

Optimized TDOA-based Drone Localization With Distributed Microphones

Rigel P. Fernandes, José A. Apolinário Jr., Julio Cesar Duarte, and José M. de Seixas

Abstract—Accurately localizing drones in complex environments remains a significant challenge, with important implications for defense, law enforcement, and autonomous systems. This study addresses the problem of estimating drone localization in environments characterized by strong reflections and background noise. We employ Time Difference of Arrival (TDOA) techniques for localization estimation and compare them with a specialized machine learning regression model. While previous works have considered neural networks for audio source localization, these methods often suffer from limited generalization across different environments. To address this, we propose a novel method that enhances the TDOA vector by incorporating both primary and secondary peaks of the cross-correlation, guided by the Zero Cyclic Sum condition. Additionally, we introduce optimization strategies that selectively reduce the number of TDOA inputs based on a least-squares cost function. We present a comparative analysis of TDOA-based optimization techniques with a machine learning method that utilizes the environment’s reverberation fingerprint as input features for training. Experimental results demonstrate that the proposed TDOA-based method achieves a localization accuracy of 0.55 ± 0.35 meters, showcasing its effectiveness and practical applicability in challenging acoustic environments.

Index Terms—Drone localization, TDOA, Zero Cyclic Sum, logistic regression.

I. INTRODUCTION

TARGET localization problem holds significant importance across different domains, ranging from academic research to defense applications and industry, including scenarios such as underwater target localization [1, 2, 3], shooter localization using drones [4, 5, 6, 7], and target localization using radiofrequency signals [8]. A recent challenge in this field is the accurate localization of quadcopter drones, which are increasingly used for diverse applications such as aerial surveillance [9], search and rescue missions [10], and environmental monitoring.

Rigel P. Fernandes and Julio Cesar Duarte are with the Program of Defense Engineering, Military Institute of Engineering (IME), Rio de Janeiro-RJ, Brazil; E-mails: rigelfernandes@gmail.com and duarte@ime.eb.br; ORCID: 0000-0001-5269-2342 and 0000-0001-6656-1247.

José A. Apolinário Jr. is with the Program of Electrical Engineering and Defense Engineering, Military Institute of Engineering (IME), Rio de Janeiro-RJ, Brazil; E-mail: apolin@ime.eb.br; ORCID: 0000-0003-1426-9636.

José M. de Seixas is with the Signal Processing Lab, COPPE/Poli, Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro-RJ, Brazil; E-mail: seixas@lps.ufrj.br; ORCID: 0000-0001-5148-7363.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES) – Finance Code 001, FAPERJ, CNPq, and the Brazilian Navy.

Submission: 2025-09-01, First decision: 2025-11-14, Acceptance: 2026-01-27, Publication: 2026-02-07.

Digital Object Identifier: 10.14209/jcis.2026.4

Previous methods in target localization have considered both simulated and real-world signals. For instance, Apolinário Jr. et al. [11] proposed a data-selective approach to estimate the localization using simulated data. In our latest work [12], we observed that TDOA measurements, estimated as the primary peak of the cross-correlations (i.e., the peak with the highest amplitude), were inadequate in highly reverberant environments, making them less reliable. Conversely, He et al. [13] proposed collaborative localization schemes, i.e., the use of drones equipped with speakers that emit pseudo-random noise (PRN) pulses to improve detection and localization accuracy. PRN pulses are orthogonal and statistically independent, allowing for clear separation and identification of signals from multiple drones, thus facilitating concurrent detection and localization. Wang and Hu [14] explored a decentralized framework for geometry calibration in acoustic transceiver networks, employing Direction-of-Arrival (DOA) and TDOA measurements. This framework addresses the challenge of inconsistencies among local coordinates by implementing a distributed consensus approach to map local coordinates to a unified global system. The proposed method has been validated through numerical simulations.

Other related works [15] focused on drone localization in indoor environments using acoustic-inertial measurements. Zhao and Chen [16] employed a Spatial-Temporal Graph Convolutional Network for localizing acoustic sources in composite panels. Zhang et al. [1] explored multifunctional wearable patches that integrate sensing and communication capabilities. Additionally, Melchiorre et al. [2] discussed AI-enhanced acoustic emission techniques for crack source localization. Xiang et al. [17], Jia and Feng [18] covered the far- and near-field source localization, with a focus on multi-station passive radar systems and the use of DOA and Time of Arrival (TOA) measurements. A factor graph-based geolocation and tracking network using the Extended Kalman Filter is presented by Jiang et al. [19], while Yang et al. [20] introduced a TDOA-FDOA-based target localization method. Additionally, Gola and Arya [3] provided a comprehensive review of underwater acoustic sensor networks, covering various aspects from applications to localization methods.

In another previous work [21], we investigated the use of TDOA measurements to estimate the DOA of drones. Building on this foundation, the present work introduces techniques to address the time delay estimation problem, particularly in scenarios where the primary peak does not correspond to the correct time delay. To improve localization accuracy, an optimization step is introduced to discard TDOA values that do not minimize the cost function. Additionally, the Zero

Cyclic Sum (ZCS) cost function is evaluated, enabling the integration of both primary and secondary peaks into the drone localization process.

Prior studies on direction-of-arrival estimation partially inspire the algorithms used in this work under the ZCS condition [21, 22] and on TDOA selection strategies to mitigate noisy or inconsistent delay measurements [23, 24]. However, unlike these earlier approaches, the present study introduces several new contributions specifically tailored for drone localization with distributed acoustic sensors. This work makes the following five-fold contributions to the problem of drone localization using distributed microphones in reverberant environments:

- 1) integration of ZCS with an LS-based algorithm, enabling the exploitation of secondary peaks of cross-correlations while simultaneously rejecting inconsistent TDOAs through LS cost minimization;
- 2) introduction and evaluation of selection strategies, Exhaustive Search (ES) and Greedy TDOA Selection (GTS), designed to progressively refine the set of delays and improve localization accuracy;
- 3) validation of these algorithms with experimental trials in a highly reverberant environment with a real drone and a distributed microphone array;
- 4) comparative analysis with a neural network approach, showing that while data-driven methods can achieve superior accuracy, our proposed solutions offer robustness and generalization without requiring prior training; and
- 5) achievement of sub-meter accuracy in highly reflective scenarios, which underscores the potential of the methods for real-time deployment in embedded systems for UAV monitoring.

The remainder of the paper is as follows. Section II formulates the drone localization problem using acoustic signals, including the underlying assumptions and a detailed description of the dataset employed in this study. Section III introduces the proposed localization method. Section IV presents and analyzes the experimental results. Finally, Section V summarizes the key findings and outlines directions for future work.

II. THE TARGET LOCALIZATION PROBLEM

This section defines the problem addressed in this work, outlines the underlying assumptions, describes the TDOA estimation challenge, presents the localization geometry, and the conventional localization algorithms. The goal here is to establish an understanding of the context and constraints within which the proposed techniques operate.

A. Problem statement and assumptions

For the three-dimensional target localization problem, it is assumed that the target, represented by the unknown position vector \mathbf{p} , corresponds to a hovering drone located within the spatial bounds defined by a distributed microphone array. We also assume exact observer positions \mathbf{r}_k during the observation time in the simulated and experimental phases. The following assumptions are made in this study:

- In this scenario, with a poor GPS signal, the estimated position $\hat{\mathbf{p}}$ of the drone can vary over time due to drift, hovering dynamics, or low-accuracy localization, leading to small deviations from the true position \mathbf{p} . By constraining the drone to hover at a fixed location, these variations remain minimal, enabling a more accurate and reliable estimate of the ground truth position compared to a moving drone;
- The microphones are synchronized, and their positions \mathbf{r}_k are known with high precision for both simulation and experimental phases;
- The acoustic signal propagates in a homogeneous medium with a constant and known speed of sound, v , considered as 343 m/s; and
- Multipath propagation and environmental reflections can cause multiple peaks in the cross-correlation functions used for TDOA estimation. Notably, the primary peak may not always correspond to the true line-of-sight propagation delay, thereby increasing the difficulty of identifying the correct TDOA and, consequently, the localization process.

