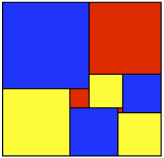
# Strong Duality and Facial Reduction in SDP: with Applications to Sensor Network Localization and Molecular Conformation

**Yuen-Lam Cheung and Henry Wolkowicz**

(Parts of this talk represent work based on Refs: [4, 5, 11, 7, 6] )

TUTTE SEMINAR

Dept. of Combinatorics and Optimization
University of Waterloo



# Motivation: Loss of Slater CQ/Facial reduction

- optimization algorithms rely on the KKT system; and require that some constraint qualification (CQ) holds (Slater's CQ for convex conic optimization)
- However, surprisingly many conic opt. instances arising from applications (QAP, GP, strengthened MC, SNL, MolecConf...) do not satisfy Slater's CQ/are degenerate
- lack of Slater's CQ results in: unbounded dual solutions; theoretical and numerical difficulties, in particularly for *primal-dual interior-point methods*.
- solution:
  - theoretical *facial reduction (Borwein-W81[4])*
  - preprocess for regularized smaller problem (Cheung-Schurr-W01[7])
  - take advantage of degeneracy (Krislock-W10[10])

# Outline: Regularization/Facial Reduction

1. Part I: Preprocessing/Regularization
   - Abstract convex program: LP, CP cases
   - Cone optimization/SDP case
2. Part II: Applications: QAP, GP, SNL, Molecular conformation ...

# Background/Abstract convex program

$$\text{(ACP)} \quad \inf_x f(x) \text{ s.t. } g(x) \preceq_K 0, x \in \Omega$$

where:

- $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ convex; $g : \mathbb{R}^n \to \mathbb{R}^m$ is $K$-convex
- $K \subset \mathbb{R}^m$ closed convex cone; $\Omega \subseteq \mathbb{R}^n$ convex set
- $a \preceq_K b \iff b - a \in K$
- $g(\alpha x + (1 - \alpha y)) \preceq_K \alpha g(x) + (1 - \alpha) g(y)$,
  $\forall x, y \in \mathbb{R}^n, \forall \alpha \in [0, 1]$ ($g$ is $K$-convex)

**Slater's CQ:** $\exists \hat{x} \in \Omega$ s.t. $g(\hat{x}) \in -\operatorname{int} K$ $\quad (g(x) \prec_K 0)$

- CQ guarantees strong duality
- essential for efficiency/stability in primal-dual interior-point methods

# Case of Linear Programming, LP

## Primal-Dual Pair: $A, m \times n$; $\mathcal{P} = \{1, \ldots, n\}$ constr. matrix/set

$$\text{(LP-P)} \quad \begin{array}{ll} \max & b^\top y \\ \text{s.t.} & A^\top y \leq c \end{array} \qquad \text{(LP-D)} \quad \begin{array}{ll} \min & c^\top x \\ \text{s.t.} & Ax = b, \; x \geq 0. \end{array}$$

## Slater's CQ for (LP-P) / Theorem of alternative

exactly one of (I),(II) holds:

$(I)$ $\exists \hat{y}$ s.t. $c - A^\top \hat{y} > 0$, $\quad \left((c - A^\top \hat{y})_i > 0, \forall i \in \mathcal{P} = \mathcal{P}^<\right)$

$(II)$ $Ad = 0, \; c^\top d = 0, \; 0 \neq d \geq 0 \qquad (*)$

## implicit equality constraints: $i \in \mathcal{P}^=$

Finding solution $0 \neq d^*$ to $(*)$ with max number of non-zeros determines:

$$d_i^* > 0 \quad \implies (c - A^\top y)_i = 0, \forall y \in \mathcal{F}^y \quad (i \in \mathcal{P}^=)$$

# Rewrite implicit-equalities to equalities/ Regularize LP

## Facial Reduction: $A^T y \leq_f c$; minimal face $f \trianglelefteq \mathbb{R}^n_+$

$$(\mathrm{LP}_{reg}\text{-P}) \qquad \begin{array}{ll} \max & b^\top y \\ \text{s.t.} & (A^<)^\top y \leq c^< \\ & (A^=)^\top y = c^= \\ \text{(equiv:} & A^T y \leq_f c \ \ \text{(min. face))} \end{array}$$

$$(\mathrm{LP}_{reg}\text{-D}) \qquad \begin{array}{ll} \min & (c^<)^\top x^< + (c^=)^\top x^= \\ \text{s.t.} & [A^< \quad A^=] \begin{pmatrix} x^< \\ x^= \end{pmatrix} = b \\ & x^< \geq 0,\ x^= \ \ \text{free} \\ \text{(equiv:} & x \in f^* \ \ \text{(dual. of face))} \end{array}$$

## Mangasarian-Fromovitz CQ (MFCQ) holds

(after deleting redundant equality constraints!)

$$\begin{pmatrix} & \underline{i \in \mathcal{P}^<} & \underline{i \in \mathcal{P}^=} \\ \exists \hat{y}: & (A^<)^\top \hat{y} < c^< & (A^=)^\top \hat{y} = c^= \end{pmatrix} \qquad (A^=)^T \text{ is onto}$$

## MFCQ holds iff dual optimal set is compact iff "stable"

Numerical difficulties if MFCQ fails; in particular for interior point methods! Modelling issue? Netlib dataset?

# Case of ordinary convex programming, CP

$$\text{(CP)} \qquad \sup_y \; b^\top y \text{ s.t. } g(y) \le 0,$$

where

- $b \in \mathbb{R}^m$; $g(y) = (g_i(y)) \in \mathbb{R}^n$, $g_i : \mathbb{R}^m \to \mathbb{R}$ convex $\forall i \in \mathbb{P}$
- Slater's CQ: $\exists \hat{y}$ s.t. $g_i(\hat{y}) < 0, \forall i$ (implies MFCQ)
- Slater's CQ fails implies implicit equality constraints exist, i.e.:
  $\mathcal{P}^= := \{i \in \mathcal{P} : g(y) \le 0 \implies g_i(y) = 0\} \ne \emptyset$
  Let $\mathcal{P}^< := \mathcal{P} \backslash \mathcal{P}^=$ and
  $g^< := (g_i)_{i \in \mathcal{P}^<}, g^= := (g_i)_{i \in \mathcal{P}^=}$

# Rewrite implicit equalities to *equalities*/ Regularize CP

**(CP) is equivalent to $g(y) \preceq_f 0$, $f$ is minimal face**

$$(\mathrm{CP_{reg}}) \qquad \begin{array}{ll} \sup & b^\top y \\ \text{s.t.} & g^<(y) \le 0 \\ & y \in \mathcal{F}^= \end{array} \qquad \text{or } (g^=(y) = 0)$$

where $\mathcal{F}^= := \{y : g^=(y) = 0\}$, and

$\mathcal{F}^= = \{y : g^=(y) \le 0\}$, so is a convex set!

