Authors' accepted manuscript
Article published in *Mechanical Systems and Signal Processing*, Volume 132 (1 October 2019), p. 546-559
https://doi.org/10.1016/j.ymssp.2019.07.010

Optimization of a spherical microphone array geometry for localizing acoustic sources using the generalized cross-correlation technique

Thomas Padois[a], Olivier Doutres[a], Franck Sgard[b], Alain Berry[c]

[a] *Department of Mechanical Engineering, École de technologie supérieure (ÉTS), Montréal, (Qc), H3C 1K3 ,Canada*
[b] *Institut de Recherche Robert-Sauvé en Santé et en Sécurité du Travail, Montréal, (Qc), H3A 3C2, Canada*
[c] *Groupe d'Acoustique de l'Université de Sherbrooke, Sherbrooke, (Qc), J1K 2R1, Canada*

Abstract

Many workers are exposed daily to excessive noise levels. In order to reduce the noise exposure at the workstation, the main noise sources have to be detected. This task can be done with a microphone array and a source localization technique. Previous studies have shown promising results with time domain beamforming based on the generalized cross-correlation technique. The objective of this work is to propose an optimal spherical microphone array geometry dedicated to this technique. A cost function based on the symmetry of the aperture angle maps is proposed and is maximized using a Nonlinear Optimization by Mesh Adaptive Direct Search. **Numerical results show that the optimized geometry improves the noise source map by reducing the side lobe amplitude without increasing the main lobe surface.** Experimental measurements are carried out in a semi-anechoic chamber with prototyped spherical microphone arrays confirming that the optimized microphone array improves the quality of the noise source map.

Keywords: optimization, microphone array geometry, source localization, generalized cross-correlation

Email address: Corresponding author : Thomas.Padois@etsmtl.ca (Thomas Padois)

1. Introduction

In Quebec, hearing loss is by far the most recorded occupational disease [1]. Many workers are exposed daily to excessive noise levels, therefore noise control solutions have to be developed to continue prevention efforts. Before proposing acoustic solutions at the workstation, the main noise sources have to be identified.

Acoustic sources can be localized using acoustic intensimetry. A camera and position tracking system can be added to monitor the position of the intensity probe. The result is a map of source positions [2]. This measurement technique works in the vicinity of the sources and is useful when the main source has been detected because it allows for localizing acoustic leaks or secondary sources. However, acoustic intensimetry is not suitable to detect and rank multiple sources in a large workplace. The other available techniques are beamforming [3] and near-field acoustic holography (NAH) [4]. These techniques have shown their capacity to accurately identify noise sources in aircraft fly-over [5], snowmobile pass-by [6] or diesel engines [7]. Beamforming and NAH have also been used to detect acoustic sources on machine tools. For example Eret and Meskell [8] have used a 25 planar microphone array to detect compressed air leaks with a beamforming technique. Zhiwei *et al.* [9] have considered NAH to detect sources on an industrial sewing machine. Bai and Chen [10] have investigated a planar microphone array to detect noise sources on a turning lathe and a milling machine. In all these cases, the microphone array was set in front of the machine.

In a workplace such as an industrial hall or machine shop, the number of sources is highly variable and can be distributed over all directions around the worker, therefore acoustic intensimetry and planar microphone arrays are not appropriate. Noël *et al.* [11] proposed a 15-channel spherical microphone array to detect acoustic sources in an industrial hall relying on a localization technique based on the cross-correlation functions of the microphone signals, which is similar to a time domain beamforming, and the resolution of an inverse problem. Similar works have been performed either to improve the computation time of the inverse problem or the performance of the Generalized Cross-Correlation (GCC) technique [12, 13]. Others works have shown that the GCC technique is able to identify sound source positions in a reverberant environment or with few microphones [14, 15].

In the present work, the GCC technique and an open spherical array are combined to detect the acoustic source positions. The spherical array has to be small enough to be easily moved from one workstation to another and the number of microphones has to be low to ensure a fast computation time. Based on these constraints, a cost function is derived to define an optimal array geometry that will lead to accurate noise source maps.

Previous studies have addressed the optimization of microphone array geometry. In the case of frequency-domain beamforming using planar microphone arrays, noise source map metrics such as the Main Lobe Width (MLW) or the Maximum Side lobe Level (MSL) can be optimized [16, 17]. Planar spiral shaped arrays are a

common tradeoff [18, 19, 20] because they allow for reducing side lobes amplitude. However, these metrics depend on the source properties (position, frequency content, source-array distance). In the case of spherical arrays based on spherical harmonics decomposition, the microphone geometry should be chosen in order to avoid sampling errors in the computation of the spherical Fourier transform [21].

The spherical harmonics are not considered here and a cost function independent of the source properties is sought. Recently, a geometrical criterion has been proposed to reject side lobes and to decrease the MLW of the GCC noise source maps [13, 23]. This criterion is based on the cross product of two vectors defined by a scan point and the two positions of a given microphone pair. In this work, the angle value between these two vectors is investigated and the concept of aperture angle map is introduced. Depending on the considered microphone pair, the features of the aperture angle map are different. **The spherical microphone array geometry is optimized based on a cost function involving the concept of aperture angle map in order to reduce side lobes level.**

The article is organized as follows. Section 2 presents the source localization technique based on the GCC. The aperture angle map and the cost function are introduced in Section 3 where numerical results are presented. Section 4 presents experimental data that illustrate the performance of the optimized microphone array.

2. Source localization technique

2.1. Microphone array signals

Consider an acoustic signal $q_q(t)$, where t is time, generated by a source located at position $\mathbf{r}_q$. The acoustic waves propagate in free-field conditions and reach an array of M microphones at locations $\mathbf{r}_m$ $(m = 1, ..., M)$. The microphones record the acoustic pressures $p_m(t)$ which can be expressed as

$$p_m(t) = \alpha_{mq} q_q(t - \Delta t_{mq}) + v_m(t), \tag{1}$$

where α_{mq} is the geometrical attenuation due to the acoustic wave propagation between the source and the microphone m and $v_m(t)$ is an uncorrelated noise due to background or measurement noise. The Time of Flight (ToF) Δt_{mq} between the source and the microphone m is given by

$$\Delta t_{mq} = \frac{1}{c_0} \|\mathbf{r}_m - \mathbf{r}_q\|_2, \tag{2}$$

where c_0 is the sound velocity and $\|\cdot\|_p$ is the p-norm of a vector or matrix (which are denoted by bold characters). The microphone array signal $\mathbf{P}$ is a matrix with dimensions $[M \times N_t]$ (where N_t is the number of time samples)

$$\mathbf{P} = \begin{bmatrix} p_1(t_1) & p_1(t_2) & \dots & p_1(t_{N_t}) \\ \vdots & \vdots & \vdots & \vdots \\ p_M(t_1) & p_M(t_2) & \dots & p_M(t_{N_t}) \end{bmatrix}. \tag{3}$$

2.2. Generalized cross-correlation

Based on the microphone signals, different source localization techniques can be used. In acoustics, the most standard techniques are beamforming and near-field holography and they can be implemented in the time or frequency domain. In this paper, time domain beamforming is chosen. The technique is based on the computation of the cross-correlation function R_{mn} which is typically recovered from an inverse Fourier Transform of the cross-spectrum C_{mn} between microphone signals

$$R_{mn}(\tau) = \sum_{i=0}^{N_f-1} C_{mn}(i) \exp\left(j2\pi \frac{i}{N_f}\tau\right), \tag{4}$$

with,

$$C_{mn}(i) = P_m(i)P_n^*(i), \tag{5}$$

where i is the frequency index, N_f the number of frequency samples and $j = \sqrt{-1}$. P_m and P_n are the fast Fourier transforms of p_m and p_n and $(\cdot)^*$ denotes the complex conjugate. The cross-correlation function measures the similarity between two time signals. For two microphones recording the same acoustic source, the maximum of the cross-correlation function estimates the time delay between the two microphones. If the cross-correlation function is computed for all microphone pairs, a matrix of size $[M \times M]$ is obtained for each time lag τ. To minimize the computation time, the size of the cross-correlation matrix $[M \times M]$ is decreased by removing the redundant microphone pairs, which means the pair (m, n) is computed but the pair

(n, m) is not. Moreover, since the diagonal of the cross-correlation matrix provides the auto-correlation functions which correspond to a null time delay, these terms are discarded because they do not provide useful information on propagation delay. Removing these terms is equivalent to the diagonal removing of the cross-spectral matrix in the frequency domain beamforming. Finally, the cross-correlation matrix becomes a vector with size $[1 \times M_p]$, where M_p is the number of unique microphone pairs. This cross-correlation vector provides a set of time delays that depend on the microphone and source positions and the sound velocity.

