Undergraduate Research Project Report

Development of an algorithm to track a moving acoustic source

Student:
Edo Barlell
ID: 318660982

Supervisor:
Assistant Professor Dr. Oksana Stalnov

Faculty of Aerospace Engineering
April 22, 2021

Abstract

Phase microphone arrays have become state-of-the-art tool in many fields of application. Such a device is commonly used for determining the location and magnitude of acoustic sources. The demand of the ability to pin-point the location of an acoustic source is constantly rising. As the use of rotor-based technologies has been spreading, the noise generated by rotors became a concerning issue. In order to understand the rotor noise and deal with it, an understanding of the source location and behavior is required. For stationary sources, algorithms working in the frequency or time domain can be applied. Unfortunately, the stationary sources locating algorithms cannot identify and locate high-speed rotating acoustic sources properly.

Objects undergoing a motion are subjected to multiple aeroacoustics effects. As the sensor array measures the signal from the acoustic source of interest, the signal is compromised and does not represent the true behavior and location of the source. Since standard beamforming methods cannot be applied directly to rotating sources such as wind turbines, UAVs and other open rotor aerial vehicles, in the frequency domain, rotating sources require a special configuration that preprocess the measured data and correct both frequency and amplitude.

In this project the relation between the acoustic pressure measured by the sensor array to the actual signal of the acoustic source will be described while considering the different factors that defect the measurement. And a different method of beamforming, suitable even for dealing with a high-speed rotating source, will be presented. With a good understanding of the cause and effect relations of the acoustic effects on the signal, it is possible to break it down and extract the desired information.

1 Background and Introduction

Phase microphone arrays is a device that uses an array of microphones at determined locations that are collecting sound pressure measurements simultaneously. As a source of an acoustic signal transmit from a certain location in space, the phase microphone array (PMA) measurements can be used to determine the exact location of the source relative to the PMA. Due to the different locations of every sensor in the array, the distance between each sensor to the acoustic source is slightly different. Under the assumption that the signal is transmitted to all the sensors through the same medium at the same speed of sound, each sensor will measure the same signal but with a certain margin in phase compered to the other sensors. The method that takes into account the difference in phase between the measurements and corresponds them with the different sensors locations is called Beamforming Method.

The principle behind the widely used beamforming algorithms is to shift the measured time signals to compensate the sound travel time from a focus point. If the focus point contains a source, a summation of the signals will reveal a maximum; otherwise, the signals mostly cancel each other out. As the frame of interest is settled and divided to small focus points, the beamforming method uses a cross correlation matrix between all the different microphone measurements in relation to every focus point and that way "map" out all the acoustic sources in that frame of interest.

Beamforming can be performed in the time domain or in the frequency domain. In the time domain, the signals are directly delayed and summed. In the frequency domain, the signals are first Fourier-transformed and then their cross-spectra are averaged. A multiplication of the resulting complex cross-spectral matrix (CSM) with steering vectors fulfills the task of phase-shifting the signals according to the focus points and subsequent summation. The direct averaging of the cross-spectra has the advantage that non-continuous sources are canceled out before the beamforming is performed. Furthermore, uncorrelated background noise can be accounted for via removal of the main diagonal of the CSM.

In the case of a moving source, a time domain formulation is often used, since the distances from focus points to a microphone constantly changes and this time depending phase-shift cannot be taken into account with pre-averaged data, as is the case in frequency domain. The most famous beamforming to identify rotating sound source is the time-domain ROSI beamforming proposed by Sijtsma et al. [1]. Nowadays ROSI is still the widely recognized beamforming for array with arbitrary microphone configuration to identify rotating sound source. Moreover, acoustic effects, such as the Doppler Effect for example, change the signal measured by the PMA in both magnitude and frequency compered to the original signal.

Wavelet transform is a method that enables reconstruction of a two-dimensional signal, such as acoustic pressure wave, as a complex matrix that describes the original signal frequency in dependency of time. With that, the connection between the time and the frequency changes is formed. By looking at the wavelet transform result at a specific time of measurement, and applying a time delay correction, the frequency related to the actual time of transmission can be obtained, and by using a frequency correction factor on that, the actual transmitting frequency is gained. Which means that by using wavelet transform, a beamforming method that combines frequency domain and time domain beamforming methods can be constructed and provide an efficient tool in the case of a moving sources.

2 Objectives

The main goal of the project was to develop an acoustic camera capable of tracking high-speed rotating acoustic sources. Such a device is essential for multiple applications in the process of research and development of acoustic related projects. therefore, it was decided that developing and assembling such a device will provide an essential tool for future work at the Aeroacoustics and Unsteady Aerodynamics Laboratory in at the Faculty of Aerospace Engineering, Faculty at the Technion – Israel Institute of Technology.

The research project focused on implementing the theoretical method of wavelet-based beamforming algorithm [2] and modify the method in terms of efficiency and reduction of the assumptions required.

3 Model

3.1 Broadband Array Design

3.1.1 Background

Array design [3] is very important when phased array measurements are to be analyzed. Depending on the required aeroacoustic application of the array, the array design may vary. In order to achieve the goals of this project, as in most aeroacoustic applications, the focused array approach is the most fitting. Focusing the array means selecting phase delays to look at a particular point in space and implies spherical wave (near-field) propagation.

A primer step to the array design would be defining the array's performance characteristics. Array resolution, spatial aliasing, and array sidelobe suppression are to be considered in relation to the task at hand. Array resolution describes the ability of the array to delineate where a source of noise is located. Spatial aliasing is the appearance of noise sources at locations other than that from which the noise is actually being generated. This side effect may be caused by a bad design or sound reflecting from surrounding objects and the array itself. Array sidelobe suppression is the ability of the array to ignore sources coming from directions other than where the array is being focused.

3.1.2 Array Design Considerations

The challenge with planar array designs is that they must be applicable for a broad range of frequencies and consider the size of the frame of interest in comparison to the array aperture. For maximum resolution, the array should be positioned perpendicular to the acoustic waves. The far-field effective aperture dimension D_f is given by

$$D_f = D \cos \phi \tag{1}$$

where D is the actual aperture dimension and ϕ is the angle relative to the perpendicular axis of the array. Hence, the maximum resolution is obtained at broadside to the array.

The near-field effective aperture dimension is similar but also depending upon the focus point. If the source is located at the point (x_s, z_s) in a coordinate system with the origin at

the center of the array, then the near-field effective aperture dimension D_n is given by

$$D_n = 2r \sin \left[\frac{\frac{\pi}{2} - \arctan\left(\frac{z_s}{x_s + D/2}\right) - \arctan\left(\frac{x_s - D/2}{z_s}\right)}{2} \right] \tag{2}$$

where $r = \sqrt{x_s^2 + z_s^2}$.

By referring to this relation, and taking into account all the focus points of interest, a general understanding of the array's aperture size requirement for optimal resolution is formed (e.g. Fig 1).

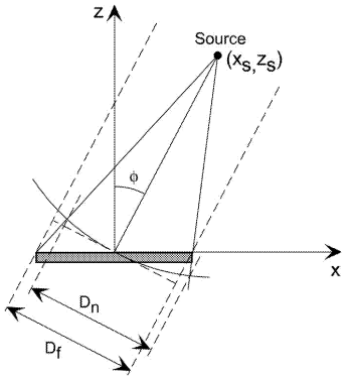


Figure 1: Effective aperture dimension D_f ,for a far-field source is shown by the dashed line representing the wave front of a plane wave cutting through center of the array aperture. Effective aperture dimension D_n, for a near-field source at (x_s, z_s) is shown by a solid curve representing the spherical wave front of a point source cutting through the center of the array aperture [3].

To get rid of spatial aliasing, the conventional approach requires a half-wavelength spacing between sensors. therefore, the lowest frequencies of interest demand a larger aperture size, and the highest frequencies drive the sensor spacing smaller. These combined requirements lead to large sensor count requirements, typically beyond what is practical from a cost standpoint.

An approach suggested to address this problem is the sparse arrays [4]. This approach removes as many sensors as possible while still filling the spacing requirement. Although this method is very effective at reducing the number of sensors required, it has been shown to cause undesirable asymmetries for near-field source mapping as well as increased sidelobe levels.

Another method of arrays is known as irregular arrays. One method of designing an irregular array is to determine sensor locations randomly. Another method is to use an algorithm that guarantees irregularity in sensor locations. The second method should be used when it is possible to specify and control sensor location. Irregular arrays provide a way to eliminate the requirement for half-wavelength sensor spacing while also avoiding spatial aliasing.

3.1.3 Spiral Array

In the development of the project presented in this report, it was decided to use an irregular array because of the advantages specified before. The array design chosen is called multi-arm spiral array [5] which is based on multiple logarithmic spiral arrays [6]. Such an array provide particularly effective sidelobe control over a broad range of frequencies with a limited number of sensors. The spiral array design equally spaces the array sensors on a logarithmic spiral, which provide sufficient sensor spacing. The polar equation for a logarithmic spiral is

$$r(\theta) = r_0 exp[\cot(\nu)\theta] \tag{3}$$

where, as shown in Fig 2, θ is the polar angle in radians, r is the distance from the origin O to the point P on the spiral, r_0 is the value of r when $\theta = 0$, and $0 \leq \nu \leq \pi/2$ is the spiral angle. The spiral angle is the constant angle at which radii from O are cut by the spiral curve. To specify equally spaced sensor locations along the spiral, an expression for the distance PN, the radius of curvature at the point P is needed. Referring again to Fig 2, if PT is tangent to the spiral at the point P, PN is normal to the spiral at the point P, and TON is perpendicular to OP, then

$$PN = \sqrt{r^2 + (ON)^2} \tag{4}$$

Since $ON = r\cot(\nu)$,

$$PN = r\sqrt{1 + \cot^2(\nu)} \equiv r_p \tag{5}$$

where r_p is the radius of curvature at the point P.

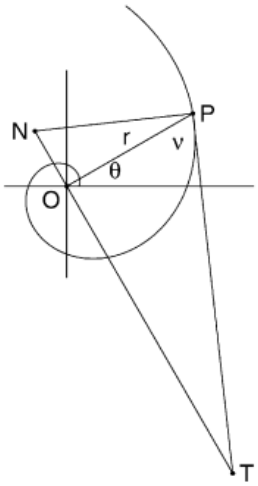


Figure 2: Geometry for a logarithmic spiral. OP is perpendicular to NT. NP is perpendicular to PT. PT is tangent to the spiral at the point P. The spiral angle ν is the constant angle at which radii from O are cut by the spiral curve [3].

