On SDP and ESDP Relaxation of Sensor Network Localization

Paul Tseng
Mathematics, University of Washington
Seattle

IASI, Roma
September 16, 2008
(joint work with Ting Kei Pong)

Talk Outline

- Sensor network localization
- SDP, ESDP relaxations: formulation and properties

Talk Outline

- Sensor network localization
- SDP, ESDP relaxations: formulation and properties
- A robust version of ESDP to handle noises
- BCGD-barrier method

Talk Outline

- Sensor network localization
- SDP, ESDP relaxations: formulation and properties
- A robust version of ESDP to handle noises
- BCGD-barrier method
- Numerical simulations
- Conclusion & Ongoing work

Sensor Network Localization

Basic Problem:

- n pts in $\Re^2$.
- Know last $n - m$ pts ('anchors') $x_{m+1}, ..., x_n$ and Eucl. dist. estimate for pairs of 'neighboring' pts

$$d_{ij} \geq 0 \quad \forall (i,j) \in \mathcal{A}$$

with $\mathcal{A} \subseteq \{(i,j) : 1 \leq i < j \leq n\}$.

- Estimate first m pts ('sensors').

Sensor Network Localization

Basic Problem:

- n pts in $\Re^2$.

- Know last $n - m$ pts ('anchors') $x_{m+1}, ..., x_n$ and Eucl. dist. estimate for pairs of 'neighboring' pts

$$d_{ij} \geq 0 \quad \forall (i,j) \in \mathcal{A}$$

with $\mathcal{A} \subseteq \{(i,j) : 1 \leq i < j \leq n\}$.

- Estimate first m pts ('sensors').

History? Graph realization, position estimation in wireless sensor network, ...

Optimization Problem Formulation

$$v_{\rm opt} := \min_{x_1,\ldots,x_m} \sum_{(i,j)\in\mathcal{A}} \left| \|x_i - x_j\|^2 - d_{ij}^2 \right|$$

Optimization Problem Formulation

$$v_{\text{opt}} := \min_{x_1,\ldots,x_m} \sum_{(i,j)\in\mathcal{A}} \left| \|x_i - x_j\|^2 - d_{ij}^2 \right|$$

- Objective function is nonconvex. m can be large ($m > 1000$).
- Problem is NP-hard (reduction from PARTITION).
- Use a convex (SDP, SOCP) relaxation. Low soln accuracy OK. Distributed methods.

SDP Relaxation

Let $X := [x_1 \ \cdots \ x_m]$.

$$Z = [\, X \quad I \,]^T [\, X \quad I \,] \quad \Longleftrightarrow \quad Z = \begin{bmatrix} Y & X^T \\ X & I \end{bmatrix} \succeq 0, \ \mathrm{rank} Z = 2$$

SDP Relaxation

Let $X := [x_1 \ \cdots \ x_m]$.

$$Z = [\, X \quad I \,]^T [\, X \quad I \,] \quad \Longleftrightarrow \quad Z = \begin{bmatrix} Y & X^T \\ X & I \end{bmatrix} \succeq 0, \ \operatorname{rank} Z = 2$$

SDP relaxation (Biswas,Ye '03):

$$\begin{aligned} v_{\mathrm{sdp}} := \quad & \min_Z \sum_{(i,j)\in\mathcal{A}, j>m} \left| Y_{ii} - 2x_j^T x_i + \|x_j\|^2 - d_{ij}^2 \right| \\ & \quad + \sum_{(i,j)\in\mathcal{A}, j\le m} \left| Y_{ii} - 2Y_{ij} + Y_{jj} - d_{ij}^2 \right| \\ & \text{s.t.} \quad Z = \begin{bmatrix} Y & X^T \\ X & I \end{bmatrix} \succeq 0 \end{aligned}$$

Adding the nonconvex constraint $\operatorname{rank} Z = 2$ yields original problem.

But SDP relaxation is still expensive to solve for m large..