To localize the source, a search area called scan zone has to be defined. The scan zone is a region of space which is sampled with N_s discrete points. For each scan point s and microphone pair (m, n), a time delay can be computed

$$\tau_{mn}^{s} = \Delta t_{ms} - \Delta t_{ns}. \tag{6}$$

For a given microphone pair and all the scan points, the previous time delay is a vector of dimension $[1 \times N_s]$. The maximum value of this vector represents the maximum time delay reachable for the considered microphone pair. Therefore, the values of the cross-correlation function evaluated at time lags above this maximum are discarded. Finally, the cross-correlation values are linearly interpolated over the scan zone. For a given microphone pair, the corresponding cross-correlation function plotted over the scan area is called Spatial Likelihood Function [13] (SLF). Each SLF corresponds to a hyperbola

and the location where the SLFs (each corresponding to a different microphone pair) intersect, is the source position. The sum of all the SLFs is called the noise source map.

A prefiltering operation $W(i)$ can be added to the cross-correlation function Eq. 4 which leads to the GCC, denoted R'_{mn},

$$R'_{mn}(\tau) = \sum_{i=0}^{Nf-1} W(i) P_m(i) P_n^*(i) \exp\left(j2\pi \frac{i}{N_f}\tau\right). \quad (7)$$

Among the various prefiltering techniques, the PHAse Transform (PHAT) [22] is the most employed method [23, 24]. The PHAT operation normalizes the magnitude of the cross spectrum in Eq. 7 but keeps its phase unchanged.

2.3. Noise source maps

In this section, numerical noise source maps obtained with Matlab are provided for an 18 microphones spherical array. These examples are used to introduce the cost function discussed in Section 3.

The microphones are located on a sphere with a 20 cm radius. Among the 18 microphones, 6 microphones are set on the equator with a 60° spacing. The other microphones are set at elevation angles 45° and -45° with the same azimuth angles as on the equator (see black points in Figure 1.a). This geometry is called "regular" because the microphones have the same azimuth angles. Consider a source in far-field, located at azimuth angle $\phi = 0$ and elevation angle $\theta = 0$ which generates acoustic waves in free-field conditions.

The sound speed is 343 m/s and the source signal is white Gaussian noise (*wgn* function in Matlab). The microphone signals are sampled at 44000 Hz and are filtered by a 2^{nd} order bandpass Butterworth filter between 100 and 10000 Hz. The duration of the source signal is 1 s and the signal to noise ratio is 0 dB (white Gaussian noise is added with *awgn* function in Matlab). The scan zone extends from -180° to 180° in the azimuth direction and from -90° to 90 ° in the elevation direction. The grid size is (101×101) which leads to 10201 scan points. No prefiltering operation is used ($W(i) = 1$ in Eq 7). The noise source map obtained is shown in Figure 1.a. It should be noticed that the scan zone is a sphere which is unwrapped for clarity. The results at the pole are distorted due to unwrapping, for instance a horizontal long line is actually a point. The peak value of the noise source map provides the source position and corresponds to the maximum of the main lobe. In the current case, the source position is correctly located. **The MLW is usually characterized by its width at-3 dB [19] along the observation directions (x and y or azimuth and elevation, here). According to the reference [25], the maximum of the two directions is defined as the MLW. Spurious lobes with a lower amplitude are also present, they are called side lobes. The amplitude difference between the main lobe peak and the closest side lobe is the MSL.**

If the microphone array geometry is slightly modified, the noise source map is different as shown in Figure 1.b. In this case, the 6 microphones located on the equator are shifted by -15° along the azimuth direction. This

small shift modifies the noise source map pattern by removing side lobes at the corners of the map and changing the side lobe pattern around the main lobe. The slices of the noise source map at $\phi = 0$ (elevation direction) and $\theta = 0$ (azimuth direction) are presented in Figure 2.a and Figure 2.b, respectively. Along the azimuth direction, the side lobe amplitude, in the vicinity of the main lobe, is decreased for the shifted configuration.The MSL difference between the regular and shifted arrays are -1.4 dB (for $\phi < 0$) and -1.5 dB (for $\phi > 0$). However, side lobes with higher amplitude appear. Along the elevation direction, the side lobe amplitude is decreased (Figure 2.b), the MSL difference are -2.3 dB (for $\theta < 0$) and -2.2 dB (for $\theta > 0$). **The MLW, based on the azimuth and elevation slices, remains the same between the two microphone configurations**.

In order to avoid for selecting a slicing direction as done for the MLW calculation, a covariance ellipse (with 95% confidence) is used to characterize the noise source map in this work. The covariance ellipse surrounds an area of the noise source map where the values are higher than a given threshold. The ellipse is obtained i) by selecting the data of the noise source map higher than the threshold and ii) by computing the covariance matrix of the data. The minor and major axes of the ellipse are given by the eigen-vectors and the radii by the eigen-values of the covariance matrix [26]. The main advantage is that the covariance ellipse takes into account the two dimensional aspect of the noise source map while the slices provide

an information along a unique direction. The choice of the threshold value allows for quantifying the main lobe or the presence of side lobes. For instance, a -3 dB threshold provides an ellipse which surrounds the main lobe, therefore it is possible to determine the main lobe surface by computing the surface ellipse. The -12 dB threshold allows for estimating the presence of side lobes. Finally, a surface reduction can be assessed by computing the relative error between two surfaces ($(S_1 - S_2)/S_1$) where S_1 and S_2 are the surfaces obtained with the regular and shifted (or optimized) array for instance. The covariance ellipses at -3 dB and -12 dB are shown in Figure 2.c (a zoom is provided in Figure 2.d). The main lobe surfaces, obtained from the covariance ellipse at -3 dB are the same with the regular and shifted arrays (see Figure 2.d). In contrast, the covariance ellipse at -12 dB is reduced with the shifted array (63% of surface reduction). To conclude, a slight change in the geometry can modify the noise source map. Therefore, an optimal geometry has to be found to decrease the side lobe amplitudes.

3. Optimization problem

3.1. Aperture angle map

The objective of the optimization problem is to find a microphone array geometry which maximizes or minimizes a cost function. In previous array geometry optimization, the MLW or the MSL were chosen as cost

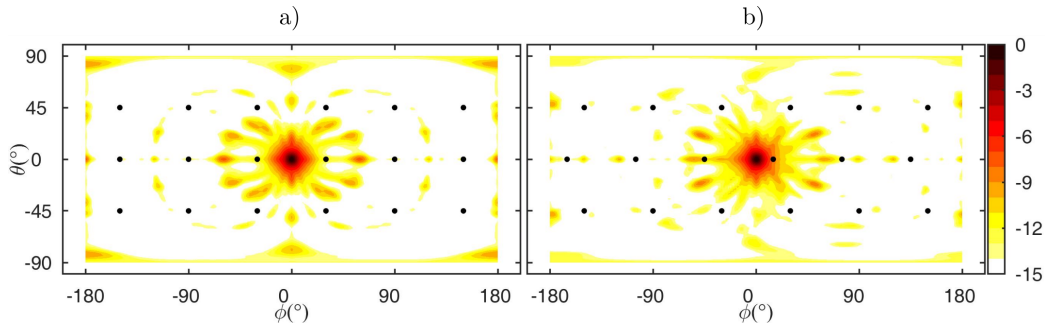

Figure 1: Noise source maps obtained with a) a regular microphone array and b) a regular microphone array where the equator microphones are shifted by -15°. The black dots are the microphone positions. The source generates white Gaussian noise and is located at $\phi = 0$ and $\theta = 0$. The colorbar is in dB. (Color online)

function [16, 17]. Reducing the MLW may affect the MSL and conversely. A multi-objective optimization can be used to tackle this problem at the expense of a more complex computation. Moreover, the MLW and MSL depend on the source position and frequency content (the higher the frequency, the smaller the MLW). Therefore, a cost function based on the resulting noise source map is not the best choice. A cost function allowing for overcoming the source influence together with an optimization procedure depending only on geometrical properties of the microphone array and scan zone are desired.

