Applications to Hard Discrete Optimization and Low Rank Solutions:
SDP Relaxations of **QQP**s; Solving **QAP**s

Henry Wolkowicz

Dept. Combinatorics and Optimization, University of Waterloo, Canada

Tues. Feb. 28, 2023,

Lectures

Summary: Lectures

- We recall **QQP**s that model hard combinatorial problems.
- In particular, we look at the quadratic assignment problem, **QAP**, arguably the hardest of the NP-hard problems.
- We exploit the special structure of the **QAP** to obtain an efficient **SDP** relaxation and low rank solutions.

Preview: Lectures

- We look at an **ADMM** approach for solving the **SDP** relaxation and taking advantage of the structure to get low rank solutions. (Work with: Oliveira/Xu [2]; Graham/Hu/Im/Li [1])
- We continue with looking at low rank solutions for matrix completion problems.

$$p^* := \quad \min \quad \text{trace } CX \quad (= \langle C, X \rangle)$$
$$\text{(PSDP)} \qquad \text{s.t.} \quad \mathcal{A}X = b \quad (b_i = \langle A_i, X \rangle)$$
$$X \succeq 0.$$

$$d^* := \quad \max \quad b^T y \qquad (= \langle b, y \rangle)$$
$$\text{(DSDP)} \qquad \text{s.t.} \quad \mathcal{A}^* y \preceq C \quad (\mathcal{A}^* y = \sum_{i=1}^{m} y_i A_i)$$

$$(\mathcal{A}^* y + Z = C, \text{ slack } Z \succeq 0)$$

$p^* \geq d^*$ is weak duality; provides a lower bound.

Strong Duality

$p^* = d^*$ and d^* is attained.

SDP arise from general quadratic approximations?

General Quadratic Approximations

Approximations from quadratic functions are stronger than from linear functions. E.g.

$x \in \{\pm 1\}$ iff $x^2 = 1$ $x \in \{0, 1\}$ iff $x^2 - x = 0$

QQPs

Let

$$q_i(y) = \frac{1}{2} y^T Q_i y + y^T b_i + c_i, \; y \in \Re^n$$

$$
\begin{aligned}
q^* = \;\; &\min && q_0(y) \\
(QQP) \;\; &\text{s.t.} && q_i(y) \leq 0 \\
& && i = 1, \ldots m
\end{aligned}
$$

Lagrangian Relaxation

Lagrangian; x Lagrange multiplier vector

$$L(y, x) = q_0(y) + \sum_{i=1}^{m} x_i q_i(y)$$

or equivalently (combine quad./lin. terms)

$$L(y, x) = \frac{1}{2} y^T (Q_0 + \sum_{i=1}^{m} x_i Q_i) y \\ + y^T (b_0 + \sum_{i=1}^{m} x_i b_i) \\ + (c_0 + \sum_{i=1}^{m} x_i c_i)$$

Weak Duality

Use hidden constraints

$$d^* = \max_{x \geq 0} \min_{y} L(y, x) \leq q^* = \min_{y} \max_{x \geq 0} L(y, x)$$

Homogenize the Lagrangian

multiply linear term by new variable y_0:

$$y_0 y^T (b_0 + \sum_{i=1}^m x_i b_i), \quad y_0^2 = 1$$

use: strong duality for TRS; hidden **SDP** constraints

$$
\begin{aligned}
d^* &= \max_{x \geq 0} \min_y \quad L(y, x) \\
&= \max_{x \geq 0} \min_{y_0^2 = 1} \quad \tfrac{1}{2} y^T (Q_0 + \sum_{i=1}^m x_i Q_i) y + t y_0^2 \\
&\qquad\qquad\qquad + y_0 y^T (b_0 + \sum_{i=1}^m x_i b_i) \\
&\qquad\qquad\qquad + (c_0 + \sum_{i=1}^m x_i c_i) - t \\
&= \max_{x \geq 0, t} \min_y \quad \tfrac{1}{2} y^T (Q_0 + \sum_{i=1}^m x_i Q_i) y + t y_0^2 \\
&\qquad\qquad\qquad + y_0 y^T (b_0 + \sum_{i=1}^m x_i b_i) \\
&\qquad\qquad\qquad + (c_0 + \sum_{i=1}^m x_i c_i) - t
\end{aligned}
$$

Hidden **SDP** Constraint in Lagrangian Dual

Hessian is $\succeq 0$

$$B := \begin{pmatrix} 0 & b_0^T \\ b_0 & Q_0 \end{pmatrix},$$

$$\mathcal{A} \begin{pmatrix} t \\ x \end{pmatrix} := - \begin{bmatrix} t & \sum_{i=1}^{m} x_i b_i^T \\ \sum_{i=1}^{m} x_i b_i & \sum_{i=1}^{m} x_i Q_i \end{bmatrix}, \quad : \mathbb{R}^{m+1} \to \mathbb{S}^{n+1}$$

and the **SDP** constraint

$$B - \mathcal{A} \begin{pmatrix} t \\ x \end{pmatrix} \succeq 0.$$

NOTE: There is NO hidden constraint needed in convex case; e.g. if all q_i are convex.

