Hard Combinatorial Problems, DNN Relaxations, Facial Reduction, and ADMM

Henry Wolkowicz

Dept. Combinatorics and Optimization, University of Waterloo, Canada

Friday, October 18, 2019, 8:30-9:10 AM
at: MOM21: The 21st Midwest Optimization Meeting
Northern Illinois University, DeKalb, Illinois

Collaborators

The work on Min-Cut is With:

Xinxin Li (Jilin University)
Ting Kei Pong (The Hong Kong Polytechnic University)
Hao Sun (University of Waterloo)

- Solving hard combinatorial/discrete optimization problems requires: efficient upper/lower bounding techniques.
- These problems are often modelled using quadratic objectives and/or quadratic constraints, i.e., QQPs.
- Lagrangian relaxations of QQPs lead to Semidefinite Programming, SDP, and SDP relaxations, e.g., Handbook on SDP [10].
- SDP relaxations are expensive to solve using interior-point approaches. This becomes *doubly* expensive when cutting planes are added, e.g., using Doubly Nonnegative, DNN, relaxations
- Strict feasibility fails for many of the SDP relaxations of these hard combinatorial problems. Facial reduction, FR, e.g., [1, 2, 3, 4] provides a means of regularizing the SDP relaxations.

- FR appears to provide a natural splitting of variable for the application of Alternating Direction Method of Multipliers, ADMM type methods; and for exploiting structure.
- Classes of Problems: Min-Cut Rendl, Lisser, Piacentini '13 [8]; Pong, Sun, Wang, W. '16 [7]; Rendl, Sotirov '16 [9], Maxcut and GP

  and QAP, Oliveira, W., Xu, '18 [5]
- Numerics (Large Scale)

Hard Combinatorial Problems

Instance /Modelling with Quadratic Functions

$$\begin{array}{rll} \min & q_0(x) & (= x^T Hx + 2g^T x + \alpha) \\ \text{s.t.} & Ax = b & \text{(linear constraint)} \\ & x \in K \subseteq \mathbb{R}^N & (K \text{ hard constraints}) \end{array}$$

Hard Constraints: e.g.,

- K is zero-one or ± 1;
  modelled with quadratic constraints, respectively,
  $\forall i : \quad q_i(x) = x_i^2 - x_i = 0$ or $q_i(x) = x_i^2 - 1 = 0$
- K is partition matrices, $x \in \mathcal{M}_m$, (GP)
- K is permutation matrices, $x \in \Pi_n$, (QAP)

Close Duality Gap

Duality Gap for QP

A Lagrangian duality gap can happen, e.g.,

$1 = \max\{-x_1^2+x_2^2 : x_2 = 1\} < \infty = \inf_\lambda \max_x -x_1^2+x_2^2-\lambda(x_2-1)$

BUT:

$$1 = \max\{-x_1^2 + x_2^2 : \boxed{(x_2-1)^2 = 0}\}$$
$$= \inf_\lambda \max_x\{-x_1^2 + x_2^2 - \lambda(x_2-1)^2\}$$

QP: Close Gap, Strong Duality (A $m \times n$, $m < n$, K compact)

Theorem (Poljak, Rendl, W. '95, [6])

$$\begin{aligned} p^* &= \max_x\{q_0(x) := x^T Hx + 2g^T x + \alpha : Ax = b, x \in K\} \\ &= \max_x\{q_0(x) : \|Ax-b\|^2 = 0, x \in K\} \\ &= \boxed{\min_\lambda} \max_x\{q_0(x) - \lambda\|Ax-b\|^2, x \in K\} \end{aligned}$$

□

Model with Quadratics, Homogenize, and Lift

Homogenize

$$x_0^2 - 1 = 0 : \quad \begin{cases} \min q_0(x, x_0) = x^T H x + 2 g^T x x_0 + \alpha x_0^2 \\ \quad Ax - b = 0 \quad \cong \quad \|Ax - bx_0\|_2^2 = 0 \end{cases}$$

Lifting (linearization): $\mathbb{R}^{N+1} \rightarrow \mathbb{S}^{N+1}$

$$y = \begin{pmatrix} x_0 \\ x \end{pmatrix}, \; Y = yy^T \in \mathbb{S}_+^{N+1}, \quad \text{symmetric, psd,} \quad Y_{00} = 1$$

$$\text{obj. fn.} \quad y^T \begin{bmatrix} \alpha & g^T \\ g & H \end{bmatrix} y = \operatorname{trace} \begin{bmatrix} \alpha & g^T \\ g & H \end{bmatrix} Y, \qquad \operatorname{rank}(Y) = 1$$

