Tracking drones with small microphone arrays

MSc thesis Barry Blekemolen

To be defended publically on 22-05-2025

Project supervisors:

Prof. dr. ir. M. Snellen
dr. ir. S. Luesutthiviboon
ir. A. Altena

Special thanks to:

A.M. Morin - DJI Mavic pilot
S. Luesutthiviboon - Parrot Bebop pilot

Abstract

This report investigates how accurately drones can be tracked by measuring the noise they emit over time with a small microphone array; the ReSpeaker Mic Array v2.0. This acoustic localization is achieved by applying beamforming to select frequency ranges. In literature this localization has been achieved and accurately so, but the microphone arrays there are generally at least a meter in diameter with many microphones. The array used in this project is about 5 cm in diameter and contains only four microphones. It is therefore cheaper in cost and in terms of calculation time.

The first experiment takes place in an anechoic chamber to compare the beamform output with reference values, specifying the expected elevation and azimuth angles. The output turns out to be fairly accurate, with an error of a couple of degrees. A second experiment is with indoor drone flights, where the reference location is only broadly known. The beamform output at high frequencies appears to be very accurate and the output at low frequencies is less accurate. This is expected due to the arrays small size which hinders beamforming at low frequencies but it is unfortunate as drones emit most of their noise at lower frequencies.

The third experiment is with outdoor drone flights in the Unmanned Valley near Katwijk aan Zee. A 500 gram heavy Parrot Bebop 2 drone with three-bladed propellers can be localized up to a distance of 100m and a 907 gram heavy DJI Mavic 2 drone with two-bladed propellers can be localized up to a distance of 80m. Both with an error between GPS and beamform output of less than 10 degrees if beamforming is applied to a higher frequency range of 2-5kHz. This only applies if the array is placed at a 45-degree angle with the ground, which improves the beamforming accuracy. The application of functional beamforming did not improve the results of this experiment.

Nomenclature

Units

dB - deciBel, logarithmic ratio used to express magnitude
Hz - Hertz, frequency of sound
m - meters, spatial distance
rpm - revolutions per minute
s - seconds, measure of time

Abbreviations

BPF - Blade Passing Frequency
DOA - Direction Of Arrival
i.e. - id est
IC - Integrated Circuit
int - interference
LED - Light Emitting Diode
LIDAR - LIght Detection And Ranging
MEMS - MicroElectroMechanical Systems
OSPL - Overall Sound Pressure Level
PALILA - PsychoAcoustic LIstening LAboratory
PSF - Point Spread Function
SPL - Sound Pressure Level
USB - Universal Serial Bus

Equation variables

B - Beamform output, complex value
c - Speed of sound in m/s
d - Distance between microphones in meter
f - Frequency in Hz
f_0 - Fundamental frequency, rotation speed of the rotor of an electric motor
FB - Functional beamform output, complex value
f_k - Frequencies considered for beamforming with index k
g - Steering vector, the phase shift needed to align signals for beamforming
g(b) - Steering vector to one specific point, used for a point spread function
L - Aperture in meter, measure of spatial microphone array size
R - Range in meter
R_{gps} - Range in meter, distance between a drone and an origin point near the array
r_j - Scan point range in meter with index j
r_n - Microphone range in meter with index n
$r_{n,j}$ - Range vector containing distance between a scan point and every microphone in meter
U - Eigenvectors obtained after spectral decomposition
X - Vector containing complex values for one frequency and all microphones
xEast - Distance between a GPS transponder and an origin point in easterly direction in meter
x_j - Scan point location in x direction in meter
x_n - Microphone location in x direction in meter
y_j - Scan point location in y direction in meter
y_n - Microphone location in y direction in meter
yNorth - Distance between a GPS transponder and an origin point in northerly direction in meter
z_j - Scan point location in z direction in meter
z_n - Microphone location in z direction in meter

zUp - Distance between a GPS transponder and an origin point in the vertical direction in meter

Greek letters

θ - Elevation angle in degrees, spherical coordinate
θ_B - Spatial beamforming beamwidth, in radians
θ_{gps} - Elevation angle in degrees of the reference GPS signal
λ - Wavelength of a sound wave in meters
ν - Functional beamforming exponent
Σ - Diagonal matrix containing eigenvalues after spectral decomposition
ϕ - Azimuth angle in degrees, spherical coordinate
ϕ_{gps} - Azimuth angle in degrees of the reference GPS signal
ξ_j - Spatial beamforming scan point with index j

List of Figures

List of Tables

Contents

CHAPTER 1

Introduction

Unmanned aerial vehicles have been around since the beginning of the 20th century. Development continued mostly for military purposes to advance national security. With technological advancement and reductions in cost and complexity drones finally became available and affordable for the general public in the 2010s. Widespread use however comes with the drawback that some operators use their drones in an unsafe area or try to record activities they should not, resulting in demands for regulation. Some commercial and national operations also led to security concerns [1]. Eventually regulation was made to reduce these concerns. This regulation banned drones from many areas, required minimum separation from objects and required operators to register themselves [2]. To enforce such regulations, it is necessary to have ways to detect violations.

The propulsion system of most drones consists of propellers. These produce noise at the blade-passing frequency, in addition to broadband noise. Therefore, one possible way to detect these drones is by their specific emitted sound profile. This noise can be measured by microphones. With an array of several microphones the position of the drone can be found using an acoustic imaging technique such as beamforming.

There is previous research into this matter using various acoustic arrays to assess their ability to localize drones [3-10]. The arrays used in these projects were effective in reaching their goal. However, these custom arrays were large, sometimes several meters in diameter, using a combined arrangement of closely spaced microphones and distantly placed microphones. These choices lead to a good resolution at lower frequencies due to the large diameter, aperture, according to the Rayleigh criterion [11, 12]. As well as a good resolution at higher frequencies due to the large number of closely spaced microphones. The closely spaced elements reduce the side-to-main lobe level, which is an issue for higher frequencies. These arrangements are seemingly optimal, however employing this on a larger scale for illegal drone detection would be expensive. Expensive in cost because of the large number of components and expensive in computational power because of the amount of data produced by the many microphones. Therefore, there is an opportunity for a more cost-effective method using small, portable, low-cost arrays, if the performance is acceptable. One such array has been identified as potentially viable [16] and will be the array used for this research. Drawbacks of this array are the low aperture and fixed square microphone placement. In addition to testing the accuracy of this array, the effect of a different array orientation and the application of an advanced beamforming technique are evaluated on their effect regarding localization accuracy [5, 11-15].

1.1 Research questions

There are three main research questions.

- Can a single small low-cost microphone array detect and localize a passing small drone?
  - What is the localization range for one small array for noise from a quadcopter drone with a weight lower than 1kg?
- What is the localization accuracy when using an array placed in a different orientation?
  - What array orientation provides the best localization accuracy?
- Does the application of advanced beamforming techniques improve the localization accuracy of the system?

1.2 Thesis setup

The second and third chapters set up a theoretical framework that serves as the basis for the later chapters. The fourth and fifth chapters apply these elements to measurements using the small array obtained at two different indoor experiments. The sixth chapter describes the setup and outcome of an outdoor experiment. The final chapter relates the results of the outdoor experiment to the research questions.

Chapter 2

Measurement and visualization of drone noise

The purpose of this thesis project is to evaluate a small microphone array in its ability to track the location of a drone by its emitted noise and possible methods of enhancement that can improve this ability. A first step towards this objective is to know the characteristics of the noise produced by drones, these are described in section 2.1. The measured noise then needs to be visualized to convert the data into information. The spectrogram plot in section 2.2 gives a clear picture of the relation between time, frequency and sound level. Due to the characteristics of drone noise, windowing will need to be applied to obtain accurate results. This process is described in section 2.3. The details of the small microphone array used in this project are given in section 2.4.

2.1 Drone noise

Noise is a term used for undesirable sound waves. Sound waves are small variations in pressure and are usually expressed in terms of frequency in Hz and magnitude in decibels (dB). For drones these pressure variations are usually created by rotating propellers [21]. Because these actions take place due to a propeller blade rotating at a certain frequency, the sound produced is also at this frequency and at its higher harmonics. It is known as the blade passing frequency (BPF), in Hz. As specified in equation 2.1, it is the product of the number of propeller blades (N) and the spin frequency of the rotor of the electric motor (f_0). Multiplying the rotor frequency by 60 expresses it in the more conventional unit rpm, revolutions per minute.

$$\text{BPF} = \frac{\text{rpm} * \text{N}}{60} = f_0 * \text{N} \tag{2.1}$$

Figure 2.1 shows a magnitude versus frequency plot of an operational DJI Phantom IV drone from an experiment by Blanchard et al. [3]. The motors were running with an rpm of 7742, resulting in $f_0 = 129\text{Hz}\ (7742/60)$ and because N = 2, the BPF = 258Hz. The experiment shows that propellers are the only relevant noise source for this drone.

The blue line shows multiple peaks. These are the harmonics of the BPF. The first four at 258Hz, 516Hz, 774Hz, and 1032Hz are the highest. The small peaks in between are odd multiples of f_0. Tonal noise is especially noticeable because the sound energy is highly concentrated. The average appears to be about 40dB lower than the peaks.

2.2 Spectrogram

Sound can be visualized with a spectrum plot as in figure 2.1. The spectrum is obtained by converting the measured pressure changes by the microphones over time to the frequency domain by means of the Fourier transformation. It provides information on the frequency content of the signal by specifying the magnitude per frequency bin. In contrast, a pressure versus time plot only gives information on the signal strength over time. This is insufficient information because the main contribution could come from background noise.

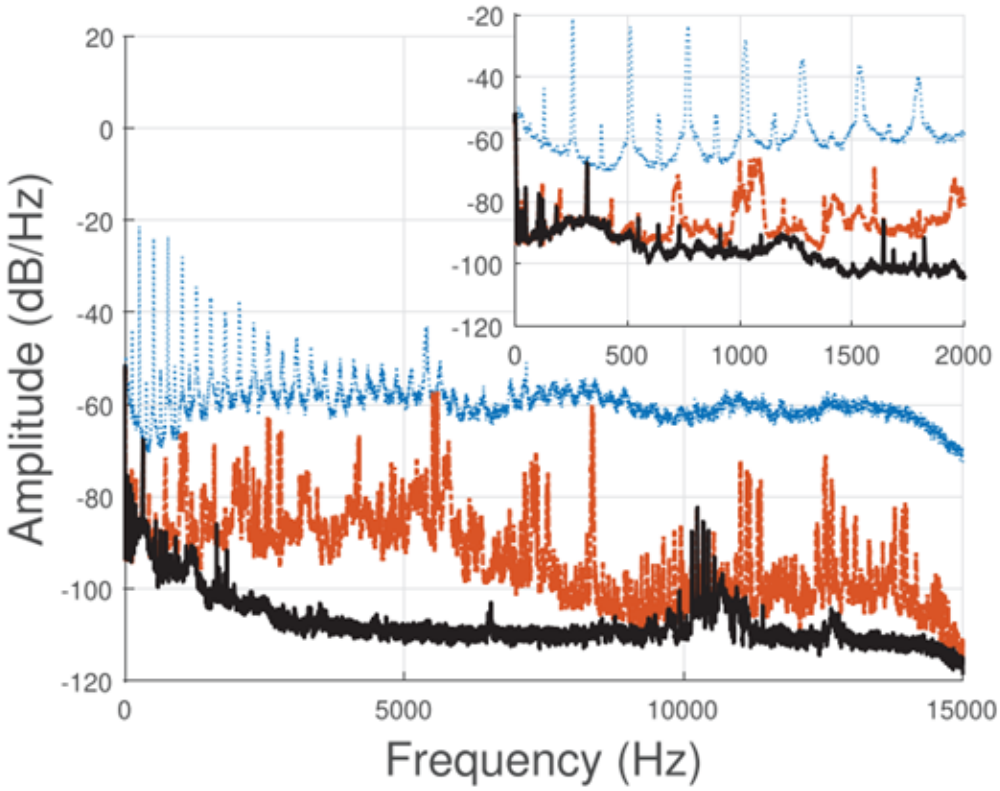


Figure 2.1: Noise spectrum of the DJI Phantom IV, Blanchard et al. Fig 4 [3]. The blue line indicates the drone noise with propellers, the orange line without propellers and the black line is the background noise in the anechoic chamber the experiment was performed in.

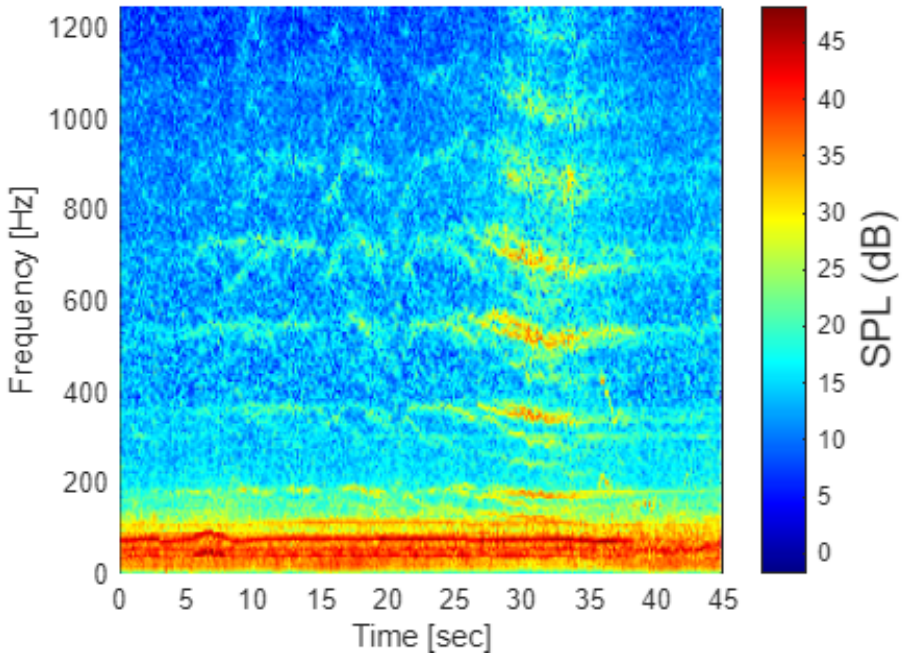


Figure 2.2: Spectrogram DJI Phantom 3 flyby

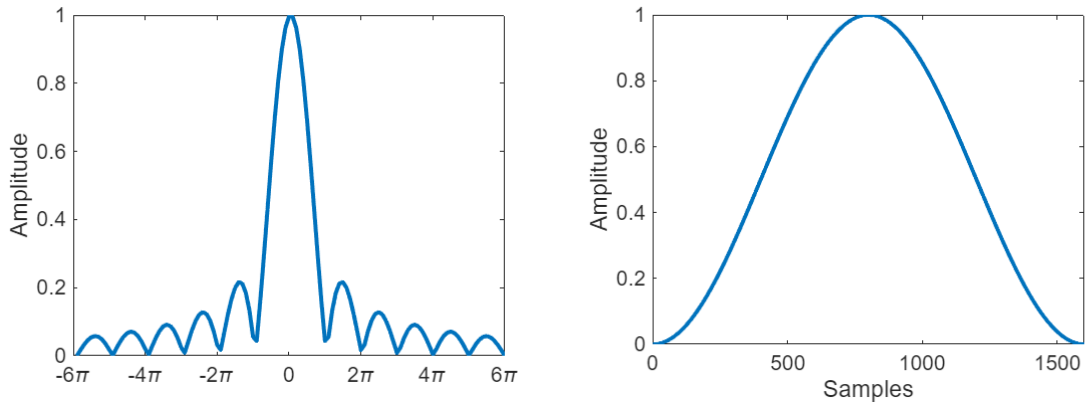
However, it is useful to see how the frequency content changes over time. A clear way to demonstrate this is with the spectrogram. Figure 2.2 shows an example. This plot uses the two spatial axes for frequency and time and it also has colors to indicate the magnitude. The colored bar to the right indicates the Sound Pressure Level (SPL) associated with each color. It is the result of stitching a lot of spectrum plots like figure 2.1 together. Every column could be a separate spectrum plot. If a spectrum is calculated for every 0.1 seconds this plot would be the result of 450 sequential spectra. This plot again shows the harmonics of drone noise, which are louder when it is closer to the array. Most notably between 30 and 35 seconds. Also visible are the high values associated with background noise in the low-frequency range. The value of this plot is that it shows which frequencies should be considered for the beamform loop used for the localization.

2.3 Windowing

The acoustic localization in this project requires taking snapshots in time and localizing the source of noise at that moment in time. As the source is a moving drone, short snapshots of 0.1 seconds are needed for an accurate result. These snapshots are essentially windows; only a part of the whole signal is being looked at. Because the edge of the window abruptly cuts off the signal, it is generally non-zero at the edges. This discontinuity leads to sinc functions in the frequency domain, see figure 2.3a, instead of a series of impulses. In addition, sinc functions are also produced by the discrete Fourier transform. These are an issue because all the smaller sidelobes beside the main lobe in the middle in figure 2.3a are spectral leakage. It is energy incorrectly distributed to other frequencies. The true magnitude at other frequencies can be completely overshadowed by sidelobes of a larger peak at another frequency. This especially applies when the signal has high peaks at certain frequencies, such as drone noise at the BPF harmonics.

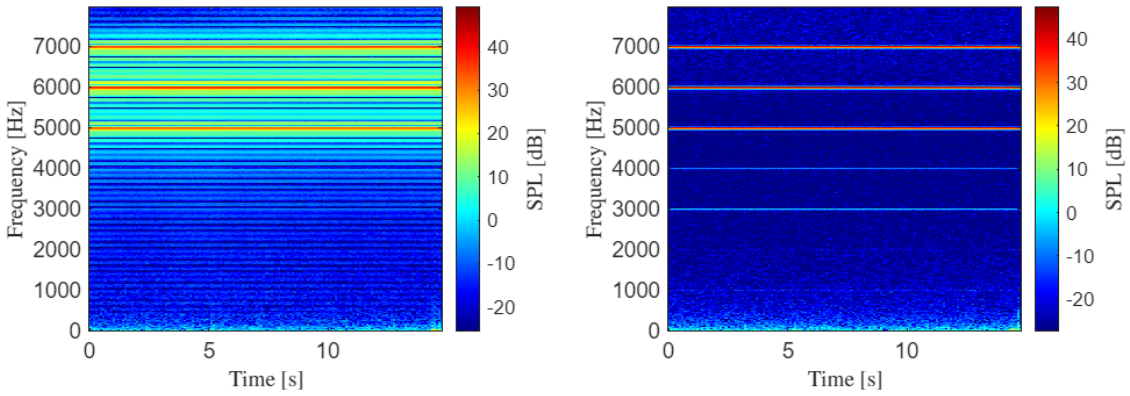
These undesirable effects can be reduced by applying a windowing function that smooths the edges of the signal to zero, removing the discontinuity. The Hann function in figure 2.3b is an example of such a function. It is applied with element-by-element multiplication with the data points of the snapshot in the time domain.

Figure 2.4a shows a spectrogram without windowing and figure 2.4b a spectrogram with windowing. The only differences are the application of the Hann function and the division with the



(a) Sinc function (b) Hann function

Figure 2.3: Spectral leakage and the function to resolve it



(a) Spectrogram without windowing (b) Spectrogram with windowing

Figure 2.4: Effect of a windowing function

root mean square value of the Hann function to compensate for the loss of energy. The windowed spectrogram shows a result closer to reality with a source that emits tones at 5kHz, 6kHz and 7kHz.

2.4 ReSpeaker Mic Array v2.0

The measurements in this project are done with the ReSpeaker Mic Array v2.0 [18]. The array is pictured in figure 2.5. The array in question has four digital microphones of the type MP34DT01TR-M, labeled (2) in the figure, which are spaced 45.7mm from each other. These microelectromechanical systems (MEMS) send their data to the integrated circuit (IC) in the middle, an XMOS XVF-3000 (1). This block sends the output to the 3.5mm (5) and USB-B (4) connections at the bottom. It also has an internal beamformer that identifies the direction of arrival (DOA) angle. Those angle values are sent to the USB and to a ring of LEDs (3) for a visual indication of where the sound comes from. The values sent to the USB can be read by running a Python script on the connected device, such as a laptop. The sampling frequency of the array is 16kHz, which means that sound up to 8000Hz is measured according to the Nyquist criterion. The IC with label (6) is a stereo codec with speaker drivers. The signal from (1) passes through here before going to (5). This port was not used for this project.

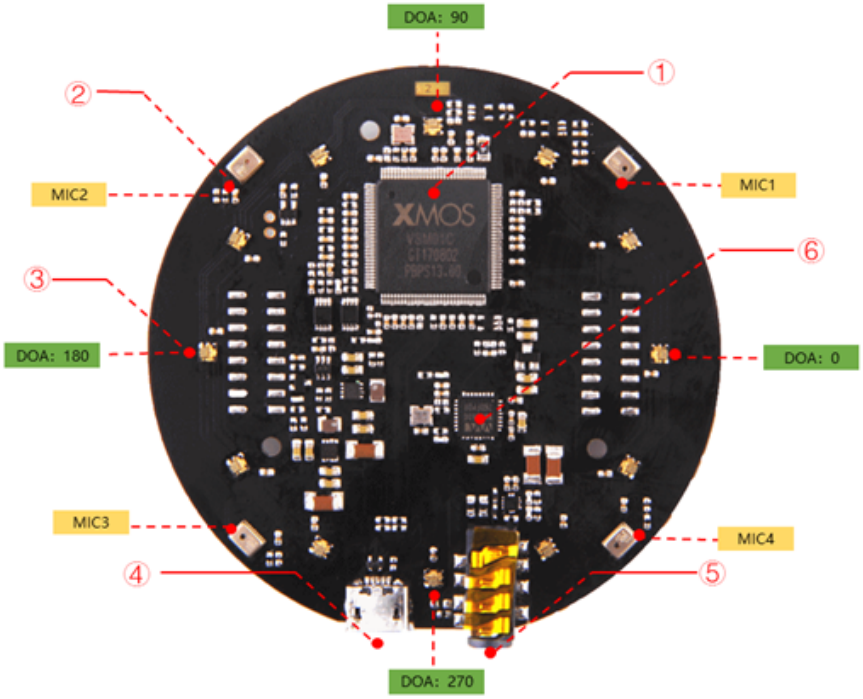


Figure 2.5: ReSpeaker Mic Array v2.0

Chapter 3

Acoustic localization using beamforming

Acoustic localization of drones is done using beamforming. In this project beamforming is done over certain frequency ranges of interest, such as the frequencies associated with the relatively loud low harmonics of the BPF. The ranges are obtained by visual inspection of a spectrogram to ensure the correct frequency bandwidth is selected. The BPF can change over time due to Doppler shift as pictured in figure 2.2 around $t = 32s$, or due to changes in motor rpm if the drone maneuvers.

Beamform plots are created to aid the localization. These are plots that share some similarities with the spectrograms in figures 2.2 and 2.4. Both spectrograms and beamform plots depict noise levels and have a colored SPL bar. But the axes of a (two dimensional) beamform plot are two spatial dimensions instead of frequency and time. It shows the measured SPL at every spatial location in the area to inform the viewer about the physical location of the noise source(s).

The first step towards understanding beamform plots is with the general explanation of the theory behind them in section 3.1. Because the interest is in the location relative to the array, the coordinate system is changed from two Cartesian spatial dimensions to spherical coordinates. A description of the change towards horizontal azimuth angles ϕ and vertical elevation angles θ is given in section 3.2. Beamforming in spherical coordinates is then applied to reference drone noise in section 3.3, showing two example beamform plots. These plots are more detailed than later beamform plots in this report. Why this is the case is explained in section 3.4. The last part in section 3.5 describes an advanced beamforming method which may improve the localization accuracy.

3.1 Theory of beamforming

The process of beamforming is shown in figure 3.1. [1] The left side of the image shows wavefronts arriving at yellow circles that represent microphones. The sound is depicted as wavefronts because sound is a longitudinal pressure wave, with the pressure varying slightly above and below the ambient air pressure. This can be picked up by a microphone, whose output is visualized with a red transverse wave in figure 3.1. These red waves can be shifted in time to align them, a step called beamsteering. The top half of figure 3.1 shows an effective application of this process. The target wavefront arrives at the top microphone first, but the signal is delayed by t_2 seconds to match it with the third microphone which will receive the signal coming from this angle t_2 seconds later. The same process is applied to the second microphone, but with t_1 seconds. The result of steering all three microphones to the same angle is signal alignment. Summing them up leads to a stronger signal. This does not mean that the microphones only pick up sound from the target angle. As the bottom half of figure 3.1 shows, there is noise or interference coming from other angles. However, because all data received by the microphones is shifted by either t_2, t_1 or unshifted, only signals coming from the target angle are aligned and constructive. The interference wavefront in the bottom half of figure 3.1 arrives at the top microphone last, but it is still delayed by t_2 seconds, potentially misaligning it further. The lower microphone receives the interference as first, but it is not delayed and will pass through first. Signals coming from this angle are not enhanced.

