Autonomous Sensing UAV for Accurate Multi-User Identification and Localization in Cellular Networks

Niccolò Paglierani *Student Member, IEEE*, Francesco Linsalata *Member, IEEE*,
Vineeth Teeda *Student Member, IEEE*, Davide Scazzoli *Member, IEEE*, Maurizio Magarini, *Member, IEEE*

***Abstract*—This paper presents an autonomous sensing framework for identifying and localizing multiple users in Fifth Generation (5G) networks using an Unmanned Aerial Vehicle (UAV) that is not part of the serving access network. Unlike conventional aerial serving nodes, the proposed UAV operates passively and is dedicated solely to sensing. It captures Uplink (UL) Sounding Reference Signals (SRS), and requires virtually no coordination with the network infrastructure. A complete signal processing chain is proposed and developed, encompassing synchronization, user identification, and localization, all executed onboard UAV during flight. The system autonomously plans and adapts its mission workflow to estimate multiple user positions within a single deployment, integrating flight control with real-time sensing. Extensive simulations and a full-scale low-altitude experimental campaign validate the approach, showing localization errors below 3 m in rural field tests and below 8 m in urban simulation scenarios, while reliably identifying each user. The results confirm the feasibility of infrastructure-independent sensing UAVs as a core element of the emerging Low Altitude Economy (LAE), supporting situational awareness and rapid deployment in emergency or connectivity-limited environments.**

***Index Terms*—Low Altitude Economy (LAE), Unmanned Aerial Vehicle (UAV), Autonomous systems, Localization, Multi-Users, 5G.**

I. Introduction

THE LOW-ALTITUDE ECONOMY (LAE) refers to the emerging ecosystem of aerial services operating within the lower portion of the airspace close to ground level. It is rapidly expanding thanks to advances in Unmanned Aerial Systems (UASs), the evolution of regulatory frameworks, and the deployment of Fifth Generation (5G) and future Sixth Generation (6G) networks [1], [2].

Recent progress in wireless communication, computing, and flight autonomy has transformed Unmanned Aerial Vehicles (UAVs) from simple mobile platforms into intelligent network nodes capable of communication, sensing, and onboard computation. This convergence is turning the low-altitude domain into a connected layer where aerial and terrestrial components cooperate to enable time-sensitive and data-driven services [1].

Within this environment, mission-oriented applications, such as autonomous navigation, drone-based logistics, precision agriculture, and infrastructure inspection, critically depend on reliable wireless connectivity and real-time signal processing [3]. A particularly relevant class of applications involves emergency and search-and-rescue operations, where drones provide rapid coverage, robust connectivity, and accurate localization even when terrestrial infrastructure is unavailable or impaired [4]–[6].

Conventional UAV assisted localization schemes typically rely on aerial serving nodes that establish links with ground User Equipments (UEs) [7]–[9]. In disaster response scenarios, however, the main goal of such UAV base stations is to maximize network coverage and connectivity rather than optimize geometric conditions for localization. Consequently, operational procedures that ensure adequate communication performance often do not align with those needed for high-accuracy user position estimation [10].

Additionally, maintaining Line-of-Sight (LoS) connectivity with all ground users, which is an essential requirement for reliable localization [11], [12], is generally infeasible for a single UAV acting as an aerial base station. Although the UAV can maneuver to establish LoS with individual UEs, geometric constraints, occlusions, and flight limitations prevent simultaneous coverage of multiple spatially distributed users. To overcome this limitation, one or more UAVs can be deployed as dedicated sensing nodes operating independently from the communication infrastructure. These non-serving UAVs do not participate in user connectivity or data relaying; instead, they act as cooperative receivers that capture Uplink (UL) signals from connected UEs and perform localization solely based on signal measurements. Because their role is limited to reception and onboard processing, such sensing UAVs can be realized with relatively low hardware complexity, enabling scalable and cost-effective deployment.

This architectural *decoupling* increases flexibility in UAV placement, enables broader spatial coverage, and can substantially improve localization accuracy. However, it also introduces technical challenges. Sensing UAVs are not time or frequency synchronized with the cellular network, so they have no knowledge of users' transmission timing or resource allocation. As a result, they cannot predict user transmissions, which complicates signal detection, UE identification, and position estimation. In addition, UL transmissions from multiple UEs may overlap in time and frequency, creating interference and degrading received signal quality [13], [14].

A notable limitation of existing studies adopting such non-serving or sensing-oriented UAV architectures lies in the highly idealized assumptions made at the signal level. Prior works often consider perfectly known pilot or beacon sequences with predefined resource allocations, rather than capturing and processing actual 5G UL waveforms [11], [15],

The authors are with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milan, Italy (e-mail name.surname@polimi.it). The authors thank the Polus Tech team for their support on the research subject and Dr. Christian Mazzucco for the deep discussion on the topics.

[16]. While such abstractions simplify algorithmic evaluation, they do not address how localization signals can be extracted in real network conditions, where synchronization, scheduling, and multiplexing constraints must be explicitly considered.

In this work, we demonstrate the *autonomous synchronization, localization and identification* of *multiple UEs* within a single flight using a *non-serving UAV* capable of *on-board processing actual 5G UL signals*, requiring only *minimal initial coordination* with the network and *no mission time control* from it.

Paper Contributions The main contributions of this paper can be summarized as follows.

- We propose a fully autonomous framework that enables a single non-serving UAV to identify and localize multiple 5G users using only UL Sounding Reference Signal (SRS). Unlike aerial serving nodes approaches, the UAV operates purely as a passive receiver, requiring no network synchronization or control. To the best of the authors' knowledge, this is the first demonstration of practical multi-user identification and localization from real 5G UL signals captured by a UAV that is not part of the access network.
- We introduce a new synchronization and user identification pipeline that exploits the SRS structure and cyclic shifts, combining coarse and fine timing synchronizations with a Matching Pursuit frequency extraction method to separate multiple users transmitting in the same band. The proposed method was shown to be robust to both delay and power spreads between UEs, as evidenced by our simulation and experimental results.
- We develop a UAV centric weighted mean shift localization algorithm that operates on SRS derived metrics collected along the UAV trajectory. Despite relying on narrowband (1.4 MHz) SRS and lightweight hardware, the approach achieves high accuracy, below 3 m in rural and below 8 m in urban conditions, outperforming state-of-the-art [11] Angle of Arrival (AoA)- and [7] Time Difference of Arrival (TDoA)-based methods by approximately 5 m and 6 m, respectively. The method is validated not only through extensive simulations but also through an extensive outdoor experimental campaign.
- We design and integrate an onboard platform combining a dual RX ADALM Pluto Software Defined Radio (SDR), a Pixhawk flight controller, and an NVIDIA Jetson Orin NX, enabling all signal processing, mission control, and waypoint updates to be performed autonomously during flight. In this architecture, *autonomy* refers to the UAV's ability to execute the sensing workflow without external intervention, including real-time UL signals processing, UE identification, position estimation, and the consequent update of navigation commands directly onboard.

Paper Organization The remainder of the paper is structured as follows. Section II reviews related works on UAV-based localization. Section III presents the system model, including UAV and user configurations and the SRS structure. In Sec. IV, we describe the proposed framework workflow, detailing synchronization, UE identification, and measurement collection stages, and highlighting the correlation metric and antenna selection. Section V reports numerical results evaluating localization and identification of multiple UEs under different scenarios and parameters, while Sec. VI presents the testbed setup and experimental validation. Finally, Sec. VII concludes the paper and outlines future work.

Paper Notation Vectors and matrices are denoted by bold lowercase ($\mathbf{x}$) and uppercase ($\mathbf{X}$) letters, respectively, with transposition indicated by $(\cdot)^T$. The notation $\mathrm{diag}(\mathbf{x})$ represents a square matrix with the elements of $\mathbf{x}$ on its main diagonal. The absolute value of a scalar x is $|x|$, and the vector norm is $\|\mathbf{x}\|$. The operators $\lfloor x \rfloor$ and $\lfloor x \rceil$ denote the floor and rounding to the nearest integer. A continuous-time signal $x(t)$ sampled with period T_s is represented as $x[n] = x(nT_s)$. Its N-point Discrete Fourier Transform (DFT) is $\mathrm{DFT}_N\{x[n]\}$, and the complex conjugate of x is x^*. Sets are denoted by calligraphic uppercase letters (e.g., $\mathcal{A}$), with cardinality $|\mathcal{A}|$.

II. Related Works

This section reviews prior research on UAV-assisted localization, focusing on experimental studies that provide practical benchmarks for real-world implementations. Many works are simulation-based and rely on idealized assumptions, while experimental studies reveal the challenges and limitations of operating UAVs in realistic cellular networks. The following discussion highlights these contributions and key constraints.

Research on UAV-assisted localization can be broadly divided into two categories. The first includes serving-UAV approaches, where the aerial platform acts as a base station or relay to provide both communication and positioning services. The second focuses on non-serving UAVs, which operate as passive receivers to estimate UE positions without participating in the network.

