PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Direction of arrival estimation of acoustic vehicular sources

Gabriela D. ROCHA(1), Felipe R. PETRAGLIA(2), Julio Cesar B. TORRES(3), Mariane R. PETRAGLIA(4)

(1)Electrical Engineering Department, Federal University of Rio de Janeiro, Brazil, gabidantas@poli.ufrj.br

(2)Electrical Engineering Department, Federal University of Rio de Janeiro, Brazil, fpetraglia@pads.ufrj.br

(3)Electrical Engineering Department, Federal University of Rio de Janeiro, Brazil, julio@poli.ufrj.br

(4)Electrical Engineering Department, Federal University of Rio de Janeiro, Brazil, mariane@pads.ufrj.br

Abstract

Road traffic is the main component of urban noise, considered by the World Health Organization as the third type of pollution that most affects the population of large urban centers. The auralization technique stands out as an intuitive tool for assessing environmental noise, but does not yet explore the sound characteristics to provide an accurate hearing experience. In this context, it is desirable for a database containing vehicle sounds to be provided to the auralization systems, but these signals must be previously located and isolated from undesirable sources of noise. This paper presents a performance evaluation and comparison of five direction of arrival (DoA) algorithms to locate vehicular noise sources. Four of the algorithms are based on estimates of the time difference of arrival (TDoA), whereas one employs the minimum variance distortionless response (MVDR) beamformer. The estimated source locations obtained by the five algorithms for data acquired by eleven circularly arranged microphones are compared with the correct positions of the noise sources.

Keywords: Acoustic virtual reality, vehicular noise, microphone array, time difference of arrival, beamforming.

1 INTRODUCTION

Noise pollution causes a strong impact in the health and well-being of city inhabitants and is stated by the World Health Organisation (WHO) as a public health issue (1). Noise maps and indicators are often employed for assessment purposes, as they provide an idea of the sound in the mapped location by representing sound features in a numerical scale. Such simplification helps in establishing a common ground for noise assessment but lacks in providing a realistic and intuitive sound representation. This problem is aggravated when evaluation is to be conducted by non-experts such as decision makers, politicians and the general public.

As an alternative, auralization technique produces an audio signal which represents, as accurately as possible, the hearing experience that a listener would have in the real scene. Auralization systems (2) are able to generate an audio signal from numerical data which are responsible for describing the whole scene, including source characteristics, sound propagation, sound reproduction and all the elements involved in these steps. A proper description influences on whether the generated sound can be distinguished from the real one or not. Therefore, an auralization system implementation requires a signals and models database to be used as an input.

Road traffic is one of the leading causes of complaint in noise monitoring studies in urban areas (3–6) and it is generated mainly by vehicles and their components, such as engine, tires and exhaust (7). Each individual source has its own spectral and directional patterns, which must be modelled in order to be enclosed to the auralization system database.

The ongoing project "Auralization of Urban Areas" is held in cooperation with the Institute of Technical Acoustics from RWTH Aachen University, whose objective is to model vehicular sources inside a simulation tool. At the first stage, the sources are modelled by acquiring vehicle noise emissions, among other noise sources, with a microphone array. Then the relevant data is selected by applying signal processing techniques to filter the recorded signals, both in time and space. However, spatial filtering requires the knowledge of position, or at least direction, of the desired source. This work focuses in determining the most appropriate method to identify

direction of arrival for vehicular noise sources.

2 METHODS FOR DIRECTION OF ARRIVAL ESTIMATION

The direction of arrival estimation methods used in this work are described in this section. The algorithms use two microphone signals in order to estimate the time difference of arrival (TDoA) between signals and the source direction of arrival (DoA). The two-microphone setup under plane-wave propagation assumption is sketched in Figure 1, which indicates the geometric relation between TDoA and DoA.

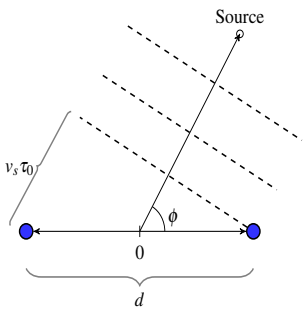


Figure 1. Two-microphone setup for delay τ_0 and direction of arrival ϕ estimation.

