Semidefinite Programming for Euclidean Distance Geometric Optimization

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

http//www.stanford.edu/ yyye

Outline

- Ad Hoc Wireless Sensor Network Localization and SDP Approximation (with Biswas 2003, www.stanford.edu/ yyye/adhocn2.pdf)
- Related Problems: Access Point Placement, Euclidean Ball Packing, Metric Distance Embedding, etc.
- Radii of High-Dimension Points and SDP Approximation (with Zhang 2003, www.stanford.edu/ yyye/radii2.pdf)

1. Ad Hoc Wireless Sensor Network Localization

- **Input** m known points $a_k \in \mathbf{R}^2$, $k = 1, ..., m$, and n unknown points $x_j \in \mathbf{R}^2$, $j = 1, ..., n$. For each pair of two points, we have a Euclidean distance upper bound $\bar{d}_{kj}$ and lower bound $\underline{d}_{kj}$ between a_k and x_j, or upper bound $\bar{d}_{ij}$ and lower bound $\underline{d}_{ij}$ between x_i and x_j.
- **Output** Position estimation for all unknown points.
- **Objective** Robust and accurate.

Related Work

- A great deal of research has been done on the topic of position estimation in ad-hoc networks, see Hightower and Boriello (2001) and Ganesan, Krishnamachari, Woo, Culler, Estrin, and Wicker (2002).
- Beacon grid: e.g., Bulusu and Heidemann (2000) and Howard, Mataric, and Sukhatme (2001).
- Distance measurement: e.g., Doherty, Ghaoui, and Pister (2001), Niculescu and Nath (2001), Savarese, Rabaey, and Langendoen (2002), Savvides, Han, and Srivastava (2001), Savvides, Park, and Srivastava (2002), Shang, Ruml, Zhang and Fromherz (2003).

Quadratic Inequalities

Two points x_1 and x_2 are within radio range r of each other, the proximity constraint can be represented as a convex second order cone inequality of the form

$$\|x_1 - x_2\| \leq r$$

Two points x_1 and x_2 are beond radio range r of each other, the “bounding away” constraint can be represented as a quadratic inequality of the form

$$\|x_1 - x_2\| \geq r$$

Unfortunately, the latter is not convex.

Doherty et al. use only the former in their convex optimization model, the others solve them as non-convex feasibility or optimization problems.

Quadratic Models

$$\begin{array}{ll} \text{minimize} & \alpha \\ \text{subject to} & (\underline{d}_{ij})^2 - \alpha \le \|x_i - x_j\|^2 \le (\bar{d}_{ij})^2 + \alpha, \ \forall i \ne j, \\ & (\underline{d}_{kj})^2 - \alpha \le \|a_k - x_j\|^2 \le (\bar{d}_{kj})^2 + \alpha, \ \forall k, j, \end{array}$$

or

$$\begin{array}{ll} \text{minimize} & \sum_{i,j:i\ne j} \alpha_{ij} + \sum_{k,j} \alpha_{kj} \\ \text{subject to} & (\underline{d}_{ij})^2 - \alpha_{ij} \le \|x_i - x_j\|^2 \le (\bar{d}_{ij})^2 + \alpha_{ij}, \ \forall i \ne j, \\ & (\underline{d}_{kj})^2 - \alpha_{kj} \le \|a_k - x_j\|^2 \le (\bar{d}_{kj})^2 + \alpha_{kj}, \ \forall k, j. \end{array}$$

Models continued

minimize α

subject to $(1-\alpha)(\underline{d}_{ij})^2 \le \|x_i - x_j\|^2 \le (1+\alpha)(\bar{d}_{ij})^2,\ \forall i \ne j,$

$(1-\alpha)(\underline{d}_{kj})^2 \le \|a_k - x_j\|^2 \le (1+\alpha)(\bar{d}_{kj})^2,\ \forall k, j,$

or

minimize $\sum_{i,j:i\ne j} \alpha_{ij} + \sum_{k,j} \alpha_{kj}$

subject to $(1-\alpha_{ij})(\underline{d}_{ij})^2 \le \|x_i - x_j\|^2 \le (1+\alpha_{ij})(\bar{d}_{ij})^2,\ \forall i \ne j,$

$(1-\alpha_{kj})(\underline{d}_{kj})^2 \le \|a_k - x_j\|^2 \le (1+\alpha_{kj})(\bar{d}_{kj})^2,\ \forall k, j.$

Models continued

If distance measures $\bar{d}_{ij} = \underline{d}_{ij} = \hat{d}_{ij}$ for $i, j \in N_1$ and $\bar{d}_{kj} = \underline{d}_{kj} = \hat{d}_{kj}$ for $k, j \in N_2$, and the rest of them have only a lower bound R, then sum problem can be formulated with mixed equalities and inequalities:

$$
\begin{array}{ll}
\text{minimize} & \sum_{i,j \in N_1,\; i<j} \alpha_{ij} + \sum_{k,j \in N_2} \alpha_{kj} \\
\text{subject to} & \|x_i - x_j\|^2 = (\hat{d}_{ij})^2 + \alpha_{ij}, \text{ for } i, j \in N_1,\; i < j, \\
& \|a_k - x_j\|^2 = (\hat{d}_{kj})^2 + \alpha_{kj}, \text{ for } k, j \in N_2, \\
& \|x_i - x_j\|^2 \geq R^2, \text{ for the rest } i < j, \\
& \|a_k - x_j\|^2 \geq R^2, \text{ for the rest } k, j, \\
& \alpha_{i,j} \geq 0,\; \alpha_{k,j} \geq 0.
\end{array}
$$

Matrix Representation

Let $X = [x_1\ x_2\ \ldots\ x_n]$ be the $2 \times n$ matrix that needs to be determined. Then

$\|x_i - x_j\|^2 = e_{ij}^T X^T X e_{ij}$ and $\|a_i - x_j\|^2 = (a_i; e_j)^T [I\ \ X]^T [I\ \ X](a_i; e_j)$,

where e_{ij} is the vector with 1 at the ith position, -1 at the jth position and zero everywhere else; and e_j is the vector of all zero except 1 at the jth position.

$$
\begin{array}{ll}
\text{min} & \alpha \\
\text{s.t.} & (\underline{d}_{ij})^2 - \alpha \le e_{ij}^T Y e_{ij} \le (\bar{d}_{ij})^2 + \alpha, \\
 & (\underline{d}_{kj})^2 - \alpha \le (a_k; e_j)^T \begin{pmatrix} I & X \\ X^T & Y \end{pmatrix} (a_k; e_j) \le (\bar{d}_{kj})^2 + \alpha, \\
 & Y = X^T X.
\end{array}
$$

SDP Relaxation

$$
\begin{aligned}
\text{min} \quad & \alpha \\
\text{s.t.} \quad & (\underline{d}_{ij})^2 - \alpha \le e_{ij}^T Y e_{ij} \le (\bar{d}_{ij})^2 + \alpha, \\
& (\underline{d}_{kj})^2 - \alpha \le (a_k; e_j)^T \begin{pmatrix} I & X \\ X^T & Y \end{pmatrix} (a_k; e_j) \le (\bar{d}_{kj})^2 + \alpha, \\
& Y \succeq X^T X.
\end{aligned}
$$

