UNIVERSITY OF BIRMINGHAM

On the use of Domain Decomposition Techniques in Structural Optimization

Michal Kočvara

School of Mathematics, The University of Birmingham

with D. Loghin and J. Turner (Birmingham)
M. Kojima (Tokyo)

Advances in Large Scale Optimization
Edinburgh, May 2012

Structural optimization

The goal is to improve behavior of a mechanical structure while keeping its structural properties.

Objectives/constraints:
weight, stiffness, vibration modes, stability,stress

Control variables:
thickness/density (VTS/SIMP)
material properties (FMO)

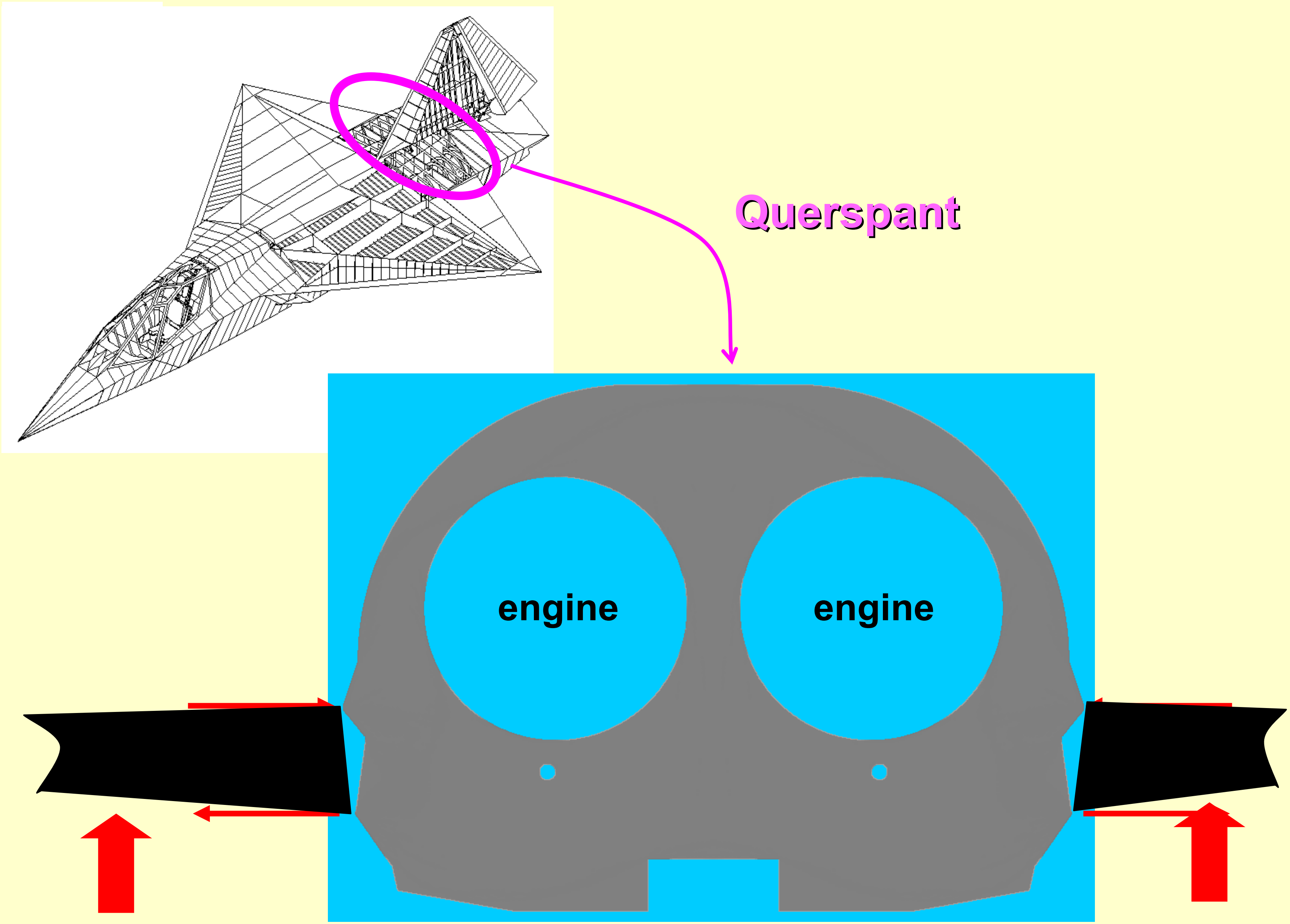
Topology optimization

$E(x) = \rho(x) E_0$ with $0 \le \underline{\rho} \le \rho(x) \le \overline{\rho}$ ... topology optimization (SIMP, VTS)

