A multi-Gaussian component EDA with restarting applied to direction of arrival tracking

André R. Gonçalves, Levy Boccato, Romis Attux, Fernando J. Von Zuben
University of Campinas (Unicamp)

IEEE CEC 2013
Cancún, Mexico

June 22, 2013
Outline

1. Introduction
2. DOA estimation
3. EDAs
4. Experiments and Results
5. Conclusions and Future works
6. References
Introduction and Motivation

Sound Source Location

Telecommunication systems

Underwater object tracking
Figure: Uniform linear array with $N$ sensors.
Signal model:

\[ y_k = A(\Phi)x_k + \eta_k, \quad k = 1, 2, \ldots, K. \]  (1)

- \( y_k \in \mathbb{C}^{N \times 1} \) is the noisy signal vector received in the \( k \)-th snapshot;
- \( x_k \in \mathbb{C}^{M \times 1} \) contains the unknown transmitted signals;
- \( \eta_k \in \mathbb{C}^{N \times 1} \) is the noise vector, \( K \) is the number of snapshots;
- \( A = [a_1 \ldots a_M] \in \mathbb{C}^{N \times M} \) is the array transfer matrix, with \( a_m = [1 e^{j\omega_m} \ldots e^{j(N-1)\omega_m}]^T \) and \( \omega_m = \pi \sin(\phi_m), \ m = 1, \ldots, M. \)
Data model

- Most of the proposed methods operate on the sample covariance matrix of the observed sensor outputs;

- Signal and noise are modeled as independent zero-mean complex stationary Gaussian random processes;
Taking into account the statistical hypotheses and the model assumption, the maximum likelihood estimate of $\phi$ is obtained by solving [Kay, 1993, Stoica and Sharman, 1990b]:

$$\hat{\phi}_{ML} = \arg \max_{\phi} J_{ML}(\phi) = \arg \min_{\phi} \text{Tr} \left\{ P_{A}^{\dagger} \hat{R}_{y} \right\},$$  \hspace{1cm} (2)

- $P_{A}^{\dagger} = I - AA^{\dagger}$ is the projection matrix related to the noise subspace.
- $\hat{R}_{y}$ is the sample covariance matrix.

Non-linear, non-quadratic and multimodal function.
Examples of cost function: 2 sources and 10 sensors

Figure: SNR = 0.

Figure: SNR = -10.
Traditional DOA estimation techniques

- **Static DOA estimation:**
  - MODE [Stoica and Sharman, 1990a], MODEX [Gershman and Stoica, 1999] (ML-based algorithms);
  - GA [Karamalis et al., 2001], PSO [Hung, 2013], DE and AIS [Boccato et al., 2012a].

- **Dynamic DOA estimation:**
  - ESPRIT algorithm [Liu et al., 1994];
  - Neural Networks [Badidi and Radouane, 2000];
  - Dopt-aiNet [Boccato et al., 2009].
EDAs are evolutionary algorithms which employ probabilistic models rather than genetic operators to conduct the search.

EDAs evolve a probability distribution on the space of solutions.
EDA framework

1. Start
2. Generate initial population
3. Sample new individuals
4. Estimate the probabilistic model
5. Evaluate population
6. Select the best individuals
7. Stop?
   - Yes: End
   - No: Go to step 3
MGcEDA: The proposed algorithm

- General purpose dynamic optimization algorithm;
- Multipopulation Gaussian-based EDA;
- Explicit diversity maintenance procedures:
  - Mutation operator;
  - Controlled restarting procedure.
MGcEDA for dynamic environments

**Flowchart:**
1. **Start**
2. **Generate Gaussian parameters** ($\mu, \Sigma$)
3. **Stop?**
   - **Yes** → **End**
   - **No** → **Apply mutation operator**
4. **Apply restarting procedure**
5. **Check overlapping components**
6. **Draw $n$ samples from Gaussian pdf**
7. **Evaluate and sort fitness**
8. **Select determ., half of the best samples**
9. **Estimate ($\mu, \Sigma$) using shrinkage**
MGcEDA: Parameters estimation