Quaegebeur *et al.* have proposed a geometrical criterion for improving noise source maps derived from GCC calculations [23]. This criterion is based solely on the scan point and microphone positions and is used in the following. Considering a scan point at position $\mathbf{r}_s$, it is possible to define two vectors originating from the scan point to a microphone pair position $(\mathbf{r}_m, \mathbf{r}_n)$. The angle between both vectors $\beta \in [-\pi/2, \pi/2]$, called aperture angle, can be

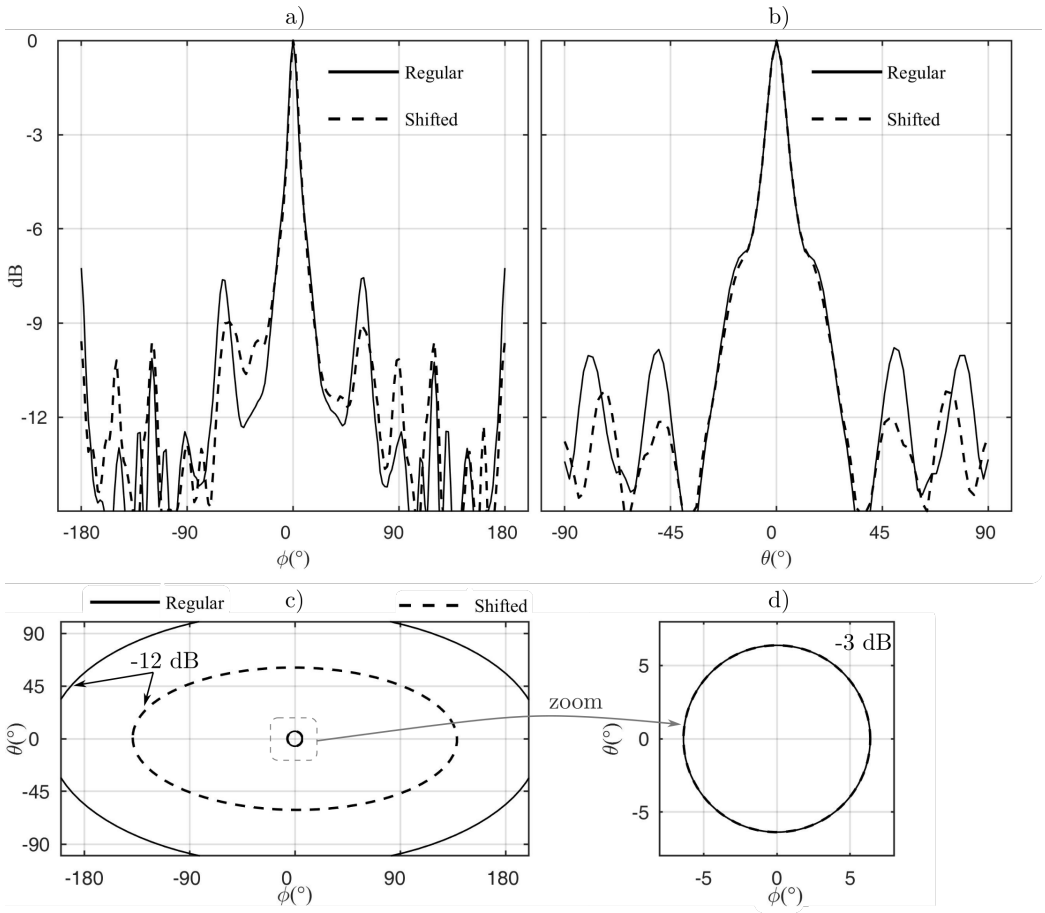

Figure 2: **Slices of the noise source maps (Figure 1) along a) azimuth and b) elevation directions. c) Covariance ellipses at -3 dB and -12 dB. d) Zoom in the covariance ellipses at -3 dB.**

computed from the normalized cross product of these vectors according to Eq. 8. Therefore, an aperture angle map $\beta^{mn}(\phi, \theta)$ can be computed for all the scan points s at angles (ϕ, θ) and a microphone pair (m, n),

$$\beta^{mn}(\phi, \theta) = \arcsin\left(\frac{\|\overrightarrow{\mathbf{r}_{ms}} \times \overrightarrow{\mathbf{r}_{ns}}\|_2}{\|\overrightarrow{\mathbf{r}_{ms}}\|_2 \|\overrightarrow{\mathbf{r}_{ns}}\|_2}\right). \tag{8}$$

Figure 3 shows aperture angle maps for two different microphone pairs, called pair A and pair B (microphone locations are shown by red crosses). Each aperture angle map is normalized by its peak value. In the region comprised between the two microphones, the aperture angle takes the largest values in Figure 3.a. A decrease in the aperture angle map value (see e.g. values at $\phi = \pm 80°$ and $\theta = 0$ in Figure 3.a) means that both vectors $(\mathbf{r}_{ms}, \mathbf{r}_{ns})$ become collinear. Both aperture angle maps present different features. A potential link with the corresponding SLF is investigated in the following section.

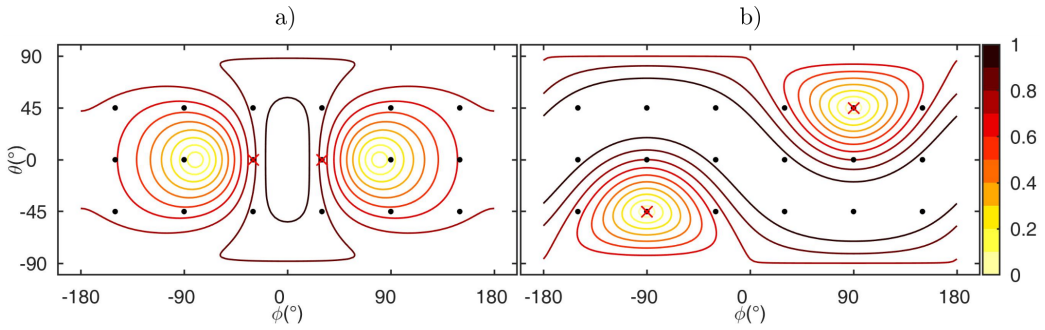

Figure 3: Aperture angle maps for the microphone a) pair A and b) pair B. The black dots are the microphone positions. The red crosses are the microphones considered. (Color online)

3.2. Spatial likelihood function

First, it is worth remembering that a SLF is obtained for each microphone pair and the noise source map is obtained by averaging the M_p SLFs. Figure 4 shows four SLFs obtained with microphone pairs A and B in the case of two different source positions. In the first case, the source is located at $\phi = 5°$

and $\theta = 5°$ where both aperture angle maps take high values (see Figure 3.a and Figure 3.b). In the second case, the source position is $\phi = -120°$ and $\theta = -25°$ where the aperture angle map values are smaller for both microphone pairs A and B. The source parameters are the same and identical to those of Section 2.3.

The SLFs consist in a main isoline (as shown in Figure 4) due to the main lobe of the cross correlation function and secondary isolines due to the side lobes of the cross correlation function. If the main isoline is large, a large main lobe may be expected or if the SLF presents secondary isolines with high amplitude, side lobes with high amplitude may be expected.

In the case of the microphone pair A and source position 1, the main isoline surrounds the right side of the SLF ($\theta > 0$) and a secondary isoline is apparent on the other side ($\theta < 0$). In the case of microphone pair B and source position 1, the main isoline follows a sine curve with a small width and the amplitude of the secondary isolines is low. In the case of source position 2, both microphone pairs provide a similar main isoline, however the amplitude of the secondary isolines for microphone pair A is higher. In each case, the SLFs follow the iso-contour lines of the aperture angle maps. These observations lead to the conclusion that the shape of individual SLFs is related to the angle aperture map (which is based onto the array geometry and the scan zone). Moreover, microphone pair B should provide more accurate SLFs than microphone pair A in these cases because the main isoline is narrower or the amplitude of the secondary isolines is lower. Based

on these observations, a cost function relying on the geometry of aperture angle map is proposed.

