Optimum experimental design for infinite dimensional inverse problems

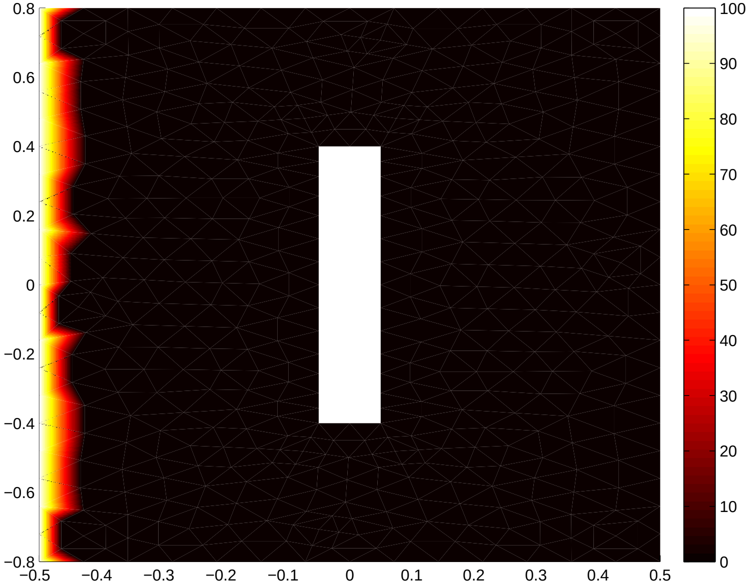
Roland Herzog[1] **Ilka Riedel**[1] **Dariusz Uciński**[2]

[1]Faculty of Mathematics, Technical University of Chemnitz, Germany

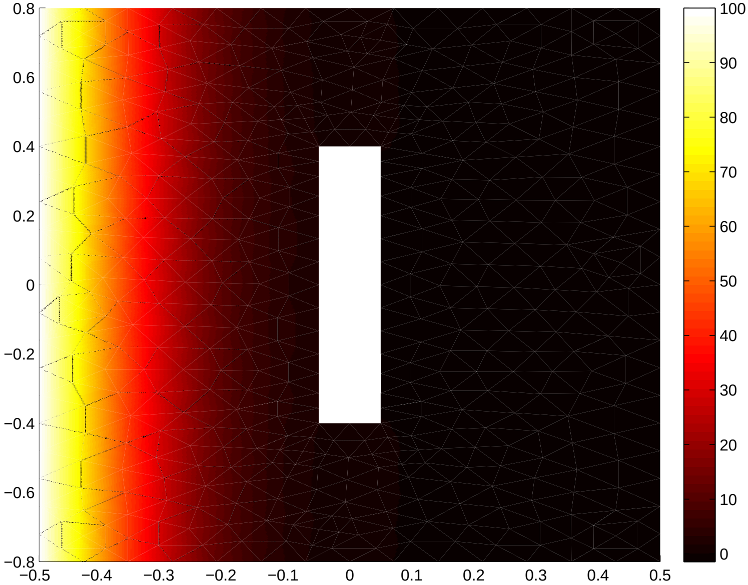
[2]Institute of Control and Computation Eng., University of Zielona Góra, Poland

Workshop on Design of Experiments, April 30–May 4 2018,
CIRM, Marseilles, France

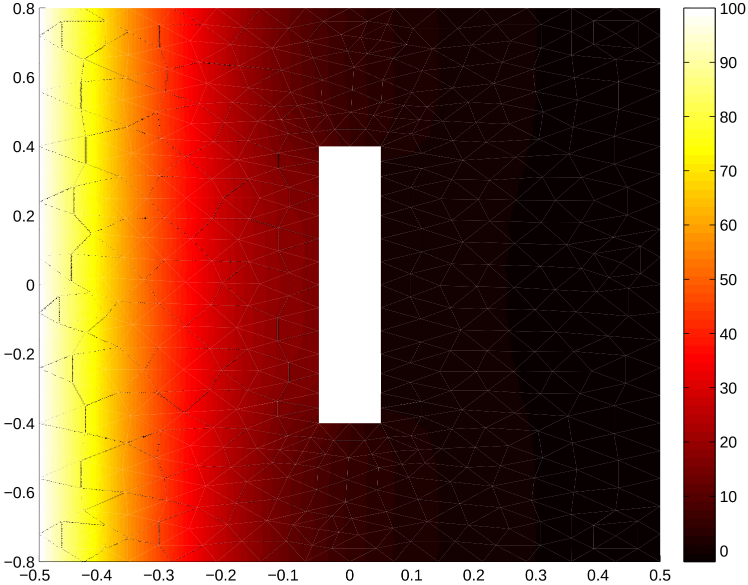
Example: Heating a metal block with a crack



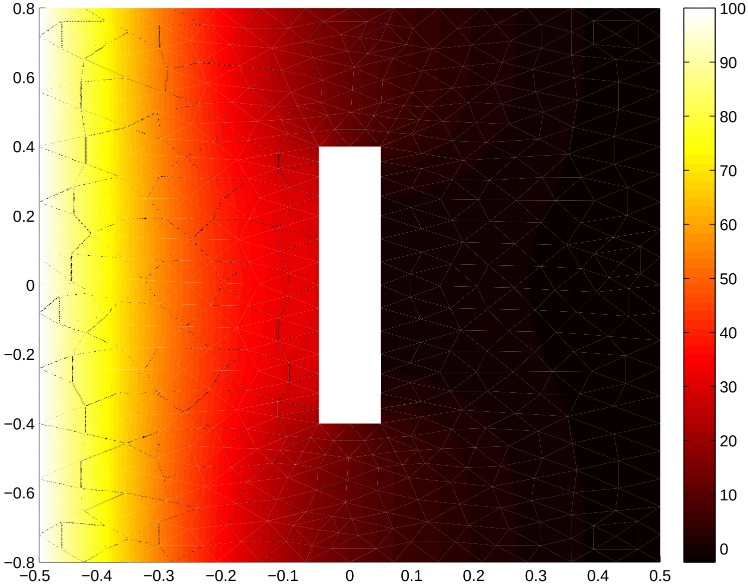
Example: Heating a metal block with a crack



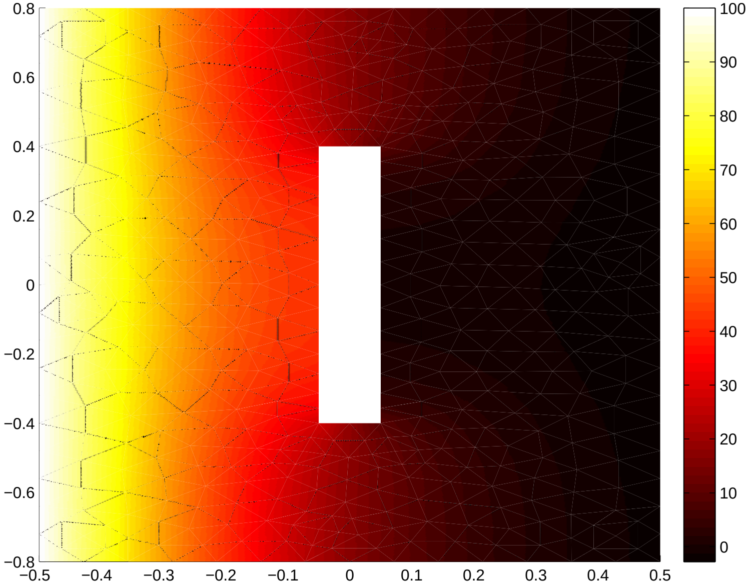
Example: Heating a metal block with a crack



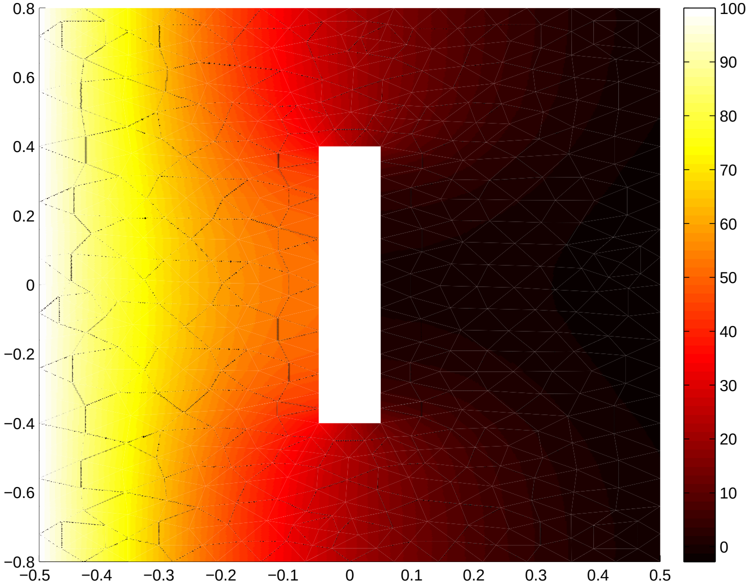
Example: Heating a metal block with a crack



