Design of Multi-Model Linear Inferential Sensors with SVM-based Switching Logic

Martin Mojto[1], Karol Ľubušký[2], Miroslav Fikar[1] and Radoslav Paulen[1]

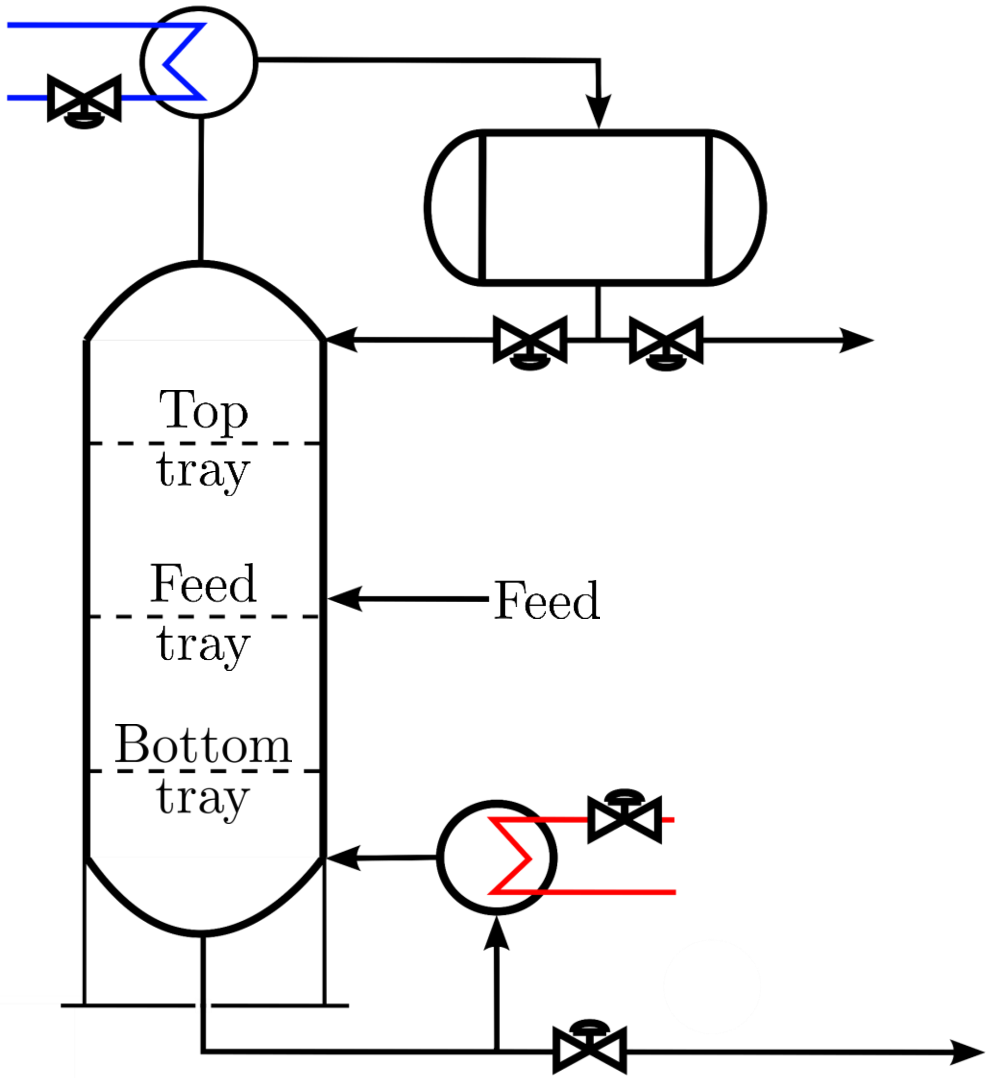
[1]Slovak University of Technology in Bratislava, Slovakia

