Robot Learning from Demonstration in the Force Domain

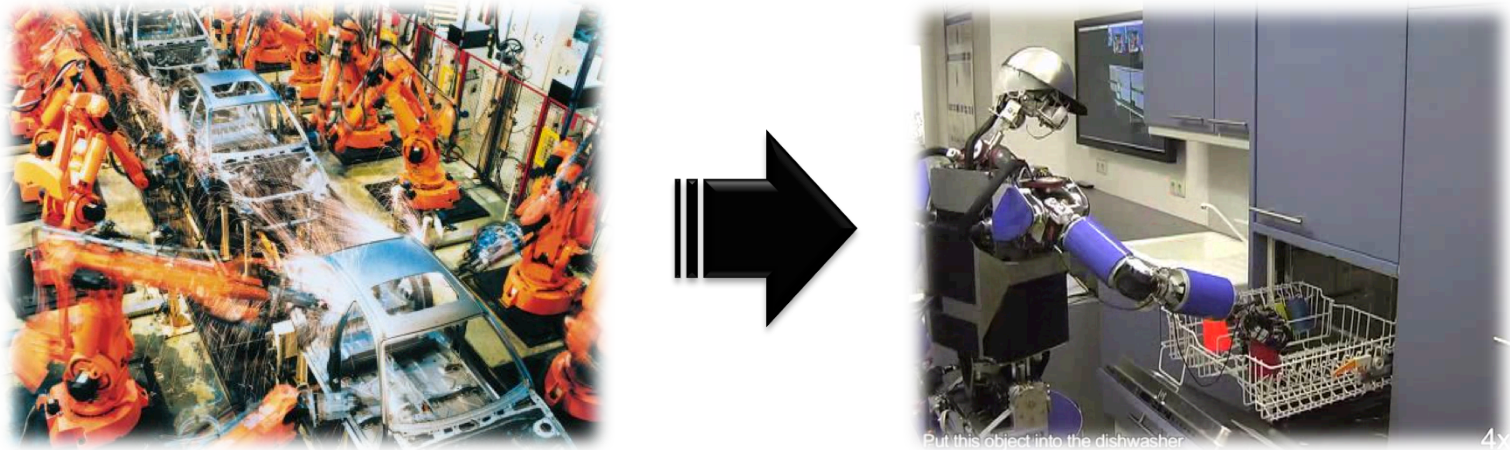
Leonel Rozo, Pablo Jiménez and Carme Torras

Institut de Robòtica i Informàtica Industrial

IJCAI 2011 Workshop on Agents Learning Interactively from Human Teachers
Barcelona, Spain – July, 2011

Introduction

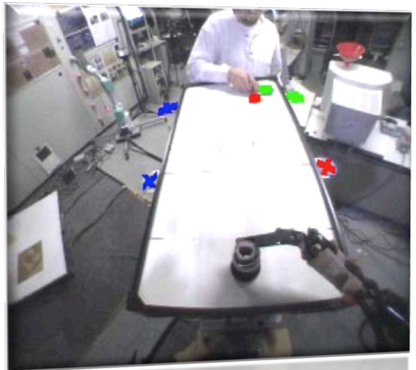
- Robots collaborating with humans at home, at work and other human-centered environments.



- It is desirable to teach robots in a *natural* and *comfortable* way.
  - Learning by demonstration (**LbD**)
  - Human-Robot Interaction (**HRI**)

Motivation

- Most of work in LbD is concerned with transferring skills to robots using:
  - Vision information
  - Kinaesthetic data
  - A priori task information

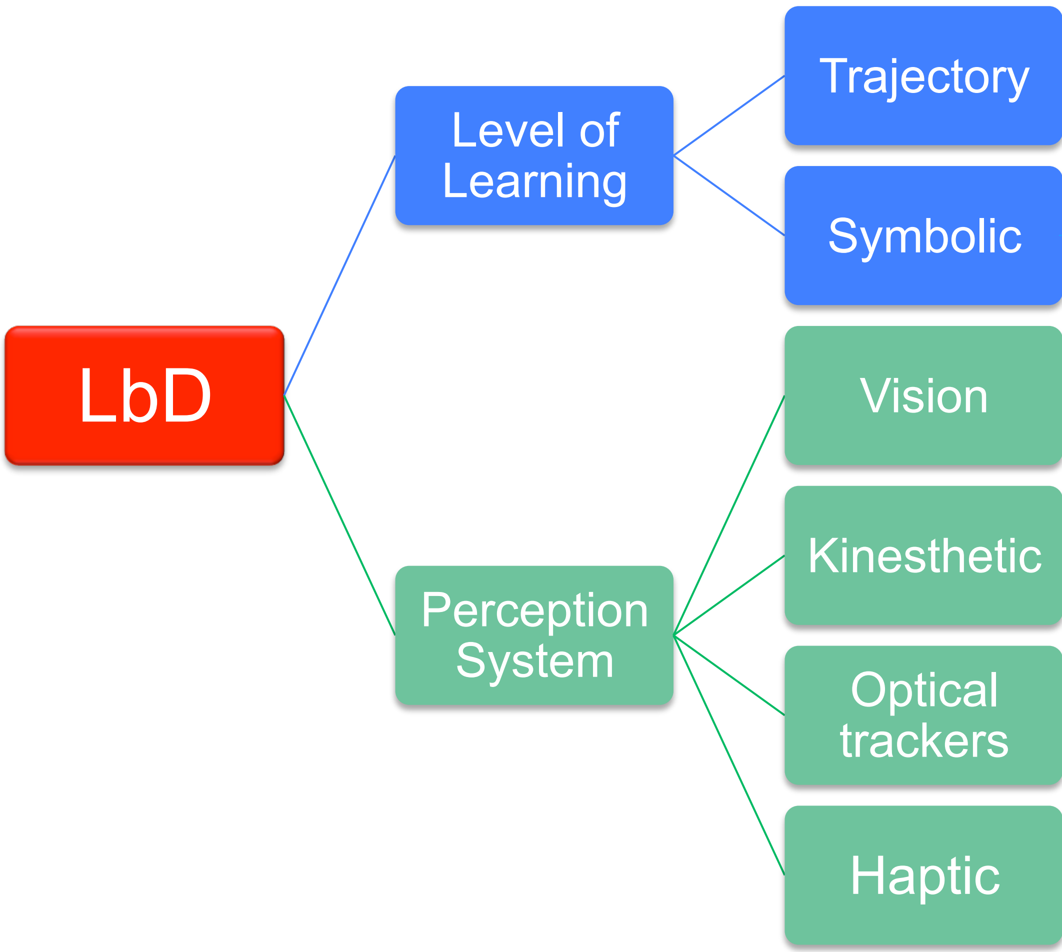


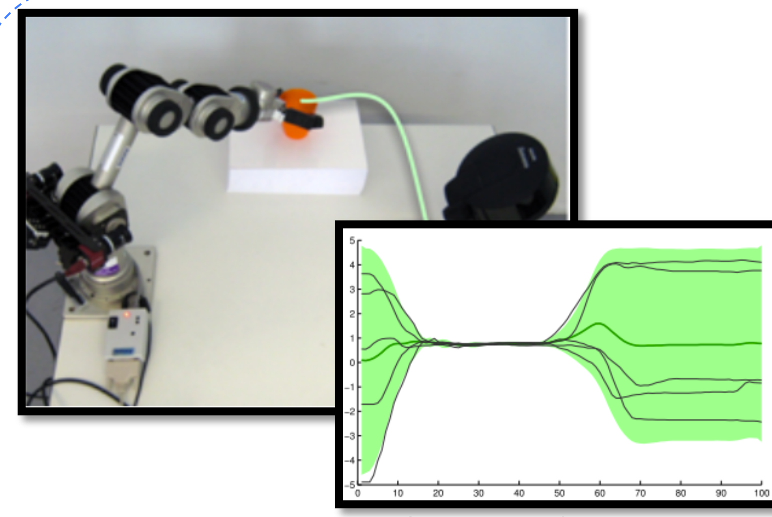
[Bentivegna et al., 2002]

