Single-Sensor Localization of Moving Sources Using Diffusion Kernels and Brownian Motion Model

Eran Zeitouni and Israel Cohen
Andrew and Erna Viterbi Faculty of Electrical & Computer Engineering
Technion–Israel Institute of Technology, Technion City, Haifa 3200003, Israel
eranze@campus.technion.ac.il; icohen@ee.technion.ac.il

***Abstract*—Recently, we introduced a single-sensor method for estimating the location and velocity of moving sources. The obtained state-of-the-art results are valid to slow sources whose velocities are gradually changing through time. In this paper, we challenge our algorithm's fundamental assumption. We apply the algorithm to sources that have rapid and random fluctuations in their velocity. The proposed algorithm is based on a supervised learning approach, using diffusion maps with a Euclidean distance-based diffusion kernel. Experimental results demonstrate the benefits of the proposed single-sensor localization method for sources with a Brownian motion model of randomly fluctuating movements.**

***Index Terms*—Source localization, direction finding, single-sensor, single-site, manifold learning, diffusion maps, passive sensing, position finding, non-cooperative localization.**

I. Introduction

The task of source localization has attracted research endeavors over the past decades. In classic and modern approaches, the location of a source is determined by a triangulation process. This process is based on variations of physical attributes between signals acquired by spatial arrays. Advanced spatial array processing approaches consist of maximum likelihood (ML) based beamformers, and subspace-based methods. While the former involves parameter estimation using the received signals [1], [2], the latter relies on measurements-based singular value decomposition (SVD) [3]–[11]. However, when only a single sensor is available, none of these approaches is capable of estimating the location of the source. A method for single-sensor source localization based on diffusion kernels was presented by Talmon et al. [12]. Using a training set, this data-driven approach aims at learning the nonlinear structure of the manifold of the data. Unfortunately, this method addresses only the localization of static sources.

Recently, we introduced in [13] an extension of [12] for moving sources, where both their locations and velocities are recovered by a single sensor. Unlike [12], which relies on a specially-tailored distance-based kernel that associates locations with observations, our kernel is Euclidean distance-based and more practical. Following an iterative process, the trajectory formed by the movement of the source during each time frame is approximated by a short linear segment. Accordingly, each frame is inspected individually for estimating the location and velocity. This approach's premise of recovering the location and velocity values of the sources relies on its capability to organize the data on the manifold according to these values. This principle relies on slow sources whose velocity changes gradually so that an unambiguous manifold structure to be observed despite movement. The approach, described in [13], has demonstrated state-of-the-art localization results for moving sources and stationary sources alike.

In this work, we investigate a more challenging setting involving the influence of environmental conditions (e.g., vibrations and wind) on the movement of the sources. This influence is modeled by including a Brownian motion term in the movement of the sources. By inspecting sources whose average speed is slow, but the velocity is rapidly and randomly changing, we examine the impact of violating the fundamental assumption of [13]: gradual changes of the velocity through time. In order to do so, we adopt the algorithm framework of [13] for executing the worst-case scenario. To emphasize the localization results in this worst-case scenario, the algorithm is executed just once (i.e., a single frame). Even when the environmental conditions change dramatically, our algorithm estimates the average velocity accurately despite failing to estimate the instantaneous velocity.

The rest of this paper is organized as follows. In Section II, we formulate the problem. In Section III, we present the computation of the diffusion kernel. In Section IV, we present the proposed algorithm for single-sensor localization for moving sources. In Section V, results of the proposed algorithm are depicted. In Section VI, we present conclusions regarding the performance of the algorithm.

