Dimensional reduction in topology optimization with vibration constraints

Michal Kočvara

School of Mathematics, The University of Birmingham

Durham Symposium 2017

A question by SDP software developer

. . . mostly already answered by Tobias Damm

Balanced truncation and (non-)linear matrix inequalities?

In the unreduced world of linear systems, one way to find Gramians is by solving system of linear matrix inequalities (LMI)

$$A^T P + PA \prec 0, \quad P \succ 0$$

or by solving the linear semidefinite optimization problem

$$\min_P \operatorname{trace}(P) \quad \text{s.t.} \;\; A^T P + PA + BB^T \preccurlyeq 0, \quad P \succ 0$$

by available SDP software (MOSEK, SeDuMi, PENSDP. . .)

A question by SDP software developer

In balanced truncation, working with Lyapunov inequalities (rather than equalities) can improve error bounds:

–D. Hinrichsen and A. J. Pritchard. “An improved error estimate for reduced-order models of discrete-time systems.” IEEE Transactions on Automatic Control 35.3 (1990): 317-320.

–H. Sandberg. “An extension to balanced truncation with application to structured model reduction.” IEEE Transactions on Automatic Control 55.4 (2010): 1038-1043.

by solving

$$\min_P \operatorname{trace}(P) \quad \text{s.t.} \quad \begin{array}{l} A^T P + PA + BB^T \preccurlyeq 0, \quad P \succ 0 \\ P = \operatorname{diag}(P_N, P_1, \ldots, P_q) \end{array}$$

–see alsoTobias Damm’s talk.

Are these techniques known/used/useful in model order reduction?

A comment by SDP software developer

Software also available for

–bilinear matrix inequalities (BMI) e.g. the static output feedback stabilization problem of the type

$$(A + BFC)^T P + P^T (A + BFC) \prec 0, \quad P \succ 0$$

–polynomial matrix inequalities (PMI) e.g.

$$Q_1 + x_1 x_3 Q_2 + x_2 x_4^3 Q_3 \succcurlyeq 0$$

(the above SOF can be reformulated as PMI without the large matrix variable)

–(general) nonlinear matrix inequalities

$$\mathcal{A}(x, Y) \succcurlyeq 0$$

PENBMI, PENNON, PENLAB (open source Matlab)

A comment by SDP software developer

For LMIs, SeDuMi is no longer state-of-the-art software.

Matlab's Robust Control Toolbox solver is slow.

Try
– MOSEK
or
– PENSDP with iterative solvers
or
– SDPLR, the low rank solver by Samuel Burer

using

YALMIP or direct interface (YALMIP can be slow for big problems!)

A comment by SDP software developer

SDP solver complexity (one iteration of PENSDP, augmented Lagrangian method)

Matrix assembly:
dense data matrices: $O(m^3 n + m^2 n^2)$
$\rightarrow$sparse data matrices: $O(m^3 + K^2 n^2)$ $\quad K = \max_i(\mathrm{nnz}(A_i))$
sparse matrices, iterative solver: $O(m^3 + Kn)$

Linear system solution:
$\rightarrow$dense Cholesky: $O(n^3)$
sparse Cholesky: $O(n^\kappa), 1 \le \kappa \le 3$
iterative solver: $O(n^2)$

SeDuMi sparse data: one iteration $O(m^3)$, total $O(m^{3.5})$

$$\min_{x\in\mathbb{R}^n} c^\top x \quad \text{subject to} \quad \sum_{i=1}^n x_i A_i - B \succeq 0 \quad (A_i, B \in \mathbb{R}^{m\times m})$$

The talk starts now…

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

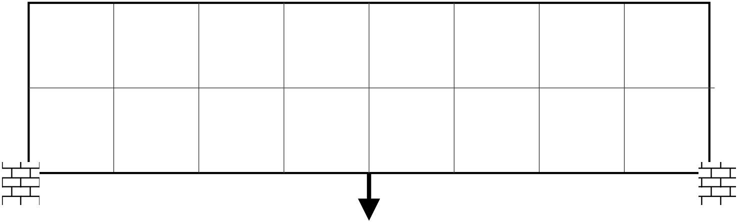
Aim:

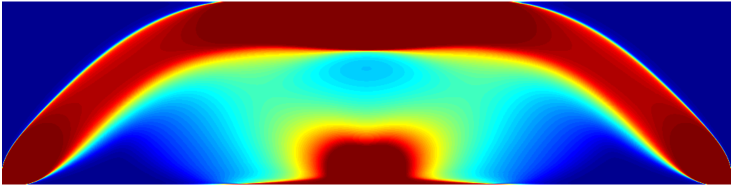
Given an amount of material, boundary conditions and external load f, find the material distribution so that the body is as stiff as possible under f.

$E(x) = \rho(x) E_0$ with $0 \leq \underline{\rho} \leq \rho(x) \leq \overline{\rho}$

E_0 a given (homogeneous, isotropic) material

Topology optimization, example





Pixels—finite elements

Color—value of variable ρ, constant on every element

Equilibrium

Equilibrium equation:

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

$$f := \sum_{i=1}^{m} f_i$$

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

$$(0 \leq)\ \underline{\rho} \leq \rho_i \leq \overline{\rho}, \quad i = 1,\ldots,m$$

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

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

. . . large-scale nonlinear non-convex problem

SDP formulation of TO

The TO problem

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

subject to

$$f^T u \le \gamma, \quad K(\rho)u = f$$

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

can be equivalently formulated as a linear SDP:

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

subject to

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

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

Helpful when vibration/buckling constraints present

SDP formulation of TO

The TO problem

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

subject to

$$f^T u \le \gamma, \quad K(\rho)u = f$$

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

can be equivalently formulated as a linear SDP:

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

subject to

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

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

Helpful when vibration/buckling constraints present

SDP formulation of TO

The TO problem

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

subject to

$$f^T u \le \gamma, \quad K(\rho)u = f$$

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

can be equivalently formulated as a linear SDP:

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

subject to

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

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

Helpful when vibration/buckling constraints present

TO with a vibration constraint

Self-vibrations of the (discretized) structure—eigenvalues of

$$K(\rho)w = \lambda M(\rho)w$$

where the mass matrix $M(\rho)$ has the same sparsity as $K(\rho)$.

Low frequencies dangerous $\rightarrow$ constraint $\lambda_{\min} \geq \hat{\lambda}$

Equivalently: $V(\hat{\lambda};\rho) := K(\rho) - \hat{\lambda}M(\rho) \succeq 0$

TO problem with vibration constraint as linear SDP:

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

subject to

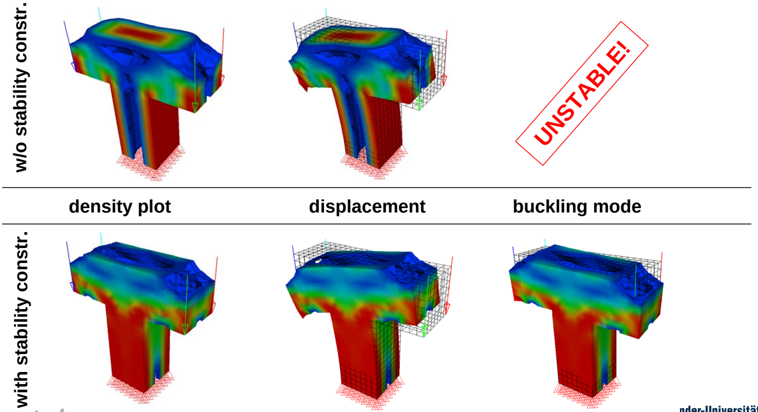
$$\begin{pmatrix} \gamma & f^T \\ f & K(\rho) \end{pmatrix} \succeq 0$$