Relaxation

Discard rank one constraint on Y

Facial Reduction, FR

Lifting Linear Equality Constraint

$$
\begin{aligned}
0 &= \|Ax - bx_0\|_2^2 = \left\| \begin{bmatrix} -b & A \end{bmatrix} \begin{pmatrix} x_0 \\ x \end{pmatrix} \right\|_2^2 \\
&= \begin{pmatrix} x_0 \\ x \end{pmatrix}^T \begin{bmatrix} -b^T \\ A^T \end{bmatrix} \begin{bmatrix} -b & A \end{bmatrix} \begin{pmatrix} x_0 \\ x \end{pmatrix} \\
&= \operatorname{trace} \begin{bmatrix} \|b\|^2 & -b^T A \\ -A^T b & A^T A \end{bmatrix} Y = 0
\end{aligned}
$$

Exposing Vector, $W \in \mathbb{S}_+^{N+1}$, with spectral decomp., and FR

$$
W := \begin{bmatrix} \|b\|^2 & -b^T A \\ -A^T b & A^T A \end{bmatrix} = \begin{bmatrix} V & U \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} V & U \end{bmatrix}^T, \ D \in \mathbb{S}_+^{N+1-r}
$$

$$
\begin{aligned}
Y \text{ feasible} \quad &\implies \quad YW = 0 \quad \text{(Strict feasibiliity (Slater) fails)} \\
&\implies \quad Y = VRV^T, \ R \in \mathbb{S}_+^r \quad \text{(facial reduction)}
\end{aligned}
$$

Hard Discrete Constraints

Zero-One

$$x_0^2 - 1 = 0 : \; \{ \; q_i(x, x_0) = x_i^2 - x_i x_0 = 0, \; \forall i$$

Lifting (linearization): $\mathbb{R}^{N+1} \to \mathbb{S}^{N+1}$

$$y = \begin{pmatrix} x_0 \\ x \end{pmatrix}, \; Y = yy^T \in \mathbb{S}_+^{N+1}, \quad \text{symmetric, psd,} \quad Y_{00} = 1$$

$$\text{constr.:} \qquad \operatorname{arrow}(Y) = e_0 := \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in \mathbb{R}^{n+1}$$

$$(\operatorname{diag}(Y) = Y_{:,0})$$

Adjoint: Arrow $\cong$ arrow*

$$\langle \operatorname{Arrow}(v), S \rangle = \langle v, \operatorname{arrow}(S) \rangle, \quad \forall v \in \mathbb{R}^{N+1}, \forall S \in \mathbb{S}^{N+1}$$

Natural Splitting? $Y \in \mathcal{P}$, $R \in \mathbb{S}^r_+$

$Y = VRV^T$

$$Y \in \mathcal{P} \subset \mathbb{S}^{N+1}_+, \qquad R \in \mathbb{S}^r_+, \quad r < N+1$$

Facial reduction generally provides a reduction in dimension and a guarantee that strict feasibility holds.
There is a natural separation of constraints where

$$Y \in \mathcal{P} \text{ polyhedral} \qquad R \in \mathbb{S}^r_+ \text{ sdp cone}$$

Instance: Minimum Cut, MC, Problem

Given: Undirected Graph $G = (\mathcal{V}, \mathcal{E})$

edge set $\mathcal{E}$ and node set $|\mathcal{V}| = n$
$m = (m_1 \ m_2 \ \dots \ m_k)^T$, $\sum_{i=1}^{k} m_i = n$; given partition into k sets

MC Problem:

partition vertex set $\mathcal{V}$ into k subsets with given sizes in m
to *minimize the cut* after removing the k-th set;

Applications

re-orderings for sparsity patterns; microchip design and circuit board, floor planning and other layout problems.
($k = 3$, vertex separator problem)

(Graph Partitioning) Model for MC

Notation

A adjacency matrix of graph $G = (\mathcal{V}, \mathcal{E})$
e ones vector, $E = ee^T$

$$B = \begin{bmatrix} E - I_{k-1} & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{S}^k$$

$m = (m_1, \ldots, m_k)^T \in \mathbb{Z}^k_+, k > 2$, $n = |\mathcal{V}| = m^T e$.
$S = \{S_1, S_2, \ldots, S_k\}$ partition of vertex set, $|S_i| = m_i > 0, \forall i$
$M = \operatorname{Diag}(m) \qquad (m = \operatorname{Diag}^*(M) = \operatorname{diag}(M))$