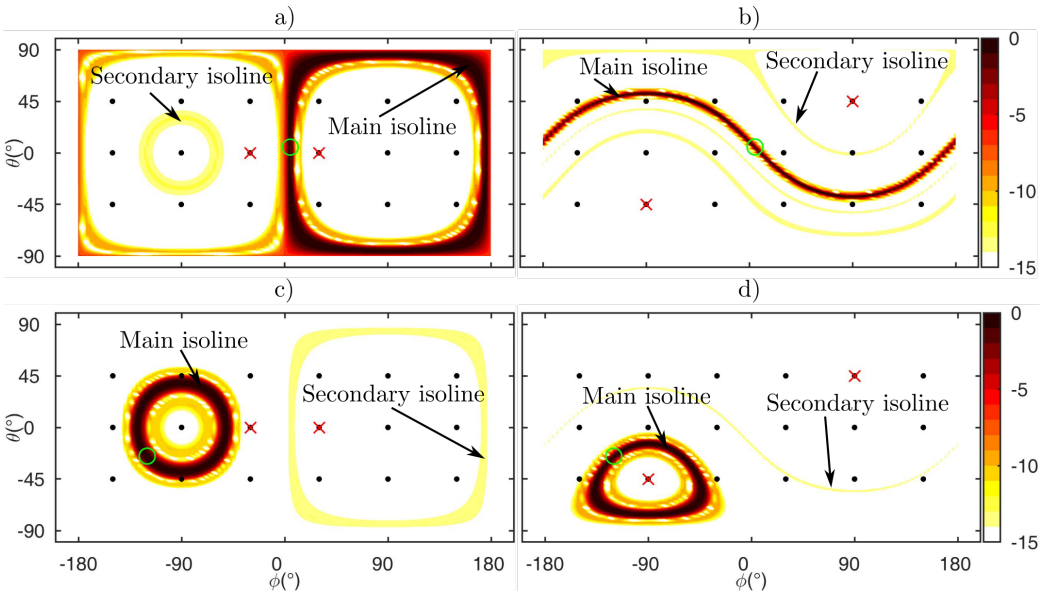

Figure 4: Spatial likelihood functions for a) pair A - source 1, b) pair B - source 1, c) pair A - source 2 and d) pair B - source 2. The black dots are the microphone positions. The red crosses are the microphones considered . The green circle is the source position. The colorbar is in dB. (Color online)

3.3. Cost function

In this section, the symmetry of the aperture angle map is investigated. With the microphone pair A, horizontal and vertical symmetries of the aperture angle map can be observed. These symmetry properties are not present in the aperture angle map obtained with the microphone pair B. To evaluate the symmetry of the aperture angle map $\beta^{mn}(\phi, \theta)$ for the microphone pair (m, n), two variables, denoted J_h and J_v, are computed

$$J_h^{mn}(\phi) = \frac{1}{N_\theta} \sum_{i=1}^{N_\theta} |\beta^{mn}(\phi, \theta_i) - \beta^{mn}(\phi, -\theta_i)|, \tag{9}$$

$$J_v^{mn}(\theta) = \frac{1}{N_\phi} \sum_{i=1}^{N_\phi} |\beta^{mn}(\phi_i, \theta) - \beta^{mn}(-\phi_i, \theta)|, \tag{10}$$

with N_ϕ and N_θ the number of scan points along each direction. The resulting vectors size are $[1 \times N_\phi]$ and $[1 \times N_\theta]$ for $J_h^{mn}(\phi)$ and $J_v^{mn}(\theta)$ respectively. if the aperture angle map is symmetric with respect to θ (respectively ϕ), then $J_h^{mn}(\phi)$ (respectively $J_v^{mn}(\theta)$) is zero.

To obtain the cost function, Eq. 9 and Eq. 10 are averaged

$$J^{mn} = \frac{1}{N} \sum_{i=1}^{N} J_h^{mn}(\phi_i) + J_v^{mn}(\theta_i), \tag{11}$$

where N is the number of point to perform the average. If the number of scan points is different along azimuth and elevation directions, the horizontal or vertical variables are interpolated to obtain the same vector length. In this case, the vector length is similar for each direction and equal to $N_\phi = N_\theta = N = 101$ scan points. Low values of the cost function refer to aperture angle maps with high symmetry relative to θ and ϕ, whereas higher values correspond to aperture angle maps with high anti-symmetry.

The cost function J^{mn} calculated for the regular array of 18 microphones presented in Section 2.3 and sorted in increasing value as a function of the microphone pairs is presented in Figure 5. Microphone pair A (considered in Figure 3.a) corresponds to the lowest value of the function. In this case,

the aperture angle map is fully symmetrical along both directions (θ and ϕ). Microphone pair B (considered in Figure 3.b) has the highest value. In this case, the aperture angle map is highly anti-symmetrical. In the following, few microphone pairs (m, n), i.e. few SLFs, are used to compute the noise source map.

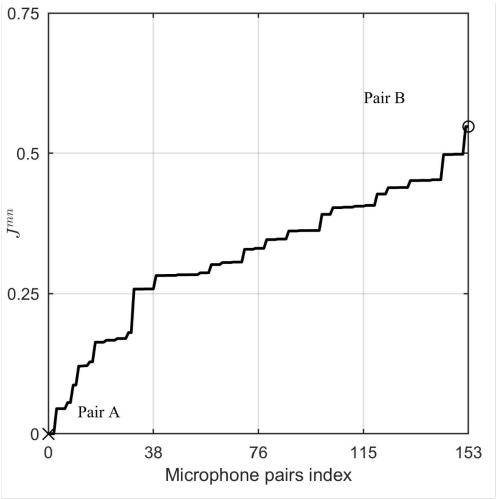

Figure 5: Cost function for the regular microphone array.

In order to highlight the relevance of the proposed cost function, the noise source map is now computed with only ten SLFs. The lowest, half the peak value (called average value of the cost function in the following) and highest values of the cost function are selected (see Figure 6.a-c). Two source positions are investigated, one at the origin and the other one randomly located ($\phi = 42°$ and $\theta = 17°$). The noise source maps obtained are compared in Figure 6.d-i.

When the source is at the origin, each map provides a peak value at this position (Figure 6.d-e). With the lowest values of the cost function, a vertical isoline with a high amplitude (-3 dB) is present as well as isolines surrounding the peak value (Figure 6.d). With the average values of the cost function, the side lobes are less extended and the amplitude is reduced (Figure 6.e). The side lobe amplitude is the lowest with the highest values of the cost function where only a cross is present at the source position (Figure 6.f).

When the source is randomly located, the noise source map obtained with the lowest values cannot localize the source position because two main lobes with the same amplitude exist (Figure 6.g). With the average values of the cost function, the source is detected but with many side lobes surrounding the source position (Figure 6.h). With the highest values, the side lobe amplitude is decreased and it is easier to localize the source position (Figure 6.i).

Therefore, microphone pairs that provide a high value of the cost function J^{mn} (Eq. 11) lead to more accurate noise source maps than those with a lower value. It should be noticed that the cost function proposed does not directly act on the main lobe or the side lobes of the noise source map. The cost function allows selecting properly the SLFs in order to limit the generation of a large main lobe many side lobes. It can be seen as an indirect method acting before the averaging process of the SLFs, which leads to the noise source map. To conclude, the optimization problem consists, for a given number of microphones, in finding the microphone positions which maximize the minimum value of the cost function $\mathbf{J}$ over all possible microphone pair

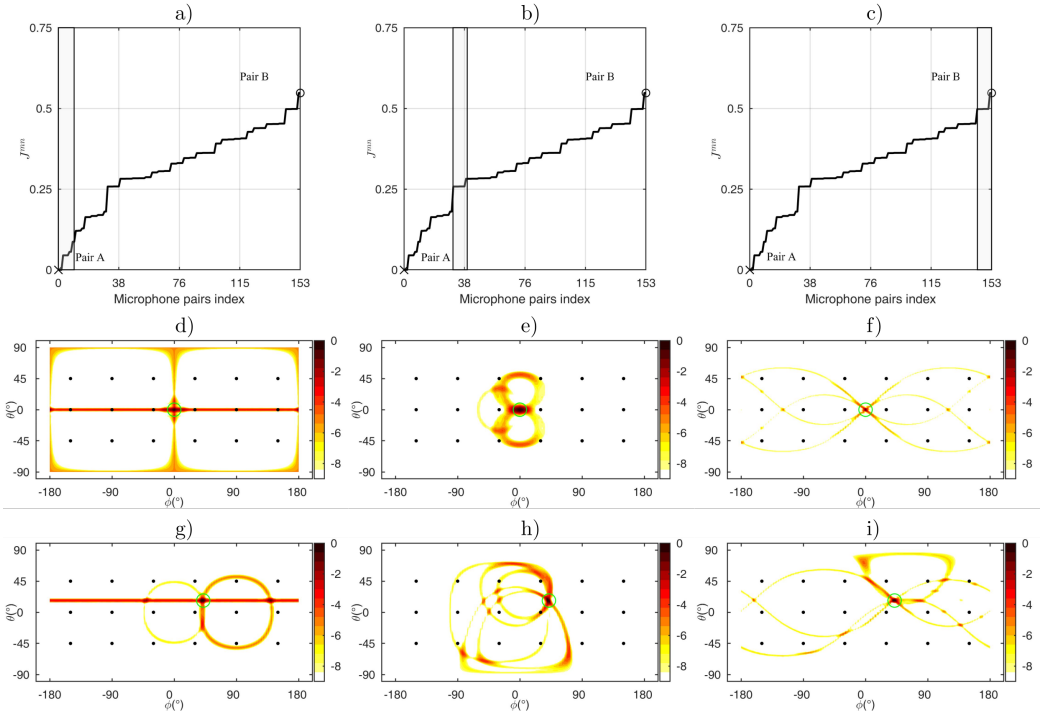

Figure 6: Cost function for the regular microphone array with a) the ten first b) the average and c) the ten last values selected. The gray patch shows the chosen microphone pairs. Noise source maps obtained with the source at origin d,e,f) and randomly located g,h,i). The black dots are the microphone positions and the green circles are the source positions. The source generates white Gaussian noise. The colorbar is in dB. (Color online)

