



# OPTIMIZATION METHODOLOGY OF MICROPHONE ARRAYS FOR ENVIRONMENTAL SOURCE LOCALIZATION USING THE GENERALIZED CROSS-CORRELATION

Lucas Carneiro\*    Alain Berry<sup>†</sup>    Thomas Padois<sup>‡</sup>

## Abstract

Microphone array design is a discipline of great interest for the scientific community and end-users because its properties have a big impact on source localization performance. This publication addresses the design methodology and geometry optimization of spherical microphone arrays for environmental source localization using time-domain beamforming algorithm, more precisely the Generalized Cross-Correlation (GCC). In the proposed methodology, a set of conceptual decisions such as array design, frequency range and number of microphones take the form of non-linear constraints in a single or multi-objective genetic algorithm for the optimization of the microphones positioning. The fitness criteria are based on the maximization of the sum of all available differences of distances, or spatial orientations, or both, between pairs of microphones. These criteria are empirically supported by the geometric intrinsic nature of the Spatial Likelihood Functions (SLFs), the constitutional functions of the GCC: indeed, under certain conditions, the geometric properties of a SLF hyperbola can be controlled by the distance and the relative spatial orientation between the associated pair of microphones. These properties can be used to improve rejection of side-lobes in the noise source map. The experiments are based on numerical simulations: first, a statistical analysis of the optimization is performed to demonstrate its repeatability. It was proved that the single-objective criteria are effective to induce side-lobe rejection and similar in the design space (they converge to a same region in the Pareto front). Latter, a scenario with distributed broadband sources is investigated. Compared to a regular microphone arrangement, the results show that an optimized array generates a smoother cartography and better source localization.

---

\*GAUS, Department of Mechanical Eng., Université de Sherbrooke, Sherbrooke QC J1K2R1 Canada, lucas.carneiro@usherbrooke.ca

<sup>†</sup>GAUS, Department of Mechanical Eng, Université de Sherbrooke, Sherbrooke QC J1K2R1 Canada

<sup>‡</sup>École de technologie supérieure, Montreal, QC H3C1K3 Canada

## 1 INTRODUCTION

In engineering science, noise control and attenuation aims the conception of human-friendly buildings, vehicles and machines from the acoustic point of view . The noise abatement usually starts with a first study requiring sources localization and its hierarchical classification. Currently, the popular techniques used for such a study are [1]:

- Noise assessment with sound level meters: the device measures the sound pressure level for a given point but does not provide the localization of the source;
- Source localization with sound intensity probe: two microphones are used to measure the pressure and acoustic velocity. Based on these quantities an intensity vector can be computed, the sound direction obtained and a source map constructed from the measurement. The technique requires multiple measurements in the near-field, which in the most part of the cases are only feasible for small machines;
- Source localization using microphone array and beamforming: the technique consists on the calculation of the sound direction from the time-lag of signals of two or more microphones. Compared to previous techniques, beamforming requires more hardware and more computational power to run. The technique is interesting because the measurement does not require multiple captures.

The source detection capability of microphone antennas is directly related to the available hardware: increasing the number of microphones may offer better source localization at the expense of a higher computational cost while improving the microphones distribution may provide better source localization without additional computational cost. Different array geometries can be used depending on the situation: while planar arrays are directive and recommended for source localization in pass-by vehicles, spherical arrays are deployed in environmental source localization.

The objective of this paper is to propose an optimization methodology of spherical arrays geometry for environmental source localization based on beamforming algorithm. The source localization technique used is the Generalized Cross-Correlation (GCC). Noël et al. [2] and Padois et al. [3] have described the implementation of the GCC algorithm for the proposed application.

The GCC technique was proposed by Knapp & Carter [4] as a maximum likelihood estimator for the time-lag of two signals. The correlation is obtained from the signals in the frequency domain. This frequency domain approach allows the insertion of a frequency weighting filter in the cross-correlation for a better time-lag estimation. Among the most important filters, the Phase Alignment Transformation (PHAT) normalizes, the microphone signal amplitude information and keeps the phase information. GCC-PHAT has been successfully used in acoustic source localization and has shown good results in the presence of reverberation [5]. More recently, it was mathematically proven that PHAT is equivalent to the maximum likelihood estimator in cases where the signal-to-noise ratio was low [6].

Since Knapp & Carter [4], the GCC has been extensively used and improved for source localization using microphone arrays of different shapes. It was shown that the noise source map is the result of the summation of the cross-correlations between individual pairs of microphones in the array. The geometric nature of each term in this summation, also called Spatial Likelihood

Function (SLF), was studied in the last two decades by Dibiase [7], Aarabi [8] and Velasco et al. [9]. Padois et al. [10] studied the mechanism of side-lobe rejection and main-lobe narrowing obtained in noise sources maps computed using geometric and harmonic means on the GCC summation. Salvati et al. [11] proposed an adaptative scanning domain that also improves source localization performance. From the technological perspective, state-of-the-art devices composed of MEMS microphones and FPGA boards for digital real-time processing have been implemented in numerous applications such as source localization in robotics [12].

Array design aspects such as the influence of array positioning and microphone distribution on the final source localization have also been investigated. For example, Rafaely [13] describes the formulations for source localization from the Fourier transform decomposition in the spherical harmonics domain and showed how this theory may be used to optimize the spherical array beam pattern and other performance parameters. Hu et al. [14] used a minimum variance estimator of the time-lag to design an array that guarantees an omnidirectional performance for source localization. Evolutionary optimization algorithms [15] or convex optimization [16] have been used for optimal geometry of planar arrays for near-field and far-field imaging enhancement as a function of beamforming beam-pattern cost functions.

In this work, specific fitness criteria for spherical array geometry optimization used with the GCC are proposed and genetic algorithms ([17], [18]) are used to solve the optimization problem. The criteria are designed to reject side-lobes generated from the GCC summation.

This article is divided in the following sections: Section 2 explains the GCC technique in the context of localization of environmental sources, the mathematical and geometrical locus of the SLFs, how SLFs and microphones can be positioned for side-lobe rejection and how geometric criteria can be built from these concepts. Section 3 details the optimization methodology. Section 4 is dedicated to the numerical results: a statistical analysis of the optimization is performed and the performances of the optimized arrays are discussed.

