JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

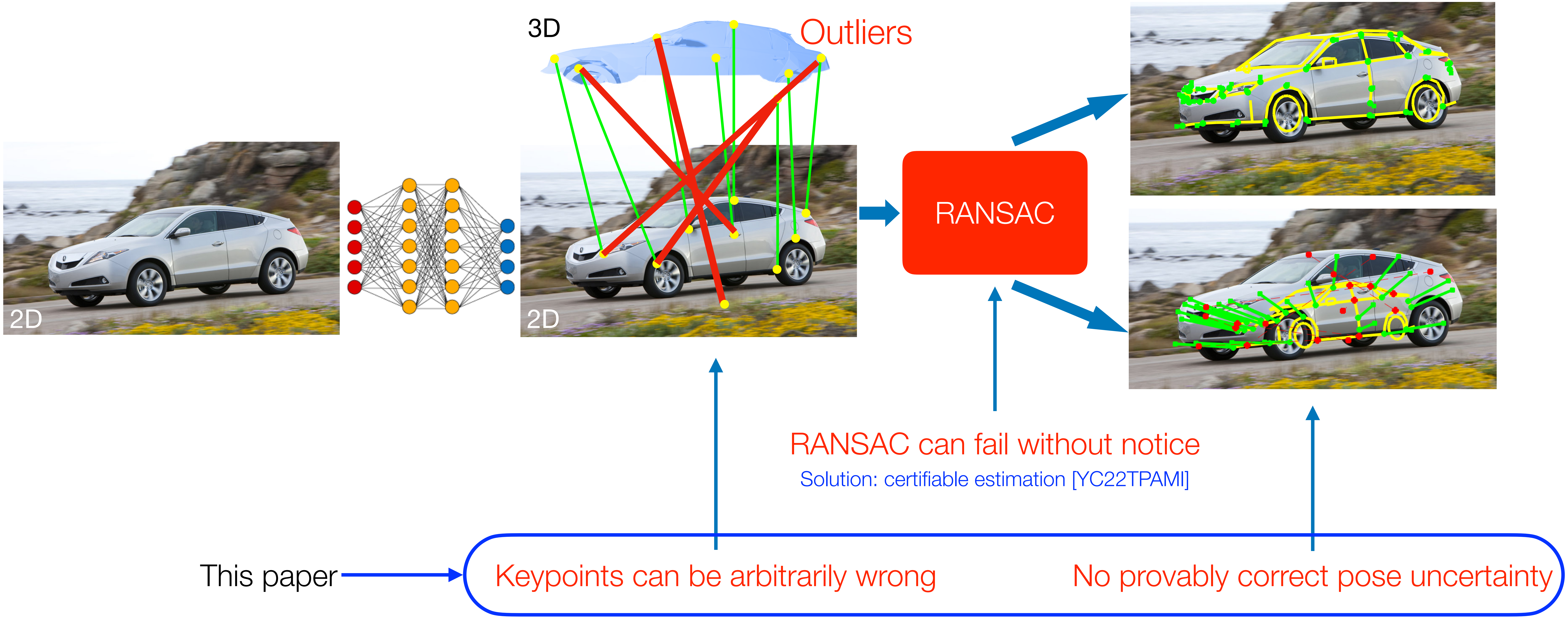
NVIDIA

Object Pose Estimation with Statistical Guarantees:
Conformal Keypoint Detection & Geometric Uncertainty Propagation

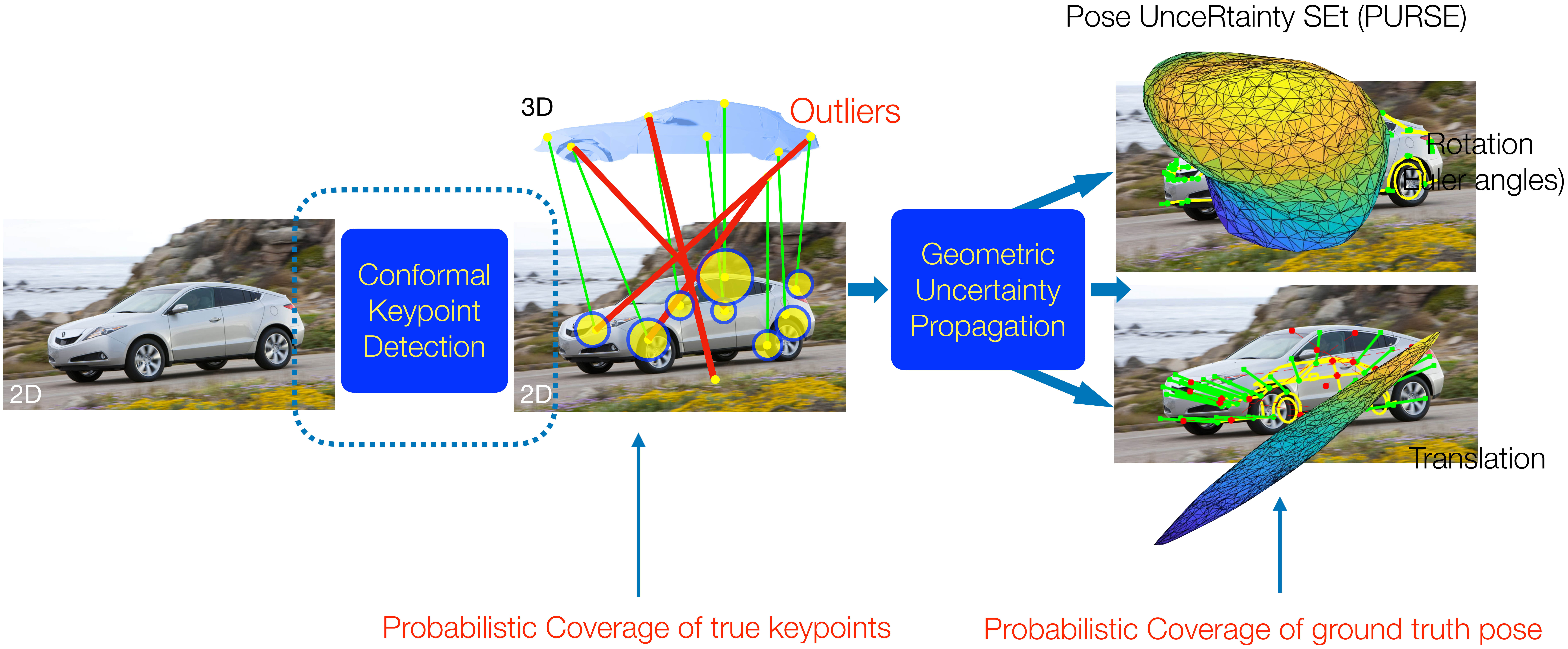
Heng Yang and Marco Pavone
NVIDIA Research

