BeBeC-2018-D17

IDENTIFICATION OF ACOUSTIC MOVING SOURCES USING A TIME-DOMAIN METHOD

Rémi Cousson[1,2], Quentin Leclère[2], Marie-Agnès Pallas[1] and Michel Bérengier[3]
[1]Univ Lyon, IFSTTAR, CEREMA, UMRAE, F-69675, Lyon, France
[2]Univ Lyon, INSA-Lyon, Laboratoire Vibrations Acoustique, F-69621 Villeurbanne, France
[3]IFSTTAR, CEREMA, UMRAE, F-44344 Bouguenais, France

Abstract

The issue of the identification of acoustic sources has been widely treated in static or quasi-static source contexts, more recently for moving sources, in a wide range of applications. The beamforming method is a reference in that situation, but it has well-known limitations, namely a poor resolution at low frequencies which makes it difficult to discriminate close sources or with too different levels, and to quantify acoustic power levels. Deconvolution methods are often used to overcome those limitations. Performed in the frequency domain, they are effective in static contexts or with quite restrictive conditions if considering moving sources. In the present investigation, an original method is presented to address those problems. It is a matching pursuit-type method, performed exclusively in the time-domain and combining the sweeping focus point beamforming approach to an iterative cleaning algorithm, with the noticeable advantage to be adapted to non-stationary signals in pass-by situations. Experimental validation setups in both linear and rotational motions are presented to prove the potential and versatility of this method.

1 INTRODUCTION

The present study's goal is to find an effective method to localize and quantify acoustic sources in both linear movement, with a road vehicle pass-by application, and rotating movement. Beamforming is well suited to the localization of moving sources by steering the focus point of the microphone array with the moving source-point, by means of time-varying delays[2]. However, it has a poor resolution in low frequency and is not very accurate to quantify sources.

To overcome these issues, deconvolution methods have been developed to improve the localization and the estimation of the power level radiated by static sources[3, 8]. Several of them have been adapted to the moving source problem. Fleury and Bulté[5] proposed a solution which was tested in simulation and real experiment in the case of an aicraft flying over. A variant in the context of underwater acoustics was implemented in the case of a ship, studied by Oudompheng[12] and Lamotte[9], involving far lower vehicle speeds. The performance of this method has been numerically assessed in the case of road vehicle pass-by conditions and with an academic experimental set-up[4]. Although it showed a clear improvement from conventional beamforming, some limitations were also pointed out in terms of separation power and in the frequency range. Road vehicle pass-by conditions differ from the cases studied in [5, 9, 12] in the vehicle speed and distance from the microphone array, ultimately resulting in a great difference in the pass-by duration.

Another domain concerned with the localization and quantification of moving sources is the characterization of rotating sources such as fans. Time domain beamforming using a rotating focusing point is the topic of several works [6, 11, 13], major drawbacks being identified as the quite heavy computational costs and the difficulty to implement standard deconvolution techniques. Other approaches are based on the construction of a virtual rotating array based on the fixed array measurements [1, 7]. Using signals reconstructed on this virtual rotating array, standard frequency domain methods can be implemented as well as deconvolution. However, these techniques require the use of specific arrays, such as ring arrays, and can be limited by the spatial sampling that is required to correctly interpolate the sound field.

The present paper proposes an original method, exclusively performed in the time domain, that takes advantage of the fact that time domain beamforming actually reconstructs the signal emitted from the focus point. The non-stationary nature of a pass-by measurement makes it interesting to operate in time domain, especially to take into account the Doppler effect. In a first section, the method is described and discussed. Then, an academic experimental set-up designed to test the method and compare it with moving source conventional beamforming followed by a deconvolution algorithm (noted CBF from this point) is presented. A third section is dedicated to an experimental application of the method to rotating sources.

2 PROBLEM FORMULATION

2.1 Identification of static sources

Conventional beamforming

The reference method to identify static sources with a phased array of microphones is conventional beamforming. It can be performed either in time-domain or frequency-domain. In either case, the location and the power of sources in a scan area are estimated thanks to the signals received by an array of sensors. In the frequency-domain, the power estimate from the i^{th} point in the sampled scan area is, at each frequency f:

$$\boldsymbol{B_i}(f) = \boldsymbol{w_i^*}(f)\boldsymbol{C}(f)\boldsymbol{w_i}(f) \tag{1}$$

where $\boldsymbol{w_i}$ is a steering vector depending on the source and propagation models, $*$ denotes the complex conjugate transpose and $\boldsymbol{C}$ is the cross-spectral matrix of microphone signals.

