Localization of Acoustic Sources Through the Fitting of Propagation Cones Using Multiple Independent Arrays

M. Compagnoni, P. Bestagini, F. Antonacci, A. Sarti, *Member, IEEE*, and S. Tubaro, *Member, IEEE*

***Abstract*—In this paper, we propose a novel acoustic source localization method that accommodates the general scenario of multiple independent microphone arrays. The method is based on a 3-D parameter space defined by the 2-D spatial location of a source and the range difference extracted from the time difference of arrival (TDOA). In this space, the set of points that correspond to a given range lie on a circle that expand as the range increases, forming a cone whose apex is the actual location of the source. In this parameter space, the lack of synchronization between arrays results in the fact that clusters of data associated to individual arrays are free to shift along the range axis. The cone constraint, in fact, enables the realignment of such clusters while positioning the cone vertex (source location), thus resulting in a joint data re-synchronization and source localization. We also propose a novel and general analysis methodology for swiftly assessing the localization error as a function of the TDOA uncertainties, which is remarkably accurate for small localization bias. With the aid of this method, simulations and experiments on real data, we show that the cone-fitting process offers excellent localization accuracy in the scenario of multiple unsynchronized arrays, as well as in simpler single-array scenarios, also in comparison with state-of-the-art techniques. We also show that the proposed method offers the desired flexibility for adapting to arbitrary geometries of microphone clusters.**

***Index Terms*—Microphone arrays, source localization, time difference of arrival.**

I. Introduction

ACOUSTIC source localization based on passive stationary sensor arrays is a problem of great interest in a variety of applications ranging from telecommunications to entertainment. Localizing acoustic sources is also a crucial aspect of acoustic source separation and echo cancellation methods. Numerous algorithms have been proposed over the past few decades, which use different types of acoustic measurements. Among them, those based on time differences of arrival (TDOAs) have proven particularly effective.

The maximum-likelihood estimator (MLE) (see [1]–[5]) is among the most popular approaches to TDOA-based localization thanks to its well-established advantage of asymptotic efficiency for a wide sample space. However, in order to apply the maximum-likelihood principle, a statistical characterization of the measurements is required. An alternate approach that does not require a prior statistical characterization of the measurements is the least squares estimator (LSE) [6]–[9], which attempts to minimize a squared error function that incorporates both source location and measurements. In [10], the authors provide exact solution procedures for efficiently estimating the source location based on squared range difference-least squares (SRD-LS) and squared range-least squares cost functions.

In the past few years, however, low-cost integrated digital arrays of microphones have begun to appear, paving the way to a new scenario in which multiple independent (unsynchronized) arrays are employed. Various examples of applications that use this pervasive sensing configuration can be found in the literature, for example in [11] and [12]. These articles describe a two-step solution for the localization problem, which accommodates multiple clusters of sensors. Source localization is based on the directions of arrival (DOAs) that are measured from two clusters of sensors. Triangulating DOAs, in fact, does not require synchronization of the two arrays, but it does imply an assumption of far field operation. If we are working in a near-field scenario, however, we need to resort to TDOAs, which normally require that the arrays be synchronized with each other. If array synchronization is not an option, synchronization of measurements is usually required.

The fact that TDOA measurements relative to different arrays might not be synchronized with each other is a rather common scenario to expect, particularly when working with numerous low-cost arrays. Furthermore, even if synchronization were possible, TDOAs obtained from distant microphones turn out to be less reliable than those obtained from closely spaced ones [13]. There is, therefore, a growing interest towards localization using TDOAs extracted from independent (i.e., unsynchronized) microphone arrays. In [14], source localization is approached through a ML estimation procedure using TDOAs measured at independent pairs of microphones, suitable also for near-field scenarios. A closed-form solution for the least squares error criterion was recently proposed in [15]. This paper is particularly interesting not just for its generality, but also because it partly addresses the case of multiple

Manuscript received August 04, 2011; revised December 22, 2011, March 05, 2012; accepted March 12, 2012. Date of publication April 03, 2012; date of current version May 07, 2012. This work was supported by the SCENIC project, within the Future and Emerging Technologies (FET) program, Seventh Framework Programme for Research of the European Commission, under FET-Open Grant 226007 The work of M. Compagnoni was supported by Imax s.r.l. —MaxMaraFashionGroup. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Boaz Rafaely.

M. Compagnoni is with the Dipartimento di Matematica "Francesco Brioschi", Politecnico di Milano, 20133 Milan, Italy (e-mail: marco.compagnoni@polimi.it).

P. Bestagini, F. Antonacci, A. Sarti, and S. Tubaro are with the Dipartimento di Elettronica ed Informazione, Politecnico di Milano, 20133 Milano, Italy (e-mail: bestagini@elet.polimi.it; antonacci@elet.polimi.it; sarti@elet.polimi.it; tubaro@elet.polimi.it).

Digital Object Identifier 10.1109/TASL.2012.2191958

sub-arrays (each with their own reference microphone and time origin). In fact, after extensively covering the single-array case, the authors propose a possible extension of their method for the multiple array scenario, with the goal of improving the robustness against noise.

In this paper, we address two critical issues within the scenario of localization based on multiple independent arrays: first we propose a localization methodology that performs the localization using a space-range representation. Second, we introduce a framework for the analytical assessment of the accuracy that can be achieved by a specific localization technique and for a given deployment of the individual arrays. As far as the source localization technique is concerned, we start with TDOA measurements and convert them into geometric constraints, which are defined in a 3-D space formed by the propagation distance (range difference extracted from the TDOA) and the geometric coordinates of the source. In this space, the set of points that correspond to a given range lie on a circle that expands as the range increases, forming a cone whose apex is the actual location of the source. Localizing the source, therefore, means finding the location of the apex of the cone through the minimization of some cost function. The cone representation is partially inspired by the light-cone principle, a well-known concept in the literature of relativistic physics [16].

We first propose an algorithm that performs 2-D source localization with independent arrays through the minimization of two cost functions with meaningful geometric interpretations in the space-range reference frame. We then specialize the concept to the case of a single array in order to show the strong relationships and the differences existing between the proposed cost functions and state-of-the-art techniques. We study the 2-D localization because it allows us to visualize the approach in the 3-D geometrical space. The method, however, can be rather straightforwardly generalized to the 3-D localization case.

In a general scenario, the deployment of the individual arrays are not expected to be optimized for localization purposes, as it might be subject to constraints of a different nature (e.g. architectural). It is therefore important to be able to accurately and swiftly predict the localization accuracy for a given geometry of clusters. In the literature, researchers typically resort to Monte-Carlo simulations or to the Cramer–Rao Lower Bound (CRLB) (e.g., [17]). The first alternative is often time-demanding, as for each location of interest of the source a huge number of simulations is required. For this reason, CRLB is used as a theoretical lower bound of the achievable accuracy under the hypothesis that no bias is introduced in the estimation. In this paper, we propose a novel method that combines the advantages of Monte Carlo simulations (algorithm-dependent prediction of the accuracy) and CRLB (low computational cost) to predict the root mean square localization error, which consists of a linear relationship between the covariance matrix of the measurement errors and localization errors that is valid under the same hypothesis of CRLB (i.e., small bias). The proposed analysis, however, goes a bit beyond CRLB, as it predicts the error achievable by a specific localization technique, assumed that it is based on the minimization of a cost function. We remark, moreover, that the proposed error propagation analysis can be used also not only in the context of source localization, but also for the estimation of other variables. For this purpose, a Matlab toolbox is available online [18].

The paper is structured as follows. Section II presents the problem formulation and introduces the adopted 3-D parameter space. Section III introduces two cost-functions for the multiple arrays case and then shows that other state-of-the-art single-array localization algorithms are easily represented in the space-range framework. In Section IV, we introduce the theoretical study of the error propagation from measurements to source location estimation. Section V presents some results from simulations and experiments aimed at validating the theoretical error analysis; at presenting the influence of the number of sub-arrays on localization accuracy; at showing a comparison between the proposed algorithms and those found in literature in different scenarios. Finally, in Section VI we draw some concluding remarks.

II. Problem Formulation

As anticipated above, for reasons of representational clarity, in this paper we focus on the localization of a single acoustic source in a 2-D geometry. The proposed framework, however, can be rather straightforwardly extended to the 3-D space. Let us consider L independent microphone sub-arrays, each composed of $N_l + 1, l = 1, \dots, L$ sensors. The ith sensor in the lth sub-array is denoted with the index $i^{(l)}$ and it is located at coordinates $(x_i^{(l)}, y_i^{(l)})$. The sub-arrays are not assumed as being synchronized with each other. The first microphone in each sub-array is the local reference microphone, i.e., the TDOAs within each sub-array are computed between the $i^{(l)}$th sensor ($i \neq 0$) and the reference one ($i = 0$), while TDOAs between microphones belonging to different sub-arrays are not directly available because of the lack of synchronization. In order to derive the cone-based representation, let us assume for the moment that the source located at (x_S, y_S) produces a pulse at time 0 and that source and microphones are synchronized. The $i^{(l)}$th microphone receives the pulse at time $t_i^{(l)}$. Correspondingly, the propagation distance from the source to the $i^{(l)}$th microphone is $z_i^{(l)} = \eta t_i^{(l)}$, η being the sound speed. In the space-range representation, the sensor $i^{(l)}$ is at coordinates $\mathbf{x}_i^{(l)} = [x_i^{(l)}, y_i^{(l)}, z_i^{(l)}]^T$. We notice that by moving the source to a position with a different range the coordinate $z_i^{(l)}$ changes, which means that the space-range representation depends on the source position.

In a homogeneous and isotropic 2-D plane, a wavefront that propagates from a point-like source describes a circle centered in the source location, whose radius grows proportionally with time. In the novel space-range representation introduced here, each propagation distance from the source corresponds to a different value of z and, therefore, the expansion of the "propagation circle" originates a cone (see Fig. 1) whose apex lies in $\mathbf{x}_s$, whose angular aperture is 45°. This cone is therefore characterized by the equation

$$(x - x_S)^2 + (y - y_S)^2 = (z - z_S)^2. \tag{1}$$

Notice that a microphone co-located at the source position would have coordinates $(x_S, y_S, 0)$.