The recorded raw data is fed into the system depicted in Figure 2. Firstly, in pre-processing block, the input signal length, sample rate and frequency spectrum is adjusted towards enhancing DoA algorithms performance. This is related to the source bandwidth and aims to avoid undesired noise. Each method provide different approaches for DoA estimation. Then, a curve fitting algorithm helps to separate multiple sources. Finally, DoA estimation method is implemented in the last block and the user must decide which among the methods is to be used in the current estimation.

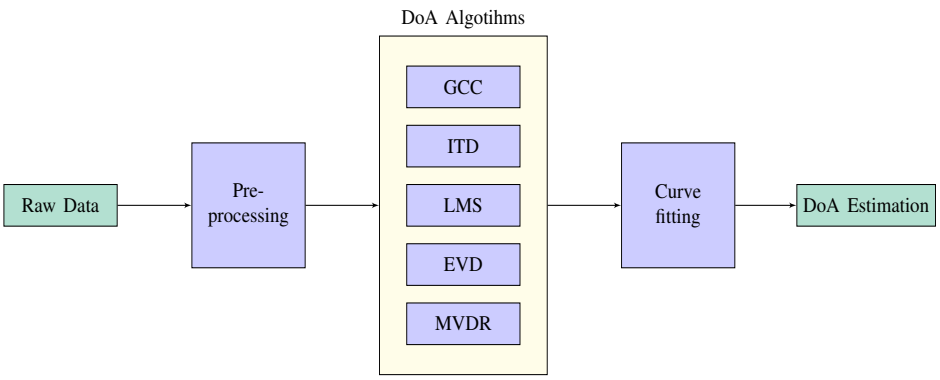


Figure 2. System schematic diagram.

2.1 Generalized Cross-Correlation

The Generalized Cross-Correlation (GCC) (8) method uses correlation function properties and the assumption that signals arriving in both microphones are identical, except for a time delay, to find a TDoA estimate. Cross-correlation function provides an idea of similarity between signals as a function of the lag τ applied to one of them. For two identical signals separated by a time delay τ_0, the correlation reaches its peak when the lag τ equals the real delay τ_0. Therefore, a peak detector can be used for delay estimation. Generalized

cross-correlation function differs from regular cross-correlation due to a factor introduced by previous signal filtering. This pre-filter helps at sharpening function peaks and consequently favouring their detection and TDoA estimation.

2.2 Interaural Time Differences

The Interaural Time Differences (ITD) algorithm (9) is inspired by the human auditory system, which is able to identify time differences between the sound at both ears and use it to locate sound sources. The algorithm generates a set of all possible delays, limited by the distance between microphones and by the chosen angular discretization for DoA. The delays are applied, in the frequency domain, to the original signals and a coincidence detector checks which phase difference compensates the real delay τ_0.

2.3 Least Mean Square

The Least Mean Square (LMS) algorithm uses adaptive finite impulse response (FIR) filters for TDoA estimation (10). One microphone signal is filtered by an FIR system, which has its coefficients adapted in order to minimize the least mean square error between the filter output signal and the second microphone signal. After adaption, the highest valued coefficients index indicate the delay between signals.

2.4 Adaptive Eingenvalue Decomposition

Alternatively, adaptive filters might be used to estimate the impulse response between the source and each microphone, as in the Adaptive Eingenvalue Decomposition (AEVD) method (11). The eigenvalue decomposition is performed in the covariance matrix containing space and temporal correlation between the two microphone signals.