II. Problem Formulation

We consider a standard conference room. The velocity of moving source as a function of time, denoted by $\dot{\mathbf{p}}(t)$, is defined as a combination of a deterministic term and a random term, as follows:

$$\dot{\mathbf{p}}(t) = a(\mathbf{p}(t)) + \mathbf{n}(t), \tag{1}$$

where $\mathbf{p}(t)$ is the absolute location of the source at time t; $a(\mathbf{p}(t))$ is the drift term that describes the deterministic com-

This work was supported by the Pazy Research Foundation and the ISF-NSFC joint research program (grant No. 2514/17).

ponent of the velocity at location $\mathbf{p}(t)$; $a(\cdot)$ is a deterministic function which maps $\mathbf{p}(t)$ to a velocity ; and $\mathbf{n}(t)$ is the Brownian motion term, which represents rapid and random perturbations that affect the movement of the sources as a result of environmental conditions (e.g., enclosure vibrations, vibrations of source's mount, and wind). The Brownian motion term is described by white Gaussian noise (WGN).

One by one, each source transmits a signal during its movement. The source signal is assumed to be a zero-mean wide-sense stationary (WSS) process. As speech is a quasi-stationary process, meaning WSS in short time frames, we can even comply with such a weaker assumption.

The relative location and the velocity of the ith source at discrete time index j are denoted by the parameters vector $\boldsymbol{\theta}_i(j) = [\rho_i(j), \phi_i(j), s_i(j), \beta_i(j)]$, where: $\rho_i(j)$ is the distance (i.e., radius) between the source and the sensor; $\phi_i(j)$ is the azimuth angle; $s_i(j)$ is the speed of the source; and $\beta_i(j)$ is the direction of movement of the source. We assume that the height difference between the source and the sensor is negligible- meaning the elevation angle is constant.

An acoustic impulse response (AIR), between a source and the sensor, is based on several aspects: room size and geometry; locations of the source and the sensor; surfaces' material; and objects' presence. Let $h_{\boldsymbol{\theta}_i(j)}(n, j)$ denote an AIR, which is defined as the response at discrete time index n to an impulse transmitted at discrete time index j, between the ith source and the sensor, with respect to $\boldsymbol{\theta}_i(j)$.

The signal received by the sensor is defined by:

$$y_i(n) = \sum_{j=-\infty}^{\infty} h_{\boldsymbol{\theta}_i(j)}(n, j) x_i(j), \tag{2}$$

where $x_i(n)$ is the signal transmitted by the ith source. We note that the system, characterized by the AIR, is time-variant since the movement of the source influences the AIR. We use a WGN signal as the source signal, as it fully excites the frequency response of the AIR. The received signal is recorded and divided into time frames. The trajectory, formed by the movement of the source during the frame, is assumed to be approximated by a linear movement segment. Each frame is examined individually in order to estimate the location and velocity.

For each frame, we manage training and test datasets. For generating the training dataset, we choose arbitrarily m known locations and velocities of the source $\overline{\boldsymbol{\Theta}} = \left\{\overline{\boldsymbol{\theta}}_1(q), \ldots, \overline{\boldsymbol{\theta}}_m(q)\right\} \subset \mathbb{R}^d$, where q is a discrete time index representing a query point along the trajectories of the sources, and d is the dimension of the parameters vector (i.e., the number of system's degrees of freedom). Let $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_{m+1}(q), \ldots, \boldsymbol{\theta}_{m+M}(q)\} \subset \mathbb{R}^d$ be a set of M arbitrary unknown source locations and velocities, corresponding to the M measurements of the test dataset. The query point q is defined as the discrete time index corresponding to the midway point of the trajectory, for minimizing the approximation error of the true trajectory by the linear segment. Note that the acoustic environment is fixed between the training and test stages. In other words, the location and velocity values of the sources are the only degrees of freedom of the controlling parameters of the AIR.

III. Diffusion Kernel

A. Feature Vectors

We extract a feature vector from each observation. The feature vector is defined based on an autocorrelation function of the observation. This choice is explained by the benefits of using a second-order statistics measure over the raw observation. The autocorrelation function expresses the location better, and is less dependent on the specific random and unknown signal transmitted by the source. From (2), the time-variant autocorrelation function of $y_i(n)$ is given by substituting the autocorrelation function of the transmitted WGN signal (i.e., $c_{\mathbf{x}_i}(\tau) = \sigma^2_{\mathbf{x}_i}\delta(\tau)$) :

$$c_{\mathbf{y}_i}(n_1, n_2) = \sigma^2_{\mathbf{x}_i} \sum_{j=-\infty}^{\infty} h_{\boldsymbol{\theta}_i(j)}(n_1, j) h_{\boldsymbol{\theta}_i(j)}(n_2, j). \tag{3}$$

As indicated by (3), each observation can be represented as a function of the location and velocity. Given a sufficiently short time interval, we assume the first two moments of the quasi-stationary speech signal would not change along the interval.