The last matrix inequality is equivalent to (Boyd et al. 1994)

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

SDP Form

$$
\begin{array}{ll}
\text{minimize} & \alpha \\
\text{subject to} & (1;0;\mathbf{0})^T Z (1;0;\mathbf{0}) = 1 \\
& (0;1;\mathbf{0})^T Z (0;1;\mathbf{0}) = 1 \\
& (1;1;\mathbf{0})^T Z (1;1;\mathbf{0}) = 2 \\
& (\underline{d}_{ij})^2 - \alpha \le (\mathbf{0}; e_{ij})^T Z (\mathbf{0}; e_{ij}) \le (\bar{d}_{ij})^2 + \alpha, \ \forall i \ne j, \\
& (\underline{d}_{kj})^2 - \alpha \le (a_k; e_j)^T Z (a_k; e_j) \le (\bar{d}_{kj})^2 + \alpha, \ \forall k, j, \\
& Z \succeq 0.
\end{array}
$$

Here $Z \in \mathbf{R}^{(n+2)\times(n+2)}$ and it has $2n + n(n+1)/2$ unknowns.

Deterministic Analysis

If there are $2n + n(n+1)/2$ point pairs each of which has accurate distance measures and other distance bounds are feasible. Then, we have the minimal value of $\alpha = 0$ in the relaxation. Moreover, if the relaxation has a unique minimal solution Z^*, we must have $Y^* = (X^*)^T X^*$ in the minimal solution Z^* and the SDP relaxation solves the original problem exactly.

A point can be determined by its distances to three known points that are not on a same line.

Probabilistic or Error Analysis

Alternatively, each x_j can be viewed a random point $\tilde{x}_j$ since the distance measures contain random errors. Then the solution to the SDP problem provides the first and second moment information on $\tilde{x}_j$, $j = 1, ..., n$ (Bertsimas and Ye 1998).

Generally, we have

$$\mathsf{E}[\tilde{x}_j] \sim \bar{x}_j, \quad j = 1, ..., n$$

and

$$\mathsf{E}[\tilde{x}_i^T \tilde{x}_j] \sim \bar{Y}_{ij}, \quad i, j = 1, ..., n.$$

Here

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

is the optimal solution of the SDP problem.

Thus,

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

represents the co-variance matrix of $\tilde{x}_j$, $j = 1, ..., n$.

Observable Measures

In certain probabilistic models, $\bar{x}_j$ is a point estimate of the mean of $\tilde{x}_j$, and $\bar{Y} - \bar{X}^T \bar{X}$ estimates the covariance of $\tilde{X}$.

Therefore,

$$\text{Trace}(\bar{Y} - \bar{X}^T \bar{X}),$$

the trace of the co-variance matrix, measures the quality of sample data d_{ij} and d_{kj}.

In particular,

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

which is the variance of $\|\tilde{x}_j\|$, helps us to detect possible outlier or defect sensors.

Simulation and Computation Results

Simulations were performed on a network of 50 sensors or nodes randomly placed in a square region of size $r \times r$ where $r = 1$. The distances between the nodes was calculated. If the distance between 2 notes was less than a given $radiorange$ between $[0, 1]$, a random error was added to it

$$\hat{d}_{ij} = d_{ij} \cdot (1 + (2 * rand - 1) * noisyfactor),$$

where $noisyfactor$ was a given number between $[0, 1]$, and then both upper and lower bound constraints were applied for that distance in the SDP model.

If the distance was beyond the given $radiorange$, only the lower bound constraint, $\geq 1.001 * radiorange$, was applied.

The average estimation error is defined by

$$\frac{1}{n} \cdot \sum_{j=1}^{n} = \|\bar{x}_j - a_j\|,$$

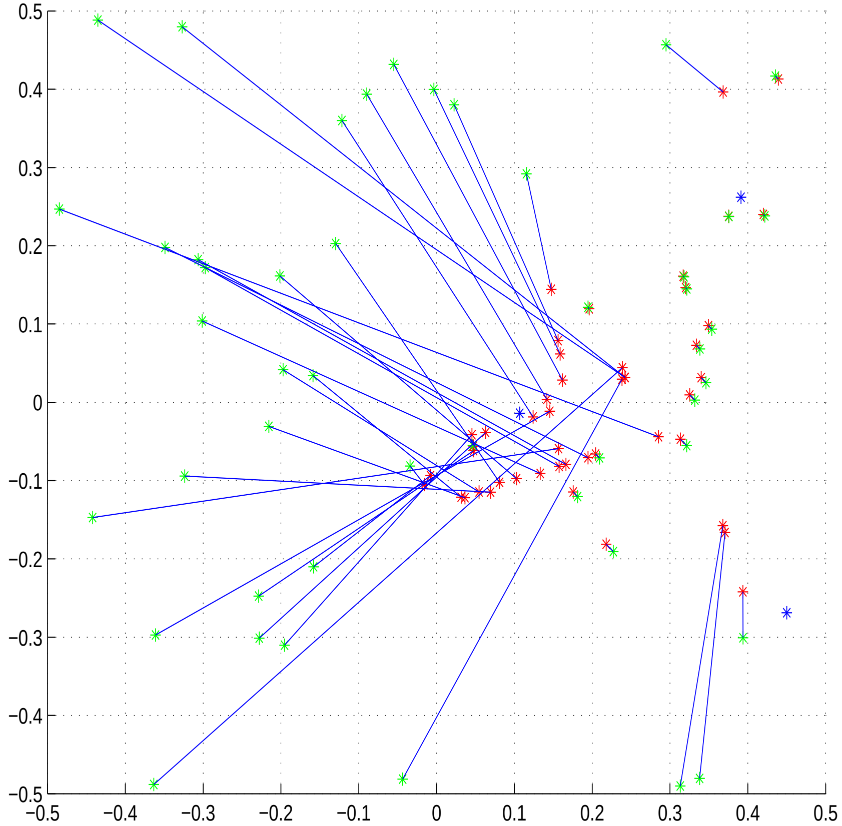
where $\bar{x}_j$ comes from the SDP solution and a_j is the true position of the jth node.

The trace of $\bar{Y} - \bar{X}^T\bar{X}$ is called the total-variance, and $\bar{Y}_{jj} - \|\bar{x}_j\|^2$ the jth individual trace.

Connectivity indicates how many of the nodes, on average, are within the radio range of a node.

SDP solvers used were DeDuMi (Sturm) and DSDP4.5 (Benson).

