Sound Source Localization from Uncertain Information Using the Evidential EM Algorithm

Xun Wang[1,2,⋆], Benjamin Quost[1], Jean-Daniel Chazot[2], and Jérôme Antoni[3]

[1] Heudiasyc, UMR CNRS 7253, Université de Technologie de Compiègne, Compiègne, France
xun.wang@hds.utc.fr

[2] Roberval, UMR CNRS 7337, Université de Technologie de Compiègne, Compiègne, France

[3] LVA, INSA Lyon, Lyon, France

Abstract. We consider the problem of sound sources localization from acoustical measurements obtained from a set of microphones. We formalize the problem within a statistical framework: the pressure measured by a microphone is interpreted as a mixture of the signals emitted by the sources, pervaded by a Gaussian noise. Maximum-likelihood estimates of the parameters of the model (locations and strengths of the sources) may then be computed via the EM algorithm. In this work, we introduce two sources of uncertainties: the location of the microphones and the wavenumber. First, we show how these uncertainties may be transposed to the data using belief functions. Then, we detail how the localization problem may be studied using a variant of the EM algorithm, known as Evidential EM algorithm. Eventually, we present simulation experiments which illustrate the advantage of using the Evidential EM algorithm when uncertain data are available.

Keywords: Localization of sound sources, Inverse problem, EM algorithm, Belief function, Evidential EM algorithm, Uncertain data.

1 Introduction

In this paper, we consider the problem of sound sources localization. We assume there exist N sound sources on the plane, and our aim is to determine their position using measurements made by an array of microphones. The sound pressure measured by each microphone is interpreted as a superimposed signal composed of N components, each of which has been emitted by a sound source. Our purpose is to estimate the locations and strengths of the sound sources.

Feder and Weinstein [9] investigated the parameter estimation problem of superimposed signals using the EM algorithm. This approach makes it possible to compute iteratively maximum-likelihood (ML) estimates of the parameters of a model which depends on unobserved (or "complete") variables. In the case of sound source estimation, each observed variable (the pressure measured by a microphone) is the sum of several complete ones (the signals propagated by the

⋆ Corresponding author.

W. Liu, V.S. Subrahmanian, and J. Wijsen (Eds.): SUM 2013, LNAI 8078, pp. 162–175, 2013.

sources towards this microphone). The EM algorithm proceeds with the complete likelihood (that is, the likelihood of the complete variables), the expectation of which is maximized at each iteration. Cirpan and Cekli [2,3] and Kabaoglu et al. [10] studied the localization of near-field sources using the EM algorithm. This work is a particular case of [9], in which the sources are located using polar coordinates. In our work, we investigated the problem of sound source localization. However, unlike in [2,3,10], we use general coordinates, and we explicitly describe the propagation process from the sources to the microphones using a specific operator.

It should be stressed out that uncertainties may pervade the measurement process, so that the sound pressures received by the microphones are not exactly known. For instance, it may be difficult to perfectly assess the positions of the microphones, for example due to the vibration of the antenna on which they are set. The medium may also be the cause to some uncertainties: in particular, the wavenumber can vary, due to a significant variation of the temperature between the sound sources and the microphones. In this paper, we present how both these sources of uncertainty may be taken into account in the localization process.

The theory of belief functions, also known as Dempster-Shafer theory, is a powerful tool for managing and mining uncertain data. The theory was developed by Dempster and Shafer [5,6,13]. The problem of statistical inference was addressed in [13], and developed by Denoeux [7]. In this latter work, the author proposed a framework in which data uncertainty is represented using belief functions. He then introduced an extension of the EM algorithm, the Evidential EM (E2M) algorithm, which makes it possible to estimate the parameters of the model from such uncertain data. In this paper, we will adopt this approach for representing the uncertainties arising from ill-known microphone locations and wavenumber, and for estimating the locations and strengths of the sound sources.

The organization of this paper is as follows. In Section 2, we give the basic description of the sound sources localization model in the case of precise data and we show how the EM algorithm may be used to solve the parameter estimation problem. In Section 3, we show how imprecise microphone locations and wavenumber may induce an uncertainty on the sound pressures measured by the microphones. Then, in Section 4, we detail our parameter estimation method for such uncertain data using the E2M algorithm. Finally, Section 5 presents simulation experiments which show the advantage of the E2M algorithm in coping with uncertain data, and Section 6 concludes the paper.

2 Sound Source Localization via the EM Algorithm

2.1 Basic Description of the Model

We assume that there are M microphones on a line, with known locations $(\theta_m, 0), m = 1, \ldots, M$. We consider N sound sources, the coordinates of the n-th source being $(\xi_n, \eta_n), n = 1, \ldots, N$. The signal received by the m-th microphone in each snapshot is the sum of components from different sources, altered by a complex-valued Gaussian distributed noise:

$$x_p = G(\xi, \eta)A + \epsilon_p. \tag{1}$$

Here $x_p = (x_{1p}, \ldots, x_{Mp})^T$ is the vector of pressures measured by the M microphones in the p-th snapshot. The M-by-N matrix $G(\xi, \eta) = (G(\xi_n, \eta_n, \theta_m))_{m=1,n=1}^{M,N}$ describes the sound propagation process: the (m, n)-th entry of this matrix is the Green function $G(\xi_n, \eta_n, \theta_m) = \frac{e^{jk\sqrt{(\xi_n-\theta_m)^2+\eta_n^2}}}{2\pi\sqrt{(\xi_n-\theta_m)^2+\eta_n^2}}$, that is the operator which transforms the signal emitted by the n-th source into the signal received by the m-th microphone. The vector $A = (A_1, \ldots, A_N)^T$ contains the strengths of the sound sources. Eventually, $\epsilon_p = (\epsilon_{1p}, \ldots, \epsilon_{Mp})^T$ is a complex Gaussian-distributed noise. Note that $Re(\epsilon_p)$ and $Im(\epsilon_p)$ are independent $N(\mathbf{0}, \frac{\sigma^2}{2} I_M)$ M-dimensional random variables (here, I_M stands for the M-by-M identity matrix).

2.2 The EM Algorithm

In this section, we give a brief introduction of the EM algorithm [4]. Let us denote by x the incomplete or observed data, with probability density function (pdf) $g(x|\Phi)$. Similarly, y stands for the complete (unknown) data, with pdf $f(y|\Phi)$. Both f and g depend on the parameter vector Φ, which is to be estimated by maximizing the log-likelihood of the observed data (observed log-likelihood)

$$L(\Phi) = \log g(x|\Phi),$$

over Φ. For this purpose, the EM algorithm proceeds with the complete log-likelihood by iterating back and forth between two steps:

– the E-step, where the expectation $Q(\Phi|\Phi') = E(\log f(y|\Phi)|x, \Phi')$ over the unknown variables is computed, knowing the parameter vector Φ' estimated at the previous iteration;
– the M-step, where ML estimates of the parameters are determined by maximizing $Q(\Phi|\Phi')$ with respect to Φ.

It may be shown that, under regularity conditions, the EM algorithm converges towards a local maximum of the observed log-likelihood [4,17].

2.3 Model Estimation Using the EM Algorithm

We now present how the problem of sound source localization may be addressed using the EM algorithm. Let $x = (x_1, \ldots, x_P)$ be the vector of observed data. For each observed vector of pressures x_p, the complete data are the contributions $y_p = (y_{1p}, \ldots, y_{Np})$ of the sound sources to these measured pressures: each vector y_{np} represents the set of pressures emitted by the n-th source and received by the microphones. Thus, x_p is related to the y_{np} by

$$x_p = \sum_{n=1}^{N} y_{np}.$$

We have $y_{np} \propto N(G_n(\xi_n, \eta_n)A_n, \frac{\sigma^2}{N} I_M)$, where $G_n(\xi_n, \eta_n)$ is the n-th column of the matrix $G(\xi, \eta)$. Since the joint pdf of the complete data over all snapshots is

$$f(y; \xi, \eta, A, \sigma) = \prod_{p=1}^{P} \prod_{n=1}^{N} f(y_{np}; \xi_n, \eta_n, A_n, \sigma),$$

we may write the complete log-likelihood:

$$\log L(y; \xi, \eta, A, \sigma) = -2MNP \log \sigma - \frac{N}{\sigma^2} \sum_{p=1}^{P} \sum_{n=1}^{N} |y_{np} - G_n(\xi_n, \eta_n)A_n|^2 . \quad (2)$$

Then, given the parameters $\Phi^l = (\xi^l, \eta^l, A^l, \sigma^l)$ estimated at iteration l, the $(l+1)$-th iteration of EM algorithm consists in the E- and M-steps.

E-step: compute $Q(\Phi|\Phi^l) = E(\log L(y; \xi, \eta, A, \sigma)|x, \xi^l, \eta^l, A^l, \sigma^l)$. For this purpose, let us remind the following theorem ([15]):

Theorem 1. *Let X and Y be n-dimensional Gaussian random vectors with expectation m_X and m_Y and with covariance matrix Σ_{XX} and Σ_{YY}. Let $\Sigma_{XY} = Cov(X,Y)$ and $\Sigma_{YX} = Cov(Y,X)$, then the conditional pdf of Y given X is*

$$N(m_Y + \Sigma_{YX}\Sigma_{XX}^{-1}(x - m_X), \Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}).$$

Since $(x_p, y_{np})^T$ are jointly Gaussian with expectation $(G(\xi, \eta)A \; G_n(\xi_n, \eta_n)A_n)^T$, assuming that y_{n_1p} and y_{n_2p} are uncorrelated for $n_1 \neq n_2$, the covariance matrix of $(x_p, y_{np})^T$ is $\begin{pmatrix} \sigma^2 I_M & \frac{\sigma^2}{N} I_M \\ \frac{\sigma^2}{N} I_M & \frac{\sigma^2}{N} I_M \end{pmatrix}$. Then by Theorem 1, we obtain the conditional expectation of y_{np}:

$$E(y_{np}|x, \xi^l, \eta^l, A^l, \sigma^l) = G_n(\xi_n^l, \eta_n^l)A_n^l + \frac{1}{N}(x_p - G(\xi^l, \eta^l)A^l); \quad (3)$$

in the following, we will denote this expectation by v_{np}^l. For each n, we have

$$\begin{aligned} &\arg \max_{\xi_n, \eta_n, A_n} Q(\Phi|\Phi^l) \\ &= \arg \min_{\xi_n, \eta_n, A_n} \sum_{p=1}^{P} E\left(|y_{np} - G_n(\xi_n, \eta_n)A_n|^2 \,|x, \xi^l, \eta^l, A^l, \sigma^l\right) \\ &= \arg \min_{\xi_n, \eta_n, A_n} \sum_{p=1}^{P} \left|v_{np}^l - G_n(\xi_n, \eta_n)A_n\right|^2 . \end{aligned} \quad (4)$$

M-step: compute $\Phi^{l+1} = (\xi^{l+1}, \eta^{l+1}, A^{l+1}, \sigma^{l+1})$ so as to maximize $Q(\Phi; \Phi^l)$.

