Relaxations of Graph Partitioning and Vertex Separator Problems using Continuous Optimization

**Henry Wolkowicz** 

(work with Ting Kei Pong, Hao Sun, Ningchuan Wang)

Dept. Combinatorics and Optimization, University of Waterloo

at: THE FIELDS INSTITUTE FOR RESEARCH IN MATHEMATICAL SCIENCES November 25- 29, 2013; Retrospective Workshop on Discrete Geometry, Optimization, and Symmetry



# Motivation: Graph Partitioning/Vertex Separator

- Model NP-hard problem using quadratic-quadratic program
- approximate/relax using eigenvalue bounds and semidefinite programming
- bounds: we follow approaches for eigenvalue and projected eigenvalue bounds in: Hadley, Rendl, W. 1990 [1, 5] Rendl, Lisser, Piacenti, (RLP) 2012 [4] and Semidefinite bounds in: W., Zhao 1996 [6].

# Background/Notation

### Given graph G and set sizes m

• G = (N, E) edge-weighted undirected graph  $N = \{1, 2, ..., n\}$  node set  $E_{ij}, ij = 1, 2, ..., n$  edge weights  $m = \begin{pmatrix} m_1 \\ ... \\ m_k \end{pmatrix}$  (pos. integer) set sizes, with  $m^T e = n$ 

Set of all Partitions

$$P_m =$$

$$\{ (S_1, \ldots, S_k) : S_i \subset N, |S_i| = m_i \forall i; \\ S_i \cap S_j = \emptyset \forall i \neq j; \cup_i S_i = N \}$$

Partition matrix  $X \in \mathbb{R}^{n \times k}$ ; col.  $X_j$  incidence vector of  $S_j$ 

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

## linear/quadratic constraints (many are redundant)

set of: zero-one; nonnegative; linear equalities; *m*-diagonal orthogonality type; *e*-diagonal orthogonality type; and gangster constraints, respectively:

$$\begin{split} \mathcal{Z} &:= & \{X \in \mathbb{R}^{n \times k} : X_{ij} \in \{0,1\}, \forall ij\} \\ &= \{X \in \mathbb{R}^{n \times k} : X_{ij}^2 = X_{ij}, \forall ij\} \\ \mathcal{N} &:= & \{X \in \mathbb{R}^{n \times k} : X_{ij} \geq 0, \forall ij\} \\ \mathcal{E} &:= & \{X \in \mathbb{R}^{n \times k} : Xe = e, X^T e = m\} \\ &= \{X \in \mathbb{R}^{n \times k} : \|Xe - e\|^2 + \|X^T e - m\|^2 = 0\} \\ \mathcal{D}_0 &:= & \{X \in \mathbb{R}^{n \times k} : X^T X = \text{Diag}(m)\} \\ \mathcal{D}_e &:= & \{X \in \mathbb{R}^{n \times k} : \text{diag}(XX^T) = e\} \\ \mathcal{G} &:= & \{X \in \mathbb{R}^{n \times k} : X_{:i} \circ X_{:j} = 0, \forall i \neq j\}, \quad \circ \text{Hadamard prod.} \end{split}$$

# **Equivalent Representations of Partition Matrices**

## The set of partition matrices in $\mathbb{R}^{nk}$ ,

$$\mathcal{M}_{m} = \mathcal{E} \cap \mathcal{Z}$$
  
=  $\mathcal{E} \cap \mathcal{D}_{0} \cap \mathcal{N}$   
=  $\mathcal{E} \cap \mathcal{D}_{0} \cap \mathcal{D}_{e} \cap \mathcal{N}$   
=  $\mathcal{E} \cap \mathcal{Z} \cap \mathcal{D}_{0} \cap \mathcal{G} \cap \mathcal{N}$ 

## Cut of a partition

- $\delta(S_i, S_j)$  set of edges between sets  $S_i, S_j$
- δ(S) = ∪<sub>i<j<k</sub>δ(S<sub>k</sub>, S<sub>j</sub>) set of edges with endpoints in distinct partition sets S<sub>1</sub>,..., S<sub>k-1</sub>
- The minimum of the cardinality |δ(S)| is denoted (objective) cut(m) = min{|δ(S)| : S ∈ P<sub>m</sub>}

