DATASET SIZE TUNING IN SCENARIO OPTIMIZATION

Speaker: **Simone Garatti**

(Politecnico di Milano, Italy – email: simone.garatti@polimi.it)

in collaboration with: **Marco C. Campi**

(University of Brescia, Italy – email: marco.campi@unibs.it)

Outline

- ❑ Scenario optimization: a brief resume
- ❑ Risk modulation and experiment design
- ❑ Support set and complexity
- ❑ Incremental scenario optimization
- ❑ Conclusions

Outline

- ❑ Scenario optimization: a brief resume
- ❑ Risk modulation and experiment design
- ❑ Support set and complexity
- ❑ Incremental scenario optimization
- ❑ Conclusions

Warning:

~~sequential decision making~~

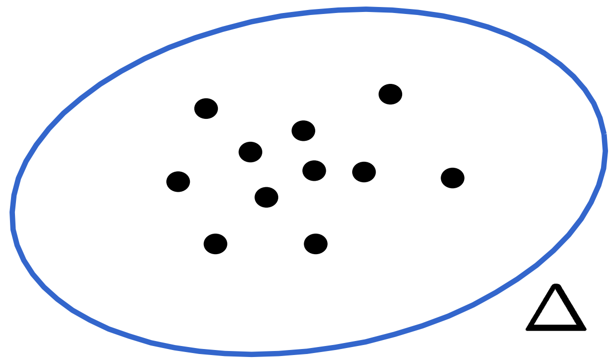
sequential acquisition of information

Scenario optimization: ingredients

Convex cost function: $f(x)$ ($x \in \mathbb{R}^d$ optimization variable)

Family of convex constraints: $\mathcal{X}_\delta$

δ stochastic parameter



i.i.d. sample of the stochastic parameter: $\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)}$

(experiments – data driven optimization)

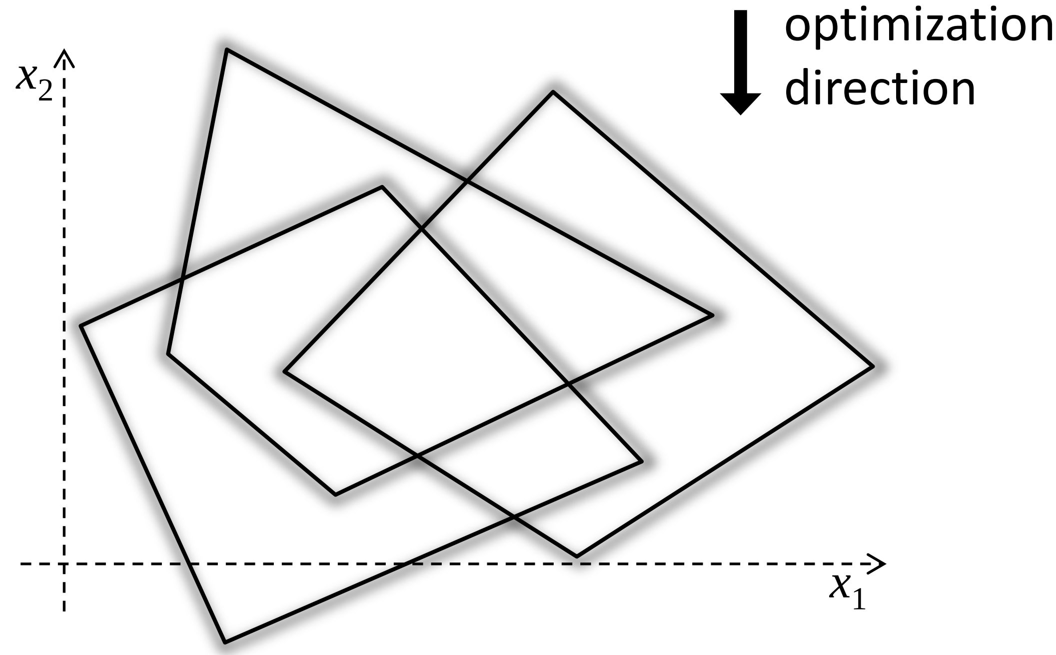
Scenario optimization

$$\delta^{(1)} \quad \rightarrow \quad \mathcal{X}_{\delta^{(1)}}$$

$$\delta^{(2)} \quad \rightarrow \quad \mathcal{X}_{\delta^{(2)}}$$

$$\vdots$$

$$\delta^{(N)} \quad \rightarrow \quad \mathcal{X}_{\delta^{(N)}}$$



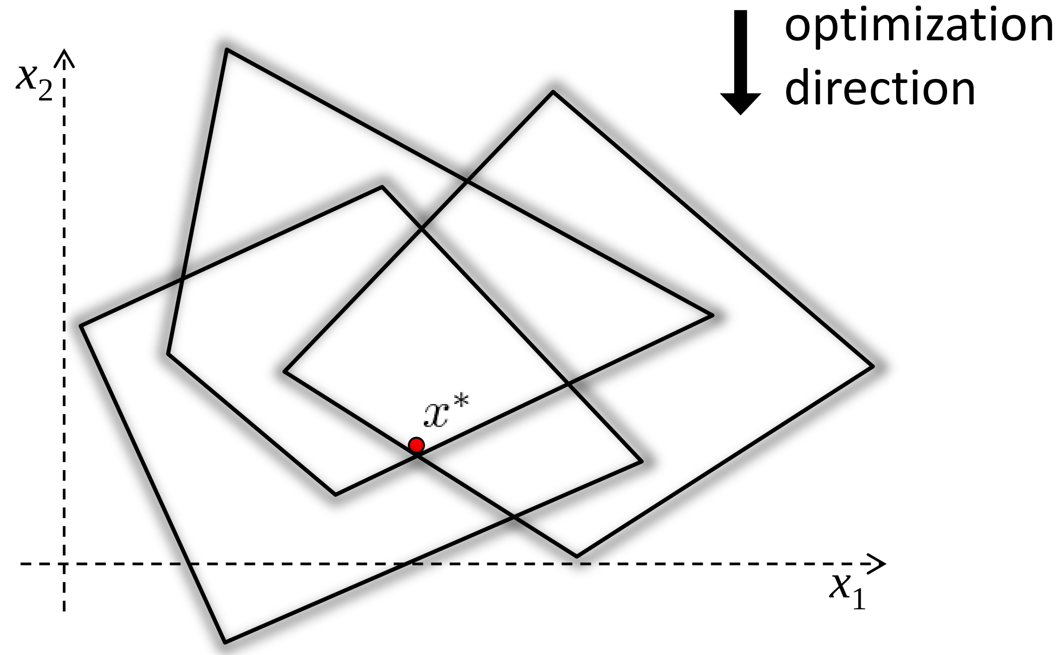
Scenario optimization

$$\delta^{(1)} \quad \rightarrow \quad \mathcal{X}_{\delta^{(1)}}$$

$$\delta^{(2)} \quad \rightarrow \quad \mathcal{X}_{\delta^{(2)}}$$

$$\vdots$$

$$\delta^{(N)} \quad \rightarrow \quad \mathcal{X}_{\delta^{(N)}}$$



scenario program

$$\min_x \; f(x)$$

$$\text{s.t.} \;\; x \in \bigcap_{i=1}^{N} \mathcal{X}_{\delta^{(i)}}$$

solution: x^*

Scenario optimization

scenario solution: x^*

main features:

- easy to compute
- data targeted to objective (direct approach)

issue: feasibility addressed empirically

Scenario optimization

scenario solution: x^*

main features:

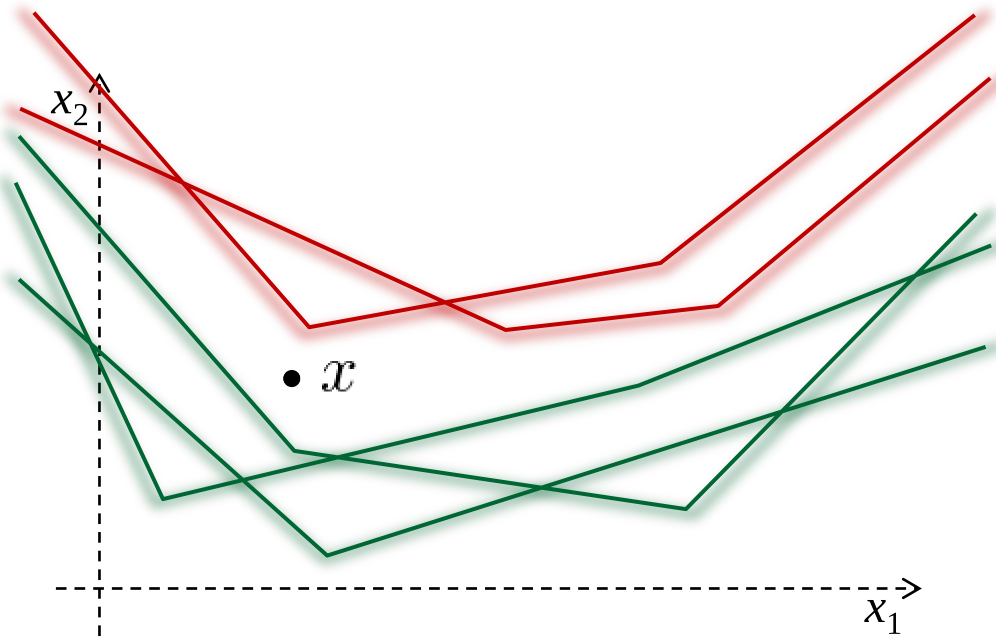
- easy to compute
- data targeted to objective (direct approach)

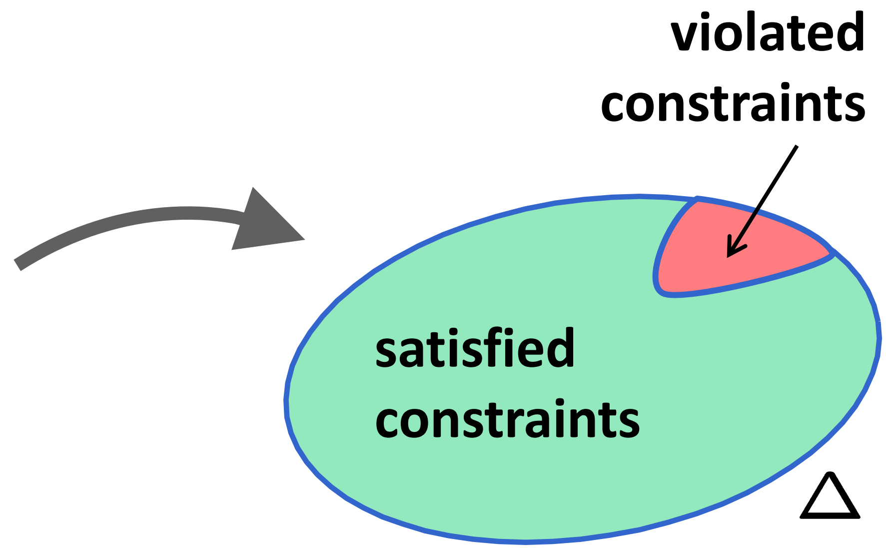
issue: feasibility addressed empirically

dependability of the scenario approach ⇨ keep control on risk

Violation (risk)

$$V(x) = \mathbb{P}\{\delta \in \Delta : \; x \notin \mathcal{X}_\delta\}$$

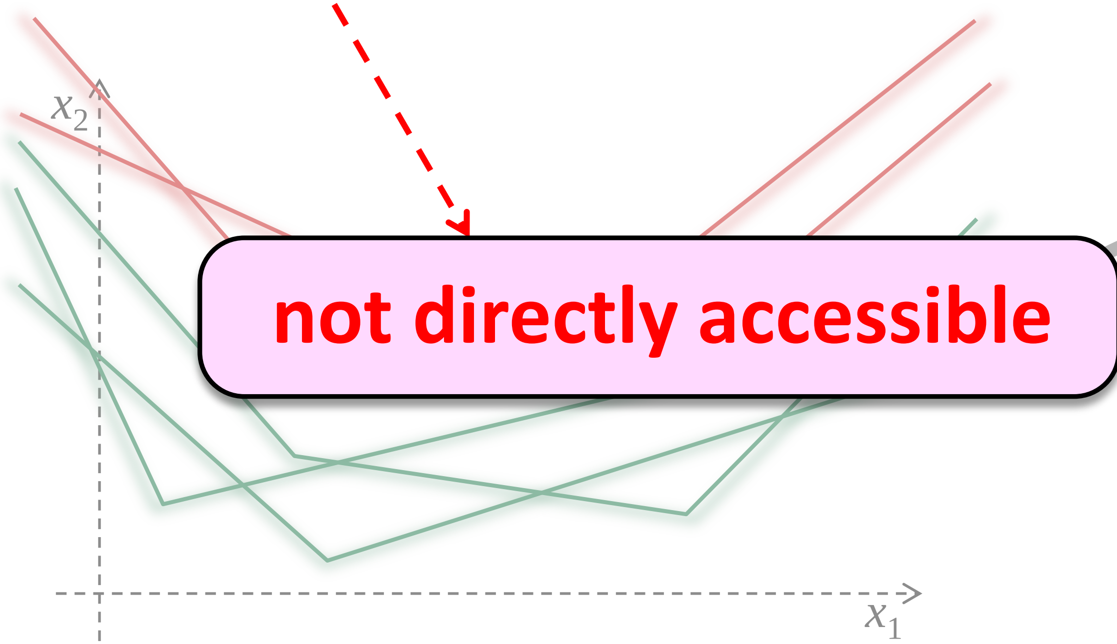


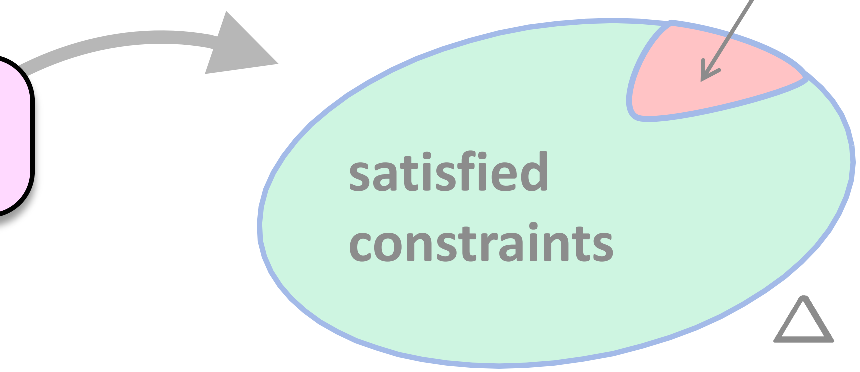


$V(x)$ = “size” of red region

Violation (risk)

