An optimal scanning sensor activation policy for parameter estimation of distributed systems

Dariusz Uciński

Institute of Control and Computation Engineering

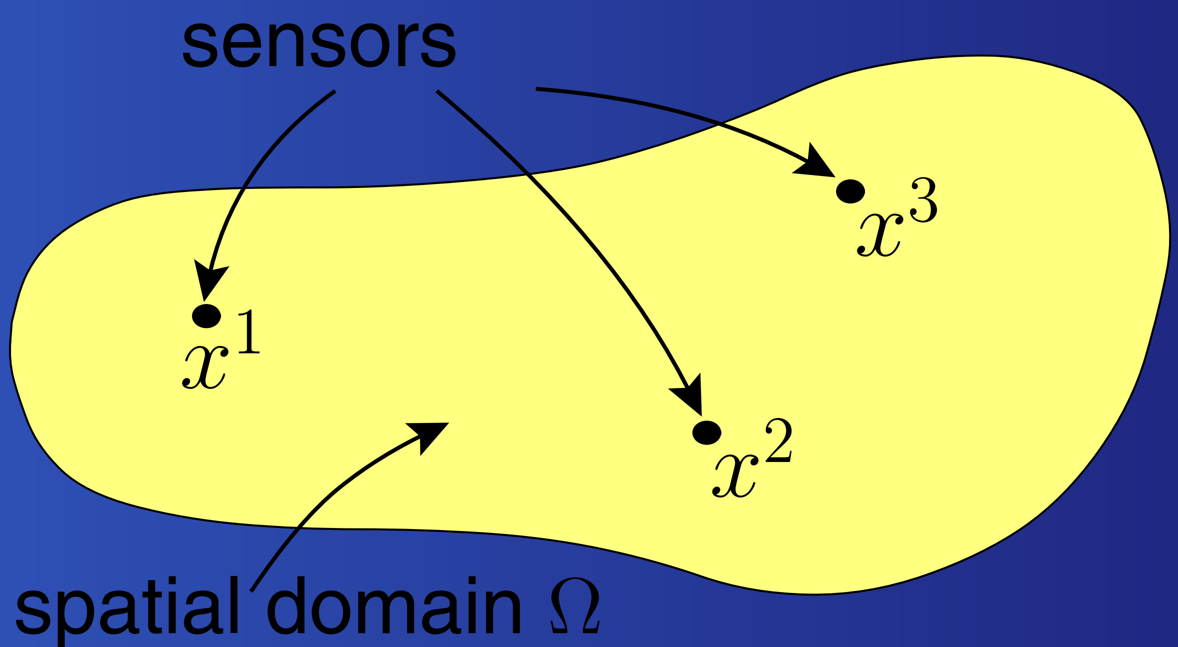
University of Zielona Góra, Poland

Spatiotemporal dynamics

Distributed parameter system—dynamic system whose state depends on both time and space; its model is known up to a vector of unknown parameters θ.

Observations—using sensors in order to estimate θ.

Subject of the talk



Problem: How to determine "optimal" sensor locations?

Motivations

- calibration of smog prediction models
- data assimilation in meteorology and oceanography
- smart material systems
- fault detection and isolation in DPSs
- groundwater resources management
- recovery of valuable minerals and hydrocarbon
- inspection in hazardous environments

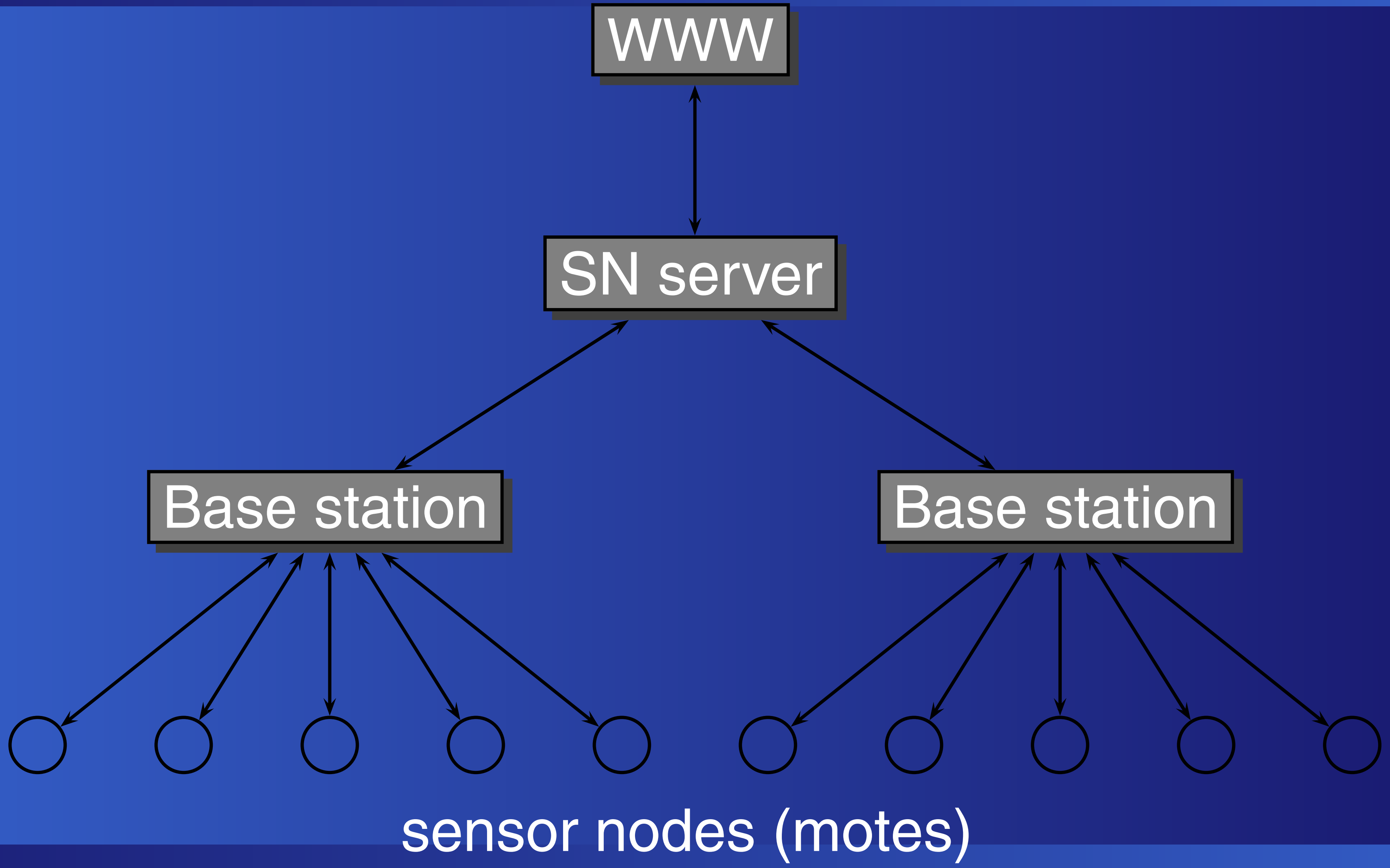
Technology: Wireless sensor networks

Sensor network—an array of sensors of diverse type interconnected by a communication network.

A large number of inexpensive, miniature and low-power SN nodes can be deployed throughout a physical space, providing dense sensing close to physical phenomena.

WSNs incorporate technologies from sensing, communication and computing.

Generic architecture



SN nodes

Motes must be low cost, low power (for long-term operation), automated (maintenance free), robust (to withstand errors and failures), and non-intrusive.

- Hardware: a microprocessor, data storage, sensors, AD-converters, a data transceiver (Bluetooth), controllers, and an energy source
- Software: TinyOS

They are already manufactured (Crossbow, Intel).

Strategies of taking measurements

- employing stationary sensors (How to determine the best sensor configuration?)

Strategies of taking measurements

- employing stationary sensors (How to determine the best sensor configuration?)
- application of spatially-movable sensors (How to design the optimal trajectories?)

Strategies of taking measurements

- employing stationary sensors (How to determine the best sensor configuration?)
- application of spatially-movable sensors (How to design the optimal trajectories?)
- scanning (How to select optimal locations of the sensors at given time instants?)

Strategies of taking measurements

- employing stationary sensors (How to determine the best sensor configuration?)
- application of spatially-movable sensors (How to design the optimal trajectories?)
- scanning (How to select optimal locations of the sensors at given time instants?)

Another problem: How to choose a minimal number of sensors?

Pros and cons

A reason for not using all the available sensors could be the reduction of the observation system complexity and the cost of operation.

- Another interpretation: mobile sensors.
- Best points are tracked.
- Technology makes it affordable.

Pros and cons

A reason for not using all the available sensors could be the reduction of the observation system complexity and the cost of operation.

- Another interpretation: mobile sensors.
- Best points are tracked.
- Technology makes it affordable.

But:

- Lack of efficient algorithms.
- Combinatorial complexity excludes naive approaches.

Existing approaches

- Conversion to a non-linear problem of state estimation (Malebranche, 1988; Korbicz *et al.*, 1988);
- Employing random fields analysis (Kazimierczyk, 1989; Sun, 1994);
- Formulation in the spirit of optimum experimental design (Uspenskii & Fedorov, 1975; Quereshi *et al.*, 1980; Rafajłowicz, 1978–1995; Kammer, 1990; 1992; Sun, 1994; Point *et al.*, 1996; Vande Wouwer *et al.*, 1999; Patan, 2004; Uciński, 1999–2007).

Existing approaches

- Conversion to a non-linear problem of state estimation (Malebranche, 1988; Korbicz *et al.*, 1988);
- Employing random fields analysis (Kazimierczyk, 1989; Sun, 1994);
- Formulation in the spirit of optimum experimental design (Uspenskii & Fedorov, 1975; Quereshi *et al.*, 1980; Rafajłowicz, 1978–1995; Kammer, 1990; 1992; Sun, 1994; Point *et al.*, 1996; Vande Wouwer *et al.*, 1999; Patan, 2004; Uciński, 1999–2007).

System description

Let $\Omega \subset \mathbb{R}^2$ be a region with boundary $\partial\Omega$.

$$\frac{\partial y}{\partial t} = \mathcal{F}(\boldsymbol{x}, t, y, \ \boldsymbol{\theta} \ ) \quad \text{in } \Omega \times T,$$

subject to appropriate I & BCs, where

- $\boldsymbol{x}$ – spatial variable, t – time, $T = (0, t_f)$;
- $y = y(\boldsymbol{x}, t)$ – state variable;
- $\mathcal{F}$ – well-posed differential operator which involves spatial derivatives;
- $\boldsymbol{\theta} \in \mathbb{R}^m$ – vector of *unknown* parameters.

System description

Let $\Omega \subset \mathbb{R}^2$ be a region with boundary $\partial\Omega$.

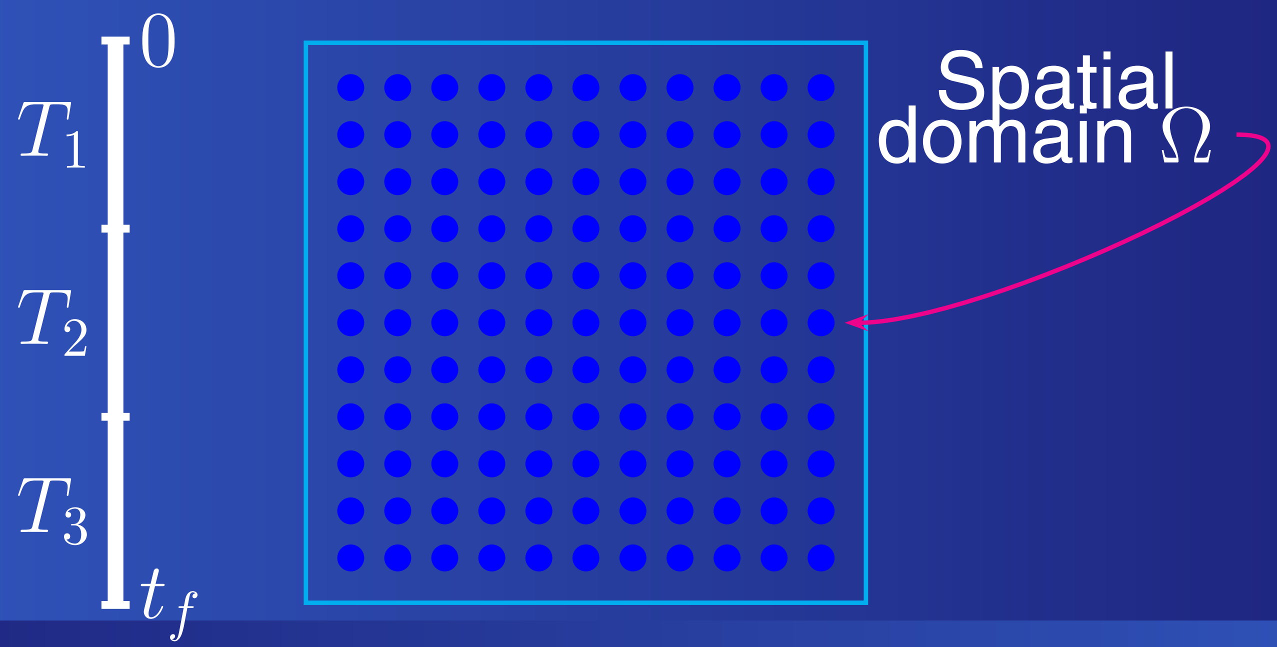
$$\frac{\partial y}{\partial t} = \mathcal{F}(\boldsymbol{x}, t, y, \boldsymbol{\theta}) \quad \text{in } \Omega \times T,$$

subject to appropriate I & BCs, where

- $\boldsymbol{x}$ – spatial variable, t – time, $T = (0, t_f)$;
- $y = y(\boldsymbol{x}, t)$ – state variable;
- $\mathcal{F}$ – well-posed differential operator which involves spatial derivatives;
- $\boldsymbol{\theta} \in \mathbb{R}^m$ – vector of *unknown* parameters.

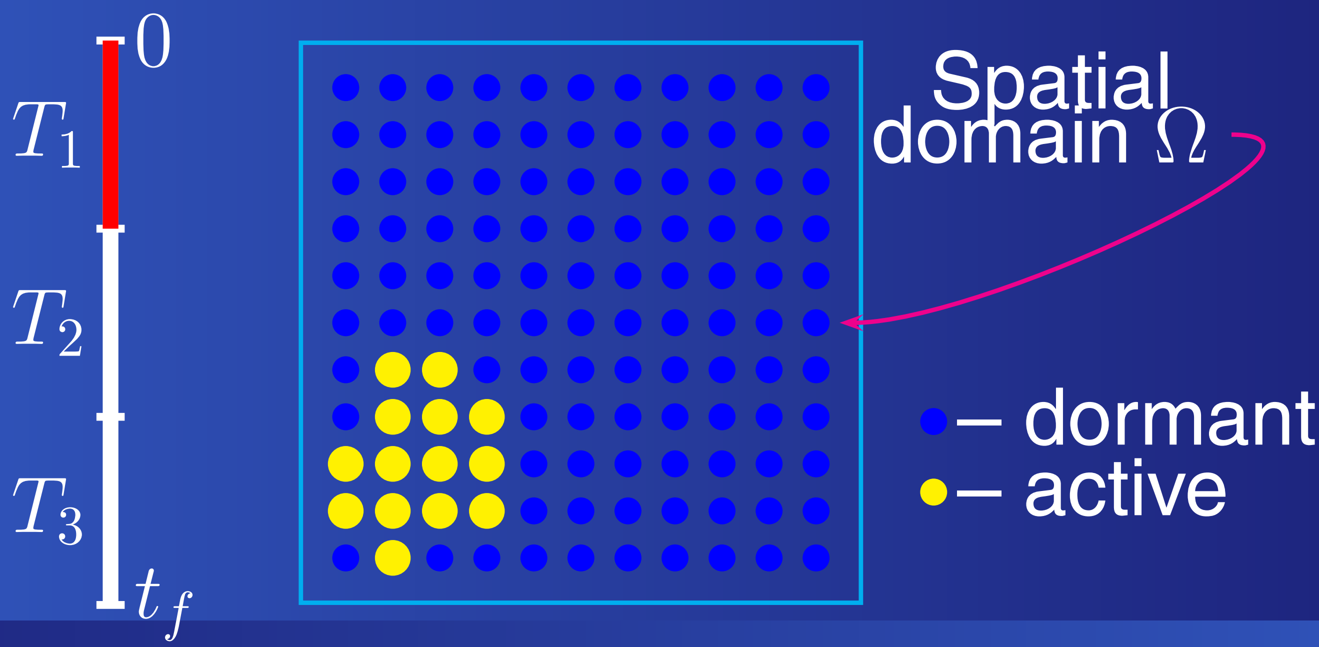
Scanning observations

Given I stationary located sensors, partition the observation horizon $T = (0, t_f)$ into subintervals T_k, $k = 1, \ldots, K$ and activate a best n-element subset of sensors over each T_k.



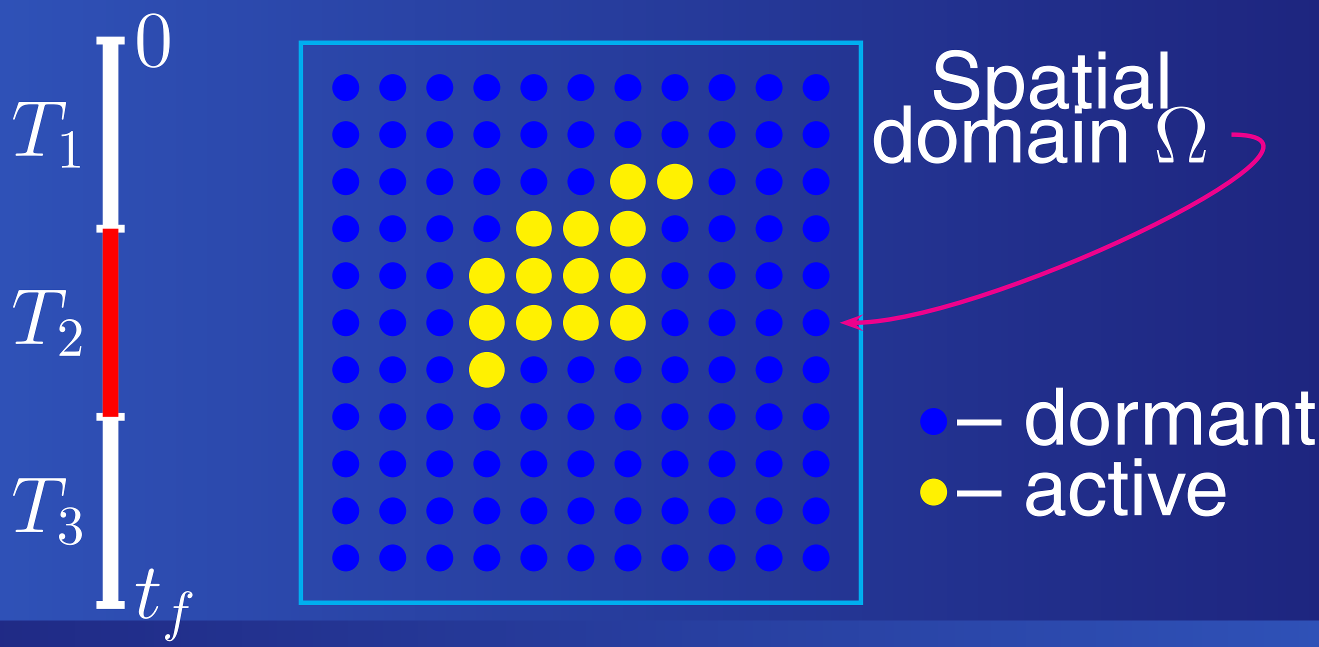
Scanning observations

Given I stationary located sensors, partition the observation horizon $T = (0, t_f)$ into subintervals T_k, $k = 1, \dots, K$ and activate a best n-element subset of sensors over each T_k.