Figure 1: Position estimations with 3 anchors, noisy factor=0, and radio range=0.2 (error:0.28, connec:5.8, trace:2.4) and 0.25 (error:0.023, connect:7.8, trace:0.16)



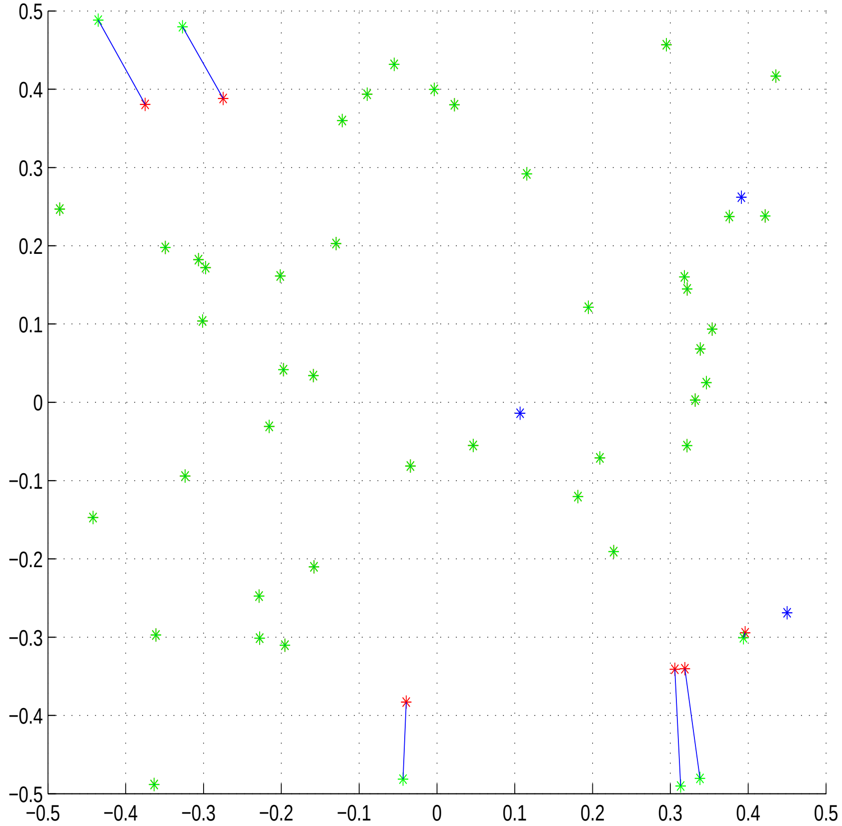


Figure 2: Correlation of square root of individual trace and error for each sensor with 3 anchors and radio range=0.25

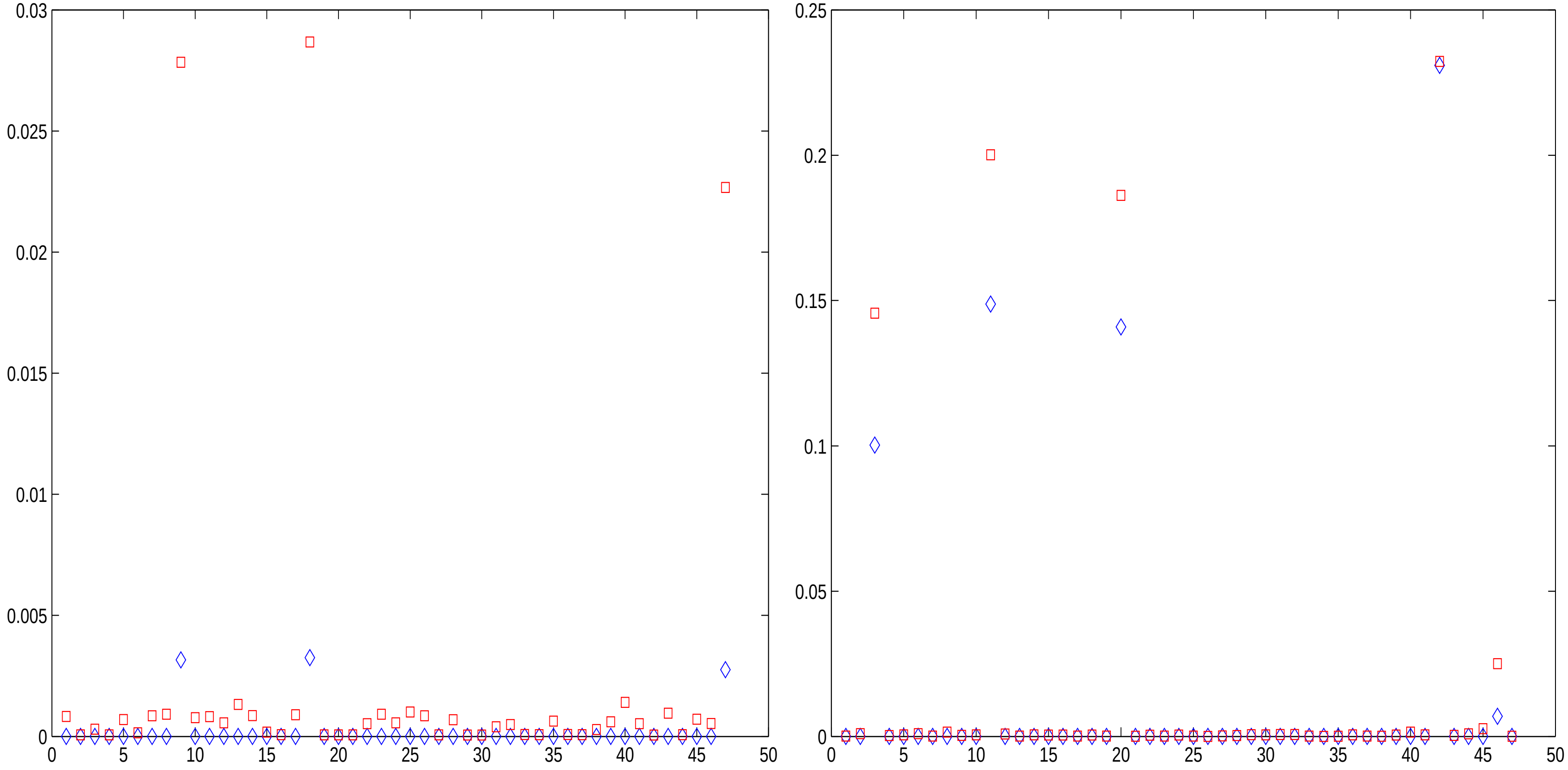


Figure 3: Position estimations with 3 anchors, noisy factor=0, and radio range=0.30 (error:0.0014, connec:10.5, trace:0.03) and 0.35 (error:0.0014, connect:13.2, trace:0.04)