Lagrangian Relaxation and Equivalent **SDP**

Dual-Primal Programs

Lagrangian Relaxation is equivalent to **SDP** (with $c_0 = 0$)

$$
\begin{aligned}
d^* = \quad & \sup \quad -t + \sum_{i=1}^{m} x_i c_i \\
(\textbf{DSDP}) \qquad & \text{s.t.} \quad \mathcal{A} \begin{pmatrix} t \\ x \end{pmatrix} \preceq B \\
& x \in \Re^m, t \in \Re
\end{aligned}
$$

As in LP, Dual of Dual; Use Opt. Strategy of Competing Player

$$
\begin{aligned}
d^* \leq p^* := \quad & \inf \quad \text{trace } BY \\
(\textbf{DD}) \qquad & \text{s.t.} \quad \mathcal{A}^* Y = \begin{pmatrix} -1 \\ c \end{pmatrix} \\
& Y \succeq 0.
\end{aligned}
$$

Quadratic Assignment Problem, QAP

- Semidefinite programming, **SDP**, relaxations extremely strong for many hard discrete optimization problems; particularly true for quadratic assignment problem, **QAP**, one of the hardest NP-complete problems; even finding an ϵ-approximation is NP-complete, Sahni-Gonzales'76. (Nugent instance size $n = 30$ finally solved Anstreicher-Brixius'00 using Condor.)

- difficulties: large dimensional relaxations; inefficiency of the current primal-dual interior point solvers for time and accuracy; high expense in adding cutting plane constraints.

- we propose: alternating direction method of multipliers, **ADMM**, to solve the **SDP** relaxation along with nonnegativity constraints (cuts), i.e., we solve the **DNN** relaxation.

- Several instances are solved to optimality by the relaxation.

- We exploit <u>facial reduction, **FR**</u>; it fits well with **ADMM**.

What is the QAP?

The quadratic assignment problem, **QAP**, in the trace formulation

$$(\text{QAP}) \qquad p^* := \min_{X \in \Pi_n} \langle AXB - 2C, X \rangle,$$

$A, B \in \mathbb{S}^n$ real symmetric $n \times n$ matrices, C real $n \times n$,
$\langle \cdot, \cdot \rangle$ denotes trace inner product, $\langle Y, X \rangle = \text{trace } YX^\top$,
and Π_n set of $n \times n$ permutation matrices (permutations ϕ)

assign n facilities to n locations; minimize total cost

flow is A_{ij} between facilities i, j and it multiplies
distance $B_{\phi(i)\phi(j)}$ to get the total cost of assigning facilities i, j
to locations $\phi(i), \phi(j)$, respectively;
then add location costs in $-\frac{1}{2} \left(C_{i\phi(i)} + C_{j\phi(j)} \right)$

Applications Include: (e.g., Nyberg et al '2012)

Koopmans-Beckmann 1957

- facility location planning: Universities, hospital layout, airport gate assignment, wiring problems/circuit boards/VLSI, typewriter keyboards (though max?)
- Bandwith minimization of a graph
- Image processing
- Scheduling
- Supply Chains
- Economics
- Molecular conformations in chemistry
- Manufacturing lines
- Includes as special case: Traveling salesman problem and Maximum cut problem

Quadratic-Quadratic Model for $X \in \Pi$

$Xe = e, X^T e = e, X \geq 0$, doubly stochastic; ($e$ – ones vector)
turn linear constraints into quadratic

Start with Quadratic-Quadratic Model for $X \in \Pi$, a QQP

$$\min_X \quad \langle AXB - 2C, X \rangle$$
$$\text{s.t.} \quad \|Xe - e\|^2 + \|X^T e - e\|^2 = 0 \quad \text{(r-c sums)}$$
$$XX^T = X^T X = I_n \quad \text{(orthogonality)}$$
$$X_{ij}X_{ik} = 0, \ X_{ji}X_{ki} = 0, \ \forall i, \ \forall j \neq k, \quad \text{(gangster)}$$
$$X_{ij}^2 - X_{ij} = 0, \ \forall i, j, \quad \text{(0 − 1)}$$
$$X \geq 0 \quad \text{(nonnegativity)}$$

Dual of Dual is SDP Relaxation

The Lagrangian dual is an SDP.
The (Lagrangian) dual of this SDP is equivalent to the SDP
relaxation of the QQP. BUT, strict feasibility (Slater) fails!