B. Localization geometry

Fig. 1 depicts the geometry of the TDOA-based localization estimation methods. Δd_{12} is an example of one measure between pairs of microphones that can be explored to localize acoustic sources. This distance is estimated using the cross-correlation function [21].

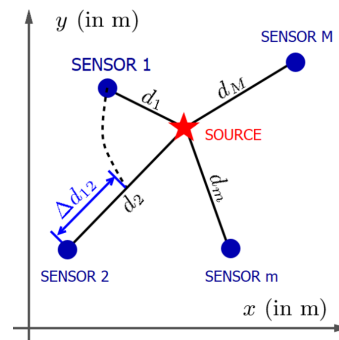


Fig. 1. Geometry of TDOA-based localization methods.

It should be noted that the target should not be distant from the center of the microphones, as discussed in [11], where a data-selective LS solution was proposed. This approach emphasizes that the accuracy of TDOA-based methods decreases when the source lies outside the microphones. Thus, ensuring that the emitter remains within or close to the microphone setup is essential for achieving reliable localization performance. This requirement does not imply prior knowledge of the emitter's position; instead, it serves as a deployment guideline, meaning that the monitored or protected region should be strategically placed inside or near the microphones to enhance estimation results.

C. Conventional localizations algorithms

The conventional least-squares (LS) solution can estimate the position of the source [11] using:

$$\hat{\mathbf{p}} = [\mathbf{I} \ \mathbf{0}](\mathbf{A}_1^T \mathbf{A}_1)^{-1} \mathbf{A}_1^T \mathbf{b}_1, \quad (1)$$

where \mathbf{A}_1 is defined as follows:

$$\mathbf{A}_1 = \begin{bmatrix} (\mathbf{p}_2 - \mathbf{p}_1)^T & \Delta d_{21} \\ (\mathbf{p}_3 - \mathbf{p}_1)^T & \Delta d_{31} \\ \vdots & \vdots \\ (\mathbf{p}_M - \mathbf{p}_1)^T & \Delta d_{M1} \end{bmatrix}, \quad (2)$$

where M denotes the total number of microphones in the acoustic array, \mathbf{p}_m is the position of the m^{th} microphone, and \mathbf{b}_1 is defined as:

$$\mathbf{b}_1 = \begin{bmatrix} b_{12} \\ b_{13} \\ \vdots \\ b_{1M} \end{bmatrix}, \quad (3)$$

which elements b_{1m} are given by:

$$b_{1m} = \frac{\|\mathbf{p}_m\|^2 - \|\mathbf{p}_1\|^2 - \Delta d_{m1}^2}{2}, \quad (4)$$

and the corresponding distance difference Δd_{ij} (in meters) is then calculated as:

$$\Delta d_{ij} = \frac{v\tau_{ij}}{f_s}, \quad (5)$$

where the TDOA between microphones i and j , denoted as τ_{ij} , is expressed in samples, v is the speed of sound, and f_s is the sampling frequency of the recorded signals.

The extended LS solution, also a TDOA-based solution, uses all the possible peaks of the cross-correlations. This solution [25] can be achieved using:

$$[\hat{\mathbf{p}}^T \ \hat{d}_1 \ \hat{d}_2 \ \dots \ \hat{d}_{M-1}] = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}, \quad (6)$$

where, in this case, \mathbf{A} is defined, with all distance measures from all cross-correlations as:

$$\mathbf{A} = \begin{bmatrix} (\mathbf{p}_2 - \mathbf{p}_1)^T & \Delta d_{21} & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ (\mathbf{p}_M - \mathbf{p}_1)^T & \Delta d_{M1} & 0 & 0 & \dots & 0 \\ (\mathbf{p}_3 - \mathbf{p}_2)^T & 0 & \Delta d_{32} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ (\mathbf{p}_M - \mathbf{p}_2)^T & 0 & \Delta d_{M2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ (\mathbf{p}_M - \mathbf{p}_1)^T & 0 & 0 & \dots & \Delta d_{M(M-1)} & \vdots \end{bmatrix}, \quad (7)$$

and vector \mathbf{b} is defined as:

$$\mathbf{b}_1 = [b_{12} \ b_{13} \dots b_{1M} \ b_{23} \ b_{24} \dots b_{2M} \dots b_{M(M-1)}]. \quad (8)$$

The key element for estimating localization is τ_{ij} , the TDOA between microphones i and j . In ideal conditions, without noise, τ_{ij} can be estimated accurately, leading to a precise calculation of the source location. However, in real-world scenarios, background noise can interfere with the

correct estimation of τ_{ij} , resulting in a noisy estimate $\hat{\tau}_{ij}$. Additionally, in environments with strong reflections, the true delay may not correspond to the maximum peak of the cross-correlation function, further complicating the localization accuracy, as denoted in Fig. 2.

D. Time Difference of Arrival Estimation Problem

Fig. 2 illustrates the cross-correlation issues that can arise in multipath environments. The cross-correlations shown pertain to the seventh and eighth microphones, with a maximum delay of 649 samples between the center (zero delay). Analyzing these cross-correlations can help researchers create new strategies to identify the most relevant peaks among numerous possibilities. Initially, a peak can be defined as a sample with a higher amplitude than its adjacent samples. However, this method may not effectively capture samples that span significant portions of the cross-correlations, potentially resulting in the selection of many closely spaced peaks. Therefore, an additional processing step is required to identify the peak whose amplitude is greater than that of its neighboring points.

III. DESIGNING AN ACOUSTIC-BASED TARGET LOCALIZATION SYSTEM

The ZCS condition is used to derive a cost function to minimize the DOA estimation error, and the TDOA selection algorithms use a least-squares solution to measure the coherence of TDOAs for the estimated localization $\hat{\mathbf{p}}$.

A. Integration with Localization Pipeline

To validate the effectiveness of the ZCS method in selecting meaningful subsets of TDOA measurements, we integrated ZCS with two distinct localization strategies: GTS and ES. The objective was to determine whether the TDOA subsets returned by ZCS contribute to more accurate localization results compared to using all available TDOAs. The localization pipeline was organized as follows:

Step 1 – TDOA Subset Selection using ZCS: ZCS was applied to the complete set of pairwise TDOA measurements to identify a subset that satisfies a predefined consistency criterion based on the behavior of ZCS in the residuals of estimated positions. This step is designed to eliminate outlier measurements and reduce the impact of multipath and synchronization errors.

Step 2 – Position Estimation: Two localization techniques were then independently applied to the selected TDOA subset. (1) ZCS + GTS: the Greedy TDOA Selection algorithm was used to compute the source position by minimizing the nonlinear least-squares error between the observed and predicted TDOA values. (2) ZCS + ES: an Exhaustive Search over a discretized grid of possible source locations was conducted to find the point that minimizes the TDOA residual error.

Step 3 – Performance Evaluation: The estimated positions obtained using both ZCS + GTS and ZCS + ES pipelines were compared against ground-truth source locations. The goal was to determine whether the TDOA subset identified by ZCS led to more accurate or robust localization.

