IEEE ICIP 2017

# Efficient estimation of target detection quality


**Juan C. SanMiguel**
Escuela Politécnica Superior
Universidad Autónoma de Madrid

www-vpu.eps.uam.es/~jcs

@jc_sanmiguel

UAM UNIVERSIDAD AUTONOMA DE MADRID

**Andrea Cavallaro**
Centre for Intelligent Sensing
Queen Mary University of London

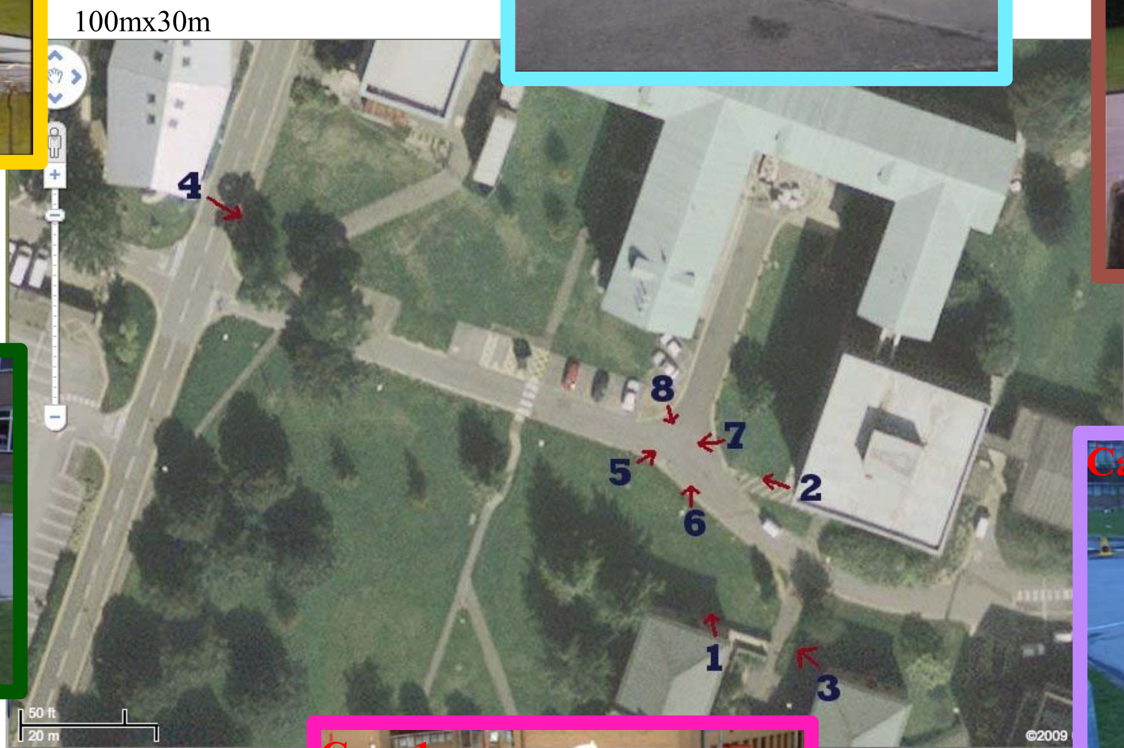
www.eecs.qmul.ac.uk/~andrea

@smartcameras

Queen Mary University of London

# Introduction: multi-camera processing







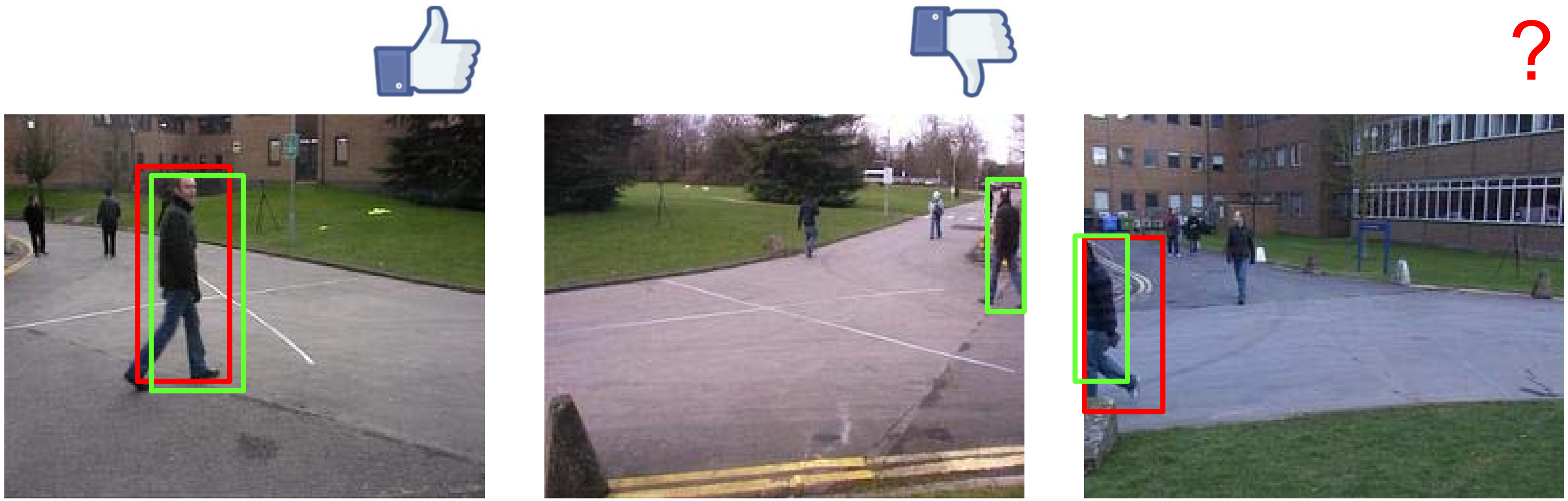
Fields of view (and data) are projected to a common plane for multi-camera processing

Detecting targets is key for many applications

Images credit: http://www.cvg.reading.ac.uk/PETS2009/

# What is detection quality?

- Detecting targets is key for many multicamera approaches
- Detection quality
  - models the miss-detection rate
  - related to the probability of a target to be detected within the FOV
  - accounts for the number of undetected targets over time



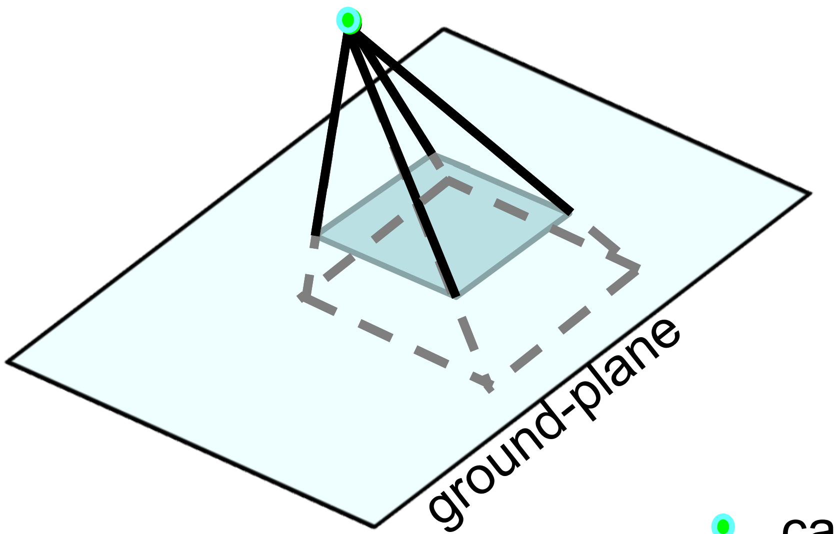

manual annotation

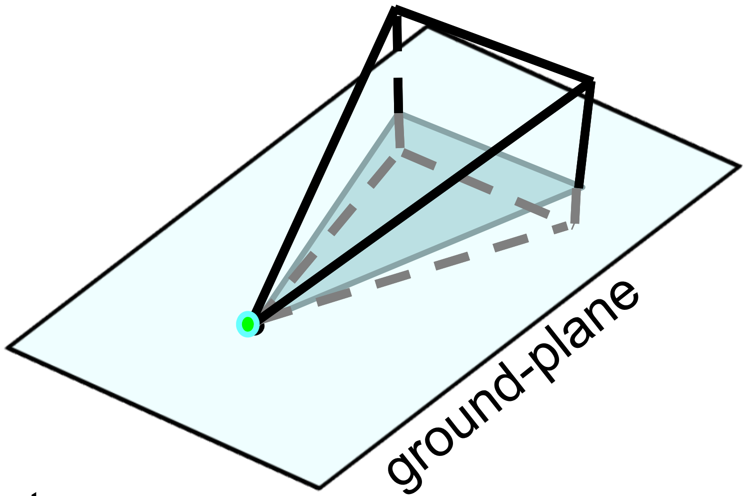
detector result

# Field Of View (FOV) models

- FOV projection
  - geometrical properties of the FOV
  - widely accepted abstraction: projection onto a plane