[1] Figure 3.1 source: http://www.labbookpages.co.uk/audio/beamforming/delaySum.html

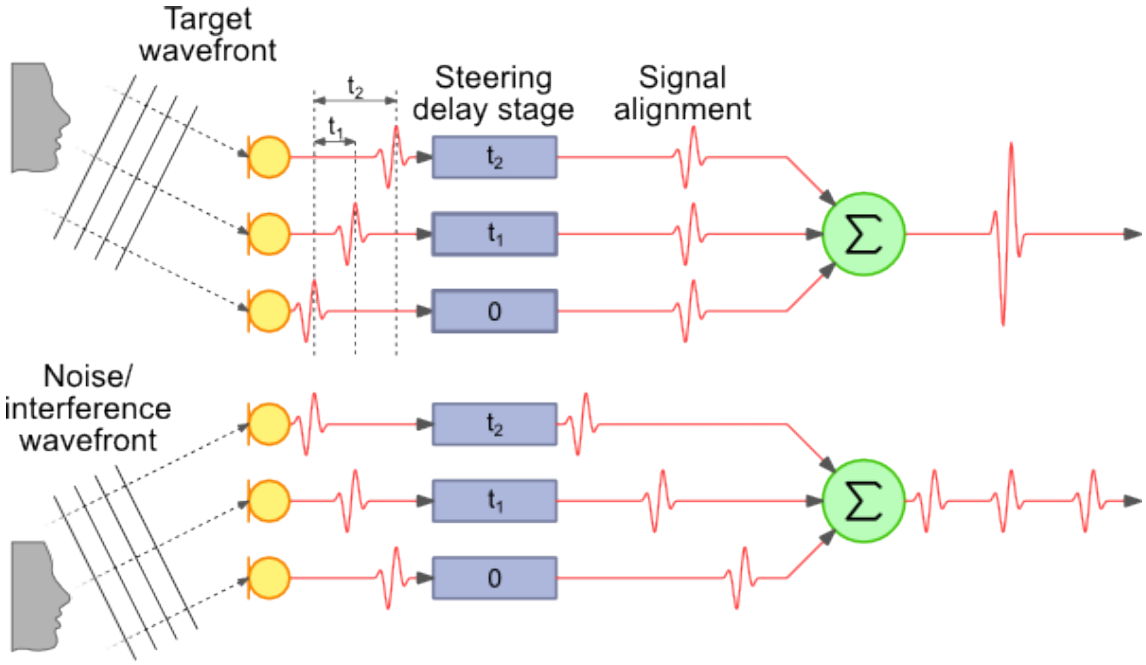


Figure 3.1: Delay and sum beamforming

The beamforming applied in this project is done in more than the one dimension pictured in figure 3.1. It is also done in the frequency domain to facilitate filtering for certain frequency ranges. This means the application of phase shifts rather than time delays, but the end result is the same. This project uses equations 3.1, 3.2 and 3.3 to obtain the beamform plots needed to locate the drone flying by.

$$\mathbf{r_{n,j}} = \sqrt{(x_j - x_n)^2 + (y_j - y_n)^2 + (z_j - z_n)^2} \tag{3.1}$$

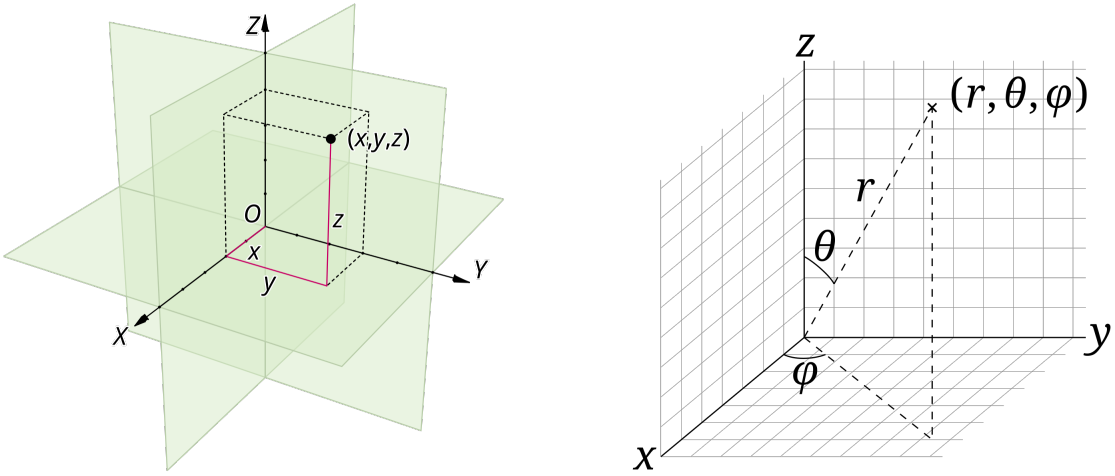
$$\mathbf{g}(\xi_j, f_k) = e^{-2\pi i f_k \frac{\mathbf{r_{n,j}}}{c}} \tag{3.2}$$

$$\mathrm{B}(\xi_j, f_k) = \frac{\mathbf{g}^*(\mathbf{X}\mathbf{X}^*)\mathbf{g}}{||\mathbf{g^4}||} \tag{3.3}$$

A scan grid is defined of elements x_j, y_j and z_j containing the value of the spatial distance between each scan point and an origin point in three dimensions. The locations of the microphones in the same area, relative to the same origin, are given by x_n, y_n and z_n. The whole beamforming process is done point by point throughout the area, all three equations are applied to a single scan point before moving on to the next point. As such, equation 3.1 produces a vector with its size defined by the number of microphones. The elements in this range vector $(\mathbf{r_{n,j}})$ are the distances to the scan point for each microphone.

Equation 3.2 then divides the range by the speed of sound to obtain the time it takes a sound wave to get from the scan point to each microphone. This time is multiplied by the frequency. Beamforming is done not only point by point (ξ_j) but also frequency by frequency (f_k), as every frequency requires a different phase shift for alignment. The product of time $[s]$ and frequency $[1/s]$ is a dimensionless number that specifies the number of radians that the phase needs to shift in the complex circle. This steering vector contains the phase shift for each microphone to align them all for sound coming from the scan point.

Equation 3.3 then applies the steering vector to the measured signal. The X is the Fourier transformed signal magnitude and phase for each microphone at a single frequency. It is multiplied by its complex conjugate transpose, indicated by the (*) symbol, to obtain the cross-spectral matrix.



(a) Cartesian coordinates

(b) Spherical coordinates

Figure 3.2: Two coordinate systems

Multiplying this by the steering vector and dividing it by the norm to account for the effect of four multiplications leads to $B(\xi_j, f_k)$, which is a single complex value that represents the signal strength coming from the scan point. This process is repeated over all scan points in the area for all relevant frequencies.

3.2 Coordinate system change

The equations in the previous section can be used to obtain a three-dimensional plot that specifies the drone location. Such as the point pictured in figure 3.2a. However, scanning every point in a three-dimensional area can take a long time. It is possible to set one dimension constant and reduce the scan grid to a plane to find the location that is in line with the sound source, but this does not cover all directions. Additionally, if the source is detected at the edge of the scan grid it likely means that the source is somewhere outside of the scan grid.

There is a more efficient way. Spherical coordinates facilitate scanning an area in a sphere around the array. Setting the range constant provides information about the angles relative to the array that are in line with the sources of noise. The source will always be somewhere outside the scan grid if the set range is small, but its relative horizontal and vertical angles are known. An example is shown in figure 3.2b. This section describes the equations used to convert the scan points x_j, y_j and z_j into the scan points r_j, ϕ and θ, which are the range, azimuth angle and elevation angle respectively. The range r_j is set constant at 1 meter, leading to a new version of equation 3.1. Equations 3.2 and 3.3 are unchanged.

Factoring equation 3.1 and substituting the equivalent spherical expressions:

$$x_j = r_j cos\phi cos\theta \tag{3.4}$$

$$y_j = r_j sin\phi cos\theta \tag{3.5}$$

$$z_j = r_j sin\theta \tag{3.6}$$

Leads to this equation:

$$\mathbf{r_{n,j}} = \sqrt{\begin{array}{r} r_j^2 cos^2\phi cos^2\theta - 2r_j cos\phi cos\theta x_n + x_n^2 \\ +r_j^2 sin^2\phi cos^2\theta - 2r_j sin\phi cos\theta y_n + y_n^2 \\ +r_j^2 sin^2\theta - 2r_j sin\theta z_n + z_n^2 \end{array}} \tag{3.7}$$

This equation can be simplified by removing most of the squares using this trigonomic identity and equation for the range:

$$cos^2(a)cos^2(b) + sin^2(a)cos^2(b) + sin^2(b) = 1 \tag{3.8}$$

$$x_n^2 + y_n^2 + z_n^2 = r_n^2 \tag{3.9}$$

Leading to:

$$\mathbf{r_{n,j}} = \sqrt{r_j^2 + r_n^2 - 2r_j(x_n cos\phi cos\theta + y_n sin\phi cos\theta + z_n sin\theta)} \tag{3.10}$$

Which, for the purpose of enabling further simplification, can be rewritten to:

$$\mathbf{r_{n,j}} = r_j\sqrt{1 + (\frac{r_n}{r_j})^2 - \frac{2}{r_j}(x_n cos\phi cos\theta + y_n sin\phi cos\theta + z_n sin\theta)} \tag{3.11}$$

In all relevant scenarios the distance between the drone and the center of the array will be at least an order of magnitude greater than the distance between the microphones and the center, $r_j >> r_n$, which means that the square will be negligible, $(\frac{r_n}{r_j})^2 = 0$. Additionally, given the form of the equation and that the microphones are close to the center of the array, a Taylor series approximation can be taken around $x = 0$ for $\sqrt{1-x} = 1 - \frac{x}{2} - \frac{x^2}{8} - \frac{x^3}{16} - ....$ Only the first elements $(1 - \frac{x}{2})$ are considered because the rest is negligible. The Taylor approximation and neglecting higher order terms results in:

$$\mathbf{r_{n,j}} = r_j - (x_n cos\phi cos\theta + y_n sin\phi cos\theta + z_n sin\theta) \tag{3.12}$$

This equation is used in the beamforming algorithm for all subsequent applications in this project.

3.3 Applying 2D angle beamforming to drones

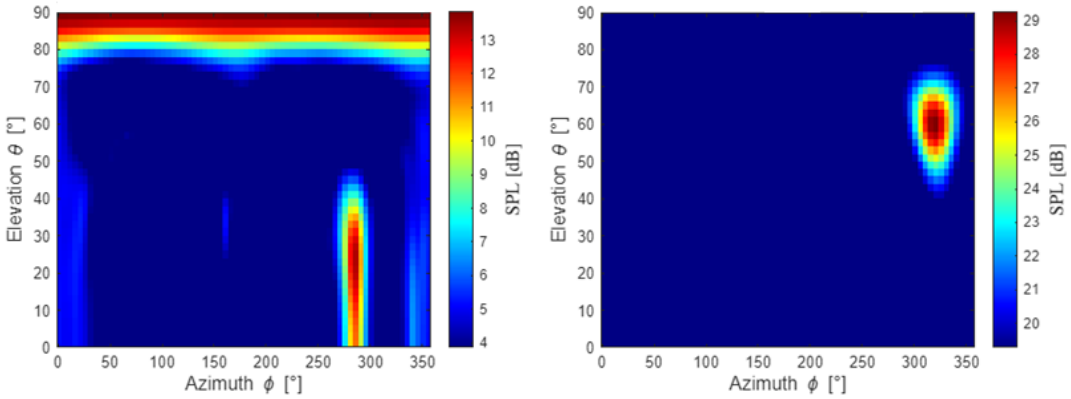
The spectrogram in figure 2.1 in section 2.2 shows the noise produced by a flyby of a DJI Phantom 3 drone. This measurement was done during a different experiment with a larger array with 58 microphones and a diameter of 3.2m and not the array pictured in figure 2.5. The beamform plots in this section will therefore have a much better spatial resolution.

Selecting three frequency bands around the first three harmonics and using these in the equations in the previous sections leads to beamform plots as pictured in figure 3.3. Figure 3.3a was made with a data snapshot at $t = 15$ seconds. The spectrogram suggests that the drone is still distant at this point. Its sound levels are at somewhat similar levels as the background noise, especially at the lowest frequency harmonic. This is visible in the beamform plot with the noticeable presence of noise at elevation angles directly above the array. Figure 3.3b is made at $t = 30$ seconds where the drone is much closer to the array. The sound levels are over 15dB higher than those of the background noise in these three frequency bands. The drone can be accurately tracked in these conditions. Tracking is achieved by saving the location of the maximum SPL value in the plot and plotting these maximum locations over time. Figure 3.4 shows the locations of the maximum values for both elevation and azimuth over time. The elevation angle relative to the array increases the closer the drone is to the array and decreases again after it passes by. The azimuth plot shows a shift from one angle to another. The large shift in elevation angle and mostly constant azimuth angle before and after the flyby suggest that the drone flew very close by the array.

3.4 Point spread function

The plots in the previous section, section 3.3 were produced on the basis of measurements obtained with a large microphone array. The size of the array is an important factor in the resolution of the beamform plot. Equation 3.13, the Rayleigh criterion, shows the related factors.

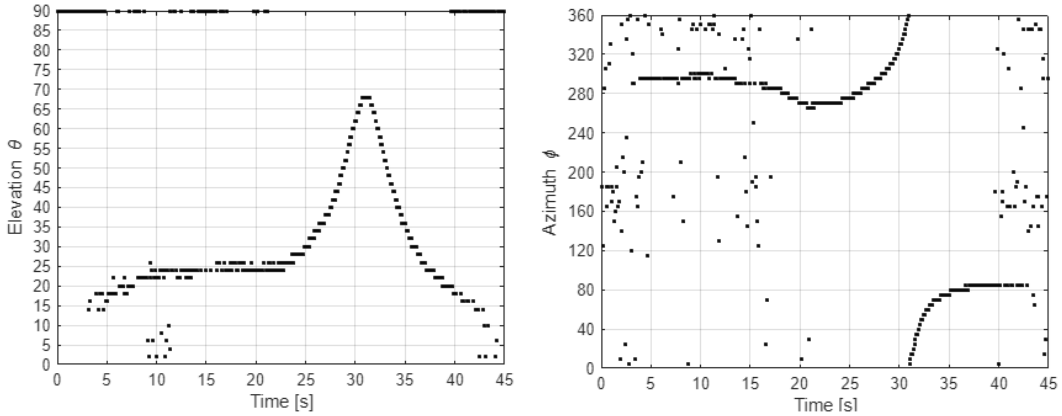
$$\theta_B = 1.22\frac{\mathrm{c}}{f\mathrm{L}} = 1.22\frac{\lambda}{\mathrm{L}} \tag{3.13}$$



(a) Drone flyby at t = 15s

(b) Drone flyby at t = 30s

Figure 3.3: Beamformed snapshots



(a) Elevation versus time

(b) Azimuth versus time

Figure 3.4: Tracking plots DJI Phantom 3 flyby

Beamwidth θ_B [rad] is the resolution limit, the best possible spatial resolution based on the given elements. A smaller beamwidth results in a plot with a better resolution. λ [m] is the wavelength. The shorter wavelengths that come with higher frequencies have a narrower spatial beamwidth. L [m] is the aperture. Increasing the aperture improves the array's ability to distinguish between noise coming from different directions and increases the resolution of the beamform plot. Unfortunately, the small array used in this project has a small aperture and it is used to measure drone noise which is mostly concentrated at lower frequencies.

The point spread function can provide some insight on the difference between the small and large arrays. A point spread function plot can be produced by multiplying the steering vector for every point, as used in the beamforming equation, with a steering vector directed to a single point. The resulting figure shows what the beamform plot would look like if there were only one source, at the specified location (b). The original equation 2.42 in the reference book 'aeroacoustic measurements' [19] can be represented by:

$$\text{PSF} = |\mathbf{g}(\mathbf{b})\mathbf{g}|^2 \tag{3.14}$$

The g is the regular steering vector in equation 3.2 and g(b) is a steering vector that refers everything to one single location. This contains no actual noise data, it is only the magnitude of the steering vector. Figure 3.5a shows the psf at 1000Hz for the array used for the drone in the previous section. Figure 3.5b shows the outcome of the same input applied to a different slightly smaller array. The scaling is chosen to be down to -20dB below the maximum value which is set at 0dB. 20dB is $1/100$ times the magnitude, it is included to obtain visible sidelobes.

Spatial sidelobes are artifacts of the beamforming method whose locations vary depending on the geometry of the measurement array. As mentioned in section 3.1, the beamform plot is constructed point by point. But when steering all microphones to a certain scan point it is possible that some of the microphones are in phase with noise from another location, resulting in a false magnitude attributed to the scan point. It is possible to apply a weighting function to the steering vector to reduce these sidelobes, but this would increase the beamwidth and it is not necessary for the small array. As sidelobes are generally lower in magnitude than the main lobe, especially with a low number of microphones and a singular source, they are not relevant for the localization effort.

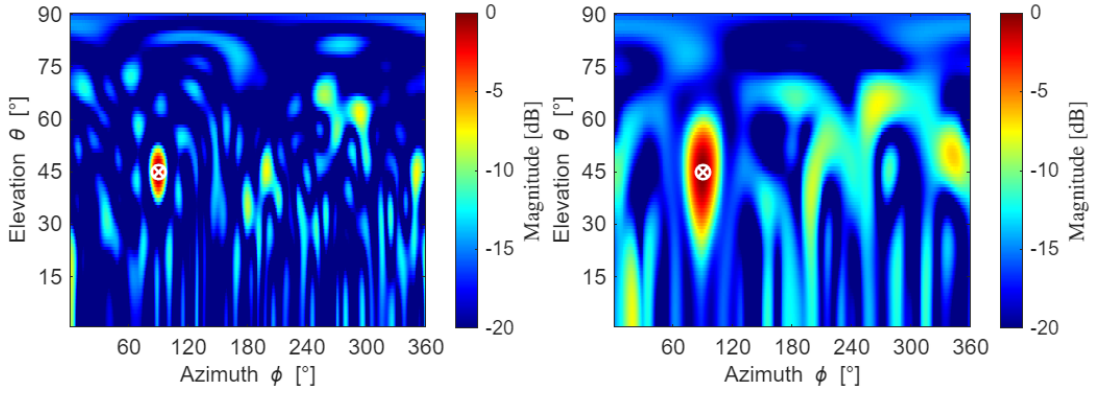
What can be relevant for this project are grating lobes. These are locations where all microphones are in phase with sound created by a source at another location. Thus, an incorrect location with potentially maximum magnitude. These are avoided when the distance (d) between microphones is less than half the wavelength (λ) of the sound. There are no grating lobes if equation 3.15 is true [20, equation 6.15].

$$\text{d} < \frac{\lambda}{2} \tag{3.15}$$

$$f = \frac{\text{c}}{\lambda} \tag{3.16}$$

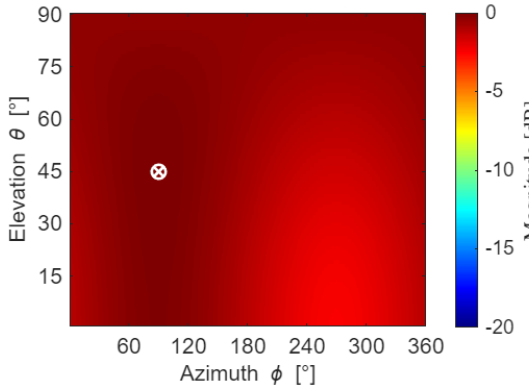
The spacing between elements is $0.0457m$. With equation 3.15 this means wavelengths smaller than 0.0914m can cause grating lobes. Using this value in equation 3.16 with a speed of sound c = 343m/s, leads to a frequency $f = 3753Hz$. Grating lobes can show up above this frequency for beamform plots based on measurements with the ReSpeaker array.

The array used in figure 3.5a has a diameter of about 3.2m and the array used in 3.5b has a diameter of about 2m. The small array used in figures 3.5c and 3.5d only spans 0.0457m or 0.06463m diagonally. Figure 3.5c shows a PSF with similar scaling as in figure 3.5a and figure 3.5b and figure 3.5d has much narrower scaling down to half the magnitude. The maximum scaling has to go to decimal fractions before a main lobe can be distinguished, essentially only showing the very tip of the highest peak. Sidelobes are not visible. It will be more difficult to localize sources with this large beamwidth.

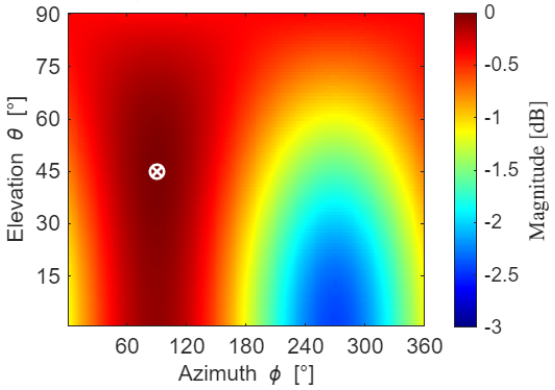


(a) PSF 1000Hz 58mic array

(b) PSF 1000Hz 32mic array



(c) PSF 1000Hz 4mic array



(d) PSF 1000Hz 4mic array with narrower scaling

Figure 3.5: Point spread functions

3.5 Functional beamforming

Functional beamforming is an alternative method of beamforming that includes an extra step. An exponent is applied after spectral decomposition which greatly reduces sidelobes, elucidating the real source contribution.

$$\mathrm{FB} = [g^* U \Sigma^{\frac{1}{\nu}} U^* g]^{\nu} \tag{3.17}$$

The g is again the steering vector. The U's are the eigenvectors of the cross spectral matrix, the XX^* part of the beamform equation, and Σ contains its eigenvalues. The exponent ν is the contribution that changes the output. This is achieved by effectively shielding the eigenvalues, that represent the source contribution, with an inverse exponent. Multiplying the eigenvalues with the eigenvectors recreates the original matrix which is then subjected to the exponent to lower the overall values. If the peak of a sidelobe is -3dB lower than the main peak, after $\nu = 100$, it is now -300dB lower than the main peak. The main peak is also sharpened, i.e. a reduction in beamwidth. [14]

Using $\nu = 1$ leads to a result that is identical to the conventional frequency domain beamforming discussed previously. A value of $\nu = 100$ leads to good functional beamforming results [13].

Functional beamforming can identify and elevate additional sources as those are shielded in the eigenvalues. For the purpose of this project the main attraction of functional beamforming is the reduction in beamwidth.

Chapter 4

Measurements with the ReSpeaker array in an anechoic chamber

A description of the small array was provided earlier in section 2.4. This chapter describes the setup of an indoor experiment with the small array in an anechoic chamber and its results. The setup is described in section 4.1. The results are obtained by applying the beamforming theory from chapter 3 and varying four different parameters. The parameters are as follows:

The effect of the choice of frequency band in section 4.2. The effect of the source location in section 4.3. The effect of the array orientation in section 4.4 and the effect of the chosen beamforming method in section 4.5.

The objective of the experiment is to gain information on the capabilities of the array to determine the viability of an outdoor test and assess the added value of placing the array at a 45-degree angle.

4.1 Experiment setup

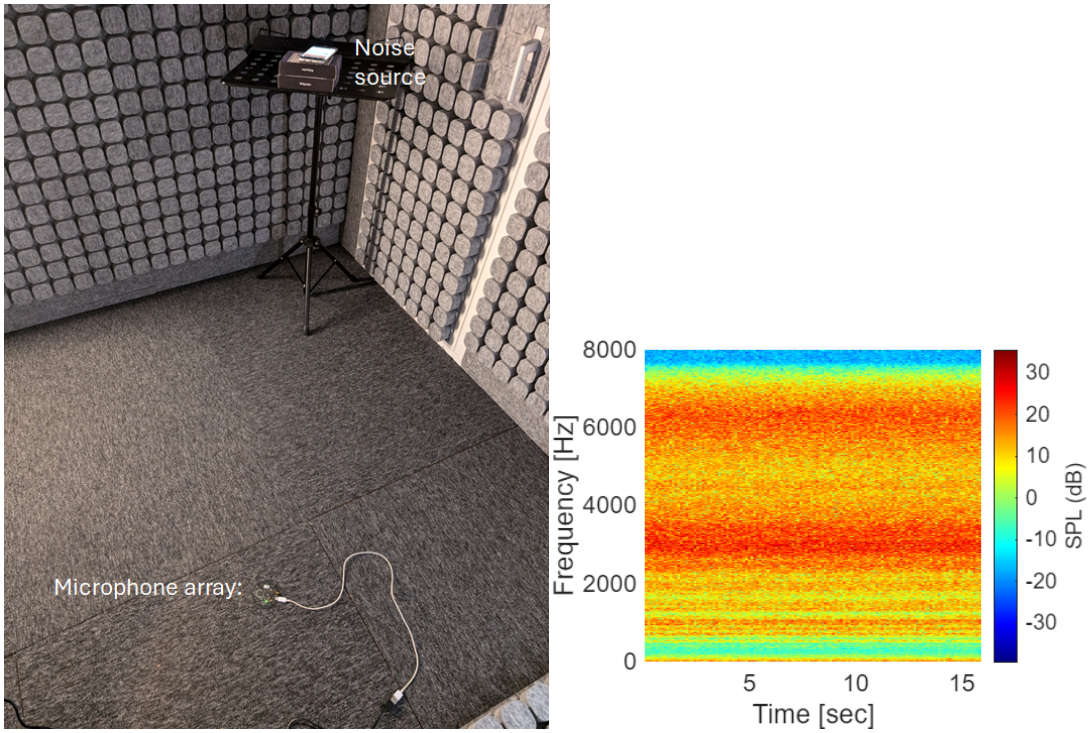
An experiment was performed in the PsychoAcoustic LIstening LAboratory (PALILA) in the Aerospace engineering faculty at the TU Delft university. It is a small anechoic chamber. Conducting a sound measurement in such an environment ensures that there is no interference from outside noise sources nor reflected sound (echoes) from the walls or ceiling above a certain threshold frequency.

The array was placed on the ground near the middle of the room and the source of noise elevated on a stand near the wall, facing the array. The noise was chosen to be white noise from a website [1], produced by a Samsung Galaxy s7 Edge phone. White noise is broadband noise, covering many frequencies, with ideally equal intensity at all frequencies. This facilitates beamforming at many different frequencies. Figure 4.1 shows the setup of one of the measurements performed in the PALILA and the noise spectrum produced by the phone on the stand during one of the measurements. The spectrum indicates the presence of a low-pass filter on the circuit board, filtering out frequencies over 7.5kHz. It is also noticeable that the sound levels were not as equally distributed as expected from a white noise signal.

4.2 Effect of frequency on beamform plots

Chapter 3 mentioned the effect of frequency on beamform plots. In particular, the beamwidth scaling with frequency in equation 3.13. This section shows the effect the choice of frequency range has on beamforming with data from the small array. First four different bands of equal width are shown and compared with their PSF in section 4.2.1. Afterwards, the effect of a narrow range and a grating lobe is shown in section 4.2.2. Lastly, the more ideal result of a wider frequency range is shown in section 4.2.3.

[1] Noise signal source: onlinetonegenerator.com



(a) One of the experiment setups

(b) Noise used in section 4.2

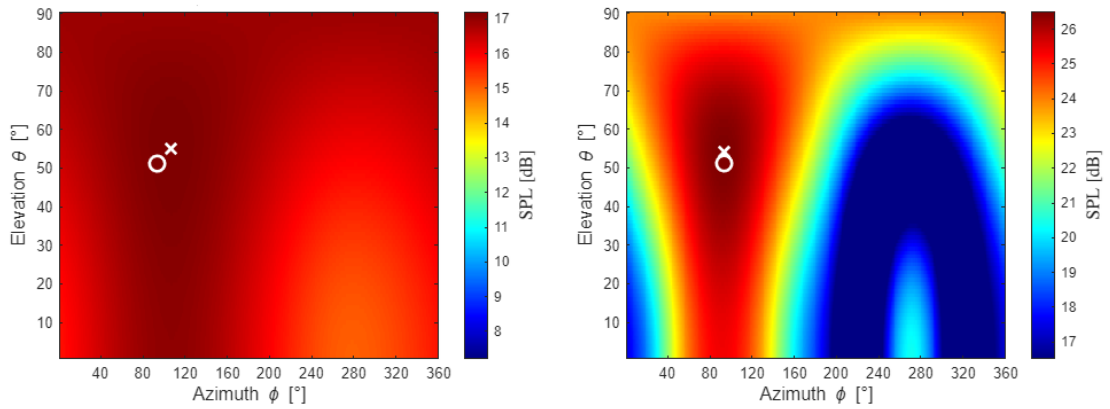
Figure 4.1: PALILA experiment conditions

4.2.1 Four different frequency bands

Frequency is a major factor in the signal beamwidth, as seen in equation 3.13. A higher frequency tone has a smaller wavelength with a smaller beamwidth, leading to a more accurate localization. Figure 4.2 shows beamform plots that cover four different frequency ranges to determine the effect of frequency for this small array. These are compared with point spread functions that cover the same frequency ranges in figure 4.3. Every beamform plot is marked by a white circle and a white cross. The white cross marks the location of the maximum value in the beamform plot. The white circle is the reference maximum location. The reference azimuth angle is obtained from the Python script mentioned in section 2.4 and the reference elevation angle is calculated from distances obtained with a measuring tape. The azimuth angle was crudely verified by visual inspection comparing the relative direction of the phone with the expected DOA angles as specified by the green labels in figure 2.5. The DOA angles in the green labels refer to the azimuth angle output of the internal beamformer. All beamform plots cover a snapshot range of 0.1 seconds at t = 4 seconds, mimicking the tracking effort in section 3.3 where the maximum is recorded for every snapshot of 0.1 seconds.