$$V(x) = \mathbb{P}\{\delta \in \Delta : \; x \notin \mathcal{X}_\delta\}$$





$V(x)$ = "size" of red region

Experiment design in scenario optimization

Problem: choose N so that $V(x^*) \leq \epsilon$

↑
violation of the scenario solution

Theorem

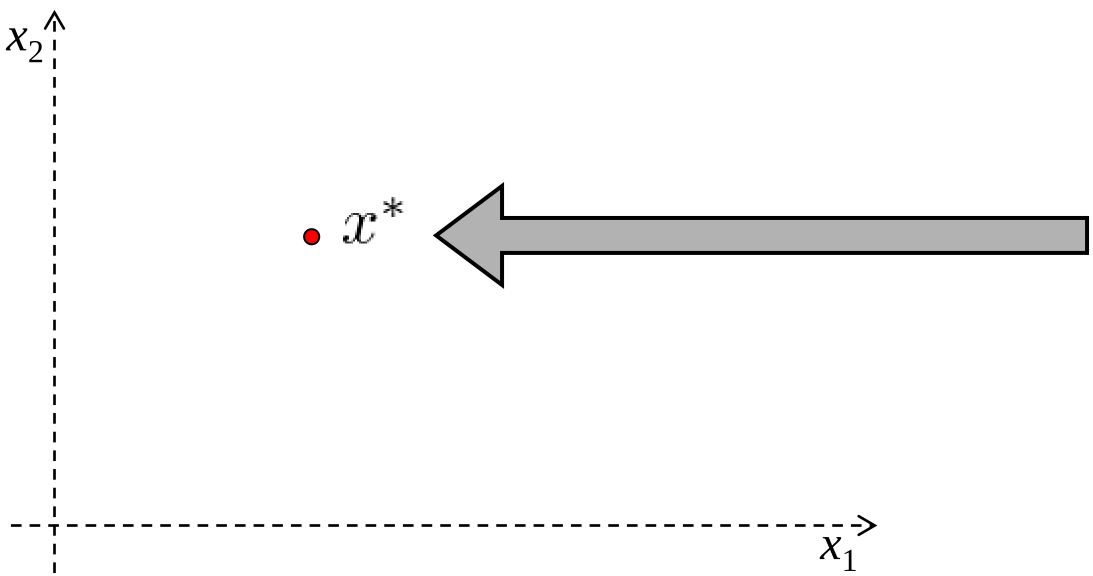
If N is big enough so that $\sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \leq \beta$

then $V(x^*) \leq \epsilon$ with confidence $1-\beta$

Violation of the scenario solution

Problem: choose N so that $V(x^*) \leq \epsilon$

↑ violation of the scenario solution

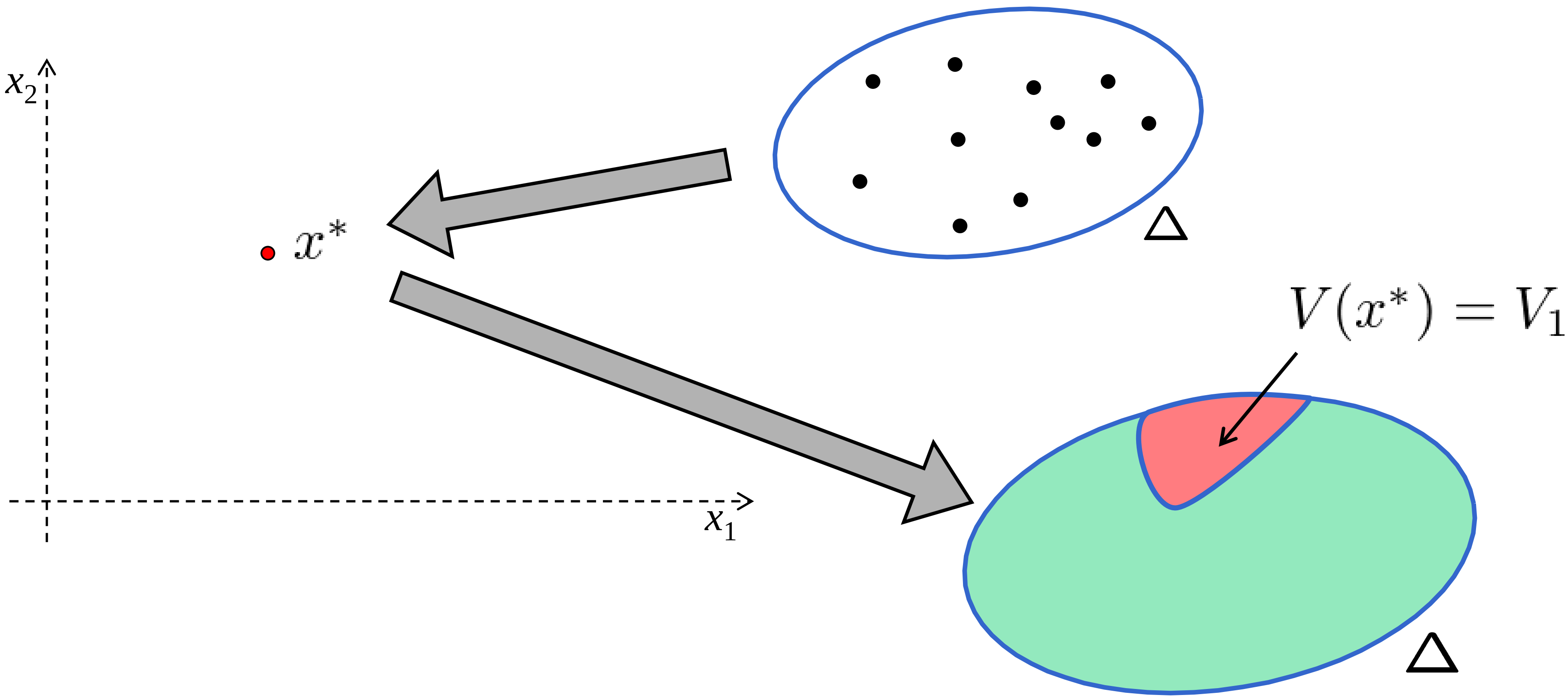


$$\min_{x} \quad f(x)$$
$$\text{s.t.} \quad x \in \bigcap_{i=1}^{N} \mathcal{X}_{\delta(i)}$$

