Maintaining a relevant dataset for data-driven MPC using Willems' fundamental lemma extensions

Alexandre Faye-Bédrin Stanislav Aranovskiy
Paul Chauchat Romain Bourdais

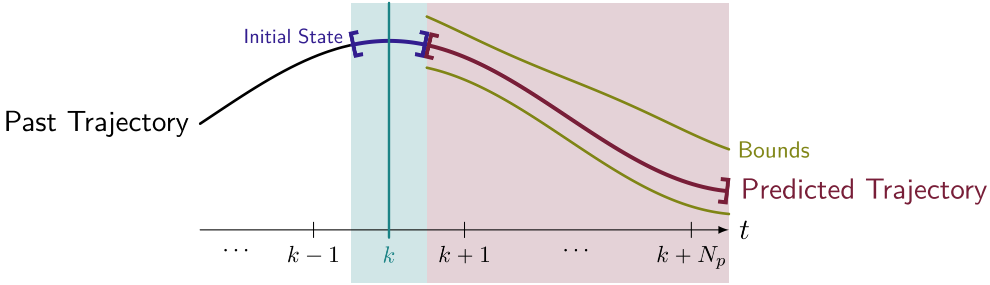
CentraleSupélec IETR

Réunion CT CPNL – 16th November 2023

Model Predictive Control

MPC over horizon N_p: at time k, solve

$$\min_u J(y,u) \quad \text{with} \quad \begin{cases} x_{i+1} = f(x_i, u_i) \\ (u,x) \in \mathbb{X} \\ x_k \end{cases}$$



Data-Driven MPC?

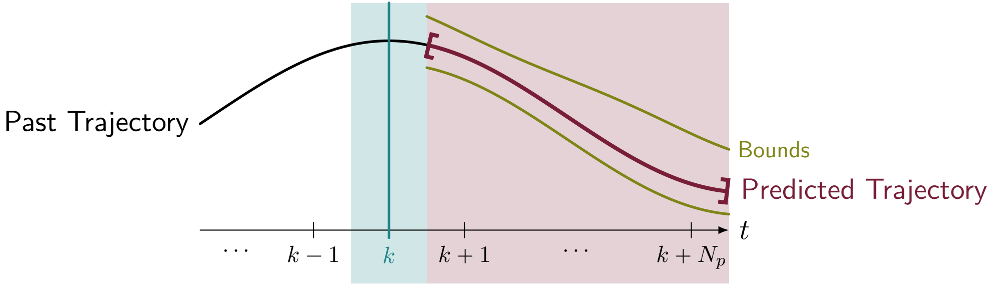
Data-Driven MPC over horizon N_p: at time k, solve

$$\min_u J(y,u) \quad \text{with} \quad \begin{cases} x_{i+1} = f(x_i, u_i) \\ (u,x) \in \mathbb{X} \\ x_k \end{cases}$$

Data! but which and how?

$(u, y) \in \mathbb{X}$

Initial Conditions?



Outline

Trajectory

In the behavioral framework, a system is a set of trajectories.
Let system $\mathcal{S}$, with input u and output y.

Definition (Trajectory)

Let τ_L a sequence of L samples of (u, y): $\tau_L = \begin{bmatrix} u_{[k,k+L-1]} \\ y_{[k,k+L-1]} \end{bmatrix}$.
τ_L is a trajectory of $\mathcal{S}$ when consistent with $\mathcal{S}$ behaviour.
We write: $\tau_L \in \mathcal{S}_{|L}$

Trajectory

In the behavioral framework, a system is a set of trajectories.
Let system $\mathcal{S}$, with input u and output y.

Definition (Trajectory)

Let τ_L a sequence of L samples of (u, y): $\tau_L = \begin{bmatrix} u_{[k,k+L-1]} \\ y_{[k,k+L-1]} \end{bmatrix}$.
τ_L is a trajectory of $\mathcal{S}$ when consistent with $\mathcal{S}$ behaviour.
We write: $\tau_L \in \mathcal{S}_{|L}$

Let $\mathcal{S}_{\mathcal{L}}$ a LTI system: we have $\mathcal{S}_{\mathcal{L}|L} \subseteq (\mathbb{R}^{n_u+n_y})^L$
We can arrange a collection of trajectories $(\tau_{L,i})_{i=\{1...N\}}$ as a matrix:

$$\mathcal{T}_L = \begin{bmatrix} \tau_{L,1} & \tau_{L,2} & \cdots & \tau_{L,N} \end{bmatrix}$$

