

## Acoustic target tracking through a cluster of mobile agents

Drioli, Carlo; Giordano, Giulia; Salvati, Daniele; Blanchini, Franco; Foresti, Gian Luca

**DOI**

[10.1109/TCYB.2019.2908697](https://doi.org/10.1109/TCYB.2019.2908697)

**Publication date**

2019

**Document Version**

Accepted author manuscript

**Published in**

IEEE Transactions on Cybernetics

**Citation (APA)**

Drioli, C., Giordano, G., Salvati, D., Blanchini, F., & Foresti, G. L. (2019). Acoustic target tracking through a cluster of mobile agents. *IEEE Transactions on Cybernetics*, 51 (2021)(5), 2587-2600.  
<https://doi.org/10.1109/TCYB.2019.2908697>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

# Acoustic Target Tracking Through a Cluster of Mobile Agents

Carlo Drioli<sup>1</sup>, *Member, IEEE*, Giulia Giordano, Daniele Salvati<sup>2</sup>, Franco Blanchini, *Senior Member, IEEE*,  
and Gian Luca Foresti<sup>3</sup>, *Senior Member, IEEE*

**Abstract**—This paper discusses the problem of tracking a moving target by means of a cluster of mobile agents that is able to sense the acoustic emissions of the target, with the aim of improving the target localization and tracking performance with respect to conventional fixed-array acoustic localization. We handle the acoustic part of the problem by modeling the cluster as a sensor network, and we propose a centralized control strategy for the agents that exploits the spatial sensitivity pattern of the sensor network to estimate the best possible cluster configuration with respect to the expected target position. In order to take into account the position estimation delay due to the frame-based nature of the processing, the possible positions of the acoustic target in a given future time interval are represented in terms of a compatible set, that is, the set of all possible future positions of the target, given its dynamics and its present state. A frame-by-frame cluster reconfiguration algorithm is presented, which adapts the position of each sensing agent with the goal of pursuing the maximum overlap between the region of high acoustic sensitivity of the entire cluster and the compatible set of the sound-emitting target. The tracking scheme iterates, at each observation frame, the computation of the target compatible set, the reconfiguration of the cluster, and the target acoustic localization. The reconfiguration step makes use of an opportune cost function proportional to the difference of the compatibility set and the acoustic sensitivity spatial pattern determined by the mobile agent positions. Simulations under different geometric configurations and positioning constraints demonstrate the ability of the proposed approach to effectively localize and track a moving target based on its acoustic emission. The Doppler effect related to moving sources and sensors is taken into account, and its impact on performance is analyzed. We compare the localization results with conventional static-array localization and positioning of acoustic sensors through genetic algorithm optimization, and results demonstrate the sensible improvements in terms of localization and tracking performance. Although the method is discussed here with respect to acoustic target tracking, it can be effectively adapted to video-based localization and

tracking, or to multimodal information settings (e.g., audio and video).

**Index Terms**—Acoustic target tracking, microphone arrays, mobile agents, set-theoretic position estimation, source localization, sparse sensor networks.

## I. INTRODUCTION

**M**ULTIROBOTS' target detection and tracking are an active research topic within the broader field of cooperative multirobots systems [1]. Its range of applications includes ground and aerial surveillance in civilian and military operations, broadcasting of sports and entertainment events, and underwater acoustic monitoring, to name a few.

Coordination and control strategies of the units in the cluster often depend on the nature of the robotic platform, on the nature of the sensing devices they can rely on, and on the nature and number of moving targets [2]. Commonly addressed robotic platform include unmanned aerial vehicles (UAVs), unmanned ground vehicles (UGVs), and unmanned underwater vehicles (UUVs). Among these, UAVs are rapidly gaining popularity for research and applications in cooperative robotics, due to the availability of reliable, stable, and cost-effective devices which are also able to carry a variety of sensors. Recent research concerning multirobots target tracking based on UAV clusters can be found in [3], in which cooperative control of fixed-wings UAVs is addressed for the tracking of ground mobile targets emitting radio-frequency signals; in [4], the path planning strategy for video-based tracking of ground targets is especially aimed at maximizing the visibility of the target in an urban area setting; thus, when ground obstacles of different shapes and sizes may cause occlusion; in [5], a study on the localization and tracking of a generic target through a cluster of UAVs with acoustic sensors is illustrated.

In general, the studies on multirobots target tracking seem to have mainly addressed configurations based on video and range sensors [2] while neglecting the potentialities of exploiting acoustic sensing, despite the wide knowledge available in the field of multichannel audio signal processing. Acoustic source localization (ASL) through microphone arrays is a rather mature research field, in which a typical scenario consists of estimating a sound source position in space by analyzing the sound field sensed by a microphone array [6]–[10]. While most of the literature in the past has addressed sensor arrays with a static position and fixed geometry, the interest

Manuscript received September 22, 2018; revised December 30, 2018, March 9, 2019, and March 22, 2019; accepted March 26, 2019. This work was supported in part by the Research Project PNRM “Prescriptive Situational Awareness for Cooperative Auto-Organizing Aerial Sensor Networks” (PRESNET) and in part by the Research Project PNRM “Augmented Reality for Mobile Applications” (RA2M). This paper was recommended by Associate Editor J. Su. (*Corresponding author: Carlo Drioli.*)

C. Drioli, D. Salvati, F. Blanchini, and G. L. Foresti are with the Department of Mathematics, Computer Science and Physics, University of Udine, 33100 Udine, Italy (e-mail: carlo.drioli@uniud.it; daniele.salvati@uniud.it; franco.blanchini@uniud.it; gianluca.foresti@uniud.it).

G. Giordano is with the Delft Center for Systems and Control, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: g.giordano@tudelft.nl).

This paper has supplementary downloadable material available at <http://ieeexplore.ieee.org>, provided by the author.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCYB.2019.2908697

toward reconfigurable arrays and dynamically varying geometry is rapidly increasing [11]. Fields of application in which the localization of acoustic sources through sensor networks is desirable include audio surveillance [12] and acoustic monitoring [13]; autonomous robots [14], [15]; human–robot interaction [16]; and animal ecology [17].

When the acoustic source position may change during the localization process, a number of factors related to the source motion must be taken into account to effectively track the position of the target. Under given hypotheses on the target nature and on its motion constraints, it is possible to pair the localization signal-processing methods with tracking techniques based on motion dynamics modeling, which allow to forecast the target position to a certain extent. Recent approaches to acoustic tracking may also exploit the timbre characteristics of the source to recognize its acoustic signature and improve the tracking [18].

Here, we are interested in the specific problem of controlling a multirobots cluster of acoustic sensing units to localize and track a moving sound-emitting target, under the assumption of delayed measurements of the target acoustic emission. We are moreover interested in discussing the mobile agents positioning strategies in terms of the intersection of the whole cluster acoustic sensitivity pattern and of the compatible sets of the target. The choice of compatible sets framework is motivated by recent investigations that have reported the effectiveness of set-theoretic control methods for problems involving object tracking and robotic control in different scenarios [19]–[21].

To address the acoustic localization step in the target tracking and mobile agents control scheme, we might refer to the techniques for ASL through microphone networks. The optimal static positioning of microphones in sensor networks has been addressed over a few decades [11], [22], [23], however, the tracking of moving sources through reconfigurable and/or moving microphone networks has been addressed only recently. In [24], an evolutionary strategy for the optimal static placement of microphones in an indoor search region is proposed. It is based on a spatial-likelihood function built on a spatial map of the acoustic power for a given sensor network configuration. In [25], the case of ASL by a moving sensors pair is discussed. The target source localization in this case is achieved through a particle filtering algorithm; however, no strategy for the sensor positioning is suggested. In [26] and [27], examples of robotic applications are provided, in which the dynamic configuration of the sparse sensor network is achieved through particle swarm optimization methods.

We propose here a centralized control strategy for the mobile agents of the cluster, based on the computation of the compatible sets of the tracked source and on the exploitation of the acoustic sensitivity spatial pattern determined by the instantaneous positions of the agents carrying the microphones. The aim of the proposed method is to improve the target localization and tracking performance with respect to conventional fixed-array acoustic localization, by allowing the sensor network geometry to dynamically adapt, provided the sensors are carried by mobile robotic agents. With respect to other studies in the literature, this paper’s main novelty is the combination of the forecasting setting for the target dynamics

and of a model of the mobile agents cluster acoustic sensitivity, to define an effective centralized control strategy for the agents of the cluster. The main objective of the strategy is to increase as much as possible the target localization and tracking performance by the effective positioning of the agents carrying the sensors. With respect to a fixed sensor network, or to a cluster of agents moving with fixed geometry, the proposed positioning algorithm has several advantages: for example, it can provide a way to keep the agents at the minimum distance from their target, which is safe against losing it. From the acoustic point of view, this can be advantageous since it allows sensing of acoustic data with the smallest possible time latency, the highest signal-to-noise ratio (SNR), and the best localization properties. In terms of path planning, the proposed control strategy for mobile agents based on acoustic sensing can be seen as a component of a more general planning problem with several constraints, including collision avoidance, shortest path selection, and obstacle sensing. Considering these additional constraints is out of the scope of this paper.

From the point of view of array signal processing, the method can also be seen as a novel dynamic sensor network positioning and reconfiguration approach in the context of moving ASL and tracking. Practical applications of the proposed method are foreseen in situations where sensors are mounted on pools of autonomous vehicles (e.g., UGVs or UAVs), whose goal is to localize and track a moving acoustic target, while maximizing the amount and quality of the global information obtained by fusing the partial information from each single device in the pool. This scenario also includes the case in which the acoustic emission encodes further higher level information of interest, such as in acoustic events classification or as in speech recognition. In such cases, localization and tracking can be followed by signal enhancement through beamforming and by classification/recognition steps.

The centralized control method is discussed here with respect to acoustic target tracking, however, we stress the fact that it can be effectively adapted to video-based localization and tracking, or to full information (audio and video) sensor networks. An example task related to such a different scenario is target coverage maximization by a visual sensor network, in which this strategy might be employed to control the agents carrying the cameras with overlapping fields of view, so to provide maximum target visual coverage and best image resolution while keeping a safe distance.

## II. PROBLEM FORMULATION

We aim at designing a centralized control strategy for a cluster of mobile agents in the following problem setting.

- 1) The mobile robotic agents have known dynamics and their average reachable speed is larger than the maximum target speed.
- 2) The moving target has known dynamics, with bounded process noise and magnitude-bounded random control input.
- 3) Each mobile agent carries a single microphone with omnidirectional sensing characteristics.
- 4) The sound emitting target can be assimilated to an acoustic point source.

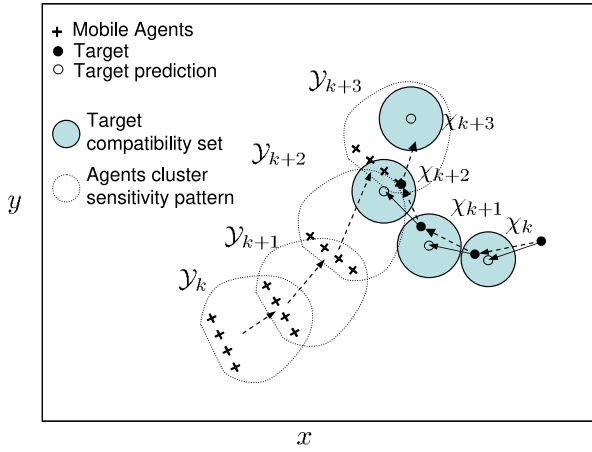


Fig. 1. Schematic example of target tracking and mobile agents positioning control (fixed linear cluster geometry).