In serving-UAV systems, several experimental works have investigated UAV-assisted localization using time-based measurements. Esrafilian *et al.* [7] presented one of the first real-world implementations of UAV-aided ToA localization in a 5G system using OpenAirInterface (OAI), combining UAV motion with UL SRS measurements to jointly estimate UE and UAV positions via least-squares optimization. Similarly, Albanese *et al.* [8] introduced SARDO, an autonomous UAV-based search-and-rescue system that localizes victims via their mobile phones without terrestrial infrastructure [6]. The UAV carries a lightweight LTE base station and estimates positions using Demodulation Reference Signal (DMRS)-based pseudo-trilateration and neural networks. Both studies demonstrate the feasibility of UAV-assisted localization with real cellular signals but share limitations of Time-of-Arrival (ToA) methods, including timing offsets from asynchronous UE clocks and the need for wide bandwidths, which may be impractical in emergencies. The work in [5] explores the potential of an Integrated Sensing and Communication (ISAC) system to improve search-and-rescue operations, detailing methods for processing echoed signals from downlink transmissions to enhance target detection and classify possible victims by fusing Synthetic Aperture Radar (SAR) imagery with triangulation results from UL signals [17], [18].

TABLE I
SUMMARY OF EXPERIMENTAL RELATED WORKS.

Ref.	Technique	Accuracy	HW Comp.	Scenario
[5]	RSSI & SAR	$\leq$10 m	Very High	180×180 m^2
[7]	TDoA	$\leq$15 m	Very High	70×70 m^2
[8]	Trilateration	$\leq$50 m	Very High	300×300 m^2
[11]	AoA Triang.	$\leq$50 m	High	100×100 m^2
[15]	W-RSS	$\leq$10 m	Low	25×25 m^2
[16]	Deep Learn.	20 m	High	170×220 m^2

Other experimental approaches have explored non-serving UAVs. Scazzoli *et al.* [11] demonstrated an AoA-based method where a UAV receives Physical Random Access Channel (PRACH) signals from a single UE, achieving about 25 m accuracy. However, the system requires multiple antennas, and its AoA estimates become unreliable during UAV rotations, as Doppler shifts distort the measured signal phase. PRACH transmissions are also unpredictable, limiting their use for continuous non-serving UAV localization. Petitjean *et al.* [15] proposed a Received Signal Strength Indicator (RSSI)-based method with a UAV equipped with a Universal Software Radio Peripheral (USRP) SDR to estimate the bearing of a ground transmitter. While effective in LoS conditions, it is sensitive to multipath and UAV altitude variations, and may converge to local RSSI maxima in complex environments. Masrur and Güvenç [16] introduced a deep-learning framework trained on simulated data for 3D source localization, achieving around 18 m average error in real-world trials, but requiring many samples. Table I summarizes the main experimental works in the literature. The reported complexity corresponds to the requirements of a hardware implementation, whereas accuracy values correspond to measurements obtained in rural scenarios.

Additional simulated studies highlight the benefit of trajectory optimization in UAV deployments. For example, [10] analyzes how different UAV mobility patterns affect data collection, evaluating a time–coverage efficiency metric. They show that circular and square trajectories offer the best trade-off between area coverage and efficiency. Similarly, [19] evaluates several UAV flight paths for device discovery in post-disaster scenarios, showing that the most effective trajectory depends on the damage distribution.

These studies highlight the potential of UAV-assisted localization but also show practical limits, often due to partial autonomy, strict synchronization assumptions, and the need for complex or costly hardware. In contrast, our work introduces a fully autonomous sensing UAV that performs synchronization, identification, localization, and mission planning entirely on-board using lightweight, low-cost hardware, enabling accurate multi-user localization with non-serving UAVs in real 5G environments.

III. SYSTEM MODEL

In this section, we introduce the system model. We describe the scenario, the UAV deployment, the communication between ground UEs and the network, and the properties of the UL signals.

Fig. 1. Operational scenario where a sensor UAV acts as non-serving receiver to localize UEs transmitting UL signals towards the network.

A. Scenario and operational workflow

We consider the scenario in Fig. 1, where a sensor UAV is deployed in an emergency environment to localize N_{UE} UEs. The position of each UE $u \in \mathcal{U}$, in the set of UEs $\mathcal{U} = \{I_j \mid j = 1, \ldots, N_{\mathrm{UE}}\}$, is represented by the vector $\mathbf{m}_u = [x_u, y_u, z_u]^T$, while the position of UAV evolves over time and is denoted, at the rth measurement time step, as $\mathbf{p}[r] = [x^{\mathrm{UAV}}[r], y^{\mathrm{UAV}}[r], z^{\mathrm{UAV}}[r]]^T$. The UAV therefore collects a set of R measurements during its trajectory, corresponding to successive time instants along the flight path. The UEs are connected to the network through a Next Generation Node B (gNB), exchanging radio frames that also include the transmission of SRS. Figure 2 illustrates the overall mission workflow, which will be detailed in the following sections.

A downlink telemetry connection is established with a ground control station (GCS) allowing the network to track estimated UE positions, identities, and UAV status parameters such as position and velocity. At the beginning of the mission, the network provides the UAV with the SRS configuration parameters (detailed in Sec. III-B) for all active UEs, enabling UEs localization and identification from the received signals. The UAV is also given an a priori area, defined by its center and extension, to guide its operation.

The UAV is equipped with N_{ant} antennas, collected in the set $\mathcal{A} = \{A_\ell \mid \ell = 1, \ldots, N_{\text{ant}}\}$, which are spaced to maximize spatial diversity. This latter was achieved, for the specific UAV platform used (detailed in Sec. VI-A), by mounting the antennas on the landing legs, spaced approximately 30 cm apart and oriented at a 90° angle with respect to each other.

After the mission initialization phase, the workflow consists of three main stages using the received SRSs: *(i)* coarse and fine time synchronization, *(ii)* UE identification and measurement extraction, and *(iii)* localization of the identified UEs. The synchronization step relies on a recursive update and is computationally negligible, with complexity $\mathcal{O}(1)$ per received sample, and no further re-synchronization is required thereafter. The UE identification, based on the Matching Pursuit algorithm [20] and DFT operations, dominates the overall cost with complexity $\mathcal{O}(LU_b \log L)$ per valid measurement, where L is half the symbol DFT length and U_b the number of UEs in the SRS band $\mathcal{K}_b$. Finally, the localization stage has complexity $\mathcal{O}(U_b\, i_{\max}^{(\mathrm{loc})})$ per valid measurement, with $i_{\max}^{(\mathrm{loc})}$ denoting the maximum number of iterations.

Fig. 2. Overview of the proposed framework building blocks.

B. Sounding Reference Signals (SRS)

SRSs are pilot signals used in 5G mainly to evaluate the UL channel quality and, starting from 3GPP Release 16, also for positioning purposes [12], [21]. They are therefore highly configurable in both time and frequency domains. Unlike other UL signals transmitted by different UEs, whose time–frequency resources are dynamically assigned by the gNB scheduler in an implementation-dependent manner, SRSs can be configured with fixed resources. For example, they may be set with a periodicity T_{per} as short as a single Orthogonal Frequency-Division Multiplexing (OFDM) slot, and with an arbitrary subcarrier starting position k_0, occupying a minimum bandwidth m_{SRS} of 4 Resource Blocks (RBs). A smaller bandwidth offers several advantages in rescue scenarios, such as enhancing the Signal-to-Noise Ratio (SNR) for the same transmitted power, reducing receiver complexity, improving the multiplexing capability of the network, and facilitating parallel search operations among multiple UAVs.

Another distinctive feature is that their modulation symbol sequence $Q[k]$ is allocated onto subcarriers in multiples of 2, up to 12, leaving the remaining subcarriers empty. In the remainder of this manuscript, we will consider an allocation step $K_{TC} = 2$, as further explained in Sec. IV-B3. The overall sequence length is then $M_{SRS} = 12m_{SRS}/K_{TC}$. An example of SRS allocated on the resource grid is depicted on the left of Figure 3. The sequence $Q[k]$ is defined as

$$Q_\alpha[k] = e^{j\alpha n}\bar{Q}_q[k], \qquad k = 0, 1, \ldots, M_{SRS} - 1, \tag{1}$$

where α is the applied cyclic shift and $\bar{Q}_q[k]$ is the base sequence generated starting from a cell identifier q. The base sequence can be expressed as:

$$\bar{Q}_q[k] = e^{-j\frac{\pi q k(k+1)}{N_{ZC}}}, \tag{2}$$

where N_{ZC} is the largest prime number such that $N_{ZC} < M_{SRS}$. In other words, $\bar{Q}_q[k]$ is the cyclic extension of the corresponding Zadoff-Chu (ZC) sequence of length N_{ZC}. ZC sequences of prime length are Constant Amplitude Zero Auto-Correlation (CAZAC), meaning they exhibit constant amplitude and zero autocorrelation, and their Inverse Discrete Fourier Transform (IDFT) results in another ZC sequence with the same properties. However, the cyclic extension partially degrades these ideal characteristics. Whenever $M_{SRS} < 36$, special flat-spectrum sequences with the same desirable properties as ZC sequences are used instead. Finally, a cyclic shift is applied to the base sequence[1], defined as $\alpha = (2\pi w)/8$, with $w = 0, 1, \ldots, 7$.

