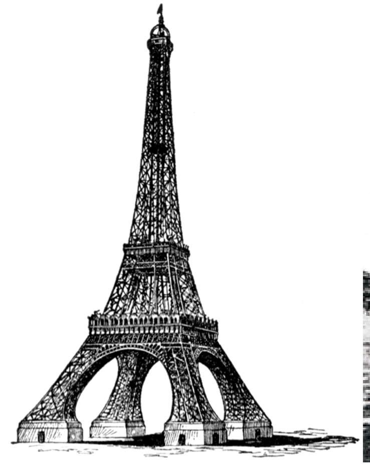
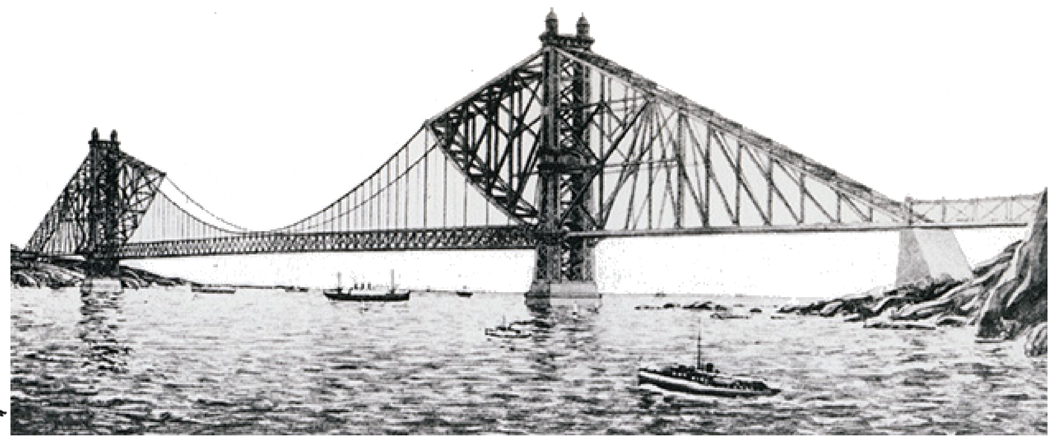




A Randomized Coordinate Descent Method for Large-Scale Truss Topology Design

Peter Richtárik and Martin Takáč

Peter.Richtarik@ed.ac.uk, M.Takac@sms.ed.ac.uk

1. Introduction

truss - a mechanical construction made of elastic **bars** linked to each other at **nodes** (these are fixed or free)

load - external forces acting at the free nodes, causing deformation of the truss

compliance - potential energy stored in the truss after deformation

Goal of TTD: Given a grid structure of nodes and forces acting on them, construct a truss of a given total weight of **minimum compliance.**

Applications of TTD: Railroad bridges, electric masts, ...

2. The Algorithm

Consider the following optimization problem

$$\min_{x \in \mathbb{R}^n} F(x) \equiv f(x) + \Psi_1(x^{(1)}) + \cdots + \Psi_n(x^{(n)}), \quad \text{(P)}$$

with f convex and $\forall x, \tau$ and i satisfying

$$|\nabla_i f(x) - \nabla_i f(x + \tau e_i)| \le L_i |\tau|$$

(gradient of f is coordinate-wise Lipschitz) and Ψ_i convex, **nonsmooth and simple.**

Algorithm 1: UCDC

1 choose initial point $x_0 \in \mathbb{R}^n$

for $k = 1, 2, \dots$ **do**

2 choose $i \in \{1, 2, \dots, n\}$ with prob. $\frac{1}{n}$

3 $\tau^* = \arg\min_{\tau \in \mathbb{R}} \nabla_i f(x_k)\tau + \frac{L_i}{2}\tau^2 + \Psi_i(x_k^{(i)} + \tau)$

4 $x_{k+1} := x_k + \tau^* e_i$

Theorem (R-T [5]): Choose initial point x_0 and target confidence ρ. If target accuracy satisfies $0 < \epsilon < F(x_0) - F^*$ then after

$$k \ge \frac{2nC}{\epsilon}\left(1 + \log\frac{1}{\rho}\right) + 2 - \frac{2nC}{F(x_0) - F^*}$$

iterations we get

$$\mathbf{Prob}[F(x_k) - F(x^*) \le \epsilon] \ge 1 - \rho,$$

where $C = \max\{R_L^2(x_0), F(x_0) - F^*\}$, $R_L(x_0) = \max_x\{\|x - x^*\|_L : F(x) \le F(x_0)\}$, $\|x\|_L = (\sum_{i=1}^n L_i (x^{(i)})^2)^{\frac{1}{2}}$ and x^* solves (P).