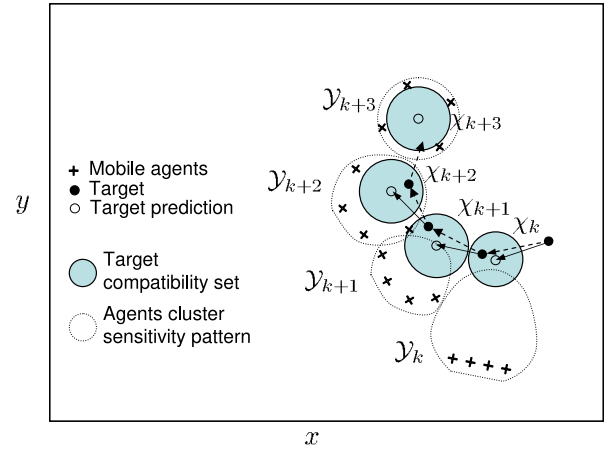


Fig. 2. Schematic example of target tracking and robotic agents positioning and reconfiguration (arbitrary cluster geometry).

- 5) The position measurements of the target are delayed by a known, generally nonconstant, time lag.
- 6) The environment in which the sensors and the target move is assumed free from obstacles.

We consider a standard second-order dynamical model for each mobile agent of the cluster, which can be written in the following discrete-time state-space representation:

$$\begin{cases} \mathbf{x}_{k+1}^{(m)} = A^{(m)}\mathbf{x}_k^{(m)} + B^{(m)}\mathbf{u}_k^{(m)} \\ \mathbf{r}_k^{(m)} = C^{(m)}\mathbf{x}_k^{(m)} \end{cases} \quad (1)$$

where \$\mathbf{r}\_k^{(m)}\$ is the \$m\$th agent position at the discrete-time instant \$k\$; \$\mathbf{x}^{(m)}\$ is the state of agent dynamics; \$\mathbf{u}^{(m)}\$ is the control input; and \$A\$, \$B\$, and \$C\$ are the state matrix, input matrix, and output matrix, respectively. We will assume that the \$m\$ mobile agents have identical dynamics, that is, \$A^{(m)} = A\_M\$, \$B^{(m)} = B\_M\$, and \$C^{(m)} = C\_M\$ for \$m = 1, \dots, M\$.

Moreover, we assume the following continuous-time dynamical model for the motion of the sound-emitting target:

$$\ddot{\mathbf{r}}(t) + \lambda \dot{\mathbf{r}}(t) = \mathbf{f}(t) \quad (2)$$

with \$\mathbf{f} = [f\_x, f\_y]^T\$, \$|\mathbf{f}| < F\$ being the control input (a magnitude bounded force) and \$\lambda^- < \lambda < \lambda^+\$ being the parameter that models the damping of the system. The vector \$\mathbf{r} = [r\_x, r\_y]^T\$ is the target position in the two spatial coordinates, and \$\dot{\mathbf{r}} = [\dot{r}\_x, \dot{r}\_y]^T\$ and \$\ddot{\mathbf{r}} = [\ddot{r}\_x, \ddot{r}\_y]^T\$ are the corresponding speed and acceleration. Similar to the mobile agent dynamics model, a discrete-time state-space representation can be derived for the motion dynamics of the target.

Even if the environment in which the sensing mobile agents and the target move is assumed free from obstacles, the agents may only move within a subset of the free space due to their dynamics and possibly due to structural constraints. In the following, we will call, respectively, \$\mathcal{Y}\_k\$, and \$\mathcal{X}\_k\$, \$k = 0, 1, 2, \dots\$, the acoustic sensitivity pattern determined by the agents cluster configuration and the compatibility set of the target, with respect to a generic time interval \$k\$ (whose length is typically the size of the audio analysis buffer of the acoustic front-end).

The problem setting is illustrated in Figs. 1 and 2. If a new estimate of the target position and velocity is available at discrete-time \$k\$, the reconfiguration control procedure

computes the set \$\mathcal{X}\_{k+1}\$ of possible states of the target at discrete-time \$k + 1\$, corresponding to the next observation instant. This computation may also take into account the sound propagation (including possible Doppler effects) and the acoustic measurement delay, if known. The control system then decides the position of the robotic mobile agents carrying the sensors, to be reached at time \$k + 1\$, on the basis of the set \$\mathcal{Y}\_k\$ of the possible acoustic sensitivity pattern compatible with its state at time \$k\$. Fig. 1 represents the target tracking and agents positioning in the case of fixed linear cluster geometry (a typical configuration in microphone array systems), Fig. 2 represents the tracking in the case of arbitrary, time-varying cluster geometry.

### III. DYNAMIC RECONFIGURATION OF THE CLUSTER OF MOBILE AGENTS

#### A. Cluster Acoustic Sensitivity Spatial Pattern Computation

From the acoustic point of view, the cluster of acoustic sensing agents can be modeled as a sparse array of \$M\$ microphones located in \$\mathbf{r}\_k^{(m)} = [x\_k^{(m)}, y\_k^{(m)}]\$, \$m = 1, \dots, M\$ at time \$k\$, and with no structural position constraints within the search space. The environment is assumed reverberant and affected by stationary noise, which corrupts the acoustic signal emitted by the moving target. The localization of an acoustic source located in \$\mathbf{r}\_k^s\$ at time \$k\$ is performed through a steered response power algorithm with phase normalization (SRP-PHAT) [28]–[30]. This procedure requires the computation of a generalized cross-correlation (GCC) function between each microphone pair

$$\Phi_k(\mathbf{r}) = \sum_{a=1}^{M-1} \sum_{b=a+1}^M R_{ab,k}(\tau_{ab}(\mathbf{r})) \quad (3)$$

where \$\mathbf{r}\$ is a generic point of a discrete search grid \$\Gamma \subset \mathbb{N}^3\$; \$\tau\_{ab}(\mathbf{r})\$ is a mapping between the spatial point \$\mathbf{r}\$ and the time-difference of arrival (TDOA) related to the microphones \$a\$ and \$b\$; and \$R\_{ab,k}(\tau(\mathbf{r}))\$ is the GGC-PHAT function, defined as

$$R_{ab,k}(\tau_{ab}(\mathbf{r})) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{X_{a,k}(\omega)X_{b,k}^*(\omega)}{|X_{a,k}(\omega)X_{b,k}^*(\omega)|} e^{j\omega\tau_{ab}(\mathbf{r})} d\omega \quad (4)$$

where  $\omega$  is the angular frequency,  $X_{m,k}(\omega)$  is the Fourier transform of the signal that reached microphone  $m$  at time  $k$ ,  $(\cdot)^*$  denotes the complex conjugate,  $j$  denotes the imaginary unit, and  $|\cdot|$  denotes the absolute value. The standard mapping from a spatial point  $\mathbf{r}$  and the TDOAs is

$$\tau_{ab}(\mathbf{r}) = \left\lfloor \frac{f_s}{c} (\|\mathbf{r} - \mathbf{r}_a\| - \|\mathbf{r} - \mathbf{r}_b\|) \right\rfloor \quad (5)$$

for each microphone pair  $(a, b)$ , with  $a \in \{1, \dots, M-1\}$  and  $b \in \{a+1, \dots, M\}$ , and with  $f_s$  being the audio digital sampling frequency,  $c$  is the speed of sound propagation, and  $\lfloor \cdot \rfloor$  denoting the floor function that maps a real number to the largest integer less than or equal to that number.

Given this acoustic array processing setting, the sound-emitting target position can be finally estimated by picking the maximum value of the function  $\Phi_k(\mathbf{r})$  on all points of the search grid:

$$\hat{\mathbf{r}}_k^s = \underset{\mathbf{r} \in \Gamma}{\operatorname{argmax}} \Phi_k(\mathbf{r}). \quad (6)$$

The conventional localization algorithm searches a uniformly spaced grid, not taking into account the spatial accuracy characteristics of the sensor network. To compute a sensitivity measure of a given spatial configuration of the sensors we rely on the geometrically sampled grid (GSG) algorithm proposed in [30], which provides a measure of the sensor network localization accuracy in the surrounding region. The GSG array sensitivity is defined as a function  $\delta(\mathbf{r})$  that provides the number of discrete hyperboloids related to sensor pairs and intersecting in the position  $\mathbf{r}$ . The sensitivity function provides a measure of the density of the TDOA information over the spatial search grid and, thus, defines a measure of localization accuracy by identifying those areas for which the sensing system is more accurate. In the implementation adopted here, the number of intersections of the hyperbolas is weighted by a term which is inversely proportional to the squared distance, in order to take into account the decrement of acoustic energy reaching the sensors as the distance increases. If  $\delta(\mathbf{r})$  is the number of intersections in  $\mathbf{r}$ , counted using the numerical procedure provided in [30], we define its distance-weighted version as

$$\delta_w(\mathbf{r}) = \frac{\delta(\mathbf{r})}{1 + (1/M) \sum_{m=1}^M \|\mathbf{r} - \mathbf{c}_m\|} \quad (7)$$

where the denominator is augmented by the average distance from  $\mathbf{r}$  to each midpoint  $\mathbf{c}_m$ ,  $m = 1, 2, \dots, M$ , of the segment connecting two neighboring microphones. Given a sensitivity map  $\delta_w(\mathbf{r})$ ,  $\mathbf{r} \in \Gamma$ , we also define a sensing region  $\mathcal{Y}$  as the 2-D region for which  $\delta_w(\mathbf{r})$  has values above a certain threshold  $\hat{\delta}_w$ , that is,

$$\mathcal{Y} = \left\{ \mathbf{r} \in \Gamma \mid \delta_w(\mathbf{r}) \geq \hat{\delta}_w \right\}. \quad (8)$$

The region  $\mathcal{Y}$  represents the region in which the sensor network is more sensitive. As discussed in [30], a reasonable choice for the threshold is  $\hat{\delta}_w = (\max[\delta_w(\mathbf{r})] + \min[\delta_w(\mathbf{r})])/2$ .

Fig. 3 shows the sensitivity function  $\delta_w(\mathbf{r})$  for some typical acoustic sensor network configurations.

The sensitivity function  $\delta_w(\mathbf{r})$  and the acoustic sensitivity pattern  $\mathcal{Y}(\mathbf{r})$  will be used, along with the target position forecast model described in the next section, to design the cost

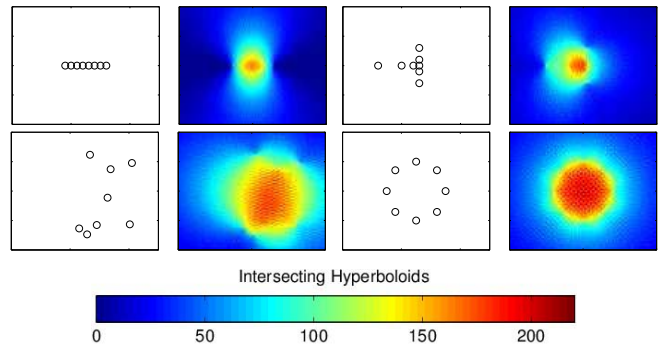


Fig. 3. Sensitivity maps corresponding to four different microphone array configurations. In the right column plots, red color is associated with high sensitivity, yellow with mid-range sensitivity, and blue with low sensitivity (the colormap shows the mapping between number of intersecting hyperboloids and colors).

function upon which the position update strategy of the sensing agents is based.

