Support-Size and Rank of CLP Solutions and Applications

Yinyu Ye
Department of Management Science and Engineering
Stanford University
Stanford, CA 94305, U.S.A.

http://www.stanford.edu/~yyye

Chapters 3.1-2, 6.4-5
LP Optimality Conditions and Solution Support

\[
\begin{aligned}
\left\{ \begin{array}{l}
\mathbf{c}^T \mathbf{x} - \mathbf{b}^T \mathbf{y} = 0 \\
(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in (\mathcal{R}_+^n, \mathcal{R}_+^m, \mathcal{R}_+^n) : \\
A \mathbf{x} = \mathbf{b} \\
-A^T \mathbf{y} - \mathbf{s} = -\mathbf{c}
\end{array} \right. ;
\end{aligned}
\]

or

\[
\begin{aligned}
\mathbf{x} \cdot \mathbf{s} &= 0 \\
A \mathbf{x} &= \mathbf{b} \\
-A^T \mathbf{y} - \mathbf{s} &= -\mathbf{c}.
\end{aligned}
\]

Let \( \mathbf{x}^* \) and \( \mathbf{s}^* \) be optimal solutions with zero duality gap. Then

\[
|\text{supp}(\mathbf{x}^*)| + |\text{supp}(\mathbf{s}^*)| \leq n.
\]

There are \( \mathbf{x}^* \) and \( \mathbf{s}^* \) such that the support sizes of \( \mathbf{x}^* \) and \( \mathbf{s}^* \) are maximal, respectively.

There are \( \mathbf{x}^* \) and \( \mathbf{s}^* \) such that the support size of \( \mathbf{x}^* \) and \( \mathbf{s}^* \) are minimal, respectively.

If there is \( \mathbf{s}^* \) such that \( |\text{supp}(\mathbf{s}^*)| \geq n - d \), then the support size for \( \mathbf{x}^* \) is at most \( d \).
**LP Strict Complementarity Theorem**

**Theorem 1** If (LP) and (LD) are both feasible, then there exists a pair of strictly complementary solutions \( x^* \in F_p \) and \((y^*, s^*) \in F_d\) such that

\[
x^* \cdot s^* = 0 \quad \text{and} \quad |\text{supp}(x^*)| + |\text{supp}(s^*)| = n.
\]

Moreover, the supports

\[
P^* = \{ j : x^*_j > 0 \} \quad \text{and} \quad Z^* = \{ j : s^*_j > 0 \}
\]

are invariant for all strictly complementary solution pairs.

Given (LP) or (LD), the pair of \( P^* \) and \( Z^* \) is called the strict complementarity partition. \( \{ x : A_{P^*}x_{P^*} = b, \ x_{P^*} \geq 0, \ x_{Z^*} = 0 \} \) is called the primal optimal face, and \( \{ y : c_{Z^*} - A_{Z^*}^Ty \geq 0, \ c_{P^*} - A_{P^*}^Ty = 0 \} \) is called the dual optimal face.

\[
\begin{align*}
\text{minimize} & \quad 2x_1 + x_2 + x_3 \\
\text{subject to} & \quad x_1 + x_2 + x_3 = 1, \ (x_1, x_2, x_3) \geq 0,
\end{align*}
\]

where \( P^* = \{2, 3\} \) and \( Z^* = \{1\} \).
### Uniqueness Theorem for LP

Given an optimal solution \( \mathbf{x}^* \), how to certify the uniqueness of \( \mathbf{x}^* \)?

**Theorem 2** An LP optimal solution \( \mathbf{x}^* \) is unique if and only if the size of \( \text{supp}(\mathbf{x}^*) \) is maximal among all optimal solutions and the columns of \( A_{\text{supp}(\mathbf{x}^*)} \) are linear independent.

It is easy to see both conditions are necessary, since otherwise, one can find an optimal solution with a different support size. To see sufficiency, suppose there there is another optimal solution \( \mathbf{y}^* \) such that \( \mathbf{x}^* - \mathbf{y}^* \neq \mathbf{0} \). We must have \( \text{supp}(\mathbf{y}^*) \subset \text{supp}(\mathbf{x}^*) \), since, otherwise, \((0.5\mathbf{x}^* + 0.5\mathbf{y}^*)\) remains optimal and its support size is greater than that of \( \mathbf{x}^* \) which is a contradiction. Then we see

\[
0 = A\mathbf{x}^* - A\mathbf{y}^* = A(\mathbf{x}^* - \mathbf{y}^*) = A_{\text{supp}(\mathbf{x}^*)}(\mathbf{x}^* - \mathbf{y}^*)_{\text{supp}(\mathbf{x}^*)}
\]

which implies that columns of \( A_{\text{supp}(\mathbf{x}^*)} \) are linearly dependent.

