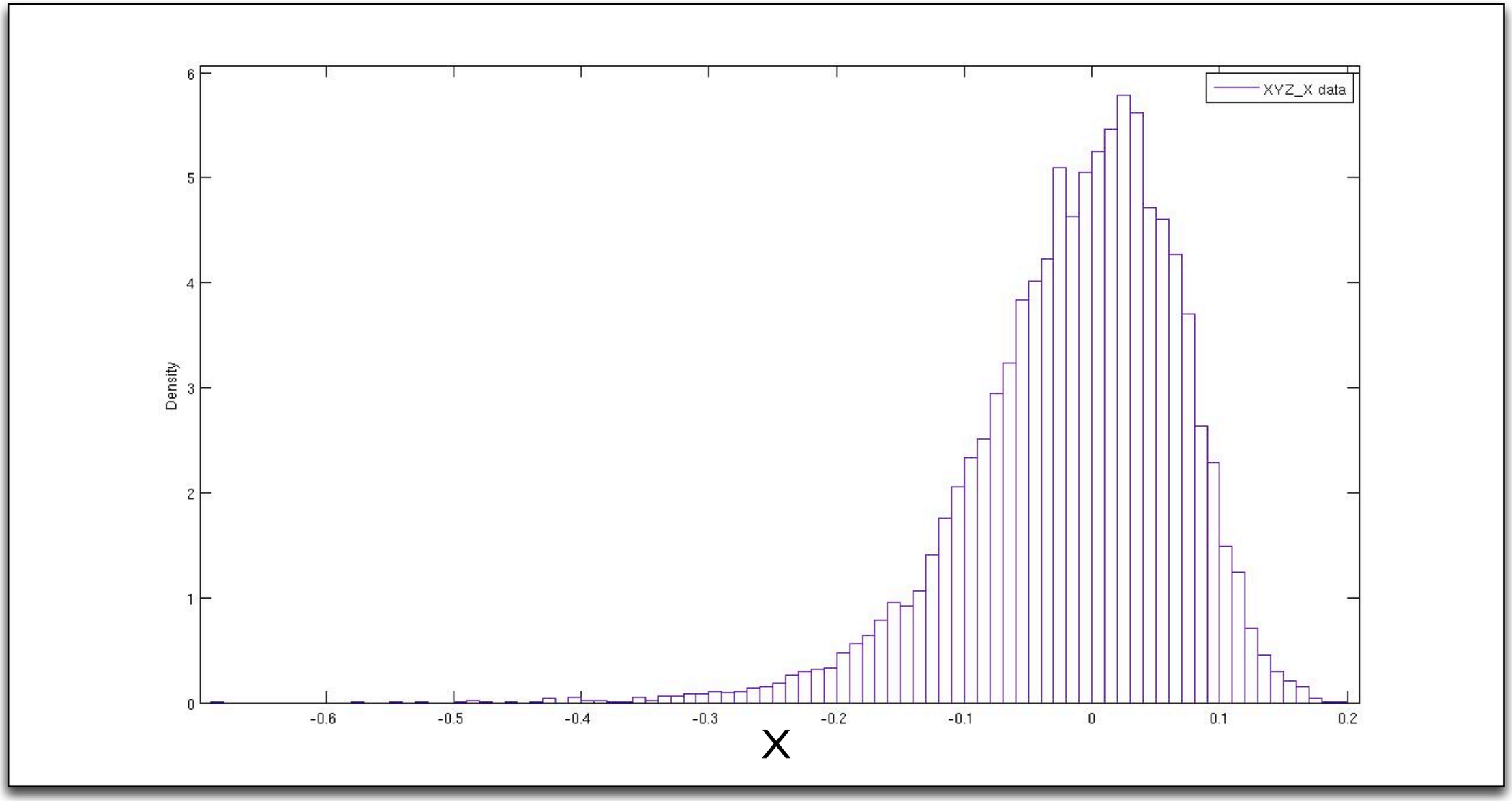
- Uncertainty Modeling - Gaussian approximation





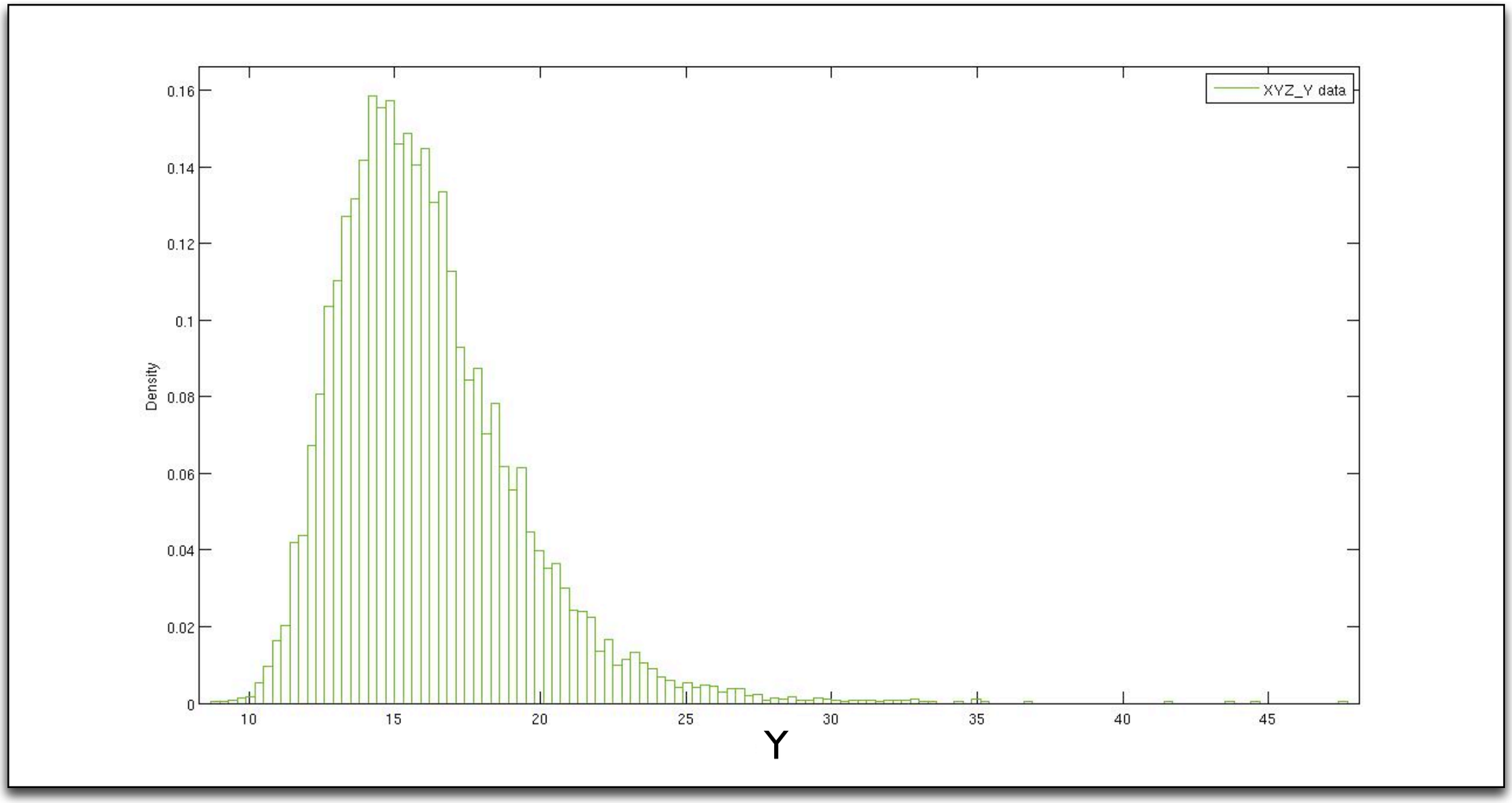
Inverse Scaling Parametrization

- Uncertainty Modeling - Gaussian approximation



Inverse Scaling Parametrization

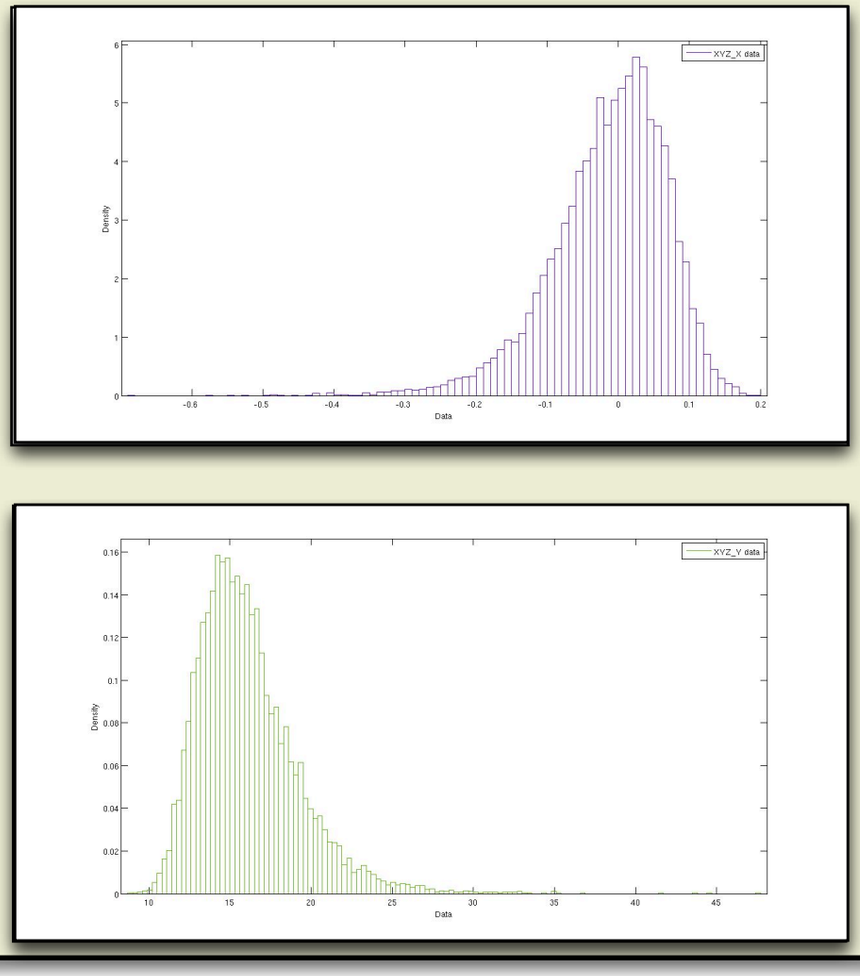
- Uncertainty Modeling - Gaussian approximation



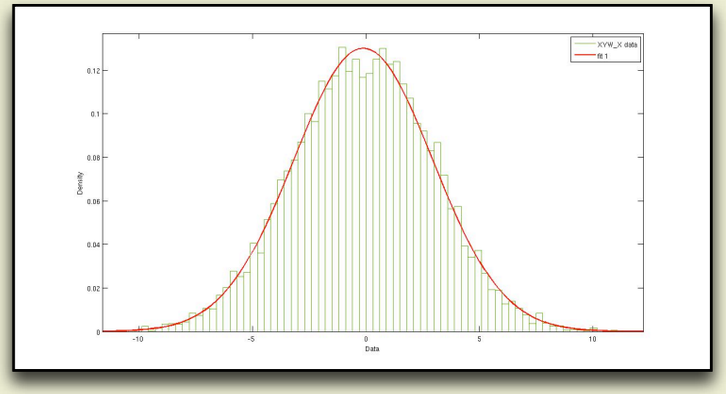
Inverse Scaling Parametrization

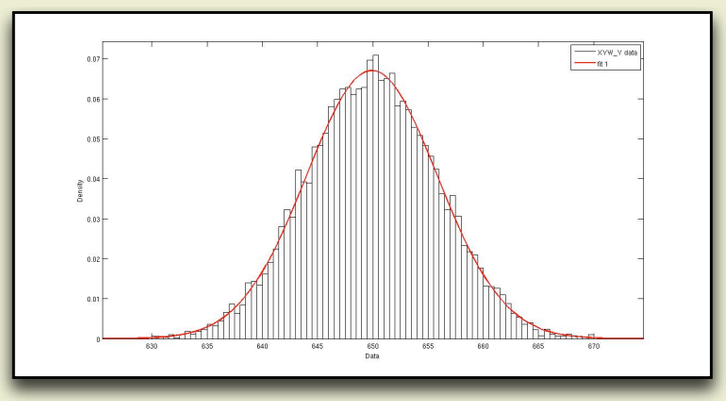
- Uncertainty Modeling - Gaussian approximation