Violation of the scenario solution

Problem: choose N so that $V(x^*) \leq \epsilon$

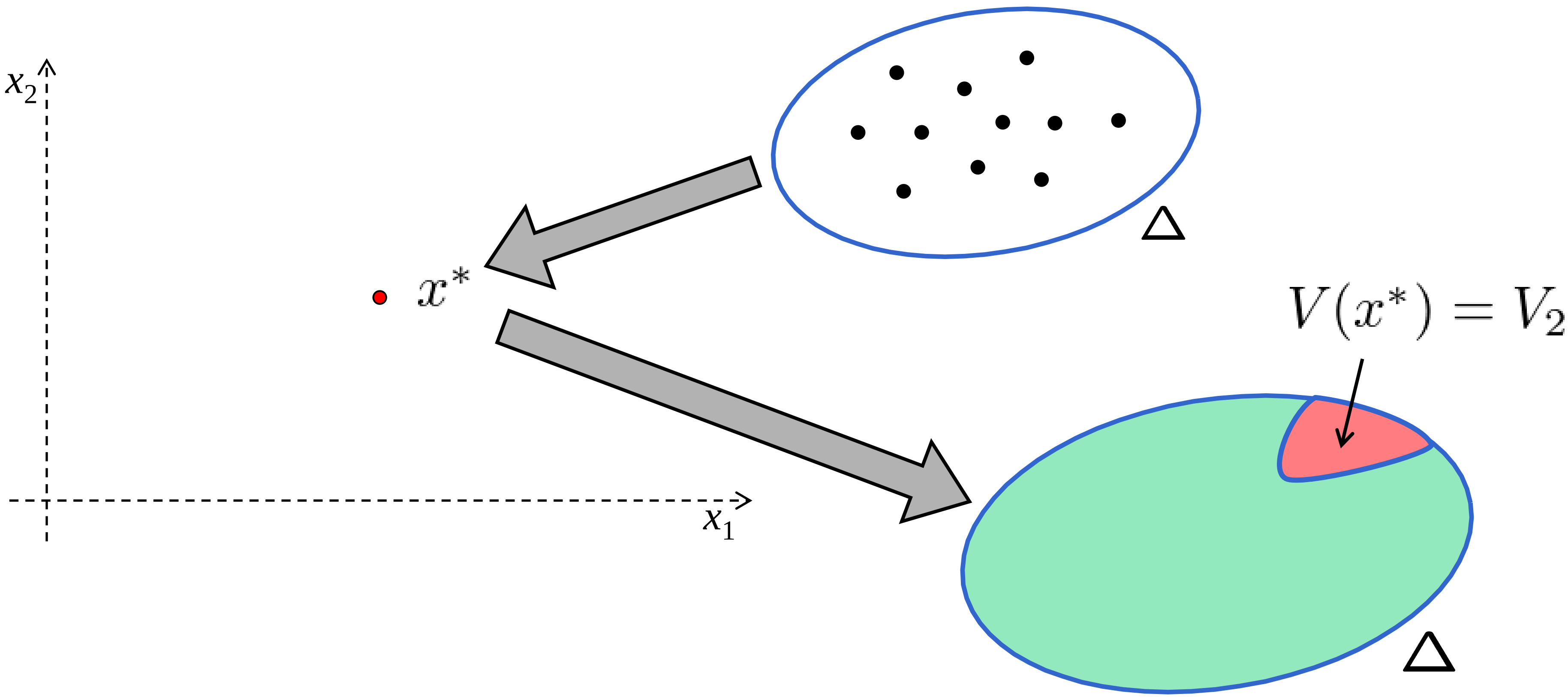
violation of the scenario solution



Violation of the scenario solution

Problem: choose N so that $V(x^*) \leq \epsilon$

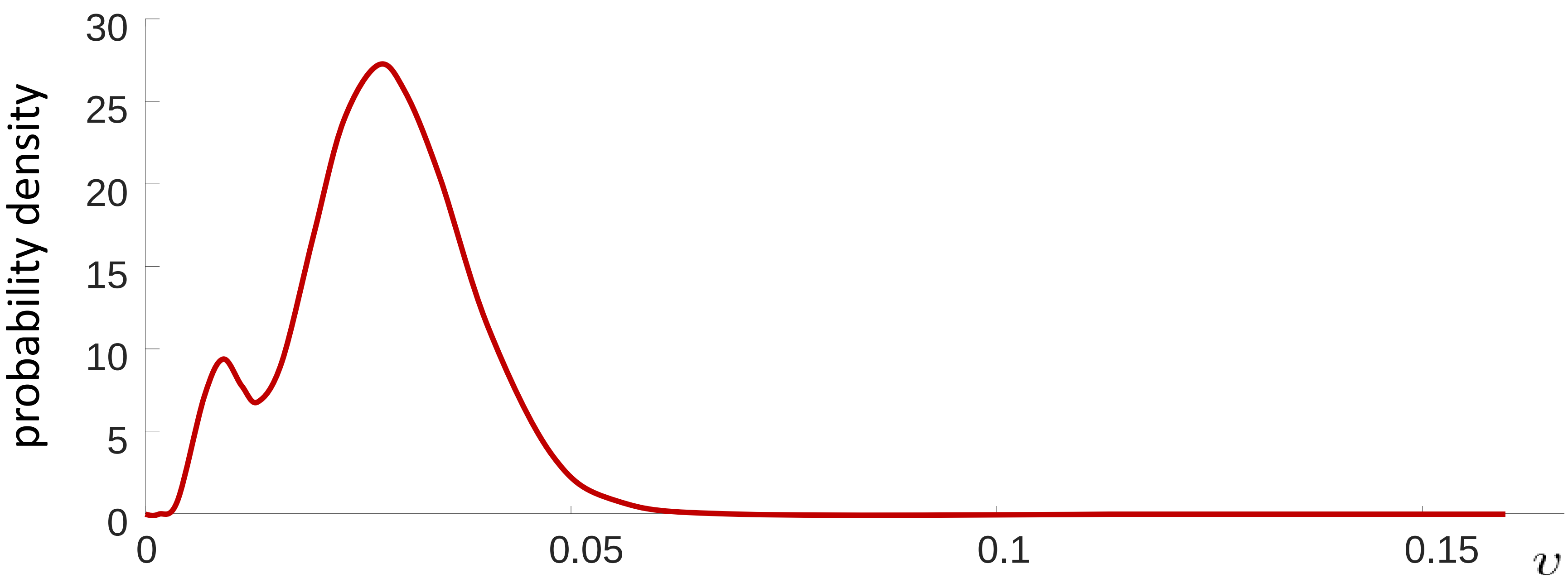
violation of the scenario solution



Violation of the scenario solution - concentration

$$V(x^*) = V(x^*(\delta^{(1)}, \delta^{(2)}, \dots, \delta^{(N)}))$$

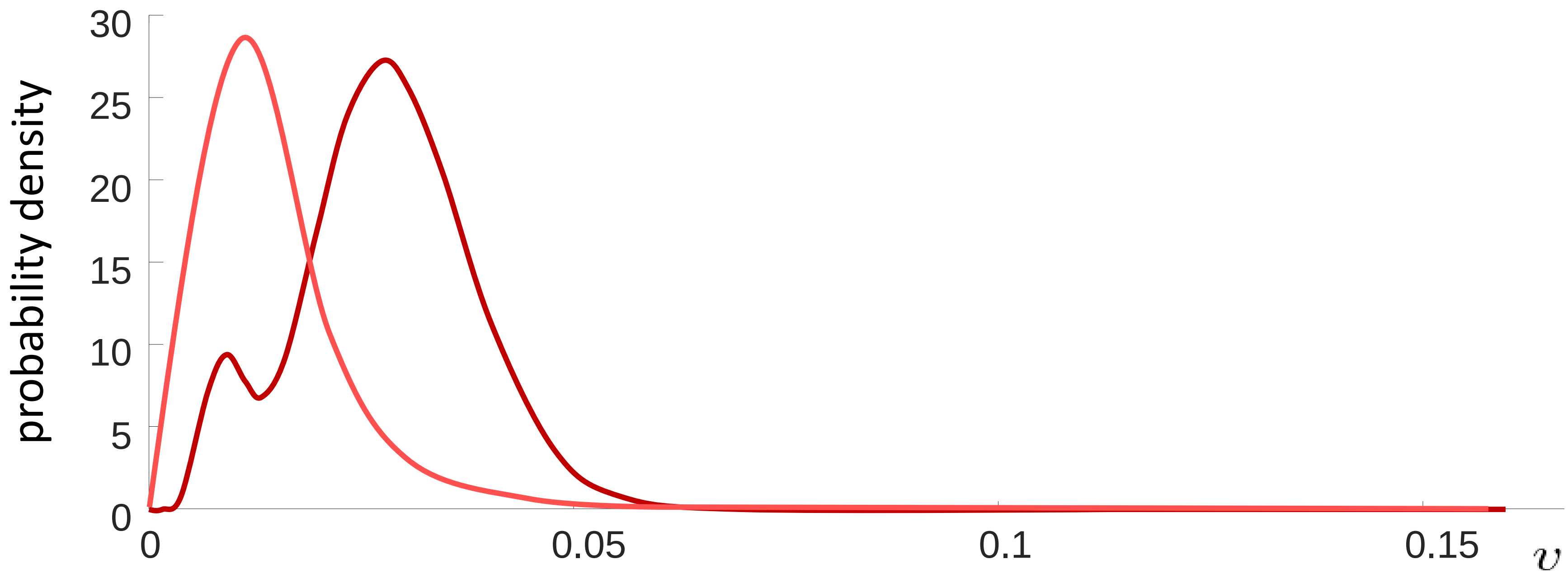
$V(x^*)$ is a random variable



Violation of the scenario solution - concentration

$$V(x^*) = V(x^*(\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)}))$$

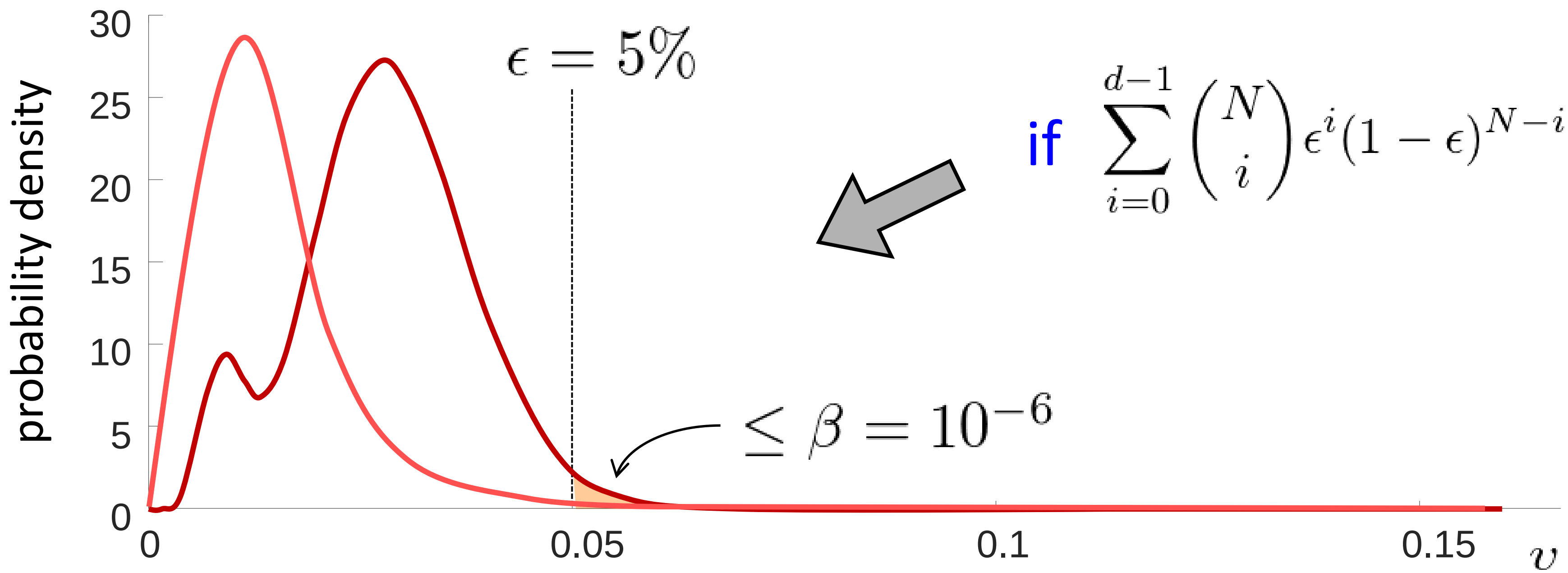
$V(x^*)$ is a random variable



Violation of the scenario solution - concentration

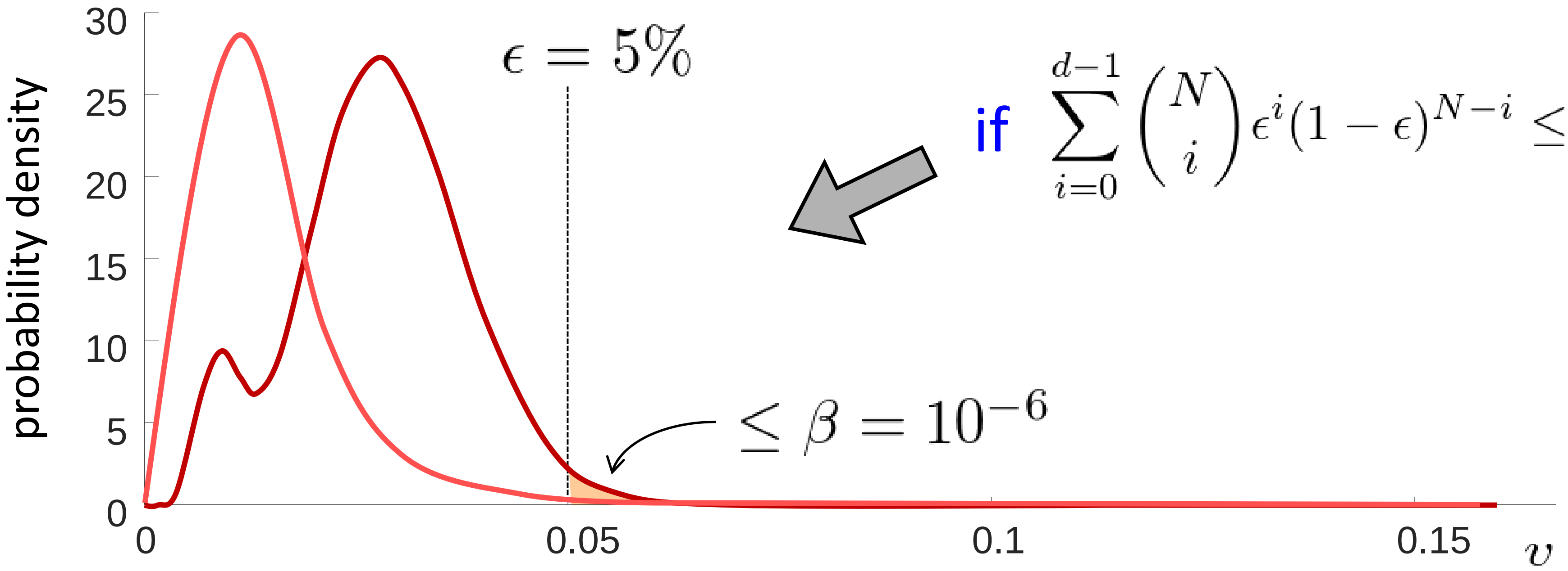
$$V(x^*) = V(x^*(\delta^{(1)}, \delta^{(2)}, \dots, \delta^{(N)}))$$

$V(x^*)$ is a random variable



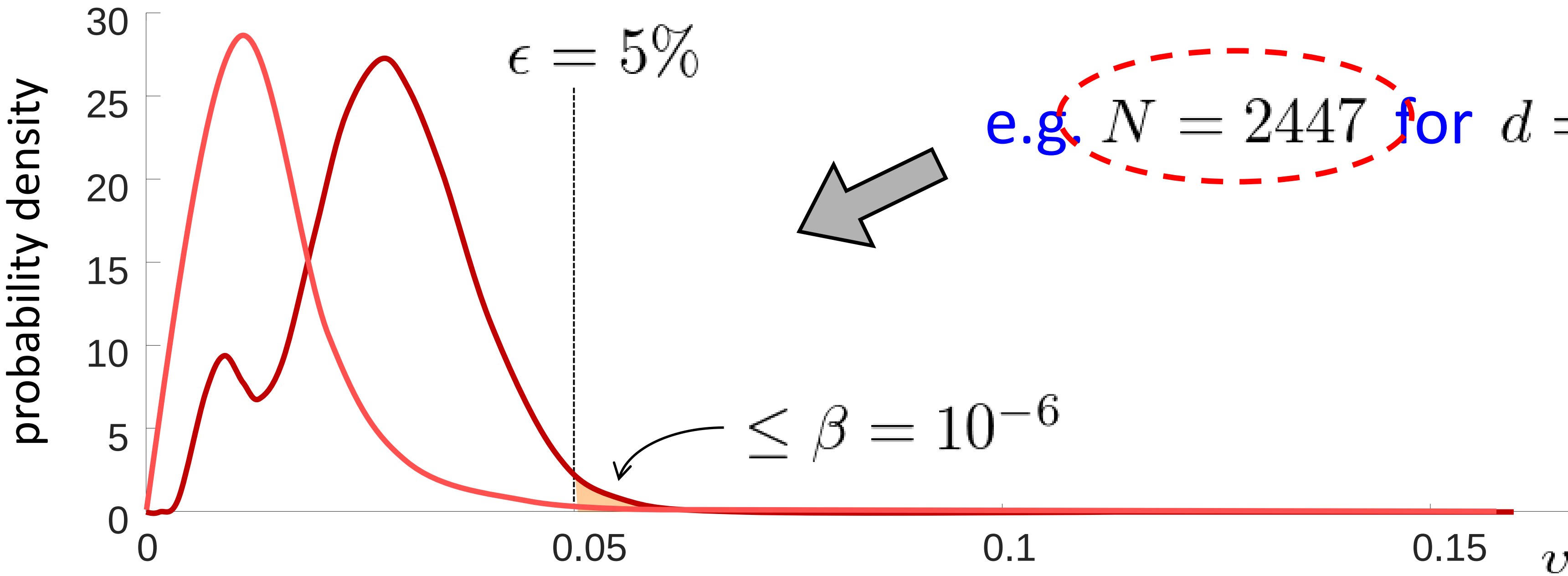
if $\displaystyle\sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \leq \beta$

The dataset size issue



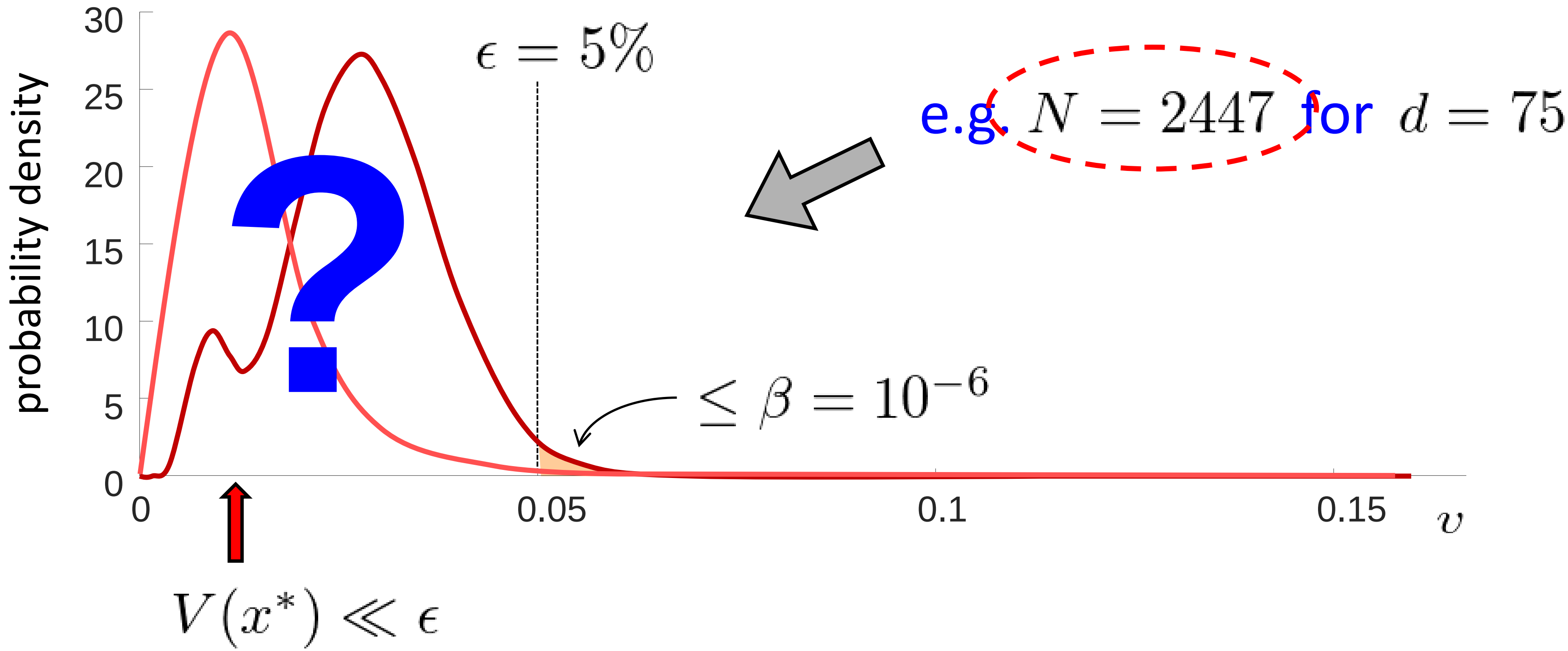
if $\sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \leq \beta$