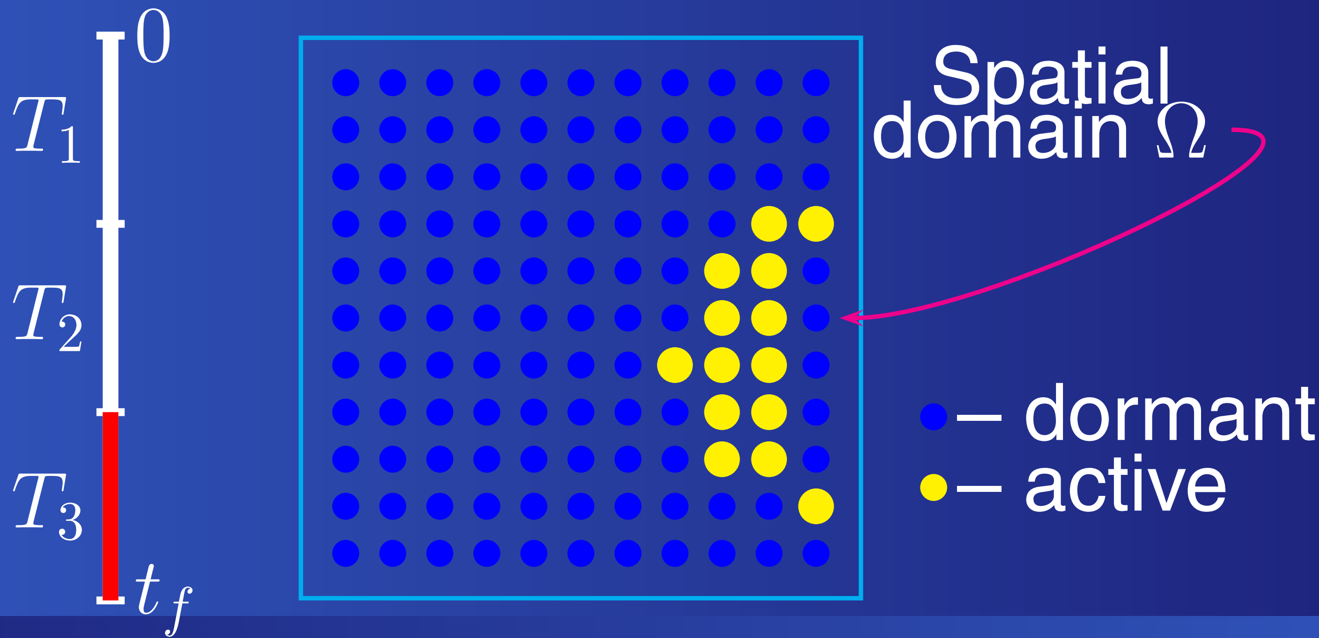
Scanning observations

Given I stationary located sensors, partition the observation horizon $T = (0, t_f)$ into subintervals T_k, $k = 1, \dots, K$ and activate a best n-element subset of sensors over each T_k.



Scanning observations

Given I stationary located sensors, partition the observation horizon $T = (0, t_f)$ into subintervals T_k, $k = 1, \dots, K$ and activate a best n-element subset of sensors over each T_k.



Output equation

At Stage k we have time subinterval T_k and

$$z_{ki}(t) = v_{ki}\left[y(\boldsymbol{x}^i, t; \boldsymbol{\theta}) + \varepsilon(\boldsymbol{x}^i, t)\right]$$

for $t \in T_k$ and $i = 1, \dots, I$, where $\varepsilon(\,\cdot\,,\,\cdot\,)$ – white Gaussian measurement noise,

$$v_{ki} = \begin{cases} 1 & \text{if the } i\text{-th sensor is active over } T_k \\ 0 & \text{otherwise} \end{cases}$$

Least-squares criterion

The LS estimate of $\boldsymbol{\theta}$ is the one which minimizes

$$\mathcal{J}(\boldsymbol{\theta}) = \sum_{i=1}^{I} \sum_{k=1}^{K} v_{ki} \int_{T_k} \big[z_{ki}(t) - \hat{y}(\boldsymbol{x}^i, t; \boldsymbol{\theta}) \big]^2 \, \mathrm{d}t$$

where $\hat{y}(\,\cdot\,,\,\cdot\,;\boldsymbol{\theta})$ stands for the solution to the state equation corresponding to a given value of $\boldsymbol{\theta}$.

Optimal measurement problem

Determine sensors to be activated so as to maximize the identification accuracy.

Optimal measurement problem

Determine sensors to be activated so as to maximize the identification accuracy. $\longleftarrow$ What is it?

Optimal measurement problem

Determine sensors to be activated so as to maximize the identification accuracy.

Make use of the Cramér-Rao inequality:

$$\operatorname{cov}\hat{\boldsymbol{\theta}} = E\Big\{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T\Big\} \succeq M^{-1}$$

We have $\operatorname{cov}\hat{\boldsymbol{\theta}} = M^{-1}$ provided that an estimator is efficient.

Optimal measurement problem

Determine sensors to be activated so as to maximize the identification accuracy.

Make use of the Cramér-Rao inequality:

$$\operatorname{cov}\hat{\boldsymbol{\theta}} = E\Big\{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T\Big\} \succeq \boldsymbol{M}^{-1}$$

We have $\operatorname{cov}\hat{\boldsymbol{\theta}} = \boldsymbol{M}^{-1}$ provided that an estimator is efficient. But what is $\boldsymbol{M}$?

Fisher Information Matrix (FIM)

$$M(v) = \sum_{k=1}^{K} \sum_{i=1}^{I} v_{ki} M_{ki}$$

$$M_{ki} = \int_{T_k} g(x^i, t) \; g^{\mathsf{T}}(x^i, t) \; \mathrm{d}t$$

where $g(x^i, t) = \big(\nabla_{\theta} y\big)(x^i, t; \theta^0)$ are sensitivity coefficients

Fisher Information Matrix (FIM)

$$\boldsymbol{M}(\boldsymbol{v}) = \sum_{k=1}^{K} \sum_{i=1}^{I} v_{ki} \boldsymbol{M}_{ki}$$

$$\boldsymbol{M}_{ki} = \int_{T_k} \boldsymbol{g}(\boldsymbol{x}^i, t)\, \boldsymbol{g}^{\mathsf{T}}(\boldsymbol{x}^i, t)\, \mathrm{d}t$$

where $\boldsymbol{g}(\boldsymbol{x}^i, t) = \big(\nabla_{\boldsymbol{\theta}} y\big)(\boldsymbol{x}^i, t; \boldsymbol{\theta}^0)$ are sensitivity coefficients

Ultimate formulation

Find $\boldsymbol{v} = (v_{11}, \dots, v_{1I}, \dots, v_{K1}, \dots, v_{KI})$ s.t.

$$\mathcal{P}(\boldsymbol{v}) = \log\det\left(\sum_{k=1}^{K}\sum_{i=1}^{I} v_{ki}\boldsymbol{M}_{ki}\right) \to \max$$

subject to the constraints

$$\sum_{i=1}^{I} v_{ki} = n, \qquad k = 1, \dots, K$$

$$v_{ki} \in \{0, 1\}, \quad i = 1, \dots, I, \quad k = 1, \dots, K$$

Branch-and-bound

Consider two sensors of which only one may be active, and two time stages. Any solution can be represented as a 2×2 matrix with binary entries

	$i = 1$	$i = 2$
$k = 1$	0	1
$k = 2$	1	0

i.e., rows correspond to consecutive time stages, and columns are associated with individual sensors.

Branch-and-bound

Thus we have four admissible solutions:

0	1
0	1

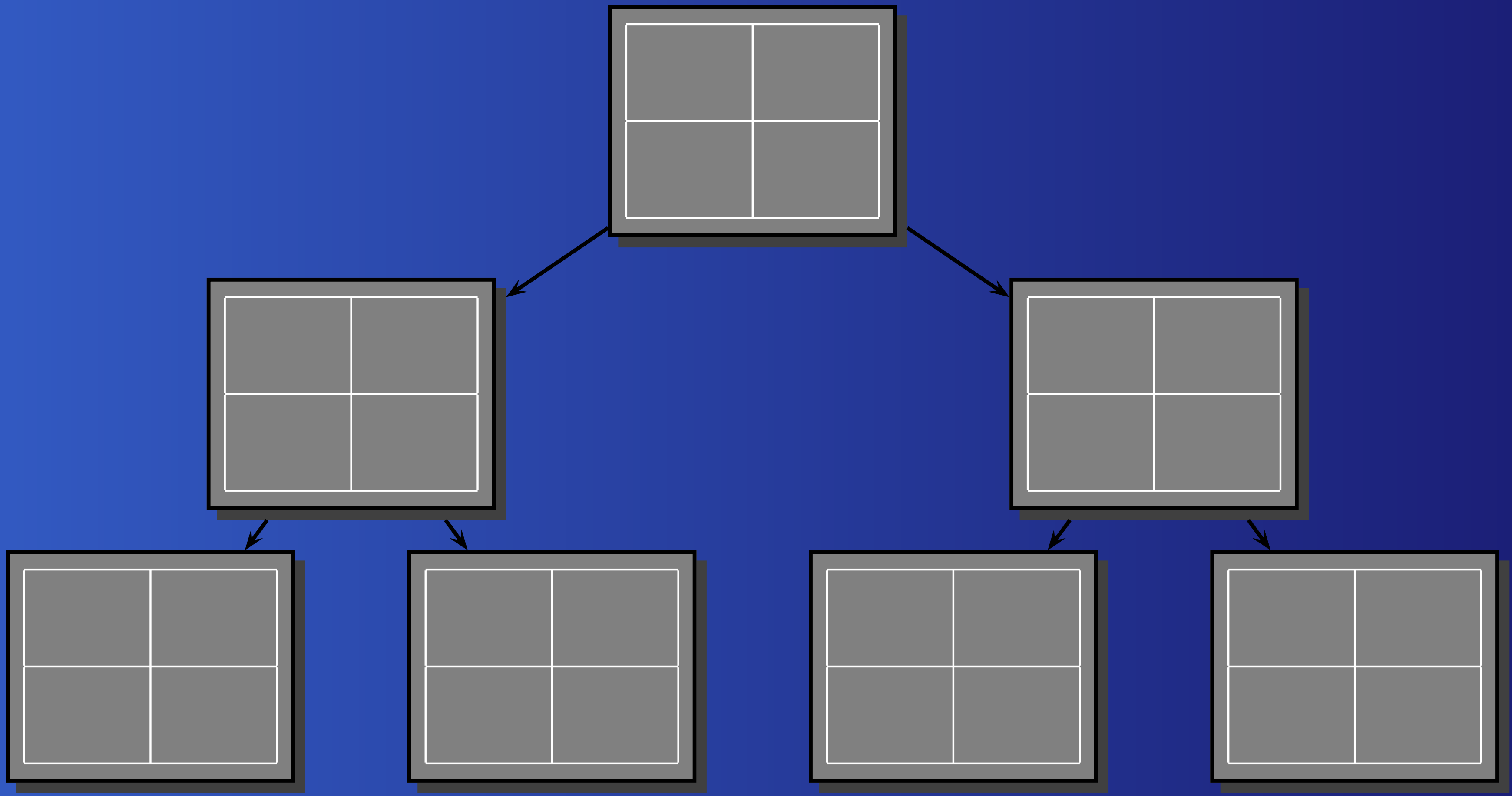
0	1
1	0

1	0
0	1

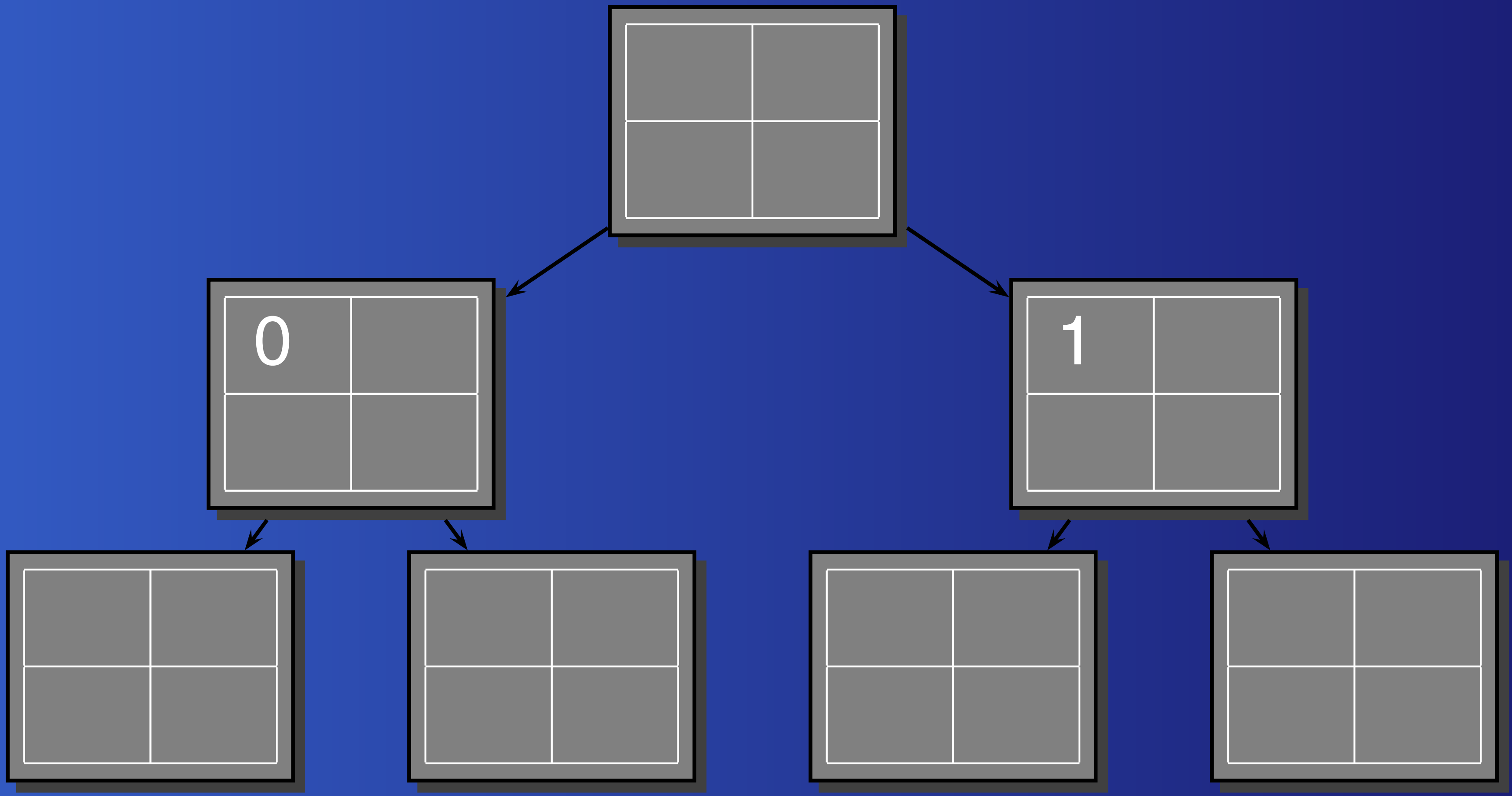
1	0
1	0

How to automatically enumerate them?

