



The Integrated Sensing and Communication Revolution for 6G: Vision, Techniques, and Applications

Downloaded from: <https://research.chalmers.se>, 2025-02-06 10:09 UTC

Citation for the original published paper (version of record):

Gonzalez-Prelcic, N., Keskin, M., Kaltiokallio, O. et al (2024). The Integrated Sensing and Communication Revolution for 6G: Vision, Techniques, and Applications. *Proceedings of the IEEE*, 112(7): 676-723. <http://dx.doi.org/10.1109/JPROC.2024.3397609>

N.B. When citing this work, cite the original published paper.

© 2024 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, or reuse of any copyrighted component of this work in other works.

The Integrated Sensing and Communication Revolution for 6G: Vision, Techniques, and Applications

Nuria González-Prelcic, *Senior Member, IEEE*, Musa Furkan Keskin, *Member, IEEE*, Ossi

Kaltiokallio, *Member, IEEE*, Mikko Valkama, *Fellow, IEEE*, Davide Dardari, *Senior Member, IEEE*, Xiao Shen, Yuan Shen, *Senior Member, IEEE*, Murat Bayraktar, *Student Member, IEEE*, and Henk Wymeersch, *Fellow, IEEE*

Abstract—Future wireless networks will integrate sensing, learning and communication to provide new services beyond communication and to become more resilient. Sensors at the network infrastructure, sensors on the user equipment, and the sensing capability of the communication signal itself provide a new source of data that connects the physical and radio frequency environments. A wireless network that harnesses all these sensing data can not only enable additional sensing services, but also become more resilient to channel-dependent effects like blockage and better support adaptation in dynamic environments as networks reconfigure. In this paper, we provide a vision for integrated sensing and communication (ISAC) networks and an overview of how signal processing, optimization and machine learning techniques can be leveraged to make them a reality in the context of 6G. We also include some examples of the performance of several of these strategies when evaluated using a simulation framework based on a combination of ray tracing measurements and mathematical models that mix the digital and physical worlds.

Index Terms—integrated sensing and communications, network sensing, sensing-aided communication, radio positioning, radio SLAM, RIS-aided localization, joint monostatic sensing and

communication, near field ISAC, distributed joint sensing and communication.

I. INTRODUCTION

Early work on integrated sensing and communication targeted the development of cooperation strategies to enable spectral coexistence between communication and sensing [1]. At the same time, there was also interest in sharing hardware among radar and communication systems to reduce cost, weight and size, what motivated the initial designs of joint sensing and communication systems [2]. Nowadays, the number of potential avenues to integrating sensing and communication and their related benefits have exploded [3]. On the one hand, communication operation at higher frequencies with large arrays and bandwidths has led to waveforms and signal processing algorithms in the transceiver which are naturally well-suited for sensing. On the other hand, the diversification and sophistication of devices in the wireless network have resulted in the creation of wireless networks where the communicating devices are also sensing devices, which opens challenges related to sensor and communication data fusion. Good examples of sensing/communicating devices are the connected vehicles to be supported by cellular networks (already a use case in the 5G standard), equipped with a wide variety of sensors including cameras, radars, or LIDAR. All these technological advancements bring new opportunities for integrating sensing and communication with motivations and benefits that go beyond conventional ones. Integrated sensing and communication (ISAC) has emerged as a renewed research area that aims to develop all these opportunities by exploiting the similarities between the required hardware, the waveforms, the signal processing algorithms and the machine learning strategies to be exploited, or the sensing and communication channels, and define new applications and frontiers for the future wireless communication and sensing systems.

In the cellular industry, the 6G Roadmap elaborated by the NextG Alliance (an industry initiative to advance North American technology for future cellular networks) considers joint sensing and communication a key technology for 6G [4]. Similar considerations are being made by the European counterpart to NextG, the 6G Smart Networks and Services Industry Association (6G-IA) [5]. Moreover, the European Telecommunications Standards Institute (ETSI) has

The work by N. González-Prelcic and M. Bayraktar was supported in part by the National Science Foundation under grant no. 2147955 and is supported in part by funds from federal agency and industry partners as specified in the Resilient & Intelligent NextG Systems (RINGS) program. The work by M. F. Keskin and H. Wymeersch was supported by the Swedish Research Council (VR grant 2022-03007) and by Hexa-X-II, part of the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101095759. The work by D. Dardari was supported by the HORIZON-JU-SNS-2022 project TIMES (Grant no. 101096307) and the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on "Telecommunications of the Future" (PE00000001 - program "RESTART") cofunded by the EU. The work by O. Kaltiokallio and M. Valkama was supported in part by the Academy of Finland under the grants #352754 and #359095, and in part by the Business Finland under the project "6G-ISAC". The work by Y. Shen and X. Shen was supported by the National Natural Science Foundation of China under Grant 62271285 and Shanghai AI Laboratory.

N. González-Prelcic and M. Bayraktar are with the Department of Electrical and Computer Engineering, University of California San Diego, USA (ngprelcic@ucsd.edu;mbayraktar@ucsd.edu).

M. F. Keskin and H. Wymeersch are with the Department of Electrical Engineering, Chalmers University of Technology, Sweden (furkan@chalmers.se;henkw@chalmers.se).

O. Kaltiokallio and Mikko Valkama are with the Unit of Electrical Engineering, Tampere University, Finland (ossi.kaltiokallio@tuni.fi; mikko.valkama@tuni.fi).

D. Dardari is with the Department of Electrical, Electronic, and Information Engineering "Guglielmo Marconi", University of Bologna, and WiLAB-CNIT, Italy (davide.dardari@unibo.it).

X. Shen and Y. Shen are with the Department of Electronic Engineering and the Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084 China (shenx20@mails.tsinghua.edu.cn; shenyuan_ee@tsinghua.edu.cn).

also launched an Industry Specification Group (ISG) for ISAC, which is developing a roadmap and prioritization of sensing types and ISAC use cases that can potentially be covered in future 6G releases of the 3rd Generation Partnership Project (3GPP) [6]. The performance requirements envisioned by the industry for some application verticals are, however, stringent and can only be met with continued research that develops advanced solutions.

In this paper, we describe different frameworks for integrating sensing and communications in future-generation cellular systems, discuss the different features to be exploited at different frequency bands, and present an overview of the recent techniques and advances that can make ISAC a reality in 6G. We focus on a communication-centric perspective for ISAC with tight integration of waveform and time and frequency resources for sensing and communication, versus other approaches where integration only appears at the site or spectrum level [7]. In this communication-centric vision, we also review how sensing can assist the network operation. Previous overview/tutorial papers do not clearly focus on a communication-centric perspective of ISAC which includes a comprehensive survey of all the ISAC techniques relevant for 6G and beyond. For example, [8] describes a radar-centric approach, where only low rate communication is considered by embedding communication signals into radar waveforms. The survey in [9] discusses radar and communication waveforms for ISAC, without covering specific algorithms in the transceiver to enable the different network sensing modes. The work in [10] focuses on automotive applications, describing radar-centric and communication-centric waveforms, but not covering the detailed description of specific algorithms or learning strategies for sensing or operation modes in the near field, to name some limitations. The overview in [11] includes communication-centric and radar-centric perspectives, but the study is not comprehensive and the level of detail is very limited; for example, it does not include a treatment of positioning, and relevant aspects in monostatic sensing like the impact of self-interference are not described. Technologies for sensor-aided communication, a new form of ISAC which may play a critical role to increase resilience and adaptability in 6G networks, are not discussed in any of these surveys. Finally, [3], includes an study of performance limits and trade-offs, aspects of waveform design from a communication and radar perspective, and only a one-page description of receiver algorithms for ISAC. We can conclude that none of the previous works provides a complete communication-centric overview with emphasis on the techniques at the receiver that enable network sensing and sensor-aided communication, as this paper does. In addition, the role of machine learning tools in the design of the physical layer of ISAC networks is hardly considered in previous surveys.

The structure of the technical sections in this paper is shown in Fig. 1. We start by describing our vision of the ISAC network, potential sensing modes, the concept of sensing-aided communication, and the status of ISAC services –mainly positioning– developed within the most recent standardization efforts in 3GPP. Then we focus on ongoing research, reviewing first approaches for joint bistatic and multi-static sensing and communication that provide the sensing information directly

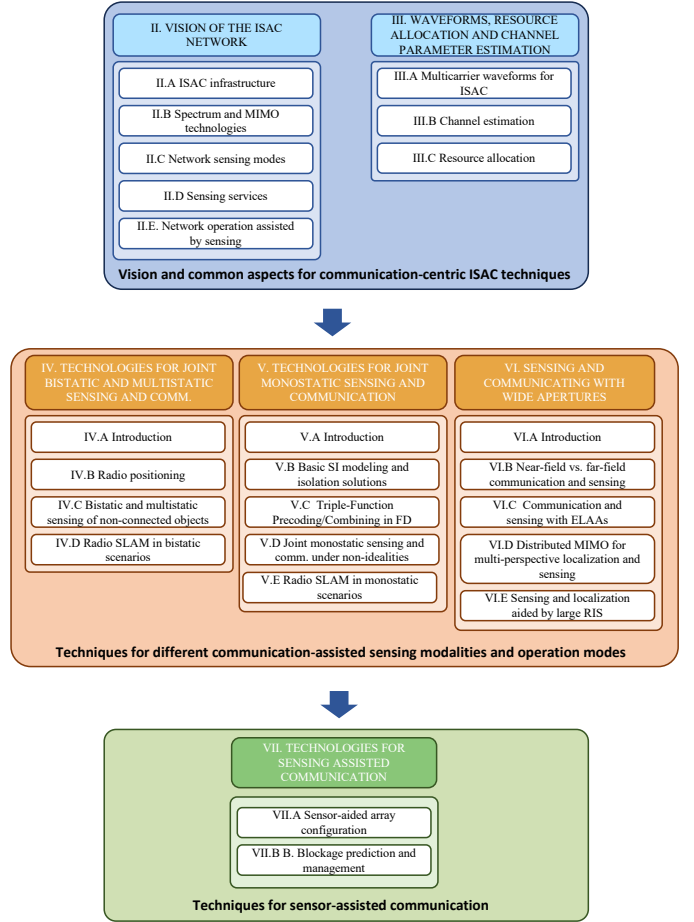


Fig. 1. Paper outline, covering the technical sections.

from the fronthaul link, downlink, uplink, or even sidelink signal, exploiting different geometric transformations between some or all the channel parameters and the user's position and orientation. These approaches are especially relevant at millimeter-wave (mmWave) and sub-THz bands, where the large arrays and bandwidths provide good angular and delay resolvability, and it is also easier to map the channel parameters to the objects in the propagation environment due to the channel sparsity. Next, we revisit joint monostatic sensing and communication strategies, where the exploitation of full duplex (FD) circuits that provide an appropriate isolation between the transmitter and receiver enables simultaneous transmission of the communication signal and reception of the reflections on potential targets, which can be processed in a radar-like operation to provide position and velocity information. In a different section, we also discuss how novel wide aperture technologies such as large reconfigurable intelligent surface (RIS) and distributed MIMO can provide a potential avenue to further increase position estimation accuracy by exploiting information from additional relevant paths. To complete the perspective of the sensing capabilities of the cellular network, we will describe several approaches for radio-simultaneous localization and mapping (SLAM), the process of simultaneously locating the user and creating a map of the environment. Finally, we will describe the opposite setting,

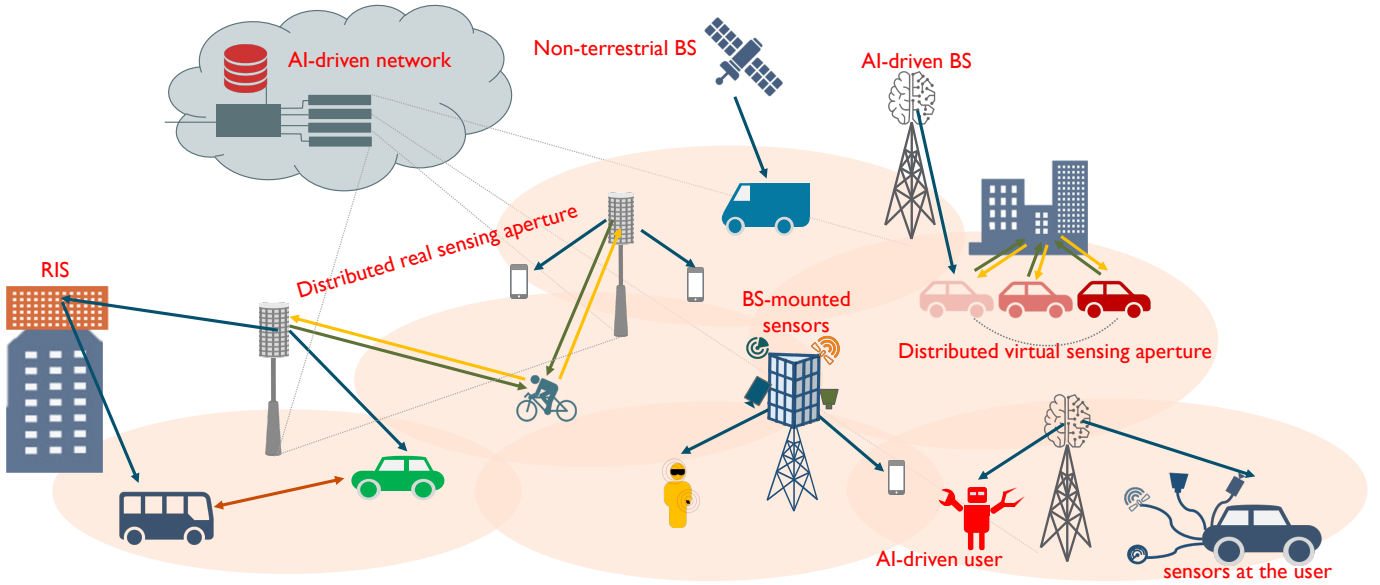


Fig. 2. Illustration of the different elements in the infrastructure of an ISAC network.

sensor-aided communication, where sensing information (e.g., user's position) can be leveraged to enhance network operation, for example significantly reducing the overhead associated with link configuration and reconfiguration or enabling early blockage detection. Throughout the paper, we will make the case that time is right for communication and sensing to be considered together, and why communication and sensing will likely be THE defining physical layer feature of 6G.