Now calculate the arc length

$$l(\theta) = \int_0^\theta dl = \int_0^\theta r_p d\theta = \int_0^\theta r\sqrt{1+\cot^2(\nu)}d\theta = \int_0^\theta r_0 exp[\cot(\nu)\theta]\sqrt{1+\cot^2(\nu)}d\theta \quad (6)$$

Make the substitution $h = \cot(\nu)$ and integrate to get

$$l(\theta) = r_0\sqrt{1+h^2}\int_0^\theta exp(h\theta)\,d\theta = \frac{r_0\sqrt{1+h^2}}{h}[exp(h\theta)-1] \quad (7)$$

The length of the entire spiral, l_{max}, will depend on θ_{max}, the maximum value of θ, which depends on r_{max}, the maximum value of r. From the original equation for the spiral

$$r_{max}(\theta_{max}) = r_0 exp(h\theta_{max}) \quad (8)$$

so

$$\theta_{max} = \frac{1}{h}\ln\left(\frac{r_{max}}{r_0}\right) \quad (9)$$

and

$$l_{max} = \frac{r_0\sqrt{1+h^2}}{h}[exp(h\theta_{max})-1] = \frac{r_0\sqrt{1+h^2}}{h}\left(\frac{r_{max}}{r_0}-1\right) \quad (10)$$

With those equations at hand, the one additional variable the array design is dependent on is the amount of sensors available. Usually the desire is to use the minimum amount of sensors required to get the defined array performances. The maximum radius is typically chosen based on desired array resolution (as explained in the previous sub-chapter). The spiral angle controls the number of revolutions the spiral will make around the origin.

The process for determining the sensor coordinates of a spiral array is as follows:

1. Select values for r_0,r_{max},ν and M, where M is the number of sensors in the array.
2. Compute l_{max} using Eq. (10).
3. Compute $\Delta l = l_{max}/(M-1)$, the delta arc-length between sensors.
4. Compute the sensor coordinates for $m = 1, 2, ..., M$

$$l_m = \Delta l\,(m-1) \quad (11)$$

Determine θ_m by rewriting Eq. (7) for $l(\theta)$ as

$$\theta(l) = \frac{1}{h}\ln\left(1+\frac{hl}{r_0\sqrt{1+h^2}}\right) \quad (12)$$

and substituting $l = l_m$ to get

$$\theta_m = \frac{1}{h}\ln\left(1+\frac{hl_m}{r_0\sqrt{1+h^2}}\right) \quad (13)$$

and finally

$$r_m = r_0 exp(h\theta_m). \quad (14)$$

Spiral arrays provide the advantages of an irregular array, it is able to support a relatively large aperture with a modest number of sensors, and maintaining sidelobe suppression over a broad range of frequencies. Multi-arm spiral arrays add additional benefits including nearly circular symmetry and additional sidelobe suppression. A multi-arm spiral array combines a spiral array with odd-number-of-element circular arrays and therefore provide the benefits of both. Circular array with an odd number of elements equally distributed around the circumference of the circle is formed for each element of the logarithmic spiral array. Equal arc-length spacing for logarithmic spirals results in linear spacing with respect to radius. More densely packing the sensors at smaller radii has the effect of broadening the mainlobe of the array while lowering the average sidelobe level. For broadband array utility, it seems that balanced weighting across all spatial lags proves most beneficial for peak sidelobe control.

One strategy that works particularly well is sampling such that each sensor samples an equal aperture area. A circular aperture is divided up into concentric annuli of equal area and the logarithmic spiral is sampled at the radial center of each annulus. An odd-number-of-element circular array is then formed for each sample thereby dividing each annulus into N equal-area sectors where N is the number of odd elements per circular array. The result is shown in Fig 3. The multi-arm spiral array design parameters are:

1. The innermost radius of the spiral
2. The outermost radius of the spiral
3. The spiral angle
4. The number of elements per spiral
5. The spiral sampling strategy
6. The (odd) number of elements per circle

These design parameters provide a class of arrays that in general work quite well to produce arrays that have plateau-like sidelobes across a broad range of frequencies.

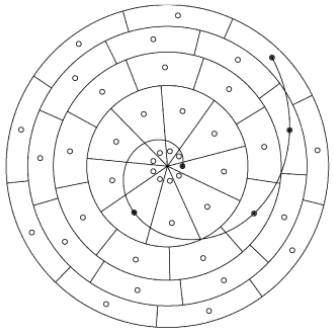


Figure 3: A multi-arm logarithmic spiral array. The circular aperture is divided up into concentric annuli of equal areas, and each annulus is divided into N equal-area sectors. Each equal-area section of the array aperture contains a single sensor [3].

3.2 Conventional Beamforming

3.2.1 Background

Conventional Beamforming [7] is the widespread term referring to two basic beamforming methods, which provide the fundamentals of beamforming that most of the advanced beamforming methods relay upon. Time-Domain (TD) and Frequency-Domain (FD) beamforming methods are the base of noise source identification and location. The TD beamforming is used to locate acoustic sources mostly of moving objects such as airplanes and wind turbines, where the accuracy of the location is key. As for the FD bemforming, it is used to locate acoustic sources of stationary objects, where the emphasis is on locating sources of specific frequency. Overall, the FD beamforming offers many advantages over TD beamforming, however it is heavily effected by frequency distortions associated with moving objects.

3.2.2 Time-Domain Beamforming

From a physical point of view, Time-Domain Beamforming is a simple and straight forward method in which the measured signals are directly delayed and summed.

Consider an array of N-sensors, the microphones in the array are located at $\overrightarrow{\mathrm{x_n}}\,(n = 1, 2, ..., N)$. For convenience sake, the origin of the coordinate system is located at the center of the array. If a monopole source is located at $\overrightarrow{x_s}$ and transmitting a signal $q\,(\tau) = q_0 e^{i\omega\tau}$, then the acoustic pressure recorded by the n^{th} microphone in the array would be

$$p\,(\overrightarrow{\mathrm{x_n}}, t) = \frac{q\,(\tau)}{4\pi|\overrightarrow{\mathrm{x_n}} - \overrightarrow{\mathrm{x_s}}|} = \frac{q\,(t - |\overrightarrow{\mathrm{x_n}} - \overrightarrow{\mathrm{x_s}}|/c)}{4\pi|\overrightarrow{\mathrm{x_n}} - \overrightarrow{\mathrm{x_s}}|} \tag{15}$$

where $p\,(\overrightarrow{\mathrm{x_n}}, t)$ is the acoustic pressure in Pa of the n^{th} microphone at the reception time, $q\,(\tau)$ is the source strength of the monopole in kg/s^2, t is the time at which the sound arrives to the observer location (reception time), c is the speed of sound in m/s, and $\overrightarrow{x_n}$ and $\overrightarrow{x_s}$ are the observer and the source position vectors in meters, respectively.

The argument of the source strength represents the time at which the sound was emitted and thus it is known as the emission time, denoted by τ. The time for a sound signal to propagate from the source to the observer location is given by the ratio of the distance between them and the speed of sound, the time delay is described as

$$\Delta t_n = \frac{|\overrightarrow{\mathrm{x_n}} - \overrightarrow{\mathrm{x_s}}|}{c} \tag{16}$$

Now, let the location of the source be unknown, and let $p_n\,(t)$ be the time series of the acoustic pressure recorded by the n^{th} microphones of the array at a sampling frequency f_s. Then, a beamforming algorithm is used to determine the source location $\overrightarrow{x_s}$. A frame of interest is settled and divided to small focus points $\overrightarrow{x_b}$, which are assumed to be potential locations of the source. Then, the time delay and amplitude of the microphone data are adjusted in such a way as if they were originated at $\overrightarrow{x_b}$ and received at the n^{th} sensor location $\overrightarrow{x_n}$, and then added up as follows

$$z\,(t) = \frac{1}{N}\sum_{n=1}^{N} w_n p_n\,(\overrightarrow{x_n}, t - \Delta_m) \tag{17}$$

where $z(t)$ is termed the beamforming output, and Δ_m is the time delay given by

$$\Delta_m = -\frac{|\overrightarrow{x_n} - \overrightarrow{x_b}|}{c} \tag{18}$$

and the term w_n in Eq. (17) is an amplitude adjustment factor given by

$$w_n = 4\pi|\overrightarrow{x_n} - \overrightarrow{x_b}| \tag{19}$$

The beamforming is then implemented by replacing Eq. (15),(18), and (19) into Eq. (17), yielding

$$\begin{aligned} z(t) &= \frac{1}{N}\sum_{n=1}^{N} 4\pi|\overrightarrow{x_n} - \overrightarrow{x_b}| p_n\left(\overrightarrow{x_n}, t + \frac{|\overrightarrow{x_n} - \overrightarrow{x_b}|}{c}\right) \\ &= \frac{1}{N}\sum_{n=1}^{N} 4\pi|\overrightarrow{x_n} - \overrightarrow{x_b}| \frac{q\left(t - \frac{|\overrightarrow{x_n} - \overrightarrow{x_s}|}{c} + \frac{|\overrightarrow{x_n} - \overrightarrow{x_b}|}{c}\right)}{4\pi|\overrightarrow{x_n} - \overrightarrow{x_s}|} \end{aligned} \tag{20}$$

Eq. (20) shows that when the potential source location coincides with the actual source location, i.e., $\overrightarrow{x_b} = \overrightarrow{x_s}$ the beamforming output equals the strength of the actual source, i.e. $z(t) = q(t)$. However, when the potential source location does not coincide with the actual sourcenlocation, i.e., $\overrightarrow{x_b} \neq \overrightarrow{x_s}$ the beamforming output will have a value less than the actual source strength. A visual example for that is shown in Fig. 4

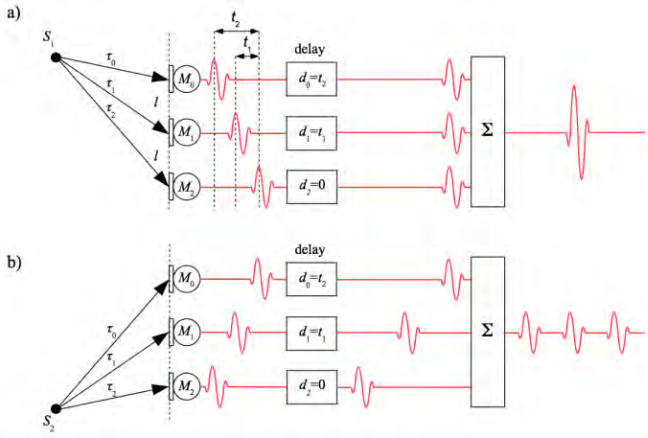


Figure 4: In case a. the source is located at S_1, and in case b. the source is located at S_2. In both cases the source is assumed to be at $\overrightarrow{x_b} = S_1$, therefore the time delay applied in both cases is the same for each sensor and calculated for $\overrightarrow{x_b} = S_1$. The delayed signals are then summed and normalized by the number of sensors. In case a. the final value would be equals the strength of the actual source, and in case b. it will less of it [8].

