OSE SEMINAR 2014

Quadratic reformulation techniques for 0-1 quadratic programs

Ray Pörn

CENTER OF EXCELLENCE IN
OPTIMIZATION AND SYSTEMS ENGINEERING
ÅBO AKADEMI UNIVERSITY

ÅBO NOVEMBER 14th 2014

ose OPTIMIZATION AND SYSTEMS ENGINEERING

Åbo Akademi University

Structure of the presentation

Joint work with PhD student Otto Nissfolk and prof. Tapio Westerlund.

- Background
- Convexification and some examples
- Quadratic and semidefinite programming
- Quadratic Convex Reformulation (QCR method)
- Nondiagonal perturbation (NDQCR)
- Numerical experiments

0-1 Quadratic Program (QP)

A standard 0-1 QP has the form:

$$\begin{aligned} \min \quad & x^T Q x + q^T x \\ s.t. \quad & Ax = a \\ & Bx \leq b \\ & x \in \{0,1\}^n \end{aligned}$$

Q, A, B are matrices and q, a, b are vectors of appropriate dimensions.

Some applications include:

- Max-Cut of a graph (unconstrained)
- Knapsack problems (inequality constrained)
- Graph bipartitioning
- Task allocation
- Quadratic assignment problems
- Coulomb glass
- Gray-scale pattern problems, taixxc instances from QAPLIB

Convexity

The following are equivalent ($Q = Q^T$):

- The quadratic function $f(x) = x^T Qx$ is convex on R^n.
- The matrix Q is positive semidefinite (psd, $Q \succcurlyeq 0$).
- All eigenvalues of Q are non-negative ($\lambda_i \geq 0$).

A sufficient condition for convexity: A **diagonally dominant** matris is psd.

Definition: A matrix Q is diagonally dominant if

$$|Q_{ii}| \geq \sum_{i \neq j} |Q_{ij}| \quad \forall i$$

Convexification of 0-1 QPs

Basic approach: If Q is indefinite, add sufficient large quadratic terms to the diagonal and subtract the same amount from the linear terms.

Recall that: $x_i \in \{0,1\} \;\Leftrightarrow\; x_i^2 = x_i$

Example 1

$$f(x) = x^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} x = x_1^2 + 6x_1x_2 + 2x_2^2$$

$$f(x) = x^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} x = x^T \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} x - \begin{bmatrix} 2 \\ 3 \end{bmatrix}^T x = 3x_1^2 + 6x_1x_2 + 5x_2^2 - 2x_1 - 3x_2$$

Indefinite

Positive semidefinite

Convexification of 0-1 QPs

Example 2: a) Diagonal dominance, b) Minimum eigenvalue, c) Best diagonal

$$\begin{array}{ll} \min & x^T Q x \\ s.t. & x \in \{0,1\}^4 \end{array} \qquad Q = \begin{bmatrix} 1 & 2 & -3 & 2 \\ 2 & 2 & -3 & 4 \\ -3 & -3 & 2 & 0 \\ 2 & 4 & 0 & -2 \end{bmatrix} \qquad \mathrm{eig}(Q) = \begin{bmatrix} -5.17 \\ -1.04 \\ 0.95 \\ 8.26 \end{bmatrix}$$

a) Diagonal dominance

$$\hat{Q} = \begin{bmatrix} 7 & 2 & -3 & 2 \\ 2 & 9 & -3 & 4 \\ -3 & -3 & 6 & 0 \\ 2 & 4 & 0 & 6 \end{bmatrix} \quad \hat{q} = \begin{bmatrix} 6 \\ 7 \\ 4 \\ 8 \end{bmatrix} \quad \mathrm{eig}(\hat{Q}) = \begin{bmatrix} 1.66 \\ 4.90 \\ 6.88 \\ 14.56 \end{bmatrix} \qquad \begin{array}{ll} \min & x^T \hat{Q} x - \hat{q}^T x \\ s.t. & x \in [0,1]^4 \end{array}$$