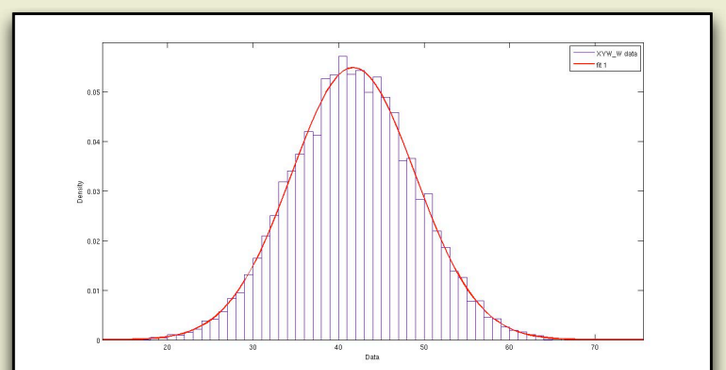
MonteCarlo simulation



Inverse Scaling Representation

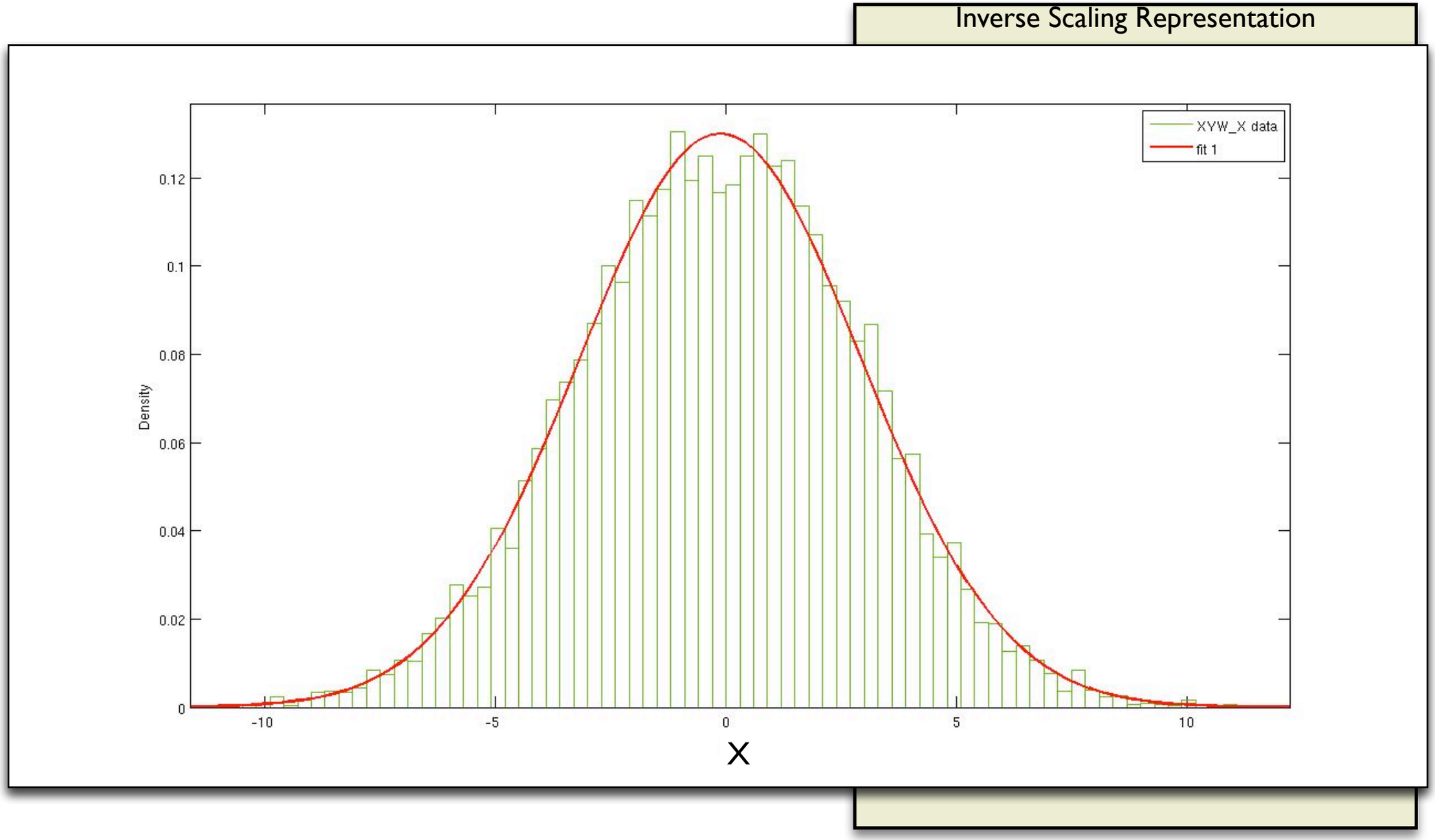






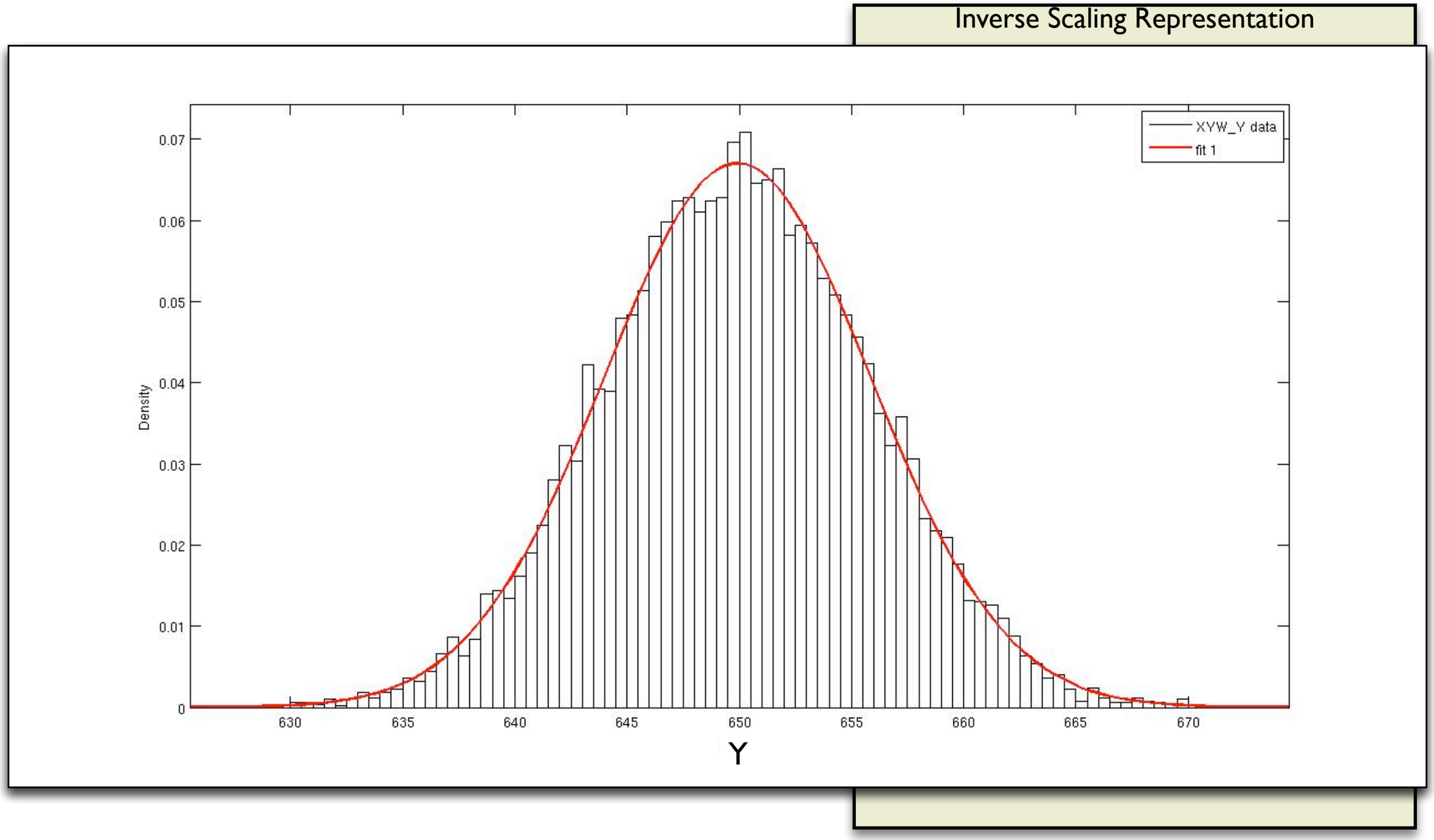
Inverse Scaling Parametrization

- Uncertainty Modeling - Gaussian approximation



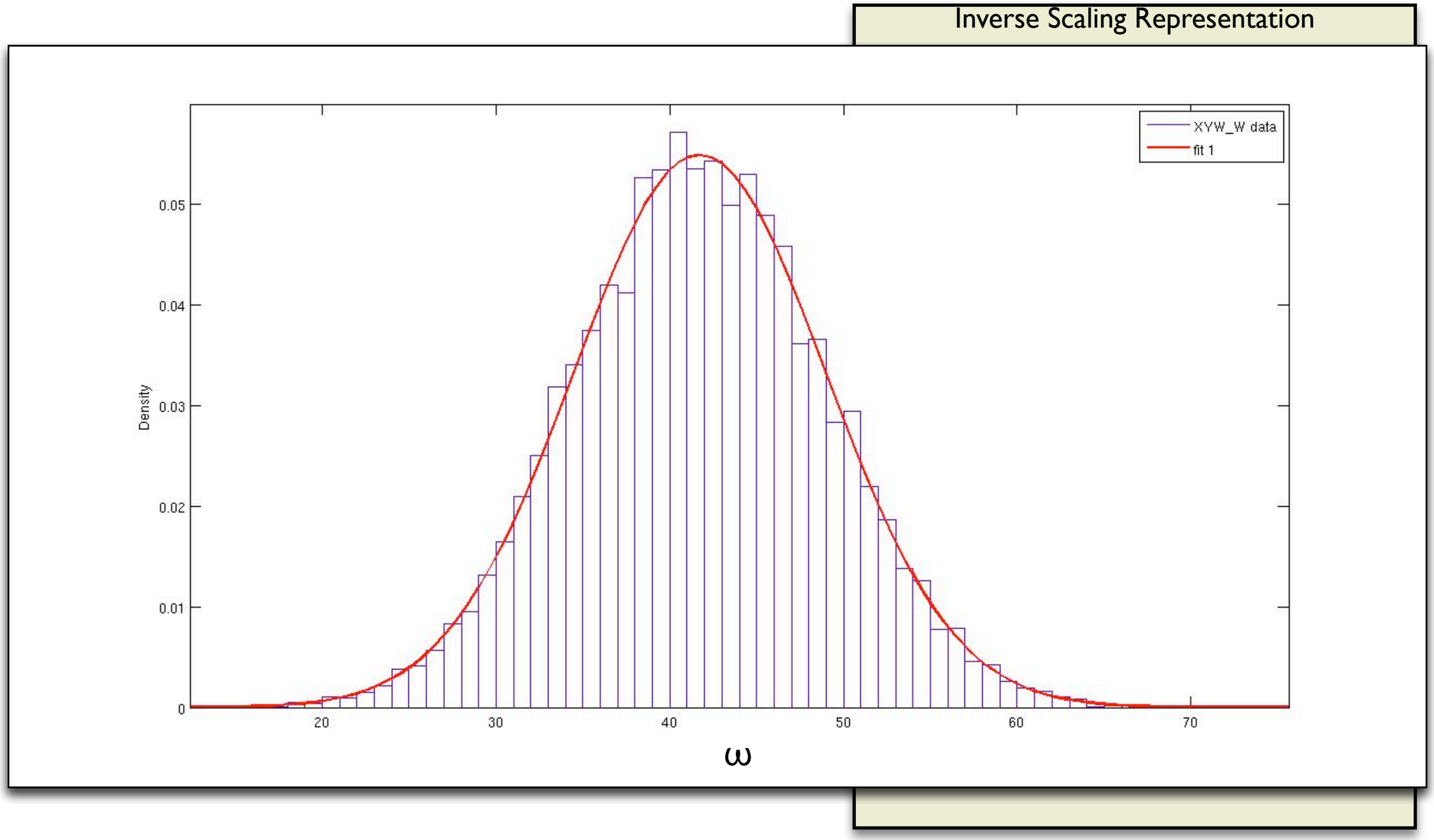
Inverse Scaling Parametrization

- Uncertainty Modeling - Gaussian approximation



Inverse Scaling Parametrization

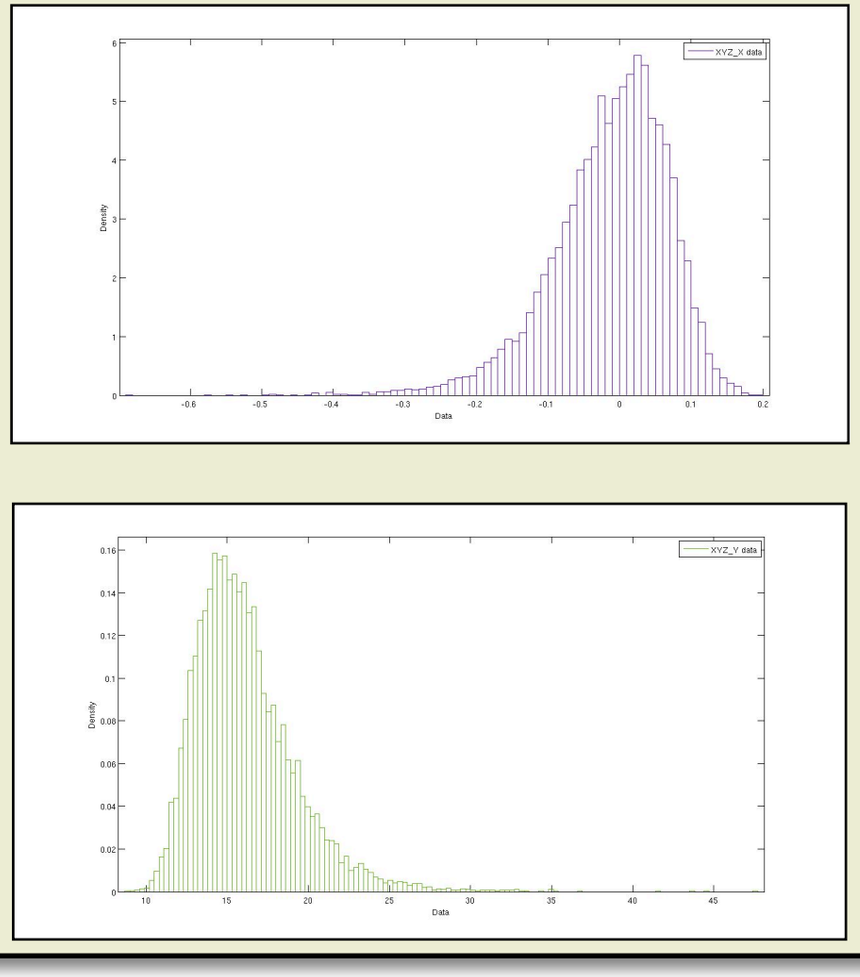
- Uncertainty Modeling - Gaussian approximation



