DOI: 10.61782/fa.2025.0346

FORUM ACUSTICUM
EURONOISE 2025

ENHANCING THE GEOMETRICAL CALIBRATION OF MICROPHONE ARRAYS: SOLUTIONS AND EXPERIMENTAL VALIDATION

Matthieu Hartenstein[1*] **Paul Luizard**[3,4] **Hélène Moingeon**[2]
Cédric Pinhède[2] **Marc Pachebat**[2] **Christian Ollivon**[1]
François Ollivier[1] **Fabrice Silva**[2]

[1] Institut Jean le Rond ∂'Alembert, Sorbonne Université, CNRS UMR 7190, Paris, France
[2] Aix Marseille Université, CNRS, Centrale Méditerranée, LMA UMR 7031, Marseille, France
[3] Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, Grenoble, France
[4] Audio Communication Group, Technische Universität Berlin, Germany

ABSTRACT

In acoustic imaging, the need for a high spatial resolution requires the deployment of arrays comprising a substantial number of microphones. The topographical report of each sensor constitutes a laborious process prone to imprecisions. Over the past two decades, there has been a surge of interest in geometrical calibration methods based on the estimation of acoustic time-of-flights (TOF). In 2019, Vanwynsberghe et al. presented numerical and experimental tests demonstrating the higher precision of their Robust MultiDimensional Unfolding (RMDU) algorithm compared to the state-of-the-art methods. However, they report experimental root-mean-square errors (RMSE) of 2 cm, which limits the usable band of microphone arrays to frequencies below a few thousand Hertz. The present study shows that locally oversampling the cross-correlation functions significantly enhances the TOF estimation. The framework is tested on a surrounding 3.6 m-diameter quasi-spherical array of 960 microphones installed in an anechoical room. The comparison of a subset of estimated positions with high-precision optical reference measurements results in a RMSE of 0.75 cm. Numerical simulations in similar experimental conditions yield a maximum RMSE of 0.2 cm. This difference is partly explained by the discrepancy between the reported position of the microphones and the position of their acoustic center.

Keywords: Geometrical calibration, microphone arrays, experimental validation.

1. INTRODUCTION

In more and more microphone array applications, the need for a high spatial sampling resolution can demand the deployment of hundreds, of microphones. If not unfeasible, the report by hand of the microphone position is cumbersome and subject to imprecisions. Since twenty years and the pionneering work from Birchfield [1], a handful of methods have been proposed to estimate the location of microphones based on acoustic measurements. The signal processing framework that allows to retrieve positions from audio recordings can be divided into two steps.

First, Propagation Delays (PDs) are estimated from measured acoustic signals. The PD estimation step of a geometric calibration experiment is usually done by computing a correlation function between source and microphones or between pairs of microphones, and by finding the peaks of this correlation function. Depending on the involved geometric calibration method, the delays of interest can be Time Of Flights (TOF) between the sources at different positions and array sensors, or Time Differences Of Arrival (TDOA) between the sensors that constitute the array. In a TOF-based calibration frameworks, $N_{\text{src}}N_{\text{mic}}$ source-microphone PDs have to be estimated (where N_{src} and N_{mic} are respectively the num-

*Corresponding author: matthieu.hartenstein@sorbonne-universite.fr.

ber of sources and microphones), while a TDOA-based framework demands the computation of N_{mic}^2 PDs. To the knowledge of the authors, most experimental studies in the literature make use of less than 30 sources. The present work focuses on the geometrical calibration of arrays of several hundreds of microphones. In order to lighten the PD computation step case, the choice of a TOF-framework is therefore made.

The second step of the geometrical calibration frameworks consists in mapping source-microphone TOFs into source and microphone positions. This is generally done by minimizing the difference between the reconstructed source and microphone positions $(\mathbf{x}_s^{\text{src}})_{s\in[0,\dots,N_{\text{src}}]} \in (\mathbb{R}^3)^{N_{\text{src}}}$ and $(\mathbf{x}_s^{\text{mic}})_{m\in[0,\dots,N_{\text{mic}}]} \in (\mathbb{R}^3)^{N_{\text{mic}}}$ and the distances $(\delta_{sm})_{(s,m)\in[0,\dots,N_{\text{src}}]\times[0,\dots,N_{\text{mic}}]}$ inferred from the estimated TOFs $(\tau_{sm})_{(s,m)\in[0,\dots,N_{\text{src}}]\times[0,\dots,N_{\text{mic}}]}$ ($\delta_{sm} = c_0\tau_{sm}$ with c_0 the speed of sound)