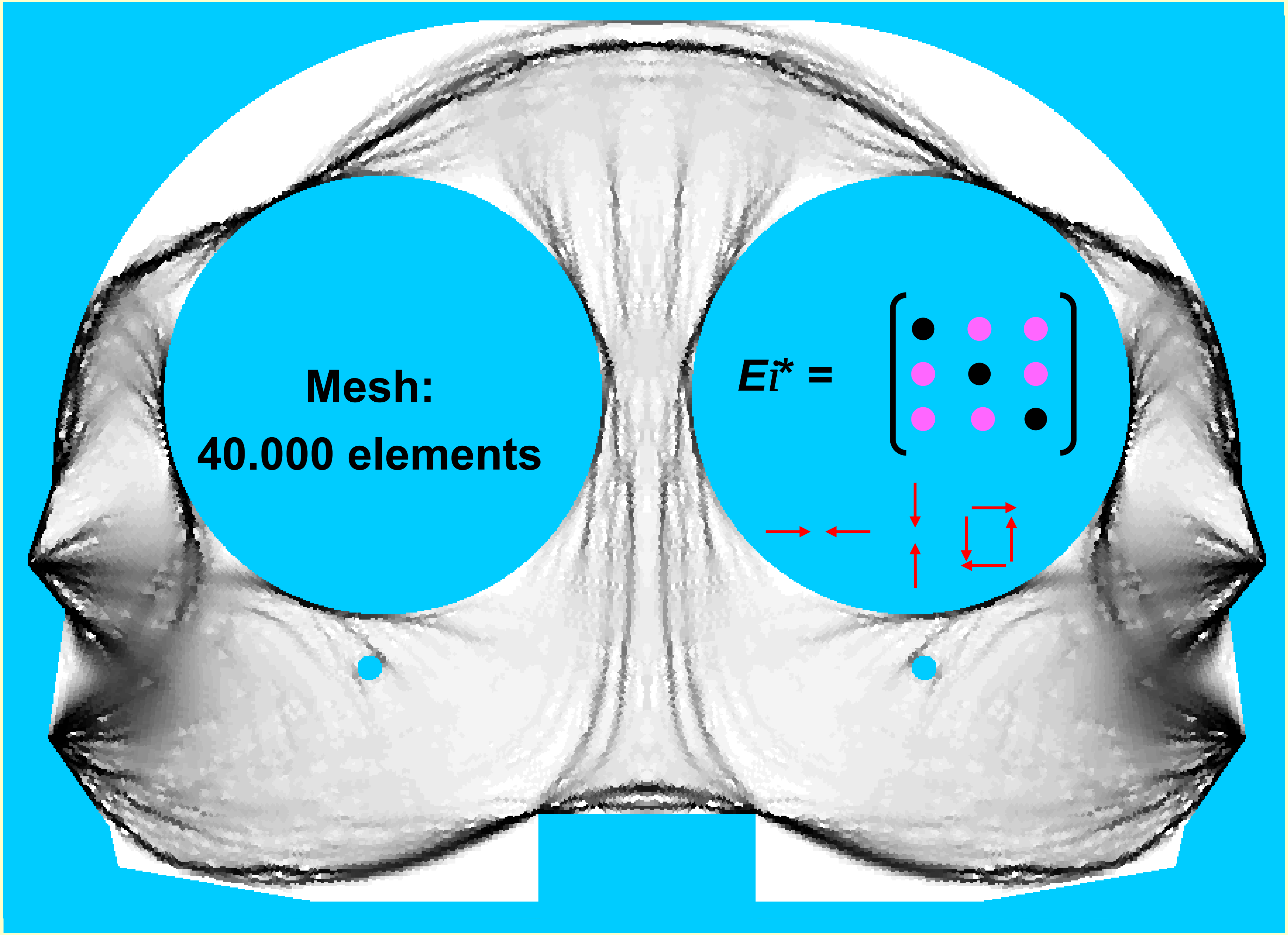
E_0 a given (homogeneous, isotropic) material

Aim:

Given an amount of material, boundary conditions and external load f, find the material distribution so that the body is as stiff as possible under f.



Querspant
engine
engine


Mesh:
40.000 elements
Ei* =

Equilibrium

Equilibrium equation:

$$K(\rho)u = f, \qquad K(\rho) = \sum_{i=1}^{m} K_i := \sum_{i=1}^{m}\sum_{j=1}^{G} B_{i,j}\rho_i E_0 B_{i,j}^\top$$

Standard finite element discretization:

Quadrilateral elements

$\rho \ldots$ piece-wise constant

$u \ldots$ piece-wise bilinear (tri-linear)