## 2 Time domain source localization technique

### 2.1 The Generalized Cross-Correlation

Consider an array of microphones distributed in space and subjected to sound waves. The summation of the microphone signals in the array can be expressed in the form of a likelihood estimator combining the delay of each pair of microphones  $(m,n)$  as a cross-correlation function [4]:

$$R_{x_m, x_n}(\tau) = E \{x_m(t)x_n(t + \tau)\}, \quad (1)$$

where  $x_m$  and  $x_n$  are the microphone signals. The argument  $\tau$  that maximizes this estimator provides the estimation of time-lag for this pair of microphones. It is considered a scanning position  $\mathbf{r}_k$  where the source is searched. The time-lag  $\tau = \Delta t_{mk} - \Delta t_{nk}$ , where  $\Delta t_{ik}$  is the time delay between a scan point in  $\mathbf{r}_k$  and the microphone  $i$ , may be intuitively understood as the time-lag between the acoustic signals when they reach the microphones pair  $(m,n)$ . Because signals are a finite observation of time, the cross-correlation can only be estimated and  $E$  denotes expectation. The beamformer output may be expressed as an output power  $Y_{\mathbf{r}_k}^{BF}$ , the summation of all possible cross-correlations between the unique pairs obtained from  $M$  microphones in the

array:

$$Y_{\mathbf{r}_k}^{BF} = \sum_{m=1}^M \sum_{n=m+1}^M R_{x_m, x_n}(\Delta t_{mk} - \Delta t_{nk}). \quad (2)$$

The estimated source positions will be provided by a set of  $\tau_{m,n} = [|\Delta t_{m1} - \Delta t_{n1}|, \dots, |\Delta t_{mk} - \Delta t_{nk}|]$  that maximizes the Equation 2. Note that  $\tau_{m,n} = \tau_{n,m}$ , therefore pairs are computed only once.

In order to improve the accuracy of the estimator, a frequency weighting filter,  $\Psi(f)$ , can be introduced in the cross-correlation computation. Equation 1 can be expressed as the following inverse discrete Fourier transform:

$$R_{x_m, x_n}(\tau) = \sum_{f=0}^{N_f-1} \Psi(f) C_{X_m, X_n}(f) e^{\frac{j2\pi f\tau}{N_f}}, \quad (3)$$

where  $C_{x_m, x_n}$  is the power cross-spectrum of microphone signals  $x_m$  and  $x_n$  and  $f$  the frequency index comprised between 0 and  $N_f - 1$ . PHAT is the most frequently used filter in source localization, the inverse of the absolute value of  $C_{X_m, X_n}$ :

$$\Psi(f) = \frac{1}{|C_{X_m, X_n}|} = \frac{1}{|X_m(f)X_n(f)^*|}. \quad (4)$$

The computation of Equation 3 presents some constraints:  $R_{x_m, x_n}$  is a function of the observation time  $t$  and might present time-lags  $\tau$  surpassing the maximum physically allowable time-lag  $\max(\tau_{m,n})$  between a pair of microphones in the array and the scanning zone. Henceforth the cross-correlation needs to be literally computed as in Equation 3, truncated in  $\tau \leq \max(\tau_{m,n})$  and interpolated in  $\tau_{m,n}$ , respecting the physical constraints of the problem [3]. A truncated cross-correlation  $\check{R}_{x_m, x_n}(\tau_{m,n})$  is obtained for each pair of microphones and is also called Spatial Likelihood Function (SLF) [19].

For each scanning position  $\mathbf{r}_k$ , the summation of all SLFs available in an array (as in Equation 2) creates the corresponding noise source map. In the case of spherical arrays and environmental source localization the scanning space is usually spherical. Hence, the scanning positions  $\mathbf{r}_k$  are obtained from a discretization of the scanning space in coarse grid elements of size  $\Delta\phi - \Delta\theta$  (azimuth and elevation), projected in a distance equivalent to the separation between the sources and the array, if the waves are spherical. In this paper, it will be supposed the hypothesis of sources in the far-field and planar waves. A scanning position  $\mathbf{r}_k$  may be more precisely understood as a scanning direction in the direction  $\phi - \theta$  because the distance of the projection to the array is irrelevant.

## 2.2 The Spatial-Likelihood Function

As long as the beamformer output  $Y_{\mathbf{r}_k}^{BF}$  is a function of the scanning position  $\mathbf{r}_k$ , the mathematical definition of the SLF can be strictly related to  $\check{R}_{x_m, x_n}(\tau_{m,n}(\mathbf{r}_k))$ . Two geometrical spaces are defined:  $\Omega$ , the space of all scanning positions  $\mathbf{r}_k$  belonging to the scanning zone and  $\mathbf{H}$ , the space of positions  $\mathbf{r}_h$  satisfying the condition  $\tau_{m,n}(\mathbf{r}_h) = \tau_{m,n}(\mathbf{r}_s)$ ,  $S$  being a real source. The geometric locus of the SLF belongs to that space  $\mathbf{H}$ , according to the first argument of the

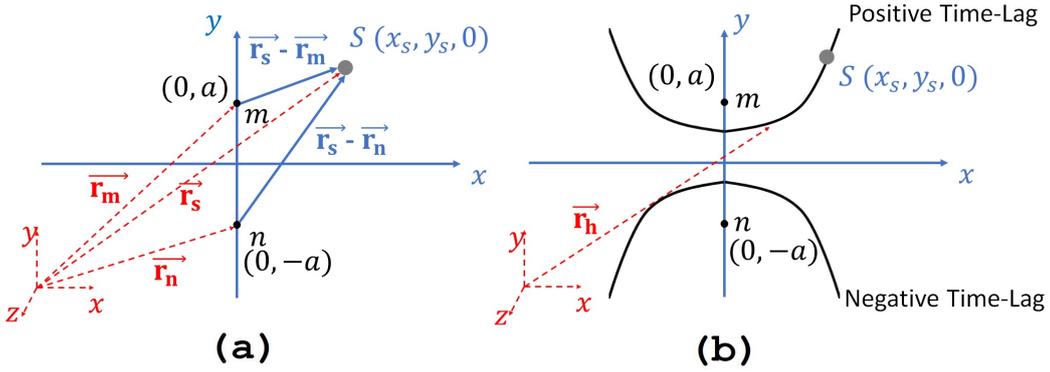


Figure 1: (a) The global coordinate system is defined in dashed-red while the local coordinate system related to the pair of microphones  $(m, n)$  (separated by a distance  $2a$ ) is in solid-blue.  $\mathbf{r}_m$ ,  $\mathbf{r}_n$  and  $\mathbf{r}_s$  are the coordinates of microphones  $(m, n)$  and the source, respectively. (b) Symmetric hyperbolas are obtained from the possible combinations of pairs of microphones, which generates opposite time-lag values.  $\mathbf{r}_h$  is the coordinate vector representing points belonging to the hyperbola.

