An *in-context learning* framework for control

Riccardo Busetto[1,3], Valentina Breschi[2], Marco Forgione[1], Dario Piga[1], Simone Formentin[3]

[1]*IDSIA Dalle Molle Institute for Artificial Intelligence*
[2]*Department of Electrical Engineering, Eindhoven University of Technology*
[3]*Dipartimento di Elettronica, Bioingegneria e Informazione, Politecnico di Milano*

Contact us

In-context learning approach

- Design of **a unified model, estimator or controller from** data $\mathcal{D} = \{D^{(i)}\}_{i=1}^{M}$ of a **class of *similar* systems** $\mathcal{S} = \cup_{i=1}^{M} S^{(i)}$
- **Transformers** are trained **for the class**, but can **adapt to the specific system** extrapolating the **context** from **data**

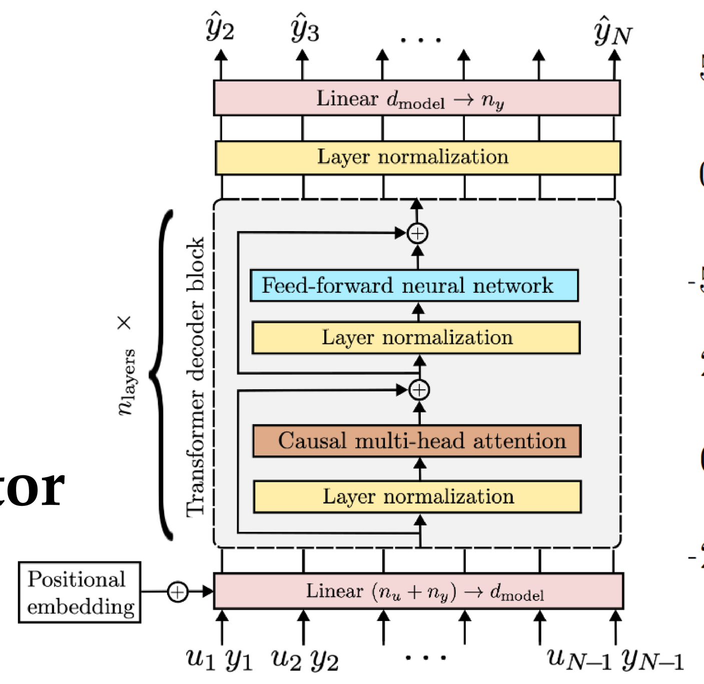
One system ⟷ **One model**

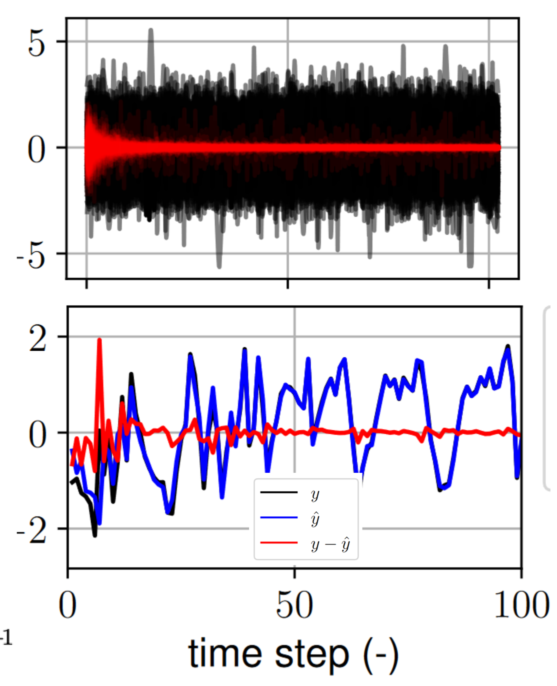
One model ⟷ **All systems!**

Paper link

System Identification

- **Goal:** one-step prediction $\hat{y}_{k+1} = \mathcal{M}_\phi(u_{1:k}, y_{1:k}), \;\; \forall S \in \mathcal{S}$
- **Context** is understood from input/output data $D^{(i)} = (u_{1:N}^{(i)}, y_{1:N}^{(i)})$
- **Learning** from (potentially infinite) set of similar systems through a **simulator**
- **Excellent performance** on nonlinear systems (Class of WH systems)





