BeBeC-2024-S09

COMPARISON OF MICROPHONE ARRAY METHODS FOR SMALL UAV FLIGHT PATH RECONSTRUCTION

Gert Herold, Adam Kujawski, Simon Jekosch, Ennes Sarradj
Technische Universität Berlin, Department of Engineering Acoustics
Einsteinufer 25, D-10587 Berlin, Germany

Abstract

Microphone arrays offer a valuable means to measure the sound emission patterns of small unmanned aerial vehicles, such as multicopter drones. A requirement for this is knowledge about the 3D position of a drone, or, in case of moving drones, their flight path. With suitable data processing, this information can be derived from microphone array measurement data as well.

Based on an outdoor measurement of a drone flyby, this contribution compares several signal processing configurations that involve frequency domain beamforming to reconstruct the flight path. In addition to discussing visual representations of processing results, the performance of the configurations is assessed using suitable error metrics.

It is demonstrated that a reliable flight path reconstruction is achievable using microphone array measurements and that the initially most promising configuration does not necessarily yield the best overall result.

1 INTRODUCTION

Small unmanned aerial vehicles (UAVs), or drones, are becoming increasingly common in various fields such as journalism, inspections, agriculture, and delivery. As their numbers grow, so does the noise they generate, which has become a growing concern. The noise from drones varies based on their specific applications and operational constraints.

Regulations limit noise exposure, especially in residential areas. Ensuring that drones comply with these limits before deployment is challenging because their noise emission characteristics during typical modes of operation are not fully known. Although standardized methods for measuring drone noise have been developed [9], there is still no standardized framework for determining the *emission* characteristics of drones in operational modes like cruise flight.

Previous research has proposed and tested a framework for drone noise emission [6, 8]. It was designed and tested only for drones that can be flown indoors, typically weighing up to 5 kg and flying at low speeds. However, to apply these methods to a broader range of drone models, outdoor measurements with simple equipment setups are needed. Konzel and Greenwood [10] conducted outdoor measurements of drone flyovers, focusing on evaluating metrics for incident sound. However, calculating emission characteristics requires more complex data processing, including back-propagating the measured sound to a fixed distance from the drone, which necessitates knowing the drone's position at all times.

This paper presents and investigates a method to reconstruct a drone's position during flyby measurements with a microphone array consisting of three orthogonal line arrays. Since the data processing involved allows for numerous variations in the parameters and algorithms, several configurations have been tested and are compared in this study.

2 MATERIALS & METHODS

2.1 Measurements

Measurements were conducted on an open grass field with a microphone array consisting of 32 microphones (GRAS 40 PK 1/4 inch free-field), which were arranged in three linear sub-arrays (see Fig. 1). Half of the microphones are mounted in a vertical line array on a tripod and are located between 1.5 m and 9.3 m above the ground. The remaining 16 microphones are distributed in two crossing line arrays (in the direction of flight and across) and mounted on SAE circular ground planes [1] (see Fig. 2 left) in a "side-set" configuration. Interferences due to ground reflections are expected to be negligible up to about 5 kHz for the ground-based sub-arrays [5].

In this contribution, one flyby of a small quadcopter drone is evaluated:

- drone model: DJI Mavic Pro
- weight: 734 g
- diagonal rotor axes distance: 33.5 cm
- dimensions without propeller.: $(22 \times 27.5 \times 8.4)$ cm
- propeller diameter: 20 cm

The drone was manually operated via remote control. The pilot was instructed to follow a straight path, ensuring the drone passed over and parallel to the array with the line array aligned in the x direction directly beneath it.

Figure 3 shows spectrograms of the flyby recorded at two different microphones. The rotor tones and their harmonics can be clearly seen. The sound pressure level reaches its maximum at around 2.5 s across the entire frequency range, indicating that the drone is closest to the microphone at this point. In the spectrogram of the microphone positioned 1.5 m above the ground, an interference pattern caused by ground reflections is visible [2]. All subsequent calculations in this paper are based on this 5-second flyby recording with the microphone array.