The update equation for the strength A_n of the sources is obtained from (4), for $n = 1 \dots, N$:

$$A_n^{l+1} = \arg\min_{A_n} \left|e_n^l - G_n(\xi_n, \eta_n)A_n\right|^2 = \frac{G_n(\xi_n, \eta_n)^H e_n^l}{G_n(\xi_n, \eta_n)^H G_n(\xi_n, \eta_n)}, \tag{5}$$

in which we write

$$e_n^l = \frac{1}{P}\sum_{p=1}^{P} v_{np}^l. \tag{6}$$

Therefore, the estimates for the source location are obtained by minimizing

$$\frac{1}{P}\sum_{p=1}^{P}\left|v_{np}^l - G_n(\xi_n, \eta_n)\frac{G_n(\xi_n, \eta_n)^H e_n^l}{G_n(\xi_n, \eta_n)^H G_n(\xi_n, \eta_n)}\right|^2,$$

which gives

$$\begin{aligned}\xi_n^{l+1}, \eta_n^{l+1} &= \arg\min_{\xi_n,\eta_n}\left|e_n^l - G_n(\xi_n, \eta_n)\frac{G_n(\xi_n, \eta_n)^H e_n^l}{G_n(\xi_n, \eta_n)^H G_n(\xi_n, \eta_n)}\right|^2 \\ &= \arg\max_{\xi_n,\eta_n}(e_n^l)^H \frac{G_n(\xi_n, \eta_n)G_n(\xi_n, \eta_n)^H}{G_n(\xi_n, \eta_n)^H G_n(\xi_n, \eta_n)} e_n^l. \end{aligned} \tag{7}$$

Finally, by computing and maximizing $E\left(\log L(y; \xi^{l+1}, \eta^{l+1}, A^{l+1}, \sigma)|x, \xi^l, \eta^l, A^l, \sigma^l\right)$ with respect to σ, we obtain the estimate of the variance σ^2:

$$(\sigma^2)^{l+1} = \frac{1}{MP}\sum_{p=1}^{P}\sum_{n=1}^{N}\left[\frac{M(N-1)(\sigma^l)^2}{N^2} + \left|v_{np}^l - G_n(\xi_n^{l+1}, \eta_n^{l+1})A_n^{l+1}\right|^2\right]. \tag{8}$$

We can see that (7) is a 2-parameter optimization problem. If we assume further that all noise sources are on a line ($\eta_n = z$), it boils down to a single-parameter optimization problem, which is easy to solve.

Eventually, the strategy for estimating the parameters of the model using the EM algorithm may be summarized as follows:

1. For $l = 0$, pick starting values for the parameters $\xi^0, \eta^0, A^0, \sigma^0$.
2. For $l \geq 1$:
   - obtain e_n^l from (6),
   - obtain $\xi_n^{l+1}, \eta_n^{l+1}$, for $n = 1, \dots, N$ from (7);
   - obtain $A_n^{l+1}, n = 1, \dots, N$ by substituting $\xi_n^{l+1}, \eta_n^{l+1}$ back into (5);
   - obtain $(\sigma^2)^{l+1}$ from (8).
3. Continue this process until convergence: stop when the relative increase of the observed data (incomplete data) log-likelihood is less than a given threshold κ:

$$\frac{\log L(x; \Phi^{l+1}) - \log L(x; \Phi^l)}{\log L(x; \Phi^l)} < \kappa. \tag{9}$$

We remark that the computation of $(\sigma^2)^{l+1}$ from (8) may be skipped if we just care the situation of the sound sources, since the estimates of the location and the strength do not depend on σ^2.

3 Uncertainty Representation Using Belief Functions

3.1 Uncertain Measurements

In practice, uncertainty may pervade the measurement process, so that the sound pressures measured by microphones are not precise. For example, it may be difficult to give an exact location for the microphones due to the vibration of the antenna. The medium may also be the cause to some uncertainties: e.g., the wavenumber may be ill-known, due to a significant variation of the temperature and thus of the sound velocity between the sound sources and the microphones.

In the next subsection, we will give a short introduction of the belief functions theory, which is a powerful tool for representing and managing uncertain information. Then we will describe how the uncertainties on the microphone locations and wavenumber may be transferred to the observed data in the belief function framework.