### B. Target Motion Modeling

Given the following general discrete-time description of the target dynamics:

$$\begin{cases} \mathbf{x}_{k+1} = a_{k+1,k}(\mathbf{x}_k, \mathbf{u}_k) + \eta_k \\ \mathbf{z}_k = h_k(\mathbf{x}_k) + \epsilon_k \end{cases} \quad (9)$$

with  $\eta_k$  and  $\epsilon_k$  being stochastic vector variables, and  $a_{k+1,k}$  and  $h_k$  being vector functions, the state estimation problem has been studied with the set-theoretic approach [31], [32]. Comparisons with the Bayesian state estimation approach can be found in [33] and [34]. The recursive estimation of the state according to this framework aims to define a set of states consistent with the set of possible values of the input control and with uncertainty about the starting conditions, the process noise and the measurement noise. This approach is also known as “set-membership” estimation [35]. Unlike the stochastic setting, the uncertainty is defined in this case in terms of bounds imposed on the range of the control and noise signals. If the reference system is still described by the equations in (9), with  $\mathbf{u}$  being the bounded input control, and  $\eta$  and  $\epsilon$  being, respectively, the bounded process noise and the bounded measurement noise, the goal is to calculate a recursive estimate of  $\mathcal{X}_k^e$ , the set to which the state  $\mathbf{x}$  needs to belong at time  $k$  based on the observed measurements and on the predictions up to instant  $k$ .

Given the set estimate  $\mathcal{X}_k^e$ , the calculation of the new set  $\mathcal{X}_{k+1}^e$  at instant  $k+1$  occurs in two phases.

1) A prediction phase

$$\mathcal{X}_{k+1}^p = a_{k+1,k}(\mathcal{X}_k^e, \mathcal{U}_k) + \mathcal{H}_k \quad (10)$$

where  $\mathcal{U}_k$  and  $\mathcal{H}_k$  are the bounding sets for  $\mathbf{u}_k$  and  $\eta_k$ , respectively, which generates the set  $\mathcal{X}_{k+1}^p$  of all possible states obtained by the evolution of  $\mathcal{X}_k^e$  according to the dynamics of the process.

2) A fusion phase in which the new bounding set is the prediction set refined by the measurement-compatible set  $\mathcal{X}_{k+1}^z$ :

$$\mathcal{X}_{k+1}^e = \mathcal{X}_{k+1}^p \cap \mathcal{X}_{k+1}^z \quad (11)$$

where denoting by  $\mathcal{E}_{k+1}$  the bounding set for  $\epsilon_{k+1}$  and  $\mathbf{z}_{k+1}$  the measurement at time  $k+1$ , the measurement-compatible set is

$$\mathcal{X}_{k+1}^z = \{\mathbf{x} : \mathbf{z}_{k+1} = h_{k+1}(\mathbf{x}) + \epsilon \text{ for some } \epsilon \in \mathcal{E}_{k+1}\}. \quad (12)$$

In general, the intersection of sets provided by the iteration state observer may not return a set described by a simple analytical form. As a result, the calculation of the sets  $\mathcal{X}$  is usually not simple and it is often accomplished by using simpler out-bounding geometric objects (typically ellipsoids, hyperboloids, or polytopes) [32].

For specific classes of systems and under given assumptions, however, it is possible to derive a particularly advantageous form for the compatibility sets. We shall examine here the case in which the motion of a point system is described by a dynamic equation, linear or nonlinear, affine in the control input and with force control bounded in amplitude and constant between observation instants.

We consider the problem of estimating the 2-D position  $\mathbf{r} = [r_x, r_y]^T$  of an object whose motion is modeled by a linear dynamical system of the second order, described by (2), with  $\mathbf{f} = [f_x, f_y]^T$ , control input  $|\mathbf{f}| < \bar{F}$  (an amplitude bounded force), and  $\lambda^- < \lambda < \lambda^+$  being an uncertain damping parameter. If the state is defined as  $\mathbf{x} = [\mathbf{r}^T, \dot{\mathbf{r}}^T]^T$ , we obtain the following continuous-time state-space system:

$$\begin{cases} \dot{\mathbf{x}}(t) = \begin{bmatrix} \mathbf{0}_2 & \mathbf{I}_2 \\ \mathbf{0}_2 & -\lambda \mathbf{I}_2 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} \mathbf{0} \\ \mathbf{f}(t) \end{bmatrix} + \eta(t) \\ \mathbf{z}(t) = [\mathbf{I}_2 & \mathbf{0}_2] \mathbf{x}(t) + \epsilon(t) \end{cases} \quad (13)$$

with the column vectors  $\eta(t) \in \mathbb{R}^4$  and  $\epsilon(t) \in \mathbb{R}^2$  being, respectively, the process noise and the measurement noise. The corresponding approximated discrete-time system ( $T = t_{k+1} - t_k$  being the sampling period) is defined by

$$\begin{cases} \mathbf{x}_{k+1} = \begin{bmatrix} \mathbf{I}_2 & \frac{1-e^{-\lambda T}}{\lambda} \mathbf{I}_2 \\ \mathbf{0}_2 & e^{-\lambda T} \mathbf{I}_2 \end{bmatrix} \mathbf{x}_k + T \begin{bmatrix} \mathbf{0} \\ \mathbf{f}_k \end{bmatrix} + \eta_k \\ \mathbf{z}_k = [\mathbf{I}_2 & \mathbf{0}_2] \mathbf{x}_k + \epsilon_k. \end{cases} \quad (14)$$

Working in a 4-D space would still be a problem. We show that, fortunately, we can work in 2-D, by slightly adapting the approach. We preliminary note that in view of the first two equations in the discrete-time system above, if the velocity  $\dot{\mathbf{r}}$  and position  $\mathbf{r}$  at instant  $k$  were known, the displacement in a single sampling interval, without considering the force and the noise, would be given by

$$\mathbf{r}_{k+1} = \frac{1 - e^{-\lambda T}}{\lambda} \dot{\mathbf{r}}_k + \mathbf{r}_k. \quad (15)$$

We can take into account the effect of the force  $\mathbf{f}$  directly in this equation. Indeed, the force has an unknown direction and an unknown intensity with maximum value  $\bar{F}$ . Therefore, given  $\dot{\mathbf{r}}_k$  and  $\mathbf{r}_k$  at time  $k$ , at time  $k+1$  the object will be inside a circle of center  $\mathbf{r}_{k+1}$  and radius  $\rho$ :

$$\mathcal{C}(\mathbf{r}_{k+1}, \rho) = \left\{ r : \|\mathbf{r} - \mathbf{r}_{k+1}\| \leq \rho = \frac{\lambda T - 1 + e^{-\lambda T}}{\lambda^2} \bar{F} \right\}. \quad (16)$$

Note that the force  $|\mathbf{f}(t)| \leq \bar{F}$  can vary in the interval, according to the following result, whose proof is reported in the Appendix.

*Proposition 1:* If  $\dot{\mathbf{r}}_k$  and  $\mathbf{r}_k$  are exactly measured, then the set (16) is the exact prediction of all possible states at the next step. Hence, the prediction for the next step is

$$\mathbf{r}_{k+1} \in \mathcal{C}(\mathbf{c}_{k+1}, \rho), \text{ with } \begin{cases} \mathbf{c}_{k+1} = \frac{1-e^{-\lambda T}}{\lambda} \dot{\mathbf{c}}_k + \mathbf{c}_k \\ \rho = \frac{\lambda T - 1 + e^{-\lambda T}}{\lambda^2} \bar{F}. \end{cases} \quad (17)$$

The considered prediction set is optimal: the circle  $\mathcal{C}(\mathbf{c}_{k+1}, \rho)$  is the smallest region in which one is sure to find the target. In other words, regardless of the target escaping strategy, given its current position  $r_k$ , speed  $\dot{r}_k$ , and maximum acceleration  $\bar{F}$  (escaping capability), the target will be in this circle. However, any position inside  $\mathcal{C}(\mathbf{c}_{k+1}, \rho)$  is possible.

When we want to consider the measurement noise as well, the equations read:

$$\begin{cases} \mathbf{C}'_{k+1} = \mathbf{c}_{k+1} + \mathbf{C}_{k+1} = \frac{1-e^{-\lambda T}}{\lambda} (\dot{\mathbf{c}}_k + \mathbf{V}_k) + (\mathbf{c}_k + \mathbf{C}_k) \\ \rho = \frac{\lambda T - 1 + e^{-\lambda T}}{\lambda^2} \bar{F} \end{cases} \quad (18)$$

where  $\mathbf{C}_k$  is the set due to the additive noise on the position measurement,  $\mathbf{V}_k$  is the set due to the estimation error of the velocity, and  $\mathbf{C}'_k$  is the possible target positions set. Since it is assumed that the noise has bounded magnitude, these sets are in turn represented as circles centered on the estimated location and having radius equal to the maximum module  $\bar{V}$ ,  $\bar{C}$  for the velocity noise and the position noise, respectively. We can, therefore, transfer the component due to measurement noise in the computation of the radius of the circle:

$$\begin{cases} \mathbf{c}_{k+1} = \frac{1-e^{-\lambda T}}{\lambda} \dot{\mathbf{c}}_k + \mathbf{c}_k \\ \rho' = \rho + \frac{\lambda - e^{-\lambda T}}{\lambda} \bar{V} + \bar{C} = \frac{\lambda T - 1 + e^{-\lambda T}}{\lambda^2} \bar{F} + \frac{1-e^{-\lambda T}}{\lambda} \bar{V} + \bar{C}. \end{cases} \quad (19)$$

To compute the center  $\mathbf{c}_{k+1}$  of the circular set, the speed  $\dot{\mathbf{c}}_k$  needs to be known or suitably estimated: the Appendix gives an upper bound  $\bar{V}$  for the speed estimation error achieved using a filter.

Fig. 4 (top panel) shows a sequence of observations generated by the dynamic system, and the sets  $\mathcal{X}^p$  for the next two steps ( $k-1 \rightarrow k$  and  $k \rightarrow k+1$ ). We adopted the same value of  $\lambda$  used for the simulation of the process; the speed used for the prediction of the center of the circle is the exact one. Fig. 4 (bottom panel) shows the same sequence when the velocity used for the prediction of the circle is estimated (with the Euler approximation, see the Appendix) instead of being the exact one. The error on the estimated velocity causes an error in the estimated center of the prediction circle and, in this case, the observation falls outside the circle. This does not happen if the radius is enlarged taking into account the speed estimation error.

Fig. 5 (top panel) shows a single step of the observer ( $k \rightarrow k+1$ ) when the value  $\lambda = 1$  is used for the dynamics of the process, and different values of  $\lambda$  are used for the calculation of  $\mathcal{X}^p$  (i.e.,  $\lambda_1 \leq \lambda_2 \leq \lambda_3$  and  $\lambda_1 = 0.5$ ,  $\lambda_2 = 1$ ,  $\lambda_3 = 2$ ). Note that, in this case, the union of the sets is still a circle. Fig. 5 (bottom panel) shows the same step for  $\lambda = 0.5$  and  $\lambda_1 = 0.1$ ,  $\lambda_2 = 0.5$ ,  $\lambda_3 = 1$ . In this case, the union of the sets is no longer a circle.

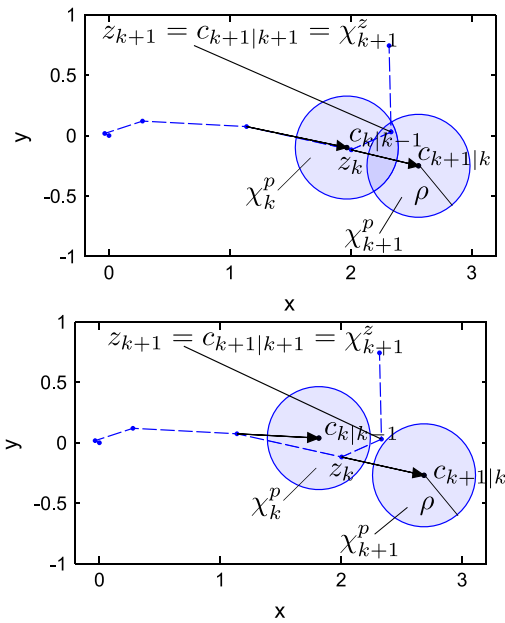


Fig. 4. Sets  $\mathcal{X}^p$  for a few steps when using the exact speed (top), and when the speed of the target is estimated (bottom). In both cases,  $\lambda = 1.5$ .

