

Optimal experimental design for spatiotemporal system identification (Part 2)

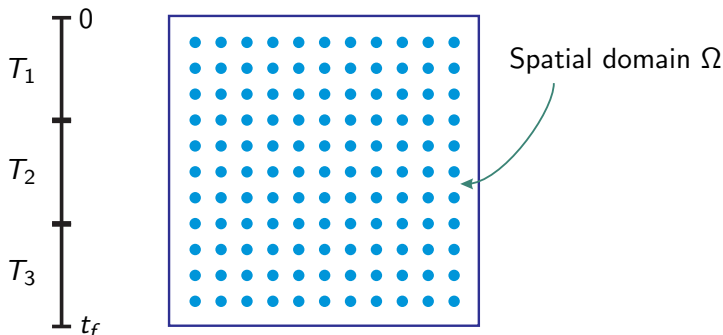
Dariusz Uciński

Institute of Control and Computation Engineering
University of Zielona Góra

minicourse at the 9th German-Polish Conference on
Optimization: Methods and Applications
Białystok, 19 September 2026

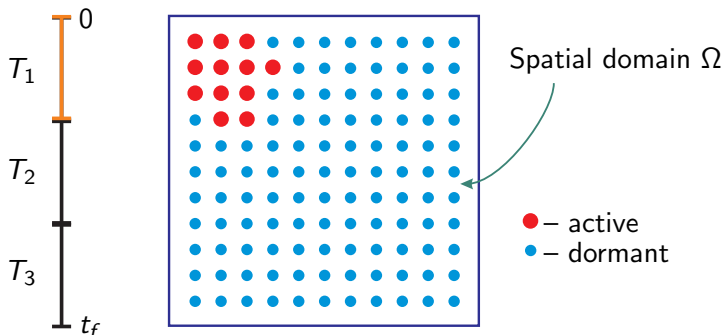
Scanning observations

Given N' sensors whose spatial positions are fixed *a priori*, partition the observation horizon $T = (0, t_f)$ into adjacent subintervals T_k , $k = 1, \dots, K$ and activate a best N -element subset of sensors over each T_k .



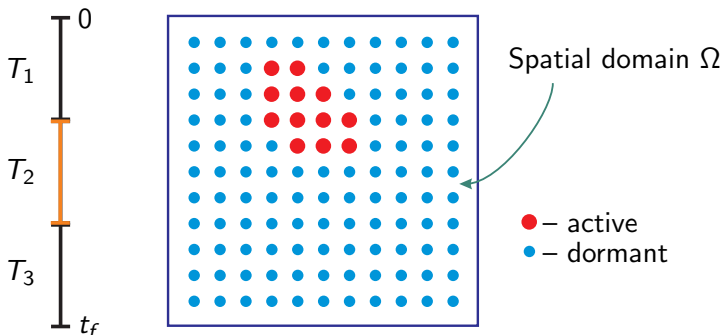
Scanning observations

Given N' sensors whose spatial positions are fixed *a priori*, partition the observation horizon $T = (0, t_f)$ into adjacent subintervals T_k , $k = 1, \dots, K$ and activate a best N -element subset of sensors over each T_k .



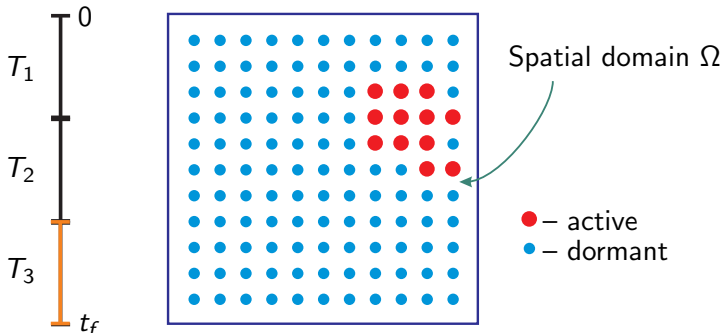
Scanning observations

Given N' sensors whose spatial positions are fixed *a priori*, partition the observation horizon $T = (0, t_f)$ into adjacent subintervals T_k , $k = 1, \dots, K$ and activate a best N -element subset of sensors over each T_k .



Scanning observations

Given N' sensors whose spatial positions are fixed *a priori*, partition the observation horizon $T = (0, t_f)$ into adjacent subintervals T_k , $k = 1, \dots, K$ and activate a best N -element subset of sensors over each T_k .



Advantages and drawbacks

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

- ▶ Another interpretation: several mobile sensors.
- ▶ Sensors track points providing best information about the DPS.
- ▶ Technology makes it quite natural.

Advantages and drawbacks

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

- ▶ Another interpretation: several mobile sensors.
- ▶ Sensors track points providing best information about the DPS.
- ▶ Technology makes it quite natural.

But:

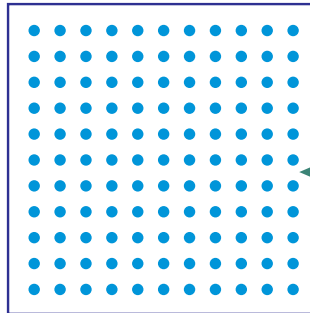
- ▶ Lack of efficient algorithms for parameter estimation.
- ▶ Combinatorial complexity excludes direct (naive) approaches.

Start with a simpler case: Stationary network

Given N' sensors whose spatial positions are fixed *a priori*, activate a best N -element subset of sensors over the observation horizon T .



Spatial domain $\Omega \cup \partial\Omega$



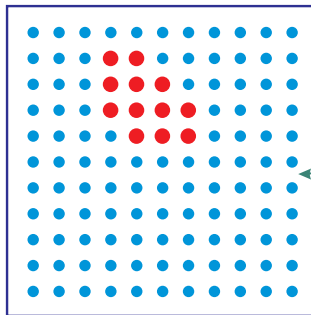
N' nodes

Start with a simpler case: Stationary network

Given N' sensors whose spatial positions are fixed *a priori*, activate a best N -element subset of sensors over the observation horizon T .



Spatial domain $\Omega \cup \partial\Omega$



N' nodes

- – active
- – dormant

Which N -element subset of nodes to activate?

System description

Consider a spatial domain $\Omega \subset \mathbb{R}^2$ and the PDE

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

subject to the appropriate initial and boundary conditions.

Notation:

- ▶ x – spatial variable, t – time, $T = [0, t_f]$;
- ▶ $y = y(x, t)$ – state vector;
- ▶ $\mathcal{F}$ – given operator including spatial derivatives;
- ▶ $\boldsymbol{\theta} \in \mathbb{R}^m$ – vector of *unknown* (constant) parameters.

Example: Air pollution

Consider domain $\Omega \subset \mathbb{R}^2$ with sufficiently smooth boundary Γ .