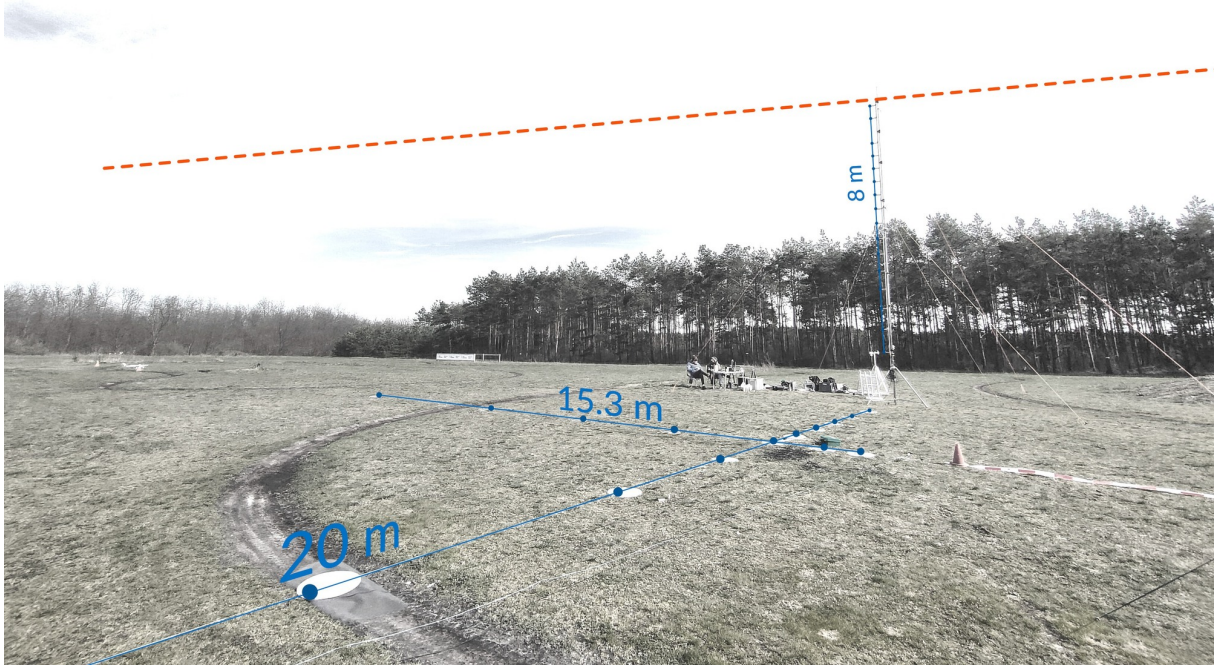

Figure 1: Measurement setup with microphone array (blue) and intended flight path (orange).

Figure 2: Left: Two microphone ground planes with windscreen and camera for flyover recording. Right: Drone used for the measurement evaluated here.

2.2 Data processing

The method for the trajectory detection is based on the processing described in earlier publications [6, 7]. For several of the processing steps, it is possible to apply different algorithms or parameters. A comprehensive study of alternatives is beyond the scope of this paper. Nonetheless, some variations are investigated here. The respective options are denoted by bold abbreviations (**abbr**) enclosed in brackets.

The primary objective of the data processing is to determine the 3D position of the drone at each point in time. For this, it is assumed that the drone consistently represents the predominant