WED-AM-069

Object Pose Estimation: three challenges



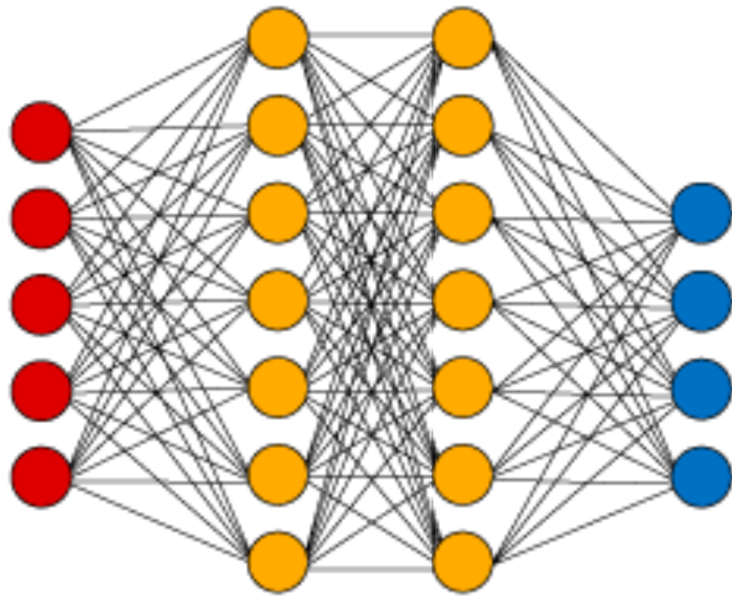
Object Pose Estimation with Statistical Guarantees



Conformal Keypoint Detection

$x \in \mathbb{R}^{H \times W \times 3}$

f



$y = (y_1, \ldots, y_k, \ldots, y_K) \in (\mathbb{R}^2)^K$

$q^\star$

$f_k(x)$

Nonconformity function

$r(y, f(x)) = \max\{\phi(y_k, f_k(x))\}_{k=1}^{K}$

$p_k^\star = \max\{f_k(x)\}$

$q_k^\star = \arg\max_{q \in x}\{f_k(x)\}$

$\phi = p_k^\star \|y_k - q_k^\star\|$

Calibration

Scores: $\{\alpha_{\pi(i)}\}_{i=1}^{n}$

New sample: x_{l+1}

Prediction

Quantile: $\alpha_{\pi(\lfloor (n+1)\epsilon \rfloor)}$

User: ϵ

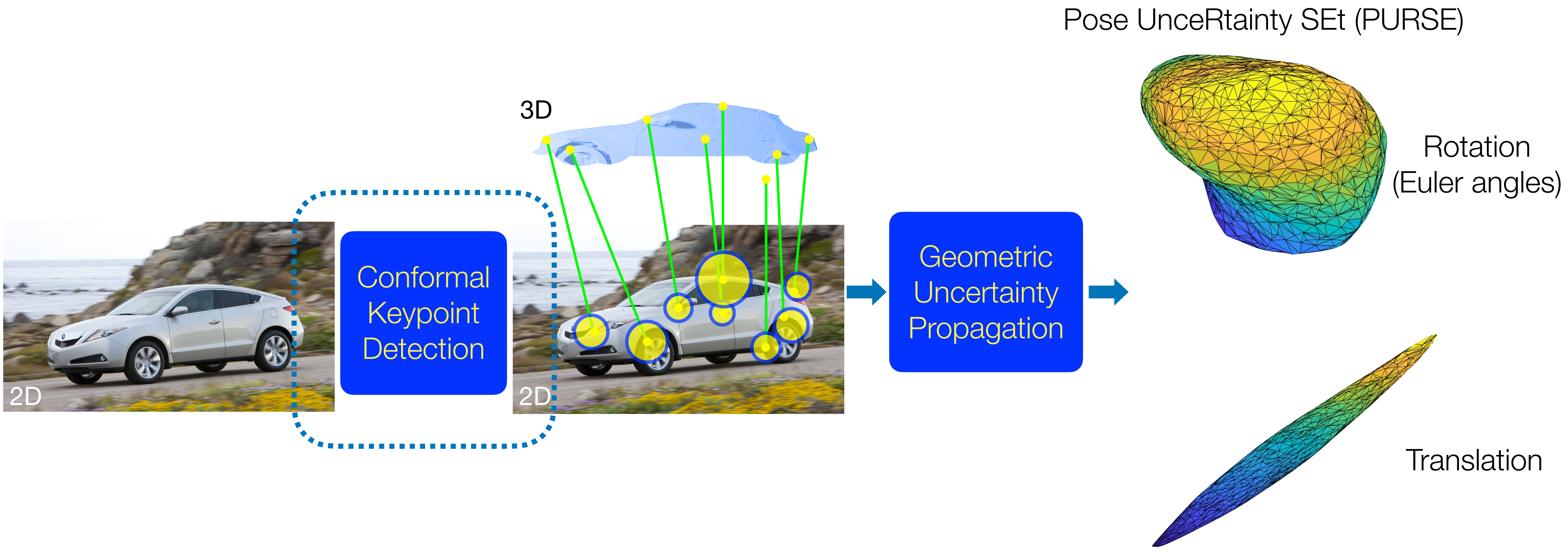
$\max\{\phi_1, \ldots, \phi_K\} \leq \alpha \Leftrightarrow \phi_k \leq \alpha, \forall k$

