Single-Sensor Localization of Moving Sources Using Diffusion Kernels

Eran Zeitouni

M.Sc. Seminar
Supervisor: Prof. Israel Cohen

June 2021

Outline

- Introduction and Motivation
- Problem Formulation
- Proposed Method
- Results
- Conclusion
Outline

- Introduction and Motivation
  - Application Areas
  - Goal
  - Challenges
  - Main Contribution
  - Approaches
- Problem Formulation
- Proposed Method
- Results
- Conclusion
Application Areas

- Conversation quality improvement.
- Gunfire localization.
- Noise pollution sources.
- Wildlife researches.
Application Areas

- Conversation quality improvement.
- Gunfire localization.
- Noise pollution sources.
- Wildlife researches.
Application Areas

- Conversation quality improvement.
- Gunfire localization.
- Noise pollution sources.
- Wildlife researches.
Application Areas

- Conversation quality improvement.
- Gunfire localization.
- Noise pollution sources.
  - Car’s Cabin.
  - Industrial Noise.
  - Traffic Noise Violators.
- Wildlife researches.
Application Areas

- Conversation quality improvement.
- Gunfire localization.
- Noise pollution sources.
- Wildlife researches.

[Super Senses: The Secret Power of Animals, BBC]  [pinimg.com]
Goal

- Recovery of both location and velocity of a moving acoustic source, using a single sensor.
- Passive sensor = microphone.

Hello!

I DON'T KNOW WHERE YOU ARE
BUT I WILL FIND YOU

(pngitem.com)
Challenges

- How fast can the source move?
- How does the direction affect the performance?
- Accuracy tradeoff: estimated location Vs. estimated velocity?
- What is the ideal frame length?
- What is the ideal number of features?
- How many training observations are needed?
Challenges (cont’d)

- Under which circumstances?
  - SNR
  - Reverberation time
  - Environmental conditions changes

- What is the effect of rapid and random movements on the performance?

- How do we quantify the estimation error and the accuracy for physical quantities of different units?
Main Contribution

- Stationary sources – an additional degree-of-freedom.
- Moving sources – a degree-of-freedom (velocity).
  - Deterministic movement model.
  - Brownian motion model.
- Time variant system.
- Different kernel.
  - Different extension of the eigenvectors.
  - Different extension of the diffusion kernel.
- Different estimation error and accuracy measures.
[Talmon et al., 2011]

- Diffusion maps—data-driven approach.
- Single-sensor.
- Completely stationary setting—both source and sensor.
- Single degree-of-freedom: azimuth angle.
- Source signal: white Gaussian noise (WGN).
- Feature Vector: autocorrelation function.
- Mahalanobis distance-based kernel.
- Supervised.
Introduction and Motivation

Approach

- [Talmon et al., 2011] (cont’d)
  - Monotonic with respect to the angle:
Outline

- Introduction and Motivation
- Problem Formulation
  - Setting and Assumptions
- Proposed Method
- Results
- Conclusion
Setting and Assumptions

- Stationary single-sensor.
- Multipath must be exploited.

[https://resonance-audio.github.io/resonanceaudio/images/concepts/reflections.png]
Setting and Assumptions

- Stationary single-sensor.
- Multipath must be exploited.
- Unknown acoustic impulse response (AIR).
- Time-variant AIR.
- Source signal for both training and test stages: WGN.
- Feature vector: autocorrelation function.
Variability of the AIR in specific enclosures relies on a small number of parameters.
Variability of the AIR in specific enclosures relies on a small number of parameters.

Slow and gradually changing velocity components.

Small changes to the AIR along time frames.

The trajectory, formed by the movement of the source during the frame, is approximated by a linear movement segment.

Each frame is inspected individually for estimation of the location and velocity.

Constant elevation angle.
The velocity of the source is defined as:

\[ \dot{p}(t) = a(p(t)) + n(t) \]

- \( \dot{p}(t) \) – velocity of the source at time \( t \).
- \( p(t) \) – absolute location of the source at time \( t \).
- \( a(p(t)) \) – drift term. \( a : p(t) \rightarrow \mathbb{R}^2 \).
- \( n(t) \) – Brownian motion term.
Problem Formulation

- Each source, one at a time, transmits a signal during its movement.
- The signal, received by the sensor (corresponding to the \(i\)th source):

\[
y_i(n) = H_\theta \{x_i(n)\} = \sum_{j=-\infty}^{\infty} h_{\theta_i(j)}(n, j)x_i(j)
\]

- \(x_i(n)\) – the signal transmitted by the \(i\)th source.
- \(h_{\theta_i(j)}(n, j)\) – AIR between the \(i\)th source and the sensor, with respect to parameters vector \(\theta_i(j)\), where:

\[
\theta_i(j) = \begin{bmatrix} \rho_i(j) \\ \phi_i(j) \\ s_i(j) \\ \beta_i(j) \end{bmatrix}^T
\]

\{ Radius, Azimuth, Speed, Direction \} \rightarrow Relative location, Velocity
The received signal is divided into time frames.

