# 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

$$\left\{ (\mathbf{x},\mathbf{y},\mathbf{s}) \in (\mathcal{R}^n_+, \mathcal{R}^m, \mathcal{R}^n_+) : \begin{array}{rcl} \mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y} & = & \mathbf{0} \\ A\mathbf{x} & = & \mathbf{b} \\ -A^T\mathbf{y} - \mathbf{s} & = & -\mathbf{c} \end{array} \right\};$$

or

$$\begin{array}{rcl} \mathbf{x}.\cdot\mathbf{s} & = & \mathbf{0} \\ A\mathbf{x} & = & \mathbf{b} \\ -A^T\mathbf{y} - \mathbf{s} & = & -\mathbf{c}. \end{array}$$

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* $\mathbf{x}^* \in \mathcal{F}_p$ *and* $(\mathbf{y}^*, \mathbf{s}^*) \in \mathcal{F}_d$ *such that*

$$\mathbf{x}^*_{\cdot} \cdot \mathbf{s}^* = \mathbf{0} \quad \textit{and} \quad |\textit{supp}(\mathbf{x}^*)| + |\textit{supp}(\mathbf{s}^*)| = n.$$

*Moreover, the supports*

$$P^* = \{j : x_j^* > 0\} \quad \textit{and} \quad Z^* = \{j : s_j^* > 0\}$$

*are invariant for all strictly complementary solution pairs.*

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

$$\begin{array}{ll} \text{minimize} & 2x_1 + x_2 + x_3 \\ \text{subject to} & x_1 + x_2 + x_3 = 1,\ (x_1, x_2, x_3) \geq \mathbf{0}, \end{array}$$

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 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 supp$(\mathbf{y}^*) \subset$ 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

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

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

$$\begin{array}{rcl} C \bullet X - \mathbf{b}^T\mathbf{y} & = & 0 \\ \mathcal{A}X & = & \mathbf{b} \\ -\mathcal{A}^T\mathbf{y} - S & = & -C \\ X, S & \succeq & \mathbf{0}, \end{array} \quad \text{or} \quad \begin{array}{rcl} XS & = & \mathbf{0} \\ \mathcal{A}X & = & \mathbf{b} \\ -\mathcal{A}^T y - S & = & -C \\ X, S & \succeq & \mathbf{0} \end{array}$$

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 $V \in R^{r \times n}$ 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

$$\mathbf{b} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \; K = \mathcal{S}^3_+.$$

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 $V^*A_i(V^*)^T$, $i = 1, ..., m$, are linearly independent, where $X^* = (V^*)^T V^*, V^* \in \mathcal{R}^{r\times n}$, 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 $S^* \bullet (V^*)^T V^* = 0$ which implies that $S^*(V^*)^T = \mathbf{0}$. Consider any matrix

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

where $U \in \mathcal{S}_+^r$ and

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

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

$$V^* A_i (V^*)^T \bullet W = 0, \ \ i = 1, ..., m, \ \mathbf{0} \neq W \in \mathcal{S}^r.$$

Now consider

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

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

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

$$V^* A_i (V^*)^T \bullet (I - U) = 0, \ \ i = 1, ..., m,$$

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 $\mathbf{a}_k \in \mathbf{R}^d$, $d_{ij} \in N_x$, and $\hat{d}_{kj} \in N_a$, find $\mathbf{x}_i \in \mathbf{R}^d$ such that

$$\|\mathbf{x}_i - \mathbf{x}_j\|^2 = d_{ij}^2, \ \forall\,(i,j) \in N_x,\ i < j,$$
$$\|\mathbf{a}_k - \mathbf{x}_j\|^2 = \hat{d}_{kj}^2, \ \forall\,(k,j) \in N_a,$$

$(ij)$ ($(kj)$) connects points $\mathbf{x}_i$ and $\mathbf{x}_j$ ($\mathbf{a}_k$ and $\mathbf{x}_j$) with an edge whose Euclidean length is $d_{ij}$ ($\hat{d}_{kj}$).

Does the system have a localization or realization of all $\mathbf{x}_j$'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?

## Matrix Representation

Let $X = [\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_n]$ be the $d \times n$ matrix that needs to be determined and $\mathbf{e}_j$ be the vector of all zero except $1$ at the $j$th position. Then

$$\mathbf{x}_i - \mathbf{x}_j = X(\mathbf{e}_i - \mathbf{e}_j) \quad \text{and} \quad \mathbf{a}_k - \mathbf{x}_j = [I\ \ X](\mathbf{a}_k; -\mathbf{e}_j)$$

so that

$$\|\mathbf{x}_i - \mathbf{x}_j\|^2 = (\mathbf{e}_i - \mathbf{e}_j)^T X^T X(\mathbf{e}_i - \mathbf{e}_j)$$

$$\|\mathbf{a}_k - \mathbf{x}_j\|^2 = (\mathbf{a}_k; -\mathbf{e}_j)^T [I\ \ X]^T [I\ \ X](\mathbf{a}_k; -\mathbf{e}_j) =$$

$$(\mathbf{a}_k; -\mathbf{e}_j)^T \begin{pmatrix} I & X \\ X^T & X^T X \end{pmatrix} (\mathbf{a}_k; -\mathbf{e}_j).$$

Or, equivalently,

$$
\begin{aligned}
&(\mathbf{e}_i - \mathbf{e}_j)^T Y (\mathbf{e}_i - \mathbf{e}_j) = d_{ij}^2,\ \forall\, i, j \in N_x,\ i < j,\\
&(\mathbf{a}_k; -\mathbf{e}_j)^T \begin{pmatrix} I & X \\ X^T & Y \end{pmatrix} (\mathbf{a}_k; -\mathbf{e}_j) = \hat{d}_{kj}^2,\ \forall\, k, j \in N_a,\\
&Y = X^T X.
\end{aligned}
$$

## SDP Relaxation

Change

$$Y = X^T X$$

to

$$Y \succeq X^T X.$$

This matrix inequality is equivalent to

$$\begin{pmatrix} I & X \\ X^T & Y \end{pmatrix} \succeq \mathbf{0}.$$

This matrix has rank at least $d$; if it's $d$, then $Y = X^T X$, and the converse is also true.

## SDP Standard Form

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

Find a symmetric matrix $Z \in \mathbf{R}^{(d+n)\times(d+n)}$ such that

$$\begin{aligned}
&Z_{1:d,1:d} = I \\
&(\mathbf{0}; \mathbf{e}_i - \mathbf{e}_j)(\mathbf{0}; \mathbf{e}_i - \mathbf{e}_j)^T \bullet Z = d_{ij}^2, \ \forall\, i, j \in N_x,\ i < j, \\
&(\mathbf{a}_k; -\mathbf{e}_j)(\mathbf{a}_k; -\mathbf{e}_j)^T \bullet Z = \hat{d}_{kj}^2, \ \forall\, k, j \in N_a, \\
&Z \succeq \mathbf{0}.
\end{aligned}$$

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

$$
\begin{aligned}
&(1;0;\mathbf{0})(1;0;\mathbf{0})^T \bullet Z = 1,\\
&(0;1;\mathbf{0})(0;1;\mathbf{0})^T \bullet Z = 1,\\
&(1;1;\mathbf{0})(1;1;\mathbf{0})^T \bullet Z = 2,\\
&(\mathbf{0};\mathbf{e}_i - \mathbf{e}_j)(\mathbf{0};\mathbf{e}_i - \mathbf{e}_j)^T \bullet Z = d_{ij}^2, \ \forall\, i,j \in N_x,\ i<j,\\
&(\mathbf{a}_k;-\mathbf{e}_j)(\mathbf{a}_k;-\mathbf{e}_j)^T \bullet Z = \hat{d}_{kj}^2, \ \forall\, k,j \in N_a,\\
&Z \succeq \mathbf{0}.
\end{aligned}
$$

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

is a feasible rank-2 solution for the relaxation, where $\bar{X} = [\bar{\mathbf{x}}_1\ \bar{\mathbf{x}}_2\ \ldots\ \bar{\mathbf{x}}_n]$ and $\bar{\mathbf{x}}_j$ is the true location of sensor $j$.