$$\mathcal{L}_{\text{MDU}}(\mathbf{X}^{\text{mic}}, \mathbf{X}^{\text{src}}) = \sum_{s=1}^{N_{\text{src}}}\sum_{m=1}^{N_{\text{mic}}} \left(\delta_{sm} - ||\mathbf{x}_m^{\text{mic}} - \mathbf{x}_s^{\text{src}}||_2\right)^2. \tag{1}$$

Minimizing the cost function in Eq. (1) is referred in the literature as Multi Dimensional Unfolding [2, 3]. Several methods have been proposed for solving the MDU problem, either by use of an iterative algorithm (*e.g.* the FrSM method from Ref. [4]), or of a closed-form solution (C-CF and L-CF in respectively Refs. [5] and [6]). A possible drawback of a nonlinear least-squares cost function such as the one in Eq. (1) is its sensitivity to outliers. In fact, the TOFs computed in the first step of the geometrical calibration experiment correspond to source-microphone distances only if the corresponding peak identified in the GCC function corresponds to the incident propagation path. In the case where unwanted reflections happened during the acquisition of acoustic signals, the TOF estimation is subject to lead to TOF outliers. To cope with this issue, Vanwynsberghe *et al.* [7] propose the Robust Multi Dimensional Unfolding (RMDU) algorithm. In RMDU, potential outliers are identified together with the source and microphone positions. Ref. [7] compares the performance of the RMDU algorithm with state-of-the-art methods in numerical simulations. RMDU is shown to lead to the best performance with a RMS positioning error of about $2\,\text{mm}$ when $1\,\%$ of outliers contaminate the TOF population. Nevertheless, a real-life experiment presented in Ref. [7] evidences an error positioning an order of magnitude ($2\,\text{cm}$) above the one obtained in numerical simulations. For comparison, Ref. [8] has shown that a RMS uncertainty of $5\,\text{mm}$ in sensor positioning significantly reduces the usable frequency band of the microphone array when reconstructing the directivity of a sound source.

In the present study, an effort is made to reduce the effect of different sources of TOF bias on the performance of the RMDU algorithm. Sec. 2 shows that oversampling the GCC function can improve the TOF estimation in cases where the measurement noise level is reasonably low, and that the outlier-rejecting property of the RMDU algorithm should enable to handle the TOF outliers obtained in case of higher noise levels. Sec. 3 presents the results of the geometrical calibration of a surrounding $3.6\,\text{m}$-diameter quasi-spherical array of 960 MEMS microphones. Finally, Sec. 4 concludes this study.

2. GEOMETRICAL CALIBRATION METHOD

2.1 TOF estimation

2.1.1 The generalized cross-correlation function

The estimation of source-microphone TOFs demands the prior computation of a correlation function between the signal sent to source and the signal recorded by microphone. The Generalized Cross-Correlation (GCC) function between two time signals x and y writes [9]

$$r_{xy}(\tau) = \int_{\mathbb{R}} \Psi(f) S_{xy}(f) e^{j\pi f\tau} df, \tag{2}$$

where S_{xy} is the cross-spectrum between x and y, and Ψ is an appropriately chosen frequency weight function. In this work, and as in numerous study on the geometric calibration of microphones (see, e.g., Refs. [4, 7]) the choice of the so-called PHAse Transform (PHAT) weight function $\Psi(f) = \dfrac{1}{|S_{xy}(f)|}$ is made. In the case where x and y are signals observed at two different positions due to the propagation of a wave in space, r_{xy} should present a peak at the delay corresponding to the propagation distance between the two corresponding points in space.