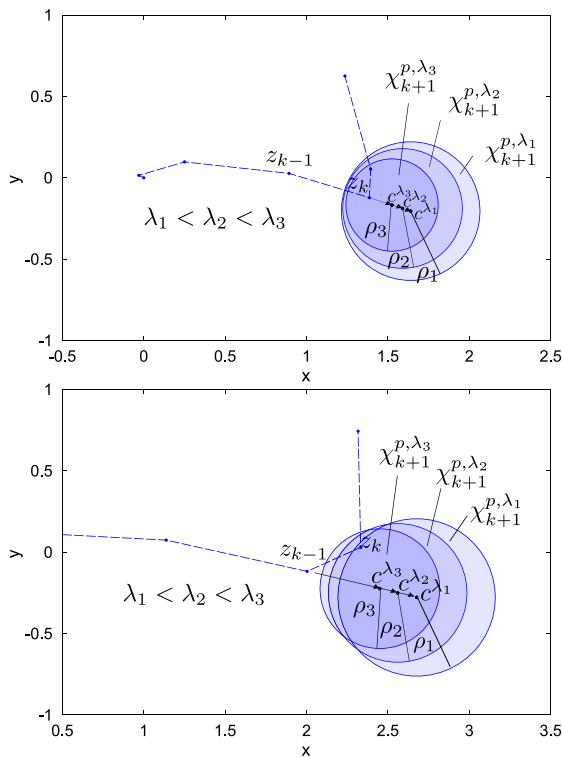


Fig. 5. Sets  $\mathcal{X}^p$  for a single step, exact velocity and different values of  $\lambda$  (top:  $\lambda_1 = 0.5$ ,  $\lambda_2 = 1$ , and  $\lambda_3 = 2$ ; bottom:  $\lambda_1 = 0.1$ ,  $\lambda_2 = 0.5$ , and  $\lambda_3 = 1$ ).

### C. Position and Geometry Adaptation of the Mobile Agents Cluster

The target localization and agents cluster reconfiguration scheme is illustrated in Fig. 6. At each iteration of the process, the position of each mobile agent in the cluster must be updated by some strategy that takes into account the predicted target position, the present sensed signal, and the position

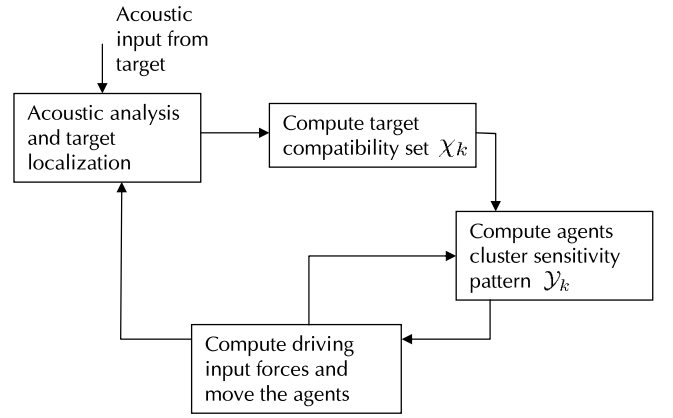


Fig. 6. Target localization and mobile agents cluster reconfiguration scheme.

of all other mobile agents. Recent algorithmic frameworks that can be used to address this class of problems include the general pattern search methods [36], or other more specific techniques gaining popularity in swarm robotics [37], such as particle swarm optimization or glowworm swarm optimization [26], [27].

The algorithm designed here to update the sensing agents position aims at reaching the maximum superposition between the forecasting set of possible target positions and the high sensitivity regions of the agents cluster as computed by the acoustic sensitivity function. The algorithm, exploiting the principles of the general pattern search method, attempts at selecting the optimal new cluster configuration by displacing one agent at a time, and assessing each time the beneficial effect on a reward function designed to predict the performance of the acoustic localization. During the search step at frame  $k$ , the displacement of sensing agent  $m$  is determined by an exploratory step  $u_k^{(m)} \in (\Delta_k^{(m)} P_k^{(m)})$ , whose direction is defined by the pattern  $P_k^{(m)} = BC_k^{(m)}$  [ $B$  being the bias matrix and  $C_k^{(m)}$  being the generating matrix] and whose magnitude is determined by  $\Delta_k^{(m)}$  [36].

1) *Reward Function Design and Assessment*: Since it is reasonable to assume that the best acoustic localization performance would be reached when the cluster of sensing agents has maximum sensitivity in the region where the target is expected to be found, we define the reward function as

$$f_k = f(\mathcal{X}_k, \mathcal{Y}_k, \delta_{w,k}(\cdot)) = \frac{|\mathcal{X}_k \cap \mathcal{Y}_k|}{N_{\mathcal{X},k}} \cdot \sum_{\mathbf{r} \in (\mathcal{X}_k \cap \mathcal{Y}_k)} \delta_{w,k}(\mathbf{r}) \quad (20)$$

where  $\mathcal{X}_k$  is the target prediction set at time  $k$ ,  $|A|$  denotes the cardinality of the set  $A$ ,  $N_{\mathcal{X},k} = |\mathcal{X}_k|$ ,  $\mathcal{Y}_k$  is the cluster sensing region, and  $\delta_{w,k}(\mathbf{r})$  is the sensitivity pattern. Note that the sensitivity function  $\delta_{w,k}(\mathbf{r})$ , if not weighted by the average distance between  $\mathbf{r}$  and the sensors, would increase its value as the sensors move far away due to the increment of the number of intersecting hyperbolas in  $\mathbf{r}$ . When referring to a target prediction set with circular symmetry, this would imply that a circular sensor cluster surrounding the target prediction set would increase its sensitivity with respect to the target as the sensors move far away, which would be unrealistic. By

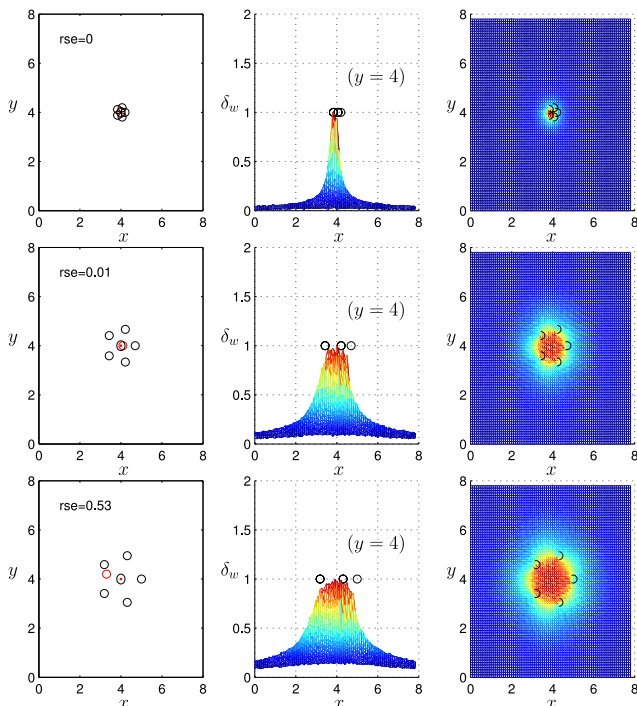


Fig. 7. Acoustic map related to a circular cluster configuration, for different radii.

weighting the sensitivity pattern by the inverse of the distance from the acoustic target, we take into account also the sound intensity attenuation, which has the effect of compensating the increment of intersections. Fig. 7 shows the effect of weighting the sensitivity pattern by the distance.

Next, we should assess the reward function to ensure that it correctly predicts the localization accuracy of an acoustic target in terms of a desired performance measure. In other words, we want the reward function to provide a maximum for the sensing agents cluster configuration that provides the smallest localization root-mean-square error (RMSE) or largest accuracy rate (AR). In that case, we might effectively use the reward function in an iterated positioning procedure instead of the actual localization process, which is known to be computationally intensive. Since the reward function is built upon the sensing agents cluster sensitivity function, the search for its maximum in terms of sensing agent positions, for a given target prediction set, has no simple analytical solution. We thus conducted numerical investigations to gain insights on this issue for prediction sets of circular shape. We provide in Fig. 8 an analysis conducted on a circular cluster having center equal to that of a static prediction circular set, in which the target source is randomly positioned at each new repetition of the experiment. The localization step was repeated on 100 random source positions within the circle, for each cluster radius tested, that is, 20 samples in the range  $[0.01, 1]$  m, when the target prediction set radius was set to 0.2, 0.3, 0.4, 0.5, and 0.6 m. It can be seen how the localization RMSE with this circular symmetry setting is lower on average when the sensors are positioned on the circumference of the prediction set, and how the reward function has a maximum on that value.

2) *Iterative Cluster Optimization and Target Tracking Algorithm:* At each new processing frame, the acoustic target

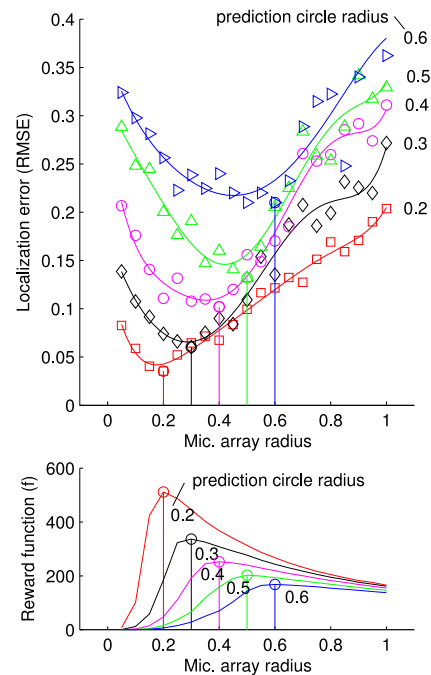


Fig. 8. Analysis of the reward function properties compared to acoustic localization performance, for a circular prediction set and target sources randomly positioned within the circular set.

gets localized through the microphone cluster, using its present configuration. Next, the best positions that the sensing agents should occupy at the next frame, given the prediction of the possible states of the target, are computed through the iterative optimal positions search. During the agents optimal positions search, the exploratory displacement of the  $m$ th sensing agent requires that  $\mathcal{Y}$ ,  $\delta_w(\mathbf{r})$ , and the reward function  $f$ , are updated given the new cluster configuration. We will call these updates  $\mathcal{Y}^{(m)}$ ,  $\delta_w^{(m)}(\mathbf{r})$ , and  $f^{(m)}$ , respectively. The general pattern search algorithm (GPSA) used for the cluster optimization is described in Algorithm 1.

3) *Positioning Algorithm Assessment in Static Conditions:* The cluster geometry optimization algorithm was first assessed with respect to a number of target prediction sets with different shapes and in static conditions, that is, when the target prediction set is not moving during position adaptation. The iterative procedure is thus aimed at finding the best sensing agent positions in terms of the reward function. Fig. 9 illustrates three examples in which the agents are adapted to a circular shaped set with radius 0.5 m (first row), and to two noncircular sets resulting from the union of circular sets with different radii (second and third row). In Fig. 10, the reward function is shown for each case.

