ESDP Relaxation of Sensor Network Localization Analysis, Extensions and Algorithm

Ting Kei Pong
Mathematics, University of Washington
Seattle

ISMP 2009
August 2009
(Joint work with Paul Tseng)

Talk Outline

- Sensor network localization.
- ESDP relaxation: properties and accuracy certificate.
- A robust version of ESDP for the noisy case.
- LPCGD algorithm and numerical simulations.
- Conclusion and extension.

Sensor Network Localization

Basic Problem:

- n pts $\underbrace{x_1, ..., x_m}_{\text{sensors}}, \underbrace{x_{m+1}, ..., x_n}_{\text{anchors}}$ in $\mathbb{R}^2$.

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

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

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

- Estimate the first m pts ('sensors') $x_1, ..., x_m$.

Optimization Problem Formulation

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

Optimization Problem Formulation

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

- Objective function is nonconvex. m can be large ($m > 1000$).
- Problem is NP-hard (reduction from PARTITION).
- Use a convex relaxation.
- Fast, distributed algorithm.

SDP Relaxation

Let $X = [x_1 \cdots x_m]$. $\qquad Y = X^T X \iff Z = \begin{bmatrix} Y & X^T \\ X & I \end{bmatrix} \succeq 0, \ \mathrm{rank}\, Z = 2$.

SDP Relaxation

Let $X = [x_1 \cdots x_m]$. $\qquad Y = X^T X \Longleftrightarrow Z = \begin{bmatrix} Y & X^T \\ X & I \end{bmatrix} \succeq 0, \ \mathrm{rank}\, Z = 2.$

SDP relaxation (Biswas, Ye '03):

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

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

ESDP Relaxation

ESDP relaxation (Wang, Zheng, Boyd, Ye ’07):

$$
\begin{aligned}
v_{\text{esdp}} := \quad & \min_{Z} \sum_{(i,j)\in\mathcal{A}^a} \left| y_{ii} - 2x_j^T x_i + \|x_j\|^2 - d_{ij}^2 \right| \\
& \quad + \sum_{(i,j)\in\mathcal{A}^s} \left| y_{ii} - 2y_{ij} + y_{jj} - d_{ij}^2 \right| \\
\text{s.t.} \quad & Z = \begin{bmatrix} Y & X^T \\ X & I \end{bmatrix} \\
& \begin{bmatrix} y_{ii} & y_{ij} & x_i^T \\ y_{ij} & y_{jj} & x_j^T \\ x_i & x_j & I \end{bmatrix} \succeq 0 \quad \forall (i,j) \in \mathcal{A}^s \\
& \begin{bmatrix} y_{ii} & x_i^T \\ x_i & I \end{bmatrix} \succeq 0 \quad \forall i = 1, \cdots, m.
\end{aligned}
$$

ESDP Relaxation

ESDP relaxation (Wang, Zheng, Boyd, Ye '07):

$$
\begin{aligned}
v_{\text{esdp}} := \quad & \min_Z \sum_{(i,j)\in\mathcal{A}^a} \left| y_{ii} - 2x_j^T x_i + \|x_j\|^2 - d_{ij}^2 \right| \\
& \quad + \sum_{(i,j)\in\mathcal{A}^s} \left| y_{ii} - 2y_{ij} + y_{jj} - d_{ij}^2 \right| \\
\text{s.t.} \quad & Z = \begin{bmatrix} Y & X^T \\ X & I \end{bmatrix} \\
& \begin{bmatrix} y_{ii} & y_{ij} & x_i^T \\ y_{ij} & y_{jj} & x_j^T \\ x_i & x_j & I \end{bmatrix} \succeq 0 \quad \forall (i,j) \in \mathcal{A}^s \\
& \begin{bmatrix} y_{ii} & x_i^T \\ x_i & I \end{bmatrix} \succeq 0 \quad \forall i = 1, \cdots, m.
\end{aligned}
$$

In simulations, ESDP is nearly as strong as SDP relaxation, and solvable much faster by IP method.

