Distributed Learning Gaussian Process Tracking with Uncertainty Quantification (Part of SIGNeTs Project)

Xingchi Liu, Chenyi Lyu, and Lyudmila Mihaylova
Department of Automatic Control and Systems Engineering, The University of Sheffield
Email: xingchi.liu@sheffield.ac.uk, clyu5@sheffield.ac.uk, l.s.mihaylova@sheffield.ac.uk

Outline

- Distributed learning Gaussian process tracking
- Case study: Tracking over wireless sensor networks
- Multi-class Gaussian process classification based on acoustic-seismic classification identification data set
- Summary

The University Of Sheffield.

Background Knowledge

- Model-free tracking approach: Learning the model dynamics directly from data assuming temporal correlation in the target trajectory

Input (time step) → Unknown function → Output (object state)

- At time step t, the Gaussian process (GP)-based motion tracker is written as:

$$x = f^x(t), \; f^x(t) \sim \mathrm{GP}^x(m^x(t), k^x(t, t')),$$
$$z^x = x + \epsilon^x, \; \epsilon^x \sim \mathcal{N}(0, \sigma_x^2),$$

Gaussian Process Regression

The University Of Sheffield.

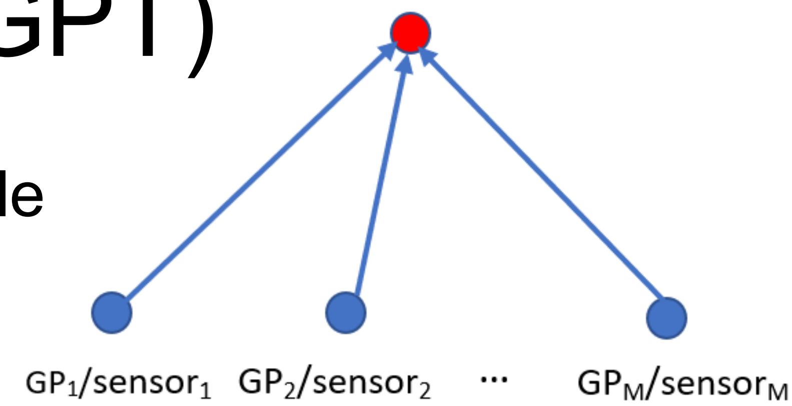
- The GP regression equations (predicted mean and variance) at an input location t_* can be written as

$$\mu_* = m(t_*) + K_*^\intercal \Sigma^{-1}(\mathbf{z} - m(t_*)),$$
$$\sigma_*^2 = k(t_*, t_*) - K_*^\intercal \Sigma^{-1} K_*,$$

- Scalability issue: High computational complexity $O(N^3)$ due to kernel matrix inversion
- Distributed GP : Divide-and-conquer and product of expert

Distributed Gaussian Process-based Tracking (DGPT)

The University Of Sheffield.



- Multiple local GPs instead of a single global GP

Prediction aggregation method:

– Product of experts (PoE)

$$p(f_*^x \mid t_*, \mathcal{D}) = \prod_{i=1}^{M} p_i(f_*^x \mid t_*, \mathcal{D}^{(i)})$$

$$\mu_*^{\text{PoE}} = (\sigma_*^{\text{PoE}})^2 \sum_{i=1}^{M} \sigma_i^{-2}(t_*)\mu_i(t_*),$$

$$(\sigma_*^{\text{PoE}})^{-2} = \sum_{i=1}^{M} \sigma_i^{-2}(t_*).$$

– Bayesian committee machine (BCM)

$$p(f_*^x \mid t_*, \mathcal{D}) = \frac{\prod_{i=1}^{M} p_i(f_*^x \mid t_*, \mathcal{D}^{(i)})}{p^{M-1}(f_*^x \mid t_*)}$$

$$\mu_*^{\text{BCM}} = (\sigma_*^{\text{BCM}})^2 \sum_{i=1}^{M} \sigma_i^{-2}(t_*)\mu_i(t_*),$$

$$(\sigma_*^{\text{BCM}})^{-2} = \sum_{i=1}^{M} \sigma_i^{-2}(t_*) + (1-M)\sigma_{**}^{-2}.$$

Hyperparameter Online Learning

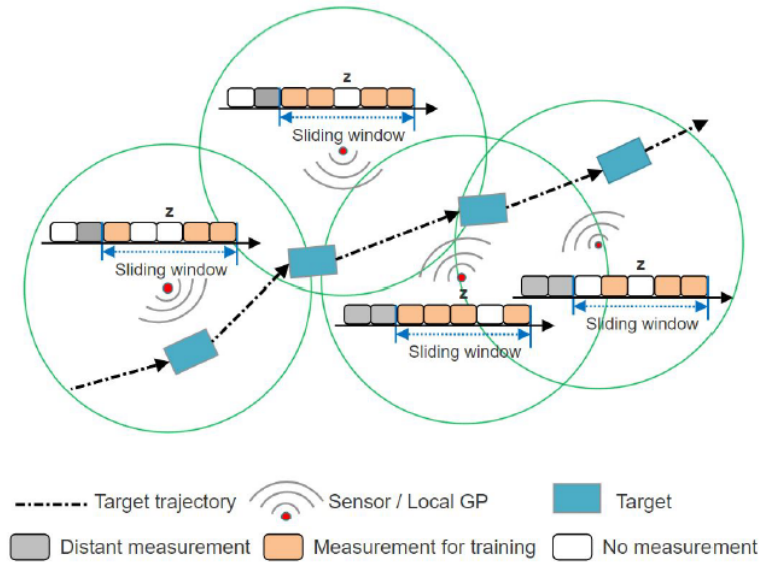


Fig. 1: A distributed point tracking system with 4 sensors. The length of sliding window in this example is 5 time steps