2.5 Minimum Variance Distortionless Response

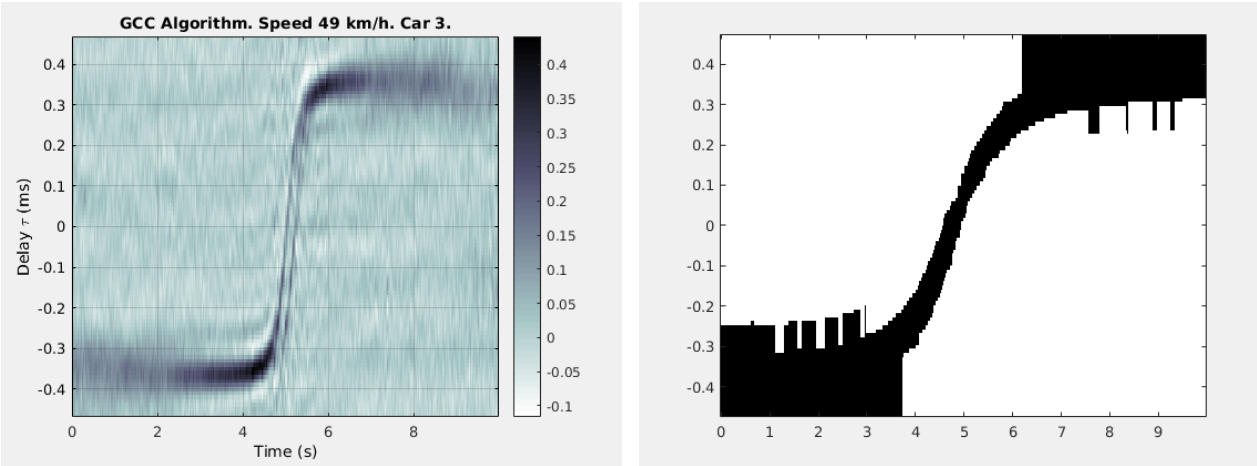
The Minimum Variance Distortionless Response (MVDR) method is an adaptive beamforming approach which seeks to minimize the variance of the recorded signal subject to the restriction that the signal is not distorted. Given that the noise and the desired signal are uncorrelated, the variance of the recorded signal is equal to the sum of the variances of the desired signal and the noise. Therefore, the MVDR algorithm aims to minimize this sum, hence mitigating the noise effect (12).

2.6 DoA Esimation Process

In general, the output of the DoA block is not yet a direction of arrival, but a two-dimensional function of time, t, and of inter-microphone delay, τ, as illustrated in Figure 3a. The function peaks indicate the TDoA associated with the direction of arrival. The originally proposed DoA estimation methods would simply find the function maximum value for each time instant t and define the related direction as the DoA estimation. Such a strategy is sufficient in single-source scenarios, where the desired sound source is isolated or prevails over others, which is not the case with vehicular noise sources.

Traffic noise emissions are mainly accredited to the vehicle engine, tires and exhaust, and such multiple-source contribution is visible in the output function of the DoA estimation block, as shown in Figure 3a, where higher values are indicated in darker colours. Two parallel peak regions are visible in the figure, separated by about 0.19 s. At a 49 km/h speed this is equivalent to a distance around 2.5 m, which is similar to the wheelbase of vehicles. Therefore, combined with literature reports that tire noise predominates over other emitted noises (7), this coincidence of source separation leads one to assume that the two visible peaks come from the front and rear wheels of the car.

The curve fitting block is added to the system as an effort to distinguish between multiple emissions. The two-dimensional function generated by the chosen DoA algorithm is first treated as an image, in which morphological and thresholding operations are applied, resulting in the binary image shown in Figure 3b. From



(a) DoA algorithm (GCC) block output.

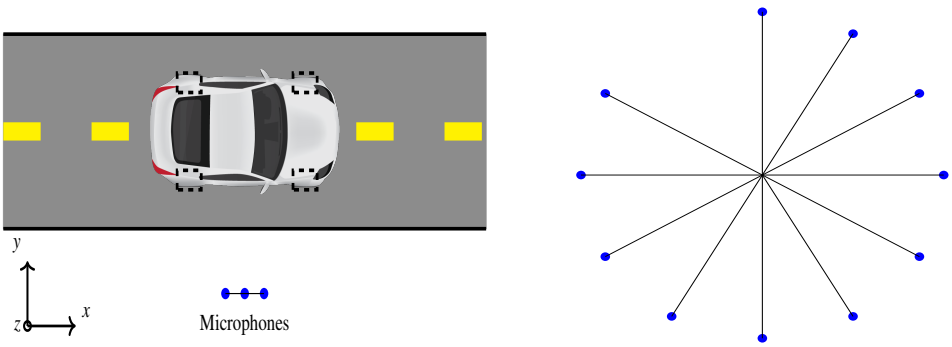
(b) Binary image used to fit TDoA curve.

Figure 3. Illustration of the DoA estimation procedure.

such binary image, nonzero data are separated into two vectors, representing the two visually distinguishable sources. Finally, a curve is fitted to the data in each of the two vectors, using a model derived for the known TDoA behaviour. Thus, the curve fitting block provides TDoA estimates for two sources in parallel motion.

