Decentralized Stochastic Planning for Nonparametric Bayesian Models

Silvia Ferrari
Professor of Engineering and Computer Science
Department of Mechanical Engineering and Materials Science
Duke University

2011 ONR MURI: Nonparametric Bayesian Models to Represent Knowledge and Uncertainty for Decentralized Planning

Review Meeting, MIT, Boston (MA)
January 24, 2014

Research Problem and Motivation

Problem: BNP-based sensor planning and control for modeling agent behaviors in the presence of significant uncertainties.

- Learn BNP models of dynamic agents
- BNP model size and parameters are both learned from data

Motivation: BNP (e.g. DP-GP and DDP-GP) models can be used to learn trajectory and velocity fields from data:

- Dirichlet process mixture (DP-GP) infers number of trajectory field classes [Roy, MIT]
- Dependent Dirichlet process mixture (DDP-GP) extends to temporally evolving trajectory fields [Fisher, MIT]

Technical Challenges: BNP-based sensor planning and control requires an information value function that can be updated in real time, as the BNP model acquires new data.

- Information theoretic function for BNP models
- Computationally tractable update of expected information value (for real time implementations)
- Dynamic and geometric constraints on sensor state and control
- BNP-based planning and control
- Decentralized BNP-based planning and control (for multiple, cooperative sensors)

Technical Accomplishments – Year 2

- Developed **information value functions** for DP-GP models of dynamic targets
  (1) Represent expected uncertainty reduction in target position and velocity field
  (2) Update iteratively over time, as the DP-GP model learns target behavior from data obtained in real time **[Carin, Roy, How]**

- Developed **planning/control** algorithms for DP-GP models of dynamic targets
  (1) Demonstrated on camera intruder problem with continuous state and control (static sensors) **[How, Carin]**
  (2) Extension to multiple mobile sensors tracking multiple moving targets

- Developed **decentralized planning/control** algorithms for DP-GP models of dynamic targets for multiple sensors with continuous state and control, tracking multiple moving targets **[Leonard]**

- Demonstrated new **hybrid system approximate dynamic programming** (ADP) algorithm on a benchmark optimal control problem.
  - Important for performing distributed learning control through ADP, for teams of heterogeneous autonomous static and mobile agents involving both discrete and continuous state and control variables

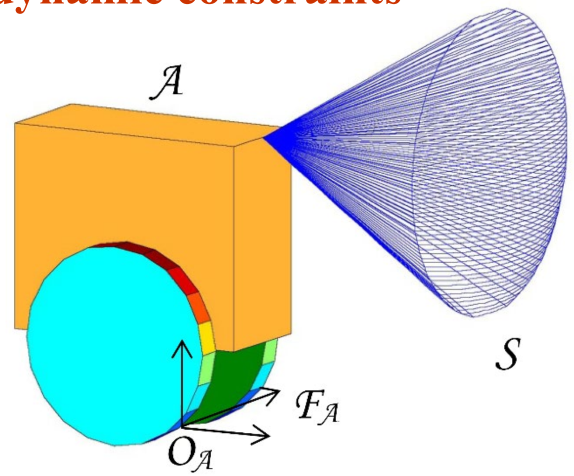
Problem Formulation and Assumptions

Problem: Sensor planning/control for modeling dynamic **target behaviors** via DP-GP

Consider one or more sensors with configuration $\mathbf{q}(t)$, and control inputs $\mathbf{u}(t)$, such that:

$\dot{\mathbf{q}}(t) = \mathbf{A}\mathbf{q}(t) + \mathbf{B}\mathbf{u}(t)$, with I.C. $\mathbf{q}(t_0) = \mathbf{q}_0$ **"sensor dynamic constraints"**

- Using the available **control inputs** $\mathbf{u}(t)$, the position and size (mode) of the sensor's FOV or visibility region S can be controlled to obtain measurements in the sensor workspace $\mathcal{W}$
- The **sensor objective** is to learn a model of target behavior from data, in the form of a DP-GP (or other BNP).



Targets are non-cooperative, independent and obey a time-invariant velocity field:

$$\dot{\mathbf{x}}_j(t) = \mathbf{f}_j[\mathbf{x}_j(t)] \equiv \mathbf{v}_j(t), \quad j = 1,\ldots,N(t) \quad \rightarrow \{\mathcal{F}, \boldsymbol{\pi}\}$$ **"target model"**

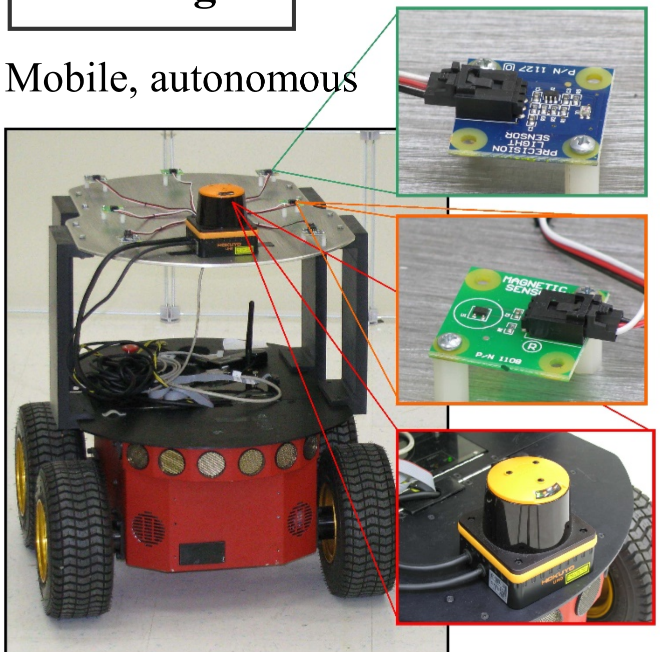
- Target j follows $\mathbf{v}_j$ with $\Pr\{C_j = i\} = \pi_i, \forall i,j; \quad \sum_{i=1}^{M} \pi_i = 1$
- When target j first enters W, its initial position is known without error; Perfect measurement-target association; velocity field is of class C^1.