Relaxations for QAP, e.g., Finke-Burkhard-Rendl '87; Rendl-W '87 (and others)

Eigenvalue Bound (apply Hoffman-Wielandt inequality)

$$\min_X \quad \langle AXB, X \rangle$$
$$\text{s.t.} \quad XX^T = I_n$$
$$(X^T X = I_n \text{ NOT redundant in Lagr. relax.})$$

Significantly Stronger Projected Eigenvalue Bound Hadley-Rendl-W. '89

(Used by Brixius-Anstreicher '01 to 'finally' solve Nugent $n = 30$, with help from Condor.)

$$\min_X \quad \langle AXB, X \rangle$$
$$\text{s.t.} \quad XX^T = I_n, \quad Xe = e, X^T e = e$$

More Significantly Stronger but Expensive SDP Bound

FR - SDP bound Zhang-Karisch-Rendl-W. '98

New Derivation of **FR**, **SDP** Relax. in ZKRW, '98

Start new derivation with **QQP** (with fewer constraints)

$\min_X \quad \langle AXB - 2C, X \rangle$
$\text{s.t.} \quad X_{ij}X_{ik} = 0, \; X_{ji}X_{ki} = 0, \; \forall i, \forall j \neq k,$ (gangster)
$\qquad X_{ij}^2 - X_{ij} = 0, \; \forall i, j,$ (0 − 1)
$\qquad \sum_{i=1}^n X_{ij}^2 - 1 = 0, \; \forall j, \; \sum_{j=1}^n X_{ij}^2 - 1 = 0, \; \forall i.$ (r-c sums)

Gangster constraints

The first set of constraints, the elementwise orthogonality of the row and columns of X, are the gangster constraints. They are particularly strong constraints and enable many of the other constraints (such as orthogonality $XX^T = I$, $X^T X = I$, row and columns sums are 1) to be redundant. In fact, after the facial reduction, **FR**, many of these constraints also become redundant.

The Lagrangian Dual (quadratic function!)

Lagrangian

$$
\begin{aligned}
\mathcal{L}_0(X, U, V, W, u, v) &= \langle AXB - 2C, X \rangle + \\
&\sum_{i=1}^{n} \sum_{j \neq k} U_{jk}^{(i)} X_{ij} X_{ik} + \\
&\sum_{i=1}^{n} \sum_{j \neq k} V_{jk}^{(i)} X_{ji} X_{ki} + \\
&\sum_{i,j} W_{ij}(X_{ij}^2 - X_{ij}) + \\
&\sum_{j=1}^{n} u_j \left(\sum_{i=1}^{n} X_{ij}^2 - 1 \right) + \\
&\sum_{i=1}^{n} v_i \left(\sum_{j=1}^{n} X_{ij}^2 - 1 \right).
\end{aligned}
$$

Dual problem is maximization of dual functional d_0

$$
\max \; d_0(U, V, W, u, v) := \min_X \mathcal{L}_0(X, U, V, W, u, v)
$$

hidden/implicit constraints for inner minimization - Hessian is positive semidefinite, i.e. SDP constraints arise.

Simplify the Dual by Homogenization with $x_0, y = (x_0, x)$

$$
\begin{aligned}
\mathcal{L}_1(X, x_0, U, V, W, w_0, u, v) \;=\; & y^\top \left[L_Q + \mathcal{B}_1(U) + \mathcal{B}_2(V) + \right. \\
& \text{Arrow}(w, w_0) + \mathcal{K}_1(u) + \\
& \left. \mathcal{K}_2(v) \right] y - e^\top(u + v) - w_0,
\end{aligned}
$$

where
$$
\mathcal{K}_1(u) = \text{blkdiag}(0, u \otimes I), \quad \mathcal{K}_2(v) = \text{blkdiag}(0, I \otimes v)
$$
$$
\text{Arrow}(w, w_0) = \begin{bmatrix} w_0 & -\frac{1}{2}w^\top \\ -\frac{1}{2}w & \text{Diag}(w) \end{bmatrix}
$$

$$
\mathcal{B}_1(U) = \text{blkdiag}(0, \tilde{U}), \quad \mathcal{B}_2(V) = \text{blkdiag}(0, \tilde{V}).
$$

And, $\tilde{U}$ and $\tilde{V}$ are $n \times n$ block matrices.

