27th INTERNATIONAL CONFERENCE ON EMERGING TECHNOLOGIES AND FACTORY AUTOMATION

ETFA 2022 STUTTGART, GERMANY SEPTEMBER 6th-9th 2022

Inverse Optimal Control for the identification of human objective: a preparatory study for physical Human-Robot Interaction

Paolo Franceschi[1,2], Nicola Pedrocchi[1] , Manuel Beschi[2]

[1] CNR-STIIMA, via Alfonso Corti 12, 20148, Milano, Italy
2 University of Brescia, Via Branze, 38, 25123, Brescia, Italy

27th INTERNATIONAL CONFERENCE ON EMERGING TECHNOLOGIES AND FACTORY AUTOMATION

ETFA 2022 STUTTGART, GERMANY SEPTEMBER 6th-9th 2022

UNIVERSITÀ DEGLI STUDI DI BRESCIA

STIIMA Sistemi e Tecnologie Industriali Intelligenti per il Manifatturiero Avanzato Consiglio Nazionale delle Ricerche

Who am I

https://www.stiima.cnr.it/
Research fellow (2018-present)
Main activities:

- Autonomous/Cooperative manipulation and assembly of heavy and bulky objects
- Workspace optimization to increase complex task's robustness

https://drimi.unibs.it/
Ph.D. student (2020-present)
Main activities:

- game theoretical approach to physical Human-Robot Interaction
- Human-Robot role arbitration

The DrapeBot project

https://www.drapebot.eu/
The DrapeBot project aims at human-robot collaborative draping.

My main activities:

- Develop a control framework to enable human-robot co-transportation of carbon fiber plies

Method:

- Game theoretic description of the task to take into account the interaction

Required components:

- Human cost function
- Human desired trajectory
- Arbitration policy

The DrapeBot project

https://www.drapebot.eu/

The DrapeBot project aims at human-robot collaborative draping.

My main activities:

- Develop a control framework to enable human-robot co-transportation of carbon fiber plies

Method:

- Game theoretic description of the task to take into account the interaction

Required components:

- Human cost function - Identification
- Human desired trajectory
- Arbitration policy

Background: optimal control

Optimal control: find a control input u for a dynamical system over time period such that an objective function J is optimized

$$\min_{u} J(x(t), u(t), t)$$
$$s.t \quad \dot{x} = f(x(t), u(t), t)$$
$$x(0) = x_0$$

Background: optimal control

Optimal control: find a control input u for a dynamical system over time period such that an objective function J is optimized

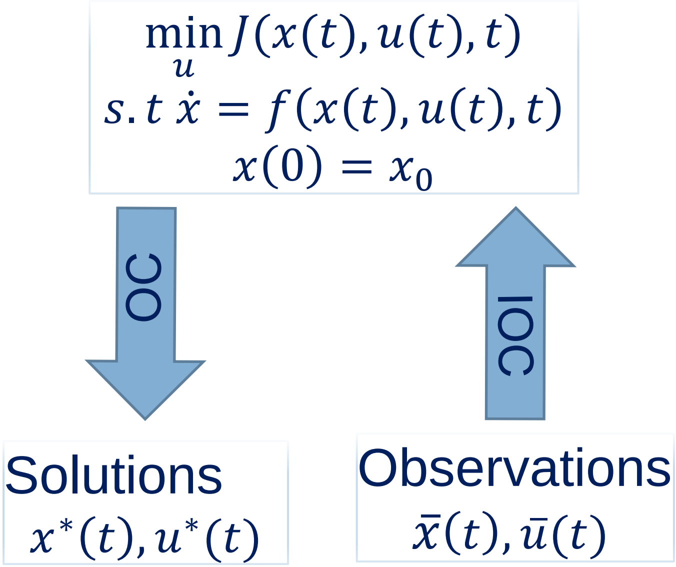
$$\min_{u} J(x(t), u(t), t)$$
$$s.t\ \dot{x} = f(x(t), u(t), t)$$
$$x(0) = x_0$$

OC

Solutions
$x^*(t), u^*(t)$

Background: Inverse optimal control

Inverse Optimal control: given observed state $\bar{x}(t)$ and control histories $\bar{u}(t)$, given the system dynamics in $\dot{x} = f(x(t), u(t))$, recover the cost function $J(x(t), u(t))$ that produced such control histories.



WHY IOC ?