Application Example

Multi-agent Monitoring and Surveillance

Sensor agent

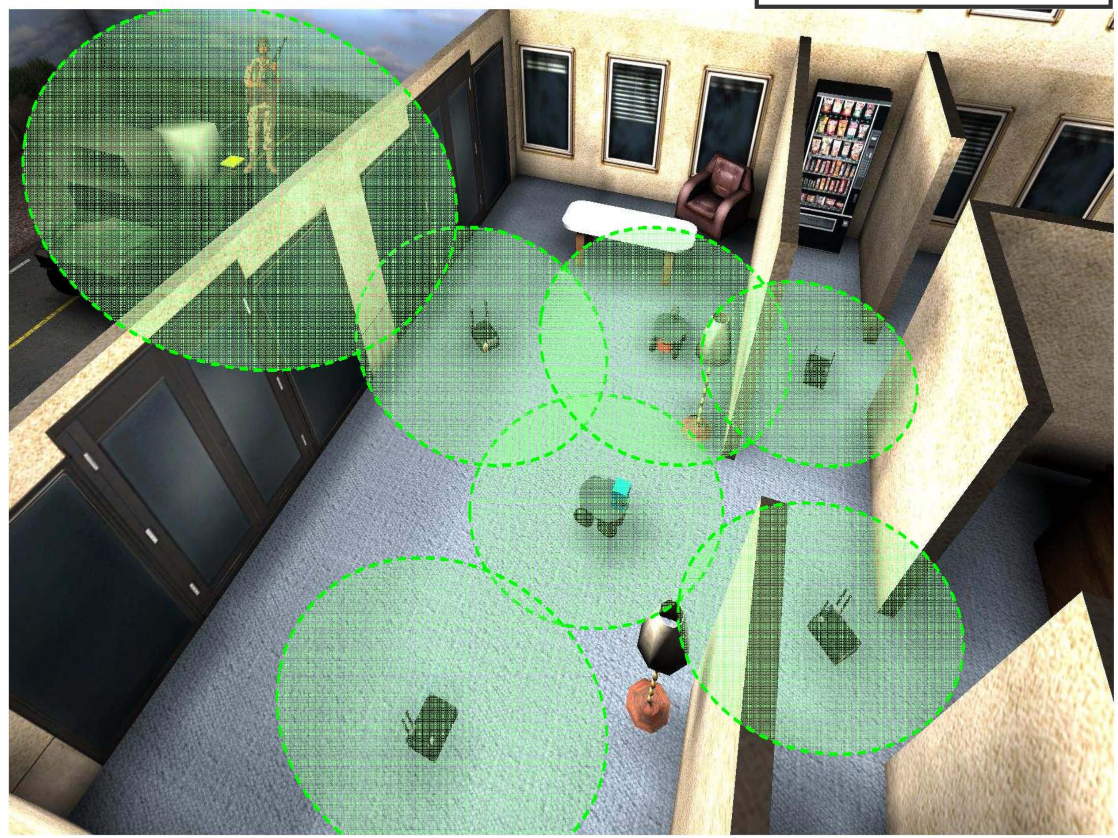
Mobile, autonomous



Wireless communication

Sensors

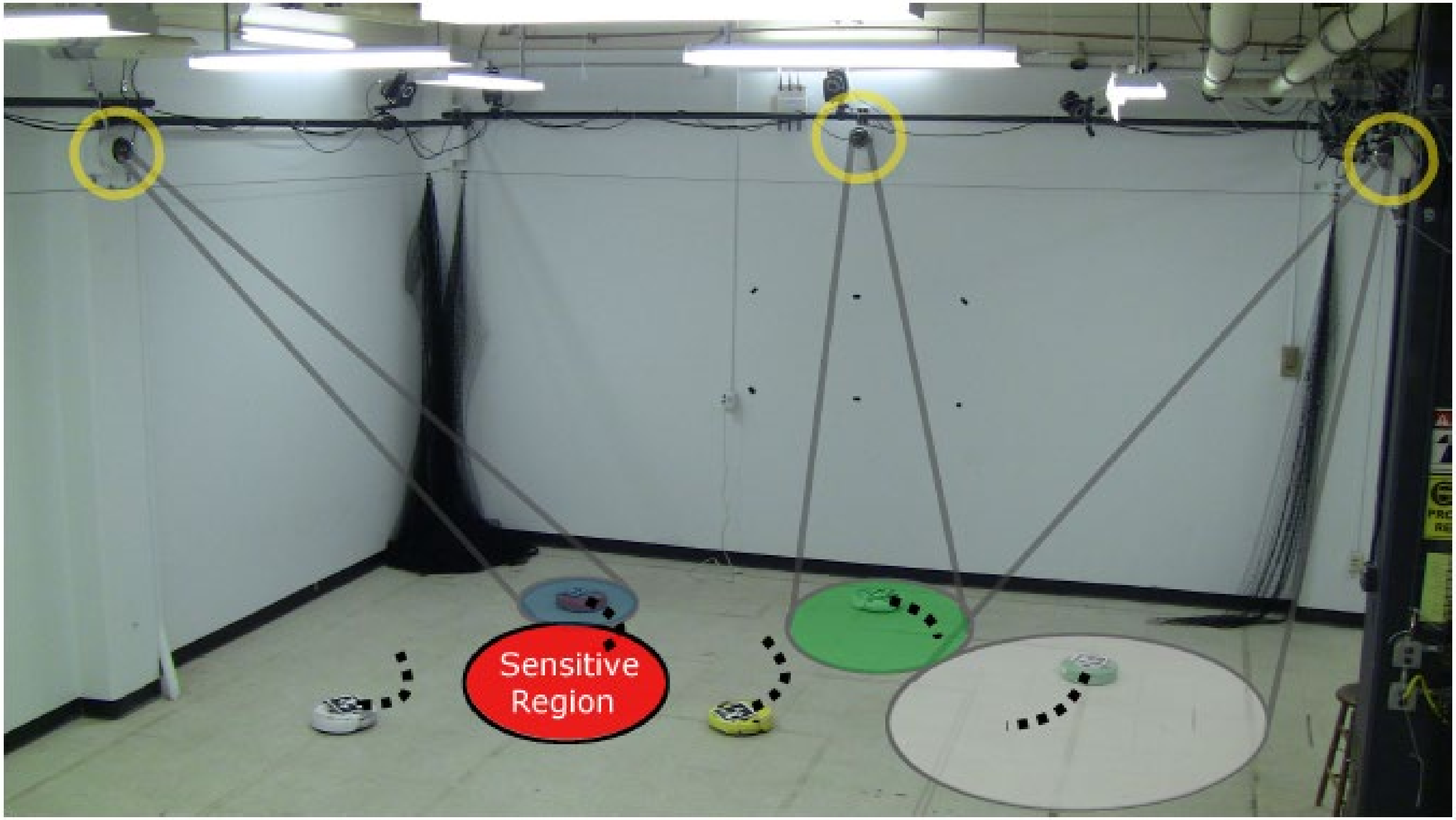
Target agents



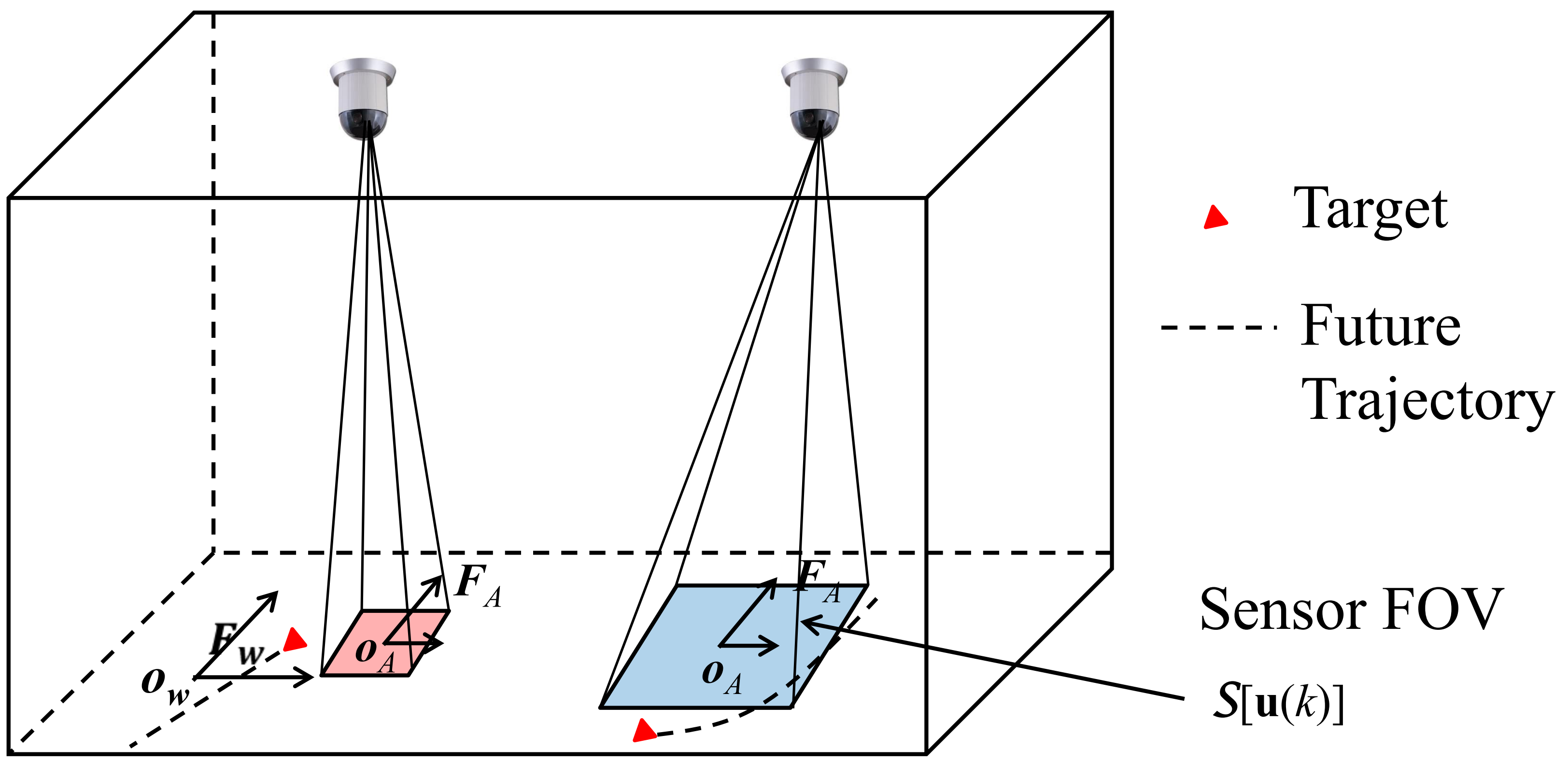
- Microscopic level: local agent model, communication, and inference algorithms; agent actions, constraints.
- Macroscopic level: mission objectives (performance) and constraints; field-level inference and situational awareness; CBBA task allocation.