Conventional beamforming works as a spatial filter but has a poor resolution in low frequency and is not very accurate to quantify the power of extended sources or in a multi-source environment. Deconvolution methods have been developed to overcome those limitations and are briefly presented in the next section.

Deconvolution

Most deconvolution methods are applied to a beamforming output that has been performed in the frequency domain. For time-invariant conditions and decorrelated sources, the deconvolution consists in inversing the following system:

$$\boldsymbol{B}(f) = \boldsymbol{H}(f)\boldsymbol{S}(f) \quad with \quad \boldsymbol{S} \geq 0 \tag{2}$$

where $\boldsymbol{B}$ is the vector of quadratic beamforming outputs for the set of focus points on the grid representing the sampled scan area, $\boldsymbol{H}$ is an energetic transfer function matrix and $\boldsymbol{S}$ is a vector representing the unknown quadratic source strengths. $\boldsymbol{H}$ depends on the model of sources and propagation chosen according to the physical context. When considering point sources, the columns of $\boldsymbol{H}$ are called *Point Spread Functions* (PSF).

Various algorithms are available in the literature and can be used to find a solution to Eq. (2), for instance DAMAS[3], CLEAN[8] or NNLS (non-negative least square)[10]. While deconvolution can be very effective for static sources, its application to moving sources is more challenging. This problem must be addressed in a different way, as it is presented in section 2.2.

2.2 Identification of moving sources

Sources with a linear motion

The signal received by a microphone m of the array from a source s linearly moving at the constant speed V is:

$$p_m\left(t + \frac{R_{sm}(t)}{c}\right) = \frac{Q(t)}{R_{sm}(t)(1 - M\cos\theta_{sm}(t))^2} \tag{3}$$

where $Q(t) = \frac{q'(t)}{4\pi}$, $q'(t)$ is the derivative of the source mass flow, M is the Mach number ($M = \frac{V}{c}$), R_{sm} is the propagation distance between the position of source s and sensor m, θ_{sm} is the angle between the source movement direction and the source-sensor direction and c is the speed of sound.

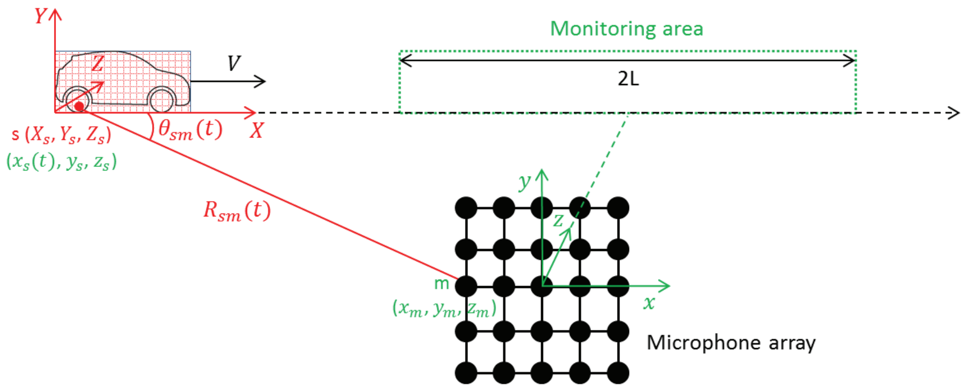


Figure 1: Pass-by configuration in the fixed (green) and moving (red) coordinate systems.

Beamforming focuses the array on a point F at position $[X_F, Y_F, Z_F]$ on the source-grid constituted of N_F points, which samples the area scanned on the vehicle. With the hypothesis that the source is moving along the x axis, the beamforming output at this point is:

$$b(t,F,\{p_m\}) = \frac{W(t+\frac{X_F}{V})}{N_m} \sum_{m=1}^{N_m} R_{Fm}(t)(1-M\cos\theta_{Fm}(t))^2 p_m\left(t+\frac{R_{Fm}(t)}{c}\right) \tag{4}$$

$b(t,F,\{p_m\})$ can be considered as an estimate of the time signal emitted from the moving focus point F. The focus point moves along with the vehicle, so knowledge of the vehicle kinematics is essential. To avoid an unacceptable broadening of the beamforming response main lobe when focusing at high angles from the normal to the array, a monitoring area in front of the array is defined (Fig. 1). $W(t)$ is a centred apodization window which is null for all $|t| > L/V$ where L is half the length of the monitoring area (Fig 1).