Indeed, by taking into consideration an additional assumption- slow speed and gradually varying velocity- we restrict the AIR to vary slightly along the time frames. Accordingly, we can acquire the familiar convolution-based version of (3), as in [12], [14]:

$$c_{\mathbf{y}_i}(\tau) = h_{\boldsymbol{\theta}_i}(\tau) * h_{\boldsymbol{\theta}_i}(-\tau) * c_{\mathbf{x}_i}(\tau). \tag{4}$$

As implied by (4), the time differences of the autocorrelation function of the measured signal completely rely upon the variations of the AIR- meaning, the evolution of the location and velocity of the ith source.

Let $\mathbf{c}_i$ denote a feature vector, composed of the first D elements of the autocorrelation function of the observation $\mathbf{y}_i$. In such a fashion, a feature vector is extracted for each signal received by the sensor. Let $\Gamma = \{\mathbf{c}_i\}_{i=1}^{M}$ and $\overline{\Gamma} = \{\overline{\mathbf{c}}_i\}_{i=1}^{m}$ denote the sets of the feature vectors with respect to the unlabeled parameters in $\boldsymbol{\Theta}$ and the labeled parameters in $\overline{\boldsymbol{\Theta}}$, respectively.

B. Diffusion Kernel Computation

Despite the movement of the sources, we assume the high-dimensional feature vectors lie on a nonlinear and low-dimensional manifold $\mathcal{M}$ of dimension d ($d \ll D$). This hypothesis is justified by the combination of a quasi-stationary source signal, short time intervals, a slow speed and a gradually changing velocity of the source, and a fixed sensor. This combination leads to small variations in the feature vectors along the intervals, and consequently allows recovery of the degrees of freedom of the system: the location and velocity of the sources. The structure of the manifold is assumed to be locally linear in the vicinity of each observation. For

measuring affinity between observations and complying with the manifold's structure, we use Euclidean distance in the proximity of each observation. Greater distances, on the other hand, are ignored as the Euclidean distance is unreliable for large scales. These assumptions are explained in detail in [13].

Let $\mathbf{W}$ be an $m \times m$ affinity matrix between all the feature vectors in $\overline{\Gamma}$. The affinity matrix consists of a Gaussian kernel with a scale parameter ε, and its ijth element is defined by:

$$\mathbf{W}^{(ij)} = \begin{cases} \exp\left\{-\frac{\|\overline{\mathbf{c}}_i - \overline{\mathbf{c}}_j\|^2}{\varepsilon}\right\} & \text{, if } \overline{\mathbf{c}}_i \in \overline{\mathcal{N}}_j \text{ or } \overline{\mathbf{c}}_j \in \overline{\mathcal{N}}_i \\ 0 & \text{, otherwise,} \end{cases} \tag{5}$$

where $\overline{\mathcal{N}}_j$ is the set of the k-nearest-neighbors of $\overline{\mathbf{c}}_j$ in $\overline{\Gamma}$.

IV. Localization Based on Diffusion Mapping

A. Manifold Parameterization

By using a diagonal matrix $\mathbf{D}$ with $\mathbf{D}^{(ii)} = \sum_{j=1}^{m} \mathbf{W}^{(ij)}$, we normalize the affinity matrix $\mathbf{W}$, and obtain the transition matrix

