AVERAGE NUMBER OF MISTAKES IN SEQUENTIAL RISK-AVERSE SCENARIO DECISION-MAKING

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
UNIVERSITY OF BRESCIA
POLITECNICO MILANO 1863

AVERAGE NUMBER OF MISTAKES IN SEQUENTIAL RISK-AVERSE SCENARIO DECISION-MAKING

Simone Garatti
(Politecnico di Milano, MI)

Marco C. Campi
(University of Brescia, IT)

63rd Conference on Decision and Control 2024 - "Learning-Based Control V: Safety and Convergence Guarantees" December 19

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
UNIVERSITY OF BRESCIA
POLITECNICO MILANO 1863

AVERAGE NUMBER OF MISTAKES IN SEQUENTIAL RISK-AVERSE SCENARIO DECISION-MAKING

Simone Garatti
(Politecnico di Milano, MI)

Marco C. Campi
(University of Brescia, IT)

Algo Carè
(University of Brescia, IT)

ALGO.CARE@UNIBS.IT

63rd Conference on Decision and Control 2024 - "Learning-Based Control V: Safety and Convergence Guarantees" December 19

Finanziato dall'Unione europea NextGenerationEU

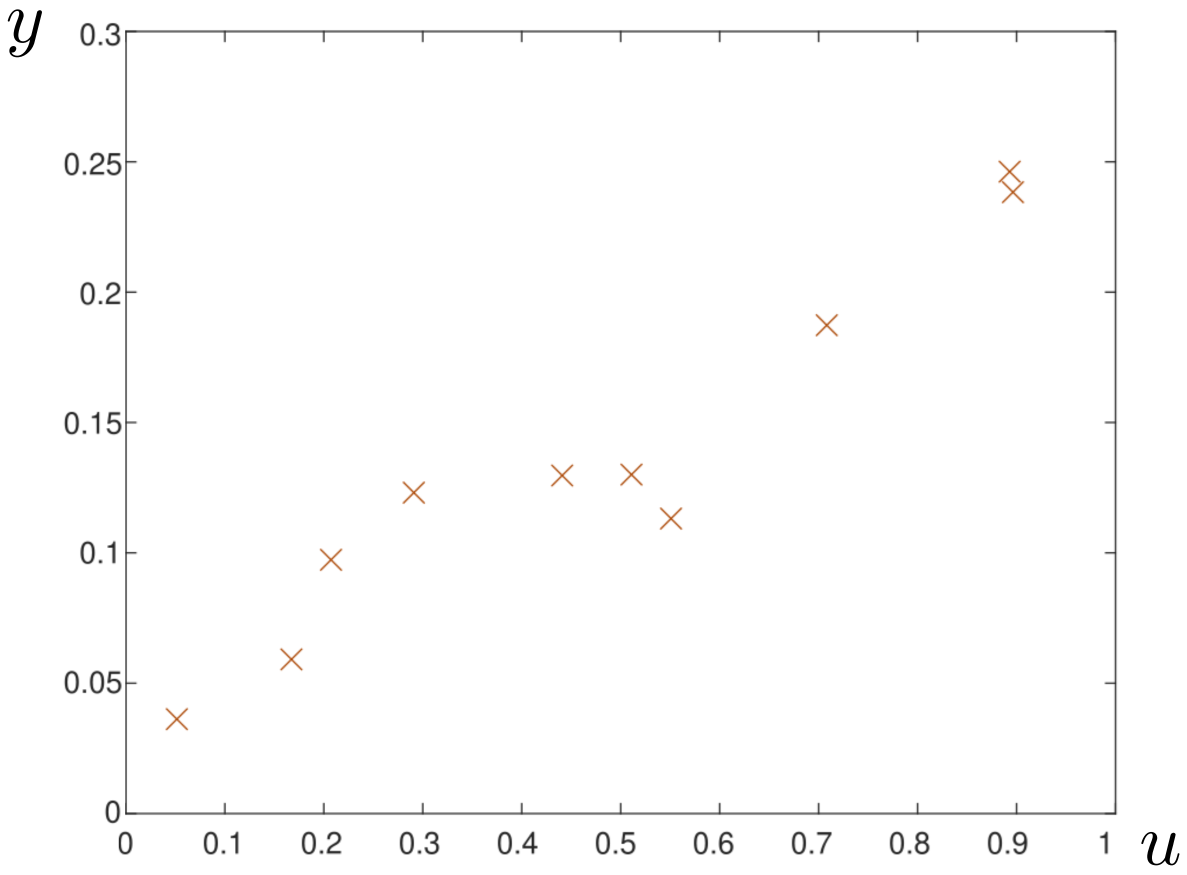
Ministero dell'Università e della Ricerca

Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA

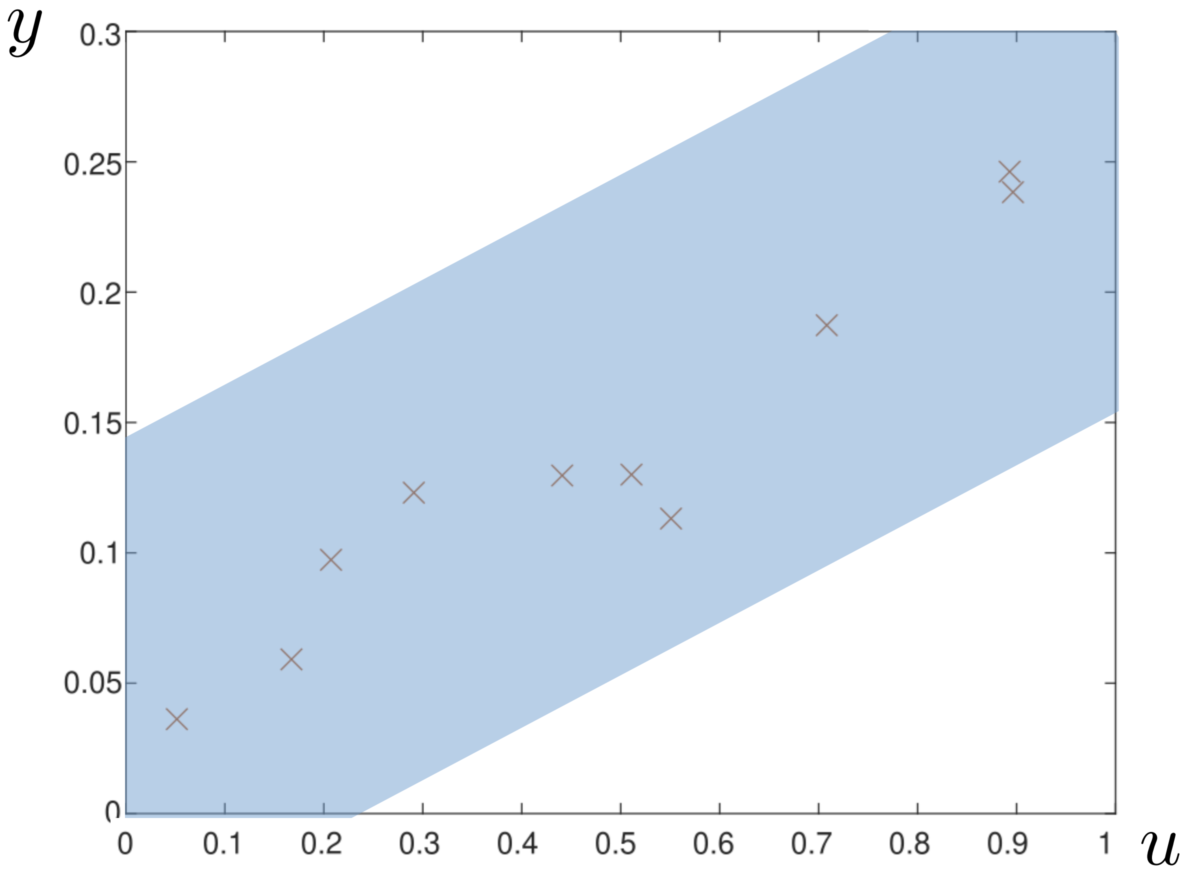
UNIVERSITY OF BRESCIA

POLITECNICO MILANO 1863

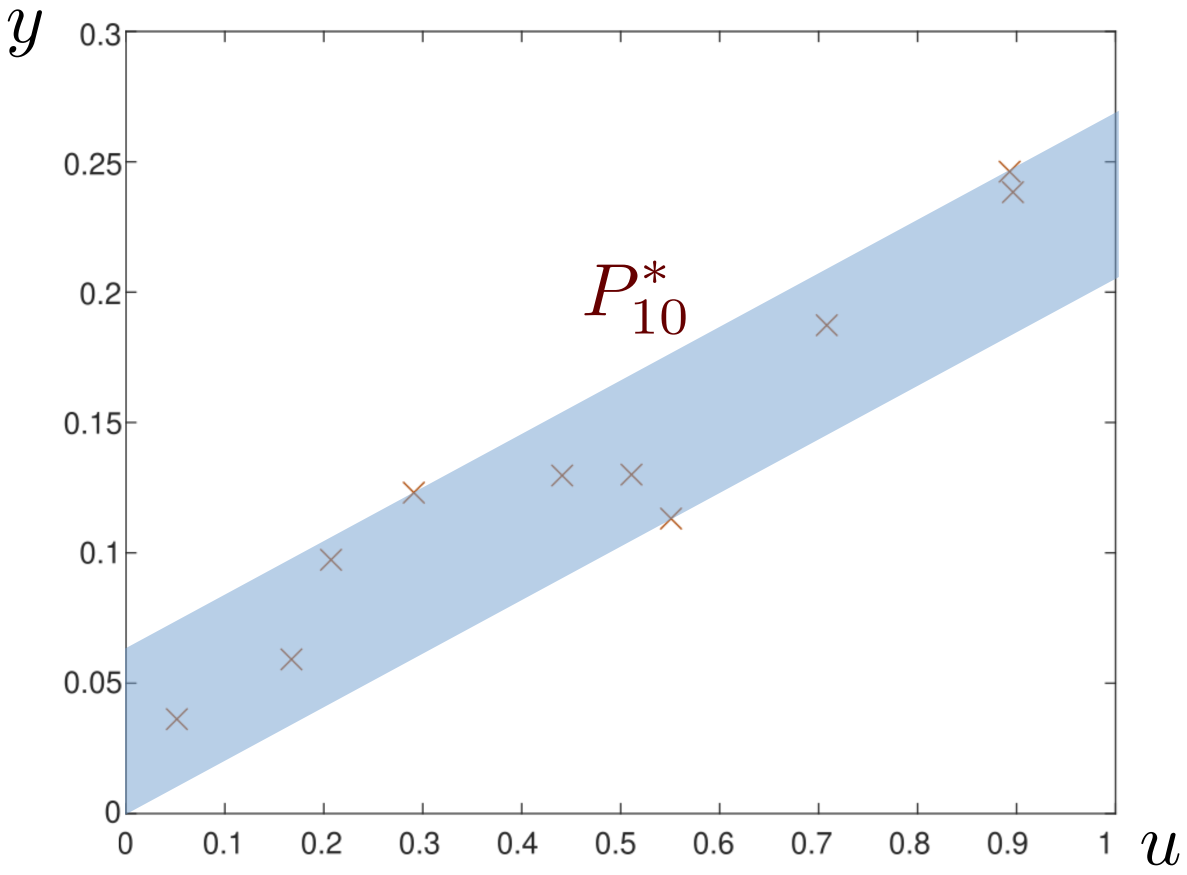
Train a simple predictor on N=10 data points (i.i.d.)



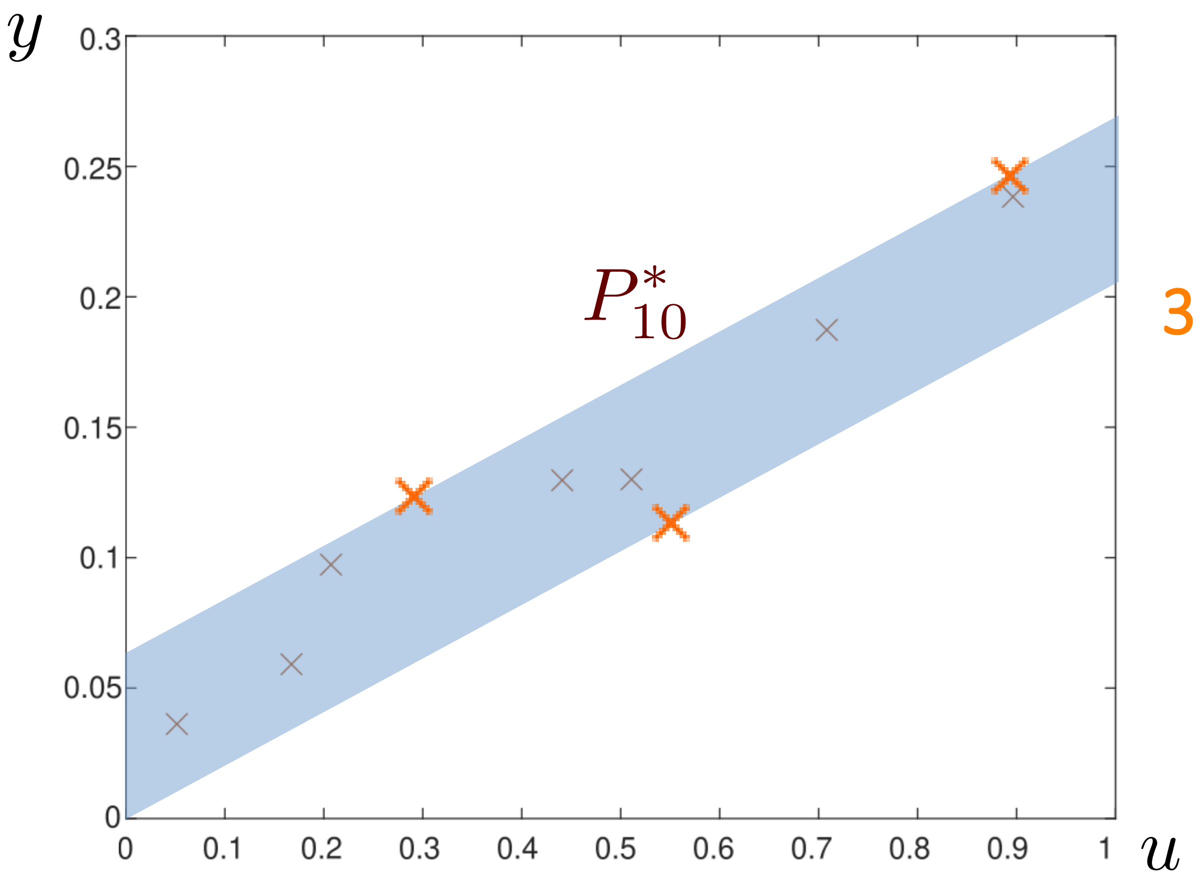
Train a simple predictor on N=10 data points (i.i.d.)



Train a simple predictor on N=10 data points (i.i.d.)

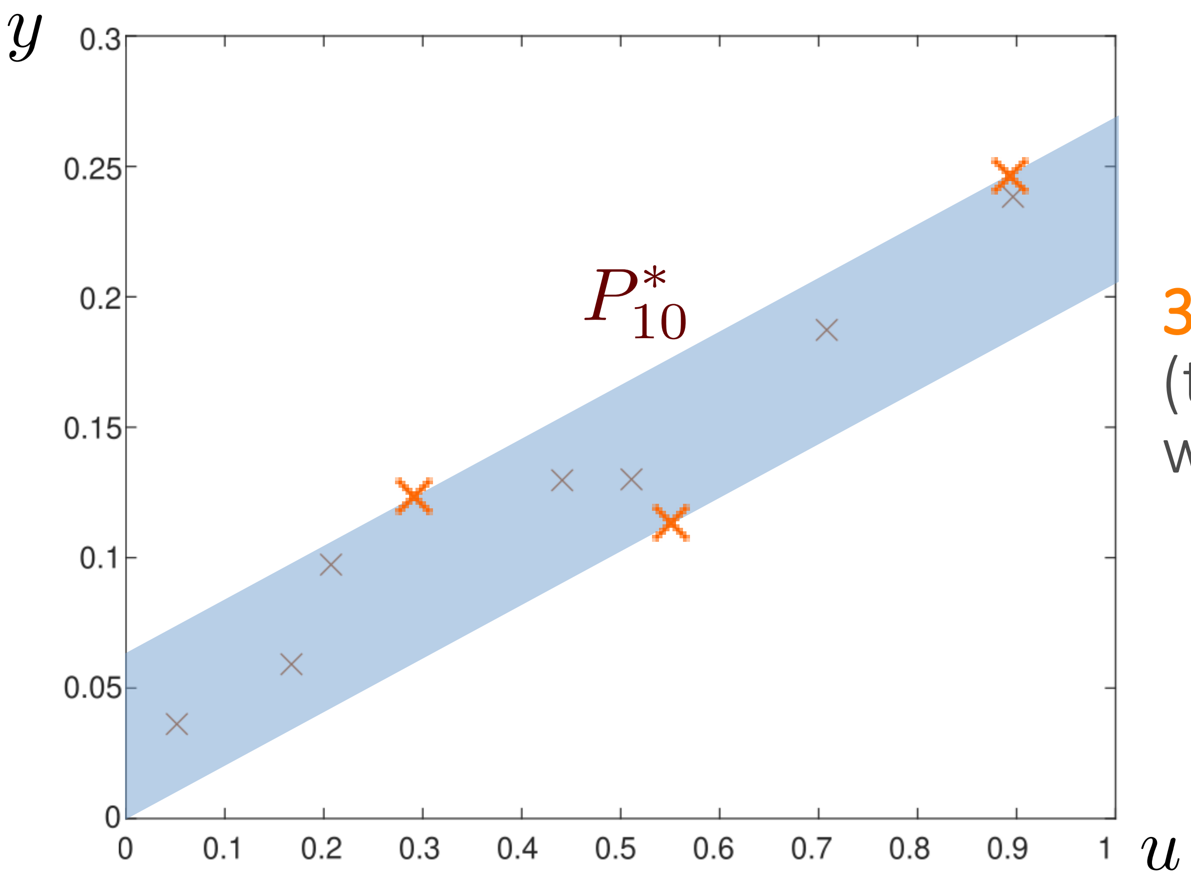


Train a simple predictor on N=10 data points (i.i.d.)



