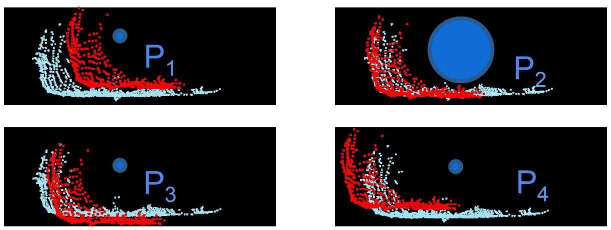
Combining 3D Shape, Color, and Motion for Robust Anytime Tracking

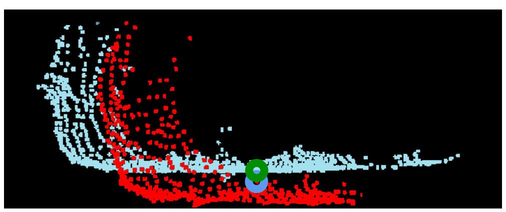
David Held, Jesse Levinson, Sebastian Thrun, and Silvio Savarese

Stanford University

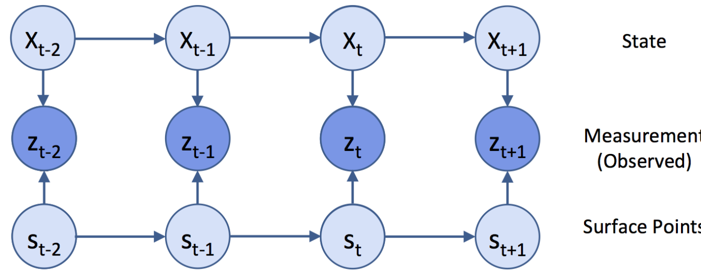
Goal: Fast and Robust Velocity Estimation

Our Approach: Alignment Probability



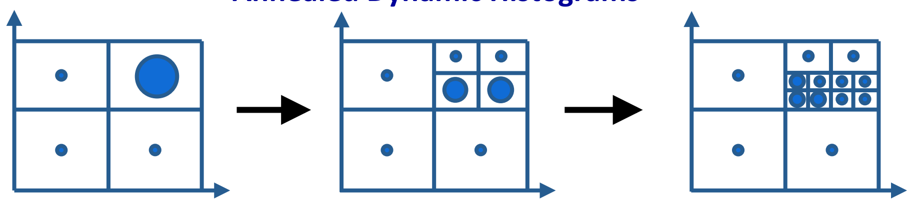


- Spatial Distance
- Color Distance (if available)
- Probability of Occlusion



$$p(x_t \mid z_1 \ldots z_t) = \eta \, p(z_t \mid x_t, z_{t-1}) p(x_t \mid z_1 \ldots z_{t-1})$$

Annealed Dynamic Histograms



Raw Points **ADH Tracker** **ICP Baseline**

Combining 3D Shape, Color, and Motion for Robust Anytime Tracking

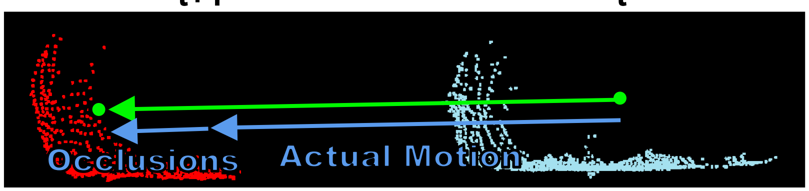
David Held, Jesse Levinson, Sebastian Thrun, and Silvio Savarese

Stanford University

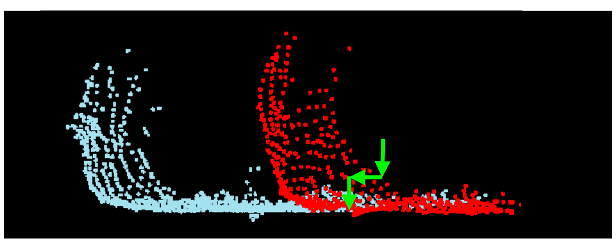
Goal: Fast and Robust Velocity Estimation

Baseline: Centroid Kalman Filter

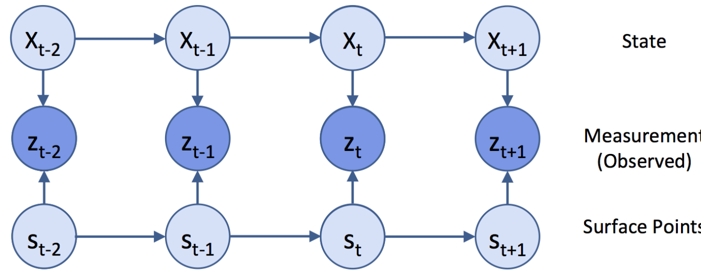
t+1 **t**



Baseline: ICP

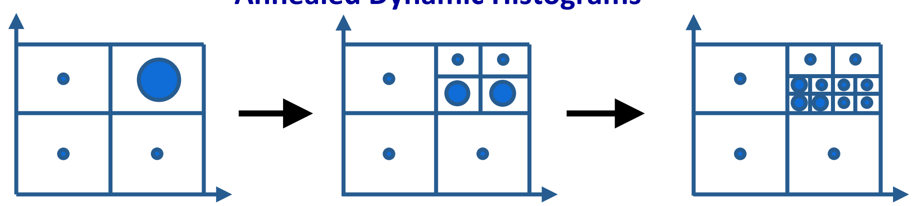


Local Search ⟶ Poor Local Optimum!



$$p(x_t \mid z_1 \ldots z_t) = \eta \, p(z_t \mid x_t, z_{t-1}) p(x_t \mid z_1 \ldots z_{t-1})$$

Annealed Dynamic Histograms



Raw Points **ADH Tracker** **ICP Baseline**

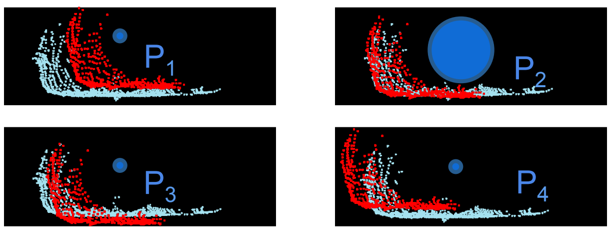
Combining 3D Shape, Color, and Motion for Robust Anytime Tracking

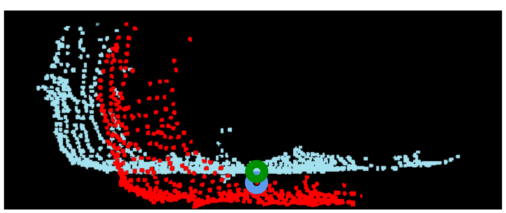
David Held, Jesse Levinson, Sebastian Thrun, and Silvio Savarese

Stanford University

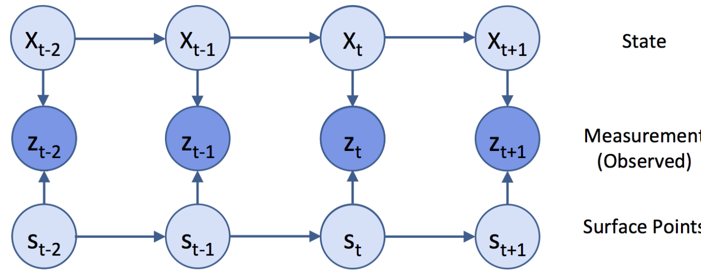
Goal: Fast and Robust Velocity Estimation

Our Approach: Alignment Probability



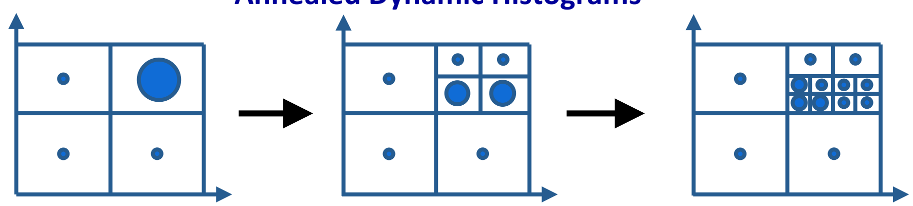


- Spatial Distance
- Color Distance (if available)
- Probability of Occlusion



$$p(x_t \mid z_1 \ldots z_t) = \eta \, p(z_t \mid x_t, z_{t-1}) p(x_t \mid z_1 \ldots z_{t-1})$$

Annealed Dynamic Histograms



Raw Points **ADH Tracker** **ICP Baseline**

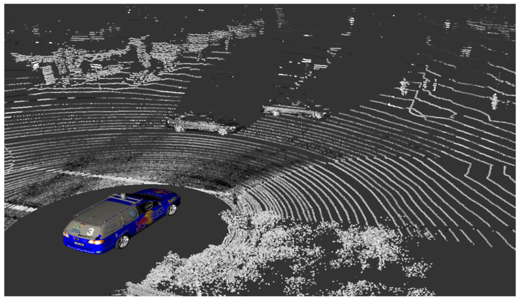
Motivation

Quickly and robustly estimate the speed of nearby objects

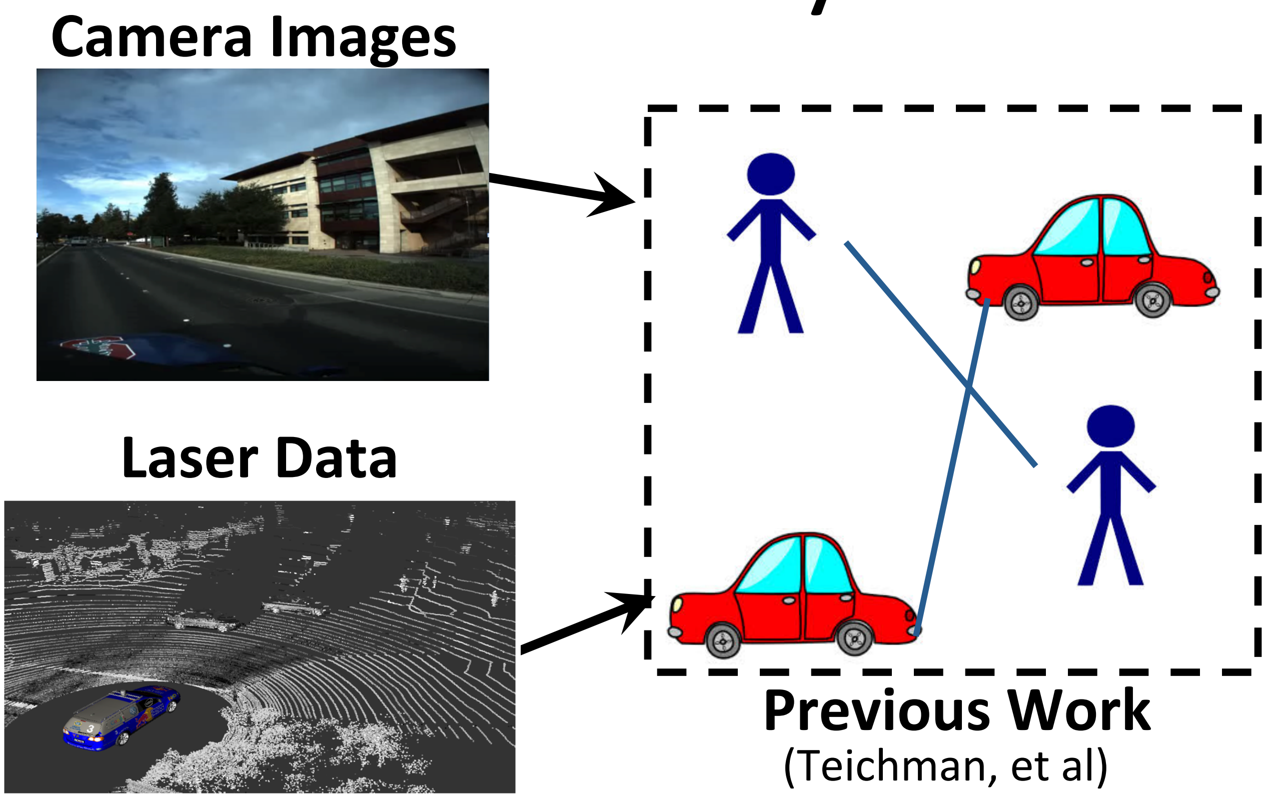
System

Camera Images

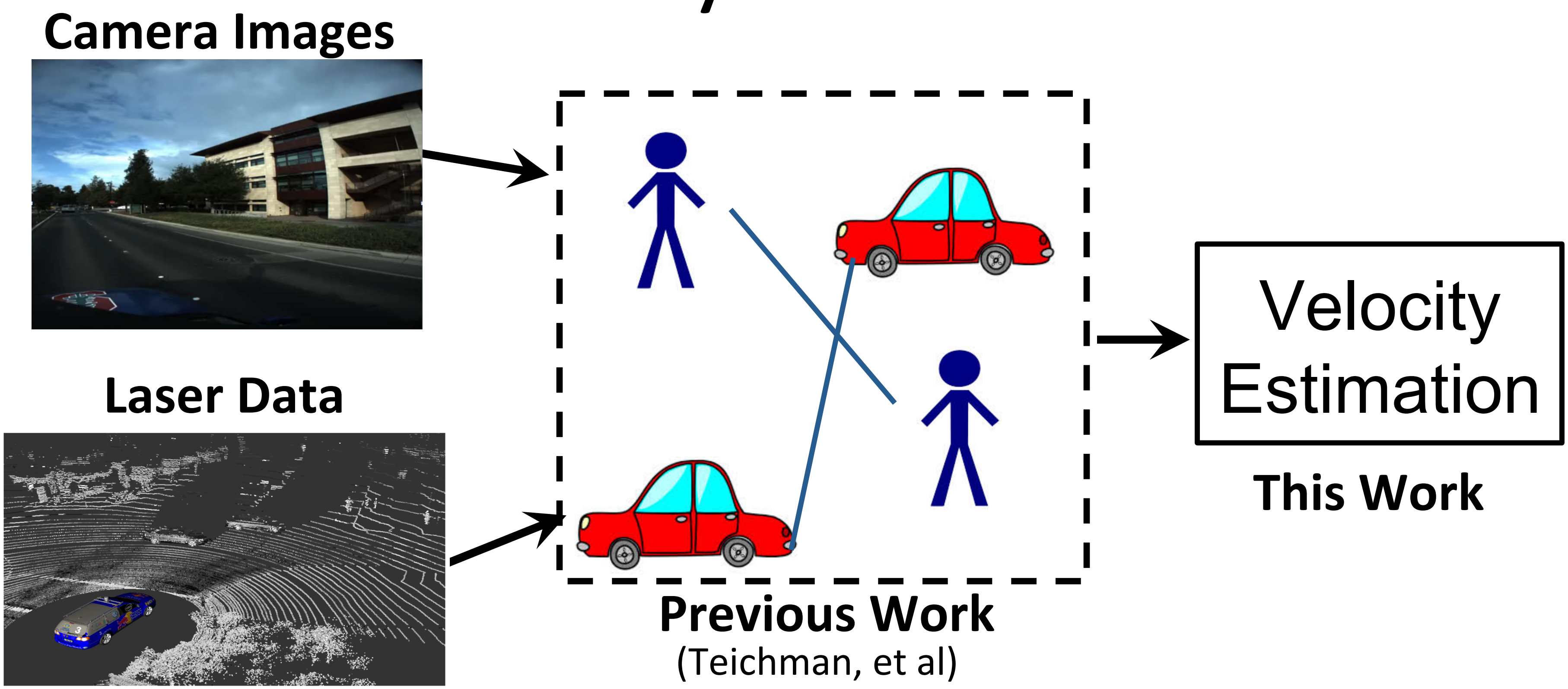
Laser Data



System

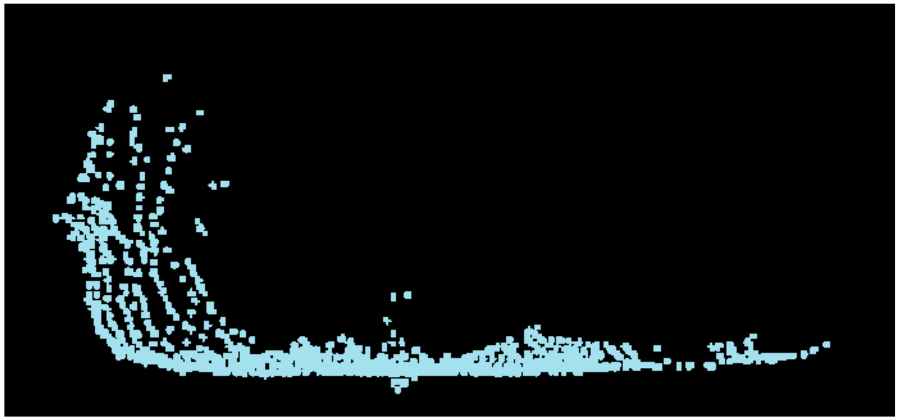


System


Camera Images
Laser Data
Previous Work
(Teichman, et al)
Velocity Estimation
This Work