2.1.2 Estimating propagation delays from the GCC

Ref. [7] emphasizes that acoustic propagation paths between a sound source and a microphone can be such that the incident wave can appear as a weaker peak than the one associated to a reflected wave in the GCC function. Figure 1 shows an example of GCC function where the contributions from different reflected paths arrive in phase on the sensor of interest, leading to a higher peak than the peak associated to the incident path. This example shows that, in presence of reflections, identifying the delay where

the GCC function is maximal can lead to a wrong TOF estimate. However, the contribution that propagates via the incident path will always appear prior to the contribution that arises from reflections. Based on this observation, the approach depicted in Fig. 1 is proposed and used in the remaining of this chapter. Firstly, The maximum value of the GCC function is identified (peak arising at $0.5\,\mathrm{m}$ and corresponding to contributions from the reflected path in Fig. 1). The peaks of the GCC functions are sought for in a time window centered on the maximal peak of the GCC function. This window has a duration equivalent to an equivalent maximal difference in propagation distance $L_{\max}$ ($L_{\max}$ =$0.5\,\mathrm{m}$ in Fig. 1). A threshold value η ($\eta = 0.5$ in the remaining of this study) is defined, and a peak in the GCC is defined as a sample whose value is higher than a fraction η of the GCC's maximal value (dashed green horizontal line), and whose two nearest samples present lower values than the present sample. In Fig. 1, only two peaks are identified : the contribution of the incident path at $0.4\,\mathrm{m}$, and the contribution of the reflected paths at $0.5\,\mathrm{m}$. From the identified peaks in the GCC, the earliest one is deemed to be the incident peak. The corresponding TOF is identified.

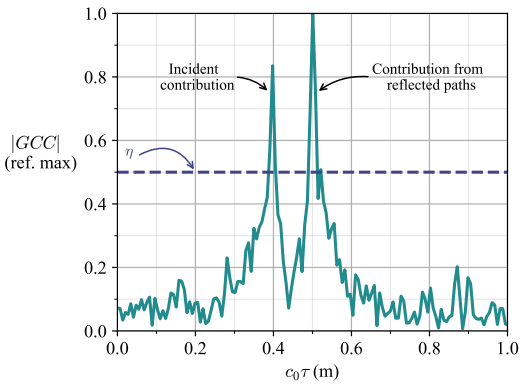


Figure 1: Shape of the GCC function (blue curve) between a source signal and a microphone signal when contributions coming from several reflected paths arrive in phase. In the case presented here, the contribution from reflected paths is higher than the incident contribution. To overcome this issue, the first GCC peak whose amplitude is higher than a fraction $\eta = 0.5$ of the maximum GCC peak is sought for.

2.1.3 Oversampling the GCC function

The temporal resolution $1/F_s$ (where F_s is the sampling frequency) of the signals used to compute the GCC function introduces a stochastic bias that follows a uniform distribution on the interval $[-F_s/2, F_s/2]$ in the TOF estimation process. To reduce the impact of this bias, we propose to locally over-sample the GCC function in the time interval where the TOF is sought for.

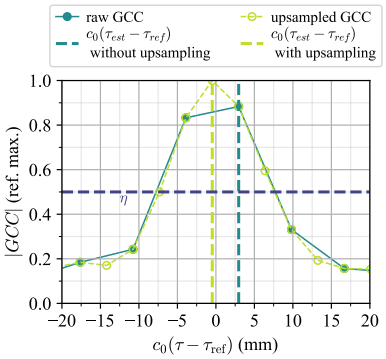
In the particular case shown in Fig. 2 (a), the theoretical peak of the GCC should be observed at $c_0(\tau - \tau_{\text{ref}}) = 0$. The maximal sample of the observed raw GCC function - whose abscissa is indicated as a solid blue line - is located at an equivalent distance of about $2.5\,\mathrm{mm}$ from the theoretical peak. The oversampled GCC (dashed line and circles) presents a maximum at an abscissa (light green dashed vertical line) less than approximately $1\,\mathrm{mm}$ close to the theoretical peak. In this case, oversampling the GCC by a factor 2 therefore allows to divide the error on estimated TOF by more than 2.

The approach described in the last paragraphs presents limitations in presence of sensor noise. Figure 2 (b) shows a particular case where the sensor noise level is acceptable, and where the overall shape of the GCC function is preserved. Even in cases where the sensor's SNR is reasonable, noise can accentuate some side lobes of the GCC function. In Fig. 2 (b), the side lobe appearing prior to the main GCC lobe at an equivalent propagation distance of $4\,\mathrm{cm}$ has an amplitude higher than the peak detecion threshold $\eta = 0.5$ that was defined in Sec. 2.1.2 to enhance the robustness of the peak detection method to reflections. This phenomenon leads to TOF estimation errors of the same order as the extent of the time-window in which the incident peak is sought for. Figure 2 (c) shows another example, where the noise level at a particular sensor is extreme. In this case, the incident peak of the GCC is drown into noise, and the maximum peak can appear randomly over the whole extent of the GCC function. In the case depicted in Fig. 2 (c), an equivalent distance error of a few tens of centimeters is made in the TOF estimation. In the cases depicted in Figs. 2 (b) and (c), oversampling the GCC does not improve the TOF estimation.

