On SDP and ESDP Relaxation of Sensor Network Localization

Paul Tseng
Mathematics, University of Washington
Seattle

7th Int. Conf. Num. Optim. Num. Lin. Algeb., Lijiang
August 9, 2009
(joint work with Ting Kei Pong)

Talk Outline

- Sensor network localization
- SDP, ESDP relaxations: properties and soln accuracy certificate

Talk Outline

- Sensor network localization
- SDP, ESDP relaxations: properties and soln accuracy certificate
- A robust version of ESDP to handle noises
- Log-barrier penalty CGD method

Talk Outline

- Sensor network localization
- SDP, ESDP relaxations: properties and soln accuracy certificate
- A robust version of ESDP to handle noises
- Log-barrier penalty CGD 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/rigidty, Euclidean matrix completion, position estimation in wireless sensor network, ...

Optimization Problem Formulation

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

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 ($m \geq 1000$).
- Problem is NP-hard (reduction from PARTITION).
- Local improvement heuristics can fail badly.

Optimization Problem Formulation

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

- Objective function is nonconvex. m can be large ($m \geq 1000$).
- Problem is NP-hard (reduction from PARTITION).
- Local improvement heuristics can fail badly.
- Use a convex (SDP, SOCP) relaxation (& local improvement).
  Low soln accuracy OK. Distributed computation preferred.

SDP Relaxation

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

SDP Relaxation

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

SDP relaxation (Biswas,Ye '03):

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

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

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

ESDP Relaxation

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

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

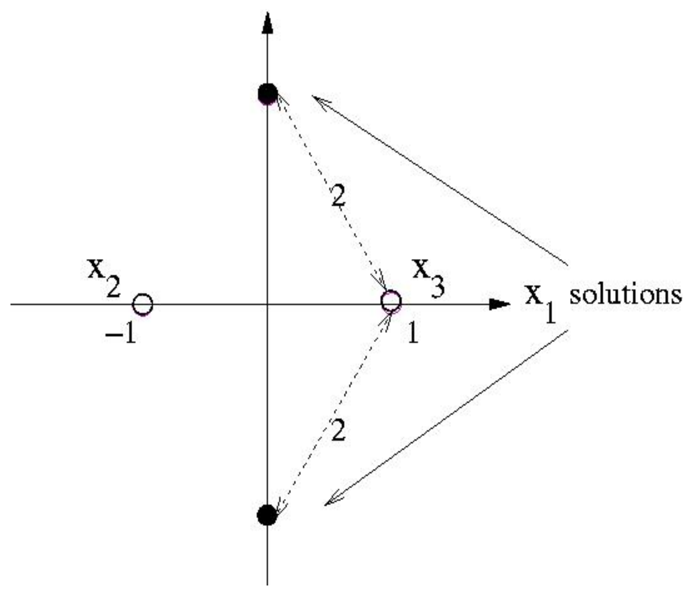
$0 \le \upsilon_{\mathrm{esdp}} \le \upsilon_{\mathrm{sdp}} \le \upsilon_{\mathrm{opt}}$. In simulation, ESDP is nearly as strong as SDP, and solvable much faster by IP method.

Example 1

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

Problem:

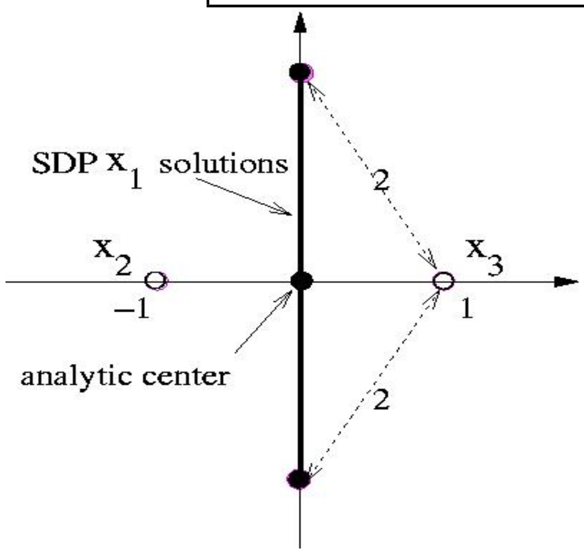
$$0 = \min_{x_1 \in \Re^2} \; \left| \left\| x_1 - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\|^2 - 4 \right| + \left| \left\| x_1 - \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right\|^2 - 4 \right|$$