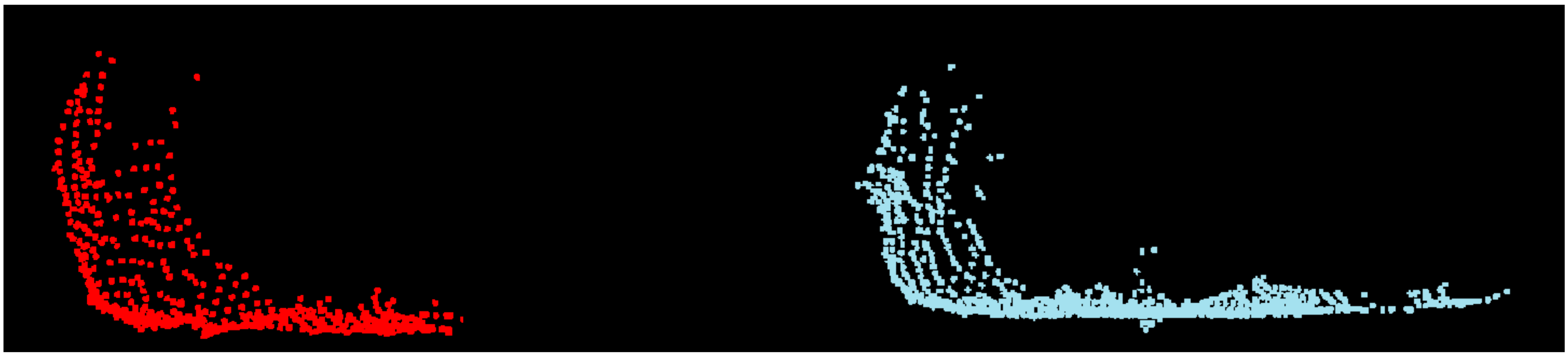
Velocity Estimation

t



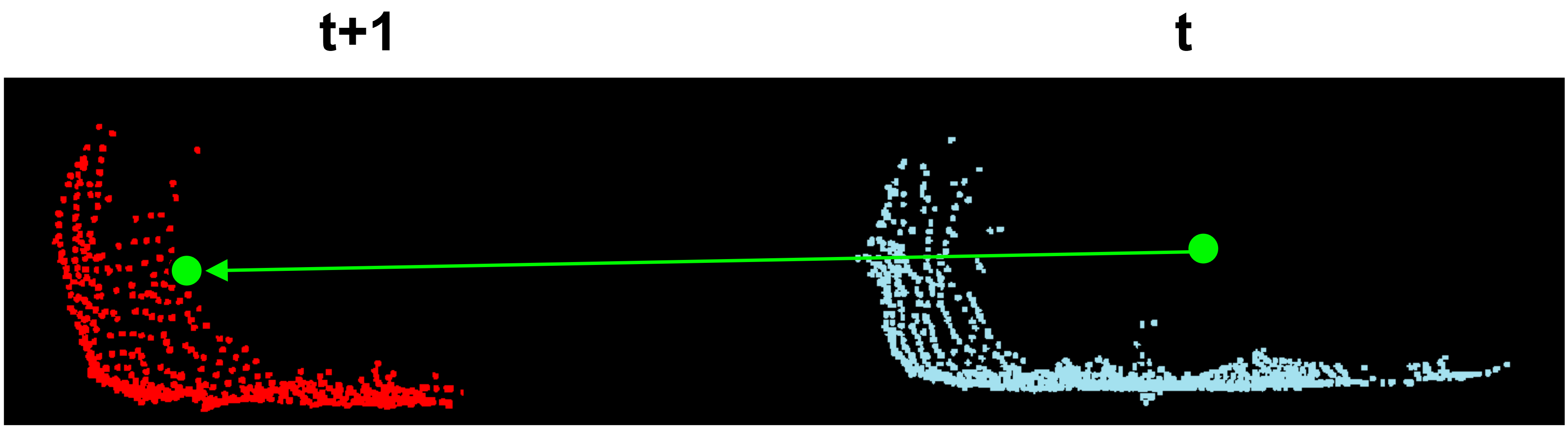
Velocity Estimation

t+1 **t**

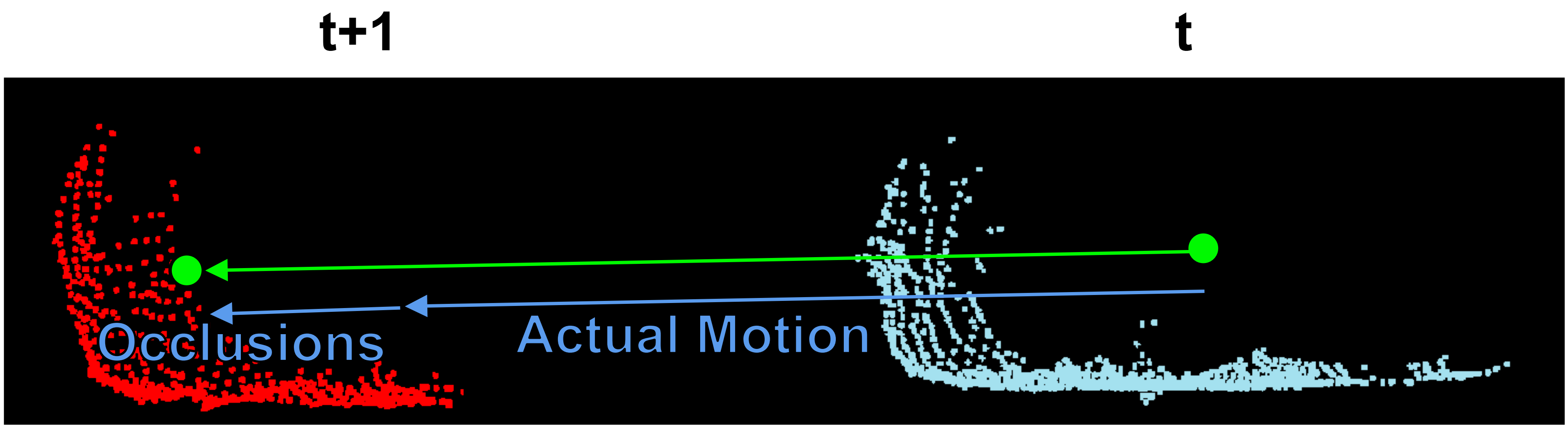


Velocity Estimation

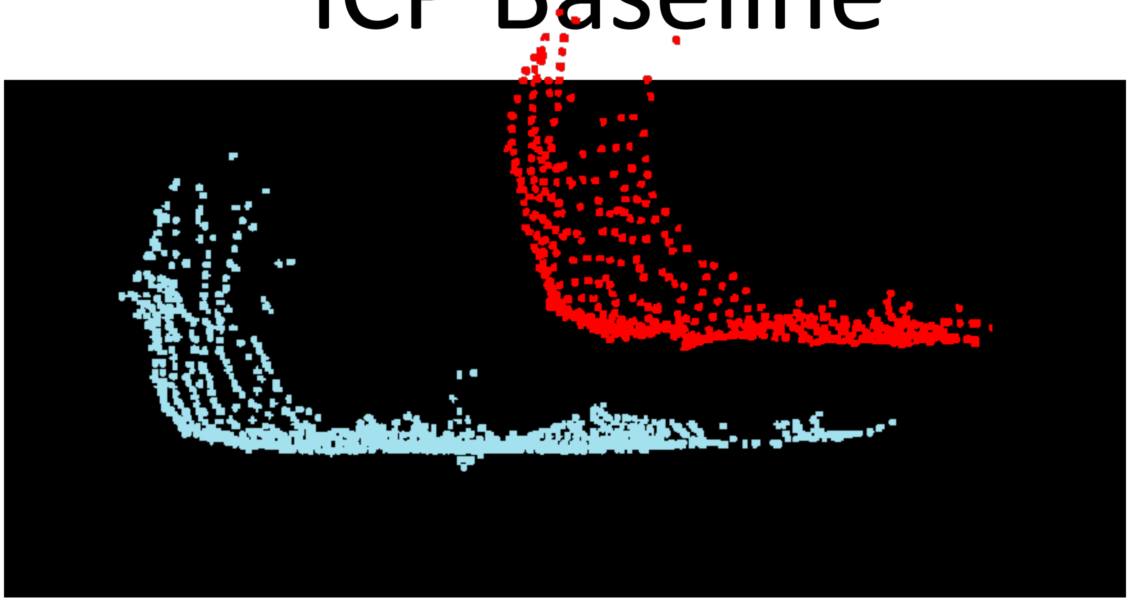
Velocity Estimation



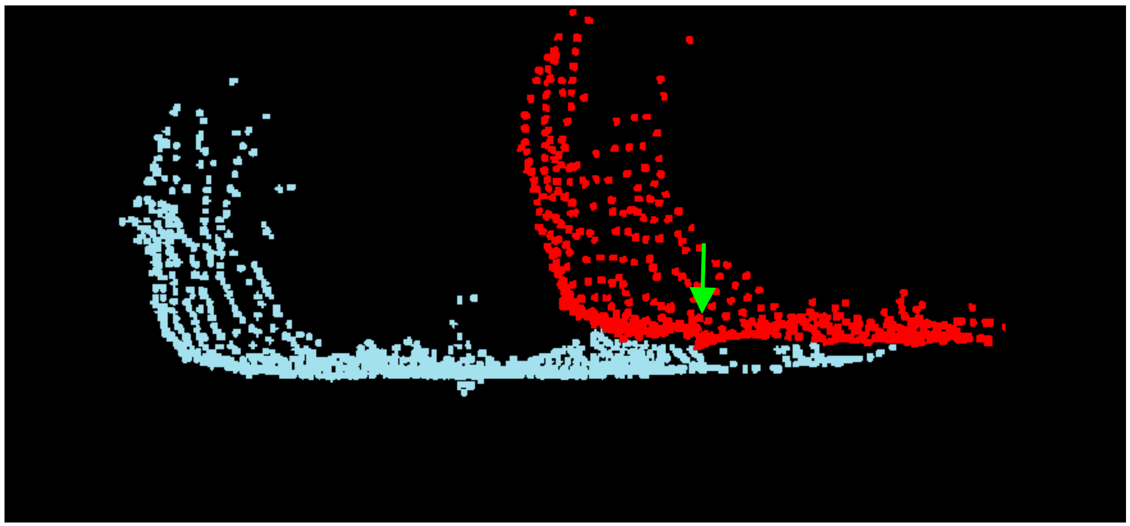
Velocity Estimation



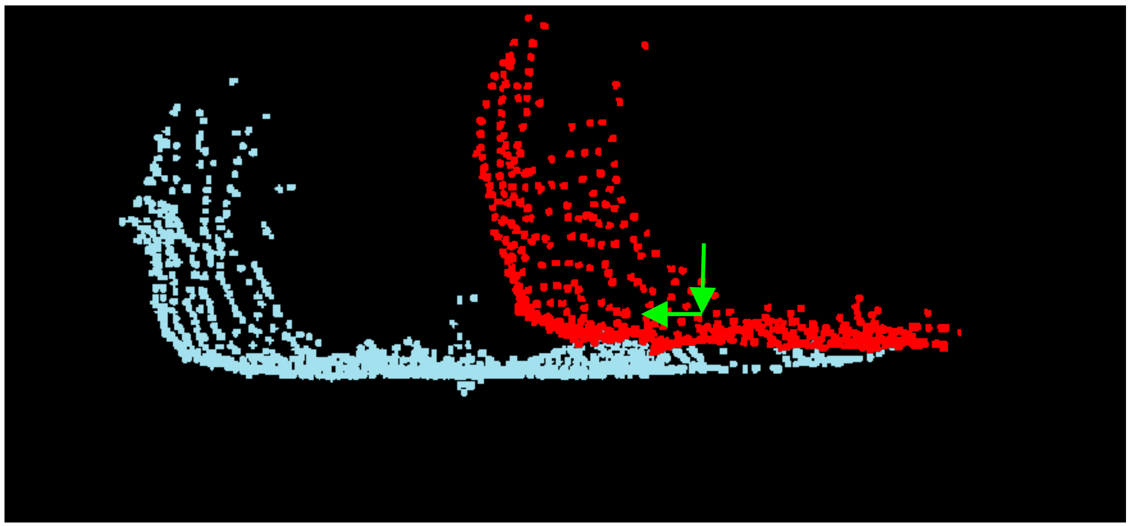
ICP Baseline



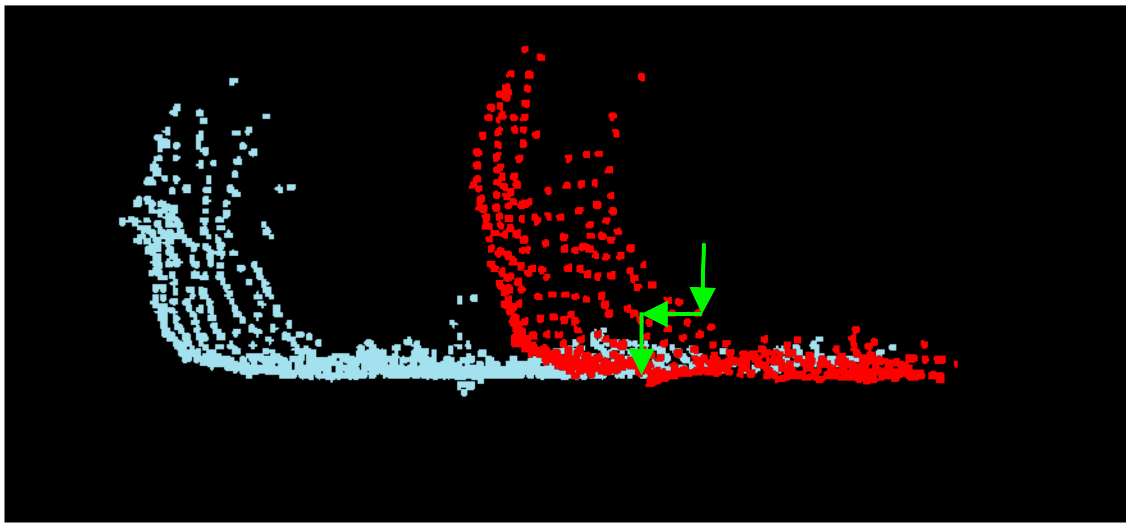
ICP Baseline



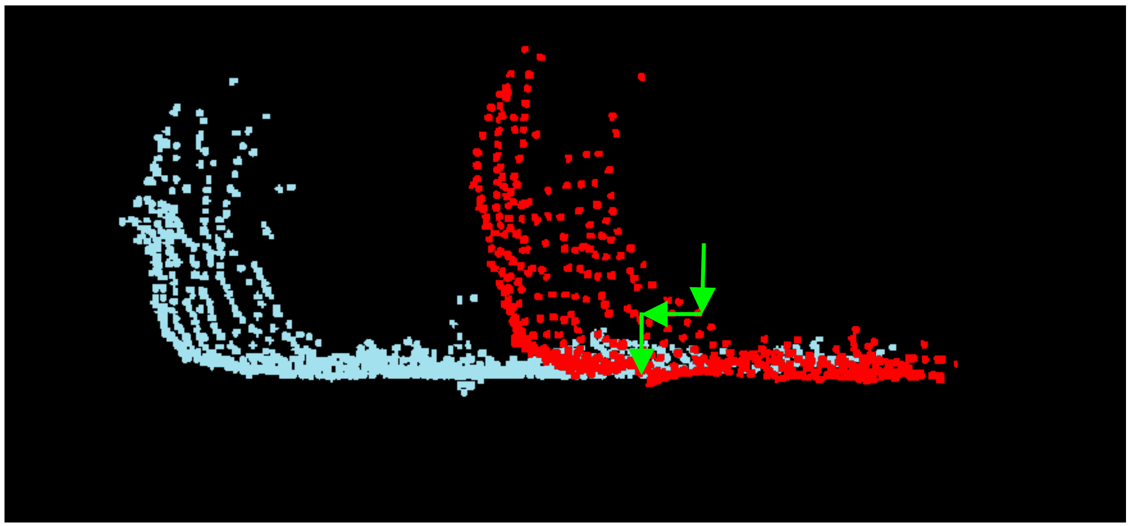
ICP Baseline



ICP Baseline

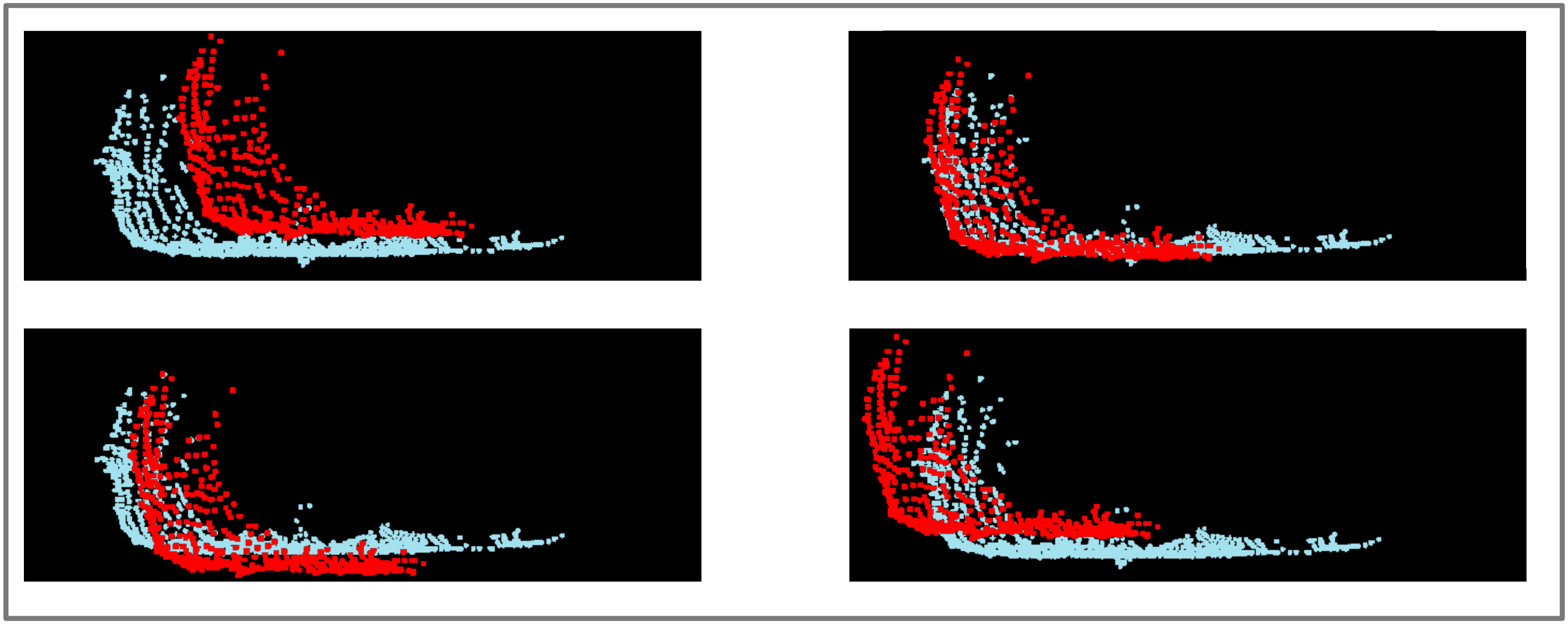


ICP Baseline



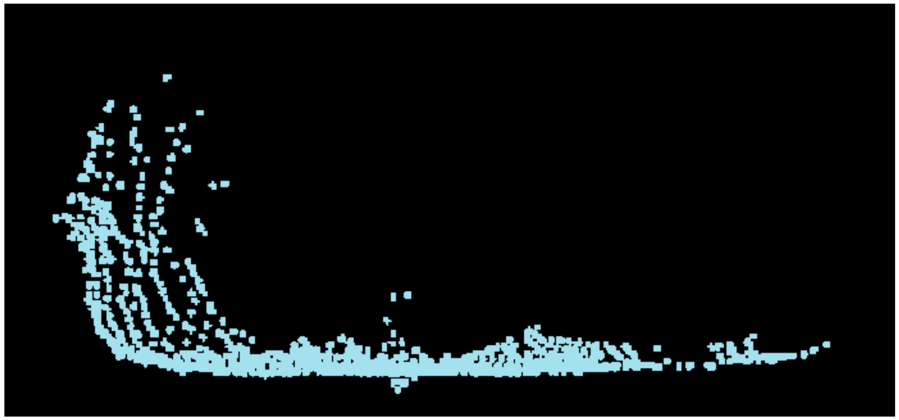
Local Search ⟶ Poor Local Optimum!