The expression (4) corresponds to inversing Eq. (3) for each microphone and averaging the result over all the sensors.

Sources with a rotating motion

The time domain focusing method described in the previous section for a source traveling with a linear motion is directly applicable to a source traveling with a rotating motion. Equation (3) is still valid and only windowing parameters have to be adapted in Eq. (4). Of course, geometrical parameters $R_{Fm}(t)$ and $\theta_{Fm}(t)$ have to take into account the rotating motion, and the Mach number depends on the distance between the source point and the rotation axis. This formulation of the beamforming applied to rotating source has been used in several papers (see for instance [6, 11, 13]), the originality of this work is to implement it in the frame of the CLEANT methodology, which is detailed in the next section.

CLEANT procedure

CLEANT is an iterative deconvolution approach which operates in the time domain. Using conventional moving source beamforming, it gradually subtracts the source time contributions from the microphone signals. The “dirty map” and “clean map” specific to the CLEAN method[8] are of a different nature in this CLEANT procedure: they are arrays of source time signals associated with the main source locations. The term “map”, although inappropriate, is used in the following for these arrays so as to highlight the parallels with CLEAN.

The beamforming mentioned in the algorithm corresponds to the moving source beamforming as described in 2.2. The output is composed of the signals $b(t,F,\{p_m\})$ reconstructed for each point F of the source-grid during the time interval T_F, completed with zero-padding over the recording interval $\mathscr{T}$.

The algorithm performs the following steps:

1. Initialization of the dirty “map”D and the clean “map” Γ at iteration $i=0$:

$$\begin{aligned} &D^{(0)}(t,F) = b(t,F,\{p_m\}) \quad \forall F \in [\![1;N_F]\!] \\ &\Gamma^{(0)}(t,F) = 0 \\ &p_m^{res(0)}(t) = p_m(t) \end{aligned}$$

2. At iteration i ($i \geq 1$), detection of the signal $D^{(i-1)}(t,\hat{F})$ of the dirty "map", having the maximum energy:

$$\hat{F} = \underset{F}{\operatorname{argmax}} \left(\int_{\mathscr{T}} \left| D^{(i-1)}(t,F) \right|^2 dt \right)$$

3. The dominant source contribution is removed from the microphone signals:

$$p_m^{res(i)}\left(t + \frac{R_{\hat{F}m}}{c}\right) = p_m^{res(i-1)}\left(t + \frac{R_{\hat{F}m}}{c}\right) - \gamma \frac{D^{(i-1)}\left(t,\hat{F}\right)}{R_{\hat{F}m}(t)\left(1 - M\cos\theta_{\hat{F}m}(t)\right)^2} \quad \forall m \in [\![1;N_m]\!]$$

where $\gamma \in [0,1]$ is the CLEANT loop gain and p_m^{res} is the signal on microphone m cleaned from the dominant source time signal contribution.

4. Clean "map" is updated:

$$\Gamma^{(i)}(t,\hat{F}) = \Gamma^{(i-1)}(t,\hat{F}) + \gamma D^{(i-1)}(t,\hat{F})$$

5. Dirty "map" is updated:

$$D^{(i)}(t,F) = b(t,F,\{p_m^{res(i)}\}) \quad \forall F \in [\![1;N_F]\!]$$

6. A new iteration starts back at step 2 unless one of the stopping criteria is reached.

The stopping criteria are defined as follows:

- At each iteration, the total amount of energy left in the dirty map should decrease. However, when all the significant sources have been removed, only noise is left and this energy can increase from this point: the algorithm can stop when the energy at an iteration is higher than at the previous one.
- A maximum number of iteration can be set by the user: depending on the gain factor chosen and the possible knowledge of the number of sources expected, the user can shorten the duration of the method without waiting for the first criterion to be reached.

The CLEANT loop gain $\gamma \in [0,1]$ is chosen to remove only a portion of the dominant source estimate found at each iteration. Indeed, the estimate of the source at this location is likely to include contributions from the beamforming sidelobes of nearby sources. For an optimum robustness of the algorithm, one should choose a quite small value for γ. However, the smaller the loop factor, the longer the algorithm, and values around $\gamma = 1/2$ are usually good enough in most cases. Even $\gamma = 1$ gives good results in the simplest cases (e.g. simulation of a single source).

