Localization of random acoustic sources in an inhomogeneous medium

Shahram Khazaie *, Xun Wang, Pierre Sagaut
Aix Marseille Univ, CNRS, Centrale Marseille, M2P2, 13451 Marseille Cedex 13, France

A R T I C L E   I N F O

Article history:
Received 2 February 2016
Received in revised form 20 July 2016
Accepted 3 August 2016
Handling Editor: Y. Auregan
Available online 25 August 2016

A B S T R A C T

In this paper, the localization of a random sound source via different source localization methods is considered, the emphasis being put on the robustness and the accuracy of classical methods in the presence of uncertainties. The sound source position is described by a random variable and the sound propagation medium is assumed to have spatially varying parameters with known values. Two approaches are used for the source identification: time reversal and beamforming. The probability density functions of the random source position are estimated using both methods. The focal spot resolutions of the time reversal estimates are also evaluated. In the numerical simulations, two media with different correlation lengths are investigated to account for two different scattering regimes: one has a correlation length relatively larger than the wavelength and the other has a correlation length comparable to the wavelength. The results show that the required sound propagation time and source estimation robustness highly depend on the ratio between the correlation length and the wavelength. It is observed that source identification methods have different robustness in the presence of uncertainties. Advantages and weaknesses of each method are discussed.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Sound source estimation using recorded signals by various active or passive sensors is one of the main problems in acoustical engineering, oceanic engineering and seismology. Modeling the wave propagation requires the knowledge about the properties of the propagation medium along with the source parameters. The spatial variation of the medium properties is one of the main factors which plays an important role on the applicability of the localization method. In elastic media for instance, geological processes, experimental observations, and well log data are among the direct evidences of the existence of spatial variations [1,2]. In ocean acoustics, the sound propagation speed varies with depth, pressure, salinity and temperature [3]. Typical sound speed profiles show drastic variations especially near the sea surface. Hence, the environmental parameters are spatially varying and only their statistical information might be provided [4]. Another source of uncertainty comes from the randomness of the source position. In acoustics, Grosveld [5] investigated the acceleration response and noise reduction of plates which are excited by random acoustic sources. In seismology, Husen et al. [6] combined a probabilistic earthquake location and a 3D velocity model to more precisely identify the hypocenter location. In this

* Corresponding author.
E-mail addresses: shahram.khazaie@univ-amu.fr (S. Khazaie), xun.wang@univ-amu.fr (X. Wang), pierre.sagaut@univ-amu.fr (P. Sagaut).
presentation, spatial variations of the medium properties are taken into account using a single realization of a random function of space. The source position is modeled via a random variable.

A comprehensive review of the source localization techniques along with the advantages and drawbacks of each method is introduced in [7]. Classical beamforming [8–10] and Near-field acoustical holography (NAH) [11–13] are two approaches dedicated to the source localization problem, which assume that the medium is homogeneous, i.e., Green’s function of wave equation has an analytical solution. NAH back-propagates the acoustic field in the wavenumber domain from the measurement plane to the source plane. This approach has a high spatial resolution for source identification but works only for near-field source. The beamforming method estimates the sound source location via the delay of signal arrival measured by the microphones. As already mentioned, in many real applications such as in geophysical media or ocean environments, the homogeneous assumption does not hold and would cause large errors in the source localization result. In heterogeneous media, however, the wave equation has no analytical solution so that the sound propagation process has to be numerically simulated. To this end, a variant of the classical beamforming is employed to estimate the sound source, in which the closed form of Green’s function is replaced by a numerical version. Another source localization method in the time domain that is frequently used in both acoustic and elastic media is time reversal [14–19]. It is based on the symmetric nature of the wave equation with respect to the time variable in non-dissipative media. The received sound waves are reversed in time and reinjected into the same medium to refocus on the initial source location within a resolution. An experimental comparison between time reversal and beamforming methods has been recently done in aeroacoustics [20]. Mimani et al. [21] introduced an improved version of time reversal in aeroacoustics and made the same comparison. One of the main differences between our work and the last two cited papers lies on the fact that we consider a randomly heterogeneous propagation medium and we discuss the influence of two different scattering regimes.

In this paper, the localization of random sound source propagating through randomly-generated inhomogeneous media is discussed. The emphasis is put on the robustness of the methods and the number of samples required to get a converged solution of random sound source. The sound source location is chosen randomly following a given probability distribution. Each realization of the random sound source emits the sound waves and the generated wave field is measured by a linear array of microphones. Then, the source location is estimated via beamforming and time reversal methods. The estimates of both approaches reconstruct the probability density functions (PDF) of the random sound source position and a comparison between the results is done. The present analysis has relevance to nondestructive testing of materials, structural health monitoring, ultrasonic medical imaging and underwater communication, among others. In all these applications, one can consider the random nature of the source position with an aim of identifying the propagation of the initial uncertainty on the direct and then on the source localization problem.

The organization of this paper is as follows. Section 2 begins with a description of the proposed model. In particular, the inhomogeneous medium and random source are explicitly introduced. Then, in Section 3, the source localization methods used in this paper (time reversal and beamforming) are presented and compared. In Section 4, numerical simulations are presented, in which two media with different correlation lengths are investigated. The estimation efficiency of the sound source using the proposed approaches is justified. Finally, Section 5 concludes the paper.