To analyze the frequency content of the sound emitted by the source, the general approach is to take the beamforming output time signal and convert it to the frequency domain using discrete Fourier transform (DFT) as

$$Z(\overrightarrow{x_b}, f_k) = DFT\{z(\overrightarrow{x_b}, t)\} \tag{21}$$

Then, the mean square value of this transformed signal is estimated using the following expression:

$$z^2_{rms}\left(\overrightarrow{x_b}, f_k\right) = \frac{1}{2}\frac{1}{N_b}\sum_{N_b}\left|Z\left(\overrightarrow{x_b}, f_k\right)\right|^2 = \frac{1}{2}\langle\left|Z\left(\overrightarrow{x_b}, f_k\right)\right|^2\rangle \tag{22}$$

where $\langle\rangle$ denotes average, and N_b is the number of data blocks used in the estimation of the average msv. Finally, the decibel values of the average msv beamforming output are computed and plotted as acoustic maps.

3.2.3 Frequency-Domain Beamforming

In this section, the concept of delay-and-sum beamforming is extended to the frequency domain. In frequency domain beamforming, the microphone time data signals are to be transformed into the frequency domain. The equations that allow that transform, are the so-called Discrete Fourier Transform (DFT) given as

$$y(n) = \frac{1}{N}\sum_{k=1}^{N-1} Y(k)e^{i\omega k\Delta t} \tag{23}$$

where,

$$Y(\omega) = \sum_{n=0}^{N-1} y(n)e^{-i\omega k\Delta t} \tag{24}$$

Eq. (23) and Eq. (24) are known as the Inverse Discrete Fourier Transform (IDFT) and the Discrete Fourier transform (DFT) equations, respectively. Eq. (24) is the starting point of any frequency domain analysis as it allows representing any time domain discrete signal of finite duration as a combination of sinusoidal basis functions. The property that makes this transformation possible is the orthogonality of the basis functions which can be expressed as:

$$\sum_{n=0}^{N-1}(e^{i\omega k\Delta t})\cdot(e^{-i\omega k'\Delta t}) = N\delta_{kk'} \tag{25}$$

where $\delta_{kk'}$ is the Dirac function.

With the transformed signals the FD beamforming process may start. Similarly to the TD beamforming,assume the source of interest is a monopole located at $\overrightarrow{x_s}$. Taking the Fourier Transform of the sound field of a monopole source in Eq. (15) yields the acoustic field in the frequency domain as

$$P(\overrightarrow{x}, f) = \frac{Q(f)}{4\pi|\overrightarrow{x} - \overrightarrow{x_s}|}\cdot e^{-i2\pi f\Delta t} \tag{26}$$

where $Q(f)$ is the Fourier Transform of the monopole source strength $q(t)$, and the exponential term $e^{-i2\pi f\Delta t}$ accounts for the phase difference between the source and the microphone signal which is related to the propagation time

$$\Delta t = \frac{|\overrightarrow{x} - \overrightarrow{x_s}|}{c} \tag{27}$$

Therefore, Eq. (26) is the frequency domain version of Eq. (15). For simplicity, Eq. (26) can be re-written as

$$P(\overrightarrow{x}|\overrightarrow{x_s}, f) = Q(f)\cdot G(\overrightarrow{x}|\overrightarrow{x_s}, f) \tag{28}$$

where P,Q, and G are complex, and G is given as

$$G(\overrightarrow{x}|\overrightarrow{x_s},f) = \frac{e^{-i2\pi f\frac{|\overrightarrow{x}-\overrightarrow{x_s}|}{c}}}{4\pi|\overrightarrow{x}-\overrightarrow{x_s}|} \tag{29}$$

Eq. (29) is the free-field Green's function of a monopole source.

As shown in the previous subsection, time domain beamforming is achieved by delaying the microphone signals, adjusting their amplitude and summing the signals over the array microphones to steer the array to a particular location $\overrightarrow{x_b}$ where a potential source might be present. The equivalent of a time delay in the time domain is a phase shift in the frequency domain. Therefore, beamforming in the frequency domain is achieved by applying phase shifts and amplitude adjustments to the microphone signals to steer the array to a particular location $\overrightarrow{x_b}$.

The source strength estimated from the signal of the n^{th} microphone can be obtained from Eq 28 as

$$Q_n(f) = \frac{G(\overrightarrow{x}|\overrightarrow{x_s},f)^* P(\overrightarrow{x}|\overrightarrow{x_s},f)}{|G(\overrightarrow{x}|\overrightarrow{x_s},f)|^2} \tag{30}$$

where the $*$ denotes complex conjugate. Therefore, the beamforming output from the N microphones of the array can be obtained from the following expression,

$$Z(\overrightarrow{x_b},f) = \frac{1}{N}\sum_{n=1}^{N} Q_n(f) \tag{31}$$

Eq. (31) is then used to compute the mean square value as,

$$Z^2_{rms}(\overrightarrow{x_v},f) = \frac{1}{2}\langle|Z(\overrightarrow{x_v},f)|\rangle = \frac{1}{2N^2}\langle\sum_{m=1}^{N}\sum_{n=1}^{N} Q_m(f)Q_n(f)^*\rangle \tag{32}$$

Substituting Eq. (30) into Eq. (32) yields,

$$Z^2_{rms}(\overrightarrow{x_v},f) = \frac{1}{2N^2}\langle\sum_{m=1}^{N}\sum_{n=1}^{N} \frac{G_m(\overrightarrow{x}|\overrightarrow{x_s},f)^*}{|G_m(\overrightarrow{x}|\overrightarrow{x_s},f)|^2} P_m(\overrightarrow{x}|\overrightarrow{x_s},f)P_n(\overrightarrow{x}|\overrightarrow{x_s},f)^* \frac{G_n(\overrightarrow{x}|\overrightarrow{x_s},f)}{|G_n(\overrightarrow{x}|\overrightarrow{x_s},f)|^2}\rangle \tag{33}$$

which is the expression for the beamforming output for conventional frequency domain beamforming.

3.3 Wavelet Beamforming

3.3.1 Rotating Source Beamforming Issue

In the particular case, when the source is stationary, the sampling period in all the signals is the same. In other words, if the source emits pulses of sound at equally spaced intervals of time, these pulses will arrive to the microphone with the same equally spaced intervals of time. This does not apply when dealing with a moving source. During the pulses interval emission time the source changes it's location, which cause a change in the time delay between the source and the sensor. As the next pulses transmit the signals, they will arrive to the microphone with an increasing spaced intervals or decreasing spaced intervals whether the source is moving farther or closer to the sensor respectively.

In addition, due the movement of the source the acoustic signal is subjected to changes in frequency and phase due the Doppler effect, and changes in magnitude due convective amplification. By attempting conventional TD beamforming, even if the changing time delays are considered, the effects in frequency and magnitude corrupt the summed signal and make the method ineffective. As for FD beamforming, block averaging of Fourier transform results in the frequency domain would smear the resultant acoustic images across the moving trajectories.

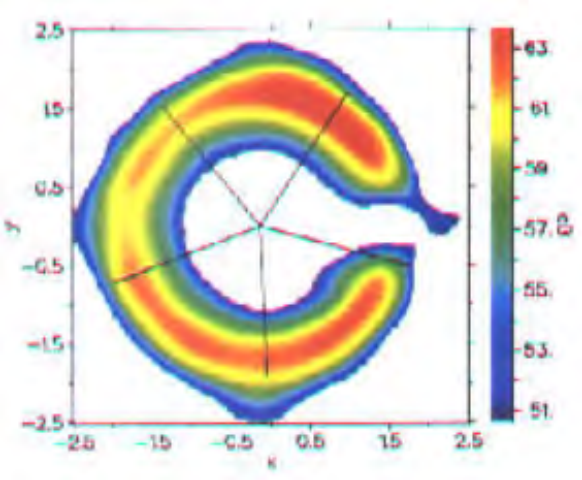


Figure 5: Application of conventional beamforming assuming non-rotating sources on rotating sources of helicopter in hover [1].

This issue can be solved even with the conventional beamforming methods by approaching deconvolution type post-processing on the signals. Although post-processing the signals may solve the problem, it costs a lot of computation time and overall it is not an efficient approach to address the issue. A better approach would be Adaptive beamforming method [9], such method can save significant amounts of post-processing time for a deconvolution method. Therefore, adaptive beamforming is considered as a promising signal processing method. However, this method is susceptible to background noise and interference. All of these difficulties provided the motivation used to develop the Wavelet-Based beamforming'[2] method.

3.3.2 Acoustic Pressure of Rotating Source

As seen in sub-section 3.2.2 and 3.2.3, at the beginning of any beamforming algorithm a good understanding of the acoustic pressure field is desired (Eq. (15) and (26)). Consider a rotating monopole acoustic source, transmitting a signal $q(t) = q_0 e^{i\omega t}$ in a uniform and stationary medium. We assume stationary medium in order to neglect convective propagation effects. Furthermore, the velocity of the source is defined by $\overrightarrow{v}(t)$ and the source location at time 0 is defined as $\overrightarrow{x_0}$. The associated sound pressure field p can be described by the associated governing equation as follows:

$$\left(\frac{1}{c_0^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) p = \frac{\partial}{\partial t}\left(q(t)\delta\left(\overrightarrow{x_n} - \int_0^t \overrightarrow{v}(\tau')\cdot d\tau' - \overrightarrow{x_0}\right)\right) \tag{34}$$

where c_0 is the speed of sound, δ is the Dirac delta function, and $\overrightarrow{x_n}$ is the space coordinates at which the pressure field is measured. Next, by using the relationship between the velocity potential and acoustic pressure, $p = -\partial\phi/\partial t$. Eq. (34) can be rewritten as:

$$\left(\frac{1}{c_0^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) \phi = -q(t)\delta\left(\overrightarrow{x_n} - \int_0^t \overrightarrow{v}(\tau')\cdot d\tau' - \overrightarrow{x_0}\right) \tag{35}$$

The solution for the velocity potential based on the free field Green's function is given by

$$\phi(\overrightarrow{x_n}, t) = \int_{-\infty}^{\infty} \frac{-q(t - \frac{|\overrightarrow{x_n} - \overrightarrow{x}'|}{c_0})\delta(F(\overrightarrow{x}'))}{4\pi|\overrightarrow{x_n} - \overrightarrow{x}'|} d\overrightarrow{x}' \tag{36}$$

where $F(\overrightarrow{x}')$ is defined as follows:

$$F(\overrightarrow{x}') = \overrightarrow{x}' - \int_0^{t - \frac{|\overrightarrow{x_n} - \overrightarrow{x}'|}{c_0}} \overrightarrow{v}(\tau')\cdot d\tau' - \overrightarrow{x_0} \tag{37}$$

Considering $F(\overrightarrow{x}') = 0$ when $\overrightarrow{x}' = \overrightarrow{x_s}$, where $\overrightarrow{x_s}$ is the source location at time τ, where

$$\tau = t - \frac{|\overrightarrow{x_n} - \overrightarrow{x_s}|}{c_0} \tag{38}$$