## The Dual of the SDP Relaxation in 2D

$$\begin{array}{ll} \text{min} & w_1 + w_2 + 2w_3 + \sum_{i<j \in N_x} w_{ij} d_{ij}^2 + \sum_{k,j \in N_a} \hat{w}_{kj} \hat{d}_{kj}^2 \\ \text{s.t.} & w_1(1;0;\mathbf{0})(1;0;\mathbf{0})^T + w_2(0;1;\mathbf{0})(0;1;\mathbf{0})^T + w_3(1;1;\mathbf{0})(1;1;\mathbf{0})^T + \\ & \sum_{i<j \in N_x} w_{ij}(\mathbf{0};\mathbf{e}_i - \mathbf{e}_j)(\mathbf{0};\mathbf{e}_i - \mathbf{e}_j)^T + \sum_{k,j \in N_a} \hat{w}_{kj}(\mathbf{a}_k;-e_j)(\mathbf{a}_k;-\mathbf{e}_j)^T \succeq \mathbf{0} \end{array}$$

$w_{ij}$ 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 $\mathbf{0}$ 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: $\bar{Z} \bullet \bar{U} = 0$ or $\bar{Z}\bar{U} = \mathbf{0}$;*
2. *Rank$(\bar{Z})$ + Rank$(\bar{U}) \leq 2 + n$;*
3. *Rank$(\bar{Z}) \geq 2$ and 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.*

## Theoretical Analyses on Sensor Network Localization

A sensor network is $2$-universally-localizable (UL) if there is a unique localization in $\mathbf{R}^2$ and there is no $x_j \in \mathbf{R}^h,\ j = 1, ..., n$, where $h > 2$, such that

$$
\begin{aligned}
\|x_i - x_j\|^2 &= d_{ij}^2,\ \forall\, i,j \in N_x,\ i < j, \\
\|(a_k; \mathbf{0}) - x_j\|^2 &= \hat{d}_{kj}^2,\ \forall\, k,j \in N_a.
\end{aligned}
$$