Paper in preparation.. stay tuned

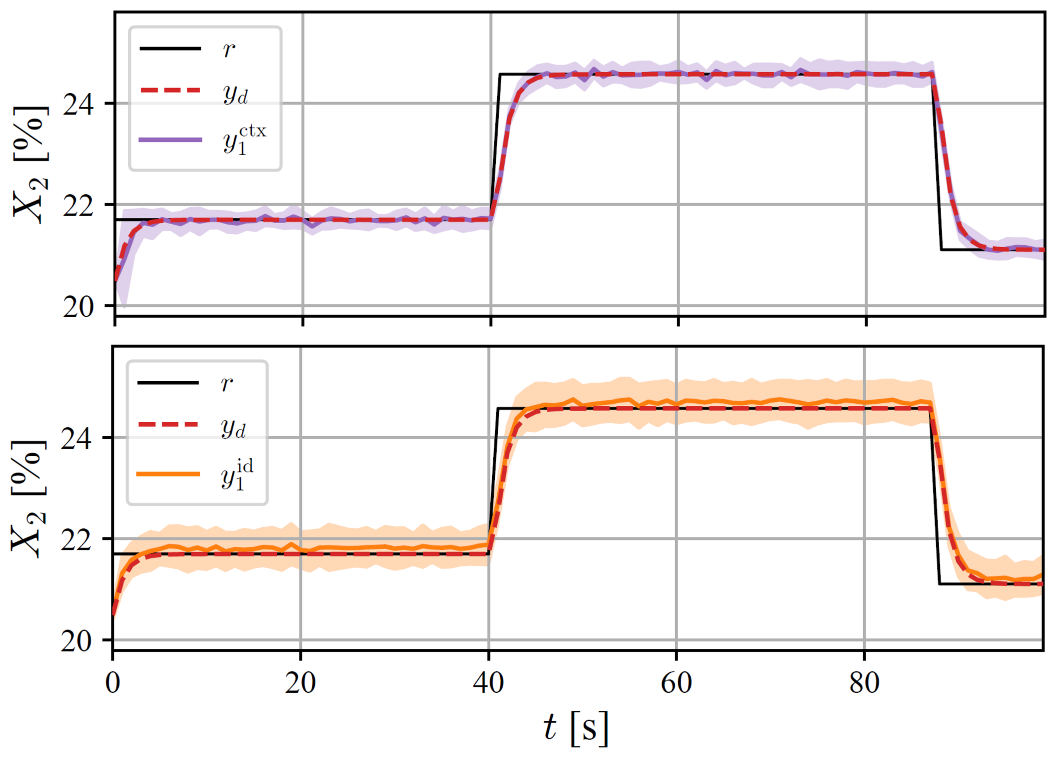
Control

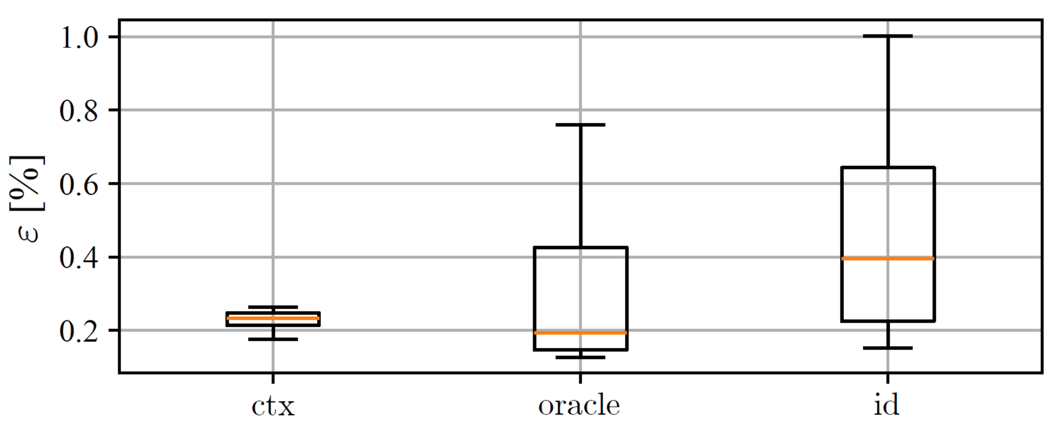
$$\hat{u}_k = \mathcal{C}_\phi(e_{1:k}, u_{0:k-1})$$