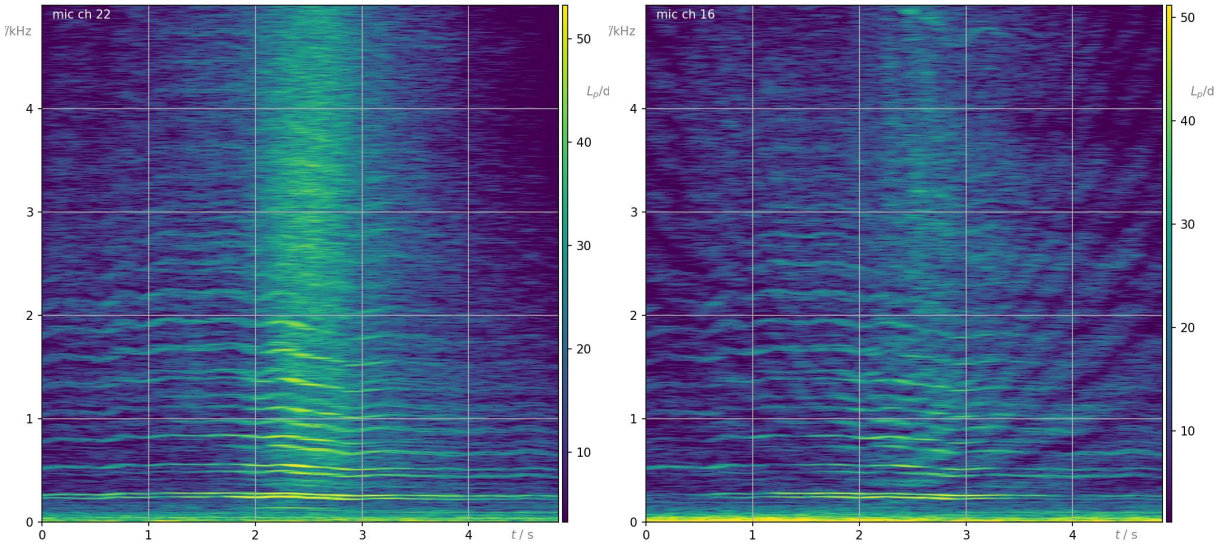

Figure 3: Spectrograms at the center ground microphone (left) and at the bottom microphone of the vertical line sub-array, 1.5 m above ground (right).

sound source and is sufficiently compact to be treated as a point source. In order to track its position, the recorded signals are divided into short equal-length time segments. In each segment, the drone's movement is considered to be minimal. This enables the application of frequency-domain beamforming methods, which are usually employed for stationary sources. For each of the segments, a short-time cross-spectral matrix (CSM) is estimated using Welch's method [15].

$$\boldsymbol{C} = \frac{1}{K} \sum_{k=1}^{K} \boldsymbol{p}_k \, \boldsymbol{p}_k{}^{\mathrm{H}} \, . \tag{1}$$

Every $\boldsymbol{p}_k \in \mathbb{C}^M$ contains the complex spectral data from one FFT block for each of the M microphones. K is the number of individual cross-spectra which can be obtained from one segment.

In contrast to typical frequency-domain beamforming applications and due to the short duration of the segments, this number is relatively low (usually below 50). However, it is assumed that this is sufficient for obtaining qualitatively meaningful results that do not require quantitative accuracy.

The next step requires the shifting of phases and amplitudes according to the distances of the microphones to the discretized focus points $r_{s,m}$ using a model describing the propagation of sound. This is typically done using a steering vector $\boldsymbol{h}$ for a focus point, whose entries can be calculated via

$$h_m = \frac{1}{r_{s,m} \sqrt{M \sum_{l=1}^{M} r_{s,l}{}^{-2}}} \mathrm{e}^{-\mathrm{j}k(r_{s,m}-r_{s,0})} \, , \qquad m = 1 \ldots M \, , \tag{2}$$

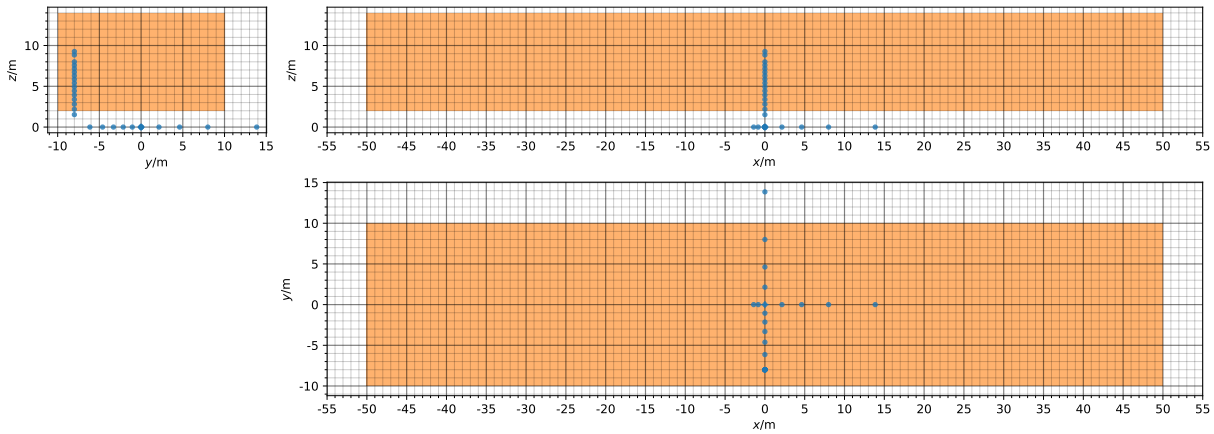

Figure 4: Monitored area for trajectory tracking (orange). Microphones in blue.

where $s = 1 \dots N$ is the index of discretized focus points. The above equation (2) is referred to as "formulation IV" in [12]. In addition to this (**SV4**), formulations I (**SV1**) and III (**SV3**) from the same reference are used in this paper.