$$\textbf{optimal value} = -5.93$$

b) Minimum eigenvalue

$$\hat{Q} = \begin{bmatrix} 6.17 & 2 & -3 & 2 \\ 2 & 7.17 & -3 & 4 \\ -3 & -3 & 7.17 & 0 \\ 2 & 4 & 0 & 3.17 \end{bmatrix} \quad \hat{q} = \begin{bmatrix} 5.17 \\ 5.17 \\ 5.17 \\ 5.17 \end{bmatrix} \quad \mathrm{eig}(\hat{Q}) = \begin{bmatrix} 0 \\ 4.13 \\ 6.12 \\ 13.43 \end{bmatrix}$$

$$\begin{array}{ll} \min & x^T \hat{Q} x - \hat{q}^T x \\ s.t. & x \in [0,1]^4 \end{array} \qquad \textbf{optimal value} = -5.34$$

Convexification of 0-1 QPs

c) ***The best diagonal****. The QCR method allows computation of the diagonal that gives the highest optimal value of the relaxation.*

$$\hat{Q} = \begin{bmatrix} 2.93 & 2 & -3 & 2 \\ 2 & 4.28 & -3 & 4 \\ -3 & -3 & 6.83 & 0 \\ 2 & 4 & 0 & 6.20 \end{bmatrix} \qquad \hat{q} = \begin{bmatrix} 1.93 \\ 2.28 \\ 4.83 \\ 8.20 \end{bmatrix} \qquad \mathrm{eig}(\hat{Q}) = \begin{bmatrix} 0 \\ 1.31 \\ 6.71 \\ 12.21 \end{bmatrix}$$

$$\begin{array}{ll} \min & x^T \hat{Q} x - \hat{q}^T x \\ s.t. & x \in [0,1]^4 \end{array} \qquad \textbf{optimal value} = \mathbf{-4.08}$$

$$\begin{array}{ll} \min & x^T Q x \\ s.t. & x \in \{0,1\}^4 \end{array} \qquad \textbf{optimal value} = \mathbf{-3}$$

Bounding: $\mathbf{-5.93 \leq -5.34 \leq -4.08 \leq -3}$

Semidefinite relaxation of 0-1 QPs

$$\begin{array}{ll} \min & x^T Q x + q^T x \\ s.t. & Ax = a \\ & Bx \leq b \\ & x \in \{0,1\}^n \end{array}$$

Relaxation into a positive semidefinite matrix variable

$$X = xx^T \;\mapsto\; X - xx^T \succcurlyeq 0 \;\Leftrightarrow\; \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \succcurlyeq 0$$

A quadratic expression in x is linear in X: $x^T Q x = Q \bullet X = \sum_i \sum_j Q_{ij} X_{ij}$

Binary condition: $x_i \in \{0,1\} \;\Leftrightarrow\; x_i^2 - x_i = 0 \;\Leftrightarrow\; X_{ii} = x_i$

Semidefinite relaxation:

$$\begin{array}{ll} \min & Q \bullet X + q^T x \\ s.t. & Ax = a \\ & Bx \leq b \\ & \operatorname{diag}(X) = x \\ & \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \succcurlyeq 0 \end{array}$$

Deriving the dual problem

Lagrangian relaxation of 0-1 QP:

$$f(x,\lambda,\mu,\delta) = x^T Qx + q^T x + \lambda^T (Ax - a) + \mu^T (Bx - b) + \sum_{i=1}^{n} \delta_i (x_i^2 - x_i)$$
$$= x^T \underbrace{(Q + \mathrm{Diag}(\delta))}_{\bar{Q}} x + \underbrace{(q + A^T\lambda + B^T\mu - \delta)}_{\bar{q}}{}^T x \underbrace{-\lambda^T a - \mu^T b}_{\bar{c}}$$

Lagrangian dual problem:

$$\sup_{\delta,\lambda,\mu} \inf_{x \in R^n} \quad x^T \bar{Q} x + \bar{q}^T x + \bar{c}$$

which equals a semidefinite program

$$\begin{array}{ll} \max & t \\ s.t. & \begin{bmatrix} -t + \bar{c} & \frac{1}{2}\bar{q}^T \\ \frac{1}{2}\bar{q} & \bar{Q} \end{bmatrix} \succcurlyeq 0 \\ & \delta \in R^n, \lambda \in R^m, \mu \in R^k_+ \end{array}$$