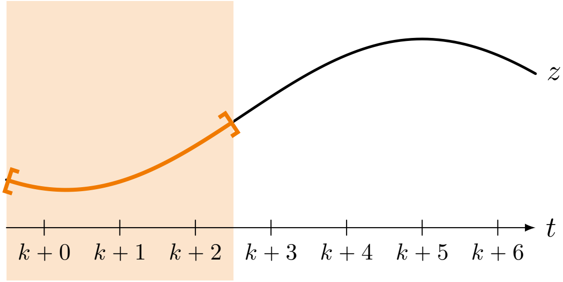
and then

$$Im(\mathcal{T}_L) \subseteq \mathcal{S}_{|L}$$

The Hankel Matrix

With $L = 3$ and $N = 7$:

$$H_L(z_{[k,k+N-1]}) = \begin{pmatrix} z_{k+0} & z_{k+1} & z_{k+2} & z_{k+3} & z_{k+4} \\ z_{k+1} & z_{k+2} & z_{k+3} & z_{k+4} & z_{k+5} \\ z_{k+2} & z_{k+3} & z_{k+4} & z_{k+5} & z_{k+6} \end{pmatrix}$$



Collection from a Single Trajectory

Take $(u^d, y^d) \in \mathcal{S}_{\mathcal{L}|N}$, with $N \geqslant L$.

- $\mathcal{T}_L = \begin{bmatrix} H_L(u^d) \\ H_L(y^d) \end{bmatrix}$ is a "trajectory collection" matrix: $Im(\mathcal{T}_L) \subseteq \mathcal{S}_{\mathcal{L}|L}$
- Can we have $Im(\mathcal{T}_L) = \mathcal{S}_{\mathcal{L}|L}$?

Willems' Lemma

Theorem (Willems et al. 2005)

Let $\mathcal{S}_{\mathcal{L}}$ a LTI, controllable system of order n_x. Let $(u^d, y^d) \in \mathcal{S}_{\mathcal{L}|N}$, with u^d Persistently Exciting of order $L + n_x$. Then:

$$\begin{bmatrix} u \\ y \end{bmatrix} \in \mathcal{S}_{\mathcal{L}|L} \iff \exists \alpha \in \mathbb{R}^{N-L+1}, \begin{bmatrix} u \\ y \end{bmatrix} = \begin{bmatrix} H_L(u^d) \\ H_L(y^d) \end{bmatrix} \alpha$$

Willems' Lemma

Theorem (Willems et al. 2005)

Let $\mathcal{S}_{\mathcal{L}}$ a LTI, controllable system of order n_x. Let $(u^d, y^d) \in \mathcal{S}_{\mathcal{L}|N}$, with u^d Persistently Exciting of order $L + n_x$. Then:

$$\begin{bmatrix} u \\ y \end{bmatrix} \in \mathcal{S}_{\mathcal{L}|L} \iff \exists \alpha \in \mathbb{R}^{N-L+1}, \begin{bmatrix} u \\ y \end{bmatrix} = \begin{bmatrix} H_L(u^d) \\ H_L(y^d) \end{bmatrix} \alpha$$

Criterion (Persistence of Excitation)

u^d, with $u_k^d \in \mathbb{R}^{n_u}$, is Persistently Exciting of order K if:

$$rank(H_K(u^d)) = K n_u$$

Shorthand: $PE(u^d) = K$

Linear vs Affine

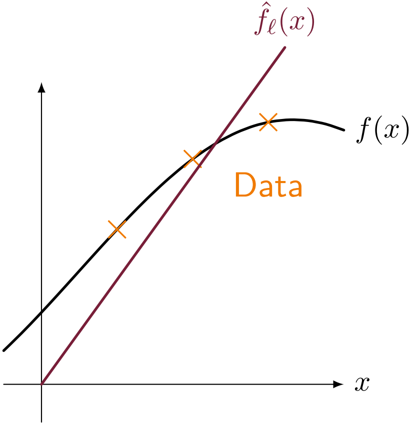


Figure 1: Linear approximation

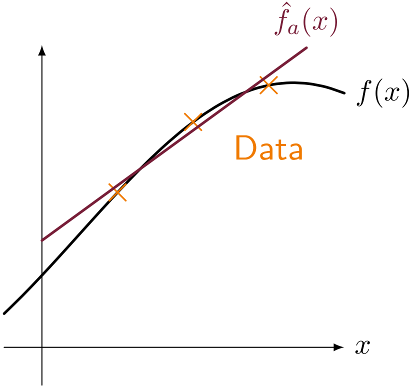


Figure 2: Affine approximation

$$\|\hat{f}_\ell - f\| \quad > \quad \|\hat{f}_a - f\|$$

Willems' Lemma for Affine Systems

Theorem (Berberich et al. 2022, based on Willems et al. 2005)

