



# Geometric calibration of very large microphone arrays in mismatched free field

Charles Vanwynsberghe, Pascal Challande, François Ollivier, Jacques Marchal, Régis Marchiano

## ► To cite this version:

Charles Vanwynsberghe, Pascal Challande, François Ollivier, Jacques Marchal, Régis Marchiano. Geometric calibration of very large microphone arrays in mismatched free field. Journal of the Acoustical Society of America, Acoustical Society of America, 2019, 145 (1), pp.215-227. 10.1121/1.5083829 . hal-02003732

**HAL Id: hal-02003732**

**<https://hal-ensta-bretagne.archives-ouvertes.fr/hal-02003732>**

Submitted on 28 Feb 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Geometric calibration of very large microphone arrays in mismatched Free Field

Charles Vanwynsberghe,<sup>1</sup> Pascal Challande,<sup>2</sup> François Ollivier,<sup>2</sup> Jacques Marchal,<sup>2</sup> and Régis Marchiano<sup>2</sup>

<sup>1</sup>*Lab-STICC UMR 6285, CNRS, ENSTA Bretagne, 2 rue François Verny, 29200 Brest, France<sup>a</sup>*

<sup>2</sup>*Sorbonne Université, CNRS, Institut Jean Le Rond d'Alembert, 4 place Jussieu, 75005 Paris, France*

Large microphone arrays are an efficient means for source localization thanks to a wide aperture and a great number of sensors. When such arrays are deployed *in situ*, accurate geometric calibration becomes essential to obtain the microphone positions. In free field, the classic procedures rely on measured Times Of Arrival (TOA) or Time Differences of Arrival (TDOA) between the microphones and several controlled sources. However free field model mismatches, such as reflectors, generate outliers which severely deteriorate the positioning accuracy. This paper introduces a unified framework for robust calibration using TOA or TDOA, by exploiting an outlier-aware noise model. Thanks to the largeness of the array, the existing outliers are sparse and can be identified by a Lasso regression. From this, three iterative robust solvers<sup>a</sup> are proposed: (i) for TOA by Robust Multi Dimensional Unfolding, a particular variation of Robust Multi Dimensional Scaling (ii) for TDOA by data predenoising based on sparse and low-rank matrix decomposition, and (iii) for TDOA by jointly identifying the outliers and the geometry. The relevance of outlier-aware approaches is asserted by numerical and experimental tests. Compared with the baseline least-square approaches, the proposed robust solvers significantly improve the positioning accuracy in a free field mismatched by reflectors.

Copyright 2019 Acoustical Society of America. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the Acoustical Society of America. The following article appeared in *The Journal of the Acoustical Society of America* 145, 215 and may be found at <https://asa.scitation.org/doi/full/10.1121/1.5083829>.

---

<sup>a</sup> Python code available online at <https://github.com/cvanwynsberghe>.

<sup>a)</sup> [charles.vanwynsberghe@ensta-bretagne.fr](mailto:charles.vanwynsberghe@ensta-bretagne.fr)

## I. INTRODUCTION

A microphone array is a classical tool to solve a source localization problem. The latter consists in finding the origin and amplitude of the sources which produce the measured sound field. In addition to conclusions from antenna theory (Williams, 1999), it is experimentally well-known that increasing both the aperture (Sachar *et al.*, 2001) and the number of microphones (Weinstein *et al.*, 2007) improve the global performances of localization. Thanks to the outcome of microphones based on the technology called Microelectromechanical Systems (MEMS), developing a large microphone array with several hundreds of microphones becomes easier and low-cost compared with older technologies (Hafizovic *et al.*, 2012; Koyano *et al.*, 2016; Vanwynsberghe *et al.*, 2015). At the same time, original multichannel processing methods anticipate and rely on the use of a great number of sensors: three examples are dereverberation (Chardon *et al.*, 2015), source separation (FitzGerald *et al.*, 2016) and monitoring (Lai *et al.*, 2013). However for most applications, the position of the sensors must be known with a sufficient accuracy (Chen *et al.*, 2015; Himawan *et al.*, 2008).

If a large array is to be deployed *in situ*, the use of conventional measurement methods (eg with a laser-based system), to identify the array shape would imply a process all the more tedious as the number of sensors increases or the array extends. Since our purpose is to deploy hundreds of microphones over tens of meters, an efficient alternative process is required that should be fast and easy to implement. The process described in the paper answers these needs. It requires a limited number of acoustic experiments and provides the precise overall geometry of very large arrays.

Positioning sensors classically relies on the measurement of Time of Arrivals (TOA) or Time Differences of Arrival (TDOA). Both have been used for the geometric calibration of microphone arrays in Free Field. Performing this inverse problem needs to find both sensor and source positions. An overall review of geometric calibration was done in (Plinge *et al.*, 2016), and a more specific one on T(D)OA approaches by Wang *et al.* (Wang *et al.*, 2016, tab. 1). The strategy to choose depends on the available setup:

- the sources are known and synchronously recorded with the microphones. This setup is the most convenient and allows to measure TOAs by cross correlation between source and sensor signals;
- the hardware or setup prohibits the synchronous access to source signals. Thus, only cross correlations between microphone signals are available, which give access to TDOAs.

In both cases, the geometric inversion can be written as a non-linear problem in the least square form. It is solvable by an iterative algorithm (Gaubitch *et al.*, 2013; Hennecke and Fink, 2011; Ono *et al.*, 2009), with the risk of a convergence in a local minimum depending on the initial guess of the geometry. To overcome this drawback, an explicit closed-form solution can be computed (Crocco *et al.*, 2012; Le *et al.*, 2016), with certain specific conditions on the setup to be respected.

#### A. Related works

In free field experiments, the literature exhibits several *in situ* validations with a great number of microphones, by TOA (from 128 (Khanal *et al.*, 2013) up to 448 (Sachar *et al.*, 2005)) or TDOA (121 (Le *et al.*, 2016)). All of them evidence that the estimation accuracy increases as the number of sensors and sources increases. At first sight, these approaches fit naturally well for an application to very large arrays.

A Free Field model is generally an oversimplified assumption to describe real environments. It is mismatched by the presence of diffusors, refractors, and more importantly reflectors such as walls, ceiling and floor. Thus a multi-path propagation from source to sensors exists, and affects T(D)OA measurements that supposedly determines the direct path. If the measured path includes one or more reflections, it can lead to a strongly biased value of T(D)OA from the measured cross correlations (Perrodin *et al.*, 2012), as it is depicted in figure 1. In the present context, we define the measured T(D)OA as an outlier if it comes from a multipath rather than the direct propagation to the sensors. However, a straightforward least-square approach of geometric calibration intrinsically supposes that noise follows a normal distribution (Hennecke and Fink, 2011), so it is highly sensitive to outlying observations.

To alleviate this limitation, several solutions propose to bring robustness by RANdom SAMple Consensus (RANSAC) (Batstone *et al.*, 2016; Burgess *et al.*, 2015; Hennecke *et al.*, 2009; Zhayida *et al.*, 2014). It identifies inliers by an iterative scheme:

1. one subset of observations is randomly selected. Since the geometric inversion by T(D)OA is easily overdetermined with a sufficient number of sources and sensors, it still remains solvable with a subset of minimal cardinality. A first estimation of positions is given by a closed-form solution;
2. second, the unselected observations are iteratively included in the subset if they fit well with the estimated model from step 1. In the end we obtain the *consensus* set corresponding to inliers of the first estimation.

Steps 1 and 2 are repeated several times, and the preferred solution corresponds to the one with the *consensus* set of highest cardinality. Despite this method being efficient, it is based on a random subset sampling. The probability to obtain a *consensus* without outlier is a function of the number of trials (steps 1 and 2), the number of outliers, and the cardinality of the dataset (Fischler and Bolles, 1981); it converges to 1 with an infinity of iterations. Consequently, the RANSAC complexity increases as the number of sensors to locate is great.

## B. Our contribution

This paper proposes a computationally efficient method to tackle robustness with very large arrays. As mentioned previously, RANSAC has no upper bound on the number of iterations which fully guarantees a geometric calibration without outlier. Rather, we propose to replace the inlier subset selection based on random sampling, by an identification of the outliers amplitudes through a convex optimization.

To do so, the key point is to rely on an outlier-aware noise model, and suppose that the outlying subset is sparse. The related error amplitudes are considered as a second unknown matrix to be estimated, jointly with positions. Then a simple alternating optimization procedure can successively (a) compute the position matrix based on observations exempt from outlying errors, and (b) estimate the error amplitude matrix from the residuals, with a shrinkage operation enforcing sparsity. Pioneer studies on outlier-aware procedure were applied for robust linear regression (She and Owen, 2012), rank reduction by Robust Principal Component Analysis (RPCA) (Candès *et al.*, 2011), and interesting variants of low-rank and sparse decomposition of matrices applied to video processing (Bouwman *et al.*, 2017). It was used in Robust MultiDimensional Scaling (RMDS), in order to retrieve point positions from spurious distances between these points (Forero and Giannakis, 2012). RMDS has been applied to the geometric calibration of large microphone arrays (Vanwysberghe *et al.*, 2016). Indeed, in diffuse field the measure of coherence gives access to pairwise distance between microphones. However mismatch on diffuseness occurs in reverberant rooms: below the Schroeder cutoff frequency, the diffuse model becomes no more valid because the density of room modes is too weak (Taghizadeh *et al.*, 2014). Mismatch leads to outliers in observed distances (Taghizadeh *et al.*, 2015; Vanwysberghe *et al.*, 2016), and need to be removed through an automatic process to obtain the array geometry. That study evidenced the efficiency of an outlier-aware approach dedicated to large arrays. The present paper extends this previous work, and focuses on the case of a mismatched free field. In this way, the two papers together aim at proposing a unified outlier-aware framework of geometric calibration. Because of the alternating

minimization scheme, the solvers are iterative, and require an initialization with a first guess of the array geometry to reconstruct.

This paper is organized as follows. First, section II defines the problem and notations. Section III deals with the TOA problem as a specific case of MDS called MultiDimensional Unfolding (MDU) (Dokmani *et al.*, 2015), in which only distances from source to sensor points are known; as a result the Robust MDU (RMDU) is proposed, and derived from RMDS (Forero and Giannakis, 2012). Section IV focuses on TDOA calibration, where two strategies are considered. The first one (sec. IV A) relies on a low-rank and sparse decomposition of noisy TDOA observations (Velasco *et al.*, 2016). Then the low-rank part is the denoised data, which can be used with *any* TDOA calibration method from the state of the art. The second strategy (sec. IV B) proposes a joint iterative estimation of positions and outlier matrices; to our knowledge no such algorithm has been proposed in the literature. In section V, the robustness of the proposed methods are analyzed through Monte Carlo simulations, and compared to least-square baseline methods. Finally, section V presents a real experiment with a large-aperture array of 256 microphones. The results validate the interest of the proposed robust algorithms, and outperform iterative least-square methods in presence of free field model mismatches. The code for all the described algorithms is available online at <https://github.com/cvanwynsberghe>.