II. VISION OF THE ISAC NETWORK

A. ISAC infrastructure

The ISAC network provides an integrated combination of sensing and communication. It offers sensing and communication as services to applications that are run in, around and by the network. A smart meter may subscribe to communication as a means to send back meter measurements, a bicycle commuter may subscribe to sensing to enhance their situational awareness, while an automated vehicle may subscribe to sensing and communication as part of an automated driving package. An ISAC network, and pertinent components of infrastructure, are illustrated in Fig. 2. We summarize the different components of the infrastructure here.

The foundation of an ISAC network is a diverse and heterogeneous cellular communication infrastructure. There is conventional terrestrial base station (BS) infrastructure in the form of macro/micro/pico cells, which are typically mounted on towers, rooftops or lightpoles. To expand coverage and increase sensing accuracy and probability of target detection, important for higher frequencies like mmWave, there are low-power RIS to generate favorable reflections between the BSs and the user equipment (UE). There are also relays or BSs that connect to the cellular infrastructure using the same spectrum as the network, with integrated access and backhaul. The

infrastructure need not be terrestrial; ISAC also supports non-terrestrial components where the BSs are untethered to the ground in the form of satellites or unmanned aerial vehicles. Wired networks, typically realized via cable, fiber or point-to-point microwave links form another piece of infrastructure in the cellular network. These fixed communication links are used as part of backhaul to network BSs together with the core network, and also to implement front-haul, where a BS is realized in two pieces as a remote radio head connecting antennas to a distantly located baseband unit.

The sensing part of an ISAC network is realized with several infrastructural components. Sensors in the form of cameras, radars, and LIDARs are present at some of the UE in the network. These sensors use spectrum that is different from the communication spectrum. But the sensors need not just be on the UEs as there is a compelling case to co-locate sensing with the infrastructure to offer a birds-eye-view of the environment [12]. Sensing is also facilitated by reusing the communication signal for radio sensing. A network with integrated sensing and communication reuses the existing communication waveforms possibly with more capable hardware (for example a FD BS to enable monostatic sensing), and additional network components to fuse data collected in the environment.

The learning portion of an ISAC network is realized through the combination of data and computation. Data refers to the storage of past sensor data and communication performance data, which is collected over time. Computation refers to the training of machine learning models on the data, the updating of models based on new data, and the execution of inference operations using the trained models. As shown in Fig. 2, the data and computation are distributed in different components of the network including the UE, the BS (so-called network edge), and in the core network. The computation capabilities will vary significantly at these different components, as will the extent

of data stored and shared with other network components.

One aspect that makes ISAC networks interesting is the vastly different communication, sensing, computation and learning capabilities of the devices in the network. It includes devices that directly serve people like smart phones, watches and glasses; low-capability connected devices like smart meters and location tags; and high-capability connected devices like automotive vehicles, aerial vehicles and robots.

Fundamentally, the ISAC network infrastructure is itself smart. Many elements of the infrastructure and the network itself can be driven by protocols and algorithms which exploit both models and data. The learning (shown via data and computation in Fig. 2) should be viewed not just as a sensing service provided to devices in the network, but also as a fundamental component of the network's self-optimization capability. For example, the AI-driven BS could use data to optimize how it balances sensing and communication needs with users in its coverage area. The network central processor could use data to reconfigure how all the different infrastructure components work together to serve broader sensing and communication functions over a larger geographic area.

B. Spectrum and MIMO technologies

In addition to the infrastructure, the ability to perform sensing is strongly related to the resolution (i.e., the ability to resolve multipath) provided in different dimensions, in particular bandwidth (providing delay resolution) and array aperture (providing angle resolution). For that reason, in this section, we briefly review the different frequency bands and multiple-input multiple-output (MIMO) architectures.

1) *Spectrum Considerations*: In 6G, there are several bands under discussion. The combination of the frequency band and the available bandwidth are important from the sensing perspective.

- *Frequency range 1 (FR1)*: This band spans from around 400 MHz to 7 GHz. In this band, bandwidths between 5 MHz and 100 MHz can be supported. The main benefits of the low carrier frequency are a low path loss leading to large coverage, and small Doppler frequencies, supporting high mobility. On the other hand, the small bandwidth leads to poor delay resolution. Moreover, the propagation tends to be less geometric (i.e., the channel does not have a clear geometric relation to the environment and is more statistical in nature), due to weak shadowing and complex multipath propagation.
- *Frequency range 2 (FR2)*: This band spans from around 24 GHz to 70 GHz, with supported bandwidths ranging from 50 to 400 MHz. Due to the higher path loss, use of this band must be combined with directional arrays. This implies that while resolution is good, coverage is limited and only applications with moderate mobility can be supported. In terms of the propagation, shadowing is more pronounced leading to fewer propagation paths and a more geometric channel.
- *Upper mid-band*: This band lies between 7 GHz and 24 GHz, and is sometimes referred to as the golden band

or even FR3. This band has not been studied extensively, but is expected to provide a good trade-off between data rate and wide coverage. Initial studies claim that it is possible to maintain the same area coverage as in FR1 while achieving a significant improvement in throughput thanks to the exploitation of extremely large arrays and increased bandwidth [13].

- *Sub-THz*: The sub-THz bands spans from 100 GHz to 300 GHz. This band is envisioned for extremely high data rates in nearly static conditions. Thanks to large bandwidths and large arrays for fine, high-gain beams, resolution is expected to be high, but range is likely very short (tens of meters). The channel is characterized by diffusive, rather than specular reflections, as the wavelength gets close to the roughness of materials. This provides opportunities of sub-THz for imaging and mapping applications.

In summary, each band features clear benefits and drawbacks for sensing. Consequently, judicious selection and aggregation of different bands will be important to support the wide variety of sensing services. **Multi-band networks, with the possibility to combine or switch between a variety of bands (ranging from sub-6 GHz to THz and visible light), are promising in this respect [14], but require further study in terms of transceiver and antenna design, propagation, and resource allocation [15].** Finally, we emphasize that geometric models will be needed to evaluate sensing capabilities within and across the 6G bands, relying on ray-tracing or common data-bases, rather than conventional stochastic channel models. Such a common data-set is important not just for standardization, but also for academic research.

2) *MIMO architectures*: Multiantenna communication is a distinctive feature in current cellular networks at both FR1 and FR2. MIMO architectures are however radically different at different frequency bands, due to different hardware constraints, antenna scales, and channel bandwidths. At FR1, it is possible to operate with small arrays and one radio frequency (RF) chain per antenna element, so that all the signal processing operations are performed in digital. At mmWave frequencies, power consumption considerations and circuit technologies introduce different hardware constraints [16], [17]. For example, space limitations and excessive power consumption when operating with high resolution converters prevent from using an RF chain per antenna. This has led to specific MIMO architectures to operate at mmWave, which include analog beamforming, hybrid precoding and combining and low resolution architectures that keep one RF chain per antenna but significantly reducing the number of bits in the analog-to-digital converter (ADC)s and/or digital-to-analog converter (DAC)s. The MIMO architecture heavily impacts the received signal model and the techniques used to extract the channel parameters, later used for localization or sensing. In addition, the design of precoders and combiners for joint sensing and communication purposes also depends on the specific MIMO architecture.

In an analog architecture, beamforming is performed in the analog domain by configuring a set of phase shifters. Configuring these phases for analog beamforming requires several stages of beam training at both sides of the link. Assuming

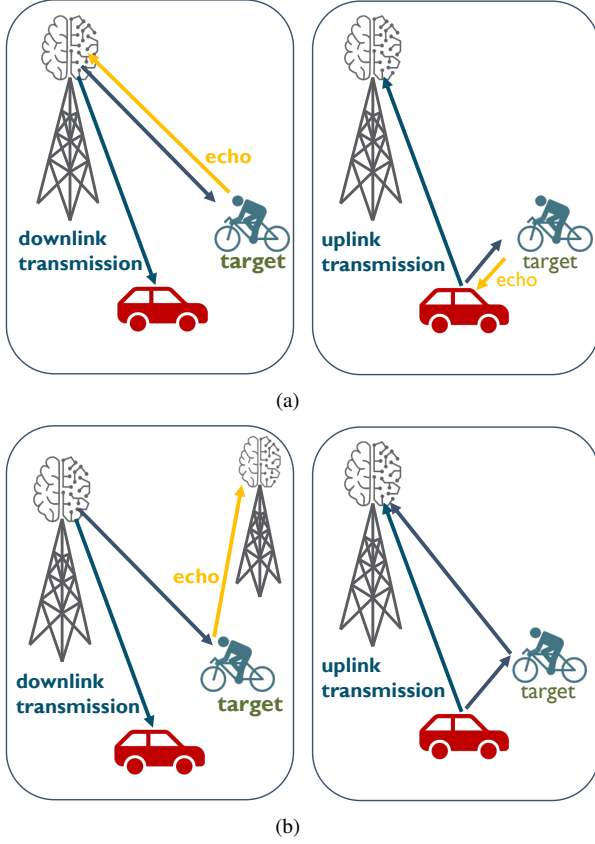


Fig. 3. Illustration of different network sensing modes. (a) Monostatic sensing: the transmitter and receiver are co-located, share a common clock, and knowledge of the transmitted signal so the echoes can be processed in radar-like operation. It can be performed by the BS or a mobile user. (b) Bistatic sensing: the transmitter and receiver are physically separated. It can be performed in the uplink or the downlink, and the transmitter and receiver can be two BSs, or one BS and one user.

a number of N_T transmit antennas and N_R receive antennas, the beamforming operation is represented by multiplication by a beamforming vector $\mathbf{f} \in \mathcal{F}^{N_T \times 1}$ at the transmitter and a weight beamforming vector $\mathbf{w} \in \mathcal{W}^{N_R \times 1}$ at the receiver, with \mathcal{F} and \mathcal{W} the sets of possible phases at the transmitter and receiver. The performance that can be achieved with analog beamforming, both for communication and sensing, is limited by the lack of amplitude tunability and the number of bits used to quantize the phases. In general, it is not a suitable approach for multistream, multiuser, or multitarget scenarios. Moreover, the spatial processing performed with analog beamforming is frequency flat, since the the same set of phase shifters are used for the entire band. The design of analog beamformers for joint sensing and communication purposes heavily depends on the sensing mode. For example, Section V provides the details of an analog beamforming design for monostatic network sensing that incorporates sensing and communication metrics in addition to the phase shifter constraints and self-interference mitigation requirements.