ESDP Relaxation

ESDP relaxation (Wang, Zheng, Boyd, Ye '06):

$$
\begin{aligned}
\upsilon_{\text{esdp}} := \quad & \min_{Z} \sum_{(i,j)\in\mathcal{A},\, j>m} \left| Y_{ii} - 2x_j^T x_i + \|x_j\|^2 - d_{ij}^2 \right| \\
& + \sum_{(i,j)\in\mathcal{A},\, j\le m} \left| Y_{ii} - 2Y_{ij} + Y_{jj} - d_{ij}^2 \right| \\
\text{s.t.} \quad & Z = \begin{bmatrix} Y & X^T \\ X & I \end{bmatrix} \\
& \begin{bmatrix} Y_{ii} & Y_{ij} & x_i^T \\ Y_{ij} & Y_{jj} & x_j^T \\ x_i & x_j & I \end{bmatrix} \succeq 0 \quad \forall (i,j) \in \mathcal{A},\, j \le m \\
& \begin{bmatrix} Y_{ii} & x_i^T \\ x_i & I \end{bmatrix} \succeq 0 \qquad \forall i \le m
\end{aligned}
$$

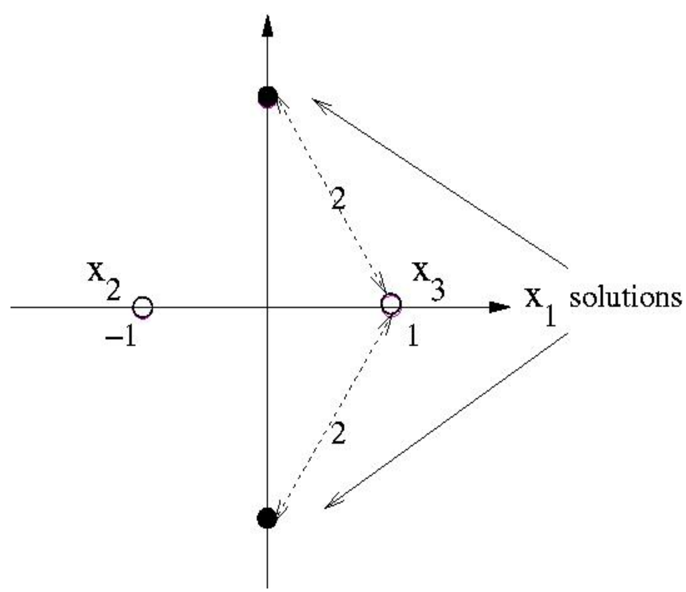
$0 \le \upsilon_{\text{esdp}} \le \upsilon_{\text{sdp}} \le \upsilon_{\text{opt}}$. In simulation, ESDP is nearly as strong as SDP, and solvable much faster by IP method.

An Example

$n = 3$, $m = 1$, $d_{12} = d_{13} = 2$

Problem:

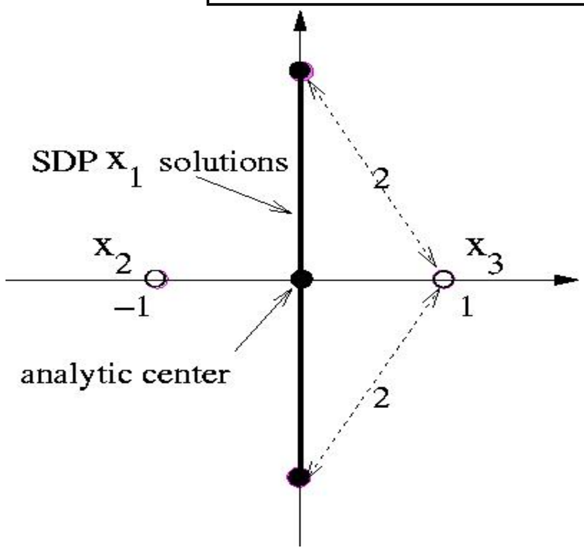
$$0 = \min_{x_1 \in \Re^2} \quad |\|x_1 - (1,0)\|^2 - 4| + |\|x_1 - (-1,0)\|^2 - 4|$$



