On SDP and ESDP Relaxation of Sensor Network Localization

Paul Tseng
Mathematics, University of Washington
Seattle

Southern California Optimization Day, UCSD
March 19, 2009
(joint work with Ting Kei Pong)
Talk Outline

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

- Sensor network localization
- SDP, ESDP relaxations: properties and 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 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 $\mathbb{R}^2$.

- Know last $n - m$ pts (‘anchors’) $x_{m+1}, \ldots, x_n$ 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 \( \mathbb{R}^2 \).
- Know last \( n - m \) pts (‘anchors’) \( x_{m+1}, \ldots, x_n \) and Eucl. dist. estimate for pairs of ‘neighboring’ pts \( d_{ij} \geq 0 \) \( \forall (i, j) \in A \)
  
  with \( A \subseteq \{(i, j) : 1 \leq i, j \leq n\} \).
- Estimate first \( m \) pts (‘sensors’).

History? Graph realization/rigidity, Euclidean matrix completion, position estimation in wireless sensor network, ...
Optimization Problem Formulation

\[
\nu_{opt} := \min_{x_1, \ldots, x_m} \sum_{(i, j) \in A} \left| \|x_i - x_j\|^2 - d_{ij}^2 \right|
\]
### Optimization Problem Formulation

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

- 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

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

- 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?
Let $X := [x_1 \cdots x_m]$. \quad Y = X^T X \iff Z = \begin{bmatrix} Y & X^T \\ X & I \end{bmatrix} \succeq 0, \quad \text{rank} \, Z = 2
SDP Relaxation

Let \( X := [x_1 \cdots x_m] \). \( Y = X^T X \iff Z = \begin{bmatrix} Y & X^T \\ X & I \end{bmatrix} \succeq 0, \text{rank} Z = 2 \)

SDP relaxation (Biswas, Ye ’03):

\[

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

\]

Adding the nonconvex constraint \( \text{rank} Z = 2 \) yields original problem.

But SDP relaxation is still expensive to solve for \( m \) large.
ESDP Relaxation

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

\[ \nu_{\text{esdp}} := \min_Z \sum_{(i,j) \in A, i \leq m < j} \left| y_{ii} - 2x_j^T x_i + \|x_j\|^2 - d_{ij}^2 \right| + \sum_{(i,j) \in A, i < j \leq m} \left| y_{ii} - 2y_{ij} + y_{jj} - d_{ij}^2 \right| \]

s.t. \[ Z = \begin{bmatrix} Y & X^T \\ X & I \\ y_{ii} & y_{ij} & x_i^T \\ y_{ij} & y_{jj} & x_j^T \\ x_i & x_j & I \end{bmatrix} \succeq 0 \quad \forall (i,j) \in A, i < j \leq m \]

\[ 0 \leq \nu_{\text{esdp}} \leq \nu_{\text{sdp}} \leq \nu_{\text{opt}} \]. 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:

\[ 0 = \min_{x_1 \in \mathbb{R}^2} \|x_1 - \begin{bmatrix} 1 \\ 0 \end{bmatrix}\|^2 - 4 + \|x_1 - \begin{bmatrix} -1 \\ 0 \end{bmatrix}\|^2 - 4 \]
SDP/ESDP Relaxation:

\[
0 = \min_{x_1 = [\alpha \beta]^T \in \mathbb{R}^2, \ y_{11} \in \mathbb{R}} |y_{11} - 2\alpha - 3| + |y_{11} + 2\alpha - 3|
\]

s.t. \[
\begin{bmatrix}
  y_{11} & \alpha & \beta \\
  \alpha & 1 & 0 \\
  \beta & 0 & 1
\end{bmatrix} \succeq 0
\]

If solve SDP/ESDP by IP method, then likely get analy. center \( y_{11} = 3, \ x_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \)
Example 2

\( n = 4, m = 1, d_{12} = d_{13} = 2, d_{14} = 1 \)

Problem:

\[
0 = \min_{x_1 \in \mathbb{R}^2} ||x_1 - [\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}]||^2 - 4 + ||x_1 - [\begin{smallmatrix} -1 \\ 0 \end{smallmatrix}]||^2 - 4 + ||x_1 - \left[ \begin{smallmatrix} 1 \\ \sqrt{3} \end{smallmatrix} \right]||^2 - 1
\]
SDP/ESDP Relaxation:

\[ 0 = \min_{x \in \mathbb{R}^2, y_{11} \in \mathbb{R}} |y_{11} - 2\alpha - 3| + |y_{11} + 2\alpha - 3| + |y_{11} - 2\alpha - 2\sqrt{3}\beta + 3| \]

s.t. \[
\begin{bmatrix}
  y_{11} & \alpha & \beta \\
  \alpha & 1 & 0 \\
  \beta & 0 & 1 \\
\end{bmatrix} \succeq 0
\]