3 EXPERIMENT

The scenario of the experiment conducted to acquire vehicle noise emissions is illustrated in Figure 4a. It consists in 11 microphones arranged in a 0.25-m-diameter circle that are responsible for recording sound pressure information. The array configuration is depicted in Figure 4b. Only eleven microphones were used, rather than twelve equally spaced, due to a limitation in the available signal acquisition module, which only supported eleven inputs.



(a) Array positioning towards vehicle path.

(b) Microphone position in 0.25-m-diameter circular array.

Figure 4. Schematic depiction of the experiment.

Four different passenger cars were used in the experiment trials, as detailed in Table 1, as well as three different drivers. Each trial consisted in one car passing by the microphone array at constant speed of around 30, 50, 60 or 70 km/h or accelerating. The exact speed was measured with a GPS equipped device placed inside the

vehicle. The accelerating tests aimed to gather information from engine noise emissions, since tire noise predominance was observed in previous constant speed tests. The experiment took place at the Brazilian National Institute of Metrology (INMETRO), in a quiet location, resulting in a negligible background noise level.

Table 1. Vehicles used in pass-by tests

Car ID	Car Model	Transmission
1	Volkswagen Gol 1.0	Manual
2	Jeep Renegade 1.8	Automatic
3	Mitsubishi ASX 2.0	Automatic
4	Hyundai Creta 1.6	Automatic

4 RESULTS

Each DoA algorithm provides a different function as output, which is used for curve fitting and displayed as background image. GCC block output is the generalized cross-correlation function between microphone signals, while ITD provides a histogram map summed over frequency and the adaptive methods, LMS and EVD, return the filter coefficients map. MVDR block output is the cross power spectral density function between the microphone signals.

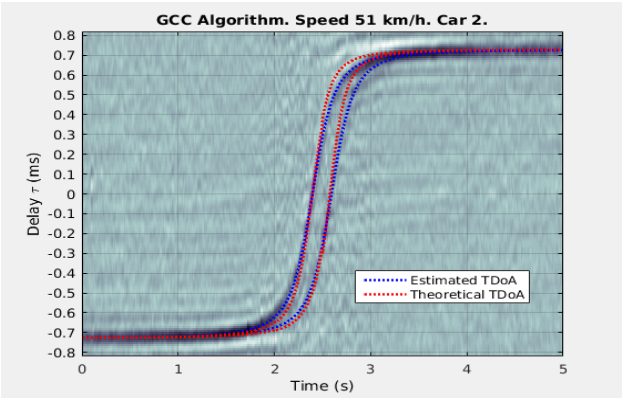
Typical outputs of the TDoA estimation algorithms are shown in grayscale on the bottom layer of each image of Figure 5. The theoretical and estimated TDoA curves are superimposed on these images, with the estimated curves obtained by the curve fitting algorithm applied to each grayscale image. As the example in Figure 5 suggests, GCC and MVDR were, in general, the most successful methods in providing TDoA curves which approached the theoretical model and the visible function peaks.

From the estimated TDoA curves, it is possible to obtain the distance between the sources for the instant they are in front of the microphone array, which is indicative of the distance of the front-rear axle. This value can be compared to the actual wheelbase as a measure of the performance of the algorithm. The estimated wheelbase values extracted from the TDoA curves are given in Table 2 for each vehicle, together with the corresponding actual values. The ITD method was not able to distinguish the two sources and, therefore, the estimated direction of arrival was almost the same for both sources, resulting in the estimated separation practically equal to zero observed in the table. A similar behaviour can be observed in the results of the LMS algorithm, where the estimated values were significantly smaller than the actual ones for the four vehicles. The EVD algorithm, in addition to presenting unstable and sensitive data behaviour, provided larger distance estimates between noise sources than expected. The estimates obtained with GCC and MVDR were close to the true wheelbase values and did not show significant changes in the accuracy for different vehicles.