following relationship [9]:

$$\check{R}_{x_m, x_n}(\tau_{m, n}(\mathbf{r}_k)) = \begin{cases} \check{R}_{x_m, x_n}(\tau_{m, n}(\mathbf{r}_S)) & \forall \mathbf{r}_k \in \mathbf{H} \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

Considering the time-lag between acoustic signals,  $\mathbf{r}_h$  in space  $\mathbf{H}$  will satisfy the following condition:

$$\tau_{m, n}(\mathbf{r}_h) = (\|\mathbf{r}_m - \mathbf{r}_h\| - \|\mathbf{r}_n - \mathbf{r}_h\|) / c_0, \quad (6)$$

where  $\mathbf{r}_m$  and  $\mathbf{r}_n$  are the positions of the pair of microphones  $(m, n)$  and  $c_0$  the speed of sound.

In order to study the locus of the SLF, it is appropriate to consider a local coordinate system centered between a pair of microphones. Cartesian coordinates are considered with microphones located at  $y = -a$  and  $y = +a$ , as in Figure 1.a. The time-lag between signals allow the following equation, the ensemble of points  $(x, y)$  belonging to  $\mathbf{r}_h$ , with a time-lag equal to  $\tau_{m, n}(\mathbf{r}_S)$ , allows the following equation to be written [8]:

$$\frac{4y^2}{(c_0 \tau_{m, n}(\mathbf{r}_S))^2} - \frac{4x^2}{4a^2 - (c_0 \tau_{m, n}(\mathbf{r}_S))^2} = 1. \quad (7)$$

Therefore the locus of positions corresponding to a given time-lag  $\tau_{m, n}(\mathbf{r}_S)$  is an hyperbola in the plane  $x - y$  of the type [20]:

$$\frac{y^2}{A^2} - \frac{x^2}{C^2} = 1. \quad (8)$$

The space  $\mathbf{H}$  satisfying Equation 6 is not restricted to a plane. The 3D locus of the SLF is obtained by the rotation of the hyperbola 7 around the  $y$ -axis defining an hyperboloid. Nonetheless, as long as the source maps are two-dimensional, the two dimensional representation (Figure 1.b) is sufficient.

### 2.3 Geometric Criteria

The proposed fitness criteria for array geometry optimization consist in positioning pairs of microphones in order to reduce the side-lobe generation. Considering the application to free field source localization, the detailed comprehension of the mathematical and geometric nature of the SLFs allows us to coin fitness criteria and design constraints to achieve this goal. The following assumptions are made:

- The sources are considered to be static and behave as perfect monopoles;
- The sources are exterior to the array and in the far field. Under these conditions, the time-lag  $\tau_{m,n}$  is purely a function of the distance between the microphones and the relative orientation of the pair to the source.

The fitness criteria induce a mechanism called ‘spatial whitening’. In this article, ‘spatial whitening’ consists in omnidirectionally orient the SLFs around the point of origin of a source, as described in Figure 2 and in Figure 3.

Three cases are presented: a noise source map obtained from aligned pairs of microphones of same separation (Figure 2.a,b), two orthogonal pairs of same separation (Figure 2.c,d) and orthogonal pairs with different separations (Figure 3). The left part of the figures shows an arrangement of 4 microphones and a source. The hyperbolic shapes of the SLFs are represented in the plan for the assigned pairs of microphones and intersect the spherical scanning space. The right part of the figures illustrates the projection of the tridimensional SLFs (the hyperboloids) projected in the spherical scanning space with azimuth-elevation angles  $(\phi - \theta)$ . The projection is equivalent to the noise source map.

While in Figure 2.a,b the pairs of microphones (1,2) and (3,4) are aligned, causing an unfavorable SLF summation, in Figure 2.c and d those same pairs are orthogonal. It can be observed in Figure 2.d that the respective SLFs also intersect in a pin-location except for SLFs (1,4) - (2,3) and (1,3) - (2,4), which are superimposed. Those SLFs belong to the parallel pairs of microphones in the array. In the case of Figure 3, one of the microphones is moved, increasing its separation from the other microphones. This results in a loss of parallelism between microphones and separates SLF (1,4) from (2,3) and (1,3) from (2,4). It can be concluded that the reduction of the array redundancy in terms of microphones spacing and orientation induces ‘spatial whitening’ near the source.

Therefore, a first possibility to induce ‘spatial whitening’ is by positioning pairs of microphones in such a way that, for a given global coordinate system, the orientations of the pairs of microphones relative to the global system will be as different as possible. This procedure should be equivalent to maximizing the sum of all available differences of orientation between pairs of microphones, as follows:

$$[\mathbf{e}_1, \dots, \mathbf{e}_{C_p}] \text{ where } C_p = \binom{M}{2} \Rightarrow FF_{orient} = \operatorname{argmax} \left( \sum_{c_p=1}^{C_p} \Theta(\mathbf{e}_{c_p}, \mathbf{e}'_{c_p}) \right). \quad (9)$$

$M$  is the total number of available microphones,  $C_p$  is the number of pairs of microphones,  $\mathbf{e}_{c_p}$  is the unitary orientation vector for a pair of microphones  $(m,n)$ :

$$\mathbf{e}_{c_p} = \frac{\mathbf{r}_m - \mathbf{r}_n}{\|\mathbf{r}_m - \mathbf{r}_n\|}. \quad (10)$$