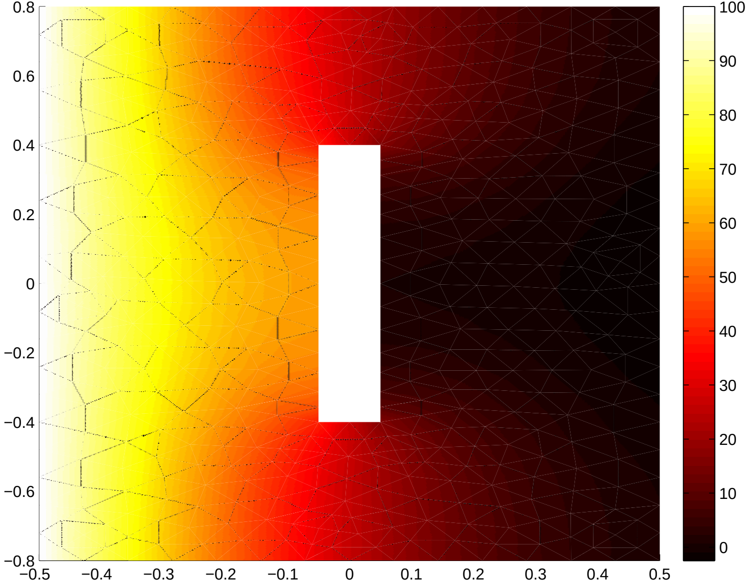
Example: Heating a metal block with a crack



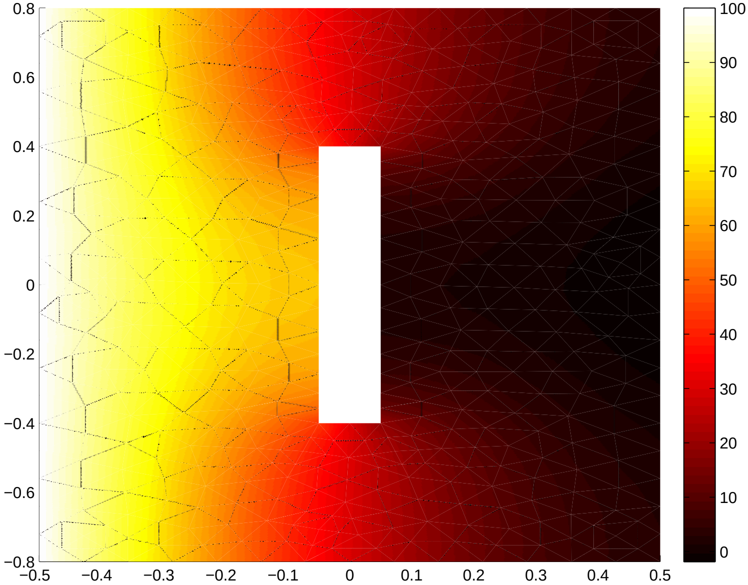
Example: Heating a metal block with a crack



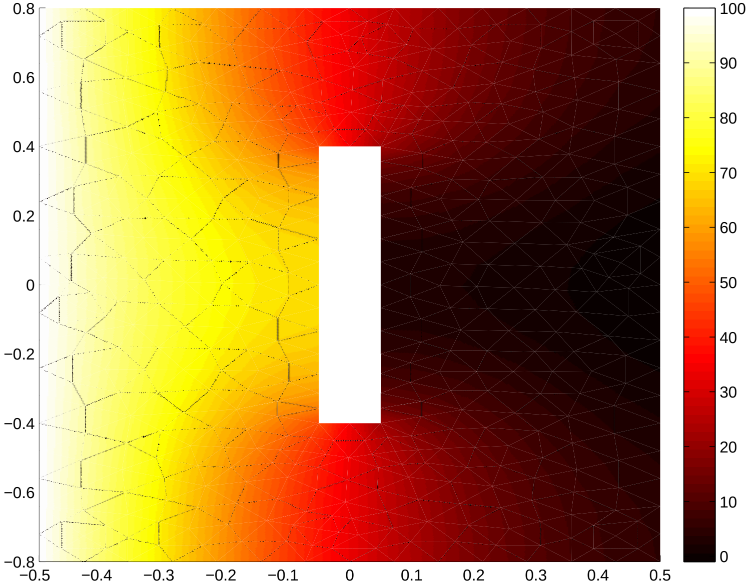
Example: Heating a metal block with a crack



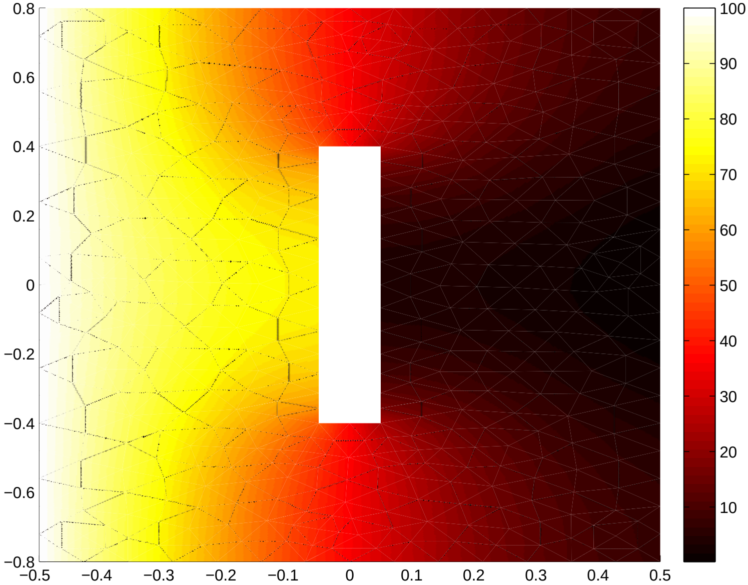
Example: Heating a metal block with a crack



Example: Heating a metal block with a crack



Example: Heating a metal block with a crack



Example ctd': Mathematical model

Let $T(\chi, t)$ be the temperature at spatial point χ and time t.
Its evolution is governed by the partial differential equation (PDE)

$$\frac{\partial T}{\partial t} = \alpha \Delta y \quad \text{for } x \in \Omega \text{ and } t \in [0, t_f]$$

α being a parameter, subject to the initial condition

$$T = T_0 \quad \text{in } \Omega \text{ at } t = 0$$

and the boundary conditions

$$T = 100 \qquad \text{on the left side of } \Omega \qquad \text{(Dirichlet condition)}$$

$$\frac{\partial T}{\partial n} = -10 \qquad \text{on the right side of } \Omega \qquad \text{(Neumann condition)}$$

$$\frac{\partial T}{\partial n} = 0 \qquad \text{on all other boundaries} \qquad \text{(Neumann condition)}$$

Example ctd': Calibration and measurement design

If thermal conductivity α and initial condition T_0 are unknown, which is rather typical, we could observe the temperature evolution at a fixed point and tune α and T_0 so that our model fits the data as best as it can.

Example ctd': Calibration and measurement design

If thermal conductivity α and initial condition T_0 are unknown, which is rather typical, we could observe the temperature evolution at a fixed point and tune α and T_0 so that our model fits the data as best as it can.

Problem

Where to place the pyrometer so as to get the most valuable information about α and T_0?

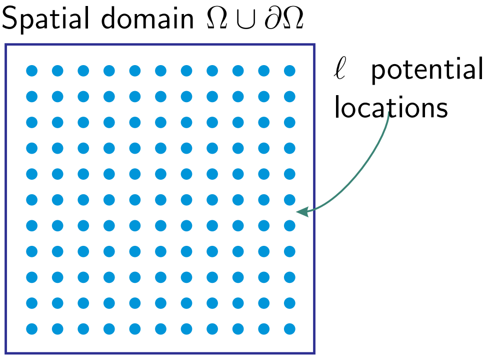
Such situations (PDEs and related sensor location problems) are common in engineering practice.

Other motivating examples

- environment observation and forecasting systems
    - air pollution monitoring,
    - groundwater resources management and river monitoring,
    - military applications: surveillance and inspection in hazardous environments,
- fault detection and isolation,
- emerging smart materials,
- intelligent building monitoring,
- habitat monitoring,
- intelligent traffic systems,
- computer-assisted tomography,
- recovery of valuable minerals and hydrocarbon from underground permeable reservoirs,
- and many more ...