In the end, all the non null reconstructed source time signals are stored, as well as the associated positions, and become available for any analysis, such as a power spectral density estimate for instance. Then, using the clean "map" $\Gamma(t,F)$, we can create a deconvoluted display of the source locations and of their levels on the source-grid, at any frequency.

3 EXPERIMENTAL APPLICATION OF THE METHOD: CASE OF A PENDULUM

This section presents an academic experimental set-up used to test the method and compare it to CBF in real conditions. In order to carry out a scaled experiment of a vehicle in pass-by condition, involving sources in linear motion, at an approximately constant speed, a pendulum with a long rod carrying a noise source has been used. The circular movement can be locally approximated by a linear movement with the condition that the monitoring area is small with respect to the rod length. The parameters of the experiment are described below.

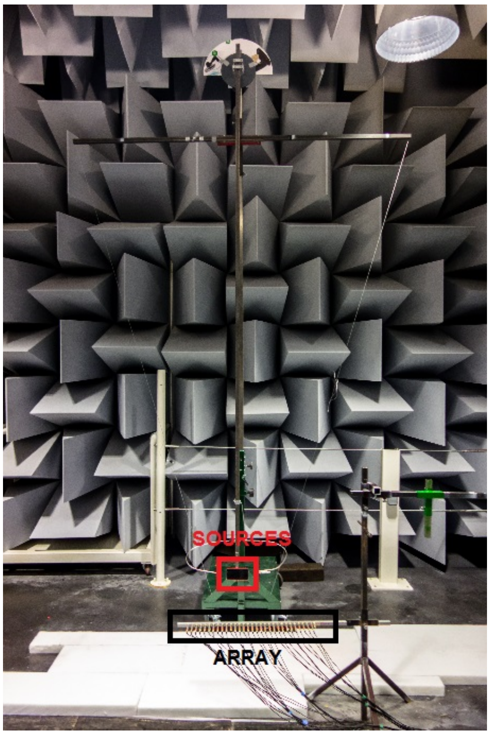


Figure 2: Global view of the experimental set-up.

3.1 Experimental parameters and configurations

The pendulum is made with a three meter-long rigid rod that revolves around an axis supported by a weighted structure (Fig. 2). An optical system is used to get information on kinematics (position, speed). The microphone array has 30 microphones, regularly spaced by 2 cm. Measurement is in direct proximity of the pendulum position at rest, which is used as a reference. Thus, the horizontal array is centered on this reference position and located at the height of the rod end, which supports the sources. A microphone, providing a reference signal for further research, is also fixed to the rod end and placed in direct proximity of one of the sources.

The sound device, located at 0.5 m from the array, is a smartphone equipped with two loudspeakers 128 mm apart, which has several advantages for this experiment. First, the signals are amplified inside the phone, so there is no need for a cable which would be impractical. Second, the two loudspeakers are independent and can play two different signals at the same time: two tone noises at different frequencies, or two uncorrelated white noises for instance. The last advantage of the device is its small dimensions which make it easy to fix at the rod end, and its lightness with no risk to modify the movement.

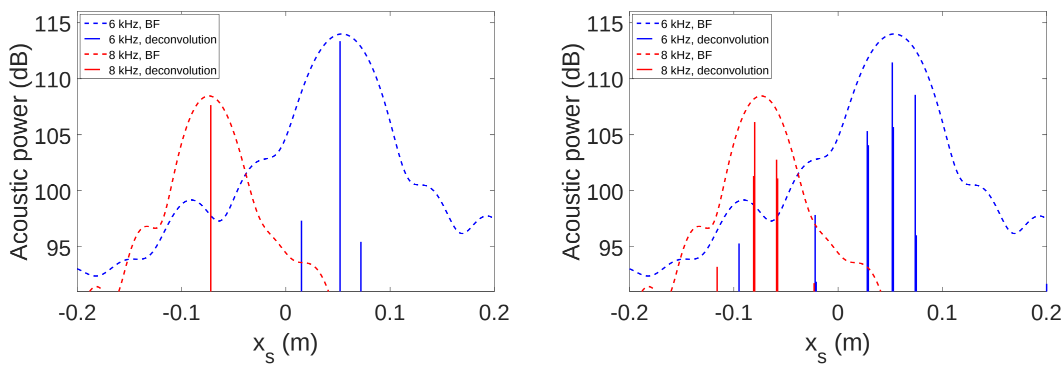
Finally, acoustic panels have been placed on the floor between the pendulum and the array to remove ground reflections.

