

## On the use of global optimization methods for acoustic source mapping

Malgoezar, Anwar; Snellen, Mirjam; Merino Martinez, Roberto; Simons, Dick; Sijtsma, P

**DOI**

[10.1121/1.4973915](https://doi.org/10.1121/1.4973915)

**Publication date**

2017

**Document Version**

Final published version

**Published in**

Journal of the Acoustical Society of America

**Citation (APA)**

Malgoezar, A., Snellen, M., Merino Martinez, R., Simons, D., & Sijtsma, P. (2017). On the use of global optimization methods for acoustic source mapping. *Journal of the Acoustical Society of America*, *141*, 453–465. <https://doi.org/10.1121/1.4973915>

**Important note**

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

**Copyright**

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

**Takedown policy**

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

## On the use of global optimization methods for acoustic source mapping

Anwar M. N. Malgoezar, Mirjam Snellen, Roberto Merino-Martinez, and Dick G. SimonsPieter Sijtsma

Citation: *J. Acoust. Soc. Am.* **141**, 453 (2017); doi: 10.1121/1.4973915

View online: <http://dx.doi.org/10.1121/1.4973915>

View Table of Contents: <http://asa.scitation.org/toc/jas/141/1>

Published by the [Acoustical Society of America](#)

---

### Articles you may be interested in

[Sound field reconstruction using compressed modal equivalent point source method](#)

*J. Acoust. Soc. Am.* **141**, (2017); 10.1121/1.4973567

[Adaptive and compressive matched field processing](#)

*J. Acoust. Soc. Am.* **141**, (2017); 10.1121/1.4973528

[Large-region acoustic source mapping using a movable array and sparse covariance fitting](#)

*J. Acoust. Soc. Am.* **141**, (2017); 10.1121/1.4974054

[A learning method for localizing objects in reverberant domains with limited measurements](#)

*J. Acoust. Soc. Am.* **141**, (2017); 10.1121/1.4973807

---

# On the use of global optimization methods for acoustic source mapping

Anwar M. N. Malgoezar,<sup>a)</sup> Mirjam Snellen, Roberto Merino-Martinez, and Dick G. Simons  
*Delft University of Technology, Faculty of Aerospace Engineering, Kluyverweg 1, 2629 HS Delft,  
The Netherlands*

Pieter Sijtsma

*Advanced AeroAcoustics, Prinses Margrietlaan 13, 8091 AV Wezep, The Netherlands*

(Received 3 August 2016; revised 21 December 2016; accepted 22 December 2016; published online 20 January 2017)

Conventional beamforming with a microphone array is a well-established method for localizing and quantifying sound sources. It provides estimates for the source strengths on a predefined grid by determining the agreement between the pressures measured and those modeled for a source located at the grid point under consideration. As such, conventional beamforming can be seen as an exhaustive search for those locations that provide a maximum match between measured and modeled pressures. In this contribution, the authors propose to, instead of the exhaustive search, use an efficient global optimization method to search for the source locations that maximize the agreement between model and measurement. Advantages are two-fold. First, the efficient optimization allows for inclusion of more unknowns, such as the source position in three-dimensional or environmental parameters such as the speed of sound. Second, the model for the received pressure field can be readily adapted to reflect, for example, the presence of more sound sources or environmental parameters that affect the received signals. For the work considered, the global optimization method, Differential Evolution, is selected. Results with simulated and experimental data show that sources can be accurately identified, including the distance from the source to the array.

© 2017 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4973915>]

[JFL]

Pages: 453–465

## I. INTRODUCTION

Beamforming is a widely applied method for imaging sound sources. To perform beamforming, use is made of an array of microphones. Beamforming is based on differences in arrival times (or phase differences) of sound at the different microphones.

When it is assumed that sound sources behave like acoustic monopoles, often a least-squares approach is used for obtaining an estimate for the source strength at a potential source location.<sup>1</sup> Since both source strength and source location are unknown, the general approach is to define a grid of potential source locations and estimate the sound pressure level for each grid point. By depicting these estimates in a so-called source map, an image is established where high levels indicate the presence of a sound source. This approach can be considered as an exhaustive search, estimating the source strength for all grid points and identifying those with high values as source locations.

The above described approach is subject to a number of drawbacks. First, it restricts the optimization problem to a limited number of unknowns due to computational constraints induced by the approach of an exhaustive search. Typically, beamforming is applied in searches for the source in two dimensions, often assuming a scan plane parallel to the array at a known distance. Second, the assumption of a single monopole limits the suitability of beamforming for situations with multiple sound sources.

In this contribution, we propose to consider the search for source locations and source pressure amplitudes as a global optimization problem. In this way, the procedure of estimating the source strength for each grid point is abandoned and focus is only put on identifying the actual source locations and source strengths.

The presence of sidelobes, indicating relatively high beamforming output levels without a source being present, will, however, hamper the optimization as they act as local optima against which the global optimum needs to be found. In literature, a number of mathematical methods are presented that allow for optimization problems with many unknowns and with the ability to escape local optima. These methods are generally denoted as global optimization methods. Well-known examples of these types of methods are genetic algorithms,<sup>2–4</sup> simulated annealing,<sup>5,6</sup> and ant colony optimization.<sup>7,8</sup> In contrast to local search methods, e.g., gradient methods, these global optimization methods have the capability to escape local optima. This ability is essential for the application considered with sidelobes present. In this paper, we use the method of Differential Evolution (DE)<sup>9,10</sup> for obtaining the source positions and source strengths. This method is a variant of the genetic algorithm. This type of global optimization method mimics the natural evolution of species. They use populations of solutions, where promising solutions are given a higher probability to reproduce than bad solutions.

The approach taken in this work is in line with that of, for example, the “Deconvolution Approach for the Mapping of Acoustic Sources” (DAMAS)<sup>11,12</sup> technique and the

<sup>a)</sup>Electronic mail: a.m.n.malgoezar@tudelft.nl

approach of Cross-Spectral Matrix (CSM) inversion,<sup>13,14</sup> where the aim is to obtain maximum information about the acoustic sources by maximizing the agreement between the measured and modeled microphone pressure field. In DAMAS, the delay-and-sum beamformer result is used as a starting point. The idea is that the delay-and-sum beamforming output can be considered as the summation of point-spread functions (PSFs) of all sources present, weighted by the source strengths, where the PSF is defined as the beamformer response to a point source with unit strength at a given position of a grid. The deconvolution is based on solving the resulting inverse problem for retrieving the source distribution that resulted in the observed delay-and-sum beamformer output. In general, unless the scan grid is very small, the problem is ill-conditioned, preventing a direct inversion of the matrix, and an iterative scheme is used to solve the inverse model. In general, a significant number of iterations can be required. Depending on the chosen grid size, this can become time consuming. In Refs. 13 and 14 alternative approaches were developed to solve the inverse problem. Two of these are known as sparsity constrained Deconvolution Approach for the Mapping of Acoustic Sources (SC-DAMAS) and a sparsity preserving Covariance Matrix Fitting (CMF) approach. In both techniques, sparsity is maximized by searching for the minimum amount of sources for which the deconvolution holds. The assumption is that the amount of sources is considerably smaller than the amount of scan points. The SC-DAMAS method employs the output of the delay-and-sum beamformer with which agreement has to be maximized, whereas the CMF approach directly optimizes the agreement with the measured covariance matrix. The latter has the advantage to skip the delay-and-sum beamforming and directly estimate the source locations and strengths from the covariance matrix. Similarly, in the field of ocean acoustics the sparse problem is handled by using compressive sensing for beamforming to estimate the direction-of-arrival (DOA). This is done by using convex optimization and promoting sparsity.<sup>15,16</sup> To prevent mismatch of the DOAs and not be restricted by a predefined grid, the compressive beamforming method was reformulated to be continuous and grid-free.<sup>17</sup>

Likewise, the work presented in this contribution is grid-free and is also based on a direct comparison between modeled and measured pressure fields, similar to the CMF approach. The locations of the sources are sought for by using a global optimization method. In this way, estimates for source positions and source strengths are obtained as a solution of the optimization and do not need to be obtained from a delay-and-sum beamformer result. The efficiency of the global optimization compared to an exhaustive search allows for an increased number of unknowns. In this contribution, for example, localization of the source in three dimensions is investigated, but also its potential in estimating the sound speed, i.e., a property of the propagation medium. Including environmental parameters can be of interest also for complex wind tunnel measurements in aeroacoustic research. The proposed approach is comparable to the approach of matched field processing and matched field

inversion that is well-known in the field of underwater acoustics.<sup>3,4,6,18–20</sup>

In Sec. II, the theory for the representation of the signal at the acoustic array is presented, together with the so-called energy functions considered. The method of DE is introduced and explained in Sec. III. The energy function is used as a measure for the difference between measured and modeled pressure fields and, as such, needs to be minimal. In Sec. IV, the three test cases, consisting both of synthetic and measured data, considered in the current work to apply the proposed approach to are presented. Global optimization methods typically require tuning of the algorithm to ensure good performance, i.e., having a high probability to find the global optimum at a limited number of forward calculations. Section V discusses how this tuning was done for this work. Section VI presents the results of applying the proposed inversion strategy to the three test cases. Section VII contains the conclusions.