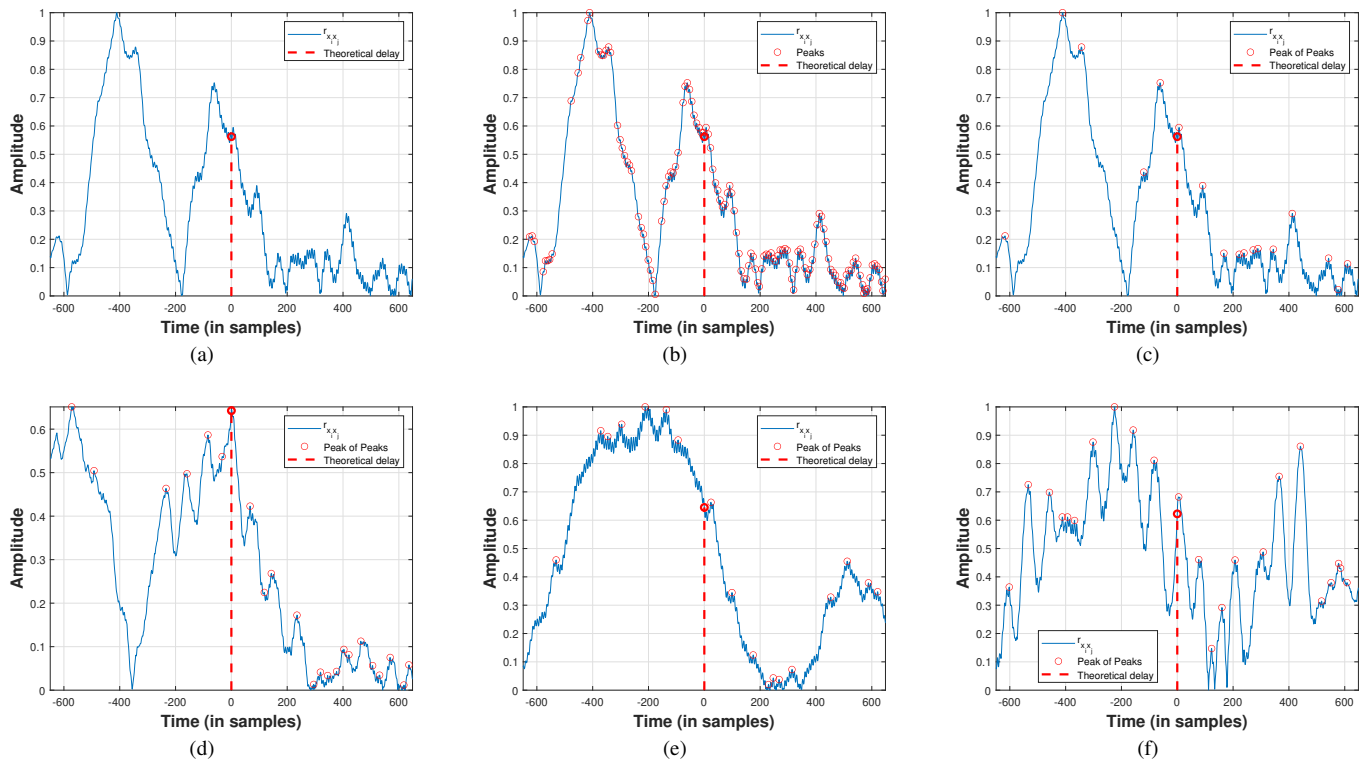


Fig. 2. Peak extraction process. (a) cross-correlation from microphones 7 and 8 with the drone hovering in position 5, (b) the same cross-correlation with the peaks detected, i.e., a peak is the sample that has a higher amplitude than the two nearest neighbors, (c) simple peak extraction method (sample amplitude greater than the other two closest peaks), the true peak estimation is the 4th highest peak, (d) example where the theoretical delay is the peak with the highest amplitude. (e) example where the theoretical delay is closer to the 7th peak (f) example where the theoretical delay is closer to the 9th peak

Additionally, this evaluation allowed us to assess whether ZCS consistently selects high-quality subsets that minimize the localization error across different algorithms.

Therefore, a total of seven TDOA-based localization strategies are evaluated. These strategies result from combining the above-mentioned different cost functions (LS or ZCS) with different TDOA selection mechanisms (ES and GTS). All combinations of techniques devised in this work are described below:

The Conventional LS estimates the source position directly using one microphone as reference and the respective TDOAs. No selection or robustness mechanism is applied.

The Extended LS is an enhanced LS formulation capable of incorporating all available TDOAs, even when they lead to an overdetermined system, improving solution stability and reducing sensitivity to measurement inconsistencies.

LS-ES applies the LS cost function to all C combinations of n elements. The solution yielding the minimum LS cost is selected. This method explores the complete solution space, offering high accuracy at the expense of increased computational complexity.

LS-GTS iteratively removes the TDOA that most reduces the LS cost function. Unlike ES, which tests every combination, GTS performs a structured and computationally efficient reduction of the TDOA set while preserving the most consistent delays.

ZCS-LS uses the ZCS condition, which is first applied to filter out delays that violate loop consistency. The reduced set

of TDOAs is then refined through LS localization, improving performance under multipath and reverberation conditions.

ZCS-LS-ES combines ZCS-based selection with exhaustive LS evaluation across all candidate peak combinations to consider not only the primary peaks, but also the secondary peaks of the cross-correlations. After considering primary and secondary peaks, an ES stage is applied to evaluate all C combinations of n elements.

ZCS-LS-GTS combines ZCS-based selection with exhaustive LS evaluation across all candidate peak combinations to consider not only the primary peaks, but also the secondary peaks of the cross-correlations. After considering primary and secondary peaks, a GTS stage is applied to iteratively evaluate and discard the delays that do not contribute to the LS cost function.

These strategies allow comparing the effectiveness of different TDOA selection criteria, as well as the benefits of incorporating acoustic loop consistency before LS-based estimation.

B. Optimization methods

The proposed acoustic-based target localization method builds upon TDOA-based optimization techniques. The approach utilizes TDOA values estimated between pairs of microphones, as illustrated in Fig. 1. These TDOAs, calculated using the cross-correlation function [21], serve as the primary cues for inferring the position of the acoustic source.

To address inaccuracies inherent in noisy and reverberant environments, this work introduces and evaluates three

algorithms that operate as TDOA selection and refinement strategies: GTS, ES, and ZCS. The core of these methods lies in selecting an enhanced subset of TDOA measurements that most likely correspond to the true acoustic delays between microphone pairs. This selection is guided by cost functions that compare the observed TDOAs to those predicted from theoretical delays, assessing the coherence of delays according to possible solutions in the localization space.

Our objective is to minimize the average squared error between observed and theoretical TDOAs, or, in the case of ZCS, enforce geometric consistency via closed-loop delay relationships. Given an initial set of TDOAs extracted from both primary and secondary peaks of the cross-correlation, the proposed algorithms apply distinct strategies to filter out unreliable measurements. The resulting refined TDOA subset is expected to better align with the actual drone position, thereby improving localization accuracy in such complex acoustic environments.

C. Greedy TDOA Selection

The GTS algorithm iteratively evaluates the impact of removing individual TDOA measurements on the accuracy of the localization estimates. Starting with the full set of $N = \binom{M}{2} = \frac{M(M-1)}{2}$ TDOAs, denoted by vector $\hat{\tau} = [\hat{\tau}_{12} \dots \hat{\tau}_{1M} \hat{\tau}_{23} \dots \hat{\tau}_{2M} \dots \hat{\tau}_{M(M-1)}]$, the algorithm removes one TDOA at a time in a greedy fashion. After each removal, it recalculates the average squared error between the estimated TDOAs, $\hat{\tau}_{ij}$, and the theoretical delays, τ_{ij} , using the cost function defined in Eq. 9.

$$\xi = \frac{1}{N} \|\hat{\tau} - \tau\|_2^2, \quad (9)$$

If the removal of a TDOA leads to a reduction in the cost function, that TDOA is discarded. This process continues iteratively and sequentially across all TDOAs until no further improvement is achieved by discarding additional delays. Although the GTS algorithm is computationally efficient, its heuristic nature may lead to suboptimal results. The GTS algorithm is similar to the Iterative Least Squares (ILS) algorithm proposed by Lyon Freire [24]; the latter removes the TDOA with the largest LS cost function in every iteration, while the former removes a TDOA only if the LS cost function decreases (a more conservative strategy). The complete procedure is detailed in Algorithm 1.

A simplified variant of the GTS algorithm described in Algorithm 1 is the one-pass GTS, which performs a single iteration over the N available time delays. In this approach, each delay is individually evaluated and removed only if its exclusion results in a decrease in the LS cost function. This variant offers the advantage of reduced computational complexity, as it avoids the need to evaluate all delays in multiple iterations. Additionally, it explores a different trajectory across the LS cost surface, potentially uncovering alternative solutions that the standard iterative GTS procedure might not reach.

