



Ultra-massive MIMO for future cell-free heterogeneous networks - **MiFuture**

Positioning and sensing algorithms - initial version

Deliverable 2.3

Grant Agreement Number: 101119643

Project Acronym: MiFuture

Project Title: ultra-massive MIMO for future cell-free heterogeneous networks

Call: HORIZON-MSCA-2022-DN-01

Type of action: HORIZON TMA MSCA Doctoral Networks - Industrial Doctorates

Granting authority: European Research Executive Agency

Project start date: 01/01/2024

Dissemination Level		
PU	Public	x
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Editor (name/partner):	Simona Lohan / TAU
Authors (name/partner):	Andres Reyes / UC3M Ariele Sapienza / TAU Fatih Ayten / TAU Mohd Adnan / UA Omid Aghda / UNL Prabhat Gupta / UC Radovan Juran/ TAU Ruben de Miguel Gil/ UC3M Ertug Pihitili / TAU Mohammad Abu Aqoulah / UAB
Internally reviewed by (name/partner):	Simona Lohan/TAU

Document history

Version	Date	Author	Notes
1.0	27.09.2025	Simona Lohan	Table of content created
1.1.1	02.10.2025	Ariele Sapienza	Radio SLAM with cm-level accuracy
1.1.2	16.10.2025	Omid Aghda	Centimetre positioning and channel estimation
1.1.3	03.11.2025	Fatih Ayten	Section 3.2 completed
1.1.4	05.11.2025	Ariele Sapienza	Section 4.1 completed
1.1.5	07.11.2025	Ertug Pihitili	Completing sensing-related aspects
1.1.6	8.11.2025	Ruben de Miguel Gil	additions in Sections 3-4
1.1.6	11.11.2025	Prabhat Gupta, Mohd Adnan	additions in Sections 2-5
1.2	13.11.2025	Simona Lohan	integration and first revisions
1.2.1	20.11.2025	Mohammad Abu Aqoulah	adding Section 2.2
1.2.2	21.11.2025	Simona Lohan	full revision and submitted to the project manager
1.3	tba	Eduardo Alonso	Final revision and portal check

Short Abstract: This deliverable includes an initial version of the algorithms to improve the sensing and positioning up to cm-level accuracy. It addresses the following aspects: an overview of radio-based positioning, sensing and joint positioning and sensing (i.e., the so called radio-based Simultaneous Localization and Mapping (SLAM)), as well as initial results based on radio (OFDM-based) sensing, radio positioning with various measurements (time, angle, carrier phases) and radio SLAM with both fixed and mobile landmarks and in both non-cooperative and cooperative manner.

Contents

1	Introduction	2
1.1	Recent works on radio sensing with cm-level accuracy	3
1.2	Recent works on radio positioning with cm-level accuracy	4
1.3	Recent works on joint radio positioning and sensing / simultaneous localization and mapping (SLAM) with cm-level accuracy	7
2	Sensing-related algorithms developped in MiFuture	9
2.1	Channel estimation and tracking	9
2.2	Sensing bounds and algorithms for perceptive cell-free networks with hybrid intelligent surfaces	10
2.2.1	Estimation Algorithms	11
2.2.2	Sensing Bounds	12
2.2.3	Role within Hybrid Intelligent Surfaces	13
2.3	Clutter Suppression in Bistatic millimeter-wave (mmWave) Integrated Sensing and Communications (ISAC) Systems	13
2.4	ISAC for Resource Management	16
3	Positioning-related algorithms developped in MiFuture	19
3.1	Machine Learning for Carrier Phase Positioning	19
3.2	3D Indoor Positioning-Aided Beamforming for mmWave Distributed MIMO Systems	22
4	Radio SLAM-related algorithms developped in MiFuture	26
4.1	Data Association Problem	26
4.2	SLAM performance in non-collaborative approach	28
4.3	Collaborative SLAM	28
5	Conclusions and future research	33
5.1	Main findings	33
5.2	Open Challenges and Future Research Directions	34
	List of Acronyms	35
	References	37

Chapter 1

Introduction

As wireless communication systems evolve toward 6G and beyond, positioning and sensing are emerging as core functionalities to complement the traditional data transmission core functionalities. These capabilities are essential for enabling new applications, such as intelligent and autonomous systems, smart cities, enhanced network services, context-aware services, improved security in wireless communications, etc.

Wireless positioning refers to the ability to determine the location of devices or users within target accuracy metrics, while **wireless sensing** involves extracting some environmental information (e.g., object detection and/or localization, motion tracking) using radio signals. The integration of these two aspects is known in the literature as Joint Positioning and Sensing (JPAS) or, as a more widespread notation, as **radio simultaneous localization and mapping (SLAM)**. The radio SLAM concept leverages the radio infrastructure to simultaneously localize users and to 'map' the surrounding environment (e.g., detect and locate fixed landmarks or moving objects in the surroundings of the target devices).

Some of the key performance indicators for positioning and sensing in future networks include, but are not limited to:

- Accuracy: Sub-meter or centimeter-level precision, especially in urban and indoor environments.
- Latency: Real-time or near-real-time updates for dynamic applications.
- Reliability: Robustness against multipath, interference, and non-line-of-sight conditions.
- Coverage: Ubiquitous availability across diverse environments.
- Energy Efficiency: Low power consumption for battery-constrained devices.

The focus of this deliverable is mostly on the accuracy metric, with targets of reaching cm-level accuracy or below.

Research in Orthogonal frequency division multiplexing (OFDM)-based radio sensing, wireless positioning, and radio SLAM is driven by the increasing ubiquity of wireless infrastructure and the limitations of traditional sensing technologies. Unlike cameras or Light Detection and Ranging (LIDAR), radio signals can penetrate walls and operate effectively in Non-line-of-sight (NLOS) conditions, making them ideal for environments where visual sensing fails. OFDM, being the backbone of modern cellular standards such as 4G, 5G, and 6G, offers fine-grained multipath resolution, which is crucial for accurate sensing and localization.

The impact of this research spans multiple domains. For example, in smart environments, it enables ambient intelligence through activity recognition and fall detection. In robotics and autonomous vehicles, radio SLAM provides robust navigation in Global navigation satellite system (GNSS)-denied or visually challenging settings. Industrial applications benefit from structural monitoring and intrusion detection, while emergency services can use radio-based positioning to locate individuals in disaster zones where conventional methods fail. In industrial settings, radio-based sensing, positioning, and radio SLAM technologies may offer

transformative capabilities for automation, safety, and asset tracking. These systems can enable real-time monitoring of machinery, detect anomalies in structural integrity, and support autonomous navigation of robots and drones in complex environments such as warehouses or manufacturing plants. Unlike vision-based systems, radio sensing can operate reliably in dusty, dark, or cluttered conditions, making it ideal for harsh industrial environments. Additionally, integrating these capabilities into existing wireless infrastructure reduces deployment costs and simplifies maintenance. Last but not least, in aviation, the potential impact of this research is equally significant. Radio SLAM and positioning can enhance aircraft navigation in g-denied environments such as hangars, tunnels, or during taxiing on runways with limited satellite visibility. Onboard radio sensing systems can also support collision avoidance, monitor cabin occupancy, and assist in autonomous ground operations. Furthermore, radio-based technologies can be used for structural health monitoring of aircraft components, detecting stress or damage through changes in signal reflections. These applications contribute to safer, more efficient, and intelligent aviation systems.

1.1 Recent works on radio sensing with cm-level accuracy

OFDM is the de facto ISAC baseline because its multi-carrier structure reuses communication resources for range-Doppler processing with well researched and understood performance. In OFDM for sensing, the fundamental range resolution is set by the signal bandwidth, $\Delta R \approx \frac{c}{2B}$, e.g. a 100 MHz 5G New Radio (NR) signal can enable ≈ 1.5 m [1]. Likewise, prototyping experiments have been demonstrated for OFDM-based NR waveforms for sensing with and high resolution processing. For example, [2], implemented an “integrated super-resolution sensing and communication” (ISSAC) testbed using 5G NR OFDM waveforms and validated delay/Doppler estimation capabilities in experiments. However, as is widely known, OFDM suffers major degradation in high mobility scenarios.

Orthogonal time frequency space (OTFS) has emerged as a strong ISAC alternative when high mobility and large Doppler spreads are present. By modulating the symbols in the delay-Doppler (DD) domain, OTFS aligns sensing parameters (delay/range, Doppler/velocity) with the resources grid, enabling sparse DD estimation and robust processing under fractional Doppler and the channel’s multipath components appear as separable impulses, where the delay component corresponds to the range and the Doppler shift corresponds to the velocity of targets or obstacles. Additionally, this structure allows the receiver to interpret channel effects in terms of the sparse aforementioned impulses or taps, rather than seeing a rapidly time-varying frequency-domain channel per subcarrier. Current research shows why OTFS better tolerates Doppler than OFDM while keeping the same bandwidth-limited range resolution [3], and below centimeter sensing accuracy if the system uses the terahertz band [4].

Therefore, in radar sensing mode, OFDM’s Doppler sensitivity likewise translates to degraded velocity estimation and cross-term interference. Meanwhile, because both OFDM and OTFS use the same total bandwidth, the delay (for range estimation) resolution is bandwidth-limited in exactly the same way: $\Delta R \approx \frac{c}{2B}$. Thus OTFS does not sacrifice range resolution, yet gains Doppler robustness, especially in high mobility channels.

Achieving centimeter-level sensing accuracy is a major challenge in wireless localization, especially when relying on traditional methods like received signal strength indicator (RSSI). As shown in [5], RSSI-based fingerprinting lacks the resolution needed for fine-grained positioning, making it unsuitable for applications that demand high precision. To overcome these limitations, recent researches has explored more detailed signal features, particularly channel state information (CSI), which captures both amplitude and phase across subcarriers and antenna pairs. CSI enables richer multipath profiling and more stable measurements, offering a pathway toward more accurate localization [6]. That is why in the following we highlights several state-of-the-art approaches that leverage CSI for sensing, alongside other advanced signal processing techniques.

A key aspect of this research area is the dual use of communication signals for both channel estimation and sensing. In 5G NR systems, reference signals such as Demodulation reference signal (DMRS), channel

state information reference signal (CSI-RS), and sounding reference signal (SRS), originally designed for channel estimation, are now being repurposed for sensing tasks. These signals are structured within OFDM resource grids and offer valuable spatial and temporal information. Their design and allocation directly affect the performance of sensing algorithms, especially in scenarios where precise localization is required [7].

While OFDM has been widely adopted in communication systems, it struggles in doubly dispersive channels, those affected by both time and frequency selectivity, especially in high-mobility environments. This is where OTFS modulation offers a significant advantage. OTFS operates in the Delay-Doppler (DD) domain, which naturally aligns with physical parameters like range and velocity. Because the channel in this domain tends to be sparse, it becomes easier to isolate and estimate individual paths. This sparsity, combined with the direct mapping to physical motion, makes OTFS particularly well-suited for positioning tasks [7]. Recent work [8] has shown that estimating fractional delay and Doppler values, rather than relying on quantized taps, can lead to more accurate and robust channel estimation. This precision directly supports high-resolution localization, especially in dynamic scenarios where traditional OFDM-based methods fall short.

Intelligent reconfigurable surface (IRS)s (also called Reconfigurable Intelligent Surface (RIS)) also offer new opportunities. A twin-IRS setup, as proposed in [9], uses spatial diversity and tensor modelling to extract multipath parameters with high precision. This approach has demonstrated centimetre-level localization accuracy, showing how advanced channel modelling can directly support positioning goals. Machine Learning (ML) is another area gaining traction. Recurrent neural network (RNN)-based models that incorporate CSI, user movement, and signal-to-noise ratio (SNR) have outperformed traditional fingerprinting methods, pushing indoor localization closer to the centimetre mark [5].

1.2 Recent works on radio positioning with cm-level accuracy

With cell-free networks and a user-centric approach to network division, optimizing resources becomes crucial for enhancing scalability, improving efficiency, and ensuring Quality of Service (QoS) for each user. This introduces challenges such as different cooperation levels required among Access Point (AP), and algorithms designed to optimize multiple objectives, particularly the User Equipment (UE)'s Spectral Efficiency (SE) [10, 11]. Leveraging user location can significantly improve resource allocation by dynamically adjusting resource distribution. Knowing UE's location helps to optimize the beamforming weights by forming narrow beams for each user, steering them appropriately, and assigning power coefficients based on the channel information [12], thereby maximizing the efficient use of limited spectrum. With location information, we make sure that APs with better channel conditions are activated to serve the user and not just the APs that are closer to the user. Knowing the user clusters helps to optimize the resources and ensures that a single AP does not get overloaded, the buffer size does not increase, and each user receives a fair share of resources. This also helps to reduce power and allocate it in an adaptive manner instead of allocating the same power to each user. The other main plus point of having the location estimate is that the APs can be pre-selected, and a smooth handover procedure can be implemented while never compromising the QoS.

For better positioning accuracy and optimized resource allocation, AP selection becomes crucial, catering for the geometry and reducing front-haul complexity. According to the user distribution, the contribution from some APs can lead to higher user interference, decreasing the Signal to Interference and Noise Ratio (SINR). The authors in [13] propose a method of AP selection along with signal detection capable of enhancing UE's throughput, thus reducing the amount of data load for the fronthaul link, where they propose a solution that considers channel quality and effective channel gain and sequentially allocates the APs to the users. In [14], the authors show that the AP selection scheme effectively provides downlink SE results while granting load balance with large power reductions. In [15], the authors analyse the impact of different UE velocities on downlink SE and how it affects performance. They emphasize the importance of pilot

training up to a certain velocity and address the effects of channel ageing caused by mobility. To mitigate channel ageing, precise user tracking is necessary. User tracking helps implement the different techniques of distributed beamforming. The most common effective beamforming approach is the Minimum mean square error (MMSE) beamforming, which offers a balance between Maximum Ratio Transmission (MRT) and Zero Forcing (ZF) as it behaves like MRT at low Signal to Noise Ratio (SNR) and approaches ZF at high SNR [16].

If we consider the downlink power optimization problem jointly linked with beamforming and a desired SINR to be achieved, the location estimate helps to design better precoders for distributed beamforming, aiding in better power and resource allocation using different optimization approaches for a non-convex problem [17, 16]. [18] shows how different positioning errors affect the beam direction for different numbers of transmitted resource blocks. Higher frequency gives a better location estimate, thus a more directed beam which reduces power and increases the throughput of the network. The amount of accuracy required for the position estimate depends on the application it is being implemented for:

- Autonomous Driving ($< 10\text{ cm}$): In an urban environment with a high density of objects and no Line of Sight (LOS), assuming a high number of connected sensors and overabundance of information an accuracy of $< 10\text{ cm}$ is required for 6G based 3D localization [19].
- Industrial Robotics ($10\text{--}30\text{ cm}$) For industrial automation, 3GPP release 16 and 17 have set a standard of less than 30cm of accuracy with a standard Time of Arrival (TOA) approach [20].
- Augmented Reality (AR) and Teleoperation ($1\text{--}10\text{ cm}$)- Teleoperation is one of the crucial applications for 6G and beyond. This helps patients to choose doctors without being location-bound, and more importantly, it provides medical facilities to remote areas. An accuracy of $1\text{--}10\text{ cm}$ is required to successfully implement this application [21].

Carrier phase positioning (CPP) is one of the key enablers for achieving the cm-level localization accuracy demanded by emerging 6G applications [22, 23]. Originally developed within the GNSS community, CPP exploits the phase of the received carrier signal.. This capability makes CPP a promising approach for terrestrial positioning, where high-frequency operation, dense access point deployment, and precise synchronization can further enhance phase resolution and enable sub-centimeter accuracy. As positioning accuracy becomes a critical enabler for communication-localization co-design, CPP stands out as a vital technique to bridge the gap between satellite-based and terrestrial localization systems.

Despite these opportunities, the practical implementation of CPP in cellular networks introduces several challenges. The most critical one is the integer ambiguity problem, which arises because the observed carrier phase is indistinguishable from any integer multiple of the signal wavelength [24]. This ambiguity leads to multiple possible range solutions and necessitates specialized algorithms for ambiguity resolution. Moreover, CPP performance strongly depends on precise synchronization among transmitters and receivers, as even minor phase misalignments or hardware impairments can severely degrade positioning accuracy. Multipath propagation and NLOS conditions further complicate phase interpretation, while network scalability and fronthaul latency pose additional implementation difficulties in distributed architectures. To address these issues, a range of approaches have been investigated, including differential phase measurements, hybrid time-phase estimation, multipath-assisted ambiguity resolution, and recent deep learning-based solutions [25, 26].

Most existing CPP studies, such as [27, 28, 29], have considered hybrid localization schemes that fuse carrier phase measurements with other positioning observables like TOA or angle of arrival (AoA). For instance, [27] introduces CPP-enhanced time/angle-based positioning schemes, while [28] proposes a novel CPP framework for 5G New Radio (NR) networks, where UE positions estimated from time-difference-of-arrival (TDoA) are combined with phase-derived position changes over time. Similarly, [29] investigates the CPP problem in a cellular LOS scenario and establishes a fundamental performance bound that explicitly

accounts for integer ambiguity. These hybrid schemes demonstrate that CPP can significantly enhance positioning accuracy when jointly exploited with complementary time- or angle-based measurements.

Recently, CPP has emerged as a distinct research direction, aiming to achieve precise localization using solely carrier phase observations. Unlike hybrid methods that treat phase as auxiliary information, this paradigm fully leverages the fine-grained phase data as the sole input for estimating the UE position. Such an approach offers opportunities for low-latency and low-overhead localization, as it eliminates the need for synchronization-intensive time-based observables. In distributed and phase-coherent Multiple-input multiple-output (MIMO) systems, carrier phase localization has been shown to offer promising accuracy under ideal synchronization conditions [26, 30]. Specifically, [26] introduces a carrier phase only UE positioning framework employing two Deep learning (DL)-based solutions under perfect hardware assumptions, while [30] proposes a deep learning-driven localization method leveraging the hyperbola intersection principle, achieving improved robustness against antenna element failures and outperforming conventional approaches. These studies highlight the potential of CPP to serve as a core building block for high precision, scalable, and learning-driven localization in 6G networks.