## II. THE ENERGY FUNCTIONS CONSIDERED

Consider a wave field generated by  $L$  acoustic monopoles where each monopole is located at  $\mathbf{x}_{S,l}$  and  $l = 1, \dots, L$ . Let the position of the microphone  $m$  be given as  $\mathbf{x}_m$ , where  $m = 1, \dots, M$  and  $M$  is the total number of microphones in the array. The model for the  $M \times 1$  array output vector  $\mathbf{y}$  for a radial frequency  $\omega$  is

$$\mathbf{y}(\omega) = \sum_{l=1}^L \mathbf{a}_l(\mathbf{x}_{S,l}, \omega) s_l(\omega), \quad (1)$$

where  $\mathbf{a}_l = [a_{l,1}, \dots, a_{l,M}]^T$  is the steering vector,  $s_l$  is the acoustic waveform of source  $l$ , and  $[\cdot]^T$  denotes the transpose of the vector. The element of  $\mathbf{a}_l$  for microphone  $m$  is given by<sup>1</sup>

$$a_{l,m} = \frac{1}{r_{l,m}} e^{-j\omega r_{l,m}/c}, \quad (2)$$

with  $c$  the speed of sound and  $r_{l,m} = |\mathbf{x}_{S,l} - \mathbf{x}_m|$  the distance between the source and microphone. Equation (1) can also be written in matrix-vector notation

$$\mathbf{y}(\omega) = \mathbf{A}(\omega)\mathbf{s}(\omega), \quad (3)$$

with  $\mathbf{A}(\omega) = [\mathbf{a}_1(\mathbf{x}_{S,1}, \omega), \dots, \mathbf{a}_L(\mathbf{x}_{S,L}, \omega)]$  the  $M \times L$  steering matrix and  $\mathbf{s}(\omega) = [s_1(\omega), \dots, s_L(\omega)]^T$  the signal waveforms. The model for the CSM of the received signals at the microphones is then given as an ensemble-average of data blocks

$$\mathbf{C}_{\text{model}}(\omega) = \text{IE} [\mathbf{y}(\omega)\mathbf{y}^H(\omega)], \quad (4)$$

where  $\text{IE}[\cdot]$  denotes the expectation operator,  $(\cdot)^H$  denotes the complex conjugate transpose of the argument. Using both Eqs. (3) and (4), and assuming that the propagation conditions are constant for the duration of a given measurement,  $\mathbf{C}_{\text{model}}$  can be written as

$$\mathbf{C}_{\text{model}} = \mathbf{A}\mathbf{P}\mathbf{A}^H, \quad (5)$$

where

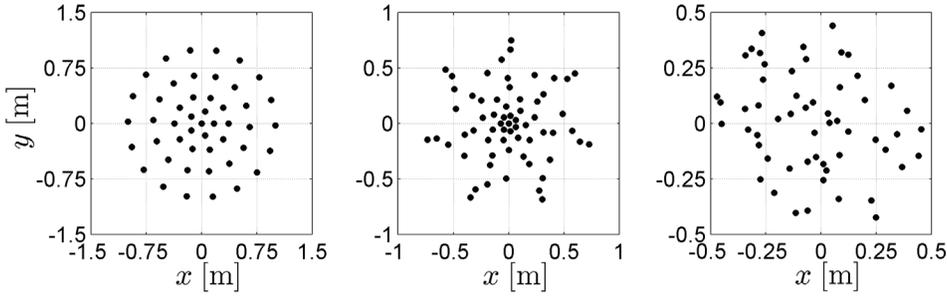


FIG. 1. Array geometry for test case 1 at the left, test case 2 in the middle, and test case 3 at the right.

$$\mathbf{P} = \mathbb{E}[\mathbf{s}(\omega)\mathbf{s}^H(\omega)]. \quad (6)$$

For uncorrelated sources, the ensemble-averaged  $\mathbf{P}$  is a diagonal matrix, as the cross terms of the sources average to zero. Each element of the diagonal matrix presents the power of a source.

The approach taken in delay-and-sum beamforming makes use of Eq. (1) with  $L$  equal to 1, i.e., considering a single source only. Each grid point is treated as a potential source location and the source strength at the grid point  $\mathbf{x}'$  is estimated as

$$\tilde{y}(\mathbf{x}', \omega) = \frac{1}{\|\mathbf{a}\|^4} \mathbf{a}^H(\mathbf{x}', \omega) \mathbf{C}_{\text{meas}}(\omega) \mathbf{a}(\mathbf{x}', \omega), \quad (7)$$

where  $\mathbf{C}_{\text{meas}}$  is the measured covariance matrix and  $\tilde{y}$  is the beamformer output, i.e., the estimate for the source strength at grid point  $\mathbf{x}'$ . Equation (7) is known as conventional beamforming.

For conventional beamforming, with a scan plane parallel to the array the resolution of the source plot is limited by the width of the main lobe. The spatial resolution of a planar microphone array can be approximated by<sup>21,22</sup>

$$\Delta \ell = 1.22 \frac{c z_{\text{array}}}{D f}, \quad (8)$$

with  $z_{\text{array}}$  the average distance between the sources and the array plane,  $D$  the diameter of the array, and  $f = \omega/2\pi$  the frequency.

In the present study a different approach is followed to locate and quantify acoustic sources. We start from Eq. (3) and alternatively Eq. (5) as the forward model. A predefined scan-grid is not used and the amount of sources  $L$  is considered to be known beforehand. An objective function, sometimes denoted as energy function, is defined such that it provides a measure for the difference between the measured CSM and that predicted from Eqs. (3) or (5), given a set of values for the unknown parameters. For this research, both synthetic and measured covariance matrix are data considered for  $\mathbf{C}_{\text{meas}}$ . The synthetic data were obtained from benchmark cases that were generated in the framework of the Benchmarking Array Analysis Methods workshop in Dallas, 2015.<sup>23</sup>

A well-known energy function is the Bartlett processor<sup>15–17,24,25</sup> given by

$$E_{\text{Bartlett}}(\mathbf{g}) = \frac{\mathbf{y}(\mathbf{g}, \omega)^H \mathbf{C}_{\text{meas}}(\omega) \mathbf{y}(\mathbf{g}, \omega)}{\|\mathbf{y}(\mathbf{g}, \omega)\|^2 \text{tr}(\mathbf{C}_{\text{meas}}(\omega))}, \quad (9)$$

with  $\mathbf{C}_{\text{meas}}(\omega)$  the measured CSM at frequency  $\omega$ , and  $\mathbf{y}(\mathbf{g}, \omega)$  the prediction for the pressures at the microphones using Eqs. (1) and (2). Vector  $\mathbf{g}$  contains the trial values for the unknown parameters. For example, in the case of one source it could have the form of  $\mathbf{g} = \mathbf{g}(\mathbf{x}_{s,1}, s_1)$  which would be four parameters considering only the spatial position (three coordinates) and amplitude of the source. The term  $\text{tr}(\cdot)$  denotes the trace of a matrix. A drawback of this energy function is that the source amplitude,  $s_i$ , information does not affect its value. Therefore, it will not be possible to estimate the source amplitude when using this function.

An alternative energy function, which includes the estimation of the source amplitude, is defined as follows:<sup>26</sup>

$$E_{\text{CSM}}(\mathbf{g}) = \sum \left\{ [\text{Re}(\mathbf{C}_{\text{meas}}) - \text{Re}(\mathbf{C}_{\text{model},\mathbf{g}})]^2 + [\text{Im}(\mathbf{C}_{\text{meas}}) - \text{Im}(\mathbf{C}_{\text{model},\mathbf{g}})]^2 \right\}, \quad (10)$$

where  $\mathbf{C}_{\text{model},\mathbf{g}}$  is the modeled covariance matrix corresponding to parameter vector  $\mathbf{g}$ , calculated using Eqs. (1)–(5). The summation is done over all  $M \times M$  elements of the matrices containing the differences between  $\mathbf{C}_{\text{meas}}$  and  $\mathbf{C}_{\text{model},\mathbf{g}}$ . The covariance matrices are defined for a specific frequency  $\omega$ .

### III. THE METHOD OF DE

DE is a method that optimizes a problem by iteratively trying to improve candidate solutions with regard to a given measure of quality.<sup>9,10</sup> The subsequent iterations are denoted as generations. DE makes use of a population of candidate solutions per generation and creates new candidate solutions by combining existing ones, and then keeping improved candidate solutions. For creating the new candidate solutions for the next generation, promising solutions of the current population are selected. Still, to allow for escaping local optima, also less good solutions have a probability of being selected

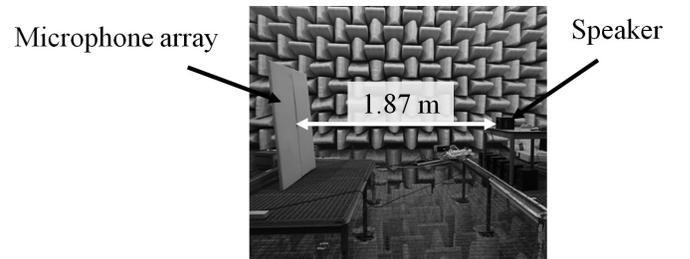


FIG. 2. Experimental setup in the anechoic chamber with the microphone array in the left and the speaker in the right.