Algorithm 1 GTS using LS cost - GTS-LS

Input: Full set of N TDOA estimations, $\hat{\tau}$, and minimal number of peaks desired p
Output: Reduced TDOA set τ^*
 $\tau^* \leftarrow \hat{\tau}$
 Compute initial mean LS cost: ξ
while cost decreases and $|\tau^*| > p$ **do**
 $cost_improved \leftarrow \text{false}$
 $best_cost \leftarrow \xi$
 for each active TDOA index k where $m_k = 1$ **do**
 Remove temporarily TDOA number k from $\hat{\tau}$, forming a new subset τ'
 Compute LS cost function ξ' using τ'
 if $\xi' < \xi$ **then**
 $\xi \leftarrow \xi'$
 $best_index \leftarrow k$
 $cost_improved \leftarrow \text{true}$
 end if
 Revert $m_k \leftarrow 1$
 end for
 if $cost_improved$ **then**
 Remove TDOA number k from τ^* , updating the τ^* TDOA vector
 $\xi \leftarrow best_cost$
 end if
end while
return τ^*

D. Exhaustive Subset Search

The ES algorithm systematically evaluates every possible combination of n TDOAs, $n < N$. For each candidate subset, the average squared error between the theoretical and observed delays is computed. The subset minimizing this cost function is selected as the optimal configuration. Although computationally intensive, this brute-force approach serves as a benchmark to assess the best performance of the localization algorithms. Algorithm 2 describes the ES procedure [26].

Algorithm 2 ES [26]

Input: Full set of N TDOAs $\hat{\tau}$ and theoretical delays τ , subset size n
Output: Optimal TDOA subset τ^* according to Eq. 9 criterion
 Generate all combinations \mathcal{C} of n elements from $\{1, 2, \dots, N\}$
 Initialize minimum cost: $\xi_{\min} \leftarrow \infty$
for each combination $\tau'_c \in \mathcal{C}$ **do**
 Compute cost ξ_c using Eq. 9 over τ'_c
 if $\xi_c < \xi_{\min}$ **then**
 $\xi_{\min} \leftarrow \xi_c$
 $\tau^* \leftarrow \tau'_c$
 end if
end for
return τ^*

E. Zero-Cyclic Sum

The ZCS method is based on the principle of cyclic consistency in TDOA measurements. In a geometrically consistent set of theoretical delays, the sum of TDOAs in any closed loop must be zero. ZCS evaluates all possible loops within the microphone network, composed of M elements, and computes the residual sum of each loop using the observed TDOAs. A cost function is derived from the deviation of these sums as follows:

$$f = \mathbf{f}^T \mathbf{f} = \|\mathbf{f}\|^2 \quad (10)$$

where $\mathbf{f} = \mathbf{D}\mathbf{v}$. The vector \mathbf{v} contains the observed TDOA values for all microphone pairs, and \mathbf{D} encodes the loop-consistency constraints among them. For the example with $M = 4$ microphones, \mathbf{D} is an $L \times N$ matrix that maps each candidate TDOA in \mathbf{v} to a set of cyclic sum equations, enabling efficient computation of the ZCS cost.

$$\mathbf{D} = \begin{bmatrix} 1 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & 1 \\ 1 & 0 & -1 & 1 & 0 & 1 \end{bmatrix}. \quad (11)$$

The (C) candidate delays (the number of primary and secondary peaks) for each cross-correlation function, $r_{x_i x_j}$, are the elements of each row of the data matrix, \mathbf{V} . For $M = 4$, which results in $N = 6$, the $N \times C$ matrix \mathbf{V} with all candidate delays, where each row corresponds to one cross-correlation and contains its most significant peaks, arranged in descending order of amplitude, is defined as follows:

$$\mathbf{V} = \begin{bmatrix} \hat{\tau}_{12,1} & \hat{\tau}_{12,2} & \hat{\tau}_{12,3} & \cdots & \hat{\tau}_{12,C} \\ \hat{\tau}_{13,1} & \hat{\tau}_{13,2} & \hat{\tau}_{13,3} & \cdots & \hat{\tau}_{13,C} \\ \hat{\tau}_{14,1} & \hat{\tau}_{14,2} & \hat{\tau}_{14,3} & \cdots & \hat{\tau}_{14,C} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \hat{\tau}_{34,1} & \hat{\tau}_{34,2} & \hat{\tau}_{34,3} & \cdots & \hat{\tau}_{34,C} \end{bmatrix}, \quad (12)$$

such that each possible set of delays is a combination of delays from each row, i.e., one per cross-correlation, denoted as the $N \times 1$ vector \mathbf{v}_c . It should be noted that the first column of matrix \mathbf{V} contains the primary peak of each cross-correlation, while the subsequent columns contain the secondary peaks.

The optimal subset of TDOAs is the one that minimizes the total loop inconsistency. This approach is grounded in geometry and does not require access to the ground-truth position, making it especially valuable in practical settings where the ground truth is unavailable. Algorithm 3 describes the ZCS-ES-LS procedure.

IV. EXPERIMENTS AND RESULTS

This section presents the experimental setup, procedures, and results. It includes both simulated data and real-world signals captured from a DJI Phantom 4 drone for the acoustic-based localization methods investigated in this study.

Algorithm 3 ZCS-ES-LS - Exhaustive search using ZCS and LS ($M=4$)

```

Compute all  $C$  candidate delays for every  $r_{x_i x_j}$ 
for  $k = 1 : N$  do
    Compute  $r_{x_i x_j}, i, j = 12$  to  $34$ 
    Obtain  $C$  candidate delays (larger peaks of  $r_{x_i x_j}$ )
     $\mathbf{V}_{k,:} \leftarrow [\tau_{ij,1} \ \tau_{ij,2} \ \cdots \ \tau_{ij,C}]$ 
end for
Create a combination of time delays and compute ZCS
for  $k = 1 : S$  do
     $\mathbf{P}_{:,k} \leftarrow \text{map } \tau_{ij,i} \text{ in } \mathbf{V}_{k,:}$ 
     $f = \text{ZCS}(\mathbf{P}_{:,k})$  Eq. 10
     $\mathbf{P}_{N+1,k} = f$ 
end for
Compute LS cost function of the  $Z$  time delay vectors with lowest ZCS
for  $k = 1 : Z$  do
     $\xi = \text{LS}(\mathbf{P}_{1:N,k})$  Eq. 9
     $\mathbf{P}_{N+2,k} = \xi$ 
end for
Choose the time delay vector with the lowest  $\xi$  (LS cost function)
    
```

A. Dataset

The dataset consists of five 20-second acoustic signals recorded using eight Behringer 800 microphones distributed around the signal of interest (SOI) emitter, i.e., an acoustic signal produced by the propellers of a DJI Phantom 4 drone. The signals were sampled at a frequency of $f_s = 44.1$ kHz.

Fig. 3 (a) depicts a 100 ms time-frequency spectrogram of SOI, the drone noise captured in the scenario represented by Fig. 4. Conversely, Fig. 3 (b) shows the spectrogram of background noise recorded, i.e., a recording in the absence of the drone noise, capturing environmental noise only. Its spectral content is weaker and more diffuse, with low-energy frequency components depicted by cooler colors (green to blue). The absence of well-defined harmonic structures or dominant peaks indicates a typical ambient noise floor.

Fig. 4 illustrates the setup used to record the noise generated by the drone. The recording positions of the drone are described in the Tab. I.

TABLE I
POSITIONS (COORDINATES IN METERS) OF THE HOVERING DRONE

#	1	2	3	4	5
x	1.6	1.6	1.6	4.1	4.1
y	3.3	2.3	1.3	3.3	2.3

The setup consists of $M = 8$ omnidirectional microphones distributed across the terrain to record the acoustic signals emitted by a hovering drone. The microphones are spatially deployed to allow localization estimation. In both simulation and real-world experiments, the DJI Phantom IV drone hovers at a fixed altitude of approximately 30 cm from the floor, generating a consistent acoustic signature predominantly in the low-frequency band due to its propeller noise.