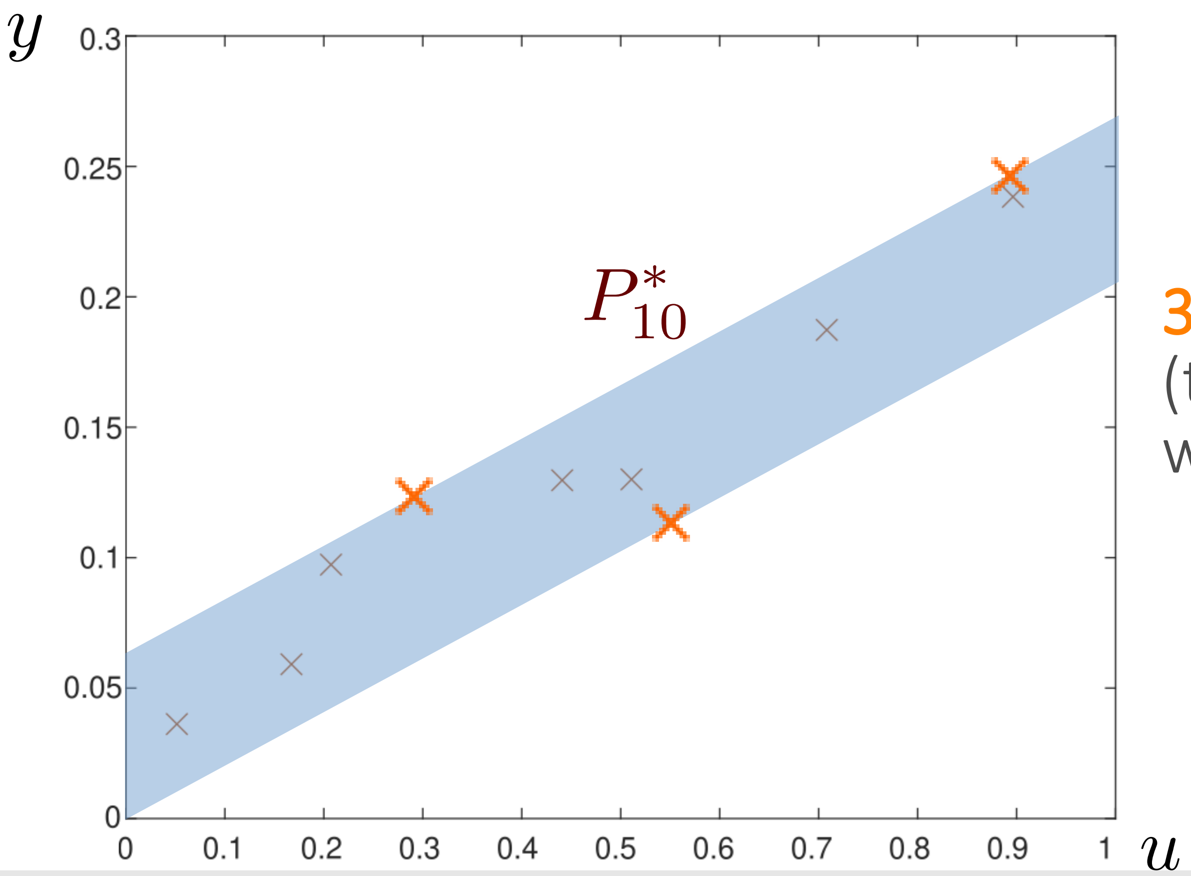
3 support points

Train a simple predictor on N=10 data points (i.i.d.)



3 support points
(this happens with probability 1)

Train a simple predictor on N=10 data points (i.i.d.)



3 support points
(this happens with probability 1)

S. Garatti, M.C. Campi, A. Carè
"On a Class of Interval Predictor Models with Universal Reliability"
Automatica 2019

$$(u_1, y_1), \ldots, (u_{10}, y_{10}), (u_{11}, y_{11})$$

$$(u_1, y_1), \ldots, (u_{10}, y_{10}), (u_{11}, y_{11})$$

Training

$$(u_1, y_1), \ldots, (u_{10}, y_{10}), (u_{11}, y_{11}) \qquad P_{10}^*$$

Training

$$(u_1, y_1), \ldots, (u_{10}, y_{10}), {\color{blue}(u_{11}, y_{11})} \qquad P_{10}^*$$

Future point

S. Garatti, M.C. Campi, A. Carè
"On a Class of Interval Predictor Models with Universal Reliability"
Automatica 2019

Training

$$(u_1, y_1), \ldots, (u_{10}, y_{10}), (u_{11}, y_{11}) : (u_{11}, y_{11}) \notin P^*_{10}$$

Future point

prediction error

Training

$$\{(u_1, y_1), \ldots, (u_{10}, y_{10}), (u_{11}, y_{11}) : (u_{11}, y_{11}) \notin P^*_{10}\}$$

Future point

prediction error

S. Garatti, M.C. Campi, A. Carè
"On a Class of Interval Predictor Models with Universal Reliability"
Automatica 2019

Training

$$Prob\{(u_1, y_1), \ldots, (u_{10}, y_{10}), (u_{11}, y_{11}) : (u_{11}, y_{11}) \notin P^*_{10}\}$$

Future point

prediction error

S. Garatti, M.C. Campi, A. Carè
"On a Class of Interval Predictor Models with Universal Reliability"
Automatica 2019

Training

$$Prob\{(u_1, y_1), \ldots, (u_{10}, y_{10}), (u_{11}, y_{11}) : (u_{11}, y_{11}) \notin P^*_{10}\}$$

Future point

prediction error

$$PErr_{10}$$

S. Garatti, M.C. Campi, A. Carè
“On a Class of Interval Predictor Models with Universal Reliability”
Automatica 2019

Training

$$Prob\{(u_1, y_1), \ldots, (u_{10}, y_{10}), (u_{11}, y_{11}) : (u_{11}, y_{11}) \notin P^*_{10}\}$$

Future point

prediction error

$$PErr_{10} = \frac{3}{11}$$

S. Garatti, M.C. Campi, A. Carè
"On a Class of Interval Predictor Models with Universal Reliability"
Automatica 2019

Training

$$Prob\{(u_1, y_1), \ldots, (u_{10}, y_{10}), (u_{11}, y_{11}) : (u_{11}, y_{11}) \notin P^*_{10}\}$$

Future point

prediction error

$$PErr_{10} = \frac{3}{11}$$

$$= \frac{\text{\#support points}}{N+1}$$

S. Garatti, M.C. Campi, A. Carè
"On a Class of Interval Predictor Models with Universal Reliability"
Automatica 2019

Our predictor

Training:

Our predictor

Training:

Data-driven decision

Optimization:

Our predictor

Training:

$$\min_{(\theta_0,\theta_1,w)\in\mathbb{R}^3} \quad w$$

Data-driven decision

Optimization:

Our predictor

Training:

$$\min_{(\theta_0,\theta_1,w)\in\mathbb{R}^3} \quad w$$

Data-driven decision

Optimization:

$$\min_{x\in\mathcal{X}\subseteq\mathbb{R}^d} \quad c(x)$$

Our predictor

Training:

$$\min_{(\theta_0,\theta_1,w)\in\mathbb{R}^3} \quad w$$

subject to:

$$|y_i - (\theta_0 + \theta_1 u_i)| \le w,$$

$$i = 1,\ldots,N$$

Data-driven decision

Optimization:

$$\min_{x\in\mathcal{X}\subseteq\mathbb{R}^d} \quad c(x)$$

Our predictor

Training:

$$\min_{(\theta_0,\theta_1,w)\in\mathbb{R}^3} \quad w$$

subject to:

$$|y_i - (\theta_0 + \theta_1 u_i)| \leq w,$$

$$i = 1, \ldots, N$$

Data-driven decision

Optimization:

$$\min_{x\in\mathcal{X}\subseteq\mathbb{R}^d} \quad c(x)$$

subject to:

$$x \in \mathcal{X}_{\delta_i},$$

$$i = 1, \ldots, N$$

Our predictor

Training:

$$\min_{(\theta_0,\theta_1,w)\in\mathbb{R}^3} \quad w$$

subject to:

$$|y_i - (\theta_0 + \theta_1 u_i)| \le w,$$

$$i = 1,\ldots,N$$

Data-driven decision

Optimization:

$$\min_{x\in\mathcal{X}\subseteq\mathbb{R}^d} \quad c(x)$$

subject to:

"scenario"

$$x \in \mathcal{X}_{\delta_i},$$

$$i = 1,\ldots,N$$

Our predictor

Training:

$$\min_{(\theta_0,\theta_1,w)\in\mathbb{R}^3} \quad w$$

subject to:

$$|y_i - (\theta_0 + \theta_1 u_i)| \leq w,$$

$$i = 1, \ldots, N$$

Predictor: $P_N^* = \{|y - (\theta_0^* + \theta_1^* u)| \leq w^*\}$

Data-driven decision

Optimization:

$$\min_{x\in\mathcal{X}\subseteq\mathbb{R}^d} \quad c(x)$$

subject to:

"scenario"

$$x \in \mathcal{X}_{\delta_i},$$

$$i = 1, \ldots, N$$

Our predictor

Training:

$$\min_{(\theta_0,\theta_1,w)\in\mathbb{R}^3} \quad w$$

subject to:

$$|y_i - (\theta_0 + \theta_1 u_i)| \leq w,$$

$$i = 1,\ldots,N$$

Predictor: $P_N^* = \{|y - (\theta_0^* + \theta_1^* u)| \leq w^*\}$

Data-driven decision

Optimization:

$$\min_{x\in\mathcal{X}\subseteq\mathbb{R}^d} \quad c(x)$$

subject to:

$$x \in \mathcal{X}_{\delta_i},$$

$$i = 1,\ldots,N$$

"scenario"

Solution (decision): x^*

Our predictor

Training:

$$\min_{(\theta_0,\theta_1,w)\in\mathbb{R}^3} \quad w$$

subject to:

$$|y_i - (\theta_0 + \theta_1 u_i)| \leq w,$$

$$i = 1, \ldots, N$$

Predictor: $P_N^* = \{|y - (\theta_0^* + \theta_1^* u)| \leq w^*\}$

Complexity: 3 support points

Data-driven decision

Optimization:

$$\min_{x\in\mathcal{X}\subseteq\mathbb{R}^d} \quad c(x)$$

subject to:

"scenario"

$$x \in \mathcal{X}_{\delta_i},$$

$$i = 1, \ldots, N$$

Solution (decision): x^*

Our predictor

Training:

$$\min_{(\theta_0,\theta_1,w)\in\mathbb{R}^3} \quad w$$

subject to:

$$|y_i - (\theta_0 + \theta_1 u_i)| \le w,$$

$$i = 1,\dots,N$$

Predictor: $P_N^* = \{|y - (\theta_0^* + \theta_1^* u)| \le w^*\}$

Complexity: 3 support points

Data-driven decision

Optimization:

$$\min_{x\in\mathcal{X}\subseteq\mathbb{R}^d} \quad c(x)$$

subject to:

"scenario"

$$x \in \mathcal{X}_{\delta_i},$$

$$i = 1,\dots,N$$

Solution (decision): x^*

Our predictor

Training:

$$\min_{(\theta_0,\theta_1,w)\in\mathbb{R}^3} \quad w$$

subject to:

$$|y_i - (\theta_0 + \theta_1 u_i)| \le w,$$

$$i = 1,\dots,N$$

Predictor: $P_N^* = \{|y - (\theta_0^* + \theta_1^* u)| \le w^*\}$

Complexity: 3 support points

Data-driven decision

Optimization:

$$\min_{x\in\mathcal{X}\subseteq\mathbb{R}^d} \quad c(x)$$

subject to: "scenario"

$$x \in \mathcal{X}_{\delta_i},$$

$$i = 1,\dots,N$$

Solution (decision): x^*

Complexity: at most d support constraints

Our predictor

Training:

$$\min_{(\theta_0,\theta_1,w)\in\mathbb{R}^3} \quad w$$

subject to:

$$|y_i - (\theta_0 + \theta_1 u_i)| \leq w,$$

$$i = 1, \ldots, N$$

Predictor: $P_N^* = \{|y - (\theta_0^* + \theta_1^* u)| \leq w^*\}$

Complexity: 3 support points

Misprediction:

$(u_{N+1}, y_{N+1}) : |y_{N+1} - (\theta_0^* + \theta_1^* u_{N+1})| > w^*$

Data-driven decision

Optimization:

$$\min_{x\in\mathcal{X}\subseteq\mathbb{R}^d} \quad c(x)$$

subject to:

"scenario"

$$x \in \mathcal{X}_{\delta_i},$$

$$i = 1, \ldots, N$$

Solution (decision): x^*

Complexity: at most d support constraints

Our predictor

Training:

$$\min_{(\theta_0,\theta_1,w)\in\mathbb{R}^3} w$$

subject to:

$$|y_i - (\theta_0 + \theta_1 u_i)| \leq w, \quad i = 1,\ldots,N$$

Predictor: $P_N^* = \{|y-(\theta_0^* + \theta_1^* u)| \leq w^*\}$

Complexity: 3 support points

Misprediction:

$$(u_{N+1}, y_{N+1}) : |y_{N+1} - (\theta_0^* + \theta_1^* u_{N+1})| > w^*$$

Data-driven decision

Optimization:

$$\min_{x\in\mathcal{X}\subseteq\mathbb{R}^d} c(x)$$

subject to:

"scenario"

$$x \in \mathcal{X}_{\delta_i}, \quad i = 1,\ldots,N$$

Solution (decision): x^*

Complexity: at most d support constraints

Constraint violation:

$$\delta_{N+1} : x^* \notin \mathcal{X}_{\delta_{N+1}}$$

Our predictor

Training:

$$\min_{(\theta_0,\theta_1,w)\in\mathbb{R}^3} \quad w$$

subject to:

$$|y_i - (\theta_0 + \theta_1 u_i)| \leq w,$$

$$i = 1, \dots, N$$

Predictor: $P_N^* = \{|y - (\theta_0^* + \theta_1^* u)| \leq w^*\}$

Complexity: 3 support points

Misprediction:

$$(u_{N+1}, y_{N+1}) : |y_{N+1} - (\theta_0^* + \theta_1^* u_{N+1})| > w^*$$

Guarantee: $PErr_N = \frac{3}{N+1}$

Data-driven decision

Optimization:

$$\min_{x\in\mathcal{X}\subseteq\mathbb{R}^d} \quad c(x)$$

subject to:

"scenario"

$$x \in \mathcal{X}_{\delta_i},$$

$$i = 1, \dots, N$$

Solution (decision): x^*

Complexity: at most d support constraints

Constraint violation:

$$\delta_{N+1} : x^* \notin \mathcal{X}_{\delta_{N+1}}$$

Our predictor

Training:

$$\min_{(\theta_0,\theta_1,w)\in\mathbb{R}^3} \quad w$$

subject to:

$$|y_i - (\theta_0 + \theta_1 u_i)| \le w,$$

$$i = 1, \ldots, N$$

Predictor: $P_N^* = \{|y - (\theta_0^* + \theta_1^* u)| \le w^*\}$

Complexity: 3 support points

Misprediction:

$$(u_{N+1}, y_{N+1}) : |y_{N+1} - (\theta_0^* + \theta_1^* u_{N+1})| > w^*$$

Guarantee: $PErr_N = \frac{3}{N+1}$

Data-driven decision

Optimization:

$$\min_{x\in\mathcal{X}\subseteq\mathbb{R}^d} \quad c(x)$$

subject to: "scenario"

$$x \in \mathcal{X}_{\delta_i},$$

$$i = 1, \ldots, N$$

Solution (decision): x^*

Complexity: at most d support constraints

Constraint violation:

$$\delta_{N+1} : x^* \notin \mathcal{X}_{\delta_{N+1}}$$

G. Calafiore, M.C. Campi *"Uncertain convex programs: randomized solutions and confidence levels"* Mathematical Programming 2005

Our predictor

Training:

$$\min_{(\theta_0,\theta_1,w)\in\mathbb{R}^3} \quad w$$

subject to:

$$|y_i - (\theta_0 + \theta_1 u_i)| \leq w,$$

$$i = 1,\ldots,N$$

Predictor: $P_N^* = \{|y - (\theta_0^* + \theta_1^* u)| \leq w^*\}$

Complexity: 3 support points

Misprediction:

$$(u_{N+1}, y_{N+1}) : |y_{N+1} - (\theta_0^* + \theta_1^* u_{N+1})| > w^*$$

Guarantee: $PErr_N = \frac{3}{N+1}$

Data-driven decision

Optimization:

$$\min_{x\in\mathcal{X}\subseteq\mathbb{R}^d} \quad c(x)$$

subject to: "scenario"

$$x \in \mathcal{X}_{\delta_i},$$

$$i = 1,\ldots,N$$

Solution (decision): x^*

Complexity: at most d support constraints

Constraint violation:

$$\delta_{N+1} : x^* \notin \mathcal{X}_{\delta_{N+1}}$$

Guarantee: $PErr_N \leq \frac{d}{N+1}$

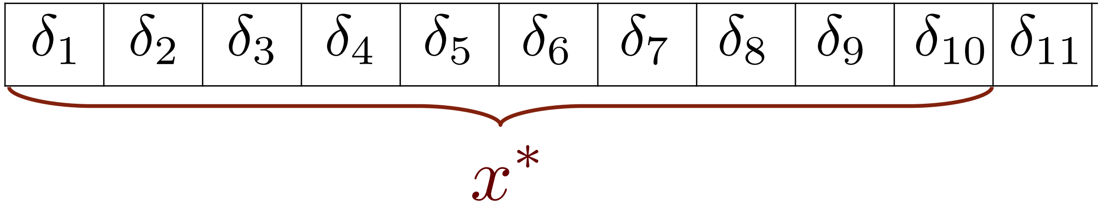
G. Calafiore, M.C. Campi *"Uncertain convex programs: randomized solutions and confidence levels"* Mathematical Programming 2005

Application to sequential decision-schemes

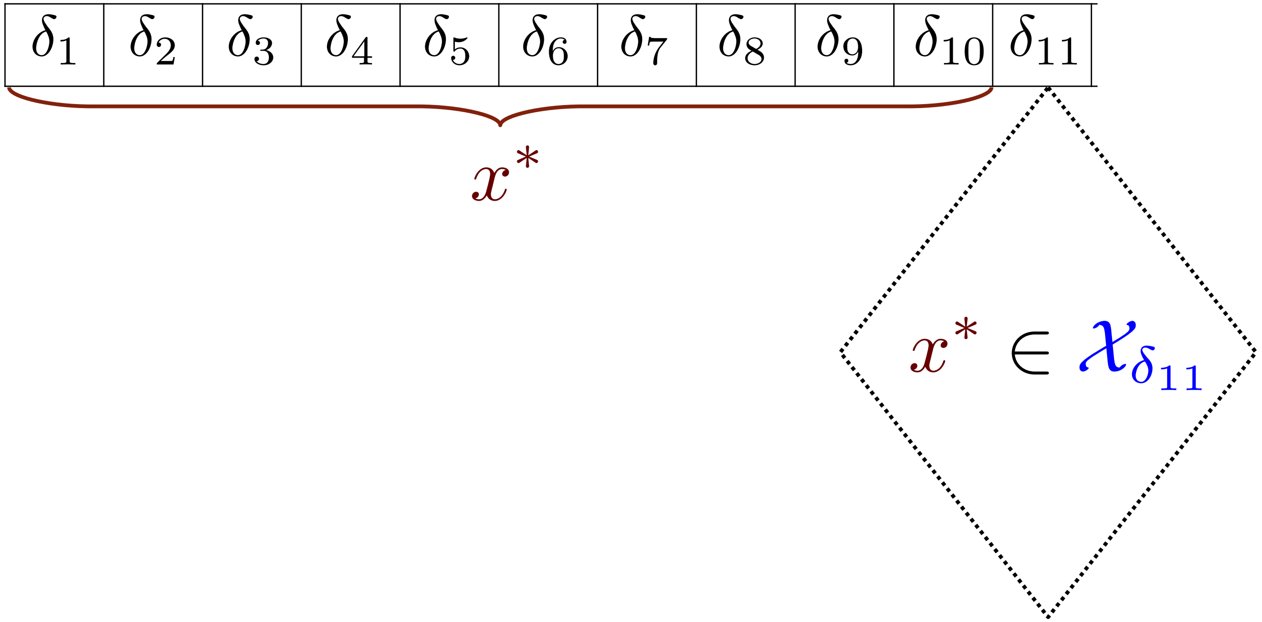
δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9	δ_{10}

$$\underbrace{\qquad\qquad\qquad\qquad}_{x^*}$$

Application to sequential decision-schemes



Application to sequential decision-schemes

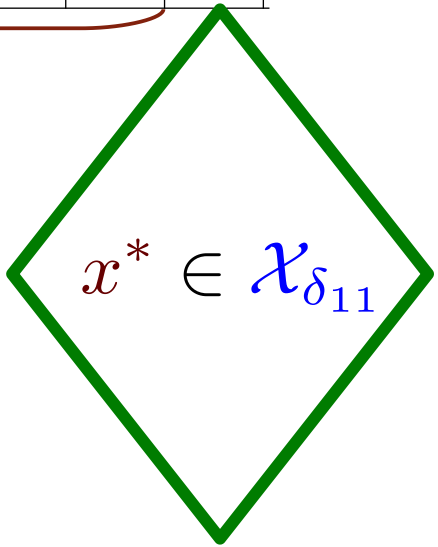


Result: ……………………………………………..