Accurate user positioning is also essential for reliable communication and beamforming. However, traditional positioning methods continue to face challenges when dealing with wideband and multi-subcarrier signals that are typical of modern wireless networks. In dense urban and outdoor microcell environments, the effects of multipath propagation and shadowing significantly degrade localization performance [31]. To overcome these issues, recent research has focused on applying machine learning techniques to Distributed MIMO (D-MIMO) architectures, enabling data driven position estimation based on received signal features. The work [32] proposed a Gaussian Process Regression (GPR) framework that maps the uplink Received Signal Strength (RSS) data to user positions while mitigating shadowing effects through probabilistic modeling. This approach has shown notable improvements in localization accuracy compared to conventional deterministic methods. To further enhance computational efficiency, the authors [33] introduced a two-phase clustering and regression strategy, where fingerprint data are first grouped using affinity propagation, and a localized GPR model is trained within each cluster. This method maintains high positioning accuracy while reducing training and inference complexity. The authors [34] developed a delay-Doppler domain particle filter that performs direct positioning using the CSI of OFDM signals, eliminating the need for strict inter-base-station synchronization. The work [35] applied an Extended Kalman filter (EKF) to sequentially estimate user position and velocity from combined TOA and Angle-of-Arrival (AoA) measurements in a D-MIMO environment, demonstrating robust tracking capability and reduced computational complexity compared to particle-filter-based methods.

With the evolution toward 6G networks, the demand for ultra-high precision positioning has become even more critical. 6G aims to support emerging applications such as autonomous driving, drone coordination, and immersive extended reality, all of which require cm-level accuracy and ultra-low latency localization. However, achieving this level of precision introduces new challenges due to severe path loss, blockage, and the high sensitivity of mmWave and sub-terahertz frequency bands to NLOS propagation. To address these limitations, researchers have started to combine geometric signal processing with data-driven inference to improve positioning robustness and beam alignment. The authors in [36] proposed an EFIM-guided hybrid beamforming method that jointly optimizes the beam pattern based on the Equivalent Fisher Information Matrix, thereby improving both positioning accuracy and communication reliability. The work [37] investigated RIS-assisted angular beamforming, where controllable reflections are used to reconstruct blocked LOS paths, enhancing localization in dense environments. As 6G networks push toward higher frequencies and greater localization precision, mmWave communication has become a key enabling technology. Operating in the 30-300 GHz range, mmWave provides ultra-high data rates, low latency, and fine-grained spatial resolution, which are essential for both high-throughput communication and cm-level positioning. Its capability to form narrow, directional beams allows the extraction of rich spatial information, making it highly effective under challenging NLOS conditions. To further enhance system performance, D-MIMO architec-

tures have been introduced, where spatially separated AP collaborate through a centralized processing unit. This distributed design improves spatial diversity and coverage, significantly benefiting localization accuracy and link robustness. Recent studies have demonstrated promising progress in this direction. In [38], they proposed a beam prediction and adaptation hybrid beamforming method that leverages user position information to predict optimal beams, thereby reducing search overhead and improving spectral efficiency. The authors in [39] developed a fast in-band beam tracking algorithm that jointly estimates the user's channel and position to dynamically refine beam directions during mobility, ensuring stable connectivity with minimal latency. In [40], they introduced an EKF-based beam tracking framework that continuously updates angular and range states using real-time channel estimates, enabling fast and stable beam alignment under high mobility conditions in mmWave networks. The authors in [41], have explored joint channel position estimation frameworks that integrate localization and beamforming into a single optimization process, achieving sub-meter accuracy in dynamic environments. The integration of mmWave communication with D-MIMO architectures has not only improved positioning accuracy but also opened the path for advanced beam management strategies that exploit location information directly within the beamforming process.

The integration of these high-precision localization techniques with intelligent beam control has led to the emergence of Positioning-Aided Beamforming (PA-BF) a unified framework that aligns communication and sensing objectives within the same spatial domain. In this paradigm, user location and motion information are not treated as auxiliary data but as core parameters that actively influence beam selection, alignment, and tracking. By leveraging accurate positioning estimates, the network can predict optimal beam directions and dynamically adjust them to follow user mobility, thereby reducing beam-training overhead and improving reliability under NLOS and high mobility conditions. This joint design transforms conventional beamforming into a location-aware process that enhances spectral efficiency, minimizes latency, and maintains stable connectivity even in dense urban or fast changing environments. In D-MIMO and mmWave architectures, PA-BF serves as the key enabler that converts precise localization into actionable beam management intelligence laying the foundation for adaptive, context-aware, and energy-efficient wireless communication envisioned in 6G systems.

1.3 Recent works on joint radio positioning and sensing / simultaneous localization and mapping (SLAM) with cm-level accuracy

Radio SLAM builds upon the SLAM foundations in robotics [42], where the workhorse algorithm for nearly two decades has relied on an EKF to jointly estimate the unknown state of the robot and environmental landmarks. However, the EKF-SLAM algorithm has several shortcomings when using monostatic measurements and other sensing modalities such as Lidar and camera, typical in robotic SLAM:

1. Quadratic complexity in the number of landmarks [43] making real-time operation challenging for large maps with potentially thousands of landmarks;
2. Very optimistic uncertainty estimates which can cause the EKF filter to diverge [44];
3. High sensitivity to data association (DA) errors (i.e., the incorrect matching of measurements and estimated landmarks) which can lead to catastrophic failure of the algorithm [43].

Due to the aforementioned drawbacks, alternative approaches to solve the problem have been developed and nowadays one of the most widely used algorithms is the GraphSLAM approach [45, 46]. However the work [47] is combining EKF-SLAM algorithm with dynamic covariance scaling (DCS) to improve the capability in radio-SLAM to handle DA errors.

The early radio SLAM algorithms have utilized the EKF for inference [48, 49], but due to the notoriously challenging DA problem and the inability of the EKF to handle DA uncertainty, also the radio SLAM

research field has evolved past the EKF algorithm. To this end, the state-of-the-art radio SLAM algorithms can account for DA uncertainty and the problem is typically solved either by using belief propagation (BP) on factor graphs [50] or Random finite set (RFS) theory combined with Monte Carlo methods [51].

Usually, when dealing with SLAM with dynamic environment, the majority of work are filtering out the moving objects. In this way the target is identified and classical SLAM algorithm can be applied [52]. The dynamic environment is categorized according to how much the environment is moving. In [53, 54] the algorithm used is the Dynamic Pose Graph (DPG) SLAM that allows to manage environment with low dynamic object. The technique employed integrate information about the moving objects in the scene without the necessity of actually estimating the pose of the objects or relying on the 3D model. In [54] the goal is to identify the stationary portion of the environment, for future uses combined with up to date dynamic object. The technique used is the DPG. It uses the prior information that some of the structure points belong to moving objects and these points can be associated in subsequent frames. There are works like [55, 56] that track the dynamic objects as separately using traditional multi-target tracking.

The work [57] is the first example of the integration of Detection and Tracking of Moving Objects (DATMO) and SLAM to addresses the critical challenge in autonomous navigation through dynamic environments, such as urban areas populated with vehicles, pedestrians, and other moving agents. Traditional SLAM algorithms typically assume a static world model. SLAM with DATMO approach extends the probabilistic framework of SLAM by incorporating the detection and tracking of dynamic elements in the environment. This is achieved by separating sensor measurements into static and dynamic components, allowing the system to simultaneously build a map of static landmarks while identifying and predicting the behavior of moving objects. A motion-based detector is employed to recognize dynamic elements, which are then tracked using Multiple Hypothesis Tracking and modeled through a constant-velocity assumption or more complex motion models when needed. Technologically, the system is implemented using laser scanners for environment perception and odometry for motion estimation. The overall estimation process is handled through an EKF. This separation of SLAM and DATMO posteriors ensures computational efficiency and improves robustness by avoiding contamination of the static map with dynamic features. The combined SLAM with DATMO framework demonstrated significant improvements in both mapping fidelity and object tracking reliability compared to traditional SLAM systems, making it highly suitable for applications such as autonomous driving and intelligent driver-assistance systems. There is a drawback for complex motion model where is difficult to decompose the measurements coming from stationary and moving objects. This require to analyze the data over a sufficient period of time using complex feature extraction algorithms. SLAM with DATMO decomposes the estimation problem into two separate estimators. By maintaining separate posteriors for stationary objects and moving objects, the resulting estimation problems are much lower dimensional than SLAM with generalized objects.

Another technique proposed in the literature is the simultaneous localization and mapping and Multi-Object Tracking (SLAMMOT) [58]. Conventional Multi-Object tracking (MOT) methods rely on a known and accurate ego-pose, which is not always available or reliable in unknown environments. To address these challenges, SLAMMOT systems have been developed to jointly estimate the state of the ego-vehicle and surrounding moving objects in a tightly-coupled framework. One of the central challenges in SLAMMOT is the handling of dynamic objects that may become occluded, missed, or incorrectly associated across frames. To overcome this, SLAMMOT method introduces a confidence-guided implicit data association strategy. There is a dynamic associations integrated into a joint graph optimization backend that simultaneously updates the states of the ego-vehicle and the tracked objects. By incorporating dynamic object information as a resource rather than interference, SLAMMOT not only enhances robustness in dynamic scenes but also improves overall pose estimation accuracy.

Chapter 2

Sensing-related algorithms developped in MiFuture

2.1 Channel estimation and tracking

Channel estimation plays a pivotal role in the performance of OTFS modulation, particularly in high-mobility scenarios. A key advantage of OTFS lies in its representation of the wireless channel in the DD domain, where each channel tap is associated with a specific delay and Doppler shift. This representation not only facilitates robust communication under doubly dispersive conditions but also provides a physically meaningful interpretation of the propagation environment. Each DD tap corresponds to a distinct propagation path, with the delay indicating the distance to a reflecting object and the Doppler shift revealing its relative velocity. This inherent structure makes OTFS channel estimation particularly attractive for ISAC applications, where the same waveform is used for both communication and environmental sensing. By accurately estimating the DD channel, it becomes possible to extract valuable situational awareness information, such as object location and motion, directly from the communication signal [59]. As illustrated in the Figure 2.1, the DD domain channel effectively models the physical environment, capturing both motion and spatial characteristics [59]. Based on this structure, this modulation has great potential to be used in ISAC applications [60, 61]. This dual functionality underscores the potential of OTFS as a foundational technology for future wireless systems that integrate sensing and communication.

Several recent studies have explored channel estimation techniques for fractional delay and Doppler in radar and ISAC applications. In [62], the authors used Time-Frequency (TF) spreading codes, previously applied in Gabor Division Spread Spectrum (GDSS) systems, to approximate the ambiguity function of the GDSS signal as a Two-Dimensional (2D) sinc function. They proposed two algorithms: the first uses 2D sinc interpolation via optimization for higher accuracy, while the second employs quadratic interpolation to provide an approximate closed-form solution that can serve as an initial point for the first algorithm. The approach presented in the paper offers a good trade-off between complexity and performance. However, it is limited to radar scenarios with a single tap and does not address multi-tap channels. Additionally, it lacks comparison with other methods and only evaluates the Root Mean Square Error (RMSE) of the two proposed algorithms. In [63], the authors prove that the Range-Velocity (RV) matrix used in conventional radar sensing is equivalent to the DD channel representation. They further demonstrate an equivalence between the Fast-Time-Slow-Time (FTST) and Delay-Time (DT) matrices, which supports the relevance of OTFS in ISAC systems. Based on this, the authors proposed an approximate closed-form solution for estimating fractional delay and Doppler using pulse compression and 2D correlation between the received and transmitted signals in the DD domain. However, this method is only applicable when no two channel taps share the same integer delay or Doppler. In [64], a 2D correlation-based framework was used to propose the Maximum Likelihood based on Multi-Frame Superposition (ML-MFS) method for estimating

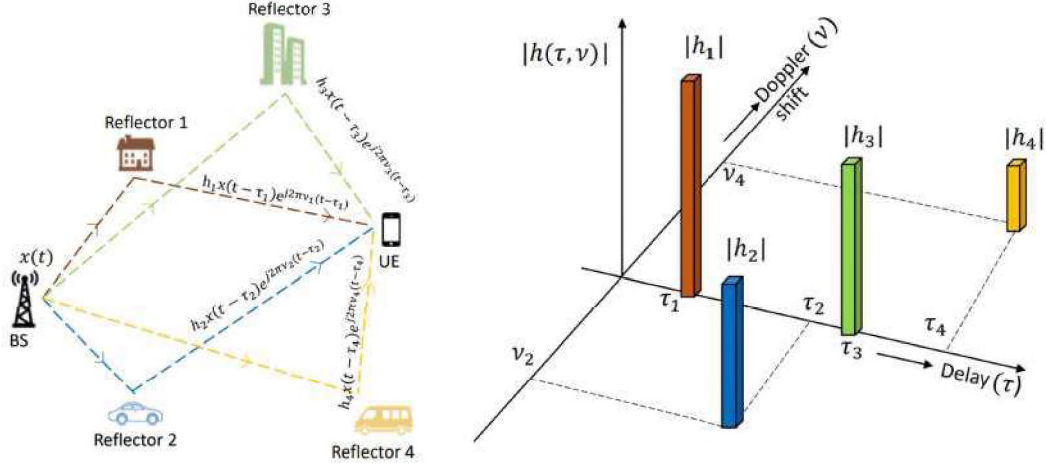


Figure 2.1: Illustration of the DD domain channel in OTFS, showing how each tap corresponds to a physical path with specific delay and Doppler characteristics [59]. You can see that the signal transmitted from the Base Station (BS) travels through four distinct paths to reach the user device. Each path is characterized by its own delay (related to distance), Doppler shift (related to velocity), and gain (related to attenuation). On the right, the channel is represented in the DD domain as multiple coefficients, each located at a specific DD position and associated with a value corresponding to the attenuation of that particular path.

channel parameters, including fractional delay and Doppler, in ISAC applications. This method improves performance with lower complexity compared to [65], which applies ML after 2D correlation. To further enhance performance, the authors introduced Whale Optimization Multi-Frame Superposition (WOA-MFS). However, the method assumes channel parameters remain constant over a super-frame, which may not hold in high-Doppler scenarios. Moreover, its reliance on multiple frames makes it unsuitable for delay-sensitive applications that require rapid symbol detection. Inspired by these works, we aim to develop a delay and Doppler estimation algorithm that not only maintains low complexity but also performs effectively in Line of Sight (LoS) scenarios, making it suitable for ISAC applications and object tracking.

In [8], the authors proposed a low-complexity fractional delay and Doppler channel estimation method for OTFS modulation in both LoS and NLOS scenarios. Their approach is based on a well-known classical frequency estimation technique [66]. They leverage the fact that, in the presence of fractional delay and Doppler, which is the case in realistic simulations as opposed to the idealized integer case, the received signal exhibits characteristics similar to a frequency estimation problem. Utilizing this insight, they apply the 2D-based Aboutanios and Mulgrew (A&M) algorithm to estimate the channel parameters [66]. The proposed method outperforms other well-known off-grid Sparse Bayesian Learning (SBL)-based approaches while maintaining significantly lower computational complexity. In Figure 2.2, a comparison is presented between the proposed method in [8] and other off-grid techniques, including GESBI [67], OGSBI [68], T-GEESBI [69], the conventional Embedded Pilot (EP) method [70], and the Genie Bound calculated in [69]. It is seen from the figure that their proposed method outperforms the other approaches while achieving significantly lower complexity.

2.2 Sensing bounds and algorithms for perceptive cell-free networks with hybrid intelligent surfaces

Future cellular networks are increasingly shifting toward *ultra-massive* MIMO (UM-MIMO) deployments, where the physical aperture of the antenna array becomes comparable to the link distance. In this regime—

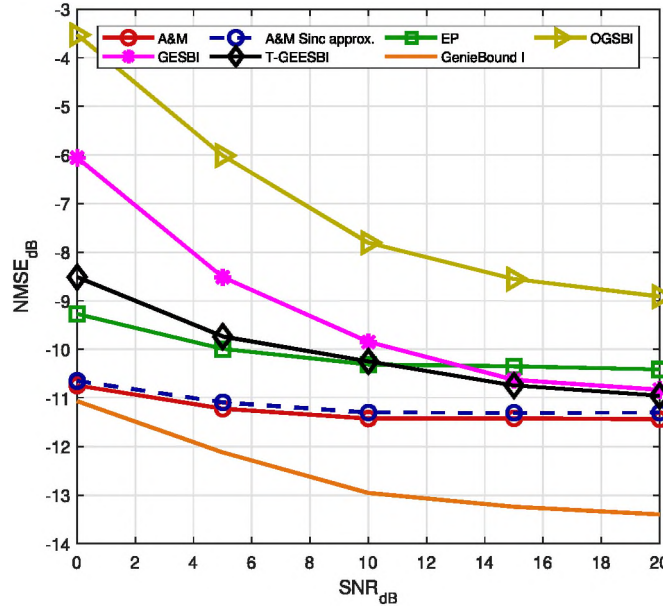


Figure 2.2: NMSE versus SNR for estimated DD, comparing off-grid methods, EP scheme, and Genie Bound I.

commonly referred to as the radiative near field—the classical plane-wave assumption breaks down, and the channel exhibits spherical wavefronts, distance-dependent beam focusing, and spatial non-stationarity across the array elements [71]. These characteristics fundamentally alter the underlying sensing problem and require estimation techniques that jointly recover the users' range and angle.

To understand the sensing capability of UM-MIMO systems in this regime, it is essential to analyze both: (i) algorithmic approaches for near-field parameter estimation, (ii) fundamental accuracy bounds such as the non-parametric and parametric Cramér–Rao bounds (CRBs), and (iii) the role of hybrid-intelligent Surfaces (HIS) in enhancing the robustness and accuracy of these sensing pipelines. This section therefore provides an overview of the near-field estimators considered in our study—with a particular focus on Partial Relaxation (PR) approaches—together with the corresponding CRBs and their operational relevance within HIS-enabled UM-MIMO networks.

2.2.1 Estimation Algorithms

Several high-resolution algorithms have recently been revisited for joint distance–angle estimation in radiative near-field MIMO systems. Among them, the two-dimensional MUSIC (2D-MUSIC) method has received renewed interest due to its ability to exploit the spherical wavefront model and jointly estimate range and azimuth using the noise subspace of the sample covariance matrix. Recent results in [72] demonstrate that properly designed 2D-MUSIC schemes can achieve excellent near-field estimation accuracy and even approach the parametric Cramér–Rao bound (CRB). This confirms 2D-MUSIC as a strong and reliable baseline for near-field sensing.