Advection-diffusion equation

$$\frac{\partial y}{\partial t}(\mathbf{x}, t) = \nabla \cdot [\kappa(\mathbf{x}, \theta) \nabla y(\mathbf{x}, t)] + \nu(\mathbf{x}) \cdot \nabla y(\mathbf{x}, t) + f(\mathbf{x}, t, \theta)$$
$$y(\mathbf{x}, 0) = y_0(\mathbf{x}), \quad y(\mathbf{x}, t) \Big|_{\partial\Omega} = 0$$

Notation:

- $y(\mathbf{x}, t)$ – contaminant concentrat. at point $\mathbf{x}$ and time t ,
- $\kappa(\mathbf{x}, \theta)$ – diffusivity coefficient, parameterized by θ ,
- $f(\mathbf{x}, t, \theta)$ – source term
- $\nu(\mathbf{x}) = (\nu_1(\mathbf{x}), \nu_2(\mathbf{x}))$ – wind velocity field

Output equation

At a gauged site we have

$$z^i(t) = y(\mathbf{x}^i, t; \boldsymbol{\theta}) + \varepsilon(\mathbf{x}^i, t), \quad t \in \mathcal{T}$$

where $\varepsilon(\cdot)$ – white Gaussian measurement noise.

Output equation

At a gauged site we have

$$z^i(t) = y(\mathbf{x}^i, t; \boldsymbol{\theta}) + \varepsilon(\mathbf{x}^i, t), \quad t \in T$$

where $\varepsilon(\cdot)$ – white Gaussian measurement noise.

Least-squares criterion

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

$$\mathcal{J}(\boldsymbol{\theta}) = \frac{1}{2} \sum_{i=1}^{N'} v_i \int_T [z_i(t) - \hat{y}(\mathbf{x}^i, t; \boldsymbol{\theta})]^2 dt$$

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

$$v_i = \begin{cases} 1 & \text{if sensor } i \text{ is active} \\ 0 & \text{if sensor } i \text{ is dormant} \end{cases}$$

Optimal measurement problem

Determine $v_1, \dots, v_{N'}$ (binary flags representing sensor activation) so as to maximize the identification accuracy

Optimal measurement problem

Determine $v_1, \dots, v_{N'}$ (binary flags representing sensor activation) so as to maximize the **identification accuracy**. ←
What is it?

Optimal measurement problem

Determine $v_1, \dots, v_{N'}$ (binary flags representing sensor activation) so as to maximize the **identification accuracy**. ←
What is it?

Cramér-Rao inequality

$$\text{cov } \hat{\theta} = E\left\{(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T\right\} \succeq M^{-1}$$

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

Fisher Information Matrix (FIM)

Up to a constant multiplier, we have

$$M = \sum_{i=1}^{N'} v_i \mathbf{r}_i, \quad \mathbf{r}_i = \int_0^{t_f} \mathbf{g}(\mathbf{x}^i, t) \mathbf{g}^T(\mathbf{x}^i, t) dt$$

where $\mathbf{g}(\mathbf{x}^i, t) = \left(\frac{\partial y(\mathbf{x}^j, t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}^T$ is the vector of **sensitivity coefficients**, $\boldsymbol{\theta}^0$ stands for a preliminary estimate of $\boldsymbol{\theta}$.

Problem of best sensor activation

Find $v_1, \dots, v_{N'}$ so as to maximize the **D-optimality criterion**

$$\ln \det \left(\sum_{i=1}^{N'} v_i \mathbf{r}_i \right)$$

subject to

$$\sum_{i=1}^{N'} v_i = N$$
$$v_i \in \{0, 1\}, \quad i = 1, \dots, N'$$

Relaxed problem of best sensor activation

Find $v_1, \dots, v_{N'}$ so as to maximize the **D-optimality criterion**

$$\ln \det \left(\sum_{i=1}^{N'} v_i \mathbf{r}_i \right)$$

subject to

$$\sum_{i=1}^{N'} v_i = N$$

$$v_i \in [0, 1], \quad i = 1, \dots, N'$$

Problem simplification

To circumvent the curse of dimensionality for large SNs, we relax our problem: **Instead of seeking an optimal configuration of activated sensors, we are going to look for a spatial density of sensors** in X , a subdomain of $\Omega \cup \partial\Omega$ where measurements are allowed.

Problem simplification

To circumvent the curse of dimensionality for large SNs, we relax our problem: **Instead of seeking an optimal configuration of activated sensors, we are going to look for a spatial density of sensors** in X , a subdomain of $\Omega \cup \partial\Omega$ where measurements are allowed.

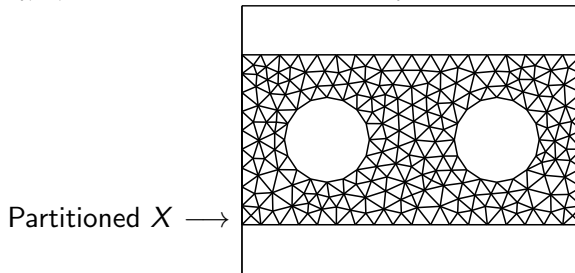
Mathematically, this density can be identified with a probability measure $\xi(dx)$ on the space $(X, \mathcal{B})$, where $\mathcal{B}$ is the σ -algebra of all Borel subsets of X .

Is it useful?

Definitely yes! Given ξ , partition X into small subdomains X_i and allocate to each of them

$$N_i = \left\lceil N \int_{X_i} \xi(d\mathbf{x}) \right\rceil$$

of sensors ($\lceil \rho \rceil$ is the smallest integer $\geq \rho$).



Necessary corrections

Thus our aim is to find a probability measure ξ over X which maximizes $\Psi[\mathbf{M}(\xi)]$ for

$$\mathbf{M}(\xi) = \int_X \mathbf{r}(\mathbf{x}) \xi(d\mathbf{x}), \quad \mathbf{r}(\mathbf{x}) = \frac{1}{t_f} \int_0^{t_f} \mathbf{g}(\mathbf{x}, t) \mathbf{g}^T(\mathbf{x}, t) dt$$

It is sensible to introduce $\omega(d\mathbf{x})$ as the maximal possible ‘number’ of sensors per $d\mathbf{x}$:

$$\xi(d\mathbf{x}) \leq \omega(d\mathbf{x}) \quad \text{s.t.} \quad \int_X \omega(d\mathbf{x}) \geq 1$$

Characterization of solutions

Theorem

If $\omega(d\mathbf{x}) = \varrho(\mathbf{x}) d\mathbf{x}$, where $\varrho(\mathbf{x})$ is a positive continuous function, then a (Ψ, ω) -optimal design ξ^* exists s.t. $\xi^*(d\mathbf{x}) = h(\mathbf{x}) d\mathbf{x}$ and

$$h(\mathbf{x}) = \begin{cases} \varrho(\mathbf{x}) & \text{on some subdomain } X^* \\ 0 & \text{on } X \setminus X^* \end{cases}$$

A necessary and sufficient condition for this optimality is

$$\min_{\mathbf{x} \in X^*} \phi(\mathbf{x}, \xi^*) \geq \max_{\mathbf{x} \in X \setminus X^*} \phi(\mathbf{x}, \xi^*)$$