TO primal formulation

$$\min_{\rho\in\mathbb{R}^m,\ u\in\mathbb{R}^n} f^T u$$

subject to

$$\underline{\rho} \le \rho_i \le \overline{\rho}, \quad i = 1,\dots,m$$

$$\sum_{i=1}^m \rho_i \le 1$$

$$K(\rho)u = f$$

…large-scale nonlinear non-convex problem

Reduced primal formulation

Using $u = K(\rho)^{-1} f$ $\quad (\underline{\rho} > 0!)$ we obtain

$$\min_{\rho \in \mathbb{R}^m} f^\top K(\rho)^{-1} f$$

subject to

$$\underline{\rho} \le \rho_i \le \overline{\rho}, \quad i = 1, \dots, m$$

$$\sum_{i=1}^{m} \rho_i \le 1$$

…large-scale nonlinear convex problem

Reduced primal formulation

$$\min_{\rho\in\mathbb{R}^m} f^\top K(\rho)^{-1} f$$

subject to

$$\underline{\rho} \le \rho_i \le \overline{\rho}, \quad i = 1, \dots, m$$

$$\sum_{i=1}^{m} \rho_i \le 1$$

Observation: The finite element mesh is very fine but (often) regular.

Can we use techniques like multigrid or domain decomposition?

Domain decomposition for TO, “simple” approach

Solve (TO) by first or second-order method
(MMA or IPOPT/PENNON):

in each iteration we have to solve a linear system

$$Hd = g$$

where H is either the Hessian or the stiffness matrix.

Solve this system by multigrid/domain decomposition (Maar-Schultz, 2000)

We’d like to solve the “full” constrained optimization problem by MGM/DD

DD for TO, “simple” approach extended

UNIVERSITY OF BIRMINGHAM

Work with the original TO formulation

$$\min_{\rho\in\mathbb{R}^m,\ u\in\mathbb{R}^n} f^T u$$

subject to

$$0 \le \rho_i \le \overline{\rho}, \quad i = 1,\dots,m$$

$$\sum_{i=1}^m \rho_i \le 1$$

$$K(\rho)u = f$$

write down KKT conditions, perturb and apply Newton (Interior Point)

solve the Newton system by GMRES-DD

DD for TO, “simple” approach extended

KKT:

$$-\mathrm{Res}^{(1)} := K(\rho)u - f = 0$$

$$-\mathrm{Res}^{(2)} := \sum_{i=1}^{m} \rho_i - 1 = 0$$

$$-\mathrm{Res}^{(3)} := -\frac{1}{2}u^T K_i u - \lambda - \varphi_i + \psi_i = 0, \quad i = 1, \dots, m$$

$$\varphi_i \rho_i = 0, \quad i = 1, \dots, m$$

$$\psi_i(\overline{\rho} - \rho_i) = 0, \quad i = 1, \dots, m$$

$$\rho_i \geq 0, \quad \overline{\rho} - \rho_i \geq 0, \quad \varphi_i \geq 0, \quad \psi \geq 0$$

We will perturb the complementarity constraints by “penalty” parameters $s, r > 0$:

$$-\mathrm{Res}^{(4)} := \varphi_i \rho_i - s = 0, \quad i = 1, \dots, m$$

$$-\mathrm{Res}^{(5)} := \psi_i(\overline{\rho} - \rho_i) - r = 0, \quad i = 1, \dots, m$$

DD for TO, "simple" approach extended

In every step of the Newton method solve

$$\begin{bmatrix} K(\rho) & 0 & B(u) & 0 & 0 \\ 0 & 0 & e^T & 0 & 0 \\ B(u)^T & e & 0 & I & -I \\ 0 & 0 & \Phi & X & 0 \\ 0 & 0 & \Psi & 0 & \widetilde{X} \end{bmatrix} \begin{bmatrix} d_u \\ d_\lambda \\ d_x \\ d_\varphi \\ d_\psi \end{bmatrix} = \begin{bmatrix} \mathrm{Res}^{(1)} \\ \mathrm{Res}^{(2)} \\ \mathrm{Res}^{(3)} \\ \mathrm{Res}^{(4)} \\ \mathrm{Res}^{(5)} \end{bmatrix}.$$

Here $B(u) = (K_1 u, K_2 u, \ldots, K_m u)$, e is a vector of all ones and

$$X = \mathrm{diag}(\rho), \quad \widetilde{X} = \mathrm{diag}(\overline{\rho} - \rho), \quad \Phi = \mathrm{diag}(\varphi), \quad \Psi = \mathrm{diag}(\psi)$$