SDP/ESDP Relaxation:

$$0 = \min_{\substack{x_1=(\alpha,\beta)\in\Re^2 \\ Y_{11}\in\Re}} |Y_{11} - 2\alpha - 3| + |Y_{11} + 2\alpha - 3|$$

$$\text{s.t.} \quad \begin{bmatrix} Y_{11} & \alpha & \beta \\ \alpha & 1 & 0 \\ \beta & 0 & 1 \end{bmatrix} \succeq 0$$



If solve SDP/ESDP by IP method, then likely get analy. center.

Properties of SDP & ESDP Relaxations

Assume each $i \le m$ is conn. to some $j > m$ in the graph $(\{1, ..., n\}, \mathcal{A})$.

Fact 0:

- $\mathrm{Sol(SDP)}$ and $\mathrm{Sol(ESDP)}$ are nonempty, closed, convex, bounded.
- If

$$d_{ij} = \|x_i^{\mathrm{true}} - x_j^{\mathrm{true}}\| \ \forall\, (i,j) \in \mathcal{A}$$

"noiseless case"

($x_i^{\mathrm{true}} = x_i \ \forall\, i > m$), then

$$\upsilon_{\mathrm{opt}} = \upsilon_{\mathrm{sdp}} = \upsilon_{\mathrm{esdp}} = 0$$

and

$$Z^{\mathrm{true}} := \begin{bmatrix} X^{\mathrm{true}} & I \end{bmatrix}^T \begin{bmatrix} X^{\mathrm{true}} & I \end{bmatrix}$$

is a soln of SDP and ESDP (i.e., $Z^{\mathrm{true}} \in \mathrm{Sol(SDP)} \subseteq \mathrm{Sol(ESDP)}$).

Let $$\mathrm{tr}_i[Z] := Y_{ii} - \|x_i\|^2, \quad i = 1, ..., m.$$ "ith trace"

Fact 1 (Biswas,Ye '03, T '07, Wang et al '06): For each i,

$$\mathrm{tr}_i[Z] = 0 \; \exists Z \in \mathrm{ri}(\mathrm{Sol}(\mathrm{ESDP})) \quad \Longrightarrow \quad x_i \text{ is invariant over Sol(ESDP)}$$
$$\text{(so } \; x_i = x_i^{\mathrm{true}} \text{ in noiseless case)}$$

Still true with "ESDP" changed to "SDP".

Let $\mathrm{tr}_i[Z] := Y_{ii} - \|x_i\|^2, \quad i = 1, ..., m.$ "ith trace"

Fact 1 (Biswas,Ye '03, T '07, Wang et al '06): For each i,

$$\mathrm{tr}_i[Z] = 0 \; \exists Z \in \mathrm{ri}(\mathrm{Sol}(\mathrm{ESDP})) \quad \Longrightarrow \quad x_i \text{ is invariant over } \mathrm{Sol}(\mathrm{ESDP})$$

$$(\text{so } \; x_i = x_i^{\text{true}} \; \text{ in noiseless case})$$

Still true with "ESDP" changed to "SDP".

Fact 2 (Pong, T '08): Suppose $v_{\text{opt}} = 0$. For each i,

$$\mathrm{tr}_i[Z] = 0 \; \forall Z \in \mathrm{Sol}(\mathrm{ESDP}) \quad \Longleftrightarrow \quad x_i \text{ is invariant over } \mathrm{Sol}(\mathrm{ESDP}).$$

Proof is by induction, starting from sensors that neighbor anchors.
(Q: True for SDP?)

In practice, there are measurement noises:

$$d_{ij}^2 = \|x_i^{\text{true}} - x_j^{\text{true}}\|^2 + \delta_{ij} \quad \forall (i,j) \in \mathcal{A}.$$

When $\delta := (\delta_{ij})_{(i,j)\in\mathcal{A}} \approx 0$, does $\mathrm{tr}_i[Z] = 0$ (with $Z \in \mathrm{ri}(\mathrm{Sol}(\mathrm{ESDP}))$) imply $x_i \approx x_i^{\text{true}}$?