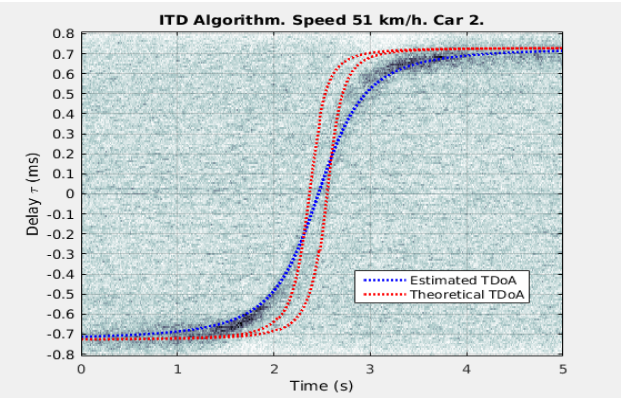
Table 2. Wheelbase estimated by the distance between fitted curves and real values.

	Car ID			
	1	2	3	4
GCC	2.4 ± 0.84	2.11 ± 1.32	2.79 ± 1.57	2.92 ± 0.91
ITD	0.08 ± 0.1	0.06 ± 0.09	0.07 ± 0.08	0.05 ± 0.07
LMS	0.6 ± 0.18	0.56 ± 0.56	1.39 ± 1	1.02 ± 0.89
EVD	4.41 ± 2.54	9.89 ± 21.69	5.19 ± 3.57	8.02 ± 9.13
MVDR	2.08 ± 0.43	2.19 ± 1.09	2.62 ± 0.61	2.17 ± 0.45
Real	2.47	2.57	2.67	2.59

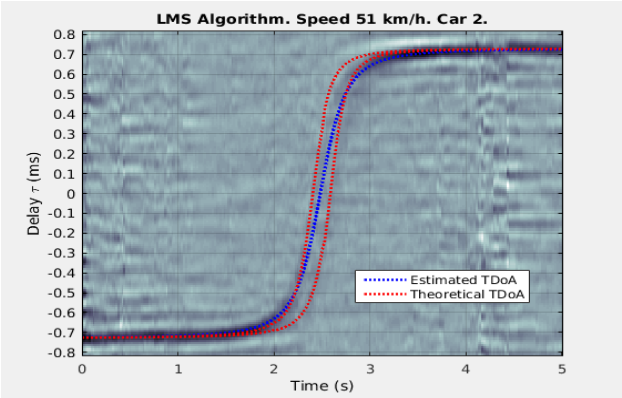
Table 3 presents the mean absolute error between measured and estimated speeds, chosen as an additional



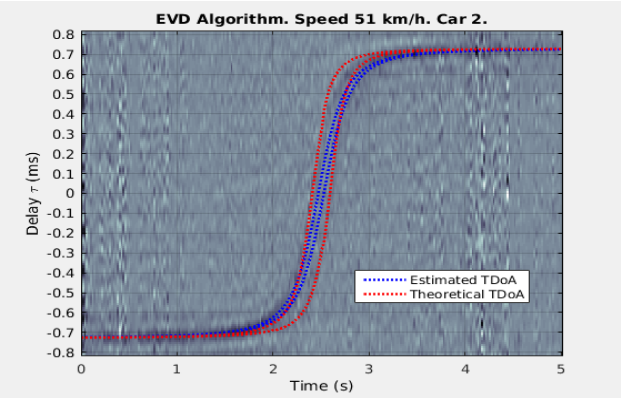
(a) GCC



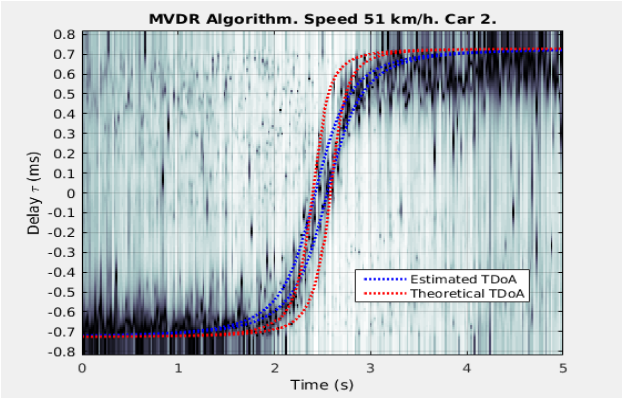
(b) ITD



(c) LMS



(d) AEVD



(e) MVDR

Figure 5. TDoA estimated (blue) and theoretical model (red) curves for each method, for Car 2 at 51 km/h. Background grayscale images are the outputs of the TDoA estimation algorithms.

performance comparison criterion. The estimated values were obtained from the curve fitting coefficients. The GCC method resulted in the smallest errors for all speeds, whereas the ITD provided the largest errors for the three lower speed ranges and the EVD for the higher speed range. All of the methods presented a better

performance for lower speeds and achieved the lowest speed estimation errors in the 30 - 40 km/h range.