The subcarrier allocation of K_{TC} ensures that, after the IDFT, the time domain signal consists of K_{TC} identical halves, each lasting half a symbol. Therefore, the time domain waveform $x_\alpha[n]$, under the $K_{TC} = 2$ assumption including the Cyclic Prefix (CP), can be expressed as

$$x_\alpha[n] = \begin{cases} s_\alpha[n+L], & -L_{CP} \le n < 0, \\ s_\alpha[n], & 0 \le n < L, \\ s_\alpha[n-L], & L \le n < N, \\ 0, & \text{otherwise}, \end{cases} \tag{3}$$

where L_{CP} is the CP length, $L = N/K_{TC}$, N denotes the number of symbol IDFT points, and

$$s_\alpha[n] = \frac{1}{N}\sum_{k=0}^{M_{SRS}-1} Q_\alpha[k] e^{j\frac{2\pi}{N}(kK_{TC}+k_0)n}. \tag{4}$$

The visual representation of (3) is shown in the right portion of Fig. 3. It is important to note that, due to the CAZAC properties of ZC sequences, two signals with different cyclic shifts applied $s_{\alpha_i}[n]$ and $s_{\alpha_j}[n]$ satisfy

$$\sum_{n=0}^{L-1} s_{\alpha_i}[n]\, s^*_{\alpha_j}[n] \approx 0, \qquad \text{when } i \neq j. \tag{5}$$

Before the UAV mission starts, the network partitions with the C_{SRS} and B_{SRS} parameters the available SRS bandwidth into a set of mutually exclusive subbands with the minimum length of 4 RBs $\mathcal{B} = \{\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_B\}$, where each subband $\mathcal{K}_b = \{k_{0,b} + mK_{TC} \mid m = 0, 1, \ldots, M_{\text{SRS}} - 1\}$ corresponds to a contiguous set of frequency resources. These subbands are disjoint, i.e., $\mathcal{K}_b \cap \mathcal{K}_{b'} = \emptyset, \quad \forall b \neq b'$, and their union spans the allocated SRS bandwidth $\mathcal{B}$.

Fig. 3. Illustration of the SRS mapping: (left) placement on the OFDM grids with $K_{TC} = 2$, and (right) the corresponding repeated structure in the time domain.

[1]A single omnidirectional antenna port SRS transmission from each UE is considered.

Fig. 4. OFDM DFT window for (top) an unsynchronized receiver and (bottom) a synchronized receiver.

Each subband $\mathcal{K}_b$ can be assigned to one or more UEs, denoted by the set $\mathcal{U}_b \subseteq \mathcal{U}$, where every UE is characterized by a unique cyclic shift α_u within that subband. The SRS periodicity T_{per} is set to be the same for all UEs, and all SRS-related configuration parameters are communicated to the UAV at the beginning of the mission. We remark that each UE is associated with a unique pair $(\mathcal{K}_b, \alpha_u)$, which uniquely identifies the UE. In this work, we assume that a single UAV samples all SRS subbands and processes them in parallel.

C. Received Signal Model

The signal received at the UAV, located at position $\mathbf{p}[r]$ at the rth time step, on its antenna ℓ, from the set of UEs assigned to subband $\mathcal{K}_b$, $u \in \mathcal{U}_b = \{I_j \in \mathcal{U} : I_j \to \mathcal{K}_b\}$, each experiencing a set of propagation paths $\mathcal{P}_{u,\ell} = \{1, \ldots, P_{u,\ell}\}$ and transmitting over the same bandwidth $\mathcal{K}_b$, can be expressed in the time domain as

$$y_\ell(t) = \sum_{u=1}^{U_b} y_{u,\ell}(t) = \sum_{u=1}^{U_b} \sum_{p=1}^{P_{u,\ell}} a_{p,u,\ell} z_u(t - \tau_{p,u,\ell}) e^{j2\pi(\nu_{p,u,\ell} + \nu_{u,\text{cfo}})t} \tag{6}$$

where $U_b = |\mathcal{U}_b|$ denotes the number of transmitting UEs in subband $\mathcal{K}_b$, $P_{u,\ell}$ the number of propagation paths of UE u, and $z_u(t)$ its transmitted signal. The ToA of the pth path is $\tau_{p,u,\ell} = \tau'_{p,u,\ell} - TA_u$, where $\tau'_{p,u,\ell}$ is the physical propagation delay and TA_u is the applied timing advance. For each UE, the propagation delays are ordered in ascending order, i.e., $\tau_{1,u} < \tau_{2,u} < \cdots < \tau_{P_{u,\ell},u}$. The Doppler shift and Carrier Frequency Offset (CFO) are denoted by $\nu_{p,u,\ell}$ and $\nu_{u,\text{cfo}}$, respectively, while the complex path gain is $a_{p,u,\ell} = \tilde{a}_{p,u,\ell} e^{j2\pi f_c \tau'_{p,u,\ell}}$, with $\tilde{a}_{p,u,\ell}$ the complex baseband channel coefficient and f_c the carrier frequency. All these parameters, except CFO, implicitly depend on the UAV position $\mathbf{p}[r]$. For simplicity in the notation, we will, from now on, consider the received signal at a single antenna and a fixed SRS subband, omitting the subscripts ℓ and b.

IV. UEs Synchronization, Identification, and Localization

In this section, we describe the signal processing procedures that enable the sensing UAV to autonomously synchronize, identify, and localize multiple UEs from their UL SRSs. The synchronization and UE identification stages run independently on each antenna mounted on the UAV. Localization then exploits antenna diversity by merging the outputs of all antennas.

A. Coarse Time Synchronization via SRS Structure

In 5G systems, the transmitted data stream consists of consecutively concatenated OFDM symbols. To correctly recover the transmitted information, the gNB receiver must apply a DFT to each received symbol over a precisely aligned DFT window. Any misalignment causes interference between adjacent symbols, degrading signal quality, as illustrated in Fig. 4. Since multiple UEs are generally located at different positions, their signals arrive at the gNB with varying propagation delays depending on their relative distances. To ensure alignment, the network periodically instructs each UE to apply a Timing Advance (TA) correction so that all signals arrive within the same DFT window at the gNB.

If the sensing UAV is co-located at the gNB position, it will receive all UL transmissions in mutual synchrony. However, when the UAV operates away from the gNB location, the relative propagation delays differ, and the received streams become partially desynchronized. Additionally, the sensing UAV has no knowledge of the UEs' transmission start times.

To obtain a coarse time alignment, we exploit the inherent repetition structure of the SRS. We define the detection metric

$$M[n] = \frac{|P[n]|^2}{R[n]^2}, \tag{7}$$

where

$$\begin{aligned} P[n] &= \sum_{l=0}^{L-1} y[n+l]\, y^*[n+l+L], \\ R[n] &= \frac{1}{2}\left(\sum_{l=0}^{L-1} |y[n+l]|^2 + \sum_{l=0}^{L-1} |y[n+l+L]|^2 \right). \end{aligned} \tag{8}$$

The metric $M[n]$ is bounded between 0 and 1. This procedure provides coarse synchronization with any detected SRS, is robust against CFO (since the phase rotations cancel out in $P[n]$), and can be computed efficiently using recursive updates.

The ToAs of the SRSs received at the UAV from different UEs are generally mutually misaligned, since the UAV is not located at the gNB site where the TA aligns the streams. For convenience, we select one arrival time $\tau_{\bar{p},\bar{u}}$ as reference, with $\bar{u} \in \mathcal{U}$ and $\bar{p} \in \mathcal{P}_{\bar{u}}$, and express all others relative to it

$$\Delta\tau_{p,u} = \tau_{p,u} - \tau_{\bar{p},\bar{u}}. \tag{9}$$

To reliably detect the presence of SRSs, the metric should have values close to unity (i.e., $M[n] \approx 1$) at sufficiently high SNR. Evaluating $M[n]$ on the reference arrival index $n_0 = \lceil \tau_{\bar{p},\bar{u}}/T_s \rceil$ and neglecting noise, Appendix A proves that

$$M[n_0] \approx \frac{\sum_{p,u} |a_{p,u}|^2 f_0(\Delta\tau_{p,u})}{\sum_{p,u} |a_{p,u}|^2 g_0(\Delta\tau_{p,u})}, \tag{10}$$

where $f_0(\Delta\tau_{p,u}), g_0(\Delta\tau_{p,u}) \in [0,1]$ depend on the temporal misalignment $\Delta\tau_{p,u}$. Figure 5a shows that both satisfy $f_0 = g_0 = 1$ whenever

$$-T_{\text{CP}} \le \Delta\tau_{p,u} \le 0, \tag{11}$$

(a) Behavior of f_0 and g_0 as functions of the relative SRS arrival time $\Delta\tau_{p,u}$.

(b) Detection metric when the condition in (11) holds $\forall p, u$ (blue), and when it does not (black).

Fig. 5. Impact of delay spread on the synchronization detection metric.

with T_{CP} denoting the cyclic prefix duration. Outside this interval, it holds that $g_0(\Delta\tau_{p,u}) > f_0(\Delta\tau_{p,u})$, reducing $M[n_0]$. Figure 5b illustrates the two outcomes. If the relative arrival times $\Delta\tau_{p,u}$ $\forall\, p, u$ lie within the interval defined in (11), the metric attains its maximum value (blue curve) and remains 1 until at least one arrival time falls outside this interval. Conversely, if the condition is never satisfied, the metric remains strictly below 1 (black curve).

In practice, signals from distant UEs typically exhibit much lower power than those from nearby UEs, i.e., $|a_{p,u_{\text{near}}}|^2 \gg |a_{p,u_{\text{far}}}|^2$. Thus, their contribution to (10) is often negligible. As shown later through numerical results, the proposed metric remains robust even when multiple UEs are received with comparable power levels, up to practical distances on the order of kilometers.