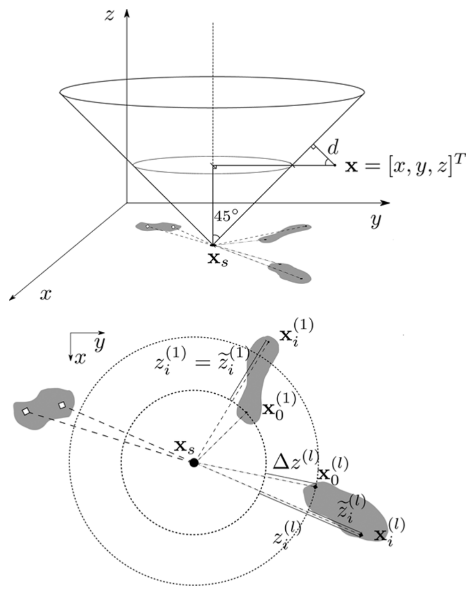

Fig. 1. Propagation cone in the space-range reference frame: x and y describe the object location, while the range z is proportional to the TDOAs with respect to the global reference microphone. The cone apex is located at $\mathbf{x}_S$, while the point $\mathbf{x}_i$ represents the ith microphone. This point is bound to lie on the cone surface but measurement errors displace points from the surface of the cone. We therefore define the distance d from $\mathbf{x}$ to the cone and use it in the cone fitting procedure. The relationship between measurements obtained by TDOAs ($\tilde{z}_i^{(l)}$) and resynchronized measurements ($\tilde{z}_i^{(l)} + \Delta z^{(l)}$) at different microphones clusters is also shown from a top view.

We now remove the assumption of synchronization between the source and the microphones and between microphones belonging to different arrays. The lack of synchronization between source and microphones means that the origin of the z-axis does not correspond to the source, but it is arbitrarily placed. Furthermore, the lack of synchronization between sub-arrays implies that TDOAs relative to microphones belonging to different sub-arrays are not available.

In order to use a representation that addresses both issues, we set the zero of the z-axis to correspond to the global reference microphone ($i = 0, l = 1$). This way the actual range differences $z_i^{(l)}$ are proportional to the TDOAs between the global reference microphone and the $i^{(l)}$th sensor, and the proportionality factor is the sound speed η. However, only $\tilde{z}_i^{(l)}, l = 1, \ldots, L$ $, i = 0, \ldots, N_l - 1$ are available, which are referenced to the local reference microphone. In order to go from $\tilde{z}_i^{(l)}$ to $z_i^{(l)}$ we define the difference of distance of propagation from the source to the global reference sensor and to the local reference sensor of the lth sub-array as $\Delta z^{(l)}, l = 1, \ldots, L$. The range difference of the $i^{(l)}$th sensor referred to the global reference sensor is therefore

$$z_i^{(l)} = \tilde{z}_i^{(l)} + \Delta z^{(l)}.$$

Fig. 1 represents the geometry of the problem. In particular, the top plot represents a scenario in which three sub-arrays (denoted with the shaded clouds) are present, and the bottom plot shows the corresponding situation in the 2-D geometry, where also the terms $\Delta z^{(l)}$ are represented. With this geometry in mind, we define the following problem.

Source Localization Problem: Given a set of independent (i.e., unsynchronized) arrays, find in the space-range reference frame the apex of the cone that best fits the $\mathbf{x}_i^{(l)}$ points. The first two coordinates of the apex of the cone are the searched source location.

In the next section we propose the localization algorithm based on this cone-fitting approach.

III. Localization of Acoustic Sources

In order to approach the above source localization problem, we introduce two error definitions that tell us how points in the space-range reference frame fit a prescribed cone. These definitions will then be used for constructing two different cone-based localization functions in the most general case of multiple independent arrays. After then, we will show that working with a single array, which is the situation that is addressed the most in the literature, can be treated as a special case of the multiple arrays one. Finally, we will discuss the problem of efficiently minimizing the derived cost functions.

A. Cone-Based Error Definitions

As discussed in Section II, the source localization problem can be formulated in terms of propagation cone fitting [see (1)].

We assume the TDOA measurement $\tau_i^{(l)}$ at sensor $i^{(l)}$ to be affected by noise

$$\hat{\tau}_i^{(l)} = \tau_i^{(l)} + \varepsilon_i^{(l)}. \quad (2)$$

$\tau_i^{(l)}$ being the nominal TDOA and $\varepsilon_i^{(l)}$ being an additive noise term. We assume errors associated to TDOAs measured on independent pairs of microphones to be uncorrelated. In reverberant environments TDOAs could be affected by noise that cannot be modeled as additive. In fact, TDOAs coming from reflective paths can be confused with the direct-path ones. In such scenarios, TDOA-disambigaution techniques (see for example [19]) could be used to discriminate between direct-path and reflective-path TDOAs before performing localization. We consider only the case of direct-path TDOAs. Because of the measurement noise, range differences $\tilde{z}_i^{(l)}$ are affected by noise as well, therefore we cannot expect the points $\mathbf{x}_i^{(l)}$ to lie on the cone, as assumed in (1). The localization problem, therefore, needs to be treated as a problem of fitting a cone from noisy measurements of points lying on it.

In order to assess how well a generic point $[x, y, z]^T$ lies on a cone with apex in $\mathbf{x}_S = [x_S, y_S, z_S]^T$ and with aperture 45° we consider two different metrics, which are the most natural to be used in such a fitting problem [20]:

- *Cone Equation*: a definition of the error that tells us how well (1) is satisfied is

$$\varepsilon_e = (x - x_S)^2 + (y - y_S)^2 - (z - z_S)^2. \quad (3)$$

- *Distance from the cone*: a definition of the error that accounts for the distance of the point $[x, y, z]$ from the cone surface is

$$\varepsilon_a = \sqrt{(x - x_S)^2 + (y - y_S)^2} - (z - z_S). \quad (4)$$

We notice that (4) is proportional to the distance d of a generic point $\mathbf{x}$ from the cone surface. With reference to Fig. 1 we obtain

$$d = \frac{\sqrt{2}}{2}\left|\sqrt{(x-x_S)^2+(y-y_S)^2}-(z-z_S)\right|. \tag{5}$$

The next step consists in defining a cost function that measures the ℓ^2-norm of any of the above errors and localizes the source as the point $\mathbf{x}_S$ that minimizes this cost function.

B. Multiple Independent Arrays

We now address the problem of localizing the acoustic source using multiple independent arrays. As mentioned above, we assume all of the TDOAs of an array to be referred to the local reference microphone. This means that we have all the $\tilde{z}_i^{(l)}$ measurements at our disposal. However, in order to be able to use the cone-based cost function, we need the terms $z_i^{(l)}$ obtained by referring all the measurements to a single global reference microphone. In order to attain this goal, we need to determine the displacements $\Delta z^{(l)}, l=1,\ldots,L$ between local reference microphones of different arrays and the global reference microphone. We accomplish this task by expressing $\Delta z^{(l)}$ as a function of the unknown $\mathbf{x}_S$.

We notice that

$$\begin{aligned}\Delta z^{(l)} = &\sqrt{\left(x_0^{(l)}-x_S\right)^2+\left(y_0^{(l)}-y_S\right)^2}+\\ &-\sqrt{\left(x_0^{(1)}-x_S\right)^2+\left(y_0^{(1)}-y_S\right)^2}\\ &\qquad l=2,\ldots,L\end{aligned} \tag{6}$$

while

$$\Delta z^{(l)} = 0, \quad l=1. \tag{7}$$

The expressions of the errors (3) and (4) for the point $\mathbf{x}_i$ therefore become

$$\begin{aligned}\varepsilon_{e,i}^{(l)} = &\left(x_i^{(l)}-x_S\right)^2+\left(y_i^{(l)}-y_S\right)^2\\ &-\left(\tilde{z}_i^{(l)}+\Delta z^{(l)}-z_S\right)^2\end{aligned} \tag{8}$$

$$\begin{aligned}\varepsilon_{a,i}^{(l)} = &\sqrt{\left(x_i^{(l)}-x_S\right)^2+\left(y_i^{(l)}-y_S\right)^2}\\ &-\left(\tilde{z}_i^{(l)}+\Delta z^{(l)}-z_S\right)\end{aligned} \tag{9}$$

and the source location can be estimated as

$$\hat{\mathbf{x}}_S = \underset{\mathbf{x}_S}{\operatorname{argmin}}\left(J_{\varepsilon_e}\right) = \underset{\mathbf{x}_S}{\operatorname{argmin}}\left(\boldsymbol{\varepsilon}_e^T\boldsymbol{\varepsilon}_e\right) \tag{10}$$

or

$$\hat{\mathbf{x}}_S = \underset{\mathbf{x}_S}{\operatorname{argmin}}\left(J_{\varepsilon_a}\right) = \underset{\mathbf{x}_S}{\operatorname{argmin}}\left(\boldsymbol{\varepsilon}_a^T\boldsymbol{\varepsilon}_a\right) \tag{11}$$

where

$$\boldsymbol{\varepsilon}_k = \left[\varepsilon_{k,0}^{(1)} \quad \cdots \quad \varepsilon_{k,N_1}^{(1)} \quad \varepsilon_{k,0}^{(2)} \quad \cdots \quad \varepsilon_{k,N_L}^{(L)}\right]^T$$

and the subscript k is replaced by a or e according to which error definition is adopted. We notice that by substituting $\Delta z^{(l)}, l=1,\ldots,N_l$ with their expressions as a function of the source position in (11) and (10) allows us to keep the number of unknown parameters to three ($[x_S, y_S, z_S]^T$), instead of $L+2$ (x_S, y_S, z_S and $\Delta z^l, l=2,\ldots,L$).