**Corollary 1** If all optimal solutions of an LP has the same support size, then the optimal solution is unique.
Solution Rank for SDP

\[ C \cdot X - b^T y = 0 \quad X S = 0 \]
\[ A X = b \quad A X = b \]
\[ -A^T y - S = -C \quad or \quad -A^T y - S = -C \]
\[ X, S \succeq 0, \quad X, S \succeq 0 \]

Let \( X^* \) and \( S^* \) be optimal solutions with zero duality gap. Then

\[ \text{rank}(X^*) + \text{rank}(S^*) \leq n. \]

Hint of the Proof: for any symmetric PSD matrix \( P \in S^n \) with rank \( r \), there is a factorization \( P = V^T V \) where \( V \in \mathbb{R}^{r \times n} \) and columns of \( V \) are nonzero-vectors and orthogonal to each other.

There are \( X^* \) and \( S^* \) such that the ranks of \( X^* \) and \( S^* \) are maximal, respectively.

There are \( X^* \) and \( S^* \) such that the ranks of \( X^* \) and \( S^* \) are minimal, respectively.

If there is \( S^* \) such that \( \text{rank}(S^*) \geq n - d \), then the maximal rank of \( X^* \) is at most \( d \).
**SDP Strict Complementarity?**

Given a pair of SDP and (SDD) where the complementarity solution exist, is there a solution pair such that

\[ \text{rank}(X^*) + \text{rank}(S^*) = n? \]

\[
C = \begin{pmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix},
A_1 = \begin{pmatrix}
0 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0
\end{pmatrix},
A_2 = \begin{pmatrix}
0 & -1 & 0 \\
-1 & 0 & 0 \\
0 & 0 & 2
\end{pmatrix}
\]

and

\[
b = \begin{pmatrix}
0 \\
0
\end{pmatrix};
K = \mathcal{S}_3^+.
\]

The maximal solution rank of either the primal or dual is one.
Uniqueness Theorem for SDP

Given an SDP optimal and complementary solution $X^*$, how to certify the uniqueness of $X^*$?

**Theorem 3** An SDP optimal and complementary solution $X^*$ is unique if and only if the rank of $X^*$ is maximal among all optimal solutions and $V^* A_i (V^*)^T$, $i = 1, \ldots, m$, are linearly independent, where $X^* = (V^*)^T V^*$, $V^* \in \mathbb{R}^{r \times n}$, and $r$ is the rank of $X^*$.

It is easy to see why the rank of $X^*$ being maximal is necessary.

Note that for any optimal dual slack matrix $S^*$, we have $S^* \cdot (V^*)^T V^* = 0$ which implies that $S^* (V^*)^T = 0$. Consider any matrix

$$X = (V^*)^T U V^*$$

where $U \in S^r_+$ and

$$b_i = A_i \cdot (V^*)^T U V^* = V^* A_i (V^*)^T \cdot U, \ i = 1, \ldots, m.$$  

One can see that $X$ remains an optimal SDP solutions for any such $U \in S^r_+$, since it makes $X$ feasible and remain complementary to any optimal dual slack matrix. If $V^* A_i (V^*)^T$, $i = 1, \ldots, m$, are not
linearly independent, then one can find

\[ V^* A_i (V^*)^T \bullet W = 0, \quad i = 1, \ldots, m, \quad 0 \neq W \in S^r. \]

Now consider

\[ X(\alpha) = (V^*)^T (I + \alpha \cdot W)V^*, \]

and then we can choose \( \alpha \neq 0 \) such that \( X(\alpha) \succeq 0 \) is another optimal solution.

To see sufficiency, suppose there there is another optimal solution \( Y^* \) such that \( X^* - Y^* \neq 0 \). We must have \( Y^* = (V^*)^T U V^* \) for some \( I \neq U \in S^r_+ \). Then we see

\[ V^* A_i (V^*)^T \bullet (I - U) = 0, \quad i = 1, \ldots, m, \]

contradicts that they are linear independent.

**Corollary 2** If all optimal solutions of an SDP has the same rank, then the optimal solution is unique.
Recall Sensor Localization Problem (SNL)

Given $a_k \in \mathbb{R}^d$, $d_{ij} \in N_x$, and $\hat{d}_{kj} \in N_a$, find $x_i \in \mathbb{R}^d$ such that

$$\|x_i - x_j\|^2 = d_{ij}^2, \quad \forall (i, j) \in N_x, \ i < j,$$

$$\|a_k - x_j\|^2 = \hat{d}_{kj}^2, \quad \forall (k, j) \in N_a,$$

$(i, j)$ ($(k, j)$) connects points $x_i$ and $x_j$ ($a_k$ and $x_j$) with an edge whose Euclidean length is $d_{ij}$ ($\hat{d}_{kj}$).