$$\mathbf{P} = \mathbf{D}^{-1}\mathbf{W}. \tag{6}$$

Let $\{\lambda_j\}_{j=0}^{m-1}$ and $\{\boldsymbol{\psi}_j\}_{j=0}^{m-1}$ be the eigenvalues and eigenvectors of the transition matrix $\mathbf{P}$. Note that $\lambda_0 = 1$ and its corresponding eigenvector $\boldsymbol{\psi}_0$ relate to the trivial case [15]. The eigenvectors of $\mathbf{P}$ are assumed to establish the reparameterization of the controlling parameters of the m training observations. Thus, let Ψ_d be the diffusion mapping of the training feature vectors into the embedded Euclidean space $\mathbb{R}^d$, which is spanned by d eigenvectors corresponding to the d largest eigenvalues (trivial case excluded). Ψ_d is defined as

$$\Psi_d : \overline{\mathbf{c}}_i \to \left[\lambda_1 \boldsymbol{\psi}_1^{(i)}, \ldots, \lambda_d \boldsymbol{\psi}_d^{(i)}\right]^T. \tag{7}$$

B. Extension for New Observations

Given an additional set of M new observations, associated with unknown locations and velocities of the sources, we seek to map them as well into the embedded manifold. For bypassing another spectral decomposition, we add M new rows to the affinity matrix $\mathbf{W}$ as follows:

$$\mathbf{W}^{(\tilde{i}j)} = \begin{cases} \exp\left\{-\frac{\|\mathbf{c}_i - \overline{\mathbf{c}}_j\|^2}{\varepsilon}\right\} & \text{, if } \overline{\mathbf{c}}_j \in \mathcal{N}_{\tilde{i}} \\ 0 & \text{, otherwise,} \end{cases} \tag{8}$$

where $\tilde{i} = m+i$, and $\mathcal{N}_{\tilde{i}}$ is the set of the k-nearest-neighbors of $\mathbf{c}_i$ in $\overline{\Gamma}$. As opposed to the construction process of the affinity matrix in (5), the affinity in (8) is executed with respect to the training set only. Accordingly, the new entries of the transition matrix $\mathbf{P}$ are given by

$$\mathbf{P}^{(\tilde{i}j)} = \left(\sum_{j=1}^{m} \mathbf{W}^{(\tilde{i}j)}\right)^{-1} \mathbf{W}^{(\tilde{i}j)}. \tag{9}$$

The new entries of the extended eigenvectors can be represented as a weighted interpolation of the original entries of the eigenvectors, as follows:

$$\boldsymbol{\psi}_l^{(\tilde{i})} = \frac{1}{\lambda_l} \sum_{j=1}^{m} \mathbf{P}^{(\tilde{i}j)} \boldsymbol{\psi}_l^{(j)}. \tag{10}$$

Subsequently, Ψ_d maps the new feature vectors of the unlabeled observations to their corresponding representation of dominating parameters in the embedded manifold:

$$\Psi_d : \mathbf{c}_i \to \left[\lambda_1 \boldsymbol{\psi}_1^{(\tilde{i})}, \ldots, \lambda_d \boldsymbol{\psi}_d^{(\tilde{i})}\right]^T. \tag{11}$$

C. Recovery of the Independent Controlling Parameters

The unknown location and velocity of the test observation are estimated using its labeled neighbors on the manifold, by a weighted interpolation:

$$\hat{\boldsymbol{\theta}}_i(q) = \sum_{j:\, \boldsymbol{\Psi}_d(\overline{\mathbf{c}}_j) \in \tilde{\mathcal{N}}_i} \gamma_j\left(\mathbf{c}_i\right) \overline{\boldsymbol{\theta}}_j(q), \tag{12}$$

where $\tilde{\mathcal{N}}_i$ consists of the $\tilde{k}$-nearest training observations $\overline{\mathbf{c}}_j$ of $\mathbf{c}_i$, in the embedded space. The interpolation coefficients $\{\gamma_j\}_{j=1}^{\tilde{k}}$, which satisfy $\sum_{j=1}^{\tilde{k}} \gamma_j\left(\mathbf{c}_i\right) = 1$, are in proportion to the distance between $\boldsymbol{\Psi}_d\left(\mathbf{c}_i\right)$ and each of its embedded $\tilde{k}$-nearest labeled neighbors:

$$\gamma_j\left(\mathbf{c}_i\right) = \frac{\exp\left(-\frac{\|\boldsymbol{\Psi}_d(\mathbf{c}_i) - \boldsymbol{\Psi}_d(\overline{\mathbf{c}}_j)\|^2}{\varepsilon_{\gamma_i}}\right)}{\sum_{l:\, \boldsymbol{\Psi}_d(\overline{\mathbf{c}}_l) \in \tilde{\mathcal{N}}_i} \exp\left(-\frac{\|\boldsymbol{\Psi}_d(\mathbf{c}_i) - \boldsymbol{\Psi}_d(\overline{\mathbf{c}}_l)\|^2}{\varepsilon_{\gamma_i}}\right)}, \tag{13}$$

where ε_{γ_i} is defined as the minimal distance between $\boldsymbol{\Psi}_d\left(\mathbf{c}_i\right)$ and its nearest neighbor. Following the considerations discussed in [13], we define the normalized estimation error as

$$\mathbf{e}\left(\mathbf{c}_i\right) = \left[e_i^{(1)}, \ldots, e_i^{(d)}\right], \tag{14}$$

where each coordinate is assigned for a different physical quantity. Its jth element is defined by:

$$e_i^{(j)} = \frac{\mid \boldsymbol{\theta}_i^{(j)}(q) - \hat{\boldsymbol{\theta}}_i^{(j)}(q) \mid}{\mid \boldsymbol{\theta}_i^{(j)}(q) \mid}. \tag{15}$$

Note that the recovery of the location and velocity is based on the midway point, as indicated by (12). As pointed out by (15), the approximation error of the true trajectory by the linear segment is minimized based on that point.

D. Accuracy Measure

According to the definition of the estimation error (14), we define the root mean square error (RMSE) as:

$$\text{RMSE} = \sum_{j=1}^{d} \alpha_j \sqrt{\frac{1}{M} \sum_{i=1}^{M} \left(e_i^{(j)}\right)^2}, \tag{16}$$

where $\{\alpha_j\}$ are weights that represent the significance of the estimation error for each of the physical quantities (For simplicity, we define $\{\alpha_j\} = \frac{1}{d}$).

V. Experimental Results

In this section, we demonstrate the capability of the proposed single-sensor localization algorithm for estimating the location and velocity of an acoustic source, despite the influence of environmental conditions on the movement of the sources.

We describe the simulated setup used for conducting the experimental study, based on the image method [16]. Room dimensions were defined as $6 \times 5.8 \times 3\,\mathrm{m}^3$, and an omnidirectional sensor was positioned at $(3, 1, 1.8)$ m. The reverberation time of the room was set as $T_{60} = 0.3$ sec, simulating moderate reverberation conditions. In each location of the source, a 1 sec long signal of a zero-mean and unit-variance WGN, sampled at $f_s = 16$ kHz, is transmitted from the source. After the convolution of the transmitted signal with the AIR, it is acquired by the sensor. We obtain a total of $m + M$ observations, where m out of them are randomly selected for the training set, and the remaining M samples are allocated for the test set. The length of the autocorrelation-based feature vector is fixed to $D = 800$ lags. The directional standard deviations of the Brownian motion term are defined in direct proportion to the Brownian motion coefficient, according to

$$\left[\ \sigma_{\tilde{x}},\quad \sigma_{\tilde{y}}\ \right]^T = \left[\ \eta \cdot v_{\tilde{x}_{max}}/\xi \quad ,\eta \cdot v_{\tilde{y}_{max}}/\xi\ \right]^T, \qquad (17)$$

where $v_{\tilde{x}_{max}}$ and $v_{\tilde{y}_{max}}$ are the maximal horizontal and vertical speed components drawn in the experiment, respectively; η is the Brownian motion coefficient; and $\xi = 4$ is defined as the four-sigma confidence level. We position all sources at a distance of 1 m from the sensor and at an azimuth angle of $45°$. Their movement is initialized with a speed of 0.5m/sec, in directions of movement which are drawn according to a uniform distribution $U\,[45, 85]\,°$. Training and test sets of 720 observations each are generated. In each simulation, we introduce a Brownian motion term to the movement of all sources. The simulation is repeated each time with a different Brownian motion coefficient value, ranging from 0.05 to 0.8, followed by a scenario free of Brownian motion (for comparison to [13]).