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

**Theorem 5** *The SDP relaxation is exact for all universally-localizable networks.*

Figure 1: One sensor-Two anchors: Not Universally Localizable

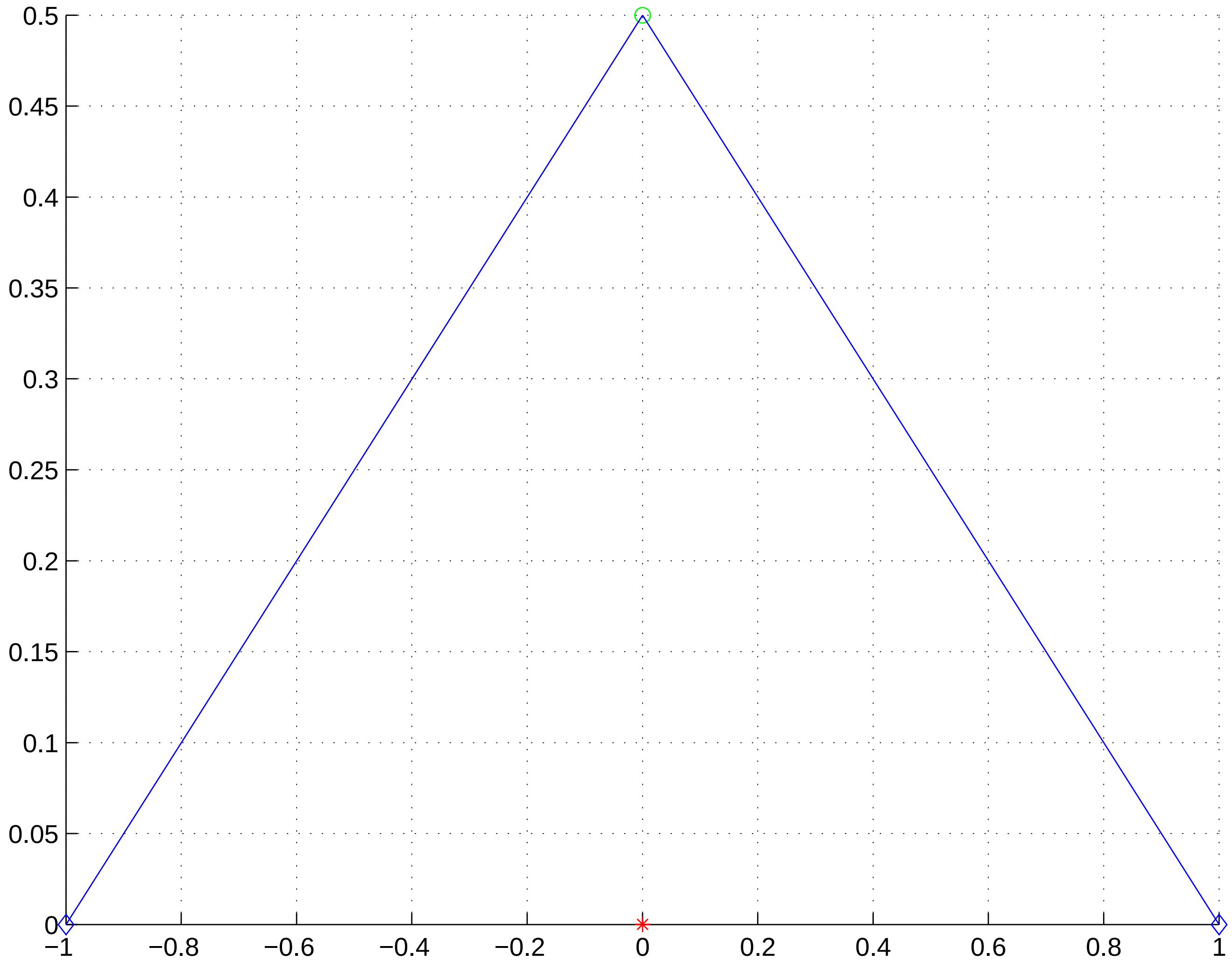


Figure 2: Two sensor-Three anchors: Universally Localizable

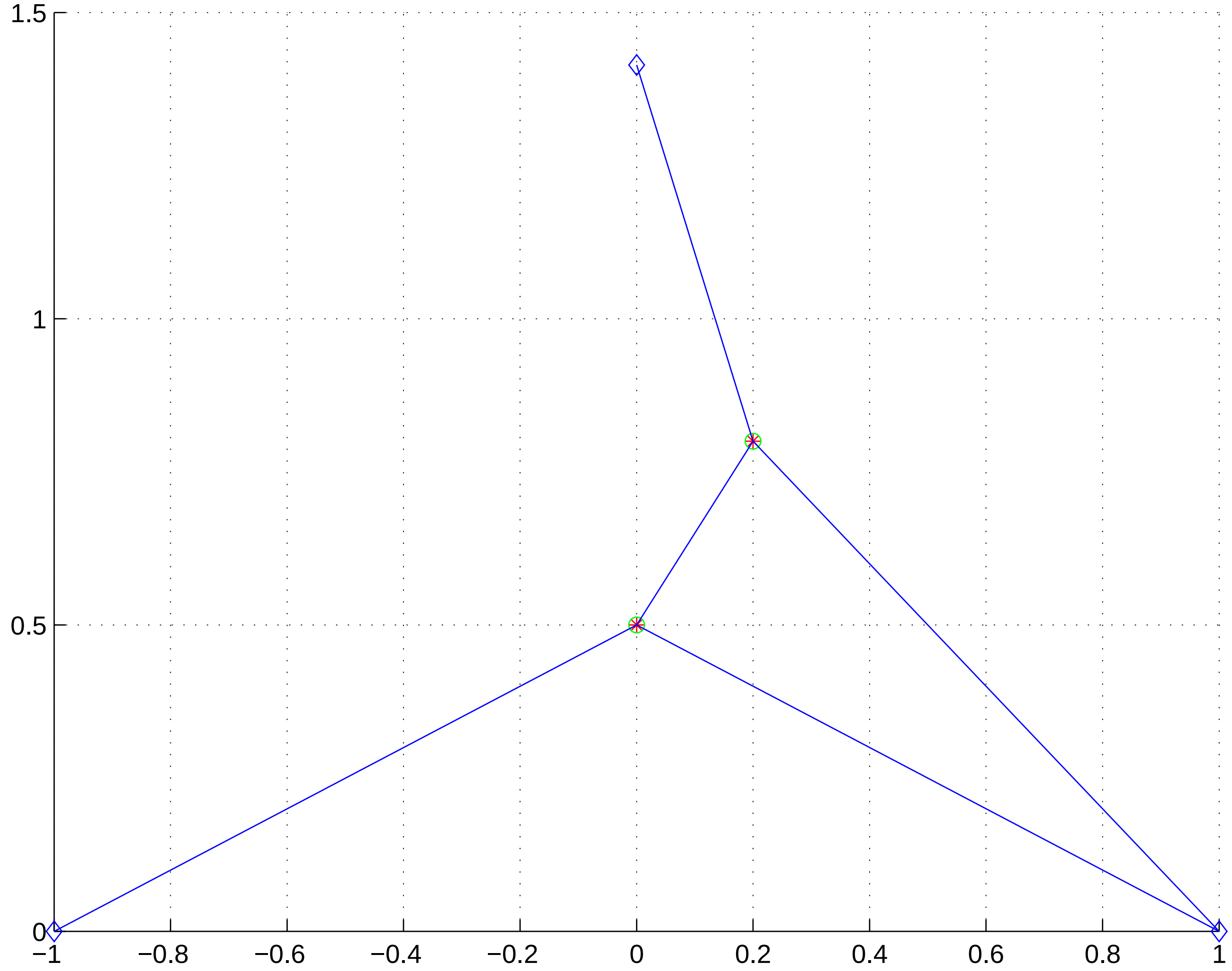


Figure 3: Two sensor-Three anchors: Universally Localizable (but not Strongly)