Does the system have a localization or realization of all $x_j$’s? Is the localization unique? Is there a certification for the solution to make it reliable or trustworthy? Is the system partially localizable with certification?
Let \( X = [x_1 \ x_2 \ \ldots \ x_n] \) be the \( d \times n \) matrix that needs to be determined and \( e_j \) be the vector of all zero except 1 at the \( j \)th position. Then

\[
x_i - x_j = X(e_i - e_j) \quad \text{and} \quad a_k - x_j = [I \ X](a_k; -e_j)
\]

so that

\[
\|x_i - x_j\|^2 = (e_i - e_j)^T X^T X (e_i - e_j)
\]

\[
\|a_k - x_j\|^2 = (a_k; -e_j)^T [I \ X]^T [I \ X](a_k; -e_j) = (a_k; -e_j)^T \begin{pmatrix} I & X \\ X^T & X^T X \end{pmatrix} (a_k; -e_j).
\]
Or, equivalently,

\[
(e_i - e_j)^T Y (e_i - e_j) = d_{ij}^2, \ \forall \ i, j \in N_x, \ i < j,
\]

\[
(a_k; -e_j)^T \begin{pmatrix} I & X \\ X^T & Y \end{pmatrix} (a_k; -e_j) = \hat{d}_{kj}^2, \ \forall \ k, j \in N_a,
\]

\[
Y = X^T X.
\]
Change

\[ Y = X^T X \]

to

\[ Y \succeq X^T X. \]

This matrix inequality is equivalent to

\[
\begin{pmatrix}
I & X \\
X^T & Y
\end{pmatrix} \succeq 0.
\]

This matrix has rank at least \( d \); if it’s \( d \), then \( Y = X^T X \), and the converse is also true.
Find a symmetric matrix $Z \in \mathbb{R}^{(d+n) \times (d+n)}$ such that

$$Z = \begin{pmatrix} I & X \\ X^T & Y \end{pmatrix}.$$ 

If every sensor point is connected, directly or indirectly, to an anchor point, then the solution set must be bounded.
Sensor Localization SDP Relaxation in 2D

\[(1; 0; 0)(1; 0; 0)^T \bullet Z = 1,\]
\[(0; 1; 0)(0; 1; 0)^T \bullet Z = 1,\]
\[(1; 1; 0)(1; 1; 0)^T \bullet Z = 2,\]
\[(0; e_i - e_j)(0; e_i - e_j)^T \bullet Z = d_{ij}^2, \forall i, j \in N_x, i < j,\]
\[(a_k; -e_j)(a_k; -e_j)^T \bullet Z = \hat{d}_{kj}^2, \forall k, j \in N_a,\]
\[Z \succeq 0.\]

\[
\bar{Z} = \begin{pmatrix}
I & \bar{X} \\
\bar{X}^T & \bar{X}^T \bar{X}
\end{pmatrix} = (I, \bar{X})^T (I, \bar{X})
\]

is a feasible rank-2 solution for the relaxation, where \(\bar{X} = [\bar{x}_1 \bar{x}_2 \ldots \bar{x}_n]\) and \(\bar{x}_j\) is the true location of sensor \(j\).
The Dual of the SDP Relaxation in 2D

\begin{align*}
\text{min} & \quad w_1 + w_2 + 2w_3 + \sum_{i<j \in N_x} w_{ij} d_{ij}^2 + \sum_{k,j \in N_a} \hat{w}_{kj} \hat{d}_{kj}^2 \\
\text{s.t.} & \quad w_1 (1; 0; 0)(1; 0; 0)^T + w_2 (0; 1; 0)(0; 1; 0)^T + w_3 (1; 1; 0)(1; 1; 0)^T + \\
& \quad \sum_{i<j \in N_x} w_{ij} (0; e_i - e_j)(0; e_i - e_j)^T + \sum_{k,j \in N_a} \hat{w}_{kj} (a_k; -e_j)(a_k; -e_j)^T \succeq 0
\end{align*}

\(w_{ij}\) and \(\hat{w}_{kj}\): tensional forces on edge \(ij\); dual objective is the potential energy of the network.

Since the primal is feasible, the minimal value of the dual is not less than 0. Note that all 0 is an minimal solution for the dual. Thus, there is no duality gap.
**Duality Theorem for SNL**

**Theorem 4** Let $\bar{Z}$ be a feasible solution for SDP and $\bar{U}$ be an optimal slack matrix of the dual. Then,

1. **complementarity condition** holds: $\bar{Z} \odot \bar{U} = 0$ or $\bar{Z}\bar{U} = 0$;

2. $\text{Rank}(\bar{Z}) + \text{Rank}(\bar{U}) \leq 2 + n$;

3. $\text{Rank}(\bar{Z}) \geq 2$ and $\text{Rank}(\bar{U}) \leq n$.