## II. PROBLEM SETUP

### A. Notations and setup

We consider the geometric problem in a 3-dimensional space. The array is made of  $M$  microphones, whose coordinates are given by the vectors  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M$ , or by the matrix  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_M] \in \mathbb{R}^{3 \times M}$ . One at a time,  $I$  fixed sources successively produce a signal at positions  $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_I$ , also written with the matrix  $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_I] \in \mathbb{R}^{3 \times I}$ . Finally let  $d_{im} = \|\mathbf{s}_i - \mathbf{x}_m\|$  the distance between the  $i$ -th source and the  $m$ -th microphone.

Notational convention throughout the paper uses lower (upper) bold letters for column vectors (matrices).  $\cdot^\top$  is the transposition operator,  $\|\cdot\|_F$  the Frobenius norm, and  $[\cdot]$  (resp.  $\begin{bmatrix} \cdot \\ \cdot \end{bmatrix}$ ) is horizontal (resp. vertical) matrix concatenation.  $\tilde{\mathbf{A}}$  refers to the estimation of the matrix  $\mathbf{A}$  obtained by a solver. Finally,  $\text{diag}(\mathbf{v})$  denotes the diagonal matrix whose diagonal is the vector  $\mathbf{v}$ , and  $\mathbf{1}_{M \times N}$  ( $\mathbf{0}_{M \times N}$ ) the  $M \times N$  matrix filled with ones (zeros).

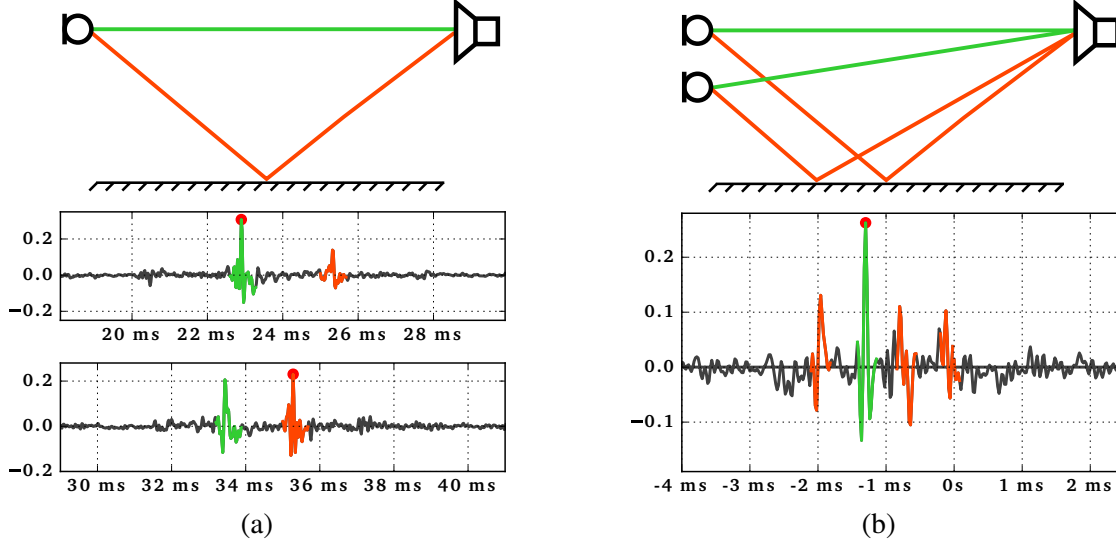


FIG. 1. The measured cross correlation reveals the direct (green) and indirect (orange) paths, the ground being the significant reflector in the field. The peaks (red dot) appearing at indirect paths are a source of outliers. Data from experiment in sec. VI. (a) TOA case; (b) TDOA case. (Color online)

### B. Use of cross correlation and impact of reflectors

The measured TOAs are denoted  $\tau_{im}$  which is the argument of the maximum peak of the cross correlation between source  $i$  and sensor  $m$ . In order to reduce the impact of reverberation, the Generalized Cross Correlation (GCC) is favored. It is a weighted inverse Fourier transform of the cross spectral density  $C_{im}$ , that is:

$$r_{\text{PHAT}}(\tau) = \int_{\mathbb{R}} \Psi(f) C_{im}(f) e^{j2\pi f \tau} df \quad (1)$$

where  $\Psi$  is a weighting function. One usual choice is the PHase Transform window (PHAT) (Knapp and Carter, 1976): it whitens the amplitude to preserve an identical contribution of all the components of the signal spectrum, ie  $\Psi(f) = \frac{1}{|C_{im}(f)|}$ . If the minimal and maximal frequencies of the emitted source are known,  $\Psi(f)$  should equal 0 outside the interval, otherwise incoherent noise contributes as much as the reference signal, and degrades SNR. Experiment shows that a chirp is an efficient choice for geometric calibration, thanks to its high signal to noise rate (Khanal et al., 2013; Sachar et al., 2005). Note that the TDOAs obtained from the  $i$ -th source emission (denoted  $\tau_{imn}$ ) are measured in the same fashion, by replacing  $C_{im}$  with the cross spectral density between sensors  $C_{mn}$  in GCC-PHAT formula (1).

Fig. 1 illustrates a result from real data, provided from the experiments described in sec. V. Here the path from the ground reflection appears in TOAs as a second delayed peak (orange) 2 ms after the direct path (green). If the second peak has the highest amplitude it leads to an outlying observation, with 68 cm of bias. The same situation appears with TDOA measurements. However false detection is not trivial since 4 path combinations exist with one reflector. In Fig. 1b, the peak corresponding to the direct path is the second one, yet the order depends on the relative positions of sensors and reflector. Moreover, the dataset cardinality is a function of  $M$ :  $MI$  for  $\{\tau_{im}\}$  set, and  $M(M-1)I/2 \approx M^2I/2$  for  $\{\tau_{imn}\}$ . So with a very large array with hundreds of sensors, the amount of data grows so fast that it makes visual checking unrealistic, even with TOAs.

### III. OUTLIER-AWARE TOA GEOMETRIC CALIBRATION

In a least square approach, estimating the microphone positions is done by minimizing the residual between the observations  $\tau_{im}$  and a theoretical TOA:

$$(\tilde{\mathbf{X}}, \tilde{\mathbf{S}}) = \underset{\mathbf{X}, \mathbf{S}}{\operatorname{argmin}} \sum_{i=1}^I \sum_{m=1}^M [\tau_{im} - t_{im}(\mathbf{X}, \mathbf{S})]^2 \quad (2)$$

where  $t_{im}$  is the modeled TOA between microphone  $m$  and source  $i$

$$t_{im}(\mathbf{X}, \mathbf{S}) = \frac{d_{im}}{c_0} = \frac{\|\mathbf{s}_i - \mathbf{x}_m\|}{c_0}, \quad (3)$$

and  $c_0$  is the known constant sound speed. Although  $\tilde{\mathbf{X}}$  is the estimation of interest, problem (2) also computes  $\tilde{\mathbf{S}}$  since sensors are located in relation to the sources. This least square approach corresponds to the maximum likelihood with observations modeled as  $\tau_{im} = t_{im} + \epsilon_{im}$  (Biswas and Thrun, 2004). Kuang *et al* showed that the TOA calibration problem in three dimensions is well-defined if  $I \geq 4$ ,  $M \geq 4$  and  $I + M \geq 10$  (Kuang *et al.*, 2013), so with a very large array the problem is widely overdetermined. In this paper, the algorithm from Khanal *et al.* (Khanal *et al.*, 2013) is chosen as the iterative baseline method. It is, to our knowledge, the only iterative approach experimentally validated with large arrays.

#### A. Robust Multi Dimensional Unfolding: a specific case of RMDS

The maximum likelihood problem (2) is sensitive to outliers because of the chosen noise model. Indeed, normal distribution rather characterizes small and homogeneous amplitudes of noise. Typically, it is consistent to depict errors due to the finite sampling of cross correlations, which approx-



imates the values of  $\tau_{im}$  as multiples of the sampling period. But errors with large amplitude due to reflections, as depicted in sec II B, severely exceed the extent of the chosen normal distribution.

Rather, the noise model is extended with a supplementary term:

$$\tau_{im} = t_{im} + \epsilon_{im} + o_{im} \quad (4)$$

where  $o_{im}$  has an arbitrary amplitude so that it can absorb the outlying errors. Let  $\mathbf{O} \in \mathbb{R}^{I \times M}$  be the matrix with elements  $o_{im}$ . This matrix is unknown, but is sparse if the number of outlying observations is much smaller than  $IM$ . In this way, the whole problem can be rewritten as

$$(\tilde{\mathbf{X}}, \tilde{\mathbf{S}}, \tilde{\mathbf{O}}) = \underset{\mathbf{X}, \mathbf{S}, \mathbf{O}}{\operatorname{argmin}} f_{\text{TOA}}(\mathbf{X}, \mathbf{S}, \mathbf{O}) \quad (5)$$

with

$$\begin{aligned} f_{\text{TOA}}(\mathbf{X}, \mathbf{S}, \mathbf{O}) &= \sum_{i=1}^I \sum_{m=1}^M [\tau_{im} - t_{im}(\mathbf{X}, \mathbf{S}) - o_{im}]^2 + \lambda \|\mathbf{O}\|_1 \\ &= \sum_{i=1}^I \sum_{m=1}^M [\tau_{im} - t_{im}(\mathbf{X}, \mathbf{S}) - o_{im}]^2 + \lambda \sum_{i=1}^I \sum_{m=1}^M |o_{im}| \end{aligned} \quad (6)$$

and  $\lambda$  a regularization parameter. This cost function is a combination of

- the term of residuals between observed and estimated values, discarded from errors  $o_{im}$ . Thus it consistently comes back to a least square approach exempt from outliers;
- a regularized term of  $\ell_1$ -norm on matrix  $\mathbf{O}$  that promotes its sparsity. Ideally the sparsest representation is obtained by a constraint with a  $\ell_0$  norm on  $\mathbf{O}$ , which refers the number of nonzero elements in the matrix. But it involves a combinatorial problem, making the algorithm intractable in high dimension. Rather, a representation with  $\ell_1$ -norm leads to a closest convex relaxation replacing the  $\ell_0$ -norm, which still preserves sparsity but making problem solvable in polynomial time. For more details, see (Zhang *et al.*, 2015, sec. III). Convex relaxation on  $\mathbf{O}$  written in the form as equation (6) is called Lasso.

Problem (5) can be seen as a specific case of MDS since  $c_0 \tau_{im}$  are distances, linking points from the source subset to the sensor subset, but inner distances from each subset are unknown. This case is known as MDU, see (DokmaniÄŸ *et al.*, 2015, fig. 4). An outlier-aware MDS was proposed in (Forero and Giannakis, 2012) with the same structure as problem (5). Based on this effort, the next section proposes the Robust MDU (RMDU), suitable for TOA geometric calibration.