[2]Slovnaft, a.s., Bratislava, Slovakia

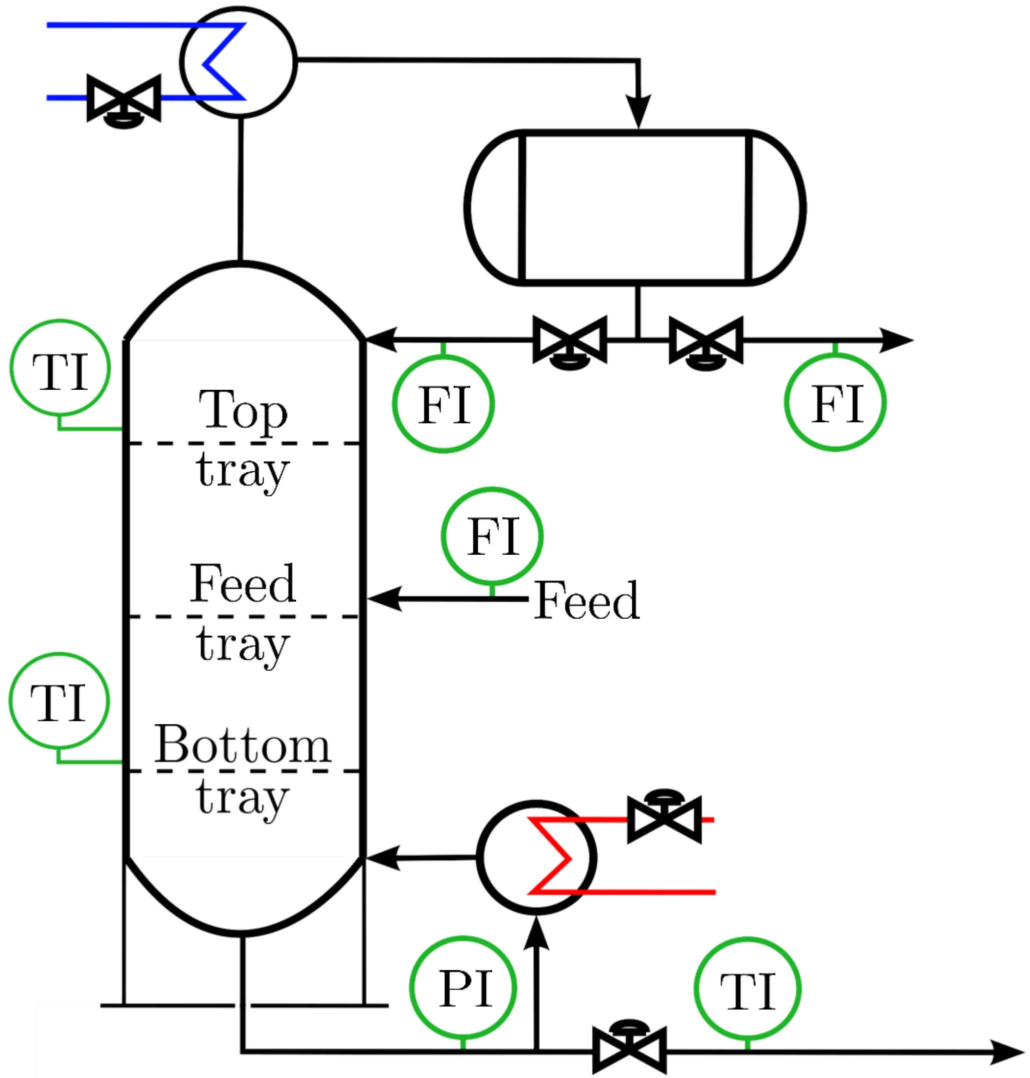
IFAC World Congress

July 9–14, 2023

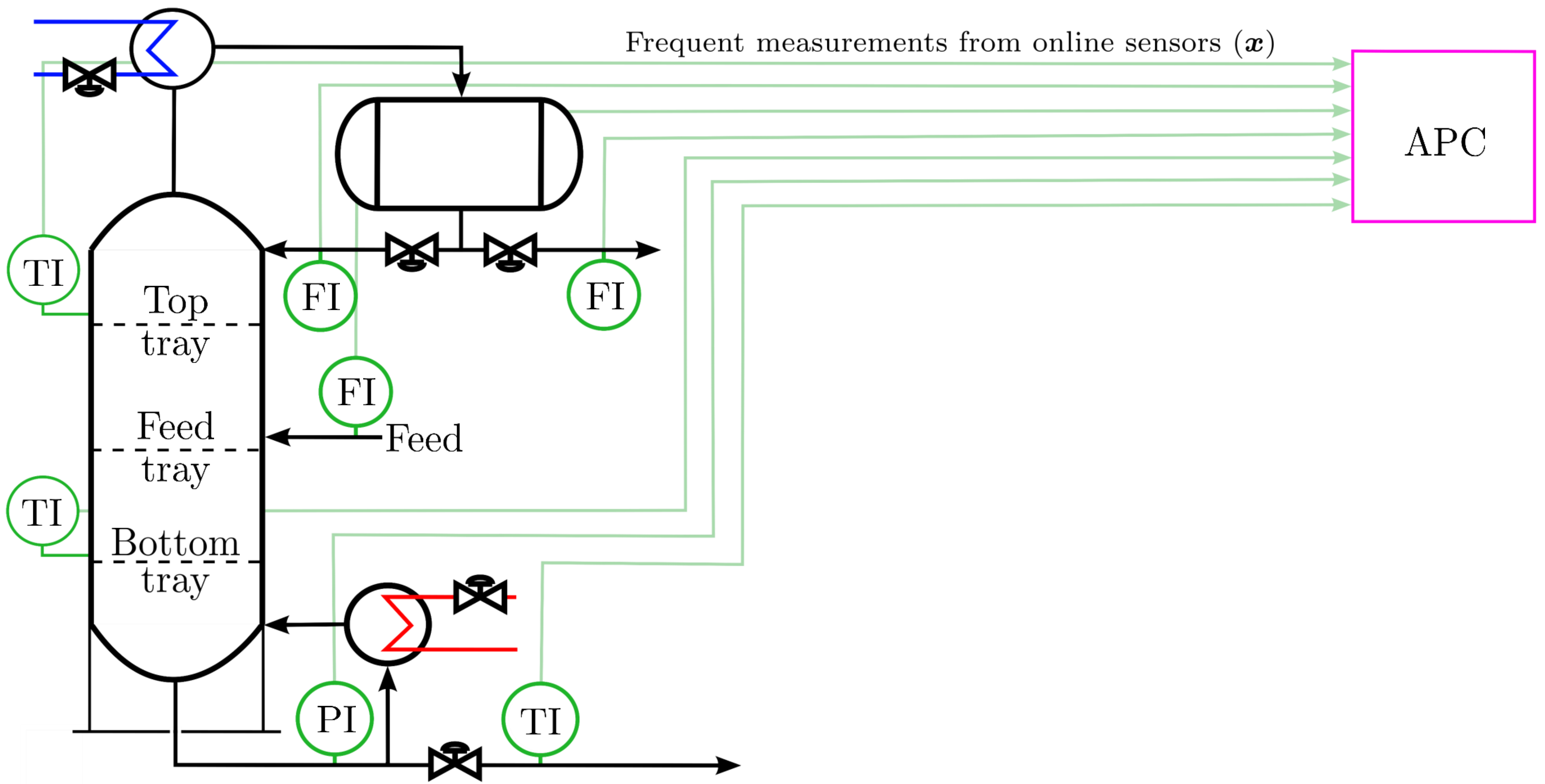
Motivation: Inferential Sensors



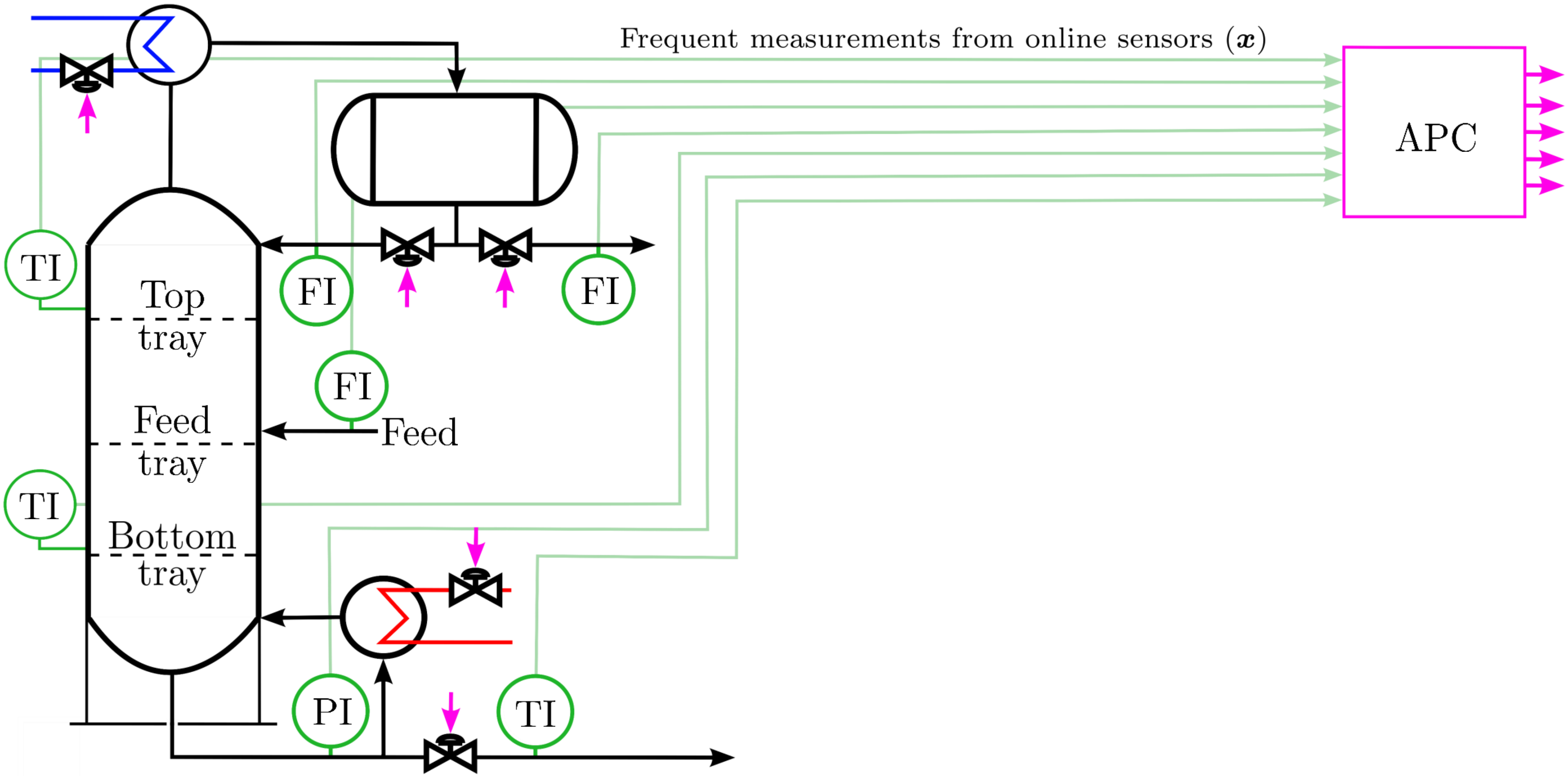
Motivation: Inferential Sensors



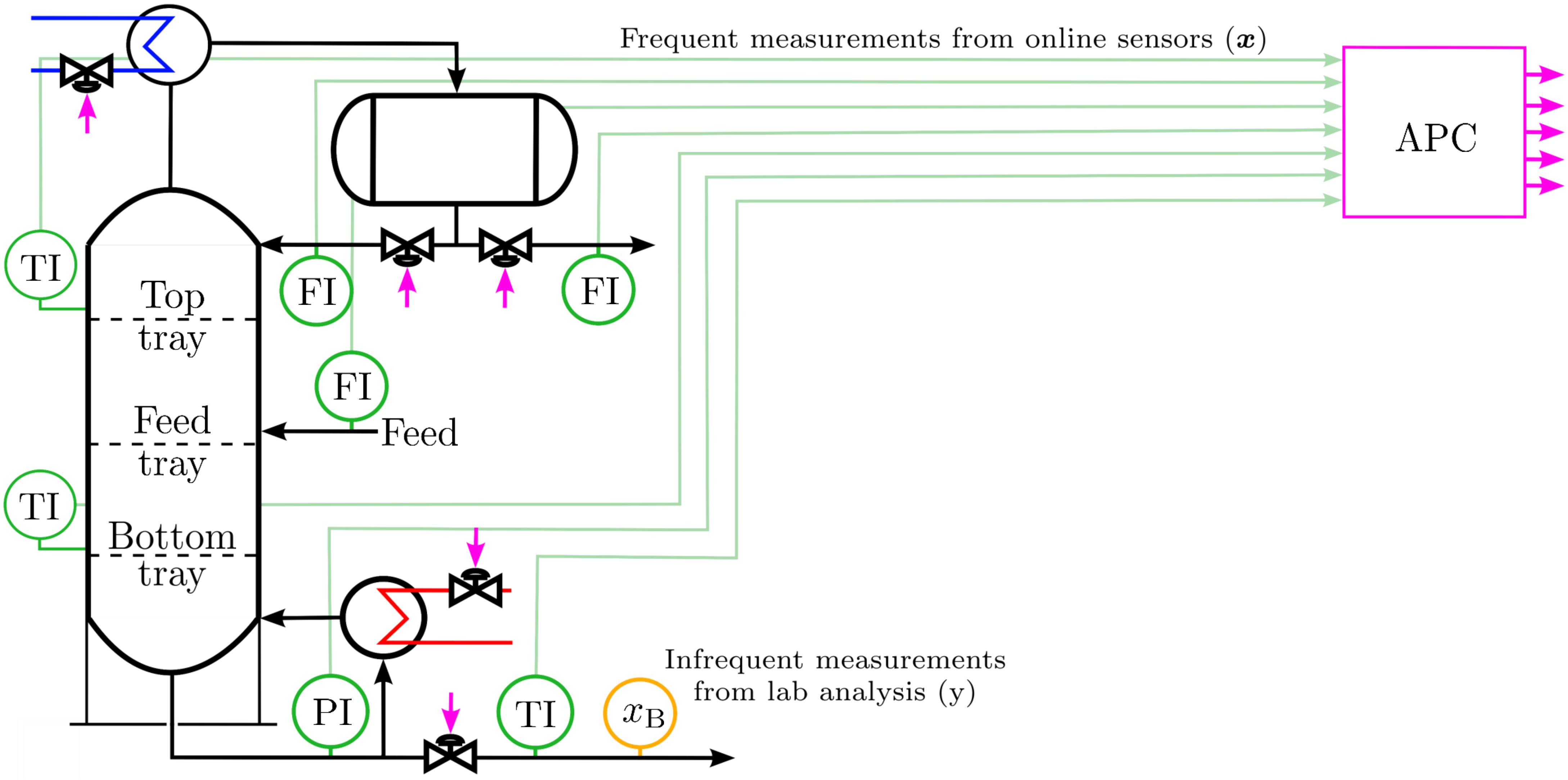
Motivation: Inferential Sensors



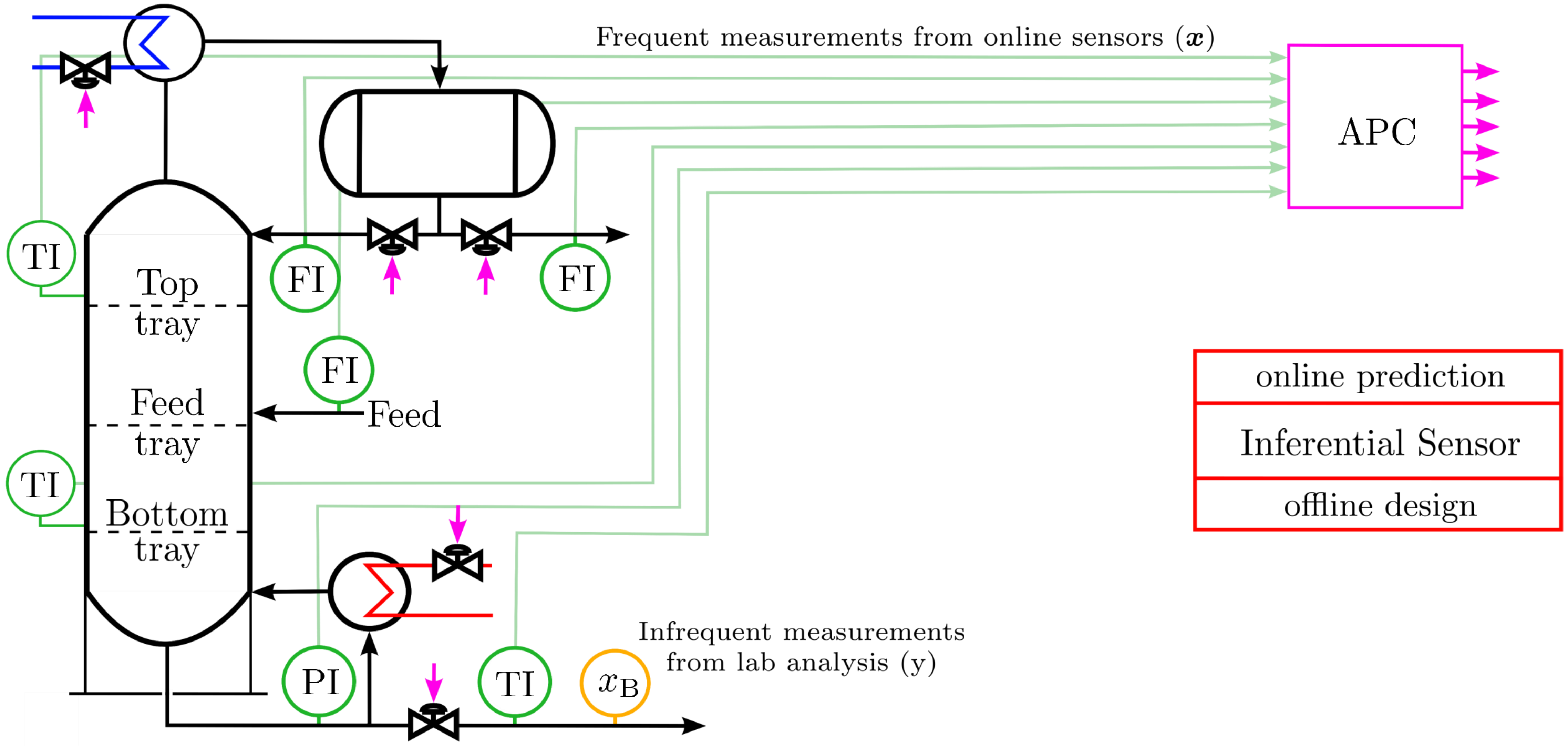
Motivation: Inferential Sensors



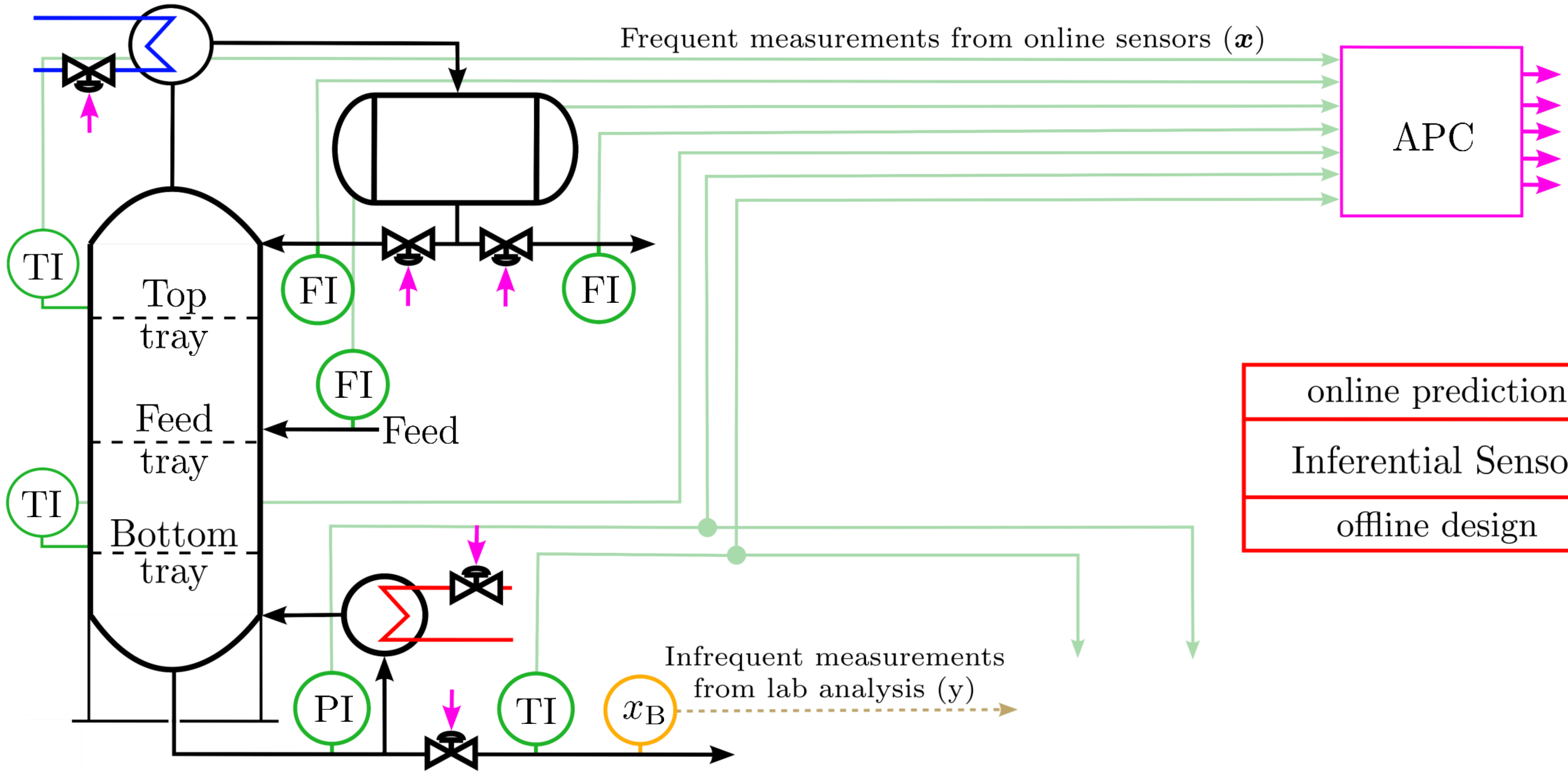
Motivation: Inferential Sensors



Motivation: Inferential Sensors

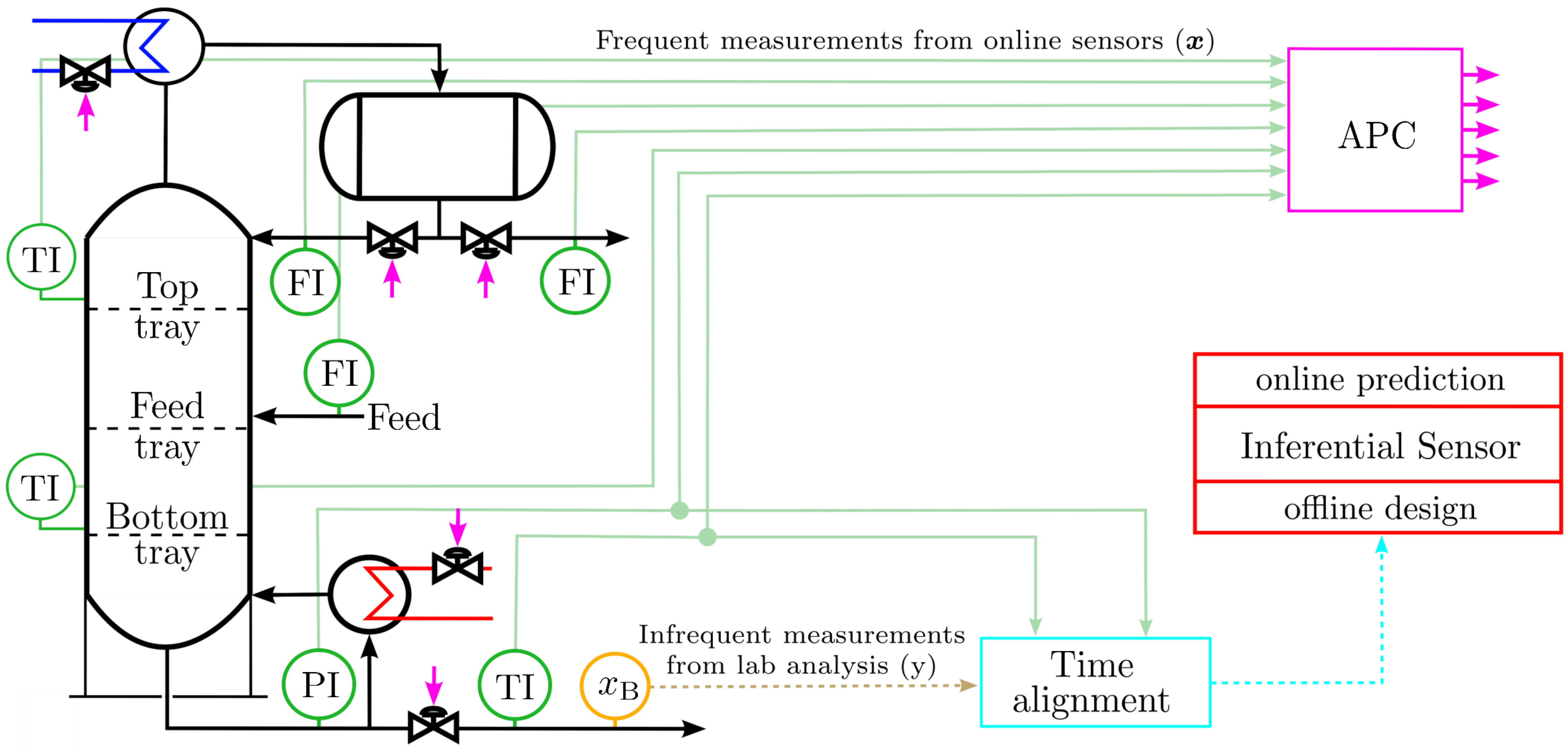


Motivation: Inferential Sensors

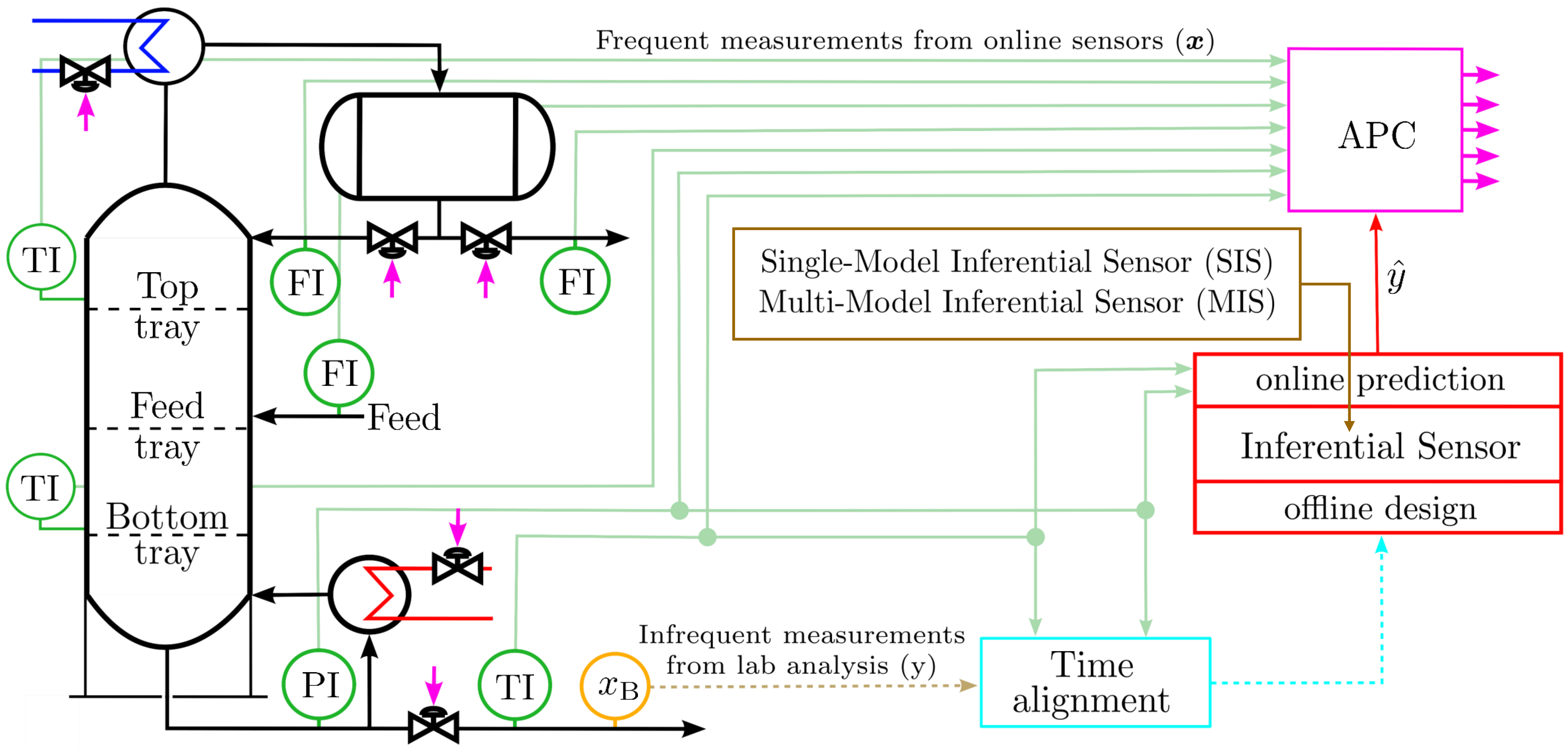


online prediction
Inferential Sensor
offline design

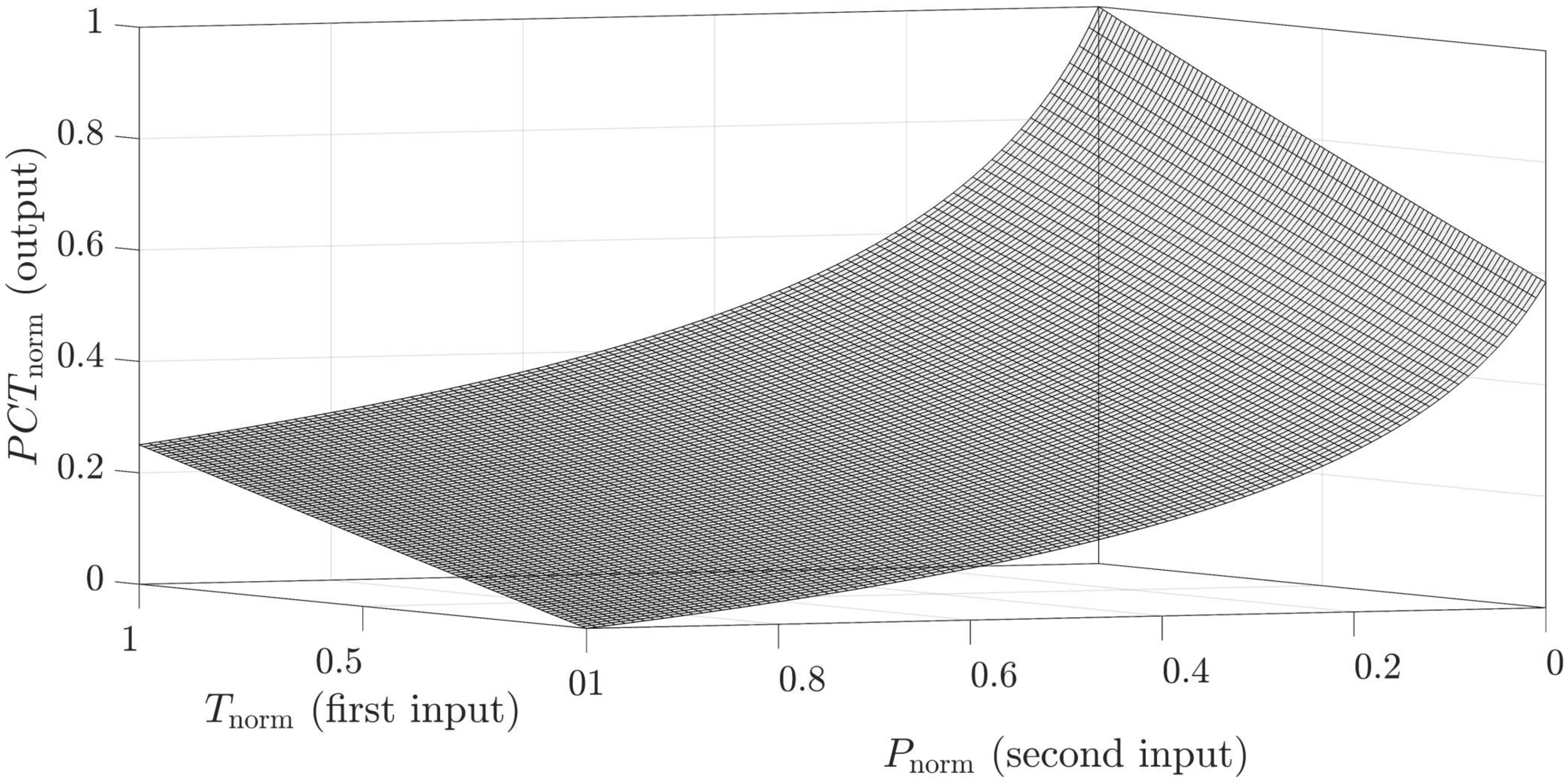
Motivation: Inferential Sensors



Motivation: Inferential Sensors

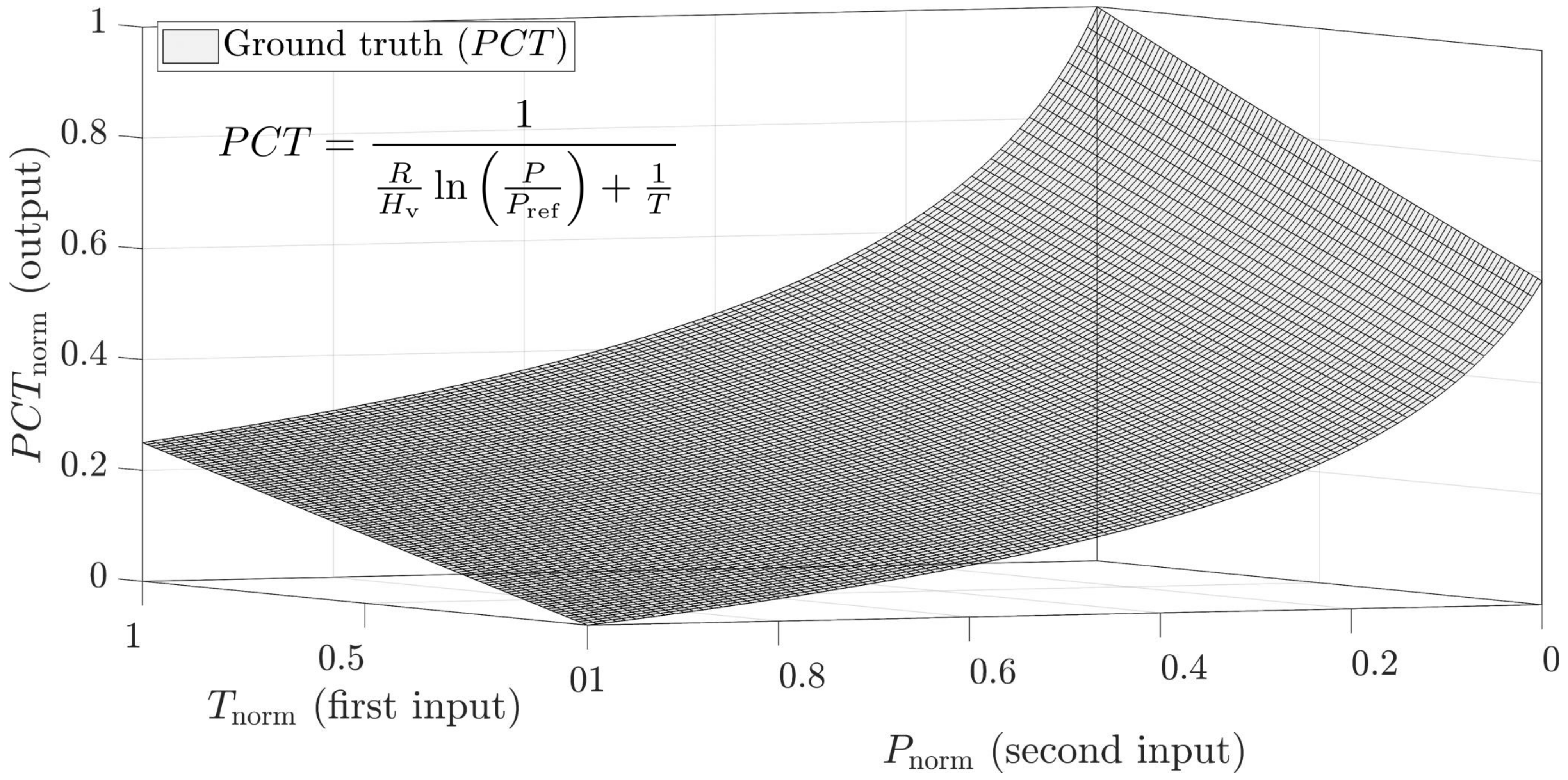