Table 3. Mean absolute error between measured and estimated speeds and corresponding standard deviations.

	Speed (km/h)			
	30 - 40	40 - 50	50 - 60	60 - 70
GCC	2.17 ± 1.51	3.06 ± 1.69	3.86 ± 1.47	3.74 ± 1.62
ITD	12.62 ± 2.74	16.49 ± 1.34	23.06 ± 8.61	18.88 ± 6.51
LMS	3.05 ± 1.42	4.59 ± 2.32	6.09 ± 0.8	7.67 ± 4.74
EVD	9 ± 6.71	17.26 ± 8.37	17.98 ± 9.43	27.54 ± 20.87
MVDR	9.26 ± 1.9	15.11 ± 1.55	12.79 ± 1.8	14.88 ± 4.1

Building a rich database, containing vehicular sources of multiples types moving at different speeds, requires a solid performance from our system for all possible scenarios. Therefore, from Tables 2 and 3, GCC and MVDR methods proved to be suitable choices for DoA estimation of vehicular noise sources.

5 CONCLUSIONS

This work sought a method capable of locating typical noise sources of an urban environment. Five direction of arrival estimation methods were tested for this purpose. The methods were originally proposed for single-source speech signal tracking and were modified to fit the urban noise application by identifying multiple sources. Both the Interaural Time Differences and the Least Mean Square methods were unable to distinguish different noise sources in a vehicle. The Adaptive Eigenvalue Decomposition algorithm results demonstrated a great variation in performance for different test conditions. The Generalized Cross-Correlation and the Minimum Variance Distortionless Response algorithms presented the best performances among the five tested methods for the chosen evaluation criteria. All methods showed better performances at lower speeds.

ACKNOWLEDGEMENTS

The authors thank the financial support of CAPES/DAAD PROBRAL Program for developing this work through project number 88881.198848/2018-01.

REFERENCES

1. Organization WH, Others. Burden of disease from environmental noise: Quantification of healthy life years lost in Europe. In: Burden of disease from environmental noise: quantification of healthy life years lost in Europe; 2011. p. 126.

2. Vorländer M. Auralization: fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality. Springer Science & Business Media; 2007.

3. Zannin PH, Calixto A, Diniz FB, Ferreira JA. A survey of urban noise annoyance in a large Brazilian city: the importance of a subjective analysis in conjunction with an objective analysis. Environmental Impact Assessment Review. 2003;23(2):245–255.

4. Hammer MS, Swinburn TK, Neitzel RL. Environmental noise pollution in the United States: developing an effective public health response. Environmental health perspectives. 2013;122(2):115–119.

5. Seong JC, Park TH, Ko JH, Chang SI, Kim M, Holt JB, et al. Modeling of road traffic noise and estimated human exposure in Fulton County, Georgia, USA. Environment international. 2011;37(8):1336–1341.

6. Nugent C, Blanes N, Fons J, Sáinz de la Maza M, Ramos M, Domingues F, et al. Noise in europe 2014. European Enviroment Agency. 2014;10:2014.

7. Ouis D. Annoyance from road traffic noise: a review. Journal of environmental psychology. 2001;21(1):101–120.

8. Knapp C, Carter G. The generalized correlation method for estimation of time delay. IEEE Transactions on Acoustics, Speech, and Signal Processing. 1976;24(4):320–327.

9. Liu C, Wheeler BC, O'Brien Jr WD, Bilger RC, Lansing CR, Feng AS. Localization of multiple sound sources with two microphones. The Journal of the Acoustical Society of America. 2000;108(4):1888–1905.

10. Reed F, Feintuch P, Bershad N. Time delay estimation using the LMS adaptive filter–static behavior. IEEE Transactions on Acoustics, Speech, and Signal Processing. 1981;29(3):561–571.

11. Benesty J. Adaptive eigenvalue decomposition algorithm for passive acoustic source localization. the Journal of the Acoustical Society of America. 2000;107(1):384–391.

12. Vorobyov S. Principles of minimum variance robust adaptive beamforming design. Signal Processing. 2013;93(1):3264–3277.