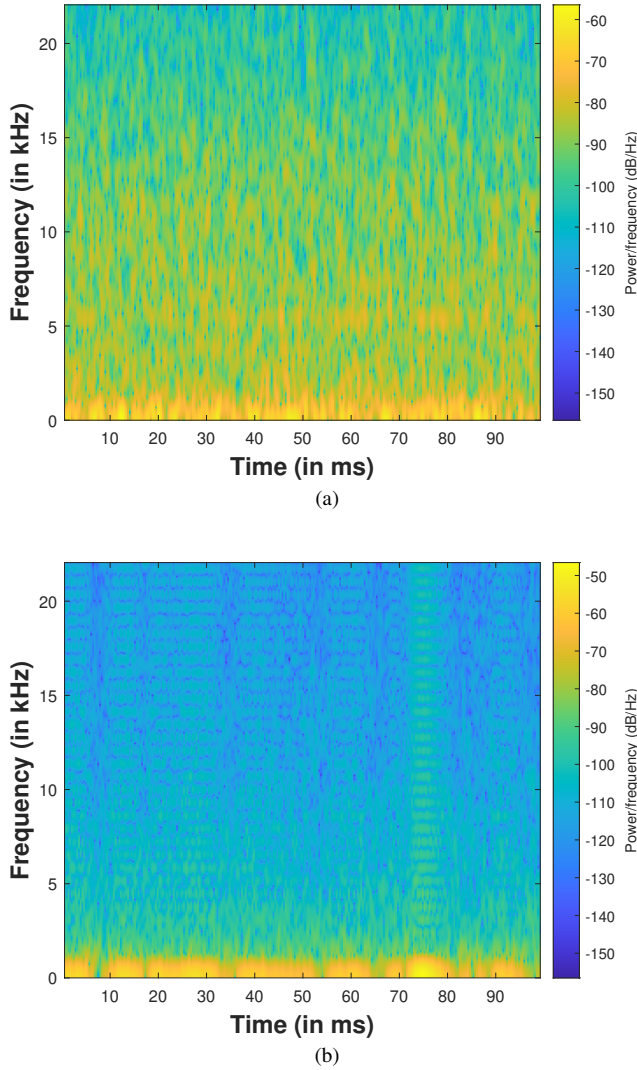


Fig. 3. Phantom IV drone spectrogram: (a) drone noise (b) background noise, i.e., all noises surrounding microphones except the drone noise.

Given $M = 8$ microphones, a total of $N = \binom{8}{2} = 28$ microphone pairs are formed. One TDOA measurement is estimated from each pair of microphones forming a vector $\hat{\tau}_{28 \times 1}$.

This setup enables the assessment of different TDOA selection algorithms (GTS, ES, and ZCS) under both controlled (simulated) and realistic (recorded) conditions, providing a robust framework for evaluating and improving the reliability of drone localization based on passive acoustic sensing. All recordings were made with the microphones placed in the coordinates presented in Tab. II.

TABLE II
MICROPHONE COORDINATES (IN METERS)

Mic	1	2	3	4	5	6	7	8
x	0.0	0.0	1.6	6.6	8.2	8.2	6.6	1.6
y	0.6	4.0	4.6	4.6	0.4	0.6	0.0	0.0

To quantify the performance of the localization methods,

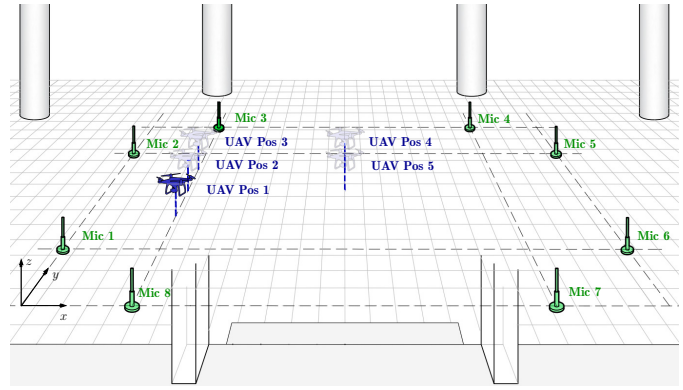


Fig. 4. Set-up for the drone noise recording.

the localization error, e_{loc} , is defined as the Euclidean distance between the true position of the drone, \mathbf{p} , and the estimated position, $\hat{\mathbf{p}}$:

$$e_{loc} = \|\mathbf{p} - \hat{\mathbf{p}}\|_2. \quad (13)$$

For multiple drone positions (K), the mean localization error (MLE) is computed as

$$\text{MLE} = \frac{1}{K} \sum_{k=1}^K \|\mathbf{p}_{\text{true},k} - \hat{\mathbf{p}}_k\|_2, \quad (14)$$

and the corresponding standard deviation (σ_e) is given by

$$\sigma_e = \sqrt{\frac{1}{K} \sum_{k=1}^K (\|\mathbf{p}_{\text{true},k} - \hat{\mathbf{p}}_k\|_2 - \text{MLE})^2}. \quad (15)$$

B. Simulation results

In the simulated environment, white Gaussian noise is added to each TDOA measurement to emulate realistic conditions with noisy TDOA estimations. This setup enables the assessment of the robustness of the localization algorithms under controlled conditions of noise and outliers.

Fig. 5 depicts the behavior of localization error using the extended LS solution under TDOAs with additive noise and introducing one and two outliers using all TDOAs. The simulated TDOA values were derived from the actual microphone geometry, assuming the drone was hovering at position five. Fig. 5 (a) shows the effect as the additive Gaussian noise increases, the localization error grows gradually, reflecting a predictable degradation in estimation accuracy due to uncertainty in the TDOAs. Fig. 5 (b) denotes that introducing one outlier slightly increases the estimation error, i.e., this demonstrates that this system of equations has sufficient redundancy, and this one outlier is diluted among the many correct delays. Fig. 5 (c) depicts the error rising sharply with low additive TDOA error due to the presence of two outliers. These results highlight the importance of TDOA selection mechanisms that are not only noise-aware but also resilient to outliers, as even a small number of outliers can have a disproportionately large effect on the localization accuracy. The mean TDOA estimation error over 10 samples is approximately 0.5 m; however, the presence of two outliers increases the average error to 1.4 m, as illustrated in Fig. 5 (c).

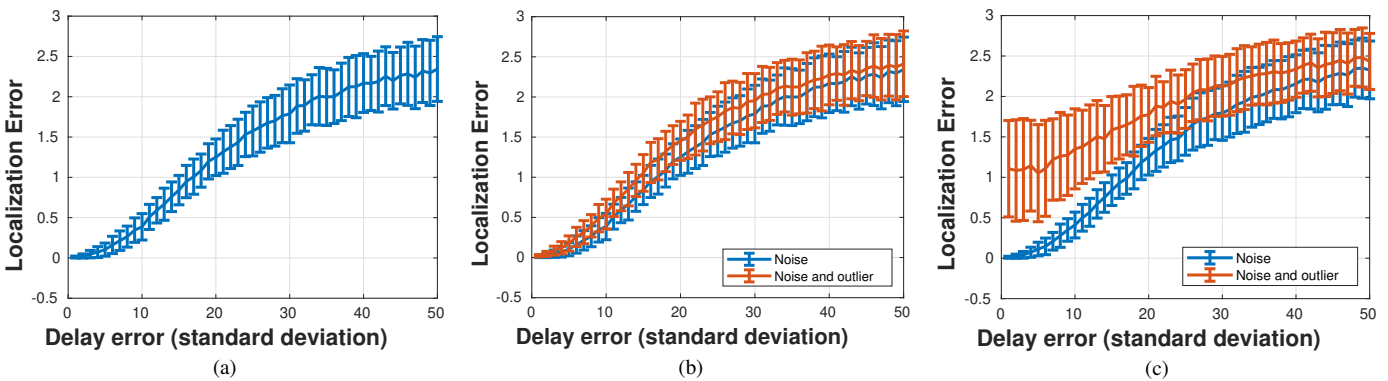


Fig. 5. Results obtained using simulated delays with additive noise, based on the dataset’s geometry and assuming the drone is hovering at position five. (a) Localization error as a function of the standard deviation of the additive noise. (b) Localization error when introducing a single outlier, while varying the noise standard deviation. (c) Localization error when introducing two outliers, while varying the noise standard deviation.