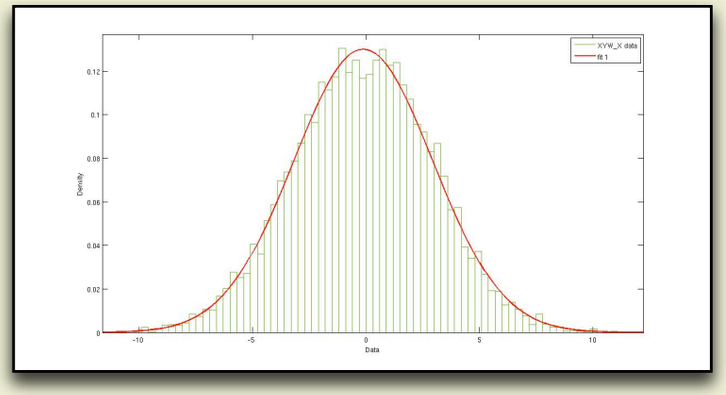
Inverse Scaling Parametrization

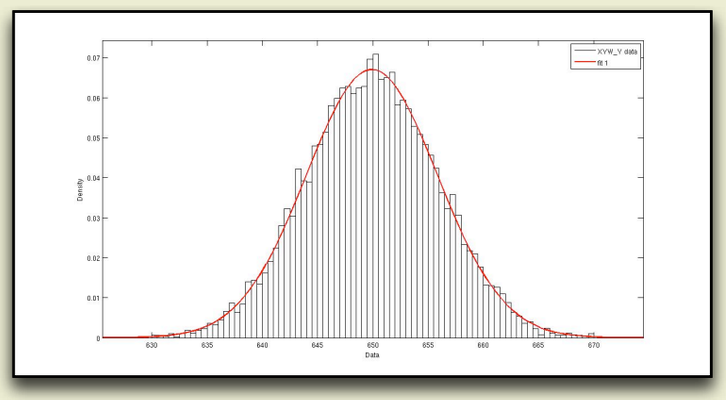
- Uncertainty Modeling - Gaussian approximation

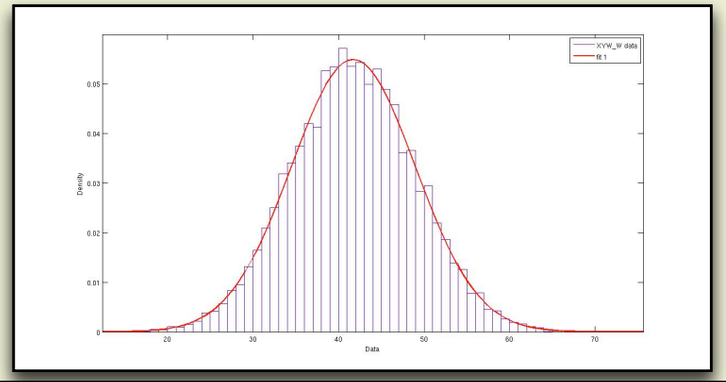
MonteCarlo simulation



Inverse Scaling Representation







MonoSLAM with Inverse Scaling

MonoSLAM with Inverse Scaling

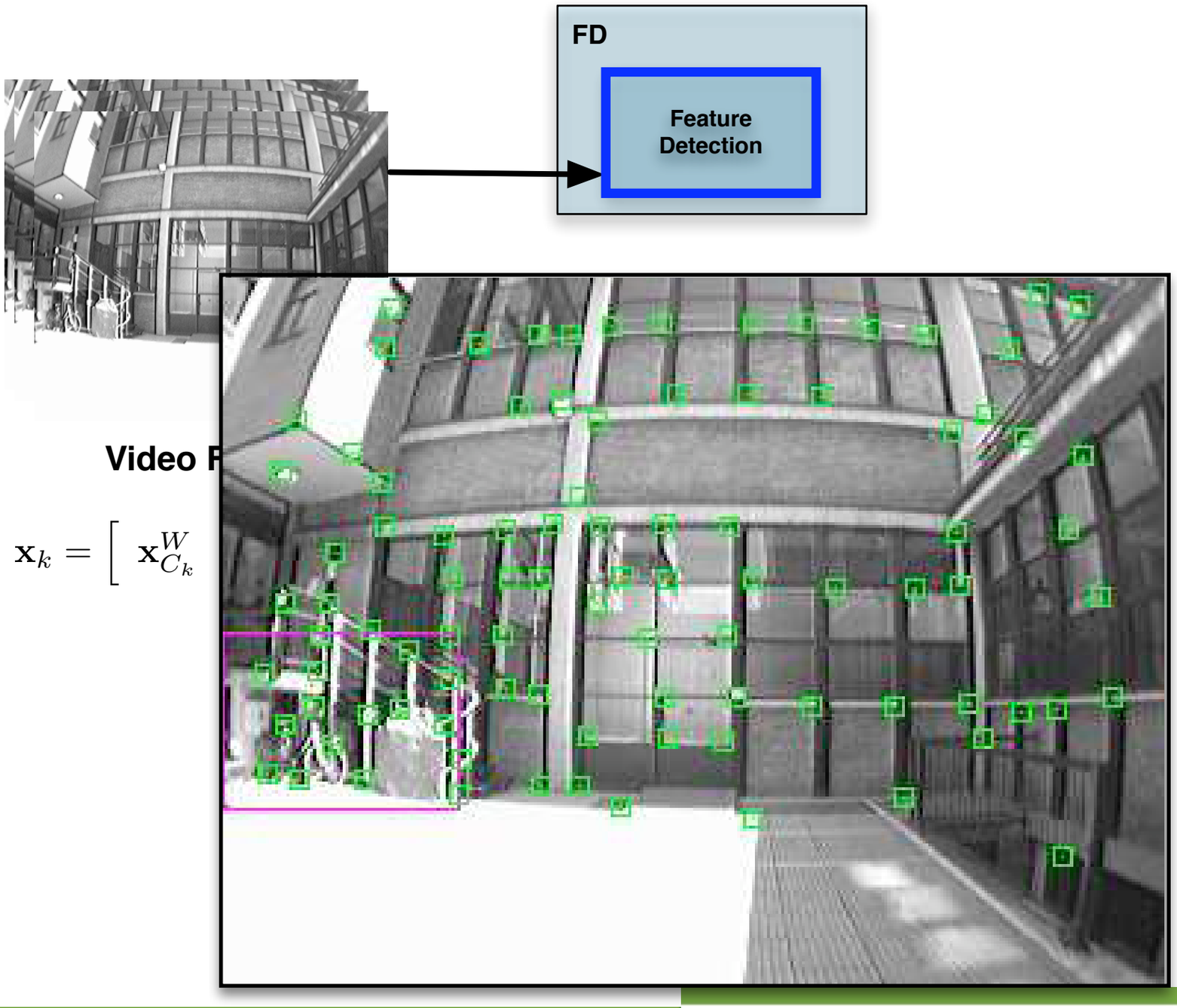
- Extended Kalman Filter

Video Frame

$$\mathbf{x}_k = \left[\begin{array}{cccccc} \mathbf{x}_{C_k}^W & \mathbf{v}_k^{C_k} & \mathbf{x}_{F_1\,k}^W & \cdots & \mathbf{x}_{F_n\,k}^W & \cdots & \mathbf{x}_{F_N\,k}^W \end{array} \right]^T$$

MonoSLAM with Inverse Scaling

- Extended Kalman Filter



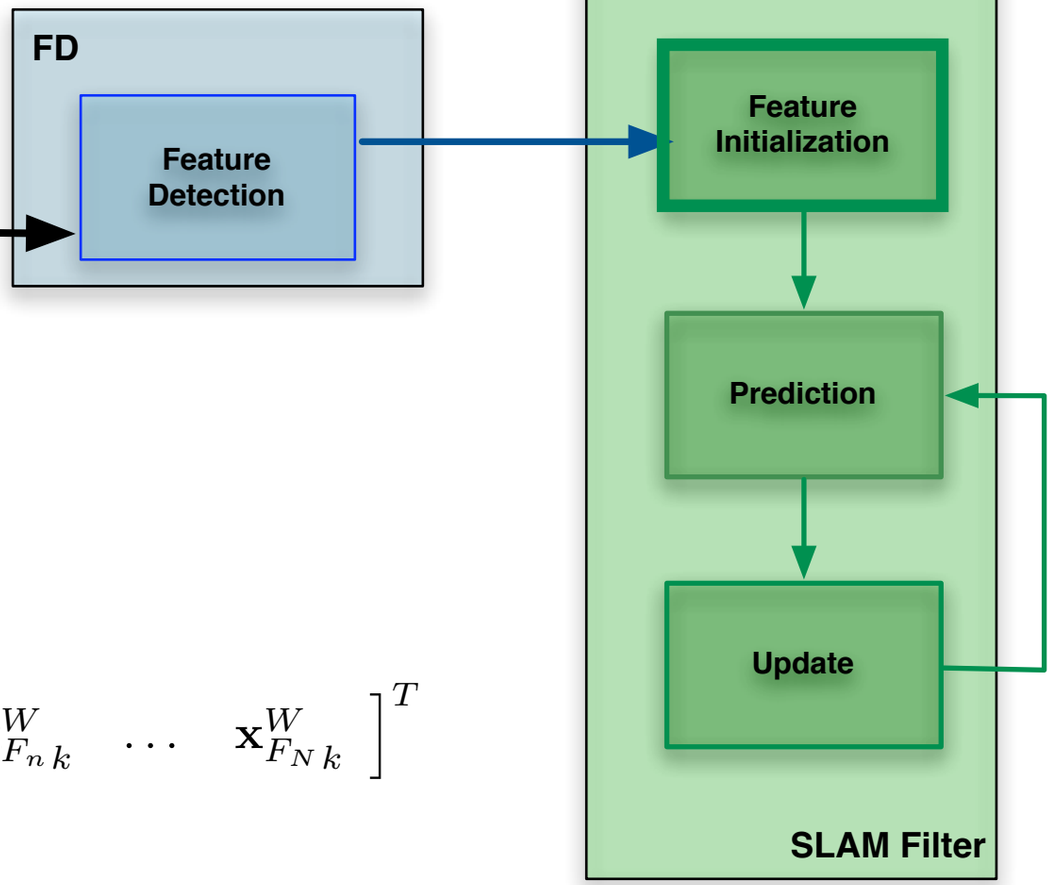
Video R

$$\mathbf{x}_k = \left[\mathbf{x}_{C_k}^W \right.$$

MonoSLAM with Inverse Scaling

- Extended Kalman Filter

Video Frame



$$\mathbf{x}_k = \begin{bmatrix} \mathbf{x}_{C_k}^W & \mathbf{v}_k^{C_k} & \mathbf{x}_{F_1\,k}^W & \cdots & \mathbf{x}_{F_n\,k}^W & \cdots & \mathbf{x}_{F_N\,k}^W \end{bmatrix}^T$$

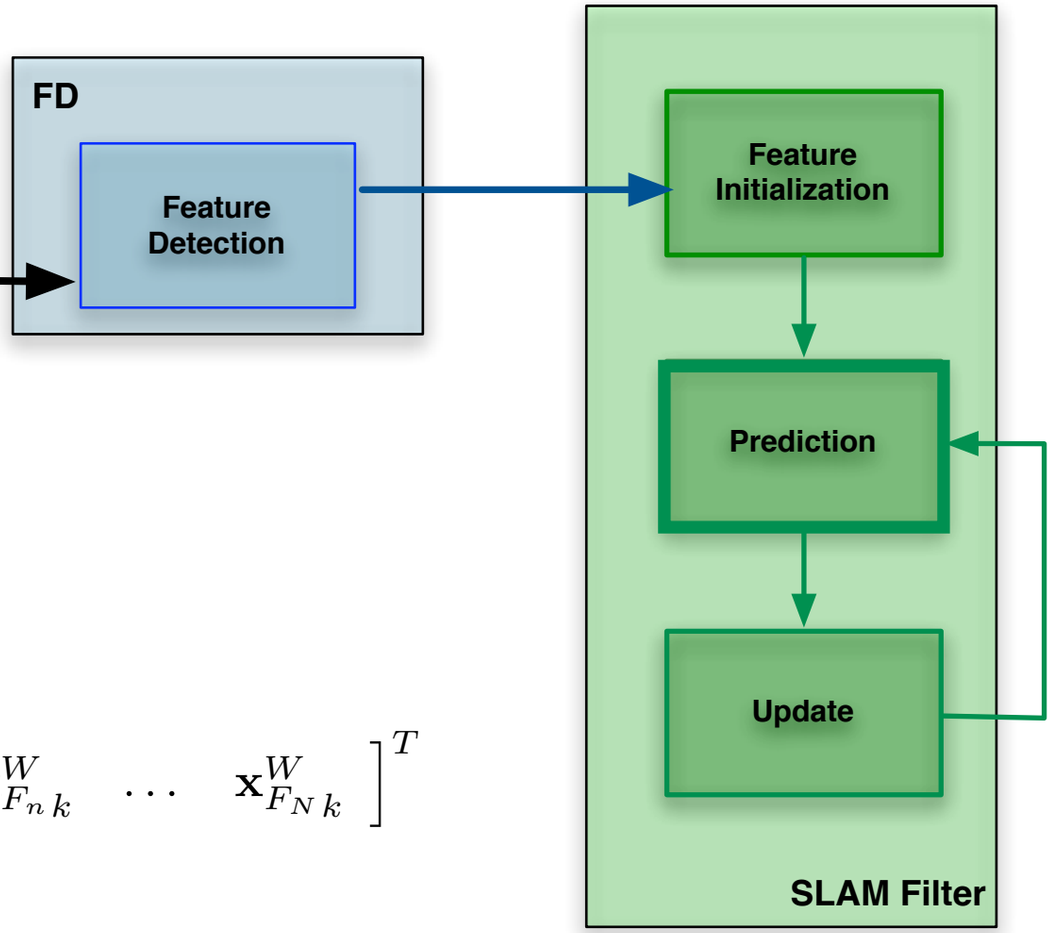
$$\begin{pmatrix} x \\ y \\ z \\ \omega \end{pmatrix} = \begin{pmatrix} u - cc_x \\ v - cc_y \\ fc \\ \hat{\omega} \end{pmatrix}$$

$$\hat{\omega} = fc/(2 * min_d)$$

MonoSLAM with Inverse Scaling

- Extended Kalman Filter

Video Frame



$$\mathbf{x}_k = \begin{bmatrix} \mathbf{x}_{C_k}^W & \mathbf{v}_k^{C_k} & \mathbf{x}_{F_1\,k}^W & \dots & \mathbf{x}_{F_n\,k}^W & \dots & \mathbf{x}_{F_N\,k}^W \end{bmatrix}^T$$