Properties of ESDP

Assume

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

"noiseless case"

($x_j^{\text{true}} = x_j \;\forall\, j > m$)

Fact 0:

$$Z^{\text{true}} := [X^{\text{true}} \;\; I]^T [X^{\text{true}} \;\; I] = \begin{bmatrix} (X^{\text{true}})^T X^{\text{true}} & (X^{\text{true}})^T \\ X^{\text{true}} & I \end{bmatrix}$$

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

Properties of ESDP

Assume

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

"noiseless case"

($x_j^{\text{true}} = x_j \;\forall\, j > m$)

Fact 0:

$$Z^{\text{true}} := [X^{\text{true}} \;\; I]^T [X^{\text{true}} \;\; I] = \begin{bmatrix} (X^{\text{true}})^T X^{\text{true}} & (X^{\text{true}})^T \\ X^{\text{true}} & I \end{bmatrix}$$

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

x_i is invariant over $\mathrm{Sol(ESDP)} \Rightarrow x_i = x_i^{\text{true}}$.

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

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

Fact 1 (Wang et al '07): For each i,

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

$$(\text{so } \ x_i = x_i^{\text{true}})$$

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

Fact 1 (Wang et al '07): For each i,

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

$$(\text{so } \ x_i = x_i^{\text{true}})$$

Is this suff. condition for invariance also necessary?

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

Fact 1 (Wang et al '07): For each i,

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

Is this suff. condition for invariance also necessary?

Fact 2 (P, Tseng '09): For each i,

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

In practice, there are measurement noises:

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

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 is near the true position of sensor i?

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 is near the true position of sensor i?

No.

Fact 3 (P, Tseng '09): For $|\delta_{ij}| \approx 0$,

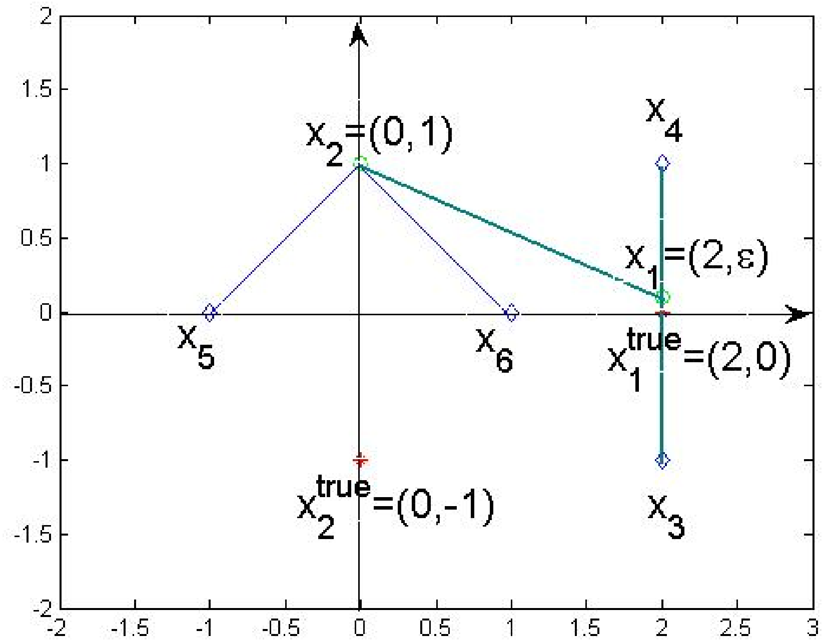
$$\text{tr}_i[Z] = 0 \quad \text{for some } Z \in \text{ri}(\text{Sol}(\text{ESDP})) \not\Longrightarrow \quad \|x_i - x_i^{\text{true}}\| \approx 0.$$

Proof is by counterexample.