$$F^\epsilon(x_{l+1}) = \left\{ y \in (\mathbb{R}^2)^K \mid \|y_k - q_{l+1,k}^\star\| \leq \frac{\alpha_{\pi(\lfloor (n+1)\epsilon \rfloor)}}{p_{l+1,k}^\star}, \forall k \right\}$$

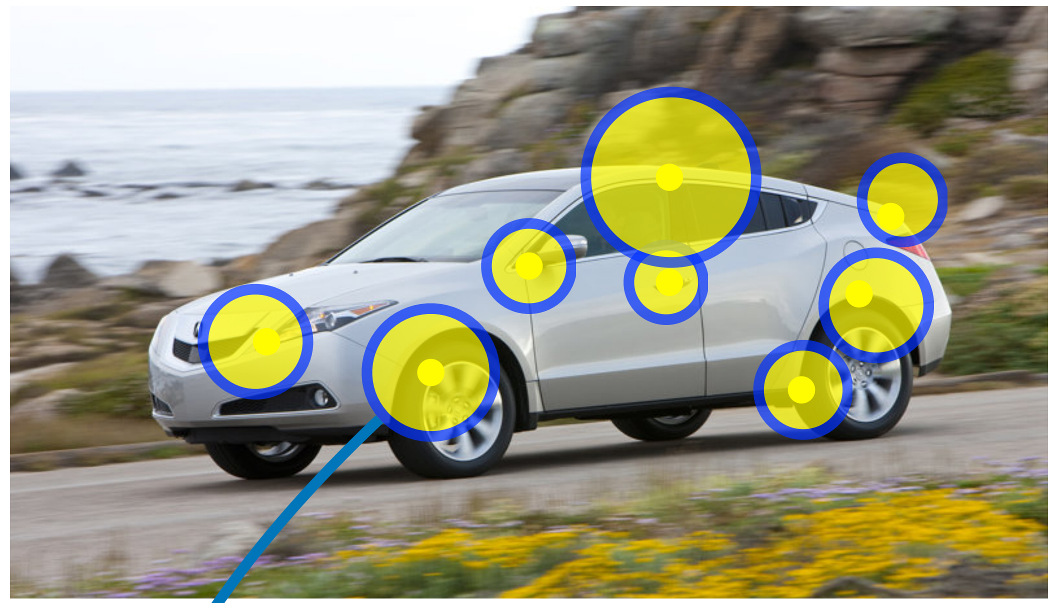
Center

Radius

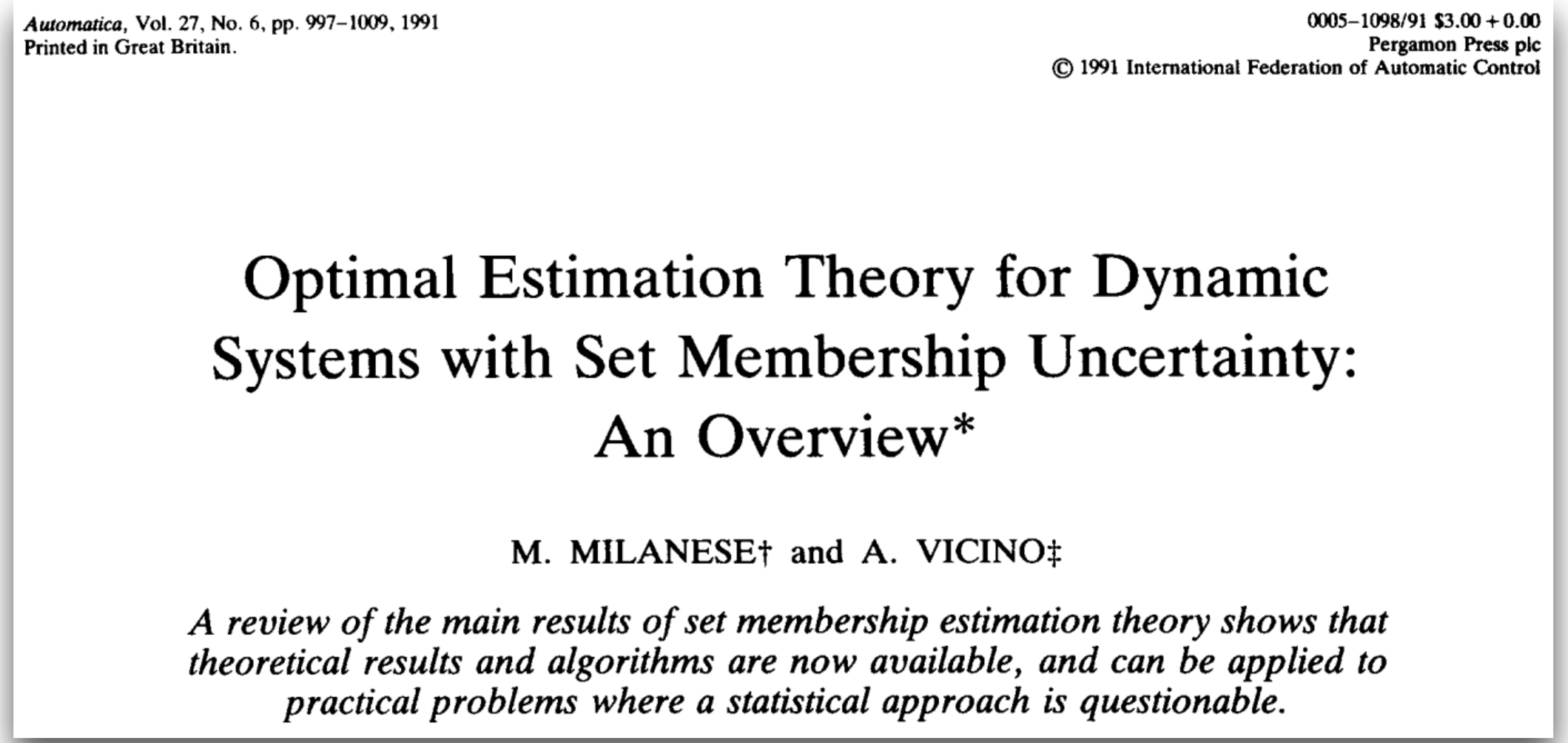
Object Pose Estimation with Statistical Guarantees