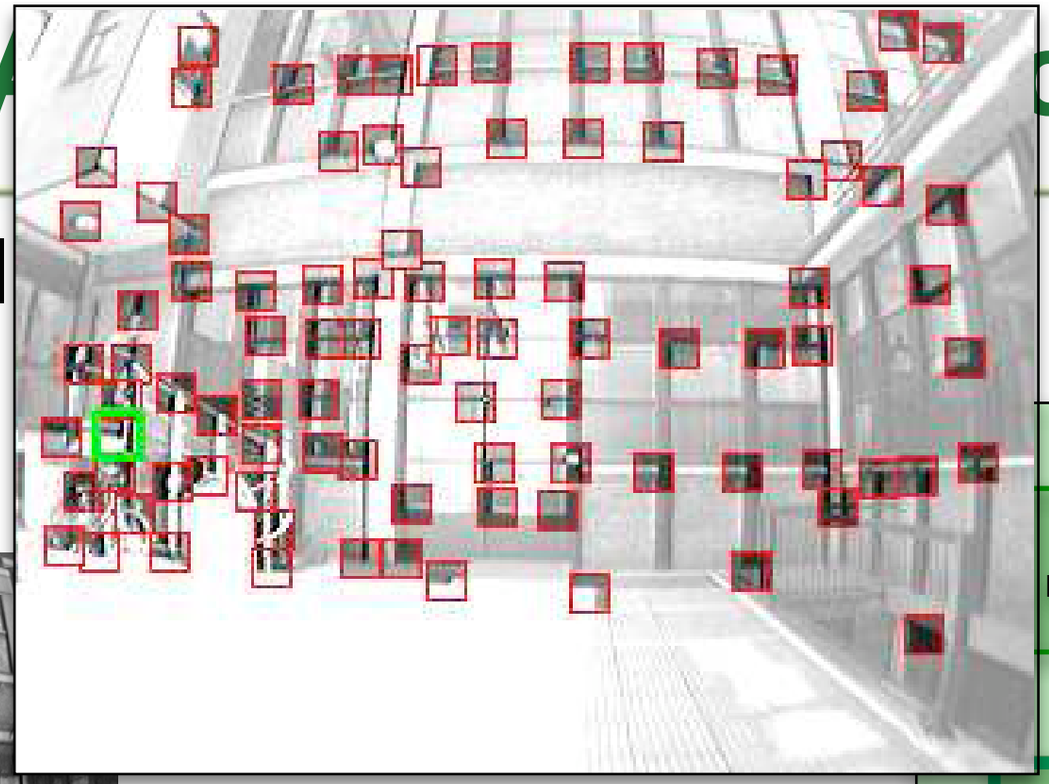
$$\hat{\mathbf{x}}_k = \begin{bmatrix} \mathbf{x}_{C_{k-1}}^W \oplus \mathbf{x}_{C_k}^{C_{k-1}} \\ \mathbf{v}_k^{C_k} \\ \mathbf{x}_{F_1\,k-1}^W \\ \vdots \\ \mathbf{x}_{F_N\,k-1}^W \end{bmatrix}, \quad \begin{aligned} \mathbf{v}_k^{C_k} &= \mathbf{v}_{k-1}^{C_{k-1}} + \mathbf{a} \cdot \Delta t, \\ \mathbf{x}_{C_k}^{C_{k-1}} &= \mathbf{v}_{k-1}^{C_{k-1}} \cdot \Delta t. \end{aligned}$$

$$\hat{\mathbf{P}}_k = \mathbf{J}_1 \mathbf{P}_{k-1} \mathbf{J}_1^T + \mathbf{J}_2 \mathbf{Q} \mathbf{J}_2^T$$

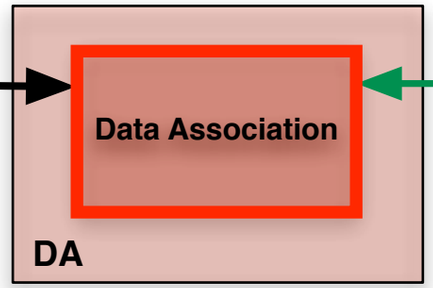
$$\mathbf{J}_1 = \begin{bmatrix} \mathbf{J}_x & \mathbf{J}_v & \dots & \mathbf{J}_{F_n} \end{bmatrix}, \quad \mathbf{J}_2 = \begin{bmatrix} \mathbf{J}_{a_k} \end{bmatrix}$$

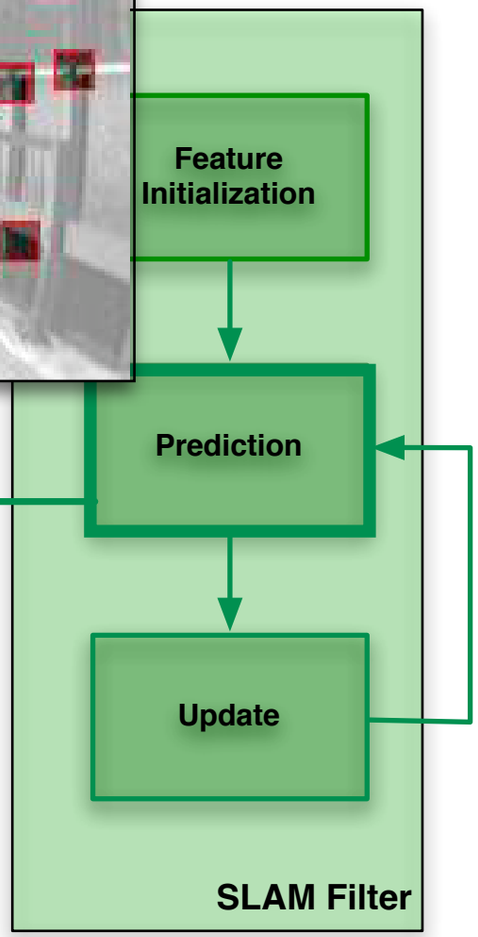
MonoSLA caling

- Extended Kal



Video Frame





$$\mathbf{x}_k = \begin{bmatrix} \mathbf{x}_{C_k}^W & \mathbf{v}_k^{C_k} & \mathbf{x}_{F_1\,k}^W & \dots & \mathbf{x}_{F_n\,k}^W & \dots & \mathbf{x}_{F_N\,k}^W \end{bmatrix}^T$$

$$\hat{\mathbf{x}}_k = \begin{bmatrix} \mathbf{x}_{C_{k-1}}^W \oplus \mathbf{x}_{C_k}^{C_{k-1}} \\ \mathbf{v}_k^{C_k} \\ \mathbf{x}_{F_1\,k-1}^W \\ \vdots \\ \mathbf{x}_{F_N\,k-1}^W \end{bmatrix}, \quad \begin{aligned} \mathbf{v}_k^{C_k} &= \mathbf{v}_{k-1}^{C_{k-1}} + \mathbf{a} \cdot \Delta t, \\ \mathbf{x}_{C_k}^{C_{k-1}} &= \mathbf{v}_{k-1}^{C_{k-1}} \cdot \Delta t. \end{aligned}$$

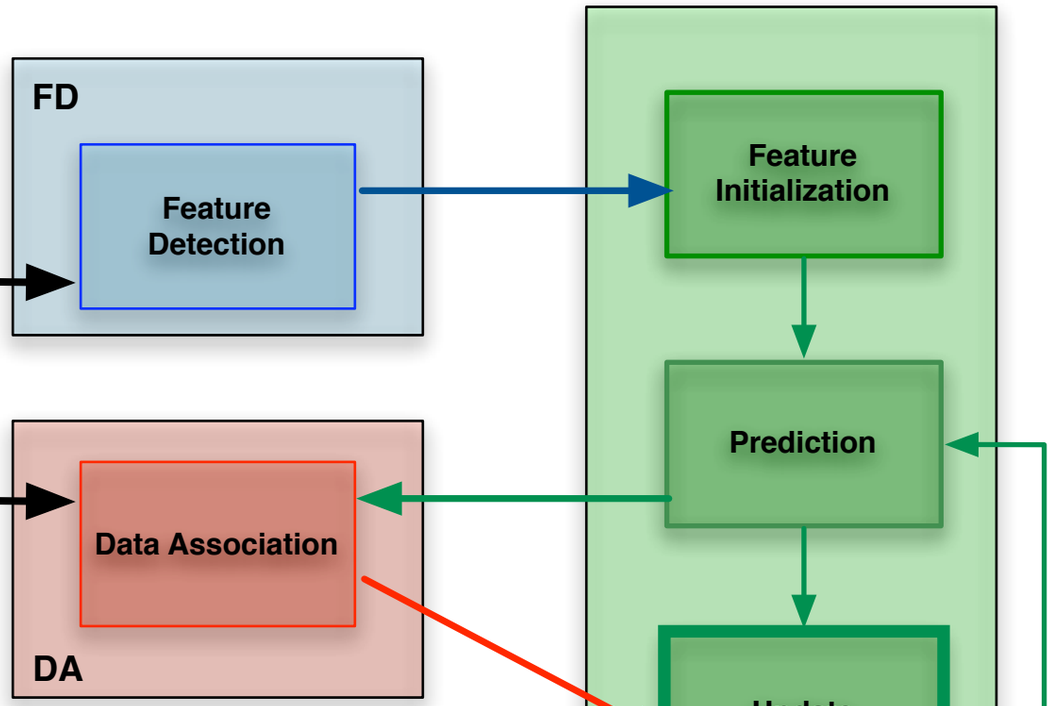
$$\hat{\mathbf{P}}_k = \mathbf{J}_1 \mathbf{P}_{k-1} \mathbf{J}_1^T + \mathbf{J}_2 \mathbf{Q} \mathbf{J}_2^T$$

$$\mathbf{J}_1 = \begin{bmatrix} \mathbf{J}_x & \mathbf{J}_v & \dots & \mathbf{J}_{F_n} \end{bmatrix}, \quad \mathbf{J}_2 = \begin{bmatrix} \mathbf{J}_{a_k} \end{bmatrix}$$

MonoSLAM with Inverse Scaling

- Extended Kalman Filter

Video Frame



$$\mathbf{x}_k = \begin{bmatrix} \mathbf{x}_{C_k}^W & \mathbf{v}_k^{C_k} & \mathbf{x}_{F_1\,k}^W & \dots & \mathbf{x}_{F_n\,k}^W & \dots & \mathbf{x}_{F_N\,k}^W \end{bmatrix}^T$$

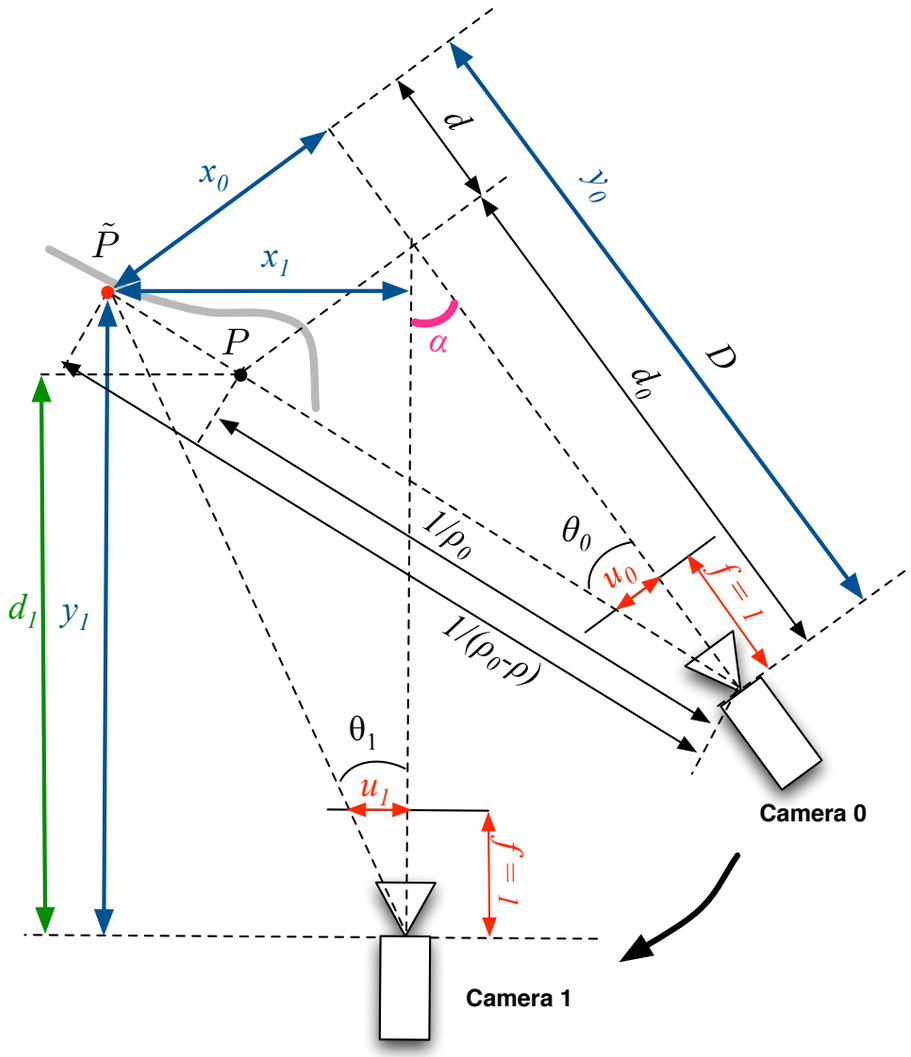
$$\mathbf{h}_n^{C_k} = \mathbf{M}\left(\mathbf{R}_W^{C_k}\left(\begin{bmatrix} x_{F_n}^W \\ y_{F_n}^W \\ z_{F_n}^W \end{bmatrix} - \omega_{F_n}^W \mathbf{r}_{C_k}^W\right)\right) \quad \mathbf{h}_{k,n} = \begin{bmatrix} \frac{h_{x,n}^{C_k}}{h_z^{C_k}} \\ \frac{h_{y,n}^{C_k}}{h_z^{C_k}} \end{bmatrix}.$$

$$\mathbf{S} = \mathbf{H}_k \hat{\mathbf{P}}_k \mathbf{H}_k^T + \mathbf{W}_k \mathbf{R}_k \mathbf{W}_k^T$$
$$\mathbf{K} = \hat{\mathbf{P}}_k \mathbf{H}_k^T \mathbf{S}^{-1}$$
$$\mathbf{P}_k = \hat{\mathbf{P}}_k - \mathbf{K}\mathbf{S}\mathbf{K}^T$$
$$\mathbf{x}_k = \hat{\mathbf{x}}_k + \mathbf{K}\left(\mathbf{z}_k - \mathbf{h}_k\right)$$