In practice, there are measurement noises:

$$d_{ij}^2 = \|x_i^{\text{true}} - x_j^{\text{true}}\|^2 + \delta_{ij} \quad \forall (i,j) \in \mathcal{A}.$$

When $\delta := (\delta_{ij})_{(i,j)\in\mathcal{A}} \approx 0$, does $\text{tr}_i[Z] = 0$ (with $Z \in \text{ri}(\text{Sol}(\text{ESDP}))$) imply $x_i \approx x_i^{\text{true}}$? No! 

Fact 3 (Pong, T '08): For $\delta \approx 0$ and for each i,

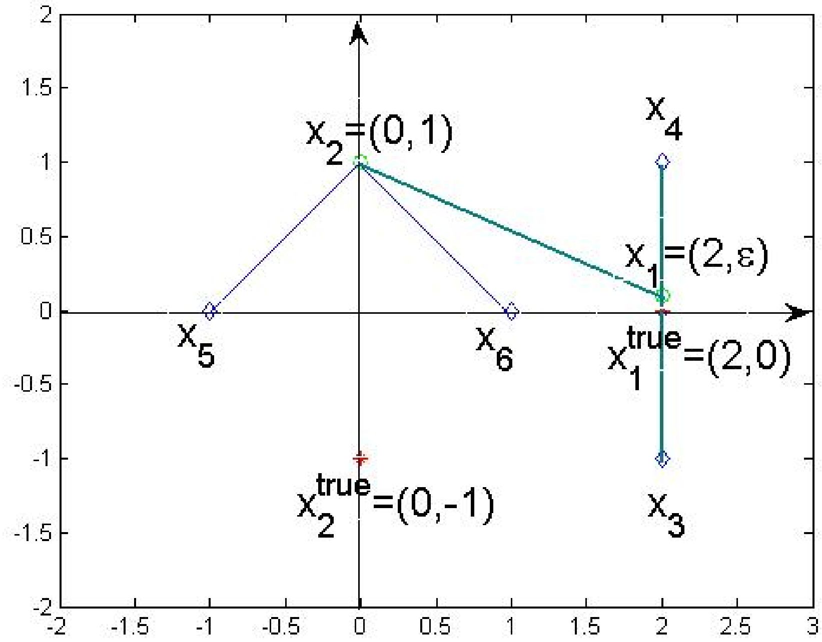
$$\text{tr}_i[Z] = 0 \; \exists Z \in \text{ri}(\text{Sol}(\text{ESDP})) \quad \not\Longrightarrow \quad x_i \approx x_i^{\text{true}}.$$

Still true with “ESDP” changed to “SDP”.

Proof is by counter-example.

An example of sensitivity of ESDP solns to measurement noise:

Input distance data: $\epsilon > 0$
$d_{12} = \sqrt{4 + (1-\epsilon)^2}, d_{13} = 1 + \epsilon, d_{14} = 1 - \epsilon, d_{25} = d_{26} = \sqrt{2}$; $m = 2$, $n = 6$.



Thus, even when $Z \in \mathrm{Sol(ESDP)}$ is unique, $\mathrm{tr}_i[Z] = 0$ fails to certify accuracy of x_i in the noisy case!

Robust ESDP

Fix any $\rho_{ij} > |\delta_{ij}|\ \forall (i,j) \in \mathcal{A}$ ($\rho > |\delta|$).