3.2 Belief Functions

Let X be a variable taking values in a finite domain Ω. Uncertain information about X may be represented by a *mass function* $m^{\Omega} : 2^{\Omega} \to [0,1]$, where 2^{Ω} stands for the power set of Ω, such that $\sum_{A \subseteq \Omega} m(A) = 1$. Any subset A of Ω such that $m(A) > 0$ is called *focal element* of m. A mass function m may also be represented by its associated *belief* and *plausibility functions*. Both are defined for all $A \subseteq \Omega$, by:

$$Bel(A) = \sum_{B \subseteq A} m(B), \quad Pl(A) = \sum_{B \cap A \neq \phi} m(B).$$

We can interpret $Bel(A)$ as the degree to which the evidence supports A, while $Pl(A)$ can be interpreted as an upper bound on the degree of support that could be assigned to A if further evidence was available. Eventually, note that the function $pl : \Omega \to [0,1]$ such that $pl(\omega) = Pl(\{\omega\})$ is the *contour function* associated to m^{Ω}.

Belief Functions on the Real Line. Here we consider the case in which the domain $\Omega_X = \mathcal{R}$. In this case, a mass density can be defined as a function m from the set of closed real intervals to $[0, +\infty)$ such that $m([u,v]) = f(u,v)$ for all $u \leq v$, where f is a two-dimensional probability density function with support in $\{(u,v) \in \mathcal{R}^2 : u \leq v\}$. The intervals $[u,v]$ such that $m([u,v]) > 0$ are called *focal intervals* of m. The contour function pl corresponding to m is defined by the integral:

$$pl(x) = \int_{-\infty}^{x} \int_{x}^{+\infty} f(u,v) dv du.$$

One important special case of continuous belief functions are Bayesian belief functions, for which focal intervals are reduced to points. Then the two-dimensional pdf has the following form: $f(u,v) = p(u)\delta(u-v)$, where p is a univariate pdf and δ is the Dirac delta function. If we assume further that p is a Gaussian pdf, then $pl(x)$ is Gaussian contour function.

3.3 Uncertain Data Model Using Belief Functions

We consider here the same model as described in Section 2. We have M microphones with coordinates $(\theta_m, 0), m = 1, \ldots, M$, and N sound sources, with coordinates $(\xi_n, \eta_n), n = 1, \ldots, N$. The complete data are the signals emitted by the sources: $y_p = (y_{1p}, y_{2p}, \ldots, y_{Np})$, in which $y_{np} = G_n(\xi_n, \eta_n)A_n + \epsilon_{np}$, such that $x_p = \sum_{n=1}^{N} y_{np}$ for all $p = 1, \ldots, P$. In this section, we present how the uncertainties on the microphone locations and on the wavenumber may be transposed to the data. A variance estimation technique via first-order Taylor expansion [1] is used here.

Assume X_1 and X_2 are two independent real-valued random variables and $E(X_1) = \mu_1$, $E(X_2) = \mu_2$, $\mathrm{Var(X_1)} = \sigma_1^2$, $\mathrm{Var(X_2)} = \sigma_2^2$. Furthermore, we assume $f(x_1, x_2)$ is second-order differentiable and has real-valued inputs and complex outputs. By first-order Taylor expansion in (μ_1, μ_2) we have

$$f(x_1, x_2) \approx f(\mu_1, \mu_2) + \frac{\partial f}{\partial x_1}(\mu_1, \mu_2)(x_1 - \mu_1) + \frac{\partial f}{\partial x_2}(\mu_1, \mu_2)(x_2 - \mu_2),$$

and by the independence of X_1 and X_2, we obtain

$$E\left(|f(X_1, X_2) - f(\mu_1, \mu_2)|^2\right) \approx \left|\frac{\partial f}{\partial X_1}(\mu_1, \mu_2)\right|^2 \sigma_1^2 + \left|\frac{\partial f}{\partial X_2}(\mu_1, \mu_2)\right|^2 \sigma_2^2.$$

Let us remind that the p-th pressure measured by the m-th microphone is

$$x_{mp} = \sum_{n=1}^{N} y_{mnp} = \sum_{n=1}^{N} G(\xi_n, \eta_n, \theta_m)A_n = \sum_{n=1}^{N} \frac{e^{jk\sqrt{(\xi_n - \theta_m)^2 + \eta_n^2}}}{2\pi\sqrt{(\xi_n - \theta_m)^2 + \eta_n^2}} A_n,$$

for $m = 1, \ldots, M, p = 1, \ldots, P$. Assume that the imprecise knowledge of the microphone locations and of the wavenumber is expressed by variances σ_θ^2 (assumed to be the same for all the microphones) and σ_k^2. First, we remark that

$$\frac{\partial x_{mp}}{\partial k} = \sum_{n=1}^{N} \frac{jA_n}{2\pi} e^{jkr_{mn}}, \tag{10}$$

$$\frac{\partial x_{mp}}{\partial \theta_m} = \sum_{n=1}^{N} \frac{A_n}{2\pi} \left[\frac{jk}{r_{mn}} e^{jkr_{mn}} - \frac{1}{r_{mn}^2} e^{-jkr_{mn}}\right] \left(-\frac{\xi_n - \theta_m}{r_{mn}}\right), \tag{11}$$

$m = 1, \ldots, M$, where $r_{mn} = \sqrt{(\xi_n - \theta_m)^2 + \eta_n^2}$ is the distance from the n-th source to the m-th microphone.