While MUSIC remains a powerful subspace-based estimator, its performance is sensitive to finite-snapshot conditions, model mismatch, and colored or structured interference. To overcome these limitations, the *Partial Relaxation (PR)* framework has been proposed as an eigenvalue-based alternative that relaxes only the nuisance parameters while keeping the physical steering-vector structure intact. The PR approach was originally introduced in [73] and has been shown to systematically outperform classical MUSIC

and ESPRIT-type methods across a wide range of SNRs and array sizes. Its main advantage stems from reducing multi-source searches into tractable single-source optimizations without compromising estimation resolution.

Motivated by its strong theoretical and empirical performance, we extend PR concepts to radiative near-field estimation. In particular, we examine two representative PR variants. The first is a *covariance-fitting* estimator (PR-CF), inspired by the JADE framework in [74], where user signals are treated as unknown and the estimator fits the received covariance matrix directly. Covariance fitting has demonstrated strong robustness and high accuracy in joint delay–angle estimation tasks, making it a natural candidate for near-field applications. The second is a pilot-assisted maximum-likelihood variant (PR-ML), designed for scenarios where the user pilots are known at the receiver and can be exploited to perform a structured ML search over the near-field parameter space. PR-ML provides a reference estimator under known-signal conditions.

In summary, our estimator set consists of: (i) 2D-MUSIC as a classical high-resolution baseline validated in near-field literature; (ii) PR-CF as a powerful covariance-fitting estimator for unknown transmitted signals; and (iii) PR-ML as a pilot-aided PR estimator that leverages known training sequences for enhanced estimation performance.

2.2.2 Sensing Bounds

To quantify the fundamental limits of near-field sensing and benchmark the performance of the considered estimators, we evaluate a comprehensive set of Cramér–Rao Bounds (CRBs). These bounds differ in the structural assumptions imposed on the channel model and in the amount of prior knowledge available at the receiver.

At one extreme lies the *non-parametric* CRB, which treats the channel vector as an arbitrary deterministic parameter without exploiting any geometric near-field structure. This bound characterizes the performance limit of estimators that operate without a physical propagation model and therefore serves as a useful baseline.

At the other extreme, the *parametric* near-field CRB incorporates the spherical-wave array response through the range–angle parameterization of the channel. This produces a much tighter bound that reflects the information carried by the physical near-field geometry. Depending on the signal assumptions, both deterministic-signal and Gaussian-signal parametric CRBs are considered, following the classical formulations in [75].

Beyond these classical bounds, we also incorporate *Partial-Relaxation (PR)*–specific constrained CRBs (CCRBs). The CCRB is a lower bound derived for estimation problems in which certain nuisance parameters are “relaxed” while the structural array model is preserved. In our study, the PR-CCRB corresponds to the scenario where the transmitted user signals are unknown but the near-field steering-vector structure is maintained. This bound reflects the best achievable performance of PR-type estimators operating under unknown-signal conditions and captures the effect of relaxing part of the signal model while retaining the geometric near-field constraints.

Finally, when training pilots (known signals) are available, we also evaluate the PR-CRB for known pilots, which represents the accuracy limit of our PR-ML estimator under fully known training sequences.

By jointly examining the non-parametric CRB, the classical parametric CRBs (deterministic and Gaussian), the PR-CCRB for unknown signals, and the PR-CRB for known pilots, we obtain a unified view of the theoretical limits associated with each estimator and each information scenario. This enables a comprehensive performance comparison between 2D-MUSIC, PR-CF, and PR-ML across different signal and structural assumptions.

2.2.3 Role within Hybrid Intelligent Surfaces

The estimation algorithms and sensing bounds described above are intended to support a broader vision of *hybrid intelligent surfaces* in future UM-MIMO-based cellular networks. In such systems, communication, sensing and learning are tightly integrated: high-rate data transmission is combined with fine-grained user localization, environment mapping and situation awareness. Accurate near-field sensing is a key enabler for these capabilities, as it allows the network to exploit the full spatial degrees of freedom offered by ultra-massive arrays and to adapt its operation in a user-centric manner.

By systematically comparing subspace-based, covariance-fitting and pilot-assisted PR estimators against a rich set of CRBs, our study highlights the trade-offs between estimator complexity, required prior knowledge and achievable accuracy in realistic near-field conditions. This insight is instrumental for designing hybrid-intelligent procedures that can select or combine estimators depending on the operating regime and service requirements, ultimately contributing to more reliable and efficient sensing in next-generation cellular-free and UM-MIMO deployments.

2.3 Clutter Suppression in Bistatic mmWave ISAC Systems

As wireless networks evolve into *perceptive networks* through the integration of sensing, achieving reliable sensing performance is vital for effectively addressing the promising use cases of ISAC systems. In this regard, clutter poses a significant challenge in ISAC systems, and its suppression is of paramount importance. Here, clutter refers to static objects in the sensing environment that mask target parameters and degrade sensing performance. To mitigate the effects of static clutter, a novel two-dimensional clutter suppression algorithm is presented, and its effectiveness is demonstrated through computer simulations compared with the state-of-the-art methods.

Considering a MIMO-OFDM ISAC system operating at mmWave frequencies, each multipath cluster in mmWave systems corresponds to a physical object; therefore, clutter can be modeled as a multipath cluster [76]. Consequently, since the ISAC transmitter emits an OFDM waveform across K active subcarriers and M_s symbols to perform the sensing task, the overall target and clutter channel can be expressed as $\mathbf{H}k, m = \mathbf{H}^t k, m + \mathbf{H}_{k, m}^c \in \mathbb{C}^{N_R \times N_T}$, where $\mathbf{H}_{k, m}$ denotes the channel corresponding to the k -th subcarrier of the m -th OFDM symbol.

$$\begin{aligned} \mathbf{H}_{k, m} = \mathbf{H}_{k, m}^t + \mathbf{H}_{k, m}^c = & \sqrt{\alpha_{\text{RCS}, t}} e^{-j2\pi k \Delta f \tau_t} e^{j2\pi m T_s f_{D_t}} \times \mathbf{a}_{\text{Rx}}(\theta_r) \mathbf{a}_{\text{Tx}}^H(\theta_t) + \sqrt{\frac{1}{N_{cl}}} \sum_{n=1}^{N_{cl}} \sqrt{\alpha_{\text{RCS}, t} \alpha_{\text{RCS}, n}} \\ & \times e^{-j2\pi k \Delta f \tau_n} e^{j2\pi m T_s f_{D_c}} \mathbf{a}_{\text{Rx}}(\theta_c + \phi_n) \mathbf{a}_{\text{Tx}}^H(\theta_t), \end{aligned} \quad (2.1)$$

where Δf is the subcarrier spacing, T_s is the OFDM symbol duration, θ_t denotes the azimuth AoD, and τ_t and τ_n are the delays associated with the target and each ray in the clutter channel, respectively, with all clutter rays assumed to share the same delay τ_n . The quantities f_{D_t} and f_{D_c} represent the Doppler shifts of the target and clutter, respectively. The angular spread in the clutter channel is modeled by representing the azimuth AoA of the n -th clutter ray as $\phi_n \sim \mathcal{N}(\theta_c, \sigma_{\text{AS}}^2)$, where θ_c is the mean azimuth AoA and σ_{AS} denotes the angular spread [77].

Although the radar cross section (RCS) is a deterministic quantity measured in square meters (m^2), statistical approaches are commonly employed to model the behavior of objects with complex shapes. Considering objects with irregular geometries and larger sizes, such as walls or buildings, these may contain multiple scattering points distributed across their surfaces, all of which can backscatter the incident wave toward the receiver. The RCS of an object containing N_{cl} scattering points can be expressed as [78]

$$\alpha_{\text{RCS}, c} = \left| \sum_{n=1}^{N_{cl}} \sqrt{\alpha_n} e^{j\psi_n} \right|^2, \quad (2.2)$$

where $\alpha_{\text{RCS},c}$ denotes the total RCS of the object, $\sqrt{\alpha_n}e^{j\psi_n}$ represents the complex reflectivity of the n -th scattering point, and $\alpha_{\text{RCS},n} = |\sqrt{\alpha_n}e^{j\psi_n}|^2$ is the RCS of the n -th point on the object. As a result, based on (2.2), the clutter RCS can be modeled using the Swerling-I model [79], where $\alpha_{\text{RCS},n} \sim \mathcal{CN}(0, \alpha_{\text{RCS},c}/N_d)$ represents the RCS of the n -th ray in the clutter channel. Furthermore, since clutter arises from isotropic scattering of the incident signal, the total RCS contribution of the clutter is modeled as $\alpha_{\text{RCS},t}\alpha_{\text{RCS},n}$. The received array steering vector is given by $\mathbf{a}_{\text{Rx}}(\theta_r) = [1, e^{j\pi \sin \theta_r}, \dots, e^{j\pi(N_R-1)\sin \theta_r}]^T \in \mathbb{C}^{N_R \times 1}$, and the corresponding transmit steering vector $\mathbf{a}_{\text{Tx}}(\theta_t) \in \mathbb{C}^{N_T \times 1}$ is defined analogously.

In an OFDM frame, the Doppler shift manifests as phase variations across M_s OFDM symbols, while the delay corresponds to phase shifts across K subcarriers. These relationships allow the Doppler and delay to be represented as steering vectors in the slow- and fast-time domains, respectively [80]. Consequently, the Doppler steering vector is given by

$$\mathbf{d}(f_D) = [1, e^{j2\pi T_s f_D}, \dots, e^{j2\pi T_s (M_s-1)f_D}]^T, \quad (2.3)$$

where $\mathbf{d}(f_D) \in \mathbb{C}^{M_s \times 1}$. Accordingly, the space-time steering vector can be defined as $\Psi(f_D, \theta) = \mathbf{d}(f_D) \otimes \mathbf{a}_{\text{Rx}}(\theta) \in \mathbb{C}^{M_s N_R \times 1}$. Thereafter, the received OFDM signal, after removing the transmit symbols, is expressed as

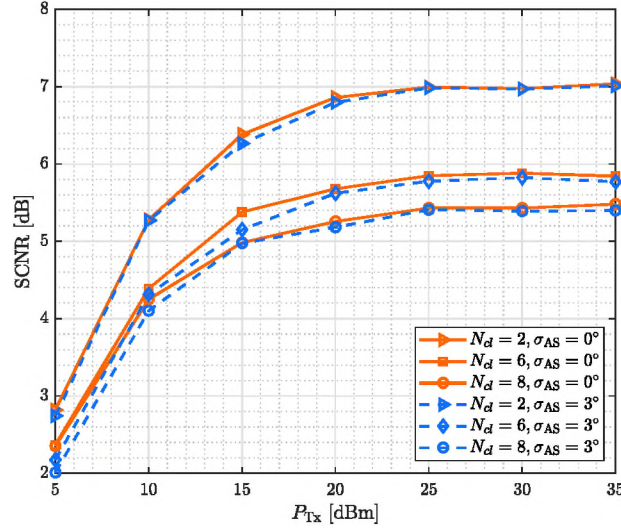
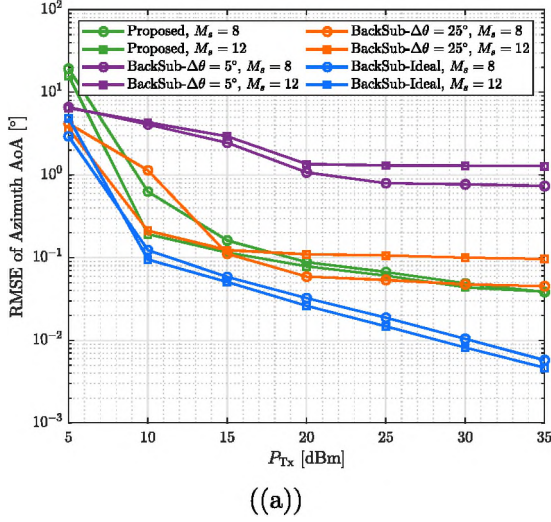
$$\begin{aligned} \hat{\mathbf{y}}_k &= \hat{\mathbf{y}}_k^t + \hat{\mathbf{y}}_k^c \\ &= \left(\Psi(f_D, \theta_r) \mathbf{a}_{\text{Tx}}^H(\theta_t) e^{-j2\pi k \Delta f \tau_t} + \sqrt{\frac{1}{N_d}} \sum_{n=1}^{N_d} \Psi(f_D, \theta_c + \phi_n) \mathbf{a}_{\text{Tx}}^H(\theta_t) e^{-j2\pi k \Delta f \tau_c} \right) \hat{\mathbf{w}}_{\text{Tx}} + \mathbf{z}_k, \end{aligned} \quad (2.4)$$

where $\hat{\mathbf{y}}_k^t$, $\hat{\mathbf{y}}_k^c$, and $\hat{\mathbf{y}}_k \in \mathbb{C}^{M_s N_R \times 1}$ denote the target, clutter, and total received signals for the k -th subcarrier, respectively. The noise term $\mathbf{z}_k \sim \mathcal{CN}(0, \sigma_N^2 \mathbf{I}_{M_s N_R})$ represents the vector of additive white Gaussian noise (AWGN) samples.

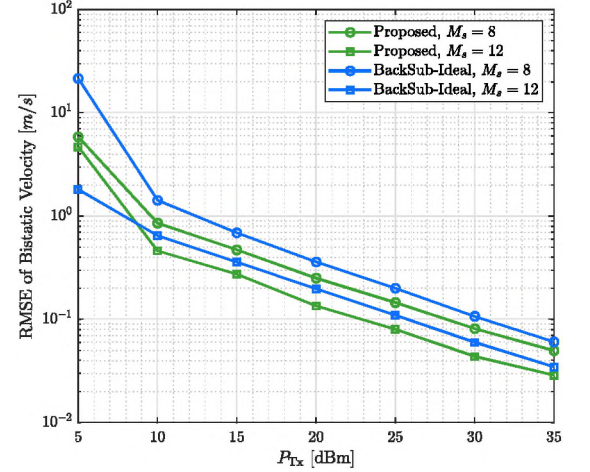
Following the space-time representation in (2.4), the proposed two-dimensional root-multiple signal classification (MUSIC)-based clutter suppression algorithm is employed to enhance sensing parameter estimation. Furthermore, based on the proposed algorithm, a space-time receive filter is designed to improve the signal-to-clutter-plus-noise ratio (SCNR). In the proposed algorithm, the target and clutter components are first jointly detected using the minimum description length (MDL) criterion [81]. The algorithm is then initialized based on the detected objects in the sensing environment, which are assumed to correspond to the sensing target and clutter. Further details regarding the algorithmic implementation and simulation parameters are provided in [82]. Extensive computer simulation results are presented below.

The SCNR results for different N_d and σ_{AS} (angular spread) values are demonstrated for varying P_{Tx} values in Fig. 2.3. The SCNR improves as P_{Tx} increases across all cases and tends to saturate when P_{Tx} exceeds 25 dBm. As the number of rays in the clutter channel increases, the SCNR decreases, indicating that the clutter signal becomes more dominant and interferes more with the target signal. Moreover, for all N_d values, only a slight performance degradation is observed when the angular spread increases from $\sigma_{\text{AS}} = 0^\circ$ to $\sigma_{\text{AS}} = 3^\circ$, validating the robustness of the proposed algorithm against angular spread introduced by the clutter channel.

In Figs. 2.4(a) and 2.4(b), the RMSE performance of the proposed method is compared with the background subtraction (BackSub) benchmark algorithm presented in [83], where 2D-rootMUSIC is employed for target parameter estimation. Ideally, the BackSub algorithm first captures a reference signal $\mathbf{y}_{\text{ref}} = \hat{\mathbf{y}}_k^c + \mathbf{z}_k^{\text{ref}}$ under the simplifying assumption that no actual target is present in the sensing area. When a target appears, the algorithm subtracts $\mathbf{y}_{\text{ref}}$ from the received signal $\hat{\mathbf{y}}_k$ to remove the clutter effect. In contrast, the proposed method estimates target parameters directly from $\hat{\mathbf{y}}_k$, eliminating the need for a reference signal. As an additional more realistic baseline, Fig. 2.4(a) considers also a scenario in which a point-scatterer object sharing the same RCS as the target, characterized by $\{\theta_x = \theta_r + \Delta\theta, v_x = 5 \text{ m/s}\}$, is present during AoA

Figure 2.3: SCNR for different values of N_d and σ_{AS} under varying P_{Tx} .

((a))



((b))

Figure 2.4: (a) Root mean squared error (RMSE) of azimuth AoA estimation for the proposed method and two BackSub algorithm cases: a point-scatter object scenario with an AoA difference of $\Delta\theta$ from the target, and an ideal reference signal acquisition case, with varying P_{Tx} and different M_s values. (b) RMSE of bistatic velocity estimation for the proposed method versus the ideal benchmark case of the BackSub algorithm, with varying P_{Tx} and different M_s values.

reference signal acquisition. This yields $\mathbf{y}_{\text{ref}} = \hat{\mathbf{y}}_k^x + \hat{\mathbf{y}}_k^c + \mathbf{z}_k^{\text{ref}}$. It can be observed that the proposed method consistently outperforms the BackSub algorithm as transmit power P_{Tx} increases for $\Delta\theta = 5^\circ$. Although BackSub performance improves with greater azimuth AoA separation, especially at $M_s = 12$ and $\Delta\theta = 25^\circ$, the proposed method maintains superior performance at higher P_{Tx} values. Interestingly, increasing M_s leads to worse performance in the more realistic evaluation case of the BackSub algorithm, as the increased contribution of clutter hinders effective suppression, while the presence of the point-scatterer object during reference signal acquisition further obscures the target parameters during estimation. In the ideal case, BackSub achieves better AoA estimation accuracy than the proposed method.

As shown in Fig. 2.4(b), the proposed method achieves a lower RMSE in bistatic velocity estimation compared to the baseline, demonstrating robustness to strong background clutter. Better performance is

observed for $M_s = 12$ compared to $M_s = 8$, since increasing the coherent processing interval (CPI) of the OFDM waveform, given by $M_s T_s$, enhances Doppler estimation accuracy. Furthermore, as the azimuth AoA and bistatic velocity are jointly estimated, extending the CPI also improves AoA estimation, as illustrated in Fig. 2.4(a).