The acoustic beamforming in its classic form (**CB**) is calculated via:

$$b(\boldsymbol{x}_s) = \boldsymbol{h}^{\mathrm{H}}(\boldsymbol{x}_s)\, \boldsymbol{C}\, \boldsymbol{h}(\boldsymbol{x}_s), \tag{3}$$

yielding squared sound pressures at reference $r_{s,0}$.

For improved spatial resolution, Dougherty [3] introduced Functional Beamforming (**FB**):

$$b(\boldsymbol{x}_s) = \left(\boldsymbol{h}^{\mathrm{H}}(\boldsymbol{x}_s)\, \boldsymbol{C}^{\frac{1}{\nu}}\, \boldsymbol{h}(\boldsymbol{x}_s)\right)^{\nu}, \tag{4}$$

where the exponent ν aids in suppressing artifacts such as side lobes. Both classic and functional beamforming are used in the evaluations.

As the sidelobes are dependent on the array geometry and the position of the source relative to the array, they can also be suppressed by geometrically averaging separate beamforming results that use the same focus grid but different arrays. This multiplicative beamforming approach (**mult**) was described by Porteous et al. [11] and is used in this context with the three line sub-arrays.

Frequency domain beamforming is typically performed independently for each discrete frequency. Results across a broader frequency range are then obtained by summing narrow-band results within the desired range. For the calculations done here, only frequencies ranging from 400 Hz to 4400 Hz are considered. Beamforming results at lower frequencies are affected by environmental noise and static pressure fluctuations (e. g. due to wind) detected by the microphones. At higher frequencies, stronger sidelobes and spatial aliasing become a concern.

As absolute quantitative information is not important for the localizing task, a possible processing step is the normalization (**norm**) of narrow band beamforming results before the summation. This ensures equal weighting of all frequencies in the final result. In the case of predominantly tonal sources, this approach allows for the utilization of broadband components for the localization, albeit at the potential expense of a lower overall signal-to-noise ratio.

Each beamforming application yields a 3D sound map representing the source distribution at one time segment. Figure 4 shows the area which is scanned for sources. Within each sound map, the maximum is detected and its position stored for further processing. The following section aims to identify the best-performing configuration among those investigated for the evaluated measurement. Important measurement and processing parameters are listed in Table 1. All calculations for this paper were conducted with the help the software package *Acoular* [13, 14], which specializes in microphone array signal processing.

Table 1: Measurement and signal processing parameters. Abbreviations of variants in bold.

Microphones (sub-arrays)	32 (7; 9; 16)
Array extent	(15.3 m; 21.9 m; 9.3 m)
Sampling frequency	51 200 Hz
Monitored area extent	$(l_x, l_y, l_z) = (100\,\text{m}, 20\,\text{m}, 10\,\text{m})$
Resolution of focus area	0.1 m
FFT for CSM	1024 samples
	von Hann window
	50 % overlap
Averaging time	0.2 s
Time step size	0.05 s
Evaluated frequency range	[400 Hz, 4400 Hz]
Beamforming	Classic Beamforming (**CB**) / Functional Bf. $\nu = 8$ (**FB**)
further processing options	3-line geom. averaging (**mult**)
	spectrum normalization (**norm**)
CSM main diagonal	used (**wd**) / removed (**rd**)
Steering vector	reference distance $r_{s,0} = 1$ m
	Formulations [12] I (**SV1**) / III (**SV3**) / IV (**SV4**)

3 RESULTS

Figures 5, 6, and 7 show exemplary sound maps (spatial distribution of sound pressure levels relative to the respective maximum) for three distinct processing configurations at two points in time. The left plots show the moment where evaluation starts and the drone is positioned on the left in the representations with the x axis present. The right plots feature an evaluation when the drone is nearly directly above the array center at $t = 2.5$ s.

In the case of the classic beamforming evaluation (Fig. 5), a lower dynamic range than in the other configurations can be observed. This potentially makes it more challenging to reliably detect maxima under poorer measurement conditions, e. g. if background noise levels are higher.