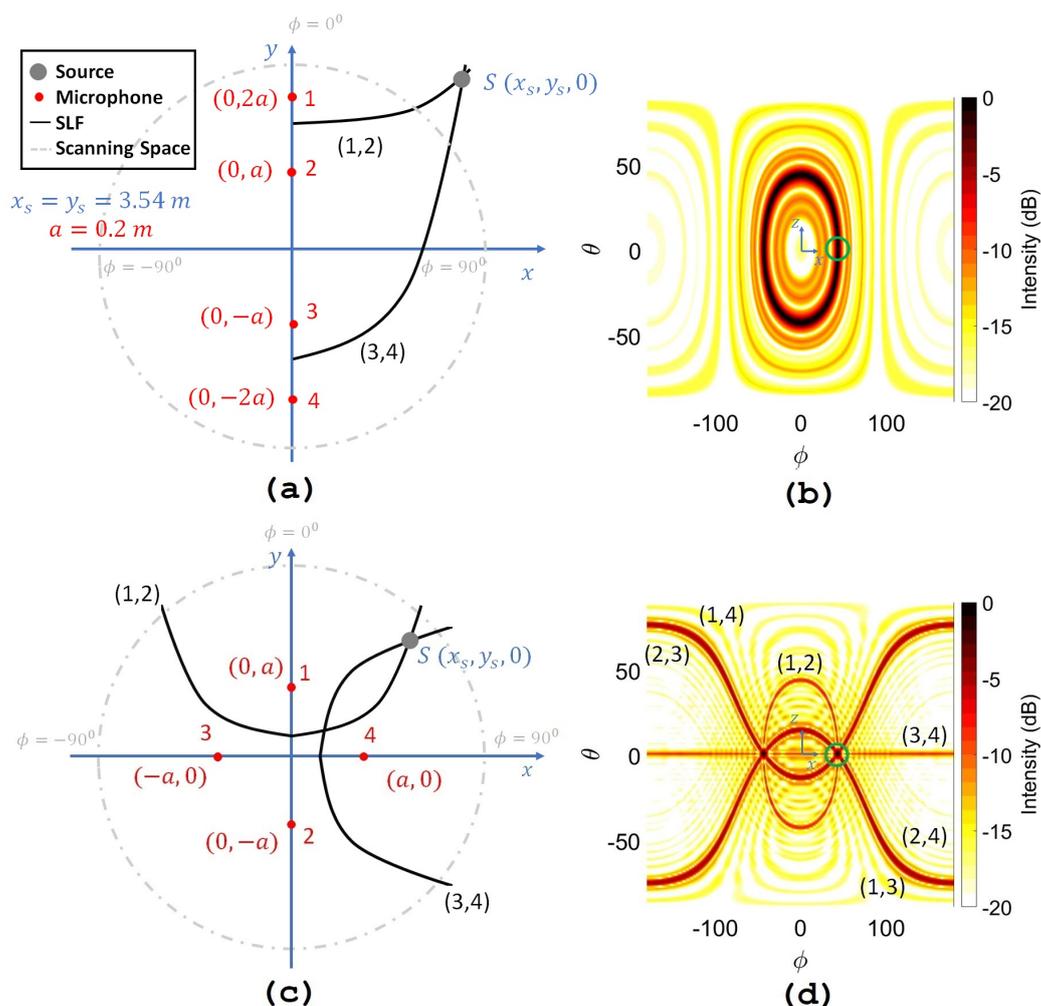


Figure 2: (a) Planar arrangement of aligned pairs of microphones. For convenience, only 2 of the 6 SLFs are represented. (b) Representation of the summation of all 6 possible SLFs in spherical coordinates. It is observed that all SLFs have the same y-axis symmetry and the source localization is impossible (the source is identified by a green circle). (c) Planar arrangement of orthogonal pairs of microphones. (d) Representation of the summation of all 6 possible SLFs in spherical coordinates. It is observed that most SLFs intersect orthogonally except for SLFs (1,4) - (2,4) and (1,3) - (2,4) (superimposed).

$\mathbf{e}'_{c_p}$  is the vector with better alignment among all microphone pairs to  $\mathbf{e}_{c_p}$ .  $\Theta$  is the angle between  $\mathbf{e}_{c_p}$  and  $\mathbf{e}'_{c_p}$  and is calculated from a scalar product.

Another optimization criterion is maximizing the sum of all available differences of microphone separations among all possible pairs, as follows:

$$[d_1, \dots, d_{C_p}] \text{ where } C_p = \binom{M}{2} \Rightarrow FF_{dist} = \operatorname{argmax} \left( \sum_{c_p=1}^{C_p} |d_{c_p} - d'_{c_p}| \right), \quad (11)$$

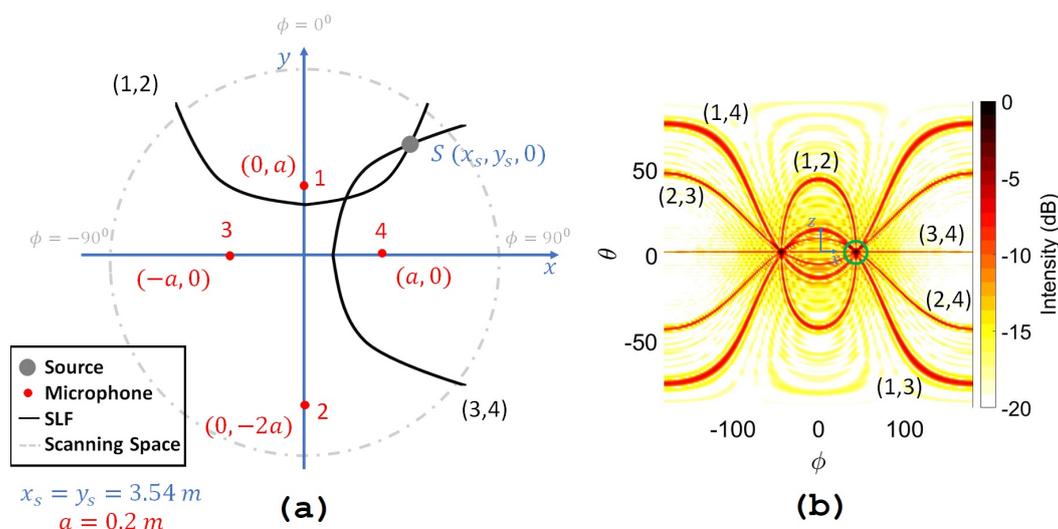


Figure 3: (a) Planar arrangement of proposed pairs of microphones. For convenience, only 2 of the 6 SLFs are represented. (b) Representation of the summation of all 6 possible SLFs in spherical coordinates. SLFs (1,4) - (2,4) and (1,4) - (2,4) are no longer merged (the actual source is identified by a green circle).

where  $d_{c_p}$  is the separation (or the  $l^2$ -norm) of a pair of microphones  $(m,n)$ :  $\|\mathbf{r}_m - \mathbf{r}_n\|$ .  $d'_{c_p}$  is the closest separation to  $d_{c_p}$  among all other pairs of microphones.