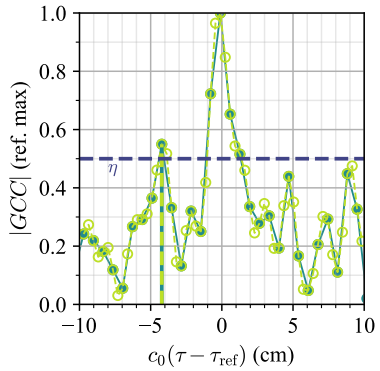
2.2 Reconstructing microphone positions

2.2.1 The RMDU algorithm

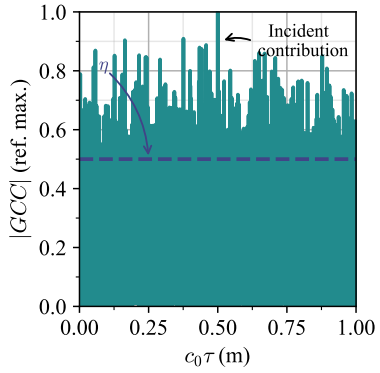
In their work, *Vanwynsberghe et al.* [7] propose the Robust Multi Dimensional Unfolding (RMDU) algorithm.



(a) Oversampling the GCC



(b) GCC outlier, case 1



(c) GCC outlier, case 2

Figure 2: (a) : illustration of the potential benefits of the GCC oversampling approach. (b) and (c) : Cases where measurement noise comes to bias the GCC function, and where oversampling does not improve the TOF estimation.

Their approach consists in finding matrices of source and microphone positions $\mathbf{X}^{\text{mic}} = [\mathbf{x_m^{\text{mic}}}] \in \mathbb{R}^{3\times N_{\text{mic}}}$ and $\mathbf{X}^{\text{src}} = [\mathbf{x_m^{\text{src}}}] \in \mathbb{R}^{3\times N_{\text{src}}}$, and of outlier errors $\mathbf{O} = [o_{sm}]_{(s,m)\in[0,N_{\text{src}}-1]\times[0,N_{\text{src}}-1]}$ that minimize the following cost function

$$\mathcal{L}_{\text{RMDU}}(\mathbf{O},\mathbf{X}^{\text{mic}},\mathbf{X}^{\text{src}}) = \sum_{s=1}^{N_{\text{src}}}\sum_{m=1}^{N_{\text{mic}}}\left(\delta_{sm}-||x_m^{\text{mic}}-x_s^{\text{src}}||_2-o_{sm}\right)^2+\lambda||\mathbf{O}||_1. \tag{3}$$

The terms o_{sm} model large errors due to TOA outliers. The ℓ_1 regularization term in Eq. 3 ensures sparsity in the matrix of outlier errors.

The cost function in Eq. (3) is minimized in terms of its parameters using an iterative framework. Each iteration n of RMDU implies a soft-thresholding step (sgn is the sign operator)

$$o_{sm}^{(n+1)} = S_\lambda(\delta_{sm}-||x_m^{\text{mic},(n)}-x_s^{\text{src},(n)}||_2), \text{ where} \tag{4}$$

$$S_\lambda(u) = sgn(u)(|u|-\lambda/2) \text{ if } |u|>\lambda/2, \text{ else } 0. \tag{5}$$

The soft-thresholding operator S_λ in Eq. (4) shrinks the absolute values of distance errors larger than a threshold $\lambda/2$, and vanishes for errors lower than $\lambda/2$. $\lambda/2$ is thus to be understood as the minimal error for which a TOA will be deemed outlier.