An immediate result from the theorem is the following:

**Corollary 3** If an optimal dual slack matrix has rank $n$, then every solution of the SDP has rank 2 so that the solution is unique, that is, the SDP relaxation solves the original problem **exactly**.
A sensor network is \(2\)-universally-localizable (UL) if there is a unique localization in \(\mathbb{R}^2\) and there is no \(x_j \in \mathbb{R}^h, j = 1, \ldots, n\), where \(h > 2\), such that

\[
\|x_i - x_j\|^2 = d_{i,j}^2, \quad \forall \ i, j \in N_x, \ i < j,
\]

\[
\|(a_k; 0) - x_j\|^2 = \hat{d}_{k,j}^2, \quad \forall \ k, j \in N_a.
\]

The latter says that the problem cannot be localized in a higher dimension space where anchor points are simply augmented to \((a_k; 0) \in \mathbb{R}^h, k = 1, \ldots, m\).

**Theorem 5**  The SDP relaxation is exact for all universally-localizable networks.
Figure 1: One sensor-Two anchors: Not Universally Localizable
Figure 2: Two sensor-Three anchors: Universally Localizable
Figure 3: Two sensor-Three anchors: Universally Localizable (but not Strongly)
Figure 4: Two sensor-Three anchors: Not Universally Localizable
Figure 5: Two sensor-Three anchors: Universally Localizable
Theorem 6  The following SNL problems are Universally-Localizable:

- If every edge length is specified, then the sensor network is $2$-universally-localizable (Schoenberg 1942).

- There is a sensor network (trilateral graph), with $O(n)$ edge lengths specified, that is $2$-universally-localizable (So 2007).

- If one sensor with its edge lengths to at least three anchors (in general positions) specified, then it is $2$-universally-localizable (So and Y 2005).
ULPs can be localized in polynomial time

**Theorem 7** (So and Y 2005) The following statements are equivalent:

1. The sensor network is 2-universally-localizable;

2. The max-rank solution of the SDP relaxation has rank 2;

3. The solution matrix has $Y = X^T X$ or $\text{Tr}(Y - X^T X) = 0$.

When an optimal dual (stress) slack matrix has rank $n$, then the problem is 2-strongly-localizable-problem (SLP). This is a sub-class of ULP.

Example: if one sensor with its edge lengths to three anchors (in general positions) are specified, then it is 2-strongly-localizable.
One-Sensor Three-Anchor Example

Given three anchors \( \mathbf{a}_k \in \mathbb{R}^2, \ k = 1, 2, 3 \), who are not co-linear, and the three (exact) Euclidean distances, \( d_k \), from a sensor to the three anchors, find the sensor position \( \mathbf{x} \in \mathbb{R}^2 \) such that

\[
\| \mathbf{a}_k - \mathbf{x} \|^2 = d_k^2, \ k = 1, 2, 3,
\]

Denote by \( \bar{\mathbf{x}} \) the true position of the sensor that is the position we like to compute.

Does the system of multivariate quadratic equations have a solution? Is the solution unique even it has?
Convex Relaxation: SOCP

Relax “=” to “≤”): find $\textbf{x}$ such that $\|\textbf{a}_k - \textbf{x}\| \leq d_k$, $k = 1, 2, 3$.

\[
\begin{align*}
\max & \quad \textbf{0}^T \textbf{x} \\
\text{s.t.} & \quad \delta_1 = d_1 \\
& \quad \textbf{x} + s_1 = \textbf{a}_1 \\
& \quad \delta_2 = d_2 \\
& \quad \textbf{x} + s_2 = \textbf{a}_2 \\
& \quad \delta_3 = d_3 \\
& \quad \textbf{x} + s_3 = \textbf{a}_3 \\
& \quad (\delta_k; s_k) \in \text{SOCP}, \ k = 1, 2, 3.
\end{align*}
\]

This problem is in the standard SOCP dual form.
Since \( a_k - x = [I \ x](a_k; -1) \) (\( I \) here is a \( 2 \times 2 \) identity matrix) so that

\[
\|a_k - x\|^2 = (a_k; -1)^T[I \ x]^T[I \ x](a_k; -1) = (a_k; -1)^T \begin{pmatrix} I & x \\ x^T & x^T x \end{pmatrix} (a_k; -1).
\]

The original three quadratic equations can be written as

\[
(a_k; -1)(a_k; -1)^T \bullet \begin{pmatrix} I & x \\ x^T & y \end{pmatrix} = d_k^2, \ \forall \ k, j \in N_a,
\]

\[
y = x^T x.
\]

Relax \( y = x^T x \) to \( y \succeq x^T x \), which is equivalent to matrix positive semi-definiteness:

\[
\begin{pmatrix} I & x \\ x^T & y \end{pmatrix} \succeq 0.
\]

