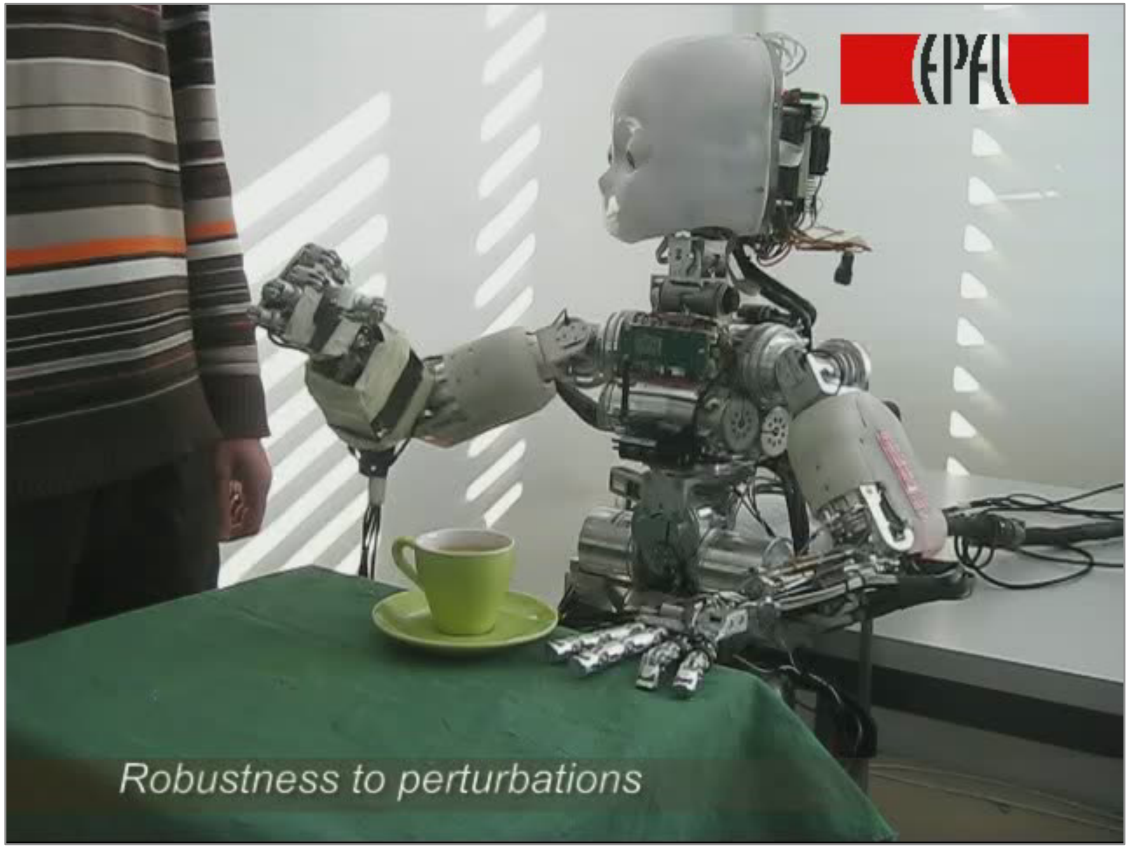
Programming by Demonstration: Some recent challenges

Aude G. Billard
LASA Laboratory
EPFL

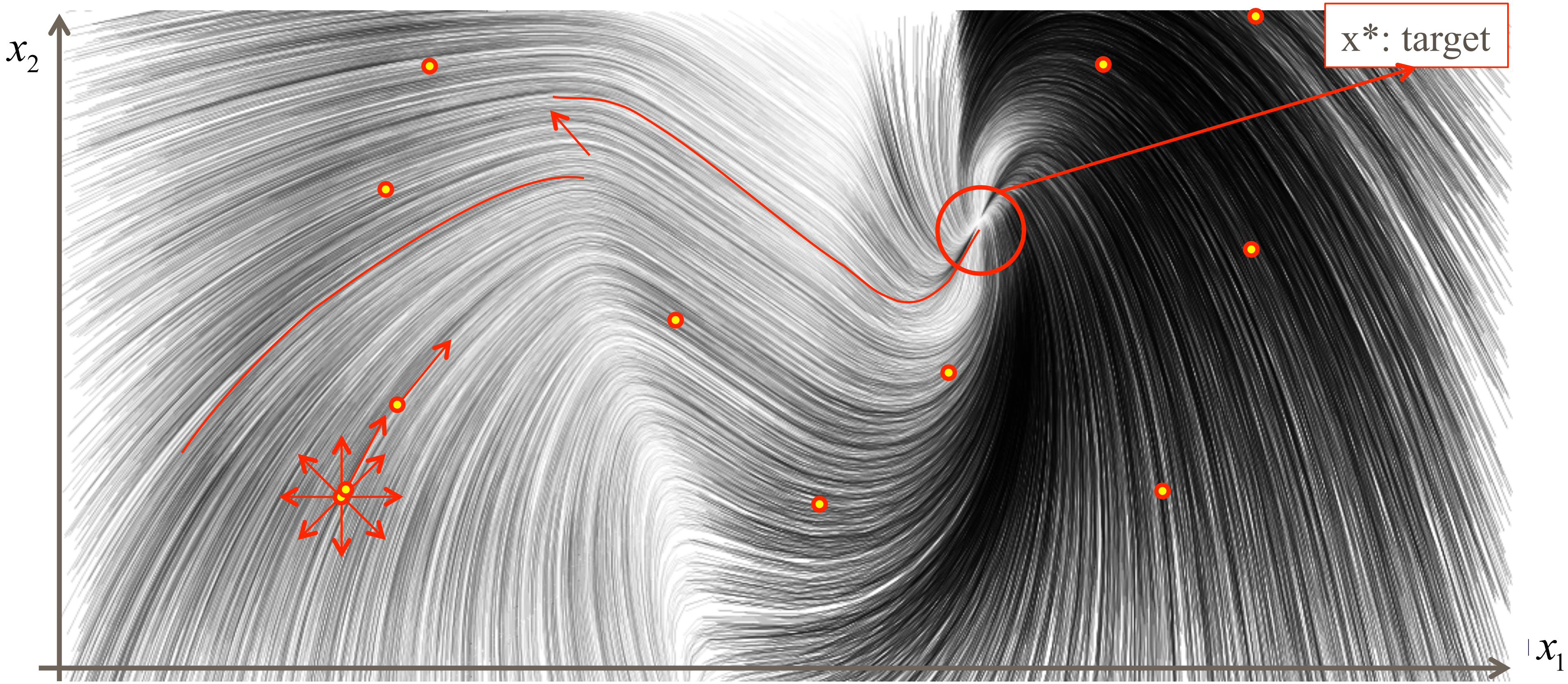
Generalizing: Learning a control law



Learning a control law that ensures that you reach the target even if perturbed and that you follow a particular dynamics

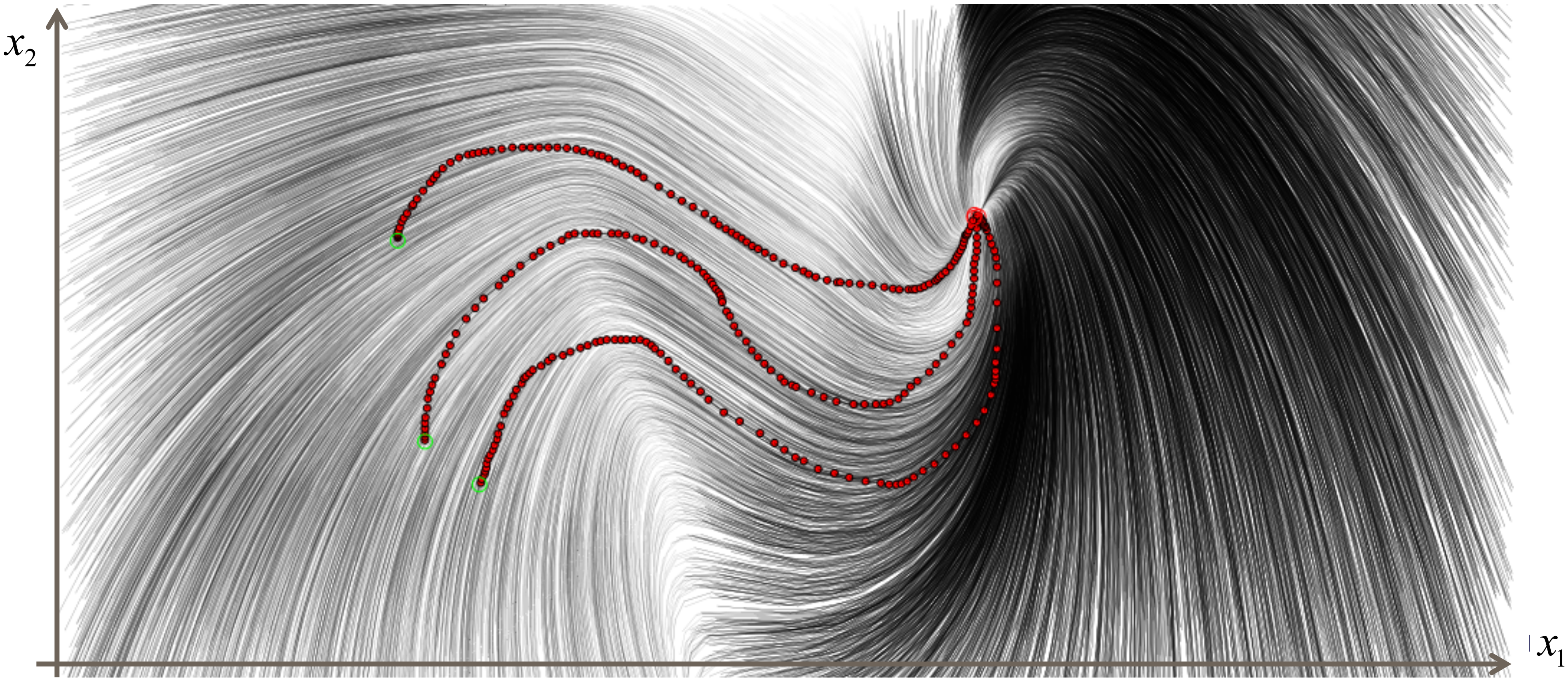
Learn a control law from examples

Time-invariant DS $\dot{x} = f(x)$ *with stable attractor* $\dot{x}^* = f(x^*) = 0.$



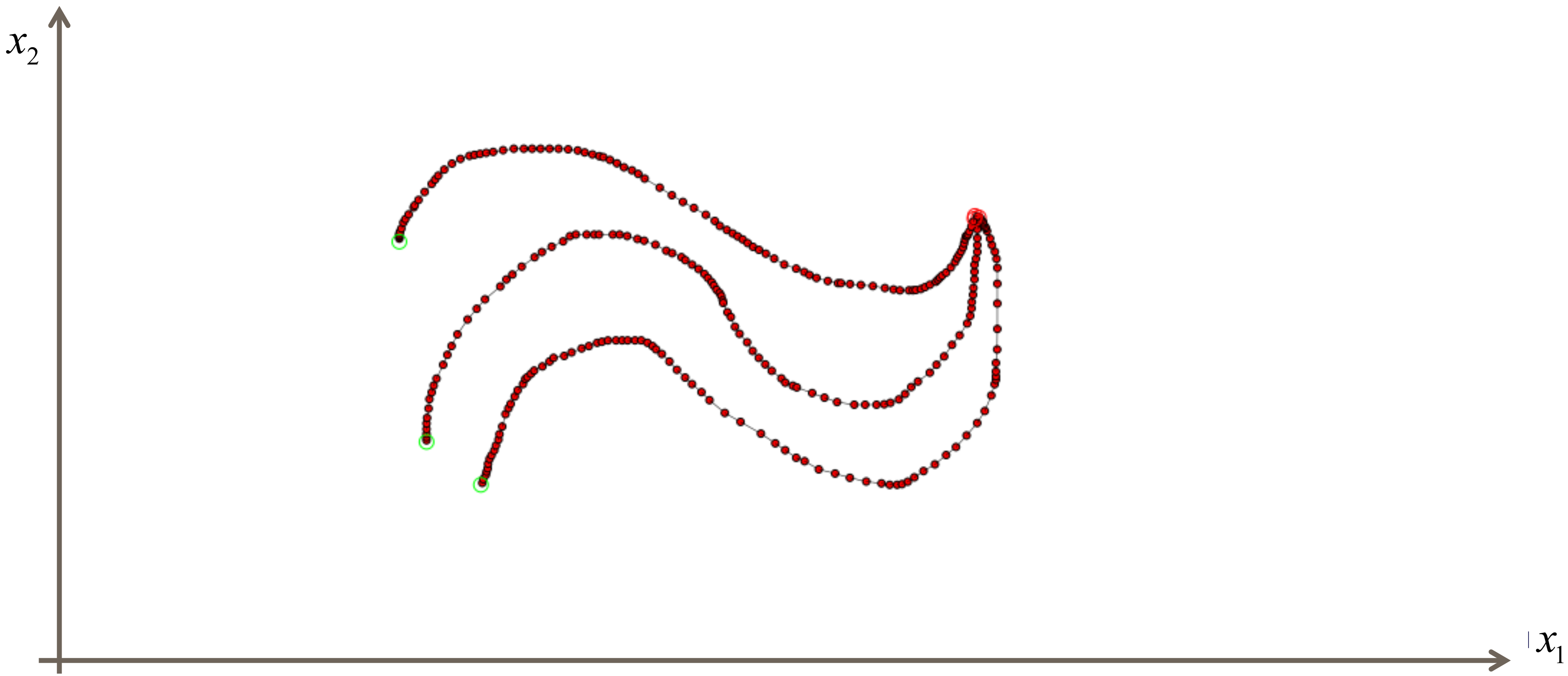
Learn a control law from examples

Make N observations of the state of the system $\{\dot{x}^i, x^i\}, i = 1 \ldots N$.



Learn a control law from examples

Make N observations of the state of the system $\{\dot{x}^i, x^i\}, i = 1 \ldots N$.

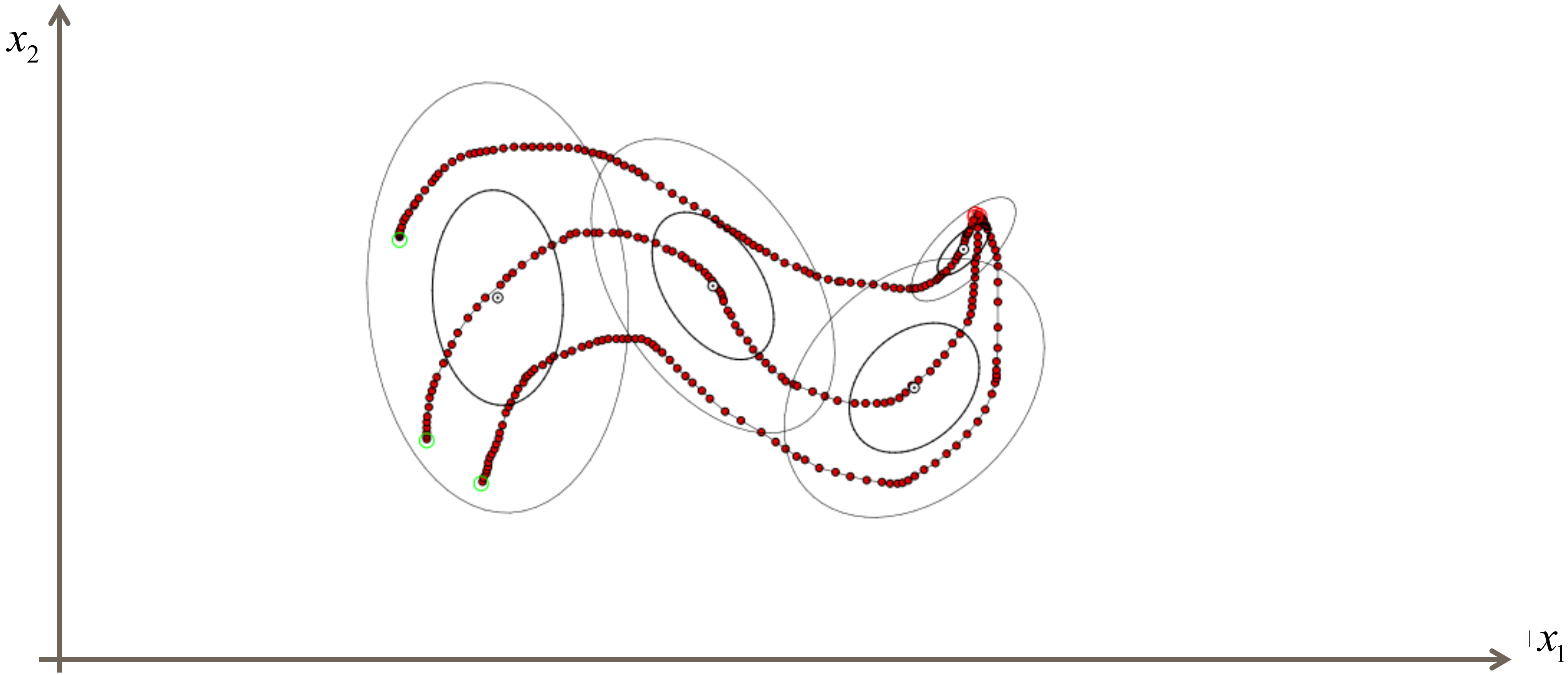


Learn a control law from examples

Make N observations of the state of the system $\left\{\dot{x}^i, x^i\right\}, i = 1 \dots N$.