The dataset size issue

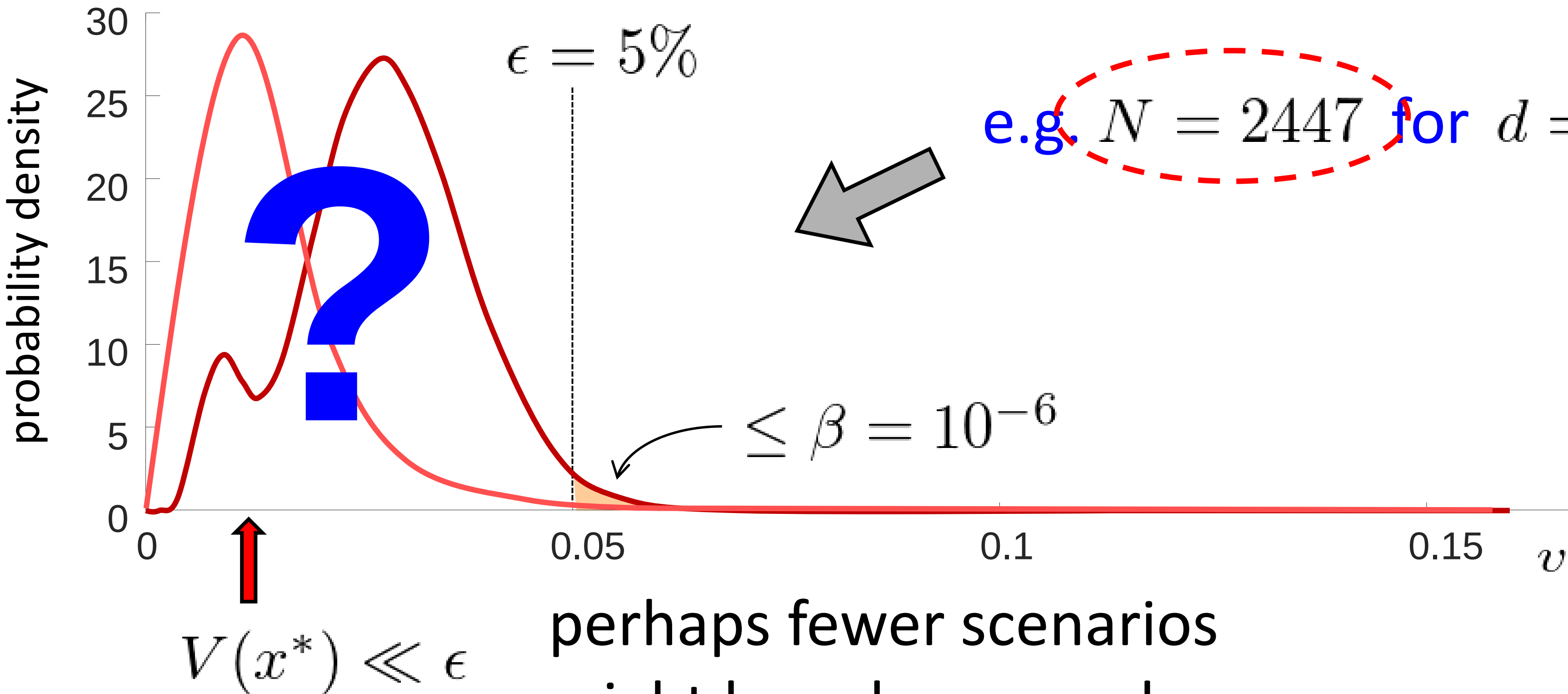


e.g. $N = 2447$ for $d = 75$

The dataset size issue



The dataset size issue



e.g. $N = 2447$ for $d = 75$

perhaps fewer scenarios might have been used...

waste of scenarios (data)

The dataset size issue (cont'd)

Collecting scenarios can be:

1. expensive

NASA experiment: 12000$ for each single scenario!

The dataset size issue (cont’d)

Collecting scenarios can be:

1. expensive

    NASA experiment: 12000$ for each single scenario!

2. time-consuming

    3 scenarios per day…

    2447 scenarios = more than 2 years!

Goal

to devise a guaranteed (e.g. risk < 5%) scheme where the sample size is learned on the way so as to avoid any waste of scenarios

difficulties:

violation is

- problem dependent
- dataset dependent
- not directly accessible

Goal

to devise a guaranteed (e.g. risk $< 5\%$) scheme where the sample size is learned on the way so as to avoid any waste of scenarios

difficulties:

violation is

- problem dependent
- dataset dependent
- not directly accessible

→ tool to evaluate violation without using additional info

Support set and its cardinality

Support set: $\left\{\delta^{(i_1)}, \delta^{(i_2)}, \ldots, \delta^{(i_k)}\right\}$ such that

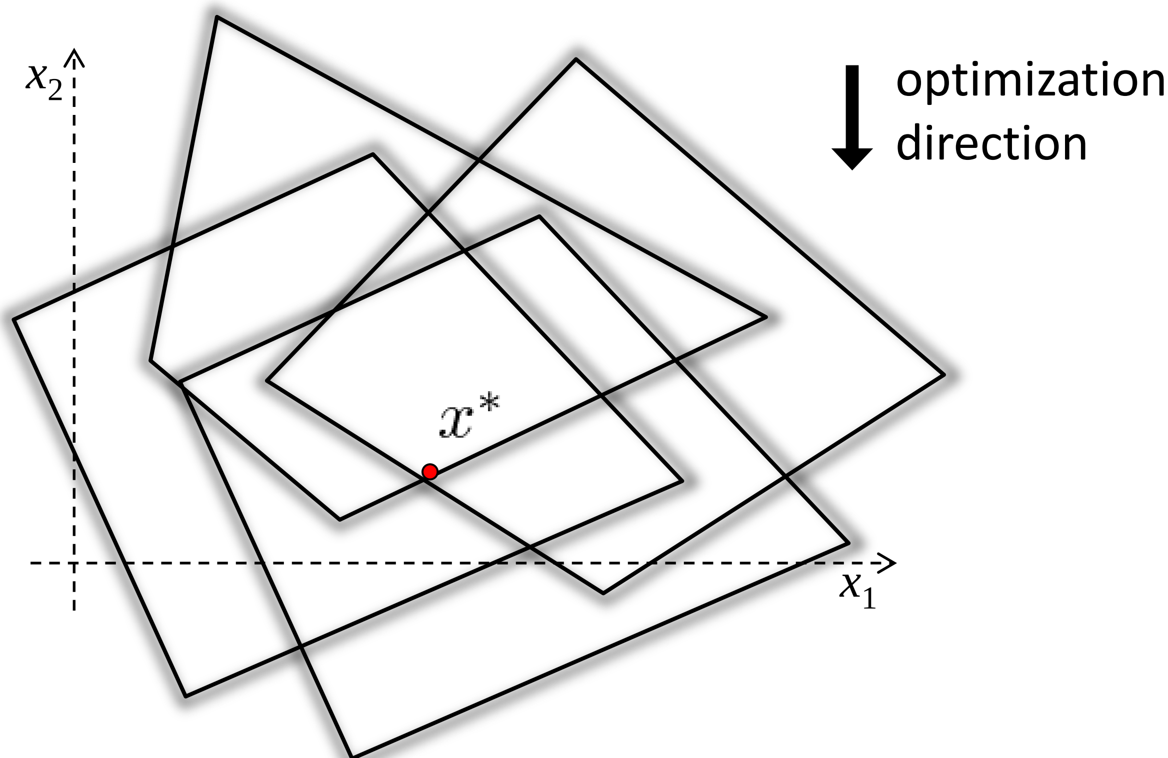
1. $x^*\left(\delta^{(i_1)}, \delta^{(i_2)}, \ldots, \delta^{(i_k)}\right) = x^*\left(\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)}\right)$
2. to keep x^* unchanged, no $\delta^{(i_j)}$ can be further removed

Support set and its cardinality

Support set: $\left\{\delta^{(i_1)}, \delta^{(i_2)}, \ldots, \delta^{(i_k)}\right\}$ such that

1. $x^*\left(\delta^{(i_1)}, \delta^{(i_2)}, \ldots, \delta^{(i_k)}\right) = x^*\left(\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)}\right)$

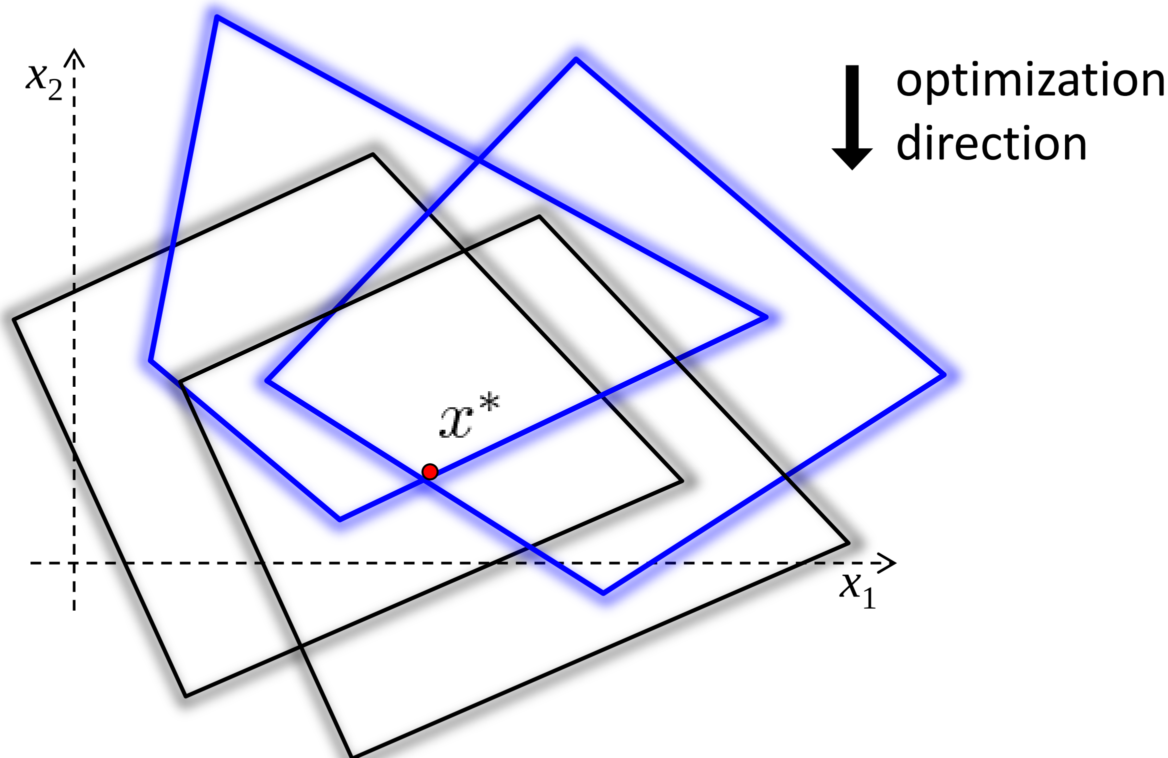
2. to keep x^* unchanged, no $\delta^{(i_j)}$ can be further removed



Support set and its cardinality

Support set: $\left\{\delta^{(i_1)}, \delta^{(i_2)}, \ldots, \delta^{(i_k)}\right\}$ such that

1. $x^*\left(\delta^{(i_1)}, \delta^{(i_2)}, \ldots, \delta^{(i_k)}\right) = x^*\left(\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)}\right)$
2. to keep x^* unchanged, no $\delta^{(i_j)}$ can be further removed

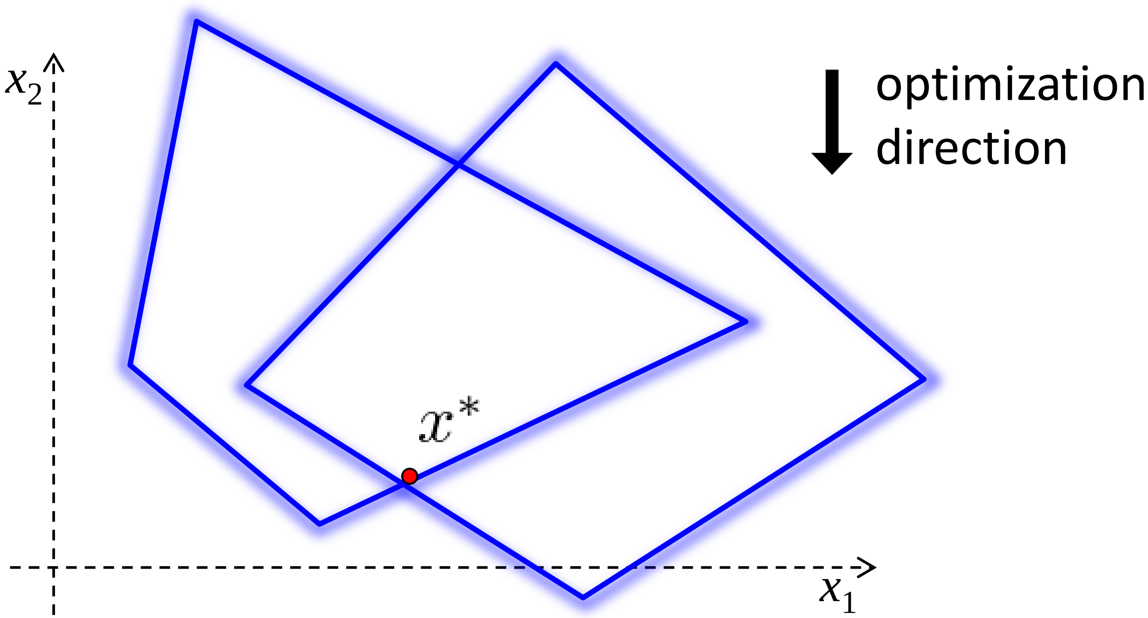


Support set and its cardinality

Support set: $\left\{\delta^{(i_1)}, \delta^{(i_2)}, \ldots, \delta^{(i_k)}\right\}$ such that