An example of sensitivity of ESDP solns to measurement noise:

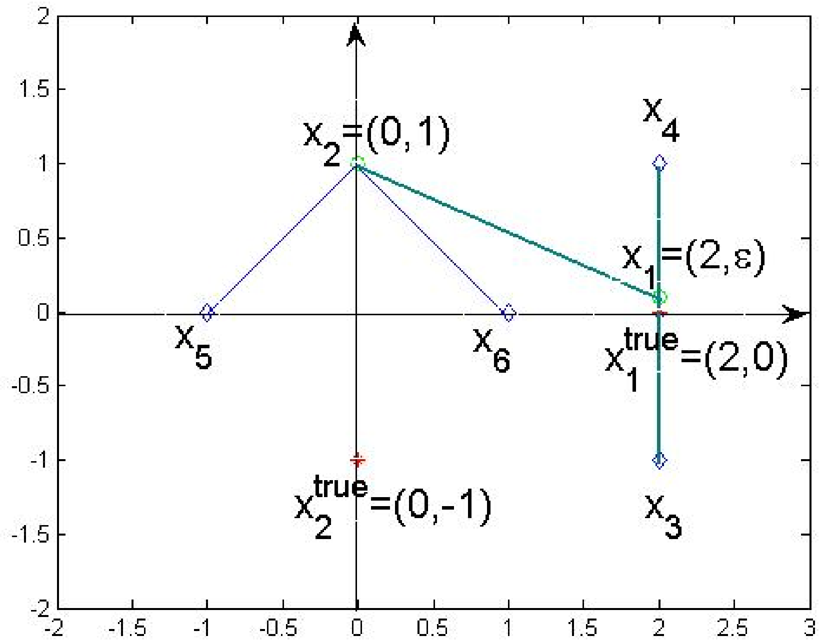
Input distance data: $\epsilon > 0$

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



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

For each $(i,j) \in \mathcal{A}$, fix $\rho_{ij} > |\delta_{ij}|$ ($\rho > |\delta|$).

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

$$
\begin{aligned}
\begin{bmatrix} y_{ii} & y_{ij} & x_i^T \\ y_{ij} & y_{jj} & x_j^T \\ x_i & x_j & I \end{bmatrix} &\succeq 0 \quad \forall (i,j) \in \mathcal{A}^s \\
\begin{bmatrix} y_{ii} & x_i^T \\ x_i & I \end{bmatrix} &\succeq 0 \quad \forall i = 1, \cdots, m \\
|y_{ii} - 2x_j^T x_i + \|x_j\|^2 - d_{ij}^2| &\le \rho_{ij} \quad \forall (i,j) \in \mathcal{A}^a \\
|y_{ii} - 2y_{ij} + y_{jj} - d_{ij}^2| &\le \rho_{ij} \quad \forall (i,j) \in \mathcal{A}^s.
\end{aligned}
$$

Robust ESDP

For each $(i,j) \in \mathcal{A}$, fix $\rho_{ij} > |\delta_{ij}|$ $(\rho > |\delta|)$.

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

$$
\begin{array}{rll}
\begin{bmatrix} y_{ii} & y_{ij} & x_i^T \\ y_{ij} & y_{jj} & x_j^T \\ x_i & x_j & I \end{bmatrix} & \succeq 0 & \forall (i,j) \in \mathcal{A}^s \\
\begin{bmatrix} y_{ii} & x_i^T \\ x_i & I \end{bmatrix} & \succeq 0 & \forall i = 1, \cdots, m \\
|y_{ii} - 2x_j^T x_i + \|x_j\|^2 - d_{ij}^2| & \leq \rho_{ij} & \forall (i,j) \in \mathcal{A}^a \\
|y_{ii} - 2y_{ij} + y_{jj} - d_{ij}^2| & \leq \rho_{ij} & \forall (i,j) \in \mathcal{A}^s.
\end{array}
$$

Note: $Z^{\mathrm{true}} = [X^{\mathrm{true}} \ \ I]^T [X^{\mathrm{true}} \ \ I] \in \mathrm{Sol}(\rho\mathrm{ESDP})$.