Let $\mathcal{S}_{\mathcal{A}}$ *an* *affine*, *controllable system of order* n_x. *Let* $(u^d, y^d) \in \mathcal{S}_{\mathcal{A}|N}$, *with* $PE(u^d) = L + n_x + 1$ *Then:*

$$\begin{bmatrix} u \\ y \end{bmatrix} \in \mathcal{S}_{\mathcal{A}|L} \iff \exists \alpha \in \mathbb{R}^{N-L+1}, \begin{cases} \begin{bmatrix} u \\ y \end{bmatrix} = \begin{bmatrix} H_L(u^d) \\ H_L(y^d) \end{bmatrix} \alpha \\ \sum_i \alpha_i = 1 \end{cases}$$

Definition (Affine system)

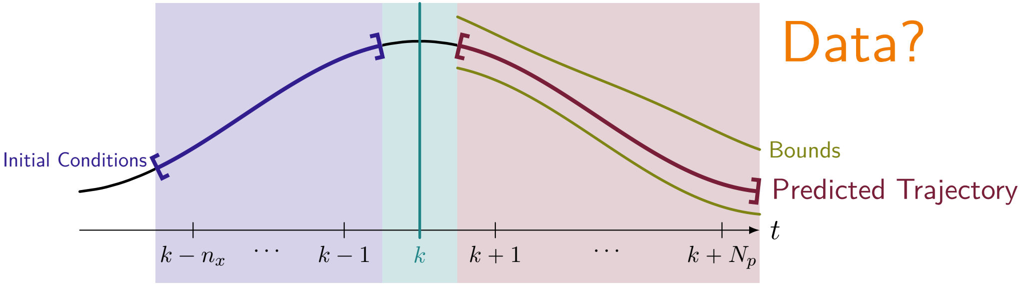
An affine system $\mathcal{S}_{\mathcal{A}}$ is like a LTI system with constant disturbance:

$$\begin{cases} x_{k+1} &= Ax_k + Bu_k + e \\ y_k &= Cx_k + Du_k + r \end{cases}$$

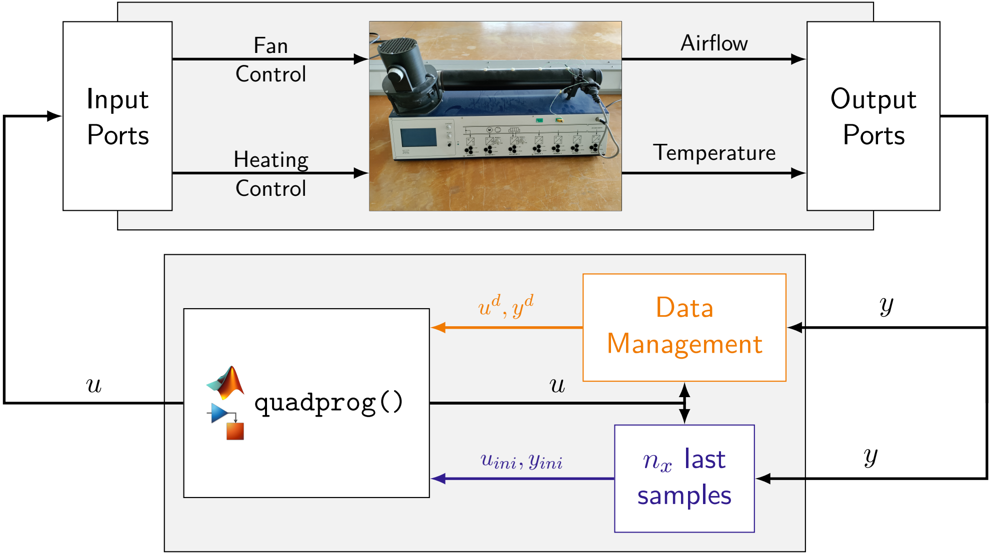
Data-Driven Model Predictive Control

Data-Driven MPC over horizon N_p: at time k, solve

$$\min_u J(y,u) \quad \text{with} \quad \begin{cases} (u,y) \in \mathcal{S}_{\mathcal{A}|L}(u^d, y^d) \\ (u,y) \in \mathbb{X} \\ (u_{ini}, y_{ini}) \end{cases}$$