Application Benchmark

Raven Testbed at MIT



Problem: learn target dynamics and track targets by controlling the cameras' pan-tilt-zoom (PTZ) variables (control inputs).

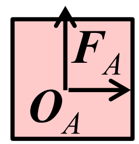


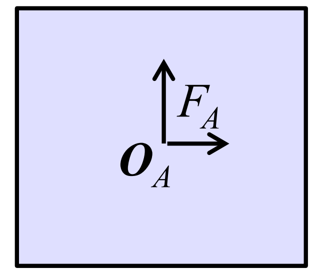
➢ **Camera control: u**
- **Continuous control: Position of o_A**
- **Discrete control: Zoom level l**

Pan-Tilt-Zoom (PTZ) Camera Model and Assumptions

- **Zoom level:** $l = 1(\text{in})$ | $l = 2(\text{out})$

Camera FoV





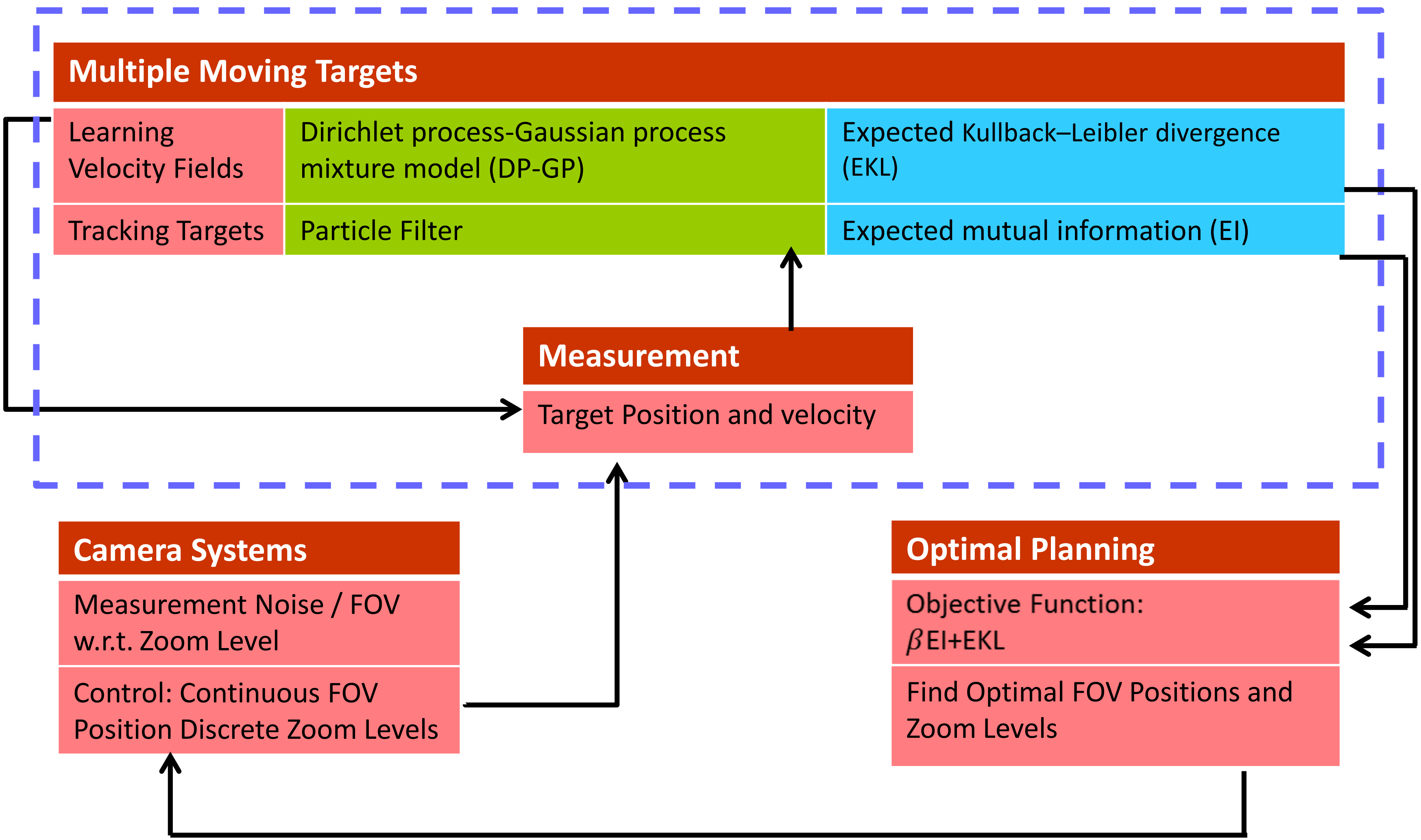
- **Measurement Model:**

$$\mathbf{m}(t) \equiv \begin{bmatrix} \mathbf{y}_j(t) \\ \mathbf{z}_j(t) \end{bmatrix} \equiv \begin{bmatrix} \mathbf{x}_j(t) \\ \mathbf{v}_j(t) \end{bmatrix} + \begin{bmatrix} \mathbf{n}_x \\ \mathbf{n}_v \end{bmatrix}, \qquad \begin{matrix} \mathbf{n}_x \sim \mathcal{N}(\mathbf{0},\ \sigma_x^2(l_i)\mathbf{I}_2) \\ \mathbf{n}_v \sim \mathcal{N}(\mathbf{0},\ \sigma_v^2(l_i)\mathbf{I}_2) \end{matrix}$$

- **Assumptions**
  - Camera is fully controlled by the command (control vector), i.e., $\mathbf{u}(\mathrm{k}) = [\mathbf{q}^{\mathrm{T}}(k)\ \ l(k)]^{\mathrm{T}}$, where $\mathbf{q}(k)$ is the position of $\boldsymbol{O}$ in $\mathcal{W}$ at time k and $l(k)$ is the zoom level.
  - $\mathbf{u}(k) \in \mathcal{U}(k)$, where $\mathcal{U}(k)$ is the admissible control space at time k.

Algorithmic Architecture

➢ **Active Sensing: Camera-Intruder Problem**



BNP-based Objective Function

KL-MI Information value function for active sensing control:

- Objective function

$$J[\mathbf{u}(k)] = \hat{\varphi}(\mathcal{F}; \mathbf{m}(k+1) | \mathcal{M}(k), \mathcal{E}(k), \mathbf{u}(k))$$

Expected KL-divergence: reward to update DP-GP model

$$+\beta\hat{\psi}\left(\mathbf{x}_j(k+1); \mathbf{m}(k+1) \middle| \mathcal{M}(k), \mathcal{E}(k), \mathbf{u}(k)\right)$$

Expected mutual information: reward to track targets

where

$$\mathcal{M}(k) = \{\mathbf{m}(\ell) | \mathbf{m}(\ell) \neq \emptyset, k' < \ell \leq k\},$$

measurement history not used in updating DP-GP model

$$\mathcal{E}(k) = \{\mathbf{m}(\ell) | \mathbf{m}(\ell) \neq \emptyset, 0 \leq \ell \leq k'\},$$

measurement history used in updating DP-GP model

parameters of Gaussian process

β: relative weight

BNP-based Camera Control Design

- **Centralized Control problem**

$$\mathbf{u}(k) = \underset{\mathbf{u}(k)}{\operatorname{argmax}} \sum_j \hat{\varphi}_j \, [\mathcal{F};\, \mathbf{m}_j(k+1) \mid \mathcal{M}_j(k), \varepsilon(k), \mathbf{u}(k)]$$