Let $\mathrm{Sol}(\rho\mathrm{ESDP})$ denote the set of $Z = \begin{bmatrix} Y & X^T \\ X & I \end{bmatrix}$ satisfying

$$
\begin{aligned}
\begin{bmatrix} Y_{ii} & Y_{ij} & x_i^T \\ Y_{ij} & Y_{jj} & x_j^T \\ x_i & x_j & I \end{bmatrix} &\succeq 0 \quad \forall (i,j) \in \mathcal{A}, j \le m \\
\begin{bmatrix} Y_{ii} & x_i^T \\ x_i & I \end{bmatrix} &\succeq 0 \quad \forall i \le m \\
|Y_{ii} - 2x_j^T x_i + \|x_j\|^2 - d_{ij}^2| &\le \rho \quad \forall (i,j) \in \mathcal{A}, j > m \\
|Y_{ii} - 2Y_{ij} + Y_{jj} - d_{ij}^2| &\le \rho \quad \forall (i,j) \in \mathcal{A}, j \le m
\end{aligned}
$$

Note: $Z^{\mathrm{true}} = \begin{bmatrix} X^{\mathrm{true}} & I \end{bmatrix}^T \begin{bmatrix} X^{\mathrm{true}} & I \end{bmatrix} \in \mathrm{Sol}(\rho\mathrm{ESDP})$.

Let

$$Z^{\rho,\delta} := \arg\min_{Z \in \mathrm{Sol}(\rho\mathrm{ESDP})} \quad - \sum_{(i,j)\in\mathcal{A}, j\le m} \ln\det \begin{bmatrix} Y_{ii} & Y_{ij} & x_i^T \\ Y_{ij} & Y_{jj} & x_j^T \\ x_i & x_j & I \end{bmatrix} - \sum_{i\le m} \ln\det \begin{bmatrix} Y_{ii} & x_i^T \\ x_i & I \end{bmatrix}$$

Let

$$Z^{\rho,\delta} := \arg\min_{Z \in \mathrm{Sol}(\rho\mathrm{ESDP})} \; - \sum_{(i,j)\in\mathcal{A}, j\le m} \ln\det \begin{bmatrix} Y_{ii} & Y_{ij} & x_i^T \\ Y_{ij} & Y_{jj} & x_j^T \\ x_i & x_j & I \end{bmatrix} - \sum_{i\le m} \ln\det \begin{bmatrix} Y_{ii} & x_i^T \\ x_i & I \end{bmatrix}$$

Fact 4 (Pong, T '08): $\exists\ \eta > 0$ and $\bar{\rho} > 0$ such that for each i,

$$\mathrm{tr}_i[Z^{\rho,\delta}] < \eta \;\; \exists |\delta| < \rho \le \bar{\rho}e \implies \lim_{|\delta|<\rho\to 0} x_i^{\rho,\delta} = x_i^{\mathrm{true}}$$

$$\mathrm{tr}_i[Z^{\rho,\delta}] > \frac{\eta}{10} \;\; \exists |\delta| < \rho \le \bar{\rho}e \implies x_i \text{ not invar. over } \mathrm{Sol}(\mathrm{ESDP}) \text{ when } \delta = 0$$

Moreover,

$$\|x_i^{\rho,\delta} - x_i^{\mathrm{true}}\| \le \sqrt{2|\mathcal{A}| + m}\sqrt{\mathrm{tr}_i[Z^{\rho,\delta}]} \quad \forall\ |\delta| < \rho.$$

BCGD-Barrier Method

Compute $Z^{\rho,\delta}$? Let $h_a(t) := \frac{1}{2}(t-a)_+^2 + \frac{1}{2}(-t-a)_+^2$ and

$$
\begin{aligned}
f_\mu(Z) \quad := \quad & \sum_{(i,j)\in\mathcal{A}, j>m} h_{\rho_{ij}}(y_{ii} - 2x_j^T x_i + \|x_j\|^2 - d_{ij}^2) \\
& + \sum_{(i,j)\in\mathcal{A}, j\le m} h_{\rho_{ij}}(y_{ii} - 2y_{ij} + y_{jj} - d_{ij}^2) \\
& -\mu \sum_{(i,j)\in\mathcal{A}, j\le m} \ln\det \begin{bmatrix} y_{ii} & y_{ij} & x_i^T \\ y_{ij} & y_{jj} & x_j^T \\ x_i & x_j & I \end{bmatrix} \\
& -\mu \sum_{i\le m} \ln\det \begin{bmatrix} y_{ii} & x_i^T \\ x_i & I \end{bmatrix}
\end{aligned}
$$