2.4 ISAC for Resource Management

For Resource Management ISAC plays a vital role in improving the SE and the fairness in the system. It enables the joint utilization of communication and sensing functionalities within the same spectral and spatial resources. With the help of ISAC we can gain environmental awareness by knowing the user's location and the mobility pattern of the user. This makes the allocation of resources dynamic, allowing the network to adjust the power levels, optimize distributed beamforming by selecting the right set of APs and make the handover process smoother. It can also preallocate the resources as the user movement is predicted, which helps to reduce latency and satisfy the strict scheduler timing constraints while still providing the necessary QoS. In a dense environment with a lot of APs and users, the resources are rare, and with ISAC, the SE is improved a lot, as we do not waste resources to detect and then communicate separately, and it also improves the interference management in these scenarios. The channel can be detected by geometric means without adding to the signalling overhead. It optimizes power and minimizes energy consumption as it adapts to location-aware beamforming.

But with these capabilities also come problems. The main issue is to deal with the positioning error, as a false location can make the beam steer in a different direction or choose the wrong cluster of APs to beamform, which would increase interference and also increase the blocking probability, which would reduce the overall data rate of the network. Figure 2.5 shows a simple illustration of how the wrong APs can be chosen when there is a positioning error. In [84], the authors show that with sensing capabilities, the overhead for beamforming can be reduced by a significant amount, which boosts the data rate of the network. They deploy a grid-based solution to showcase the error in the position estimate and compare the data rates for different accuracy in the positioning errors. There is a tradeoff between the location error and beamforming overhead. Solving this problem is crucial as the overhead would increase with the user's movement. A simple solution is required to compensate for the location error and to also optimize the resources and guarantee the desired throughput, and choose the right cluster of APs.

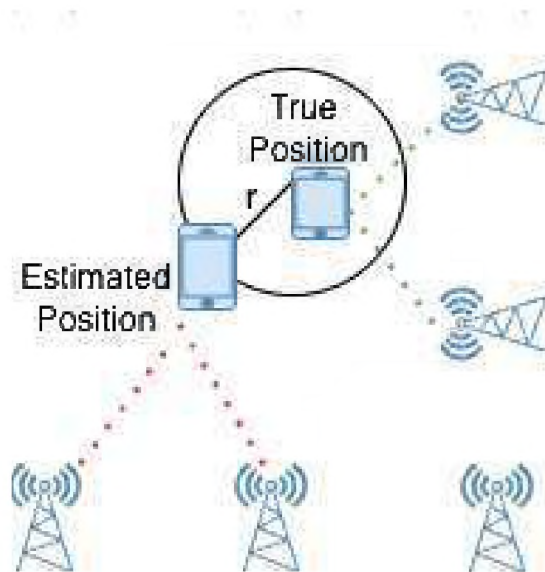


Figure 2.5: Effect of positioning error in AP selection

To compensate for the position error and to allocate the right set of APs for distributed beamforming, we deploy an eigenvalue-based approach. The benefit of having eigen-beamforming is that the precoder can pick the best eigenbeams, which compensate for the error and reduce the overhead significantly. Also, by knowing the user's location and the estimated error covariance, a more dynamic preselection and AP grouping can be carried out, again by utilising the concept of eigenvalues. This makes the SINR equation not only have interference from the other users, but also from the APs which are in the vicinity of the user but are not contributing. If $\hat{H}$ is the channel estimated using the location of the user in the downlink, the eigenvalue decomposition can be given as $\hat{H} = USV^H$, where V^H are the downlink beamforming vectors and S are the diagonal eigenvalues. Now, if the position is acquired using TOA, the ranging accuracy in the ideal case can be given as

$$\sigma_n = \sqrt{E[\Delta d^2]} \geq \frac{c}{B} \frac{1}{2\sqrt{2\pi}\sqrt{SNR}} \quad (2.5)$$

where $\Delta D = c\Delta\tau$, c being the speed of light, and $\Delta\tau$ is the time difference from the actual time taken to the time taken due to noise and interference that results in positioning errors. For different position accuracy requirements, we can calculate the SNR for a certain Bandwidth B . Now, if this is the minimum SNR which is required, the minimum eigenvalue can be calculated with the help of the total power which is transmitted by an AP. This can be given as: $S_{min} = \frac{SNR_{min} * Noise}{Power_{AP}}$. Now we can form AP clusters that have eigenvalues above S_{min} . This would help choose APs which would not affect the throughput of the network even with positioning errors. For M APs such that $m \in \{1, 2, \dots, M\}$ the cluster can be formed as $m \in M_k$ for user k . The set M_k has the APs which could serve and are in the vicinity of the user k , but not all of them are utilized depending on the SINR requirements, thus the other APs belonging to this set are the interference for user k . We denote the APs serving the user by the set S_k . This brings us to a more interference-aware SINR equation for the user k for the network with K users, such that $k \in \{1, 2, \dots, K\}$:

$$SINR_k = \frac{\sum_{m \in S_k} p_{mk} |H_{mk}^H g_{mk}|^2}{\sum_{\substack{m' \in M_k \\ m' \notin S_k}} \sum_{j \neq k} p_{m'j} |H_{m'k}^H g_{m'j}|^2 + \sigma^2} \quad (2.6)$$

Here, p_{mk} is the power transmitted by AP m to user k , H_{mk} is the channel and g_{mk} are the beamforming vectors that are calculated from MMSE beamforming using the eigenvectors V^H and σ^2 is the noise variance. This optimizes power, reduces interference that increases the SE of the network.

A power optimization problem of minimizing the overall network power p_k , which is the sum of powers across all the APs can be given as:

$$\min_{p \in \mathbb{R}_+^K} \sum_{k=1}^K p_k \quad (2.7)$$

$$\text{subject to } SINR_k(p, g) \geq SINR_{desired,k}, \quad \forall k = 1, \dots, K, \quad (2.8)$$

$$p_{min,k} \leq p_k \leq p_{max,k}, \quad \forall k = 1, \dots, K. \quad (2.9)$$

$$\frac{p_k}{M} \leq p_{mk} \quad (2.10)$$

In figure 2.6 power optimization graph is illustrated which shows how choosing the right set of APs can help reduce the overall power of the network compared to when all the APs try to do the beamforming to the user k , during errors in the position estimation.

Further, the average SE of the network can be calculated, and the APs could be preallocated depending on the movement of the user and the fairness of the system, such that they work to increase the overall throughput of the network.

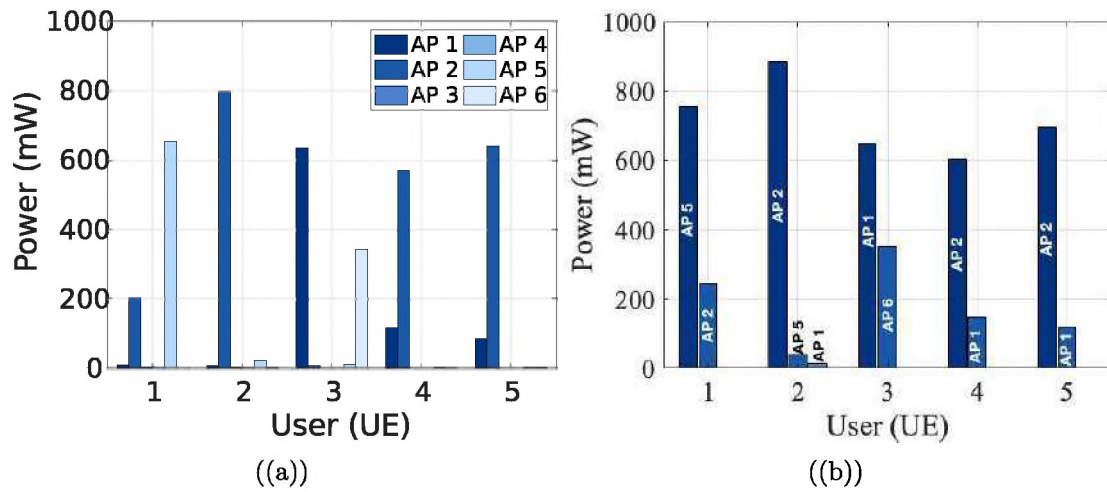


Figure 2.6: Optimized per-AP transmit power: (a) all APs active, (b) optimized AP selection based on dominant AP-UE channel eigenvalues.

Chapter 3

Positioning-related algorithms devellopped in MiFuture

3.1 Machine Learning for Carrier Phase Positioning

The rapid proliferation of data-driven applications and the need for accurate, low-latency services in 5G and beyond have driven the exploration of artificial intelligence (AI) techniques to overcome the limitations of traditional wireless algorithms, particularly in dynamic and complex environments. In localization and positioning, ML enables systems to extract robust features from noisy measurements and adapt to environmental variability, offering superior accuracy and responsiveness compared to conventional model-based methods [85].

As surveyed in detail in [86], AI and ML methods have been increasingly adopted in both range-based and fingerprint-based localization approaches, enhancing accuracy and robustness in dynamic environments. For instance, supervised and unsupervised learning techniques are widely employed to mitigate non-line-of-sight (NLOS) errors in systems based on time-of-arrival (ToA) and ultra-wideband (UWB) ranging [87], while deep learning and clustering algorithms significantly improve fingerprinting-based indoor positioning systems by reducing storage and computation overheads [88]. Moreover, reinforcement learning is applied in trajectory learning and autonomous robot navigation, showcasing the versatility of ML in handling multi-modal data and adapting to environmental variations [89]. Among the various localization techniques and ML-based algorithms, this section focuses on the benefits of incorporating ML into the CPP process.

We consider an uplink scenario consisting of a UE and I distributed and mutually phase-synchronized APs. An example scenario with ten APs is shown in Fig. 3.1.

The UE has an unknown position $\mathbf{x}_{\text{ue}} \in \mathbb{R}^2$ whereas the APs have known positions $\mathbf{x}_{\text{ap},i} \in \mathbb{R}^2$ for $i \in \{0, \dots, I-1\}$, with the scenario being extendable to $\mathbb{R}^3$. When the UE transmits a unit-power narrowband pilot symbol s occupying a bandwidth W with a transmit power P , the received signal at the i -th AP over the LOS link is $y_i = \sqrt{E}\rho_i \exp(-j(\frac{2\pi}{\lambda}d_i - \theta))s + v_i$, where $E = P/W$ is the symbol energy, ρ_i is the path loss from the UE to AP i , $d_i = \|\mathbf{x}_{\text{ue}} - \mathbf{x}_{\text{ap},i}\|$ is the Euclidean distance, θ is the common phase offset between the UE and the AP network, and $v_i \sim \mathcal{CN}(0, N_0)$ represents additive white complex Gaussian noise. By processing y_i , the resulting phase observations can be formulated as

$$r_i = -\frac{2\pi}{\lambda}d_i + \theta + 2\pi z_i + n_i, \quad (3.1)$$

where $z_i \in \mathbb{Z}$ is the integer ambiguity, and we model $n_i \sim \mathcal{N}(0, \sigma_i^2)$. From (3.1), we can take one AP as reference, e.g., $i = 0$, and compute *differential measurements* as

$$\delta_m = -\frac{\lambda}{2\pi}(r_m - r_0) = \Delta_m + k_m\lambda + w_m, m \in \{1, \dots, I-1\}, \quad (3.2)$$

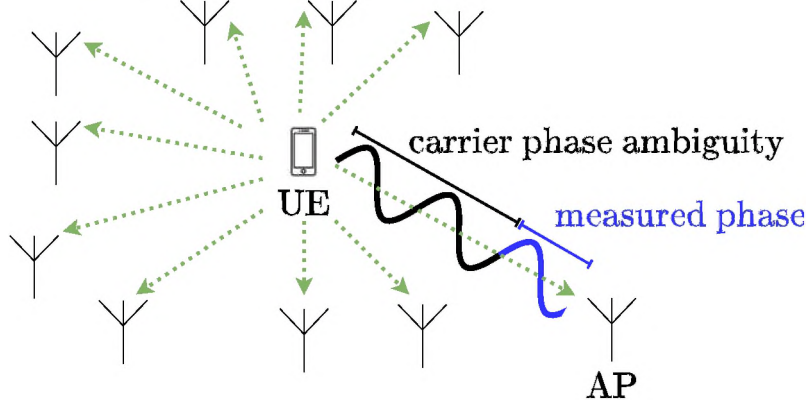


Figure 3.1: Uplink positioning with distributed antenna points (APs) where only the carrier phase measurements at different APs are used to estimate the UE position.

where $\Delta_m = d_m - d_0$, $k_m = z_0 - z_m$ and $w_m = (\lambda/2\pi)(n_0 - n_m)$. It is also assumed that the maximum delay difference between the reference AP and the other APs is much smaller than $1/W$. Based on Fisher information theory, the lower bound on the error covariance of $\mathbf{w} = [w_1, \dots, w_{I-1}]^\top$ can now be shown to read

$$\mathbf{\Sigma}_{\text{diff}} = \frac{\lambda^2 N_0}{8\pi^2 E} (\mathbf{D} + \frac{\mathbf{1}\mathbf{1}^\top}{\rho_0^2}), \quad (3.3)$$

where $\mathbf{D}$ is a diagonal matrix whose m -th diagonal element is given by $[\mathbf{D}]_{m,m} = 1/\rho_m^2$, $\mathbf{1}$ is a vector of $I-1$ ones and the superscript $(\cdot)^\top$ denotes the transpose operation.

As a baseline, we will devise and use the maximum likelihood estimation (MLE). The vector of differential measurements in (3.2) can be expressed as

$$\boldsymbol{\delta} = \mathbf{h}(\mathbf{x}_{\text{ue}}) + \mathbf{w}, \quad (3.4)$$

where $\boldsymbol{\delta} = [\delta_1, \dots, \delta_{I-1}]^\top$, $\mathbf{h}(\mathbf{x}_{\text{ue}}) = [h_1(\mathbf{x}_{\text{ue}}), \dots, h_{I-1}(\mathbf{x}_{\text{ue}})]^\top$, $h_m(\mathbf{x}_{\text{ue}}) = \frac{-\lambda}{2\pi} (\text{mod}(-2\pi\|\mathbf{x}_{\text{ue}} - \mathbf{x}_{\text{ap},m}\|/\lambda, 2\pi) - \text{mod}(-2\pi\|\mathbf{x}_{\text{ue}} - \mathbf{x}_{\text{ap},0}\|/\lambda, 2\pi))$ and $\text{mod}(a, b)$ denotes the modulo operation of a with respect to b . Given that $\mathbf{w}$ follows a multivariate normal distribution with zero-mean and covariance matrix $\mathbf{\Sigma}_{\text{diff}}$, the likelihood function is $\mathcal{L}(\boldsymbol{\delta}|\mathbf{x}_{\text{ue}}) = \mathcal{N}(\boldsymbol{\delta}; \mathbf{h}(\mathbf{x}_{\text{ue}}), \mathbf{\Sigma}_{\text{diff}})$. The MLE for the UE position can be obtained by minimizing the negative log-likelihood, given by

$$\hat{\mathbf{x}}_{\text{ue}} = \arg \min_{\mathbf{x}_{\text{ue}}} (\boldsymbol{\delta} - \mathbf{h}(\mathbf{x}_{\text{ue}}))^\top \mathbf{\Sigma}_{\text{diff}}^{-1} (\boldsymbol{\delta} - \mathbf{h}(\mathbf{x}_{\text{ue}})). \quad (3.5)$$

Since the integer ambiguities are unknown to the AP network, the resulting MLE cost function becomes highly non-smooth, exhibiting numerous local minima and requiring a dense grid search to identify the global minimum. An example of the MLE cost function with ten randomly distributed APs is depicted in Fig. 3.2. As the grid becomes increasingly dense, the computational burden grows significantly, rendering the MLE approach impractical for real-time, latency-critical applications such as dynamic UE tracking or autonomous vehicle localization.

In order to decrease the complexity, one can utilize ML approaches. We propose two neural network-based approaches for indoor CPP:

- The first method, called direct phase-only CPP, uses a single multi layer perceptron (MLP) that takes the differential phase measurements $\boldsymbol{\delta}$ as input and directly estimates the UE position $\hat{\mathbf{x}}_{\text{ue}}$. The MLP consists of an input layer of dimension $I-1$, followed by seven fully connected layers with A , $2A$, $4A$, $8A$, $4A$, $2A$, and A neurons, respectively, and an output layer of dimension 2.

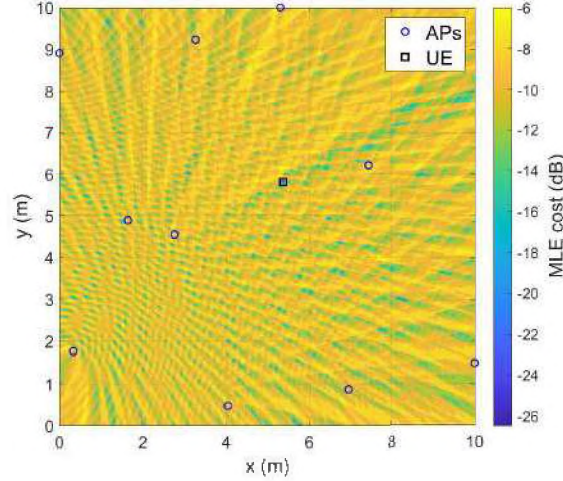
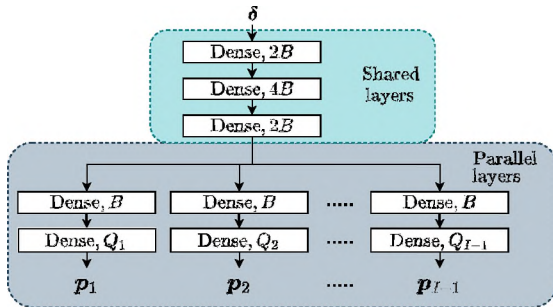
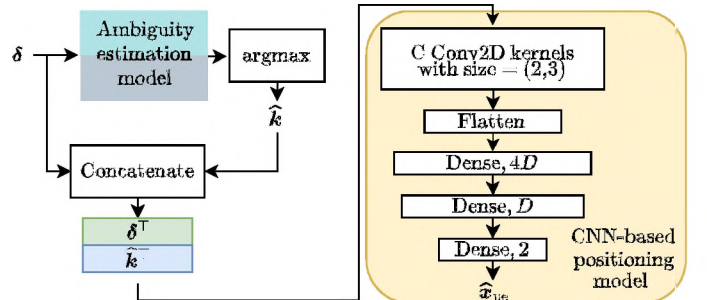


Figure 3.2: Visualization of the MLE-based localization cost over a discretized 10m x 10m 2D area.