A hybrid precoding architecture provides an interesting performance/complexity trade-off. In this case, the precoding and combining operations are divided between the analog and digital domains, with a number of RF chains much lower

than the number of antennas. In other words, $N_{\text{RF},T} < N_T$ and $N_{\text{RF},R} < N_R$. This way, the precoding matrix can be represented by $\mathbf{F} = \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}$, where \mathbf{F}_{RF} is the analog precoding matrix and \mathbf{F}_{BB} is the digital or baseband precoding matrix. Analogously, $\mathbf{W} = \mathbf{W}_{\text{RF}}\mathbf{W}_{\text{BB}}$, with \mathbf{W}_{RF} the analog combiner and \mathbf{W}_{BB} the digital combiner. The additional digital precoding/combining stage enables frequency selective spatial processing. Moreover, the higher number of degrees of freedom in a hybrid design allows mutibeam solutions, making it suitable for multistream, multiuser, or multitarget scenarios. The sensing and communication performance provided by hybrid designs is close to that obtained with all-digital solutions. However, the hardware constraints of the analog counterparts complicate the channel parameter estimation process and the optimization problems to be solved to design the hybrid precoders and combiners.

An alternative to the analog and hybrid architectures is the low resolution architecture. It is a fully digital architecture where low resolution DACs and ADCs are employed to reduce power consumption and cost. Performance is compromised because of the high quantization noise. The investigation of specific designs and its performance in the context of ISAC systems is very limited [18], [19] and will not be further discussed in this paper.

Operation in the upper-mid-band will likely be driven by BSs equipped with extremely large arrays and a hybrid MIMO architecture with a very high number of RF chains [13]. MIMO configurations will likely vary within this band, to accommodate different channel features and bandwidths as moving from lower to higher carrier frequencies. Multiband array designs will integrate different types of antenna arrays for each band [20]. Open challenges include multiband array configuration both for communication and sensing.

C. Network sensing modes

The purpose of this section is to clarify the different types of sensing and relate them to other concepts used throughout this paper.

1) *Sensing, Positioning, Localization*: Sensing in 6G networks is a highly overloaded term. Sensing comprises receiving a radio signal or a set of radio signals and processing these radio signals to extract information relevant for a service. The received radio signals in general depend on the geometric state of the transmitter, receiver, and the environment (e.g., radar sensing), though not all sensing services rely on this geometry (e.g., pollution monitoring). Hence, localizing of connected users and passive objects relies on sensing information. In 3GPP, positioning refers to localizing UEs from uplink, downlink, or sidelink transmissions. Localization then extends positioning to also include the estimation of the position of a passive object / target (as in device-free localization), which includes also the detection of the presence of these objects. In this paper, we will use localization and positioning interchangeably, and when we refer to sensing, we intend radar-like sensing, whereby we detect objects and determine their state.

Sensing measurements and information derived from them can be fused with sensors external to 6G, such as cameras,

lidar, or radar to provide a more detailed or complementary view. For example, the cellular system could provide additional sensing information to automotive sensors, allowing vehicles to see around corners.

2) *Mono-, Bi-, and Multistatic Sensing*: Sensing is conventionally broken down into three types, though many other forms of sensing exist, which are not covered in this paper.

- *Monostatic sensing*: In monostatic sensing, illustrated in Fig. 3(a), the transmitter and receiver are co-located, share a common clock, and knowledge of the transmitted signal. Hence, sensing can be based on pilot or data signals. The sensing measurements (e.g., time-of-arrival (ToA), angle-of-arrival (AoA) and corresponding detected objects) are in relation to the coordinate system of the transmitter. Hence, the transmitter may be a BS or a mobile UE. In the former case, sensed objects are tracked, based on downlink transmissions, in the frame of reference of the static BS. In the latter case, the sensed objects are tracked, based on up/sidelink transmissions, in the frame of reference of the mobile UE. In this case, the UE can over time build a map of the environment, with respect to its original position and orientation, a process known as SLAM.
- *Bistatic sensing*: In bistatic sensing, the transmitter and receiver are physically separate, as shown in Fig. 3(b). If the transmitter or receiver is a UE, no synchronization can be assumed and sensing is based on pilot signals. Also, sensing measurements must account for the unknown location of the UE, leading to a SLAM problem in a global coordinate system. If both transmitter and receiver are BSs, time or even phase synchronization may be assumed, as well as knowledge of the transmitted data. In case such synchronization is not available, the line-of-sight (LoS) path can serve as a reference for all later multipath components.
- *Multistatic sensing*: In multistatic sensing, there are several transmitters and/or several receivers, all physically separated. Pilot signals or some form of multiplexing is needed in case there are several transmitters, to avoid interference. As in bistatic sensing, different levels of synchronization may be available (time or phase synchronization), leading to different ways to fuse measurements from the different receivers.

D. Sensing services

1) *Evolution from positioning to sensing, a standardization perspective*: Positioning a UE has been standardized over multiple 3GPP releases. Fig. 4 shows the history of 3GPP positioning Radio Access Network (RAN) standardization and evolution of foreseeable sensing standardization in 3GPP. As can be seen from the figure, specification for estimating UE's location has been built over many years. One aspect of this evolution is the change on the accuracy requirement. It started with positioning in release 9 LTE networks, with the aim of meeting the regulatory requirement of 50 m accuracy in positioning a UE. Regulatory requirement was the mainstay for building the positioning specifications in nearly all LTE

releases. Positioning signals, measurements, procedure and architecture were specified to meet this requirement. The new radio (NR) in 5G supports larger bandwidth than LTE. The increase in bandwidth in 5G NR also improved positioning accuracy requirements in 5G NR releases. The first 5G NR release had positioning requirements down to 3 m for indoor use cases which tightened further to 1 m in release 17 for industrial indoor Internet of Things (IoT) use cases [21].

As a location estimation problem, sensing seems to be an evolution of the positioning with objective of locating a passive scatterer. However, the protocol and architectural landscape of sensing will be significantly different and may not be seen as an evolution of positioning. New signals for sensing may be standardized in 6G releases of 3GPP if existing signals in specifications do not meet sensing requirements.

2) *Wireless network based use cases and requirements*: 3GPP has begun the standardization of sensing with a study on use cases [22] and subsequently building specification on the requirements [23]. Network sensing can enable new services and use cases for various verticals including smart homes, smart factories or vehicle-to-everything (V2X). There are 32 use cases proposed in [22]. In initial phases of standardization selective use cases will be prioritized. A possible prioritization of use cases can be the following,

- **Smart home/building intrusion detection** - Intrusion detection in buildings or surroundings of smart home.
- **Transport use cases** - Examples include intrusion detection of animal/human on highway, sensing aided automotive maneuvering and navigation, parking space determination or blind spot detection.
- **Industry use cases** - Detection and tracking of autonomous ground vehicles in factories, autonomous mobile robot collision avoidance, or integrated sensing and positioning in a factory hall.
- **Unmanned Aerial Vehicles (UAV) use cases** - Some examples include UAV flight trajectory tracking and UAV detection near smart grid equipment.

3) *3GPP Propagation modeling for sensing*: The current version of channel models in 38.901 does not support sensing evaluation in detail and sensing specific parameters and aspects need to be added to the channel models. For example, modelling of sensing targets in terms of their physical scattering surface as radio cross section (RCS) and modelling the mobility of the sensing targets has to be included in a new channel model that supports sensing. In addition, realizations of the current 3GPP channel model never generate reflections fulfilling the geometric relationships imposed by the laws of physics, yet these relationships are generally exploited for localization or sensing. Finally, the possibility of tracking an object requires a spatially consistent channel model depicting movement of the object consistently with respect to the evolution of multipath, phases of the signals, etc. Recent work has started to address these challenges [24], [25], but the number of contributions is still scarce.

E. Network operation assisted by sensing

Networks can sense their surroundings to provide sensing data interesting for the users, the cities, or the road infrastruc-

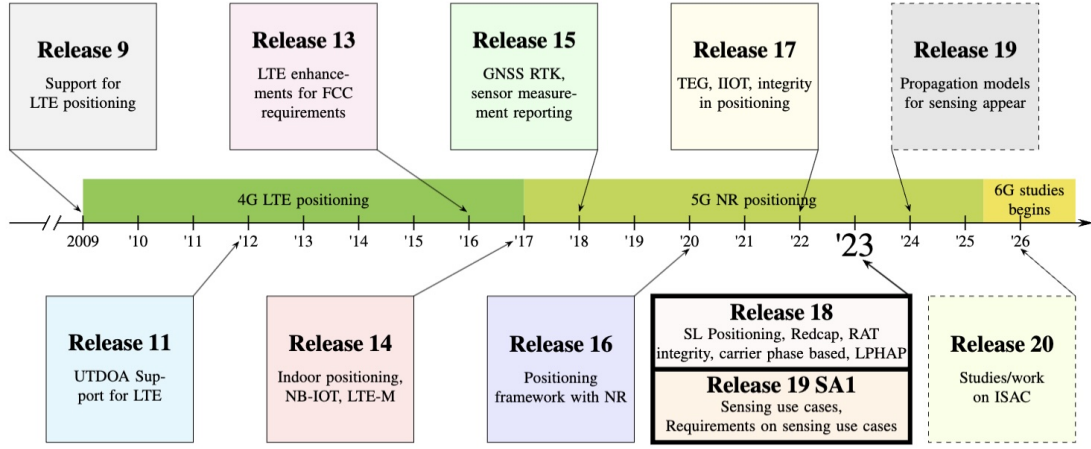


Fig. 4. Evolution of positioning across 3GPP releases and beginning of sensing. Dashed box show undecided release of 3GPP.

ture for example. Moreover, an ISAC network can also exploit these sensing data to become more resilient. A wireless network that harnesses such data can improve its adaptability to changes in the propagation environment and become more resilient to channel-dependent effects like blockage [12], [26]. The huge amount of information that this type of network can collect is the basis for exploiting machine learning (ML) algorithms to assist communication. ML can help in creating representations of the environment that fuse sensing data with digital maps and models for the communication system. Moreover, ML can also provide intelligent recommendations that exploit sensing data for network adaptation to a dynamic environment [27]–[30].

We envision a network that can create a joint map that combines measurements of both the physical world and the radio world, which we call an ISAC-map. Radio maps have been used extensively for cell network planning based on propagation simulation tools and drive testing [31]. Those maps normally only capture average received signal strength or signal-to-interference ratio (SIR) as a function of location for the purpose of network configuration and densification. The ISAC-map we envision goes well beyond the idea of radio maps, capturing the distribution of objects in the real world and their inferred properties like type, size, trajectory, and so on. This could be obtained using the radio localization and sensing capabilities of the network itself or the wealth of data that can be obtained with conventional sensors on the UE and/or on the infrastructure [12], [32], for example on lamp poles [33]. The ISAC map will include all sensing information relevant to network operation: location of users, speed and position of blockers, information about static scatterers of the environment, etc. This information can be superimposed on a digital map of the network coverage area to also leverage information about landmarks in the digital map relevant for wireless propagation. Conceptually, an ISAC map is a semantic representation of the propagation environment useful for network operation. In addition, past configurations of the network that provided good performance, for example a reduced beam codebook associated to a given location, can also be fused with the ISAC map. An example of the ISAC map for the urban scenario in Fig 5(a) is illustrated in Fig. 5(b)—including the technology used to obtain

the sensing data—while Fig. 5(c) shows some use cases that exploit the ISAC map for enhancing communication operation. For example, the location of user 1 can be used to create a beam codebook adapted to the statistical behavior of the propagation environment around that location [28]. In addition, information about the moving target 1 can be exploited to predict the blockage that user 1 experiences, so the network can proactively find an alternative path and mitigate its impact. Finally, the automotive radar signal created at user 2 can be tracked by the radar receiver deployed at the BS, and the information about the radar channel can be exploited as a prior to reduce the training overhead of establishing the link between user 2 and the BS. In Section VII, we consider in detail two particular aspects of the network operation that can greatly benefit from the exploitation of sensing data in the ISAC map: the configuration of the antenna arrays required for directional communication and blockage prediction and management.