Once the synchronization metric $M[n]$ is computed, the sensing UAV must identify the point at which it reaches its maximum value and exceeds a predefined threshold M_{th}. This point provides a coarse estimate of the ToA of the received SRSs and allows the receiver to achieve an initial time alignment. When received SRSs are synchronized, the timing metric exhibits a sharp rise followed by a plateau whose duration equals the CP length L_{CP}. This behavior arises because, when n is within the cyclic prefix interval, the two identical halves of the preamble align perfectly, contributing coherently to the correlation and yielding a constant maximum value of the metric rather than a single narrow peak. Therefore, to improve the detection, the metric $M[n]$ is convolved with a rectangular window of length equal to L_{CP}. This operation enhances the peak sharpness by transforming the plateau into a more localized maximum, facilitating reliable detection. Additionally, if the receiving signals are very unsynchronized, $M[n]$ may oscillate before the main peak. Filtering attenuates these side fluctuations and sharpens the main correlation peak. The resulting filtered metric is

$$M_F[n] = \frac{1}{L_{CP}} \sum_{n'=0}^{L_{CP}-1} M[n-n']. \tag{12}$$

The initial detection instant n_{sync} is determined by

$$M_F[n_{\text{sync}}] \geq M_{\text{th}} \wedge D[n_{\text{sync}}-1] > \delta \wedge D[n_{\text{sync}}] < -\delta, \tag{13}$$

where δ is a small positive quantity and $D[n] = M_F[n+1] - M_F[n]$ is the discrete derivative of the metric. For moderate SNRs (< 20 dB) within the investigated subband, and when (11) holds $\forall p, u$, the SNR can be estimated as

$$\hat{\text{SNR}} \approx \frac{\sqrt{M_F(n_{sync})}}{1-\sqrt{M_F(n_{sync})}}. \tag{14}$$

Consequently, the detection threshold M_{th} can be selected according to the minimum SNR level that must be reliably detected [22].

B. Fine Synchronization and UE Identification

Once this initial timing is established, the UAV must: (i) separate the incoming signals, (ii) identify the transmitting UEs, (iii) localize them, (iv) refine the synchronization error. Specifically, the residual synchronization error for UE u and path p is defined as

$$\varepsilon_{p,u} = n_{p,u} - n_{sync}, \tag{15}$$

where n_{sync} denotes the initial coarse synchronization instant and $n_{p,u} = \lceil \tau_{p,u}/T_s \rfloor$ the actual ToA.

1) Separation of UL Signals: This step consists of discriminating the individual UE signals within the received composite waveform, thereby enabling UE identification and subsequent processing.

The received SRS signal can be expressed as

$$y_{\text{SRS}}[n] = \sum_{u=1}^{U}\sum_{p=1}^{P_u} a_{p,u} x_u[n - L_{\text{CP}} - \varepsilon_{p,u}]\, e^{j2\pi(\nu_{p,u}+\nu_{u,\text{cfo}})\, nT_s}, \tag{16}$$

with $n = 0, \ldots, L-1$. Exploiting the repeated time structure, this can be rewritten (neglecting a common phase term) as

$$y_{\text{SRS}}[n] = \sum_{u=1}^{U}\sum_{p=1}^{P_u} a_{p,u}\, s_u[(n-\varepsilon_{p,u}) \bmod L] e^{j2\pi(\nu_{p,u}+\nu_{u,\text{cfo}})\, nT_s}. \tag{17}$$

Applying an L-point DFT, we obtain the received spectrum $Y_{\text{SRS}}[k']$ for $k' = 0, \ldots, L-1$. We index the M_{SRS} active subcarriers through the mapping $\kappa[m] \triangleq \frac{k_0 + mK_{\text{TC}}}{K_{\text{TC}}}$, $m = 0, \ldots, M_{\text{SRS}}-1$, such that the received spectrum on the mth active subcarrier is

$$Y_{\text{SRS}}[m] = Y_{\text{SRS}}(\kappa[m]) = \sum_{u=1}^{U}\sum_{p=1}^{P_u} a_{p,u} S(m) e^{j2\pi m\left(\frac{\alpha_u}{8} - \frac{\varepsilon_{p,u}}{L}\right)}. \tag{18}$$

We then multiply the received spectrum by the known reference sequence

$$c[m] = Y_{\text{SRS}}[m]\, S^*[m] = \sum_{u=1}^{U}\sum_{p=1}^{P_u} a_{p,u}\, e^{j2\pi m f_{p,u}}, \tag{19}$$

where $f_{p,u} = \alpha_u/8 - \varepsilon_{p,u}/L$. This compact expression shows that each user–path pair received by the UAV manifests as a complex exponential in the frequency domain, enabling the application of frequency estimation techniques for refined ToA and cyclic shift estimation. It is worth recalling that each pair $(\mathcal{K}_b, \alpha_u)$, representing the allocated SRS bandwidth and cyclic shift respectively, uniquely identifies a single UE. Therefore, by estimating the received cyclic shift and determining the corresponding occupied bandwidth, the UAV can discriminate

Fig. 6. Example of the computed $C[k]$ for three different UEs applying cyclic shifts 0, 4, and 2, respectively.

among the active UEs and associate each received signal with its transmitting UE.

Applying an L-point DFT on $c[m]$ yields

$$C[k]=\sum_{u,p} a_{p,u} e^{-(j\,2\pi\,\eta_{p,u}[k]\,(M_{\mathrm{SRS}}-1))} \frac{\sin(\pi M_{\mathrm{SRS}} \eta_{p,u}[k])}{\sin(\pi \eta_{p,u}[k])}, \tag{20}$$

where $\eta_{p,u}[k] = f_{p,u} - k/L$. Figure 6 illustrates an example of the computed $C[k]$ obtained from SRS signals received by three different UEs applying different cyclic shifts.

2) Peaks Extraction: The dominant frequency components of the received signal are iteratively extracted using a Matching Pursuit algorithm [20] over a sinusoidal dictionary, i.e. the DFT bins. If the dictionary resolution is coarse, oversampling and parabolic interpolation refine the estimates. At each iteration, the strongest component is removed from the residual, and its amplitude and equivalent power are estimated. The process continues until the Bayesian Information Criterion (BIC) [23] indicates that adding further components no longer improves the model. The BIC is computed as

$$\mathrm{BIC}^{(i)} = i \log M_{\mathrm{SRS}} + M_{\mathrm{SRS}} \log \left(\frac{1}{M_{\mathrm{SRS}}} \sum_{m=0}^{M_{\mathrm{SRS}}-1} |c^{(i)}[m]|^2 \right). \tag{21}$$

The procedure is summarized in Algorithm 1.

3) User Identification via Cyclic Shift: Note that, due to the different cyclic shifts applied, the frequencies associated with different UEs are typically well separated, which facilitates their extraction. We can convert the estimated frequencies $\hat{f}^{(i)}$ into the corresponding estimated cyclic shifts as

$$\hat{\alpha}^{(i)} = \left\lceil U(\hat{f}^{(i)} + 1/(2U)) \right\rceil \bmod L, \tag{22}$$

where U denotes the total number of cyclic shifts assigned to the UEs (i.e. the number of users) within the considered band.

Algorithm 1 Matching Pursuit over DFT dictionary

Require: SRS spectrum $c[m]$, half symbol length L, M_{SRS}
Initialize $c^{(0)}[m] = c[m]$, $i = 1$
while BIC decreases **do**
  Compute DFT: $C^{(i)}[k] = FFT_L(c^{(i-1)}[m])$
  Find peak: $\hat{k}^{(i)} = \arg\max_k |C^{(i)}[k]|$
  Estimate: $\hat{f}^{(i)} = \hat{k}^{(i)}/L$
  Estimate: $\hat{a}^{(i)} = \frac{1}{M_{\mathrm{SRS}}} \sum_{m=0}^{M_{\mathrm{SRS}}-1} c^{(i-1)}[m] e^{-j2\pi m \hat{f}^{(i)}}$
  Estimate : $\hat{\gamma}^{(i)} = 20 \log_{10} |\hat{a}^{(i)}|$
  Update: $c^{(i)}[m] = c^{(i-1)}[m] - \hat{a}^{(i)} e^{j2\pi m \hat{f}^{(i)}}$
  Compute $\mathrm{BIC}^{(i)}$ as in (21)
  $i \leftarrow i + 1$

If the cyclic shifts are uniformly (i.e., maximally) spaced and U is even, to avoid a misdetection of the user–path tuple (u, p), the synchronization error $\varepsilon_{p,u}$, must satisfy

$$\varepsilon_{p,u} \in (-L/(2U),\, L/(2U)]\,. \tag{23}$$

Consequently, our goal is to maximize $L = N/K_{\mathrm{TC}}$ by setting $K_{\mathrm{TC}} = 2$, which corresponds to its minimum allowed value. Furthermore, depending on the size of the considered scenario, fewer cyclic shifts than the maximum per band (i.e., $U_{\max} = 8$) can be allocated, reducing UE identification errors. Once the cyclic shift $\hat{\alpha}^{(i)}$ is estimated, the identity of the transmitting UE can be unambiguously determined by pairing it with the corresponding frequency band $\mathcal{K}_b$.