Motivation: MIS Design



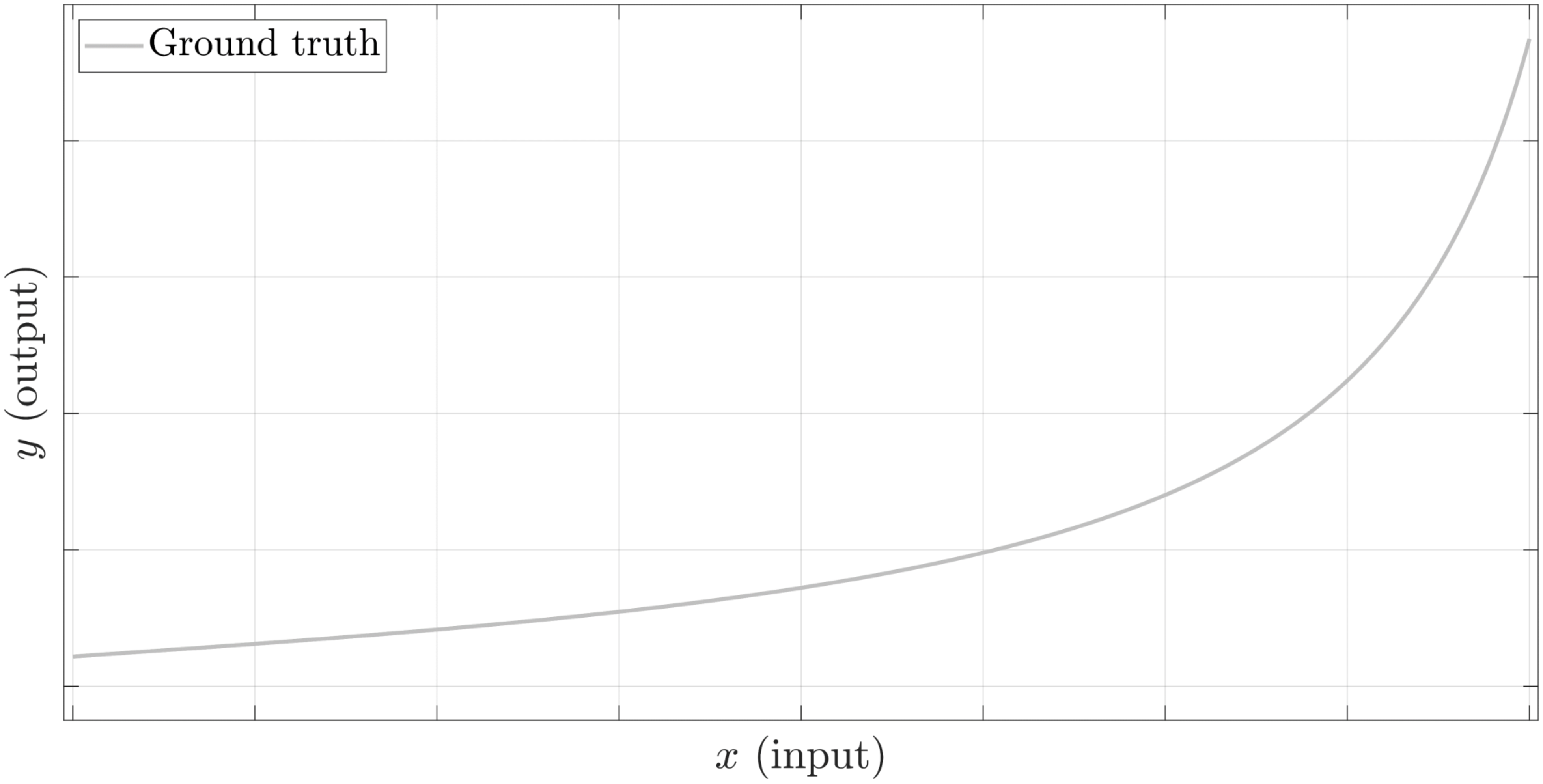
Motivation: MIS Design

Case study: 3D

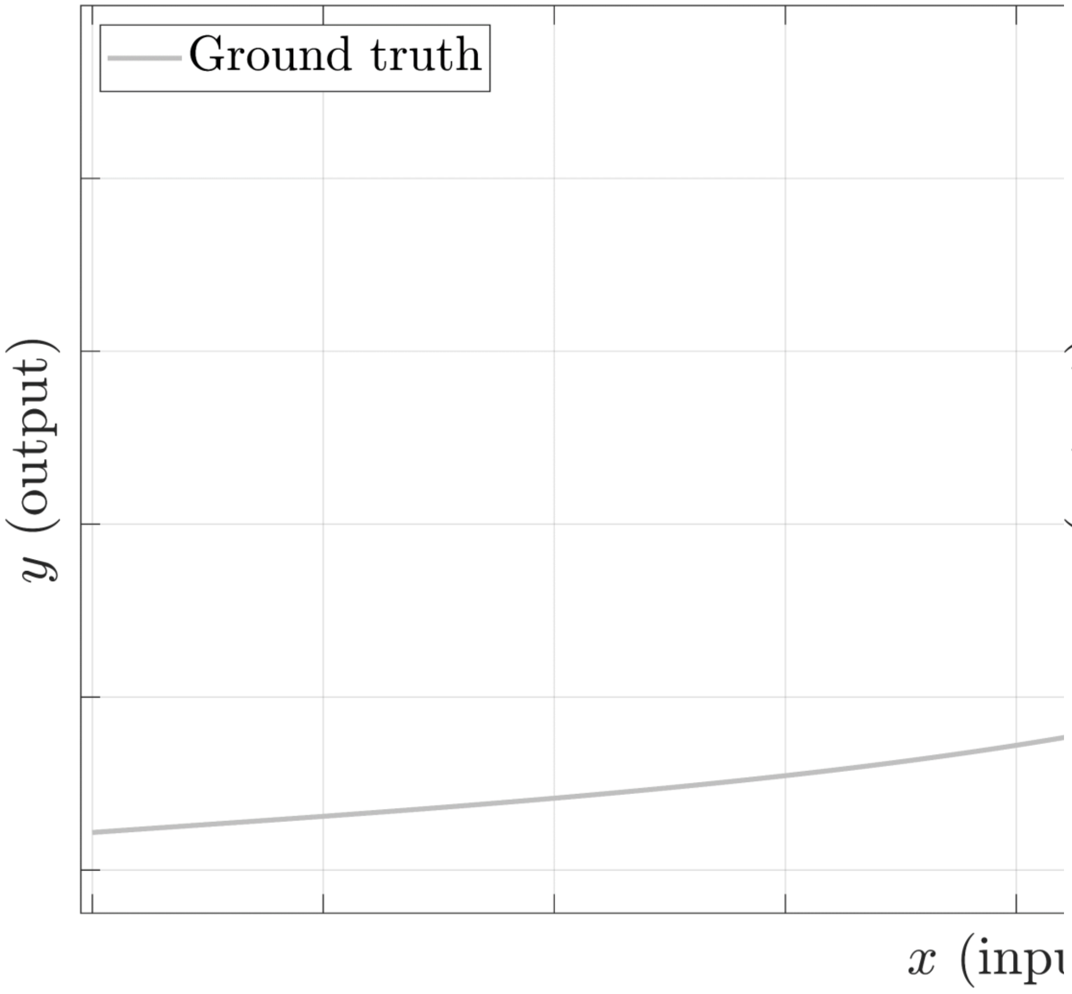


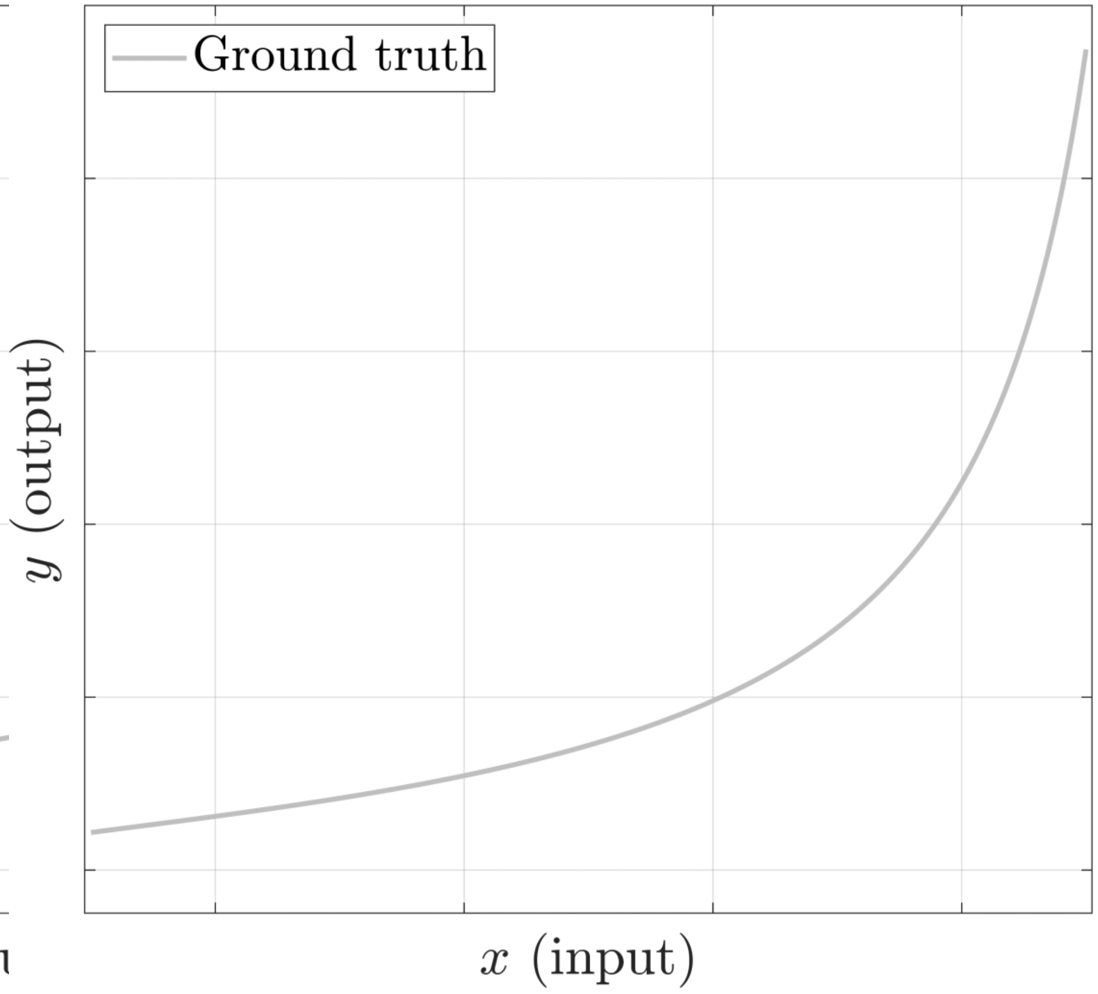
Motivation: MIS Design

Case study: 3D $\longrightarrow$ 2D

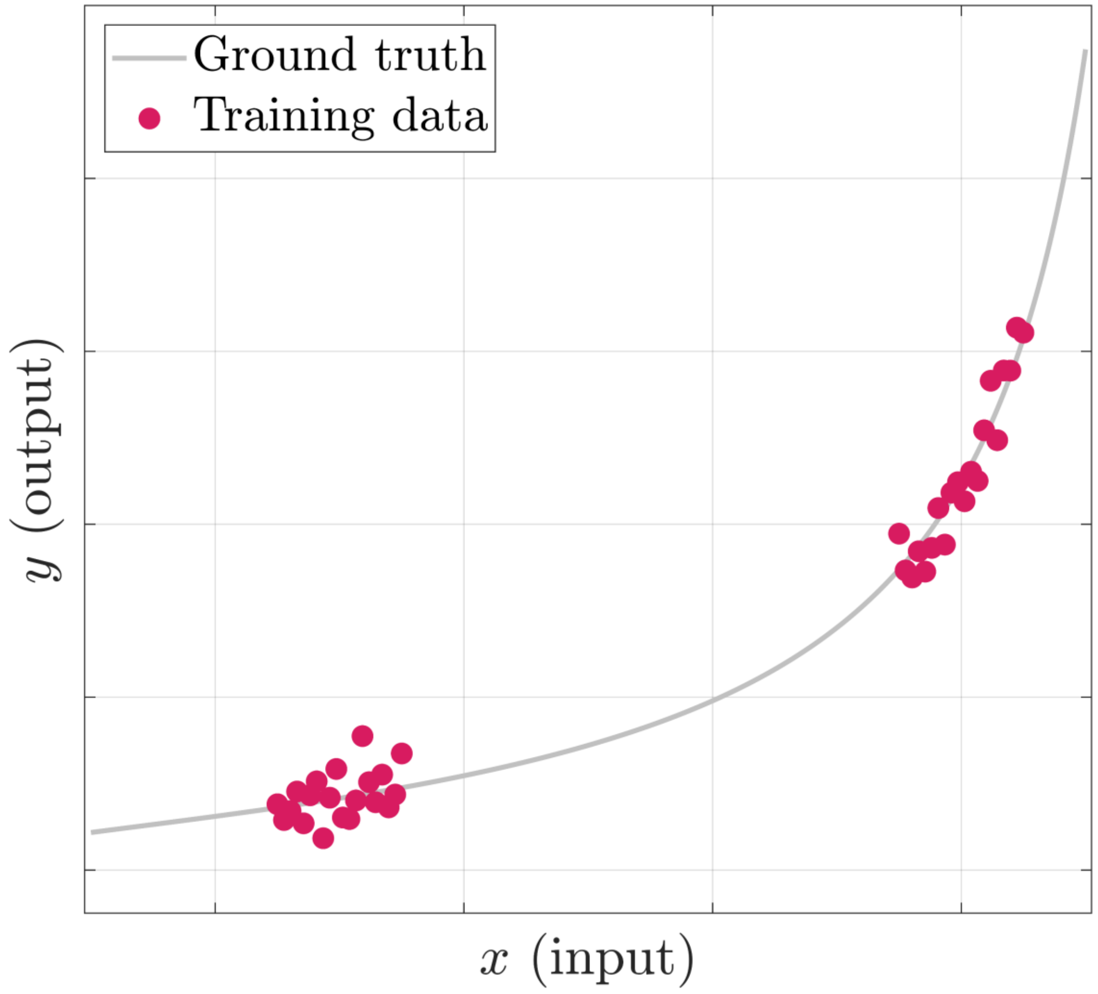


State-of-the-Art Approach



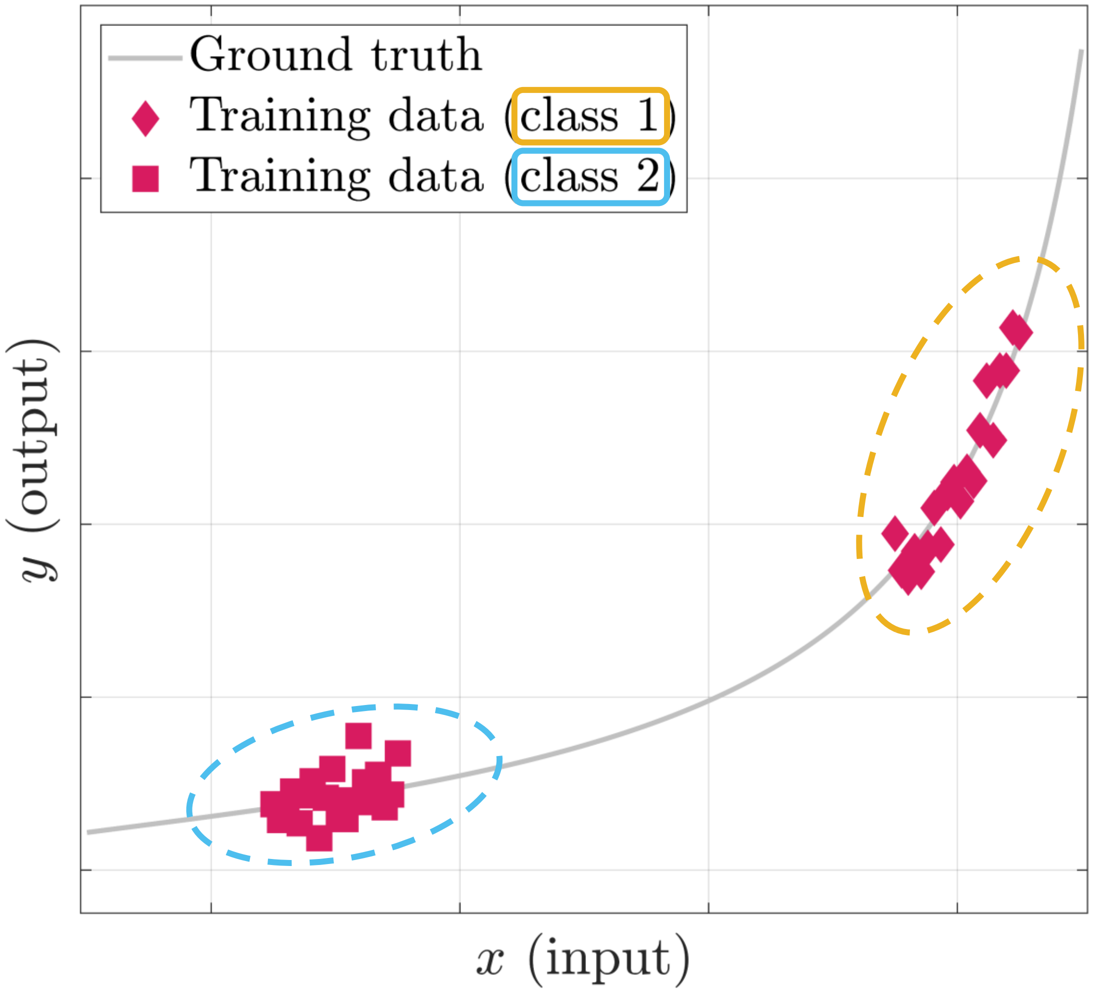


State-of-the-Art Approach



State-of-the-Art Approach

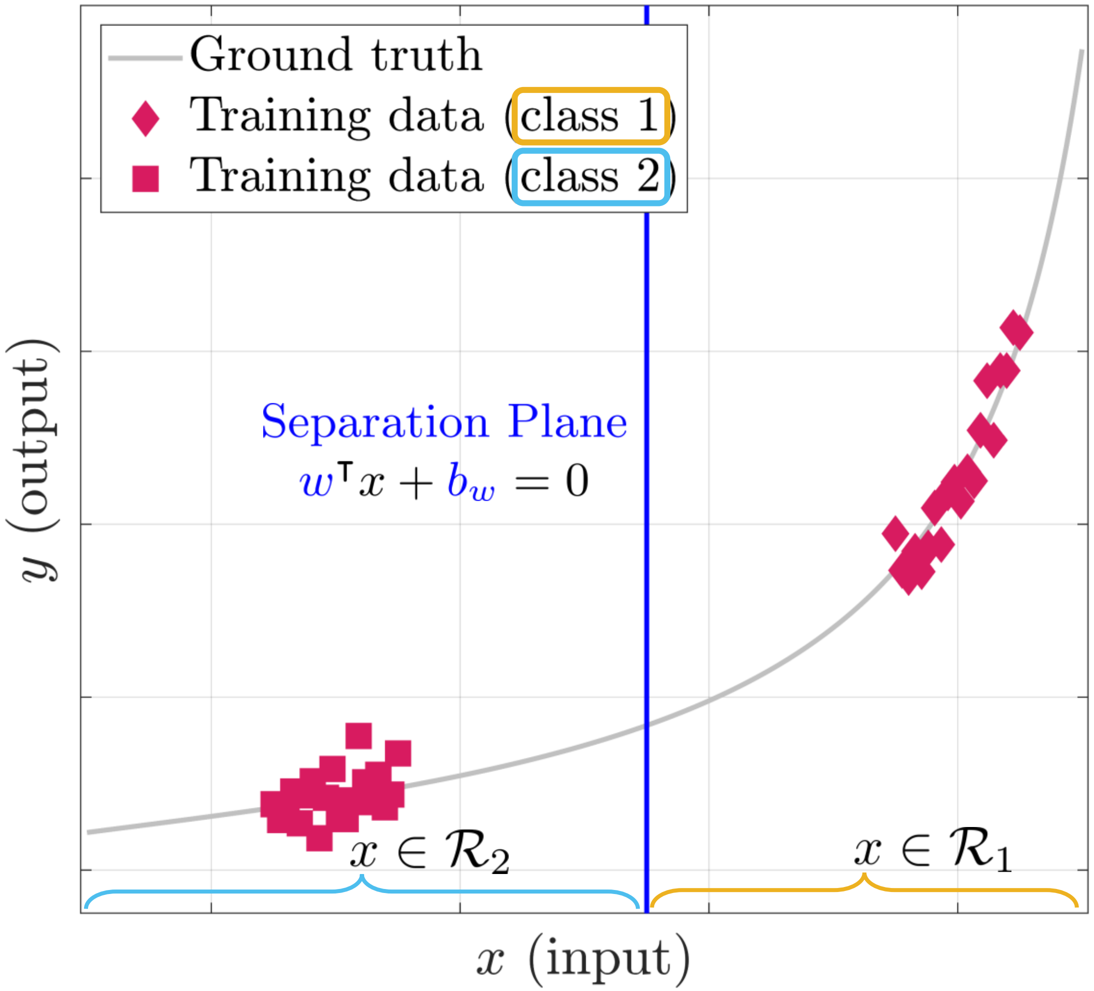
1. A priori labeling of the training dataset



State-of-the-Art Approach

1. A priori labeling of the training dataset

2. Data classification
   - Method: Support Vector Machine (SVM)
   - Linear separation plane: $w^\intercal x + b_w = 0$



State-of-the-Art Approach

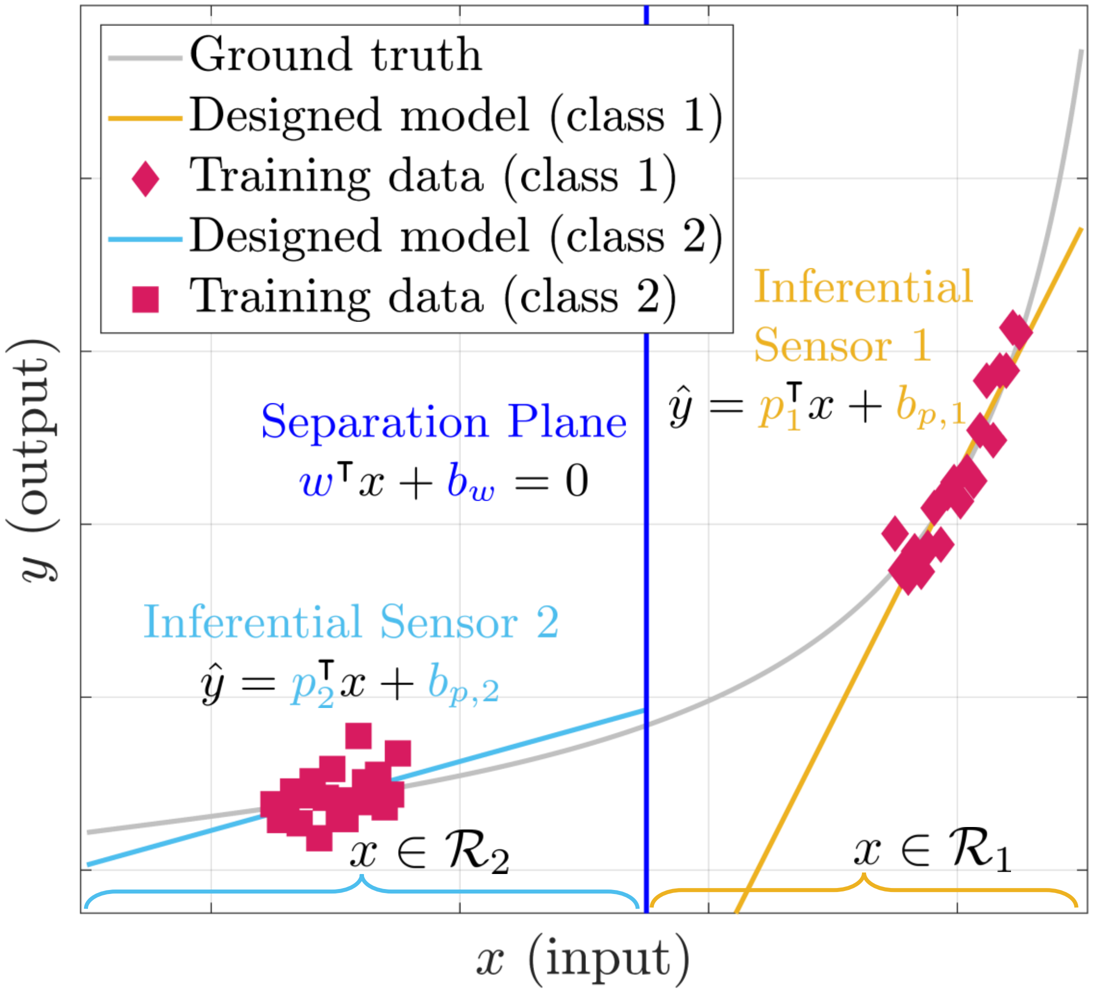
1. A priori labeling of the training dataset