## B. Update rules and algorithm

Forero *et al* propose to minimize the non-linear cost function  $f_{\text{TOA}}$  by alternating minimization (Forero and Giannakis, 2012), so that updates on  $\mathbf{O}$  and  $(\mathbf{X}, \mathbf{S})$  are independent. It relies on the Majorization-Minimization (MM) technique. The update sequence proves to make a non-increasing residual  $f_{\text{TOA}}(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}, \mathbf{O}^{(k)})$  as iteration  $k$  increases (Forero and Giannakis, 2012). MM relies on the minimization of a *surrogate* function which approximates  $f_{\text{TOA}}$  and holds 3 properties: (a) it majors  $f_{\text{TOA}}$ , (b) it is tangent to  $f_{\text{TOA}}$  at the current iteration point  $(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}, \mathbf{O}^{(k)})$ , and (c) it is quadratic as a function of parameters to update. Point (c) makes MM particularly easy to implement in the algorithm since the updates correspond to the condition where the surrogate derivatives are null. For thorough details on MM procedure, see (Sun *et al.*, 2017). The expression of the surrogate in the general case is derived in (Forero and Giannakis, 2012). Here, the resulting update rules for the RMDU case at iteration  $k$  are described so as to obtain  $(\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)}, \mathbf{O}^{(k+1)})$ .

The first step solves the sparse problem nested in a formulation named *Least Absolute Shrinkage and Selection Operator* (Lasso) (Tibshirani, 1996). As shown in equation (6)  $f_{\text{TOA}}$  can be split as a sum of the *IM* Lasso subproblems  $\arg\min_{o_{im}} (\tau_{im} - t_{im}(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}) - o_{im})^2 + \lambda |o_{im}|$ . As shown in (Zibulevsky and Elad, 2010), the minimizer of these subproblems are given by a proximal operation called soft-thresholding on the elements  $(\tau_{im} - t_{im}(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}))$ . Thus, the first update  $\mathbf{O}^{(k+1)}$  is achieved element-wise as:

$$o_{im}^{(k+1)} = S_{\lambda} (\tau_{im} - t_{im}(\mathbf{X}^{(k)}, \mathbf{S}^{(k)})) \quad (7)$$

where element-wise operator  $S_{\lambda}$  shrinks all the components of  $\mathbf{O}$  with amplitude smaller than  $\lambda/2$  by

$$S_{\lambda}(u) = \begin{cases} u - \lambda/2 & \text{for } u > \lambda/2 \\ 0 & \text{for } u \in [-\lambda/2; \lambda/2] \\ u + \lambda/2 & \text{for } u < -\lambda/2 \end{cases} \quad (8)$$

As expected, a stronger regularization ( $\lambda$  increasing) favors a higher sparsity of  $\mathbf{O}^{(k+1)}$ .

The update of the geometric variables is done simultaneously, and for convenience we define  $\mathbf{Z} \in \mathbb{R}^{3 \times (M+I)}$  that contains all unknown positions as  $\mathbf{Z} = [\mathbf{X} | \mathbf{S}]$ . In this way,  $\mathbf{Z}^{(k+1)}$  is the minimizer of the quadratic surrogate function from the MM procedure in RMDS. Demonstrated in (Forero and Giannakis, 2012), it writes:

$$\mathbf{Z}^{(k+1)} = \mathbf{Z}^{(k)} \mathbf{L}_1(\mathbf{O}^{(k+1)}, \mathbf{Z}^{(k)}) \mathbf{L}_1^{\dagger} \quad (9)$$

TABLE I. RMDU for Robust TOA geometric calibration

**Require:**  $\{\tau_{im}\}, \mathbf{Z}^{(0)} = [\mathbf{X}^{(0)}|\mathbf{S}^{(0)}], \lambda, \kappa$   
**for**  $k = 0, 1, 2, \dots$  **do**  
     Compute  $\mathbf{O}^{(k+1)}$  via eq. (7)  
     Compute  $\mathbf{Z}^{(k+1)} = [\mathbf{X}^{(k+1)}|\mathbf{S}^{(k+1)}]$  via eqs. (9)-(12)  
     **if**  $\frac{\|\mathbf{X}^{(k+1)} - \mathbf{X}^{(k)}\|_F}{\|\mathbf{X}^{(k)}\|_F} < \kappa$  **then**  
         break  
     **end if**  
**end for**  
**return**  $(\tilde{\mathbf{X}}, \tilde{\mathbf{S}}, \tilde{\mathbf{O}}) = (\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)}, \mathbf{O}^{(k+1)})$

This relation is closely related to the Guttman transform used in Smacof (de Leeuw and Mair, 2009) for regular MDS. To adapt this update for the MDU case, one should consider  $f_{\text{TOA}}$  as the initial RMDS cost function associated to  $\mathbf{Z}$ , weighted such that the inner distances from the source and sensor subsets are discarded (*ie* to 0), keeping only the distances linking the two subsets (weighted to 1). In this case, matrix  $\mathbf{L}_1(\mathbf{O}, \mathbf{Z})$  writes

$$\mathbf{L}_1(\mathbf{O}, \mathbf{Z}) = \text{diag} \left( \left[ \begin{array}{c} \frac{\mathbf{A}(\mathbf{O}, \mathbf{Z})^\top \mathbf{1}_{I \times 1}}{\mathbf{A}(\mathbf{O}, \mathbf{Z}) \mathbf{1}_{M \times 1}} \end{array} \right] \right) + \left[ \begin{array}{c|c} \mathbf{0}_{M \times M} & \mathbf{A}(\mathbf{O}, \mathbf{Z})^\top \\ \hline \mathbf{A}(\mathbf{O}, \mathbf{Z}) & \mathbf{0}_{I \times I} \end{array} \right] \quad (10)$$

with  $\mathbf{A}(\mathbf{O}, \mathbf{Z}) \in \mathbb{R}^{I \times M}$  the matrix filled with entries

$$\begin{aligned}
 [\mathbf{A}(\mathbf{O}, \mathbf{Z}=[\mathbf{X}|\mathbf{S}])]_{im} &= \frac{\tau_{im} - o_{im}}{t_{im}(\mathbf{X}, \mathbf{S})} \quad \text{if } \tau_{im} > o_{im} \text{ and } t_{im}(\mathbf{X}, \mathbf{S}) > 0 \\
 &= 0 \quad \text{otherwise}
 \end{aligned} \quad (11)$$

Finally,  $\mathbf{L}^\dagger$  is the Moore-Penrose pseudo inverse of the  $(M + I) \times (M + I)$  matrix  $\mathbf{L}$  :

$$\mathbf{L} = \sum_{m=1}^M \sum_{i=1}^I (\mathbf{e}_m - \mathbf{e}_{M+i})(\mathbf{e}_m - \mathbf{e}_{M+i})^\top \quad (12)$$

with  $\mathbf{e}_n \in \mathbb{R}^{(M+I) \times 1}$  the column vector filled with 1 at  $n$ -th element, and zeros elsewhere.

RMDU is summarized in algorithm I. Since it is iterative, it needs a first guess of the geometry matrices  $(\mathbf{X}^{(0)}, \mathbf{S}^{(0)})$ . The choice of this initial condition has an important impact due to the numerous local minima in  $f_{\text{TOA}}$ . Sensitivity to first guess was studied for classic MDS in (Borg

and Mair, 2017), and will be analyzed for RMDU by simulations in section V. Finally, algorithm I iterates until a sufficient convergence of the solution. Since the sensor positions are the unknowns of interest, the convergence criterion computes the relative difference between the two last updates of  $\mathbf{X}$ , and the algorithm stops when this criterion goes below the threshold  $\kappa$ .

#### IV. OUTLIER-AWARE TDOA GEOMETRIC CALIBRATION

This section deals with the second feasible approach of microphone localization. The experimental conditions are supposed to be identical to those in the previous section, except that only the signals from microphones are known. In practice, such an approach becomes interesting when the signal feeding the sources cannot be synchronously measured together with the sensor signals. In this way any source of opportunity from the ambient sound, such as people speeches or a wireless speakers, can be used (Zhayida *et al.*, 2014).

Let  $t_{imn}$  be the theoretical TDOA between microphones  $m$  and  $n$  related to source  $i$ :

$$t_{imn}(\mathbf{X}, \mathbf{S}) = t_{im}(\mathbf{X}, \mathbf{S}) - t_{in}(\mathbf{X}, \mathbf{S}) \quad (13)$$

$$= \frac{\|\mathbf{s}_i - \mathbf{x}_m\| - \|\mathbf{s}_i - \mathbf{x}_n\|}{c_0} \quad (14)$$

This relation is true for a phased array. With asynchronous arrays, internal delays exist on each channel, and add  $M$  extra unknowns. In the state of the art, authors generally consider ad hoc (*ie* asynchronous) arrays (Gaubitch *et al.*, 2013), but it is out of the scope of the present paper. As it becomes a fundamental hypothesis for the next proposed method, this paper focuses only on very large phased arrays, with no internal delays.

Similar to the formulation in equation (2), the least-square problem of TDOA geometric writes

$$(\tilde{\mathbf{X}}, \tilde{\mathbf{S}}) = \underset{\mathbf{X}, \mathbf{S}}{\operatorname{argmin}} \sum_{i=1}^I \sum_{m,n=1}^M [\tau_{imn} - t_{imn}(\mathbf{X}, \mathbf{S})]^2 \quad (15)$$

In order to bring robustness, the two proposed calibration methods commonly rely on the same outlier-aware noise model of TDOA, in a same fashion as equation (4). Indeed, it writes:

$$\tau_{imn} = t_{imn}(\mathbf{X}, \mathbf{S}) + \epsilon_{imn} + o_{imn} \quad (16)$$

where  $\epsilon_{imn}$  follows a centered distribution with a small standard deviation, and  $o_{imn}$  denotes the unknown outlier term of high amplitude.

### A. First approach: TDOA predenoising based on rank constraint

In this section, we propose a first approach which runs through two steps:

1. the measured TDOA set  $\{\tau_{imn}\}$  is processed through a predenoising algorithm in order to discard outliers. As a matter of fact, we will see that denoising occurs on each of the  $I$  subsets related to each calibration source:  $\{\tau_{1mn}\}, \dots, \{\tau_{imn}\}, \dots, \{\tau_{Imn}\}$ ;
2. second step solves geometric calibration with any state of the art algorithm, since identification of outliers and of geometry are separated.