Then, we utilize the following Jacobian property of Dirac δ function:

$$\delta(F(\overrightarrow{x}')) = \left|\frac{\partial F(\overrightarrow{x_s})}{\partial \overrightarrow{x}'}\right|^{-1} \delta(\overrightarrow{x}' - \overrightarrow{x_s}) \tag{39}$$

therefore, the solution of the velocity potential is

$$\phi(\overrightarrow{x_n}, \overrightarrow{x_s}, t) = \frac{-q(t - \frac{|\overrightarrow{x_n} - \overrightarrow{x_s}|}{c_0})}{4\pi|\overrightarrow{x_n} - \overrightarrow{x_s}|} \cdot \left|\frac{\partial F(\overrightarrow{x_s})}{\partial \overrightarrow{x}'}\right|^{-1} \tag{40}$$

An association can be made between $\overrightarrow{x_n}$, the space coordinates at which the pressure field is measured, to the defined location of the n^{th} sensor of the array. For convenience, $R_n = |\overrightarrow{x_n} - \overrightarrow{x_s}|$ is defined to be the distance between the n^{th} sensor and the source's location

at a specific moment τ. Where τ is the generation time of the signal at $\overrightarrow{x_s}$, that will be received by the n^{th} sensor at location $\overrightarrow{x_n}$ at time t. Eq. (40) can be rewritten as

$$\phi(\overrightarrow{x_n}, \overrightarrow{x_s}, t) = \frac{-q(\tau)}{4\pi R_n} \cdot \alpha(\overrightarrow{x_n}, \overrightarrow{x_s}, \tau) \tag{41}$$

where $\alpha(\overrightarrow{x_n}, \overrightarrow{x_s}, \tau)$ is due to the Doppler effect, that is,

$$\alpha(\overrightarrow{x_n}, \overrightarrow{x_s}, \tau) = \left| \frac{\partial F(\overrightarrow{x_s})}{\partial \overrightarrow{x}'} \right|^{-1} = \frac{d\tau}{dt} = \left(1 + \frac{\overrightarrow{v}(\tau)(\overrightarrow{x_s} - \overrightarrow{x_n})}{c_0 |\overrightarrow{x_n} - \overrightarrow{x_s}|} \right)^{-1} \tag{42}$$

Then, the solution of acoustic pressure is given by

$$p(\overrightarrow{x_n}, \overrightarrow{x_s}, t) = \frac{\partial}{\partial t} \left(\frac{q(\tau)}{4\pi R_n} \cdot \alpha(\overrightarrow{x_n}, \overrightarrow{x_s}, \tau) \right) \tag{43}$$

Given the form of a point source $q(t) = q_0 e^{i\omega t}$, the associated acoustic pressure can be written as

$$p(\overrightarrow{x_n}, \overrightarrow{x_s}, t) = \frac{1}{4\pi} \left(\frac{\left(\frac{\partial q(\tau)}{\partial \tau} \cdot \frac{\partial \tau}{\partial t} \cdot \alpha + q(\tau) \cdot \frac{\partial \alpha}{\partial \tau} \cdot \frac{\partial \tau}{\partial t} \right) \cdot R_n - \frac{\partial R_n}{\partial \tau} \cdot \frac{\partial \tau}{\partial t} \cdot q(\tau) \cdot \alpha}{R_n^2} \right) \tag{44}$$

$$= \frac{\alpha q(\tau)}{4\pi R_n} \left(i\omega \cdot \alpha - \left(\frac{\alpha}{R_n} \cdot \frac{\partial R_n}{\partial \tau} - \frac{\partial \alpha}{\partial \tau} \right) \right) \tag{45}$$

$$= \frac{i\omega q(\tau) \alpha^2}{4\pi R_n} \left(1 + \frac{i}{\omega} \left(\frac{1}{R_n} \cdot \frac{\partial R_n}{\partial \tau} - \frac{1}{\alpha} \cdot \frac{\partial \alpha}{\partial \tau} \right) \right) \tag{46}$$

$$= \frac{i\omega q(\tau)}{4\pi R_n} \cdot \alpha^2(\overrightarrow{x_n}, \overrightarrow{x_s}, \tau) \cdot \beta(\overrightarrow{x_n}, \overrightarrow{x_s}, \tau) \tag{47}$$

where β is caused by the nearfield amplification effect, and can be described as

$$\beta(\overrightarrow{x_n}, \overrightarrow{x_s}, \tau) = 1 + \frac{i}{\omega} \left(\frac{\partial}{\partial \tau} \ln \frac{R_n}{\alpha} \right) = 1 + i \frac{\alpha}{\omega} \left(\frac{c_0(v^2/c_0^2 + \alpha^{-1} - 1)}{R_n} + \frac{\partial \overrightarrow{v(\tau)}}{\partial \tau} \cdot \frac{(\overrightarrow{x_s} - \overrightarrow{x_n})}{c_0 R_n} \right) \tag{48}$$

Nevertheless, it is important to remember that when given a rotating source of frequency ω, the n^{th} sensor will perceive a different imaging frequency ω^n owing to the Doppler effect, that is,

$$\omega^n = \omega \cdot \alpha(\overrightarrow{x_n}, \overrightarrow{x_s}, \tau) \tag{49}$$

3.3.3 Wavelet Transform

The first method that comes to mind when approaching signal analysis is the Fourier Transform, which provides frequency information of a signal that represents frequencies and their magnitude. However, it does not provide information about when in time the frequencies exist. This transform is therefore ideal for stationary signals. Short-Time Fourier Transform (STFT) was developed to overcome the poor time resolution of the Fourier Transform. It

gives a time-frequency representation of the signal. By assuming some portions of the non-stationary signal are stationary, the STFT adds up the Fourier transform of each stationary portion along the signal. Where the FT of a signal $f(t)$ is represented by:

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt \tag{50}$$

and STFT by

$$F(\tau, \omega) = \int_{-\infty}^{\infty} f(t)w(t-\tau)e^{-i\omega t}dt \tag{51}$$

where $w(t-\tau)$ is a window function of fixed length that moves along the signal from start to end, and divide the signal to portions in which the signal is assumed to be stationary. And τ is the translation parameter. The problem with this method is that fixed length of the window means time and frequency resolutions are fixed for the entire length of the signal. Which means that the information about what frequencies exist at what time instance is unattainable, only information about what frequency bands exist at what time intervals. The relation between the resolutions is described by

$$\Delta t \Delta f \geq \frac{1}{4\pi} \tag{52}$$

where higher resolution in either means a lower resolution in the other. Low frequency components often last a long period of time, so a high frequency resolution is required. Where high frequency components often appear as short bursts, invoking the need for a higher time resolution.

The Wavelet Transform results in analyzing a signal into different frequencies at different resolutions, known as multi-resolution analysis.

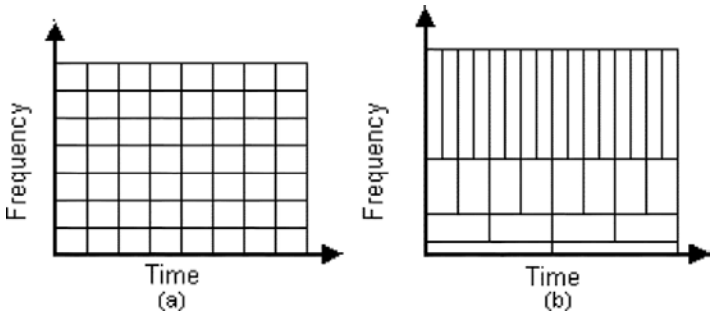


Figure 6: Where (a) is STFT and (b) is Wavelet transform. STFT uses constant resolution windows for the entire signal where WT provides good time resolution and poor frequency resolution at high frequencies and good frequency resolution and poor time resolution at low frequencies [10].

The WT is presented as follows

$$F(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} f(t)\psi^* \left(\frac{t-\tau}{s}\right) dt \tag{53}$$

where s is the scale parameter ($1/frequency$) and ψ^* is the complex conjugate of the wavelet function $\psi(t)$. The wavelet $\psi(t)$ is a small wave that will be the new basis function, as opposed to sine and cosine in the FT for example, and act as a window function. Now it is possible to change the width of the wavelet and its central frequency as it moves across the signal by changing s, this is called scaling. Expanded wavelet is better at resolving low frequency components of the signal with bad time resolution (large values of s). Shrunken wavelet is better at resolving high frequency components of the signal with good time resolution (small values of s). Therefore, all windows that are used are the dilated or compressed and shifted versions of the mother wavelet $\psi(t)$.

A practical question is which wavelet to use for a particular application. In the absence of a suitable unifying theory for wavelet behaviors, the choice of a particular wavelet for a particular problem may even appear arbitrary. For that purpose the generalized Morse wavelets are shown [11] to effectively unify many wavelet types. These wavelets are defined in the frequency domain as

$$\Psi_{\mathcal{P},\gamma}(\omega) = \int_{-\infty}^{\infty} \psi_{\mathcal{P},\gamma}(t) e^{-i\omega t} dt = U(\omega) a_{\mathcal{P},\gamma} \omega^{\frac{\mathcal{P}^2}{\gamma}} e^{-\omega^\gamma} \tag{54}$$

where $a_{\mathcal{P},\gamma}$ is used to normalize the wavelet, $U(\omega)$ denotes the unit step function, and $\mathcal{P}$ and γ are parameters that control the form of the wavelet. Reference [10] has conducted a parametric study of parameters in terms of the so-called Heisenberg area, and this work adopts the default values used by MATLAB, i.e., $\gamma = 3$ and $\mathcal{P}^2 = 60$, which yield the most symmetric Gaussian and most nearly time-frequency concentrated wavelets, and therefore hopefully ensure a satisfactory performance. For brevity, the subscripts $\mathcal{P}$ and γ will be omitted in the rest of this paper when no confusion will arise.