III. WAVEFORMS, RESOURCE ALLOCATION AND CHANNEL PARAMETER ESTIMATION IN ISAC NETWORKS

A. Multicarrier waveforms for ISAC

1) *Fundamentals*: Multicarrier modulation forms the basis for the physical-layer waveform in various contemporary and emerging wireless systems. Good examples are WLAN/WiFi networks, digital video broadcasting systems, and the latest generations of mobile cellular networks, i.e., 4G LTE/LTE-Advanced and 5G NR [34]. While there are many alternative multicarrier modulation schemes [35], [36], the so-called orthogonal frequency-division multiplexing (OFDM) principle [37] is by far the most commonly adopted approach – including all the previously noted commercial systems. Powered by the involved subchannel or subcarrier structure, multicarrier modulation—and OFDM in particular—allows for efficient mitigation of channel time-dispersion in the form of computationally efficient channel equalizers. In general, OFDM enables flexible and reconfigurable physical-layer, in terms of multicarrier symbol durations while supporting also backwards compatibility and coexistence of LTE and NR. Complementary filtering and windowing [38], [39] can also be added, either at the transmitter (TX) or the receiver (RX)

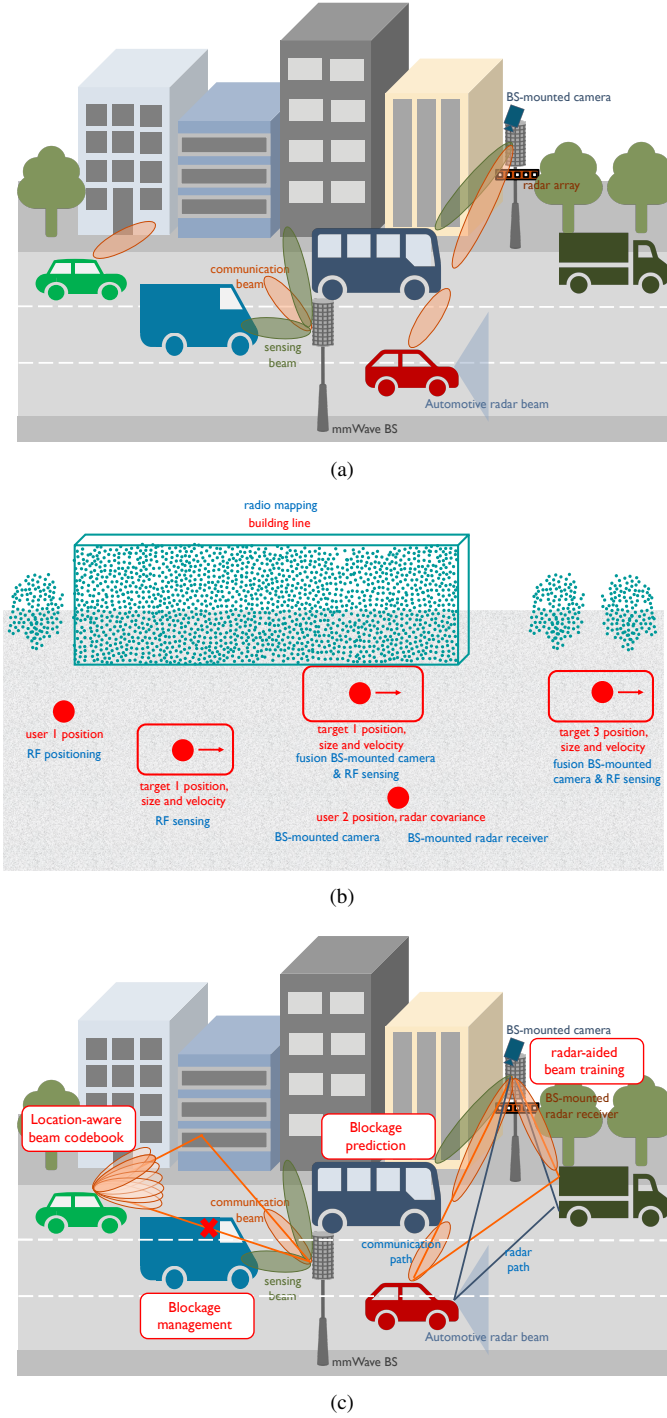


Fig. 5. Illustration of the sensing assisted communication concept: (a) Joint sensing and vehicular communication system supported by a cellular network in an urban scenario. (b) ISAC map corresponding to the urban scenario in (a), showing the detected users, targets and scatterers and the technology employed for detection. (c) Application of the ISAC map for sensing-assisted communication: location aware beam codebook design for user 1, blockage prediction created by target 2, and radar-aided beam training for user 2.

or both, in an essentially transparent manner [40], to enhance the waveform spectral containment. Additionally, OFDM and its multiple access variant called orthogonal frequency-division multiple access (OFDMA) are particularly well-suited for MIMO communications, facilitating efficient frequency-

dependent precoding or beamforming. OFDMA also allows for harnessing efficiently the channel state information (CSI) available at the transmitter, in the form of channel fading responses and interference levels, for link adaptation and scheduling in adaptive modulation and coding based multiuser systems, while rate adaptation in power-domain through, e.g., water-filling is also technically feasible.

OFDM/OFDMA has also its challenges and limitations. One particular implementation concern is related to the highly dynamic envelope of the transmit waveform, commonly quantified through the crest factor (CF) or the peak-to-average power ratio (PAPR). Such highly dynamic envelope is problematic from the power amplifiers (PAs) point of view, as the PA power-efficiency is commonly improved if operating closer towards the saturating region. Such operation point, however, also implies highly nonlinear PA behavior, thus efficient PA linearization through digital predistortion (DPD) is commonly needed – especially in cellular BSs. There exist also different precoded OFDM schemes, most notably the DFT-spread OFDM (DFT-s-OFDM), where the precoding across the subcarriers helps to reduce the PAPR—especially with contiguous spectrum allocations. Such DFT-s-OFDM approach is supported in the uplink of LTE/LTE-Advanced and NR. OFDM is also known to be sensitive to oscillator phase noise, carrier frequency offsets, and the Doppler spread of the mobile radio channel—all primarily because of the long symbol duration of the multicarrier system. These hold particularly when interpreted from the data communications and the related demodulation and decoding perspectives. An alternative multicarrier scheme called orthogonal time-frequency-space (OTFS) modulation offers increased robustness, by design, against the Doppler phenomenon [41].

When it comes to sensing and localization in the spirit of ISAC, multicarrier waveforms and MIMO-OFDM in particular are attractive for several reasons. In general, multicarrier waveforms allow for flexible injection of known reference signals in time, frequency and space, to facilitate efficient channel parameter estimation. Such is the key aspect, both from the communications receiver as well as the sensing receiver perspectives. Additionally, while the ordinary receiver implementations build commonly on OFDM symbol-wise FFT processing, extending this to two-dimensional (2D) FFT/IFFT pairs over multiple symbol durations provides basis for accurate delay/range and Doppler/velocity estimation. Such processing leads to the basic delay and Doppler resolutions of the form $\Delta\tau = N/\Delta f$ and $\Delta f_D = \Delta f/M$, respectively, where N and M refer to the transform sizes in frequency and time, respectively, while Δf refers to the subcarrier spacing. One may also straightforwardly, e.g., combine the individual range profiles obtained for the different consecutive OFDM symbols. Importantly, wider bandwidths improve the delay estimation and thereon ranging capability, while longer observation intervals in time allow for improved Doppler and thus velocity estimation.

Especially in the basic 2D-transform based implementations, the involved cyclic prefix (CP) length limits directly the sensing range such that all the involved target reflections and dominant scattering components are within the CP duration.

With 30 kHz subcarrier spacing adopted commonly in the current C-band (3.5 GHz) 5G NR networks, this still leads to target distances in the order of 350 m. However, when the networks evolve towards mmWave bands, the symbol durations and the corresponding CP lengths are reduced, thus this may become a more obvious limitations if not properly handled. Additionally, the long symbol durations of OFDM waveforms may easily lead to inter-carrier interference (ICI), calling for attention in devising OFDM-based ISAC and sensing systems. When properly handled, such phenomenon can also be turned from a foe to a friend, and described and demonstrated later in this article, in Section V.

The ambiguity function of multicarrier waveforms, measuring the capability to separate multiple coexisting targets, e.g., in range or velocity domains, is impacted by the sidelobes stemming from the FFT processing together with the involved cyclic prefix (CP). Additionally, for example the frequency-sparsity of certain known reference signals, such as the positioning reference signal (PRS) allowing for simultaneous yet orthogonal transmission from multiple nodes, may impose further ambiguity challenges. The ambiguity as well as the ultimate target parameter estimation performance can be impacted through waveform optimization, for which the subcarrier structure of MIMO-OFDM forms an excellent basis. Representative example works are for example [42]–[45]. These aspects are also discussed further in this same section, under Subsection III-C.

2) *Unified Communication/Localization/Sensing Signal Model with MIMO-OFDM*: In this part, we provide the unified MIMO-OFDM receive signal model that covers communications, localization and sensing (including both monostatic and bistatic configurations), to be employed throughout the rest of the paper. Extending the model in [46] to the case of time-varying channels, the received signal $\mathbf{y}_{n,m} \in \mathbb{C}^{N_{\text{RF},\text{R}} \times 1}$ at subcarrier n and symbol m can be written as¹

$$\mathbf{y}_{n,m} = \mathbf{W}_{\text{RF}}^H \mathbf{H}_{n,m} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}[n, m] \mathbf{x}_{n,m} + \mathbf{z}_{n,m}, \quad (1)$$

where $\mathbf{z}_{n,m}$ is the additive white Gaussian noise (AWGN), $\mathbf{x}_{n,m} \in \mathbb{C}^{N_s \times 1}$ contains the transmit symbols of N_s data streams at subcarrier n and symbol m , $\mathbf{F}_{\text{BB}}[n, m] \in \mathbb{C}^{N_{\text{RF},\text{T}} \times N_s}$ is the digital baseband precoder at subcarrier n and symbol m , $\mathbf{F}_{\text{RF}} \in \mathbb{C}^{N_{\text{T}} \times N_{\text{RF},\text{T}}}$ is the analog RF precoding matrix applied in the time domain for the entire bandwidth, $\mathbf{W}_{\text{RF}} \in \mathbb{C}^{N_{\text{R}} \times N_{\text{RF},\text{R}}}$ denotes the analog combining matrix at the RX, and $\mathbf{H}_{n,m} \in \mathbb{C}^{N_{\text{R}} \times N_{\text{T}}}$ is the channel at subcarrier n and symbol m , given by²

$$\mathbf{H}_{n,m} = \sum_{\ell=0}^{L-1} \alpha_{\ell} e^{-j2\pi n \Delta f \tau_{\ell}} e^{j2\pi m T_{\text{sym}} \nu_{\ell}} \mathbf{a}_{\text{R}}(\phi_{\ell}) \mathbf{a}_{\text{T}}^T(\theta_{\ell}). \quad (2)$$

¹The self-interference (SI) term in monostatic sensing is omitted from (1) for ease of exposition, yet it will be duly considered in (30).

²To provide a more generic channel model, it is possible to account for the impact of filters involved in pulse shaping, analog-to-digital (A/D) conversion, and matched filtering through the incorporation of complex coefficients on a per-subcarrier, per-path basis, as shown in [46, (4)]. Nevertheless, in some practical implementations, subcarriers located within the roll-off region of the combined filter in the frequency domain may be left unused. This approach ensures that the frequency response of the filter remains flat over the active subcarriers, and justifies the adoption of the simplification in (2).

In (2), $\mathbf{a}_{\text{T}}(\theta) \in \mathbb{C}^{N_{\text{T}} \times 1}$ and $\mathbf{a}_{\text{R}}(\phi) \in \mathbb{C}^{N_{\text{R}} \times 1}$ denote the array steering vectors at the TX and RX, α_{ℓ} , τ_{ℓ} , ν_{ℓ} , $\phi_{\ell} = [\phi_{\text{az},\ell}, \phi_{\text{el},\ell}]$ and $\theta_{\ell} = [\theta_{\text{az},\ell}, \theta_{\text{el},\ell}]$ denote the complex channel gain, delay (including clock offset), Doppler shift (including carrier frequency offset), AOA and AOD of the ℓ -th path/target, respectively. For localization, we assume that $\ell = 0$ indicates the LOS path, implying that α_0 involves the impact of one-way attenuation of the LOS path, while α_{ℓ} for $\ell > 0$ includes the combined attenuation of the first and second legs of the ℓ -th reflected/scattered path and the corresponding reflection/scattering coefficient. For sensing, α_{ℓ} covers the radar cross section (RCS) of the ℓ -th target and the two-way attenuation in monostatic sensing (bistatic RCS of the ℓ -th target and the combined attenuation of the first and second legs associated with the ℓ -th target, in bistatic sensing).

We note that for communications, $\mathbf{H}_{n,m}$ in (2) can be usually modeled as frequency selective yet time-invariant (i.e., not doubly-selective as in localization and sensing), since the channel coherence time is such that the impact of Doppler can be neglected [47]. For localization and sensing, high-mobility applications might necessitate even more comprehensive Doppler modeling that accounts for not only *slow-time* (i.e., inter-symbol) phase shifts represented by $e^{j2\pi m T_{\text{sym}} \nu_{\ell}}$, but also *fast-time* (i.e., intra-symbol) phase progressions. Although fast-time effects can be neglected in low- and medium-mobility scenarios (e.g., target/UE radial velocities below 30 m/s) with standard 5G NR FR2 parameters [48], they can lead to ICI in high-mobility scenarios [49] and must be taken into account explicitly (see (37) for further details).