C. Special Case: The Single Array

In this section, we consider the case of having one array ($L=1$). This situation is widely addressed in the literature, as, for example, in [9], [10], and [21]. We show that the cone-based representation is useful also in this case and it sheds light on the relationship existing between the proposed approach and the methods existing in the literature.

We notice from (6) that $\Delta z^{(l)}=0$ when $l=1$; thus, the term $\Delta z^{(l)}$ disappears from (8) and (9). We omit for reasons of compactness the apex l in the notation. The two error definitions (8) and (9) become

$$\varepsilon_{e,i} = (x_i-x_S)^2+(y_i-y_S)^2-(\tilde{z}_i-z_S)^2 \tag{12}$$

$$\varepsilon_{a,i} = \sqrt{(x_i-x_S)^2+(y_i-y_S)^2}-(\tilde{z}_i-z_S). \tag{13}$$

The source is then localized as

$$\hat{\mathbf{x}}_S = \underset{\mathbf{x}_S}{\operatorname{argmin}}\left(J_{\varepsilon_e}^{(s)}\right) = \underset{\mathbf{x}_S}{\operatorname{argmin}}\left(\boldsymbol{\varepsilon}_e^{(s)T}\boldsymbol{\varepsilon}_e^{(s)}\right) \tag{14}$$

or

$$\hat{\mathbf{x}}_S = \underset{\mathbf{x}_S}{\operatorname{argmin}}\left(J_{\varepsilon_a}^{(s)}\right) = \underset{\mathbf{x}_S}{\operatorname{argmin}}\left(\boldsymbol{\varepsilon}_a^{(s)T}\boldsymbol{\varepsilon}_a^{(s)}\right) \tag{15}$$

where

$$\boldsymbol{\varepsilon}_k^{(s)} = \left[\varepsilon_{k,0} \quad \varepsilon_{k,1} \quad \cdots \quad \varepsilon_{k,N}\right]^T$$

and the subscript k is replaced by a or e according to which error definition is adopted. The index (s) denotes that we are working with a single array.

With reference to (12) it is interesting to notice that the cost function $J_{\varepsilon_e}^{(s)}$ matches that of SRD-LS if we set $\varepsilon_{e,0}=0$. Indeed, the terms that appear in (14) include also the error on the reference microphone, while this is discarded in SRD-LS. Differently from SRD-LS, in [21] and [22] the authors add the distance between source and reference sensor as an additional unknown. These similarities give support to the validity of our approach, showing that other techniques, although independently developed, can admit a geometrical reinterpretation based on the propagation cone idea.

D. Minimization

In order to efficiently minimize the four cost functions defined above, we use an iterative Taylor series expansion.

At this purpose we expand $\boldsymbol{\varepsilon}_e$ and $\boldsymbol{\varepsilon}_a$ with a first-order Taylor series expansion about the reference point $\mathbf{x}_{S,0} = [x_{S,0}, y_{S,0}, z_{S,0}]^T$ as

$$\boldsymbol{\varepsilon}_k \simeq \boldsymbol{\varepsilon}_k|_{\mathbf{x}_{S,0}} + \nabla\boldsymbol{\varepsilon}_k|_{\mathbf{x}_{S,0}} \cdot (\mathbf{x}_S - \mathbf{x}_{S,0}) \tag{16}$$

where $\mathbf{x}_{S,0}$ is the initial guess of the source position

$$\nabla\boldsymbol{\varepsilon}_k = \begin{bmatrix} \frac{\partial\varepsilon_{k,0}^{(1)}}{\partial x_S} & \frac{\partial\varepsilon_{k,0}^{(1)}}{\partial y_S} & \frac{\partial\varepsilon_{k,0}^{(1)}}{\partial z_S} \\ \vdots & \vdots & \vdots \\ \frac{\partial\varepsilon_{k,N_1}^{(1)}}{\partial x_S} & \frac{\partial\varepsilon_{k,N_1}^{(1)}}{\partial y_S} & \frac{\partial\varepsilon_{k,N_1}^{(1)}}{\partial z_S} \\ \frac{\partial\varepsilon_{k,0}^{(2)}}{\partial x_S} & \frac{\partial\varepsilon_{k,0}^{(2)}}{\partial y_S} & \frac{\partial\varepsilon_{k,0}^{(2)}}{\partial z_S} \\ \vdots & \vdots & \vdots \\ \frac{\partial\varepsilon_{k,N_L}^{(L)}}{\partial x_S} & \frac{\partial\varepsilon_{k,N_L}^{(L)}}{\partial y_S} & \frac{\partial\varepsilon_{k,N_L}^{(L)}}{\partial z_S} \end{bmatrix}$$

$$\mathbf{x}_S - \mathbf{x}_{S,0} = \begin{bmatrix} x_S - x_{S,0} \\ y_S - y_{S,0} \\ z_S - z_{S,0} \end{bmatrix} \tag{17}$$

and the subscript k is replaced by a or e according to which error definition is adopted. If we use (16) in (10), or (11), we readily obtain the update equation of the iterative minimization procedure

$$\hat{\mathbf{x}}_{S,v+1} = \hat{\mathbf{x}}_{S,v} - \left(\nabla\boldsymbol{\varepsilon}_k^T\big|_{\hat{\mathbf{x}}_{S,v}} \cdot \nabla\boldsymbol{\varepsilon}_k|_{\hat{\mathbf{x}}_{S,v}}\right)^{-1} \cdot \nabla\, \boldsymbol{\varepsilon}_k^T\big|_{\hat{\mathbf{x}}_{S,v}} \cdot \boldsymbol{\varepsilon}_k \tag{18}$$

where the symbol v is the iteration number. We assume

$$\hat{\mathbf{x}}_S = \hat{\mathbf{x}}_{S,v+1} \tag{19}$$

when $|\hat{\mathbf{x}}_{S,v+1} - \hat{\mathbf{x}}_{S,v}|$ is smaller or equal to a given threshold. The cone-based cost functions exhibit a smooth behavior; therefore, we resort to a randomly chosen starting point $\mathbf{x}_{S,0}$ within the area of interest. This iterative procedure turns out to converge quite rapidly to a solution. Through numerical simulations we show in Section V that convergence time is comparable to that of state-of-the-art techniques such as [10] or [15]. This makes this iterative estimation method suitable for real-time applications.

IV. Analysis of the Error Propagation

In this section, we present a novel approach to error propagation analysis that allows us to predict the impact of measurement errors on the minimization process. This formulation is based on introductory concepts of Catastrophe Theory [23], and can be applied to a wide range of situations and cost functions. In Section V, we will assess the accuracy of this analytical approach through numerical simulations. We also claim that it is possible to prove the equivalence of our error prediction with other different theoretical approaches to the characterization of the (asymptotic) first-order efficiency of a given estimator (see for example [24, Sec. 4.4]). This analysis, however, is beyond the scope of this paper, and therefore will not be included in it.

Let $f(\mathbf{x};\mathbf{c})$ be a generic cost function, with variables $\mathbf{x} = [x_1,\ldots,x_M]^T$ and parameters $\mathbf{c} = [c_1,\ldots,c_N]^T$. The parameters correspond to the experimental measurements (in our case the range differences) and the variables represent the object of the estimation (in our case the position of the source). Let $\mathbf{x}_0$ be the true location that we intend to determine and $\mathbf{c}_0$ the related error-free measurements. In a real situation we are given noisy measurements $\bar{\mathbf{c}} = \mathbf{c}_0 + \delta\mathbf{c}$, $\delta\mathbf{c}$ being the additive noise. Consequently, the new position of the minimum of $f(\mathbf{x};\bar{\mathbf{c}})$ becomes $\bar{\mathbf{x}} = \mathbf{x}_0 + \delta\mathbf{x}$. Assuming the error $\delta\mathbf{c}$ to be sufficiently small, we want to determine $\delta\mathbf{x}$ through a truncated Taylor expansion of $f(\mathbf{x},\mathbf{c})$.

The second-order Taylor series expansion of $f(\mathbf{x};\bar{\mathbf{c}})$, centered at $(\mathbf{x}_0;\mathbf{c}_0)$ can be written as

$$\begin{aligned} f(\mathbf{x};\bar{\mathbf{c}}) \simeq{}& f\big|_{\mathbf{x}_0,\mathbf{c}_0} + (\nabla_{\mathbf{x}} f)^T\big|_{\mathbf{x}_0,\mathbf{c}_0}\delta\mathbf{x} \\ &+ (\nabla_{\mathbf{c}} f)^T\big|_{\mathbf{x}_0,\mathbf{c}_0}\delta\mathbf{c} \\ &+ \frac{1}{2}\delta\mathbf{x}^T\mathbf{H}_{\mathbf{x},\mathbf{x}}(f)|_{\mathbf{x}_0,\mathbf{c}_0}\delta\mathbf{x} \\ &+ \delta\mathbf{c}^T\mathbf{H}_{\mathbf{c},\mathbf{x}}(f)|_{\mathbf{x}_0,\mathbf{c}_0}\delta\mathbf{x} \\ &+ \frac{1}{2}\delta\mathbf{c}^T\mathbf{H}_{\mathbf{c},\mathbf{c}}(f)|_{\mathbf{x}_0,\mathbf{c}_0}\delta\mathbf{c} \end{aligned} \tag{20}$$

where

$$\nabla_{\mathbf{x}} f = [f_{x_1},\ldots,f_{x_M}]^T, \quad \nabla_{\mathbf{c}} f = [f_{c_1},\ldots,f_{c_N}]^T$$
$$f_{x_i} = \frac{\partial f}{\partial x_i}, \quad f_{c_j} = \frac{\partial f}{\partial c_j}$$

$$\mathbf{H}_{\mathbf{x},\mathbf{x}}(f) = \begin{bmatrix} f_{x_1x_1} & \cdots & f_{x_1x_M} \\ \vdots & \ddots & \vdots \\ f_{x_Mx_1} & \cdots & f_{x_Mx_M} \end{bmatrix} \tag{21}$$

$$\mathbf{H}_{\mathbf{c},\mathbf{c}}(f) = \begin{bmatrix} f_{c_1c_1} & \cdots & f_{c_1c_N} \\ \vdots & \ddots & \vdots \\ f_{c_Nc_1} & \cdots & f_{c_Nc_N} \end{bmatrix} \tag{22}$$

$$\mathbf{H}_{\mathbf{x},\mathbf{c}}(f) = \begin{bmatrix} f_{x_1c_1} & \cdots & f_{x_1c_N} \\ \vdots & \ddots & \vdots \\ f_{x_Mc_1} & \cdots & f_{x_Mc_N} \end{bmatrix} \tag{23}$$

and

$$f_{x_ix_j} = \frac{\partial^2 f}{\partial x_i \partial x_j}, \quad f_{x_ic_j} = \frac{\partial^2 f}{\partial x_i \partial c_j}$$
$$f_{c_ic_j} = \frac{\partial^2 f}{\partial c_i \partial c_j}.$$

Notice that $(\nabla_{\mathbf{x}} f)^T|_{\mathbf{x}_0,\mathbf{c}_0} = 0$, as the function with the true parameters $\mathbf{c}_0$ has a minimum in $\mathbf{x}_0$.