SDP Version of Lagrangian Dual

We let $L_Q := \begin{bmatrix} 0 & -\operatorname{vec}(C)^\top \\ -\operatorname{vec}(C) & B \otimes A \end{bmatrix}$.

$$\max \ -e^\top(u+v) - w_0$$
$$\text{s.t. } L_Q + \mathcal{B}_1(U) + \mathcal{B}_2(V) + \operatorname{Arrow}(w, w_0) + \mathcal{K}_1(u) + \mathcal{K}_2(v) \succeq 0.$$

where $X \succeq 0$ denotes $X \in \mathbb{S}^n_+$, is $n \times n$ positive semidefinite

Now take the dual of the dual to get **SDP** relaxation.
(Note the similarity to a linear program, LP.)

$$\min \ \operatorname{trace} L_Q Y$$
$$\text{s.t. } \operatorname{adjoints}(Y) = \operatorname{RHS}$$
$$Y \succeq 0$$

FR : $Y = \widehat{V} R \widehat{V}^T$ Greatly Simplifies **SDP** Relaxation

Dual of Dual after **FR**

$$(SDP_R) \qquad \begin{aligned} p_R^* := \min_R \quad & \langle L_Q, \widehat{V} R \widehat{V}^\top \rangle \\ \text{s.t.} \quad & \mathcal{G}_J(\widehat{V} R \widehat{V}^\top) = E_{00} \\ & R \succeq 0, \end{aligned}$$

- Gangster operator $\mathcal{G}$ *shoots* holes in the matrix $Y = \widehat{V} R \widehat{V}^\top$.
- J is the index set that guarantees the diagonal blocks are diagonal and the off-diagonal blocks have zero diagonal.
- (Some of these block constraints are redundant as are the previous block constraints.)
- with V (now) providing an orthonormal basis for $e^\perp$ and

$$\widehat{V} = \begin{bmatrix} 1 & 0 \\ \frac{1}{n} e & V \otimes V \end{bmatrix} \qquad \text{(does facial reduction, } \mathbf{FR})$$

Explicit Primal-Dual Strictly Feasible Points

> **Lemma (ZKRW , Explicit Primal Strictly Feasible Point)**
>
> *The matrix $\widehat{R}$ defined by*
>
> $$\widehat{R} := \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \frac{1}{n^2(n-1)} \left(nI_{n-1} - E_{n-1}\right) \otimes \left(nI_{n-1} - E_{n-1}\right) \end{array} \right] \in \mathcal{S}_{++}^{(n-1)^2+1}$$
>
> *is (strictly) feasible for (SDP_R).* □

Dual

Dual Program

$$d_Y^* := \max_Y \quad \langle E_{00}, Y \rangle \qquad\qquad (= Y_{00})$$
$$\text{s.t.} \quad \widehat{V}^\top \mathcal{G}_J(Y) \widehat{V} \preceq \widehat{V}^\top L_Q \widehat{V}$$

Lemma (ZKRW , Explicit Dual Strictly Feasible Point)

The matrices $\widehat{Y}, \widehat{Z}$, with $M > 0$ sufficiently large,

$$\widehat{Y} := M \left[\begin{array}{c|c} n & 0 \\ \hline 0 & I_n \otimes (I_n - E_n) \end{array} \right] \in \mathcal{S}_{++}^{(n-1)^2+1},$$

$$\widehat{Z} := \widehat{V}^\top L_Q \widehat{V} - \widehat{V}^\top \mathcal{G}_J(\widehat{Y}) \widehat{V} \in \mathcal{S}_{++}^{(n-1)^2+1},$$

are (strictly) feasible and slack variables for the dual, respectively.

New **ADMM** Algorithm for the **SDP** Relaxation

KEEP! both Y (for nonneg. and gangster) and R (for $\succeq 0$)

rewrite SDP_R equivalently as

$$\min_{R,Y} \left\{ \langle L_Q, Y \rangle \text{ s.t. } \mathcal{G}_J(Y) = E_{00},\ Y = \widehat{V} R \widehat{V}^\top,\ R \succeq 0 \right\}$$

Therefore we can work with the augmented Lagrangian

$$\mathcal{L}_A(R, Y, Z) = \langle L_Q, Y \rangle + \langle Z, Y - \widehat{V} R \widehat{V}^\top \rangle + \frac{\beta}{2} \| Y - \widehat{V} R \widehat{V}^\top \|_F^2.$$

(R, Y, Z) are the primal reduced, primal, and dual variables, and this denotes the current iterate

$\mathbb{S}_{r+}^n$ denotes the matrices in $\mathbb{S}_+^n$ with rank at most r.

ADMM with Augmented Lagrangian

Updates for (R_+, Y_+, Z_+):

1. $R_+ = \operatorname{argmin}_{R \in \mathbb{S}_{r+}^n} \mathcal{L}_A(R, Y, Z)$
2. $Y_+ = \operatorname{argmin}_{Y \in \mathcal{P}_i} \mathcal{L}_A(R_+, Y, Z)$
3. $Z_+ = Z + \gamma \cdot \beta(Y_+ - \widehat{V} R_+ \widehat{V}^\top)$

Polyhedral Sets

$\mathcal{P}_1 = \{Y \in \mathbb{S}^{n^2+1} : \mathcal{G}_J(Y) = E_{00}\}$ gangster constraints.
$\mathcal{P}_2 = \mathcal{P}_1 \cap \{0 \leq Y \leq 1\}$ (polytope with nonnegativity)