- For each $(i,j) \in \mathcal{A}$ with $j > m$ (resp. $j \leq m$),
$$h_{\rho_{ij}}(y_{ii} - 2x_j^T x_i + \|x_j\|^2 - d_{ij}^2) = 0 \quad \Longleftrightarrow \quad |y_{ii} - 2x_j^T x_i + \|x_j\|^2 - d_{ij}^2| \leq \rho_{ij}$$
$$(\text{resp. } h_{\rho_{ij}}(y_{ii} - 2y_{ij} + y_{jj} - d_{ij}^2) = 0 \quad \Longleftrightarrow \quad |y_{ii} - 2y_{ij} + y_{jj} - d_{ij}^2| \leq \rho_{ij}).$$

- f_μ is partially separable, strictly convex & diff. on its domain.

- For each fixed $\rho > |\delta|$, $\operatorname{argmin} f_\mu \to Z^{\rho,\delta}$ as $\mu \to 0$.

- In the noiseless case ($\delta = 0$), if we set $\rho = 0$, then $\operatorname{argmin} f_\mu \to Z$ as $\mu \to 0$, for some $Z \in \operatorname{ri}(\operatorname{Sol}(\text{ESDP}))$.

- For each $(i,j) \in \mathcal{A}$ with $j > m$ (resp. $j \le m$),

$$h_{\rho_{ij}}(y_{ii} - 2x_j^T x_i + \|x_j\|^2 - d_{ij}^2) = 0 \quad \Longleftrightarrow \quad |y_{ii} - 2x_j^T x_i + \|x_j\|^2 - d_{ij}^2| \le \rho_{ij}$$

$$(\text{resp. } h_{\rho_{ij}}(y_{ii} - 2y_{ij} + y_{jj} - d_{ij}^2) = 0 \quad \Longleftrightarrow \quad |y_{ii} - 2y_{ij} + y_{jj} - d_{ij}^2| \le \rho_{ij}).$$

- f_μ is partially separable, strictly convex & diff. on its domain.
- For each fixed $\rho > |\delta|$, $\operatorname{argmin} f_\mu \to Z^{\rho,\delta}$ as $\mu \to 0$.
- In the noiseless case ($\delta = 0$), if we set $\rho = 0$, then $\operatorname{argmin} f_\mu \to Z$ as $\mu \to 0$, for some $Z \in \operatorname{ri}(\operatorname{Sol}(\operatorname{ESDP}))$.

Idea: Approx. min f_μ by block-coordinate gradient descent (BCGD). (T, Yun '06)

BCGD-Barrier Method:

Given Z in $\mathrm{dom} f_\mu$, compute gradient $\nabla_{Z_i} f_\mu$ of f_μ w.r.t. $Z_i := \{x_i, y_{ii}, y_{ij} : (i,j) \in \mathcal{A}\}$ for each i.

- If $\|\nabla_{Z_i} f_\mu\| \geq \max\{\mu, 10^{-6}\}$ for some i, update Z_i by moving along the Newton direction $-\left(\nabla^2_{Z_i Z_i} f_\mu\right)^{-1} \nabla_{Z_i} f_\mu$ with Armijo stepsize rule.
- Decrease μ when $\|\nabla_{Z_i} f_\mu\| < \max\{\mu, 10^{-6}\} \ \forall\, i$.

$\mu_{\mathrm{initial}} = 100$, $\mu_{\mathrm{final}} = 10^{-9}$. Decrease μ by a factor of 10 each time.

Coded in Fortran. Computation easily distributes.

Simulation Results

- Compare ρESDP as solved by BCGD-barrier with ESDP as solved by Sedumi 1.05 Sturm (with the interface to Sedumi coded by Wang et al).