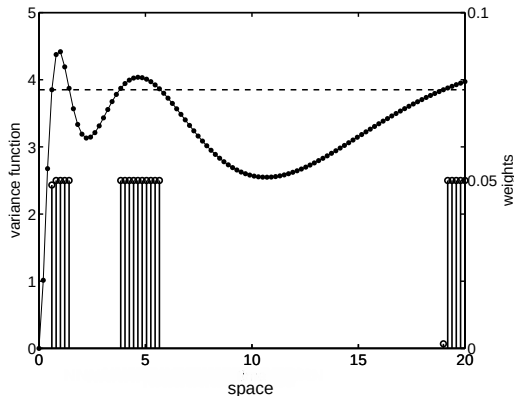
How to compute $\phi(\mathbf{x}, \xi^*)$? (cf. Part 1)

$$\phi(\mathbf{x}, \xi^*) = \frac{1}{t_f} \int_0^{t_f} \mathbf{g}^\top(\mathbf{x}, t) \mathbf{M}^{-1}(\xi^*) \mathbf{g}(\mathbf{x}, t) dt$$

How to compute $\phi(\mathbf{x}, \xi^*)$? (cf. Part 1)

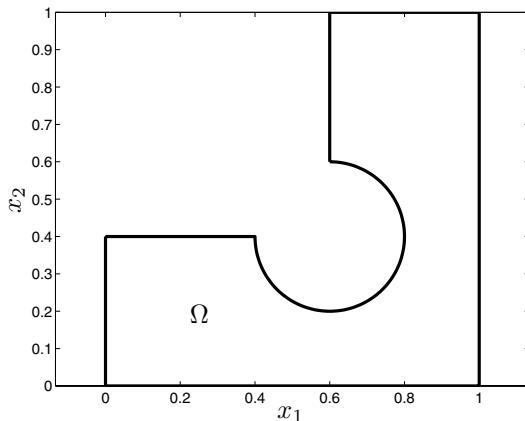
$$\phi(\mathbf{x}, \xi^*) = \frac{1}{t_f} \int_0^{t_f} \mathbf{g}^\top(\mathbf{x}, t) \mathbf{M}^{-1}(\xi^*) \mathbf{g}(\mathbf{x}, t) dt$$

Optimality conditions amount to a separability of its values:



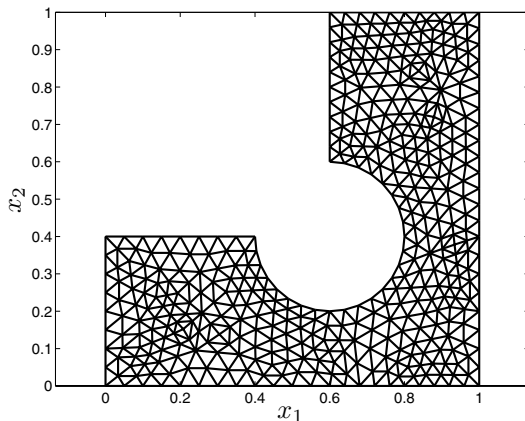
Fine partition of X

Partition X into a union of small disjoint subdomains X_i , $i = 1, \dots, n$, i.e., $X = \bigcup_{i=1}^n X_i$:



Fine partition of X

Partition X into a union of small disjoint subdomains X_i ,
 $i = 1, \dots, n$, i.e., $X = \bigcup_{i=1}^n X_i$:



Weight optimization problem

Observe that ξ assigns each subdomain X_i a weight $p_i = \xi(X_i)$. The knowledge of the p_i 's suffices to determine an optimal distribution of sensor nodes. We have

$$\mathbf{M}(\xi) \approx \mathbf{M}(\mathbf{p}) := \sum_{i=1}^n p_i \mathbf{M}_i$$

where $\mathbf{p} = [p_1, \dots, p_n]^T$, $\mathbf{M}_i = \mathfrak{T}(\mathbf{x}^i)$, $\mathbf{x}^i$ is the centre of gravity of X_i .

Ultimate formulation

Problem

Find a vector of weights $\mathbf{p}$ to maximize

$$\Phi(\mathbf{p}) = \log \det \left(\sum_{i=1}^n p_i \mathbf{M}_i \right)$$

subject to

$$\mathbf{0} \preceq \mathbf{p} \preceq \mathbf{b}$$
$$\mathbf{1}^\top \mathbf{p} = 1$$

where $\mathbf{b} = [\omega(X_1), \dots, \omega(X_n)]^\top$.

Numerically, it solved using **simplicial decomposition** combined with Torsney's multiplicative weight optimization algorithm.

Problem statement once again

Ultimate formulation

Given a vector $\mathbf{b} \succ \mathbf{0}$ satisfying

$$\mathbf{1}^\top \mathbf{b} \geq 1$$

find a vector of weights $\mathbf{p} = (p_1, \dots, p_n)$ to maximize

$$\Phi[\mathbf{M}(\mathbf{p})] = \log \det(\mathbf{M}(\mathbf{p}))$$

subject to

$$\mathbf{0} \preceq \mathbf{p} \preceq \mathbf{b}$$

$$\mathbf{1}^\top \mathbf{p} = 1$$

Properties

- 1 The performance index Φ is concave over the canonical simplex

$$S_n = \{\mathbf{p} \succeq \mathbf{0} \mid \mathbf{1}^\top \mathbf{p} = 1\}$$

Properties

- 1 The performance index Φ is concave over the canonical simplex

$$S_n = \{\mathbf{p} \succeq \mathbf{0} \mid \mathbf{1}^\top \mathbf{p} = 1\}$$

- 2 It is differentiable at points yielding nonsingular FIMs, with

$$\phi(\mathbf{p}) := \nabla \Phi(\mathbf{p}) = [\text{trace}\{\mathbf{M}(\mathbf{p})^{-1} \mathbf{M}_1\}, \dots, \text{trace}\{\mathbf{M}(\mathbf{p})^{-1} \mathbf{M}_n\}]^\top$$

Properties

- 1 The performance index Φ is concave over the canonical simplex

$$S_n = \{\mathbf{p} \succeq \mathbf{0} \mid \mathbf{1}^\top \mathbf{p} = 1\}$$

- 2 It is differentiable at points yielding nonsingular FIMs, with

$$\phi(\mathbf{p}) := \nabla \Phi(\mathbf{p}) = [\text{trace}\{\mathbf{M}(\mathbf{p})^{-1} \mathbf{M}_1\}, \dots, \text{trace}\{\mathbf{M}(\mathbf{p})^{-1} \mathbf{M}_n\}]^\top$$

- 3 The constraint set P is a rather nice convex set (e.g., fast algorithms of orthogonal projection on P exist). Numerous computational methods can potentially be employed, e.g., the conditional gradient method or a gradient projection method. But, if the number of the support points is large, they may lead to unsatisfactory long computational times.