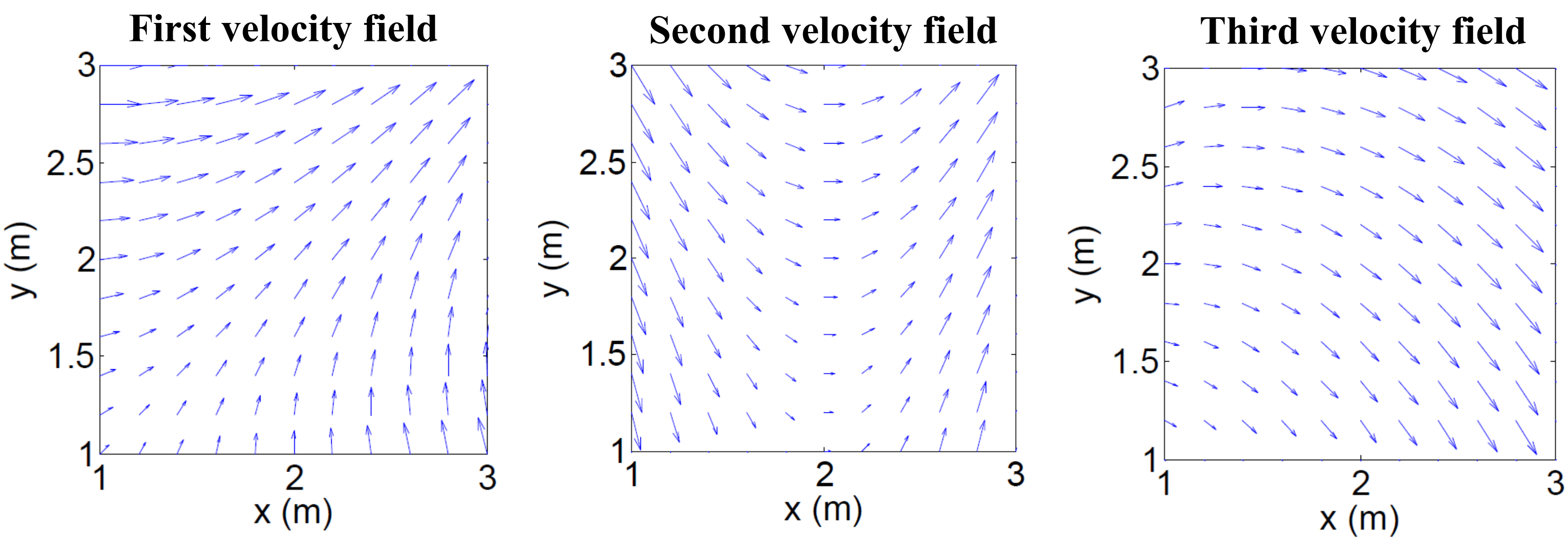
- 1. $\mathbf{y}_j(k+1) \approx \mathbf{x}_j(k+1)$ 2. Target position estimation: $p[\mathbf{x}_j(k+1)]$
- Approximation: → Weighted points covering problem
  - Expected benefit for $\mathcal{F}$ from a velocity measurement $\mathbf{v}_j(k+1)$ at $\mathbf{x}_j(k+1)$

    $$\hat{V}_j[\mathcal{F}; \mathbf{v}_j(k+1) \mid \mathcal{M}_j(k), \mathcal{E}(k), \mathbf{x}_j(k+1), \mathbf{u}(k)]$$

  - $\hat{\varphi}_j = \int \hat{V}_j \, [\mathbf{x}_j(k+1)] p[\mathbf{x}_j(k+1)] d\mathbf{x}_j(k+1)$
  - Sample $\{\boldsymbol{\chi}_j^1, \dots, \boldsymbol{\chi}_j^P\} \sim p[\hat{\mathbf{x}}_j(k+1)]$
  - $\hat{\varphi}_j \, [\mathbf{u}(k)] \approx 1/P \sum_{\chi^p_j \in S[\mathbf{u}(k)]} \hat{V}_j[\chi^p_j(k+1)]$
- Approximation of ψ : same samples $\{\boldsymbol{\chi}_j^1, \dots, \boldsymbol{\chi}_j^P\}$
- Computational complexity $PN(k) \log[PN(k)]$ (One camera case)

Camera-Intruder Example Problem

- **Workspace:** $\mathcal{W} = \{\mathbf{x} \in \mathbb{R}^2 \mid 1 \leq x \leq 3, 1 \leq y \leq 3\}$
- **Velocity fields:** $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$

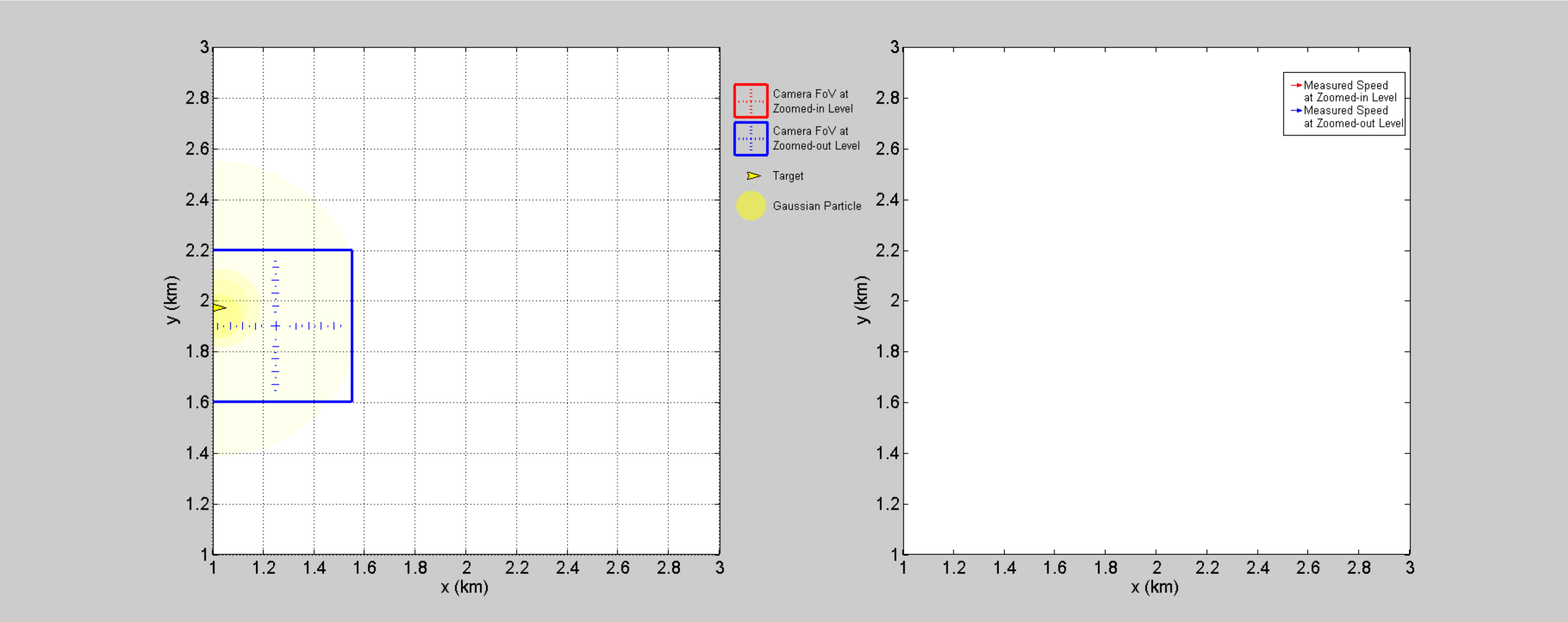


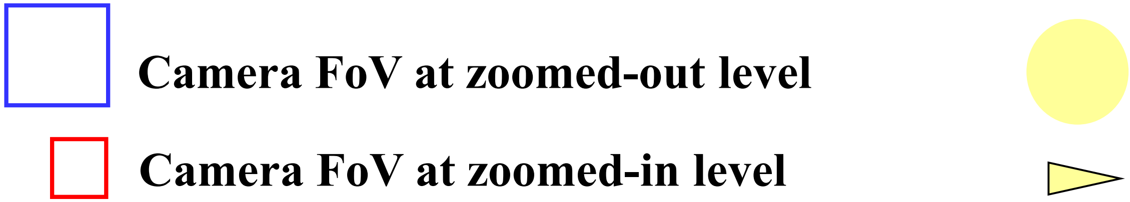
$$\mathbf{f}_1(\mathbf{x}) = \begin{bmatrix} -2 & 4 \\ 3 & -1 \end{bmatrix}\mathbf{x} \qquad \mathbf{f}_2(\mathbf{x}) = \begin{bmatrix} 2 & -2 \\ -6 & -3 \end{bmatrix}\mathbf{x} \qquad \mathbf{f}_3(\mathbf{x}) = \begin{bmatrix} 0 & 1 \\ 5 & 0 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 0 \\ -10 \end{bmatrix}$$

