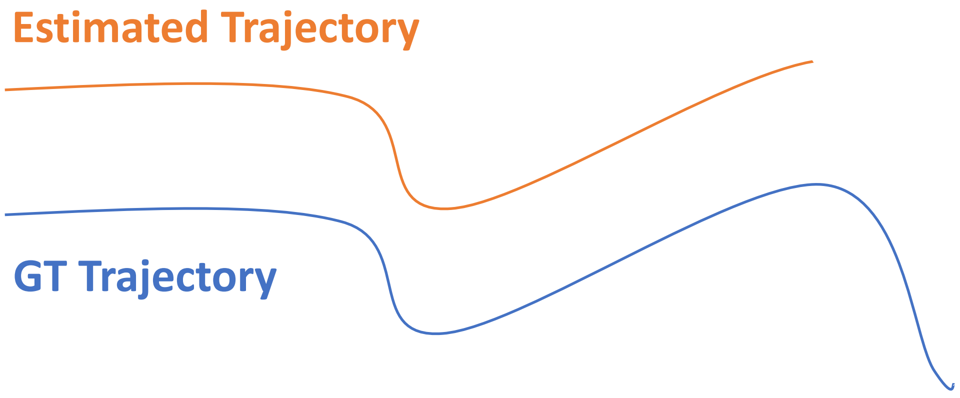
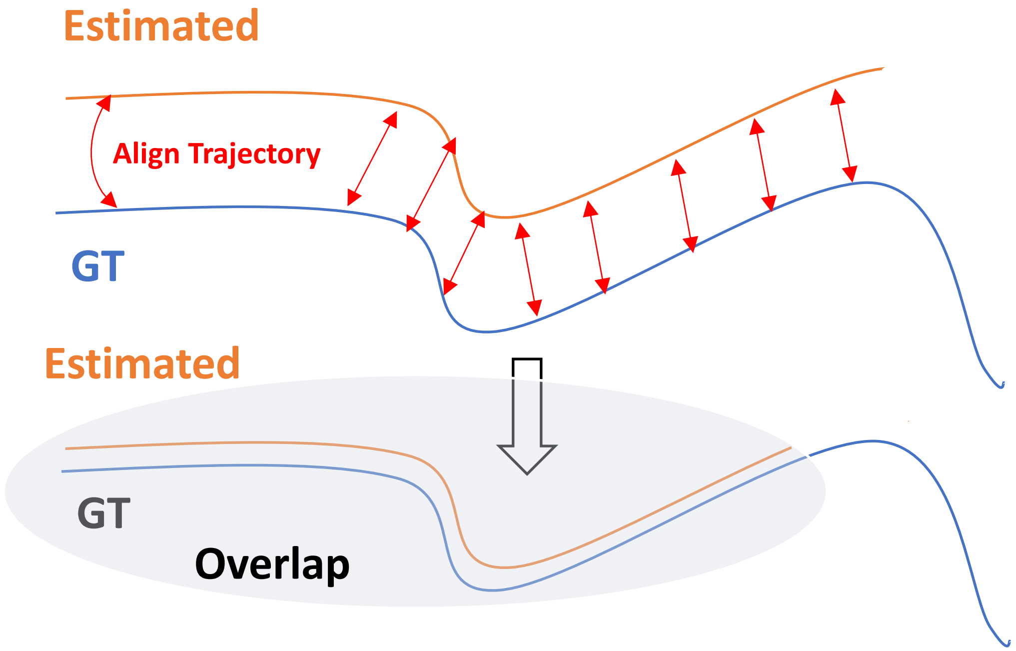
Why we want to Robustness Metrics? Consider traditional ATE,

Accuracy Metric:



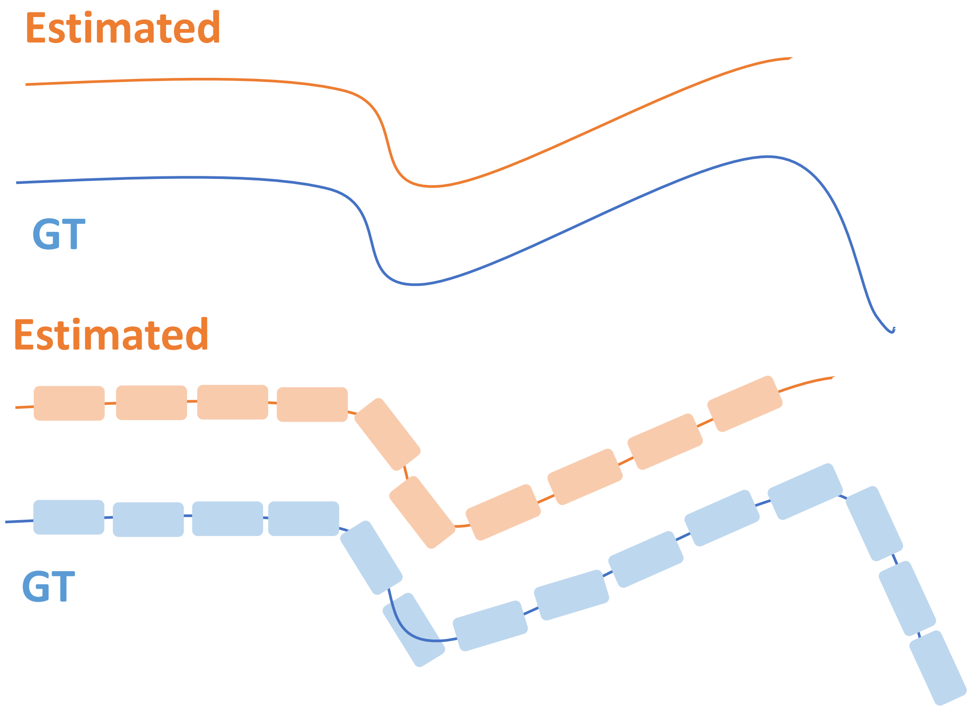


$$\mathrm{ATE}_{\mathrm{rot}} = \left(\frac{1}{N}\sum_{i=0}^{N-1} \|\angle(\Delta \mathrm{R}_i)\|^2\right)^{\frac{1}{2}},$$

$$\mathrm{ATE}_{\mathrm{pos}} = \left(\frac{1}{N}\sum_{i=0}^{N-1} \|\Delta \mathbf{p}_i\|^2\right)^{\frac{1}{2}},$$

Only consider overlap area, Not consider completeness **(Recall)**

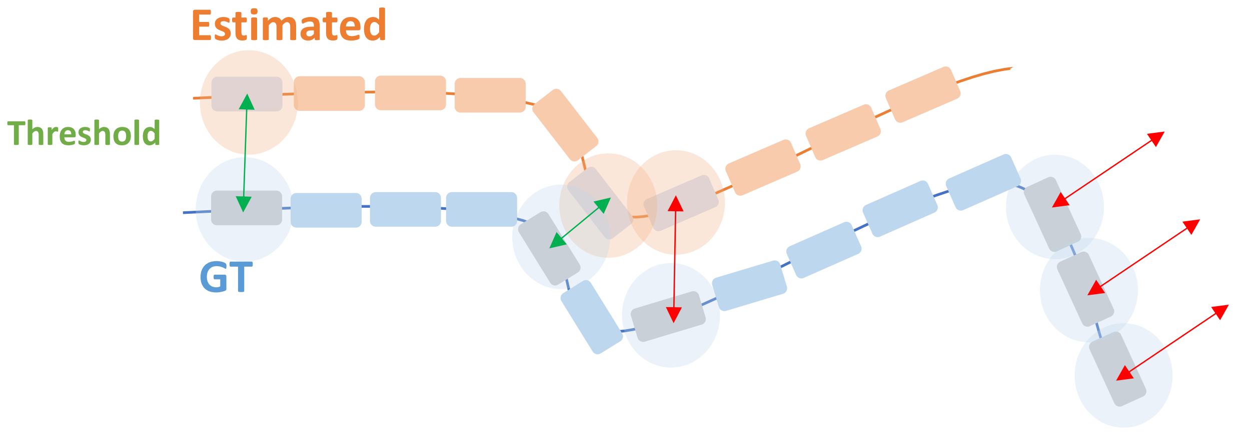
Robustness Metrics consider both precision and recall (completeness)!