Tracking Probability



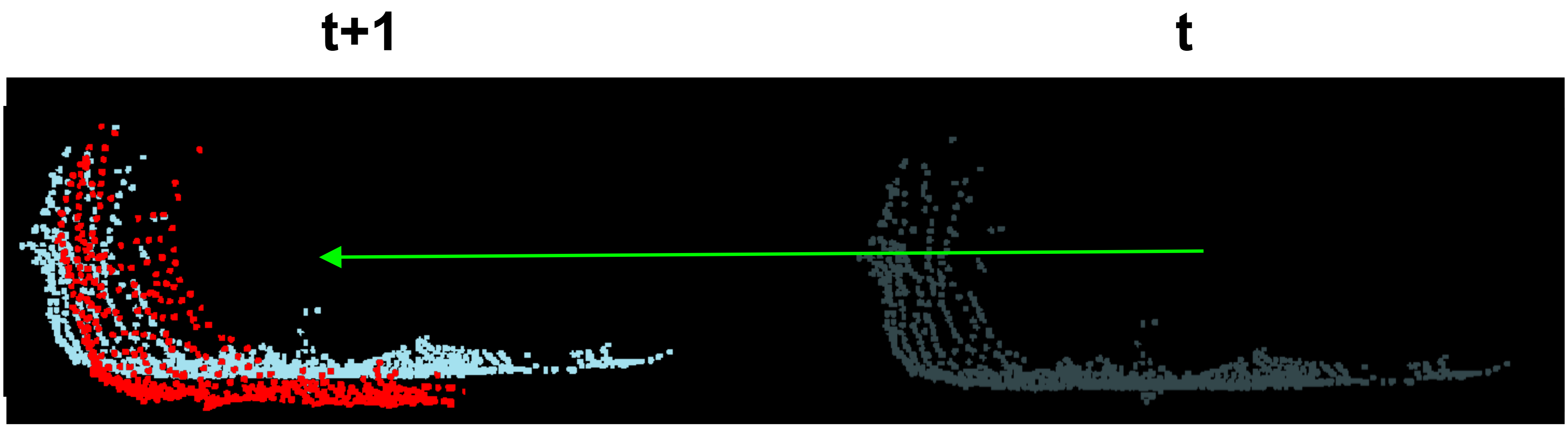
Velocity Estimation

t

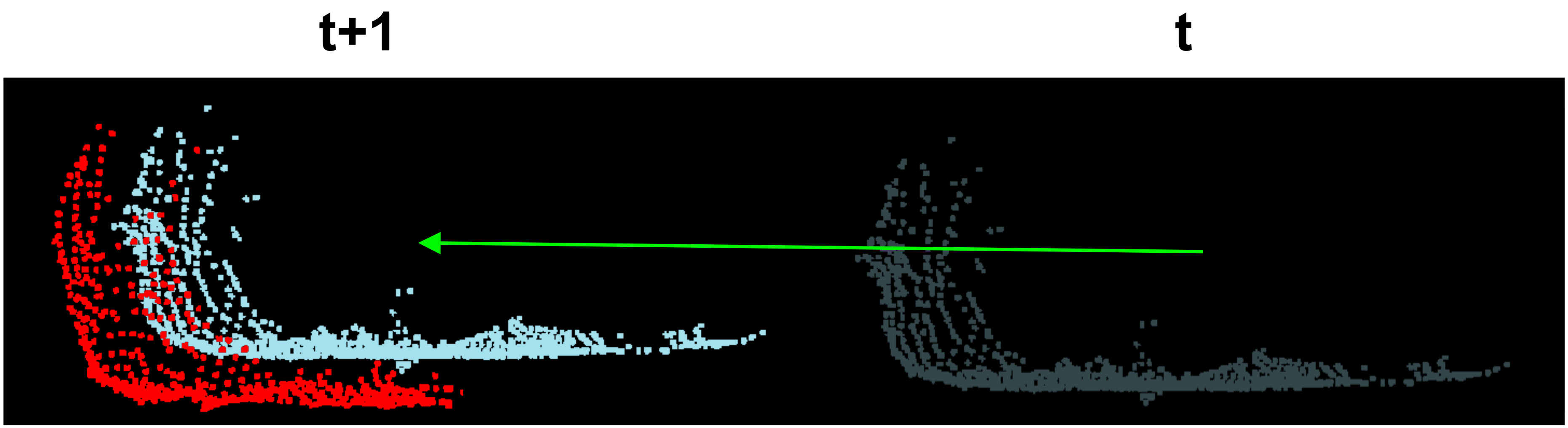


Velocity Estimation

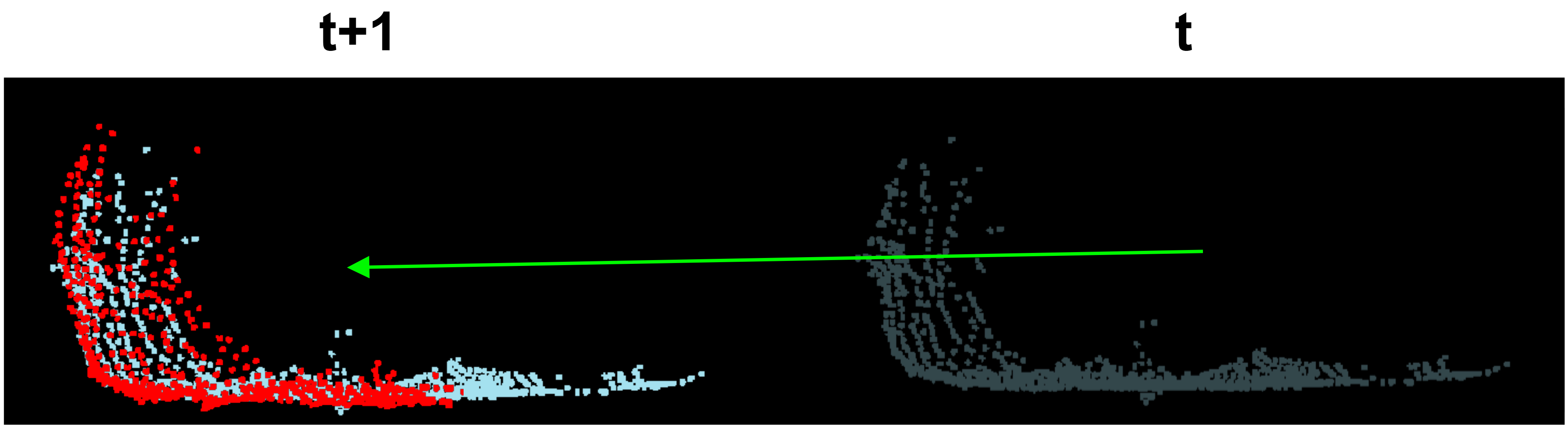
Velocity Estimation



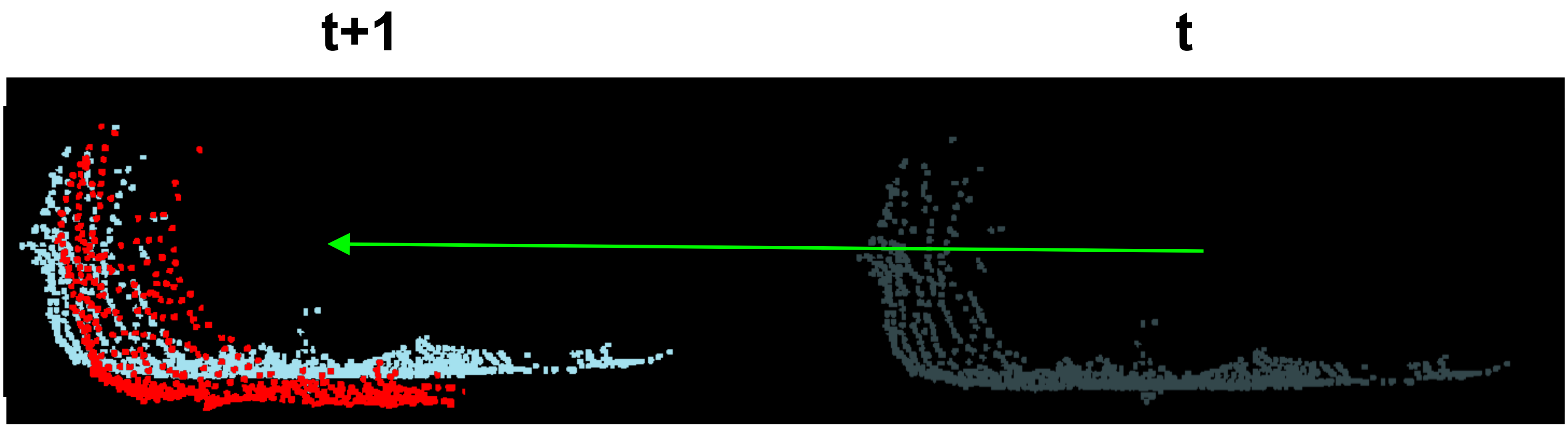
Velocity Estimation



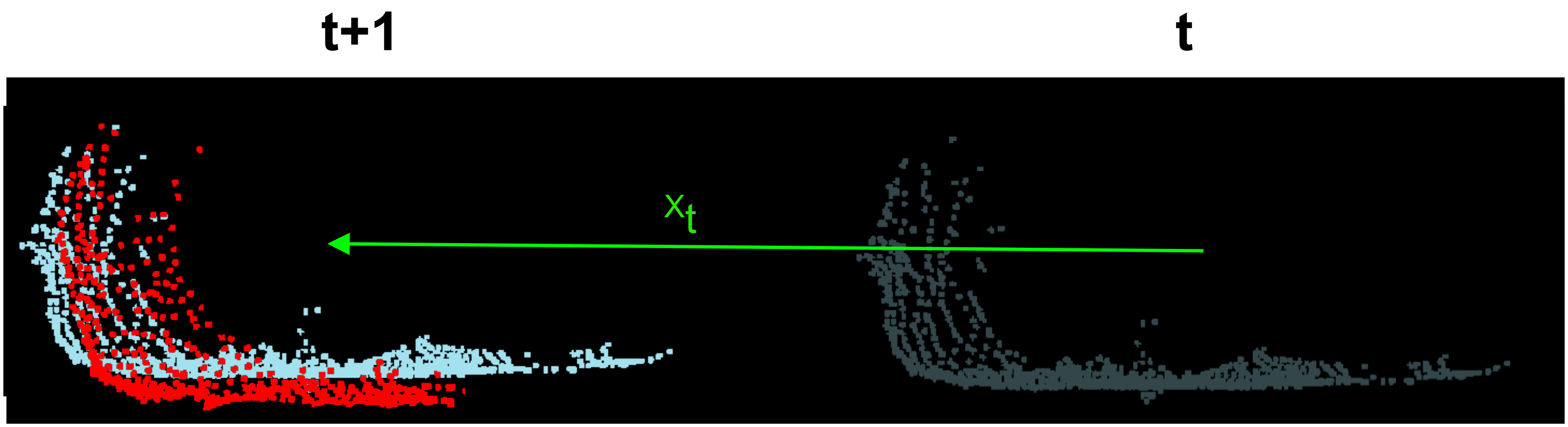
Velocity Estimation



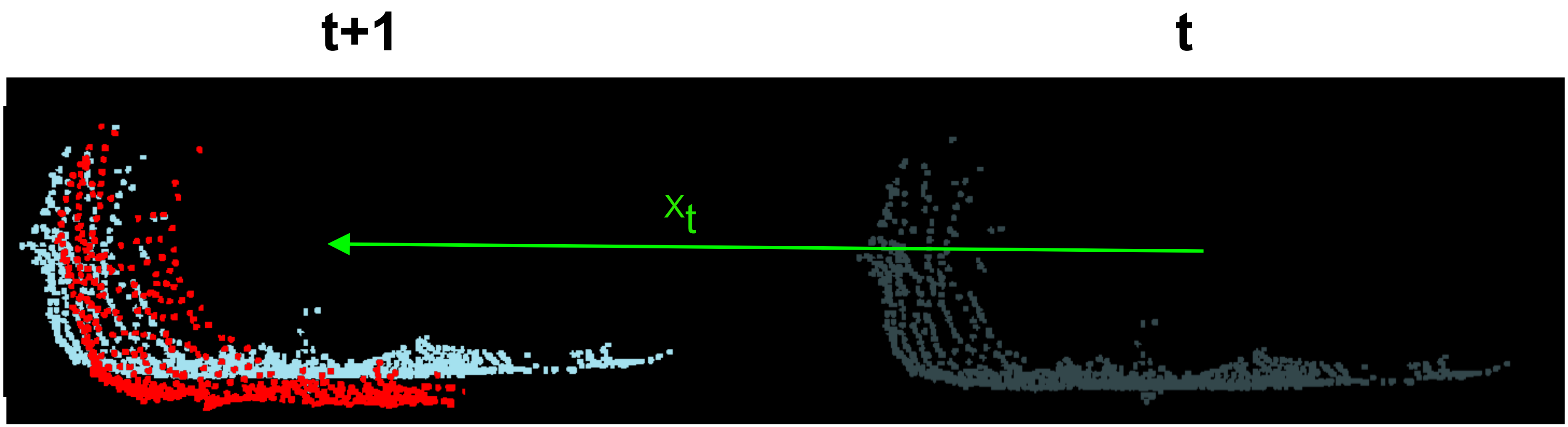
Velocity Estimation



Velocity Estimation



Velocity Estimation



$$p(x_t \mid z_1 \ldots z_t)$$

Tracking Probability

$$p(x_t \mid z_1 \dots z_t) = \eta \, p(z_t \mid x_t, z_{t-1}) \, p(x_t \mid z_1 \dots z_{t-1})$$

Measurement Model

Motion Model

Tracking Probability

$$p(x_t \mid z_1 \ldots z_t) = \eta \, p(z_t \mid x_t, z_{t-1}) \boxed{p(x_t \mid z_1 \ldots z_{t-1})}$$

Measurement Model

Motion Model

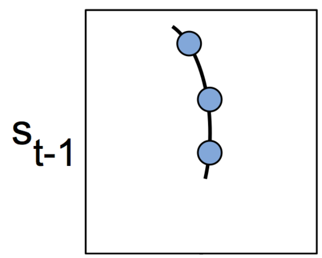
Constant velocity Kalman filter

Tracking Probability

$$p(x_t \mid z_1 \ldots z_t) = \eta\, p(z_t \mid x_t, z_{t-1})\, p(x_t \mid z_1 \ldots z_{t-1})$$

Measurement Model

Motion Model

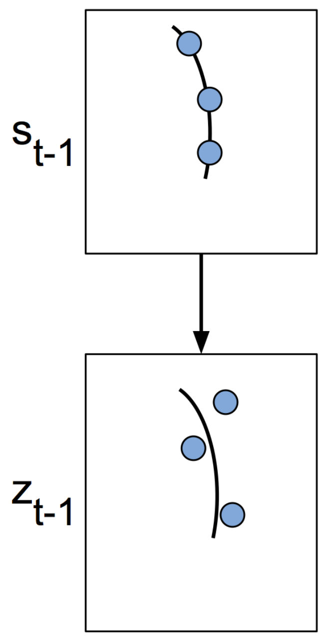


Tracking Probability

$$p(x_t \mid z_1 \ldots z_t) = \boxed{\eta\, p(z_t \mid x_t, z_{t-1})}\, p(x_t \mid z_1 \ldots z_{t-1})$$

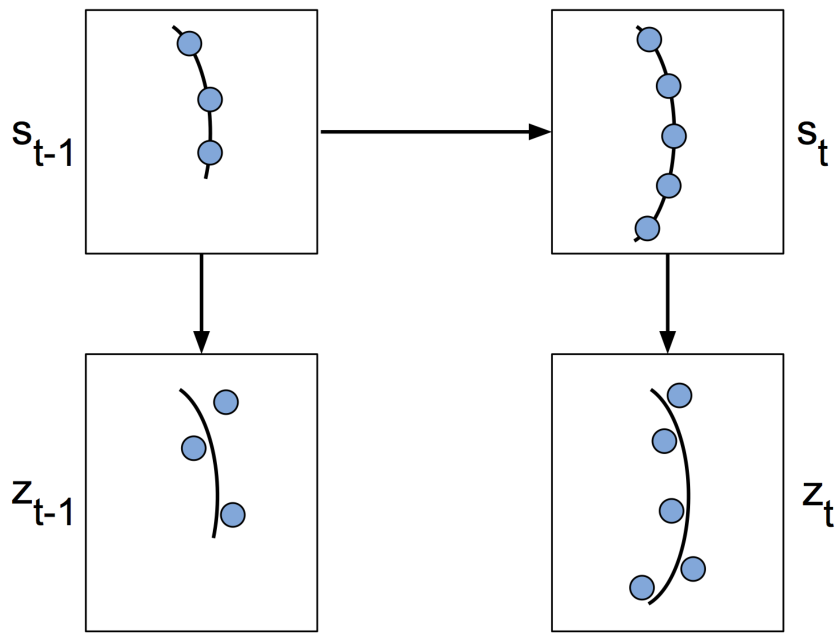
Measurement Model

Motion Model



Tracking Probability

$$p(x_t \mid z_1 \ldots z_t) = \boxed{\underbrace{\eta\, p(z_t \mid x_t, z_{t-1})}_{\text{Measurement Model}}}\, \underbrace{p(x_t \mid z_1 \ldots z_{t-1})}_{\text{Motion Model}}$$

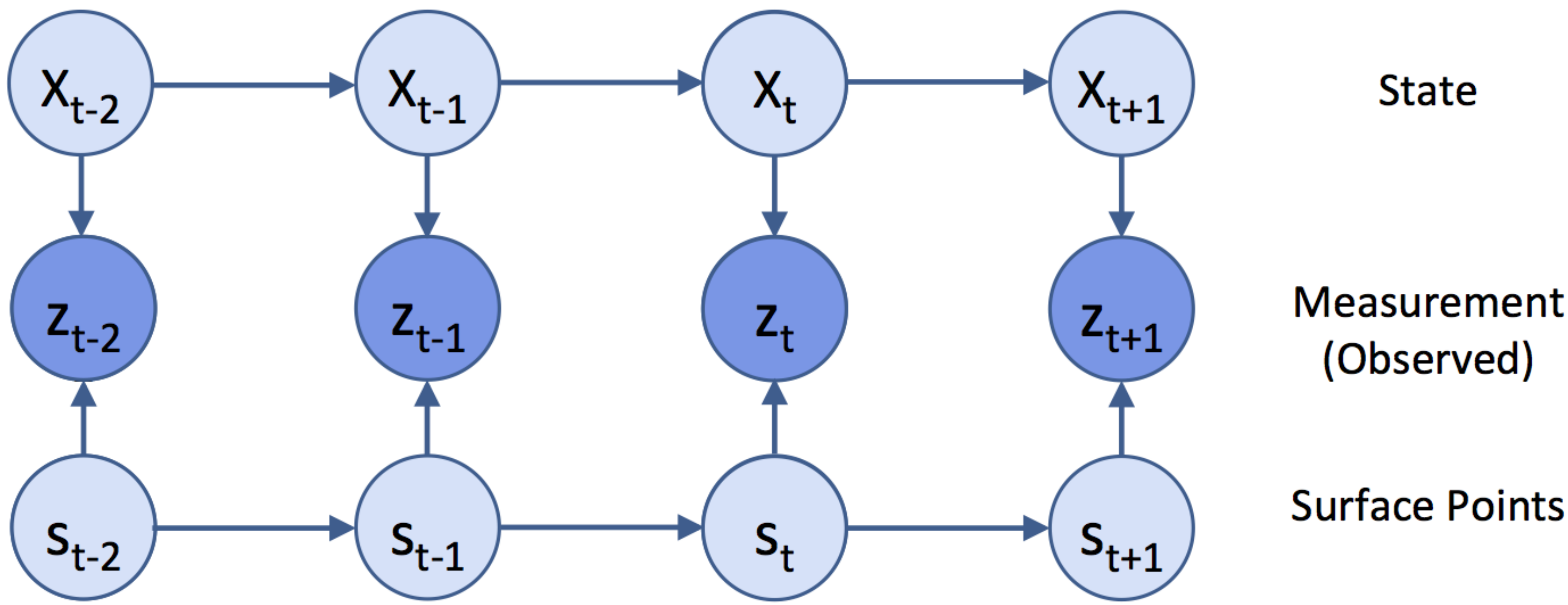


Tracking Probability

$$p(x_t \mid z_1 \ldots z_t) = \boxed{\eta\, p(z_t \mid x_t, z_{t-1})}\, p(x_t \mid z_1 \ldots z_{t-1})$$

Measurement Model

Motion Model

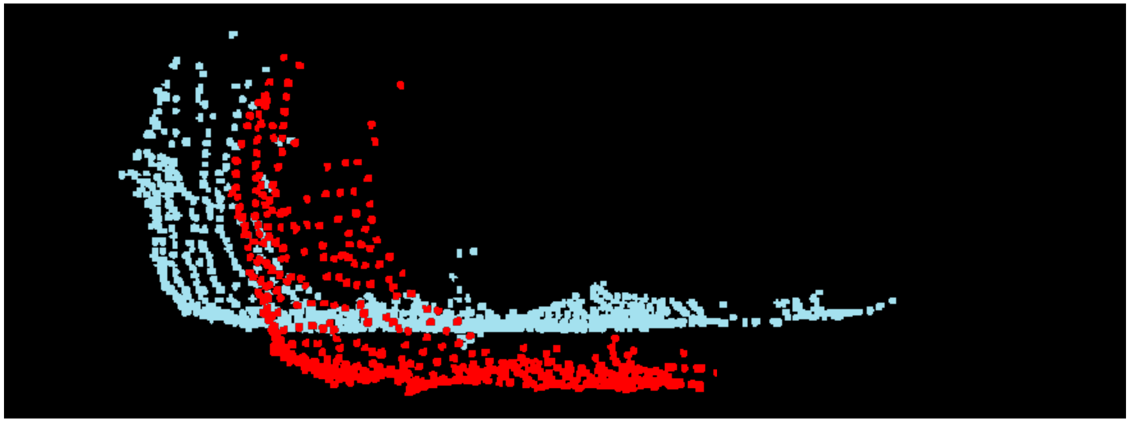


Tracking Probability

$$p(x_t \mid z_1 \ldots z_t) = \boxed{\eta\, p(z_t \mid x_t, z_{t-1})}\, p(x_t \mid z_1 \ldots z_{t-1})$$

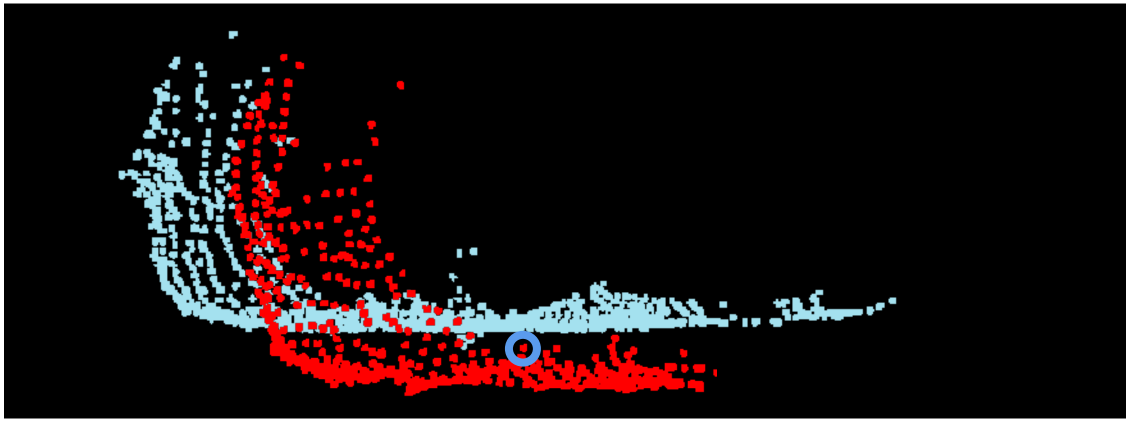
Measurement Model

Motion Model



Tracking Probability

$$p(x_t \mid z_1 \ldots z_t) = \underbrace{\boxed{\eta\, p(z_t \mid x_t, z_{t-1})}}_{\text{Measurement Model}} \underbrace{p(x_t \mid z_1 \ldots z_{t-1})}_{\text{Motion Model}}$$