$$V(\hat{\lambda};\rho) \succeq 0$$

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

Case studies: Tc12

50.000 design variables (sizing), 4 LC + global stability constraints



Dimensions in Semidefinite Optimization

$$\min_{x \in \mathbb{R}^n} c^\top x$$

subject to

$$\sum_{i=1}^{n} x_i A_i^{(k)} - B^{(k)} \succeq 0\,, \quad k = 1, \ldots, p$$

where

$$x \in \mathbb{R}^n, \quad A_i^{(k)},\ B^{(k)} \in \mathbb{R}^{m \times m}$$

Majority of SDP software

BAD . . . n large, m large   many variables, big matrix

OK . . . n small, m large   rare

GOOD . . . n large, m small   many variables, small matrix

GOOD . . . n large, m small, p large   many small matrix constraints

Dimensions in Semidefinite Optimization

$$\min_{x \in \mathbb{R}^n} c^\top x$$

subject to

$$\sum_{i=1}^{n} x_i A_i^{(k)} - B^{(k)} \succeq 0\,, \quad k = 1, \ldots, p$$

where

$$x \in \mathbb{R}^n, \quad A_i^{(k)},\ B^{(k)} \in \mathbb{R}^{m \times m}$$

So we may want to replace

BAD . . . n large, m large, p=1

by

GOOD . . . n large, m small, p large   many small matrix constraints

SDP formulation of TO by DD

Both

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

and

$$V(\hat{\lambda}; \rho) \succeq 0$$

are large matrix constraints dependent on many variables
...bad for existing SDP software

Can we replace them by several smaller constraints equivalently?

Chordal decomposition

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

See also:

L. Vandenberghe and M. Andersen, Chordal graphs and semidefinite optimization. Foundations and Trends in Optimization 1:241–433, 2015

Chordal decomposition

$G(N, E)$ – graph with $N = \{1, \dots, n\}$ and max. cliques $C_1, \dots, C_p$.

$\mathbb{S}^n(E) = \{Y \in \mathbb{S}^n : Y_{ij} = 0 \; (i, j) \notin E \cup \{(\ell, \ell), \; \ell \in N\}$

$\mathbb{S}_+^{C_k} = \{Y \succeq 0 : \; Y_{ij} = 0 \text{ if } (i, j) \notin C_k \times C_k\}$

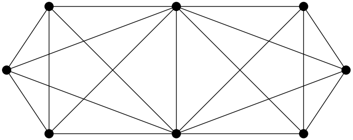
> Theorem 1: $G(N, E)$ is chordal if and only if
> for every $A \in \mathbb{S}^n(E)$, $A \succeq 0$, it holds that
> $\exists Y^k \in \mathbb{S}_+^{C_k} \; (k = 1, \dots, p)$ s.t. $A = Y^1 + Y^2 + \dots + Y^p$.

Every psd matrix is a sum of psd matrices that are non-zero only on maximal cliques.

So $A(x) \succeq 0$ replaced equivalently by $Y^k(x) \succeq 0$, $k = 1, \dots, p$.

Graph representation of matrix sparsity

Chordal sparsity graph, overlapping blocks



$$\begin{bmatrix} \times & \times & \times & \times & \times & & & \\ \times & \times & \times & \times & \times & & 0 & \\ \times & \times & \times & \times & \times & & & \\ \times & \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times & \times \\ & & & \times & \times & \times & \times & \times \\ & 0 & & \times & \times & \times & \times & \times \\ & & & \times & \times & \times & \times & \times \end{bmatrix}$$

Chordal decomposition

Theorem 1: $G(N, E)$ is chordal if and only if
for every $A \in \mathbb{S}^n(E)$, $A \succeq 0$, it holds that
$\exists Y^k \in \mathbb{S}_+^{C_k}$ $(k = 1, \ldots, p)$ s.t. $A = Y^1 + Y^2 + \ldots + Y^p$.