For the  $i$ -th measured subset  $\{\tau_{imn}\}$  the noise model from equation (16) is rewritten in the matrix form:

$$\mathcal{T}_i = \mathbf{T}_i + \mathbf{E}_i + \mathbf{O}_i \quad (17)$$

whose entries are  $[\mathcal{T}_i]_{mn} = \tau_{imn}$ ,  $[\mathbf{T}_i]_{mn} = t_{imn}$ ,  $[\mathbf{E}_i]_{mn} = \epsilon_{imn}$  and  $[\mathbf{O}_i]_{mn} = o_{imn}$ . These four matrices belong to the set of  $M \times M$  skew-symmetric hollow matrices, noted  $\mathbb{A}_h^M$ . Velasco *et al.* demonstrated that the model from equation (17) leads to a structured decomposition (Velasco *et al.*, 2016). Indeed,  $\mathbf{O}_i$  is a sparse matrix and  $\mathbf{T}_i$  is such that  $\text{rank}(\mathbf{T}_i) = 2$ . With a large array, it is low-rank since  $M \gg 2$ . Note that, at this stage, the hypothesis of synchronous channels is crucial to guarantee this theoretical rank.

Note that the Robust Component Analysis (RPCA) exploits the similar low-rank + sparse structure of a matrix: although the resolution algorithms are different (Bouwman and Zahzah, 2014; Candès *et al.*, 2011; Huang *et al.*, 2012), RPCA is used to recover a low-rank matrix  $\mathbf{T}_i$  from data  $\mathcal{T}_i$  as it is done with regular PCA. Yet it is still efficient if a sparse matrix noise  $\mathbf{O}_i$  is inherently added in the data measurement.

In the present case, finding the decomposition is obtained by the following program:

$$\begin{aligned} (\tilde{\mathbf{T}}_i, \tilde{\mathbf{O}}_i) = & \underset{\mathbf{T}_i \in \mathbb{A}_h^M, \mathbf{O}_i \in \mathbb{A}_h^M}{\text{argmin}} \quad \|\mathbf{T}_i + \mathbf{O}_i - \mathcal{T}_i\|_F^2 \\ \text{s.t.} \quad & \|\mathbf{O}_i\|_0 < 2K \text{ and } \text{rank}(\mathbf{T}_i) = 2 \end{aligned} \quad (18)$$

where  $K$  is the sparsity of  $\mathbf{O}_i$  such that  $2K \ll M^2$ . Generally RPCA aims at finding an approximation of  $\mathcal{T}_i$  in a reduced dimension, with no accurate knowledge of the rank value. Conversely, it is given in the problem (18) as a known constraint. Zhou *et al.* (Zhou and Tao, 2011) show that

a solution is given by alternated projections, minimizing problem (18) splits into two steps:

$$\begin{cases} \mathbf{T}_i^{(k+1)} = \underset{\mathbf{T}_i \in \mathbb{A}_h^M}{\operatorname{argmin}} \left\| \mathbf{T}_i + \mathbf{O}_i^{(k)} - \mathcal{T}_i \right\|_F^2 & \text{s.t. rank}(\mathbf{T}_i) = 2 \\ \mathbf{O}_i^{(k+1)} = \underset{\mathbf{O}_i \in \mathbb{A}_h^M}{\operatorname{argmin}} \left\| \mathbf{T}_i^{(k+1)} + \mathbf{O}_i - \mathcal{T}_i \right\|_F^2 & \text{s.t. } \|\mathbf{O}_i\|_0 < 2K \end{cases} \quad (19)$$

If a projection onto the subset of TDOAs matrices exists (giving  $\mathbf{T}_i^{(k+1)}$  update), then the alternated projection framework given in (Lewis and Schreier, 2013) converges to a local minimum (Zhou and Tao, 2011). Velasco *et al.* (Velasco *et al.*, 2016) prove that the first low-rank update is given by:

$$\mathbf{T}_i^{(k+1)} = \frac{(\mathcal{T}_i - \mathbf{O}_i^{(k)}) \mathbf{1}_{M \times M} + \mathbf{1}_{M \times M} (\mathcal{T}_i - \mathbf{O}_i^{(k)})}{M} \quad (20)$$

Afterwards  $\mathbf{O}_i^{(k+1)}$  is computed by projection onto the  $2K$ -sparse matrix manifold:

$$\mathbf{O}_i^{(k+1)} = \mathcal{H}_{2K} \left( \mathcal{T}_i - \mathbf{T}_i^{(k+1)} \right) \quad (21)$$

$\mathcal{H}_{2K}(\mathbf{A})$  being the hard-thresholding operator such that:

$$[\mathcal{H}_{2K}(\mathbf{A})]_{mn} = \begin{cases} [\mathbf{A}]_{mn} & \text{for the } 2K \text{ entries in } \mathbf{A} \text{ with highest amplitude} \\ 0 & \text{elsewhere} \end{cases} \quad (22)$$

This update is also proposed in the TDOA denoising algorithm by Velasco *et al.* (Velasco *et al.*, 2016). Thus, it needs a careful choice of the cardinality  $K$  for *each* of the  $I$  outlier matrices  $\mathbf{O}_i$ . However  $K$  is difficult to guess in practice: it directly depends on the reflectors in the experiment field – *eg* their number, position and impedance. Moreover, a measured matrix  $\mathcal{T}_i$  contains the residual noise  $\mathbf{E}_i$  in all entries. So for any value  $K$ , hard-thresholding on  $\mathbf{O}_i^{(k)}$  always results in a matrix with  $2K$  non-null entries at convergence. Indeed, if  $K$  is higher than the optimal value, the noise entries in  $\mathbf{E}_i$  with the highest amplitude leak into  $\mathbf{O}_i$ , as revealed by Zhou *et al* (Zhou and Tao, 2013). Consequently it becomes difficult to make a good empirical choice of this parameter with real data.

### 1. Relaxation and update rules

Rather, based on (Zhou and Tao, 2013) we propose to transform the problem (18) with a cardinality constraint by a relaxation using  $\ell_1$ -norm penalty:

$$\begin{aligned} (\tilde{\mathbf{T}}_i, \tilde{\mathbf{O}}_i) = & \underset{\mathbf{T}_i \in \mathbb{A}_h^M, \mathbf{O}_i \in \mathbb{A}_h^M}{\operatorname{argmin}} \quad \|\mathbf{T}_i + \mathbf{O}_i - \mathcal{T}_i\|_F^2 + \lambda \|\mathbf{O}_i\|_1 \\ \text{s.t.} \quad & \operatorname{rank}(\mathbf{T}_i) = 2 \end{aligned} \quad (23)$$

so the cost function is reduced to a Lasso problem with rank constraint. The key point is to change the proximal mapping onto the sparse matrix set, by using  $S_\lambda$  (soft-thresholding) rather than  $\mathcal{H}_{2K}$  (hard-thresholding). Indeed, parameter  $\lambda$  sets the amplitude bound between residuals and outlying errors. It automatically sets the  $K$ -sparsity of  $\mathbf{O}_i$ , and thus prevents *leakage* from  $E_i$ . Besides, tuning  $\lambda$  is easier than  $K$  in practice, since it can be empirically determined, *eg* with a L-curve (Vanwysberghe *et al.*, 2016). This modification results in a new update of  $\mathbf{O}_i^{(k+1)}$

$$o_{imn}^{(k+1)} = S_\lambda \left( \tau_{imn} - t_{imn}^{(k+1)} \right) \quad (24)$$

with  $S_\lambda$  as defined in equation (8). Note that this relaxed program does not change the low-rank update  $\mathbf{T}_i^{(k+1)}$ , given by equation (20).

### 2. Final algorithm

Predenoising requires to run algorithm II for each of the  $I$  matrices  $\mathcal{T}_i$ . The denoised estimations  $\tilde{\mathbf{T}}_i$  are then stacked together, and form a set of  $IM^2$  values, which can be the input to any known TDOA geometric calibration. The overall approach is summarized in algorithm III. Although that point seems appealing, the whole approach consists of a processing chain with  $I + 1$  independent algorithms. Besides, a good empirical choice of regularization in predenoising would require the tuning of  $I$  parameters  $\lambda$ . It may become a tedious task with a lot of calibration sources, even though we experimentally notice that one choice of  $\lambda$  for all remains efficient – cf sec. V.

It evidences the interest of proposing a unified approach. To that end, the next section proposes the second outlier-aware TDOA calibration procedure, by jointly estimating the outliers and geometry in one go: rather than chaining  $I + 1$  algorithms which separately give the matrices  $\tilde{\mathbf{O}}_i$  and  $\tilde{\mathbf{X}}$ , we will see that one algorithm can nest all estimations in a unique alternating minimization.

TABLE II. TDOA denoising by matrix decomposition

**Require:**  $\lambda, \kappa$ , Initialize  $\mathbf{O}_i^{(0)}$  to zero and  $\mathbf{T}_i^{(0)} = \mathcal{T}_i$

**for**  $k = 0, 1, 2, \dots$  **do**

    Compute  $\mathbf{T}_i^{(k+1)}$  via (20)

    Compute  $\mathbf{O}_i^{(k+1)}$  by soft-thresholding via (24)

**if**  $\frac{\|\mathbf{T}_i^{(k+1)} - \mathbf{T}_i^{(k)}\|_F^2}{\|\mathbf{T}_i^{(k)}\|_F^2} < \kappa$  **then**

        break

**end if**

**end for**

**return**  $(\tilde{\mathbf{T}}_i, \tilde{\mathbf{O}}_i) = (\mathbf{T}_i^{(k+1)}, \mathbf{O}_i^{(k+1)})$

TABLE III. TDOA geometric calibration – first approach

**Require:**  $\mathcal{T}_1, \dots, \mathcal{T}_i, \dots, \mathcal{T}_I$

**for**  $i = 1, \dots, I$  **do**

    Apply algorithm II with input  $\mathcal{T}_i$ , and get  $\tilde{\mathbf{T}}_i$

**end for**

With denoised input  $\tilde{\mathbf{T}}_1 \dots, \tilde{\mathbf{T}}_I$ , apply *one* TDOA geometric calibration

**return**  $\tilde{\mathbf{X}}$

### B. Second approach: identifying geometry and outliers at the same time

In this section, the problem formulation is inspired by RMDU: it puts the three unknowns  $(\mathbf{X}, \mathbf{S}, \mathbf{O})$  into the same cost function. Like equation (6) we write a Lasso problem fitting with the model (16) as:

$$(\tilde{\mathbf{X}}, \tilde{\mathbf{S}}, \tilde{\mathbf{O}}) = \underset{\mathbf{X}, \mathbf{S}, \mathbf{O}}{\operatorname{argmin}} f_{\text{TDOA}}(\mathbf{X}, \mathbf{S}, \mathbf{O}) \quad (25)$$

where

$$f_{\text{TDOA}}(\mathbf{X}, \mathbf{S}, \mathbf{O}) = \sum_{i=1}^I \sum_{m,n=1}^M [\tau_{inm} - t_{inm}(\mathbf{X}, \mathbf{S}) - o_{inm}]^2 + \lambda \sum_{i=1}^I \sum_{m,n=1}^M |o_{inm}| \quad (26)$$

and  $\mathbf{O}$  becomes a tensor that contains all TDOAs matrices:  $\mathbf{O} = [\mathbf{O}_1, \dots, \mathbf{O}_i, \dots, \mathbf{O}_I]$ . Unlike previous section, all the outliers are included in the same cost function. Hence the sparsity structure of tensor  $\mathbf{O}$  is jointly controlled through one same regularization.