The Functional Beamforming result (Fig. 6) has a significantly increased signal-to-noise ratio. If the drone is far from the array, however, its representation is considerably elongated, with sidelobes still clearly visbile. Figure 7 features the sound map of an evaluation with multiplicative

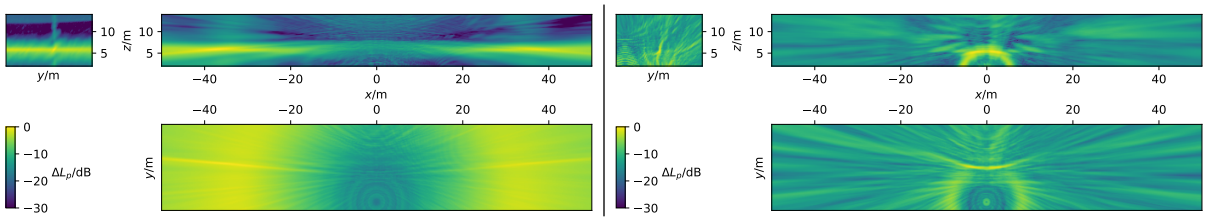

Figure 5: Classic beamforming (CB, dr, SV4): exemplary short-time sound maps (slices through planes with source position). Left: $t = 0$, $x_d = -36\,\mathrm{m}$, Right: $t = 0$, $x_d = 0.6\,\mathrm{m}$.

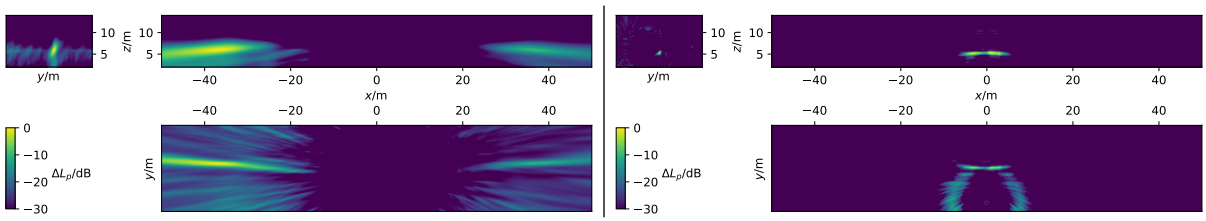

Figure 6: Functional beamforming (FB, dr, SV4): exemplary short-time sound maps. Left: $t = 0$, $x_d = -36\,\mathrm{m}$, Right: $t = 0$, $x_d = 0.6\,\mathrm{m}$.

beamforming and spectral normalization. The maps exhibit a high dynamic, with less visible artifacts compared to Functional Beamforming.

In all maps, the representation of the source close to the array center lacks a clear and unambiguous localization. Multiple maxima are visible, and from visual observation of a single snapshot alone, a clear distinction which of those represents the actual source position is not possible. This limitation arises from the imaging properties of the microphone array. While a geometry with three lines was chosen for practical reason and easy setup, line arrays possess good resolution capabilities in only one spatial direction. Consequently, they exhibit distinct point spread functions in their perpendicular planes, which is suboptimal for source detection in one of these planes. A possible mitigation of the ambiguities can be achieved by considering successive time steps together, provided that the drone does not perform a sudden change in its flight path.

In total, 21 distinct data processing configurations were evaluated. Following each processing, the respective maximum locations were connected. The resulting trajectories are collectively plotted in Figure 8. While some deviations are visible, an overall agreement of the resulting flight path can be observed.

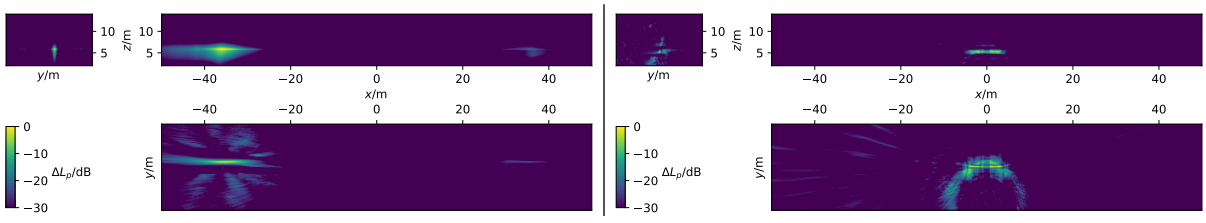

Figure 7: Sub-array multiplied and spectrum-normalized beamforming (CB, dr, SV4, norm, mult): short-time sound maps. Left: $t = 0$, $x_d = -36\,\mathrm{m}$, Right: $t = 2.5\,\mathrm{s}$, $x_d = 0.6\,\mathrm{m}$.