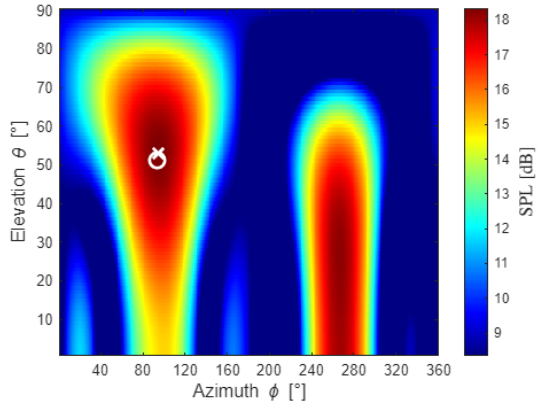
The plots show the narrower beamwidth that comes with higher frequencies. An increase in frequency shrinks the red area with similar sound levels. The azimuth angles of the cross and the circle are a good match except for the first scenario with the 500-1500Hz band as the beamwidth is too large there. However, the elevation angles are slightly off in all scenarios. This appears to be an effect of the larger beamwidth in elevation. The red area in the plots is more stretched out in the vertical direction.

What also stands out is the emergence of additional noise sources at higher frequencies. The fact that they also appear in the PSF plots in figure 4.3 means that these are artifacts of the beamforming method called grating lobes and sidelobes, as discussed in the previous chapter.

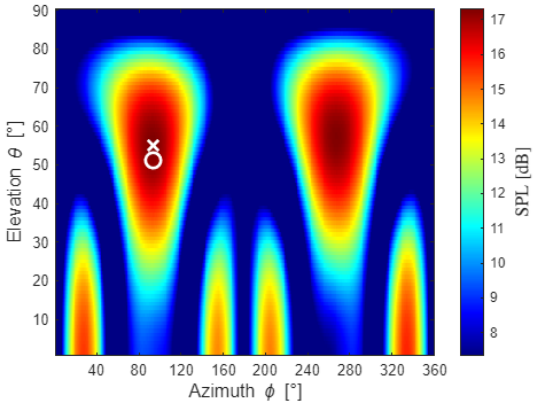


(a) Beamforming 500-1500Hz

(b) Beamforming 2500-3500Hz

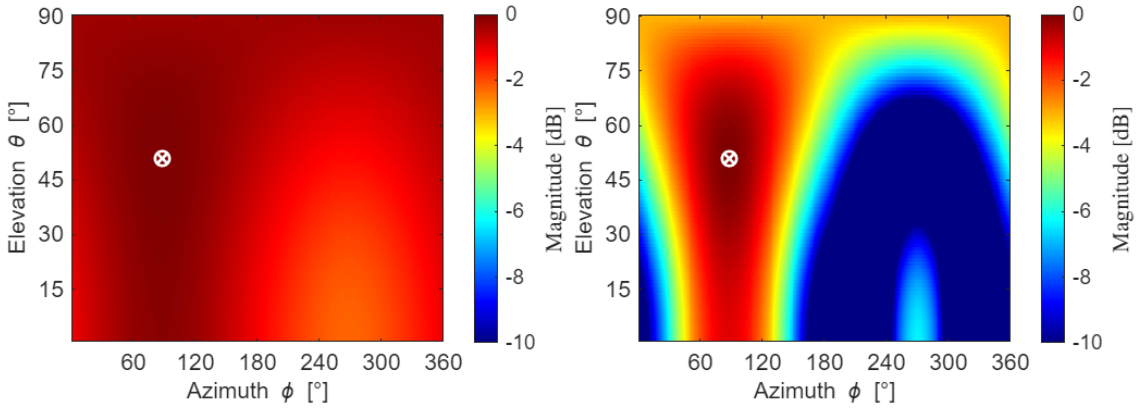


(c) Beamforming 4500-5500Hz



(d) Beamforming 6500-7500Hz

Figure 4.2: Beamform plots ReSpeaker array at various ranges



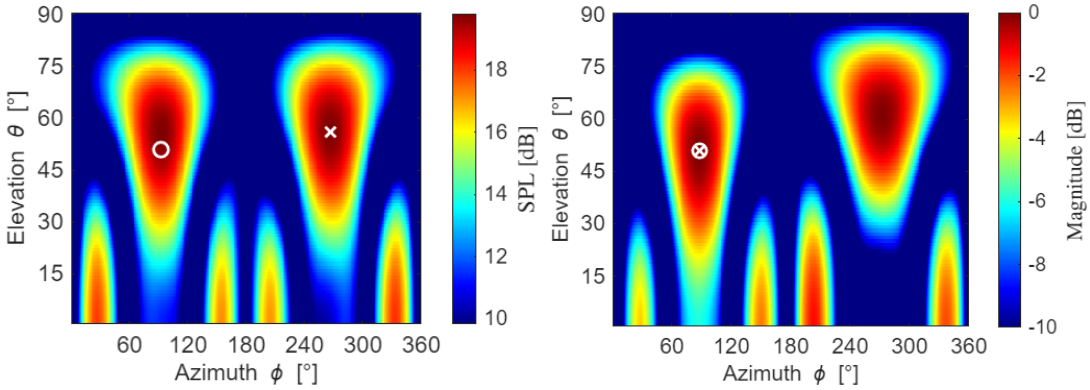
(a) PSF 500-1500Hz

(b) PSF 2500-3500Hz

(c) PSF 4500-5500Hz

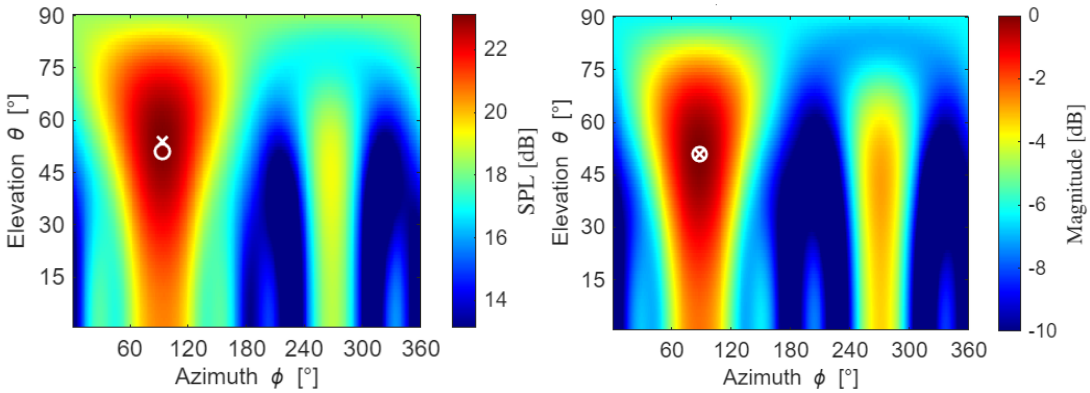
(d) PSF 6500-7500Hz

Figure 4.3: PSF plots ReSpeaker array at various ranges



(a) Beamforming 6500-7000Hz (b) PSF 6500-7000Hz

Figure 4.4: Narrower band with misdirection



(a) Beamforming 2000-7000Hz (b) PSF 2000-7000Hz

Figure 4.5: Wider frequency band with reduced side and grating lobes

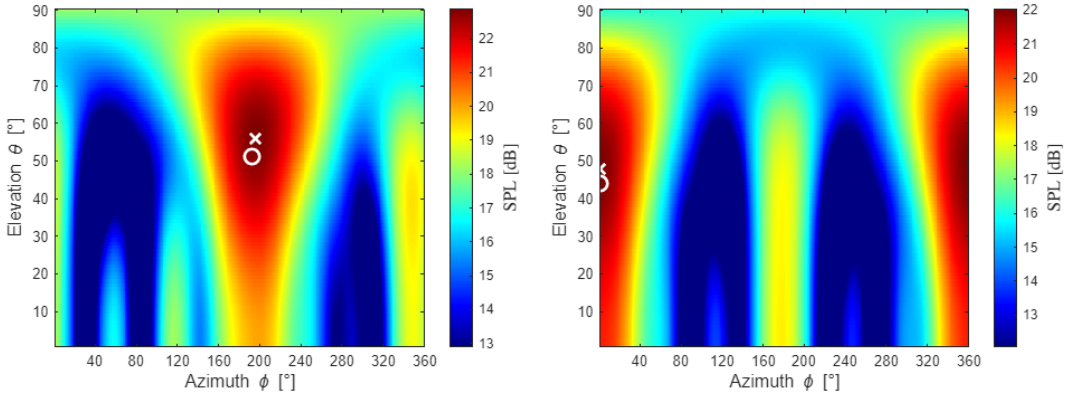
4.2.2 Narrower frequency band

Figure 4.4 shows the result of beamforming a slightly too narrow high frequency range. The maximum is detected in the wrong location due to a grating lobe. The magnitude at the circle is 19.68dB and the maximum magnitude in the same lobe is 19.81dB. The magnitude of the global maximum in the grating lobe is 19.82dB at the same elevation as the other local maximum. The difference is 0.0055dB. The magnitude of the maximum in the grating lobe is -0.02dB in the theoretical PSF plot, which is still essentially equal to the maximum at the expected location. Beamforming a continuous narrow range at higher frequencies where grating lobes are possible should be avoided.

4.2.3 Wider frequency band

Using a frequency band that is 5kHz wide as opposed to the 0.5kHz band in the previous section leads to a beamform plot that is significantly less affected by grating lobes or sidelobes, see figure 4.5. Beamforming over a wide range and then averaging out the resulting values leads to a plot where all the side lobes and grating lobes associated with different frequencies at different locations average out.

Each frequency comes with its own sidelobes. It is therefore useful to beamform over a wider range, especially for higher frequencies. The magnitude of the maximum in the grating lobes of the higher frequency ranges can equal the magnitude of the real source. This can lead to misdirection if only the maximum value of the plot is considered.



(a) Beamforming 2000-7000Hz

(b) Beamforming 2000-7000Hz

Figure 4.6: Beamforming sources in other directions

4.3 Effect of source direction

This section investigates if the array gives similar results in multiple directions to establish that the array can be used to locate sources in more than one direction. It was achieved by moving the source of noise to different places in the room. The resulting beamform plots will have their maximum locations at different angles. Figure 4.6 shows two beamform plots with sound sources at locations different from the source used in section 4.2. The array appears to be almost equally capable in all three directions, with small variations in elevation and azimuth offset.

4.4 Effect of array orientation

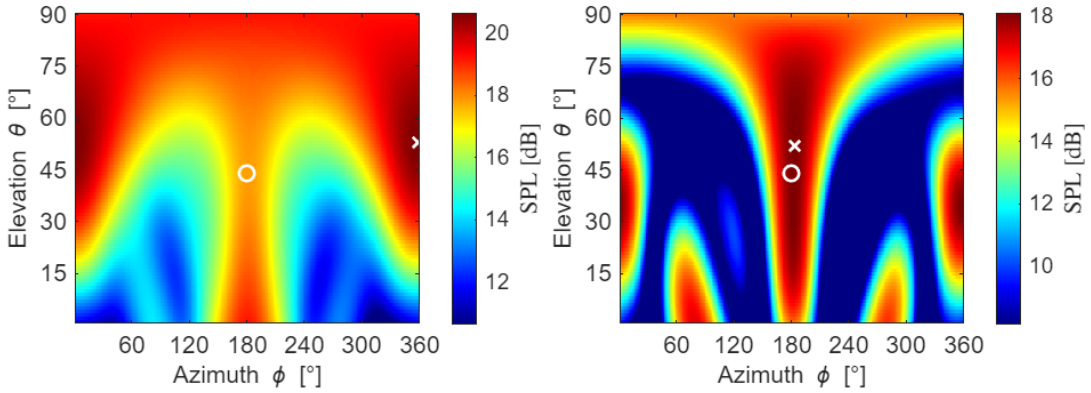
The elevation angle was slightly misplaced in the previous plots. It is possible that increasing the vertical distance between the microphones may lead to better results. The initial idea behind this was an increase in aperture in the vertical direction. A greater aperture leads to a smaller beamwidth, as given by equation 3.13, which would solve the issue. A wooden implement was produced that supports the array in a 45-degree orientation to facilitate this. The beamforming results of measurements with the array supported by this element are shown in section 4.4.1. Because this addition could compromise the previous omnidirectional detection ability, a measurement is done with the array facing away from the sound source in section 4.4.2.

4.4.1 Array placed at a 45-degree angle

The measurement of the array placed in a 45-degree orientation was performed with a horizontal and vertical distance of about 1 meter between the sound source and the array. The objective was to obtain a beamform plot with the main source in the center of the plot. But for an array angled at 45 degrees towards the source, this places the source directly overhead. Unfortunately, this leads to an ambiguous azimuth angle, as indicated in figure 4.7. Figure 4.7a shows the beamformed range of 2-7kHz, which has the reference circle at 180° azimuth and the actual maximum at 360° azimuth. Figure 4.7b shows a subset of this range, 6-7kHz, where the maximum is located near the expected azimuth angle. The internal beamformer in the ic on the array that was used as azimuth reference was insistent on 180° but it did have a habit of getting stuck on a single value sometimes.

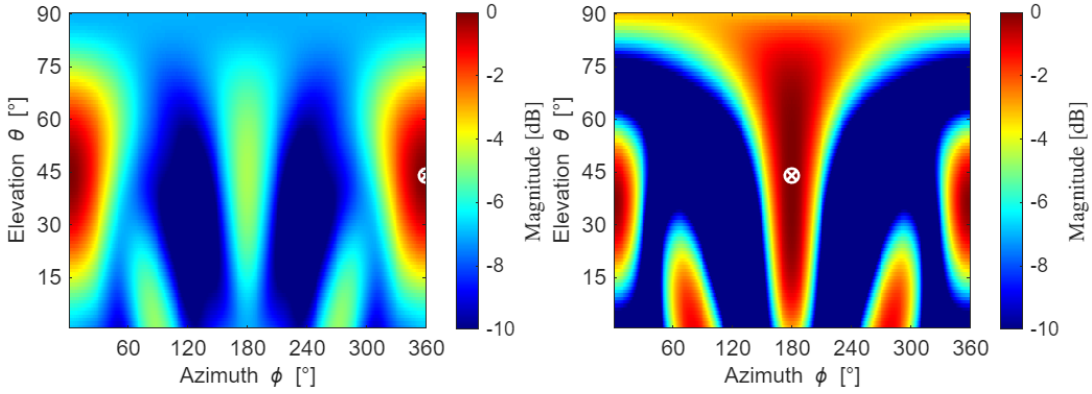
The PSF plots in 4.8 show the expected outcomes matching the beamformed maximums. The beamform plot in figure 4.7a is a combination of detected sources at either 180° or 360°. A PSF plot matching the white reference circle in figure 4.7a looks like figure 4.8b but without the sidelobes.

These inconsistent results necessitate another test for a valid assessment of the accuracy of the 45-degree array.



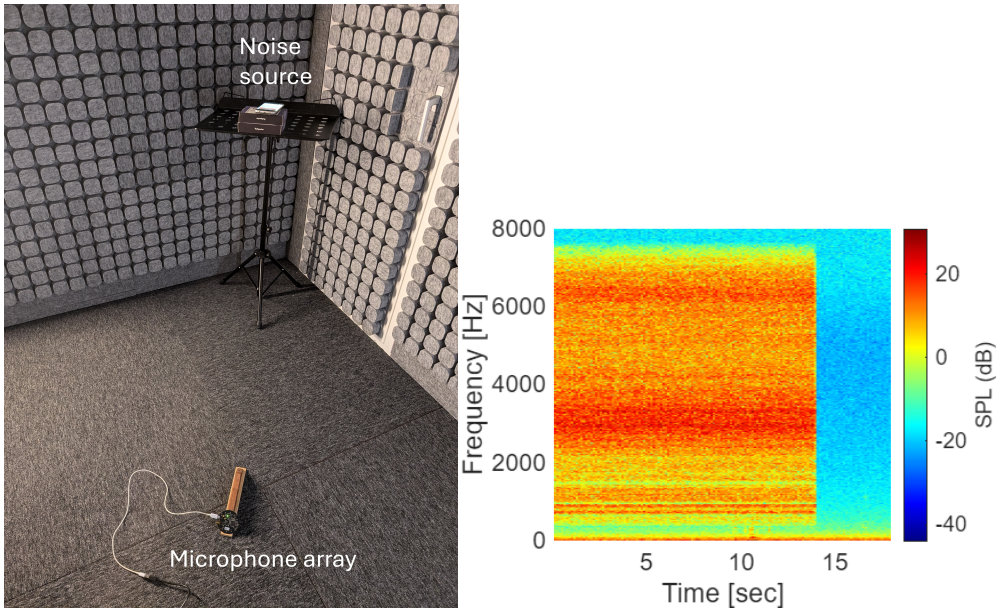
(a) Beamforming 2000-7000Hz (b) Beamforming 6000-7000Hz

Figure 4.7: Beamforming a 45-degree array measurement



(a) PSF 2000-7000Hz, 360° azimuth (b) PSF 6000-7000Hz, 180° azimuth

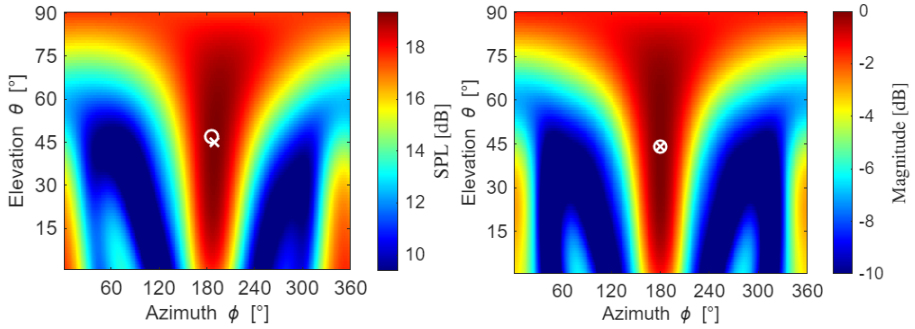
Figure 4.8: PSF 45-degree array measurement



(a) Experiment setup for section 4.4.2

(b) Noise during experiment in section 4.4.2

Figure 4.9: Setup of a 45-degree array facing away from the source



(a) Beamforming reverse orientation, 2000-7000Hz

(b) PSF 2000-7000Hz

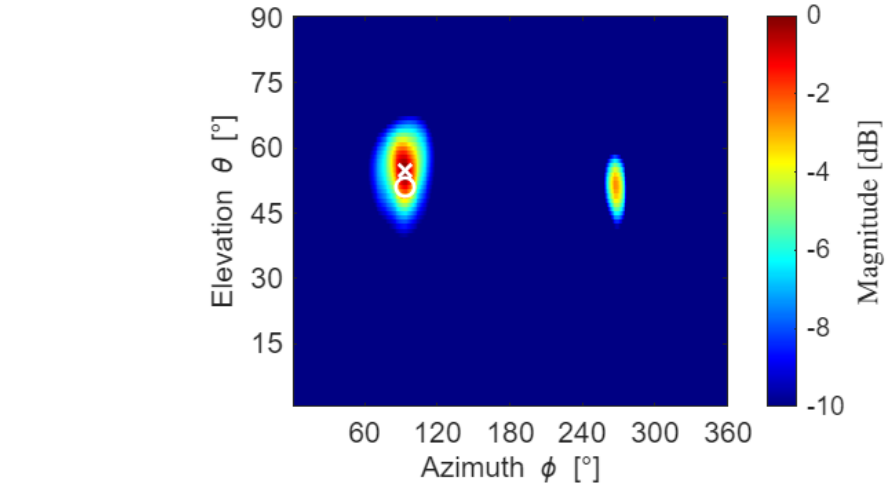
Figure 4.10: Beamform output 45-degree array facing away from the source

4.4.2 Reversed orientation

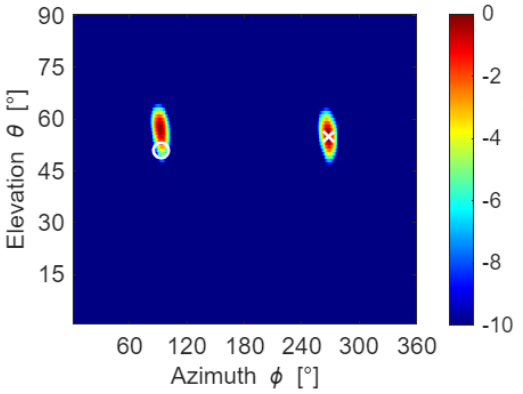
One measurement was done with the array in a reversed position to assess the beamform accuracy with noise coming from behind the angled array. The setup is shown in figure 4.9a and the measured noise in figure 4.9b. The noise signal cuts off after 14 seconds but this has no effect on a beamformed snapshot at an earlier time. The beamform plot in figure 4.10a looks nearly identical to the PSF plot in 4.10b and the white circle and the white cross are a close match. This outcome suggests that the array is very capable at measuring sound coming from behind the circuit board. At least from a source almost perpendicular to it. There is no ambiguity in azimuth as the microphones at the top are closest to the source both horizontally and vertically.

4.5 Effect of beamforming method

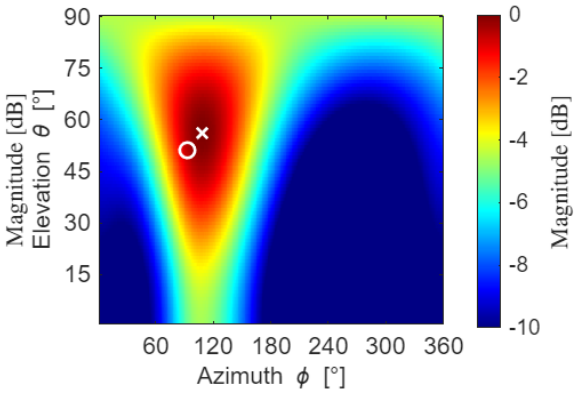
This section uses the same data as section 4.2, but this time the result is shown after applying functional beamforming. Section 3.5 mentioned the removal of sidelobes and a reduction in beamwidth. Figure 4.11 clearly shows a great reduction in beamwidth and complete removal of



(a) 2000-7000Hz



(b) 6500-7000Hz



(c) 500-1500Hz

Figure 4.11: Functional beamforming applied to three different frequency ranges

the sidelobes. Figures 4.11a, 4.11b and 4.11c are the functional beamforming versions of figures 4.5a, 4.4a and 4.2a respectively. Figures 4.11a and 4.11b show that grating lobes are not adversely affected by functional beamforming. Their sound levels do not decrease relative to the main source and figure 4.11b still has the maximum at the grating lobe similar to figure 4.4. Figure 4.11c shows a much clearer peak than figure 4.2a, which was entirely covered in shades of red.

Unfortunately, the potential increase in accuracy did not materialize. The distance between the white circle and the white cross is barely affected. There are only changes up to 1° between conventional and functional beamforming and not always in favor of the functional beamforming output. Table 4.1 compares the output shown in figure 4.11 with the output from conventional beamforming in the three earlier sections with the horizontal array. The most accurate result is marked in green. This is not applied to the azimuth angles in the $6500 - 7000Hz$ measurements because those are far away in a grating lobe.

	Conventional beamforming		Functional beamforming	
Frequency band	azimuth	elevation	azimuth	elevation
reference white circle	93	51	93	51
2000-7000Hz	+0	+3	+0	+4
6500-7000Hz	+174	+5	+175	+4
500-1500Hz	+14	+4	+15	+5

Table 4.1: Accuracy comparison of conventional and functional beamforming applied to measurements in the PALILA of a single source with the ReSpeaker mic array v2.0

Chapter 5

Indoor drone flight

A second experiment was conducted in the Cyberzoo facility in the aircraft hangar at the TU Delft aerospace engineering faculty. It is a cage of 10 by 10 meters surrounded by screens and nets. It was created for the purpose of testing drones and it can be filled with various obstacles. The purpose of testing here for this project was to measure the noise of a drone flying by the array and to obtain a valid assessment of the 45-degree array orientation. The former facilitates localization plots over time and the latter determines if there is a reason to bring the wooden implement for the later outdoor measurement. The setup and reference angles are described in section 5.1. The results of the experiment are shown graphically in section 5.2.

5.1 Experiment setup

The experiment in chapter 4 was conducted by measuring with one array at a time. The experiment in this chapter uses two arrays simultaneously for a direct comparison of the same data. Both arrays were placed next to each other at the edge of the green 8x8m field pictured in figure 5.1. The results shown in this chapter attempt to localize a Parrot Bebop drone flying in a straight line at the intersection between the green and the blue area on the opposite side, at an altitude of around 5 meters. Because the array was placed off-center by 0.7m, the distance varies between 9.28m, 28.3° elevation, at the farthest corner, to 8m and 32° elevation near the middle and 8.65m and 30.0° elevation at the other corner. The difference in azimuth angle between the two distant corners, with respect to the arrays, is 53°.

5.2 Experiment results

The spectrogram in figure 5.2 shows the harmonics of the BPF produced by the propellers of the Parrot Bebop drone cover almost the entire frequency spectrum. The bottom is dominated by background noise from other equipment in the hangar and the top is filtered out by the low pass filter on the circuit board. In the frequency range that is covered by drone noise, the first four harmonics of the BPF in figure 5.2 have the highest sound levels, which is consistent with the figures in sections 2.1 and 2.2. Beamforming is applied to four narrow frequency bands that cover these first four harmonics: $380 - 410$Hz, $770 - 810$Hz, $1160 - 1210$Hz and $1540 - 1610$Hz. The latter two ranges cover wider frequency bands to capture the complete harmonic over time. The spectral width of the harmonics increasing with frequency is clear by comparing the thin red lines below 1000Hz in figure 5.2 with the green bands near 3000Hz. The harmonics appear to merge entirely at even higher frequencies. Multiples of the first BPF harmonic cover a multiple of the frequency bandwidth of the first BPF harmonic. As such, the high frequency range that is also beamformed for comparison is a wider band covering $5000 - 7000$Hz.