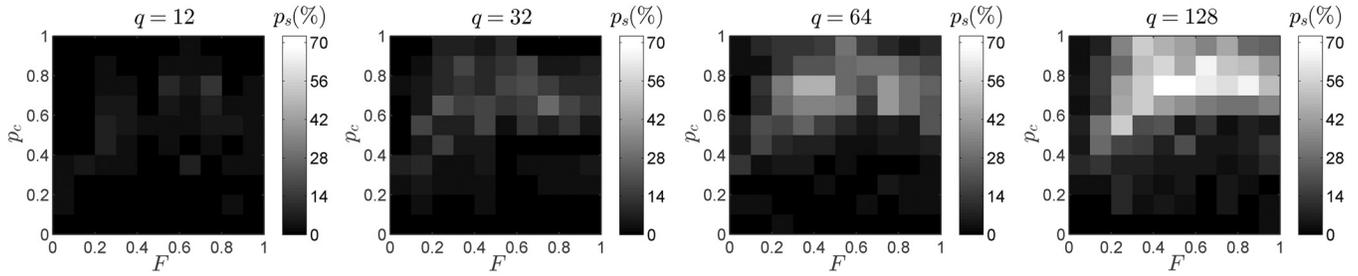


FIG. 3.  $p_s$  as a function of  $p_c$  and  $F$  for four different population sizes  $q = 12, 32, 64,$  and  $128$  with  $N_G = 600$  using the highest frequency,  $f = 6000$  Hz, for the Bartlett energy function.

for creating new candidate solutions. This probability decreases for subsequent generations.

DE can be classified as a metaheuristic method making few or no assumptions about the problem being optimized and can search very large spaces of candidate solutions. DE is used for multidimensional real-valued functions and does not require the user to calculate the gradient of the problem being optimized.

DE starts with an initial population of randomly chosen parameter value combinations. The population consists of  $q$  members (thus,  $q$  vectors  $\mathbf{g}$  per generation), each containing trial values for the unknown parameters. At each generation, a partner population,  $\mathbf{h}_{k,u}$ , is created from the population members  $\mathbf{g}_{k,u}$  as

$$\mathbf{h}_{k,u} = \mathbf{g}_{k,r_1} + F(\mathbf{g}_{k,r_2} - \mathbf{g}_{k,r_3}), \quad (11)$$

with  $u, r_1, r_2, r_3 \in \{1, 2, \dots, q\}$ , integer and mutually exclusive,  $F$  a scalar multiplication factor between 0 and 1, and  $k$  indicates the current generation. The values for  $r_1, r_2, r_3$  are chosen randomly. A higher value of  $F$  indicates an increased difference between original parameter values  $\mathbf{g}_{k,u}$  and those contained in the partner population  $\mathbf{h}_{k,u}$ .

The next step is to calculate its descendant  $\mathbf{d}_{k,u}$  by applying crossover to  $\mathbf{g}_{k,u}$  and  $\mathbf{h}_{k,u}$  with a probability  $p_c$ . For each parameter  $v$  of  $\mathbf{d}_{k,u}$  we get

$$d_{k,u,v} = \begin{cases} g_{k,u,v} & \text{if } r \geq p_c \\ h_{k,u,v} & \text{if } r < p_c, \end{cases} \quad (12)$$

with  $r$  a realization from the uniform distribution, with values between 0 and 1. Setting the value of  $p_c$  high means that more values are replaced by the partner population, while a low value of  $p_c$  results in generations that differ only slightly regardless of the value of  $F$ .

To create the new generation  $k + 1$  from the previous generation  $k$ , the member  $\mathbf{g}_{k,u}$  is replaced by  $\mathbf{d}_{k,u}$  only if it yields a smaller value for the energy function  $E$  as

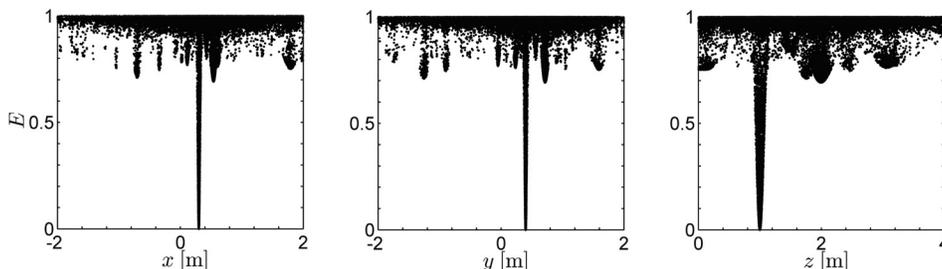


FIG. 4. Energy landscape of the Bartlett energy function as function of the spatial coordinates  $(x, y, z)$  for the optimal setting of  $F = 0.45, p_c = 0.75$  at the population size of  $q = 128$ .

$$\mathbf{g}_{k+1,u} = \begin{cases} \mathbf{d}_{k,u} & \text{if } E(\mathbf{d}_{k,u}) < E(\mathbf{g}_{k,u}) \\ \mathbf{g}_{k,u} & \text{if } E(\mathbf{d}_{k,u}) \geq E(\mathbf{g}_{k,u}). \end{cases} \quad (13)$$

Doing this for all members  $u$  in the population, we obtain the next generation  $k + 1$ . This process is repeated for  $N_G$  generations. For decreasing energy values a member would converge to the correct parameter values (in this problem the positions and strengths of all the sources).

The performance of global optimization methods, i.e., their success in localizing the global optimum in an efficient way, is dependent on a number of so-called setting parameters. For DE these are

- Population size  $q$ ,
- Multiplication factor  $F$ ,
- Crossover probability  $p_c$ ,
- Number of generations  $N_G$ .

These settings must be set beforehand to suitable values, and can be problem specific, to maximize the probability of localizing the global optimum. In this work, the best values for the parameters are determined first, see Sec. V.

## IV. THE TEST CASES CONSIDERED

### A. Test case 1: A single monopole sound source

In this test case, a single monopole sound source is considered, located at  $\mathbf{x}_S = (0.3 \text{ m}, 0.4 \text{ m}, 1.0 \text{ m})$  with source amplitude of 1 Pa. The array consists of 48 microphones. Figure 1 shows the array geometry. The data provided are simulated and consist of the CSM of the microphone measurements at the frequencies 500 to 6000 Hz in 500 Hz steps. In Sec. VI, conventional beamforming and inversion will be applied for each frequency.

### B. Test case 2: Four monopole sound sources

In this case, four uncorrelated monopole sources are considered at corners of a  $0.2 \text{ m} \times 0.2 \text{ m}$  square emitting

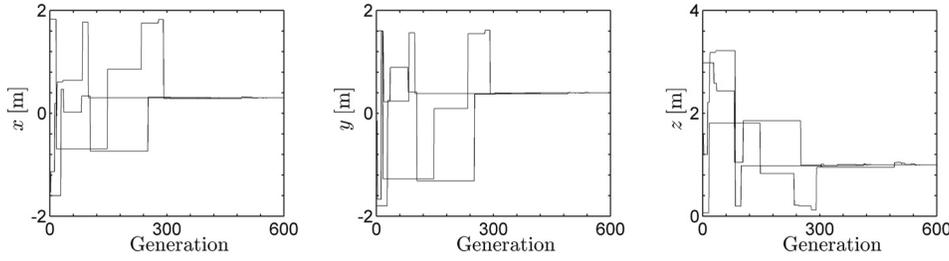


FIG. 5. Convergence of the spatial coordinates ( $x, y, z$ ) belonging to three runs for the optimal setting of  $F=0.45$ ,  $p_C = 0.75$  at the population size of  $q = 128$ .

white noise. The distance between the center of the source plane and the center of the 64 microphone array is 0.75 m. The source locations are denoted as

$$\begin{aligned} \mathbf{x}_{S,1} &= (0.1, 0.1, 0.75), \\ \mathbf{x}_{S,2} &= (0.1, -0.1, 0.75), \\ \mathbf{x}_{S,3} &= (-0.1, 0.1, 0.75), \\ \mathbf{x}_{S,4} &= (-0.1, -0.1, 0.75), \end{aligned}$$

all with equal and known power.<sup>23</sup>

In this case, the array consists of a seven arm logarithmic spiral arrangement with an aperture of 1.5 m. All the sources have the same power and are simulated as white noise in the time domain. The CSM contains values of 513 frequencies in 50 Hz steps. In Sec. VI, nine frequencies will be selected for beamforming and inversion.

### C. Test case 3: One speaker in an anechoic chamber

In addition to the test cases with simulated data, a simple experiment was performed in the anechoic chamber at the Faculty of Applied Sciences of Delft University of Technology. The walls, ceiling, and floor of this room are covered in wedges made from glass wool in order to prevent sound reflections, leaving a space of 8 m  $\times$  8 m  $\times$  8 m inside.

A 56 microphone array with a random distribution and an aperture of approximately 1 m was employed. The acoustic source for this experiment was a small speaker located at a distance of 1.87 m from the array plane and aligned with the array center, as shown in Fig. 2. The array plane formed an angle of 4° with the vertical, which was accounted for in the microphone positions. Hence, the expected sound source position with respect to the array reference system is  $\mathbf{x}_S = (0 \text{ m}, 0 \text{ m}, 1.87 \text{ m})$ .

The overall A-weighted sound pressure level inside the anechoic chamber with the assembled experimental setup, measured using a calibrated Brüel & Kjær 2231 modular precision sound level meter [Brüel & Kjær (Sound and Vibration Measurement A/S), Nærum, Denmark], was found to be lower than 20 dBA.