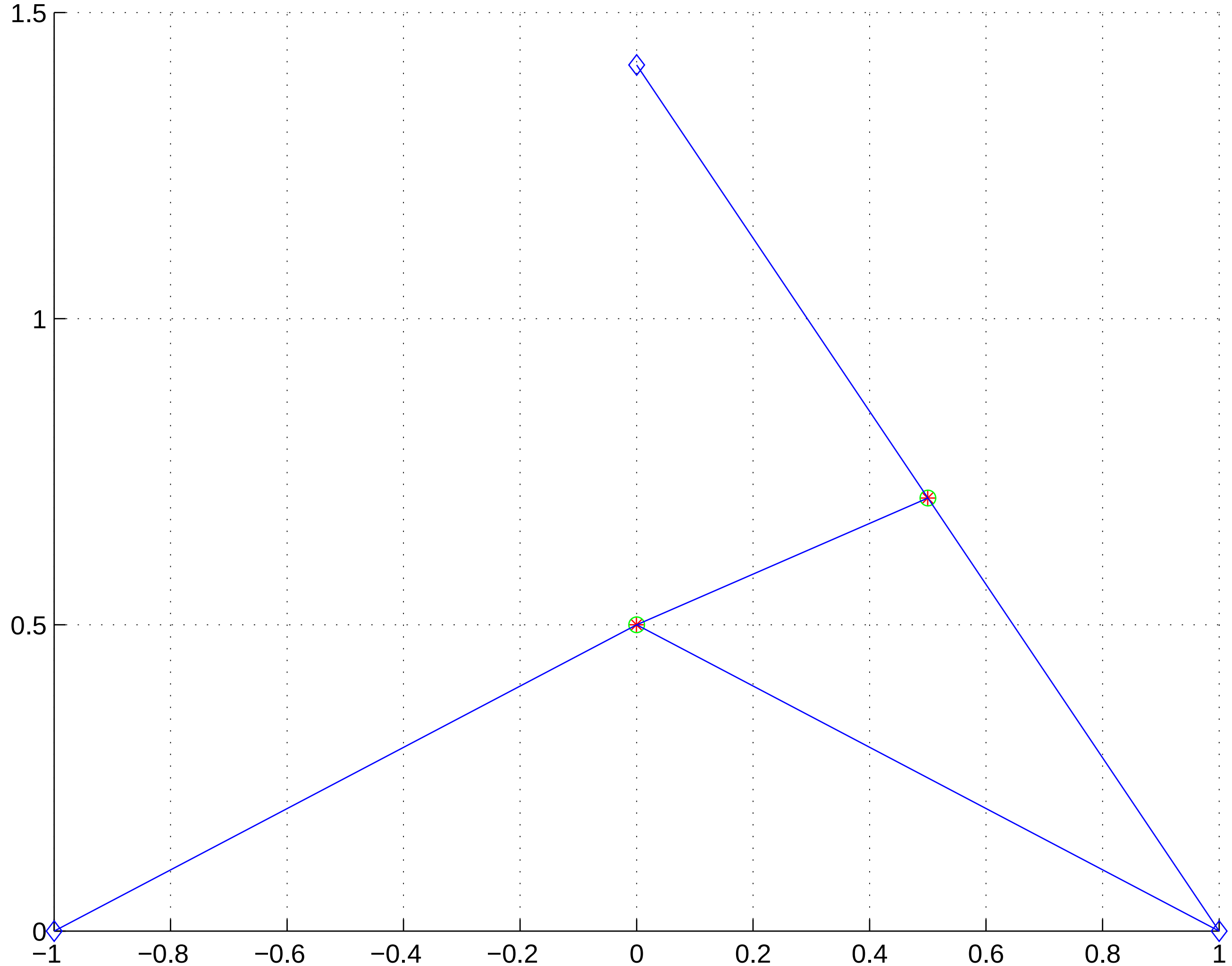


Figure 4: Two sensor-Three anchors: Not Universally Localizable

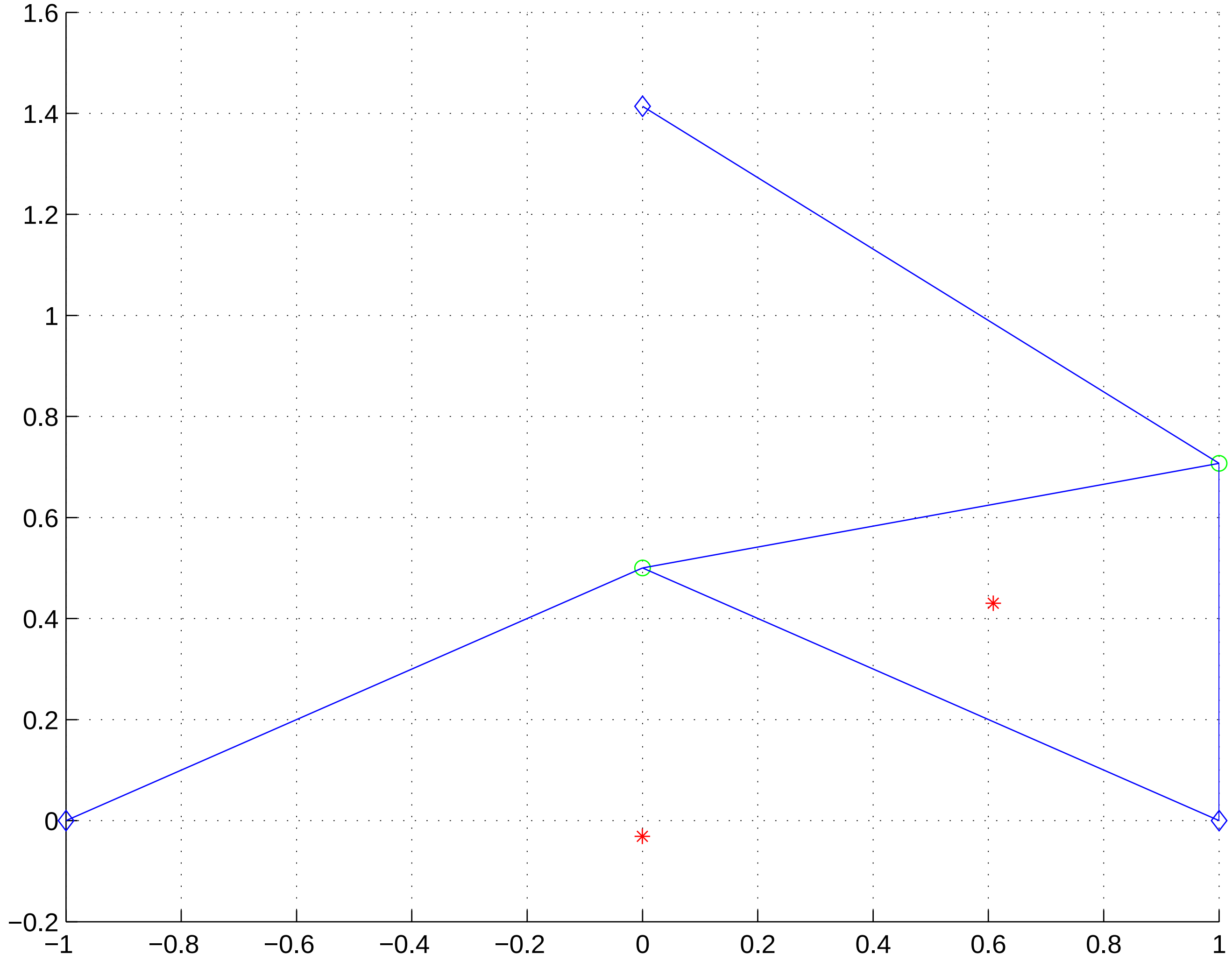
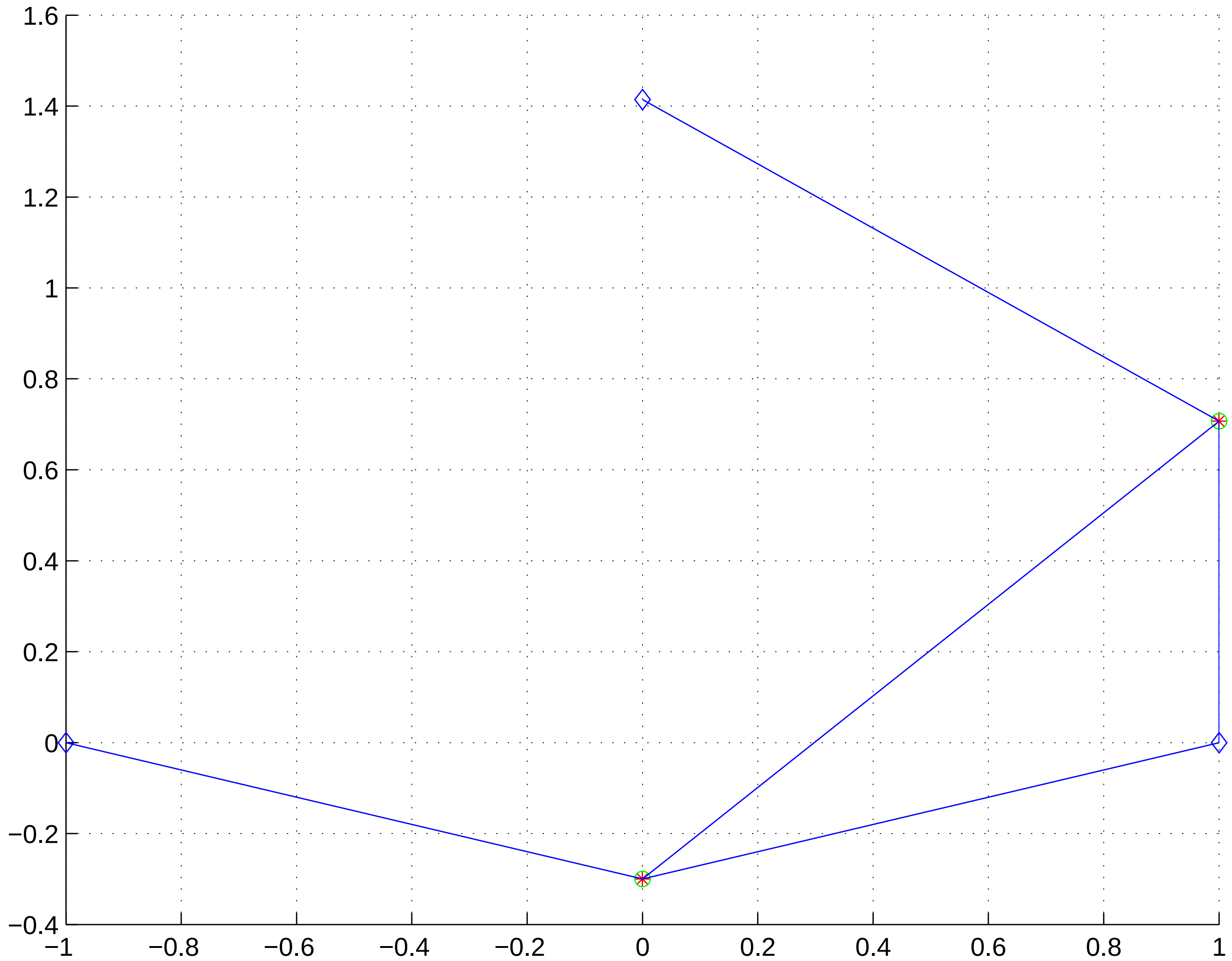


Figure 5: Two sensor-Three anchors: Universally Localizable



## Universally-Localizable Problems (ULP)

**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 $Tr(Y - X^T X) = 0$.*

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

$$
\begin{array}{lllll}
\max & \mathbf{0}^T\mathbf{x} & & & \\
\text{s.t.} & & \delta_1 & = d_1 & \\
 & \mathbf{x}+ & \mathbf{s}_1 & = \mathbf{a}_1 & \\
 & & \delta_2 & = d_2 & \\
 & \mathbf{x}+ & \mathbf{s}_2 & = \mathbf{a}_2 & \\
 & & \delta_3 & = d_3 & \\
 & \mathbf{x}+ & \mathbf{s}_3 & = \mathbf{a}_3 & \\
 & & (\delta_k; \mathbf{s}_k) & \in SOCP,\ k = 1, 2, 3. &
\end{array}
$$

This problem is in the standard SOCP dual form.

# Convex Relaxation: SDP

Since $\mathbf{a}_k - \mathbf{x} = [I \;\; \mathbf{x}](\mathbf{a}_k; -1)$ ($I$ here is a $2 \times 2$ identity matrix) so that

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

The original three quadratic equations can be written as

$$(\mathbf{a}_k; -1)(\mathbf{a}_k; -1)^T \bullet \begin{pmatrix} I & \mathbf{x} \\ \mathbf{x}^T & y \end{pmatrix} = d_k^2, \; \forall\, k, j \in N_a,$$
$$y = \mathbf{x}^T\mathbf{x}.$$

Relax $y = \mathbf{x}^T\mathbf{x}$ to $y \succeq \mathbf{x}^T\mathbf{x}$, which is equivalent to matrix positive semi-definiteness:

$$\begin{pmatrix} I & \mathbf{x} \\ \mathbf{x}^T & y \end{pmatrix} \succeq \mathbf{0}.$$

Denote this matrix by $Z$. Then the relaxed problem can be written as SDP in the standard form.