2. Acoustic wave propagation in randomly inhomogeneous media with randomly-generated source position

The acoustic wave equation for the pressure and velocity fields (p and \(v\), respectively) propagating in a non-dissipative medium is considered as:

\[
\rho(x) \frac{\partial v(x,t)}{\partial t} + \nabla p(x,t) = 0, \tag{1a}
\]

\[
\frac{1}{\kappa(x)} \frac{\partial p(x,t)}{\partial t} + \nabla \cdot v(x,t) = 0, \tag{1b}
\]

in which \(\rho(x)\) and \(\kappa(x)\) are, respectively, the local density and bulk modulus of the heterogeneous medium, and \(x = (x,y)\) stands for the 2D spatial coordinate. The local angular frequency and the wavenumber are related via the dispersion equation \(\omega(x) = c(x)|k|\), in which \(c(x) = \sqrt{\kappa(x)/\rho(x)}\) is the local velocity of propagation. The corresponding wave equation for the velocity potential field \(\Phi_v\) reads

\[
\left[\Delta - \frac{1}{c^2(x)} \frac{\partial^2}{\partial t^2}\right] \Phi_v(x,t) = S(x_0,t), \quad (x,x_0,t) \in \mathbb{R}^2 \times \mathbb{R}^+ \times \mathbb{R}^2, \tag{2}
\]

in which the term \(S(x_0,t)\) represents the acoustic source. Note that the velocity field \(v\) and its potential \(\Phi_v\) are related via \(v = \nabla \Phi_v\). The experimental estimation of the medium parameters, i.e., \(\kappa(x)\) and \(\rho(x)\), is highly limited and in general can be done locally on some limited points. The source parameters such as position \(x_0 = (x_0,y_0)\) and amplitude are also a priori unknown and are often considered as the parameters to be identified. In this paper, the medium properties are considered as a single realization of a sufficiently large (compared to the wavelength) randomly heterogeneous medium and the source position is modeled using a random variable. In the next two sections we will deal with the probabilistic description of the medium and source parameters.
2.1. Modeling of the medium

For the sake of simplicity, the mass density is considered as homogeneous and deterministic so that \( \rho(\mathbf{x}) = \rho \). A probabilistic approach will be used to model the inhomogeneous positive-valued bulk modulus field \( \kappa(\mathbf{x}): \mathbb{R}^2 \rightarrow \mathbb{R}^+ \). In this paper, the randomness in the medium properties is nothing more than the spatial variation of them. The bulk modulus is generated from a random field following a Gamma probability distribution. The Gamma distribution results from solving a maximum entropy principle problem (see [22–24] for more details). More precisely, we model \( \kappa(\mathbf{x}) \) as second-order statistically homogeneous random field with a priori known mean value \( \kappa \) and autocorrelation function (ACF) \( \kappa(\mathbf{x} - \mathbf{x'}) = \mathbb{E}[\kappa(\mathbf{x}) - \kappa] \kappa(\mathbf{x'}) - \kappa) \). The variance of the corresponding random field is therefore \( \sigma^2 = R_\kappa(0) \). It should be pointed out that the hypothesis of statistical homogeneity implies that the average field \( \kappa \) does not depend on the space position and that its ACF depends only on \( \kappa(\mathbf{x} - \mathbf{x'}) \). We further assume that the correlation function has an isotropic structure which means that it depends only on the distance between the points \( r = |\mathbf{x} - \mathbf{x'}| \). Finally, we define the scale at which the values of the stochastic field \( \kappa(\mathbf{x}) \) fluctuate that is called the correlation length \( \varepsilon_c \). For the sake of simplicity, we consider only a single length scale to describe the fluctuations of the bulk modulus. However, in multiphase media, more complex stochastic models that take into account several correlation lengths may be considered (see [25] for instance). This parameter can be imagined as the typical size of the heterogeneities and is defined here as

\[
\varepsilon_c = \frac{1}{\sigma^2} \int R_\kappa(r) dr. \tag{3}
\]

In some propagation regimes (radiative transfer and diffusion regimes for instance), the Wigner transform of the wave field (being a sort of localized energy density) turns out to be statistically stable. This means that it converges, in probability, to its deterministic limit and becomes almost independent of the particular realization of the random medium. This statistical stability results from the assumption of ergodicity of the underlying random fields so that a single realization of the medium suffices to estimate the corresponding statistics (see [26] and the references therein for more discussion). Since in this study the boundaries are considered as reflecting, the propagation medium can be imagined as an ergodic cavity and thus a diffusion regime will be established provided that the amplitude of the fluctuations are weak. That is why only a single realization of a randomly heterogeneous is employed to model the propagation medium.

In numerical simulations, one should discretize the random field over the numerical domain which is denoted herein-after by \( \Omega \). In this presentation, in order to represent the random field \( \kappa(\mathbf{x}) \), a Karhunen–Loève (KL) expansion is used (see [27–29] for instance).

2.2. Modeling of the source

We model the randomness in the source position via two independent Gaussian random variables for each of the coordinates, i.e.

\[
f_W(w) = \frac{1}{\sigma_W \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{w - \mu_W}{\sigma_W} \right)^2 \right], \quad W \in (X_0, Y_0), \tag{4}
\]

wherein \( X_0 \) and \( Y_0 \) denote the random variables corresponding to the \( x \) and \( y \) components of the source position. Since the values of the mean and standard deviation of the source position are given and based on the maximum entropy principle, the best PDF to be used is the Gaussian. The objective of the subsequent sections is to calculate the probability densities of the estimated source location.