#### $\mathcal{G}$ has a vertex separator

graph *G* has a vertex separator if there exists  $S \in P_m$  with  $\delta(S) = \emptyset$ , i.e., cut(m) = 0. (see (RLP) [4], Hager, Hungerford 2013 [2] for relationship with

bandwidth of graph and other applications)

# Trace Representation of Cut Problem

• 
$$B := \begin{bmatrix} ee^T - I_{k-1} & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{S}^k$$
,  
 $\mathbb{S}^k \cdot k \times k$  symm. matrices with trace inner-product.  
•  $A = (a_{ij})$  - adjacency matrix,  $a_{ij} = \begin{cases} 1 & \text{if } E_{ij} \neq 0 \\ 0 & \text{otherwise} \end{cases}$   
•  $L := \text{Diag}(Ae) - A = \sum_{ij \in E(G)} (e_i - e_j)(e_i - e_j)^T$  -  
Laplacian ( $e_i$  unit vectors)

### Quadratic objective for cut(m)

Proposition RLP [4, Prop. 2] For partition  $S \in P_m$ , and associated partition matrix  $X \in \mathcal{M}_m$ , the cardinality of the partition is  $|\delta(S)| = \frac{1}{2} \operatorname{trace} AXBX^T = \frac{1}{2} \operatorname{trace}(-L)XBX^T$ 

# **Basic Eigenvalue Bound**

## Relaxed problem

$$\begin{array}{rcl} \operatorname{cut}(m) & \geq & p_{eig}^{*}(m) \\ & & := & \min & \frac{1}{2}\operatorname{trace} AXBX^{T} & (A \text{ or } -L) \\ & & \text{ s.t. } & X \in \mathcal{D}_{O} \end{array}$$
$$\mathcal{D}_{O} = \{X \in \mathbb{R}^{n \times k} : X^{T}X = M := \operatorname{Diag}(m)\} \\ (\text{orthogonal type cols for } X) \end{array}$$

### Hoffman-Wielandt '53 [3] bound/Theorem

*C*, *D* symmetric order *n*, *k*, resp.,  $k \le n$ . Then min {trace  $CXDX^T : X^TX = I_k$ } = min { $\sum_{i=1}^k \lambda_i(D)\lambda_{\phi(i)}(C) : \phi : N \to \{1, \dots, k\}$  is an injection }. minimum attained for  $X = (p_{\phi(1)}, \dots, p_{\phi(k)}) Q^T$ , where  $p_{\phi(i)}$ normalized eigenvector to  $\lambda_{\phi(k)}(C)$  and cols of  $Q = [q_1 \dots q_k]$  contains normalized eigenvectors  $q_i$  of  $\lambda_i(D)$ .

# Basic Eigenvalue Bound II

## Lemma (RLP)

*k*-ordered eigs of  $\tilde{B} := M^{1/2} B M^{1/2}$  satisfy

$$\lambda_1( ilde{B}) \leq \lambda_2( ilde{B}) \leq \ldots \leq \lambda_{k-2}( ilde{B}) < \lambda_{k-1}( ilde{B}) = 0 < \lambda_k( ilde{B}).$$

## Basic Eigenvalue Bound, apply Hoffman-Wielandt Theorem

Let  $-\lambda_1(L) \ge -\lambda_2(L) \ge -\lambda_n(L)$  denote ordered *n* eigenvalues of -L;  $-\lambda(L)$  denotes corresponding vector of eigenvalues. Pad the 0 eigenvalue of  $\tilde{B}$  with further zeros to get an ordered vector of length *n* and denote it by  $\hat{\lambda}(\tilde{B})$ . Then

$$\operatorname{cut}(m) \ge 0 > p_{eig}^* = -\lambda(L)^T \hat{\lambda}(\tilde{B})$$

# Two Projected Eigenvalue Bound

### Relaxed problem

$$\begin{array}{rcl} \operatorname{cut}(m) & \geq & p_{\operatorname{projeig}}^{*}(m) \\ & \coloneqq & \min & \frac{1}{2}\operatorname{trace} AXBX^{T} & (A \text{ or } -L) \\ & & \text{ s.t. } & X \in \mathcal{D}_{\mathsf{O}} \cap \mathcal{E} \end{array}$$