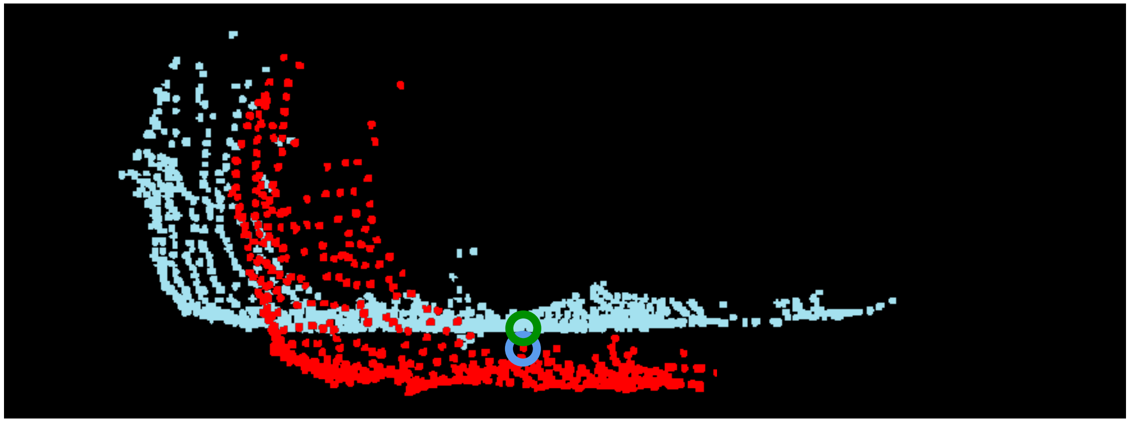


Tracking Probability

$$p(x_t \mid z_1 \ldots z_t) = \boxed{\eta\, p(z_t \mid x_t, z_{t-1})}\, p(x_t \mid z_1 \ldots z_{t-1})$$

Measurement Model

Motion Model

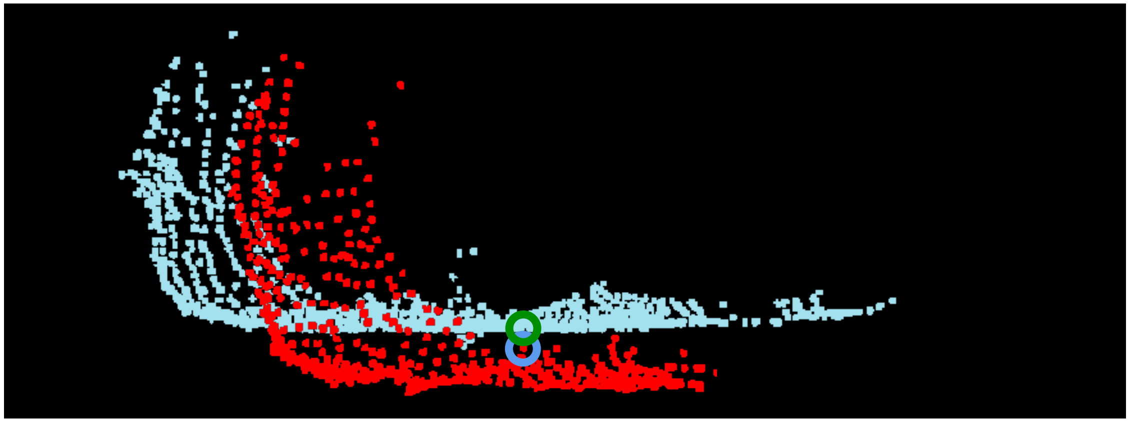


Tracking Probability

$$p(x_t \mid z_1 \ldots z_t) = \boxed{\eta\, p(z_t \mid x_t, z_{t-1})}\, p(x_t \mid z_1 \ldots z_{t-1})$$

Measurement Model

Motion Model



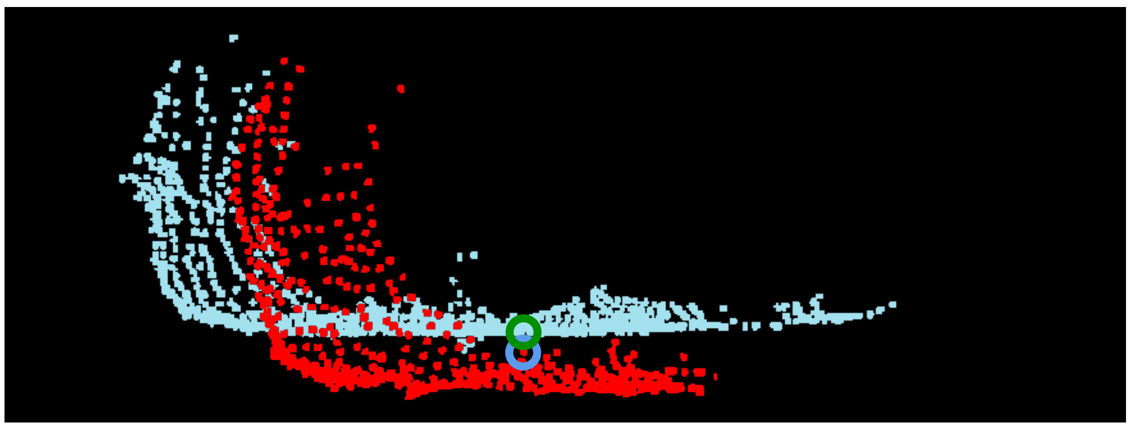
$$= \eta \left(\prod_{z_i \in z_t} \exp\left(-\tfrac{1}{2}(\boxed{z_i} - \boxed{\bar{z}_j})^T \Sigma^{-1} (z_i - \bar{z}_j)\right) + k \right)$$

Tracking Probability

$$p(x_t \mid z_1 \dots z_t) = \boxed{\eta\, p(z_t \mid x_t, z_{t-1})}\, p(x_t \mid z_1 \dots z_{t-1})$$

Measurement Model

Motion Model



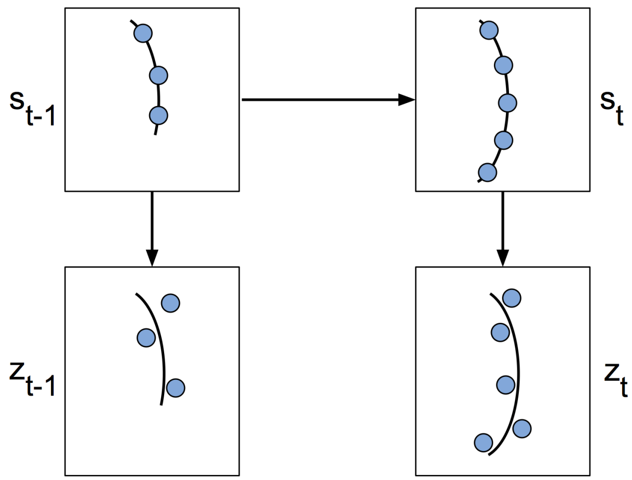
$$= \eta \left(\prod_{z_i \in z_t} \exp\left(-\frac{1}{2} (\boxed{z_i} - \boxed{\bar{z}_j})^T \Sigma^{-1} (z_i - \bar{z}_j) \right) + \boxed{k} \right)$$

Tracking Probability

$$p(x_t \mid z_1 \ldots z_t) = \boxed{\eta\, p(z_t \mid x_t, z_{t-1})}\, p(x_t \mid z_1 \ldots z_{t-1})$$

Measurement Model

Motion Model



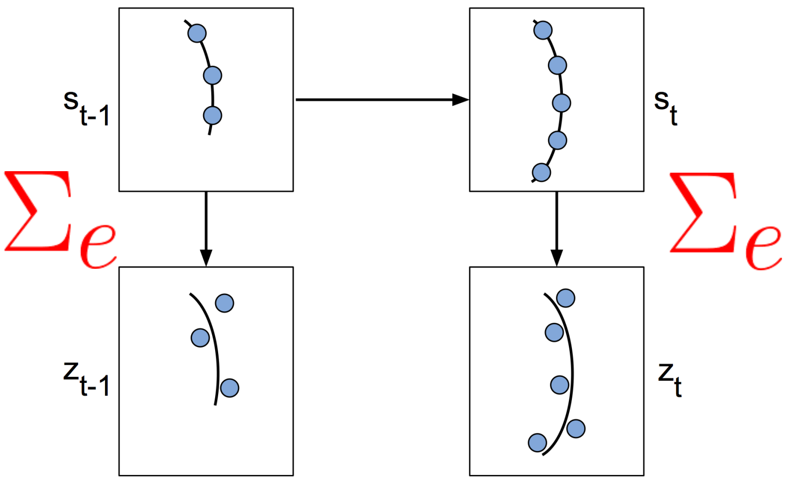
$$= \eta \left(\prod_{z_i \in z_t} \exp\left(-\frac{1}{2}(\boxed{z_i} - \boxed{\bar{z}_j})^T \Sigma^{-1}(z_i - \bar{z}_j)\right) + \boxed{k} \right)$$

Tracking Probability

$$p(x_t \mid z_1 \dots z_t) = \boxed{\eta \, p(z_t \mid x_t, z_{t-1})} \, p(x_t \mid z_1 \dots z_{t-1})$$

Measurement Model

Motion Model



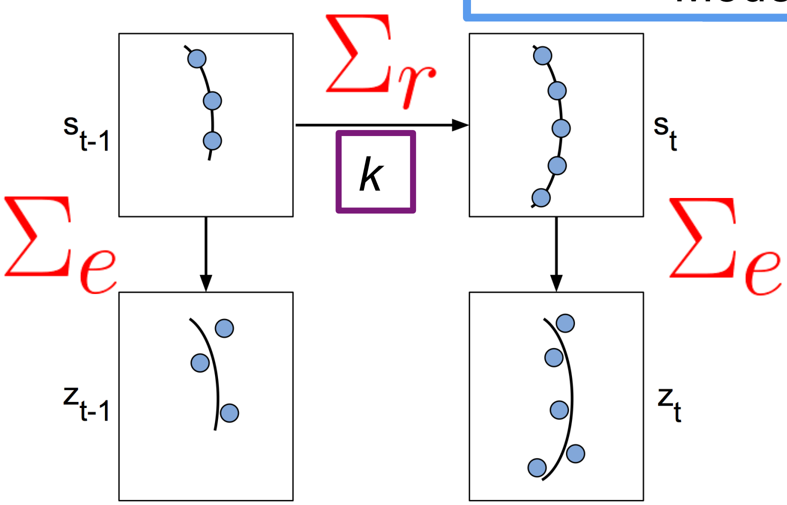
$$= \eta \left(\prod_{z_i \in z_t} \exp\left(-\tfrac{1}{2}(\boxed{z_i} - \boxed{\bar{z}_j})^T \Sigma^{-1} (z_i - \bar{z}_j)\right) + \boxed{k} \right)$$

Tracking Probability

$$p(x_t \mid z_1 \ldots z_t) = \boxed{\eta\, p(z_t \mid x_t, z_{t-1})}\, p(x_t \mid z_1 \ldots z_{t-1})$$

Measurement Model

Motion Model



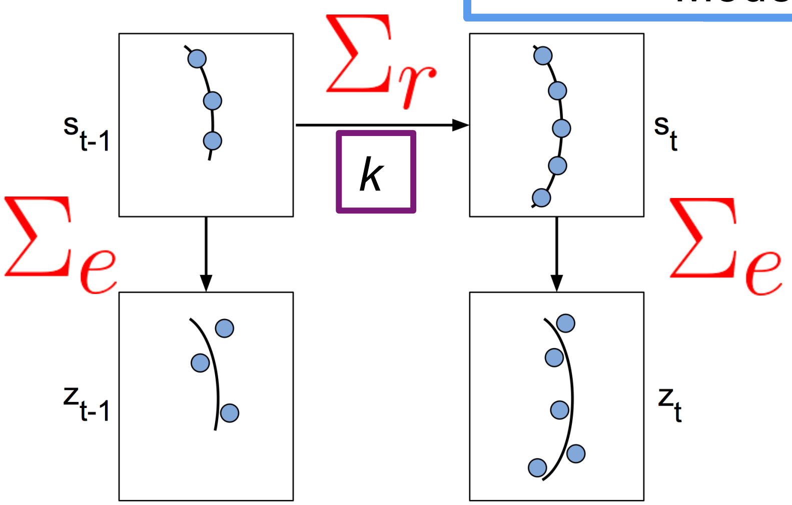
$$= \eta \left(\prod_{z_i \in z_t} \exp\left(-\frac{1}{2}(\boxed{z_i} - \boxed{\bar{z}_j})^T \Sigma^{-1}(z_i - \bar{z}_j)\right) + \boxed{k}\right)$$

Tracking Probability

$$p(x_t \mid z_1 \ldots z_t) = \boxed{\eta\, p(z_t \mid x_t, z_{t-1})}\; p(x_t \mid z_1 \ldots z_{t-1})$$

Measurement Model

Motion Model

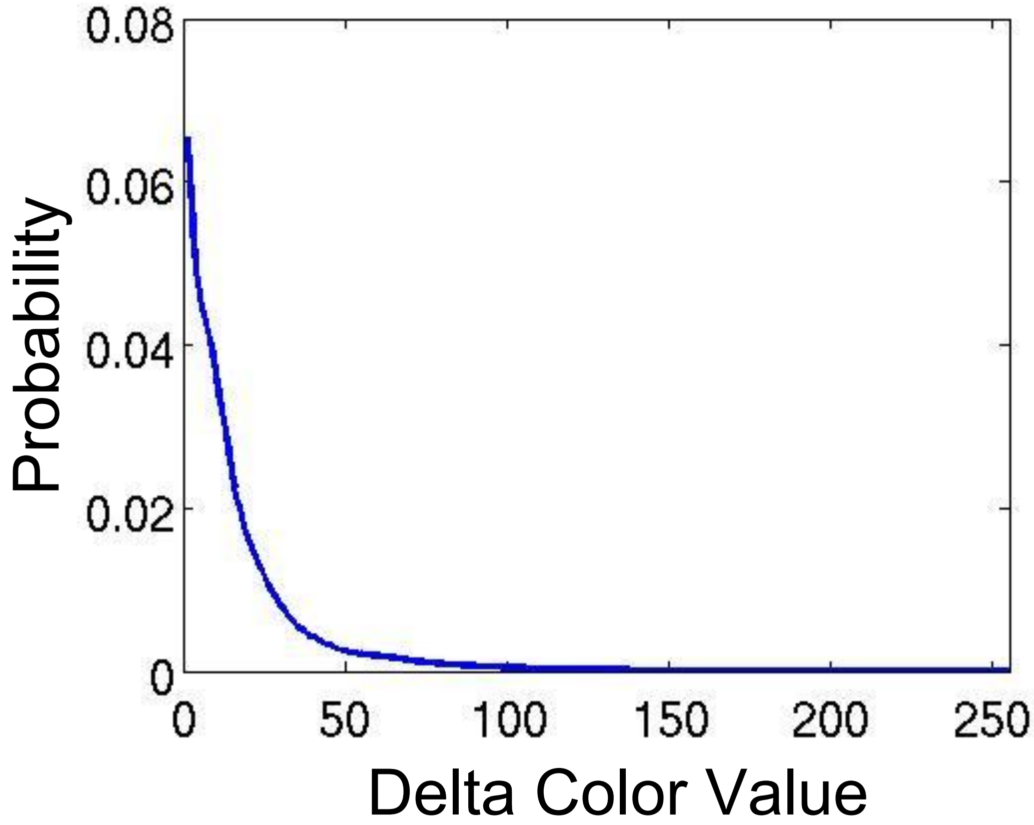


$$\Sigma = 2\underbrace{\Sigma_e}_{\text{Sensor noise}} + \underbrace{\Sigma_r}_{\text{Sensor resolution}}$$

$$= \eta \left(\prod_{z_i \in z_t} \exp\left(-\frac{1}{2}(\boxed{z_i} - \boxed{\bar{z}_j})^T \Sigma^{-1}(z_i - \bar{z}_j)\right) + \boxed{k}\right)$$


EMERGENCY
911
EMERGENCY

Color Probability



Including Color

$$p(z_t \mid x_t, z_{t-1}) = \eta \prod_{z_i \in z_t} p_s(z_i \mid \bar{z}_j)$$

Including Color

$$p(z_t \mid x_t, z_{t-1}) = \eta \prod_{z_i \in z_t} p_s(z_i \mid \bar{z}_j)$$

$$\eta \left(\prod_{z_i \in z_t} \exp(-\tfrac{1}{2}(z_i - \bar{z}_j)^T \Sigma^{-1} (z_i - \bar{z}_j)) + k \right)$$

Including Color

$$p(z_t \mid x_t, z_{t-1}) = \eta \prod_{z_i \in z_t} p_s(z_i \mid \bar{z}_j) \boxed{p_c(z_i \mid \bar{z}_j)}$$

$$\eta \left(\prod_{z_i \in z_t} \exp(-\tfrac{1}{2}(z_i - \bar{z}_j)^T \Sigma^{-1} (z_i - \bar{z}_j)) + k \right)$$

Including Color

$$p(z_t \mid x_t, z_{t-1}) = \eta \prod_{z_i \in z_t} p_s(z_i \mid \bar{z}_j) \boxed{p_c(z_i \mid \bar{z}_j)}$$

$$p(C)p(z_t \mid \bar{z}_j, C) + P(\neg C)p(z_t \mid \bar{z}_j, \neg C)$$

Including Color

$$p(z_t \mid x_t, z_{t-1}) = \eta \prod_{z_i \in z_t} p_s(z_i \mid \bar{z}_j) \boxed{p_c(z_i \mid \bar{z}_j)}$$

$$p(C)p(z_t \mid \bar{z}_j, C) + P(\neg C)p(z_t \mid \bar{z}_j, \neg C)$$

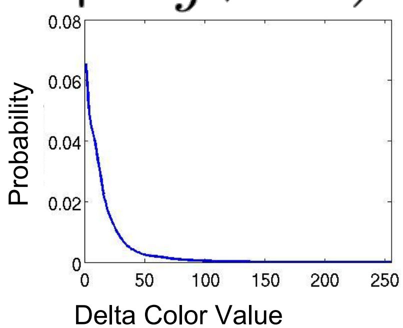
$$p_c \exp(\tfrac{-r^2}{2\sigma_c^2})$$

Including Color

$$p(z_t \mid x_t, z_{t-1}) = \eta \prod_{z_i \in z_t} p_s(z_i \mid \bar{z}_j) \boxed{p_c(z_i \mid \bar{z}_j)}$$

$$p(C)p(z_t \mid \bar{z}_j, C) + P(\neg C)p(z_t \mid \bar{z}_j, \neg C)$$

$$p_c \exp(\frac{-r^2}{2\sigma_c^2})$$

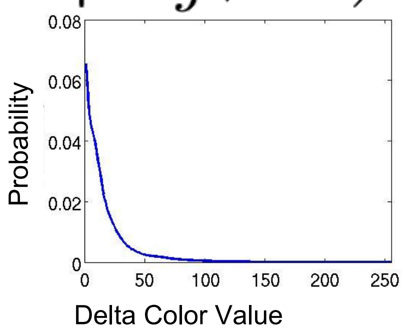


Including Color

$$p(z_t \mid x_t, z_{t-1}) = \eta \prod_{z_i \in z_t} p_s(z_i \mid \bar{z}_j) \boxed{p_c(z_i \mid \bar{z}_j)}$$

$$p(C)p(z_t \mid \bar{z}_j, C) + P(\neg C)p(z_t \mid \bar{z}_j, \neg C)$$

$$p_c \exp(\frac{-r^2}{2\sigma_c^2})$$



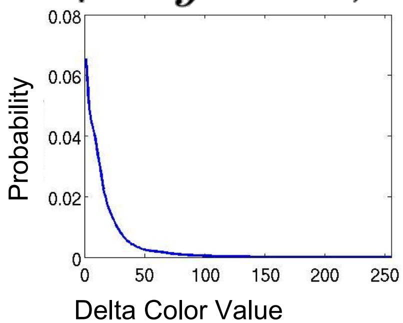
$$1 - p_c \exp(\frac{-r^2}{2\sigma_c^2})$$