All previous movement experiments, described in [13], have focused on a single degree of freedom. The introduction of the Brownian motion term to the source's movement, however, may lead to independent and random variability in additional parameters.

Figure 1 illustrates the total estimation error of the proposed algorithm for various Brownian motion conditions, with respect to three intrinsic dimension d (i.e., the number of dimensions of the low-dimensional manifold) cases: one (direction), two (direction and speed), or three (direction, radius and azimuth). As expected, as the Brownian motion term gets more dominant, the estimation error grows, for all cases. The reason for the significant estimation error values is the revoke of our assumption: approximation of the trajectories by short linear segments. The assumption is no longer valid due to the rapid and random changes in the movement of the sources. These changes prevent us from collecting a sufficiently long enough observation, and consequently a meaningful feature vector during each time frame.

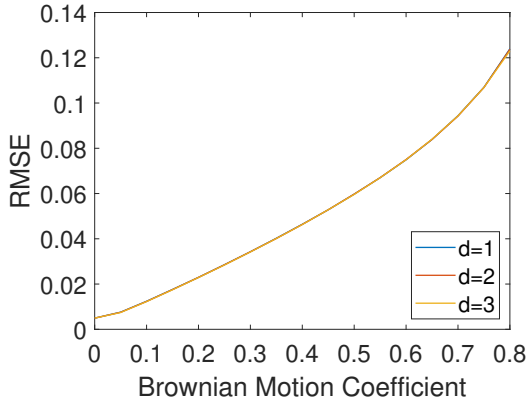

Fig. 1: Performance of the proposed algorithm as a function of Brownian motion coefficient, for various intrinsic dimension cases.

Regarding the determination of the intrinsic dimension value based on the various cases, we note that the examined scenario is unambiguous upon the parameters vector. In addition, the Brownian motion adds small variability compared to the significant variability triggered by the direction. Moreover, there is no case that dominates all the others- all the more so, the difference between the results is negligible, as depicted in Fig. 1. Thus, we determine $d = 1$ as the dimension of the low-dimensional embedded manifold.

For emphasizing the impact of the Brownian motion, we elaborate on the scenario of a Brownian motion coefficient of 0.8. The individual RMSE values (composing the total RMSE all together) that are associated with the velocity (23.99% for speed, and 25.03% for direction) are colossal. On the other hand, the individual RMSE values of the location (0.24% for radius, and 0.32% for azimuth) are negligible.

For demonstrating the localization results of the location and velocity through time, we focus on an arbitrary source as a representative example of the aforementioned scenario of a Brownian motion coefficient of 0.8. The source is characterized by a speed of 0.5 m/sec (common to all) and a direction of about $52°$ (drawn). Figure 2 depicts a comparison through time between the true and the estimated trajectories of all four unknown parameters of the source: the radius (a), azimuth angle (b), speed (c), and direction (d). The true trajectories of all location and velocity parameters describe fluctuations in the location and velocity values through time, which reflect the nature of the Brownian motion.

The estimated trajectories are divided into 3 cases- common to all is the nearest training neighbors in the manifold, which are determined according to the midway point. The estimated trajectories of the radius and the azimuth recover the true values well despite the Brownian motion. However, the efforts to recover the instantaneous speed and direction of the source are unsuccessful. The first estimated trajectory, denoted as LM,