The primal and dual

$$
\begin{array}{ll}
\min & Q \bullet X + q^T x \\
s.t. & Ax = a \\
& Bx \leq b \\
& \mathrm{diag}(X) = x \\
& \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \succcurlyeq 0
\end{array}
$$

$$
\begin{array}{ll}
\max & t \\
s.t. & \begin{bmatrix} -t + \bar{c} & \frac{1}{2}\bar{q}^T \\ \frac{1}{2}\bar{q} & \bar{Q} \end{bmatrix} \succcurlyeq 0 \\
& \delta \in R^n, \lambda \in R^m, \mu \in R^k_+
\end{array}
$$

Solution give optimal values on the multipliers : $\delta^*, \lambda^*, \mu^*$.

These are used to construct the ”best” diagonal perturbation of matrix Q according to

$$Q^* = Q + \mathrm{Diag}(\delta^*).$$

Strengthening

*Inclusion of constraints may improve bounding quality. There are **many ways** to include or construct quadratic constraints.*

1) *Add new redundant quadratic constraints (some examples)*

$$x_i x_j \geq 0, \qquad x_i x_j \geq x_i + x_j - 1, \qquad x_i x_j \leq x_i, \qquad x_i x_j \leq x_j$$

2) *Combine and multiply existing linear constraints (some examples)*

$$p^T x = s \implies x_i p^T x = x_i s \quad \forall i$$
$$p^T x = s \implies (1 - x_i) p^T x = (1 - x_i) s \quad \forall i$$

$$\begin{cases} p^T x = s \\ r^T x = t \end{cases} \implies p^T x r^T x = st \implies x^T (p r^T) x = st$$

$$Ax = a \implies \|Ax - a\|^2 = 0 \implies x^T A^T A x = a^T a$$

Our strengthening

Original 0-1 QP

$$
\begin{aligned}
\min \quad & x^T Q x + q^T x \\
s.t. \quad & Ax = a \\
& Bx \le b \\
& x \in \{0,1\}^n
\end{aligned}
$$

Strengthened SDP relaxation

$$
\begin{aligned}
\min \quad & Q \bullet X + q^T x \\
s.t. \quad & Ax = a \\
& Bx \le b \\
& \operatorname{diag}(X) = x \\
& A^T A \bullet X = a^T a \\
& X_{ij} \ge 0, \quad X_{ij} \ge x_i + x_j - 1 \quad \forall\, i \ne j \\
& X_{ij} \le x_i, \quad X_{ij} \le x_j \quad \forall\, i \ne j \\
& \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \succeq 0 \\
& x \in R^n, \; X \in S^n
\end{aligned}
$$

- Multipliers from all quadratic constraints are used to convexify the objective function so that the lower bound becomes as high as possible.

- Multipliers: $\delta \in R^n, \alpha \in R, \quad S, T, U, V \ge 0$

Convexified 0-1 QP problem

$$
\begin{array}{ll}
\min & x^T Q x + q^T x \\
s.t. & Ax = a \\
 & Bx \leq b \\
 & x \in \{0,1\}^n
\end{array}
$$

$$
\begin{aligned}
\bar{Q}^* &= Q + \mathrm{Diag}(\delta^*) + \alpha^* A^T A - S^* - T^* + U^* + V^* \\
\bar{q}^* &= q + A^T \lambda^* + B^T \mu^* - \delta^* \\
\bar{c}^* &= -\lambda^{*T} a - \mu^{*T} - \alpha^* a^T a
\end{aligned}
$$

(MIQP)

$$
\begin{array}{ll}
\min & x^T \overline{Q}^* x + \overline{q}^{*T} x + \overline{c}^* + 2\sum_{i=1}^{n} \sum_{j=i+1}^{n} (S^*_{ij} + T^*_{ij}) y_{ij} - 2\sum_{i=1}^{n} \sum_{j=i+1}^{n} (U^*_{ij} + V^*_{ij}) z_{ij} \\
s.t. & Ax = a \\
 & Bx \leq b \\
 & y_{ij} \geq 0, \quad y_{ij} \geq x_i + x_j - 1 \\
 & z_{ij} \leq x_i, \quad z_{ij} \leq x_j \\
 & x \in \{0,1\}^n \\
 & y_{ij}, \quad z_{ij} \in R_+ \ (i < j)
\end{array}
$$

NDQCR method

Non-diagonal quadratic convex reformulation technique (NDQCR)
Given a general QP01 problem.