Application to sequential decision-schemes

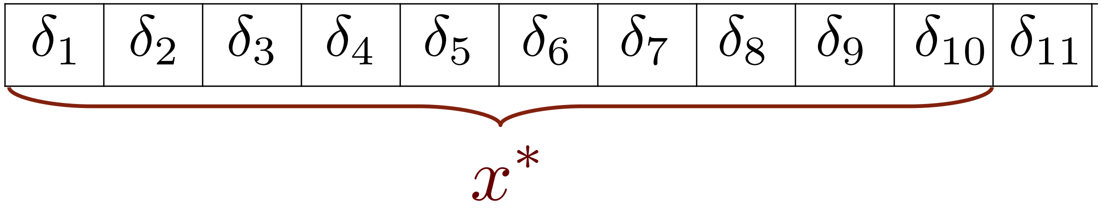
δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9	δ_{10}	δ_{11}

x^* (spans δ_1 to δ_{10})



Result: ✔ ……………………………………………..

Application to sequential decision-schemes



Result: ……………………✔…………………………………

Application to sequential decision-schemes

δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9	δ_{10}	δ_{11}

$$\underbrace{\delta_1 \cdots \delta_{10}}_{x^*}$$

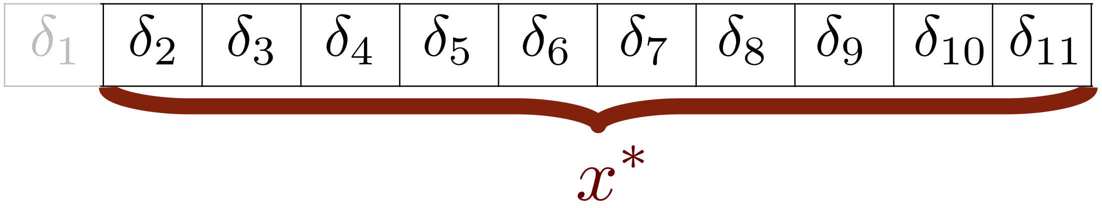
Result: ……………………………………………….. ✔

Application to sequential decision-schemes

δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9	δ_{10}	δ_{11}

Result: ✔ ………………………………………..

Application to sequential decision-schemes



Result: ✔ ……………………………………………..

Application to sequential decision-schemes

δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9	δ_{10}	δ_{11}

$\underbrace{\delta_2 \cdots \delta_{11}}_{x^*}$

Result: ………………………………………….. ✔

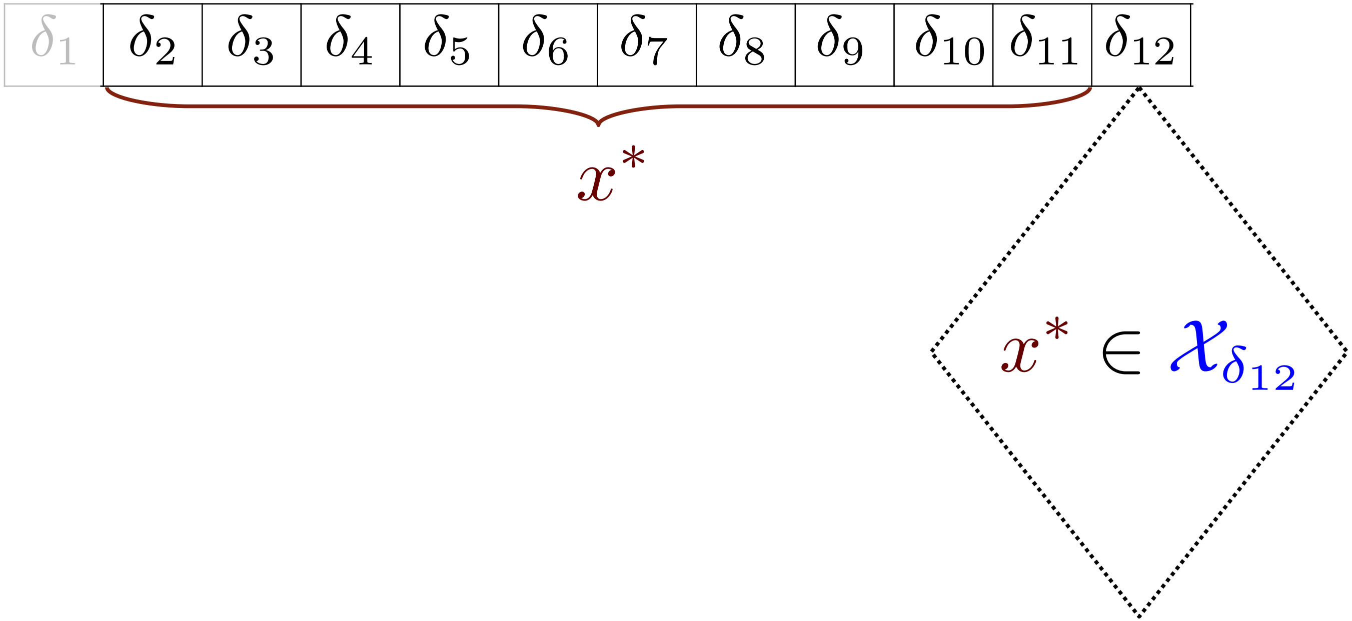
Application to sequential decision-schemes

δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9	δ_{10}	δ_{11}	δ_{12}

x^*

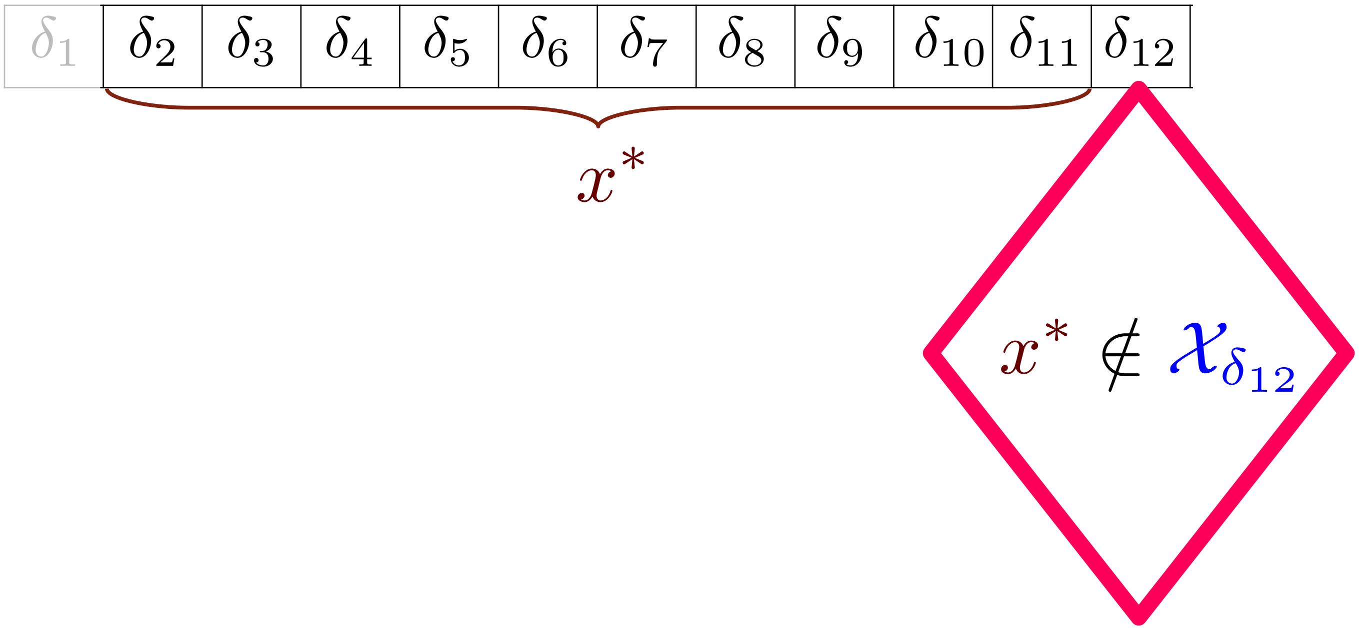
Result: ✔ ……………………………………………..

Application to sequential decision-schemes



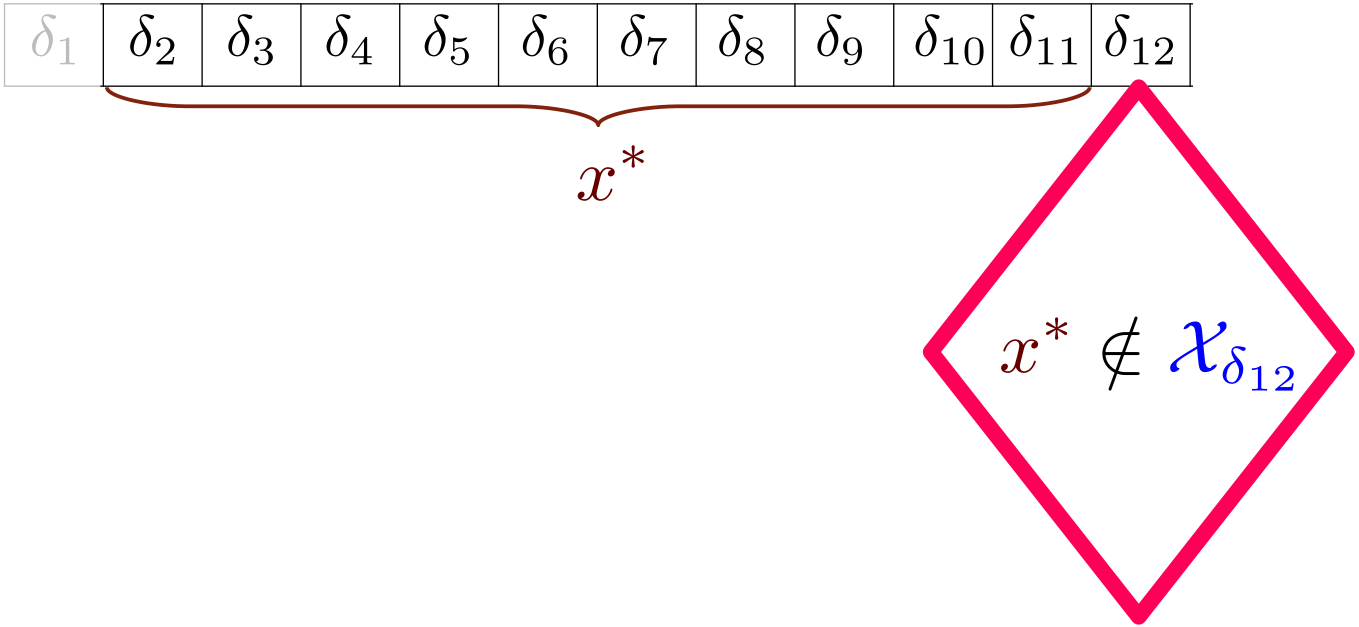
Result: ✔ ……………………………………………….

Application to sequential decision-schemes



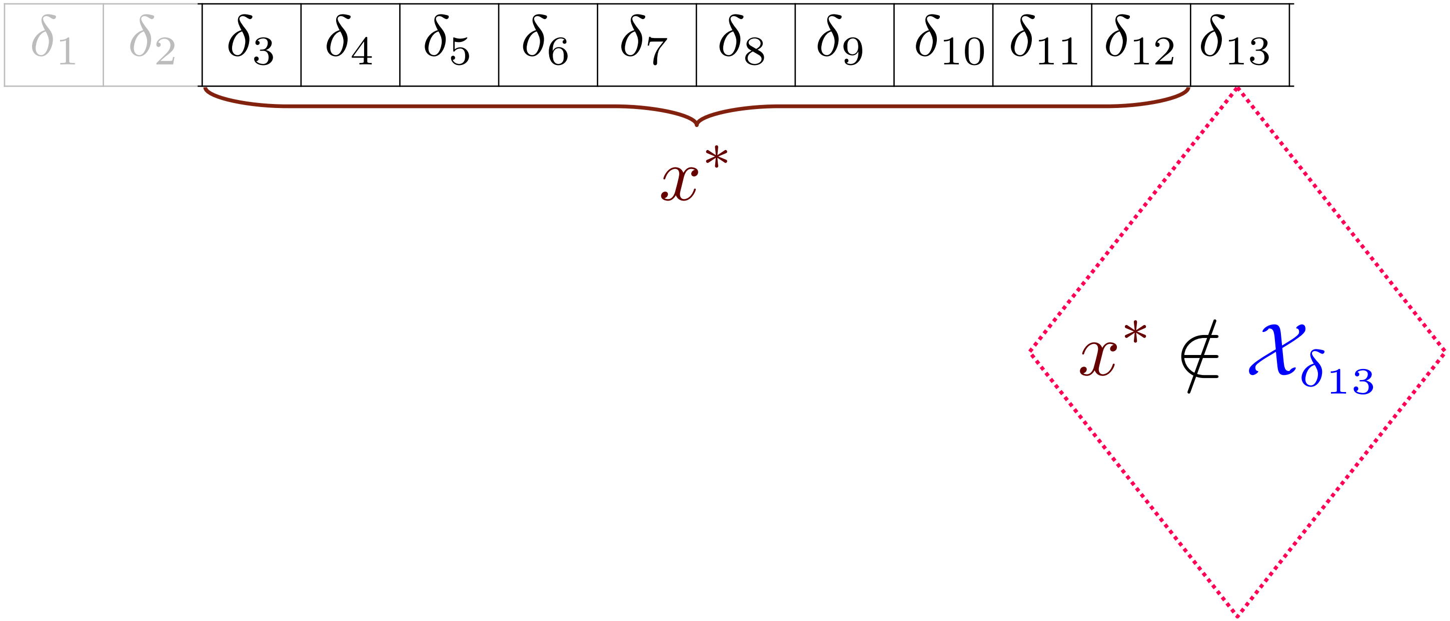
Result: ✔ ………………………………………………..

Application to sequential decision-schemes



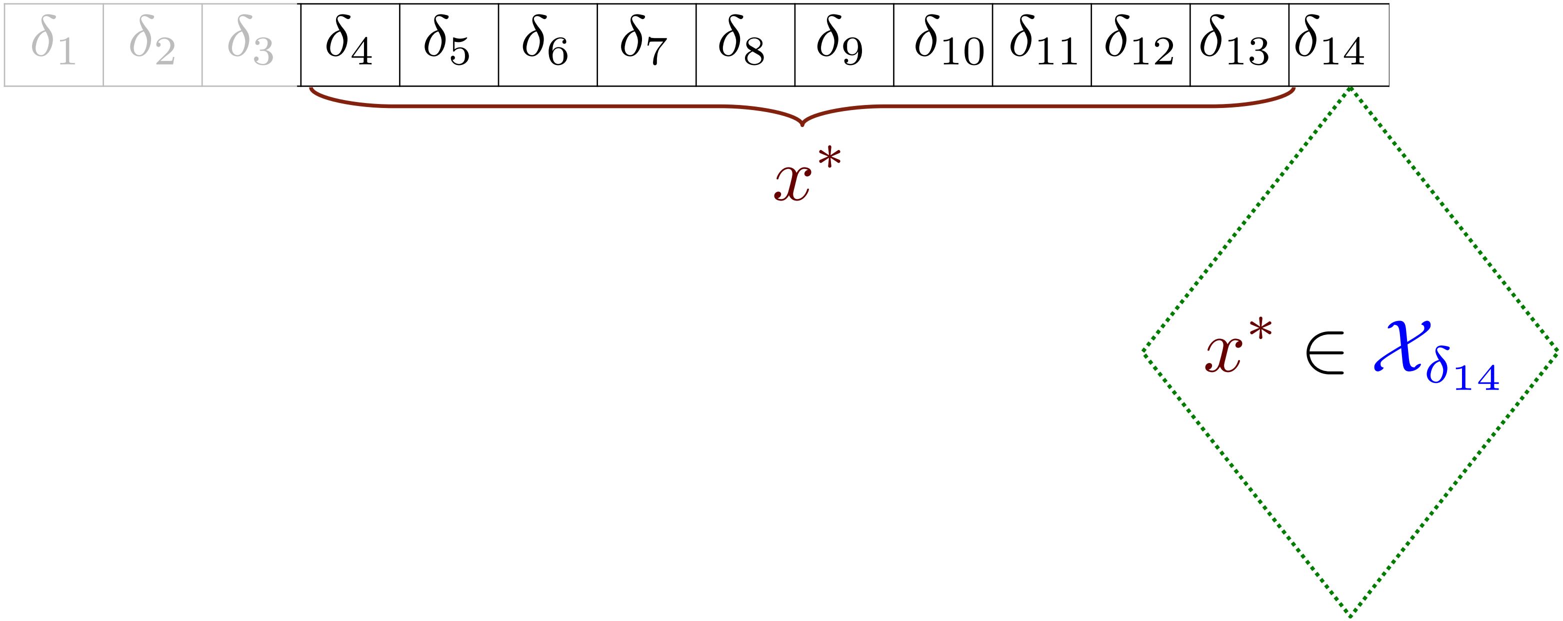
Result: ✔ ✘ ……………………………………………..