For this example, the speaker was emitting sound at a single frequency of 5000 Hz. The sampling frequency used for the microphone array was 50 kHz and the recording time was 60 s. In order to obtain the time-averaged CSM, the acoustic data were separated in 49 time blocks with a 50% data overlap.

The sound pressure level at the array center microphone was 80.25 dB.

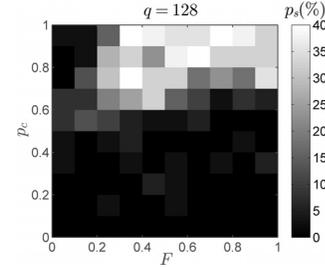


FIG. 6.  $p_s$  as a function of  $p_C$  and  $F$  for  $q = 128$  and  $N_G = 600$  using the highest frequency,  $f = 6000$  Hz, for the CSM energy function. Note the different color bar scale.

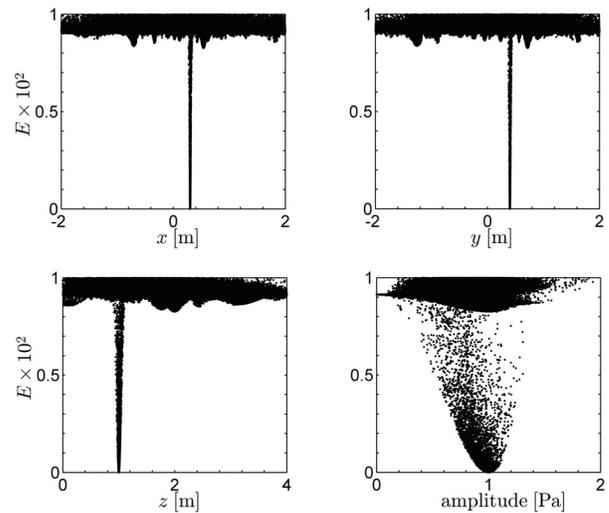


FIG. 7. Energy landscape of the CSM energy function as a function of the spatial coordinates ( $x, y, z$ ) and the source amplitude for the optimal setting of  $F = 0.35$ ,  $p_C = 0.75$  at the population size of  $q = 128$ .

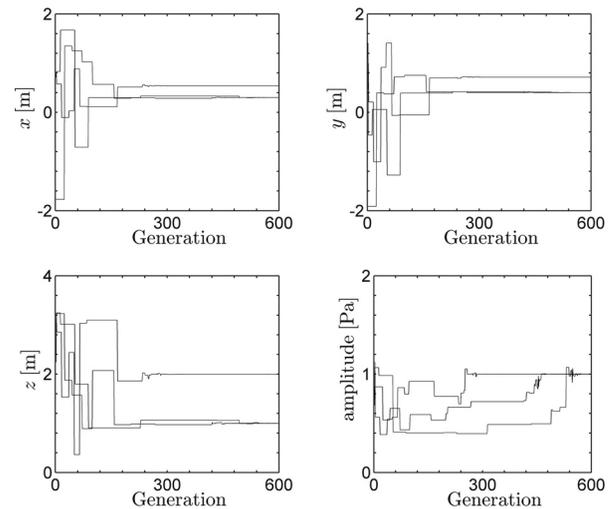


FIG. 8. Convergence of the spatial coordinates ( $x, y, z$ ) and source amplitude belonging to three runs for the optimal setting of  $F = 0.35$ ,  $p_C = 0.75$  at the population size of  $q = 128$ .

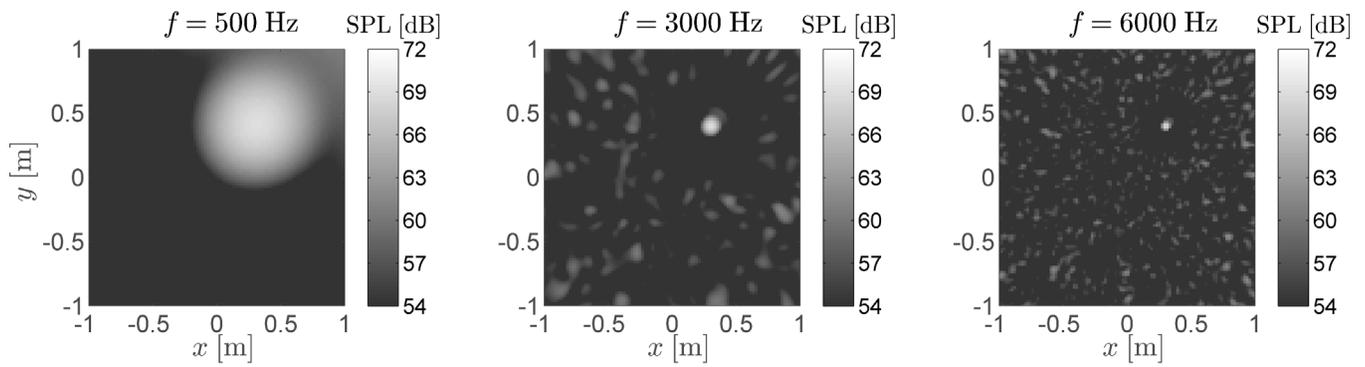


FIG. 9. Beamforming of the monopole source for 500, 3000, and 6000 Hz.

## V. FINDING THE OPTIMAL SETTINGS FOR THE DE ALGORITHM

For DE to find good solutions, its setting parameters need to be selected appropriately.<sup>4</sup> For this test the number of generations was set to  $N_G = 600$ . Data from test case 1 was used at the maximum frequency of 6000 Hz which will exhibit more sidelobes and of higher value than the lower frequencies, and ensures the DE settings to be appropriate also for the lower frequencies. For each combination of DE

setting parameters, 50 independent runs were performed. The fraction of successful runs out of these 50 is denoted as  $p_s$ , and serves as an estimate for the probability of success.

### A. Bartlett energy function

For the Bartlett energy function, a run is considered successful if any of the elements of the final population has a value for the objective function lower than 0.1. Various population sizes were considered. Figure 3 shows the percentage

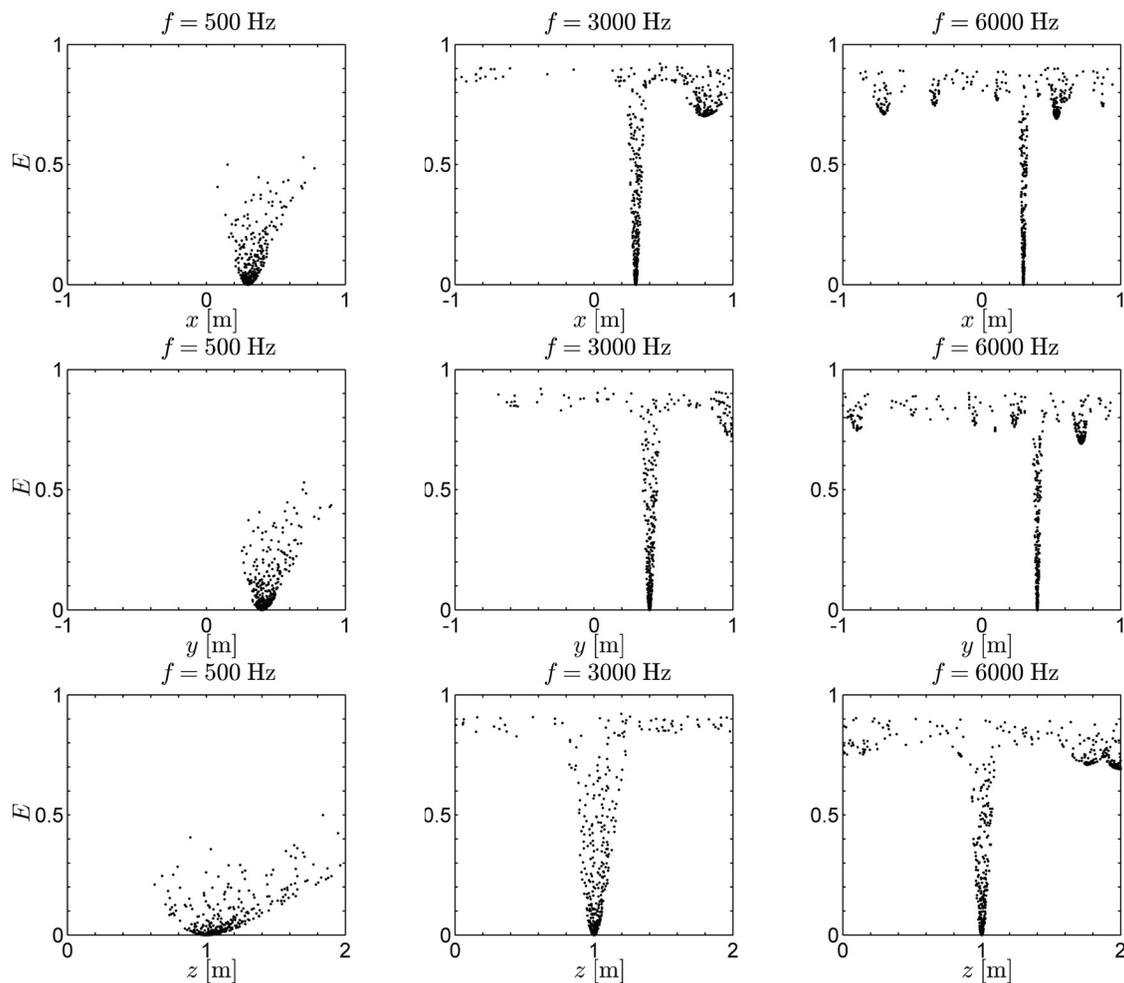


FIG. 10. Inversion using Bartlett as the energy function for 500, 3000, and 6000 Hz.