For each time frame, we collect datasets:

- Training dataset generated from $m$ labeled arbitrary source locations and velocities:
  \[
  \overline{\Theta} = \{\overline{\theta}_1(q), \ldots, \overline{\theta}_m(q)\} \subset \mathbb{R}^d
  \]

- Test dataset generated from $M$ unknown arbitrary source locations and velocities:
  \[
  \Theta = \{\theta_{m+1}(q), \ldots, \theta_{m+M}(q)\} \subset \mathbb{R}^d
  \]

- $q$ – query point along the trajectories. Defined as the midway point.
- $d$ – intrinsic dimension. Number of system’s degrees-of-freedom.
Outline

- Introduction and Motivation
- Problem Formulation
- Proposed Method
  - Feature Vector
  - Diffusion Kernel
  - Localization Based on Diffusion Mapping
- Results
- Conclusion
Feature Vector

- Must represent the propagation paths of the signal and the AIR.
- The feature vector is defined based on an autocorrelation function of the received signal:

\[ c_{y_i}(n_1, n_2) = \mathbb{E}[y_i(n_1)y_i(n_2)] = \sum_{j,l=-\infty}^{\infty} h_{\theta_i(j)}(n_1, j) h_{\theta_i(l)}(n_2, l) c_{x_i}(j - l) \]

- Source signal is wide sense stationary (WSS).
- Time-variant system.
Proposed Method | Diffusion Kernel

---

**Feature Vector**

- Source signal is a WGN, hence: \( c_{x_i}(\tau) = \sigma_{x_i}^2 \delta(\tau) \).

- By substitution, we yield:

\[
c_{y_i}(n_1, n_2) = \sigma_{x_i}^2 \sum_{j=-\infty}^{\infty} h_{\theta_i}(j)(n_1, j)h_{\theta_i}(j)(n_2, j)
\]

- Assumptions—recall:
  - Short time intervals.
  - Slow speed and gradually changing direction and speed values (Brownian motion ignored).
We introduce small changes to the AIR along the time frames, thus:

\[ c_{y_i}(\tau) = h_{\theta_i}(\tau) * h_{\theta_i}(-\tau) * c_{x_i}(\tau) = \]

\[ = \sigma_{x_i}^2 h_{\theta_i}(\tau) * h_{\theta_i}(-\tau) \]

The time differences of \( c_{y_i} \) depend purely on the variations of the AIR \( \rightarrow \) on the evolution of \( \theta_i \)!
For each received signal, we generate a feature vector of $D$ elements of the autocorrelation function:

$$c_{yi}^{(j)} = c_{yi}(n, n + j) = \mathbb{E}[y_i(n)y_i(n + j)] , j = 0, ..., D - 1$$

Tradeoff considerations of Feature vector’s length, $D$.

Let $\Gamma = \{c_i\}_{i=1}^M$ denote the feature vectors with respect to the unlabeled parameters in $\Theta$.

Let $\overline{\Gamma} = \{\overline{c}_i\}_{i=1}^m$ denote the feature vectors with respect to the labeled parameters in $\overline{\Theta}$. 

Proposed Method: Diffusion Kernel
Feature vectors are represented as points in a high-dimensional space.
Proposed Method: Diffusion Kernel

Choice of an Affinity Measure

- Feature vectors are represented as points in a high-dimensional space.

- They are spread on a low-dimensional nonlinear manifold $\mathcal{M}$. 
Feature vectors are represented as points in a high-dimensional space.

- They are spread on a low-dimensional nonlinear manifold $\mathcal{M}$.
- Linearity is maintained in the vicinity of each one.
Feature vectors are represented as points in a high-dimensional space.

- They are spread on a low-dimensional nonlinear manifold $\mathcal{M}$.
- Linearity is maintained in the vicinity of each one.
- Distances between the feature vectors must be measured along the manifold.
Why distances must be measured along the manifold and not by the Euclidean distance?