Application to sequential decision-schemes



Result: ✔ ✘ ✘ ……………………………………………….

Application to sequential decision-schemes



Result: ✔ ✘ ✘ ✔ ………………………………………

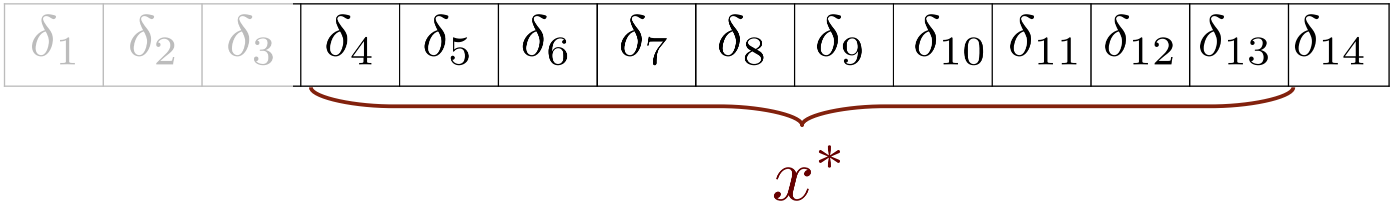
Application to sequential decision-schemes

δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9	δ_{10}	δ_{11}	δ_{12}	δ_{13}	δ_{14}

x^*

Result: ✔ ✘ ✘ ✔ ...

Application to sequential decision-schemes



Rate of errors

Result: ....✔ ✘ ✘ ✔ ....

Application to sequential decision-schemes

$$\delta_1 \;\; \delta_2 \;\; \delta_3 \;\; \underbrace{\delta_4 \;\; \delta_5 \;\; \delta_6 \;\; \delta_7 \;\; \delta_8 \;\; \delta_9 \;\; \delta_{10} \;\; \delta_{11} \;\; \delta_{12} \;\; \delta_{13}}_{x^*} \;\; \delta_{14}$$

$$\text{Rate of errors} \rightarrow PErr_N$$

Result: ✔ ✘ ✘ ✔ ...

Application to sequential decision-schemes

δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9	δ_{10}	δ_{11}	δ_{12}	δ_{13}	δ_{14}

x^* (over $\delta_4, \ldots, \delta_{13}$)

$$\text{Rate of errors} \rightarrow PErr_N \leq \frac{d}{N+1}$$

Result: ✔ ✘ ✘ ✔ ...

Example 1: Sequence of **investments**

F.A. Ramponi, M.C. Campi "*Expected shortfall: Heuristics and certificates*"
European Journal of Operational Research,2018

Example 1: Sequence of **investments**

δ_i rate of return at day i

F.A. Ramponi, M.C. Campi "*Expected shortfall: Heuristics and certificates*"
European Journal of Operational Research,2018

Example 1: Sequence of **investments**

δ_i rate of return at day i

x^* optimized portfolio of investments with guaranteed loss threshold

F.A. Ramponi, M.C. Campi "*Expected shortfall: Heuristics and certificates*"
European Journal of Operational Research,2018

Example 1: Sequence of **investments**

δ_i rate of return at day i

x^* optimized portfolio of investments with guaranteed loss threshold

$x^* \notin \mathcal{X}_{\delta_{N+1}}$ "shortfall"

F.A. Ramponi, M.C. Campi "*Expected shortfall: Heuristics and certificates*"
European Journal of Operational Research,2018

Example 1: Sequence of **investments**

δ_i rate of return at day i

x^* optimized portfolio of investments with guaranteed loss threshold

$x^* \notin \mathcal{X}_{\delta_{N+1}}$ "shortfall"

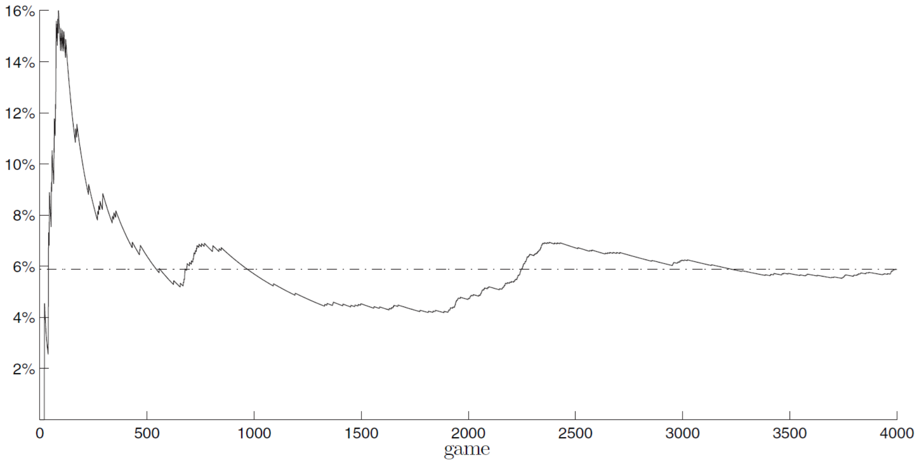
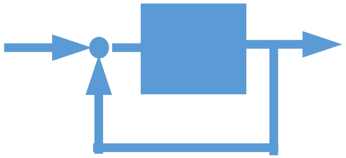


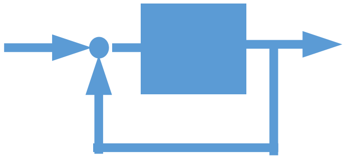
Fig. 6. Sliding window. Solid line (–) = average number of times when $L_{j+N+1}(x^*_{N,j}) > \bar{L}_{N,j}$; dashed-dotted line (–·) = 5.9% obtained from Theorem 4.1.

F.A. Ramponi, M.C. Campi "*Expected shortfall: Heuristics and certificates*"
European Journal of Operational Research,2018

Example 2: receding-horizon **control**



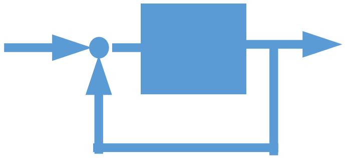
Example 2: receding-horizon **control**



δ_i realization of disturbances etc.

x^* control inputs

Example 2: receding-horizon **control**

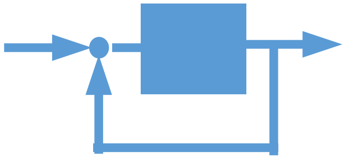


δ_i realization of disturbances etc.

x^* control inputs

$x^* \notin \mathcal{X}_{\delta_{N+1}}$ violation of the control constraints

Example 2: receding-horizon **control**



δ_i realization of disturbances etc.

x^* control inputs

$x^* \notin \mathcal{X}_{\delta_{N+1}}$ violation of the control constraints

Automatica 50 (2014) 3009–3018

Contents lists available at ScienceDirect

Automatica

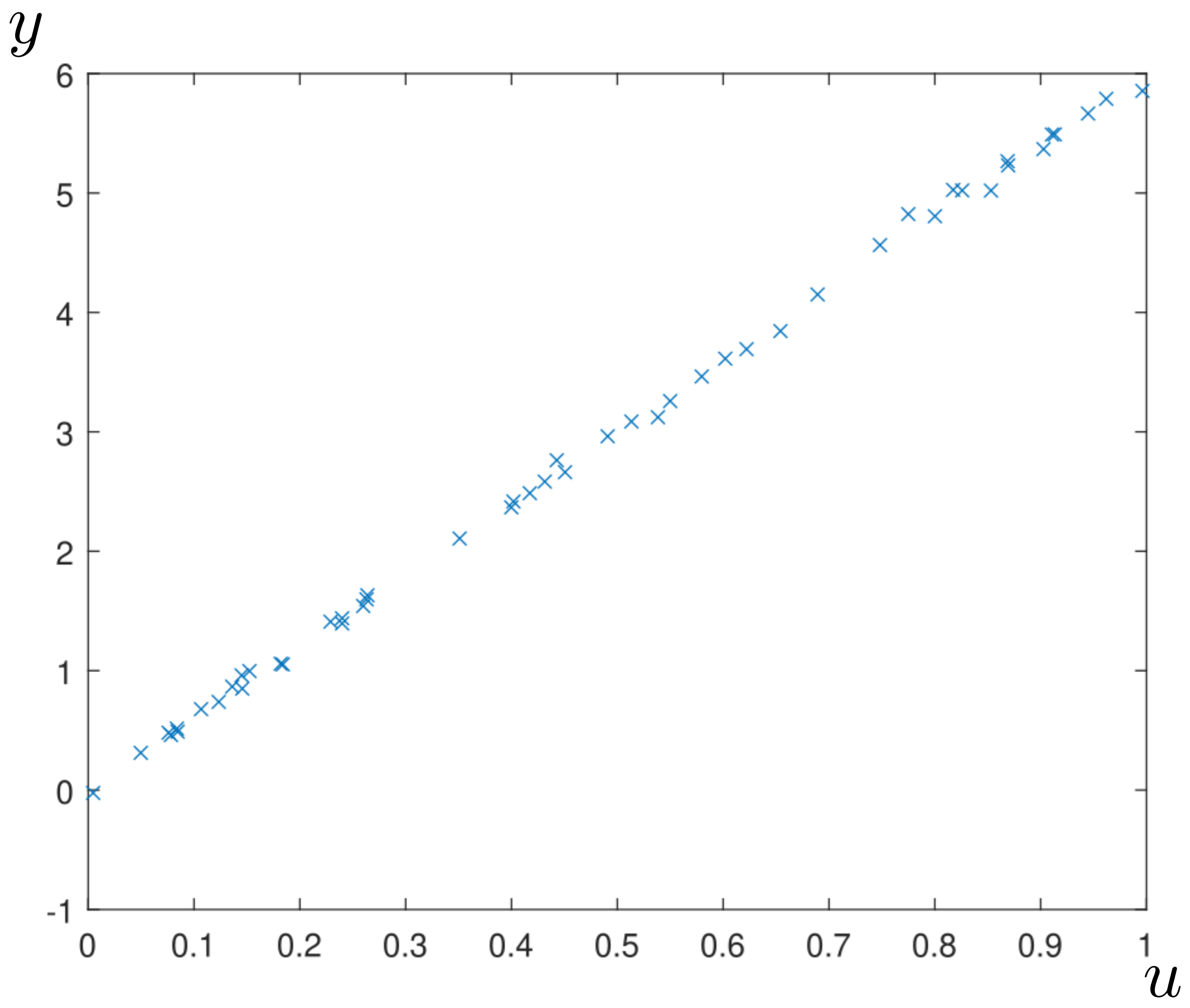
journal homepage: www.elsevier.com/locate/automatica

ELSEVIER

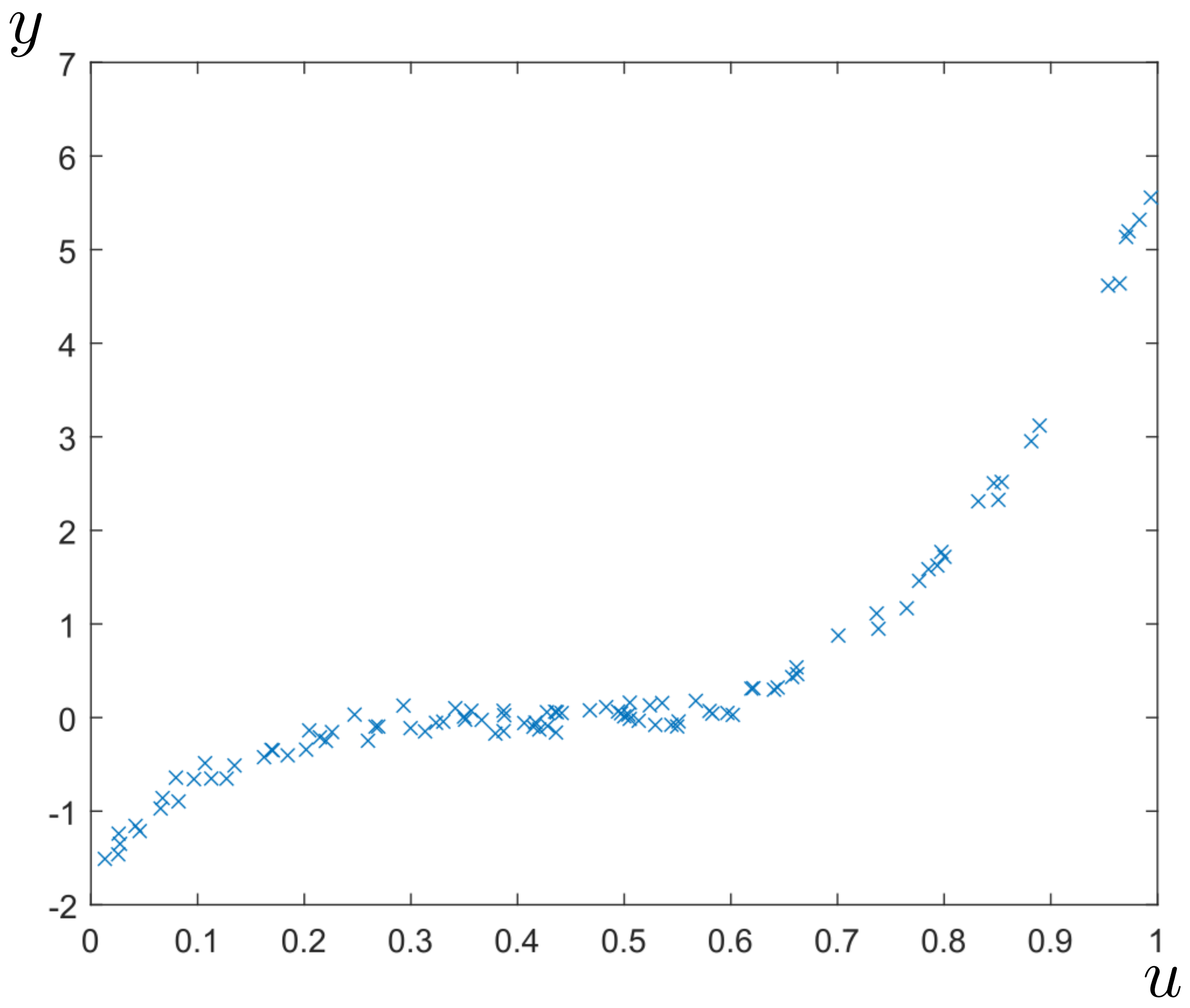
The scenario approach for Stochastic Model Predictive Control with bounds on closed-loop constraint violations☆

Georg Schildbach[a,1], Lorenzo Fagiano[a,b], Christoph Frei[c], Manfred Morari[a]

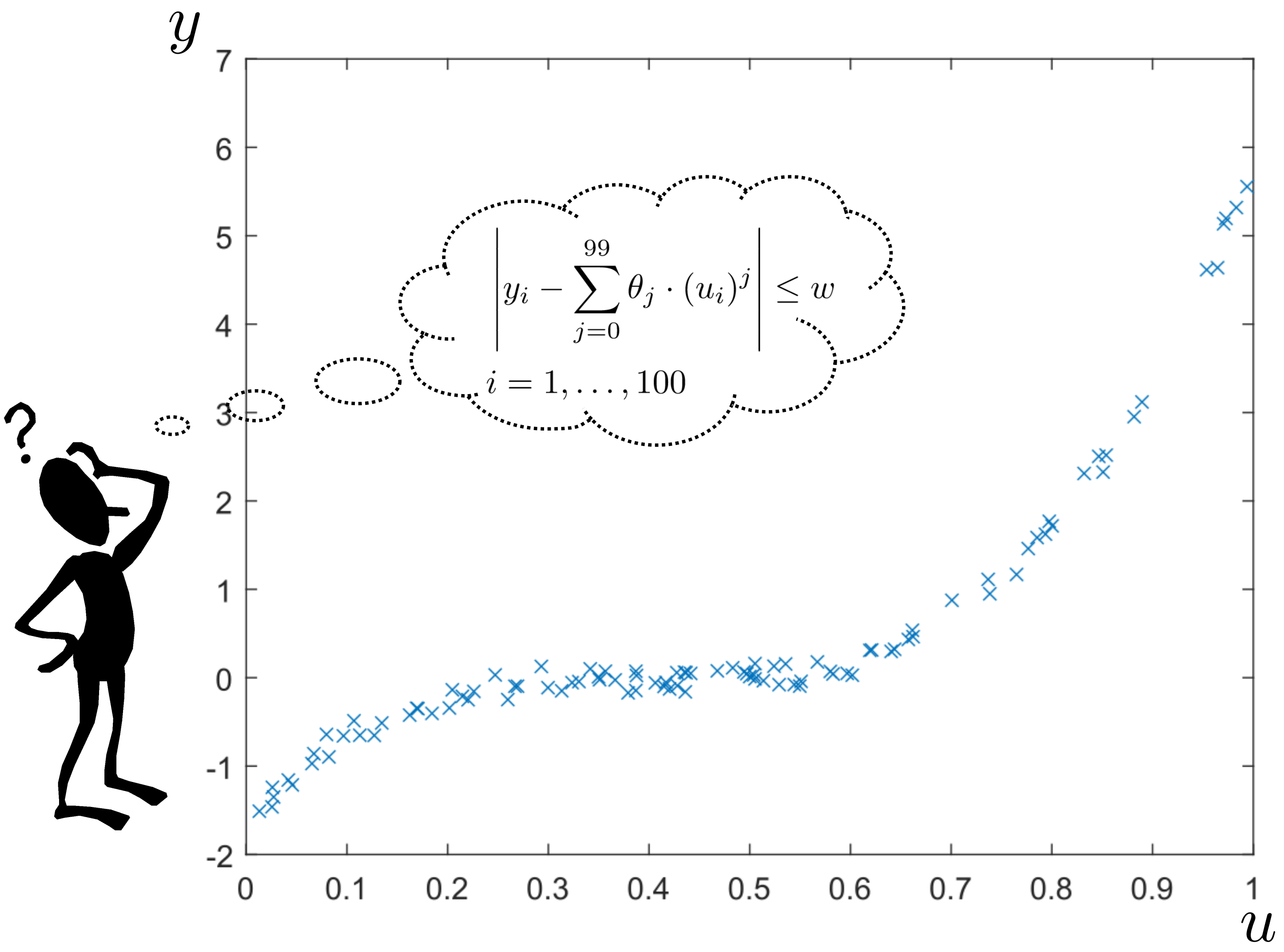
[a] Automatic Control Laboratory, Swiss Federal Institute of Technology Zurich, Physikstrasse 3, 8092 Zurich, Switzerland
[b] ABB Switzerland Ltd., Corporate Research, Segelhofstrasse 1, Baden-Daettwil, Switzerland
[c] Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB T6G 2G1, Canada