Learn $p\left(\dot{x}, x\right) = \sum_{k=1}^{K} w_k N\left(\dot{x}, x; \mu^k, \Sigma^k\right)$: joint density (mixture of Gaussians)

describing the distribution of velocity in state space. $p(x) \sim N(x; \mu, \Sigma)$

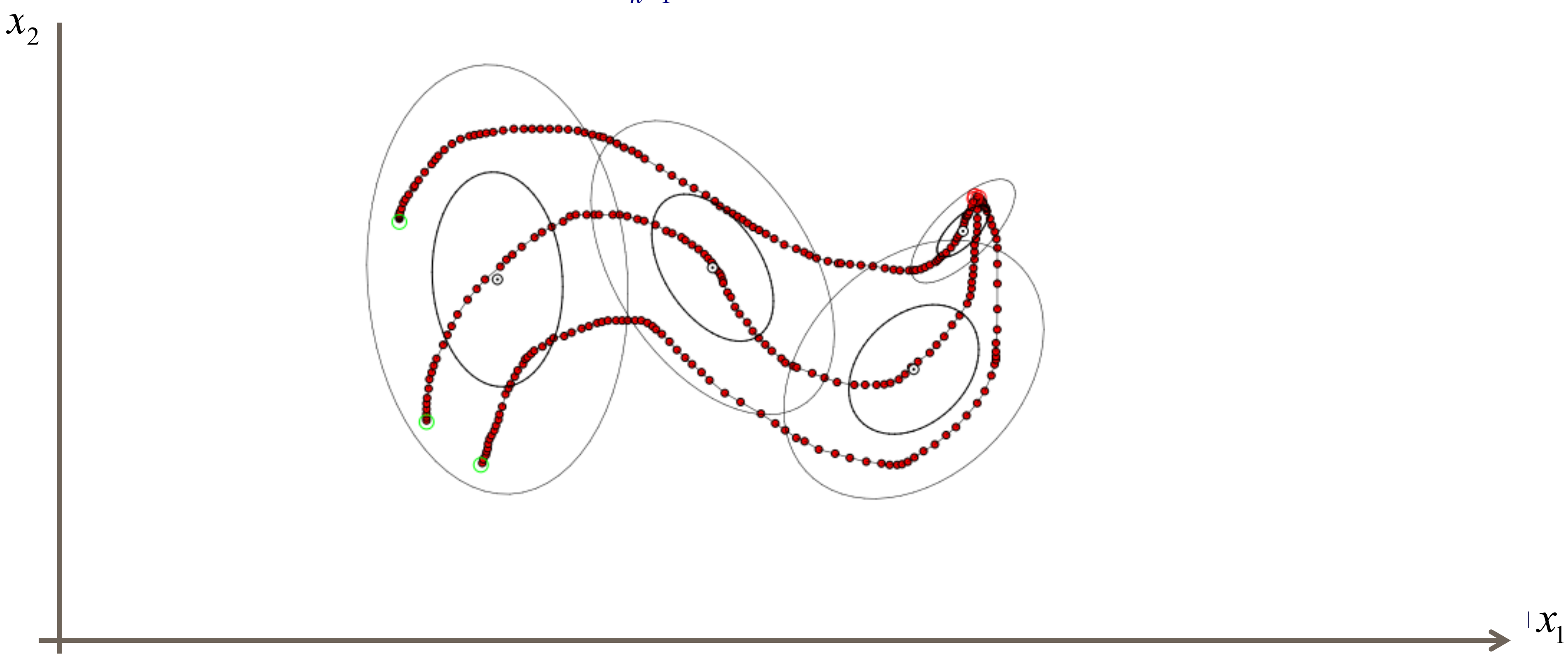


Learn a control law from examples

Make N observations of the state of the system $\left\{\dot{x}^i, x^i\right\}, i = 1 \ldots N$.

Learn $p\left(\dot{x}, x\right) = \sum_{k=1}^{K} w_k N\left(\dot{x}, x; \mu^k, \Sigma^k\right)$: joint density (mixture of Gaussians)

Compute $f\left(x\right) = E\left\{p\left(\dot{x} \mid x\right)\right\} = \sum_{k=1}^{K} h^k\left(x\right)\left(A^k x + b^k\right)$ (analytical expression for f)

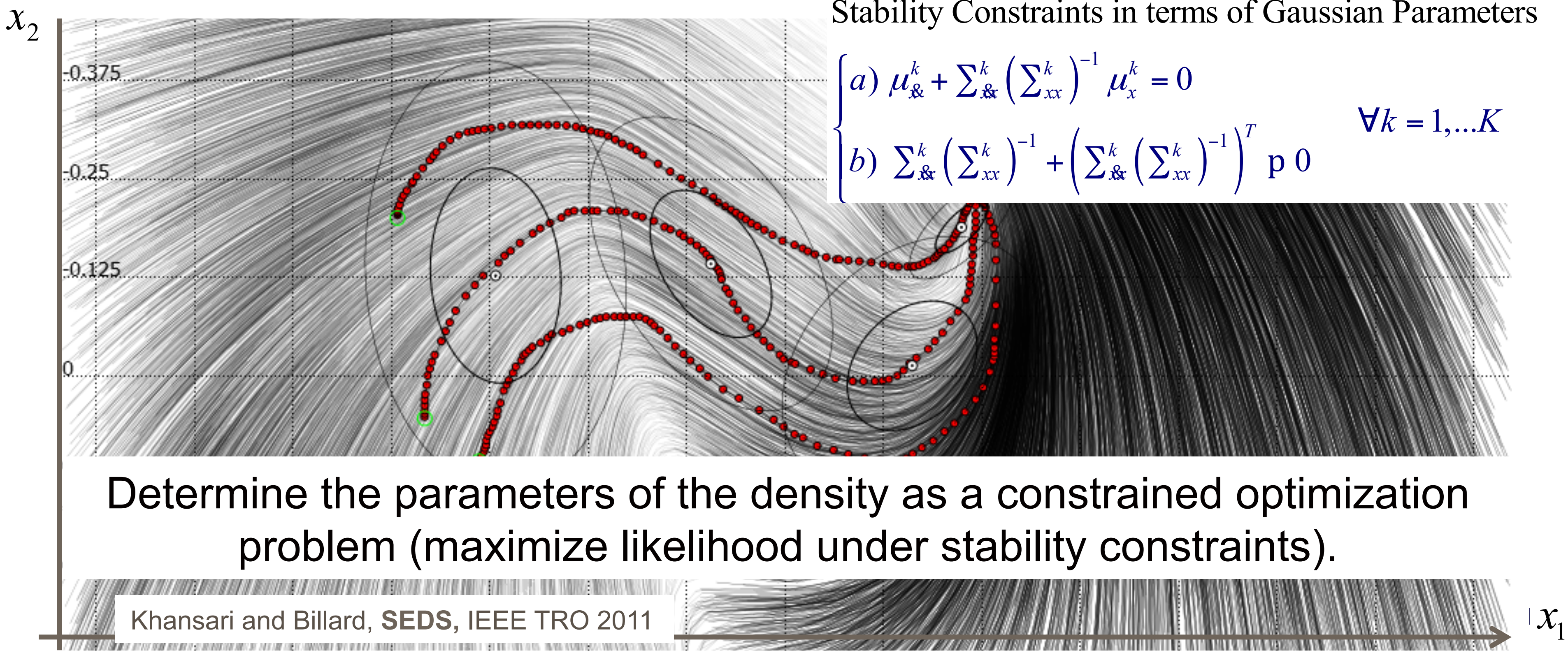


Learn a control law from examples

Make N observations of the state of the system $\left\{\dot{x}^i, x^i\right\}, i = 1 \ldots N$.

Learn $p(\dot{x}, x) = \sum_{k=1}^{K} w_k N\left(\dot{x}, x; \mu^k, \Sigma^k\right)$: joint density (mixture of Gaussians)

Compute $f(x) = E\left\{p(\dot{x} \mid x)\right\} = \sum_{k=1}^{K} h^k(x)\left(A^k x + b^k\right)$ (analytical expression for f)



Stability Constraints in terms of Gaussian Parameters

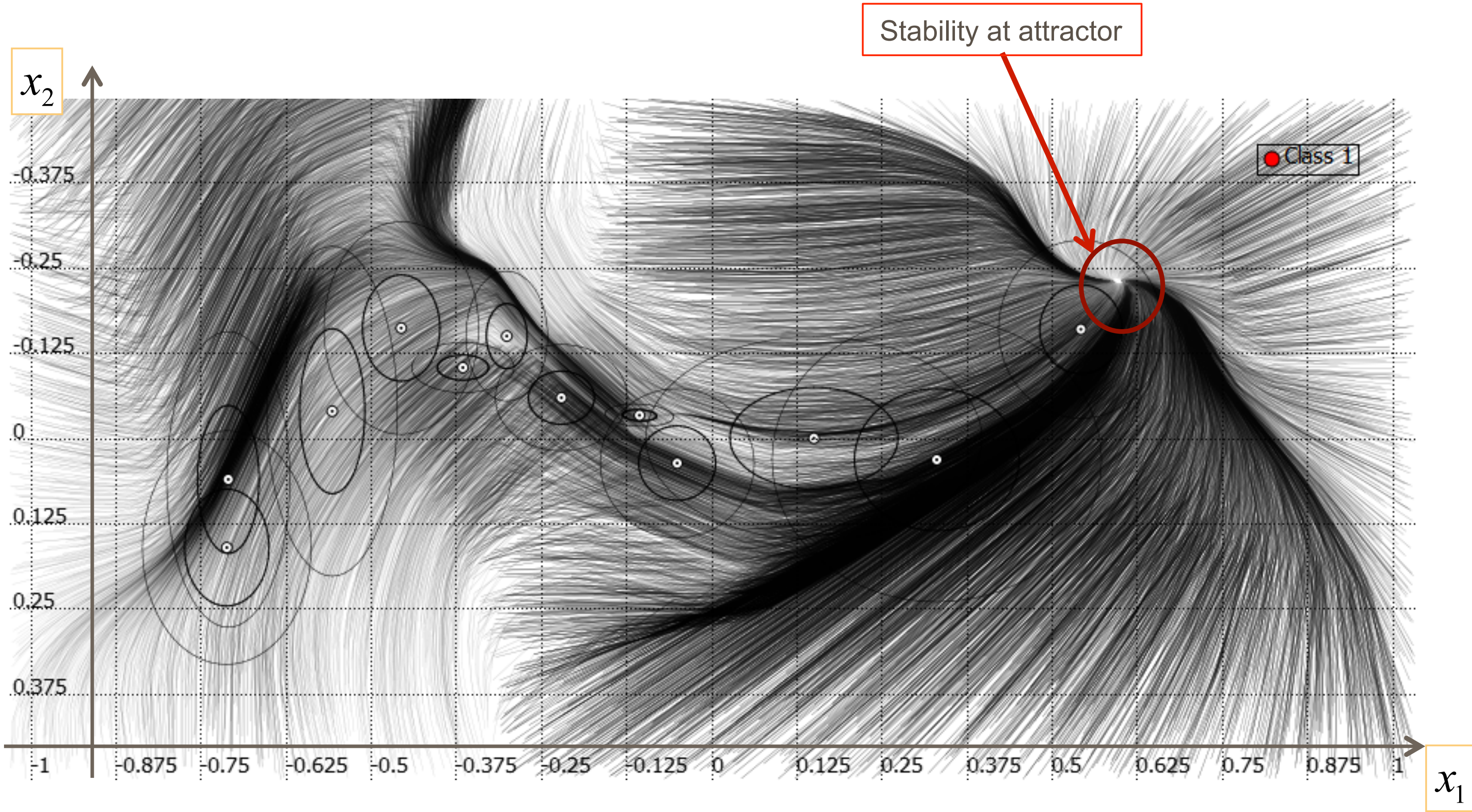
$$\begin{cases} a)\ \mu_{\dot{x}}^k + \Sigma_{\dot{x}x}^k \left(\Sigma_{xx}^k\right)^{-1} \mu_x^k = 0 \\ b)\ \Sigma_{\dot{x}x}^k \left(\Sigma_{xx}^k\right)^{-1} + \left(\Sigma_{\dot{x}x}^k \left(\Sigma_{xx}^k\right)^{-1}\right)^T \succ 0 \end{cases} \quad \forall k = 1, \ldots K$$

Determine the parameters of the density as a constrained optimization problem (maximize likelihood under stability constraints).

Khansari and Billard, **SEDS**, IEEE TRO 2011

Generalizing: Learning a control law

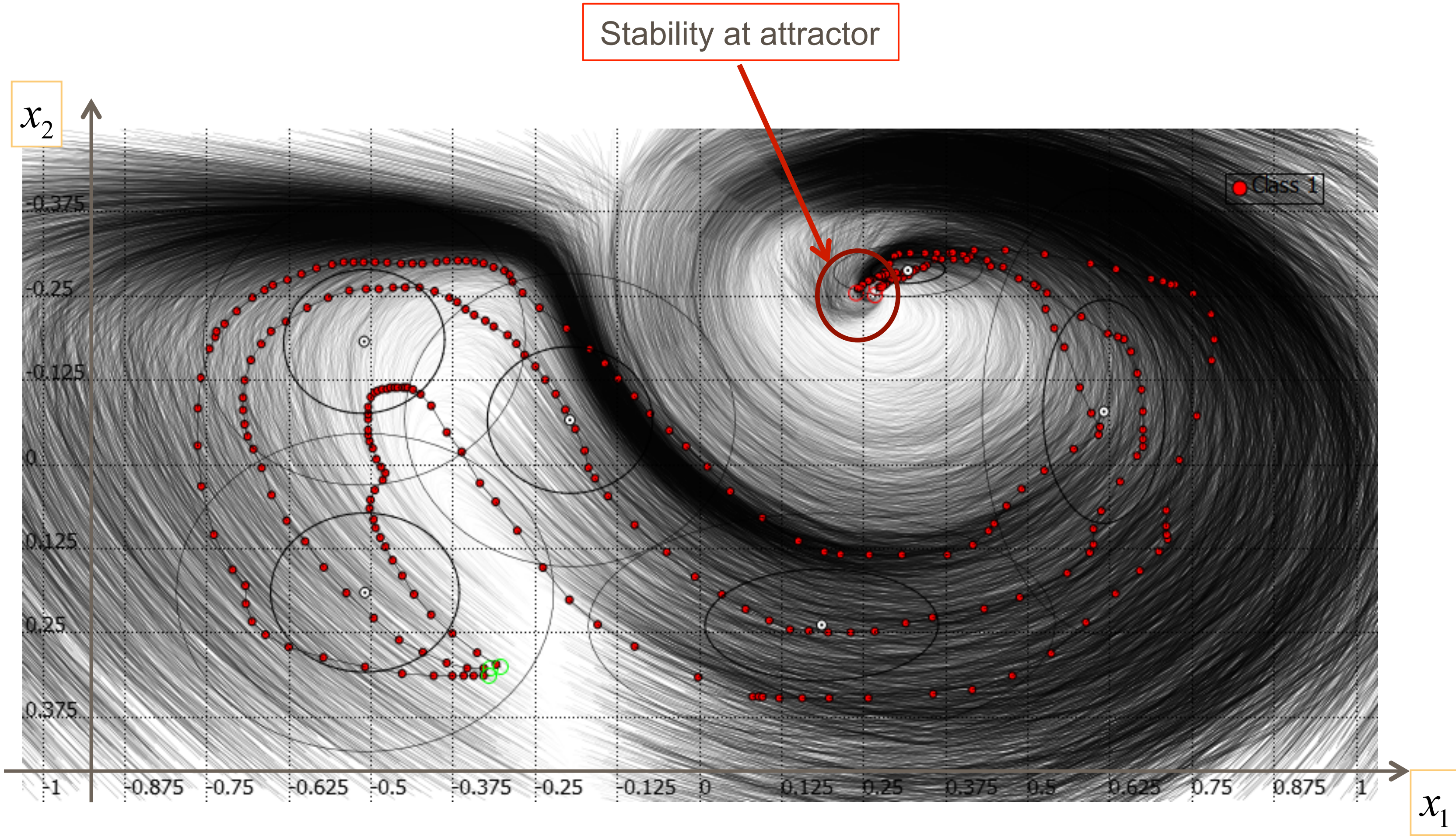
Other examples of complex dynamics that can be estimated through SEDS.