1. Explicit solution for R

Let $\widehat{V}$ be normalized such that $\widehat{V}^\top \widehat{V} = I$. Then:

$$
\begin{aligned}
R_+ &= \operatorname{argmin}_{R \succeq 0} \langle Z, Y - \widehat{V} R \widehat{V}^\top \rangle + \frac{\beta}{2} \| Y - \widehat{V} R \widehat{V}^\top \|_F^2 \\
&= \operatorname{argmin}_{R \succeq 0} \left\| Y - \widehat{V} R \widehat{V}^\top + \frac{1}{\beta} Z \right\|_F^2 \quad \text{(complete square)} \\
&= \operatorname{argmin}_{R \succeq 0} \left\| R - \widehat{V}^\top (Y + \frac{1}{\beta} Z) \widehat{V} \right\|_F^2 \quad \text{(use } \widehat{V}^\top \widehat{V} = I) \\
&= \mathcal{P}_{\mathbb{S}_+^{(n-1)^2+1}} \left(\widehat{V}^\top (Y + \frac{1}{\beta} Z) \widehat{V} \right),
\end{aligned}
$$

where we then apply the Eckart-Young-Mirsky Theorem and project onto the face of the **SDP** cone of desired rank ($\leq$ number of positive eigenvalues of the argument).

2. Explicit solution for Y

$$
\begin{aligned}
Y_+ &= \operatorname{argmin}_{\mathcal{G}_J(Y)=E_{00}} \langle L_Q, Y \rangle + \langle Z, Y - \widehat{V} R_+ \widehat{V}^\top \rangle + \\
&\qquad \frac{\beta}{2} \| Y - \widehat{V} R_+ \widehat{V}^\top \|_F^2 \\
&= \operatorname{argmin}_{\mathcal{G}_J(Y)=E_{00}} \left\| Y - \widehat{V} R_+ \widehat{V}^\top + \frac{L_Q + Z}{\beta} \right\|_F^2 \\
&= E_{00} + \mathcal{G}_{J^c} \left(\widehat{V} R_+ \widehat{V}^\top - \frac{L_Q + Z}{\beta} \right).
\end{aligned}
$$

major advantage of using **ADMM** : we can easily add
$0 \le \widehat{V} R \widehat{V}^\top \le 1$ to solve the DNN!:

$$
p^*_{RY} := \min_{R,Y} \{ \langle L_Q, Y \rangle : \mathcal{G}_J(Y) = E_{00}, \, 0 \le Y \le 1, \, Y = \widehat{V} R \widehat{V}^\top, \, R \succeq 0 \}
$$

Update Y_+ Becomes:

$$
Y_+ = E_{00} + \min \left(1, \max \left(0, \mathcal{G}_{J^c} (\widehat{V} R_+ \widehat{V}^\top - \frac{L_Q + Z}{\beta}) \right) \right)
$$

Lower bound from Inaccurate Solutions

$(R^{out}, Y^{out}, Z^{out})$ output

Find a lower bound from a feasible solution.
(For inaccurate optimality.)

Lemma

Let $\mathcal{R} := \{R \succeq 0\}, \quad \mathcal{Y} := \{Y : \mathcal{G}_J(Y) = E_{00}, 0 \leq Y \leq 1\}$
$\mathcal{Z} := \{Z : \widehat{V}^\top Z \widehat{V} \preceq 0\}$ and
*$g(Z) := \min_{Y \in \mathcal{Y}} \{\langle L_Q + Z, Y \rangle\}$, be the **ADMM** dual function.*
*Then the dual of **ADMM** satisfies weak duality and is:*

$$
\begin{aligned}
d_Z^* &:= \max_{Z \in \mathcal{Z}} g(Z) \\
&\leq p_{RY}^* \quad \text{(optimal value of \textbf{ADMM})}
\end{aligned}
$$

Proof

Proof.

The dual problem can be derived as

$$
\begin{aligned}
d_Z^* &:= \max_Z \min_{R \in \mathcal{R}, Y \in \mathcal{Y}} \langle L_Q, Y \rangle + \langle Z, Y - \widehat{V} R \widehat{V}^\top \rangle \\
&= \max_Z \min_{Y \in \mathcal{Y}} \langle L_Q, Y \rangle + \langle Z, Y \rangle + \min_{R \in \mathcal{R}} \langle Z, -\widehat{V} R \widehat{V}^\top \rangle \\
&= \max_Z \min_{Y \in \mathcal{Y}} \langle L_Q, Y \rangle + \langle Z, Y \rangle + \min_{R \in \mathcal{R}} \langle \widehat{V}^\top Z \widehat{V}, -R \rangle \\
&= \max_{Z \in \mathcal{Z}} \min_{Y \in \mathcal{Y}} \langle L_Q + Z, Y \rangle, \\
&= \max_{Z \in \mathcal{Z}} g(Z)
\end{aligned}
$$