- Observe nature and learn how it behaves
  - J. Mainprice, R. Hayne and D. Berenson, "Predicting human reaching motion in collaborative tasks using Inverse Optimal Control and iterative re-planning," *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015.
  - Westermann, K. et al. “Inverse optimal control with time-varying objectives: application to human jumping movement analysis.” *Sci Rep* 10, (2020).
  - T. Maillot et al, "How pilots fly: An inverse optimal control problem approach," *52nd IEEE Conference on Decision and Control*, 2013
  - I. A. Faruque, F. T. Muijres, K. M. Macfarlane, A. Kehlenbeck, and J. S. Humbert, “Identification of optimal feedback control rules from micro-quadrotor and insect flight trajectories” *Biological cybernetics*, 2018
  - M. C. Priess, R. Conway, J. Choi, J. M. Popovich, and C. Radcliffe, “Solutions to the inverse lqr problem with application to biological systems analysis” *IEEE Transactions on Control Systems Technology*, 2015
- Teach to a robot a behavior rather than a trajectory – Learning-By-Demonstration
  - Park, Taesung, and Sergey Levine. "Inverse optimal control for humanoid locomotion." *Robotics science and systems workshop on inverse optimal control and robotic learning from demonstration*. 2013.
  - Finn, Chelsea, Sergey Levine, and Pieter Abbeel. "Guided cost learning: Deep inverse optimal control via policy optimization." *International conference on machine learning*. PMLR, 2016.
  - Kalakrishnan, M. “Learning objective functions for autonomous motion generation”, *University of Southern California*, Los Angeles, CA, 2014

Linear-Quadratic case

Linear-Quadratic formulation simplifies both Optimal and Inverse Optimal Control problems, and its applicability to many problems explains its spread

$\dot{x} = Ax + Bu$

Linear system, with A and B state and input matrices

$J(x,u) = \int w^T \phi \, dt$

Quadratic cost function, with ϕ and w vectors of features and weights

$J(x,u) = \int (x^T Q x + u^T R u) \, dt$

Classical LQ definition of the cost function, with matrices Q and R weighting state and control features

Linear-Quadratic case

The Optimal Control problem is now

$$\min_u J = \int (x^T Q x + u^T R u)\, dt$$
$$s.t \quad \dot{x} = Ax + Bu$$
$$x(0) = x_0$$

And the feedback solution is

$$u^*(t) = -Kx(t)$$

With feedback gain matrix K

$$K = R^{-1} B^T P$$

With matrix P solution of the Algebraic Riccati Equation (ARE)

$$0 = A^T P + P A^T - P B R^{-1} B^T P + Q$$

Linear-Quadratic IOC

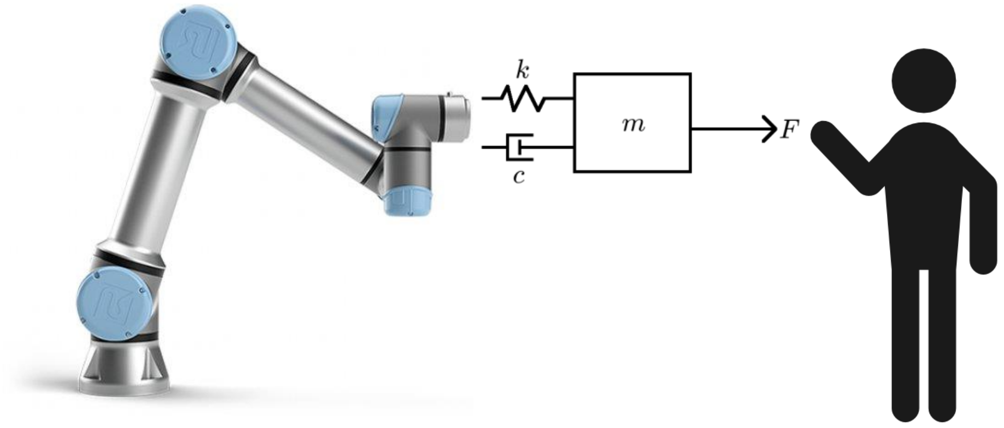
The LQ Inverse Optimal Control problem can be solved exploiting the ARE, if the matrix K is known (Least-squares, Recursive least squares, EKF methods can be exploited)
In the LQ IOC case, the problem can be reformulated as minimization of a residual in a constrained Quadratic Problem form

$$\min_{w} |r|^2 = \frac{1}{2}\, w^T H w$$
$$s.t \quad I\, w \geq 0$$
$$R > 0$$

In this, as an intuition matrix H contains a reformulation of the ARE*, while the constraints include the standard Optimal Control requirements (semi-positiveness of Q and positiveness of R)