4) Fine synchronization: The timing synchronization described in Sec. IV-A allows us to coarsely synchronize to the received SRS signals. This operation is required only once, since the SRS periodicity of the transmitting UEs is known by the UAV. This allows us to maintain synchronization at subsequent receptions without recomputing the timing metric from scratch every SRS reception.

Starting from the initial synchronization instant n_{sync}, the synchronization points are updated at each SRS reception as

$$n_q = n_{sync} + qT_{\mathrm{per}} - \tilde{\varepsilon}(q), \tag{24}$$

where $q \in \mathbb{N}$, T_{per} is the SRS periodicity and $\tilde{\varepsilon}(q)$ is the filtered residual synchronization error, updated through

$$\tilde{\varepsilon}(q) = \beta\, \hat{\varepsilon}_{\bar{p},\bar{u}}(q) + (1-\beta)\, \tilde{\varepsilon}(q-1), \tag{25}$$

with $\beta \in [0, 1]$ controlling the weighting between the latest and previous estimates.

The residual synchronization error arises for two main reasons. First, due to noise and mutual SRS interference, the initial n_{sync} is typically not perfectly aligned with the SRS peak but falls somewhere between replicas. Second, even after fine synchronization, the UAV's and UEs' clock drift over time, introducing an increasing offset.

Since fine synchronization can only be accurately aligned with respect to a single received SRS (the others will remain slightly misaligned, except when the UAV is located in correspondence with the gNB site), the most powerful received SRS is typically chosen as the reference. The residual error associated with this reference path at the qth SRS reception is then estimated as

$$\hat{\varepsilon}_{\bar{p},\bar{u}}(q) = \left\lfloor L(\hat{f}^{(1)}(q) - \hat{\alpha}^{(1)}(q)/8) \right\rfloor, \tag{26}$$

where the first iteration of the frequency estimation process always corresponds to the strongest received replica.

C. Measurement Cleaning and False Positive Handling

After frequency estimation, the UAV obtains vectors of estimated metrics $\tilde{\boldsymbol{\gamma}} = [\hat{\gamma}^{(1)}, \ldots, \hat{\gamma}^{(i_{end})}]^T$ and corresponding cyclic shifts $\tilde{\boldsymbol{\alpha}} = [\hat{\alpha}^{(1)}, \ldots, \hat{\alpha}^{(i_{end})}]^T$, where each element represents a detected user-path pair (u, p) and i_{end} is the maximum number of iterations. To prepare for localization, only the strongest component per UE is retained. For each UE $u \in \mathcal{U}_b$, define $\mathcal{J}_u = \{j \mid \hat{\alpha}^{(j)} = \alpha_u\}$, $\hat{\gamma}_u^{\max} = \max_{j \in \mathcal{J}_u} \hat{\gamma}^{(j)}$, and discard all other replicas. The cleaned measurement vector

Algorithm 2 Weighted Mean-Shift Localization for UE u

Require: Per-antenna sequences $\{\gamma_{\mathrm{SRS,ant}}^{(u,\ell)}[r]\}$, UAV positions $\{\mathbf{p}_{u,r}\}$, threshold γ_{th}, kernel bandwidths h_x, h_y, tolerance δ, max iterations $i_{\max}$
Ensure: Estimated position $\hat{\mathbf{m}}_u$

Select best antenna: $\ell_u^\star[r] \leftarrow \arg\max_{\ell\in\mathcal{A}} \gamma_{\mathrm{SRS}}^{(u,\ell)}[r]$
Form measurement sequence: $\gamma_{\mathrm{SRS}}^{(u)}[r] \leftarrow \gamma_{\mathrm{SRS}}^{(u,\ell_u^\star[r])}[r]$
Normalize: $\tilde{\gamma}^{(u)}[r] \leftarrow \gamma_{\mathrm{SRS}}^{(u)}[r]/\max_r \gamma_{\mathrm{SRS}}^{(u)}[r]$
Select valid samples: $\mathcal{R}_u \leftarrow \{r \mid \tilde{\gamma}^{(u)}[r] \geq \gamma_{\mathrm{th}}\}$
Set weights: $w_{u,r} \leftarrow \tilde{\gamma}^{(u)}[r]$
Initialize $\hat{\mathbf{m}}_u^{(0)}$ (e.g., weighted centroid), $i \leftarrow 0$
repeat
    Update position $\hat{\mathbf{m}}_u^{(i+1)}$ as in (29)
    $i \leftarrow i+1$
until $\|\hat{\mathbf{m}}_u^{(i+1)} - \hat{\mathbf{m}}_u^{(i)}\| < \delta$ **or** $i \geq i_{\max}$
return $\hat{\mathbf{m}}_u \leftarrow \hat{\mathbf{m}}_u^{(i+1)}$

is $\boldsymbol{\gamma}_{\mathrm{SRS}} = [\,\hat{\gamma}_1^{\max}, \ldots, \hat{\gamma}_U^{\max}\,]$. Missed detections due to low received power can be flagged for subsequent processing.

False positives may occur when other received signals exceed the detection threshold M_{th}, typically from narrowband signals with slow variations. In this case, the algorithm applies the nominal SRS preprocessing and a second DFT. Under the SRS hypothesis, structured peaks appear at specific frequency bins; under random data, the spectrum is noise-like. The BIC-based model selection automatically distinguishes these cases: if the first iteration is selected, a false positive is declared, and the synchronization metric $M[n]$ is recomputed on newly acquired samples.

D. UEs Localization

The accurate estimation of UE positions relies fundamentally on the correlation metric $M[n]$ introduced in Section IV-A, which captures the temporal and frequency structure of the SRS signals. Since each UAV is equipped with N_{ant} antennas, at every measurement instant r the receiver obtains, for each UE u, the antenna-specific metric vector $\boldsymbol{\gamma}_{\mathrm{SRS},ant}^{(u)}[r] = [\,\gamma_{\mathrm{SRS}}^{(u,1)}[r],\ \ldots,\ \gamma_{\mathrm{SRS}}^{(u,N_{ant})}[r]\,]^T$.

To reduce antenna-dependent small-scale fading fluctuations while retaining the large-scale spatial trend relevant for localization, we select, for each UE u, at the measurement instant r, the antenna whose SRS-derived metric exhibits the highest metric

$$\ell_u^\star[r] = \arg\max_{\ell\in\mathcal{A}}\ \gamma_{\mathrm{SRS},ant}^{(u,\ell)}[r], \tag{27}$$

and define the selected measurement sequence as $\gamma_{\mathrm{SRS}}^{(u)}[r] \triangleq \gamma_{\mathrm{SRS},ant}^{(u,\ell_u^\star[r])}[r]$. By stacking the selected sequences over time, we obtain $\boldsymbol{\Gamma}_{\mathrm{SRS}} = \begin{bmatrix}\boldsymbol{\gamma}_{\mathrm{SRS}}^{(1)} & \cdots & \boldsymbol{\gamma}_{\mathrm{SRS}}^{(U)}\end{bmatrix}^T \in \mathbb{R}^{U\times R}$.

Each row is normalized independently:

$$\tilde{\boldsymbol{\gamma}}_{\mathrm{SRS}}^{(u)} = \frac{\boldsymbol{\gamma}_{\mathrm{SRS}}^{(u)}}{\max_r \gamma_{\mathrm{SRS}}^{(u)}[r]}, \qquad u = 1, \ldots, U, \tag{28}$$

and unreliable measurements are discarded using $\mathcal{R}_u = \{\, r \mid \tilde{\gamma}_{\mathrm{SRS}}^{(u)}[r] \geq \gamma_{\mathrm{th}}\,\}$. Each valid sample is then mapped to the UAV position and localization weight: $\mathbf{p}_{u,r} = \mathbf{p}^{(2D)}[r], \quad w_{u,r} = \tilde{\gamma}_{\mathrm{SRS}}^{(u)}[r]$.

Fig. 7. 3D view of the urban scenario. In green, the mission area.

The localization is performed using a weighted mean-shift procedure

$$\hat{\mathbf{m}}_u^{(i+1)} = \frac{\sum\limits_{r\in\mathcal{R}_u} \mathbf{p}_{u,r} w_{u,r} e^{-\frac{1}{2}(\mathbf{p}_{u,r}-\hat{\mathbf{m}}_u^{(i)})^\top \mathbf{H}^{-1}(\mathbf{p}_{u,r}-\hat{\mathbf{m}}_u^{(i)})}}{\sum\limits_{r\in\mathcal{R}_u} w_{u,r} e^{-\frac{1}{2}(\mathbf{p}_{u,r}-\hat{\mathbf{m}}_u^{(i)})^\top \mathbf{H}^{-1}(\mathbf{p}_{u,r}-\hat{\mathbf{m}}_u^{(i)})}} \tag{29}$$

where $\mathbf{H} = \mathrm{diag}\big([h_x^2,\ h_y^2]^T\big)$, and h_x and h_y denote the kernel bandwidths along the East–West and North–South directions, respectively. The algorithm iterates until $\|\hat{\mathbf{m}}_u^{(i+1)} - \hat{\mathbf{m}}_u^{(i)}\| < \delta$, and the final estimate is denoted by $\hat{\mathbf{m}}_u$. The proposed localization workflow is summarized in Alg. 2.