3. Source identification methods

In this work, the source localization problem is solved via time reversal and beamforming. These two methods are briefly introduced in Sections 3.1 and 3.2, respectively, and are subsequently compared in Section 3.3.

3.1. Time reversal

The reconstruction of the initial source(s) can be carried out using time-reversal technique introduced by Fink [30] at the beginning of 1990s. This technique is inspired from the symmetric nature of the wave equation, of any type, with respect to the time variable in non-dissipative media. For instance, the acoustic wave equation (Eq. (2)) remains unchanged for the time transformation \( t \rightarrow -t \). It should be pointed out that Eq. (2) in dissipative (attenuating) media has another term with a first-order time derivative which implies that the waves are \( a \text{ priori} \) no longer reversible in time. Nonetheless, a modified version of time reversal makes the source localization feasible in dissipative media (see [31] for instance). The time reversal technique consists of two steps. In the first step, an acoustic source emits the waves into a medium. These waves are then recorded via an array of transducers and in the second step, they are time-reversed and re-injected into the same propagation medium. The resolution of the refocused spot depends mainly on the amount of information from the initial source(s).
received by the transducers. In this context, the multipathing effect resulting from the wave scattering by the heterogeneities and also the reflecting boundaries can improve the resolution of the focal spot and hence the accuracy of the source localization. For instance, in the case of an open homogeneous medium, if the energy of the initial source is partially detected over a linear array of transducers of size \( a \), the time-reversed (phase conjugated in the Fourier domain) signals injected to the same medium by transducers will be refocused on the source position with a resolution of \( \lambda L/a \), in which \( \lambda \) is the wavelength and \( L \) is the propagation distance between the source and the array. This is the famous diffraction limit in homogeneous media. In heterogeneous media, due to the multipathing effect, the resolution will be enhanced [32–35]. This phenomenon is called the super-resolution and is theoretically elaborated in [34]. In the next section we introduce the basic elements related to the beamforming method.

3.2. Beamforming

In this section, the classical beamforming for homogeneous medium is introduced first. Then, its variant for inhomogeneous medium is presented. The idea of the conventional (Bartlett) beamforming is to steer the microphone array to search the source position. The Delay-and-Sum beamformer, for measured sound pressures \( p = (p_1, \ldots, p_M)^T \), is defined as

\[
B = w^T p = \sum_{m=1}^{M} p_m w_m,
\]

in which \( w = (w_1, \ldots, w_M)^T \) is the steering vector. The output power of the Delay-and-Sum beamformer is thus

\[
|B|^2 = w^T p p^T w.
\]

Here, the measurement \( p \) follows the vector \( G(x_0) \) of Green’s functions. Each element of \( G(x_0) \) represents the sound emitted from the source \( x_0 \) and received by each microphone. The beamforming technique is to decide the steering vector which maximizes the output power

\[
\hat{w} = \arg \max_w E\left( w^T p p^T w \right) = \arg \max_w (|w^T G(x_0)|^2 + \sigma^2 |w|^2).
\]

In order to obtain a non-trivial solution, \( \hat{w} \) is constrained to \( |\hat{w}| = 1 \), so that the solution of Eq. (7) is

\[
\hat{w} = \frac{G(x_0)}{|G(x_0)|}\sqrt{G(x_0)^T G(x_0)}.
\]

By substituting Eq. (8) back into Eq. (6), the source location estimate reads [9]

\[
\hat{x}_0 = \arg \max_{x_0} \left( \frac{p^T G(x_0) G(x_0)^T p}{G(x_0)^T G(x_0)} \right).
\]

The source amplitude \( A \) can also be quantified as

\[
\hat{A} = \left( G(x_0)^T G(x_0) \right)^{-1} G(x_0)^T p.
\]

It is remarkable that Eqs. (9) and (9) are maximum likelihood estimates of the source location and amplitude [36].

In the case of homogeneous medium, Green’s function can be analytically solved:

\[
G(x_0) = \left( \frac{e^{ik \| x_0 - x_i \|}}{4\pi \| x_0 - x_i \|}, \ldots, \frac{e^{ik \| x_0 - x_M \|}}{4\pi \| x_0 - x_M \|} \right)^T.
\]

in which \( j = \sqrt{-1} \) is the imaginary unity. Therefore, by substituting Eq. (11) back into Eq. (9) and solving the optimization problem with respect to \( x_0 \), the sound source can be estimated. However, in the inhomogeneous medium, Green’s function has no analytical solution. In this work, a numerical Green’s function defined on a set of discrete source locations is constructed. It is assumed that we know \textit{a priori} that the source belongs to a region \( \Gamma \). Then, this region is discretized by a grid having 5 points, denoted as \( \Gamma^d \). For each point \( x_s \) \((s = 1, \ldots, S)\) in \( \Gamma^d \), the sound propagation from this assumed source location \( x_s \) to each microphone \( x_m \) is numerically computed. Then, the numerical Green’s function based on the region \( \Gamma^d \) is constructed, denoted by \( G(x_s) \), \( s = 1, \ldots, S \). Finally, the source location is estimated among the \( S \) points by

\[
\hat{x}_0 = \arg \max_{x_s} \left( \frac{p^T G(x_s) G(x_s)^T p}{G(x_s)^T G(x_s)} \right), \quad s = 1, \ldots, S.
\]