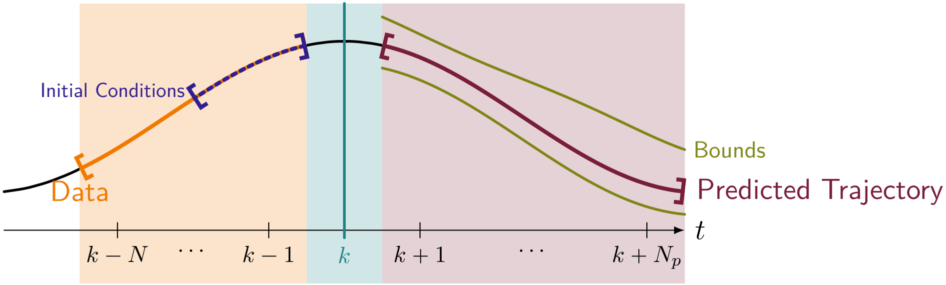
Experimental Setup



Experiment 1

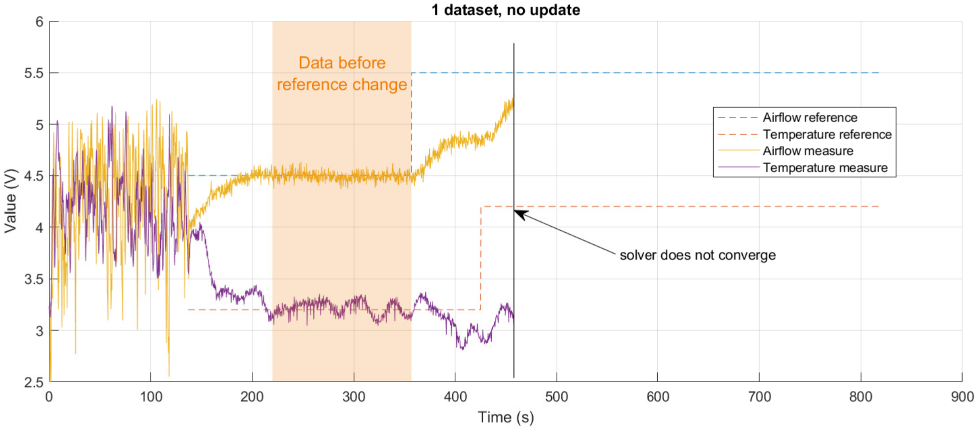
Basic Data Choice

At time k, take last N samples for u^d, y^d



Experiment 1

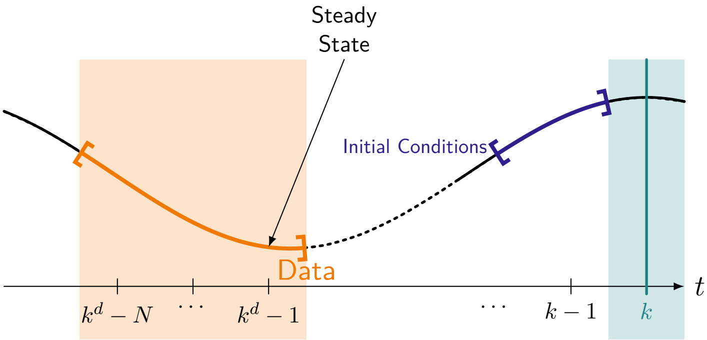
Need for Excitation



Experiment 2

Existing Heuristic[1]

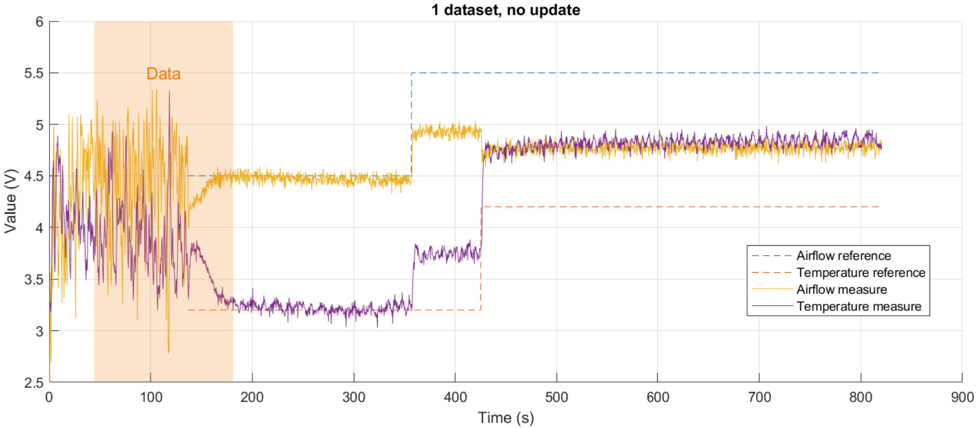
Update u^d, y^d until $\|y - y^s\| \leqslant \varepsilon$ (steady state attained)



[1] Baros et al. 2022; Berberich et al. 2021, 2022.

Experiment 2

Need for Updates



Data Management

What about noise?