Slater's CQ holds for $(CP_{reg})$ $\quad \exists \hat{y} \in \mathcal{F}^= : g^<(\hat{y}) < 0$

modelling issue again?

# Faithfully convex case

## Faithfully convex function $f$ (Rockafellar70[14])

$f$ affine on a line segment only if affine on complete line containing the segment (e.g. analytic convex functions)

$$f(y) = f_s(Ay + b) + v^T y + \alpha, \quad f_s \text{ strictly convex}$$

(cone of directions of constancy: $D_f^{=}(x) = D_f^{=} = \left(\mathcal{N}(A) \cap v^{\perp}\right)$ )

## $\mathcal{F}^{=} = \{y : g^{=}(y) = 0\}$ is an affine set

Then:

$\mathcal{F}^{=} = \{y : Vy = V\hat{y}\}$ for some $\hat{y}$ and full-row-rank matrix $V$.

And, MFCQ holds for

$$(\mathrm{CP_{reg}}) \qquad \begin{array}{llll} \sup & b^\top y & & \\ \text{s.t.} & g^{<}(y) & \leq & 0 \\ & Vy & = & V\hat{y} \end{array}$$

# Semidefinite Programming, SDP

## $K = \mathcal{S}^n_+$ nonpolyhedral cone!

$$\text{(SDP-P)} \quad v_P = \sup_{y \in \mathbb{R}^m} b^\top y \text{ s.t. } g(y) := \mathcal{A}^* y - c \preceq_{\mathcal{S}^n_+} 0$$

$$\text{(SDP-D)} \quad v_D = \inf_{x \in \mathcal{S}^n} \langle c, x \rangle \text{ s.t. } \mathcal{A} x = b,\ x \succeq_{\mathcal{S}^n_+} 0$$

where

- PSD cone $\mathcal{S}^n_+ \subset \mathcal{S}^n$ space of (real) symmetric matrices
- $c \in \mathcal{S}^n$, $b \in \mathbb{R}^m$
- $\mathcal{A} : \mathcal{S}^n \to \mathbb{R}^m$ is a linear transformation, with adjoint $\mathcal{A}^*$
- $\mathcal{A} x = (\operatorname{trace} A_i x) \in \mathbb{R}^m$ ; $\quad \mathcal{A}^* y = \sum_{i=1}^m y_i A_i \in \mathcal{S}^n$

# Slater's CQ/Theorem of Alternative

## Slater's CQ

exactly one of (I),(II) holds:

$$(I)\quad \exists \hat{y}\ \text{ s.t. } s = c - \mathcal{A}^*\hat{y} \succ 0 \quad \text{(is positive definite)}$$

$$(II)\quad \mathcal{A}d = 0,\ \operatorname{trace} cd = 0,\ 0 \neq d \succeq 0 \qquad (*)$$

# Faces of Cones - Useful for Charact. of Opt.

## Face

A convex cone $F$ is a face of $K$, denoted $F \unlhd K$, if
$x, y \in K$ and $x + y \in F \implies x, y \in F$
$(\{0\} \neq F \lhd K$ proper face)

## Conjugate Face

If $F \unlhd K$, the conjugate face (or complementary face) of $F$ is
$F^c := F^\perp \cap K^* \unlhd K^*$, $\quad (K^* = \{\phi : \langle \phi, k \rangle \geq 0, \forall k \in K\})$
If $x \in \mathrm{ri}(F)$, then $F^c = \{x\}^\perp \cap K^*$

## Minimal Faces (Intersection of Faces is a Face)

$f_P :=$ face $\mathcal{F}_P^s \unlhd K$, $\qquad \mathcal{F}_P^s$ is primal feasible set
$f_D :=$ face $\mathcal{F}_D^x \unlhd K^*$, $\qquad \mathcal{F}_D^x$ is dual feasible set

(also: face of a face is a face)

# Regularization Using Minimal Face

**Borwein-W81[4], $f_P = \text{face}\,\mathcal{F}_P^s$ (minimal face)**

(SDP-P) is equivalent to the regularized

(SDP$_{reg}$-P) $\quad v_{RP} := \sup_y \{\langle b, y\rangle \,:\, \mathcal{A}^* y \preceq_{f_P} c\}$

$\dim f_P = t(r)$ (triangle number $t(r) = r(r+1)/2$) and

$\max_{S \in f_P} \operatorname{rank} S = r < n$, if $f_P \trianglelefteq \mathcal{S}^n_+$ (implicit rank reduction)

**Lagrangian Dual DRP Satisfies Strong Duality:**

(SDP$_{reg}$-D) $\quad v_P = v_{RP} = v_{DRP} := \inf_x \{\langle c, x\rangle \,:\, \mathcal{A}x = b,\, x \succeq_{f_P^*} 0\}$

and $v_{DRP}$ is attained

smaller cone in primal $f_P \subseteq K$; larger cone in dual $K^* \subseteq f_P^*$

# (SYMMETRIC) Subspace form

**Assume Linear Feasibility for $\tilde{s}, \tilde{y}, \tilde{x}$; with data $A, b, c, K$**

$$\mathcal{A}^*\tilde{y} + \tilde{s} = c \qquad \mathcal{A}\tilde{x} = b$$

$$\mathcal{L}^\perp = \mathcal{R}(\mathcal{A}^*) \text{ (range)} \qquad \mathcal{L} = \mathcal{N}(\mathcal{A}) \text{ (nullspace)}$$