3. Optimization Formulations of Truss Topology Design

# nodes (grid size)	$r \times c$
# free nodes	m
# fixed nodes	$rc - m$
# potential bars	n
bar weights	$w \in \mathbb{R}^n$
forces acting at the free nodes	$d \in R^{2m}$
displacement associated with bar i	$b_i \in R^{2m}$

Linearization of the physics involved gives the following formula:

$$\text{Compl}_d(w) = \tfrac{1}{2} d^T v, \quad \text{where} \sum_{i=1}^n w_i b_i b_i^T v = d. \quad (1)$$

The problem of **minimizing compliance** subject to $\sum_i w_i = 1$ can be equivalently written as

$$\max_v \{d^T v \ : \ |b_i^T v| \le 1, i = 1, \dots, n\}, \quad (2)$$

the dual of which is equivalent to

$$\min_q \{\|q\|_1 : Bq = d\}, \text{ where } B = (b_1, \dots, b_n). \quad (3)$$

Problem (2) can be reformulated as

$$\min_v \{\max_i |b_i^T v| \ : \ d^T v = 1\}. \quad (4)$$

After elimination of one variable from $d^T v = 1$ we get an unconstrained problem of the form:

$$\min_{\tilde{v}} \{\max_i |\tilde{b}_i^T \tilde{v} - c_i|\} \quad (5)$$

(see [1, 4] for more details).

4. Approach 1 (Penalization)

Instead of solving problem (3) one may penalize the function with

$$f_\gamma^1(q) \equiv \frac{\gamma}{2}\|Bq - d\|_2^2, \quad \gamma > 0$$

and instead solve:

$$\min_q \{\|q\|_1 + f_\gamma^1(q)\}. \quad (6)$$

$\boxed{\text{UCDC runs with } f = f_\gamma^1 \text{ and } \Psi_i \equiv |\cdot|.}$

5. Approach 2 (Smoothing)

The objective of (5) can be approximated to any accuracy by the smooth convex function (see [3,4] for details)

$$f_\xi^2(\tilde{v}) = \xi \log\left[\frac{1}{2n}\sum_{i=1}^n \left(e^{(\tilde{b}_i^T \tilde{v} - c_i)/\xi} + e^{-(\tilde{b}_i^T \tilde{v} - c_i)/\xi}\right)\right].$$

Moreover,

$$0 \le \max_i |\tilde{b}_i^T \tilde{v} - c_i| - f_\xi^2(\tilde{v}) \le \xi \log 2n, \ \forall \tilde{v}.$$

$\boxed{\text{UCDC runs with } f = f_\xi^2 \text{ and } \Psi_i \equiv 0}$

Potential numerical problems with small ξ can be avoided by minimizing $e^{-M/\xi} f_\xi(\tilde{v})$, where $M = \max_i |\tilde{b}_i^T \tilde{v}_0 - c_i|$ and v_0 is our initial point.

6. Comparison of Approaches

	Approach 1	Approach 2
f	f_γ^1	f_ξ^2
$\Psi_i(\cdot)$	$\lvert \cdot \rvert$	0
L_i	$\gamma b_i^T b_i$	$\frac{2}{\xi}\|B_i\|_\infty^2$
work/iteration	$O(4)$	$O(2n/m)$
computing L_i	$O(4)$	$O(2n/m)$

(B_i is i-th row of B)