Then, we can transfer the uncertainties on the microphone locations and on the wavenumber to the measured pressure. We assume that our imprecise knowledge on the actual measured pressure may be represented using a Gaussian contour function $N(\mu_{mp}^*, \sigma_{mp}^2)$ with expectation $\mu_{mp}^* = x_{mp}$ and variance

$$\sigma_{mp}^2 \approx \left|\frac{\partial x_{mp}}{\partial k}\right|^2 \sigma_k^2 + \left|\frac{\partial x_{mp}}{\partial \theta_m}\right|^2 \sigma_\theta^2. \tag{12}$$

Note that this variance is a function of ξ_n, η_n and A_n for $n = 1, \ldots, N$.

4 Sound Source Localization from Credal Data

4.1 Likelihood Function of a Credal Sample

Let Y be a discrete random vector taking values in Ω_Y, with probability function $p_Y(y|\Phi)$. Let y denote a realization of Y, referred to as the *complete* data. In some cases, y is not precisely observed, but it is known for sure that $y \in A$ for some $A \subseteq \Omega_Y$. The likelihood function given such imprecise data is:

$$L(\Phi; A) = p_Y(A; \Phi) = \sum_{y \in A} p_Y(y|\Phi).$$

More generally, our knowledge of y may be not only imprecise, but also uncertain; it can be described by a mass function m on Ω_Y with focal elements $A_1, \ldots, A_r$ and corresponding masses $m(A_1), \ldots, m(A_r)$. To extend the likelihood function, Denoeux [7] proposes to compute the weighted sum of the terms $L(\Phi; A_i)$ with coefficients $m(A_i)$, which leads to the following expression:

$$L(\Phi; m) = \sum_{i=1}^{r} m(A_i) L(\Phi; A_i) = \sum_{y \in \Omega_Y} p_Y(y|\Phi) pl(y). \tag{13}$$

The likelihood function $L(\Phi; m)$ thus only depends on m through its associated contour function pl. Therefore, we will write indifferently $L(\Phi; m)$ or $L(\Phi; pl)$.

The above definitions can be straightforwardly transposed to continuous case. Assume that Y is a continuous random vector with probability density function $p_Y(y|\Phi)$ and let $pl : \Omega_X \to [0, 1]$ be the contour function of a continuous mass function m on Ω_X. The likelihood function given pl can be defined as:

$$L(\Phi; pl) = \int_{\Omega_Y} p_Y(y; \Phi) pl(y) dy, \tag{14}$$

assuming this integral exists and is nonzero.

4.2 The Evidential EM Algorithm

Here we remind how the classical EM algorithm may be extended so as to estimate the parameters of the model when the data at hand are imprecise. An extensive presentation of this approach may be found in [7]. The new method is called Evidential EM (E2M) algorithm, which maximizes the generalized criterion introduced in the previous section.

At iteration l, the E-step of the E2M algorithm consists in computing

$$Q(\Phi; \Phi^l) = E_{\Phi^l} \left[\log L(\Phi; y) | pl(x) \right]. \tag{15}$$

Note that this expectation is now computed with respect to the imprecise sample known through the contour function $pl(x)$. The M-step is unchanged and requires the maximization of $Q(\Phi; \Phi^l)$ with respect to Φ. As in the EM algorithm, the E2M algorithm alternately repeats the E- and M-steps defined above until the relative increase of the observed-data likelihood becomes smaller than a given threshold.

4.3 Sound Source Localization via the E2M Algorithm

In this section, we present the main results which lead to the update equations of the parameter estimates ξ, η, A and σ^2. For this purpose, we give a lemma which will be used later. Due to page limitation, we omit the proof.

Lemma 1. *If X follows M-dimensional complex Gaussian distribution with mean μ_1 and covariance matrix $\sigma_1^2 I_M$, that is $f(x) = \phi(x; \mu_1, \sigma_1^2 I_M)$, and if X is known through the Gaussian contour function $pl(x) = \phi(x; \mu_2, \sigma_2^2 I_M)$, then*

$$f(x)pl(x) = \phi(x; \frac{\sigma_1^2\mu_2 + \sigma_2^2\mu_1}{\sigma_1^2+\sigma_2^2}, \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2} I_M)\phi(\mu_1; \mu_2, (\sigma_1^2+\sigma_2^2)I_M)$$

and

$$f(x|pl(x)) = \phi\left(x; \frac{\sigma_2^2\mu_1 + \sigma_1^2\mu_2}{\sigma_1^2+\sigma_2^2}, \frac{\sigma_1^2\sigma_2^2 I_M}{\sigma_1^2+\sigma_2^2}\right).$$

The log-likelihood of the complete data y is

$$\log L(y; \xi, \eta, A, \sigma) = -2MNP\log\sigma - \frac{N}{\sigma^2}\sum_{p=1}^{P}\sum_{n=1}^{N}\left|y_{np} - G_n(\xi_n, \eta_n)A_n\right|^2.$$