SDP/ESDP Relaxation:

$$0 = \min_{\substack{x_1=\left[\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right]\in\Re^2 \\ y_{11}\in\Re}} |y_{11} - 2\alpha - 3| + |y_{11} + 2\alpha - 3|$$

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



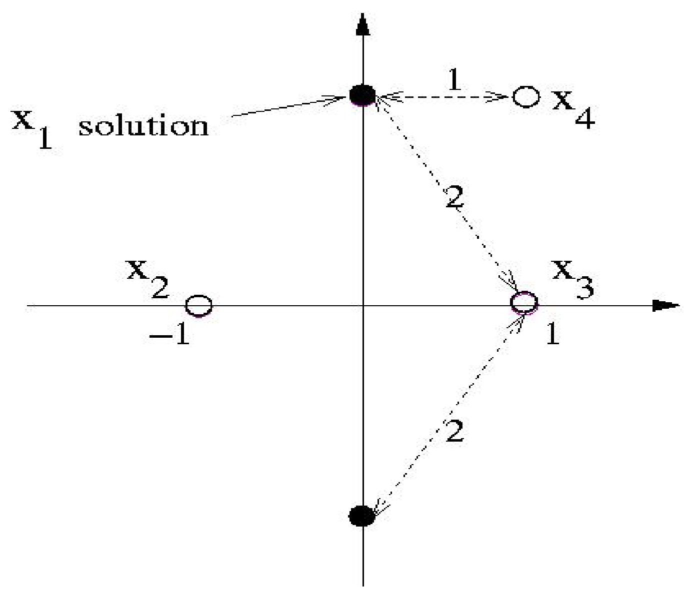
If solve SDP/ESDP by IP method, then likely get analy. center $y_{11} = 3$, $x_1 = \left[\begin{smallmatrix}0\\ 0\end{smallmatrix}\right]$

Example 2

$n = 4$, $m = 1$, $d_{12} = d_{13} = 2$, $d_{14} = 1$

Problem:

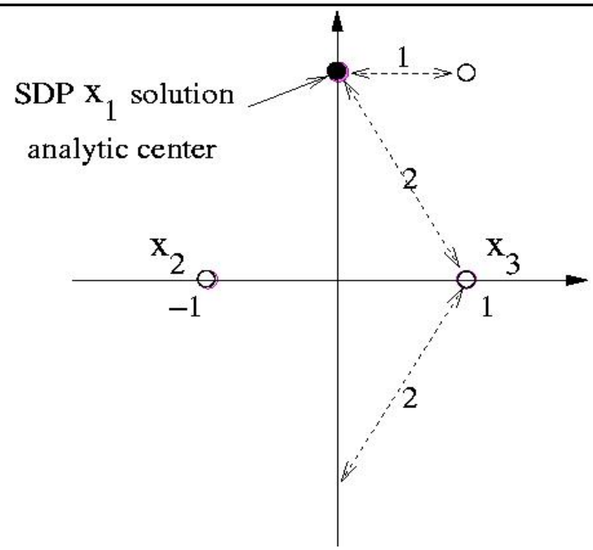
$$0 = \min_{x_1 \in \Re^2} \quad \left| \left\| x_1 - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\|^2 - 4 \right| + \left| \left\| x_1 - \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right\|^2 - 4 \right| + \left| \left\| x_1 - \begin{bmatrix} 1 \\ \sqrt{3} \end{bmatrix} \right\|^2 - 1 \right|$$



SDP/ESDP Relaxation:

$$0 = \min_{\substack{x_1 = \left[\begin{smallmatrix}\alpha \\ \beta\end{smallmatrix}\right] \in \Re^2 \\ y_{11} \in \Re}} |y_{11} - 2\alpha - 3| + |y_{11} + 2\alpha - 3| + |y_{11} - 2\alpha - 2\sqrt{3}\beta + 3|$$

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



SDP/ESDP has unique soln $y_{11} = 3$, $x_1 = \left[\begin{smallmatrix}0 \\ \sqrt{3}\end{smallmatrix}\right]$

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.
- 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,

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

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,

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

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

Fact 2 (Pong, T '09): Suppose $v_{\mathrm{opt}} = 0$. For each i,

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

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

Proof idea:

- If $(i,j) \in \mathcal{A}$ and x_i, x_j are invar. over $\mathrm{Sol(ESDP)}$, then $\mathrm{tr}_i[Z] = \mathrm{tr}_j[Z]$ $\forall Z \in \mathrm{Sol(ESDP)}$.