- The second approach, termed integer ambiguity-aided CPP, consists of a two-stage neural architecture. In the first stage, an MLP-based ambiguity estimator (shown in Fig. 3.3(a)) processes δ and outputs a probability distribution over possible integer values for each differential ambiguity. These estimated integers are then concatenated with the original phase measurements and passed to the convolutional neural network (CNN)-based positioning model depicted in Fig. 3.3(b). The CNN applies convolutional filtering to the combined input to extract spatial features, which are subsequently processed by dense layers to estimate the final UE position.



((a)) Proposed MLP-based ambiguity estimation model.



((b)) Proposed CNN-based positioning model incorporating MLP-based ambiguity estimation.

Figure 3.3: Overview of the second approach: (a) MLP-based ambiguity estimation and (b) CNN-based positioning model.

To assess the performance of the proposed positioning schemes, we conduct simulations within a square evaluation area of 100 m². A total of 20 antenna points are uniformly randomly distributed across the area to emulate a distributed/cell-free MIMO network architecture. A single UE is placed uniformly at random within the same region. Two carrier frequencies are considered for the system: 800 MHz and 1.8 GHz. The neural network models used in the evaluation are designed with the following layer configurations: dense layers are assigned $A = B = D = 128$ neurons, while the convolutional layer employs $C = 32$ learnable filters. Simulation parameters include a noise power spectral density of -174 dBm/Hz, a receiver noise figure of 13 dB, and an uplink transmit power of 0 dBm. A free-space path loss model is assumed. Uplink pilot signals follow the 5G-NR standard with 15 kHz subcarrier spacing and a transmit bandwidth of 180 kHz,

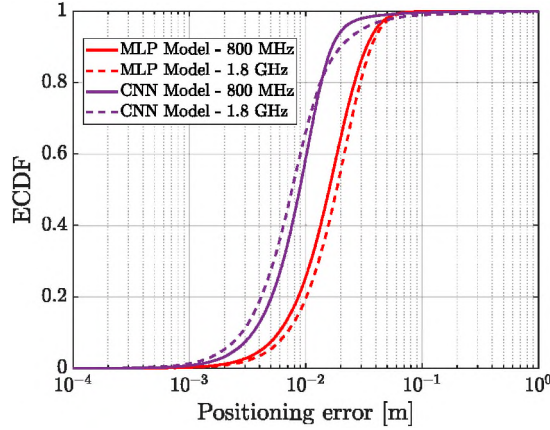


Figure 3.4: ECDF of the proposed approaches.

configured for both frequency options.

The empirical cumulative distribution function (ECDF) of the positioning error for both neural network models is presented in Fig. 3.4. The CNN-based model demonstrates superior accuracy, achieving 95th-percentile errors of 2.08 cm at 800 MHz and 3.08 cm at 1.8 GHz. In contrast, the MLP-based model exhibits higher errors of 4.01 cm at 800 MHz and 4.34 cm at 1.8 GHz, with a noticeable performance degradation at higher frequencies. These results demonstrate the promising potential of deep learning for achieving accurate and efficient phase-only positioning in future cell-free wireless systems, especially in environments where conventional time-of-arrival methods may be impractical.

3.2 3D Indoor Positioning-Aided Beamforming for mmWave Distributed MIMO Systems

The mmWave communication offers ultra-high data rates but suffers from severe path loss and sensitivity to blockage. To sustain reliable connectivity, both transmitter and receiver must employ highly directional beams. However, performing exhaustive beam search across all angular combinations results in excessive training overhead, latency, and energy consumption in NLOS environments. Recent research highlights that integrating user positioning information into the beamforming process can drastically reduce this overhead and improve overall link robustness. In the context of mmWave non-orthogonal multiple access systems, [90] demonstrated that incorporating positioning data to guide beam training confines the search to the most promising spatial directions, thereby reducing beam alignment time, improving reliability, and enhancing energy efficiency in NLOS and mobile conditions. Similarly, [91] proposed a 3D positioning-aided beamforming method that integrates uplink-derived positioning information into downlink beamforming optimization. Their results show a clear correlation between positioning accuracy and spectral efficiency with approximately 1 m positioning accuracy, the 3D PA-BF system achieves nearly the same performance as exhaustive beam search, whereas 10 m and 50 m errors introduce about 15% and > 30% degradation, respectively.

To estimate the user position, we consider an indoor D-MIMO architecture operating at a carrier frequency of $f_c = 28$ GHz. The environment is modeled as a rectangular room with dimensions $20 \times 15 \times 4$ m³. 5 Base Station (BS)s are deployed, where four are positioned near the room corners at a height of approximately 3 m and one ceiling-mounted BS at the center to improve vertical coverage. 5 UEs are placed at a height of 1.2 m, distributed across the room to ensure that each UE maintains LOS visibility with at least three BS. Each BS operates with an OFDM waveform consisting of 100 MHz bandwidth.

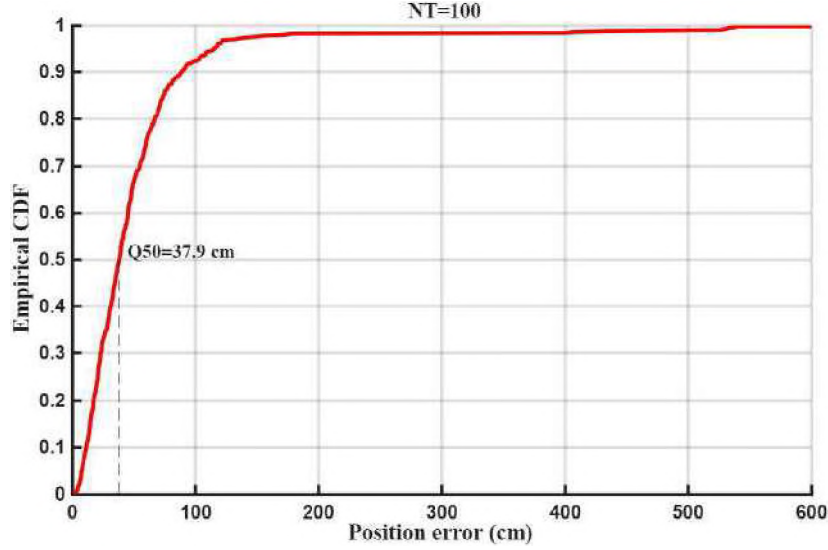


Figure 3.5: CDF of 3D positioning error obtained over 100 Monte Carlo trials

The noise power spectral density is $N_0 = -174$ dBm/Hz, the total noise power

$$P_n = N_0 + 10 \log_{10}(\text{BW}) \text{ dBm} \quad (3.6)$$

and wavelength $\lambda = c/f_c$, where c is the speed of light.

The geometric distance and angles between each BS and UE pair are expressed as

$$d_{m,u} = \sqrt{(x_b - x_u)^2 + (y_b - y_u)^2 + (z_b - z_u)^2} \quad (3.7)$$

$$\text{az}_{m,u} = \tan^{-1} \left(\frac{y_u - y_b}{x_u - x_b} \right) \quad (3.8)$$

$$\text{el}_{m,u} = \tan^{-1} \left(\frac{z_u - z_b}{\sqrt{(x_u - x_b)^2 + (y_u - y_b)^2}} \right) \quad (3.9)$$

where (x_b, y_b, z_b) and (x_u, y_u, z_u) represent the coordinates of the BS and UE, respectively.

The received power is modeled using the Free-Space Path Loss (FSPL) equation with log-normal shadowing as

$$\text{FSPL}(d) = 20 \log_{10} \left(\frac{4\pi d}{\lambda} \right) \quad (3.10)$$

$$P_r = P_t - \text{FSPL}(d) + X, \quad X \sim \mathcal{N}(0, \sigma_{\text{shad}}^2) \quad (3.11)$$

where $P_t = 10$ dBm and $\sigma_{\text{shad}} = 4$ dB. The instantaneous SNR for each BS-UE link is given by

$$\text{SNR}_{m,u} = P_r - P_n \quad (3.12)$$

A link is retained if $P_r \geq -95$ dBm and $\text{SNR} \geq -5$ dB, ensuring that only reliable high-SNR connections contribute to position estimation.

Each valid link produces a time-of-flight (ToF) measurement derived from the phase variation of OFDM pilot tones

$$H_k = e^{-j2\pi f_k \tau_0} \quad (3.13)$$

where $\tau_0 = d_{m,u}/c$ is the true propagation delay between the BS and UE.

By unwrapping the pilot phase and performing a linear least-squares fit across the pilot subcarriers, the estimated delay is obtained as

$$\hat{\tau} = -\frac{1}{2\pi} \frac{d\angle H_k}{df_k} \quad (3.14)$$

where $\angle H_k$ denotes the unwrapped phase of the frequency response at the k^{th} pilot tone.

The timing uncertainty is modeled as

$$\sigma_\tau = \frac{\kappa}{\text{BW}\sqrt{\text{SNR}}} \quad \sigma_r = c\sigma_\tau \quad (3.15)$$

and serves as the weighting factor for range fusion during position estimation.

Each BS employs an Uniform Planar Array (UPA) for AoA estimation using the three-dimensional MUSIC algorithm. The space-frequency steering vector is defined as

$$\mathbf{a}_{\text{sf}}(\alpha, \epsilon, \tau) = \mathbf{b}(\tau) \otimes \mathbf{a}(\alpha, \epsilon) \quad (3.16)$$

where $\otimes$ denotes the Kronecker product, $\mathbf{b}(\tau) = [e^{-j2\pi f_k \tau}]_{k=1}^{K_{\text{pilot}}}$ is the frequency-domain steering vector, and $\mathbf{a}(\alpha, \epsilon)$ represents the spatial UPA steering vector parameterized by azimuth α and elevation ϵ .

Given the space-frequency covariance matrix $\mathbf{R}_{\text{sf}}$, the MUSIC pseudospectrum is expressed as

$$P(\alpha, \epsilon, \tau) = \frac{1}{\mathbf{a}_{\text{sf}}^H \mathbf{E}_n \mathbf{E}_n^H \mathbf{a}_{\text{sf}}} \quad (3.17)$$

where $\mathbf{E}_n$ denotes the noise subspace obtained from the eigen-decomposition of $\mathbf{R}_{\text{sf}}$. The spectral peak of $P(\alpha, \epsilon, \tau)$ yields the 3D MUSIC estimates $(\hat{\alpha}, \hat{\epsilon}, \hat{\tau}_{\text{MUSIC}})$.

Finally, a hybrid Weighted Least-Squares (WLS) estimator fuses the ToF and AoA measurements by minimizing the cost function

$$J(\mathbf{x}) = \sum_i \left(\frac{\|\mathbf{x} - \mathbf{b}_i\| - c\hat{\tau}_i}{\sigma_{r,i}} \right)^2 + \sum_i \left(\frac{\hat{\alpha}_i - \alpha_i(\mathbf{x})}{\sigma_\alpha} \right)^2 + \sum_i \left(\frac{\hat{\epsilon}_i - \epsilon_i(\mathbf{x})}{\sigma_\epsilon} \right)^2 \quad (3.18)$$

where $\mathbf{x}$ is the unknown UEs position, $\mathbf{b}_i$ denotes the i^{th} BSs position, c is the speed of light, and $(\hat{\tau}_i, \hat{\alpha}_i, \hat{\epsilon}_i)$ are the corresponding measurements with standard deviations $(\sigma_r, \sigma_\alpha, \sigma_\epsilon)$.

The position estimate is refined iteratively using a damped Gauss-Newton method

$$\mathbf{x}_{k+1} = \mathbf{x}_k - (\mathbf{J}^T \mathbf{J} + \mu \mathbf{I})^{-1} \mathbf{J}^T \mathbf{e} \quad (3.19)$$

where $\mathbf{J}$ is the Jacobian of residuals, μ is a damping factor, and $\mathbf{e}$ represents the residual vector. Convergence is achieved when $\|\mathbf{x}_{k+1} - \mathbf{x}_k\| < 10^{-4}$. The resulting estimated position $\hat{\mathbf{x}}$ provides the foundation for beam direction initialization in the PA-BF stage.

A Monte-Carlo simulation with NT=100 independent trials was conducted to evaluate the robustness of the proposed framework under LOS indoor conditions. In each trial, random shadowing and noise were regenerated to emulate realistic link variability among the distributed BSs. Fig. 3.5 presents the empirical Cumulative Distribution Function (CDF) of the positioning error. The median (Q50) error is 37.9 cm, the 90th percentile (Q90) is 88.8 cm, and the 95th percentile (Q95) is 115.6 cm. Nearly all localization estimates lie within 1 m of the ground truth, validating the centimeter-level accuracy of the hybrid ToF and AoA fusion scheme. The results show that the proposed framework achieves a mean localization accuracy of 30-40 cm with limited variation across users. Central UE achieves the lowest errors due to symmetrical BSs visibility, while corner UEs show slightly higher variance caused by wider angular spread. Despite these differences, all UEs maintain sub-meter accuracy across most trials. Fig. 3.6 shows the error heatmap, depicting per UE localization error (cm) over all trials. The majority of errors remain within the 0-100 cm range (dark blue regions), confirming high positional stability. Only a few isolated high-error instances (yellow bands exceeding 400 cm) appear when certain BS-UE links fall below the SNR threshold.

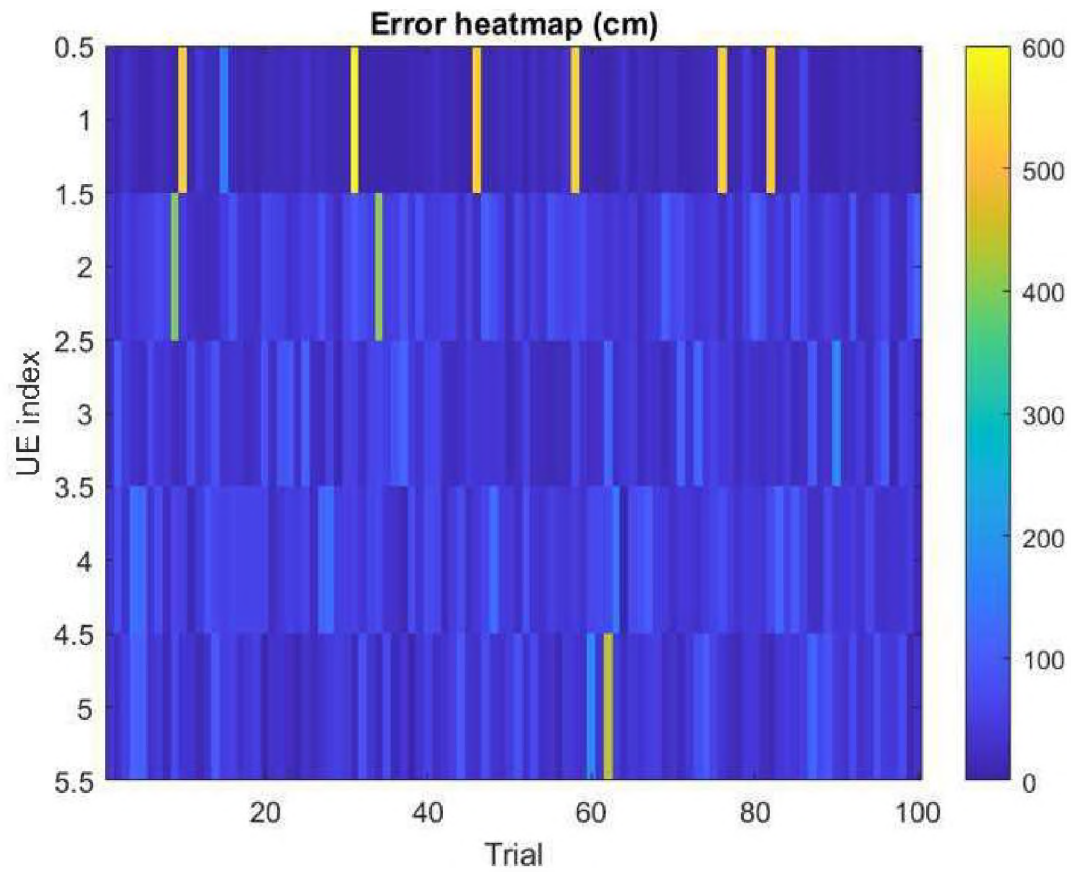


Figure 3.6: Heatmap of 3D positioning error (cm) across 100 Monte Carlo trials for five user equipments.

Chapter 4

Radio SLAM-related algorithms devellopped in MiFuture

Radio SLAM refers to the process of using radio signals, such as 5G or 6G, to simultaneously estimate a device's location and build a map of the surrounding environment (i.e., estimate the landmarks' positions). Unlike the traditional SLAM methods from robotics that rely on visual or inertial sensors, radio SLAM leverages the spatial and temporal characteristics of the radio wave propagation, including multipath reflections and, possibly, signal strength variations. This approach is particularly valuable in environments where visual sensing is unreliable or unavailable, such as dark, smoky, or cluttered industrial spaces, and can be integrated into existing wireless infrastructure for low-cost deployment. A central challenge in radio SLAM is the DA problem, which involves correctly matching observed radio features (e.g., signal reflections, channel impulse responses) to physical landmarks or previously mapped signal sources. Unlike visual features, radio signals are often ambiguous due to multipath propagation, dynamic environments, and NLOS conditions. This makes it difficult to determine whether a received signal corresponds to a known location, a new feature, or a transient change in the environment. Solving the data association problem is critical for maintaining accurate localization and consistent map updates, and typically requires probabilistic models, machine learning techniques, or fusion with other sensor modalities to improve robustness. The DA aspects are addressed in Section 4.1.