Denote this matrix by \( Z \). Then the relaxed problem can be written as SDP in the standard form.
SDP Standard Form

\[
\begin{align*}
\text{max} & \quad 0 \cdot Z \\
\text{s.t.} & \quad (1; 0; 0)(1; 0; 0)^T \cdot Z = 1, \\
& \quad (0; 1; 0)(0; 1; 0)^T \cdot Z = 1, \\
& \quad (1; 1; 0)(1; 1; 0)^T \cdot Z = 2, \\
& \quad (a_k; -1)(a_k; -1)^T \cdot Z = d_k^2, \text{ for } k = 1, 2, 3, \\
Z & \succeq 0.
\end{align*}
\]

Note that \( Z \) has rank at least 2; if it’s 2, then \( y = x^T x \), and the converse is also true. In particular, unknown

\[
\bar{Z} = \begin{pmatrix} I & \bar{x} \\ \bar{x}^T & \bar{x}^T \bar{x} \end{pmatrix} = (I, \bar{x})^T (I, \bar{x})
\]

is a rank-2 solution for the relaxation.

If we can prove the optimal dual matrix has a rank-1 solution, then the max-rank of any primal matrix solution would be 2 (and it is unique).
Assign the dual variables to

\[(1; 0; 0)(1; 0; 0)^T \bullet Z = 1, (w_1)\]
\[(0; 1; 0)(0; 1; 0)^T \bullet Z = 1, (w_2)\]
\[(1; 1; 0)(1; 1; 0)^T \bullet Z = 2, (w_3)\]
\[(a_k; -1)(a_k; -1)^T \bullet Z = d_k^2, (\lambda_k) \text{ for } k = 1, 2, 3.\]

The Dual would be

\[
\min \quad w_1 + w_2 + 2w_3 + \sum_{k=1}^{3} \lambda_k d_k^2
\]
\[
\begin{array}{ccc}
    w_1 & + & w_3 \\
    & w_3 & + \\
    w_3 & & w_2 + w_3
\end{array}
\]
\[
+ \sum_{k=1}^{3} \lambda_k a_k a_k^T - \sum_{k=1}^{3} \lambda_k a_k
\]
\[
\begin{pmatrix}
    - (\sum_{k=1}^{3} \lambda_k a_k)^T \\
    \sum_{k=1}^{3} \lambda_k
\end{pmatrix} \succeq 0.
\]

Does the dual has a rank-1 slack matrix, \(S\), with zero-objective value?
If we choose \((w., \lambda.)\)’s such that
\[
\bar{S} = (-\bar{x}; 1)(-\bar{x}; 1)^T,
\]
then, \(\bar{S} \succeq 0\) and \(\bar{S} \bullet \bar{Z} = 0\) so that \(\bar{S}\) is an optimal slack matrix for the dual and its rank is 1.

We only need to consider choosing \(\lambda\)’s such that
\[
\sum_{k=1}^{3} \lambda_k a_k = \bar{x} \quad \text{or} \quad \sum_{k=1}^{3} \lambda_k (a_k - \bar{x}) = 0
\]
\[
\sum_{k=1}^{3} \lambda_k = 1.
\]

This system always has an unique solution as long as \(a_k\)’s are not co-linear.

Then we choose (unique) \(w\)’s such that
\[
\begin{pmatrix}
w_1 + w_3 & w_3 \\
w_3 & w_2 + w_3
\end{pmatrix}
= \bar{x}_1 \bar{x}_1^T - \sum_{k=1}^{3} \lambda_k a_k a_k^T
\]
\( \lambda_k \)'s are nontrivial stresses/forces the edges between \( a_k \) and solution \( x \), respectively, and all stresses are balanced or at the equilibrium state.

Even if \( a_k \) is co-linear, the system

\[
\sum_{k=1}^{3} \lambda_k (a_k - \bar{x}) = 0
\]
\[
\sum_{k=1}^{3} \lambda_k = 1
\]

may still have a solution \( \lambda \).?
Figure 6: Dual Stresses – A 3-D Toy; provided by Anstreicher
Figure 7: **Dual Stresses** – A Needle Tower; provided by Anstreicher
In most applications, we may not be lucky and need an effort to search a rank-minimal SDP solution for SDP:

\[
(SDP) \quad \min \quad C \cdot X \\
\text{subject to} \quad A_i \cdot X = b_i, \ i = 1, 2, \ldots, m, \ X \succeq 0,
\]

where \( C, A_i \in S^n \).

Or simply for the SDP feasibility problem:

Solve \( A_i \cdot X = b_i, \ i = 1, 2, \ldots, m, \ X \succeq 0, \)
Theorem 8  (Carathéodory's theorem)

- If there is a minimizer for (LP), then there is a minimizer of (LP) whose support size $r$ satisfying $r \leq m$.