 $\mathcal{D}_{O} = \{ X \in \mathbb{R}^{n \times k} : X^{T}X = M := \text{Diag}(m) \}$  (orthog type)

 $\mathcal{E} = \{ X \in \mathbb{R}^{n \times k} : Xe = e, X^Te = m \}$  (linear row/col sums)

# Special Parametrization of $X \in \mathcal{E}$

$$\tilde{m} = \sqrt{m}; \quad n \times n, k \times k \text{ orthogonal matrices } P, Q$$
$$P = \begin{bmatrix} \frac{1}{\sqrt{n}}e & V \end{bmatrix} \in \mathcal{O}_n, \quad Q = \begin{bmatrix} \frac{1}{\sqrt{n}}\tilde{m} & W \end{bmatrix} \in \mathcal{O}_k. \quad (*)$$

## LEMMA: Rendl and W. 1990 [5]

Let  $\tilde{M} = \text{Diag}(\tilde{m})$ . Suppose that  $X \in \mathbb{R}^{n \times k}$  and  $Z \in \mathbb{R}^{(n-1) \times (k-1)}$  are related by

$$X = P \begin{bmatrix} 1 & 0 \\ 0 & Z \end{bmatrix} Q^T \tilde{M}. \qquad (*)$$

Then the following holds:

 $X \in \mathcal{E} .$ 2  $X \in \mathcal{N} \Leftrightarrow VZW^T \geq -\frac{1}{n}e\tilde{m}^T$ 3  $X \in \mathcal{D}_0 \Leftrightarrow Z \in \mathcal{O}_{(n-1) \times (k-1)}$ 

Conversely, if  $X \in \mathcal{E}$ , then there exists Z such that the representation (\*) holds.

## THEOREM

V, W as above,  $\hat{X} := \frac{1}{n} em^T \in \mathbb{R}^{n \times k}$   $\mathcal{Q} : \mathbb{R}^{(n-1) \times (k-1)} \to \mathbb{R}^{n \times k}, \ \mathcal{Q}(Z) = VZW^T \tilde{M}$ Then:  $\hat{X} \in \mathcal{E}$ , and  $\mathcal{Q}$  is invertible  $\mathbb{R}^{(n-1) \times (k-1)} \leftrightarrow \mathcal{E} - \hat{X}$ Equivalently,  $\mathcal{E}$  can be parametrized using  $\hat{X} + VZW^T \tilde{M}$ .

## Thus, two objective functions

 $\begin{array}{l} \frac{1}{2} \operatorname{trace} AXBX^{T} = \\ \frac{1}{2} \operatorname{trace} (A\hat{X}B\hat{X}^{T} + (V^{T}AV)Z(W^{T}\tilde{M}B\tilde{M}W)Z^{T} + 2V^{T}A\hat{X}B\tilde{M}WZ^{T}) \\ \text{and} \\ \frac{1}{2} \operatorname{trace} ((-L)XBX^{T}) = \frac{1}{2} \operatorname{trace} (V^{T}(-L)V)Z(W^{T}\tilde{M}B\tilde{M}W)Z^{T}. \end{array}$ 

# Two Projected Eigenvalue Bounds

Let  $\hat{A} = V^T A V, \hat{L} = V^T (-L) V, \hat{B} = W^T \tilde{M} B \tilde{M} W,$   $\alpha = \text{trace } A \hat{X} B \hat{X}^T, C = 2 V^T A \hat{X} B \tilde{M} W.$ Then:

$$\operatorname{cut}(m) \ge p_{\operatorname{projeig},A}^* = \frac{1}{2} \left\{ \alpha + \min_{\phi \text{ injective}} \left\{ \sum_{i=1}^k \lambda_i(\hat{B}) \lambda_{\phi(i)}(\tilde{A}) \right\} + \min_{\substack{0 \le \hat{X} + VZW^T \tilde{M}}} \operatorname{trace} CZ^T \right\}$$
  
$$\ge p_{eig}^*$$

$$\begin{array}{ll} \mathsf{cut}(m) \geq p_{\textit{projeig},L}^{*} &=& \frac{1}{2} \min_{\phi \text{ injective}} \left\{ \Sigma_{i=1}^{k} \lambda_{i}(\hat{B}) \lambda_{\phi(i)}(\tilde{L}) \right\} \\ &\geq & p_{\textit{eig}}^{*}, \end{array}$$

and note eigenvalues of  $V^T L V$  are n - 1 nonzero eigenvalues of L.

