

# Development and Characterization of a 2D, 700–2600 MHz Receiver Array, for the Estimation of Direction of Arrival (DoA) in the Frequency Domain

DIEGO GALLARDO  
RICARDO FINGER  
DAVID MONASTERIO  
SEBASTIÁN JORQUERA  
FRANCO CUROTTO  
JUAN RIQUELME  
JOSÉ PIZARRO  
LEONARDO BRONFMAN

Universidad de Chile, Department of Astronomy, Santiago, Chile

**Abstract**—We present a  $4 \times 4$  antenna array receiver designed for real-time Direction of Arrival (DoA) estimation of radio frequency sources within the 700–2600 MHz range. Our device, called RadioVision, overlays detected radio source markers onto optical images captured by a video camera. It features a heterodyne receiver with a 70 MHz intermediate frequency bandwidth, 16 ADCs operating at 140 MSPS, and an FPGA for real-time FFT and linear algebra computations. We have implemented the 2D U-ESPRIT algorithm in the frequency domain using Polyphase Filter Banks (PFBs) and pipelined FFTs, which reduces the number of needed ADCs, simplifies the calibration process, and filters out unwanted frequency bands. Experiments conducted in both rural and urban environments demonstrated the system’s effectiveness. We achieved a mean squared error for DoA estimation of  $1.78^\circ$  at 90 m and  $3.79^\circ$  at 180 m in rural areas, and approximately  $5.76^\circ$  in urban settings at 60 m. This work represents the first on-field

Manuscript received XXXXX 00, 0000; revised XXXXX 00, 0000; accepted XXXXX 00, 0000.

This work was supported by the Chilean Agencia Nacional de Investigación y Desarrollo (ANID) through its grants Basal ACE210002 and FB210003.

Authors’ Addresses: Diego Gallardo, Ricardo Finger, David Monasterio, Sebastián Jorquera, Franco Curotto, Juan Riquelme, José Pizarro and Leonardo Bronfman are with the Department of Astronomy, Universidad de Chile, Santiago 8370451, Chile, E-mail: (diego.gallardo@raig.uchile.cl; rfinger@u.uchile.cl; david.monasterio@raig.uchile.cl; sebastian.jorquera@ug.uchile.cl; francocurotto@gmail.com; jriquelme@das.uchile.cl; jose.pizarro@raig.uchile.cl; leo@das.uchile.cl). (Corresponding author: Diego Gallardo).

Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>.

0018-9251 © 2024 IEEE

performance report of modern DoA algorithms, such as the 2D U-ESPRIT.

**Index Terms**—Antenna array, Digital processing, DoA estimation, FPGA, U-ESPRIT 2D.

## I. Introduction

DIRECTION of arrival (DoA) estimation using antenna arrays is crucial in many applications, such as mobile telecommunications, radar, and object tracking [1], [2], [3]. While there are numerous algorithms for DoA estimation, such as MUSIC [4], ESPRIT [5], U-ESPRIT [6], and U-ESPRIT 2D [7], most have only been validated through simulations, and may not perform well in real-life scenarios. In those situations, DoA estimation algorithms must address challenges such as the non-zero bandwidth of the receiver, potential imbalances in phase and amplitude due to imperfections in antennas and receivers, achieving high computing speed for real-time applications, interference from other sources -particularly base stations-, and multipath effects caused by reflections and scattering of signals.

The Bartlett beamformer was the first method used for source localization with antenna arrays. It employs an electronic sweep to find the direction of maximum power reception, indicating the source’s location [8], [9]. Over time, it evolved into the adaptive beamformer [10], which offers enhanced functionality. While Bartlett beamformer is widely applicable and easy to understand, it has significant limitations. Notably, its precision is contingent on the size of the antenna array, irrespective of factors such as the signal-to-noise ratio (SNR) or the temporal data collection process [2]. Due to these shortcomings, subspace-based methods emerged, such as MUSIC, ESPRIT, and U-ESPRIT 2D, which can achieve higher resolution independent of array size, given adequate data collection and SNR.

Sub-space methods have demonstrated remarkable performance over the years, achieving simulated source localization accuracy of less than one degree, even with signal-to-noise ratios less than 0 dB and antenna array sizes as small as  $4\lambda$  [7]. However, much of the research on these methods has focused on numerical simulations that exclude significant effects from reality. For example, the impact of floor reflections is typically excluded, which can be problematic in any application where the antenna array is not located on a high-rise tower. Moreover, ground reflections can be particularly troublesome in algorithms that estimate the number of sources [11], [12], [13], [14]. Another problem that must be solved is the presence of multiple unwanted sources, such as mobile telephone base stations. Indeed, most algorithms limit the number of sources for which the direction of arrival can be determined. For instance, if we have a one-dimensional array of  $N$  antennas, MUSIC and ESPRIT can determine only  $N - 1$  sources. If  $d$  interferers are also considered, this number decreases to  $N - 1 - d$ . In this sense, a base station with multiple incoherent emitters ( $d \geq N - 1$ ,

for example) can completely ruin the detection capability of sub-space methods. One last challenge arises from phase and amplitude imbalances presented in the non-zero bandwidth of the receiving system. When implementing sub-space algorithms in the time domain, as commonly done in the literature [4], [5], [6], [7], [15], [16], [17], [18], [19], calibrating these errors may necessitate the use of finite impulse response filters, adding complexity to both model design and resource utilization.

Although there are studies that have sought to experimentally verify subspace algorithms [16], [17], [18], [19], these have been conducted in controlled laboratory environments and do not explore realistic situations involving large distances in the open field, obstacles such as trees and buildings, and the presence of interference. Furthermore, the studies cited in [16], [17], [18], [19] only validate the subspace algorithms on the horizontal axis, failing to address the vertical axis, which presents challenges due to ground reflections.

Given the challenges and limitations of the existing methods, we implemented subspace methods in the frequency domain. While this is a known theoretical technique, it has not been widely considered until now due to the higher computational cost and slowness in the Fast Fourier Transform (FFT) calculation. However, technological advances have made it possible to calculate the FFT in real time using field-programmable gate arrays (FPGAs). This implementation offers several advantages, including the elimination of half of the ADCs necessary to form I-Q inputs (or the elimination of direct digital synthesizers (DDS) otherwise), the calibration of amplitude and phase imbalances across the entire frequency range, the identification and flagging of unwanted frequency bands, and, if the modulation type allows it, the precise calculation of the carrier frequency, which could lead to a better estimation of the DoAs. This proposed solution is a significant step forward in the field of DoA estimation, addressing fundamental challenges and opening up new possibilities for real-time applications.

In this article, we present the design, construction, and testing of RadioVision: a system for real-time localization of radio frequency (RF) sources in the 700–2600 MHz range. For signal reception, the device incorporates an analog system comprising a  $4 \times 4$  antenna array and a heterodyne receiver with double-sideband (DSB) mixers and a variable local oscillator (LO), allowing the selection of 70 MHz sub-bands within the 700–2600 MHz range. For digital and real-time calculation of DoAs, the signals are digitized by 16 ADCs operating at 140 MSPS. The device implements polyphase filter banks (PFBs) and the frequency domain version of the U-ESPRIT 2D algorithm on an FPGA, which enables its use outside of a laboratory environment. We demonstrate its applicability in open terrain and buildings with significant electromagnetic interference. Furthermore, we validate its effectiveness using non-synthetic laboratory sources, specifically cell phones operating across different telecommunications bands. We highlight that RadioVision is a robust device that serves

TABLE I

Cell phone bands tendered in Chile. UL: up-link, DL: down-link, BW: bandwidth.

Band	UL (MHz)	DL (MHz)	UL BW (MHz)
28 APT (LTE)	703–748	758–803	45
5 CLR (UMTS)	824–849	869–894	25
8 E-GSM (UMTS)	902–912	947–957	10
10 AWS (LTE)	1710–1770	2110–2170	60
2 PCS (LTE)	1850–1910	1930–1990	60
7 IMT-E (LTE)	2505–2565	2625–2685	60

as a test bench for multiple algorithms, such as MUSIC, ESPRIT, U-ESPRIT, and several source number estimators, which, until now, have only been tested through simulations or in the laboratory.

We introduce the notation used throughout the paper to make the presentation as straightforward as possible. Uppercase boldface characters refer to matrices, while lowercase boldface characters refer to vectors. We use  $(\mathbf{A})_{mn}$  to denote the  $(m, n)$ -th entry of  $\mathbf{A}$ , and we start the numbering with  $m = 0$  and  $n = 0$  (i.e.,  $(\mathbf{A})_{00}$  is the element in the first row and the first column). The notation  $\mathbf{A}^T$ ,  $\mathbf{A}^*$  and  $\mathbf{A}^H$  denote, respectively, the transpose, the conjugate, and the transpose conjugate of  $\mathbf{A}$ . We use  $j = \sqrt{-1}$  to represent imaginary numbers,  $\Re\{x\}$  and  $\Im\{x\}$  to denote the real and imaginary parts of  $x$ , and  $\otimes$  to denote the Kronecker product. Finally, some useful matrices used throughout the paper are  $\mathbf{I}_N$ , which denotes the  $N \times N$  identity matrix,  $\mathbf{\Pi}_N$ , which is the  $N \times N$  exchange matrix with ones in its antidiagonal and zeros elsewhere, and  $\mathbf{0} = [0 \ 0 \ \dots \ 0]^T$ , which is the zero-vector whose dimension can be deduced from context.

## II. System Description

The primary objective of this article is to validate and evaluate the performance of RadioVision: a system that implements the frequency domain version of U-ESPRIT 2D and is capable of locating multiple narrowband ( $\leq 70$  MHz) radio sources distributed across a wide range of frequencies (700–2600 MHz). To demonstrate the effectiveness and practicality of the proposed instrument, we have chosen the localization of cell phones as a representative scenario. The UMTS and LTE frequency bands we considered to validate RadioVision are those tendered in Chile, as shown in Table I.

While the proposed system has many applications, we are particularly interested in locating cell phones in various scenarios -stationary or moving, in open fields or behind walls- and across multiple operation modes -texting, calling, or standby. To achieve this, RadioVision performs rapid DoA estimation with sub-millisecond response times, aligning with the beacon duration of cell phones. Furthermore, since cell phones switch between operating bands, the system must be able to switch between different bands quickly.