During the localization process, the UAV first moves to the center of the designated polygonal search area and then executes an initial perimeter scan. The collected measurements are immediately processed onboard to provide a first coarse estimate for each UE position. Then, the UAV autonomously transitions to a refinement phase, during which it performs a small-scale, high-resolution search (typically hexagonal) centered around each preliminary estimate. This phase yields dense measurement clusters that further improve the localization accuracy in real time. Once all UEs are refined, the UAV transmits the final estimates to the GCS and executes an autonomous return-to-home sequence, completing the multi-UE localization mission.

V. Simulation Results

We report here the numerical results obtained from the evaluation of the proposed framework for both UEs identification and localization. A realistic environment was reconstructed from *OpenStreetMap*. We considered two environments: (i) an urban scenario near the Giuriati Sports Campus in Milan and (ii) a rural scenario, both covering an area of approximately 120 m × 80 m. The urban scenario is shown in Fig. 7. The geographic information was imported into MATLAB, from which the scenario boundaries and building heights were extracted.

The MATLAB ray-tracing toolbox is used to determine the LoS and Non-LoS (NLoS) visibility conditions between the UAV position and all UEs. Channel realizations are generated using the QUADRIGA MATLAB library, leveraging the pre-computed LoS/NLoS conditions to ensure geometrical consistency. Two propagation models were adopted, namely the `3GPP_38.901_RMa_LOS/NLOS` and `3GPP_38.901_UMa_LOS/NLOS` scenarios, depending on

(a) $U_b = 2$. (b) $U_b = 4$.

Fig. 8. Misidentification probability for $U_b = 2$ and $U_b = 4$ as a function of the inter-UE received power disparity and delay spreads, for an SRS bandwidth of 1.4 MHz.

(a) 1.4 MHz bandwidth. (b) 13 MHz bandwidth.

Fig. 9. Misidentification probability as a function of delay spread for two SRS bandwidth configurations. Increasing the bandwidth sharpens the correlation peaks and reduces the likelihood of adjacent peaks leakage, improving identification robustness.

the considered environment. Although these models were originally designed for terrestrial gNB deployments, they provide a reasonable approximation for UAV communications up to 25 m. A Doppler component was then applied for all the UEs according to the UAV velocity and orientation. The maximum UAV speed is fixed at 6 m/s. The UEs antennas were modeled as half-wave dipoles with random orientation to emulate realistic UE conditions.

As discussed in the next section, the UAV was equipped with two ADALM-PLUTO SDR stock antennas, mounted approximately 30 cm apart and oriented 90° angle with respect to each other. To approximate their 3D radiation pattern, the measured horizontal and vertical cuts were imported and combined through 2D interpolation. We set the receiver noise figure to 7 dB and the carrier frequency to $f_c = 2.4$ GHz. The power threshold and synchronization metric threshold were both set to $\gamma_{th} = M_{th} = 0.6$. The kernel bandwidths h_x and h_y were set to half of the width and height of the spatial region of interest, respectively. The convergence tolerance was fixed at 10^{-6}, with a maximum of 50 iterations. As exchanged traffic between UEs and the UAV, we loaded dummy traffic SDR recordings with active SRS using the OAI platform, as better described in the next section. Finally, to emulate realistic hardware conditions, a timing clock drift of 0.05 ppm was introduced at the UEs, with drift compensation applied every second, in accordance with the maximum synchronization tolerance defined by the 3GPP specifications [12]. The UAV

(a) (b)

Fig. 10. CDFs of the localization error obtained with the proposed method and benchmark approaches in (a) rural and (b) urban environments.

flight grid was discretized with a spatial resolution of 1 m and UAV height $z_b^{UAV} = 25$ m. The UAV flight path was described in Section IV-D. Throughout the entire UAV trajectory, measurements were assumed to be collected as soon as they became available, with a minimum spacing of 1 m between consecutive samples, determined by the scenario resolution.

We first evaluate the reliability of UE identification, quantifying the probability of misidentification under varying conditions. Figure 8a and Figure 8b report the misidentification probability as a function of the power spread and delay spread between UEs in the same band b when $U_b = 2$ and $U_b = 4$, respectively. As predicted by (10), UEs whose signals are received with similar power levels at the UAV degrade the synchronization metric, which in turn reduces the identification reliability, consistently with the bounds in (23). Figures 9a and 9b show the misidentification probability versus the delay spread, averaged over the power spread, for SRS bandwidths of 1.4 MHz and 13 MHz, respectively. As expected, the error probability decreases when fewer UEs are present. Moreover, increasing the bandwidth sharpens the sinc-shaped correlation peaks in Figure 6, improving time resolution and reducing inter-bin leakage, which lowers the chance of misidentification. For $U_b = 8$, a local peak appears around 1 μs. This behavior is due to the fact that, in such delay spreads, the orthogonality condition in (37) is violated more frequently when many UEs are multiplexed within a single band. The resulting correlation metric $M[n]$ degrades, and the likelihood of identification errors increases.

We compared the proposed localization approach against two state-of-the-art methods [7], [11], while keeping the synchronization and identification procedures identical across all implementations. The first benchmark follows a TDoA-based scheme that exploits the UAV motion around the targets, whereas the second performs triangulation using the AoA estimated at each UAV position. For a fair comparison, the benchmark methods operated over a 40 MHz bandwidth to mitigate multipath effects and noise on the LoS component, whereas our approach used narrower 1.4 MHz individual SRS bands (4 RBs, 5G numerology 1) up to a total of 20 MHz overall. Each simulation deployed four UEs across each single band, while the benchmark methods localized a single UE at a time using a 40 MHz transmission, due to their higher

(a) Hardware components and connections. (b) Assembled F450 UAV. (c) On field deployment.

Fig. 11. Hardware, physical realization, and experimental testbed of the autonomous UAV-based localization platform. (a) Onboard hardware architecture. (b) Assembled F450 drone with payload and antenna system. (c) On field deployment and experimental setup for UAV-based localization of six UEs.

sensitivity to interference. Although this configuration was necessary for the benchmark schemes, it required higher UE transmit power to maintain the same SNR, leading to faster UE battery depletion in real deployments. It also increased receiver complexity and reduced the number of measurements per UE, thereby suppressing the system's multiplexing capability.

To evaluate the positioning method, we use the Localization Error (LE) metric, defined for the uth UE as $\mathrm{LE}_u = \sqrt{\|\hat{\mathbf{m}}_u - \mathbf{m}_u\|^2}$. The Average Localization Error (ALE) is then computed as the mean LE over all localized UEs. The cumulative distribution functions (CDFs) of the positioning error obtained with the three methods, along with their corresponding ALE values, are shown in Figs. 10a and 10b. As depicted, our approach achieves a localization improvement of about 5–6 m compared to [7], [11].

At low flight altitudes AoA tends to approach the horizon, which limits its localization capability. In real deployments, it also requires multi-element arrays with precise geometry and calibration, which is difficult to support on small aerial platforms. TDoA based localization faces similar constraints: accurate multilateration needs sub-nanosecond synchronization, and even small timing offsets create meter-level errors. Resolving closely spaced multipath components also demands large bandwidth, yet reflected paths often arrive only a few nanoseconds apart, which complicates delay estimation at moderate bandwidths. Increasing bandwidth alleviates this but requires high sampling rates and significant onboard processing, incompatible with lightweight UAV platforms. To address the limitations of angle and delay based estimators, both sensitive to multipath and platform constraints, our method employs a signal consistency metric combined with kernel based weighting of the measurements collected along the UAV trajectory.

VI. Experimental Campaign

The UAV control and telemetry architecture used for experimental validation of the proposed UEs identification and localization system is described below.

A. Testbed Configuration

The hardware configuration and software integration required to enable an autonomous drone capable of detecting one or more UEs within a defined area of interest are detailed here. The system consists of an autonomous UAV unit and a GCS. The setup of the former is shown in Fig. 11a, while the latter consists of a laptop connected via Wi-Fi. During flight, the UAV transmits telemetry data to the GCS using a User Datagram Protocol (UDP) channel, with data serialization handled through Google Protocol Buffers (Protobuf). The telemetry stream includes position, altitude, velocity, and system status parameters. The received data are processed and visualized in real-time using a custom Live Drone GPS Tracker application developed with Dash and Leaflet. The autonomous UAV unit is further subdivided into a SDR, a Central Processing Unit (CPU), and the UAV's flight controller. Each component is described in detail below.

The core Radio Frequency (RF) component of the autonomous multi-UE localization system is based on the Analog Devices ADALM-Pluto Revision C, a low-cost Commercial off-the-Shelf (CotS) SDR. To enhance spatial measurement diversity and fully exploit the capabilities of its integrated RFIC, the Pluto SDR was modified to expose its dual-channel receive functionality, as shown in Fig. 11a. This customized SDR module, equipped with the stock dipole antennas, was mounted on the UAV platform, as shown in Fig. 11b. The use of stock antennas provides a wide field of view, which is particularly advantageous during the initial broad-area search phase. To account for differences in channel gain, measurements were performed using a known source generated by an R&S SMA100B CW generator. The channels were found to be within 0.3 dB of each other when measuring the same source, so calibration was not required. The SDR exposes its functionalities and captures I/Q samples via USB using the Industrial Input/Output (IIO) protocol.