[Calinon & Billard, 2008]

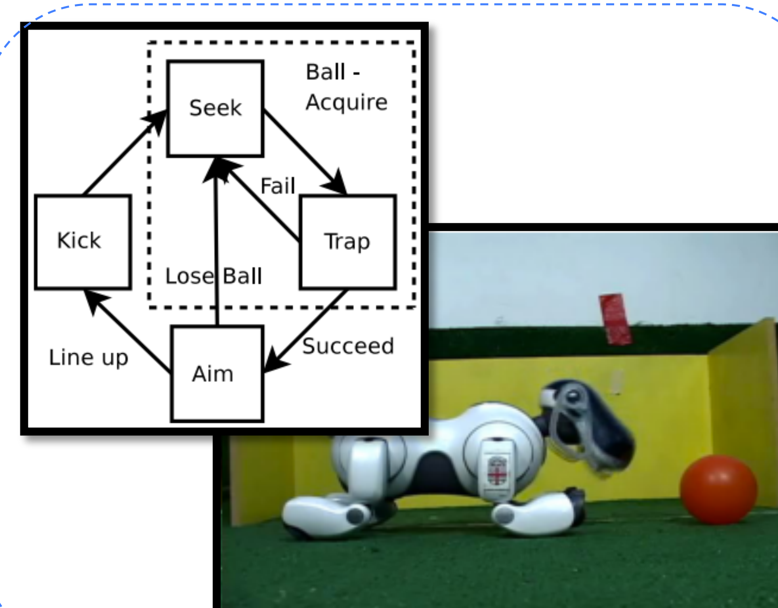
- However, there are several common tasks where forces/torques are relevant to accomplish them.

Related work





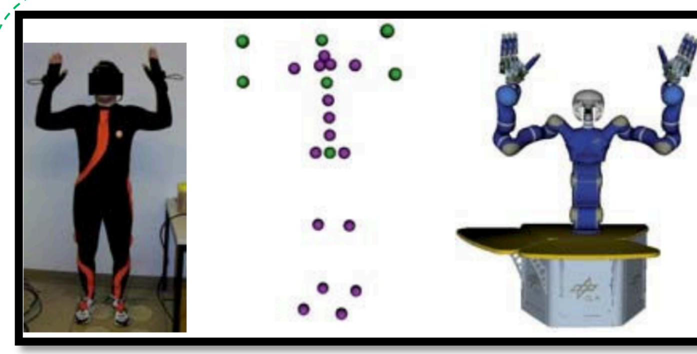
[Schneider & Ertel, 2010]



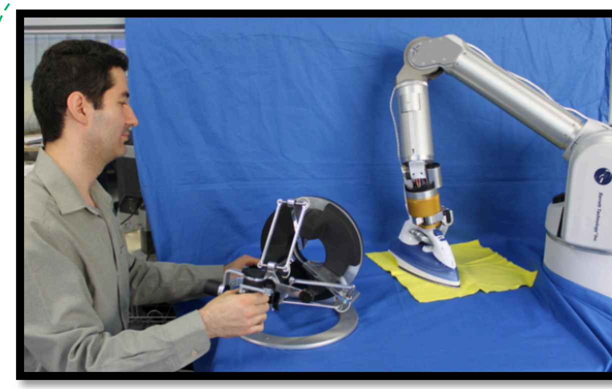
[Grollman & Jenkins, 2010]

Related work

- LbD
  - Level of Learning
    - Trajectory
    - Symbolic
  - Perception System
    - Vision
    - Kinesthetic
    - Optical trackers
    - Haptic

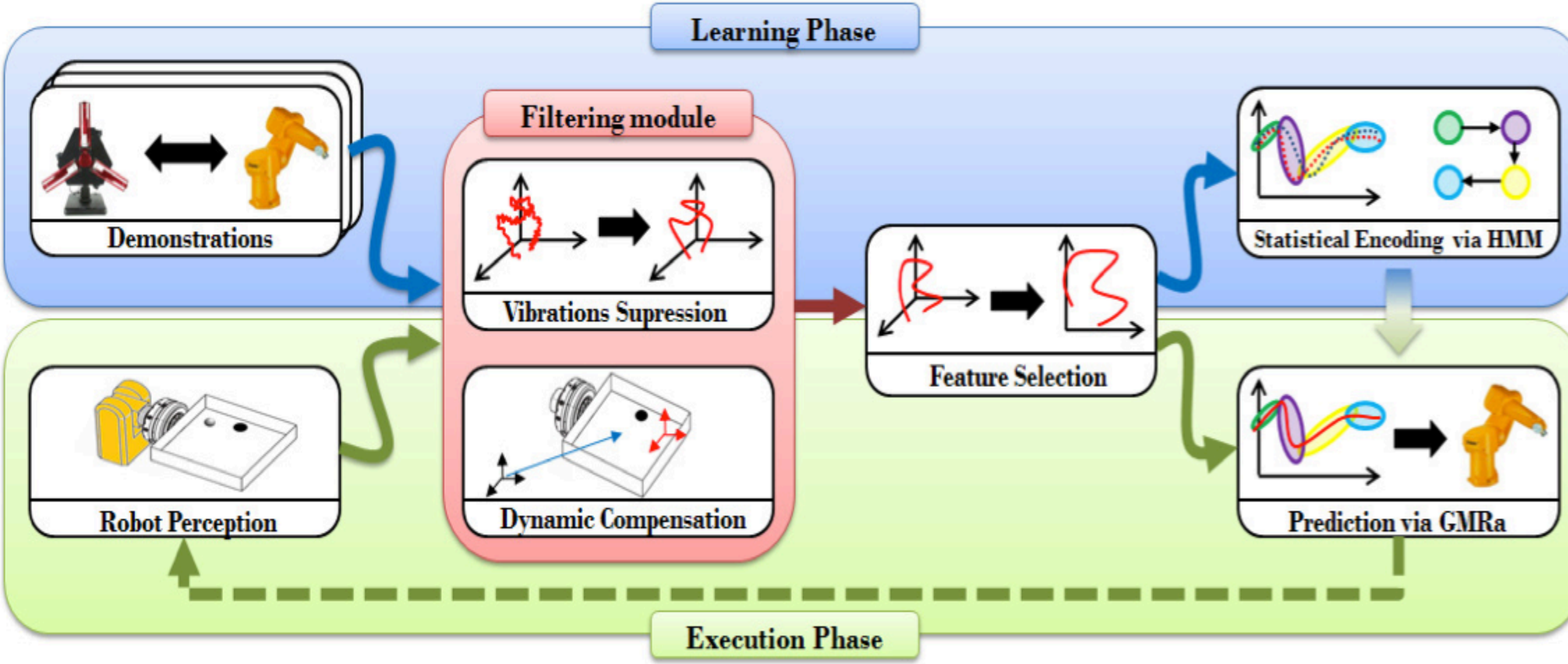


[Lee & Ott, 2010]