- Shape of the FOV projection
  - square (aerial views)
    - enable fast computation of detection quality
  - triangular (pinhole camera)
    - more common
    - does not lend itself to fast computation

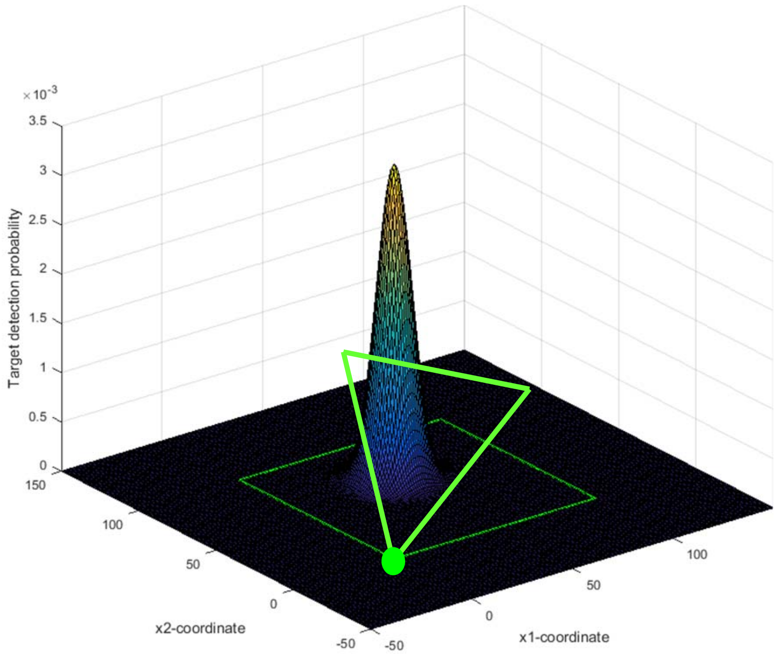


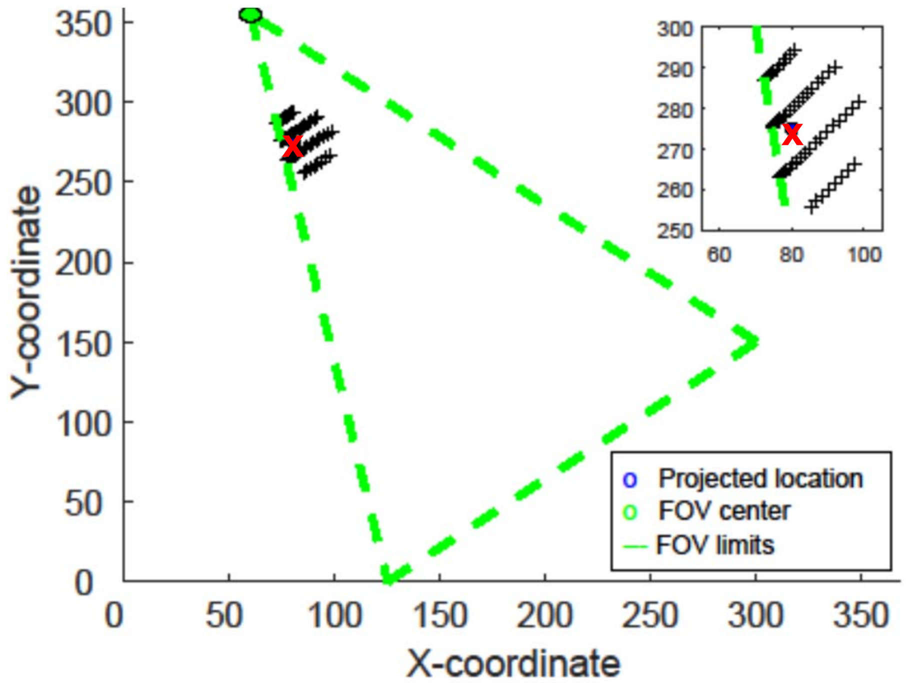




camera center
projected camera FOV

# Detection quality for triangular FOVs

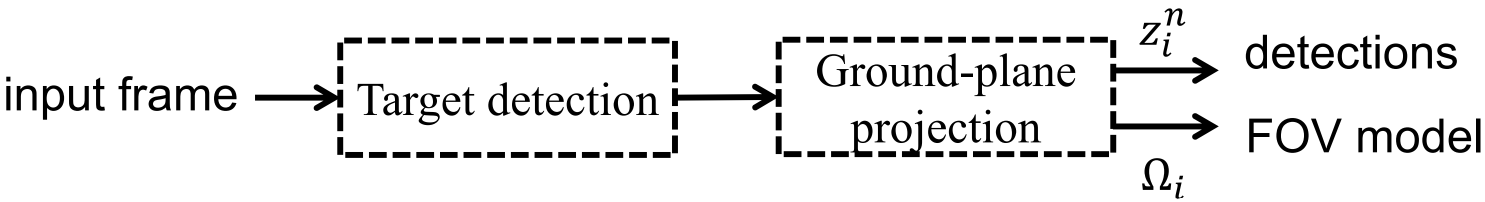
- Main ideas of our proposal
  - mapping the problem to a definite integral considering uncertainty
  - integration domain: represents the FOV
  - integral: numerically approximated by combining quadrature-based integration and importance sampling