- **Euclidean distance**: \( D_{\text{EUC}}(c_i, c_j) = \|c_i - c_j\| \).
  - Measures affinity between observations in their original space.
  - No assumption regarding an existence of a manifold.

[Tenenbaum et al., 2000]
Why distances must be measured along the manifold and not by the Euclidean distance?

- **Euclidean distance:** \( D_{\text{EUC}}(c_i, c_j) = \| c_i - c_j \| \).
  - Measures affinity between observations in their original space.
  - No assumption regarding an existence of a manifold.
  - Reliable for flat/linear manifolds, or when observations are scattered uniformly all over the space.
How can we measure distance along the manifold?

- Geodesic distance is the locally shortest path along the manifold, but the manifold structure is unknown.
- The Geodesic distance can be approximated by the diffusion distance.
[Coifman and Lafon, 2006].

The manifold can be referred as a graph:
- The feature vectors are graph nodes/vertices.
- The weights of the edges are defined according to a kernel function:

\[ W^{(ij)} = w(\mathbf{c}_i, \mathbf{c}_j) = \exp \left\{ -\frac{(D_{\text{kernel}}^{(ij)})^2}{\varepsilon} \right\} \]
Diffusion Maps

- Two common choices for kernel’s distance measure:
  - Mahalanobis distance-based
  - Euclidean distance-based
Proposed Method: Diffusion Kernel

Diffusion Maps

- **Mahalanobis distance-based kernel**
  - [Talmon et al., 2011; Talmon et al., 2011; Kushnir et al., 2012; Laufer-Goldshtein et al., 2013].
  - Motivation: direct relation between the observations and the latent parameters:
    \[
    \|\bar{\theta}_k - \bar{\theta}_l\|^2 \approx (\bar{c}_k - \bar{c}_l)^T [\Sigma_k + \Sigma_l]^{-1} (\bar{c}_k - \bar{c}_l)
    \]
  - Definition:
    \[
    W^{(kl)} = \frac{\pi}{d_{kl}} \exp \left\{ -\frac{(\bar{c}_k - \bar{c}_l)^T \left[ \hat{\Sigma}_k + \hat{\Sigma}_l \right]^{-1} (\bar{c}_k - \bar{c}_l)}{\varepsilon} \right\}
    \]
  - Practical drawbacks:
    - Singular local covariance matrices.
    - Additional observations shall be generated for each training observation.

\[
\hat{\Sigma}_i = \frac{1}{L} \sum_{j=1}^{L} (c_{i,j} - \mu_i) (c_{i,j} - \mu_i)^T
\]
Diffusion Maps

- Euclidean distance-based kernel
  - [Laufer–Goldshtein et al., 2016].
  - Reliable for small neighborhoods.
  - Meaningless for larger distances.
  - Definition:

\[
W^{(i,j)} = \begin{cases} 
\exp \left\{ - \frac{||c_i - c_j||^2}{\varepsilon} \right\} , & \text{if } c_i \in \mathcal{N}_j \text{ or } c_j \in \mathcal{N}_i \\
0 , & \text{otherwise}
\end{cases}
\]
By normalizing the affinity matrix $W$, we obtain the transition matrix:

$$P^{(ij)} \equiv P_r(c_j \mid c_i) = \left( \sum_{j=1}^{m} W^{(ij)} \right)^{-1} W^{(ij)}$$

- Defines a Markov process—a discrete diffusion process over the data.
By eigenvalue decomposition of $P$, we obtain the eigenvalues $\{\lambda_j\}_{j=0}^{m-1}$ and the eigenvectors $\{\psi_j\}_{j=0}^{m-1}$.

Let $\Psi_d$ be the diffusion mapping of the observations into a low-dimensional Euclidean space $\mathbb{R}^d$:

$$\Psi_d : \bar{c}_i \rightarrow \left[ \lambda_1 \psi_1^{(i)}, \ldots, \lambda_d \psi_d^{(i)} \right]^T$$

- A parameterization of the manifold.
- Reveal of the latent parameters of the observations.
Diffusion Maps

- Diffusion distance:
  - Recall– the diffusion distance approximates the distance along the manifold.
  - The diffusion distance can be approximated by the Euclidean distance in the embedded space:

\[
D_{DIFF} (\bar{c}_i, \bar{c}_j) \approx \| \Psi_d (\bar{c}_i) - \Psi_d (\bar{c}_j) \|
\]
So far, we discussed mapping of the training observations.

Given a new set of $M$ observations of unknown locations and velocities, we seek to embed them as well in the low-dimensional space.