Let $K = \begin{pmatrix} K^{(1)}_{1,1} & K^{(1)}_{1,2} & 0 \\ K^{(1)}_{2,1} & \boxed{K^{(1)}_{2,2} + K^{(2)}_{1,1}} & K^{(2)}_{1,2} \\ 0 & K^{(2)}_{2,1} & K^{(2)}_{2,2} \end{pmatrix}$ with $K^{(1)}, K^{(2)}$ dense.

Then $K \succeq 0 \Leftrightarrow K = Y^1 + Y^2$ such that

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

Even if $K^{(1)}, K^{(2)}$ not dense, we just assume that S is dense.

Chordal decomposition

Let $A \in \mathbb{S}^n$, $n \geq 3$, with a sparsity graph $G = (N, E)$.
Let $N = \{1, 2, \ldots, n\}$ be partitioned into $p \geq 2$ overlapping sets

$$N = I_1 \cup I_2 \cup \ldots \cup I_p \,.$$

Define $I_{k,k+1} = I_k \cap I_{k+1} \neq \emptyset \,, \quad k = 1, \ldots, p-1 \,.$

Assume $A = \sum_{k=1}^{p} A_k$, with A_k only non-zero on I_k.

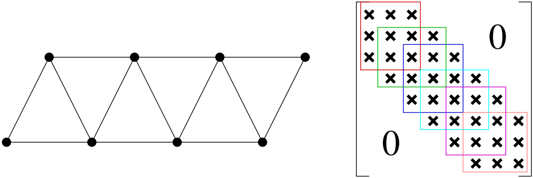
Corollary 1: $A \succeq 0$ if and only if
$\exists S_k \in \mathbb{S}^{I_{k,k+1}}, k = 1, \ldots, p-1$ s.t.

$$A = \sum_{k=1}^{p} \widetilde{A}_k \text{ with } \widetilde{A}_k = A_k - S_{k-1} + S_k \quad (S_0 = S_p = [\,])$$

and $\widetilde{A}_k \succeq 0$ $(k = 1, \ldots, p)$.

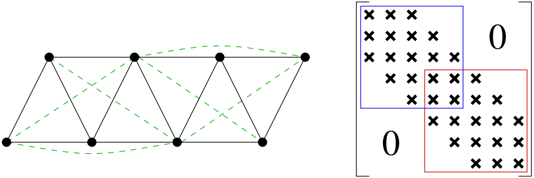
We can choose the partitioning $N = I_1 \cup I_2 \cup \ldots \cup I_p$!

Using the original theorem:



6 max. cliques of size 3, 5 additional 2×2 variables

Using the corollary:



2 “cliques” of size 5, 1 additional 2×2 variable

We can choose the partitioning $N = I_1 \cup I_2 \cup \ldots \cup I_p$!

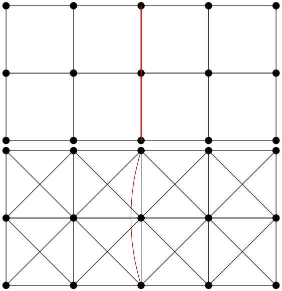
When we know the sparsity structure of A, we can choose a "regular" partitioning.

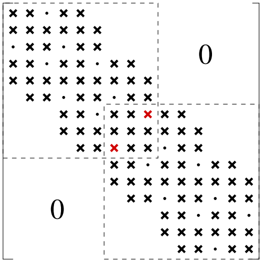
SDP formulation of TO by DD

$$\begin{pmatrix} K(\rho) & f \\ f^\top & \gamma \end{pmatrix} \succeq 0 \qquad \text{and} \qquad V(\hat{\lambda};\rho) \succeq 0$$

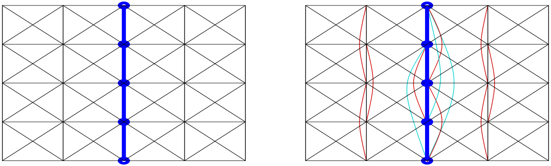
are large matrix constraints dependent on many variables.