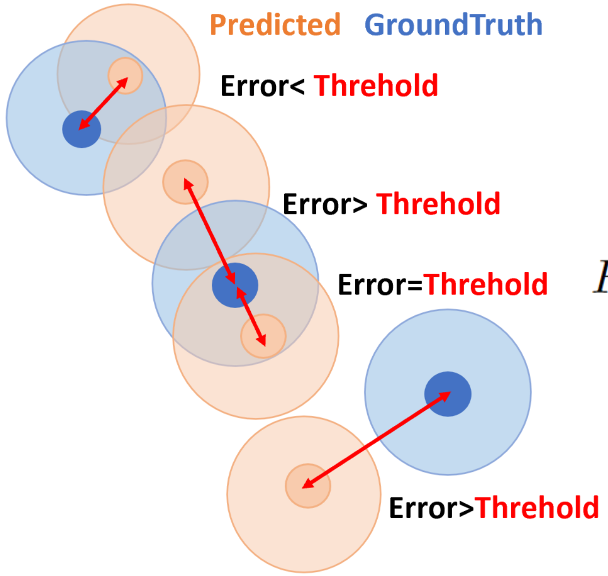


$$\mathrm{RPE}_{trans} = \|(\hat{p}_i - \hat{p}_{i-1}) - (p_i - p_{i-1})\|$$

$$\mathrm{RPE}_{rot} = \mathrm{Angle}((\hat{R}_i \hat{R}_{i-1}^T)(R_i R_{i-1}^T)^T)$$