The difference of distances or orientations are the values used to compute the fitness function, obtained from a summation over all pairs of microphones in the array.

### 3 Optimization solver : Genetic algorithm

Equations 9 and 11 are non-linear with respect to optimization variables and do not have closed-form solutions. Additionally, the number of optimization variables (coordinates of the microphones in the array) is twice or three times the total number of microphones. The search space is hence multidimensional and increases rapidly with the number of microphones.

Several global optimization techniques have been proposed to solve such problems. In the last decades a large number of methods based on Genetic Algorithms (GA) have been proposed, with successful applications on engineering problems of many types [18]. GA is inspired from evolutionary principles where a certain initial solution population is distributed on a search space. The individuals possess intrinsic characteristics named genes. Each individual is first ranked according to its 'fitness' with respect to the optimization criterion and a directly proportional reproduction factor is attributed. A crossover is then realized between individuals with larger reproduction factors. During a crossover, individuals share genes in order to produce a new population. Nonetheless the reproduction and crossover steps may be carefully controlled using mutation and non-dominant mechanisms, keeping a stable diversity and the optimal exploitation of the search space.

Like other evolutionary algorithms, GA allows solving multi-objective problems either with

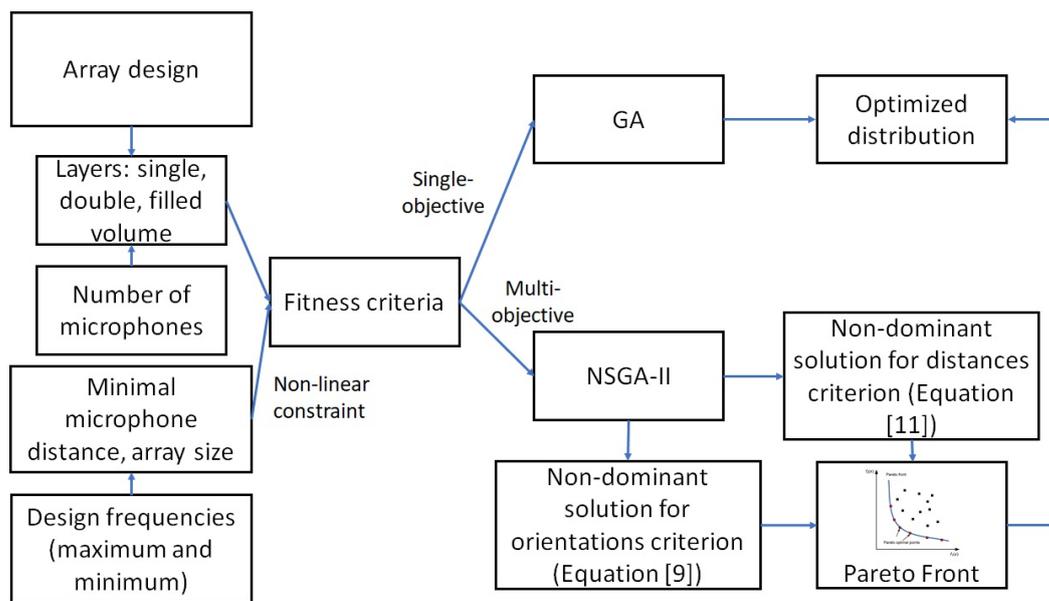


Figure 4: Block diagram of the proposed optimization methodology. NSGA stands for Non-dominated Sorting Genetic Algorithm [18], the algorithm implemented in MATLAB toolbox.

aggregation and non-aggregation methods<sup>1</sup>. Nonetheless, GA tends to be more robust for a larger variety of problems. Computational time not being a constraint for the optimization to be performed, GA has been designated as the paradigm of choice.

The proposed approach consists in a series of scripts written in MATLAB and compatible with MATLAB's Global Optimization Toolbox, which includes GA for both single and multi-objective optimization. As input parameters the routine accepts:

- The number of microphones in the array;
- The array design: the microphones may be organized in a single or double layer distribution, meaning that the microphones will be distributed on concentric surfaces. Otherwise a more general distribution is accepted, where the microphones are distributed in-between the layers;
- Design frequencies: the low-frequency limit of the array determines its size. Two microphones are enforced to be installed at a distance equal to the diameter. Wavelengths smaller than the smallest wavelength of interest (high-frequency limit) are not important. Its size determines the shortest possible separation between microphones. These conditions are set as a non-linear constraint on the fitness criteria.

Figure 4 describes the block diagram representing the optimization. As GA is a paradigm for function minimization, instead of maximizing the non-negative Equations 9 and 11, the scripts

<sup>1</sup>Aggregation: once dominant solutions are merged into one solution, hence a non-dominant solution. Dominant solutions are solutions for which the improvement of one fitness function can not be achieved without the degradation of the other.

seek to minimize the opposite of those. In this paper, for convenience, the absolute value of the criterion will be assumed for discussion.

In a single-objective optimization the optimized distribution is obtained immediately after the convergence. In non-aggregative multi-objective optimization, dominant (rank 1) solutions are used to construct a tradeoff surface (commonly known as the Pareto front) at each iteration: the final solution can be finally obtained from a compromise between both fitness criteria (Equations 9 and 11). While compromising design objectives is a tricky task for some optimization problems, it was observed that the Pareto front of the optimized array is often convex, justifying the choice of the non-aggregative GA method in this work.

## 4 NUMERICAL RESULTS

### 4.1 Array optimization

In order to validate the optimization routine and the assumptions of this work, a spherical single-layer microphone array is considered as reference. The array is 0.4 m in diameter and has 25 microphones regularly spaced at 5 azimuth and elevation directions, as seen from the microphone distribution in Figure 8.a.