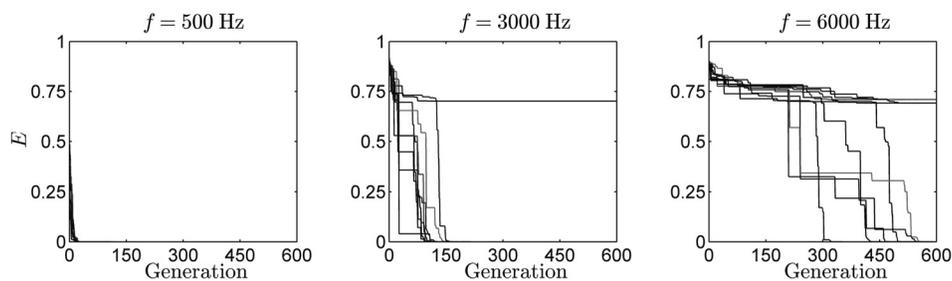


FIG. 11. Energy as a function of generation for three different frequencies and ten independent runs.

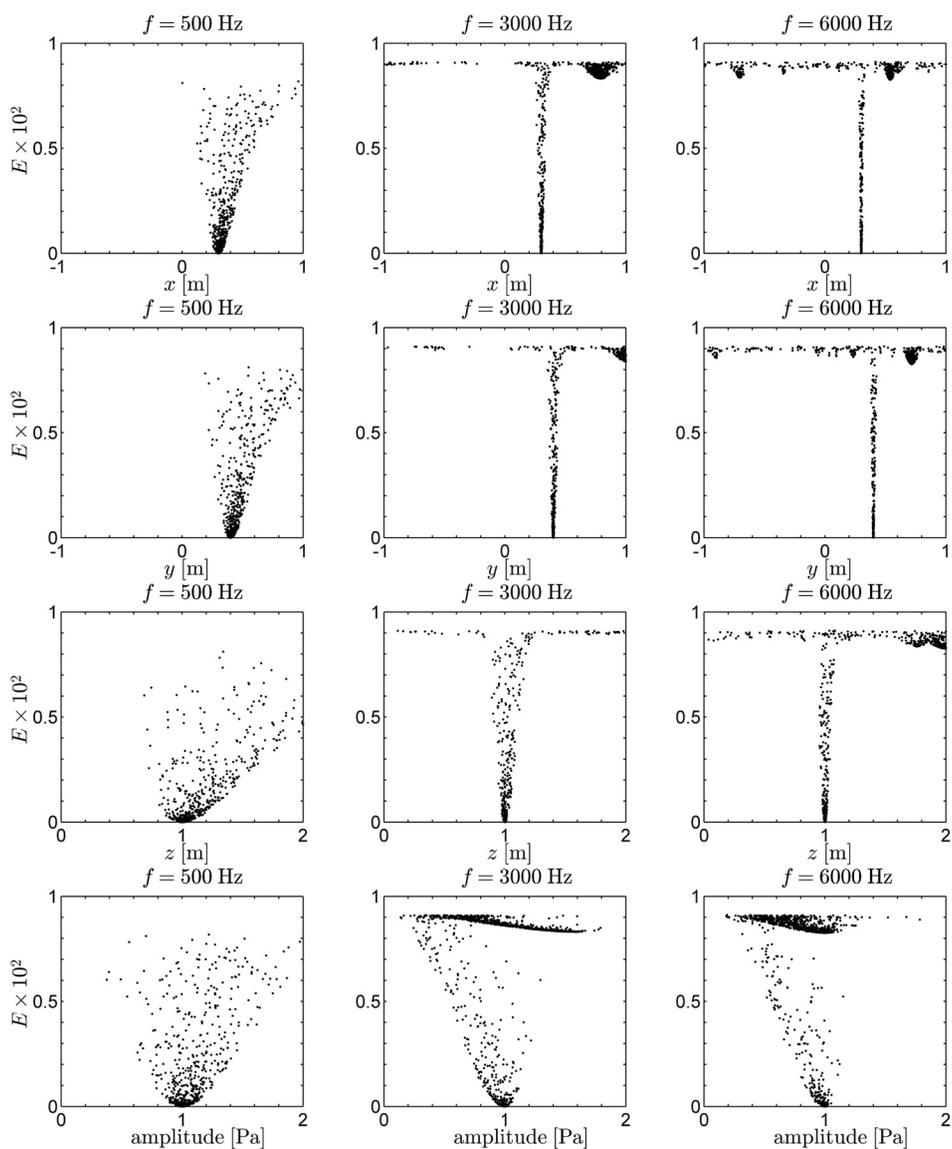


FIG. 12. Inversion using CSM as the energy function for 500, 3000, and 6000 Hz including the amplitude.

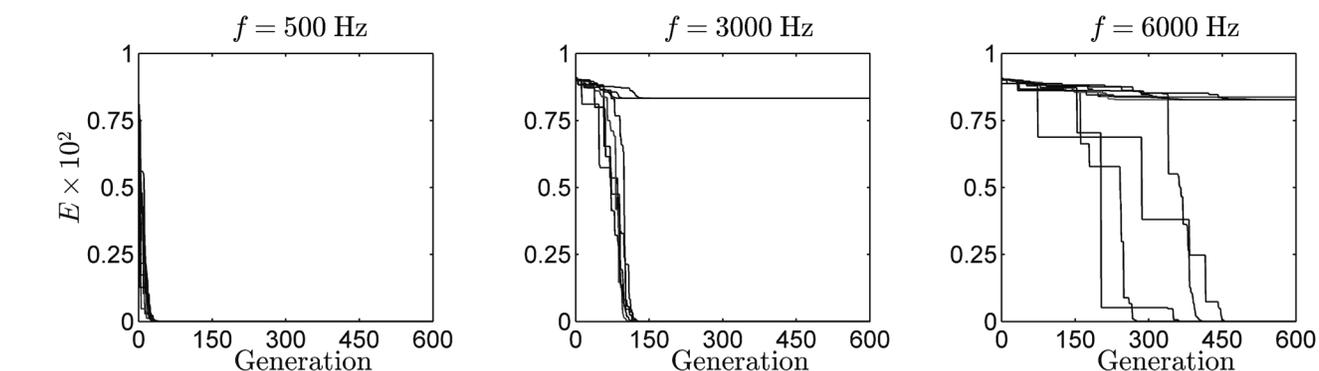


FIG. 13. Energy ( $E$ ) as a function of generation for three different frequencies and ten independent runs using CSM.