- Suppose $\exists i \leq m$ such that x_i is invar. over $\mathrm{Sol(ESDP)}$ but $\mathrm{tr}_i[\bar{Z}] > 0$ for some $\bar{Z} \in \mathrm{Sol(ESDP)}$. Consider maximal $\bar{\mathcal{I}} \subset \{1, \ldots, m\}$ such that x_i is invar. over $\mathrm{Sol(ESDP)}$ and $\mathrm{tr}_i[\bar{Z}] > 0$ $\forall i \in \bar{\mathcal{I}}$.

- Then x_i is not invar. over $\mathrm{Sol(ESDP)}$ $\forall i \in \mathcal{N}(\bar{\mathcal{I}})$.
  So $\exists Z \in \mathrm{ri(Sol(ESDP))}$ with $x_i \neq \bar{x}_i$ $\forall i \in \mathcal{N}(\bar{\mathcal{I}})$.

- Let $Z^\alpha = \alpha\bar{Z} + (1-\alpha)Z$ with $\alpha > 0$ suff. small.
  Can rotate x_i^α $\forall i \in \bar{\mathcal{I}}$ and Z^α still remains in $\mathrm{Sol(ESDP)}$. $\Rightarrow\Leftarrow$

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 $x_i \approx x_i^{\mathrm{true}}$?

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 $x_i \approx x_i^{\mathrm{true}}$? No!

Fact 3 (Pong, T '09): For $\delta \approx 0$ and for each i,

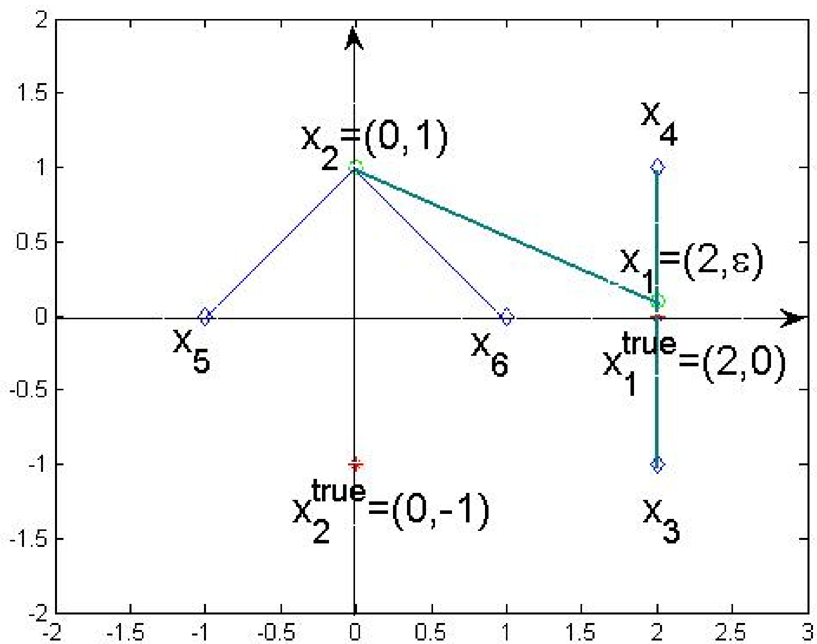
$$\mathrm{tr}_i[Z] = 0 \; \exists Z \in \mathrm{ri}(\mathrm{Sol}(\mathrm{ESDP})) \quad \not\Longrightarrow \quad x_i \approx x_i^{\mathrm{true}}.$$

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

Proof is by counter-example.

An example of sensitivity of ESDP solns to measurement noise:

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



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

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

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

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

Let

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

Fact 4 (Pong, T '09): $\exists\ \eta > 0$ and $\bar{\rho} > 0$ such that for each i,

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

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

Moreover,

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

Log-barrier Penalty CGD Method

Efficiently compute $Z^{\rho,\delta}$? Let

$$h_a(t) := \frac{1}{2}(t-a)_+^2 + \frac{1}{2}(-t-a)_+^2$$

$(|t| \le a \iff h_a(t) = 0)$ and

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

- f_μ is partially separable, strictly convex & diff. on its domain.
- For each fixed $\rho > |\delta|$, $\mathrm{argmin} f_\mu \to Z^{\rho,\delta}$ as $\mu \to 0$.

- f_μ is partially separable, strictly convex & diff. on its domain.
- For each fixed $\rho > |\delta|$, $\mathrm{argmin} f_\mu \to Z^{\rho,\delta}$ as $\mu \to 0$.