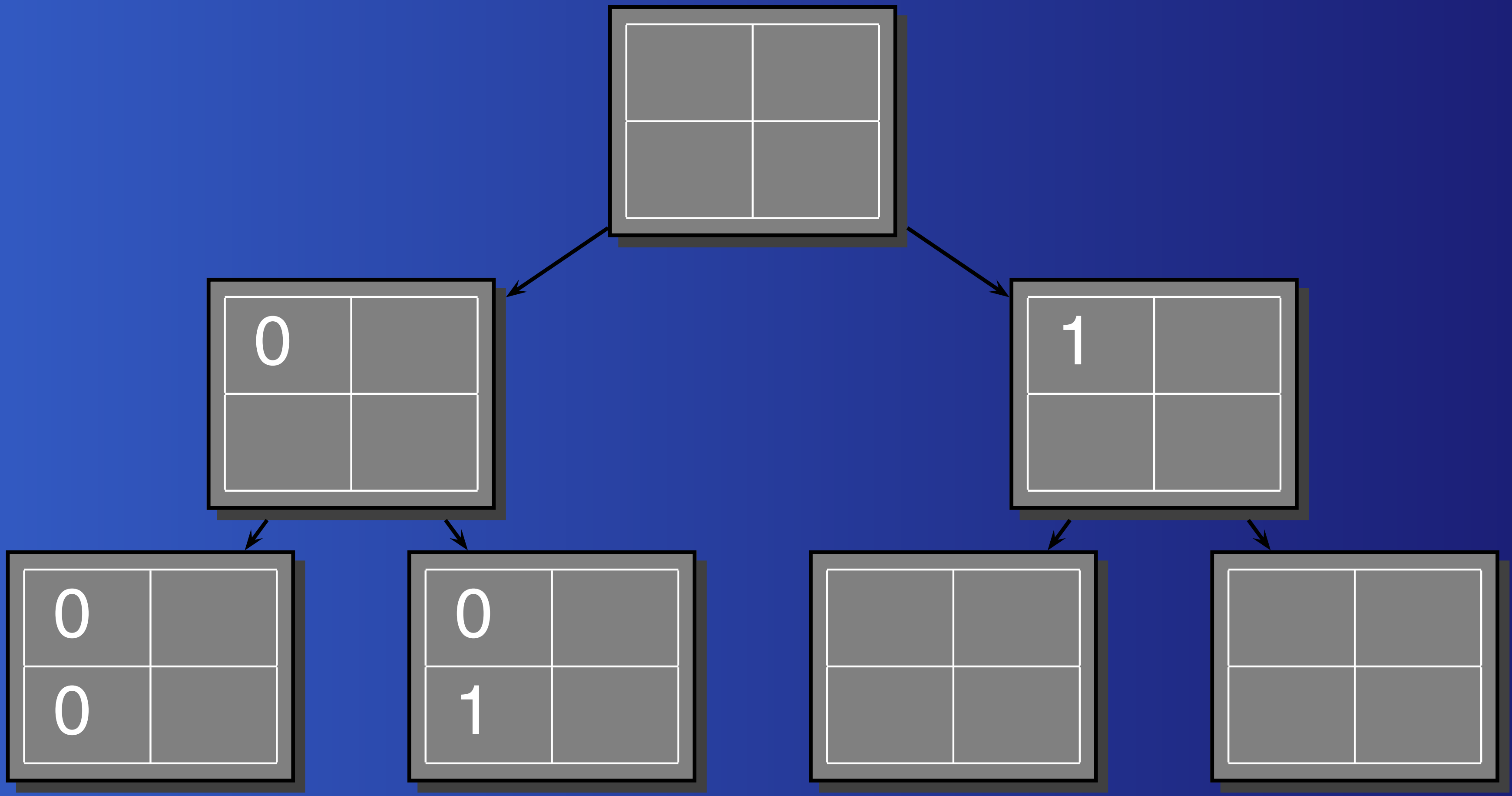
Branching



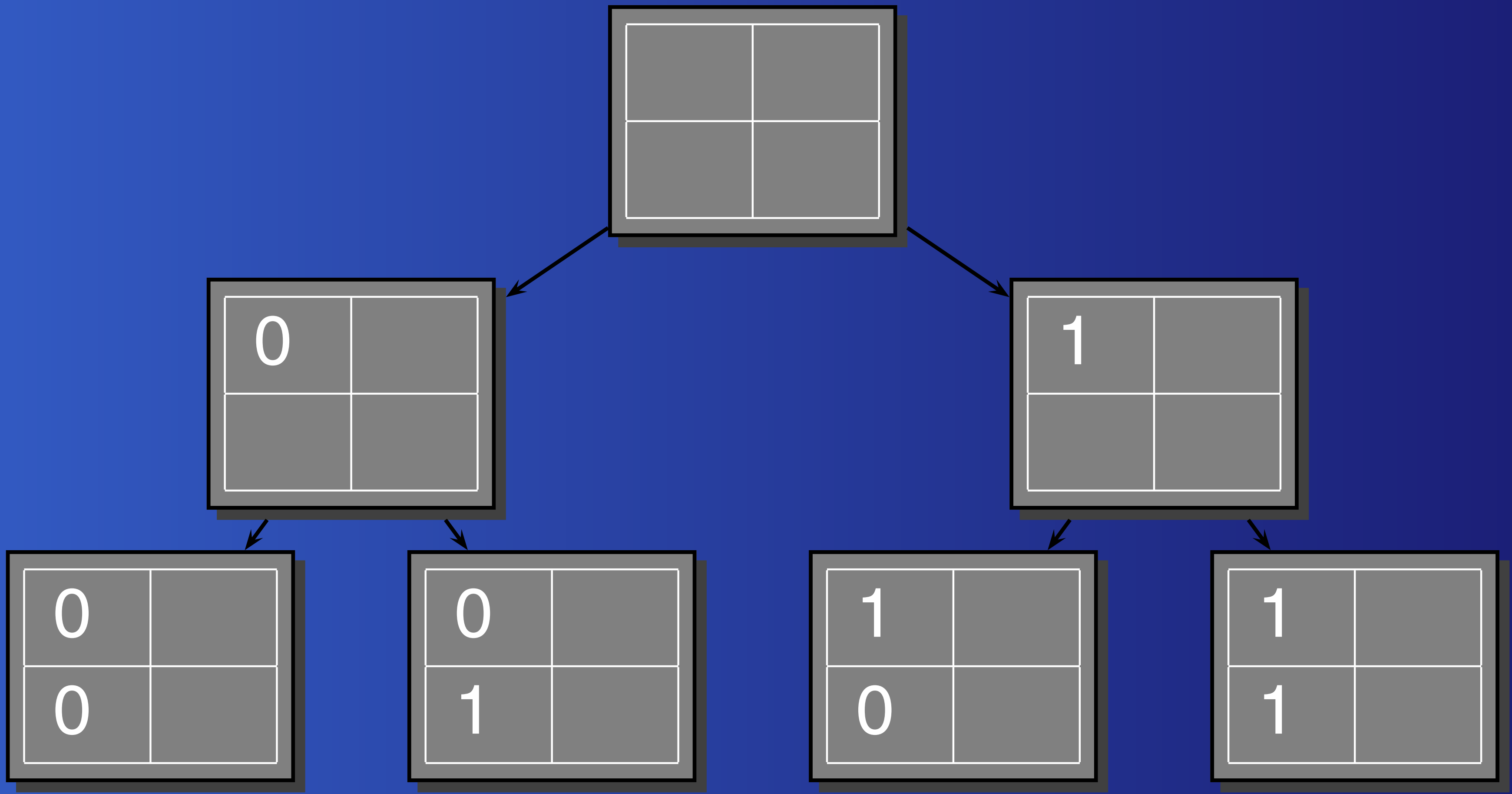
Branching



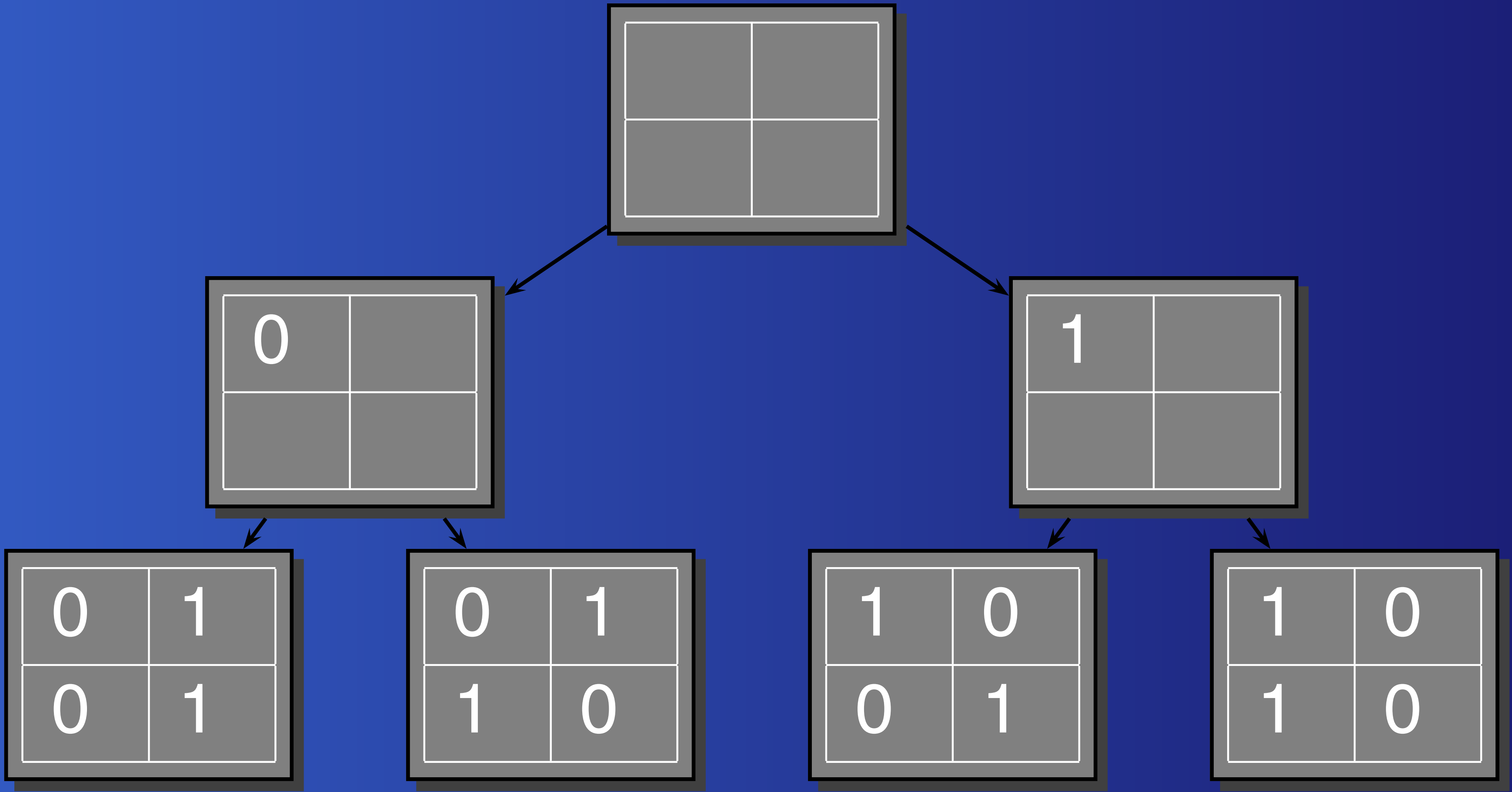
Branching



Branching

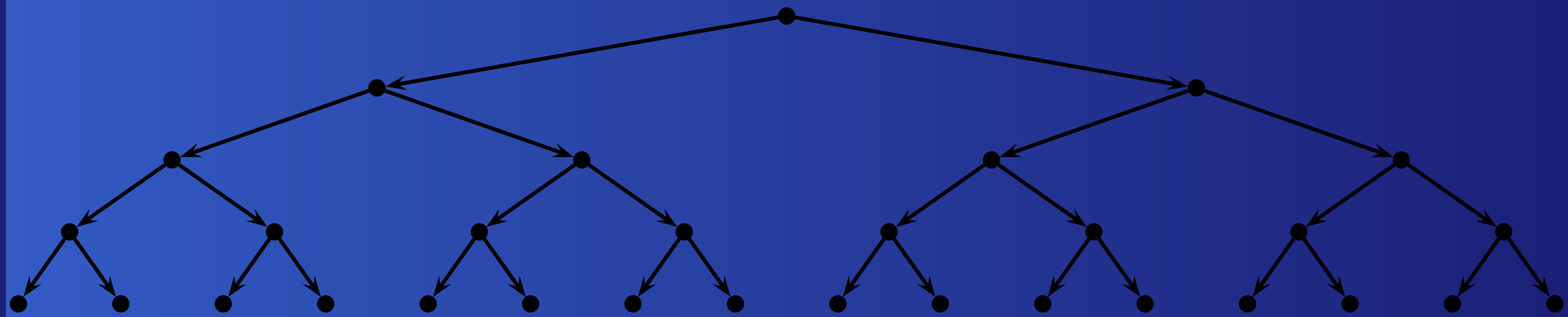


Branching



BB trees

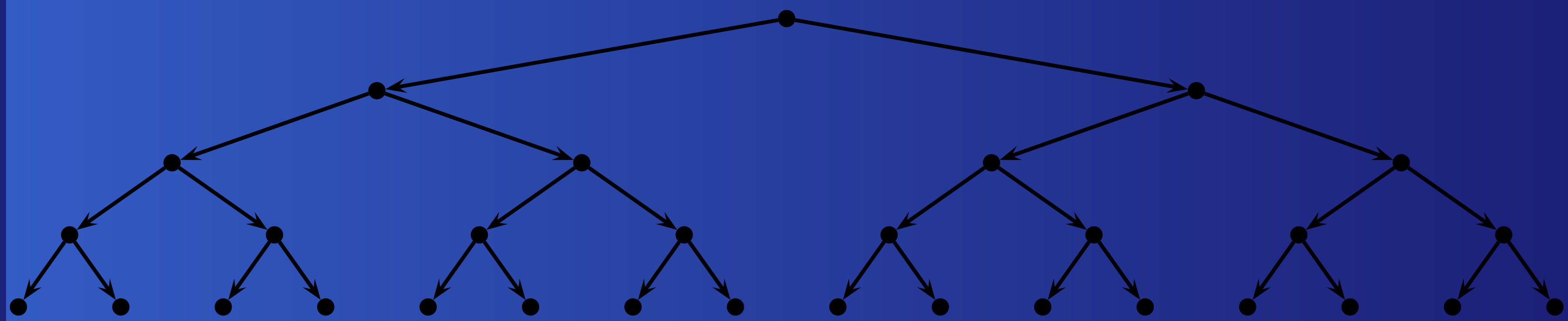
Clearly, binary trees can be larger:



For example, for 1000 sites, 500 active sensors and 10 stages, the number of admissible solutions is $\binom{1000}{500}^{10} > 10^{2994}$.

BB trees

Clearly, binary trees can be larger:



For example, for 1000 sites, 500 active sensors and 10 stages, the number of admissible solutions is $\binom{1000}{500}^{10} > 10^{2994}$. How to reduce such huge numbers?

Idea of bounding

Economize computations by eliminating subtrees that have no chance of containing an optimal solution.

Idea of bounding

Economize computations by eliminating subtrees that have no chance of containing an optimal solution.

Idea: Assume that given a node of the BB tree, we are able to cheaply find an upper bound to the maximum value of the objective function which can be obtained for the terminal nodes being its descendant nodes.

Idea of bounding

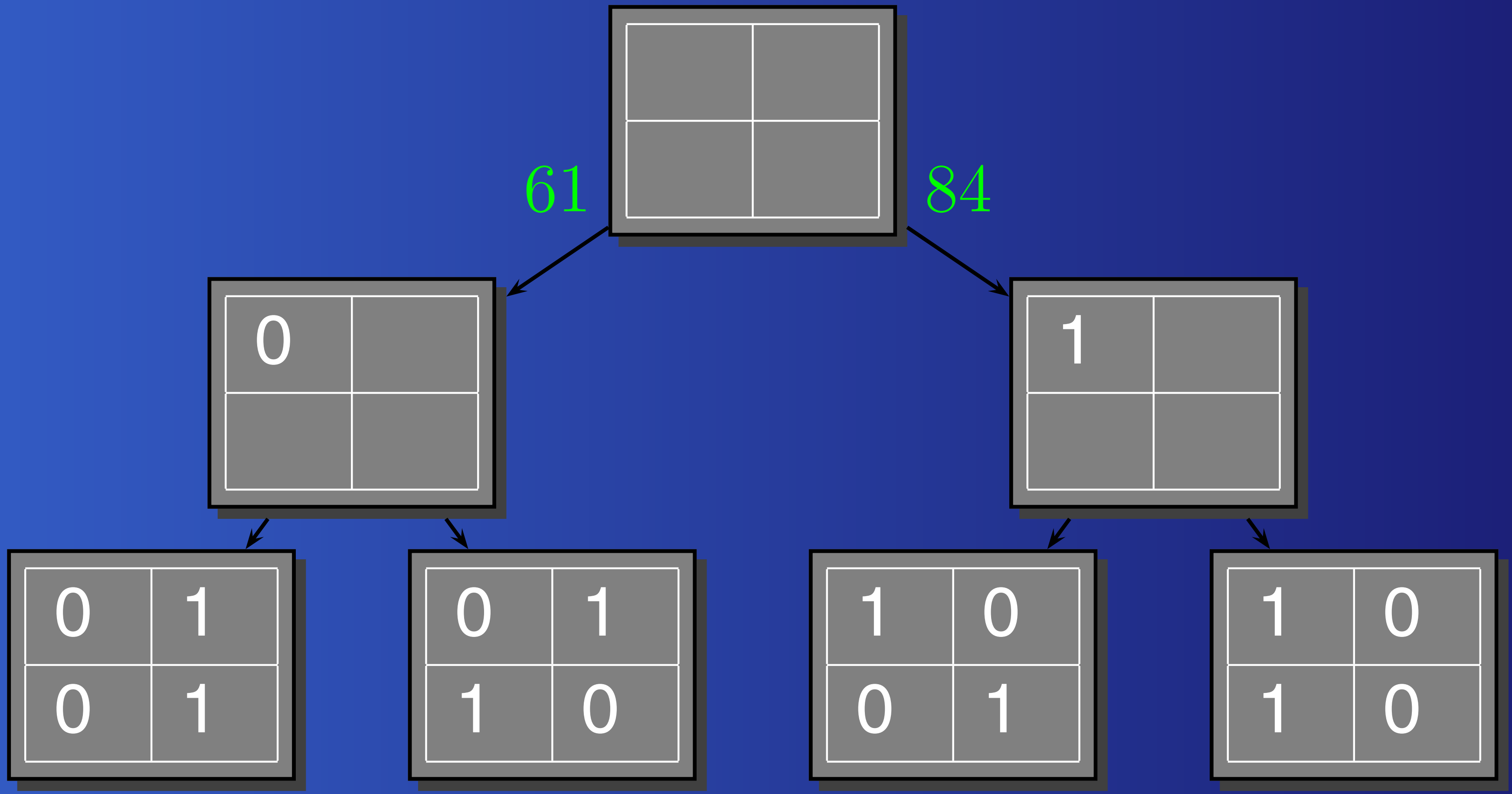
Economize computations by eliminating subtrees that have no chance of containing an optimal solution.

Idea: Assume that given a node of the BB tree, we are able to cheaply find an upper bound to the maximum value of the objective function which can be obtained for the terminal nodes being its descendant nodes.

Moreover, assume that we also know a lower bound to the maximum value of the objective function over all admissible solutions.