Now we study $\nabla_{\mathbf{x}} f(\mathbf{x};\bar{\mathbf{c}}) = 0$ to search for the new minimum $\bar{\mathbf{x}}$, using the Taylor series expansion (20). We obtain

$$\mathbf{H}_{\mathbf{x},\mathbf{x}}(f)|_{\mathbf{x}_0,\mathbf{c}_0}\delta\mathbf{x} + \mathbf{H}_{\mathbf{x},\mathbf{c}}(f)|_{\mathbf{x}_0,\mathbf{c}_0}\delta\mathbf{c} = 0 \tag{24}$$

and finally we can write

$$\delta\mathbf{x} = -\mathbf{H}_{\mathbf{x},\mathbf{x}}(f)|_{\mathbf{x}_0,\mathbf{c}_0}^{-1} \cdot \mathbf{H}_{\mathbf{x},\mathbf{c}}(f)|_{\mathbf{x}_0,\mathbf{c}_0}\delta\mathbf{c}. \tag{25}$$

A. Discussion

The above analysis relies on the hypothesis that we can truncate the Taylor series expansion of f to the second order. This is, of course, true if $\delta\mathbf{x}$ and $\delta\mathbf{c}$ can be assumed as being sufficiently small. Related to this, in [24, Sec. 4.4] a demonstration is given, which confirms that higher order terms are not relevant for the asymptotic analysis of the error.

It is also important to notice that (25) does not envision the possibility of estimation bias, as $\delta\mathbf{x}$ and $\delta\mathbf{c}$ are linearly related. For the cost functions that we propose in this paper, the small error hypothesis is sufficient to prevent bias. However, as we will see in Section V, some localization techniques presented in the literature do introduce a noticeable estimation bias for small measurement errors. If this case happens, a different error propagation analysis is in order.

Finally, (25) is valid under the hypothesis that $\det(\mathbf{H}_{\mathbf{x},\mathbf{x}}(f))|_{\mathbf{x}_0,\mathbf{c}_0} \neq 0$. Mathematically, this condition means that f should have an isolated non-degenerate minimum at $(\mathbf{x}_0; \mathbf{c}_0)$.

In Appendix A, we provide the expressions of $\mathbf{H}_{\mathbf{x},\mathbf{x}}(f)$ and $\mathbf{H}_{\mathbf{x},\mathbf{c}}(f)$ for the cost functions proposed in this paper and for SRD-LS [10].

In Section V, we validate this error propagation analysis.

We remark, finally, that the proposed error propagation analysis is valid not only for the analytical assessment of the accuracy source localization, but also for other estimation problems, under the hypothesis of small bias. For this reason, a Matlab toolbox is available online [18], which includes also a number of demos that show the accuracy of the error prediction.

B. Statistical Error Analysis

In a real scenario we cannot assume the measurement noise $\delta\mathbf{c}$ to be known. However, some statistical information could be available or could be estimated from data. It is therefore important to find a relation between statistical descriptors of the noise $\delta\mathbf{c}$ and of $\delta\mathbf{x}$. In this paragraph, we provide examples for the case of zero-averaged Gaussian error on $\delta\mathbf{c}$. We reformulate the expression in (25) as

$$\delta\mathbf{x} = \mathbf{A}\delta\mathbf{c} \tag{26}$$

where $\mathbf{A} = -\mathbf{H}_{\mathbf{x},\mathbf{x}}(f)|^{-1}_{\mathbf{x}_0,\mathbf{c}_0} \cdot \mathbf{H}_{\mathbf{x},\mathbf{c}}(f)|_{\mathbf{x}_0,\mathbf{c}_0}$. The relationship between the covariance matrix $\mathbf{M}_\mathbf{x}$ of the localization error, and the covariance matrix $\mathbf{M}_\mathbf{c}$ of the measurements is

$$\mathbf{M}_\mathbf{x} = \mathbf{A}\mathbf{M}_\mathbf{c}\mathbf{A}^T \tag{27}$$

where

$$\mathbf{M}_\mathbf{x} = \begin{bmatrix} \sigma_x^2 & \sigma_x\sigma_y & \sigma_x\sigma_z \\ \sigma_x\sigma_y & \sigma_y^2 & \sigma_y\sigma_z \\ \sigma_x\sigma_z & \sigma_y\sigma_z & \sigma_z^2 \end{bmatrix} \tag{28}$$

and

$$\mathbf{M}_\mathbf{c} = \begin{bmatrix} \sigma_{c1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{c2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{cN}^2 \end{bmatrix} \tag{29}$$

under the assumption of statistical independence of measurement errors. As a consequence, given an array geometry and the noise variance of the TDOAs, we can estimate the corresponding error on the coordinates x_S and y_S of the source.

V. Results

In this section, we validate the theoretical error analysis under the small-bias hypothesis. After that we evaluate the performance of the cone-based algorithms.

We start introducing the metrics adopted for accuracy evaluation, and then we describe the simulation setup in both cases of single and multiple arrays.

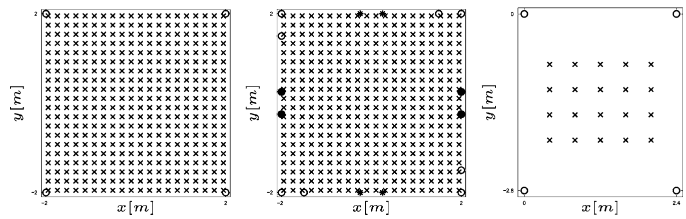

Fig. 2. Setup used for simulations with (a) a single array, (b) multiple arrays, and (c) real experiments. Crosses represent sources, the other symbols represent microphones. In the multiple array case (b), white circles represent microphones positions used during the main simulations while filled circles and asterisks represent the added microphones used for the six and eight sub-arrays cases.

A. Setup and Evaluation Metrics

The setup used for simulations and experiments is shown in Fig. 2. In particular, Fig. 2(a) shows the setup for single-array simulations, Fig. 2(b) shows the setup for multiple-array simulations, and Fig. 2(c) shows the setup used for real data experiments. Circles and asterisks denote microphones, while crosses mark possible positions of the source. In particular, the area covered by the microphones in the simulations is a square of 4 m $\times$ 4 m, while the area covered in the real experiment is a rectangle of 2.8 m $\times$ 2.4 m. We assume that microphone locations are calibrated, i.e., their positions are not affected by errors.

The accuracy of the localization algorithms is evaluated using the following metrics:

— *average bias on the x coordinate of the localized source*:

$$b_x = \left| \frac{1}{n} \sum_{i=1}^{n} (x_S - \hat{x}_{S,i}) \right| \tag{30}$$

where n is the number of noisy measurements tested for each source location, x_S is the x coordinate of the source, and $\hat{x}_{S,i}$ is its estimation based on the ith realization. This is the measure of the average x-coordinate distance between the estimated source and real one.

— *RMSE on x*

$$R_x = \left(\frac{1}{n-1} \sum_{i=1}^{n} (\hat{x}_{S,i} - \bar{x}_S)^2 \right)^{\frac{1}{2}} \tag{31}$$

where n is the number of noisy measurements tested for each source location, and

$$\bar{x}_S = \frac{1}{n} \sum_{i=1}^{n} \hat{x}_{S,i}. \tag{32}$$

All the geometries in Fig. 2 are symmetrical. We also introduced asymmetries in the setup and we verified that the variation of accuracy is quite modest. We resorted nonetheless to symmetrical deployments in order to maximize the area covered by the source location.

In the next subsection, we proceed by validating the error propagation analysis proposed in Section IV. Later on, we compare the accuracy of the proposed cost functions with state-of-the-art techniques. In particular, for the case of a single array we compare $J_{\varepsilon_e}^{(s)}$ and $J_{\varepsilon_a}^{(s)}$ with the squared-range-difference-

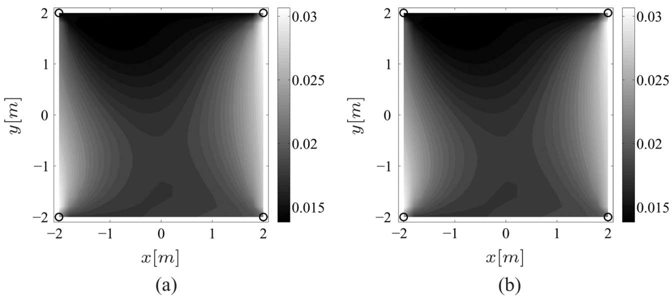

Fig. 3. Comparison for the geometry in 2(a) between CRLB 3(a) and RMSE of MLE [4] 3(b) foreseen by (27). We notice that the RMSE of MLE perfectly matches the CRLB, thus validating the theoretical analysis. (a) CRLB. (b) MLE.