## SDP Standard Form

$$
\begin{array}{ll}
\max & \mathbf{0} \bullet Z \\
\text{s.t.} & (1;0;0)(1;0;0)^T \bullet Z = 1, \\
& (0;1;0)(0;1;0)^T \bullet Z = 1, \\
& (1;1;0)(1;1;0)^T \bullet Z = 2, \\
& (\mathbf{a}_k;-1)(\mathbf{a}_k;-1)^T \bullet Z = d_k^2, \ \text{for } k = 1,2,3, \\
& Z \succeq \mathbf{0}.
\end{array}
$$

Note than $Z$ has rank at least $2$; if it's $2$, then $y = \mathbf{x}^T\mathbf{x}$, and the converse is also true. In particular, unknown

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

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).

## The Dual of SDP

Assign the dual variables to

$$(1;0;0)(1;0;0)^T \bullet Z = 1,\ (w_1)$$
$$(0;1;0)(0;1;0)^T \bullet Z = 1,\ (w_2)$$
$$(1;1;0)(1;1;0)^T \bullet Z = 2,\ (w_3)$$
$$(\mathbf{a}_k;-1)(\mathbf{a}_k;-1)^T \bullet Z = d_k^2,\ (\lambda_k) \text{ for } k=1,2,3.$$

The Dual would be

$$\begin{array}{ll} \min & w_1 + w_2 + 2w_3 + \sum_{k=1}^3 \lambda_k d_k^2 \\ \text{s.t.} & \begin{pmatrix} \begin{pmatrix} w_1 + w_3 & w_3 \\ w_3 & w_2 + w_3 \end{pmatrix} + \sum_{k=1}^3 \lambda_k \mathbf{a}_k \mathbf{a}_k^T & -\sum_{k=1}^3 \lambda_k \mathbf{a}_k \\ -(\sum_{k=1}^3 \lambda_k \mathbf{a}_k)^T & \sum_{k=1}^3 \lambda_k \end{pmatrix} \succeq \mathbf{0}. \end{array}$$

Does the dual has a rank-1 slack matrix, $S$, with zero-objective value?

## An Optimal Dual Slack Matrix

If we choose $(w., \lambda.)$'s such that

$$\bar{S} = (-\bar{\mathbf{x}}; 1)(-\bar{\mathbf{x}}; 1)^T,$$

then, $\bar{S} \succeq \mathbf{0}$ 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

$$\begin{array}{c} \sum_{k=1}^{3} \lambda_k \mathbf{a}_k = \bar{\mathbf{x}} \\ \sum_{k=1}^{3} \lambda_k = 1. \end{array} \quad \text{or} \quad \begin{array}{c} \sum_{k=1}^{3} \lambda_k (\mathbf{a}_k - \bar{\mathbf{x}}) = \mathbf{0} \\ \sum_{k=1}^{3} \lambda_k = 1. \end{array}$$

This system always has an unique solution as long as $\mathbf{a}_k$'s are not co-linear.

Then we choose (unique) $w.$'s such that

$$\begin{pmatrix} w_1 + w_3 & w_3 \\ w_3 & w_2 + w_3 \end{pmatrix} = \bar{\mathbf{x}}_1 \bar{\mathbf{x}}_1^T - \sum_{k=1}^{3} \lambda_k \mathbf{a}_k \mathbf{a}_k^T$$

## Dual Interpretation

$\lambda_k$'s are nontrivial stresses/forces the edges between $\mathbf{a}_k$ and solution $\mathbf{x}$, respectively, and all stresses are balanced or at the equilibrium state.

Even if $\mathbf{a}_k$ is co-linear, the system

$$\sum_{k=1}^{3} \lambda_k(\mathbf{a}_k - \bar{\mathbf{x}}) = \mathbf{0}$$
$$\sum_{k=1}^{3} \lambda_k = 1$$

may still have a solution $\lambda$.?