Geometric Uncertainty Propagation



$$F^{\epsilon}(x_{l+1}) = \left\{ y \in (\mathbb{R}^2)^K \mid \| y_k - q^{\star}_{l+1,k} \| \leq r_k, \forall k \right\}$$



Automatica, Vol. 27, No. 6, pp. 997–1009, 1991
Printed in Great Britain.
0005–1098/91 $3.00 + 0.00
Pergamon Press plc

Optimal Estimation Theory for Dynamic Systems with Set Membership Uncertainty: An Overview*

M. MILANESE† and A. VICINO‡

A review of the main results of set membership estimation theory shows that theoretical results and algorithms are now available, and can be applied to practical problems where a statistical approach is questionable.

With $(1-\epsilon)$ probability:

$$(y_k - q_k^{\star})^T \Lambda_k (y_k - q_k^{\star}) \leq 1, \forall k$$

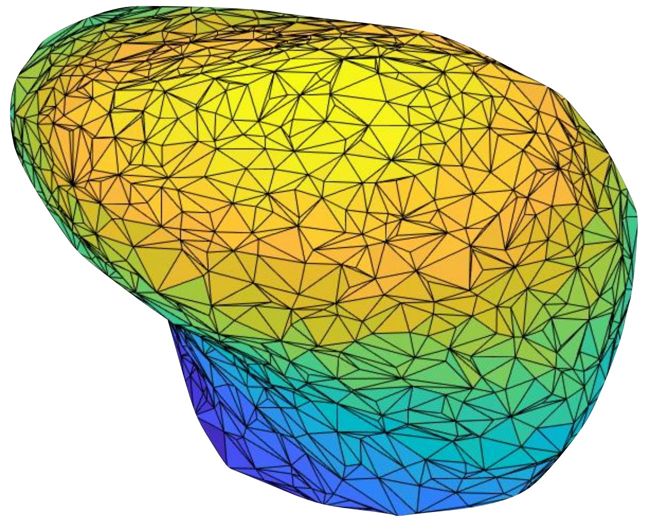
$$y_k = \Pi(RY_k + t)$$
$$s := [\text{vec}(R)^T, t^T]^T$$

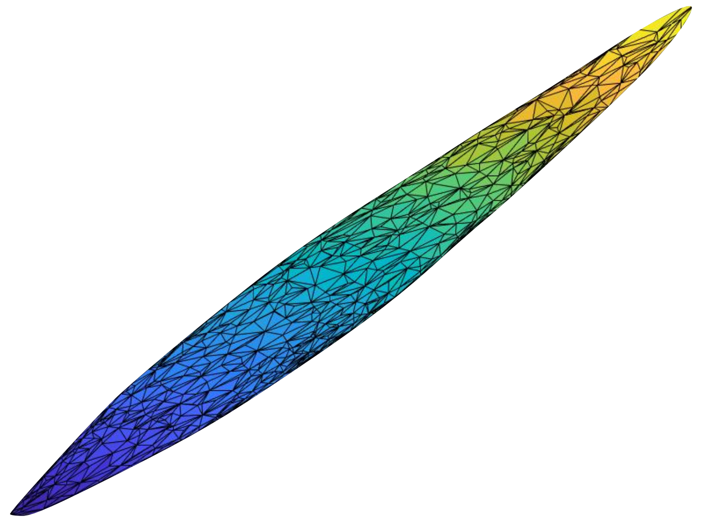
$$s^T A_k s \leq 0$$
$$b_k^T s > 0$$

With $(1-\epsilon)$ probability:

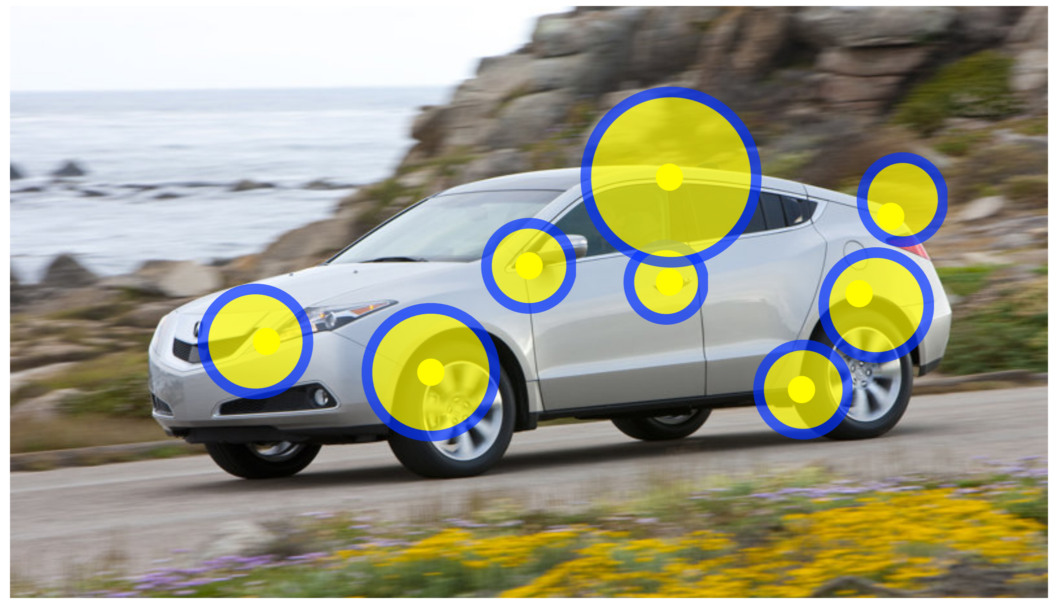
$$s \in S^{\epsilon} := \left\{ s \in \text{SE}(3) \mid s^T A_k s \leq 0,\ b_k^T s > 0, \forall k \right\}$$

Pose UnceRtainty SEt (PURSE)





Characterizing Pose Uncertainty Set



Pose UnceRtainty SEt (PURSE)

With $(1-\epsilon)$ probability:

$$s \in S^{\epsilon} := \left\{ s \in \text{SE}(3) \mid s^T A_k s \leq 0,\ b_k^T s > 0, \forall k \right\}$$

Nonconvex set !!

Algorithm: Random Sample Averaging (RANSAG)

1. Random choose 3 (out of K) prediction sets (ball or ellipse)
2. Random sample 2D keypoints $\{\hat{y}_{k_i}\}_{i=1}^{3}$ inside the prediction sets
3. $\{(R_j, t_j)\}_{j=1}^{4} = \text{SolveP3P}\left(\{\hat{y}_{k_i} \leftrightarrow Y_{k_i}\}_{i=1}^{3} \right)$
4. If $\{(R_j, t_j)\}_{j=1}^{4} \cap S^{\epsilon} \neq \emptyset$: return success; else: go back to step 1

- Check membership
- Generate samples and average

$$\max_{(R,t)\in S^{\epsilon}} \lambda \|R - R_{\text{avg}}\|_F^2 + (1-\lambda)\|t - t_{\text{avg}}\|^2$$

- Error bound

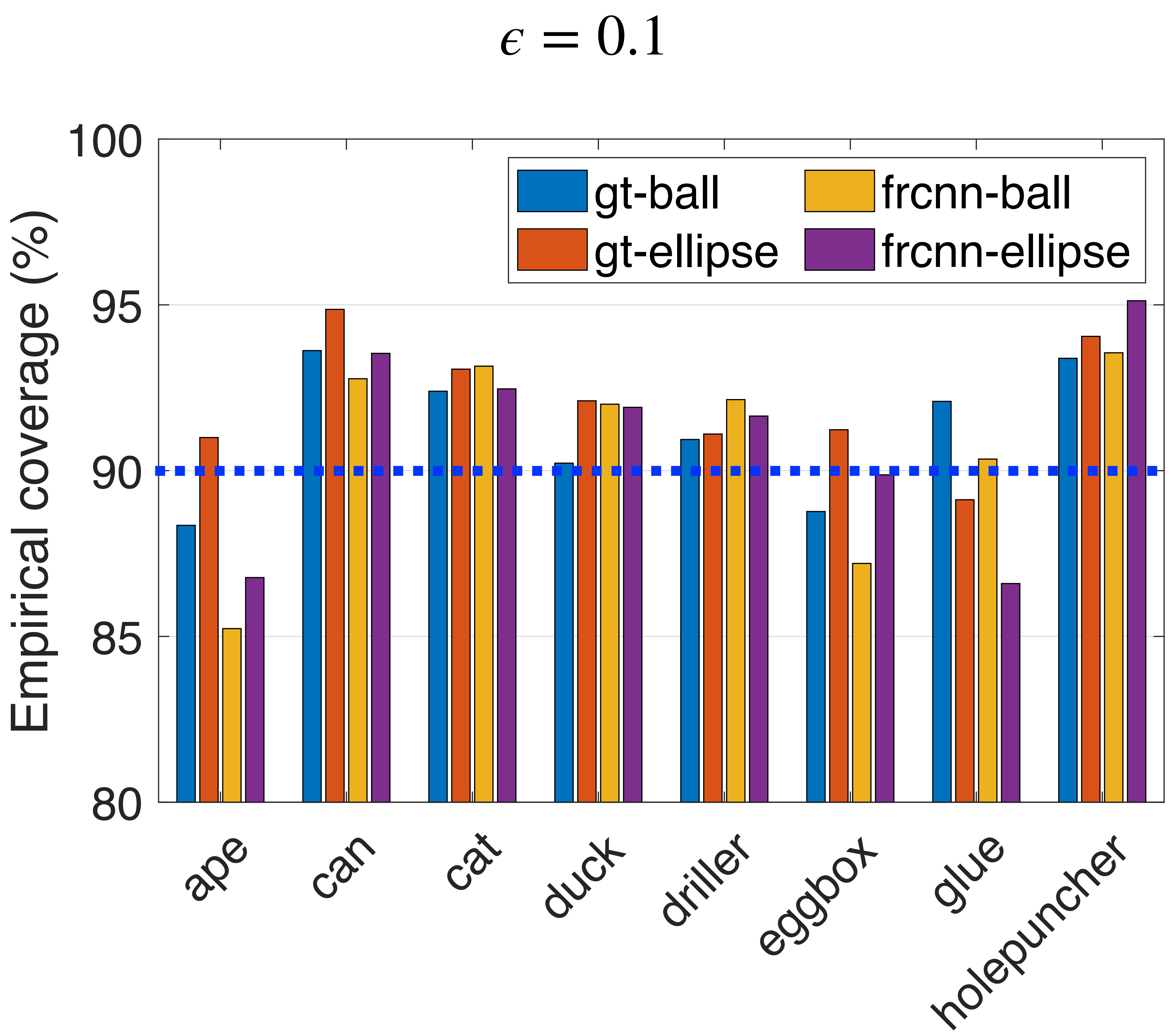
Semidefinite relaxation (sum of squares)
Exact global optimizer at order two