2. Data classification
   - Method: Support Vector Machine (SVM)
   - Linear separation plane: $w^\intercal x + b_w = 0$

3. Individual sensor training

$$\hat{y} = \begin{cases} p_1^\intercal x + b_{p,1}, & \text{if } x \in \mathcal{R}_1 \\ p_2^\intercal x + b_{p,2}, & \text{if } x \in \mathcal{R}_2 \end{cases}$$

calculation of model parameters
$p_1, b_{p,1}, p_2, b_{p,2}$

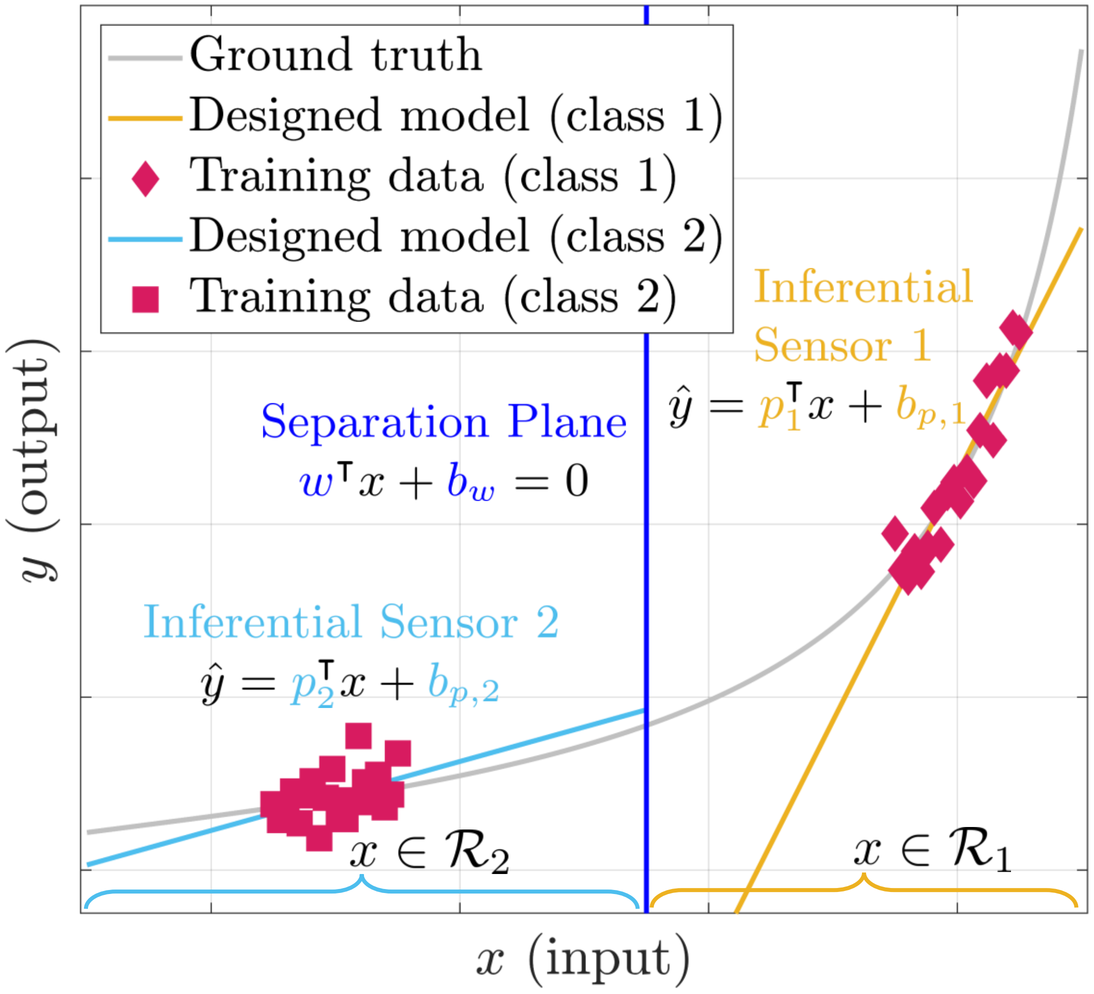


1. A priori labeling of the training dataset
2. Data classification
   - Method: Support Vector Machine (SVM)
   - Linear separation plane: $w^\intercal x + b_w = 0$
3. Individual sensor training

$$\hat{y} = \begin{cases} p_1^\intercal x + b_{p,1}, & \text{if } x \in \mathcal{R}_1 \\ p_2^\intercal x + b_{p,2}, & \text{if } x \in \mathcal{R}_2 \end{cases}$$

calculation of model parameters $p_1, b_{p,1}, p_2, b_{p,2}$

MIS-SotA



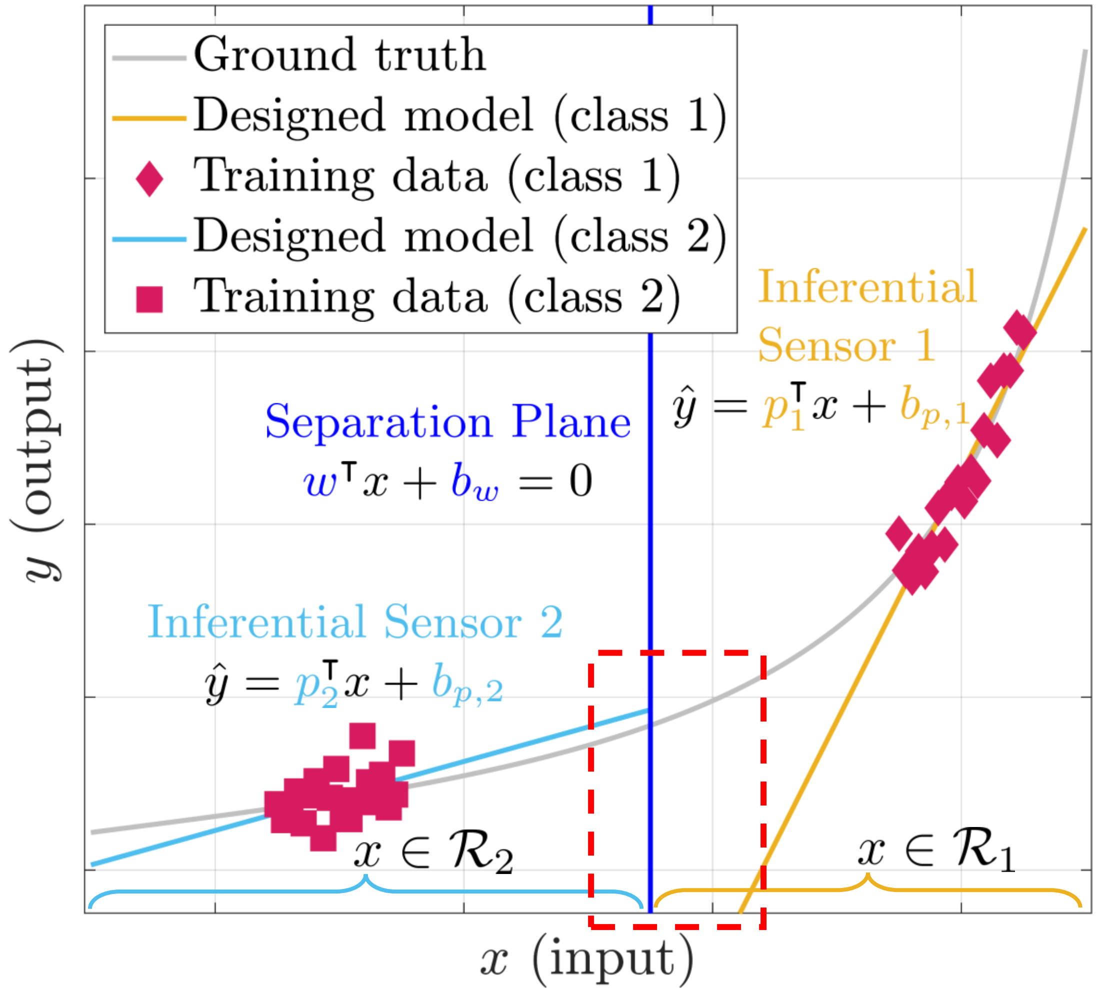
State-of-the-Art Approach

1. A priori labeling of the training dataset
2. Data classification
   - Method: Support Vector Machine (SVM)
   - Linear separation plane: $w^\intercal x + b_w = 0$
3. Individual sensor training

$$\hat{y} = \begin{cases} p_1^\intercal x + b_{p,1}, & \text{if } x \in \mathcal{R}_1 \\ p_2^\intercal x + b_{p,2}, & \text{if } x \in \mathcal{R}_2 \end{cases}$$

calculation of model parameters $p_1, b_{p,1}, p_2, b_{p,2}$

MIS-SotA



Main Challenges:

!! Discontinuity of MIS models

State-of-the-Art Approach

1. A priori labeling of the training dataset
2. Data classification
   - Method: Support Vector Machine (SVM)
   - Linear separation plane: $w^\intercal x + b_w = 0$
3. Individual sensor training

$$\hat{y} = \begin{cases} p_1^\intercal x + b_{p,1}, & \text{if } x \in \mathcal{R}_1 \\ p_2^\intercal x + b_{p,2}, & \text{if } x \in \mathcal{R}_2 \end{cases}$$

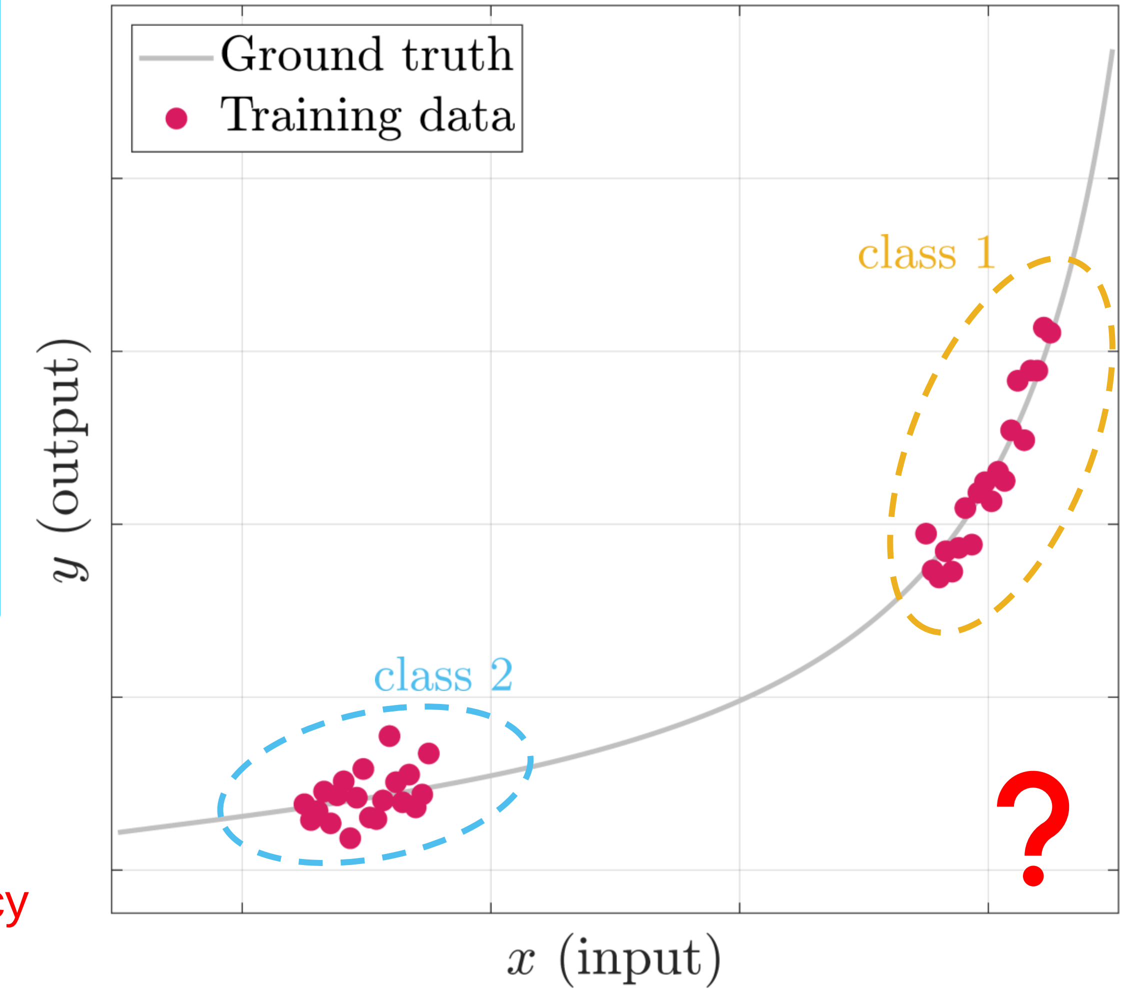
calculation of model parameters $p_1, b_{p,1}, p_2, b_{p,2}$

MIS-SotA

Main Challenges:

!! Discontinuity of MIS models

!! Unknown impact of a priori labeling on the MIS accuracy



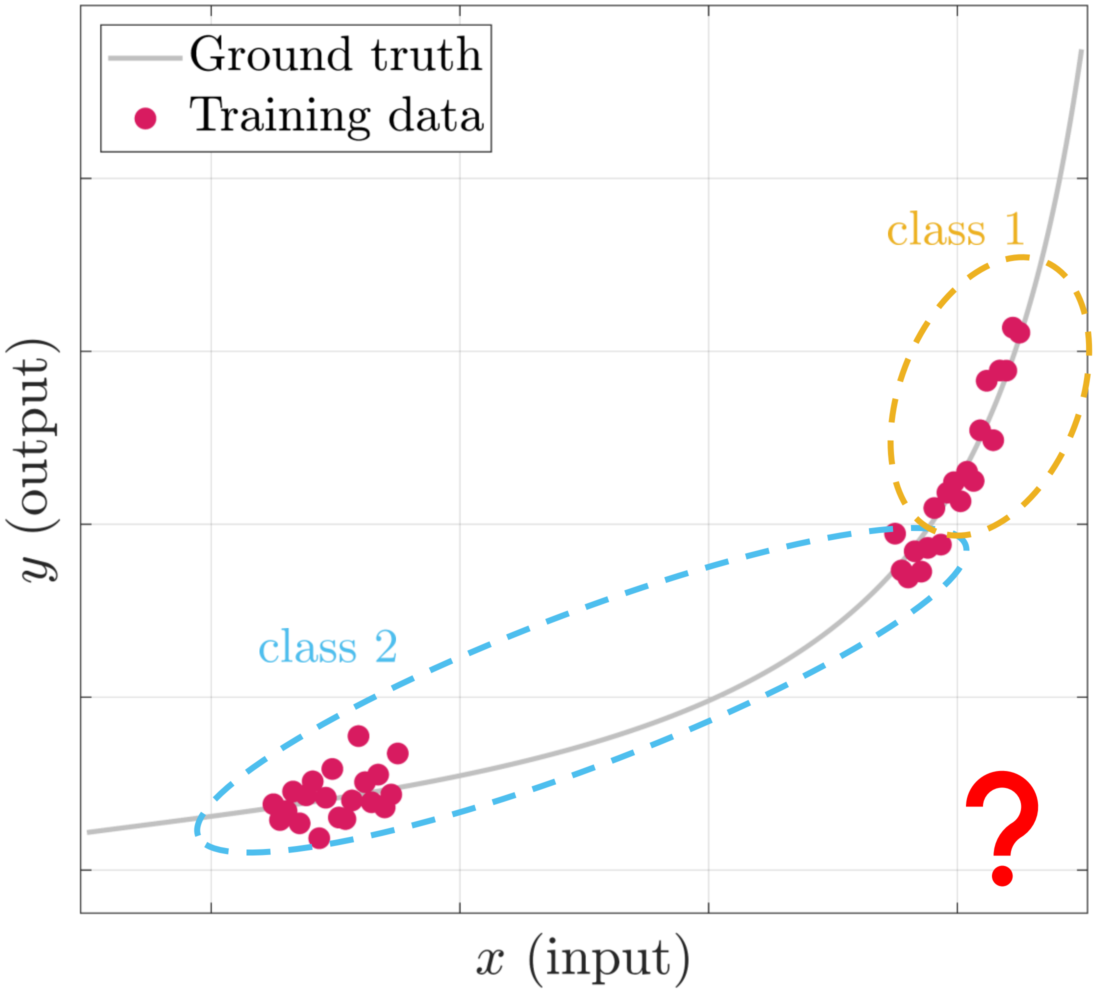
State-of-the-Art Approach

1. A priori labeling of the training dataset
2. Data classification
   - Method: Support Vector Machine (SVM)
   - Linear separation plane: $w^\intercal x + b_w = 0$
3. Individual sensor training

$$\hat{y} = \begin{cases} p_1^\intercal x + b_{p,1}, & \text{if } x \in \mathcal{R}_1 \\ p_2^\intercal x + b_{p,2}, & \text{if } x \in \mathcal{R}_2 \end{cases}$$

calculation of model parameters $p_1, b_{p,1}, p_2, b_{p,2}$

MIS-SotA



Main Challenges:

!! Discontinuity of MIS models

!! Unknown impact of a priori labeling on the MIS accuracy

State-of-the-Art Approach Enhancement

1. A priori labeling of the training dataset
2. Data classification & sensor training

$$\min_{\substack{e\geq 0, w, b_w \\ p_1, b_{p,1}, p_2, b_{p,2}}} \sum_{i=1}^{n} \left[z_i \underbrace{\mathrm{SSE}_{1,i}}_{\text{sum of squared errors}} + (1-z_i)\mathrm{SSE}_{2,i} \right] + \alpha \|w\|_2^2 + \beta \underbrace{\|e\|_1}_{\text{slack variables}}$$

$$\text{s.t.} \quad (2z_i - 1)\left(w^\intercal x_i + b_w\right) \geq 1 - e_i$$

$$\forall i = \{1, 2, \ldots, n\}$$

MIS-SotA (enhanced)

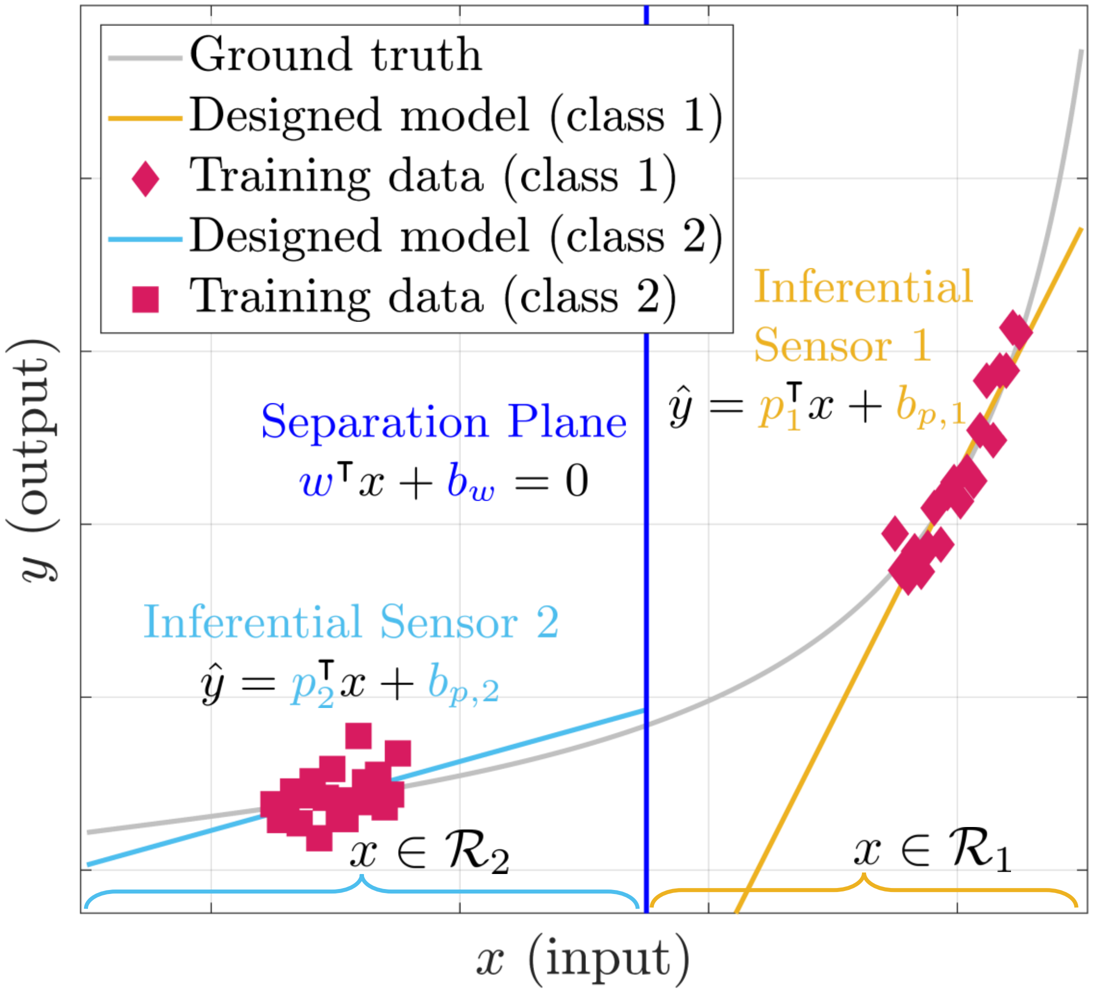
Binary variables

$$z_i = \begin{cases} 1, & \text{if } x_i \in \mathcal{R}_1 \\ 0, & \text{if } x_i \in \mathcal{R}_2 \end{cases}$$

number of measurements in a class

$$n = n_1 + n_2$$

$$\forall i = \{1, 2, \ldots, n\}$$



1. A priori labeling of the training dataset

2. Data classification & sensor training

$$\min_{\substack{e\geq 0, w, b_w\\ p_1, b_{p,1}, p_2, b_{p,2}}} \sum_{i=1}^{n} \left[z_i \mathrm{SSE}_{1,i} + (1 - z_i)\mathrm{SSE}_{2,i} \right] + \alpha \|w\|_2^2 + \beta \|e\|_1$$