A simplified schematic of RadioVision is shown in Fig. 1. The system features a  $4 \times 4$  antenna array to receive

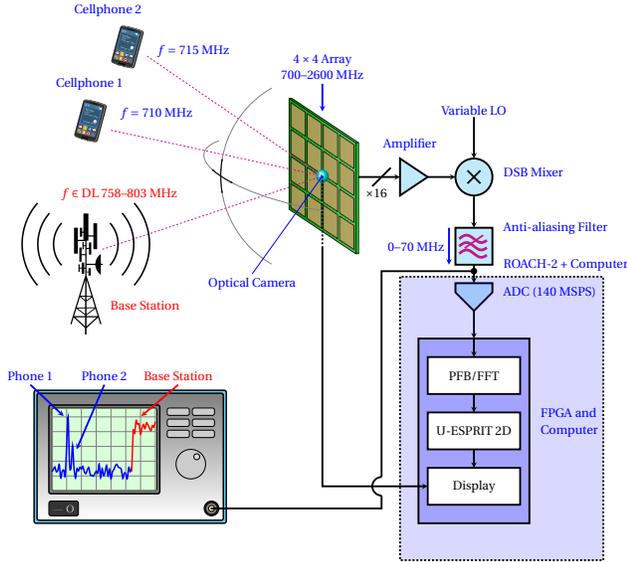


Fig. 1. Simplified schematic of RadioVision. As an example, we consider two cell phones in the UL range of 703–748 MHz and a base station in the DL range of 758–803 MHz. The variable LO allows us to choose the bands from Table I. For this example, we set the LO at 700 MHz.

phone signals, covering all up-links (UL) listed in Table I. The signals are then amplified and down-converted to baseband using DSB mixers and anti-aliasing filters, allowing band selection by adjusting the LO. The instrument incorporates a ROACH-2 platform (Reconfigurable Open Architecture Computing Hardware) developed by the CASPER group (Collaboration for Astronomy Signal Processing and Electronics Research). The board’s key components include the Xilinx Virtex-6 XC6VVSX475T FPGA and 16 HMCAD1511 ADCs with a sampling rate of 140 MSPS. Collected samples pass through PFBs and FFT calculators implemented in the FPGA, allowing real-time spectra calculation, facilitating amplitude and phase calibration, and flagging.

After calculating the DoAs, we display the positions as markers on an optical image taken with a camera located at the array’s center. This approach makes RadioVision easier to operate and debug. Furthermore, with this approach, we can conduct tests with moving sources and compare the estimated DoAs with the actual positions after obtaining the latter with the camera recording. In this manner, the integration with the camera allows us to evaluate RadioVision’s performance in multiple source positions, in contrast to what is done in [16], [17], [18], [19], which keeps the sources in fixed positions.

### III. Signal Model

In this section, we develop the mathematical model for the signals an antenna array receives in the presence of multiple sources. Since the literature has covered this development extensively [4], [5], [6], [7], [15], [20], we will emphasize only the critical aspects of our system.

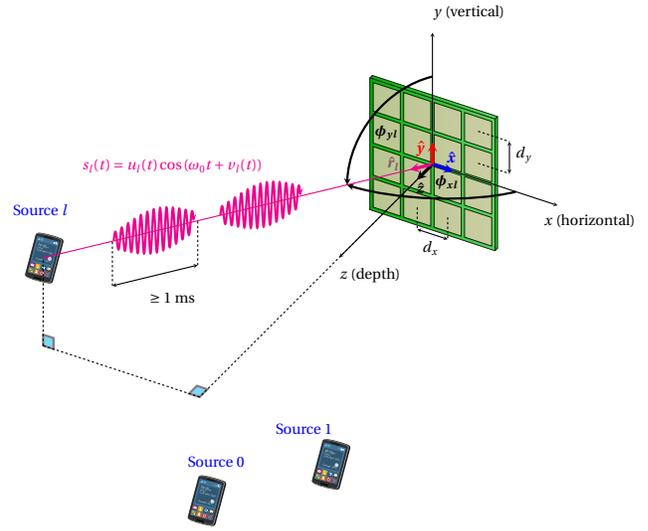


Fig. 2. Illustration of a  $4 \times 4$  ( $P \times Q$ ) rectangular array surrounded by three ( $r$ ) sources (phones).  $d_x$  and  $d_y$  are the spacings in the  $x$  and  $y$  axes, respectively.  $\phi_{xl}$  and  $\phi_{yl}$  are the angles of arrival of the  $l$ -th source measured relative to the  $x$  and  $y$  axes, respectively. All sources share the same carrier frequency  $\omega_0$ . Each source  $l$  has its own amplitude and phase modulations  $u_l(t)$  and  $v_l(t)$ , with  $l = 0, 1, \dots, r - 1$ . Depending on the mode of operation, phones can radiate packets with durations as short as 1 ms.

To determine an expression for the signals arriving at the ADCs, let us consider a uniform rectangular array (URA) of  $P \times Q$  elements ( $4 \times 4$  in the case of RadioVision) and  $r$  sources (cell phones in our application), as shown in Fig. 2. We will consider the origin of the coordinate system at the center of the antenna array and use the indices  $p \in \{0, 1, \dots, P - 1\}$  and  $q \in \{0, 1, \dots, Q - 1\}$  so that the  $(p, q)$ -th antenna is the one centered at  $((-\frac{P-1}{2} + p) d_x, (-\frac{Q-1}{2} + q) d_y)$ , where  $d_x$  and  $d_y$  are the antenna spacings in the  $x$  and  $y$  axes, respectively.

We consider the sources narrowband, non-coherent, and located within an isotropic and non-dispersive medium in the far field. This allows us to consider their radiation as plane waves when impinging on the array. Although our system can choose sub-bands by varying the LO, RadioVision receives only one narrow 70 MHz band at a time. Therefore, we will consider that the  $r$  sources share the same carrier frequency  $\omega_0$ . With all this, we can express the signals  $\{s_l\}_{l=0}^{r-1}$  radiated by the sources as  $s_l(t) = u_l(t) \cos(\omega_0 t + v_l(t))$ , where  $u_l(t)$  and  $v_l(t)$  are the amplitude and phase modulations of  $s_l(t)$ , which vary slowly compared to  $\cos(\omega_0 t)$  (since  $\{s_l\}_{l=0}^{r-1}$  are narrowband signals).

The position of the  $l$ -th source is determined by the angles  $\phi_{xl} = \cos^{-1}(\hat{\mathbf{r}}_l \cdot \hat{\mathbf{x}})$  and  $\phi_{yl} = \cos^{-1}(\hat{\mathbf{r}}_l \cdot \hat{\mathbf{y}})$ , where  $\hat{\mathbf{x}}$  and  $\hat{\mathbf{y}}$  are the standard unit vectors of the Cartesian coordinate system, and  $\hat{\mathbf{r}}_l$  is the unit vector pointing towards the  $l$ -th source (see Fig. 2). Note that when the source is at the center, we have  $\phi_{xl} = \phi_{yl} = 90^\circ$ . By convention, it is best to define  $\theta_{xl} = 90^\circ - \phi_{xl}$  and  $\theta_{yl} = 90^\circ - \phi_{yl}$ , such that  $\theta_{xl} = \sin^{-1}(\hat{\mathbf{r}}_l \cdot \hat{\mathbf{x}})$  and

$\theta_{yl} = \sin^{-1}(\hat{\mathbf{r}}_l \cdot \hat{\mathbf{y}})$ . Hence, for a source at the center,  $\theta_{xl} = \theta_{yl} = 0^\circ$ .

Considering all the definitions above, and following the traditional delayed-propagation model shown in [4], [5], [6], [7], [15], [20], we get that the  $l$ -th signal impinging on the  $(p, q)$ -th antenna is given by  $x_{pql}(t) = s_l(t - \tau_{pq}(\theta_{xl}, \theta_{yl}))$ , where

$$\begin{aligned} \tau_{pq}(\theta_{xl}, \theta_{yl}) &= \tau_{pql} = \tau_{pl,x} + \tau_{ql,y} \\ &= \left( \frac{P-1}{2} - p \right) \frac{d_x \sin \theta_{xl}}{c} + \left( \frac{Q-1}{2} - q \right) \frac{d_y \sin \theta_{yl}}{c} \end{aligned} \quad (1)$$

is the propagation delay between the  $(p, q)$ -th antenna and the array's center, and  $c$  is the speed of light. Then, since  $u_l(t)$  and  $v_l(t)$  vary slowly compared to  $\cos(\omega_0 t)$ , we can approximate

$$\begin{aligned} x_{pql}(t) &= u_l(t - \tau_{pql}) \cos(\omega_0(t - \tau_{pql}) + v_l(t - \tau_{pql})) \\ &\approx u_l(t) \cos(\omega_0(t - \tau_{pql}) + v_l(t)) \\ &= \frac{1}{2} \left( e^{-j(\mu_{pl} + \nu_{ql})} \xi_l(t) + e^{j(\mu_{pl} + \nu_{ql})} \xi_l^*(t) \right), \end{aligned} \quad (2)$$

where we have defined  $\mu_{pl} = \omega_0 \tau_{pl,x}$ ,  $\nu_{ql} = \omega_0 \tau_{ql,y}$ , and  $\xi_l(t) = u_l(t) e^{j(\omega_0 t + v_l(t))}$ .

The  $P \times Q$  antennas receive the sum of the  $r$  signals centered at  $\omega_0$ , each one modified by the radiation pattern response at  $(\theta_{xl}, \theta_{yl})$ . Subsequently, all the received signals pass through identical heterodyne receivers, composed of amplifiers, low-pass filters, and mixers with an LO frequency  $\omega_{LO}$ , for posterior digitization by  $P \times Q$  ADCs. If we assume that all receiver chains are equal, the analog signals sensed by the ADCs can be written as

$$z_{pq}(t) = \frac{1}{2} \sum_{l=0}^{r-1} \left( e^{-j(\mu_{pl} + \nu_{ql})} \chi_l(t) + e^{j(\mu_{pl} + \nu_{ql})} \chi_l^*(t) \right), \quad (3)$$

where  $\{\chi_l(t)\}_{l=0}^{r-1}$  are baseband signals centered at  $\omega_0 - \omega_{LO}$ , and they comprise the radiation pattern and front-end modifications over the RF signals  $x_{pql}(t)$ .