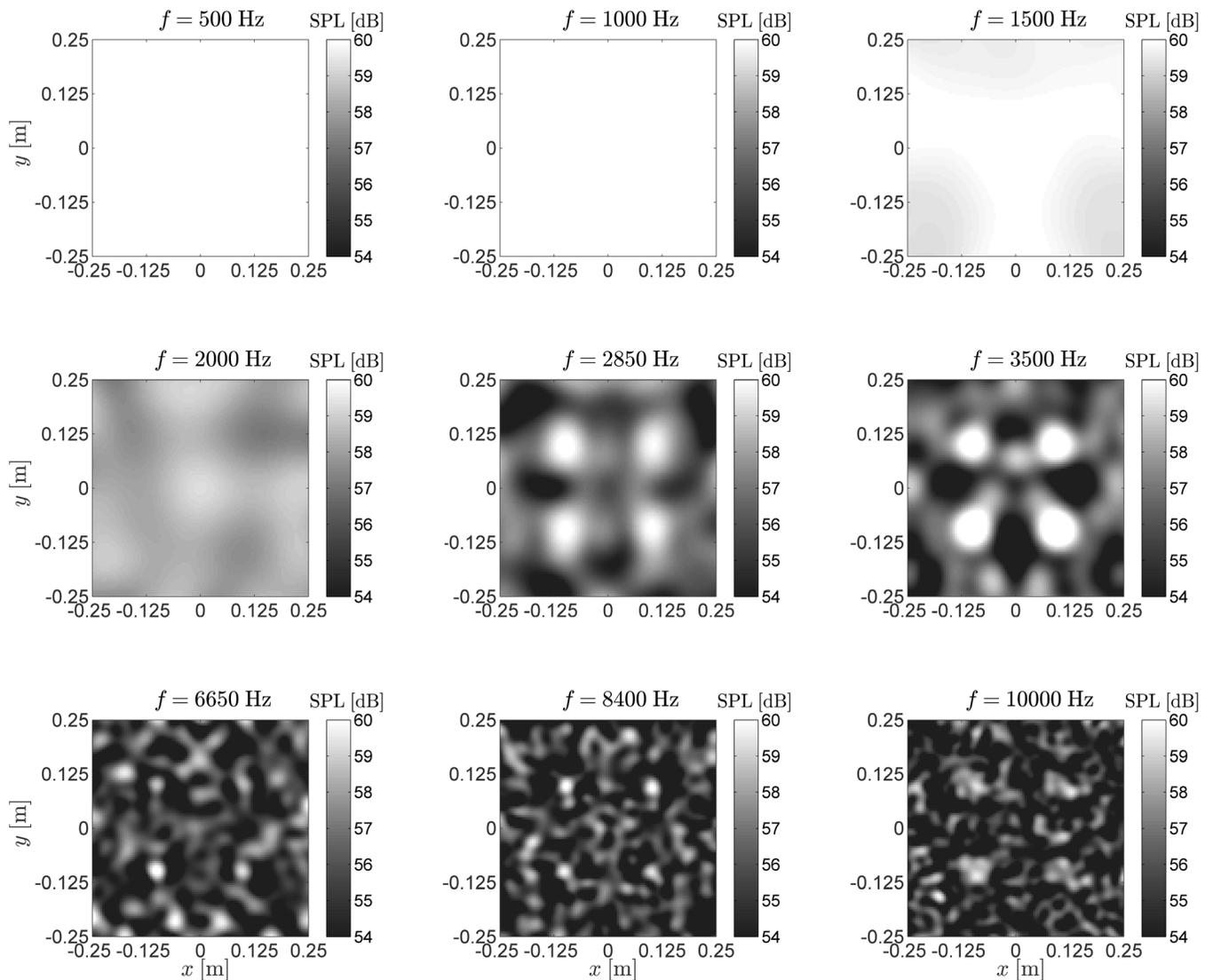


FIG. 14. Conventional beamforming source plots of four monopole sources for various frequencies. The corresponding resolutions according to Eq. (8) are 0.47, 0.21, 0.14, 0.11, 0.074, 0.060, 0.032, 0.025, and 0.021 m for the increasing frequencies.

of converged runs as a function of both  $p_C$  and  $F$  for  $q = 12, 32, 64,$  and  $128$ . In order to find suitable values, the DE performance is assessed for values between 0 and 1 with steps of 0.1 for  $p_C$  and  $F$ .

From Fig. 3 it can be seen that  $p_s$  has the largest value at the population size of  $q = 128$  for  $F = 0.45$  and  $p_C = 0.75$  with a value of around 70%. In Fig. 4, the energy landscape can be seen as a function of the spatial coordinates of the source for all the runs. Regions of local optima can be seen especially for the  $z$ -coordinate. This explains the success-rate of 70%. Because the microphone array is two-dimensional, the sensitivity in the direction pointing away from the array, i.e., the  $z$ -direction, is worse than that in the lateral directions.<sup>1</sup> Additional runs or generations could prevent solutions from being stuck in local optima. For the inversion tests,  $p_s = 70\%$  will be considered adequate since multiple independent inversion runs are carried out for each test case.

Figure 5 shows the convergence of the spatial position of the source for three runs. It can be seen that the correct

source position (0.3, 0.4, 1.0) m is found well within the 600 generations.

## B. CSM energy function

When using the CSM energy function given in Eq. (10), to also determine the source strength,  $s_l$ , a similar procedure is followed as in Sec. V A. For the CSM energy function, a run is considered successful if any of the elements of the final population has a value for the objective function lower than  $10^{-3}$ , due to the small values, in general, for this energy function. The result for test case 1 for a frequency of 6000 Hz is given in Fig. 6. A slight shift of optimal values can be seen to  $F = 0.35$  and  $p_C = 0.75$ , and a significant drop to 40% in  $p_s$ . The energy landscape for this parameter setting is given in Fig. 7. Figure 8 illustrates the convergence behaviour, showing one run to have converged to a local optimum, with values for the  $x, y,$  and  $z$  positions deviating from the true positions.

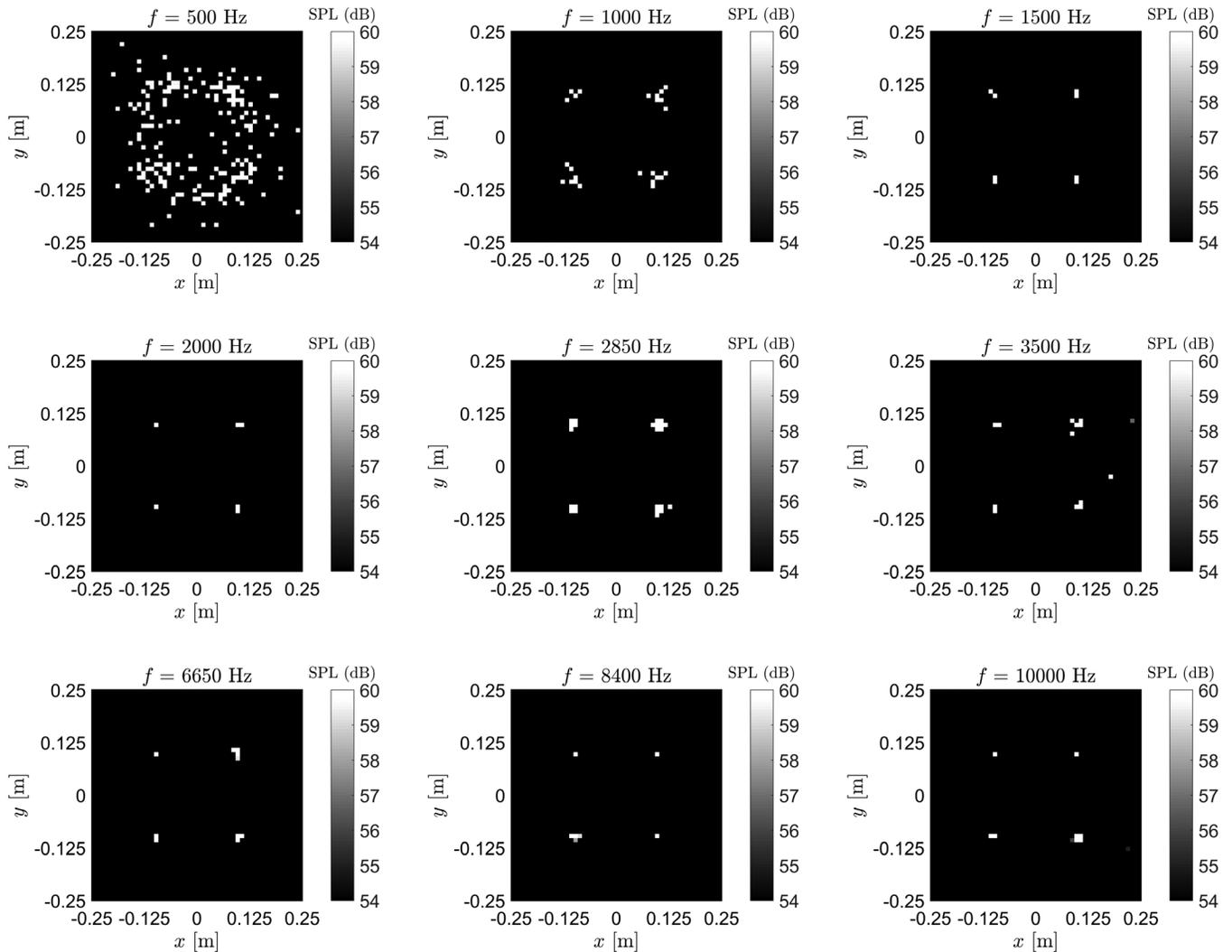


FIG. 15. Inversion using the CSM energy function for four monopoles for various frequencies. Results are shown for all 50 runs and are the solutions with the lowest value for the energy per run.

## VI. RESULTS

### A. Test case 1: Single monopole

As a reference, we start with applying conventional beamforming, for which we define a scan grid at the source location parallel to the array. Beamforming is performed for 12 different frequencies from 500 to 6000 Hz in steps of 500 Hz. Figure 9 shows the results for 500, 3000, and 6000 Hz. As expected, it can be seen that the source is well localized with the resolution improving for higher frequencies. The appearance of sidelobes also increases with higher frequencies.

Figure 10 shows the results for the three frequencies by using the proposed inversion method, employing the Bartlett energy function. To obtain this result, the settings of DE were set to  $q = 128$ ,  $N_G = 600$ ,  $p_C = 0.75$ , and  $F = 0.45$ . The number of independent runs was selected to be 50. For all frequencies, it can be seen that the source position is retrieved correctly, since the values for  $x$ ,  $y$ , and  $z$  that correspond to the lowest energy values are in agreement with the true source position. To show the rate of convergence, Fig. 11 is presented, where the energy is given as a function of the number of generations for the frequencies 500, 3000, and 6000 Hz. For 500 and 3000 Hz, convergence to the

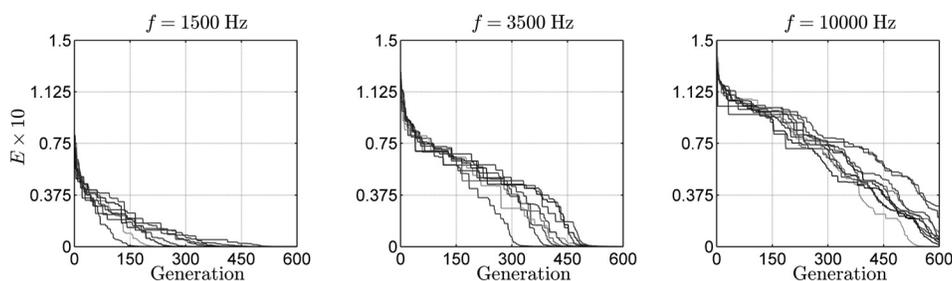


FIG. 16. Energy as a function of generation for a frequency of 1500, 3500, and 10000 Hz (10 independent runs).

correct position is achieved well within the 600 generations. For 6000 Hz, three out of ten runs are seen not to have reached zero energy value at 600 generations.