3.3. Comparison of time reversal and beamforming

In this section, the relation and difference between time reversal and beamforming are both discussed. It is assumed that \( T \) time signals are measured by each microphone, denoted as \( p_m(t), t = 1, \ldots, T, m = 1, \ldots, M \). Given a frequency \( f \), the sound
pressure in the frequency domain used in beamforming is obtained from a discrete Fourier transform

\[ p_m(f) = \sum_{t=1}^{T} p_m(t)e^{-2\pi j ft}, \quad m = 1, \ldots, M. \]  

Taking the conjugate of both sides gives

\[ p_m^*(f) = \sum_{t=1}^{T} p_m(t)e^{2\pi j ft} = a \sum_{t=1}^{T} p_m(T+1-t)e^{-2\pi j ft}, \]  

which corresponds to the discrete Fourier transform of the reversed signal \( p_m(T+1-t), \ t = 1, \ldots, T, \) and \( a = e^{2\pi j f(T+1)} \) is a coefficient independent of \( t. \) Then, Eq. (9) is rewritten as

\[ \mathbf{x}_0 = \arg \max_{\mathbf{x}_0} \left( \frac{p_m^*(\mathbf{G}^T(\mathbf{x}_0))}{|G(\mathbf{x}_0)|^2} \right)^2 = \arg \max_{\mathbf{x}_0} \left( \frac{\sum_{m=1}^{M} G_m(\mathbf{x}_0)p_m(f)^2}{\sum_{m=1}^{M} |G_m(\mathbf{x}_0)|^2} \right), \]  

where \( |G_m(\mathbf{x}_0)| \) stands for Green’s function at the location of the \( m \)-th microphone. It is clear that the numerator in Eq. (15) corresponds to the time reversal in the frequency domain. In the case of far-field source and homogeneous medium, \( G_m(\mathbf{x}_0) \) can be approximated by a constant (independent of \( m \)), such that the beamforming and time reversal methods are identical. However, in the inhomogeneous case (especially for high heterogeneity), the influence of the denominator in Eq. (15) cannot be ignored, such that both source localization methods return different results.

4. Numerical results

4.1. Introduction

The numerical simulations of acoustic wave propagation in 2D randomly heterogeneous media in this work are carried out via the spectral element-based code, SEMLAB.\(^1\) We integrated the time reversal and beamforming methods into this code. More discussion about the applications of the spectral element method (SEM) in 2D acoustics and in 2D and 3D elastodynamics can be found in [37–40]. In this section, the parameters of the source and the sound propagation medium used in the numerical simulations are introduced first. Then, the simulation time in the SEM and the computational cost of both source localization methods are presented.

4.1.1. Source and medium parameters

The coordinates of the source are assumed as two independent Gaussian random variables as follows:

\[ X_0 \sim \mathcal{N}(\mu_{X_0} = 15 \text{ m}, \sigma_{X_0} = 1 \text{ m}); \quad Y_0 \sim \mathcal{N}(\mu_{Y_0} = 10 \text{ m}, \sigma_{Y_0} = 1 \text{ m}). \]  

The source time function is a Ricker pulse (i.e. the second derivative of a Gaussian) with a central frequency of \( f_0 = 1 \text{ Hz} \) and a delay time of \( t_0 = 1.5 \text{ s}. \) The mass density \( \rho \) is assumed to be deterministic and equal to 1. The bulk modulus is considered as a realization of a Gamma random field with an average of \( \kappa = 4 \) and a coefficient of variation equal to \( \delta = 0.5 \) (standard deviation \( \sigma_\kappa = 4\delta_\kappa = 2 \)). As a result, the dominant wavelength is \( \lambda_0 = 2 \text{ m}. \) It should be pointed out that the assumption of low-variance medium is considered in this paper since large variances induce a localization regime where the transport of the energy would not occur. The latter avoids the transition to a diffusion regime. The size of the propagation medium is \( 20 \times 20 \text{ m} \) which is divided by \( 40 \times 40 \) elements. Since in the framework of the SEM at least five control points per minimum wavelength should be used to accurately estimate the wave field [37,38,41], each of the elements is considered to contain \( 7 \times 7 \) GLL points. Hence, the total number of points of the mesh is \((6 \times 40 + 1)^2 = 58081.\) Note that the shape functions are the Lagrange polynomials of order 6. Perfectly reflecting (Neumann) boundary conditions are used which, in general, enhance the results using time reversal since the wave energy remains always in the propagation medium. The explicit Newmark scheme is used to solve the time integrations in the framework of the SEM. In addition, the Courant number is 0.4 corresponding to a time discretization of \( \Delta t = 0.005 \text{ s}. \) Two different media with correlation lengths of \( \epsilon_c = 1 \text{ m} \) and \( \epsilon_c = 10 \text{ m} \) are considered. These two cases model two randomly heterogeneous media in which \( \epsilon_c \sim \lambda \) (mesoscopic scenario) and \( \epsilon_c \gg \lambda \) (macroscopic scenario), respectively. In this work we are interested in these two scenarios and not in microscopic scenario where \( \epsilon_c \ll \lambda \) since in this case the random scattering is weak and thus the mean free path of the waves are large. These imply in particular that the transition to a diffusion regime occurs less quickly than the considered

\(^1\) This code can be found here: http://www.mathworks.com/matlabcentral/fileexchange/6154-semlab.
scenarios which makes this case less interesting. Both media are assumed to have an exponential correlation kernel
\[ R_\kappa(x, x') = \sigma^2 \exp \left( -\frac{|x-x'|}{\ell_c} \right). \] (17)

The influence of the correlation kernel on the scattering regime is discussed theoretically in [42]. In the regimes considered in this study, the choice of this function has not a considerable effect on the identification results. In the case \( \ell_c = 1 \) m, since the correlation length and the dominant wavelength are almost comparable, the interactions between the waves and the heterogeneities of the medium are more efficient compared to the case where \( \ell_c = 10 \) m. Fig. 1 shows the particular realizations of the propagation media used in this study. Fig. 2 depicts the snapshots of the wave field at \( t=4 \) s for both propagation media. It can be observed that for \( \ell_c = 1 \) m (right figure), the wave field is much more scattered by the heterogeneities and thus the wave front is more distorted in comparison with the case \( \ell_c = 10 \) m (left figure).