Weak duality follows by exchanging the max and min. $\qquad\square$

Implementation: Lower bound from Inaccurate Solutions

- $Z \in \mathcal{Z} \implies g(Z)$ is a lower bound; therefore use projection $g(\mathcal{P}_{\mathcal{Z}}(Z^{out}))$ as lower bound,

- to get $\mathcal{P}_{\mathcal{Z}}(\tilde{Z})$: Let $\bar{V} = (\widehat{V}, \widehat{V}_\perp)$ be an orthogonal matrix; let
$\bar{V}^\top Z \bar{V} = W = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}$; Then:

$$\widehat{V}^\top Z \widehat{V} \preceq 0 \Leftrightarrow \widehat{V}^\top Z \widehat{V} = \widehat{V}^\top \bar{V} W \bar{V}^\top \widehat{V} = W_{11} \preceq 0.$$

Hence,

$$\begin{aligned}
\mathcal{P}_{\mathcal{Z}}(\tilde{Z}) &= \mathsf{argmin}_{Z \in \mathcal{Z}} \, \|Z - \tilde{Z}\|_F^2 \\
&= \mathsf{argmin}_{W_{11} \preceq 0} \, \|\bar{V} W \bar{V}^\top - \tilde{Z}\|_F^2 \\
&= \mathsf{argmin}_{W_{11} \preceq 0} \, \|W - \bar{V}^\top \tilde{Z} \bar{V}\|_F^2 \\
&= \begin{bmatrix} \mathcal{P}_{\mathcal{S}_-}(\tilde{W}_{11}) & \tilde{W}_{12} \\ \tilde{W}_{21} & \tilde{W}_{22} \end{bmatrix},
\end{aligned}$$

Upper Bound from Feasible Solution

$(R^{out}, Y^{out}, Z^{out})$ output of **ADMM**

- obtain best rank-one approximation of Y from largest eigenvalue and corresponding eigenvector: $\lambda vv^{\top}$.
- Reshape to get square matrix X^{out} as an approximate optimal permutation matrix.
- Since X permutation matrix implies trace $X^T X = n$, a constant, we get

$$\|X^{out} - X\|_F^2 = -2\operatorname{trace} X^T X^{out} + \text{ constant.}$$

Take advantage of the Birkoff, von Neumann Theorem: permutation matrices are extreme points of the doubly stochastic matrices.

- Solve the linear program

$$\max_{X} \left\{ \langle X^{out}, X \rangle : Xe = e, \ X^{\top}e = e, \ X \geq 0 \right\}$$

Low Rank Solutions

CHEAT on Rank (Apply Eckert/Young Theorem)

Project R onto a rank-one matrix,

$$R_+ = \mathcal{P}_{\mathcal{S}_+^{(n-1)^2+1} \cap \mathcal{R}_1} \left(\widehat{V}^\top \left(Y + \frac{Z}{\beta} \right) \widehat{V} \right),$$

where $\mathcal{R}_1 = \{R : \operatorname{rank}(R) = 1\}$ denotes the set of rank-one matrices. For a symmetric matrix W with largest eigenvalue $\lambda > 0$ and corresponding eigenvector w, we have

$$\mathcal{P}_{\mathcal{S}_+^{(n-1)^2+1} \cap \mathcal{R}_1} = \lambda w w^\top.$$

Often provides better feasible solutions/upper bounds.
PROVED OPTIMALITY in 7 instances.

Different choices for $V, \widehat{V}$

matrix $\widehat{V}$ is essential; sparse $\widehat{V}$ helps in projection using a sparse eigenvalue code. From several, the most successful (from Pong, Sun, Wang, W. '14):

$$
V = \left[\begin{bmatrix} \begin{bmatrix} I_{\lfloor \frac{n}{2} \rfloor} \otimes \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \end{bmatrix} \\ 0_{(n-2\lfloor \frac{n}{2} \rfloor), \lfloor \frac{n}{2} \rfloor} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} I_{\lfloor \frac{n}{4} \rfloor} \otimes \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \end{bmatrix} \\ 0_{(n-4\lfloor \frac{n}{4} \rfloor), \lfloor \frac{n}{4} \rfloor} \end{bmatrix} \begin{bmatrix} \dots \end{bmatrix} \begin{bmatrix} \widehat{V} \end{bmatrix} \right]_{n \times n-1}
$$

i.e., the block matrix consisting of t blocks formed from Kronecker products along with one block $\widehat{V}$ to complete the appropriate size so that $V^{\top}V = I_{n-1}$, $V^{\top}e = 0$.