Note that in the standard DoA literature, only terms of the form  $e^{-j(\mu_{pl} + \nu_{ql})} \chi_l(t)$  are considered. This simplification arises because phasors are used, and the need for both in-phase (I) and quadrature (Q) components is omitted in the mathematical formulation of the algorithms and simulations. However, in the implementation of a real system, both terms  $e^{-j(\mu_{pl} + \nu_{ql})} \chi_l(t)$  and  $e^{j(\mu_{pl} + \nu_{ql})} \chi_l^*(t)$  appear in the received signals, as shown in (3). This discrepancy invalidates the standard procedures followed in the literature [4], [5], [6], [7], which rely on received signals expressed solely in terms of  $e^{-j(\mu_{pl} + \nu_{ql})} \chi_l(t)$ .

To be able to apply the standard procedures, three methods allow us to obtain only terms of the form  $e^{-j(\mu_{pl} + \nu_{ql})} \chi_l(t)$  from the real signals stated in (3). The first method is analog and uses I-Q mixers. In this

case, more complex mixers are needed, and the post-intermediate frequency (IF) hardware and the number of ADCs are also doubled. The second method is digital and consists of using digital down-converters [21], [22]. For this, DDSs must be implemented to form cosines and sines, in addition to multiple digital filters, which implies an increase in the use of digital resources. As can be seen, both methods require increasing the system's complexity and resource usage. In this sense, the third method, which consists of using the FFT, is optimal since the benefits that come with the use of the FFT are multiple, as mentioned in Sections I and II.

#### IV. 2D U-ESPRIT with FFT applied to the time domain

Let us take the expression obtained in (3) and consider the digitization of the signals by the ADCs, where we will write the discrete-time dependence with  $(n)$  instead of  $(t)$ . Given this, we can express the  $P \times Q$  signals entering the digital system as

$$\mathbf{Z}(n) = \frac{1}{2} \sum_{l=0}^{r-1} \mathbf{a}_{Pl} \mathbf{a}_{Ql}^T \chi_l(n) + \frac{1}{2} \sum_{l=0}^{r-1} \mathbf{a}_{Pl}^* \mathbf{a}_{Ql}^H \chi_l^*(n), \quad (4)$$

where  $\mathbf{Z}(n)$  is a  $P \times Q$  matrix whose  $(p, q)$ -th entry is  $z_{pq}(n)$ , and  $\{\mathbf{a}_{Pl}\}_{l=0}^{r-1}$  and  $\{\mathbf{a}_{Ql}\}_{l=0}^{r-1}$  are  $P \times 1$  and  $Q \times 1$  vectors given by  $\mathbf{a}_{Pl} = [e^{-j\mu_{0l}} \ e^{-j\mu_{1l}} \ \dots \ e^{-j\mu_{(P-1)l}}]^T$  and  $\mathbf{a}_{Ql} = [e^{-j\nu_{0l}} \ e^{-j\nu_{1l}} \ \dots \ e^{-j\nu_{(Q-1)l}}]^T$ .

After computing the  $N$ -sized FFT of each  $z_{pq}(n)$ , we can write

$$\begin{aligned} \hat{\mathbf{Z}}(k) &= \frac{1}{2} \sum_{l=0}^{r-1} \mathbf{a}_{Pl} \mathbf{a}_{Ql}^T \sum_{n=0}^{N-1} \chi_l(n) e^{-j \frac{2\pi nk}{N}} \\ &\quad + \frac{1}{2} \sum_{l=0}^{r-1} \mathbf{a}_{Pl}^* \mathbf{a}_{Ql}^H \sum_{n=0}^{N-1} \chi_l^*(n) e^{-j \frac{2\pi nk}{N}}, \end{aligned} \quad (5)$$

being  $\hat{\mathbf{Z}}(k)$  a  $P \times Q$  matrix whose  $(p, q)$ -th entry is the  $k$ -th frequency bin of the FFT of the sequence  $\{z_{pq}(n)\}_{n=0}^{N-1}$ . Then, we introduce the centrosymmetric computations performed in U-ESPRIT [6], [7] to obtain

$$\hat{\mathbf{Z}}'(k) = \mathbf{Q}_P^H \hat{\mathbf{Z}}(k) \mathbf{Q}_Q^* = \frac{1}{2} \sum_{l=0}^{r-1} \boldsymbol{\Psi}_l \hat{\chi}_{1l}(k) + \frac{1}{2} \sum_{l=0}^{r-1} \boldsymbol{\eta}_l \hat{\chi}_{2l}, \quad (6)$$

where  $\mathbf{Q}_P$  and  $\mathbf{Q}_Q$  are defined by the general formulas

$$\mathbf{Q}_{2K} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_K & j\mathbf{I}_K \\ \mathbf{\Pi}_K & -j\mathbf{\Pi}_K \end{bmatrix} \quad (7)$$

if  $P$  or  $Q$  are even numbers, and

$$\mathbf{Q}_{2K+1} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_K & \mathbf{0} & j\mathbf{I}_K \\ \mathbf{0}^T & \sqrt{2} & \mathbf{0}^T \\ \mathbf{\Pi}_K & \mathbf{0} & -j\mathbf{\Pi}_K \end{bmatrix} \quad (8)$$

if  $P$  or  $Q$  are odd numbers, while  $\{\hat{\chi}_{1l}(k)\}_{k=0}^{N-1}$  and  $\{\hat{\chi}_{2l}(k)\}_{k=0}^{N-1}$  are the FFTs of  $\{\chi_l(n)\}_{n=0}^{N-1}$  and  $\{\chi_l^*(n)\}_{n=0}^{N-1}$ , respectively, and  $\Psi_l$  and  $\eta_l$  are  $P \times Q$  real matrices given by

$$\Psi_l = (\mathbf{Q}_P^H \mathbf{a}_{Pl})(\mathbf{Q}_Q^H \mathbf{a}_{Ql})^T \quad (9)$$

$$\eta_l = (\mathbf{Q}_P^H \mathbf{a}_{Pl}^*)(\mathbf{Q}_Q^H \mathbf{a}_{Ql}^*)^T. \quad (10)$$

If we consider  $\chi_l(n) = f_l(n) + jg_l(n)$ , we can write (6) as

$$\begin{aligned} \hat{\mathbf{Z}}'(k) &= \sum_{l=0}^{r-1} \frac{\Psi_l + \eta_l}{2} \sum_{n=0}^{N-1} f_l(n) e^{-j \frac{2\pi nk}{N}} \\ &+ j \sum_{l=0}^{r-1} \frac{\Psi_l - \eta_l}{2} \sum_{n=0}^{N-1} g_l(n) e^{-j \frac{2\pi nk}{N}} \quad (11) \\ &= \sum_{l=0}^{r-1} \mathbf{M}_{1l} \hat{f}_l(k) + j \sum_{l=0}^{r-1} \mathbf{M}_{2l} \hat{g}_l(k), \end{aligned}$$

where  $\mathbf{M}_{1l} = (\Psi_l + \eta_l)/2$ ,  $\mathbf{M}_{2l} = (\Psi_l - \eta_l)/2$ , and  $\{\hat{f}_l(k)\}_{k=0}^{N-1}$  and  $\{\hat{g}_l(k)\}_{k=0}^{N-1}$  are the FFTs of  $\{f_l(n)\}_{n=0}^{N-1}$  and  $\{g_l(n)\}_{n=0}^{N-1}$ , respectively. Then, to align with the standard ESPRIT-like mathematical developments, we must write each matrix  $\hat{\mathbf{Z}}'(k)$  in vector form, which is accomplished with the  $\text{vec}(\cdot)$  operation:  $\hat{\mathbf{z}}'(k) = \text{vec}(\hat{\mathbf{Z}}'(k))$ . In this way, the  $(p, q)$ -th entry of  $\hat{\mathbf{Z}}'(k)$  is mapped into the  $(P(q-1) + p)$ -th position of the  $PQ \times 1$  complex vector  $\hat{\mathbf{z}}'(k)$ . After applying the  $\text{vec}(\cdot)$  operation to the expression of  $\hat{\mathbf{Z}}'(k)$  given in (11), we get

$$\begin{aligned} \hat{\mathbf{z}}'(k) &= \sum_{l=0}^{r-1} \text{vec}(\mathbf{M}_{1l}) \hat{f}_l(k) + j \sum_{l=0}^{r-1} \text{vec}(\mathbf{M}_{2l}) \hat{g}_l(k) \\ &= [\text{vec}(\mathbf{M}_{10}) \text{vec}(\mathbf{M}_{11}) \dots \text{vec}(\mathbf{M}_{1(r-1)})] \hat{\mathbf{f}}(k) \\ &+ j [\text{vec}(\mathbf{M}_{20}) \text{vec}(\mathbf{M}_{21}) \dots \text{vec}(\mathbf{M}_{2(r-1)})] \hat{\mathbf{g}}(k) \\ &= \mathbf{S}_{m1} \hat{\mathbf{f}}(k) + j \mathbf{S}_{m2} \hat{\mathbf{g}}(k), \quad (12) \end{aligned}$$

where  $\mathbf{S}_{m1} = [\text{vec}(\mathbf{M}_{10}) \text{vec}(\mathbf{M}_{11}) \dots \text{vec}(\mathbf{M}_{1(r-1)})]$  and  $\mathbf{S}_{m2} = [\text{vec}(\mathbf{M}_{20}) \text{vec}(\mathbf{M}_{21}) \dots \text{vec}(\mathbf{M}_{2(r-1)})]$  are  $PQ \times r$  real matrices, and  $\hat{\mathbf{f}}(k) = [\hat{f}_0(k) \dots \hat{f}_{r-1}(k)]^T$  and  $\hat{\mathbf{g}}(k) = [\hat{g}_0(k) \dots \hat{g}_{r-1}(k)]^T$  are  $r \times 1$  real vectors corresponding to the  $k$ -th frequency bin of the FFTs of the vector sequences  $\mathbf{f}(n) = [f_0(n) \dots f_{r-1}(n)]^T$  and  $\mathbf{g}(n) = [g_0(n) \dots g_{r-1}(n)]^T$ . Then, we can stack all vectors  $\hat{\mathbf{z}}'(k)$  from  $k = 0$  to  $k = N/2 - 1$  (which corresponds to taking half the FFT spectrum) to form the matrix

$$\begin{aligned} \hat{\mathbf{z}} &= [\hat{\mathbf{z}}'(0) \dots \hat{\mathbf{z}}'(N/2 - 1)] \\ &= \mathbf{S}_{m1} [\hat{\mathbf{f}}(0) \dots \hat{\mathbf{f}}(N/2 - 1)] + \mathbf{S}_{m2} [\hat{\mathbf{g}}(0) \dots \hat{\mathbf{g}}(N/2 - 1)] \\ &= \mathbf{S}_{m1} \mathbf{F}_N \mathbf{W}_H + j \mathbf{S}_{m2} \mathbf{G}_N \mathbf{W}_H, \quad (13) \end{aligned}$$

being  $\mathbf{F}_N = [\mathbf{f}(0) \dots \mathbf{f}(N-1)]$ ,  $\mathbf{G}_N = [\mathbf{g}(0) \dots \mathbf{g}(N-1)]$  and  $\mathbf{W}_H$  the  $N \times N/2$  FFT matrix obtained after removing half the spectrum, given by

$$\mathbf{W}_H = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & W_N & \dots & W_N^{N/2-1} \\ \vdots & \vdots & \dots & \vdots \\ 1 & W_N^{N-1} & \dots & W_N^{(N/2-1)(N-1)} \end{bmatrix}, \quad (14)$$

where  $W_N = e^{-j \frac{2\pi}{N}}$ . The rationale for considering only half of the FFT spectrum lies in the fact that, for real-valued signals, the real part of the FFT is even, while the imaginary part is odd. Consequently, the phase contributions from each half of the spectrum, derived from the imaginary to the real part ratio, are opposite in sign. This would result in erroneous DoA estimates if both halves were considered.