- GP hyperparameters online update

$$\log p(\mathbf{z}|X,\theta) = -\frac{1}{2}\mathbf{z}^\intercal \Sigma^{-1}\mathbf{z} - \frac{1}{2}\log|\Sigma| - \frac{n}{2}\log 2\pi,$$

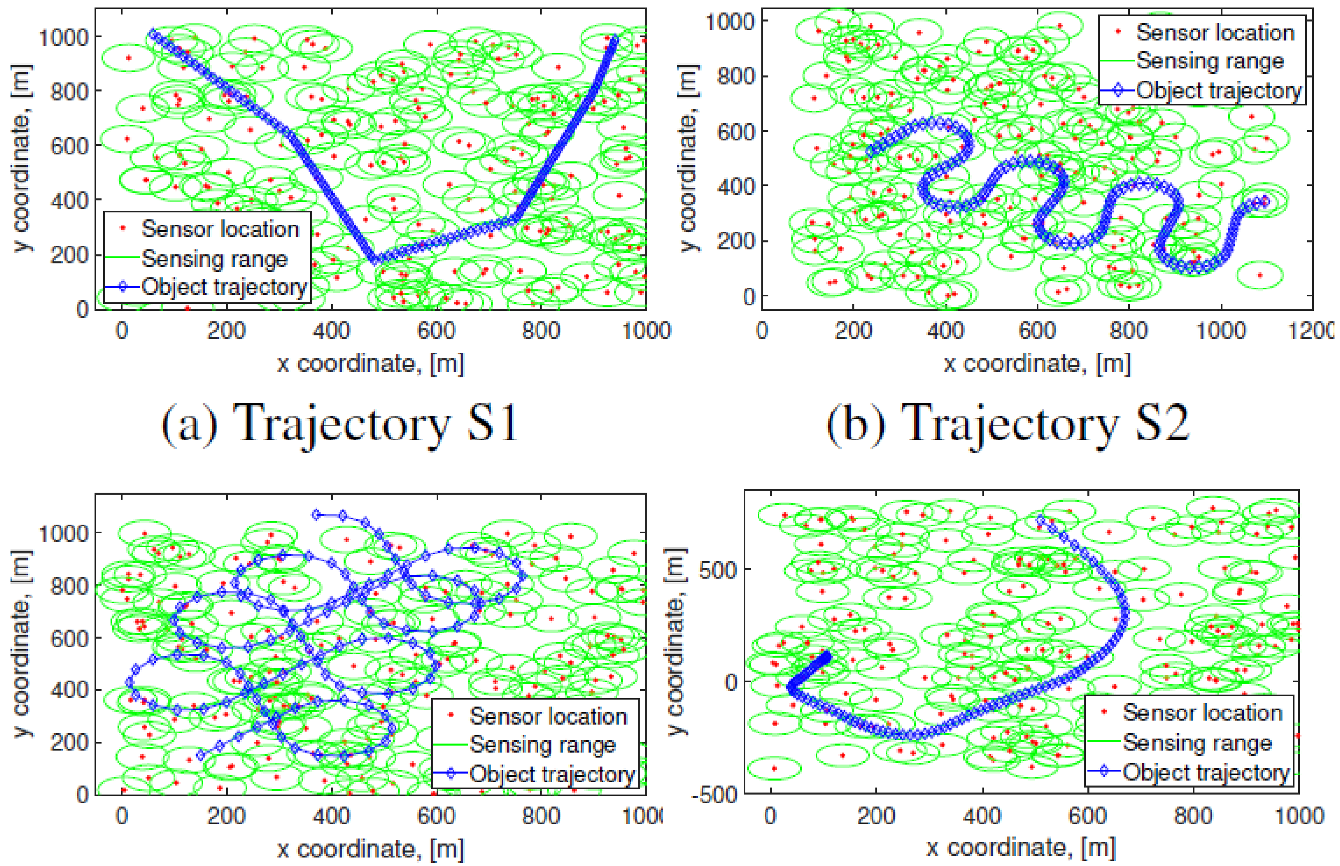
- Sliding window-based measurements
- Maximum likelihood estimation with consistent initial values of hyperparameters

Numerical Results

- Tracking performances under 4 trajectories and 3 noise levels
- Schemes:
  - DGPT-RBCM $(\mathcal{O}(Nn^2))$
  - DGPT-GPoE $(\mathcal{O}(Nn^2))$
  - Standard GP $(\mathcal{O}(N^3))$
  - Sparse GP $\left(\mathcal{O}(NM^2)\right)$
- Performance metric: Normalized root mean squared error (NRMSE), standard deviation of RMSEs over 100 MC runs

Numerical Results

- Testing scenarios



(a) Trajectory S1 (b) Trajectory S2

(c) Trajectory S3 (d) Trajectory S4

Fig. 2: Object trajectories

Numerical Results

- Superior performance is observed when the measurement noise level is relatively low (noise level 1 & 2), competitive performance is observed at high noise level