The fact that the heuristic search leads to this particular arrangement where the sensors lie around the prediction region might sound obvious, but it is not. As said before, the minimization of (20) depends on the rather complex spatial distribution of hyperboloid intersections generated by the sensors (even for the relatively small number of sensors considered here). A rigorous analysis of the hyperbolic solution for sensor arrays has been conducted in a few cases, showing that the area in which the hyperboloid intersections concentrate



### Algorithm 1 Target Tracking and Sensing Agents Position Search Algorithm (GPSA)

---

$M$ : number of sensing agents  
 $\mathbf{r}_0^{(m)}$ ,  $m = 1, \dots, M$ : initial positions  
Initialize  $C_0^{(m)}$  and  $\Delta_0^{(m)}$   
**for**  $k = 1, 2, \dots$  **do**  
  Estimate the target position  $\hat{\mathbf{r}}_k^t$  by eq. (6)  
  Compute the target prediction set  $\mathcal{X}_{k+1}$   
  Compute  $f^{(0)}(k+1)$  by eq. (20)  
  **for**  $m = 1$  to  $M$  **do**  
    Compute move  $u_k^{(m)} \in \Delta_k^{(m)} P_k^{(m)}$   
    Compute  $\hat{\mathbf{r}}^{(m)}$  by eq. (1)  
    Compute  $\mathcal{Y}_{k+1}^{(m)}$ , and  $\delta_{w,k+1}^{(m)}(\mathbf{r})$ , due to changed positions  
    Compute  $f^{(m)}(k+1)$  using  $\mathcal{Y}_{k+1}^{(m)}$ , and  $\delta_{w,k+1}^{(m)}(\mathbf{r})$   
    **if**  $\rho_k^{(m)} = (f(k+1) - f^{(m)}(k+1)) < 0$  **then**  
       $\mathbf{r}_{k+1}^{(m)} = \hat{\mathbf{r}}^{(m)}$   
    **else**  
       $\mathbf{r}_{k+1}^{(m)} = \mathbf{r}_k^{(m)}$   
    **end if**  
    Compute  $C_{k+1}^{(m)}$  and  $\Delta_{k+1}^{(m)}$   
  **end for**  
   $f(k+1) = f^{(M)}(k+1)$   
**end for**

---

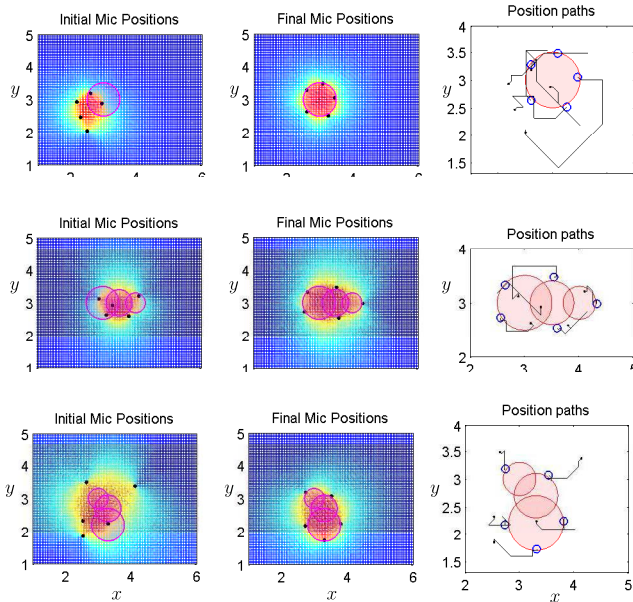


Fig. 9. GPSA optimization with respect of a circular target region with radius 0.5 m (first row), of a target region given by the union of three aligned circular regions (second row), and of a target region given by the union of three nonaligned circular regions (third row). Left, middle, and right plots are, respectively, the initial sensing agent positions, the final agent positions after 15 GPSA iterations, and the sensing agent paths during optimization.

can be localized in the central region of triangular arrays, as in [38], or of square arrays, as in [39]. However, to the best of our knowledge, no general analytical description of the spatial pattern of the intersections is already known, nor it is trivial to derive, even for simple symmetric array shapes. However, it is known that for each pair of acoustic sensors in a network, its higher TDOA resolution area is the (cone-shaped) region perpendicular to that pair, and the resolution decreases as the

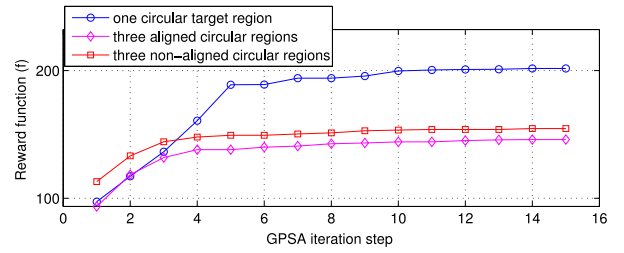


Fig. 10. GPSA optimization procedure: reward function at each iteration for the three cases (single circular region with radius 0.5 m, union of three aligned circular regions, and union of three nonaligned circular regions).

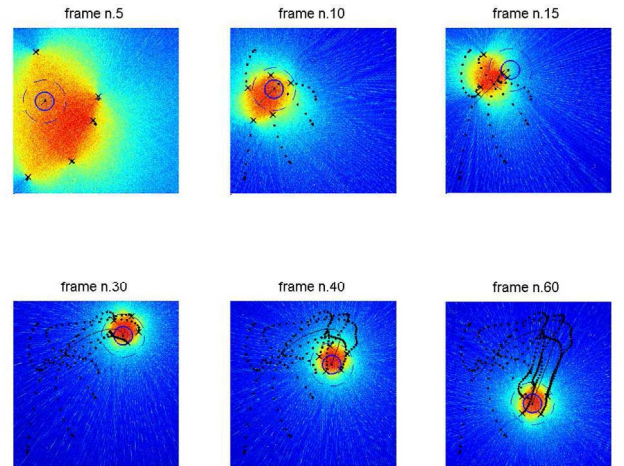


Fig. 11. Example of a target tracking sequence obtained through GPSA (selected frames): the target is represented by a black dot, its trajectory by a black continuous line, the target prediction set by the blue circle, and the sensing agent positions by the black crosses.

incidence angle increases (see [40]). It is then intuitive that, for a round-shaped sensed region, the best sensor arrangement that allows each pair to look frontally at the region of interest is an equally spaced, symmetric arrangement around that area. It is thus reasonable that, if the sensed area is a disk, a circular arrangement of the sensors is obtained.

4) *Sensor Cluster Reconfiguration During Tracking*: Finally, Fig. 11 illustrates the result of the tracking and of the cluster dynamic reconfiguration obtained through the GPSA algorithm when the acoustic target moves through the search region with a linear motion first, a turn to the right, and a linear motion at the end.

5) *Low Complexity Positioning Algorithm for Tracking*: It has been shown in Section III-C3 that when the prediction set has circular symmetry, the iterative positioning procedure converges to equally spaced positions on the prediction set circumference. This suggests that a low-complexity alternative to the tracking and positioning algorithm sketched above would be to drive the sensing agents on a set of equally spaced positions on the circumference of the next frame prediction set. This simplified tracking algorithm, formalized in Algorithm 2, will be also assessed in the next sections when the experimental assumptions will imply that the prediction sets are circular.

In a real-time tracking application, the algorithms presented require that three main tasks are carried out at each analysis

**Algorithm 2** Low Complexity Tracking and Mobile Agent Positioning Algorithm

---

```

M: number of sensing agents
 $\mathbf{r}_0^{(m)}$ ,  $m = 1, \dots, M$ : initial positions
for  $k = 1, 2, \dots$  do
  Estimate the target position  $\hat{\mathbf{r}}_k^s$  by eq. (6)
  Compute the target prediction set  $\mathcal{X}_{k+1}$ 
  Move the sensing agents to equally spaced positions on the
  circumference of the target prediction set  $\mathcal{X}_{k+1}$ 
end for

```

---

frame (whose duration is typically in the range of 10–30 ms), namely the acoustic target localization, the computation of the target prediction set, and the computation of the agent positioning. Acoustic localization, if performed through efficient methods as the ones mentioned earlier, can be normally achieved within a fraction of a typical analysis frame. Since the localization is performed on the data collected during the previous frame, localization estimates are in general assumed to be delayed by the duration of one frame. The computation of the target positions forecast and prediction sets is also computationally light. The computational cost of agents positioning, on the other hand, depends on the strategy adopted and may vary greatly. When using the low complexity positioning algorithm, it is safe to assume that all three tasks can be completed within one typical analysis frame using standard computing hardware. Using the time consuming GPSA algorithm in real time, on the other hand, might turn to be feasible only under specific hardware requirements.

#### IV. EXPERIMENTAL RESULTS

In this section, we report the results of experiments related to the 2-D localization and tracking of a sound-emitting target moving in a noisy and reverberant environment. The sensed acoustic data are generated through the numerical simulation of the acoustic propagation originating from an acoustic point source moving in the environment, and of the sensing agents displacement and signal acquisition. The sensing agents were allowed to move in the  $XY$  plane, while their height was kept constant at 1.7 m. We compare the performance of the proposed reconfiguration-based localization with a conventional SRP-PHAT localization operated through a fixed microphone linear array [28], and with the sparse sensor configuration method based on genetic algorithm optimization reported in [11]. The former method, although quite simplistic if compared to the one proposed here, is, however, still in use today in many practical situations where ASL is involved, and thus is a useful reference benchmark. The latter, a state-of-the-art sparse sensor configuration method, is more similar to ours, as the positioning is determined based on heuristic algorithms aimed at minimizing a fitness function built on a set of properties of the beampattern of the sensor network. With respect to the method we propose, however, it only relies on shape characteristics of the sensors beampattern (i.e., the main lobe amplitude and the greatest secondary lobe amplitude) and it does not exploit the target prediction sets. Localization performance is reported in terms of AR and RMSE for different sensing cluster reconfiguration strategies, for different

values of the SNR and reverberation time (RT60), and for different parameter settings of the target and mobile agent dynamics. Note that we have made the assumption that the noise corrupting the data from the acoustic source is spatially uncorrelated. More complex acoustic scenarios might be considered in principle, for example, in which other concurrent clutter signals are present as well. In such cases, it might be necessary to address the problem with a multisource localization approach, integrated with a clutter signals rejection method.

The localization performances are evaluated by means of Monte Carlo simulations, using ten run trials for each different condition under test. For each run trial, the tracking is conducted for several subsequent frames, and the localization performance measure for each different condition is averaged over all the analysis frames in the ten run trials. To simulate reverberant audio data due to the room acoustics, an image source method (ISM) model was used [41]. The room dimensions in the simulations were set to  $x = 9$  m,  $y = 6$  m, and  $z = 3$  m. The tests were conducted with different SNR values, which were obtained by adding mutually independent white Gaussian noise to each channel. A cluster of 5 sensing mobile agents was used. A speech signal was used as acoustic emission of the target, and its position in the search area was updated at each iteration using trajectories generated by feeding the dynamical model in (2).

The audio sampling frequency was set to  $f_s = 44.1$  kHz, and the signal analysis frame length was set to 4096 samples. The spatial resolution of the search grid was set to 0.01 m, and the sensitivity map resolution to 0.25 m.

Three different scenarios were investigated. First, we refer to a situation in which the target compatibility sets can be represented by circular 2-D regions, with constant radius over time. This condition can represent a situation in which the target dynamics is known and do not change over time (i.e., the  $\lambda$  parameter is constant), and the propagation time taken by the acoustic wave to propagate from the acoustic source to the sensors is also constant. Then, we look at possible situations in which the target compatibility sets are still circular, but their radius may vary in time. This behavior may represent the case in which acoustic propagation time from the source to the sensor cluster changes consistently from frame-to-frame, as it can be the case of sensing agents rapidly approaching (or receding) the target. Third, we look at a case in which the target compatibility sets should be represented by noncircular shaped regions, as it is the case of  $\lambda$  parameter assuming non-constant values in a finite range  $[\lambda_m, \lambda_M]$  (see the discussion referring to Fig. 5).