DD for TO, "simple" approach extended

Reduce to the Schur complement

$$Z \begin{bmatrix} d_u \\ d_\lambda \end{bmatrix} = \mathrm{Res}^{(Z)} \, ,$$

with

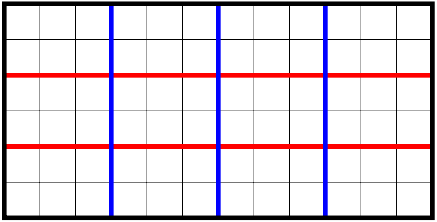
$$Z = \begin{bmatrix} K(\rho) & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} B(u) \\ e^T \end{bmatrix} (X^{-1}\Phi + \widetilde{X}^{-1}\Psi)^{-1} \begin{bmatrix} B(u)^T & e \end{bmatrix}$$

and

$$\mathrm{Res}^{(Z)} = \begin{bmatrix} \mathrm{Res}^{(1)} \\ \mathrm{Res}^{(2)} \end{bmatrix} + \begin{bmatrix} B(u) \\ e^T \end{bmatrix} (X^{-1}\Phi + \widetilde{X}^{-1}\Psi)^{-1} \widetilde{\mathrm{Res}}^{(3)} \, .$$

Structure of Z = structure of $K(\rho)$ plus one full column/row

DD for TO, “simple” approach extended



Z has block diagonal structure + boundary strip:

$$\begin{pmatrix} Z_{1,1} & & & & Z_{1,\Gamma} \\ & Z_{2,2} & & & Z_{1,\Gamma} \\ & & \ddots & & \vdots \\ & & & Z_{k,k} & Z_{k,\Gamma} \\ Z_{\Gamma,1} & Z_{\Gamma,2} & \cdots & Z_{\Gamma,k} & Z_{\Gamma,\Gamma} \end{pmatrix}$$

Schur complement solved by preconditioned GMRES, small systems with $Z_{i,i}$ solved directly (in parallel (in the future))

$H_{1/2}$ preconditioner (Arioli-Loghin, 2009) ($H_{1/2} = L_0(L_0^{-1}L_1)^{1/2}$)

DD for TO, “simple” approach extended

UNIVERSITY OF BIRMINGHAM

Interior point method:

Set initial values. Do until convergence:

1. Solve system $Zd = \mathrm{Res}^{(Z)}$
2. Find the step length α
3. Update the solution
$$z = z + \alpha d, \quad z = (u, \lambda, \rho, \varphi, \psi)^T$$
4. Update the penalty parameters
$$s = s/3, \quad r = r/3$$
and return to Step 1.

Linesearch:

$$\alpha = \min\{\alpha_l, \alpha_u, 1\}\,.$$

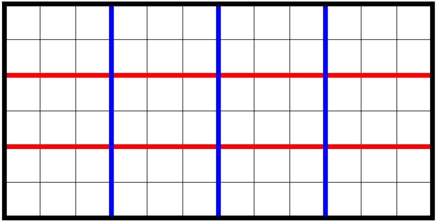
$$\alpha_l = 0.9 \cdot \min_{i:(d_\rho)_i<0} \{-\frac{\rho_i}{(d_\rho)_i}\}, \quad \alpha_u = 0.9 \cdot \min_{i:(d_\rho)_i>0} \{\frac{\overline{\rho} - \rho_i}{(d_\rho)_i}\}\,.$$

Numerical experiments

no of doms	no of elems	IP steps	Nwt steps	aver. GMRES per Nwt
4×4	128	6	15	26
4×4	512	6	18	26
4×4	2048	6	19	26
4×4	8192	6	21	20
4×4	32768	6	25	17

no of doms	no of elems	IP steps	Nwt steps	aver. GMRES per Nwt
2×2	8192	6	21	11
4×4	8192	6	21	20
8×8	8192	6	22	35
16×16	8192	6	23	49

“Full” DD for TO



Subdomains $1, 2, \ldots, d$, vector of variables

$$\rho = (\rho^{(1)}, \rho^{(2)}, \ldots, \rho^{(d)}) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \ldots \times \mathbb{R}^{m_d}$$

Idea

For $i = 1, \ldots, d$:

Solve TO w.r.t. $\rho^{(i)}$, keeping the other variables $\rho^{(j)}, j \neq i$, fixed

DD for constrained optimization

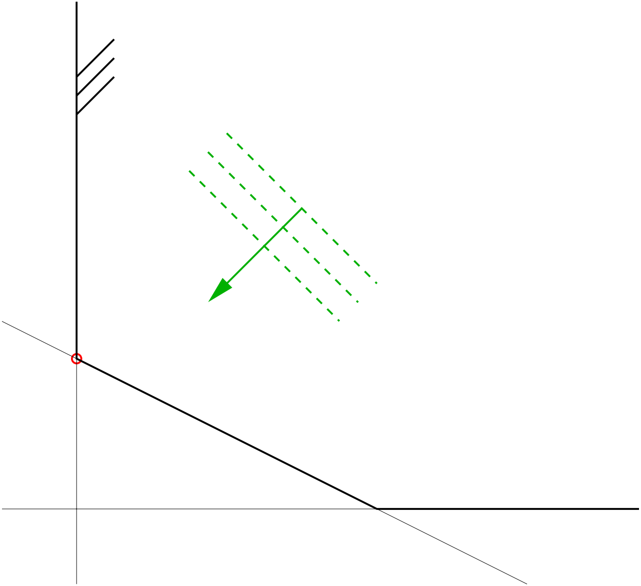
UNIVERSITY OF BIRMINGHAM

Is this a good idea? NOT, in general!