Characterization of optimal designs

Theorem

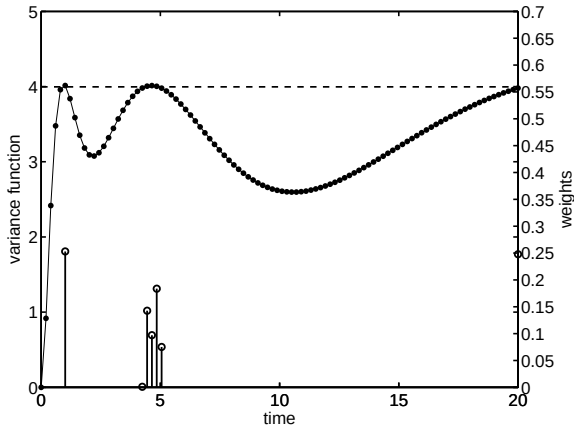
Suppose that the matrix $\mathbf{M}(\mathbf{p}^*)$ is nonsingular for some $\mathbf{p}^* \in P$. The vector $\mathbf{p}^*$ constitutes a global maximum of Φ over P if, and only if, there exists a number λ^* such that

$$\phi_i(\mathbf{p}^*) \begin{cases} \geq \lambda^* & \text{if } p_i^* = b_i \\ = \lambda^* & \text{if } 0 < p_i^* < b_i \\ \leq \lambda^* & \text{if } p_i^* = 0 \end{cases}$$

for $i = 1, \dots, n$.

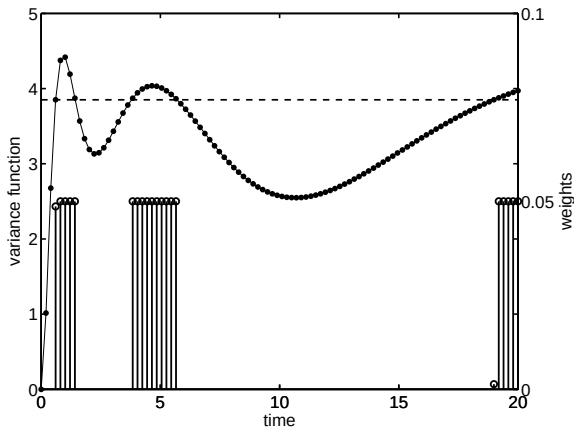
Clear interpretation

Optimal ϕ for a case with 4 estimated parameters (no upper bound constraints on the weights):



Clear interpretation

Optimal ϕ for a case with 4 estimated parameters (constraint $\mathbf{p} \preceq 0.05 \mathbf{1}$ imposed):



Simplicial decomposition

Simplicial Decomposition (SD) stands for a class of methods for solving large-scale continuous problems in mathematical programming with convex feasible sets (von Hohenbalken, 1977). It iterates by alternately solving

Simplicial decomposition

Simplicial Decomposition (SD) stands for a class of methods for solving large-scale continuous problems in mathematical programming with convex feasible sets (von Hohenbalken, 1977). It iterates by alternately solving

- 1 a linear programming subproblem (the so-called *column generation problem*) which generates an extreme point of the polyhedron, and

Simplicial decomposition

Simplicial Decomposition (SD) stands for a class of methods for solving large-scale continuous problems in mathematical programming with convex feasible sets (von Hohenbalken, 1977). It iterates by alternately solving

- 1 a linear programming subproblem (the so-called *column generation problem*) which generates an extreme point of the polyhedron, and
- 2 a nonlinear *restricted master problem* (RMP) which finds the maximum of the objective function over the convex hull (a simplex) of previously defined extreme points.

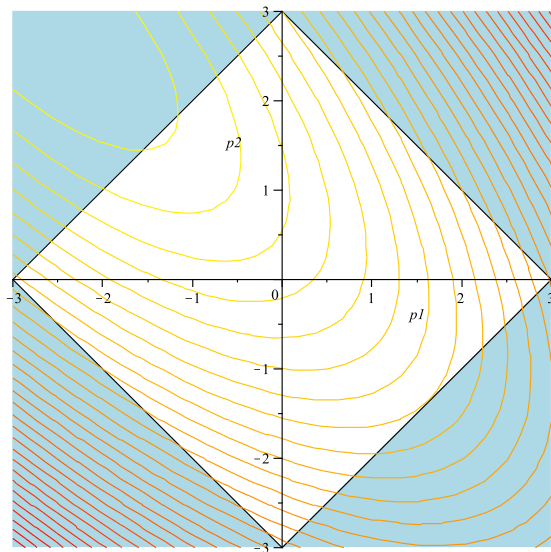
Simplicial decomposition

Simplicial Decomposition (SD) stands for a class of methods for solving large-scale continuous problems in mathematical programming with convex feasible sets (von Hohenbalken, 1977). It iterates by alternately solving

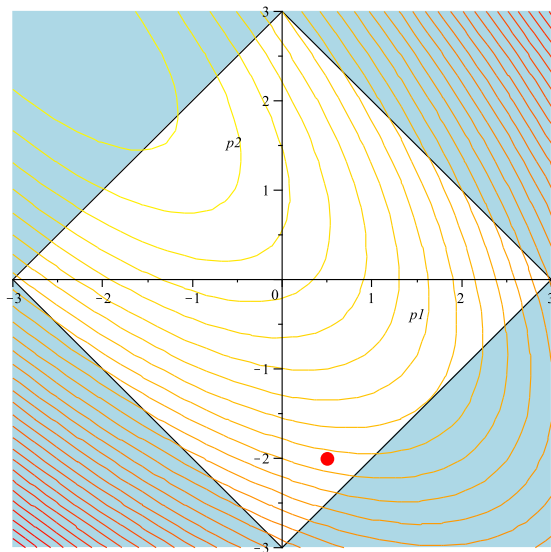
- 1 a linear programming subproblem (the so-called *column generation problem*) which generates an extreme point of the polyhedron, and
- 2 a nonlinear *restricted master problem* (RMP) which finds the maximum of the objective function over the convex hull (a simplex) of previously defined extreme points.

Its principal characteristic is that the sequence of successive solutions to the master problem tends to a solution to the original problem in such a way that the objective function strictly monotonically approaches its optimal value.