4.1 Data Association Problem

In EKF-SLAM algorithm one of the main drawbacks is the capacity of handling data associations errors. In [47] the authors try to solve this problem using a robust cost function that is correcting the measurement noise covariance in the update step of the EKF algorithm. Instead of minimizing the negative log posterior which is proportional to the sum of the likelihood and the prior, we propose to minimize:

$$\tilde{q}(\mathbf{x}_k) = \frac{1}{2} \|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\|_{\mathbf{P}_{k|k-1}^{-1}}^2 + \frac{1}{2} \sum_{i=0}^{n_k} \rho(\varepsilon_i^2), \quad (4.1)$$

where $\varepsilon_i^2 \triangleq \|\mathbf{z}_k^i - \mathbf{g}(\mathbf{x}_k^{\gamma(i)})\|_{(\mathbf{R}_k^i)^{-1}}^2$ is the error and $\rho(\cdot)$ is a robust cost function. The chosen robust cost function is the DCS method, it has a low computational overhead and has shown good performance in graph-based SLAM. DCS provides a closed-form solution for computing the scaling factor individually for each measurement which depends on the original error term. The DCS cost function is given by [92]

$$\rho(\varepsilon_i^2) \triangleq s_i^2 \varepsilon_i^2 + \psi(1 - s_i)^2, \quad (4.2)$$

where $s_i \in [0, 1]$ is a switch variable that controls the influence of measurement i and ψ is a user defined parameter. Agarwal *et al.* [92] solve for the optimal s_i in closed-form and show that $\rho(\varepsilon_i^2) \leq \psi$ holds for every constraint. In theory, s_i can be freely chosen within the interval that satisfies the inequality, but the authors propose to use the value that minimizes $\rho(\varepsilon_i^2)$ while not exceeding 1, which yields [92]

$$s_i = \begin{cases} 1 & \text{if } \varepsilon_i^2 \leq \psi, \\ \frac{2\psi}{\psi + \varepsilon_i^2} & \text{otherwise.} \end{cases} \quad (4.3)$$

This constraint makes DCS a quadratic cost function for small errors ($\varepsilon_i^2 \leq \psi$), and similar to Geman-McClure for large errors ($\varepsilon_i^2 > \psi$). Using DCS, the proposed objective function can now be expressed as

$$\tilde{q}(\mathbf{x}_k) = \frac{1}{2} \|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\|_{\mathbf{P}_{k|k-1}^{-1}}^2 + \frac{1}{2} \sum_{i=0}^{n_k} \|\mathbf{z}_k^i - \mathbf{g}(\mathbf{x}_k^{\gamma(i)})\|_{(\tilde{\mathbf{R}}_k^i)^{-1}}^2, \quad (4.4)$$

where $\tilde{\mathbf{R}}_k^i = s_i^{-2} \mathbf{R}_k^i$ is the scaled measurement covariance. Using DCS, the prediction step of the EKF-SLAM algorithm remains unchanged (see e.g., [43]), whereas the update step of the proposed robust EKF-SLAM algorithm can be computed as:

$$\tilde{\mathbf{S}}_k = \mathbf{G}_x \mathbf{P}_{k|k-1} \mathbf{G}_x^\top + \tilde{\mathbf{R}}_k, \quad (4.5)$$

$$\tilde{\mathbf{K}}_k = \mathbf{P}_{k|k-1} \mathbf{G}_x^\top \tilde{\mathbf{S}}_k^{-1}, \quad (4.6)$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \tilde{\mathbf{K}}_k (\mathbf{z}_k - \mathbf{g}(\hat{\mathbf{x}}_{k|k-1})), \quad (4.7)$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \tilde{\mathbf{K}}_k \tilde{\mathbf{S}}_k \tilde{\mathbf{K}}_k^\top. \quad (4.8)$$

In the paper [47] incorrect data association is forced to verify the robustness of the algorithm against DA error. The DA variable is known and, incorrect DAs are forced with a specified probability and severity to systematically investigate the robustness of the proposed filter to DA errors. The errors are modelled using a Gaussian mixture distribution.

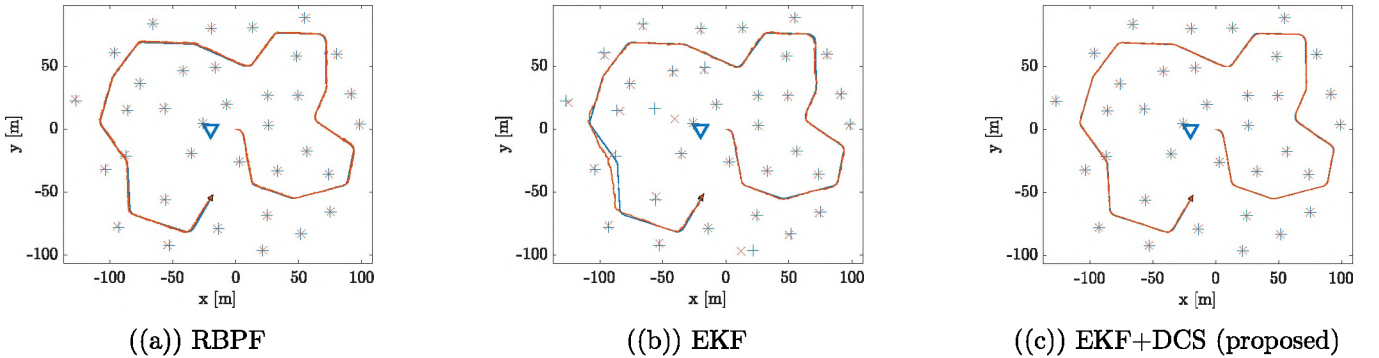


Figure 4.1: Bistatic radio SLAM illustration for RBPF (a), standard EKF (b) and the proposed EKF-SLAM algorithm (c). The probability of correct DA is set to 0.9. In each figure, the blue triangle denotes the known BS location. Additionally, solid blue and red lines indicate the true and estimated UE trajectories, respectively, while the blue and red markers denote the true and estimated landmarks.

As illustrated in 4.1, the standard EKF is severely impacted by the DA errors even though the number of correct DAs is relatively high, of 0.9. The estimated UE positions and landmarks do not align very well with the ground truth values, especially towards the end of the trajectory. In contrast, the proposed method is robust against DA errors resulting in much better SLAM performance. The estimated UE trajectory closely follows the ground truth trajectory and the estimated landmarks are very close to the the actual

landmarks. The algorithm has similar performance to Rao-Blackwellized particle filter (RBPF) but they are computationally less expensive, because for achieving this good performance the number of particle used are very high.

4.2 SLAM performance in non-collaborative approach

The performance metrics of the SLAM algorithm with non-collaborative UEs are evaluated via the RMSE. For the UE position, heading and clock bias are evaluated; for the mapping the landmark position the accuracy is evaluated. It can be seen in Fig. 4.2 that our proposed EKF+DCS algorithm highly outperforms the traditional EKF and slightly outperforms the RBPF algorithm; however it has a much lower complexity than RBPF algorithm.

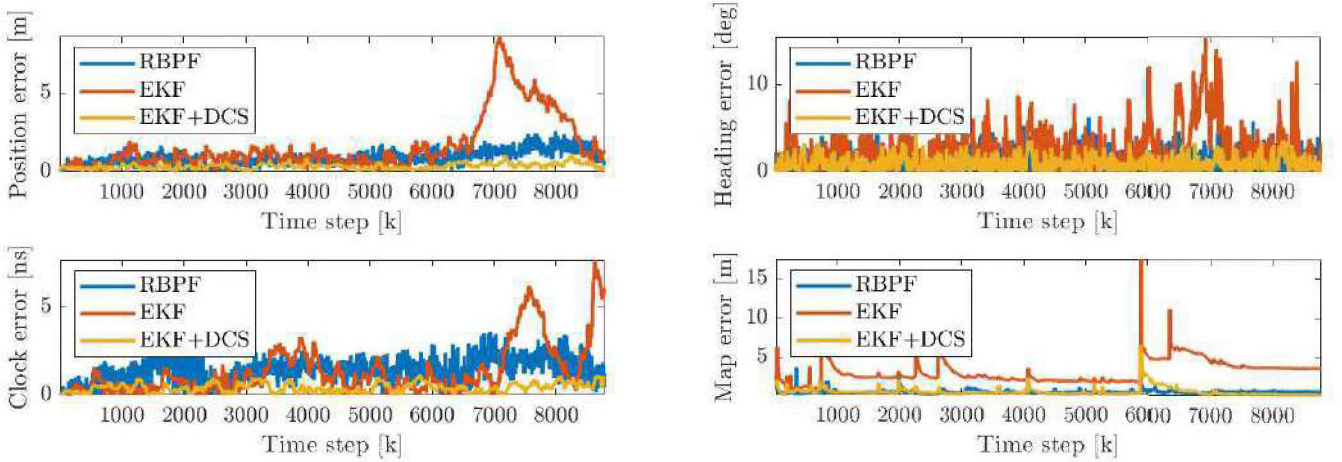


Figure 4.2: Performance metrics comparison of the results reported in Fig. 4.1. The averaged over the time steps RMSEs of each methods are reported for position, heading, clock bias and mapping.

4.3 Collaborative SLAM

The SLAM principles that apply to a single UE can be extended by deploying multiple instances and their performance utilized collaboratively through information sharing between them.

Collaborative SLAM is an extension of traditional SLAM where multiple UEs jointly estimate their own positions while building and maintaining a shared map of the environment, as depicted in the Figure 4.3 below. By exchanging information about observations, features, or estimated states, the UEs can improve accuracy, robustness, and coverage. Collaborative SLAM is enabling collective situational awareness and cooperative perception through sensing, communication, and data fusion. For example in vehicular scenarios it has potential to extend the sensing capabilities of vehicles by adding another layer of modality for extended robustness and safety.

To illustrate the benefit of collaboration in SLAM, the *Cramér–Rao Posterior Bounds* (PCRB) [93] were derived and simulated for a multi-UE system. They represent a theoretical lower limit on the estimation accuracy of unknown parameters when both the measurement data and the prior knowledge are taken into account. In other words, it defines the smallest achievable error covariance for any unbiased estimator, given both the sensor observations and the prior uncertainty.

Let the posterior covariance matrix of the estimated state be denoted as:

$$\mathbf{P} = \mathbb{E}[(\hat{\mathbf{x}} - \mathbf{x})(\hat{\mathbf{x}} - \mathbf{x})^\top]. \quad (4.9)$$

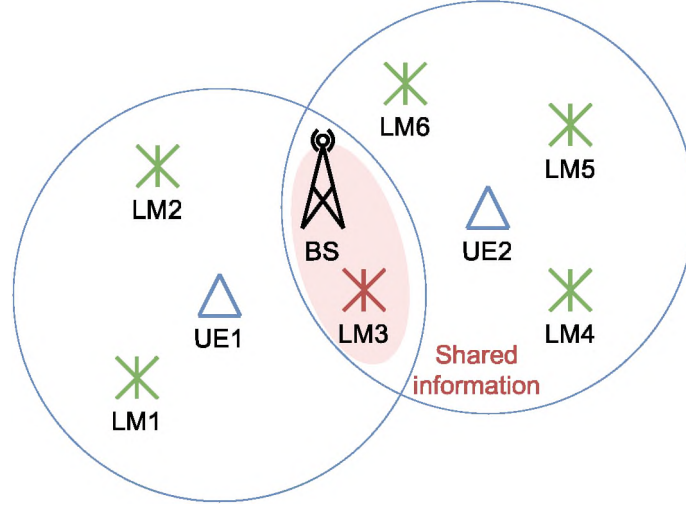


Figure 4.3: The red landmark (LM) 3 is within the blue circle filed of view of both UE1 and UE2.

According to the PCRB, this covariance is bounded as:

$$\mathbf{P} \succeq \mathbf{J}^{-1}, \quad (4.10)$$

where $\mathbf{J}$ is the posterior Fisher Information Matrix (FIM).

For position components, if $\mathbf{x} = [x, y, \dots]^\top$, the position error bound can be extracted from the corresponding 2×2 submatrix:

$$\mathbf{P}_{\text{pos}} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}. \quad (4.11)$$

The per-axis position bounds are obtained directly from the diagonal elements as:

$$\sigma_x = \sqrt{[\mathbf{P}_{\text{pos}}]_{1,1}}, \quad \sigma_y = \sqrt{[\mathbf{P}_{\text{pos}}]_{2,2}}. \quad (4.12)$$

Consequently, the overall position error bound can be expressed as:

$$\text{RMSE}_{\text{pos}} = \sqrt{\sigma_x^2 + \sigma_y^2} = \sqrt{[\mathbf{P}_{\text{pos}}]_{1,1} + [\mathbf{P}_{\text{pos}}]_{2,2}}. \quad (4.13)$$

Considering a set of independent measurements $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_K$ that each depend on the same unknown parameter vector $\mathbf{x}$, if their joint likelihood factorizes as:

$$p(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_K | \mathbf{x}) = \prod_{k=1}^K p(\mathbf{z}_k | \mathbf{x}), \quad (4.14)$$

then the corresponding FIMs are additive. The total FIM can be expressed as:

$$\mathbf{J}_{\text{total}}(\mathbf{x}) = \sum_{k=1}^K \mathbf{J}_k(\mathbf{x}), \quad (4.15)$$

In the context of Collaborative Radio SLAM, each user equipment (UE) contributes its own measurements $\mathbf{z}^{(m)}$ that provide information about both its own state and possibly shared parameters of the map of set of landmarks. Assuming that the measurements of different UEs are conditionally independent given the

global state $\mathbf{x}$, the joint likelihood follows equation 4.14 and therefore the global Fisher Information Matrix is additive across UEs (Equation 4.15).

This property naturally motivates the definition of the multi-UE system PCRB, in which the collective estimation accuracy bound is obtained by combining the information contributions from all participating UEs. In this manner, collaboration effectively increases the available information and tightens the achievable lower bound on the estimation error covariance.

The Figure 4.4(a) depicts a close-up of the map used in simulations, where numbered blue crosses denote the landmarks, the blue triangle is BS position, and squares denote waypoints of a default UE ground-truth path.

Figure 4.4(b) then shows an example of 4 UEs deployed in that environment. Each path is defined by a set of waypoints $\mathbf{w}_i = [x_i, y_i]^T$. To introduce randomness in their positions, each waypoint is noised by a random displacement vector generated in polar coordinates. This setup reflects a real-world situation where vehicles follow the same path while changing lanes or maintaining lateral offsets.

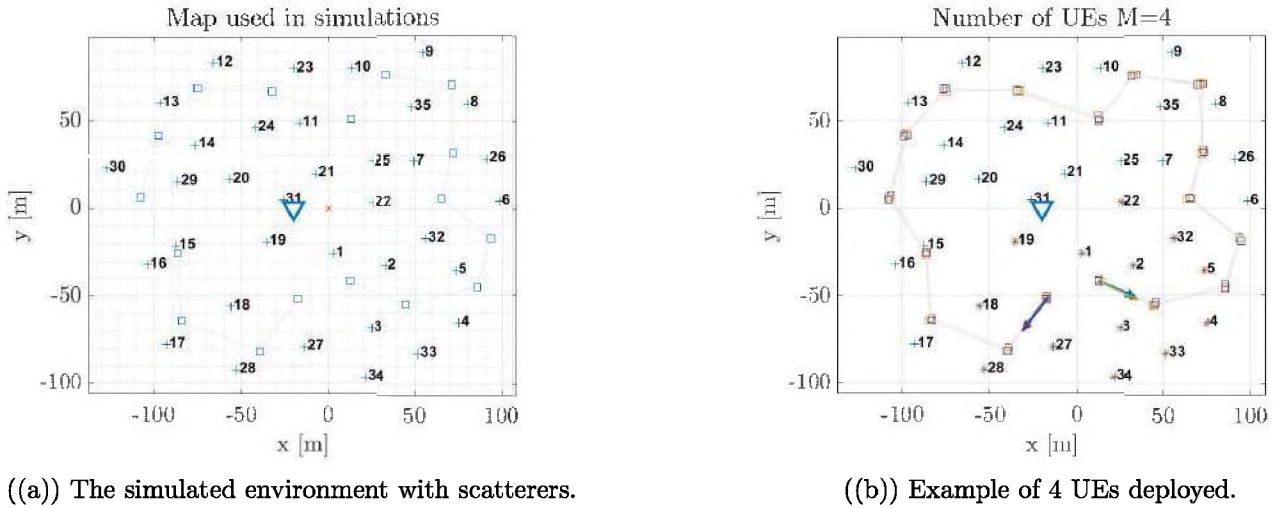


Figure 4.4: The ground-truth map environment and waypoints of paths.

First, a random angle θ_i is drawn uniformly from $[0, 2\pi)$:

$$\theta_i \sim \mathcal{U}(0, 2\pi). \quad (4.16)$$

Then, a random radial distance r_i is sampled such that the resulting displacement is uniformly distributed over a disk of radius $r_{\max} = \text{wp_displacement_max}$ (2 m in this document):

$$r_i = r_{\max} \sqrt{u_i}, \quad u_i \sim \mathcal{U}(0, 1). \quad (4.17)$$

The Cartesian displacement vector is obtained by converting from polar to Cartesian coordinates:

$$\delta_i = \begin{bmatrix} r_i \cos \theta_i \\ r_i \sin \theta_i \end{bmatrix}. \quad (4.18)$$

Finally, the noised waypoint is given by:

$$\mathbf{w}_i^{\text{new}} = \mathbf{w}_i + \delta_i. \quad (4.19)$$

In matrix form, for all waypoints,

$$\mathbf{W}_{\text{new}} = \mathbf{W} + \begin{bmatrix} r \odot \cos \theta \\ r \odot \sin \theta \end{bmatrix}, \quad (4.20)$$

where $\odot$ denotes the element-wise multiplication.

As shown in Figure 4.5, the average position error (in meters) is plotted as a function of the number of UEs, denoted by M , over 100 Monte Carlo simulation runs using map depicted in Figure 4.4 above, with noised waypoints as described in Equations 4.16–4.20 and half the UEs going through the map clockwise, half counter-clockwise. The vertical axis employs a logarithmic scale illustrating the positioning accuracy across several orders of magnitude.

The results indicate that the average position error decreases sharply as M increases, demonstrating significant improvements in localization accuracy when more UEs are involved. Beyond approximately $M = 100$, the rate of improvement slows down, and the error stabilizes around 10^{-2} m, suggesting diminishing returns in accuracy gains as the number of UEs becomes large. This trend highlights the benefit of cooperative estimation with multiple UEs, while also revealing the saturation effect in performance beyond a certain network size. Presented results accumulate an average of mean errors, i.e. if scenario had 4 UEs, an average of 4 per-UE mean values was taken.