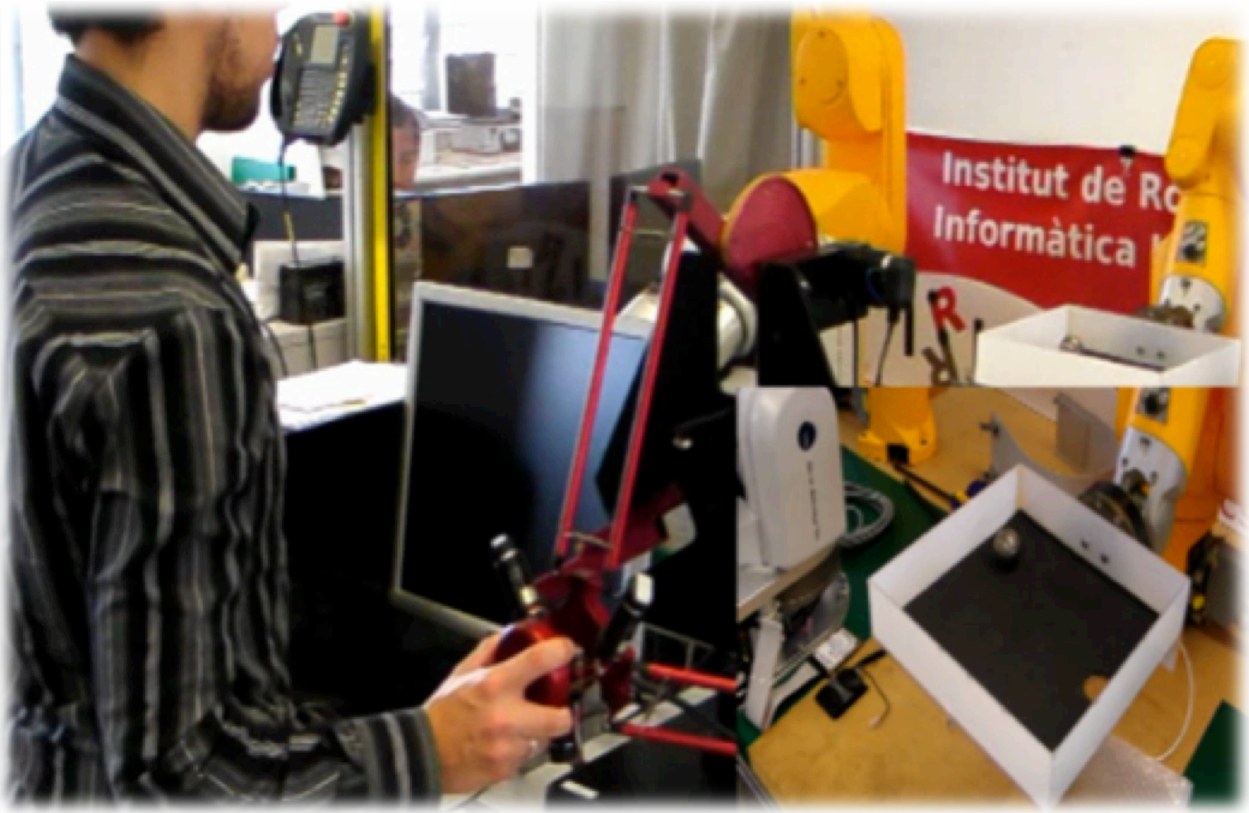
[Kormushev et al., 2011]

Our approach



The experimental setup

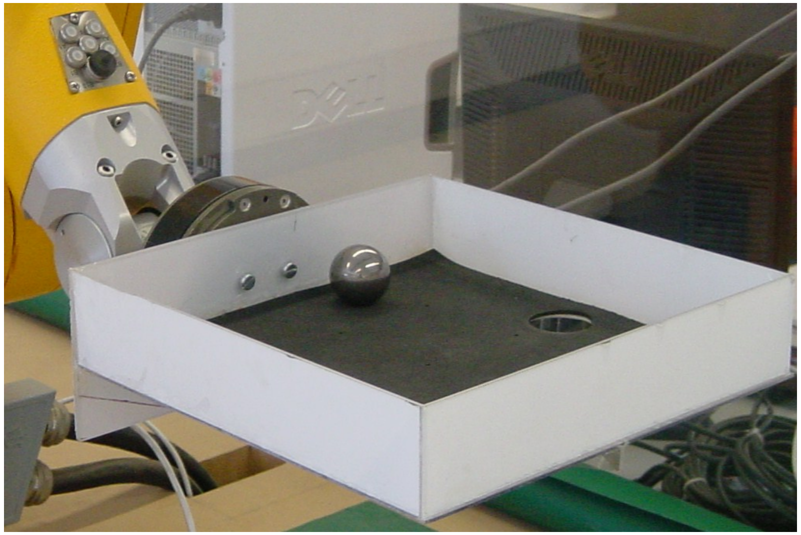
- The teacher repeatedly demonstrates the task, which consists of taking the ball out of the box through the hole, following a specific motion strategy.



Starting at some predefined initial positions, the ball is driven towards the wall adjacent to the hole, and then forced to roll along this wall to the hole.

Bidirectional communication

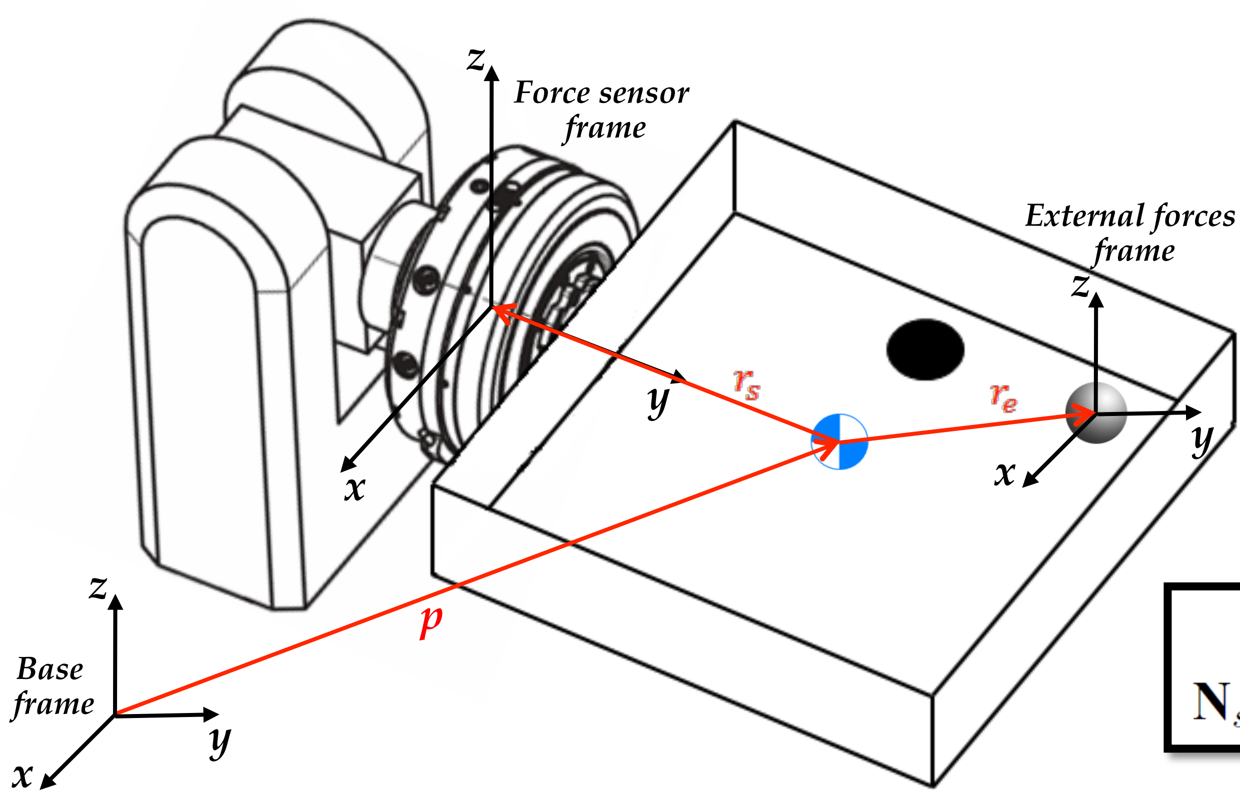
$$\boxed{\mathbf{F/N}_s = \mathbf{F/N}_b + \mathbf{F/N}_m + \varepsilon}$$