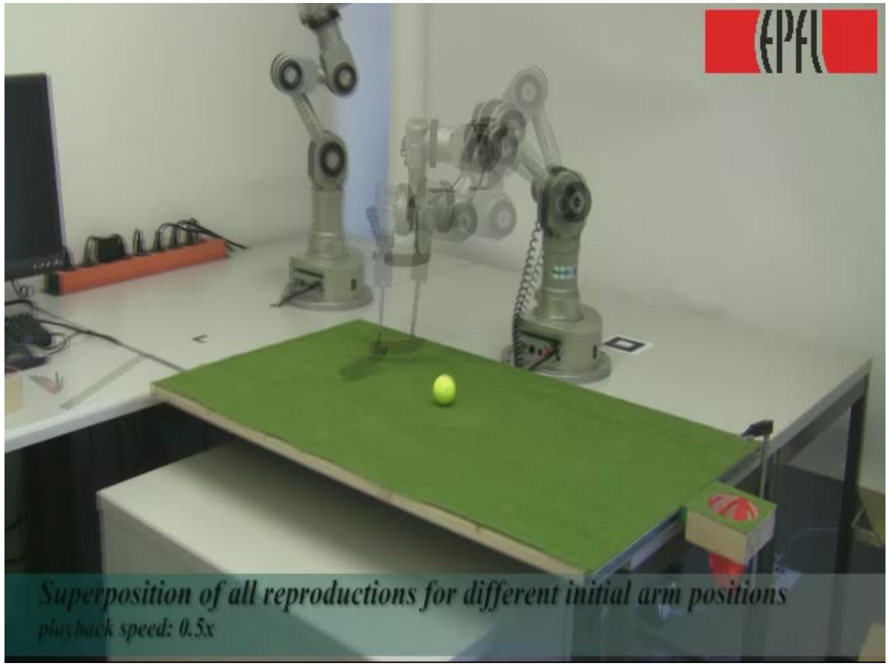
Khansari and Billard, **SEDS,** IEEE TRO 2011

Generalizing: Learning a control law

Other examples of complex dynamics that can be estimated through SEDS.



Learning motion with non-zero velocity at target

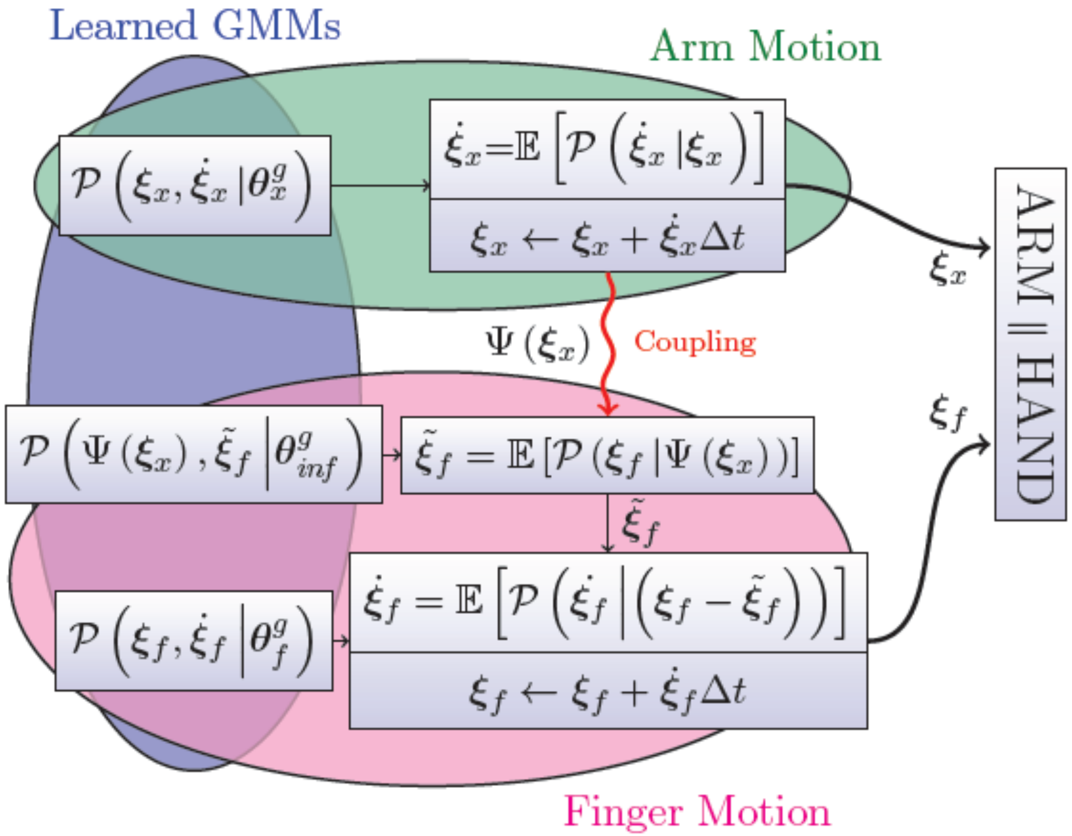


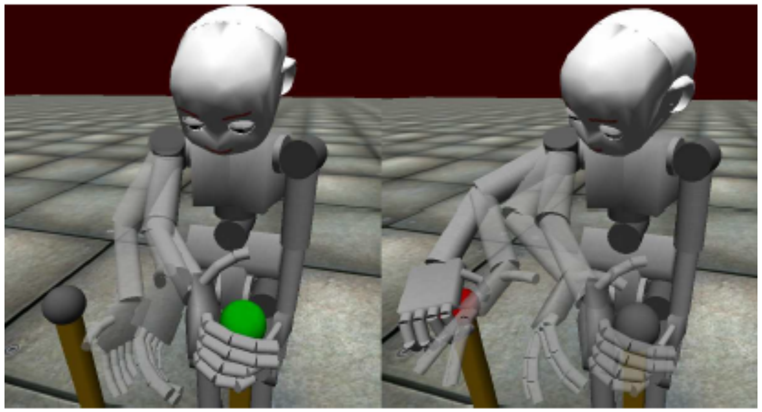
Extend the SEDS model with modulation in speed at target

Learning coupling across dynamical systems

Learn separately stable control laws to control for arm and fingers.

Couple the two systems to allow adequate adaptation to perturbations.



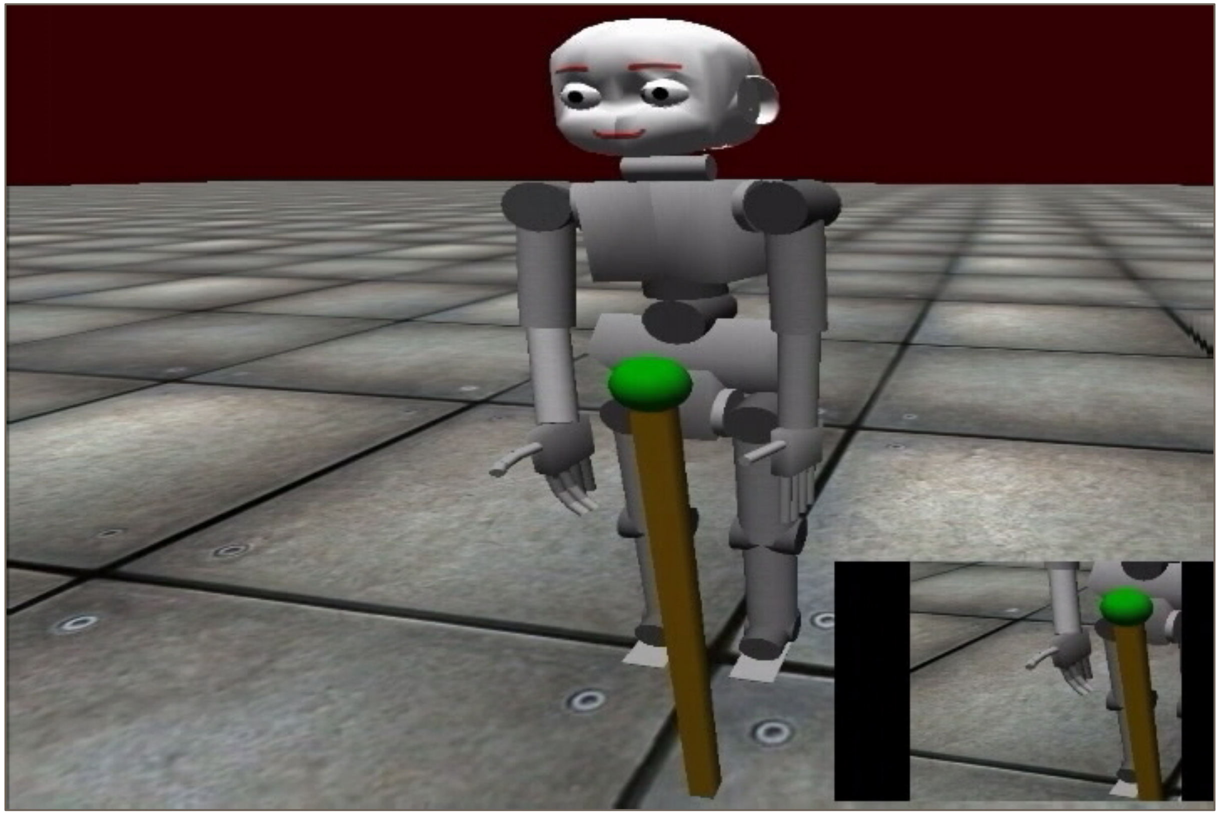


Shukla and Billard, RSS 2011

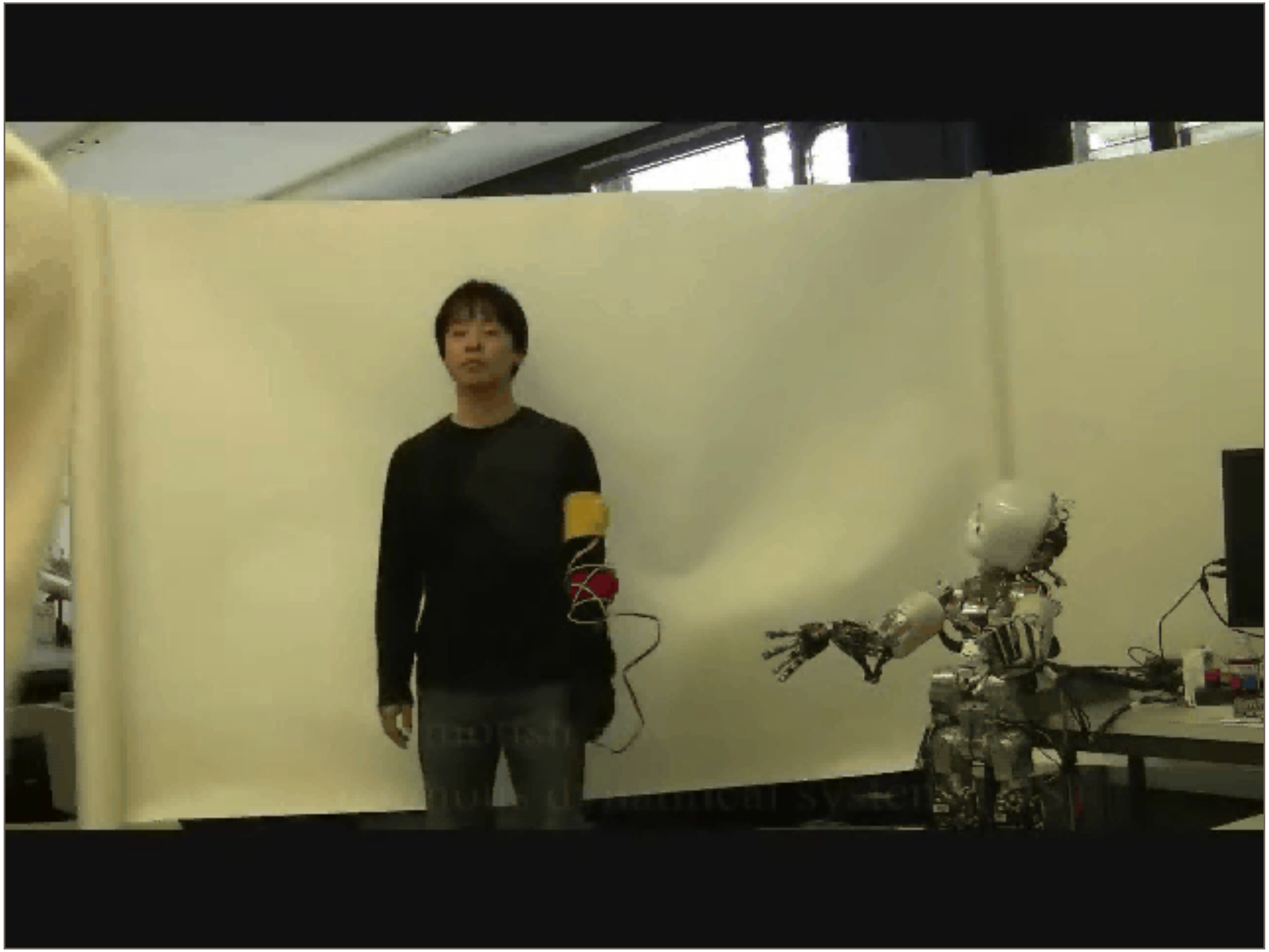
Learning coupling across dynamical systems

Learn separately stable control laws to control for arm and fingers.

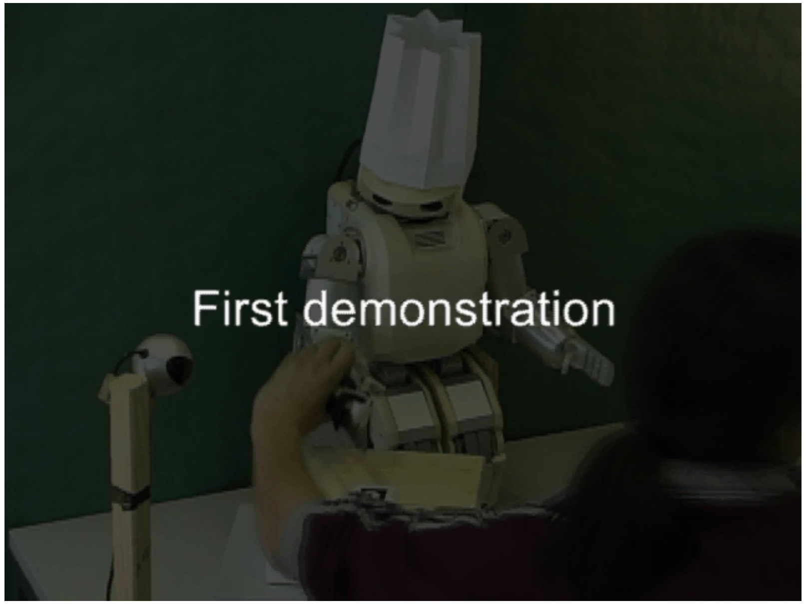
Couple the two systems to allow adequate adaptation to perturbations.

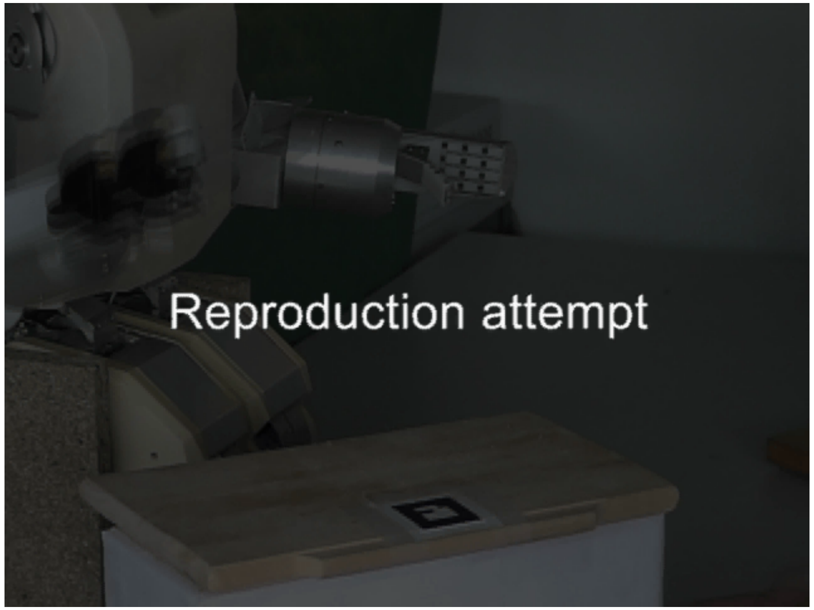


Catching Objects in Flight



What to Imitate?