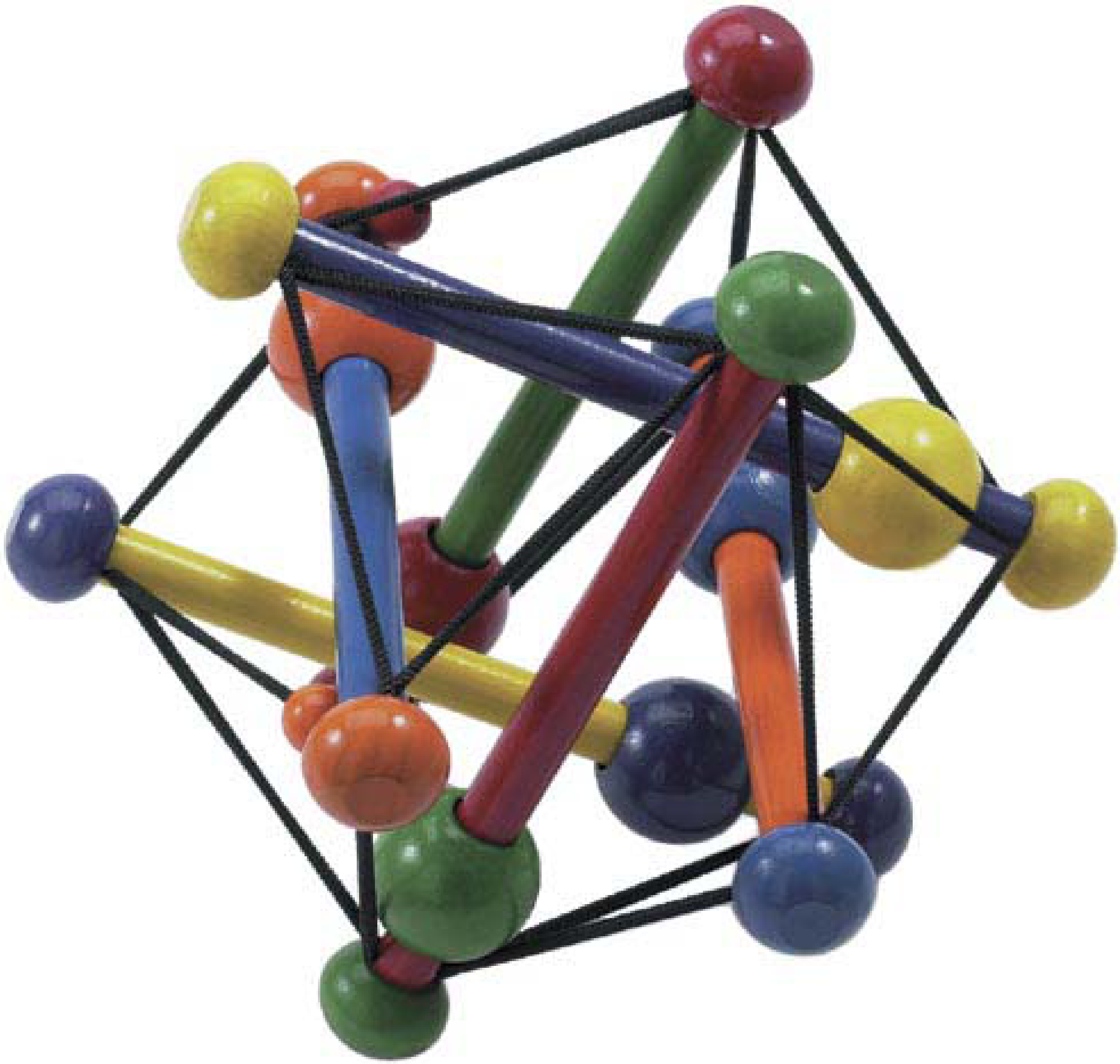


Figure 6: Dual Stresses – A 3-D Toy; provided by Anstreicher

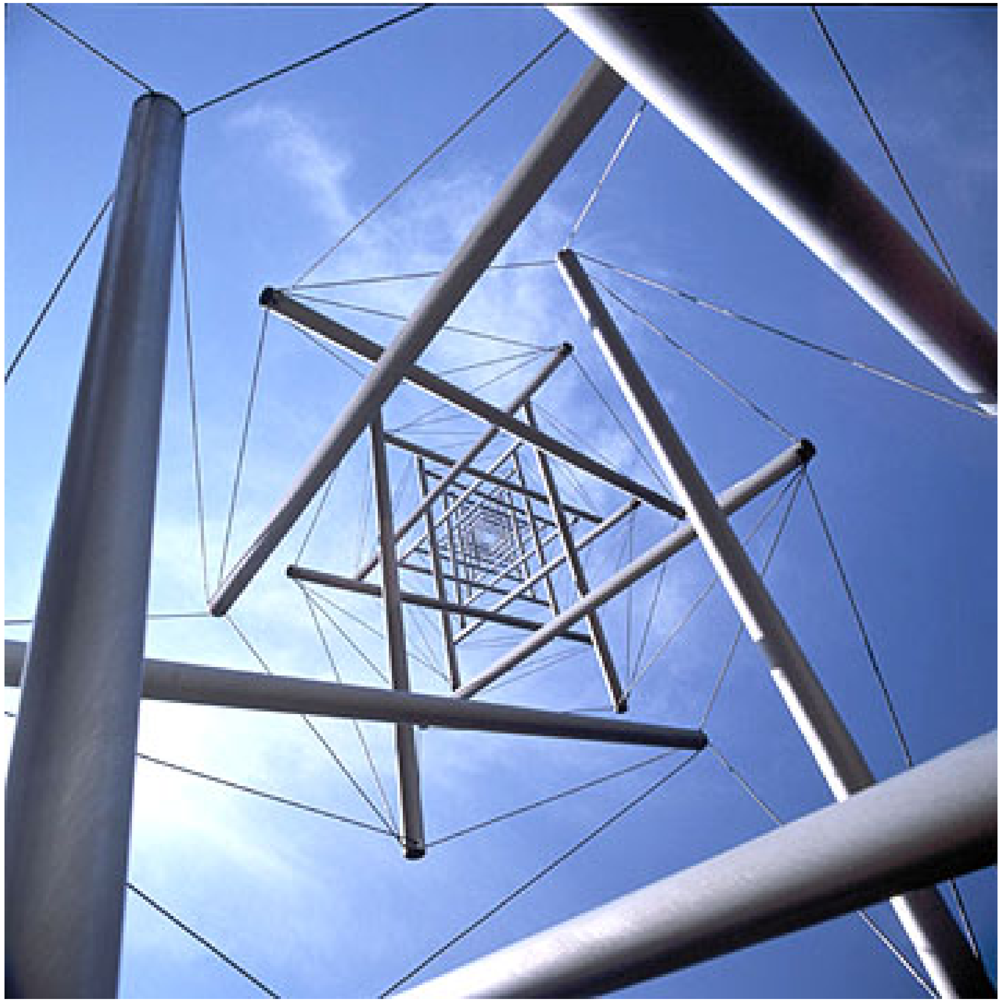


Figure 7: Dual Stresses – A Needle Tower; provided by Anstreicher

## Rank-Reduction for SDP

In most applications, we may not be lucky and need an effort to search a rank-minimal SDP solution for SDP:

$$(SDP) \quad \text{min} \quad C \bullet X$$
$$\text{subject to} \quad A_i \bullet X = b_i, i = 1, 2, ..., m, \ X \succeq 0,$$

where $C, \ A_i \in \mathcal{S}^n$.

Or simply for the SDP feasibility problem:

$$\text{Solve} \quad A_i \bullet X = b_i, i = 1, 2, ..., m, \ X \succeq 0,$$

# A Bound on Support/Rank

**Theorem 8** *(Carathéodory's theorem)*

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

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

$$\begin{aligned} (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T \bullet X &= 1, \ \forall i < j = 1, 2, 3, 4, \\ X &\succeq 0, \end{aligned}$$

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 $\mathbf{x}^0$ and, without loss of generality, assume $\mathbf{x}^0 > \mathbf{0}$, and let $k = 0$ and $A^0 = A$.

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

3. Eliminate the the variable(s) in $\mathbf{x}^{k+1}$ and column(s) in $A^k$ corresponding to $x_j^{k+1} = 0$, 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 $X^*$ be an optimal solution. Then, if the rank of $X^*$, $r$, satisfies the inequality, we need do nothing. Thus, we assume $r(r+1)/2 > m$, and let

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

Then consider

$$\begin{array}{ll} \text{Minimize} & VCV^T \bullet U \\ \text{Subject to} & VA_iV^T \bullet U = b_i, \; i = 1, ..., m \\ & U \succeq 0. \end{array} \qquad (1)$$

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

$$VCV^T \bullet I = C \bullet V^T V = C \bullet X^* = z^*.$$

Moreover, for any feasible solution of (1) one can construct a feasible matrix solution for (**??**) using

$$X(U) = V^T UV \quad \text{and} \quad C \bullet X(U) = VCV^T \bullet U. \tag{2}$$

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)

$$\begin{array}{rll} z^* := & \text{Maximize} & \mathbf{b}^T\mathbf{y} = \sum_{i=1}^m b_i y_i \\ & \text{Subject to} & VCV^T \succeq \sum_{i=1}^m y_i VA_iV^T. \end{array} \tag{3}$$

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

$$I \bullet (VCV^T - \sum_{i=1}^m y_i^* VA_iV^T) = 0$$

so that we have

$$VCV^T - \sum_{i=1}^{m} y_i^* VA_iV^T = \mathbf{0}.$$

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

$$VA_iV^T \bullet W = 0,\ i = 1, ..., m$$

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 $W \neq \mathbf{0}$ 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

$$U(\alpha^*) \succeq \mathbf{0}$$

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

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

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 $O(m^2 n)$ arithmetic operations and finds the least eigenvalue of $W$, which is a polynomial time.

## II. The Principle-Component or Eigenvalue Reduction

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

$$\bar{X} = \sum_{i=1}^{r} \lambda_i \mathbf{v}_i \mathbf{v}_i^T$$

where

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

Then, let

$$\hat{X} = \sum_{i=1}^{d} \lambda_i \mathbf{v}_i \mathbf{v}_i^T$$

# III. Continuous Randomized Reduction

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

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

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

Then, let random matrix

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

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

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

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

$$E[\hat{X}] = VE[R]V^T = VV^T = \bar{X}.$$

## Approximate Low-Rank SDP Theorem

For simplicity, consider the SDP feasibility problem

$$A_i \bullet X = b_i \qquad i = 1, \ldots, m, \quad X \succeq \mathbf{0}$$

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

$$x_1 + x_2 + x_3 = 1,$$
$$\begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \succeq \mathbf{0}.$$

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

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

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

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

$$\alpha(m,n,d) = \begin{cases} 1 + \dfrac{12\ln(4mr)}{d} & \textit{for } 1 \leq d \leq 12\ln(4mr) \\ 1 + \sqrt{\dfrac{12\ln(4mr)}{d}} & \textit{for } d > 12\ln(4mr) \end{cases}$$

*and*

$$\beta(m,n,d) = \begin{cases} \dfrac{1}{e(2m)^{2/d}} & \textit{for } 1 \leq d \leq 4\ln(2m) \\ \max\left\{\dfrac{1}{e(2m)^{2/d}},\ 1 - \sqrt{\dfrac{4\ln(2m)}{d}}\right\} & \textit{for } d > 4\ln(2m) \end{cases}$$

## 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 upp 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?

# IV. $\{-1, 1\}$ Randomized Reduction

Let $X$ be an SDP solution with rank $r$ and

$$X = VV^T.$$

Then, let random vector

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

where

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

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

$$\mathsf{E}[\hat{x}_i \hat{x}_j] = \frac{2}{\pi} \arcsin(\bar{X}_{ij}), \quad i, j = 1, 2, \ldots, n.$$

## Max-Cut Problem

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.

