Dual Interpretations and Duality Applications (continued II)

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

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

Recall the Facility Location Problem

Let $\mathbf{c}_j$ be the location of client $j = 1, 2, ..., m$, and $\mathbf{y}$ be the location decision of a facility to be built.

$$\text{minimize}_{\mathbf{y}} \quad \sum_j \|\mathbf{y} - \mathbf{c}_j\|_2.$$

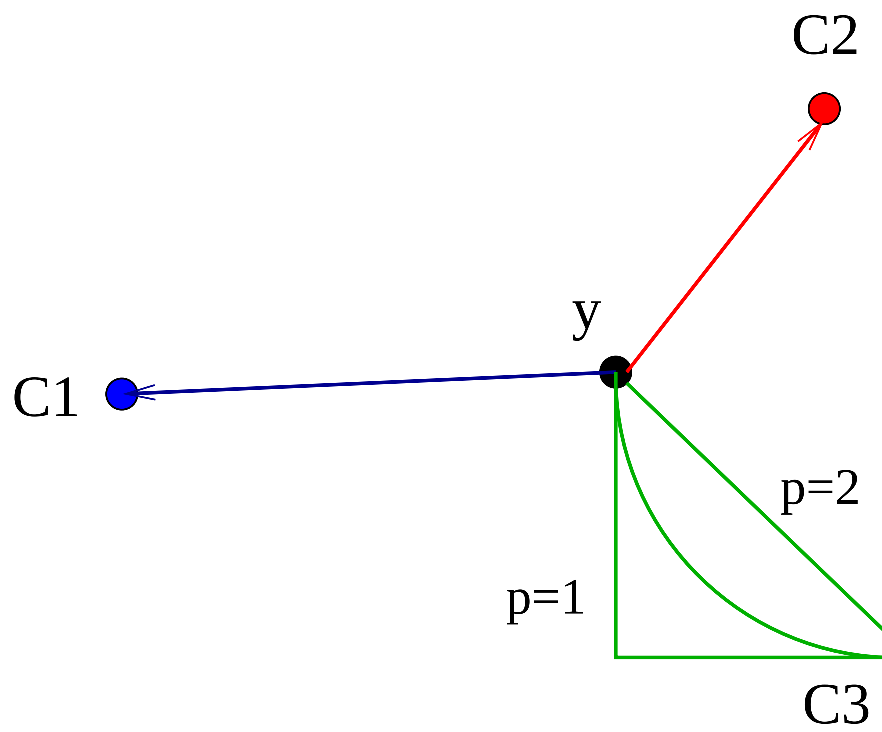

Figure 1: Facility Location at Point $\mathbf{y}$.

Conic Formulation of the Facility Location Problem

$$\begin{array}{ll} \text{minimize} & \sum_j \delta_j \\ \text{subject to} & \mathbf{y} + \mathbf{x}_j = \mathbf{c}_j, \ (\mathbf{z}_j) \ (\delta_j; \mathbf{x}_j) \in SOCP, \ \forall j. \end{array}$$

The Dual:

$$\begin{array}{ll} \text{maximize} & \sum_j \mathbf{c}_j^T \mathbf{z}_j \\ \text{subject to} & \sum_j \mathbf{z}_j = \mathbf{0} \ (\mathbf{y}) \ (1; \mathbf{z}_j) \in SOCP, \ \forall j. \end{array}$$

Let $\mathbf{y}*$ be the optimal location. Then the dual is equivalent to

$$\begin{array}{ll} \text{maximize} & \sum_j (\mathbf{c}_j - \mathbf{y}^*)^T \mathbf{z}_j \\ \text{subject to} & \sum_j \mathbf{z}_j = \mathbf{0} \ (\mathbf{y}) \\ & (1; \mathbf{z}_j) \in SOCP, \ ((\delta_j; \mathbf{x}_j)) \ \forall j. \end{array}$$

The optimality condition would have

$$\mathbf{z}_j^* = (\mathbf{c}_j - \mathbf{y}^*) / \|(\mathbf{c}_j - \mathbf{y}^*)\|, \ \forall j$$

Portfolio Management

Let $\mathbf{r}$ denote the expected return vector and V denote the co-variance matrix of an investment portfolio, and let $\mathbf{x}$ be the investment proportion vector. Then, one management model is:

$$\begin{array}{ll} \text{minimize} & \mathbf{x}^T V \mathbf{x} \\ \text{subject to} & \mathbf{r}^T \mathbf{x} \geq \mu, \ \mathbf{e}^T \mathbf{x} = 1, \ \mathbf{x} \geq \mathbf{0}. \end{array}$$

where $\mathbf{e}$ is the vector of all ones. This is a quadratic program.

Let $V = R^T R$ and $\mathbf{z} = R\mathbf{x}$ (Cholesky Factor Matrix). Then the problem can be written as

$$\begin{array}{ll} \text{minimize} & y_0 \\ \text{subject to} & \mathbf{r}^T \mathbf{x} \geq \mu, \\ & \mathbf{e}^T \mathbf{x} = 1, \\ & R\mathbf{x} - \mathbf{y} = \mathbf{0}, \\ & \mathbf{x} \geq \mathbf{0}, \ \|\mathbf{y}\|_2 \leq y_0 \end{array}$$

which is a mixed linear and second-order cone program.

The Dual of Portfolio Management

Write the problem in the standard form:

$$\begin{array}{ll} \text{minimize} & y_0 \\ \text{subject to} & \mathbf{r}^T\mathbf{x} - s = \mu, \ (\lambda) \\ & \mathbf{e}^T\mathbf{x} = 1, \ (\gamma) \\ & R\mathbf{x} - \mathbf{y} = \mathbf{0}, \ (\mathbf{z}) \\ & \mathbf{x} \geq \mathbf{0}, \ s \geq 0, \ (y_0; \mathbf{y}) \in SOCP \end{array}$$

The dual would be

$$\begin{array}{ll} \text{maximize} & \mu\lambda + \gamma \\ \text{subject to} & -\mathbf{r}\lambda - \mathbf{e}\gamma - R^T\mathbf{z} \geq \mathbf{0} \ (\mathbf{x}) \\ & \lambda \geq 0, \ (s) \\ & (1; \mathbf{z}) \in SOCP \ ((y_0; \mathbf{y})) \end{array}$$

Robust Portfolio Management

In real applications, $\mathbf{r}$ and V may be estimated under various scenarios, say $\mathbf{r}_i$ and V_i for $i = 1, ..., m$.

$$\begin{array}{ll} \text{minimize} & \max_i \mathbf{x}^T V_i \mathbf{x} \\ \text{subject to} & \min_i \mathbf{r}_i^T \mathbf{x} \geq \mu, \\ & \mathbf{e}^T \mathbf{x} = 1, \ \mathbf{x} \geq \mathbf{0}. \end{array}$$

$$\begin{array}{ll} \text{minimize} & y_0 \\ \text{subject to} & \mathbf{r}_i^T \mathbf{x} \geq \mu, \ \forall i \\ & \|R_i \mathbf{x}\|_2 \leq y_0, \ \forall i \\ & \mathbf{e}^T \mathbf{x} = 1, \ \mathbf{x} \geq \mathbf{0}, \end{array}$$

where R_i is the Cholesky Factor Matrix of V_i. This can again be reduced to a mixed linear and second-order cone program in the standard form.