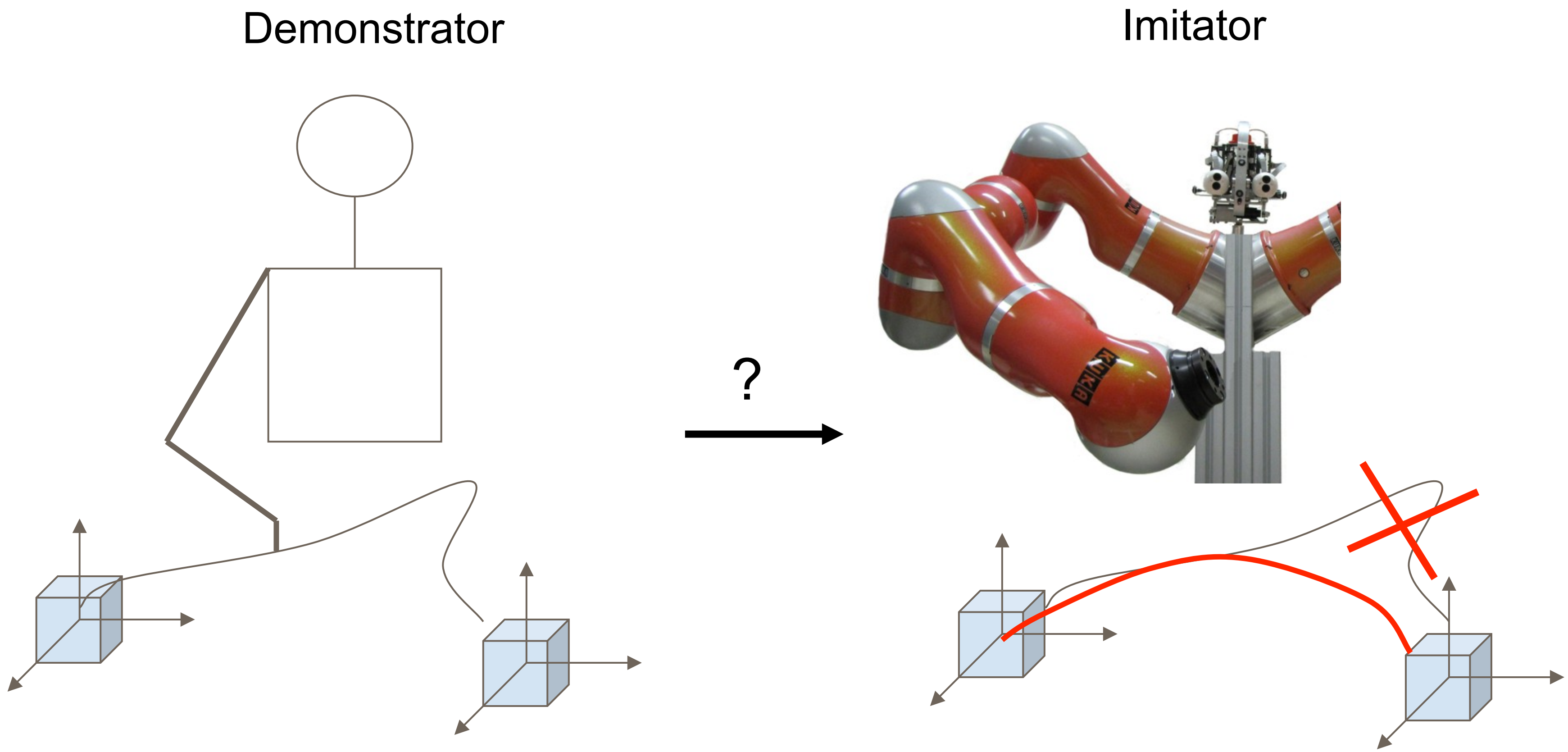


Learning a skill is more than simply replaying a trajectory.
It requires to understand what a skill is.

To learn this, one needs to show several demonstrations
to *generalize* across sets of examples.

Billard et al, Rob. and Aut. Systems, 2005; Calinon et al. IEEE SMC 2007

How to Imitate?

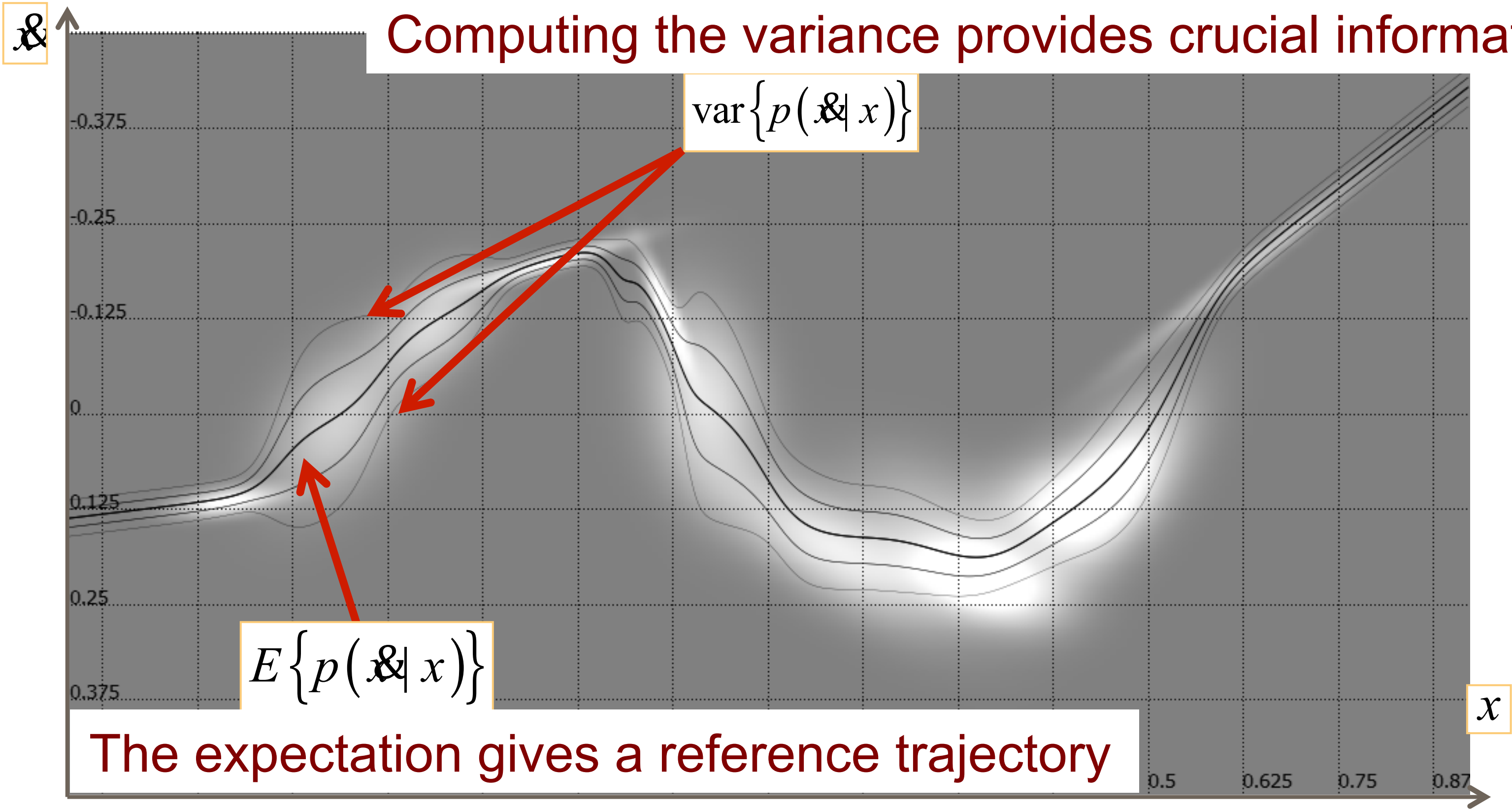


→ **Find the closest solution according to some cost function**

Statistical model of the data

Key Idea: The world is uncertain; learn about its uncertainty through probabilistic modeling of information.

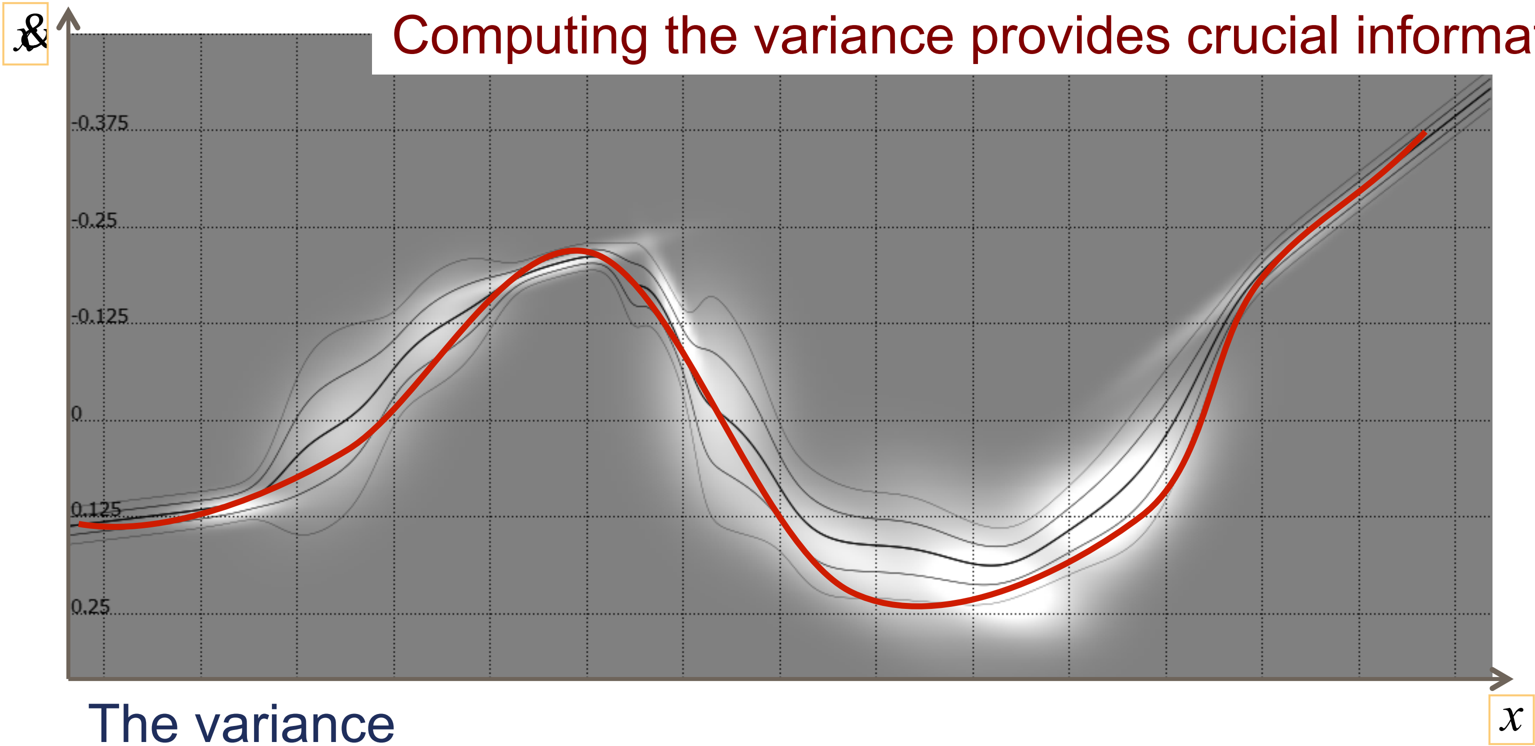
$$p\left(\dot{x},x\right)=\sum_{i=1}^{K} w_i N\left(\dot{x},x;\mu_i,\Sigma_i\right)$$



Generalizing

To generate new trajectories that depart from the *reference trajectory* while remaining within the total variance.

Computing the variance provides crucial information



The variance

→ provides a notion of feasible space of solutions

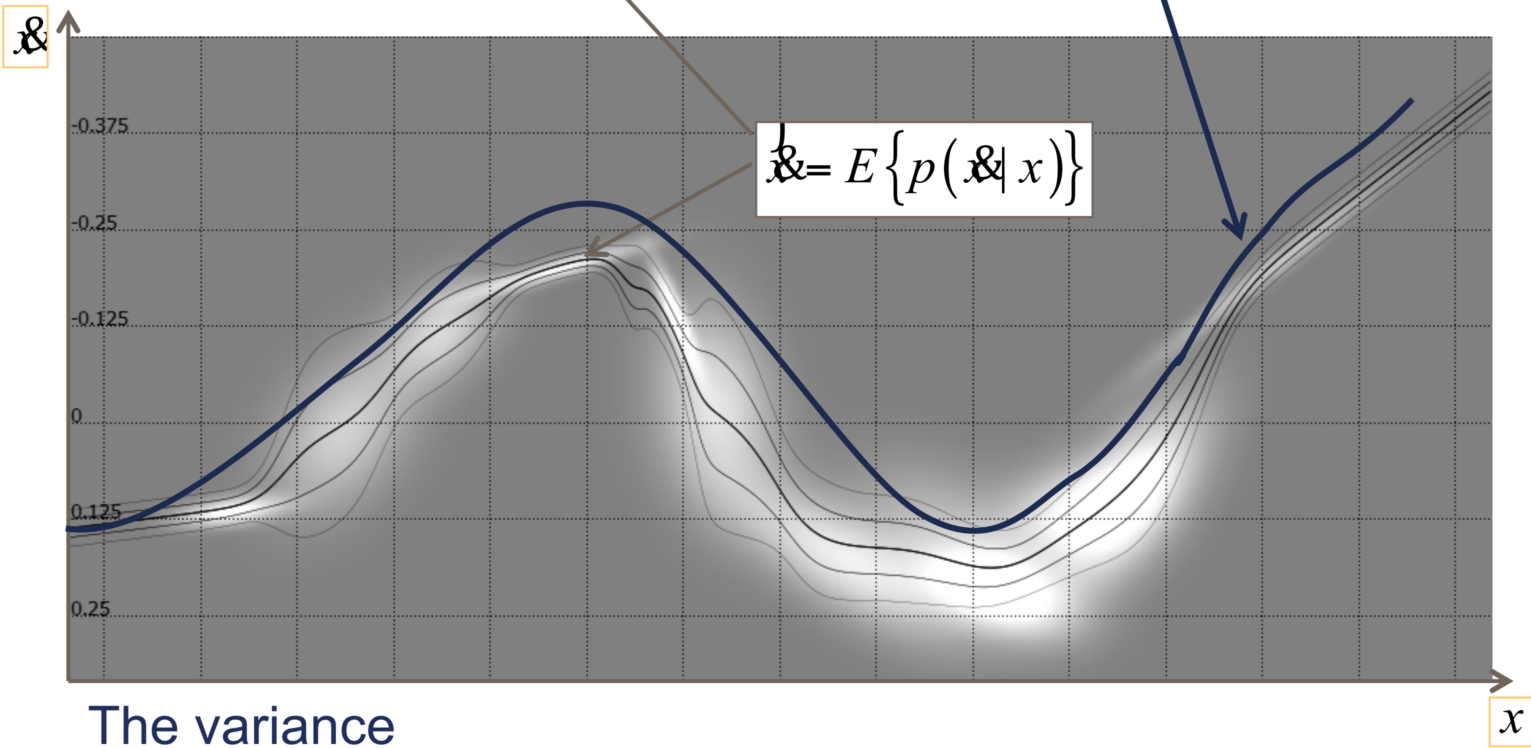
→ is used to compute new path in the face of changes in the context