B. Channel estimation

1) *Why channel estimation for localization or sensing is different*: Technologies for both positioning and sensing usually involve the estimation and exploitation of some or all of the multipath channel parameters described in (2). Channel estimation is more challenging, however, when the estimated parameters are used for localization or sensing. First, the required estimation accuracy is higher than that required when the only objective is the design of the communication system. For example, the precoder or combiner designs for communications based on channel estimates are relatively robust to small variations in the AoA or AoDs, while a high accuracy localization algorithm exploiting angular measurements will require very precise estimations (as an example, if we target a localization error of 1 m for a user 50 meter away from a BS, the angle estimation accuracy should be approximately 1 degree). This pushes the limits of the estimation algorithms, increasing complexity and length of the training sequence, which impacts the overall overhead of the system. Second, while for communications channel estimation is usually performed in the frequency domain without need of explicitly extracting the delays, these are key parameters for localization and sensing; moreover, many localization/sensing techniques need precise estimation of the absolute delays, which requires the consideration of an additional parameter in the estimation process, the clock offset between the TX and RX [50]. Third, for communication, it can be assumed

that the channel is not varying within the coherence time $T_c = 1/\sigma_D$, with σ_D the Doppler spread, so the doubly selective channel model in (2) can be simplified to a time invariant frequency selective channel model, where the impact of the Doppler frequencies can be neglected, i.e. $\sigma_D T_{\text{sym}} \ll 1$, with σ_D the Doppler spread [50]. In contrast, for sensing, the channel has to be observed over a longer period of time, so the Doppler shifts can also be estimated and exploited for velocity estimation. The joint estimation of these space-time-frequency parameters leads to higher computational complexity solutions than the usual channel estimator for communications-only systems. Finally, an additional element to be considered in some practical systems during the channel estimation process is the impact of pulse shaping, lowpass filtering after downconversion and matched filtering. In this case, the channel model in (2) has to be modified to introduce a time domain function which represents all the filtering stages which impact the baseband equivalent model [46]. This function contributes to the entanglement of the channel parameters and complicates its estimation. In summary, channel estimation for joint sensing and communication requires very high resolution and accuracy, which increases complexity and training overhead.

2) *Techniques for channel estimation:* Recent work on channel estimation has attempted to provide low complexity solutions, very high resolution parameter estimation, or both. Different frequency bands and system architectures lead to different features and structure in the MIMO channel matrices, and many of the channel estimation algorithms have been specifically designed to exploit particular features. Most of the techniques share, however, a common process to sound the channel. First, a number of pilot sequences are transmitted using a given number of training precoders and combiners as spatial filters, and the corresponding received sequences, which follow (1), are collected. These training precoders/combiners have to be designed to sound the channel in the spatial dimension. The collected measurements are later exploited in the estimation process to extract the multipath parameters. The estimation techniques considered in the literature on MIMO communication or MIMO joint sensing and communication can be classified into three main categories: based on maximum likelihood estimation [51]–[55], exploiting compressed sensing [46], [56]–[58], or subspace-based estimators. **Some types of techniques might be more suitable for a particular frequency band than others, as discussed in the next paragraphs.**

When operating at sub-6 GHz or with relatively small or moderate size antenna arrays, techniques that exploit the idea of maximum likelihood estimation become a solution that can provide high resolution. The conventional ML estimator is optimal [51] but results in high complexity. Alternative techniques based on expectation maximization (EM) are also effective to provide high accuracy channel estimates but their complexity is still high [52]. In contrast, the space-alternating generalized expectation maximization (SAGE) algorithm [53] and its variations—such as in [54], [55]—can provide super resolution at a moderate complexity.

Channel estimation at mmWave is more challenging than at low frequencies [16]. First, channel estimation is usually performed before array configuration. Since the precoders and

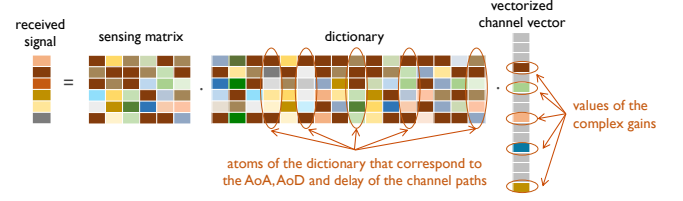


Fig. 6. Illustration of the channel estimation algorithms exploiting sparsity. The MIMO channel matrices are expanded in terms of a sparsifying dictionary and vectorized to create the sparse vector to be estimated. Identifying the AoD/AoA and delay boils down to the selection of the columns (atoms) in the dictionary which represent the sparse vector.

combiners at this stage have not been adapted to the channel yet, the directional beam patterns of the TX and RX are not aligned, and the estimation has to be performed at low or very low SNR. Second, since a hybrid MIMO architecture is commonly used at mmWave, the channel is observed through the lens of the analog combiner, without direct access to the outputs of every antenna. This way, the analog combiner acts as a compression stage for the receive signal. Finally, the large antenna arrays used at both ends of the link heavily increase the dimensionality of the channel matrices, making unfeasible many of the techniques used at lower frequencies or with smaller arrays. Literature on channel estimation at mmWave exploits the sparse nature of the channel to develop suitable solutions at this frequency band. These solutions assume a frequency selective channel model and may provide sufficient information for localization as a byproduct of communication. Different compressed sensing based techniques—including greedy sparse recovery and nuclear norm or atomic norm minimization—have been proposed in the recent literature [46], [56]–[58]. For example, greedy solutions considering frequency selective channels can operate either in the time domain or the frequency domain. In both cases, the channel estimation problem can be formulated as the recovery of a sparse vector. For the frequency domain approaches, the dictionaries are built as a Kronecker product of the array steering vectors at the transmitter and at the receiver evaluated on a grid for the angle of departure (AoD) and the angle of arrival (AoA) [57]. In the time domain approaches, the delay domain also has to be considered when building the dictionary. In this case, and assuming uniform linear array (ULA)s at both ends, the received signal for the k -th training frame can be written as [46]

$$\mathbf{y}_k = \Phi_{\text{td}}^{(k)} (\mathbf{I} \otimes \bar{\mathbf{A}}_{\text{tx}} \otimes \mathbf{A}_{\text{rx}}) \Gamma \mathbf{h}_{\text{vec}} + \mathbf{z}_k, \quad (3)$$

where \mathbf{h}_{vec} is the sparse vector containing the time domain complex channel gains after vectorization of the channel matrix, $\Phi_{\text{td}}^{(k)}$ is the sensing matrix built from the k -th training precoder, the k -th training combiner and the pilot symbols, $\bar{\mathbf{A}}_{\text{tx}}$ is the conjugate of the dictionary for the angle of departure, which contains the transmit steering vectors evaluated on a grid of potential AoDs, \mathbf{A}_{rx} is the dictionary for the angle of arrival, containing the receive steering vectors evaluated on a grid of potential AoAs, and Γ is a dictionary that represents the sparsity in the delay domain. The observation stacking all the measurements for a number K of training frames can be denoted \mathbf{y}_{td} , while the overall sensing matrix that stacks $\Phi_{\text{td}}^{(k)}$

for all k is denoted as Φ_{td} . The overall sparsifying dictionary is defined as

$$\Psi_{\text{td}} = (\mathbf{I} \otimes \bar{\mathbf{A}}_{\text{tx}} \otimes \mathbf{A}_{\text{rx}}) \Gamma. \quad (4)$$

Each column in Ψ_{td} corresponds to a given combination of AoD, AoA and delay. Estimating these parameters is equivalent to identifying the support of the sparse vectorized channel, as illustrated in Fig. 6. Finding the support and the gains in \mathbf{h}_{vec} is equivalent to solving the following problem

$$\min \|\mathbf{h}_{\text{vec}}\|_1 \quad \text{such that} \quad \|\mathbf{y}_{\text{td}} - \Phi_{\text{td}} \Psi_{\text{td}} \mathbf{h}_{\text{vec}}\|_2 \leq \epsilon, \quad (5)$$

which is the ℓ_1 relaxation of a sparse recovery problem. For both frequency and time domain estimations, highly overcomplete dictionaries for the angular and delay domains have to be exploited to achieve high resolution [46], [57], making some of the approaches proposed for communications impractical for sensing or localization. This is because a more stringent resolution requirement results in a larger dictionary, which may lead to prohibitive computational complexity or memory requirements. In this context, to reduce complexity, new greedy solutions have been recently proposed to operate with a multidimensional dictionary built as the product of independent and smaller dictionaries instead of a large dictionary based on a Kronecker product [50], [59].

Subspace-based techniques have been proposed to estimate the doubly selective MIMO channel with high resolution at different frequency bands, providing Doppler shifts information in addition to delays and angles to enable sensing applications beyond localization. For example, ESPRIT-based channel estimation [60]–[62] provides good resolution for localization and sensing at moderate complexity. The main limitation of state-of-the-art techniques based on ESPRIT for channel parameter estimation is that they can only operate when the channel model does not include any filtering effect as in (2). An alternative approach that combines the strengths of beamspace ESPRIT for angular estimation with a dictionary-based sparse recovery solution that targets delay estimation, and can operate when the channel model includes the filtering effect, has been proposed in [63].

3) *Spatial designs for channel estimation*: The training precoders and combiners used to sound the channel and build the observations for channel estimation could be directly created from the beam codebooks used in the communication network. However, an enhanced design for training can help to reduce the overhead of channel estimation or to increase the accuracy of the estimation. For example, to reduce the number of measurements when estimating the channel by exploiting a sparse recovery algorithm, it is interesting to design the training precoders and combiners—that lead to a specific Φ_{td} —so that the product $\Phi_{\text{td}} \Psi_{\text{td}}$ exhibits a low mutual coherence [64], [65]. It is also possible to consider the accuracy of the estimation as the metric that drives the design of the spatial filters. For example, the works in [66]–[68] design a new codebook for accurate angle estimation in a downlink localization scenario. The example below shows the significantly better performance provided by the new design.

Example 1. We consider a 5G/6G downlink localization scenario with the parameters $f_c = 28$ GHz, $\Delta f = 120$ kHz,

$N = 1024$ and $M = 20$. For ease of illustration, we consider a downlink multiple-input single-output (MISO) scenario with LOS-only propagation in a two-dimensional setup, where $N_{\text{T}} = 16$ and $N_{\text{R}} = 1$. In this case, using (1) and (2), the received signal at the single-antenna UE is given by

$$y_{n,m} = \alpha e^{-j2\pi n \Delta f \tau} e^{j2\pi m T_{\text{sym}} \nu} \mathbf{a}_{\text{T}}^T(\theta) \mathbf{f}_m x_{n,m} + z_{n,m}, \quad (6)$$

where $\mathbf{f}_m \in \mathbb{C}^{N_{\text{T}} \times 1}$ is the RF beamformer at the BS, with controllable amplitude and phase per antenna (i.e., analog active phased array [69]) and $z_{n,m} \sim \mathcal{CN}(0, \sigma^2)$ denotes additive noise. The BS transmits unit-amplitude pilots $x_{n,m}$ over N subcarriers and M symbols, and the UE aims to estimate the AoD θ from $y_{n,m}$ in (6). The goal herein is to design the precoder $\mathbf{F} = [\mathbf{f}_0 \dots \mathbf{f}_{M-1}] \in \mathbb{C}^{N_{\text{T}} \times M}$ that maximizes the accuracy of AoD estimation under an a-priori knowledge on θ (i.e., how to optimally allocate the pilot resources over time to achieve the highest accuracy in AoD estimation). This a-priori knowledge is quantified by an AoD uncertainty interval $\mathcal{U} = [\theta - \Delta\theta, \theta + \Delta\theta]$. We evaluate the performance of two codebooks used to construct \mathbf{F} :

- **Conventional Codebook**: conventional directional codebook employed in 5G NR mmWave systems [21], [70], given by

$$\mathbf{F}^{\text{dir}} = [\mathbf{a}_{\text{T}}(\theta_1), \dots, \mathbf{a}_{\text{T}}(\theta_{2G})]^* \quad (7)$$

- **New Codebook**: recently proposed directional/derivative codebook [66]–[68] (similar to sum/difference beams used in monopulse radar [71]) given by

$$\mathbf{F}^{\text{dir/der}} = [\mathbf{a}_{\text{T}}(\theta_1), \dots, \mathbf{a}_{\text{T}}(\theta_G), \dot{\mathbf{a}}_{\text{T}}(\theta_1), \dots, \dot{\mathbf{a}}_{\text{T}}(\theta_G)]^* \quad (8)$$

Here, $\dot{\mathbf{a}}_{\text{T}}(\theta) = \partial \mathbf{a}_{\text{T}}(\theta) / \partial \theta$, $\{\theta_g\}_{g=1}^{2G}$ represent uniformly sampled grid points from \mathcal{U} , and each column of \mathbf{F}^{dir} and $\mathbf{F}^{\text{dir/der}}$ is normalized to have unit norm.