2.2.2 *Performance of the MDU and RMDU frameworks*

In Ref. [4] (resp. [10]), Khanal *et al.* (resp. Le *et al.*) report on experimental Root-Mean-Square (RMS) positioning errors of 1 cm for the FrSM algorithm (resp. 4 cm to 5 cm for C-CF and L-CF). In Ref. [7], the performance of the RMDU algorithm is compared with the one of the FrSM method in numerical simulations. RMDU is shown to outperform FrSM with a RMS positioning error of about 2 mm when 1 % of outliers contaminate the TOF population (RMS error of 6 cm for FrSM). A real-life experiment presented in Ref. [7] yields an RMS error of 2 cm. The mismatch of an order of magnitude between the simulated and experimental performances of RMDU evidences that some sources of uncertainty lack in the simulations from Ref. [7]. Moreover, the advertised experimental error is comparable to the acoustic wavelength in audible applications, which limitates the bandwidth of the calibrated array.

2.2.3 *Proposed use of the algorithm*

Sec. 2.1.3 has shown that, in noisy environments, the incident peak of the GCC function can be drown into noisy

components. In such case, the TOF estimation can be significantly impacted, leading to outliers in the source-microphone TOF matrix. Originally designed for the geometric calibration of microphone arrays in mismatched free-field, the RMDU algorithm can in fact detect outliers of any nature in the TOF matrix. The parameter λ in the RMDU algorithm controls the level of TOF error from which a source-microphone pair is considered to be an outlier. For a high value of λ, no outlier will be sought for, and the RMDU problem reduces to classical MDU. In this particular case, the outlier term in the cost function of the RMDU algorithm Eq. (3) is forced to vanish, and the RMDU algorithm reduces to a differentiable simple least-squares problem. For lower values of λ, RMDU allow more outliers to exist in the solution. In this case, the outlier term in the cost function is not null anymore, and the optimization problem behind RMDU is not differentiable. To be able to detect outliers, our use of the RMDU algorithm is the following

1. Initialize the microphone positions randomly on a sphere of same radius as the array to be calibrated, and the source positions at the positions where they are expected to be.
2. Use the RMDU algorithm with a parameter λ larger than the duration of the GCC function. In these conditions, and at each iteration, the soft-thresholding step of the RMDU algorithm (Eq. (4)) detects no outlier in the solution. The RMDU cost function in Eq. (3) converges more easily to a minimum, even with a coarse initialization. At this stage, most microphone positions are correctly estimated, but the RMDU algorithm is not able to detect the TOF outliers. The corresponding estimated microphone positions highly deviate from their true values.
3. Use the result of the first RMDU trial to initialize a second RMDU trial, using a lower value of λ. In these conditions, the optimization problem is not differentiable and will converge less quickly. However, if the solution of the first trial is close enough to the true solution for most sources and microphones, the second RMDU trial can be expected to lead to a consistent solution.

Figure 2 (a) highlights the fact that the maximal error made in the TOF estimation under ideal conditions is the half-distance $1/(2F_s)$ between two time samples. In the RMDU algorithm, a source-microphone pair is deemed outlier if the estimated distance TOF for this pair deviates more than $\lambda/2$ from the estimated sensor and microphone positions at iteration n. An adequate parameter λ for the second RMDU run should thus be of the order of magnitude of the sound propagation distance equivalent to the distance between two time samples

$$\lambda \approx \frac{c_0}{F_s}. \tag{6}$$

3. EXPERIMENTAL VALIDATION

3.1 Setup

3.1.1 Disposition of microphones and sources

Figure 3 depicts the setup used in the experiment. The array consists in 960 MEMS microphones placed on a spheroidal structure of radius $1.8\,\mathrm{m}$ and is located in an anechoical room. Nine sources are disposed around and inside the array. One of the sources is approximately placed at the center of the array. The remaining eight sources are distributed at a diversity of positions outside the array. The temperature of the room was reported to T_C =22.9 °C during the experiment, which gives a speed of sound $c_0 = \sqrt{\frac{\gamma R T_K}{M}} = 345\,\mathrm{m/s}$,.