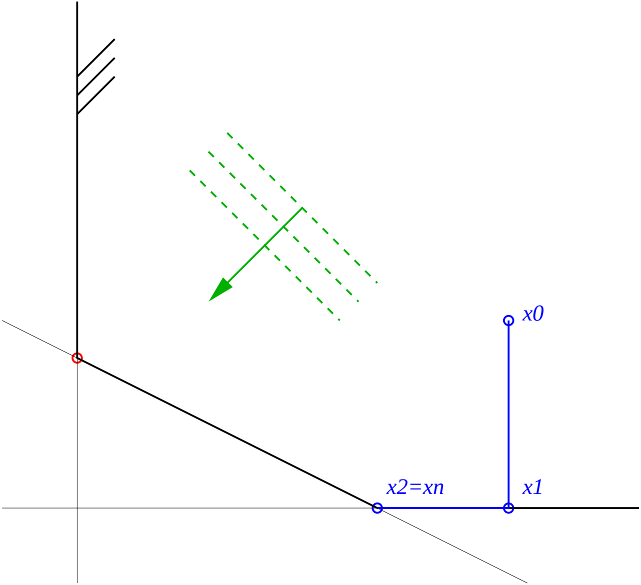
It is well-known that block Gauss-Seidel method does not work for constrained optimization problems (with non-separable constraints).

We can get stuck at a boundary point, far from any critical point.

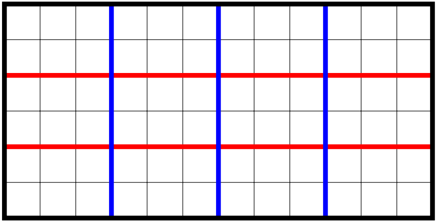
LP example



LP example



"Full" DD for TO



Subdomains $1, 2, \ldots, d$, vector of variables

$$\rho = (\rho^{(1)}, \rho^{(2)}, \ldots, \rho^{(d)}) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \ldots \times \mathbb{R}^{m_d}$$

Idea

1. For $i = 1, \ldots, d$:
   Solve TO w.r.t. $\rho^{(i)}$, keeping the other variables $\rho^{(j)}, j \neq i$, fixed
2. Run one "correction" step

Optimality condition method

Optimality conditions for TO: (for $\underline{\rho} \le \rho_i$)

$$
\begin{aligned}
K(\rho)u &= f \\
u^T K_i u - \lambda + \mu_i &= 0 \\
\lambda(\sum \rho_i - 1) &= 0 \\
\mu_i(\rho_i - \underline{\rho}) &= 0 \\
\lambda \ge 0,\ \mu_i &\ge 0
\end{aligned}
$$

OC iterative method:

1. Given ρ, find u s.t. $K(\rho)u = f$
2. Compute λ s.t. $\sum \max\left\{\dfrac{\rho_i u^T K_i u}{\lambda}, \underline{\rho}\right\} = 1$
3. Update ρ by $\rho_i^{\text{new}} = \max\left\{\frac{\rho_i u^T K_i u}{\lambda}, \underline{\rho}\right\}$

OC method converges but is slow.

A two-step DD algorithm

UNIVERSITY OF BIRMINGHAM

1. Run 5 steps of the OC method
2. Repeat till convergence:
2a. One step of symmetric block Gauss-Seidel
2b. Two steps of OC

OC works like a “centering step”, gets us away from the “bad” points.

The idea works:

Convergence after 15–30 iterations of the two-step method
depending on the number of sub-domains

SDP formulation of TO by DD

The TO problem

$$\min_{\rho\in\mathbb{R}^m} f^\top K(\rho)^{-1} f$$

subject to

$$\sum \rho_i \leq 1, \quad \underline{\rho} \leq \rho_i \leq \overline{\rho}, \quad i = 1,\dots,m$$

can be equivalently formulated as a linear SDP:

$$\min_{\rho\in\mathbb{R}^m,\, \gamma\in\mathbb{R}} \gamma$$

subject to

$$\begin{pmatrix} \gamma & f^T \\ f & K(\rho) \end{pmatrix} \succeq 0$$

$$\sum \rho_i \leq 1, \quad \underline{\rho} \leq \rho_i \leq \overline{\rho}, \quad i = 1,\dots,m\,.$$

Helpful when vibration/buckling constraints present

SDP formulation of TO by DD

$$\begin{pmatrix} \gamma & f^T \\ f & \sum \rho_i K_i \end{pmatrix} \succeq 0$$

is a large matrix constraint dependent on many variables. . . bad for existing SDP software

Can we replace it by several smaller constraints equivalently?