- Avoiding another spectral decomposition.
Nyström method:

- We add $M$ new rows to the affinity matrix $W$:

$$W^{(\tilde{i},j)} = \begin{cases} \exp\left\{-\frac{||c_i - \overline{c}_j||^2}{\varepsilon}\right\}, & \text{if } \overline{c}_j \in \mathcal{N}_{\tilde{i}} \\ 0, & \text{otherwise} \end{cases},$$

where $\tilde{i} = m + i$.

- Accordingly, the new entries of the transition matrix $P$ are given by:

$$P^{(\tilde{i},j)} = \left(\sum_{j=1}^{m} W^{(\tilde{i},j)}\right)^{-1} W^{(\tilde{i},j)}.$$
Nyström method (cont’d):

- The new entries of the extended eigenvectors:

\[
\psi_l^{(\tilde{i})} = \frac{1}{\lambda_l} \sum_{j=1}^{m} P^{(\tilde{i},j)} \psi_l^{(j)}
\]

- Thus:

\[
\Psi_d : c_i \rightarrow \left[ \lambda_1 \psi_1^{(\tilde{i})}, \ldots, \lambda_d \psi_d^{(\tilde{i})} \right]^T
\]

- Solves extension issues of [Laufer–Goldshtein et al., 2016].
Recovery of Controlling Parameters

- Estimation of the location and velocity of the new observations:

\[ \hat{\theta}_i(q) = \sum_{j: \Psi_d(\bar{c}_j) \in \tilde{N}_i} \gamma_j(c_i) \bar{\theta}_j(q) \]

- The interpolation coefficients:

\[ \gamma_j(c_i) = \frac{\exp \left( - \frac{||\Psi_d(c_i) - \Psi_d(\bar{c}_j)||^2}{\varepsilon_{\gamma_i}} \right)}{\sum_{l: \Psi_d(\bar{c}_l) \in \tilde{N}_i} \exp \left( - \frac{||\Psi_d(c_i) - \Psi_d(\bar{c}_l)||^2}{\varepsilon_{\gamma_i}} \right)} \]

- Satisfying \( \sum_{j=1}^{\tilde{k}} \gamma_j(c_i) = 1. \)
Estimation error measures, suggested in former works, must be modified—

- Various physical quantities—different units.
- No longer a scalar.
- Canceling out the units of the individual estimation errors.

The normalized estimation error:

\[ e(c_i) = \left[ e_i^{(1)}, \ldots, e_i^{(d)} \right] \]

- A vector of length \(d\).
- Its \(j\)th element is given by:

\[
e_i^{(j)} = \frac{\theta^{(j)}_i(q) - \hat{\theta}^{(j)}_i(q)}{|\theta^{(j)}_i(q)|}
\]
The accuracy is measured by:

\[
RMSE = \sum_{j=1}^{d} \alpha_j \sqrt{\frac{1}{M} \sum_{i=1}^{M} \left( e_i^{(j)} \right)^2},
\]

- A linear combination of the root mean square error (RMSE) of each one of the \( d \) elements of the error vector.
- \( \{\alpha_j\} \) – the coefficients which represent the weight of each physical quantity (For simplicity, we define as \( \{\alpha_j\} = \frac{1}{d} \)).
Outline

- Introduction and Motivation
- Problem Formulation
- Proposed Method
- Results
  - Simulated Setup
  - Stationary Sources
  - Deterministically Moving Sources
  - Brownian Motion Induced Moving Sources
- Conclusion
All simulations rely on an efficient implementation [Habets, 2006] of the image method [Allen and Berkley, 1979].

Omnidirectional microphone.

In all experiments:
- Room dimensions: $6 \times 5.8 \times 3$ [m$^3$].
- Microphone location: $(3, 1, 1.8)$ m.
- Source signal: a zero-mean and unit-variance WGN.
- $f_s = 16$ kHz.

Unless otherwise noted:
- $T_{60} = 0.3$ sec.
- Source signal duration: 1 sec.
- $D = 800$ lags.
- $d = 1$. 
One-Dimensional Case

- Former single-sensor algorithm [Talmon et al., 2011]– recall:
  - Mahalanobis distance–based kernel.
  - One-dimensional stationary scenario.

- Objectives:
  - Comparison to literature results.