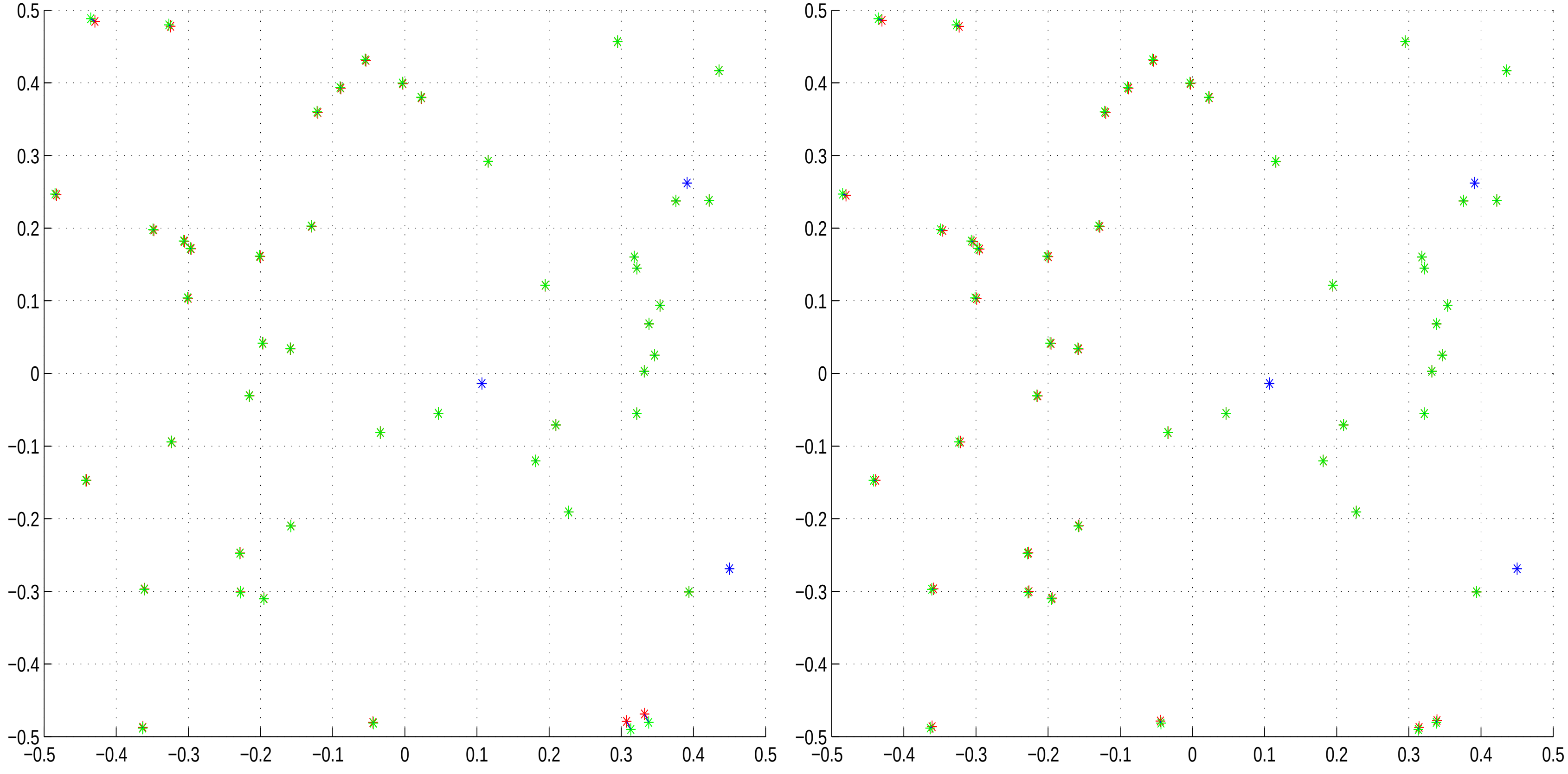


Figure 4: Position estimations with 7 anchors, noisy factor=0, and radio range==0.2 (error:0.054, connec:5.8, trace:0.54) and 0.25 (error:0.012, connect:7.8, trace:0.14)

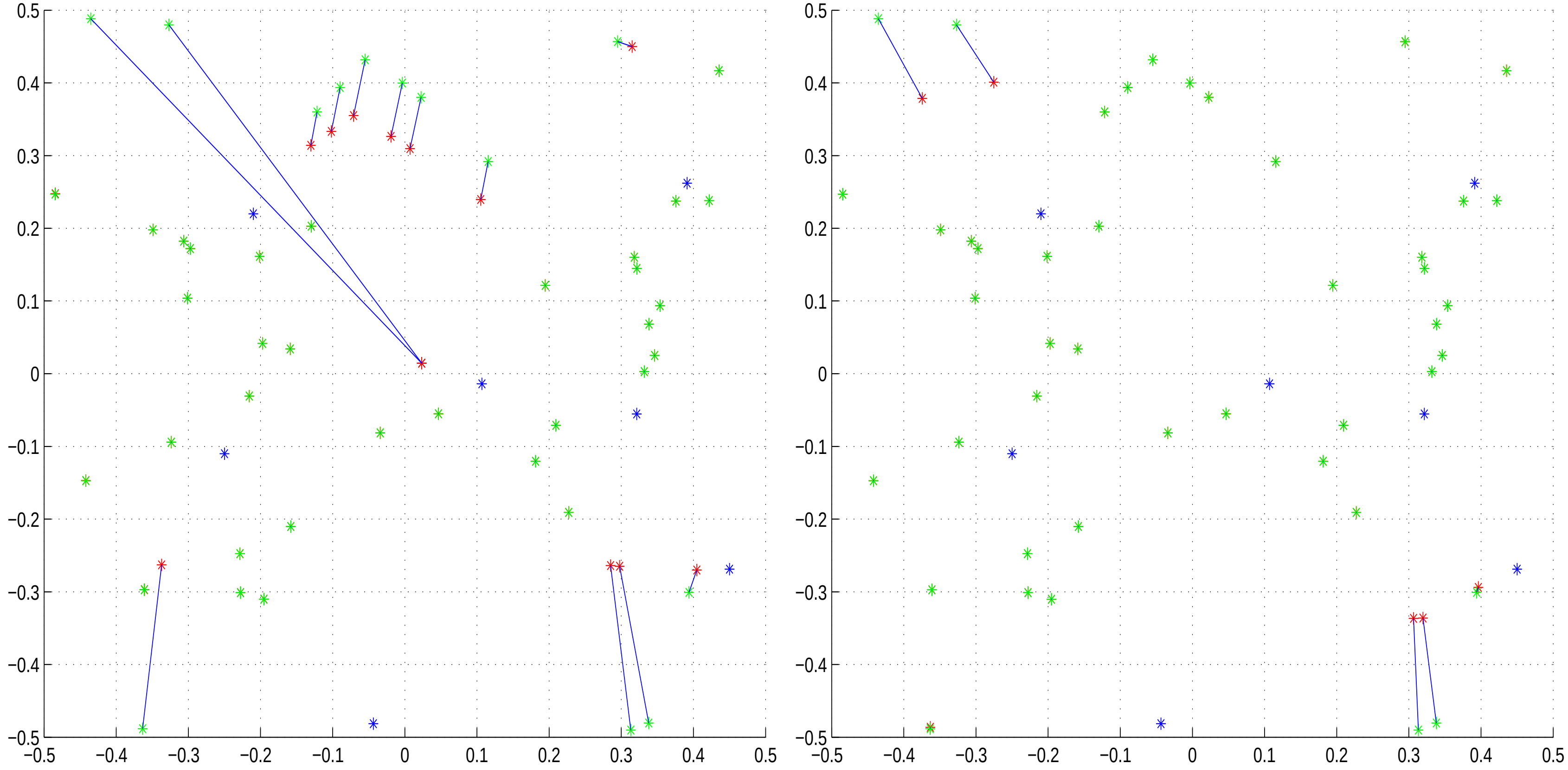
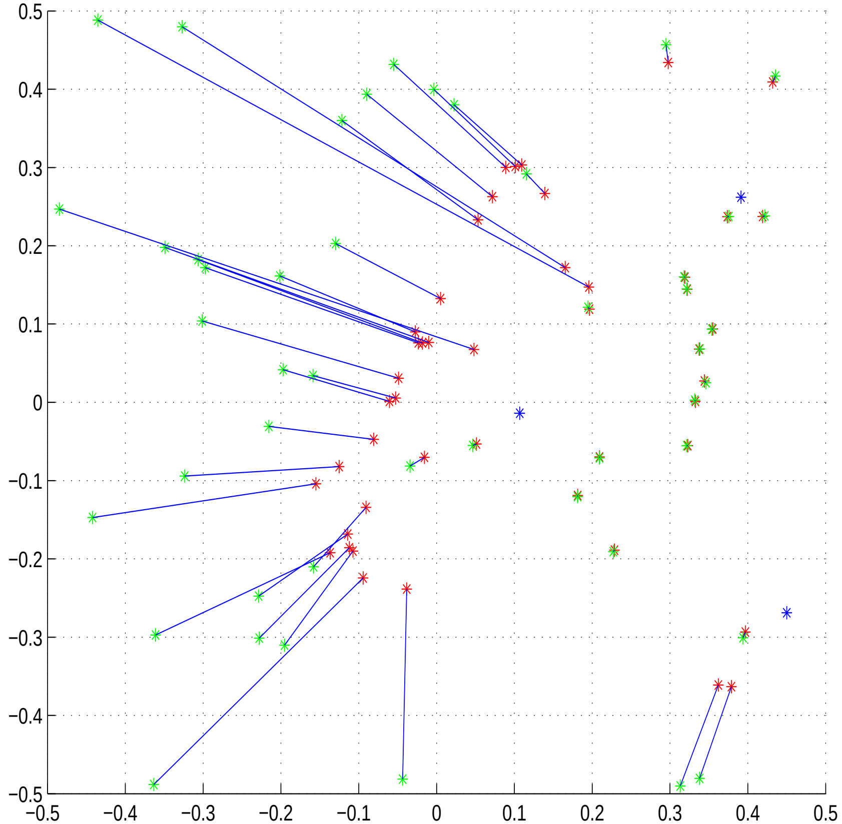