let  $Q \in \mathbb{R}^{k-1 \times k-1}$  be orthog. with cols consisting of eigenvectors of  $\hat{B}$  corresponding to eigenvalues of  $\hat{B}$  in nondecreasing order; let  $P_A, P_L \in \mathbb{R}^{n-1 \times k-1}$  have orthonormal cols consisting of k-1 eigenvectors of  $\hat{A}, \hat{L}$ , respectively, corresponding to eigenvalues in nonincreasing order where the columns correspond to the largest k - 2 followed by the smallest. Then the minimal scalar product terms in  $p_{projeig,A}^*, p_{projeig,L}^*$  are attained by resp.

 $Z_A = P_A Q^T, Z_L = P_L Q^T.$ 

Get two approx. solutions using Q:

 $X_A = \hat{X} + V Z_A W^T \tilde{M}, \quad X_L = \hat{X} + V Z_L W^T \tilde{M},$ 

# Feasible Solutions; Upper Bounds

## Using an approx. solution $\bar{X}$

Find nearest (Frobenius norm) feas. soln (use strong polytime LP)

Recall:  $X \in \mathcal{E} \cap \mathcal{Z}$  implies that  $Xe = e, X^Te = m$ , and  $X^TX = \text{Diag}(m)$ . Therefore:

$$\begin{aligned} \|\bar{X} - X\|_F^2 &= \operatorname{trace}\left(\bar{X}^T\bar{X} + X^TX - 2\bar{X}^TX\right) \\ &= n + n + 2\operatorname{trace}\left(-\bar{X}^TX\right). \end{aligned}$$

Finding nearest feasible solution; a strong polytime LP

Solve the transportation problem:

$$\begin{array}{ll} \max &= \operatorname{trace} \bar{X}^T X \\ \text{s.t.} & X e = e \\ & X^T e = m \\ & X \ge 0 \end{array}$$

# Node-Arcs for a Random Adjacency Matrix

node <i>i</i>											
1	2	3	4	5	7	8	9	10	11	12	13
2	3	4	8	9	10	11	12	13	14		
3	6	7	8	9	10	11	12	13	14		
4	7	8	9	11	13	14					
5	6	7	9	10	12	13					
6	7	9	10	12	13						
7	8	10	12	13							
8	9	10	11	12	14						
9	10	13	14								
10	11	12	14								
11	12										
12	13	14									
Table: Existing edges node <i>i</i> to node <i>i</i>											

# Random Ex.; Proj. Eigenvalue Lower Bound





total edges: 61

Bounds, Feas. Sol.,  $m = (4 \ 2 \ 1 \ 6), k = 4, n = 13$ 



*imax* = 35; *k* = 6

## n is 144 and m is [28 17 28 32 34 5]

best projection lower and upper bounds are: 5092 5495 relative gap is: 0.076131

## *n* is 94 and *m* is [3 17 14 32 19 9]

best projection lower and upper bounds are 1672 1890 relative gap is 0.1224

#### *imax* = 35; k = 8

*n* is 188 and *m* is [31 27 26 34 7 6 35 22] best projection lower and upper bounds are 7558 8285 relative gap is 0.091776

## An equivalent quadratically constrained quadratic problem

$$\operatorname{cut}(m) \ge p_{SDP}^* = \min \quad \frac{1}{2} \operatorname{trace} AXBX^T \qquad (A \text{ or } (-L))$$
  
s.t.  $X \circ X = X$   
 $\|Xe - e\|^2 = 0$   
 $\|X^Te - m\|^2 = 0$   
 $X_{;i} \circ X_{;i} = 0 \ \forall i \neq j.$ 

where o is the Hadamard (elementwise) product

## **Quadratic Model**

We can use the various equality (quadratic) constraints in the representation and use the quadratic objective function. The Lagrangian relaxation for this quadratic-quadratic problem is equivalent to a semidefinite program, SDP. The dual of this is the SDP relaxation. Adding redundant constraints can help.