At the other end of the hardware stack is the flight controller, responsible for navigating the UAV to a specified location. The flight controller comprises a PixHawk 2.4.8 mounted on a standard F450 UAV frame. This choice ensures compatibility with ArduPilot, enabling autonomous waypoint-guided control and real-time flight data access. Additionally, testing of flight control scripts can be performed through software-in-the-loop simulations. The PixHawk controller exposes these functionalities via its telemetry ports, using the Micro Air Vehicle Link (MAVLink) protocol over a UART physical link.

We selected the Nvidia Jetson Orin NX development board as the CPU that controls the drone. Located between the SDR and the flight controller, the CPU is tasked with collecting

data from both ends, processing it, and then providing new waypoints to the flight controller. The choice of a Jetson enables offloading to the GPU for processing-intensive tasks. The SDR is controlled via Python using Analog Devices' provided packages, and, in the same framework, the flight controller is handled by the python DroneKit library.

The GCS consists of a computer that establishes a wireless link with the onboard processing unit of the UAV, implemented using an NVIDIA Jetson module. Communication between the GCS and the UAV is achieved via Wi-Fi, with mission parameters configured remotely through a secure Secure Shell (SSH) connection.

B. Experimental Results

The experimental campaign was conducted at the HeliFly field near Milan, as shown in Fig. 11c, which offered a wide open area suitable for UAV-based data collection. A GCS was deployed to manage flight operations and telemetry. Six UEs were used in total: three placed inside a car on the back seats and three positioned outdoors on a chair at approximately 1 m height. Initially, a complete 5G network built on the OAI platform was planned. However, the parameter $B_{\mathrm{SRS}} > 0$, required to frequency-multiplex multiple UEs SRS, was not yet supported. As a workaround, a realistic dataset was generated offline.

Dummy UL traffic was recorded in the laboratory using an OAI-based 5G setup with SRS enabled and configured as follows [12], [21]: $T_{\mathrm{per}} = 40$ ms, numerology $= 1$, $C_{\mathrm{SRS}} = 0$, $B_{\mathrm{SRS}} = 0$, and $K_{\mathrm{TC}} = 2$. The recording was repeated three times with different cyclic shifts $\alpha = \{0, 4, 2\}$, corresponding to three distinct UEs. The three signals were then superimposed, each affected by a random CFO drawn according to the 3GPP tolerance of 0.5 ppm, and additional timing offsets of $1.5T_s$ ns and $-1.5T_s$ ns were applied to the second and third signals, respectively. These offsets emulate realistic timing misalignments and partial interference due to propagation differences and timing-advance adjustments, while maintaining proper UE synchronization.

To emulate frequency-domain multiplexing, a duplicate of the composite signal was placed on adjacent subcarriers to form a second SRS sub-band. Transmit powers were set to -5 dBfs, -8 dBfs, and -15 dBfs for the three UEs in each sub-band. This procedure generated two equivalent groups of three users each, assigned to two distinct SRS sub-bands. Although the UEs were co-located in this setup, the UAV treated them as six independent targets, each with its own measurement set. UEs 1, 2, and 3 transmit on subband $\mathcal{K}_0$, while UEs 4, 5, and 6 transmit on subband $\mathcal{K}_1$.

The UAV flew at an altitude of 25 m to ensure safe clearance and independently monitored both sub-bands, each associated with three UEs. Throughout the mission, the UAV maintained a continuous telemetry link with the GCS, transmitting flight data and UE position updates as soon as they became available. During the flight, the system continuously logged UE identifiers, received power levels, and precise UAV coordinates (latitude, longitude, altitude), along with auxiliary data such as battery level and attitude parameters. A schematic overview of the experimental setup is shown in Fig. 11c.

Fig. 12. Measured $\tilde{\gamma}^{(u)}_{SRS}[r]$ and localization results obtained during the experimental campaign: (a) UE 1 with clear LoS condition, (b) UE 4 placed inside a vehicle under NLoS conditions. Markers indicate the estimated (purple) and true (light blue) UE positions.

TABLE II
AVERAGE EXPERIMENTAL RESULTS.

UE ID	Flight Duration [min]		Localization Accuracy [m]	
	Initial	Refined	Initial	Refined
UE1	1.55	3.35	2.8 ± 1.6	1.8 ± 0.6
UE2	1.55	4.55	2.8 ± 1.6	1.8 ± 0.6
UE3	1.55	6.15	2.8 ± 1.6	1.7 ± 0.5
UE4	1.55	8.05	10.1 ± 1.5	8.2 ± 1.2
UE5	1.55	8.35	10.3 ± 1.3	8.3 ± 1.1
UE6	1.55	10.05	9.9 ± 1.4	8 ± 1.0

The carrier frequency was set within the Short Range Device (SRD) Industrial, Scientific, and Medical (ISM) band at $f_c = 866$ MHz. The maximum transmitted power of the ADALM-Pluto SDR was below 10 dBm, in compliance with SRD regulations, while the receiver gain onboard the UAV was set to 50 dB.

The mission began with the simultaneous powering of the autonomous UAV platform and the GCS. Once the onboard computer initialized, an SSH connection was established from the GCS to configure mission parameters. The target localization area was defined by a central geographic coordinate pair and a configurable polygon of $120\,\mathrm{m} \times 95\,\mathrm{m}$, with a flight altitude of 25 m. Upon command initiation, the UAV autonomously took off, navigated to the center of the area, and executed a perimeter sweep maintaining a 20 m margin from each side. After completing this trajectory, the system performed an initial localization estimate for each detected UE, followed by a refinement phase based on a hexagonal search pattern centered around the preliminary estimate, as described in Section IV. Finally, the UAV executed an autonomous

return-to-home sequence.

Measured $\tilde{\gamma}_{SRS}[r]$ and UE position estimates obtained for UE 1 (outdoor) and UE 4 (in-car) are shown in Figs. 12a and 12b, respectively. For the outdoor UE, benefiting from a clear LoS channel, measurements were consistently collected across the entire scenario, yielding localization accuracies below 2.5 m after refinement and below 4.5 m after the initial estimate, across multiple flight repetitions.

Conversely, the UEs inside the vehicle experienced harsh propagation conditions. The metallic roof blocked the direct path, forcing the received signal to propagate mainly through windows and ground reflections, leading to lower power levels compared to the LoS case. Nevertheless, localization accuracy remained below 10 m after both estimation phases. This outcome is consistent with expectations, as received power maxima do not align directly above the vehicle but rather originate from reflection points near it.

As visible in Fig. 12a, measurement density is non-uniform. This effect stems from a hardware limitation of the ADALM-Pluto SDR, which, using a USB 2.0 interface, cannot continuously stream all received samples to the host computer. As a result, the effective measurement sampling interval was approximately 1 s. This limitation could be largely mitigated using a higher-performance SDR capable of continuous data streaming, enabling denser and more uniform spatial sampling.

The overall mission time averaged 2.15 min to complete the first localization estimates for all UEs within both SRS subbands. Each hexagonal refinement pattern required an additional 1.30 min per UE. A summary of the main experimental results is reported in Table II. Although not implemented in this work, future extensions may allow cooperative missions involving multiple UAVs, where one platform could transmit an initial estimate to another for refinement, further reducing total mission time.

VII. Conclusion

This work demonstrated that a single sensing UAV, operating independently from the communication infrastructure, can reliably capture and interpret real 5G UL SRS transmissions to identify and localize multiple UEs in a fully autonomous manner. Beyond validating the proposed synchronization and identification pipeline, the results highlight an important insight: non-serving UAVs are sufficient to achieve meter-level localization without requiring large arrays, wide bandwidths, or strict time synchronization. The experimental campaign confirmed the method's robustness under hardware limitations typical of lightweight UAV platforms.

Our findings indicate that sensing-oriented UAVs can complement existing terrestrial networks by providing on-demand situational awareness in areas that lack coverage or where infrastructure is temporarily unavailable. Future work will focus on extending the proposed framework to cooperative multi-UAV operation, enabling distributed sensing, parallel UE tracking, and reduced mission times. Coordinated trajectories and shared measurements among multiple platforms have the potential to further enhance scalability, robustness, and situational awareness in complex or large-scale environments.

Fig. 13. Comparison between the exact correlation metric $M[n]$ (red), and its approximation (blue).

Appendix A
Expansion of the Detection Metric

Using the notation introduced in Section IV-A, the detection metric evaluated at the reference arrival point n_0 can be expressed as

$$M[n_0] = \frac{|P[n_0]|^2}{|R[n_0]|^2}. \tag{30}$$

The correlation term $P[n_0]$ expands as

$$P[n_0] = \sum_{l=0}^{L-1} \left(\sum_{u=1}^{U} \sum_{p=1}^{P_u} a_{p,u}\, y(lT_s - \Delta\tau_{p,u}) e^{j2\pi\nu_{u,\text{cfo}} lT_s} \right) \times \left(\sum_{v=1}^{U} \sum_{q=1}^{P_v} a_{q,v}^*\, y^*((l+L)T_s - \Delta\tau_{q,v}) e^{-j2\pi\nu_{v,\text{cfo}}(l+L)T_s} \right), \tag{31}$$

where we assumed $\nu_{u,\text{cfo}} \gg \nu_{p,u}$.