FE mesh, matrix $K(\rho)$ and its sparsity graph:





Chordal decomposition



$$\begin{pmatrix} K_{II}^{(1)} & K_{I\Gamma}^{(1)} & 0 & 0 \\ K_{\Gamma I}^{(1)} & K_{\Gamma\Gamma}^{(1)}+K_{\Gamma\Gamma}^{(2)} & K_{\Gamma I}^{(2)} & 0 \\ 0 & K_{I\Gamma}^{(2)} & K_{II}^{(2)} & f \\ 0 & 0 & f^\top & \gamma \end{pmatrix} = \begin{pmatrix} K_{II}^{(1)} & K_{I\Gamma}^{(1)} & 0 & 0 \\ K_{\Gamma I}^{(1)} & K_{\Gamma\Gamma}^{(1)}+S & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & K_{\Gamma\Gamma}^{(2)}-S & K_{\Gamma I}^{(2)} & 0 \\ 0 & K_{I\Gamma}^{(2)} & K_{II}^{(2)} & f \\ 0 & 0 & f^\top & \gamma \end{pmatrix}$$

Even though $K^{(1)}$ and $K^{(2)}$ are sparse, we need to assume that S is dense.

In this way, we can control the number and size of the maximal cliques and use the chordal decomposition theorem.

New result: For the matrix inequality

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

the additional matrix variables S are rank-one; this further reduces the size of the solved SDP problem.

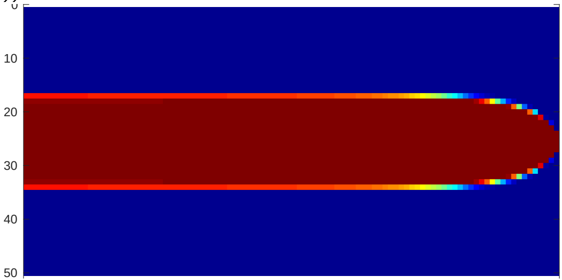
Numerical experiments

SDP codes tested: PENSDP, SeDuMi, SDPT3, Mosek

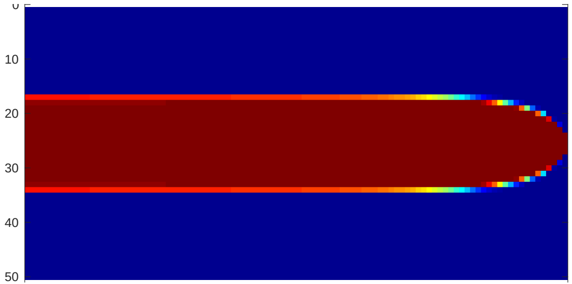
Results shown for Mosek: not the fastest for the original problem but has highest speedup

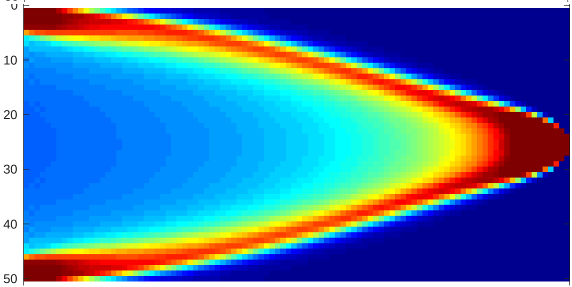
Mosek:
– new version 8 much more reliable than version 7
– called from YALMIP
– difficult (for me) to control any options

Numerical experiments



Numerical experiments





Numerical experiments

Regular decomposition, 40x20 elements, Mosek 8.0
Basic problem (no vibration constraints)

no of doms	no of vars	size of matrix	no of iters	CPU total	CPU per iter	speedup total	speedup /iter
1	801	1681	53	2489	47	1	1
2	844	882	66	778	12	3	4
8	1032	243	57	49	0.86	51	55
32	1492	73	55	11	0.19	235	244
50	1764	51	54	8	0.14	323	329
200	3544	19	45	5	0.10	553	470
34	22997	11...260	42	1206	29	2	2