Including Color

$$p(z_t \mid x_t, z_{t-1}) = \eta \prod_{z_i \in z_t} p_s(z_i \mid \bar{z}_j) \boxed{p_c(z_i \mid \bar{z}_j)}$$

$$p(C)p(z_t \mid \bar{z}_j, C) + P(\neg C)p(z_t \mid \bar{z}_j, \neg C)$$

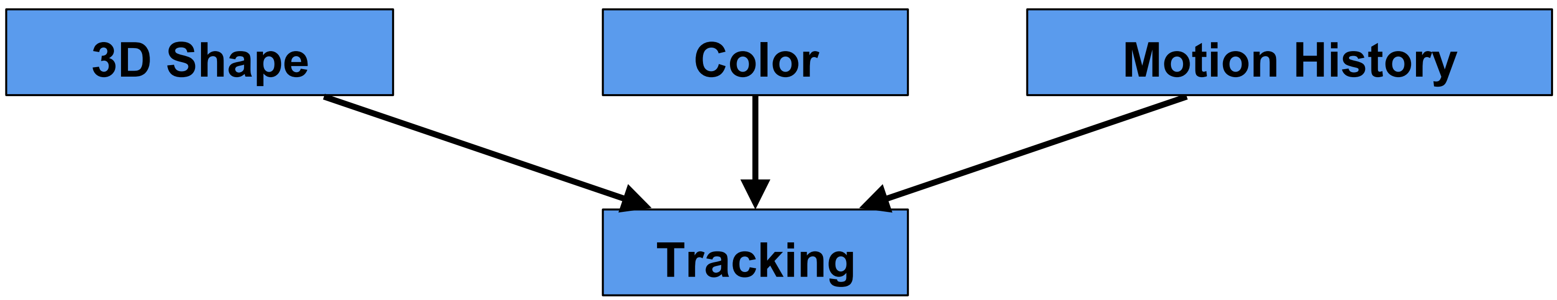
$$p_c \exp(\frac{-r^2}{2\sigma_c^2})$$



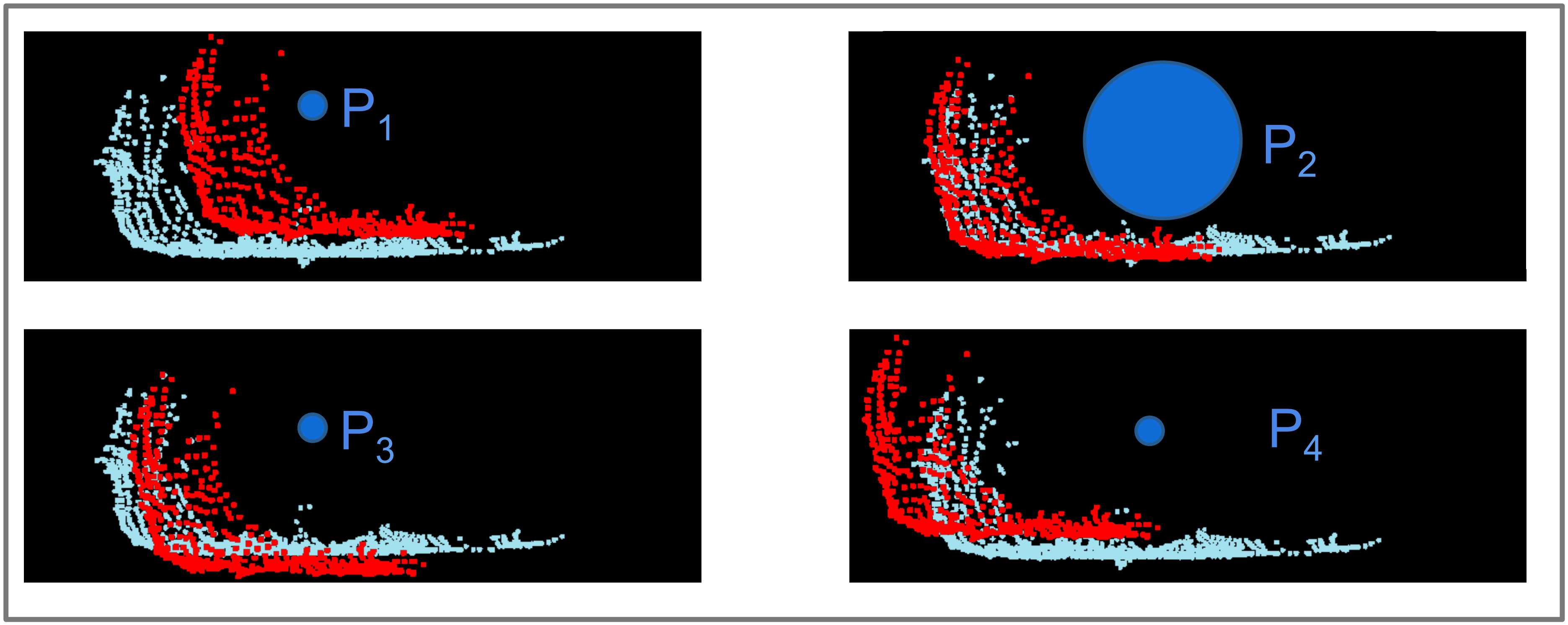
$$1 - p_c \exp(\frac{-r^2}{2\sigma_c^2})$$

$$\frac{1}{255}$$

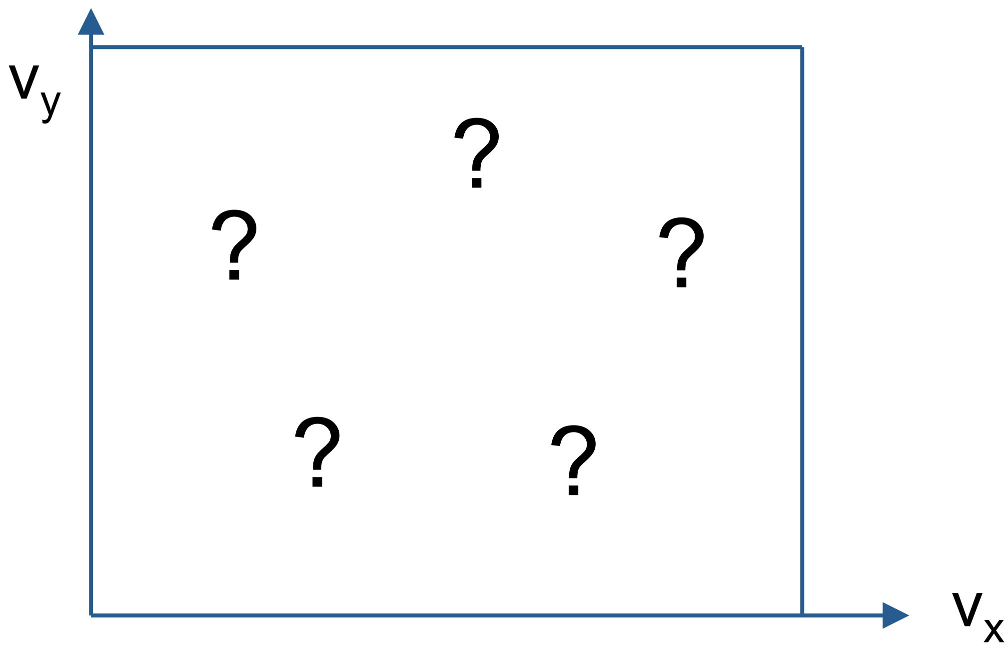
Probabilistic Framework



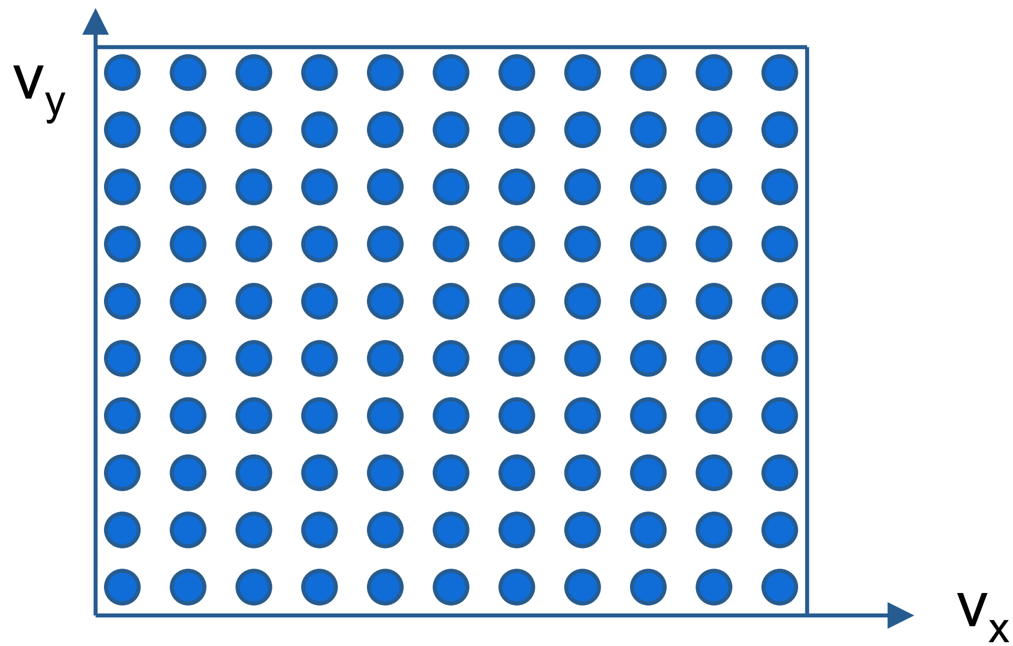
Tracking Probability



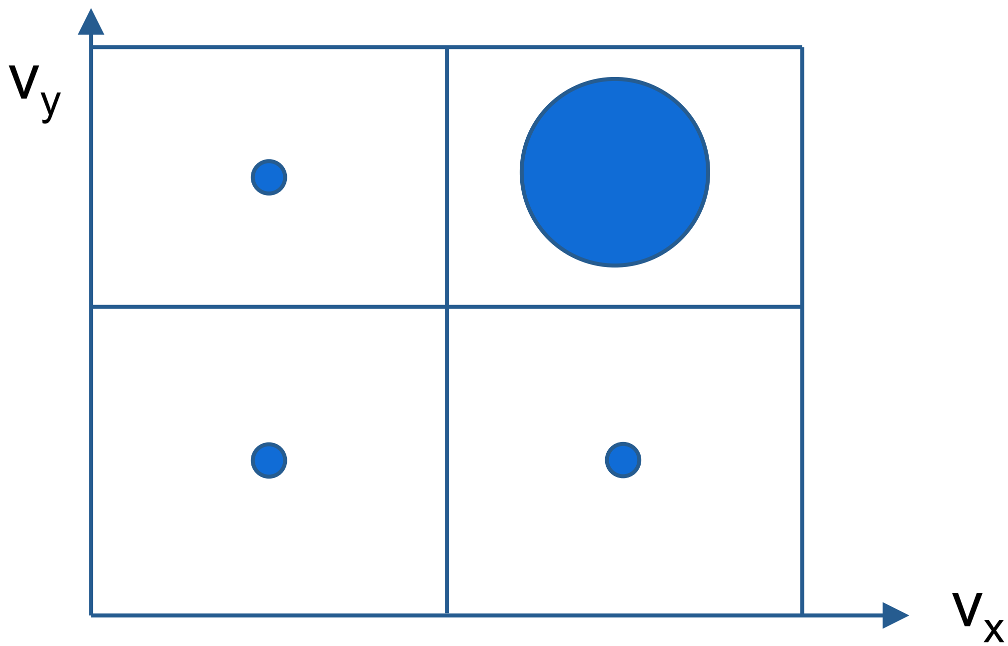
Tracking Probability



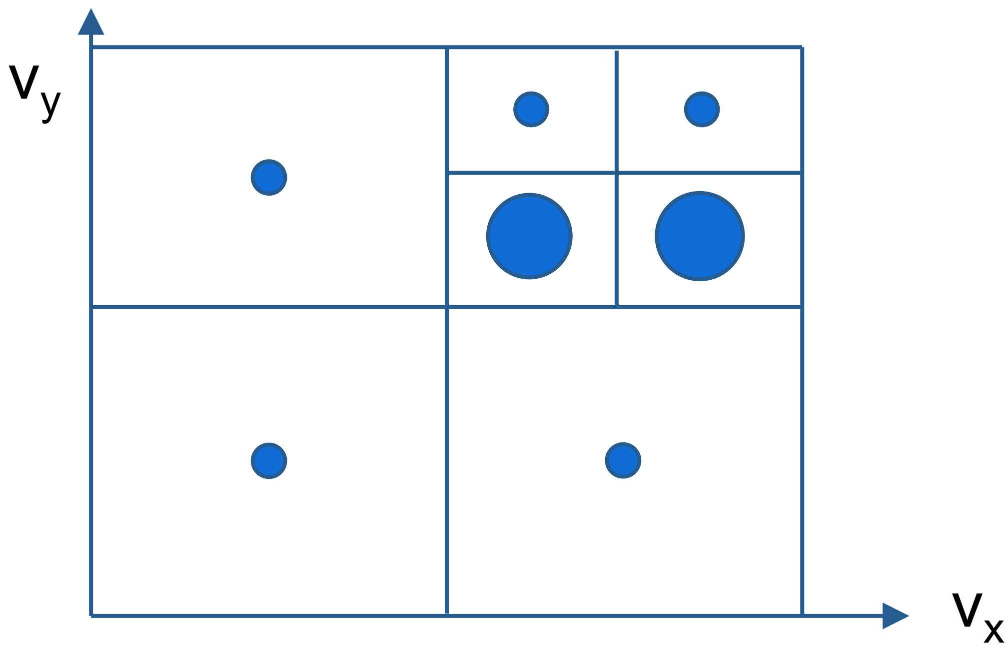
Tracking Probability



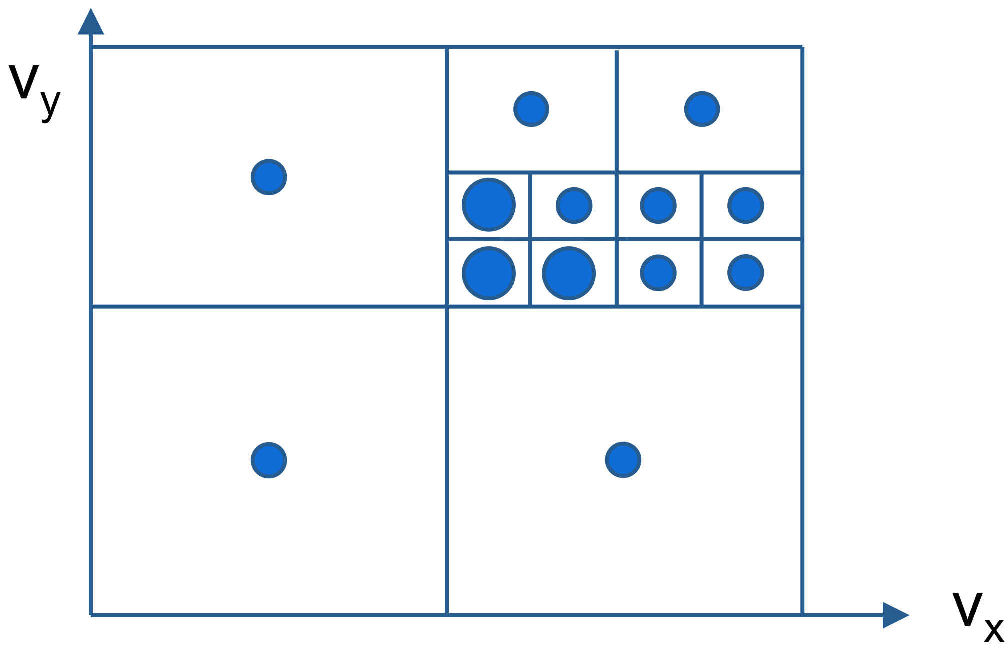
Dynamic Decomposition



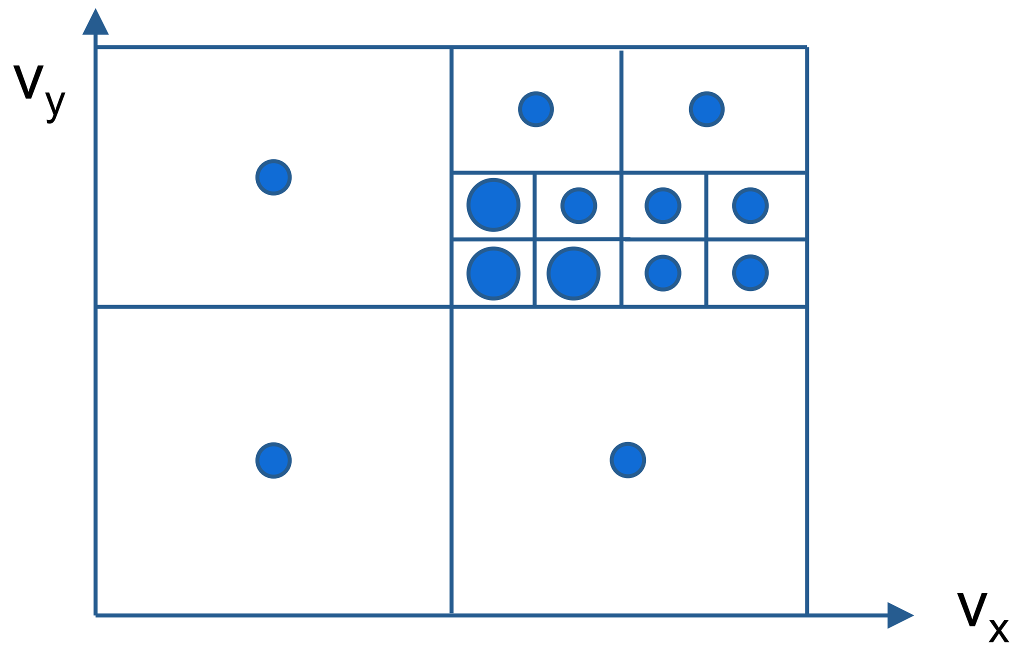
Dynamic Decomposition



Dynamic Decomposition

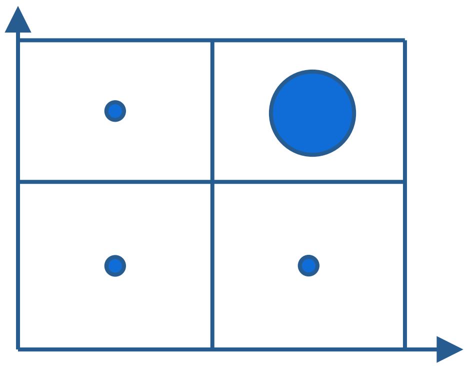


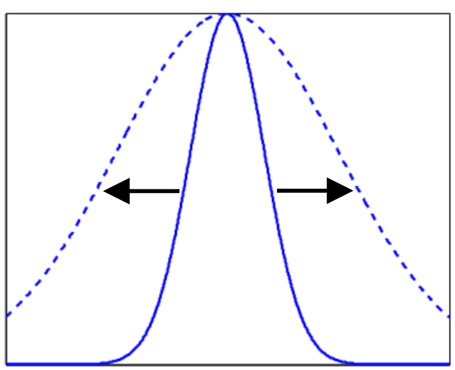
Dynamic Decomposition



Derived from minimizing KL-divergence between approximate distribution and true posterior