Let

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

Fact 4 (P, Tseng '09): $\exists\eta > 0,\ \bar\rho > 0$ such that for each i,

$$\mathrm{tr}_i(Z^{\rho,\delta}) < \eta, \text{ for some } 0 < \rho < \bar\rho e \quad \Longrightarrow \quad \lim_{|\delta|<\rho\to 0} x_i^{\rho,\delta} = x_i^{\mathrm{true}}.$$

Moreover,

$$\|x_i^{\rho,\delta} - x_i^{\mathrm{true}}\| \le \sqrt{2|\mathcal{A}^s| + m}\,\left(\mathrm{tr}_i[Z^{\rho,\delta}]\right)^{\frac{1}{2}} \quad 0 \le |\delta| < \rho.$$

LPCGD Algorithm

Let $h_a(t) := \frac{1}{2}(t-a)_+^2 + \frac{1}{2}(-t-a)_+^2$ ($|t| \le a \iff h_a(t) = 0$).

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

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

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

Use Block Coordinate Gradient Descent

LPCGD Algorithm:

Given $Z \in \operatorname{dom} f_\mu$, 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 $\|\nabla_{Z_i} f_\mu\| \geq \max\{\mu, 1e-7\}$ for some i, update Z_i by moving along the Newton direction $-\left(\partial^2_{Z_i Z_i} f_\mu\right)^{-1} \nabla_{Z_i} f_\mu$ with Armijo stepsize rule.

- Decrease μ when $\|\nabla_{Z_i} f_\mu\| < \max\{\mu, 1e-7\}$ for all i.

$\mu_{\text{initial}} = 10$, $\mu_{\text{final}} = 1e-14$. Decrease μ by a factor of 10.

Coded in Fortran. Computation easily distributes.

Simulation Results

- Compare ρESDP, as solved by LPCGD, and ESDP, as solved by Sedumi (with the interface to Sedumi coded by Wang et al.).

- Uniformly generate $\{x_1^{\text{true}}, ..., x_n^{\text{true}}\}$ in $[-.5, .5]^2$, $m = .9n$. Two pts are neighbors iff $\text{dist} < rr$. Set

$$d_{ij} = \|x_i^{\text{true}} - x_j^{\text{true}}\| \cdot |1 + \sigma\epsilon_{ij}|,$$

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

- Sensor i is judged as "accurately positioned" if

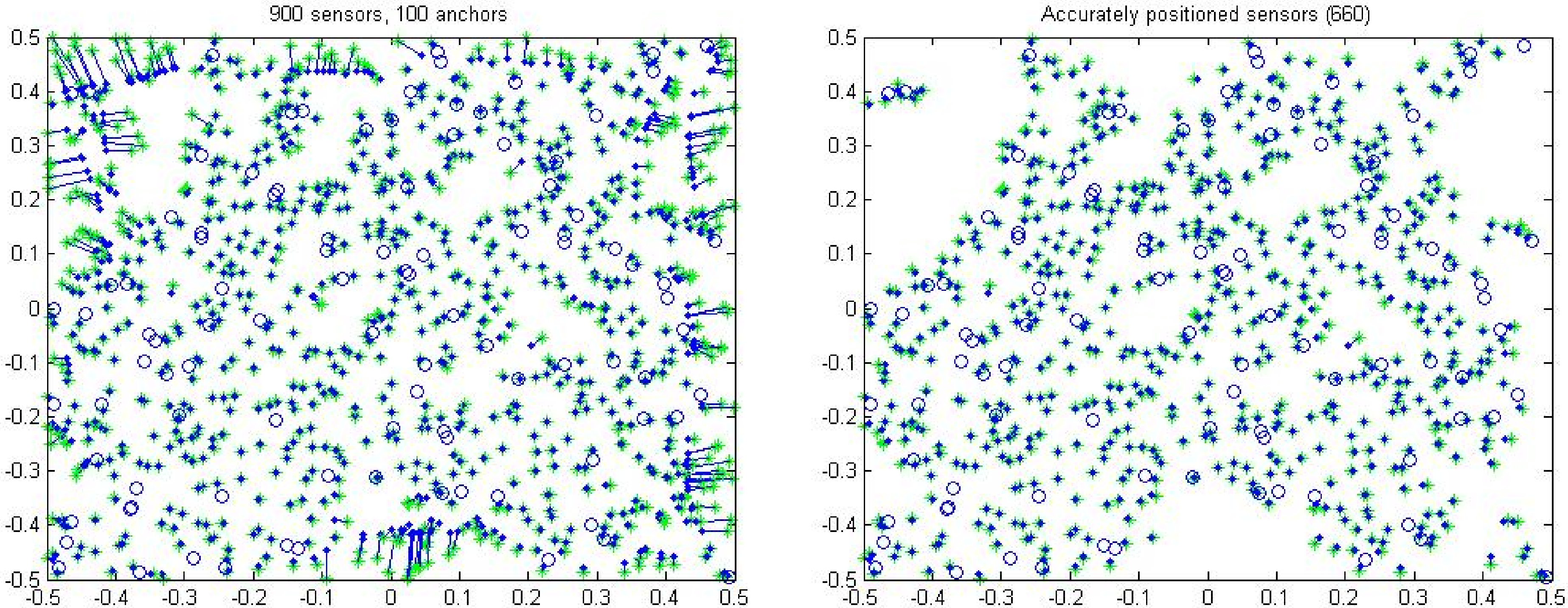
$$\text{tr}_i[Z^{\text{found}}] < (.01 + 30\sigma)\bar{d}_i^2.$$