### 4.1.2. Simulation time

The results obtained by both source identification methods largely depend on the simulation time. This implies that the propagation time should be large enough for the waves to reach the transducers in the direct problem and to refocus to the source in the inverse process. As a result, due to high computational costs, we should properly choose the propagation time. Here we discuss why when \( \ell_c = 1 \) m, the effective phase velocity of the wave field is less than that in the case where \( \ell_c = 10 \) m as it has been observed in Fig. 2. Following [2,43], we first consider Green’s function of Eq. (2) in \( k-\omega \) (wave vector–angular frequency) domain by taking its Fourier transform with respect to \( x' \). Then taking an ensemble averaging results in the following equation for the mean Green’s function being the lowest-order approximation of the Dyson equation:

\[ E(G(k, \omega)) = \frac{1}{k_0^2 - |k|^2 - M(k, \omega)} \] (18)

wherein \( k_0 = \omega/c \) is the wavenumber in homogeneous background (with \( c \) being the average of the velocity field \( c(x) \)) and \( M(k, \omega) \) is the mass operator which depends on the type of the correlation model. The latter for an exponential correlation
function reads:

$$M(\mathbf{k}, \omega) = \frac{4\sigma_x^2 k_0^4}{(k_0^2 - \frac{i}{\ell_c})^2 - |\mathbf{k}|^2}.$$  \hspace{1cm} (19)

Inserting (19) into (18) results in

$$E(\mathbf{G}, \omega) = \frac{(k_0^2 - \frac{i}{\ell_c})^2 - |\mathbf{k}|^2}{(k_{eff_1}^2 - |\mathbf{k}|^2)(k_{eff_2}^2 - |\mathbf{k}|^2)}.$$  \hspace{1cm} (20)

We define also the dimensionless parameter $\zeta_0 = k_0\ell_c = 2\pi\ell_c/k_0$. When $8\sigma_x^2\zeta_0^2/(1 + 4\zeta_0^2) \ll 1$ holds (e.g. in low-variance media $\sigma_x^2 \ll 1$ or in low-frequencies $\zeta_0 \ll 1$), the following approximations for $k_{eff_1}$ and $k_{eff_2}$ being the poles of Eq. (20) will be obtained:

$$k_{eff_1} = \pm \left( k_0 + \sigma_x^2 k_0 \frac{2\zeta_0^2}{1 + 4\zeta_0^2} + i\sigma_x^2 k_0 \frac{4\zeta_0^4}{1 + 4\zeta_0^2} \right).$$  \hspace{1cm} (21a)

$$k_{eff_2} = \pm \left( k_0 + \frac{i}{\ell_c} \right) + O(\sigma^2).$$  \hspace{1cm} (21b)

in which the negative values correspond to the non-physical waves. The limit of the real part of $k_{eff_1}$ in low-variance media ($\sigma_x^2 \ll 1$) tends to the wavenumber of the homogeneous background (i.e. $k_0$). However, the same limit for $k_{eff_2}$ will be $k_0 + 2\sigma_x^2\zeta_0^2/k_0 > k_0$. This shows why in a heterogeneous medium the phase velocity of the mean field, being inversely proportional to the wavenumber, is smaller than that of the homogeneous background. Moreover, in inhomogeneous media, increasing $\ell_c$ will decrease the effective phase velocity. That is why the effective propagation velocity when $\ell_c = 10$ m is larger than that of the case where $\ell_c = 1$ m. For the cases where $\ell_c = 10$ m and $\ell_c = 1$ m, we, respectively, consider a propagation time of $t = 20$ s and $t = 45$ s (4000 and 9000 time steps with a sampling frequency of 200 Hz).

4.1.3. Computational cost

In this numerical experiment, the assumed source region for beamforming is $\Gamma = \{(x, y) : x \in [12, 18], y \in [7, 13]\}$. We remind that the length of the elements is 0.5 m which is discretized by 7 GLL points. Hence, the testing grid $\Gamma^q$ has $(12 \times 6 + 1)^2 = 5329$ points. Since 1000 realizations of the random sound source are considered, the full source localization computation for beamforming needs 6329 sound propagation simulations. On the other hand, time reversal simulates the sound propagation in both direct and inverse problems for each source, thus 2000 simulations are needed. Therefore, in this experiment the computational cost of beamforming is around 3 times higher than time reversal.

In the rest of this section, the sound sources are numerically estimated using both methods.