1. $x^*\left(\delta^{(i_1)}, \delta^{(i_2)}, \ldots, \delta^{(i_k)}\right) = x^*\left(\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)}\right)$

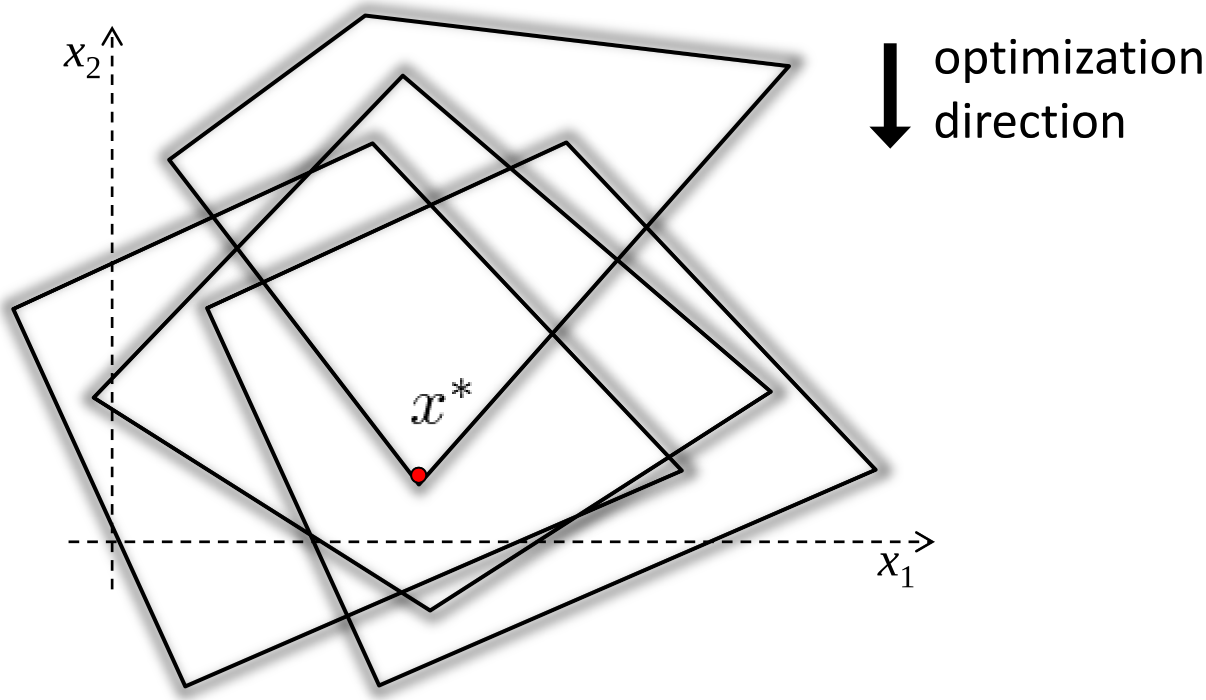
2. to keep x^* unchanged, no $\delta^{(i_j)}$ can be further removed



Support set and its cardinality

Support set: $\left\{\delta^{(i_1)}, \delta^{(i_2)}, \ldots, \delta^{(i_k)}\right\}$ such that

1. $x^*\left(\delta^{(i_1)}, \delta^{(i_2)}, \ldots, \delta^{(i_k)}\right) = x^*\left(\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)}\right)$
2. to keep x^* unchanged, no $\delta^{(i_j)}$ can be further removed

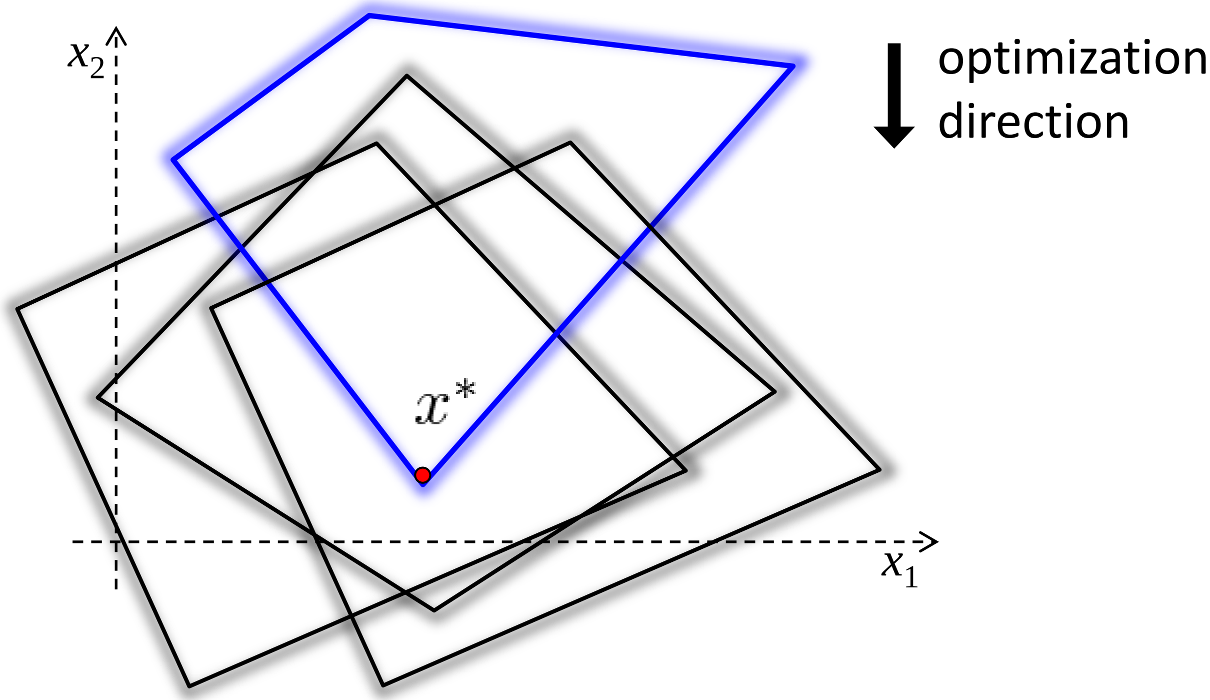


Support set and its cardinality

Support set: $\left\{\delta^{(i_1)}, \delta^{(i_2)}, \ldots, \delta^{(i_k)}\right\}$ such that

1. $x^*\left(\delta^{(i_1)}, \delta^{(i_2)}, \ldots, \delta^{(i_k)}\right) = x^*\left(\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)}\right)$

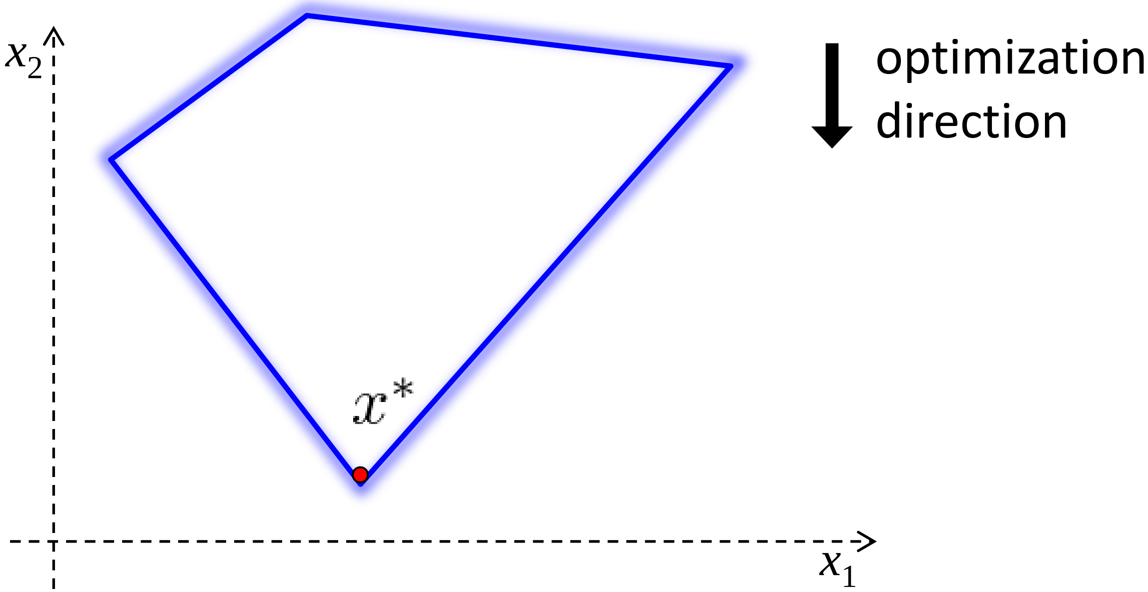
2. to keep x^* unchanged, no $\delta^{(i_j)}$ can be further removed



Support set and its cardinality

Support set: $\left\{\delta^{(i_1)}, \delta^{(i_2)}, \ldots, \delta^{(i_k)}\right\}$ such that

1. $x^*\left(\delta^{(i_1)}, \delta^{(i_2)}, \ldots, \delta^{(i_k)}\right) = x^*\left(\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)}\right)$
2. to keep x^* unchanged, no $\delta^{(i_j)}$ can be further removed

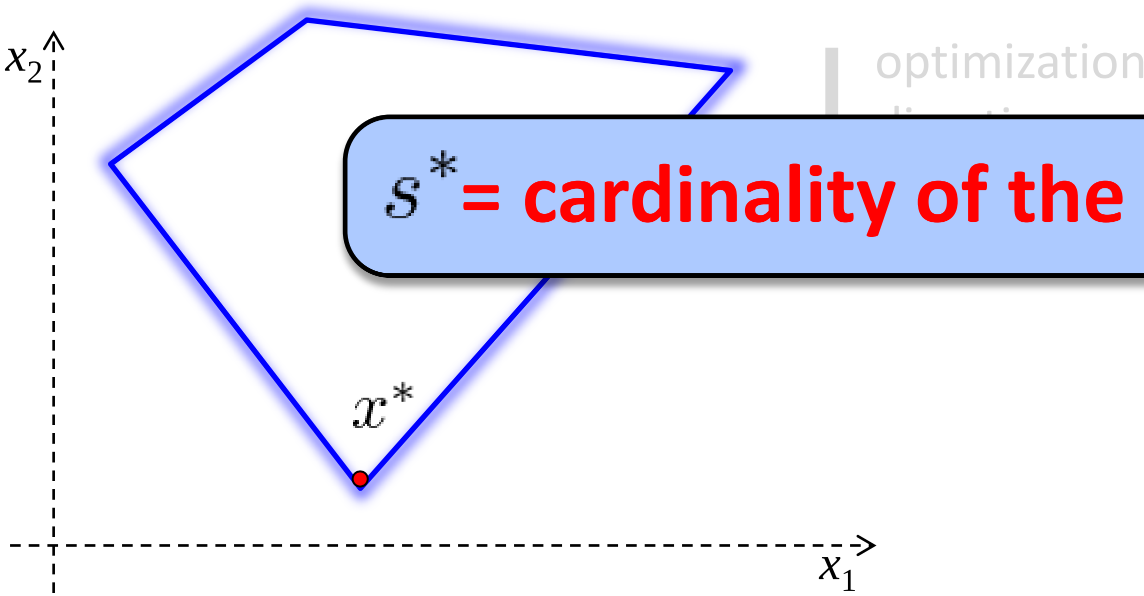


Support set and its cardinality

Support set: $\left\{\delta^{(i_1)}, \delta^{(i_2)}, \ldots, \delta^{(i_k)}\right\}$ such that

1. $x^*\left(\delta^{(i_1)}, \delta^{(i_2)}, \ldots, \delta^{(i_k)}\right) = x^*\left(\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)}\right)$