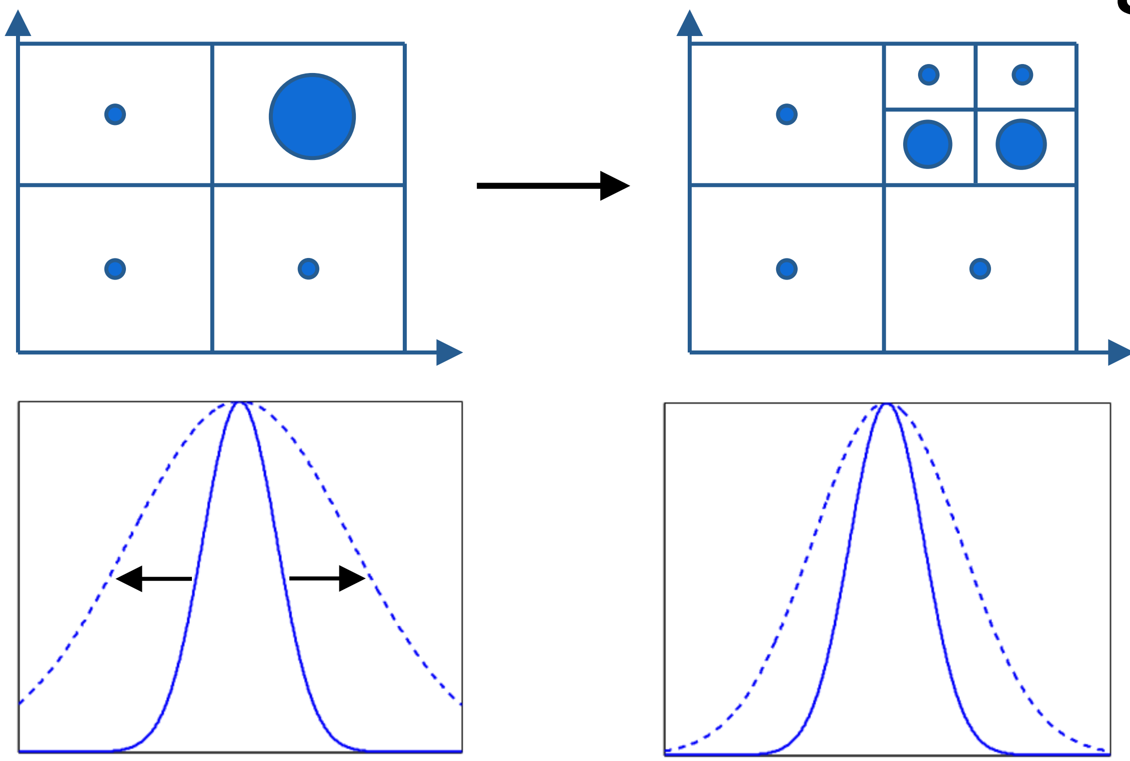
Annealing





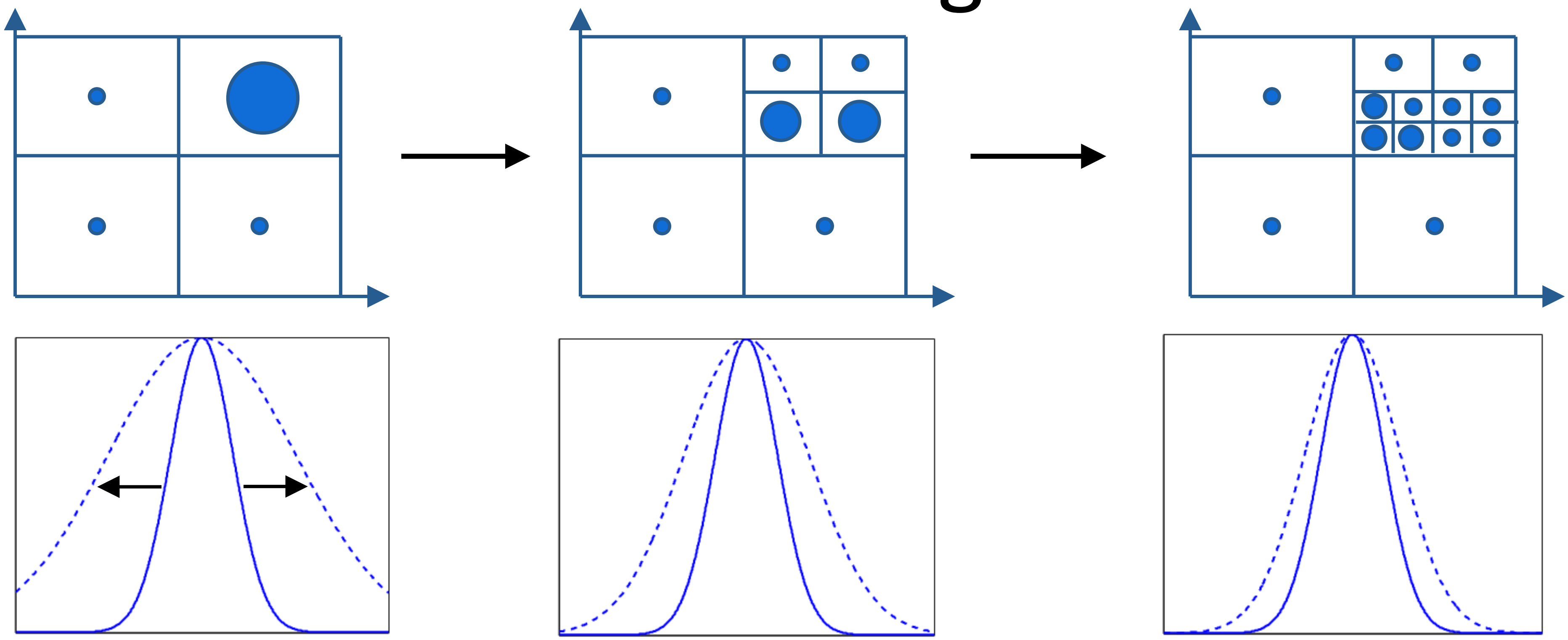
Inflate the
measurement model

Annealing



Inflate the measurement model

Annealing



Inflate the
measurement model

Algorithm

1. For each hypothesis
   A. Compute the probability of the alignment

$$p(x_t \mid z_1 \ldots z_t) = \eta \, \underbrace{p(z_t \mid x_t, z_{t-1})}_{\text{Measurement Model}} \, \underbrace{p(x_t \mid z_1 \ldots z_{t-1})}_{\text{Motion Model}}$$

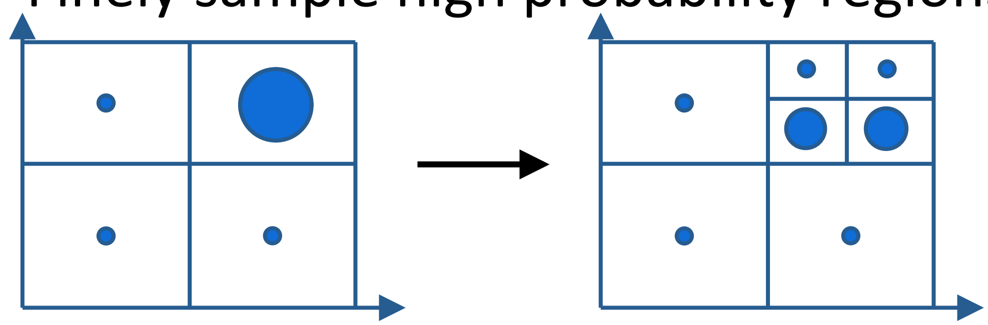
Algorithm

1. For each hypothesis

   A. Compute the probability of the alignment

$$p(x_t \mid z_1 \ldots z_t) = \eta \, \underbrace{p(z_t \mid x_t, z_{t-1})}_{\text{Measurement Model}} \, \underbrace{p(x_t \mid z_1 \ldots z_{t-1})}_{\text{Motion Model}}$$

   B. Finely sample high probability regions

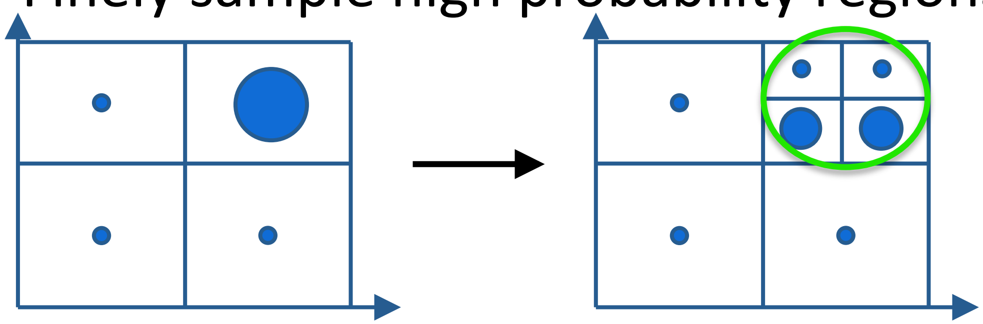


Algorithm

1. For each hypothesis
   A. Compute the probability of the alignment

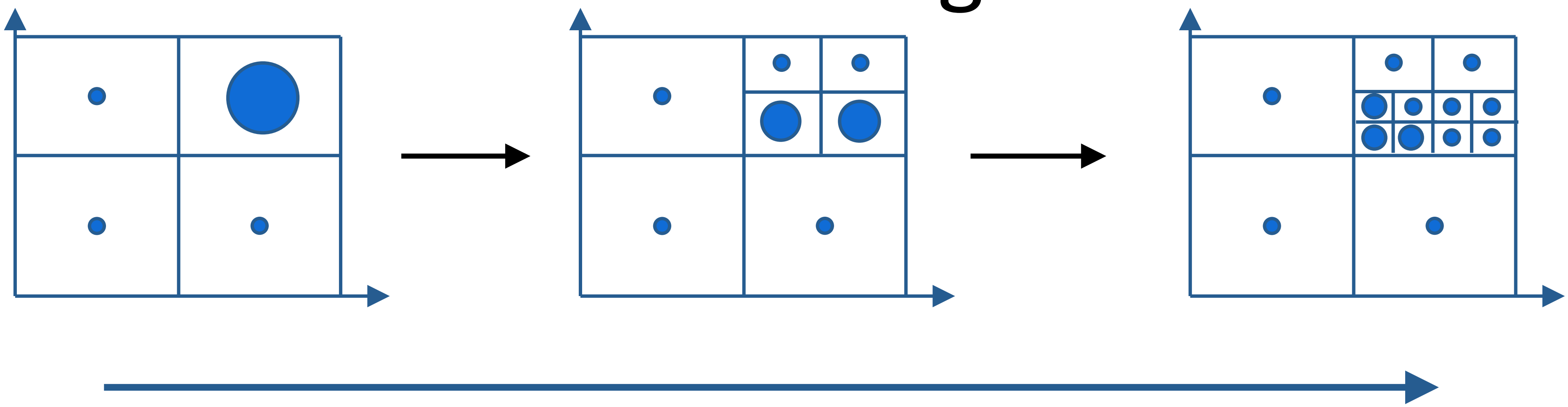
$$p(x_t \mid z_1 \ldots z_t) = \eta \, \underbrace{p(z_t \mid x_t, z_{t-1})}_{\text{Measurement Model}} \, \underbrace{p(x_t \mid z_1 \ldots z_{t-1})}_{\text{Motion Model}}$$

   B. Finely sample high probability regions



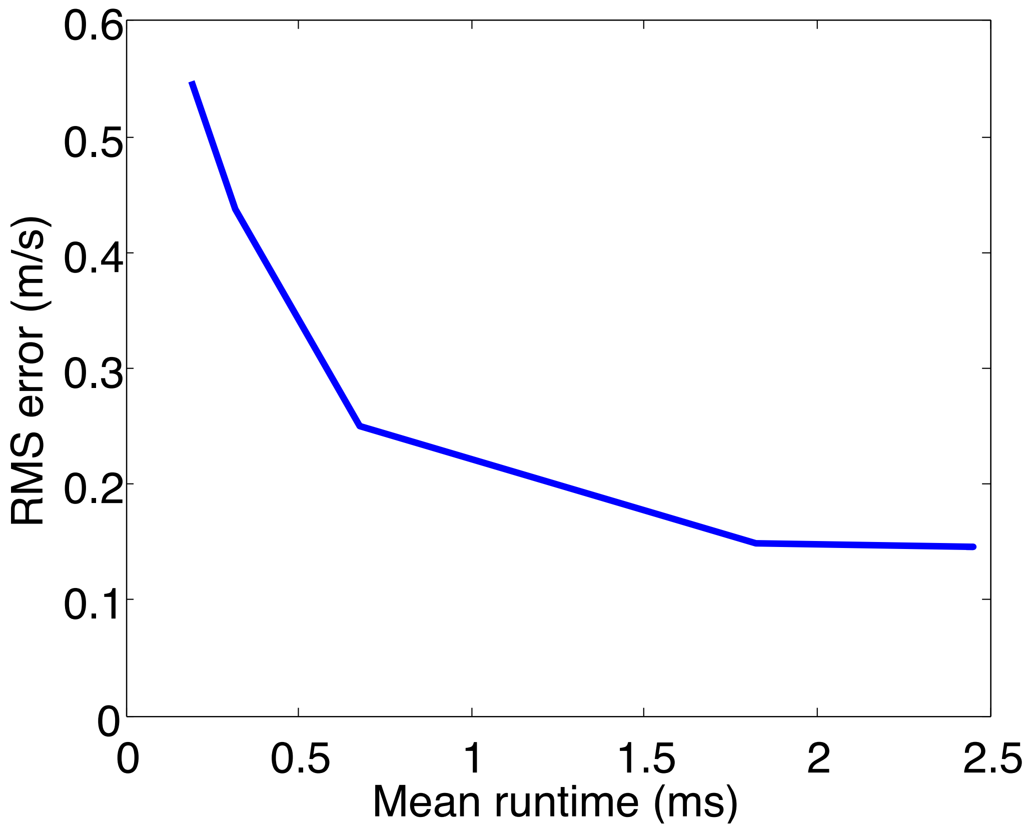
   C. Go to step 1 to compute the probability of new hypotheses

Annealing

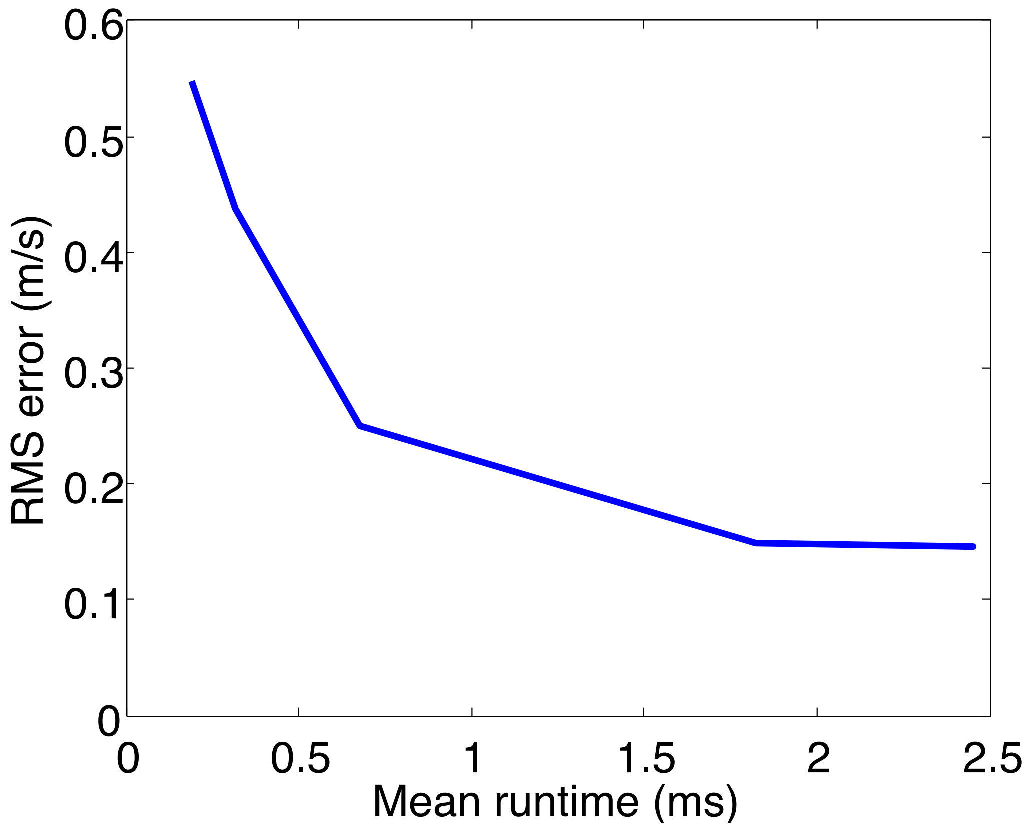


More time
More accurate

Anytime Tracker



Anytime Tracker



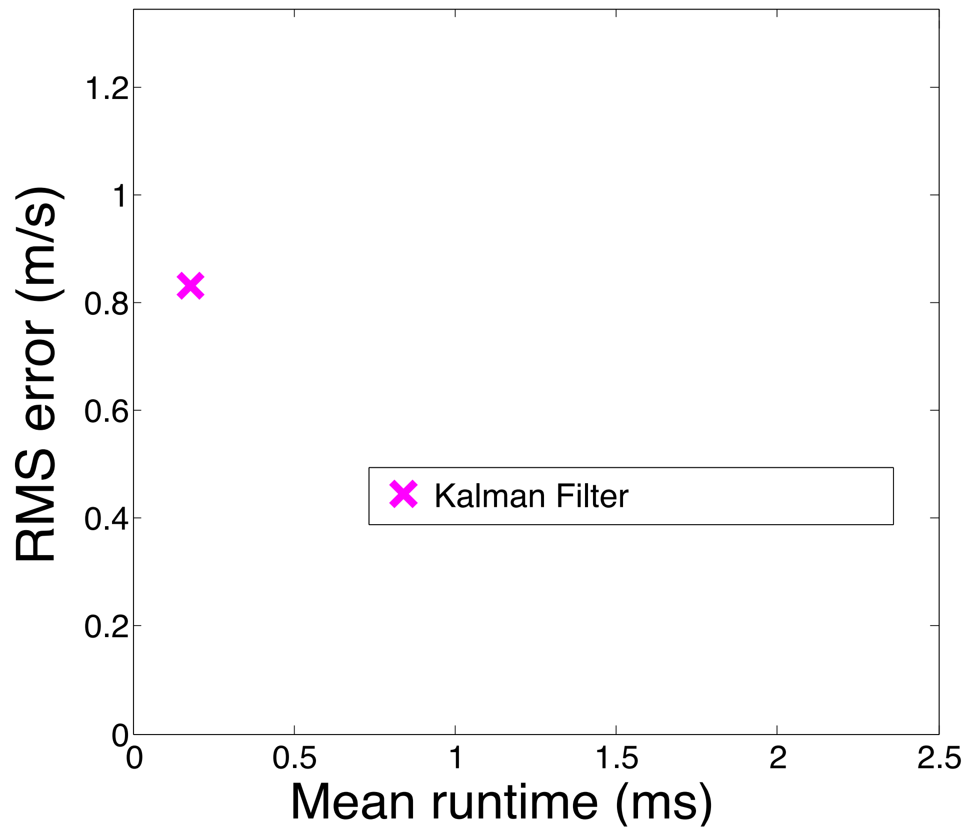
Choose runtime based on:

Total runtime requirements

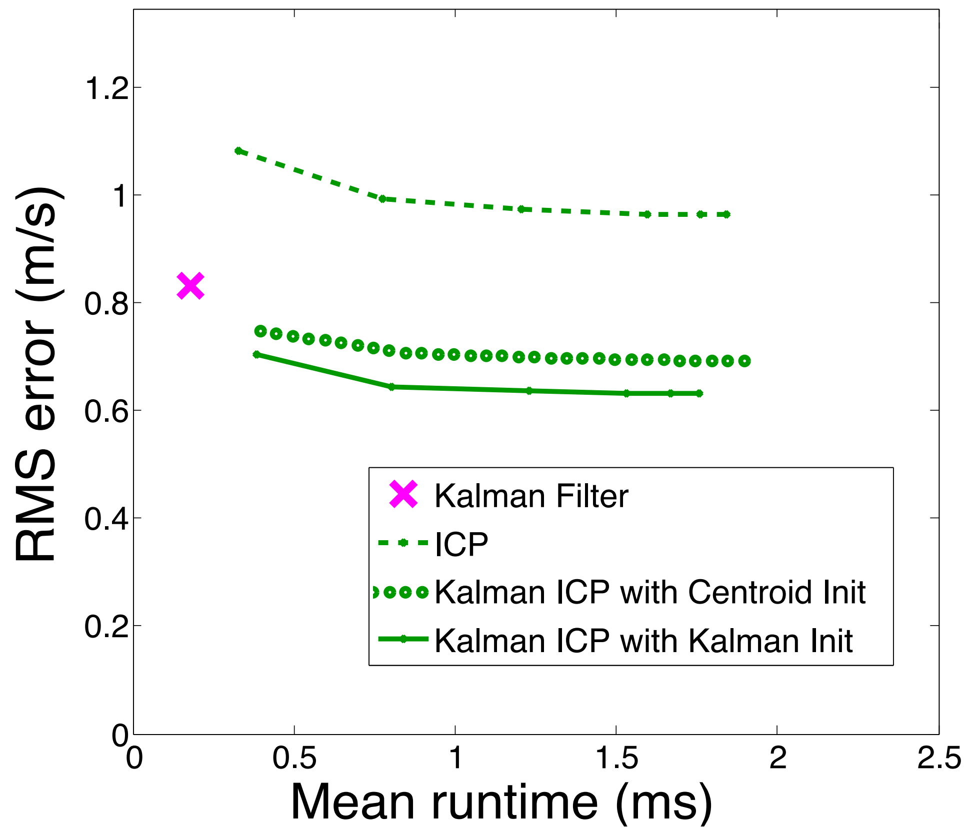
Importance of tracked object

...