1. Strengthen the problem by including a set of RLT inequalities and squared norm constraints.
2. Solve the semidefinite relaxation (SDPr) and its dual (SDPd).
3. Collect the multiplier values and form problem MIQP.
4. Solve problem MIQP using any suitable solver.

NDQCR versus QCR

Example 3:

$$\begin{aligned} \min \quad & x^T Q x + q^T x \\ & Ax = a \\ & x \in \{0,1\}^5 \end{aligned}$$

$$Q = \begin{bmatrix} 0 & -24 & 2 & 18 & -12 \\ -24 & 0 & -3.5 & 18 & -42 \\ 2 & -3.5 & 0 & 20 & 2 \\ 18 & 18 & 20 & 0 & -44 \\ -12 & -42 & 2 & -44 & 0 \end{bmatrix}, \quad q = \begin{bmatrix} -9 \\ -7 \\ 2 \\ 23 \\ 12 \end{bmatrix}, \quad A^T = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \quad a = 2$$

i) α and δ perturbations (QCR method)

ii) α, δ and S perturbations

iii) α, δ and T perturbations

iv) α, δ and U perturbations

v) α, δ and V perturbations

Strategy	i)	ii)	iii)	iv)	v)
$v(*)$	-88.02	-80	-82.23	-82.20	-83.84

NDQCR versus QCR

Best reformulation – strategy (ii)

Multipliers

$$S^* = \begin{bmatrix} 0 & 1.99 & 1.40 & 56.96 & 12.66 \\ 1.99 & 0 & 0 & 32.40 & 0 \\ 1.40 & 0 & 0 & 22.38 & 0 \\ 56.96 & 32.40 & 22.38 & 0 & 6.36 \\ 12.66 & 0 & 0 & 6.36 & 0 \end{bmatrix}, \quad \delta^* = \begin{bmatrix} -15.89 \\ 4.78 \\ 1.00 \\ -18.07 \\ -25.22 \end{bmatrix}, \quad \alpha^* = 113.32$$

Matrices

$$\overline{Q}^* = Q + \mathrm{Diag}\,(\delta^*) + \alpha^* A^T A - S^* = \begin{bmatrix} 97.44 & 87.33 & 0.60 & 74.36 & 88.66 \\ 87.33 & 118.10 & -3.50 & 98.92 & 71.32 \\ 0.60 & -3.50 & 1.00 & -2.38 & 2.00 \\ 74.36 & 98.92 & -2.38 & 95.26 & 62.96 \\ 88.66 & 71.32 & 2.00 & 62.96 & 88.10 \end{bmatrix} \quad \overline{q}^* = q - \delta^* = \begin{bmatrix} 6.89 \\ -11.78 \\ 1.00 \\ 41.07 \\ 37.22 \end{bmatrix}$$

$$\overline{c}^* = -\alpha^* a^T a = -453.29$$

Convexified QP

$$\boxed{\begin{array}{ll} \min & x^T\, \overline{Q}^* x + \overline{q}^{*T} x + \overline{c}^* + 2 \displaystyle\sum_{(i,j)\in I} S^*_{ij} y_{ij} \\ \text{s.t.} & x_1 + x_2 + x_4 + x_5 = 2 \\ & y_{ij} \geq 0, \;\; y_{ij} \geq x_i + x_j - 1 \quad \forall\ (i,j) \in I \\ & x \in \{0,1\}^5 \end{array}}$$

Boolean least squares

The problem is to identify a binary signal $x \in \{0,1\}^n$ from a collection of noisy measurements.