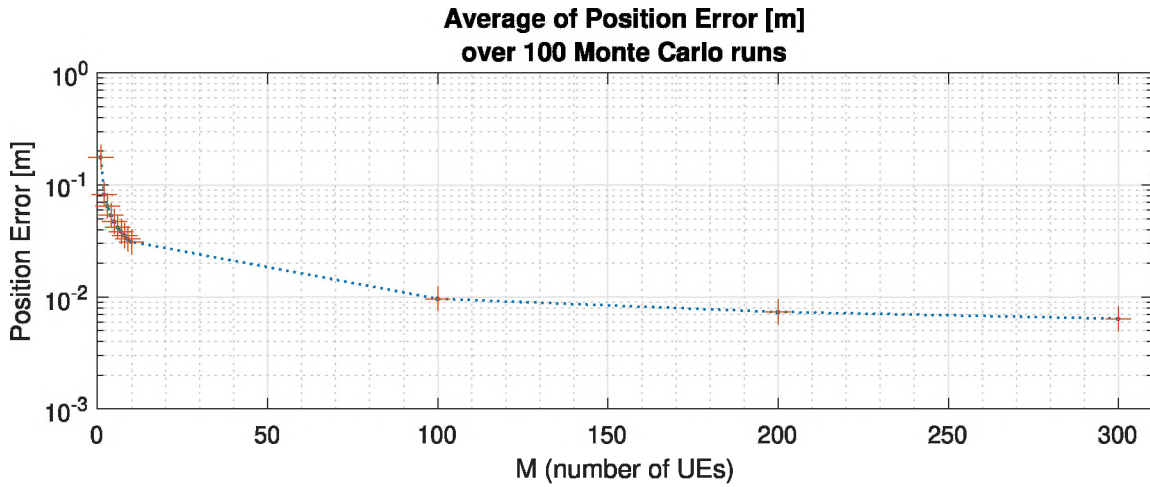


Figure 4.5: Average of UE lower PCRB position error over 100 Monte Carlo runs.

Figure 4.6 illustrates the average landmark position error as a function of the number of UEs M , evaluated over 100 Monte Carlo runs. Similar to the trend observed in Figure 4.5, the landmark position error decreases significantly as the number of UEs increases, confirming that cooperative sensing improves the accuracy of landmark estimation. This behavior suggests that a larger number of UEs provides more reliable geometric constraints for landmark localization, though the incremental improvement becomes less pronounced with larger number of UEs.

Figure 4.7 suggests that more scattering points improve the theoretically achievable estimation error bound.

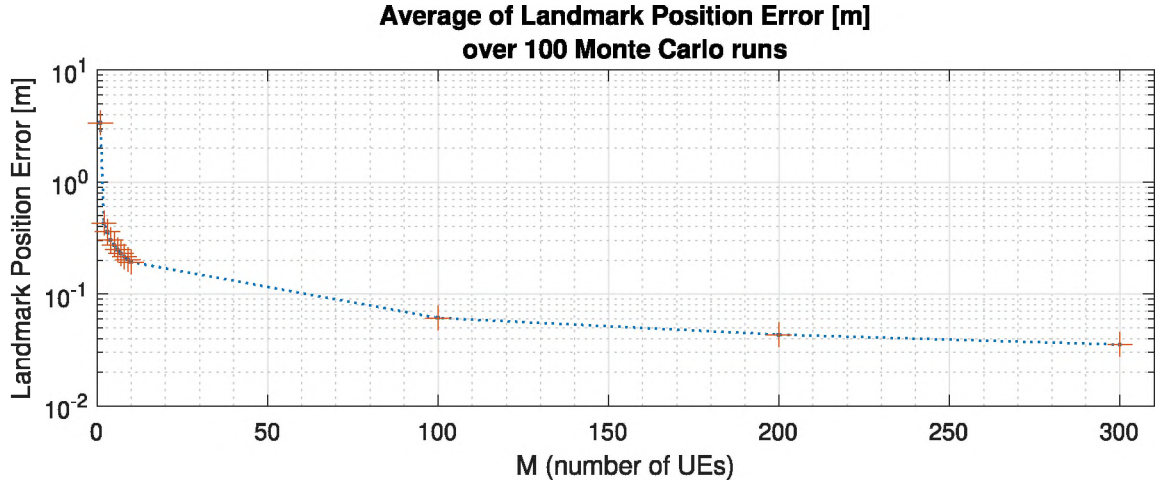


Figure 4.6: Average of Landmark lower PCRB position error over 100 Monte Carlo runs.

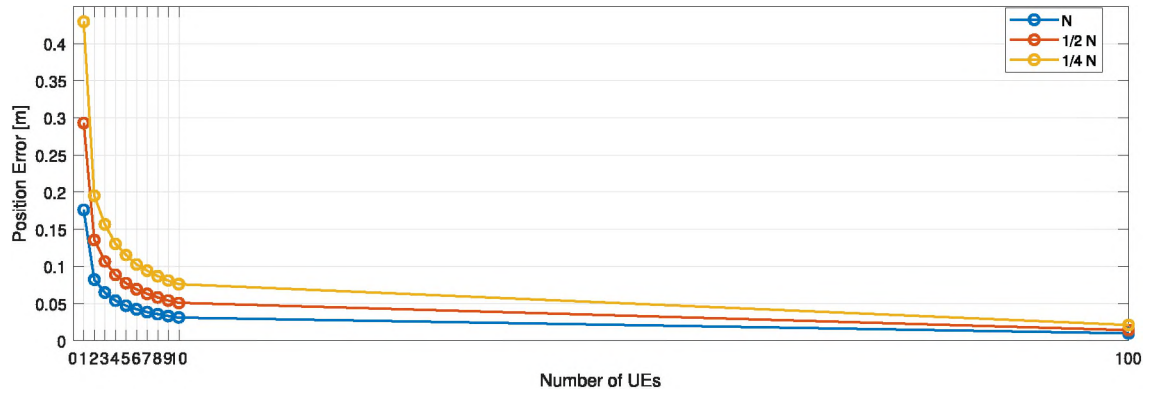
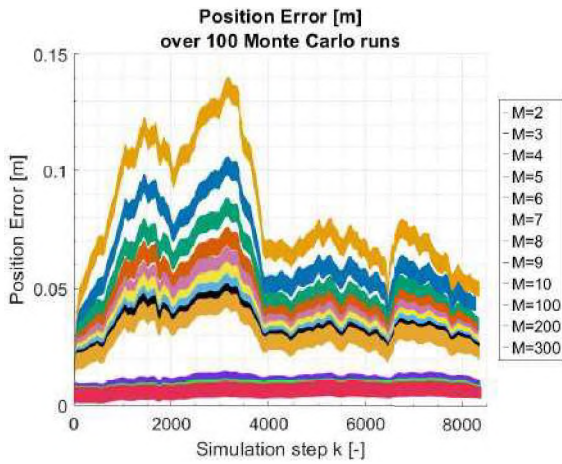
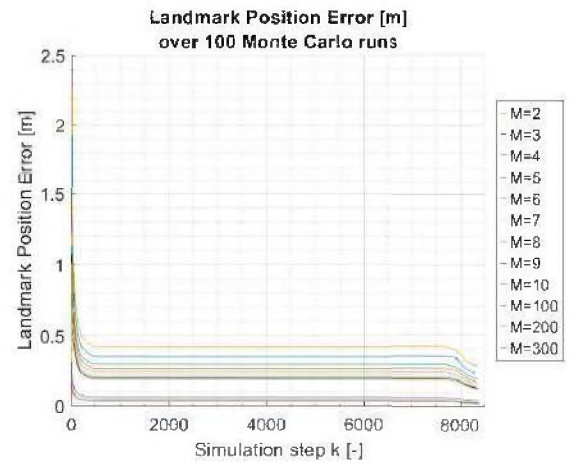


Figure 4.7: Effect of number of landmarks on the bounds ($N=35$) on an example of UE position estimation.



((a)) Position estimation of UE.



((b)) Position estimation of LM.

Figure 4.8: The lower bounds as time series.

Chapter 5

Conclusions and future research

5.1 Main findings

Our MiFuture research in radio-based positioning and sensing demonstrates that **centimeter-level accuracy** is achievable for both **user equipment (UE) localization** and **environment mapping** under realistic conditions. Theoretical lower error bounds in **Collaborative Radio SLAM** confirm that, when multiple UEs cooperate, estimation errors for both UE positions and scatterer locations can approach the centimeter level. Importantly, these bounds improve significantly with the number of collaborating devices or users, converging toward the theoretical limit when hundreds of UEs are deployed. This scalability highlights the strong potential of collaborative approaches for high-precision positioning in dense networks.

Performance enhancements remain feasible even in the presence of **data association errors**, which are common in dynamic environments in radio SLAM. Combining **Extended Kalman Filter** algorithms with **dynamic covariance scaling** has proven effective in mitigating these errors. This hybrid approach enables robust correction mechanisms, ensuring high accuracy for both positioning and mapping tasks despite imperfect measurements.

Accurate positioning also unlocks **network-level optimizations**. When user location is known, resource allocation and handover strategies can be proactively managed. Access points (APs) can be selected to meet SNR requirements for positioning while simultaneously improving SE. Furthermore, precise location information simplifies **power management optimization**, which in turn enhances **beamforming** in distributed cell-free networks. These benefits illustrate the strong interplay between positioning accuracy and overall network performance.

Moreover, our studies further demonstrate that **carrier-phase-based localization**, combined with **AI/ML algorithms**, can achieve centimeter-level accuracy in **distributed or cell-free MIMO systems**. This capability is a critical enabler for future technologies such as autonomous mobility, smart manufacturing, and advanced 6G services, where ultra-reliable and precise positioning is indispensable.

Last but not least, accurate user positioning is essential for **mmWave distributed MIMO (D-MIMO)** systems to maintain precise beam alignment and avoid link degradation caused by misdirected beams. A positioning framework leveraging **range and angular measurements** has shown promising results, achieving centimeter-level accuracy and providing a robust foundation for **position-aided beamforming (PA-BF)** in next-generation networks. Additionally, accurate near-field range-angle estimation—together with its corresponding CRB benchmarks—further strengthens both sensing and positioning capabilities in future cell-free and ultra-massive MIMO systems.

In summary, centimeter-level positioning and mapping accuracy is no longer theoretical—it is becoming practical through collaborative SLAM, advanced filtering techniques, AI-driven carrier-phase localization, and integration with network optimization strategies. These advancements position radio-based sensing and SLAM as key pillars for enabling high-performance, resilient, and intelligent 6G systems.

5.2 Open Challenges and Future Research Directions

Despite significant progress in radio positioning, sensing, and SLAM, several open challenges remain that require further investigation.

First, there is the issue of **scalability and complexity**: for example, collaborative Radio SLAM relies on multiple UEs sharing information, which introduces communication overhead and computational complexity. Efficient distributed algorithms and low-latency coordination mechanisms are needed to scale to dense networks.

Secondly, the **robustness in dynamic environments** needs to be addressed in more detail. Mobility, blockage, and multipath propagation can degrade positioning accuracy. Future research should focus on adaptive filtering and learning-based approaches that maintain centimeter-level precision under highly dynamic conditions.

Thirdly, the **integration with AI/ML** poses several generalization and explainability problems. For example, while AI-driven carrier-phase localization shows promise, challenges remain in model generalization, training data diversity, and real-time inference. Developing lightweight, energy-efficient models for edge deployment is a key research direction.

Fourthly, the **location-based resource optimization and beamforming** require tight integration between positioning and communication layers. Future work should explore joint optimization frameworks for D-MIMO and cell-free architectures.

Fifthly, the **standardization and interoperability** aspects have not yet been addressed within MiFuture research so far. Achieving global adoption of radio SLAM and high-precision positioning demands standardized protocols and interoperability across vendors and technologies, especially for emerging 6G systems.

Last but not least, the **security and privacy** aspects are very important and need further consideration. For example, high-accuracy positioning raises concerns about user privacy and resilience against spoofing or jamming attacks. Research into secure localization protocols and privacy-preserving SLAM is essential.

In summary, addressing these challenges will require a multidisciplinary approach, as the one adopted within MiFuture, combining advanced signal processing, machine learning, network optimization, and security mechanisms. These efforts will further pave the way for robust, scalable, and secure positioning solutions that underpin future 6G services and beyond.

List of Acronyms

2D	Two-Dimensional
AI	artificial intelligence
A&M	Aboutanios and Mulgrew
AoA	Angle-of-Arrival
AP	Access Point
AR	Augmented Reality
AWGN	additive white Gaussian noise
BP	belief propagation
BS	Base Station
BS	Base Station
CDF	Cumulative Distribution Function
CNN	convolutional neural network
CPP	carrier phase positioning
CSI	channel state information
CSI-RS	channel state information reference signal
DA	data association
DATMO	Detection and Tracking of Moving Objects
DCS	dynamic covariance scaling
DD	Delay-Doppler
DL	Deep learning
D-MIMO	Distributed MIMO
DMRS	Demodulation reference signal
DNN	Deep neural network
DPG	Dynamic Pose Graph
DT	Delay-Time
ECDF	empirical cumulative distribution function
EKF	Extended Kalman filter
EP	Embedded Pilot
FSPL	Free-Space Path Loss

FTST Fast-Time-Slow-Time

GDSS Gabor Division Spread Spectrum

GNSS Global navigation satellite system

GPR Gaussian Process Regression

ICI Inter-carrier interference

IRS Intelligent reconfigurable surface

ISAC Integrated Sensing and Communications

ISI Inter-symbol interference

LIDAR Light Detection and Ranging

LOS Line of Sight

LoS Line of Sight

MDL minimum description length

MIMO Multiple-input multiple-output

ML Machine Learning

MLE maximum likelihood estimation

ML-MFS Maximum Likelihood based on Multi-Frame Superposition

MLP multi layer perceptron

MMSE Minimum mean square error

mmWave millimeter-wave

MOT Multi-Object tracking

MRT Maximum Ratio Transmission

MUSIC multiple signal classification

NLoS Non Line of Sight

NLOS Non-line-of-sight

NR New Radio

OFDM Orthogonal frequency division multiplexing

OTFS Orthogonal time frequency space

PA-BF Positioning-Aided Beamforming

QoS Quality of Service

RBPF Rao-Blackwellized particle filter

RCS radar cross section

RFS Random finite set

RIS Reconfigurable Intelligent Surface

RMSE Root Mean Square Error

RMSE Root mean squared error

RNN recurrent neural network

RSS Received Signal Strength

RSSI received signal strength indicator

RV Range-Velocity

SBL Sparse Bayesian Learning

SE Spectral Efficiency

SINR Signal to Interference and Noise Ratio

SLAM simultaneous localization and mapping

SLAMMOT simultaneous localization and mapping and Multi-Object Tracking

SNR Signal to Noise Ratio

SNR signal-to-noise ratio

SRS sounding reference signal

TD_{oA} time-difference-of-arrival

TF Time-Frequency

TOA Time of Arrival

ToF time-of-flight

UE User Equipment

UPA Uniform Planar Array

WLS Weighted Least-Squares

WOA-MFS Whale Optimization Multi-Frame Superposition

ZF Zero Forcing

References

- [1] G. K. Carvajal, M. F. Keskin, C. Aydogdu, O. Eriksson, H. Herbertsson, H. Hellsten, E. Nilsson, M. Rydström, K. Vänas, and H. Wymeersch, “Comparison of automotive fmcw and ofdm radar under interference,” in *2020 IEEE Radar Conference (RadarConf20)*, pp. 1–6, 2020.
- [2] C. Zhang, Z. Zhou, H. Wang, and Y. Zeng, “Integrated super-resolution sensing and communication with 5g nr waveform: Signal processing with uneven cps and experiments: (invited paper),” in *2023 21st International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, pp. 681–688, 2023.
- [3] W. Yuan, L. Zhou, S. K. Dehkordi, S. Li, P. Fan, G. Caire, and H. V. Poor, “From ofds to dd-isac: Integrating sensing and communications in the delay doppler domain,” *IEEE Wireless Communications*, vol. 31, no. 6, pp. 152–160, 2024.
- [4] Y. Wu, C. Han, and Z. Chen, “Dft-spread orthogonal time frequency space system with superimposed pilots for terahertz integrated sensing and communication,” *IEEE Transactions on Wireless Communications*, vol. 22, no. 11, pp. 7361–7376, 2023.
- [5] J. Yu, H. M. Saad, and R. M. Buehrer, “Centimeter-level indoor localization using channel state information with recurrent neural networks,” in *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pp. 1317–1323, 2020.
- [6] F. Zafari, A. Gkelias, and K. K. Leung, “A survey of indoor localization systems and technologies,” *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019.
- [7] J. A. Zhang, M. L. Rahman, K. Wu, X. Huang, Y. J. Guo, S. Chen, and J. Yuan, “Enabling joint communication and radar sensing in mobile networks—a survey,” *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 306–345, 2022.
- [8] O. A. Aghda, O. B. H. Belkacem, J. Guerreiro, N. Souto, M. Szczachor, and R. Dinis, “A&m based fractional delay and doppler channel estimation in ofds,” *Research Square*, vol. Preprint, September 2025.
- [9] Y. Lin, S. Jin, M. Matthaiou, and X. You, “Channel estimation and user localization for irs-assisted mimo-ofdm systems,” *IEEE Transactions on Wireless Communications*, vol. 21, no. 4, pp. 2320–2335, 2022.
- [10] Z. H. Shaik, E. Björnson, and E. G. Larsson, “Cell-free massive mimo with radio stripes and sequential uplink processing,” in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, 2020.
- [11] Z. H. Shaik, E. Björnson, and E. G. Larsson, “Mmse-optimal sequential processing for cell-free massive mimo with radio stripes,” *IEEE Transactions on Communications*, vol. 69, no. 11, pp. 7775–7789, 2021.