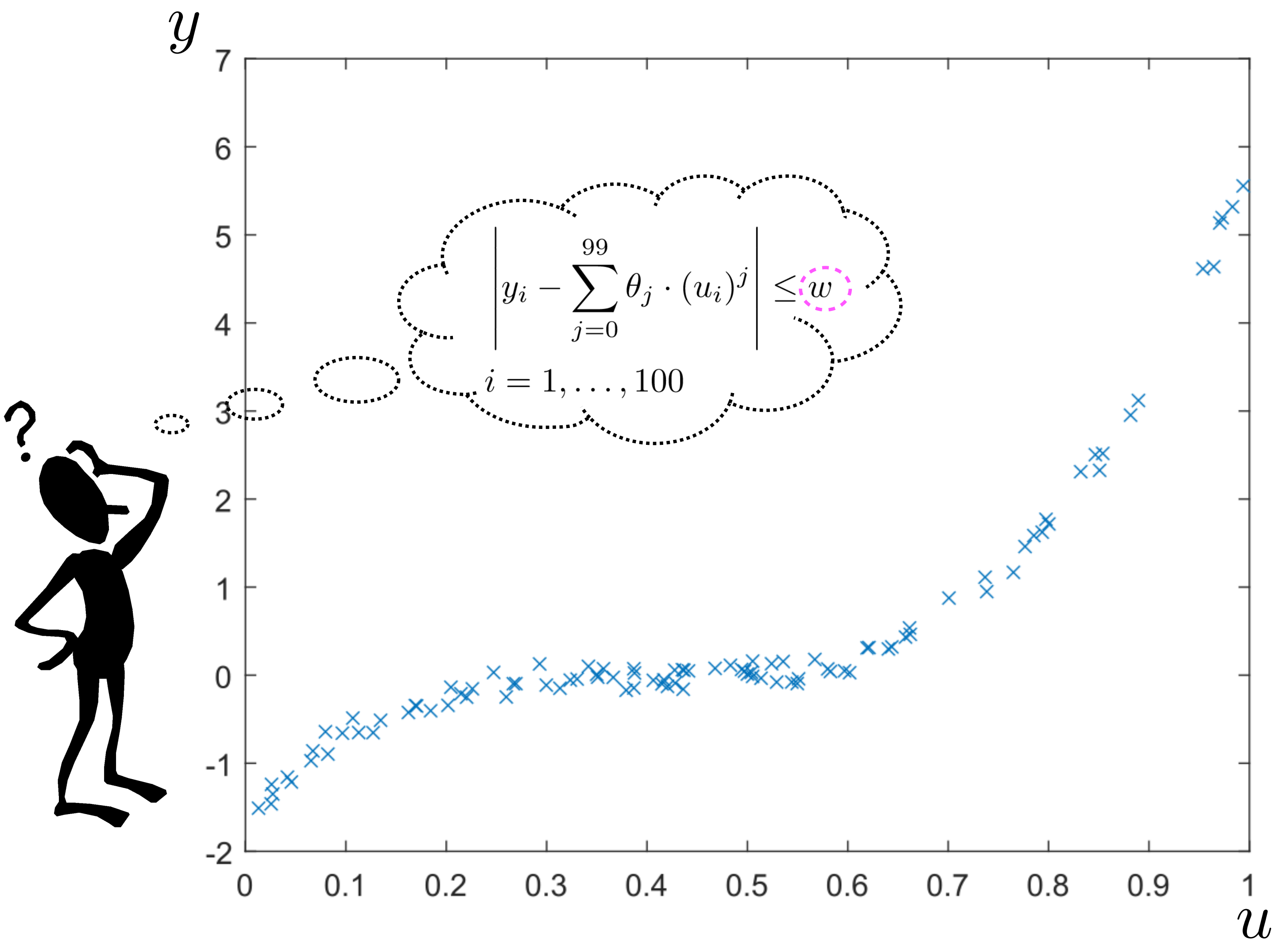
y
6
5
4
3
2
1
0
-1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
u

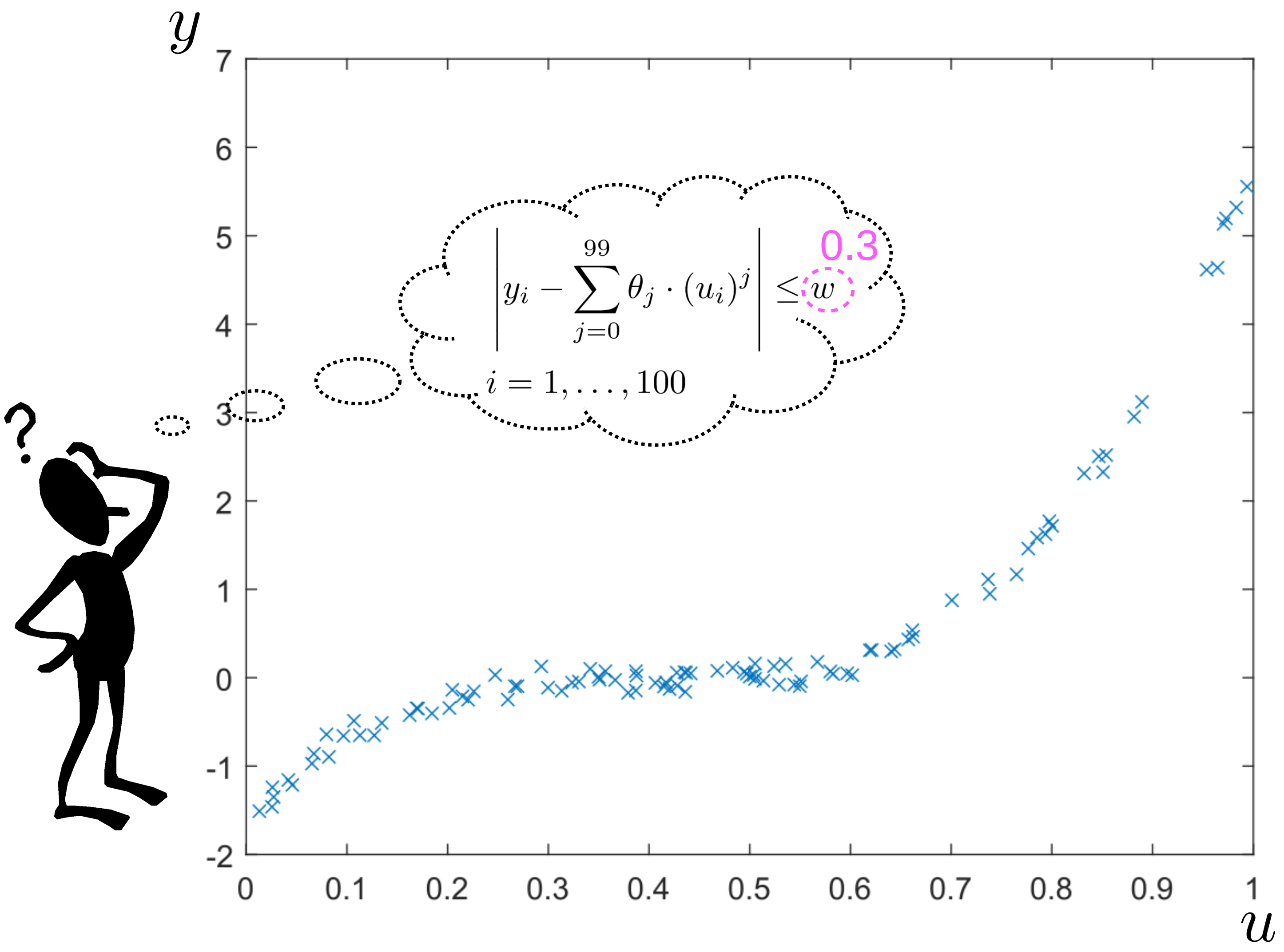


y
7
6
5
4
3
2
1
0
-1
-2
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
u

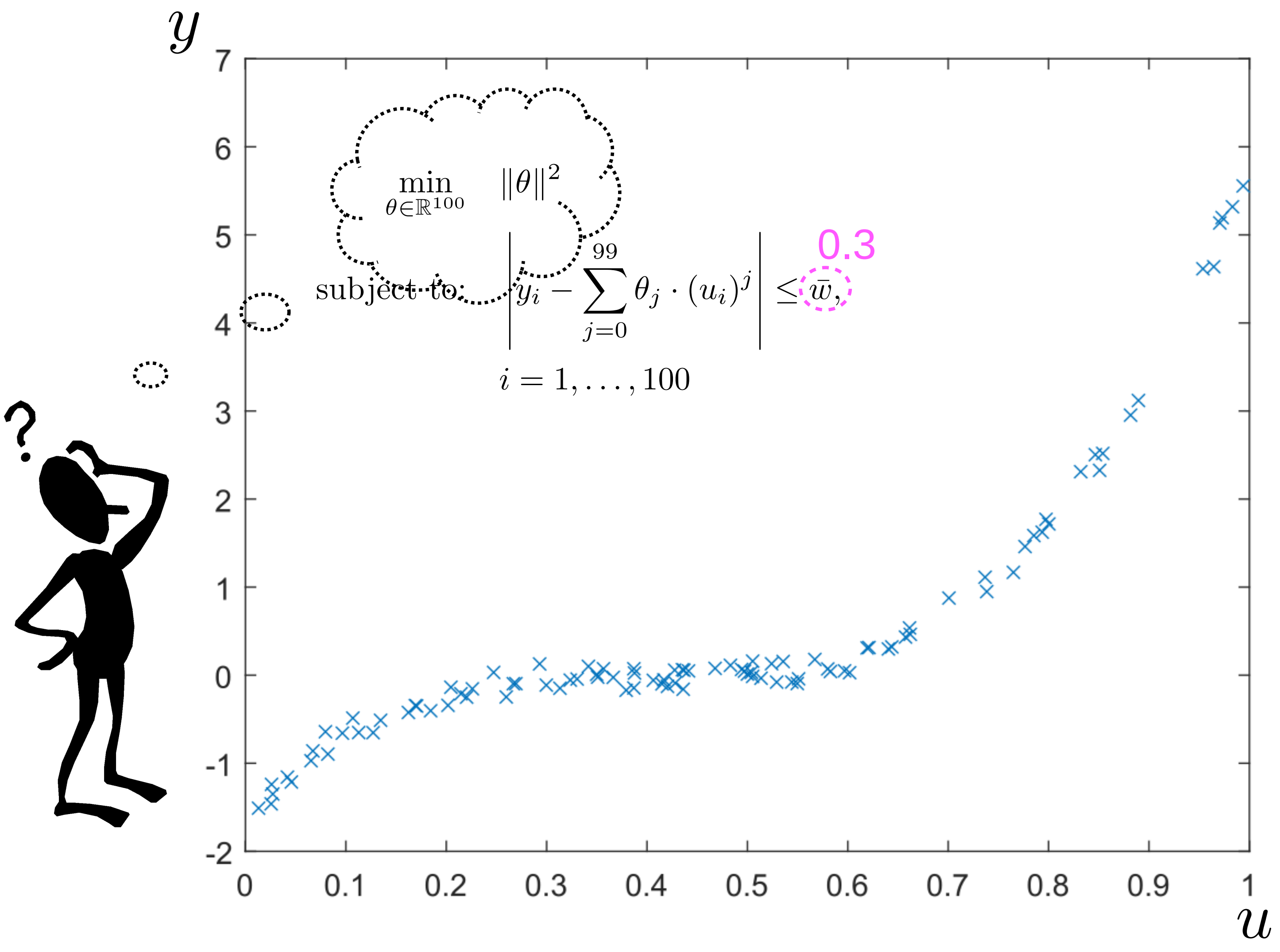


y
7
6
5
4
3
2
1
0
-1
-2
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
u
$$\left| y_i - \sum_{j=0}^{99} \theta_j \cdot (u_i)^j \right| \leq w$$
$i = 1, \ldots, 100$

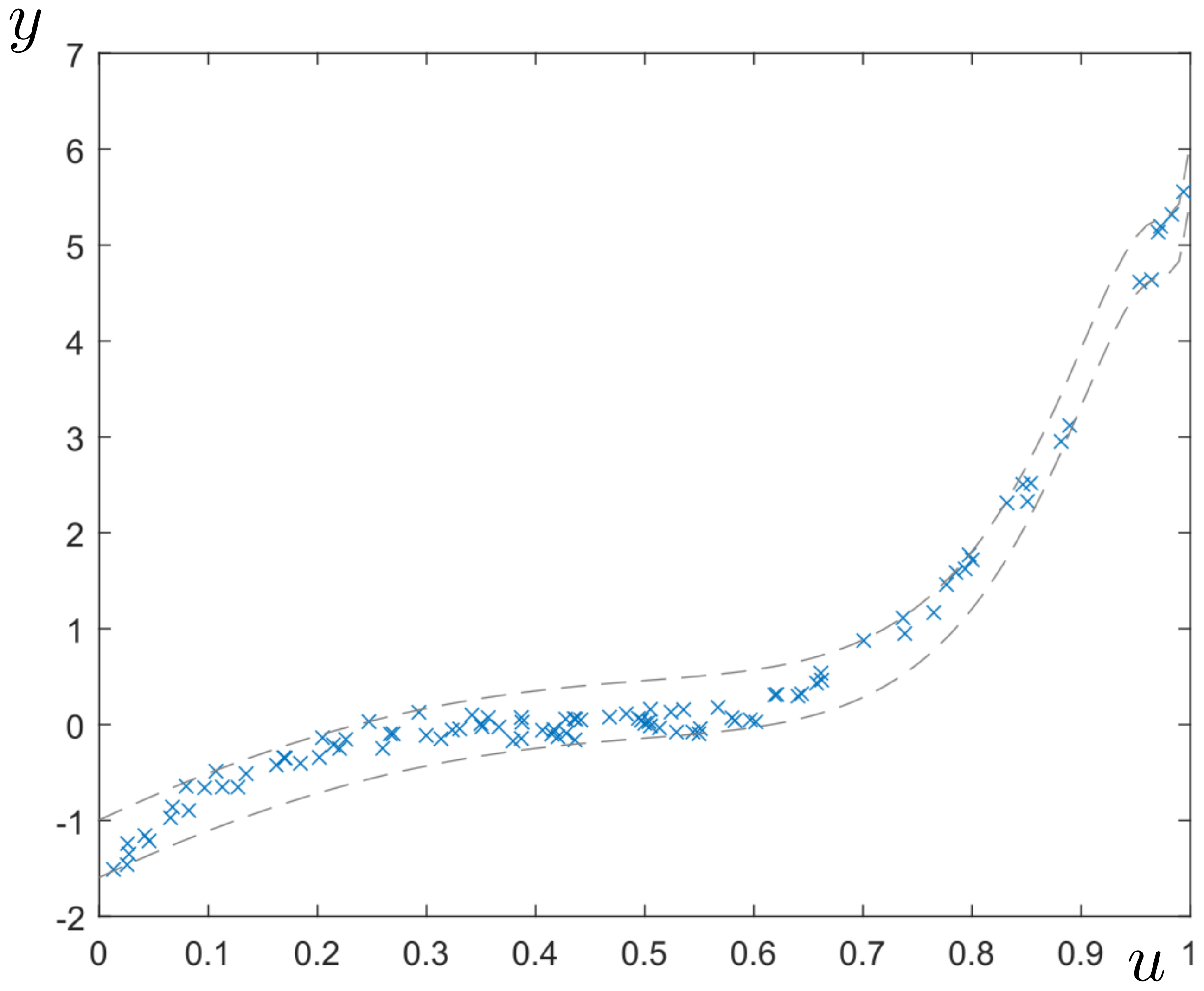

y
$\left| y_i - \sum_{j=0}^{99} \theta_j \cdot (u_i)^j \right| \leq w$
$i = 1, \ldots, 100$
7
6
5
4
3
2
1
0
-1
-2
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
u