After aligning with the standard ESPRIT-like formulations, we can compute the real part of the correlation matrix as

$$\begin{aligned} \mathbb{R}\{\mathbf{R}_{zz}\} &= \frac{1}{2} (\mathbf{R}_{zz} + \mathbf{R}_{zz}^*) \\ &= [\mathbf{S}_{m1} \mathbf{F}_N \quad \mathbf{S}_{m2} \mathbf{G}_N] \begin{bmatrix} \mathbf{S}_{wr} & \mathbf{S}_{wi} \\ -\mathbf{S}_{wi} & \mathbf{S}_{wr} \end{bmatrix} \begin{bmatrix} \mathbf{F}_N^T \mathbf{S}_{m1}^T \\ \mathbf{G}_N^T \mathbf{S}_{m2}^T \end{bmatrix} \\ &= \mathbf{S}_m \mathbf{\Gamma} \mathbf{S}_m^T, \quad (15) \end{aligned}$$

where we have defined  $\mathbf{S}_m = [\mathbf{S}_{m1} \mathbf{F}_N \quad \mathbf{S}_{m2} \mathbf{G}_N]$ ,  $\mathbf{S}_w = \mathbf{W}_H \mathbf{W}_H^H$ ,  $\mathbf{S}_{wr} = \mathbb{R}\{\mathbf{S}_w\}$ ,  $\mathbf{S}_{wi} = \mathbb{I}\{\mathbf{S}_w\}$ , and

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{S}_{wr} & \mathbf{S}_{wi} \\ -\mathbf{S}_{wi} & \mathbf{S}_{wr} \end{bmatrix}. \quad (16)$$

Now, since the signals are not correlated,  $\mathbf{F}_N$  and  $\mathbf{G}_N$  have full rank  $r$ . Therefore, the column space of  $\mathbf{S}_{m1} \mathbf{F}_N$  equals the column space of  $\mathbf{S}_{m1}$ , and the column space of  $\mathbf{S}_{m2} \mathbf{G}_N$  equals the column space of  $\mathbf{S}_{m2}$ . Hence, the column space of  $[\mathbf{S}_{m1} \mathbf{F}_N \quad \mathbf{S}_{m2} \mathbf{G}_N]$  includes the column space of  $\mathbf{S}_{m1} + \mathbf{S}_{m2}$ . In consequence, there are  $r$  eigenvectors of  $\mathbb{R}\{\mathbf{R}_{zz}\}$  that span the column space of  $\mathbf{S}_{m1} + \mathbf{S}_{m2}$ . Now, notice that  $\mathbf{S}_{m1} + \mathbf{S}_{m2}$  is just  $\mathbf{\Psi} = [\text{vec}(\Psi_0) \dots \text{vec}(\Psi_{r-1})]$ . Hence, if we concatenate those aforementioned  $r$  eigenvectors as the columns of a matrix  $\mathbf{V}_r$ , we can write  $\mathbf{\Psi} = \mathbf{V}_r \mathbf{T}$ , where  $\mathbf{T}$  is a non-singular  $r \times r$  real matrix. From now on, we can follow the standard ESPRIT-like procedure documented in [7].

Specifically, we must consider that each  $vec(\Psi_l)$  follows the invariance relationships [7]

$$\begin{aligned} \tan\left(\frac{\mu_l}{2}\right) \mathbf{K}_{\mu 1} vec(\Psi_l) &= \mathbf{K}_{\mu 2} vec(\Psi_l) \\ \tan\left(\frac{\nu_l}{2}\right) \mathbf{K}_{\nu 1} vec(\Psi_l) &= \mathbf{K}_{\nu 2} vec(\Psi_l), \end{aligned} \quad (17)$$

where  $\mu_l = (\omega_0 d_x \sin \theta_{xl})/c$ ,  $\nu_l = (\omega_0 d_y \sin \theta_{yl})/c$ ,  $\mathbf{K}_{\mu 1}$  and  $\mathbf{K}_{\mu 2}$  are  $(P-1)Q \times PQ$  matrices given by

$$\mathbf{K}_{\mu 1} = \mathbf{I}_Q \otimes \Re\{\mathbf{Q}_{P-1}^H \mathbf{J}_P \mathbf{Q}_P\} \quad (18)$$

$$\mathbf{K}_{\mu 2} = \mathbf{I}_Q \otimes \Im\{\mathbf{Q}_{P-1}^H \mathbf{J}_P \mathbf{Q}_P\}, \quad (19)$$

and  $\mathbf{K}_{\nu 1}$  and  $\mathbf{K}_{\nu 2}$  are  $(Q-1)P \times PQ$  matrices given by

$$\mathbf{K}_{\nu 1} = \Re\{\mathbf{Q}_{Q-1}^H \mathbf{J}_Q \mathbf{Q}_Q\} \otimes \mathbf{I}_P \quad (20)$$

$$\mathbf{K}_{\nu 2} = \Im\{\mathbf{Q}_{Q-1}^H \mathbf{J}_Q \mathbf{Q}_Q\} \otimes \mathbf{I}_P, \quad (21)$$

where  $\mathbf{J}_P$  and  $\mathbf{J}_Q$  are  $(P-1) \times P$  and  $(Q-1) \times Q$  selection matrices given by

$$\mathbf{J}_P = \begin{bmatrix} \mathbf{0} & \mathbf{I}_{P-1} \end{bmatrix}, \mathbf{J}_Q = \begin{bmatrix} \mathbf{0} & \mathbf{I}_{Q-1} \end{bmatrix}. \quad (22)$$

After concatenating the  $r$  invariance relationships (17) (from  $l=0$  to  $l=r-1$ ) into a matrix form, we obtain

$$\mathbf{K}_{\mu 1} \Psi \Omega_\mu = \mathbf{K}_{\mu 2} \Psi \quad (23)$$

$$\mathbf{K}_{\nu 1} \Psi \Omega_\nu = \mathbf{K}_{\nu 2} \Psi, \quad (24)$$

and after using  $\Psi = \mathbf{V}_r \mathbf{T}$ , we get

$$\begin{aligned} \mathbf{K}_{\mu 1} \mathbf{V}_r (\mathbf{T} \Omega_\mu \mathbf{T}^{-1}) &= \mathbf{K}_{\mu 2} \mathbf{V}_r \\ \mathbf{K}_{\nu 1} \mathbf{V}_r (\mathbf{T} \Omega_\nu \mathbf{T}^{-1}) &= \mathbf{K}_{\nu 2} \mathbf{V}_r, \end{aligned} \quad (25)$$

where  $\Omega_\mu$  and  $\Omega_\nu$  are diagonal matrices such that  $(\Omega_\mu)_{ll} = \tan \frac{\mu_l}{2}$  and  $(\Omega_\nu)_{ll} = \tan \frac{\nu_l}{2}$ . Finally, we can solve the two systems of equations (25) to obtain  $(\mathbf{T} \Omega_\mu \mathbf{T}^{-1})$  and  $(\mathbf{T} \Omega_\nu \mathbf{T}^{-1})$ , which are similar matrices to  $\Omega_\mu$  and  $\Omega_\nu$ , and therefore share the eigenvalues, which in turn contain the DoAs via the expressions  $\tan \frac{\mu_l}{2}$  and  $\tan \frac{\nu_l}{2}$ . The practical application of the development just described is summarized below.

- 1) Calculate the  $N$ -sized FFT of the  $P \times Q$  data streams received by the antenna array and eliminate half of the spectrum.
- 2) Perform the centrosymmetric operations stated in (6), via the matrices  $\mathbf{Q}_P$  and  $\mathbf{Q}_Q$ .
- 3) Calculate the real part of the correlation matrix  $\mathbf{R}_{zz}$  of the antenna array using each frequency bin as a sample.
- 4) Considering  $r$  sources (assumed as a known number in this work), compute the eigendecomposition of  $\Re\{\mathbf{R}_{zz}\}$  and select the  $r$  eigenvectors associated with the  $r$  greatest eigenvalues. Concatenate those  $r$  eigenvectors into a matrix  $\mathbf{V}_r$ .
- 5) Solve the linear equations  $\mathbf{K}_{\mu 1} \mathbf{V}_r \mathbf{X} = \mathbf{K}_{\mu 2} \Psi$  and  $\mathbf{K}_{\nu 1} \mathbf{V}_r \mathbf{Y} = \mathbf{K}_{\nu 2} \Psi$  to find  $\mathbf{X}$  and  $\mathbf{Y}$ .

- 6) Form the complex matrix  $\mathbf{X} + j\mathbf{Y}$  and find its complex eigenvalues  $\{\lambda_l\}_{l=0}^{r-1}$ .
- 7) Obtain the estimated DoAs  $(\theta_{xl}, \theta_{yl})$  for the  $l$ -th signal as:

$$\theta_{xl} = \sin^{-1} \left( \frac{2c}{\omega_0 d_x} \tan^{-1}(\Re\{\lambda_l\}) \right) \quad (26)$$

$$\theta_{yl} = \sin^{-1} \left( \frac{2c}{\omega_0 d_y} \tan^{-1}(\Im\{\lambda_l\}) \right) \quad (27)$$

## V. Hardware Design

In this section, we show the design and implementation of RadioVision. The system comprises an analog and digital part, as shown in Fig. 3. Each of these are described in the following subsections.