### Alternatively: directly by lifting process

linearize quadratic terms using the matrix

$$Y_X := \begin{pmatrix} 1 \\ \operatorname{vec}(X) \end{pmatrix} (1 \operatorname{vec}(X)^T),$$

vec (X) is vector formed from the columns of X.  $Y_X \succeq 0$  and is rank one, the hard constraint that is relaxed.

# **SDP** Relaxation

### From direct lifting (can use A or -L?)

trace  $AXBX^T = \langle AXB, X \rangle = \text{vec}(X)^T(\text{vec} AXB) =$   $\text{vec}(X)^T(B \otimes A)\text{vec}(X) = \text{trace}(B \otimes A)(\text{vec}(X)\text{vec}(X)^T)$ The objective function becomes trace  $AXBX^T = \text{trace} L_A Y_X$ ,  $L_A := \begin{bmatrix} 0 & 0 \\ 0 & B \otimes A \end{bmatrix}$  $B \otimes A$  is the Kronecker product

### Relax the rank one restriction

 $\begin{array}{ll} \operatorname{cut}(m) \geq p_{SDP}^* := \min & \operatorname{trace} L_A Y \\ \mathrm{s.t.} & \operatorname{arrow}(Y) = e_0 \\ & \operatorname{trace} D_1 Y = 0 \\ & \operatorname{trace} D_2 Y = 0 \\ & \mathcal{G}_J(Y) = 0 \\ & Y_{00} = 1 \\ & Y \succeq 0, \end{array}$ 

# Linear Transformations

#### arrow operator

acting  $(kn + 1) \times (kn + 1)$  matrix Y

$$\operatorname{arrow}(\mathsf{Y}) := \operatorname{diag}(\mathsf{Y}) - (\mathsf{0}, \mathsf{Y}_{\mathsf{0},\mathsf{1}:\mathsf{kn}})^{\mathsf{T}}$$

represents the 0, 1 constraints; guarantees diagonal and 0-th row (or column) are identical;

## Gangster operator $\mathcal{G}_J : \mathcal{S}_{kn+1} \rightarrow \mathcal{S}_{kn+1}$

shoots "holes" in a matrix

$$(\mathcal{G}_J(\mathsf{Y}))_{ij} := \left\{ egin{array}{cc} \mathsf{Y}_{ij} & ext{if } (i,j) ext{ or } (j,i) \in J \\ \mathsf{0} & ext{otherwise,} \end{array} 
ight.$$

 $J := \{(i,j) : i = (p-1)n + q, j = (r-1)n + q, \\ \text{for } \begin{cases} p < r, p, r \in \{1, \dots, k\} \\ q \in \{1, \dots, n\} \end{cases} \}$ 

represents the (Hadamard) orthogonality of the cols

# Linear Transformations

### The norm constraints



### Loss of Slater's condition

all  $D_1$ ,  $D_2$ ,  $Y \succeq 0$ , both trace  $YD_1 = 0$ , trace  $YD_2 = 0$ ; therefore, range of Y subset intersection of nullspaces of  $D_1$ ,  $D_2$ . feasible set of (RGP) has no strictly feasible points; implies numerical difficulties for interior-point methods. Fix: apply facial reduction. Facial Reduction;  $Y = \hat{V}Z\hat{V}^T \in \mathbb{S}^{kn+1}, Z \in \mathbb{S}^{(n-1)(k-1)+1}$ 



Range of  $\hat{V}$  forms basis for range (any)  $\hat{Y} \in \text{relint } F$ 

$$\hat{V} := \begin{bmatrix} 1 & 0 \\ \frac{1}{n}m \otimes e_n & V_k \otimes V_n \end{bmatrix}$$

# Constraints for $X \in \mathcal{E}$ eliminated; $Z \in \mathbb{S}^{(n-1)(k-1)+1}$

min trace  $\hat{V}^T L_A \hat{V} Z$ s.t. arrow $(\hat{V} Z \hat{V}^T) = 0$  $\mathcal{G}_J (\hat{V} Z \hat{V}^T) = 0$  $(\hat{V} Z \hat{V}^T)_{00} = 1$  $Z \succeq 0$ 

Slater's CQ now holds (strict feasibility). But are we done? Are the constraints onto?