Sensor Network Localization (SNL) and Graph Realization

Given a graph $G = (V, E)$ and sets of non–negative weights, say $\{d_{ij} : (i,j) \in E\}$, the goal is to compute a realization of G in the Euclidean space $\mathbf{R}^d$ for a given low dimension d, i.e.

- to place the vertices of G in $\mathbf{R}^d$ such that
- the Euclidean distance between every pair of adjacent vertices (i,j) equals (or bounded) by the prescribed weight $d_{ij} \in E$.

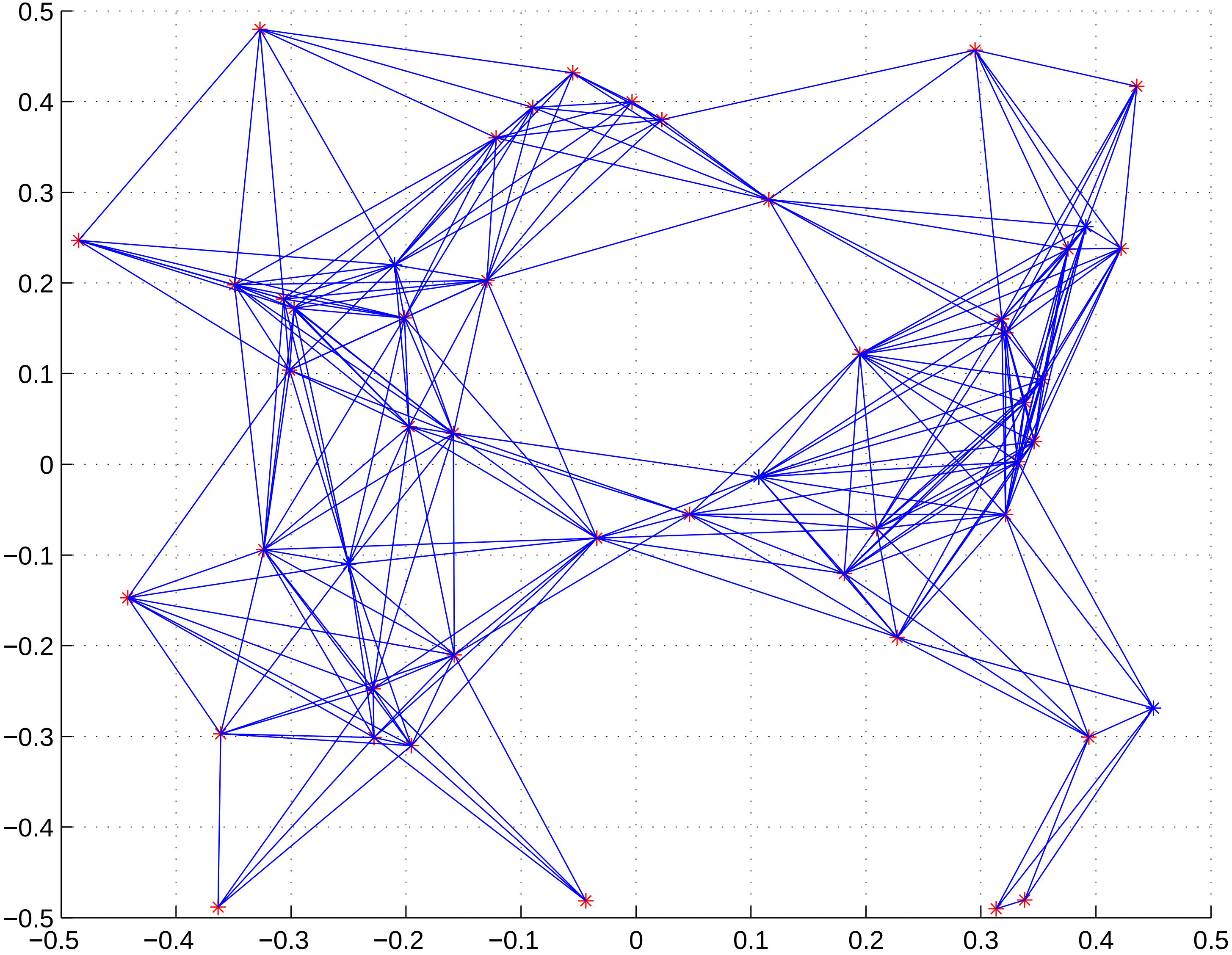

Figure 2: 50-node 2-D Graph Realization

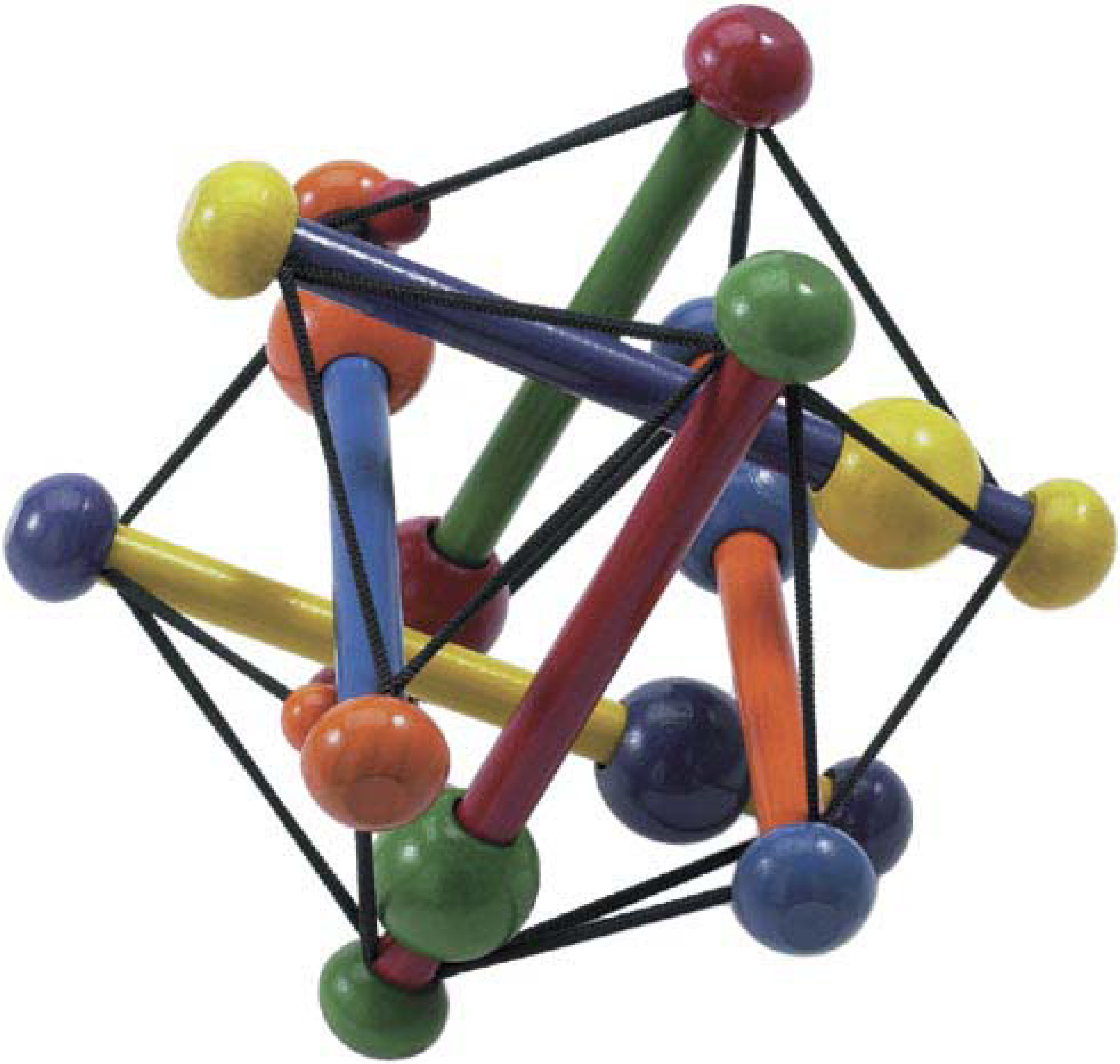

Figure 3: A 3-D Tensegrity Graph Realization

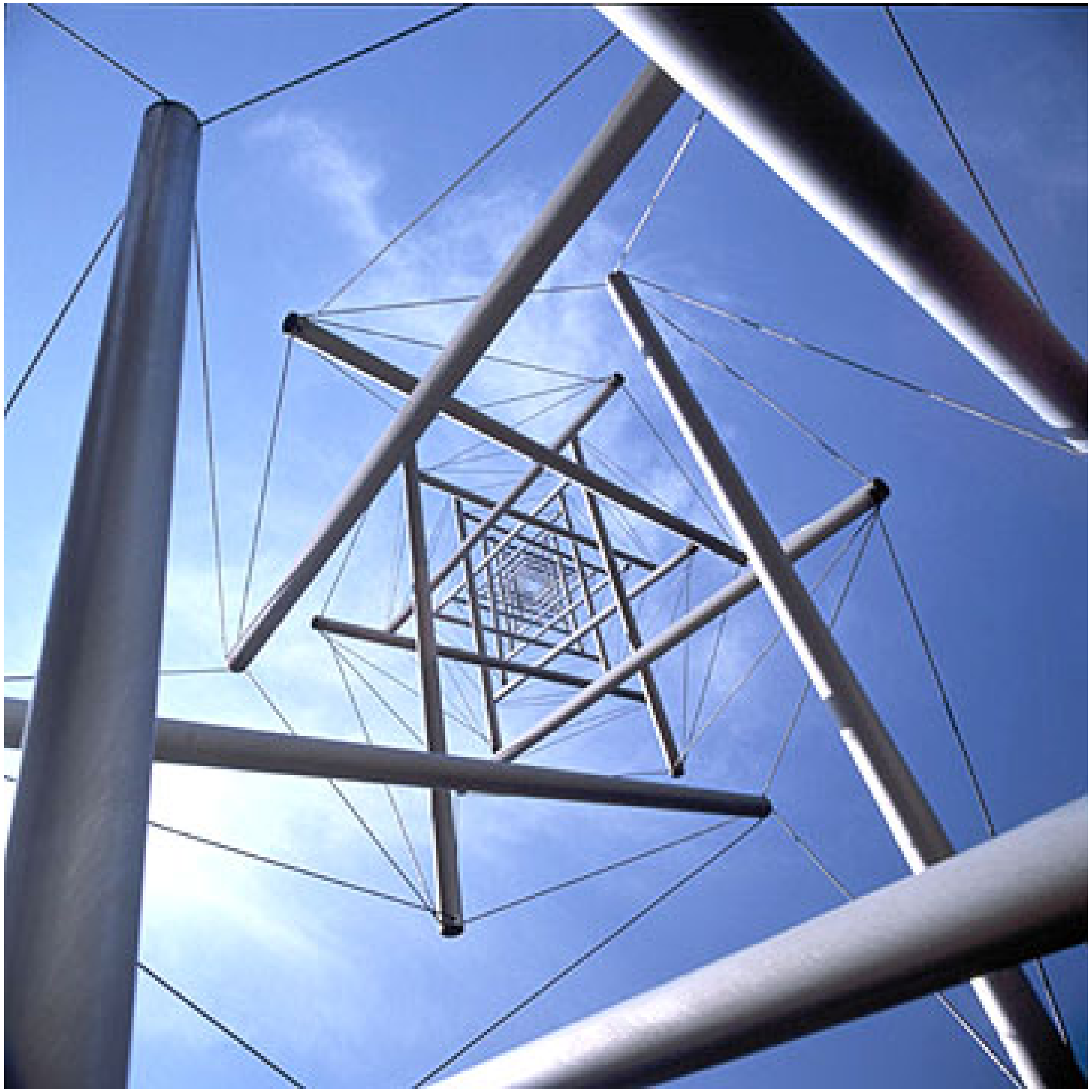

Figure 4: A 3-D Needle Tower

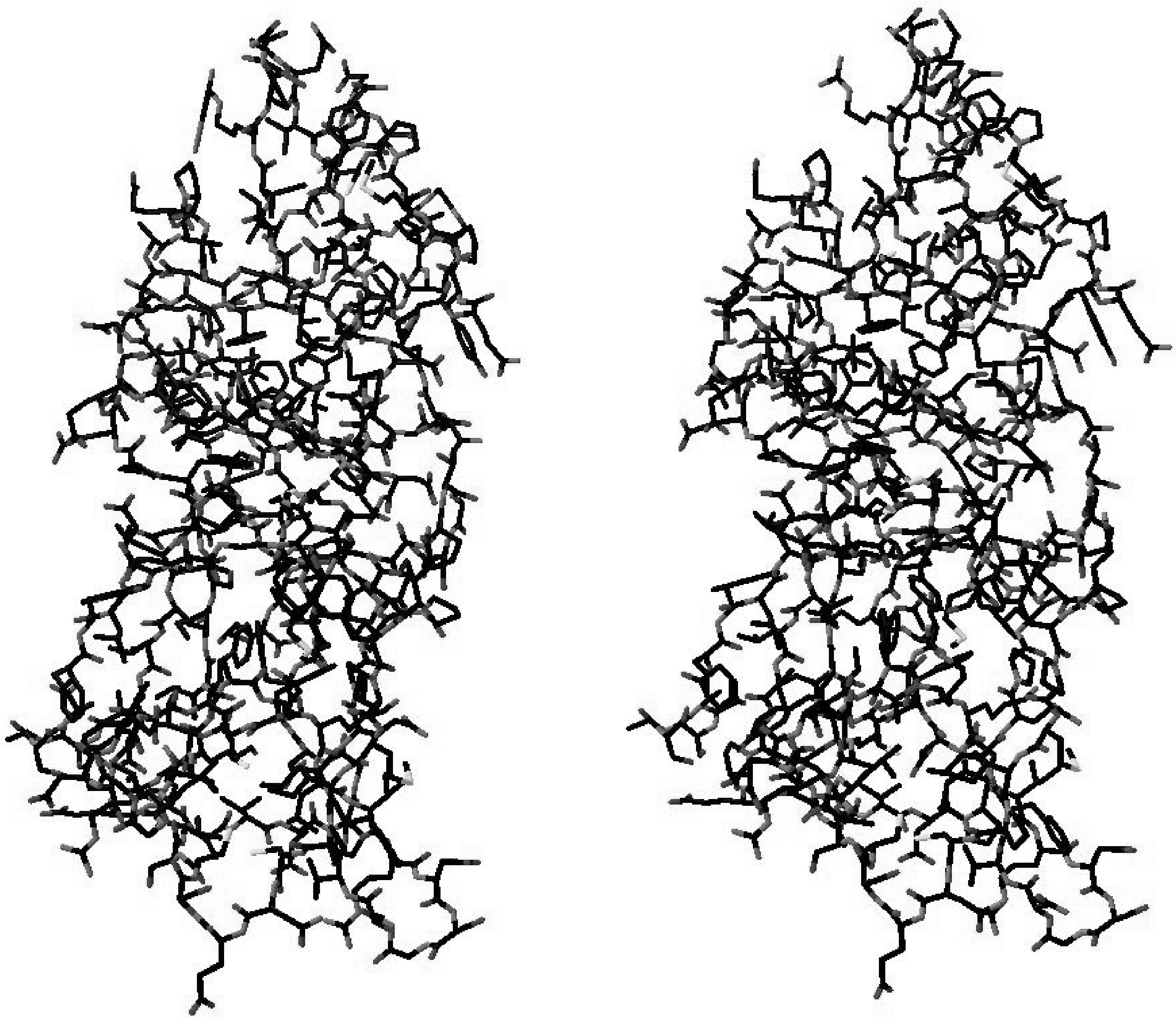

Figure 5: Molecular Conformation: 1F39(1534 atoms) with 85% of distances below $6\mathbf{r}$A and 10% noise on upper and lower bounds

A Distance Geometry Model: System of Quadratic Equations

System of nonlinear equations for $\mathbf{x}_i \in R^d$:

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

where $\mathbf{a}_k$ are possible points whose locations are known, often called anchors.