Quadratic Program for MC

Notation

- the set of edges between two sets of nodes

$$\delta(S_i, S_j) := \{uv \in \mathcal{E} : u \in S_i, v \in S_j\}$$

- cut of a partition S

$$\delta(S) := \cup \left\{\delta(S_i, S_j) \,:\, 1 \le i < j \le k-1\right\}$$

- the set of partition matrices (cols of incidence vectors)

$$\mathcal{M}_m = \left\{X \in \mathbb{R}^{n\times k} : Xe = e,\ X^T e = m, X_{ij} \in \{0,1\}\right\}$$

$$X_{ij} = \begin{cases} 1 & \text{if } i \in S_j \\ 0 & \text{otherwise.} \end{cases}$$

- objective of MC: minimize cardinality of the cut $|\delta(S)|$:

$$\begin{array}{rcc} \mathrm{cut}(m) = & \min & \frac{1}{2}\,\mathrm{trace}\, AXBX^T \\ & \text{s.t.} & X \in \mathcal{M}_m, \end{array}$$

Quadratic-Quadratic Model

Include Many Redundant Constraints

$$
\begin{array}{rclll}
\mathrm{cut}(m) & = & \min & \frac{1}{2}\,\mathrm{trace}\, AXBX^T & \\
 & & \text{s.t.} & X \circ X = x_0 X & \in \{0,1\} \\
 & & & \|Xe - x_0 e\|^2 = 0 & \text{row sums } = 1 \\
 & & & \|X^T e - x_0 m\|^2 = 0 & \text{column sums} \\
 & & & X_{:i} \circ X_{:j} = 0, \ \forall i \neq j & \text{col. elem. orth.} \\
 & & & X^T X - M = 0 & \text{scaled orth.} \\
 & & & \mathrm{diag}\,(XX^T) - e = 0 & \text{unit norm rows} \\
 & & & x_0 e_n^T X e_k - n = 0 & n \text{ vertices} \\
 & & & x_0^2 = 1 & \text{homog.}
\end{array}
$$

- e_j is the vector of ones of dimension j; $M = \mathrm{Diag}\,(m)$.
- $u \circ v$ Hadamard (elementwise) product.

Facial Reduction, FR

Lifting/Block Appropriately/ $x = \operatorname{vec}(X)$

$$Y = \begin{pmatrix} x_0 \\ x \end{pmatrix} \begin{pmatrix} x_0 \\ x \end{pmatrix}^T =: \begin{bmatrix} Y_{00} & Y_{0\,1:nk}^T \\ Y_{1:nk\,0} & \overline{Y} \end{bmatrix},$$

$$Y_{1:nk\,0} := \begin{bmatrix} Y_{(10)} \\ Y_{(20)} \\ \vdots \\ Y_{(k0)} \end{bmatrix}, \quad \overline{Y} := \begin{bmatrix} \overline{Y}_{(11)} & \overline{Y}_{(12)} & \cdots & \overline{Y}_{(1k)} \\ \overline{Y}_{(21)} & \overline{Y}_{(22)} & \cdots & \overline{Y}_{(2k)} \\ \vdots & \ddots & \ddots & \vdots \\ \overline{Y}_{(k1)} & \ddots & \ddots & \overline{Y}_{(kk)} \end{bmatrix}$$

Objective

$$\frac{1}{2} \operatorname{trace} AXBX^T = \frac{1}{2} \operatorname{trace} L_A Y, \text{ where } L_A := \begin{bmatrix} 0 & 0 \\ 0 & B \otimes A \end{bmatrix}.$$

SDP Constraints (lifting/linearization)

The arrow constraint

$$\operatorname{arrow}(Y) := \operatorname{diag}(Y) - \begin{bmatrix} 0 \\ Y_{1:nk\,0} \end{bmatrix} = e_0,$$

e_0 first (0-th) unit vector
(redundant in the final SDP relaxation)

DNN, doubly nonnegative

$$Y \in \mathrm{DNN} \cap \{Y \in \mathbb{S}^{nk+1} : 0 \leq Y \leq 1\}$$

DNN is doubly nonnegative cone, i.e., intersection of positive semidefinite cone and nonnegative orthant.

SDP Constraints cont...

Trace constraints (from linear equality constraints

$$\operatorname{trace} D_1 Y = 0, \qquad D_1 := \begin{bmatrix} n & -e_k^T \otimes e_n^T \\ -e_k \otimes e_n & (e_k e_k^T) \otimes I_n \end{bmatrix},$$

$$\operatorname{trace} D_2 Y = 0, \qquad D_2 := \begin{bmatrix} m^T m & -m^T \otimes e_n^T \\ -m \otimes e_n & I_k \otimes (e_n e_n^T) \end{bmatrix},$$

e_j vector of ones of dimension j; $D_i \succeq 0, i = 1, 2$; nullspaces of these matrices yield the facial reduction $Y = VRV^T$.

