

## Using global optimization methods for three-dimensional localization and quantification of incoherent acoustic sources

von den Hoff, B.; Merino Martinez, R.; Simons, D.G.; Snellen, M.

**DOI**

[10.1121/10.0010456](https://doi.org/10.1121/10.0010456)

**Publication date**

2022

**Document Version**

Final published version

**Published in**

JASA Express Letters (online)

**Citation (APA)**

von den Hoff, B., Merino Martinez, R., Simons, D. G., & Snellen, M. (2022). Using global optimization methods for three-dimensional localization and quantification of incoherent acoustic sources. *JASA Express Letters (online)*, 2(5), Article 054802. <https://doi.org/10.1121/10.0010456>

**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.

# Using global optimization methods for three-dimensional localization and quantification of incoherent acoustic sources

Cite as: JASA Express Lett. 2, 054802 (2022); <https://doi.org/10.1121/10.0010456>

Submitted: 02 March 2022 • Accepted: 19 April 2022 • Published Online: 11 May 2022

 Bieke von den Hoff,  Roberto Merino-Martínez,  Dick G. Simons, et al.



## ARTICLES YOU MAY BE INTERESTED IN

### [Machine learning in acoustics: Theory and applications](#)

The Journal of the Acoustical Society of America **146**, 3590 (2019); <https://doi.org/10.1121/1.5133944>

### [Distinguishing multiple surface ships using one acoustic vector sensor based on a convolutional neural network](#)

JASA Express Letters **2**, 054803 (2022); <https://doi.org/10.1121/10.0010492>

### [Very low frequency three-dimensional beamforming for a miniaturized aperture acoustic vector sensor array](#)

JASA Express Letters **2**, 054805 (2022); <https://doi.org/10.1121/10.0010535>





Read Now!

JASA  
THE JOURNAL OF THE  
ACOUSTICAL SOCIETY OF AMERICA

Special Issue:  
Lung Ultrasound



# Using global optimization methods for three-dimensional localization and quantification of incoherent acoustic sources

Bieke von den Hoff,<sup>a)</sup>  Roberto Merino-Martínez,  Dick G. Simons,  and Mirjam Snellen   
Aircraft Noise and Climate Effects, Faculty of Aerospace Engineering, Delft University of Technology, Kluyverweg 1,  
2629 HS Delft, The Netherlands

*B.vondenHoff@tudelft.nl, R.MerinoMartinez@tudelft.nl, D.G.Simons@tudelft.nl, M.Snellen@tudelft.nl*

**Abstract:** Complex acoustic systems typically present three-dimensional distributions of noise sources. Conventional acoustic imaging methods with planar microphone arrays are unsuitable for three-dimensional acoustic imaging, given the computational demands and the incapability to explicitly account for the presence of multiple sources. This paper proposes the use of global optimization methods to solve these shortcomings. An experiment with three incoherent speakers proved that this method can accurately determine the three-dimensional location and the respective sound level of each individual source. In addition, super-resolution is achieved beyond half the Rayleigh resolution limit. © 2022 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

[Editor: Charles C. Church]

<https://doi.org/10.1121/10.0010456>

**Received:** 2 March 2022 **Accepted:** 19 April 2022 **Published Online:** 11 May 2022

## 1. Introduction

Recent studies by the European Union under the Environmental Noise Directive and the World Health Organization show that a large part of the European population is exposed to noise levels that can be harmful for physical and mental health.<sup>1–3</sup> Noise exposure can lead to annoyance, hypertension, sleep disturbance, hearing impairment, decreased performance, and ultimately coronary artery diseases.<sup>4,5</sup> Thus, it is essential to reduce noise levels, which requires accurate knowledge of the location and individual contributions of all noise sources within a system.

To localize and quantify noise sources, phased microphone arrays are typically used. The pressures emitted by a complex sound source can be measured by the microphones in the array and then imaged with a beamforming method. The most well-known method is conventional beamforming (CB).<sup>6</sup> Although robust and easy to implement, CB suffers from high sidelobe levels and limited spatial resolution.<sup>7–9</sup> Moreover, CB requires an exhaustive search of source locations within a scanning grid. Given the predominant use of planar arrays and their poor resolution in the depth direction, normally two-dimensional scan grids are placed parallel to the array plane.<sup>10,11</sup> Also, CB assumes that a single monopole source is present per grid point, potentially causing an overestimation of the source amplitude when other sources are present elsewhere.<sup>12</sup>

Over the past years, a multitude of research was invested in a vast range of acoustic imaging methods to overcome these limitations.<sup>13</sup> Most methods are based on enhancing the CB source map through deconvolution, examples of which include DAMAS<sup>14</sup> and CLEAN methods.<sup>15–17</sup> Modifications to improve the CB algorithm also exist, such as functional beamforming<sup>18,19</sup> and orthogonal beamforming.<sup>20</sup> This paper, however, approaches the localization and quantification of acoustic sources as a global optimization (GO) problem.<sup>21</sup>

By using the GO approach, the use of a scan grid is not necessary, and a search in three-dimensional (3D) space can be initiated. As was shown in previous research, this approach can be used to determine the location of multiple sources simultaneously but also to determine propagation parameters of the medium, such as the sound speed.<sup>22–25</sup> In previous research, simulations of a single point source and multiple point sources were analyzed. Additionally, a single speaker in an anechoic room at a single frequency was used to test the algorithm. GO was even applied on an airfoil measurement in an aeroacoustic wind tunnel, although it was concluded that the energy function would have to be changed from point source to a distributed line source.<sup>25</sup>

These cases proved the ability of the algorithm to correctly identify sound sources by explicitly accounting for the simultaneous presence of multiple sources in simulations. However, for the case with multiple sources, all sources had the same normal distance to the array. In reality, sources might be located at different normal distances from the array.