One can equivalently represent it as

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

which becomes a system of multi-variable-quadratic equations.

Nonlinear Least-Squares Optimization

Nonlinear least-squares or quartic polynomial minimization:

$$\min \quad \sum_{i,j \in N_x} (\|\mathbf{x}_i - \mathbf{x}_j\|^2 - d_{ij}^2)^2 + \sum_{k,j \in N_a} (\|\mathbf{a}_k - \mathbf{x}_j\|^2 - d_{kj}^2)^2$$

or

$$\min \quad \sum_{i,j \in N_x} (\|\mathbf{x}_i - \mathbf{x}_j\| - d_{ij})^2 + \sum_{k,j \in N_a} (\|\mathbf{a}_k - \mathbf{x}_j\| - d_{kj})^2$$

Either one is a non-convex optimization problem.

For simplicity, we assume $d = 2$ in the following analysis.

SOCP Relaxation for SNL

System of SOCP Feasibility for $\mathbf{x}_i \in R^2$:

$$\|\mathbf{x}_i - \mathbf{x}_j\| \le d_{ij},\ \forall\,(i,j) \in N_x,\ i < j,$$
$$\|\mathbf{a}_k - \mathbf{x}_j\| \le d_{kj},\ \forall\,(k,j) \in N_a,$$

where $\mathbf{a}_k$ are points whose locations are known.

Consider the case where a single unknown point $\mathbf{x}_1$ is connected to three anchors $\mathbf{a}_k,\ k = 1, 2, 3$ on R^2:

$$\|\mathbf{a}_k - \mathbf{x}\| \le d_k,\ k = 1, 2, 3$$