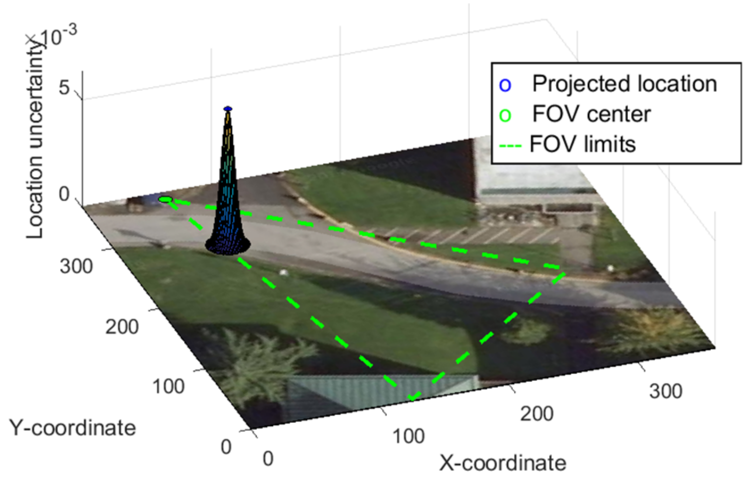




camera FOV  camera center  target location

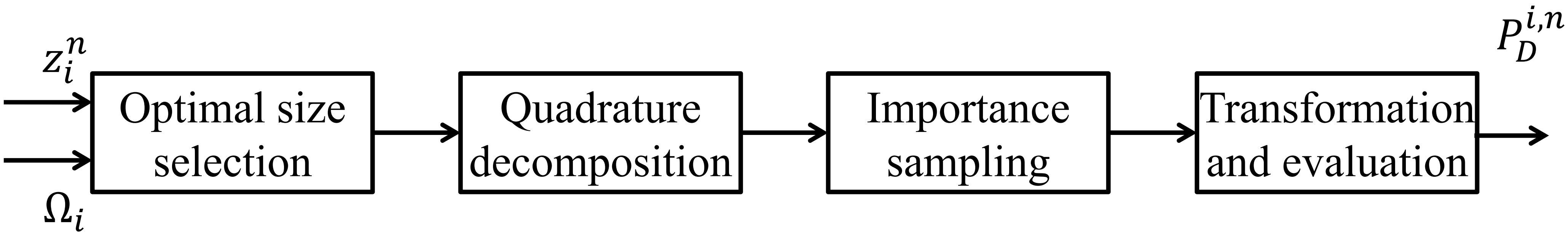
# Proposal: overview



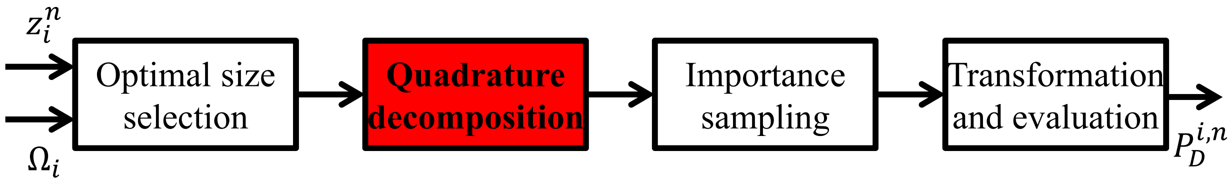

input to estimate detection quality


- Pipeline to estimate detection quality via integration



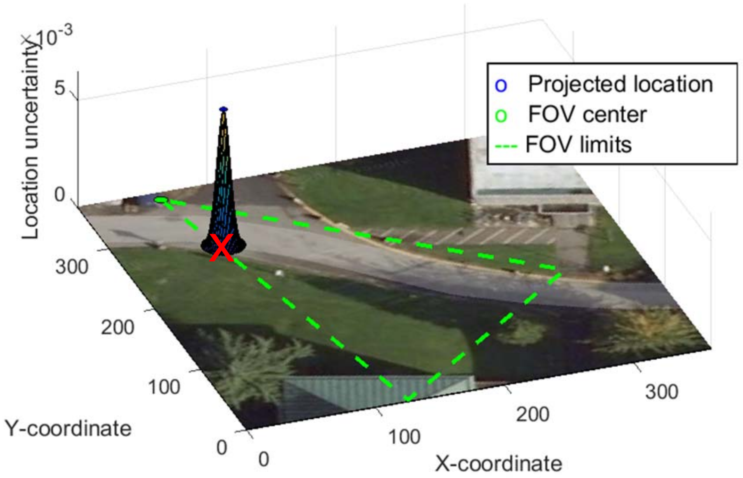
# Proposal (1/4)

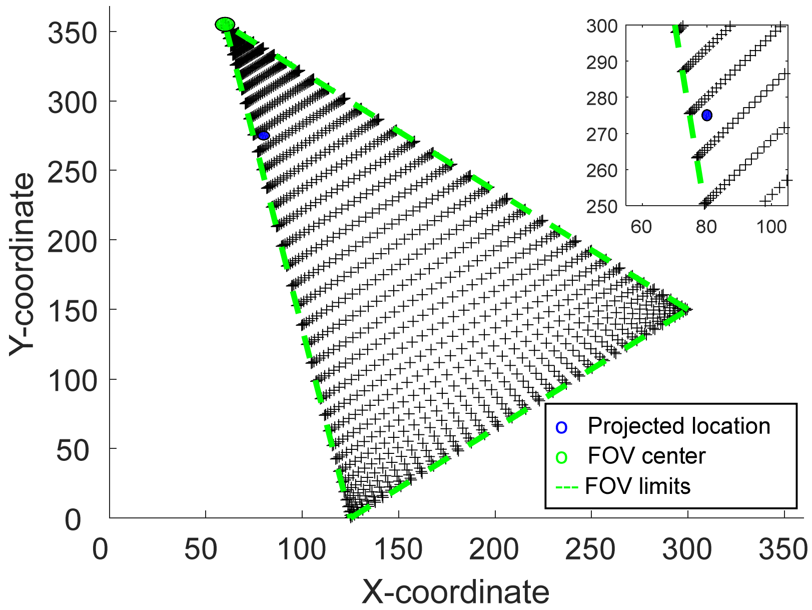



- Quadrature decomposition
  - Integration of detection uncertainty $f(z_i)$
  - Gaussian quadrature: tabulated weights $w_r$ and normalized $\hat{z}_r$

$$\int_{\Omega_i} f(z_i)d\boldsymbol{x} \approx \sum_{r=1}^{N} w_r\, f(\hat{z}_r)$$

a weighted sum of N function values at specified points $\hat{z}_r$ within the domain of integration