<sup>a)</sup> Author to whom correspondence should be addressed.

Therefore, this research compares the ability to find multiple sources at the same normal distance and at different normal distances, to assess whether GO with differential evolution (DE) is able to locate real sound sources in a 3D space.

Additionally, this research investigates the ability of DE to separate sources below the Rayleigh resolution limit in a multi-source environment.

The paper is organized as follows: Sec. 2 explains the method of DE, including the energy function used and the choice of its optimal settings. The experimental setup used is described in Sec. 3. The results and discussion of DE applied to the experimental case are presented in Sec. 4. Last, Sec. 5 contains the conclusions of this research.

## 2. Method of DE

Generally, in GO, a population of candidate solutions, which contain the unknown parameters, is entered into the energy function, which quantifies the error between the measured data and the predicted data for a candidate solution. A population of size  $q$  contains  $q$  candidate solution members ( $\mathbf{m}$ ) per generation. By updating the candidate solutions through a number of iterations, the energy function is minimized. If the algorithm uses robust settings for the search, it converges to the (global) optimum. Possible methods for updating a candidate solution include simulated annealing,<sup>26,27</sup> genetic algorithms,<sup>28,29</sup> and ant colony algorithms.<sup>30,31</sup> Previously, the DE method from the group of genetic algorithms was used as a GO algorithm for experiments with acoustic sources.<sup>21,22</sup>

The candidate solution member contains possible values for the unknown parameters, henceforth called *search parameters*. For acoustic source identification, these are the source location coordinates  $(x, y, z)$  and source amplitude  $s$ . The source amplitude is squared in the search to  $s^2$ , such that the sound pressure level (SPL) can easily be calculated as  $10 \log_{10}(0.5s^2/p_{ref}^2)$ , where  $p_{ref} = 20 \mu\text{Pa}$ . Considering a single source, a candidate solution member contains four search parameters; thus,  $\mathbf{m} = (x, y, z, s^2)$ . When multiple sources are present, the size of a candidate solution member should be extended accordingly. For the experimental case considered featuring three sources, the candidate solution member is  $\mathbf{m} = (x_1, y_1, z_1, s_1^2, x_2, y_2, z_2, s_2^2, x_3, y_3, z_3, s_3^2)$ .

For the explanation of DE, a general reference is made to an *energy function*  $E$ , which could be any energy equation (see Sec. 2.1). The first population of candidate solutions is randomly determined within preset search boundaries, but to iteratively minimize the energy function, the candidate solutions have to be updated with DE. This approach follows Darwin's evolution theorem, in which natural selection takes place over multiple generations on the basis of how fit a member of a population is. The updated candidate solution members are so-called descendants  $\mathbf{d}$ , created from a member of the original population  $\mathbf{m}$  with a partner population member  $\mathbf{s}$ . The partner population should be equal in population size to the original population, and it is created from a combination of members of the original population. A member of the partner population is created according to

$$\mathbf{s}_{\gamma,r_1} = \mathbf{m}_{\gamma,r_2} + F(\mathbf{m}_{\gamma,r_3} - \mathbf{m}_{\gamma,r_4}). \quad (1)$$

The generation  $\gamma$  and population member index  $r$  are used to indicate which  $\mathbf{m}$  or  $\mathbf{s}$  is meant. The partner population is made from the same generation  $\gamma$  but from different original members  $r$ , which are mutually exclusive. The partner member is thus created by the values of an original member and  $F$  times the difference between two other original members. Multiplication factor  $F$  defines how different the partner member should be from the first original member as a larger  $F$  increases the difference between  $\mathbf{s}_{\gamma,r_1}$  and  $\mathbf{m}_{\gamma,r_2}$ . This increased difference between the partner and the original member leads to a stronger exploration of the search space as more substantially different combinations are formed and vice versa for smaller  $F$ .  $F$  can be set anywhere between 0 and 1 for optimal performance.<sup>32</sup>

A descendant  $\mathbf{d}$  from a current population member  $\mathbf{m}$  with a partner population member  $\mathbf{s}$  is found by crossover per search parameter between these members from the same generation. There are different types of crossover like single point crossover, two-point crossover, uniform crossover, or exponential crossover.<sup>33</sup> Uniform crossover has been used in aeroacoustic implementations.<sup>25</sup> It is applied by comparing a random number  $b$  from a uniform distribution on the interval  $[0,1]$  to a predefined crossover probability  $p_c$  lying between 0 and 1. By decreasing  $p_c$ , more often  $\mathbf{m}$  is chosen, and thus the combinations of promising search parameter values are less exploited and vice versa.

The population of descendants is now assessed by the energy function. When a descendant has a lower energy compared to the original population member, it is included in the next generation and vice versa.