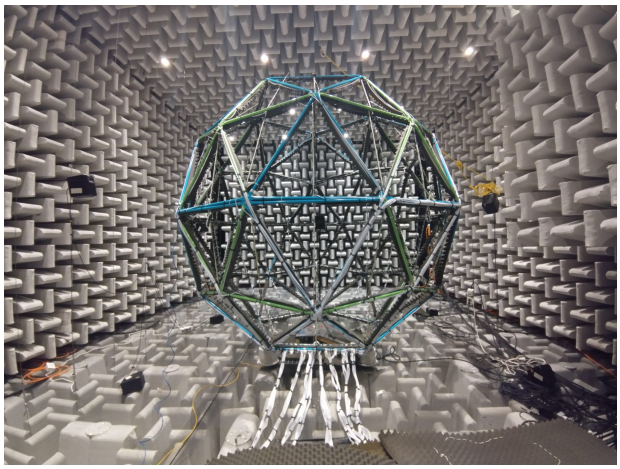


Figure 3: Picture of the experimental setup for the geometric calibration of a microphone array.

3.1.2 Reference geometric measurements

Prior to the geometric calibration experiment, a total station shown in Fig. 4 was used. This optical device allows

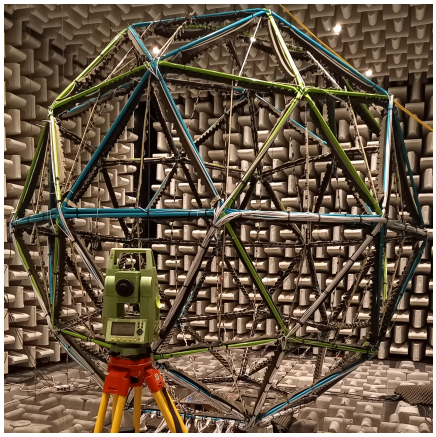


Figure 4: Picture of the total station in front of the array to be calibrated.

to manually report the relative positions of points in space by sequentially targetting these points with the help of an optical viewfinder. Due to the cumbersomeness of the total station report, only 25 MEMS microphone positions were measured. Moreover, two of the reported positions deviate by more than a meter from their expected positions. These points are eliminated from the report.

3.2 Signal recording and processing

3.2.1 Acquisition system

The 9 sound sources shown in Fig. 3 sequentially played a calibration signal. Both the signal sent to each source and the signals of the 960 MEMS microphones composing the array were recorded using the Megamicros acquisition system [7]. A sampling frequency of 50 kHz was used. A preliminar acoustic calibration of the microphone array brought to light a source-microphone delay of N_{delay} =8 samples in the acquisition system. Numerical simulations that are not presented here for the sake of concision showed that introducing such a delay in the RMDU algorithm can lead to RMS positioning errors of several centimeters. This delay was thus compensated for in the acquired signals before the GCC computation.

3.2.2 Signal synthesis and processing

The calibration signal sent to the source was a Synchronized Swept-Sine signal [11] of duration 30 s, sampled at 50 kHz and spanning the frequency range 100 Hz to 10 000 Hz. First, transfer functions between the theoretical expression of the calibration signal and the source and microphone recorded signals are computed using the method described in [11]. This allows to get rid of the non-linear contributions introduced by the calibration sources. The linear transfer functions between the theoretical swept-sine signal and the source and microphone recordings are used to compute the GCC (Eq. (2)).

3.2.3 Parameter of the TOF estimation and position reconstruction

The GCC is locally oversampled at a virtual sampling frequency of 500 kHz. The first trial of the RMDU algorithm is performed with an outlier threshold $\lambda = 50$ m. In this trial, the RMDU algorithm is initialized using approximate positions of the microphone bars on the spheroidal structure. The parameter λ used in the second RMDU trial is fixed to λ =5 mm, according to Eq. (6).

3.3 Results

3.3.1 Estimated source and microphone positions

Fig. 5 (a) shows the source and microphone positions estimated with the proposed geometric calibration strategy. The 9 estimated source positions are depicted as big dark blue dots. One of them is approximately located at the center of the array. Four of them are retrieved around the array at the floor level. The remaining four are retrieved around the array at the equator level. This qualitatively corresponds to the true source positions seen on the picture of the experimental setup shown in Fig. 3. As in the true disposition of sensors in Fig. 3, the 960 estimated microphone positions are qualitatively aligned by groups of eight. The average distance between the estimated microphone position and the centroïd of the set of positions is of 1.72 m. The true radius of the outer spheroidal surface in the experimental setup shown in Fig. 3 is 1.8 m. In practice, the microphones are placed inside this stucture, which explains the 8 cm lower average distance to center.