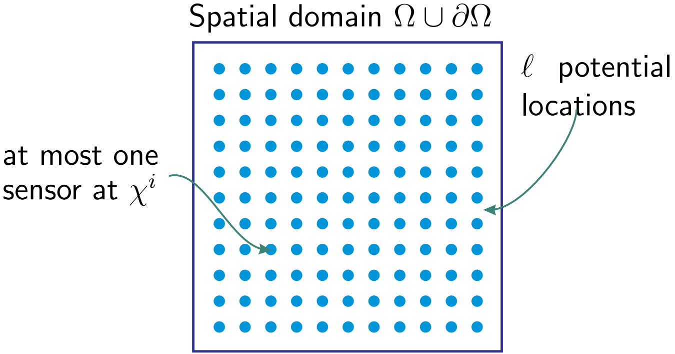
Observations

Given m sensors and a finite set $\{\chi^1, \dots, \chi^\ell\} \subset \Omega \cup \partial\Omega$ of points at which they can be placed, consider their fixed configuration.



Observations

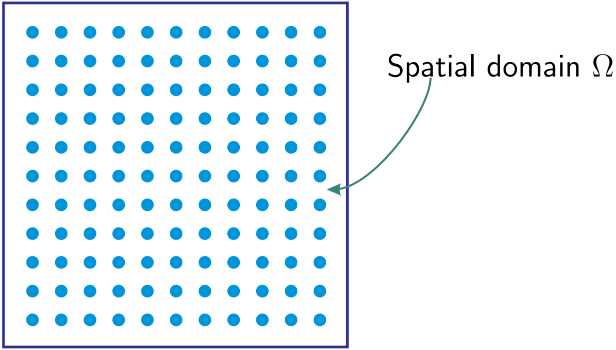
Given m sensors and a finite set $\{\chi^1, \ldots, \chi^\ell\} \subset \Omega \cup \partial\Omega$ of points at which they can be placed, consider their fixed configuration.



Observations

Equivalently, we can treat $\{\chi^1, \ldots, \chi^\ell\} \subset \Omega \cup \partial\Omega$ as gauged sites at which ℓ sensors reside and only m from among them are to be activated.

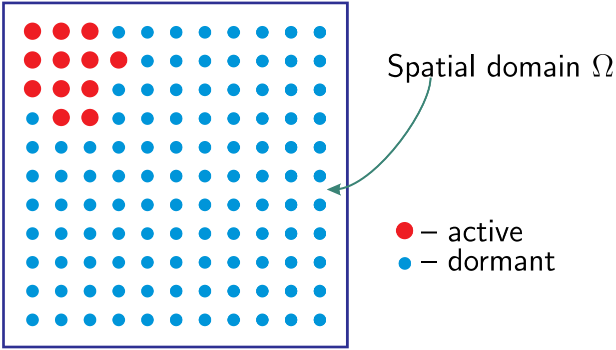
Combinatorial problem: Which m-element subset to select?



Observations

Equivalently, we can treat $\{\chi^1, \dots, \chi^\ell\} \subset \Omega \cup \partial\Omega$ as gauged sites at which ℓ sensors reside and only m from among them are to be activated.

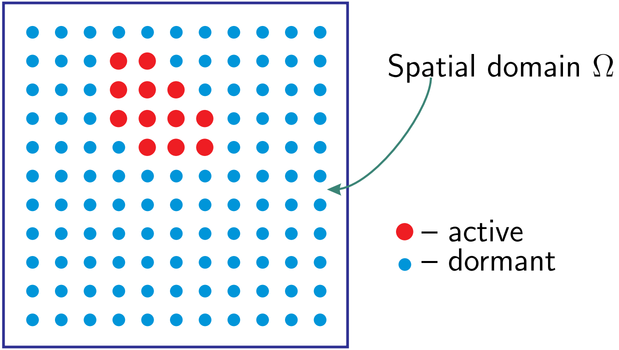
Combinatorial problem: Which m-element subset to select?



Observations

Equivalently, we can treat $\{\chi^1, \ldots, \chi^\ell\} \subset \Omega \cup \partial\Omega$ as gauged sites at which ℓ sensors reside and only m from among them are to be activated.

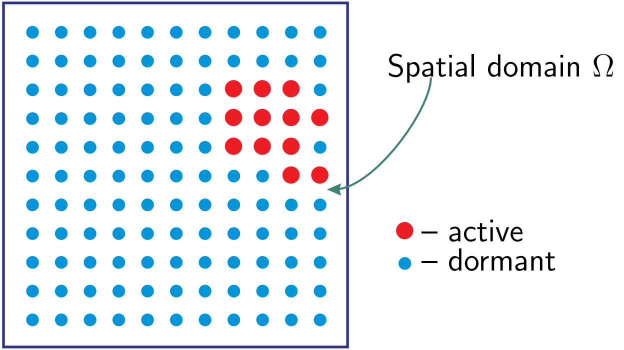
Combinatorial problem: Which m-element subset to select?



Observations

Equivalently, we can treat $\{\chi^1, \ldots, \chi^\ell\} \subset \Omega \cup \partial\Omega$ as gauged sites at which ℓ sensors reside and only m from among them are to be activated.

Combinatorial problem: Which m-element subset to select?



Spatiotemporal dynamics

Distributed parameter system—dynamic system whose state depends on both time and space; its model (a partial differential equation) is known up to a vector of unknown parameters $\boldsymbol{\theta}$.

Observations—using sensors in order to estimate $\boldsymbol{\theta}$.

Spatiotemporal dynamics

Distributed parameter system—dynamic system whose state depends on both time and space; its model (a partial differential equation) is known up to a vector of unknown parameters $\boldsymbol{\theta}$.

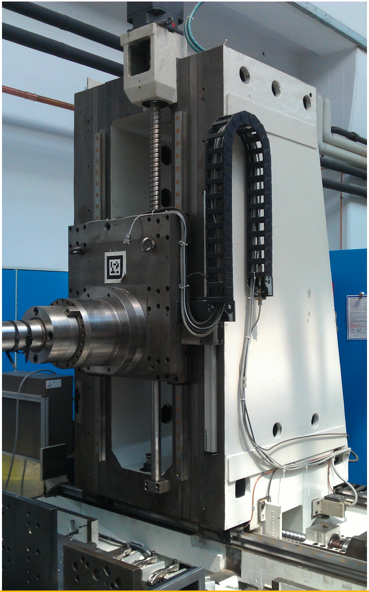
Observations—using sensors in order to estimate $\boldsymbol{\theta}$.

Crucial difficulty: In the previous example, the component T_0 of $\boldsymbol{\theta} = (\alpha, T_0)$ is an element of an infinite dimensional functional space (here $L^2(\Omega)$).

Approaches to infinite dimensional inverse problems

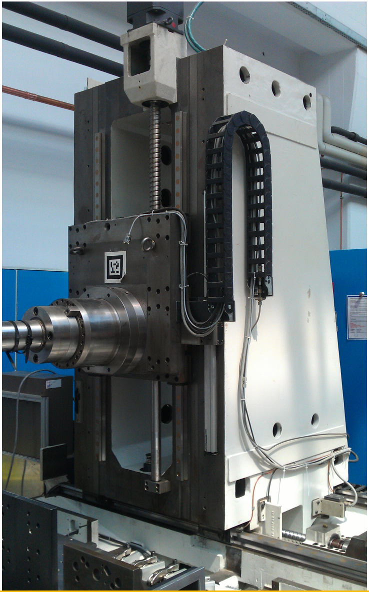
1. Stuart, *Inverse problems: A Bayesian perspective*, Acta Numerica, 2010
2. Alexanderian, Petra, Stadler, & Ghattas, *A-optimal design of experiments for infinite-dimensional bayesian linear inverse problems with regularized* ℓ_0*-sparsification*, SIAM J. Sci. Comput., 2014
3. Alexanderian, Gloor, & Ghattas, *On Bayesian A- and D-optimal experimental designs in infinite dimensions*, Bayesian Anal., 2016
4. Alexanderian, Petra, Stadler, & Ghattas, *A fast and scalable method for A-optimal design of experiments for infinite-dimensional bayesian nonlinear inverse problems*, SIAM J. Sci. Comput., 2016
5. Gejadze & Shutyaev, *On computation of the design function gradient for the sensor-location problem in variational data assim.*, SIAM J. Sci. Comput., 2012
6. Gejadze, Le Dimet, & Shutyaev, *On analysis error covariances in variational data assimilation*, SIAM J. Sci. Comput., 2008
7. Gejadze, Le Dimet, & Shutyaev, *On optimal solution error covariances in variational data assimilation problems*, J. Computat. Phys., 2010
8. Gejadze, Copeland, Le Dimet, & Shutyaev, *Computation of the analysis error covariance in variational data assimilation problems with nonlinear dynamics*, J. Comput. Phys., 2011
9. Haber, Horesh, & Tenorio, *Numerical methods for experimental design of large-scale linear ill-posed inverse problems*, Inv. Problems, 2008
10. Haber, Horesh, & Tenorio, *Numerical methods for the design of large-scale nonlinear discrete ill-posed inverse problems*, Inv. Problems, 2010
11. Haber, Magnant, Lucero, & Tenorio, *Num. methods for A-optimal designs with a sparsity constraint for ill-posed inv. problems*, Comput. Optim. Applics., 2012

Example: Deformation of machine tools



The prototype machine column Auerbach ACW 360 deforms during its operation due to waste heat from two external electical drives (the one at the top moves the sledge holding the main spindle and the one at the bottom moves the entire machine column).