Simulation Results

				$\rho\mathbf{ESDP}_{\text{LPCGD}}$	$\mathbf{ESDP}_{\text{Sedumi}}$
n	m	σ	rr	**cpu**/m_{ap}/err_{ap}	**cpu(cpus)**/m_{ap}/err_{ap}
1000	900	0	.06	7/662/1.7e-3	182(104)/669/2.1e-3
1000	900	.01	.06	5/660/2.2e-2	119(42)/720/3.1e-2
2000	1800	0	.06	26/1762/3.1e-4	1157(397)/1742/3.9e-4
2000	1800	.01	.06	20/1699/1.4e-2	966(233)/1746/2.4e-2
10000	9000	0	.02	77/7844/2.3e-3	16411(1297)/6481/2.5e-3
10000	9000	.01	.02	63/8336/1.0e-2	16368(1264)/8593/8.7e-3

- cpu(sec) times are on an HP DL360 workstation, running Linux 3.5. ESDP is solved by Sedumi; cpus:= time taken in running Sedumi.
- Take $\rho_{ij} = d_{ij}^2 \cdot ((1-2\tilde{\sigma})^{-2} - 1)$; $\tilde{\sigma} = \max\{\sigma, 1e-6\}$.
- $m_{\text{ap}} := \#$ accurately positioned sensors.
  $err_{\text{ap}} := \max_{i \text{ accurate. pos.}} \|x_i^{\text{found}} - x_i^{\text{true}}\|$.

900 sensors, 100 anchors, $rr = 0.06$, $\sigma = 0.01$, solving ρESDP by LPCGD. x_i^{true} denoted by green asterisks, x_i^{LPCGD} denoted by blue dots, anchors denoted by circles. x_i^{true} and x_i^{LPCGD} joined by blue straight line.