positions $(\mathbf{r}_m, \mathbf{r}_n)$

$$\underset{(\mathbf{r}_m,\mathbf{r}_n)}{\text{maximize}} \min\left(\mathbf{J}\right), \tag{12}$$

with

$$\mathbf{J} = [J^{1,2}, \ldots, J^{M-1,M}]. \tag{13}$$

3.4. Optimization solver

Now that a cost function has been defined, it is necessary to choose an optimization solver. Optimization is a dedicated topic in mathematics and this section does not aim to propose an improved solver but rather use an existing adequate and efficient solver.

Integer variables for microphone angular positions (in degrees) are considered, therefore a 1° precision is assumed. The derivatives of the cost function are not analytically accessible thus the cost function can be seen as a "black box". A suitable solver is the Nonlinear Optimization by Mesh Adaptive Direct Search which allows for solving black box problems [28].

To avoid obtaining two microphones at the same location or all the microphones on the same side of the sphere, constraints have to be introduced. First, boundaries of the optimization space are defined to avoid microphone positions beyond $\phi = \pm 180°$ and $\theta = \pm 90°$. Also, each microphone is constrained to be located into a region depicted in Figure 7. Finally an inequality constraint is added. The separation between two microphones along azimuth or elevation directions has to be larger than 35°. This condition prevents two microphones to be located on the same boundary of the rectangular regions (Figure 7). No microphones are allowed beyond $\theta = \pm 70°$. These regions are left empty in order to set the tripod which holds the lower surface of the spherical microphone array and a 360° camera which is placed over the top of the sphere (not used in this work).

Five distinct initial solutions are given as input to the optimization solver

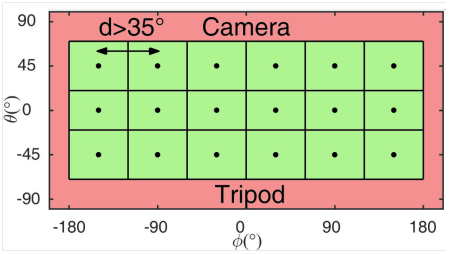

Figure 7: Regular microphone array geometry. The black lines delimit regions where individual microphones can move during the optimization. The black dots are the microphones. In red, the microphone positions unattainable and conversely in green. (Color online)

in order to compare the optimized solutions. The arbitrary selected initial microphone array geometries are

- Array 1 : Regular microphone array geometry (Figure 1.a and Figure 7)
- Array 2 : Regular microphone array geometry shifted by 10° along the azimuthal direction
- Array 3 : Regular microphone array geometry shifted by -10° along the azimuthal direction
- Array 4 : Regular microphone array geometry shifted by 20° along the elevation direction
- Array 5 : Regular microphone array geometry shifted by -20° along the elevation direction.

The cost function J^{mn} derived for the initial regular microphone array is compared with the average and min/max of the cost function derived for

the 5 optimized microphone arrays obtained by solving Eq. 12 (Figure 8). The five optimized arrays provide a similar trend with a low min/max deviation. For each optimized array, the minimum value is greater than 0.2 which means that the symmetrical aperture angle maps have been removed. Unfortunately, despite a similar trend for each optimized array the minimum value is different which means that the optimization solver fails to provide a global optimum. To investigate how these local solutions affect the results, the noise source map and the corresponding microphone locations are compared for each optimized microphone array.

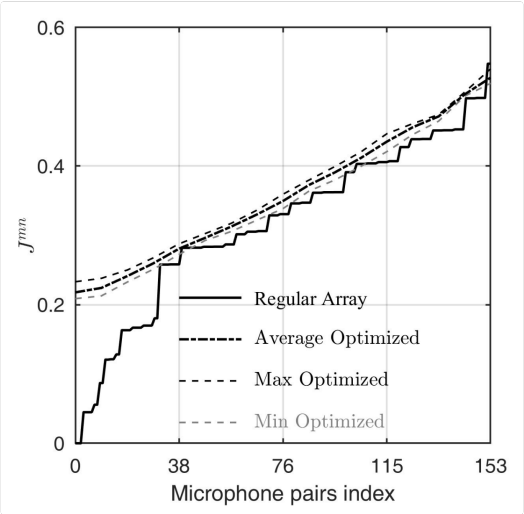

Figure 8: Cost function for the regular microphone array and for the optimized microphone arrays.

Figure 9 presents the different noise source maps obtained with the optimized array geometries and the covariance ellipses for the thresholds (-3 dB and -12 dB). The side lobes are less apparent for

the optimized microphone arrays than for the regular microphone array, the surface reduction is at least 85% for the -12 dB threshold. The side lobes at the periphery of the noise source map and surrounding the main lobe in the case of the regular array vanish for the optimized arrays. Because of the regular array symmetry, the covariance ellipse at -3 dB reduces to a circle. The optimization randomizes the microphone positions and therefore the covariance ellipses at -3 dB are no more circles with the optimized arrays which mean that the MWL is sometimes larger. However, the main lobe surface provided by the optimized arrays is at least 5% smaller. Therefore, even though different solutions are provided by the optimization solver, the side lobes are removed and the main lobe surfaces are smaller for the five optimized microphone arrays.

Slices of the noise source maps for optimized geometries are compared with those of the regular geometry in Figure 10. Again, only the average value and min/max deviation are plotted. In both directions, the side lobes are essentially removed. The MSL difference (obtained with the regular array and the average values) are -7.3 dB (for $\phi < 0$), -7.5 dB (for $\phi > 0$), -6.8 dB (for $\theta < 0$) and -5.8 dB (for $\theta > 0$). Therefore, solving the optimization problem Eq. 12 improves the noise source map by removing side lobes. In the following, the optimized array geometry obtained from the initial array, called Array 1, is considered.