$$\begin{array}{ll} \min & \|Dx - d\|^2 \\ \text{s.t.} & x \in \{0,1\}^n \end{array}$$

$$\begin{array}{ll} \min & D^T D \bullet X - 2d^T Dx + d^T d \\ \text{s.t.} & \operatorname{diag}(X) = x \\ & \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \succcurlyeq 0 \\ & X_{ij} \geq 0,\ X_{ij} \geq x_i + x_j - 1 \\ & X_{ij} \leq x_i,\ X_{ij} \leq x_j \end{array}$$

	MIQP		SDP		Total
Size (n)	Gap	Time	Gap	Time	time
40	0.00 %	1.2	21.47 %	0.4	1.6
60	0.00 %	16.3	27.33 %	0.5	16.8
80	0.00 %	175.0	31.50 %	0.6	175.6
100	3.29 %	1849.9	37.59 %	0.8	1850.7

Table 1: Average results for BLS with QCR

	MIQP		SDP		Total
Size (n)	Gap	Time	Gap	Time	time
40	0.00 %	0.6	2.78 %	14.1	14.8
60	0.00 %	3.8	5.87 %	41.0	44.8
80	0.00 %	17.4	7.72 %	107.8	125.2
100	0.00 %	236.5	11.13 %	261.8	498.4

Table 3: Average results for BLS with NDQCR

Coulomb glass problem

Given *n* sites in the plane. *k* of these sites are filled with electrons. Find the configuration that has minimal energy.

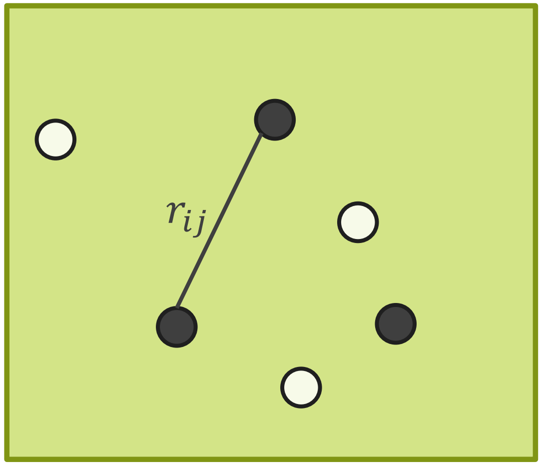
Energy = Coulomb interaction + site specific energy

Variables: $x \in \{0,1\}^n$

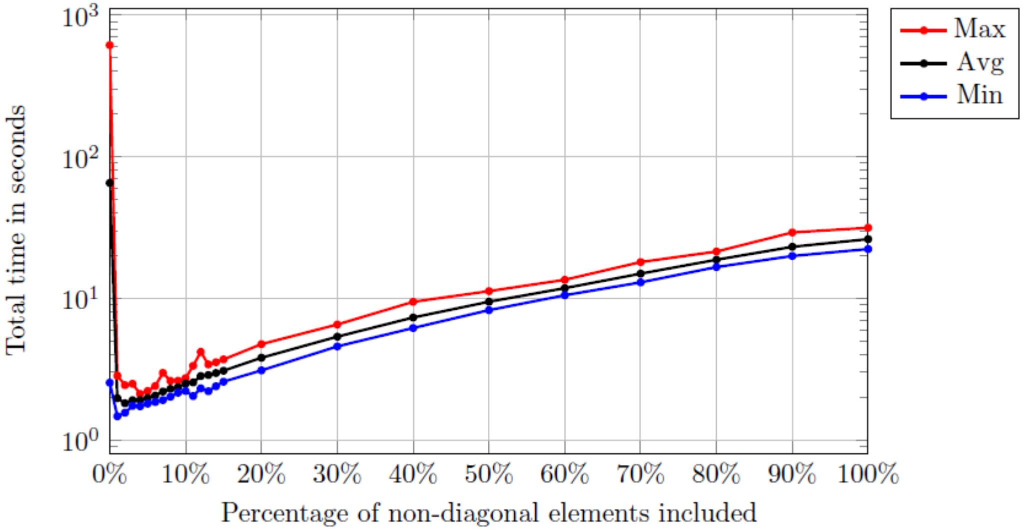
$$x_i = \begin{cases} 0, & \text{if site } i \text{ is empty} \\ 1, & \text{if site } i \text{ is filled} \end{cases}$$