The test procedure consisted in dropping the pendulum without initial speed from a variable height. Two distinct heights were chosen to obtain movements with various speeds. Two types of source signals were used : i) a tone at 6 kHz on one speaker and another at 8 kHz on the other one; ii) two uncorrelated white noises, one per speaker. One measurement was made with the pendulum at rest to get a reference absolute position of the sources and an accurate estimate of the actual sound source levels.

In all of the measurements, both CLEANT and CBF have been implemented on a single snapshot that lasted 200 ms. For CLEANT, a gain factor of $\gamma = 0.7$ has been chosen.

3.2 Experimental results and comparison with beamforming method

Only the results corresponding to the highest source speed tested are presented here, informing on the methods performance and flagrant differences in the most difficult conditions investigated. The maximal speed of the sources in this configuration is 3.5 m/s (speed of the pendulum when reaching the lowest position) and corresponds to approximatively 30 km/h at road pass-by vehicle scale. We recall that the CBF applied here is adapted to moving sources, with a sweeping focus point steered on the sources (time domain) and that the PSF for the deconvolution (frequency domain) is the same as in the static case.



(a) CLEANT (b) CBF

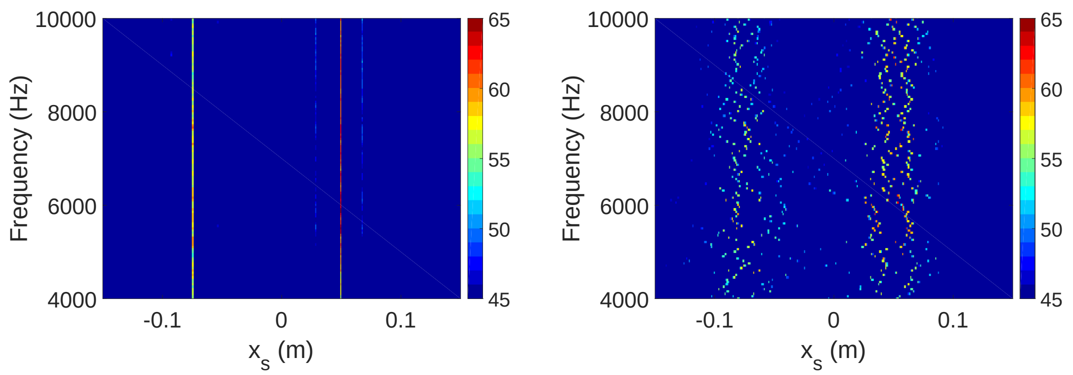
Figure 3: Comparison of deconvolution output for tone noise moving sources.

A first comparison is made for tone sound sources, with a 6 kHz (resp. 8 kHz) tone on the right (resp. left) speaker. Beamforming output without deconvolution points out both sources, with slightly disymetrical and flattened sidelobe patterns (Fig. 3). The deconvolution involved in CBF gives several significant source contributions instead of one expected, within the beamforming main lobe extent (Fig. 3b). All of them must be taken into account to get the overall level found in the static case (not presented here), but the position of the real source is then ambiguous. On the other hand, CLEANT yields two main sources, with spurious sidelobes for the right source but at noticeably lower levels. This removes the ambiguity of the position compared to CBF, but the level is slightly underestimated in this moving source context. The source characteristics estimated are given in Table 1 (weighted averaged source position and overall level are given in the case of important secondary source contribution).

Table 1: Source position and level estimates with moving tone sources.

Method	**Source 1 (8 kHz)**		**Source 2 (6 kHz)**		**Source gap**
	Position (mm)	Level (dB)	Position (mm)	Level (dB)	(mm)
BF	-74	109	+53	114	127
CBF	-73	109	+53	115	126
CLEANT	-72	108	+52	113	124

The beamforming map for moving white noise sources is blurrier than in the case of static sources. This results in very dispersed source contributions in the deconvolution map obtained with CBF (Fig. 4b). The moving source context degrades the result so much that it is not possible to determine a clear location of any of the two sources. CLEANT gives a much clearer map with two well localized sources (-74 mm and +50 mm) and almost no dispersion, the gap between sources being 124 mm (1 mm narrower than in the static context). The total sound level over the frequency range amounts to 82 dB for the right source with CLEANT and to 78 dB for the left source. The levels are slightly underestimated compared to the static case (83 dB and 79 dB), probably because of the proximity of the two sources in the moving context. This level is not calculated for CBF because of the excessive dispersion.