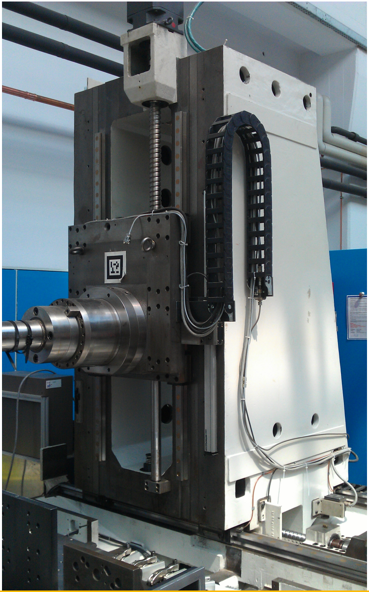
Example: Deformation of machine tools



The inhomogeneous transient temperature field displaces the tool centre point (TCP) and thus reduces production accuracy and product quality.

Key idea: Predict the machine tool deformation based on measurements from temperature sensors and then correct it on line.

Example: Deformation of machine tools



Key problem: Temperature sensors may be placed only on the surface of the machine column. But which locations are best?

Thermo-mechanical system

Heat equation

$$\rho\, c_p\, \dot{T} - \operatorname{div}(\lambda \nabla T) = 0 \quad \text{in } \Omega \times (0, t_f)$$

$$\lambda \frac{\partial}{\partial n} T + \alpha(\chi)\,(T - T_{\text{ref}}) = r(\chi, t) \quad \text{on } \partial\Omega \times (0, t_f)$$

$$T(\chi, 0) = T_0(\chi) \quad \text{in } \Omega$$

T	temperature
r	thermal surface load
ρ	density
c_p	specific heat at constant pressure
λ	thermal conductivity
α	coefficient of heat transfer
T_{ref}	ambient temperature
T_0	initial temperature

Thermo-mechanical system

Linear elasticity (BCs omitted)

$$-\operatorname{div}\boldsymbol{\sigma}\big(\boldsymbol{\varepsilon}(\boldsymbol{u}),T(t_f)\big) = \mathbf{0} \quad \text{in } \Omega$$

$$\boldsymbol{\sigma}\big(\boldsymbol{\varepsilon}(\boldsymbol{u}),T(t_f)\big) = \boldsymbol{\sigma}^{\mathsf{el}}(\boldsymbol{\varepsilon}(\boldsymbol{u})) + \boldsymbol{\sigma}^{\mathsf{th}}(T(t_f))$$

$$\boldsymbol{\sigma}^{\mathsf{el}}(\boldsymbol{\varepsilon}(\boldsymbol{u})) = \frac{E}{1+\nu}\boldsymbol{\varepsilon}(\boldsymbol{u}) + \frac{E\nu}{(1+\nu)(1-2\nu)}\operatorname{trace}(\boldsymbol{\varepsilon}(\boldsymbol{u}))\,\mathsf{id}$$

$$\boldsymbol{\sigma}^{\mathsf{th}}(T(t_f)) = -\frac{E}{1-2\nu}\,\beta\,\big(T(t_f) - T_{\mathsf{ref}}\big)\,\mathsf{id}_3$$

$$\boldsymbol{\varepsilon}(\boldsymbol{u}) = \frac{1}{2}(\nabla\boldsymbol{u} + \nabla\boldsymbol{u}^\top)$$

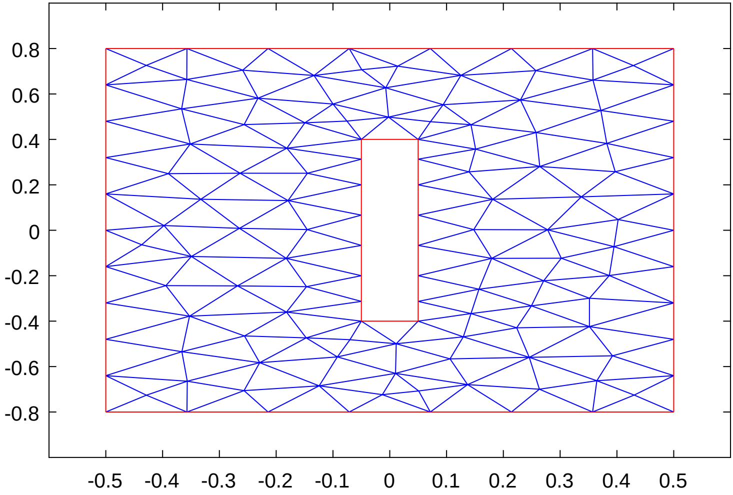
$\boldsymbol{u}$	displacement
$\boldsymbol{\sigma}$	stress
$\boldsymbol{\varepsilon}$	strain
ν, E	Poisson's ratio and modulus of elasticity, resp.
β	thermal volumetric expansion coefficient

Semi-discretization (method of lines)

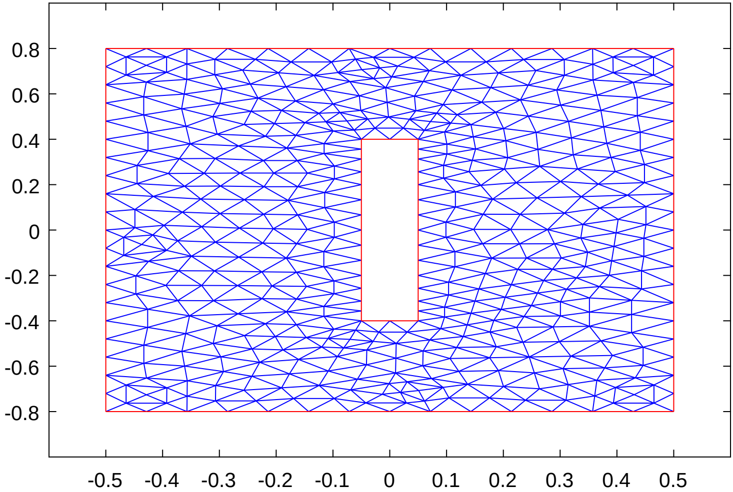
The finite-element method (FEM) is a technique for solving PDEs by first discretizing these equations in their space dimensions. The discretization is carried out locally over small regions of simple but arbitrary shape (the finite elements). Each vertex is called a node.

The solution should be simple on each element. Piecewise linear functions are a good choice.

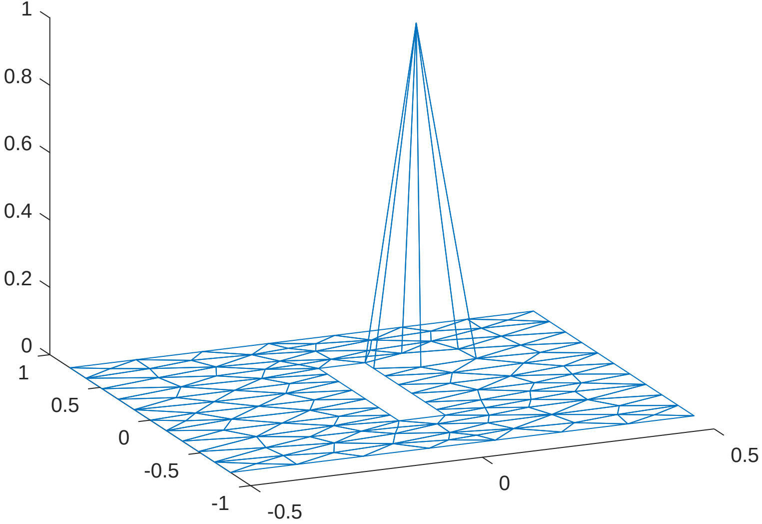
Triangulation of the spatial domain



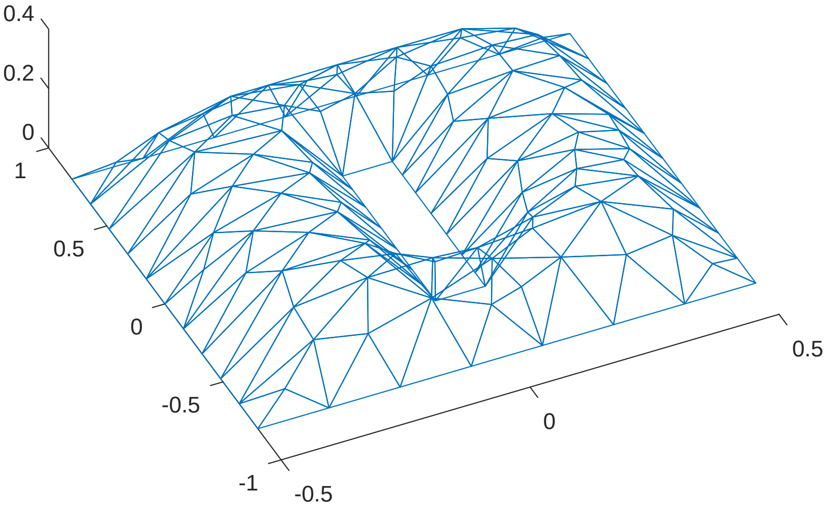
Triangulation of the spatial domain



Basis functions



Piecewise linear function



Resulting approximations

Tetrahedral finite elements and piecewise linear continuous Lagrange basis functions $\{\varphi_j\}_{j=1}^{n}$ are used here (specifically, $n = 25\,615$). The basis functions correspond to nodal points $\{\chi_j\}_{j=1}^{n}$ such that