Range-space sparsity

S. Kim, M. Kojima, M. Mevissen and M. Yamashita, Exploiting Sparsity in Linear and Nonlinear Matrix Inequalities via Positive Semidefinite Matrix Completion, Mathematical Programming, 2011

Based on:

A. Griewank and Ph. Toint, On the existence of convex decompositions of partially separable functions, MPA 28, 1984

J. Agler, W. Helton, S. McCulough and L. Rodnan, Positive semidefinite matrices with a given sparsity pattern, LAA 107, 1988

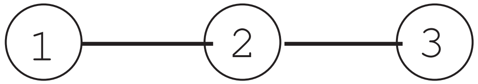
$G(N,E)$: a chordal graph with $N = \{1,\ldots,n\}$ and the max. cliques of $C_1,\ldots,C_\ell$. $E^\bullet = E \cup \{(i,i) : i \in N\}$.

$\mathbb{S}^n(E^\bullet) = \{\boldsymbol{Y} \in \mathbb{S}^n : Y_{ij} = 0\ (i,j) \notin E^\bullet\}$.

$\mathbb{S}^C_+ = \{\boldsymbol{Y} \succeq \boldsymbol{O} : Y_{ij} = 0$ if $(i,j) \notin C \times C\}$ for $\forall C \subseteq N$.

Theorem (Agler, Helton, McCulough and Rodman 1988)
Suppose $\boldsymbol{M} \in \mathbb{S}^n(E^\bullet)$. $\boldsymbol{M} \succeq \boldsymbol{O}$ iff

$$\boldsymbol{M} = \boldsymbol{Y}^1 + \boldsymbol{Y}^2 + \cdots + \boldsymbol{Y}^\ell \text{ for } \exists \boldsymbol{Y}^k \in \mathbb{S}^{C_k}_+ \ (k = 1,\ldots,\ell).$$



$C_1 = \{1,2\},\ C_2 = \{2,3\}$. $\boldsymbol{M} : \mathbb{R}^m \to \mathbb{S}^3(E^\bullet)$.

$$\boldsymbol{M}(\boldsymbol{u}) = \begin{pmatrix} M_{11}(\boldsymbol{u}) & M_{12}(\boldsymbol{u}) & 0 \\ M_{21}(\boldsymbol{u}) & M_{22}(\boldsymbol{u}) & M_{23}(\boldsymbol{u}) \\ 0 & M_{32}(\boldsymbol{u}) & M_{33}(\boldsymbol{u}) \end{pmatrix}$$

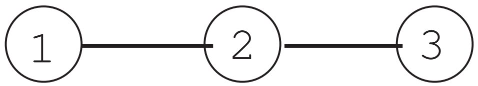
$G(N,E)$: a chordal graph with $N=\{1,\ldots,n\}$ and the max. cliques of $C_1,\ldots,C_\ell$. $E^\bullet = E \cup \{(i,i) : i \in N\}$.

$\mathbb{S}^n(E^\bullet) = \{\boldsymbol{Y} \in \mathbb{S}^n : Y_{ij} = 0 \ (i,j) \notin E^\bullet\}$.

$\mathbb{S}^C_+ = \{\boldsymbol{Y} \succeq \boldsymbol{O} : Y_{ij} = 0$ if $(i,j) \notin C \times C\}$ for $\forall C \subseteq N$.

Theorem (Agler, Helton, McCulough and Rodman 1988)
Suppose $\boldsymbol{M} \in \mathbb{S}^n(E^\bullet)$. $\boldsymbol{M} \succeq \boldsymbol{O}$ iff

$$\boldsymbol{M} = \boldsymbol{Y}^1 + \boldsymbol{Y}^2 + \cdots + \boldsymbol{Y}^\ell \text{ for } \exists \boldsymbol{Y}^k \in \mathbb{S}^{C_k}_+ \ (k=1,\ldots,\ell).$$



$C_1 = \{1,2\},\ C_2 = \{2,3\}$. $\boldsymbol{M} : \mathbb{R}^m \to \mathbb{S}^3(E^\bullet)$.

$$\boldsymbol{M}(\boldsymbol{u}) = \begin{pmatrix} Y^1_{11} & Y^1_{12} & 0 \\ Y^1_{12} & Y^1_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & Y^2_{22} & Y^2_{23} \\ 0 & Y^2_{32} & Y^2_{33} \end{pmatrix}$$

$$\boldsymbol{M}(\boldsymbol{u}) \succeq \boldsymbol{O}$$
$$\Updownarrow$$

$$\left.\begin{array}{l} M_{11} = Y^1_{11},\ M_{12} = Y^1_{12}, \\ M_{22} = Y^1_{22} + Y^2_{22}, \\ M_{23} = Y^2_{23},\ M_{33} = Y^2_{33}, \\ \square \succeq \boldsymbol{O},\ \square \succeq \boldsymbol{O} \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} \begin{pmatrix} M_{11}(\boldsymbol{u}) & M_{12}(\boldsymbol{u}) \\ M_{21}(\boldsymbol{u}) & Y^1_{22} \end{pmatrix} \succeq \boldsymbol{O}, \\ \begin{pmatrix} M_{22}(\boldsymbol{u}) - Y^1_{22} & M_{23}(\boldsymbol{u}) \\ M_{32}(\boldsymbol{u}) & M_{33}(\boldsymbol{u}) \end{pmatrix} \succeq \boldsymbol{O} \end{array}\right.$$