The same array was optimized according to the criteria of Equations 9 and 11 respectively named ‘Orientations Criterion’ and ‘Distances Criterion’ and also for the ‘Multi-Objective Criterion’ considering the simultaneous optimization of both criteria. A non-linear-constraint defining a minimal distance of 0.04 m between microphones was imposed, as wavelengths smaller than this value are not of interest. An initial population at least 8 times larger than the number of variables was used, securing a sufficient diversity along the optimization. The initial solution-set was composed of randomly distributed arrays and the reference array. The stopping criterion was the relative change in the fitness function over generations, known as ‘stall generations’. Hence, no limitation was imposed in terms of computational time. Figure 8.b to d show the microphone distribution of each optimization case.

Trial optimizations were performed to verify the convergence of the fitness criteria to target values with acceptable deviation (Figure 5) and the best obtained results were chosen for analyze. It was observed that all criteria do not result in an unique geometrical solution: in general the microphone distributions vary from one trial optimization to another for the same criterion. However the values of the fitness criteria after convergence remain relatively constant. Also, it can be seen on Figure 5.a,b that the optimal solution in terms of microphone pair orientation is slightly sub-optimal in terms of microphone pair separation, and vice-versa. Compared to optimal solutions for either separation or orientation, the regular array shows poor values of the fitness function.

For the presented cases, Figure 6.a depicts in ascending order the separation and the difference of separation of a given pair of microphones to the pair of nearest separation. The summation of the differences of separation over all available pairs of microphones results in the ‘Distances Criterion’ of the array. In the case of a regular spaced array, it is pertinent to note that the separation curve follows a step-shape: for a given pair of microphones, a similar pair with same separation is always available and results in a ‘Distances Criterion’ of 0. For the same array, microphones pairs of small separation (between 0 – 0.1 m, 0.12 – 0.2 m and 0.22 – 0.3 m) are rare while, for any of the proposed optimized arrays, a broad distribution of separations

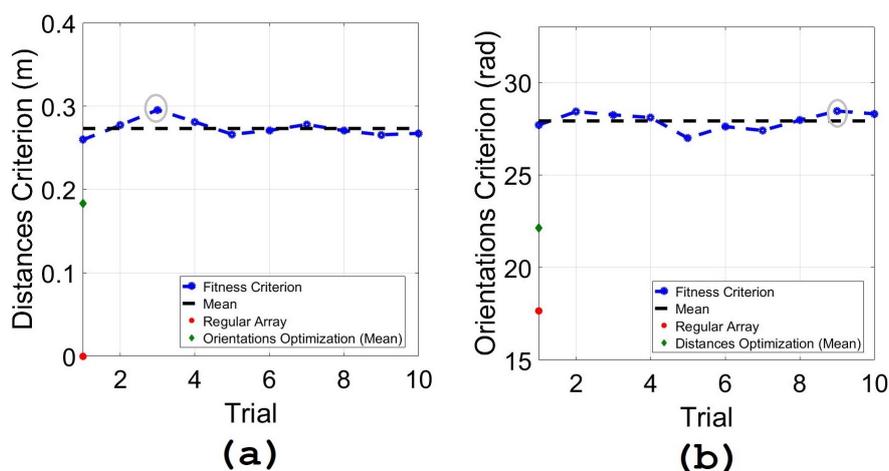


Figure 5: Repeatability analysis performed with a set of trial optimizations. (a) Distances Optimization. (b) Orientation Optimization. For each case, the best trial obtained is identified with an ellipse.

between 0.04 – 0.4 m is noticed.

In a similar fashion Figure 6.b depicts the orientation relative to an arbitrary direction and the angular separation of each pair of microphones. The summation of the angular separation results in the ‘Orientations Criterion’ of the array. It is noticed that the optimized cases present a strictly increasing angular separation between pairs of microphones. In contrast, the regular spaced array presents a redundancy in terms of microphones pairs orientations, with several microphones presenting the same orientation.

In the multi-objective problem an estimation of the Pareto Front is calculated from the linear regression of the statistical data obtained from the single-objective optimizations (Figure 7). It becomes evident that both fitness criteria are somehow coupled: since the separation in the design space between both single-objective optimization solutions is small, optimizing the array for one criteria is almost equivalent to optimizing the array for the another. This is confirmed in the multi-objective optimization: as little space is left for trade-off decisions on the Pareto Front, the multi-objective solution is close to the the single-objective solutions in the design space .

The previous findings are confirmed when the noise source maps of the respective arrays are evaluated with the GCC-PHAT. A numerical simulation was carried with a 94 dB (SPL) broadband monopole source at  $\phi = \theta = 0^\circ$  (Figure 8). The regular array shows strong side-lobe lines converging vertically, horizontally and with a circular pattern (Figure 8.a). In the case of the optimized arrays, a systematic side-lobe suppression was observed: the source map intensity decreases omnidirectionally from the source location. This observation may be perceived as a strong evidence of the proposed ‘spatial whitening’ mechanism.

#### 4.2 Sources with different pressure levels

A numerical simulation was carried to evaluate the performances of different arrays in the case of source localization of sources with different pressure levels. A regular array as the one from