Figure 5: Position estimations with radio range 0.3, noisy factor=0.01, and number of anchors=3 (error:0.083, trace: 3.7) and 6 (error: 0.015, trace:0.25)



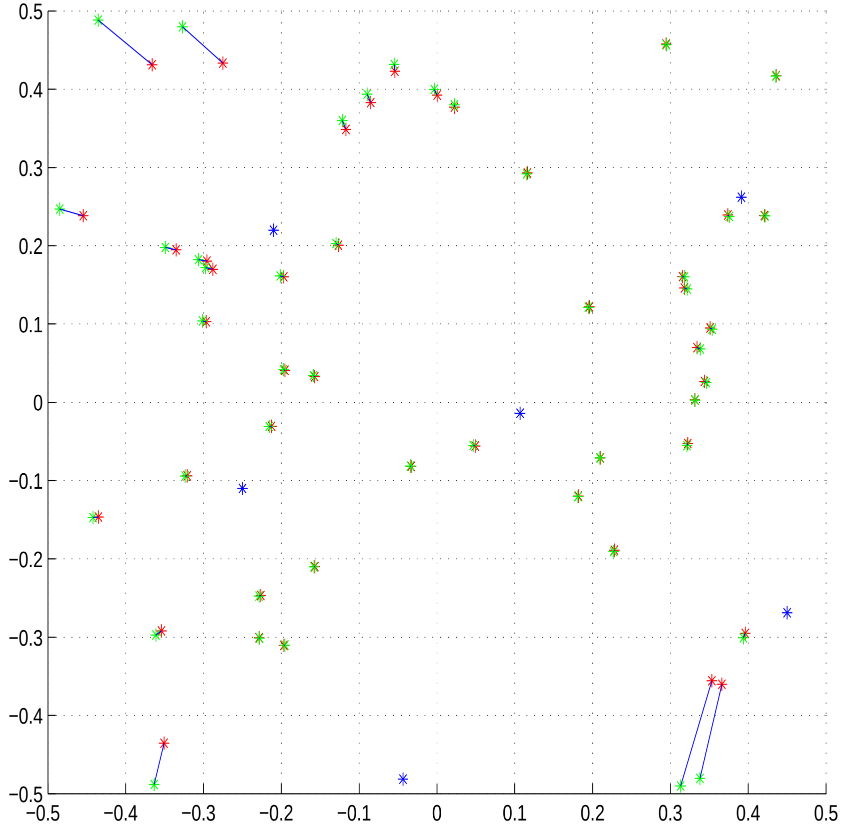
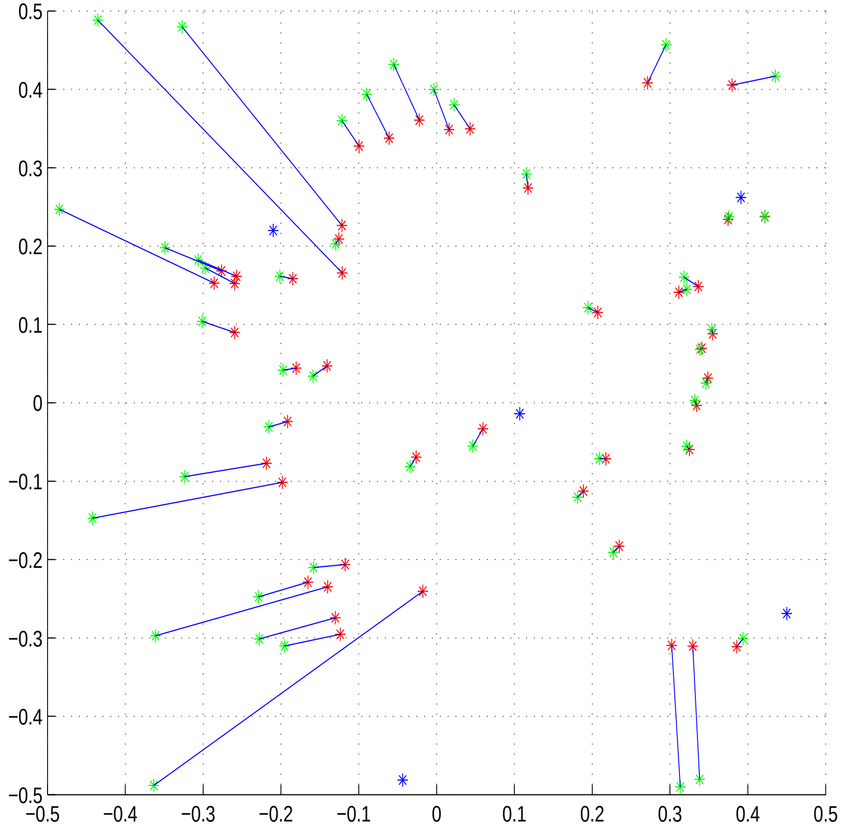