The same procedure is repeated using the CSM energy function given in Eq. (10) with  $q = 128$ ,  $N_G = 600$ ,  $p_C = 0.75$ , and  $F = 0.35$ . In Fig. 12 the source can be seen to be well localized, with minimum energy corresponding to the correct position. For 3000 and 6000 Hz, more runs are seen to be stuck in a local minimum. This can also be seen in Fig. 13, where at the last generation some runs have not converged to the minimum energy possible yet. The value for the amplitude is obtained correctly at 1 Pa. From Fig. 6 it is expected that the rate of success is lower than that of the Bartlett energy function. This is confirmed by Fig. 13, where fewer runs reach zero energy for higher frequencies. The runs that do reach close to zero energy do it relatively quickly within the 600 generations. The runs which seem to be stuck in local optima are unable to get to a much lower energy throughout the generations, indicating the need for carrying out sufficient independent runs. Using the given parameter settings for DE the runtime in this case is around an hour per frequency for non-optimized code.

## B. Test case 2: Four monopoles

Test case 2 concerns the situation with 4 monopoles present. For the frequencies of 500, 1000, 1500, 2000, 2850, 3500, 6650, 8400, and 10 000 Hz conventional beamforming is applied, see Eq. (7). The result can be seen in Fig. 14. For the first four frequencies, the sources cannot be separated. This can be understood from Eq. (8): with the distance to the source being 0.75 m, the resolution will be too low. For the frequencies of 2850 Hz and higher, the source can be properly identified. The sidelobes become more prominent for higher frequencies and the sources become harder to localize.

It should be noted that, although indeed at sufficiently high frequencies, four sources can be identified in the source plots, the localization is affected by the mismatch between the modeled situation with a single noise source only and the actual situation with four sources presented simultaneously. This is not the case when using Eq. (10) and modeling the CSM according to Eq. (5) for four sources. This model is then employed in the inversion method with the parameters for DE found in Sec. VB. In this case, only the CSM energy function is used because of its additional advantage of estimating the source amplitude. The result can be seen in Fig. 15, indicating that from 1000 Hz on the sources can be identified accurately with good resolution. Upon inspection of Fig. 16, it can also be seen that many runs converge to zero energy, even at frequencies as high as 10 000 Hz. It can be noted that the CSM energy function works better compared with the single monopole case. Whereas at 10 000 Hz nine out of ten runs reach close to zero energy, this is only seven out of ten for the single monopole case at 6000 Hz. A disadvantage of this method is the need to know the amount of sources beforehand.

In Fig. 16 the convergence behaviour is seen for 10 different runs at frequencies of 1500, 3500, and 10 000 Hz. It can be seen that for the first two frequencies, 600 generations

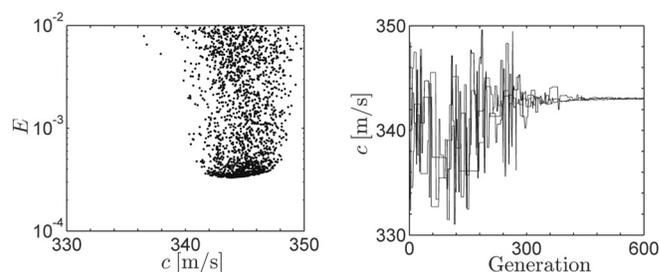


FIG. 17. Estimation of the speed of sound for the 4 monopole case at  $f = 6650$  Hz.

are sufficient. For 10 000 Hz it can be observed that the energy is getting closer to zero around 600 generations for most runs. From this, it can be concluded that the chosen DE setting parameters in general work well for this problem. However, for frequencies from  $\sim 10$  000 Hz on, a higher amount of generations is recommended. The runtime for the four sources was within 3 h per frequency for non-optimized code.

As a further illustration of using optimization methods, the problem is extended to include estimation for parameters other than those directly related to the acoustic source. To this end, the 4 monopole case is taken for  $f = 6650$  Hz and the speed of sound is considered as the additional unknown parameter. This was set to 343 m/s for the simulation. The result from the optimization is presented in Fig. 17. For low energies, it provides the values for the speed of sound to be in the range of 342 to 346 m/s. For the best three runs, it presents the value for the speed of sound converging to 343 m/s. This shows that, despite a relatively broad range of sound speeds corresponding to low energies, the best runs result in a sound speed of 343 m/s. Including these types of environmental parameters can be of interest also for complex wind tunnel experiments in aeroacoustic research.

## C. Test case 3: One speaker in an anechoic chamber

Conventional beamforming was applied again to the acoustic data from the microphone array to obtain a reference baseline result. Figure 18 presents the source map obtained for 5000 Hz. The position of the source is found with the expected spatial resolution but several sidelobes are displayed as well.

Figure 19 presents the results for 5000 Hz by using the proposed inversion method with the Bartlett energy function. Use was made of Bartlett since this energy function showed a high probability of success for cases with a single source present. For this situation with a single source, the source

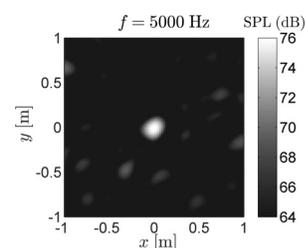


FIG. 18. Beamforming result for the speaker sound source at 5000 Hz using the randomly distributed array.

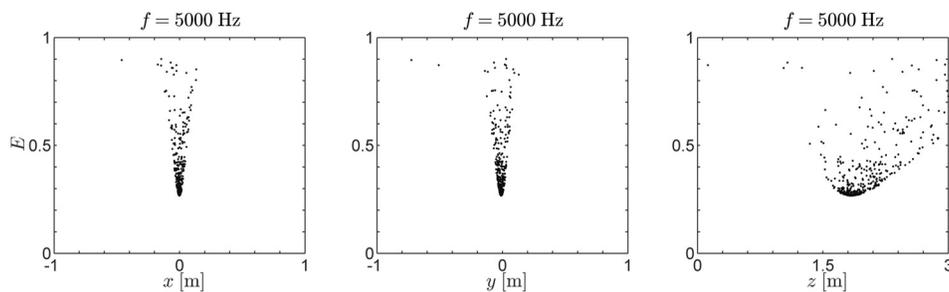


FIG. 19. Inversion results for the speaker source case at 5000 Hz using the Bartlett energy function.

strength follows directly from Eq. (7) for the source position derived through inversion. To achieve this result, DE was set to  $q = 128$ ,  $N_G = 600$ ,  $p_C = 0.75$ , and  $F = 0.4$ . The number of independent runs was selected to be 50. It can be observed that the source position is obtained correctly, since the values for  $x$ ,  $y$ , and  $z$  that correspond to the lowest energy values are in agreement with the actual source position.

To study the convergence rate for each coordinate in this case, Fig. 20 is presented, where the parameter value for three independent runs is given as a function of the number of generations. It can be observed that after approximately 50 generations, the result has rapidly converged to the correct solution for the three coordinates.

In Fig. 20 the convergence behaviour of the energy is presented for 10 different runs (out of the 50 used) at 5000 Hz. It can be clearly stated that 600 generations are more than sufficient to obtain convergence. However, in this case, an asymptotic energy value of approximately 0.25 is reached instead of zero, reflecting the always present imperfections of the measurement such as small errors in the microphone positions.

## VII. SUMMARY AND CONCLUSION

In this work, an inversion method is presented using the global optimization method DE to localize sound sources with a microphone array and determine the corresponding

source strengths. For this purpose, two energy functions were formulated. The first energy function used was the Bartlett processor. The second energy function is based upon modeling the CSM. The Bartlett energy function has the disadvantage of not determining the source strength, while the CSM energy function has no such disadvantage.

For both energy functions, the best parameters for DE were determined and subsequently used for source identification for three test cases: one simulated monopole source, four incoherent simulated monopole sources, and one speaker emitting at a single frequency in an anechoic room. For the single simulated monopole case, the performance of the localization of the source was better by using the Bartlett energy function. Still, both energy functions performed better compared with conventional beamforming by having few to no sidelobes. The improvement when using inversion as source identification was also clearly seen for the case with four sources. While conventional beamforming had trouble identifying the source at low and high frequencies, the inversion method was able to localize the sources accurately and with high resolution. Extending the optimization for four sources to include the estimation of the speed showed that in addition to the source locations, also the sound speed could be estimated. Moreover, for the experimental case with a speaker, the  $x$ ,  $y$ , and  $z$  positions of the source were correctly obtained using the Bartlett energy function.