- **Probability of choosing every velocity field:** $\boldsymbol{\pi} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$

Simulation Results

- **Active Camera Control** for mobile intruder BNP modeling and tracking





Camera FoV at zoomed-out level

Camera FoV at zoomed-in level

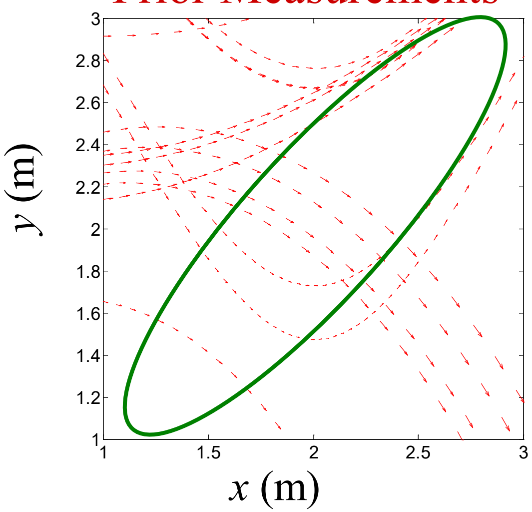
Gaussian particle

Target

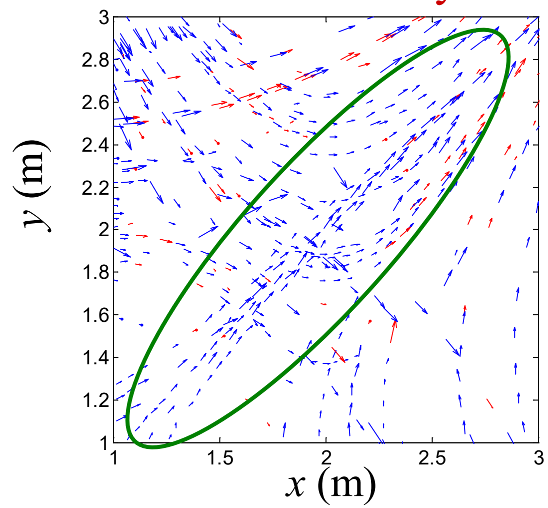
Results: Measurement Comparison

Goal: Control camera PTZ to maximize the expected reduction in uncertainty of future sensor measurements (information value) / optimize the DP-GP target model.

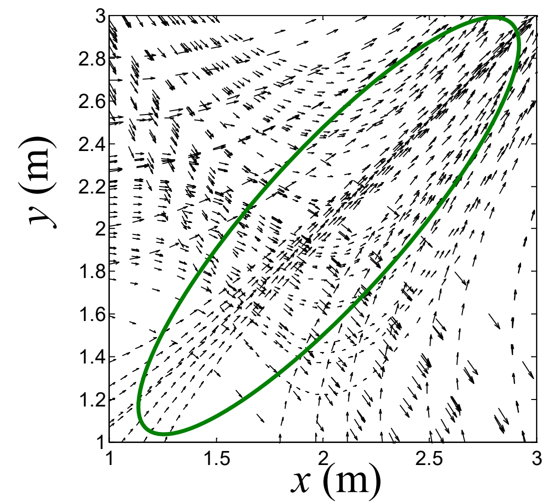
Prior Measurements



KL-MI Policy

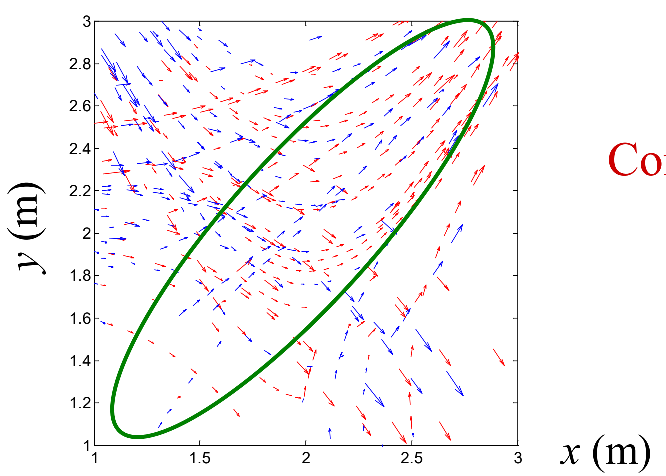


Actual Trajectories



Camera measurements at zoomed-out level

Camera measurements at zoomed-in level

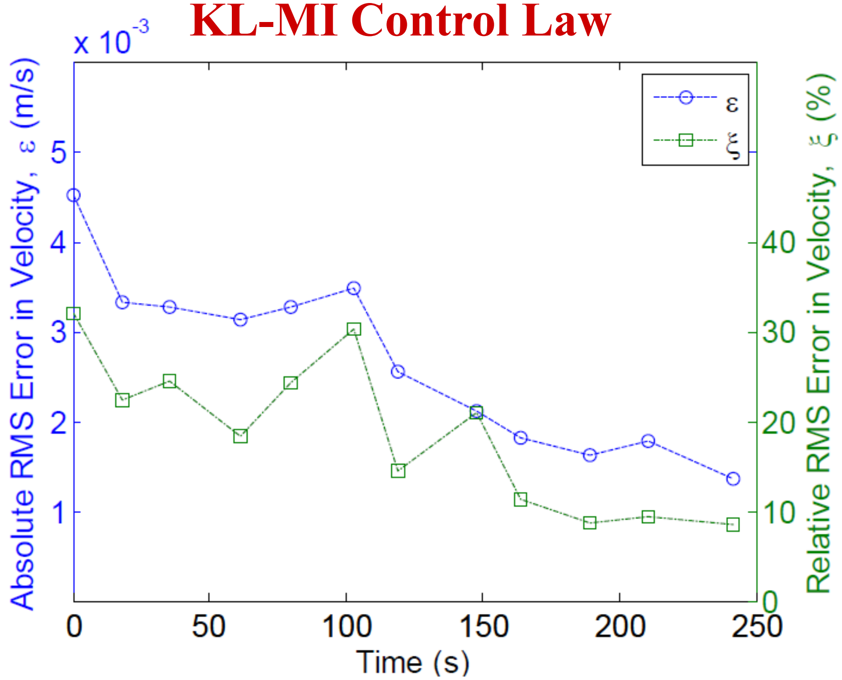


Comparison with MI Policy

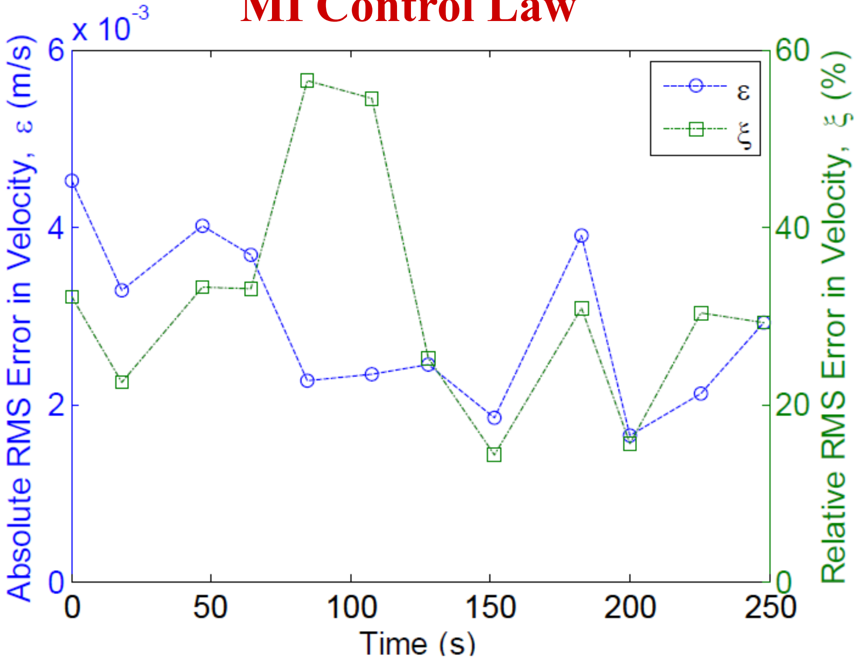
Results: Performance Comparison

Performance Evaluation: error time histories of DP-GP target velocity estimate

KL-MI Control Law



MI Control Law



Absolute error:

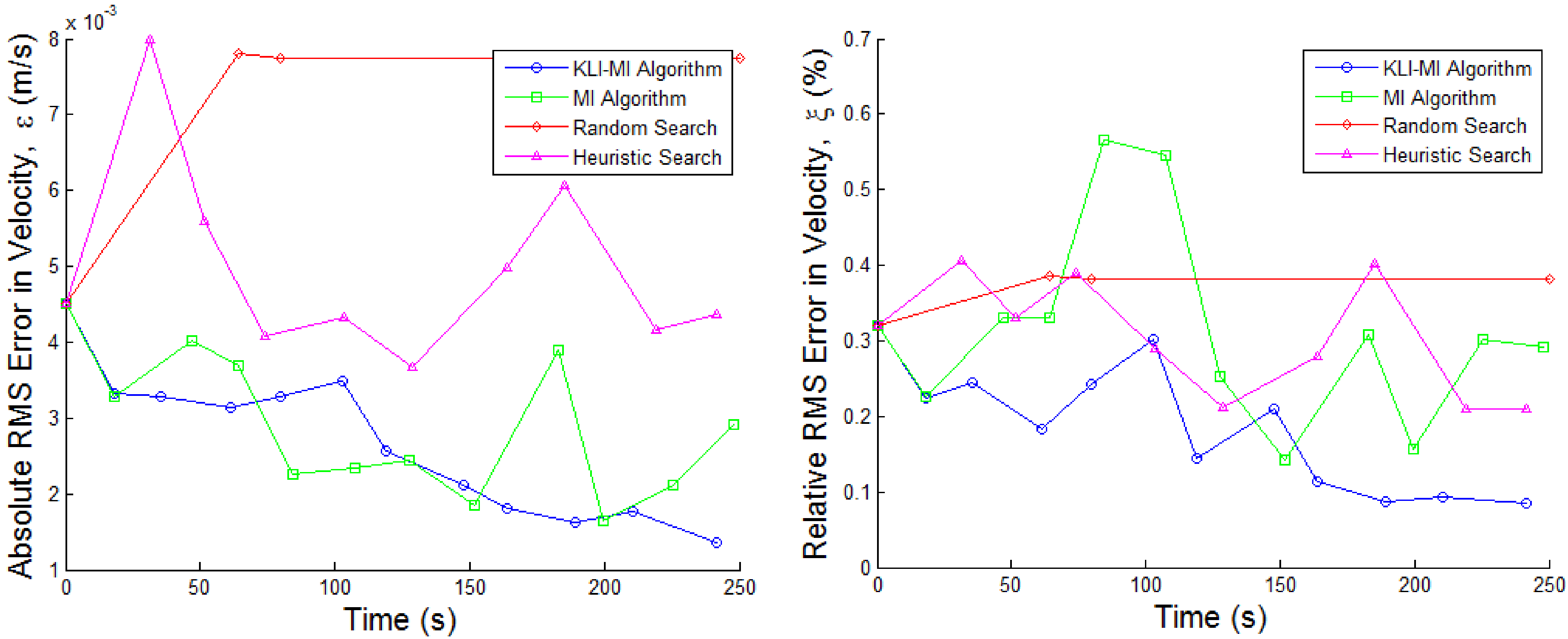
$$\varepsilon = \frac{1}{N_A}\sum_{i=1}^{N_A}\sum_{j=1}^{N_G} w_{ij}\sqrt{\frac{1}{N_{T_i}}\sum_{k=1}^{N_{T_i}} \| \mathbf{v}_i(k) - \boldsymbol{\mu}_j(k) \|_2^2}$$

Relative error:

$$\xi = \frac{1}{N_A}\sum_{i=1}^{N_A}\sum_{j=1}^{N_G} w_{ij}\sqrt{\frac{1}{N_{T_i}}\sum_{k=1}^{N_{T_i}} \| 1 - \frac{\boldsymbol{\mu}_j(k)}{\mathbf{v}_i(k)} \|_2^2}$$

Results: Performance Comparison

Performance Evaluation: error time histories of DP-GP target velocity estimate

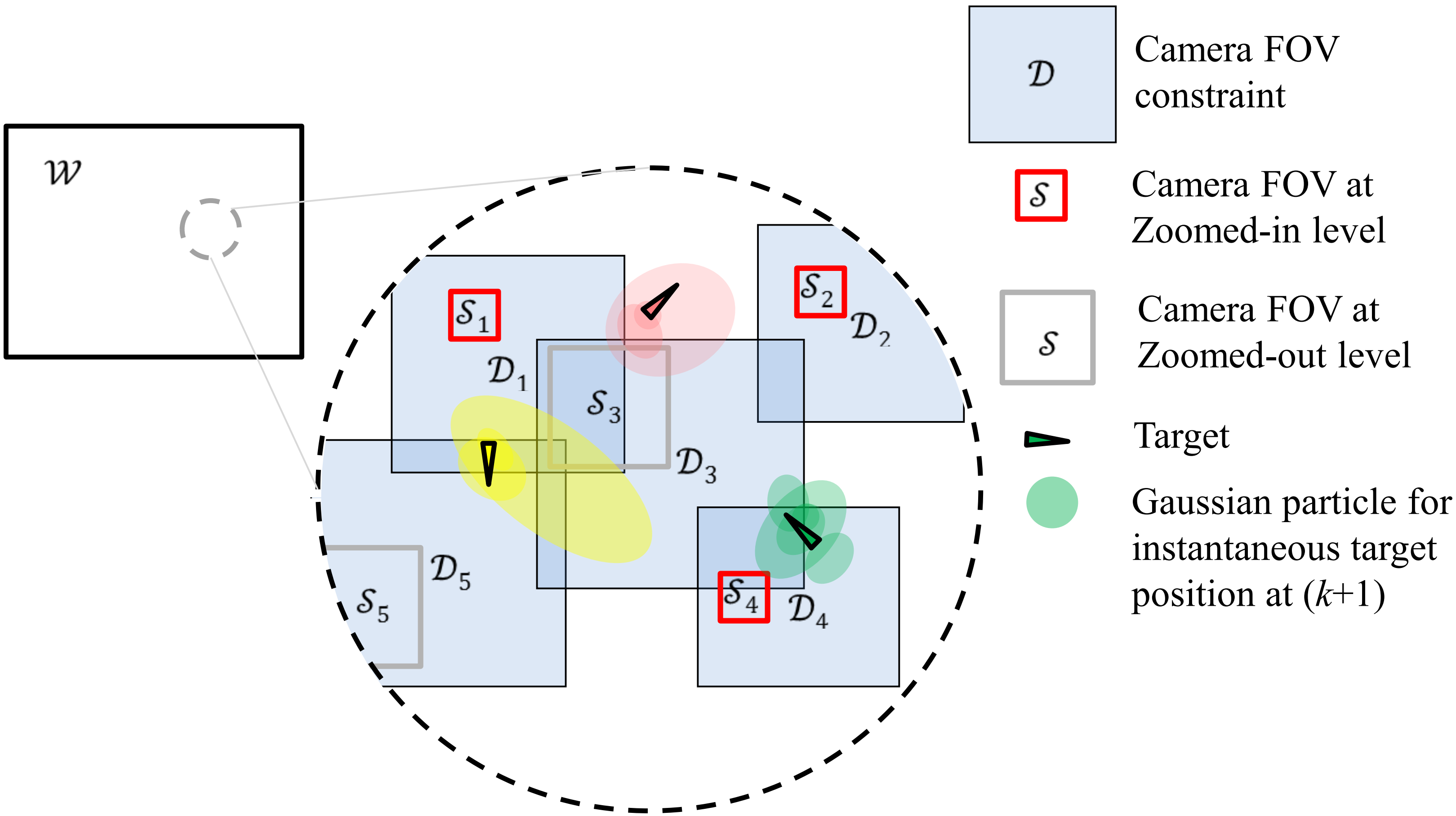


Random Search: camera PTZ levels are controlled by a random search algorithm.