2. to keep x^* unchanged, no $\delta^{(i_j)}$ can be further removed



s^* = **cardinality of the support set**

Scenario Optimization: wait & judge

$V(x^*)$ is a real random variable

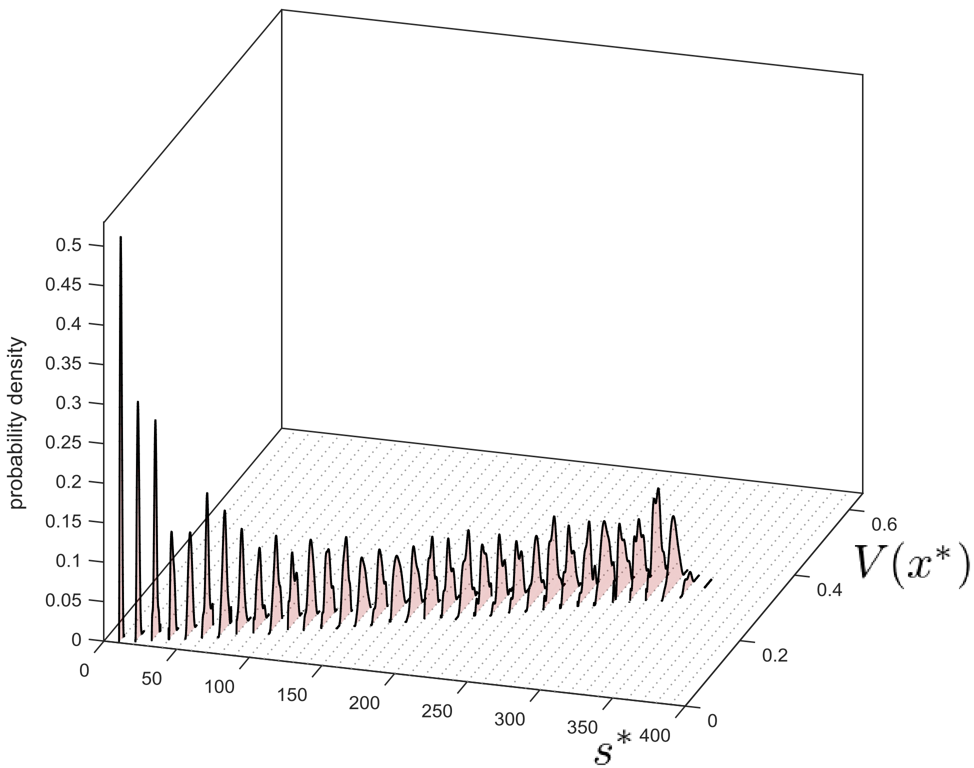
s^* is an integer random variable

} bivariate perspective

Scenario Optimization: wait & judge

$V(x^*)$ is a real random variable
s^* is an integer random variable

} bivariate perspective

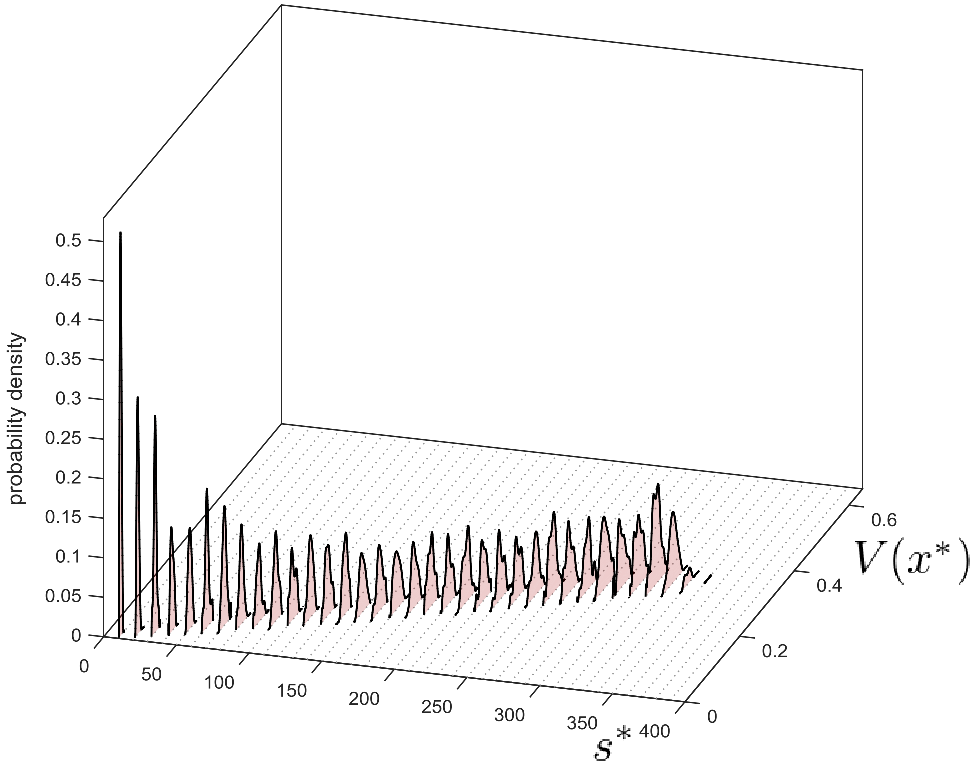


Scenario Optimization: wait & judge

$V(x^*)$ is a real random variable

s^* is an integer random variable

} bivariate perspective

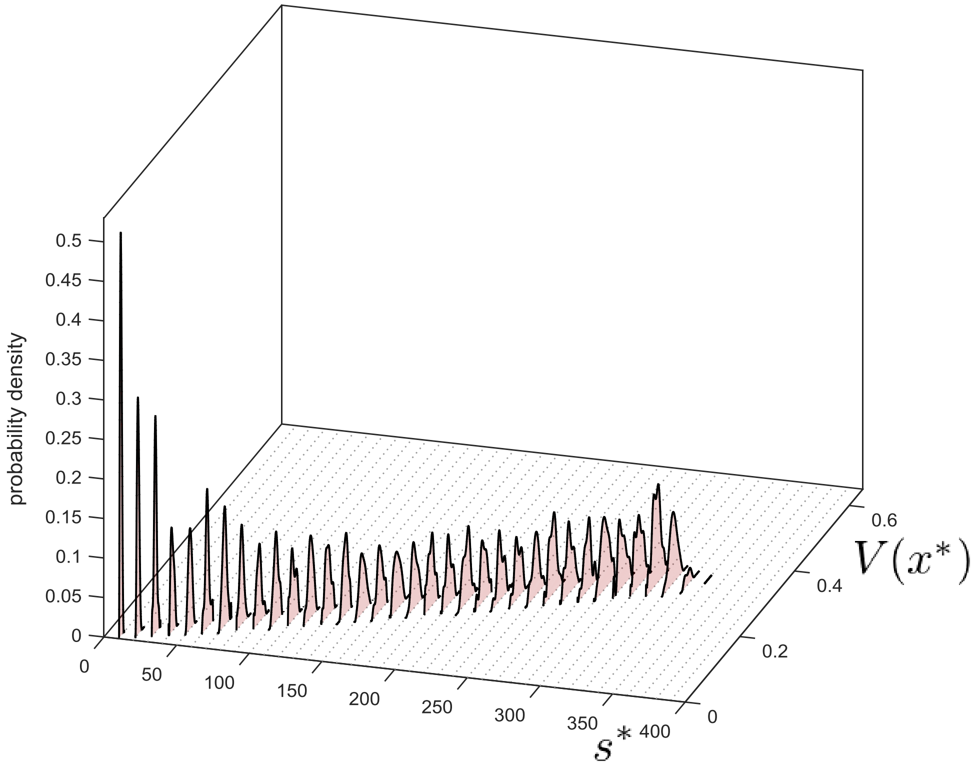


bivariate probability distribution is always concentrated!

Scenario Optimization: wait & judge

$V(x^*)$ is a real random variable
s^* is an integer random variable

} bivariate perspective



bivariate probability distribution is always concentrated!

→ **high linear correlation**

$V(x^*)$ can be accurately estimated from s^* ----→ **observed once the program is solved !!**

Incremental Scenario Optimization

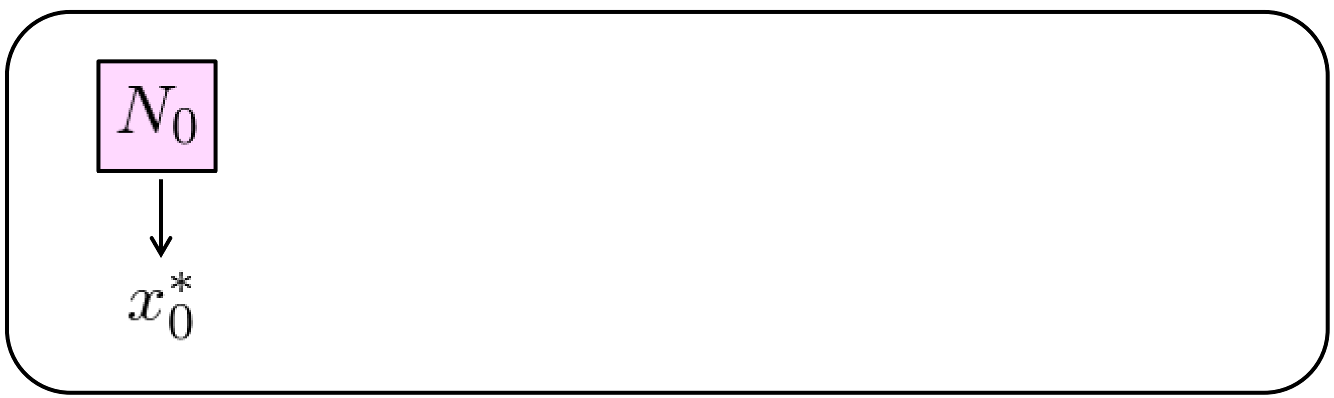
1. collect $\delta^{(1)}, \delta^{(2)}, \dots, \delta^{(N_0)}$

2. compute $$x_0^* = \arg\min_x \; f(x) \quad \text{s.t.} \quad x \in \mathcal{X}_{\delta^{(i)}}, \; i = 1, \dots, N_0$$

3. compute s_0^*

4. if $s_0^* = 0$ return $x^* = x_0^*$

5. else ...



Incremental Scenario Optimization

1. add scenarios $\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N_0)}, \delta^{(N_0+1)}, \ldots, \delta^{(N_1)}$

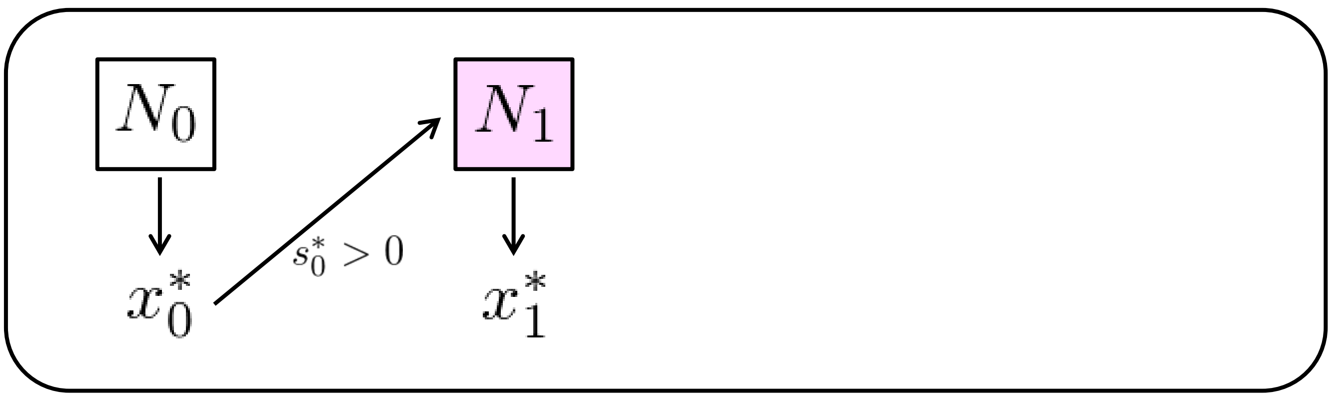
2. compute $x_1^* = \arg\min_x \quad f(x)$

$$\text{s.t.} \quad x \in \mathcal{X}_{\delta^{(i)}}, \ i = 1, \ldots, N_1$$

3. compute s_1^*

4. if $s_1^* \leq 1$ return $x^* = x_1^*$

5. else ...



Incremental Scenario Optimization

1. add scenarios $\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N_0)}, \delta^{(N_0+1)}, \ldots, \delta^{(N_1)}$

2. compute $x_1^* = \arg\min_x \; f(x)$

$\text{s.t.}\ x \in \mathcal{X} \quad i = 1 \quad , N_1$

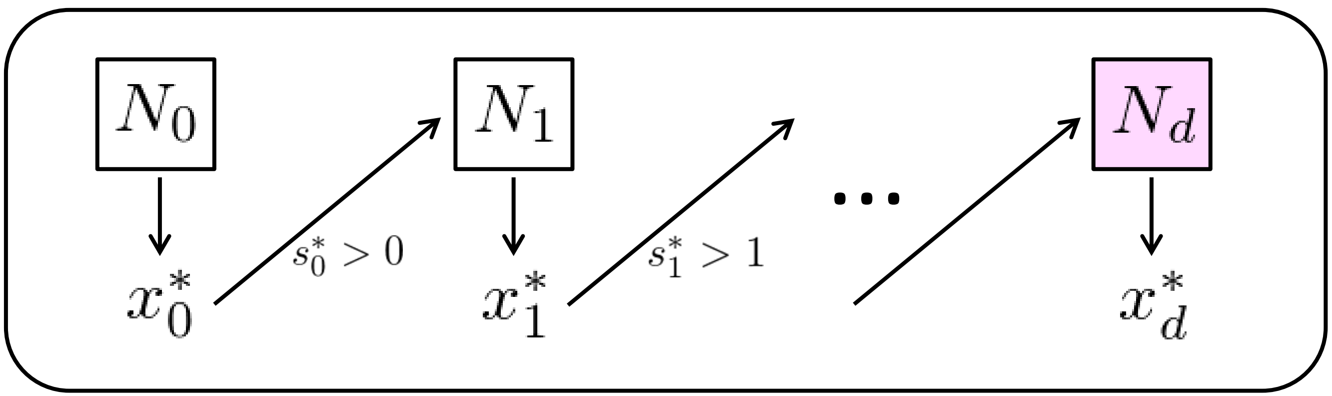
3. compute

d steps at most

(Helly's theorem ➡ $s^* \leq d$)

4. if $s_1^* \leq 1$ return $x^* = x_1^*$

5. else ...