- Filtering the noise due to vibrations
  - Low-pass filter
    - Constrained least squares
    - Filter order: 75
    - Cutoff frequency: 7.5 Hz

Bidirectional communication

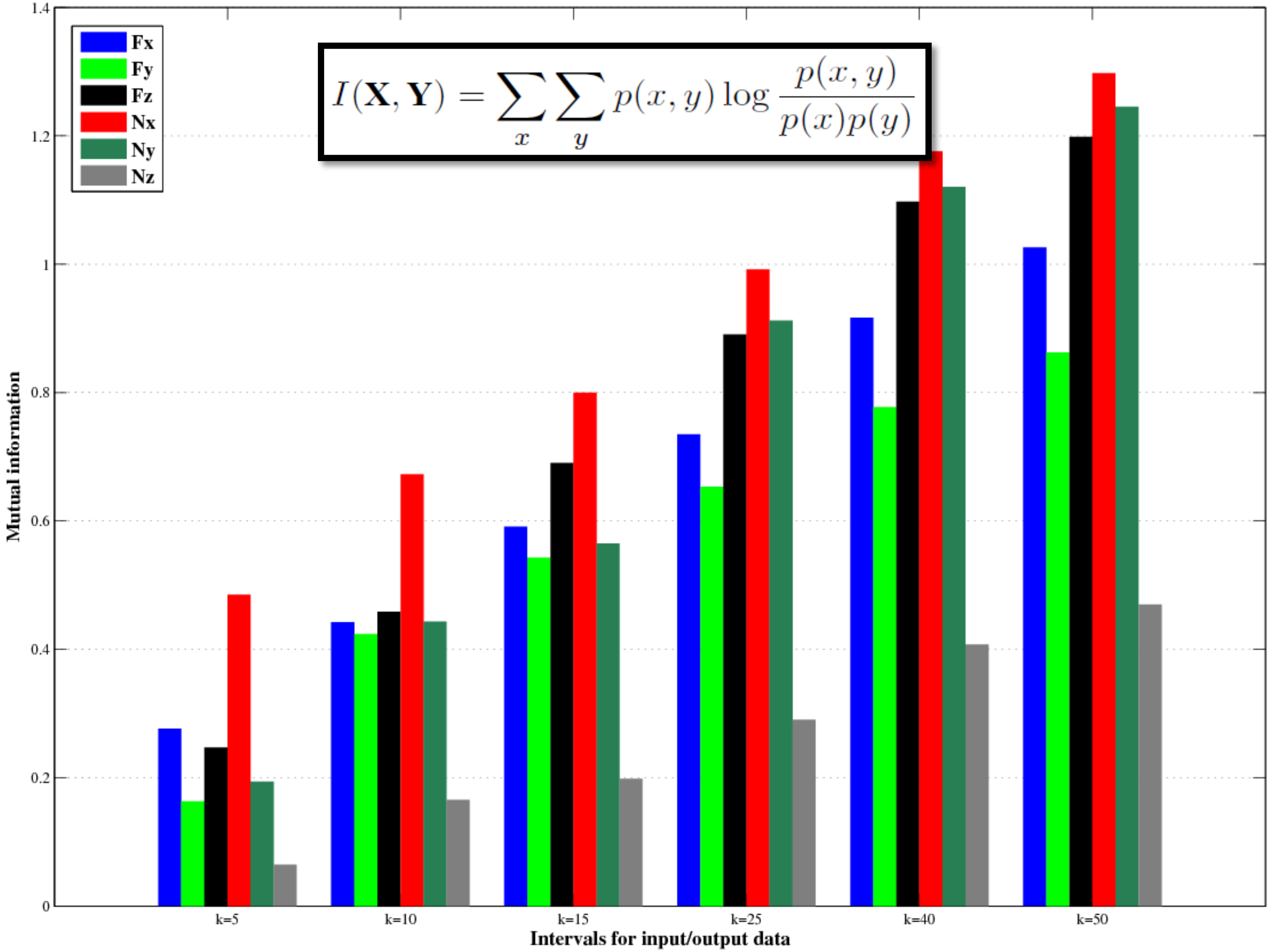
- Compensating the dynamics of the box
  - Model the forces/torques generated by the container dynamics and remove them from sensor readings.



$$
\begin{aligned}
\Sigma\mathbf{F} &= m\ddot{\mathbf{p}} = m\mathbf{g} + \mathbf{F}_e + \mathbf{F}_s \\
\Sigma\mathbf{N} &= \mathbf{I}\ddot{\mathbf{r}} + \dot{\mathbf{r}} \times \mathbf{I}\dot{\mathbf{r}} \\
&= \mathbf{N}_s + \mathbf{r}_s \times \mathbf{F}_s \\
&\quad + \mathbf{N}_e + \mathbf{r}_e \times \mathbf{F}_e
\end{aligned}
$$

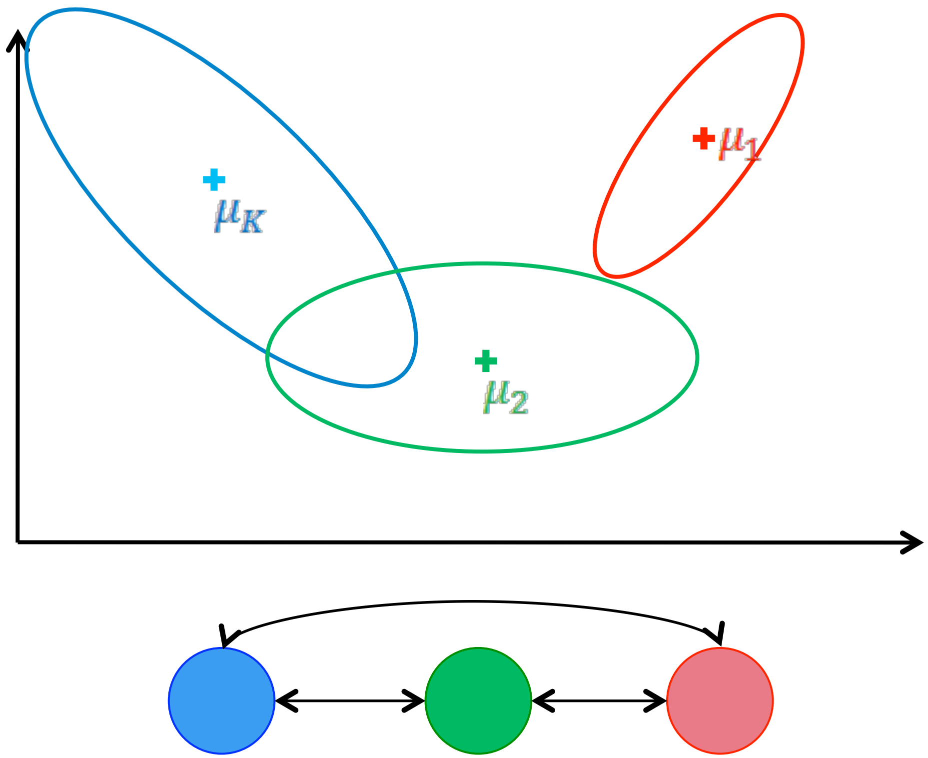
$$
\boxed{
\begin{aligned}
\mathbf{F}_s &= -m\mathbf{g} - \mathbf{F}_e \\
\mathbf{N}_s + \mathbf{r}_s \times \mathbf{F}_s &= -\mathbf{N}_e - \mathbf{r}_e \times \mathbf{F}_e
\end{aligned}
}
$$

What to Imitate?