**Equivalent P-D Pair in Subspace Form, (e.g. N&N94[12])**

$\widetilde{\text{Particular solution}} + \underline{\text{solution of homogeneous equation}}$

$$\text{(SDP-P)} \quad v_P = c\tilde{x} - \inf_s \left\{ s\tilde{x} : s \in (\tilde{s} + \mathcal{L}^\perp) \cap K \right\}.$$

$$\text{(SDP-D)} \quad v_D = \tilde{y}b + \inf_x \{ \tilde{s}x : x \in (\tilde{x} + \mathcal{L}) \cap K^* \}.$$

# Minimal subspaces

## Recall

minimal faces: $f_P = \text{face}\,\mathcal{F}_P^s, \quad f_D = \text{face}\,\mathcal{F}_D^x$

## Minimal Subspaces/Linear Transformations

min. subsp.: $\mathcal{L}_{PM}^{\perp} := \mathcal{L}^{\perp} \cap (f_P - f_P), \quad \mathcal{L}_{DM} := \mathcal{L} \cap (f_D - f_D)$

min. Lin. Tr.: $\mathcal{A}_{PM}^{*}, \quad \mathcal{A}_{DM}$

# Regularization using minimal subspace

**Assume $K$ Facially Dual Complete, FDC (Pataki07[13], 'nice')**

i.e. $F \lhd K \implies K^* + F^\perp$ is closed. (e.g. $\mathcal{S}^n_+$, $\mathbb{R}^n_+$, SOC).

**$\mathcal{L}^\perp_{PM} = \mathcal{L}^\perp \cap (f_P - f_P)$**

$$v_{RP} = c\tilde{x} - \inf_s \left\{ s\tilde{x} : s \in (\tilde{s} + \mathcal{L}^\perp_{MP}) \cap K \right\} \qquad (\mathbb{RP})$$

**Lagrangian Dual DRP Satisfies Strong Duality:**

$$v_P = v_{RP} = v_{DRP} = \tilde{y}b + \inf_x \{\tilde{s}x : x \in (\tilde{x} + \mathcal{L}_{MP}) \cap K^*\} \quad (\mathbb{DRP})$$

and $v_{DRP}$ is attained

# SDP Regularization process

## (recall) alternative to Slater CQ

$$\mathcal{A}d = 0, \ \operatorname{trace} cd = 0, \ 0 \neq d \succeq_{\mathcal{S}^n_+} 0 \qquad (*)$$

## Determine a proper face $f \lhd \mathcal{S}^n_+$

let $d$ solve $(*)$ with $d = \begin{bmatrix} P & Q \end{bmatrix} \begin{bmatrix} d_+ & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P^\top \\ Q^\top \end{bmatrix} = Pd_+P^\top$

and: $0 \prec d_+ \in \mathcal{S}^{n-\bar{n}}_+$; with $[P\ Q] \in \mathbb{R}^{n \times n}$ orthogonal

Then

$$c - \mathcal{A}^* y \succeq_{\mathcal{S}^n_+} 0 \quad \Longleftrightarrow \quad \langle c - \mathcal{A}^* y, d^* \rangle = 0,$$

$$\mathcal{F}^s_P \subseteq f = \mathcal{S}^n_+ \cap \{d\}^\perp = \boxed{Q \mathcal{S}^{\bar{n}}_+ Q^T} \lhd \mathcal{S}^n_+$$

(implicit rank reduction, $\bar{n} < n$)

# Backwards Stable Regularization of SDP

**to check Theorem of Alternative:**
$\mathcal{A}d = 0,\ \langle c, d\rangle = 0,\ 0 \neq d \succeq_{\mathcal{S}^n_+} 0, (*)$

- use stable auxiliary problem

$$(AP) \qquad \min_d \ \delta \ \text{ s.t. } \quad \left\| \begin{bmatrix} \mathcal{A}d \\ \langle c, d\rangle \end{bmatrix} \right\|_2 \leq \delta, \quad \text{trace}(d) = \sqrt{n}, \quad d \succeq 0.$$

- MFCQ holds for both (AP) and its dual.
  Implies polytime accurate solution??
  But, strict complementarity can fail (resulting in possible inaccurate search directions).
- at most $n-1$ iterations to satisfy Slater's CQ for the SDP.

# Connections; complementarity partitions,

Connections: (i) finite duality gap, (ii) Slater's CQ, (iii) strict compl. in Tuncel-W07[16].

DEFINITION: Pair faces $F_1 \trianglelefteq K, F_2 \trianglelefteq K^*$ form *complementarity partition of* $K, K^*$ if $F_1 \subseteq F_2^c$ (equiv. $F_2 \subseteq F_1^c$) Partition is *strict* if $(F_1)^c = F_2$ or $(F_2)^c = F_1$.

## New Connection

THEOREM: Let $\delta^* = 0, d^* \neq 0$ solve (AP). with $d^*$ being a maximal (not full) rank; Then corresponding reduced problem satisfies Slater's CQ if, and only if, strict complementarity holds for (AP); if, and only if, the two faces

$f_P^0 := \text{face}(\{d \in \mathcal{S}^n : \mathcal{A}d = 0, \langle c, d\rangle = 0, d \succeq 0\})$
$f_D^0 := \text{face}(\{w \in \mathcal{S}^n : w = \mathcal{A}_c^* z \succeq 0, \text{ for some } z\})$

form a strict complementarity partition of $\mathcal{S}_+^n$.

# Regularizing SDP

**Minimal face containing $\mathcal{F}_P^S := \left\{ s : s = c - \mathcal{A}^* y \succeq_{\mathcal{S}_+^n} 0 \right\}$**

$$f_P = Q \mathcal{S}_+^{\bar{n}} Q^\top$$

for some $n \times n$ orthogonal matrix $U \in [P \; Q]$

**(SPD-P) is equivalent to**

$$\begin{array}{rl} \sup\limits_y & b^\top y \\ \text{s.t.} & g^{\prec}(y) := \; Q^\top(\mathcal{A}^* y - c)Q \; \preceq 0 \\ & g^{=}(y) := \begin{bmatrix} P^\top z P \\ Q^\top z P (+P^\top z Q) \end{bmatrix} = 0. \end{array}$$

MFCQ holds after removing redundant equality constraints:
$\exists \hat{y}$ s.t. $g^{\prec}(y) \prec 0$ and $g^{=}(y) = 0$.