Generalizing

Cost function $H(\dot{x}) = \left\| \dot{x} - \hat{\dot{x}} \right\|$

$$\min H(\dot{x}) = \left\| \dot{x} - \hat{\dot{x}} \right\|$$

$$\text{u.c.} \quad J\dot{\theta} = \dot{x} \quad \text{(inverse kinematics)}$$



The variance

→ provides a notion of feasible space of solutions

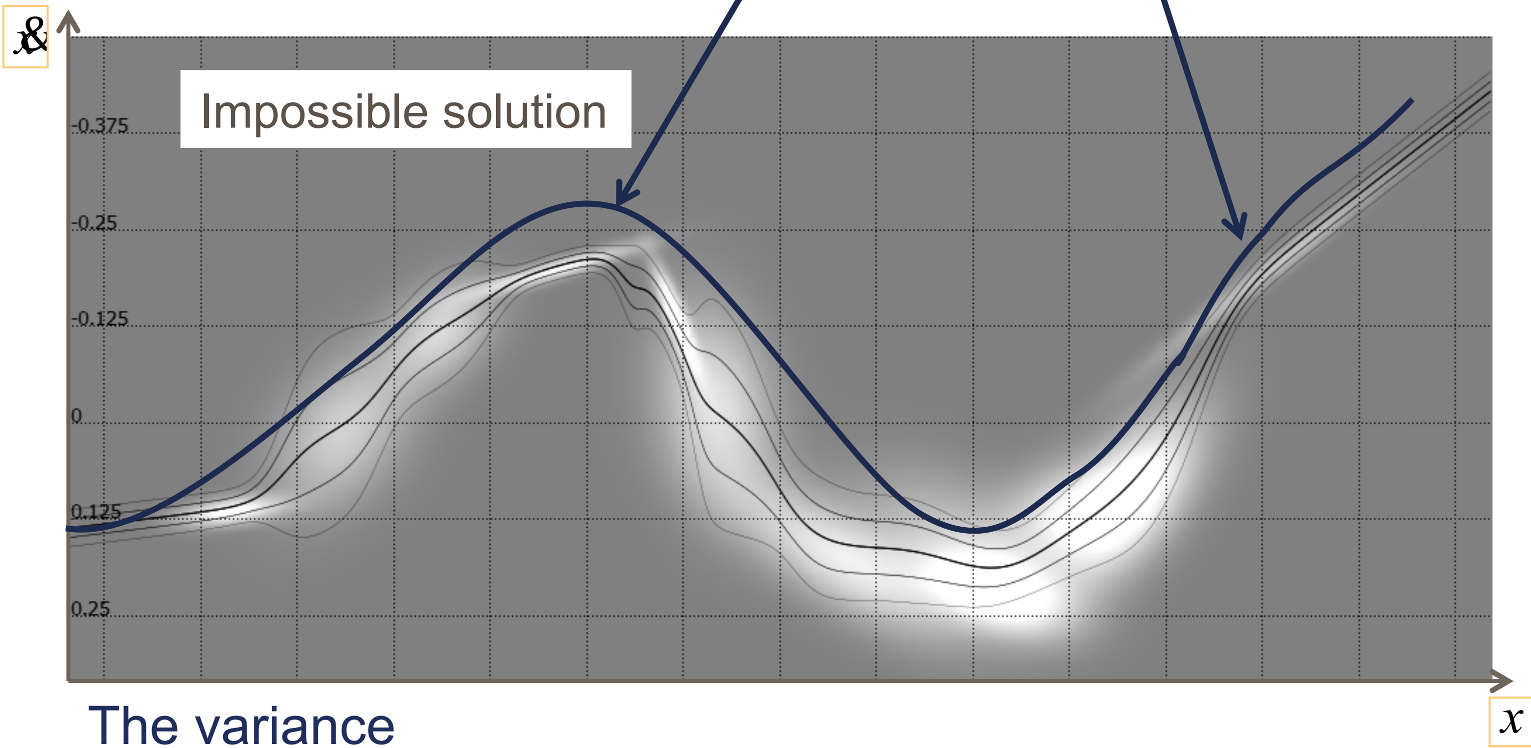
→ is used to compute new path in the face of changes in the context

Generalizing

Cost function $H(\dot{x}) = \left\| \dot{x} - \hat{\dot{x}} \right\|$

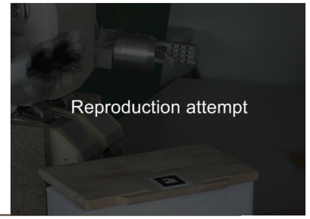
$$\min H(\dot{x}) = \left\| \dot{x} - \hat{\dot{x}} \right\|$$

$$\text{u.c.} \quad J\dot{\theta} = \dot{x} \ \text{(inverse kinematics)}$$



The variance
→ provides a notion of feasible space of solutions
→ is used to compute new path in the face of changes in the context

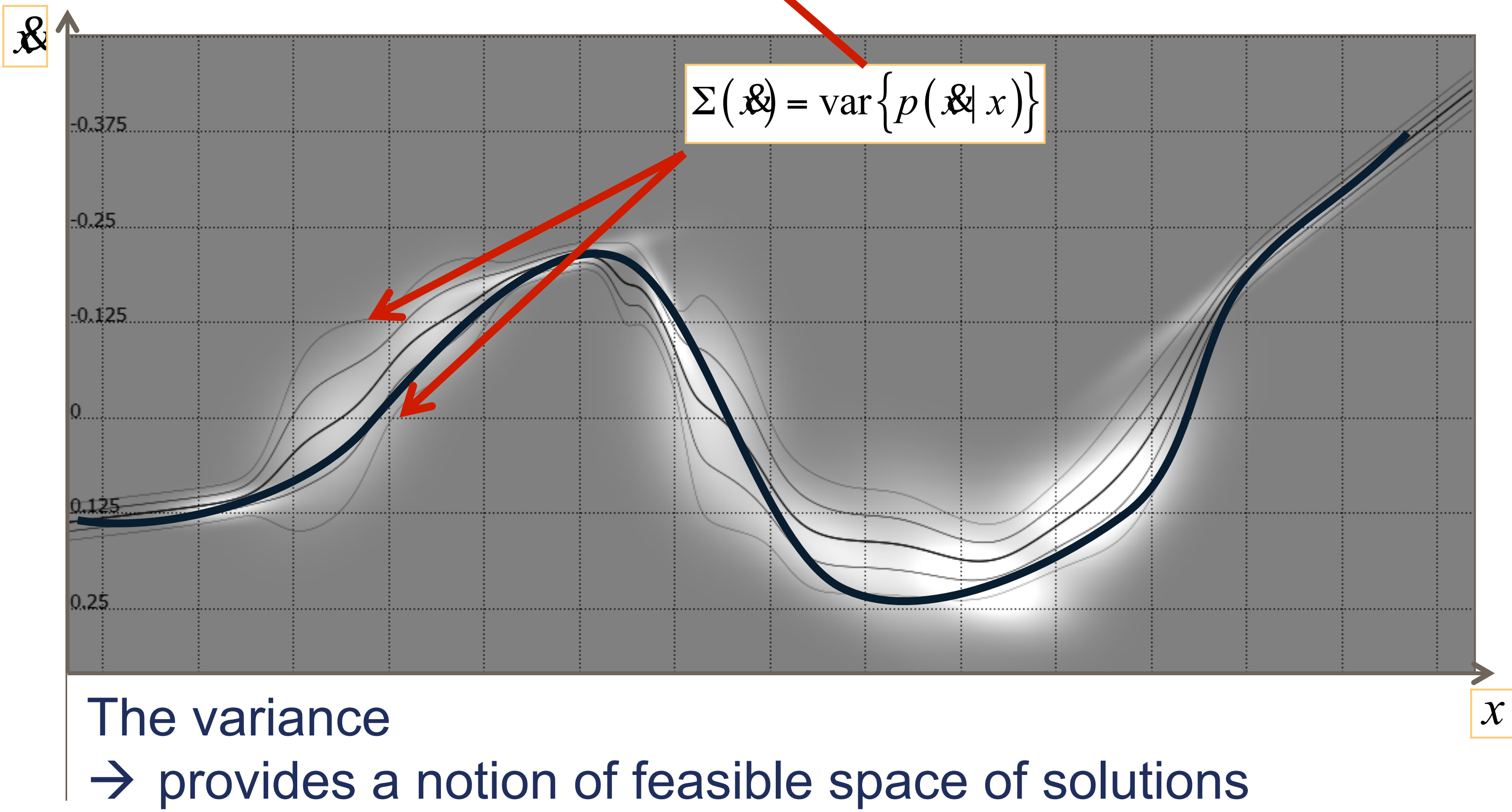
Generalizing



Cost function $H(\dot{x}) = \left(\dot{x} - \hat{\dot{x}}\right)^T \Sigma(\dot{x})^{-1} \left(\dot{x} - \hat{\dot{x}}\right)$

$\min H(\dot{x})$

u.c. $J\dot{\theta} = \dot{x}$ (inverse kinematics)



The variance

→ provides a notion of feasible space of solutions

→ is used to compute new path in the face of changes in the context

Adaptive Grasping

Grasping usually solved by searching for the optimal placement of fingers onto an object.

Adaptive Grasping

Grasping usually solved by searching for the optimal placement of fingers onto an object.

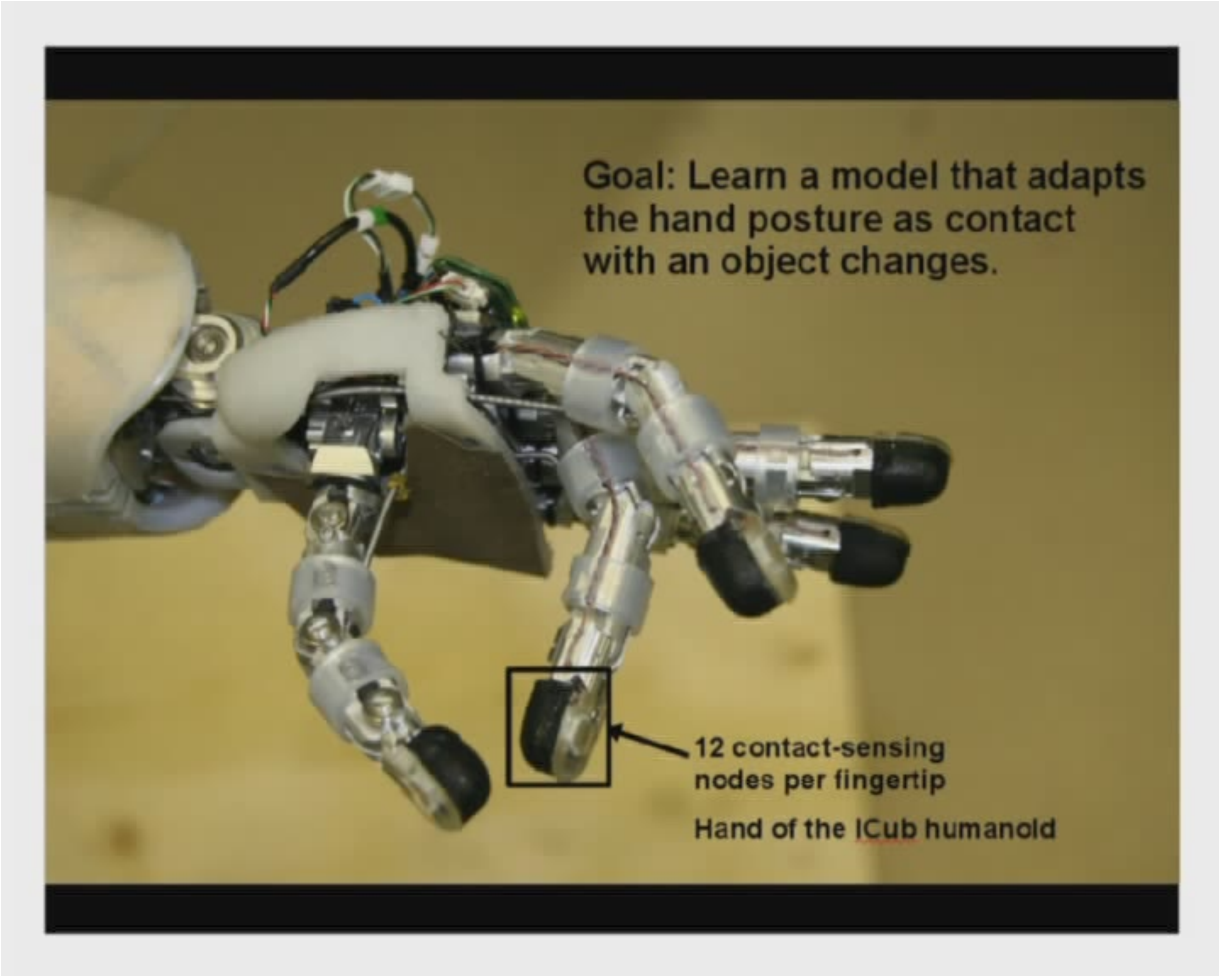
Knowing the extent to which one can adapt this grasp is useful for safe manipulation.

Learn how comply with external perturbations while maintaining a firm grasp.