### A. Analog system

The analog system comprises two parts: the antenna array and the receiver itself. Since our system is intended for phone localization, an antenna array that works over the entire mobile phone range is required. For this, we used the antenna presented in [23], which has several advantages. First, it covers all the up-links from the UMTS and LTE bands tendered in Chile (700–2600 MHz). Second, it has a wide average half-power beamwidth of  $72^\circ$ , which defines the field of view of the localization system. Third, it has a half-space radiation pattern, which is necessary for a two-dimensional array (because of phase ambiguities between the front and back of the array), and prevents interference and coupling with the reception system. Finally, its small size (7.5 cm) allows operation without phase ambiguities (or grating lobes) across the entire bandwidth in the field of view.

The receiver can be decomposed into three parts, enumerated as (1)–(3) in Fig. 3. The first part (1) is a series capacitor with a cut-off frequency of 100 MHz, which acts as a high-pass filter and prevents AM and FM radio from entering the system and eventually reaching the IF (0–70 MHz). The second part (2) has the typical elements expected on a heterodyne receiver, which are RF and IF amplifiers, a DSB-mixer to obtain the baseband signal (up to 70 MHz), and a low-pass filter to prevent aliasing in the ADCs. It also has a 1-to-16 splitter to distribute the LO signal. The conversion gains of the 16 receiver chains are shown in Fig. 4 for all the LO configurations needed to address the up-links from Table I. Note that the placement of the LO frequency is important for the specific task of locating phones. In the context of DSB mixers, where there is no differentiation between the upper and lower sidebands (USB and LSB), the location of phones (which operate in the UL) can be significantly influenced by base stations (which operate in the DL) if both UL and DL are overlapped in the IF. For example, as can be seen in Fig. 4, if the LO is placed at 2.5 GHz, the USB is in the desired UL of 2.505–2.565 GHz, but the LSB falls in the 2.4 GHz WiFi Band. As both bands are mixed in a DSB

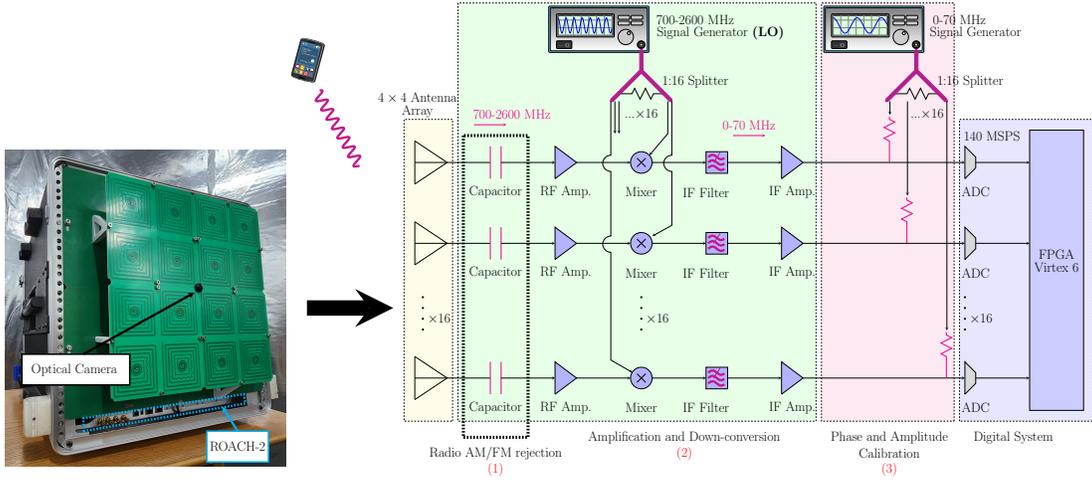


Fig. 3. Schematic representation of RadioVision hardware.

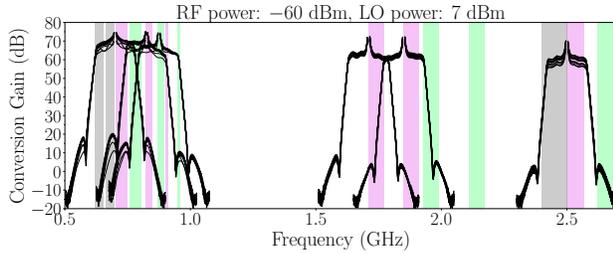


Fig. 4. Conversion gain of the 16 receiver chains. The purple bands correspond to the up-links, the green bands to the down-links, and the gray bands to the digital dividend ( $\approx 600$ – $700$  MHz) and WiFi ( $\approx 2400$ – $2500$  MHz).

mixer, it will not be possible to differentiate which source comes from which band, and therefore, the DoAs cannot be estimated correctly. If the LO were placed at 2.57 GHz to avoid WiFi, we would still have the same problem since now the desired UL would be in the LSB, but the DL would be in the USB. This demonstrates the importance of being able to flag channels to avoid interference (2.4 GHz WiFi Band or base station DL in the case of 2.5 GHz).

In the third part (3), a 1-to-16 splitter distributes a signal that sweeps from 0 to 70 MHz. This signal is used to calibrate ADC phase imbalances digitally.

Note that the order of the components in the receiver is relevant. For example, placing the anti-aliasing filter before the IF amplifier helps eliminate leakage into the IF and intermodulation products, preventing them from saturating the IF amplifier. It is also worth mentioning that the complete receiver was placed inside a Faraday cage since, due to the large gain of the system (60–70 dB), interference may be picked up by the PCBs.

## B. Digital system

The digital system, implemented in the ROACH-2 platform, consists mainly of 16 HMCAD1511 ADCs and the Virtex-6 XC6VSX475T FPGA. The signals are

digitized to 8 bits by the ADCs operating at 140 MSPS, which allows processing signals with a bandwidth of 70 MHz (in accordance with the anti-aliasing filter). Then, the FPGA performs most of the calculations required in the algorithm described in Section IV. Specifically, the digital system implemented in the FPGA comprises three fundamental parts: the PFB-FFT stage ( $\mathbf{W}_H$  from Section IV), the centrosymmetric matrix multiplication stage ( $\mathbf{Q}_P$  and  $\mathbf{Q}_Q$ ), and the correlation matrix accumulation and computation stage ( $\mathbb{R}\{\mathbf{R}_{zz}\}$ ).

For the PFB-FFT implementation, we considered a 128-sized FFT and a Hamming Windowing function in the PFB [24], [25], [26], [27], [28]. For the implementation of the centrosymmetric operations, notice that the matrices  $\mathbf{Q}_P$  and  $\mathbf{Q}_Q$  consist mainly of ones and zeros (scaled by  $1/\sqrt{2}$ ). Hence, we emphasize that there is no necessity to implement complicated matrix multiplications in the FPGA. Finally, for the computation of the correlation matrix, note that from (15) we can write

$$\begin{aligned} \mathbb{R}\{\mathbf{R}_{zz}\} &= \mathbb{R}\{\hat{\mathbf{z}}\hat{\mathbf{z}}^H\} = \mathbb{R}\left\{\sum_{k=0}^{N-1} \hat{\mathbf{z}}(k)\hat{\mathbf{z}}(k)^H\right\} \\ &= \sum_{k=0}^{N-1} \mathbb{R}\{\hat{\mathbf{z}}(k)\}\mathbb{R}\{\hat{\mathbf{z}}(k)\}^T + \sum_{k=0}^{N-1} \mathbb{I}\{\hat{\mathbf{z}}(k)\}\mathbb{I}\{\hat{\mathbf{z}}(k)\}^T. \end{aligned} \quad (28)$$

Hence, there is no need to form a large matrix  $\hat{\mathbf{z}}$  and then perform a computationally expensive matrix product  $\hat{\mathbf{z}}\hat{\mathbf{z}}^H$ . Instead, we only need to accumulate individual multiplications between the real and imaginary parts of the FFT channels generated at each clock cycle. Also, we would like to emphasize that using just one spectrum in the accumulation is not mandatory. If there are  $S$  spectra in a time frame shorter than the minimum length of the phone signals (i.e., 1 ms referring to Fig. 2), the accumulation can be done from  $k = 0$  up to  $k = NS - 1$ .

## VI. Software Design

### A. Task Partitioning: FPGA and Computer Work Division

The DoA computation pipeline is a two-stage system consisting of an FPGA and a computer. RadioVision's 16 ADCs capture data continuously at 140 MSPS, giving a 2.24 GB/s data rate that needs to be processed without any data loss. This hard real-time constraint is satisfied using an FPGA for the ADCs calibrations, channel flagging, and correlation matrix computation. To increase the signal-to-noise ratio further, successive correlation matrices are averaged in the FPGA to obtain a final single  $16 \times 16$  symmetric matrix with a reduced data rate. The half of this final matrix is sent to a computer (since it is symmetric), where the linear algebra steps of the algorithm take place.

### B. Approach to the Multiple Processes of the Computer Program

Given that RadioVision is a real-time system, the software must continuously handle multiple operations in parallel, such as reading data from the FPGA, performing the required linear algebra computations, and simultaneously generating and displaying the augmented reality image. To achieve this concurrency level and ensure a smooth, continuous display of the augmented reality overlay, the RadioVision software leverages Python's multiprocessing library.

#### 1. Data Reception and DoA calculation

The software code begins with a loop that requests the correlation matrix from the FPGA, which is transmitted to the computer using a TCP/IP protocol. To efficiently manage this data flow, a non-blocking queue with a capacity of 20 matrices is initialized using the multiprocessing library. The non-blocking nature of the queue ensures that the process does not freeze when it reaches its capacity. Therefore, if the queue becomes full and a new request is made, the other processes continue to operate without interruption. However, any new data will not be entered into the queue and will be lost.

#### 2. First Eigendecomposition

In parallel, we pass the received correlation matrices to a separate process dedicated to performing the linear algebra operations outlined in Section IV. Specifically, given that  $\mathbb{R}\{\mathbf{R}_{zz}\}$  represents the numpy array of  $16 \times 16$  elements, the first step in this process is to compute the eigendecomposition of  $\mathbb{R}\{\mathbf{R}_{zz}\}$ , which is carried out using the Python linear algebra library. Subsequently, we must select the  $r$  eigenvectors corresponding to the  $r$  largest eigenvalues, where  $r$  represents the number of signal sources.

#### 3. Source Number Criteria

For the source number estimation, we tried to implement the methods documented in [11], [12], [13], [14],

but none worked stably. For testing with continuous wave controlled sources, the inability to estimate the source number is not a problem since we can force the value of  $r$  into the algorithm. However, when treating sporadic sources such as cell phones, it is necessary to address the problem and correctly estimate the number of sources. By implementing the following strategy, we could only differentiate when there is or there is not a source.