Because beamforming was performed with 0.1-second snapshots, the frequency resolution is 10Hz. The selected low frequency ranges took this into account and the first $380 - 410$Hz range therefore covers three frequency bins. Although the number of bins is small, because this is the first harmonic, the sound level is relatively high. Most of the signal strength coming from low frequencies leads

(a) Setup overview (b) Array placement

Figure 5.1: Cyberzoo setup, the arrays are placed on the intersection of the green and blue floor sections. The data is transferred through USB cables to the laptops on the chairs. The drone flies by above the green/blue intersection on the far side. The numbers refer to the microphones.

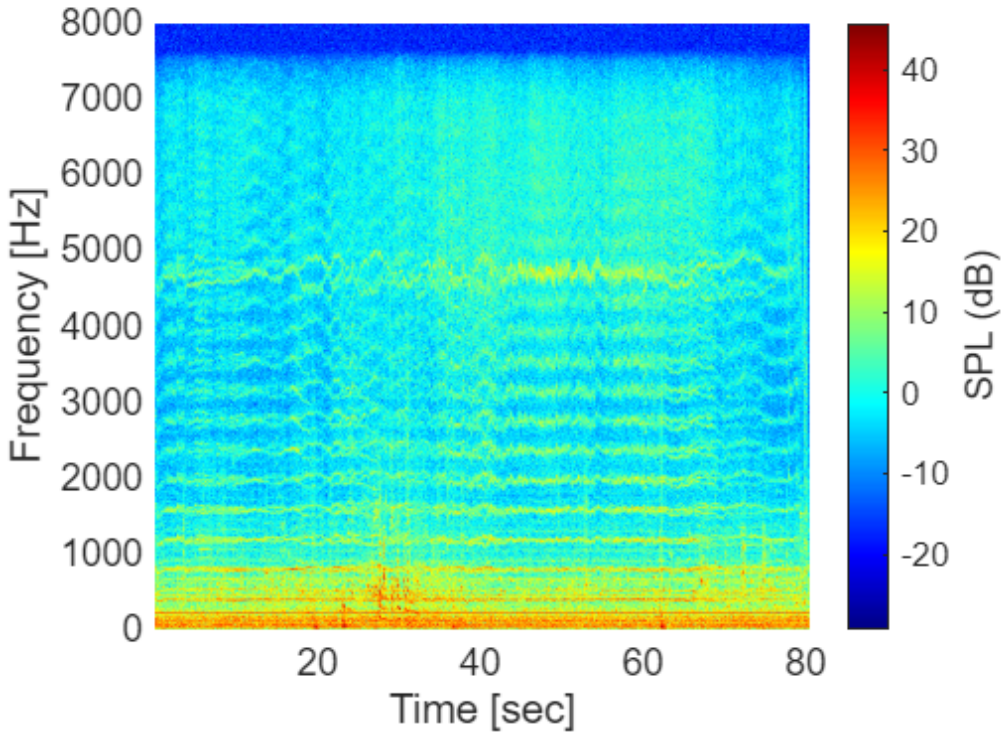
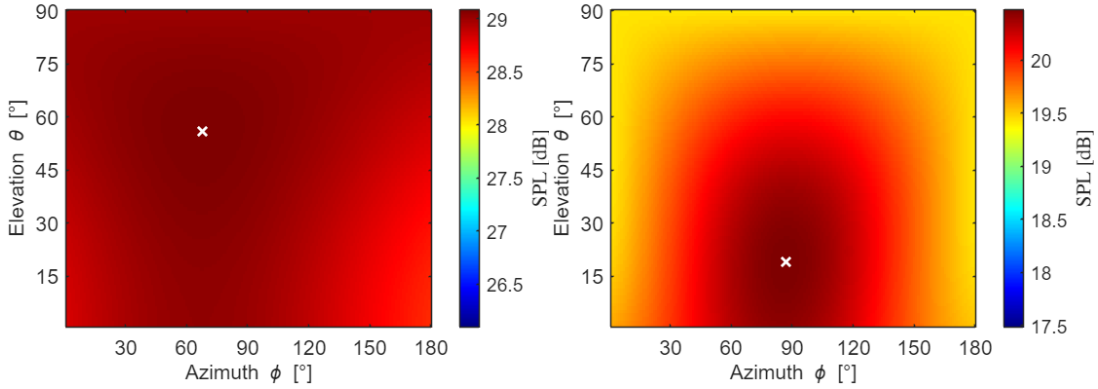
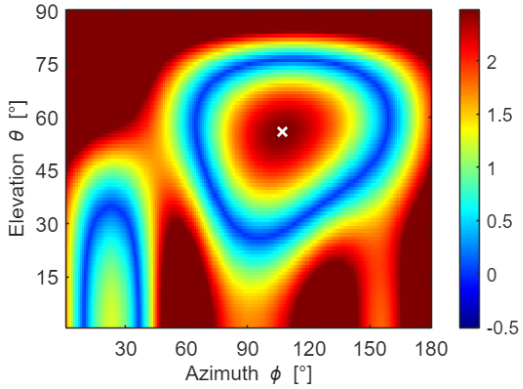


Figure 5.2: Spectrogram Parrot Bebop Cyberzoo flyby

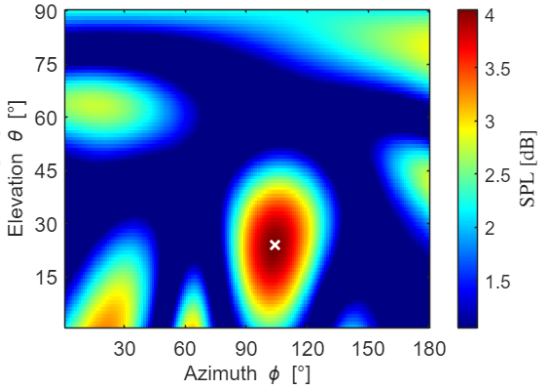


(a) Horizontal array t = 50s, low frequencies

(b) 45-degree array t = 50s, low frequencies



(c) Horizontal array t = 50s, high frequencies



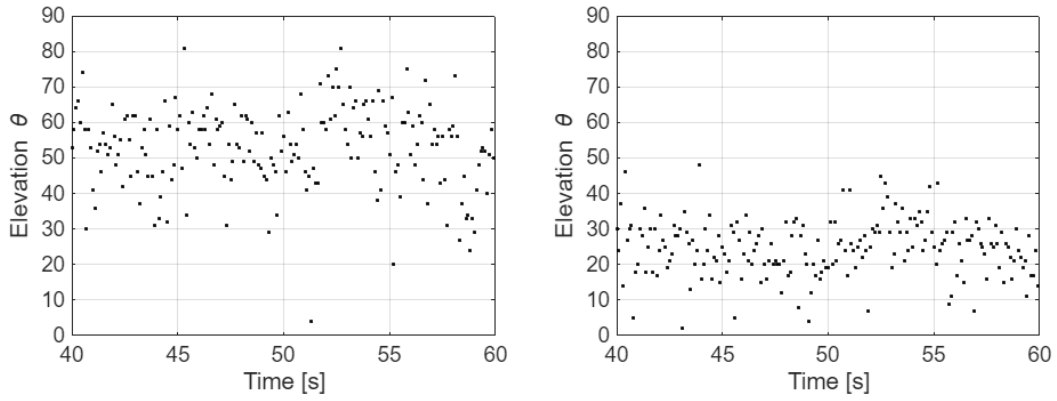
(d) 45-degree array t = 50s, high frequencies

Figure 5.3: Beamform plots at t = 50s of the indoor Parrot Bebop drone flyby. Low frequency range: 380 − 410Hz, 770 − 810Hz, 1160 − 1210Hz and 1540 − 1610Hz. High frequency range: 5000 − 7000Hz

to a large beamwidth, as visualized in the beamformed snapshots in figures 5.3a and 5.3b. The improved spatial resolution at higher frequencies in section 4.2 was due to nearly equal signal strength at all frequencies. All the beamform plots covering drone noise measured by the small array will look like the red square in figure 5.3a if the lower frequency harmonics are taken into consideration due to their relatively high sound levels, especially with a larger scaling in magnitude. The SPL colorbars in these plots only cover a 3dB range as opposed to the 10dB range in chapter 4. The plots in section 3.3 based on other data avoided this by having a large array aperture.

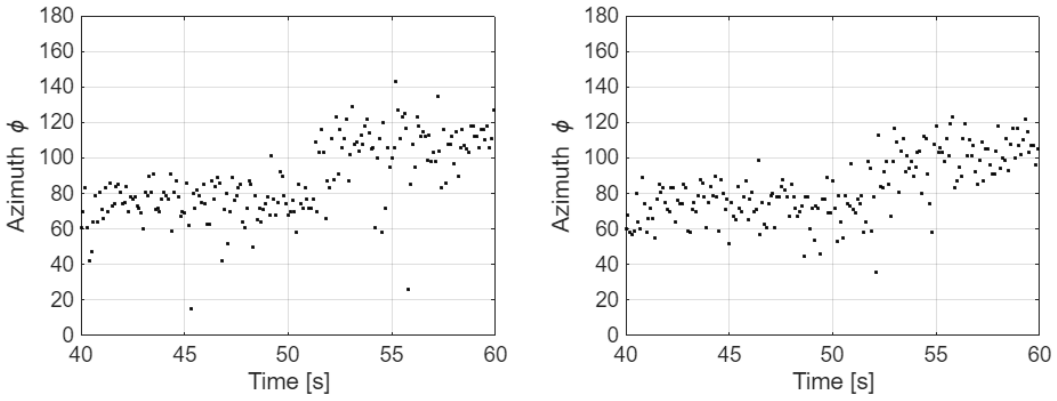
The locations of the white crosses in figure 5.3 are saved and plotted as dots in the tracking plots in figure 5.4 and figure 5.5, together with 199 other dots for each array orientation. There are only 200 dots because the lateral drone movement took less than 20 seconds. There are no changes in azimuth angle during the other 60 seconds of the flight covered by the spectrogram in 5.2.

The elevation angle dots in these plots should be around 30° and the azimuth angle dots should shift by about 50° over the course of the flight. Figures 5.4a and 5.5a show that the array placed flat on the ground have an offset in elevation tracking of about 20°. In contrast, figures 5.4b and 5.5 with data from the array placed at a 45-degree angle show more accurate elevation angles of around 30°. The array placed at a 45-degree orientation is clearly better in this aspect. Both arrays appear capable of accurately tracking the azimuth angles, showing a shift of about 50°. But the dots of the 45-degree array are slightly closer together. The 45-degree array may be slightly better in azimuth tracking, although the difference is very small. It is clear that for an indoor flight it is best to beamform higher frequency ranges when using a small microphone array.



(a) Elevation horizontal array

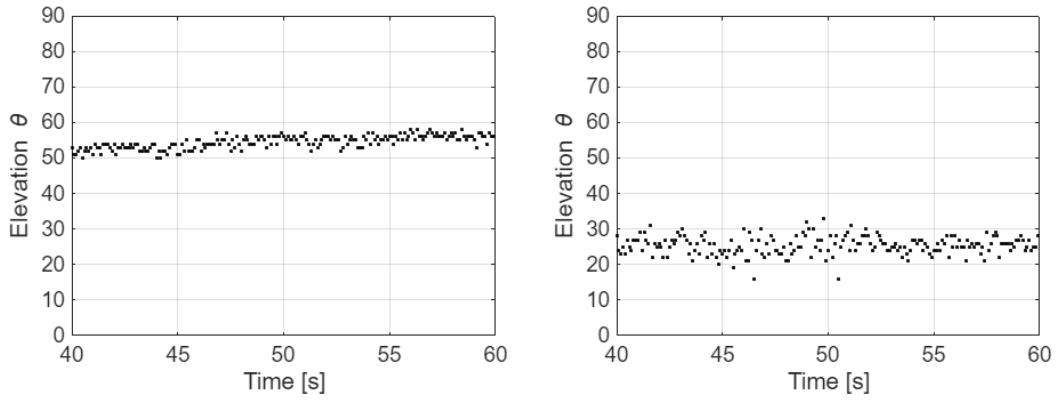
(b) Elevation 45-degree array



(c) Azimuth horizontal array

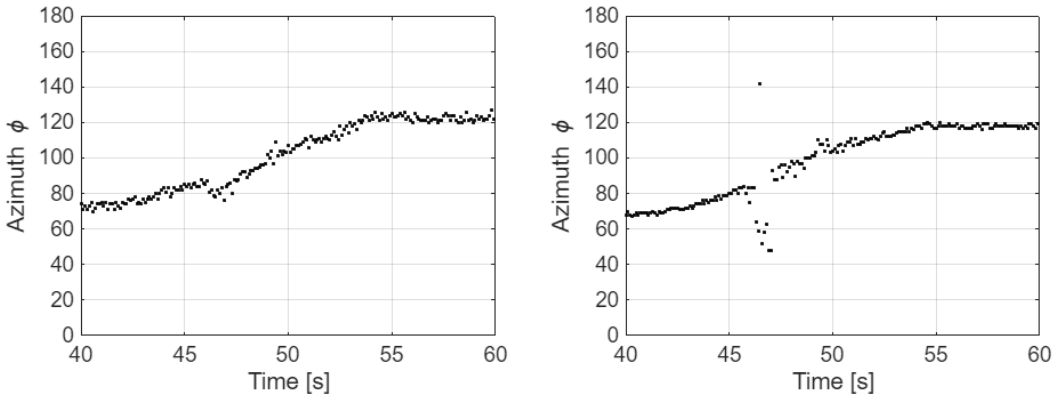
(d) Azimuth 45-degree array

Figure 5.4: Tracking plots of the Parrot Bebop drone flyby phase in the Cyberzoo, obtained by beamforming the low frequency bands: 380-410Hz, 770-810Hz, 1160-1210Hz, 1540-1610Hz



(a) Elevation horizontal array

(b) Elevation 45-degree array



(c) Azimuth horizontal array

(d) Azimuth 45-degree array

Figure 5.5: Tracking plots of the Parrot Bebop drone flyby phase in the Cyberzoo, obtained by beamforming the high frequency band: 5000-7000Hz

The implement that facilitates measuring with an array oriented at 45 degrees is certainly worth bringing along for the outside measurement because this orientation appears to result in better overall localization performance than the horizontal array.

CHAPTER 6

Outdoor drone flight

The third experiment was an outdoor drone flight in the Unmanned Valley near Katwijk aan Zee. The benefit of an outdoor drone flight is the availability of a GPS signal that can be used as a reference location. Comparing acoustic localization with GPS gives the option of a quantified measure of accuracy. This allows answers to the research questions. The outdoor experiment was performed on the afternoon of October 3rd 2024 at the Unmanned Valley.

The setup of the experiment is explained first in section 6.1. Then the details of four different drone flights conducted during the experiment are discussed in turn. The flight with a Parrot Bebop drone is discussed in section 6.2 and the flight with a DJI Mavic drone is discussed in section 6.3. The results of those two flights are shown first because they were not affected by interference from other aircraft. The other two flights had some noise from aircraft on approach to Schiphol Airport. A second flight with the Parrot Bebop that was hindered by interference is shown in section 6.4. A second DJI Mavic flight with interference is shown in section 6.5. Lastly, the outcome of this experiment is discussed and compared with the results of a similar experiment in section 6.6.

6.1 Experiment setup

The environmental conditions during the experiment are pictured in figure 6.1. Two ReSpeaker arrays were placed in the field, one in a horizontal position and the other angled at 45 degrees, facing the field as pictured in figures 6.2 and 6.3. The results of the indoor measurement in the chapter 5 predicted good results for this extra orientation. The arrays were placed on a towel just over six meters distant from a concrete slab used as drone take-off location. The reason for placing the arrays in the grass field was to reduce the effect of potential reflected sound from the surrounding solid concrete. The coordinate system shown in this picture is not in line with the cardinal directions. As figure 6.4 shows, the reference stripe in line with the array, visible in figure 6.2 is at an 18° angle with the nearest latitude line. The GPS coordinates from the drones will need to be rotated to the coordinate system used for the measurement arrays. A compass was not consulted during the experiment.

The two drones used in this experiment are a Parrot Bebop II and a DJI Mavic 2. Some specifications are listed in table 6.1. Linear interpolation was applied to the GPS signal to match it with the beamforming rate. Otherwise matching the beamform output with every third Parrot Bebop point would build up an error of 0.1s every 1000 snapshots. More data points of the consistent GPS signal from the DJI Mavic allows for a better localization analysis as the beamform output can have outliers.

The GPS data is saved as latitude, longitude and altitude. It was converted to Cartesian coordinates and corrected with equations 6.1 and 6.2 to match the coordinate system used for the array. The next step is a conversion to spherical coordinates, relative to the position of the array, which was achieved with equations 6.3, 6.4 and 6.5. The increased range that an outdoor flight allows was calculated with equation 6.3. The vertical elevation angle of the drone relative to the array was calculated with equation 6.4 and the horizontal azimuth angle was calculated with equation 6.5. These coordinates match those used in the beamforming process.

	Parrot Bebop 2	DJI Mavic 2
Weight [g]	500	907
height [mm]	89	84
width [mm]	328	242
length [mm]	382	322
number of propellers	4	4
blades per propeller	3	2
GPS logging rate [s]	0.03337	0.200

Table 6.1: Details of the drones used in the experiment

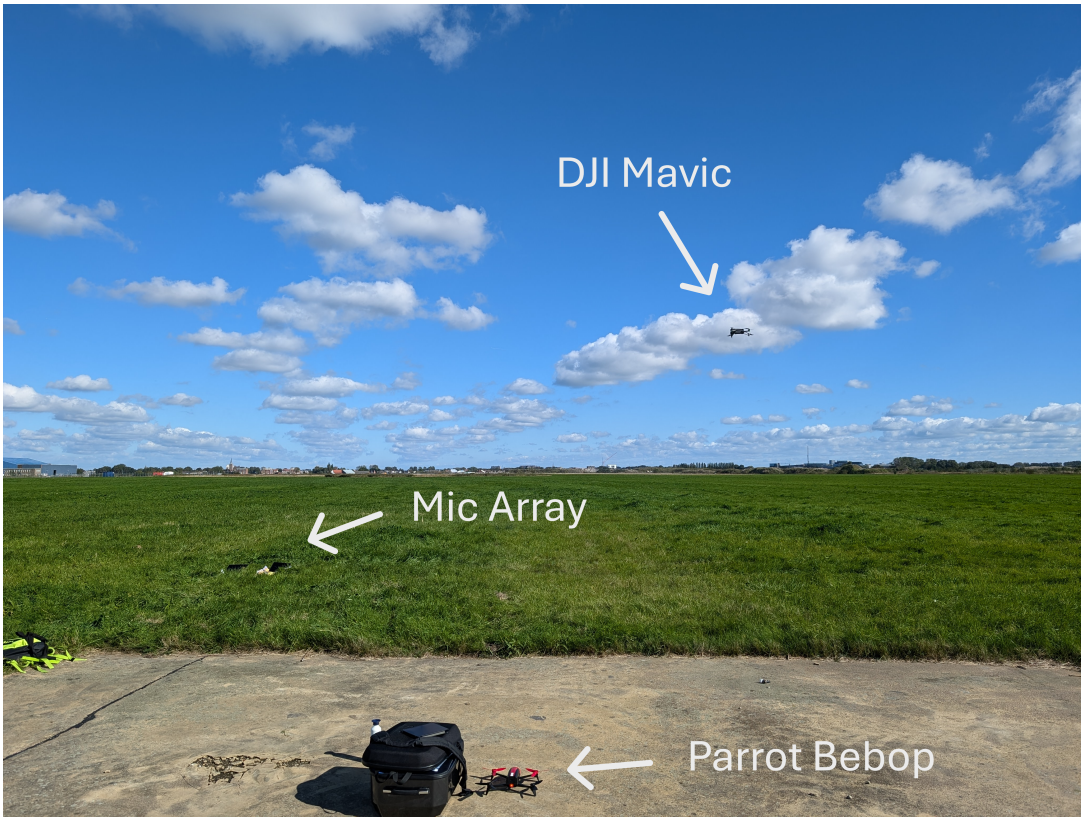


Figure 6.1: Experiment setup, from ground to drone

$$xEastb = xEast * cos(18°) + yNorth * sin(18°) \tag{6.1}$$

$$Ynorthb = -xEast * sin(18°) + yNorth * cos(18°) \tag{6.2}$$

$$R_{gps} = \sqrt{xEastb^2 + yNorthb^2 + zUp^2} \tag{6.3}$$

$$\theta_{gps} = asin(\frac{zUp}{R_{gps}}) \tag{6.4}$$

$$\phi_{gps} = sign(yNorthb) * acos(\frac{xEastb}{\sqrt{xEastb^2 + yNorthb^2}}) \tag{6.5}$$

Figure 6.2: Experiment setup, from drone to ground

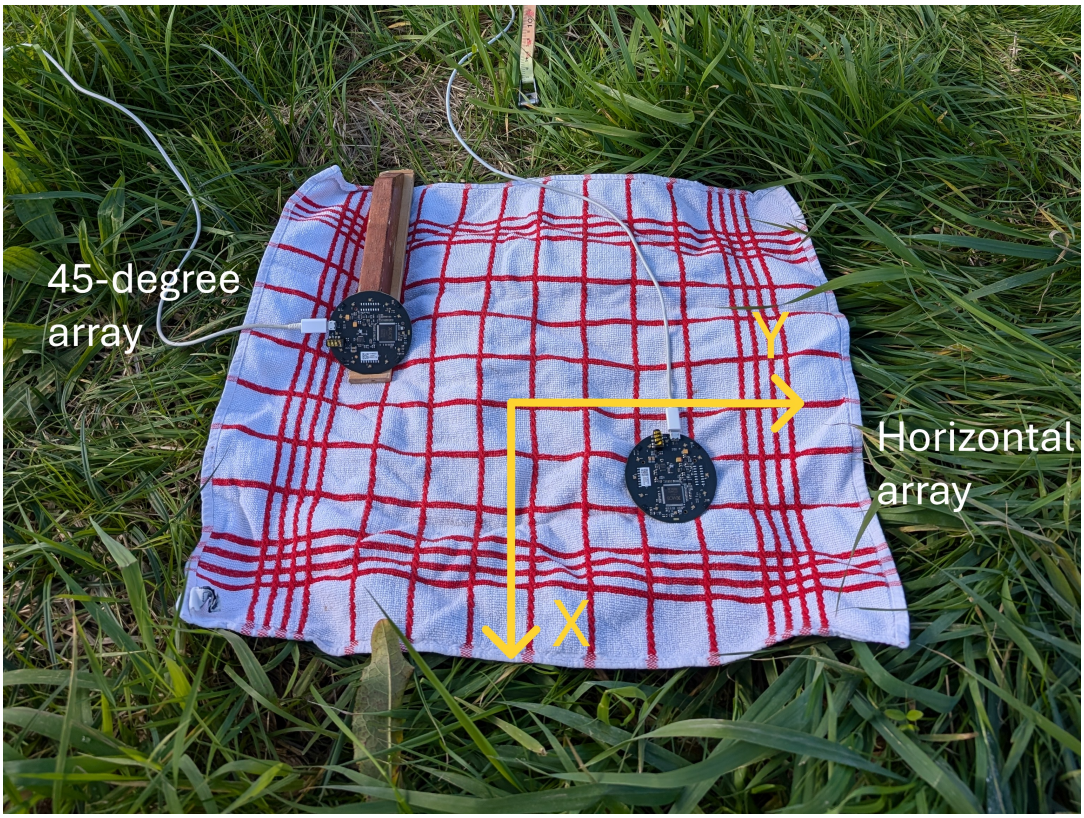


Figure 6.3: Microphone array setup

Figure 6.4: Satellite overview to determine array orientation

6.2 Parrot Bebop

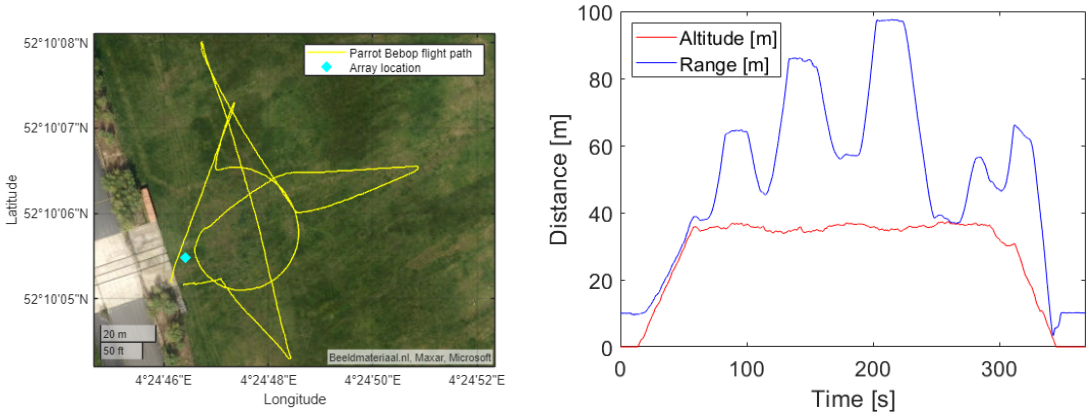
The specifics of the flight with the Parrot Bebop drone are visualized with several plots. This information is then used to apply beamforming in section 6.2.1. Applying beamforming over multiple sequential snapshots can be visualized with tracking plots showing elevation versus time and azimuth versus time, shown in section 6.2.2. The tracking plots shown are all with conventional frequency domain beamforming applied to measurements with the 45-degree array. The other orientation and beamforming options were calculated, but their results are only shown in section 6.2.3, which contains boxplots visualizing the error between the GPS and the acoustic localization for all options.

6.2.1 Flight details and beamform range

Figure 6.5 shows the flight path, altitude and range of the flight for an overview of the relative angles and distances. The spectrograms in figure 6.6 show the measured sound levels during the flight of the Parrot Bebop. The spectrogram is cut into three parts to show the noise levels at the higher frequencies. As those are mostly around 0dB they would be a dark blue if plotted in same graph as the sound levels in figure 6.6e. The spectrogram of the indoor drone flight in chapter 5 showed visible harmonics throughout the spectrum, but there was no wind indoors and the drone was much closer to the array.