- Simulation setup:
  - Sources init.:
    - Radius: 1m from the microphone.
    - Azimuth angles: $U[25, 64]°$.
  - $M = m = 720$ observations.
Results

Stationary Sources

One-Dimensional Case

- **Euclidean distance**— monotonic with respect to the azimuth in a small region only—unsuitable for the job.
  - However, it is adequate in the vicinity of each observation—useful for our kernel.
- **Diffusion distance**— monotonic with respect to the azimuth for the entire range.
Results

Stationary Sources

One-Dimensional Case

- **Euclidean distance**—monotonic with respect to the azimuth in a small region only—unsuitable for the job.
  - However, it is adequate in the vicinity of each observation—useful for our kernel.
- **Diffusion distance**—monotonic with respect to the azimuth for the entire range.
- Diffusion mapping follows the azimuth successfully—latent parameter can be recovered.
Results - Stationary Sources

One-Dimensional Case

- **Euclidean distance**— monotonic with respect to the azimuth in a **small region only**— unsuitable for the job.
  - However, it is adequate in the vicinity of each observation— useful for our kernel.
- **Diffusion distance**— monotonic with respect to the azimuth for the **entire range**.
- Diffusion mapping follows the azimuth successfully— latent parameter can be recovered.
- Comparison to [Talmon et al., 2011]:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Ours</th>
<th>[Talmon et al., 2011]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diffusion Kernel Type</td>
<td>Euclidean distance–based</td>
<td>Mahalanobis distance–based</td>
</tr>
<tr>
<td>Training Set Size</td>
<td>720</td>
<td>420 (5220, including local observations)</td>
</tr>
<tr>
<td>Angles Range [°]</td>
<td>39</td>
<td>60</td>
</tr>
<tr>
<td>Azimuth Angles Allocation</td>
<td>Uniform distribution over the entire range</td>
<td>60 Predefined angles, 1° apart.</td>
</tr>
<tr>
<td>RMSE [°] ; ( e_i =</td>
<td>\theta_i - \hat{\theta}_i</td>
<td>)</td>
</tr>
</tbody>
</table>

- Euclidean distance— monotonic with respect to the azimuth in a **small region only**— unsuitable for the job.
  - However, it is adequate in the vicinity of each observation— useful for our kernel.
- Diffusion distance— monotonic with respect to the azimuth for the **entire range**.
- Diffusion mapping follows the azimuth successfully— latent parameter can be recovered.
- Comparison to [Talmon et al., 2011]:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Ours</th>
<th>[Talmon et al., 2011]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diffusion Kernel Type</td>
<td>Euclidean distance–based</td>
<td>Mahalanobis distance–based</td>
</tr>
<tr>
<td>Training Set Size</td>
<td>720</td>
<td>420 (5220, including local observations)</td>
</tr>
<tr>
<td>Angles Range [°]</td>
<td>39</td>
<td>60</td>
</tr>
<tr>
<td>Azimuth Angles Allocation</td>
<td>Uniform distribution over the entire range</td>
<td>60 Predefined angles, 1° apart.</td>
</tr>
<tr>
<td>RMSE [°] ; ( e_i =</td>
<td>\theta_i - \hat{\theta}_i</td>
<td>)</td>
</tr>
</tbody>
</table>
Two–Dimensional Case

- Examining the feasibility to organize observations monotonically—two latent parameters of different units.

- Simulation setup:
  - Sources init.:
    - Radius: $U[1,1.3]$m from the microphone.
    - Azimuth angles: $U[25,64]$°.
  - $d = 2$.
  - $M = m = 2880$ observations.
Two-Dimensional Case

- Diffusion mapping perceives both the radius and the azimuth successfully—thus they can be recovered.

(a) Color-coded according to radius values

(b) Color-coded according to azimuth values
**Results**

**Stationary Sources**

# Two-Dimensional Case

- Diffusion mapping perceives both the radius and the azimuth successfully—thus they can be recovered.
- RMSE of 1.05%—consists of:
  - 0.4% for radius,
  - 1.7% for azimuth.
- Visual demonstration of the localization performance.
The Experiments

- **Objective:** recovery of location and velocity of deterministically moving sources.

- The performance of our algorithm is examined by various experiments:

  - **Hyperparameters:**
    - Training set size
    - Number of features
    - Frame length

  - **Variables:**
    - Speed
    - Direction

  - **Conditions:**
    - Signal to noise ratio (SNR)
    - Reverberation time
    - Environmental conditions changes

- For isolating the different factors affecting the performance—$d = 1$ in all scenarios.

- The single controlling parameter dictates variations of another two parameters along the movement of the sources.