*please refer to the workshop article, and to *J. Inga et al., "Solution sets for inverse non-cooperative linear-quadratic differential games," IEEE Control Systems Letters, vol. 3, no. 4, pp. 871–876, 2019.* for a full mathematical tractation

System modeling



$$M_i\, \ddot{z} + C_i \dot{z} + K_i z = F$$

Impedance equation

$$\begin{Bmatrix} \dot{z} \\ \ddot{z} \end{Bmatrix} = \begin{pmatrix} 0 & J_a \\ -M_i^{-1} K_i & -M_i^{-1} C_i \end{pmatrix} \begin{Bmatrix} \dot{z} \\ \ddot{z} \end{Bmatrix} + \begin{pmatrix} 0 \\ M^{-1} \end{pmatrix} u$$

$$\dot{x} = A x + B u$$

State-space reformulation

$$u = F = -Kx$$

Full-state feedback human control*

Having written the problem in linear state space formulation* allows to apply all the LQ (Inverse) Optimal Control techniques presented above

*the human control described as full-state feedback is common in literature and well approximate human behavior.
See as example Li, Yanan, et al. "Differential game theory for versatile physical human–robot interaction." Nature Machine Intelligence 1.1 (2019): 36-43.

Experiments

Data are collected by:

- Three subjects perform a reaching task directly imposing a force to the robot tip to a set-point along one direction
- The set-point randomly changes at each record $z_{sp} = [0.4, 0.9]$
- Three different impedance control tuning ($M_i = 10, C_i = \{25, 50, 75\}, K_i = 0$)
- 12 records are taken for each impedance tuning, for each subject
- A UR5 robot equipped with a Robotiq FT300 force sensor are used for the test, the robot provides positions records and the sensor the force records

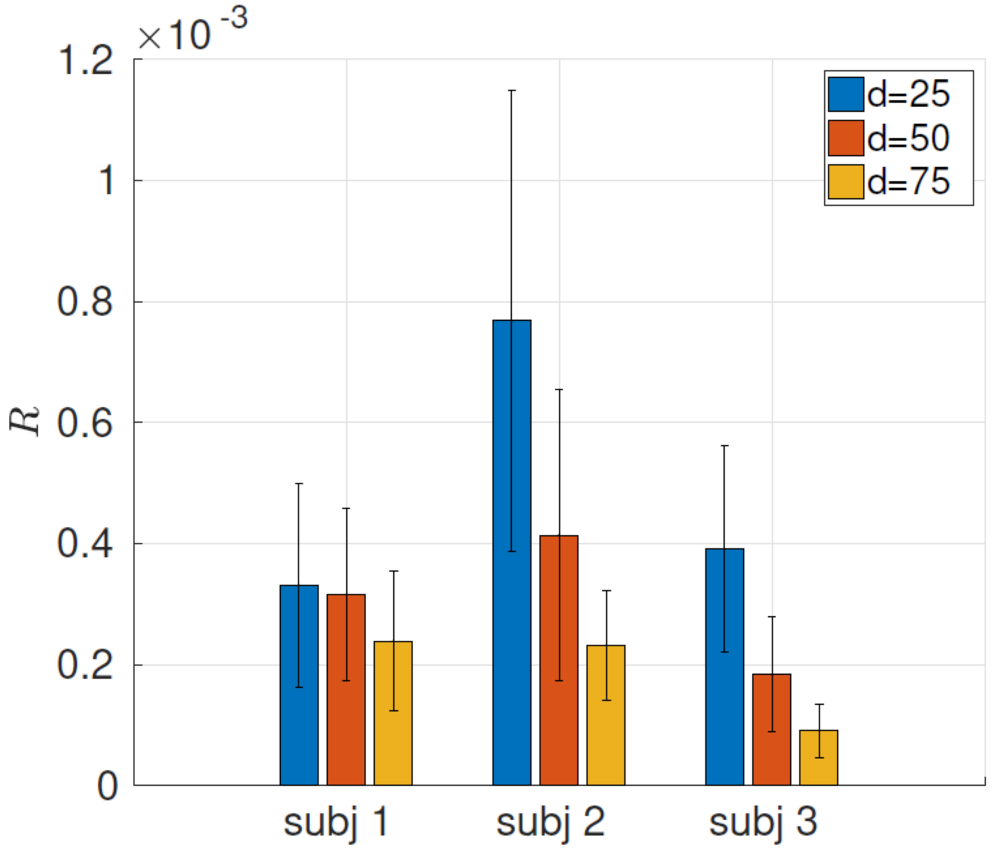
Note 1: the human cost function matrices, as typically happens, are assumed to be nonzero only on the diagonal terms, resulting in $Q = diag([q1, q2])$, and $R = r$.
Note 2: the solution to the optimal control problem is the same for any λJ with λ scalar. The constraint $q1 = 1$ is added to the QP problem to make results comparable.