Scenario	Approach	Noise Level 1				Noise Level 2				Noise Level 3			
		Predicted NRMSE		Updated NRMSE		Predicted NRMSE		Updated NRMSE		Predicted NRMSE		Updated NRMSE	
		X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
S1	Standard GP	1.74%	1.71%	1.20%	1.14%	1.83%	2.19%	1.36%	1.55%	2.73%	3.50%	2.09%	2.65%
	Sparse GP	2.79%	1.95%	2.25%	1.18%	1.94%	2.00%	1.55%	1.31%	2.43%	2.54%	2.02%	1.91%
	DGPT-RBCM	1.20%	1.37%	0.78%	0.84%	1.79%	1.92%	1.27%	1.30%	3.06%	3.57%	2.45%	2.75%
	DGPT-GPoE	1.22%	1.76%	1.03%	1.04%	1.99%	1.99%	1.82%	1.72%	3.69%	3.52%	3.56%	3.68%
S2	Standard GP	6.36%	4.87%	4.61%	3.72%	5.67%	5.12%	4.01%	3.96%	6.78%	5.50%	5.04%	4.31%
	Sparse GP	4.82%	6.28%	2.70%	5.25%	5.05%	6.03%	3.88%	4.92%	5.50%	5.54%	4.25%	4.41%
	DGPT-RBCM	2.70%	3.58%	1.83%	2.50%	2.93%	7.29%	1.98%	5.82%	3.86%	10.01%	2.89%	8.19%
	DGPT-GPoE	2.40%	5.05%	1.86%	4.15%	2.47%	7.74%	2.23%	7.26%	3.60%	10.01%	3.55%	9.63%
S3	Standard GP	4.75%	5.89%	2.71%	3.81%	5.53%	6.25%	3.19%	4.02%	7.05%	7.24%	4.23%	4.63%
	Sparse GP	5.75%	5.85%	3.48%	4.25%	5.87%	5.87%	3.64%	4.34%	6.31%	6.13%	4.16%	4.67%
	DGPT-RBCM	6.62%	6.44%	3.24%	4.18%	6.69%	6.78%	3.48%	4.29%	7.45%	7.48%	4.07%	4.74%
	DGPT-GPoE	5.41%	5.99%	3.18%	4.47%	5.56%	6.06%	3.41%	4.57%	6.05%	6.40%	3.99%	4.98%
S4	Standard GP	2.21%	1.30%	1.59%	0.90%	3.30%	1.52%	2.58%	1.09%	4.54%	2.17%	3.61%	1.68%
	Sparse GP	3.89%	2.21%	3.59%	2.00%	3.49%	2.27%	3.26%	2.13%	3.52%	2.25%	3.33%	2.21%
	DGPT-RBCM	2.71%	1.82%	1.85%	1.07%	3.73%	2.42%	2.72%	1.49%	5.64%	4.10%	4.58%	2.98%
	DGPT-GPoE	3.33%	1.55%	3.06%	1.31%	4.88%	2.67%	4.61%	2.37%	8.16%	5.03%	7.86%	4.72%

Numerical Results

- Impact of uncertainties on tracking

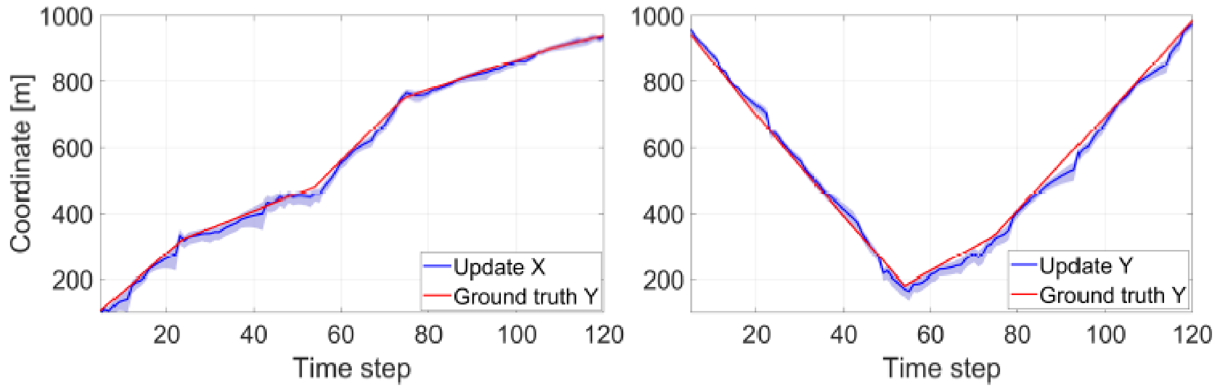


Fig. 3: Tracking uncertainty in both coordinates. The results are collected in S1 under noise level three.