This process is repeated for a limited number of generations  $N_{gen}$ , which should be selected high enough to ensure convergence. The outcomes are used to assess the energy landscape, and in particular the existence of local optima. This is then repeated for a number of independent runs ( $N_{runs}$ ). Therefore, the so-called *setting parameters* contain  $q, F, p_c, N_{gen}$ , and  $N_{runs}$ .  $F$  and  $p_c$  can be found through a sensitivity analysis where convergence is analyzed. The number of forward calculations is determined by

$$N_{forward} = qN_{gen}N_{runs}. \quad (2)$$

The number of forward calculations is not dependent on the number of search parameters. This number can be compared to the number of calculations necessary for acoustic imaging methods, such as CB, where it is dependent on the number of grid points.

### 2.1 Energy function for DE

The energy function that is evaluated is based on minimizing the difference between the measured cross-spectral matrix (CSM) and the modeled CSM. Therefore, this method uses both the phase and amplitude information. The measured CSM is constructed by Fourier transforming a time window of the measured time domain pressures of each microphone to the frequency domain. For a single frequency, the Fourier transform of the pressure of each microphone is placed into the pressure vector  $\mathbf{P}_{meas}$ . The measured CSM is then estimated by

$$\mathbf{C}_{meas}(f) = \frac{1}{2} \langle \mathbf{P}_{meas} \mathbf{P}_{meas}^* \rangle, \quad (3)$$

where  $\langle \rangle$  indicates the ensemble average over time. The factor  $\frac{1}{2}$  is introduced as the effective pressure is taken for the CSM. The asterisk indicates the complex conjugate transpose.

The modeled CSM explicitly accounts for the presence of multiple sources. This requires one to simultaneously search for the source amplitudes. The CSM is modeled based on the fact that, in the absence of background noise or reflections, the measured pressure on a microphone is the result of a pressure wave caused by a source of amplitude  $s$  that is phase shifted and attenuated over a distance  $r$  to an observer.

Often this is combined into the modeled pressure vector  $\mathbf{P}_{mod} = \mathbf{sg}$ , where  $\mathbf{g}(\mathbf{m}, f)$  is the steering vector containing the phase shifts corresponding to the distance between the source position and the different microphone locations. The elements of the steering vector for a single microphone  $n$ , a candidate solution member  $\mathbf{m}$ , and frequency  $f$  are

$$g_n(\mathbf{m}, f) = \frac{1}{r_n} e^{-2\pi i f r_n / c}, \quad (4)$$

where  $r_n$  is the distance between microphone  $n$  and a source location included in  $\mathbf{m}$  and where  $c$  is the speed of sound. Then the modeled CSM can be represented by

$$\mathbf{C}_{mod}(\mathbf{m}, f) = \frac{1}{2} \mathbf{P}_{mod} \mathbf{P}_{mod}^* = \frac{1}{2} \mathbf{s} \mathbf{g} \mathbf{g}^* s = \frac{1}{2} s^2 \mathbf{g} \mathbf{g}^*. \quad (5)$$

When multiple incoherent sources are present, the CSM is modeled for each source separately and then summed to obtain  $\mathbf{C}_{mod}$ . The energy of a solution,  $E_{CSM}$ , is finally computed as

$$E_{CSM}(\mathbf{m}, f) = \sum_{\text{elements of CSM}} |\mathbf{C}_{mod}(\mathbf{m}, f) - \mathbf{C}_{meas}(f)|^2. \quad (6)$$

### 2.2 Choice of optimal DE settings

The settings for the DE algorithm vary per publication due to the variety of applications and also the computational budget available.<sup>21,23–25</sup> In general, decreasing  $F$  means a decreased exploration of the search space and therefore a need for a larger population size  $q$  when  $N_{gen}$  is kept constant. This would, however, lead to an increase in forward calculations and hence computational cost. The same logic can be applied for decreasing  $p_c$ , which decreases the exploitation of promising search parameters of the partner population. Thus, an optimal balance between  $F$  and  $p_c$  should be found for a certain population size and number of generations.

For this paper, an analysis was performed to assess the energy of the final solution as a function of  $F$  and  $p_c$  for a baseline case. The baseline case is the most complicated case used, which has three different sources in a 3D domain, i.e., case 2 of Sec. 3. The case was analyzed only at  $f = 4992$  Hz for  $N_{gen} = 600$  and  $N_{runs} = 5$ . At this frequency, it is possible to recognize the sources separately with CB, and the image is not corrupted with a high level of sidelobes; thus, it is expected that DE will converge well. Two possibilities for the population size, i.e.,  $q = 64$  and  $q = 128$ , were considered. The multiplication factor  $F$  was increased from 0.2 to 1 in steps of 0.2, while the crossover probability  $p_c$  was increased from 0.45 to 0.85 in steps of 0.1. The search was initiated in a 3D space with bounds for  $x$  from  $-2$  to  $2$  m, for  $y$  from  $-2$  to  $2$  m, for  $z$  from  $0$  to  $4$  m, and for  $s^2$  from  $0$  to  $0.5$  Pa<sup>2</sup>. These boundaries are used throughout the whole research.

The final energy of the best run, i.e., having the lowest final energy, per combination of search parameters was evaluated.  $q = 128$  reached the lowest energy for  $F = 0.4$  and  $p_c = 0.55$ .  $q = 64$  has a minimum energy for  $F = 0.6$  and  $p_c = 0.75$ . Although using  $q = 128$  results in a slightly lower energy, it requires more forward calculations than  $q = 64$ . As both solutions converge, the less computationally expensive solution is chosen. Thus, the final DE setting parameters are  $q = 64$ ,  $F = 0.6$ , and  $p_c = 0.75$ . All runs converged to the same final solution.