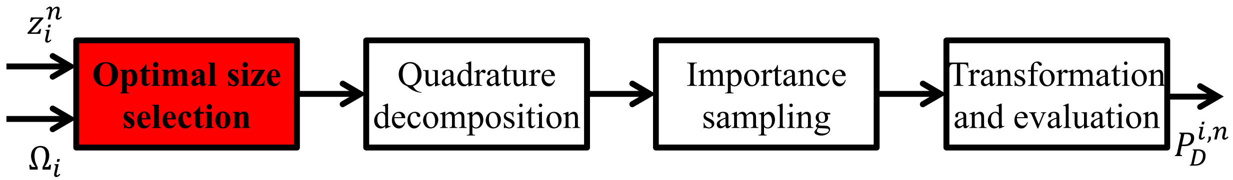



target location $z_i$

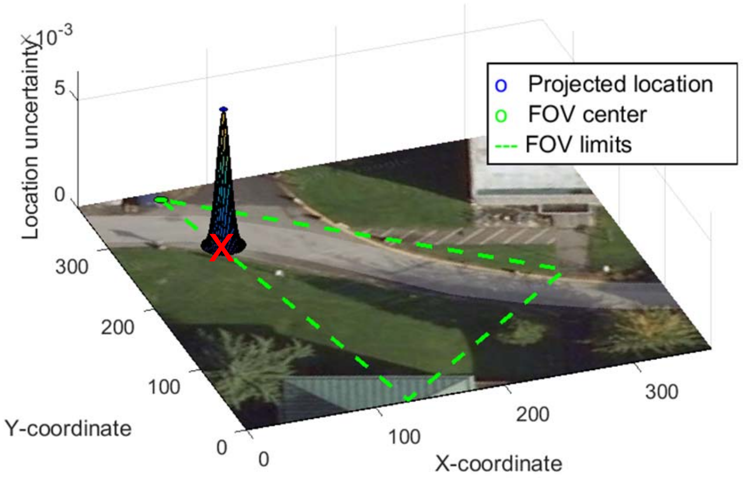
camera FOV $\Omega_i$

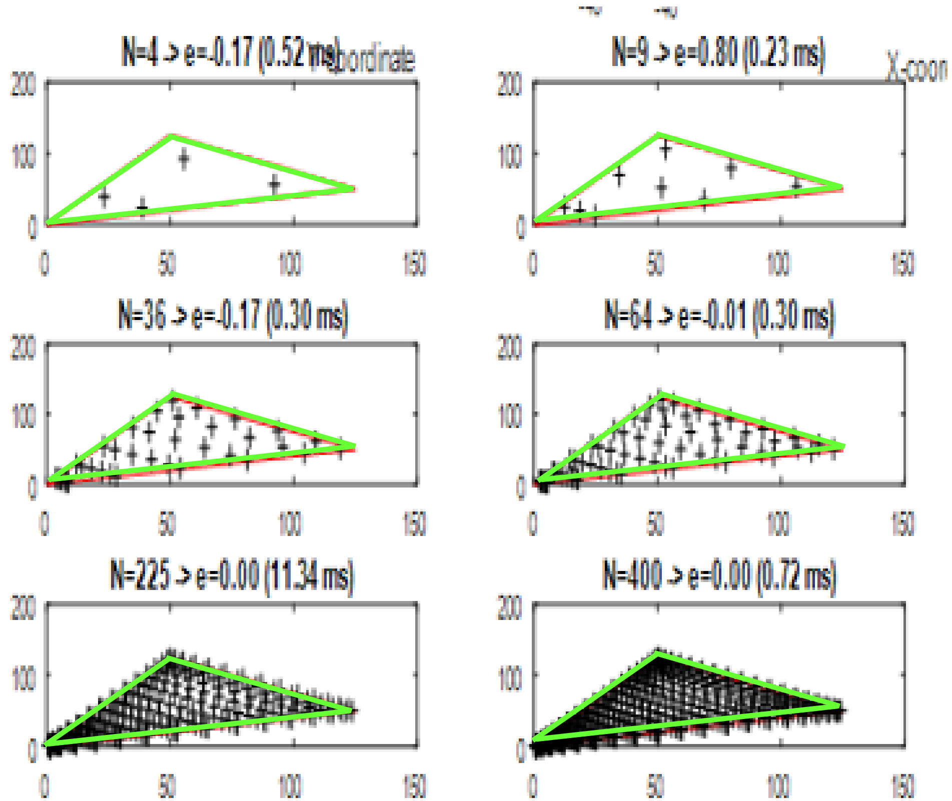
# Proposal (2/4)




- Optimal size selection
  - Number of samples to integrate according to
    - FOV size $FOV_i$
    - detection uncertainty $\Sigma$
  - Heuristic minimization approach

$$N^* = \underset{N}{\operatorname{argmin}} \left| N - \frac{FOV_i}{\Sigma} \right|$$

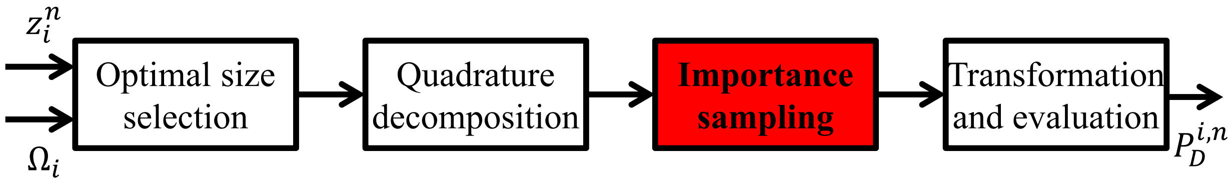



target location $z_i$

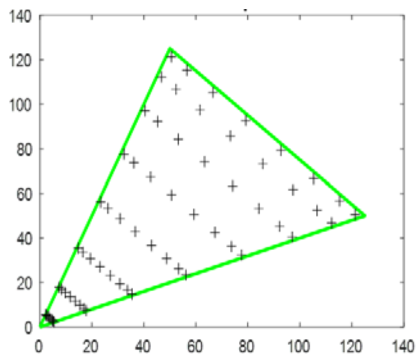
camera FOV $\Omega_i$

# Proposal (3/4)

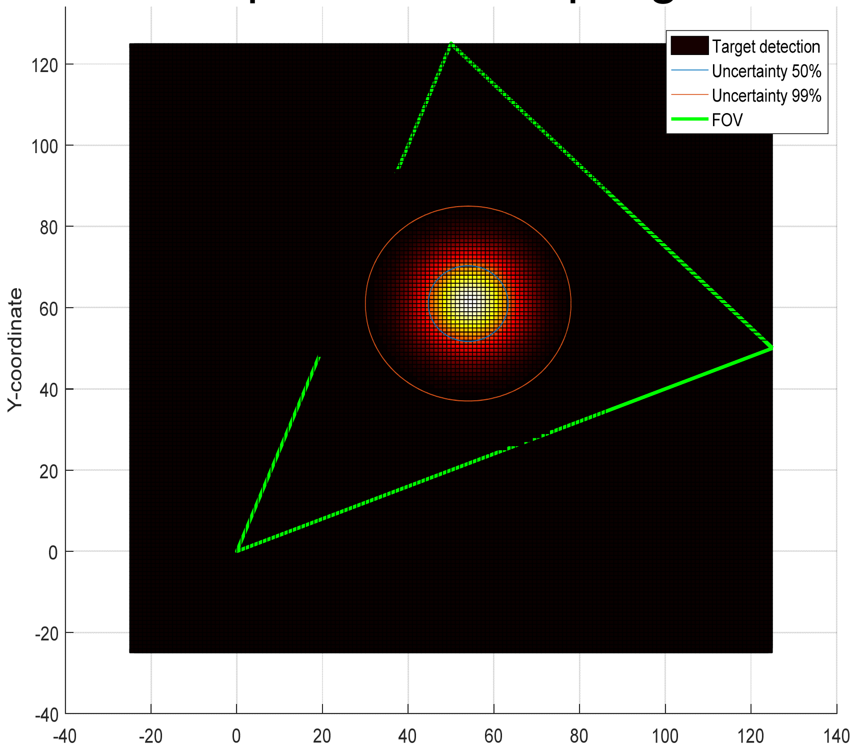



- Importance sampling
  - place higher density of samples where integrand is large
  - location uncertainty for targets handled by covariance matrix
    - such covariance may guide the location of quadrature samples
    - it can be approximated by a polygon: an ellipse