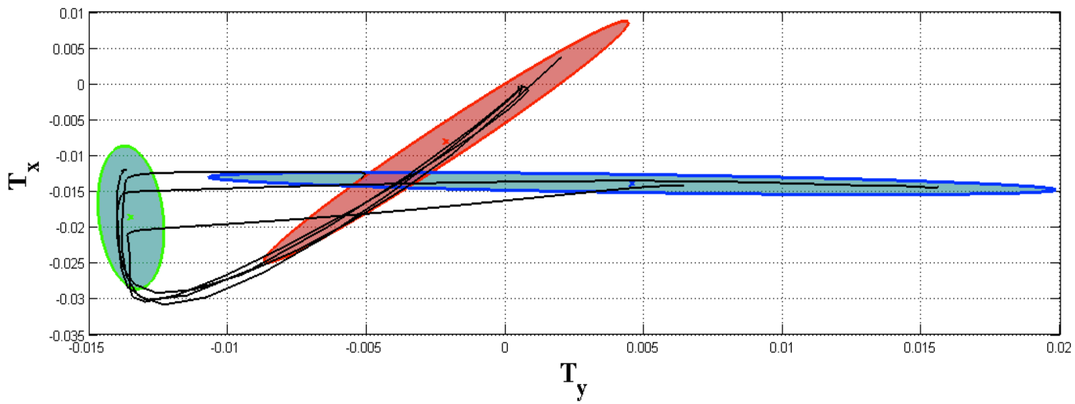
Learning the manipulation task

- Hidden Markov Model (HMM)
  - The learning framework's goal is not to learn merely a trajectory or a task with predefined states.
  - We propose to use a HMM to encode the teacher demonstrations using an ergodic topology of *N* states.
  - Selected **haptic variables** are used as *inputs* for this framework.
  - *Outputs* are the **velocity commands** at the robot joint space.

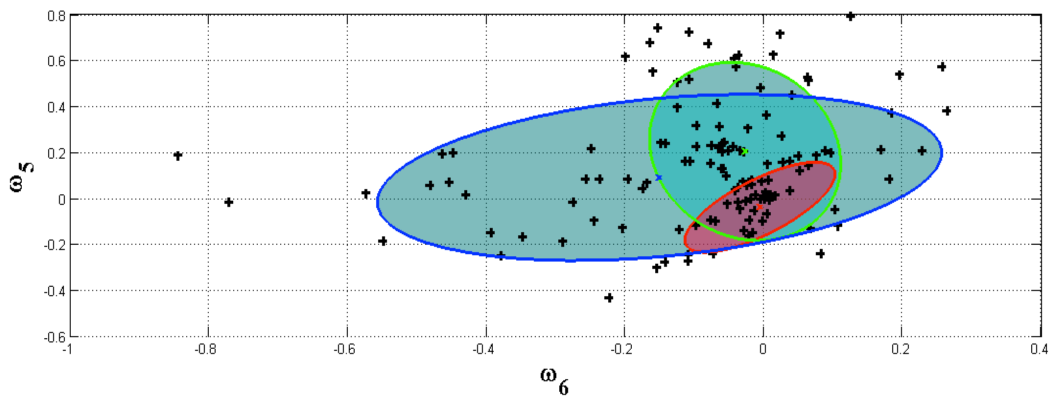


Learning the manipulation task

- Hidden Markov Model (HMM)
  - With all demonstrations, we encode the joint distribution *P*(***T***,***ω***) through an ergodic HMM defined as $\lambda=(\mathbf{A},\mathbf{B},\pi)$.



$$\mathbf{A} = \{a_{ij}\} \quad \pi = \{\pi_i\}$$
$$\mathbf{B} = \{b_j(k)\}$$

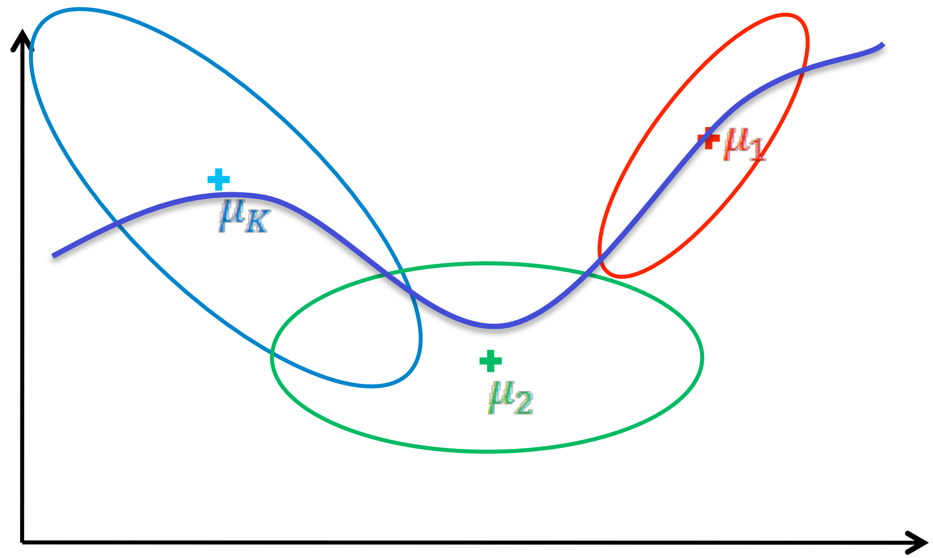


$$1 \leq i,j \leq N$$

Reproduction of the task

- Gaussian Mixture Regression (GMR)
  - Typically, using a Gaussian Mixture Model based encoding, a generalized form of the data can be recovered using GMR.
  - This classical approach does not take into account temporal information if this is not considered as an input.