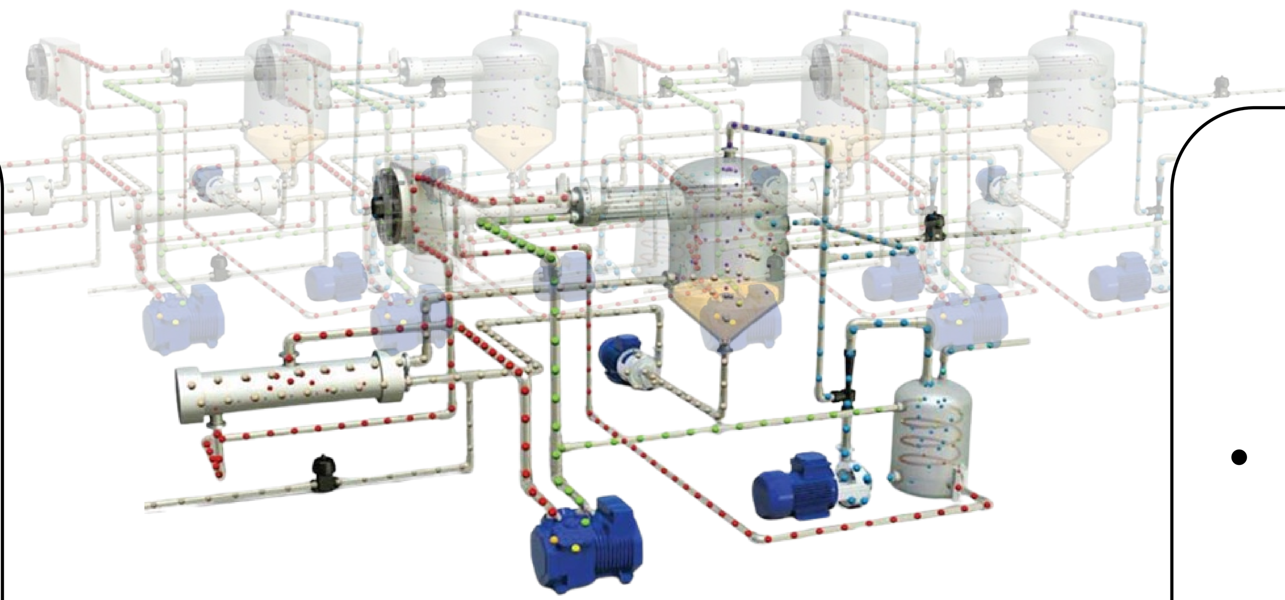
- **Goal:** model reference control

$$\min_\phi \mathbb{E}_{p(\mathcal{D})} \left[\sum_{k=1}^{N-1} \left\| u_k^\star - \mathcal{C}_\phi(e_{1:k}, u_{0:k-1}) \right\|_2^2 \right]$$

- **Context** is understood from errors $e_{1:k}$ and inputs $u_{0:k-1}$
- **Matching error is reduced** with respect to NMPC

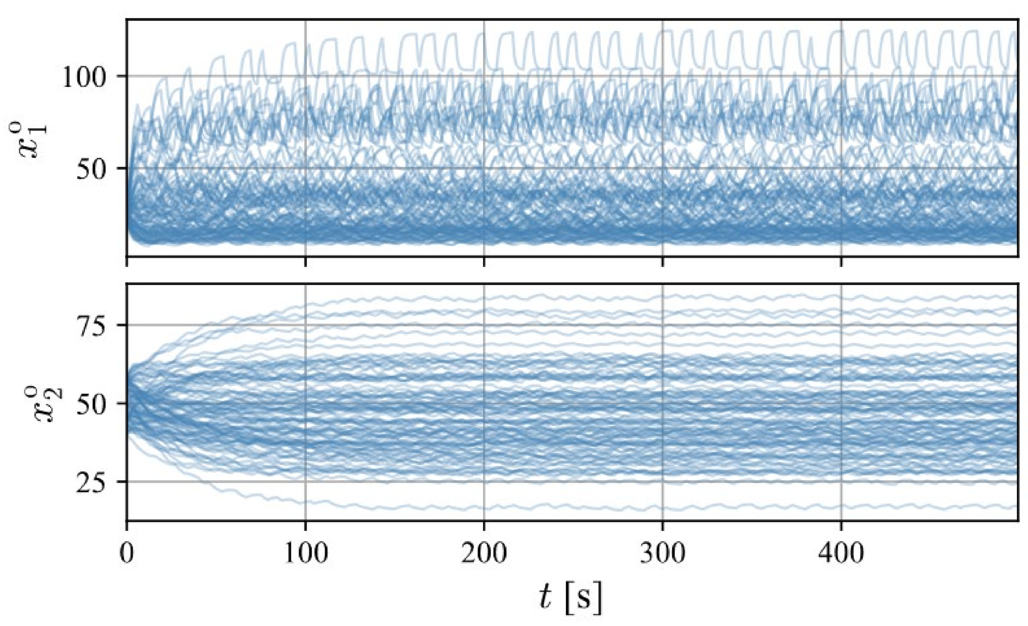






Example: evaporation process

- Nonlinear benchmark
- Class $\mathcal{S}$ described by 19 parameters, each perturbed by 5-20%