3.3.4 Wavelet Transform Application in Beamforming

As stated in the beginning of subsection 3.3.2, a beamforming process begins by studying the variability of a localized sound pressure field in time and frequency. In this case, bandpass type continuous wavelet transform [11] is used for that purpose, and obtain P_n as

$$P_n(\overrightarrow{x_n}, \overrightarrow{x_s}, s, t) = \lim_{\Delta t \to 0} P_n(\overrightarrow{x_n}, \overrightarrow{x_s}, s, t + \Delta t) = \lim_{\Delta t \to 0} \frac{1}{s} \int_{-\infty}^{\infty} p(\overrightarrow{x_n}, \overrightarrow{x_s}, t + \tau') \psi^* \left(\frac{\Delta t - \tau'}{s} \right) d\tau' \tag{55}$$

where $p(\overrightarrow{x_n}, \overrightarrow{x_s}, t + \tau')$ denotes the sensor samples at time $t + \tau'$, where t is usually equal to $\tau + R_n/c_0$, ψ is defined in Eq. (54), and $P_n(\overrightarrow{x_n}, \overrightarrow{x_s}, s, t)$ is the corresponding continuous wavelet transform result of the n^{th} sensor at time t. From Eq. (55), we can further have

$$P_n(\overrightarrow{x_n}, \overrightarrow{x_s}, s, t) = i\omega C_n(\overrightarrow{x_n}, \overrightarrow{x_s}, \tau) e^{i\omega\tau} \cdot q_0 \cdot \lim_{\Delta t \to 0} \frac{1}{s} \int_{-\infty}^{\infty} e^{i\omega^n \tau'} \psi^* \left(\frac{\Delta t - \tau'}{s} \right) d\tau' \tag{56}$$

where ω^n is defined in Eq. (49). And C_n is defined as

$$C_n(\overrightarrow{x_n}, \overrightarrow{x_s}, \tau) = \frac{\alpha^2(\overrightarrow{x_n}, \overrightarrow{x_s}, \tau) \cdot \beta(\overrightarrow{x_n}, \overrightarrow{x_s}, \tau)}{4\pi R_n} \tag{57}$$

The result of Eq. (56) is a function of time t and the scale parameter s (of wavelets). The latter is related to the angular frequency ω as follows,

$$\omega(s) = 2\pi \frac{f_0}{s} \cdot T_s \tag{58}$$

where f_0 is the peak frequency and T_s is the sampling time interval. Thus, $P_n(\overrightarrow{x_n}, \overrightarrow{x_s}, s, \tau + R_n/c_0)$ can be rewritten as $P_n(\overrightarrow{x_n}, \overrightarrow{x_s}, s(\omega), \tau + R_n/c_0)$. From Eqs. 55 and 56, the wavelet transform output at the frequency ω^n and time $\tau + R_n/c_0$ can be gained.

Therefore,

$$Y_n(\overrightarrow{x_n}, \overrightarrow{x_s}, \omega^n, \tau + R_n/c_0) = P_n(\overrightarrow{x_n}, \overrightarrow{x_s}, s(\omega^n), \tau + R_n/c_0) \tag{59}$$

where $Y(\overrightarrow{x_b}, \omega, \tau)$ represents the time-frequency outputs of the array. In classical delay-and-sum beamforming, this output is actually independent of the imaging gridpoint (i.e. small focus point) $\overrightarrow{x_b}$. Nevertheless, given a rotating source of the frequency ω, the n^{th} sensor will perceive a different imaging frequency ω^n owing to the Doppler effect, as shown in Eq. (49). Hence, $Y(\overrightarrow{x_b}, \omega, \tau)$ becomes to be dependent of $\overrightarrow{x_b}$ and $\overrightarrow{x_n}$. As a result, it is better to denote the output from the n^{th} sensor as $Y_n(\overrightarrow{x_n}, \overrightarrow{x_b}, \omega^n, \tau + R_n/c_0)$, where c_0 is the speed of sound and Y_n is the time-frequency result from the n^{th} sensor at $\overrightarrow{x_n}$ by applying continuous wavelet transform.

The array beamforming for high-speed moving sources of imaging frequency ω in the position $\overrightarrow{x_s}$ at an imaging time τ (from the source perspective) can be written as

$$\mathfrak{b}(\overrightarrow{x_b}, \omega, \tau) = \overrightarrow{w}^*(\overrightarrow{x_b}, \tau) A(\overrightarrow{x_b}, \omega, \tau) \overrightarrow{w}(\overrightarrow{x_b}, \tau) \tag{60}$$

where $(\cdot)^*$ denotes the complex conjugate, A is the cross spectral matrix, and the vector $\overrightarrow{w}(\overrightarrow{x_b}, \tau)$ is array weight. Theoretically, Eq. (60) acts as a spatial filter that rejects background noise and interference from other sources. The weight vector is calculated from the so-called array propagation vector and, therefore, takes the following form as

$$\overrightarrow{w}(\overrightarrow{x_b}, \tau) = \frac{C(\overrightarrow{x_b}, \tau)}{||C(\overrightarrow{x_b}, \tau)||} \tag{61}$$

where $||\cdot||$ denotes the $L2$-norm, that is,

$$||C(\overrightarrow{x_b}, \tau)|| = \left(\sum_{n=1}^{N} C_n(\overrightarrow{x_n}, \overrightarrow{x_b}, \tau) C_n^*(\overrightarrow{x_n}, \overrightarrow{x_b}, \tau) \right)^{\frac{1}{2}} \tag{62}$$

where $C_n(\overrightarrow{x_n}, \overrightarrow{x_b}, \tau)$ is the n^{th} element of the vector $C(\overrightarrow{x_b}, \tau)$, $\overrightarrow{x_n}$ is the associated location of the n^{th} sensor, and N is the number of array sensors. To further illustrate the inherent physics, consider a point type source and the corresponding array propagation vector is

$$C_n(\overrightarrow{x_n}, \overrightarrow{x_b}, \tau) = \frac{\alpha^2(\overrightarrow{x_n}, \overrightarrow{x_b}, \tau) \cdot \beta(\overrightarrow{x_n}, \overrightarrow{x_b}, \tau)}{4\pi |\overrightarrow{x_n} - \overrightarrow{x_b}|} \tag{63}$$

which is the modification of Eq. (57), where $|\overrightarrow{x_n} - \overrightarrow{x_b}|$ represents the distance between the n^{th} sensor (at $\overrightarrow{x_n}$) of the array and the b-th gridpoint (at $\overrightarrow{x_b}$) on the imaging plane. Furthermore, the cross spectral matrix, $A(\overrightarrow{x_b}, \omega, \tau)$, in Eq. (60) is defined as

$$A(\overrightarrow{x_b}, \omega, \tau) = Y(\overrightarrow{x_b}, \omega, \tau) Y^*(\overrightarrow{x_b}, \omega, \tau) \tag{64}$$

which brings us back to Eq. (59), and the substitution of the general $Y(\overrightarrow{x_b}, \omega, \tau)$, time-frequency outputs of the array, by $Y_n(\overrightarrow{x_n}, \overrightarrow{x_b}, \omega^n, \tau + |\overrightarrow{x_n} - \overrightarrow{x_b}|/c_0)$, the time-frequency result from the n^{th} sensor by applying continuous wavelet transform.

4 Numerical Simulation and Experimental Demonstration

The Wavelet-based beamforming algorithm is the same for both the numerical simulation and the experiment. Generally the algorithm can be broken down to these steps:

1. Acquire PMA data (either measured or simulated).
2. Preforming a wavelet transform on said data around a moment of interest.
3. Setting a frame of interest for potential source locations, and calculating the de-dopplerization, frequency and time correction parameters.
4. Constructing the CSM from the time-frequency wavelet transform result of all the sensors for the corrected frequency.
5. Using the correction factors to construct the correction Green's function weight vectors.
6. Image the beamforming outputs.

4.1 Numerical Simulation

As stated before, the algorithm itself does not change between the simulation and the experiment, the numerical simulation is of the generated PMA data.

In order to preform any PMA data simulation a sensor array must be simulated. By following the process of broadband array design presented in section 3.1, and taking into account the considerations stated in that section, it was decided that a multi-arm logarithmic spiral array would fit the experimental requirements and the equipment limitations. The simulated array is constructed of 63 sensors positioned on 7 different logarithmic spirals thus providing the array for the data simulation and also the model for the physical array used in the experiment.

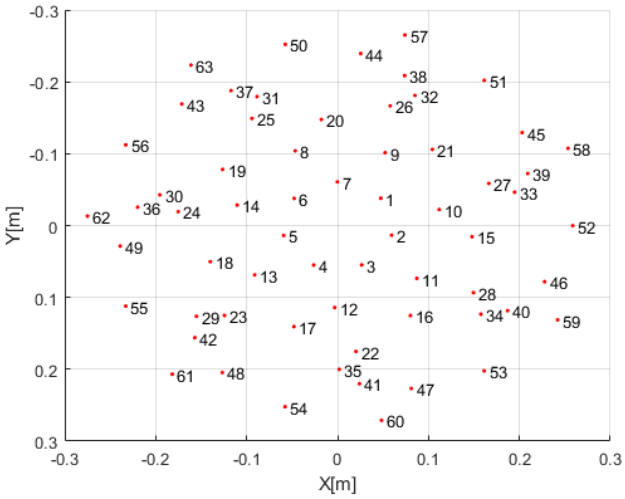


Figure 7: The microphones positions in the array from the observer's view.

In section 3.3.2 the acoustic pressure of rotating source is presented and in Eq. (47) a formula for the acoustic pressure observed at a certain location $\overrightarrow{x_n}$ in relation to a source signal transmitted from $\overrightarrow{x_s}$ is given. By defining the source's signal and location and applying this

model, a simulation of the pressure measured at the location of each and every microphone in the simulated PMA can be obtained as the data measured by that sensor.

Assuming there are two acoustic sources, each located at the tip of a propeller blade of a two blade propeller. In this simulation we will define the propeller radius to be $r = 0.2[m]$, the rotation frequency to be $f = 36[Hz]$ in a clock-wise rotation, and the signal transmission frequency to be $F = 3000[Hz]$. The propeller rotation plane will be defined as the XY plane, with the origin at the center of rotation and the Z axis directed towards the PMA plane. The planes are parallel and concentric with a distance of $z = 0.8[m]$ between them. The simulated data is of a 1[sec] of sampling time at a sampling frequency $F_s = 40[kHz]$ under standard atmospheric conditions.

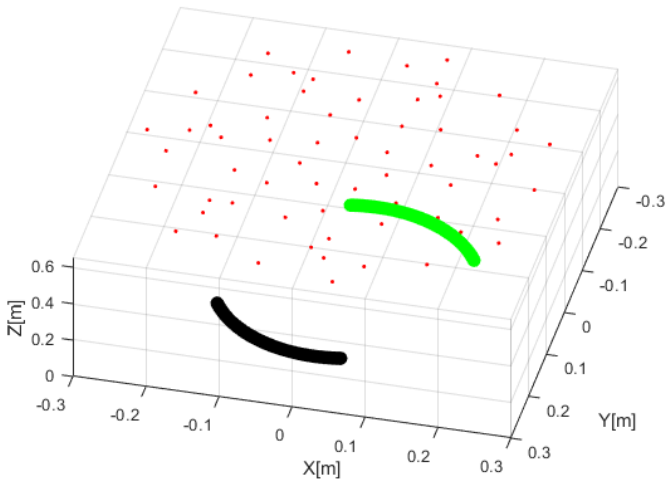
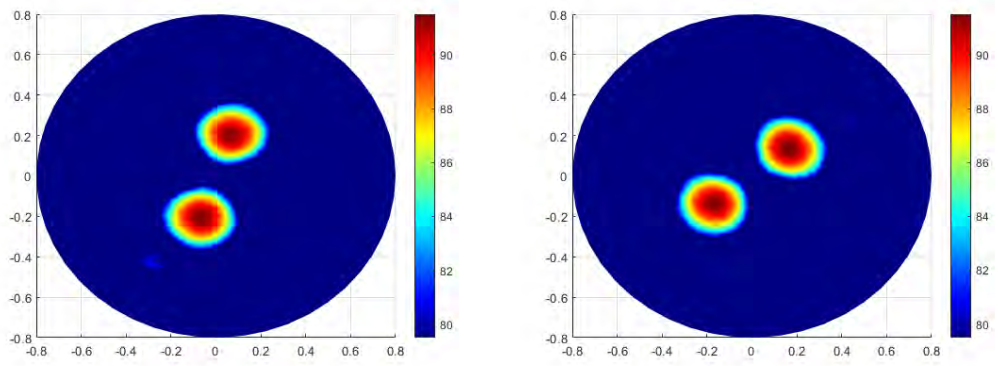


Figure 8: The PMA position relative to the propeller rotation plane.