To estimate the case  $r = 0$  reliably, we should note that all eigenvalues of the correlation matrix are close to 0 when there are no signals [4], [5]. Therefore, if we sort the 16 eigenvalues as  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{16}$ , it is expected that  $\lambda_1$  should be not much greater than  $\lambda_{16}$ , and consequently we can define a threshold  $u_1$ , such that if  $\lambda_1/\lambda_{16}$  (dB)  $\geq u_1$ , we can assume that  $r \neq 0$ .

For the case  $r \geq 2$ , we can do something similar, noting that in the presence of more than one signal, the rest of the eigenvalues that follow  $\lambda_1$  begin to grow in value. Therefore, we can define a second threshold  $u_2$ , such that if  $\lambda_1/\lambda_2 \geq u_2$ , we can assume that  $r < 2$ . In conclusion, we can reliably estimate both when  $r = 0$  and  $r = 1$ . For  $r = 0$ , we must ensure that the  $\lambda_1/\lambda_{16} \geq u_1$  criterion is not met, while for  $r = 1$  both the criteria  $\lambda_1/\lambda_{16} \geq u_1$  and  $\lambda_1/\lambda_2 \geq u_2$  must met.

Finally, when we know the number of sources in advance (such as in controlled laboratory conditions), we can set  $r$ . We can further evaluate  $\lambda_1/\lambda_{16}$ ,  $\lambda_2/\lambda_{16}$ , ...,  $\lambda_r/\lambda_{16}$  to estimate the confidence in detecting each of the sources.

#### 4. Solution of the system of equations and Second Eigendecomposition

Once we have identified the number  $r$ , we pass the first  $r$  eigenvectors and eigenvalues as  $16 \times 1$  and  $16 \times 1$  numpy arrays  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$  and  $\mathbf{L} = [\lambda_1 \lambda_2 \dots \lambda_r]$  to another parallel process. In this process, we form the  $16 \times r$  numpy array  $\mathbf{V}_r = [\mathbf{v}_1 \mathbf{v}_2 \dots \mathbf{v}_r]$  and we solve the system of equations (25) by applying the pseudo inverse available in the linear algebra library as

$$\begin{aligned} \mathbf{X} &= (\mathbf{K}_{\mu 1} \mathbf{V}_r)^+ \mathbf{K}_{\mu 2} \mathbf{V}_r \\ \mathbf{Y} &= (\mathbf{K}_{\nu 1} \mathbf{V}_r)^+ \mathbf{K}_{\nu 2} \mathbf{V}_r, \end{aligned} \quad (29)$$

where  $+$  denotes the pseudo inverse. After obtaining  $\mathbf{X}$  and  $\mathbf{Y}$ , we perform in this same process the eigendecomposition of  $\mathbf{X} + j\mathbf{Y}$  (again via the linear algebra library) and then obtain the estimated DoAs from its eigenvalues, as stated in Section IV.

#### 5. Moving Average

In the case of having a single source with multiple detections (for example, a cellphone on call), we can use post-processing techniques to improve the result of the estimates. Indeed, measurements with high values of  $\lambda_1/\lambda_{16}$  can be considered reliable, and measurements with low values of  $\lambda_1/\lambda_{16}$  can be considered unreliable due to the proximity to noise. With this, a moving average can be formed to obtain a single DoA from a cloud of

DoAs. Specifically, we can average  $N$  DoAs to form the moving average

$$y(n) = \frac{1}{W(n)} \sum_{k=n-N+1}^n x(k)w(k) \quad (30)$$

$$W(n) = \sum_{k=n-N+1}^n w(k),$$

where  $\{x(k)\}$  is the sequence of DoAs calculated from  $N$  consecutive correlation matrices, and  $\{w(k)\}$  are the weights, given by  $\lambda_1(k)/\lambda_{16}(k)$ . If any of the criteria  $\lambda_1(k)/\lambda_{16}(k) \geq u_1$  or  $\lambda_1(k)/\lambda_2(k) \geq u_2$  are not met, the value of  $w(k)$  is considered as 0, so it does not contribute to the average.

Note that if many consecutive values of  $w(k)$  are null, the value of  $W(n)$  will also be null. More generally, if multiple values of  $w(k)$  are unreliable, with  $\lambda_1(k) \approx \lambda_{16}(k)$  and low SNR, the value of  $W(n)$  will be low and equally unreliable. Therefore, similar to the criteria  $u_1$  and  $u_2$ , we define a threshold  $u_3$ , such that if  $W(n) > u_3$ , the moving average is considered reliable. Otherwise, we discard the value. In this work, we have set  $u_3 = 0$  for simplicity, but other values could be explored in the future.

Finally, we can calculate separate moving averages for the DoAs in the directions of  $x$  and  $y$ . Specifically, denoting the DoAs as  $\theta_x(k)$  and  $\theta_y(k)$ , both weighted by  $w(k)$ , given by  $\lambda_1(k)/\lambda_{16}(k)$ , we can define two moving averages as

$$\Theta_x(n) = \frac{1}{W_x(n)} \sum_{k=n-N_x+1}^n \theta_x(k)w(k) \quad (31)$$

$$W_x(n) = \sum_{k=n-N_x+1}^n w(k),$$

and

$$\Theta_y(n) = \frac{1}{W_y(n)} \sum_{k=n-N_y+1}^n \theta_y(k)w(k) \quad (32)$$

$$W_y(n) = \sum_{k=n-N_y+1}^n w(k),$$

where  $N_x$  and  $N_y$  represent the number of samples used for the  $x$  and  $y$  direction estimations, respectively.

Experimentally, we observed that at least when RadioVision is close to the ground ( $\approx 1$  m), the values in the  $y$  direction tend to have a higher standard deviation, which we attribute to ground reflections. For this reason, in this paper, we used different sample sizes for the moving averages, with  $N_x = 10$  and  $N_y = 30$ .

## 6. Image Formation

For image formation, we execute a parallel process that utilizes Python's OpenCV library to capture frames from the optical camera and overlay the calculated DoA

values as markers. These values can represent (1) the instantaneous DoAs computed from the matrix received from the FPGA, (2) the results passed through the moving average filter, or (3) a combination of both. This approach provides flexibility in visualizing the raw data or the smoothed results, ensuring that the final augmented reality image accurately reflects the system's real-time DoA estimations.

## 7. Band Change

Finally, the code includes an option to scan all frequency bands listed in Table I at intervals of  $x$  seconds, where  $x$  is a user-programmable variable. This functionality is implemented using the PyVISA Python library, which facilitates communication with external equipment responsible for generating the LO tone. The equipment is connected to the main RadioVision computer via Ethernet, allowing the system to automatically cycle through the different frequency bands at the defined intervals.

## VII. Field Measurement Tests

To validate our system, we conducted several experiments related to source localization. Specifically, experiments were performed with continuous-wave (CW) sources (i.e., antennas connected to laboratory sources) and cellphones in call. Additionally, the experiments were carried out in various environmental conditions. In the following subsections, we present both the setup and the results of each test.

### A. Tests with two synthetic CW sources (T1)

For this experiment, we conducted our tests at the Astronomy Department in the Universidad de Chile, where our laboratory is located. It is important to note that numerous base stations are present in this area, so we encountered interference from the nearby base stations and individuals' phones within the building.

RadioVision was positioned at a height of approximately 1 m, and two antennas were arranged 10 meters apart, as seen in Fig. 5, which shows the optical image captured by RadioVision's camera. Sources 1 and 2 were fed with sine waves of 1870 and 1855 MHz, respectively, both with powers of  $-5$  dBm, while the LO was set at 1840 MHz. Sources 1 and 2 were located at  $(-8.8^\circ, -6.4^\circ)$  and  $(22.6^\circ, -4.7^\circ)$ , respectively. For DoA estimation,  $r = 2$  was manually forced into the algorithm (i.e., the number of sources was not estimated but rather given as known). The experiment lasted 50 seconds, and the results associated with the DoA estimation are shown in Fig. 6. As can be seen, the values obtained are stable over time for both  $x$  and  $y$  axes, with low standard deviations and root-mean-square errors (RSME), as shown in Table II, which demonstrates the accuracy of the algorithm. Note that although the net error in the  $x$  axis is larger in source 2, this is because the angle is farther from 0 for that source ( $22.6 > 8.8$ ).

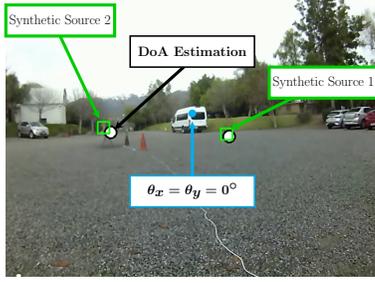


Fig. 5. Source placement in the astronomy department.

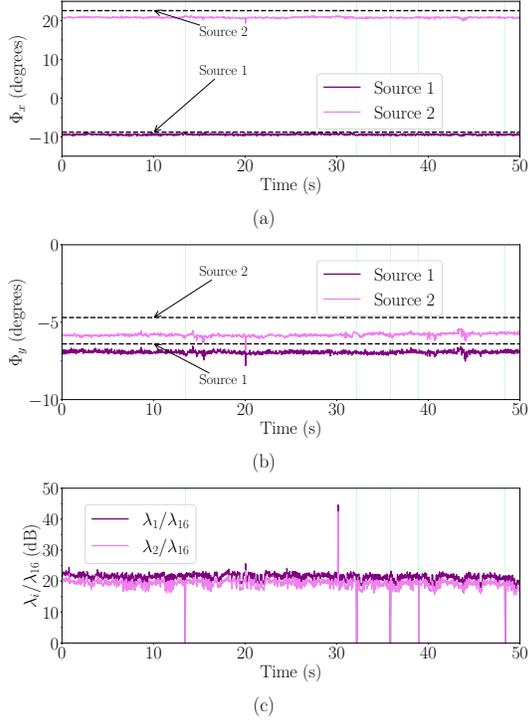


Fig. 6. DoA estimation of two sources located at  $(-8.8^\circ, -6.4^\circ)$  and  $(22.6^\circ, -4.7^\circ)$ . The experiment lasted 50 seconds, in which 1550 accumulation matrices were collected for DoA estimation. (a) DoAs in the horizontal axis. (b) DoAs in the vertical axis. (c)  $\lambda_1/\lambda_{16}$  and  $\lambda_2/\lambda_{16}$  criteria.