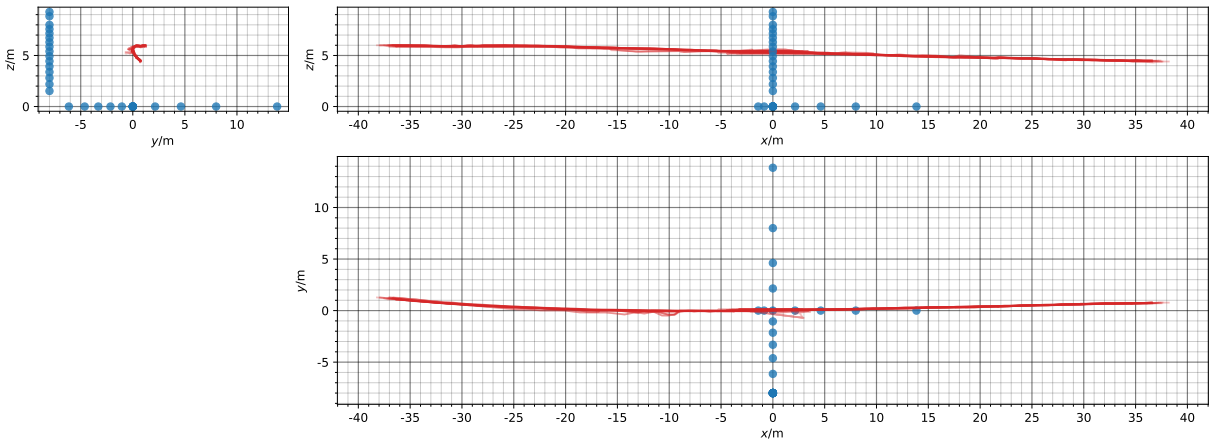

Figure 8: Trajectories reconstructed from all 21 different processing chains (red).

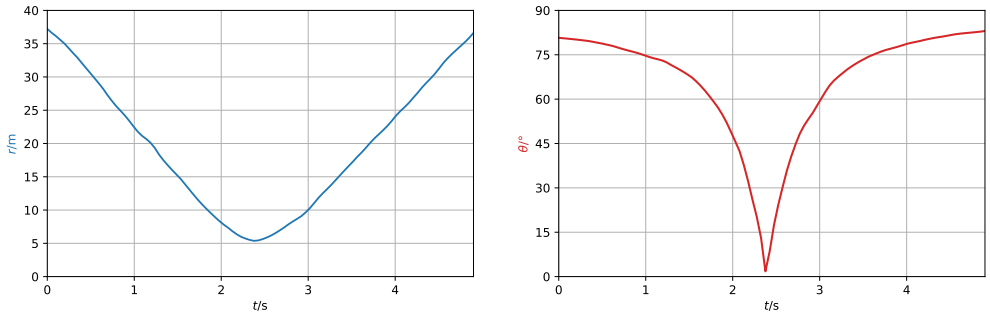

Figure 9: Time-dependent distance r of the Kalman-filtered "best fit" trajectory to the center of the ground microphone array and relative angle θ to its center vertical axis.

A straight-forward way of quantitatively assessing the performance of the configurations would be to compare each result to a ground truth. However, no other means of tracking the drone's position with sufficient precision was available for these measurements. Therefore, a "probable" ground truth was determined using a two step-process:

1. For each time step, determine the median position of all 21 variants.
2. Process the data using a Kalman filter [4] to ensure the calculation of a physically plausible flight path.

Figure 9 shows the resulting smooth trajectory in terms of the distance to the array center and the relative angle towards the z-axis (perpendicular to the ground) over time. These two representations form the basis for evaluating error metrics introduced in an earlier study [8]. For each configuration, the reconstructed radial position r_c and angular position θ_c are compared against the nominal trajectory. The relative distance error is quantified as a ratio via