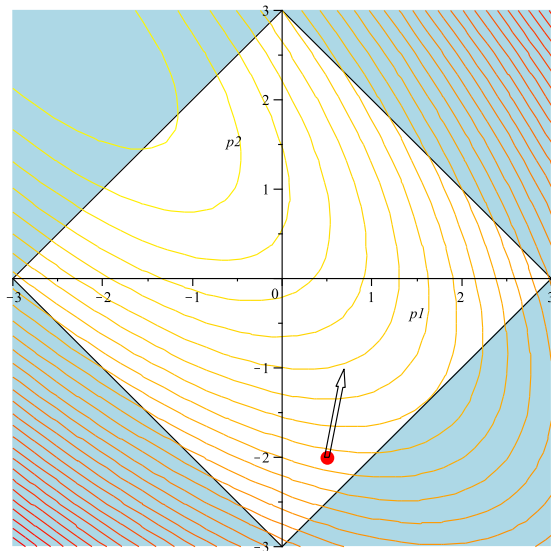
Simplicial decomposition



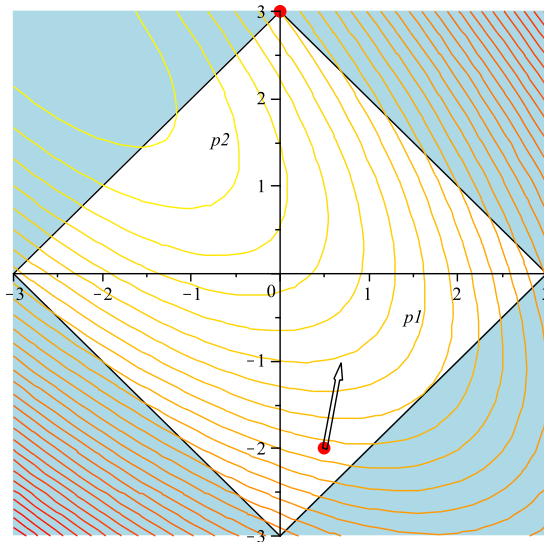
Simplicial decomposition



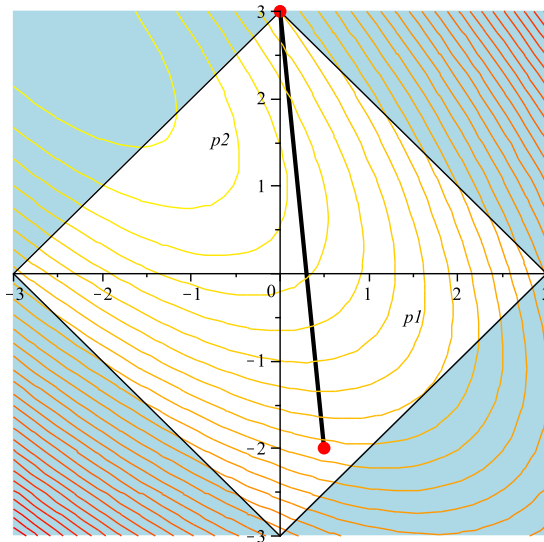
Simplicial decomposition



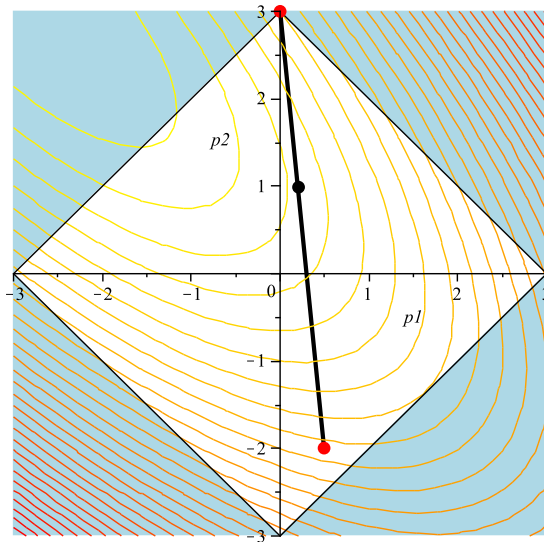
Simplicial decomposition



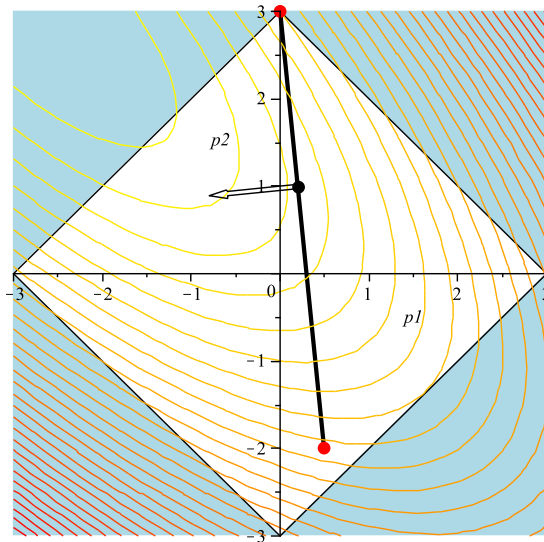
Simplicial decomposition



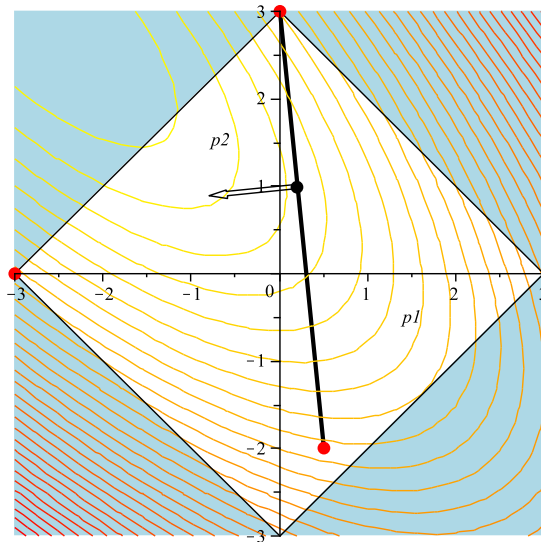
Simplicial decomposition



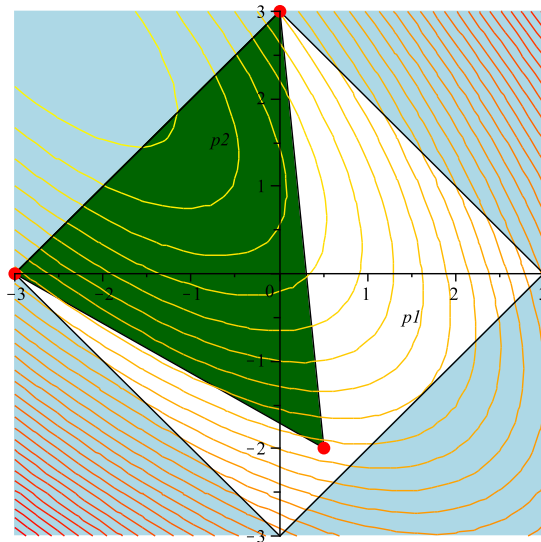
Simplicial decomposition



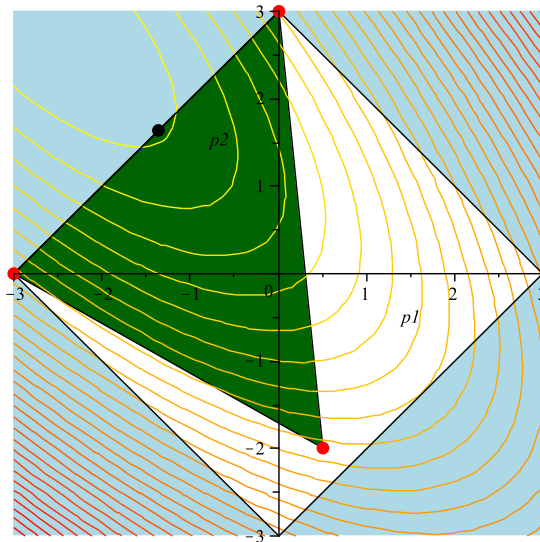
Simplicial decomposition



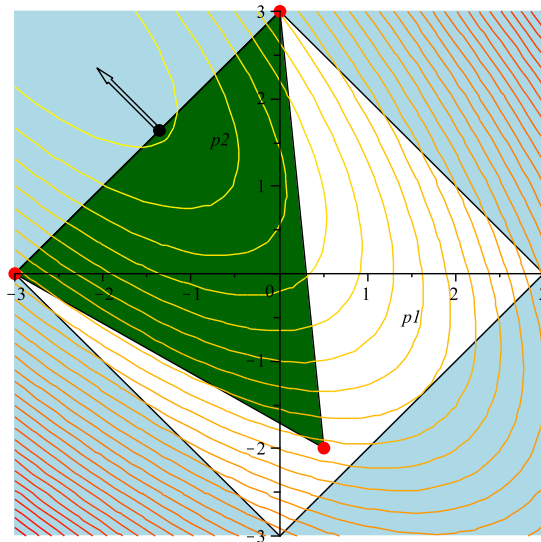
Simplicial decomposition



Simplicial decomposition



Simplicial decomposition



Algorithm SD

Step 0: (**Initialization**) Guess an initial solution $\mathbf{p}^{(0)} \in P$ such that $\mathbf{M}(\mathbf{p}^{(0)})$ is nonsingular. Set $I = \{1, \dots, n\}$, $Q^{(0)} = \{\mathbf{p}^{(0)}\}$ and $k = 0$.

Algorithm SD

Step 0: (**Initialization**) Guess an initial solution $\mathbf{p}^{(0)} \in P$ such that $\mathbf{M}(\mathbf{p}^{(0)})$ is nonsingular. Set $I = \{1, \dots, n\}$, $Q^{(0)} = \{\mathbf{p}^{(0)}\}$ and $k = 0$.

Step 1: (**Termination check**) Set

$$\begin{aligned} I_{\text{ub}}^{(k)} &= \{i \in I \mid p_i^{(k)} = b_i\}, \\ I_{\text{im}}^{(k)} &= \{i \in I \mid 0 < p_i^{(k)} < b_i\}, \\ I_{\text{lb}}^{(k)} &= \{i \in I \mid p_i^{(k)} = 0\}. \end{aligned}$$

If

$$\phi_i(\mathbf{p}^{(k)}) \begin{cases} \geq \lambda & \text{if } i \in I_{\text{ub}}^{(k)}, \\ = \lambda & \text{if } i \in I_{\text{im}}^{(k)}, \\ \leq \lambda & \text{if } i \in I_{\text{lb}}^{(k)} \end{cases}$$

for some $\lambda \in \mathbb{R}_+$, then STOP and $\mathbf{p}^{(k)}$ is optimal.

Step 2: (Solution of the column generation subproblem)

Compute

$$\mathbf{q}^{(k+1)} = \arg \max_{\mathbf{p} \in P} \phi(\mathbf{p}^{(k)})^T \mathbf{p}$$

and set

$$Q^{(k+1)} = Q^{(k)} \cup \{\mathbf{q}^{(k+1)}\}.$$

Step 2: (Solution of the column generation subproblem)

Compute

$$\mathbf{q}^{(k+1)} = \arg \max_{\mathbf{p} \in P} \phi(\mathbf{p}^{(k)})^T \mathbf{p}$$

and set

$$Q^{(k+1)} = Q^{(k)} \cup \{\mathbf{q}^{(k+1)}\}.$$

Step 3: (Solution of the restricted master subproblem)

Find

$$\mathbf{p}^{(k+1)} = \arg \max_{\mathbf{p} \in \text{co}(Q^{(k+1)})} \Phi(\mathbf{M}(\mathbf{p}))$$

and purge $Q^{(k+1)}$ of all extreme points with zero weights in the resulting expression of $\mathbf{p}^{(k+1)}$ as a convex combination of elements in $Q^{(k+1)}$.

Increment k by one and go back to Step 1.

Column generation problem

Basically, it is an extremely simple **linear programming problem**:

$$\begin{array}{ll}\text{maximize} & \mathbf{c}^\top \mathbf{p} \\ \text{subject to} & \mathbf{0} \preceq \mathbf{p} \preceq \mathbf{b}, \mathbf{1}^\top \mathbf{p} = 1\end{array}$$

where $\mathbf{c} = \phi(\mathbf{p}^{(k)})$.

Column generation problem

Basically, it is an extremely simple **linear programming problem**:

$$\begin{aligned} & \text{maximize} && \mathbf{c}^T \mathbf{p} \\ & \text{subject to} && \mathbf{0} \preceq \mathbf{p} \preceq \mathbf{b}, \mathbf{1}^T \mathbf{p} = 1 \end{aligned}$$

where $\mathbf{c} = \phi(\mathbf{p}^{(k)})$.

Just pick the consecutive largest components c_i of $\mathbf{c}$ and set the corresponding weights q_i as their maximal allowable values b_i . The process is repeated until the sum of the assigned weights exceeds one. Then the value of the last weight which was set in this manner should be corrected so that the weights sum up to one. The remaining (i.e., unassigned) weights are set as zeros.

Solution of the restricted master problem

Suppose that in the $(k + 1)$ -th iteration of SD, we have

$$Q^{(k+1)} = \{\mathbf{q}_1, \dots, \mathbf{q}_r\}$$

possibly with $r < k + 1$ (owing to the deletion mechanism of uninformative points). Step 3 of Algorithm SD involves maximization of

$$\Phi[\mathbf{M}(\mathbf{p})] = \log \det(\mathbf{M}(\mathbf{p}))$$

over

$$\text{co}(Q^{(k+1)}) = \left\{ \mathbf{p} = \sum_{j=1}^r w_j \mathbf{q}_j \mid \mathbf{w} \succeq \mathbf{0}, \mathbf{1}^\top \mathbf{w} = 1 \right\}$$

Solution of the restricted master problem

Equivalent formulation of the RMP

Find the sequence of weights $\mathbf{w} \in \mathbb{R}^r$ to maximize

$$\Psi(\mathbf{w}) = \log \det(\mathbf{H}(\mathbf{w}))$$

subject to the constraints

$$\mathbf{1}^\top \mathbf{w} = 1$$

$$\mathbf{w} \succeq \mathbf{0}$$

Notation:

- $\mathbf{H}(\mathbf{w}) = \sum_{j=1}^r w_j \mathbf{H}_j$
- $\mathbf{H}_j = \mathbf{M}(\mathbf{q}_j)$

Optimality condition for RMP

Theorem

Suppose that the matrix $\mathbf{H}(\mathbf{w}^*)$ is nonsingular for some $\mathbf{w}^* \in S_r$. The vector $\mathbf{w}^*$ constitutes a global solution to the RMP if and only if

$$\psi_j(\mathbf{w}^*) \begin{cases} = m & \text{if } w_j^* > 0 \\ \leq m & \text{if } w_j^* = 0 \end{cases}$$

for each $j = 1, \dots, r$, where

$$\psi_j(\mathbf{w}) = \text{trace}[\mathbf{H}(\mathbf{w})^{-1} \mathbf{H}_j], \quad j = 1, \dots, r$$

Multiplicative algorithm for the RMP

Step 0: (Initialization)

Set $\mathbf{w}^{(0)} = (1/r)\mathbf{1}$ and $\ell = 0$.

Multiplicative algorithm for the RMP

Step 0: (Initialization)

Set $\mathbf{w}^{(0)} = (1/r)\mathbf{1}$ and $\ell = 0$.