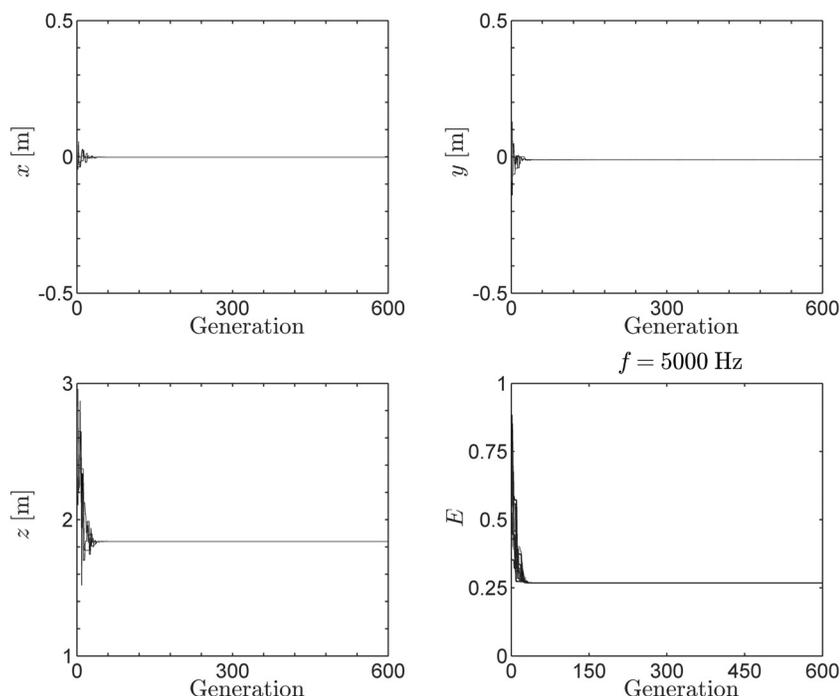


FIG. 20. Convergence of the spatial coordinates ( $x$ ,  $y$ ,  $z$ ) belonging to three runs for the optimal setting of  $F = 0.4$ ,  $p_C = 0.75$  at the population size of  $q = 128$  and the case with the speaker emitting at 5000 Hz. In the last figure the energy ( $E$ ) is given as a function of generation for ten independent runs.

The main advantage of the proposed inversion approach is that by employing global optimization methods that are efficient, in the sense that they require a limited number of forward calculations while still having a high probability of locating the global optimum, also for many unknowns, it becomes possible to search for all parameters of relevance by including them in the steering vector formulation. In addition, also parameters, such as those of the propagating medium, can be included as unknowns.

Moreover, the energy function can be adapted to the situation at hand. It can account for multiple sound sources, reflections, or refractions of the sound, thus representing the actual measurement environment to a large extent.

## LIST OF SYMBOLS

$\mathbf{A}$	$M \times L$ steering matrix consisting of $\mathbf{a}_l$ as columns
$\mathbf{a}_l$	steering vector corresponding to the $l$ th source
$a_{l,m}$	element of $\mathbf{a}_l$ corresponding to the $m$ th microphone
$\mathbf{C}$	CSM of microphone signals
$c$	speed of sound in air
$D$	diameter of microphone array
$\mathbf{d}_{k,u}$	descendant member $u$ of the population at generation $k$
$E$	energy function used for optimization
$\mathbf{IE}$	expectation operator
$F$	scalar multiplication factor between 0 and 1
$\mathbf{g}_{k,u}$	member $u$ of the population at generation $k$
$g_{k,uv}, h_{k,uv}, d_{k,uv}$	parameter $v$ of member $u$ of the population at generation $k$
$\mathbf{h}_{k,u}$	partner member $u$ of the population at generation $k$
$j$	$\sqrt{-1}$
$k$	(current) generation $k$
$L$	number of acoustic monopoles
$l$	acoustic monopole index
$M$	number of microphones
$m$	microphone index
$N_G$	number of generations
$\mathbf{P}$	signal sample covariance matrix
$p_c$	crossover probability
$q$	total members of the generation
$r_{l,m}$	distance between source $l$ and microphone $m$
$r_{l,0}$	distance between source $l$ and array center location
$r$	random value from the uniform distribution between 0 and 1
$\mathbf{s}$	acoustic waveform source vector
$s_l$	acoustic waveform amplitude of source $l$
$u, r_1, r_2, r_3$	mutually exclusive member of a population
$v$	parameter number for DE
$\mathbf{x}_l$	position of the $l$ th acoustic monopole
$\mathbf{x}_m$	position of the $m$ th microphone

$\mathbf{x}'$	scan point location
$\mathbf{y}$	microphone array output vector
$\tilde{\mathbf{y}}$	beamformer output
$z_{\text{array}}$	average distance between the array plane to the source(s)
$\Delta\ell$	spatial resolution of microphone array
$\omega, f$	radial and temporal frequency, $\omega = 2\pi f$

<sup>1</sup>B. D. van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Mag.* **5**(2), 4–24 (1988).

<sup>2</sup>J. J. Grefenstette, "Optimization of control parameters for genetic algorithms," *IEEE Trans. Systems, Man, Cybernetics A* **16**(1), 122–128 (1986).

<sup>3</sup>P. Gerstoft, "Inversion of seismoacoustic data using genetic algorithms and a *posteriori* probability distributions," *J. Acoust. Soc. Am.* **95**(2), 770–782 (1994).

<sup>4</sup>M. Snellen and D. G. Simons, "An assessment of the performance of global optimization methods for GEO-acoustic inversion," *J. Comput. Acoust.* **16**(02), 199–223 (2008).

<sup>5</sup>S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, "Optimization by simulated annealing," *Science* **220**(4598), 671–680 (1983).

<sup>6</sup>C. E. Lindsay and N. R. Chapman, "Matched-field inversion for geoacoustic model parameters using simulated annealing," *IEEE J. Ocean. Eng.* **18**, 224–231 (1993).

<sup>7</sup>M. Dorigo, V. Maniezzo, and A. Colnori, "Ant system: Optimization by a colony of cooperating agents," *IEEE Trans. Systems, Man, Cybernetics B* **26**(1), 29–41 (1996).

<sup>8</sup>M. Dorigo, G. Di Caro, and L. M. Gambardella, "Ant algorithms for discrete optimization," *Artificial Life* **5**(2), 137–172 (1999).

<sup>9</sup>R. Storn, "On the usage of differential evolution for function optimization," in *IEEE Biennial Conference of the North American Fuzzy Information Processing Society (NAFIPS)*, Berkeley, CA (1996), pp. 519–523.

<sup>10</sup>R. Storn and K. Price, "Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optim.* **11**(4), 341–359 (1997).

<sup>11</sup>T. F. Brooks and W. M. Humphreys, "A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays," *J. Sound Vib.* **294**, 856–879 (2006).

<sup>12</sup>T. F. Brooks and W. M. Humphreys, "Extension of DAMAS phased array processing for spatial coherence determination (DAMAS-C)," AIAA paper 2006–2654.

<sup>13</sup>T. Yardibi, J. Li, P. Stoica, and L. N. Cattafesta III, "Sparsity constrained deconvolution approaches for acoustic source mapping," *J. Acoust. Soc. Am.* **123**(5) 2631–2642 (2008).

<sup>14</sup>T. Yardibi, J. Li, P. Stoica, N. S. Zawodny, and L. N. Cattafesta III, "A covariance fitting approach for correlated acoustic source mapping," *J. Acoust. Soc. Am.* **127**(5), 2920–2931 (2010).

<sup>15</sup>G. F. Edelmann and C. F. Gaumont, "Beamforming using compressive sensing," *J. Acoust. Soc. Am.* **130**, EL232–EL237 (2011).

<sup>16</sup>A. Xenaki, P. Gerstoft, and K. Mosegaard, "Compressive beamforming," *J. Acoust. Soc. Am.* **136**, 260–271 (2014).

<sup>17</sup>A. Xenaki and P. Gerstoft, "Grid-free compressive beamforming," *J. Acoust. Soc. Am.* **137**, 1923–1935 (2015).

<sup>18</sup>D. G. Simons and M. Snellen, "Multi-frequency matched-field inversion of benchmark data using a genetic algorithm," *J. Comput. Acoust.* **6**(1–2), 135–150 (1998).

<sup>19</sup>C. F. Mecklenbräcker, P. Gerstoft, J. F. Böhme, and P.-J. Chung, "Hypothesis testing for geoacoustic environmental models using likelihood ratio," *J. Acoust. Soc. Am.* **105**(3), 1738–1748 (1999).

<sup>20</sup>M. Siderius, M. Snellen, D. G. Simons, and R. Onken, "Environmental assessment in the strait of Sicily: Measurement and analysis techniques for determining bottom and oceanographic properties," *IEEE J. Ocean. Eng.* **25**(3), 364–386 (2000).

<sup>21</sup>F. R. S. Rayleigh, "Investigations in optics, with special reference to the spectroscope," *Philos. Mag.* **8**(49), 261–274 (1879).

<sup>22</sup>R. P. Dougherty, "Beamforming in acoustic testing," in *Aeroacoustic Measurements*, edited by T. J. Mueller (Springer, Berlin, 2002), pp. 62–97.

<sup>23</sup>Benchmarking Array Analysis Methods 2015, <http://www.b-tu.de/fg-akustik/lehre/aktuelles/arraybenchmark> (Last viewed 21 April 2015).

<sup>24</sup>E. Sarradj, "Three-dimensional acoustic source mapping with different beamforming steering vector formulations," *Adv. Acoust. Vib.* **2012**, 1–12 (2012).

<sup>25</sup>D. F. Gingras and P. Gerstoft, "Inversion for geometric and geocoustic parameters in shallow water: Experimental results," *J. Acoust. Soc. Am.* **97**(6), 3589–3598 (1995).

<sup>26</sup>D. Blacodon and G. Élias, "Level estimation of extended acoustic sources using an array of microphones," AIAA paper 2003–3199.