##### A. Circular Target Prediction Sets, Constant Radius

If we assume that the target dynamics is known and do not change over time, and if we neglect the propagation delays due to the changing target-cluster distance, it is reasonable to assume that the prediction set of the possible target positions can be modeled as a circular set with constant radius depending on its dynamics parameter  $\lambda$  and on the maximum possible input force. By further assuming that the target dynamics is known (or can be estimated), we report here

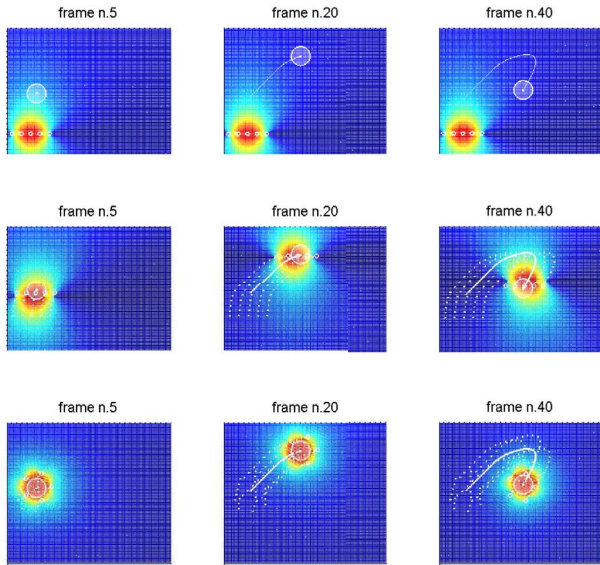


Fig. 12. Acoustic tracking and agents cluster reconfiguration. First row: control strategy C0 (static ULA); second row: strategy C1 (translating ULA); third row: strategy C4 (reconfigurable sensor cluster, with sensing agents positioning provided by Algorithm 2).

the localization performance of the proposed sensitivity-based reconfiguration and tracking scheme. The results are discussed by referring to the following control strategies: C0 [static uniform linear sensor array (ULA)], where the array linear geometry and its position are hold constant; C1 (translating ULA sensor array), where the array linear geometry is constant while the array can translate (although not rotate) along  $x$  and  $y$ -axes; C2–C4 (reconfigurable sensor network), where the sensors are carried by robotic mobile agents whose position can change according to some sensing agents positioning strategy. In this section, strategy C2 refers to agent positioning driven by the GAs-based strategy for sensor network configuration [11], strategy C3 refers to agent positioning through the GPS Algorithm 1, and strategy C4 will refer to agent positioning through Algorithm 2. We stress again that strategies C0–C2 do not make any use of the sensitivity map of the array and of the circular prediction sets (although these are included in the illustrations for comparison purposes), whereas strategies C3–C4 do exploit such features for agent positioning.

Fig. 12 illustrates three different sensor cluster control strategies: for each row, the plots report the array configuration and position, and the corresponding sensitivity pattern, during the tracking of an acoustic target moving along a piecewise linear trajectory.

Table I reports the parameters related to the dynamics of the moving objects in the simulation.

The results in Fig. 13 show the RMSE and AR of the localization performance for the different control strategies, averaged on ten repetitions each (the box plots show the median, the lower and upper quartiles, and the lower and upper extremes). Environmental conditions were set to  $\text{SNR} = 5$  dB and  $\text{RT60} = 0.7$  s (i.e., moderately adverse), and the parameters related to the sensor motion dynamics were set

TABLE I  
MOTION-RELATED SIMULATION PARAMETERS

Parameter	Value
$\lambda$ (target)	1 kg/s
$F$ (target)	360 N
$\lambda$ (sensor)	1 – 100 kg/s
$F$ (sensor)	2160 N
target velocity	1 – 20 Km/h
sensor velocity	1 – 20 Km/h
target–sensors distance	0.2 – 1.0 m

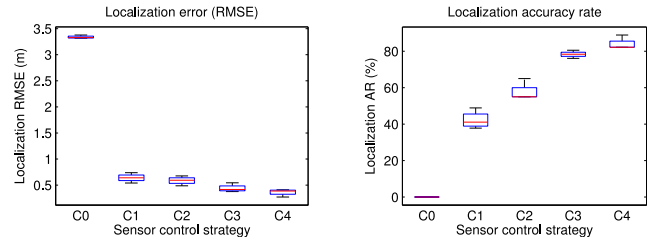


Fig. 13. Localization performances for the different control strategies: static linear array (C0), translating linear array (C1), reconfigurable circular sensor cluster by GAs-based algorithm (C2), reconfigurable circular sensor cluster by GPSA (C3), and reconfigurable sensor cluster by prediction circle targeting (C4). Environmental conditions are set to  $\text{SNR} = 5$  dB,  $\text{RT60} = 0.7$  s, and mobile agent dynamics are set to  $F_{\text{max}_{\text{mic}}} = 6F_{\text{max}_{\text{target}}}$  and  $\lambda_{\text{mic}} = 100\lambda_{\text{target}}$ .

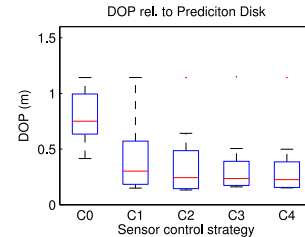


Fig. 14. DOP referred to the prediction disk, for the different agents positioning strategies.

to  $F_{\text{max}_{\text{mic}}} = 6F_{\text{max}_{\text{target}}}$  and  $\lambda_{\text{mic}} = 100\lambda_{\text{target}}$ . Moreover, the input force magnitude used in this experiment to drive the target was constrained to not exceed  $F_{\text{max}_{\text{target}}}/2$ , so that, at each new simulation frame, the actual target position falls within the circular prediction set, far from its boundary. This is to avoid to deal, in this first experiment, with the problem of localization error propagation from one frame to the next, causing the target to repeatedly fall outside the prediction set.

The accuracy of localization also depends on the limited number of TDOA values that a given spatial configuration of sensors yields, and a common way to relate the geometrical error due to the position of transmitter and receivers is provided by the dilution of precision (DOP) measure [42]. Since, in our method, the source is searched in a prediction region, it is interesting to look at the DOP related to this region for each different strategy. Fig. 14 shows that the strategies in which the sensors follow the target are the most effective in terms of the average DOP within the prediction disk.

### B. Factors Impacting on Performance

There are other factors that may impact on the localization performance, that is, the mobile agent dynamics, their number

and initial positioning with respect to the target, and the environmental acoustic conditions. In the supplementary material accompanying this paper, we show how the proposed method performs with respect to these factors.

### C. Doppler Effect

In an acoustic setting, when either the source, the receiver, or both move with respect to the other, the receiver may observe a change in the frequency of the source due to the Doppler effect. The frequency shift is a function of the relative velocity of the two along the source–receiver axis, and is perceived as an increase in frequency when the source and the receiver move toward each other and as a decrease in frequency when they move away from each other. In the context of ASL, the Doppler effect can be a problem when the source and/or the receiver move, as pointed out for example in [43], because it may alter the phase relations among the signals at the array, for example, when a different Doppler shift is observed at each sensor. However, it can also be exploited constructively, as it may provide useful cues to estimate the velocity of an acoustic source in motion [43] or even to localize it [44]. We briefly discuss here the case of a possibly negative impact on the localization performance and how this problem can be addressed. We, therefore, focus on a specific condition in which this effect can be highlighted, specifically when a different Doppler shift is observed at different channels of the array. To achieve this, we refer to a static linear uniform array, sufficiently large so that a source that moves parallel to it from left to right is seen to move away from the sensors to the left and to approach from the sensors to the right. The Doppler effect at each channel of the array is simulated by first computing the relative velocity of each receiver–source pair projected along their axis, by computing the related frequency shift, and by finally processing the signal through an interpolation operation as the one suggested in [45].

The results of the simulation are reported in Fig. 15 in which we compare, for different target velocities, the static ULA array with Doppler simulation (a), the static ULA array when the Doppler effect is not taken into account (b), and the reconfigurable circular agent cluster driven by strategy C4, with Doppler simulation (c). It can be seen how the Doppler effect introduces a degradation in the localization performance when the target–sensor relative velocities are different for each channel, as in the static ULA case with nearby moving source (a). In the agent cluster strategy, the simulated Doppler effect has a limited impact since it affects the localization only in the first few frames, when the cluster has not yet reached the prediction circle. After this time interval, since the velocity of the sensors approaches the velocity of the target, the Doppler effect vanishes. We finally note that in the other examples proposed in this paper, the Doppler effect has no relevant impact, since the source–receivers relative velocities are either all the same (static ULA array) or extremely low (reconfigurable circular sensing cluster (C4), and reconfigurable cluster with adaptive radius (that we will refer to as C4a)). In any case, the Doppler simulation operation is an invertible process, so that in principle it can be used to address Doppler shift compensation in those situations that may require it.

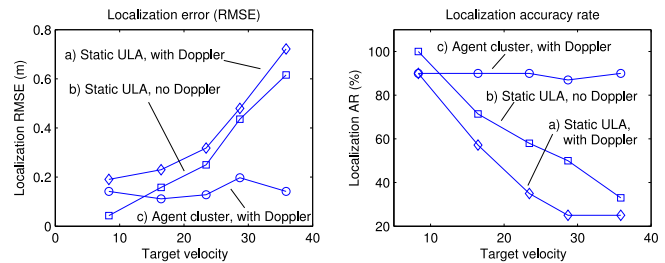


Fig. 15. Localization performance when the Doppler effect is taken into account.

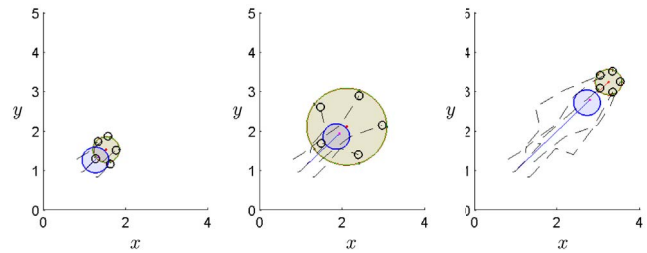


Fig. 16. Mobile agents cluster reconfiguration when modeling the loss of sensors data in a processing frame. Frames 3, 5, and 7 are shown. Data loss occurs at frame 5 (middle plot), and the circular prediction set radius increases accordingly to account for a doubled prediction interval.

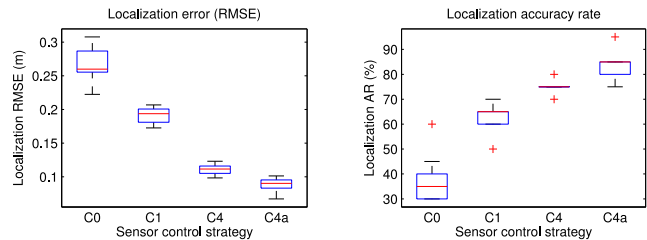


Fig. 17. Localization performances when the data are lost at frames 6, 11, 12, 13, and 15. Average error for static linear array (C0), translating linear array (C1), reconfigurable circular cluster (C4), and reconfigurable cluster with adaptive radius (C4a). Red plus signs are outlier data points.

### D. Circular Target Prediction Sets, Time-Varying Radius

In this section, we investigate a scenario involving time-varying prediction set radii, which can account for the possibility of data loss at certain frames. In these conditions, the prediction set of the possible target positions can be modeled as a circular set whose radius might become larger when data loss happens, to compensate for the longer time interval on which the target position must be predicted (e.g., two frames instead of one). Fig. 16 illustrates such a situation with an example in which data loss occurs at frame 5 and consequently the circular prediction set radius increases temporarily.