- We would like to keep (u^d, y^d) updated...
- Proposed heuristic: update u^d, y^d whenever $u_{[k-N,k-1]}$ is PE
  - Matrix rank criterion not suited to noisy data (can give bad SNR)

  $\hookrightarrow$ Data updated too often... To avoid this:

Criterion (PE)

$PE(u^d) = K$:

$$rank(H_K(u^d)) = n_u \times K$$

$\Longrightarrow$

Criterion ("Enough" PE)

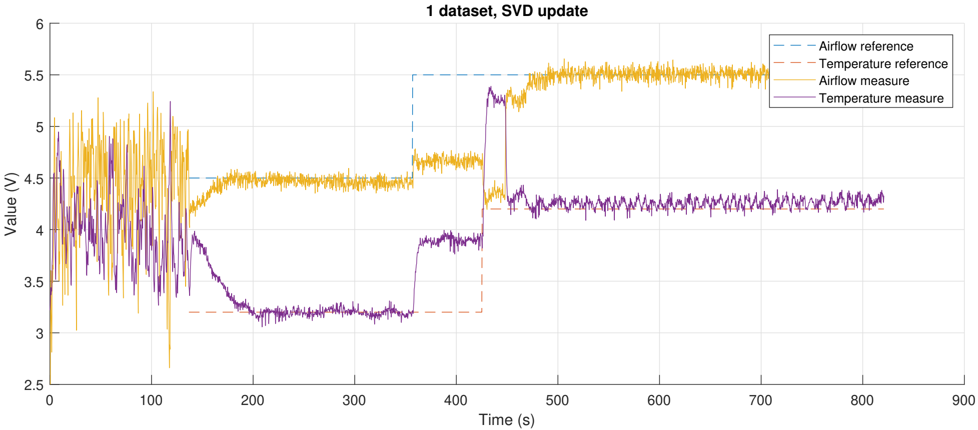
$PE(u^d) = K$:

$$\sigma_{min}(H_K(u^d)) \geqslant \varsigma$$

$\sigma_{min}(\cdot)$ *smallest singular value*

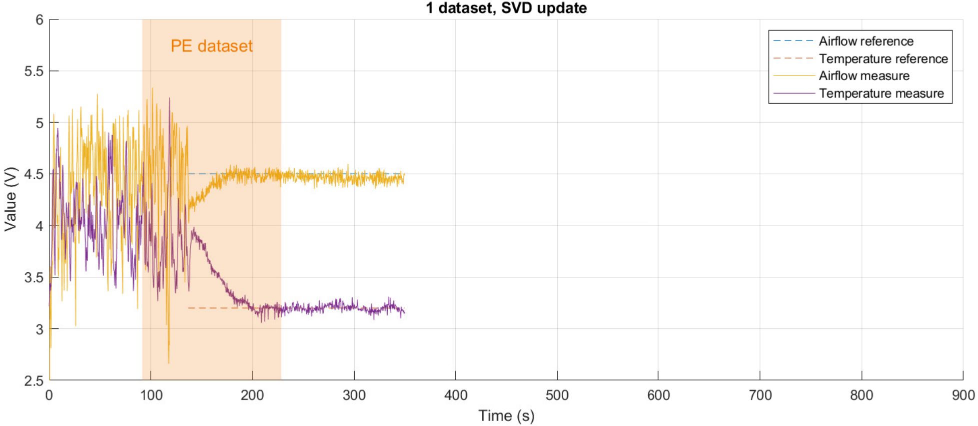
Experiment 3: Unique Dataset

(Sometimes) Improved Performance



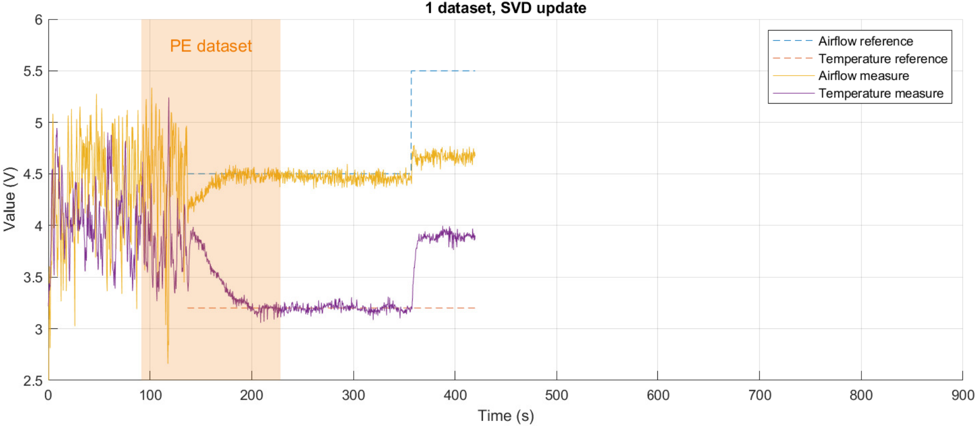
Experiment 3: Unique Dataset

(Sometimes) Improved Performance



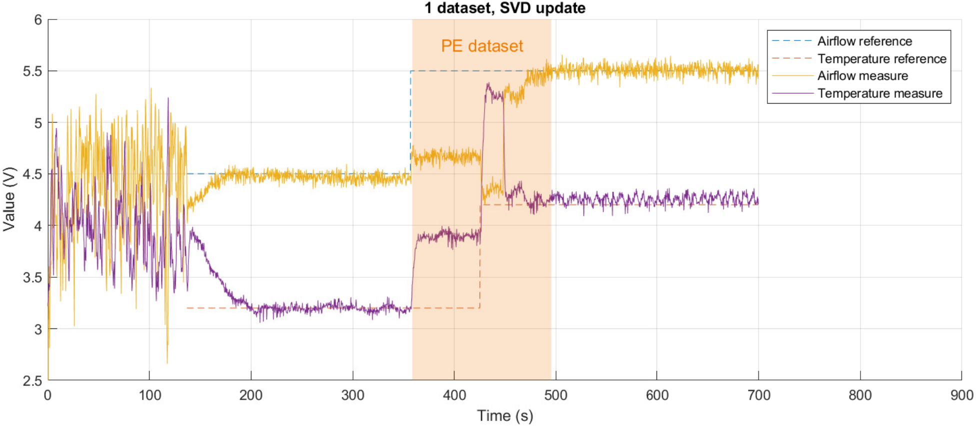
Experiment 3: Unique Dataset

(Sometimes) Improved Performance



Experiment 3: Unique Dataset

(Sometimes) Improved Performance



Data Management

Problem and Solution

- Conflict:
  - **Problem (Control)**
    u^d *close to setpoint*
    $\Longrightarrow$ Experiment 1: u^d not PE anymore
  - **Theorem (Willems' Lemma)**
    u^d *is PE* $\Longrightarrow$ *it works*
    $\Longrightarrow$ Experiment 2: u^d not close enough
- Proposed approach: combine 2 datasets
  - (u_ℓ^d, y_ℓ^d) PE
  - (u_a^d, y_a^d) tracking the operating point

Two datasets

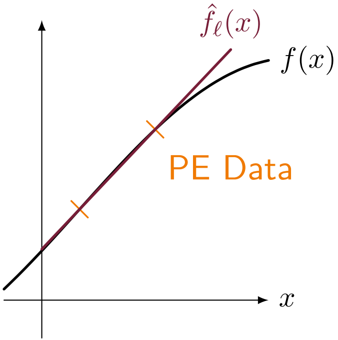


Figure 3: PE dataset: "slope"

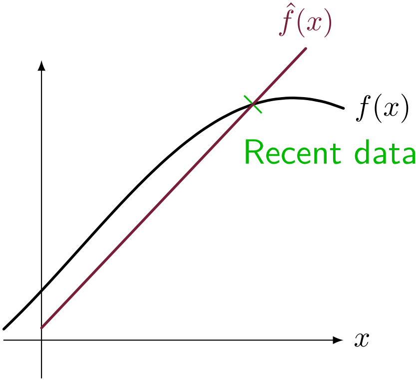


Figure 4: Recent dataset: "Offset"