$$\beta_i = \frac{p(i)p(T|i)}{\sum_{l=1}^{N} p(l)p(T|l)}$$

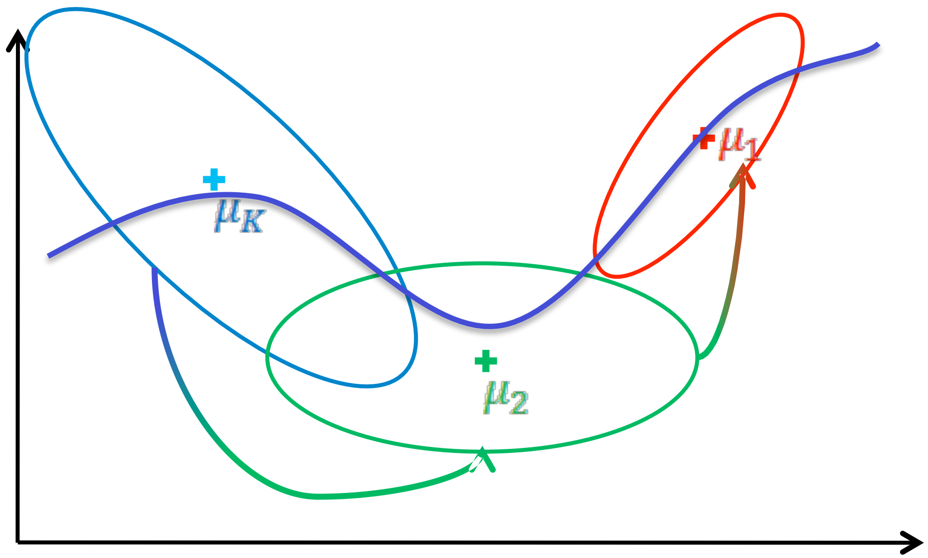


$$\hat{\omega} = \sum_{i=1}^{N} \beta_i \left[\mu_{\omega,i} + \Sigma_{\omega T,i}(\Sigma_{TT,i})^{-1}(T - \mu_{T,i})\right]$$

Reproduction of the task

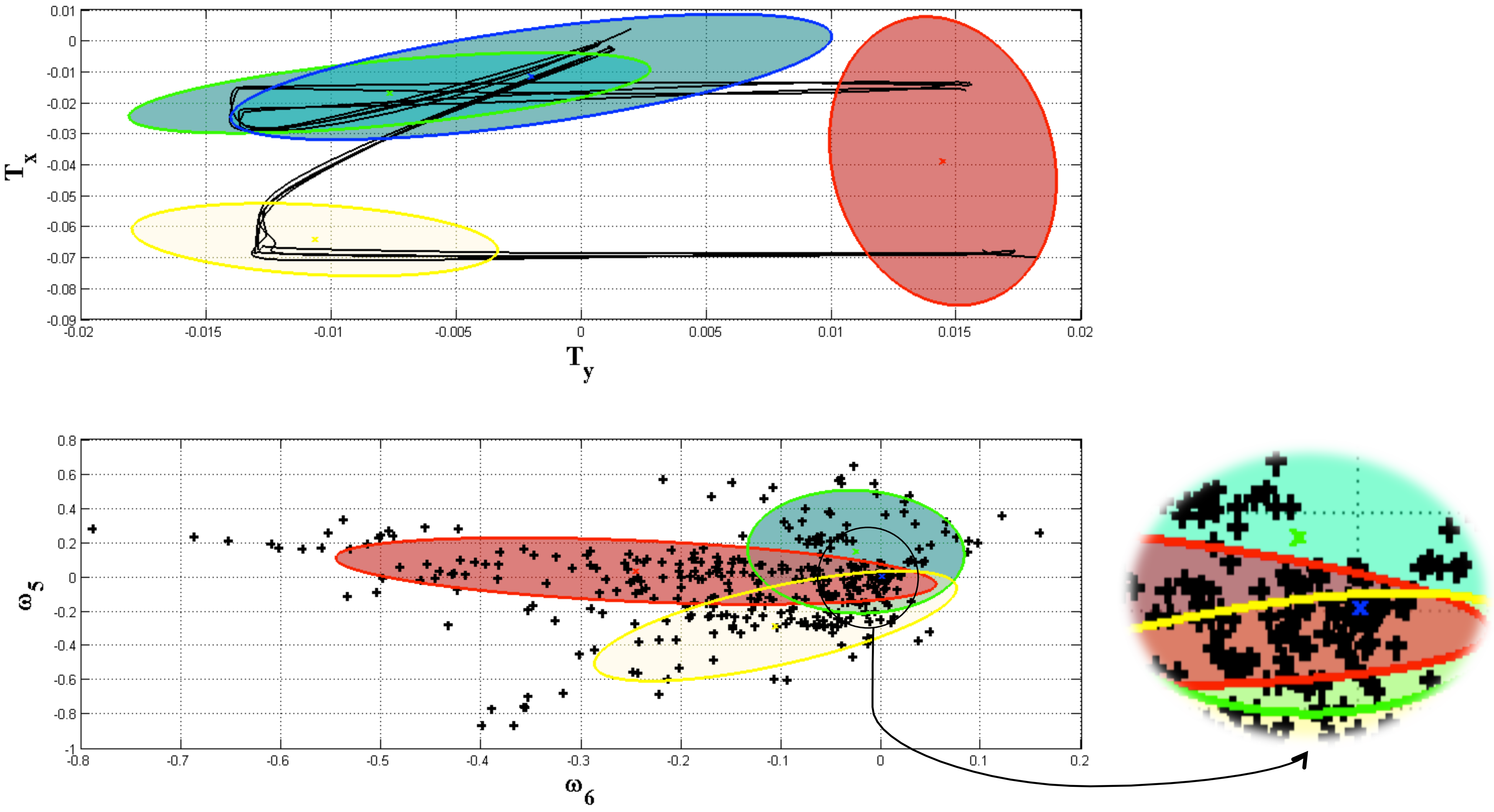
- Gaussian Mixture Regression-alpha (GMRa)
  - GMRa [Calinon et al., 2010] can provide a better estimation by using a weight that takes into account both haptic and sequential information.
  - The predicted velocity is based on current and past observations, useful when more than one output exists for a given input.

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1})$$



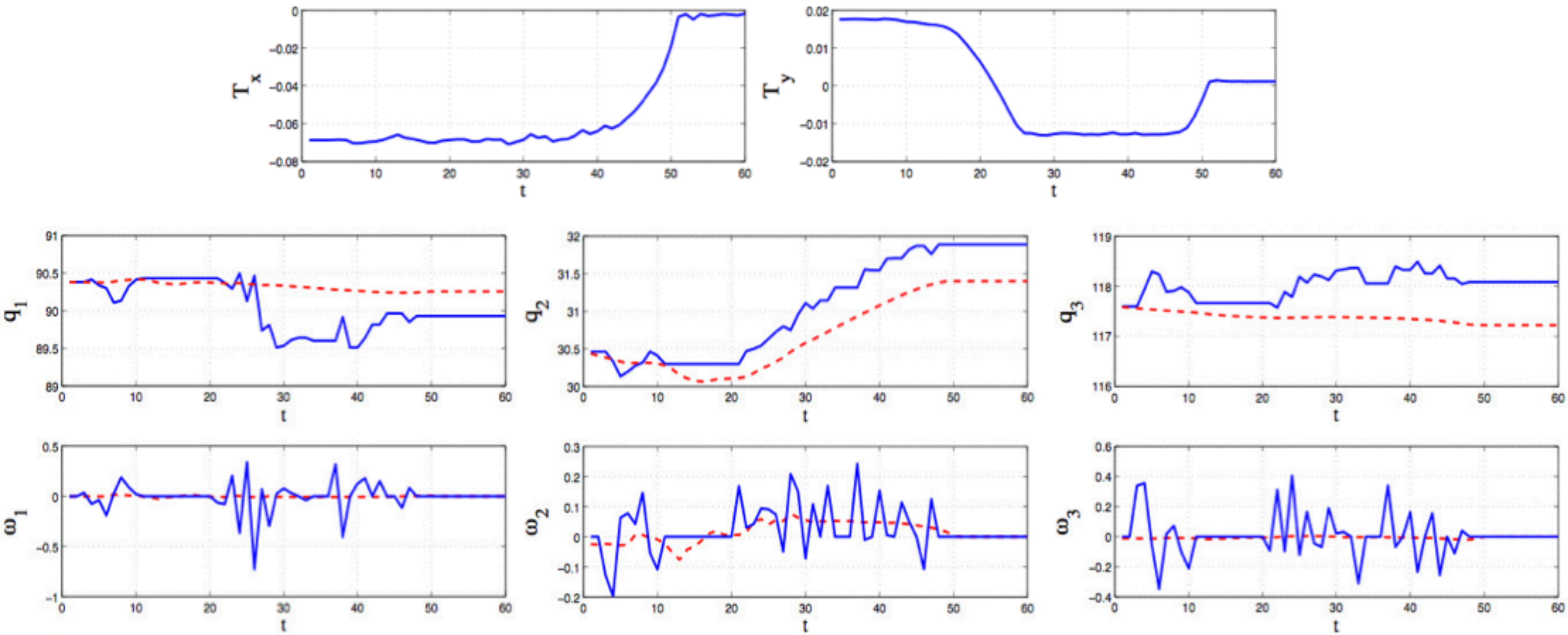
$$\hat{\omega} = \sum_{i=1}^{N} \alpha(i) \left[\mu_{\omega,i} + \Sigma_{\omega T,i}(\Sigma_{TT,i})^{-1}(T - \mu_{T,i})\right]$$