Figure 6: Position estimations with 5 anchors, radio range=0.3, and noisy factor= 0.05 (error: 0.05, trace 0.71) and 0.10 (error:0.07, trace: 0.96)



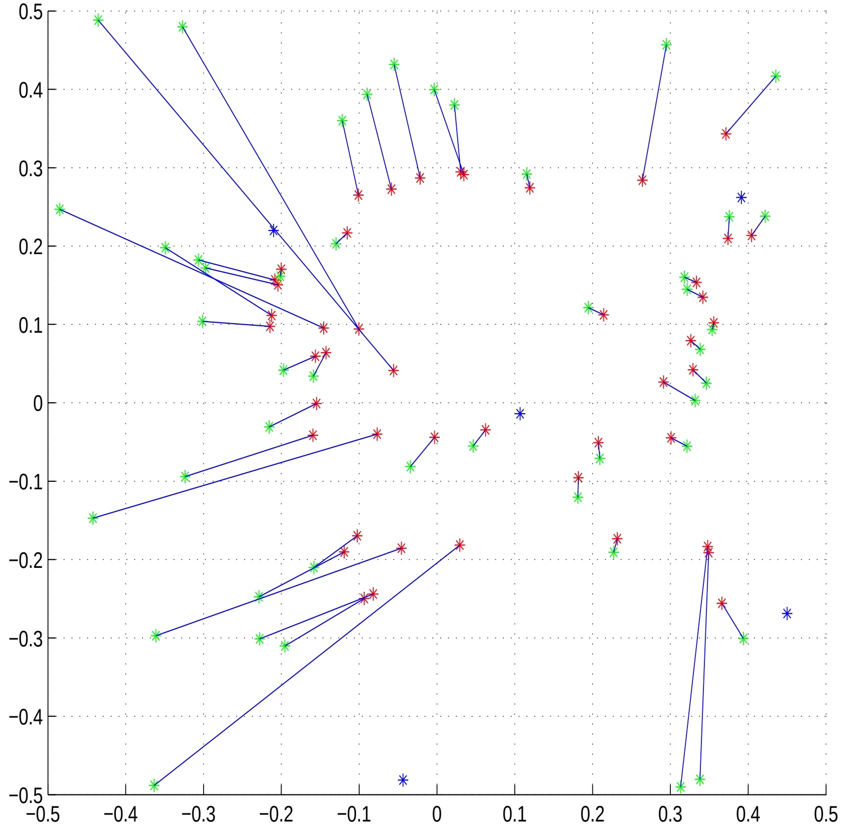
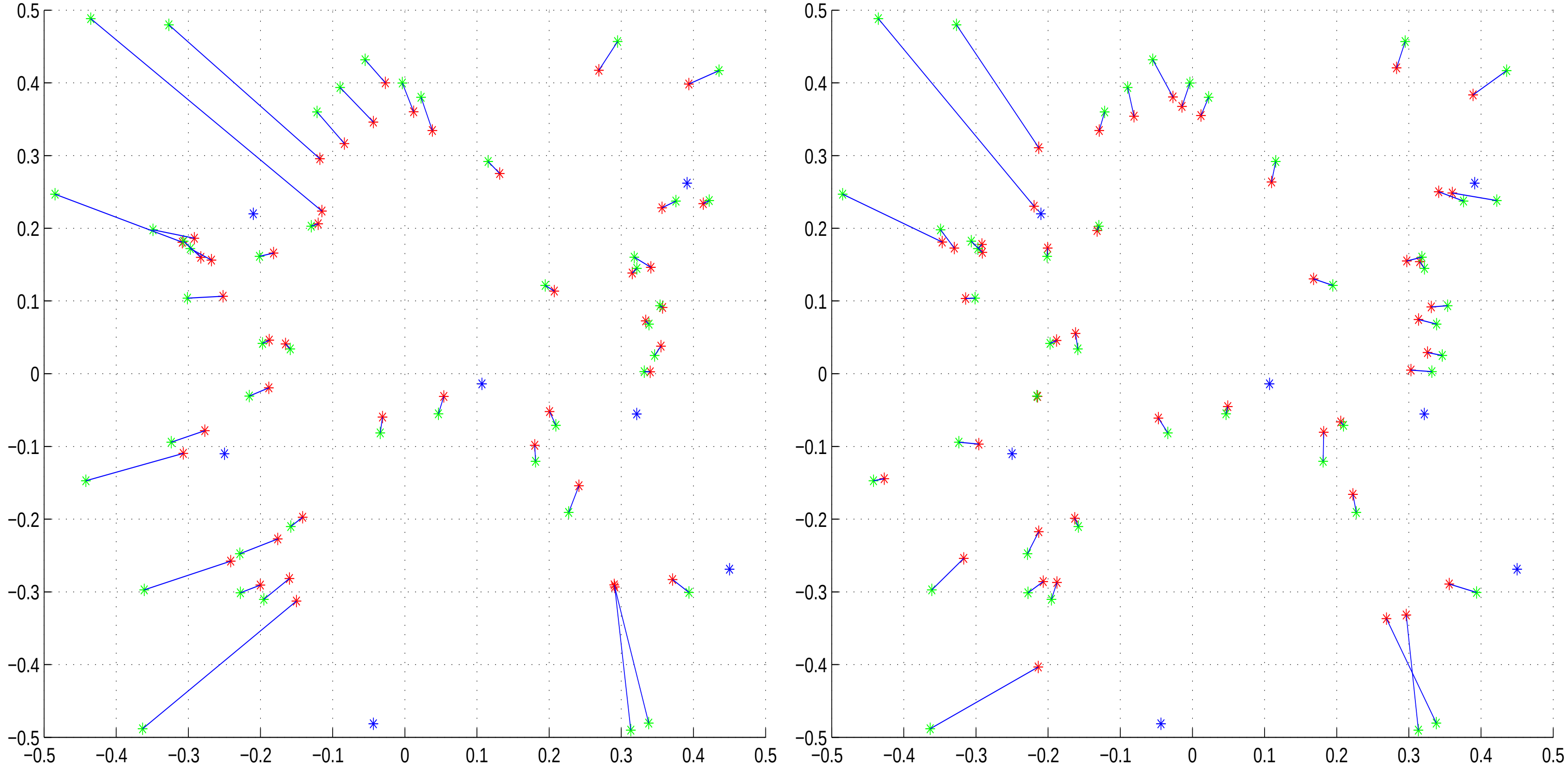


Figure 7: Position estimations with 7 anchors, noisy-factor=0.1, and radio range= 0.30 (error: 0.081, trace: 0.78) and 0.35 (error: 0.065, trace: 0.76)



Work in Progress: Active constraint generation

In the SDP problem, the dimension of the matrix is $n + 2$ and the number of constraints is in the order of $O(n + m)^2$. Typically, each iteration of interior-point algorithm SDP solvers need to factorize and solve a dense matrix linear system whose dimension is the number of constraints. The current interior-point algorithm SDP solvers can handle such a system whose dimension is about $10,000$.