Since the partial knowledge of the actual pressure measured by the m-th microphone is represented by a Gaussian contour function $pl(x_{mp}) = N(x_{mp}; \mu^*_{mp}, \sigma^2_{mp})$, by the second equation of Lemma 1

$$f(x_{mp}|pl(x_{mp})) = \phi\left(x_{mp}; \frac{\sigma^2_{mp}[G(\xi,\eta)A]_m + \sigma^2\mu^*_{mp}}{\sigma^2_{mp}+\sigma^2}, \frac{\sigma^2_{mp}\sigma^2}{\sigma^2_{mp}+\sigma^2}\right),$$

where $[v]_m$ stands for the m-th element of the vector v. Then, by Theorem 1,

$$f(y_{mnp}|x_{mp}) = \phi\left(y_{mnp}; [G_n(\xi_n,\eta_n)A_n]_m + \frac{1}{N}(x_{mp} - [G(\xi,\eta)A]_m), \frac{N-1}{N^2}\sigma^2\right),$$

from which we can deduce

$$\begin{aligned} & f(y_{mnp}|x_{mp})f(x_{mp}|pl(x_{mp})) \\ &= N^2\phi\left(x_{mp}; \left[G(\xi,\eta)A - N(y_{np} - G_n(\xi_n,\eta_n)A_n)\right]_m, (N-1)\sigma^2\right) \\ &\phi\left(x_{mp}; \frac{\sigma^2_{mp}[G(\xi,\eta)A]_m + \sigma^2\mu^*_{mp}}{\sigma^2_{mp}+\sigma^2}, \frac{\sigma^2_{mp}\sigma^2}{\sigma^2_{mp}+\sigma^2}\right). \end{aligned}$$

Then, the conditional probability of y_{mnp} given $pl(x_{mp})$ is

$$f(y_{mnp}|pl(x_{mp})) = \int f(y_{mnp}|x_{mp})f(x_{mp}|pl(x_{mp}))dx_{mp}, \tag{16}$$

and by the first equation in Lemma 1, we can write

$$f(y_{mnp}|pl(x_{mp})) = \phi\left(y_{mnp}; v_{mnp}, \frac{1}{N^2}\frac{N\sigma^2\mu^*_{mp} + (N-1)\sigma^4}{\sigma^2_{mp} + \sigma^2}\right),$$

where $v_{mnp} = [G_n(\xi_n, \eta_n)A_n]_m + \frac{1}{N}\frac{\sigma^2\left(\mu^*_{mp} - [G(\xi,\eta)A]_m\right)}{\sigma^2_{mp}+\sigma^2}$. Then we obtain

$$E(y_{mnp}|pl(x_{mp}); \xi^l, \eta^l, A^l, \sigma^l) = v^l_{mnp},$$

where v^l_{mnp} is obtained by replacing the parameters ξ, η, A, σ by $\xi^l, \eta^l, A^l, \sigma^l$ in v_{mnp}. Eventually, for each n, for obtaining the estimates of ξ_n, η_n, A_n, the E2M algorithm amounts to solve, at iteration l,

$$\begin{aligned} &\arg\max_{\xi_n,\eta_n,A_n} E\left(\log L(y;\xi,\eta,A,\sigma)|pl(x);\xi^l,\eta^l,A^l,\sigma^l\right) \\ &= \arg\min_{\xi_n,\eta_n,A_n} \sum_{p=1}^{P}\sum_{m=1}^{M}\left|v^l_{mnp} - [G_n(\xi_n,\eta_n)A_n]_m\right|^2. \end{aligned} \tag{17}$$

The M-step of the E2M algorithm thus corresponds to the following computations:

1. From (17), we obtain the estimate of A_n at $(l+1)$-th step

$$A^{l+1}_n = \frac{G_n(\xi_n,\eta_n)^H e^l_n}{G_n(\xi_n,\eta_n)^H G_n(\xi_n,\eta_n)}, \tag{18}$$

$n = 1, \ldots, N$, where

$$e^l_n = (e^l_{1n}, e^l_{2n}, \ldots, e^l_{Mn})^T \tag{19}$$

and $e^l_{mn} = \frac{1}{P}\sum_{p=1}^{P} v^l_{mnp}$.

2. Therefore, the update equation for the source location are obtained by

$$\begin{aligned} &\arg\min_{\xi_n,\eta_n} \sum_{m=1}^{M}\left|e^l_{mn} - \frac{[G_n(\xi_n,\eta_n)]_m G_n(\xi_n,\eta_n)^H}{|G_n(\xi_n,\eta_n)|^2}e^l_n\right|^2 \\ &= \arg\min_{\xi_n,\eta_n} (e^l_n)^H \frac{G_n(\xi_n,\eta_n)G_n(\xi_n,\eta_n)^H}{|G_n(\xi_n,\eta_n)|^2}e^l_n. \end{aligned} \tag{20}$$

3. Finally, we compute $E\left(\log L(y;\xi^{l+1},\eta^{l+1},A^{l+1},\sigma)|pl(x);\xi^l,\eta^l,A^l,\sigma^l\right)$ and maximize it with respect to σ, then the estimate of variance σ^2 is given by

$$\begin{aligned} (\sigma^2)^{l+1} = \frac{1}{MP}\sum_{p=1}^{P}\sum_{n=1}^{N}\sum_{m=1}^{M}\Bigg[&\frac{N(\sigma^l)^2\sigma^2_{mp} + (N-1)(\sigma^l)^4}{N^2((\sigma^l)^2 + \sigma^2_{mp})} \\ &+ \left|v^l_{mnp} - [G_n(\xi^{l+1}_n,\eta^{l+1}_n)A^{l+1}_n]_m\right|^2\Bigg]. \end{aligned} \tag{21}$$