Results

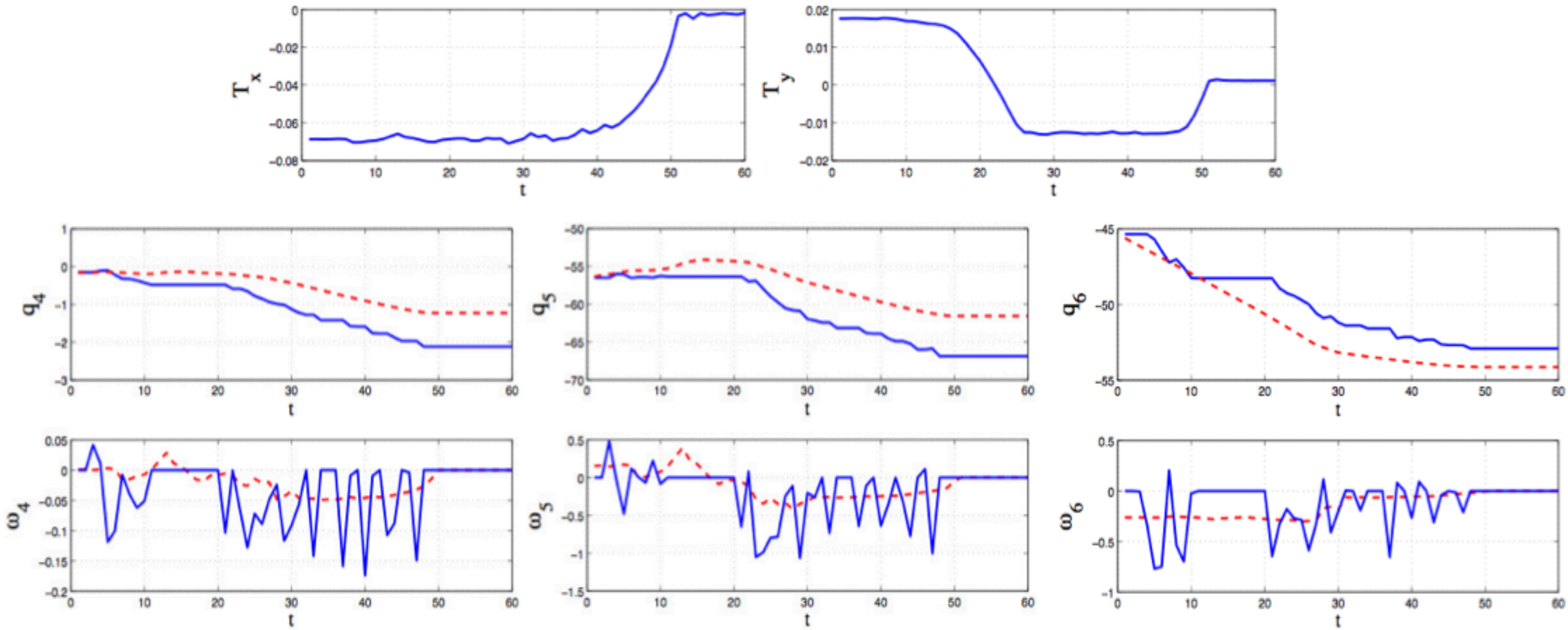


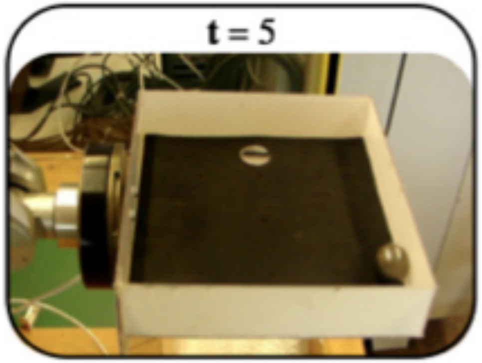
0.01
0
-0.01
-0.02
-0.03
-0.04
-0.05
-0.06
-0.07
-0.08
-0.09
-0.02
-0.015
-0.01
-0.005
0
0.005
0.01
0.015
0.02
T_x
T_y
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-0.8
-0.7
-0.6
-0.5
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
ω_5
ω_6