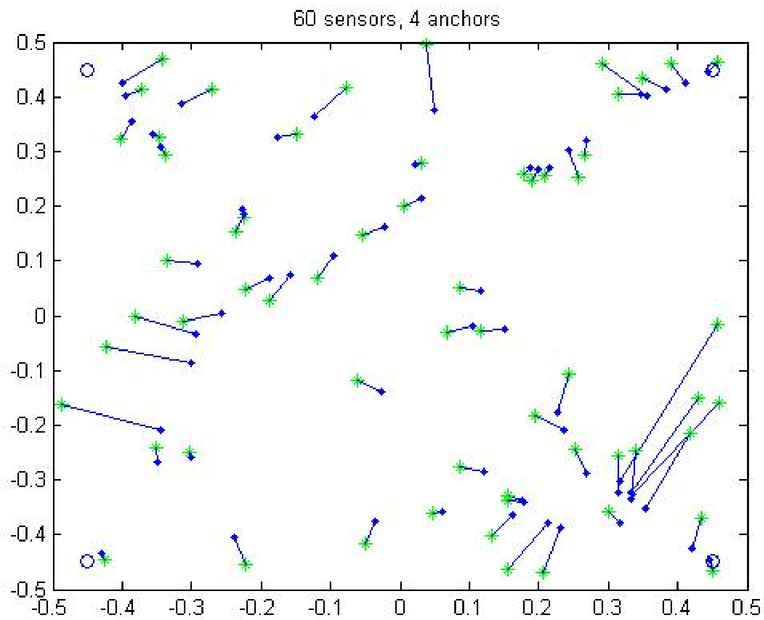
Solution Refinement

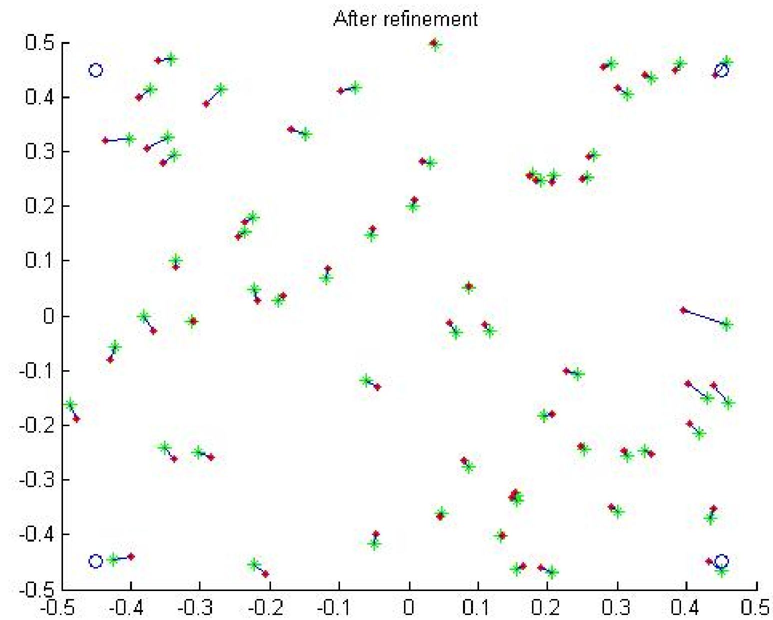
Update $x_i \leftarrow x_i - \alpha \nabla_{x_i} \hat{f}(X), \forall i$, where $\hat{f}(X) := \sum_{(i,j)\in\mathcal{A}} (\|x_i - x_j\| - d_{ij})^2$.

Solution Refinement

Update $x_i \leftarrow x_i - \alpha \nabla_{x_i} \hat{f}(X), \forall i$, where $\hat{f}(X) := \sum_{(i,j)\in\mathcal{A}} (\|x_i - x_j\| - d_{ij})^2$.

60 sensors, 4 anchors, $rr = 0.3$, $\sigma = 0.1$, solving ρESDP by LPCGD. x_i^{true} denoted by green asterisks, x_i^{LPCGD} denoted by blue dots, x_i^{refine} denoted by red dots, anchors denoted by circles. x_i^{true} and $x_i^{\text{LPCGD}}/x_i^{\text{refine}}$ joined by blue straight line.





Conclusion & Extension

- ESDP is sensitive to dist. measurement noise. Lack soln accuracy certificate.

- ρESDP has soln accuracy certificate when $\rho > |\delta|$. Can $\rho > |\delta|$ be relaxed? seems not.

- ESDP/ρESDP solns can be refined by performing gradient descent. RMSD is improved when noise is high ($\sim 10\%$).

- Extension of our analysis to SOS relaxation (Nie '06)? Adding range (upper bd/lower bd) constraints?

Thanks for coming! :)