- For simplicity, the sources move linearly.
Results

Deterministically Moving Sources

Sensitivity to Training Set Size

- **Simulation setup:**
  - **Sources init.:**
    - Radius: 1 m from the microphone.
    - Azimuth: 45°.
    - Speed: 0.5 m/sec.
  - $M = m$, varies from 120 to 2880 observations.
Bigger training set size = better accuracy.
Results

Deterministically Moving Sources

Sensitivity to Number of Features

- Simulation setup:
  - Sources init.:
    - Radius: 1m from the microphone.
    - Azimuth: 45°.
    - Speed: 0.5 m/sec.
  - $M = m = 720$ observations.
  - $D$: 50 to 5600 lags.
Sensitivity to Number of Features

- Curse of dimensionality—Peaking phenomenon.
- Small number of features—bad resolution.
- High number of features—insufficient number of observations.
- Optimal results for $D=800$. 

![Graph showing RMSE vs. number of feature-vector's dimensions]
**Simulation setup:**

- **Sources init.:**
  - Radius: $1$ m from the microphone.
  - Azimuth: $45^\circ$.
  - Speed: $0.0625$ to $1$ m/sec (different value for each simulation).
  - Direction: $U[45,85]^\circ$.

- $M = m = 720$ observations.
Sensitivity to Speed

- Total estimation error – dictated by the direction.
Sensitivity to Speed

- **Total estimation error**—dictated by the **direction**.
  - **Very slow sources**—struggle of perceiving variations between different directions during a bare movement.

---

Results

**Deterministically Moving Sources**

- (a) Embedding for 0.0625 m/sec
- (b) Embedding for 0.5 m/sec

![Graphs and data points showing sensitivity to speed](image-url)

**Graph Legend:****

- Total
- Radius
- Azimuth
- Direction

**Axes:**
- **X-axis:** Direction [°]
- **Y-axis:** RMSE

**Speed [m/sec]:**

- 0.1
- 0.2
- 0.3
- 0.4
- 0.5
- 0.6
- 0.7
- 0.8
- 0.9
- 1
Sensitivity to Speed

- **Total estimation error**—dictated by the **direction**.
  - Very slow sources—struggle of perceiving variations between different directions during a bare movement.
  - **Faster sources**—improvement due to a more noticeable movement by the sources.
Total estimation error—dictated by the direction.

- Very slow sources—struggle of perceiving variations between different directions during a bare movement.
- Faster sources—improvement due to a more noticeable movement by the sources.
- Fast sources—decay due to two factors:
  - Greater speeds = sparser manifolds.
  - The quasi-stationarity assumption is revoked.
Sensitivity to Speed

- **Radius & azimuth**–
  Faster speed ⇒ a wider range of possible location values.
- The results demonstrate:
  - Tradeoff between the accuracy of the estimated location and the estimated direction.
  - The influence of conflicting forces:
    - The validity of our assumptions regrading velocity attributes.
    - The time variance property of the system.
Results

Deterministically Moving Sources

Sensitivity to Direction

Simulation setup:

- Sources init.:
  - Radius: 1m from the microphone.
  - Azimuth: 45°.
  - Speed: $U[0.25, 0.5]$ m/sec
  - Direction: 5° to 90° (different value for each simulation).
- $M = m = 720$ observations.
Sensitivity to Direction

- Total estimation error & radius estimation error – dictated by the speed.

![Graph showing sensitivity to direction with RMSE on the y-axis and direction in degrees on the x-axis. The graph includes lines for total, radius, azimuth, and speed.]
Sensitivity to Direction

- **Total estimation error** & **radius estimation error**—dictated by the **speed**.
  - **Speed estimation error**—ruled by a varying extent of spatial aliasing.

![Graphs](a) Original affinity matrix–spatially aliased case
(b) Original affinity matrix–non-spatially aliased case

---

**Results** Deterministically Moving Sources

Eran Zeitouni  Single-Sensor Localization of Moving Sources  June 2021
Sensitivity to Direction

- Total estimation error & radius estimation error—dictated by the speed.
  - Speed estimation error—ruled by a varying extent of spatial aliasing.
Sensitivity to Direction

- Total estimation error & radius estimation error – dictated by the speed.
  - Speed estimation error – ruled by a varying extent of spatial aliasing.
- Azimuth estimation error – wider direction with respect to 45° (the degenerated scenario) ⇒ a wider range of possible azimuth values.
**Results**  Deterministically Moving Sources