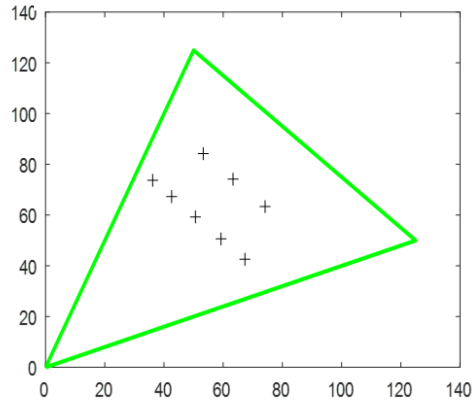
Initial quadrature points



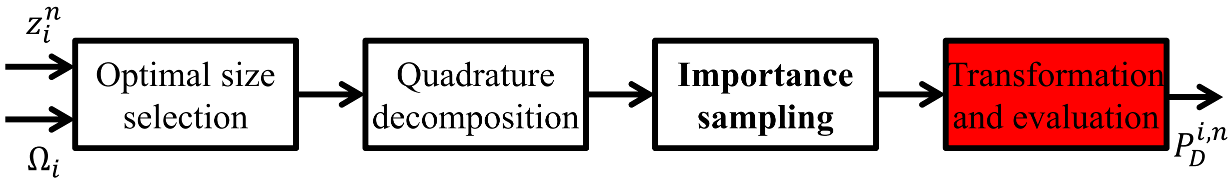

Importance sampling




Selected points

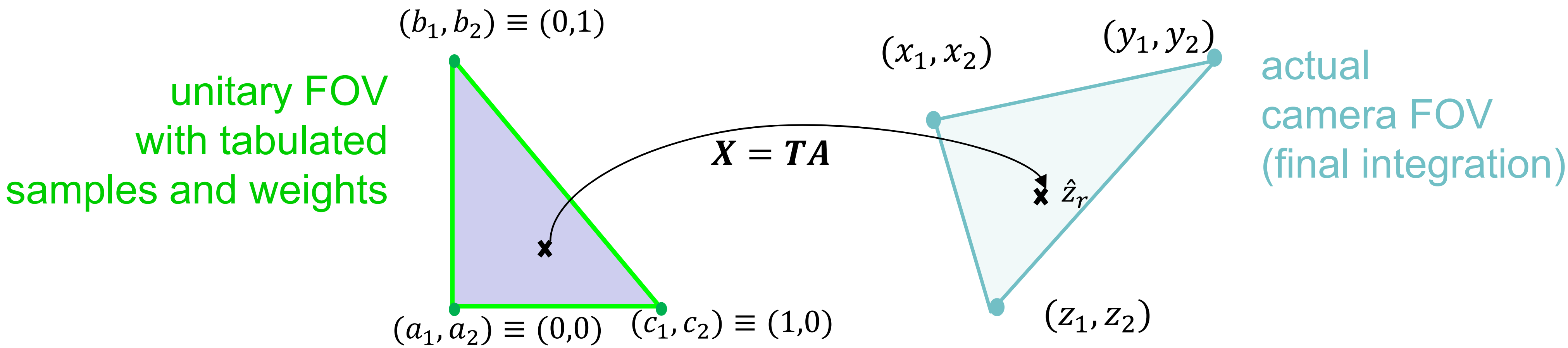


# Proposal (4/4)



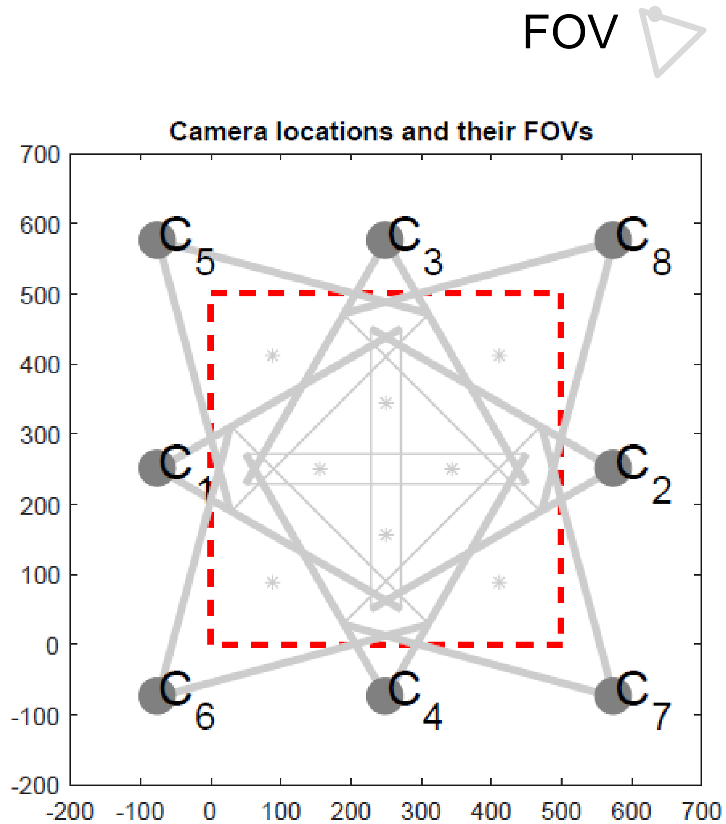

- Transformation and evaluation
  - adapt the coordinates of selected samples to the desired FOV
  - affine transformation $X = T \cdot A$ to obtain samples in FOV
    - transformation $T$ is found by $T = X \cdot A^{-1}$

$$\begin{bmatrix} t_1 & t_2 & t_3 \\ t_4 & t_5 & t_6 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ 1 & 1 & 1 \end{bmatrix}^{-1}$$

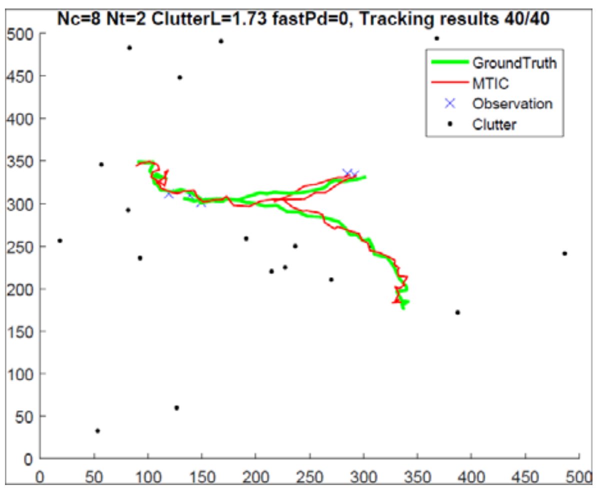


# Evaluation on multi-target tracking

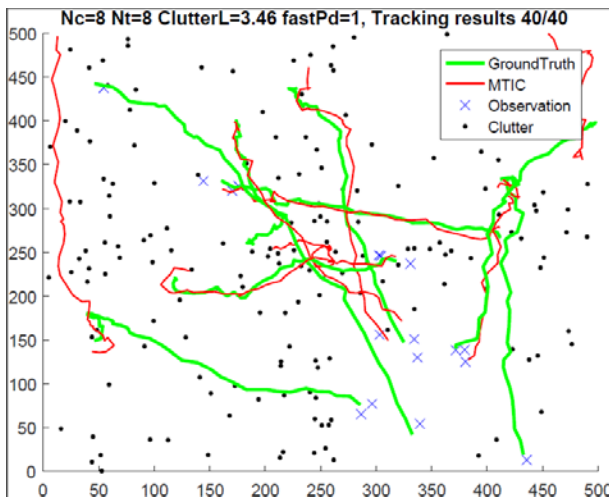
- Task
  - to estimate detection quality $P_D$
  - monitored área: 500x500
  - 8 cameras
  - triangular FOV projection
  - target state vector: position and velocity
  - Gaussian models for uncertainty
  - results over $6 \cdot 10^5$ runs