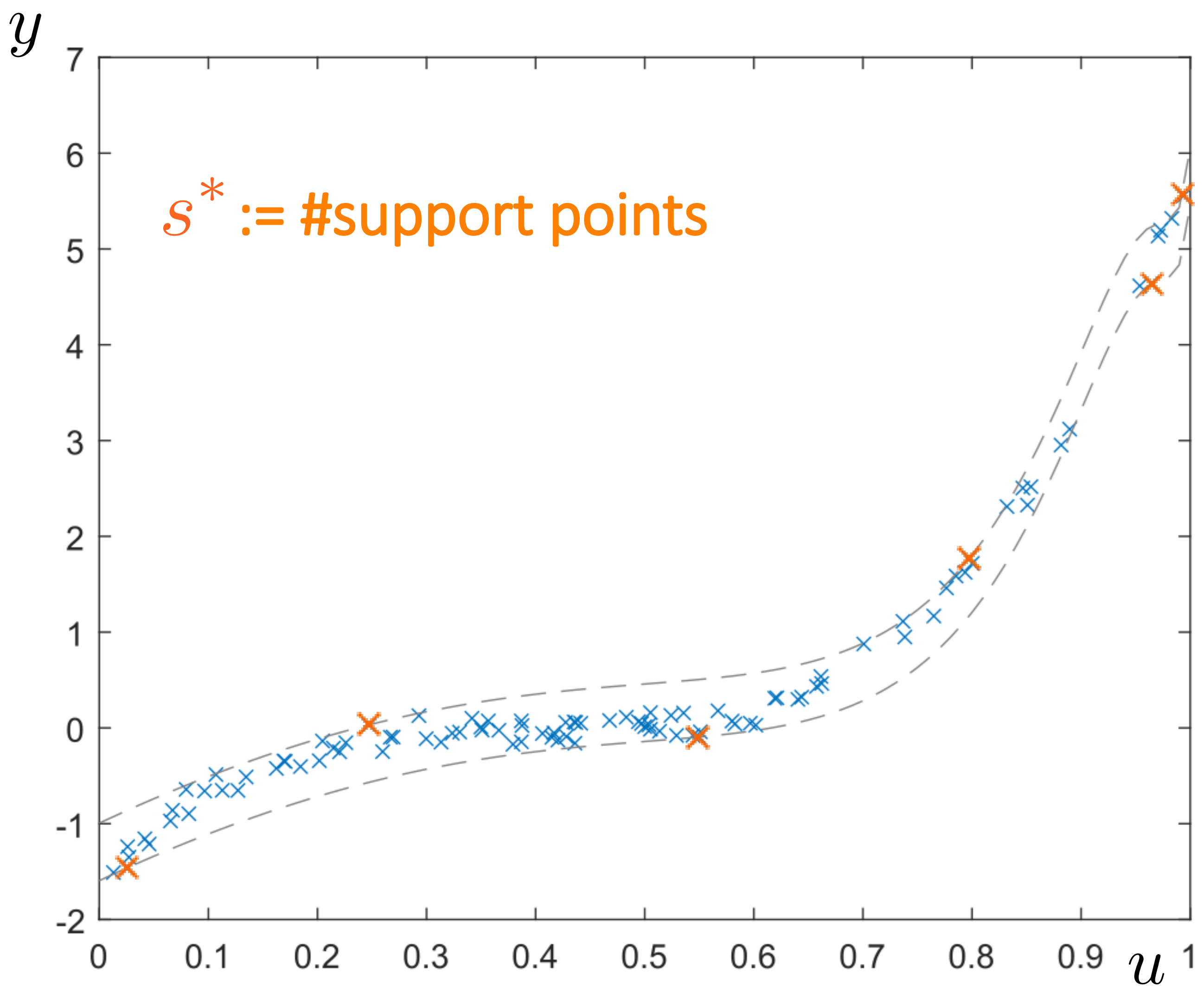


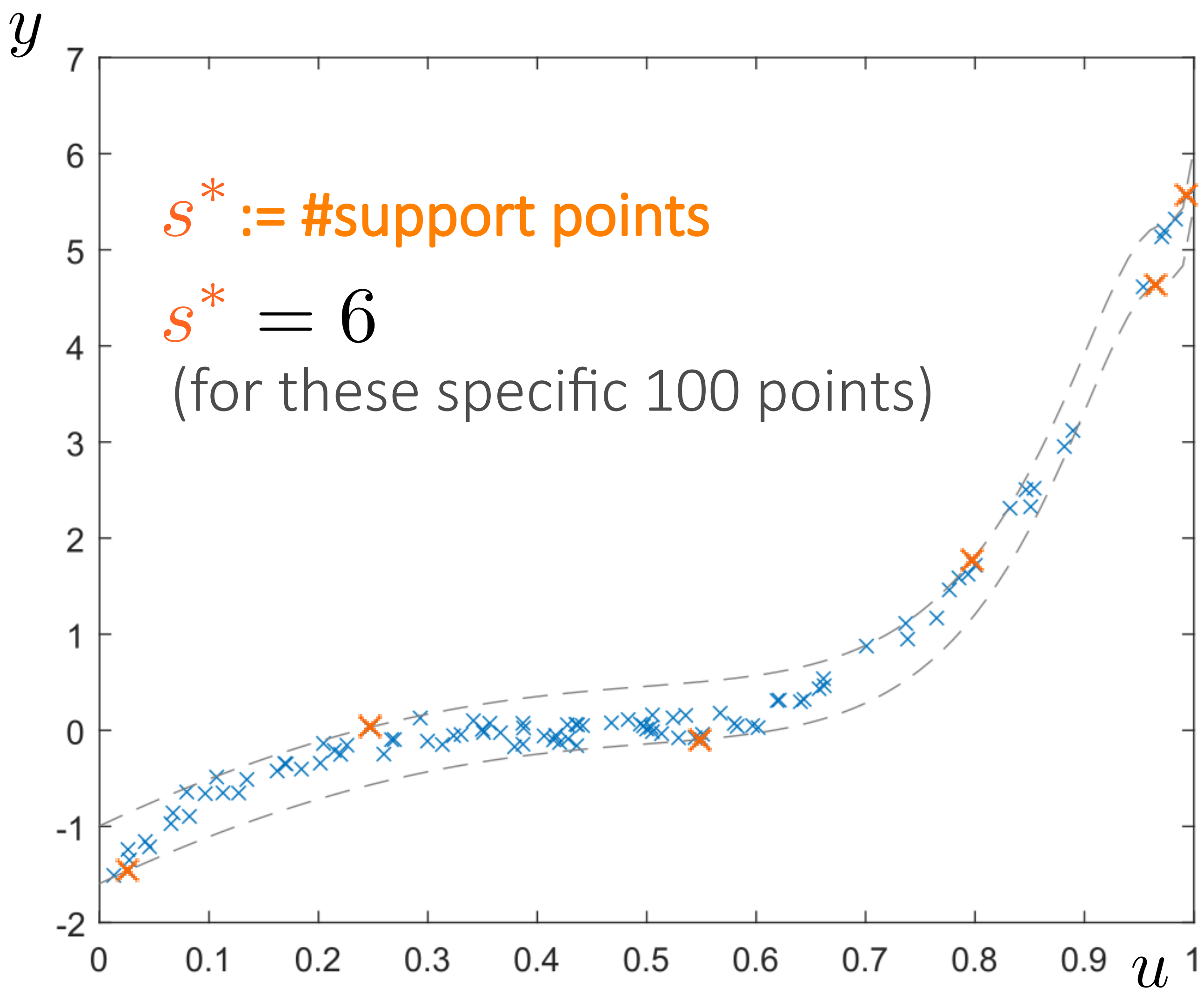
y
7
6
5
4
3
2
1
0
-1
-2
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
u
$\left| y_i - \sum_{j=0}^{99} \theta_j \cdot (u_i)^j \right| \leq w$
0.3
$i = 1, \ldots, 100$
?



$\min_{\theta \in \mathbb{R}^{100}} \quad \|\theta\|^2$
subject to: $\left| y_i - \sum_{j=0}^{99} \theta_j \cdot (u_i)^j \right| \leq \bar{w},$
$i = 1, \ldots, 100$
0.3
y
u


y
7
6
5
4
3
2
1
0
-1
-2
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
u


y
7
6
5
4
3
2
1
0
-1
-2
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
u
s^* := #support points

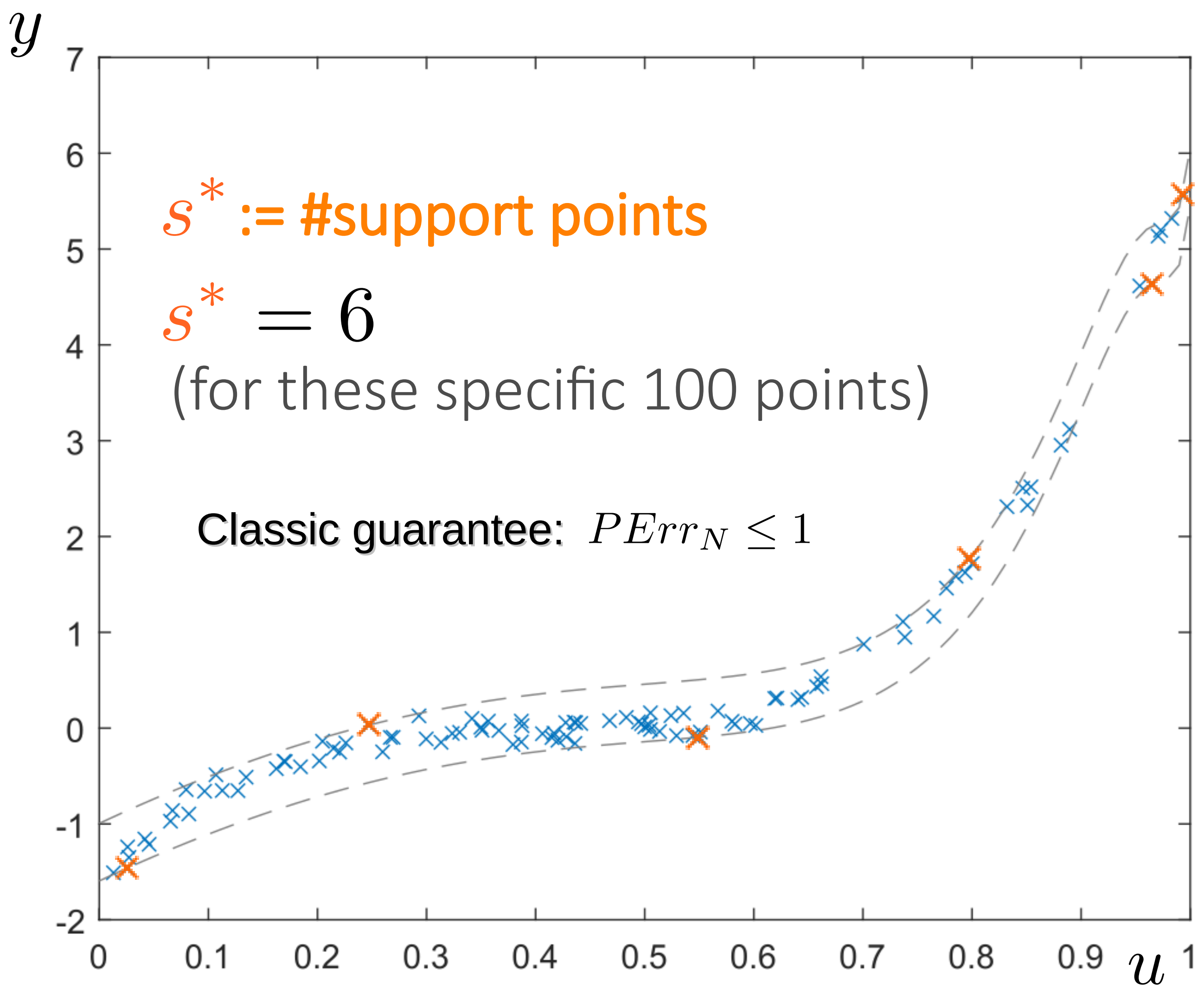

y
7
6
5
4
3
2
1
0
-1
-2
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
u
s^* := #support points
$s^* = 6$
(for these specific 100 points)

s^* := #support points

$s^* = 6$

(for these specific 100 points)

Classic guarantee: $PErr_N \leq 1$



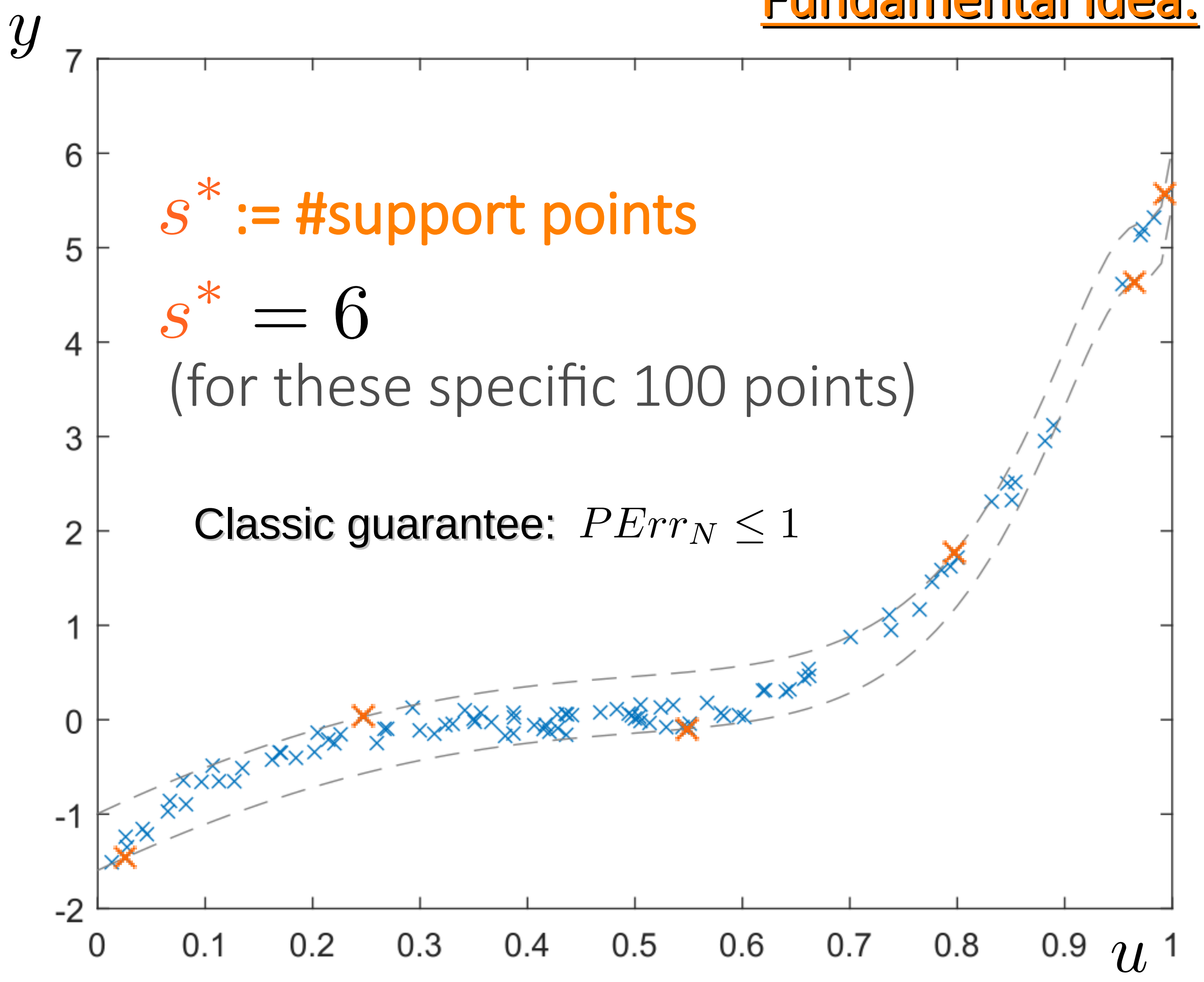
Fundamental idea:

s^* := #support points

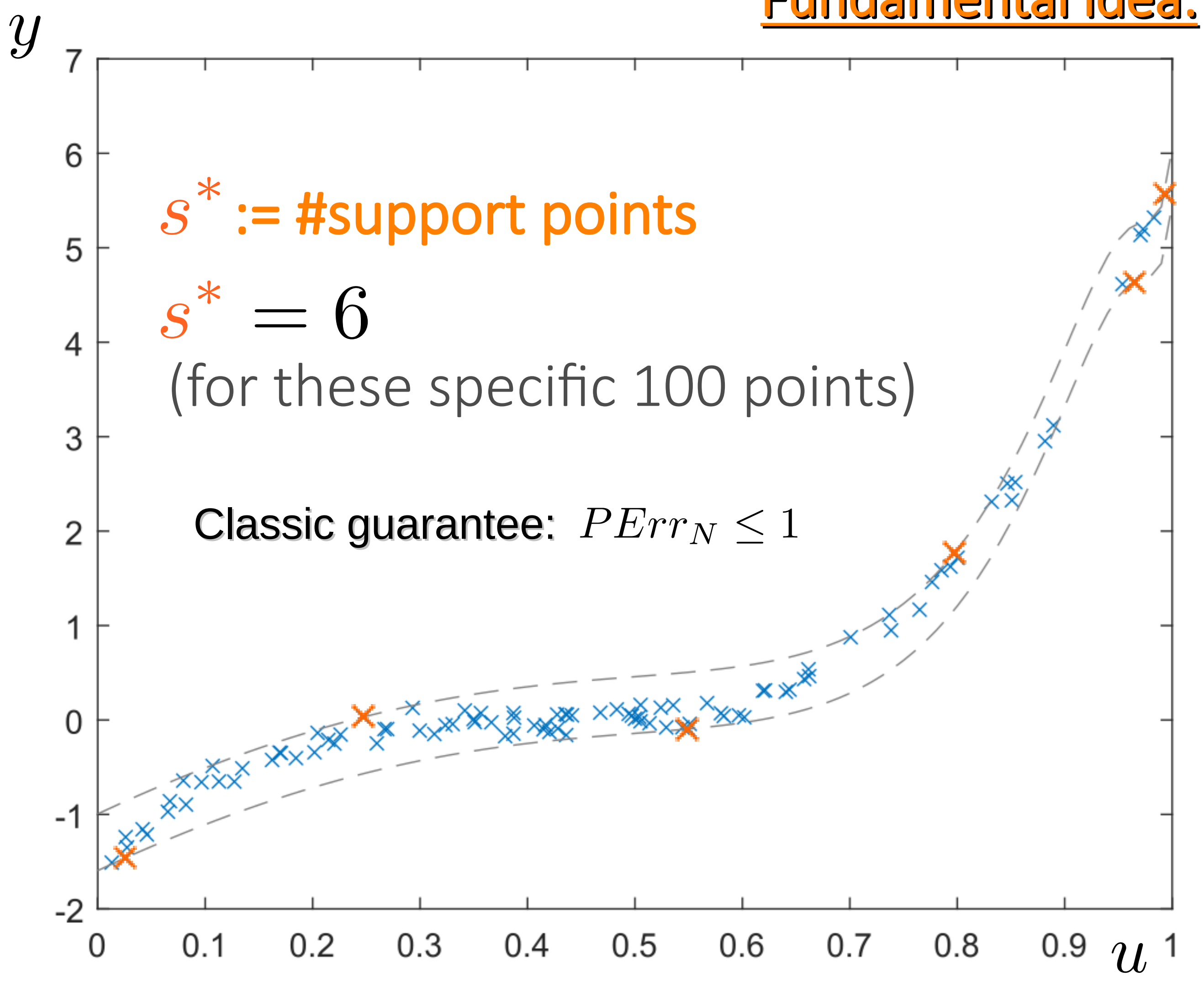
$s^* = 6$

(for these specific 100 points)