4.2. Case: $\ell_c = 10$ m

In this section the numerical results using both time reversal and beamforming will be discussed for the case where $\ell_c = 10$ m. The 1000 realizations of source position are estimated using both methods. For each source position, the sound propagation process is simulated and the velocity field is measured via 21 microphones located at $(5, y_m)$. $y_m = 0, 1, 2, \ldots, 20$. Note that the spacing between the microphones is considered to be equal to $\lambda_0/2 = 1$ m following the Nyquist–Shannon sampling law. Instead of deciding the source location by eyes, here the source estimate using time reversal is decided by computing the maximum value of the recorded wave field during the inverse step. Consequently, the identified source locations lie always on the GLL points of the rectangular mesh. The simulation time for both direct and inverse problems is considered to be 20 s. The 1000 source realizations, characterized with independent normally distributed $x$- and $y$-coordinates parameterized in Section 2.2, are shown in Fig. 3(a). The plots (b) and (c) depict the estimated source locations using time reversal and beamforming, respectively. Since the beamforming method is based on the numerical Green’s function defined on the GLL points, the source estimates are also located at these discrete points. It should also be noted that the test grid of source location, as well as the grid resolution, used for both methods are the same, which are GLL points of the given region.

Fig. 4 shows the convergence of the mean (plots (a) and (b)) and standard deviations (plots (c) and (d)) of the $x$- and $y$-coordinates of the identified sources in terms of the number of source realizations. In these figures, the solid and dashed curves indicate the results obtained using time reversal and beamforming methods, respectively. The horizontal solid lines represent the analytical values. Fig. 5 depicts the corresponding relative errors (RE) in percent. It should be noted that in this paper, we used a moving average whose width is 50 samples in order to smoothen the curves. Fig. 5(a) and (b) shows that the average value estimates of the $x$- and $y$-coordinates obtained using beamforming have larger error than using time reversal. Fig. 5(c) and (d) shows that the standard deviations of the $x$- and $y$-coordinates using both methods have more or
less the same estimation error. We emphasize that the analytical values are not exactly the values given in Section 2.2. The former are calculated from the 1000 source realizations.

A more general information is the PDF of the estimated source positions. Hence, we also compare the PDF of each coordinate obtained from the 1000 source realizations with that of the identified sources. In Fig. 6, the dashed, dash-dot and dotted curves indicate the estimated PDF of the source coordinates using, respectively, 1000, 500 and 100 samples. The solid curves show the analytical PDF calculated over 1000 source realizations. The dashed and the solid curves are almost
perfectly overlapped. The former obviously describe the analytical PDF better compared to the dash-dot and dotted curves especially around the mean values. The plots (c) and (d) display the corresponding results of beamforming which indicate that the dashed curves (1000 samples) satisfactorily estimate the PDFs.

In what follows, in order to compare more precisely the results obtained using both methods, the percentiles of the source coordinates are computed. More precisely, 1st, 25th, 75th and 99th percentiles of the $x$- and $y$-coordinates are computed, which describe the estimation of the left far, left near, right near and right far tails of the PDFs, respectively. Fig. 7 displays the REs in percent of these estimates. The 1st percentile of the $x$- and $y$-coordinates converge, respectively, at about (250,350) and (600,450) in the sense of RE being less than 1%. The 25th percentile of the $x$- and $y$-coordinates converge at about 50 samples for both methods. The convergence of the 75th percentile of both coordinates occurs at (50,50) and (150,300), respectively. However, it can be observed that the RE corresponding to the beamforming method is larger than that of the time reversal. The 99th percentiles of the $x$- and $y$-coordinates converge at around (700,400) and (250,250) samples, respectively. Generally speaking, the far percentiles (1st and 99th) need more data than the near percentiles (25th and 75th) to guarantee the convergence.

Higher-order statistics of the corresponding random variables could be of importance. Skewness as a measure of symmetry and kurtosis as a measure of the tailedness of the PDFs are estimated. Remember that for any univariate normal PDF, the analytical values are 0 and 3 for skewness and kurtosis, respectively. Fig. 8 shows these estimates for both $x$- and $y$-coordinates in terms of the simulation numbers along with the corresponding analytical values which are obtained from the 1000 realizations of the random sound source. Fig. 8 shows that all the estimates converge to their analytical values except the kurtosis of $x$-coordinate by beamforming (the dashed curve in plot (c)), which stabilizes after 500 number of simulations but underestimates the actual kurtosis.

Fig. 5. Relative errors on average ((a) and (b)) and standard deviation ((c) and (d)) of $x$ (left) and $y$ (right) components of the source position in terms of the number of simulations. Solid and dashed curves are related to time reversal and beamforming, respectively. Case: $\zeta_i = 10$ m, $t = 20$ s.

\[2\] Here, the ordered pair $(N_1, N_2)$ indicates that time reversal and beamforming converge, respectively, at $N_1$ and $N_2$ samples.
Regarding the time reversal method, the resolution of the refocused wave field is one of the main concerns. As mentioned before, during the inverse simulations, we identify the point at which the field value is maximized. Fig. 9 shows the snapshots of the wave field at the identified point following x and y directions. The resolution in each direction is defined as an interval which covers the maximum point and its left and right endpoints at which the values of the wave field become negligible (less than $10^{-3}$). Fig. 10 shows the PDFs of the resolutions calculated following x (left plot) and y (right plot) directions. The maxima (i.e. the most probable resolutions) are observed at about $\text{res}_x = 1.5$ m and $\text{res}_y = 2.5$ m, respectively, in x and y directions.