## Sensitivity to SNR

- **Simulation setup:**
  - Additive WGN is introduced to the received signal.
  - Sources init.:
    - Radius: 1m from the microphone.
    - Azimuth: 45°.
    - Speed: 0.5 m/sec.
    - Direction: $U[45, 85]^\circ$.
  - SNR: 0 to 30dB, followed by a scenario free of noise (different value for each simulation).
  - $M = m = 720$ observations.
Results

Deterministically Moving Sources

Sensitivity to SNR

- Harsher conditions $\Rightarrow$ higher total estimation error.
- Significantly high estimation error for SNR of 0 dB.
- From SNR of 20 dB onward– the accuracy improves insignificantly.
Sensitivity to Reverberation Time

- **Simulation setup:**
  - **Sources init.**:
    - Radius: **1m** from the microphone.
    - Azimuth: **45°**.
    - Speed: **0.5 m/sec**.
    - Direction: \( U[45,85]^\circ \).
  - Reverberation time: **0.128 to 1 sec** (different value for each simulation).
  - \( M = m = 720 \) observations.
Sensitivity to Reverberation Time

- Nearly no reverberation—colossal estimation error.
- Low to moderate reverberation time—gradual improvement.
- Moderate to high reverberation time—steady decay.
- Validation of the implied hypothesis of our work—The reflections are vital for accurate localization using a sole sensor.

![Graph showing the relationship between RMSE and Reverberation Time.]
Results

Sensitivity to Environmental Conditions Changes

- Training set according to reverberation time of 0.4 sec.
- Test set according to varying reverberation time (in each simulation).
- Bigger difference of the reverberation levels between the stages ⇒ larger estimation error.
- The error deviates from the values achieved ideally– unless the reflections do not play a role in practice (e.g., as at 0.128 sec).

![Graph showing RMSE vs. Reverberation Time]
Sensitivity to Frame Length

Simulation setup:

○ Source signal duration: 2 sec.
○ Sources init.:
  • Radius: 1m from the microphone.
  • Azimuth: \( U[45, 85]^{\circ} \).
  • Speed: 0.5 m/sec.
  • Direction: 45°.
○ Frame length:
  2 sec to 0.1 sec (different value for each simulation)
○ \( M = m = 720 \) observations.
Results

Deterministically Moving Sources

Sensitivity to Frame Length

- Simulation setup:
  - Source signal duration: 2 sec.
  - Sources init.:
    - Radius: 1 m from the microphone.
    - Azimuth: $U[45, 85]^\circ$.
    - Speed: 0.5 m/sec.
    - Direction: 45°.
  - Frame length: 2 sec to 0.1 sec (different value for each simulation)
  - $M = m = 720$ observations.

The average total estimation error:

$$\overline{\text{RMSE}} = \frac{1}{N} \sum_{r=1}^{N} \text{RMSE}^{(r)}$$
Sensitivity to Frame Length

- Long frames – inaccurate approximation by a linear segment, unless the source has a constant velocity.
  - Evident even in such an ideal case.
- Short frames – unable to capture the movement properly, regardless the trajectory.
- The optimal accuracy is achieved by frames of 0.5 sec.

Results
Deterministically Moving Sources

![Graph showing sensitivity to frame length]
Sensitivity to Frame Length

- Localization results of the azimuth and radius:
  - Optimal frame length.
  - Arbitrary source.
  - Degenerated speed and direction.

(a) Radius

(b) Azimuth
So far, we ignored the Brownian motion term.

Recall the velocity of a source as a function of time:

\[ \dot{p}(t) = a(p(t)) + n(t) \]

- \( \dot{p}(t) \) – velocity of the source at time \( t \).
- \( p(t) \) – absolute location of the source at time \( t \).
- \( a(p(t)) \) – drift term.
- \( n(t) \) – Brownian motion term.

Objective: examination of the impact of violating the fundamental assumption– by sources that change their velocity rapidly and randomly.

- Worst-case scenario– the same algorithm framework.
  For emphasis– the algorithm is executed just once.
The Brownian motion term—

\[ n(t) = \begin{bmatrix} n_{\tilde{x}}(t) \\ n_{\tilde{y}}(t) \end{bmatrix} \]

Described by WGN with the variance vector 

\[ \begin{bmatrix} \sigma_{\tilde{x}}^2 \\ \sigma_{\tilde{y}}^2 \end{bmatrix} \]