Sauser, Argall and Billard, Autonomous Robots, 2012

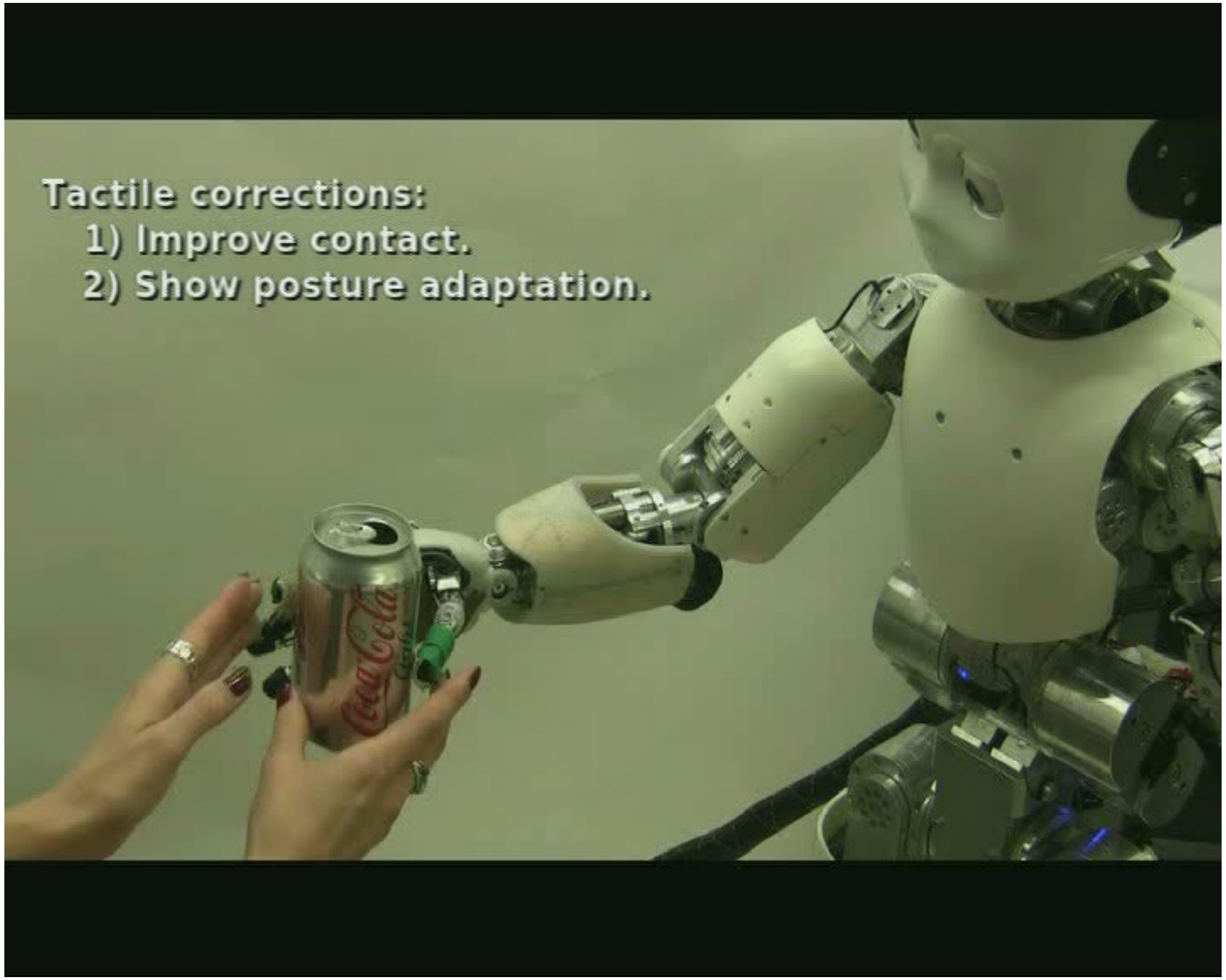
Adaptive Grasping

Teaching through tactile sensing



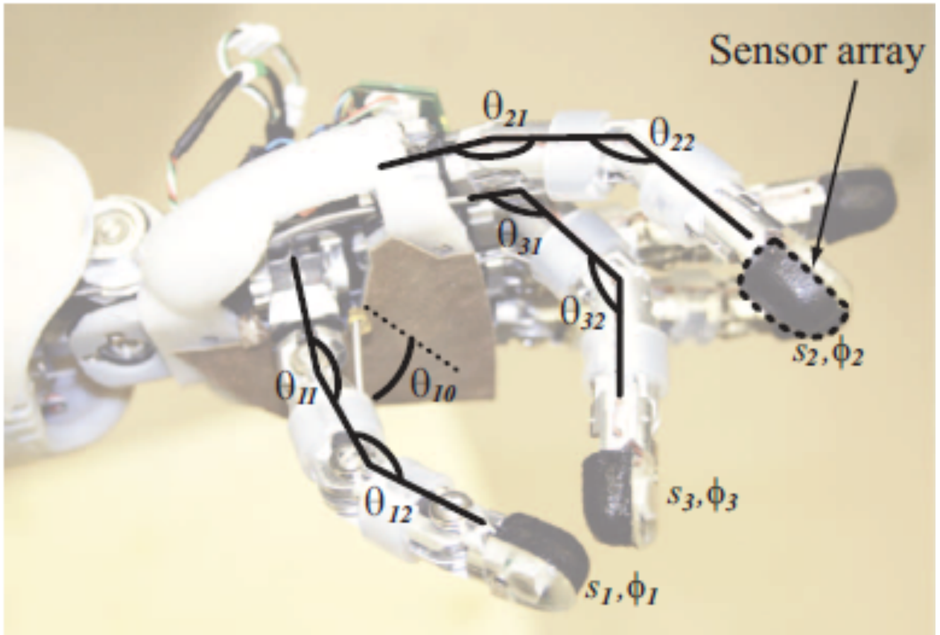
Adaptive Grasping

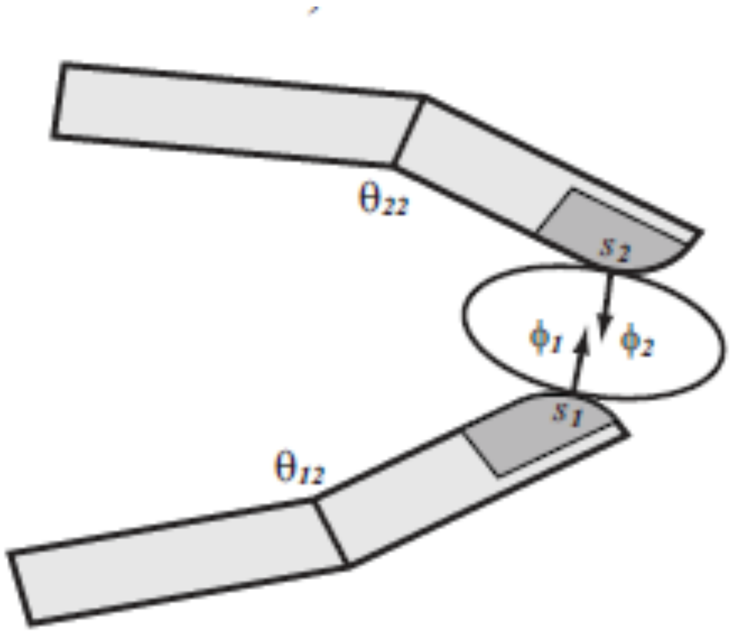
Teaching through tactile sensing



Adaptive Grasping

Learn a probabilistic mapping $p(\phi, s, \theta)$ between contact signature of the object (normal force ϕ and tactile response s) and fingers' posture θ.





Adaptive Grasping

Learn a probabilistic mapping $p(\phi, s, \theta)$ between contact signature of the object (normal force ϕ and tactile response s) and fingers' posture θ.

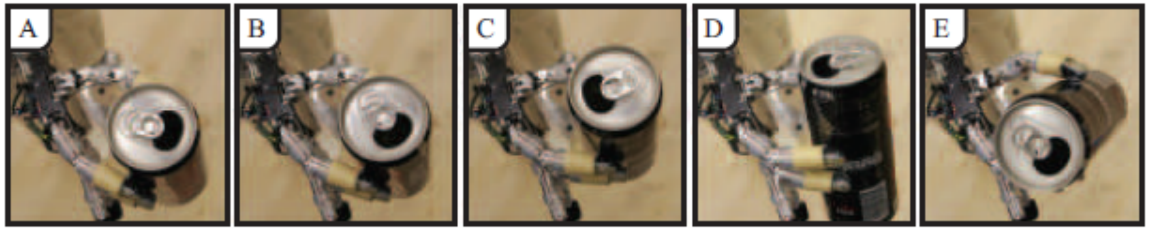
Make N observations of the state of the system $\xi^i = \{\phi^i, s^i, \theta^i\} \in \mathbb{R}^{47}$, $i = 1 \ldots N$.

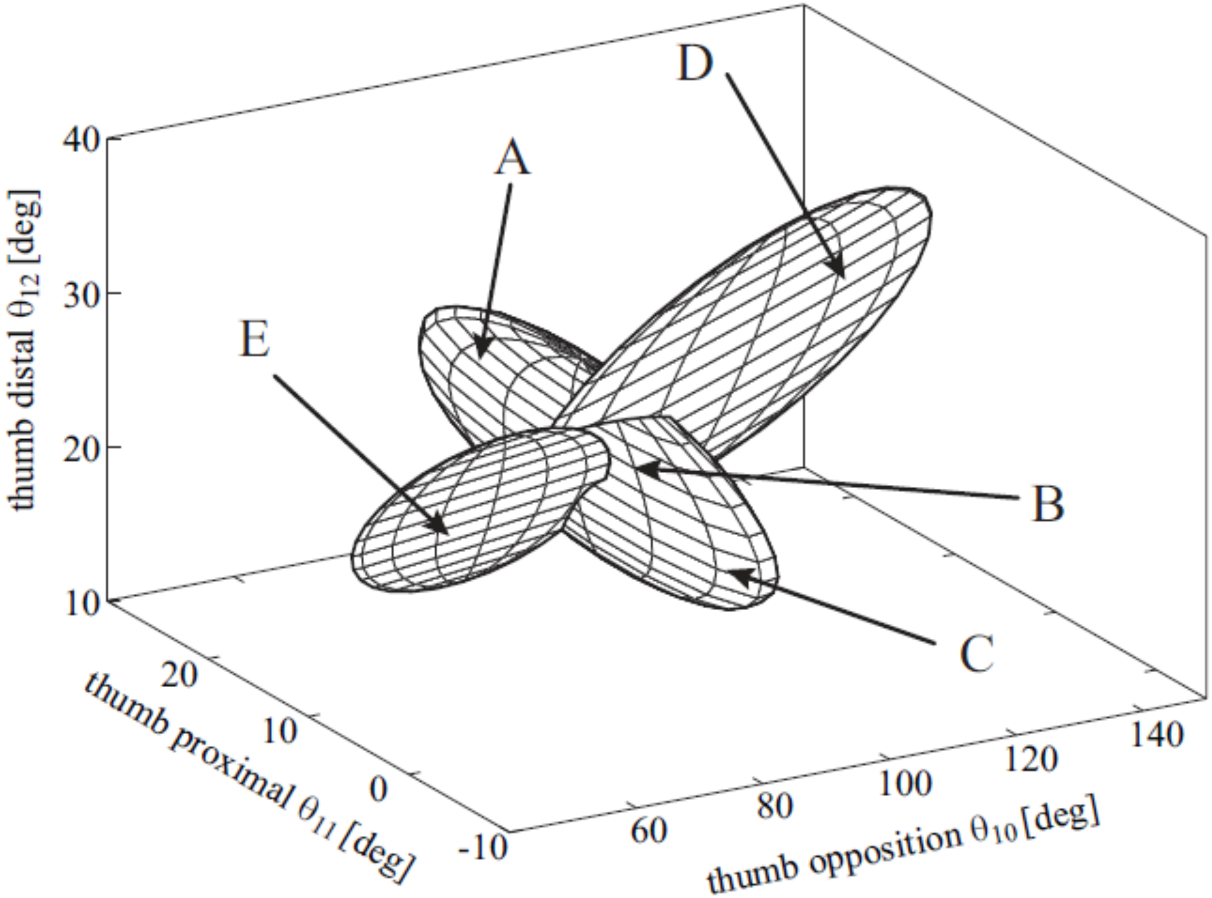
$p(\xi) = \sum_{k=1}^{K} w_k N(\xi; \mu^k, \Sigma^k)$: joint density (mixture of Gaussians) describing the observations and the correlation across the variables of the system.

Can be used to predict the appropriate joint posture when perceiving a change in contact signature:

$$\hat{\theta} = E\{p(\theta \mid s, \phi)\}, \quad \hat{s} = E\{p(s \mid \theta, \phi)\}$$

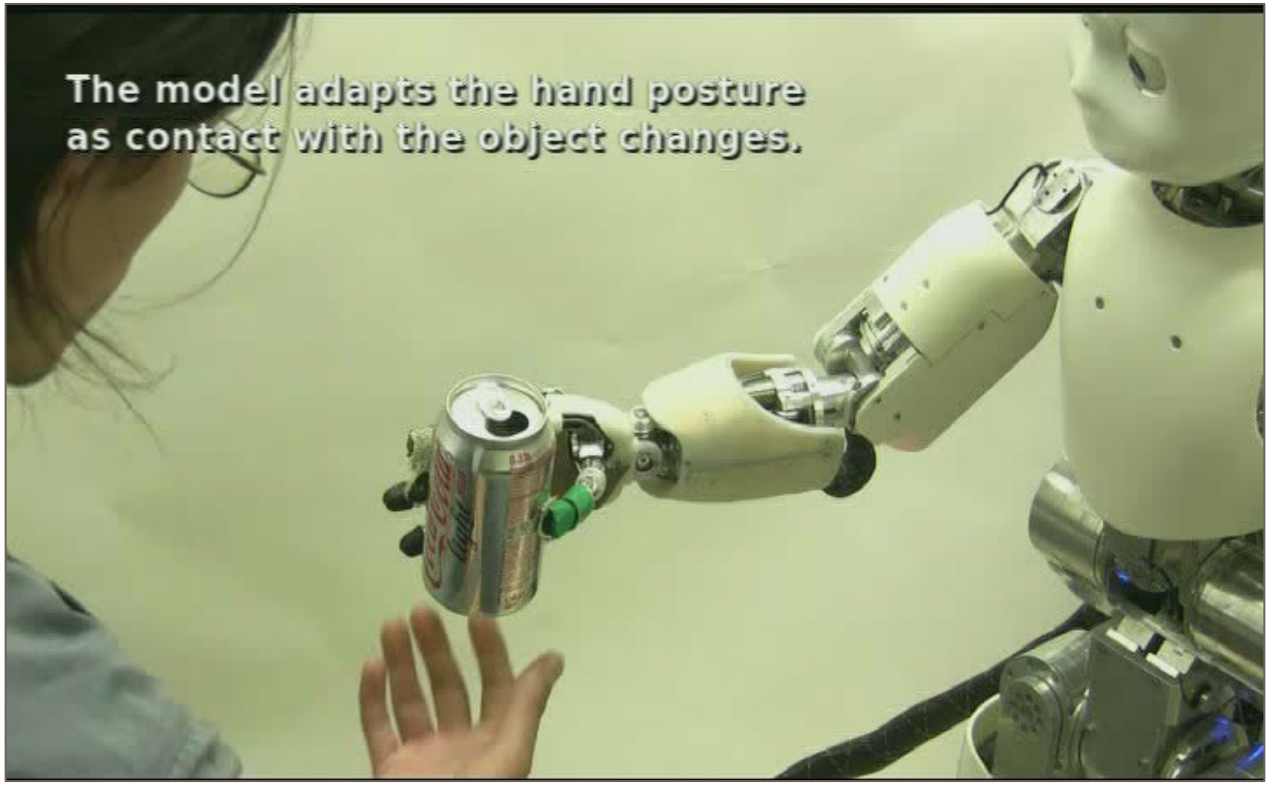
Adaptive Grasping





Adaptive Grasping

After Training

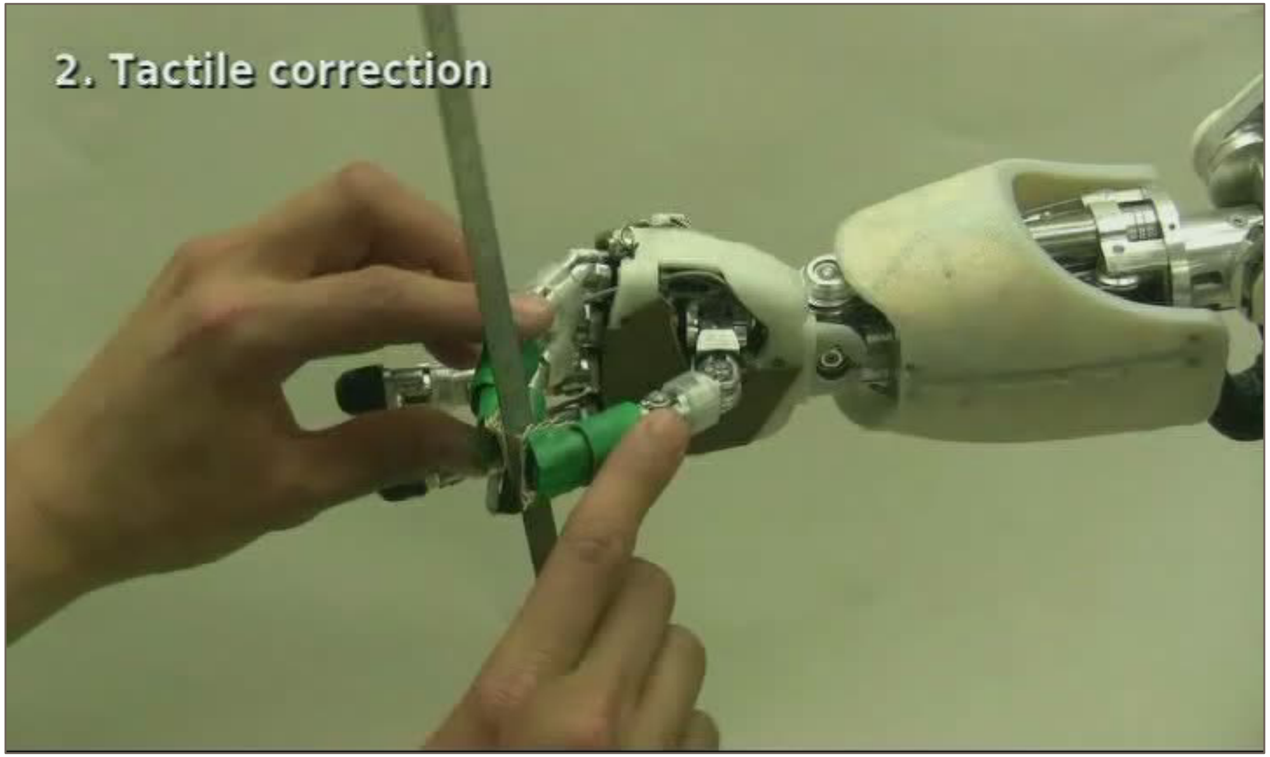


Adaptive Grasping

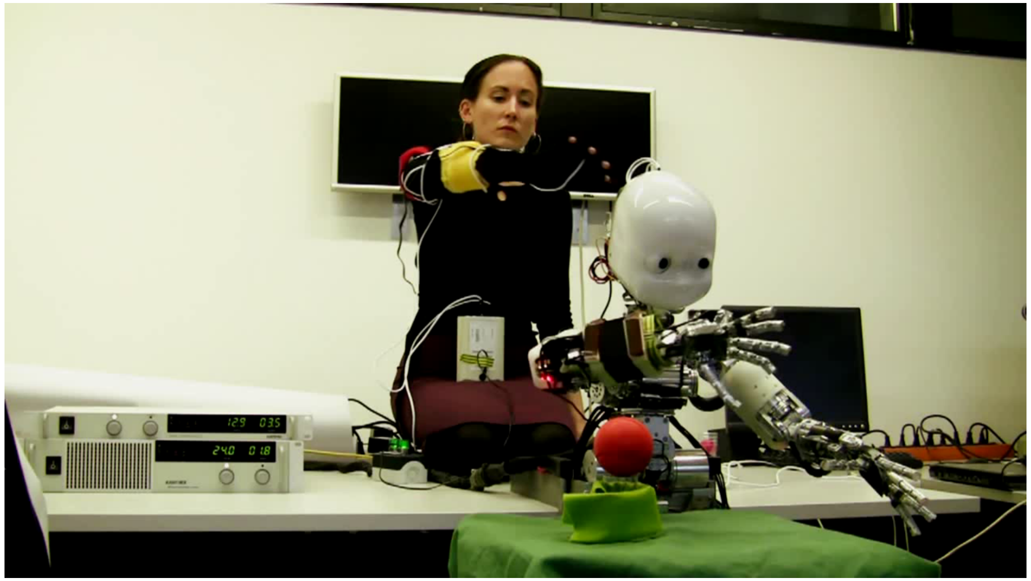
Another Example

Adaptive Grasping

Another Example

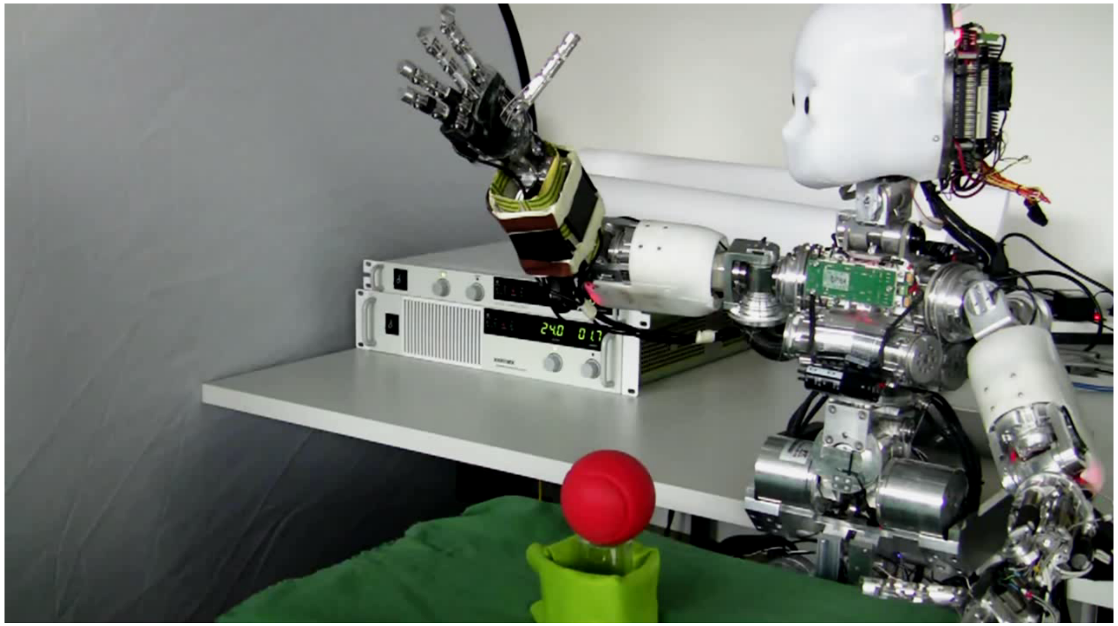


Adaptive Manipulation



<u>Teaching through teleoperation</u>

using Interface for direct joint motion transfer (Xsens motion sensors)