A set of numerical simulations were conducted in which data are lost at frames 6, 11, 12, 13, and 15. Fig. 17 shows the average error for different array positioning strategies, that is, static linear array (C0), translating linear array (C1), reconfigurable circular sensing cluster (C4), and reconfigurable cluster with adaptive radius (that we will refer to as C4a). It can be seen that, when the data loss occurs, a reconfigurable circular cluster with adaptive radius (C4a) provides the best performance.

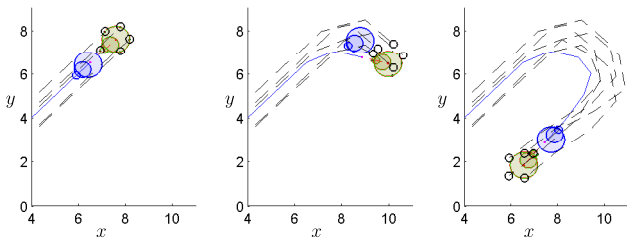


Fig. 18. Mobile agents cluster reconfiguration when modeling the uncertainty of target dynamics through the union of circular prediction sets.

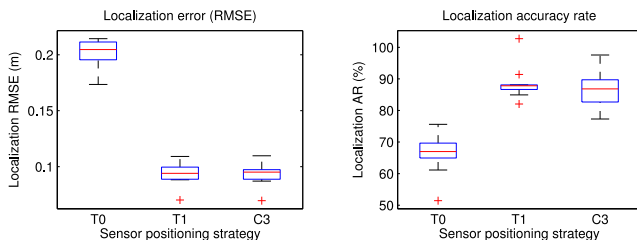


Fig. 19. Localization performances when different positioning strategies are adopted in case of uncertainty on the target dynamics. *T0*: sensing agents positioned on the circle corresponding to the mean target  $\lambda$ ; *T1*: sensing agents positioned on the minimum circle covering the union of target prediction circles; and *C3*: sensing agents positioned through GPSA iterative algorithm. Red plus signs are outlier data points.

### E. Noncircular Target Prediction Sets

We finally report the localization and tracking performance in the case of target uncertain dynamics, that is, when we assume that the parameter  $\lambda$  of the target falls within a known range. In this case, as we have discussed before, the prediction set of the possible target positions can be modeled as the union of the circular sets corresponding to  $\lambda$  values falling in the assumed range. Fig. 18 shows an example of such setting, when the damping parameter  $\lambda$  is assumed to vary in the range [4, 48], and the prediction set is modeled through the union of three circular sets corresponding to  $\lambda$  equal to 4, 20, and 48.

It is worth noting that, in this case, no simple mobile agent positioning strategy is easily found as the prediction set resulting from the union of circular sets is in general not a circular set, nor it is necessarily symmetric. The adoption of an iterative optimization strategy is thus justified in this case. Fig. 19 reports the performance of the GPSA iterative optimization (*C3*) compared to a trivial strategy *T0* where the sensing agents are positioned on the circle corresponding to the mean value of the allowed  $\lambda$  range, thus resulting in possible prediction errors, and to a strategy *T1* in which the sensing agents are positioned on the minimum circumference covering the union of the circular sets related to the dynamical parameter range assumed for the target.

The strategies *T1* and *C3* outperform strategy *T0*, as expected, however, strategy *C3* based on iterative optimization does not seem to provide substantial advantage with respect to strategy *T1*, even though the latter results in a sensor network configuration that fits the target prediction set less accurately.

## V. CONCLUSION

The problem of controlling a reconfigurable cluster of mobile agents with the aim of tracking an acoustic target in

noisy environments has been discussed. This paper has been focused on the exploitation of the target position estimation and forecasting and of the acoustic sensitivity pattern of the cluster, to drive the positioning of the sensing agents of the cluster at each tracking step. Different sensing agent positioning strategies have been proposed and assessed through numerical simulations. Moreover, the application of the proposed tracking strategies has been explored in different situations that translate into different prediction set characteristics, including circular and noncircular shape, and time-variability of the shape during tracking. Experimental results have shown substantial performance improvements in the target location and tracking performance with respect to conventional fixed-array acoustic localization.

A number of issues remain to be investigated in the future studies. These include the collision avoidance between target and sensing agents, and the trajectory planning to avoid collisions with obstacles in the environment.

Finally, we note that this strategy can in principle be applied to a multitarget tracking problem by setting the prediction set to be sensed by the sensor cluster as the union of the prediction sets of the different targets, and by then operating a conventional multisource acoustic localization in that region.

## APPENDIX

### PROOF OF PROPOSITION 1

To get an analytical formula for the estimation set, given the initial position and speed, we note that, in view of the superposition principle for linear systems, the predicted state is achieved due to the superposed effect of: 1) the uniform motion of the agent starting from its actual initial position at its actual initial speed, with no force acting on it  $f = 0$  (this provides the center of the prediction circle  $c_{k+1}$ ) and 2) the motion due to the action of the force, starting from zero initial position and with zero speed.

Consider the Laplace transform of (2) to get

$$s^2 \mathbf{r}(s) + \lambda s \mathbf{r}(s) = \mathbf{f}(s) \quad (21)$$

and therefore

$$\mathbf{r}(s) = \frac{1}{s(s + \lambda)} \mathbf{f}(s). \quad (22)$$

The inverse Laplace transform can be bounded in norm as follows:

$$\begin{aligned} \|r(T)\| &= \left\| \int_0^T \frac{1}{\lambda} [1 - e^{-\lambda\tau}] \mathbf{f}(t - \tau) d\tau \right\| \\ &\leq \int_0^T \frac{1}{\lambda} |1 - e^{-\lambda\tau}| \|\mathbf{f}(t - \tau)\| d\tau \\ &\leq \bar{F} \int_0^T \frac{1}{\lambda} [1 - e^{-\lambda\tau}] d\tau = \frac{\lambda T - 1 + e^{-\lambda T}}{\lambda^2} \bar{F} = \rho. \end{aligned}$$

This expression of the radius is nonconservative (optimal). Indeed the target can apply a constant force  $\mathbf{f}$  of maximum amplitude  $\bar{F}$  to get  $\|r(T)\| = \rho$ , exactly, although we cannot predict the direction. For smaller forces  $\|r(T)\|$  will be smaller. We conclude that the target will be necessarily inside  $\mathcal{C}(c_{k+1}, \rho)$ , but its position can be any point in this circle and we cannot predict it.

*Speed Estimation Error:* It is possible to compute an estimation of the velocity from the acceleration, by using an integrator filter with a single pole in  $-1/\tau$ :

$$\hat{\mathbf{v}}(s) = \frac{s}{1 + \tau s} \cdot \frac{1}{s(s + \lambda)} \mathbf{f}(s). \quad (23)$$

By comparing the estimated velocity with the exact velocity, given by

$$\mathbf{v}(s) = s \cdot \frac{1}{s(s + \lambda)} \mathbf{f}(s) \quad (24)$$

we have the estimation error  $\delta \mathbf{v}(s) = \hat{\mathbf{v}}(s) - \mathbf{v}(s)$ :

$$\begin{aligned} \delta \mathbf{v}(s) &= \left( \frac{s}{1 + \tau s} - s \right) \cdot \frac{1}{s(s + \lambda)} \mathbf{f}(s) \\ &= \frac{-\tau s}{1 + \tau s} \cdot \frac{1}{(s + \lambda)} \mathbf{f}(s). \end{aligned} \quad (25)$$

This estimation procedure can be easily transposed in the discrete-time domain, using the Euler approximation.

An analytic expression can be derived for the set of admissible values of the estimated velocity. In the time domain, the product of two functions becomes the convolution of the respective inverse Laplace transforms

$$\delta \mathbf{v}(T) = \int_0^T P(\sigma) \mathbf{f}(t - \sigma) d\sigma \quad (26)$$

where  $P(t; \lambda, \tau) = \mathcal{L}^{-1}\{(-\tau s/1 + \tau s) \cdot (1/(s + \lambda))\}$ . Taking the absolute value, we have that

$$|\delta \mathbf{v}(T)| \leq \int_0^T |P(\sigma)| |\mathbf{f}(t - \sigma)| d\sigma \leq \bar{F} \int_0^T |P(\sigma)| d\sigma = \bar{V}. \quad (27)$$

Finally, by computing the integral, it is possible to obtain an upper bound for the speed estimation error. ■