3.3.2 Cumulative density histogram

Figure 5 (b) shows the cumulative density histogram of the error between the 23 valid total station measurements and the corresponding positions estimated with the proposed geometric calibration strategy. More than 80 % (19 microphones) of the errors are smaller than 9 mm, and the remaining microphones (4 microphones) present errors in the range 1 cm to 1.5 cm.

3.3.3 *RMS and standard deviation of the estimation error*

Figure 5 (b) indicates the RMS and standard deviation of the distances between the positions of the microphone measured with the total station and the ones reconstructed with the geometric calibration strategy. The RMS error is of 7.6 mm. The RMS $\pm$ one standard deviation of this error is represented as dashed green vertical lines. The standard deviation of the error population is of 3.4 mm.

3.3.4 *Bias in the microphone position estimation*

According to the bias-variance decomposition of the RMS error [12], the bias between the total station measurements and the positions reconstructed with RMDU is

$$bias(\epsilon_{\mathbf{x}}^{\text{mic}}) = \sqrt{RMS(\epsilon_m^{\text{mic}})^2 - std(\epsilon_m^{\text{mic}})^2} = 6.8\,\text{mm}. \tag{7}$$

The RMS deviation between the estimated microphone positions and the ones measured using the total station is therefore dominated by its bias contribution. In fact, the microphone positions estimated with T(D)OA-based geometric calibration methods correspond to the positions of their equivalent acoustic centers (commonly defined [13] as 'the point from which spherical wavefronts appear to diverge' if these microphones were used as sound sources). At low frequencies where the acoustic wavelength is higher than the characteristic dimension of the sensor of interest, this acoustic center is expected to be located in front of the sensor's baffle [13]. As an example, the low-frequency acoustic center of a point source on a spherical baffle of radius a is located about $a/2$ in front of the baffle. The MEMS microphones that were used to build the array calibrated in this section are baffled by a rectangular plate of about 5×10 mm. Their acoustic center can thus be expected to be located a few millimeters in front of the MEMS microphone's port that was targeted in the total station measurements. This partly explains the high bias observed between the total station measurements and the estimated microphone positions.

3.3.5 *Note on the position alignement*

Comparing the total station measurements with the estimated microphone positions demands a prior alignment step, which was done using Procrustes analysis [14]. The last paragraph showed that there is an instrisic bias between the geometric measurements and the microphone positions retrieved with a T(D)OA-based geometric calibration method. Trying to align these positions can introduce additional bias and variance contributions in the RMS error between the estimated microphone positions and the total station measurements.

4. CONCLUSION

This work illustrated the effect of the temporal resolution and of measurement noise on the geometrical calibration of microphone arrays. It was shown that oversampling the correlation function prior to searching its peaks can improve the TOF estimation, and that the RMDU algorithm proposed in Ref. [7] enables to reject the TOF outliers introduced by measurement noise.

The geometrical calibration of a 3.6 m-diameter quasi-spherical array of 960 microphones was performed. A lower experimental RMS positioning error of 7.5 mm than the one of 2 cm reported in Ref. [7] was obtained. Finally, the mismatch between optically reported microphone positions and the positions of the microphones' acoustic centers reconstructed with RMDU explained the discrepancy of about 5 mm between the simulated RMS error obtained in Ref. [7] and the experimental RMS error of the present study.

5. ACKNOWLEDGMENTS

This research is funded by the French National Research Agency (ANR-21-CE42-0017, 2021-25).

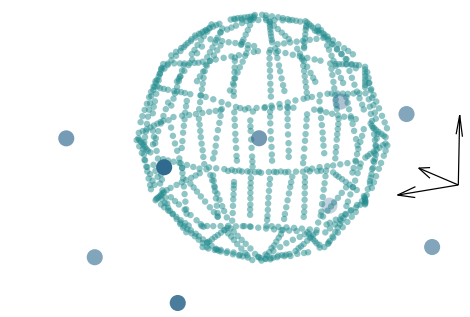
6. REFERENCES

[1] S. Birchfield, "Geometric microphone array calibration by multidimensional scaling," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 5, (Hong Kong, China), 2003.