Step 1: (Termination check)

If

$$\frac{1}{m}\psi(\mathbf{w}^{(\ell)}) \preceq \mathbf{1}$$

then STOP.

Multiplicative algorithm for the RMP

Step 0: (Initialization)

Set $\mathbf{w}^{(0)} = (1/r)\mathbf{1}$ and $\ell = 0$.

Step 1: (Termination check)

If

$$\frac{1}{m}\psi(\mathbf{w}^{(\ell)}) \preceq \mathbf{1}$$

then STOP.

Step 2: (Multiplicative update)

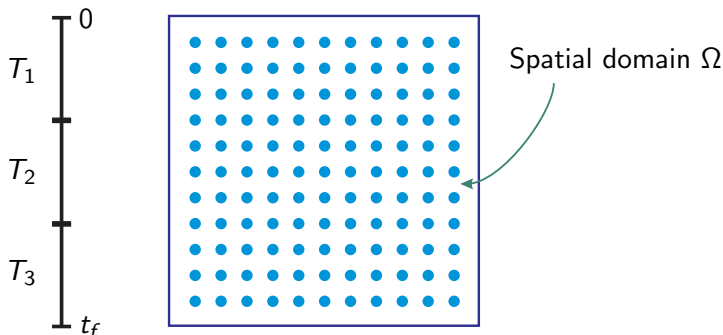
Evaluate

$$\mathbf{w}^{(\ell+1)} = \frac{1}{m}\psi(\mathbf{w}^{(\ell)}) \cdot \mathbf{w}^{(\ell)}$$

Increment ℓ by one and go to Step 1.

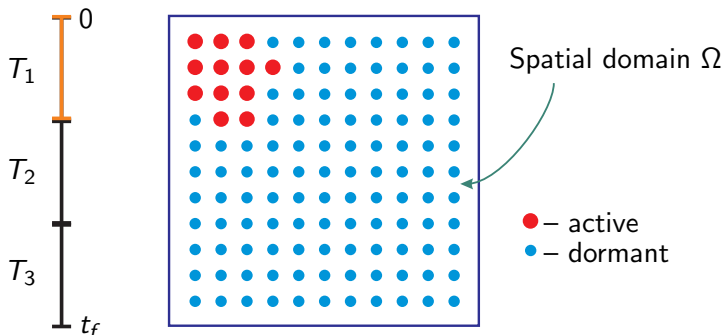
Extension to scanning observations?

Given N' sensors whose spatial positions are fixed *a priori*, partition the observation horizon $T = (0, t_f)$ into adjacent subintervals T_k , $k = 1, \dots, K$ and activate a best N -element subset of sensors over each T_k .



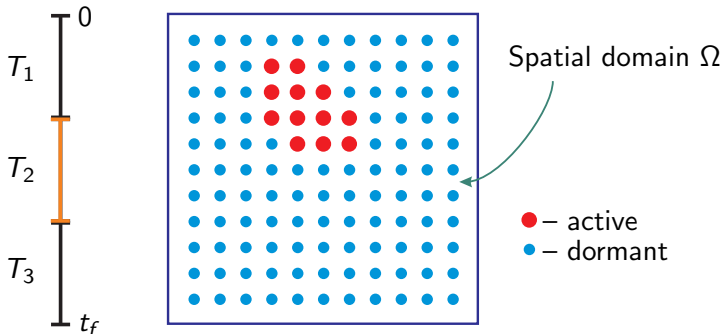
Extension to scanning observations?

Given N' sensors whose spatial positions are fixed *a priori*, partition the observation horizon $T = (0, t_f)$ into adjacent subintervals T_k , $k = 1, \dots, K$ and activate a best N -element subset of sensors over each T_k .



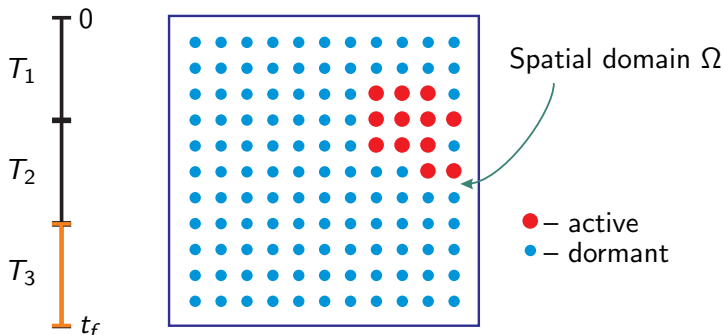
Extension to scanning observations?

Given N' sensors whose spatial positions are fixed *a priori*, partition the observation horizon $T = (0, t_f)$ into adjacent subintervals T_k , $k = 1, \dots, K$ and activate a best N -element subset of sensors over each T_k .



Extension to scanning observations?

Given N' sensors whose spatial positions are fixed *a priori*, partition the observation horizon $T = (0, t_f)$ into adjacent subintervals T_k , $k = 1, \dots, K$ and activate a best N -element subset of sensors over each T_k .



Extension to scanning observations?

For K scanning stages we have K vectors of weights

$$\mathbf{p}_k = (p_k^1, \dots, p_k^n), \quad k = 1, \dots, K$$

Problem

Find vectors $\mathbf{p}, \dots, \mathbf{p}_K$ to maximize

$$\log \det \left(\sum_{k=1}^K \mathbf{M}_k(\mathbf{p}_k) \right)$$

subject to

$$\begin{aligned} \mathbf{0} &\preceq \mathbf{p}_k \preceq \mathbf{b}, \quad k = 1, \dots, K \\ \mathbf{1}^\top \mathbf{p}_k &= 1, \quad k = 1, \dots, K \end{aligned}$$

We solve it using a **block-coordinate ascent method**.

Illustrative example

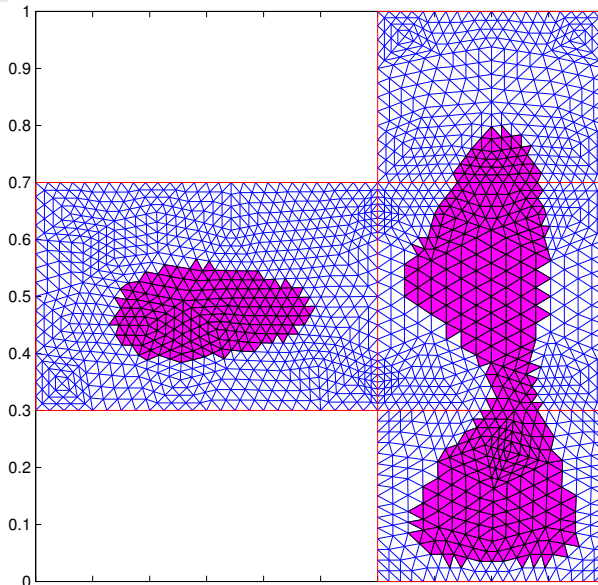
Consider $T = (0, 1)$

$$\begin{aligned}\frac{\partial y}{\partial t} &= \nabla \cdot (\kappa \nabla y) + u && \text{in } \Omega \times T \\ y(\mathbf{x}, 0) &= 5 && \text{in } \Omega \\ y(\mathbf{x}, t) &= 5(1 - t) && \text{on } \partial\Omega \times T\end{aligned}$$