Sizing of N_0, N_1, …, N_d

Find $N_0, N_1, \ldots, N_d$ such that

1. as small as possible
2. $V(x^*) \leq \epsilon$ with confidence $1 - \beta$

Sizing of $N_0, N_1, \ldots, N_d$

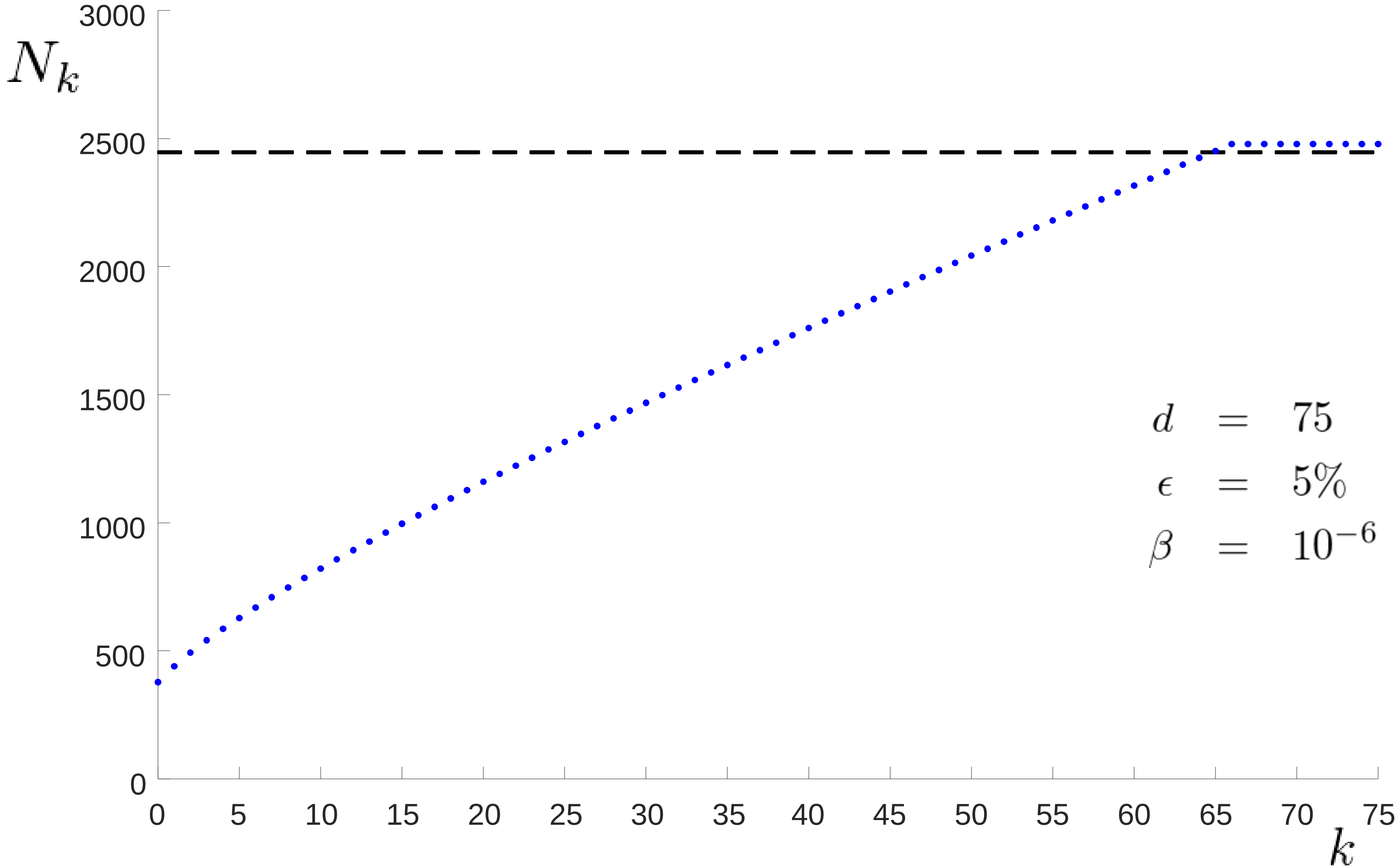
Find $N_0, N_1, \ldots, N_d$ such that

1. as small as possible
2. $V(x^*) \leq \epsilon$ with confidence $1 - \beta$

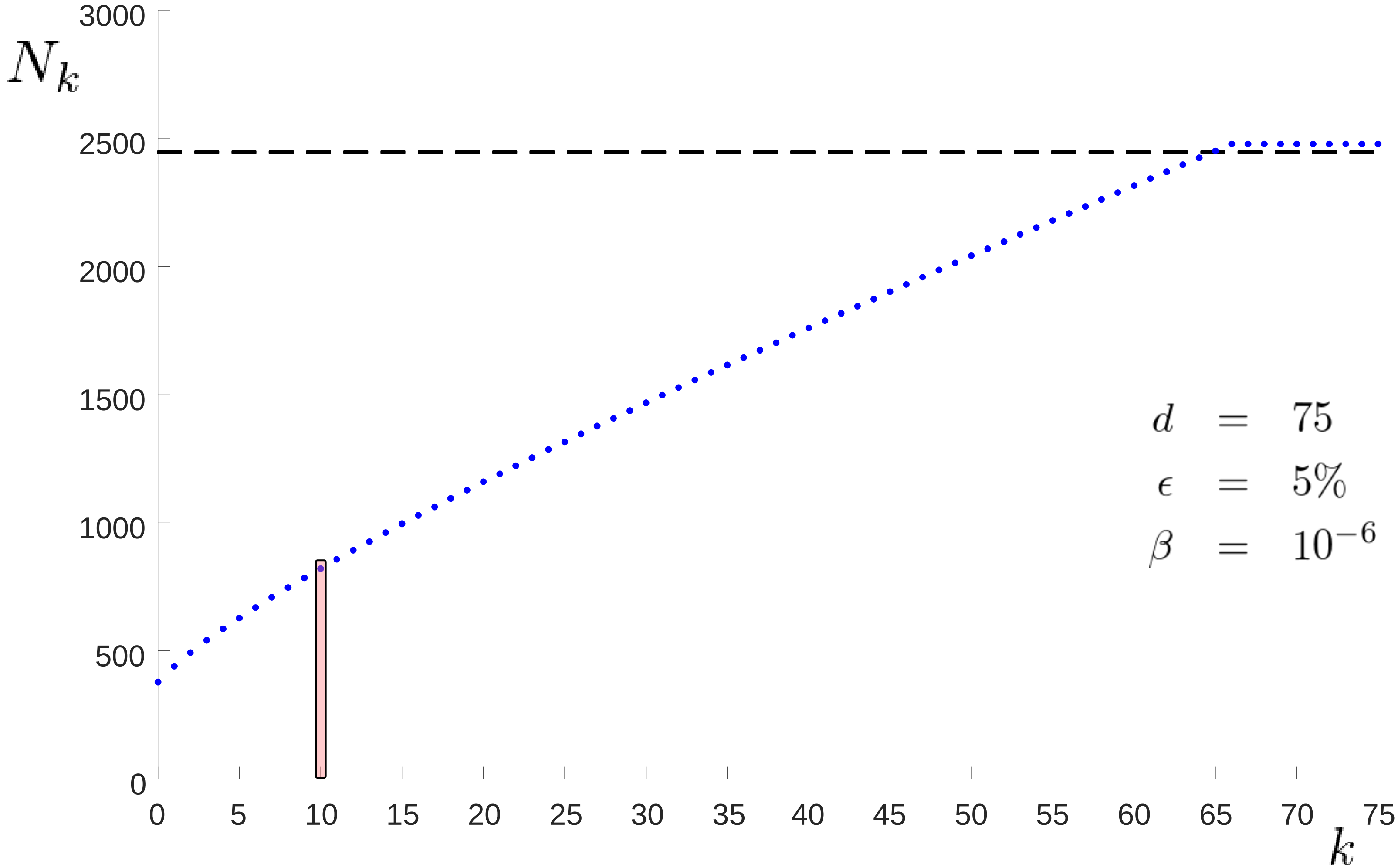
Theorem

$$N_j = \min \left\{ N : \quad N \geq \bar{M}_j \quad \text{and} \quad \frac{\beta}{(d+1)(\bar{M}_j + 1)} \sum_{m=j}^{\bar{M}_j} \binom{m}{j} (1-\epsilon)^{m-j} \geq \binom{N}{j} (1-\epsilon)^{N-j} \right\}$$

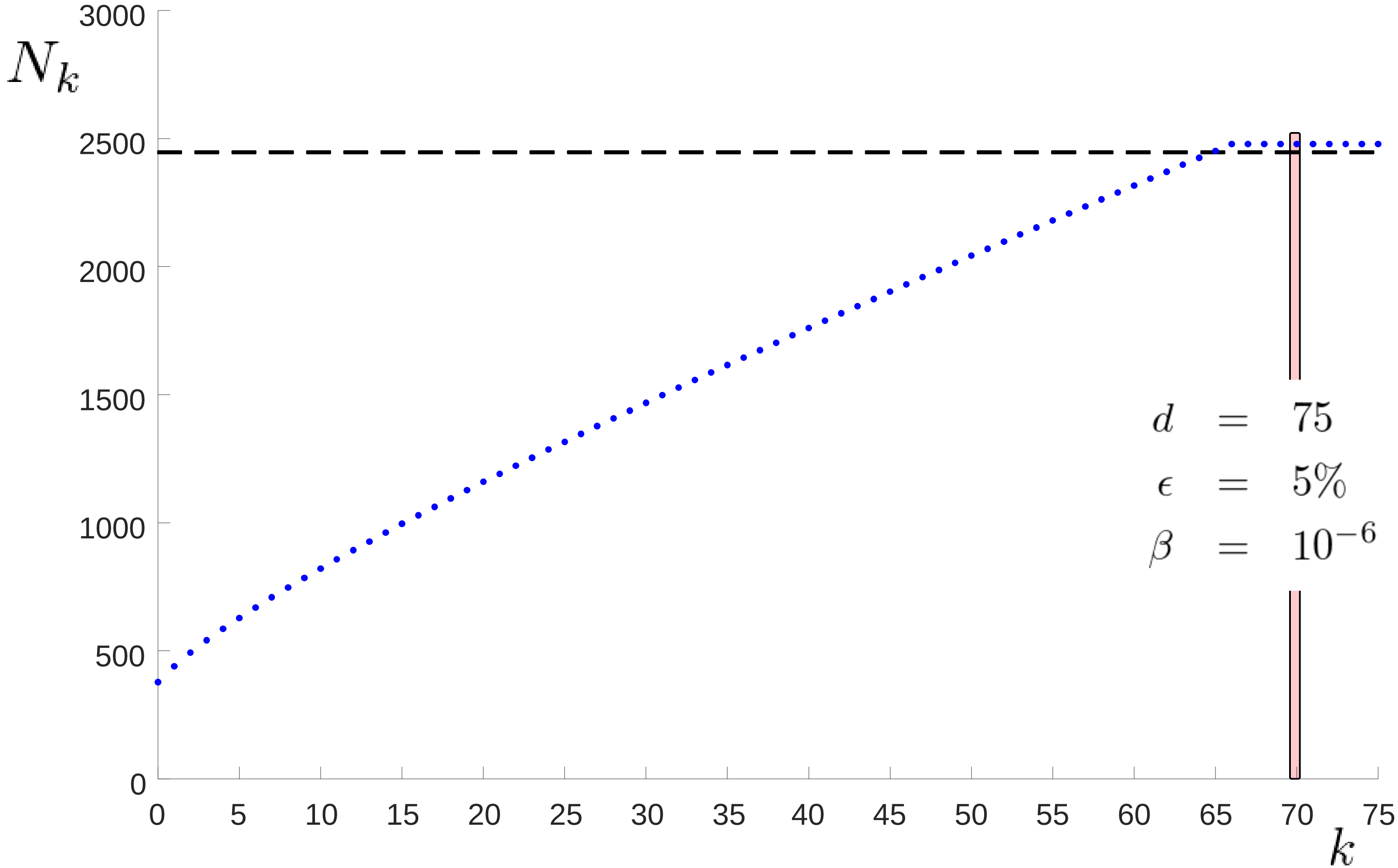
Sizing of N_0, N_1, …, N_d (cont'd)



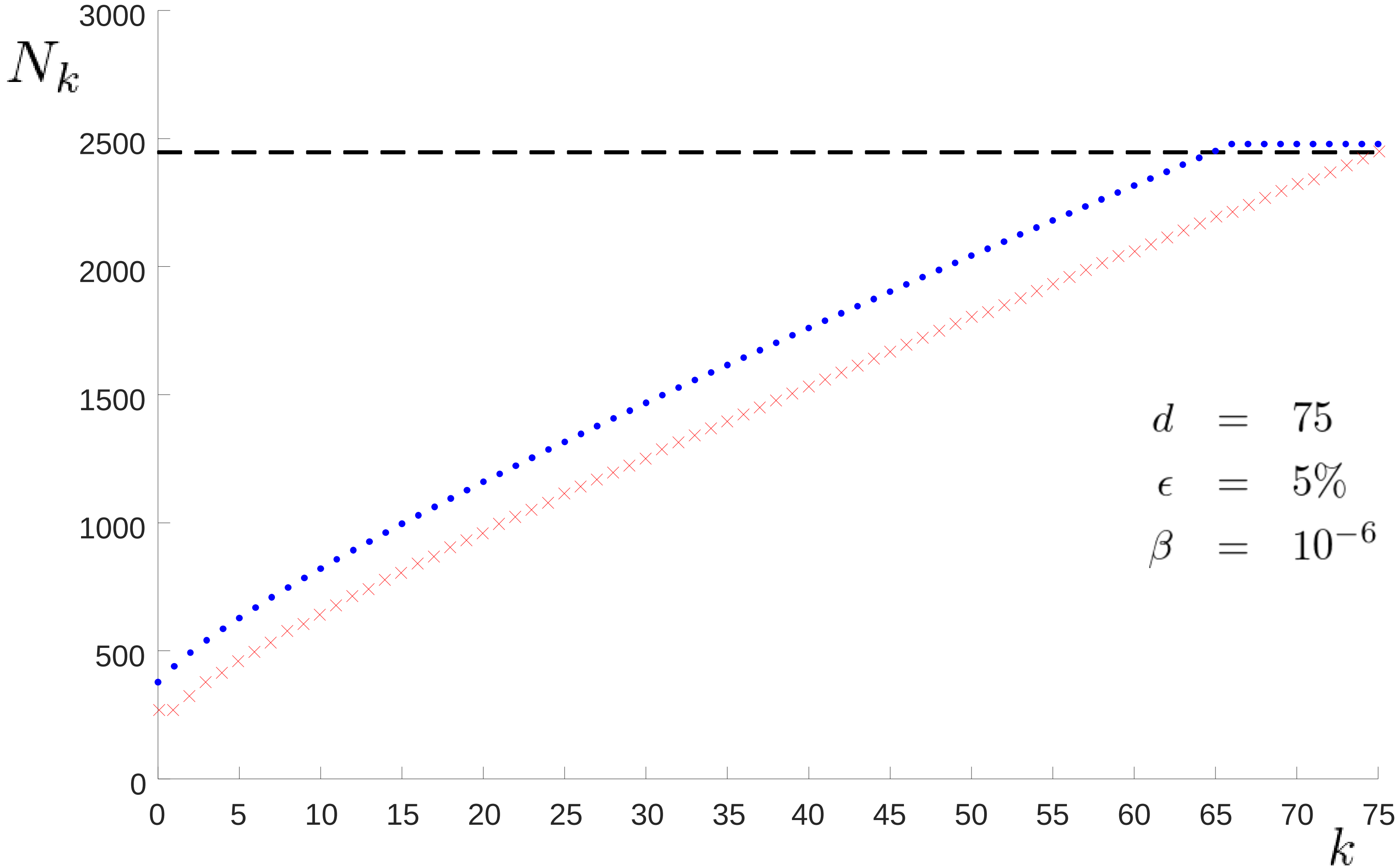
Sizing of N_0, N_1, …, N_d (cont'd)