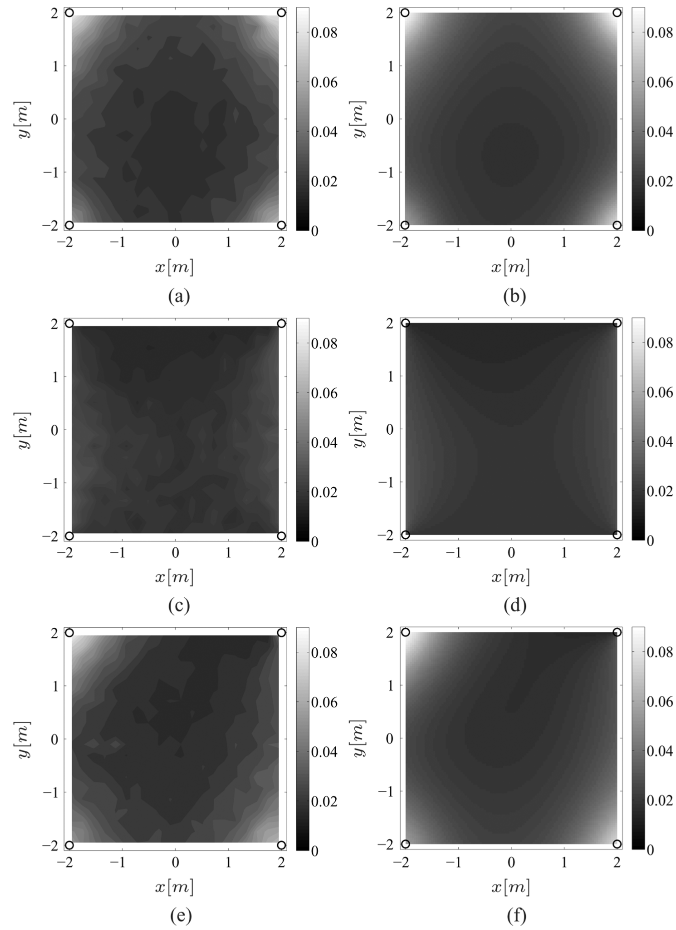

Fig. 4. Comparison between (left) simulated and (right) predicted R_x for a single-array setup. Measures are expressed in meters. We notice from the comparison between (a) and (d) that $J_{\varepsilon_a}^{(s)}$ is close to the CRLB. (a) $J_{\varepsilon_e}^{(s)}$ (simulation), (b) $J_{\varepsilon_e}^{(s)}$ (theoretical), (c) $J_{\varepsilon_a}^{(s)}$ (simulation), (d) $J_{\varepsilon_a}^{(s)}$ (theoretical), (e) SRD-LS (simulation), and (f) SRD-LS (theoretical).

based least squares technique (SRD-LS, [10]), which is suitable for arbitrary configurations of synchronized microphones and it is characterized by state-of-the-art performance. In order to test the robustness of the algorithm against reverberations, we also show the average root mean square error of proposed and state-of-the-art algorithms as a function of the reverberation time. At this purpose, the setup in Fig. 2(a) has been simulated in a room with variable reflection coefficients. Early reflections, up to the tenth-order, have been simulated using the beam tracing algorithm [25]. We also compare the different localization techniques under the case of a variable amount of noise in the TDOA measurements. We finally proceed with the evaluation of our approach for source localization using multiple independent arrays. We do so for both cost functions J_{ε_e} and J_{ε_a}, through a comparison with a state-of-the-art technique proposed by Gillette and Silverman in [15] (GS). Finally, we show the feasibility of the proposed localization techniques with experimental results.

B. Validation of the Error Propagation Analysis

In order to validate the theoretical analysis, we consider the single array setup. The global reference microphone is located in (2, 2). It is well known in the literature (see for example [26]) that maximum-likelihood estimation techniques (e.g., [4]) asymptotically attain the Cramer-Rao Lower Bound. We validate, therefore, the theoretical analysis by showing that the RMSE foreseen by (27) for the log-likelihood used in [4] is equal to the CRLB. For reasons of compactness we omit the derivation of the matrices in (25) for the ML technique in [4].

Fig. 3 shows the results. We notice that, as envisioned by the theory, the theoretical analysis of MLE exactly matches the CRLB, thus validating the error propagation analysis.

C. Single Array

In this section, we adopt the same setup of the previous one. For each position of the source, the set of nominal range differences has been corrupted by 200 repetitions of a uniformly distributed zero-averaged Gaussian noise of standard deviation $\sigma = 3$ cm. This is a consolidated approach to testing, also adopted in [22]. With this data all techniques under examination produced results whose bias b_x turned out to be small and uniformly distributed over the whole area of interest. The average values of the bias over the whole set of 400 source positions and for all the repetitions of the experiment can be found in Table I.

TABLE I
AVERAGE VALUE OF b_x FOR THE THREE ALGORITHMS IN THE SINGLE-ARRAY SETUP

Algorithm	b_x [m]
$J_{\varepsilon_e}^{(s)}$	1.6e-3
$J_{\varepsilon_a}^{(s)}$	1.1e-3
SRD-LS	1.5e-3

Notice that our cost functions produce an average bias that is comparable with that of the SRD-LS method. The spatial distribution of the RMS error for $J_{\varepsilon_e}^{(s)}$, $J_{\varepsilon_a}^{(s)}$ and SRD-LS is shown in Fig. 4. If we compare the two cone-based algorithms, we notice that the two methods are comparable when the source is far from the microphones, while $J_{\varepsilon_a}^{(s)}$ performs better in the proximity of the microphones.

Notice that the values of the bias are small with respect to the RMS error; therefore, the theoretical analysis method proposed in Section IV can be safely applied. In particular Fig. 4(b), (d), and (f) show the RMSE predicted with (27), while Fig. 4(a), (c), and (e) show the spatial distribution of R_x obtained with simulations. Notice that the theoretical error and the simulations are well matched, as confirmed by values of the average and peak differences between RMSE collected in Table II.

TABLE II
AVERAGE AND PEAK R_x VALUES FOR THE THREE ALGORITHMS IN THE SINGLE-ARRAY SETUP

Algorithm	R_x [m] (Simulation) mean	R_x [m] (Simulation) max	R_x [m] (Theoretical) mean	R_x [m] (Theoretical) max
$J_{\varepsilon_e}^{(s)}$	27e-3	85.7e-3	26.4e-3	90e-3
$J_{\varepsilon_a}^{(s)}$	20.2e-3	31.5e-3	20e-3	31.6e-3
SRD-LS	25.5e-3	85.8e-3	24.8e-3	87.6e-3

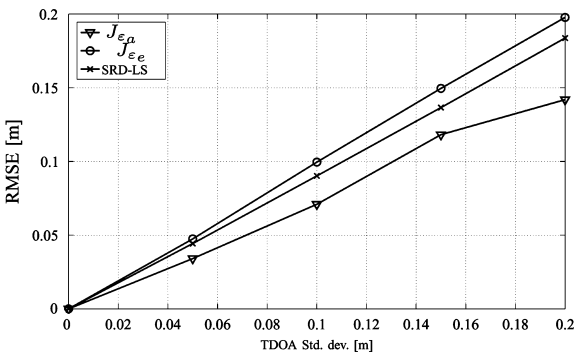

Fig. 5. Root mean square error as a function of the standard deviation of the error introduced on the TDOAs.

We can thus conclude that:

- the theoretical error propagation analysis performs correctly for the tested cost functions;
- the minimization of all the examined cost functions ($J_{\varepsilon_e}^{(s)}$, $J_{\varepsilon_a}^{(s)}$ and SRD-LS) do converge to the global minimum. We should remember that the error propagation analysis only considers cost functions and not the minimization procedure;
- the RMSE of $J_{\varepsilon_a}^{(s)}$ matches quite well that of CRLB and MLE shown in Fig. 3(a) and (b), suggesting that this method, even if it is based on a different idea, has an accuracy comparable with the maximum-likelihood approach in [4];
- SRD-LS and $J_{\varepsilon_e}^{(s)}$ attain the same accuracy. This has a mathematical explanation in the fact that they match except for the presence of the error on the reference microphone in $J_{\varepsilon_e}^{(s)}$ and it is also confirmed by the similarity of the matrices in the theoretical error analysis reported in (38) and (36).

Fig. 5 shows the RMS localization error averaged on all the points of the grid as a function of the standard deviation of the TDOA error, ranging from 0 to 0.2 meters. We notice that the accuracy of J_{ε_e} and SRD-LS are comparable, while J_{ε_a} is more robust.

In order to show the robustness of the proposed algorithms against reverberations, we also performed simulations in a room with variable reverberations. This has been accomplished by simulating that the setup in Fig. 2(a) is placed at the center of a 5 m $\times$ 5 m room, whose walls are characterized by reflection coefficients ranging from 0.5 to 0.94, corresponding to a reverberation time from 0.2 s to 1.0 s. The impulse responses (simulated using the fast beam tracing algorithm [25]) from each possible source location to each microphone have been convolved with the source signal. TDOAs are then extracted from the reverberant signal using the GCC-PHAT algorithm and are fed to the localization algorithms. Fig. 6 shows the average localization error for all the points on the grid in Fig. 2(a) as a function of the reverberation time T60, measured in seconds. We notice that the accuracies of SRD-LS and J_{ε_e} are comparable, while J_{ε_a} more accurately localizes the source, even in the presence of strong reverberations.

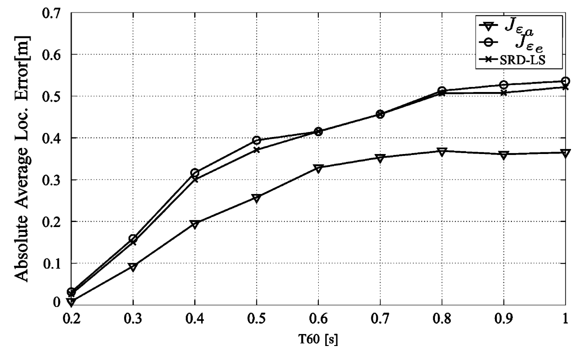

Fig. 6. Absolute value of the average localization error as a function of the reverberation time.