Range-space sparsity

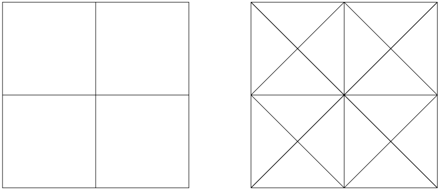
S. Kim, M. Kojima, M. Mevissen and M. Yamashita, Exploiting Sparsity in Linear and Nonlinear Matrix Inequalities via Positive Semidefinite Matrix Completion, Mathematical Programming, 2011

For a general LMI $A \succeq 0$:
1. Check whether the graph associated with A is chordal. If not, extend it.
2. Find the maximum cliques of the (extended) graph
3. Use the above theorem

Doesn't really work for TO

Range-space sparsity in TO

TO: the graph of $K(\rho)$ is given by the finite element discretization. The maximum cliques are known!

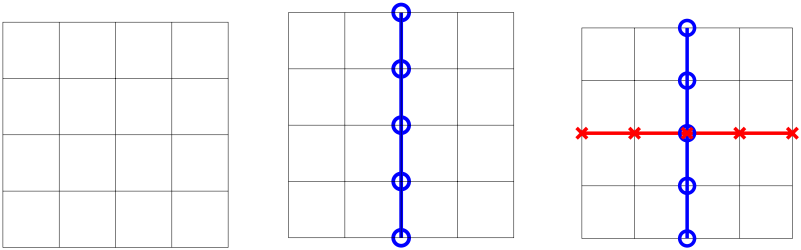


Finite element mesh and the graph of the associated stiffness matrix

BUT: Graph of $K(\rho)$ is not chordal.

We will define its chordal extension using the sub-domains

Hierarchical Type II decomposition



Hierarchical Type II decomposition:
four matrix constraints and three new matrix variables of sizes 5×5 (level 1), 3×3 and 3×3 (level 2)

Open question

UNIVERSITY OF BIRMINGHAM

$$K = \begin{pmatrix} K^{(1)}_{1,1} & K^{(1)}_{1,2} & 0 \\ K^{(1)}_{2,1} & K^{(1)}_{2,2} + K^{(2)}_{1,1} & K^{(2)}_{1,2} \\ 0 & K^{(2)}_{2,1} & K^{(2)}_{2,2} \end{pmatrix} \succcurlyeq 0$$

replaced by

$$\begin{pmatrix} K^{(1)}_{1,1} & K^{(1)}_{1,2} \\ K^{(1)}_{2,1} & S \end{pmatrix} \succcurlyeq 0, \qquad \begin{pmatrix} K^{(1)}_{2,2} + K^{(2)}_{1,1} - S & K^{(2)}_{1,2} \\ K^{(2)}_{2,1} & K^{(2)}_{2,2} \end{pmatrix} \succcurlyeq 0$$

Chordal extension $\Rightarrow$ variable S should be dense.

But we get the same results with S assumed sparse!

Hypothesis: A "strengthened/modified" version of the theorem applies in our case.

Numerical experiments

UNIVERSITY OF BIRMINGHAM

Type II decomposition, 80x40 elements, PENSDP

no of doms	no of vars	size of matrix	Nwt steps	CPU total	CPU per iter	CPU/iter Hess	CPU/iter Chol
1	3200	6560	150	15838	105.6	104.6	0.84
2	3483	3362	158	8226	52.6	49.4	2.66
8	4614	882	151	1851	12.3	11.3	0.89
32	6912	242	166	612	3.7	3.0	0.70
128	11652	72	203	338	1.7	0.8	0.86
200	14094	50	210	262	1.3	0.4	0.88
800	27024	18	234	355	1.5	0.2	1.29