## 3. Experimental setup

For the experiment, the anechoic chamber of the Faculty of Applied Sciences at Delft University of Technology was used. The room has dimensions  $8 \text{ m} \times 8 \text{ m} \times 8 \text{ m}$  and is covered with  $1 \text{ m}$  long glass wool wedges. The chamber is entirely decoupled from the building and has its own support, creating a fully vibration- and (virtually) sound-free room. There is no structural floor present, but there is a wire net on which a rigid floor can be installed if necessary. To set up the

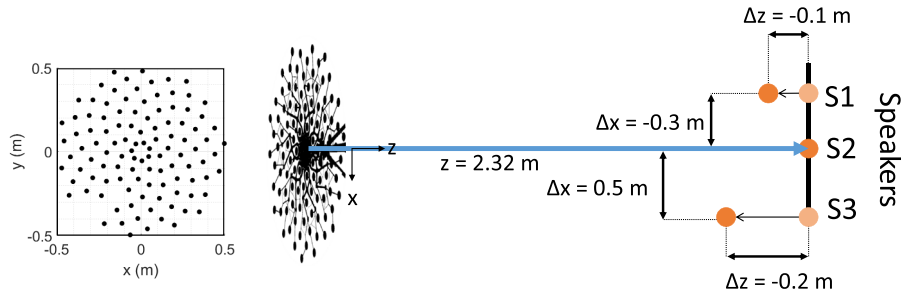


Fig. 1. (Left) The microphone array distribution as seen from behind. (Right) Top view of the speakers with respect to the microphone array in the anechoic room of Delft University of Technology. The speakers are indicated by orange dots, with lighter orange for the first position of S1 and S3. Their second position, darker orange, is set relative to the  $z$ -location of S2 to generate out-of-plane positions.

equipment for the experiment, a small grid was placed as floor to support the tripod of the microphone array. The sound sources were placed on a separate support system that did not need a rigid floor.

The acoustic measurements were recorded with a CAE Systems (Gütersloh, Germany) Bionic-M microphone array of 1 m in diameter, consisting of 112 microelectromechanical system (MEMS) digital microphones. Each microphone has been calibrated for phase and amplitude over the full frequency range. The microphone distribution is shown on the left of Fig. 1. Three microphones were ill-performing, and those were taken out of the analysis and hence are not present in the overview. The sound sources were placed in front of the array, on a wooden bar. As sound sources, three Visaton (Haan, Germany) FRWS 5 speakers were used. These speakers were coupled to an amplifier with white noise signals as input, which were generated in MATLAB. Each of the speakers had its own input signal to create incoherent signals. All speakers were set at the same overall SPL of 60 dBA measured at 1 m in front of the speaker baffle at the start of the measurement campaign. An amplitude degradation of source 1 was found throughout the campaign. The combination of three speakers in a 3D domain aims to simulate a complex noise source where multiple closely spaced elements cause noise.

The reference system is defined from the center of the microphone array. The bar with the three speakers is located in the  $xz$ -plane; thus, all speakers have the same  $y$ -coordinate of 0 m. One speaker is placed exactly opposite the center of the microphone array at a distance of 2.32 m, i.e.,  $(x, y, z) = (0, 0, 2.32)$ . This is speaker 2 (S2), which is the middle speaker. Speaker 1 (S1) is located in negative  $x$ -direction at  $x = -0.3$  m, and speaker 3 (S3) is located in positive  $x$ -direction at  $x = 0.5$  m. S1 and S3 can be moved in  $z$ -direction to create two different test cases: (1) all sources at the same  $z$ -distance at  $z = 2.32$  m and (2) all sources at different  $z$ -distance for which S1 was moved 10 cm toward the array, to  $z = 2.22$  m, while S3 was moved 20 cm forward to  $z = 2.12$  m. See Fig. 1 for a top view of the experimental setup with the speaker numbering. Note that for all frequencies above 796 Hz, the array is in the far-field according to the Fresnel distance.<sup>34</sup> Generally, this makes it more difficult to search for source distance  $z$ .

The microphone array is connected to a data acquisition system with a sampling frequency of 48 kHz. From the 15 s of measured signal, a part of 3 s is selected. This part is subdivided into time blocks of 2048 samples (0.043 s), which are averaged over 70 time blocks in the CSM. Thus, the resulting frequency resolution is 23.4 Hz. The results in this paper are presented for frequencies from 100 to 6000 Hz in steps of 2 times the frequency resolution to save computational cost.

#### 4. Results and discussion

Before the more complex multi-source cases were analyzed, the sound emissions of each speaker were analyzed individually. This analysis provides an insight into the performance of DE on the experiments, which can help in understanding the results of the multi-source cases. The single source analysis is only described here briefly. The case resulted in a perfect match for  $x$  and  $y$  and an accurate estimation of the source spectrum. However, an overestimation is present for  $z$  over a wide range of frequencies. This overestimation has a maximum around  $f = 4008$  Hz of 24 cm but reduces to 4 cm for  $f \geq 7195$  Hz. In this experiment, it was not possible to determine whether this offset for  $z$  is an effect of the algorithm, the directionality of the source, or perhaps an artifact of the microphone array or other experimental errors. As this offset is already present for the single source results, it was taken into account for the multi-source cases.