Inverse Scaling Parametrization

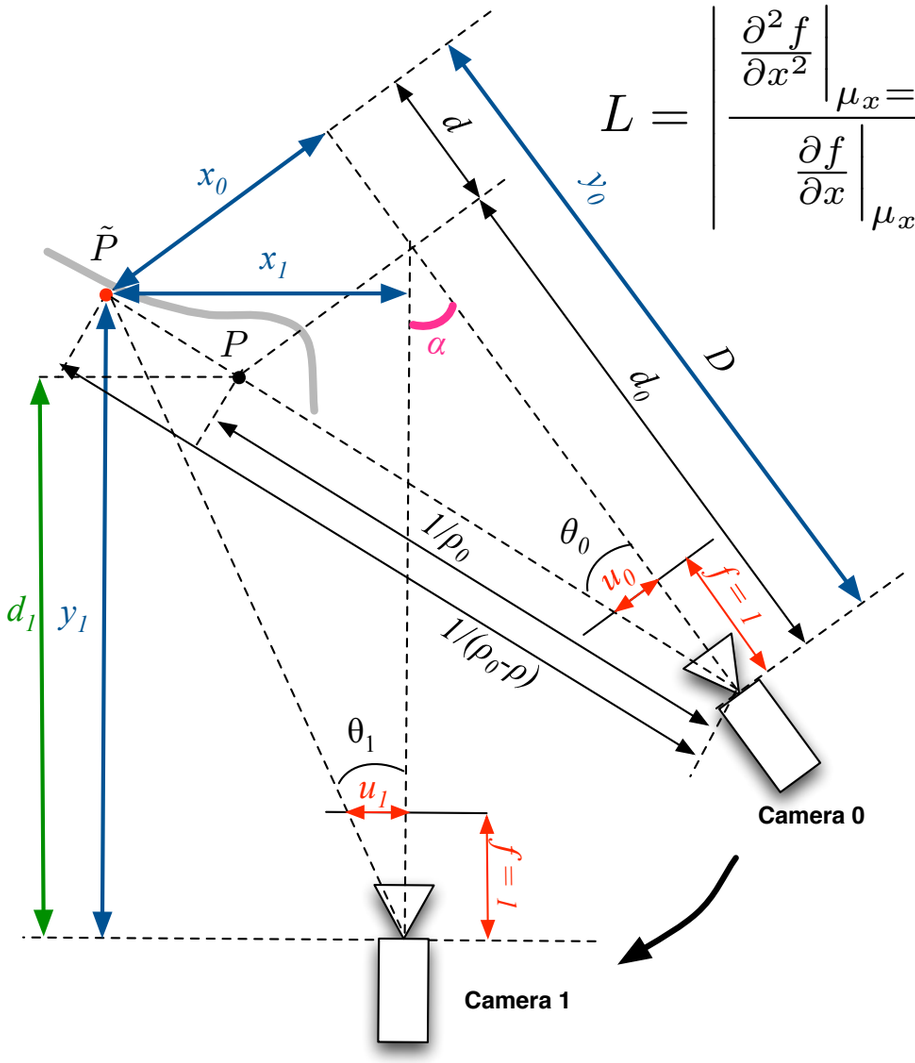
Inverse Scaling Parametrization

- Improvements w.r.t. Unified Inverse Depth:
  - Measurement model non-linearity

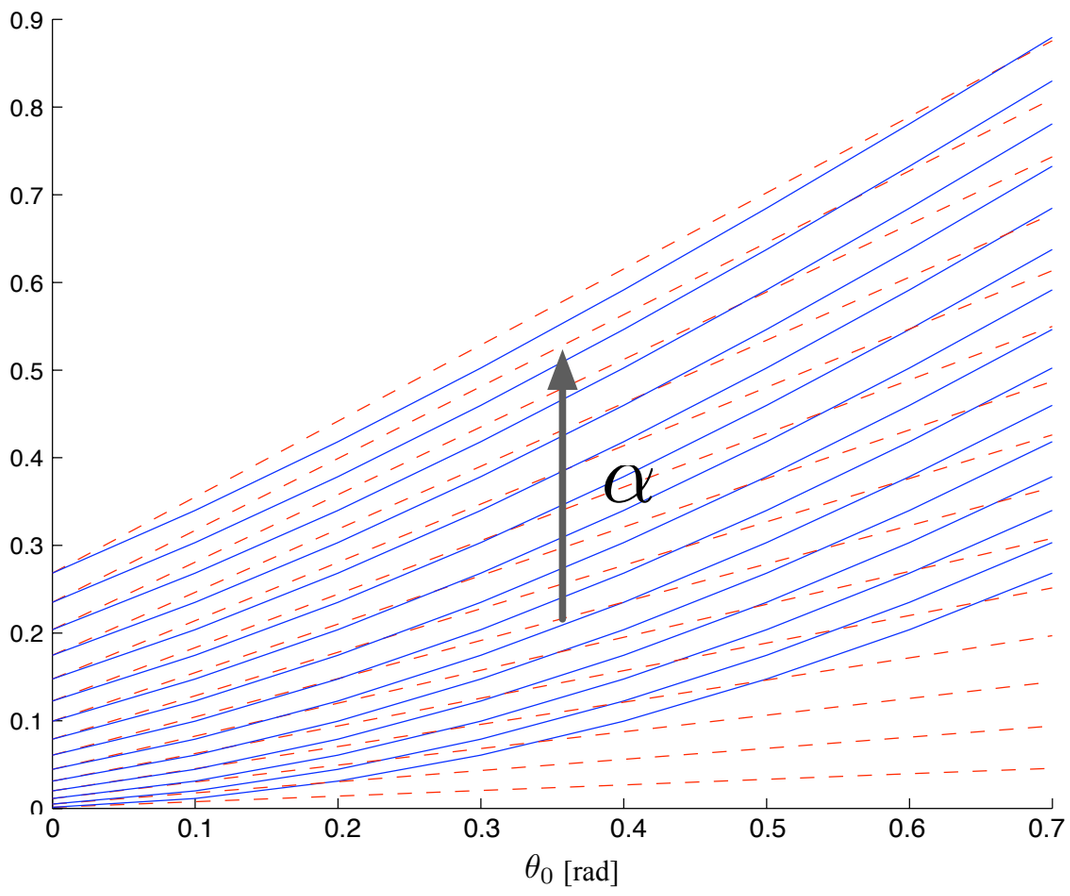


Inverse Scaling Parametrization

- Improvements w.r.t. Unified Inverse Depth:
  - Measurement model non-linearity

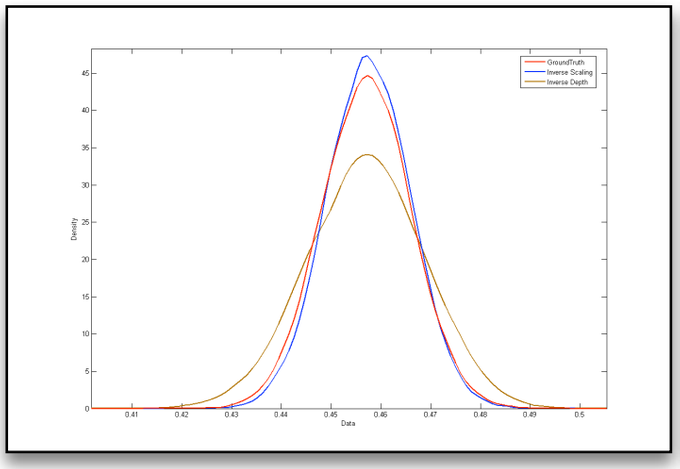


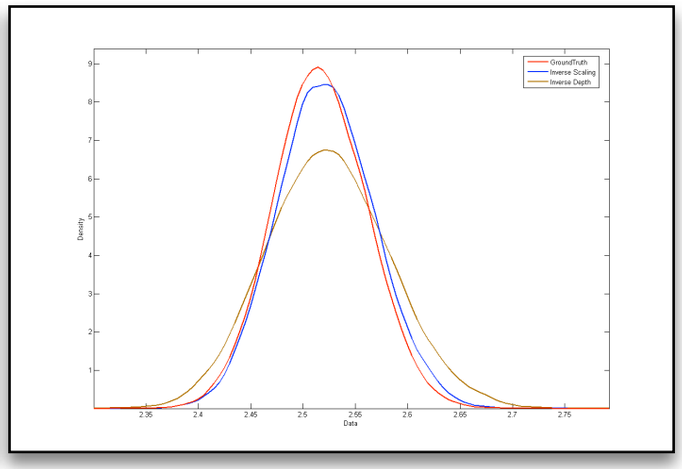
$$L = \left| \frac{\left.\frac{\partial^2 f}{\partial x^2}\right|_{\mu_x=0} 2\sigma_x}{\left.\frac{\partial f}{\partial x}\right|_{\mu_x=0}} \right|,$$

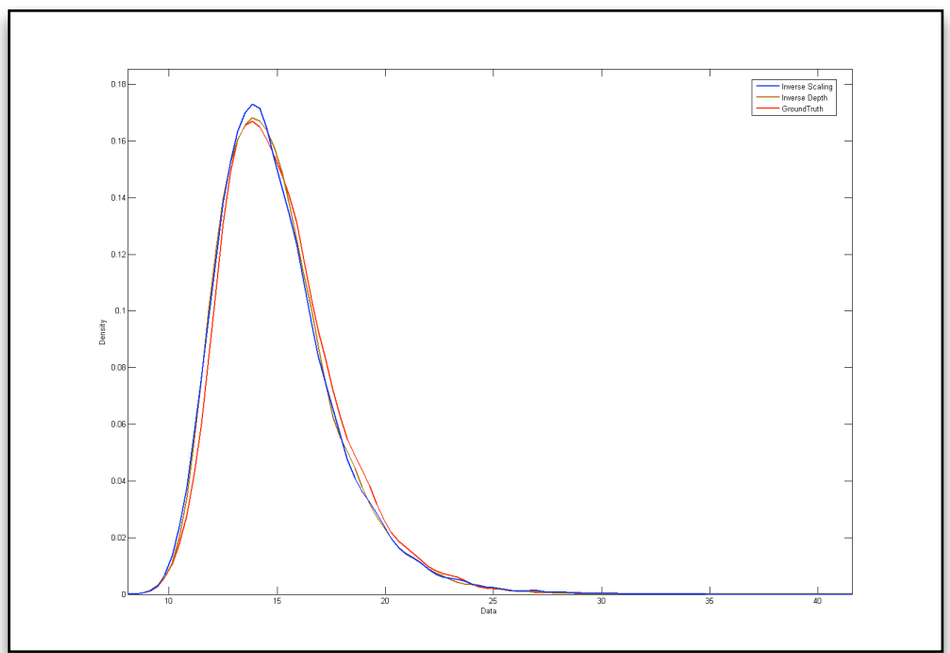


- J. Civera et al. - "*Inverse depth to depth conversion for monocular* SLAM" - ICRA 2007

Inverse Scaling Parametrization



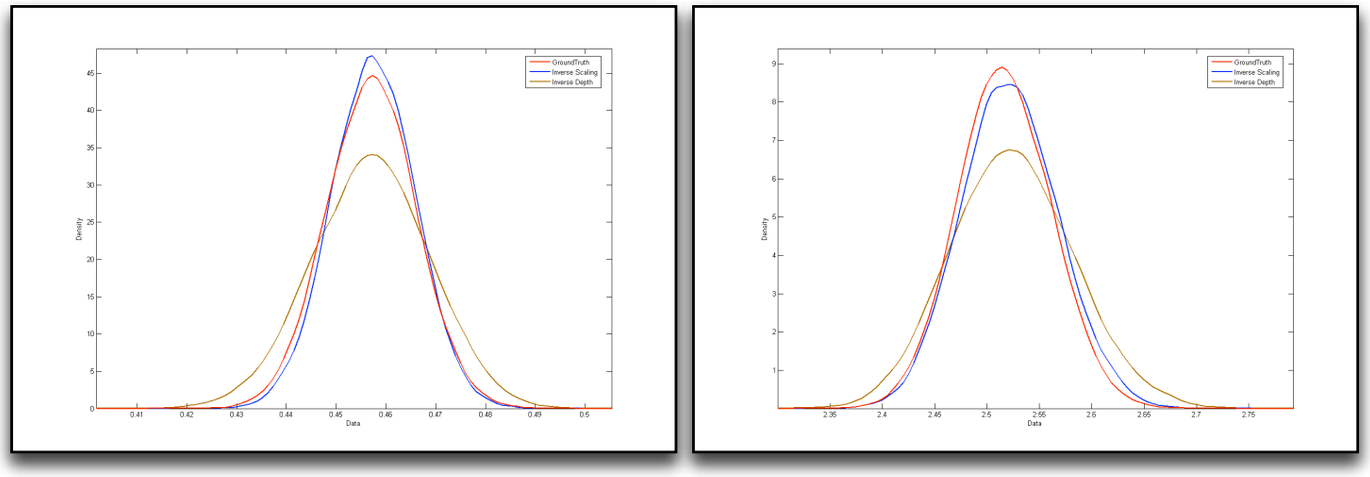




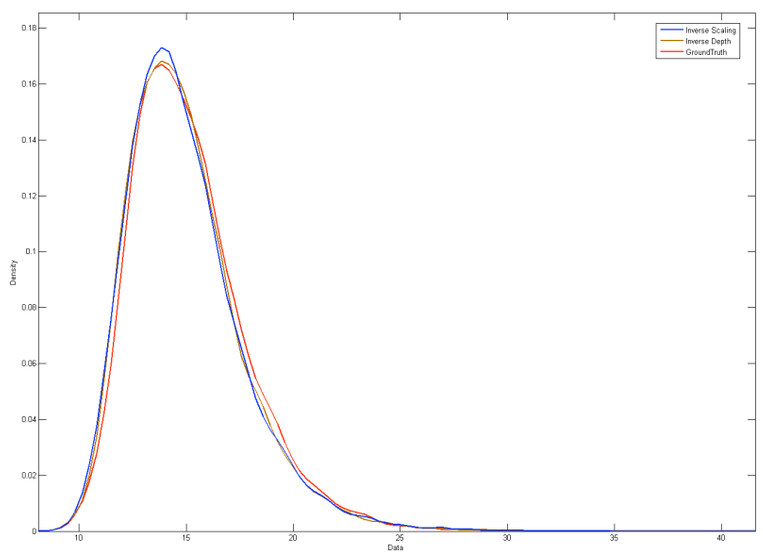
Inverse Scaling Parametrization

- Improvements w.r.t. Unified Inverse Depth:
  - Uncertainty Modeling - Comparison with Inverse Depth