The -20log(Range) and OSPL plots in figure 6.6b, d and f show the relation between the measured sound level and the distance of the drone. The trends should align. If the drone is closer to the array, indicated by a higher black line, the sound level should increase, which is most notably the case in 6.6d. If the trends do not match at all, the sound levels are likely not due to the drone and there is interference from a different noise source. This does not appear to be the case here. The lower sound levels near the edges of the plot are likely due to the drone not needing as much thrust, producing lower noise levels and shifting the BPF to very low frequencies.

Figure 6.5a shows that the drone performed multiple maneuvers. Stopping and turning requires a different amount of thrust than flying in a straight line. It was also quite windy and gusty in the open field located about 3.5km from the North Sea, affecting the amount of thrust required to maintain position. Varying the thrust means changing the rotation speed of the rotors, thus



(a) Satellite overview

(b) Altitude and range

Figure 6.5: Parrot Bebop flight details

changing the BPF, as shown in equation 2.1. This means that the harmonics cover a wider frequency band. The first four harmonics in figure 6.6e cover frequency bands 100s of Hz wide as opposed to the first four harmonics in figure 5.2 in chapter 5 that covered very narrow bands.

As the first four harmonics of the BPF cover wide frequency bands, beamforming is applied to the whole $300 - 1800$Hz range to cover the entire harmonic frequency bands over time. The spectrograms in figure 6.6a and figure 6.6c show the merged higher harmonics. Beamforming is also applied to the $2000 - 5000$Hz band and the $5000 - 7500$Hz bands. Beamforming was tested on the $5500 - 7000$Hz band but the result had some distortion from grating lobes. The functional beamforming version of this tracking plot looked similar.

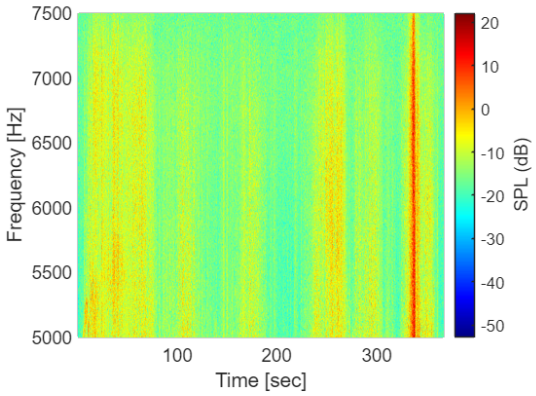
6.2.2 Flight tracking plot

The angle locations of the maximums of the beamforming result of every snapshot are plotted as a dot in figures 6.7, 6.8 and 6.9. The red line in these figures is the expected GPS reference angle. The yellow line is the range, which has an inverse relation with the red GPS line in the elevation plot. Flying farther away leads to a lower relative elevation angle.

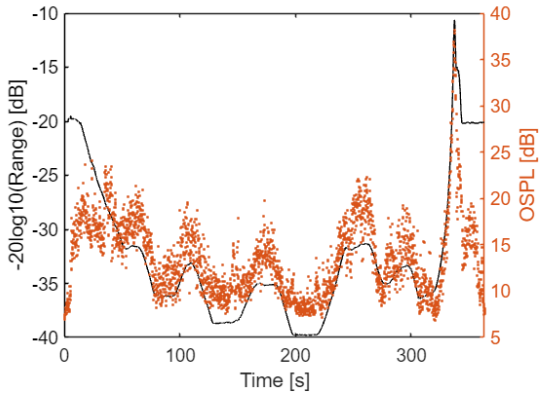
In general, the black dots representing the acoustic localization and the red line representing the reference GPS data are mostly in similar locations, though there is some offset and there are some outliers. There are anomalies before the 50s time range and near the 250s time range. The deviations at the start could potentially be explained by the strong wind gusts visible in figure 6.6e at that time, spatially displacing the noise.

6.2.3 Localization error boxplots

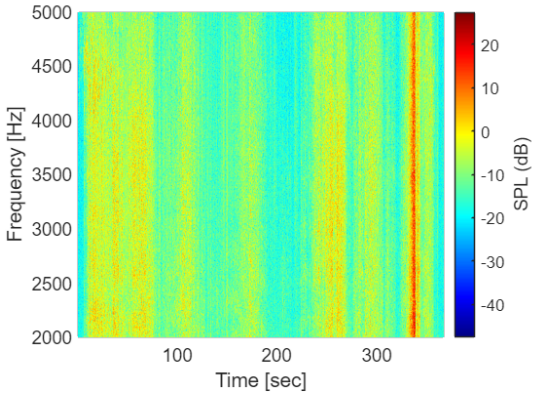
The values of the GPS angle and the beamformed angle are subtracted to obtain an error value to quantify the accuracy of the localization. The localization is accurate if this error is small. It is visualized with boxplots in figures 6.10 and 6.11. These boxplots contain the error values for all combinations of array orientation, frequency range selection and beamforming method. The boxplots matching the tracking plots in the previous section are in figure 6.11. Except for the red boxes as the functional beamforming tracking plot was omitted. The boxplots are grouped by distance between the drone and the array, as the distance is a major factor in sound level. The horizontal bars of each box indicate the 25th, 50th and 75th percentile of the error values. The red plus signs represent outliers. The elevation error of the boxes representing the closest distance range is large because the drone had some issues recording altitudes below sea level.



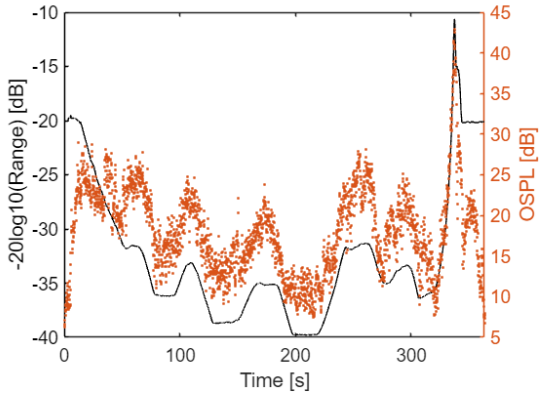
(a) Spectrogram high frequency beamform range



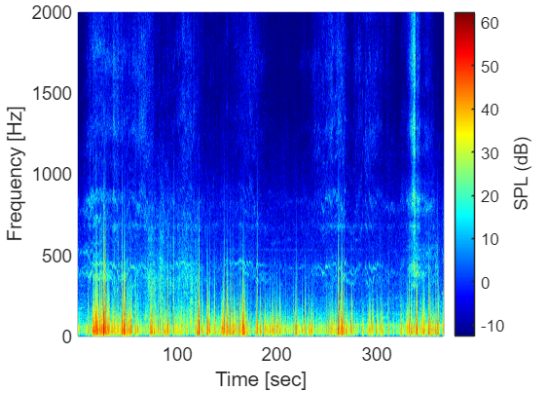
(b) -20log(Range) & OSPL 5000-7500Hz



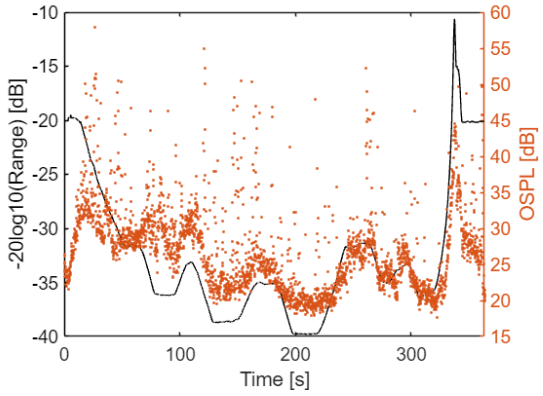
(c) Spectrogram intermediate frequency beamform range



(d) -20log(Range) & OSPL 2000-5000Hz

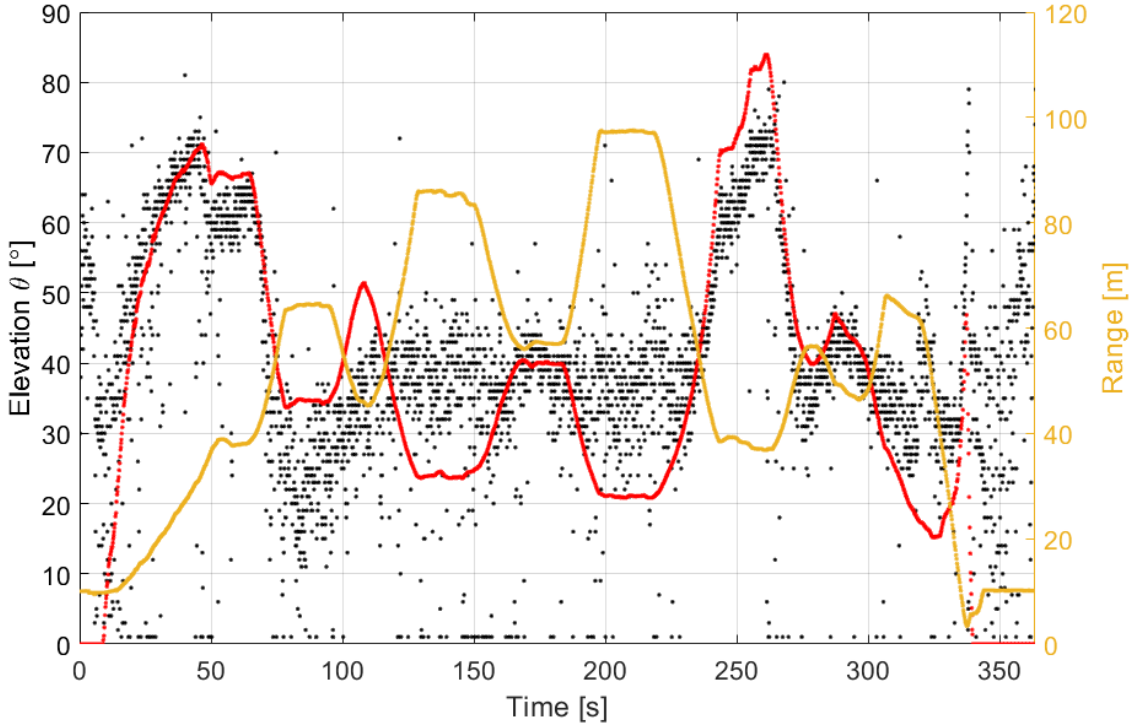


(e) Spectrogram low frequency beamform range

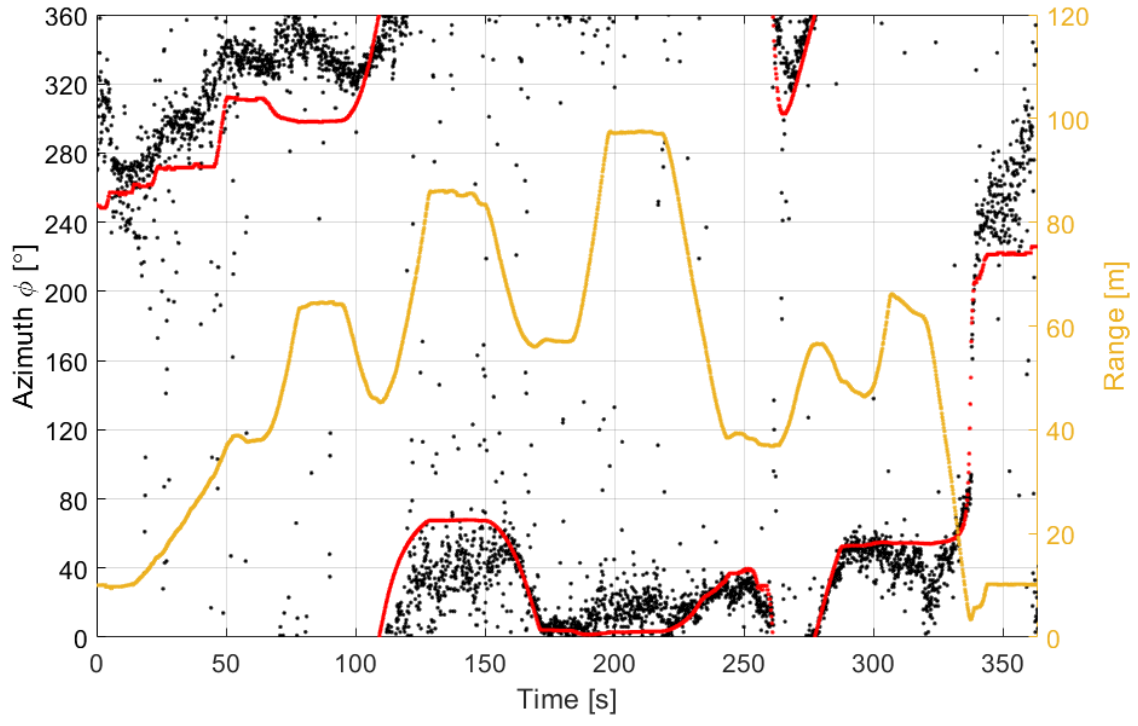


(f) -20log(Range) & OSPL 300-1800Hz

Figure 6.6: Parrot Bebop flight noise details

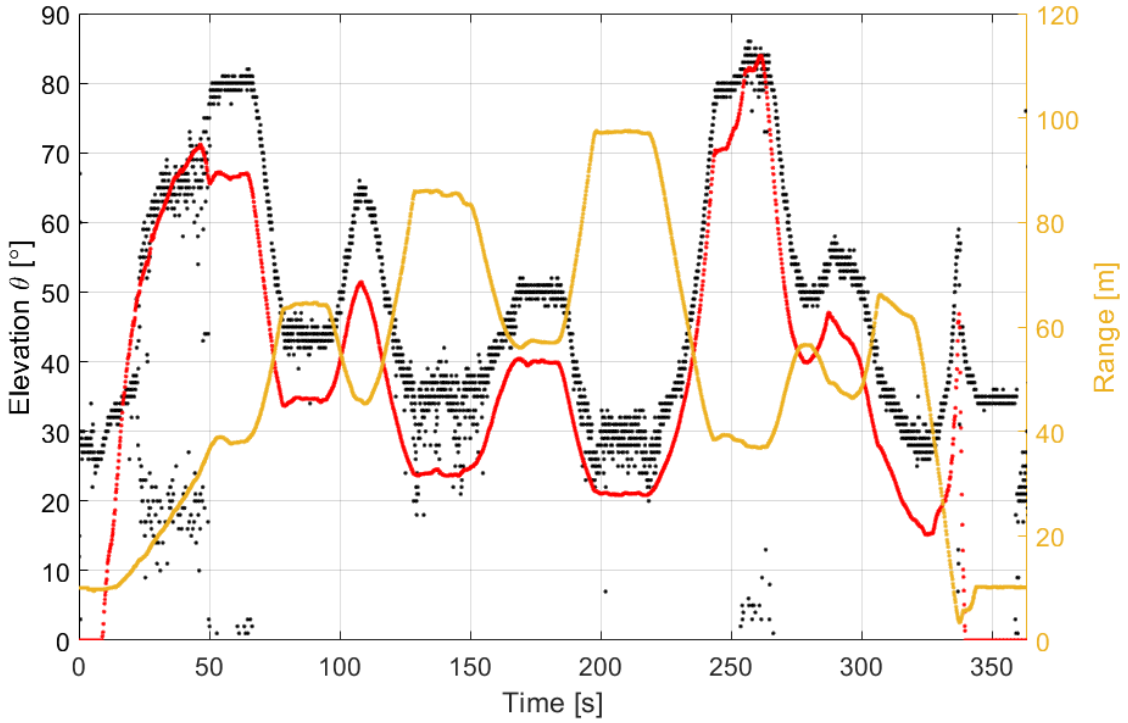


(a) Elevation angle

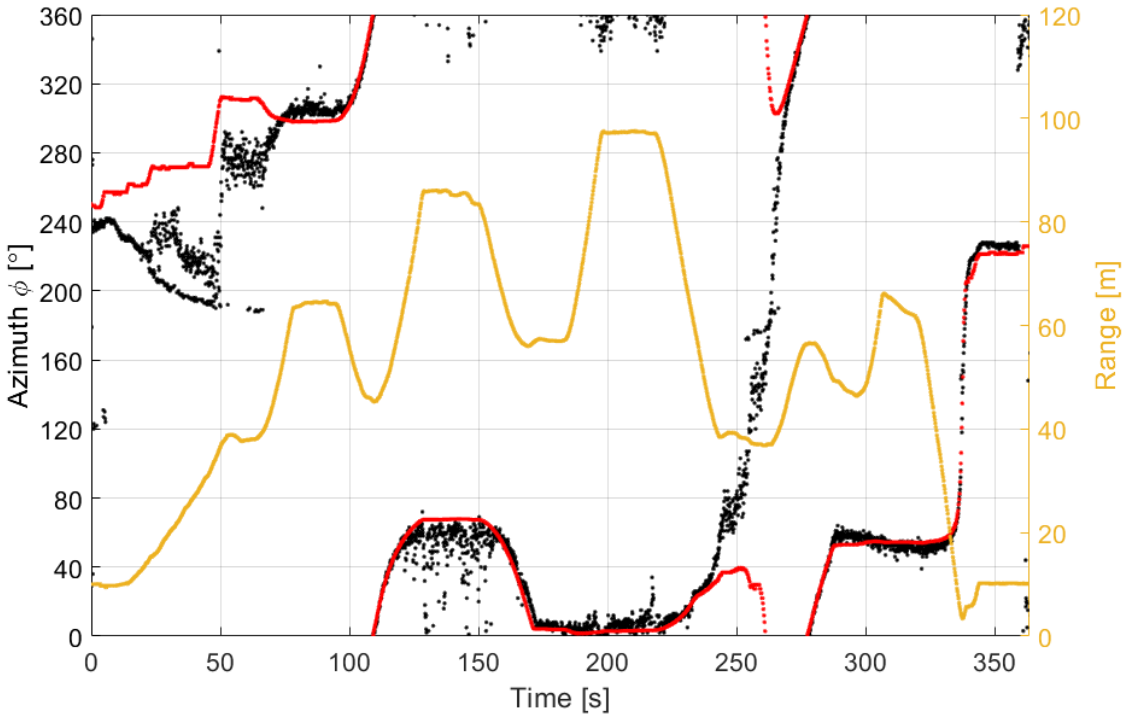


(b) Azimuth angle

Figure 6.7: Beamformed snapshots of an outdoor Parrot Bebop flyby measured with the 45 degree array. Conventional frequency domain beamforming applied to a frequency range of 300-1800Hz. The black dots indicate the snapshot beamform maximum, the red line is the GPS reference location and the yellow line is the distance between the drone and the array.

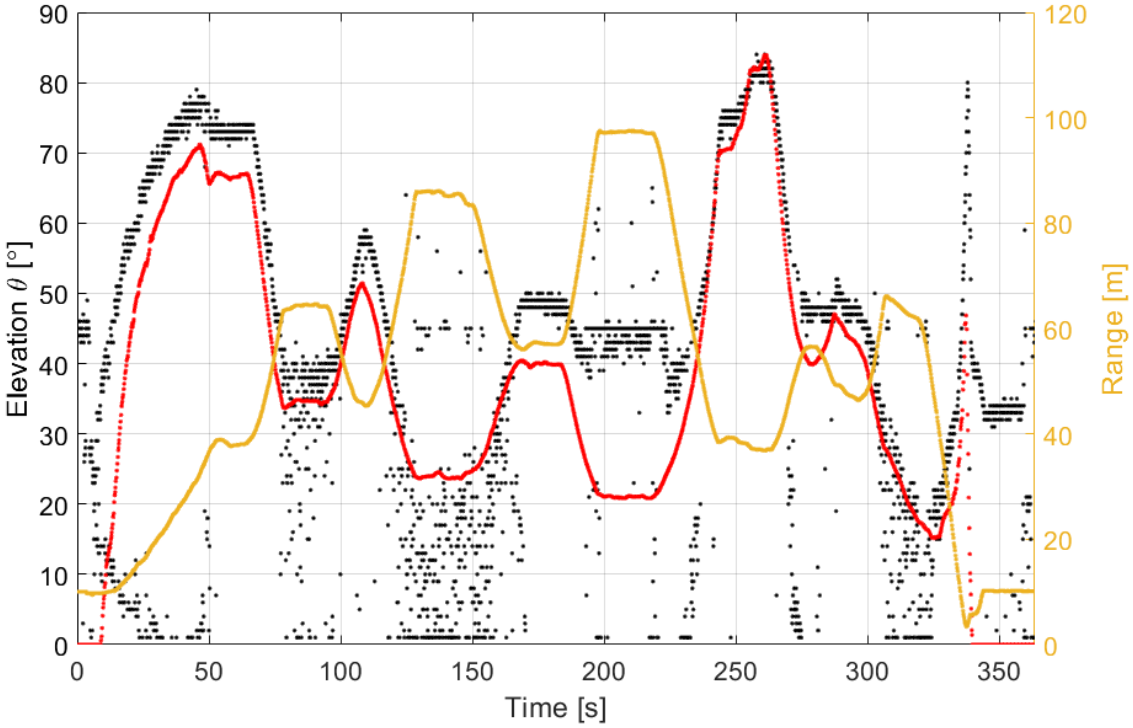


(a) Elevation angle

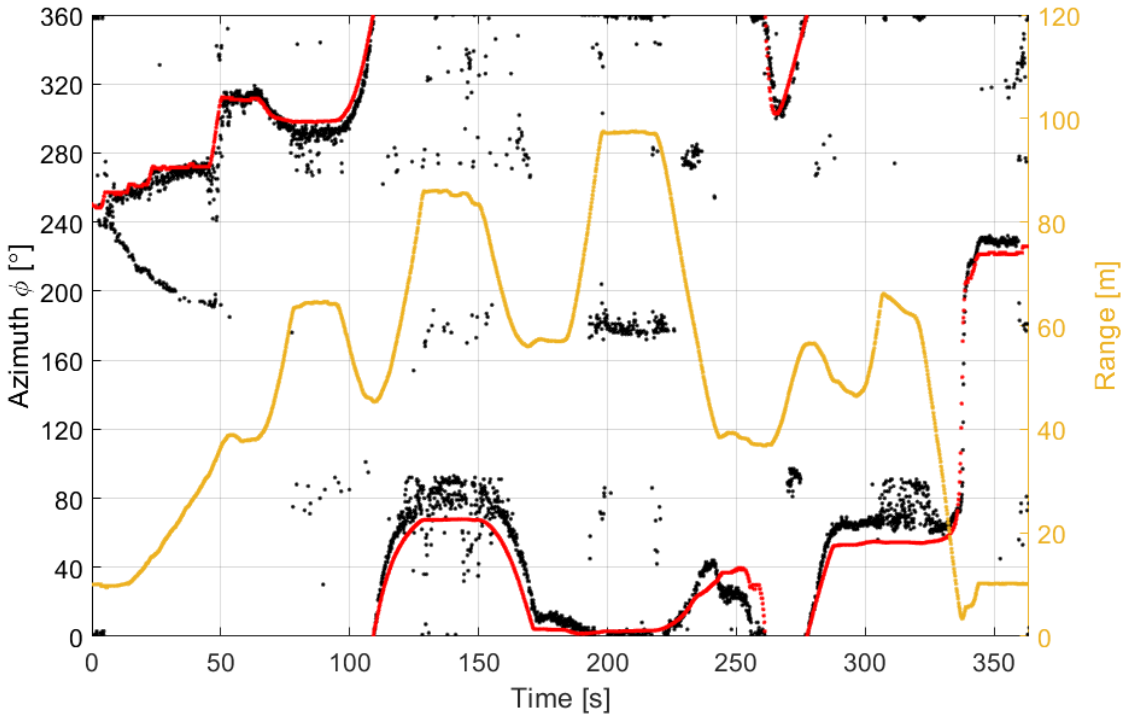


(b) Azimuth angle

Figure 6.8: Beamformed snapshots of an outdoor Parrot Bebop flyby measured with the 45 degree array. Conventional frequency domain beamforming applied to a frequency range of 2000-5000Hz. The black dots indicate the snapshot beamform maximum, the red line is the GPS reference location and the yellow line is the distance between the drone and the array.

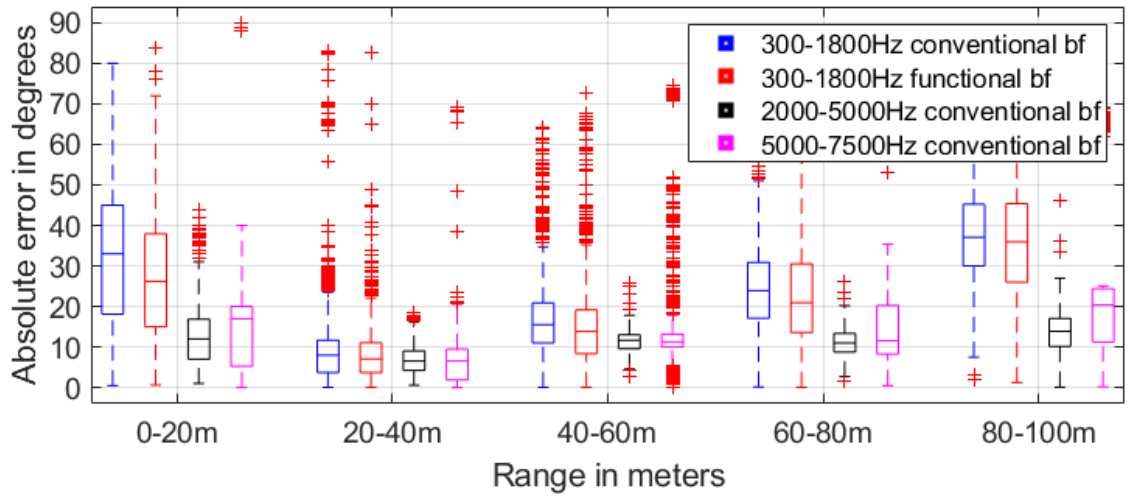


(a) Elevation angle

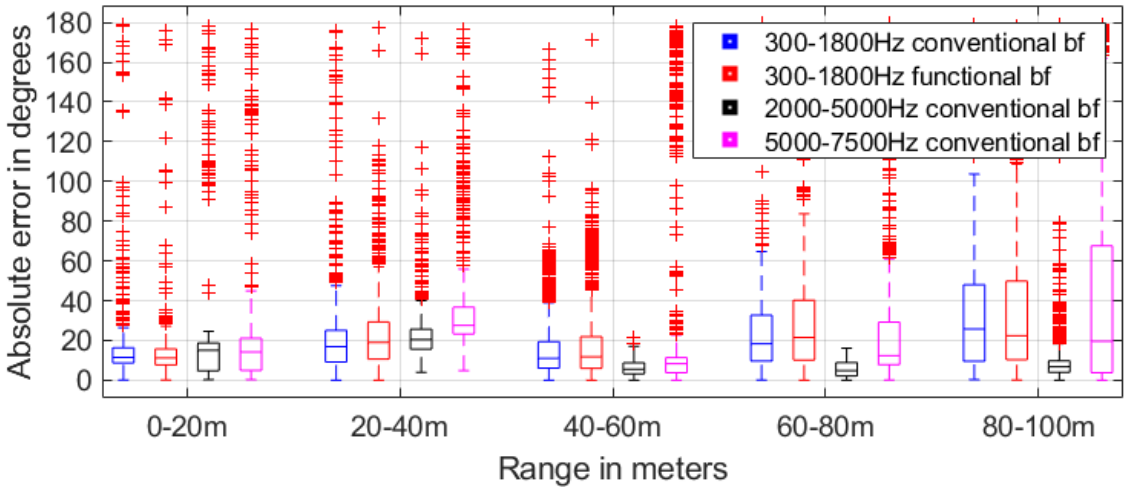


(b) Azimuth angle

Figure 6.9: Beamformed snapshots of an outdoor Parrot Bebop flyby measured with the 45 degree array. Conventional frequency domain beamforming applied to a frequency range of 5000-7500Hz. The black dots indicate the snapshot beamform maximum, the red line is the GPS reference location and the yellow line is the distance between the drone and the array.