- If there is a minimizer for (SDP), then there is a minimizer of (SDP) whose rank $r$ satisfying $\frac{r(r+1)}{2} \leq m$. Moreover, such a solution can be found in polynomial time.

How Sharp is the Rank Bound? The rank bound is sharp: consider $n = 4$ and the SDP problem:

$$(e_i - e_j)(e_i - e_j)^T \bullet X = 1, \forall i < j = 1, 2, 3, 4,$$

$$X \succeq 0,$$

Applications: Finding the extreme eigenvalue of a symmetric matrix and the singular value of any matrix are convex optimization!
The Null-Space Support-Reduction for LP

1. Start at any feasible solution $x^0$ and, without loss of generality, assume $x^0 > 0$, and let $k = 0$ and $A^0 = A$.

2. Find any $A^k d = 0, \ d \neq 0$, and let $x^{k+1} = x^k + \alpha d$ where $\alpha$ is chosen such as $x^{k+1} \geq 0$ and at least one of $x^{k+1}$ equals 0.

3. Eliminate the the variable(s) in $x^{k+1}$ and column(s) in $A^k$ corresponding to $x^{k+1}_j = 0$, and let the new narrower matrix be $A^{k+1}$.

4. Set $k = k + 1$ and return to step 2.

This process is called rounding, or purification, procedure in linear programming.
I. The Null-Space Rank-Reduction: A Constructive Proof

Let $X^\ast$ be an optimal solution. Then, if the rank of $X^\ast$, $r$, satisfies the inequality, we need do nothing. Thus, we assume $r(r + 1)/2 > m$, and let

$$V^TV = X^\ast, \quad V \in R^{r \times n}.$$ 

Then consider

Minimize $VCV^T \bullet U$

Subject to $VA_iV^T \bullet U = b_i, \ i = 1, \ldots, m$

$U \geq 0$. \hspace{1cm} (1)

Note that $VCV^T$, $VA_iV^T$ and $U$ are $r \times r$ symmetric matrices and, in particular,

$$VCV^T \bullet I = C \bullet V^TV = C \bullet X^\ast = z^\ast.$$
Moreover, for any feasible solution of (1) one can construct a feasible matrix solution for (2) using

\[ X(U) = V^T U V \quad \text{and} \quad C \bullet X(U) = VCV^T \bullet U. \]  

(2)

Thus, the minimal value of (1) is also \( z^* \), and \( U = I \) is a minimizer of (1).

Now we show that any feasible solution \( U \) to (1) is a minimizer for (1); thereby \( X(U) \) of (2) is a minimizer for the original SDP. Consider the dual of (1)

\[ z^* := \text{Maximize} \quad b^T y = \sum_{i=1}^{m} b_i y_i \]

Subject to \( VCV^T \succeq \sum_{i=1}^{m} y_i V A_i V^T \).

(3)

Let \( y^* \) be a dual maximizer. Since \( U = I \) is an interior optimizer for the primal, the strong duality condition holds, i.e.,

\[ I \bullet (VCV^T - \sum_{i=1}^{m} y_i^* V A_i V^T) = 0 \]
so that we have

\[ VCV^T - \sum_{i=1}^{m} y_i^* V A_i V^T = 0. \]

Then, any feasible solution of (1) satisfies the strong duality condition so that it must be also optimal.

Consider the system of homogeneous linear equations

\[ VA_i V^T \cdot W = 0, \quad i = 1, \ldots, m \]

where \( W \) is a \( r \times r \) symmetric matrices (does not need to be definite). This system has \( r(r + 1)/2 \) real number variables and \( m \) equations. Thus, as long as \( r(r + 1)/2 > m \), we must be able to find a symmetric matrix \( W \neq 0 \) to satisfy all \( m \) equations. Without loss of generality, let \( W \) be either indefinite or negative semidefinite (if it is positive semidefinite, we take \( -W \) as \( W \)), that is, \( W \) has at least one negative eigenvalue, and consider

\[ U(\alpha) = I + \alpha W. \]

Choosing \( \alpha^* = 1/|\bar{\lambda}| \) where \( \bar{\lambda} \) is the least eigenvalue of \( W \), we have

\[ U(\alpha^*) \succeq 0 \]
and it has at least one 0 eigenvalue or \( \text{rank}(U(\alpha^*)) < r \), and

\[
VA_iV^T \bullet U(\alpha^*) = VA_iV^T \bullet (I + \alpha^*W) = VA_iV^T \bullet I = b_i, \ i = 1, \ldots, m.
\]

That is, \( U(\alpha^*) \) is a feasible and so it is an optimal solution for (1). Then,

\[
X(U(\alpha^*)) = V^T U(\alpha^*) V
\]

is a new minimizer for (1), and \( \text{rank}(X(U(\alpha^*))) < r \).