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

Log-barrier Penalty CGD 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 $\|\nabla_{Z_i} f_\mu\| \geq \max\{\mu, 10^{-7}\}$ 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 $\|\nabla_{Z_i} f_\mu\| < \max\{\mu, 10^{-6}\} \ \forall\, i$.

$\mu_{\mathrm{initial}} = 10$, $\mu_{\mathrm{final}} = 10^{-14}$. Decrease μ by a factor of 10 each time.

Coded in Fortran. Compute Newton direc. by sparse Cholesky.
Computation easily distributes.

Simulation Results

- Compare ρESDP as solved by LPCGD method 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 LPCGD method 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 $\|x_i^{\text{true}} - x_j^{\text{true}}\| \leq rr$. Set

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

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

Simulation Results

- Compare ρESDP as solved by LPCGD method 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 $\|x_i^{\text{true}} - x_j^{\text{true}}\| \le rr$. Set

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

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

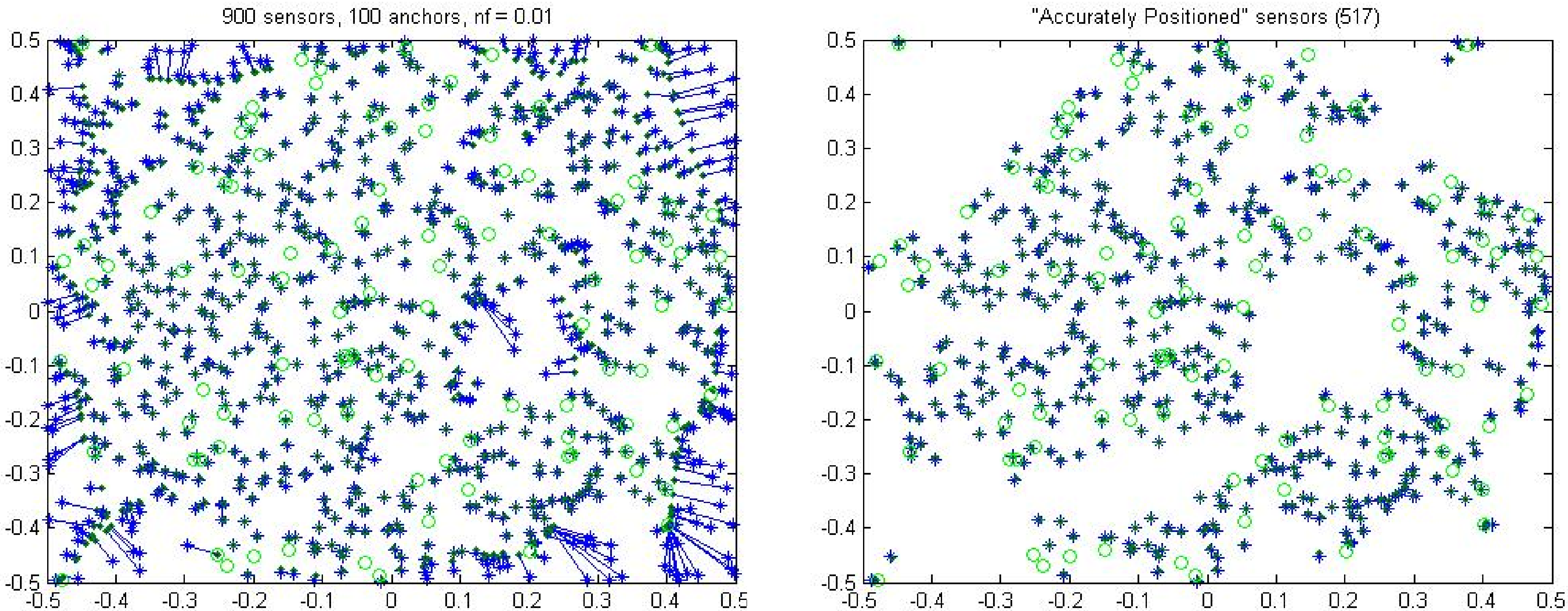
- Sensor i is judged as “accurately positioned” if