Experiment 4: Combined Datasets

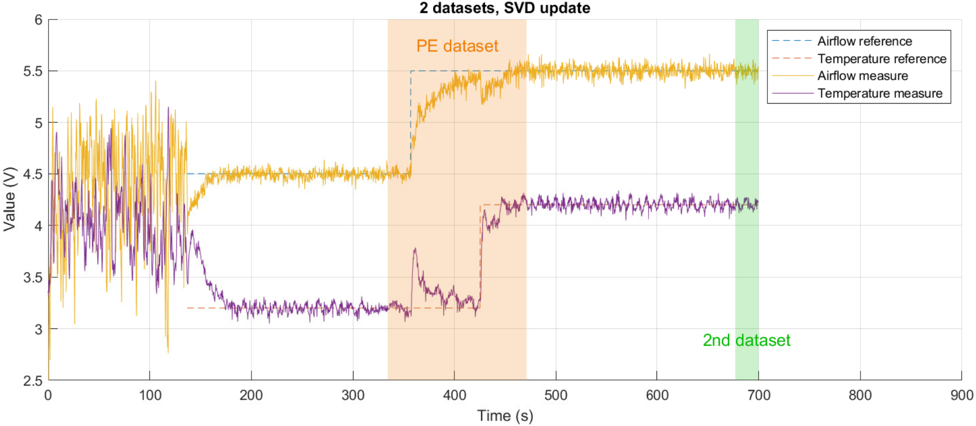
Improved Performance

Experiment 4: Combined Datasets

Improved Performance

Experiment 4: Combined Datasets

Improved Performance



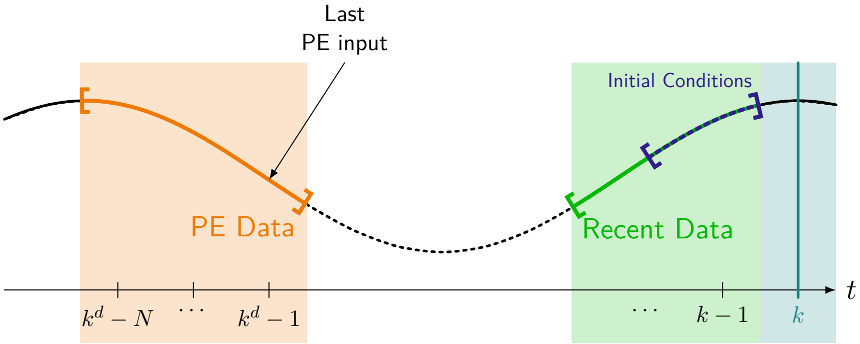
Limits & Ongoing Work

- Choice of threshold ς not easy
  - Compromise: excitation (good SNR) or novelty (better fit –hopefully)?
  - May change with setpoint...
- Limited complexity impact
- Hopefully: stability guarantee[2]

[2] extending the result of Berberich et al. 2022

Takeaway Messages

1. 2 datasets $\implies$ 2 desired properties
2. Singular value threshold for good SNR



Complexity

- Algorithm complexity:
  - SVD of $M \in \mathbb{R}^{n\times m}$: $O(mn\min(m,n))$
  - `quadprog(H,f)` with `f`$\in \mathbb{R}^p$: at best $O(p^3)$

$\hookrightarrow$ SVD has limited impact

- Optimized variable (1 dataset): $\alpha_1 \in \mathbb{R}^{N-L+1}$
  - $\dim(\alpha_1) = N - L + 1 \geqslant (L + n_x + 1)n_u$ (see PE criterion)

  $\hookrightarrow$ Complexity $O(N^3)$ or $O(L^3)$
- Optimized variable (2 datasets): $\alpha_2 \in \mathbb{R}^{N_a+N_\ell-2L+2}$
  - $N_\ell - L + 1 > (L + n_x + 1)n_u$ as before
  - $N_a \geqslant L$

  $\hookrightarrow$ $\dim(\alpha_2) \geqslant (L + n_x + 1)n_u + 1$

$\hookrightarrow$ No big change

References I

- Baros, Stefanos et al. (Apr. 1, 2022). "Online Data-Enabled Predictive Control". In: Automatica 138, p. 109926. DOI: 10.1016/j.automatica.2021.109926. (Visited on 02/03/2023).
- Berberich, Julian et al. (Apr. 2021). "Data-Driven Model Predictive Control with Stability and Robustness Guarantees". In: IEEE Transactions on Automatic Control 66.4, pp. 1702–1717. DOI: 10.1109/TAC.2020.3000182. (Visited on 11/28/2022).
- — (Sept. 2022). Linear Tracking MPC for Nonlinear Systems Part II: The Data-Driven Case. DOI: 10.1109/TAC.2022.3166851. (Visited on 01/11/2023). preprint.
- Willems, Jan C. et al. (Apr. 1, 2005). "A Note on Persistency of Excitation". In: Systems & Control Letters 54.4, pp. 325–329. DOI: 10.1016/j.sysconle.2004.09.003. (Visited on 01/12/2023).

Experiment 1b

Some fails

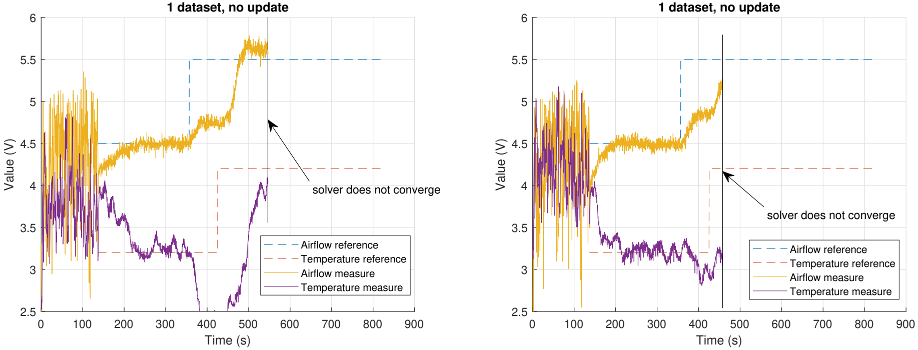


Figure 5: Data continuously updated.

Experiment 3b

Similar Performance

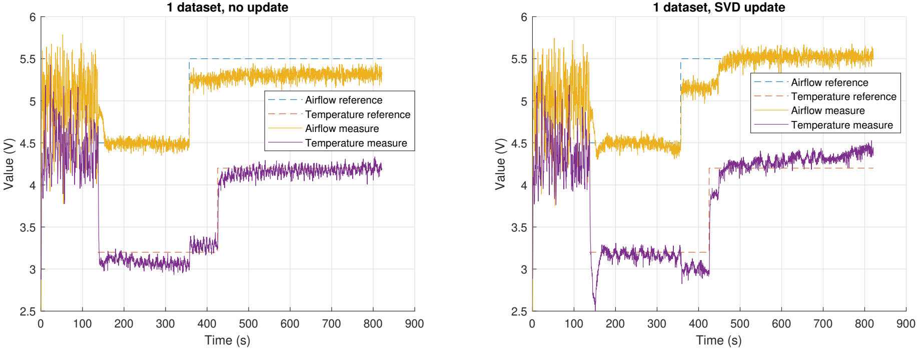


Figure 6: Left: Data frozen at the first setpoint. Right: Update according to criterion.