This process can be repeated till the system of homogeneous linear equations has only all zero solution, which is necessarily given by \( r(r+1)/2 \leq m \). The total number of such reduction steps is bounded by \( n - 1 \) and each step uses no more than \( O(m^2n) \) arithmetic operations and finds the least eigenvalue of \( W \), which is a polynomial time.
II. The Principle-Component or Eigenvalue Reduction

Let $\tilde{X}$ be an SDP solution with rank $r$ and

$$\tilde{X} = \sum_{i=1}^{r} \lambda_i v_i v_i^T$$

where

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n.$$ 

Then, let

$$\hat{X} = \sum_{i=1}^{d} \lambda_i v_i v_i^T$$
Let $\tilde{X}$ be an SDP solution with rank $r$ and

$$\tilde{X} = VV^T$$

where $V \in \mathbb{R}^{n \times r}$ is any factorization matrix of $\tilde{X}$

Then, let random matrix

$$R = \sum_{i=1}^{d} \xi_i \xi_i^T, \quad \xi_i \in N(0, \frac{1}{d} I); \quad \text{or} \quad \xi_i \in \text{Binary}(0, \frac{1}{d} I)$$

that is, each entry either 1 or $-1$ in the latter case. Then assign

$$\hat{X} = VRV^T.$$ 

Note that $(V\xi_i)(V\xi_i)^T \in N(0, \frac{1}{d} \tilde{X})$ and

$$E[\hat{X}] = VE[R]V^T = VV^T = \tilde{X}.$$
For simplicity, consider the SDP feasibility problem

\[ A_i \bullet X = b_i \quad i = 1, \ldots, m, \quad X \succeq 0 \]

where \( A_1, \ldots, A_m \) are positive semidefinite matrices and scalars \((b_1, \ldots, b_m) \geq 0\).

We try to find an approximate \( \hat{X} \succeq 0 \) of rank at most \( d \):

\[ \beta(m, n, d) \cdot b_i \leq A_i \bullet \hat{X} \leq \alpha(m, n, d) \cdot b_i \quad \forall i = 1, \ldots, m. \]

Here, \( \alpha \geq 1 \) and \( \beta \in (0, 1] \) are called the distortion factors. Clearly, the closer are both to 1, the better.
The Main Theorem

**Theorem 9** Let \( r = \max \{ \text{rank}(A_i) \} \) and \( \bar{X} = VV^T \) be a feasible solution. Then, for any \( d \geq 1 \), the randomly generated

\[
\hat{X} = V \left[ \sum_{i=1}^{d} \xi_i \xi_i^T \right] V^T, \quad \xi_i \in N(0, \frac{1}{d} I)
\]

\[
\alpha(m, n, d) = \begin{cases} 
  1 + \frac{12 \ln(4mr)}{d} & \text{for } 1 \leq d \leq 12 \ln(4mr) \\
  1 + \sqrt{\frac{12 \ln(4mr)}{d}} & \text{for } d > 12 \ln(4mr)
\end{cases}
\]

and

\[
\beta(m, n, d) = \begin{cases} 
  \frac{1}{e(2m)^{2/d}} & \text{for } 1 \leq d \leq 4 \ln(2m) \\
  \max \left\{ \frac{1}{e(2m)^{2/d}}, 1 - \sqrt{\frac{4 \ln(2m)}{d}} \right\} & \text{for } d > 4 \ln(2m)
\end{cases}
\]
Some Remarks and Open Questions

- There is always a low-rank, or sparse, approximate SDP solution with respect to a bounded relative residual distortion. As the allowable rank increases, the distortion bounds become smaller and smaller.

- The lower distortion factor is independent of \( n \) and the rank of \( A_i \)s.

- The factors can be improved if we only consider one-sided inequalities.

- This result contains as special cases several well-known results in the literature.

- Can the distortion upper bound be improved such that it’s independent of rank of \( A_i \)?

- Is there deterministic rank-reduction procedure? Choose the largest \( d \) eigenvalue component of \( X \)?

- General symmetric \( A_i \)?

- In practical applications, we see much smaller distortion, why?
Let $X$ be an SDP solution with rank $r$ and

$$X = V V^T.$$ 

Then, let random vector

$$\mathbf{u} \in N(0, I) \quad \text{and} \quad \hat{x} = \text{Sign}(V \mathbf{u})$$

where

$$\text{Sign}(x) = \begin{cases} 
1 & \text{if } x \geq 0 \\
-1 & \text{otherwise}. 
\end{cases}$$