To illustrate the performance of the optimized array geometry, an acoustic

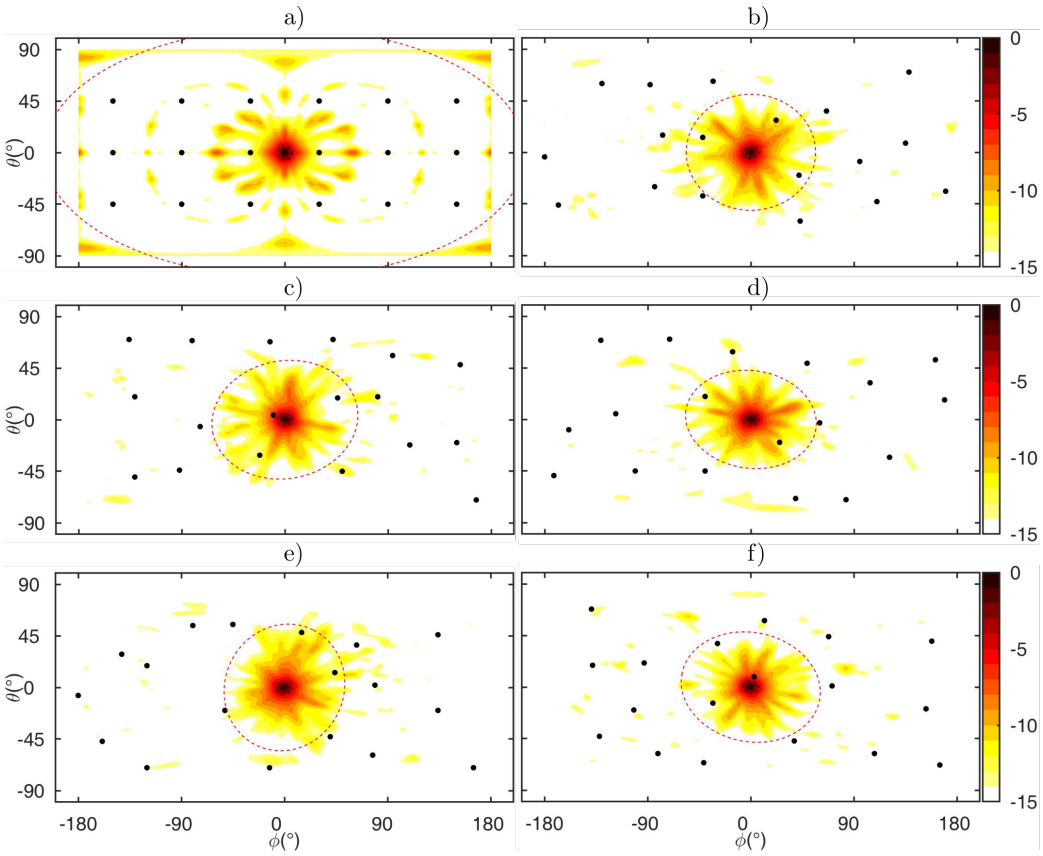

Figure 9: **Noise source map obtained with the a) regular microphone array b) optimized array with initial Array 1, c) optimized array with initial Array 2, d) optimized array with initial Array 3, e) optimized array with initial Array 4 and f) optimized array with initial Array 5. The covariance ellipses at -3 dB (black) and -12 dB (red) are displayed with dashed lines. The black dots are the microphone positions. The source generates white Gaussian noise and is located at $\phi = 0$ and $\theta = 0$. The colorbar is in dB. (Color online)**

scenario with five acoustic sources is simulated. The sources are uncorrelated white Gaussian noises and have the same amplitude. Three array geometries are compared: regular, optimized and a randomly generated array. The noise source maps obtained with the three microphone arrays are presented

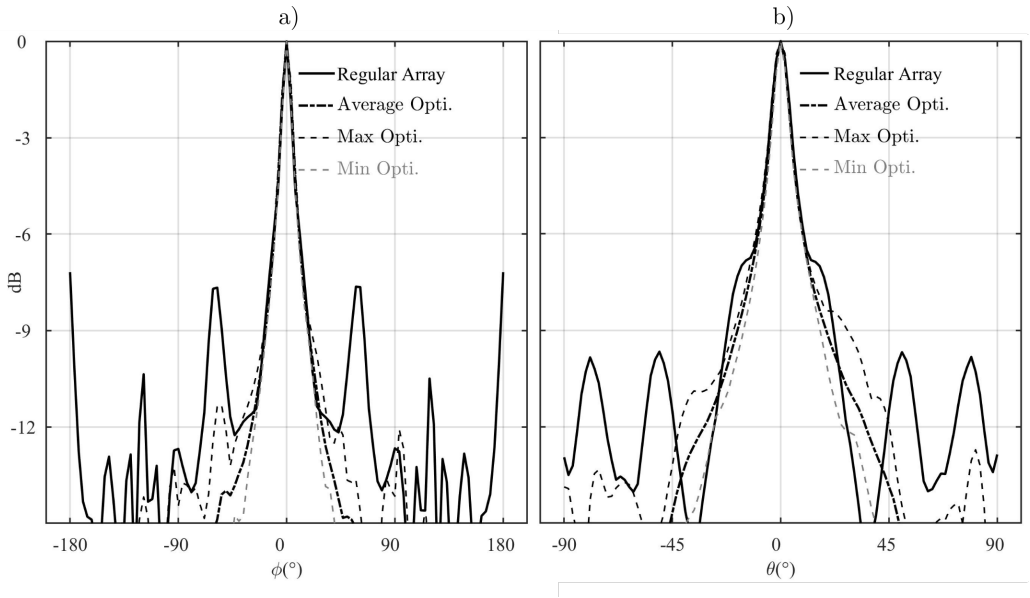

Figure 10: Slices of the noise source maps obtained with the regular array and optimized arrays along a) azimuthal and b) elevation directions.

in Figure 11. Each microphone array correctly detects the source positions. With the regular geometry, several side lobes are present. In a practical situation, these side lobes may be erroneously interpreted as real sources. A random microphone array improves the noise source map as some side lobes are removed. The minimum cost function value in this case is 0.08 which is less than half the value obtained with the optimization solver. The optimized geometry provides the most accurate noise source map with only few side lobes.

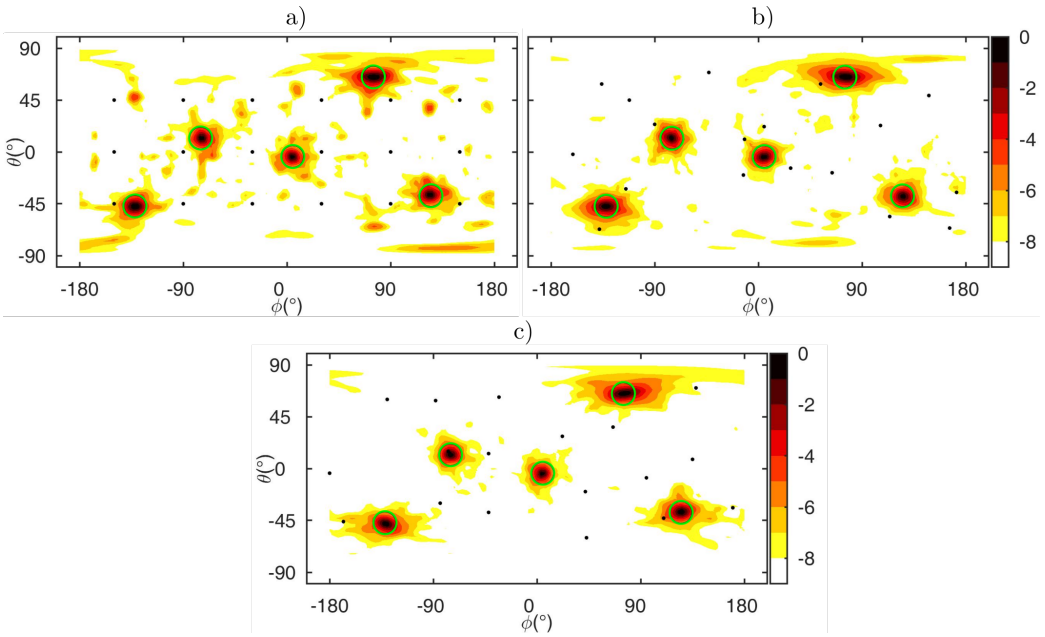

Figure 11: Noise source map obtained with the a) regular microphone array b) a random array, c) optimized array. The black dots are the microphone positions. Five sources generate uncorrelated white Gaussian noises, they are located with green circles. The colorbar is in dB. (Color online)

4. Experimental results

The goal of the experiment was to compare the noise source maps provided with the regular and the optimized array obtained with initial Array 1 (Figure 9.b). Experiments were carried out in the semi-anechoic chamber of the ICAR laboratory at École de technologie supérieure where two speakers were set in front of the spherical microphone arrays (Figure 12).

A spherical support with a 7,5 cm radius was prototyped in Acrylonitrile Butadiene Styrene (ABS) by 3D printing. Tapped holes were drilled in the spherical support to attach rods with 9,5 cm length. The microphones were

mounted at 3 cm from the end of the rods which leads to a total radius of 20 cm for the spherical microphone array (see Figure 12).

In order to test different microphone array geometries, 252 holes were drilled into the spherical support. For $\theta = 0$, holes with 10° increment were drilled along the azimuth direction. For $10^\circ \leq \theta \leq 50^\circ$ and $-10^\circ \geq \theta \geq -50^\circ$, holes with 20° increment were drilled along the azimuth direction. Finally, for $50^\circ \leq \theta \leq 70^\circ$ and $-50^\circ \geq \theta \geq -70^\circ$, holes with 40° increment were drilled along the azimuth direction. Consequently, both geometries (regular and optimized) had to be adjusted to fit onto the spherical support. However, both geometries were set onto the same spherical support which has allowed for simultaneous measurements (see Figure 12).

As the holes drilled into the spherical support were followed a regular pattern a simple shift along the azimuthal direction has allowed for adjusting the microphone positions with the regular array (for convenience a line of three microphone has been aligned at $\phi = 0$, see Figure 13.a). However, the elevation positions have slightly changed from -45° (for the numerical simulations) to -40° (for experimental measurements).

For the optimized array, the microphone positions were set to the closest positions onto the spherical support (see Figure 13.b). Consequently, the microphone positions were slightly changed for both array geometries.