SDP/ESDP has unique soln \( y_{11} = 3 \),

\( x_1 = \begin{bmatrix} 0 \\ \sqrt{3} \end{bmatrix} \)
Properties of SDP & ESDP Relaxations

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

Fact 0:

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

$$d_{ij} = \|x^\text{true}_i - x^\text{true}_j\| \quad \forall (i, j) \in A$$

"noiseless case"

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

$$\nu_{\text{opt}} = \nu_{\text{sdp}} = \nu_{\text{esdp}} = 0$$

and

$$Z^\text{true} := [X^\text{true} \quad I]^T [X^\text{true} \quad I]$$

is a soln of SDP and ESDP (i.e., $Z^\text{true} \in \text{Sol}(\text{SDP}) \subseteq \text{Sol}(\text{ESDP})$).
Let
\[ \text{tr}_i[Z] := y_{ii} - \|x_i\|^2, \quad i = 1, \ldots, m. \]

“ith trace”

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

\[ \text{tr}_i[Z] = 0 \exists Z \in \text{ri}(\text{Sol(ESDP)}) \implies x_i \text{ is invariant over Sol(ESDP)} \]

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

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

Let \( \text{tr}_i[Z] := y_{ii} - \|x_i\|^2, \quad i = 1, \ldots, m. \) \hspace{1cm} \text{“ith trace”}

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

\[
\text{tr}_i[Z] = 0 \exists Z \in \text{ri(Sol(ESDP))} \implies x_i \text{ is invariant over Sol(ESDP)}
\]

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

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

**Fact 2** (Pong, T ’09): Suppose \( \nu_{opt} = 0 \). For each \( i \),

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

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 \(\text{Sol}(\text{ESDP})\), then \(\text{tr}_i[Z] = \text{tr}_j[Z]\) \(\forall Z \in \text{Sol}(\text{ESDP})\).

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

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

- Let \(Z^\alpha = \alpha \tilde{Z} + (1 - \alpha)Z\) with \(\alpha > 0\) suff. small.
  Can rotate \(x_i^\alpha \forall i \in \bar{I}\) and \(Z^\alpha\) still remains in \(\text{Sol}(\text{ESDP})\). \(\Rightarrow\Rightarrow\)
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 \( x_i \approx x_i^{\text{true}} \)?
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 ’09):** For \( \delta \approx 0 \) and for each \( i \),

\[ \text{tr}_i[Z] = 0 \; \exists Z \in \text{ri}(\text{Sol}(\text{ESDP})) \iff x_i \approx x_i^{\text{true}}. \]

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: $\epsilon > 0$

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

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

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

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

\[
\begin{align*}
|y_{ii} - 2x_j^T x_i + ||x_j||^2 - d_{ij}^2| &\leq \rho_{ij} \quad \forall (i, j) \in A, i \leq m < j \\
|y_{ii} - 2y_{ij} + y_{jj} - d_{ij}^2| &\leq \rho_{ij} \quad \forall (i, j) \in A, i < j \leq m \\
\begin{bmatrix} y_{ii} & y_{ij} & x_i^T \\
y_{ij} & y_{jj} & x_j^T \\
x_i & x_j & I \end{bmatrix} &\succeq 0 \quad \forall (i, j) \in A, i < j \leq m
\end{align*}
\]

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

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

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

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

\[
\begin{align*}
\text{tr}_i[Z^{\rho,\delta}] < \eta & \implies \exists |\delta| < \rho \leq \bar{\rho}e \implies \lim_{|\delta| < \rho \to 0} x_{i}^{\rho,\delta} = x_{i}^{\text{true}} \\
\text{tr}_i[Z^{\rho,\delta}] > \frac{\eta}{10} & \implies \exists |\delta| < \rho \leq \bar{\rho}e \implies x_i \text{ not invar. over Sol(ESDP) when } \delta = 0
\end{align*}
\]

Moreover,

\[
\|x_{i}^{\rho,\delta} - x_{i}^{\text{true}}\| \leq \sqrt{2|A| + m\sqrt{\text{tr}_i[Z^{\rho,\delta}]} \quad \forall |\delta| < \rho.
\]
Log-barrier Penalty CGD Method

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

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

$(|t| \leq a \iff h_{\alpha}(t) = 0)$ and

$$f_{\mu}(Z) := \sum_{(i,j) \in A, i \leq m < j} h_{\rho_{ij}}(y_{ii} - 2x_j^T x_i + \|x_j\|^2 - d_{ij}^2)$$

$$+ \sum_{(i,j) \in A, i < j \leq m} h_{\rho_{ij}}(y_{ii} - 2y_{ij} + y_{jj} - d_{ij}^2)$$

$$+ \mu \sum_{(i,j) \in A, i < j \leq m} -\ln \det \begin{bmatrix} y_{ii} & y_{ij} & x_i^T \\ y_{ij} & y_{jj} & x_j^T \\ x_i & x_j & I \end{bmatrix}$$
• $f_\mu$ is partially separable, strictly convex & diff. on its domain.