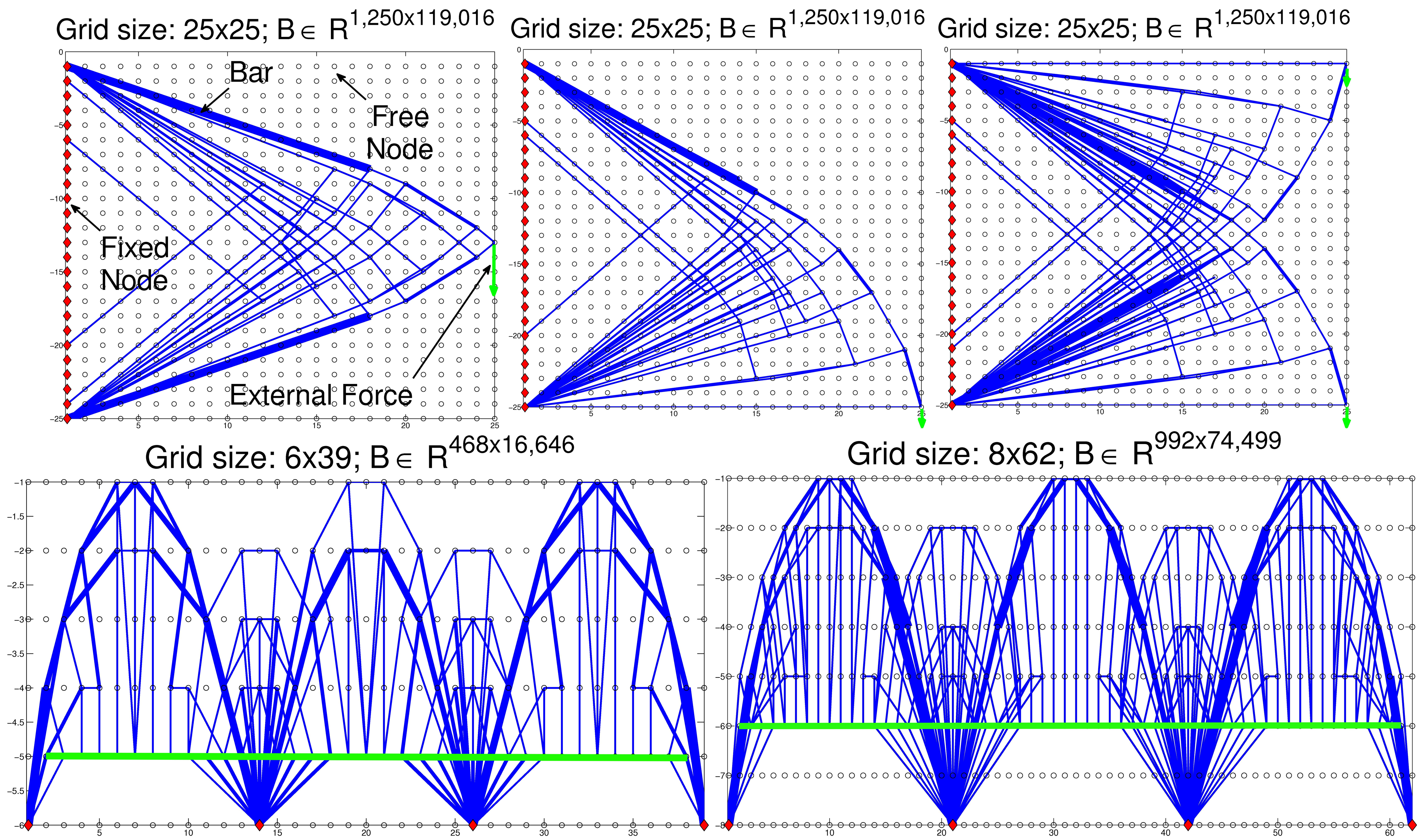
Each iteration is extremely cheap!

7. Data structure

While the size of B grows fast with the grid size, each column has **at most 4 nonzeros!**

grid $r \times c$	$B \in \mathbb{R}^{2m \times n}$	nonzero elm.
5×5	50×196	712
25×25	$1,250 \times 119,016$	473,712
100×100	$20,000 \times 30,398,795$	121,555,778
125×125	$31,250 \times 74,220,244$	296,819,224

8. Numerical Examples



Bridge trusses: there is a downward external force acting on every node covered by the green line.

9. References

[1] Nemirovski, A., Ben-Tal, A.: Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001.

[2] Nesterov, Yu.: Smooth minimization of non-smooth functions, Mathematical Programming, 103 (2005), pp. 127-152.

[3] Nesterov, Yu.: Efficiency of coordinate descent methods on huge-scale optimization problems, CORE Discussion Paper #2010/2, January 2010.

[4] Richtárik, P.: Simultaneously solving seven optimization problems in relative scale, Optimization online, 2009.

[5] Richtárik, P., Takáč, M.: Iteration complexity of randomized block-coordinate descent methods for minimizing a composite function, 2011.