Fig. 7 showcases the beampatterns of both directional and derivative beams. The incorporation of derivative beams $\dot{\mathbf{a}}_{\text{T}}(\theta)$ alongside standard directional beams $\mathbf{a}_{\text{T}}(\theta)$ is motivated by the need for the UE to detect subtle deviations around the intended direction θ and also supported by the theoretical CRLB analysis [66], [67]. The sharp curvature around θ in the beampattern of $\dot{\mathbf{a}}_{\text{T}}(\theta)$ ensures that slight perturbations in angle result in significant changes in amplitude, which enables precise mapping of angles based on complex amplitude measurements. Hence, high-accuracy AoD estimation and localization in 5G/6G systems can be achieved by a judicious combination of directional and derivative beams. It is worth emphasizing that these localization-optimal beams are different from those used in communications (i.e., directional beams for sweeping an angular region of interest as in (7)).

To evaluate the AoD estimation performance of the codebooks in (7) and (8), we construct \mathbf{F} in (6) by selecting its columns from these codebooks. The time sharing of the columns of \mathbf{F}^{dir} over M symbols follows a uniform strategy, while that of $\mathbf{F}^{\text{dir/der}}$ is optimized based on the CRLB criterion [66]. Fig. 8 shows the AoD root mean square error (RMSE) performances with respect to the AoD of the UE for $\Delta\theta = 1^\circ$ at $\text{SNR} = |\alpha|^2 N_{\text{T}} / \sigma^2 = 0$ dB, using the maximum-likelihood

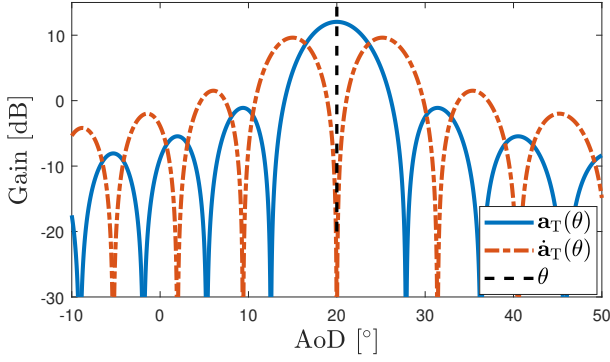


Fig. 7. 16-element ULA beampatterns of directional $\mathbf{a}_T(\theta)$ and derivative $\dot{\mathbf{a}}_T(\theta)$ beams for the UE located at an AoD $\theta = 20^\circ$ with respect to the BS. The directional beam ensures the necessary SNR for AoD estimation, whereas the derivative beam assists the UE in detecting subtle deviations from the targeted direction θ , as indicated by its pronounced curvature around θ . The combined use of directional and derivative beams allows for high-accuracy tracking of the UE in 5G/6G mmWave scenarios.

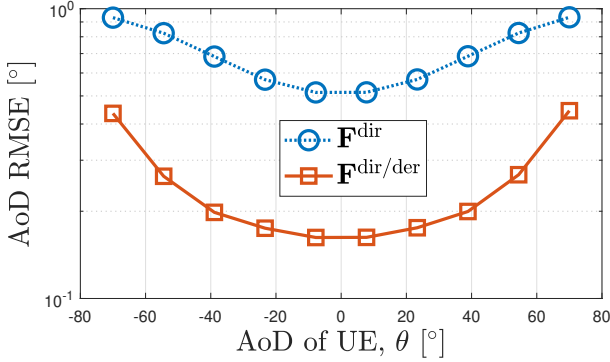


Fig. 8. AoD estimation performance with respect to the AoD of the UE, achieved by the considered codebooks in (7) and (8).

estimator [68, Eq. (11)]. We observe substantial improvements in AoD estimation accuracy with the use of $\mathbf{F}^{\text{dir/der}}$ compared to the traditional 5G codebook \mathbf{F}^{dir} , suggesting significant potential for achieving extreme location accuracy in 6G through innovative beam designs and resource allocation.

C. Resource allocation

1) *Fundamentals*: In the ISAC architecture, the sensing and communication (S&C) functions are simultaneously performed based on the unified waveforms to improve the spectrum efficiency as well as reduce the hardware costs, where the radio resources are allocated to achieve the optimal tradeoff between the S&C performance. Therefore, the design of resource allocation schemes are evaluated and guided by the performance metrics of the dual functions in ISAC systems, which are discussed in detail as follows.

The communication performance of the ISAC system is usually measured by the maximum achievable rate of reliable information transmission over the channel, i.e., the channel capacity, which can be further represented by the maximum mutual information between the unified waveform and the communication symbols, or simply the signal-to-noise-plus-interference ratio (SINR) based on the Shannon formula. As

for the sensing performance, while it can be evaluated from the detection perspective where the existence of the target is determined based on the received signals, we refer sensing to recovering the target information from the noisy measurements in this section. Then the sensing performance can be measured by the estimation error of target states, which is characterized via the Fisher information analysis in the literature [72]. Namely, let $\hat{\mathbf{s}}$ denote the unbiased estimator of the actual target state vector \mathbf{s} , which may include the position \mathbf{p} , the orientation ψ , the velocity v , and other states of interest. Then according to the information inequality, there exists

$$\mathbb{E}\{(\hat{\mathbf{s}} - \mathbf{s})(\hat{\mathbf{s}} - \mathbf{s})^T\} \succeq \mathbf{J}^{-1}(\mathbf{s}) \quad (9)$$

where $\mathbf{J}(\mathbf{s})$ denotes the Fisher information matrix (FIM) given by

$$\mathbf{J}(\mathbf{s}) = \mathbb{E}\left\{\left[\frac{\partial}{\partial \mathbf{s}} \ln f(\mathbf{y}; \mathbf{s})\right] \left[\frac{\partial}{\partial \mathbf{s}} \ln f(\mathbf{y}; \mathbf{s})\right]^T\right\}, \quad (10)$$

and the likelihood function $f(\mathbf{y}; \mathbf{s})$ is determined by the signal model (1). When the target position \mathbf{p} is concerned, the squared position error bound (SPEB) can be applied to measure the sensing performance in ISAC systems, which is given by

$$\mathcal{S}(\mathbf{p}) = \text{tr}\{\mathbf{J}_e^{-1}(\mathbf{p})\}. \quad (11)$$

The notation $\mathbf{J}_e(\mathbf{p})$ denotes the equivalent Fisher information matrix (EFIM) for \mathbf{p} , which is obtained by calculating the Schur complement of submatrix in the original Fisher information matrix (FIM) $\mathbf{J}(\mathbf{s})$ [73].

2) *Problem formulations*: Note that it remains an open challenge to establish a universal theoretical framework to jointly evaluate the performance of dual-functions in ISAC systems, which can integrate the classical results derived from information theory and estimation theory, respectively. As a consequence, the resource allocation in ISAC no longer follows a unified problem formulation as in localization networks. In particular, the wireless resources including time, frequency, space, power, and code, are allocated to solve the optimization problems with various objective functions, which can be mainly classified into three categories in the literature, i.e., the sensing-oriented formulation, the communication-oriented formulation, and the joint formulation.

- For the sensing-oriented formulations, the power, space, and other kinds of resources are allocated to optimize the objective functions derived based on sensing metrics, such as the SPEB and the detection probability. For example, the sensing-oriented joint waveform, precoding, and combining design in monostatic ISAC systems with signal model (1) can be modeled as

$$\begin{aligned} & \underset{p(\mathbf{x}), \mathbf{W}_c, \mathbf{W}_s, \mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}}{\text{minimize}} && \mathbb{E}_{\mathbf{x}}\{\mathcal{S}(\mathbf{p}; \mathbf{W}_s, \mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}})|\mathbf{x}\} \\ & \text{subject to} && \mathbf{h}_c(\mathbf{x}, \mathbf{W}_c, \mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}) \geq \gamma_c \\ & && \mathbf{f}(\mathbf{x}, \mathbf{W}_c, \mathbf{W}_s, \mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}) \leq \mathbf{0} \end{aligned} \quad (12)$$

where the objective function denotes the average SPEB conditioned on the random transmitted symbols $\mathbf{x}_{n,m}$ [74], the function \mathbf{h}_c denotes the communication constraints, e.g., the communication SINR or the sum-rate, and $p(\mathbf{x})$

denotes the distribution for transmitted symbols. The functions \mathbf{f} still denote the power and structure constraints for the symbols \mathbf{x} , the precoding matrices $\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{RB}}$ at the transmitter, and the combining matrices $\mathbf{W}_c, \mathbf{W}_s$ at the communication and sensing receivers. Note that the joint problem (12) is hard to solve due to the non-convexity and tight coupling among optimization variables. The semi-definite relaxation (SDR) and successive convex approximation (SCA) methods are applied to provide high-quality solutions with acceptable computation costs [75]–[77]. In addition, the above formulation reduce to the resource allocation in localization and sensing networks with the communication constraints \mathbf{h}_c removed, where more theoretical insights as well as efficient schemes are provided in this scenario. For example, the sparsity property of power allocation has been revealed in the localization networks, indicating that an optimal power allocation strategy requires only three anchor points to localize an agent [78]. Furthermore, the robust strategies are incorporated in both power allocation and spatial design to account for the uncertainties associated with network parameters essential for the design of resource allocation schemes [79], [80].

- In the communication-oriented resource allocation for ISAC, the communication metrics such as the mutual information and the network throughput are maximized through efficient resource management strategies including power allocation and beamforming design, i.e.,

$$\begin{aligned} & \underset{p(\mathbf{x}), \mathbf{W}_c, \mathbf{W}_s, \mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}}{\text{maximize}} && I(\mathbf{y}^c; \mathbf{x}) \\ & \text{subject to} && \mathbf{h}_s(\mathbf{x}, \mathbf{W}_s, \mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}) \leq \gamma_s \\ & && \mathbf{f}(\mathbf{x}, \mathbf{W}_c, \mathbf{W}_s, \mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}) \leq \mathbf{0} \end{aligned} \quad (13)$$

where the objective function $I(\mathbf{y}^c; \mathbf{x})$ denotes the mutual information between the received communication signals $\mathbf{y}_{n,m}^c$ and the transmitted symbols $\mathbf{x}_{n,m}$, which can be calculated by the Shannon formula based on the precoding matrices $\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}$ and the combining matrix \mathbf{W}_c . The functions \mathbf{h}_s denote the sensing constraints, e.g., the sensing accuracy or the deviation of the actual sensing beam from an ideal sensing beam pattern [81], [82]. To solve (13), auxiliary variables are introduced to decompose the original problem, after which the convex relaxation techniques can be applied to provide efficient suboptimal solutions [83], [84]. Furthermore, the two-fold tradeoff consisting of the subspace tradeoff and the deterministic-random tradeoff is revealed in terms of the communication-oriented waveform design with optimal sensing performance constraint, which provides insights for the design and analysis of practical systems [85].

- In contrast to the above discussions, the S&C requests hold equal status in the joint formulation of resource allocation for ISAC. For example, the objective function can be designed to involve both S&C performance measures, in

which sense the optimization problem can be modeled as

$$\begin{aligned} & \underset{p(\mathbf{x}), \mathbf{W}_c, \mathbf{W}_s, \mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}}{\text{maximize}} && w_c R_c + w_s R_s \\ & \text{subject to} && \mathbf{f}(\mathbf{x}, \mathbf{W}_c, \mathbf{W}_s, \mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}) \leq \mathbf{0} \end{aligned} \quad (14)$$

where the estimation rate R_s is introduced as an analog to the communication rate R_c , which measures the reduction in entropy of the target states after estimation [86], and w_c, w_s denote the weights for S&C performance, respectively. The concept termed value of service (VoS) can also be applied to design a proper objective function for (14), where the communication and sensing VoS is defined based on the classical S&C metrics, respectively, e.g., sensing CRLB and communication SINR. Then the radio resources are allocated to optimize the weighted sum of VoS from the S&C functions [87]. Additionally, the objective function can refer to the total resource consumption of the ISAC system in the joint formulation, where the S&C performance are guaranteed by certain constraints [88].