The spherical support located at $1,3$ m above the floor was held by a tripod. Brüel&Kjaer microphones type 4935 were used and the acoustic signal was recorded using a Brüel&Kjaer 3038B front end and Pulse software at

the sampling rate 32768 Hz during 1 s. The source signals were uncorrelated white noises generated with a PXI-4461 card and amplified by two active speakers Eris Presonus E5. The source-array distance was set to 1.5 m.

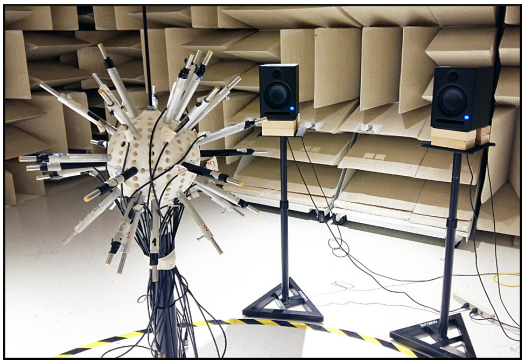

Figure 12: Picture of the spherical microphone array. (Color online)

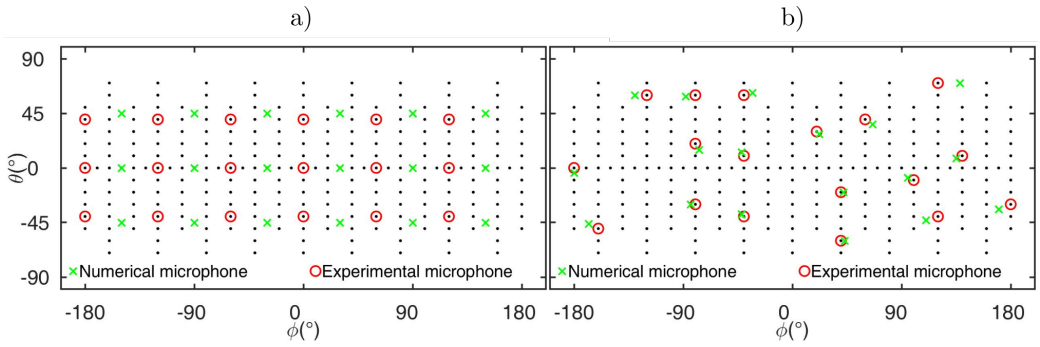

Figure 13: Numerical versus experimental microphone positions. The black dots are the positions available onto the 3D printed spherical support. (Color online)

In a first step, the source at origin was activated only to reproduce the noise source maps obtained numerically (Figure 9.a-b), the results are presented in Figure 14. With both array geometries the source position is detected. However, the regular array provides several side lobes surrounding

the main lobe whereas the optimized array removes them. Although the noise source map obtained with the experimental arrays are slightly different to numerical results of Figure 9.a-b due to slightly different microphone positions, the conclusions are the same: the optimized array provides a more accurate noise source map.

To assess the performance of the optimized array, the covariance ellipses (at -3 dB and -12 dB) provided by both array geometries are displayed in Figure 14. The optimized array allows for reducing by 66% the ellipse surface in the case of -12 dB threshold which means that some side lobes have been removed. For the main lobe surface (at -3 dB), the reduction is 14%, which means that the optimized array narrows it.

With the optimized array, a side lobe at ($\phi = 0$, $\theta = -60°$) is present below the main lobe. This side lobe does not appear onto the noise source map obtained numerically and can be attributed to a floor reflection. Indeed, if the source-image reflection theory is considered, the position of the reflected source is given by $\arctan(2 \times \text{height/source-array distance}) = \arctan(2 \times 1.3/1.5) = 60°$. In this case, the optimized array is able to detect the floor reflection whereas this lobe is merged with side lobes in the case of regular array.

In the last experiment, both speakers generated uncorrelated white noises with an amplitude smaller by 3 dB for the right speaker. The separation angle between the speakers was set to

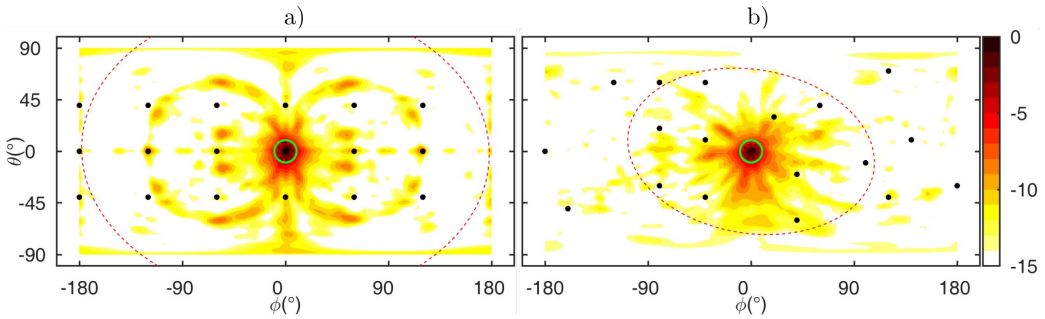

Figure 14: **Noise source maps obtained with the a) regular microphone array and b) optimized array with initial Array 1. The covariance ellipses at -3 dB (black) and -12 dB (red) are displayed with dashed lines. The black dots are the microphone positions. The source was generated white noise and was located at $\phi = 0$ and $\theta = 0$. The green circle is the source position. The colorbar is in dB. (Color online)**

$30°$ which is slightly larger than the regular array resolution (Figure 2.a-b). The noise source maps obtained with both array geometries are shown in Figure 15 together with the covariance ellipses (at -12 dB). The covariance ellipse at -3 dB is not computed because two sources with different levels are considered here. Again, the optimized array provides a better noise source map with fewer side lobes. With the regular array, the side lobes amplitude is larger and some of them could be interpreted as sources. It is especially the case at the opposite of the loudest source ($\phi = 150°$ and $\theta = 0$) where side lobes intersect due to the symmetry of the array geometry. The surface reduction is 56% for the threshold -12 dB. Again, floor reflections are visible with the optimized array whereas they are more difficult to detect with the regular array. To highlight the performance

of the optimized array, the slices along the elevation direction ($\theta = 0$) are shown in Figure 16.a. With the regular array, several side lobes with amplitudes higher than -12 dB are present whereas the optimized array limits them below -12 dB. The MSL difference are -6.8 dB (for $\phi < 0$) and -4.3 dB (for $\phi > 0$). The amplitude of the lowest source is under-estimated with both array geometries. Although the optimized array provides an amplitude closer to -3 dB the difference between both arrays in estimating the lowest source level is too small to be noticeable.

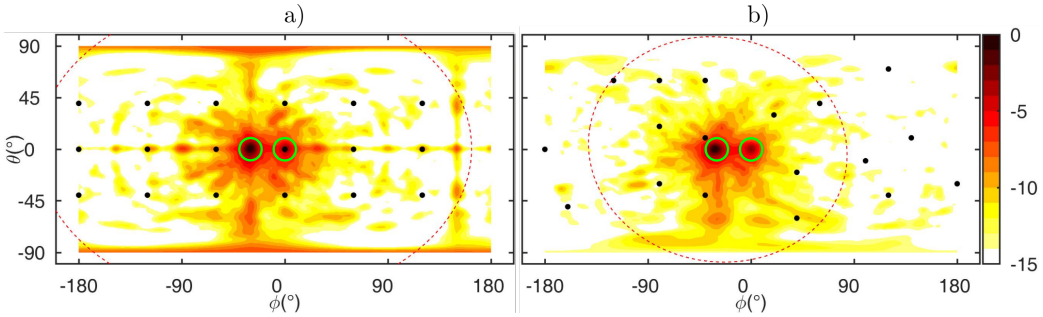

Figure 15: **Noise source maps obtained with the a) regular microphone array and b) optimized array with initial Array 1. The covariance ellipses -12 dB (red) are displayed with dashed lines. The black dots are the microphone positions. The sources were generating uncorrelated white noises with an amplitude difference of 3 dB and were located at $\phi = 0$ and $\theta = 0$ and $\phi = -30°$ and $\theta = 0$. The green circles are the source position. The colorbar is in dB. (Color online)**

5. Conclusion

Before using a microphone array-based source localization technique, the signal processing and the array geometry have to be chosen. In this work, a spherical microphone array and a source localization technique based on the

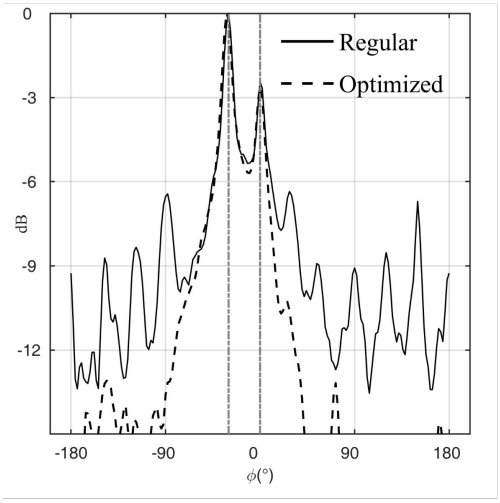

Figure 16: Slices of the noise source maps along the elevation direction. The gray dotted dashed lines are the source positions at $\phi = 0$ and $\phi = -30°$.

generalized cross-correlation of the microphone signals have been considered. The simplest microphone configuration is the regular one where the microphones have the same azimuth and elevation spacing. In this case, the noise source map obtained has many side lobes which may introduce artifacts and prevent the localization of sources with weaker amplitude.