(a) CLEANT *(b) CBF*

Figure 4: Comparison of deconvolution output for white noise moving sources.

This experiment with moving sources shows the improvement brought by CLEANT compared to CBF. The deconvolution map is precise and gives clear positions for the sources. The sources levels are estimated within 1 dB lower than in a static context.

4 EXPERIMENTAL APPLICATION ON ROTATING SOUND SOURCE

A validation experiment is also conducted for the case of a rotating source. A special prototype of rotating source is used, already presented in the frame of a previous study [6]. A 2-channels generator, an amplifier and two loudspeakers are embedded on an equilibrated rotating disc (see close-up view in Fig. 5, right). The disc is mounted on an electric motor and can be operated up to a rotation speed of 900 rpm. A tachometer is used to determine the angular location of the disc as a function of time. A 30 microphone array is placed parallel to the source plane, at a distance of 15 cm (see Fig. 5,left).

Figure 5: Experimental setup : rotating source in front of a 30 microphone array (left). Close up view of the rotating source with the two loudspeakers (left).

The CLEANT algorithm proposed in this work is tested for 6 different source configurations: 3 different rotation speeds (0, 300 and 900 rpm), for either 1 or 2 sources activated. The results obtained for the configuration (2 sources, 900 rpm) are given in Fig. 6 at different iterations of the algorithm, in order to illustrate the efficiency of CLEANT in these conditions. After only 2 iterations, the two sources are correctly localized and quantified. The position of the sources are perfectly localized (within the limitation of the source grid resolution) for all source configurations.

The ability of the method to correctly quantify source strength in any operating condition is illustrated in Table 2, in which the acoustic power level identified for each source is given for all source configurations. The identified acoustic powers for each of the two sources for different rotation speeds show very little dispersion, less than 1dB. For configurations with 2 active sources, source number 2 is systematically found about 1 dB louder than source number 1, which shows the ability of the method to correctly rank the noise sources even in the case of a very small difference of level.

5 CONCLUSIONS

CLEANT is a new approach for the deconvolution of a beamforming output using a matching pursuit inspired algorithm, operating exclusively in time domain and available for moving

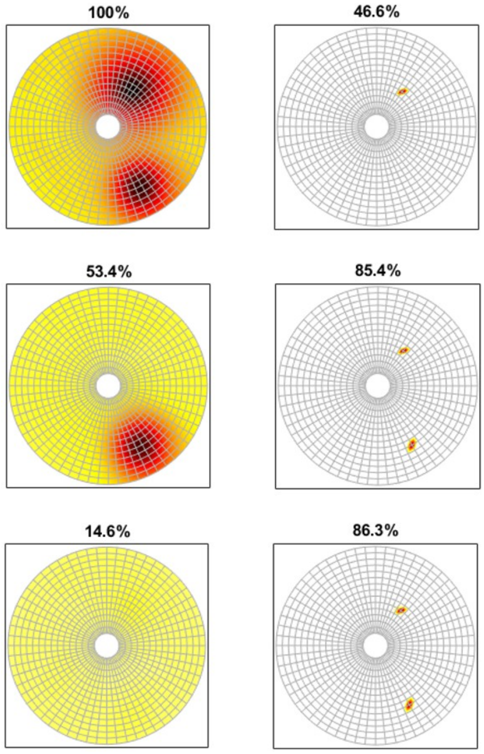


Figure 6: Source rotation speed: 900 rpm. Evolution of Dirty (left) and Clean (right) maps, iteration # 1 (top) to 3 (bottom). Wideband integration [0.5-5] kHz. Loop gain = 1. Percentage values above dirty maps (left) represent the ratio of residual acoustic pressure that is used to obtain the map. Percentage above clean maps (right) represents the part of the measured pressure that is recovered by the identified sources.