Notations

- Given a vector $z \in \mathbb{R}^{n_z}$:
  - z_k the value of z at time k
  - $z_{[k,k+\ell]} = \begin{pmatrix} z_k \\ z_{k+1} \\ \vdots \\ z_{k+\ell} \end{pmatrix} \in \mathbb{R}^{(\ell+1)n_z}$ the "stacked window"
  - $z_k(t)$ predicted/calculated value of z_{t+k} from information available at time t

  $\hookrightarrow$ $z_{[k_1,k_2]}(t)$ "what to expect over the time span $[t+k_1, t+k_2]$, based on knowledge at time t"
- Given a matrix $M \succ 0$ and a vector z: $\|z\|_M^2 := z^\top M z$
- $\otimes$ is the Kronecker product

Data-Driven Model Predictive Control

MPC + Willems' Lemma

Problem (Stabilization of $\mathcal{S}$ close to y^r, see Berberich et al. 2022)

At each time k, solve

$$\min_{\alpha, u^s, y^s} \left\| \begin{bmatrix} \bar{u} \\ \bar{y} \end{bmatrix} - \begin{bmatrix} \mathbb{1} \otimes u^s \\ \mathbb{1} \otimes y^s \end{bmatrix} \right\|_P^2 + \|y^s - y^r\|_S^2$$

with $u^d = u_{[k-N,k-1]}$, $y^d = y_{[k-N,k-1]}$ *and*

Willems' Lemma | *Initial conditions* | *Terminal constraint*

$$\begin{cases} \begin{bmatrix} \bar{u} \\ \bar{y} \end{bmatrix} = \begin{bmatrix} H_L(u^d) \\ H_L(y^d) \end{bmatrix} \alpha \\ \sum_i \alpha_i = 1 \end{cases} \quad \begin{bmatrix} \bar{u}_{[-n,-1]} \\ \bar{y}_{[-n,-1]} \end{bmatrix} = \begin{bmatrix} u_{[k-n,k-1]} \\ y_{[k-n,k-1]} \end{bmatrix} \quad \begin{bmatrix} \bar{u}_{[N_p-n,N_p]} \\ \bar{y}_{[N_p-n,N_p]} \end{bmatrix} = \begin{bmatrix} \mathbb{1} \otimes u^s \\ \mathbb{1} \otimes y^s \end{bmatrix}$$

Apply $u_k = \bar{u}_0(k)$.

Data-Driven Model Predictive Control

Problem (Stabilization of $\mathcal{S}$ close to y^r, see Berberich et al. 2022)

At each time k, solve

$$\min_{\alpha, w^s} \|\bar{w} - \mathbb{1} \otimes w^s\|_P^2 + \|y^s - y^r\|_S^2$$

with $\begin{bmatrix} \bar{w} \\ 1 \end{bmatrix} = \begin{bmatrix} H_L(w^d) \\ \mathbb{1}^\top \end{bmatrix} \alpha$ *and* $\begin{cases} H_{L,init}(w^d)\alpha = w_{[k-n,k-1]} \\ H_{L,end}(w^d)\alpha = \mathbb{1} \otimes w^s \end{cases}$

Apply $u_k = \bar{u}_0(k)$.

Notation: $w = \begin{bmatrix} u \\ y \end{bmatrix}$, $H_L(w^d) = \begin{bmatrix} H_L(u^d) \\ H_L(y^d) \end{bmatrix}$

Data-Driven Model Predictive Control: Two Datasets

Problem

At each time k, solve

$$\min_{\alpha, w^s} \|\bar{w} - \mathbb{1} \otimes w^s\|_P^2 + \|y^s - y^r\|_S^2$$

with

$$\begin{bmatrix} H_L(w_\ell^d) & H_L(w_a^d) \\ \mathbb{0}^\top & \mathbb{1}^\top \\ \mathbb{1}^\top & \mathbb{0}^\top \end{bmatrix} \alpha = \begin{bmatrix} \bar{w} \\ 1 \\ 0 \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} H_L(w_\ell^d) & H_L(w_a^d) \end{bmatrix}_{init} \alpha = w_{[k-n,k-1]} \tag{2}$$

$$\begin{bmatrix} H_L(w_\ell^d) & H_L(w_a^d) \end{bmatrix}_{end} \alpha = \mathbb{1} \otimes w^s \tag{3}$$