The simulated data is then transferred to the algorithm for processing, but first, we need to define the source frequency we are looking for as well as a moment of interest in time. By doing so the algorithm will image an acoustic photo that map out all the acoustic sources present in the frame of imaging at the time defined transmitting in the stated frequency. First we will observe the acoustic images gained at different moments in time at the same frequency of interest, in order to check the algorithm capability to follow the sources location in time. In order to achieve that the frequency of interest will be defined exactly as the actual transmission frequency simulated, $F = 3000[Hz]$.

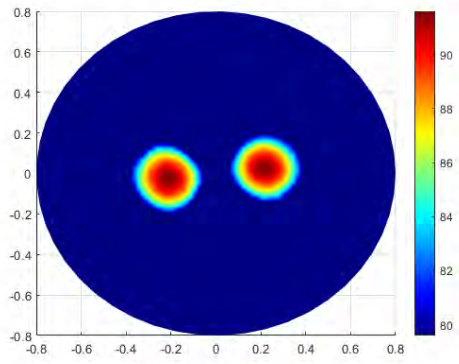
As can be seen in Fig. 9, the algorithm prints very clear images of the two sources locations. The images are clear due to the fact that the frequency of interest that we were looking for is exactly the simulated signal transmission frequency. Furthermore, the data is simulated and therefore there are no sound reflections or any other background noises. As for checking the algorithm capability to follow the sources location in time, the clock-wise rotation of the sources around the center, as well as the constant distance between them are visible.

Naturally, when sampling real data we can not be as precised and define the frequency

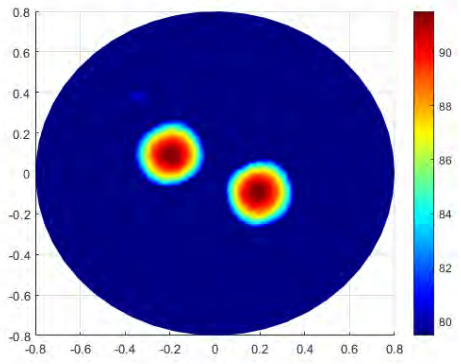


(a) t=0.3 seconds

(b) t=0.3025 seconds



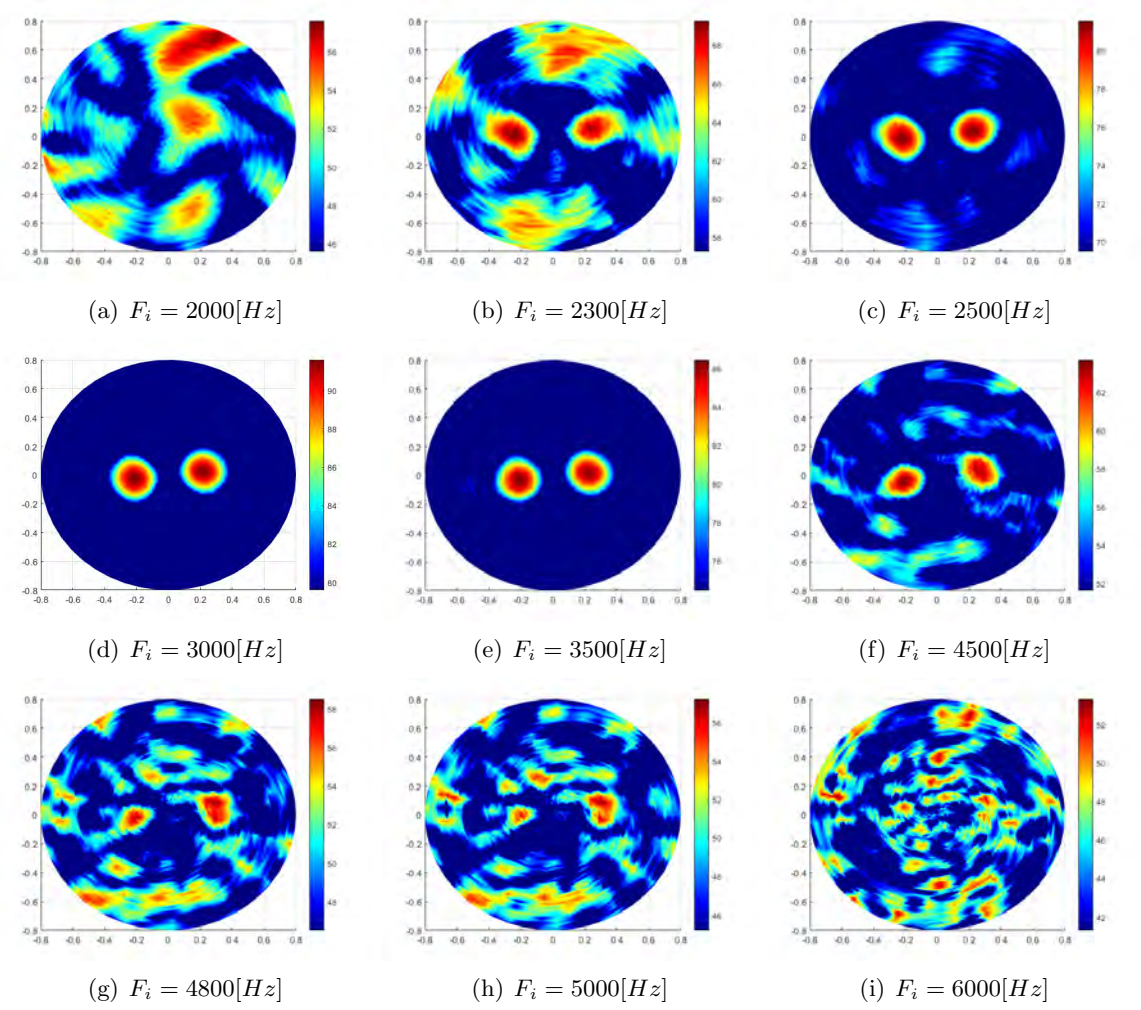
(c) t=0.305 seconds



(d) t=0.3075 seconds

Figure 9: Acoustic images of two rotating sources simulated data. The images are for the same frequency of interest $F = 3000[Hz]$, at different moments in time.

of interest exactly as the frequency of transmission. The signal might get objected to sound reflections or other background noises and other effects with minor influences that are not accounted for in the model. Usually in order to approximate the signals transmission frequency, a Fast Fourier Transform spectrum of the signals provide a decent approximation. To check the algorithm's ability to construct a decent acoustic image with only an approximated value for the transmission frequency, we will image the acoustic photo at the same moment in time, $t = 0.305 seconds$, at different frequencies of interest.



(a) $F_i = 2000[Hz]$ (b) $F_i = 2300[Hz]$ (c) $F_i = 2500[Hz]$

(d) $F_i = 3000[Hz]$ (e) $F_i = 3500[Hz]$ (f) $F_i = 4500[Hz]$

(g) $F_i = 4800[Hz]$ (h) $F_i = 5000[Hz]$ (i) $F_i = 6000[Hz]$

Figure 10: Acoustic images of two rotating sources simulated data. The images are for the same moment in time at different frequencies of interest, where image (d) shows the case when the frequency of interest is equal exactly to the signal's transmission frequency.

As we observe the images in Fig. 10, we can see that even without an accurate estimation for the sources frequency we still get somewhat accurate results. With this conclusion we understand that a Fast Fourier Transform spectrum of the signals can provide us with a proper estimation that will lead to sufficient results.

Although in images (b)-(f) the sources location is clear enough to locate their positions, by looking at the Sound Pressure Level (SPL) of the signals as shown in the color-bars it is clear that as the frequency of interest differ from the original, the SPL of the signal decreases. Even though this is a simulated data and there are no background noises or sound reflections from the environment, as we look at images for very different frequencies we can see many other "sources" appearing in the photo. As the signals are transmitted, interactions between them creates these sort of "sources" at different frequencies at different SPL. When we look

at an image for a different frequency of interest than the original we see more and more of them. Also, as stated before, at those images the SPL of the original signal in that particular frequency decrease to the part that it's relative SPL to that of the other "sources" at that frequency is not significant. This is very similar in concept to the concept described in Fig. 4 at the Time-Domain Beamforing section although now it is apparent in the frequency domain and it is not dependent on the time delay or source location. Another effect we see is that the smaller the frequency of interest taken is relative to the original, the other "sources" and noise look more like large stains and the larger the frequency of interest taken is relative to the original they look smaller and more sharp. This effect is caused by the principle of the wavelet transform. As can be seen in Fig. 6, at the Wavelet Transform section, the wavelet transform provides good time resolution and poor frequency resolution at high frequencies and good frequency resolution and poor time resolution at low frequencies. Therefore, at lower frequencies the image looks a bit more smudged, similar to a regular optic camera taking a long-exposure photo of a moving object. While at higher frequencies the positioning is very precise in time and location but includes signals of broader range of frequencies.

4.2 Experimental Demonstration

After analyzing the results from the simulation and verifying compatibility with the theoretical based expected results we proceed to an experimental demonstration. The PMA used in the experiment was created based on our design (Fig. 7) and was constructed of 63 calibrated Bruel & Kjaer type 4958 microphones. The PMA was stationed parallel to the rotation plane of a 14", two blades propeller, rotating in a counter-clock-wise motion at a frequency of $f = 40[Hz]$. The propeller rotation plane and the PMA plane are concentric and set at $z = 0.65[m]$ apart and the data is sampled at a sampling frequency of $F_s = 80[kHz]$ for $t_s = 1[sec]$.

Figure 11: The experimental set-up in the Aeroacoustics and Unsteady Aerodynamics Laboratory at the Technion – Israel Institute of Technology.

The algorithm process the data as if the acoustic noise generated from the rotor blade tips are two monopole acoustic sources. Even though this data is measured and not simulated, and therefore the noise sources may have also dipole and quadrupoles properties. For acoustic

arrays, this usually does not add much to the monopole description. Therefore, if the array is sufficiently far away, the source will be detected as if it were a monopole [12].