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

- For each fixed $\rho > |\delta|$, $\arg\min f_\mu \rightarrow Z^{\rho,\delta}$ as $\mu \rightarrow 0$.

**Idea**: Minimize $f_\mu$ approx. by block-coordinate gradient descent (BCGD). (T, Yun '06)
Log-barrier Penalty CGD Method:

Given $Z$ in $\text{dom}f_\mu$, compute gradient $\nabla Z_i f_\mu$ of $f_\mu$ w.r.t. $Z_i := \{x_i, y_{ii}, y_{ij} : (i, j) \in A\}$ 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 stepsizes rule.

- Decrease $\mu$ when $\|\nabla Z_i f_\mu\| < \max\{\mu, 10^{-6}\}$ $\forall i$.

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

Simulation Results

- Compare $\rho_{\text{ESDP}}$ 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 $\rho$ESDP 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 $x_1^{\text{true}}, \ldots, x_n^{\text{true}}$ uniformly distributed in $[-.5, .5]^2$, $m = .9n$. $(i, j) \in A$ whenever $\|x_i^{\text{true}} - x_j^{\text{true}}\| < rr$. 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 $\rho_{\text{ESDP}}$ 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 $x_1^{\text{true}}, \ldots, x_n^{\text{true}}$ uniformly distributed in $[-.5, .5]^2$, $m = .9n$. $(i,j) \in A$ whenever $\|x_i^{\text{true}} - x_j^{\text{true}}\| < rr$. 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

$$\text{tr}_i[Z^{\text{found}}] < (.01 + 30\sigma)d_{ij}^{\text{avg}}.$$
<table>
<thead>
<tr>
<th>$n$</th>
<th>$m$</th>
<th>$\sigma$</th>
<th>$rr$</th>
<th>$\rho_{ESDP}^{LPCGD}$</th>
<th>$ESDP^{Sedumi}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>900</td>
<td>0</td>
<td>.06</td>
<td>7/662/1.7e-3</td>
<td>182(104)/669/2.1e-3</td>
</tr>
<tr>
<td>1000</td>
<td>900</td>
<td>.01</td>
<td>.06</td>
<td>5/660/2.2e-2</td>
<td>119(42)/720/3.1e-2</td>
</tr>
<tr>
<td>2000</td>
<td>1800</td>
<td>0</td>
<td>.06</td>
<td>26/1762/3.1e-4</td>
<td>1157(397)/1742/3.9e-4</td>
</tr>
<tr>
<td>2000</td>
<td>1800</td>
<td>.01</td>
<td>.06</td>
<td>20/1699/1.4e-2</td>
<td>966(233)/1746/2.4e-2</td>
</tr>
<tr>
<td>10000</td>
<td>9000</td>
<td>0</td>
<td>.02</td>
<td>77/7844/2.3e-3</td>
<td>16411(1297)/6481/2.5e-3</td>
</tr>
<tr>
<td>10000</td>
<td>9000</td>
<td>.01</td>
<td>.02</td>
<td>63/8336/1.0e-2</td>
<td>16368(1264)/8593/8.7e-3</td>
</tr>
</tbody>
</table>

- 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.
  $err_{ap} := \max_i \text{accurate. pos.} \| x_i - \tilde{x}_i \|$.
900 sensors, 100 anchors, $rr = 0.06$, $\sigma = 0.01$, solve $\rho$ESDP by LPCGD method. $x_i^{true}$ (shown as *) and $x_i^{\rho,\delta}$ (shown as ●) are joined by blue line segment; anchors are shown as ○.
60 sensors, 4 anchors at corners, $rr = 0.3$, $\sigma = 0.1$. $x^\text{true}_i$ (shown as *) and $x^{\rho,\delta}_i$ (shown as ◦) are joined by blue line segment; anchors are shown as o. **Left:** Soln of $\rho$ESDP found by LPCGD method. **Right:** After local gradient improvement.
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).

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

- SDP, ESDP, $\rho$ESDP solns can be further refined by local improvement. This improves the rmsd when noise level is high (10%).

- Approximation bounds? Extension to maxmin dispersion problem?

Thanks for coming! 😊