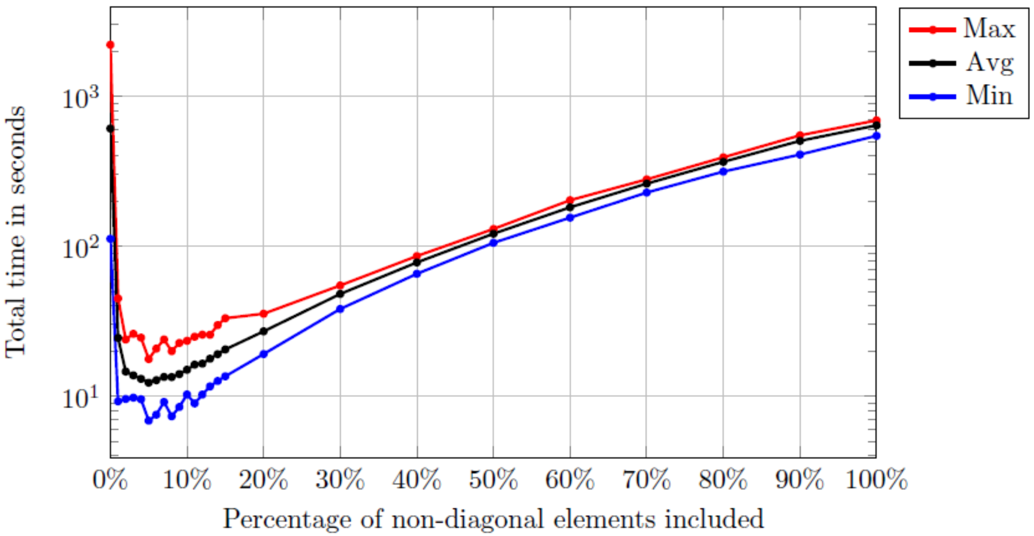
$$\boxed{\begin{array}{ll} \min & x^T Q x + q^T x \\ s.t. & e^T x = k \\ & x \in \{0,1\}^n \end{array}}$$

$$Q_{ij} = \begin{cases} 0, & \text{if } i = j \\ \dfrac{1}{2r_{ij}}, & \text{if } i \neq j \end{cases}$$



NDQCR on Coulomb glass problems (n=50, n=100)





- $k = \frac{n}{2}$ in all experiments
- Constraints $X_{ij} \geq 0$ and $X_{ij} \geq x_i + x_j - 1$ are included for indeces corresponding to the *p*% largest elements of Q.
- Even a small fraction of non-diagonal elements has a large impact on the total solution time.
- 2% - 10% non-diagonal elements result in fastest solution times.

The taixxc instances from QAPLIB

These instances are a special type of QAP problem where the flow matrix F is binary and rank-1.