The first multi-source case is the case with the three speakers placed at the same  $z$ -distance. The results, presented in Fig. 2, are shown solely for the run resulting in the lowest energy function value with the colors defining the results from different speakers: S1 is yellow, S2 is orange, and S3 is blue. The solid black horizontal lines indicate the true location, while the Rayleigh resolution limit is indicated by the solid vertical gray line. It represents the minimum frequency (3236 Hz) at which the two closest spaced sources, i.e., S1 and S2, are separable in a CB source map at  $z \leq 2.32$  m.<sup>35</sup> In the results for  $z$ , the true location is indicated as well as the  $z$  estimated by DE for the single source analysis to determine whether multi-source cases find the same offset. For the source amplitude, the reference lines were determined by the average SPL in dB measured on the array for a single speaker, which is transferred back to the source



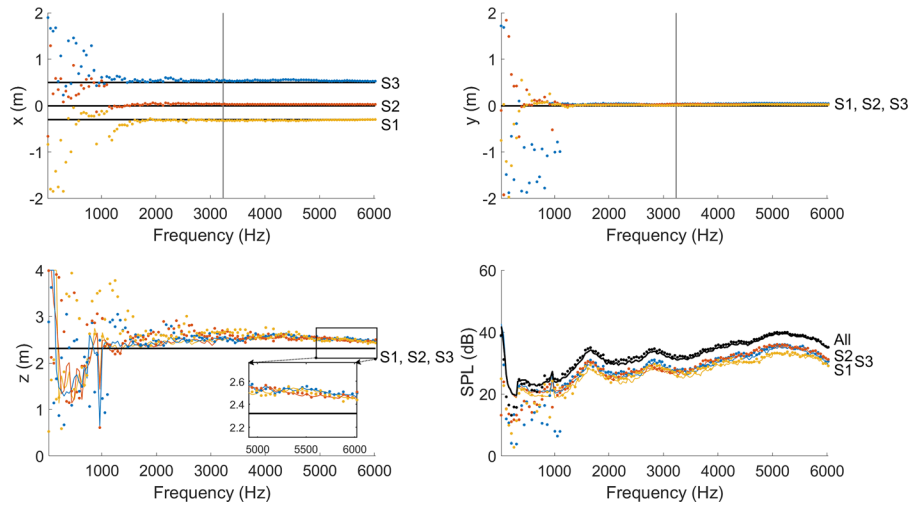


Fig. 2. Estimated location ( $x, y, z$ ) and SPL for each speaker per frequency for the case with the same normal distance  $z$ . The gray vertical line represents the Rayleigh resolution limit, while the black horizontal lines represent the true source coordinates. For  $z$ , the colored lines indicate the values of a single source search, while for SPL, the colored lines indicate the measured SPL in a single source measurement. S1 is yellow, S2 is orange, and S3 is blue.

by correcting for the spherical spreading. Additionally, the predicted spectra are summed and presented in black and compared with the total measured SPL, corrected to  $z = 2.32$  m.

The results in Fig. 2 show that DE is able to overcome the Rayleigh resolution limit. The  $x$  and  $y$  location of each sound source are correctly determined for  $f \geq 1523$  Hz and  $f \geq 1148$  Hz, respectively, which correspond to roughly one-half and one-third of the Rayleigh resolution frequency of 3236 Hz. Figure 2 also shows that S1 is the weakest source for most frequencies, while S2 and S3 are of comparable strength. Note that although  $y$  is the same for all speakers, it does not simplify the search for DE as it still has to find three individual source positions that are not linked to each other. The results for  $z$  show similarities with the single speaker results in terms of offset from the true values. All the three locations are only found at a similar  $z$  for  $f \geq 2508$  Hz. The results for SPL show that the individual spectra are determined quite well. However, the total spectrum is slightly overestimated, which can be explained from the overestimation of  $z$ . The average offset for  $f \geq 1523$  Hz per search parameter is also presented in Table 1. The values  $\mu$  are calculated as modeled value minus true value per parameter and consecutively averaged over the frequency range considered. Thus, a positive  $\mu$  represents an overestimation. This frequency range was chosen as  $f \geq 1523$  Hz, because this is the minimum frequency at which both  $x$  and  $y$  are accurately modeled. Including offsets for lower frequencies would increase  $\mu$  and not represent the offset trend well.

Thus, DE with the CSM energy function is successful in quantifying individual sources even under the Rayleigh resolution limit, which is not possible with other energy functions. This is due to the capability of the CSM energy function to account for multiple sources in the model.

This case has already proven two main advantages of DE over CB: (1) its ability to localize sources under the Rayleigh resolution limit and (2) its ability to correctly quantify individual sources when multiple sources are present.

The most complicated case presented in this paper has three out-of-plane sources as presented in Fig. 1. The Rayleigh frequency used is the same, although now the distance to the source varies per speaker, yet  $z = 2.32$  m is the worst case scenario. The results for this case are shown in Fig. 3. Note that the total measured SPL is still corrected with  $z = 2.32$  m, as it is not possible to correct the combined spectrum for individual distances.

Table 1. Mean offset between the modeled parameter value and true parameter value for  $f \geq 1523$  Hz. For the same  $z$  case, the mean offset of total  $\Delta$ SPL is 0.709 dB. For the different  $z$  case, the mean offset of total  $\Delta$ SPL is 0.315 dB.