Automatic decomposition using software SparseCoLO
by Kim, Kojima, Mevissen and Yamashita (2011); see page 16

Numerical experiments

Regular decomposition, 40x20 elements, Mosek 8.0

Problem with vibration constraints

no of matrices	no of vars	size of matrix	no of iters	CPU total	CPU per iter	speedup total	speedup /iter
2	801	1681	64	3894	61	1	1
16	1746	243	59	127	2.15	31	28
64	3384	73	54	27	0.50	144	122
100	4263	51	55	25	0.45	155	136
400	9258	19	37	18	0.49	216	125

and without again, for comparison:

no of matrices	no of vars	size of matrix	no of iters	CPU total	CPU per iter	speedup total	speedup /iter
1	801	1681	53	2489	47	1	1
8	1032	243	57	49	0.86	51	55
32	1492	73	55	11	0.19	235	244
50	1764	51	54	8	0.14	323	329
200	3544	19	45	5	0.10	553	470

Numerical experiments

Regular decomposition, 120x60 elements, Mosek 8.0
Basic problem (no vibration constraints)

no of doms	no of vars	size of matrix	no of iters	CPU total	CPU per iter	speedup total	speedup /iter
1	7200	14641	178	5089762	28594	1	1
50	9524	339	85	1475	17.4	3541	1648
200	12904	99	72	209	2.9	24355	9851
450	16984	51	67	107	1.6	47568	17905
800	21764	33	61	82	1.3	62070	21271
1800	33424	19	44	77	1.6	66101	18196

estimated; 508976 sec $\approx$ 2 months

Numerical experiments

Regular decomposition, Mosek 8.0

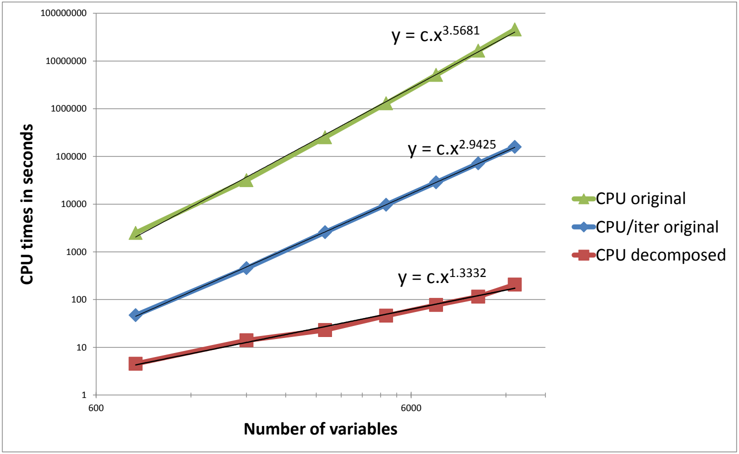
Basic problem (no vibration constraints)

"best" decomposition speedup (subdomain = 4 elements)

	ORIGINAL			DECOMPOSED			speedup
problem	no of vars	size of matrix	CPU total	no of vars	size of matrix	CPU total	
40x20	801	1681	2489	3544	19	8	311
60x30	1801	3721	31835	8164	19	25	1273
80x40	3201	6561	252355	14684	19	23	10972
100x50	5001	10201	1298087	23104	19	46	28219
120x60	7201	14641	5091862	33424	19	77	66128
140x70	9801	19881	16436180	45664	19	115	142923
160x80	12801	25921	45804946	59764	19	206	222354
complexity $c{\cdot}\text{size}^q$			$q = 3.5$			$q = 1.33$	

times estimated; 45804946 sec $\approx$ 18 months

CPU time, original versus decomposed



THE END