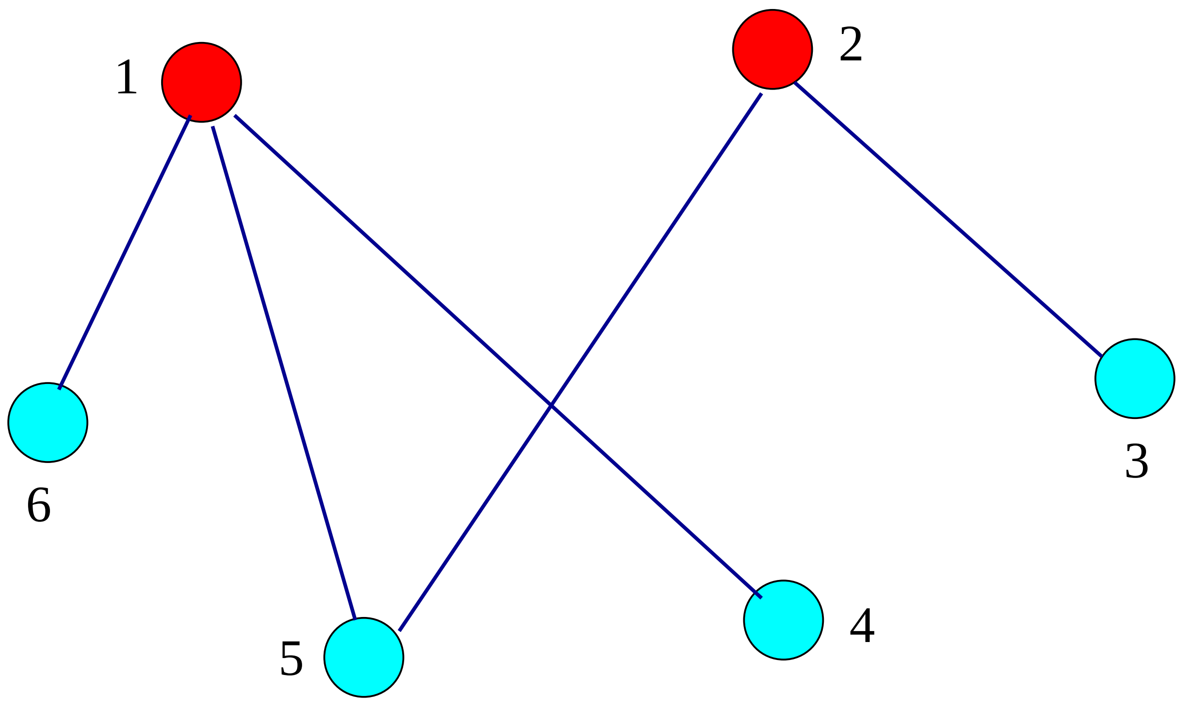



Figure 8: Illustration of the Max-Cut Problem

## Max-Cut Formulation with Binary Quadratic Minimization

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

(MC)

$$\text{Subject to} \quad (x_j)^2 = 1, \; j = 1, \ldots, n.$$

## The Coin-Toss Method: Approximation Quality

Let each node be selected to one side, or $\hat{x}_j$ be 1, independently with probability $.5$.

Or simply let random vector

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

We have

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

$$= \frac{1}{4}\sum_{i,j} w_{ij} = \frac{\text{weights of all edges}}{2} \geq \frac{1}{2} w^*.$$

# Semidefinite Relaxation for (MC)

Let $X = \mathbf{x}\mathbf{x}^T \in S^n_+$. Then the problem can be rewritten as

$$
\begin{aligned}
z^{SDP} := \quad & \text{Maximize} \quad \frac{1}{4}\sum_{i,j} w_{ij}(1 - X_{ij}) \\
& \text{Subject to} \quad X_{ii} = 1, \quad i = 1, \ldots, n, \\
& \qquad\qquad\qquad X \succeq \mathbf{0}, \ \text{rank}(X) = 1.
\end{aligned}
$$

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 $\mathbf{u} \in N(0, \bar{X})$:

$$\hat{\mathbf{x}} = \text{Sign}(\mathbf{u}), \qquad \text{E}[\hat{x}_i \hat{x}_j] = \arcsin(\bar{X}_i j)$$

**Theorem 10** *(Goemans and Williamson)*

$$\text{E}[w(\hat{\mathbf{x}})] \geq .878 z^{SDP} \geq .878 w^*.$$

## V. Objective-Guided Reduction

Construct a suitable objective for the SDP solution set

$$
\begin{array}{lll}
\text{Minimize} & R \bullet X & \\
\text{Subject to} & A_i \bullet X & = b_i, \; i = 1, \ldots, m, \\
& C \bullet X & \leq \alpha \cdot z^*, \\
& X \succeq \mathbf{0}, &
\end{array}
$$

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

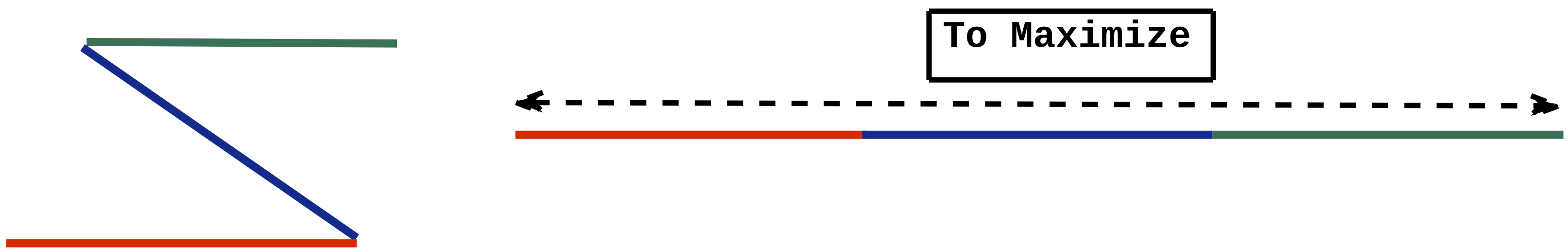
The selection of matrix $R$ is problem dependent. Examples include the $L_1$ norm function, the tensegrity graph approach, etc.

## Tensegrity (Tensional-Integrity) Objective for SNL: a Chain Graph

Anchor-free SNL: let $\mathbf{e}_i$ be the unit vector (one for the $i$th entry and zeros for the else)

$$\begin{aligned} (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T \bullet X &= d_{ij}^2, \ \forall\, (i,j) \in E,\ i < j, \\ X &\succeq \mathbf{0}. \end{aligned}$$

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.



## The Chain Graph Example

Consider:

$$\begin{array}{ll} \text{max} & \mathbf{e}_3\mathbf{e}_3 \bullet X \\ \text{s.t.} & \mathbf{e}_1\mathbf{e}_1^T \bullet X = 1, \\ & (\mathbf{e}_1 - \mathbf{e}_2)(\mathbf{e}_1 - \mathbf{e}_2)^T \bullet X = 1, \\ & (\mathbf{e}_2 - \mathbf{e}_3)(\mathbf{e}_2 - \mathbf{e}_3)^T \bullet X = 1, \\ & X \succeq \mathbf{0} \in \mathcal{S}^3, \end{array}$$

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

$$\begin{array}{ll} \text{min} & y_1 + y_2 + y_3 \\ \text{s.t.} & y_1\mathbf{e}_1\mathbf{e}_1^T + y_2(\mathbf{e}_1 - \mathbf{e}_2)(\mathbf{e}_1 - \mathbf{e}_2)^T + y_3(\mathbf{e}_2 - \mathbf{e}_3)(\mathbf{e}_2 - \mathbf{e}_3)^T - S = \mathbf{e}_3\mathbf{e}_3, \\ & S \succeq \mathbf{0} \in \mathcal{S}^3, \end{array}$$

The dual has a rank-two solution with $(y_1 = 3,\ y_2 = 3,\ y_3 = 3)$.