	Same $z$			Different $z$		
	S1	S2	S3	S1	S2	S3
$\mu_{\Delta x}$ (m)	-0.016	0.032	0.045	-0.027	0.003	0.022
$\mu_{\Delta y}$ (m)	0.017	0.029	0.042	0.026	0.039	0.039
$\mu_{\Delta z}$ (m)	0.269	0.253	0.229	0.262	0.237	0.167
$\mu_{\Delta \text{SPL}}$ (dB)	0.730	0.413	0.668	0.970	0.133	0.527

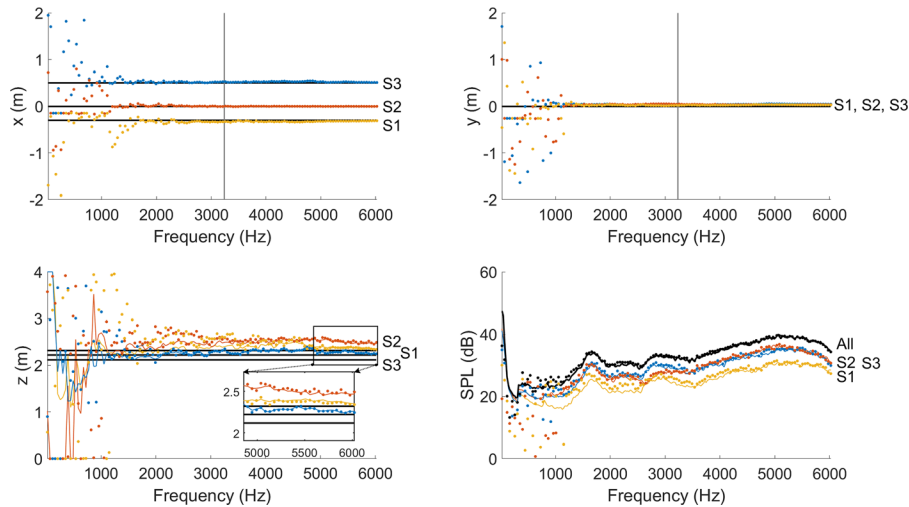


Fig. 3. Estimated location ( $x, y, z$ ) and SPL for each speaker per frequency for the case with different normal distances. The gray vertical line represents the Rayleigh resolution limit, while the black horizontal lines represent the true source coordinates. For  $z$ , the colored lines indicate the values of a single source search, while for SPL, the colored lines indicate the measured SPL in a single source measurement. S1 is yellow, S2 is orange, and S3 is blue.

Again, DE is able to localize the three different sources correctly in  $x$  and  $y$  below the Rayleigh resolution limit, and it does so for  $f \geq 1523$  Hz and  $f \geq 1242$  Hz, respectively. These new limitations correspond to roughly half the Rayleigh resolution frequency of 3236 Hz. The algorithm has more difficulty finding the correct  $z$  locations for the three sources, especially for S1, which is the weakest source. Whereas the two stronger sources already show a constant trend for  $z$  for  $f \geq 3023$  Hz, this happens only at 4992 Hz for the weakest source (S1). For all three,  $z$  is overestimated with the same offset as the single speaker cases; however, the relative  $\Delta z$  between the three sources is correct for high frequencies. This overestimation of  $z$  presents the same offset for the SPL as was seen for the previous case. Table 1 presents the average offset for  $f \geq 1523$  Hz per search parameter.

This last case proves that DE is also able to find the relative  $z$ -distances between sources at higher frequencies but that there still is an offset present in the absolute values for  $z$ . In the controlled environment used in this research, the offset of a single source can be used to correct for this. However, this can be a problem in uncontrolled environments.

A comparison was made between DE and CB to provide a first indication of the performance of DE. CB was applied to the case with different  $z$  for a frequency of 4992 Hz with a  $\Delta f$  of 23.4 Hz, as a comparison is only fair well above the Rayleigh resolution limit that CB suffers from. For this case, 3D beamforming has to be applied to find the different  $z$ , and thus a 3D grid of potential source locations was specified with the same boundaries as DE and with a grid step size of 1 cm. This results in 64 481 302 grid points, whereas the necessary forward calculations for DE are 192 000 for five runs. It is challenging to compare CB and DE in computational time since CB requires the calculation of a scan grid. This step is time consuming; however, it can be reused for every frequency. On the other hand, if DE has well-evaluated setting parameters and thus converges, a single run equals one CB source map. To give the reader an initial indication, DE took 179 s for a single run, and CB took 309 s for a single frequency on a Dell (Round Rock, TX) Precision Tower 5810  $\times$  64-based personal computer (PC), with a four-core Intel (Santa Clara, CA) Xeon central processing unit (CPU) E5-1620 at 3.50 GHz with 64 GB RAM.

Table 2 shows the offsets per method per source calculated as modeled value from CB or DE minus the true value. The offsets found with DE are generally slightly lower, hence indicating a more accurate result that holds for any

Table 2. Offsets found by 3D CB for each of the sources in the case with different  $z$  compared to the offsets found by DE at the same frequency. The offsets are calculated as modeled value minus true value, at  $f = 4992$  Hz with a  $\Delta f = 23.4$  Hz.