Simulation Results

- Compare ρESDP as solved by BCGD-barrier with ESDP as solved by Sedumi 1.05 Sturm (with the interface to Sedumi coded by Wang et al).

- Anchors and sensors $x_1^{\text{true}}, ..., x_n^{\text{true}}$ uniformly distributed in $[-.5, .5]^2$, $m = .9n$. $(i, j) \in \mathcal{A}$ whenever $\|x_i^{\text{true}} - x_j^{\text{true}}\| < rr$. Set

$$d_{ij} = \|x_i^{\text{true}} - x_j^{\text{true}}\| \cdot (1 + nf \cdot \epsilon_{ij})_+,$$

where $\epsilon_{ij} \sim N(0, 1)$.

Simulation Results

- Compare ρESDP as solved by BCGD-barrier with ESDP as solved by Sedumi 1.05 Sturm (with the interface to Sedumi coded by Wang et al).

- Anchors and sensors $x_1^{\text{true}}, ..., x_n^{\text{true}}$ uniformly distributed in $[-.5, .5]^2$, $m = .9n$. $(i, j) \in \mathcal{A}$ whenever $\|x_i^{\text{true}} - x_j^{\text{true}}\| < rr$. Set

$$d_{ij} = \|x_i^{\text{true}} - x_j^{\text{true}}\| \cdot (1 + nf \cdot \epsilon_{ij})_+,$$

  where $\epsilon_{ij} \sim N(0, 1)$.

- Sensor i is judged as "accurately positioned" if

$$\text{tr}_i[Z^{\text{found}}] < 5 \cdot 10^{-6} + 0.02\, nf.$$

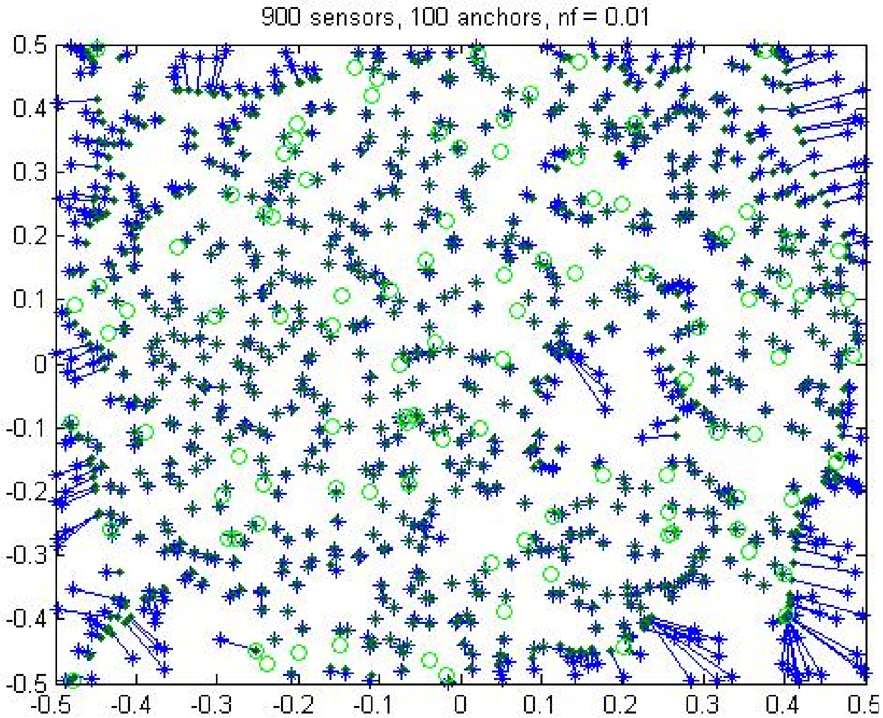
				$\rho\mathbf{ESDP}_{\text{BCGD}-\text{barrier}}$	$\mathbf{ESDP}_{\text{Sedumi}}$
n	m	nf	rr	**cpu/**m_{ap}**/**err_{ap}	**cpu(cpus)/**m_{ap}**/**err_{ap}
1000	900	0	.06	31/574/3.4e-4	189(106)/626/2.2e-4
1000	900	.001	.06	23/520/2.8e-3	170(89)/624/3.1e-3
1000	900	.01	.06	14/517/1.1e-2	128(48)/664/1.5e-2
2000	1800	0	.06	63/1626/1.7e-4	1157(397)/1689/2.7e-4
2000	1800	.001	.06	50/1596/8.5e-4	1255(503)/1653/1.3e-3
2000	1800	.01	.06	52/1602/6.2e-3	1374(417)/1689/1.2e-2