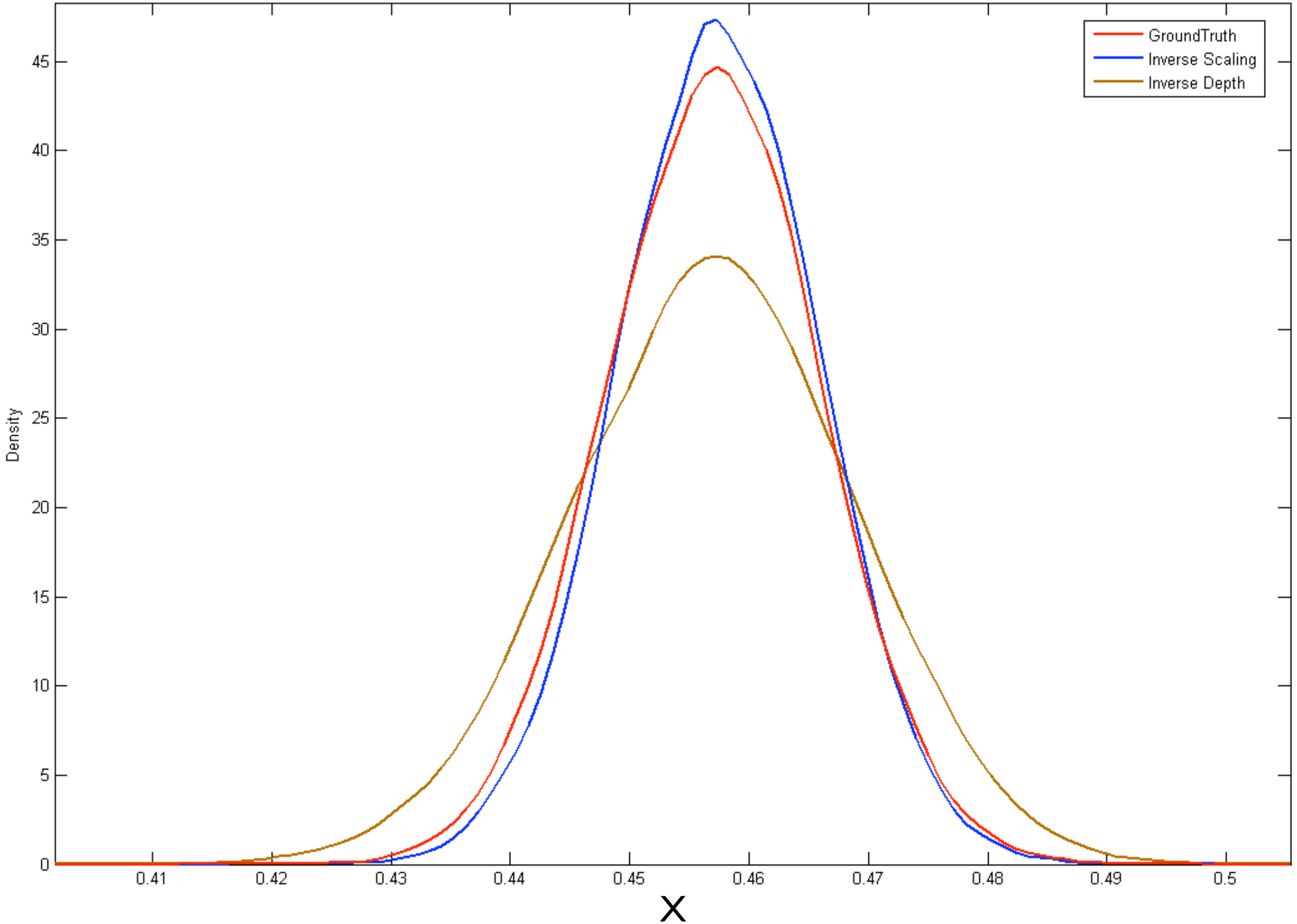
Feature - 2.5m



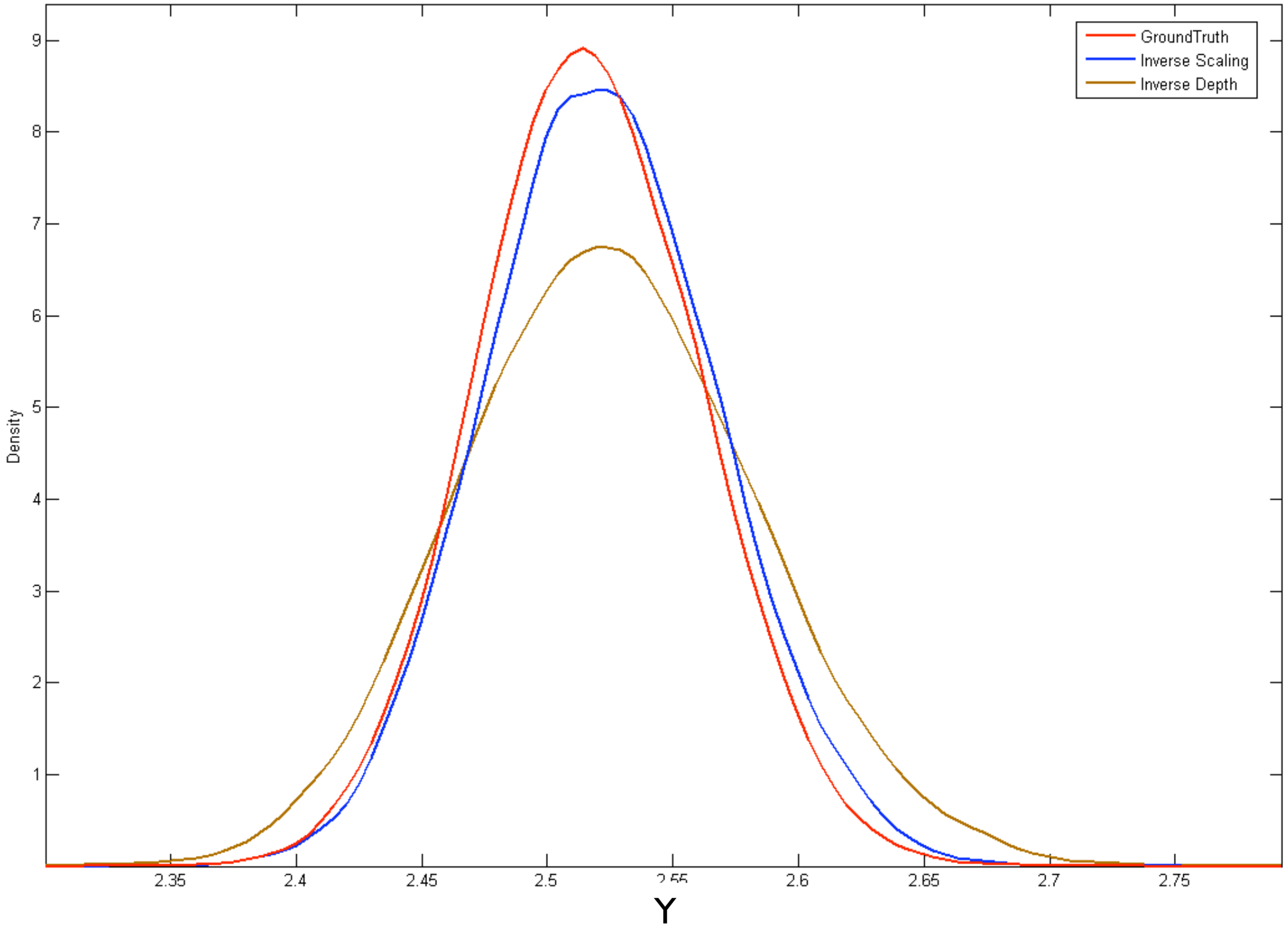
Feature - 15m



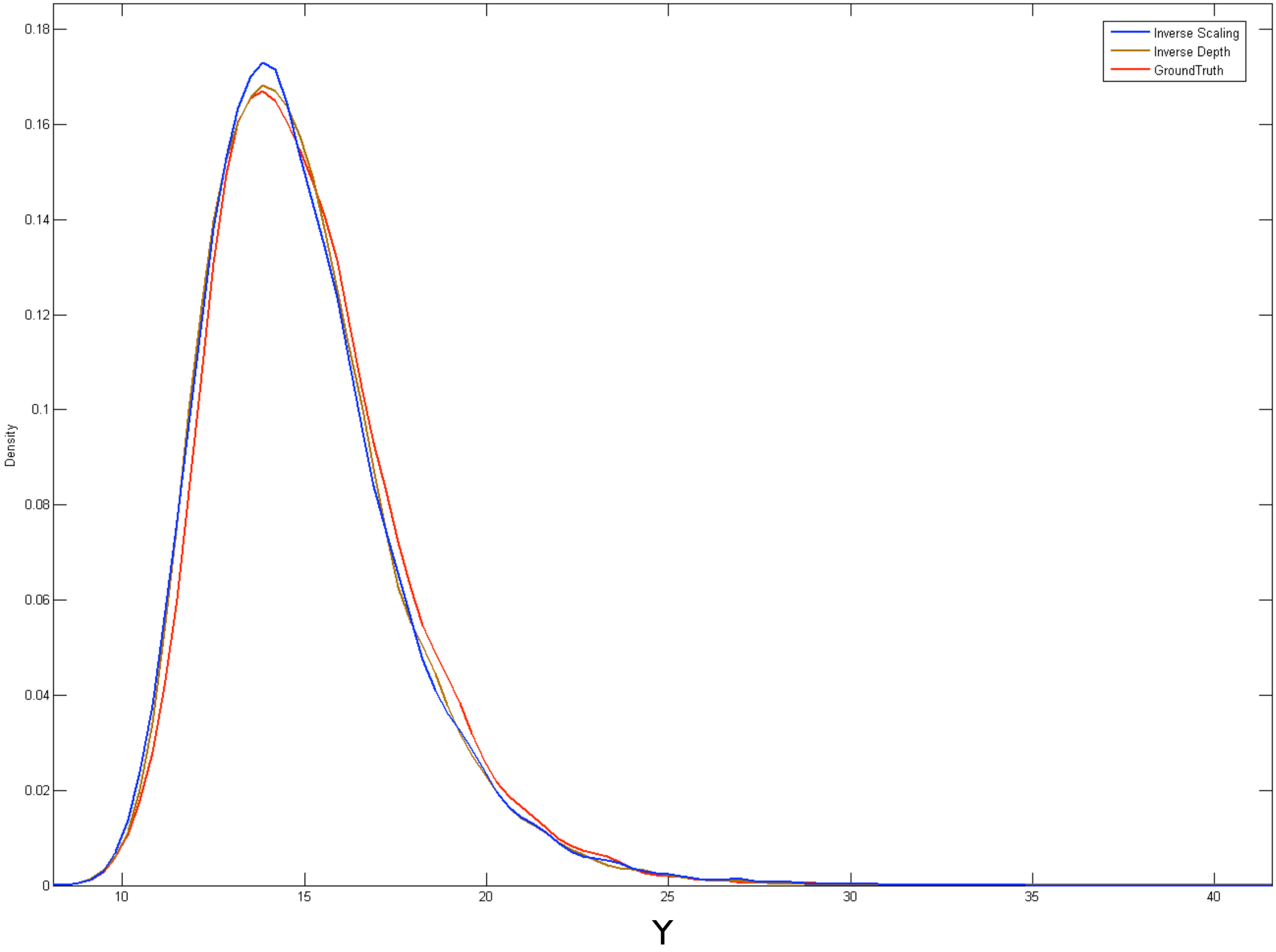
Feature - 2.5m


GroundTruth
Inverse Scaling
Inverse Depth
Density
X
0
5
10
15
20
25
30
35
40
45
0.41
0.42
0.43
0.44
0.45
0.46
0.47
0.48
0.49
0.5

Feature - 2.5m


GroundTruth
Inverse Scaling
Inverse Depth
Density
9
8
7
6
5
4
3
2
1
2.35
2.4
2.45
2.5
2.55
2.6
2.65
2.7
2.75
Y

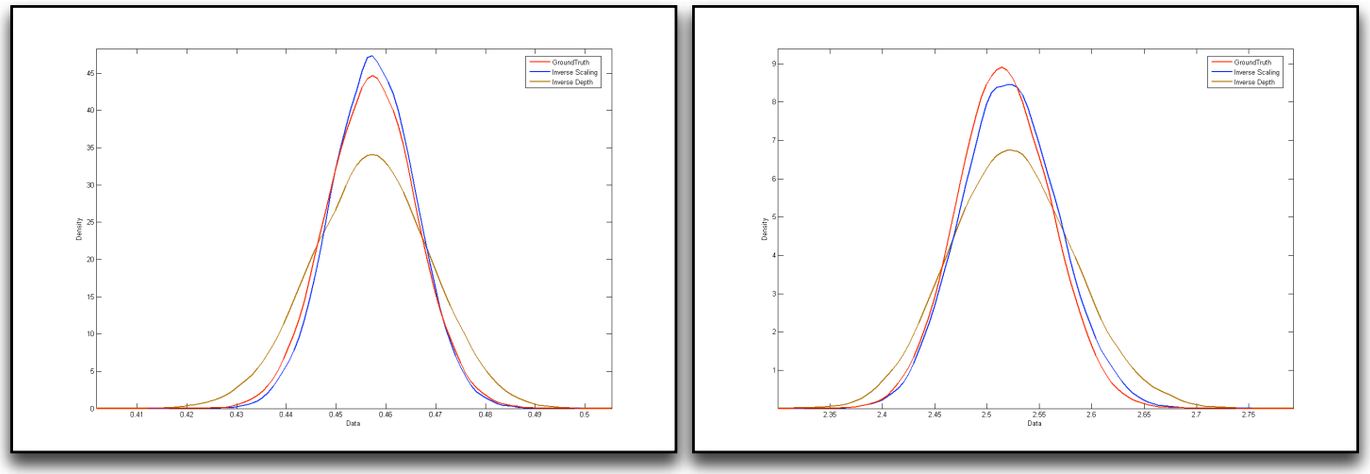
Feature - 15m


Inverse Scaling
Inverse Depth
GroundTruth
Density
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
10
15
20
25
30
35
40
Y