# Final SDP; Slater and Onto (range of G) Constraints

### Projected onto range of gangster; $\overline{J} = J \cup (0, 0)$

min trace 
$$(\hat{V}^T L_A \hat{V}) Z$$
  
s.t.  $\mathcal{G}_{\bar{J}}(\hat{V}Z \hat{V}^T) = \mathcal{G}_{\bar{J}}(E_{00})$   
 $Z \succeq 0$ 

## Dual program (also satisfies Slater)

max 
$$W_{00}$$
  
s.t.  $\hat{V}^T \mathcal{G}_{\bar{J}}(W) \hat{V} \preceq \hat{V}^T L_A \hat{V}$ 

#### **Doubly Nonnegative**

A stronger relaxation adds the nonnegativity elementwise:  $\hat{V}Z\hat{V}^{T} \ge 0.$ 

# SDP Bounds; $m = (4 \ 2 \ 1 \ 6), k = 4, n = 13$



### lower bnds: [ Proj L and A; SDP; Doubly Nonneg.]

 $\begin{bmatrix} -0.52065 & 0.76067 & 2.9057 & 4.8603 \end{bmatrix}$ rounded up:  $\begin{bmatrix} 0 & 1 & 3 & 5 \end{bmatrix}$ . Therefore, **5** is optimal value.

# Random Ex; *n* = 85, *k* = 6, *m* = [18 20 11 18 11 7]

### Proj. Eig. Bounds

## n is 85 and m is 18 20 11 18 11 7

best projection lower and upper bounds are 1518 1714 relative gap is 0.12129

### **SDP** Bounds

sdp lower and upper bounds are 1556 1726 current best lower/upper bounds are: 1556 1714 relative gap is 0.096636

- Model NP hard problems using quadratic-quadratic models
- First Relaxations lead to eigenvalue problems
- Lagrangian Relaxation leads to SDP problem and the dual is the SDP (strong) relaxation
- The Slater condition typically fails for SDP relaxations (facial reduction is needed for stability)

# **References I**

- S.W. Hadley, F. Rendl, and H. Wolkowicz, A new lower bound via projection for the quadratic assignment problem, Math. Oper. Res. 17 (1992), no. 3, 727–739. MR 93i:90080
- W.W. Hager and J.T. Hungerford, *A continuous quadratic programming formulation of the vertex separator problem*, Report, University of Florida, Gainesville, 2013.
- A.J. Hoffman and H.W. Wielandt, *The variation of the spectrum of a normal matrix*, Duke Mathematics **20** (1953), 37–39.
- F. Rendl, A. Lisser, and M. Piacentini, *Bandwidth, vertex* separators and eigenvalue optimization, Tech. report, Institut fuer Mathematik, Universitaet Klagenfurt, A - 9020 Klagenfurt, Austria, 2013.

- F. Rendl and H. Wolkowicz, A projection technique for partitioning the nodes of a graph, Ann. Oper. Res. 58 (1995), 155–179, Applied mathematical programming and modeling, II (APMOD 93) (Budapest, 1993). MR 96g:90096
- H. Wolkowicz and Q. Zhao, Semidefinite programming relaxations for the graph partitioning problem, Discrete Appl. Math. 96/97 (1999), 461–479, Selected for the special Editors' Choice, Edition 1999. MR 1 724 735

Thanks for your attention!

Relaxations of Graph Partitioning and Vertex Separator Problems using Continuous Optimization

# Henry Wolkowicz

(work with Ting Kei Pong, Hao Sun, Ningchuan Wang)

Dept. Combinatorics and Optimization, University of Waterloo

at: THE FIELDS INSTITUTE FOR RESEARCH IN MATHEMATICAL SCIENCES November 25- 29, 2013; Retrospective Workshop on Discrete Geometry, Optimization, and Symmetry