$$\varphi_j(\chi_i) = \delta_{ij} \quad \text{for } i, j \in \{1, \dots, n\}$$

We have approximations

$$T(\chi, t) \approx \sum_{j=1}^{n} x_j(t)\varphi_j(\chi), \quad T_0(\chi) \approx \sum_{j=1}^{n} x_{0,j}\varphi_j(\chi)$$

Define $x(t) = (x_1(t), \dots, x_n(t))$ and $x_0 = (x_{0,1}, \dots, x_{0,n})$.

Large-scale dynamic system

State equation

$$\begin{cases} E\,\dot{x}(t) = A(p)\,x(t) + f(t), \quad t \in [0, t_f] \\ \quad x(0) = x_0 \in \mathbb{R}^n \end{cases}$$

$E \in \mathbb{R}^{n \times n}$	a nonsingular matrix
$f(t) \in \mathbb{R}^n$	a known forcing input
$A(p) \in \mathbb{R}^{n \times n}$	a parameter-dependent matrix
$p \in \mathbb{R}^q$	the vector of unknown parameters (parameterization of the heat transfer coefficient)

Output equation (r sensors)

$$y_j = C_y\,x(t_j) + \eta_j \in \mathbb{R}^m, \quad j = 1, \ldots, N$$

$C_y \in \mathbb{R}^{m \times n}$	a matrix dependent on sensor locations
$\eta_j \sim \mathcal{N}(0, \sigma^2 \mathrm{id}_m)$	measurement noise

Data assimilation problem: Background info

The unknowns are x_0 and p. Our prior (background) information are their prior estimates x_0^{bg} and p^{bg} which are supposed to be realizations of Gaussian random vectors with means $\bar{x}_0 \in \mathbb{R}^n$ and $\bar{p} \in \mathbb{R}^q$, and covariance matrices $V_{x_0} \in \mathbb{R}^{n \times n}$ and $V_p \in \mathbb{R}^{q \times q}$, respectively, i.e., $x_0^{\mathsf{bg}} \sim \mathcal{N}(\bar{x}_0, V_{x_0})$ and $p^{\mathsf{bg}} \sim \mathcal{N}(\bar{p}, V_p)$.

Here $\bar{x}_0$ and $\bar{p}$ are unknown and interpreted as the 'true' initial state and the 'true' parameter, respectively. In turn, as for V_{x_0} and V_p, we assume that they are known and positive definite, and hence invertible.

Data assimilation problem: Objective

The number of unknowns ($n + q$) exceeds the number of measurements ($N\, m$). Consequently, regularization terms are needed expressing prior information about the unknowns:

$$\min_{x_0 \in \mathbb{R}^n,\, p \in \mathbb{R}^q} \mathcal{J}_{DA}(x_0, p) = \frac{1}{2}\|x_0 - x_0^{\mathsf{bg}}\|^2_{V_{x_0}^{-1}} + \frac{1}{2}\|p - p^{\mathsf{bg}}\|^2_{V_p^{-1}} + \frac{1}{2}\sum_{j=1}^{N}\|y_j - C_y\, x(t_j; x_0, p)\|^2_{V_y^{-1}},$$

where the term $x(t_j; x_0, p)$ is the solution to the state equation at sampling time t_j evaluated at given x_0 and p_0.

Digression: How to get V_{x_0}?

- Distribution law of T_0:

  $\mu = \mathcal{N}(\bar{T}_0, \mathcal{C})$ (Gaussian measure on Hilbert space $L^2(\Omega)$)

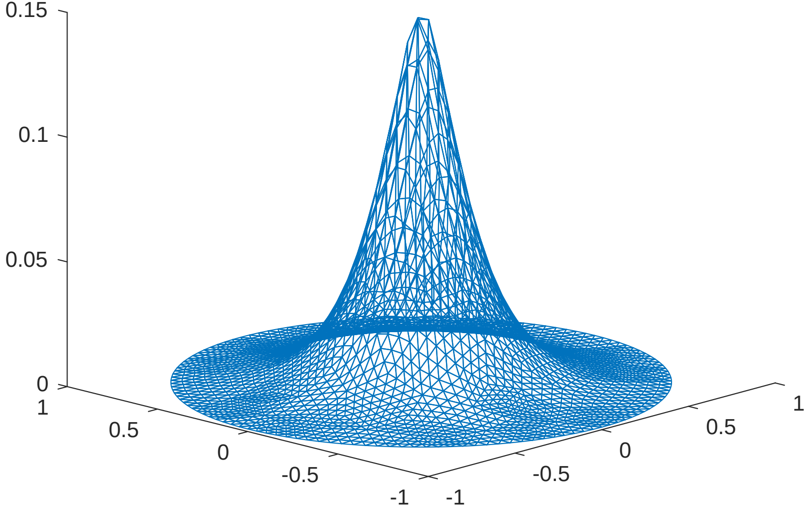
- Take covariance operator as square of inverse of Poisson-like operator:

  $$\mathcal{C} = (-\beta\Delta + \gamma I)^{-2}, \quad \beta, \gamma > 0$$

  $\mathcal{C}$ is positive, self-adjoint, of trace class; μ well-defined on $L^2(\Omega)$ (Stuart '10).

- $\gamma/\beta \propto$ correlation length; the larger β, the smaller the variance.
- $V_{x_0}^{-1}$ results from the FEM discretization of $\mathcal{C}^{-1}$.

Exemplary covariance function $c(0, \cdot\,)$



Variability of the estimates from the DA problem

DA produces estimate $\hat{\theta} = (\hat{x}_0, \hat{p})$ of the 'true' $\bar{\theta} = (\bar{x}_0, \bar{p})$.

Approximation via linearization

$$\operatorname{cov}(\hat{\theta}) \approx \Big(V_\theta^{-1} + \sum_{j=1}^{N} X(t_j)^\top C_y^\top V_y^{-1} C_y X(t_j) \Big)^{-1}$$

$$\begin{aligned} V_\theta &= \operatorname{diag}(V_{x_0}, V_p) \\ X(t) &= \big[\, X_0(t) \,\big|\, X_p(t) \,\big] \\ X_0(t) &= \frac{\partial}{\partial x_0} x(t; \bar{x}_0, \bar{p}) \in \mathbb{R}^{n \times n} \\ X_p(t) &= \frac{\partial}{\partial p} x(t; \bar{x}_0, \bar{p}) \in \mathbb{R}^{n \times q} \end{aligned}$$

Computation of sensitivities $X_0(t)$ and $X_p(t)$ is a formidable challenge. Here the adjoint approach has been adopted.

Digression: Randomized trace estimator

Randomized Gaussian trace estimator:

$$\mathrm{trace}(A^{-1}) = \frac{1}{M} \sum_{i=1}^{M} z_i \underbrace{A^{-1} z_i}_{=q_i}$$

where the z_i are M independent random vectors whose entries are i.i.d. standard normal variables. (q_i is evaluated by solving $Aq_i = z_i$.)

Covariance matrix of the QOI estimator

Our main concern is not to maximize the precision of $\hat{x}_0$ or $\hat{p}$, but rather to accurately estimate a quantity of interest z (the displacement of the tool centre point) depending on the terminal state $x(t_f)$ at time t_f,

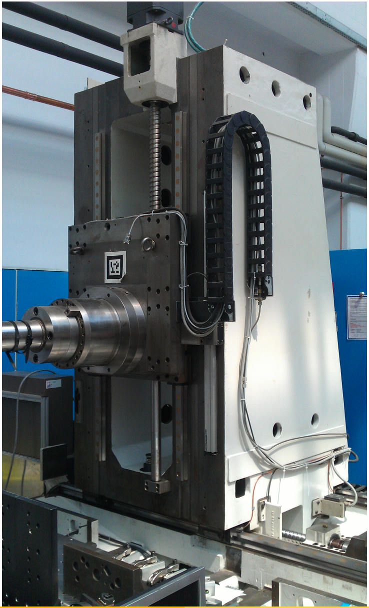
$$z = C_z\, x(t_f; \bar{\theta}) \in \mathbb{R}^r$$

through

$$\hat{z} = C_z\, x(t_f; \hat{\theta}) \in \mathbb{R}^r$$

with r small compared with the dimension n of the state variable (here $r = 3$).