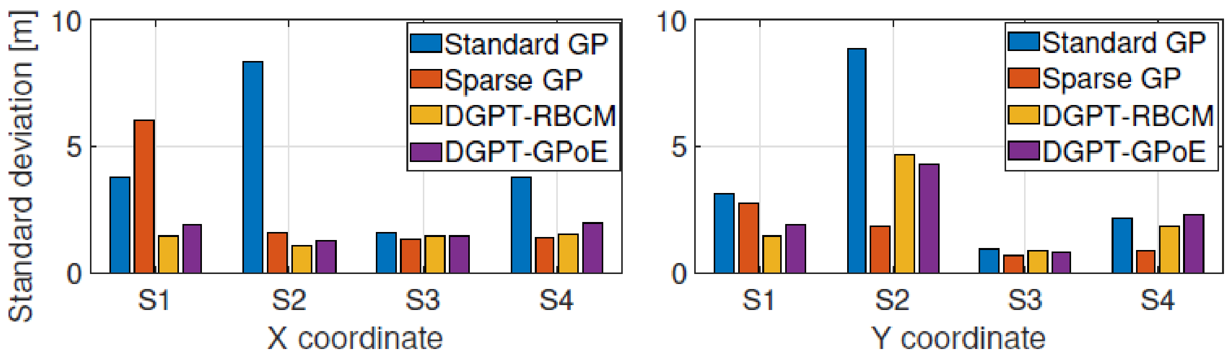
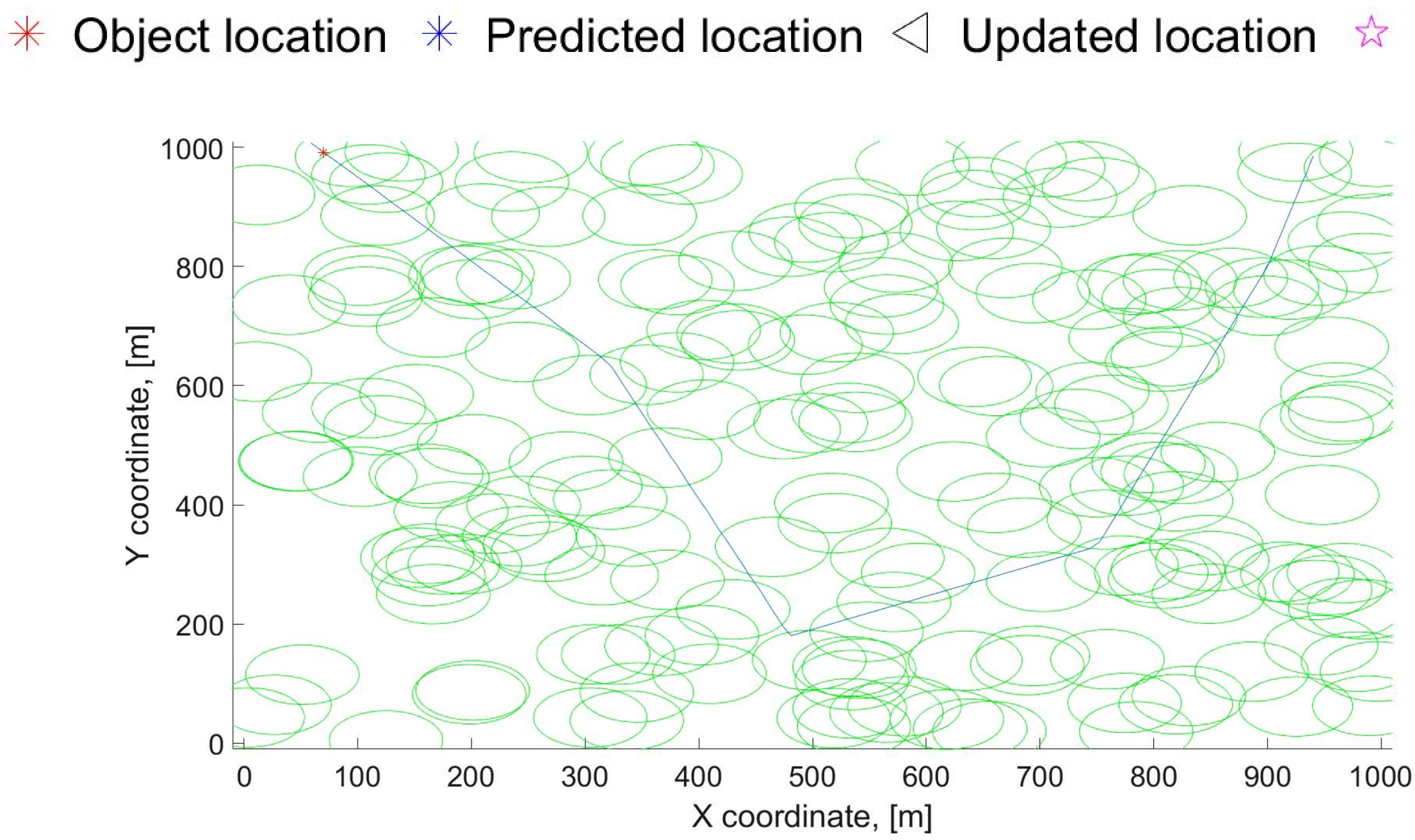


Fig. 4: Standard deviation of RMSE in both coordinates. The results are collected under noise level one.