The Standard SOCP Relaxation and Dual

$$\begin{array}{ll} \text{minimize} & 0 \\ & \delta_k = d_k, \ (\lambda_k), \ k = 1,2,3 \\ & \mathbf{y}_k + \mathbf{x} = \mathbf{a}_k, \ (\mathbf{z}_k), \ k = 1,2,3 \\ & (\delta_k; \mathbf{y}_k) \in SOCP, \ k = 1,2,3 \end{array}$$

The Dual

$$\begin{array}{ll} \text{maximize} & \sum_k (d_k \lambda_k + \mathbf{a}_k^T \mathbf{z}_k) \\ & \sum_k \mathbf{z}_k = \mathbf{0}, \\ & (-\lambda_k; -\mathbf{z}_k) \in SOCP, \ k = 1,2,3 \end{array}$$

Suppose the true sensor location is $\mathbf{b}$, the dual can be written as

$$\begin{array}{ll} \text{minimize} & \sum_k (-d_k \lambda_k + (\mathbf{a}_k - \mathbf{b})^T \mathbf{z}_k) \\ & \sum_k \mathbf{z}_k = \mathbf{0}, \\ & (\lambda_k; \mathbf{z}_k) \in SOCP, \ k = 1,2,3 \end{array}$$

Optimality Condition of the SOCP Relaxation

The conditions would be

$$\mathbf{z}_k = (\lambda_k/d_k)(\mathbf{a}_k - \mathbf{b})$$

and

$$\sum_k (\lambda_k/d_k)(\mathbf{a}_k - \mathbf{b}) = \mathbf{0}$$

Thus, λ_k represents a positive force in direction $\mathbf{a}_k - \mathbf{b}$, and the total forces should be balanced along the three directions.