Covariance matrix of the QOI estimator



Covariance matrix of the QOI estimator

Our main concern is not to maximize the precision of $\hat{x}_0$ or $\hat{p}$, but rather to accurately estimate a quantity of interest z (the displacement of the tool centre point) depending on the terminal state $x(t_f)$ at time t_f,

$$z = C_z\, x(t_f; \bar{\theta}) \in \mathbb{R}^r$$

through

$$\hat{z} = C_z\, x(t_f; \hat{\theta}) \in \mathbb{R}^r$$

with r small compared with the dimension n of the state variable (here $r = 3$).

Covariance matrix of the QOI estimator (ctd')

Linearization

$$\hat{z} \approx C_z x(t_f; \bar{\theta}) + Q\,(\hat{\theta} - \bar{\theta}),$$

with

$$Q = \left.\frac{\partial z}{\partial \theta}\right|_{\theta=\bar{\theta}} = C_z\, X(t_f; \bar{\theta}) \in \mathbb{R}^{r\times(n+q)}$$

yields

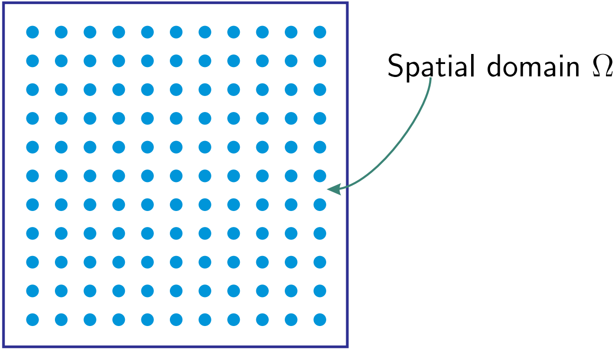
Variability of the QOI estimator

$$\operatorname{cov}(\hat{z}) = Q\, \operatorname{cov}(\hat{\theta})\, Q^\top$$

We are going to minimize its log-determinant through selection of best m sensor locations from among of a set of ℓ candidate locations.

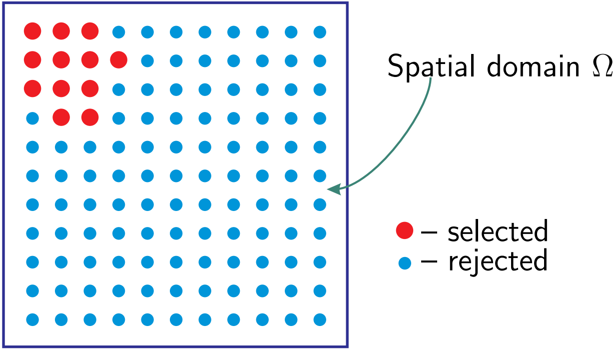
Observations

Combinatorial problem: Which m-element subset to select?



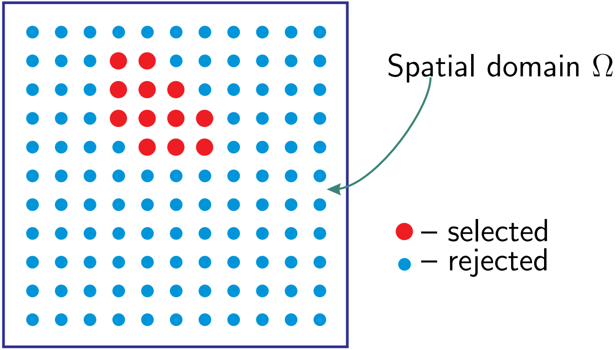
Observations

Combinatorial problem: Which m-element subset to select?



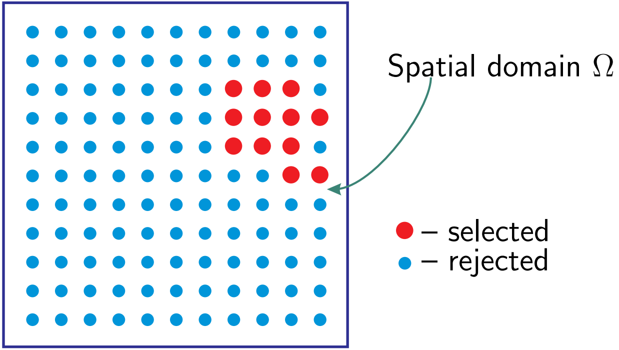
Observations

Combinatorial problem: Which m-element subset to select?



Observations

Combinatorial problem: Which m-element subset to select?



Structure of output matrix C_y

The temperature at the i-th candidate location and time instant t is linearly interpolated as $\pi_i^\top x(t)$, where the weight vector $\pi_i \in \mathbb{R}^n$ can be determined beforehand. (In the Lagrangian FE setting its components are simply the values of the corresponding shape functions at the i-th sensor location.) Let us stack the interpolation weights $\pi_i^\top$ in the matrix

$$\Pi = \begin{bmatrix} \pi_1^\top \\ \vdots \\ \pi_\ell^\top \end{bmatrix} \in \mathbb{R}^{\ell \times n}$$

Measurements at m gauged sites selected out of all ℓ available sites correspond to the measurement matrix $C_y \in \mathbb{R}^{m \times n}$ consisting of m distinct rows of matrix Π selected out of its ℓ rows. Thus the optimization of C_y can be understood as choosing best m rows of Π.

Binary decision variables

Given candidate spatial locations $\chi_1, \ldots, \chi^\ell$, introduce binary flags $w = (w_1, \ldots, w_\ell)$ satisfying

$$w_i = \begin{cases} 1 & \text{if sensor is placed at location } \chi^i \\ 0 & \text{if no sensor is at } \chi^i \end{cases}$$

Equivalently,

$$w_i = \begin{cases} 1 & \text{if the } i\text{-th row of } \Pi \text{ is included in } C_y \\ 0 & \text{otherwise} \end{cases}$$

Sensor selection as binary optimization

Sensor selection problem

Find a vector $w^\star_{\text{bin}} \in \mathbb{R}^\ell$ to minimize

$$\mathcal{J}(w) = \log\det\bigl(Q\,\mathcal{I}(w)^{-1}Q^\top\bigr)$$

subject to

$$\mathbf{1}_\ell^\top w = m$$

$$w_i \in \{0,1\}, \quad i = 1,\dots,\ell$$

Bayesian information matrix

$$\mathcal{I}(w) = V_\theta^{-1} + \sum_{i=1}^{\ell} w_i \Upsilon_i$$

$$\Upsilon_i = \frac{1}{\sigma^2}\sum_{j=1}^{N} X(t_j)^\top \pi_i\,\pi_i^\top X(t_j), \quad i = 1,\dots,\ell$$

Relaxed formulation

We allow the w_i's to take any values in $[0, 1]$, not only 0 or 1.

Relaxed sensor selection problem

Find a vector $w^\star \in \mathbb{R}^\ell$ to minimize

$$\mathcal{J}(w) = \log\det\big(Q\,\mathcal{I}(w)^{-1}Q^\top\big)$$

subject to

$$\begin{aligned} \mathbf{1}_\ell^\top w &= m \\ 0 \le w_i \le 1, &\quad i = 1, \dots, \ell \end{aligned}$$

This performance index is convex and the set of feasible solutions is the intersection of a hyperplane and a box, i.e., it is a polyhedral set. Simplicial decomposition is ideally suited to its numerical solution.

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.