Block: trace, diagonal and off-diagonal

$$\begin{aligned} \mathcal{D}_t(Y) &:= \left(\operatorname{trace} \overline{Y}_{(ij)}\right) = M \in \mathbb{S}^k; \\ \mathcal{D}_d(Y) &:= \textstyle\sum_{i=1}^k \operatorname{diag} \overline{Y}_{(ii)} = e_n \in \mathbb{R}^n; \\ \mathcal{D}_o(Y) &:= \left(\textstyle\sum_{s\neq t} \left(\overline{Y}_{(ij)}\right)_{st}\right) = \hat{M} \in \mathbb{S}^k, \end{aligned}$$

where $\hat{M} := mm^T - M$.

SDP Constraints cont...

trace $Y = n + 1$; and Gangster constraints on Y

The Hadamard product and orthogonal type constraints lead to

gangster constraints

i.e., simple constraints that restrict elements to be zero (shoot holes in the matrix) and/or restrict entire blocks.

gangster and restricted gangster constraint on Y:

$$\mathcal{G}_H(Y) = 0,$$

for specific index sets H.

SDP Relaxation with Many (some redundant) Constraints

$$
\begin{array}{rll}
\operatorname{cut}(m) \geq p^*_{\mathrm{SDP}} := & \min & \frac{1}{2} \operatorname{trace} L_A Y \\
& \text{s.t.} & \operatorname{arrow}(Y) = e_0 \\
& & \operatorname{trace} D_1 Y = 0, \ \operatorname{trace} D_2 Y = 0 \\
& & \mathcal{G}_{J_0}(Y) = 0, \ Y_{00} = 1 \\
& & \mathcal{D}_t(Y) = M, \ \mathcal{D}_d(Y) = e, \ \mathcal{D}_o(Y) = \widehat{M} \\
& & Y \in \mathbb{S}^{kn+1}_+
\end{array}
$$

Equivalent FR greatly simplified SDP; with $Y = \widetilde{V} R \widetilde{V}^T$

$$
\begin{array}{rcll}
\operatorname{cut}(m) \geq p^*_{\mathrm{SDP}} & = & \min & \frac{1}{2} \operatorname{trace} \left(\widetilde{V}^T L_A \widetilde{V} \right) R \\
& & \text{s.t.} & \mathcal{G}_{\hat{J}_{\mathcal{I}}}(\widetilde{V} R \widetilde{V}^T) = \mathcal{G}_{\hat{J}_{\mathcal{I}}}(e_0 e_0^T) \\
& & & R \in \mathbb{S}^{(k-1)(n-1)+1}_+
\end{array}
$$

Primal-Dual Strong Duality (Regularity) for FR SDP

Theorem

1. *(Generalized) slater point for the primal:*
$$\widetilde{R} = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \frac{1}{n^2(n-1)}\left(n\mathrm{Diag}\left(\hat{m}_{k-1}\right) - \hat{m}_{k-1}\hat{m}_{k-1}^T\right) \otimes \left(nI_{n-1} - E_{n-1}\right) \end{array}\right] \in \mathbb{S}_{++}^{(k-1)(n-1)+1}.$$
Moreover, Robinson regularity holds.
2. *The dual problem*
$$\begin{array}{ll} \max & \frac{1}{2} w_{00} \\ \text{s.t.} & \widetilde{V}^T \mathcal{G}^*_{\hat{J}_{\mathcal{I}}}(w)\widetilde{V} \preceq \widetilde{V}^T L_A \widetilde{V}. \end{array}$$
satisfies strict feasibility.

Motivation

Difficulties for Primal-dual interior-point Methods for SDP

- solving large problems
- obtaining high accuracy solutions
- exploiting sparsity
- adding on nonnegativity and other cutting plane constraints

First order operator splitting methods for SDP

- FR provides a natural splitting, $Y = VRV^T$
- Flexibility in dealing with additional constraints
- separable/split optimization steps are inexpensive