$$\text{tr}_i[Z^{\text{found}}] < (.01 + 30\sigma)d_{ij}^{\text{avg}}.$$

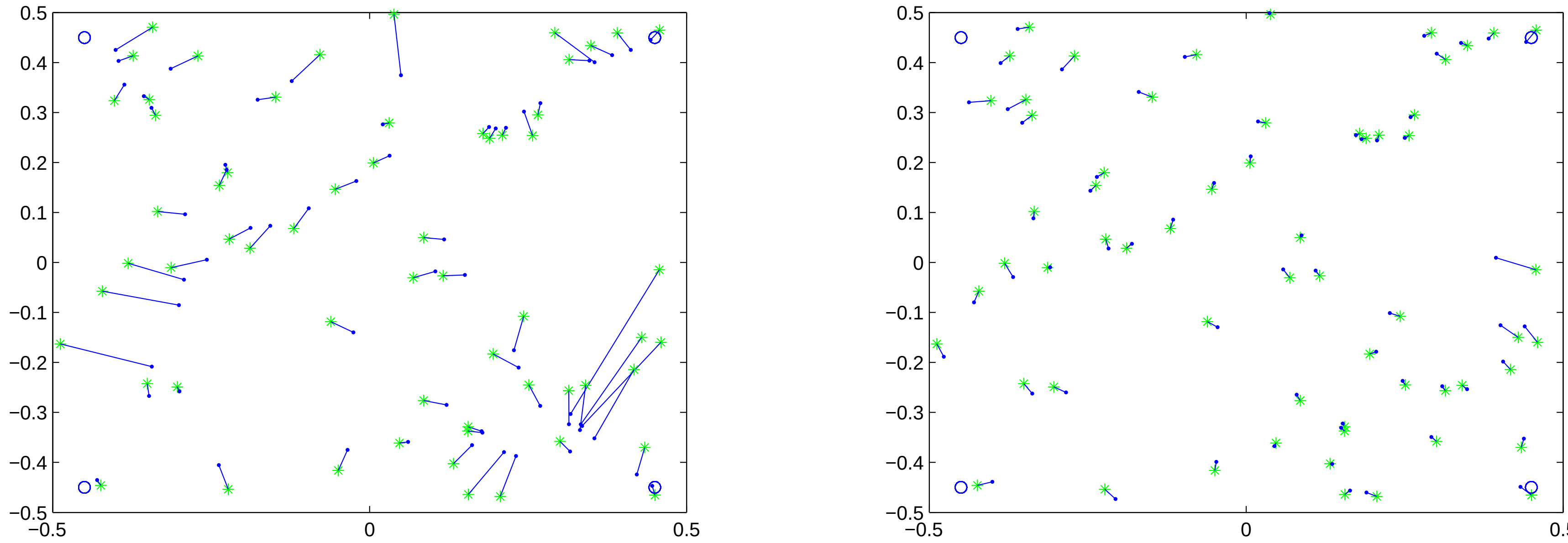
				$\rho\mathbf{ESDP}_{\text{LPCGD}}$	$\mathbf{ESDP}_{\text{Sedumi}}$
n	m	σ	rr	**cpu**/m_{ap}/err_{ap}	**cpu(cpus)**/m_{ap}/err_{ap}
1000	900	0	.06	7/662/1.7e-3	182(104)/669/2.1e-3
1000	900	.01	.06	5/660/2.2e-2	119(42)/720/3.1e-2
2000	1800	0	.06	26/1762/3.1e-4	1157(397)/1742/3.9e-4
2000	1800	.01	.06	20/1699/1.4e-2	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 a HP DL360 workstation, running Linux 3.5. ESDP is solved by Sedumi; cpus:$=$ run time for Sedumi.
- Set $\rho_{ij} = d_{ij}^2 \cdot ((1-2\sigma)^{-2} - 1)$.
- m_{ap} := # accurately positioned sensors.
  $err_{ap} := \max_{i \text{ accurate. pos.}} \|x_i - x_i^{\text{true}}\|$.

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



60 sensors, 4 anchors at corners, $rr = 0.3$, $\sigma = 0.1$. x_i^{true} (shown as $*$) and $x_i^{\rho,\delta}$ (shown as •) are joined by blue line segment; anchors are shown as ○. Left: Soln of ρESDP found by LPCGD method. Right: After local gradient improvement.



Conclusion & Ongoing work

- SDP and ESDP solns are sensitive to measurement noise. Has soln accuracy certificate under no noise only (though it works well enough in simulation).

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

- SDP, ESDP, ρESDP solns can be further refined by local improvement. This improves the rmsd when noise level is high (e.g., $\sigma = 0.1$).

- Approximation bounds? Extensions to handle lower bounds on distances (e.g., $(i, j) \notin \mathcal{A}$ imply $\|x_i^{\mathrm{true}} - x_j^{\mathrm{true}}\| > rr$)?

Thanks for coming!