$$\Delta r_{\mathrm{rel}} = \frac{r_c - r}{r}. \tag{5}$$

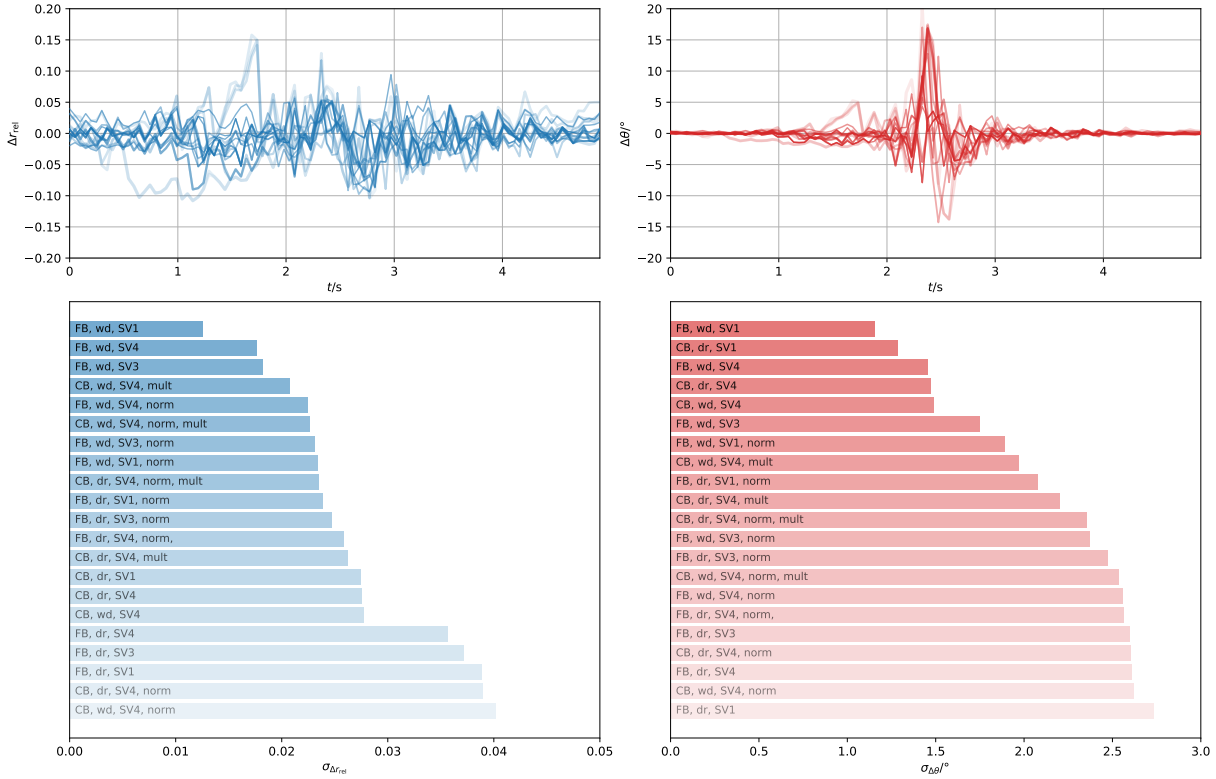

Figure 10: Relative distance error (left) and angle error (right) along the trajectory for different signal processing. The legend below at the same time shows the respective processing chains sorted by the standard deviation of the localization errors (smaller values are better).

The angle error is determined by subtracting the nominal angle from the reconstructed angle:

$$\Delta\theta = \theta_c - \theta \ . \tag{6}$$

The resulting time progression of the error metrics for the different data processing configurations are shown in Figure 10. Apart from a few exceptions, the deviations of the distance error from the ground truth stay within ±10 %. The magnitude of the angle error is highest between seconds 2 and 3, where the drone is closest to the array and where the above-mentioned ambiguities due to the point spread function may lead to a wrong localization of the drone.

The second row of Figure 10 shows the standard deviations of the respective error metric calculated for each data processing configuration and displays them in a sorted manner. With the lowest deviation in both metrics, Functional Beamforming with the full CSM including the original main diagonal and using steering vector formulation I is the most reliable configuration. Remarkably, the configuration which shows good sidelobe and noise suppression in Figure 7, only ranks 9th for relative distance error and 11th for angle error.

This indicates that optimizing for a high dynamic range in the sound map does not guarantee the best overall performance for trajectory reconstruction. It should be kept in mind, however, that the presented results are only representative for a single 5-second flyby. In order to produce

a more reliable and statistically sound ranking of the data processing configurations, a larger number of flybys has to be evaluated with the processing presented above.