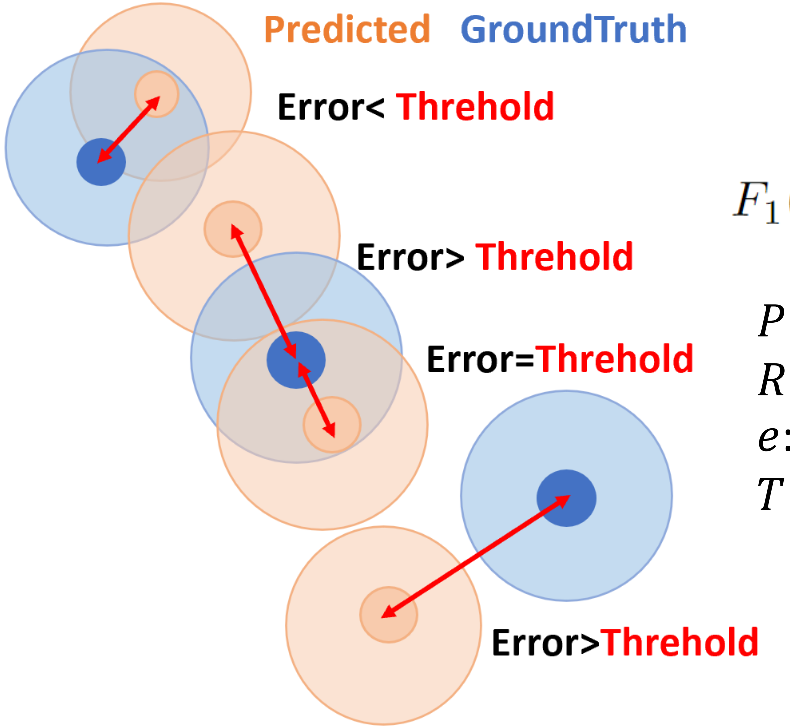
We incorporate RTE into our F_1 score





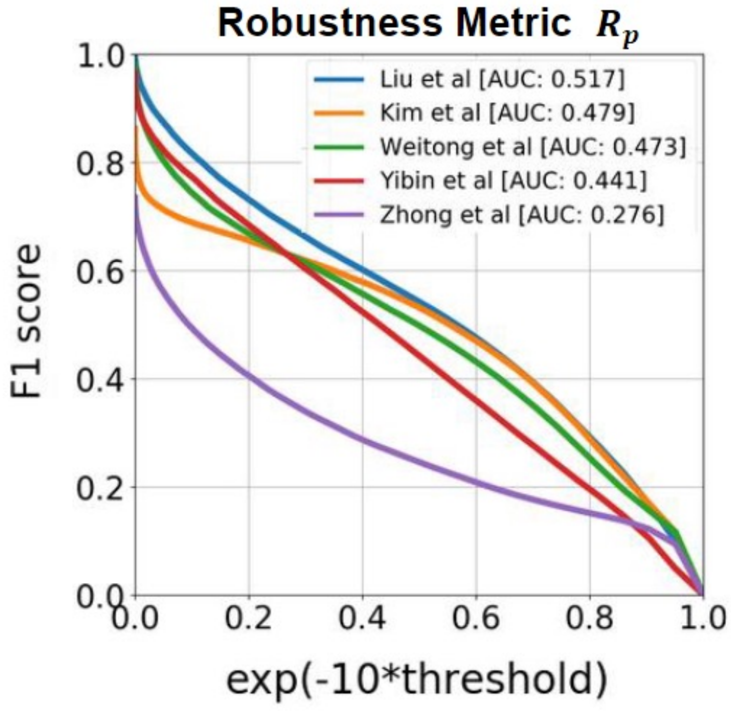
$$F_1(e) = \frac{2P(e<T)R(e<T)}{P(e<T)+R(e<T)},$$

Transform Abstract Robustness to Specific Quantifiable Value



$$F_1(e) = \frac{2P(e<T)R(e<T)}{P(e<T)+R(e<T)},$$

P: precision of RPE as percentage
R: RPE completeness
e: RPE error
T: Threshold [0,1]

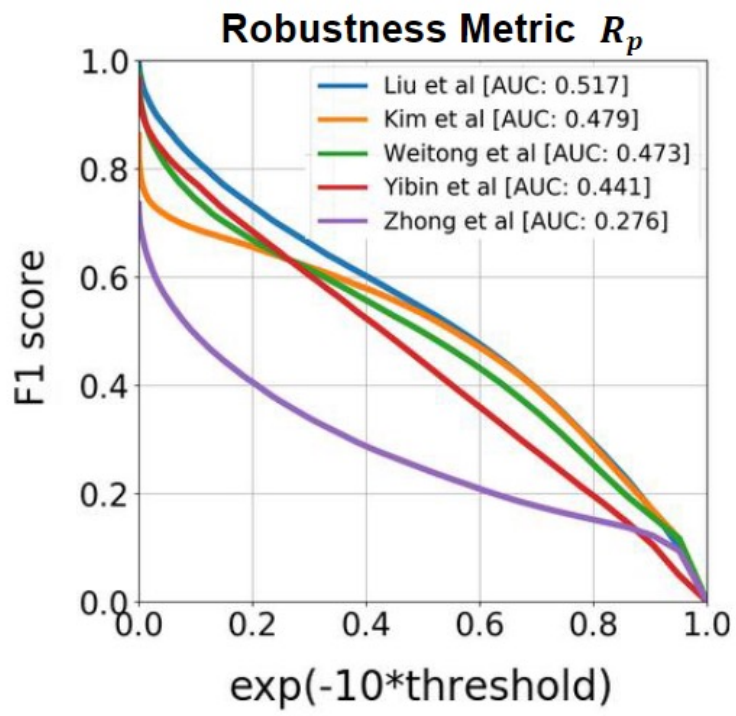


Scale up the threshold T to [0-1] by using

$$\exp(-10T)$$

Obtain the F score under different error tolearance. **The area under the curve (AUC)** is the robustness metric.

Transform Abstract Robustness to Specific Quantifiable Value



Obtain the F score in different threshold setting. **The area under the curve (AUC)** is the robustness metric.

Robustness Metric

$$F_1(e) = \frac{2P(e<T)R(e<T)}{P(e<T)+R(e<T)},$$

$$R_p = \mathrm{AUC}(\mathrm{F}_1(\mathrm{RPE}_{\mathrm{trans}}))$$

$$R_r = \mathrm{AUC}(\mathrm{F}_1(\mathrm{RPE}_{\mathrm{rot}}))$$