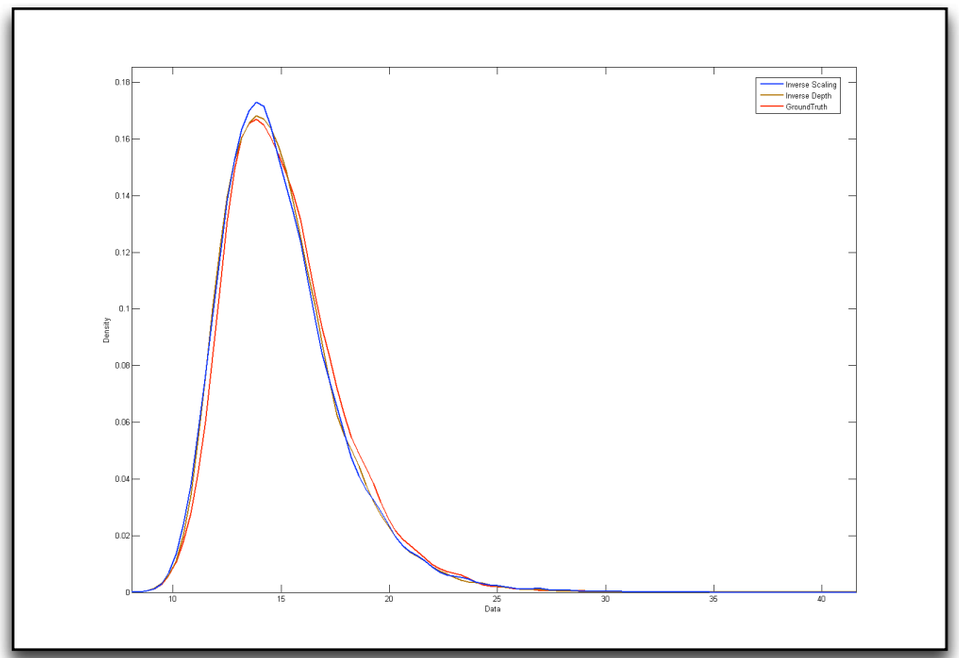
Inverse Scaling Parametrization

- Improvements w.r.t. Unified Inverse Depth:
  - Uncertainty Modeling - Comparison with Inverse Depth