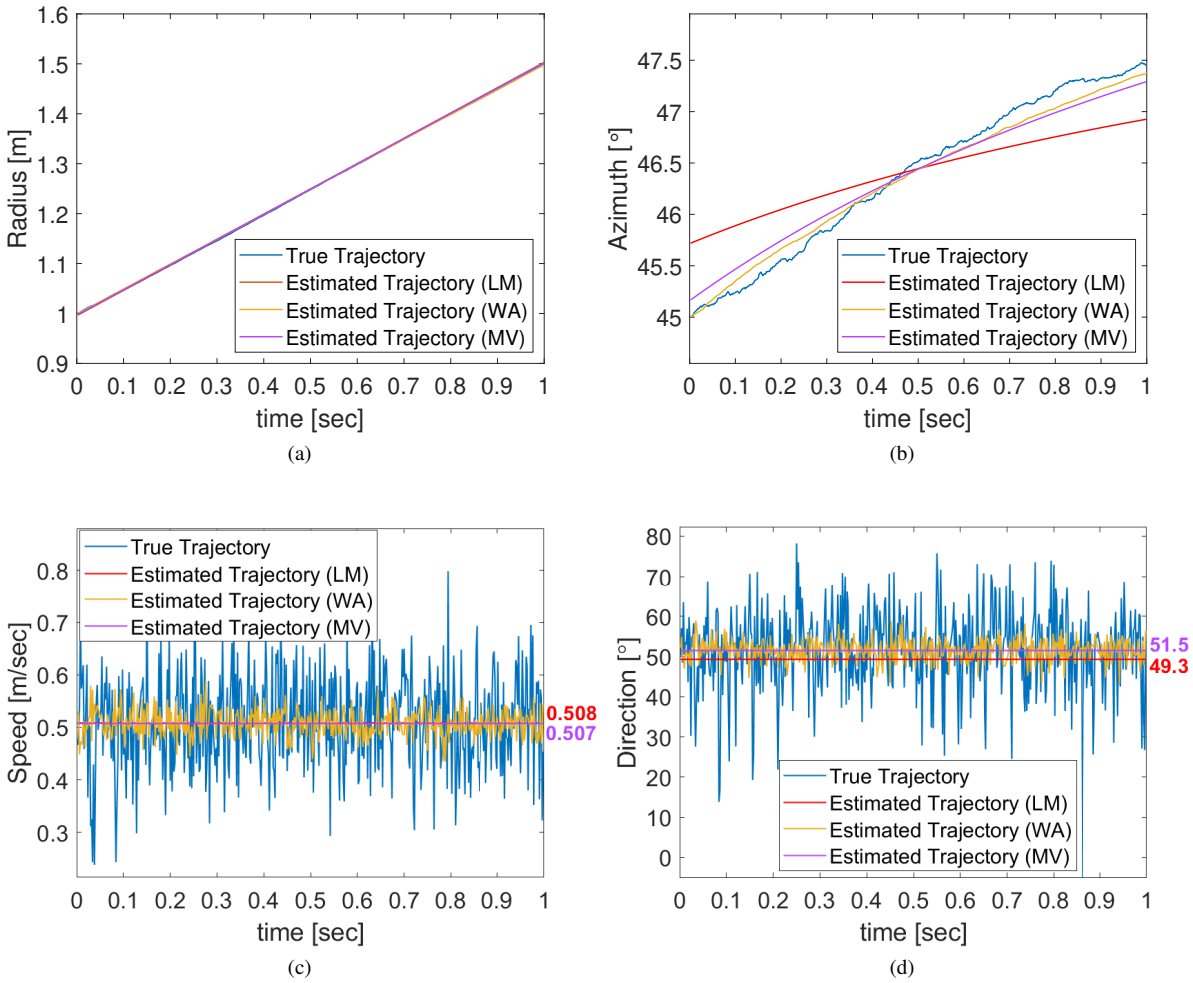

Fig. 2: Localization results of a moving source through time of each of the unknown parameters for the scenario of Brownian motion coefficient of 0.8. (a) Radius, (b) azimuth, (c) speed, and (d) direction. Blue represents true values, red represents the estimated values according to the linear model (LM), yellow represents the estimated values according to a weighted average of the trajectories (WA), and purple represents the estimated values according to a mean velocity value (MV).

is executed according to the linear model, as described in detail in Section III. This method demonstrates fair estimation results of the average velocity.