IV. TECHNOLOGIES FOR JOINT BISTATIC AND MULTISTATIC SENSING AND COMMUNICATION

A. Introduction

In this section, we will cover sensing scenarios where the transmitters and receivers are separated. Sensing where transmitter and receiver are co-located will be treated in Section V. When sensing is based on one transmitter and one receiver, it is called bistatic [89], while sensing based on several transmitters or receivers is called multistatic sensing. Since bistatic sensing can be readily implemented in communication systems, it has been covered extensively in the literature. In contrast, multistatic sensing has received rather limited treatment so-far in communications [90], [91], but is a classic topic in the radar community [92].

Bistatic sensing itself is a rich and multi-faceted field. Before we delve into the technical aspects, we first provide a brief overview of the key concepts, summarized in Fig. 9. First, the most important use of bistatic sensing is positioning, whereby a UE performs bistatic sensing with several BSs, based on which the UE location can be inferred. The positioning topic will be treated in detail in Section IV-B. Second, bistatic and multistatic sensing, which will be covered in Section IV-C involves transmitters and receivers with known locations (e.g., BSs, but possibly also UEs), to detect and localize objects in the environment, such as vehicles, pedestrians, or buildings. Third, there is the combination of sensing and positioning, known as SLAM, which will be covered in Section IV-D involves a UE determining its position, while detecting and localizing objects, based on signals to/from BSs. We note that more traditional SLAM, where the UE localizes itself and maps the environment based only on sensed backscattered signals is deferred to Section V-E.

B. Radio positioning

1) *Fundamentals of position and orientation estimation:* At its core, radio positioning aims to estimate the 3D location of

the UE in a global coordinate system, based on signals to or from one or more BSs, each of the form (1) [93], [94]. The BSs are assumed to have known positions and orientations. Positioning is thus often a two-stage process, whereby first the channel parameters, i.e., the AoA, angle-of-departure (AoD), ToA, and Doppler of the LoS path with respect to each BS are estimated, and in a second stage the UE location is estimated from the channel parameters. If the LoS path is blocked, dedicated non-line-of-sight (NLoS) detection routines can detect this phenomenon and discard the corresponding measurements [95], or the NLoS paths can be used to solve the positioning problem [96].

Positioning is generally based on dedicated pilot signals, rather than on unknown data, as this facilitates the channel parameter estimation process and provides more control to improve the resolution and accuracy (see Section III-C) [21]. Once the LoS channel parameter estimates are available, they can be related to the UE location. This relation is generally affected by nuisance parameters, e.g., the LoS ToA in (1) say τ_0 , assuming this LoS path is not blocked, is of the form

$$\tau_0 = \|\mathbf{p}_{\text{UE}} - \mathbf{p}_{\text{BS}}\|/c - \tau_{\text{bias}}, \quad (15)$$

where \mathbf{p}_{UE} is the UE position, \mathbf{p}_{BS} is a BS position, c is the speed of light and τ_{bias} is a clock bias of the UE with respect to the same BS [97]. Similarly, the Doppler measurements are affected by the carrier frequency offset (CFO) between the UE and BS, and the angles at the UE side (i.e., AoA in downlink (DL) or AoD in uplink (UL)) depend on the unknown user orientation, which is a 3D unknown. This implies that when certain measurements are used for localizing the user, the corresponding nuisance parameters must also be estimated. On the positive side, this means that there are possibilities to jointly estimate the UE location while synchronizing it to the network (thanks to the estimation of the clock bias and CFO) [98], and estimating the complete 6D UE pose [99], [100].

Mathematically, the UE positioning problem is of the form [94]

$$\mathbf{y}_{\text{meas}} = \mathbf{f}(\mathbf{x}_{\text{state}}) + \mathbf{n}, \quad (16)$$

where \mathbf{y}_{meas} comprises the estimated angles, delays, and Dopplers, $\mathbf{x}_{\text{state}}$ comprises the 3D UE location as well as any nuisance parameters (clock bias, CFO, 3D UE orientation) and $\mathbf{f}(\cdot)$ is known nonlinear mapping (e.g., containing components of the form (15), which depends on the known locations and orientations of the BSs. Recovering $\mathbf{x}_{\text{state}}$ from (16) can be done by, e.g., a least squares or maximum likelihood approach [101]. These problems are generally non-convex, due to the nonlinear relation between the measurements and the UE state, so heuristics/approximations/relaxations are employed to to find the global optimum [102]. Alternatively, prior information about the UE state can be utilized to infer $\mathbf{x}_{\text{state}}$, e.g., when applying tracking filters [103].

2) *Minimal problems*: Whereas communication to a UE in principle requires connection to only a single BS, the same does not hold for positioning, which generally needs a much larger number of connected BSs, especially when some of them may have a blocked LoS path to the UE. For that reason, understanding the minimal infrastructure needs for

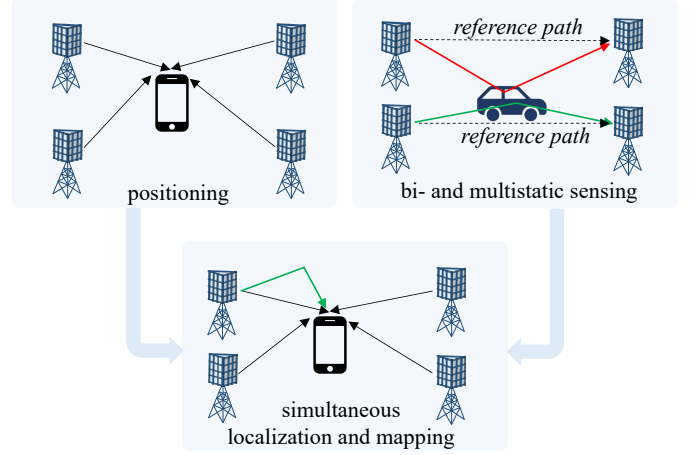


Fig. 9. Breakdown of bistatic and multistatic sensing. Operation is shown in downlink, but can equivalently occur in uplink.

positioning as well as the design of new positioning methods or technologies that can reduce the reliance on infrastructure are of great interest. These cases are called minimal problems/minimal solvers [104], in the sense that if measurements or technologies are removed, the problem can no longer be solved (e.g., in the sense of leading to an infinite number of solutions), which also facilitate outlier detection using RANSAC [105].

Example 2. Consider a scenario with 4 synchronized single-antenna BSs and a single-antenna UE, in LoS to all the BSs. This scenario is visualized in Fig. 9. Based on DL pilots, the UE estimated the ToA from each BS, which provides 4 observations of the form (15). These delay measurements are sufficient to determine the 3D UE location and the 1D clock bias.

$$\hat{\mathbf{p}}_{\text{UE}}, \hat{\tau}_{\text{bias}} = \arg \min_{\mathbf{p}_{\text{UE}}, \tau_{\text{bias}}} f(\mathbf{p}_{\text{UE}}, \tau_{\text{bias}}) \quad (17)$$

$$f(\mathbf{p}_{\text{UE}}, \tau_{\text{bias}}) = \sum_{i=1}^4 \frac{1}{2\sigma_i^2} |\hat{\tau}_{0,i} - \|\mathbf{p}_{\text{UE}} - \mathbf{p}_{\text{BS},i}\|/c + \tau_{\text{bias}}|^2 \quad (18)$$

where σ_i is the standard deviation of the LoS ToA measurement $\hat{\tau}_{0,i}$ with respect to BS i , with location $\mathbf{p}_{\text{BS},i}$. This problem can be solved iteratively from an initial guess [106]. This clock bias can also be removed by computing 3 (correlated) time-difference-of-arrival (TDoA) measurements.

From this example, the reliance of several BSs becomes apparent. This reliance can be reduced in a number of ways:

- *Additional measurements*: for instance, when augmenting ToA measurements with DL AoD measurements, the number of BS can be reduced, but at the cost of more complex multi-antenna BSs and possibly longer transmission times to support beam sweeping [21]. Such measurements are further discussed in Sections IV-B3–IV-B4. In addition, carrier phase measurements (i.e., the phase of α_ℓ) can provide extremely precise, but ambiguous location information [107], [108], as discussed in Section IV-B9.
- *Multipath exploitation*: So-far, we have considered the LoS path for positioning. While this provides the most direct

position information (see again (15)), the NLoS paths also provide information, provided they can be resolved. In particular, single-bounce NLoS paths are characterized by a single 3D incidence point. While this incidence point is unknown, the cardinality of the measurements provided by the path (e.g., delays and angles) can outweigh the unknowns, and thus improve positioning [96], [100], as will be discussed in Section IV-B4. Even when the measurements are few, multipath can be leveraged by considering the user at different time instances, see Section IV-D [109], [110].

- *New technologies*: Since positioning relies on pilot signals, there is in principle no need for using full-fledged BSs. Instead, simple beacons may be sufficient [111]. Alternatively, low-cost hardware such as RISs can be deployed to provide additional controlled multipath components [112], [113]. More on RISs in Section IV-B8.
- *New signals*: Conventional UL and DL signals can be complemented with direct links between UEs, so-called sidelinks [114]. Such links not only provide additional measurements but also support cooperative, peer-to-peer positioning. These are described in Section IV-B7.
- *New methods*: In cases where the channel is complex and the LoS cannot easily be extracted, data-driven methods can learn patterns that are beyond the realm of model-based signal processing, see Section IV-B6 [115]. Complementary to data-driven methods, advances in SLAM have provided means to perform positioning with reduced infrastructure [116], see Section IV-D.

Example 3. Consider a case with 2 BSs, each equipped with a planar array, transmitting downlink pilots to a UE. From the downlink signals, the UE can estimate the 2D-AoD from each BS, i.e., the azimuth and elevation angle. Under noise-free measurements, these angles constrain the UE to lie on the intersection of 2 lines in 3D, which is a unique point. See the left side of Fig. 10. If now the UE is equipped with a planar array as well, it can estimate the 2D-AoA from each BS, again in azimuth and elevation. This determines 2 lines in 3D away from the UE. Since the 2 lines are parameterized by 4 parameters and the UE orientation has only 3 degrees of freedom, the UE orientation can be uniquely determined. See the right side of Fig. 10.

These examples show the ability of different types of measurement to provide complementary information to the TDoA measurements described in Example 2. A set of additional examples is provided in Table I. The table provides in particular a more detailed look at the role of uncontrolled and controlled multipath.

3) *Positioning in sub-6 GHz*: Before the introduction of 5G, cellular positioning was focused exclusively on the sub-6 GHz band, so-called FR1. This band has several characteristics of relevance for positioning:

- *Limited bandwidth*: In FR1, bandwidths on the order of 5-20 MHz are available, which limits the distance resolution of delay-based measurements to around 15-60 meters [117]. Hence, if NLoS paths arrive within 15 meters of the LoS path, these paths will merge and appear

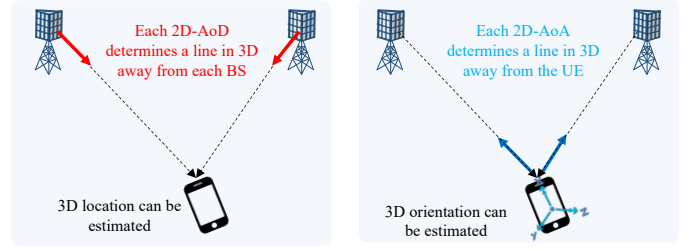


Fig. 10. Left: Example of how a UE 3D position can be estimated based on the 2D-AoD from 2 multi-antenna BSs. Right: once the 3D UE location is known, a multi-antenna UE can use the 2D-AoA to determine its 3D orientation.

Approach	3D location	3D orientation	6D pose
single-antenna BS	4 BS (TOA) and single-antenna UE	2 BS and 2D-AOA at multi-antenna UE, known location	4 BS (TOA) or 3 BS (2D-AOA at multi-antenna UE)
multi-antenna BS	2 BS (2D-AOD)	2 BS and 2D-AOA at multi-antenna UE, known location	2 BS (2D-AOD) and 2D-AOA at multi-antenna UE
multi-antenna BS with multipath exploitation	1 BS (2D-AOD, TOA), 1 IP, multi-antenna UE (2D-AOA)	1 BS (2D-AOD, TOA), 1 IP, multi-antenna UE (2D-AOA)	1 BS (2D-AOD, TOA), 1 IP, multi-antenna UE (2D-AOA)
single-antenna BS with RIS	1 BS (TOA) and 1 RIS (TOA and 2D-AOD)	1 BS (2D-AOA) and 1 RIS (2D-AOA)	1 BS (TOA, 2D-AOA) and 1 RIS (TOA, 2D-AOA)

TABLE I
EXAMPLES OF MINIMAL CONFIGURATIONS NEEDED TO SOLVE THE DL 3D LOCALIZATION PROBLEM WITHOUT A PRIORI KNOWLEDGE OF INCIDENCE POINTS (IPs) LOCATIONS.

as one path to standard signal processing methods. This means that delay-based positioning is expected to be poor in cluttered environments (on the order of tens of meters error).