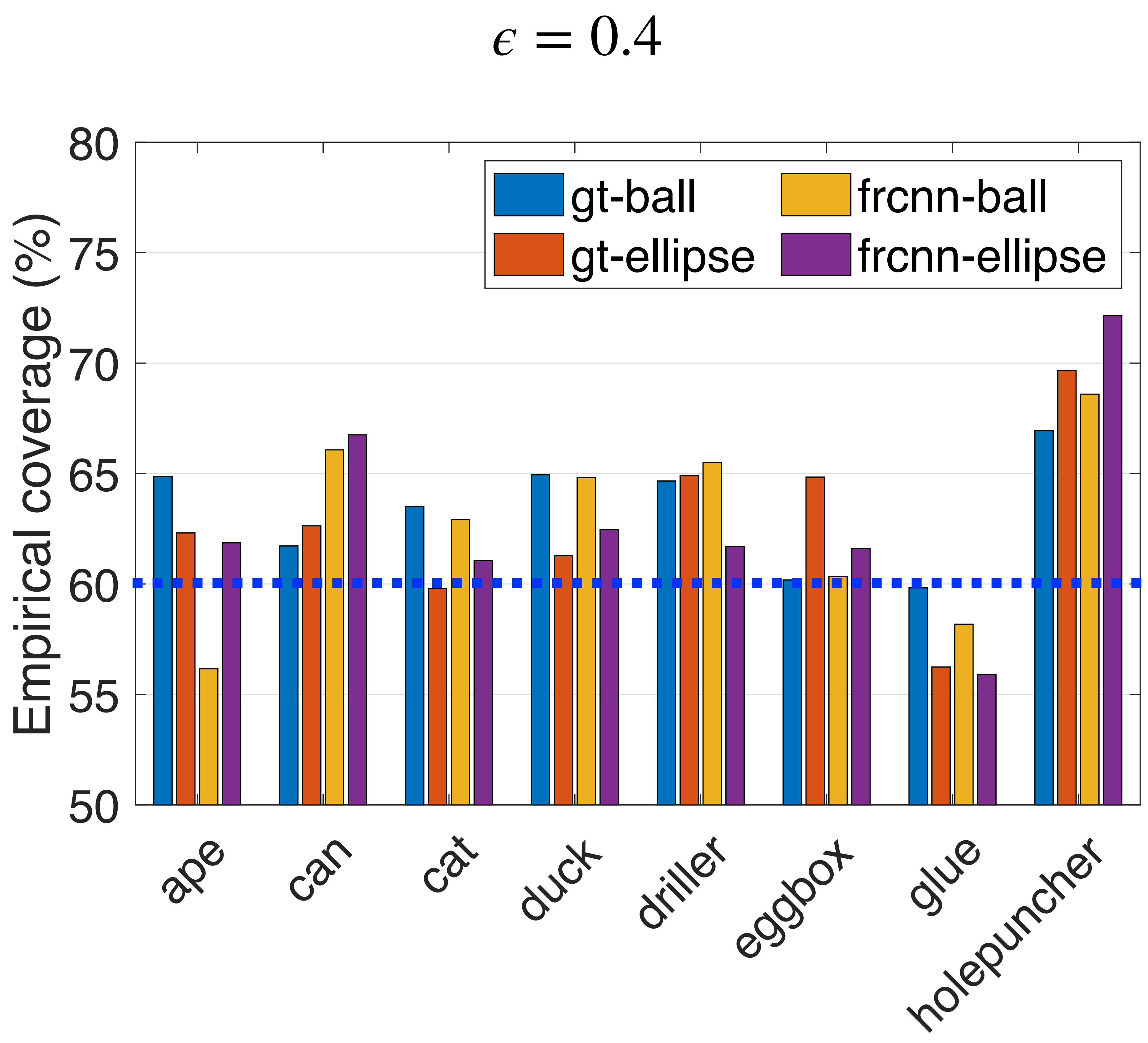
Average pose

Experiments

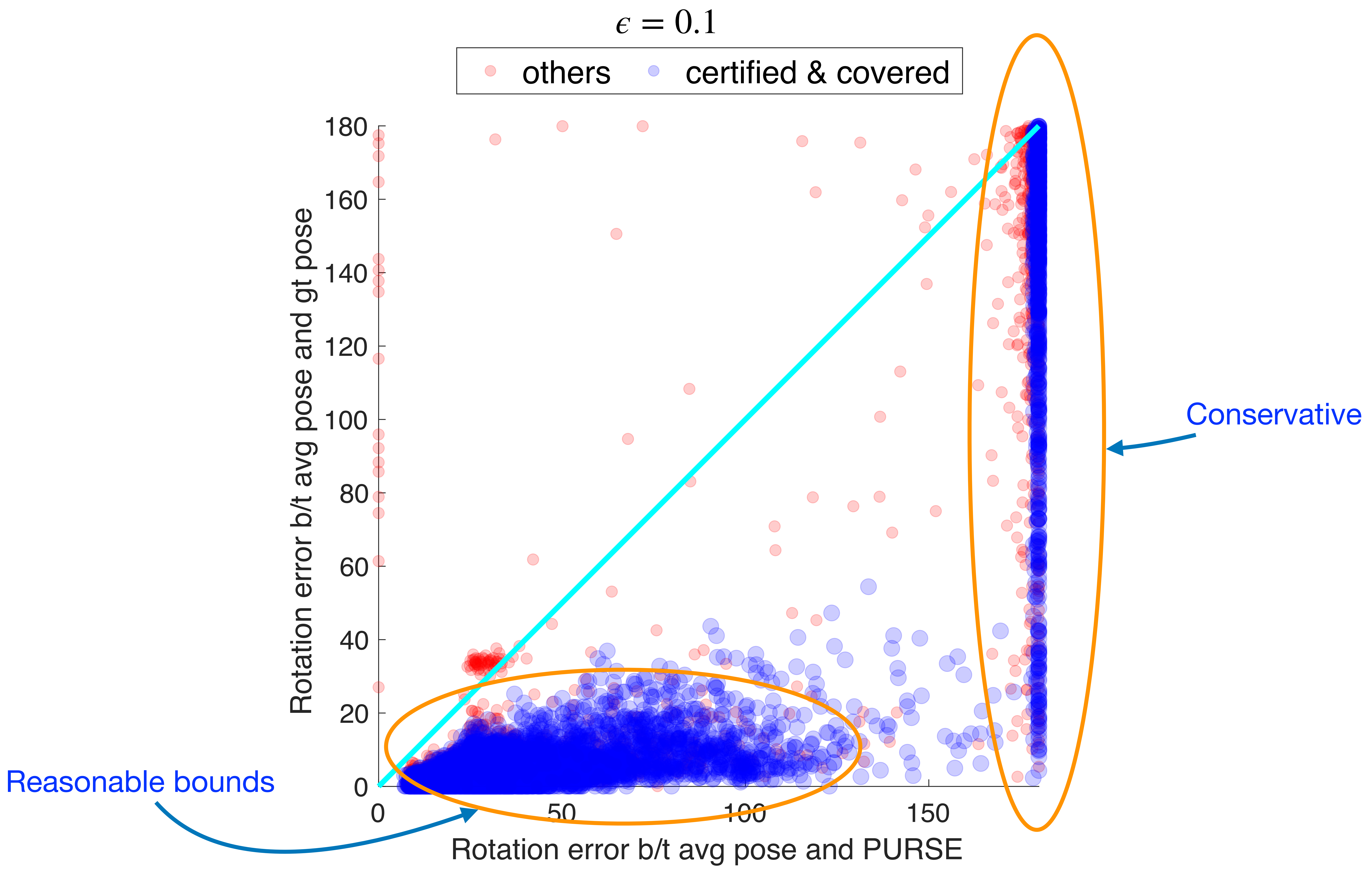
Calibration set size: 200 images; Test set size: 1214 images; $\epsilon = 0.1$ and $\epsilon = 0.4$