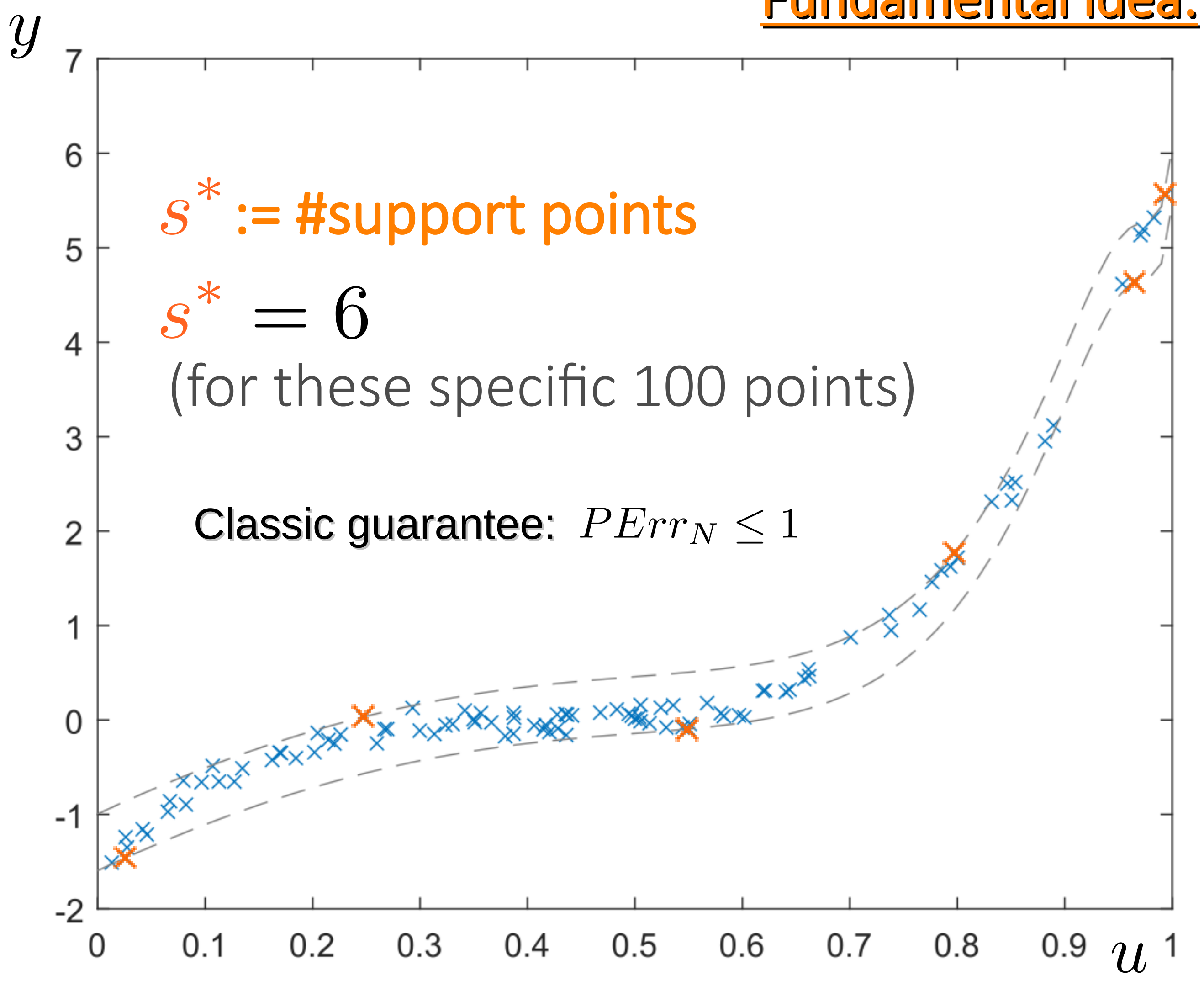
Classic guarantee: $PErr_N \leq 1$



Fundamental idea: s^* reveals



Fundamental idea: s^* reveals if our predictor



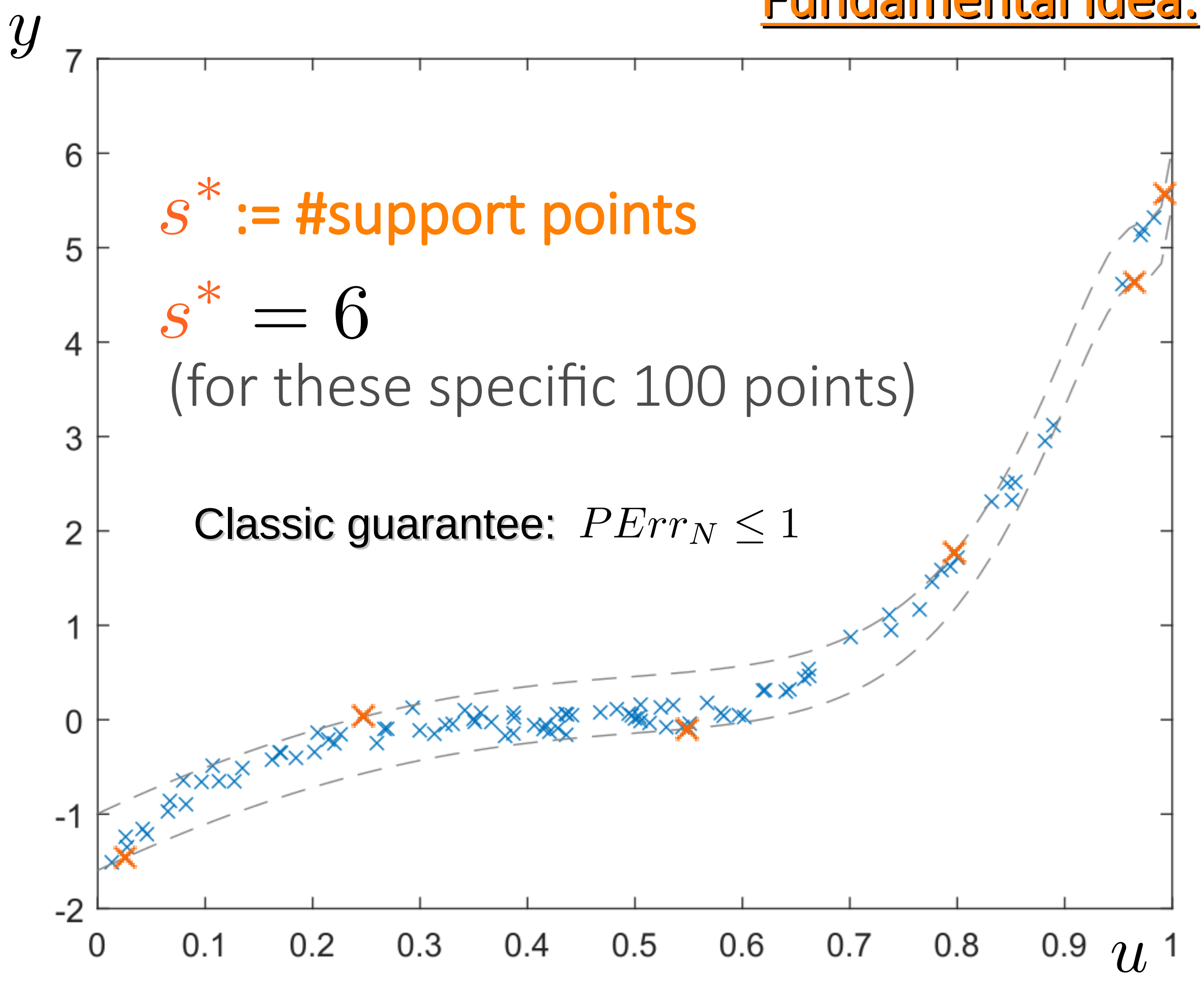
s^* := #support points

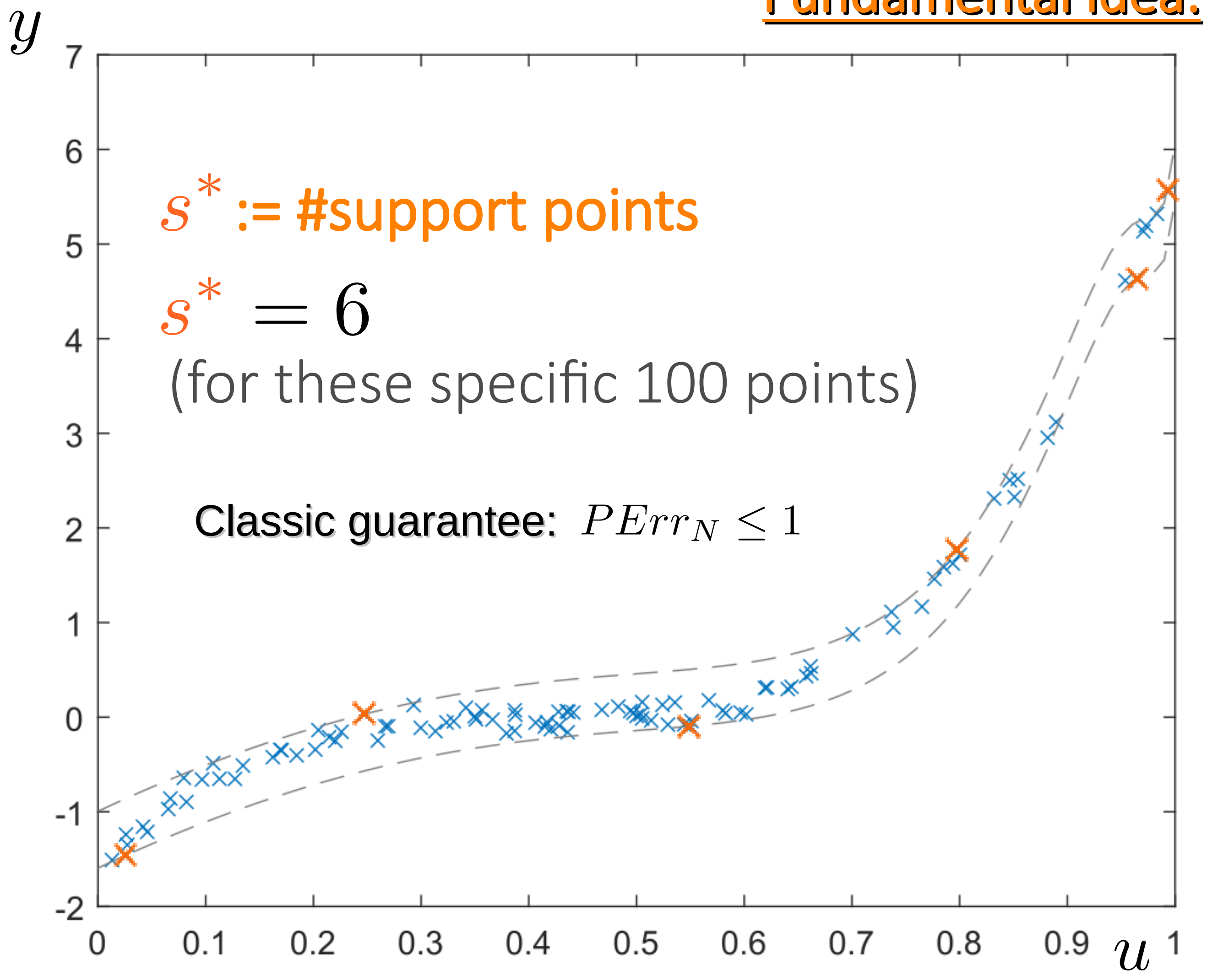
$s^* = 6$

(for these specific 100 points)

Classic guarantee: $PErr_N \leq 1$

Fundamental idea: s^* **reveals** if our predictor aligns with the **unknown** data-generating mechanism





s^* := #support points

$s^* = 6$

(for these specific 100 points)

Classic guarantee: $PErr_N \leq 1$

Fundamental idea: s^* reveals if our predictor aligns with the **unknown** data-generating mechanism

Proof of the claim:

S. Garatti, M.C. Campi

"**Risk and complexity** in scenario optimization"

Mathematical Programming 2022

A risk-averse prediction scheme

Initialization: set a complexity threshold $\bar{k}$

A risk-averse prediction scheme

Initialization: set a complexity threshold $\bar{k}$

Execution:

1) get **N data**, train the predictor and compute its s^*

A risk-averse prediction scheme

Initialization: set a complexity threshold $\bar{k}$

Execution:

1) get **N data**, train the predictor and compute its s^*

2) IF $s^* \leq \bar{k}$ THEN **USE the predictor** to predict the (N+1)-th data point

A risk-averse prediction scheme

Initialization: set a complexity threshold $\bar{k}$

Execution:

1) get **N data**, train the predictor and compute its s^*

2) IF $s^* \leq \bar{k}$ THEN **USE the predictor** to predict the (N+1)-th data point

   OTHERWISE **USE an ORACLE*** that is ALWAYS RIGHT

A risk-averse prediction scheme

Initialization: set a complexity threshold $\bar{k}$

Execution:
1) get **N data**, train the predictor and compute its s^*

2) IF $s^* \leq \bar{k}$ THEN **USE the predictor** to predict the (N+1)-th data point

OTHERWISE **USE an ORACLE*** that is ALWAYS RIGHT

***ORACLE:**

A risk-averse prediction scheme

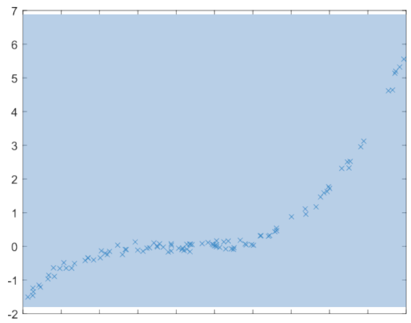
Initialization: set a complexity threshold $\bar{k}$

Execution:
1) get **N data**, train the predictor and compute its s^*

2) IF $s^* \leq \bar{k}$ THEN **USE the predictor** to predict the (N+1)-th data point

OTHERWISE **USE an ORACLE*** that is ALWAYS RIGHT

***ORACLE:** Large prediction band



A risk-averse prediction scheme

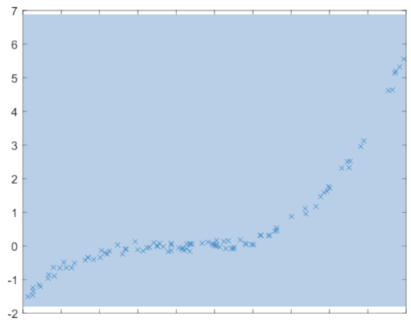
Initialization: set a complexity threshold $\bar{k}$

Execution:
1) get **N data**, train the predictor and compute its s^*

2) IF $s^* \leq \bar{k}$ THEN **USE the predictor** to predict the (N+1)-th data point

OTHERWISE **USE an ORACLE*** that is ALWAYS RIGHT

***ORACLE:** Large prediction band



"do not invest"

Ex. 1

A risk-averse prediction scheme

Initialization: set a complexity threshold $\bar{k}$

Execution:

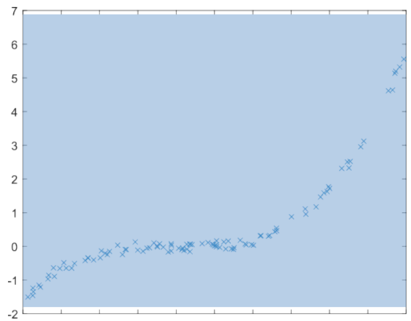
1) get **N data**, train the predictor and compute its s^*

2) IF $s^* \leq \bar{k}$ THEN **USE the predictor** to predict the (N+1)-th data point

OTHERWISE **USE an ORACLE*** that is ALWAYS RIGHT

***ORACLE:**

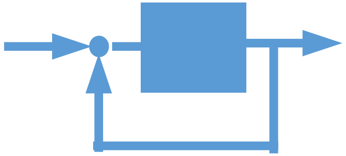
Large prediction band



"do not invest"

Ex. 1

backup control policy



Ex. 2

$$PErr_N = Prob\{(u_1, y_1), \ldots, (u_N, y_N), (u_{N+1}, y_{N+1}) :$$

$$(u_{N+1}, y_{N+1}) \notin P_N^*\}$$

$$PErr_N = Prob\{(u_1, y_1), \ldots, (u_N, y_N), (u_{N+1}, y_{N+1}) :$$

$$s^* \leq \bar{k} \text{ AND } (u_{N+1}, y_{N+1}) \notin P_N^*\}$$

$$PErr_N = Prob\{(u_1, y_1), \dots, (u_N, y_N), (u_{N+1}, y_{N+1}) :$$

$$s^* \leq \bar{k} \text{ AND } (u_{N+1}, y_{N+1}) \notin P_N^*\}$$

$$PErr_N = Prob\{(u_1, y_1), \dots, (u_N, y_N), (u_{N+1}, y_{N+1}) :$$

$$s^* \leq \bar{k} \text{ AND } (u_{N+1}, y_{N+1}) \notin P_N^*\}$$

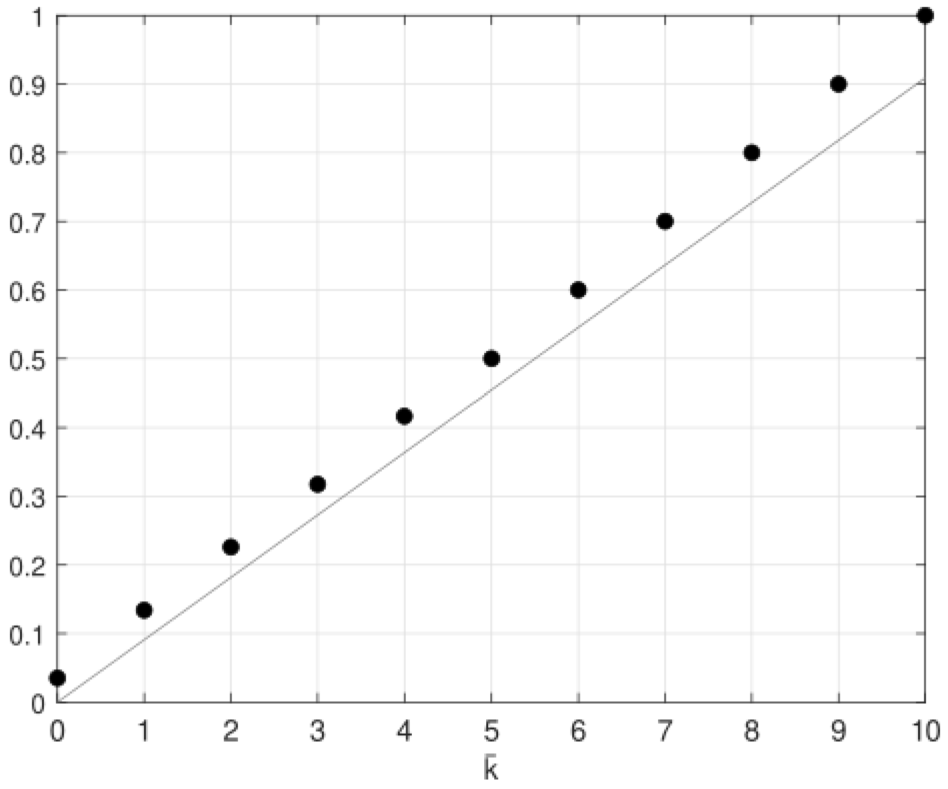
$$PErr_N \leq \frac{\bar{k}}{N+1}$$

$$PErr_N = Prob\{(u_1, y_1), \ldots, (u_N, y_N), (u_{N+1}, y_{N+1}) :$$

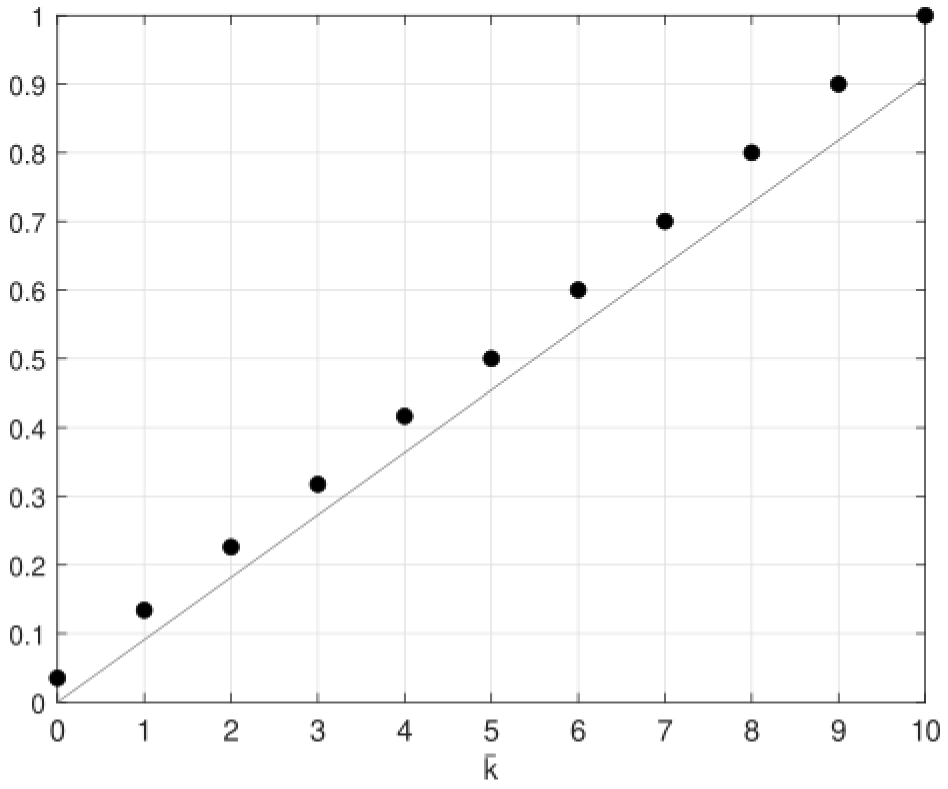
$$s^* \leq \bar{k} \text{ AND } (u_{N+1}, y_{N+1}) \notin P_N^*\}$$