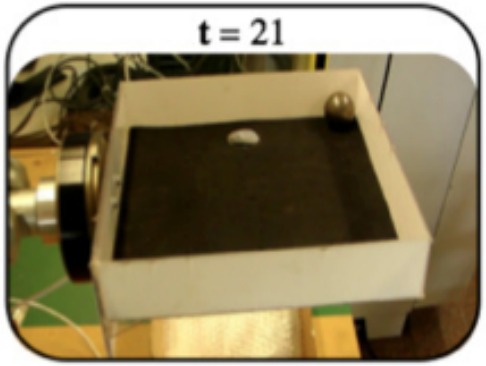
Results

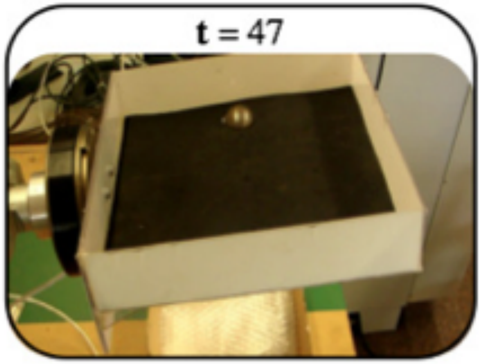


Results

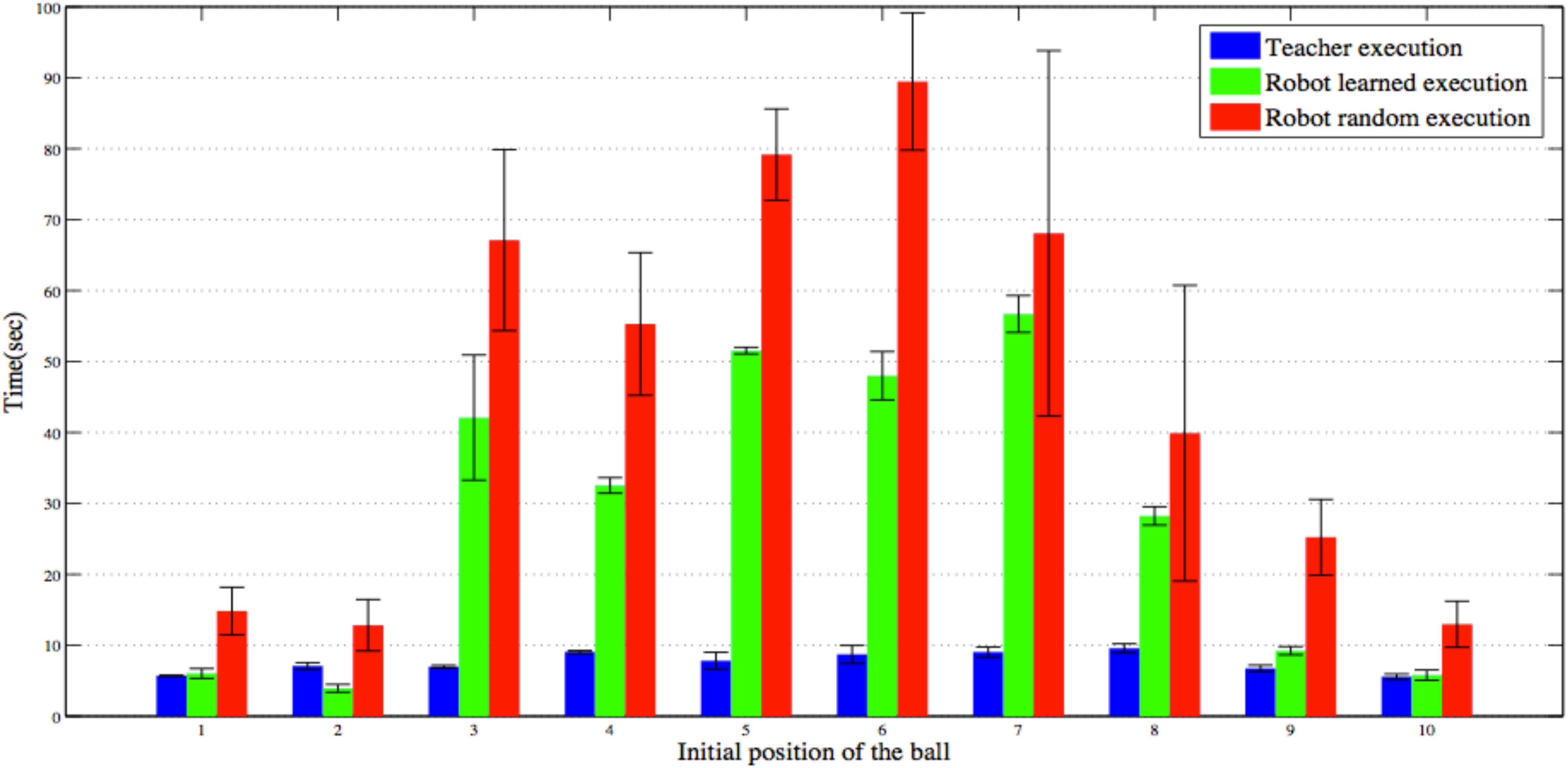




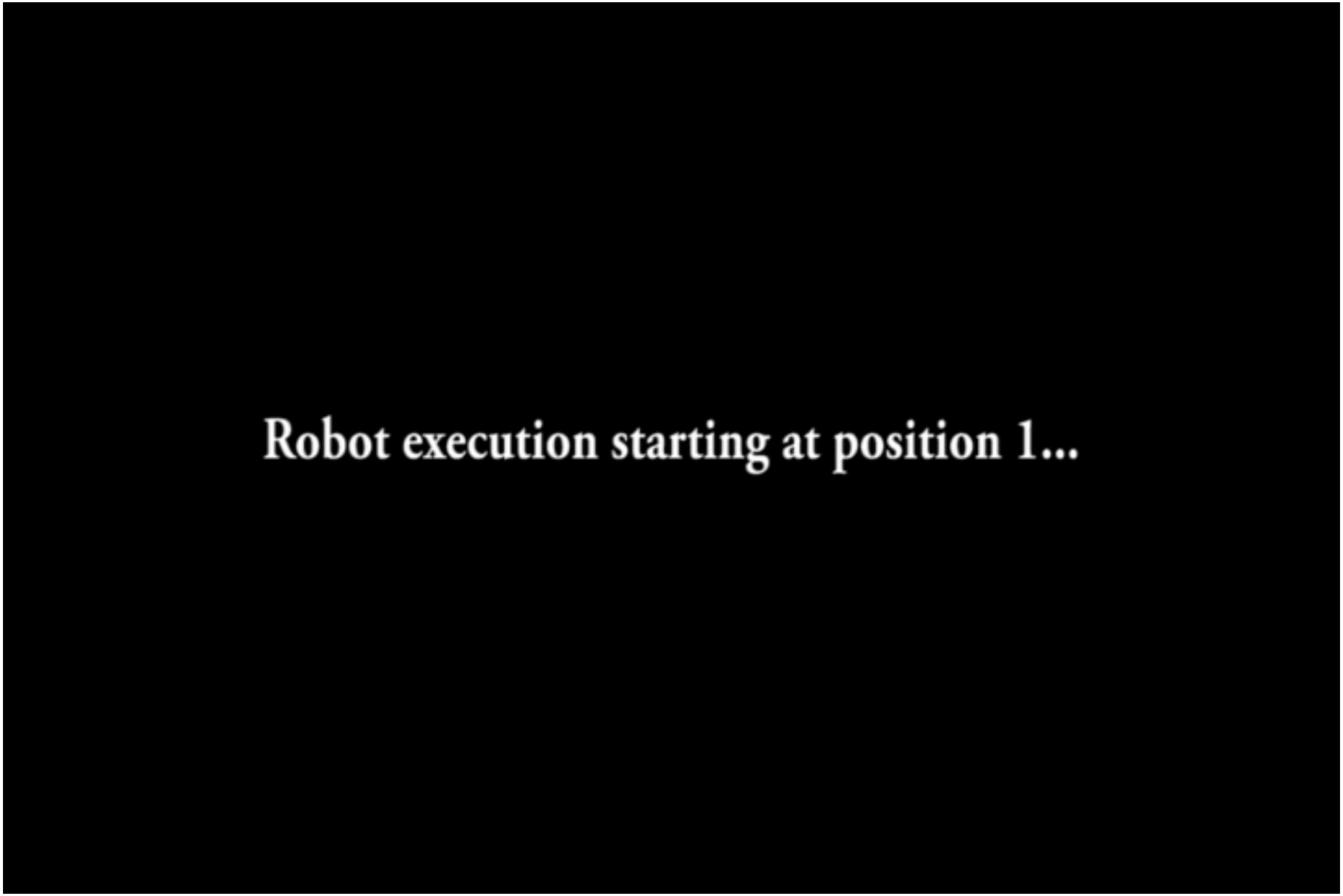




Measuring robot performance



Results



Conclusions & Future Work

- Conclusions:
  - **Force/Torque feedback** constitutes valuable input source for learning manipulation skills.
  - **MI** allows to solve the *What to Imitate?* problem.
  - **HMM/GMRa** allows to learn multi-valued functions based on sequential information, which was extended to teach a **multi-trajectory task**.
- Future work:
  - To apply a similar framework on more realistic settings.
  - To find better performance measures.
  - To take advantage of compliant robots for carrying out force/torque-based active and cooperative learning.

Thanks !