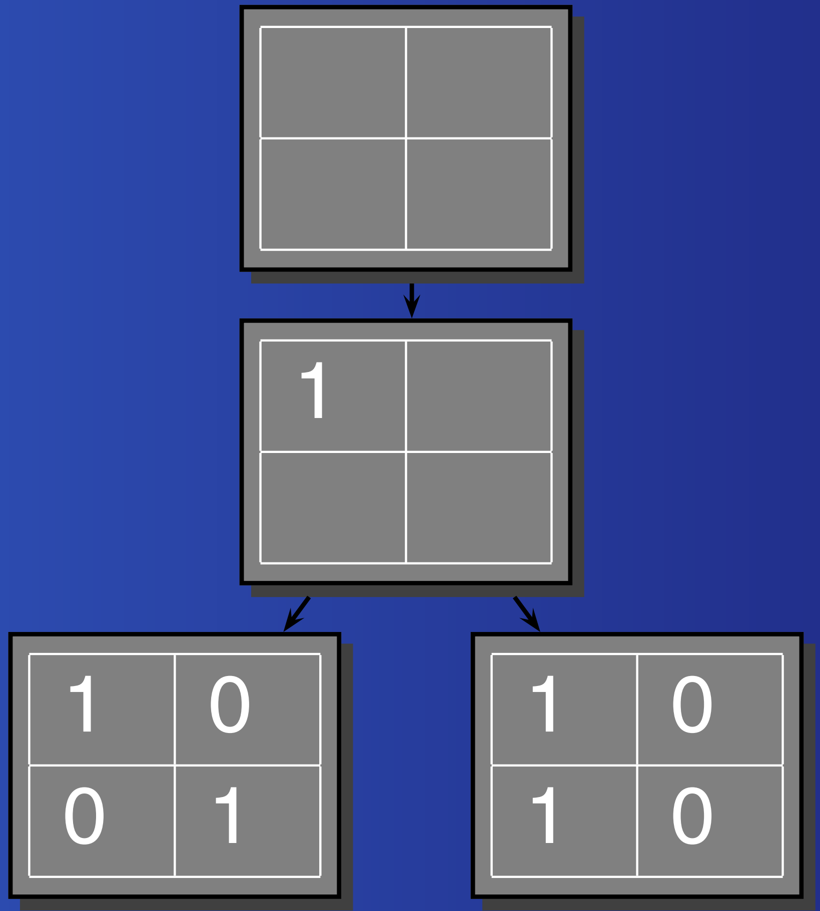
Bounding

Let $LOWER = 78$.



Pruning

We may thus discard the left subtree.



Finding upper bounds

Let

$$E = \underbrace{\{1, \ldots, K\}}_{\text{time}} \times \underbrace{\{1, \ldots, I\}}_{\text{space}}$$

For each BB node, define $E_0, E_1 \subset E$ s.t.

- $(k, i) \in E_0 \Rightarrow i$-th sensor is dormant over T_k,
- $(k, i) \in E_1 \Rightarrow i$-th sensor is active over T_k,
- $(k, i) \in E \setminus (E_0 \cup E_1) \Rightarrow$ the status of i-th sensor over T_k is not determined.

Relaxed problem

Find $\boldsymbol{v} = (v_{11}, \dots, v_{1I}, \dots, v_{K1}, \dots, v_{KI})$ s.t.

$$\mathcal{P}(\boldsymbol{v}) = \log\det\left(\sum_{k=1}^{K}\sum_{i=1}^{I} v_{ki}\boldsymbol{M}_{ki}\right) \longrightarrow \max$$

$$\sum_{i=1}^{I} v_{ki} = n, \qquad k = 1, \dots, K$$

$$v_{ki} = 0, \quad (k, i) \in E_0$$

$$v_{ki} = 1, \quad (k, i) \in E_1$$

$$0 \le v_{ki} \le 1, \quad (k, i) \in E \setminus (E_0 \cup E_1)$$

Conveniently altered formulation

Find $\boldsymbol{w} = (w_{1,1}, \dots, w_{1,q_1}, \dots, w_{L,1}, \dots, w_{L,q_L})$ s.t.

$$\mathcal{Q}(\boldsymbol{w}) = \log\det\left(\boldsymbol{A} + \sum_{l=1}^{L}\sum_{j=1}^{q_l} w_{lj}\boldsymbol{S}_{lj}\right) \to \max$$

$$\sum_{j=1}^{q_l} w_{lj} = r_l, \quad l = 1, \dots, L,$$

$$0 \le w_{lj} \le 1, \quad j = 1, \dots, q_l, \quad l = 1, \dots, L,$$

The admissible set is a polygon!

Optimality conditions

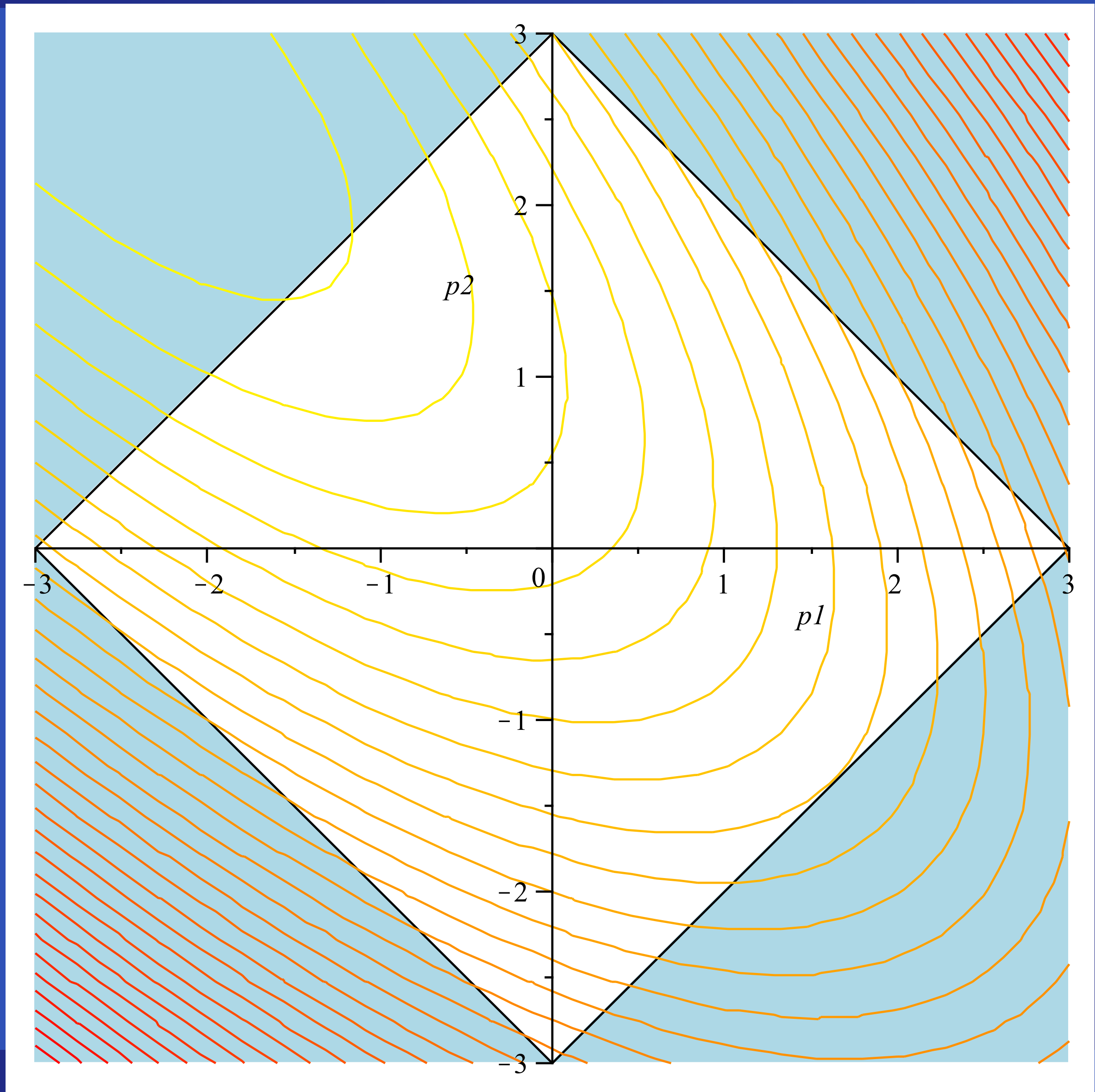
Proposition 1. *A vector* $\boldsymbol{w}^\star$ *is a global solution iff there exist numbers* $\lambda_l^\star$, $l = 1, \dots, L$ *such that*

$$\varphi(l, j, \boldsymbol{w}^\star) \begin{cases} \geq \lambda_l^\star & \text{if } w_{lj}^\star = 1 \\ = \lambda_l^\star & \text{if } 0 < w_{lj}^\star < 1 \\ \leq \lambda_l^\star & \text{if } w_{lj}^\star = 0 \end{cases}$$

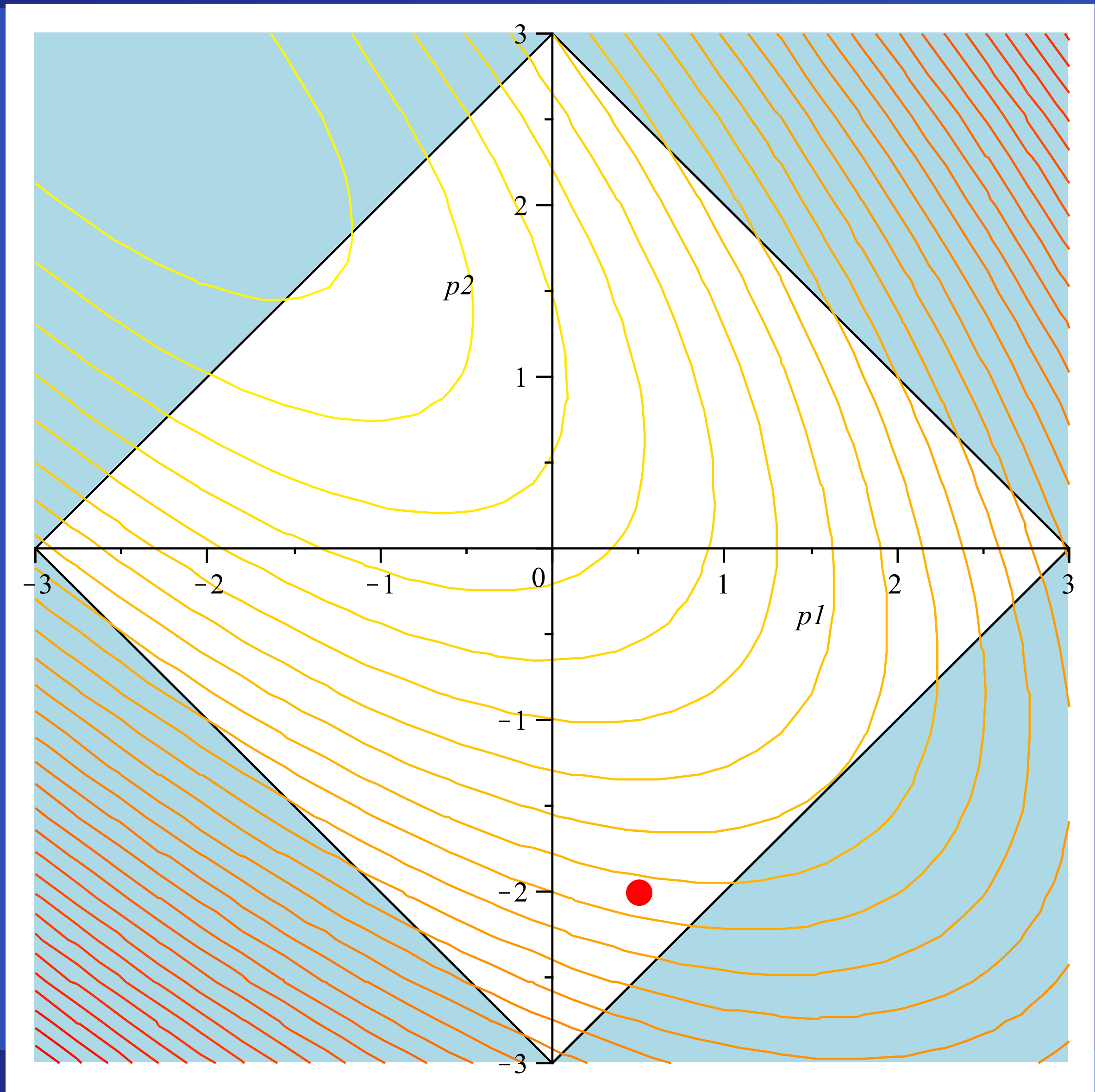
where $\varphi(l, j, \boldsymbol{w}) = \operatorname{tr}\big[\boldsymbol{G}^{-1}(\boldsymbol{w})\boldsymbol{S}_{lj}\big]$,

$$\boldsymbol{G}(\boldsymbol{w}) = \boldsymbol{A} + \sum_{l=1}^{L}\sum_{j=1}^{q_l} w_{lj}\boldsymbol{S}_{lj}$$