If $\mathbf{b}$ is in the convex-hull, this can be achieved so that the optimal solution of the SOCP relaxation is $\mathbf{x}^* = \mathbf{b}$.

What happen if NOT?

SDP Relaxation for SNL

Find a symmetric matrix $Z \in \mathbf{R}^{(2+n)\times(2+n)}$ such that

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

This is semidefinite programming feasibility system (with a null objective).

When this relaxation is exact?

One case is that the single unknown point $\mathbf{x}_1$ is connected to three anchors $\mathbf{a}_k, \; k = 1, 2, 3$.

In general, if the rank of a feasible Z is 2, then it solves the original graph relaxation problem.

Duality Theorem for SNL

Theorem 1 *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 1 *If an optimal dual slack matrix has rank n, then every solution of the SDP has rank 2, that is, the SDP relaxation solves the original problem exactly.*

Theoretical Analyses on SNL-SDP Relaxation

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

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

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

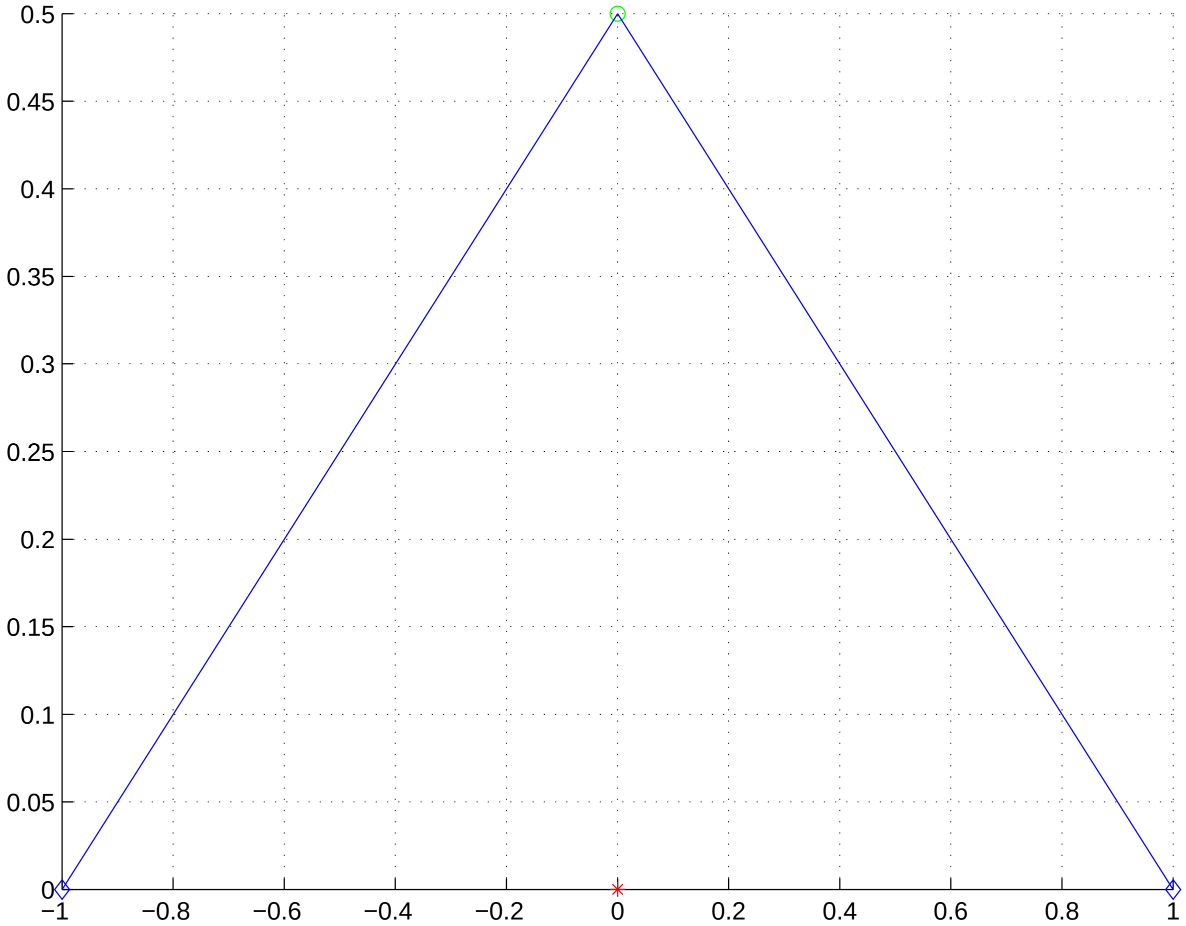

Figure 6: One sensor-Two anchors: Not Localizable

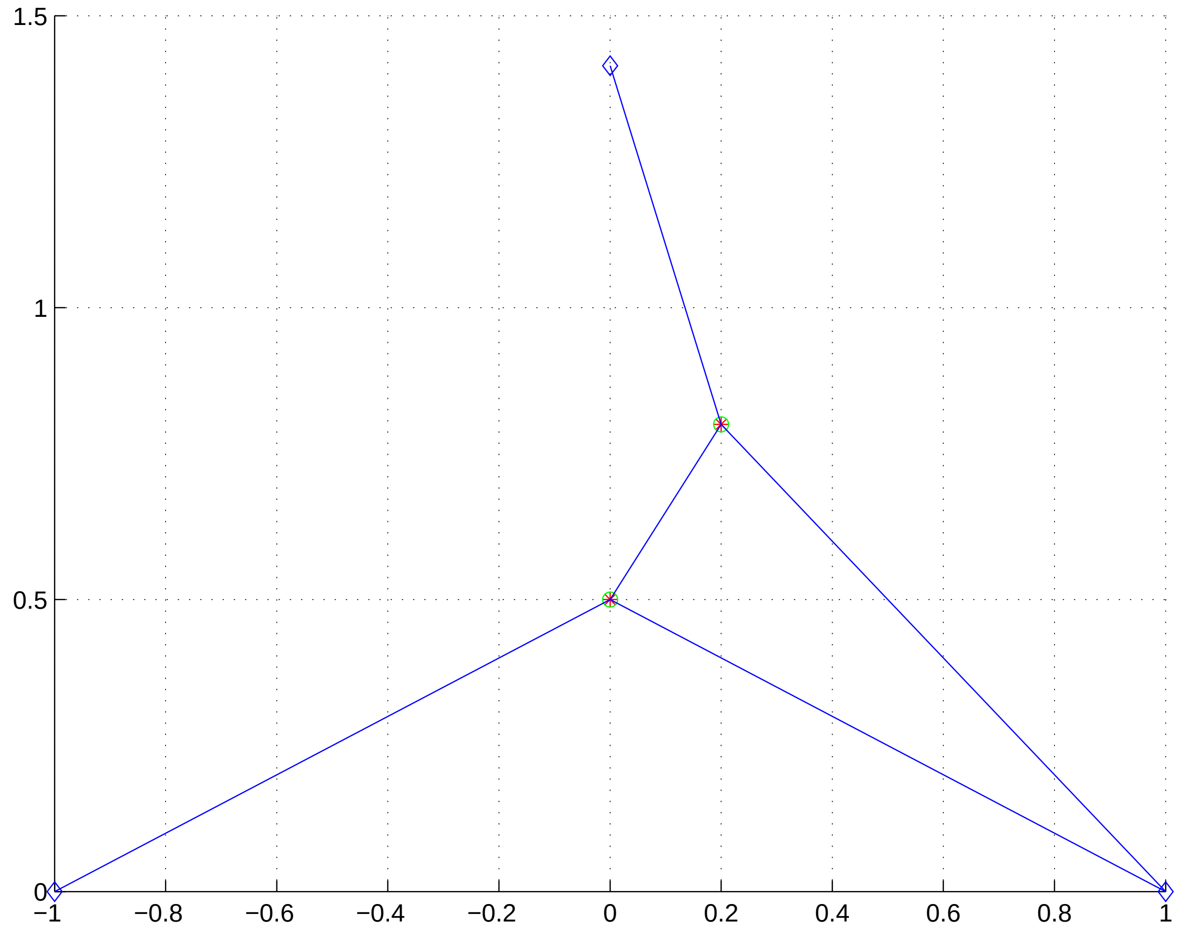

Figure 7: Two sensor-Three anchors: Strongly Localizable

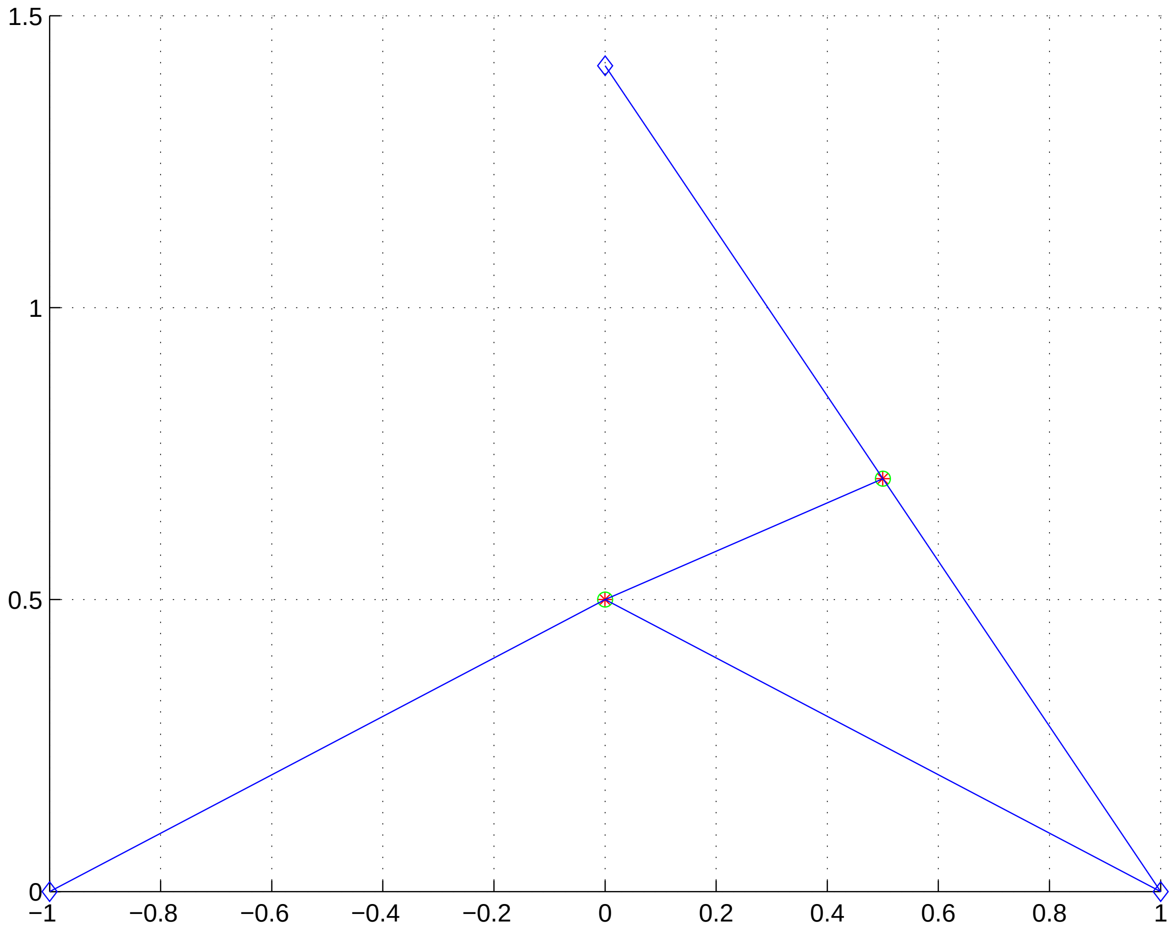