Results - Recovered Q,R

- $q_1 = 1$ by constraint
- $q_2 \cong 0$ *
- r low, visible in figure

Interesting:

1. The low value of r, compared with q1, means that a fast set-point reaching is more favorable than saving effort.
2. Cost function varies depending on the system's tuning
3. Optimal behavior → all the same

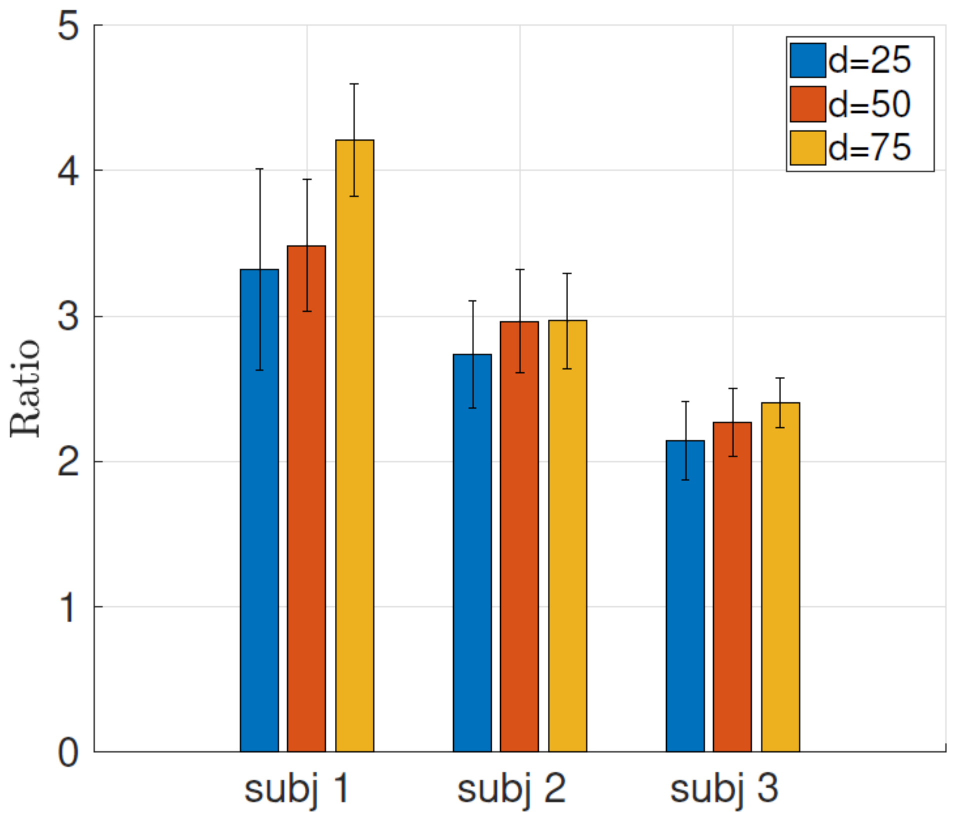


* A similar result is obtained in J. Inga, M. Eitel, M. Flad, and S. Hohmann, "Evaluating human behavior in manual and shared control via inverse optimization," in 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2018, pp. 2699–2704.