In the case that the whole trajectories of the nearest training observations are known, these localization results can be improved, as suggested by the remaining two estimation methods. The second estimated trajectory, denoted as WA, is produced by a weighted average of the trajectories of the nearest training neighbors. This method mitigates the explicit random fluctuations in the location and velocity values in each step. However, the weighted average fairly succeeds to estimate the average speed and direction values through time. Another improvement can be achieved by a combination of the two, by estimating the trajectories according to a linear model. The constant velocity is determined according to the mean value of the velocity trajectories of the weighted average method. This method, denoted as MV, improves both the accuracy of the estimated location and the estimated average velocity. It provides an accuracy of less than one percent, with respect to the average velocity, which is in practice the Brownian-motion-free instantaneous velocity.

VI. Conclusions

The problem of single-sensor localization of moving sources that change their velocity rapidly and randomly has been addressed. Using diffusion maps, the proposed supervised algorithm implements a data-driven approach for learning the nonlinear structure of the manifold of the observations. The data on the manifold can be arranged on the manifold according to the location and velocity values of the sources, in the case that their velocity is slowly changing. While the algorithm is designated for recovering the location and velocity of slow sources that change direction and speed gradually, it is even capable of successfully estimating the location and the

average velocity of sources that change their velocity rapidly and randomly but their average speed is slow.

REFERENCES

[1] M. Omologo and P. Svaizer, "Use of the crosspower-spectrum phase in acoustic event location," *IEEE Transactions on Speech and Audio Processing*, vol. 5, no. 3, pp. 288–292, 1997.

[2] J. H. DiBiase, H. F. Silverman, and M. S. Brandstein, "Robust localization in reverberant rooms," in *Microphone Arrays*, pp. 157–180, Springer, 2001.

[3] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[4] B. Laufer, R. Talmon, and S. Gannot, "Relative transfer function modeling for supervised source localization," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pp. 1–4, IEEE, 2013.

[5] B. Laufer-Goldshtein, R. Talmon, and S. Gannot, "A study on manifolds of acoustic responses," in *International Conference on Latent Variable Analysis and Signal Separation*, pp. 203–210, Springer, 2015.

[6] B. Laufer-Goldshtein, R. Talmon, and S. Gannot, "Semi-supervised sound source localization based on manifold regularization," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 8, pp. 1393–1407, 2016.

[7] B. Laufer-Goldshtein, R. Talmon, and S. Gannot, "Manifold-based bayesian inference for semi-supervised source localization," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6335–6339, IEEE, 2016.

[8] B. Laufer-Goldshtein, R. Talmon, and S. Gannot, "Semi-supervised source localization on multiple manifolds with distributed microphones," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 7, pp. 1477–1491, 2017.

[9] B. Laufer-Goldshtein, R. Talmon, and S. Gannot, "Speaker tracking on multiple-manifolds with distributed microphones," in *International Conference on Latent Variable Analysis and Signal Separation*, pp. 59–67, Springer, 2017.

[10] B. Laufer-Goldshtein, R. Talmon, and S. Gannot, "A hybrid approach for speaker tracking based on tdoa and data-driven models," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 4, pp. 725–735, 2018.

[11] B. Laufer-Goldshtein, R. Talmon, S. Gannot, *et al.*, "Data-driven multi-microphone speaker localization on manifolds," *Foundations and Trends® in Signal Processing*, vol. 14, no. 1–2, pp. 1–161, 2020.

[12] R. Talmon, I. Cohen, and S. Gannot, "Supervised source localization using diffusion kernels," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pp. 245–248, IEEE, 2011.

[13] E. Zeitouni and I. Cohen, "Single-sensor source localization of moving sources using diffusion kernels," *Applied Acoustics*, in press.

[14] R. Talmon, D. Kushnir, R. R. Coifman, I. Cohen, and S. Gannot, "Parametrization of linear systems using diffusion kernels," *IEEE Transactions on Signal Processing*, vol. 60, no. 3, pp. 1159–1173, 2011.

[15] A. Singer and R. R. Coifman, "Non-linear independent component analysis with diffusion maps," *Applied and Computational Harmonic Analysis*, vol. 25, no. 2, pp. 226–239, 2008.

[16] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.