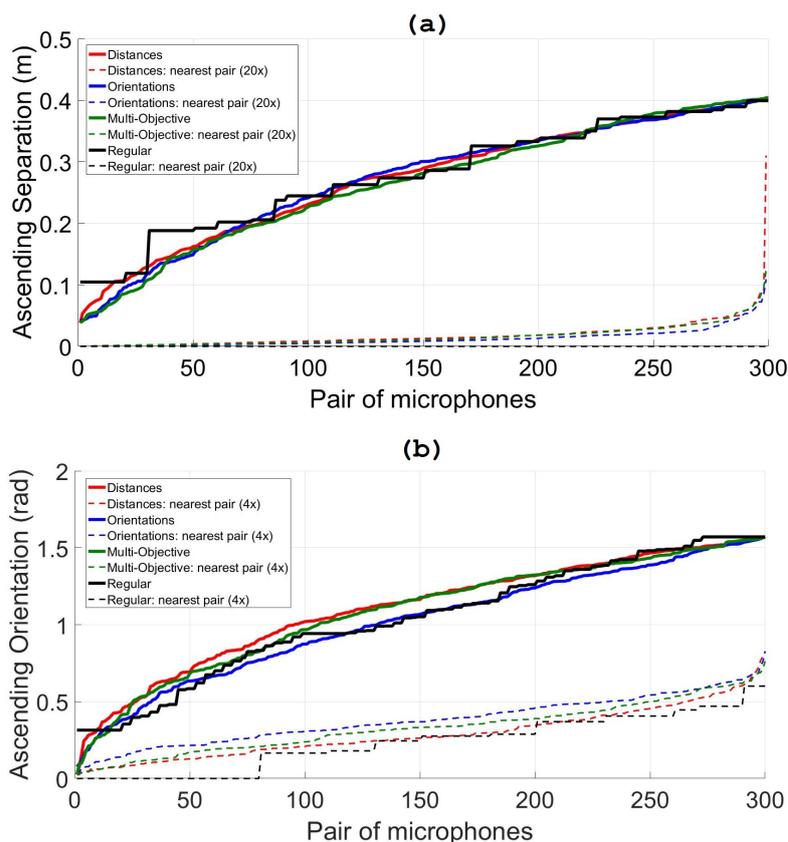


Figure 6: (a) Separation and difference of separation for each pair of microphone. The ‘Distances Criterion’ is obtained from the integration of the dashed curve. (b) Orientation and angular separation for each pair of microphone. The ‘Orientations Criterion’ is obtained from the integration of the dashed curve.

Figure 8.a was compared to an optimal array in terms of the ‘Distance Criterion’ (Figure 8.b). A set of 9 broadband monopoles (5 at 94 dB and 4 at 86 dB) was distributed in the space.

The simulation results are presented in Figure 9. From top-left to bottom-right the sources are localized alternately, starting with one of the source of lowest level. A qualitative analysis of the noise source maps reveals the expected tendency for this simulation: while the regular array provides a poor source map with main side-lobe lines and side-lobe intersections characterizing false-positives (Figure 9.a), the optimized kept the expected tendency of omnidirectionally suppressing side-lobes (Figure 9.b).

## 5 CONCLUSIONS

This study was able to numerically demonstrate the expected tendencies related to side-lobe rejection using specific geometric criteria for the distribution of microphones in an array. First, it was shown that, for a sphere, the optimization criteria based on the maximization of the sum of difference of distances and orientations between microphones pairs are similar in the design

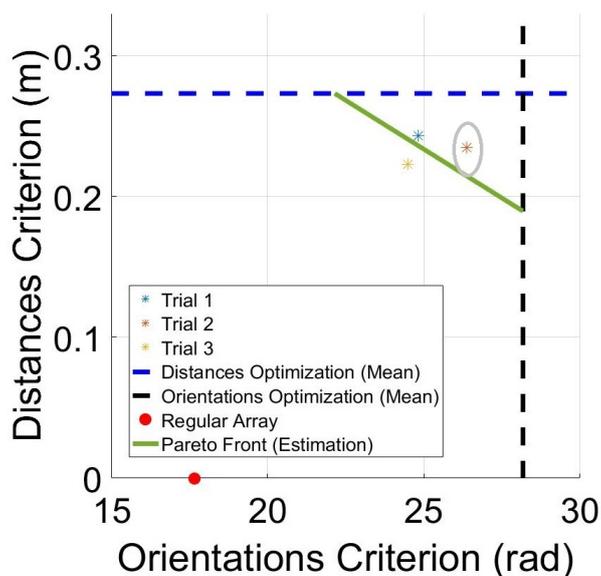


Figure 7: Estimated Pareto front for the multi-objective optimization. Best trial identified with an ellipse.

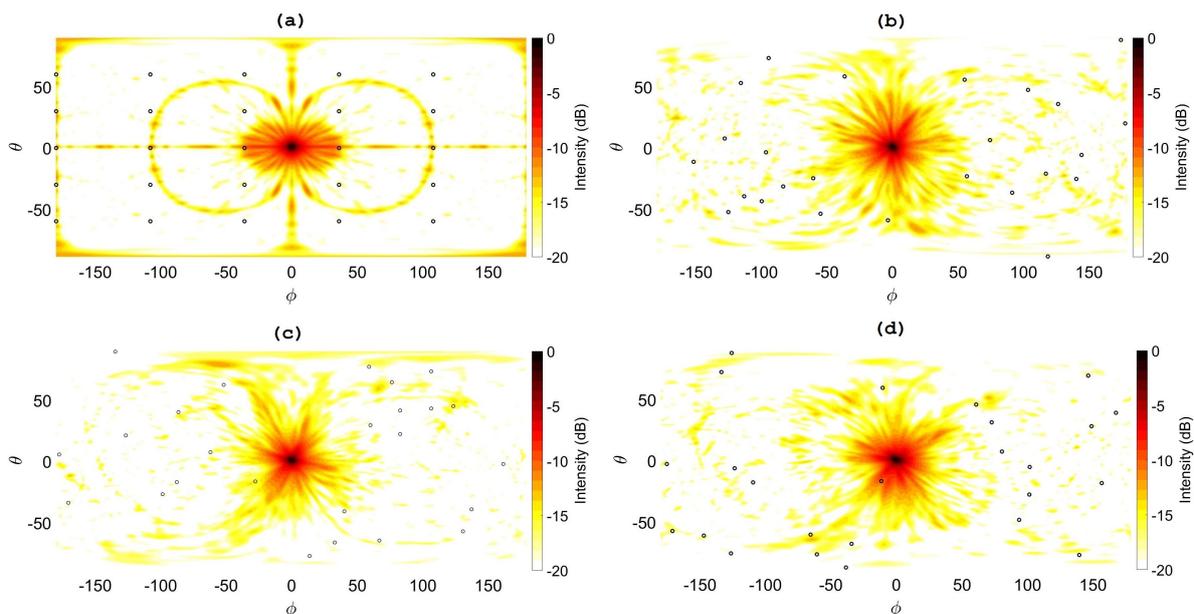


Figure 8: Microphones distributions and noise source map obtained with a numerical simulation of GCC-PHAT. Microphones positions illustrated by black dots. A broadband monopole is located in  $\phi = \theta = 0^\circ$ , at 5 m from the array. From (a) to (d) (best solutions obtained): regular array, ‘Distances Criterion’ array, ‘Orientations Criterion’ array and ‘Multi-Objective Criterion’ array.