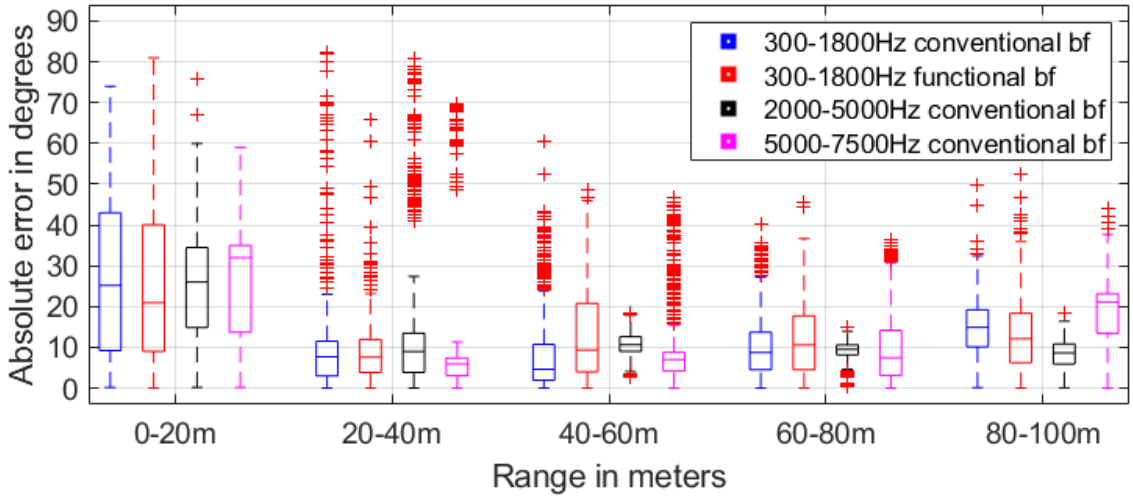


(a) Parrot Bebop elevation localization error horizontal array

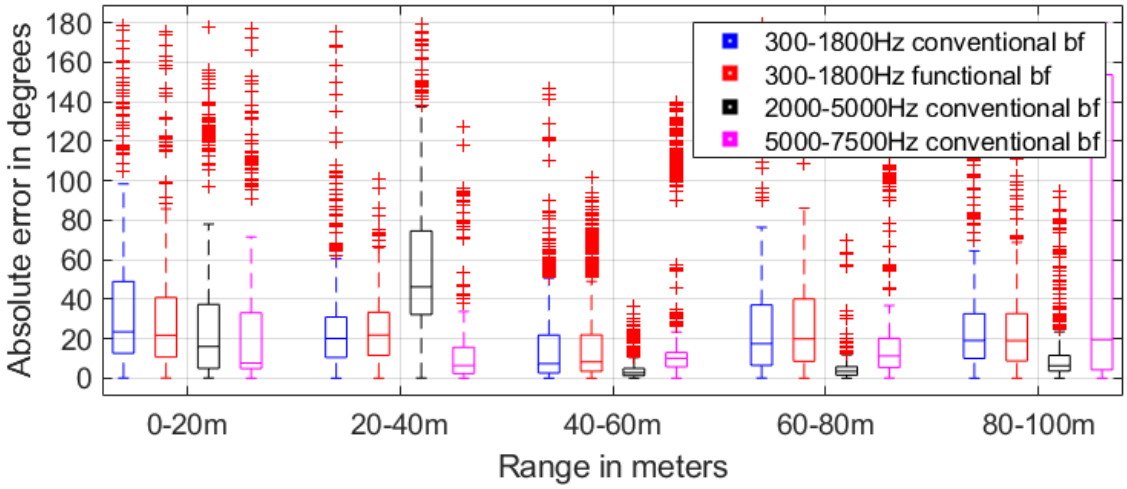


(b) Parrot Bebop azimuth localization error horizontal array

Figure 6.10: Boxplots indicating the error between the reference GPS signal and the maximum locations obtained with beamforming measurements with the horizontal array. The blue boxes represent the error values of the low frequency range with conventional beamforming. The red boxes represent the error values of the same low frequency range with functional beamforming. The black boxes are the intermediate frequency range and the magenta boxes the high frequency range.

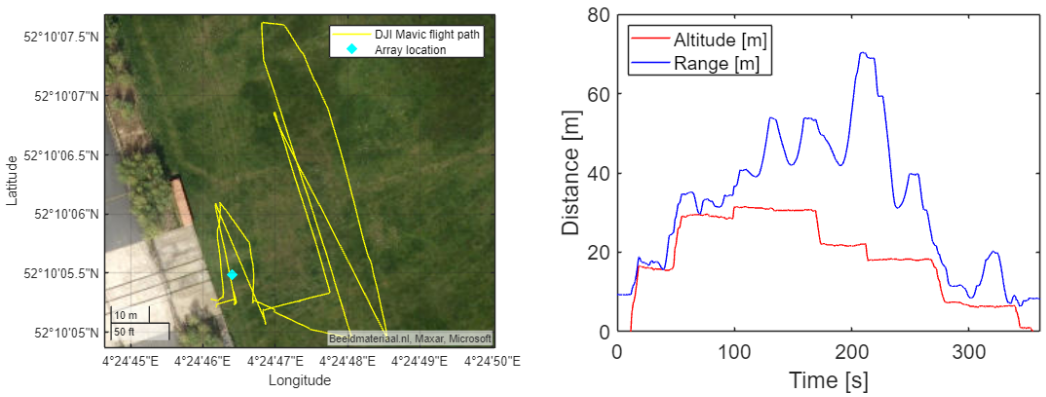


(a) Parrot Bebop elevation localization error 45-degree array



(b) Parrot Bebop azimuth localization error 45-degree array

Figure 6.11: Boxplots indicating the error between the reference GPS signal and the maximum locations obtained with beamforming measurements with the 45-degree array. The blue boxes represent the error values of the low frequency range with conventional beamforming. The red boxes represent the error values of the same low frequency range with functional beamforming. The black boxes are the intermediate frequency range and the magenta boxes the high frequency range.



(a) Satellite overview

(b) Altitude and range

Figure 6.12: DJI Mavic flight details

6.3 DJI Mavic

The setup of this section is the same as the previous section. But here the results are shown for conventional beamforming applied to measurement with the horizontal array. To avoid restating the same information, the figures containing the data are referred to directly here.

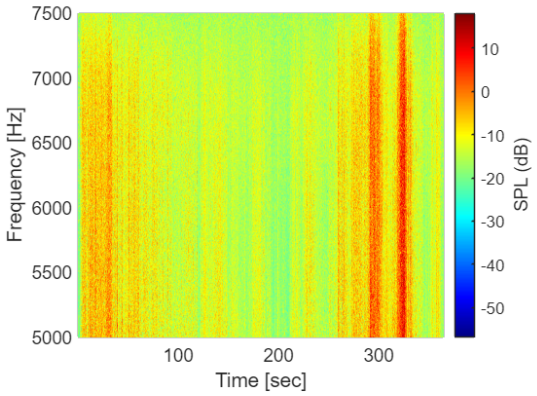
The harmonics in the spectogram in figure 6.13e look slightly different from those of the Parrot Bebop in figure 6.6e. This is because the DJI Mavic propellers only have two blades as opposed to the three-bladed Parrot Bebop propellers. As mentioned in equation 2.1, the BPF increases with the number of blades. The BPF harmonics of the DJI Mavic therefore start at a lower frequency and repeat with smaller frequency intervals. The first harmonic appears around 140Hz, but it is mostly hidden by background noise. This flight also came with several maneuvers, as seen in the satellite overview in figure 6.12a. The resulting changes in thrust, rpm and BPF lead to harmonics that vary over 100Hz with time. Therefore beamforming is again applied over a wide band, this time over $250 - 1750$Hz. Starting slightly lower to include a lower harmonic but covering the same 1.5kHz range as the low frequency range plots in the Parrot Bebop section. The intermediate and high frequency bands are the same at $2000 - 5000$Hz and $5000 - 7500$Hz. The OSPL and range plots in figure 6.13 again show outliers at the lower frequencies in 6.13f but a good match between -20log(Range) and OSPL. The drone is the dominant noise source.

The tracking plots of the DJI Mavic flight covering conventional frequency domain beamforming applied to measurements with the horizontal array are shown in figures 6.14, 6.15 and 6.16.

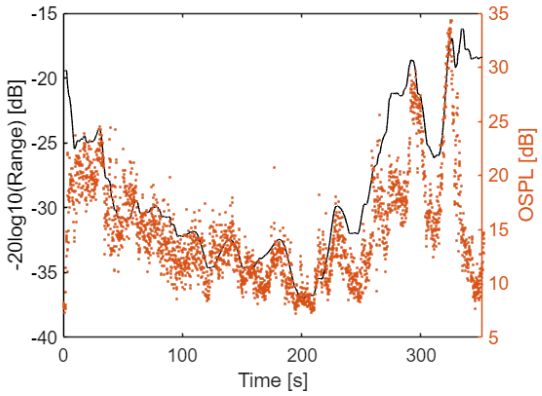
The boxplots indicating the localization error are shown in figures 6.17 and 6.18.

6.4 Second Parrot Bebop flight with interference

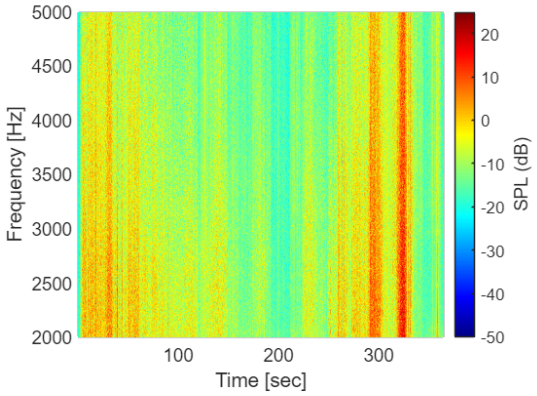
The setup of this section is similar to the previous sections. Here the tracking plots are shown for the measurement with the horizontal array of the second Parrot Bebop flight. First the flight details are shown in figure 6.19. Because there was interference from several aircraft flyovers, on approach to Schiphol, the lower frequency range is not considered for beamforming. Only conventional beamforming is applied to a wide frequency range of 3000Hz-7000Hz. The resulting tracking plot is shown in figure 6.20. The localization error between the GPS signal and the beamform output is visualized with boxplots in figure 6.21.



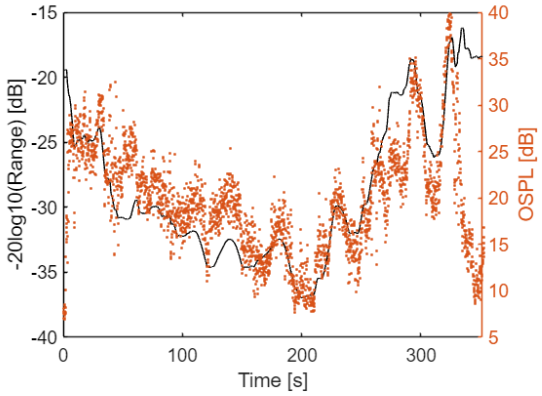
(a) Spectrogram high frequency beamform range



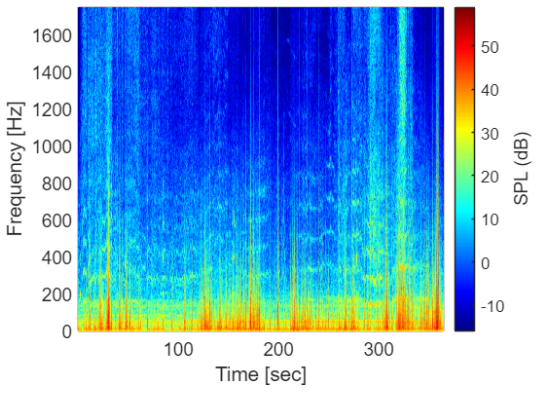
(b) -20log(Range) & OSPL 5000-7500Hz



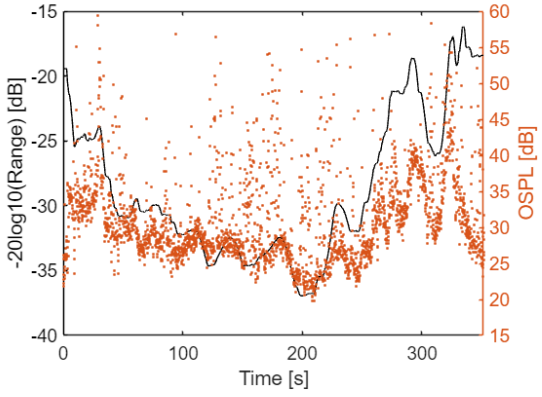
(c) Spectrogram intermediate frequency beamform range



(d) -20log(Range) & OSPL 2000-5000Hz

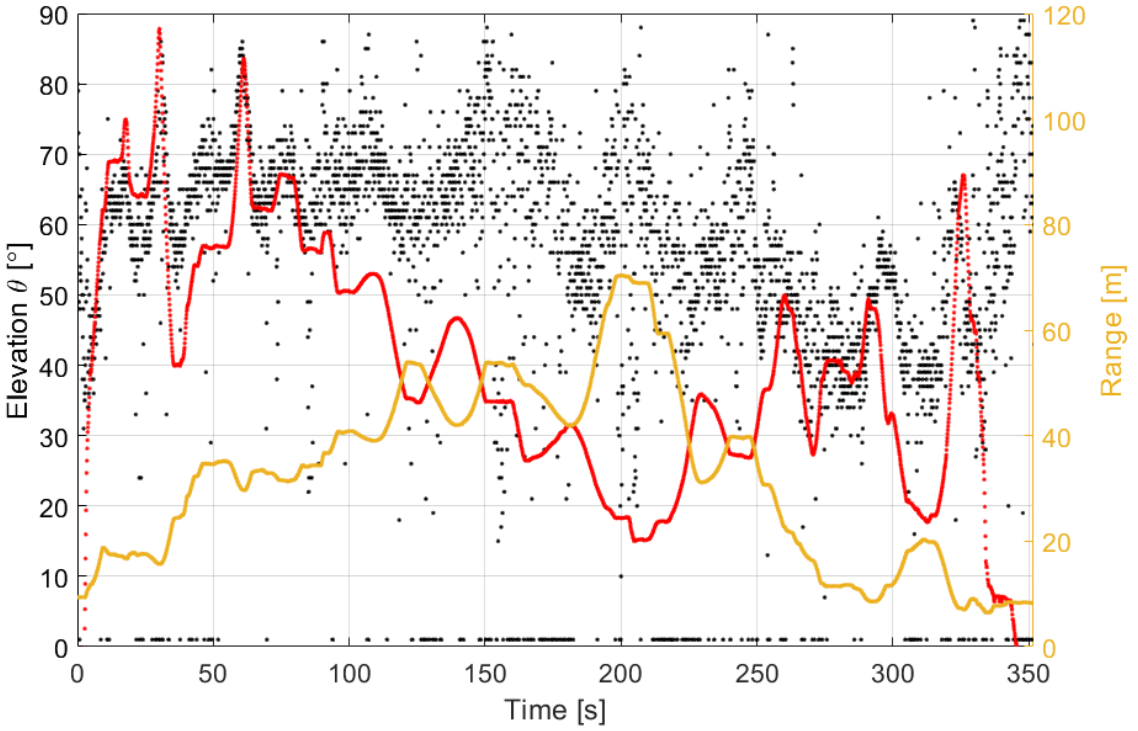


(e) Spectrogram low frequency beamform range

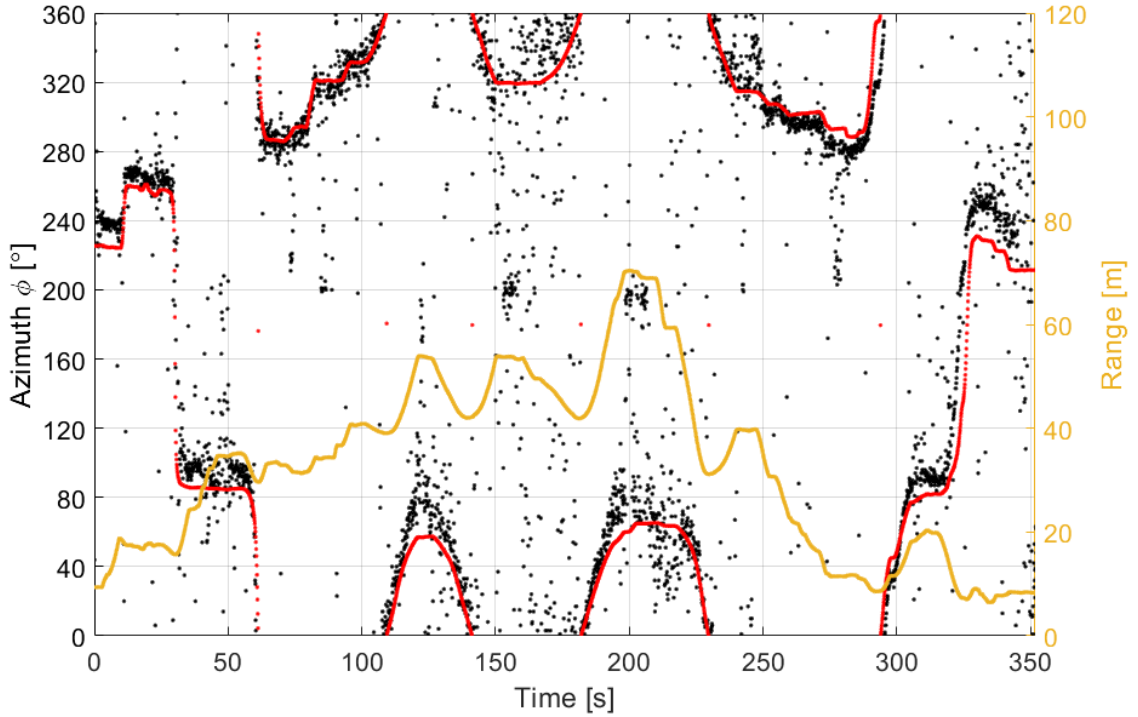


(f) -20log(Range) & OSPL 250-1750Hz

Figure 6.13: DJI Mavic flight noise details

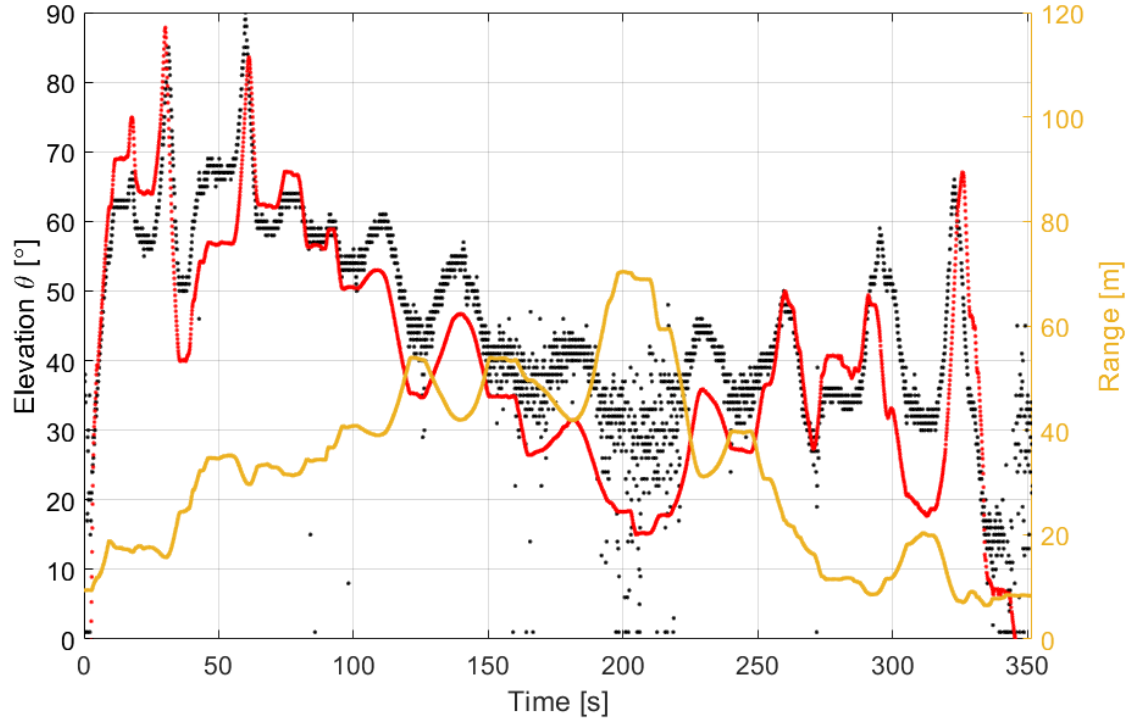


(a) Elevation angle

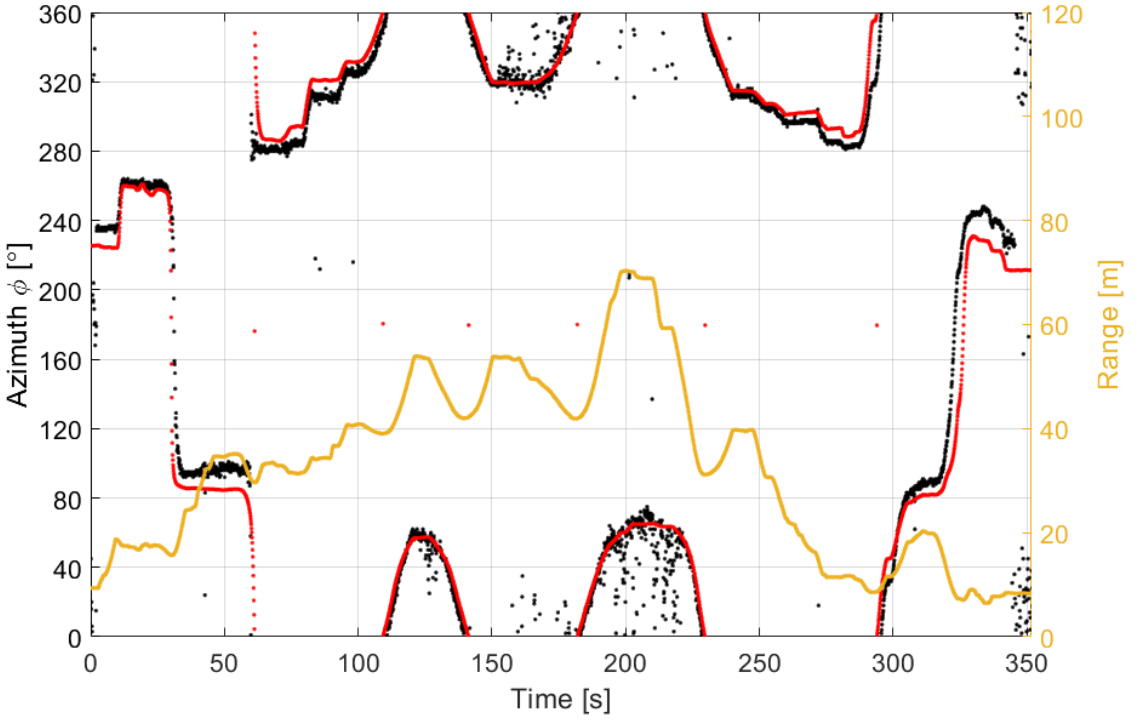


(b) Azimuth angle

Figure 6.14: Beamformed snapshots of an outdoor DJI Mavic flyby measured with the horizontal array. Conventional frequency domain beamforming applied to a frequency range of 250-1750Hz. The black dots indicate the snapshot beamform maximum, the red line is the GPS reference location and the yellow line is the distance between the drone and the array.