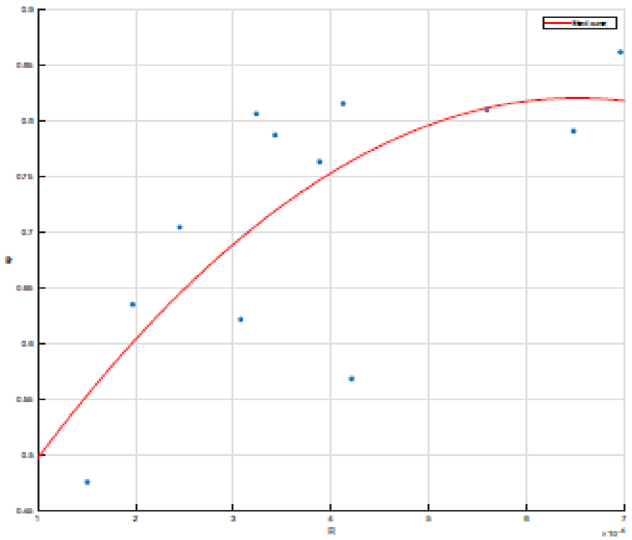
Results - risetime

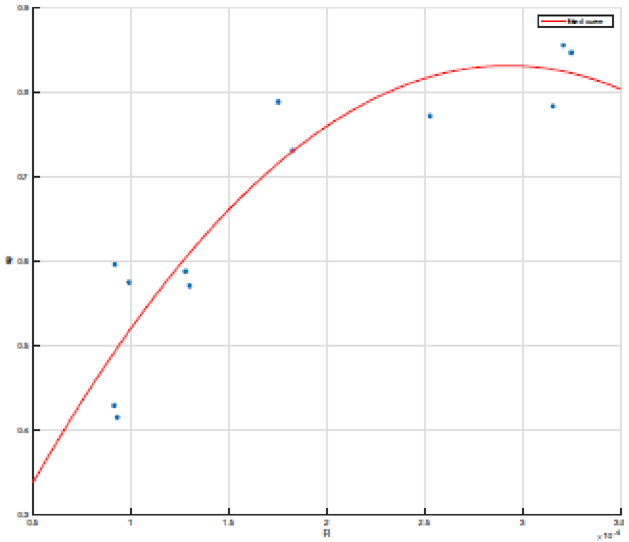
- Risetime RT (time to reach $0.95\, x_{sp}$) computed as $Ratio = \frac{RT}{x_{sp}}$

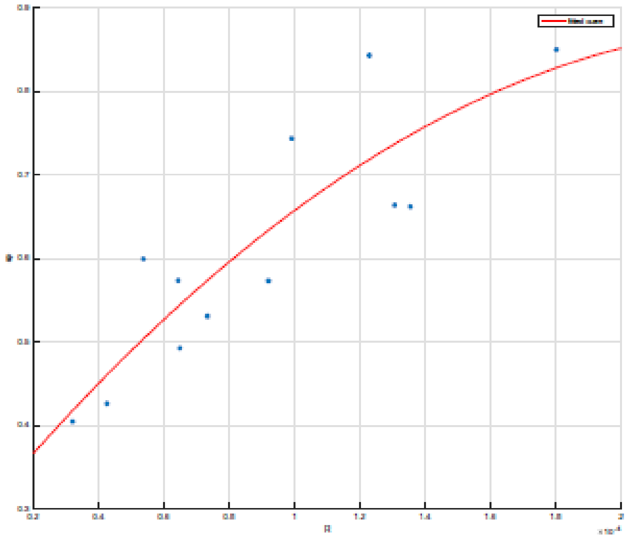
Figure shows that the rise time is almost constant for a subject independently from the system: humans tend to keep constant the required time