Finally, we could summarize the algorithm as the EM algorithm at the end of Section 2.3, but Equations (5-8) are replaced by Equations (18-21). The stopping criterion in this E2M algorithm is the same as in Equation (9), but the observed-data log-likelihood is replaced by $\log L(\Phi; pl)$, which is computed by:

$$\log L(\Phi; pl) = -MP \log \pi - \sum_{m=1}^{M} \sum_{p=1}^{P} \left[\log(\sigma^2 + \sigma_{mp}^2) + \frac{\left| \mu_{mp}^* - [G(\xi, \eta)A]_m \right|^2}{\sigma^2 + \sigma_{mp}^2} \right].$$

5 Experiments

5.1 Data Generation

We consider $M = 11$ microphones situated on the x-axis with locations $\theta_m = 0 : 0.1 : 1$, and $N = 2$ sound sources with actual coordinates $\xi_1 = 0.3$, $\xi_2 = 0.7$, and $\eta_1 = \eta_2 = 0.3$ (the locations of microphones and sound sources are displayed in Figure 1 using pink stars and black crosses respectively.). The strengths of both sources are $A_1 = 0.6$ and $A_2 = 0.8$. The theoretical wavenumber value is $k = 2\pi f/c$, where the sound frequency is $f = 500$Hz and the sound velocity is assumed to be $c = 340m/s$.

The number of snapshots (the amount of pressures measured by each microphone) is set to $P = 100$. Then, we introduce noise in the microphone locations and wavenumber, as follows. For a given value of σ_k, we generate a wavenumber value according to a Gaussian distribution with mean k and standard deviation σ_k. Similarly, we introduce noise in the microphone locations using means $\theta_1, \ldots, \theta_{11}$ and standard deviation σ_θ. Then, we obtain simulated pressures using Equation (1), in which we set $\sigma = 0.05$. The contour functions modeling the uncertainty on the pressures may be obtained using the method detailed in Section 3.3. The level of noise in the data, and consequently the amount of uncertainty, may thus be controlled through the parameters σ_θ and σ_k.

Note that the quality of the estimates obtained via the EM and E2M algorithms depend on the starting values for the parameters. Therefore, for a given dataset, we let both algorithms run using 5 different sets of starting values, retaining the solution with highest log-likelihood. Remark that the variances of the contour functions σ_{mp}^2 used in the E2M algorithm are computed using the parameters ξ_n, η_n and A_n estimated via the EM algorithm. Since the data are randomly generated, the above procedure (from data generation to model estimation) is repeated 30 times, so that mean square errors (MSE) on the parameter values and associated 95% confidence intervals may be computed.

5.2 Results

First, we set $\sigma_k = 0.5$ and $\sigma_\theta = 0.05$, and we estimate the sound source locations (ξ_n, η_n) using EM and E2M. Figure 1 shows the 30 estimation results of the source locations computed by EM (left figure) and E2M (right figure).

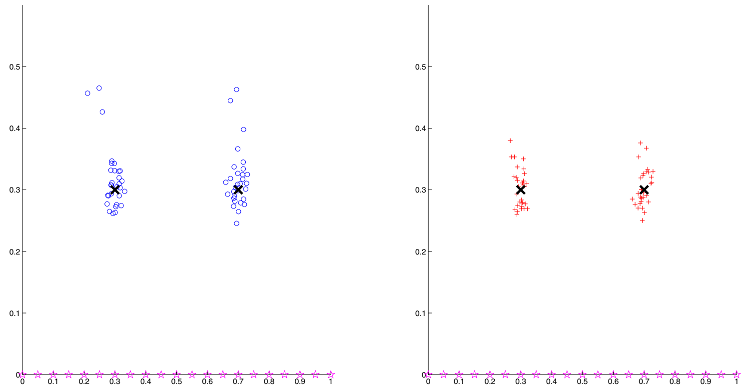

Fig. 1. Locations of the sources estimated using the EM (left) and E2M (right) algorithms, with $\sigma_k = 0.5$ and $\sigma_\theta = 0.05$

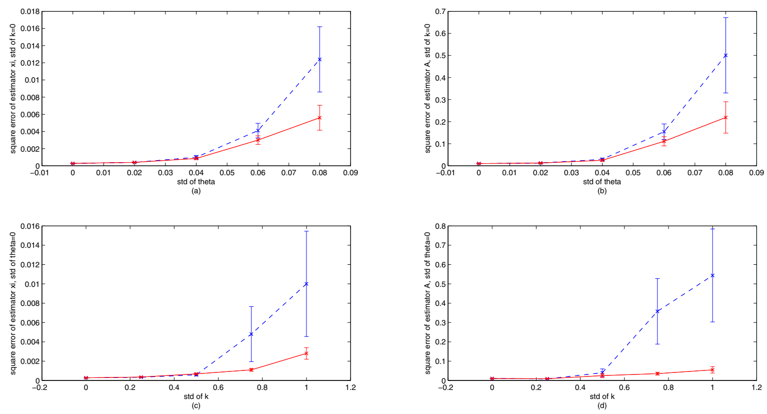

Fig. 2. MSE and 95% confidence intervals for the estimates of sound source locations (ξ, η) (left) and strengths A (right) using EM (blue dash lines) and E2M (red full lines), according to the level of uncertainty in the microphone locations (top) and wavenumber (bottom)

The estimates obtained using E2M clearly exhibit a smaller spread around the actual locations of the sources, which demonstrates its interest in terms of dealing with uncertain data.