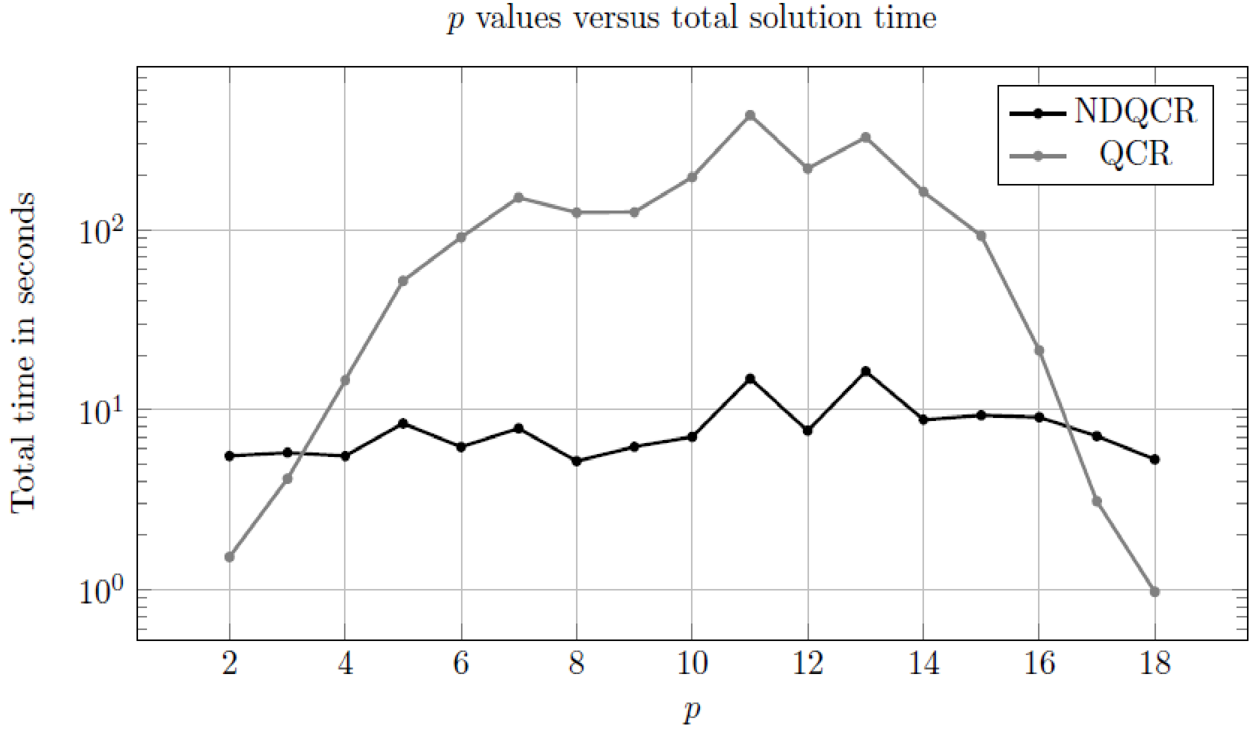
$$F = bb^T \text{ where } b \in \{0,1\}^n.$$

The objective function of QAP can be rewritten and simplified using the binary rank-1 property:

$$\begin{aligned}\operatorname{trace}(DXFX^T) &= \operatorname{trace}(DXbb^TX^T) = \operatorname{trace}(DXb(Xb)^T)\\ &= \operatorname{trace}(Dyy^T) = \operatorname{trace}(y^TDy) = y^TDy\end{aligned}$$

$$\boxed{\begin{array}{ll}\min & y^TDy\\ \text{s.t.} & e^Ty = p\\ & y \in \{0,1\}^n\end{array}}$$