Fortunately, many of those ”bounding away” constraints, i.e., the constraints between two remote nodes, are inactive or redundant at optimal solutions. Therefore, an iterative solution method can be developed.

Work in Progress: Distributed computation

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

can be decomposed into K principle blocks

$$\begin{pmatrix} I & X_1 & X_2 & \ldots & X_K \\ X_1^T & Y_{11} & Y_{12} & \ldots & Y_{1K} \\ X_2^T & Y_{21} & Y_{22} & \ldots & Y_{2K} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ X_K^T & Y_{K1} & Y_{K2} & \ldots & Y_{KK}. \end{pmatrix}$$

The kth principle block matrix is

$$\begin{pmatrix} I & X_k \\ X_k^T & Y_{kk} \end{pmatrix}$$

Then, we can solve the kth block problem, assuming all others are fixed, in a distributed fashion for $k = 1, ..., K$. That is, given other block's solutions, each of these problems can be solved locally and separately. Thereafter, we have new X_k and Y_{kk} for $k = 1, ..., K$, and Y_{ki} can be also updated to $X_i^T X_k$. These new updates are then communicated among the blocks.

2.1 Related Problems: Access Point Placement

- Due to the energy and resource constraints, nodes in a sensor network usually have very short communication ranges.
- Hierarchical sensor networks: low-energy, low-bandwidth communication protocol sensor and upper layer access point (APs) sensor that may have multiple radio capabilities.
- an AP sensor is much more expensive (tens or hundreds times more) than a sensor node, which makes a large number of APs undesirable and their placement crucial.

Access Point Placement Formulation

Let a_k be the position of sensor node k, x_j be the unknown position of AP j. Let $K(j)$ be the set of sensors served by AP j, then we have

$$\begin{array}{ll} \text{minimize} & \alpha \\ \text{subject to} & \|a_k - x_j\|^2 \le \alpha,\ \forall\, k \in K(j), j, \\ & \|x_i - x_j\|^2 \le \alpha,\ \forall i \ne j, \end{array}$$

This model will have APs placed at the positions that the maximum distance between any two APs (second constraint) and between any AP to its client sensors (first constraint) is minimized.

This problem is a convex second-order cone program.

2.2 Related Problems: Euclidean Ball Packing

The Euclidean ball packing problem is an old mathematical geometry problem with plenty modern applications in Bio-X and Chemical Structures.

Pack n balls (the jth ball has radius r_j) in a box with width and length equal $2R$ and like to minimize the height of the box:

$$
\begin{array}{ll}
\text{minimize} & \alpha \\
\text{subject to} & \|x_i - x_j\|^2 \geq (r_i + r_j)^2, \ \forall i \neq j, \\
& \|x_i - x_j\|^2 = (r_i + r_j)^2, \ \text{for some } i \neq j, \\
& -R + r_j \leq x_j(1) \leq R - r_j, \ \forall j, \\
& -R + r_j \leq x_j(2) \leq R - r_j, \ \forall j, \\
& r_j \leq x_j(3) \leq \alpha - r_j, \ \forall j,
\end{array}
$$

2.3 Related Problems: Metric Distance Embedding

Given matric distances d_{ij} for all $i \neq j$, find $x_j \in R^k$ such that

$$\begin{array}{ll} \text{minimize} & \alpha \\ \text{subject to} & (d_{ij})^2 \leq \|x_i - x_j\|^2 \leq (1+\alpha)(d_{ij})^2, \ \forall i \neq j. \end{array}$$

Want both α and k as small as possible.

3. Approximate the Minimum Radii of Projected Point Sets

- **Input.** A set P of $2n$ symmetric points in Euclidean space R^d: If $p \in P$ then $-p \in P$.
- **Objective.** To minimize the outer k-radius of P

$$R_k(P) = \min_{F \in F^k} \max_{p \in P} d(p, F),$$

where F^k is the collection of all k-dimensional subspaces of R^d, and $d(p, F)$ is the length of the projection of p onto F.

- **Mathematical formulation.** The square of $R_k(P)$ can be defined as the minimum of, over all sets of k orthogonal unit vectors $\{x_1, x_2, \cdots, x_k\}$,