1. Start
2. Generate Gaussian parameters ($\mu, \Sigma$)
3. Stop?
   - Yes: End
   - No: Apply mutation operator
4. Apply mutation operator
5. Apply restarting procedure
6. Check overlapping components
7. Draw $n$ samples from Gaussian pdf
8. Evaluate and sort fitness
9. Select best $n/2$ samples
10. Estimate ($\mu, \Sigma$) using shrinkage
MGcEDA: Parameters estimation

- Mean estimation:

\[ \mu^{(g+1)} = \sum_{i=1}^{n} w_i \cdot x_i^{(g+1)} \]  

\[ w_i = \frac{1/i}{\sum_{i=1}^{M} w_i}, \quad w_1 \geq w_2 \geq \ldots \geq w_n > 0. \]  

![Graph showing ranking vs. weight](image-url)
Covariance matrix estimation

- How to get a good covariance matrix estimate in small sample size cases?

\[
\hat{\Sigma} = (1 - \hat{\rho})\hat{S} + \hat{\rho}\hat{F} \quad (\text{Shrinkage estimator})
\]

\[
\hat{S}^{(g+1)} = \sum_{i=1}^{n} w_i \left( x_{i}^{(g+1)} - \mu^{(g)} \right) \left( x_{i}^{(g+1)} - \mu^{(g)} \right)^T
\]

\[
\hat{F} = \left( \text{Tr}(\hat{S})/p \right) \cdot I
\]

\[
\hat{\rho}_{RBLW} = \min \left( \frac{(n-2)}{n} \cdot \text{Tr}(\hat{S}^2) + \text{Tr}^2(\hat{S}) \right) \left[ \text{Tr}(\hat{S}^2) - \frac{\text{Tr}^2(\hat{S})}{p} \right], 1 \right)
\]  

[Chen et al., 2010]
Covariance matrix estimation

- How to get a good covariance matrix estimate in small sample size cases?

\[ \hat{\Sigma} = (1 - \hat{\rho})\hat{S} + \hat{\rho} \hat{F} \quad (Shrinkage \ estimator) \]  \hspace{1cm} (5)

\[ \hat{S}^{(g+1)} = \sum_{i=1}^{n} w_i \left( x_i^{(g+1)} - \mu^{(g)} \right) \left( x_i^{(g+1)} - \mu^{(g)} \right)^T \]  \hspace{1cm} (6)

\[ \hat{F} = (\text{Tr}(\hat{S})/p) \cdot I \]  \hspace{1cm} (7)

\[ \hat{\rho}_{\text{RBLW}} = \min \left( \frac{(n-2) \cdot \text{Tr}(\hat{S}^2) + \text{Tr}^2(\hat{S})}{n} \left[ \text{Tr}(\hat{S}^2) - \frac{\text{Tr}^2(\hat{S})}{p} \right], 1 \right) \]  \hspace{1cm} [Chen et al., 2010] \hspace{1cm} (8)
MGcEDA: overlapping detection

- Start
- Generate Gaussian parameters ($\mu$, $\Sigma$)
- Stop?
  - Yes: End
  - No:
    - Apply mutation operator
    - Apply restarting procedure
    - Check overlapping components
    - Draw $n$ samples from Gaussian pdf
    - Evaluate and sort fitness
    - Select determined half of the best samples
    - Estimate ($\mu$, $\Sigma$) using shrinkage
MGcEDA: Overlapping detection

- Overlapping detection ratio (ODR):

\[
ODR(\mu_1, \Sigma_1, \mu_2, \Sigma_2) = \frac{P(\mu_1 | \mu_2, \Sigma_2)}{P(\mu_2 | \mu_2, \Sigma_2)}
\]

(9)

- If \( ODR > \delta_{\text{overlap}} \), restart the component with smaller average fitness.
MGcEDA: Diversity loss and restarting

- Once lost diversity, it’s difficult to explore new regions;
- Check eigenvalues of the covariance matrix;
- Refine solutions or restart component?
  - Components need to exchange information about their fitness values.