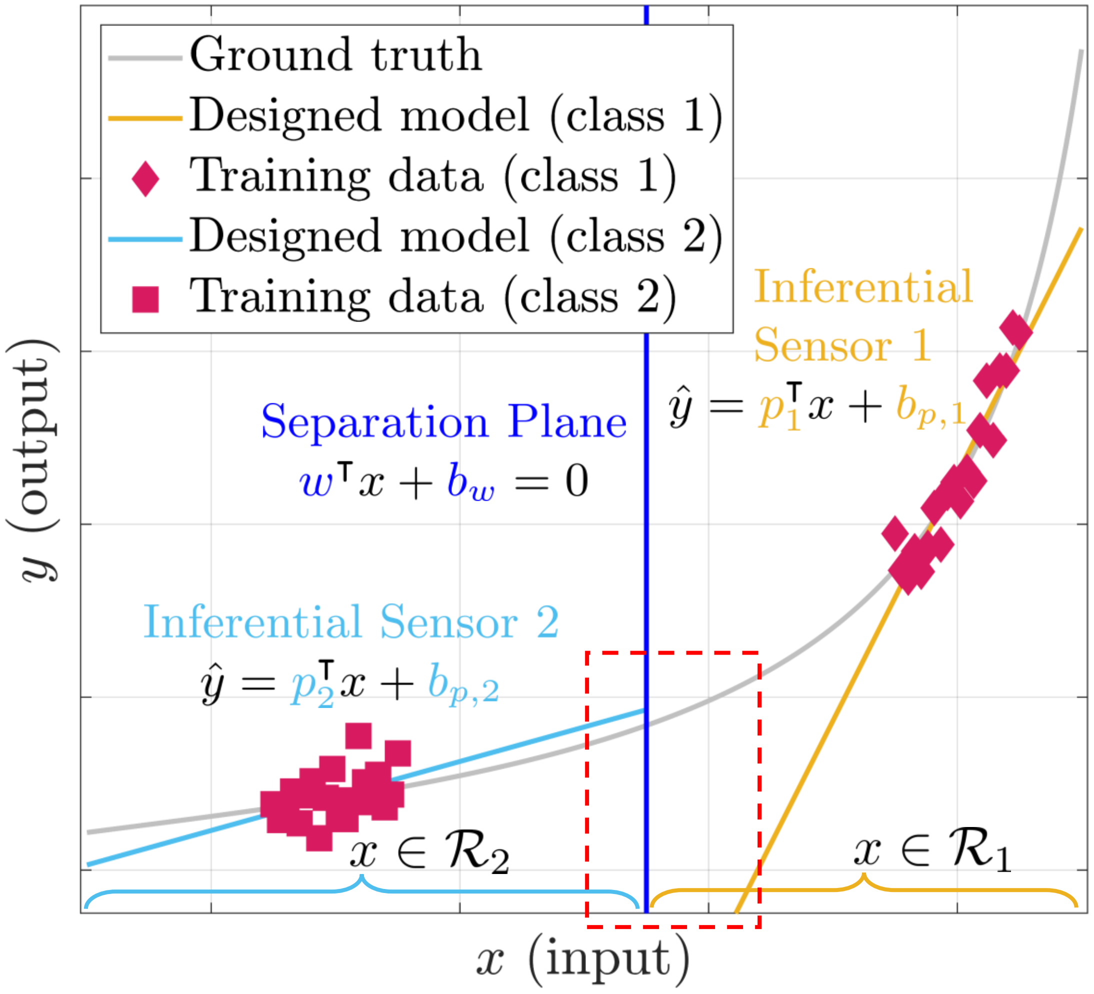
$$\text{s.t.} \quad (2z_i - 1)\left(w^\intercal x_i + b_w\right) \geq 1 - e_i$$

$$\forall i = \{1, 2, \ldots, n\}$$

!! Discontinuity of MIS models

The continuity constraints:

$$p_1 - p_2 - w = 0, \quad b_{p,1} - b_{p,2} - b_w = 0$$



Design of Continuous Sensor

1. A priori labeling of the training dataset
2. Data classification & sensor training

$$\min_{\substack{e\geq 0, w, b_w\\ p_1, b_{p,1}, p_2, b_{p,2}}} \sum_{i=1}^{n} \left[z_i \mathrm{SSE}_{1,i} + (1 - z_i)\mathrm{SSE}_{2,i} \right] + \alpha \|w\|_2^2 + \beta \|e\|_1$$

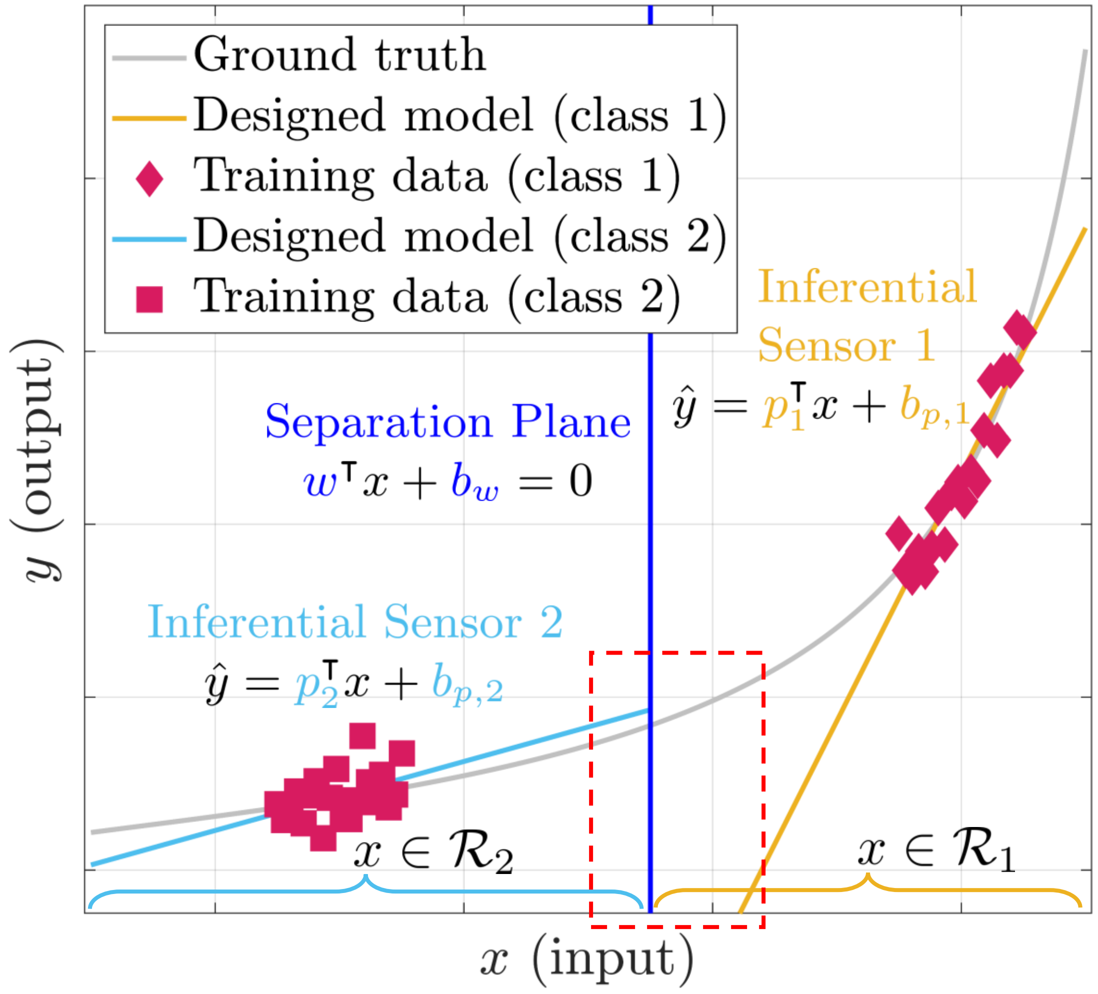
$$\text{s.t.} \quad (2z_i - 1)\left(w^\intercal x_i + b_w\right) \geq 1 - e_i$$

$$\forall i = \{1, 2, \ldots, n\}$$

!! Discontinuity of MIS models

The continuity constraints:

$$p_1 - p_2 - w = 0, \quad b_{p,1} - b_{p,2} - b_w = 0$$



Design of Continuous Sensor

1. A priori labeling of the training dataset

2. Data classification & sensor training

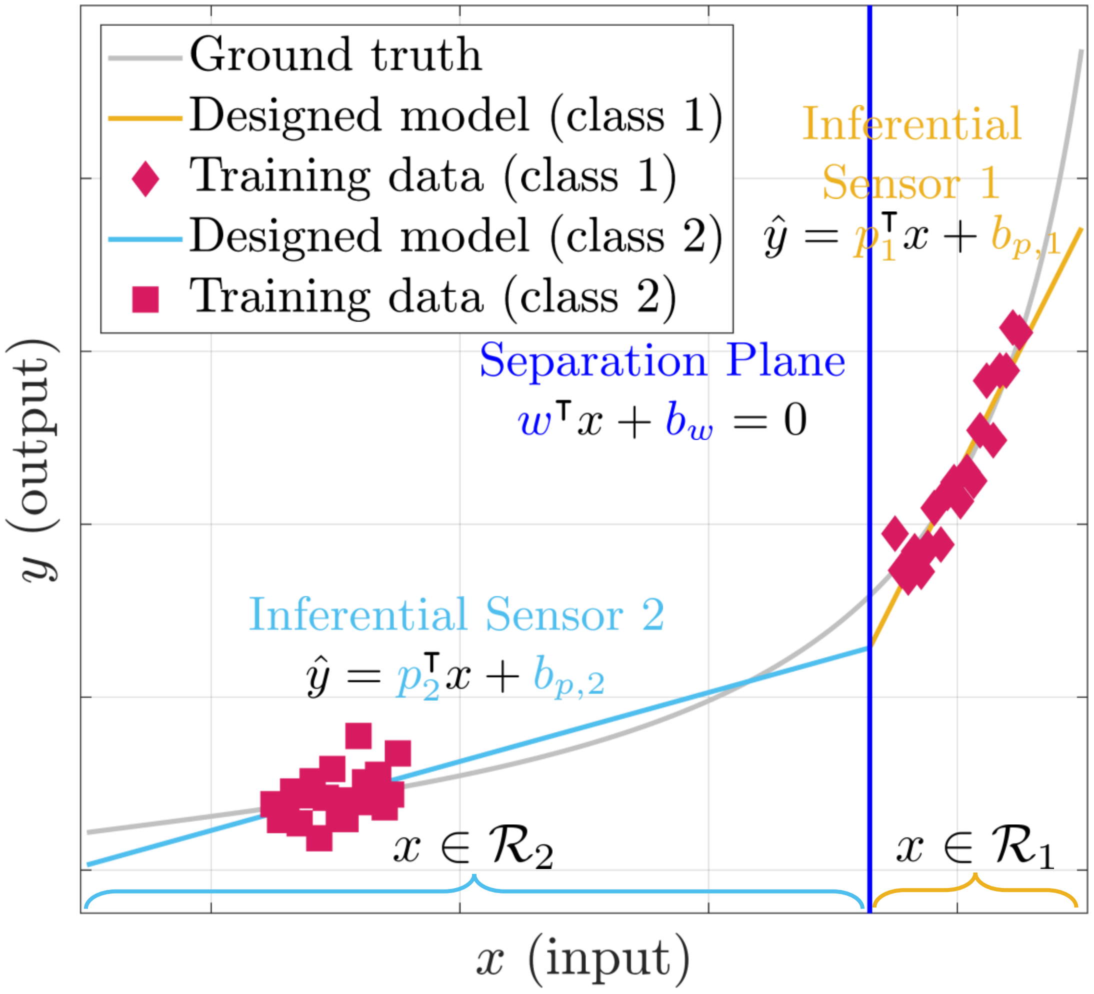
$$\min_{\substack{e\geq 0, w, b_w \\ p_1, b_{p,1}, p_2, b_{p,2}}} \sum_{i=1}^{n} \left[z_i \mathrm{SSE}_{1,i} + (1-z_i)\mathrm{SSE}_{2,i} \right] + \alpha \|w\|_2^2 + \beta \|e\|_1$$

$$\text{s.t.} \quad (2z_i - 1)\left(w^\intercal x_i + b_w\right) \geq 1 - e_i$$

$$p_1 - p_2 - w = 0, \quad b_{p,1} - b_{p,2} - b_w = 0$$

$$\forall i = \{1, 2, \ldots, n\}$$

!! Discontinuity of MIS models



Design of Continuous Sensor

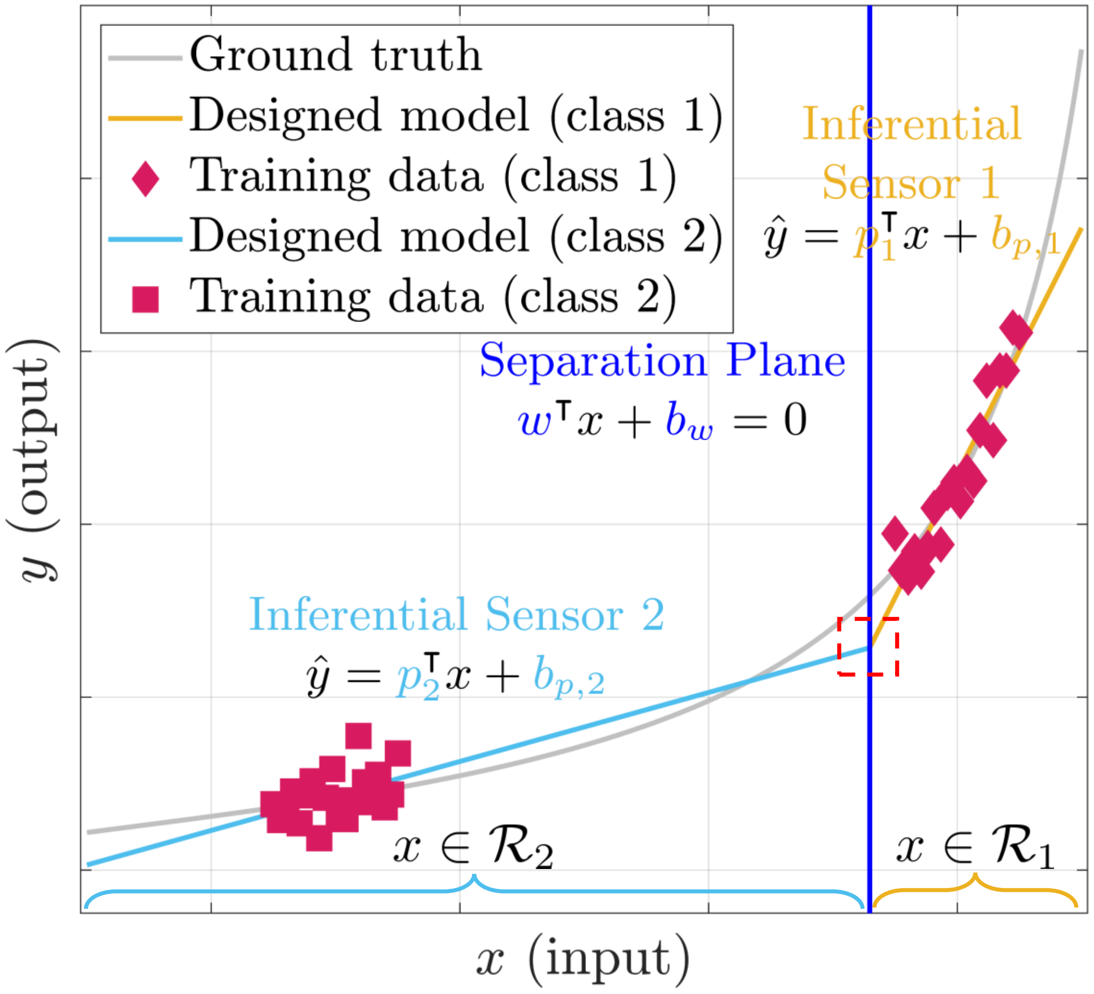
1. A priori labeling of the training dataset

2. Data classification & sensor training

$$\min_{\substack{e\geq 0, w, b_w \\ p_1, b_{p,1}, p_2, b_{p,2}}} \sum_{i=1}^{n} \left[z_i \mathrm{SSE}_{1,i} + (1 - z_i)\mathrm{SSE}_{2,i} \right] + \alpha \|w\|_2^2 + \beta \|e\|_1$$

$$\begin{aligned} \text{s.t.} \quad & (2z_i - 1)\left(w^\intercal x_i + b_w\right) \geq 1 - e_i \\ & p_1 - p_2 - w = 0, \quad b_{p,1} - b_{p,2} - b_w = 0 \\ & \forall i = \{1, 2, \ldots, n\} \end{aligned}$$

✓ Continuity of MIS models



1. A priori labeling of the training dataset
2. Data classification & sensor training

$$\min_{\substack{e\geq 0, w, b_w \\ p_1, b_{p,1}, p_2, b_{p,2}}} \sum_{i=1}^{n} \left[z_i \mathrm{SSE}_{1,i} + (1 - z_i)\mathrm{SSE}_{2,i} \right] + \alpha \|w\|_2^2 + \beta \|e\|_1$$

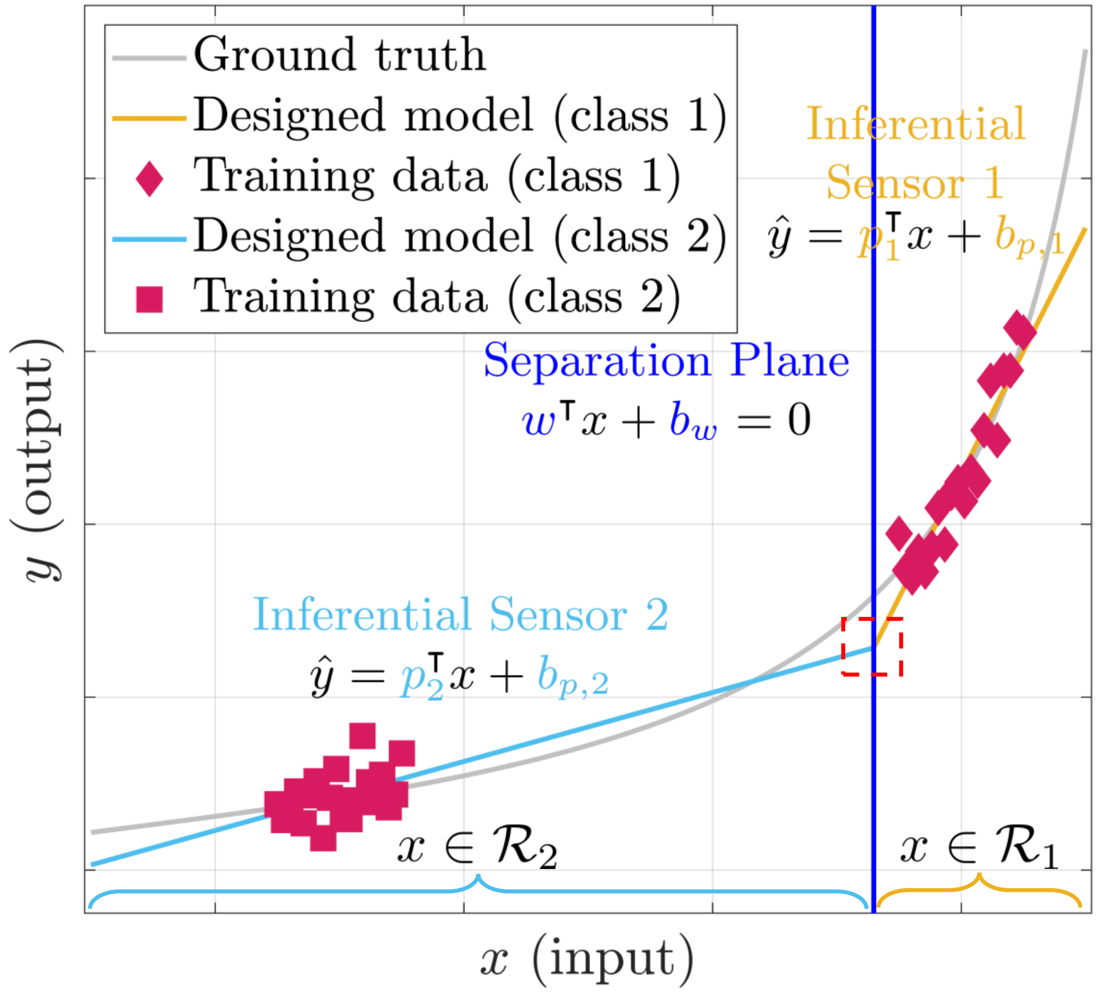
$$\text{s.t.} \quad (2z_i - 1)\left(w^\intercal x_i + b_w\right) \geq 1 - e_i$$

$$p_1 - p_2 - w = 0, \quad b_{p,1} - b_{p,2} - b_w = 0$$

$$\forall i = \{1, 2, \ldots, n\}$$

MIS-con

✓ Continuity of MIS models



1. A priori labeling of the training dataset

2. Data classification & sensor training

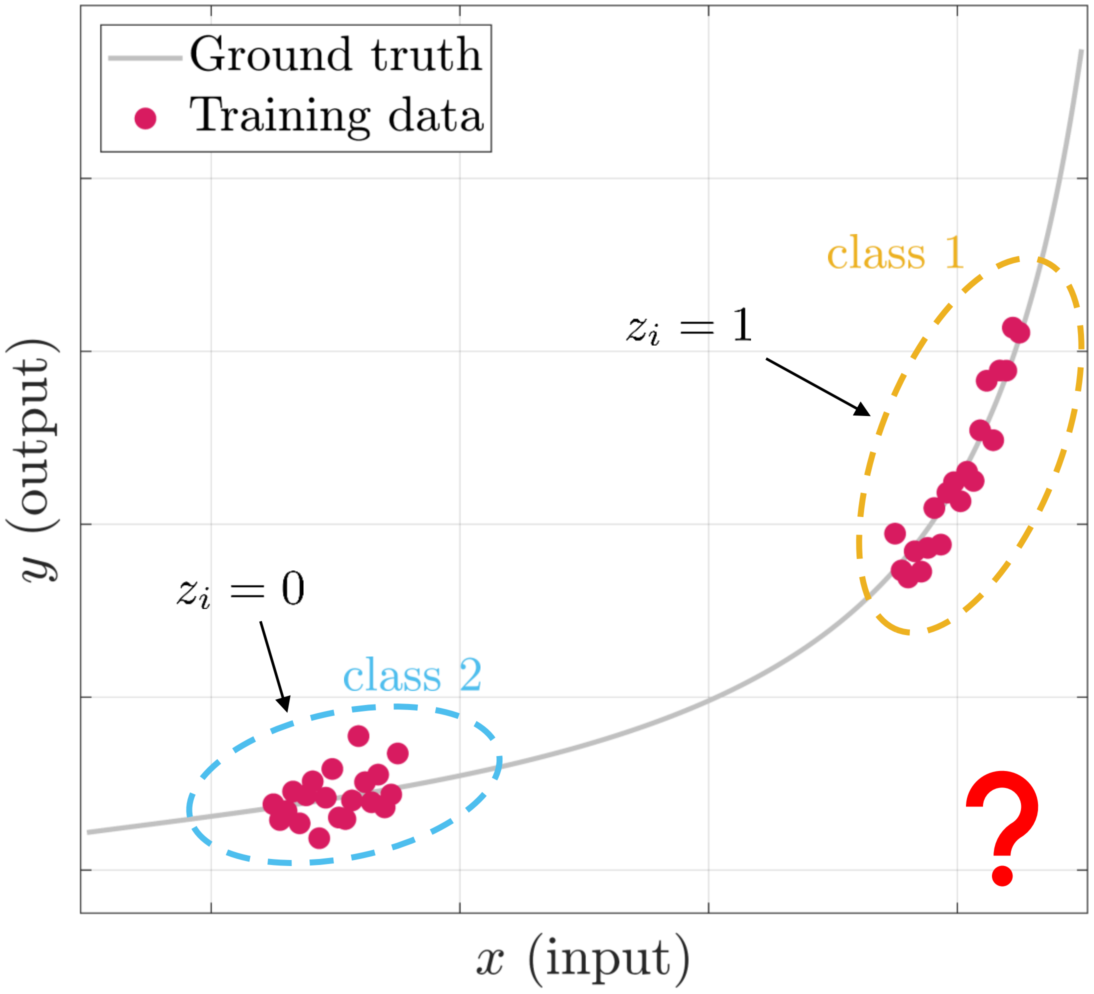
$$\min_{\substack{e\geq 0, w, b_w\\ p_1, b_{p,1}, p_2, b_{p,2}}} \sum_{i=1}^{n} \left[z_i \mathrm{SSE}_{1,i} + (1-z_i)\mathrm{SSE}_{2,i} \right] + \alpha \|w\|_2^2 + \beta \|e\|_1$$

$$\text{s.t.} \quad (2z_i - 1)\left(w^\intercal x_i + b_w\right) \geq 1 - e_i$$

$$p_1 - p_2 - w = 0, \quad b_{p,1} - b_{p,2} - b_w = 0$$

$$\forall i = \{1, 2, \ldots, n\}$$

!! Unknown impact of a priori labeling on the MIS accuracy



Design of Sensor with Optimized Data Labeling

1. A priori labeling of the training dataset
2. Data classification & sensor training

$$\min_{\substack{e\geq 0, w, b_w\\ p_1, b_{p,1}, p_2, b_{p,2}}} \sum_{i=1}^{n} \left[z_i \mathrm{SSE}_{1,i} + (1-z_i)\mathrm{SSE}_{2,i} \right] + \alpha \|w\|_2^2 + \beta \|e\|_1$$