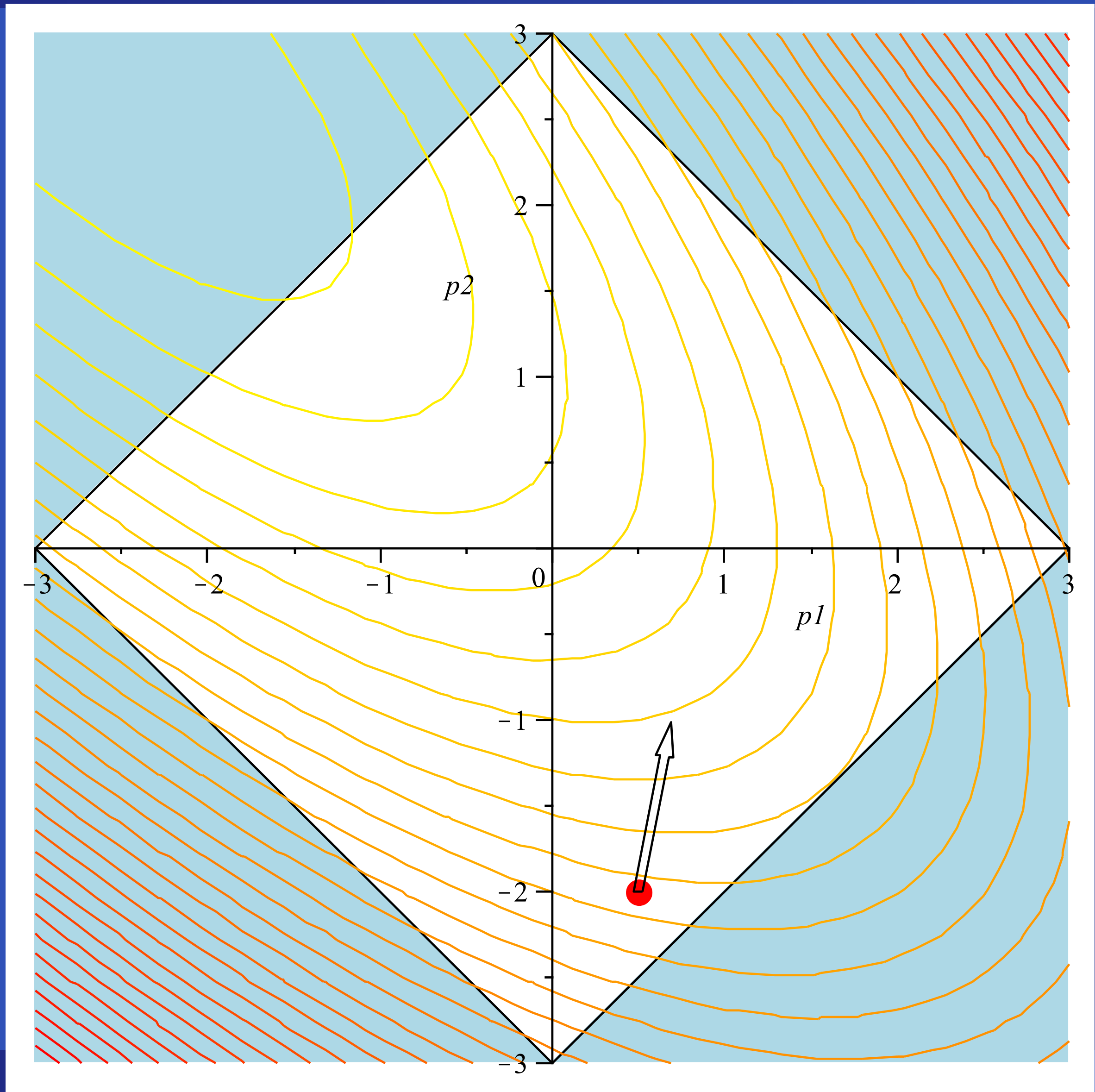
Simplicial decomposition



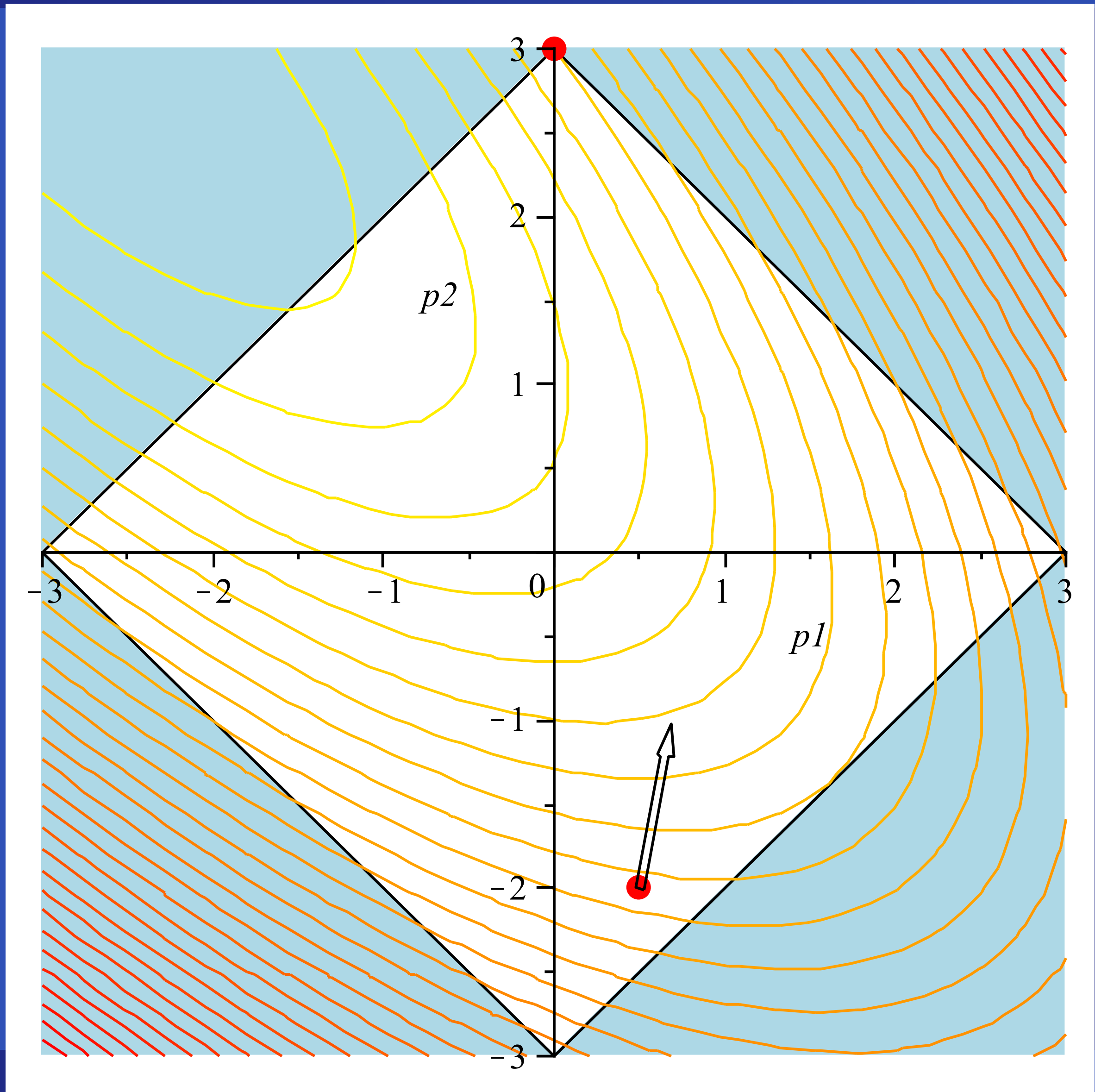
Simplicial decomposition



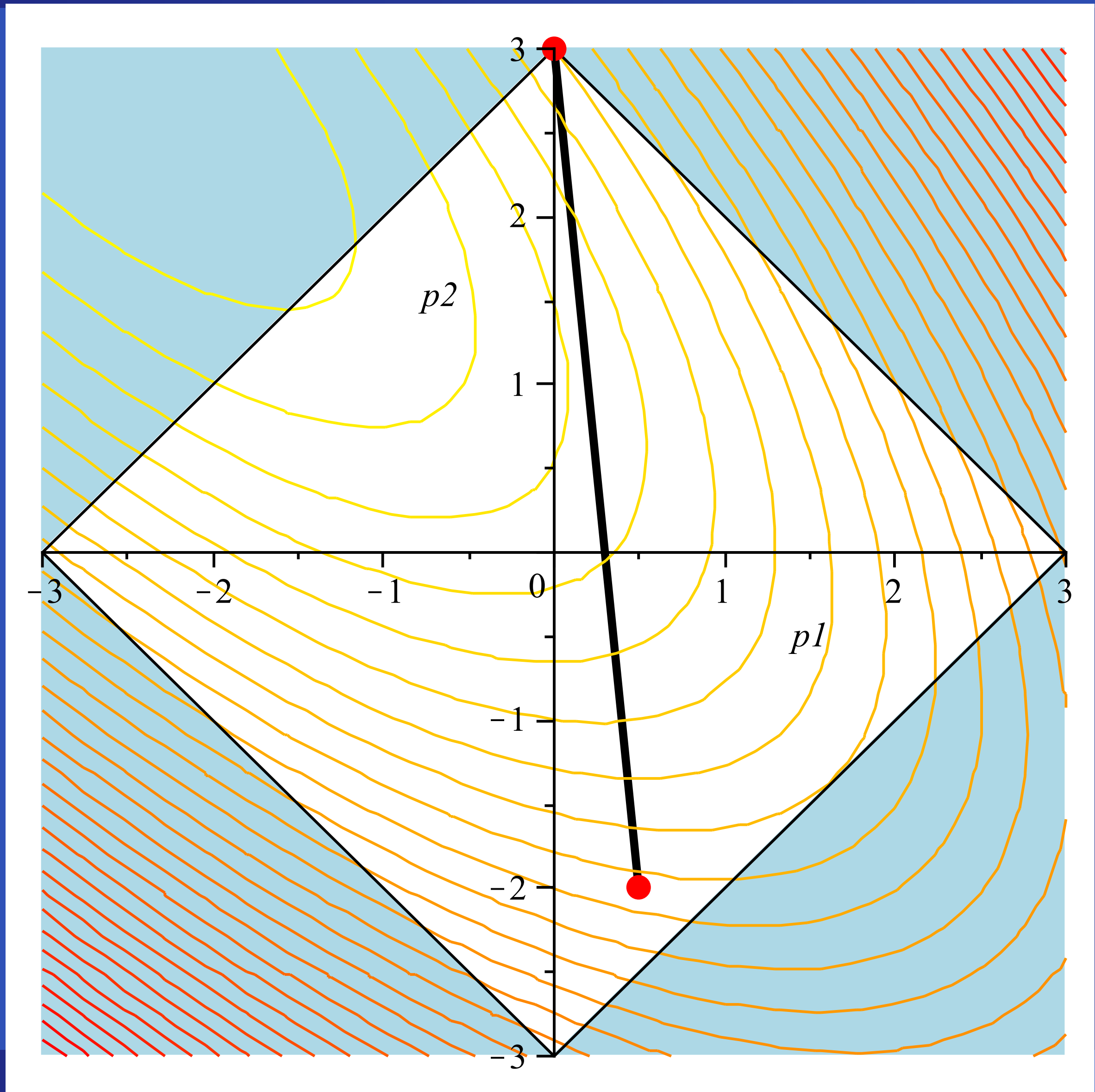
Simplicial decomposition



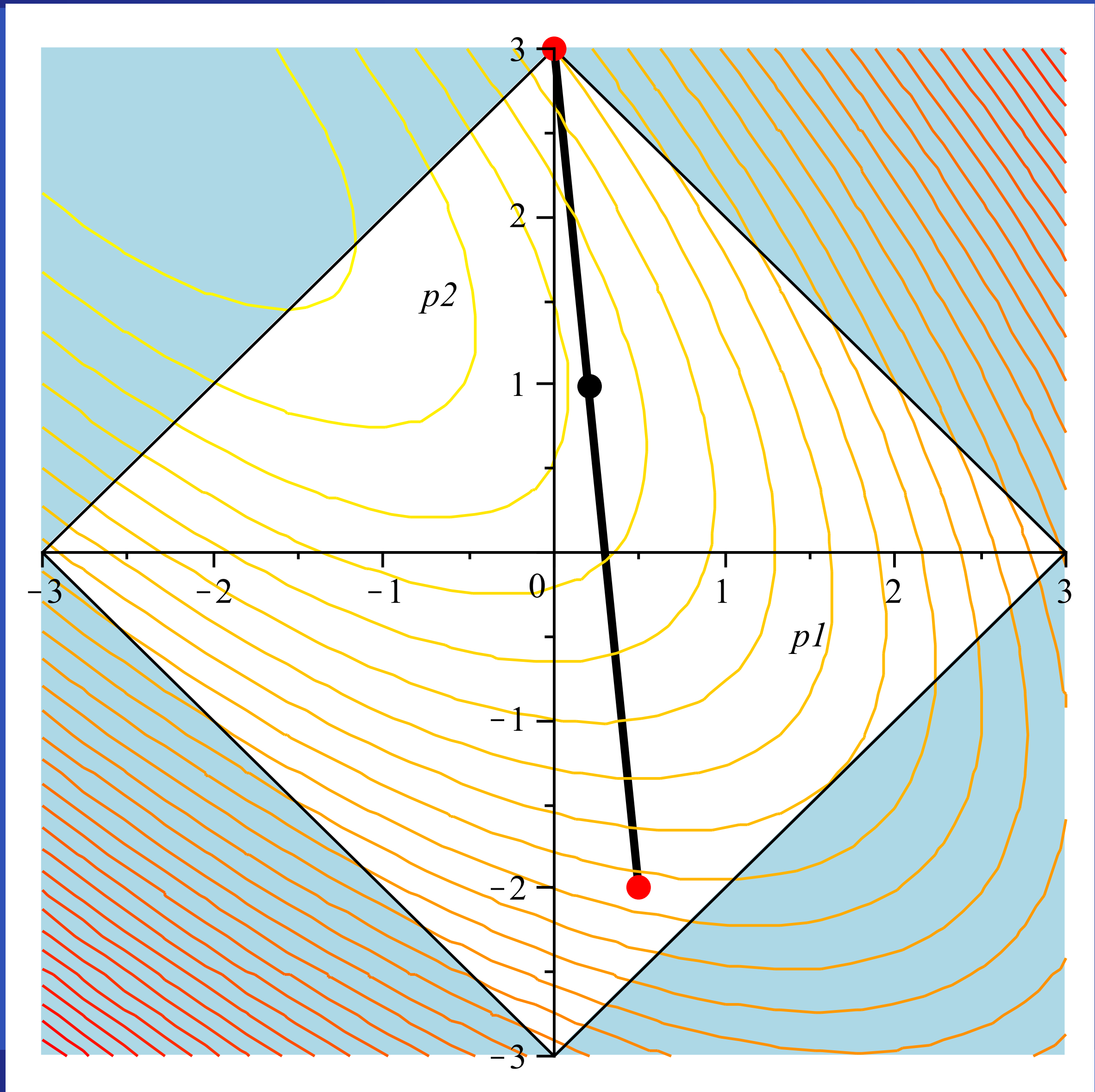
Simplicial decomposition



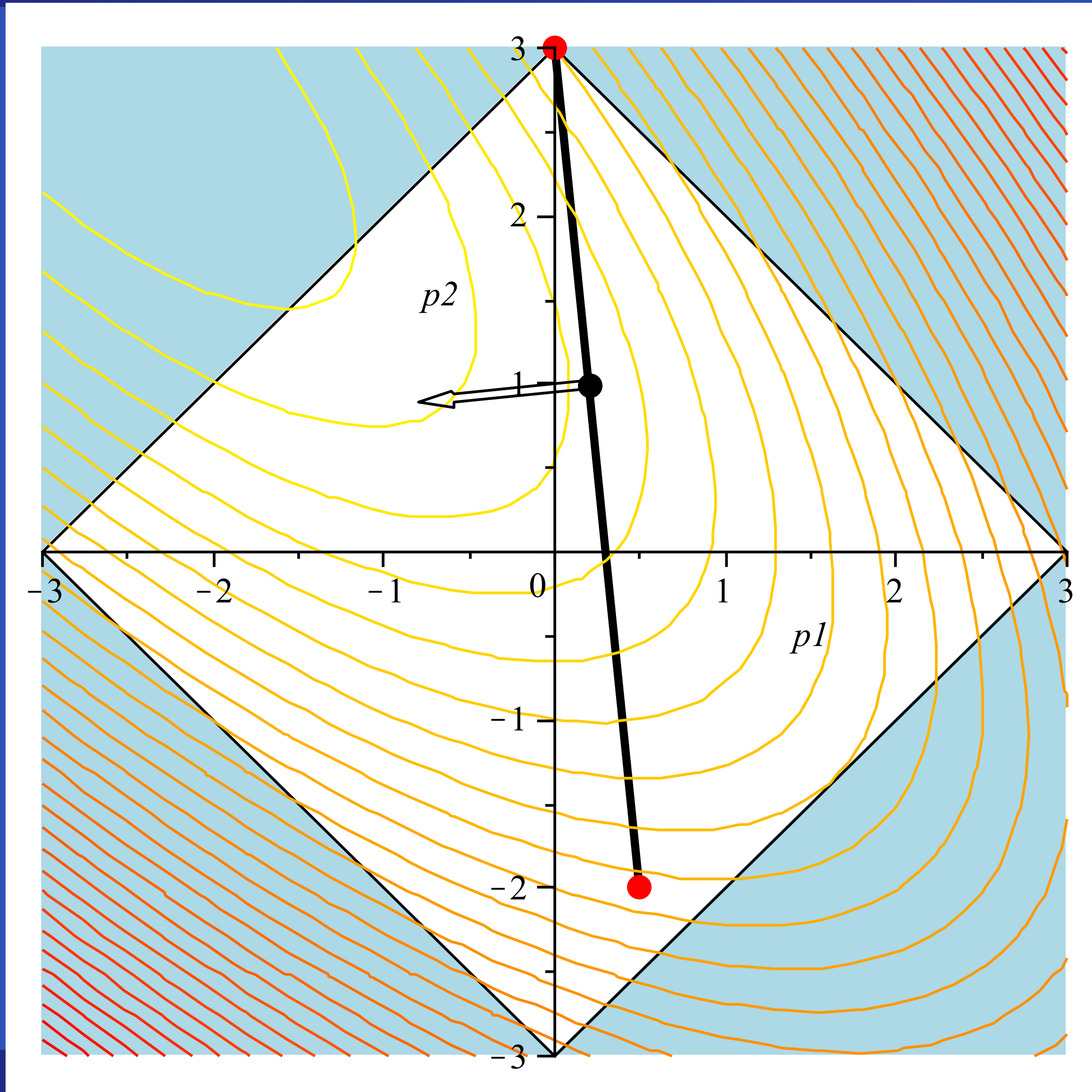
Simplicial decomposition



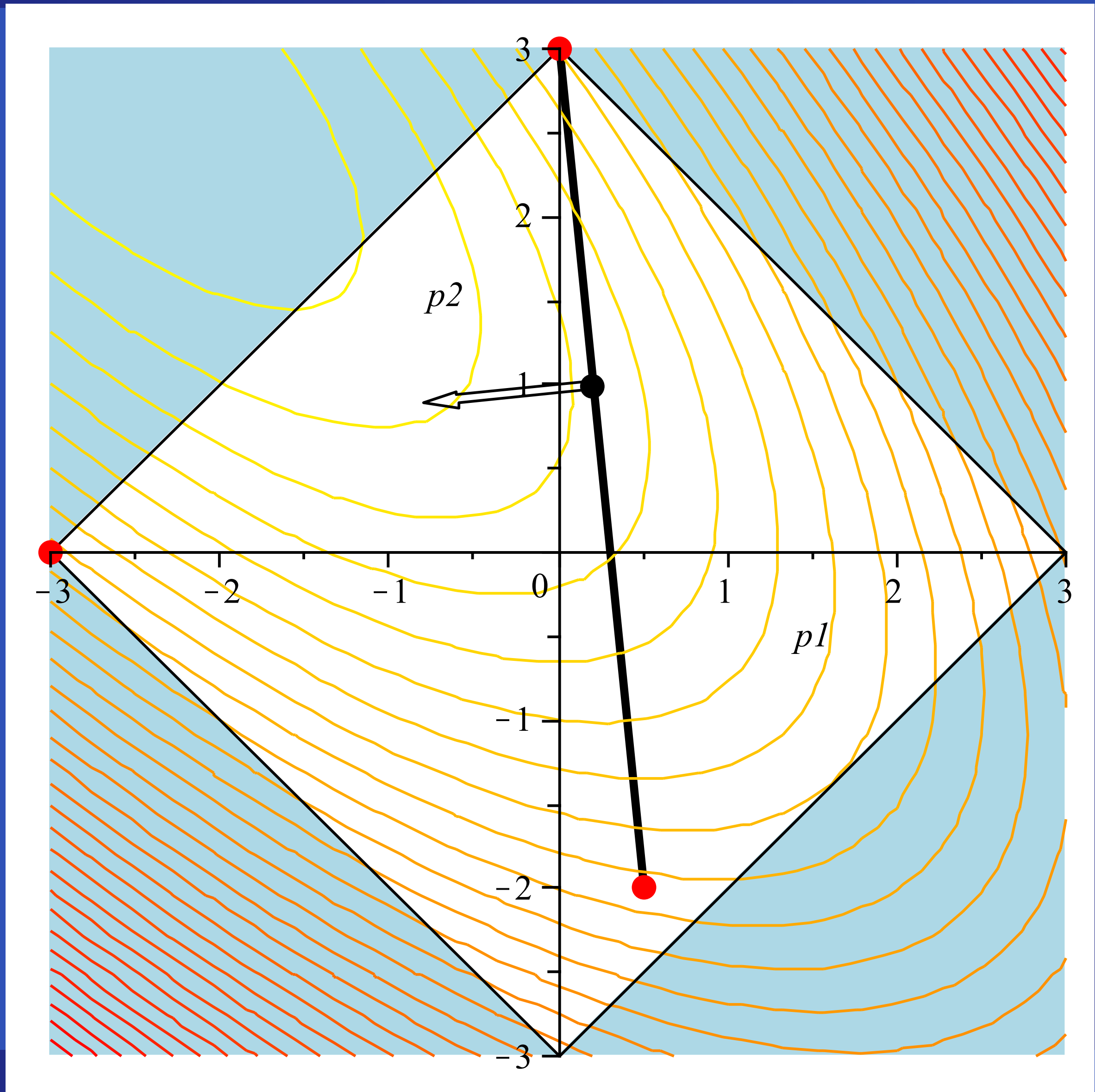
Simplicial decomposition



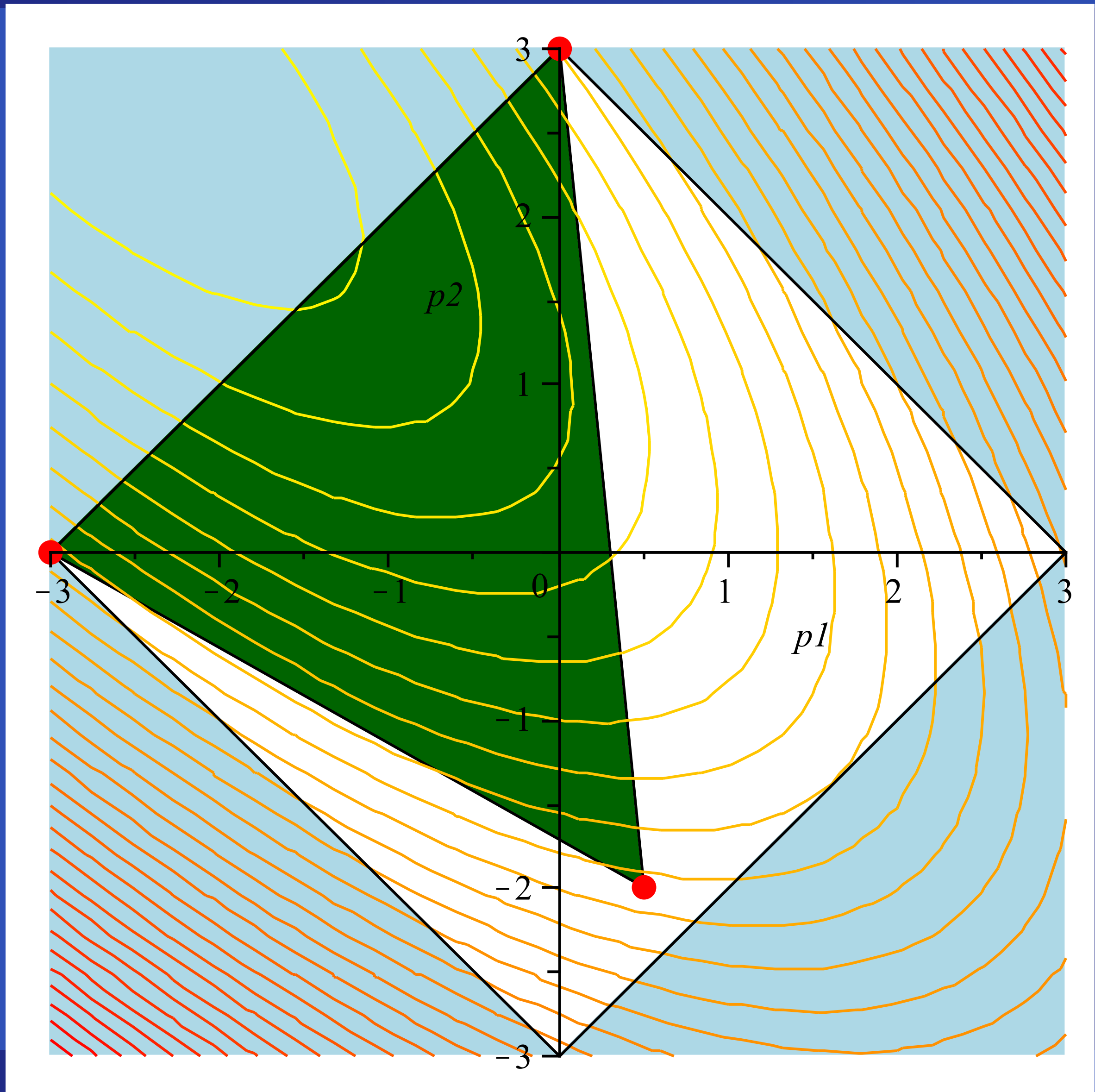
Simplicial decomposition



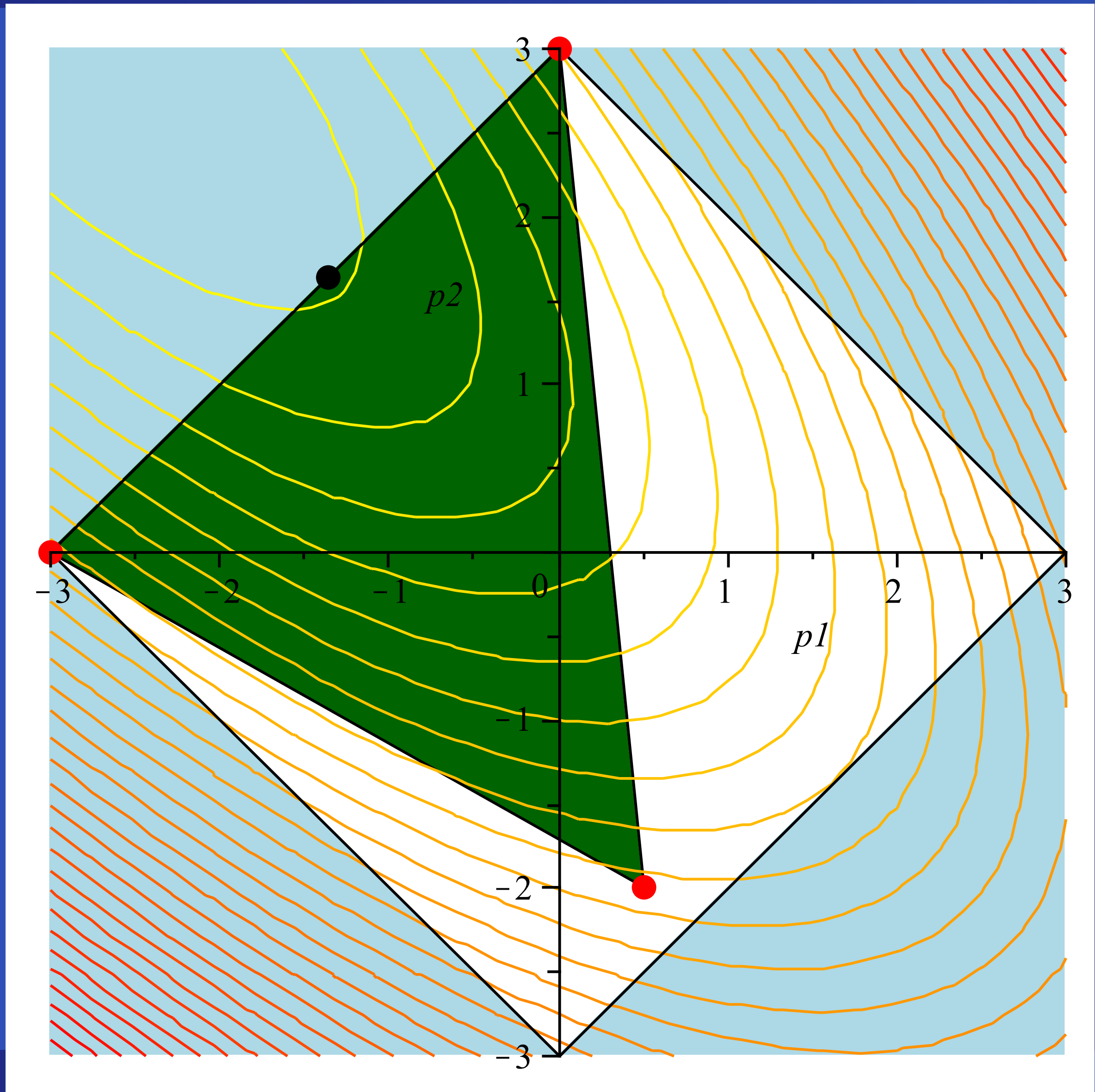
Simplicial decomposition



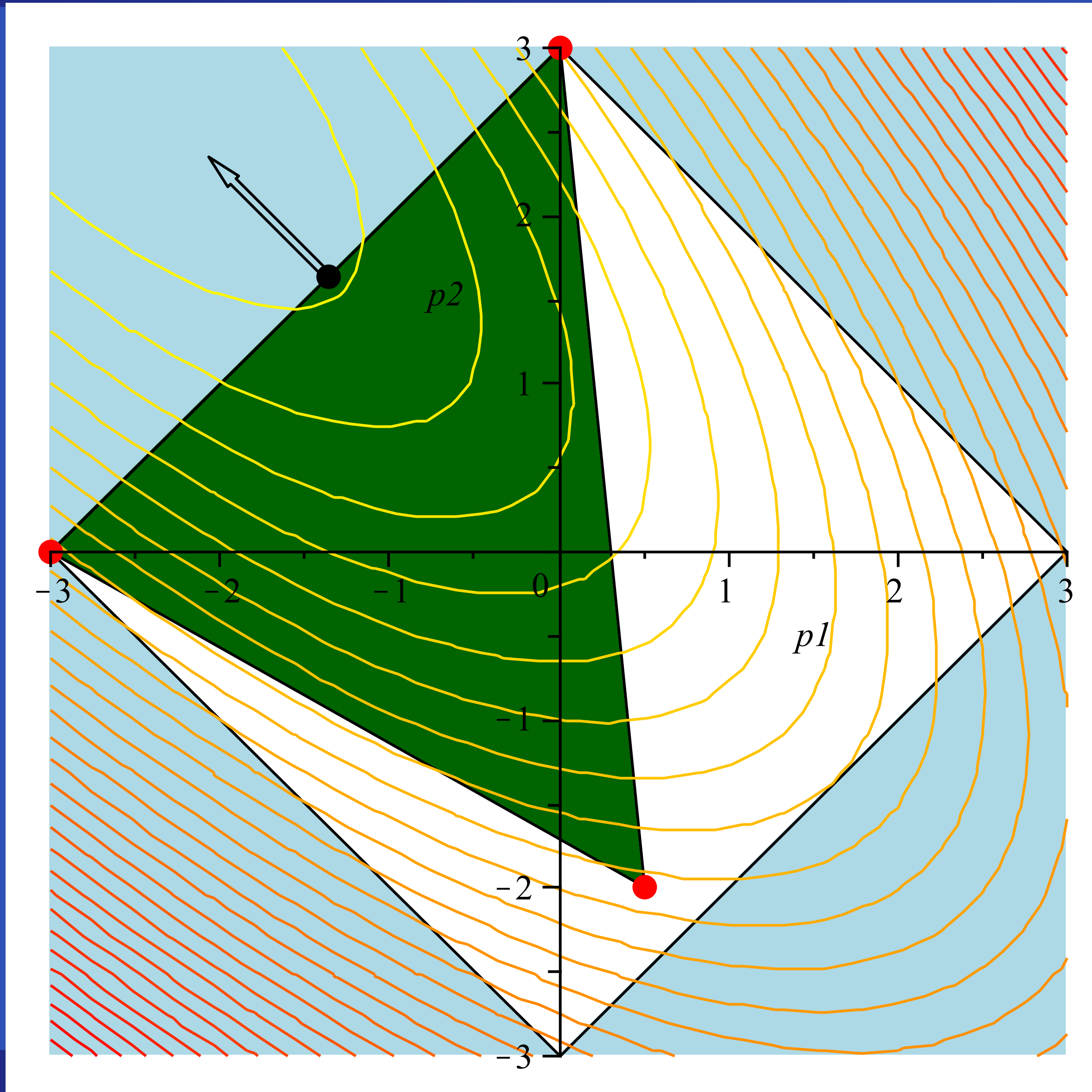
Simplicial decomposition



Simplicial decomposition



Simplicial decomposition



Simplicial decomposition

Step 0: (Initialization)

Set

$$\boldsymbol{w}^{(0)} = (\underbrace{r_1/q_1, \dots, r_1/q_1}_{q_1 \text{ times}}, \dots, \underbrace{r_L/q_L, \dots, r_L/q_L}_{q_L \text{ times}}).$$

and $Z^{(0)} = \{\boldsymbol{w}^{(0)}\}$. Select $0 < \epsilon \ll 1$, a parameter used in the stopping rule, and set $\tau = 0$.

Simplicial decomposition

Step 1: (Column generation subproblem)
Determine

$$z = \arg\max_{\boldsymbol{w} \in W} \nabla Q(\boldsymbol{w}^{(\tau)})^{\mathsf{T}} (\boldsymbol{w} - \boldsymbol{w}^{(\tau)}).$$

Step 2: (Termination check)
If $\nabla Q(\boldsymbol{w}^{(\tau)})^{\mathsf{T}} (z - \boldsymbol{w}^{(\tau)}) \leq \epsilon$, then STOP and $\boldsymbol{w}^{(\tau)}$ is optimal. Otherwise, set $Z^{(\tau+1)} = Z^{(\tau)} \cup \{z\}$.