Distributed Point Tracking in Clutter



Numerical Results

Table I: Updated NRMSEs for S1

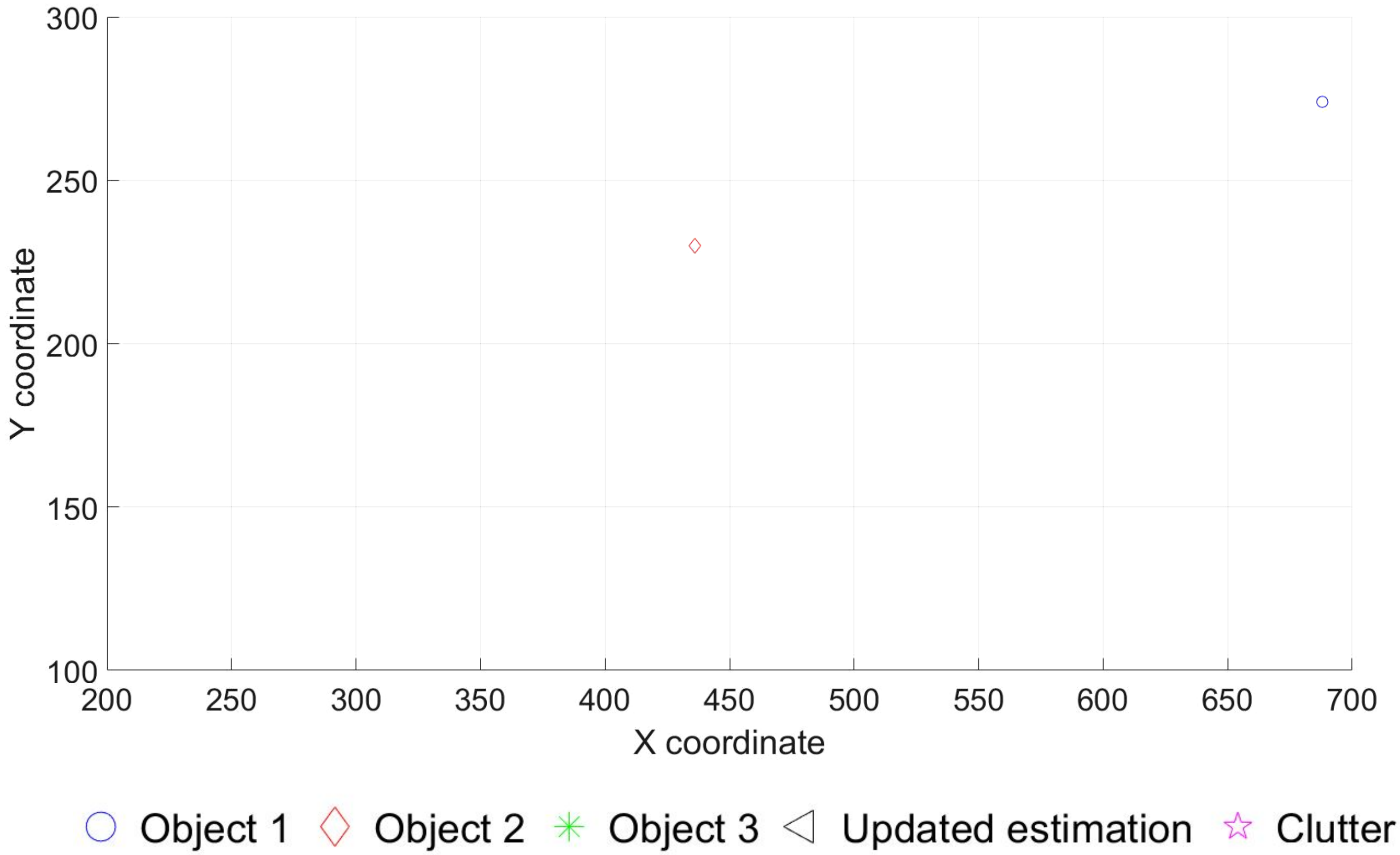
Noise level	Clutter rate	Approach					
		Standard GP		DGPT-RBCM		DGPT-GPoE	
		X	Y	X	Y	X	Y
1	0.0001	1.82%	2.46%	0.78%	0.98%	0.95%	1.13%
	0.0003	2.93%	3.16%	1.46%	1.48%	1.72%	1.64%
	0.0006	4.71%	5.03%	4.76%	3.82%	5.08%	4.05%
2	0.0001	1.79%	2.64%	0.94%	1.05%	1.17%	1.21%
	0.0003	2.76%	3.63%	1.69%	1.63%	2.00%	1.80%
	0.0006	4.74%	5.31%	5.96%	4.27%	6.29%	4.58%
3	0.0001	1.78%	2.63%	1.14%	1.17%	1.46 %	1.39%
	0.0003	2.74%	3.47%	1.90%	1.83%	2.29%	2.04%
	0.0006	4.80%	5.30%	6.19%	4.68%	6.53%	5.05%