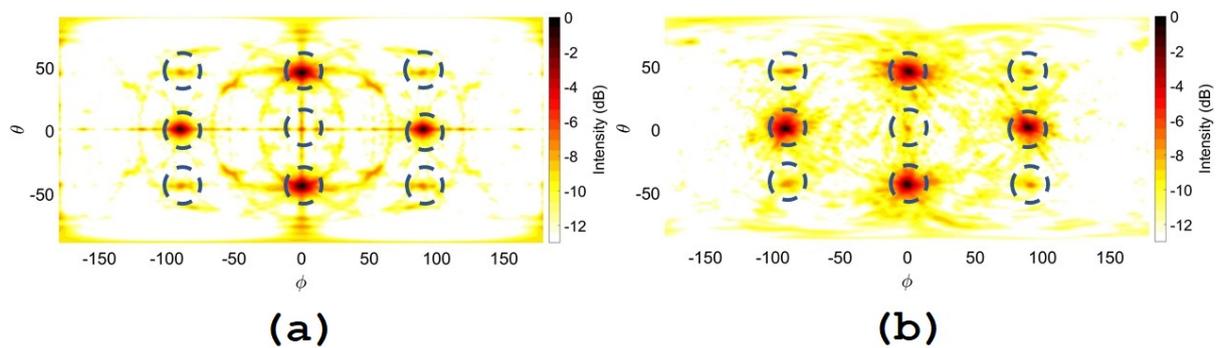


Figure 9: (a) Source localization of a regularly spaced array. (b) Source localization of an optimized array ('Distances Criterion'). Sources are conveniently identified by dashed-circles. Simulation performed with GCC-PHAT with sources 5 m away from the array.

space. Both solutions allow the suppression of side-lobes on the noise source maps obtained with the GCC-PHAT by the so-described mechanism of 'spatial whitening'. The obtained noise source map is smoother and improves the localization of sources with different levels, compared to a regularly spaced array. It was demonstrated that an optimization methodology based on genetic algorithms is computationally feasible and sufficiently flexible to accommodate different applications requirements. Future developments are proposed in sequence: validation of numerical results with similar experimental set-ups, development of specific metrics to accurately quantify the array's performances in terms of source localization and finally, validation of the optimization criteria for arrays of different sizes, shapes and applications.

## References

- [1] T. Padois, O. Doutres, F. Sgard, and A. Berry. Development of a microphone antenna incorporating an optical system to identify the position of the noisiest sound sources in an industrial setting. *Activité IRSST 2015-0075*, 2018.
- [2] C. Noël, V. Planeau, and D. Habault. A new temporal method for the identification of source directions in a reverberant hall. *Journal of Sound and Vibration*, 296(3):518–538, 2006.
- [3] T. Padois, F. Sgard, O. Doutres, and A. Berry. Acoustic source localization using a polyhedral microphone array and an improved generalized cross-correlation technique. *Journal of Sound and Vibration*, 2016.
- [4] C. Knapp and G. Carter. The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 24(4):320–327, 8 1976.
- [5] M. S. Brandstein and H. F. Silverman. A robust method for speech signal time-delay estimation in reverberant rooms. In *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 375–378, 4 1997.
- [6] C. Zhang, D. Florencio, and Z. Zhang. Why does PHAT work well in low noise, reverberative environments? *IEEE*, March 2008.

- [7] J. H. Dibiase. *A high-accuracy, low-latency technique for talker localization in reverberant environments using microphone arrays*. PhD thesis, Brown University, 2000.
- [8] P. Aarabi. Multi-sense artificial awareness. Master's thesis, University of Toronto, 1999.
- [9] J. Velasco, D. Pizarro, and J. Macias-Guarasa. Source localization with acoustic sensor arrays using generative model based fitting with sparse constraints. *Sensors*, 12:13781–13812, 10 2012.
- [10] T. Padois, O. Doutres, F. Sgard, and A. Berry. On the use of geometric and harmonic means with the generalized cross-correlation in the time domain to improve noise source maps. *The Journal of the Acoustical Society of America*, 140, 2016.
- [11] D. Salvati, C. Drioli, and G. L. Foresti. Exploiting a geometrically sampled grid in the steered response power algorithm for localization improvement. *The Journal of the Acoustical Society of America*, 141(1):586–601, 2017.
- [12] F. Grondin, D. Létourneau, F. Ferland, V. Rousseau, and F. Michaud. The manyyears open framework. *Autonomous Robots*, 34(3):217–232, 4 2013.
- [13] B. Rafaely. *Fundamentals of Spherical Array Processing*. Springer Topics in Signal Processing. Springer Berlin Heidelberg, 2015.
- [14] J. S. Hu, C. M. Tsai, C. Y. Chan, and Y. J. Chang. Geometrical arrangement of microphone array for accuracy enhancement in sound source localization. In *2011 8th Asian Control Conference (ASCC)*, pages 299–304, 5 2011.
- [15] M. R. Bai, J. H. Lin, and K. L. Liu. Optimized microphone deployment for near-field acoustic holography: To be, or not to be random, that is the question. *Journal of Sound and Vibration*, 329(14):2809 – 2824, 2010.
- [16] S. E. Nai, W. Ser, Z. L. Yu, and H. Chen. Beampattern synthesis for linear and planar arrays with antenna selection by convex optimization. *IEEE Transactions on Antennas and Propagation*, 58(12):3923–3930, 12 2010.
- [17] Y. Collette and P. Siarry. *Multiobjective Optimization: Principles and Case Studies*. Decision Engineering. Springer Berlin Heidelberg, 2011.
- [18] K. Deb. *Introduction to Evolutionary Multiobjective Optimization*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.
- [19] P. Aarabi. The fusion of distributed microphone arrays for sound localization. *EURASIP Journal on Advances in Signal Processing*, 2003(4):860465, 2003.
- [20] S. Krivoshapko and V.N. Ivanov. *Encyclopedia of Analytical Surfaces*. Springer International Publishing, 2015.