Strengthen model with redundant constraint

Set Constraints

$$\begin{aligned}\mathcal{R} &:= \{R \in \mathbb{S}_+^{(k-1)(n-1)+1} : \text{trace } R = n+1\},\\ \mathcal{Y} &:= \{Y \in \mathbb{S}^{nk+1} : 1 \geq Y(J^c) \geq 0,\\ &\qquad \mathcal{G}_{\bar{J}}(Y) = \mathcal{G}_{\bar{J}}(e_0 e_0^T)\\ &\qquad\qquad \mathcal{D}_o(Y) = \widehat{M},\ e^T Y_{(i0)} = m_i, \forall i\}\end{aligned}$$

Strengthened model

$$(\text{DNN}) \quad p^*_{DNN} = \begin{array}{ll} \min & \frac{1}{2}\text{trace } L_A Y + \mathbb{1}_{\mathcal{Y}}(Y) + \mathbb{1}_{\mathcal{R}}(R) \\ \text{s.t.} & Y = \widehat{V} R \widehat{V}^T, \end{array}$$

where $\mathbb{1}_{\mathcal{S}}(\cdot)$ is indicator function of set $\mathcal{S}$.

Splitting Method

Augmented Lagrangian Function, $\mathcal{L}_\beta(R, Y, Z) =$

$$= f_{\mathcal{R}}(R) + g_{\mathcal{Y}}(Y) + \langle Z, Y - \widehat{V}R\widehat{V}^T \rangle + \frac{\beta}{2}\left\| Y - \widehat{V}R\widehat{V}^T \right\|^2$$

- $\beta > 0$ penalty parameter for quadratic penalty term,
- (L_s diagonally scaled objective $L_s := \frac{1}{2}L + \alpha I \succ 0$)

$$f_{\mathcal{R}}(R) = \mathbb{1}_{\mathcal{R}}(R), \quad g_{\mathcal{Y}}(Y) = \operatorname{trace} L_s Y + \mathbb{1}_{\mathcal{Y}}(Y).$$

sPRSM, Strictly Contractive Peaceman-Rachford Splitting

i.e., alternate minimization of $\mathcal{L}_\beta$ in the variables Y and R interlaced by an update of the Z variable.
In particular, we update the dual variable Z both after the R-update *and* the Y-update (both of which have unique solutions).

FRSMR, FR Splitting Method with Redundancies

- Pick any $Y^0, Z^0 \in \mathbb{S}^{nk+1}$. Fix $\beta > 0$ and $\gamma \in (0,1)$. Set $t = 0$.
- For each $t = 0, 1, \ldots$, update

$$
\begin{aligned}
\bullet R^{t+1} &= \mathrm{argmin}_{R\in\mathcal{R}}\, \mathcal{L}_\beta(R, Y^t, Z^t) \\
&= \mathrm{argmin}_R\, f_{\mathcal{R}}(R) - \langle Z^t, \widehat{V}R\widehat{V}^T\rangle + \frac{\beta}{2}\left\| Y^t - \widehat{V}R\widehat{V}^T \right\|^2 \\
\bullet Z^{t+\frac{1}{2}} &= Z^t + \gamma\beta(Y^t - \widehat{V}R^{t+1}\widehat{V}^T), \\
\bullet Y^{t+1} &= \mathrm{argmin}_{Y\in\mathcal{Y}}\, \mathcal{L}_\beta(R^{t+1}, Y, Z^{t+\frac{1}{2}}) \\
&= \mathrm{argmin}_Y\, g_{\mathcal{Y}}(Y) + \langle Z^{t+\frac{1}{2}}, Y\rangle + \frac{\beta}{2}\left\| Y - \widehat{V}R^{t+1}\widehat{V}^T \right\|^2, \\
\bullet Z^{t+1} &= Z^{t+\frac{1}{2}} + \gamma\beta(Y^{t+1} - \widehat{V}R^{t+1}\widehat{V}^T).
\end{aligned}
$$

Global convergence

Theorem

Let $\{R^t\}$, $\{Y^t\}$ and $\{Z^t\}$ be the generated sequences from FRSMR. *Then $\{(R^t, Y^t)\}$ converges to an optimal solution (R^*, Y^*) of the* DNN *relaxation, $\{Z^t\}$ converges to some Z^*, and (R^*, Y^*, Z^*) satisfies the optimality conditions of the* DNN *relaxation*

$$\begin{array}{rcl} 0 & \in & -\widehat{V}^T Z^* \widehat{V} + \mathcal{N}_{\mathcal{R}}(R^*), \\ 0 & \in & L_s + Z^* + \mathcal{N}_{\mathcal{Y}}(Y^*), \\ Y^* & = & \widehat{V} R^* \widehat{V}^T, \end{array}$$