Note that $V \mathbf{u} \in N(0, X)$. It was proved by Sheppard (1900):

$$E[\hat{x}_i \hat{x}_j] = \frac{2}{\pi} \arcsin(\bar{X}_{ij}), \quad i, j = 1, 2, \ldots, n.$$
This is the Max-Cut problem on an undirected graph $G = (V, E)$ with non-negative weights $w_{ij}$ for each edge in $E$ (and $w_{ij} = 0$ if $(i, j) \notin E$), which is the problem of partitioning the nodes of $V$ into two sets $S$ and $V \setminus S$ so that

$$w(S) := \sum_{i \in S, j \in V \setminus S} w_{ij}$$

is maximized. A problem of this type arises from many network planning, circuit design, and scheduling applications.
Figure 8: Illustration of the Max-Cut Problem
Max-Cut Formulation with Binary Quadratic Minimization

\[ w^* := \text{Maximize} \quad w(x) := \frac{1}{4} \sum_{i,j} w_{ij} (1 - x_i x_j) \]

(MC)

Subject to \( (x_j)^2 = 1, \ j = 1, \ldots, n. \)
The Coin-Toss Method: Approximation Quality

Let each node be selected to one side, or $\hat{x}_j$ be 1, independently with probability $\frac{1}{2}$.

Or simply let random vector

$$ u \in N(0, I) \quad \text{and} \quad \hat{x} = \text{Sign}(u). $$

We have

$$ E[w(\hat{x})] = E\left[\frac{1}{4} \sum_{i,j} w_{ij} (1 - x_i x_j)\right] = \frac{1}{4} \sum_{i,j} w_{ij} (1 - E[x_i x_j]) $$

$$ = \frac{1}{4} \sum_{i,j} w_{ij} = \frac{\text{weights of all edges}}{2} \geq \frac{1}{2} w^*. $$
Let $X = xx^T \in S^n_+$. Then the problem can be rewritten as

$$z^{SDP} := \text{Maximize} \quad \frac{1}{4} \sum_{i,j} w_{ij} (1 - X_{ij})$$

Subject to  \quad X_{ii} = 1, \quad i = 1, \ldots, n, \quad X \succeq 0, \quad \text{rank}(X) = 1.$$

By removing the rank-one constraint, it leads to the SDP relaxation problem.

Let $\bar{X}$ be an optimal solution for (SDP). Then, generate a random vector $u \in N(0, \bar{X})$:

$$\hat{x} = \text{Sign}(u), \quad E[\hat{x}_i \hat{x}_j] = \arcsin(X_{i,j})$$

**Theorem 10** (Goemans and Williamson)

$$E[w(\hat{x})] \geq 0.878z^{SDP} \geq 0.878w^*.$$
V. Objective-Guided Reduction

Construct a suitable objective for the SDP solution set

Minimize \( R \cdot X \)

Subject to \( A_i \cdot X = b_i, \, i = 1, \ldots, m, \)
\( C \cdot X \leq \alpha \cdot z^*, \)
\( X \succeq 0, \)

where \( z^* \) is the minimal objective value of the SDP relaxation, and \( \alpha \) is a tolerance factor.

The selection of matrix \( R \) is problem dependent. Examples include the \( L_1 \) norm function, the tensegrity graph approach, etc.
Tensegrity (Tensional-Integrity) Objective for SNL: a Chain Graph

Anchor-free SNL: let $e_i$ be the unit vector (one for the $i$th entry and zeros for the else)

$$(e_i - e_j)(e_i - e_j)^T \cdot X = d_{ij}^2, \ \forall (i, j) \in E, i < j,$$

$$X \succeq 0.$$

For certain graphs, to select a subset edges to maximize and/or a subset of edges to minimize is guaranteed to finding the lowest rank SDP solution – Tensegrity Method.
Consider:

\[
\begin{align*}
\text{max} \quad & e_3 e_3 \cdot X \\
\text{s.t.} \quad & e_1 e_1^T \cdot X = 1, \\
& (e_1 - e_2)(e_1 - e_2)^T \cdot X = 1, \\
& (e_2 - e_3)(e_2 - e_3)^T \cdot X = 1, \\
& X \succeq 0 \in \mathcal{S}^3,
\end{align*}
\]

where its maximal solution \( X^* = (1, 2, 3)^T (1, 2, 3) \). The dual is

\[
\begin{align*}
\text{min} \quad & y_1 + y_2 + y_3 \\
\text{s.t.} \quad & y_1 e_1 e_1^T + y_2 (e_1 - e_2)(e_1 - e_2)^T + y_3 (e_2 - e_3)(e_2 - e_3)^T - S = e_3 e_3, \\
& S \succeq 0 \in \mathcal{S}^3,
\end{align*}
\]

The dual has a rank-two solution with \((y_1 = 3, y_2 = 3, y_3 = 3)\).
Figure 9: Dimension Reduction – Unfolding Scroll of Happiness
Figure 10: Molecular Conformation – 1F39 (1534 atoms) with 85% of distances below 6rA and 10% noise on upper and lower bounds