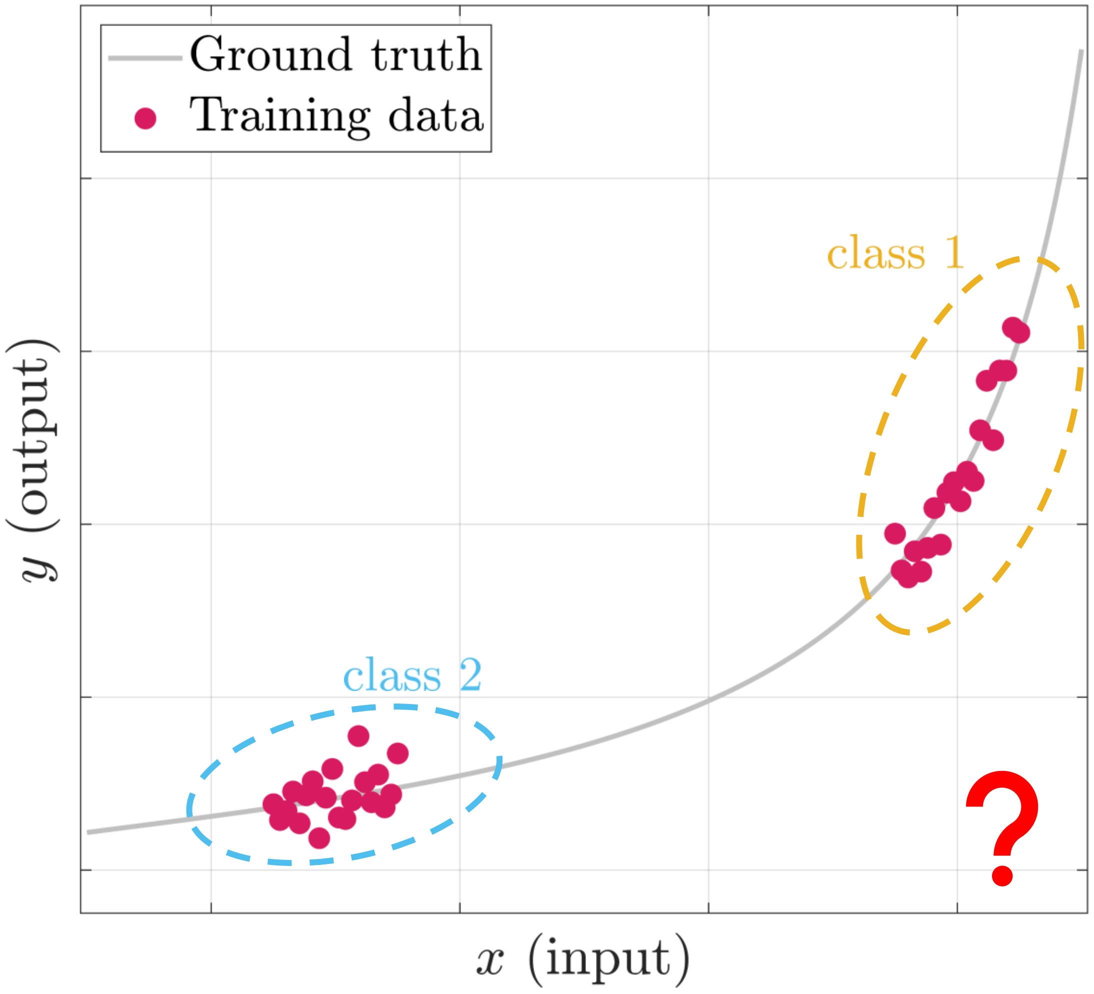
$$\text{s.t.} \quad (2z_i - 1)\left(w^\intercal x_i + b_w\right) \geq 1 - e_i$$

$$p_1 - p_2 - w = 0, \quad b_{p,1} - b_{p,2} - b_w = 0$$

$$\forall i = \{1, 2, \ldots, n\}$$

!! Unknown impact of a priori labeling on the MIS accuracy

$\Rightarrow$ z will be optimized!



Design of Sensor with Optimized Data Labeling

Data labeling & Data classification & Sensor training

simultaneously

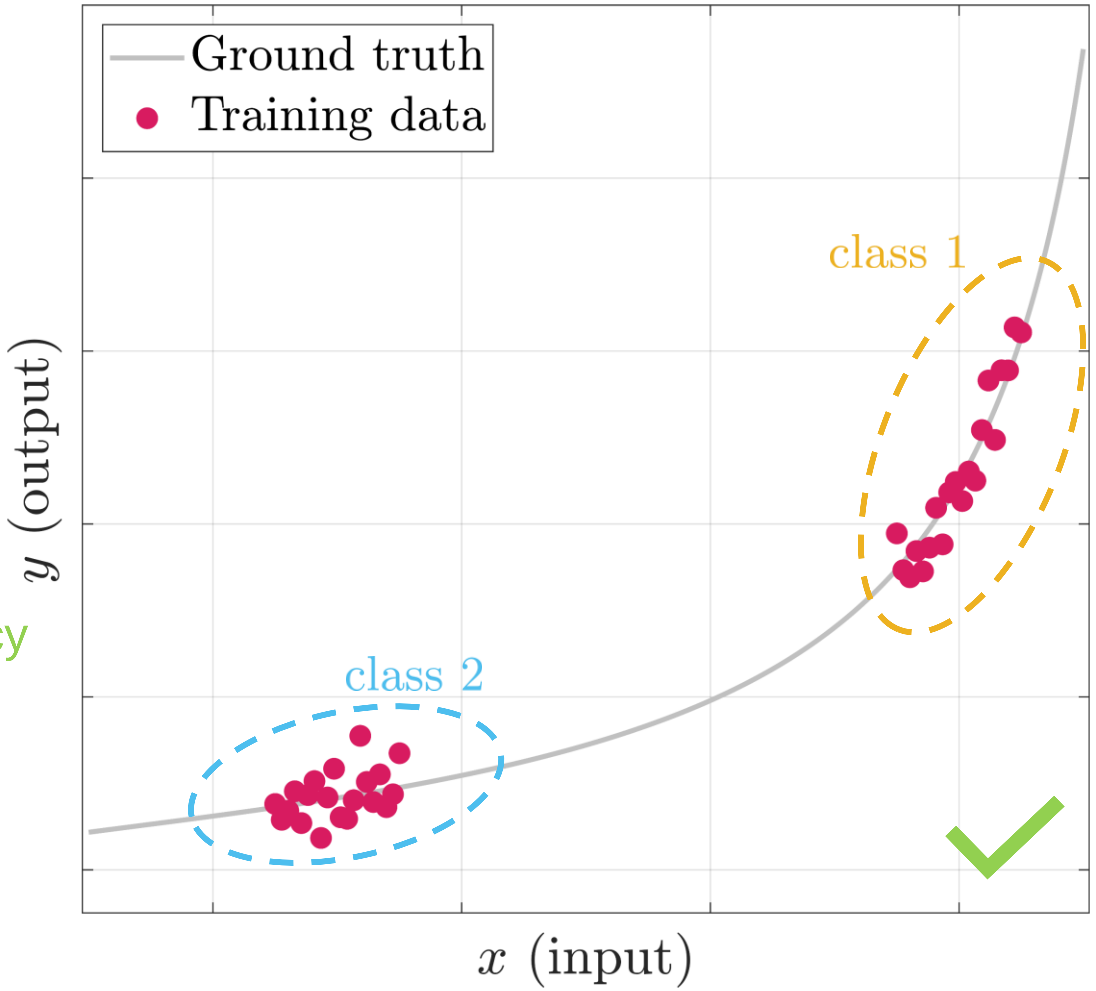
$$\min_{\substack{e\geq 0, z, w, b_w \\ p_1, b_{p,1}, p_2, b_{p,2}}} \sum_{i=1}^{n} \left[z_i \mathrm{SSE}_{1,i} + (1 - z_i)\mathrm{SSE}_{2,i} \right] + \alpha \|w\|_2^2 + \beta \|e\|_1$$

$$\text{s.t.} \quad (2z_i - 1)\left(w^\intercal x_i + b_w\right) \geq 1 - e_i$$

$$p_1 - p_2 - w = 0, \quad b_{p,1} - b_{p,2} - b_w = 0$$

$$\forall i = \{1, 2, \ldots, n\}, \quad z \in \{0, 1\}^n$$

✓ Considered impact of a priori labeling on the MIS accuracy



Data labeling & Data classification & Sensor training

simultaneously

$$\min_{\substack{e\geq 0, z, w, b_w \\ p_1, b_{p,1}, p_2, b_{p,2}}} \sum_{i=1}^{n} \left[z_i \mathrm{SSE}_{1,i} + (1-z_i)\mathrm{SSE}_{2,i} \right] + \alpha \|w\|_2^2 + \beta \|e\|_1$$

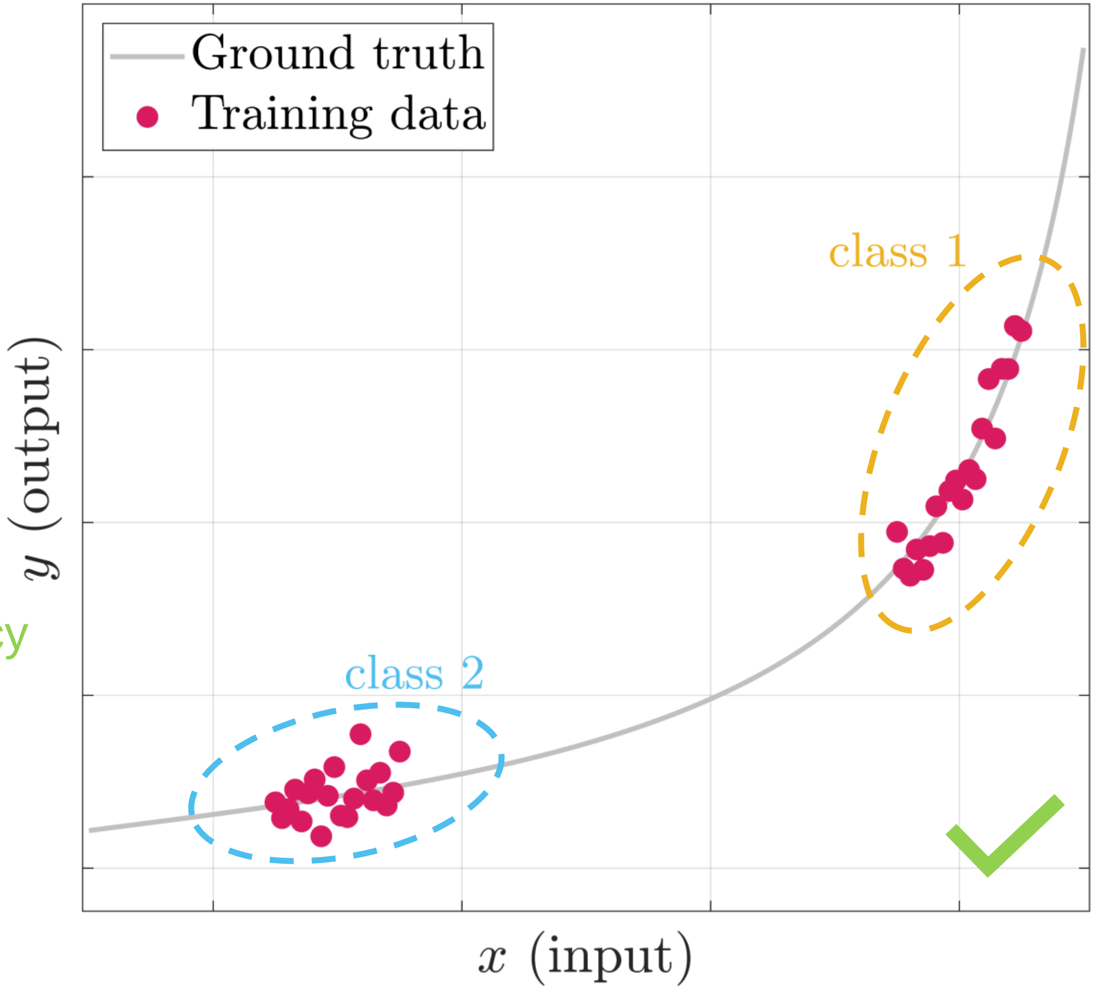
$$\begin{aligned} \text{s.t.} \quad & (2z_i - 1)\left(w^\intercal x_i + b_w\right) \geq 1 - e_i \\ & p_1 - p_2 - w = 0, \quad b_{p,1} - b_{p,2} - b_w = 0 \\ & \forall i = \{1, 2, \ldots, n\}, \quad z \in \{0,1\}^n \end{aligned}$$

MIS-con-lab

✓ Considered impact of a priori labeling on the MIS accuracy

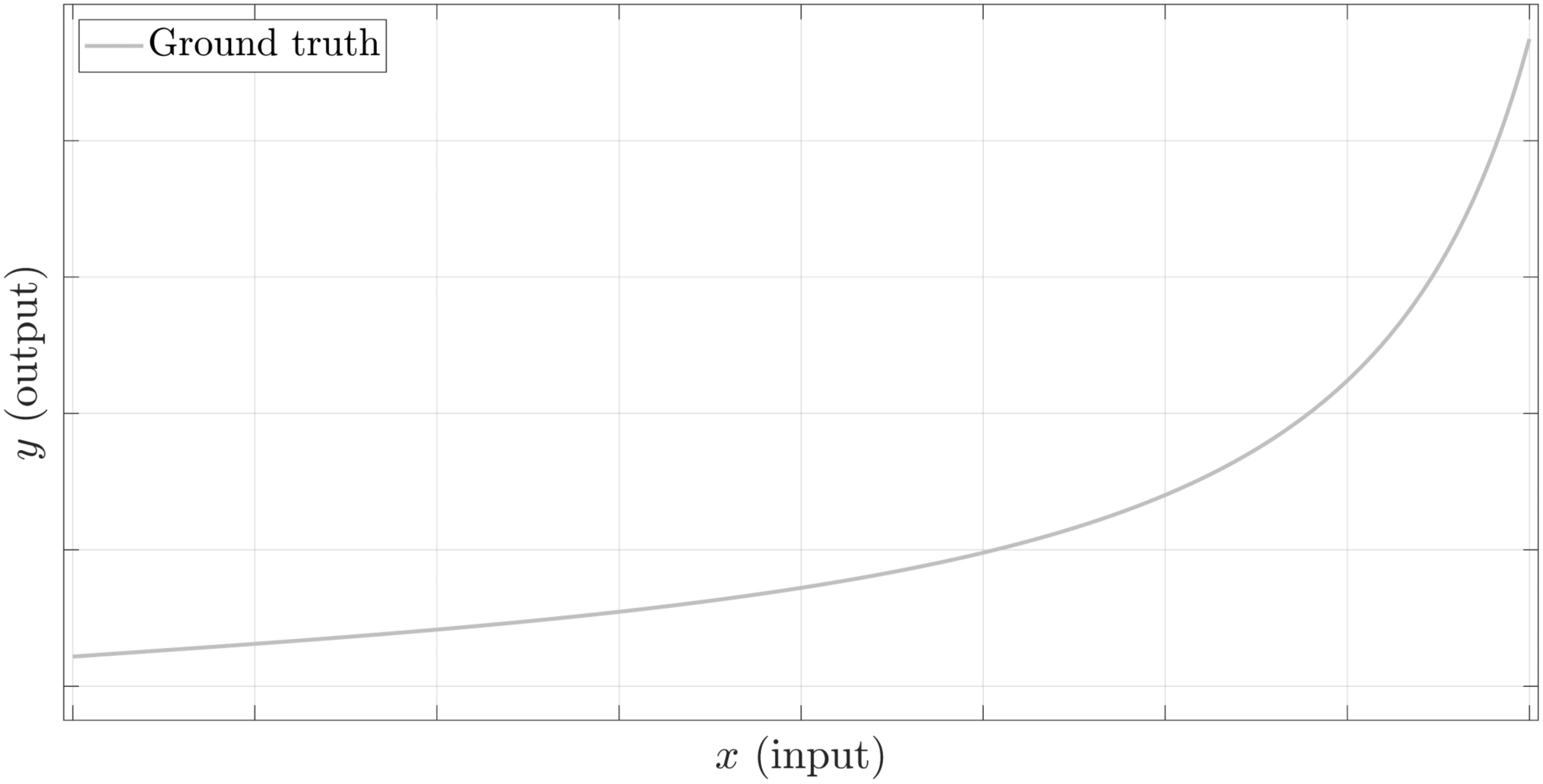
Considered solvers

- ⇒ BARON: MIS-con-lab (bar)
- ⇒ BARON with heuristic termination: MIS-con-lab (bar,ht)
- ⇒ Gurobi with big-M reformulations: MIS-con-lab (gur)