As in (Forero and Giannakis, 2012), function  $f_{\text{TDOA}}$  can be iteratively minimized by splitting in two sub-problems:

$$(\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)}) = \underset{\mathbf{X}, \mathbf{S}}{\operatorname{argmin}} f_{\text{TDOA}}(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}, \mathbf{O}^{(k)}) \quad (27)$$

$$\mathbf{O}^{(k+1)} = \underset{\mathbf{O}}{\operatorname{argmin}} f_{\text{TDOA}}(\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)}, \mathbf{O}^{(k)}) \quad (28)$$

In this way, deriving the updates becomes simpler: the first step of one iteration minimizes the least-square term in  $f_{\text{TDOA}}$  through an MM approach. Such updates have been derived by Ono *et al.* (Ono *et al.*, 2009) for the least-square case without outlier noise model. Based on their solution, we now start to solve each of the two sub-problems (27) and (28).

### 1. Update rules

We recall that the MM technique minimizes the non-convex function  $f_{\text{TDOA}}$  by iteratively building a quadratic substitute function  $g_{\text{TDOA}}$ . It should follow 3 properties: (a) majoring  $f_{\text{TDOA}}$ , ie  $g_{\text{TDOA}} \geq f_{\text{TDOA}}$ , (b) being tangent to  $f_{\text{TDOA}}$  at the current point  $(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}, \mathbf{O}^{(k)})$ , (c) being quadratic for all geometric vectors  $\mathbf{x}_m$  and  $\mathbf{s}_i$ . Ono *et al.* show that surrogate

$$g_{\text{TDOA}}(\mathbf{X}, \mathbf{S}, \boldsymbol{\mu}, \mathbf{v}, \mathbf{e}) = \sum_{i=1}^I \sum_{m,n=1}^M \|\mathbf{s}_i - \mathbf{x}_m - \mathbf{e}_{im} \mu_{imn}\|^2 + \|\mathbf{s}_i - \mathbf{x}_n - \mathbf{e}_{in} v_{imn}\|^2 \quad (29)$$

matches conditions (a), (b) and (c). The auxiliary tensors  $\boldsymbol{\mu}$ ,  $\mathbf{v}$  respectively contain variables  $\mu_{imn}$ ,  $v_{imn}$ , and  $\mathbf{e}$  contain vectors  $\mathbf{e}_{im}$  and  $\mathbf{e}_{in}$ . The tangency condition (b) holds if these three tensors are such that:

$$\mathbf{e}_{im} = \frac{\mathbf{s}_i^{(k)} - \mathbf{x}_m^{(k)}}{\|\mathbf{s}_i^{(k)} - \mathbf{x}_m^{(k)}\|}, \quad \mathbf{e}_{in} = \frac{\mathbf{s}_i^{(k)} - \mathbf{x}_n^{(k)}}{\|\mathbf{s}_i^{(k)} - \mathbf{x}_n^{(k)}\|} \quad (30)$$

and

$$\mu_{imn} = \|\mathbf{s}_i^{(k)} - \mathbf{x}_m^{(k)}\| + \frac{c_0}{2} \epsilon_{imn}, \quad v_{imn} = \|\mathbf{s}_i^{(k)} - \mathbf{x}_n^{(k)}\| - \frac{c_0}{2} \epsilon_{imn} \quad (31)$$

where  $\epsilon_{imn}$  denotes the residual term in function  $f_{\text{TDOA}}$ , at indexes  $(i, m, n)$  of the sum. At the  $k$ -th iteration it writes:

$$\epsilon_{imn} = \tau_{imn} - t_{imn}(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}) - o_{imn}^{(k)} \quad (32)$$

At this stage, this term is augmented with  $o_{imn}$  to hold the majoring property (a) of the surrogate.

Finally, finding the minimum of  $g_{\text{TDOA}}$  is trivial since the quadratic function is minimal where all partial derivatives  $\frac{\partial g_{\text{TDOA}}}{\partial \mathbf{x}_n}$  and  $\frac{\partial g_{\text{TDOA}}}{\partial \mathbf{s}_i}$  vanish. Then it can be proved that:

$$\mathbf{x}_n^{(k+1)} = \frac{1}{I} \left( \sum_{i=1}^I \mathbf{s}_i^{(k)} - \frac{1}{M} \mathbf{e}_{in} \sum_{m=1}^M v_{imn} \right) \quad (33)$$

$$\mathbf{s}_i^{(k+1)} = \frac{1}{M} \left( \sum_{m=1}^M \mathbf{x}_m^{(k)} + \frac{1}{M} \mathbf{e}_{im} \sum_{n=1}^M \mu_{imn} \right) \quad (34)$$

These updates are used to obtain the pair  $(\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)})$ . Lastly, the second stage is the solution of the Lasso sub-problem, and is obtained via a soft-thresholding of the difference between the measured and theoretical TDOAs

$$o_{imn}^{(k+1)} = S_{\lambda} (\tau_{imn} - t_{imn}(\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)})) \quad (35)$$

as it was employed in the previous proposed algorithms.

## 2. Final algorithm

The update sequence is described in algorithm IV named R-Ono. At each iteration the cost function is non-increasing since

$$f_{\text{TDOA}}(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}, \mathbf{O}^{(k)}) = g_{\text{TDOA}}(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}, \boldsymbol{\mu}, \boldsymbol{\nu}, \mathbf{e}) \quad (36)$$

$$\geq g_{\text{TDOA}}(\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)}, \boldsymbol{\mu}, \boldsymbol{\nu}, \mathbf{e}) \quad (37)$$

$$\geq f_{\text{TDOA}}(\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)}, \mathbf{O}^{(k)}) \quad (38)$$

$$\geq f_{\text{TDOA}}(\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)}, \mathbf{O}^{(k+1)}) \quad (39)$$

where (37) and (39) respectively hold because of (27) and (28); steps (36) and (38) remain valid because of the tangency and majorization of  $f_{\text{TDOA}}$  by surrogate  $g_{\text{TDOA}}$ . Given the fact that the sequence of values  $f_{\text{TDOA}}(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}, \mathbf{O}^{(k)})$  is non-increasing, and that  $f_{\text{TDOA}}$  is bounded below, the cost function value guarantees to converge. For further proofs of the convergence of the sequence of updates  $(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}, \mathbf{O}^{(k)})$ , see (Sun *et al.*, 2017, sec II). For consistency, we keep the same stopping criterion as used previously, by analyzing convergence of the estimation  $\tilde{\mathbf{X}}$ .

As a result, one may conceive R-Ono as an augmentation of method in (Ono *et al.*, 2009) with the noise model (17), in the same way as RMDS (Forero and Giannakis, 2012) augments Smacof (de Leeuw and Mair, 2009). Indeed, by discarding all  $o_{imn}$  variables, R-Ono boils down to the original algorithm.

TABLE IV. approach 2 (R-Ono): proposed robust geometric calibration by TDOA

**Require:**  $\{\tau_{imm}\}, \mathbf{X}^{(0)}, \mathbf{S}^{(0)}, \lambda, \kappa$   
Initialize tensor  $\mathbf{O}^{(0)} = \mathbf{0}_{I \times M \times M}$   
**for**  $k = 0, 1, 2, \dots$  **do**  
    Compute auxiliary tensors  $\mathbf{e}, \boldsymbol{\mu}, \mathbf{v}$  via eqs. (30)-(32)  
    Update  $\mathbf{X}^{(k+1)}$  via eq. (33)  
    Update  $\mathbf{S}^{(k+1)}$  via eq. (34)  
    Update  $\mathbf{O}^{(k+1)}$  via eq. (35)  
    **if**  $\frac{\|\mathbf{X}^{(k+1)} - \mathbf{X}^{(k)}\|_F^2}{\|\mathbf{X}^{(k+1)}\|_F^2} < \kappa$  **then**  
        break  
    **end if**  
**end for**  
**return**  $(\tilde{\mathbf{X}}, \tilde{\mathbf{S}}, \tilde{\mathbf{O}}) = (\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)}, \mathbf{O}^{(k+1)})$

## V. NUMERICAL VALIDATION OF PROPOSED METHODS

This first experimental study investigates the validity and limitations of the 3 proposed methods. Two baseline methods are chosen for the comparative study: FrSM (Khanal *et al.*, 2013) for TOA, and Ono (Ono *et al.*, 2009) for TDOA. Two aspects are considered. First performance is assessed for different quantities of outliers in the data. Second, because of the iterative nature of all the algorithms, the impact of inaccurate first guess  $(\mathbf{X}^{(0)}, \mathbf{S}^{(0)})$  is analyzed to evaluate the sensitivity to local minima.

### A. Monte-Carlo simulation

We consider the geometrical setup given in Fig. 4b with  $I = 24$  sources and  $M = 256$  sensors, from which the theoretical T(D)OAs are calculated. The dataset is successively contaminated by 0, 1, 3, 5, 8 and 11 % of outliers. They are generated by computing the path reflected by the ground. The initial guesses  $(\mathbf{X}^{(0)}, \mathbf{S}^{(0)})$  are calculated by summing a random error to ground truth positions. This error follows a centered normal distribution of standard deviation going from 0 to 1 m, with a 10 cm step.

A Monte Carlo (MC) experiment is achieved by simulating 500 realizations per configuration. Each realization generates a new set of contaminated data: for TOAs, we randomly pick  $(i, m)$  index pairs from reflected path. Likewise,  $(i, m, n)$  are randomly chosen for TDOAs but reflection combinations potentially lead to 3 spurious peaks – cf. Fig. 1b. Again, we randomly draw  $\tau_{imn}$  from one of these 3 feasible values.

Finally, although the complexity of the TDOA algorithms  $\mathcal{O}(IM^2)$  remains tractable for one process, we need to reduce  $M$  to 64 for the MC simulation. One every four microphones is selected to keep the large aperture. The algorithm parameters are the following: the convergence criterion is similar for all methods, and is set at  $\kappa = 10^{-6}$ ; the FrSM gradient step is  $\alpha = 1.0$ ; the soft regularization parameter is fixed at  $\lambda = 6 \times 10^{-5}$  s for RDMU, and  $\lambda = 1.2 \times 10^{-5}$  s for the two TDOA methods (predenoising and R-Ono).

## B. Results

The efficiency is analyzed by estimating the error between  $\tilde{\mathbf{X}}$  and the ground truth. The methods only determine the relative positions of the microphones and sources, so an optimal rotation and translation are first calculated to align the estimated geometry upon the ground truth one. The optimal transformation is obtained by solving the so called Orthogonal Procrustes problem (Gower and Dijksterhuis, 2004, chap. 4), thanks to the Kabsch algorithm (Kabsch, 1976).