# Conclusion Part I

- Minimal representations of the data regularize (P); use min. face $f_P$ and/or the min. L.T. $\mathcal{A}_{PM}$ or $\mathcal{L}^*_{PM}$ (implicit rank reduction)
- goal: preprocessing and a backwards stable algorithm to approx. solve (feasible) conic problems for which Slater's CQ (almost) fails (efficient on one-step problems, e.g. LPs)

| | | |
|---|---|---|
| $\uparrow$ $\rightarrow$ | Numer. Difficulties | $\leftarrow$ $\uparrow$ |
| Slater CQ | | str. compl. |
| $\updownarrow$ | | |
| pos. duality gap | | |
| $\uparrow$ $\leftarrow$ | str. compl. partit. homog. prob. | $\rightarrow$ $\uparrow$ |

# Part II: Applications of SDP where Slater's CQ fails

## SDP relaxations of NP-hard comb. opt. probs; row/column sum, and $0, 1$ constraints

- Quadratic Assignment (Zhao-Karish-Rendl-W96[18])
- Graph partitioning (W., Zhao99[17])
- strengthened Max-cut (Anjos-W01[3])
- General $0 - 1$ row/col sum constraints (Tuncel01[15])

## Low rank problems

- Sensor network localization (SNL) problem (Krislock-W10[10],Krislock-Rendl-W.10[9])
- Molecular conformation (Babak-Krislock-Ghodsi-W-Donaldson-Li11[2], Burkowski-Cheung-W11[6])
- Manifold learning (Alipanahi-Krislock-Ghodsi10[1])
- general SDP relaxation of low-rank matrix completion

# SNL (K-W10[10],K-R-W10[9])

## Highly (implicit) degenerate/low-rank problem

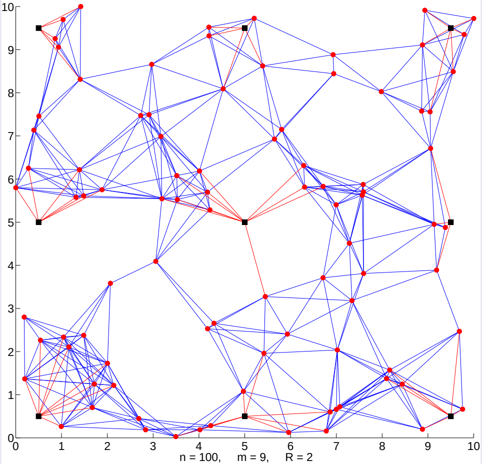
- high (implicit) degeneracy translates to low rank solutions
-fast, high accuracy solutions

## SNL - a Fundamental Problem of Distance Geometry; easy to describe - dates back to Grasssmann 1886

- $r$ : embedding dimension
- $n$ ad hoc wireless sensors $p_1, \ldots, p_n \in \mathbb{R}^r$ to locate in $\mathbb{R}^r$;
- $m$ of the sensors $p_{n-m+1}, \ldots, p_n$ are anchors (positions known, using e.g. GPS)
- pairwise distances $D_{ij} = \|p_i - p_j\|^2, ij \in E$, are known within radio range $R > 0$ (partial Euclidean distance matrix, EDM)
- $$P^T = \begin{bmatrix} p_1 & \ldots & p_n \end{bmatrix} = \begin{bmatrix} X^T & A^T \end{bmatrix} \in \mathbb{R}^{r \times n}$$

# Sensor Localization Problem/Partial EDM

## Sensors (atoms) ○ and Anchors ■; (no anchors -> molecular conformation)



# Underlying Graph Realization/Partial EDM NP-Hard

## Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \omega)$

- node set $\mathcal{V} = \{1, \ldots, n\}$
- edge set $(i,j) \in \mathcal{E}$ ; $\omega_{ij} = \|p_i - p_j\|^2$ known approximately
- The anchors form a clique (complete subgraph)
- Realization of $\mathcal{G}$ in $\mathbb{R}^r$: a mapping of nodes $v_i \mapsto p_i \in \mathbb{R}^r$ with squared distances given by $\omega$.

## Corresponding Partial Euclidean Distance Matrix, EDM

$$D_{ij} = \begin{cases} d_{ij}^2 & \text{if } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise (unknown distance),} \end{cases}$$

$d_{ij}^2 = \omega_{ij}$ are known squared Euclidean distances between sensors $p_i, p_j$; anchors correspond to a clique.

# Connections to Semidefinite Programming (SDP)

$D = \mathcal{K}(B) \in \mathcal{E}^n$, $B = \mathcal{K}^\dagger(D) \in \mathcal{S}^n \cap \mathcal{S}_C$ (centered $Be = 0$)

$P^T = \begin{bmatrix} p_1 & p_2 & \dots & p_n \end{bmatrix} \in \mathcal{M}^{r \times n}$;
$B := PP^T \in \mathcal{S}^n_+$ (Gram matrix of inner products);
$\operatorname{rank} B = r$; let $D \in \mathcal{E}^n$ corresponding $\mathrm{EDM}$; $e = \begin{pmatrix} 1 & \dots & 1 \end{pmatrix}^T$

$$
\begin{array}{rrcl}
(\text{to } D \in \mathcal{E}^n) & D & = & \left(\|p_i - p_j\|_2^2\right)_{i,j=1}^n \\
 & & = & \left(p_i^T p_i + p_j^T p_j - 2p_i^T p_j\right)_{i,j=1}^n \\
 & & = & \boxed{\operatorname{diag}(B)\, e^T + e \operatorname{diag}(B)^T - 2B} \\
 & & =: & \mathcal{D}_e(B) - 2B \\
 & & =: & \mathcal{K}(B) \quad (\text{from } B \in \mathcal{S}^n_+).
\end{array}
$$

# Euclidean Distance Matrices and SDP

## Moore-Penrose Generalized Inverse $\mathcal{K}^\dagger$

$$B \succeq 0 \quad \Longrightarrow \quad D = \mathcal{K}(B) = \operatorname{diag}(B)\,e^T + e\,\operatorname{diag}(B)^T - 2B \in \mathcal{E}$$
$$D \in \mathcal{E} \quad \Longrightarrow \quad B = \mathcal{K}^\dagger(D) = -\frac{1}{2} J \operatorname{offDiag}(D) J \succeq 0, Be = 0$$