Concerning the  $y$  axis, it can be seen that there is a greater error in source 2. The errors in this case are attributed to the proximity to the floor, which contributes to multipath reflections, especially due to the proximity between sources and receivers (10 meters in this case). A test that helps better to discriminate the error on the  $y$  axis is shown in Section 2.

Finally, note that in this experiment, we plot the raw (or instantaneous) values returned by the FPGA in each accumulation cycle of the correlation matrix. This demonstrates the effectiveness of our system under favorable conditions (synthetic sources at a short distance of 10 meters). For completeness, panel (c) of Fig. 6 shows the  $\lambda_1/\lambda_{16}$  and  $\lambda_2/\lambda_{16}$  criteria mentioned in Section VI. As can be seen, there are moments for which  $\frac{\lambda_1}{\lambda_{16}} = 0$  and/or  $\frac{\lambda_2}{\lambda_{16}} = 0$ , where there is no detection and therefore the estimated DoA value is not reliable and is omitted.

TABLE II

Source-Receiver Distance ( $D$ ), mean ( $\mu$ ), standard deviation ( $\sigma$ ), and RMSE values for the experiments shown in Figures 5, 7 and 11.

Parameter	$D$ (m)	Real DoA (deg)	$\mu$ (deg)	$\sigma$ (deg)	RMSE (deg)
T1, Source 1, $x$	10	-8.8	-9.4	0.56	0.81
T1, Source 1, $y$	10	-6.4	-6.9	0.41	0.66
T1, Source 2, $x$	10	22.6	20.8	1.19	2.19
T1, Source 2, $y$	10	-4.7	-5.8	0.34	1.15
T2, $x$	80–100	-	-	-	1.11
T2, $y$	80–100	-	-	-	1.39
T3, $x$	180	-7.6	-5.2	2.37	3.34
T3, $y$	180	2.2	3.7	1	1.79
T4, $x$	10–60	-	-	-	5.42
T4, $y$	10–60	-	-	-	1.94

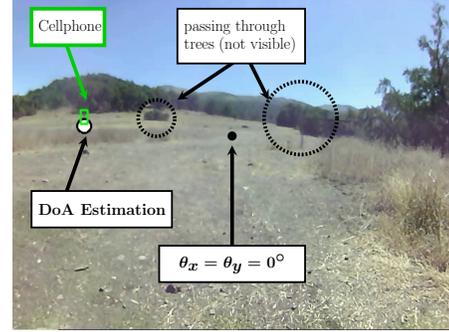


Fig. 7. Source placement in Santa Martina ranch.

## B. Tests with cellphones (T2 - T4)

For these experiments, we conducted tests in two different locations. The first location is the Santa Martina ranch, situated on a hill far from the city, with the nearest base station approximately 1000 meters away. The second location is the engineering campus of the Universidad Chile, which, in addition to its proximity to base stations, experiences significant contamination from other mobile phones. For these tests, we forced  $r = 1$  and used the source number criteria mentioned in Section VI.

### 1. Santa Martina ranch (T2 and T3)

In these experiments, RadioVision was positioned at a height of approximately 1 m, and a person making a phone call moved across the field, as shown in Fig. 7. Using a cellphone application, the call was forced to the 2 PCS band (at 1850–1910 MHz), and therefore the LO was set to 1840 MHz.

In this location, we carried out two tests. In the first (T2), the person making the telephone call moved across the field at an 80–100 m distance for 175 seconds. We used computer vision techniques to calculate the real position of the person (and therefore the source). The results of this test are shown in Figs. 8 and 9. Figs. 8(a) and (b) show the values of  $\lambda_1/\lambda_{16}$  and  $\lambda_1/\lambda_2$  for the source number criteria, with  $u_1 = 18.5$  dB and  $u_2 = 4.5$  dB. Figs. 8(c) and (d) show the estimated DoAs in the  $x$  and  $y$  axes. In light color, we show the instant value, i.e., the value calculated for each output of the correlation matrix, while in dark color we show the moving average discussed in Section VI. In the

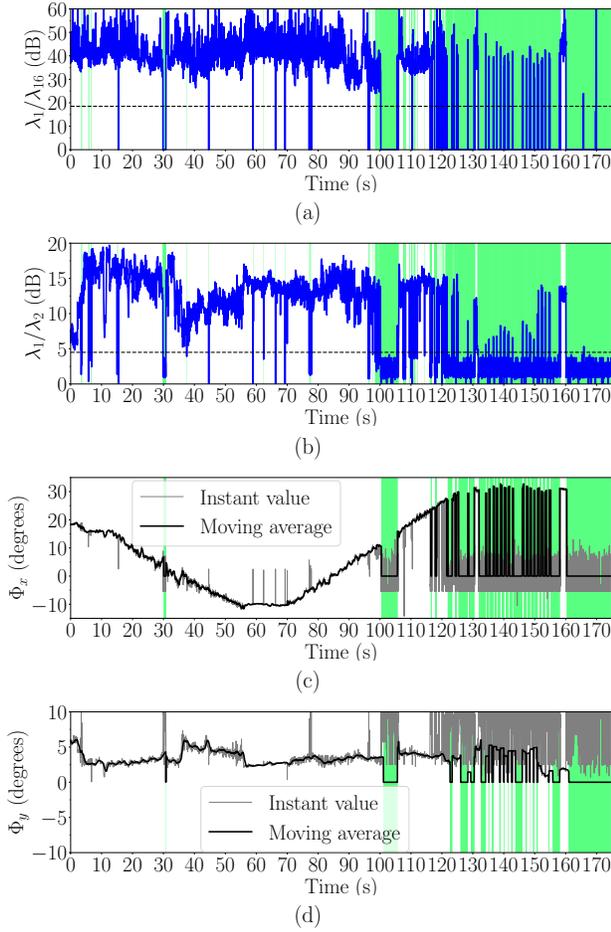


Fig. 8. Instantaneous values for the moving test in the Santa Martina ranch. The experiment lasted 175 seconds, in which 5395 accumulation matrices were collected for DoA estimation. (a)  $\lambda_1/\lambda_{16}$  criteria. The threshold  $u_1$  is shown with a dashed line. (b)  $\lambda_2/\lambda_{16}$  criteria. The threshold  $u_2$  is shown with a dashed line. (c) DoAs in the horizontal axis. (d) DoAs in the vertical axis.

green regions, the estimated DoAs are not valid. Finally, Fig. 9 shows the comparison between the estimated DoA (purple markers) and the real position captured by the camera (orange markers). Note that part of the route includes passing behind bushes. At those moments, the real position could not be captured by the optical camera. Results are very good, with errors below 5 degrees.

In the second test (T3), the person making the phone call remained still at  $(-7.6^\circ, 2.2^\circ)$  at a distance of approximately 180 m for 23 seconds. In this case, there was no direct line of sight between RadioVision and the person. The second test's results are shown in Fig. 10, where we used  $u_1 = 21$  dB and  $u_2 = 4$  dB. We increased the value of  $u_1$  since this threshold helps us discriminate whether there are sources ( $r > 0$ ) or not.

The RMSEs of both tests are shown in Table II. An increment in error can be seen due to the smaller SNR, but the error remains below 5 degrees.

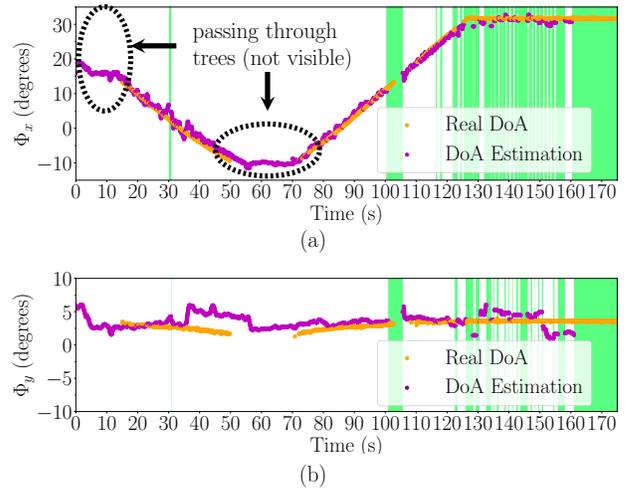


Fig. 9. Averaged values of the first test performed in the Santa Martina ranch. The experiment lasted 175 seconds, in which 5395 accumulation matrices were collected for DoA estimation. (a) DoAs in the horizontal axis. (b) DoAs in the vertical axis. The green bands correspond to the regions where  $W(n) > u_3$  is not fulfilled.

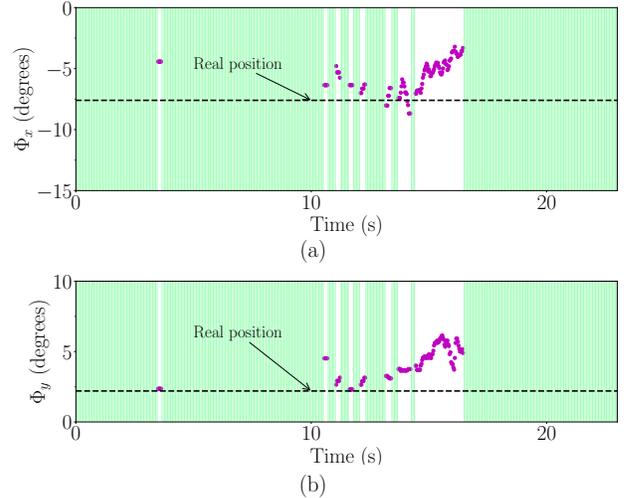


Fig. 10. Averaged values of the second test performed in the Santa Martina ranch. The experiment lasted 23 seconds, in which 730 accumulation matrices were collected for DoA estimation. (a) DoAs in the horizontal axis. (b) DoAs in the vertical axis. The green bands correspond to the regions where  $W(n) > u_3$  is not fulfilled.

## 2. Engineering Campus (T4)

We carried out the last tests on the engineering campus. This place is not only close to base stations but also very busy; therefore, it presents a lot of interference from other phones. In this case, RadioVision was positioned at a height of approximately 5 m, and a person making a phone call moved across the field of view, as shown in Fig. 11. We could not use computer vision techniques to calculate the actual position of the source since there were many other people, and much of the route was not visible in the optical image. However, we recovered some control positions after reviewing the video. For this experiment, we used the cellphone application to force

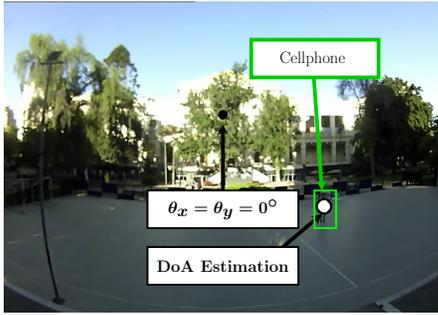


Fig. 11. Source placement in the engineering campus.