Table II: Updated NRMSEs for S2

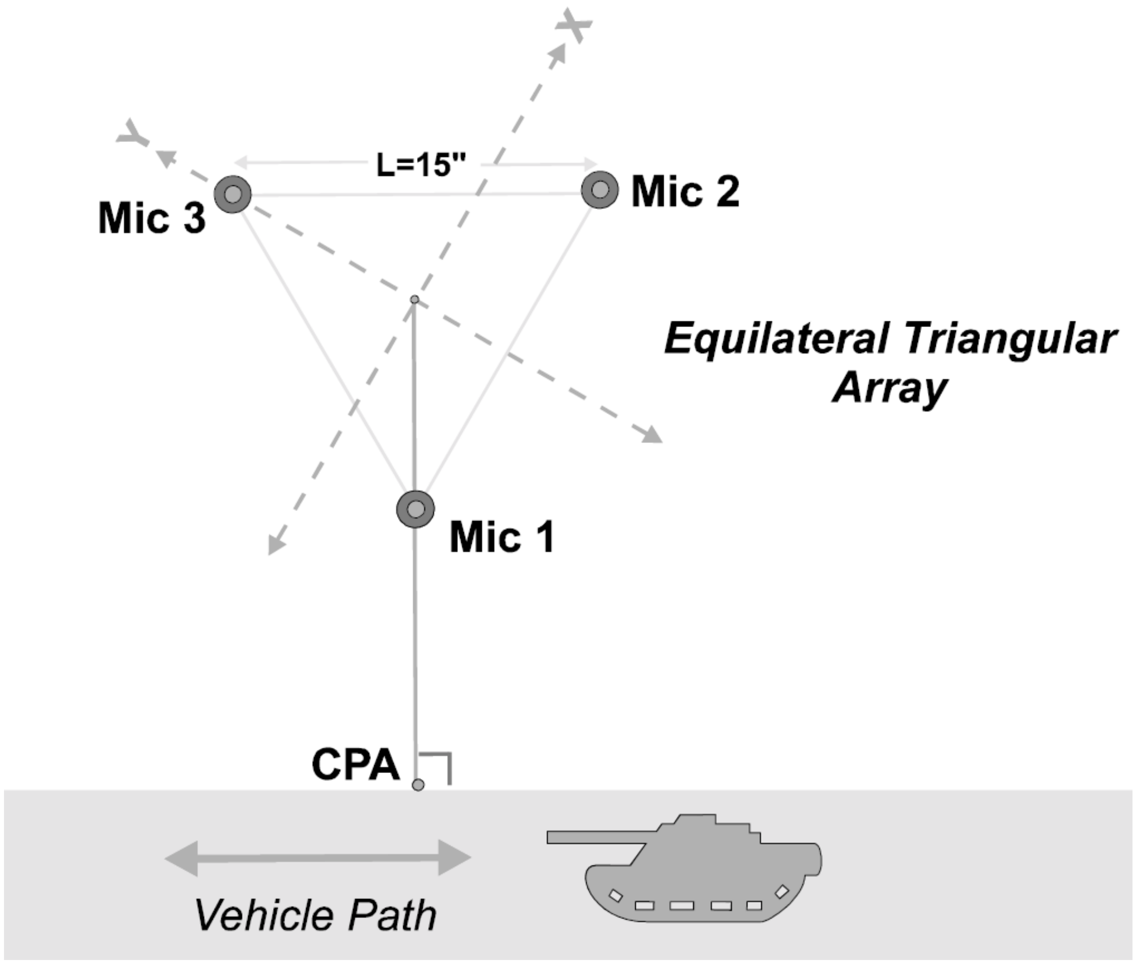
Noise level	Clutter rate	Approach					
		Standard GP		DGPT-RBCM		DGPT-GPoE	
		X	Y	X	Y	X	Y
1	0.0001	5.44%	3.19%	2.46%	2.39%	2.61%	3.55%
	0.0003	6.76%	3.77%	3.28%	3.89%	3.55%	5.13%
	0.0006	8.58%	4.42%	7.00%	10.79%	7.30%	11.31%
2	0.0001	5.44%	3.18%	2.42%	2.70%	2.56%	4.07%
	0.0003	6.83%	3.80%	3.21%	4.20%	3.49%	5.54%
	0.0006	8.58%	4.47%	7.05%	10.59%	7.34%	11.13%
3	0.0001	5.58%	3.24%	2.42%	3.39%	2.61%	4.97%
	0.0003	7.02%	3.90%	3.26%	5.07%	3.55%	6.51%
	0.0006	8.76%	4.49%	7.12%	10.92%	7.42%	11.47%

- Superior performance is observed when the clutter rate is relatively low (0.0001 & 0.0003), competitive performance is observed at high clutter rate

Multi-Target Tracking in Clutter



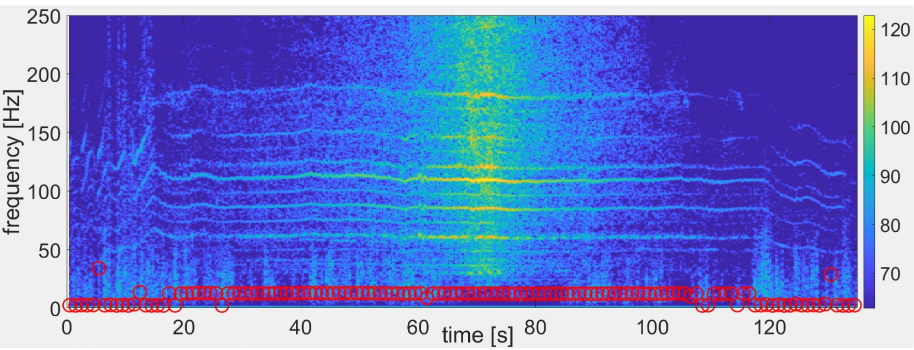
Acoustic and Seismic Classification Identification Data Set



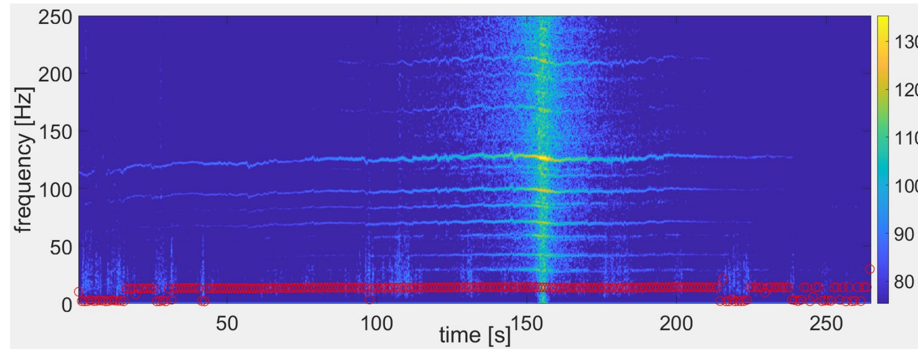
- Traveling at constant speeds (5km/h-40 km/h) with two directions
- Closest point of approach (CPA) varies from 25m-100m
- 9 different vehicle types
- 4 different test sites
- 2 different sensor systems

Fundamental Frequency Estimation[1],[2]