Overall, the results underline two key insights. First, while additive noise alone gradually degrades localization accuracy, it is manageable up to a certain threshold. Second, even a small number of outliers can severely increase the localization error. These findings reinforce the importance of incorporating strategies such as the proposed ZCS-based cost function and delay selection algorithms to mitigate the influence of erroneous TDOA values and ensure more reliable localization performance under real-world conditions.

C. Experimental results

In this section, we provide the performance of the TDOA-based LS solutions for drone localization. We aim to highlight the strengths and limitations of each method in terms of accuracy and practical applicability for the drone experiment. The LS method, which estimates the source position based on TDOA measures between microphones, is straightforward and does not require a training phase.

We also provide a comparative analysis with a Neural Networks (NNs) approach that can be used to address the challenges posed by strong reverberation in indoor environments. Apolinário et al. [12] developed an NN-based solution to estimate the position of an acoustic emitter. Unlike TDOA filtering techniques, the NN approach exploits the full reverberation fingerprint embedded in the received signals.

The NN method uses, for each pair, the ten most prominent peaks. These peaks are extracted from the cross-correlation function, capturing both the time delays and their corresponding magnitudes. These values form the feature vector \mathbf{f} . It should be noted that peaks that correspond to unrealistic delays (e.g., distances beyond twice the room’s maximum dimension) are excluded. The NN approach uses features extracted from the environment’s reverberation fingerprint, thereby allowing adaptation and improved performance under challenging scenarios with complex reverberation patterns. By evaluating all methods under the same conditions, we provide a better understanding of their relative performance and the trade-offs involved in choosing one technique over another.

The microphones are placed on tripods at fixed, known positions throughout a strong reflection environment. The drone is commanded to hover at known positions. Signals

are recorded synchronously across all channels using a multi-channel audio interface.

Fig. 6 depicts the relative frequency of TDOAs used to estimate the localization of the drone using the GTS technique, i.e., after iteratively discarding TDOAs that do not contribute to reducing the cost function. It should be noted that 15 is the largest number of TDOAs used. The minimum number of TDOAs ranges from 10 to 21, indicating that while the selection mechanism enhances the results, there exists a minimum number of TDOAs required to achieve the lowest localization error.

Tab. III presents the localization errors for drone position estimation using the seven different strategies: the Conventional LS, Extended LS, LS-ES, LS-GTS, ZCS, ZCS-LS-ES, and ZCS-LS-GTS. These methods are compared to the NN approach. The results are reported for the five different drone positions, with the error expressed as the mean localization error and the standard deviation. Tab. III presents results for these techniques. Conventional LS achieved localization errors ranging from 2.25 ± 0.84 to 3.43 ± 1.10 meters across drone positions, and Extended LS errors between 1.89 ± 0.60 and 3.30 ± 0.75 meters.

The LS-ES strategy applies the LS cost function to all possible combinations of candidate delays. This brute-force approach explores the complete solution space, but at the expense of higher computational complexity. The LS-GTS method iteratively removes delays that reduce the LS cost function. This makes it computationally lighter than LS-ES while still improving robustness compared to conventional LS. Tab. III exhibits that LS-ES and LS-GTS consistently produce good results, particularly at position 4, where the localization errors are as low as 0.68 ± 0.39 and 0.84 ± 0.40 meters, respectively. These methods benefit from the cost function described in Eq. 9, which enables rejecting noisy TDOAs.

The ZCS approach utilizes both primary and secondary correlation peaks, selecting those that satisfy loop-consistency constraints. ZCS-LS-ES combines the ZCS-LS with an exhaustive exploration for primary and secondary peaks, while ZCS-LS-GTS provides a good balance between accuracy and computational efficiency. These techniques further improve accuracy, especially at position 4, where ZCS-LS-ES achieves

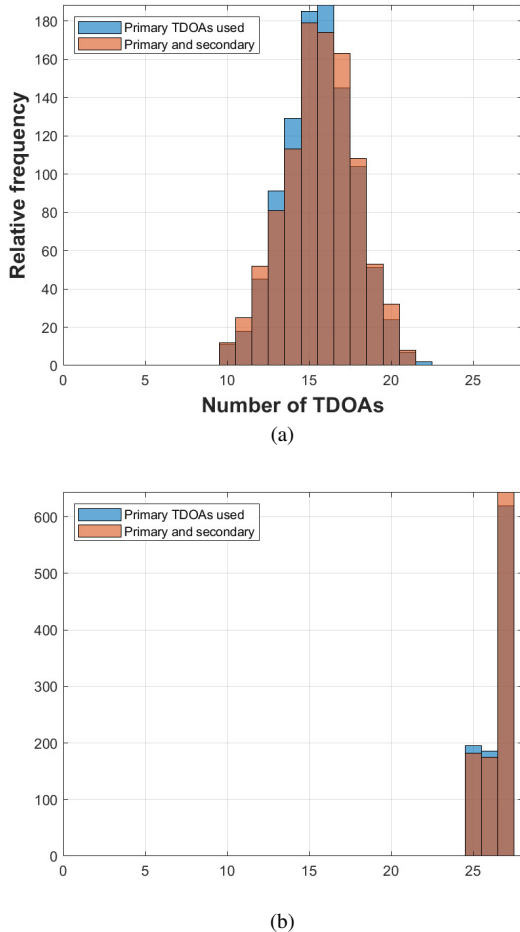


Fig. 6. Number of peaks used. (a) ILS approach, and (b) GTS technique.

the lowest error overall of 0.55 ± 0.35 meters. This result demonstrates the effectiveness of using both loop-consistency and LS cost minimization for delay selection.

The results regarding the NN were obtained employing a hold-out approach, i.e., random splitting of the dataset in 70% for training, 15% for validation, and 15% for testing, for each number of observations. Additionally, the NN was implemented as a fitting application, using the Levenberg-Marquardt training algorithm for the initial fitting procedure. Finally, the NN method outperforms all other strategies in terms of both accuracy (for drone in positions 1, 2, and 3) and consistency, with localization errors ranging from 1.02 to 1.69 meters and very low standard deviations (± 0.02) across all positions. Despite this superior performance, the NN model requires a prior dataset for training, which may not always be feasible in unknown environments.

The drone’s lack of GPS stabilization introduces slight motion during recordings, which likely contributes to the variation in estimation error across positions. The geometry and acoustic reflectivity of the environment may also explain why certain positions, like position 3, consistently exhibit higher errors across most algorithms.

The results presented in Fig. 7 highlight how the number of selected TDOAs affects the accuracy and consistency of

source localization. Each figure presents 28 candidate estimates corresponding to the lowest LS cost values, with the estimate that achieves the minimum LS cost function distinctly marked by a black plus symbol. Fig. 7 (a) shows the results using only 5 delays (ES(5)). This configuration leads to low precision, with multiple scattered estimates due to the large number of possible combinations of delays that are mapped to a large number of different solutions. In contrast, Fig. 7 (b) with 25 delays (ES(25)) shows a significant improvement, producing three dense and close clusters and reducing the spread of estimations. It should be noted that the estimate that minimizes the LS cost function does not correspond to the true location, indicating that a higher number of delays alone does not guarantee accuracy. Finally, Fig. 7 (c) using 27 delays (ES(27)) results in a single dominant cluster with only two estimates away from the cluster. Here, the most accurate localization estimate also minimizes the LS cost function, demonstrating that when nearly all relevant delays are included, the LS criterion becomes a reliable indicator of the true source position.

V. CONCLUSIONS

Accurate acoustic source localization for unmanned aerial vehicles remains a significant challenge, particularly in highly reverberant environments. With this focus, we employed and evaluated solutions based on TDOA measurements, employing both LS and ZCS cost functions to optimize the localization process. These approaches were also compared to an NN model trained on reverberation fingerprints. Simulations and real-world experiments were conducted to validate the effectiveness of the proposed techniques. Simulation results showed that while the number of selected TDOAs ranged from 10 to 21, selecting either a few or a large number did not always guarantee improved localization. In fact, a careful balance must be struck between quantity and quality.