## Theorem (Schoenberg, 1935)

A (hollow) matrix $D$ with $\operatorname{diag}(D) = 0 (D \in S_H)$ is a
Euclidean distance matrix
if, and only if,

$$B = \mathcal{K}^\dagger(D) \succeq 0. \qquad \text{(Gram matrix } B = PP^T, P\ n \times r\text{)}$$

Furthermore,

$$r = \operatorname{embdim}(D) = \operatorname{rank}\left(\mathcal{K}^\dagger(D)\right), \quad \forall D \in \mathcal{E}^n$$

# Popular Techniques; SDP Relax.; Highly Degen.

## Nearest, Weighted, SDP Approx. (relax/discard rank $B$)

- $\min_{B \succeq 0} \|H \circ (\mathcal{K}(B) - D)\|$; $\operatorname{rank} B = r$;
  typical weights: $H_{ij} = 1/\sqrt{D_{ij}}$, if $ij \in E$, $H_{ij} = 0$ otherwise.
- with rank constraint: a non-convex, NP-hard program
- SDP relaxation is convex, BUT: expensive/low accuracy/implicitly highly degenerate (cliques restrict ranks of feasible $B$s)

## Instead: (Shall) Take Advantage of Degeneracy!

clique $\alpha, |\alpha| = k$ (corresp. $D[\alpha]$) with embed. dim. $= t \leq r < k$
$\implies \operatorname{rank} \mathcal{K}^\dagger(D[\alpha]) = t \leq r \implies \operatorname{rank} B[\alpha] \leq \operatorname{rank} \mathcal{K}^\dagger(D[\alpha]) + 1$
$\implies \operatorname{rank} B = \operatorname{rank} \mathcal{K}^\dagger(D) \leq n - \boxed{(k - t - 1)} \implies$
Slater's CQ (strict feasibility) fails

# Basic Single Clique/Facial Reduction

**Matrix with fixed principal submatrix**

For $Y \in \mathcal{S}^n$, $\alpha \subseteq \{1,\dots,n\}$: $Y[\alpha]$ denotes principal submatrix formed from rows & cols with indices $\alpha$.

**$D[\alpha] \leftarrow \bar{D} \in \mathcal{E}^k$, $\alpha \subseteq 1\!:\!n$, $|\alpha| = k$**

Define $\mathcal{E}^n(\alpha, \bar{D}) := \left\{ D \in \mathcal{E}^n : D[\alpha] = \bar{D} \right\}$.

Given $\bar{D}$; find a corresponding $B \succeq 0$; find the corresponding face; find the corresponding subspace.

**if $\alpha = 1:k$; embedding dim $\operatorname{embdim}(\bar{D}) = t \le r$**

$$D = \begin{bmatrix} D[\alpha] = \bar{D} & \cdot \\ \cdot & \cdot \end{bmatrix}, \mathcal{K}^\dagger(D) = \begin{bmatrix} D[\alpha] = \mathcal{K}^\dagger(\bar{D}) + ? & ? \\ ? & ? \end{bmatrix}$$

# BASIC THEOREM for Single Clique/Facial Reduction

## THEOREM 1: Single Clique/Facial Reduction

Let: $\bar{D} := D[1:k] \in \mathcal{E}^k$, $k < n$, $\text{embdim}\,(\bar{D}) = t \le r$;
$B := \mathcal{K}^\dagger(\bar{D}) = \bar{U}_B S \bar{U}_B^T$, $\bar{U}_B \in \mathcal{M}^{k\times t}$, $\bar{U}_B^T \bar{U}_B = I_t$, $S \in \mathcal{S}^t_{++}$;
$U_B := \begin{bmatrix} \bar{U}_B & \frac{1}{\sqrt{k}} e \end{bmatrix} \in \mathcal{M}^{k\times(t+1)}$, $U := \begin{bmatrix} U_B & 0 \\ 0 & I_{n-k} \end{bmatrix}$, and
$\begin{bmatrix} V & \frac{U^T e}{\|U^T e\|} \end{bmatrix} \in \mathcal{M}^{n-k+t+1}$ orthogonal. Then:

$$\begin{aligned} \text{face}\,\mathcal{K}^\dagger\left(\mathcal{E}^n(1:k,\bar{D})\right) &= \left(U \mathcal{S}_+^{n-k+t+1} U^T\right) \cap \mathcal{S}_C \\ &= (UV)\mathcal{S}_+^{n-k+t}(UV)^T \end{aligned}$$

Note that the minimal face is defined by the subspace $\mathcal{L} = \mathcal{R}\,(UV)$. We add $\frac{1}{\sqrt{k}} e$ to represent $\mathcal{N}\,(\mathcal{K})$; then we use $V$ to eliminate $e$ to recover a centered face.

# Facial Reduction for Disjoint Cliques

## Corollary from Basic Theorem

let $\alpha_1, \ldots, \alpha_\ell \subseteq 1{:}n$ pairwise disjoint sets, wlog:
$\alpha_i = (k_{i-1}+1){:}k_i, k_0 = 0$, $\alpha := \bigcup_{i=1}^{\ell} \alpha_i = 1{:}|\alpha|$ let
$\bar{U}_i \in \mathbb{R}^{|\alpha_i| \times (t_i+1)}$ with full column rank satisfy $e \in \mathcal{R}(\bar{U}_i)$ and

$$U_i := \begin{array}{c} \\ k_{i-1} \\ |\alpha_i| \\ n-k_i \end{array} \begin{array}{c} \begin{array}{ccc} k_{i-1} & t_i+1 & n-k_i \end{array} \\ \begin{bmatrix} I & 0 & 0 \\ 0 & \bar{U}_i & 0 \\ 0 & 0 & I \end{bmatrix} \end{array} \in \mathbb{R}^{n \times (n-|\alpha_i|+t_i+1)}$$

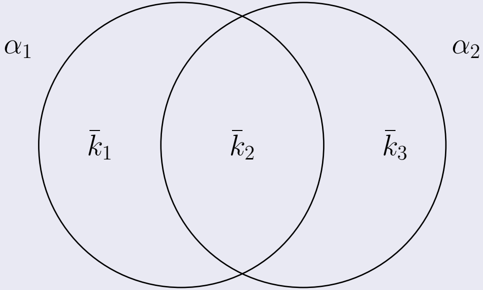
The minimal face is defined by $\mathcal{L} = \mathcal{R}(U), U = \pi_i U_i$:

$$U := \begin{array}{c} \\ |\alpha_1| \\ \vdots \\ |\alpha_\ell| \\ n-|\alpha| \end{array} \begin{array}{c} \begin{array}{cccc} t_1+1 & \ldots & t_\ell+1 & n-|\alpha| \end{array} \\ \begin{bmatrix} \bar{U}_1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & \bar{U}_\ell & 0 \\ 0 & \ldots & 0 & I \end{bmatrix} \end{array} \in \mathbb{R}^{n \times (n-|\alpha|+t+1)},$$

where $t := \sum_{i=1}^{\ell} t_i + \ell - 1$. And $e \in \mathcal{R}(U)$.