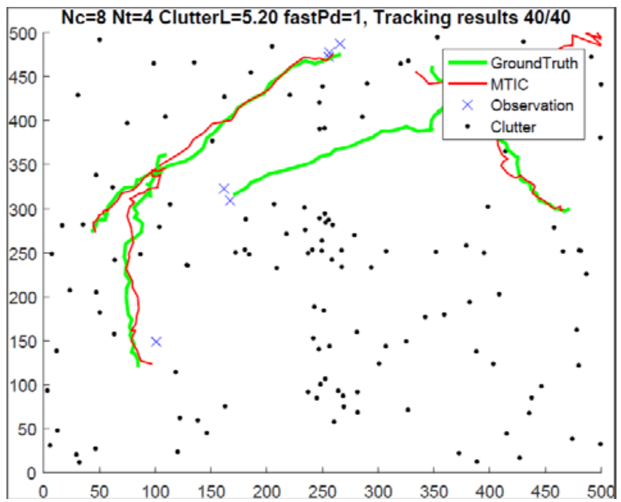

low clutter, 2 targets




medium clutter, 8 targets

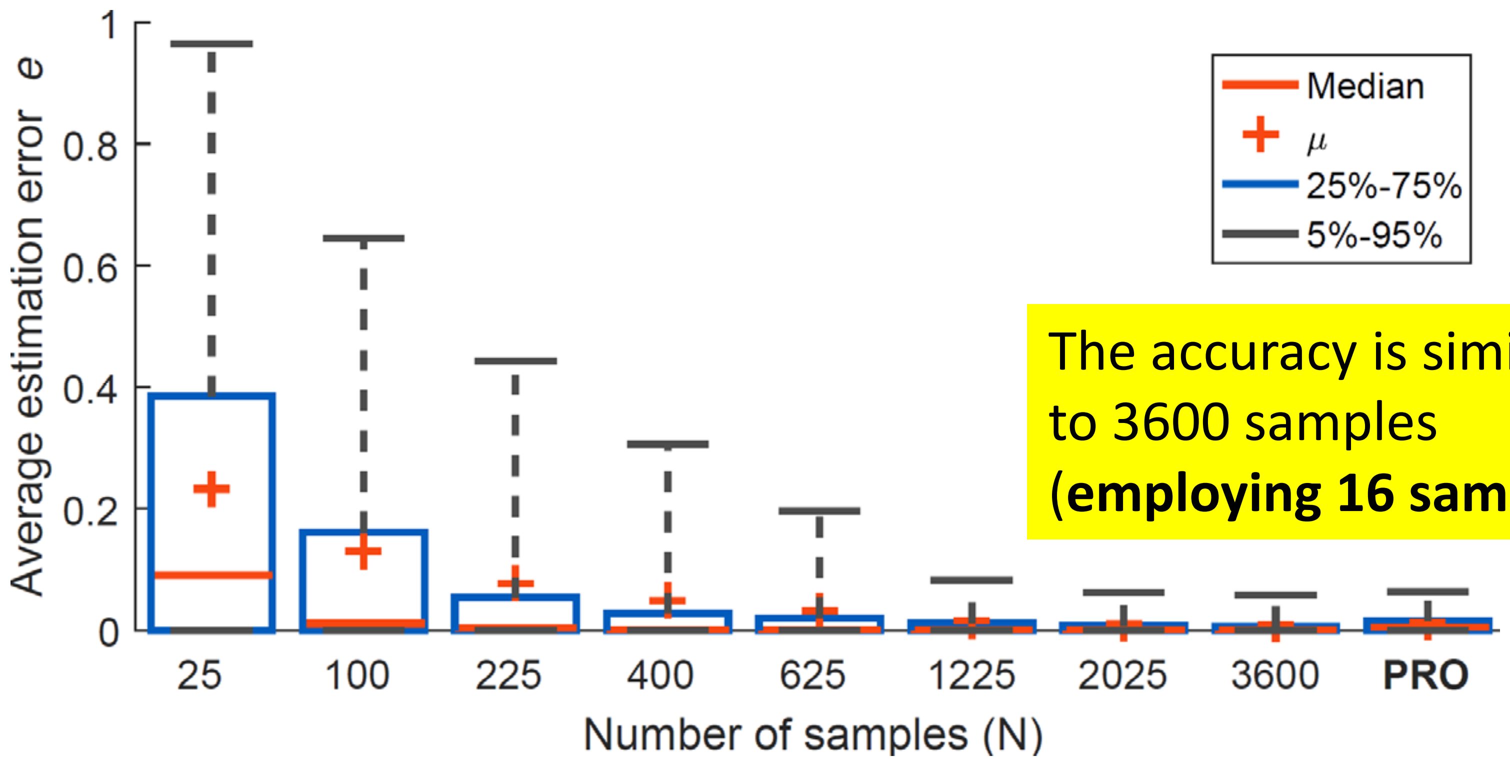



high clutter, 4 targets



# Detection quality: accuracy

- Compared vs. quadrature number integration with N samples
- Performance measure
  - average estimation error: reference value found by evaluating all locations in FOV $\Omega_i$ (high comp. cost)

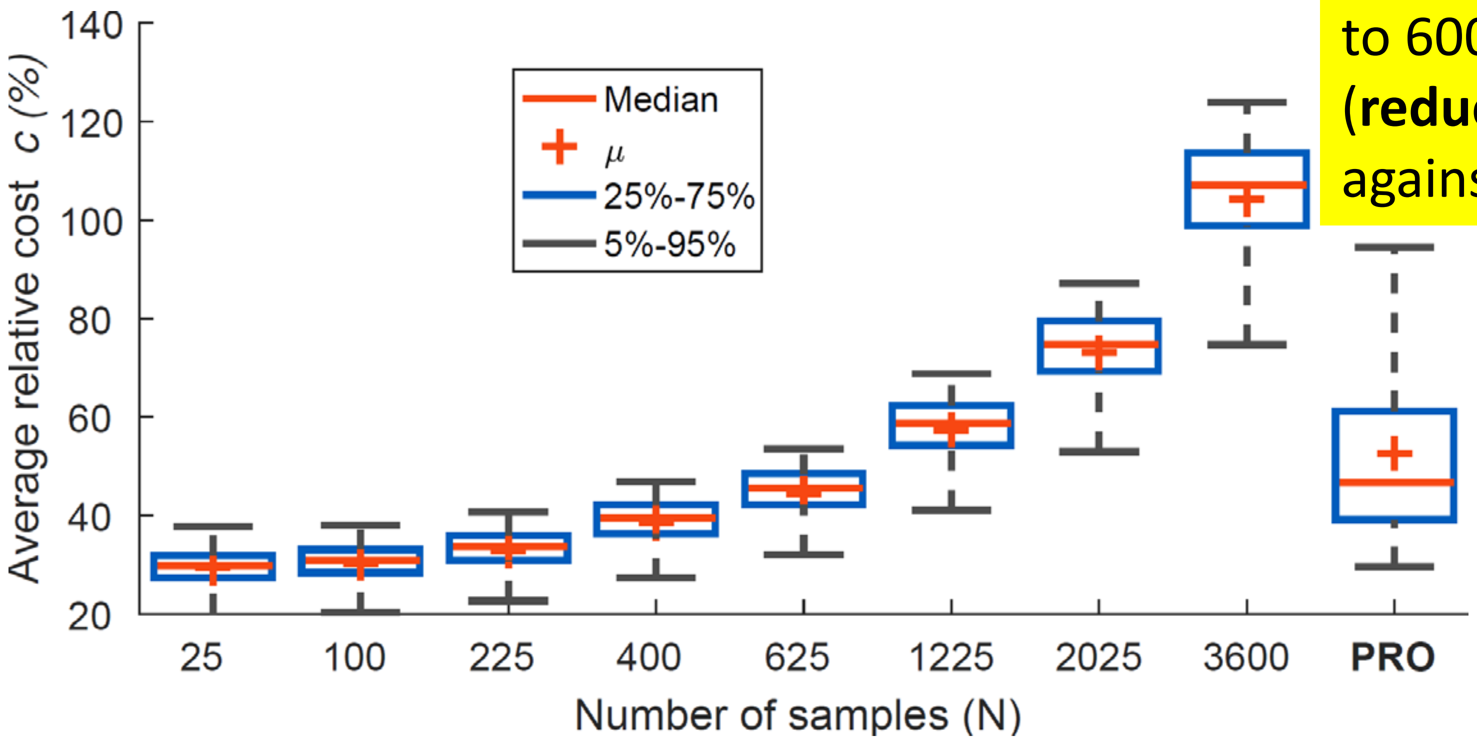



The accuracy is similar
to 3600 samples
(**employing 16 samples only!**)