Relative Performance:

\[ RP_i = \frac{1}{1 + (f_i^* - f^*)} \]  

(10)
MGcEDA: Mutation

Start

Generate Gaussian parameters (µ, Σ)

Stop?

Yes → End

No →

Draw n samples from Gaussian pdf

Evaluate and sort fitness

Select determ. half of the best samples

Estimate (µ, Σ) using shrinkage

Apply mutation operator

Apply restarting procedure

Check overlapping components

Start

Stop?

Yes → End

No →

Draw n samples from Gaussian pdf

Evaluate and sort fitness

Select determ. half of the best samples

Estimate (µ, Σ) using shrinkage

Apply mutation operator

Apply restarting procedure

Check overlapping components

Start
MGcEDA: Mutation

- Mutation in EDA is an efficient way to keep diversity in the population (fine perturbations);
- Uniform covariance matrix scaling, $\Sigma_{\text{new}} = c \cdot \Sigma$;
- Operator is applied with probability $P_{\text{mut}}$ and $c \in [1, 2]$. 
Experimental design

- Starting from a uniformly generated set of angles over the search space, $[-90^\circ, 90^\circ]$, the next set of angles are obtained by:

\[
\phi(t + 1) = \phi(t) + \mathcal{N}(0, \sigma_B^2 I)
\]  

(11)

- Two different scenarios: 2 and 4 signal sources (10 sensors).

- For each scenario:
  - Five SNRs: -10 dB, -5 dB, 0 dB, 5 dB, and 10 dB;
  - Ten values for $\sigma_B^2$: from 20 to 200.

- MGcEDA was compared with dopt-aiNet algorithm;

- Angles change at each 10,000 FEs.
Dopt-aiNet uses the same parameters considered in the original paper [Boccato et al., 2009];

MGcEDA:
- Number of components: 5;
- Total number of individuals: 100;
- Overlapping threshold: 0.9;
- Mutation probability: 0.1.
Performance measures

- Current best fitness curve;

- Offline performance:

\[
f_{\text{offline}} = \frac{1}{T} \sum_{t=1}^{T} f_i^* \tag{12}
\]

\( f_i^* \) is the best solution of all components at time \( t \).
CB fitness ($\sigma_B^2 = 100$): 2 signal sources
Offline performance ($\sigma_B^2 = 100$): 2 signal sources
### Wilcoxon test results: 2 signal sources

<table>
<thead>
<tr>
<th>$\sigma_B^2$</th>
<th>-10 dB</th>
<th>-5 dB</th>
<th>0 dB</th>
<th>5 dB</th>
<th>10 dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>4.55e-01</td>
<td>3.27e-02</td>
<td>2.12e-01</td>
<td>1.11e-03</td>
<td>6.05e-07</td>
</tr>
<tr>
<td>40</td>
<td>8.50e-02</td>
<td>9.07e-03</td>
<td>1.25e-04</td>
<td>1.25e-07</td>
<td>2.03e-09</td>
</tr>
<tr>
<td>60</td>
<td>1.12e-01</td>
<td>1.33e-02</td>
<td>3.81e-07</td>
<td>2.13e-05</td>
<td>1.20e-08</td>
</tr>
<tr>
<td>80</td>
<td>4.84e-02</td>
<td>1.68e-03</td>
<td>6.05e-07</td>
<td>2.19e-08</td>
<td>1.25e-07</td>
</tr>
<tr>
<td>100</td>
<td>1.99e-02</td>
<td>5.97e-05</td>
<td>8.35e-08</td>
<td>1.10e-08</td>
<td>3.96e-08</td>
</tr>
<tr>
<td>120</td>
<td>1.52e-03</td>
<td>1.19e-06</td>
<td>1.25e-07</td>
<td>5.53e-08</td>
<td>7.39e-11</td>
</tr>
<tr>
<td>140</td>
<td>1.25e-04</td>
<td>1.09e-05</td>
<td>1.25e-07</td>
<td>2.57e-07</td>
<td>1.33e-10</td>
</tr>
<tr>
<td>160</td>
<td>1.06e-03</td>
<td>8.35e-08</td>
<td>1.78e-10</td>
<td>9.26e-09</td>
<td>8.99e-11</td>
</tr>
<tr>
<td>180</td>
<td>2.38e-03</td>
<td>2.03e-07</td>
<td>3.96e-08</td>
<td>5.00e-09</td>
<td>9.92e-11</td>
</tr>
<tr>
<td>200</td>
<td>3.18e-03</td>
<td>2.23e-09</td>
<td>4.18e-09</td>
<td>1.09e-10</td>
<td>3.02e-11</td>
</tr>
</tbody>
</table>
CB fitness ($\sigma_B^2 = 100$): 4 signal sources
Offline performance ($\sigma_B^2 = 100$): 4 signal sources
## Wilcoxon rank sum test results