Simplicial decomposition

Step 3: (Restricted master problem)
Find

$$\boldsymbol{w}^{(\tau+1)} = \arg \max_{\boldsymbol{w} \in \mathrm{co}(Z^{(\tau+1)})} \mathcal{Q}(\boldsymbol{w})$$

and purge $Z^{(\tau+1)}$ of all extreme points with zero weight in the expression of $\boldsymbol{w}^{(\tau+1)}$ as a convex combination of elements in $Z^{(\tau+1)}$. Increment τ by one and go back to Step 1.

Restricted master problem

Find $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_s)$ s.t.

$$\mathcal{T}(\boldsymbol{\alpha}) = \log\det\left(\sum_{\ell=1}^{s} \alpha_\ell \boldsymbol{Q}_\ell\right) \to \max$$

$$\sum_{\ell=1}^{s} \alpha_\ell = 1$$

$$\alpha_\ell \geq 0, \quad \ell = 1, \dots, s$$

Computer example

Consider the heat equation

$$\frac{\partial y(x,t)}{\partial t} = \frac{\partial}{\partial x_1}\left(\kappa(x)\frac{\partial y(x,t)}{\partial x_1}\right) + \frac{\partial}{\partial x_2}\left(\kappa(x)\frac{\partial y(x,t)}{\partial x_2}\right) + 20\exp\bigl(-50(x_1 - t)^2\bigr), \quad (x,t) \in (0,1)^3,$$

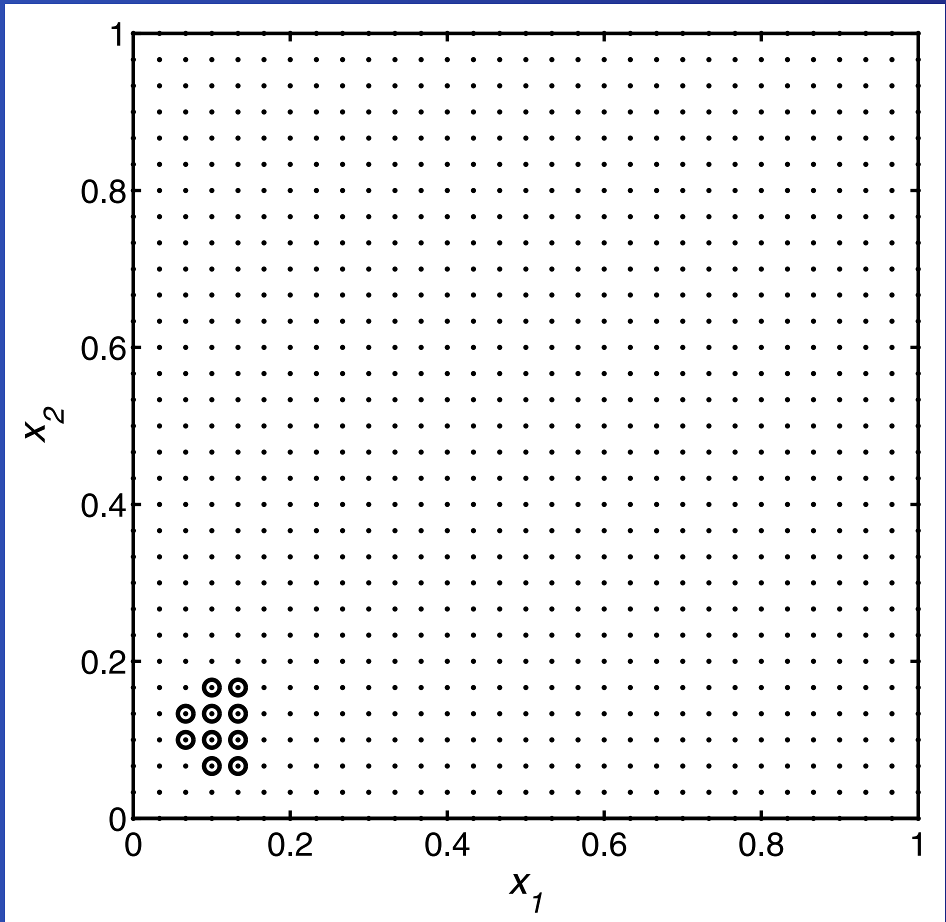
$$y(x,0) = 0, \quad x \in \Omega$$

$$y(x,t) = 0, \quad (x,t) \in \partial\Omega \times T$$

where $\kappa(x) = \theta_1 + \theta_2 x_1 + \theta_3 x_2$,
$\theta_1 = 0.1,\ \theta_2 = -0.05,\ \theta_3 = 0.2$