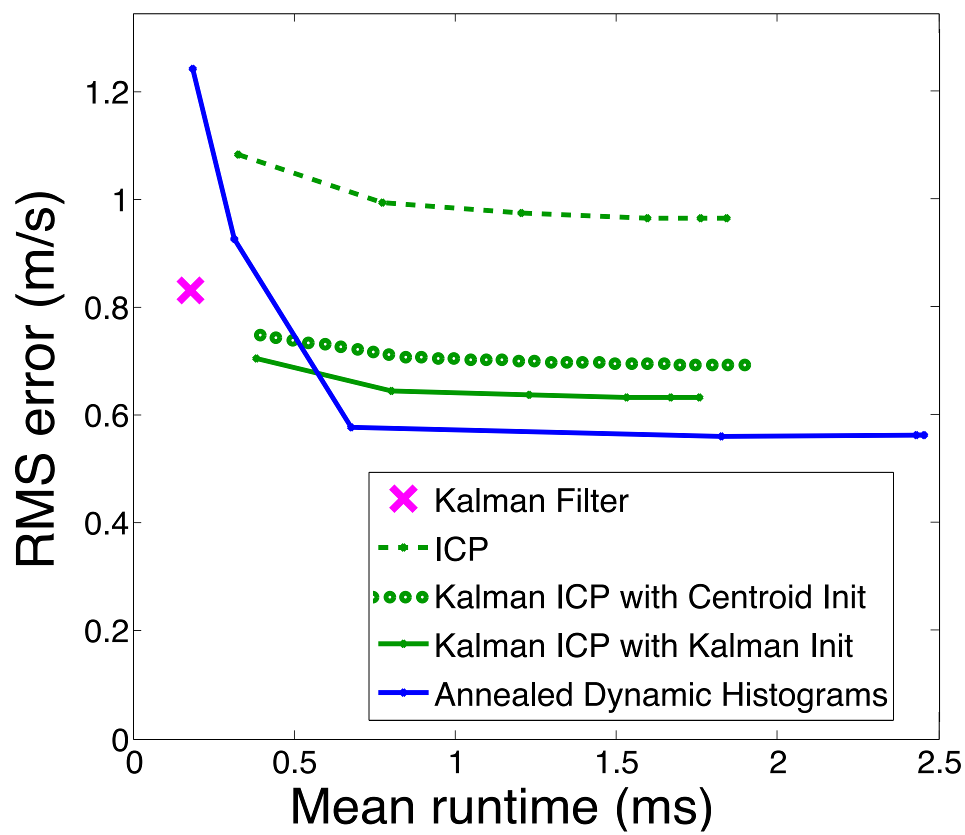
Comparisons



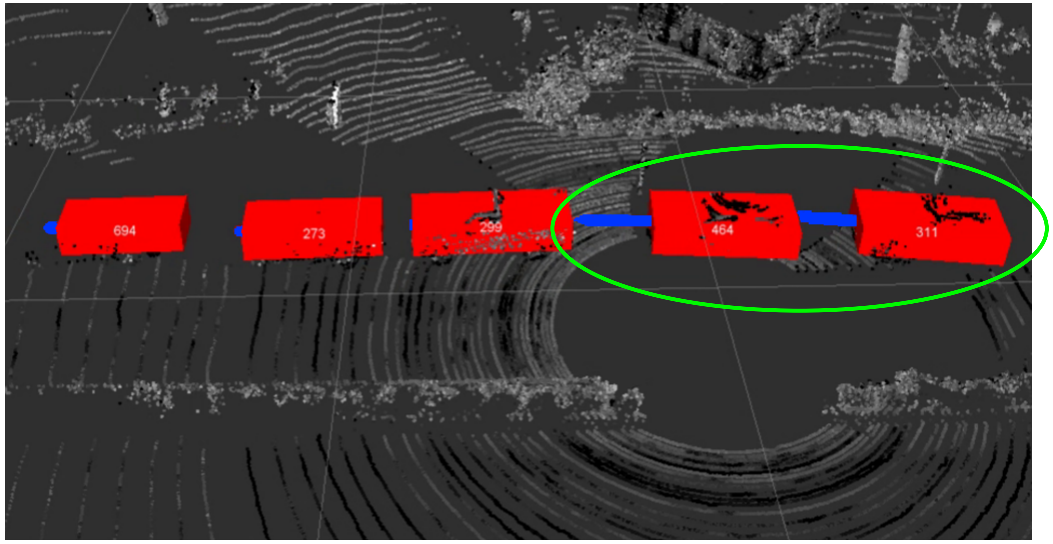
Comparisons



Comparisons

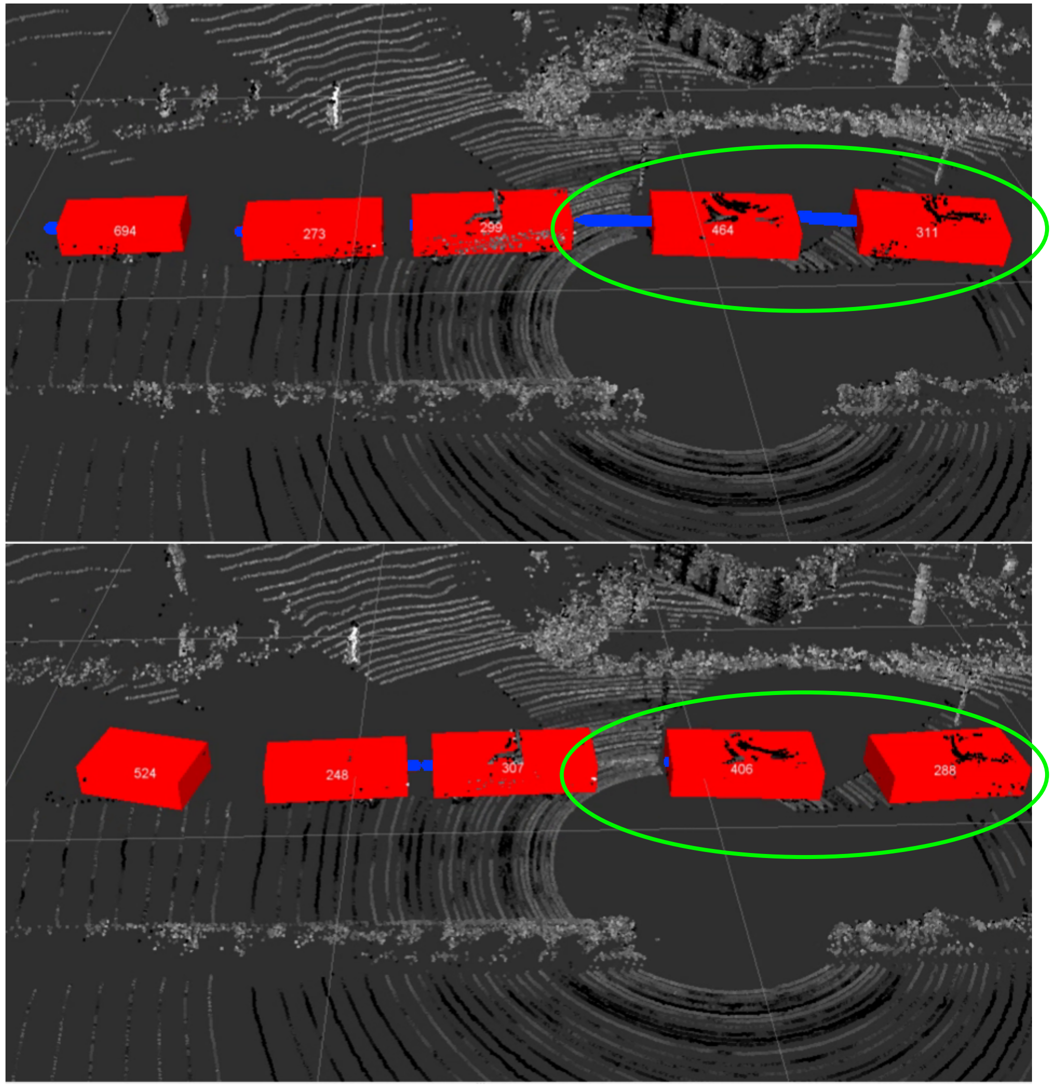


Kalman Filter

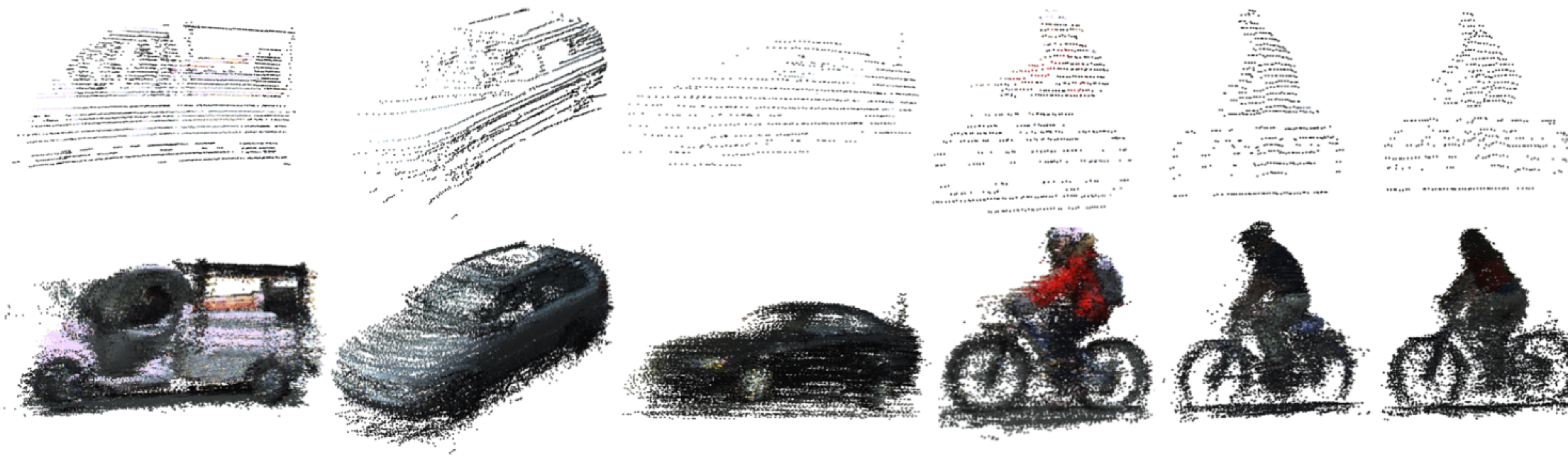


Kalman Filter

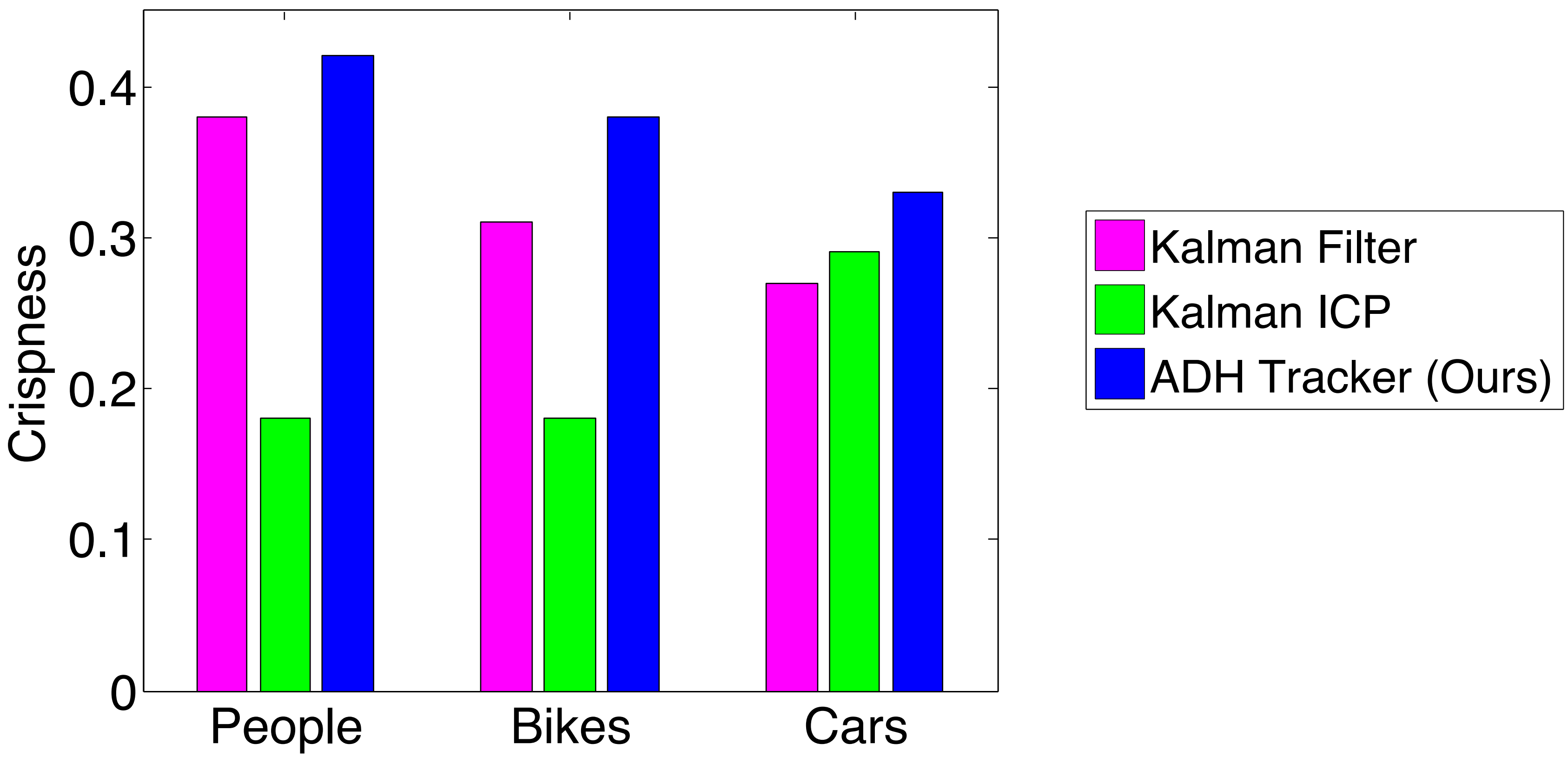
ADH Tracker (Ours)