Problem Data			Numerical Results					Timing		
#	name	true-opt	lbd	ubd	rel.gap	rel.opt.gap	rel.gap^A	iter	iter^A	time(sec)
1	chr12a	9552	9548	9552	0.04	0	0.02	11500	24800	130.04
2	chr12b	9742	9742	9742	**0**	0	0.08	10300	26700	113.96
3	chr12c	11156	11156	11156	**0**	0	0	1600	19400	17.41
4	chr15a	9896	9896	9896	**0**	0	0.28	6700	30900	126.20
5	chr15b	7990	7990	7990	**0**	0	0.03	3500	20300	70.67
6	chr15c	9504	9504	9504	**0**	0	0.08	1800	20000	28.53
7	chr18a	11098	11098	11098	**0**	0	0	2000	20600	61.64
8	chr18b	1534	1534	1794	**15.62**	15.62	75.22	5558	12600	172.94
9	chr20a	2192	2192	2192	**0**	0	0.18	3700	33700	156.45
10	chr20b	2298	2298	2298	**0**	0	0	1200	26200	58.09
11	chr20c	14142	14136	14142	0.04	0	0.15	30900	33700	1325.01
12	els19	17212548	17208748	17212548	**0.02**	0	0.35	30800	40000	1106.23
13	esc16a	68	64	74	**14.39**	8.39	41.72	399	597	10.22
14	esc16b	292	290	292	**0.69**	0	6.01	302	386	6.89
15	esc16c	160	154	166	**7.48**	3.67	34.32	399	896	8.58
16	esc16d	16	14	16	**12.90**	0	118.18	299	659	4.96
17	esc16e	28	28	28	**0**	0	69.05	100	556	3.03
18	esc16f	0	0	0	**0**	0	0	1	1	0.02
19	esc16g	26	26	28	**7.27**	7.27	69.23	300	695	6.88
20	esc16h	996	978	1100	**11.74**	9.92	31.90	1362	609	28.75
21	esc16i	14	12	14	**14.81**	0	101.96	1016	2044	25.15
22	esc16j	8	8	8	**0**	0	82.76	100	799	2.11
23	had12	1652	1652	1652	**0**	0	0	300	11600	3.92
24	had14	2724	2724	2724	**0**	0	0	400	20300	5.52
25	had16	3720	3720	3720	**0**	0	0	600	18100	12.28
26	had18	5358	5358	5358	**0**	0	0.02	1300	34700	40.66
27	had20	6922	6922	6922	**0**	0	0.13	2300	40000	106.96
28	nug12	578	568	728	**24.67**	22.95	27.86	1416	2884	15.70
29	nug14	1014	1012	1022	**0.98**	0.79	1.08	2832	19600	44.65
30	nug15	1150	1142	1280	**11.39**	10.70	16.33	2161	5812	40.45
31	nug16a	1610	1600	1610	**0.62**	0	0.62	6217	19300	138.71
32	nug16b	1240	1220	1258	**3.07**	1.44	25.41	3454	2347	80.00
33	nug17	1732	1708	1756	**2.77**	1.38	2.77	6194	6401	159.42
34	nug18	1930	1894	2022	**6.54**	4.65	12.84	9555	3988	285.40
35	nug20	2570	2508	2702	**7.45**	5.01	18.43	7065	2386	266.59
36	rou12	235528	235528	235528	**0**	0	0	3700	34200	35.98
37	rou15	354210	350216	360702	**2.95**	1.82	4.89	2531	3946	39.94
38	rou20	725522	695180	781532	**11.70**	7.43	14.93	7099	1538	281.71
39	scr12	31410	31410	31410	**0**	0	19.38	400	4268	3.93
40	scr15	51140	51140	51140	**0**	0	21.96	700	5489	12.48
41	scr20	110030	106804	132826	**21.72**	18.77	43.71	11599	9705	425.22
42	tai10a	135028	135028	135028	**0**	0	0.01	1200	21400	5.95
43	tai12a	224416	224416	224416	**0**	0	0	300	4300	2.68
44	tai15a	388214	377100	403890	**6.86**	3.96	9.03	2644	2245	39.96
45	tai17a	491812	476526	534328	**11.44**	8.29	16.25	2940	1399	64.67
46	tai20a	703482	671676	762166	**12.62**	8.01	19.03	3733	999	136.38