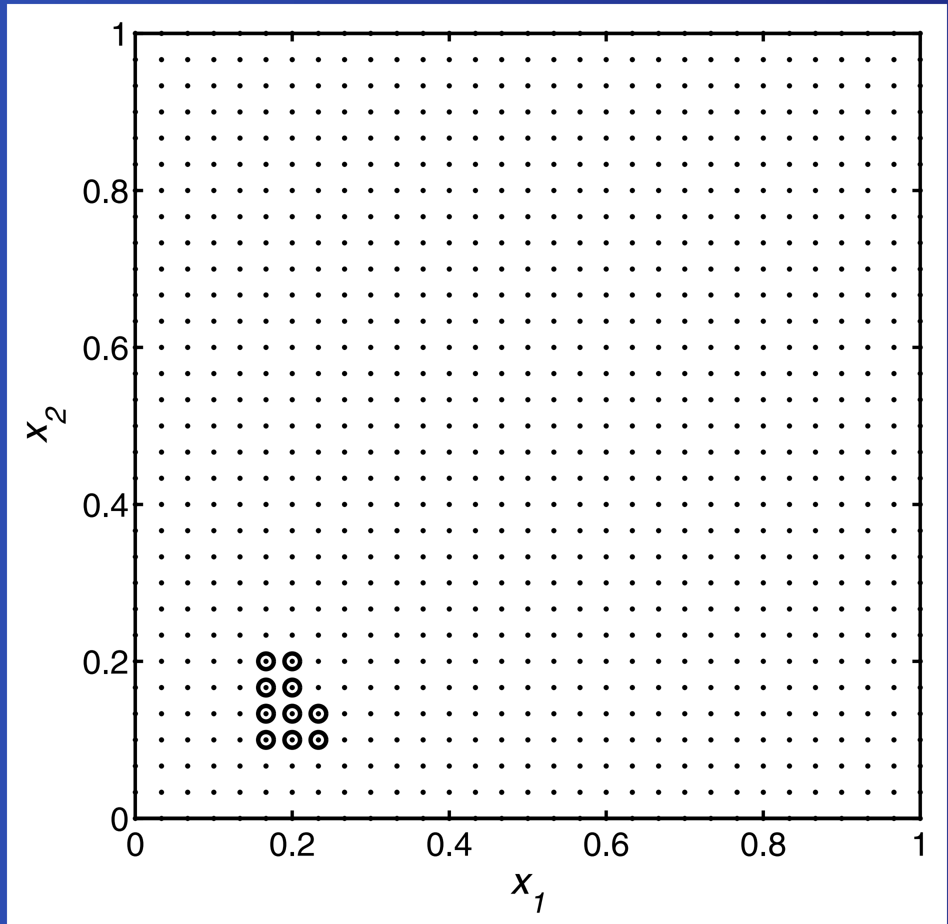
Typical behaviour

10 active sensors out of 900, Stage 1 of 6



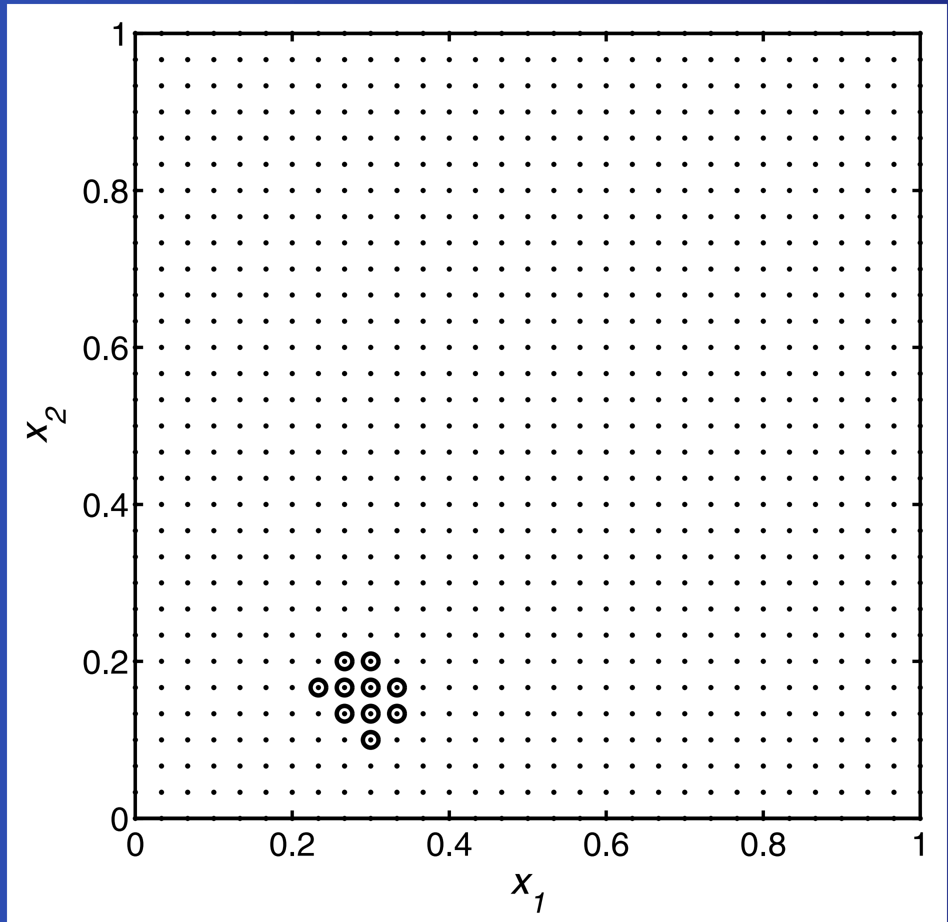
Typical behaviour

10 active sensors out of 900, Stage 2 of 6



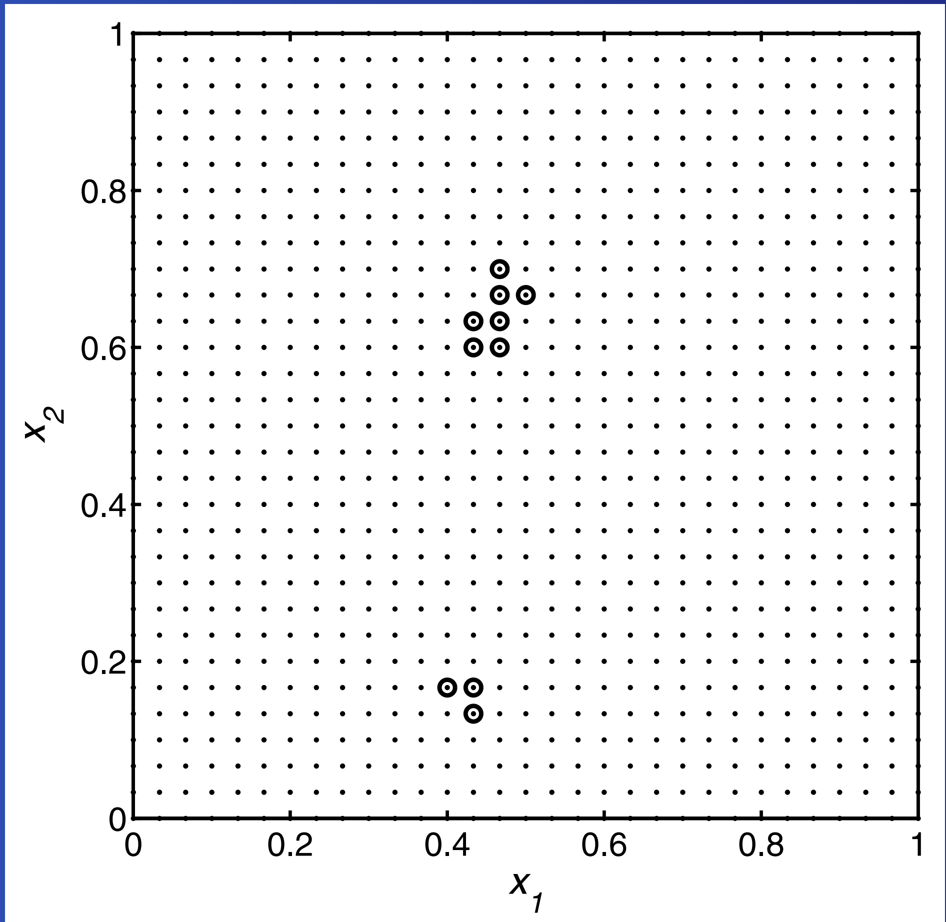
Typical behaviour

10 active sensors out of 900, Stage 3 of 6



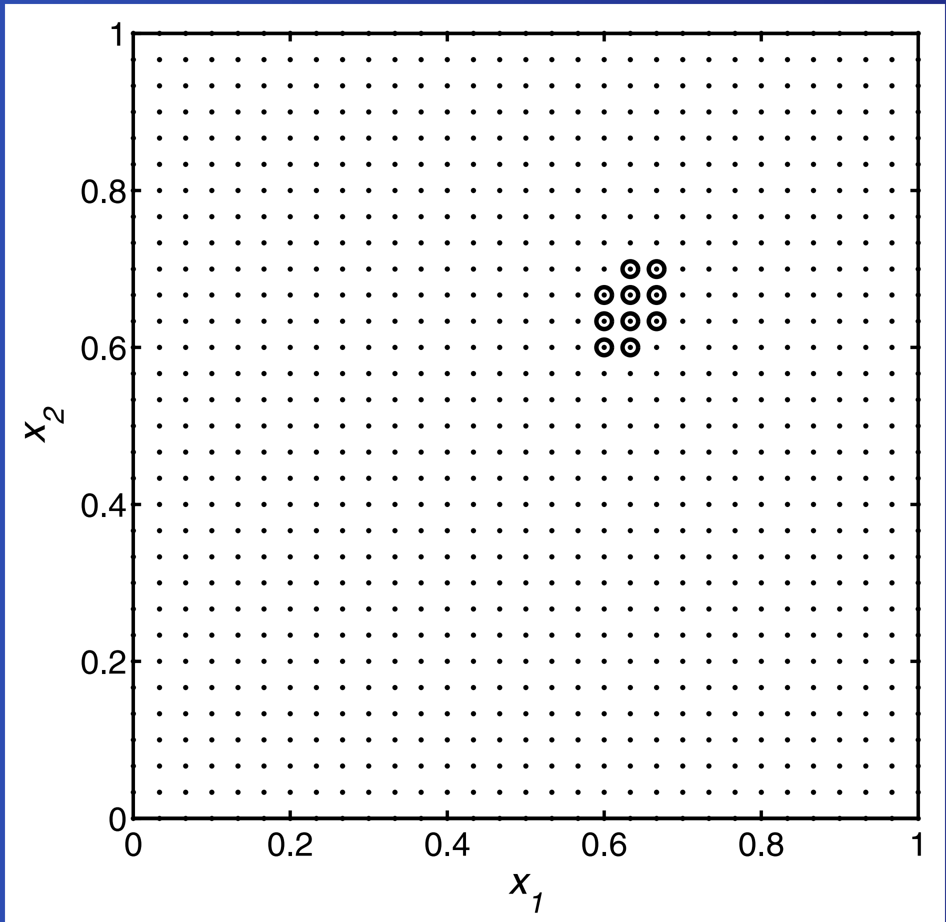
Typical behaviour

10 active sensors out of 900, Stage 4 of 6



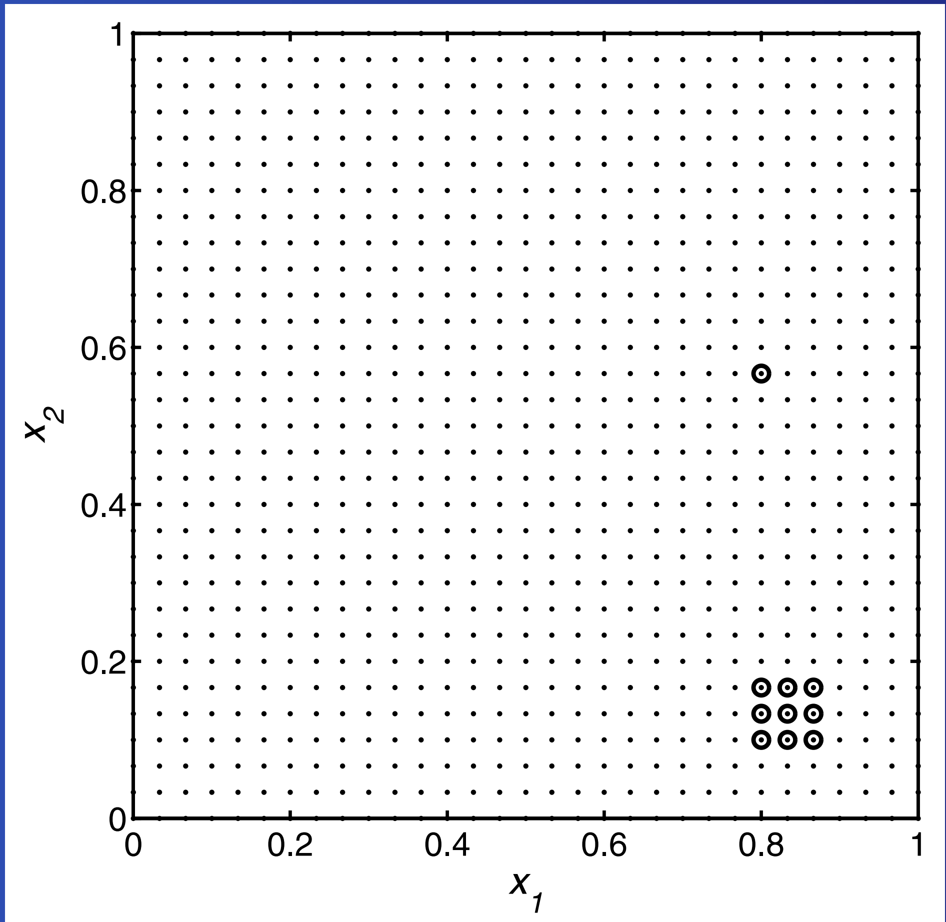
Typical behaviour

10 active sensors out of 900, Stage 5 of 6



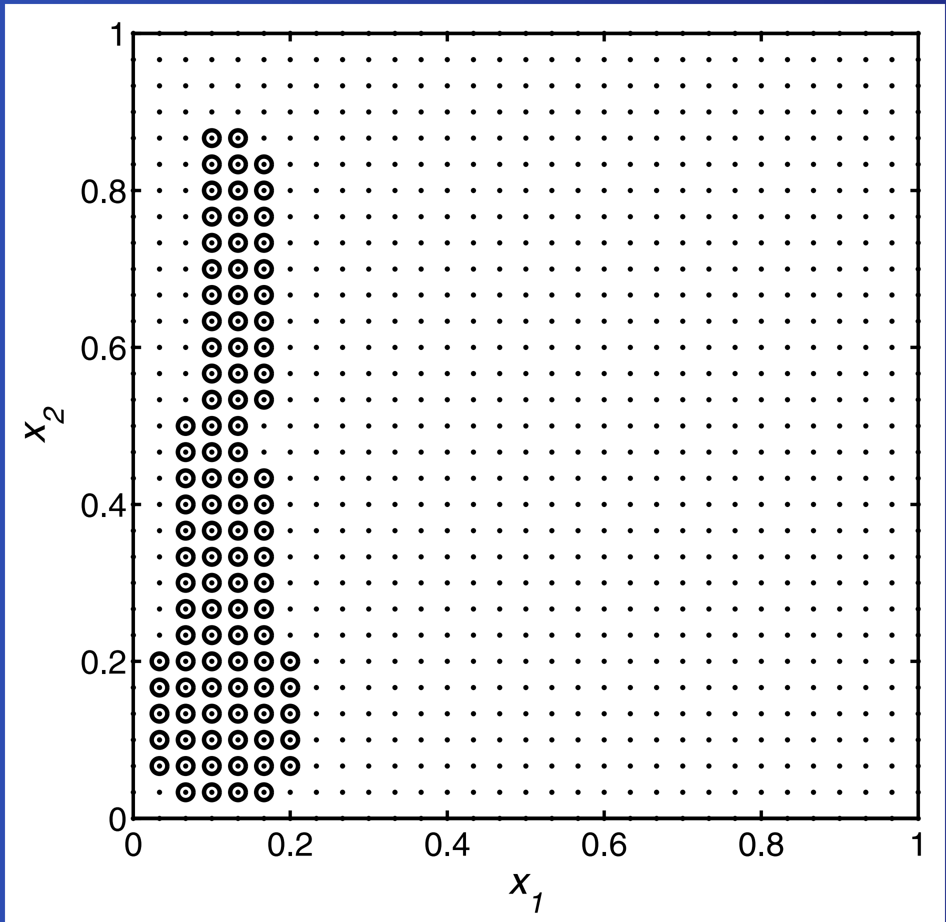
Typical behaviour

10 active sensors out of 900, Stage 6 of 6



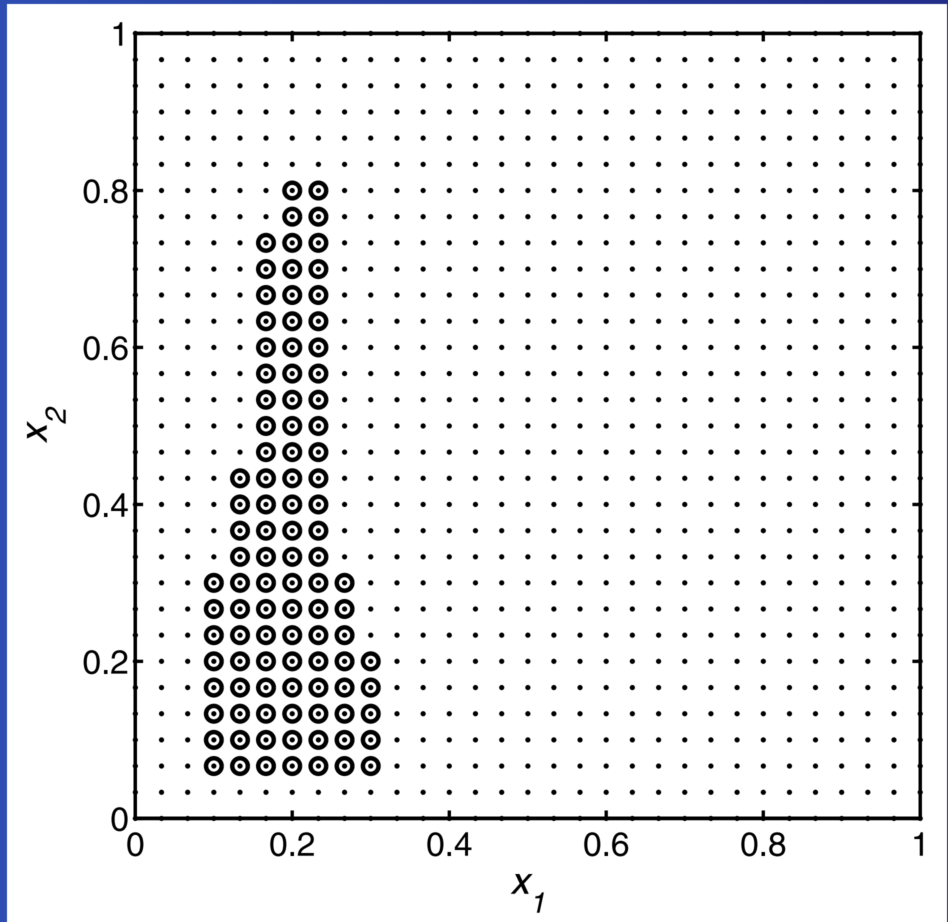
Typical behaviour

100 active sensors out of 900, Stage 1 of 6



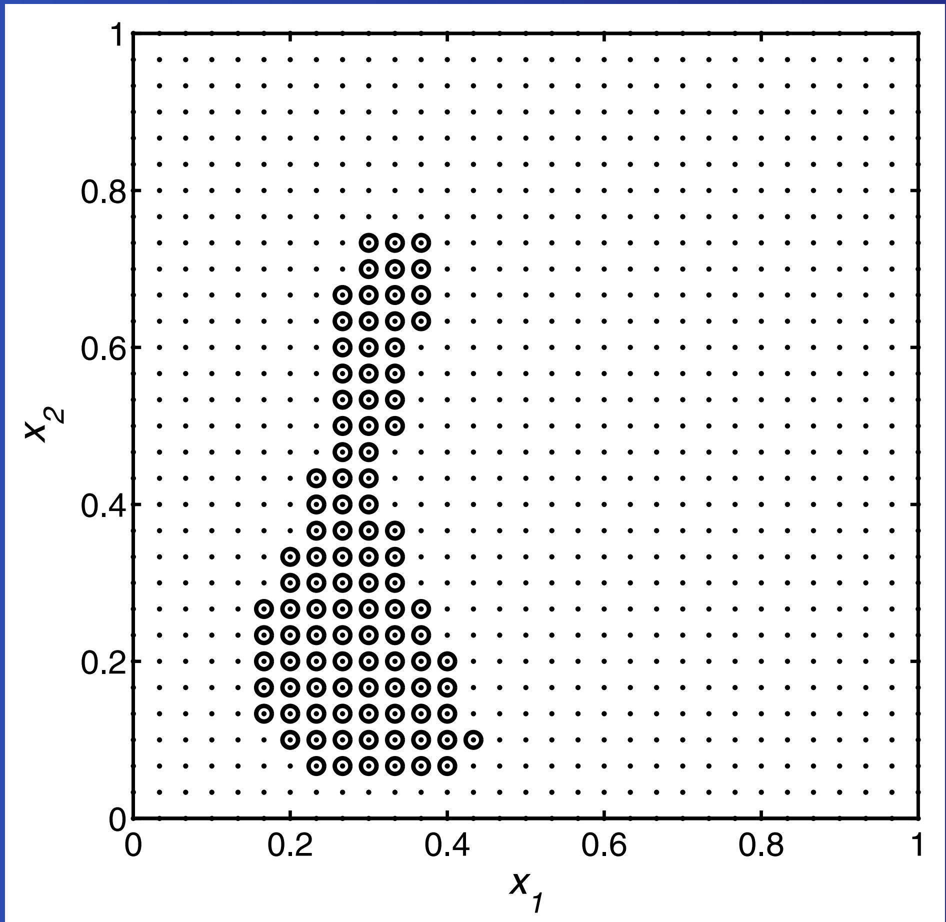
Typical behaviour

100 active sensors out of 900, Stage 2 of 6



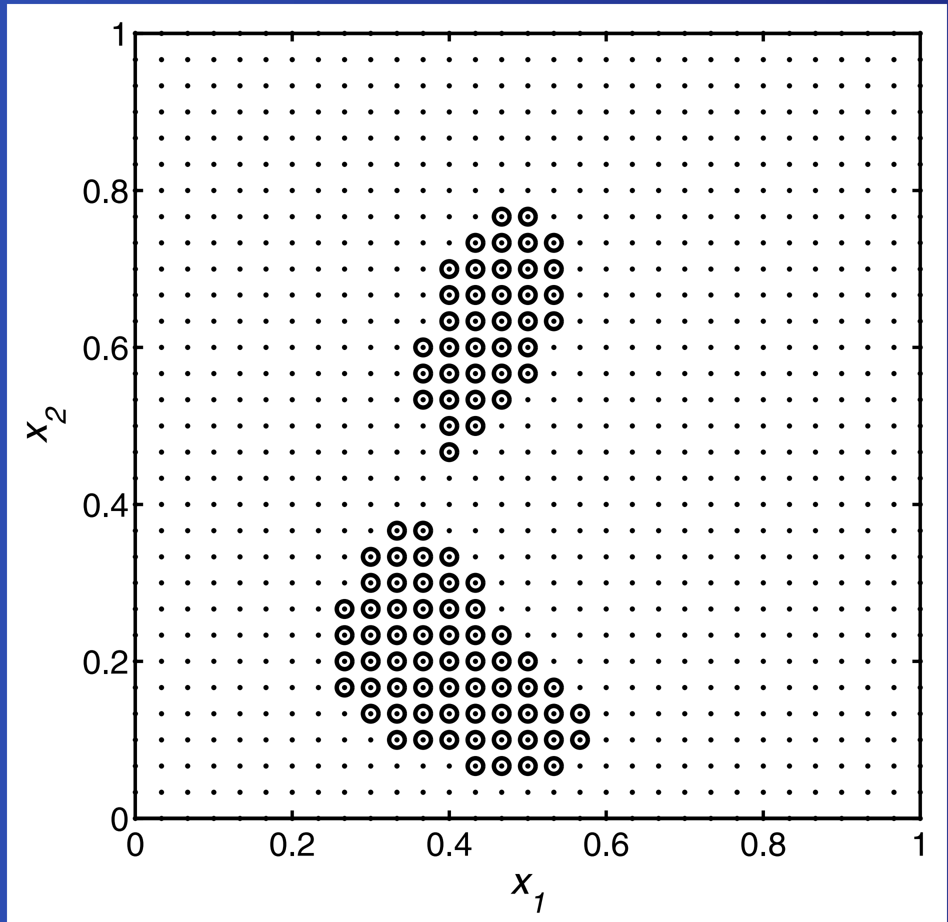
Typical behaviour

100 active sensors out of 900, Stage 3 of 6



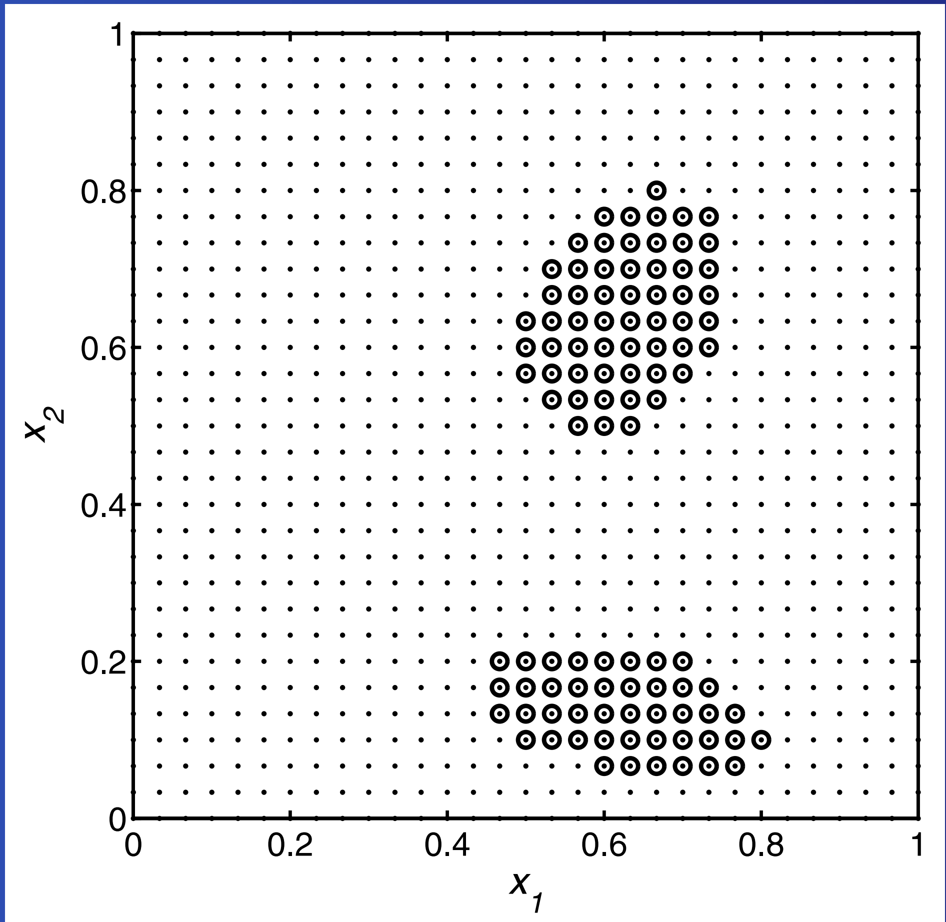
Typical behaviour

100 active sensors out of 900, Stage 4 of 6



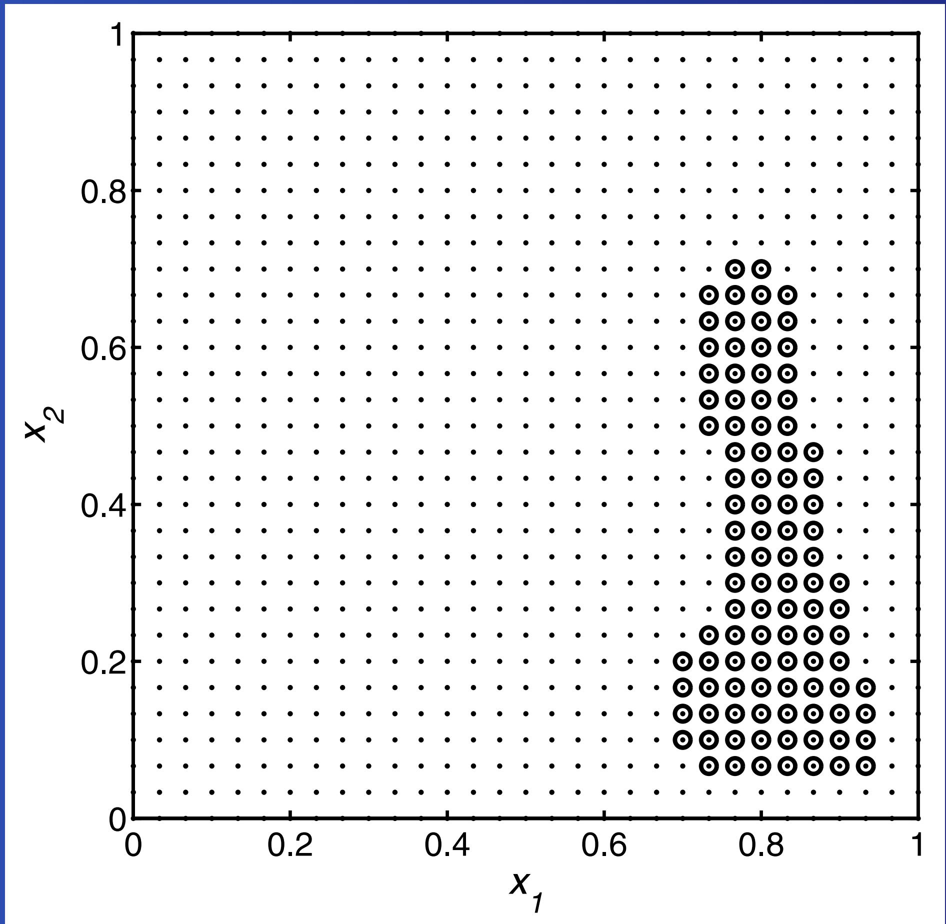
Typical behaviour

100 active sensors out of 900, Stage 5 of 6



Typical behaviour

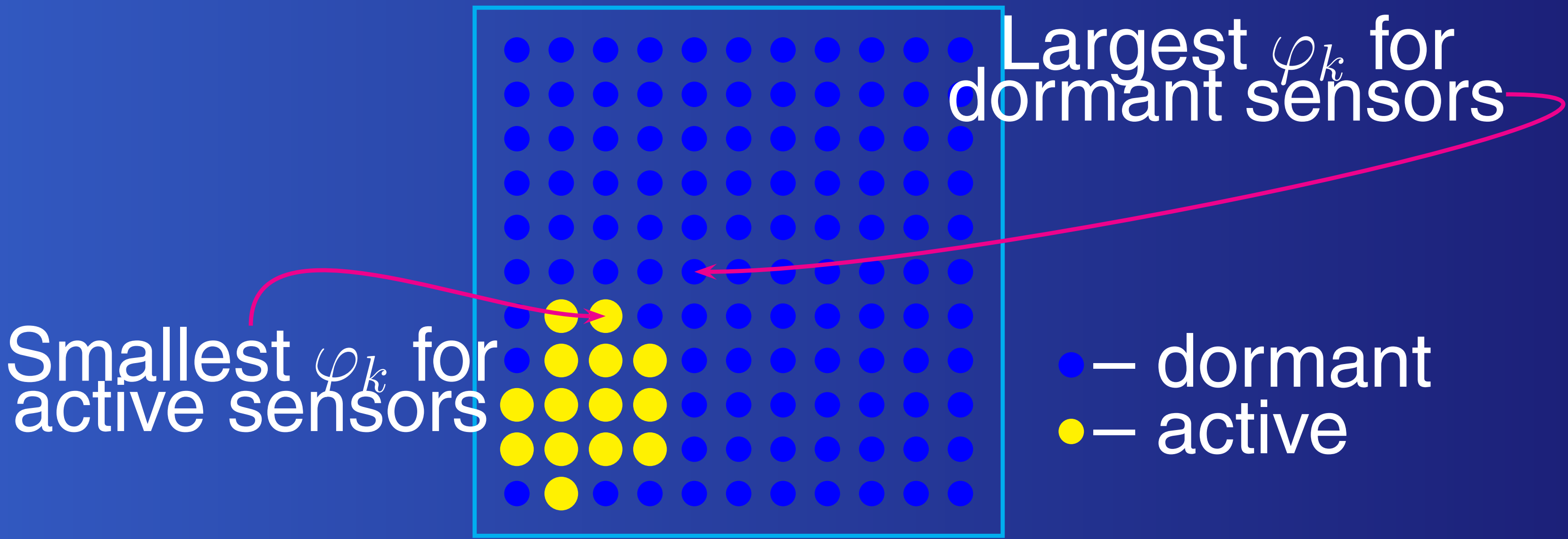
100 active sensors out of 900, Stage 6 of 6



Exchange algorithm as an alternative

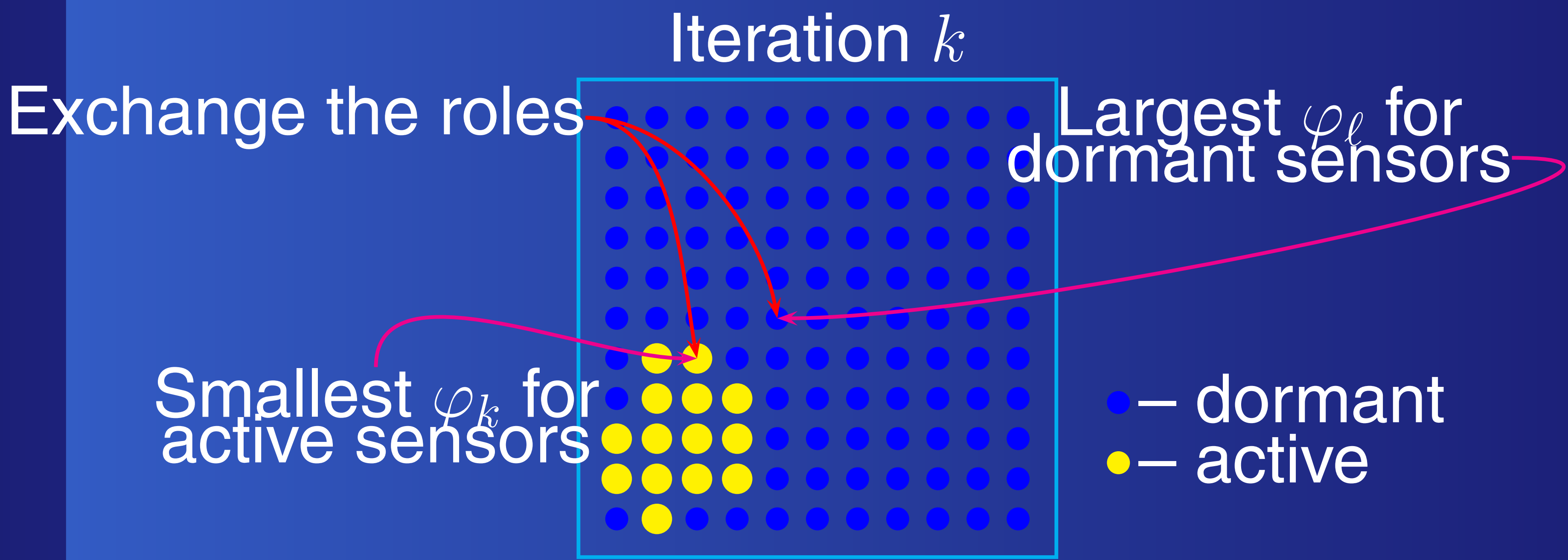
The measurement points on T_k should thus coincide with maximum points of $\varphi_k(\,\cdot\,, \xi^\star)$. This forms a basis for a numerical algorithm of exchange type:

Iteration k



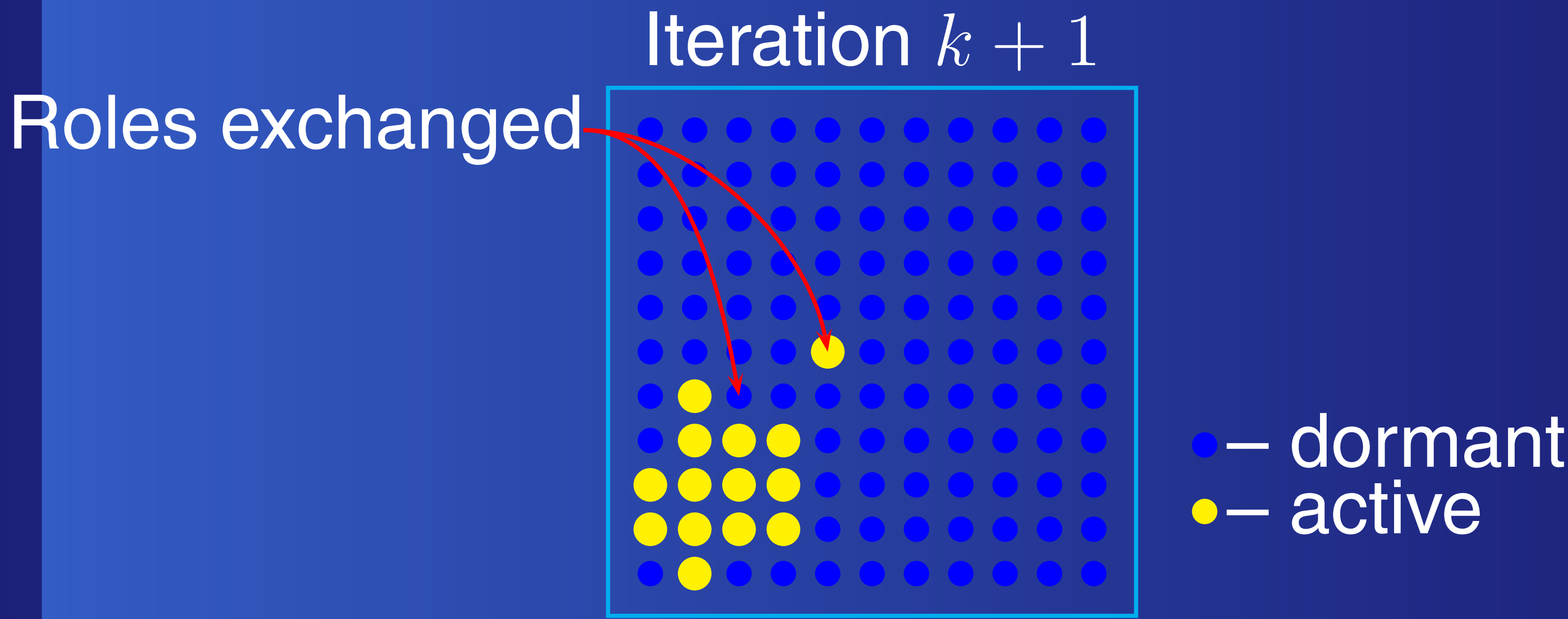
Exchange algorithm as an alternative

The measurement points on T_k should thus coincide with maximum points of $\varphi_k(\,\cdot\,, \xi^\star)$. This forms a basis for a numerical algorithm of exchange type:



Exchange algorithm as an alternative

The measurement points on T_k should thus coincide with maximum points of $\varphi_k(\,\cdot\,, \xi^\star)$. This forms a basis for a numerical algorithm of exchange type:



Conclusions

Summary of the contributions:

- Reduces the problem to a guided branch-and-bound algorithm which can be easily implemented using existing components.

Conclusions

Summary of the contributions:

- Reduces the problem to a guided branch-and-bound algorithm which can be easily implemented using existing components.
- Provides characterizations of optimal activations strategies for scanning sensors.

Conclusions

Summary of the contributions:

- Reduces the problem to a guided branch-and-bound algorithm which can be easily implemented using existing components.
- Provides characterizations of optimal activations strategies for scanning sensors.
- Works well for both high and low numbers of sensors.

Conclusions

Summary of the contributions:

- Reduces the problem to a guided branch-and-bound algorithm which can be easily implemented using existing components.
- Provides characterizations of optimal activations strategies for scanning sensors.
- Works well for both high and low numbers of sensors.
- Always produces a global optimum.

For the interested audience

Details beyond the talk are described in the book

D. Uciński (2005): *Optimal Measurement Methods for Distributed–Parameter System Identification*. — Boca Raton, FL: CRC Press, 392 p., 52 illus.

For the interested audience

Details beyond the talk are described in the book

D. Uciński (2005): *Optimal Measurement Methods for Distributed–Parameter System Identification*. — Boca Raton, FL: CRC Press, 392 p., 52 illus.

Thank you!