$$\max_{p \in P} \sum_{i=1}^{k} (p^T x_i)^2.$$

Figure 8: Radius of points

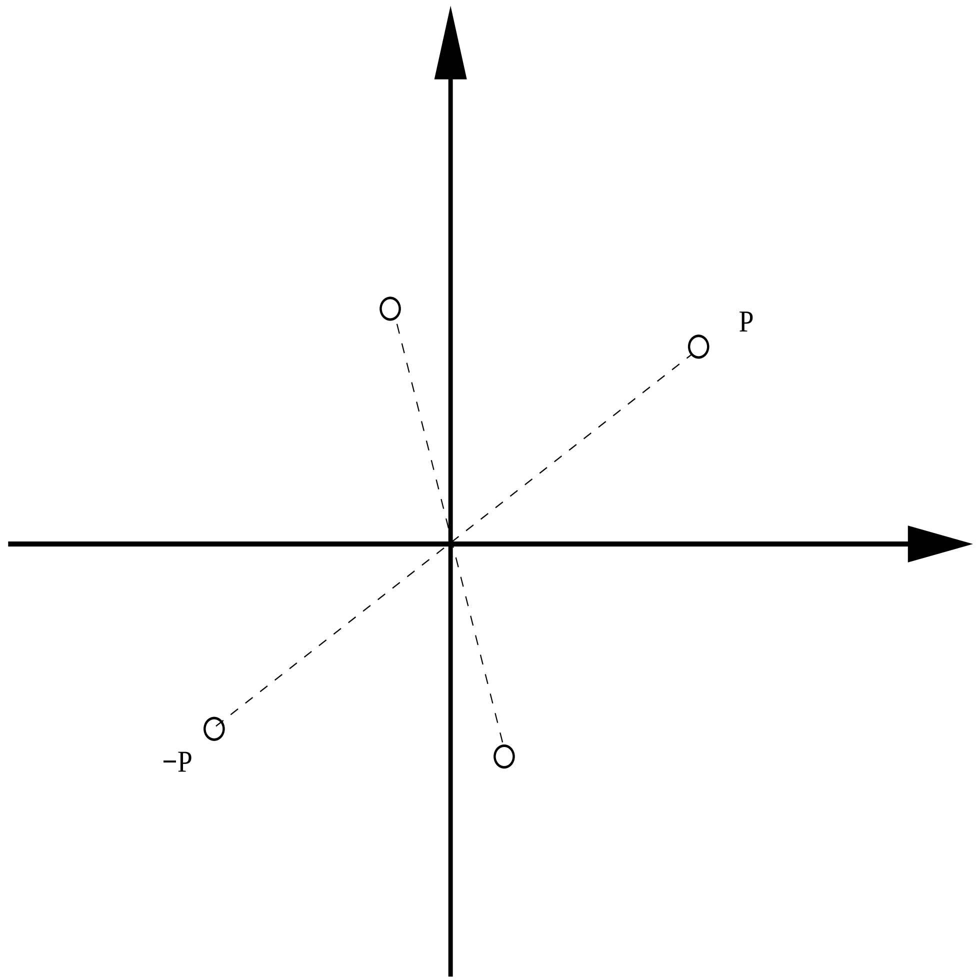
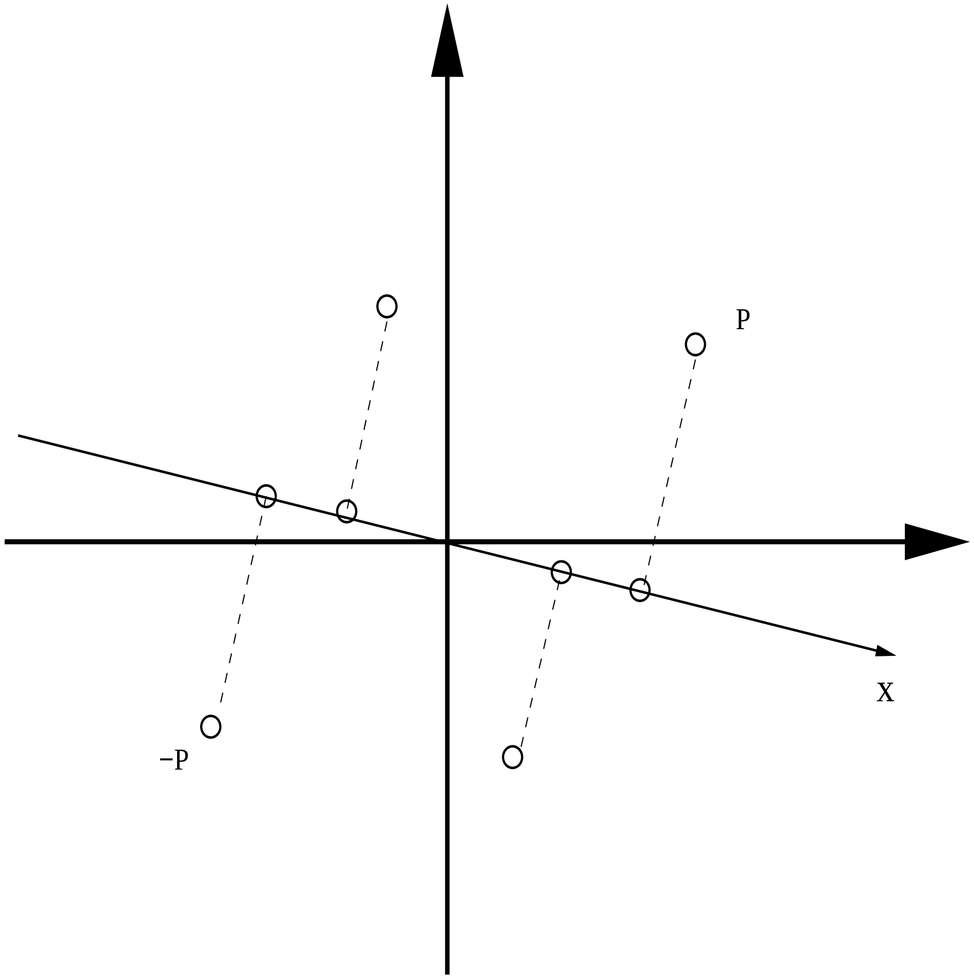


Figure 9: Radius of projected points on one-dimensional x



Previous Results

- Fundamental problem in computational geometry and has applications in data mining, statistics, clustering, etc. (Gritzmann and Klee 1993, 1994).
- When d (the dimension) is a constant: The problem is polynomial time solvable (Faigle et al 1996).
- When $d - k$ is a constant, the problem can be approximated by $(1 + \epsilon)$ (Badoiu et al 2002, Har-Peled and Varadarajan 2002).
- For $k = 1$, there is a randomized $O(\log n)$ algorithm for $R_k(P)^2$ (Implied from Nemirovskii et al. 1999).
- Nothing is known when $d - k$ varies. A few hardness results shows that it is NP-hard to approximate the problem (Briden 2000, 2002).

Varadarajan/Venkatesh/Zhang Results (FOCS2002)

- There is a poly-time algorithm that approximates $R_k(P)^2$ within a factor of $O(\log n \cdot \log d)$ for any $1 \leq k \leq d$.
- Conjecture: the problem is $O(\log n)$ approximatable.

Our Result

Their conjecture is true: there is a poly-time algorithm that approximates $R_k(P)^2$ within a factor of $O(\log n)$ for any $1 \leq k \leq d$.

- Using SDP relaxation
- Using a deterministic subspace partition based on the eigenvalue decomposition
- Using a randomized rank reduction for each subspace

Quadratic Representation

$$R_k(P)^2 = \text{Minimize} \quad \alpha$$

Subject to

$$\sum_{i=1}^{k} (p^T x_i)^2 \le \alpha, \ \forall p \in P,$$

$$\|x_i\|^2 = 1, \ i = 1, ..., k,$$

$$(x_i)^T x_j = 0, \ \forall i \neq j.$$

Classical SDP Relaxation for QCQP

An SDP of matrix dimension $k \cdot d$ and $n + k^2$ constraints.

Leaner SDP Relaxation

Consider the matrix $X = (x_1 x_1^T + x_2 x_2^T + \cdots + x_k x_k^T)$, we get a leaner SDP relaxation

$$
\begin{aligned}
\alpha_k^* = \text{Minimize} \quad & \alpha \\
\text{Subject to} \quad & pp^T \bullet X \le \alpha, \ \forall p \in P, \\
& I \bullet X = k, \\
& I - X \succeq 0, \\
& X \succeq 0.
\end{aligned}
$$

Eigenvalue Decomposition

Let X^* be an SDP optimizer with rank r. Then, considering λ_i and x_i being the eigenvalues and eigenvectors of X^*, we can compute, in "polynomial time", a set of non-negative reals $\lambda_1, \cdots, \lambda_d$ and a set of orthogonal unit vectors $x_1, \cdots, x_r$ in R^d such that

- $\sum_{i=1}^r \lambda_i = k$
- $\max_i \lambda_i \le 1$
- $X^* = \sum_{i=1}^r \lambda_i \cdot x_i x_i^T$.

Note that $r \ge k$. (Why?)

Eigenvalue and Subspace Partition

Partition λ_is into k sets, $I_1, ..., I_k$, such that

$$\sum_{i \in I_j} \lambda_i \geq \frac{1}{2}, \ \forall j = 1, ..., k.$$

Can do this quickly. How?

Eigenvectors in each I_j form a subspace which is further reduced to one basis using a random combination.

Discussion Questions

- How to round the SDP matrix into vector solutions?
- What are the duals of the SDP relaxations?
- How to interpret dual variables?
- Are there tighter SDP relaxations?
- How to solve SDP relaxations by exploiting the problem structure?