where $\mathcal{N}_S(x)$ denotes the normal cone of S at x. $\square$

1. Explicit solution for R^{t+1}

With the assumption that $\widehat{V}^T \widehat{V} = I$

$$
\begin{aligned}
R^{t+1} &= \operatorname{argmin}_{R \in \mathcal{R}} -\langle Z, \widehat{V} R \widehat{V}^T \rangle + \frac{\beta}{2} \left\| Y^t - \widehat{V} R \widehat{V}^T \right\|^2 \\
&= \mathcal{P}_{\mathcal{R}}(\widehat{V}^T (Y^t + \tfrac{1}{\beta} Z^t) \widehat{V}),
\end{aligned}
$$

where $\mathcal{P}_{\mathcal{R}}$ denotes the projection (nearest point) onto the intersection of the SDP cone $\mathbb{S}_+^{(k-1)(n-1)+1}$ and the hyperplane $\{R \in \mathbb{S}^{(k-1)(n-1)+1} : \operatorname{trace} R = n+1\}$.

(diagonalize; then project eigenvalues onto simplex)

2. Explicit solution of Y^{t+1}

The Y-subproblem yields a closed form solution by projection onto the polyhedral set $\mathcal{Y}$, i.e.,

$$Y^{t+1} = \text{argmin}_{Y \in \mathcal{Y}} \frac{\beta}{2} \left\| Y - \boxed{\widehat{V} R^{t+1} \widehat{V}^T - \frac{1}{\beta}(L_s + Z^{t+\frac{1}{2}})} \right\|^2.$$

Note that the update (projection of $\tilde{Y}$) satisfies e.g.,

$$(Y^{t+1})_{ij} = \begin{cases} 1 & \text{if } i = j = 0 \\ 0 & \text{if } ij \in J \backslash \{00\} \\ 0 & \text{if } ij \in J^c,\ Y_{ij} \leq 0 \\ \tilde{Y}_{ij} & \text{if } ij \in J^c,\ 0 < Y_{ij}. \end{cases}$$

Lower bound from Inaccurate Solutions

Theorem (Fenchel Dual)

Define modified dual functional
$g(Z) := \min_{Y \in \widetilde{\mathcal{Y}}} \langle L_s + Z, Y \rangle - (n+1)\lambda_{\max}(\widehat{V}^T Z \widehat{V})$,
with $\widetilde{\mathcal{Y}} :=$

$$\{Y \in \mathbb{S}^{nk+1} : \mathcal{G}_{\widehat{J}_0}(Y) = \mathcal{G}_{\widehat{J}_0}(e_0 e_0^T),\ 0 \leq \mathcal{G}_{\widehat{J}_0^c}(Y) \leq 1,$$
$$\mathcal{D}_o(Y) = \widehat{M},\ \mathcal{D}_t(Y) = M,\ e^T Y_{(i0)} = m_i, i = 1, \ldots, k\}.$$

Then

$$p^*_{\mathrm{DNN}} = d^*_Z := \max_Z g(Z),$$

and the latter (dual) problem is attained, i.e., strong duality holds. $\square$

The Lower Bound

Evaluating $g(Z^t)$ always yields a lower bound for the DNN relaxation optimal value

$$p^*_{\mathrm{DNN}} \geq g(Z^t)$$

Upper bound from feasible solution

Approx. output Y^{out}

- Obtain a vector $v = (v_0 \; \bar{v})^T \in \mathbb{R}^{nk+1}$, $v_0 \neq 0$ from Y^{out}
- Reshape $\bar{v}$; get $n \times k$ matrix X^{out}
- Since X implies $\operatorname{trace} X^T X = n$, a constant, we get

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

- Solve the linear program (transportation problem)

$$\hat{X} \in \operatorname{argmax} \left\{ \langle X^{\text{out}}, X \rangle : Xe = e, X^T e = m, X \geq 0 \right\}$$

- Upper bound $= \frac{1}{2} \operatorname{trace} A\hat{X}B\hat{X}^T$

Choosing the vector v for X^{out} for upper bound

rank $Y = 1 \implies$ column/eigenvector 0 yields opt. X

1. column 0 of Y^{out};
2. eigenvector corresponding to largest eigenvalue of Y^{out};
3. random sampling/repeated: sum of random weighted-eigenvalue eigenvectors of Y^{out},

   $$v = \sum_{i=1}^{r} w_i \lambda_i v_i,$$

   where ordered eigenpairs of Y^{out} and ordered weights; r here is the *numerical rank* of Y^{out}.

Numerical Tests

Tests using:

Matlab R2017a on a ThinkPad X1 with an Intel CPU (2.5GHz) and 8GB RAM running Windows 10.