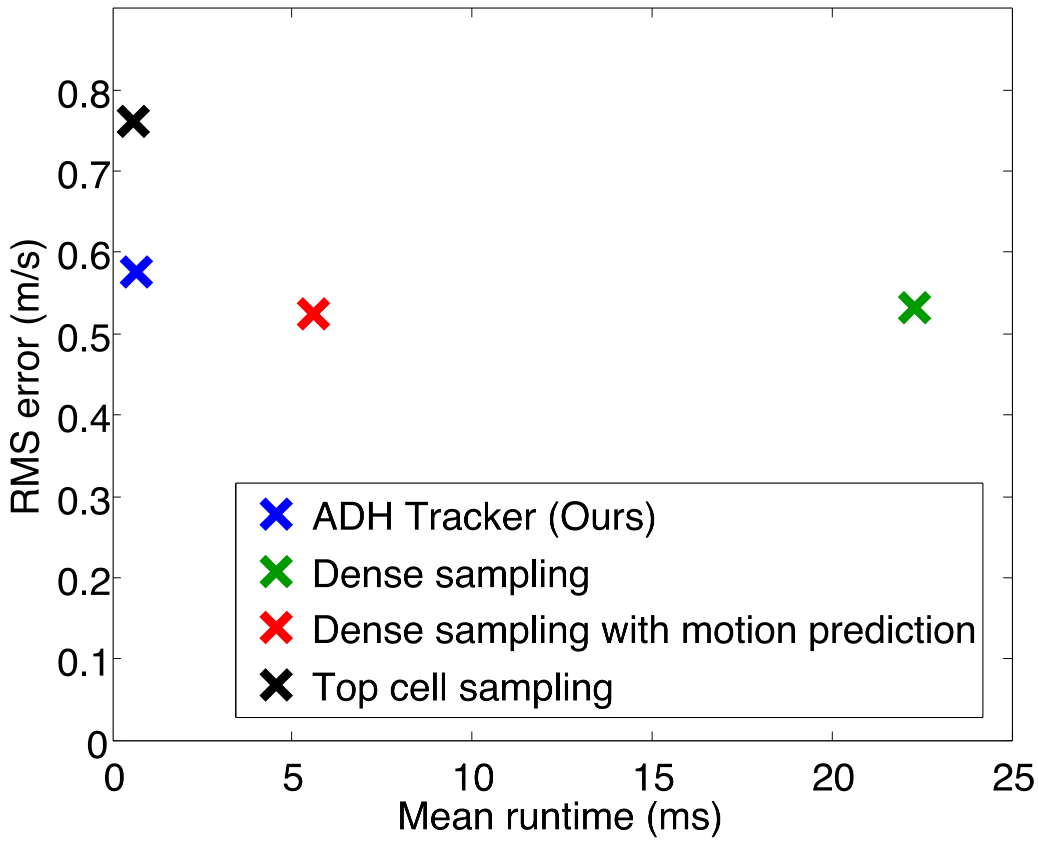
Models



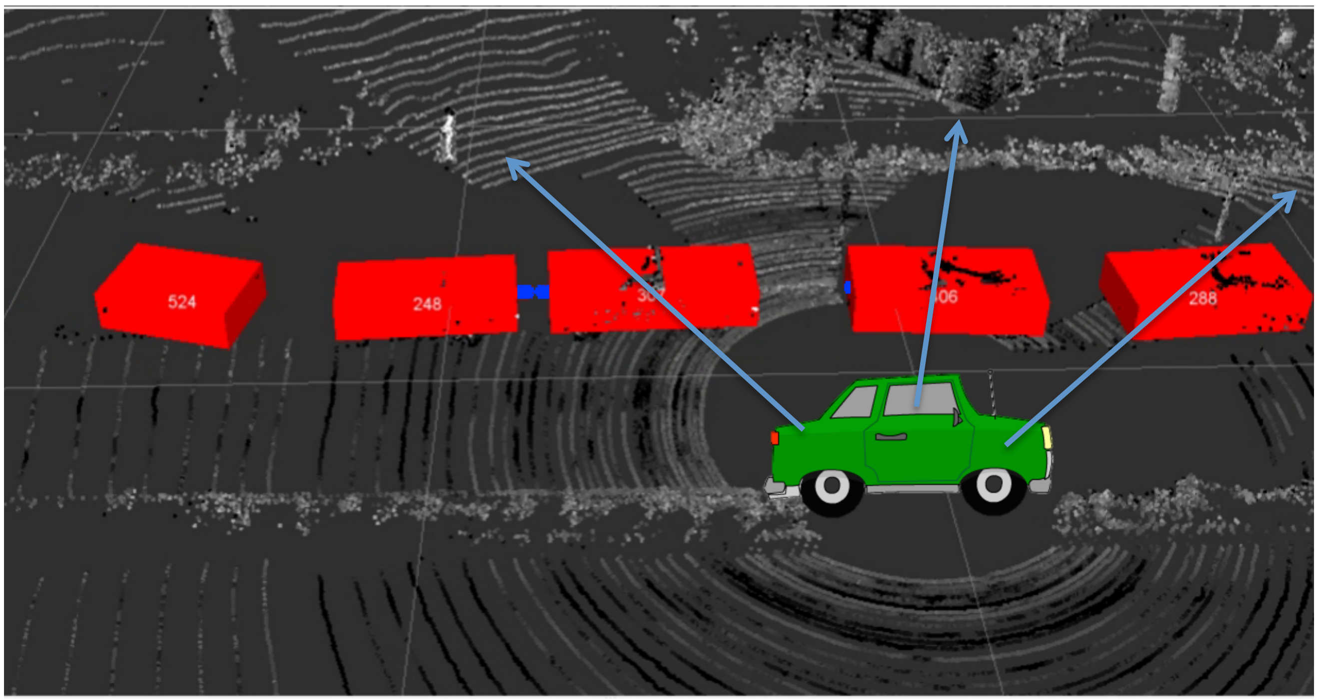
Quantitative Evaluation 2



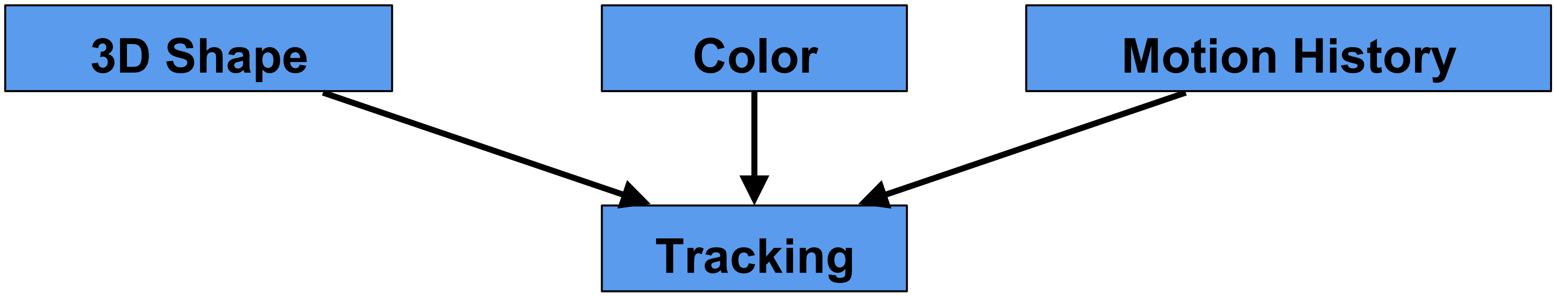
Sampling Strategies



Advantages over Radar

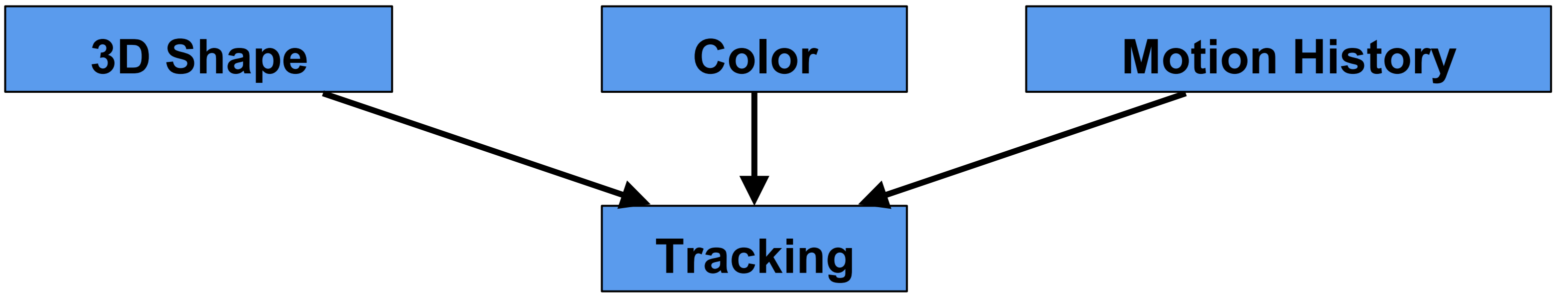


Conclusions

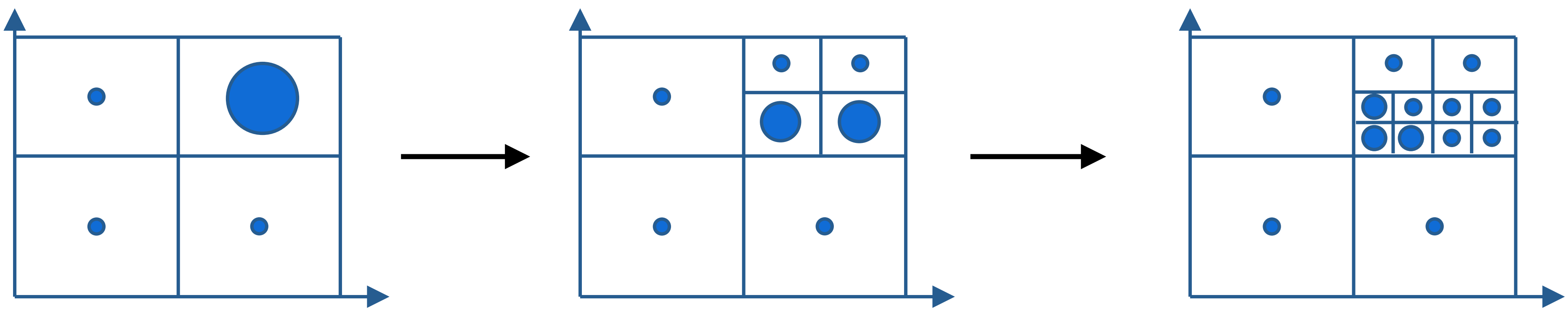


- **Robust to Occlusions, Viewpoint Changes**

Conclusions



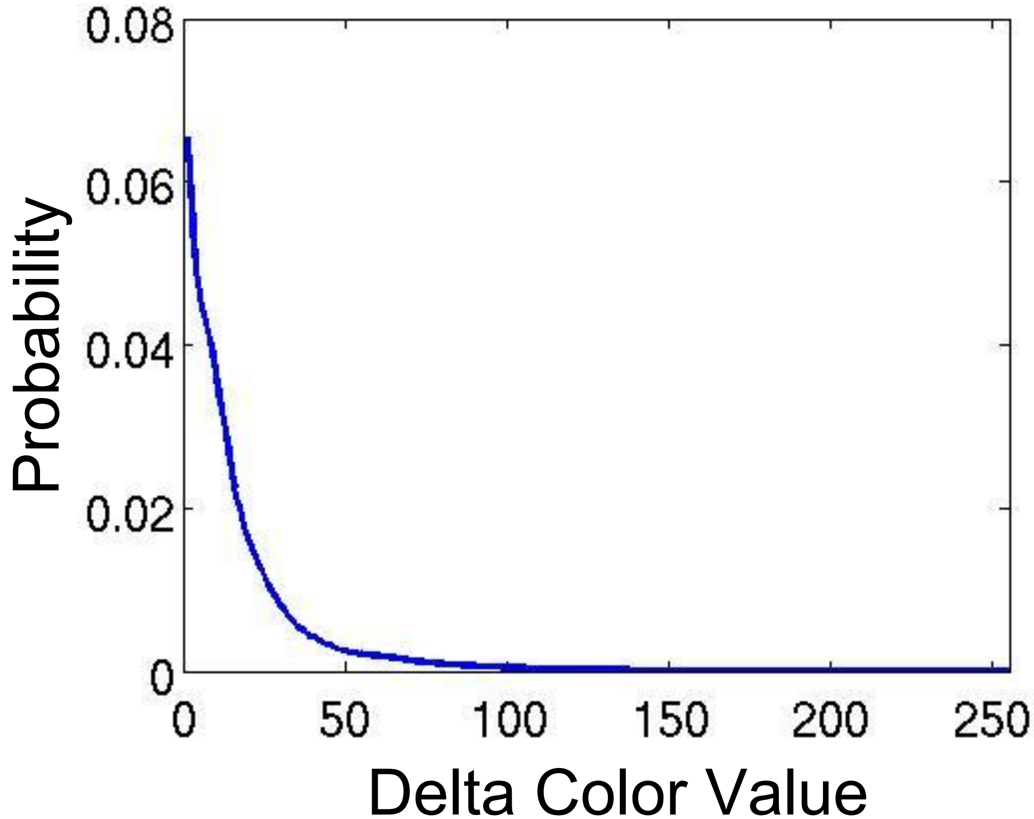
- Robust to Occlusions, Viewpoint Changes



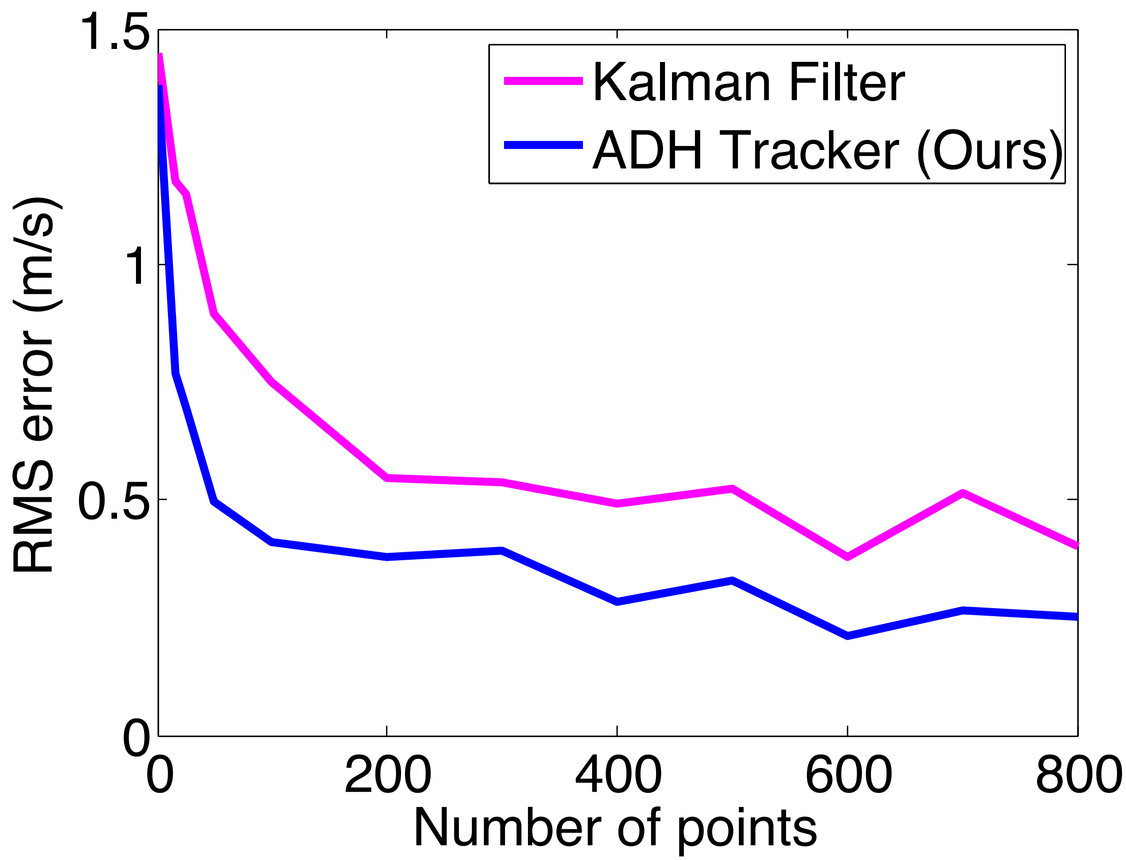
- Runs in Real-time
- Robust to Initialization Errors


EMERGENCY
911

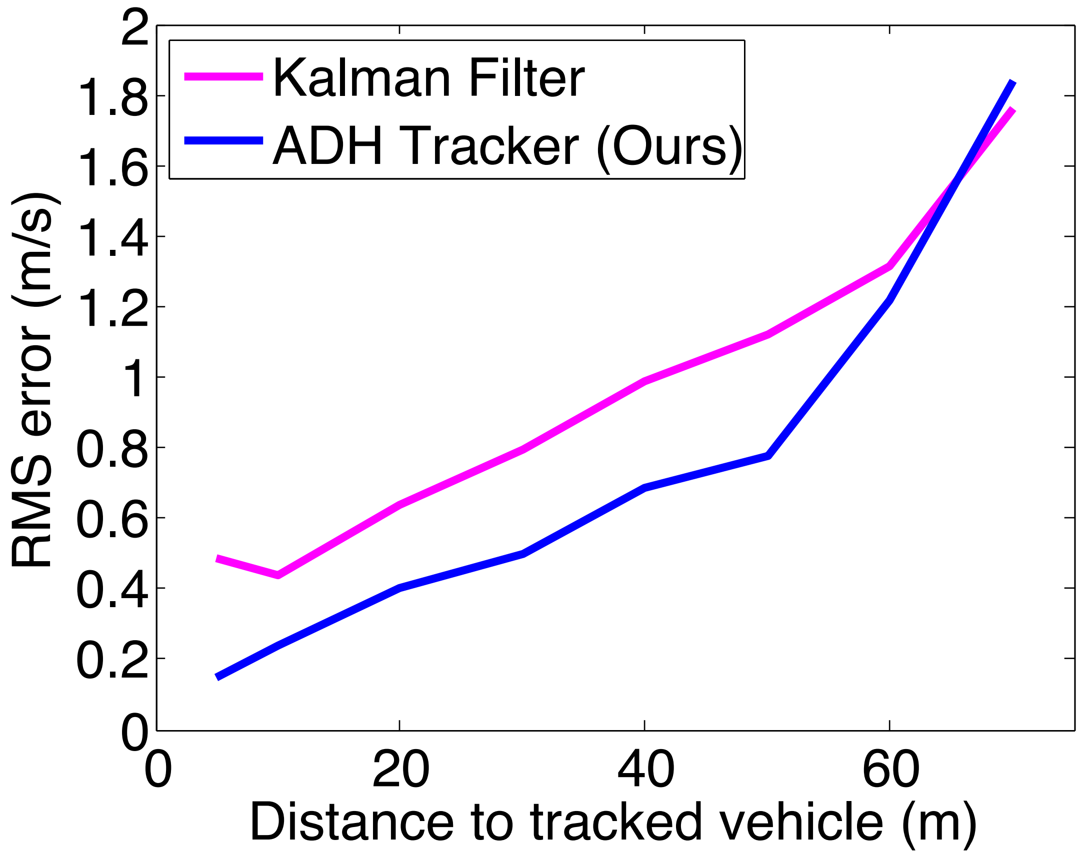
Color Probability



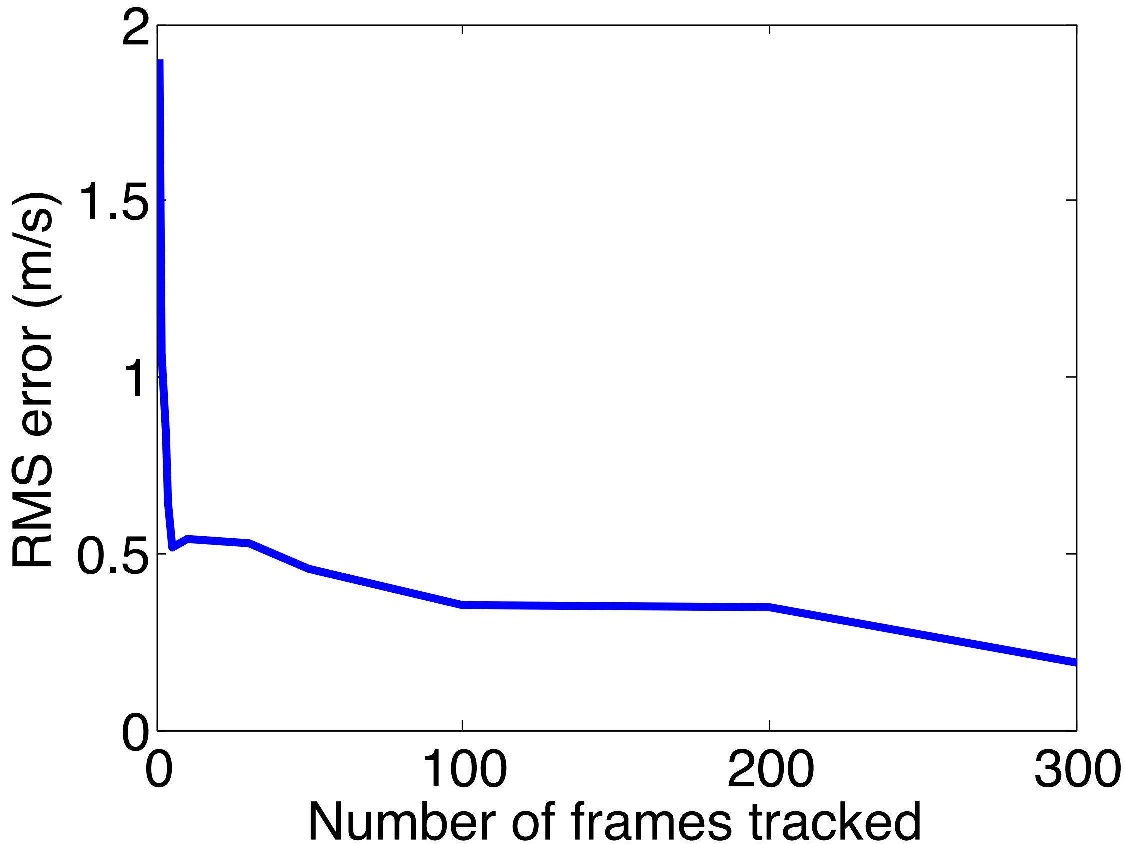
Error vs Number of Points



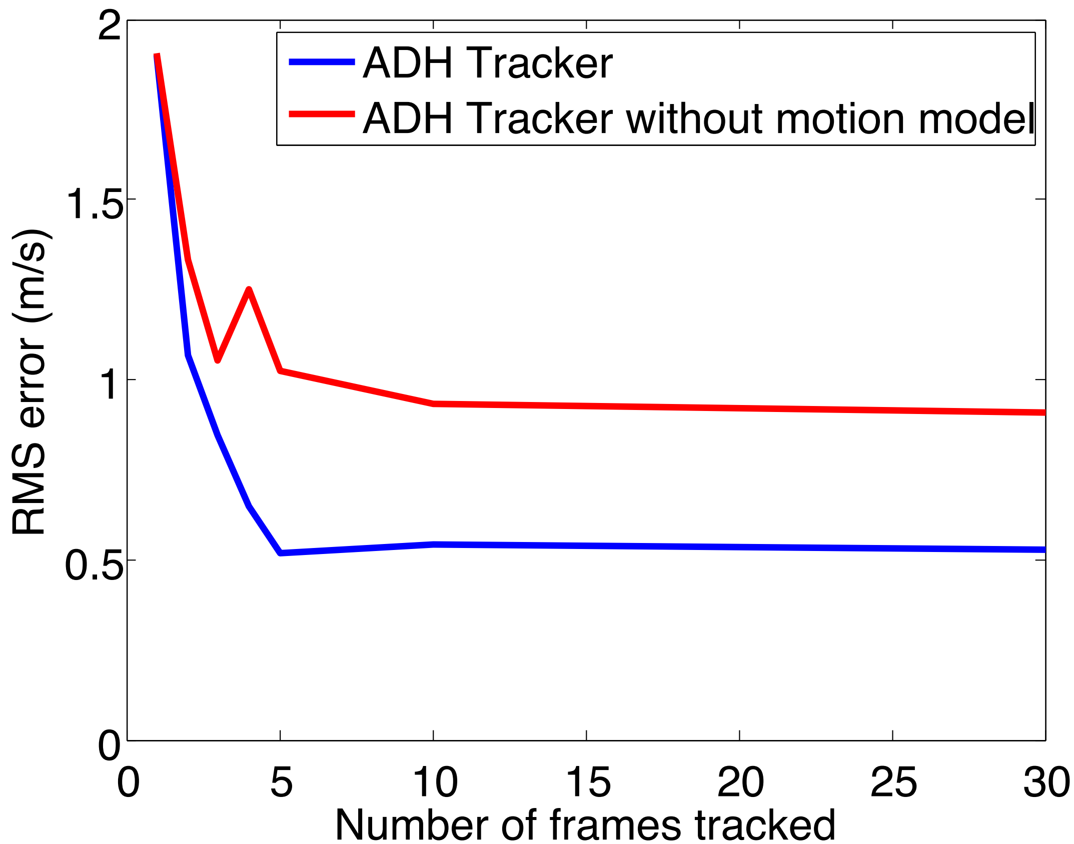
Error vs Distance

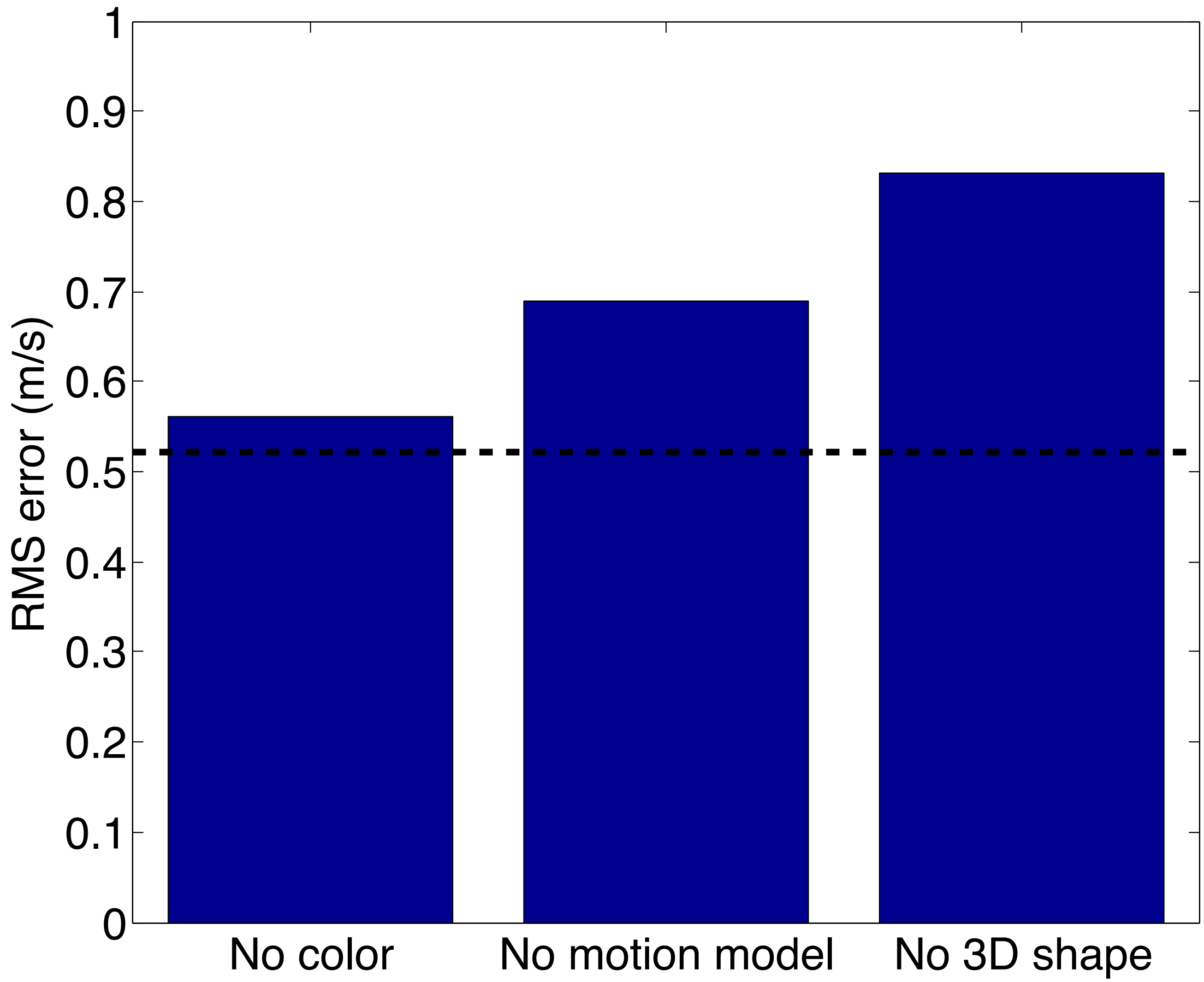


Error vs Number of Frames



Error vs Number of Frames





RMS error (m/s)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
No color
No motion model
No 3D shape