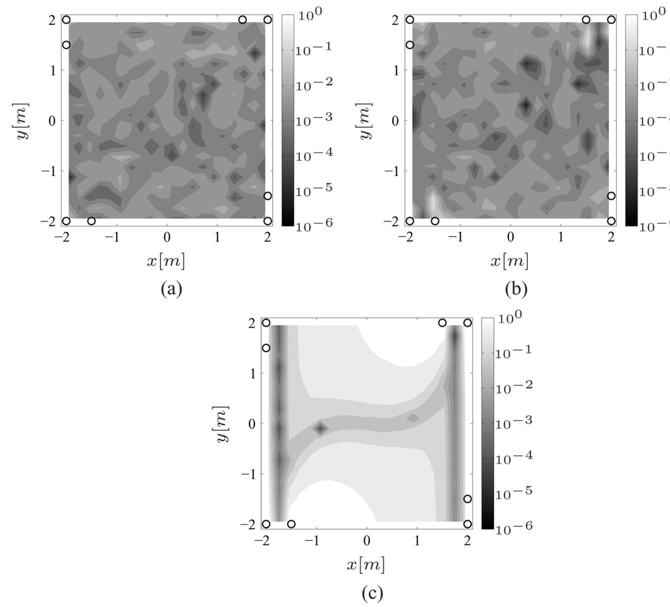

Fig. 7. b_x from simulations for multiple array setup. (a) J_{ε_e}, (b) J_{ε_a}, and (c) GS.

D. Multiple Arrays

The global reference microphone of the multiple arrays is located in (2, 2). For every source position, 200 sets of corrupted TDOAs were tested. The noise standard deviation on range differences is 1 cm.

Fig. 7 shows that the small-bias hypothesis is verified only by J_{ε_e} and J_{ε_a} algorithms, while the GS method turns out to be affected by such a large bias to prevent us from reliably applying our error propagation analysis method. The GS method, in fact, exhibits a rather uneven spatial distribution of the bias, particularly with a very scattered distribution of microphone arrays, with a limited area of modest bias.

The spatial distributions of b_x and R_x, shown in Figs. 7 and 8, respectively, confirm that both cone-based algorithms are suitable for this microphone configuration, especially when the source is not too close to the microphone pairs. The resulting

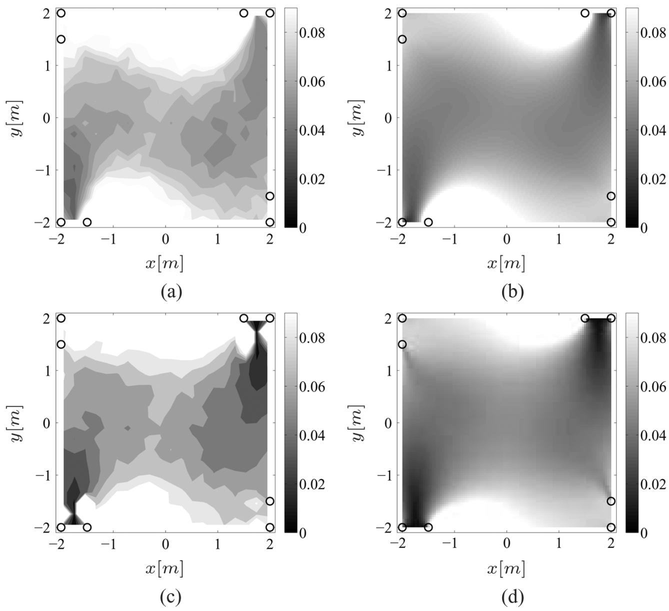

Fig. 8. Comparison between simulated (left) and theoretical (right) R_x for multiple arrays setup. Measures are expressed in meters. (a) J_{ε_e} (simulation), (b) J_{ε_e} (theoretical), (c) J_{ε_a} (simulation), and (d) J_{ε_a} (theoretical).

TABLE III
AVERAGE AND PEAK R_x VALUES FOR THE THREE ALGORITHMS IN THE MULTI-ARRAY SETUP

Algorithm	R_x [m] (Simulation)		R_x [m] (Theoretical)	
	mean	max	mean	max
J_{ε_e}	80.5e-3	239.8e-3	63.4e-3	185.8e-3
J_{ε_a}	78.4e-3	395.1e-3	57.2e-3	178e-3

TABLE IV
AVERAGE R_x ON THE CONSIDERED AREA INCREASING THE NUMBER OF THE SUB-ARRAYS. WITH RESPECT TO FIG. 2(b), MICROPHONES ARE PLACED ON WHITE CIRCLES IN THE 4 ARRAYS CASE, ON WHITE AND BLACK CIRCLES IN THE SIX ARRAYS CASE, AND ON ALL CIRCLES AND SQUARES IN THE EIGHT ARRAYS CASE

Algorithm	Average R_x [m]		
	4 arrays	6 arrays	8 arrays
J_{ε_e}	80.5e-3	52.2e-3	38.9e-3
J_{ε_a}	78.4e-3	43.7e-3	30.2e-3

RMSE, averaged over the area of interest are shown in Table III. Notice that the predicted and simulated RMSE values in the multi-array case do not match perfectly like in the single-array case. We recall, in fact, that the error propagation analysis is valid under the hypothesis of small errors for $\delta\mathbf{x}$. This is not the case for the white regions in Fig. 8(a) and (c); therefore, we can expect that for these areas the theoretical error analysis represents an approximation of the simulation error. In Table IV, we show the relationship between the number of sub-arrays and the RMSE over the area of interest. As expected, the RMSE value averaged on many simulations decreases if we use six or eight sub-arrays instead of only four.

E. Experimental Results

In order to further confirm the validity of the proposed method, we also conducted a set of experiments with real data. Here too we adopt b_x and R_x as evaluation metrics.

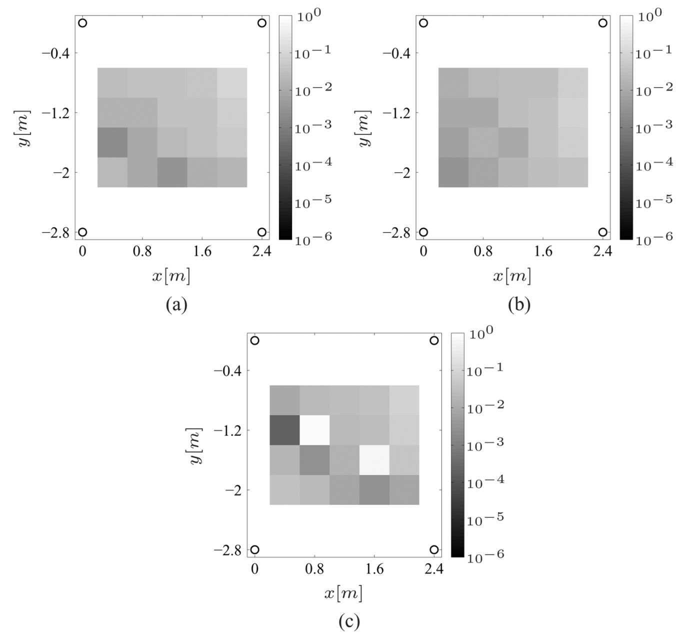

Fig. 9. b_x values obtained with different algorithms used in a real experiment. (a) J_{ε_e}, (b) J_{ε_a}, and (c) SRD-LS.

TABLE V
MEAN CONVERGENCE TIME AND R_x FOR THE TESTED ALGORITHMS

Algorithm	Time [ms]	R_x [m]
$J_{\varepsilon_e}^{(s)}$	0.804	0.0293
$J_{\varepsilon_a}^{(s)}$	0.94	0.0693
SRD-LS	1.6	0.0337

The experimental setup is described in Fig. 2(c). The TDOA measurements are associated to a source that produces a segment of white noise with a duration of 10 s in a low-reverberation room. The recorded signal was then windowed into segments of 100 ms each. We then performed source localization for each segment.

Fig. 9 shows b_x for $J_{\varepsilon_e}^{(s)}$, $J_{\varepsilon_a}^{(s)}$, and SRD-LS. The results confirm that the three cost functions perform similarly, as shown in the simulations. Table V shows the mean convergence time and the mean value of R_x for the tested algorithms.

Table V confirms that, even if the minimization techniques proposed in this paper are of iterative nature, the average convergence time turns out to be comparable with that of SRD-LS.

VI. CONCLUSION

In this paper, we proposed and discussed a method for the localization of an acoustic source using multiple unsynchronized arrays through the fitting of propagation cones in the space-range reference frame. Starting from this idea, we have proposed two different cost functions that express cone fitting, based on the cone equation and the cone aperture. The fact that the method allows us to successfully use multiple unsynchronized arrays results from the generality of the cone fitting process, which operates on range-shifting clusters of data. Experimental results showed that our approach performs well in the considered general scenario. In particular, the cone function based on the cone aperture results more accurate than state-of-the-art algorithms.

In this paper, we also proposed a technique for the prediction of the localization error, which relates the error on measurements and the localization error. This method is general enough to be applied also in other kind of estimation processes. The effectiveness of the error prediction method turns it into an effective tool for optimizing the spatial distribution of subarrays.

APPENDIX
EXAMPLES OF ERROR PROPAGATION ANALYSIS

In this appendix we provide details on error propagation analysis in the various cases examined in this manuscript. In particular, we discuss our cost functions and those discussed in [10] using (25). To avoid any possible source of confusion, we refer to $\mathbf{x}_{S0}$ as the actual position of the source and to $\mathbf{c}_0$ as the nominal measurements.