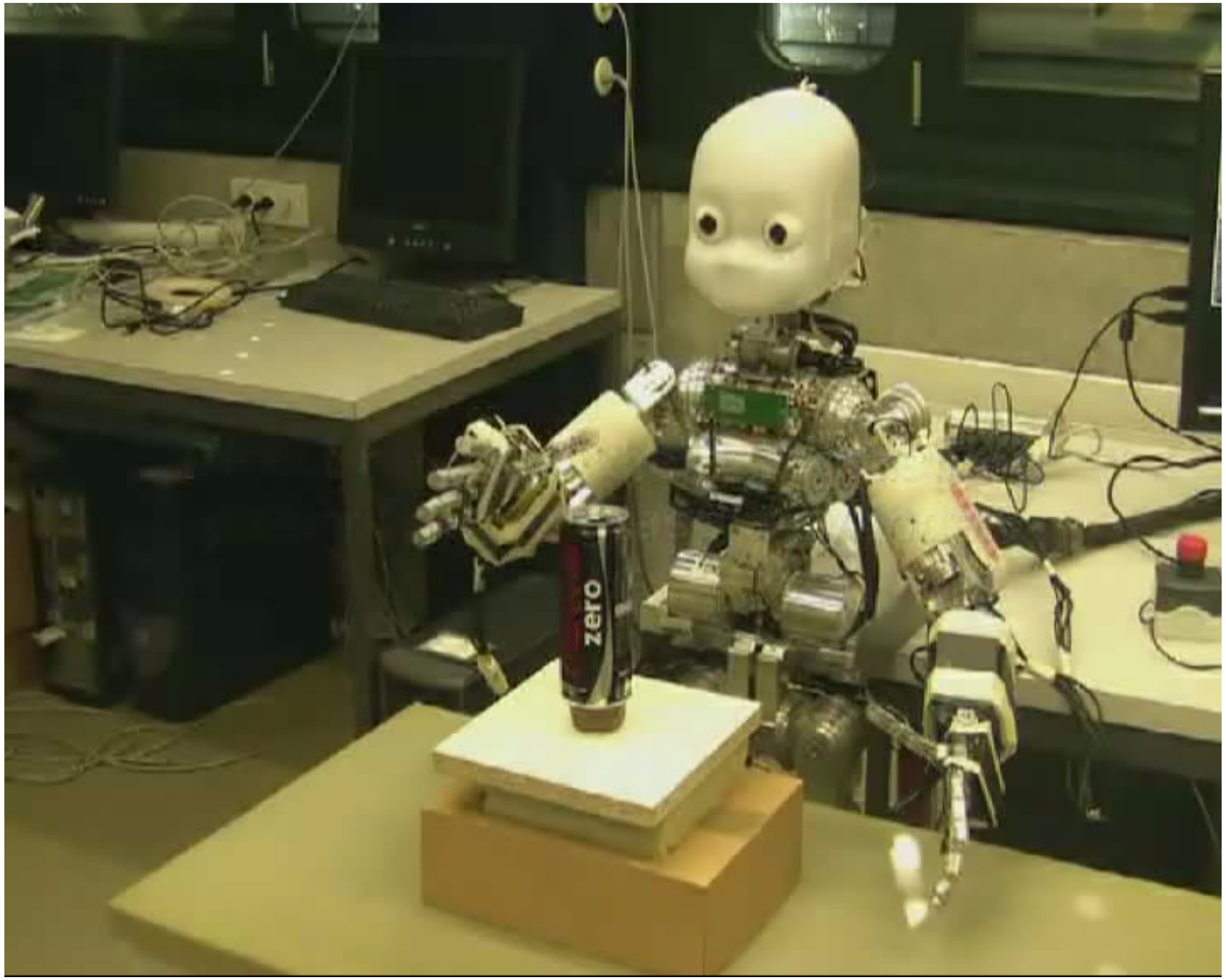


<u>Refining knowledge using tactile interface</u>

(5 touchpads mounted on robot's arm and wrist)

Adaptive Manipulation

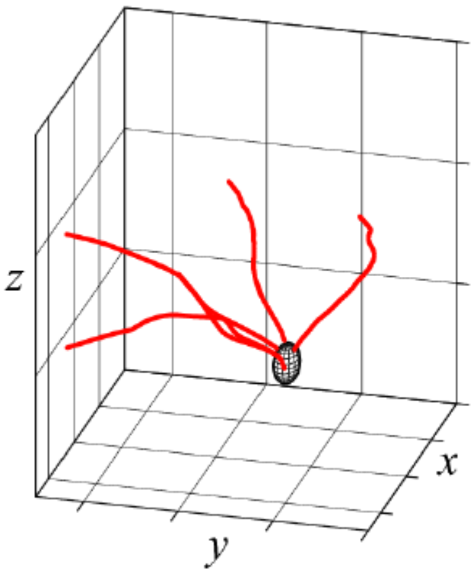
Reuse: To avoid re-learning a new task from scratch when the new task bears similarities with the old task



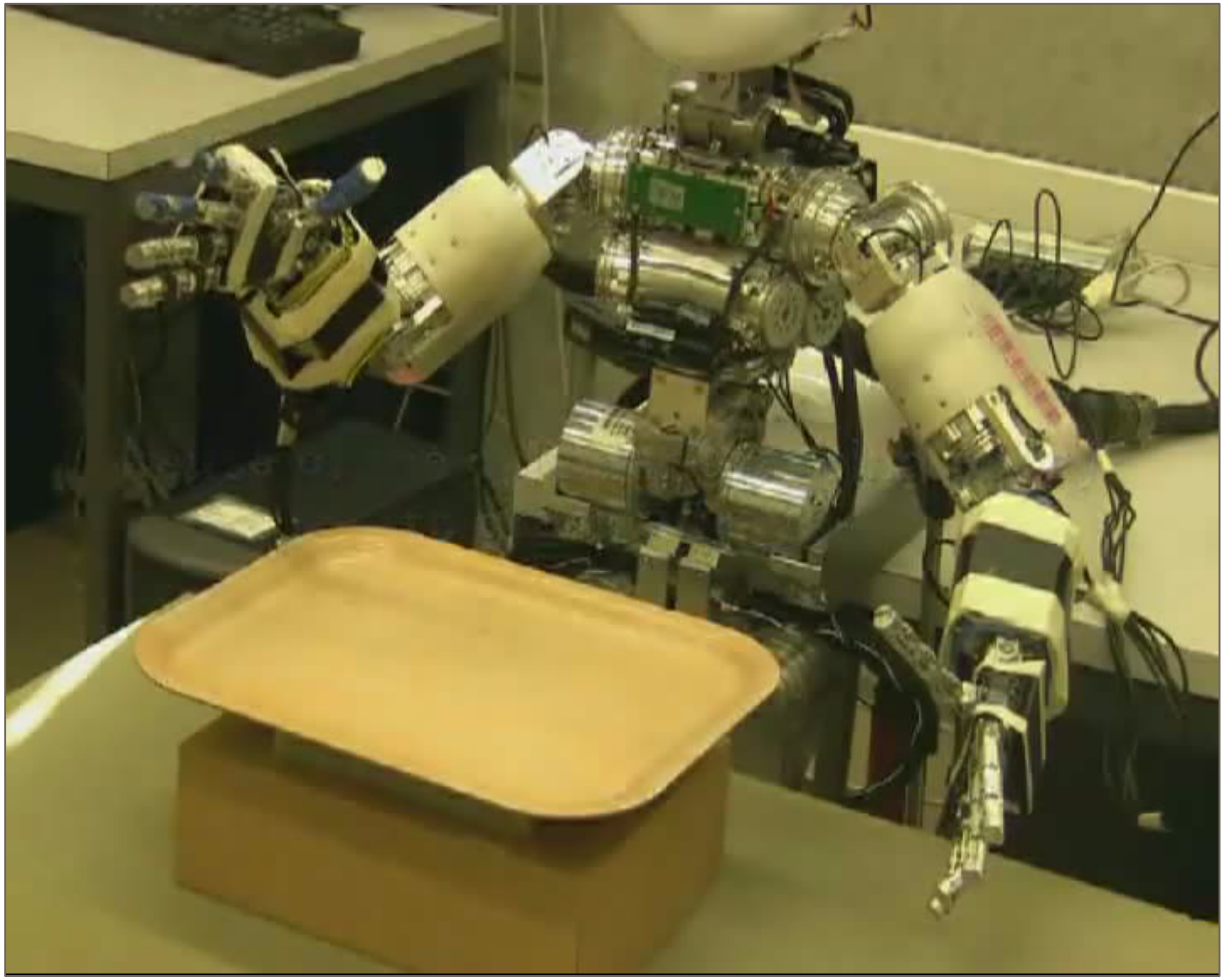
Adaptive Manipulation

Reuse preserves variability learned in the previous task. This may be a drawback → Use tactile feedback to adapt locally this variability

Before Reuse

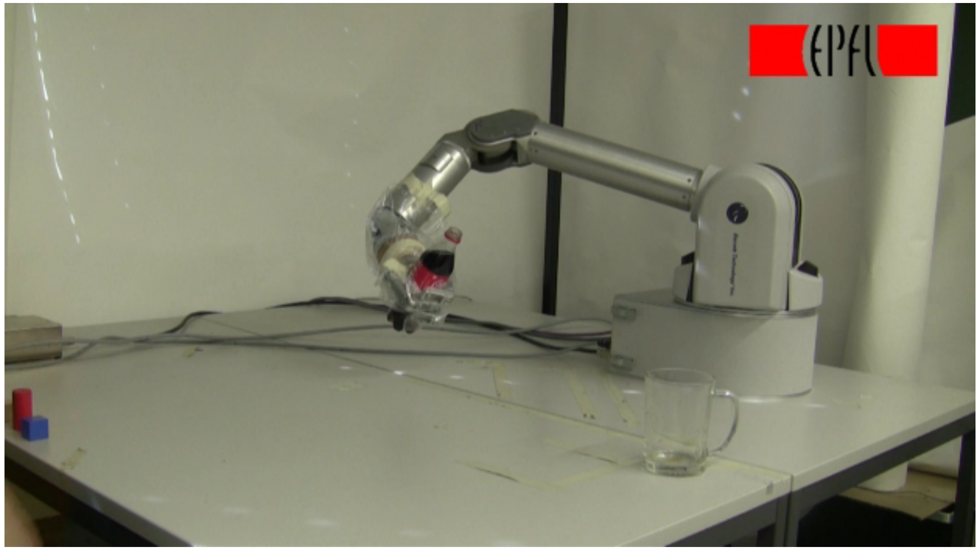


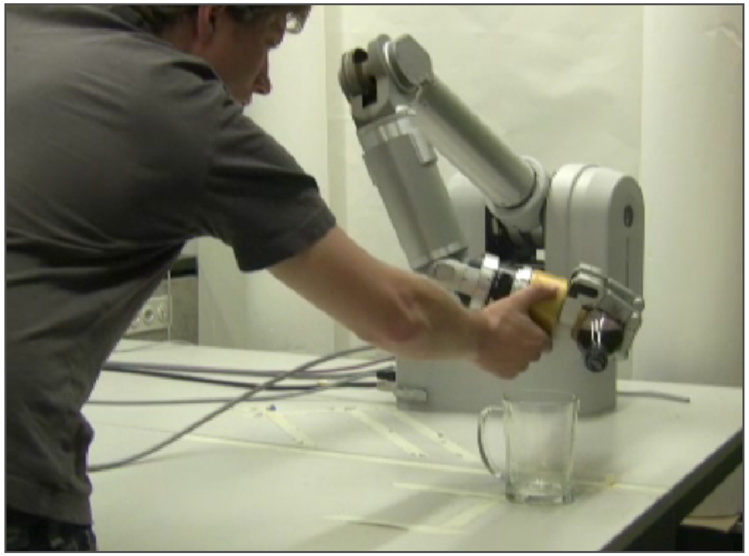
Reuse: One more example



Teaching robots to be less stiff

Being stiff is not always good → How to teach a robot to relax…





Low stiffness when carrying the liquid

High stiffness when pouring the liquid

Kronander and Billard, ICRA 2012

Teaching robots to be less stiff

Being stiff is not always good → How to teach a robot to relax…

Shaking the robot: A natural method to teach a robot to relax.