# Sets for Intersecting Cliques/Faces

$\alpha_1 := 1 : (\bar{k}_1 + \bar{k}_2); \quad \alpha_2 := (\bar{k}_1 + 1) : (\bar{k}_1 + \bar{k}_2 + \bar{k}_3)$




For each clique $|\alpha| = k$, we get a corresponding face/subspace ($k \times r$ matrix) representation. We now see how to handle two cliques, $\alpha_1, \alpha_2$, that intersect.

# Two (Intersecting) Clique Reduction/Subsp. Repres.

**THEOREM 2: Clique/Facial Intersection Using Subspace Intersection**

$\{\ \alpha_1, \alpha_2 \subseteq 1\!:\!n; \quad k := |\alpha_1 \cup \alpha_2|$

For $i = 1, 2$: $\bar{D}_i := D[\alpha_i] \in \mathcal{E}^{k_i}$, embedding dimension $t_i$;

$B_i := \mathcal{K}^\dagger(\bar{D}_i) = \bar{U}_i S_i \bar{U}_i^T$, $\bar{U}_i \in \mathcal{M}^{k_i \times t_i}$, $\bar{U}_i^T \bar{U}_i = I_{t_i}$, $S_i \in \mathcal{S}^{t_i}_{++}$;

$U_i := \begin{bmatrix} \bar{U}_i & \frac{1}{\sqrt{k_i}} e \end{bmatrix} \in \mathcal{M}^{k_i \times (t_i+1)}$; and $\bar{U} \in \mathcal{M}^{k \times (t+1)}$ satisfies

$$\mathcal{R}(\bar{U}) = \mathcal{R}\left(\begin{bmatrix} U_1 & 0 \\ 0 & I_{\bar{k}_3} \end{bmatrix}\right) \cap \mathcal{R}\left(\begin{bmatrix} I_{\bar{k}_1} & 0 \\ 0 & U_2 \end{bmatrix}\right), \text{with } \bar{U}^T \bar{U} = I_{t+1}$$

cont. . .

# Two (Intersecting) Clique Reduction, cont...

## THEOREM 2 Nonsing. Clique/Facial Inters. cont...

cont... with

$$\mathcal{R}(\bar{U}) = \mathcal{R}\left(\begin{bmatrix} U_1 & 0 \\ 0 & I_{\bar{k}_3} \end{bmatrix}\right) \cap \mathcal{R}\left(\begin{bmatrix} I_{\bar{k}_1} & 0 \\ 0 & U_2 \end{bmatrix}\right), \text{ with } \bar{U}^T\bar{U} = I_{t+1};$$

let: $U := \begin{bmatrix} \bar{U} & 0 \\ 0 & I_{n-k} \end{bmatrix} \in \mathcal{M}^{n \times (n-k+t+1)}$ and

$\begin{bmatrix} V & \frac{U^Te}{\|U^Te\|} \end{bmatrix} \in \mathcal{M}^{n-k+t+1}$ be orthogonal. Then

$$\begin{aligned} \bigcap_{i=1}^2 \text{face}\,\mathcal{K}^\dagger\left(\mathcal{E}^n(\alpha_i, \bar{D}_i)\right) &= \left(U\mathcal{S}_+^{n-k+t+1}U^T\right) \cap \mathcal{S}_C \\ &= (UV)\mathcal{S}_+^{n-k+t}(UV)^T \end{aligned}$$

# Expense/Work of (Two) Clique/Facial Reductions

## Subspace Intersection for Two Intersecting Cliques/Faces

Suppose:

$$U_1 = \begin{bmatrix} U_1' & 0 \\ U_1'' & 0 \\ 0 & I \end{bmatrix} \quad \text{and} \quad U_2 = \begin{bmatrix} I & 0 \\ 0 & U_2'' \\ 0 & U_2' \end{bmatrix}$$

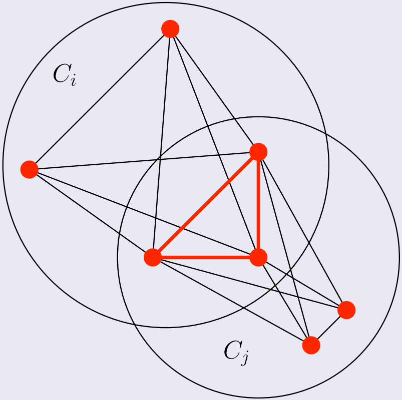
Then:

$$U := \begin{bmatrix} U_1' \\ U_1'' \\ U_2'(U_2'')^\dagger U_1'' \end{bmatrix} \quad \text{or} \quad U := \begin{bmatrix} U_1'(U_1'')^\dagger U_2'' \\ U_2'' \\ U_2' \end{bmatrix}$$

($Q_1 =: (U_1'')^\dagger U_2'', Q_2 = (U_2'')^\dagger U_1''$ orthogonal/rotation)
(Efficiently) satisfies

$$\mathcal{R}(U) = \mathcal{R}(U_1) \cap \mathcal{R}(U_2)$$

# Two (Intersecting) Clique Reduction Figure




Completion: missing distances can be recovered if desired.