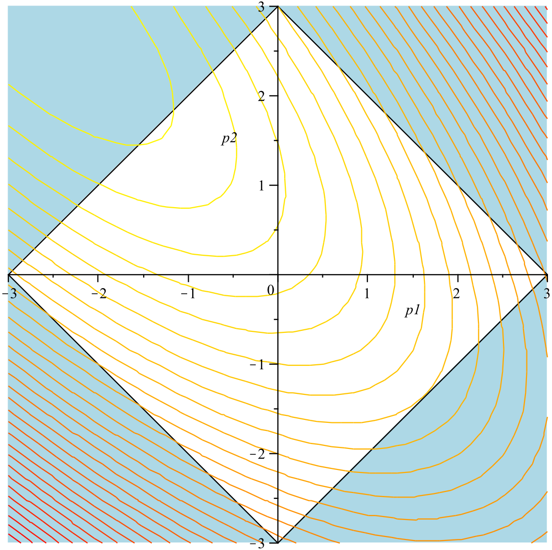
Simplicial decomposition

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

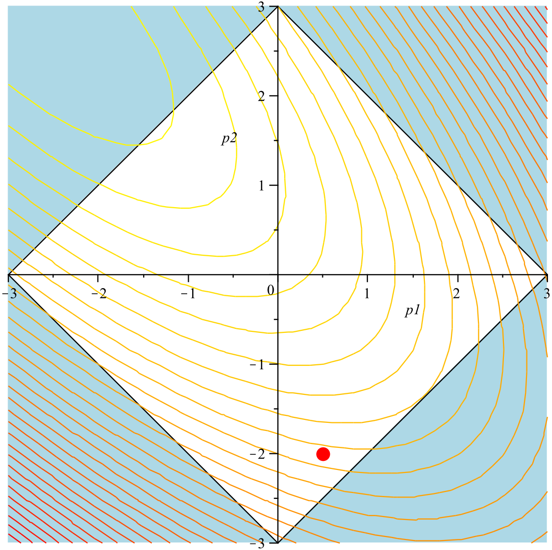
1. a linear programming subproblem (the so-called *column generation problem*) which generates an extreme point of the polyhedron, and
2. a nonlinear *restricted master problem* (RMP) which finds the maximum of the objective function over the convex hull (a simplex) of previously defined extreme points.

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.