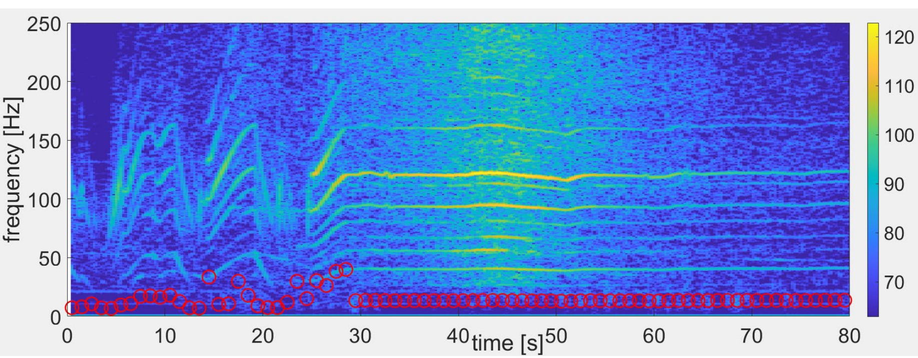
Gv1c1085



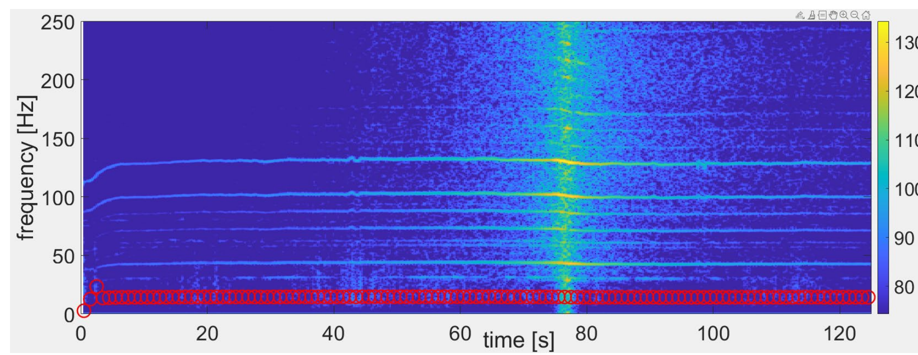
Gv1c1019



Gv2c1081



Gv2c1007

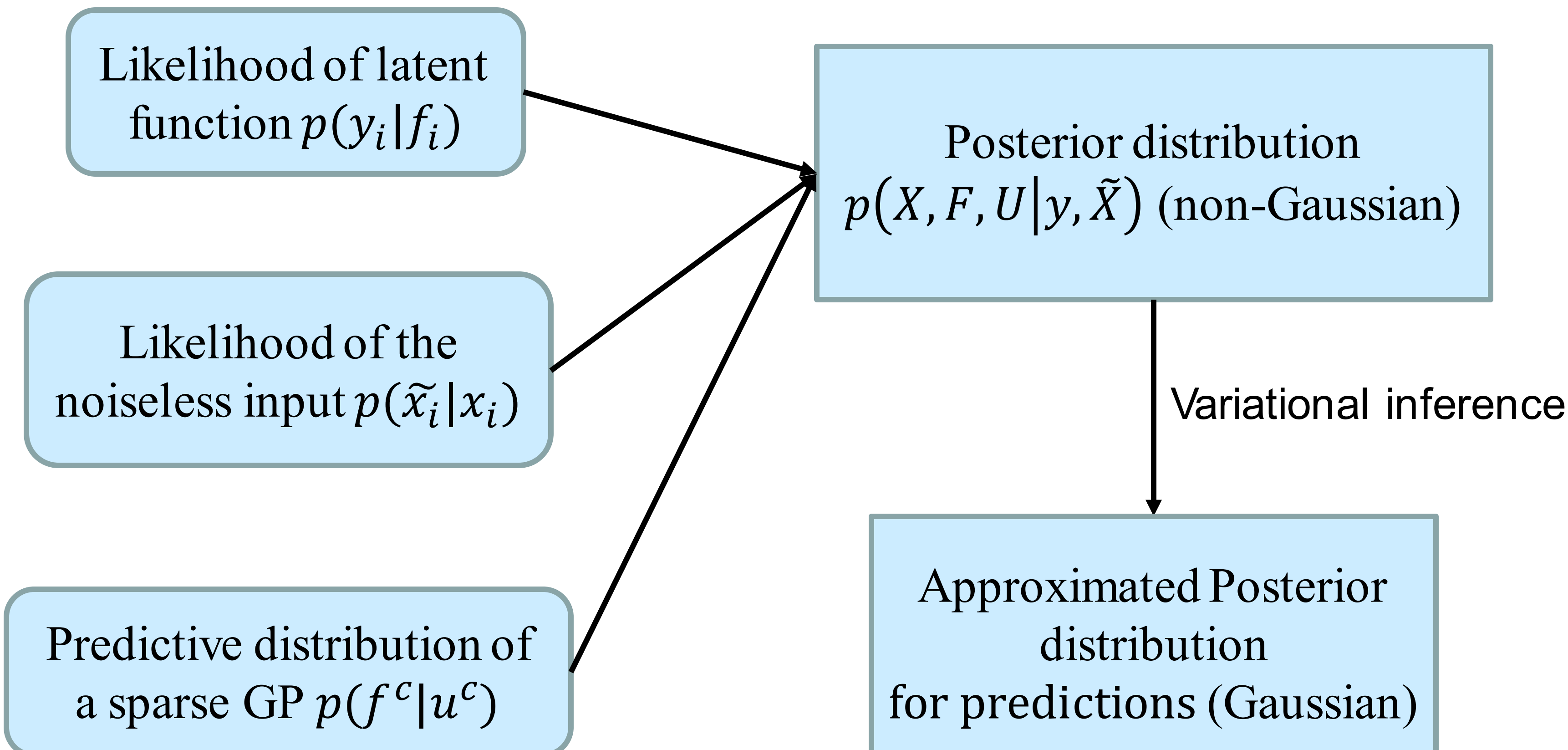


[1]. Nielsen, J. K., Jensen, T. L., Jensen, J. R., Christensen, M. G., & Jensen, S. H. (2017). Fast Fundamental Frequency Estimation: Making a Statistically Efficient Estimator Computationally Efficient. *Signal Processing*, 135, 2017, pp. 188-197.
[2]. J. K. Nielsen, M. G. Christensen, A. T. Cemgil, S. J. Godsill, and S. J. Jensen, "Bayesian Interpolation and Parameter Estimation in a Dynamic Sinusoidal Model," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 7, pp. 1986–1998, 2011.