Feature - 2.5m

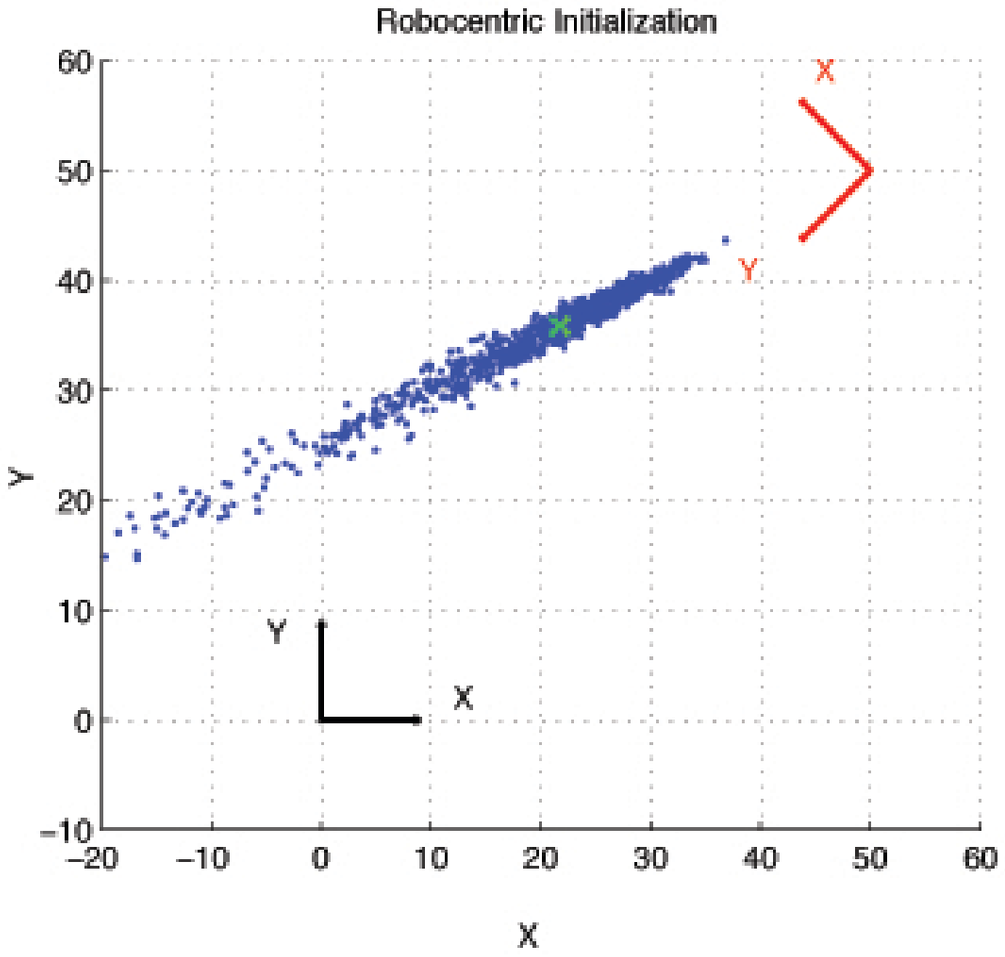


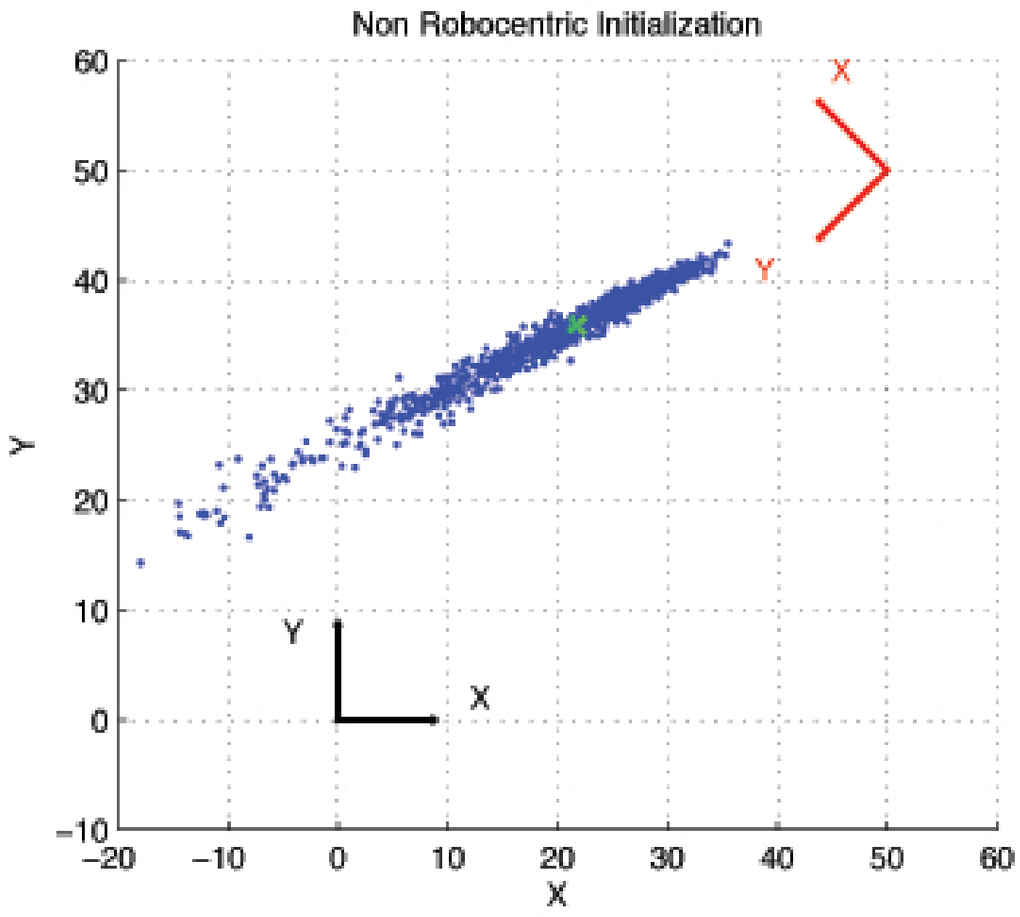
Feature - 15m



Changing Reference Frame

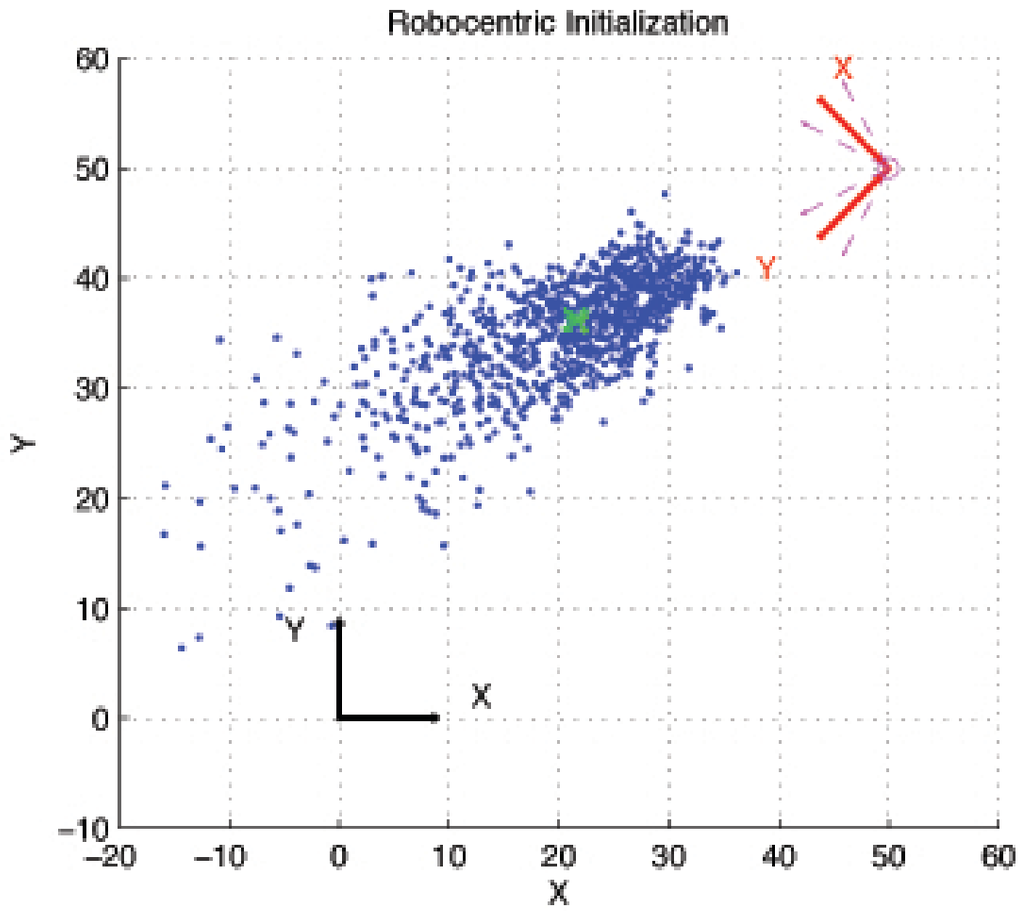
Changing Reference Frame

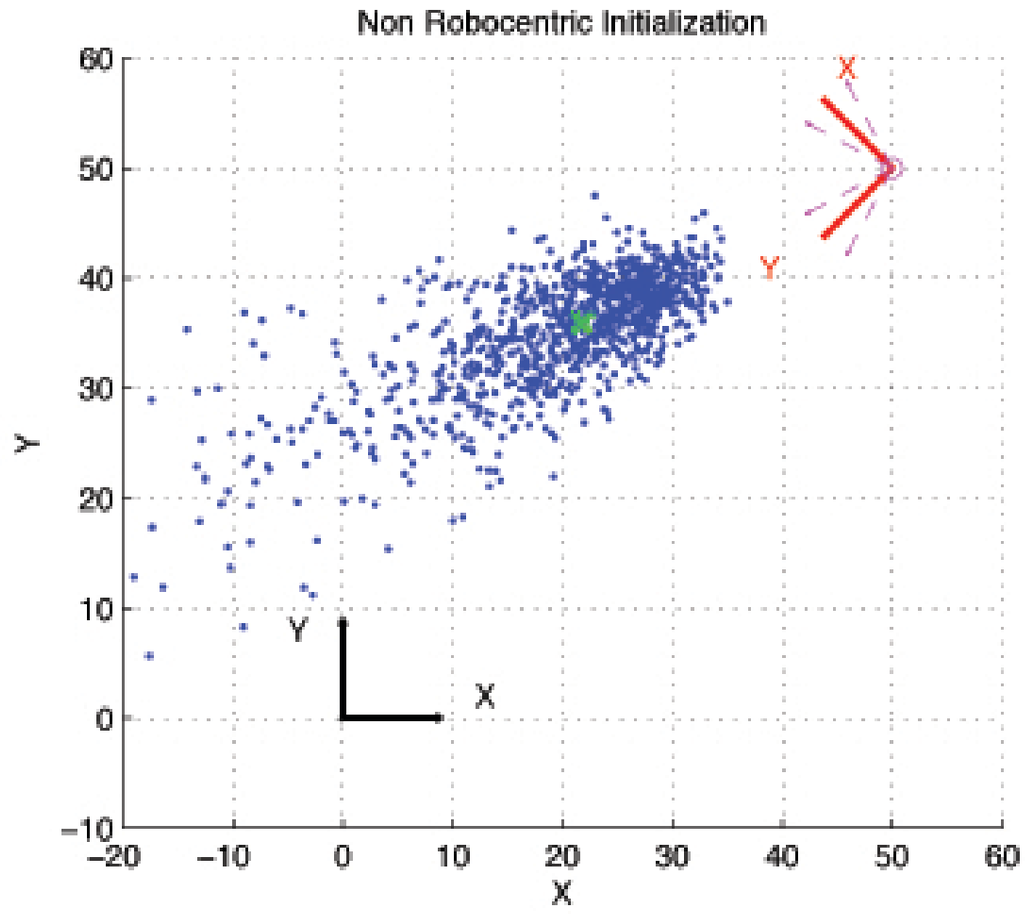




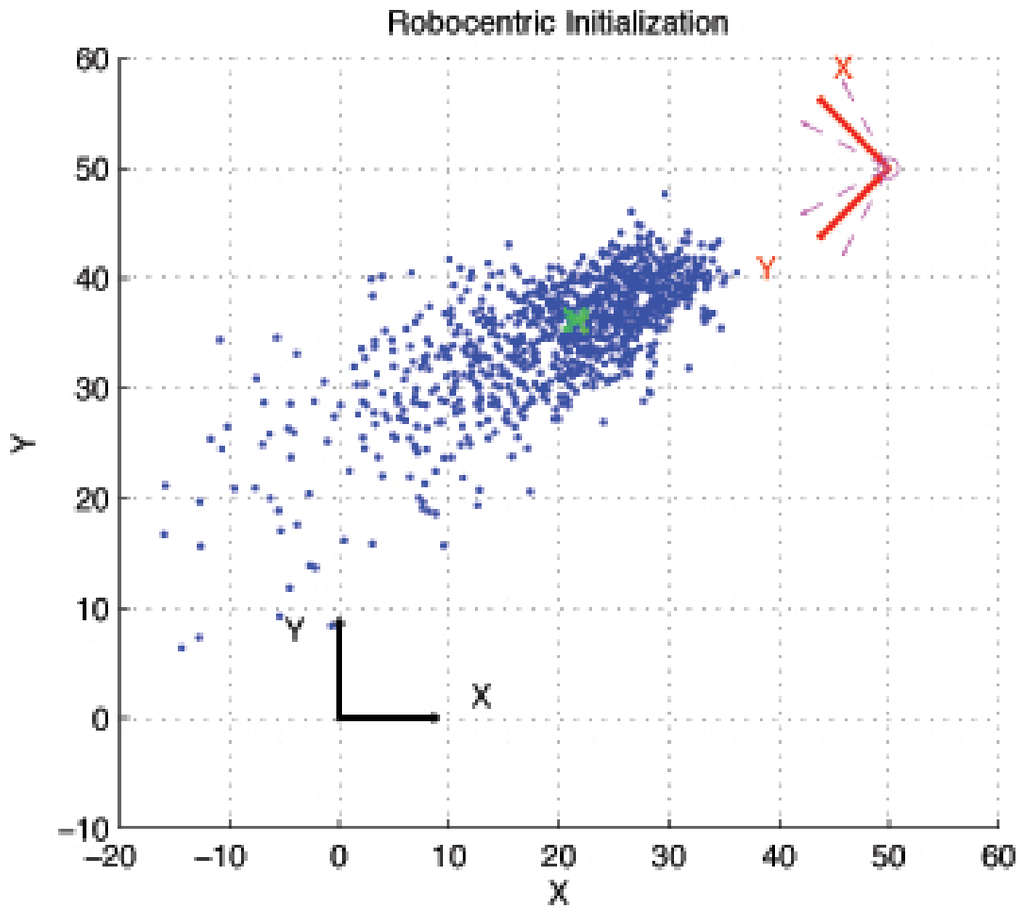
Without uncertainty into rototranslation

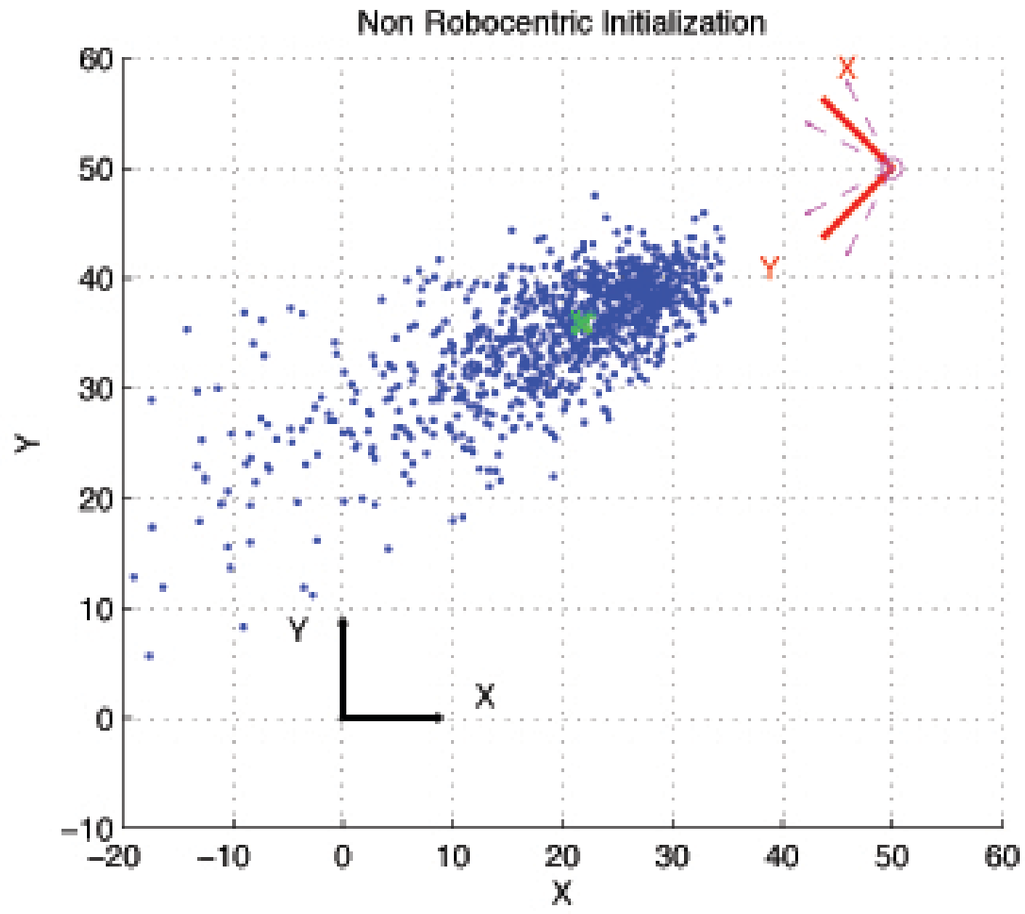
Changing Reference Frame





Changing Reference Frame



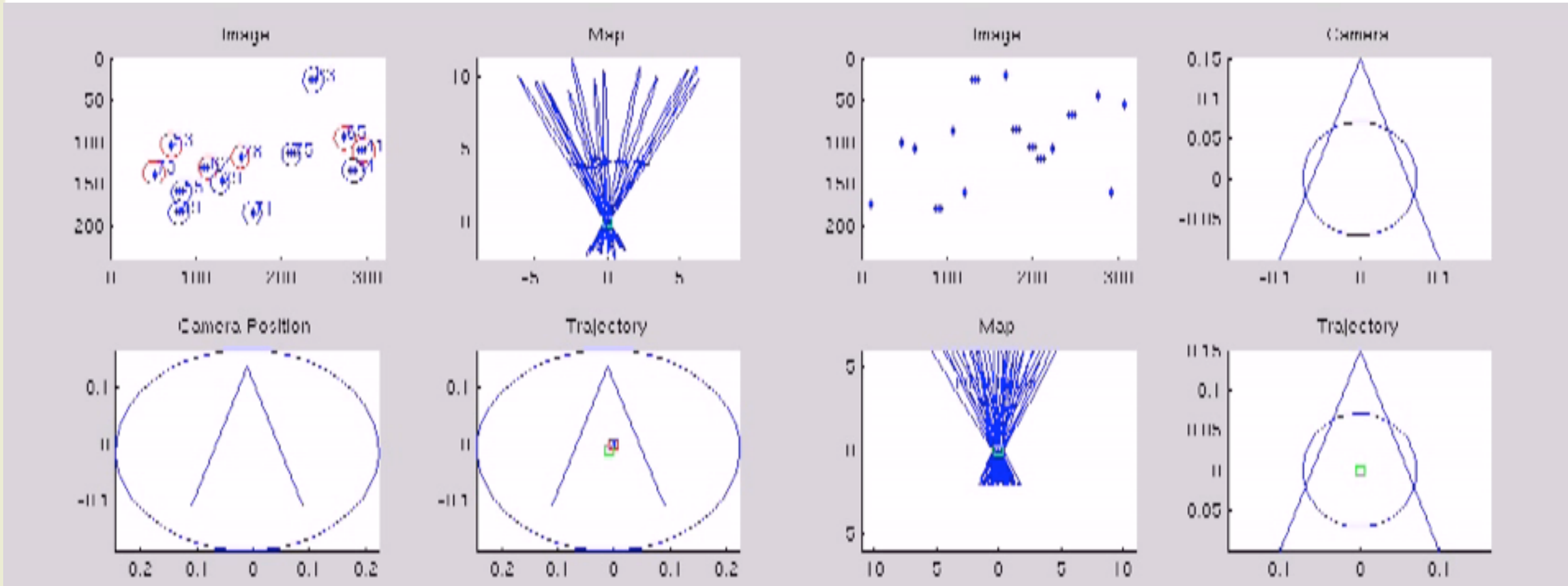


With uncertainty into rototranslation

Experimental Results

Experimental Results

- Simulated Dataset



Experimental Results

Experimental Results

MonoSLAM

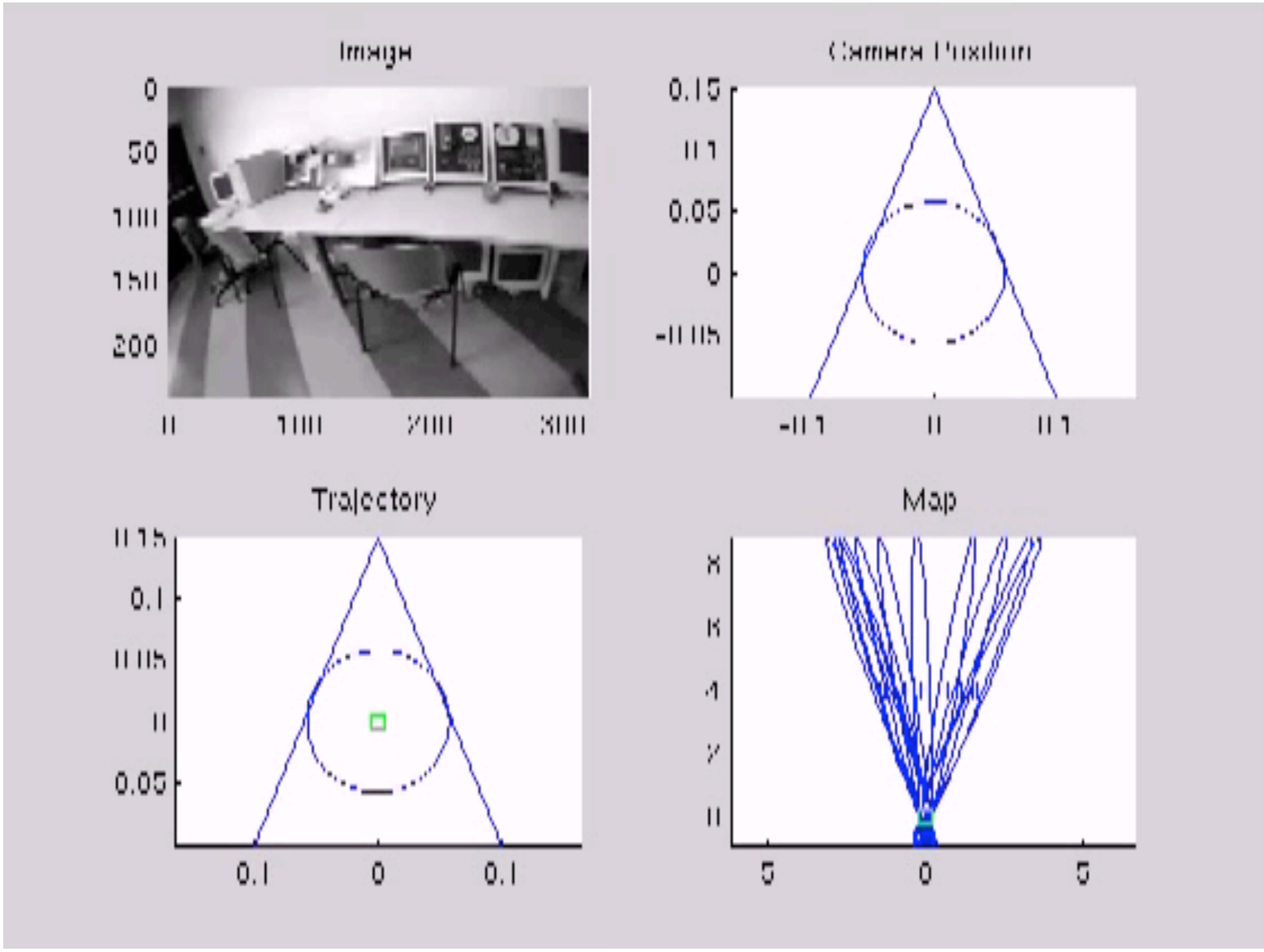
with Inverse Scaling Parametrization

Daniele Marzorati, Matteo Matteucci
Davide Migliore, Domenico G. Sorrenti

ICRA 2009

Experimental Results

Experimental Results



Conclusions

Conclusions

- Presented a new parametrization for Monocular SLAM
  - Proper Uncertainty Modeling for both low and high parallax
  - More linear than Unified Inverse Depth
  - Only 4 parameters required
  - Undelayed initialization

Conclusions

- Presented a new parametrization for Monocular SLAM
  - Proper Uncertainty Modeling for both low and high parallax
  - More linear than Unified Inverse Depth
  - Only 4 parameters required
  - Undelayed initialization

- Ongoing works:
  - MonoSLAM with Bearing Only Tracking
  - Integration on Large Maps
  - Self calibration

Conclusions

- Presented a new parametrization for Monocular SLAM
  - Proper Uncertainty Modeling for both low and high parallax
  - More linear than Unified Inverse Depth
  - Only 4 parameters required
  - Undelayed initialization

- Ongoing works:
  - MonoSLAM with Bearing Only Tracking
  - Integration on Large Maps
  - Self calibration

Thanks! Any question?