My PURSE contains your Pose

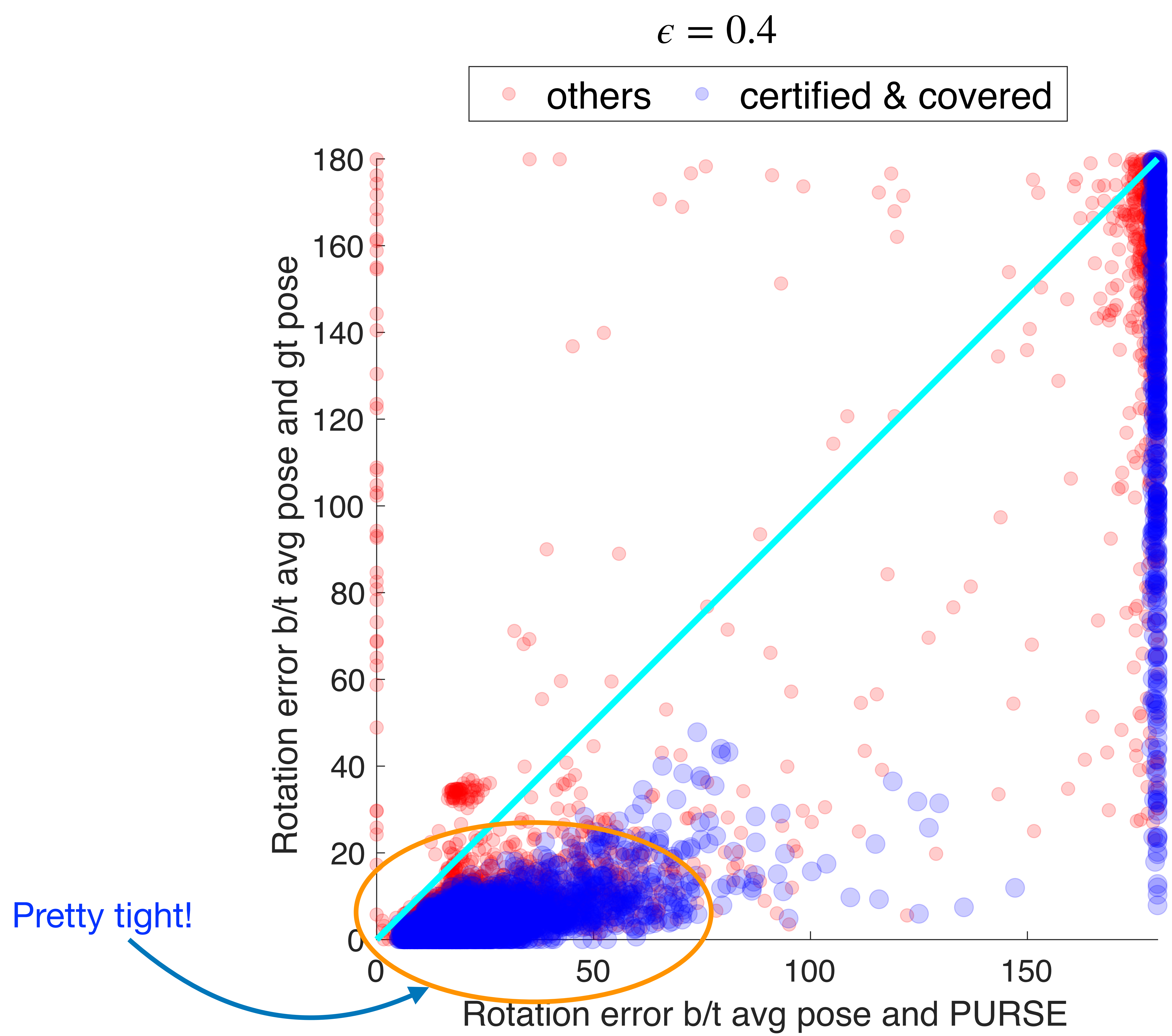




Worst-case pose errors



Worst-case pose errors

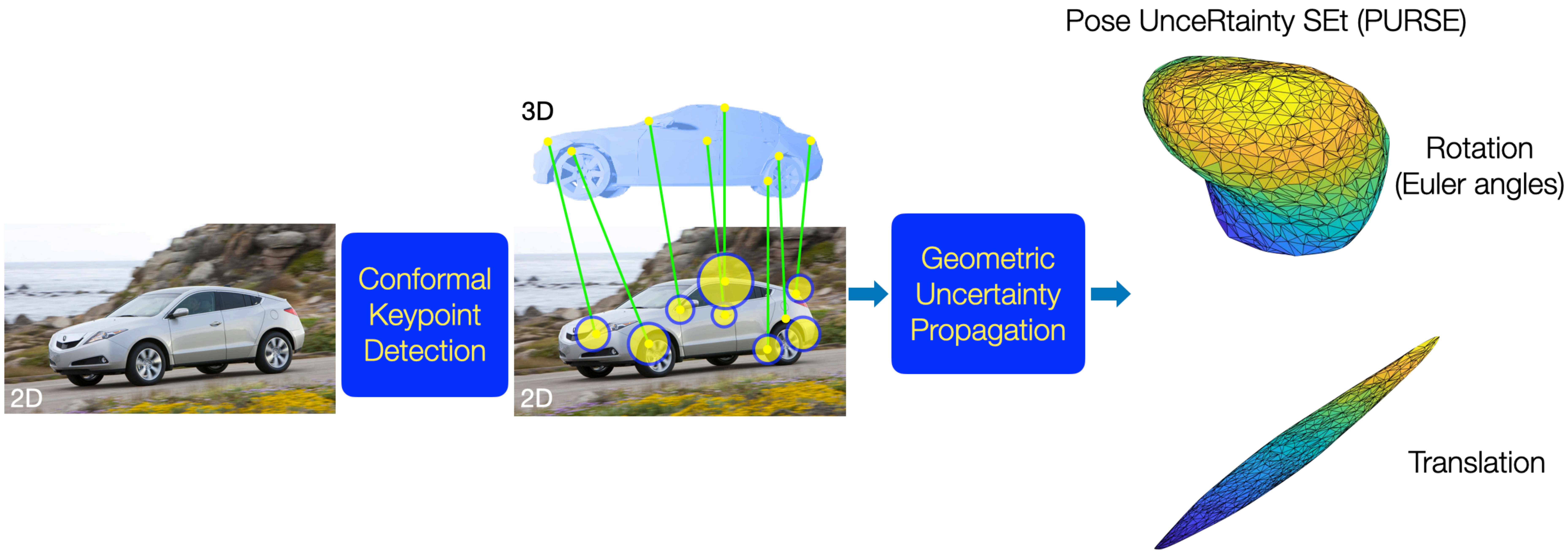


Average pose is accurate

objects	Baselines (results adapted from [26])				Conformalized heatmap							
	Tekin [33]	PoseCNN [36]	Oberweger [22]	PVNet [26]	gt-ball		gt-ellipse		frcnn-ball		frcnn-ellipse	
					$\epsilon = 0.1$	$\epsilon = 0.4$	$\epsilon = 0.1$	$\epsilon = 0.4$	$\epsilon = 0.1$	$\epsilon = 0.4$	$\epsilon = 0.1$	$\epsilon = 0.4$
ape	7.01	34.6	69.6	69.14	77.70	79.52	79.26	79.88	70.20	71.01	68.84	69.11
can	11.20	15.10	82.60	86.09	73.41	75.97	75.81	78.13	67.52	69.81	67.69	69.56
cat	3.62	10.40	65.10	65.12	87.36	90.59	89.54	90.11	74.95	80.23	68.98	78.57
duck	5.07	31.80	61.40	61.44	82.71	83.08	84.02	83.55	79.30	80.62	80.06	80.53
driller	1.40	7.40	73.80	73.06	79.32	82.54	81.22	82.04	58.48	65.92	58.06	65.67
eggbox	-	1.90	13.10	8.43	0	0	0.09	0.18	0	0	0	0.14
glue	4.70	13.80	54.90	55.37	56.49	71.08	71.69	72.93	30.03	47.18	41.96	48.26
holepuncher	8.26	23.10	66.40	69.84	81.65	82.89	83.22	84.30	74.96	77.85	76.28	78.18
average	6.16	17.20	60.90	61.06	67.33	70.71	70.61	71.39	56.93	61.58	57.73	61.25

Conclusions

- Object Pose Estimation with Statistical Guarantees
  - Conformal keypoint detection: simple circular or elliptical prediction sets
  - Geometric uncertainty propagation: Pose Uncertainty Set (PURSE)



Poster: **WED-AM-069**