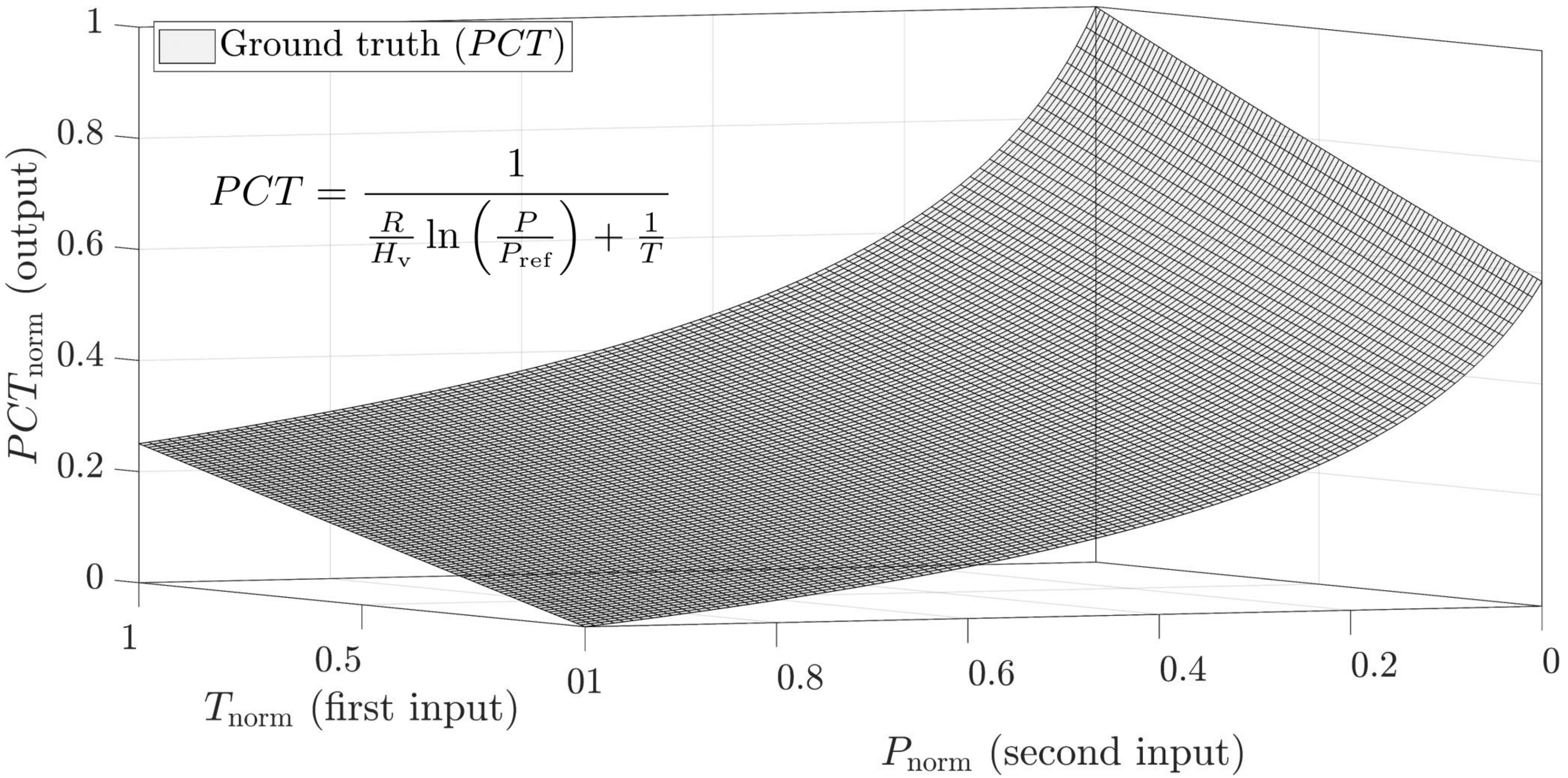
Results: Case Study

Case study: 2D

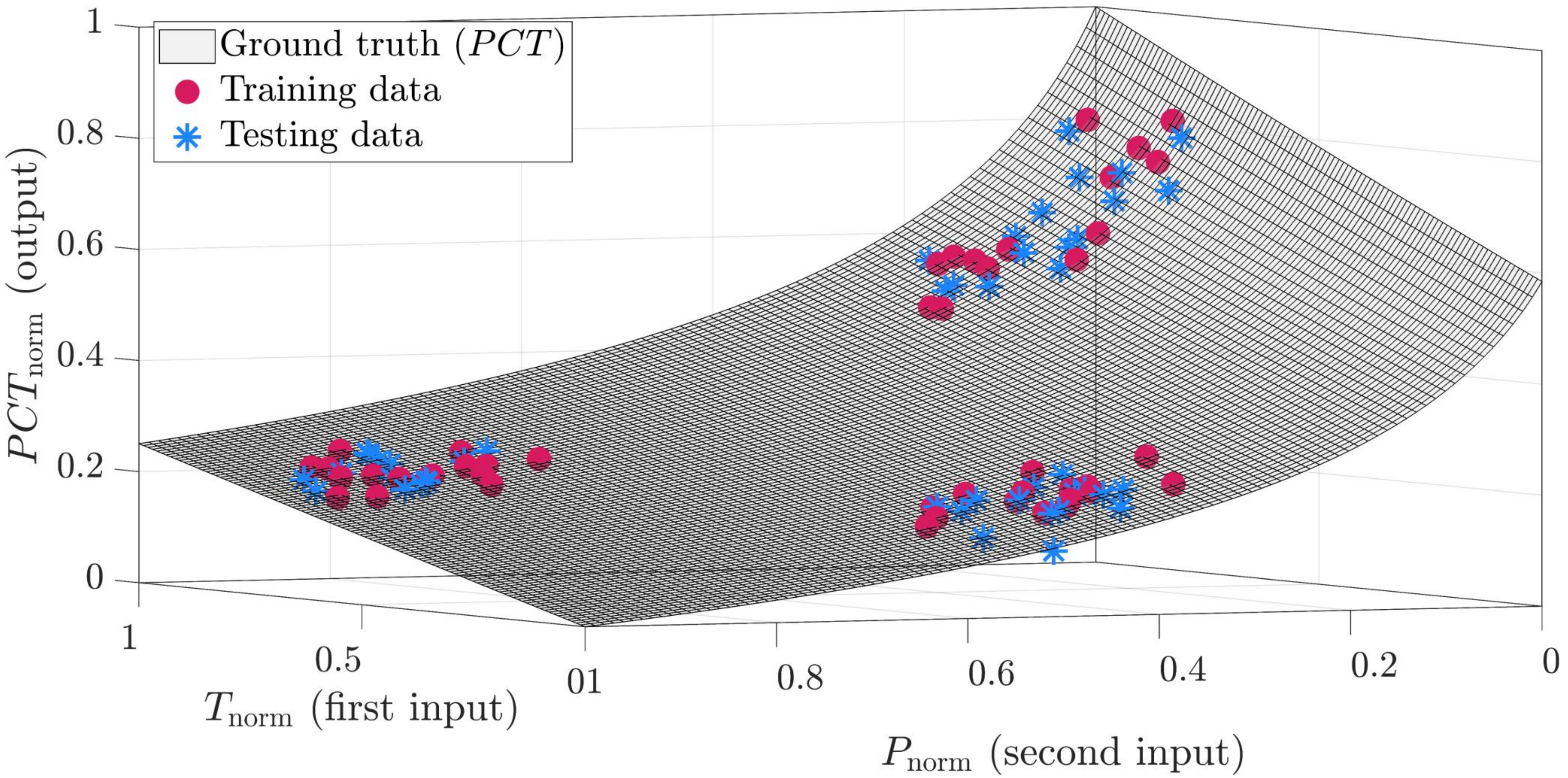


Results: Case Study

Case study: 2D $\longrightarrow$ 3D



Results: Clustered Dataset

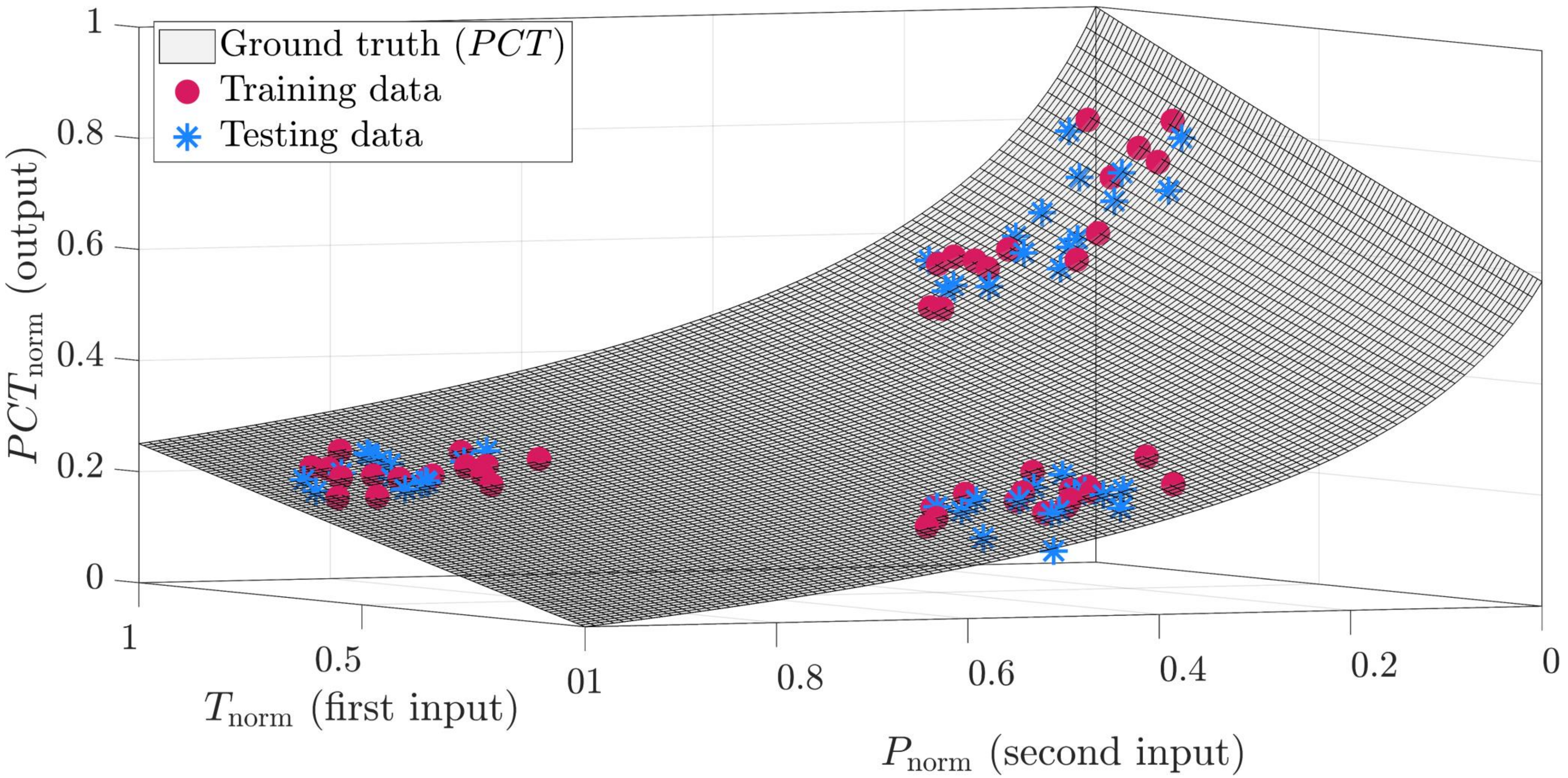


Results: Clustered Dataset

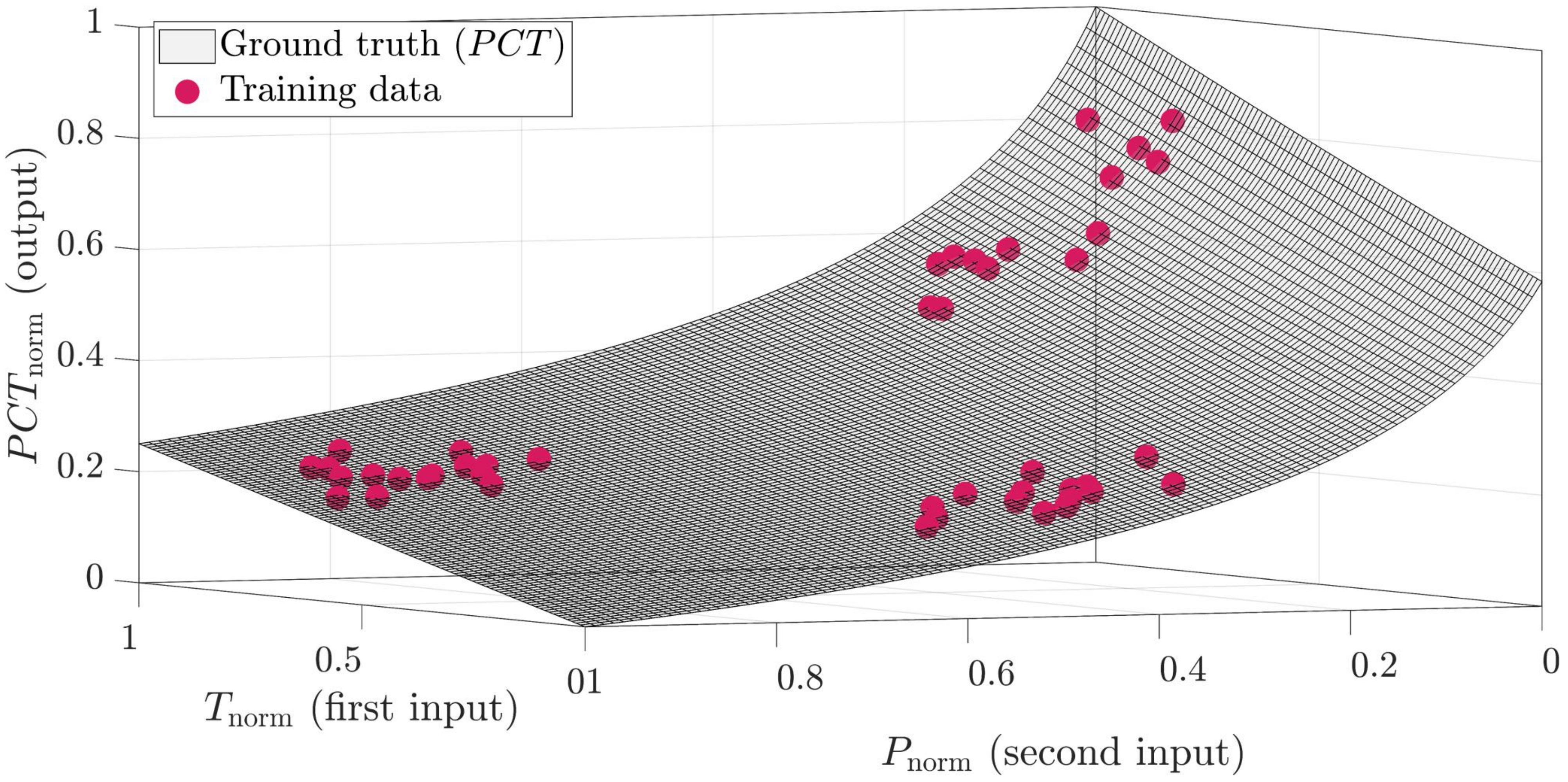
Approach	RMSE×10^2 (training)	RMSE×10^2 (testing)	CPU time [s]
SIS	4.1	3.5	0.1
MIS-SotA	0.9	1	4.3
MIS-con	0.9	1.1	4.1
MIS-con-lab (bar)	0.5	1.6	3,600.0
MIS-con-lab (bar,ht)	0.5	1.6	325.9
MIS-con-lab (gur)	0.4	0.5	2,284.0

Windows 10.0.22621, x64-based PC, Intel(R) Core (TM) i7-8750H, 2.2GHz, 6 Core(s)

Results: Clustered Dataset

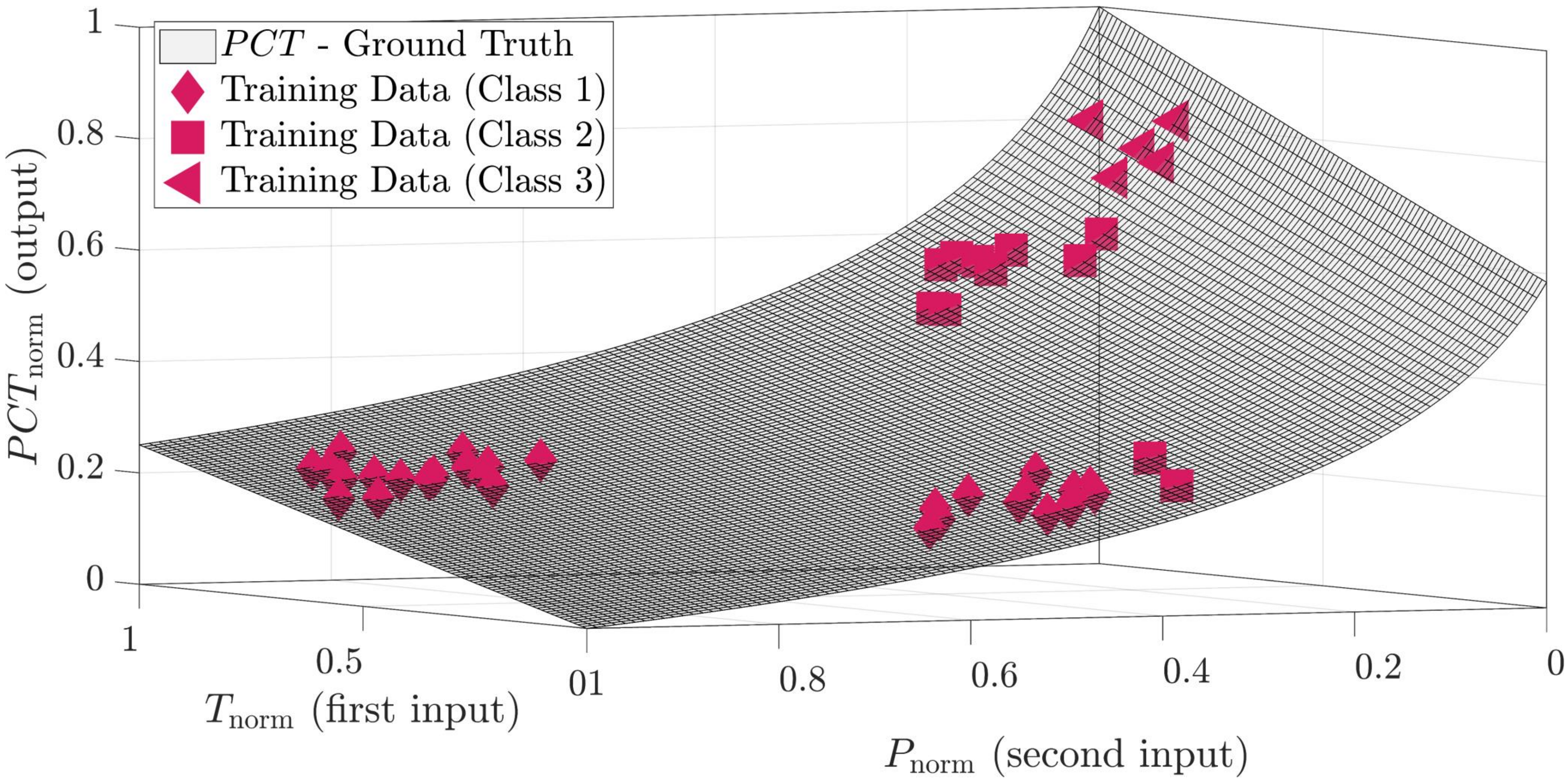


Approach: MIS-con-lab (gur)



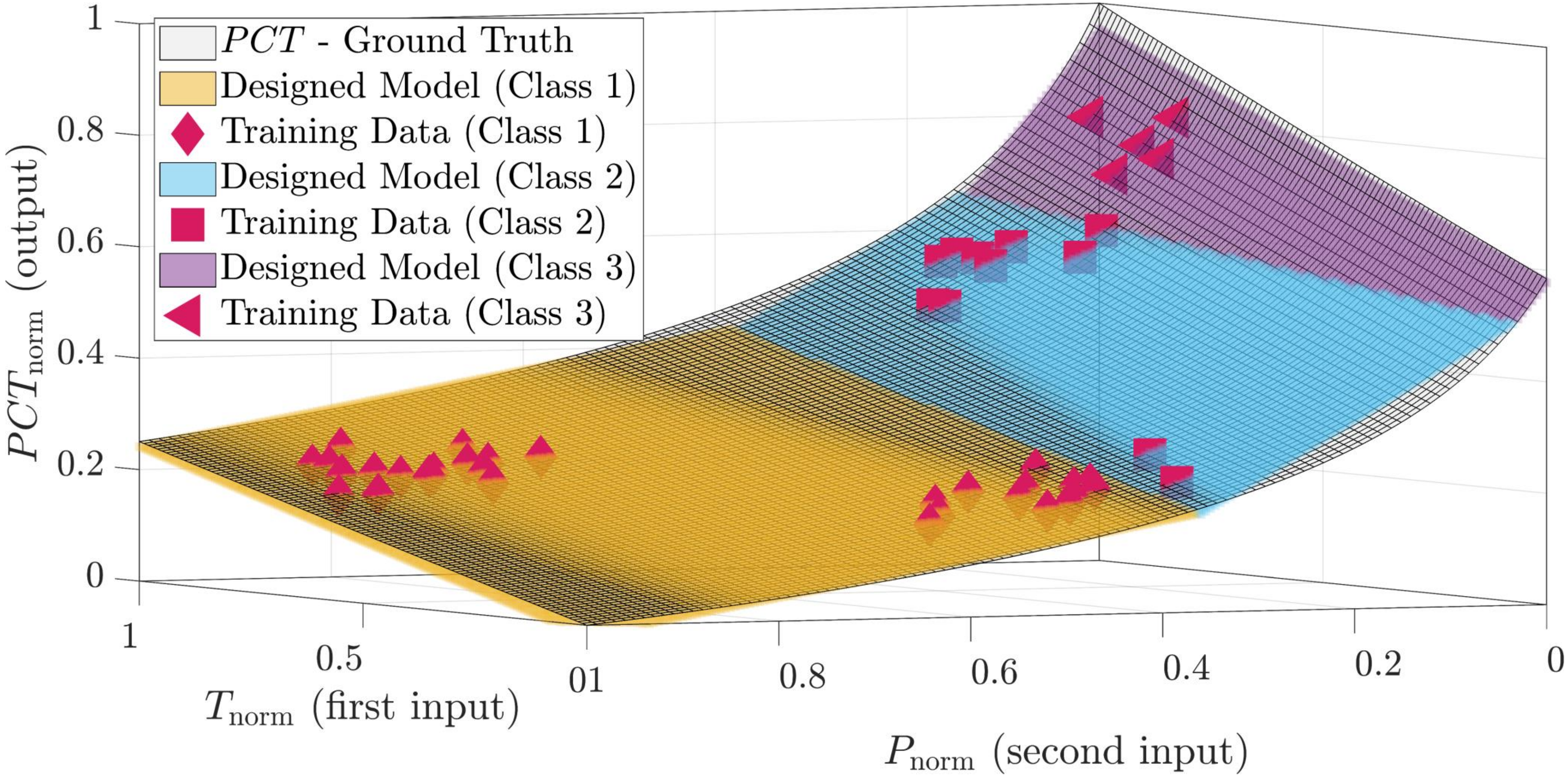
Results: Clustered Dataset

Approach: MIS-con-lab (gur)