	CB				DE			
	$\Delta x$ (m)	$\Delta y$ (m)	$\Delta z$ (m)	$\Delta \text{SPL}$ (dB)	$\Delta x$ (m)	$\Delta y$ (m)	$\Delta z$ (m)	$\Delta \text{SPL}$ (dB)
S1	0.02	-0.04	-0.20	-1.30	0.023	-0.03	-0.17	-1.20
S2	0	-0.04	-0.28	0.24	0	-0.03	-0.19	1.33
S3	-0.03	-0.05	-0.20	-1.10	-0.027	-0.05	-0.16	-0.57



frequency above the Rayleigh resolution limit of 3236 Hz. Especially for  $z$ , DE shows an increased accuracy. For  $y$ , both methods show a similar offset of 3 or 4 cm, which is most likely caused by measurement inaccuracies due to the non-rigid floor in the anechoic room.

### 5. Summary and conclusion

In this research, the applicability of DE with the CSM energy function as a GO method to identify multiple acoustic sources was investigated. Additionally, the analysis was used to assess the identification of sources below the Rayleigh resolution limit and simultaneously find sources that are at different normal distances to the planar microphone array.

The algorithm was applied to an experimental case with three speakers in an anechoic room, which emitted incoherent white noise. For the first case, the three speakers had the same normal distance  $z$  to the array. DE was able to accurately localize the speakers in  $x$  and  $y$  but overestimated  $z$  by 15 cm with respect to the true distance of 2.32 m. This offset, however, is comparable to the offset when DE was applied for a single source. The  $x$ ,  $y$ , and SPL were correctly determined even under the Rayleigh resolution limit, i.e., the spatial resolution was improved by approximately a factor of 2 with respect to CB.

For the second case, the speakers were located at different normal distances to the microphone array. Again, DE was able to accurately localize the speakers in  $x$  and  $y$ , with a 2 times better spatial resolution than conventional acoustic imaging methods. The SPL is also determined accurately, however with an overestimation for the weakest source. This offset can be explained by the overestimation that DE found in  $z$ . Again, the algorithm had difficulty determining  $z$  for each source, especially at low frequencies for the weakest source. However, for higher frequencies, the final solution was stable, and the three sources were found at the correct distances with respect to each other.

Although this method is less sensitive to differences in  $z$  than differences in  $x$  and  $y$ , it is still able to locate and quantify sources accurately with a planar array. The localization is less accurate with CB methods and can only be performed above the Rayleigh resolution limit, thus giving DE an advantage over CB. Additionally, DE requires fewer forward calculations for a single frequency than CB to find source locations with a high accuracy, as for CB the grid step should be reduced to do so, while DE can search grid free.

The advantages of DE with the CSM energy function as presented in this research are fourfold: (1) it is able to accurately quantify individual sources in a multi-source environment, (2) it is able to localize and quantify sources even under the Rayleigh resolution limit, (3) it is able to find the respective normal distances between sources in a 3D space with a planar array, and last (4) it is able to achieve a high localization accuracy without a fine predefined 3D scan grid, thereby reducing computational costs.

### References and links

- <sup>1</sup>European Environmental Agency, “Environmental noise in Europe—2020,” EEA Report 22/2019, European Environment Agency, Copenhagen, Denmark (2020).
- <sup>2</sup>World Health Organization Regional Office for Europe, “Environmental noise guidelines for the European region: Executive summary,” WHO/EURO:2018-3287-43046-60243, World Health Organization, Copenhagen, Denmark (2018).
- <sup>3</sup>World Health Organization Regional Office for Europe and Joint Research Centre European Commission, “Burden of disease from environmental noise: Quantification of healthy life years lost in Europe,” World Health Organization, Copenhagen, Denmark (2011).
- <sup>4</sup>O. Hänninen, A. B. Knol, M. Jantunen, T. A. Lim, A. Conrad, M. Rappolder, P. Carrer, A. C. Fanetti, R. Kim, J. Buekers, R. Torfs, I. Iavarone, T. Classen, C. Hornberg, and O. C. Mekel, “Environmental burden of disease in Europe: Assessing nine risk factors in six countries,” *Environ. Health Perspect.* **122**(5), 439–446 (2014).
- <sup>5</sup>J. L. Peters, C. D. Zevitas, S. Redline, A. Hastings, N. Sizov, J. E. Hart, J. I. Levy, C. J. Roof, and G. A. Wellenius, “Aviation noise and cardiovascular health in the United States: A review of the evidence and recommendations for research direction,” *Curr. Epidemiol. Rep.* **5**(2), 140–152 (2018).
- <sup>6</sup>R. Merino-Martínez, “Microphone arrays for imaging of aerospace noise sources,” Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 2018.
- <sup>7</sup>Z. Prime and C. Doolan, “A comparison of popular beamforming arrays,” in *Proceedings of Acoustics 2013*, Victor Harbor, Australia (November 17–20, 2013).
- <sup>8</sup>J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing* (Springer, Berlin, Germany, 2008), Vol. 1, pp. 1–240.
- <sup>9</sup>P. Sijtsma, “Experimental techniques for identification and characterisation of noise sources,” NLR-TP-2004-165, Royal Netherlands Aerospace Centre, Amsterdam, The Netherlands (2004).
- <sup>10</sup>S. Oerlemans, “Detection of aeroacoustic sound sources on aircraft and wind turbines,” Ph.D. thesis, University of Twente, Enschede, The Netherlands, 2009.
- <sup>11</sup>Y. Zhou, V. Valeau, J. Marchal, F. Ollivier, and R. Marchiano, “Three-dimensional identification of flow-induced noise sources with a tunnel-shaped array of MEMS microphones,” *J. Sound Vib.* **482**, 115459 (2020).
- <sup>12</sup>R. P. Dougherty, “Functional beamforming,” in *Proceedings of the 5th Berlin Beamforming Conference*, BeBeC-2014-01, Berlin, Germany (February 19–20, 2014).
- <sup>13</sup>R. Merino-Martínez, P. P. Sijtsma, M. M. Snellen, A. A. Finez, G. G. Herold, A. A. Malgoezar, A. A. Pereira, and E. E. Sarradj, “A review of acoustic imaging methods using phased microphone arrays,” *CEAS Aeronaut. J.* **10**(3), 197–230 (2019).
- <sup>14</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**(4), 856–879 (2006).