The LS-based methods demonstrated consistent improvements in localization accuracy by enabling the rejection of noisy or inconsistent TDOAs through the cost function minimization. Among the evaluated strategies, the ZCS-LS-ES method achieved the lowest localization error of 0.55 ± 0.35 meters at position 4. The ZCS approach, which incorporates secondary peaks from the cross-correlation, also showed promise in handling reverberant conditions, although it was more effective when combined with LS filtering. In the actual drone experiment, the NN method achieved more consistency, achieving average errors ranging from 1.02 to 1.69 meters. However, this performance comes at the cost of needing prior training data, which may not always be available in unknown or dynamically changing environments. The LS consistently improves results and has the significant advantage of not requiring prior training. The methods described in this work are suitable for integration into real-time embedded systems deployed at ground stations. Future work will focus on designing advanced TDOA-based localization cost functions, aiming to increase robustness against noise, multipath, and spurious delay estimates. Also, developing adaptive or learning-based cost functions could further reduce localization bias and improve convergence in challenging acoustic environments. In

TABLE III
DRONE LOCALIZATION ERROR AND STANDARD DEVIATION (IN M)

Drone Position	1	2	3	4	5
Conventional LS	2.36±0.64	2.25±0.84	2.78±0.94	3.43±1.10	2.86±0.87
Extended LS	2.22±0.35	2.40±0.67	3.30±0.75	1.89±0.60	1.98±0.25
LS-ES	2.04±0.32	2.22±0.37	2.83±0.37	0.68±0.39	1.21±0.33
LS-GTS	1.87±0.37	2.08±0.45	2.73±0.43	0.84±0.40	1.38±0.37
ZCS-LS	2.61±0.38	3.05±0.57	3.97±0.61	1.17±0.37	1.92±0.15
ZCS-LS-ES	2.11±0.35	2.27±0.37	2.85±0.38	0.55±0.35	1.07±0.36
ZCS-LS-GTS	2.00±0.39	2.20±0.43	2.79±0.44	0.66±0.40	1.21±0.41
Neural Network	1.28±0.01	1.02±0.02	1.56±0.02	1.69±0.02	1.50±0.02

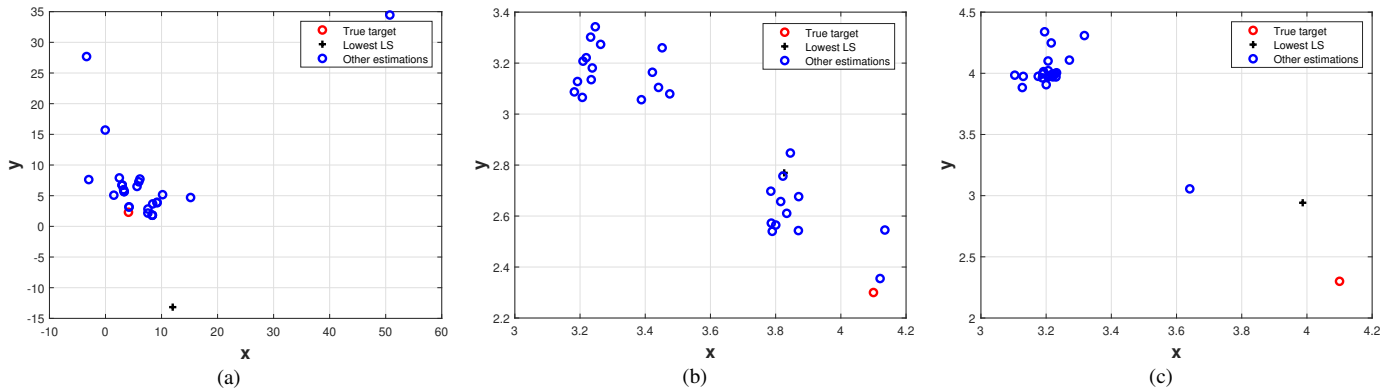


Fig. 7. Localization results obtained using the actual delays. (a) ES(5) low precision, which represents multiple solutions to the large number of possible combinations using $n = 5$ delays. (b) ES(25) higher precision forming three dense clusters and one estimation, which does not minimize the LS cost function with minimum error of localization. (c) ES(27) with 28 estimations forming one dense cluster and the other two estimations. Note that the estimation that minimizes the localization error is also the one that minimizes the LS cost function.

parallel, hybrid localization frameworks that combine acoustic TDOA estimation with visual sensing should be explored. The integration of cameras and computer vision techniques, such as object detection and visual tracking, can provide complementary estimates. Finally, real-time implementations on embedded hardware can enable the development of a functional prototype of the proposed system.

REFERENCES

[1] Q. Zhang, Y. Wang, D. Li, J. Xie, K. Tao, P. Hu, J. Zhou, H. Chang, and Y. Fu, “Multifunctional and wearable patches based on flexible piezoelectric acoustics for integrated sensing, localization, and underwater communication,” *Advanced Functional Materials*, vol. 33, no. 2, p. 2209667, 2023, doi: 10.1002/adfm.202209667.

[2] J. Melchiorre, A. Manuello Bertetto, M. M. Rosso, and G. C. Marano, “Acoustic emission and artificial intelligence procedure for crack source localization,” *Sensors*, vol. 23, no. 2, p. 693, 2023, doi: 10.3390/s23020693.

[3] K. K. Gola and S. Arya, “Underwater acoustic sensor networks: Taxonomy on applications, architectures, localization methods, deployment techniques, routing techniques, and threats: A systematic review,” *Concurrency and Computation: Practice and Experience*, vol. 35, no. 23, p. e7815, 2023, doi: 10.1002/cpe.7815.

[4] R. P. Fernandes, A. M. C. R. Borzino, A. L. L. Ramos, and J. A. Apolinário, Jr., “Investigating the potential of UAV for gunshot DoA estimation and shooter localization,” in *XXXIV Simpósio Brasileiro de Telecomunicações*

e *Processamento de Sinais*. Santarém, PA, Brazil: SBrT, 30 August - 2 September 2016, pp. 383–387, doi: 10.14209/sbrt.2016.52.

[5] R. P. Fernandes, J. A. Apolinário, Jr., and A. L. L. Ramos, “Bearings-only aerial shooter localization using a microphone array mounted on a drone,” in *2017 IEEE 8th Latin American Symposium on Circuits & Systems (LASCAS)*. Bariloche, Argentina: IEEE, 20-23 February 2017, pp. 1–4, doi: 10.1109/LASCAS.2017.7948081.

[6] R. P. Fernandes, A. L. L. Ramos, and J. A. Apolinário, Jr., “Airborne DoA estimation of gunshot acoustic signals using drones with application to sniper localization systems,” in *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security, Defense, and Law Enforcement Applications XVI*, vol. 10184. Anaheim, CA, United States: SPIE, 5 May 2017, pp. 51–57, doi: 10.1117/12.2262782.

[7] F. G. Serrenho, J. A. Apolinário, Jr., A. L. L. Ramos, and R. P. Fernandes, “Gunshot airborne surveillance with rotary wing UAV-embedded microphone array,” *Sensors*, vol. 19, no. 19, p. 4271, 2019, doi: 10.3390/s19194271.

[8] M. N. de Sousa, R. Sant’Ana, R. P. Fernandes, J. C. Duarte, J. A. Apolinário, Jr., and R. Thomä, “Improving the performance of a radio-frequency localization system in adverse outdoor applications,” *Journal on Wireless Communications and Networking*, vol. 1, p. 123, may 2021, doi: 10.1186/s13638-021-02001-6.