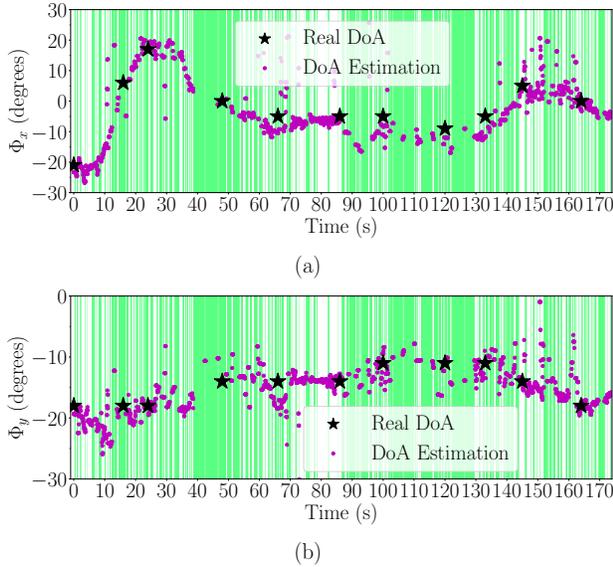


Fig. 12. Averaged values of the experiment performed in the Engineering campus. The experiment lasted 174 seconds, in which 5356 accumulation matrices were collected for DoA estimation. (a) DoAs in the horizontal axis. (b) DoAs in the vertical axis. The green bands correspond to the regions where  $W(n) > u_3$  is not fulfilled.

the call to the 10 AWS band (at 1710—1770 MHz), and therefore, the LO was set to 1700 MHz. The person making the telephone call moved across the field at a 10–60 m distance for 174 seconds. The results are shown in Fig. 12. As can be seen, in this case, there are more outliers in the scatter plot. This is due to phone emissions from multiple people on campus. For this reason, we maintained a similar value of  $u_1$  compared to the previous tests;  $u_1 = 18$  dB, but we increased  $u_2$  up to 6 dB, because this threshold helps us discriminate whether there is one source or more than one source. To give an estimated RMSE value, we considered all the values adjacent to each actual position shown in Fig. 12 within a 2-second window, obtaining RMSEs of  $5.42^\circ$  and  $1.94^\circ$  for the horizontal and vertical axes.

## VIII. Conclusions

We have developed a 2D receiver array operating in the 700–2600 MHz range for real-time Direction of Arrival (DoA) estimation. The implementation of the U-

ESPRIT 2D algorithm in the frequency domain, combined with a robust analog and digital system, has proven effective in mitigating common challenges associated with DoA estimation, such as interference from multiple sources and the calibration of phase and amplitude imbalances.

Experimental results in various environments, including urban and rural settings, validate the system’s ability to estimate the DoA of radio frequency sources accurately. The system achieved a mean 2D squared error of  $1.78^\circ$  at 90 meters and  $3.79^\circ$  at 180 meters in rural environments, while  $5.76^\circ$  in more complex urban environments. These results underscore the system’s potential for practical applications requiring precise localization of RF sources in two angular dimensions. **Moreover, this work represents the first on-field performance report of modern DoA algorithms, such as the 2D U-ESPRIT, conducted in realistic, non-laboratory environments. The evaluation explores performance in scenarios with obstacles, interference from multiple phones, and non-synthetic sources like mobile devices.**

RadioVision’s ability to process real-time data and overlay DoA markers onto optical images provides an intuitive interface for its operational use and future developments in DoA, radio spectrum awareness, and related applications.

## REFERENCES

- [1] L. C. Godara, “Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations”, *Proc. IEEE*, vol. 85, no. 8, pp. 1195–1245, 1997, doi: 10.1109/5.622504.
- [2] H. Krim and M. Viberg, “Two decades of array signal processing research: The parametric approach”, *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, 1996, doi: 10.1109/79.526899.
- [3] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, “An overview of signal processing techniques for millimeter wave MIMO systems”, *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, 2016, doi: 10.1109/JSTSP.2016.2523924.
- [4] R. Schmidt, “Multiple emitter location and signal parameter estimation”, *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, 1986, doi: 10.1109/TAP.1986.1143830.
- [5] R. Roy and T. Kailath, “ESPRIT-estimation of signal parameters via rotational invariance techniques”, *IEEE Trans. on Acoust. Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, 1989, doi: 10.1109/29.32276.
- [6] M. Haardt and J. A. Nosssek, “Unitary ESPRIT: how to obtain increased estimation accuracy with a reduced computational burden”, *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1232–1242, 1995, doi: 10.1109/78.382406.
- [7] M.D. Zoltowski, M. Haardt and C.P. Mathews, “Closed-form 2-D angle estimation with rectangular arrays in element space or beamspace via unitary ESPRIT”, *IEEE Trans. Signal Process.*, vol. 44, no. 2, pp. 316–328, 1996, doi: 10.1109/78.485927.
- [8] W. L. Stutzman and G. A. Thiele, “Antenna Theory and Design”, Third Edition, USA, John Wiley & Sons, 2013.
- [9] C. Balanis, “Antenna Theory: Analysis and Design”, Fourth Edition, USA, John Wiley & Sons, 2016.
- [10] S. P. Applebaum, “Adaptative Arrays”, *IEEE Trans. Antennas Propag.*, vol. 24, no. 5, pp. 585–598, 1976, doi: 10.1109/TAP.1976.1141417.
- [11] W. Chen, K. M. Wong and J. P. Reilly, “Detection of the number of signals: a predicted eigen-threshold approach”, *IEEE*

- Trans. Signal Process.*, vol. 39, no. 5, pp. 1088-1098, 1991, doi: 10.1109/78.80959.
- [12] Guanghan Xu, R. H. Roy and T. Kailath, "Detection of number of sources via exploitation of centro-symmetry property", *IEEE Trans. Signal Process.*, vol. 42, no. 1, pp. 102-112, 1994, doi: 10.1109/78.258125.
- [13] L. Fulai, Z. Xiyuan, L. Chun and W. Jinkuan, "A Robust Algorithm for Source Number Detection and 2-D DOA Estimation Based on Real-valued Computation", *2006 8th International Conference on Signal Processing*, Guilin, China, 2006, doi: 10.1109/ICOSP.2006.344535.
- [14] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 2, pp. 387-392, 1985, doi: 10.1109/TASSP.1985.1164557.
- [15] J. Foutz, A. Spanias and M. Banavar, "Narrowband Direction of Arrival Estimation for Antenna Arrays Synthesis", Morgan & Claypool, 2008.
- [16] A. A. Hussain, N. Tayem, M. O. Butt, A. -H. Soliman, A. Alhamed and S. Alshebeili, "FPGA Hardware Implementation of DOA Estimation Algorithm Employing LU Decomposition", *IEEE Access*, vol. 6, pp. 17666-17680, 2018, doi: 10.1109/ACCESS.2018.2820122.
- [17] A. A. Hussain, N. Tayem, A. -H. Soliman and R. M. Radaydeh, "FPGA-Based Hardware Implementation of Computationally Efficient Multi-Source DOA Estimation Algorithms", *IEEE Access*, vol. 7, pp. 88845-88858, 2019, doi: 10.1109/ACCESS.2019.2926335.
- [18] A. D. Redondo, T. Sanchez, C. Gomez, L. Betancur and R. C. Hincapie, "MIMO SDR-based implementation of AoA algorithms for Radio Direction Finding in spectrum sensing activities", *IEEE COLCOM 2015*, pp. 1-4, 2015, doi: 10.1109/ColCom-Con.2015.7152096.
- [19] B. Rares, *et al.*, "Experimental Evaluation of AoA Algorithms using NI USRP Software Defined Radios", *RoEduNet*, pp. 1-6, 2018, doi: 10.1109/ROEDUNET.2018.8514133.
- [20] C. Balanis and P. I. Ioannides, "Introduction to Smart Antennas", Morgan & Claypool, 2007.
- [21] M. Shao, P. Zhang and Z. Zhang, "Realization of digital down conversion in pulse radar receiver", *2016 IEEE MTT-S International Microwave Workshop Series on Advanced Materials and Processes for RF and THz Applications (IMWS-AMP)*, Chengdu, 2016, pp. 1-4, doi: 10.1109/IMWS-AMP.2016.7588435.
- [22] X. Zhou, Z. Xie, F. Wen and J. Yang, "An Optimal Design and Implementation Method of High Speed Digital Down Conversion Based on FPGA", *2021 International Conference on Computer Information Science and Artificial Intelligence (CISAI)*, Kunming, China, 2021, pp. 1104-1107, doi: 10.1109/CISAI54367.2021.00221.
- [23] D. Gallardo, D. Monasterio, R. Finger, F. P. Mena and L. Bronfman, "A Compact Metamaterial-Based Antenna for Multiband Phased Array Applications", *IEEE Trans. Antennas Propag.*, vol. 69, no. 12, pp. 8872-8877, 2021, doi: 10.1109/TAP.2021.3090861.
- [24] F. Curotto, *et al.*, "Digital calibration test results for Atacama Large Millimeter/submillimeter Array band 7+8 sideband separating receiver", *Journal of Astronomical Telescopes, Instruments, and Systems*, vol. 8, no. 2, 2022, doi: 10.1117/1.JATIS.8.2.024004.
- [25] R. Finger, F. Curotto, R. Fuentes, R. Duan, L. Bronfman, and D. Li, "A FPGA-based Fast Converging Digital Adaptive Filter for Real-time RFI Mitigation on Ground Based Radio Telescopes", *Publications of the Astronomical Society of the Pacific*, vol. 130, no. 984, doi: 10.1088/1538-3873/aa972f.
- [26] D. Monasterio, S. Jorquera, F. Curotto, C. Espinoza, R. Finger, and L. Bronfman, "A Proof of Concept Balanced Mixer with the use of a Digital IF Power Combiner to Improve LO Noise Rejection", *Publications of the Astronomical Society of the Pacific*, vol. 135, no. 1054, doi: 10.1088/1538-3873/ad0789.
- [27] A. Oppenheim and R. Schaffer, "Discrete-Time Signal Processing", Third Edition, Pearson, 2010.
- [28] R. Crochiere and L. Rabiner "Multirate Digital Signal Processing", Prentice Hall, 1983.