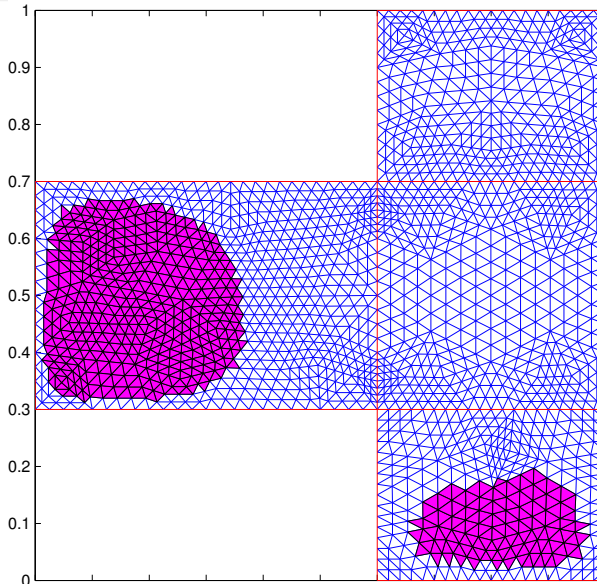
where

$$\begin{aligned}\kappa(\mathbf{x}) &= \theta_1 + \theta_2 x_1 + \theta_3 x_2 \\ \theta_1^0 &= 0.1, \quad \theta_2^0 = -0.05, \quad \theta_3^0 = 0.2 \\ u(x, t) &= 20 \exp(-50(x_1 - t)^2) \\ \omega(A) &= c \frac{|A|}{|X|}, \quad c = 4.3\end{aligned}$$

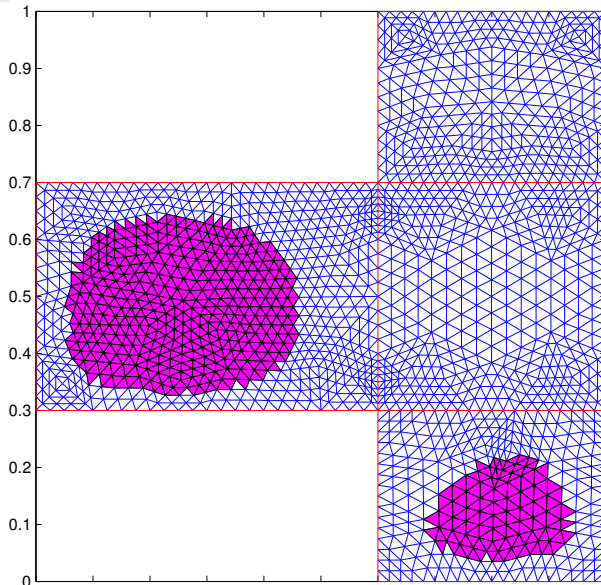
Stationary sensors: $\det(M) = 190.19$



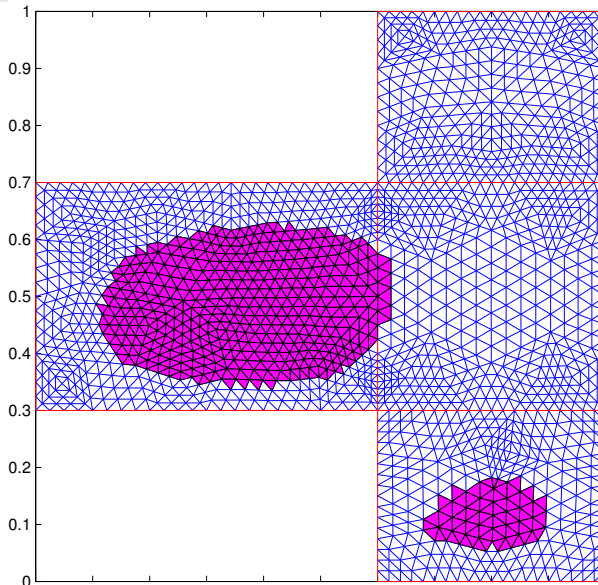
Stationary sensors: $K = 5$, $\det(M) = 460.82$



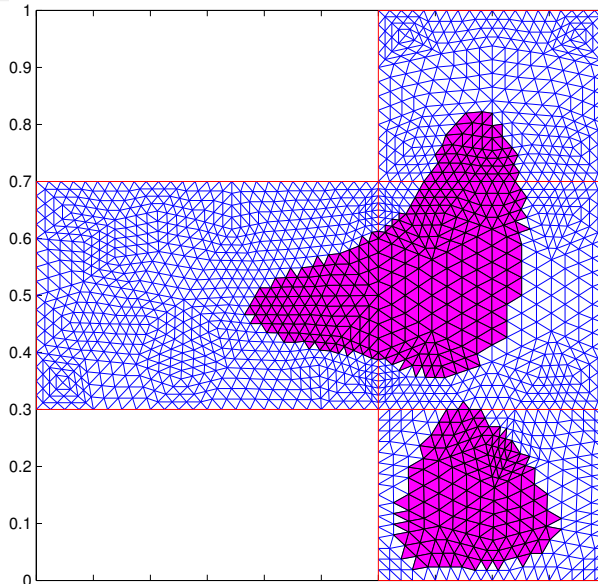
Stationary sensors: $K = 5$, $\det(M) = 460.82$



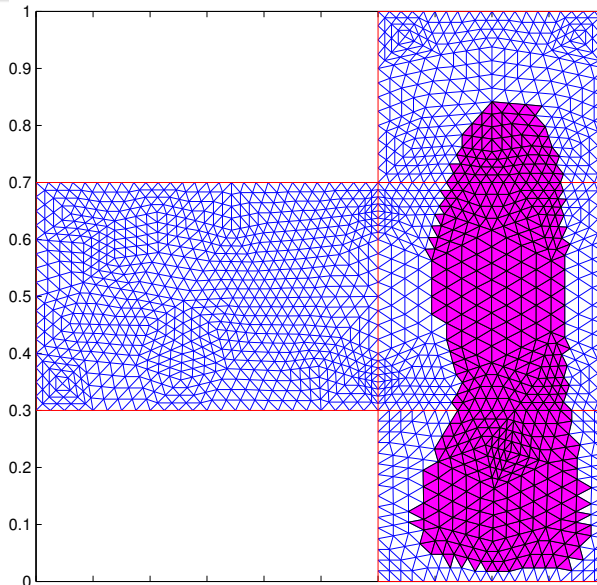
Stationary sensors: $K = 5$, $\det(M) = 460.82$



Stationary sensors: $K = 5$, $\det(M) = 460.82$



Stationary sensors: $K = 5$, $\det(M) = 460.82$



Concluding remarks at this stage

- Dramatically reduces the problem dimensionality and complexity.
- Provides characterizations of optimal activations strategies.
- Develops a very efficient block-coordinate ascent algorithm combined with simplicial decomposition for a large number of sensor nodes.