Similarly to the simulated experiments, the system's capabilities to accurately locate the sources position in time is the first concern that needs to be be checked. In Fig. 12 a Fast Fourier Transform spectrum of the signals is presented, from which, as stated before, a proper estimation of the sources transmission frequencies will be extracted in order to define the frequency of interest plotted by the algorithm. With the information from Fig. 12 and the defined experiment setup parameters the frequency of interest chosen is $F_i = 8000[Hz]$.

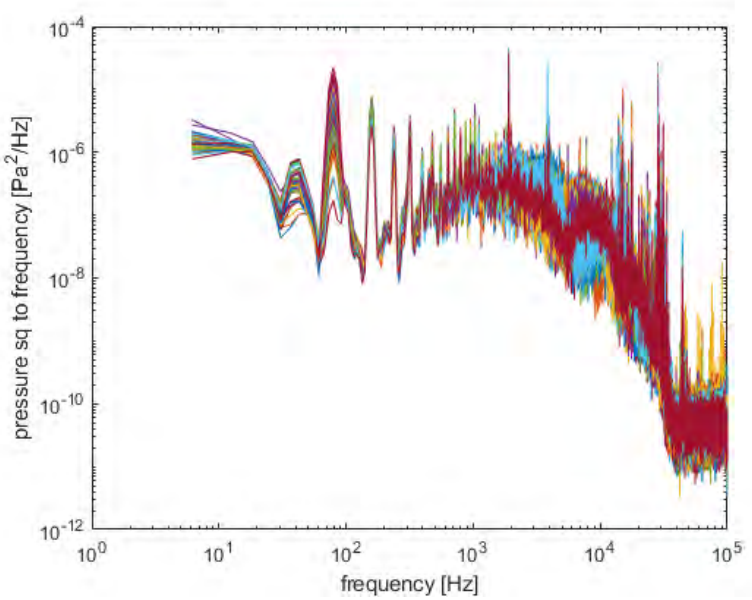
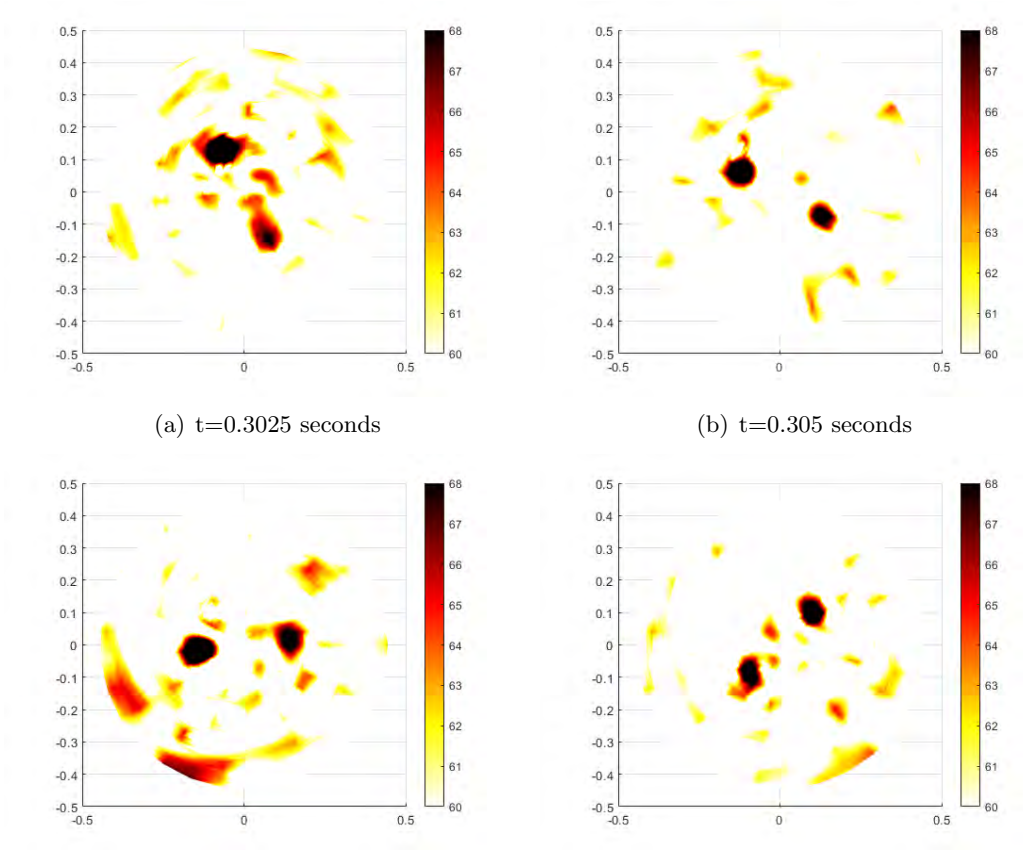


Figure 12: The power spectral density to frequency of the measured signals.

After observing Fig. 13 we see that the location of the sources is clear and the distance between them is constant and revolve the center with a counter-clock-wise motion. The system tracks the sources position in time accurately and therefore provides the essential capabilities that motivated the development of this system.

Unlike the results shown in Fig. 9, in Fig. 13 we can see a lot of background noise and the image is not as clear, also the shape and sizes of the sources images are not constant. There are multiple factors that contribute to this effect. First, the environment of the experiment was not acoustically absorbent as the experiment was executed in a regular room at the laboratory facility and not in an acoustic chamber. Therefore it is reasonable to assume that most of the noise in the images is caused by sound reflections and background noises. That is also the cause for the different shape and sizes of the dots that represent the sources location. Sound reflections from the walls cause wave interference that cause distortion to the transmitting sources signals. In addition to that, the signals de-Dopplerazation by the algorithm is based on using the inverse function of the Doppler effect model described in Eq. (42). While the Doppler effect in the simulated data is simulated and then corrected by using the same invertible function, the algorithm de-Dopplerazation of the signals can not

be as accurate with measured signals. And lastly, as explained before, the wavelet transform provides good time resolution and poor frequency resolution at high frequencies and good frequency resolution and poor time resolution at low frequencies. The transmission frequency of the sources in this case was relatively high ($F_i = 8000[Hz]$ compered to $F = 3000[Hz]$ in the simulation) and therefore the poor resolution in frequency is translated to the images being not as clear.



(a) t=0.3025 seconds

(b) t=0.305 seconds

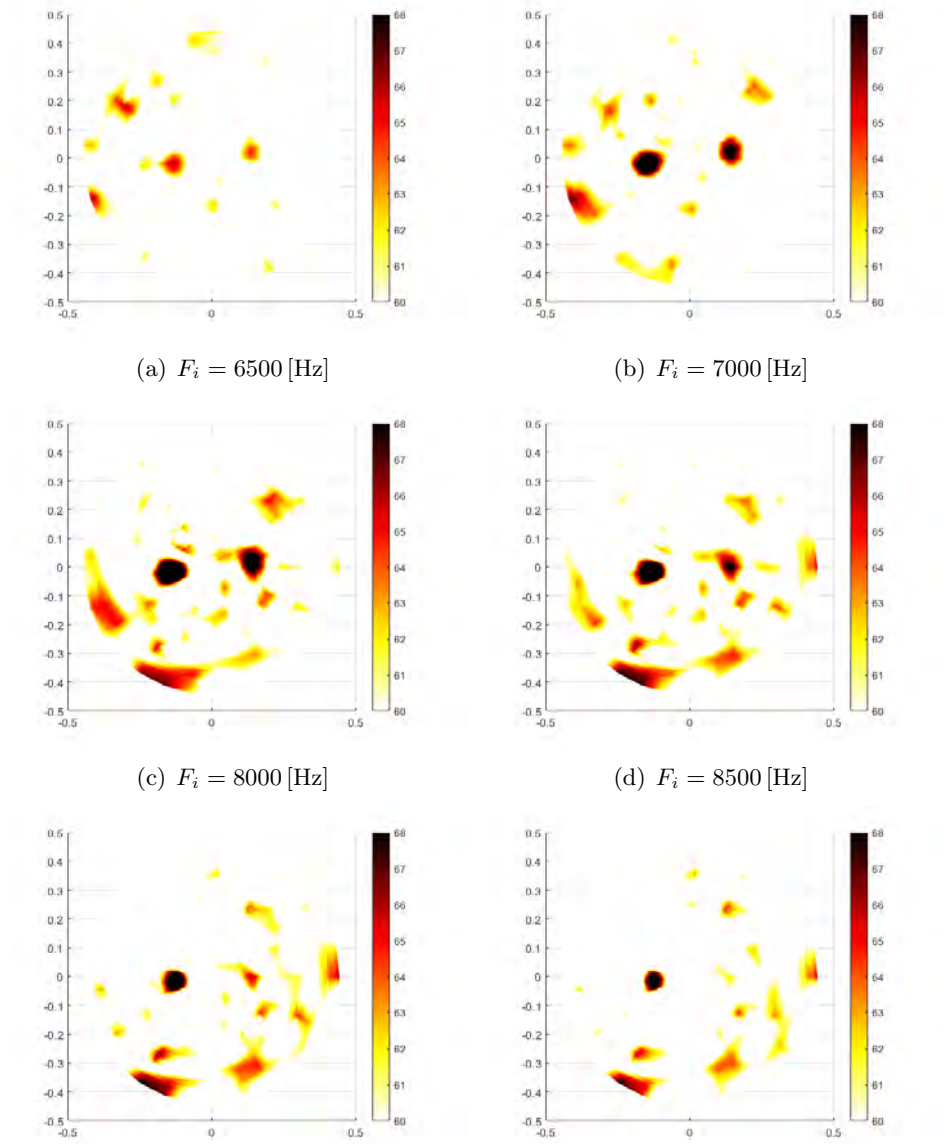
(c) t=0.3075 seconds

(d) t=0.31 seconds

Figure 13: Acoustic images of two rotating sources measured data. The images are for the same frequency of interest $F = 8000[Hz]$, at different moments in time.

By observing Fig. 10 we already came to the conclusion that even without an accurate estimation for the sources frequency we still get somewhat accurate results, but due to the poor frequency resolution in higher frequencies we would like to see if the transmission frequency estimation margin is still as forgiving at a higher frequency and also for measured data. In order to do so, in Fig. 14 and 15 we'll observe images describing both measured and simulated data of two sources transmitting at $F = 8000[Hz]$ at the same moment in time imaged by the algorithm for different frequencies of interest. The simulated data in this case is is simulated for the same conditions and parameters as the measured experiment data. Unlike the images in Fig. 10, where the scale and limits of the Sound Pressure Level (SPL)

shown by the color-bar were free to change according to the output values of the simulation, in Fig. 15 the SPL scale is defined between $60dB$ to $68dB$. That is because the measured sources SPL falls in this margin as can be seen in Fig. 13. Due to the SPL limits defined for these images we do not expect to see a lot of noise and other "sources" appearing at the images of different frequencies of interest then the original transmission frequency, unlike the images in Fig. 10.