In order to show how the simulation time can influence the results obtained by the time reversal method, computation results using a 15 s measurement (3000 time steps) are shown as follows. As it can be observed from Fig. 11(a) and (b), since the observation time is not large enough, the source reconstruction in this case fails. However, plot (c) shows that beam-forming results in much better estimations. It is obvious that time reversal is more sensitive to the simulation time, since it needs to receive more signals in the direct process to refocus the sound source in the inverse problem, in particular for this experimental setup that the microphones do not surround the source. Therefore, the simulation time in time reversal should be properly chosen according to the ratio between the correlation length $\ell_c$ and wavelength $\lambda$. On the other hand, beam-forming hardly depends on the simulation time and it can work once the sound wave propagates to and measured by the microphones.

### 4.3. Case: $\ell_c = 1$ m

In this section, the source identification in the second propagation medium with $\ell_c = 1$ m is carried out. Using time reversal and beamforming methods with a propagation time of $t=45$ s (i.e. 9000 time steps), the distributions of 1000 source realizations along with the corresponding identified source locations are shown in Fig. 12. Fig. 13 depicts the

![Fig. 6. PDFs of the estimated x (left) and y (right) components of the random source using 1000 (dashed curves), 500 (dash-dot curves) and 100 (dotted curves) samples using time reversal ((a) and (b)) and beamforming ((c) and (d)). The solid curves represent the analytical PDFs. Case: $\ell_c = 10$ m, $t=20$ s.](image)
Fig. 7. Relative errors on 1st ((a) and (b)), 25th ((c) and (d)), 75th ((e) and (f)) and 95th ((g) and (h)) percentiles of the estimated $x$ (left) and $y$ (right) components using time reversal (solid curves) and beamforming (dashed curves). Case: $\ell_c = 10$ m, $t = 20$ s.
stabilization of the mean (plots (a) and (b)) and standard deviation (plots (c) and (d)) of the estimated source coordinates towards the analytical values related to the 1000 source realizations. Fig. 14 shows the corresponding RES. The averages of x- and y-coordinates converge fast at around 50 samples for both methods within a relative error of 1%. Fig. 14(c) and (d) shows that the error of standard deviation estimate using time reversal is generally larger than beamforming.

**Fig. 8.** Skewness ((a) and (b)) and kurtosis ((c) and (d)) of the estimated x (left) and y (right) components using time reversal (solid curves) and beamforming (dashed curves). Case: $r_s = 10$ m, $t=20$ s.

**Fig. 9.** Definition of the resolution of the focal spot following x (left) and y (right) directions.

stabilization of the mean (plots (a) and (b)) and standard deviation (plots (c) and (d)) of the estimated source coordinates towards the analytical values related to the 1000 source realizations. Fig. 14 shows the corresponding RES. The averages of x- and y-coordinates converge fast at around 50 samples for both methods within a relative error of 1%. Fig. 14(c) and (d) shows that the error of standard deviation estimate using time reversal is generally larger than beamforming.
Fig. 10. PDFs of the resolution of the focal spot following x (left) and y (right) directions. Case: $\ell_c = 10$ m, $t=20$ s.

Fig. 11. 1000 realizations of the random sound source (a) and the identified source positions using time reversal (b) and beamforming (c). Case: $\ell_c = 10$ m, $t=15$ s.

Fig. 12. 1000 realizations of the random sound source (a) and the identified source positions using time reversal (b) and beamforming (c). Case: $\ell_c = 1$ m, $t=45$ s.

Fig. 15 represents the PDFs of the estimates of x (left) and y (right) components using time reversal (plots (a) and (b)) and beamforming (plots (c) and (d)), respectively. It can be observed that the PDF obtained via beamforming using 1000 source samples (dashed curves) describes better the analytical PDF (the solid curve) than 500 and 100 samples (dash-dot and dotted curves) specially near mean values. The plots (a) and (b) reveal also that time reversal does not yield good estimations of the PDFs near the mean values, especially for the y-coordinate. Fig. 16 depicts the REs of 1st, 25th, 75th and 99th percentile estimates of the x- and y-coordinates. Fig. 16(a) shows that the first percentile of the estimated x-coordinate
converges almost at 650 samples for both methods; the first percentile of the y-coordinate converges at about 450 and 500 simulation numbers using time reversal and beamforming, respectively. Plots (c) and (d) show a fast convergence at about 50 samples for both methods. Plots (e) and (f) indicate that the REs of time reversal are always higher than using beamforming, although both estimates converge within a reasonable simulation number (within 300). However, Fig. 16(g) and (h) shows that the 99th percentile estimate of the x-coordinate using beamforming cannot converge. The other three quantities converge in the sense of RE less than 1% but obviously slower than the 75th percentile. As in the case of $\ell_c = 10$ m, the far tail percentile estimates generally need more data to converge than the near tail percentiles.

Fig. 17 shows that the skewness and kurtosis estimates with respect to the simulation number. Plot (a) indicates that the skewness estimate of the x-coordinate using time reversal converges to its analytical value whereas beamforming underestimates it. By contrast, the y-coordinate is accurately estimated via beamforming but the estimate using time reversal is lower than the actual value. Plots (c) and (d) show that both methods underestimate the analytical values of the kurtosis for both source components and result in more leptokurtic distributions. Overall, the estimates of both the third and fourth order moments of the source positions in this case are less accurate than the case $\ell_c = 10$ m.

The PDFs of the focal spot resolution in both directions are plotted in Fig. 18. The most probable values for the resolutions following the x and y directions are both around $\text{res}_x = \text{res}_y = 2$ m.