[9] M. Hanzla, S. Ali, and A. Jalal, “Smart traffic monitoring through drone images via Yolov5 and Kalman filter,” in

- 2024 5th International Conference on Advancements in Computational Sciences (ICACS). IEEE, 2024, pp. 1–8, doi: 10.1109/ICACS60934.2024.10473259.
- [10] M. Kowshika, M. Ooviya, B. Pavithradevi, K. Rashika, and S. Sampath Kumar, “Unmanned aerial systems in search and rescue: a comprehensive review and future directions,” in *2024 5th International Conference on Mobile Computing and Sustainable Informatics (ICM-CIS)*. IEEE, 2024, pp. 15–18, doi: 10.1109/ICM-CIS161536.2024.00008.
- [11] J. A. Apolinário Jr., H. Yazdanpanah, A. S. Nascimento Filho, and M. L. R. de Campos, “A Data-selective LS Solution to TDOA-based Source Localization,” in *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 4400–4404, doi: 10.1109/ICASSP.2019.8682664.
- [12] J. A. Apolinário Jr., A. L. L. Ramos, R. P. Fernandes, J. C. Duarte, and M. N. de Sousa, “Exploiting reverberation fingerprint for a neural network based acoustic emitter localization,” in *2024 10th International Conference on Control, Decision and Information Technologies (CoDIT)*. Valletta, Malta: IEEE, 1-4 July 2024, pp. 1–6, doi: 10.1109/CoDIT62066.2024.10708420.
- [13] Y. He, W. Wang, L. Mottola, S. Li, Y. Sun, J. Li, H. Jing, T. Wang, and Y. Wang, “Acoustic localization system for precise drone landing,” *IEEE Transactions on Mobile Computing*, vol. 23, no. 5, pp. 4126–4144, 2023, doi: 10.1109/TMC.2023.3288299.
- [14] X. Wang and D. Hu, “Distributed self-localization for acoustic transceiver networks,” *IEEE Signal Processing Letters*, vol. 30, pp. 553–557, 2023, doi: 10.1109/LSP.2023.3274414.
- [15] Y. Sun, W. Wang, L. Mottola, J. Zhang, R. Wang, and Y. He, “Indoor drone localization and tracking based on acoustic inertial measurement,” *IEEE Transactions on Mobile Computing*, 2023, doi: 10.1109/TMC.2023.3335860.
- [16] Z. Zhao and N.-Z. Chen, “Spatial-temporal graph convolutional networks (STGCN) based method for localizing acoustic emission sources in composite panels,” *Composite Structures*, vol. 323, p. 117496, 2023, doi: 10.1016/j.compstruct.2023.117496.
- [17] Z. Xiang, H. Jin, Y. Wang, P. Ren, L. Yang, and B. Xu, “A novel modified symmetric nested array for mixed far-field and near-field source localization,” *Remote Sensing*, vol. 16, no. 15, p. 2732, 2024, doi: 10.3390/rs16152732.
- [18] W. Jia and Y. Feng, “A localization method for multistation passive radar based on DOA and TOA,” in *International Conference on Signal Processing and Communication Security (ICSPCS 2024)*, vol. 13222. SPIE, 2024, pp. 226–233, doi: 10.1117/12.3038812.
- [19] L. Jiang, N. Keerativoranan, T. Matsumoto, and J.-i. Takada, “A distributed sensor-based recursive framework for doa estimation and geolocation,” *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3424216.
- [20] Y. Yang, J. Zheng, H. Liu, K. Ho, Z. Yang, and S. Gao, “Optimal sensor placement and velocity configuration for TDOA-FDOA localization and tracking of a moving source,” *IEEE Transactions on Aerospace and Electronic Systems*, 2024, doi: 10.1109/TAES.2024.3430238.
- [21] R. P. Fernandes, J. A. Apolinário, Jr., and J. M. de Seixas, “A reduced complexity acoustic-based 3D DoA estimation with zero cyclic sum,” *Sensors*, vol. 24, no. 7, p. 2344, 2024, doi: 10.3390/s24072344.
- [22] R. P. Fernandes, J. A. Apolinário, Jr., and J. M. de Seixas, “Enhancing TDE-based drone DoA estimation with genetic algorithms and zero cyclic sum,” in *XVI Brazilian Conference on Computational Intelligence*. Salvador, BA, Brazil: Congresso Brasileiro de Inteligência Computacional, 8-11 October 2023, doi: 10.21528/CBIC2023-115.
- [23] A. M. C. R. Borzino, J. A. Apolinário Jr., and M. L. R. de Campos, “Consistent DOA estimation of heavily noisy gunshot signals using a microphone array,” *IET Radar, Sonar Navigation*, vol. 10, no. 9, pp. 1519–1527, 2016, doi: 10.1049/iet-rsn.2016.0015.
- [24] I. Lyon Freire, “Robust direction-of-arrival by matched-lags, applied to gunshots,” *The Journal of the Acoustical Society of America*, vol. 135, no. 6, pp. EL246–EL251, 2014, doi: 10.1121/1.4874223.
- [25] M. Khalaf-Allah, “An extended closed-form least-squares solution for three-dimensional hyperbolic geolocation,” in *2014 IEEE Symposium on Industrial Electronics Applications (ISIEA)*, Sept. 2014, pp. 7–11, doi: 10.1109/ISIEA.2014.8049862.
- [26] A. M. C. R. Borzino, J. A. A. Jr., and M. L. R. de Campos, “Robust DOA estimation of heavily noisy gunshot signals,” in *2015 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Brisbane, QLD, Australia, Apr. 2015, pp. 449–453, doi: 10.1109/ICASSP.2015.7178009.



Brazilian Telecommunications Society (SBRT).

Rigel Procópio Fernandes is a computer engineer at the Brazilian Navy. He holds an M.Sc. in informatics from the Federal University of Rio de Janeiro (2014), an M.Sc. in defense engineering from the Military Institute of Engineering (2017), and a D.Sc. in defense engineering from the Military Institute of Engineering (2025). His research interests include acoustic-based techniques to detect and estimate DOA and localization, underwater acoustics, military applications of digital signal processing, and computer vision. Dr. Rigel is a member of the



José A. Apolinário Jr graduated from the Military Academy of Agulhas Negras (AMAN), Brazil, in 1981 and received the following degrees in electrical engineering: B.Sc. from the Military Institute of Engineering (IME), Brazil, in 1988, M.Sc. from the University of Brasília (UnB), Brazil, in 1993, and D.Sc. from the Federal University of Rio de Janeiro (COPPE/UFRJ), Rio de Janeiro, Brazil, in 1998. He is currently a full professor in the Department of Electrical Engineering at IME, where he has previously served as Head of the Department and Vice-Rector for Study and Research. His research interests comprise many aspects of digital signal processing, including adaptive filtering, speech and audio processing, emitter localization, and array signal processing. Prof. Apolinário is a senior member of the IEEE and a senior member of the Brazilian Telecommunications Society (SBRT).



Julio Cesar Duarte is a professional with academic training and experience in the field of Computer Engineering. He graduated from the Military Engineering Institute (IME) in 1998, and subsequently obtained a master's degree in Computer Science from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio) in 2003 and a doctorate from the same institution in 2009. He completed his training by doing a post-doctoral internship at PUC-Rio in 2021. He is currently a professor in the Postgraduate Program in Systems and Computing and Dean of

Teaching and Research at IME. His academic and professional work is marked by a multidisciplinary approach, with a significant emphasis on the development of computer systems. His experience covers several areas of computing, including artificial intelligence, machine learning, deep learning, and Portuguese natural language processing.



José M. de Seixas holds a degree in Mathematics from the Pontifical Catholic University of Rio de Janeiro (1979), a degree in Electrical Engineering from the Pontifical Catholic University of Rio de Janeiro (1979), a master's degree in Electrical Engineering from the Federal University of Rio de Janeiro (1983), and a D.Sc. in Electrical Engineering from the Federal University of Rio de Janeiro (1994). He is currently a Full Professor at the Federal University of Rio de Janeiro. He has experience in Electrical Engineering, with emphasis

on Electrical, Magnetic, and Electronic Circuits. His main areas of focus include computational intelligence, high-energy calorimetry, sonar technology, signal processing, and electronic instrumentation.