Figure 8: Two sensor-Three anchors: Localizable but not Strongly

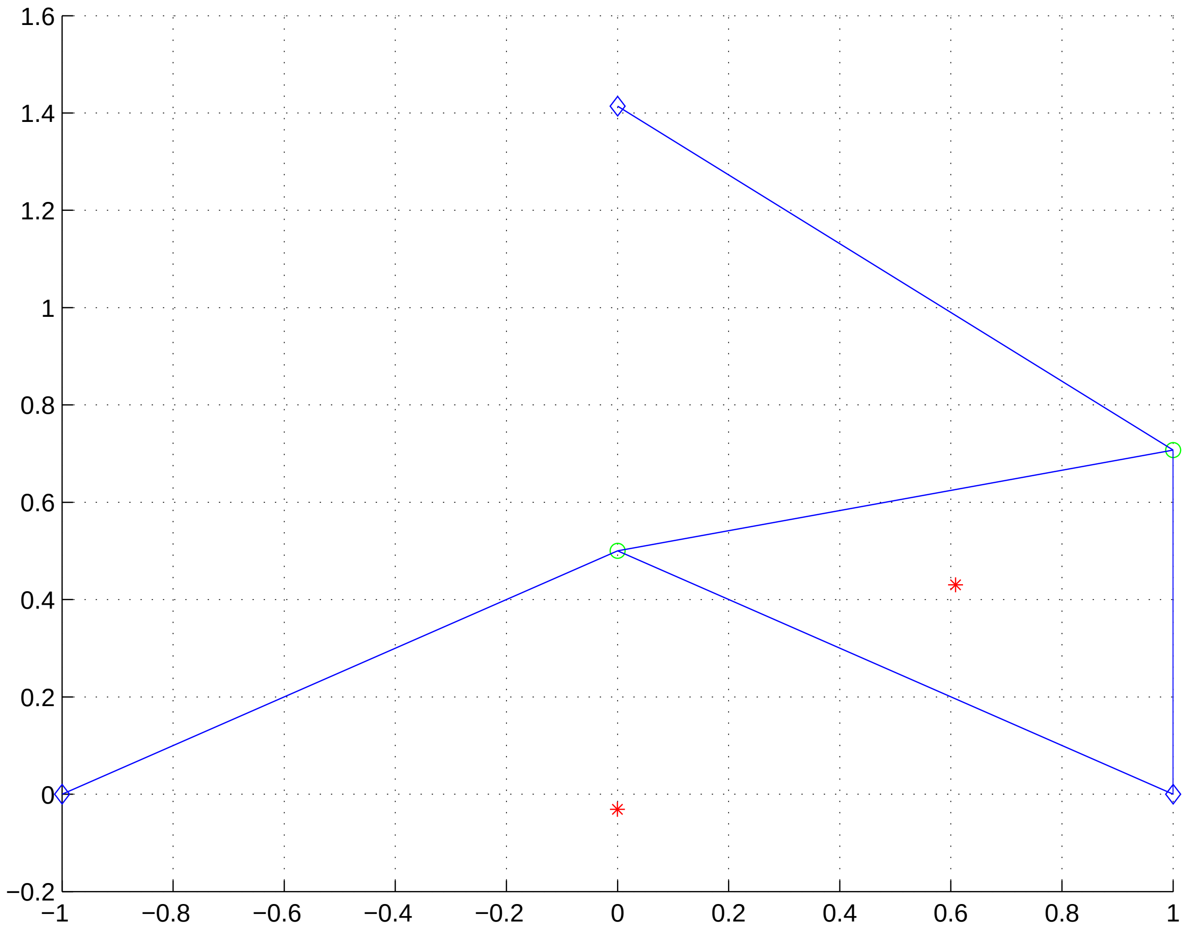

Figure 9: Two sensor-Three anchors: Not Localizable

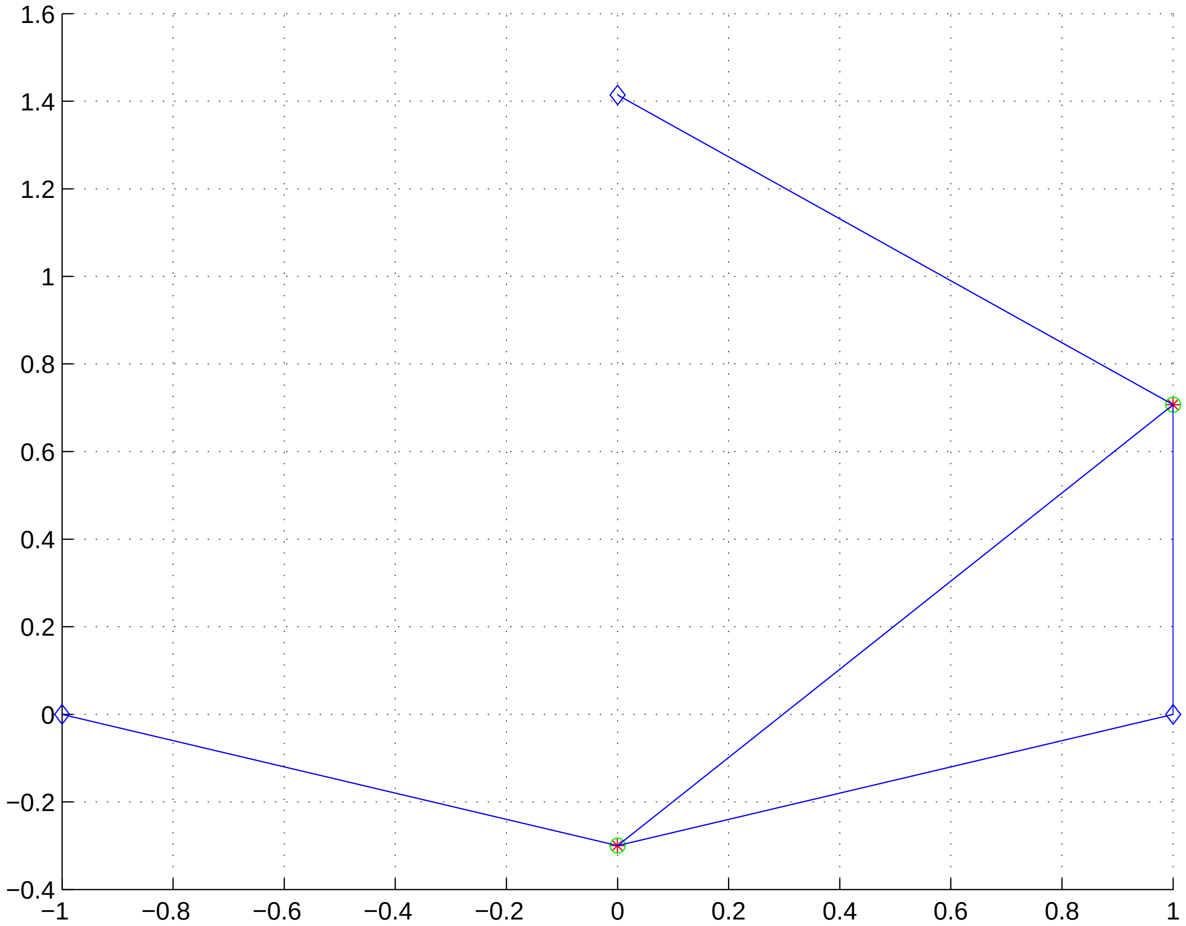

Figure 10: Two sensor-Three anchors: Strongly Localizable

Universally-Localizable Problems (ULP)

Theorem 2 *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 as Convex Optimization

Theorem 3 *(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 and three Anchors

Find $x_1 \in \mathbf{R}^2$ such that

$$\|\mathbf{a}_k - \mathbf{x}_1\|^2 = \hat{d}_{kj}^2, \text{ for } k = 1, 2, 3,$$

Let $\bar{\mathbf{x}}_1$ be the true position of the sensor.