4 CONCLUSION

As important part of the data processing for the acoustic characterization of small UAVs during cruise flight, this paper focuses on the continuous tracking of the drone using a microphone array setup. Evaluating an outdoor measurement of a quadcopter drone flyby, several data processing configurations for the source localization were tested and their performance was evaluated quantitatively using suitable error metrics. It was found that for the evaluated measurement, Functional Beamforming with steering vector formulation I outperforms the other investigated configurations. For a more robust ranking, more evaluations are needed, which can be done following the methodology described in this paper.

For future measurements in a similar environment, it is recommended to adapt the used microphone array geometry in order to avoid ambiguities in the results. It should be noted, however, that Kalman filtering or other plausibility considerations allow the usage of geometries that are not optimal for consistent localization.

These investigations have shown that it is possible to perform outdoor measurements of drone flybys with a microphone array and reconstruct the flight path purely based on the recorded acoustic signals.

References

[1] A-21 Aircraft Noise Measurement Aviation Emission Modeling. "Ground-Plane Microphone Configuration for Propeller-Driven Light-Aircraft Noise Measurement." Aerospace Standard ARP4055, SAE International, 1988. doi:10.4271/ARP4055.

[2] W. M. Carey. "Lloyd's mirror – Image interference effects." *Acoustics Today*, 5(2), 15–21, 2009. ISSN 15570223. doi:10.1121/1.3182842.

[3] R. P. Dougherty. "Functional beamforming." In *Proceedings of the 5th Berlin Beamforming Conference*, pages 1–25. Berlin, 2014.

[4] D. Duckworth, G. Lipták, M. Balatsko, N. Werner, and P. Haessig. "pykalman." https://github.com/pykalman/pykalman, 2024.

[5] J. M. Giannakis and E. Nesbitt. "Evaluation of a Correction Factor for Flyover-Noise Ground Plane Microphones." In *AIAA AVIATION 2020 FORUM*. American Institute of Aeronautics and Astronautics, VIRTUAL EVENT, 2020. ISBN 978-1-62410-598-2. doi:10.2514/6.2020-2612.

[6] G. Herold. "In-flight directivity and sound power measurement of small-scale unmanned aerial systems." *Acta Acustica*, 6(58), 13, 2022. ISSN 2681-4617. doi:10.1051/aacus/2022052.

[7] G. Herold, A. Kujawski, C. Strümpfel, S. Huschbeck, M. U. de Haag, and E. Sarradj. "Detection and Separate Tracking of Swarm Quadcopter Drones Using Microphone Array Measurements." In *Proceedings of the 8th Berlin Beamforming Conference*, pages 1–19. Berlin, Germany, 2020.

[8] G. Herold and E. Sarradj. "Microphone array based trajectory reconstruction for small UAV fly-by measurements." In *Proceedings of the 29th AIAA/CEAS Aeroacoustics Conference*, pages 1–12. San Diego, CA, 2023. ISBN 978-1-62410-704-7. doi:10.2514/6.2023-3818.

[9] ISO. "ISO 5305:2024 – Noise measurements for UAS (unmanned aircraft systems).", 2024.

[10] N. Konzel and E. Greenwood. "Ground-based Acoustic Measurements of Small Multirotor Aircraft." In *Proceedings of the Vertical Flight Society 78th Annual Forum*, pages 1–11. The Vertical Flight Society, Fort Worth, Texas, 2022. doi:10.4050/F-0078-2022-17435.

[11] R. Porteous, Z. Prime, V. Valeau, C. Doolan, and D. Moreau. "Three-dimensional beamforming of aeroacoustic sources." In *INTER-NOISE 2014*, pages 1–9. Melbourne, Australia, 2014.

[12] E. Sarradj. "Three-dimensional acoustic source mapping with different beamforming steering vector formulations." *Advances in Acoustics and Vibration*, 2012, 1–12, 2012. ISSN 1687-6261. doi:10.1155/2012/292695.

[13] E. Sarradj and G. Herold. "A Python framework for microphone array data processing." *Applied Acoustics*, 116, 50–58, 2017. ISSN 0003682X. doi:10.1016/j.apacoust.2016.09.015.

[14] E. Sarradj, G. Herold, A. Kujawski, S. Jekosch, A. J. R. Pelling, M. Czuchaj, T. Gensch, and S. Oertwig. "Acoular – Acoustic testing and source mapping software." Zenodo, 2024. doi:10.5281/ZENODO.11234980.

[15] P. D. Welch. "The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms." *IEEE Transactions on Audio and Electroacoustics*, 15(2), 70–73, 1967. ISSN 00189278. doi:10.1109/TAU.1967.1161901.