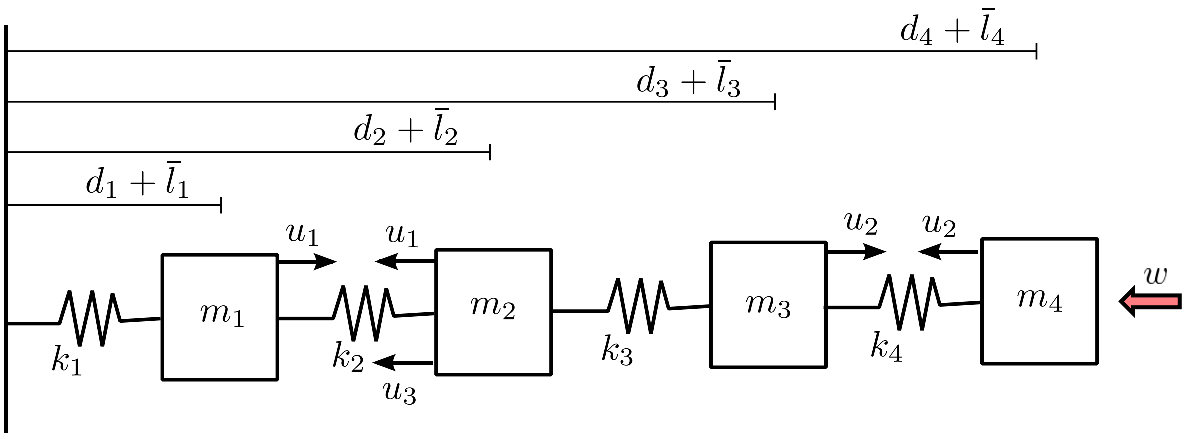
Sizing of $N_0, N_1, \ldots, N_d$ (cont'd)



Sizing of N_0, N_1, …, N_d (cont'd)



Numerical example



purified control

$$u_t = \gamma_t + \sum_{\tau=0}^{t-1} \theta_{t,\tau} w_\tau$$

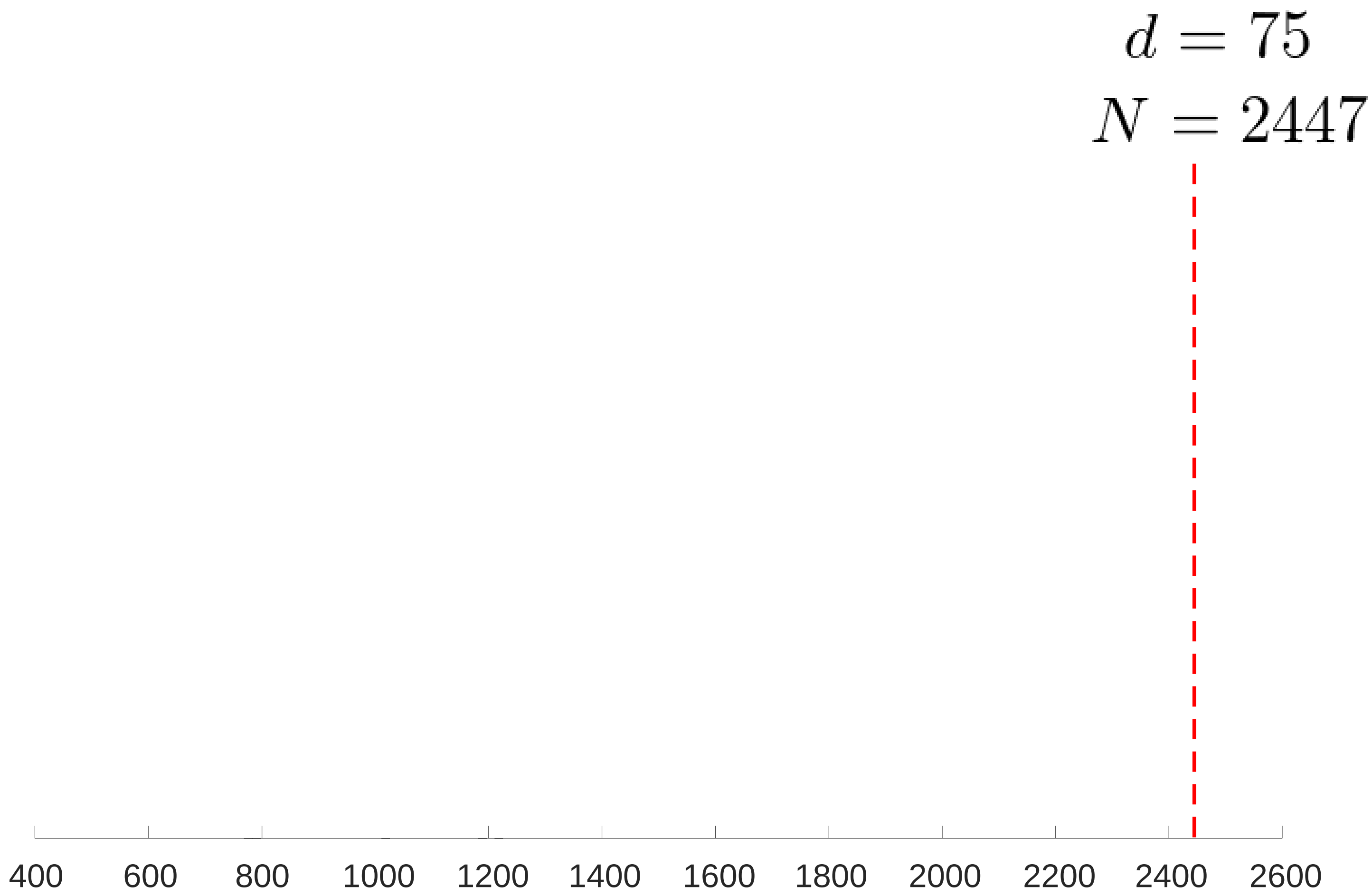
$$\min_{\gamma_t, \theta_{t,\tau}} \quad \mathbb{E}\left[\sum_{t=1}^{5} 10^{-6} \|u_{t-1}\|_2^2 + \xi_5^T Q \xi_5\right]$$

subject to:

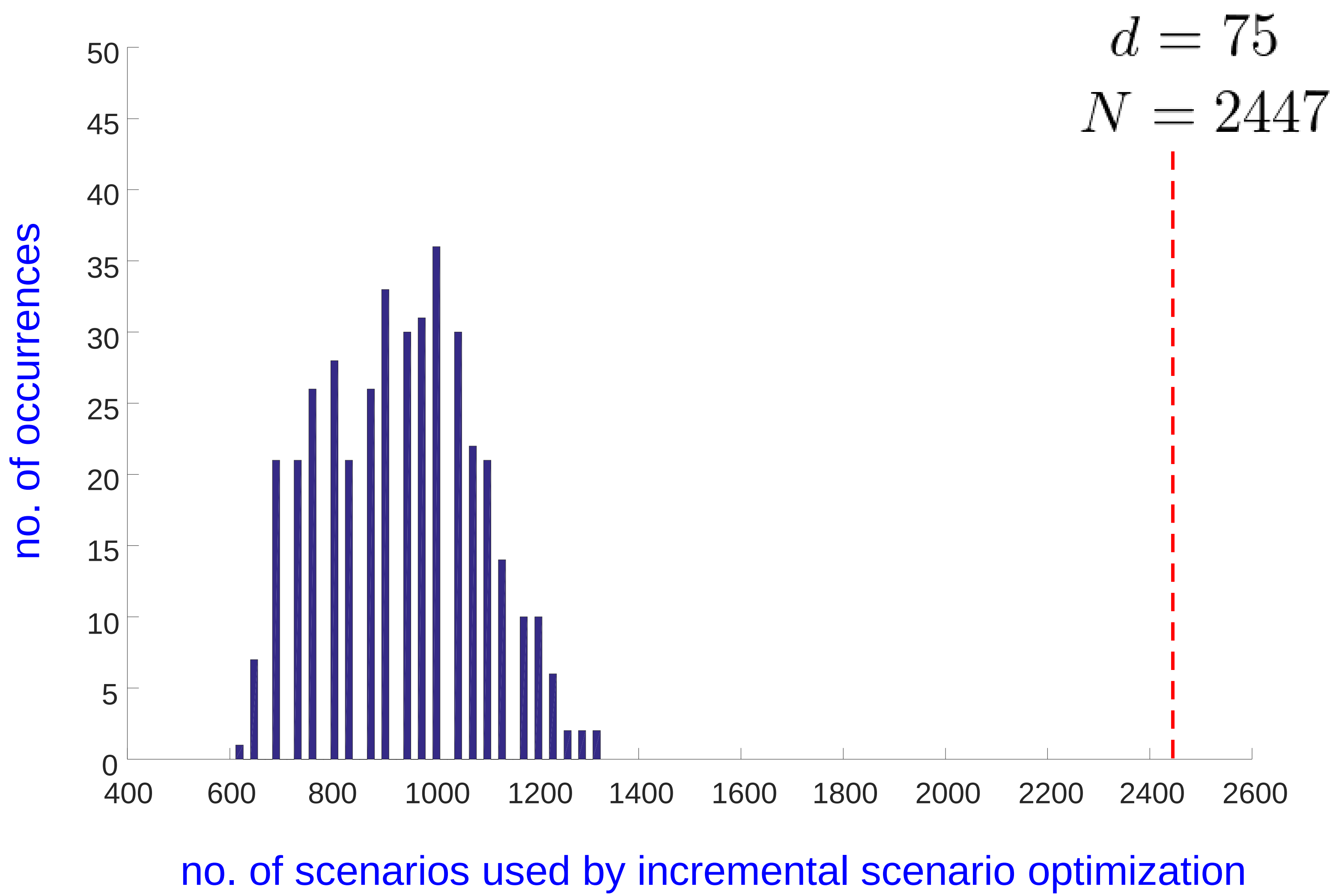
$$\sup_{t=1,\ldots,5} \left\| \begin{matrix} d_{1,t}^{(i)} \\ d_{2,t}^{(i)} - d_{1,t}^{(i)} \\ d_{3,t}^{(i)} - d_{2,t}^{(i)} \\ d_{4,t}^{(i)} - d_{3,t}^{(i)} \end{matrix} \right\|_\infty \leq 2.8, \; i = 1, \ldots, N,$$

$$\sup_{t=1,\ldots,5} \|u_{t-1}^{(i)}\|_\infty \leq 2.8, \; i = 1, \ldots, N,$$

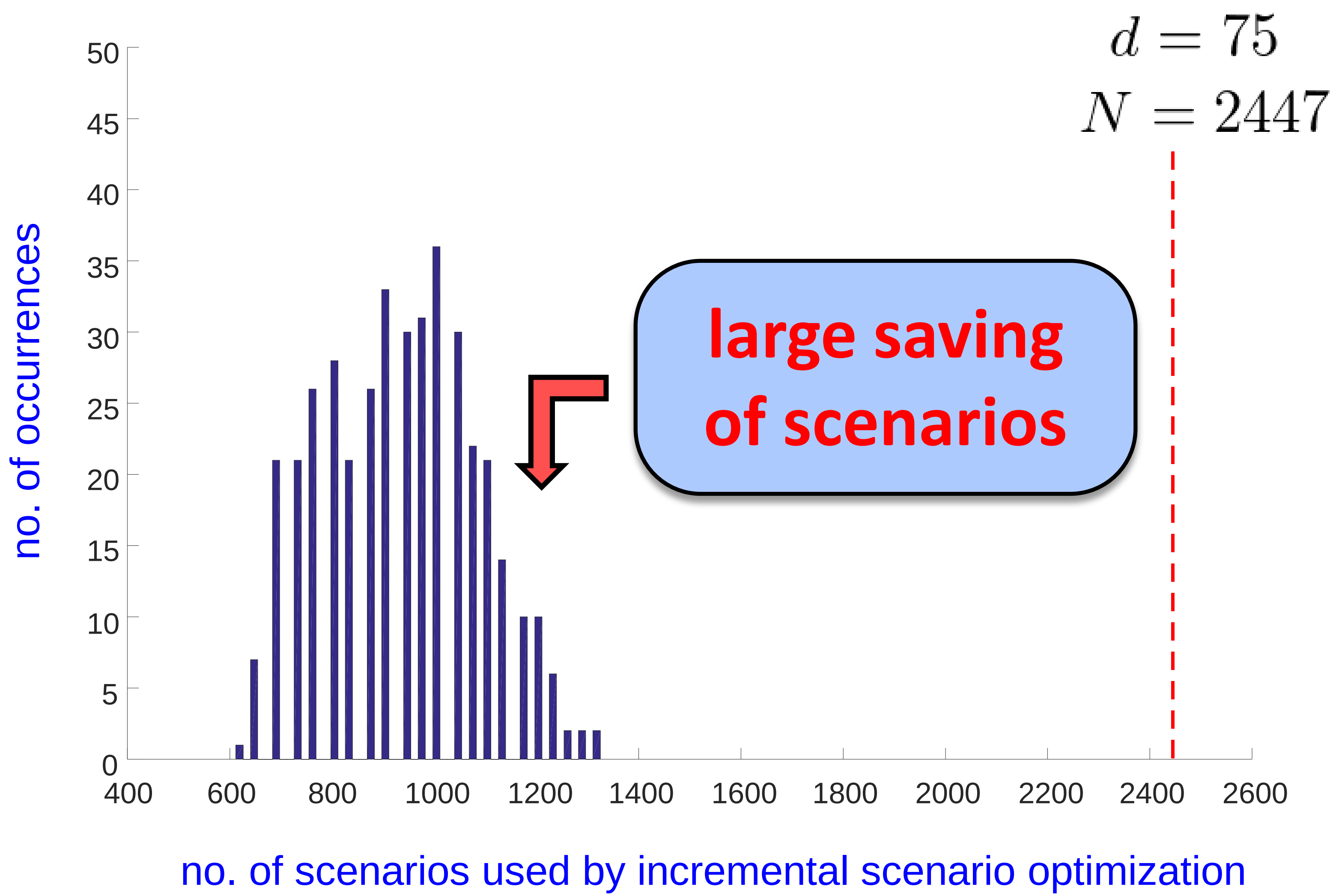
Numerical example



Numerical example



Numerical example



Conclusions

- Scenario optimization: a practical approach to data-driven optimization

- The cardinality of the support set s^* (visible) carries fundamental information on $V(x^*)$ (hidden - lack of knowledge of $\mathbb{P}$)

- Incremental scenario optimization

  → scenario theory: dataset size tuning
  large saving of data

Thank you !