Three classes of problems:

(a) random structured graphs (compare with Pong et al.)

(b) partially random graphs with various sizes classified by the number of 1's, $|\mathcal{I}|$, in the vector *m* (similar to QAP)

(c) vertex separator instances

Lifting Linear Equality Constraint

Table: Data terminology

imax	maximum size of each set
k	number of sets
n	number of nodes (sum of sizes of sets)
p	density of graph
$l = e^T m_{\text{one}}$	number of 1's in m
Iters	number of iterations
CPU	time in seconds
Bounds	best lower and upper bounds and relative gap
Residuals	*final* values of:
	$\left\| Y^{t+1} - \widehat{V} R^{t+1} \widehat{V}^T \right\| \; (\cong \Delta Z)$;
	$\left\| Y^{t+1} - Y^t \right\| \; (\cong \Delta Y)$

Numerical Tests

Comparison small structured graphs with Pong et al

Data				Lower bounds		Upper bounds		Rel-gap		Time (cpu)	
n	k	$\lvert E\rvert$	u_0	FRSMR	Mosek	FRSMR	Mosek	FRSMR	Mosek	FRSMR	Mosek
20	4	136	6	6	6	6	6	0.00	0.00	0.21	3.96
25	4	222	8	8	8	8	8	0.00	0.00	0.20	10.94
25	5	170	14	14	14	14	14	0.00	0.00	0.31	34.19
31	5	265	22	22	22	22	22	0.00	0.00	1.28	149.49

Numerics cont...

$\mathcal{I} = \emptyset$, Results for random graphs, mean 3 instances

Specifications					Iter	cpu	Bounds			Residuals	
imax	k	n	p	l			low	up	rel-gap	prim.	dual
5	6	19.0	0.49	0	333.33	0.89	38.0	38.33	0.01	4.15e-03	6.18e-03
6	7	24.67	0.44	0	500.0	3.03	60.0	61.67	0.02	4.86e-03	8.74e-03
7	8	31.0	0.37	0	966.67	9.53	68.33	71.0	0.04	8.44e-04	3.74e-04
8	9	40.0	0.31	0	833.33	22.75	100.33	110.67	0.09	1.43e-03	6.92e-04
9	10	50.33	0.23	0	1100.0	75.26	119.67	132.33	0.09	1.53e-03	6.81e-04

$k \notin \mathcal{I} \neq \emptyset$, Results for random graphs, mean 4 instances

Specifications					Iters	cpu	Bounds			Residuals	
imax	k	n	p	l			lower	upper	rel-gap	primal	dual
5	6	16.25	0.51	1.50	450.00	1.02	22.25	23.00	0.03	2.36e-03	1.64e-03
6	7	17.00	0.43	3.25	325.00	1.18	23.00	23.25	0.00	3.75e-02	5.90e-02
7	8	21.00	0.38	3.50	625.00	4.98	34.50	36.00	0.02	3.66e-03	1.95e-03
8	9	21.75	0.30	5.00	400.00	3.36	20.75	21.25	0.01	8.37e-02	9.51e-02
9	10	38.00	0.23	3.25	775.00	25.84	55.25	63.50	0.11	3.26e-03	1.37e-03

Numerics Cont...

$k \in \mathcal{I} \neq \mathcal{K}$, Results for random graphs,mean 5 instances

Specifications					Iters	cpu	Bounds			Residuals	
imax	k	n	p	l			lower	upper	rel-gap	primal	dual
5	6	13.60	0.49	2.80	160.00	0.33	22.60	22.60	0.00	2.55e-02	3.02e-02
6	7	18.00	0.42	3.40	460.00	1.99	37.80	39.00	0.02	5.66e-02	7.10e-02
7	8	22.20	0.39	3.80	560.00	3.96	57.80	60.20	0.02	1.04e-02	1.19e-02
8	9	22.60	0.30	5.20	540.00	4.92	37.20	38.00	0.01	3.48e-02	4.29e-02
9	10	31.00	0.23	4.80	700.00	16.78	61.80	68.00	0.06	1.44e-02	1.01e-02

$\mathcal{I} = \mathcal{K}$, Results for random graphs ,mean 6 instances

Specifications				Iters	Time (cpu)	Bounds			Residuals	
k	n	p	l			lower	upper	rel-gap	primal	dual
6	6.00	0.59	6.00	100.00	0.06	4.67	4.67	0.00	5.12e-03	5.10e-03
7	7.00	0.48	7.00	100.00	0.08	5.67	5.67	0.00	8.66e-02	1.27e-01
8	8.00	0.41	8.00	150.00	0.18	7.17	7.17	0.00	2.64e-01	1.68e-01
9	9.00	0.34	9.00	233.33	0.37	7.83	8.00	0.03	1.88e-01	3.99e-02
10	10.00	0.25	10.00	266.67	0.56	7.50	7.50	0.00	6.28e-02	8.71e-02

Numerics Cont...