# Two (Intersecting) Clique Explicit Delayed Completion

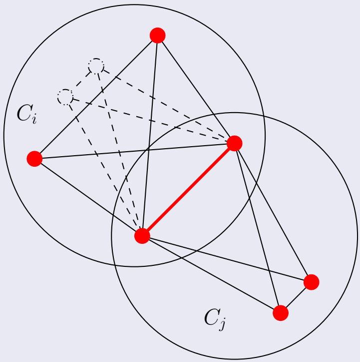
## COR. Intersection with Embedding Dim. $r$/Completion

Hypotheses of Theorem 2 holds. Let $\bar{D}_i := D[\alpha_i] \in \mathcal{E}^{k_i}$, for $i = 1, 2$, $\beta \subseteq \alpha_1 \cap \alpha_2, \gamma := \alpha_1 \cup \alpha_2, \bar{D} := D[\beta], B := \mathcal{K}^\dagger(\bar{D}), \quad \bar{U}_\beta := \bar{U}(\beta, :)$, where $\bar{U} \in \mathcal{M}^{k\times(t+1)}$ satisfies intersection equation of Theorem 2. Let $\begin{bmatrix} \bar{V} & \frac{\bar{U}^T e}{\|\bar{U}^T e\|} \end{bmatrix} \in \mathcal{M}^{t+1}$ be orthogonal. Let $\boxed{Z := (J\bar{U}_\beta \bar{V})^\dagger B((J\bar{U}_\beta \bar{V})^\dagger)^T}$. If the embedding dimension for $\bar{D}$ is $r$, <u>THEN</u> $t = r$ in Theorem 2, and $Z \in \mathcal{S}^r_+$ is the unique solution of the equation $(J\bar{U}_\beta \bar{V})Z(J\bar{U}_\beta \bar{V})^T = B$, and the exact completion is

$$\boxed{D[\gamma] = \mathcal{K}\left(PP^T\right)} \quad \text{where} \quad \boxed{P := UVZ^{\frac{1}{2}} \in \mathbb{R}^{|\gamma|\times r}}$$

# 2 (Inters.) Clique Red. Figure/Singular Case

## Two (Intersecting) Clique Reduction Figure/Singular Case




Use $R$ as lower bound in singular/nonrigid case.

# Two (Inters.) Clique Explicit Compl.; Sing. Case

## COR. Clique-Sing.; Intersect. Embedding Dim. $r-1$

Hypotheses of previous COR holds. For $i=1,2$, let $\beta \subset \delta_i \subseteq \alpha_i$, $A_i := J\bar{U}_{\delta_i}\bar{V}$, where $\bar{U}_{\delta_i} := \bar{U}(\delta_i,:)$, and $B_i := \mathcal{K}^\dagger(D[\delta_i])$. Let $\bar{Z} \in \mathcal{S}^t$ be a particular solution of the linear systems

$$\begin{array}{rcl} A_1 Z A_1^T & = & B_1 \\ A_2 Z A_2^T & = & B_2. \end{array}.$$

If the embedding dimension of $D[\delta_i]$ is $r$, for $i=1,2$, but the embedding dimension of $\bar{D} := D[\beta]$ is $r-1$, then the following holds. cont...

# 2 (Inters.) Clique Expl. Compl.; Degen. cont...

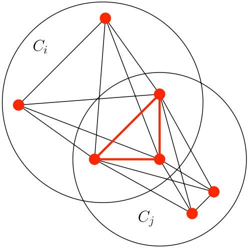
## COR. Clique-Degen. cont...

The following holds:

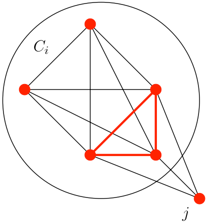
1. $\dim \mathcal{N}(A_i) = 1$, for $i = 1, 2$.
2. For $i = 1, 2$, let $n_i \in \mathcal{N}(A_i)$, $\|n_i\|_2 = 1$, and $\Delta Z := n_1 n_2^T + n_2 n_1^T$. Then, $Z$ is a solution of the linear systems if and only if
$Z = \bar{Z} + \tau \Delta Z$, for some $\tau \in \mathcal{R}$
3. There are at most two nonzero solutions, $\tau_1$ and $\tau_2$, for the generalized eigenvalue problem $-\Delta Z v = \tau \bar{Z} v$, $v \neq 0$. Set $Z_i := \bar{Z} + \frac{1}{\tau_i} \Delta Z$, for $i = 1, 2$. Then the exact completion is one of $D[\gamma] \in \{\mathcal{K}(\bar{U}\bar{V} Z_i \bar{V}^T \bar{U}^T) : i = 1, 2\}$

# Rigid Clique Union / Absorption

Clique Union




Node Absorption



# Completing SNL (Delayed use of Anchor Locations)

## Rotate to Align the Anchor Positions

- Given $P = \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} \in \mathbb{R}^{n \times r}$ such that $D = \mathcal{K}(PP^T)$
- Solve the orthogonal Procrustes problem:

$$\begin{array}{ll} \min & \|A - P_2 Q\| \\ \text{s.t.} & Q^T Q = I \end{array}$$

$P_2^T A = U \Sigma V^T$ SVD decomposition; set $Q = UV^T$;
(Golub-Van Loan/96[8], Algorithm 12.4.1)

- Set $X := P_1 Q$

# Summary: Facial Reduction for Cliques

- Using the basic theorem: each clique corresponds to a Gram matrix/corresponding subspace/corresponding face of SDP cone (implicit rank reduction)
- In the case where two cliques intersect, the union of the cliques correspond to the (efficiently computable) intersection of the corresponding faces/subspaces
- Finally, the positions are determined using a Procrustes problem

# Results - Data for Random Noisless Problems

- 2.16 GHz Intel Core 2 Duo, 2 GB of RAM
- Dimension $r = 2$
- Square region: $[0,1] \times [0,1]$
- $m = 9$ anchors
- Using only Rigid Clique Union and Rigid Node Absorption
- Error measure: Root Mean Square Deviation

$$\text{RMSD} = \left( \frac{1}{n} \sum_{i=1}^{n} \| p_i - p_i^{\text{true}} \|^2 \right)^{1/2}$$

# Results - Large $n$ (SDP size $O(n^2)$)

$n$ # of Sensors Located

| $n$ # sensors \ $R$ | 0.07 | 0.06 | 0.05 | 0.04 |
|---|---|---|---|---|
| 2000 | 2000 | 2000 | 1956 | 1374 |
| 6000 | 6000 | 6000 | 6000 | 6000 |
| 10000 | 10000 | 10000 | 10000 | 10000 |