1) Single-Path, Single-User Contribution: For the case $u = v$ and $p = q$, and considering $y(t) = x(t)$ (i.e., reception of an SRS), the expression simplifies, up to a common phase rotation, to

$$P[n_0] = \sum_{u=1}^{U} \sum_{p=1}^{P_u} |a_{p,u}|^2 \sum_{l=0}^{L-1} x_u(lT_s - \Delta\tau_{p,u})\, x_u^*((l+L)T_s - \Delta\tau_{p,u}). \tag{32}$$

Assuming normalized symbol energy $E_s = 1$, and no additional OFDM symbols before or after the targeted SRS, this can be rewritten as

$$P[n_0] = \sum_{u=1}^{U} \sum_{p=1}^{P_u} |a_{p,u}|^2 f_0(\Delta\tau_{p,u}), \tag{33}$$

where

$$f_0(\Delta\tau_{p,u}) \approx \begin{cases} \dfrac{\Delta\tau_{p,u} + T_{CP} + T_L}{T_L}, & -T_L - T_{CP} \le \Delta\tau_{p,u} < -T_{CP}, \\ 1, & -T_{CP} \le \Delta\tau_{p,u} < 0, \\ 1 - \dfrac{\Delta\tau_{p,u}}{T_L}, & 0 \le \Delta\tau_{p,u} < T_L, \\ 0, & \text{otherwise}, \end{cases} \tag{34}$$

and $T_L = LT_s$. Notably, no residual CFO appears in (34), since the contributions of the two symbol halves cancel out.

2) Multi-Path Contribution: When $u = v$ and $p \neq q$, the expression (up to a common phase term) becomes

$$P[n_0] = \sum_{u=1}^{U} \sum_{\substack{p,q \\ p\neq q}} a_{p,u} a_{q,u}^* \sum_{l=0}^{L-1} x_u(lT_s - \Delta\tau_{p,u}) \times x_u^*((l+L)T_s - \Delta\tau_{q,u})\, e^{j2\pi(\nu_{p,u}-\nu_{q,u})lT_s}. \quad (35)$$

Due to the impulsive autocorrelation property of SRSs, delayed replicas of the same signal are nearly orthogonal, and the above term vanishes, i.e., $P[n_0] \approx 0$.

3) Multi-User Contribution: For the case $u \neq v$, we obtain

$$P[n_0] = \sum_{\substack{u,v \\ u\neq v}} \sum_{p,q} a_{p,u} a_{q,v}^* \sum_{l=0}^{L-1} x_u(lT_s - \Delta\tau_{p,u}) \times x_v^*((l+L)T_s - \Delta\tau_{q,v}) e^{j2\pi\Delta\nu lT_s}. \quad (36)$$

where $\Delta\nu = \nu_{u,\text{cfo}} - \nu_{v,\text{cfo}}$. This term can be neglected under standard cellular conditions, since SRSs associated with different cyclic shifts are quasi-orthogonal and CFO values are typically limited to a few kHz. The approximation holds when

$$\Delta\tau_{p,u} - \Delta\tau_{q,v} \neq (\alpha_u - \alpha_v)T_s L/8. \quad (37)$$

The normalization term $R[n_0]$ is given by

$$R[n_0] = \frac{1}{2}\left(\sum_{l=0}^{L-1} |y[l]|^2 + \sum_{l=0}^{L-1} |y[l+L]|^2\right). \quad (38)$$

Its detailed expansion follows the same structure as above and yields the weighting function $g_0(\Delta\tau_{p,u})$ analogous to $f_0(\Delta\tau_{p,u})$. The function $g_0(\Delta\tau_{p,u})$ can be expressed in piecewise form as

$$g(\Delta\tau_{p,u}) \approx \begin{cases} \dfrac{\Delta\tau_{p,u} + T_{CP} + 2T_L}{2T_L}, & -2T_L - T_{CP} \leq \Delta\tau_{p,u} < -T_{CP}, \\ 1, & -T_{CP} \leq \Delta\tau_{p,u} < 0, \\ 1 - \dfrac{\Delta\tau_{p,u}}{2T_L}, & 0 \leq \Delta\tau_{p,u} < 2T_L, \\ 0, & \text{otherwise}. \end{cases} \quad (39)$$

Figure 13 shows the real metric compared to the approximated one. Around $\Delta\tau_{1,1} = 13\,\mu$s, a notch appears due to the violation of the orthogonality condition in (37).

References

[1] Y. Feng, C. Zhao, H. Luo, F. Gao, F. Liu, and S. Jin, "Networked ISAC based UAV tracking and handover towards low-altitude economy," *IEEE Transactions on Wireless Communications*, 2025.

[2] M. Kishk, A. Bader, and M.-S. Alouini, "Aerial base station deployment in 6G cellular networks using tethered drones: The mobility and endurance tradeoff," *IEEE Vehicular Technology Magazine*, vol. 15, no. 4, pp. 103–111, 2020.

[3] M. Khurshid, M. Shahzad, H. A. Khattak, M. I. Malik, and M. M. Fraz, "Vision-based 3-D localization of UAV using deep image matching," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 12 020–12 030, 2024.

[4] B. Jankovic, S. Jangirova, W. Ullah, L. U. Khan, and M. Guizani, "Disaster detection on the fly: Optimized transformers for UAVs," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pp. 1–14, 2025.

[5] S. Moro, F. Linsalata, M. Manzoni, M. Magarini, and S. Tebaldini, "Enhancing user localization with an integrated sensing and communication (ISAC) system: An experimental UAV search-and-rescue use case," *Remote Sensing*, vol. 16, no. 16, p. 3031, 2024.

[6] F. Linsalata, A. Albanese, V. Sciancalepore, F. Roveda, M. Magarini, and X. Costa-Perez, "OTFS-superimposed PRACH-aided localization for UAV safety applications," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.

[7] O. Esrafilian, R. Mundlamuri, F. Kaltenberger, R. Knopp, and D. Gesbert, "First results on UAV-aided user localization using ToA and OpenAirInterface in 5G NR," *arXiv preprint arXiv:2503.19529*, 2025.

[8] A. Albanese, V. Sciancalepore, and X. Costa-Pérez, "SARDO: An automated search-and-rescue drone-based solution for victims localization," *IEEE Transactions on Mobile Computing*, vol. 21, no. 9, pp. 3312–3325, 2021.

[9] D. Mandloi and R. Arya, "Q-learning-based UAV-mounted base station positioning in a disaster scenario for connectivity to the users located at unknown positions," *The Journal of Supercomputing*, vol. 79, no. 14, pp. 15 643–15 674, 2023.

[10] S. Rashed and M. Soyturk, "Analyzing the effects of UAV mobility patterns on data collection in wireless sensor networks," *Sensors*, vol. 17, no. 2, 2017.

[11] D. Scazzoli, S. Moro, V. Teeda, P. K. Upadhyay, and M. Magarini, "Experimental comparison of UAV-based RSSI and AoA localization," *IEEE Sensors Letters*, vol. 8, no. 1, pp. 1–4, 2023.

[12] N. Paglierani, F. Linsalata, O. A. Sevimay, L. Cazzella, D. Badini, M. Magarini, and U. Spagnolini, "Digital network twin-enabled synchronization and localization," *IEEE Journal on Selected Areas in Communications*, pp. 1–1, 2025.

[13] H. Rahman and C. Rosenberg, "Resource allocation for the uplink of a multi-user massive MIMO system," *IEEE Transactions on Mobile Computing*, vol. 24, no. 5, pp. 4326–4338, 2025.

[14] J. P. González-Coma, J. Rodríguez-Fernández, N. González-Prelcic, L. Castedo, and R. W. Heath, "Channel estimation and hybrid precoding for frequency selective multiuser mmWave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 2, pp. 353–367, 2018.

[15] M. Petitjean, S. Mezhoud, and F. Quitin, "Fast localization of ground-based mobile terminals with a transceiver-equipped UAV," in *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2018, pp. 323–327.

[16] S. Masrur and İ. Güvenç, "Bridging simulation and reality: A 3D clustering-Based deep learning model for UAV-based RF source Localization," in *2025 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2025, pp. 953–958.

[17] Y. Luomei, F. Xu, F. Wang, and Y. Dong, "Demonstration of simultaneous localization and imaging with multirotor-borne miniSAR," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 6548–6558, 2022.

[18] R. Guo, Z. Ren, Y. Gao, L. Wei, Y. Qin, and G. Xu, "Improvement for uav urban SAR tomography based on cylindrical wave model with elevation constraints," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, pp. 11 401–11 415, 2025.

[19] E. Christy, R. P. Astuti, B. Syihabuddin, B. Narottama, O. Rhesa, and F. Rachmawati, "Optimum UAV flying path for device-to-device communications in disaster area," in *2017 International Conference on Signals and Systems (ICSigSys)*. IEEE, 2017, pp. 318–322.

[20] J. Wang, S. Kwon, and B. Shim, "Generalized orthogonal matching pursuit," *IEEE Transactions on signal processing*, vol. 60, no. 12, pp. 6202–6216, 2012.

[21] 3GPP, "5G; NR; physical channels and modulation," 3rd Generation Partnership Project (3GPP), Technical Specification TS 38.211, 2022, release 16.

[22] T. Schmidl and D. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Transactions on Communications*, vol. 45, no. 12, pp. 1613–1621, 1997.

[23] A. A. Neath and J. E. Cavanaugh, "The Bayesian information criterion: background, derivation, and applications," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 4, no. 2, pp. 199–203, 2012.