Table: Comparisons on the bounds for MC and bounds for the cardinality of separators

Name	n	$\lvert E \rvert$	m_1	m_2	m_3	lower	upper	lower	upper	lower	upper	lower	upper
						MC by SDP_4		MC by DNN-final		Separator by SDP_4		Separator by DNN-final	
Example 1	93	470	42	41	10	0.07	1	0	1	11	11	11	11
bcspwr03	118	179	58	57	3	0.56	1	0	2	4	5	4	5
Smallmesh	136	354	65	66	5	0.13	1	0	1	6	6	6	6
can-144	144	576	70	70	4	0.90	6	0	6	5	6	5	8
can-161	161	608	73	72	16	0.31	2	0	2	17	18	17	18
can-229	229	774	107	107	15	0.40	6	0	6	16	19	16	19
gridt(15)	120	315	56	56	8	0.29	4	0	4	9	11	9	12
gridt(17)	153	408	72	72	9	0.17	4	0	4	10	13	10	13
grid3dt(5)	125	604	54	53	18	0.54	2	0	4	19	19	19	22
grid3dt(6)	216	1115	95	95	26	0.28	4	0	4	27	30	27	31
grid3dt(7)	343	1854	159	158	26	0.60	22	0	27	27	37	27	44

Conclusion

- In this paper, we discussed strategies for finding new, strengthened lower and upper bounds, for hard discrete optimization problems.
- In particular, we exploited the fact that strict feasibility fails for many of these problems and that facial reduction, FR, leads to a natural splitting approach for ADMM, sPRSM, type methods.
- The FR makes many constraints redundant and simplifies the problem. We strengthened the subproblems in the splitting by returning redundant constraints.
- A special scaling, and a random sampling provided strengthened lower and upper bounds from low approximate solutions from our approach.

J.M. Borwein and H. Wolkowicz, *Characterization of optimality for the abstract convex program with finite-dimensional range*, J. Austral. Math. Soc. Ser. A **30** (1980/81), no. 4, 390–411. MR 83i:90156

______, *Facial reduction for a cone-convex programming problem*, J. Austral. Math. Soc. Ser. A **30** (1980/81), no. 3, 369–380. MR 83b:90121

______, *Regularizing the abstract convex program*, J. Math. Anal. Appl. **83** (1981), no. 2, 495–530. MR 83d:90236

D. Drusvyatskiy and H. Wolkowicz, *The many faces of degeneracy in conic optimization*, Foundations and Trends® in Optimization **3** (2017), no. 2, 77–170.

References II

- 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.
- S. Poljak, F. Rendl, and H. Wolkowicz, *A recipe for semidefinite relaxation for* $(0,1)$*-quadratic programming*, J. Global Optim. **7** (1995), no. 1, 51–73. MR 96d:90053
- T.K. Pong, H. Sun, N. Wang, and H. Wolkowicz, *Eigenvalue, quadratic programming, and semidefinite programming relaxations for a cut minimization problem*, Comput. Optim. Appl. **63** (2016), no. 2, 333–364. MR 3457444
- F. Rendl, A. Lisser, and M. Piacentini, *Bandwidth, vertex separators and eigenvalue optimization*, Discrete Geometry and Optimization, The Fields Institute for Research in Mathematical Sciences, Communications Series, vol. 69, Springer, 2013, pp. 249–263.

References III

- F. Rendl and R. Sotirov, *The min-cut and vertex separator problem*, Comput. Optim. Appl. **69** (2018), no. 1, 159–187.
- H. Wolkowicz, R. Saigal, and L. Vandenberghe (eds.), *Handbook of semidefinite programming*, International Series in Operations Research & Management Science, 27, Kluwer Academic Publishers, Boston, MA, 2000, Theory, algorithms, and applications. MR MR1778223 (2001k:90001)

Thanks for your attention!

Hard Combinatorial Problems, DNN Relaxations, Facial Reduction, and ADMM

Henry Wolkowicz

Dept. Combinatorics and Optimization, University of Waterloo, Canada

Friday, October 18, 2019, 8:30-9:10 AM
at: MOM21: The 21st Midwest Optimization Meeting
Northern Illinois University, DeKalb, Illinois