In Fig. 2, the averages of each 500 MC series are plotted according to the initial guess error (gray scale) and outlier density. It reveals a particular difference in the initial guess between the TOA and TDOA families. Indeed, its impact is negligible on FrSM but significantly appears in RMDU beyond 60 cm of standard deviation error. However the impact from spurious data in FrSM is clearly prominent since the calibration error regularly increases up to 34 cm with 11 % of outliers in TOAs. On the other hand, the TDOAs methods show a significant disparity, thus minimization is particularly sensitive to the non-convexity of cost function  $f_{\text{TDOA}}$ , with or without the outlier-aware noise model. Still, for standard deviations smaller than 40 cm in Fig. 2b, the calibration errors are also notably reduced with robust methods. With 11 % of outliers, the baseline algorithm obtains 5.9 cm of calibration error, whereas predenoising and R-Ono respectively reach 1.6 cm and 1.8 cm.

In order to analyze the homogeneity in robustness, MC results are plotted with boxplots in Fig. 3. They still show the calibration errors, but only as a function of the number of outliers. To focus on tendencies due to the effect of spurious measurements, each boxplot contains the 2000 tests with the initial guess errors between 10 and 40 cm. For TOA, comparing FrSM and RMDU in Fig. 3a evidences a higher dispersion of RMDU. The presence of calibration errors beyond whiskers

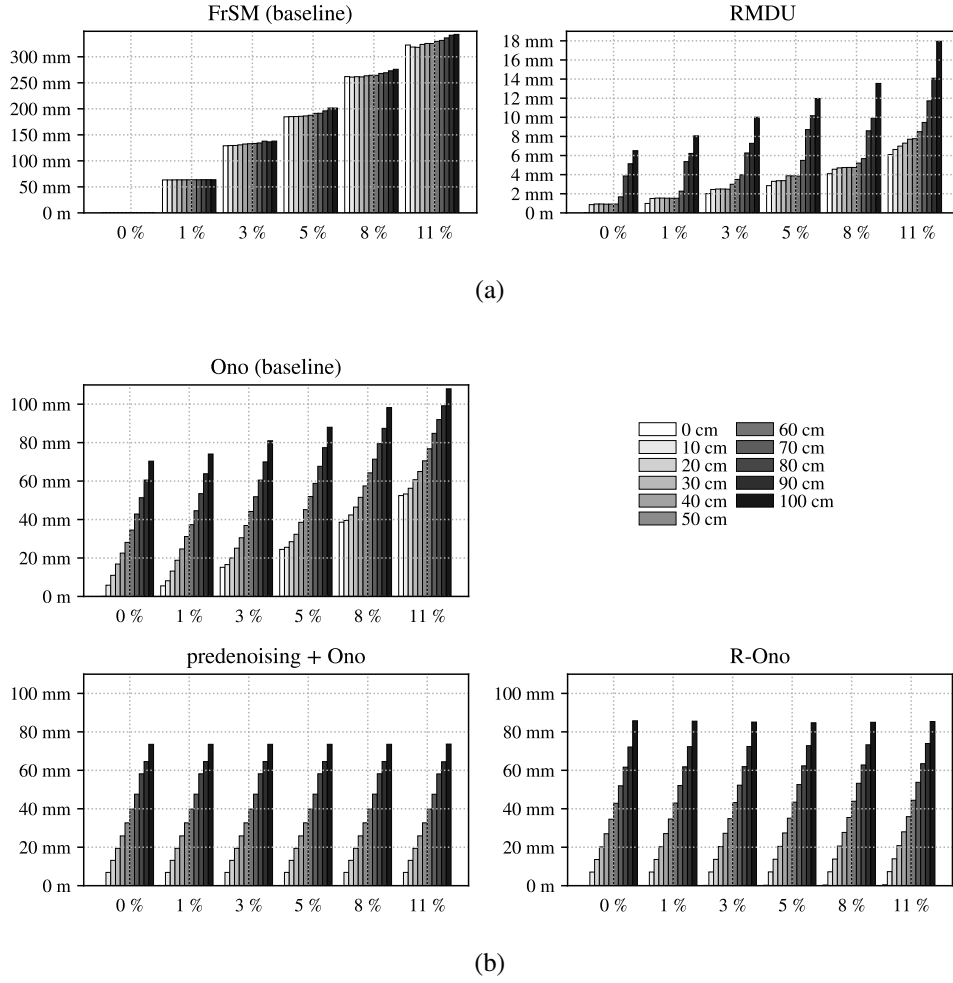


FIG. 2. Monte-Carlo simulations: average calibration error on  $\tilde{\mathbf{X}}$  on 500 realizations for each bar. 0-11% of outliers simulated from ground reflection with the geometry setup in Fig. 4b. Standard deviation error on initial guess ( $\mathbf{X}^{(0)}, \mathbf{S}^{(0)}$ ) goes from 0 to 1 m (gray scale). (a) TOA methods; (b) TDOA methods. (Color online)

("+" signs) shows that the robust approach sometimes fail to identify spurious data. Compared to 2000 realizations, such situation is seldom below 3% of outliers, but becomes more and more significant above. At 11%, the greatest error reaches 5.3 cm, nevertheless note that the worst cases with RMDU remain inferior to the best case with FrSM. As a result, RMDU is always the more efficient approach.

For TDOA in Fig. 3b, the trends also depict the constant increase of error with the pure least-square minimization (Ono baseline), up to 5.9 cm for 11 %. Conversely, the two proposed outlier-aware methods keep a stable performance whatever the percentage is: with a predenoising the

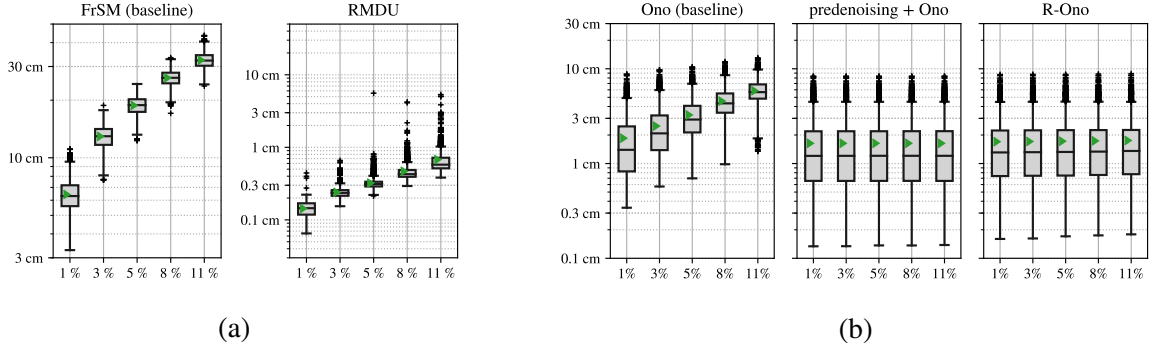


FIG. 3. Monte-Carlo simulations: boxplots of calibration errors on  $\tilde{\mathbf{X}}$ . 0-11% of outliers simulated from ground reflection with the geometry setup in Fig. 4b. Green arrows indicates the average, black mid-box segments the median, and blue box contains second and third quartiles. Whiskers represent the extreme quartiles, but their maximal size is thresholded to 1.5 times the related box size; realizations beyond this threshold are plotted with + sign. (a) TOA methods; (b) TDOA methods. Each box-and-whisker contains 2000 realizations, with error on initial guess going from 10 to 40 cm. (Color online)

average on 2000 realizations (green arrow) is 1.6 cm, slightly under the 1.8 cm average with R-Ono.

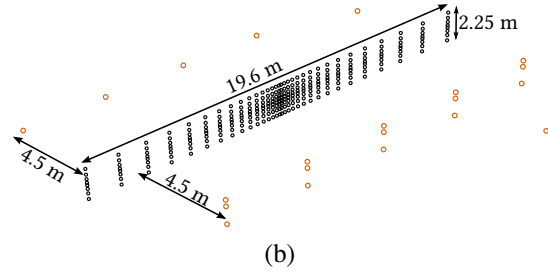
Globally, the numerical experiments reveal clear trends. First, TDOA approaches are more sensitive to the initial condition than TOA ones, so it should be considered depending on the reliability of the first guess. For that reason, one should favor the TOA approach if it is feasible in practice. Second, the MC simulation statistically validates the robustness of the three proposed methods (RMDU, predenoising and R-Ono), since all outperform the least-square iterative approaches. In spite of initial guess sensitivity, TDOA methods prove to keep more steady than TOA one when the number of spurious data varies.

## VI. RESULTS IN A REAL EXPERIMENT

This section compares all the proposed geometric calibration algorithms in a real experiment with a very large array of microphones. To do so, a phased array of 256 MEMS microphones is set up outside, as shown in Fig. 4. It uses the *Megamicros* acquisition system whose hardware is described in (Vanwynsberghe *et al.*, 2015). This outdoor experiment initially aims at achieving vehicle pass-by tests by use of acoustic imaging (Leiba *et al.*, 2017). Due to the different vehicle



(a)



(b)

FIG. 4. Experimental setup: *Megamicros* microphone array with 256 elements. (a) view of the outdoor experiment field; (b) geometry layout: microphones (black) and sources (orange) . (Color online)

sizes and speeds, the microphones are distributed on  $19.6\text{m} \times 2.25\text{m}$ . Here the test track is assumed to be the most intrusive reflector.

One after the other, each of the 24 sources successively emit an exponential chirp signal from 150 Hz to 12 kHz during 10 s. Depicted in Fig. 4b, their positions surround the array but also vary in height from 20 cm to 2.3 m so that it spans the 3D space. Both sources and microphones signal are synchronously acquired by the *Megamicros* system. GCC-PHATs are computed from time series sampled at 50 kHz. The cross power densities are obtained by the Welch method with blocks of 10240 samples, *ie* by averaging 48 blocks of 0.2 s. Examples of GCC-PHATs in this experiment are given in Fig. 1, and evidence erroneous estimated T(D)OAs because of the ground reflections.

The three outlier-aware calibration methods are similarly regularized by the soft parameter  $\lambda$ ; the latter is empirically chosen by the L-curve  $\|\mathbf{O}\|_0(\lambda)$ . The procedure is thoroughly described in (Vanwynsberghe *et al.*, 2016, sec.v-C.) for the LRMDs algorithm, in a geometric calibration in diffuse field. For the present experiment, the L-curves are plotted in Fig. 5, and the chosen value is located at the knee. As depicted, we select  $\lambda = 8.83 \times 10^{-5}$  s ( $\leftrightarrow 3.0$  cm for  $c_0 = 340$  m.s $^{-1}$ ) for TOA (RMDU), and  $\lambda = 7.6 \times 10^{-4}$  s ( $\leftrightarrow 26$  cm) for TDOA (predenoising & R-Ono).