<table>
<thead>
<tr>
<th>( \sigma_B^2 )</th>
<th>-10 dB</th>
<th>-5 dB</th>
<th>0 dB</th>
<th>5 dB</th>
<th>10 dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>1.67e-01</td>
<td>4.21e-02</td>
<td>4.51e-02</td>
<td>2.24e-02</td>
<td>2.15e-02</td>
</tr>
<tr>
<td>40</td>
<td>7.48e-02</td>
<td>5.01e-02</td>
<td>1.27e-02</td>
<td>1.17e-03</td>
<td>9.51e-06</td>
</tr>
<tr>
<td>60</td>
<td>1.84e-02</td>
<td>4.68e-02</td>
<td>1.64e-05</td>
<td>9.03e-04</td>
<td>7.74e-06</td>
</tr>
<tr>
<td>80</td>
<td>8.31e-03</td>
<td>9.07e-03</td>
<td>1.30e-03</td>
<td>6.74e-06</td>
<td>1.49e-06</td>
</tr>
<tr>
<td>100</td>
<td>4.03e-03</td>
<td>1.17e-02</td>
<td>1.44e-03</td>
<td>2.68e-06</td>
<td>4.22e-04</td>
</tr>
<tr>
<td>120</td>
<td>5.83e-03</td>
<td>3.77e-04</td>
<td>9.79e-05</td>
<td>1.11e-04</td>
<td>1.36e-07</td>
</tr>
<tr>
<td>140</td>
<td>6.10e-03</td>
<td>1.38e-02</td>
<td>8.12e-04</td>
<td>1.87e-07</td>
<td>1.02e-05</td>
</tr>
<tr>
<td>160</td>
<td>1.52e-03</td>
<td>5.57e-03</td>
<td>7.70e-04</td>
<td>1.29e-06</td>
<td>8.66e-05</td>
</tr>
<tr>
<td>180</td>
<td>6.55e-04</td>
<td>7.70e-04</td>
<td>1.61e-06</td>
<td>2.13e-05</td>
<td>1.43e-05</td>
</tr>
<tr>
<td>200</td>
<td>5.87e-04</td>
<td>3.99e-04</td>
<td>4.46e-04</td>
<td>5.61e-05</td>
<td>7.60e-07</td>
</tr>
</tbody>
</table>
Conclusions

- MGcEDA was able to quickly respond to signal source changes, using smaller number of FEs then dopt-aiNet;

- This suggest that MGcEDA is a good candidate to cope with highly dynamic optimization problems;

- The proposed algorithm also showed significant degree of robustness;

- MGcEDA has outperformed dopt-aiNet in practically all considered test sets.
Future Works

- Incorporate a priori information concerning the real DOA angles into the EDA probabilistic models by exploring a noise filtering approach [Krummenauer et al., 2010, Boccato et al., 2012b];

- Try other mutation operators such as the non-uniform ones;

- Use alternative probability distributions such as Cauchy’s and q-Gaussian;

- Include metrics defined on the fitness space in restarting control procedure; and

- Components with dynamic population.


References II


Novel eigenanalysis method for direction estimation.

Maximum likelihood methods for direction-of-arrival estimation.
A multi-Gaussian component EDA with restarting applied to direction of arrival tracking

André R. Gonçalves, Levy Boccato, Romis Attux, Fernando J. Von Zuben
University of Campinas (Unicamp)

IEEE CEC 2013
Cancún, Mexico

June 22, 2013