- <sup>15</sup>P. Sijtsma, “CLEAN based on spatial source coherence,” NLR-TP-2007-345, Royal Netherlands Aerospace Centre, Amsterdam, The Netherlands (2007).
- <sup>16</sup>P. Sijtsma, R. Merino-Martinez, A. Malgoezar, and M. Snellen, “High-resolution CLEAN-SC: Theory and experimental validation,” *Int. J. Aeroacoust.* **16**(4), 274–298 (2017).
- <sup>17</sup>S. Luesutthiviboon, A. M. N. Malgoezar, R. Merino-Martinez, M. Snellen, P. Sijtsma, and D. G. Simons, “Enhanced HR-CLEAN-SC for resolving multiple closely spaced sound sources,” *Int. J. Aeroacoust.* **18**(4), 392–413 (2019).
- <sup>18</sup>R. Merino-Martínez, M. Snellen, and D. G. Simons, “Functional beamforming applied to imaging of flyover noise on landing aircraft,” *J. Aircr.* **53**(6), 1830–1843 (2016).
- <sup>19</sup>R. P. Dougherty, “Determining spectra of aeroacoustic sources from microphone array data,” in *Proceedings of the 25th AIAA/CEAS Aeroacoustics Conference*, Delft, The Netherlands (May 20–23, 2019).
- <sup>20</sup>E. Sarradj, “A fast signal subspace approach for the determination of absolute levels from phased microphone array measurements,” *J. Sound Vib.* **329**(9), 1553–1569 (2010).
- <sup>21</sup>A. M. N. Malgoezar, M. Snellen, R. Merino-Martinez, D. G. Simons, and P. Sijtsma, “On the use of global optimization methods for acoustic source mapping,” *J. Acoust. Soc. Am.* **141**(1), 453–465 (2017).
- <sup>22</sup>P. Kumar and M. Pant, “Noisy source recognition in multi noise plants by differential evolution,” in *Proceedings of the 2013 IEEE Symposium on Swarm Intelligence (SIS)*, Singapore (April 16–19, 2013), pp. 271–275.
- <sup>23</sup>M. Snellen and D. G. Simons, “An assessment of the performance of global optimization methods for geo-acoustic inversion,” *J. Comput. Acoust.* **16**(2), 199–223 (2008).
- <sup>24</sup>E. Sarradj, G. Herold, P. Sijtsma, R. M. Martinez, A. Malgoezar, M. Snellen, T. F. Geyer, C. J. Bahr, R. Porteous, D. J. Moreau, and C. J. Doolan, “A microphone array method benchmarking exercise using synthesized input data,” in *Proceedings of the 23rd AIAA/CEAS Aeroacoustics Conference, 2017*, Denver, CO (June 5–9, 2017).
- <sup>25</sup>R. Merino-Martínez, G. Herold, M. Snellen, and R. P. Dougherty, “Assessment and comparison of the performance of functional projection beamforming for aeroacoustic measurements,” in *Proceedings of the 8th Berlin Beamforming Conference, BeBeC-2020-S07*, Berlin, Germany (March 2–3, 2020).
- <sup>26</sup>C. E. Lindsay and N. Ross Chapman, “Matched field inversion for geocoustic model parameters using adaptive simulated annealing,” *IEEE J. Oceanic Eng.* **18**(3), 224–231 (1993).
- <sup>27</sup>S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, “Optimization by simulated annealing,” *Science* **220**(4598), 671–680 (1983).
- <sup>28</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>29</sup>J. Grefenstette, “Optimization of control parameters for genetic algorithms,” *IEEE Trans. Syst. Man Cybern.* **16**(1), 122–128 (1986).
- <sup>30</sup>M. Dorigo, V. Maniezzo, and A. Colorni, “Ant system: Optimization by a colony of cooperating agents,” *IEEE Trans. Syst. Man Cybern. B Cybern.* **26**(1), 29–41 (1996).
- <sup>31</sup>M. Dorigo, G. D. Caro, and L. M. Gambardella, “Ant algorithms for discrete optimization,” *Artif. Life* **5**(2), 137–172 (1999).
- <sup>32</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>33</sup>D. Zaharie, “Influence of crossover on the behavior of differential evolution algorithms,” *Appl. Soft Comput. J.* **9**(3), 1126–1138 (2009).
- <sup>34</sup>A. O. Williams, “The piston source at high frequencies,” *J. Acoust. Soc. Am.* **23**(1), 1–6 (1951).
- <sup>35</sup>F. R. S. Lord Rayleigh, “XXXI. Investigations in optics, with special reference to the spectroscope,” *Lond. Edinb. Dublin Philosoph. Mag. J. Sci.* **8**(49), 261–274 (1879).