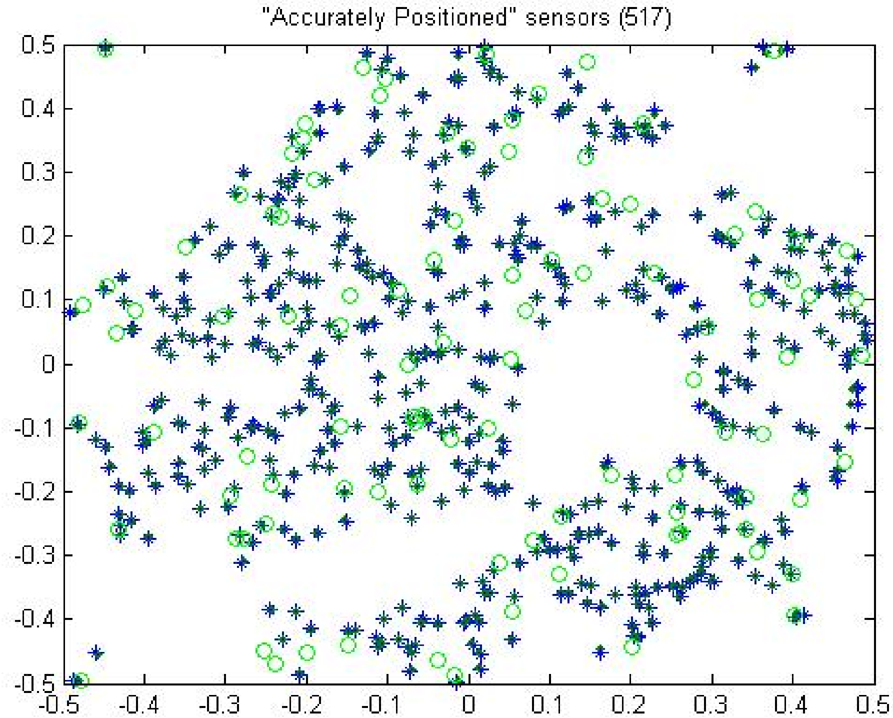
- cpu(sec) times are on a HP DL360 workstation, running Linux 3.5. ESDP is solved by Sedumi; cpus:$=$ run time for Sedumi.

- Set $\rho_{ij} = d_{ij}^2 \cdot ((1 - 2 \cdot nf)^{-2} - 1)$.

- m_{ap} := # accurately positioned sensors.
  $err_{ap} := \max_{i \text{ accurate. pos.}} \|x_i - x_i^{\text{true}}\|$.

900 sensors, 100 anchors, $rr = 0.06$, $nf = 0.01$, solve ρESDP by BCGD-barrier. x_i^{true} (shown as $*$) and $x_i^{\rho,\delta}$ (shown as $\bullet$) are joined by blue line segment; anchors are shown as $\circ$.





Conclusion & Ongoing work

- SDP and ESDP solns are sensitive to measurement noise. Lack soln accuracy certificate (though the trace test works well enough in simulation).

- ρESDP has more stable solns. Has soln accuracy certificate (which works well enough in simulation). Needs to estimate the noise level δ to set ρ. Can $\rho > |\delta|$ be relaxed?

- Approximation bounds? Extension to maxmin dispersion problem.

Thanks for coming!