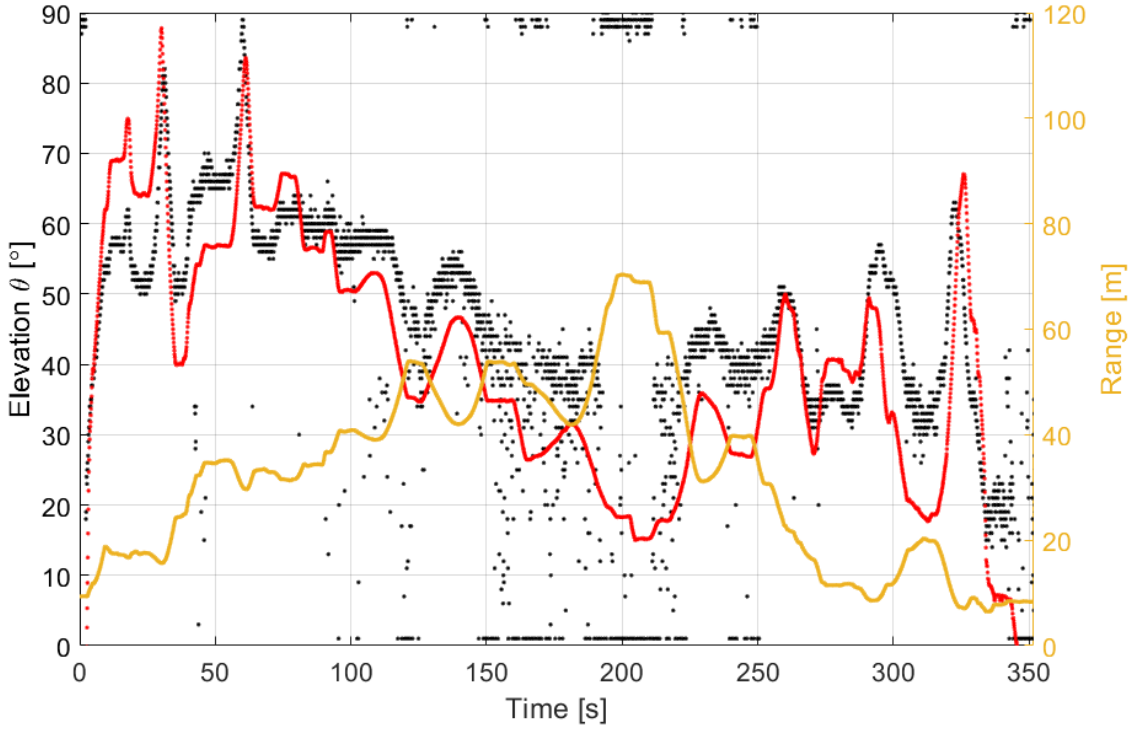


(a) Elevation angle

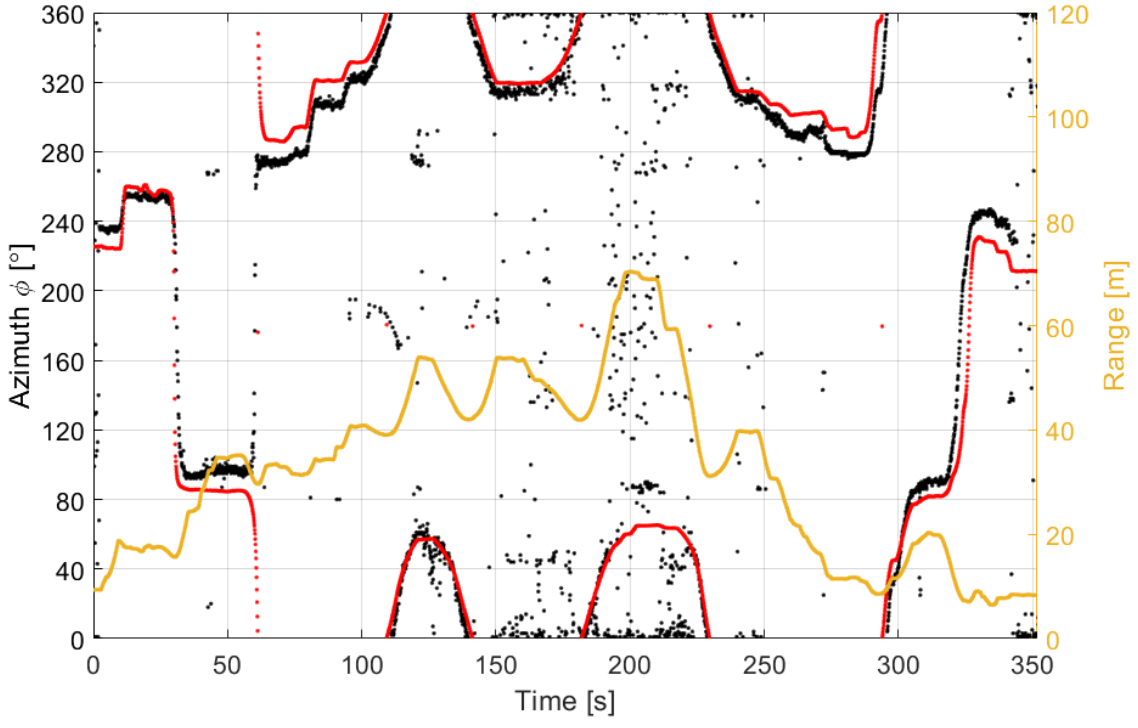


(b) Azimuth angle

Figure 6.15: Beamformed snapshots of an outdoor DJI Mavic flyby measured with the horizontal array. Conventional frequency domain beamforming applied to a frequency range of 2000-5000Hz. The black dots indicate the snapshot beamform maximum, the red line is the GPS reference location and the yellow line is the distance between the drone and the array.

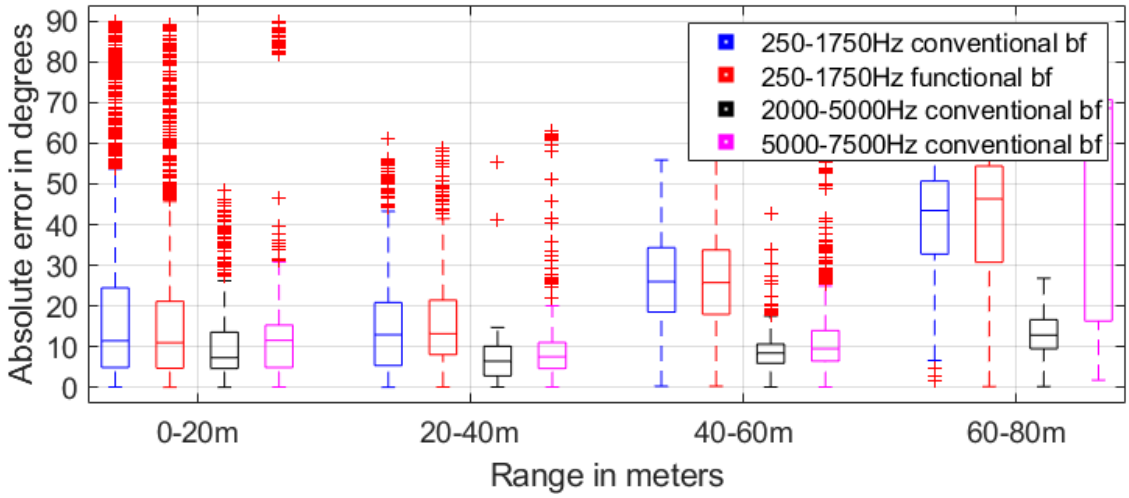


(a) Elevation angle

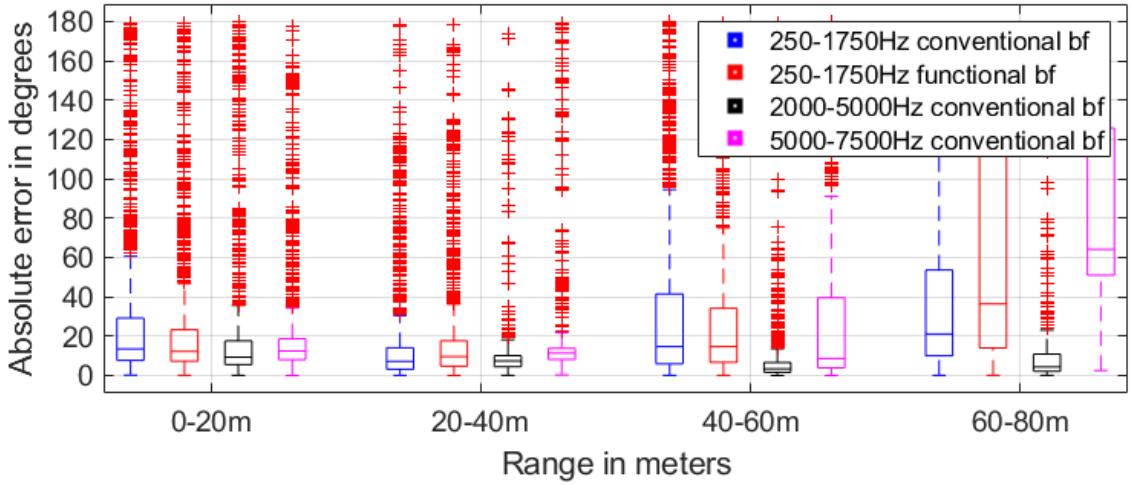


(b) Azimuth angle

Figure 6.16: Beamformed snapshots of an outdoor DJI Mavic flyby measured with the horizontal array. Conventional frequency domain beamforming applied to a frequency range of 5000-7500Hz. The black dots indicate the snapshot beamform maximum, the red line is the GPS reference location and the yellow line is the distance between the drone and the array.

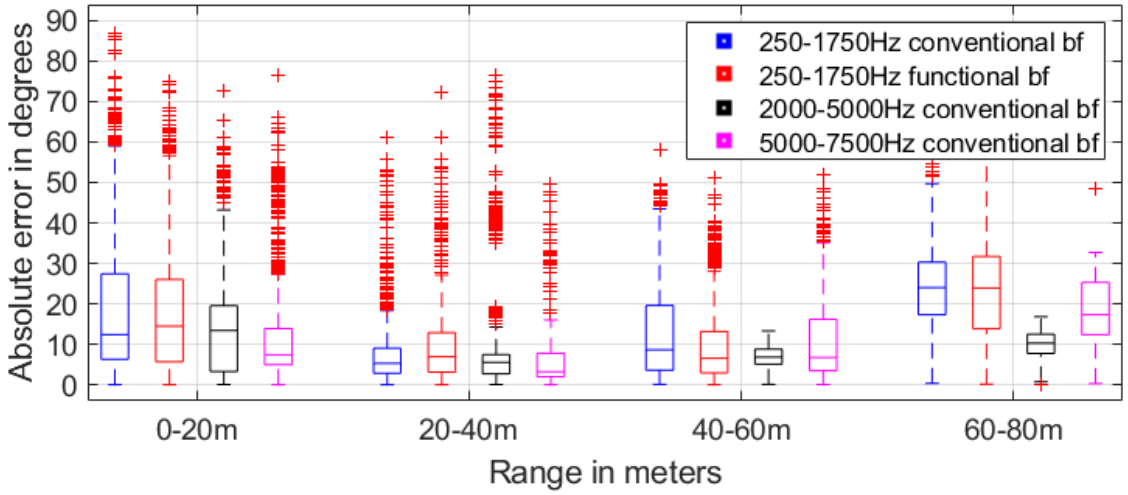


(a) DJI Mavic elevation localization error horizontal array

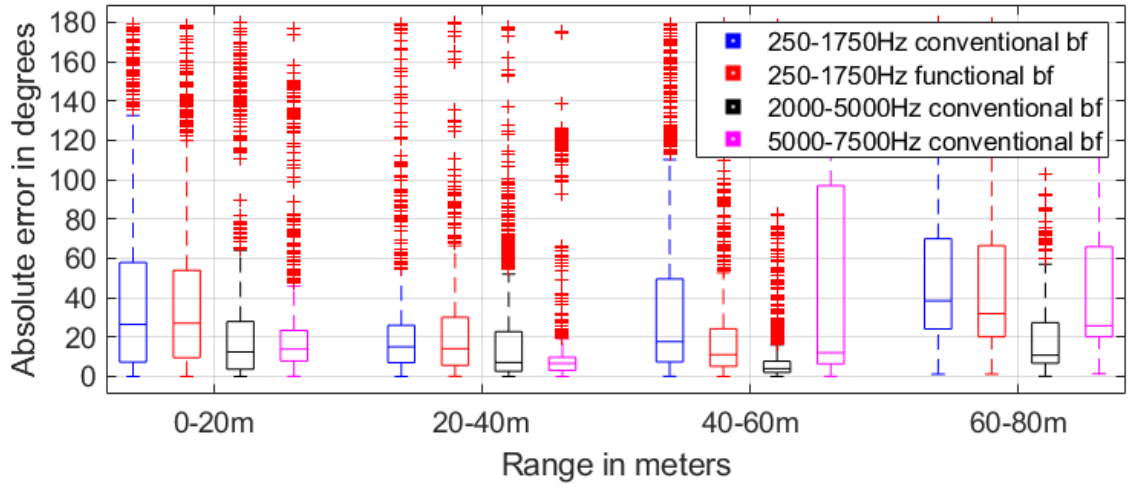


(b) DJI Mavic azimuth localization error horizontal array

Figure 6.17: Boxplots indicating the error between the reference GPS signal and the maximum locations obtained with beamforming measurements with the horizontal array. The blue boxes represent the error values of the low frequency range with conventional beamforming. The red boxes represent the error values of the same low frequency range with functional beamforming. The black boxes are the intermediate frequency range and the magenta boxes the high frequency range.



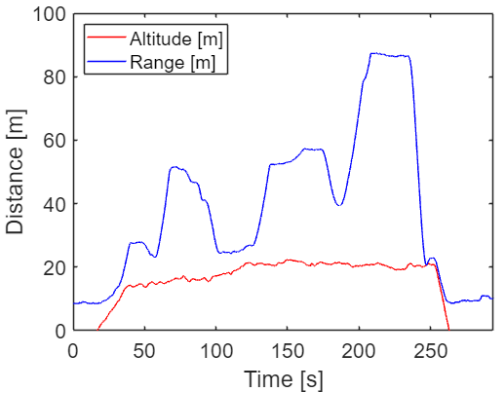
(a) DJI Mavic elevation localization error 45-degree array



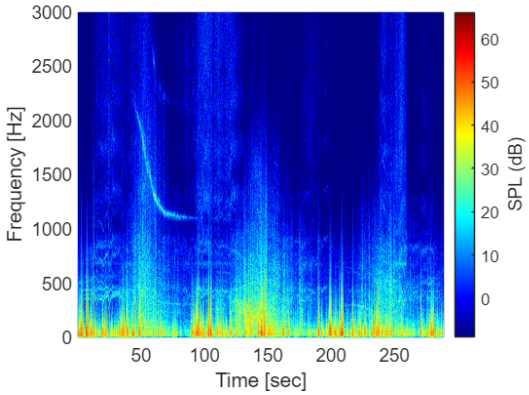
(b) DJI Mavic azimuth localization error 45-degree array

Figure 6.18: Boxplots indicating the error between the reference GPS signal and the maximum locations obtained with beamforming measurements with the 45-degree array. The blue boxes represent the error values of the low frequency range with conventional beamforming. The red boxes represent the error values of the same low frequency range with functional beamforming. The black boxes are the intermediate frequency range and the magenta boxes the high frequency range.

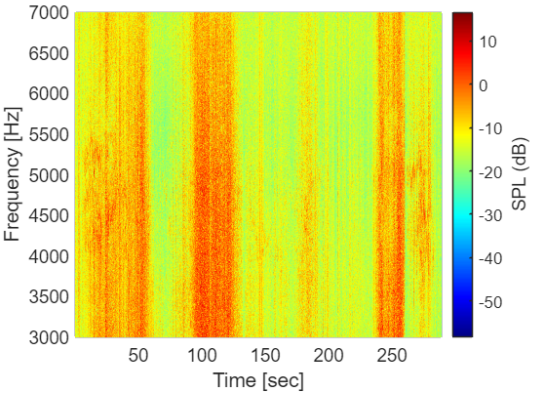
(a) Satellite overview



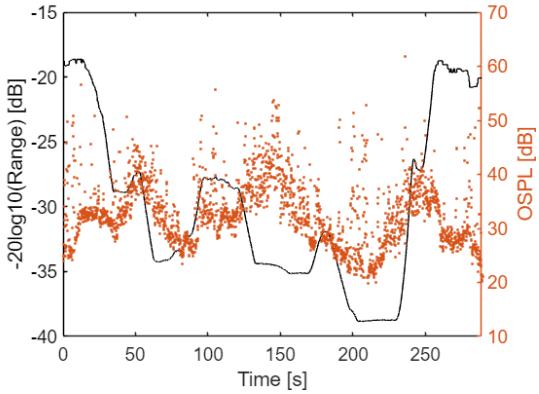
(b) Altitude and range



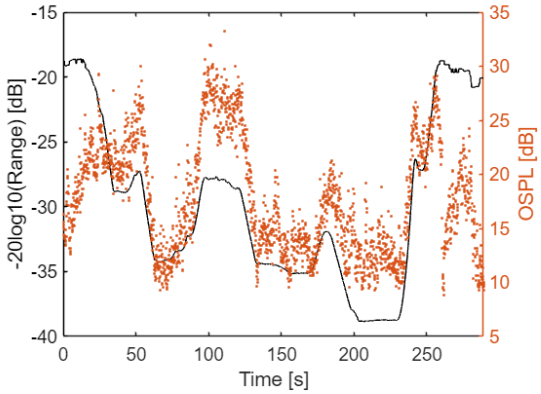
(c) Spectrogram interference region



(d) Spectrogram beamformed range

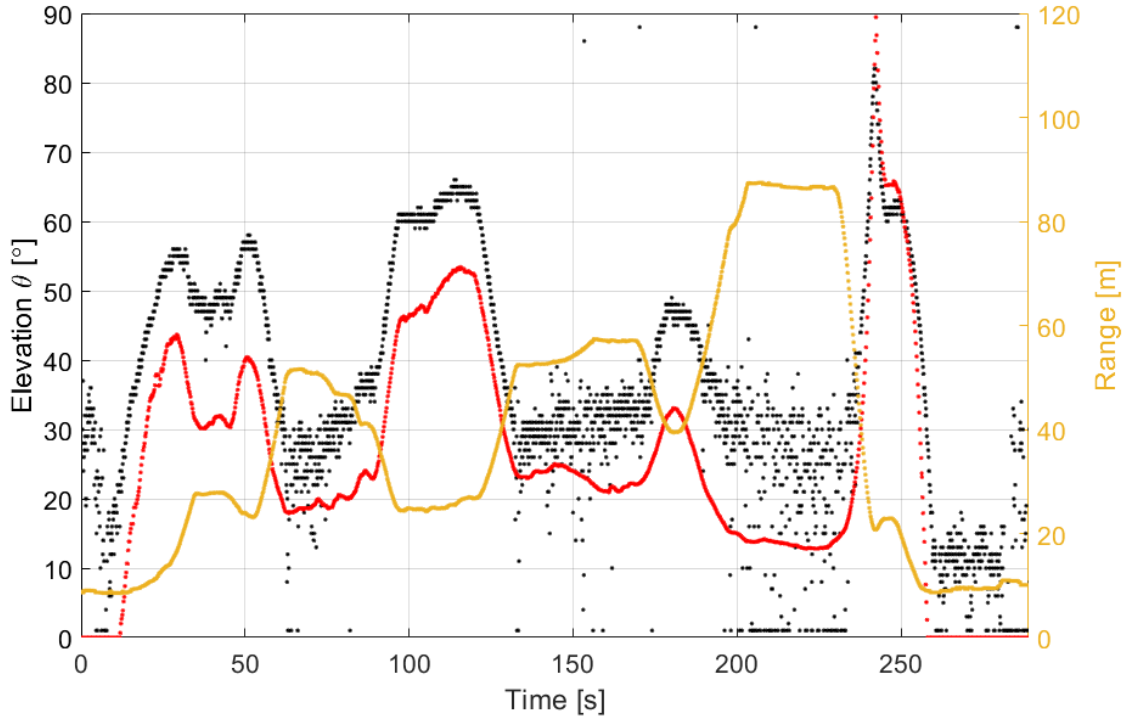


(e) -20log(Range) & OSPL 300-3000Hz

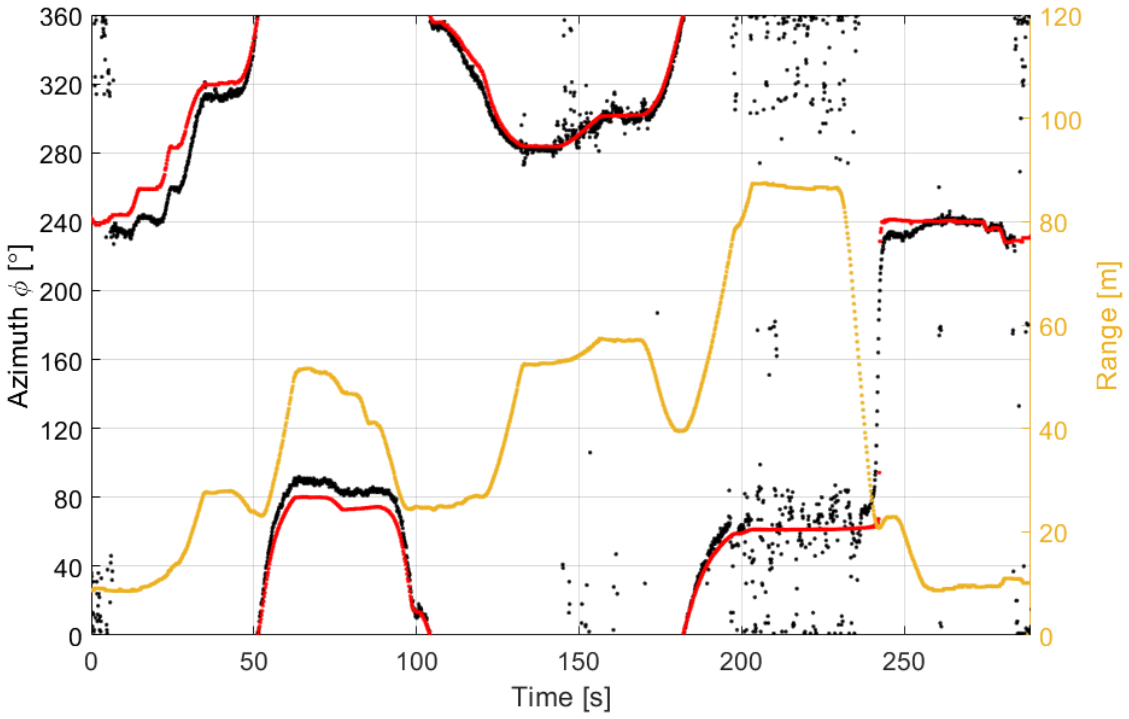


(f) -20log(Range) & OSPL 3000-7000Hz

Figure 6.19: Flight details Parrot Bebop flight with interference

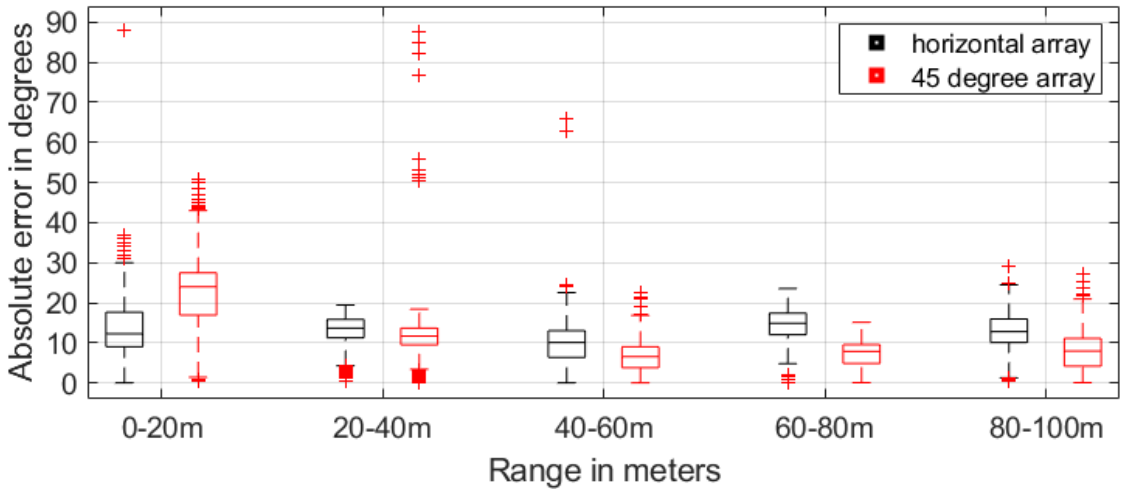


(a) Elevation angle

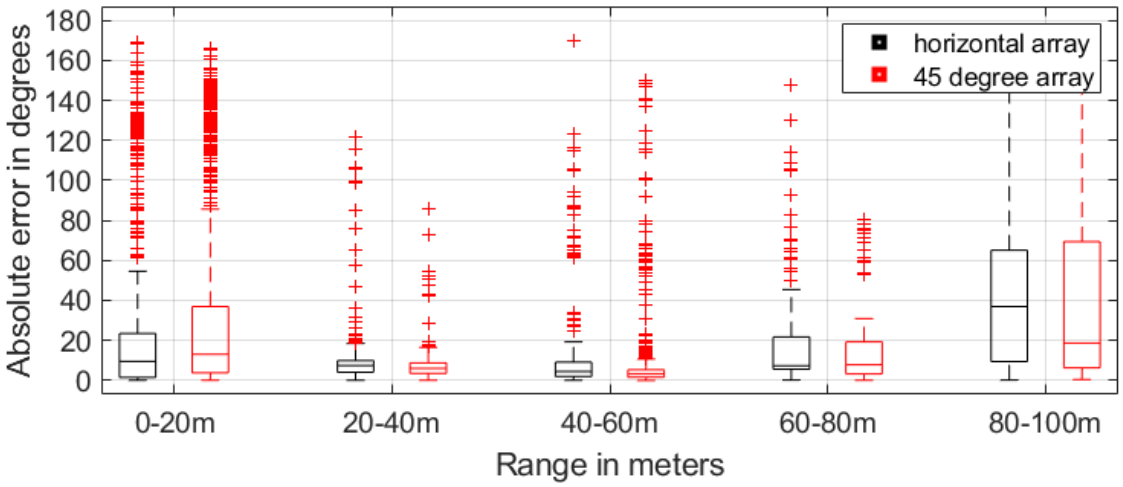


(b) Azimuth angle

Figure 6.20: Beamformed snapshots of an outdoor Parrot Bebop flyby with interference measured with the horizontal array. Conventional frequency domain beamforming applied to a frequency range of 3000-7000Hz. The black dots indicate the snapshot beamform maximum, the red line is the GPS reference location and the yellow line is the distance between the drone and the array.



(a) Parrot Bebop interference flight elevation localization error



(b) Parrot Bebop interference flight azimuth localization error

Figure 6.21: Boxplots indicating the error between the reference GPS signal and the maximum locations obtained with beamforming. The black boxes represent the error values from the measurement with the horizontal array and the red boxes represent the error values from the measurement with the 45-degree array.

Number of datapoints per distance range				
Distance range	Parrot Bebop	DJI Mavic	Parrot Bebop int	DJi Mavic int
0-20m	621	1189	702	929
20-40m	666	1119	834	988
40-60m	1072	1006	924	1576
60-80m	694	204	108	1535
80-100m	592	0	322	861
100-120m	0	0	0	128

Table 6.2: Number of datapoints per distance bin for each flight. Int refers to the two flights with low frequency interference in sections 6.4 and 6.5.

6.5 Second DJI Mavic flight with interference

The setup of this section is identical to the previous section. The tracking plots show the result of the 45-degree measurement of the second DJI Mavic flight. First the flight details are shown in figure 6.22. The lower frequency range is not considered for beamforming. Only conventional beamforming is applied to a wide frequency range of 3000Hz-7000Hz. The resulting tracking plot is shown in figure 6.23. The localization error between the GPS signal and the beamform output is visualized with boxplots in figure 6.24.

6.6 Comparison

The effects of the various factors in the previous sections are briefly compared to determine the best orientation and beamforming method in section 6.6.1. Afterwards a comparison with literature is made for a more general assessment of the use of the small array, in section 6.6.2.

6.6.1 Experiment result comparison

The error values representing the accuracy of the localization effort were visualized in boxplots in the previous sections. The values of the first distance range, 0-20m, are relatively large in all boxplots. This may be because the drones started below sealevel, which wasn't properly recorded in the GPS data. The true altitude may therefore be off by a meter. This disproportionally affects the nearby distance range, as 1 meter difference in altitude creates a large difference in elevation right next to the array.