QAPLIB: Instance of Medium Size

Problem Data			Numerical Results					Timing		
#	name	true-opt	lbd	ubd	rel.gap	rel.opt.gap	rel.gap^A	iter	iter^A	time(sec)
47	chr22a	6156	6156	6156	**0**	0	0.02	11500	40000	613.03
48	chr22b	6194	6190	6194	**0.06**	0	0.11	13500	39300	673.22
49	chr25a	3796	3796	3796	**0**	0	0	6200	35600	450.22
50	esc32a	130	104	168	46.89	25.42	106.90	15100	12400	2553.03
51	esc32b	168	132	220	49.86	26.74	92.49	1000	4144	167.59
52	esc32c	642	616	642	**4.13**	0	23.23	2500	2052	418.83
53	esc32d	200	192	220	**13.56**	9.50	41.08	670	1430	117.00
54	esc32e	2	2	18	152.38	152.38	152.38	700	3086	112.26
55	esc32g	6	6	12	**63.16**	63.16	121.21	500	999	81.81
56	esc32h	438	426	452	**5.92**	3.14	30.14	6500	17600	1097.87
57	kra30a	88900	86838	96430	**10.47**	8.13	15.91	9898	3799	1319.97
58	kra30b	91420	87858	101640	**14.55**	10.59	28.84	5480	5017	750.38
59	kra32	88700	85776	93050	**8.14**	4.79	30.03	4959	4173	870.14
60	nug21	2438	2382	2644	**10.42**	8.11	12.36	6439	5729	274.09
61	nug22	3596	3530	3678	**4.11**	2.25	12.76	7279	7573	359.10
62	nug24	3488	3402	3770	**10.26**	7.77	16.25	4543	4447	294.82
63	nug25	3744	3626	3966	**8.96**	5.76	15.37	11687	7799	864.25
64	nug27	5234	5130	5496	**6.89**	4.88	17.08	10039	8609	1010.56
65	nug28	5166	5026	5676	**12.15**	9.41	18.55	8387	7533	943.84
66	nug30	6124	5950	6610	**10.51**	7.63	20.21	11321	9036	1581.33
67	ste36a	9526	9260	9980	**7.48**	4.65	42.28	19500	27300	5262.87
68	ste36b	15852	15668	15932	**1.67**	0.50	82.03	29000	40000	7889.04
69	ste36c	8239110	8134808	8394142	**3.14**	1.86	36.15	36499	40000	9819.15
70	tai25a	1167256	1096656	1264590	**14.22**	8.00	20.55	2264	999	164.11
71	tai30a	1818146	1706872	1984536	**15.04**	8.75	15.21	4550	1599	623.39
72	tai35a	2422002	2216646	2625284	**16.88**	8.06	22.34	3161	1599	777.17
73	tai40a	3139370	2843310	3455540	**19.44**	9.59	23.43	5577	2299	5546.57
74	tho30	149936	143576	166336	**14.69**	10.37	24.33	8321	7729	1122.28
75	tho40	240516	226522	257642	**12.86**	6.88	25.19	15535	12460	17832.61

QAPLIB: Instance of Large Size

Problem Data			Numerical Results					Timing		
#	name	true-opt	lbd	ubd	rel.gap	rel.opt.gap	rel.gap^A	iter	iter^A	time(sec)
76	esc64a	116	98	260	90.25	76.39	80.97	400	1200	1085.52
77	sko42*	15812	15336	16244	**5.75**	2.70	17.24	5511	10700	6245.96
78	sko49*	23386	22654	24406	**7.45**	4.27	16.87	9484	16900	12213.03
79	sko56*	34458	33390	36468	**8.81**	5.67	15.92	5792	15100	11669.07
80	sko64*	48498	47022	50762	**7.65**	4.56	16.15	10021	21100	23033.17
81	tai50a*	4938796	4390980	5517228	**22.73**	11.06	25.79	2331	3300	1238.71
82	tai60a*	7205962	6326344	7895180	**22.06**	9.13	26.03	3799	5100	4939.96
83	tai64c	1855928	1811354	1887500	**4.12**	1.69	38.79	800	2400	1461.00
84	wil50*	48816	48126	50834	**5.47**	4.05	9.37	5384	11000	2971.40

$n = 150$

$n = 150$: X is 150×150
Y is $150^2 + 1 \times 150^2 + 1$ symmetric $150^2 + 1 = N = 22500$
$t(N) = (506295002 * 506295001)/2 = 253136250$
$253,136,250$ variables

Conclusion

- We presented ADMM framework for QAP that exploits facial reduction. (Keeping both R and Y appears to be advantageous.)

- solve large problems to extremely high accuracy while solving the DNN relaxation. This yielded improved bounds. Several problems were solved to optimality.

- The ADMM approach together with facial reduction appears to be very promising for future applications to hard discrete optimization problems.

References I

N. Graham, H. Hu, H. Im, X. Li, and H. Wolkowicz, *A restricted dual Peaceman-Rachford splitting method for a strengthened DNN relaxation for QAP*, INFORMS J. Comput. **34** (2022), no. 4, 2125–2143. MR 4485042

D.E. Oliveira, H. Wolkowicz, and Y. Xu, *ADMM for the SDP relaxation of the QAP*, Math. Program. Comput. **10** (2018), no. 4, 631–658. MR 3863706

Applications to Hard Discrete Optimization and Low Rank Solutions:
SDP Relaxations of **QQP**s; Solving **QAP**s

Henry Wolkowicz

Dept. Combinatorics and Optimization, University of Waterloo, Canada

Tues. Feb. 28, 2023,