## Applications

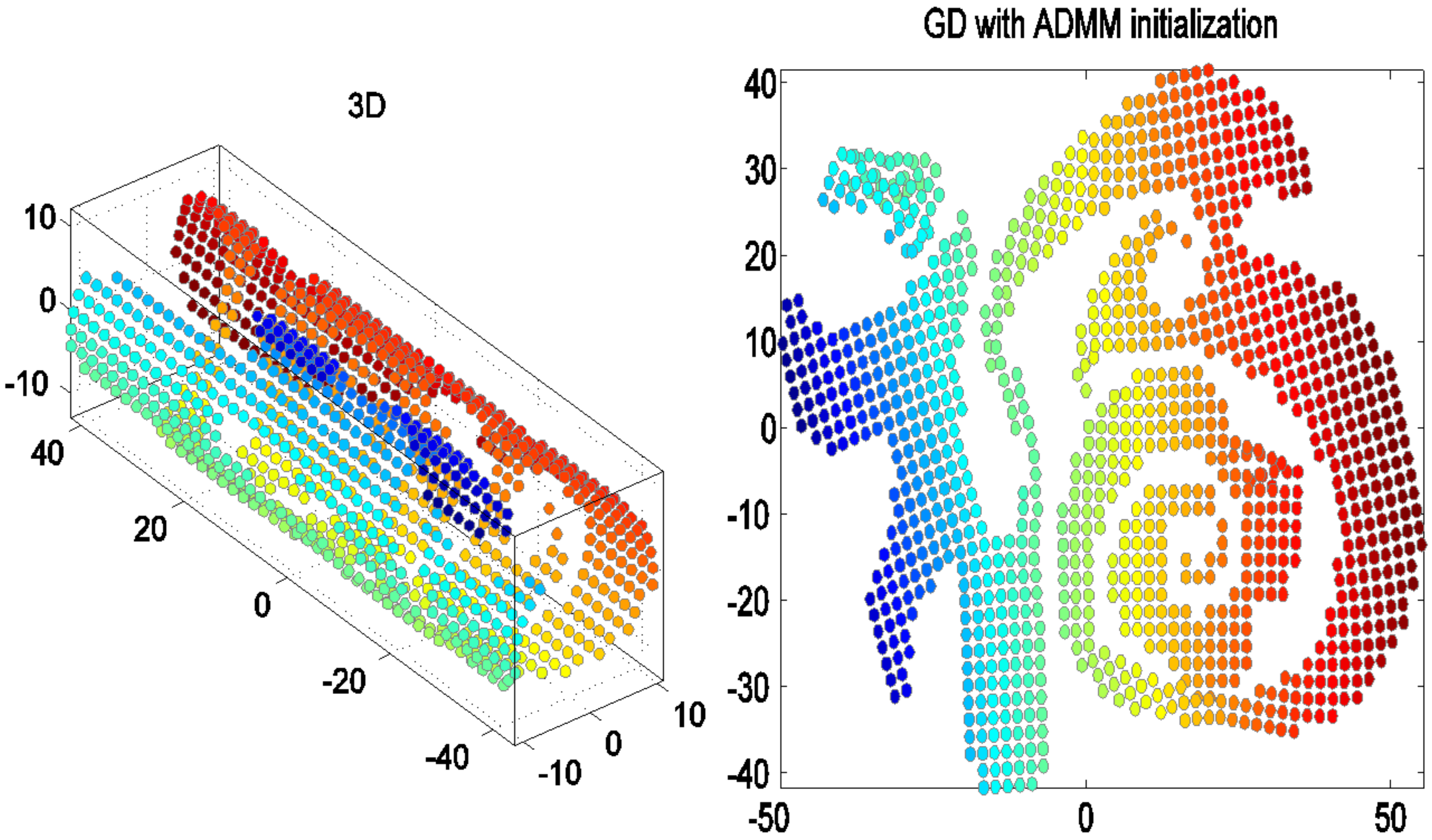



Figure 9: Dimension Reduction – Unfolding Scroll of Happiness

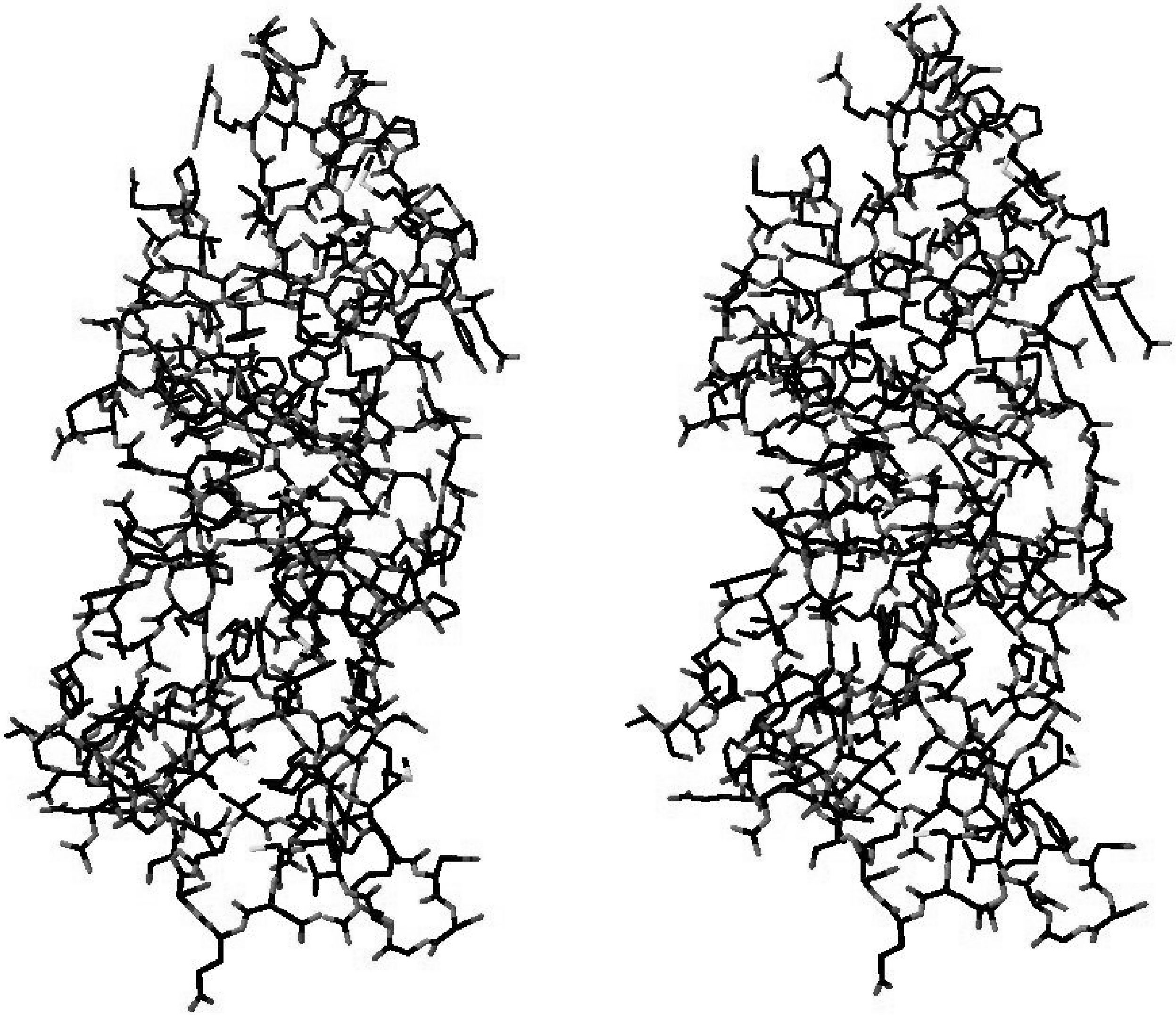


Figure 10: Molecular Conformation – 1F39(1534 atoms) with 85% of distances below $6$**r**A and 10% noise on upper and lower bounds