Teaching robots to be less stiff

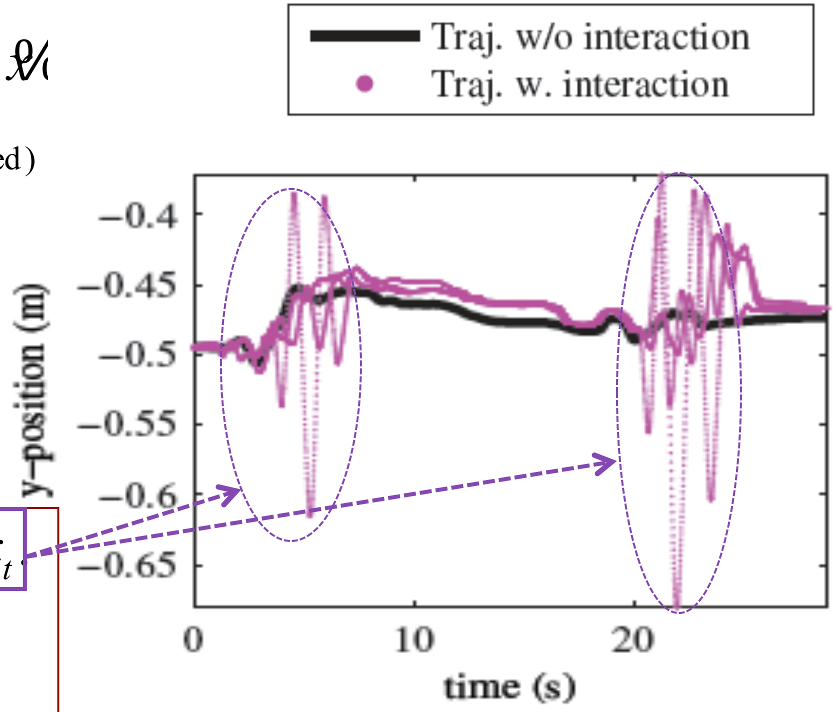
PD control law to follow a desired trajectory $\tilde{x}_t$

$$u_t = K\left(x_t - \tilde{x}_t\right) - D\left(x_t - \tilde{x}_t\right), \quad D \sim K \text{(critically damped)}$$

Adjust stiffness at each time step:

$K_t\left(x_t - \tilde{x}_t\right)$

Record perturbation from current position Δx_t



Teaching robots to be less stiff

PD control law to follow a desired trajectory $\hat{x}$

$$u_t = K\left(x_t - \hat{x}_t\right) - D\left(x_t - \hat{x}_t\right), \quad D \sim K \text{(critically damped)}$$

Adjust stiffness at each time step:

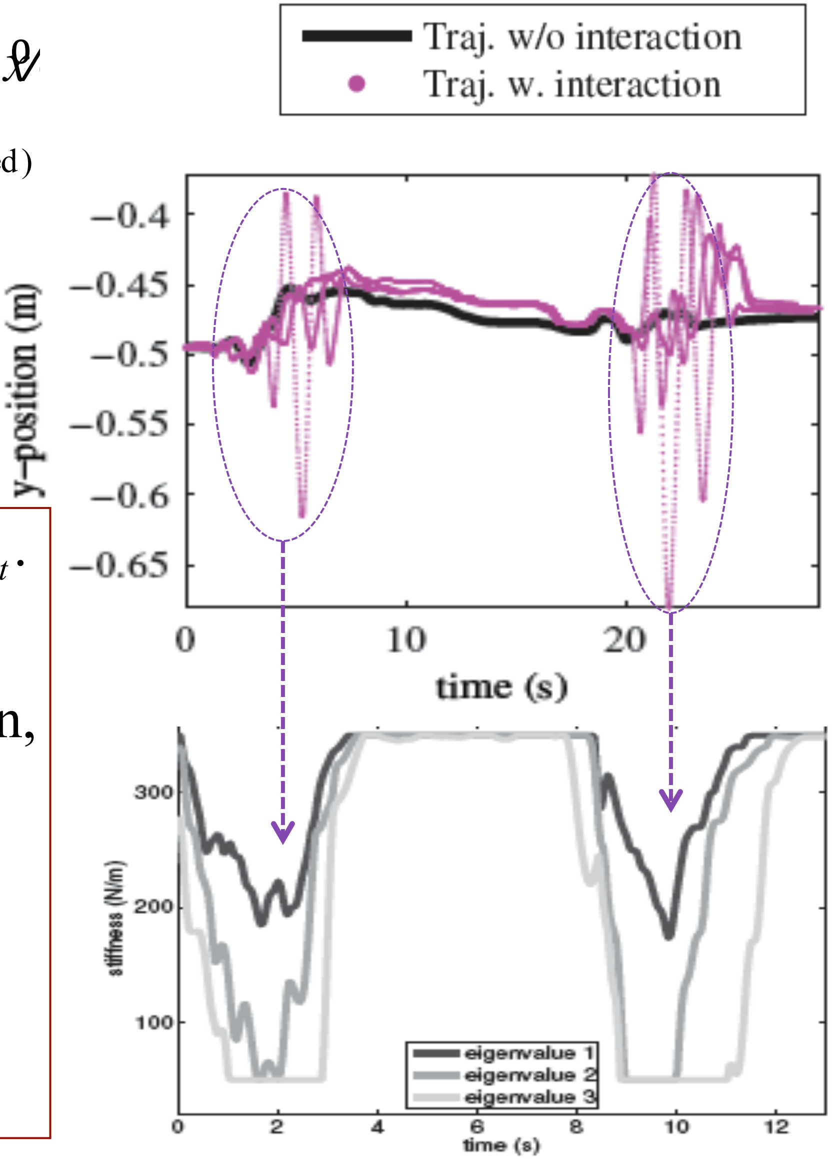
$K_t\left(x_t - \hat{x}_t\right)$

Record perturbation from current position Δx_t.
Set stiffness profile inversely proportional to variance of perturbation (the more variation, the less stiff):

Covariance matrix: $\Sigma = \Delta x\left(\Delta x\right)^T$

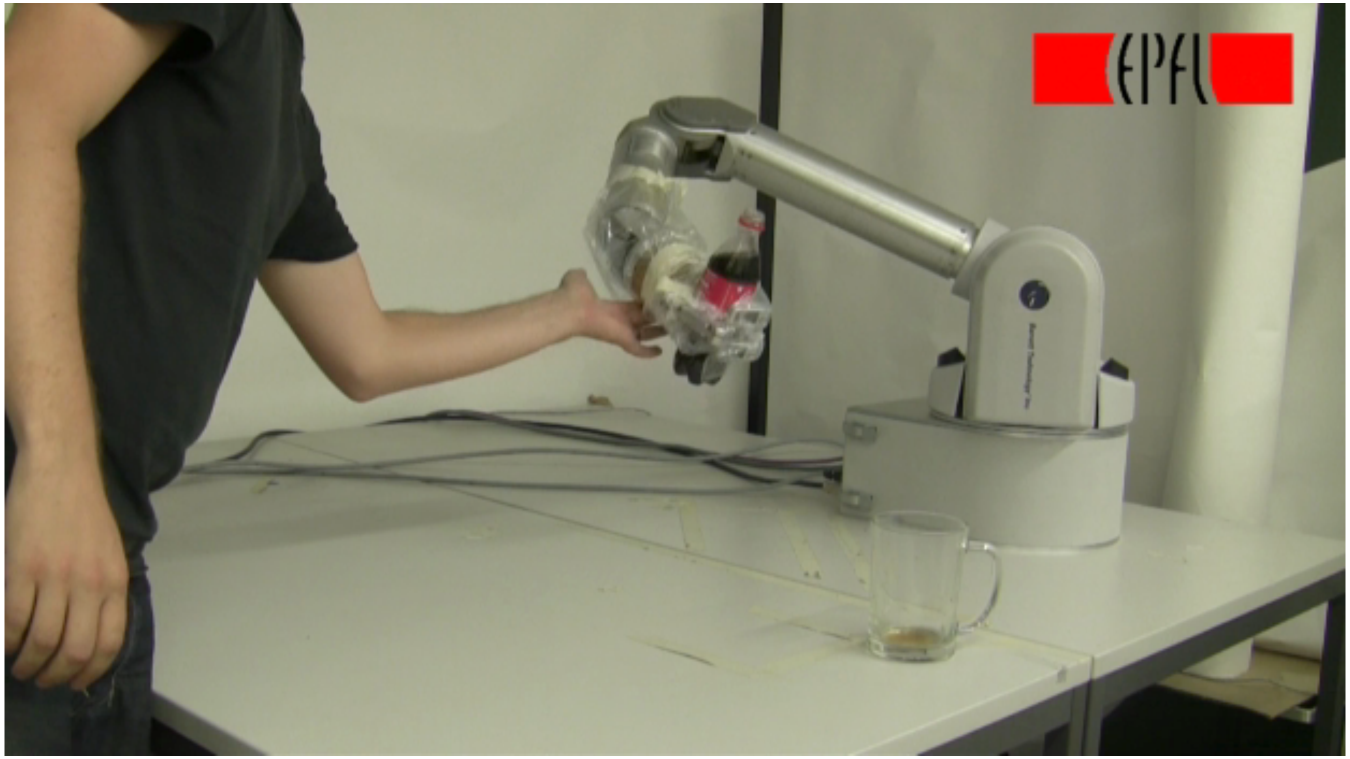
Eigenvalue decomposition: $\Sigma = U \Lambda U^T$

$\Rightarrow K_t \sim U \Lambda^{-1} U^T$

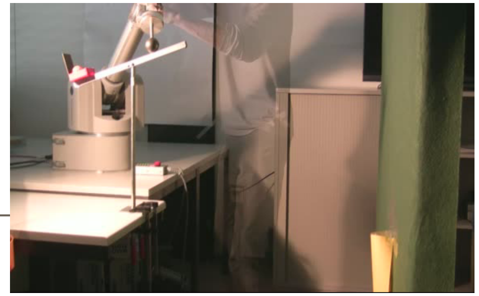


Teaching robots to be less stiff

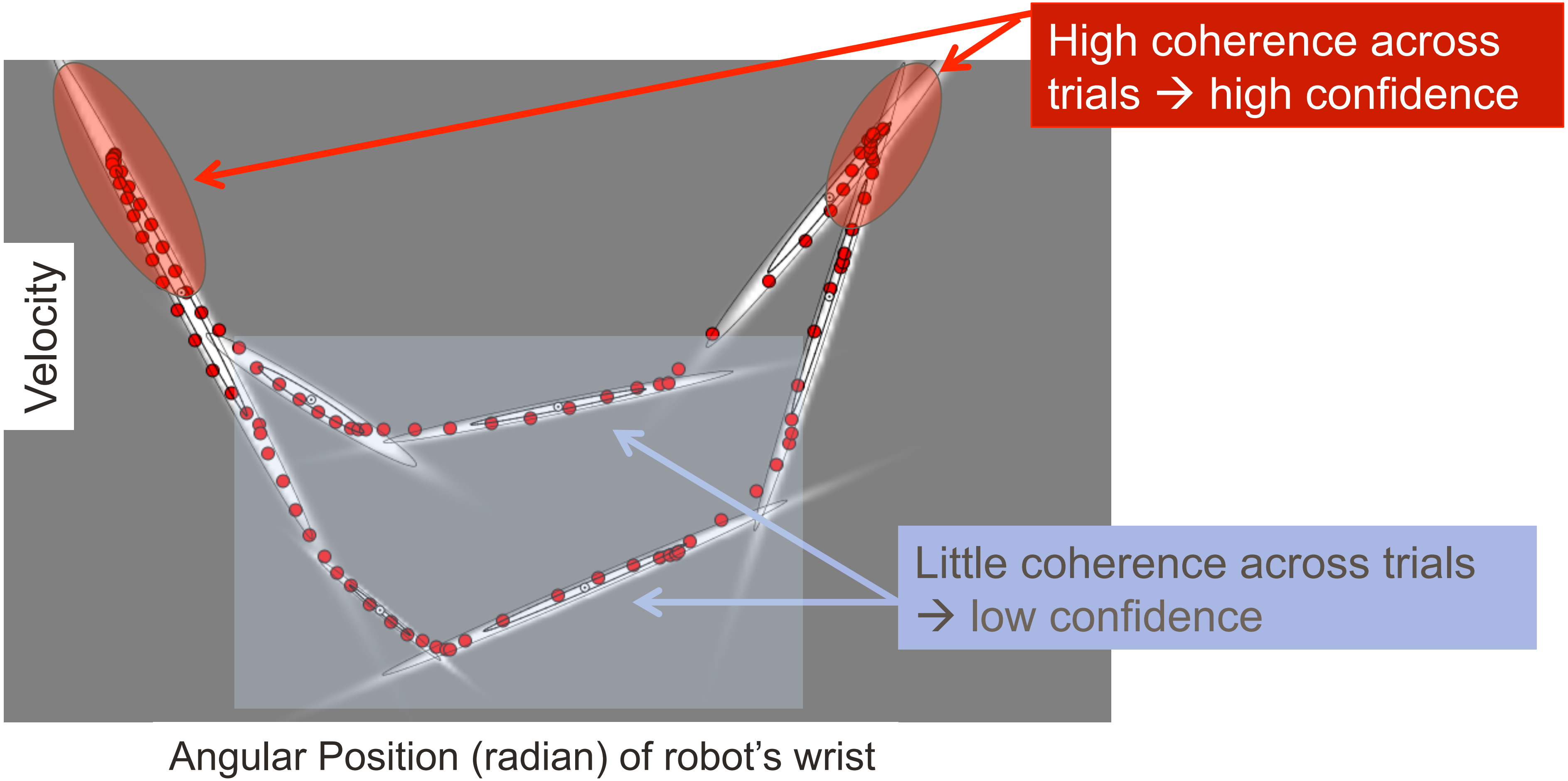
After training the robot manages to adapt naturally when required and remains stiff when required.



Learning from Bad Demonstrations

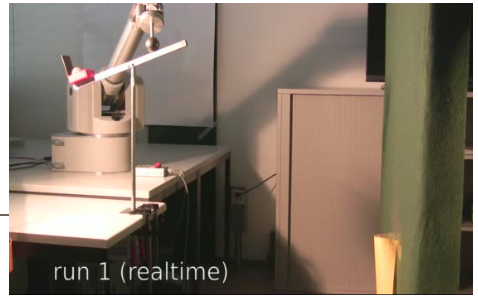


- Search around the demonstrations
- Reproduce only parts where all demonstrators agreed
- Avoid regions with high uncertainty

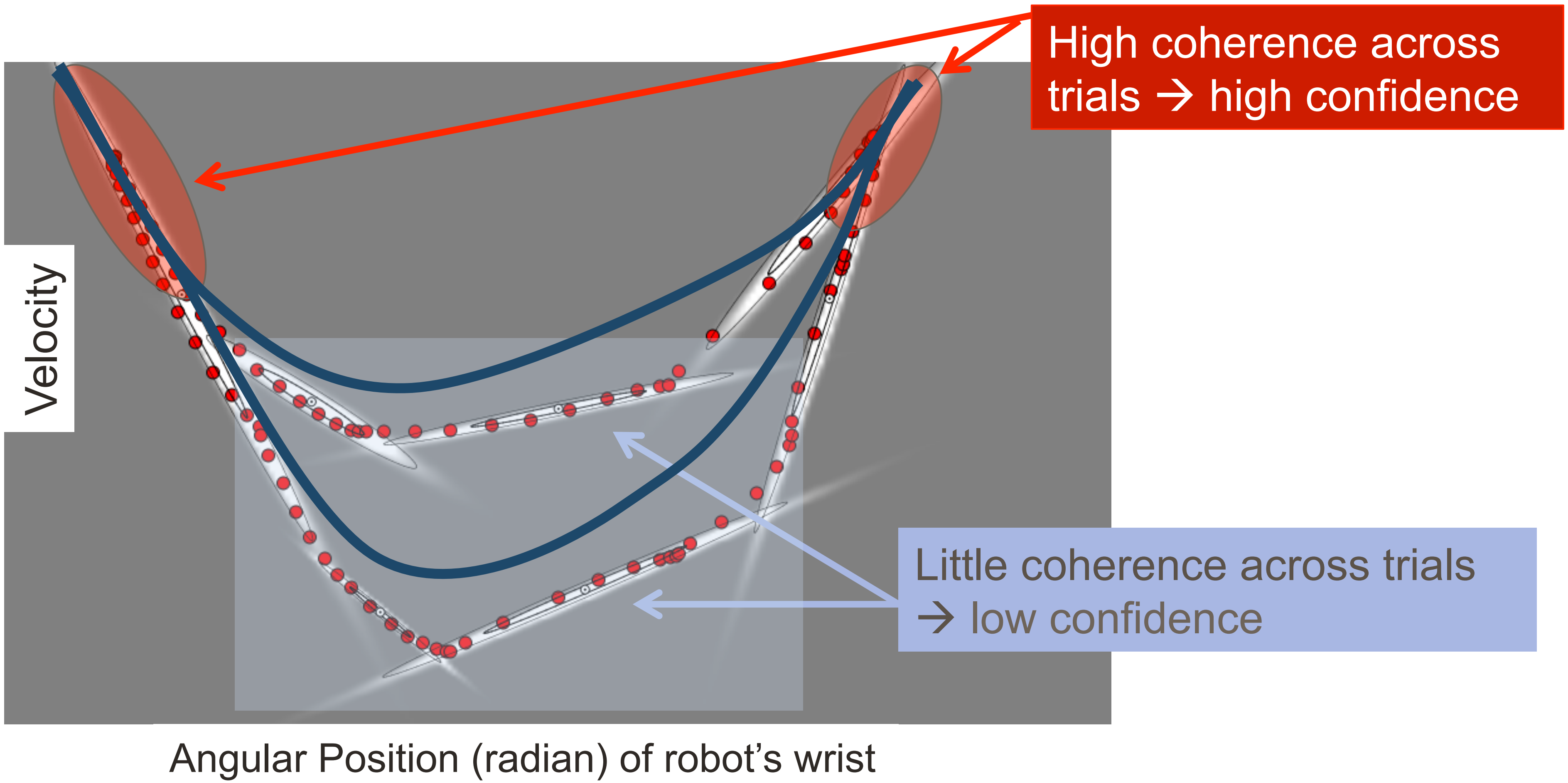


Grollman and Billard, ICRA 2012, Best Paper Award Cognitive Robotics

Learning from Bad Demonstrations



- Search around the demonstrations
- Reproduce only parts where all demonstrators agreed
- Avoid regions with high uncertainty



Conclusion

Learning from human demonstration is foremost generalizing

- Learning a generic control law
- Learning feasible regions of the state space

Observing human demonstration is not sufficient to perform the task

- Extracting key features from demonstrations
- Use these to adapt the trajectory

Demonstrations do not need to be perfect solutions to the task

→ Learning from bad demonstrations provides crucial information on what is key to perform the task.

→ More useful to know several feasible solutions to the task than a single but optimal one

The Lab

/lasa.epfl.ch