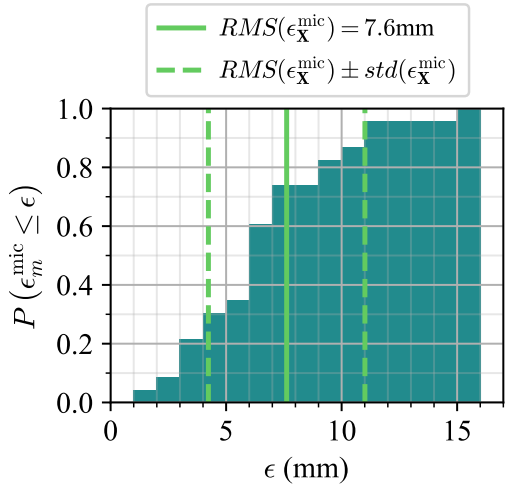
[2] I. Dokmanic, R. Parhizkar, J. Ranieri, and M. Vetterli, "Euclidean Distance Matrices: Essential Theory, Algorithms and Applications," *IEEE Signal Processing Magazine*, vol. 32, pp. 12–30, Nov. 2015.

[3] A. Plinge, F. Jacob, R. Haeb-Umbach, and G. A. Fink, "Acoustic Microphone Geometry Calibration: An overview and experimental evaluation of state-of-the-art algorithms," *IEEE Signal Processing Magazine*, vol. 33, pp. 14–29, July 2016.

[4] S. Khanal, H. F. Silverman, and R. R. Shakya, "A Free-Source Method (FrSM) for Calibrating a Large-Aperture Microphone Array," *IEEE Transactions on*



(a)



(b)

Figure 5: Results of the geometric calibration experiment. (a) : Reconstructed source (big dark blue dots) and microphone (smaller and lighter blue dots) positions. (b) : Cumulative density histogram of the error between the 23 valid total station measurements and the corresponding positions estimated with the proposed geometric calibration strategy. The RMS position error is indicated as a solid green vertical line (resp. light green). The dashed green vertical lines represent the RMS value of the error $\pm$ its standard deviation ($std(\epsilon_m^{\text{mic}}) = 3.4\,\text{mm}$).

Audio, Speech, and Language Processing, vol. 21, pp. 1632–1639, Aug. 2013.

[5] M. Crocco, A. Del Bue, and V. Murino, "A Bilinear Approach to the Position Self-Calibration of Multiple Sensors," *IEEE Transactions on Signal Processing*, vol. 60, pp. 660–673, Feb. 2012.

[6] T.-K. Le and N. Ono, "Numerical formulae for TOA-based microphone and source localization," in *14th International Workshop on Acoustic Signal Enhancement*, (Juan-les-Pins), pp. 178–182, Sept. 2014.

[7] C. Vanwynsberghe, P. Challande, F. Ollivier, J. Marchal, and R. Marchiano, "Geometric calibration of very large microphone arrays in mismatched free field," *The Journal of the Acoustical Society of America*, vol. 145, pp. 215–227, Jan. 2019.

[8] M. Hartenstein, P. Luizard, H. Moingeon, C. Pinhède, M. Pachebat, C. Ollivon, F. Ollivier, and F. Silva, "A Method for Directivity Estimation with a High-Order Non-Spherical Microphone Array," in *Proc. Forum Acusticum*, (Torino, Italy), 2023.

[9] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, pp. 320–327, Aug. 1976.

[10] T.-K. Le, N. Ono, T. Nowakowski, L. Daudet, and J. De Rosny, "Experimental validation of TOA-based methods for microphones array positions calibration," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, (Shanghai, China), pp. 3216–3220, Mar. 2016.

[11] A. Novak, P. Lotton, and L. Simon, "Synchronized Swept-Sine: Theory, Application, and Implementation," *J. Audio Eng. Soc.*, vol. 63, pp. 786–798, Nov. 2015.

[12] S. Theodoridis, *Machine learning: a Bayesian and optimization perspective*. London: Elsevier, Academic Press, 2nd edition ed., 2020.

[13] F. Jacobsen, S. Barrera Figueroa, and K. Rasmussen, "A note on the concept of acoustic center," *The Journal of the Acoustical Society of America*, vol. 115, pp. 1468–1473, Apr. 2004.

[14] W. Kabsch, "A solution for the best rotation to relate two sets of vectors," *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, vol. 32, no. 5, pp. 922–923, 1976.