The results obtained in this section are not so robust as the case where $\ell_c = 10$ m. One of the main reasons is that in the case $\ell_c = 1$ m, the sound propagation medium is more complicated due to the high degree of interaction between the waves and the heterogeneities. The same simulations with $t = 40$ s are also preformed but the convergence does not happen; the proposed simulation time $t = 45$ s is a reasonable time which considers both the estimation accuracy and the computational cost.
4.4. Concluding remarks

In this section, some concluding remarks from the results of our numerical experiments are given as follows.

1. The simulation time is crucial, especially for time reversal which needs more received signals to refocus the sound source in the inverse problem. The required simulation time depends on the ratio $\frac{\ell c}{\lambda_0}$; in the case where $\frac{\ell c}{\lambda_0} \approx 1$, it is larger than the case where $\frac{\ell c}{\lambda_0} \gg 1$.

2. The average of the random source position can be accurately estimated with a small number of data. The variance estimates are less accurate in the sense of relative error and need more data for convergence, particularly for the case $\frac{\ell c}{\lambda_0} \approx 1$ in which the medium has higher scattering degrees.

3. The PDF of the source coordinates can be replicated using the source estimates of 1000 realizations, except the case $\ell_c = 1$ m ($\ell_c/\lambda_0 \approx 1$) using the time reversal method.

4. The estimates of 1st and 99th percentiles of source distribution, which are further away from the center (mean value), need more data to guarantee the convergence (in the sense of RE less than 1%) than 25th and 75th percentiles which are nearer the center. In the case where $\ell_c = 1$ m, the former two quantities are more difficult to estimate in terms of a larger number of data to guarantee the convergence. In particular, the 99th percentile estimate of the $x$-coordinate using beamforming does not converge.

5. The skewness and kurtosis of the random source are more precisely estimated in the case $\ell_c = 10$ m than the case $\ell_c = 1$ m.

---

**Fig. 14.** Relative errors on average ((a) and (b)) and standard deviation ((c) and (d)) of $x$ (left) and $y$ (right) components of the source position in terms of the number of simulations. Solid and dashed curves are related to time reversal and beamforming, respectively. Case: $\ell_c = 1$ m, $t=45$ s.
6. From above 2–5, we can conclude that the source estimate in the case where $\ell_c = 1 \text{ m} \left( \ell_c/\lambda_0 \sim 1 \right)$ is less robust and in general needs more data to guarantee the convergence of statistics of random source.

Both time reversal and beamforming have advantages and disadvantages. In this work, a numerical version of beamforming is used to deal with the inhomogeneous medium: a set of discrete points are considered as the candidates of the sound source estimates. In this case, the final estimate of beamforming may stop at a local maximum of the beamforming output, instead of the global maximum, which therefore returns an inaccurate estimate. A denser grid of discrete points could increase the accuracy of the source estimates, but the computational cost would also be largely increased. As indicated in Section 4.1.3, the beamforming method needs a predefined region which leads to a high computational cost when no prior information of source location can be given. On the other hand, time reversal highly depends on the simulation time of sound propagation: when the received signals are not enough, this approach does not work. By contrast, beamforming can return a source estimate once the sound waves are measured by the microphones in the direct problem.

5. Conclusions and perspectives

In this paper, the localization of random sound source is investigated. A spatially varying medium is considered and the sound propagation simulation in this medium is realized using the spectral element method. The sound source is assumed to follow a Gaussian distribution. For each realization of the random sound source, its location is estimated using the time reversal and beamforming methods.

In the simulation experiment, the sound propagation simulation and the source localization are considered in two different media: a less heterogeneous medium ($\ell_c = 10 \text{ m}, \lambda_0 = 2 \text{ m}$) and a highly scattering medium with a correlation length comparable to the typical wavelength ($\ell_c = 1 \text{ m} \sim \lambda_0 = 2 \text{ m}$). The computation results show that sound propagation speed and the necessary simulation time for source localization highly depend on the ratio between the correlation length
and the dominant wavelength ($\ell_c/\lambda_0$). The statistical properties, including expectation, variance, probability density function and percentiles, are all estimated. The estimation robustness also depends on the ratio $\ell_c/\lambda_0$: as it decreases these estimates generally need more data to guarantee the convergence.

Fig. 16. Relative errors on 1st ((a) and (b)), 25th ((c) and (d)), 75th ((e) and (f)) and 99th ((g) and (h)) percentiles of the estimated $x$ (left) and $y$ (right) components using time reversal (solid curves) and beamforming (dashed curves). Case: $\ell_c = 1$ m, $t = 45$ s.
The analysis in this work raises the problem of sensitivity to uncertainties in realistic problems in which the medium is not deterministic. Future studies will therefore deal with the robust localization methods for stochastic problems based on response surface building and advanced uncertainty quantification methods, such as polynomial chaos expansion or advanced kriging methods.

Fig. 17. Skewness plots (a) and (b)) and kurtosis (plots (c) and (d)) of the estimated x (left) and y (right) component using time reversal (solid curves) and beamforming (dashed curves). Case: $e^{-1} = 1 \text{ m, } t=45 \text{ s}$.

Fig. 18. PDFs of the resolution of the focal spot following x (left) and y (right) directions. Case: $e^{-1} = 1 \text{ m, } t=45 \text{ s}$.
Acknowledgments

The authors would like to thank Régis Cottereau for helpful discussions and also two anonymous reviewers for their constructive comments that greatly contributed to improving the final version of the paper. This work has been carried out thanks to the support of the A*MIDEX grant (no. ANR-11-IDEX-0001-02) funded by the French Government “Investissements d’Avenir” program.

References