Results – R vs Set-Point



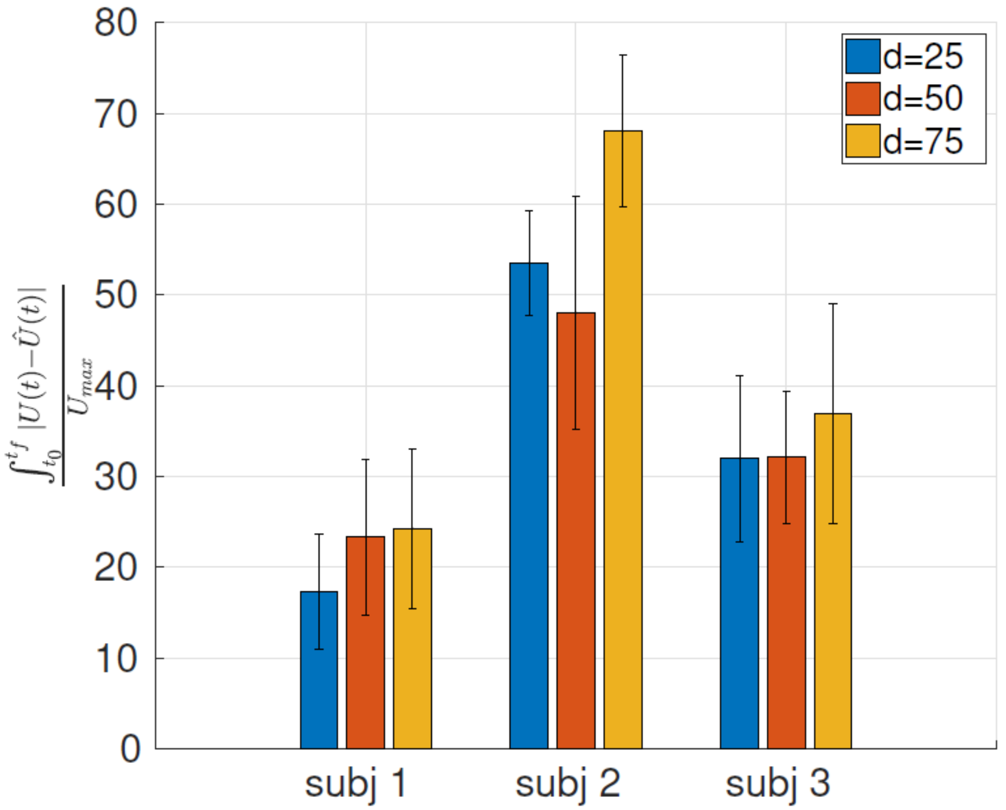




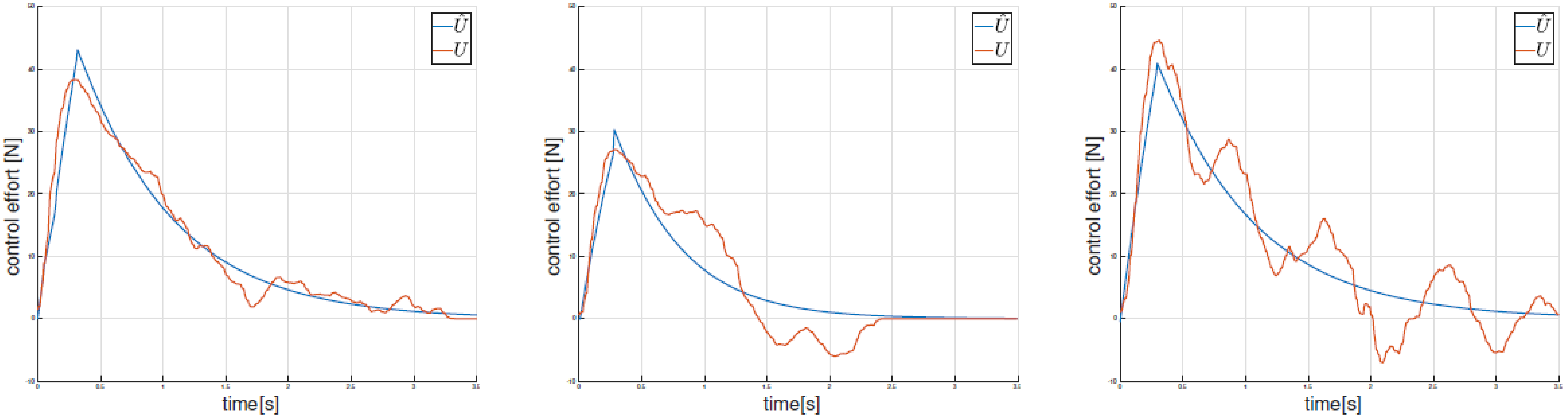
As far the set-point is, as much control weight (R) increases

Results – OC model

- To measure how well the recovered cost matches the actual control input the following index is defined $\mathcal{E}_u = \int_{t_0}^{t_f} \frac{u(t) - \hat{u}(t)}{u_{max}}\, dt$



Results – OC model



Conclusions

- Humans tend to minimize the reaching time rather than control effort
- IOC captures the essential human behavior (peak, trend)
- Humans control varies according to the system, to the task and to the subject, hence it is not generalizable
- IOC can be adopted online to make its approach general

Future works

- Robot active to provide support to the human
- Different results are expected for the LQ cost function
- Different cost functions will be implemented taking into account minimum time, jerk, energy, etc
- Analysis of human behavior in different scenarios (Cooperative, Non-Cooperative) to check the most suitable description

Thank you for your attention

p.franceschi@unibs.it, paolo.franceschi@cnr.stiima.it