The boxplots with four boxes in sections 6.2 and 6.3 show that the higher frequencies have the best accuracy. Particularly the black boxes of the $2000 - 5000$Hz range. This is likely because the sound levels in this range are higher than those in the $5000 - 7500$Hz range. The low frequencies ranges lead to poorer localization due to the large beamwidth associated with these frequencies when measuring with the ReSpeaker array.

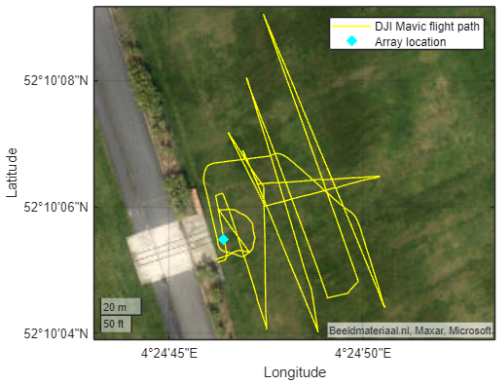
The boxplots of the second flights in sections 6.4 and 6.5 also show fairly good tracking up to 80m distance.

For all situations it appears that the elevation tracking benefits from placing the array at a 45-degree angle. The Parrot Bebop can be tracked up to almost 100m distance with an error below 10 degrees and the DJI Mavic can be tracked up to 80m distance with an error below 10 degrees.

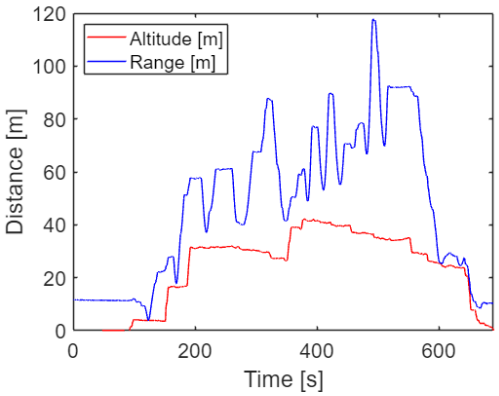
Table 6.2 shows the number of datapoints used in each distance bin, for a measure of reliability.

6.6.2 Comparison with literature

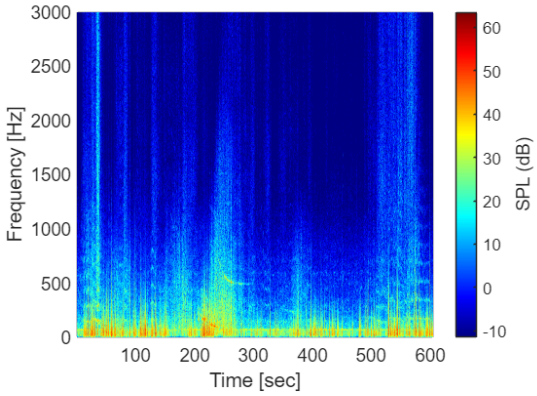
Boxplots of a similar experiment were found [21]. The localization accuracy of a larger microphone array with 31 mics is shown in figure 6.25. These values are more accurate but they correspond to a much larger DJI S900 drone, which weighs 3.3kg, has more propellers and likely produces



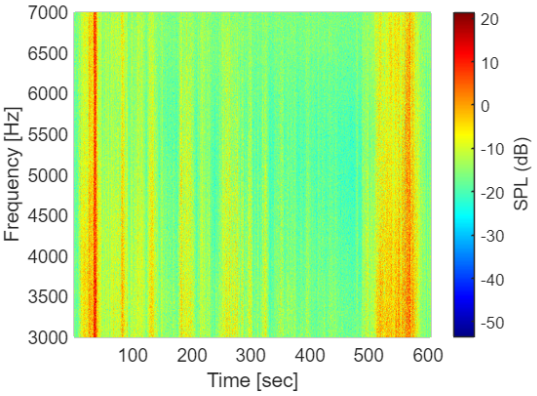
(a) Satellite overview



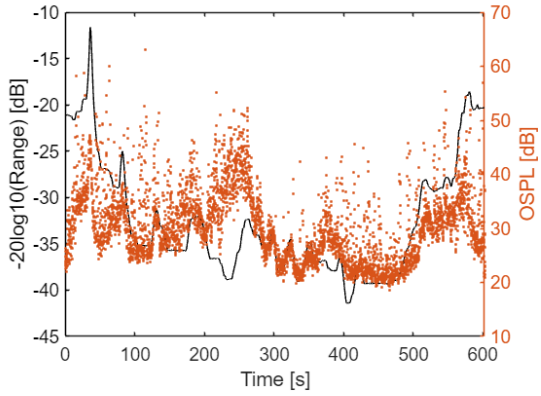
(b) Altitude and range



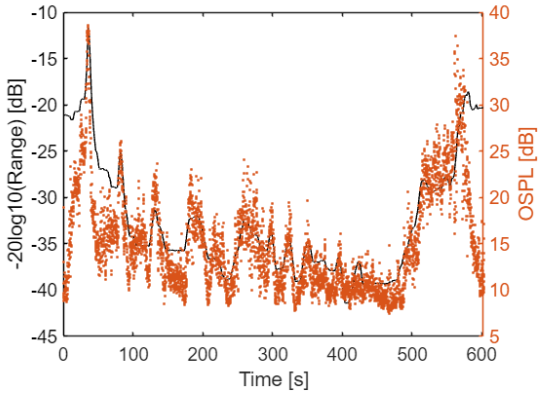
(c) Spectrogram interference region



(d) Spectrogram beamformed range

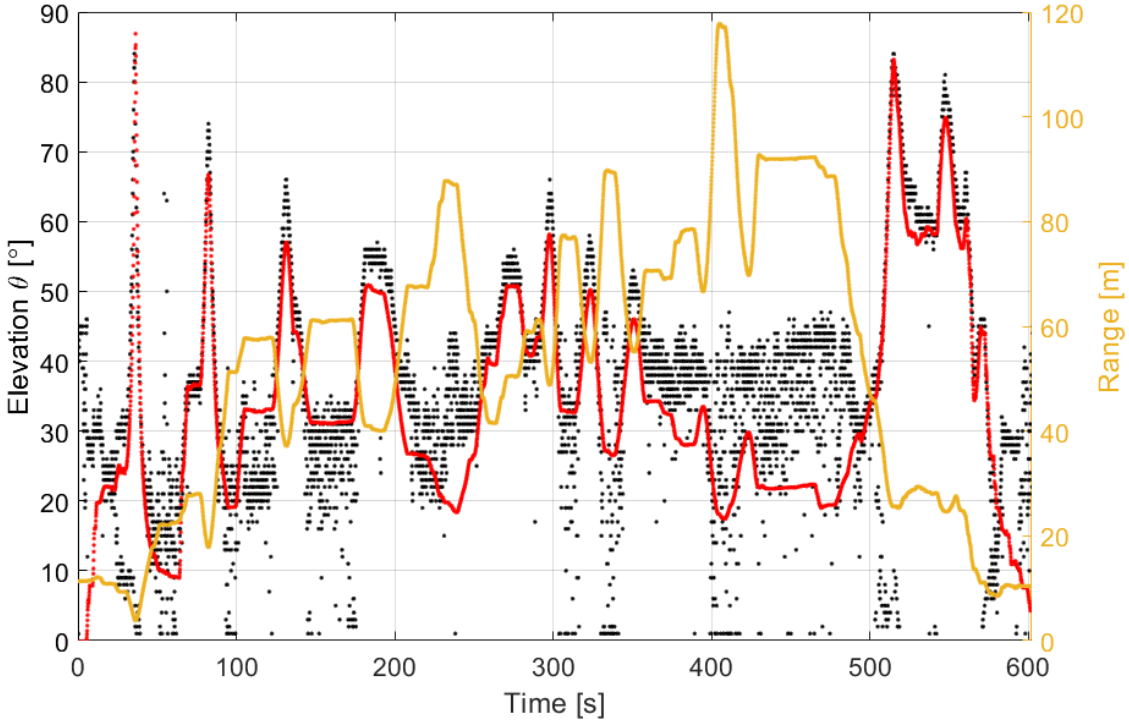


(e) -20log(Range) & OSPL 250-3000Hz

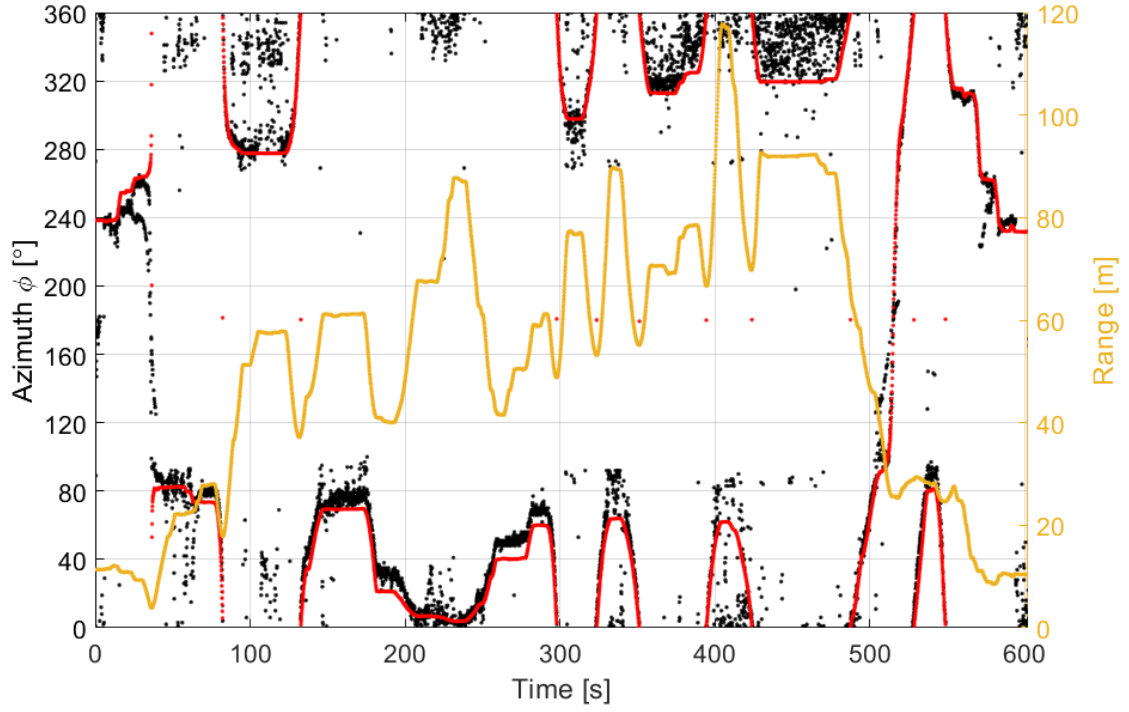


(f) -20log(Range) & OSPL 3000-7000Hz

Figure 6.22: Flight details DJI Mavic flight with interference

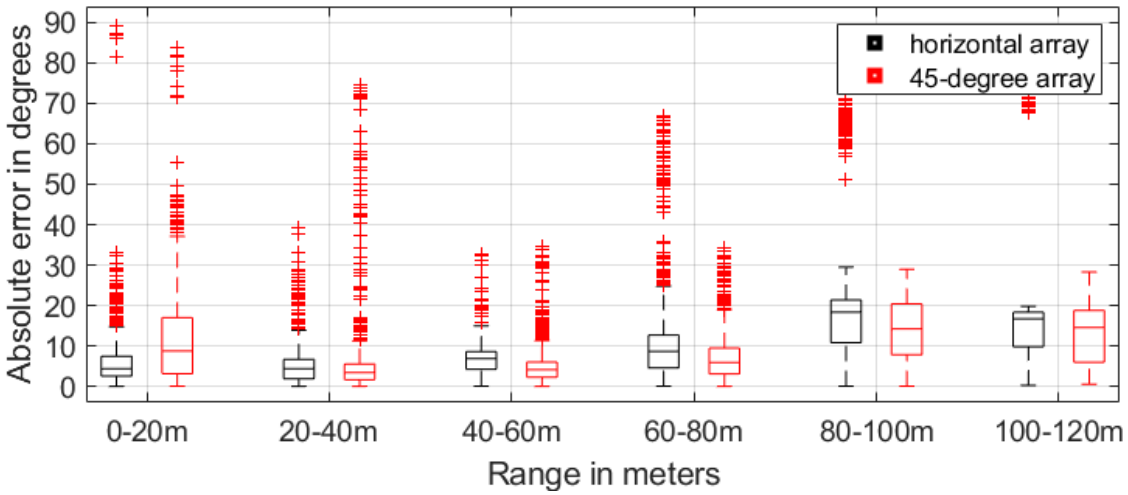


(a) Elevation angle

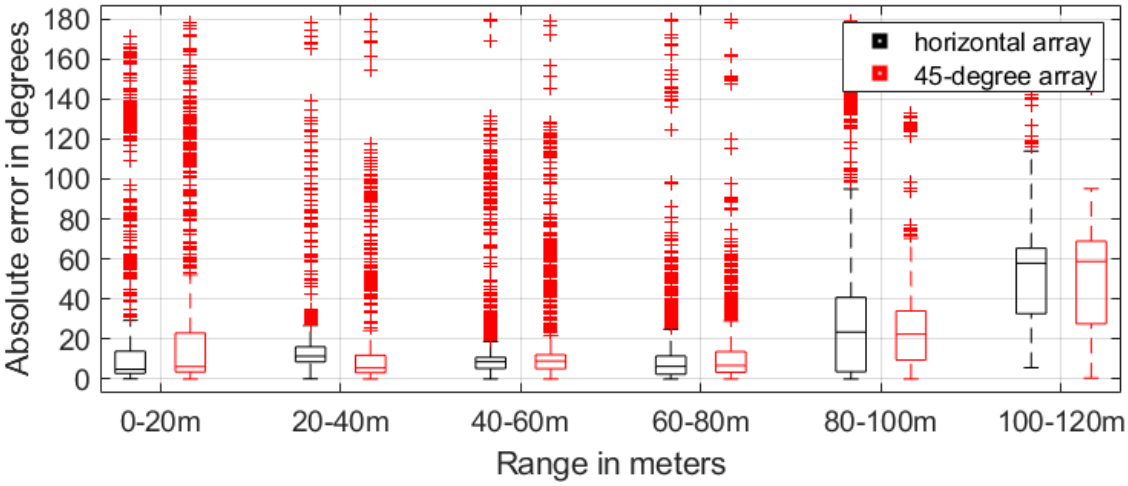


(b) Azimuth angle

Figure 6.23: Beamformed snapshots of an outdoor DJI Mavic flyby with interference measured with the 45-degree array. Conventional frequency domain beamforming applied to a frequency range of 3000-7000Hz. The black dots indicate the snapshot beamform maximum, the red line is the GPS reference location and the yellow line is the distance between the drone and the array.

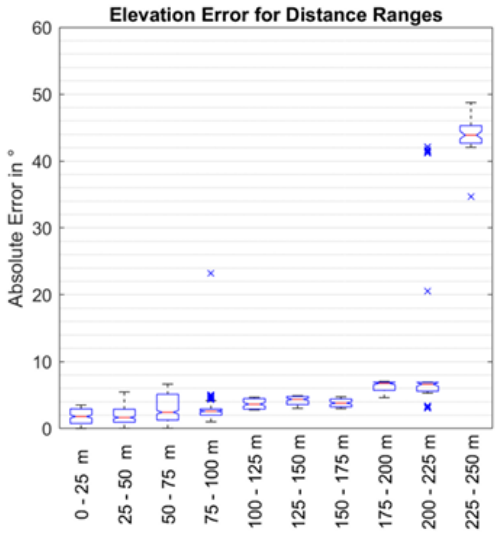


(a) DJI Mavic elevation localization error

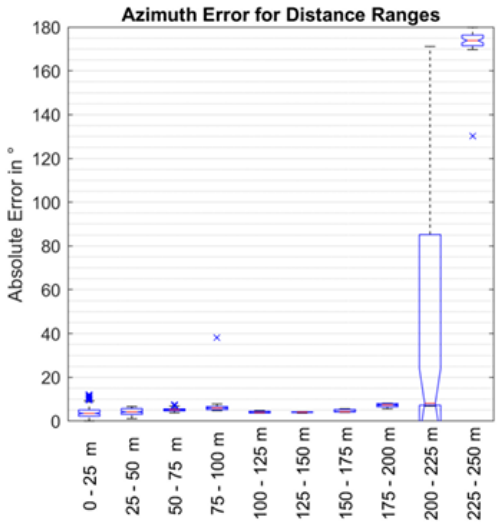


(b) DJI Mavic azimuth localization error

Figure 6.24: Boxplots indicating the error between the reference GPS signal and the maximum locations obtained with beamforming. The black boxes represent the error values from the measurement with the horizontal array and the red boxes represent the error values from the measurement with the 45-degree array.

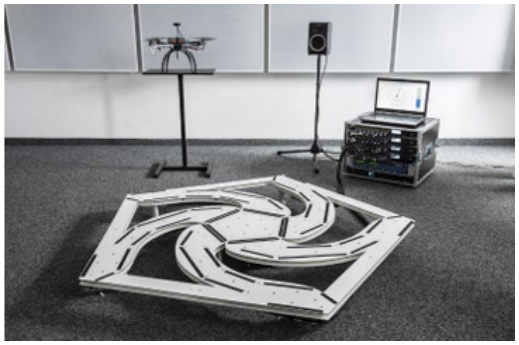


(a) Blass elevation



(b) Blass azimuth

Figure 6.25: Blass 2020 boxplots [21]



(a) Blass array (1.50m diameter)

(b) ReSpeaker array (0.05m diameter)

Figure 6.26: Two arrays

higher sound levels. They also tested a DJI Mavic 2 Pro, the same drone as the Mavic here, but they did not share boxplots of that measurement. It was mentioned that the Mavic needed to approach 100m to be localized and 50m to become the dominant noise source. This probably suggests that the accuracy of that measurement was within a couple of degrees at those distances. The results of the experiment in this project show that the ReSpeaker array is also capable of localizing drones at these distances, but with an accuracy of about 10 degrees. This places it in a different class, but it can still be a relevant option to roughly track the presence of a drone. This is an impressive result for an array this small. It is pictured one last time in figure 6.26 next to the much larger array that lead to the boxplots in figure 6.25.

Chapter 7

Conclusion

This chapter is structured based on the research questions posed in the introduction. The first is about the ReSpeaker arrays localization ability 7.1. Then the effect of an angled orientation is discussed 7.2. Thirdly, the difference between conventional and functional beamforming is evaluated 7.3. A final section 7.4 mentions potential future options.

7.1 Research question 1 localization

- Can a single small low-cost microphone array detect and localize a passing small drone?
  - What is the detection range of one small array for noise from a small drone?

The results from the previous chapter show that this small array is capable of localizing the general direction of a passing drone, with some offset. Looking at the median values, it is capable of tracking the Parrot Bebop within 10 degrees azimuth and elevation up to the maximum tested distance of 97 meters.

The localization for the DJI Mavic was slightly less accurate. The median error was within 10 degrees up to the distance bin of 60-80m.

If an error up to 10 degrees is acceptable, the Parrot Bebop can be tracked up to the maximum distance bin of 80-100m and the DJI mavic can be tracked up to the distance bin of 60-80m. If the array were to be coupled with another device that can also perform tracking, such as a LIDAR system or a camera, an error of 10 degrees could perhaps be sufficient.

7.2 Research question 2 orientation

- What is the localization accuracy when using an array placed in a differently angled orientation?
  - What array orientation has the best localization accuracy?

The localization is generally more accurate with the array placed at a 45-degree angle. This is especially true for elevation tracking.

Overall, the best localization accuracy was with the 45-degree array.

7.3 Research question 3 advanced beamforming

- Does the application of advanced beamforming techniques improve the localization accuracy of the system?

The boxplots corresponding to functional beamforming are generally very similar to the conventional beamforming boxplots. There does not appear to be a noticeable benefit from the application

of functional beamforming. Its application also increases the calculation time slightly as it involves an extra step during the frequency calculation.

There is no clear benefit from the application of functional beamforming.

7.4 Future possibilities

Using multiple small arrays synchronized with a central controller would lead to more accurate localization at low frequencies due to the increased aperture, equation 3.13. This could be necessary if the array is placed indoors or under a cover to protect it from weather effects, while trying to measure sound outside. Drones produce the highest sound levels at lower frequencies and lower frequencies are less sensitive to attenuation while traveling through any medium. To achieve this the arrays would need to be synchronized in time with an accuracy under 1ms, since beamforming relies on the difference in time between different microphones picking up the same signal. It is not certain if even this strict accuracy suffices, as with a speed of sound of 343m/s a distance of 0.0457m is covered in about 0.1ms.

An attempt was made to synchronize measurements with two laptops running Windows 11 by running a script to start the measurement based on the system clock time but these clocks are only specific up to seconds. There were also various Windows issues with connecting two laptops with an ethernet cable. This is a necessary step because the two laptops connected to the same Windows time server were out of sync by more than 3 seconds. The ReSpeaker array can instead be connected to a Raspberry Pi. Using several Raspberry Pi's and a central controller could potentially lead to a synced detection network.

Alternatively, if the target objective produces noise at sufficiently high frequencies, a single array would suffice. It could be placed under either light cover or a dynamic cover that provides shelter based on some sort of precipitation/storm detection system. Drones would also not be present in those conditions. Placing it at a 45-degree angle would improve the results based on the information obtained in this report.

If computational power is limited, tracking using staggered narrow bands at higher frequencies could be considered. For example 2500-2520, 3000-3020 etc. Harmonics at this range are all merged together so the specific range does not matter that much. Choosing bands with some spectral distance reduces the effect of grating lobes. The initial drone detection could be done by detecting a frequency profile that matches known drone sounds, for example with DroneFinderNet as mentioned in [4], and then apply beamforming at a higher frequency range for a localization that's accurate with this small array.

An alternate small off-the-shelf array may be necessary for future tests using arrays of this size. Although it is still in stock in various online stores; the ReSpeaker Mic Array v2.0 was first available in 2018, uses a USB-B port, and the component microphones and integrated circuit are no longer produced. Future research could investigate the localization accuracy of a different, newer, small array.

7.5 References

[1] Rico Merkert, James Bushell, Managing the drone revolution: A systematic literature review into the current use of airborne drones and future strategic directions for their effective control, Journal of Air Transport Management, Volume 89, 2020, 101929, ISSN 0969-6997, doi.org/10.1016/j.jairtraman.2020.101929
[2] https://www.easa.europa.eu/en/the-agency/faqs/drones-uas, visited May 8th
[3] Blanchard, Torea, J-H. Thomas, and Kosai Raoof. "Acoustic localization and tracking of a multi-rotor unmanned aerial vehicle using an array with few microphones." The Journal of the Acoustical Society of America 148.3 (2020): 1456-1467. doi.org/10.1121/10.0001930
[4] Ding, Siyi, et al. "Drone detection and tracking system based on fused acoustical and optical approaches." Advanced Intelligent Systems 5.10 (2023): 2300251. doi.org/10.1002/aisy.202300251
[5] Guo, Junfeng, Ishtiaq Ahmad, and KyungHi Chang. "Classification, positioning, and tracking of drones by HMM using acoustic circular microphone array beamforming." EURASIP Journal on Wireless Communications and Networking 2020 (2020): 1-19. doi.org/10.1186/s13638019-1632-9
[6] Herold, Gert, et al. "Detection and separate tracking of swarm quadcopter drones using microphone array measurements." Proceedings of the Berlin Beamforming Conference (BeBeC), Berlin, Germany. 2020.
[7] Altena, Anique, et al. "Comparison of acoustic localisation techniques for drone position estimation using real-world experimental data." Proceedings of the 29th International Congress on Sound and Vibration, ICSV 2023. Society of Acoustics, 2023.
[8] Itare, Nathan, et al. “Acoustic Estimation of the Direction of Arrival of an Unmanned Aerial Vehicle Based on Frequency Tracking in the Time-Frequency Plane.” Sensors 2022, Vol. 22, Page 4021. doi.org/10.3390/S22114021
[9] Blass, Martin, et al. “Towards mobile microphone array based UAV tracking.” Quiet Drones 2022 - International e-Symposium on UAV / UAS Noise
[10] Busset, Joël, et al. "Detection and tracking of drones using advanced acoustic cameras." Unmanned/Unattended Sensors and Sensor Networks XI; and Advanced Free-Space Optical Communication Techniques and Applications. Vol. 9647. SPIE, 2015. doi.org/10.1117/12.2194309
[11] Paolo Chiariotti, Milena Martarelli, Paolo Castellini, Acoustic beamforming for noise source localization – Reviews, methodology and applications, Mechanical Systems and Signal Processing, Volume 120, 2019, Pages 422-448, ISSN 0888-3270, doi.org/10.1016/j.ymssp.2018.09.019.
[12] Aldeman, M., & Raman, G. (2018). Effects of array scaling and advanced beamforming algorithms on the angular resolution of microphone array systems. Applied Acoustics, 132, 5881. doi.org/10.1016/j.apacoust.2017.11.009
[13] Merino Martinez, R., Snellen, M., & Simons, D. (2016). Functional Beamforming applied to full scale landing aircraft. In 6th Berlin Beamforming Conference: Berlin, Germany
[14] R. P. Dougherty, “Functional beamforming,” in Proceedings on CD of the 5th Berlin Beamforming Conference, 19-20 February 2014. Berlin, Germany., GFaI, e.V., Berlin, February 2014.
[15] P. Sijtsma, “CLEAN based on spatial source coherence,” International Journal of Aeroacoustics, vol. 6, pp. 357–374, 2 doi.org/10.2514/6.2007-3436
[16] Griffioen, Pieter D. “Development of a low cost acoustic detector of aircraft movements.” 2024. Delft University of Technology, MSc Thesis.
[17] Jiang, Hanbo & Zhong, Siyang & Wu, Han & Zhang, Xin & Huang, Xun & Zhou, Guocheng & Chen, Bao. (2022). Radiation Modes of Propeller Tonal Noise. Journal of Vibration and Acoustics. 144. 10.1115/1.4051864.
[18] https://wiki.seeedstudio.com/ReSpeaker_Mic_Array_v2.0/, visited May 8th.
[19] Acoustic measurements, Book by Thomas J. Mueller. doi.org/10.1007/978-3-662-05058-3
[20] Digital Signal Processing, February 2024 by Professor Dick G. Simons & Professor Mirjam Snellen.
[21] Blass, Martin & Graf, Franz. (2021). A Real-Time System for Joint Acoustic Detection and Localization of UAVs. doi.org/10.1007/s11370-024-00584-9