- In the case of a single array we have $N+1$ microphones in positions $(x_i, y_i),\ i = 0,\dots,N$, with the reference microphone at (x_0, y_0). Let us define

$$\begin{aligned}\Delta x_i &= x_i - x_{S0}\\ \Delta y_i &= y_i - y_{S0}\\ D_i &= \sqrt{\Delta x_i{}^2 + \Delta y_i{}^2}.\end{aligned}$$

  The matrices $\mathbf{H}_{\mathbf{x},\mathbf{x}}(f)|_{\mathbf{x}_{S0},\mathbf{c}_0}$ and $\mathbf{H}_{\mathbf{x},\mathbf{c}}(f)|_{\mathbf{x}_{S0},\mathbf{c}_0}$ for cost functions $J^{(s)}_{\varepsilon_e}$ and $J^{(s)}_{\varepsilon_a}$ are given in (36) and (37), respectively.
- In the case of single array, using the notation of the previous point, the SRD-LS error (see [10]) is

$$\varepsilon_{s,i} = (x_i - x_S)^2 + (y_i - y_S)^2 - (z_i + D_0)^2 \tag{33}$$

  and the relative cost function is

$$J_s = \sum_{i=1}^{N} \varepsilon_{s,i}^2. \tag{34}$$

  Let us define

$$\begin{aligned}\tilde{\Delta} x_i &= \Delta x_i - \frac{\Delta x_0}{D_0} D_i\\ \tilde{\Delta} y_i &= \Delta y_i - \frac{\Delta y_0}{D_0} D_i.\end{aligned}$$

  The matrices $\mathbf{H}_{\mathbf{x},\mathbf{x}}(f)|_{\mathbf{x}_{S0},\mathbf{c}_0}$ and $\mathbf{H}_{\mathbf{x},\mathbf{c}}(f)|_{\mathbf{x}_{S0},\mathbf{c}_0}$ for cost function J_{ε_s} are given in (38), as shown at the bottom of the page.
- The case of multiple arrays is slightly more complex. We have L arrays composed by $N+1$ microphones at positions $(x_i^{(l)}, y_i^{(l)}),\ i = 0,\dots,N,\ l = 1,\dots,L$. Each array l has its own local reference microphone at $(x_0^{(l)}, y_0^{(l)})$ and there is a global reference microphone at $(x_0^{(1)}, y_0^{(1)})$. Let us define

$$\begin{aligned}\Delta x_i^{(l)} &= x_i^{(l)} - x_{S0}\\ \Delta y_i^{(l)} &= y_i^{(l)} - y_{S0}\\ D_i^{(l)} &= \sqrt{\left(\Delta x_i^{(l)}\right)^2 + \left(\Delta y_i^{(l)}\right)^2}\\ \tilde{\Delta} x_i^{(l)} &= \Delta x_i^{(l)} - \left(\frac{\Delta x_0^{(l)}}{D_0^{(l)}} - \frac{\Delta x_0^{(1)}}{D_0^{(1)}}\right) D_i^{(l)}\\ \tilde{\Delta} y_i^{(l)} &= \Delta y_i^{(l)} - \left(\frac{\Delta y_0^{(l)}}{D_0^{(l)}} - \frac{\Delta y_0^{(1)}}{D_0^{(1)}}\right) D_i^{(l)}.\end{aligned}$$

$$\mathbf{H}_{\mathbf{x},\mathbf{x}}\left(J_e^{(s)}\right)\Big|_{\mathbf{x}_0,\mathbf{c}_0} = 8\sum_{i=0}^{N}\begin{bmatrix}\Delta x_i{}^2 & \Delta x_i\Delta y_i & -\Delta x_i D_i\\ \Delta x_i\Delta y_i & \Delta y_i{}^2 & -\Delta y_i D_i\\ -\Delta x_i D_i & -\Delta y_i D_i & D_i{}^2\end{bmatrix} \qquad \mathbf{H}_{\mathbf{x},\mathbf{c}}\left(J_e^{(s)}\right)\Big|_{\mathbf{x}_0,\mathbf{c}_0} = 8\begin{bmatrix}\Delta x_1 D_1 & \dots & \Delta x_N D_N\\ \Delta y_1 D_1 & \dots & \Delta y_N D_N\\ -D_1{}^2 & \dots & -D_N{}^2\end{bmatrix} \tag{36}$$

$$\mathbf{H}_{\mathbf{x},\mathbf{x}}\left(J_a^{(s)}\right)\Big|_{\mathbf{x}_0,\mathbf{c}_0} = 2\sum_{i=0}^{N}\begin{bmatrix}\frac{\Delta x_i{}^2}{D_i{}^2} & \frac{\Delta x_i\Delta y_i}{D_i{}^2} & -\frac{\Delta x_i}{D_i}\\ \frac{\Delta x_i\Delta y_i}{D_i{}^2} & \frac{\Delta y_i{}^2}{D_i{}^2} & -\frac{\Delta y_i}{D_i}\\ -\frac{\Delta x_i}{D_i} & -\frac{\Delta y_i}{D_i} & 1\end{bmatrix} \qquad \mathbf{H}_{\mathbf{x},\mathbf{c}}\left(J_a^{(s)}\right)\Big|_{\mathbf{x}_0,\mathbf{c}_0} = 2\begin{bmatrix}\frac{\Delta x_1}{D_1} & \dots & \frac{\Delta x_N}{D_N}\\ \frac{\Delta y_1}{D_1} & \dots & \frac{\Delta y_N}{D_N}\\ -1 & \dots & -1\end{bmatrix} \tag{37}$$

$$\mathbf{H}_{\mathbf{x},\mathbf{x}}(J_s)|_{\mathbf{x}_0,\mathbf{c}_0} = 8\sum_{i=1}^{N}\begin{bmatrix}\tilde{\Delta} x_i{}^2 & \tilde{\Delta} x_i\tilde{\Delta} y_i\\ \tilde{\Delta} x_i\tilde{\Delta} y_i & \tilde{\Delta} y_i{}^2\end{bmatrix} \qquad \mathbf{H}_{\mathbf{x},\mathbf{c}}(J_s)|_{\mathbf{x}_0,\mathbf{c}_0} = 8\begin{bmatrix}\tilde{\Delta} x_1 D_1 & \dots & \tilde{\Delta} x_N D_N\\ \tilde{\Delta} y_1 D_1 & \dots & \tilde{\Delta} y_N D_N\end{bmatrix} \tag{38}$$

$$\begin{aligned}
\mathbf{H}_{\mathbf{x},\mathbf{x}}(J_e)|_{\mathbf{x}_0,\mathbf{c}_0} &= 8\sum_{l=1}^{L}\sum_{i=0}^{N}\begin{bmatrix} \tilde{\Delta}x_i^{(l)^2} & \tilde{\Delta}x_i^{(l)}\tilde{\Delta}y_i^{(l)} & -\tilde{\Delta}x_i^{(l)}D_i^{(l)} \\ \tilde{\Delta}x_i^{(l)}\tilde{\Delta}y_i^{(l)} & \tilde{\Delta}y_i^{(l)^2} & -\tilde{\Delta}y_i^{(l)}D_i^{(l)} \\ -\tilde{\Delta}x_i^{(l)}D_i^{(l)} & -\tilde{\Delta}y_i^{(l)}D_i^{(l)} & D_i^{(l)^2} \end{bmatrix} \\
\mathbf{H}_{\mathbf{x},\mathbf{c}}(J_e)|_{\mathbf{x}_0,\mathbf{c}_0} &= 8\begin{bmatrix} \tilde{\Delta}x_1^{(1)}D_1^{(1)} & \tilde{\Delta}x_2^{(1)}D_2^{(1)} & \dots & \tilde{\Delta}x_N^{(1)}D_N^{(1)} & \tilde{\Delta}x_1^{(2)}D_1^{(2)} & \dots & \tilde{\Delta}x_N^{(L)}D_N^{(L)} \\ \tilde{\Delta}y_1^{(1)}D_1^{(1)} & \tilde{\Delta}y_2^{(1)}D_2^{(1)} & \dots & \tilde{\Delta}y_N^{(1)}D_N^{(1)} & \tilde{\Delta}y_1^{(2)}D_1^{(2)} & \dots & \tilde{\Delta}y_N^{(L)}D_N^{(L)} \\ -D_1^{(1)^2} & -D_2^{(1)^2} & \dots & -D_N^{(1)^2} & -D_1^{(2)^2} & \dots & -D_N^{(L)^2} \end{bmatrix}
\end{aligned} \tag{39}$$

$$\begin{aligned}
\mathbf{H}_{\mathbf{x},\mathbf{x}}(J_a)|_{\mathbf{x}_0,\mathbf{c}_0} &= 8\sum_{l=1}^{L}\sum_{i=0}^{N}\begin{bmatrix} \frac{\tilde{\Delta}x_i^{(l)^2}}{D_i^{(l)^2}} & \frac{\tilde{\Delta}x_i^{(l)}\tilde{\Delta}y_i^{(l)}}{D_i^{(l)^2}} & \frac{-\tilde{\Delta}x_i^{(l)}}{D_i^{(l)}} \\ \frac{\tilde{\Delta}x_i^{(l)}\tilde{\Delta}y_i^{(l)}}{D_i^{(l)^2}} & \frac{\tilde{\Delta}y_i^{(l)^2}}{D_i^{(l)^2}} & \frac{-\tilde{\Delta}y_i^{(l)}}{D_i^{(l)}} \\ \frac{-\tilde{\Delta}x_i^{(l)}}{D_i^{(l)}} & \frac{-\tilde{\Delta}y_i^{(l)}}{D_i^{(l)}} & 1 \end{bmatrix} \\
\mathbf{H}_{\mathbf{x},\mathbf{c}}(J_a)|_{\mathbf{x}_0,\mathbf{c}_0} &= 8\begin{bmatrix} \frac{\tilde{\Delta}x_1^{(1)}}{D_1^{(1)}} & \frac{\tilde{\Delta}x_2^{(1)}}{D_2^{(1)}} & \dots & \frac{\tilde{\Delta}x_N^{(1)}}{D_N^{(1)}} & \frac{\tilde{\Delta}x_1^{(2)}}{D_1^{(2)}} & \dots & \frac{\tilde{\Delta}x_N^{(L)}}{D_N^{(L)}} \\ \frac{\tilde{\Delta}y_1^{(1)}}{D_1^{(1)}} & \frac{\tilde{\Delta}y_2^{(1)}}{D_2^{(1)}} & \dots & \frac{\tilde{\Delta}y_N^{(1)}}{D_N^{(1)}} & \frac{\tilde{\Delta}y_1^{(2)}}{D_1^{(2)}} & \dots & \frac{\tilde{\Delta}y_N^{(L)}}{D_N^{(L)}} \\ -1 & -1 & \dots & -1 & -1 & \dots & -1 \end{bmatrix}
\end{aligned} \tag{40}$$

The matrices $\mathbf{H}_{\mathbf{x},\mathbf{x}}(f)|_{\mathbf{x}_{S0},\mathbf{c}_0}$ and $\mathbf{H}_{\mathbf{x},\mathbf{c}}(f)|_{\mathbf{x}_{S0},\mathbf{c}_0}$ for cost functions J_{ε_e} and J_{ε_a} are given in (39) and (40), respectively, shown at the top of the page

Acknowledgment

The authors would like to thank A. Canclini for his participation to the initial steps of this work. M. Compagnoni would also like to thank S. Turzi for many useful discussions.