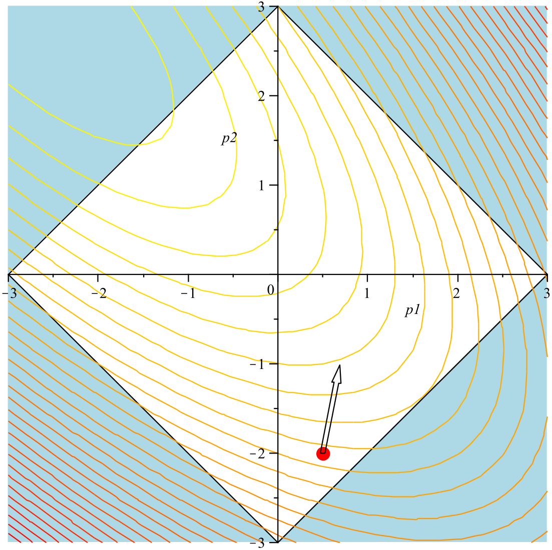
Simplicial decomposition



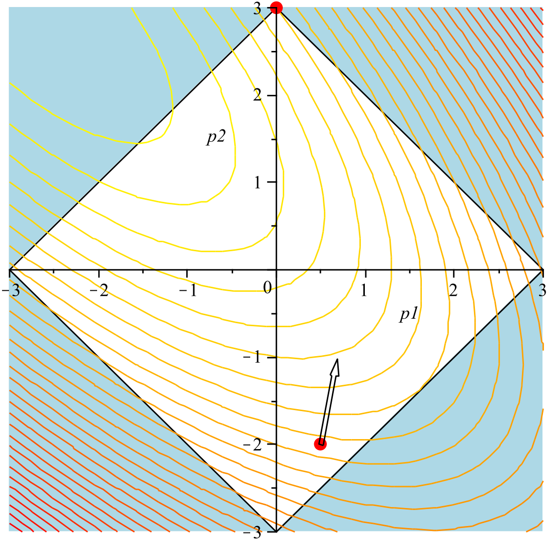
Simplicial decomposition



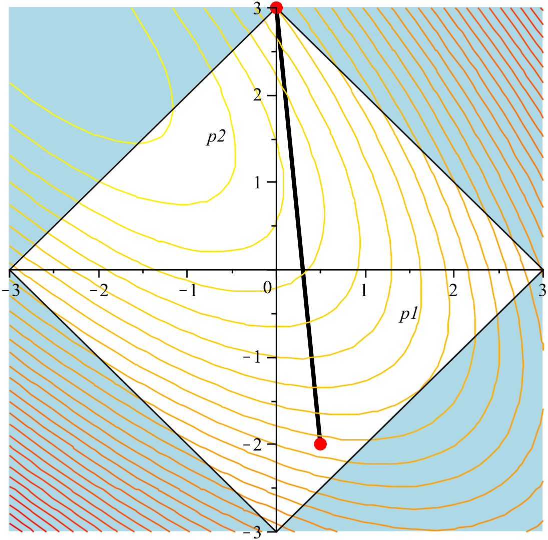
Simplicial decomposition



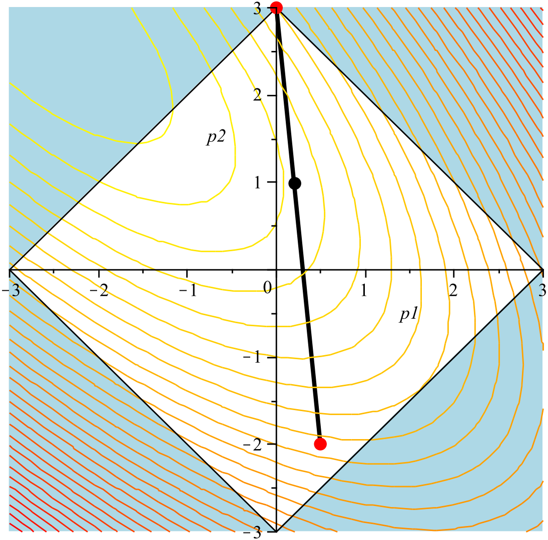
Simplicial decomposition



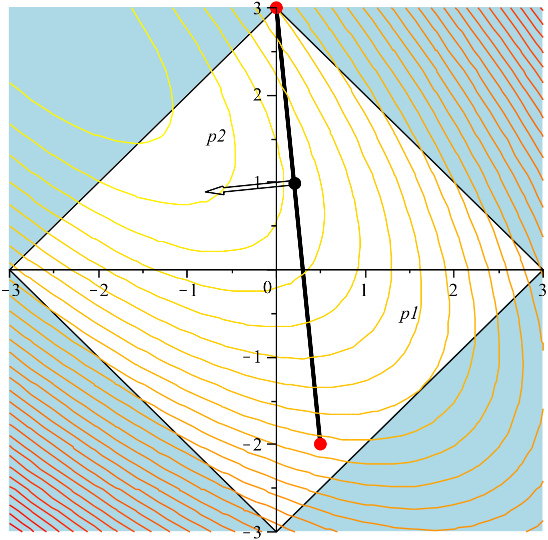
Simplicial decomposition



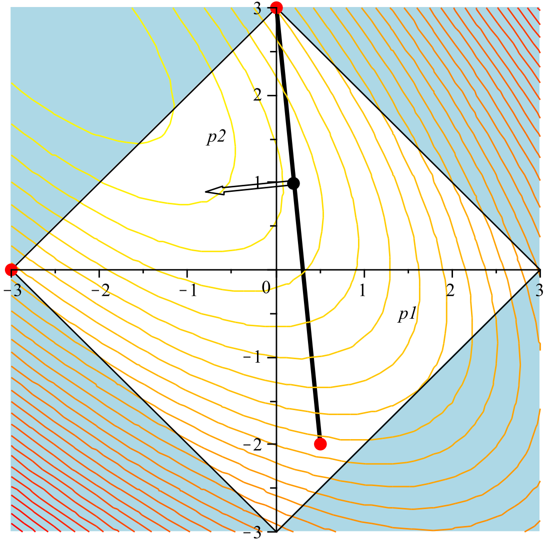
Simplicial decomposition



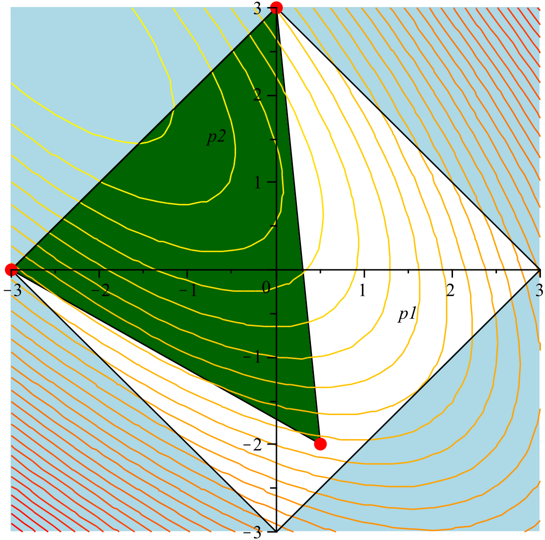
Simplicial decomposition



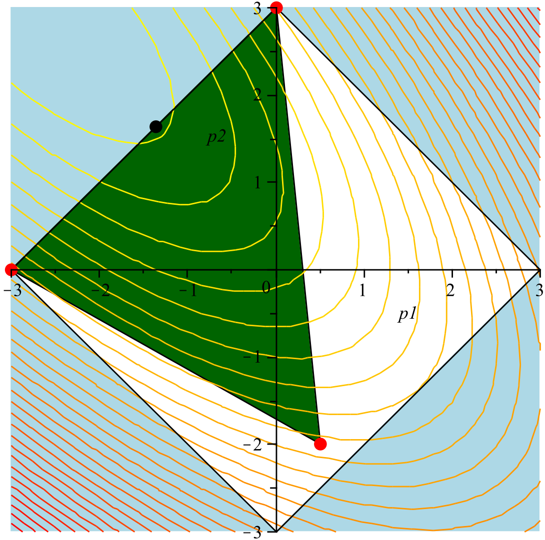
Simplicial decomposition



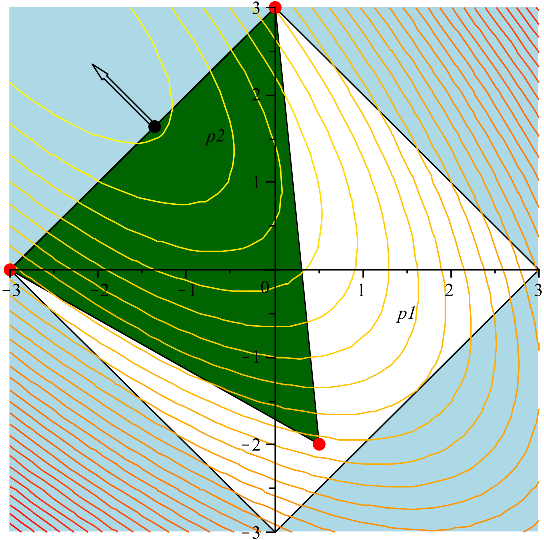
Simplicial decomposition



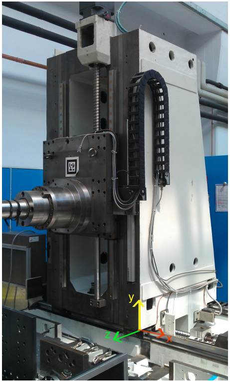
Simplicial decomposition



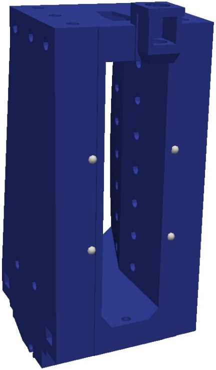
Simplicial decomposition



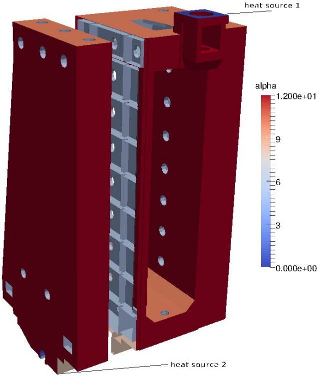
Numerical example



Machine column



CAD model with mounting points determining the TCP location



Background values of α^{bg}

Thermal model

Symbol	Meaning	Value	Units
T	temperature		K
r	thermal surface load		$\mathrm{W\,m^{-2}}$
ρ	density	7 250	$\mathrm{kg\,m^{-3}}$
c_p	specific heat at const. pressure	500	$\mathrm{J\,kg^{-1}\,K^{-1}}$
λ	thermal conductivity	46.8	$\mathrm{W\,K^{-1}\,m^{-1}}$
T_{ref}	ambient temperature	20	°C
α^{bg}	background information on α	0 to 12	$\mathrm{W\,K^{-1}\,m^{-2}}$
α	heat transfer coefficient	unknown	$\mathrm{W\,K^{-1}\,m^{-2}}$
T_0	initial temperature	unknown	K

Displacement model

Symbol	Meaning	Value	Units
$\boldsymbol{u}$	displacement		m
$\boldsymbol{\sigma}$	stress		$\mathrm{N\,m^{-2}}$
$\boldsymbol{\varepsilon}$	strain		1
ν	Poisson's ratio	0.3	1
E	modulus of elasticity	$114 \cdot 10^9$	$\mathrm{N\,m^{-2}}$
β	thermal volumetric expansion coeff.	$1.1 \cdot 10^{-5}$	$\mathrm{K^{-1}}$
L	length of the main spindle	0.993	m
σ	standard deviation of sensor noise	0.0333	K

Problem data

$$\alpha^0(\chi) = \begin{cases} 12\,\mathsf{W}\,\mathsf{K}^{-1}\,\mathsf{m}^{-2} & \text{if } \chi \in \Gamma_{\mathsf{vert}} \text{ (vertical surfaces)} \\ 10\,\mathsf{W}\,\mathsf{K}^{-1}\,\mathsf{m}^{-2} & \text{if } \chi \in \Gamma_{\mathsf{up}} \text{ (horiz. surfaces facing up)} \\ 8\,\mathsf{W}\,\mathsf{K}^{-1}\,\mathsf{m}^{-2} & \text{if } \chi \in \Gamma_{\mathsf{down}} \text{ (horiz. surf-s facing down)} \\ 5\,\mathsf{W}\,\mathsf{K}^{-1}\,\mathsf{m}^{-2} & \text{if } \chi \in \Gamma_{\mathsf{inner}} \text{ (enclosed surfaces)} \\ 0\,\mathsf{W}\,\mathsf{K}^{-1}\,\mathsf{m}^{-2} & \text{if } \chi \in \Gamma_{\mathsf{r1}} \cup \Gamma_{\mathsf{r2}} \text{ (surf. with heat sources)} \end{cases}$$

$T_0^0(\chi)$ was set as T_{ref}. The inverse covariance matrices for the parameter and for the initial state were chosen as $V_p^{-1} = \mathsf{id}_4$ and $V_{x_0}^{-1}$ for $\beta = \gamma = 1$.

$$r(\chi, t) = \begin{cases} 6700\ \mathsf{W}\,\mathsf{m}^{-2} & \text{if } \chi \in \Gamma_{r1} \text{ and } 0\ \mathsf{s} \le t \le 2400\ \mathsf{s} \\ 2700\ \mathsf{W}\,\mathsf{m}^{-2} & \text{if } \chi \in \Gamma_{r2} \text{ and } 0\ \mathsf{s} \le t \le 4800\ \mathsf{s} \\ 6700\ \mathsf{W}\,\mathsf{m}^{-2} & \text{if } \chi \in \Gamma_{r1} \text{ and } 4800\ \mathsf{s} < t \le 7200\ \mathsf{s} \\ 0 & \text{otherwise} \end{cases}$$

Algorithmic data

There are $m = 10$ sensors to be located.

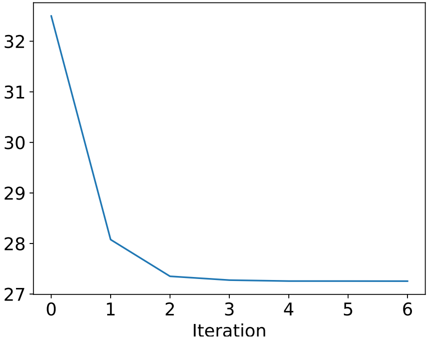
total number of mesh nodes	number of mesh cells	number of nodes on the boundary (potential sensor locations)
$n = 25\,615$	79 197	$\ell = 25\,288$

Computations were implemented on an Intel Xeon workstation with a 2.4 GHz CPU using the open-source finite element package FEniCS 2017.1.

time for computation of sensitivities	≈ 15 min
number of SDP steps	6
total time	2.5 h

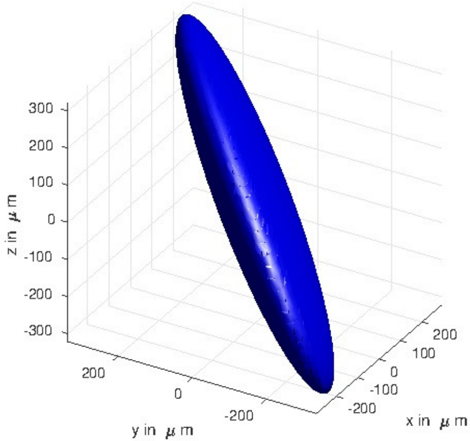
Results

(a) Optimal sensor positions ($m = 10$).



(b) Objective values vs iteration number.

Results: 90% confidence ellipsoid



Concluding remarks: Ongoing work

- The sensitivities $X(t_j) = \left[\ X_0(t_j)\ \middle|\ X_p(t_j)\ \right]$ may strongly depend on the preliminary parameter/state estimates $\theta^0 = (x_0^0, p^0)$. This is unsatisfactory when the DA problem is considered in a moving horizon context where updated estimates on the unknowns become continually available, but changes in sensor locations during the machine operation are impossible to realize.
- The method has a tendency to select candidate locations which are close to one another. Basically, this could be circumvented by imposing a minimal allowable distance constraint for gauged sites or a judicious design of the set of candidate sites, but this requires a thorough formulation.
- Yet another issue is the proper use of parametric model order reduction to reduce the overall computational effort.