Heuristic Search: camera PTZ levels are controlled such that the FoV centroid tracks the estimated position of the nearest target.

Decentralized Control for Camera-Intruder Problem

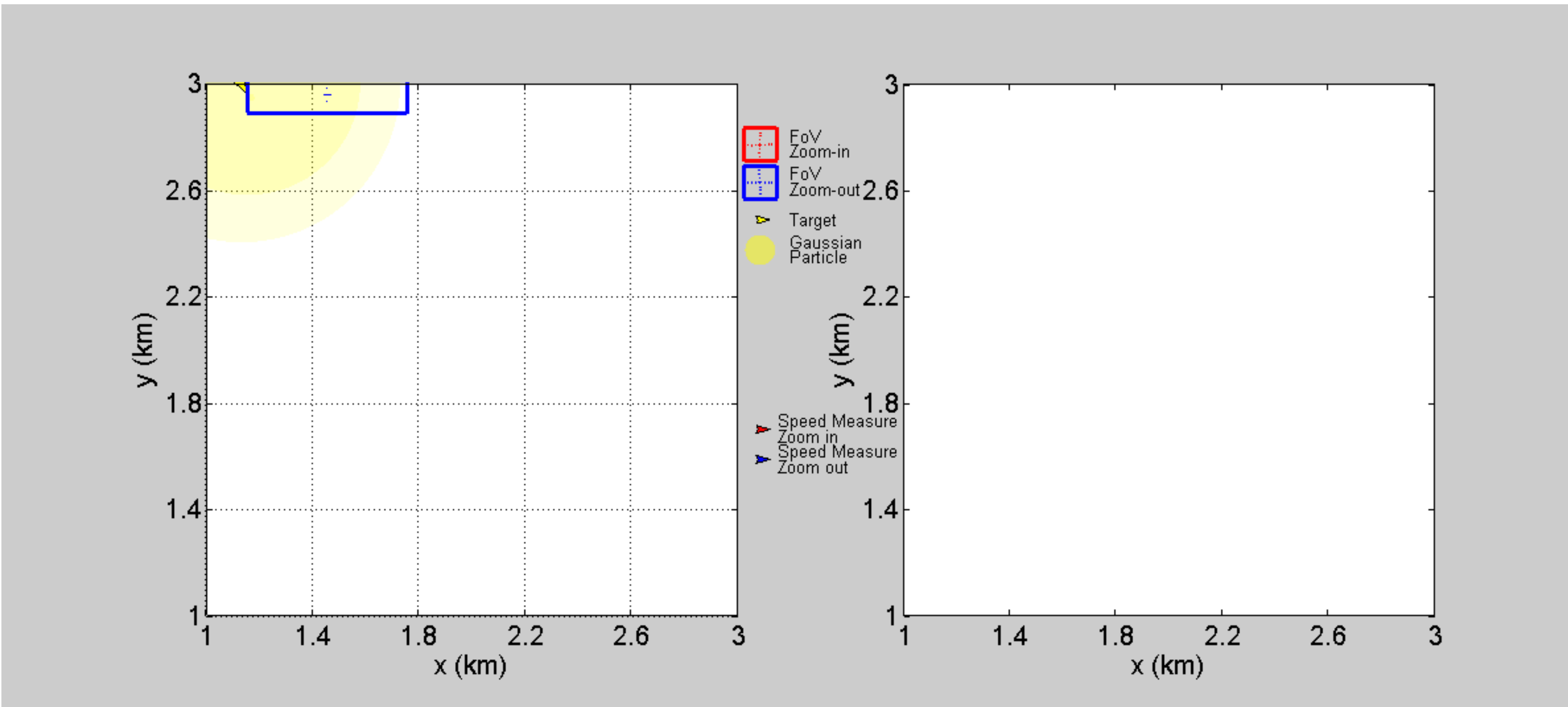
Problem Formulation: control multiple cameras for modeling multiple targets via DP-GP.



Assumptions: communications between cameras is instantaneous; each camera has a computer and wireless adaptor; each target can only be observed by a camera iff the time integral of its distribution in $\mathcal{D}$ exceeds a pre-defined threshold.

Simulation Results

- **Decentralized Active Camera Control** for mobile intruder BNP modeling and tracking, **without FOV constraints**.