## REFERENCES

- [1] C. Robin and S. Lacroix, "Multi-robot target detection and tracking: Taxonomy and survey," *Auton. Robots*, vol. 40, no. 4, pp. 729–760, Apr. 2016.
- [2] A. Khan, B. Rinner, and A. Cavallaro, "Cooperative robots to observe moving targets: Review," *IEEE Trans. Cybern.*, vol. 48, no. 1, pp. 187–198, Jan. 2018.
- [3] D. J. Pack, P. DeLima, G. J. Toussaint, and G. York, "Cooperative control of UAVs for localization of intermittently emitting mobile targets," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 39, no. 4, pp. 959–970, Aug. 2009.
- [4] J. Kim and J. L. Crassidis, "UAV path planning for maximum visibility of ground targets in an urban area," in *Proc. 13th Int. Conf. Inf. Fusion*, Jul. 2010, pp. 1–7.
- [5] V. Sridhar and A. Manikas, "Target tracking with a flexible UAV cluster array," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Washington, DC, USA, Dec. 2016, pp. 1–6.
- [6] M. R. Bai, J.-G. Ih, and J. Benesty, *Acoustic Array Systems: Theory, Implementation, and Application*, 1st ed. Singapore: Wiley, 2013.
- [7] A. Alexandridis and A. Mouchtaris, "Multiple sound source location estimation in wireless acoustic sensor networks using DOA estimates: The data-association problem," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 26, no. 2, pp. 342–356, Feb. 2018.
- [8] D. Salvati, C. Drioli, and G. L. Foresti, "A low-complexity robust beamforming using diagonal unloading for acoustic source localization," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 26, no. 3, pp. 609–622, Mar. 2018.
- [9] D. Salvati, C. Drioli, and G. L. Foresti, "Exploiting CNNs for improving acoustic source localization in noisy and reverberant conditions," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 2, no. 2, pp. 103–116, Apr. 2018.
- [10] Y. Dorfan, A. Plinge, G. Hazan, and S. Gannot, "Distributed expectation-maximization algorithm for speaker localization in reverberant environments," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 26, no. 3, pp. 682–695, Mar. 2018.
- [11] F. Le Courtois, J.-H. Thomas, F. Poisson, and J.-C. Pascal, "Genetic optimisation of a plane array geometry for beamforming. Application to source localisation in a high speed train," *J. Sound Vib.*, vol. 371, pp. 78–93, Jun. 2016.
- [12] D. Salvati and S. Canazza, "Incident signal power comparison for localization of concurrent multiple acoustic sources," *Sci. World J.*, vol. 2014, pp. 1–13, Feb. 2014.
- [13] D. Salvati, C. Drioli, and G. L. Foresti, "Incoherent frequency fusion for broadband steered response power algorithms in noisy environments," *IEEE Signal Process. Lett.*, vol. 21, no. 5, pp. 581–585, May 2014.
- [14] Y. Cho, D. Yook, S. Chang, and H. Kim, "Sound source localization for robot auditory systems," *IEEE Trans. Consum. Electron.*, vol. 55, no. 3, pp. 1663–1668, Aug. 2009.
- [15] D. Yook, T. Lee, and Y. Cho, "Fast sound source localization using two-level search space clustering," *IEEE Trans. Cybern.*, vol. 46, no. 1, pp. 20–26, Jan. 2016.
- [16] X. Li and H. Liu, "Sound source localization for HRI using FOC-based time difference feature and spatial grid matching," *IEEE Trans. Cybern.*, vol. 43, no. 4, pp. 1199–1212, Aug. 2013.
- [17] D. J. Mennill, M. Battiston, D. R. Wilson, J. R. Foote, and S. M. Doucet, "Field test of an affordable, portable, wireless microphone array for spatial monitoring of animal ecology and behaviour," *Methods Ecol. Evol.*, vol. 3, no. 4, pp. 704–712, 2012.
- [18] M. Crocco, S. Martelli, A. Trucco, A. Zunino, and V. Murino, "Audio tracking in noisy environments by acoustic map and spectral signature," *IEEE Trans. Cybern.*, vol. 48, no. 5, pp. 1619–1632, May 2018.
- [19] F. Blanchini, F. A. Pellegrino, and L. Visentini, "Control of manipulators in a constrained workspace by means of linked invariant sets," *Int. J. Robust Nonlin. Control*, vol. 14, nos. 13–14, pp. 1185–1205, 2004.
- [20] N. Ceccarelli, M. Di Marco, A. Garulli, and A. Giannitrapani, "A set theoretic approach to path planning for mobile robots," in *Proc. 43rd IEEE Conf. Decis. Control*, vol. 1, Dec. 2004, pp. 147–152.
- [21] D. Simon, J. Löfberg, and T. Glad, "Reference tracking MPC using terminal set scaling," in *Proc. IEEE 51st Annu. Conf. Decis. Control (CDC)*, Dec. 2012, pp. 4543–4548.
- [22] S. Gazor and Y. Grenier, "Criteria for positioning of sensors for a microphone array," *IEEE Trans. Speech Audio Process.*, vol. 3, no. 4, pp. 294–303, Jul. 1995.
- [23] X. Wang, M. Amin, and X. Cao, "Analysis and design of optimum sparse array configurations for adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 66, no. 2, pp. 340–351, Jan. 2018.
- [24] R. Macho-Pedroso, F. Domingo-Perez, J. Velasco, C. Losada-Gutierrez, and J. Macias-Guarasa, "Optimal microphone placement for indoor acoustic localization using evolutionary optimization," in *Proc. Int. Conf. Indoor Position. Indoor Navig. (IPIN)*, Oct. 2016, pp. 1–8.
- [25] C. E. Y. Dorfan, S. Gannot, and P. A. Naylor, "Speaker localization with moving microphone arrays," in *Proc. 24th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2016, pp. 1003–1007.
- [26] J. M. Hereford, M. Siebold, and S. Nichols, "Using the particle swarm optimization algorithm for robotic search applications," in *Proc. IEEE Swarm Intell. Symp.*, Apr. 2007, pp. 53–59.
- [27] K. N. Kaipa and D. Ghose, "Experiments using physical simulations and real robots," in *Glowworm Swarm Optimization (Studies in Computational Intelligence)*, vol. 698. Cham, Switzerland: Springer, 2017, pp. 133–155.
- [28] J. H. Dibiase, H. F. Silverman, and M. S. Brandstein, "Robust localization in reverberant rooms," in *Microphone Arrays: Signal Processing Techniques and Applications*. Heidelberg, Germany: Springer, 2001, pp. 157–180.
- [29] M. Cobos, A. Marti, and J. J. Lopez, "A modified SRP-PHAT functional for robust real-time sound source localization with scalable spatial sampling," *IEEE Signal Process. Lett.*, vol. 18, no. 1, pp. 71–74, Jan. 2011.
- [30] D. Salvati, C. Drioli, and G. L. Foresti, "Exploiting a geometrically sampled grid in the steered response power algorithm for localization improvement," *J. Acoust. Soc. America*, vol. 141, no. 1, pp. 586–601, 2017.
- [31] P. L. Combettes, "The foundations of set theoretic estimation," *Proc. IEEE*, vol. 81, no. 2, pp. 182–208, Feb. 1993.
- [32] F. Blanchini and S. Miani, *Set-Theoretic Methods in Control*. Basel, Switzerland: Birkhäuser, 2015.

- [33] M. Baum and U. D. Hanebeck, "Extended object tracking based on set-theoretic and stochastic fusion," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 4, pp. 3103–3115, Oct. 2012.
- [34] V. Klumpp, B. Noack, M. Baum, and U. D. Hanebeck, "Combined set-theoretic and stochastic estimation: A comparison of the SSI and the CS filter," in *Proc. 13th Conf. Inf. Fusion*, Edinburgh, U.K., Jul. 2010, pp. 1–8.
- [35] M. Milanese and A. Vicino, "Optimal estimation theory for dynamic systems with set membership uncertainty: An overview," *Automatica*, vol. 27, no. 6, pp. 997–1009, 1991.
- [36] T. G. Kolda, R. M. Lewis, and V. Torczon, "Optimization by direct search: New perspectives on some classical and modern methods," *SIAM Rev.*, vol. 45, no. 3, pp. 385–482, 2003.
- [37] Y. Tan and Z.-Y. Zheng, "Research advance in swarm robotics," *Defence Technol.*, vol. 9, no. 1, pp. 18–39, 2013.
- [38] M. Compagnoni, R. Notari, F. Antonacci, and A. Sarti, "A comprehensive analysis of the geometry of TDOA maps in localization problems," *Inverse Problems*, vol. 30, no. 3, 2014, Art. no. 035004.
- [39] I. Marković and I. Petrović, "Speaker localization and tracking with a microphone array on a mobile robot using von Mises distribution and particle filtering," *Robot. Auton. Syst.*, vol. 58, no. 11, pp. 1185–1196, Nov. 2010.
- [40] E. A. P. Habets and P. C. W. Sommen, "Optimal microphone placement for source localization using time delay estimation," in *Proc. Annu. Workshop Circuits Syst. Signal Process. (ProRISC)*, Veldhoven, The Netherlands, Nov. 2002, pp. 284–287.
- [41] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. America*, vol. 65, no. 4, pp. 943–950, 1979.
- [42] J. D. Bard and F. M. Ham, "Time difference of arrival dilution of precision and applications," *IEEE Trans. Signal Process.*, vol. 47, no. 2, pp. 521–523, Feb. 1999.
- [43] M. Basiri, F. Schill, P. U. Lima, and D. Floreano, "Robust acoustic source localization of emergency signals from micro air vehicles," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2012, pp. 4737–4742.
- [44] A. Schasse, C. Tendency, and R. Martin, "Source localization based on the doppler effect," in *Proc. Int. Workshop Acoust. Signal Enhanc. (IWAENC)*, Sep. 2012, pp. 1–4.
- [45] J. O. Smith, S. Serafin, J. Abel, and D. Berners, "Doppler simulation and the Leslie," in *Proc. 5th Int. Conf. Digit. Audio Effects DAFX*, Hamburg, Germany, 2002, pp. 1–8.



**Carlo Drioli** (M'03) received the Laurea degree in electronic engineering and the Ph.D. degree in electronic and telecommunications engineering from the University of Padova, Padua, Italy, in 1996 and 2003, respectively.

He is an Assistant Professor with the University of Udine, Udine, Italy. He has been a Researcher with the Centro di Sonologia Computazionale, University of Padova; a Visiting Researcher with the KTH Royal Institute of Technology, Stockholm, Sweden, with the support of the European Community

through a Marie Curie Fellowship; a Researcher with the Institute of Cognitive Sciences and Technology, Italian National Research Council, Rome, Italy; and a Research Assistant and an Adjunct Professor with the Department of Computer Science, University of Verona, Verona, Italy. His current research interests include multimedia signal processing, sound and voice processing by physical modeling, speech analysis and synthesis, and array signal processing.

Dr. Drioli is a member of the Acoustical Society of America and the International Speech Communication Association.



**Giulia Giordano** received the B.Sc. and M.Sc. (Hons.) degrees in electrical engineering and the Ph.D. degree in systems and control theory from the University of Udine, Udine, Italy, in 2010, 2012, and 2016, respectively.

She is currently an Assistant Professor with the Delft University of Technology, Delft, The Netherlands. From 2016 to 2017, she was a Researcher with Lund University, Lund, Sweden. She was awarded a Delft Technology Fellowship grant in 2018. Her current research interest includes

around dynamical systems and networks.

Dr. Giordano was a recipient of the EEIC Ph.D. Award in 2016 and the Nonlinear Analysis: Hybrid Systems Best Paper Prize in 2017.



**Daniele Salvati** received the Laurea degree in environmental engineering from the Sapienza University of Rome, Rome, Italy, in 2003, the master's degree in sound engineering from the Tor Vergata University of Rome, Rome, Italy, in 2006, and the Ph.D. degree in multimedia communication from the University of Udine, Udine, Italy, in 2012, with a research on acoustic source localization using microphone arrays.

Since 2013, he has been a Post-Doctoral Researcher with the Department of Mathematics, Computer Science and Physics, University of Udine. He was a System and Audio Consultant to many information technology companies from 2001 to 2008. His current research interests include audio and acoustic signal processing and multimedia communications.



**Franco Blanchini** (SM'03) was born in Legnano, Italy, in 1959. He received the Laurea degree in electrical engineering from the University of Trieste, Trieste, Italy, in 1984.

He is the Director of the Laboratory of System Dynamics, University of Udine, Udine, Italy. He has coauthored the book entitled *Set Theoretic Methods in Control* (Birkhäuser).

Mr. Blanchini was a recipient of the 2001 ASME Oil & Gas Application Committee Best Paper Award for his coauthored article "Experimental Evaluation

of a High-Gain Control for Compressor Surge Instability," the 2002 IFAC Prize Survey Paper Award as the author of the article "Set Invariance in Control—A survey," *Automatica*, in 1999, for which he also received the High Impact Paper Award in 2017, and the 2017 NAHS Best Paper Award for his coauthored article "A Switched System Approach to Dynamic Race Modelling," *Nonlinear Analysis: Hybrid Systems*, in 2016. He was an Associate Editor of *Automatica*, from 1996 to 2006, and the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, from 2012 to 2016. Since 2017, he has been an Associate Editor of *Automatica*. He has been involved in the organization of several international events, in particular, he was the Program Vice-Chairman of the Conference Joint CDC-ECC 2005, Seville, Spain, the Conference CDC 2008, Cancun, Mexico, and the Conference CDC 2013, Florence, Italy; the Program Chairman of the Conference ROCOND, Aalborg, Denmark, 2012.



**Gian Luca Foresti** (SM'01) received the B.Sc. and M.Sc. (Hons.) degrees in electrical engineering and the Ph.D. degree in computer science and telecommunications from the University of Genoa, Genoa, Italy, in 1990 and 1994, respectively.

He is a Full Professor of computer science with the University of Udine, Udine, Italy, where he is the Director of the Department of Mathematics, Computer Science and Physics. He is the appointed Italian Member of the NATO RTO Information System Technology Panel. He has authored over 300

papers published in international journals and international conferences and has coedited several books in the field of multimedia and computer vision. His current research interests include computer vision and image processing, multisensor data and information fusion, pattern recognition, and deep neural networks.

Mr. Foresti was a recipient of the Best IEEE Vehicular Electronics Paper in 2002, the 2010 Best Paper Award at the International Conference on Distributed Smart Cameras, and the 2016 Best Industry Related Paper Award at the International Conference on Pattern Recognition. He has been the Guest Editor of a Special Issue of the PROCEEDINGS OF THE IEEE ON VIDEO COMMUNICATIONS, Processing and Understanding for Third Generation Surveillance Systems. He was the Finance Chair of the 11th IEEE Conference on Image Processing, the General Chair of the 16th International Conference on Image Analysis and Processing and the 8th IEEE Conference on Advanced Video and Signal Based Surveillance. He is a Fellow Member of IAPR.