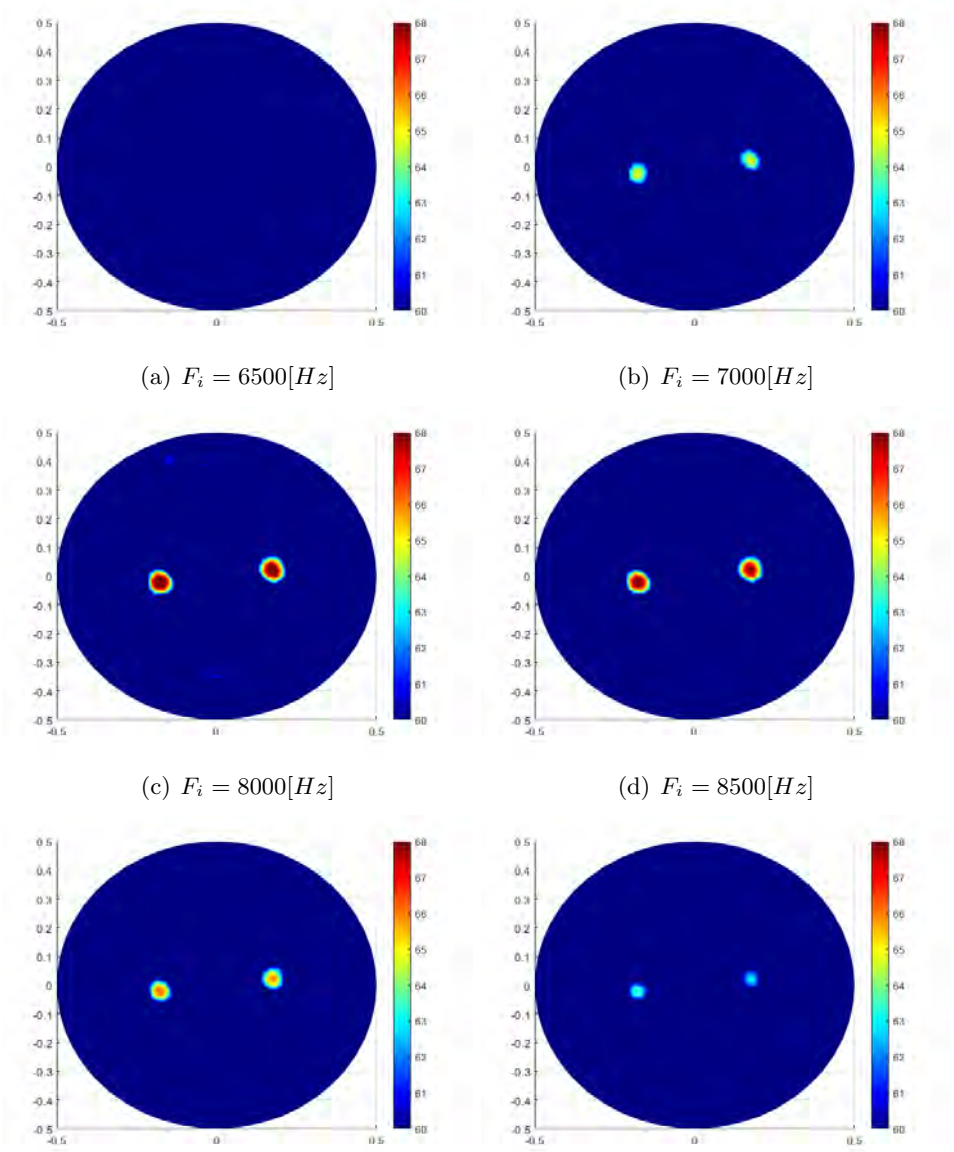


(a) $F_i = 6500\,[\text{Hz}]$ (b) $F_i = 7000\,[\text{Hz}]$

(c) $F_i = 8000\,[\text{Hz}]$ (d) $F_i = 8500\,[\text{Hz}]$

(e) $F_i = 9000\,[\text{Hz}]$ (f) $F_i = 9500\,[\text{Hz}]$

Figure 14: Acoustic images of two rotating sources measured data. The images are for the same moment in time at different frequencies of interest. Where image (c) shows the case when the frequency of interest is equal to the signal's estimated transmission frequency.



(a) $F_i = 6500[Hz]$ (b) $F_i = 7000[Hz]$

(c) $F_i = 8000[Hz]$ (d) $F_i = 8500[Hz]$

(e) $F_i = 9000[Hz]$ (f) $F_i = 9500[Hz]$

Figure 15: Acoustic images of two rotating sources simulated data. The images are for the same moment in time at different frequencies of interest. Where image (c) shows the case when the frequency of interest is equal to the signal's estimated transmission frequency.

As expected, even when dealing with sources that transmit signals at higher frequencies accurate positioning of the sources location is still possible even without an accurate estimation of the transmission frequency. Although after comparing the measured data to the simulated one, the influence of the sound reflections on the results is not neglectable. Therefore, even though an accurate estimation is not necessary in order to locate the sources position in time, it is still impotent to estimate with caution.

4.3 Method Modifications and Future Work Suggestions

In the process of developing the system we came across a few major issues that influence the systems capabilities. The system required a defined input for several parameters and took a long computing time to produce the acoustic images. Furthermore, due to the wavelet transform resolution trade-offs and the wavelet transform MATLAB function predefined properties the resolution and accuracy of the system suffered. Ideally we would want an acoustic camera that can provide accurate acoustic images with high resolution in both time and frequency in a short computation time, for a wide range of scenarios.

One of the modifications we have made in order to improve the system is the ability to extract the sources rotation movement frequency from the data, therefore it is capable of producing an accurate image even when little is known about the sources movement. Additional improvement such as this may be done in the future for estimating the vertical distance between the PMA and the sources rotation plane. Those improvements will allow easier testing of rotter noises and even fly-over experiments without many preparations and experimental parameters definitions.

Most of the computation time in the data processing is the wavelet transformation of all the signals from all the microphones. In order to shorten the computation time we have decided to only transform a small section of the data of each microphone. When an acoustic image is produced it describes a single moment in time. Therefore by only transforming a small segment of the signal around that moment of interest in time, for which the image is produced, provides all the data necessary for the process. It is important to keep in mind that there is a time delay in the measurements between different microphones, so the segment of data should be long enough to ensure that the signal transmitted from the sources at the moment of interest is received at all sensors.

Lastly, in order to improve the resolution capabilities of the wavelet transform, instead of dividing the entire frequencies spectrum to time dependent segments we apply the wavelet transform for a range of frequencies around the frequency of interest. As explained before and as can be seen in Fig. 6, the wavelet transform creates windows for a broad range of frequencies with high frequency resolution for low frequencies. By applying the transform for only a segment of the frequencies around the frequency of interest the resolution in frequency increases relative to a transform without frequency limits. Furthermore, we also added an interpolation to increase the resolution in time and frequency so the resolution would be even higher. Although the additional interpolations do cost us in computation.

In future work some more experiments and simulations may provide an even better understanding of the minimal limits and conditions that would still allow the system to produce the images properly, even faster and more accurate then the current capabilities. Where our ultimate goal would be an acoustic camera capable of producing acoustic images at close to instantaneous speed with high accuracy that requires near to no parameters predefined. With such a device, simply pointing the PMA towards a source of noise could provide a live acoustic visualization.

5 Summery

During this research we designed a PMA fitted to our aeroacustic applications and implemented the wavelet transform based beamforming method in order to produce an acoustic camera system capable of tracking high-speed rotating acoustic sources. Our PMA can is used as an acoustic camera with a relatively large aperture and it has the ability to suppress sidelobes over a broad range of frequencies. Then we applied and modified the wavelet beamforming method presented by W. Chen and X. Huang [2] into our beamforming algorithm.

By observing our results from the simulations and experiments we saw that we were able to fulfill our objective and develop an acoustic camera that is capable of tracking and producing accurate acoustic images of high-speed rotating acoustic sources. Furthermore, our camera is capable to produce those images accurately enough even under the effects of sound reflections and background noises when operating outside from an acoustic chamber environment. It does not require the sources rotational frequency to be predefined and is able to locate the sources properly even if the frequency of interest is estimated in a rather forgiving margin around the actual sources transmission frequency. All of that in a relatively short computing time, due to the algorithm modifications.

References

[1] P. Sijtsma, S. Oerlemans, and H. Holthusen, "Location of rotating sources by phased array measurements," in *7th AIAA/CEAS Aeroacoustics Conference and Exhibit*, 2001, p. 2167.

[2] W. Chen and X. Huang, "Wavelet-based beamforming for high-speed rotating acoustic source," *IEEE Access*, vol. 6, pp. 10 231–10 239, 2018.

[3] C. S. Allen, W. K. Blake, R. P. Dougherty, D. Lynch, P. T. Soderman, and J. R. Underbrink, *Aeroacoustic measurements*. Springer Science & Business Media, 2002.

[4] H. C. Pumphrey, "Design of sparse arrays in one, two, and three dimensions," *The Journal of the Acoustical Society of America*, vol. 93, no. 3, pp. 1620–1628, 1993.

[5] J. R. Underbrink, "Circularly symmetric, zero redundancy, planar array having broad frequency range applications," Mar. 20 2001, uS Patent 6,205,224.

[6] R. Dougherty and R. Stoker, "Sidelobe suppression for phased array aeroacoustic measurements," in *4th AIAA/CEAS aeroacoustics conference*, 1998, p. 2242.

[7] H. E. Camargo, "A frequency domain beamforming method to locate moving sound sources," Ph.D. dissertation, Virginia Tech, 2010.

[8] D. Król, A. Lorenc, and R. Świceiński, "Detecting laterality and nasality in speech with the use of a multi-channel recorder," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2015, pp. 5147–5151.

[9] X. Huang, L. Bai, I. Vinogradov, and E. Peers, "Adaptive beamforming for array signal processing in aeroacoustic measurements," *The Journal of the Acoustical Society of America*, vol. 131, no. 3, pp. 2152–2161, 2012.

[10] R. Ganesan, T. K. Das, A. K. Sikder, and A. Kumar, "Wavelet-based identification of delamination defect in cmp (cu-low k) using nonstationary acoustic emission signal," *IEEE Transactions on Semiconductor Manufacturing*, vol. 16, no. 4, pp. 677–685, 2003.

[11] J. M. Lilly and S. C. Olhede, "Generalized morse wavelets as a superfamily of analytic wavelets," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 6036–6041, 2012.

[12] P. Sijtsma, "Experimental techniques for identification and characterisation of noise sources," 2004.