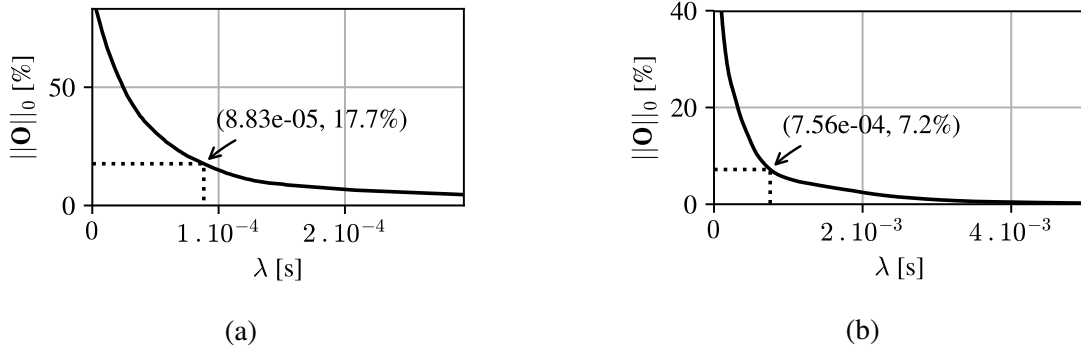


FIG. 5. Geometric calibration in the real experiment: choice of the regularization parameter  $\lambda$  by L-curve, for (a) RMDU with TOAs and (b) R-Ono with TDOAs. (Color online)

The array consists of 32 bars of 8 microphones, fixed on a rectangular frame. The structure extent is such that inherent play makes the knowledge of ground truth  $\mathbf{X}$  unfeasible with a millimeter accuracy. However the relative positions of the 8 microphones on each bar are precisely known. As a matter of fact, we evaluate the calibration error by a Procrustes analysis per bar. That is,  $m$ -th error  $\epsilon_m = \|\mathbf{x}_m - \tilde{\mathbf{x}}_m\|$  is computed after aligning the corresponding microphone's bar upon the theoretical geometry. The Kabsch method (Kabsch, 1976) is classically used to align shapes in the Procrustes analysis.

The performances of the methods are compared in Fig. 6. The 256 errors  $\epsilon_m$  are plotted by a histogram, with the root mean square error (RMSE) in red and the mean average error (MAE) in green. First it reveals that the outlier-aware model enables to outperform the baseline methods. It reduces the MAE from 4.8 cm to 1.4 cm in the TOA approach. The error decreases as well with the TDOA methods, going from 5.5 cm to 4.8 cm (predenoising) and 3.1 cm (R-Ono). Unlike MC experiments, R-Ono seems more efficient than TDOA denoising. Beyond the average performance indicated by the MAE metric, the RMSE as well as the histogram bars unveil a significant error variance from the baseline methods, whereas RMDU shows the best result, followed by R-Ono.

Globally TOA approaches outperform TDOA ones here. According to MC simulations, both are mainly differentiated by their sensitivity to accuracy on initial positions ( $\mathbf{X}^{(0)}, \mathbf{S}^{(0)}$ ). To our knowledge, that cause justifies the difference of accuracy, and makes RMDU the most efficient method.

Consequently the three proposed methods improve the estimation in this real experiment. The results reveal the importance of using a robust approach of geometric calibration. Indeed, one may suggest that the Free Field hypothesis in this outdoor place is mainly mismatched by ground



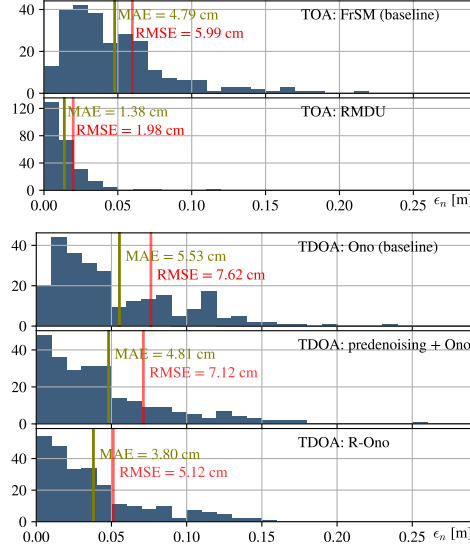


FIG. 6. A comparison of all algorithms for the geometric calibration of 256 microphones with 24 sources *in situ* (cf. Fig. 4). Histogram of 256 errors per microphone between ground truth and estimated positions. (Color online)

reflection – and also because of the array structure. Still, in this relatively convenient setting, baseline methods rapidly seem inappropriate for very large arrays.

## VII. CONCLUSION

This paper deals with the two geometric calibration approaches based on cross correlation: by TOA or TDOA. It focuses on the well-known limit of Free Field assumption. A real environment induces different mismatches like diffraction, and more importantly reflections from stiff surfaces. The latter generates outliers in measured T(D)OAs, thus estimating microphone positions becomes challenging. Indeed experiments show that baseline methods fail to retrieve the geometry of large arrays, *ie* with a great number of sensors and a large aperture.

To tackle robustness, we explore a unified outlier-aware approach by an extended noise model, and integrate it to design 3 iterative algorithms: one for TOA and two for TDOA. The sparsity assumption of strongly biased T(D)OAs regularizes problems by Lasso. The solution results in finding both unknowns: geometry and outlier amplitudes. Simulated and real experiments validate the efficiency of the proposed methods, since the estimated microphone positions are significantly more accurate compared with baseline methods.

The second motivation of this paper was to seek a unified framework: it associates simple but efficient models of geometry and noise. In future works, it could be interesting to implement and compare RANSAC methods (Batstone *et al.*, 2016; Zhayida *et al.*, 2014). In particular, we speculate a compromise between robustness and time complexity with very large array ( $M \gg 1$ ), since RANSAC is computationally more expensive. Finally, the methods could be analyzed in more complex configurations, for instance an indoor environment set up (with walls, ceiling and furniture), or a deployment of a larger array with more sensors.

## ACKNOWLEDGMENTS

This research benefits from the support of the chair “MoUViE” (Mobility and Life Quality in Urban Environment), carried by the UPMC Foundation and sponsored by donors (PSA Peugeot-Citroën, RENAULT).

<sup>1</sup>test

- Batstone, K., Oskarsson, M., and Åström, K. (2016). “Robust time-of-arrival self calibration and indoor localization using Wi-Fi round-trip time measurements,” in *Proc. IEEE Int. Conf. Communications Workshops (ICC)*, pp. 26–31, doi: [10.1109/ICCW.2016.7503759](https://doi.org/10.1109/ICCW.2016.7503759).
- Biswas, R., and Thrun, S. (2004). “A passive approach to sensor network localization,” in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS 2004)*, Vol. 2, pp. 1544–1549 vol.2, doi: [10.1109/IROS.2004.1389615](https://doi.org/10.1109/IROS.2004.1389615).
- Borg, I., and Mair, P. (2017). “The choice of initial configurations in multidimensional scaling: local minima, fit, and interpretability,” *Austrian Journal of Statistics* **46**(2), 19–32.
- Bouwman, T., Sobral, A., Javed, S., Jung, S. K., and Zahzah, E.-H. (2017). “Decomposition into low-rank plus additive matrices for background/foreground separation: A review for a comparative evaluation with a large-scale dataset,” *Computer Science Review* **23**, 1–71.
- Bouwman, T., and Zahzah, E. H. (2014). “Robust PCA via principal component pursuit: A review for a comparative evaluation in video surveillance,” *Computer Vision and Image Understanding* **122**, 22 – 34, <http://www.sciencedirect.com/science/article/pii/S1077314213002294>, doi: <http://dx.doi.org/10.1016/j.cviu.2013.11.009>.
- Burgess, S., Kuang, Y., and Åström, K. (2015). “Toa sensor network self-calibration for receiver and transmitter spaces with difference in dimension,” *Signal Processing* **107**, 33–42.

- Candès, E. J., Li, X., Ma, Y., and Wright, J. (2011). “Robust principal component analysis?,” *Journal of the ACM (JACM)* **58**(3), 11.
- Chardon, G., Nowakowski, T., de Rosny, J., and Daudet, L. (2015). “A blind dereverberation method for narrowband source localization,” *Selected Topics in Signal Processing, IEEE Journal of* **9**(5), 815–824, doi: [10.1109/JSTSP.2015.2422673](https://doi.org/10.1109/JSTSP.2015.2422673).
- Chen, H., Bao, Y., and Ser, W. (2015). “Effects of sensor position errors on farfield/nearfield wideband beamformers for microphone arrays,” *Sensors Journal, IEEE* **15**(9), 4812–4825, doi: [10.1109/JSEN.2015.2431720](https://doi.org/10.1109/JSEN.2015.2431720).
- Crocco, M., Del Bue, A., and Murino, V. (2012). “A bilinear approach to the position self-calibration of multiple sensors,” *IEEE Transactions On Signal Processing* **60**(2), 660–673, doi: [10.1109/TSP.2011.2175387](https://doi.org/10.1109/TSP.2011.2175387).
- de Leeuw, J., and Mair, P. (2009). “Multidimensional scaling using majorization: Smacof in R,” *Journal of Statistical Software* **31**(3), 1–30, doi: [10.18637/jss.v031.i03](https://doi.org/10.18637/jss.v031.i03).
- DokmaniÄĖ, I., Parhizkar, R., Ranieri, J., and Vetterli, M. (2015). “Euclidean distance matrices: Essential theory, algorithms, and applications,” *IEEE Signal Processing Magazine* **32**(6), 12–30, doi: [10.1109/MSP.2015.2398954](https://doi.org/10.1109/MSP.2015.2398954).
- Fischler, M. A., and Bolles, R. C. (1981). “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” *Commun. ACM* **24**(6), 381–395, <http://doi.acm.org/10.1145/358669.358692>, doi: [10.1145/358669.358692](https://doi.org/10.1145/358669.358692).
- FitzGerald, D., Liutkus, A., and Badeau, R. (2016). “PROJET - Spatial audio separation using projections,” in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 36–40, doi: [10.1109/ICASSP.2016.7471632](https://doi.org/10.1109/ICASSP.2016.7471632).
- Forero, P. A., and Giannakis, G. B. (2012). “Sparsity-exploiting robust multidimensional scaling,” *IEEE Transactions On Signal Processing* **60**(8), 4118–4134, doi: [10.1109/TSP.2012.2197617](https://doi.org/10.1109/TSP.2012.2197617).
- Gaubitch, N. D., Kleijn, W. B., and Heusdens, R. (2013). “Auto-localization in ad-hoc microphone arrays,” in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, pp. 106–110.
- Gower, J. C., and Dijksterhuis, G. B. (2004). *Procrustes problems*, **3** (Oxford University Press Oxford), pp. 1–248.
- Hafizovic, I., Nilsen, C.-I. C., Kjolerbakken, M., and Jahr, V. (2012). “Design and implementation of a MEMS microphone array system for real-time speech acquisition,” *Applied Acoustics* **73**(2), 132–143, doi: [10.1016/j.apacoust.2011.07.009](https://doi.org/10.1016/j.apacoust.2011.07.009).
- Hennecke, M., Plotz, T., Fink, G. A., Schmalenstroer, J., and Hab-Umbach, R. (2009). “A hierarchical approach to unsupervised shape calibration of microphone array networks,” in *Statistical*