Paper link

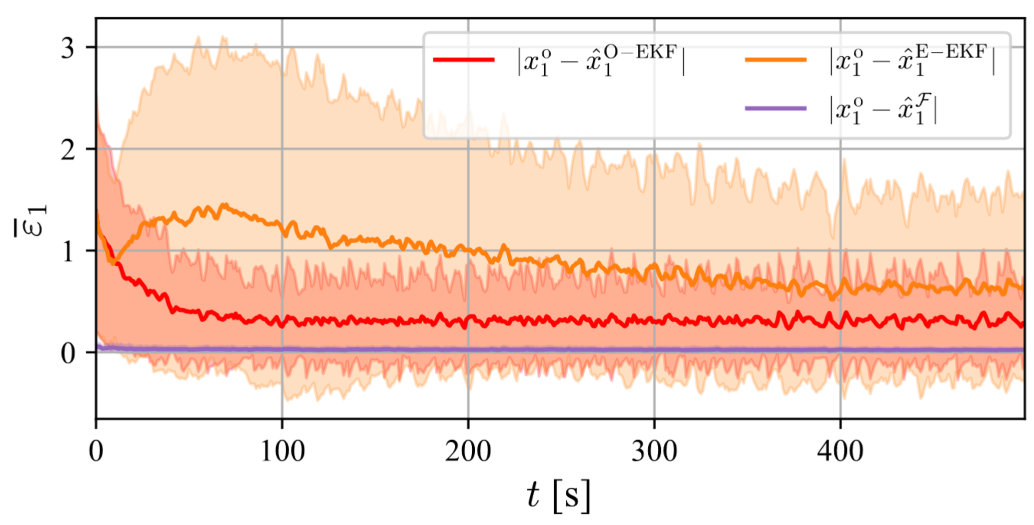
State Estimation

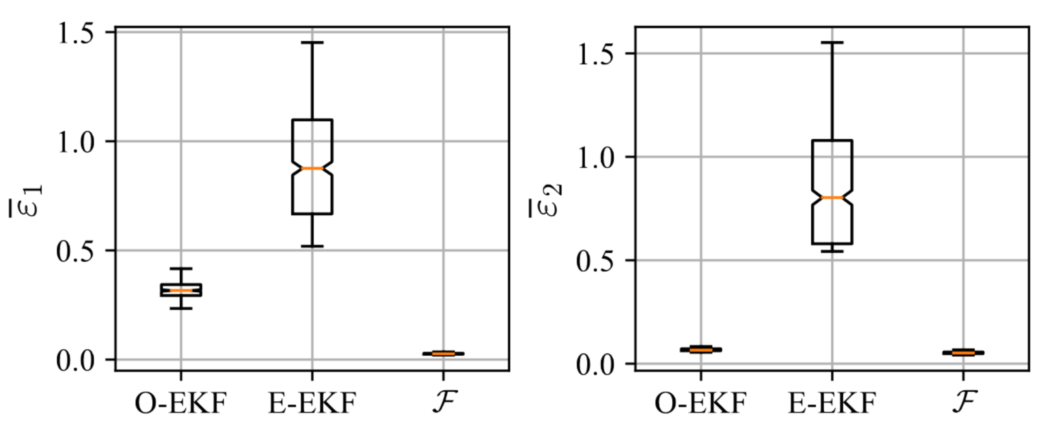
$$\hat{x}_k = \mathcal{F}_\phi(u_{1:k}, y_{1:k}),$$

- **Goal:** state estimation

$$\min_\phi \mathbb{E}_{p(\mathcal{D})} \left[\sum_{k=1}^{N-1} \left\| x_k^o - \mathcal{F}_\phi(u_{1:k}, y_{1:k}) \right\|_2^2 \right]$$

- **Context** is undestood from inputs $u_{1:k}$ and outputs $y_{1:k}$
- **Estimation error is reduced** with respect to EKF





Future works

- **Experimental validation** on applications
- Architectural modifications to **speed up learning**
- **Formal results** on properties of the approach (stability, effect of noise)