# Detection quality: cost

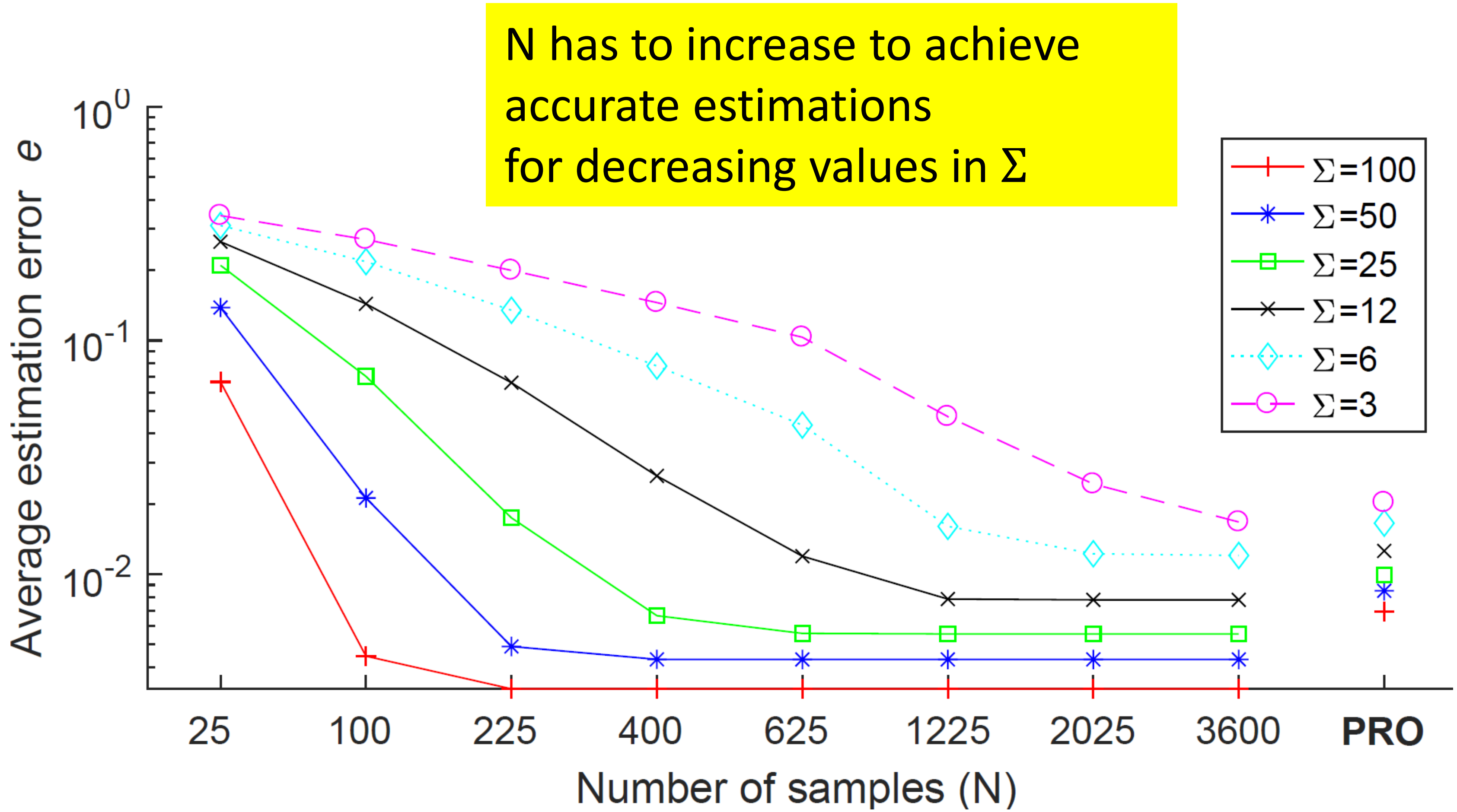
- Compared vs. quadrature number integration with N samples
- Performance measure
  - average relative cost
    - reference value: by evaluating all locations in FOV $\Omega_i$

The cost is similar
to 600-1225 samples
(**reduced ~2.5x**
against 3600 samples)



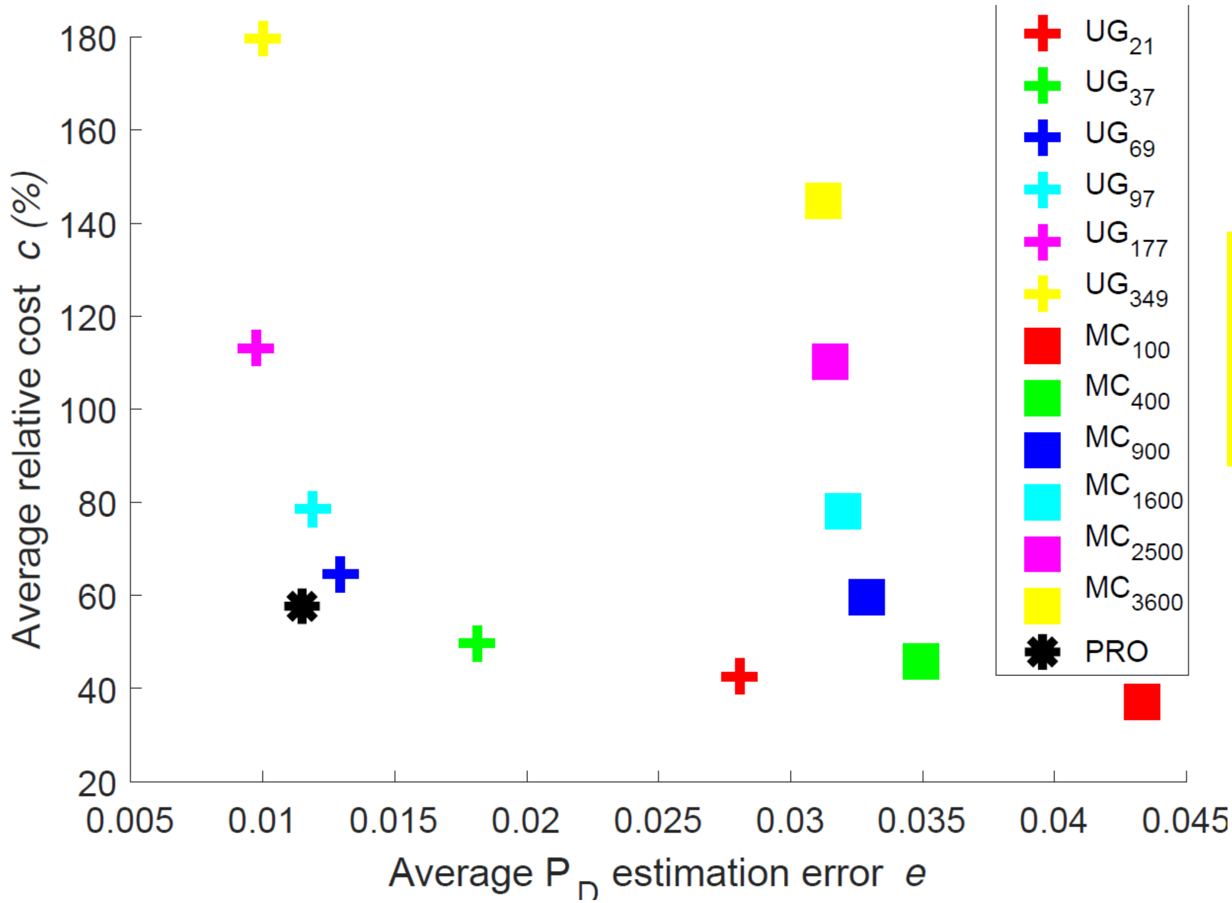
# Detection quality: uncertainty

- Uncertainty models
  - 6 uncertainty hypotheses



# Detection quality: comparison

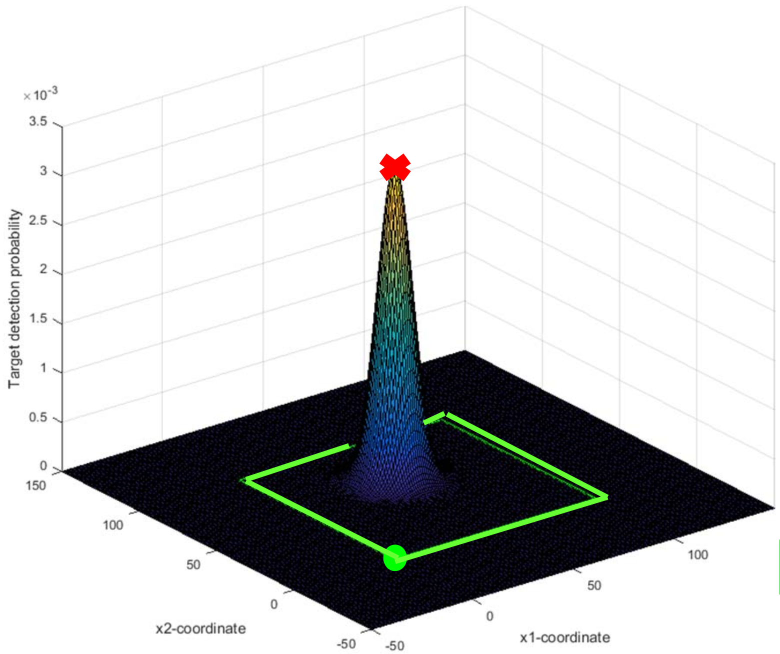
- Compared methods
  - UG: Uniform Grid sampling — low accuracy | fast computation
  - MC: Monte Carlo sample generation — high accuracy | costly