CPU Seconds

| # sensors \ $R$ | 0.07 | 0.06 | 0.05 | 0.04 |
|---|---|---|---|---|
| 2000 | 1 | 1 | 1 | 3 |
| 6000 | 5 | 5 | 4 | 4 |
| 10000 | 10 | 10 | 9 | 8 |

RMSD (over located sensors)

| $n$ # sensors \ $R$ | 0.07 | 0.06 | 0.05 | 0.04 |
|---|---|---|---|---|
| 2000 | $4e{-16}$ | $5e{-16}$ | $6e{-16}$ | $3e{-16}$ |
| 6000 | $4e{-16}$ | $4e{-16}$ | $3e{-16}$ | $3e{-16}$ |
| 10000 | $3e{-16}$ | $5e{-16}$ | $4e{-16}$ | $4e{-16}$ |

# Results - $N$ Huge SDPs Solved

## Large-Scale Problems

| # sensors | # anchors | radio range | RMSD | Time |
|---|---|---|---|---|
| 20000 | 9 | .025 | $5e{-}16$ | 25s |
| 40000 | 9 | .02 | $8e{-}16$ | 1m 23s |
| 60000 | 9 | .015 | $5e{-}16$ | 3m 13s |
| 100000 | 9 | .01 | $6e{-}16$ | 9m 8s |

## Size of SDPs Solved: $N = \binom{n}{2}$ (# vrbls)

$\mathcal{E}_n(\text{density of } \mathcal{G}) = \pi R^2$; $M = \mathcal{E}_n(|E|) = \pi R^2 N$ (# constraints)

Size of SDP Problems:

$M = [3,078,915 \quad 12,315,351 \quad 27,709,309 \quad 76,969,790]$

$N = 10^9\,[0.2000 \quad 0.8000 \quad 1.8000 \quad 5.0000]$

# Summary Part II

- instances of degeneracy/lack of Slater CQ occurs in many applications (cliques in SNL, amino acids in molecular conformation)
- SDP relaxation of SNL is highly (implicitly) degenerate: The feasible set of this SDP is restricted to a low dim. face of the SDP cone, causing the Slater constraint qualification (strict feasibility) to fail
- We take advantage of this degeneracy by finding explicit representations of intersections of faces of the SDP cone corresponding to unions of intersecting cliques
- Without using an SDP-solver (eg. SeDuMi or SDPT3), we quickly compute the exact solution to the SDP relaxation

- B. Alipanahi, N. Krislock, and A. Ghodsi, *Manifold learning by semidefinite facial reduction*, Tech. Report Submitted to Machine Learning Journal, University of Waterloo, Waterloo, Ontario, 2010.

- B. Alipanahi, N. Krislock, A. Ghodsi, H. Wolkowicz, L. Donaldson, and M. Li, *Spros: An sdp-based protein structure determination from nmr data*, Tech. report, University of Waterloo, Waterloo, Ontario, 2011, poster session at RECOMB2011.

- M.F. Anjos and H. Wolkowicz, *Geometry of semidefinite max-cut relaxations via matrix ranks*, J. Comb. Optim. **6** (2002), no. 3, 237–270, New approaches for hard discrete optimization (Waterloo, ON, 2001). MR MR1920210 (2003h:90057)

- 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
- F. Burkowski, Y-L. Cheung, and H. Wolkowicz, *Semidefinite programming and side chain positioning*, Tech. Report CORR 2011, in progress, University of Waterloo, Waterloo, Ontario, 2011.
- Y-L. Cheung, S. Schurr, and H. Wolkowicz, *Preprocessing and reduction for degenerate semidefinite programs*, Tech. Report CORR 2011-02, University of Waterloo, Waterloo, Ontario, 2011.
- G.H. Golub and C.F. Van Loan, *Matrix computations*, $3^{nd}$ ed., Johns Hopkins University Press, Baltimore, Maryland, 1996.
- N. Krislock, F. Rendl, and H. Wolkowicz, *Noisy sensor network localization using semidefinite representations and facial reduction*, Tech. Report CORR 2010-01, University of Waterloo, Waterloo, Ontario, 2010.

N. Krislock and H. Wolkowicz, *Explicit sensor network localization using semidefinite representations and facial reductions*, SIAM Journal on Optimization **20** (2010), no. 5, 2679–2708.

N. Krislock and H. Wolkowicz, *Euclidean distance matrices and applications*, Handbook of Semidefinite, Cone and Polynomial Optimization: Theory, Algorithms, Software and Applications, CORR, no. 2009-06, Springer-Verlag, Waterloo, Ontario, to appear.

Y.E. Nesterov and A.S. Nemirovski, *Interior point polynomial algorithms in convex programming*, SIAM Publications, SIAM, Philadelphia, USA, 1994.

G. Pataki, *Bad semidefinite programs: they all look the same*, Tech. report, Department of Operations Research, University of North Carolina, Chapel Hill, 2011.

- R. Tyrrell Rockafellar, *Some convex programs whose duals are linearly constrained*, Nonlinear Programming (Proc. Sympos., Univ. of Wisconsin, Madison, Wis., 1970), Academic Press, New York, 1970, pp. 293–322.
- L. Tunçel, *On the Slater condition for the SDP relaxations of nonconvex sets*, Oper. Res. Lett. **29** (2001), no. 4, 181–186. MR MR1876975 (2002k:90076)
- L. Tunçel and H. Wolkowicz, *Strong duality and minimal representations for cone optimization*, Tech. Report CORR 2008-07, University of Waterloo, Waterloo, Ontario, 2008, submitted.
- 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

Q. Zhao, S.E. Karisch, F. Rendl, and H. Wolkowicz, *Semidefinite programming relaxations for the quadratic assignment problem*, J. Comb. Optim. **2** (1998), no. 1, 71–109, Semidefinite programming and interior-point approaches for combinatorial optimization problems (Fields Institute, Toronto, ON, 1996). MR 99f:90103

# Thanks for your attention!

## Strong Duality and Facial Reduction in SDP: with Applications to Sensor Network Localization and Molecular Conformation

**Yuen-Lam Cheung and Henry Wolkowicz**

(Parts of this talk represent work based on Refs: [4, 5, 11, 7, 6] )

TUTTE SEMINAR
Dept. of Combinatorics and Optimization
University of Waterloo