Multi-Class GP Classification with Noisy Inputs

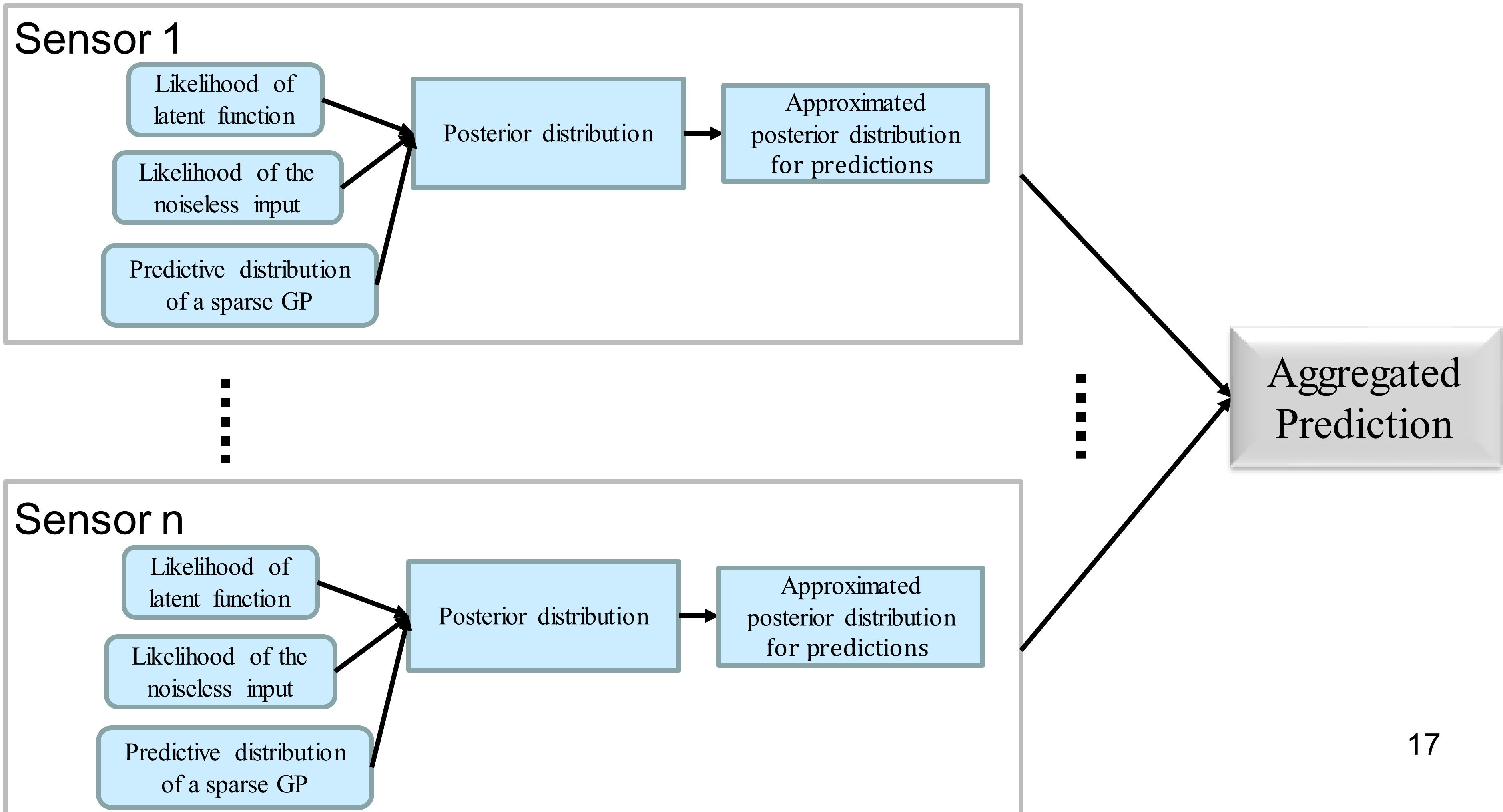
- Labeling rule (for data (label) i, C classes in total)

$$y_i = \arg\max_c \; f^c(\mathbf{x}_i)$$



Distributed Classification based on DGP

- Distributed classification



Summary

The University Of Sheffield.

- Model free tracking approach for distributed system
  - Computational complexity
  - Communication cost
- Hyperparameter online learning
- Probabilistic distributed machine learning approach for ground vehicle classification

The University Of Sheffield.

Thank you !