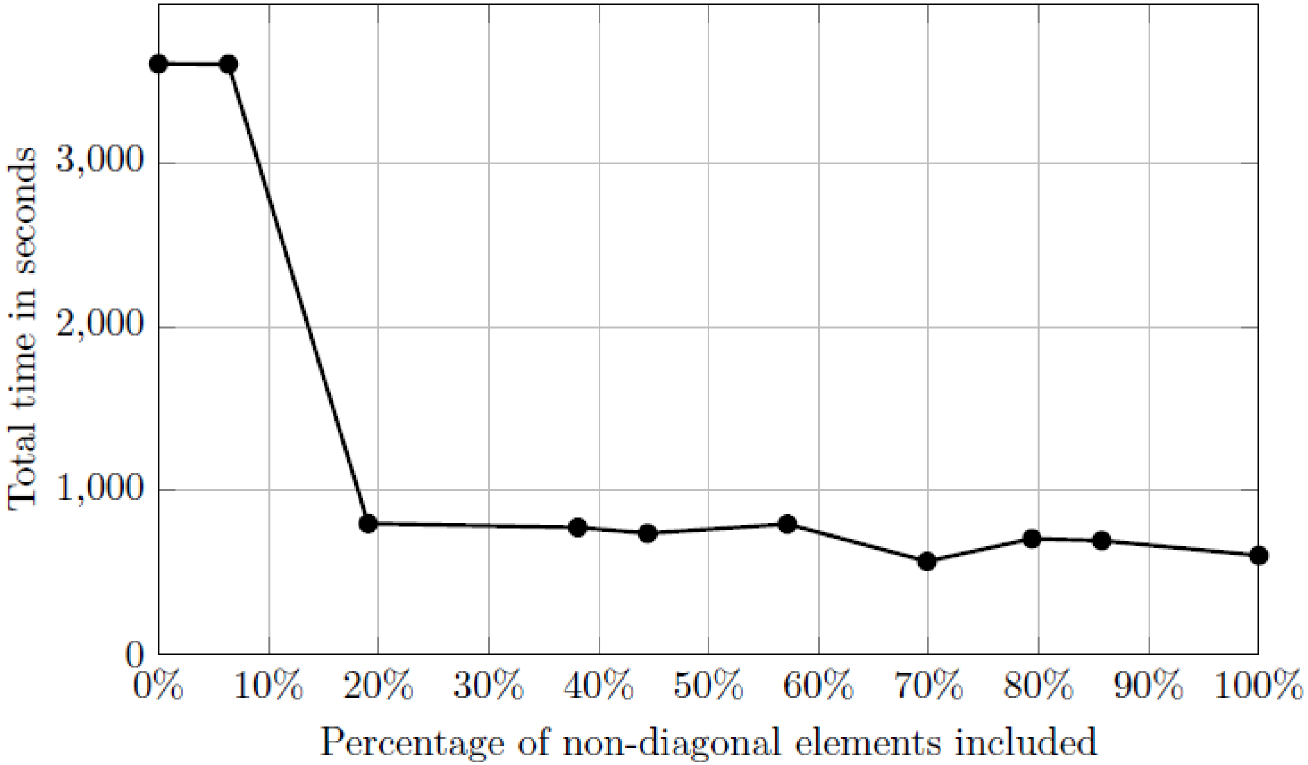
NDQCR (Y≥0) versus QCR on tai36c



$$\begin{array}{ll} \min & y^T D y \\ \text{s.t.} & e^T y = p \\ & y \in \{0,1\}^n \end{array}$$

p=11	QCR	NDQCR
CPU time	433 s	15 s
SDP gap	24 %	4 %

NDQCR on problem tai64c



Conclusions

- A technique for non-diagonal perturbation was presented.
- Non-diagonal perturbation was obtained from squared norm constraints and a set of redundant RLT inequalities.
- Gives tight bounding and fast solution for small to medium sized problems.
- Full application becomes impossible for large problems.
- The inclusion of just a few RLT inequalities may also have large impact on the solution time and bounding quality.

Future work: Construct reasonable fast heuristic procedures to find a good set of inequalities to include.

Some references on quadratic reformulation

P.L. Hammer and A.A. Rubin. Some remarks on quadratic programming with 0-1 variables. *RAIRO*, 3:67–79, 1970.

Billionnet, A., Elloumi, S., Plateau, M.-C.: Improving the performance of standard solvers for quadratic 0-1 programs by a tight convex reformulation: the QCR method. *Discrete Applied Mathematics*. **157**(6) : 1185-1197 (2009)

Billionnet, A., Elloumi, S., Lambert, A.: Extending the QCR method to the case of general mixed integer program. *Mathematical Programming*. Available online DOI: 10.1007/s10107-010-0381-7 (2010)

O. Nissfolk, R. Pörn, T. Westerlund, F. Jansson, A mixed integer quadratic reformulation of the quadratic assignment problem with rank-1 matrix, in: I. A. Karimi, R. Srinivasan (Eds.), 11th International Symposium on Process Systems Engineering, Vol. 31 of Computer Aided Chemical Engineering, Elsevier, 2012, pp. 360 – 364. doi:10.1016/B978-0-444-59507-2.50064-0.

L. Galli, A. Letchford, A compact variant of the QCR method for quadratically constrained quadratic 0-1 programs, Optimization Letters (2013) 1–12.

THANK YOU FOR YOUR ATTENTION