Camera FoV at zoomed-out level

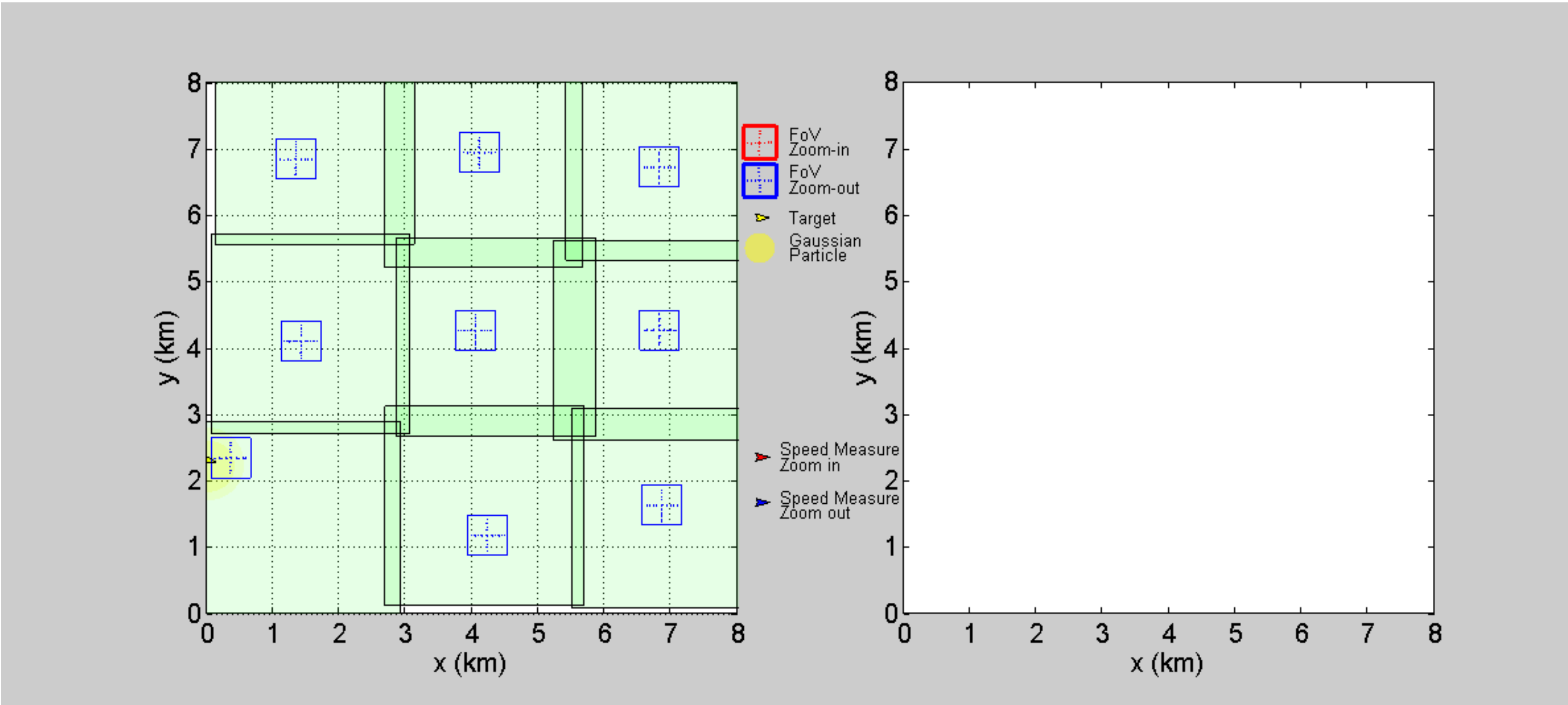
Camera FoV at zoomed-in level

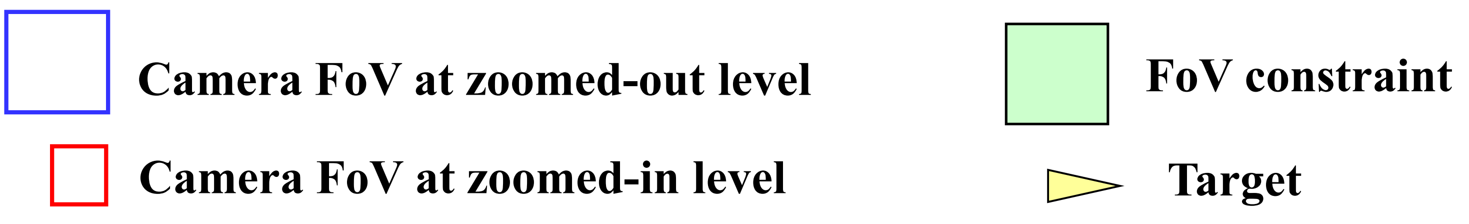
Gaussian particle

Target

Simulation Results

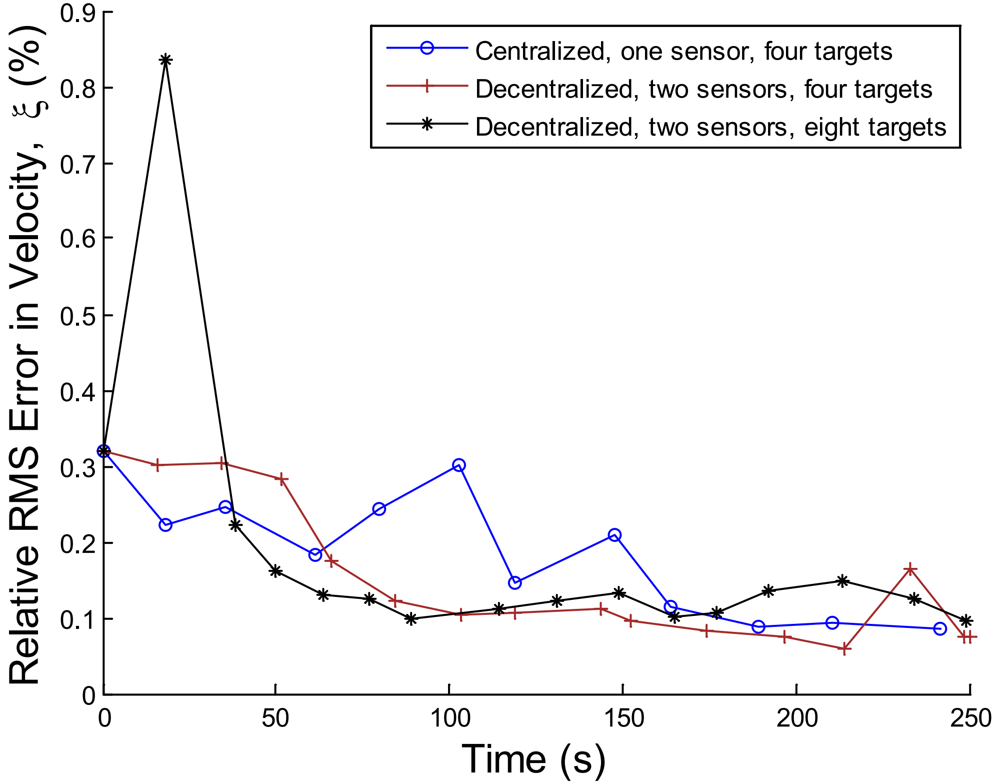
- **Decentralized Active Camera Control** for mobile intruder BNP modeling and tracking, **with FoV constraints**.





Results: Performance Comparison

Performance Evaluation: error time histories of DP-GP target velocity estimate



Heuristic Decentralized Control: camera PTZ levels are controlled by BNP-based control, using a target-camera assignment algorithm.

Decentralized BNP-based Control

- **Centralized control problem**

$$\mathbf{u}(k)=\underset{\mathbf{u}(k)}{\text{argmax}} \sum_j \hat{\varphi}_j \, [\mathcal{F};\ \mathbf{m}_j(k+1) \mid \mathcal{M}_j(k), \varepsilon(k), \mathbf{u}(k)]$$

- **Decouple:**
  - $\{\mathbf{u}_s(k)\}_j$: control set of cameras associated with target j

$$\mathbf{u}(k)=\underset{\mathbf{u}(k)}{\text{argmax}} \sum_j \hat{\varphi}_j \, [\mathcal{F};\ \mathbf{m}_j(k+1) \mid \mathcal{M}_j(k), \varepsilon(k), \{\mathbf{u}_{sj}(k)\}_j]$$

  - Decouple problem into multiple tasks:

1. Let $\mathbf{u}_{sj}(k)$ denote $\mathbf{u}_s \in \{\mathbf{u}_s(k)\}_j$, and $I_s=\{j \mid \mathbf{u}_s(k) \in \{\mathbf{u}_s(k)\}_j\}$
2. $$\sum_j \max_{\{\mathbf{u}_{sj}(k)\}_j} \hat{\varphi}_j \, [\mathcal{F};\ \mathbf{m}_j(k+1) \mid \mathcal{M}_j(k), \varepsilon(k), \{\mathbf{u}_{sj}(k)\}_j]$$

$$\text{s.t. } \forall\ \mathbf{u}_s(k) \in \{\mathbf{u}_s(k)\}_j \quad \mathbf{u}_{sj}(k) - \frac{1}{size(I_s)-1} \sum_{\ell \neq j, j \in I_s} \mathbf{u}_{s\ell}(k) = 0$$

Decouple problem into multiple tasks:

3. Augment objective function

$$\Lambda = \sum_j \{\hat{\varphi}_j[\mathcal{F};\mathbf{m}_j(k+1) \mid \mathcal{M}_j(k), \mathcal{E}(k), \{\mathbf{u}_{sj}(k)\}_j + \sum_s \lambda_{sj}{}^T[\mathbf{u}_{sj}(k) - \frac{1}{size(I_s)-1}\sum_{\ell \neq j, j \in I_s} \mathbf{u}_{s\ell}(k)]\}$$

4. Decouple, task j for target j

$$\Lambda = \sum_j \Lambda_j$$

$$\Lambda_j = \hat{\varphi}_j[\mathcal{F};\mathbf{m}_j(k+1) \mid \mathcal{M}_j(k), \mathcal{E}(k), \{\mathbf{u}_{sj}(k)\}_j + \sum_s \lambda_{sj}{}^T[\mathbf{u}_{sj}(k) - \frac{1}{size(I_s)-1}\sum_{\ell \neq j, j \in I_s} \mathbf{u}_{s\ell}(k)]$$

Summary and Future Work

Technical Accomplishments – Year 2:

- Developed **information value functions** for DP-GP models of dynamic targets
- Developed **planning/control** algorithms for DP-GP models of dynamic targets
- Developed **decentralized planning/control** algorithms for DP-GP models of dynamic targets for multiple sensors with continuous state and control, tracking multiple moving targets
- Developed **hybrid system approximate dynamic programming** (ADP) algorithm

Future Work:

- Complete decentralized BNP planning/control theory and implementation
- Extensions to DDP-GP and other BNP models
- Implementation of hybrid ADP for BNP-based control
- Decentralized BNP control convergence guarantees
- Decentralized BNP planning/control complexity analysis
- Application of decentralized BNP planning/control to mobile sensors

Acknowledgements:
This research was funded by the ONR MURI Grant N00014-11-1-0688.

?

Questions?