Under the constraint of a four-sigma confidence level, 99.99% of the realizations are within the following range:

\[
\begin{cases} 
-\eta \cdot v_{\tilde{x}_{\text{max}}} \leq n_{\tilde{x}}(t) \leq \eta \cdot v_{\tilde{x}_{\text{max}}} \\
-\eta \cdot v_{\tilde{y}_{\text{max}}} \leq n_{\tilde{y}}(t) \leq \eta \cdot v_{\tilde{y}_{\text{max}}} 
\end{cases}
\]

where:

- \( v_{\tilde{x}_{\text{max}}} \) and \( v_{\tilde{y}_{\text{max}}} \) – the maximal horizontal and vertical speed components drawn in the experiment, respectively.
- \( \eta \) – the Brownian motion coefficient.
Simulation Setup

- Sources init.:
  - Radius: **1m** from the microphone.
  - Azimuth: **45°**.
  - Speed: **0.5 m/sec**.
  - Direction: **$U[45,85]°$**.

Results

Brownian Motion Induced Moving Sources

Evolution of locations through time

Brownian motion free scenario
Results Brownian Motion Induced Moving Sources

Simulation Setup

- Sources init.:
  - Radius: 1m from the microphone.
  - Azimuth: 45°.
  - Speed: 0.5 m/sec.
  - Direction: $U[45, 85]^\circ$.

- Brownian motion coefficient: 0 to 0.8 (different value for each simulation).

- $M = m = 720$ observations.

Brownian motion induced source trajectory— a coefficient of 0.8
What is the number of intrinsic dimensions (degrees-of-freedom)?

- $d = 1$ – direction.
- $d = 2$ – direction and speed.
- $d = 3$ – direction, radius and azimuth.

For all cases: Larger Brownian motion term $\Rightarrow$ bigger estimation error.

The estimation error is colossal:

- Brownian motion coefficient of 0.425–an order-of-magnitude higher compared to the ideal case.
- Brownian motion coefficient of 0.8–2400% higher than the ideal case.
Determination of the number of intrinsic dimensions:

- Unambiguity upon the parameters vector.
- The variability by the Brownian motion is not as influential as the variability by the direction.
- No dominating case.
- Negligible difference between the results.

Conclusion:
The additional degrees-of-freedom are redundant $\Rightarrow d = 1.$
0.8 Brownian motion coefficient scenario:

- The small perturbations along the trajectory have minor influence on the location:
  - Radius—minor ambiguities due to the small range of possible values.
  - Azimuth—followed successfully.

(a) Radius

(b) Azimuth
0.8 Brownian motion coefficient scenario (cont’d):

- The instantaneous velocity is heavily affected by the Brownian motion and the significant variability of the direction:
  - Failure in maintaining monotonic organization of the observations—the instantaneous speed and direction cannot be accurately recovered.
0.8 Brownian motion coefficient scenario:

- Individual RMSE values:
  - Radius – 0.2%
  - Azimuth – 0.3%
  - Speed – 24%
  - Direction – 25%

- The accuracy of the estimated trajectories correspond well with the embeddings.
- Failure in recovering the **instantaneous** velocity through time.
- Fair success of estimating the **average** velocity through time.
Localization Results

- **0.8 Brownian motion coefficient scenario (cont’d):**
  - An arbitrary source:
    - Speed – 0.5 m/sec.
    - Direction – 52° (drawn).
  - Localization results through time:
    - LM – Linear Model.
    - WA – Weighted Average of the trajectories.
    - MV – Mean Velocity.
Outline

- Introduction and Motivation
- Problem Formulation
- Proposed Method
- Results

- Conclusion
  - Research Summary
  - Future Work
SOTA localization results for moving and stationary sources.

Data-driven approach– manifold learning.

Time-variant system.

Approximation of the trajectory of the source, formed during the frame, by a linear movement segment.

Tradeoff between the accuracy of the estimated location and the estimated direction.
Research Summary

- Struggle to distinguish between speed values of sources that move at the same direction.
- The infamous multipath is vital for accurate single-sensor localization.
- Robustness to reverberant and noisy environments—yet sensitivity to environmental changes.
- Successful localization of slow sources that change their velocity gradually.
- Successful estimation of the location and the average velocity of sources that change their velocity rapidly and randomly, but their average speed is slow.
Future Work

- Four degrees-of-freedom.
- More sophisticated movement models.
- More advanced manifold learning methods—semi-supervised manifold regularization towards tracking approach.
- Moving Sensor.
- Significant environmental changes.
- Localization of radio-frequency (RF) emitters.
- Diarization—multiple sources simultaneously.
Thanks

- Prof. Israel Cohen
- SIPL team
- Friends
- Family