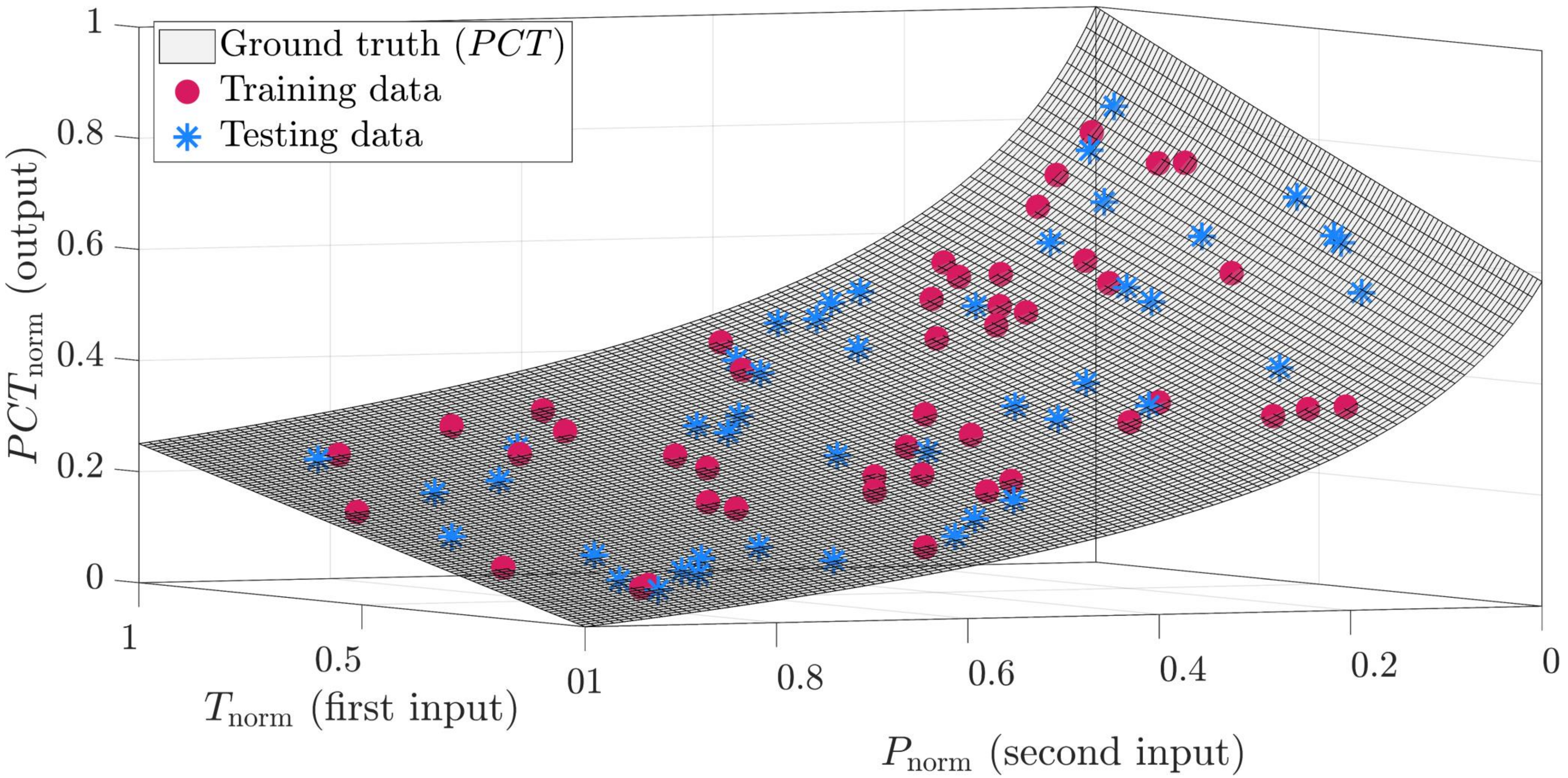


Results: Clustered Dataset

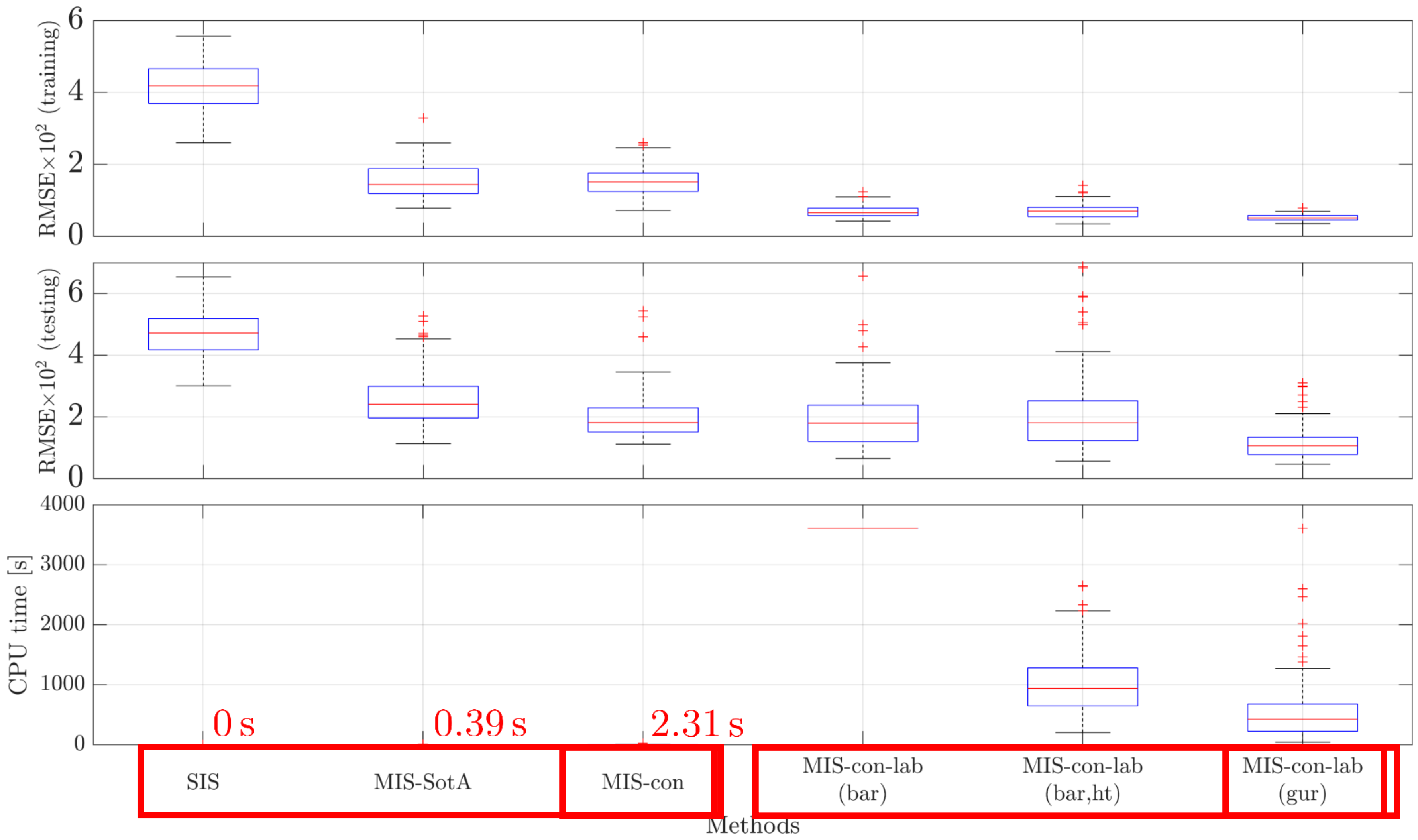
Approach: MIS-con-lab (gur)



Results: Uniformly Distributed Dataset

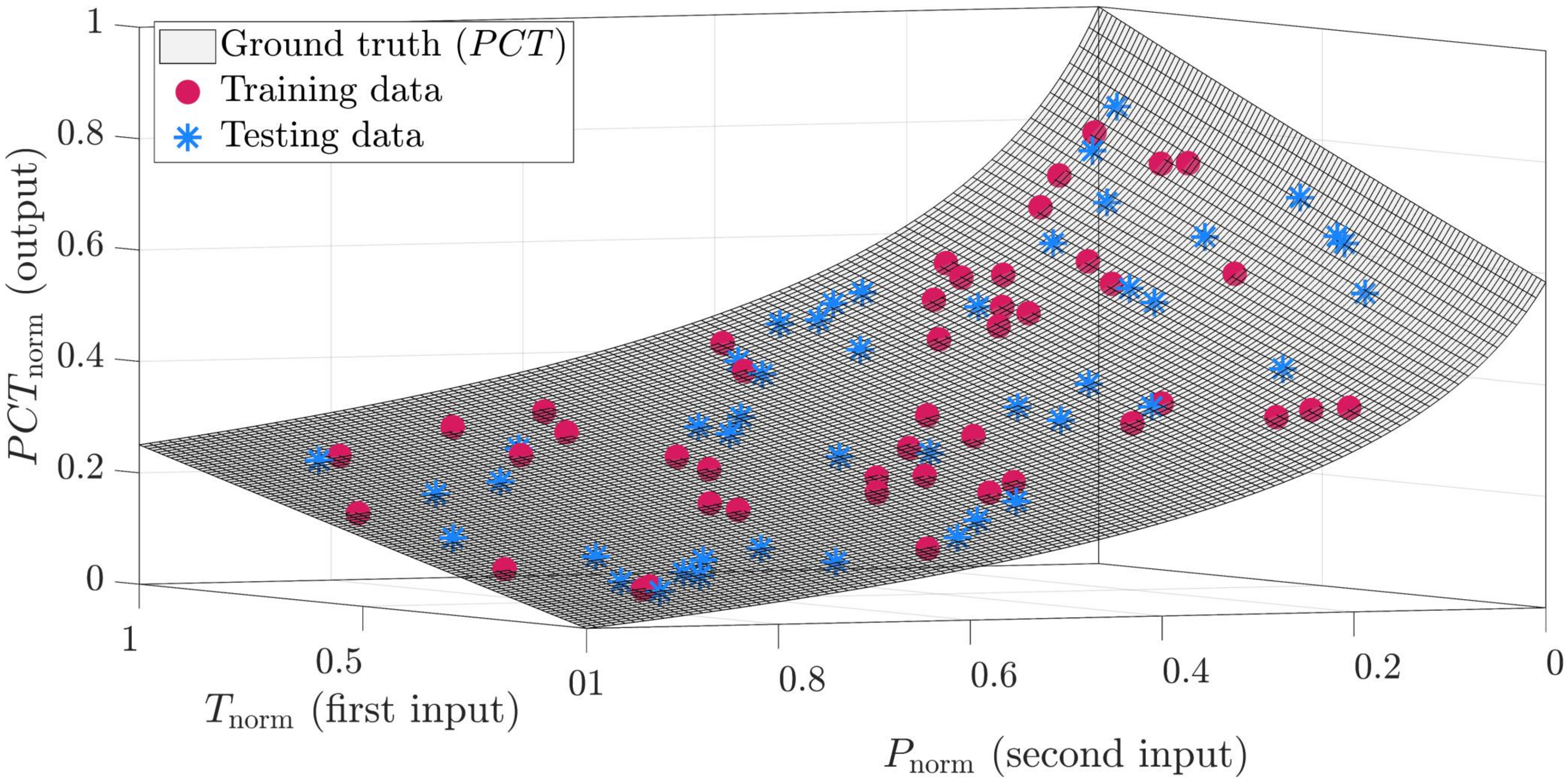


Results: Uniformly Distributed Dataset



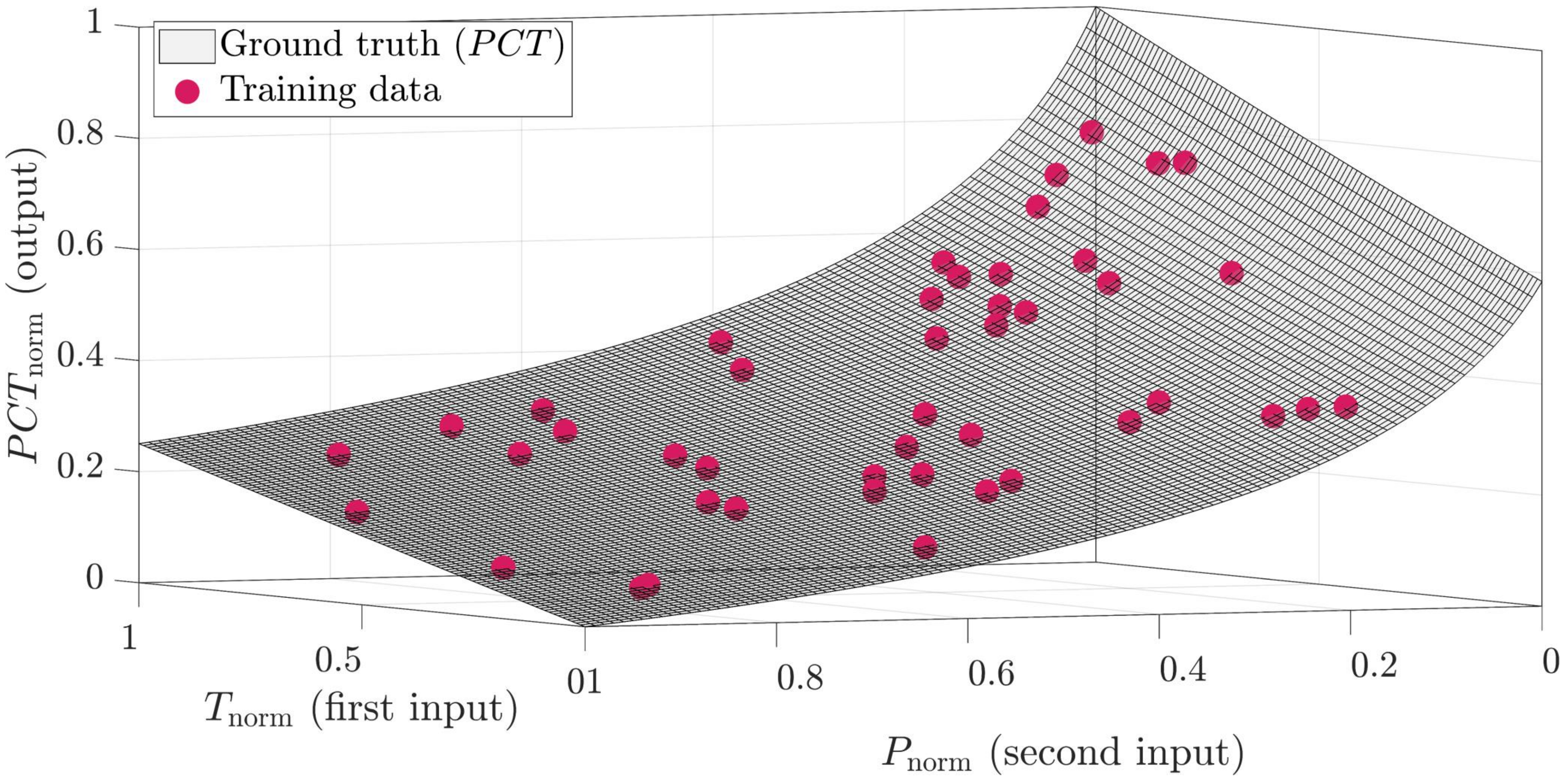
Windows 10.0.22621, x64-based PC, Intel(R) Core (TM) i7-8750H, 2.2GHz, 6 Core(s)

Results: Uniformly Distributed Dataset



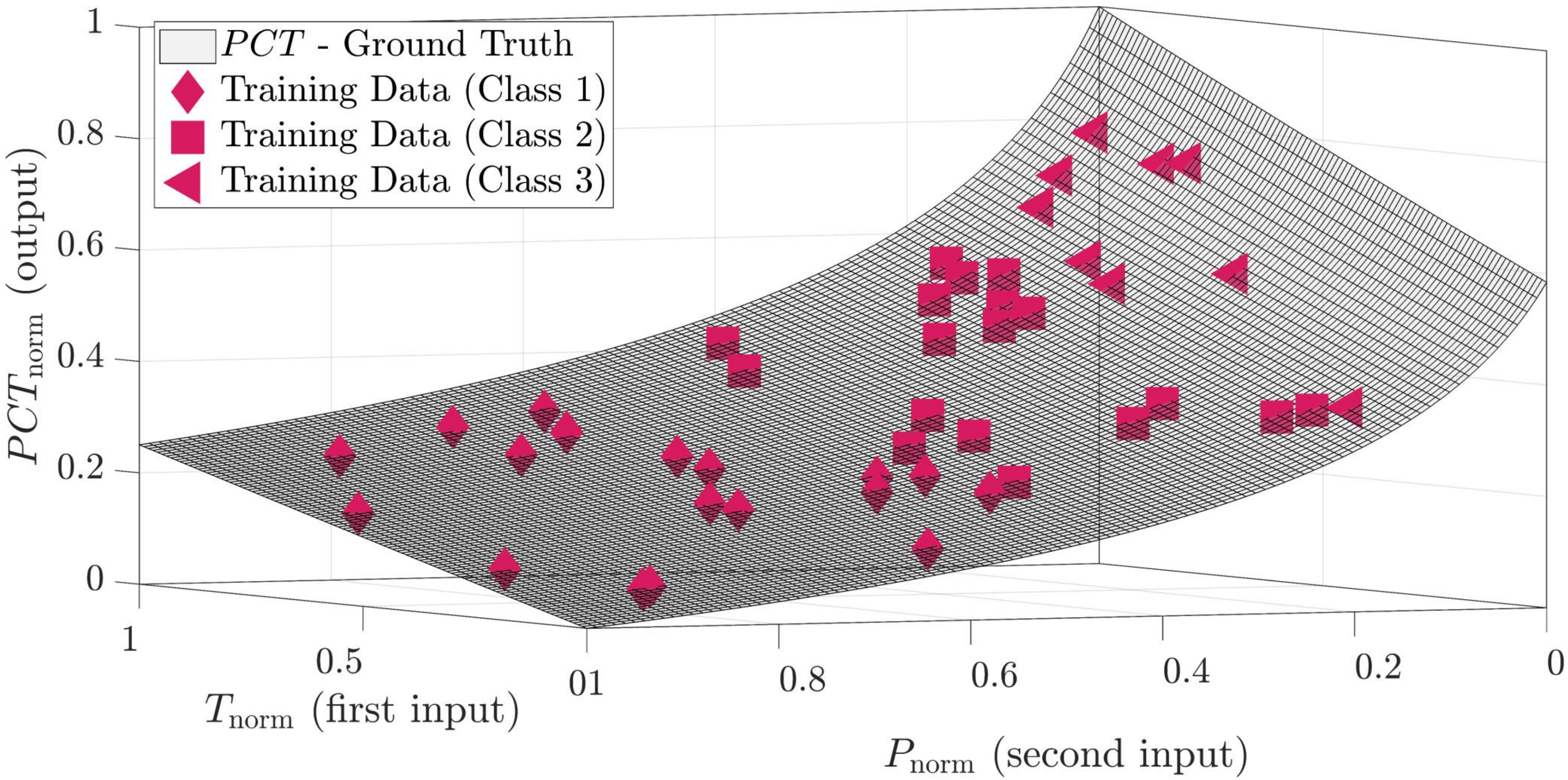
Results: Uniformly Distributed Dataset

Approach: MIS-con-lab (gur)



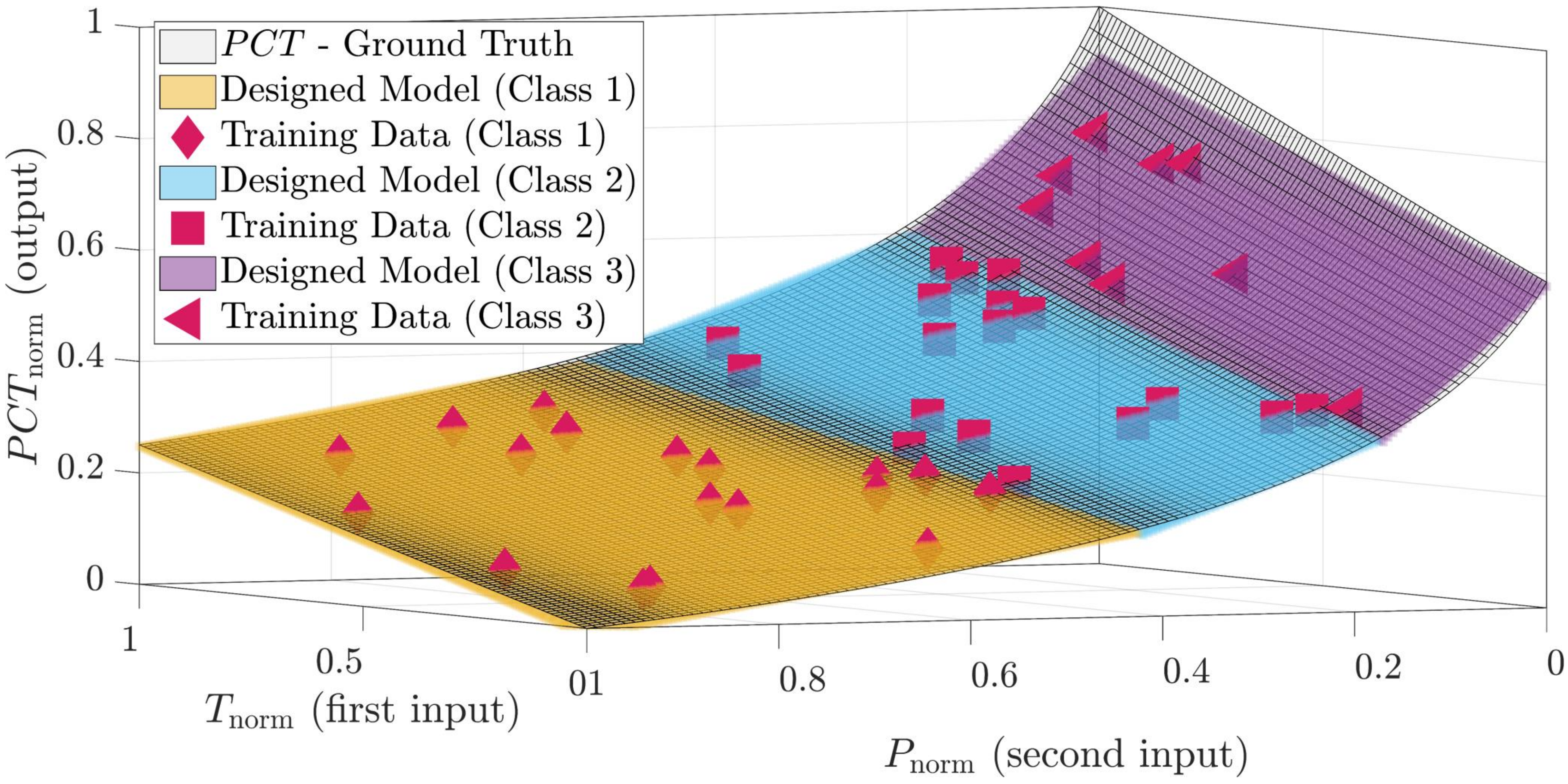
Results: Uniformly Distributed Dataset

Approach: MIS-con-lab (gur)



Results: Uniformly Distributed Dataset

Approach: MIS-con-lab (gur)



Conclusions

- Based on the studied methodology, MIS-con-lab (gur) appears to be the most suitable approach.
- MIS-con offers the optimal balance between model accuracy, computational load, and model continuity.
- Optimising the data labels results in a significant computational load and variance in the results.

for more info visit

Future Work

- Consideration of advanced methods (e.g., LASSO) into MIS to allow different model structures.
- Incorporation of various approaches for data treatment to enhance data quality.

Acknowledgments

APVV-21-0019, APVV SK-FR-2019-0004
VEGA 1/0691/21, VEGA 1/0297/22

European Commission under the grant no. 101079342 (Fostering Opportunities Towards Slovak Excellence in Advanced Control for Smart Industries)