In this work, the microphone positions of a spherical array were optimized in order to improve the source localization. A cost function independent of the source and based on the symmetry of the aperture angle map was defined. The Nonlinear Optimization by Mesh Adaptive Direct Search solver was used to maximize it. The optimized microphone array geometry allows for decreasing the side lobes amplitudes without increasing the main lobe surface.

The noise source maps obtained numerically are always more accurate than the noise source map obtained with the regular geometry.

Finally, the regular and optimized microphone geometries have been prototyped and experiments have been carried out in a semi-anechoic chamber. The experimental and numerical results are similar and confirm that the optimized microphone array improves the quality of the noise source map by removing the side lobes which allows for detecting floor reflections which is not the case with the regular array.

Acknowledgment

Authors would like to acknowledge IRSST for the financial support (No. 2015-0075).

References

[1] P. Duguay, A. Boucher and M.-A. Busque. *Statistiques sur les maladies professionnelles au Québec, Colloque IRSST Maladie professionnelles : portrait, défis et perspectives*, Novembre 2014.

[2] R. Mackenzie, L.-A. Boudreault and M. Pearson. *Validation of a sound intensity imaging system for wall ISTC calculation, with leak detection*, Vol 45(4), Proceedings of the Acoustics Week in Canada, October 11-13, Guelph, 2017.

[3] D.H.Johnson and D.E.Dudgeon. *Array Signal Processing : Concepts and Techniques*, Prentice Hall, Upper Saddle River, New Jersey, USA, Chap. 4, p.111-198, (ISBN-10:0130485136,ISBN-13:978-0130485137), 1993.

[4] J.D. Maynard, E.G. Williams and Y. Lee. *Nearfield acoustic holography: I. Theory of generalized holography and the development of NAH*, J. Acoust. Soc. Am. Vol 78(4), p.1395-1413, 1985.

[5] C. Camier, J. Provencher, T. Padois, P-A. Gauthier, A. Berry, J-F. Blais and R. Lapointe. *Fly-over source localization on civil aircrafts*, 9th AIAA/CEAS Aeroacoustics Conference (AIAA 2013-2212) Berlin, Germany, May, 2013.

[6] T. Padois and A. Berry. *Application of acoustic imaging techniques on snowmobile pass-by noise*, J. Acoust. Soc. Am., Vol 141(2), p.EL134-EL139, 2017.

[7] M. Fedia, J. Antoni, F. Taher, V. Lanfranchi and H. Mohamed. *Wavelet-based non-stationary near-field acoustical holography*, Applied Acoustics Vol 74, p.12261233, 2013.

[8] P. Eret and C. Meskell. *Microphone Arrays as a Leakage Detection Tool in Industrial Compressed Air Systems*, Advances in Acoustics and Vibration, p.1-10, 2012.

[9] L. Zhiwei, F. Tao, H. Zhigang, S. Bin, L. Ran, L. Yajun and W. Yao.

549 *The Noise Source Localization of Industrial Sewing Machine by NAH*
550 *Method*, Applied Mechanics and Materials, Vol 300-301, p.898-902.

551 [10] M. R. Bai and C.-C. Chen. *Farfield and nearfield source identification for*
552 *machine tools using Microphone array imaging systems*, Procedia Engi-
553 neering Vol 79, 37th National Conference on Theoretical and Applied
554 Mechanics (37th NCTAM 2013) & The 1st International Conference on
555 Mechanics (1st ICM), p.345-354, 2014.

556 [11] C. Noël, V. Planeau and D. Habault. *A new temporal method for the*
557 *identification of source directions in a reverberant hall*, Journal of Sound
558 and Vibration, Vol 296(3), p.518-538, 2006.

559 [12] T. Padois, O. Doutres, F. Sgard and A. Berry. *Time domain localization*
560 *technique with sparsity constraint for imaging acoustic sources*, Mechan-
561 ical System and Signal Processing, Vol 94, p.85-93, 2017.

562 [13] T. Padois, F. Sgard, O. Doutres and A. Berry. *Acoustic source localiza-*
563 *tion using a polyhedral microphone array and an improved generalized*
564 *cross-correlation technique*, Journal of Sound and Vibration Vol 386(6),
565 p.82-99, 2017.

566 [14] T. Padois, O. Doutres, F. Sgard and A. Berry. *On the use of geometric*
567 *and harmonic means with the generalized cross-correlation in the time*
568 *domain to improve noise source maps*, J. Acoust. Soc. Am. Vol 140(1),
569 p.EL56-EL61, 2016.

[15] T. Padois. *Acoustic source localization based on the generalized cross-correlation and the generalized mean with few microphones*, J. Acoust. Soc. Am., accepted, 2018.

[16] F. Le Courtois, J.-H. Thomas, F. Poisson and J.-C. Pascal. *Genetic optimisation of a plane array geometry for beamforming. Application to source localisation in a high speed train*, Journal of Sound and Vibration, Vol 371, p.78-93, 2016.

[17] M. Bjelić, M. Stanojević, D. S. Pavlović and M. Mijić. *Microphone array geometry optimization for traffic noise analysis (L)*, J. Acoust. Soc. Am. Vol 141(5), p.3101-3104, 2017.

[18] J. R. Underbrink. *Circularly symmetric, zero redundancy, planar array having broad frequency range applications*, US Patent No 6,205,224 B1, 2001.

[19] T. J. Mueller, *Aeroacoustic measurements*, Springer-Verlag Berlin Heidelberg, Chap 3, p.98-215, 2002.

[20] E.J.G. Arcondoulis, C.J. Doolan, L.A. Brooks and A.C. Zander, *A Modification to Logarithmic Spiral Beamforming Arrays for Aeroacoustic Applications*, 17th AIAA/CEAS Aeroacoustics Conference (32nd AIAA Aeroacoustics Conference), (AIAA 2011-2720), Portland, Oregon, June 2011.

[21] B. Rafaely, *Fundamentals of Spherical Array Processing*, Springer-Verlag Berlin Heidelberg, Chap 3, 2015.

[22] C. Knapp and G. C. Carter, *The generalized correlation method for estimation of time delay*, Transactions on Acoustics, Speech and Signal Processing, IEEE, Vol.24(4), p. 320-327, 1976.

[23] N. Quaegebeur, T. Padois, P-A. Gauthier and P. Masson, *Enhancement of time-domain acoustic imaging based on generalized cross-correlation and spatial weighting*, Mechanical System and Signal Processing, Vol. 75, p.512-524, 2015.

[24] T. Padois, O. Doutres and F. Sgard, *On the use of modified phase transform weighting functions for acoustic imaging with the generalized cross correlation*, J. Acoust. Soc. Am. Vol 145(3), p.1546-1555, 2019.

[25] E. Arcondoulis and Y. Liu, *An iterative microphone removal method for acoustic beamforming array design*, Journal of Sound and Vibration, Vol 442, p.552-571, 2019.

[26] R. A. Johnson and D. W. Wichern, *Applied multivariate statistical analysis (6th Ed.*, Upper Saddle River, NJ : Pearson Prentice Hall, Chap 2.3, p.60-65, 2007.

[27] S. Le Digabel. *Algorithm 909: NOMAD: Nonlinear Optimization with the MADS Algorithm*, ACM Transactions on Mathematical Software, Vol.37(4), p. 1-15, 2011.

[28] S. Le Digabel, C. Tribes and C. Audet *NOMAD User Guide Version 3.8.1*, Technical Report G-2009-37, Les cahiers du GERAD, 2009.