- [12] S. Wang, R. Liu, H. Zhou, J. Li, F. Liu, B. Sheng, D. Wang, and X. You, "Location-aware coordinated beamforming for high mobility cell-free mmwave systems," *IEEE Wireless Communications Letters*, vol. 13, no. 3, pp. 899–903, 2024.
- [13] H. T. Dao and S. Kim, "Effective channel gain-based access point selection in cell-free massive mimo systems," *IEEE Access*, vol. 8, pp. 108127–108132, 2020.
- [14] T. V. Chien, E. Björnson, and E. G. Larsson, "Optimal design of energy-efficient cell-free massive mimo: Joint power allocation and load balancing," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5145–5149, 2020.
- [15] A. Anand, C. R. Murthy, and R. Chopra, "Impact of mobility on downlink cell-free massive mimo systems," 2022.
- [16] E. Björnson and E. Jorswieck. 2013.
- [17] K. Kim, J. Myung, Y.-J. Ko, and W. Y. Lee, "Beamforming and power optimization for user fairness in cell-free mimo systems," *IEEE Access*, vol. 11, pp. 51032–51046, 2023.
- [18] J. Talvitie, T. Levanen, M. Koivisto, T. Ihalainen, K. Pajukoski, M. Renfors, and M. Valkama, "Positioning and location-based beamforming for high speed trains in 5g nr networks," in *2018 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–7, 2018.
- [19] M. Mizmizi, M. Brambilla, D. Tagliaferri, C. Mazzucco, M. Debbah, T. Mach, R. Simeone, S. Mandelli, V. Frascolla, R. Lombardi, M. Magarini, M. Nicoli, and U. Spagnolini, "6g v2x technologies and orchestrated sensing for autonomous driving," 2021.
- [20] A. Fouda, R. Keating, and H.-S. Cha, "Toward cm-level accuracy: Carrier phase positioning for iiot in 5g-advanced nr networks," 2022.
- [21] G. A. Dohler M., Saikali S., "The crucial role of 5g, 6g, and fiber in robotic telesurgery," *Robotic Surgery*, 2024.
- [22] W. Ding, W. Sun, Y. Gao, and J. Wu, "Carrier phase-based precise heading and pitch estimation using a low-cost GNSS receiver," *Remote Sensing*, vol. 13, no. 18, p. 3642, 2021.
- [23] 3GPP, "Study on Expanded and Improved NR Positioning (Release 18)," Tech. Rep. TR 38.859, June 2024.
- [24] N. González-Prelcic, M. Furkan Keskin, O. Kaltiokallio, M. Valkama, D. Dardari, X. Shen, Y. Shen, M. Bayraktar, and H. Wymeersch, "The integrated sensing and communication revolution for 6G: Vision, techniques, and applications," *IEEE Proceedings*, vol. 112, no. 7, pp. 676–723, 2024.
- [25] K. Berntorp, A. Weiss, and S. D. Cairano, "Integer ambiguity resolution by mixture Kalman filter for improved GNSS precision," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 4, pp. 3170–3181, 2020.
- [26] F. Ayten, M. C. Ilter, O. Kaltiokallio, J. Talvitie, A. Jain, E. S. Lohan, H. Wymeersch, and M. Valkama, "Phase-only positioning: Overcoming integer ambiguity challenge through deep learning." *IEEE PIMRC*, 2025, accepted for publication.
- [27] J. Nikonowicz, A. Mahmood, M. I. Ashraf, E. Björnson, and M. Gidlund, "Indoor positioning in 5G-Advanced: Challenges and solution toward centimeter-level accuracy with carrier phase enhancements," *IEEE Wireless Commun.*, vol. 31, no. 4, 2024.

- [28] S. Fan, W. Ni, H. Tian, Z. Huang, and R. Zeng, "Carrier phase-based synchronization and high-accuracy positioning in 5G new radio cellular networks," vol. 70, no. 1, pp. 564–577, 2022.
- [29] H. Wymeersch, R. Amiri, and G. Seco-Granados, "Fundamental performance bounds for carrier phase positioning in cellular networks," in *in proc. of Globecom*, pp. 7478–7483, Kuala Lumpur, Malaysia, 2023.
- [30] F. Ayten, M. C. Ilter, O. Kaltiokallio, J. Talvitie, A. Jain, E. S. Lohan, H. Wymeersch, and M. Valkama, "Failure tolerant phase-only indoor positioning via deep learning," in *In Proc. of PIMRC*, 2025.
- [31] E. D. D. Claudio, R. Parisi, and G. Jacovitti, "Space time music: Consistent signal subspace estimation for wideband sensor arrays," *IEEE Transactions on Signal Processing*, vol. 66, no. 10, pp. 2685–2699, 2018.
- [32] K. N. R. S. V. Prasad, E. Hossain, and V. K. Bhargava, "Machine learning methods for rss-based user positioning in distributed massive mimo," *IEEE Transactions on Wireless Communications*, vol. 17, no. 12, pp. 8402–8417, 2018.
- [33] S. Samira Moosavi and P. Fortier, "Fingerprinting localization method based on clustering and gaussian process regression in distributed massive mimo systems," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1–7, 2020.
- [34] B. Deutschmann, C. Nelson, M. Henriksson, G. Marti, A. Kosasih, N. Tervo, E. Leitinger, and F. Tufveson, "Accurate direct positioning in distributed mimo using delay-doppler channel measurements," in *2024 IEEE 25th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 606–610, 2024.
- [35] Z. Yang, H. Zhang, B. Di, X. Li, X. Hou, and L. Song, "Location-aware communication for ris-aided distributed mimo systems," *IEEE Transactions on Vehicular Technology*, vol. 74, no. 3, pp. 4445–4460, 2025.
- [36] W. Li, W. Yang, L. Yang, H. Xiong, and Y. Hui, "Bidirectional positioning assisted hybrid beamforming for massive mimo systems," *IEEE Transactions on Communications*, vol. 69, no. 5, pp. 3367–3378, 2021.
- [37] I. Yildirim, A. Koc, E. Basar, and T. Le-Ngoc, "Ris-aided angular-based hybrid beamforming design in mmwave massive mimo systems," in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, pp. 5267–5272, 2022.
- [38] X. Chu, Z. Lu, J. Kang, Y. Zou, H. Zhang, and X. Qiu, "Hybrid beamforming toward positioning enhancement under cellular mimo systems," *IEEE Transactions on Wireless Communications*, vol. 23, no. 10, pp. 13545–13561, 2024.
- [39] G. E. Garcia, N. Garcia, G. Seco-Granados, E. Karipidis, and H. Wymeersch, "Fast in-band position-aided beam selection in millimeter-wave mimo," *IEEE Access*, vol. 7, pp. 142325–142338, 2019.
- [40] J. K. Chege, A. Yeredor, and M. Haardt, "Probabilistic position-aided beam selection for mmwave mimo systems," 2025.
- [41] V. Va, T. Shimizu, G. Bansal, and R. W. Heath, "Online learning for position-aided millimeter wave beam training," *IEEE Access*, vol. 7, pp. 30507–30526, 2019.
- [42] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part I," *IEEE Robotics & Automation Magazine*, vol. 13, no. 2, pp. 99–110, 2006.

- [43] T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping (SLAM): part II," *IEEE Robotics & Automation Magazine*, vol. 13, no. 3, pp. 108–117, 2006.
- [44] T. Bailey, J. Nieto, J. Guivant, M. Stevens, and E. Nebot, "Consistency of the EKF-SLAM algorithm," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3562–3568, 2006.
- [45] G. Grisetti, R. Kümmerle, C. Stachniss, and W. Burgard, "A tutorial on graph-based SLAM," *IEEE Intelligent Transportation Systems Magazine*, vol. 2, no. 4, pp. 31–43, 2010.
- [46] A. Venus, E. Leitingner, S. Tertinek, F. Meyer, and K. Witrisal, "Graph-based simultaneous localization and bias tracking for robust positioning in obstructed los situations," in *Proc. 2022 56th Asilomar Conference on Signals, Systems, and Computers*, pp. 777–784, 2022.
- [47] A. Sapienza, K. Ossi, J. Talvetie, F. B. Slama, E. S. Lohan, and M. Valkama, "Extracting situational awareness in 6G cellular systems: Robust ekf-slam approach," in *59th Asilomar Conference on Signals, Systems, and Computers*, 2025.
- [48] K. Witrisal, P. Meissner, E. Leitingner, Y. Shen, C. Gustafson, F. Tufvesson, K. Haneda, D. Dardari, A. F. Molisch, A. Conti, *et al.*, "High-accuracy localization for assisted living: 5G systems will turn multipath channels from foe to friend," *IEEE Signal Processing Magazine*, vol. 33, no. 2, pp. 59–70, 2016.
- [49] P. Meissner, E. Leitingner, and K. Witrisal, "UWB for robust indoor tracking: Weighting of multipath components for efficient estimation," *IEEE Wireless Communications Letters*, vol. 3, no. 5, pp. 501–504, 2014.
- [50] E. Leitingner, F. Meyer, F. Hlawatsch, K. Witrisal, F. Tufvesson, and M. Z. Win, "A belief propagation algorithm for multipath-based SLAM," *IEEE Transactions on Wireless Communications*, vol. 18, no. 12, pp. 5613–5629, 2019.
- [51] Y. Ge, O. Kaltiokallio, H. Kim, J. Talvitie, S. Kim, L. Svensson, M. Valkama, and H. Wymeersch, "Mmwave mapping and SLAM for 5G and beyond," in *Integrated Sensing and Communications*, pp. 445–475, Springer, 2023.
- [52] H. Peng, Z. Zhao, and L. Wang, "A review of dynamic object filtering in slam based on 3d lidar," *Sensors*, vol. 24, no. 2, 2024.
- [53] M. Henein, G. Kennedy, V. Ila, and R. Mahony, "Simultaneous localization and mapping with dynamic rigid objects," 2018.
- [54] A. Walcott-Bryant, M. Kaess, H. Johannsson, and J. J. Leonard, "Dynamic pose graph slam: Long-term mapping in low dynamic environments," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1871–1878, 2012.
- [55] I. Miller and M. Campbell, "Rao-blackwellized particle filtering for mapping dynamic environments," in *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pp. 3862–3869, 2007.
- [56] J. G. Rogers, A. J. B. Trevor, C. Nieto-Granda, and H. I. Christensen, "Slam with expectation maximization for moveable object tracking," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2077–2082, 2010.
- [57] C.-C. Wang, C. Thorpe, and S. Thrun, "Online simultaneous localization and mapping with detection and tracking of moving objects: theory and results from a ground vehicle in crowded urban areas," in *2003 IEEE International Conference on Robotics and Automation (Cat. No.03CH37422)*, vol. 1, pp. 842–849 vol.1, 2003.

- [58] S. Fang and H. Li, "Lidar slamot based on confidence-guided data association," 2024.
- [59] S. K. Mohammed, R. Hadani, A. Chockalingam, and R. Calderbank, "OtfS—a mathematical foundation for communication and radar sensing in the delay-doppler domain," *IEEE BITS the Information Theory Magazine*, vol. 2, no. 2, pp. 36–55, 2022.
- [60] Z. Wei, H. Qu, Y. Wang, X. Yuan, H. Wu, Y. Du, K. Han, N. Zhang, and Z. Feng, "Integrated sensing and communication signals toward 5g-a and 6g: A survey," *IEEE Internet of Things Journal*, vol. 10, no. 13, pp. 11068–11092, 2023.
- [61] W. Yuan, Z. Wei, S. Li, J. Yuan, and D. W. K. Ng, "Integrated sensing and communication-assisted orthogonal time frequency space transmission for vehicular networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 6, pp. 1515–1528, 2021.
- [62] Y. Jitsumatsu, "2d sinc interpolation-based fractional delay and doppler estimation using time and frequency shifted gaussian pulses," in *2025 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, pp. 0060–0065, 2025.
- [63] K. Zhang, W. Yuan, S. Li, F. Liu, F. Gao, P. Fan, and Y. Cai, "Radar sensing via ofts signaling: A delay doppler signal processing perspective," in *ICC 2023 - IEEE International Conference on Communications*, pp. 6429–6434, 2023.
- [64] J. Zhu and J. Liang, "Low complexity parameter estimation approach for ofts-based isac system," in *2025 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, 2025.
- [65] L. Gaudio, M. Kobayashi, G. Caire, and G. Colavolpe, "On the effectiveness of ofts for joint radar parameter estimation and communication," *IEEE Transactions on Wireless Communications*, vol. 19, no. 9, pp. 5951–5965, 2020.
- [66] E. Aboutanios and B. Mulgrew, "Iterative frequency estimation by interpolation on fourier coefficients," *IEEE Transactions on Signal Processing*, vol. 53, no. 4, pp. 1237–1242, 2005.
- [67] Y. Shan, F. Wang, and Y. Hao, "Off-grid channel estimation using grid evolution in rectangular pulse-shaping ofts system," in *2023 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 295–300, 2023.
- [68] Z. Wei, W. Yuan, S. Li, J. Yuan, and D. W. K. Ng, "Off-grid channel estimation with sparse bayesian learning for ofts systems," *IEEE Transactions on Wireless Communications*, vol. 21, no. 9, pp. 7407–7426, 2022.
- [69] Y. Shan, F. Wang, Y. Hao, J. Yuan, J. Hua, and Y. Xin, "Off-grid channel estimation using grid evolution for ofts systems," *IEEE Transactions on Wireless Communications*, vol. 23, no. 8, pp. 9549–9565, 2024.
- [70] P. Raviteja, K. T. Phan, and Y. Hong, "Embedded pilot-aided channel estimation for ofts in delay-doppler channels," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4906–4917, 2019.
- [71] H. Zhang, L. Liu, C. Wang, and J. Yuan, "Near-field user localization and channel estimation for xl-mimo systems: Fundamentals, recent advances, and outlooks," *IEEE Wireless Communications*, vol. 32, no. 1, pp. 112–121, 2025.

- [72] D. Gürgünoğlu, A. Kosasih, P. Ramezani, Ö. T. Demir, E. Björnson, and G. Fodor, "Performance analysis of a 2d-MUSIC algorithm for parametric near-field channel estimation," *IEEE Wireless Communications Letters*, vol. 14, no. 5, pp. 1496–1500, 2025.
- [73] M. Trinh-Hoang, M. Viberg, and M. Pesavento, "Partial relaxation approach: An eigenvalue-based doa estimator framework," *IEEE Transactions on Signal Processing*, vol. 66, no. 23, pp. 6190–6204, 2018.
- [74] A. Bazzi and D. Slock, "Joint angle and delay estimation (jade) by partial relaxation," in *Proc. IEEE ICASSP*, pp. 1–5, 2019.
- [75] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1783–1795, 1990.
- [76] J. Vinogradova and G. Fodor, "On target detection in the presence of clutter in joint communication and sensing cellular networks," in *Proc. 16th Int. Conf. Signal Process. Commun. Syst. (ICSPCS)*, pp. 1–10, 2023.
- [77] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Commun. Surv. Tutor.*, vol. 20, no. 2, pp. 870–913, 2018.
- [78] R. Kell, "On the derivation of bistatic RCS from monostatic measurements," *Proc. IEEE*, vol. 53, no. 8, pp. 983–988, 1965.
- [79] M. I. Skolnik, ed., *Radar Handbook*. McGraw-Hill, 3rd ed., 2008.
- [80] X. Chen, Z. Feng, Z. Wei, X. Yuan, P. Zhang, J. A. Zhang, and H. Yang, "Multiple signal classification based joint communication and sensing system," *IEEE Trans. Wirel. Commun.*, vol. 22, no. 10, pp. 6504–6517, 2023.
- [81] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 33, no. 2, pp. 387–392, 1985.
- [82] M. E. Pihtili, J. Equi, O. Kaltiokallio, J. Talvitie, E. S. Lohan, E. Basar, and M. Valkama, "Clutter suppression in bistatic isac with joint angle and doppler estimation," 2025.
- [83] A. Ramos, M. F. Keskin, H. Wymeersch, S. Inca, and J. F. Monserrat, "Enhancing sensing-assisted communications in cluttered indoor environments through background subtraction," in *Proc. IEEE Wirel. Commun. Netw. Conf. (WCNC)*, pp. 1–6, 2024.
- [84] S. Khosravi, H. S. Ghadikolaei, J. Zander, and M. Petrova, "Location-aided beamforming in mobile millimeter-wave networks," in *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, pp. 1–7, 2022.
- [85] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3072–3108, 2019.
- [86] A. Nessa, B. Adhikari, F. Hussain, and X. N. Fernando, "A survey of machine learning for indoor positioning," *IEEE Access*, vol. 8, pp. 214945–214965, 2020.
- [87] C. Xiao, D. Yang, Z. Chen, and G. Tan, "3-d ble indoor localization based on denoising autoencoder," *IEEE Access*, vol. 5, pp. 12751–12760, 2017.

- [88] D. V. Le, N. Meratnia, and P. J. Havinga, “Unsupervised deep feature learning to reduce the collection of fingerprints for indoor localization using deep belief networks,” in *2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp. 1–7, 2018.
- [89] X. Ruan, D. Ren, X. Zhu, and J. Huang, “Mobile robot navigation based on deep reinforcement learning,” in *2019 Chinese Control And Decision Conference (CCDC)*, pp. 6174–6178, 2019.
- [90] J. Xie and Z. Zhang, “Position-aided beam training in mmwave noma systems,” *IEICE Transactions on Electronics*, vol. E108-C, pp. 153–159, Mar. 2025.
- [91] Y. Lu, M. Koivisto, J. Talvitie, M. Valkama, and E. S. Lohan, “Positioning-aided 3d beamforming for enhanced communications in mmwave mobile networks,” *IEEE Access*, vol. 8, pp. 55513–55525, 2020.
- [92] P. Agarwal, G. D. Tipaldi, L. Spinello, C. Stachniss, and W. Burgard, “Robust map optimization using dynamic covariance scaling,” in *Proc. 2013 IEEE International Conference on Robotics and Automation*, pp. 62–69, 2013.
- [93] P. Tichavsky, “Posterior cramer-rao bound for adaptive harmonic retrieval,” *IEEE Transactions on Signal Processing*, vol. 43, no. 5, pp. 1299–1302, 1995.
- [94] P.-Y. Lajoie, B. Ramtoula, F. Wu, and G. Beltrame, “Towards collaborative simultaneous localization and mapping: A survey of the current research landscape,” *Field Robotics*, vol. 2, pp. 971–1000, 2022.
- [95] M. Cui, Y. Gu, and W. Zhang, “Joint angle and delay estimation via partial relaxation and covariance fitting,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 12, pp. 8348–8362, 2020.