Table 2: Source acoustic power levels identified for all source configurations and rotation speeds. Levels are integrated in the frequency range $[0.5-5]kHz$

Speed	**0 rpm**		**300 rpm**		**900 rpm**	
Active source	S1 only	S1 and S2	S1 only	S1 and S2	S1 only	S1 and S2
S1 Level (dB)	59.4	59.2	58.8	58.7	59.4	59.1
S2 Level (dB)	n/a	60.3	n/a	59.8	n/a	59.7

sources as well. Its efficiency has been proved in various experimental situations presented in this paper. In the case of sources in linear motion, it has been compared to conventional beamforming followed by a deconvolution based on the static point spread function: CLEANT showed a great improvement, either for tone or white noise sources. In the case of rotating sources, it provided accurate results in only few iterations, with a very satisfactory quantification and source ranking ability.

The strength of this method is to be applicable to any beamforming output performed in time domain which is well adapted to the identification of moving sources, whether in linear or rotating motion.

ACKNOWLEDGEMENTS

This work was performed within the framework of the Labex CeLyA of Université de Lyon, operated by the French National Research Agency (ANR-10-LABX-0060/ ANR-11-IDEX-0007).

REFERENCES

[1] "Rotating beamforming - motion-compensation in the frequency domain and application of high-resolution beamforming algorithms." *Journal of Sound and Vibration*, 333(7), 1899–1912, 2014.

[2] B. Barsikow. "Experiences with various configurations of microphone arrays used to locate sound sources on railway trains operated by the DB AG." *Journal of Sound and Vibration*, 193(1), 283–293, 1996. URL `http://www.sciencedirect.com/science/article/pii/S0022460X96902692`.

[3] T. F. Brooks and W. M. Humphreys. "A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays." *Journal of Sound and Vibration*, 294(4–5), 856–879, 2006. ISSN 0022-460X. doi:10.1016/j.jsv.2005.12.046. URL `http://www.sciencedirect.com/science/article/pii/S0022460X06000289`.

[4] R. Cousson, M.-A. Pallas, Q. Leclère, and M. Bérengier. "Identification of acoustic moving sources in the context of a road vehicle at pass-by." page 040012. 2017. doi:10.1121/2.0000645. URL `http://asa.scitation.org/doi/abs/10.1121/2.0000645`.

[5] V. Fleury and J. Bulté. "Extension of deconvolution algorithms for the mapping of moving acoustic sources." *The Journal of the Acoustical Society of America*, 129(3), 1417–1428, 2011. ISSN 1520-8524. doi:10.1121/1.3531939.

[6] L. Geng, Q. Leclere, A. Pereira, and C. X. Bi. "Identification of rotating sound source based on microphone array measurements." In *Congrès Français de Mécanique (CFM 2015)*. Lyon, France, 2015. URL `https://hal.archives-ouvertes.fr/hal-01277227`.

[7] G. Herold and E. Sarradj. "Frequency domain deconvolution for rotating sources on an axial fan." In *Proceedings of BEBEC 2016*. Berlin, Germany, 2016.

[8] J. A. Högbom. "Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines." *Astronomy and Astrophysics Supplement Series*, 15, 417, 1974. ISSN 0365-0138. URL `http://adsabs.harvard.edu/abs/1974A%26AS...15..417H`.

[9] L. Lamotte, B. Nicolas, M. Q. Pham, and B. Oudompheng. "A theoretical and experimental comparison of the deconvolution methods for moving sources." In *Berlin Beamforming Conference (Bebec*. 2016. URL `http://www.bebec.eu/Downloads/BeBeC2016/Papers/BeBeC-2016-S7.pdf`.

[10] C. L. Lawson and R. J. Hanson. *Solving Least Squares Problems*. Prentice-Hall, 1974.

[11] O. Minck, N. Binder, O. Cherrier, L. Lamotte, and V. Pommier-Budinger. "Fan noise analysis using a microphone array." In *Fan 2012 - International Conference on Fan Noise, Technology, and Numerical Methods*. Senlis, France. URL `http://oatao.univ-toulouse.fr/5737/`.

[12] B. Oudompheng. *Array processing for the localization and the contribution of acoustic sources of a passing-by ship using a short antenna*. Theses, Université Grenoble Alpes, 2015. URL `https://tel.archives-ouvertes.fr/tel-01230591`.

[13] Sijtsma, Oerlemans, and Holthusen. "Location of Rotating Sources by Phased Array Measurements." In *7th AIAA/CEAS Aeroacoustics Conference 28-30 May 2001 Maastricht, Netherlands*. 2001.