FALSE

$$PErr_N \leq \frac{\bar{k}}{N+1}$$

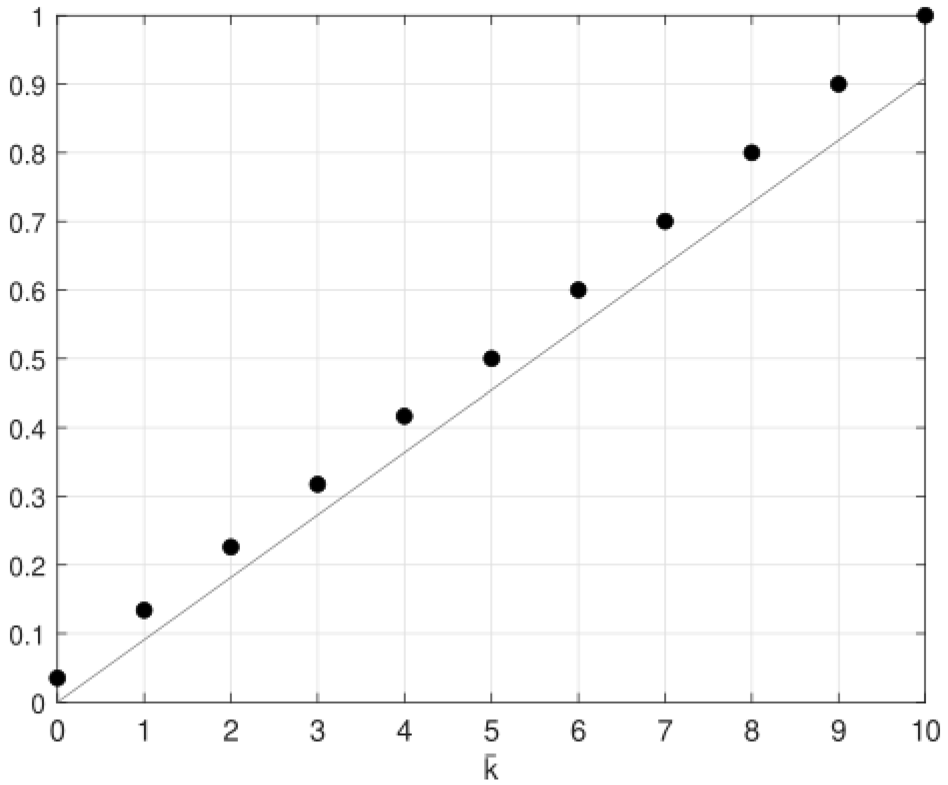


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5
6
7
8
9
10
k̄



$$\bar{k} \geq \frac{N}{2}$$

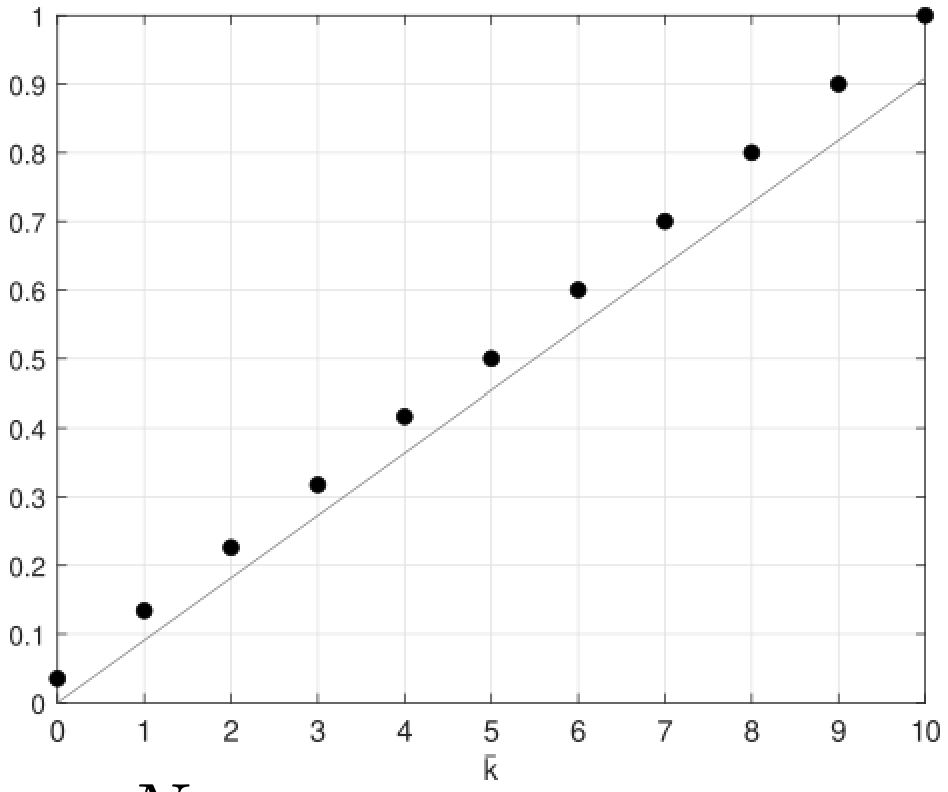
$$\mathrm{PErr}_N \le \frac{\bar{k}}{N}$$



$$\bar{k} \ge \frac{N}{2}$$

$$\mathrm{PErr}_N \leq \min_{\ell=0,1,\ldots,N} \frac{\binom{N}{\bar{k}}}{\binom{\ell}{\bar{k}}} \left(\frac{N-\ell}{N-\ell+1}\right)^{N-\ell} \frac{1}{N-\ell+1}$$

$$\mathrm{PErr}_N \leq \frac{\bar{k}}{N}$$

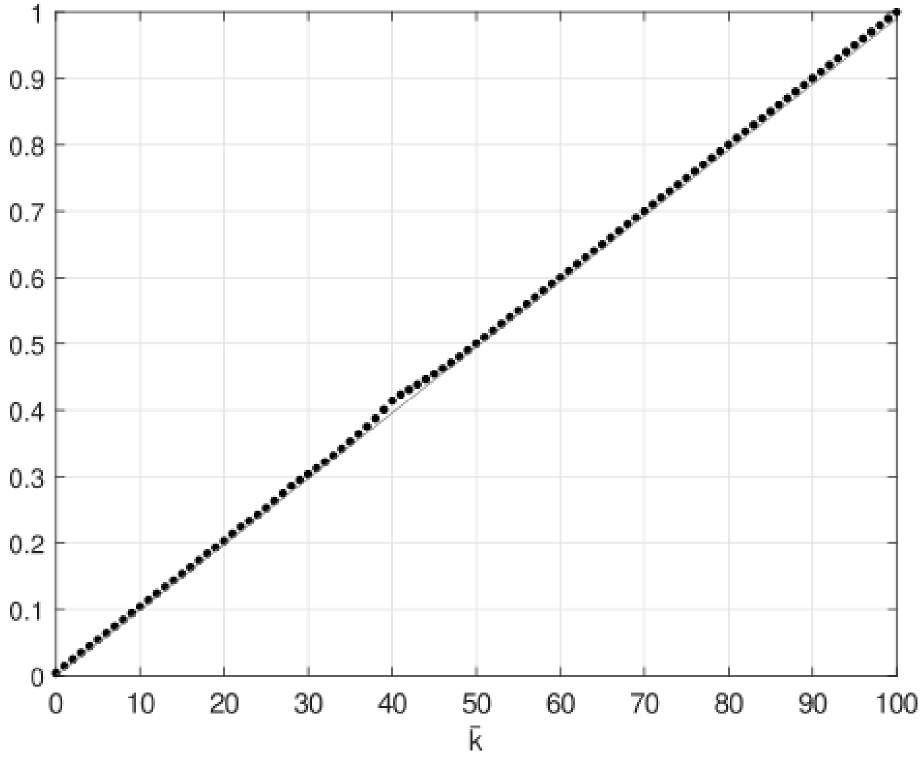


$$\bar{k} < \frac{N}{2}$$

$$\bar{k} \geq \frac{N}{2}$$

$$\mathrm{PErr}_N \leq \min_{\ell=0,1,\ldots,N} \frac{\binom{N}{\bar{k}}}{\binom{\ell}{\bar{k}}} \left(\frac{N-\ell}{N-\ell+1}\right)^{N-\ell} \frac{1}{N-\ell+1}$$

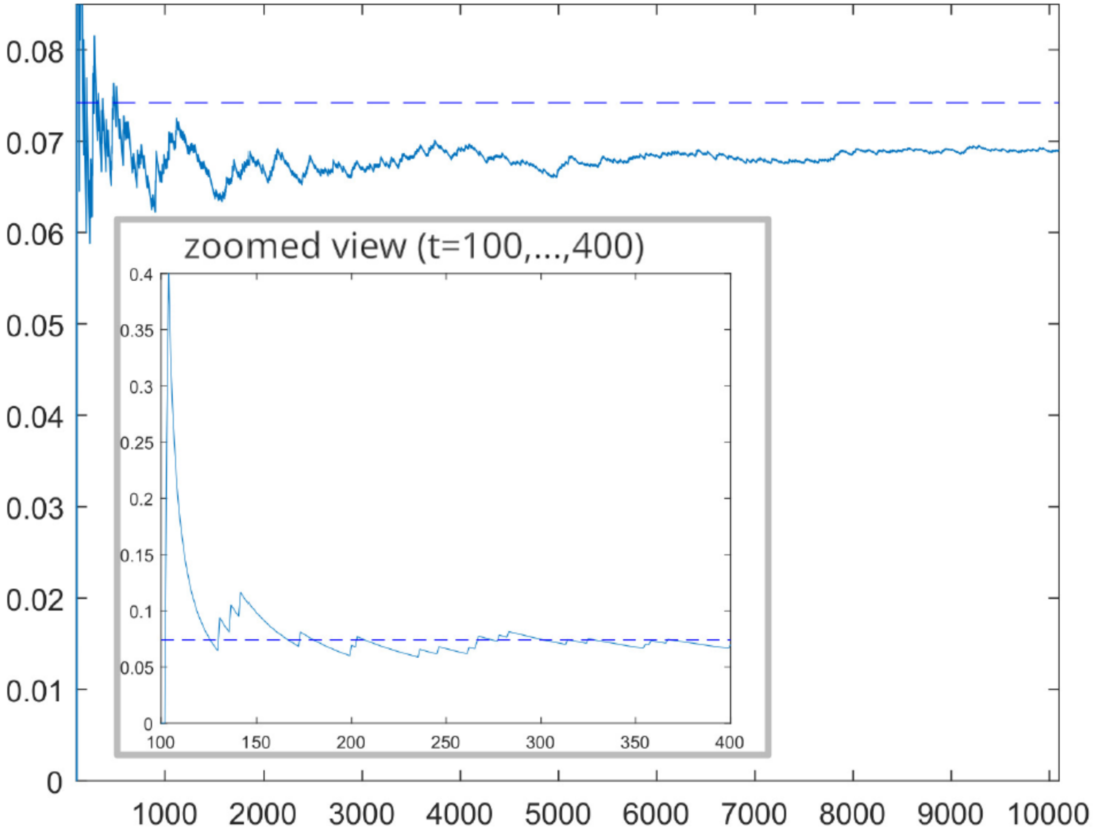
$$\mathrm{PErr}_N \leq \frac{\bar{k}}{N}$$



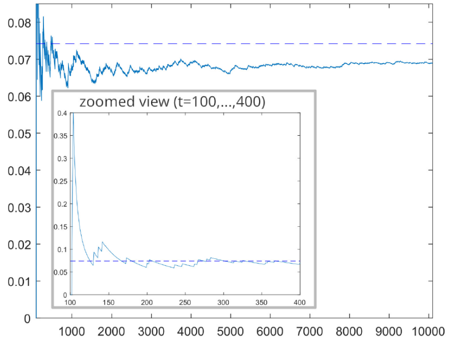
$$\bar{k} < \frac{N}{2}$$

$$\bar{k} \geq \frac{N}{2}$$

Application to sequential prediction



Application to sequential prediction

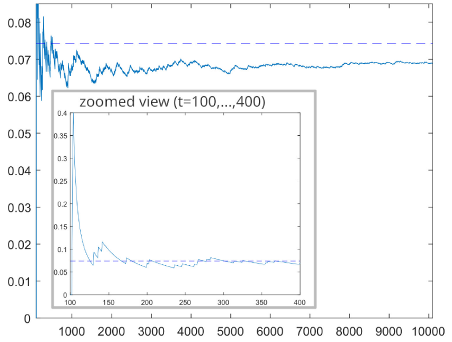


Take-home message:

To limit our mistakes,
we don't need to *postulate* that reality is simple.
By measuring the complexity (s^*) of our decisions,
we can **see** if our decisions align with reality
and *be cautious* if they don't.

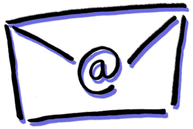
ALGO.CARE@UNIBS.IT

Application to sequential prediction



Take-home message:

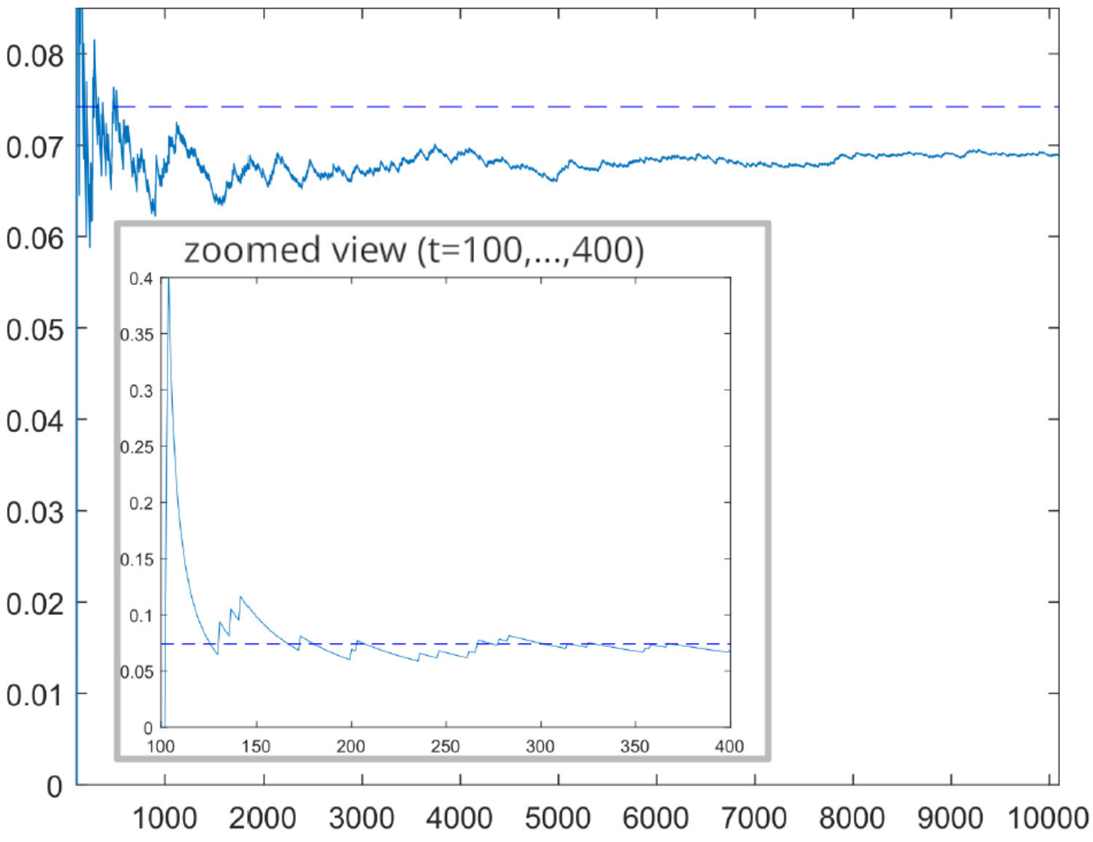
To limit our mistakes,
we don’t need to *postulate* that reality is simple.
By measuring the complexity (s^*) of our decisions,
we can **see** if our decisions align with reality
and *be cautious* if they don’t.



ALGO.CARE@UNIBS.IT

Thank you!

Application to sequential prediction



ALGO.CARE@UNIBS.IT