- Signal Processing, 2009. SSP '09. IEEE/SP 15th Workshop on*, pp. 257–260.
- Hennecke, M. H., and Fink, G. A. (2011). “Towards acoustic self-localization of ad hoc smartphone arrays,” in *Hands-free Speech Communication and Microphone Arrays (HSCMA), 2011 Joint Workshop on*, pp. 127–132.
- Himawan, I., Sridharan, S., and McCowan, I. (2008). “Dealing with uncertainty in microphone placement in a microphone array speech recognition system,” in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1565–1568, doi: [10.1109/ICASSP.2008.4517922](https://doi.org/10.1109/ICASSP.2008.4517922).
- Huang, P. S., Chen, S. D., Smaragdis, P., and Hasegawa-Johnson, M. (2012). “Singing-voice separation from monaural recordings using robust principal component analysis,” in *Proc. Speech and Signal Processing (ICASSP) 2012 IEEE Int. Conf. Acoustics*, pp. 57–60, doi: [10.1109/ICASSP.2012.6287816](https://doi.org/10.1109/ICASSP.2012.6287816).
- Kabsch, W. (1976). “A solution for the best rotation to relate two sets of vectors,” *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography* **32**(5), 922–923.
- Khanal, S., Silverman, H. F., and Shakya, R. R. (2013). “A free-source method (FrSM) for calibrating a large-aperture microphone array,” *IEEE Transactions On Audio Speech and Language Processing* **21**(8), 1632–1639, doi: [10.1109/TASL.2013.2256896](https://doi.org/10.1109/TASL.2013.2256896).
- Knapp, C. H., and Carter, G. C. (1976). “The generalized correlation method for estimation of time delay,” *IEEE Transactions on Acoustics, Speech, and Signal Processing* **24**(4), 320–327, doi: [10.1109/TASSP.1976.1162830](https://doi.org/10.1109/TASSP.1976.1162830).
- Koyano, Y., Yatabe, K., Ikeda, Y., and Oikawa, Y. (2016). “Physical-model based efficient data representation for many-channel microphone array,” in *Proc. Speech and Signal Processing (ICASSP) 2016 IEEE Int. Conf. Acoustics*, pp. 370–374, doi: [10.1109/ICASSP.2016.7471699](https://doi.org/10.1109/ICASSP.2016.7471699).
- Kuang, Y., Burgess, S., Torstensson, A., and Åström, K. (2013). “A complete characterization and solution to the microphone position self-calibration problem,” in *Proc. Speech and Signal Processing 2013 IEEE Int. Conf. Acoustics*, pp. 3875–3879, doi: [10.1109/ICASSP.2013.6638384](https://doi.org/10.1109/ICASSP.2013.6638384).
- Lai, Y. M., Balan, R., Claussen, H., and Rosea, J. (2013). “Broadband sensor location selection using convex optimization in very large scale arrays,” in *Applications of Signal Processing to Audio and Acoustics (WASPAA), 2013 IEEE Workshop on*, pp. 1–4, doi: [10.1109/WASPAA.2013.6701889](https://doi.org/10.1109/WASPAA.2013.6701889).
- Le, T.-K., Ono, N., Nowakowski, T., Daudet, L., and De Rosny, J. (2016). “Experimental validation of TOA-based methods for microphones array positions calibration,” in *Proc. Speech and Signal*

- Processing (ICASSP) 2016 IEEE Int. Conf. Acoustics*, pp. 3216–3220, doi: [10.1109/ICASSP.2016.7472271](https://doi.org/10.1109/ICASSP.2016.7472271).
- Leiba, R., Ollivier, F., Marchal, J., Misdariis, N., Marchiano, R. *et al.* (2017). “Large array of microphones for the automatic recognition of acoustic sources in urban environment,” in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, Institute of Noise Control Engineering, Vol. 255, pp. 2662–2670.
- Lewis, J., and Schreier, P. (2013). *Application Note : Low Self Noise: The First Step to High-Performance MEMS Microphone Applications*, InvenSense.
- Ono, N., Kohno, H., Ito, N., and Sagayama, S. (2009). “Blind alignment of asynchronously recorded signals for distributed microphone array,” in *Applications of Signal Processing to Audio and Acoustics, 2009. WASPAA '09. IEEE Workshop on*, pp. 161–164, doi: [10.1109/ASPAA.2009.5346505](https://doi.org/10.1109/ASPAA.2009.5346505).
- Perrodin, F., Nikolic, J., Busset, J., and Siegwart, R. (2012). “Design and calibration of large microphone arrays for robotic applications,” in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pp. 4596–4601, doi: [10.1109/IROS.2012.6385985](https://doi.org/10.1109/IROS.2012.6385985).
- Plinge, A., Jacob, F., Haeb-Umbach, R., and Fink, G. A. (2016). “Acoustic microphone geometry calibration: An overview and experimental evaluation of state-of-the-art algorithms,” *IEEE Signal Processing Magazine* **33**(4), 14–29, doi: [10.1109/MSP.2016.2555198](https://doi.org/10.1109/MSP.2016.2555198).
- Sachar, J. M., Silverman, H. F., and Patterson, W. R. (2001). “Large vs small aperture microphone arrays: performance over a large focal area,” in *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on*, Vol. 5, pp. 3049–3052 vol.5, doi: [10.1109/icassp.2001.940301](https://doi.org/10.1109/icassp.2001.940301).
- Sachar, J. M., Silverman, H. F., and Patterson, W. R. (2005). “Microphone position and gain calibration for a large-aperture microphone array,” *IEEE Transactions On Speech and Audio Processing* **13**(1), 42–52, doi: [10.1109/TSA.2004.834459](https://doi.org/10.1109/TSA.2004.834459).
- She, Y., and Owen, A. B. (2012). “Outlier detection using nonconvex penalized regression,” *Journal of the American Statistical Association* .
- Sun, Y., Babu, P., and Palomar, D. P. (2017). “Majorization-minimization algorithms in signal processing, communications, and machine learning,” *IEEE Transactions on Signal Processing* **65**(3), 794–816, doi: [10.1109/TSP.2016.2601299](https://doi.org/10.1109/TSP.2016.2601299).
- Taghizadeh, M. J., Garner, P. N., and Boursard, H. (2014). “Enhanced diffuse field model for ad hoc microphone array calibration,” *Signal Processing* **101**, 242–255, doi: [10.1016/j.sigpro.2014.02.012](https://doi.org/10.1016/j.sigpro.2014.02.012).

- Taghizadeh, M. J., Parhizkar, R., Garner, P. N., Bourlard, H., and Asaei, A. (2015). “Ad hoc microphone array calibration: Euclidean distance matrix completion algorithm and theoretical guarantees,” *Signal Processing* **107**, 123–140, <http://www.sciencedirect.com/science/article/pii/S0165168414003508>, doi: <http://dx.doi.org/10.1016/j.sigpro.2014.07.016>.
- Tibshirani, R. (1996). “Regression shrinkage and selection via the lasso,” *Journal of the Royal Statistical Society. Series B (Methodological)* 267–288.
- Vanwynsberghe, C., Challande, P., Marchal, J., Marchiano, R., and Ollivier, F. (2016). “A robust and passive method for geometric calibration of large arrays,” *The Journal of the Acoustical Society of America* **139**(3), 1252–1263, <http://scitation.aip.org/content/asa/journal/jasa/139/3/10.1121/1.4944566>, doi: <http://dx.doi.org/10.1121/1.4944566>.
- Vanwynsberghe, C., Marchiano, R., Ollivier, F., Challande, P., Moingeon, H., and Marchal, J. (2015). “Design and implementation of a multi-octave-band audio camera for realtime diagnosis,” *Applied Acoustics* **89**, 281–287, <http://www.sciencedirect.com/science/article/pii/S0003682X14002709>, doi: <http://dx.doi.org/10.1016/j.apacoust.2014.10.009>.
- Velasco, J., Pizarro, D., Macias-Guarasa, J., and Asaei, A. (2016). “TDOA matrices: Algebraic properties and their application to robust denoising with missing data,” *arXiv preprint arXiv:1601.04485* doi: [10.1109/icassp.2015.7178455](https://doi.org/10.1109/icassp.2015.7178455).
- Wang, L., Hon, T.-K., Reiss, J. D., and Cavallaro, A. (2016). “Self-localization of ad-hoc arrays using time difference of arrivals,” *IEEE Transactions on Signal Processing* **64**(4), 1018–1033, doi: [10.1109/TSP.2015.2498130](https://doi.org/10.1109/TSP.2015.2498130).
- Weinstein, E., Steele, K., Agarwal, A., and Glass, J. (2007). “A 1020-node microphone array and acoustic beamformer,” in *International Congress on Sound and Vibration (ICSV)*, doi: [10.1109/isspit.2015.7394321](https://doi.org/10.1109/isspit.2015.7394321).
- Williams, E. G. (1999). *Fourier acoustics: sound radiation and nearfield acoustical holography* (Academic press).
- Zhang, Z., Xu, Y., Yang, J., Li, X., and Zhang, D. (2015). “A survey of sparse representation: algorithms and applications,” *Access, IEEE* **3**, 490–530, doi: [10.1109/access.2015.2430359](https://doi.org/10.1109/access.2015.2430359).
- Zhayida, S., Andersson, F., Kuang, Y., and Åström, K. (2014). “An automatic system for microphone self-localization using ambient sound,” in *2014 22nd European Signal Processing Conference (EUSIPCO)*, pp. 954–958, doi: [10.1109/apsipa.2015.7415477](https://doi.org/10.1109/apsipa.2015.7415477).
- Zhou, T., and Tao, D. (2011). “Godec: Randomized low-rank & sparse matrix decomposition in noisy case,” in *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, edited by L. Getoor and T. Scheffer, ICML ’11, ACM, New York, NY, USA, pp. 33–40, doi:

[10.1109/icsmc.2009.5346879](https://doi.org/10.1109/icsmc.2009.5346879).

Zhou, T., and Tao, D. (2013). “Shifted subspaces tracking on sparse outlier for motion segmentation,” in *IJCAI*, Citeseer.

Zibulevsky, M., and Elad, M. (2010). “L1-l2 optimization in signal and image processing,” *IEEE Signal Processing Magazine* **27**(3), 76–88.