References

[1] W. Hahn and S. Tretter, "Optimum processing for delay-vector estimation in passive signal arrays," *IEEE Trans. Inf. Theory*, vol. 19, no. 5, pp. 608–614, Sep. 1973.

[2] M. Wax and T. Kailath, "Optimum localization of multiple sources by passive arrays," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-31, no. 5, pp. 1210–1217, Oct. 1983.

[3] P. Stoica and A. Nehorai, "Music, maximum likelihood and cramer-rao bound," in *Proc. 1988 Int. Conf. Acoust., Speech, Signal Process. (ICASSP '88)*, Apr. 1988, vol. 4, pp. 2296–2299.

[4] J. Chen, R. Hudson, and K. Yao, "Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field," *IEEE Trans. Signal Process.*, vol. 50, no. 8, pp. 1843–1854, Aug. 2002.

[5] N. Castaneda, M. Charbit, and E. Moulines, "Source localization from quantized time of arrival measurements," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP '06)*, May 2006, vol. 4, pp. IV-933–IV-936.

[6] R. Schmidt, "A new approach to geometry of range difference location," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-8, no. 6, pp. 821–835, Nov. 1972.

[7] H. Schau and A. Robinson, "Passive source localization employing intersecting spherical surfaces from time-of-arrival differences," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 8, pp. 1223–1225, Aug. 1987.

[8] Y. Chan and K. A. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 1905–1915, Aug. 1994.

[9] Y. Huang, J. Benesty, G. Elko, and R. Mersereati, "Real-time passive source localization: A practical linear-correction least-squares approach," *IEEE Trans. Speech and Audio Process.*, vol. 9, no. 8, pp. 943–956, Nov. 2001.

[10] A. Beck, P. Stoica, and J. Li, "Exact and approximate solutions of source localization problems," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1770–1778, May 2008.

[11] J. Jia, M. Liu, and X. Li, "Acoustic passive localization algorithm based on wireless sensor networks," in *Proc. Int. Conf. Mechatron. Autom. ICMA '09*, 2009, pp. 1145–1149.

[12] J. Jia, M. Liu, and X. Li, "Acoustic localization algorithm using wireless sensor networks," in *Proc. 2nd Int. Conf. Intell. Comput. Technol. Autom. ICICTA'09*, 2009, vol. 2, pp. 434–437.

[13] E. Habets and P. Sommen, "Optimal microphone placement for source localization using time delay estimation," in *Proc. 13th Annu. Workshop Circuits, Syst., Signal Process. (ProRISC '02)*, 2002.

[14] T. Ajdler, I. Kozintsev, R. Lienhart, and M. Vetterli, "Acoustic source localization in distributed sensor networks," in *Proc. Conf. Rec. 38th Asilomar Conf. Signals, Syst., Comput.*, 2004.

[15] M. Gillette and H. Silverman, "A linear closed-form algorithm for source localization from time-differences of arrival," *IEEE Signal Process. Lett.*, vol. 15, pp. 1–4, 2008.

[16] R. Penrose, *The Road to Reality: A Complete Guide to the Laws of the Universe*. New York: Vintage Books, 2007.

[17] Y. Huang, J. Benesty, and G. Elko, *Source Localization*. Norwell, MA: Kluwer, 2004, ch. 9, pp. 229–253.

[18] Error Propagation Toolbox, [Online]. Available: http://www.thescenicproject.eu/resources/software/EPT.zip

[19] J. Scheuing and B. Yang, "Disambiguation of TDOA estimation for multiple sources in reverberant environments," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 8, pp. 1479–1489, Aug. 2008.

[20] G. Lukcs, R. Martin, and D. Marshall, "Faithful least-squares fitting of spheres, cylinders, cones and tori for reliable segmentation," in *Proc. 5th Eur. Conf. Comput. Vis.*, 1998, pp. 671–686, Springer.

[21] J. O. Smith and J. S. Abel, "Closed-form least-squares source location estimation from range-difference measurements," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 12, pp. 1661–1669, Dec. 1987.

[22] Y. Huang and J. Benesty, *Audio Signal Processing for Next Generation Multimedia Communication Systems*. Norwell, MA: Kluwer, 2004.

[23] R. Gilmore, *Catastrophe Theory for Scientists and Engineers*. New York: Dover, 1993, reprint of the 1981 original.

[24] S. Amari and H. Nagaoka, *Methods of Information Geometry*. Providence, RI: Amer. Math. Soc., 2000.

[25] F. Antonacci, M. Foco, A. Sarti, and S. Tubaro, "Fast tracing of acoustic beams and paths through visibility lookup," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 4, pp. 812–824, May 2008.

[26] E. L. Lehmann and G. Casella, *Theory of Point Estimation*, 2nd ed. New York: Springer, 1998.

Marco Compagnoni was born in Tirano, Italy, on April 25, 1980. He received the Laurea degree (*cum laude*) in theoretical physics from the Universitá di Milano Bicocca, Milan, Italy, in 2004 and the Ph.D. degree in applied mathematics from the Politecnico di Milano, Milan, in 2009.

He is currently a Post-Doc Researcher at the Algebra and Geometry Group, Dipartimento di Matematica, Politecnico di Milano. His research interests are applied math, especially in the application of geometry in mathematical physics and in statistics, in image processing and pattern recognition and in space–time processing of audio signals.

Paolo Bestagini was born in Novara, Italy, on February 22, 1986. He received the M.Sc. degree (*cum laude*) in telecommunication engineering from the Politecnico di Milano, Milan, Italy. He is currently pursuing the Ph.D. degree at the Image and Sound Processing Group, Dipartimento di Elettronica ed Informazione, Politecnico di Milano.

He initially worked on audio signal processing for source localization. His recent research focuses on reverse engineering of multi-media data for forensics analyses.

Fabio Antonacci was born in Bari, Italy, on July 26, 1979. He received the Laurea degree in telecommunication engineering and the Ph.D. degree in information engineering from the Politecnico di Milano, Milan, Italy, in 2004 and 2008, respectively.

He is currently a Post-Doc Researcher at the Image and Sound Processing Group, Dipartimento di Elettronica ed Informazione, Politecnico di Milano. His research focuses on space–time processing of audio signals, for both speaker and microphone arrays (source localization, acoustic scene analysis, rendering of spatial sound) and on modeling of acoustic propagation (visibility-based beam tracing). He published approximately 40 articles in proceedings of international conferences and on peer-reviewed journals.

Augusto Sarti (M'04) received the M.S. and the Ph.D. degrees in electronic engineering, both from the University of Padua, Padua, Italy, in 1988 and 1993, respectively, with research on nonlinear communication systems. His graduate studies included a joint graduate program with the University of California at Berkeley, where he worked on nonlinear system theory.

In 1993, he joined the Dipartimento di Elettronica e Informazione of the Politecnico di Milano, Milan, Italy, where he is currently an Associate Professor. His research interests are in the area of digital signal processing, with particular focus on sound analysis, processing, and synthesis; space–time audio processing; geometrical acoustics; and music information retrieval. He also worked on problems of multidimensional signal processing, vision-based 3-D scene reconstruction, camera calibration, image analysis, motion planning, and nonlinear system theory. He coauthored over 180 scientific publications on international journals and congresses as well as numerous patents in the multimedia signal processing area. He coordinates the Sound and Music Computing Lab of the Image and Sound Processing Group of the Politecnico di Milano. He promoted and coordinated or contributed to numerous (20+) EC-funded projects.

Prof. Sarti is a member of the IEEE Technical Committee on Audio and Acoustics Signal Processing, and Associate Editor of the IEEE SIGNAL PROCESSING LETTERS. He has served as a guest editor for numerous special issues of international journals. He was cochairman of the 2005 Edition of the IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS); Chairman of 2009 edition of the Digital Audio Effects Conference, (DAFx); and in the organizing committees of numerous other conferences in the area of signal processing.

Stefano Tubaro (M'01) was born in Novara, Italy, in 1957. He received the Electronic Engineering degree from the Politecnico di Milano, Milan, Italy, in 1982.

He then joined the Dipartimento di Elettronica e Informazione, Politecnico di Milano, first as a Researcher of the National Research Council, then as an Associate Professor (1991) and finally as a Full Professor (2004). He initially worked on problems related to speech analysis, motion estimation/compensation for video analysis/coding, and vector quantization applied to hybrid video coding. In the past few years, his research interests have focused on image and video analysis for the geometric and radiometric modeling of 3-D scenes, and advanced algorithms for video coding and sound processing. He has authored over 150 scientific publications on international journals and congresses. He coauthored two books on digital processing of video sequences. He also coauthored several patents on image processing techniques. He coordinates the research activities of the Image and Sound Processing Group (ISPG) at the Dipartimento di Elettronica e Informazione, Politecnico di Milano, which is involved in several research programs funded by industrial partners, the Italian Government, and by the European Commission.

Prof. Tubaro has been involved in the IEEE Technical Committee of Multimedia Signal Processing (2005–2009), and is currently involved in that of Image Video and Multidimensional Signal Processing.