Numerical experiments

UNIVERSITY OF BIRMINGHAM

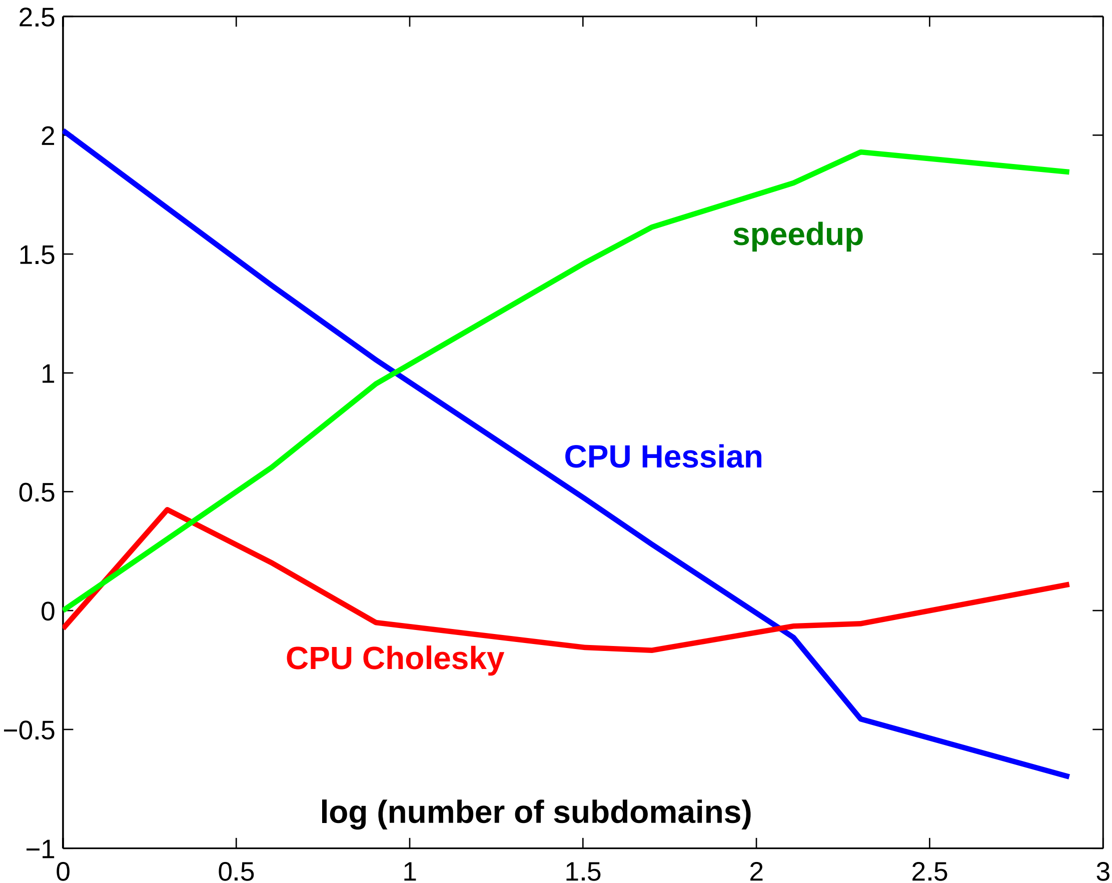
Type II decomposition, 80x40 elements, PENSDP

no of doms	no of vars	size of matrix	Nwt steps	CPU total	CPU per iter	CPU/iter Hess	CPU/iter Chol
1	3200	6560	150	15838	105.6	104.6	0.84
2	3483	3362	158	8226	52.6	49.4	2.66
8	4614	882	151	1851	12.3	11.3	0.89
32	6912	242	166	612	3.7	3.0	0.70
128	11652	72	203	338	1.7	0.8	0.86
200	14094	50	210	262	1.3	0.4	0.88
800	27024	18	234	355	1.5	0.2	1.29

Numerical experiments

Type II decomposition, 80x40 elements, PENSDP

no of doms	no of vars	size of matrix	Nwt steps	CPU total	CPU per iter	speedup total	speedup /iter	
1	3200	6560	150	15838	105.6	1	1	1
2	3483	3362	158	8226	52.6	2	2	2
8	4614	882	151	1851	12.3	9	9	7
32	6912	242	166	612	3.7	26	29	27
128	11652	72	203	338	1.7	47	63	91
200	14094	50	210	262	1.3	60	85	131
800	27024	18	234	355	1.5	45	70	364


2.5
2
1.5
1
0.5
0
−0.5
−1
0
0.5
1
1.5
2
2.5
3
speedup
CPU Hessian
CPU Cholesky
log (number of subdomains)

Numerical experiments

Type II decomposition

"best" decomposition speedup (200–400 matrices)

problem	ORIGINAL no of vars	ORIGINAL size of matrix	ORIGINAL CPU total	DECOMPOSED no of vars	DECOMPOSED size of matrix	DECOMPOSED CPU total	speedup
40x20	800	1680	538	3299	50	25	22
60x30	1800	3720	3606	10104	32	112	32
80x40	3200	6560	15838	14094	50	262	60
100x50	5000	10200	44800	18484	72	687	65
120x60	7200	14640	108000	32389	50	906	119
140x70	9800	19880	227500	36852	72	1704	130

time estimated; 227500 sec = 2 days 15 hours

THE END