To corroborate these observations, we now study the MSE of the parameters (ξ_n, η_n) and A_n estimated via both algorithms. We set $\sigma_k = 0$ and increase σ_θ from 0 to 0.08. The estimated MSE of the sound source locations and strengths are displayed in Figure 2 (top) along with 95% confidence intervals. Without surprise, the accuracy of the results obtained using both algorithms decreases as the amount of noise increases. However, E2M proves to be much more robust to the level of noise than EM: the difference clearly shows the interest of using E2M. More specifically, when σ_θ increases, the MSE of the estimates obtained via EM increase dramatically, while the MSE obtained using E2M stay under an acceptable level. The same phenomenon may be observed when $\sigma_\theta = 0$ and σ_k increases from 0 to 1. The corresponding results are displayed in Figure 2 (bottom). Again, as the level of uncertainty σ_k increases, the MSE of the estimates obtained via EM increases much more than those obtained with E2M, which remains at an acceptable level.

6 Conclusions

In this paper, we addressed the problem of sound source localization from acoustical pressures measured by a set of microphones. The problem may be solved in a statistical setting, by assuming that each pressure measured by a microphone is the sum of contributions of the various sources, pervaded by a Gaussian noise. The EM algorithm may then be used to compute maximum-likelihood estimates of the model.

However, in many applications, some parameters of the model, such as the microphone locations or the wavenumber, may be pervaded with uncertainty. In this work, we show how this uncertainty may be transposed on the measured pressures. We propose to model these uncertainties using belief functions. In this case, the parameters of the model may be estimated using a variant of the EM algorithm, known as the Evidential EM algorithm.

The results obtained on simulated data clearly show the advantage of taking into account the uncertainty on the data, in particular when the degree of noise due to the ill-known parameters (microphone locations and wavenumber) is high. Then, the results obtained using the E2M algorithm are more robust than those obtained using the EM algorithm. The generalization of our method to the general case of spatial sources, as well as its validation on real data, are left for further work.

Acknowledgments. This work has been partially funded by the European Union. Europe is committed in Picardy with the FEDER.

References

1. Bevington, P.R., Keith, D.: Data Reduction and Error Analysis for the Physical Sciences, 3rd edn. McGraw-Hill, New York (2003)
2. Cirpan, H.A., Cekli, E.: Deterministic Maximun Likelihood Approach for Localization of Near-Field Sources. International Journal of Electronics and Communications 56(1), 1–10 (2002)
3. Cirpan, H.A., Cekli, E.: Unconditional Maximum Likelihood Approach for Localization of Near-Field Sources: Algorithm and Performance Analysis. International Journal of Electronics and Communications 57(1), 9–15 (2003)

4. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society, Series B (Methodological) 39(1), 1–38 (1977)
5. Dempster, A.P.: Upper and Lower Probabilities induced by a Multivalued Mapping. Annals of Mathematical Statistics 38, 325–339 (1967)
6. Dempster, A.P.: Upper and Lower Probabilities Generated by a Random Closed Interval. Annals of Mathematical Statistics 39(3), 957–966 (1968)
7. Denoeux, T.: Maximum Likelihood Estimation from Uncertain Data in the Belief Function Framework. IEEE Transactions on Knowledge and Data Engineering 25(1), 119–130 (2013)
8. Denoeux, T.: Maximum Likelihood Estimation from Fuzzy Data Using the EM Algorithm. Fuzzy Sets and Systems 183(1), 72–91 (2011)
9. Feder, M., Weinstein, E.: Parameter Estimation of Superimposed Signals Using EM Algorithm. IEEE Transactions on Acoustics, Speech and Signal Processing 36(4) (2005)
10. Kabaoglu, N., Cirpan, H.A., Cekli, E., Paker, S.: Deterministic Maximum Likelihood Approach for 3-D Near-Field Source Localization. International Journal of Electronics and Communications 57(5), 345–350 (2003)
11. Quost, B., Denœux, T.: Clustering Fuzzy Data Using the Fuzzy EM algorithm. In: Deshpande, A., Hunter, A. (eds.) SUM 2010. LNCS (LNAI), vol. 6379, pp. 333–346. Springer, Heidelberg (2010)
12. Render, R.A., Walker, H.F.: Mixture Densities, Maximum Likelihood and the EM Algorithm. SIAM Review 26(2), 195–239 (1984)
13. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton (1976)
14. Smets, P.: Belief Functions on Real Numbers. International Journal of Approximate Reasoning 40(3), 181–223 (2005)
15. Rhodes, I.B.: A Tutorial Introduction to Estimation and Filtering. IEEE Transaction on Automatic Control AC-16(6), 688–706 (1971)
16. Williams, E.G.: Fourier Acoustic: Sound Radiation and Nearfield Acoustical Holography. Academic Press (1999)
17. Wu, J.C.F.: On the Convergence Properties of the EM Algorithm. Annals of Statistics 11(1), 95–103 (1983)