Our proposal improves both accuracy and cost of related methods

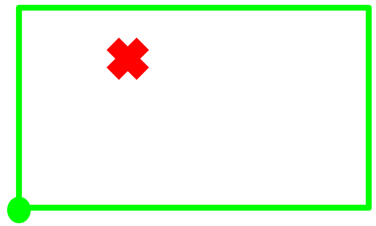
# Application to tracking

- Comparison vs. MTIC multi-camera multi-target tracker
  - Information Consensus for distributed target tracking
  - detection quality $P_D$ employed for JPDAF data association
  - cumulative distribution function of the target location




fast computation
models uncertainty
only computed for squared FOV
(unrealistic for multiple cameras)

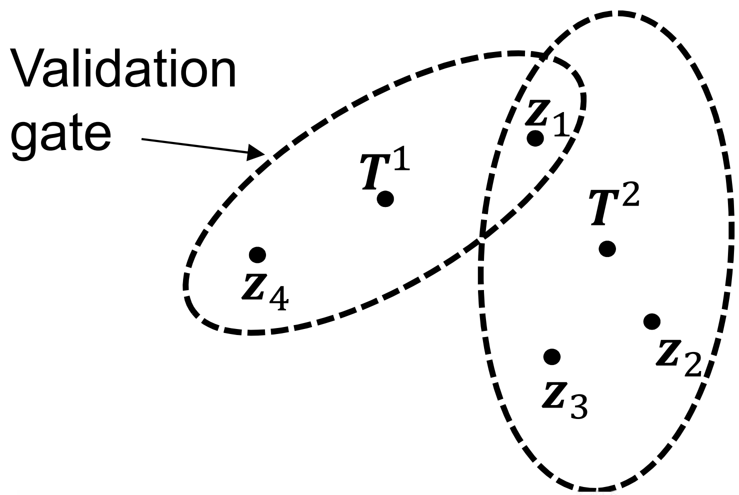
target location
camera FOV
camera center



**MTIC: Distributed multi-target tracking and data association in vision networks**
A. Kamal, J. Bappy, J. Farrell, and A. Roy-Chowdhury,
IEEE Trans. on Pattern Analysis and Machine Intelligence, 38(7):1397-1410, Jun 2016

# Evaluation: application to tracking (side note)

- Joint Probabilistic Data Association Filter for multiple targets
  - for each target $\boldsymbol{T}^j$, the mean measurement $\boldsymbol{y}^j$ is computed from all the measurements $\boldsymbol{z}_n$ and the association probabilities $\beta^{jn}$




$$\boldsymbol{y}^j = \sum_{n=1}^{l} \beta^{jn} \boldsymbol{z}_n$$

$$\beta^{jn} = \sum_{e=1}^{N_e} P(\chi_e|\boldsymbol{Z})\omega_{jn}$$

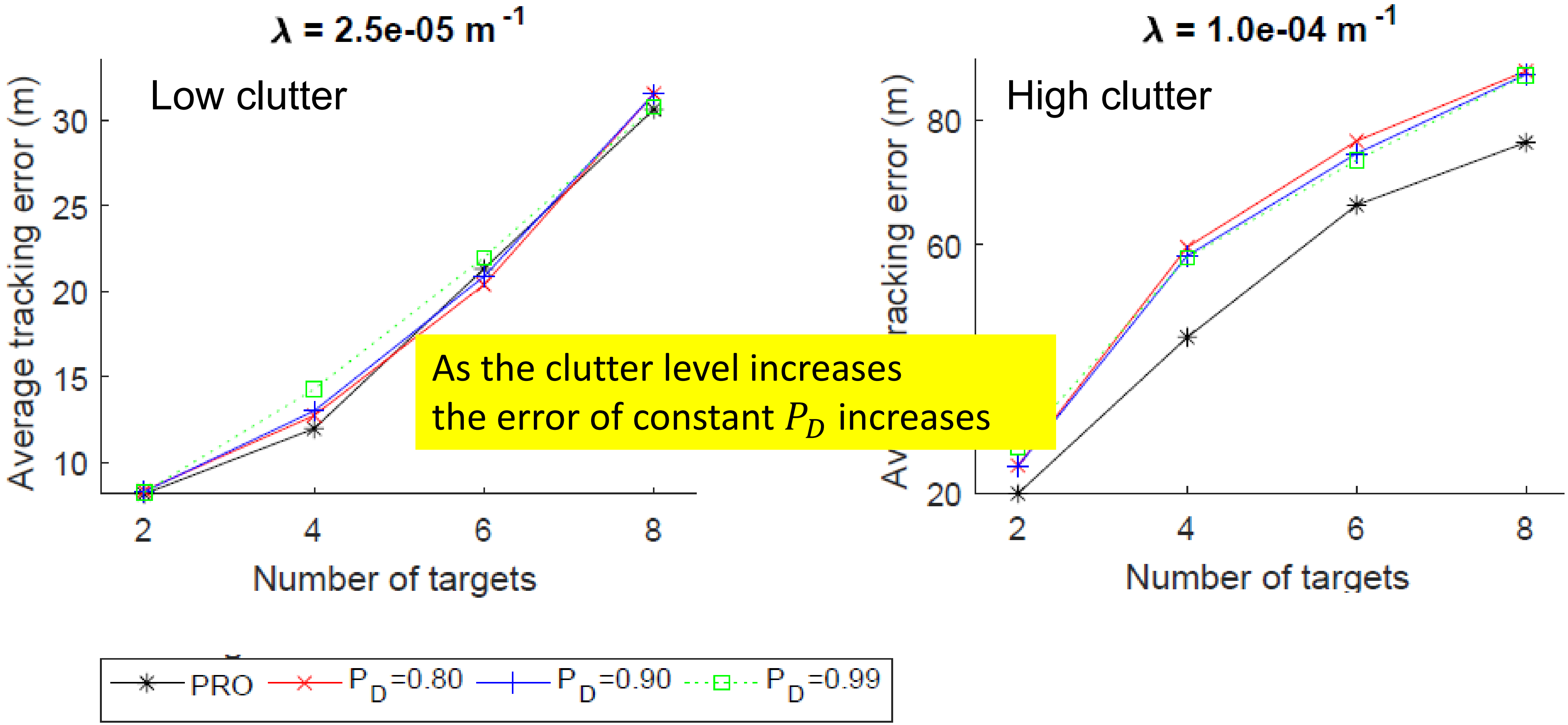
$\chi_e$: describes all possible associations between $z_n$ and $T^j$

$$P(\chi_e|Z) = \frac{C^{\varphi}}{c} \prod_{m:\tau_m=1} \frac{\exp[-\frac{1}{2}(\tilde{z}_{mi})' S_i^{-1} \tilde{z}_{mi}]}{(2\pi)^{M/2} |S_i|^{1/2}} \prod_{i:\delta_i=1} {P_D}^i \prod_{i:\delta_i=0} (1 - P_D^i), \quad e = 1, 2, \cdots, N_e$$

↓ probability of detecting target i

# Evaluation: application to tracking

- MTIC improvement: consider realistic FOVs (triangular)
- Four different levels of clutter (affect miss-detection rate)
- Comparison against typical constant $P_D$ values of literature




As the clutter level increases
the error of constant $P_D$ increases

# Conclusions

- Contribution
  - generic estimator for target detection quality
    with quadrature-based integration & importance sampling
  - closed-form function empirically derived to determine
    optimal number of integration points
  - results
    + accuracy
    - computational cost
  - application to multi-camera multi-target tracking
    - improves tracking performance & accuracy-complexity trade-off

- Future work
  - extension to color features

# WiSE-Mnet simulator for distributed computer vision

- Models: communication layers, sensing and distributed applications of camera networks (resource constraints)
- Networks capturing complex vectorial data (e.g. video)
- Includes implementations of state-of-the-art algorithms

www.eecs.qmul.ac.uk/~andrea/wise-mnet.html