SDP Relaxation Standard Form

$$
\begin{aligned}
&(1;0;0)(1;0;0)^T \bullet Z = 1, \\
&(0;1;0)(0;1;0)^T \bullet Z = 1, \\
&(1;1;0)(1;1;0)^T \bullet Z = 2, \\
&(\mathbf{a}_k;-1)(\mathbf{a}_k;-1)^T \bullet Z = \hat{d}_{k1}^2, \text{ for } k = 1,2,3, \\
&Z \succeq \mathbf{0}.
\end{aligned}
$$

$$
\bar{Z} = \begin{pmatrix} I & \bar{\mathbf{x}}_1 \\ \bar{\mathbf{x}}_1^T & \bar{x}_1^T \bar{x}_1 \end{pmatrix} = (I,\ \bar{\mathbf{x}}_1)^T (I,\ \bar{\mathbf{x}}_1)
$$

is a feasible rank-2 solution for the relaxation.

Dual Slack Matrices

$$\begin{pmatrix} \begin{pmatrix} w_1 + w_3 & w_3 \\ w_3 & w_2 + w_3 \end{pmatrix} + \sum_{k=1}^{3} \hat{w}_{k1} \mathbf{a}_k \mathbf{a}_k^T & -\sum_{k=1}^{3} \hat{w}_{k1} a_k \\ -(\sum_{k=1}^{3} \hat{w}_{k1} a_k)^T & \hat{w}_{11} + \hat{w}_{21} + \hat{w}_{31} \end{pmatrix} \succeq \mathbf{0}.$$

Does an optimal slack matrix U have rank 1 with

$$w_1 + w_2 + 2w_3 + \sum_{k=1}^{3} \hat{w}_{k1} \hat{d}_{k1}^2 = 0?$$

Optimal Dual Slack Matrix

If we choose $w_\bullet$'s such that

$$\bar{U} = (-\bar{x}_1; 1)(-\bar{\mathbf{x}}_1; 1)^T,$$

then, $\bar{U} \succeq \mathbf{0}$ and $\bar{U} \bullet \bar{X} = 0$ so that $\bar{U}$ is an optimal slack matrix for the dual and its rank is 1.

How to Select w's

We only need to consider choosing $\hat{w}$'s:

$$\sum_{k=1}^{3} \hat{w}_{k1}\mathbf{a}_k = \bar{\mathbf{x}}_1 \qquad \text{or} \qquad \sum_{k=1}^{3} \hat{w}_{k1}(\mathbf{a}_k - \bar{\mathbf{x}}_1) = \mathbf{0}$$

$$\hat{w}_{11} + \hat{w}_{21} + \hat{w}_{31} = 1. \qquad \hat{w}_{11} + \hat{w}_{21} + \hat{w}_{31} = 1.$$

This system always has a solution if $\mathbf{a}_k$ is not co-linear.

Then, select the rest

$$\begin{pmatrix} w_1 + w_3 & w_3 \\ w_3 & w_2 + w_3 \end{pmatrix} = \bar{\mathbf{x}}_1\bar{\mathbf{x}}_1^T - \sum_{k=1}^{3} \hat{w}_{k1}\mathbf{a}_k\mathbf{a}_k^T$$

Other Conditions?

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

$$\sum_{k=1}^{3} \hat{w}_{k1}(\mathbf{a}_k - \bar{\mathbf{x}}_1) = \mathbf{0}$$
$$\hat{w}_{11} + \hat{w}_{21} + \hat{w}_{31} = 1$$

may still have a solution $w_\bullet$?

Physical interpretation: $\hat{w}_{kj}$ is a stress/force on the edge and all stresses are balanced or at an equilibrium state. The objective represents the potential of the system.

Localize All Localizable Points

Theorem 4 *(So and Y 2005) If a problem (graph) contains a subproblem (subgraph) that is universally-localizable, then the submatrix solution corresponding to the subproblem in the SDP solution has rank 2. That is, the SDP relaxation computes a solution that localize all possibly localizable unknown sensor points.*

The proof is similar to the proof of Theorem 3 by removing the notes that is not localizable.

Implication: Diagonals of "co-variance" matrix

$$\bar{Y} - \bar{X}^T \bar{X},$$

$\bar{Y}_{jj} - \|\bar{x}_j\|^2$, can be used as a measure to see whether jth sensor's estimated position is reliable or not.

Uncertainty Analysis and Confidence Measure

Alternatively, each x_j's can be viewed as uncertain points from the incomplete/uncertain distance measures. Then the solution to the SDP problem provides the first and second moment estimation (Bertsimas and Y 1998).

Generally, $\bar{x}_j$ is a point estimate of x_j and $\bar{Y}_{ij}$ is a point estimate $x_i^T x_j$.

Consequently,

$$\bar{Y}_{jj} - \|\bar{x}_j\|^2,$$

which is the individual variance estimation of sensor j, gives an interval estimation for its true position (Biswas and Y 2004).

SDP Relaxation with Noise Data

When the distance measurements have noise, one can minimize the total error as

$$
\begin{array}{lll}
\min & \sum_{(i,j)\in N_x} (s_{ij})^2 + \sum_{(k,j)\in N_a} (s^a_{kj})^2 & \\
& Z_{1:2,1:2} & = I \\
& (\mathbf{0}; \mathbf{e}_i - \mathbf{e}_j)(\mathbf{0}; \mathbf{e}_i - \mathbf{e}_j)^T \bullet Z + s_{ij} & = d^2_{ij}, \ \forall\, (i,j) \in N_x,\ i < j, \\
& (\mathbf{a}_k; -\mathbf{e}_j)(\mathbf{a}_k; -\mathbf{e}_j)^T \bullet Z + s^a_{kj} & = d^2_{kj}, \ \forall\, (k,j) \in N_a, \\
& Z & \succeq \mathbf{0}.
\end{array}
$$