- *Limited array sizes*: At the UE side, arrays are generally very small, which means that there is limited angle resolution and thus no possibility to accurately estimate the UE orientation. At the BSs, since the introduction of massive MIMO in 5G, larger arrays have been considered with on the order of 64-128 antenna elements. These can provide some amount of angle resolution, provided paths are well separated in the angle domain, as seen from the BS [118].
- *Rich channel*: The challenge of limited resolution in delay and angle is further exacerbated by the richness of the channel [119]. This means that the channel matrix $\mathbf{H}_{n,m}$ comprises many clusters of paths, coming from many directions, and these clusters may be affected by shadowing, diffraction, multi-bounce reflection, and scattering. From a communication perspective, these effects are combined in statistical models, giving rise to Rayleigh, Rician, or Nakagami fading, which have only a weak relation to the underlying geometry. From a positioning perspective, such models are questionable, not only because they mask the relation to the geometry, but also because they cannot capture the site-specific nature of the channel, which is of direct importance to positioning

[120].

Due to these characteristics in FR1, conventional methods, e.g., based on FFTs or correlations often perform relatively poorly. To overcome this poor performance, two directions have been pursued: the first is based on machine learning (see Section IV-B6), e.g., in the form of fingerprinting, where the richness of the channel is considered a benefit [115]. Such methods can bring down location errors below the 10-meter mark, but come at a cost of training complexity, as labeled training data ([fingerprint, location] pairs) must be collected. The second track is based on super-resolution methods [121]. These methods are based on the principle that even if paths differ only to a very small extent, they can be resolved if the signal-to-noise ratio (SNR) is sufficiently high.

4) *Positioning in mmWave*: At mmWave bands, the situation is significantly easier from a positioning perspective. Let us reconsider the characteristics from FR1 and evaluate them from the FR2 (24-70 GHz) and sub-THz (100-300 GHz) perspective [122], [123].

- *Large bandwidth*: At FR2 significantly large bandwidths are available, up to 400 MHz, which corresponds to a distance resolution of less than one meter. Hence, delay-based positioning becomes possible in complex and relatively cluttered environments. At sub-THz, the trend is expected to continue, with bandwidths on the order of 1 or more GHz becoming available, with corresponding distance resolution below 30 cm. On the other hand, the ability to provide better accuracy also means that synchronization requirements among BSs become more strict, and thus more challenging.
- *Large arrays (normalized to the wavelength)*: At the UE side, even modest arrays of 16-32 elements can provide good angle resolution, not only providing a path towards orientation estimation but also providing an additional dimension to improve resolution. At the BS, arrays with 64 or more elements are not exceptions, supporting superior angle resolution at both ends of the link. Note that for a given physical footprint about 25 times more elements can be packed at FR2 compared to FR1 [124]. These large arrays come at a cost of shifting from digital arrays at FR1 to analog or hybrid arrays in FR2, and even simpler arrays-of-subarrays in the sub-THz regime [123]. These arrays not only constrain the signals that can be transmitted (thus affecting AoD performance) but also imply that AoA estimation should be performed in a lower-dimensional beamspace. In addition, due to hardware impairments and lack of per-device calibration, the generated beams (precoders for transmission and combiners for reception) may deviate significantly from the idealized designs. While this has little or no impact on communication (as long as the beam has a main lobe more or less in the correct direction), this precludes the use of sophisticated methods that rely on knowledge of the complex beam responses [125], [126].
- *Sparse channel*: Further complementing the large bandwidths and large array sizes are the favorable characteristics of the channel. At FR2, the channel becomes sparse,

with few clusters surviving the propagation between transmitter and receiver, in part because shadowing is so severe [127]. Multi-bounce reflections become rarer, but due to the reduced wavelength, objects appear more rough, leading to increased diffuse scattering and fewer reflections. Overall, the sparsity of the channel is beneficial since there will be a reduced requirement for multipath resolvability. At sub-THz, these effects are even more pronounced, leading to an even sparser channel, but much more sensitive to blockages, e.g., even due to foliage (at FR2) or rainy weather (at sub-THz) [128].

The compound effect of these characteristics makes positioning at mmWave attractive in support of challenging use cases, such as in the automotive industry. Due to these same factors, relatively low-complexity methods can be employed, e.g., based on FFTs, which facilitates real-time implementation. Another important consideration at mmWave is that due to the sparse channels and the need to form narrow beams to achieve sufficient SNR, (i) communication and positioning are more closely intertwined, and (ii) it is hard for a UE to connect to several mmWave BSs simultaneously. The first consideration has given rise to the concept of location-aided or context-aware communication, the most prolific example of which is location-based beam training. The second consideration is more serious and relates closely to the discussions on minimal problems in Section IV-B2. This consideration also gives rise to the topic of single-BS positioning, which also avoids the need for inter-BS synchronization [96], [99], [100].

Example 4 (Single BS Positioning). Consider a 2D scenario shown in Fig. 11, with a BS that defines the coordinate system, a UE with unknown 2D location \mathbf{p}_{UE} , 1D clock bias τ_{bias} , and 1D orientation o_{UE} , and a scatter point with unknown 2D location \mathbf{p}_{SP} . From uplink signals, the BS determines the estimates of the AoA for the LoS path (say ϕ_0), the AoA of the reflected path (say ϕ_1), the corresponding estimates of the delays (τ_0 and τ_1), as well as the corresponding estimates of the AoD (θ_0 and θ_1). The example is visualized in Fig. 11. We see immediately that $\phi_0 = o_{UE} + \theta_0 + \pi$, from which the UE orientation o_{UE} can be solved. We also immediately find that the angle $\psi = \pi - (\phi_1 - \phi_0) - (\theta_1 - \theta_0)$. All remaining angles follow from the law of sines.

A direct closed-form solution can then be obtained as follows. Let us introduce unit vectors at the BS (\mathbf{u}_i from the AoA) and at the UE (\mathbf{v}_i from the AoD) so that the RX position can be defined as

$$\mathbf{p}_{UE} = \mathbf{p}_{BS} + d_i \gamma_i \mathbf{u}_i - d_i (1 - \gamma_i) \mathbf{v}_i \quad (19)$$

where $d_i = c(\tau_i - \tau_{bias})$ denotes the propagation distance and $\gamma_i \in [0, 1]$ represents the fraction of the propagation distance along \mathbf{u}_i . We can re-arrange (19) as

$$\mathbf{p}_{UE} - c\tau_{bias}\mathbf{v}_i = \boldsymbol{\mu}_i + \gamma_i d_i \mathbf{v}_i \quad (20)$$

where $\boldsymbol{\mu}_i = \mathbf{p}_{BS} - c\tau_i \mathbf{v}_i$ and $\mathbf{v}_i = \mathbf{u}_i + \mathbf{v}_i$. Next, we solve for γ_i and substitute it back to (20) which yields the following

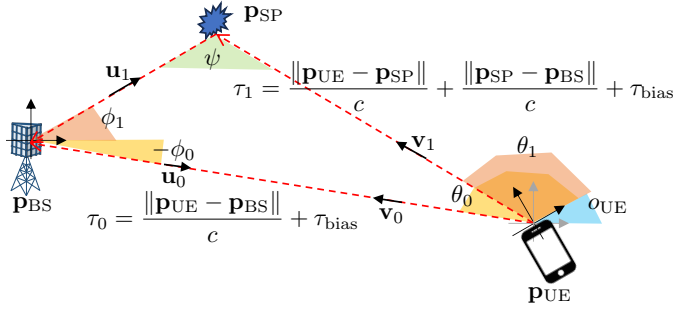


Fig. 11. Illustrative example of mmWave positioning with 1 BS by exploiting the uncontrolled multipath, via a scatter point with unknown location.

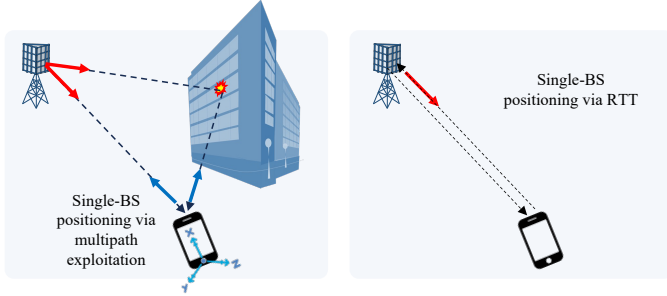


Fig. 12. Visualization of single-BS positioning approaches.

cost function

$$f(\mathbf{p}_{UE}, \tau_{bias}) = \sum_{i=0}^1 \|\mathbf{H}_i \mathbf{x}_{UE} - \boldsymbol{\mu}_i - \bar{\mathbf{v}}_i^\top (\mathbf{H}_i \mathbf{x}_{UE} - \boldsymbol{\mu}_i) \bar{\mathbf{v}}_i\|^2, \quad (21)$$

where $\mathbf{H}_i = [\mathbf{I}, -c\mathbf{v}_i]$, $\mathbf{x}_{UE} = [\mathbf{p}_{UE}^\top, \tau_{bias}]^\top$ and $\bar{\mathbf{v}}_i = \mathbf{v}_i / \|\mathbf{v}_i\|$. Now the closed-form solution can be obtained by setting the cost function's gradient to zero and solving for \mathbf{x}_{UE} . It is important to note that for the LOS path ($i = 0$), we have $\mathbf{u}_0 = -\mathbf{v}_0$ and $\bar{\mathbf{v}}_0 = \mathbf{0}$. This estimate will be affected by measurement noise, but can be utilized as a coarse estimate for further refinement, e.g., based on maximum likelihood.

A non-exhaustive list of examples of single-BS positioning is provided in Table II, one based on controlled multipath (from a RIS), one based on uncontrolled multipath, and one without any multipath. Two of the approaches are also visualized in Fig. 12. Note that methods that rely on multipath (with or without RIS) are limited by the strength of the multipath, while methods that rely on angle information suffer from large positioning errors when the UE is far away from the multi-antenna BS or RIS. Note also that the examples in Table II consider far-field propagation. As we will see in Section VI, the reliance on infrastructure can be further reduced when harnessing wavefront curvature.

5) *Positioning in sub-6 GHz vs. mmWave: performance analysis via ray-tracing data:* In this part, we carry out a comparative performance analysis of positioning in sub-6 GHz and mmWave bands using realistic ray-tracing data obtained through the REMCOM Wireless InSite@ray-tracer [129]. In the ray-tracer, we consider an urban intersection scenario involving (i) a mobile UE (corresponding to a vehicle) crossing

Scenario	Measurements	Principle	Nuisance parameters
single-antenna BS + RIS + single-antenna UE	ToA from BS, ToA and 2D-AoD from RIS	2D-AoD determines a line away from the RIS, which intersects with hyperboloid from the TDoA.	UE clock bias can be estimated from the ToA.
multi-antenna BS + single-bounce IP + multi-antenna UE	ToA and 2D-AoD from BS to UE and IP, 2D-AoD at UE from BS and IP	angle measurements determine entire geometry up to scaling, TDoA determines scaling	UE clock bias, UE 3D orientation, and IP location can be recovered.
multi-antenna BS + single-antenna UE	round-trip-time (RTT) at BS, 2D-AoA at BS	UE on intersection of sphere from RTT and line away from BS from 2D-AoA.	N/A

TABLE II

EXAMPLES OF MINIMAL CONFIGURATIONS FOR SINGLE-BS POSITIONING.

the intersection, (ii) a BS (corresponding to a road-side unit (RSU) in vehicular settings [130]) located at the center of the intersection, and (iii) four buildings with 30 m in height, located at the corners of the intersection. The UE (with antenna height 1.5 m) moves on a straight line starting from -70 m and ending at 70 m, while the BS (with antenna height 10 m) is located at 0 m (please see [131, Sec. V-B] for further details on the simulation environment). At each scenario instance (101 in total), the output of the ray-tracer consists of channel gains, delays, AOAs and AODs of the paths between the BS and the UE. To evaluate the positioning performance at sub-6 GHz and mmWave, we utilize the ray-tracer output to generate the channel matrix in (2) and the received signal in (1). For fair comparison between sub-6 GHz and mmWave, we keep the same physical aperture size at the BS, resulting in more ULA elements at mmWave.

We consider a single-BS positioning scenario via an RTT approach, as illustrated in Fig. 12, where the goal is to extract the parameters of the LOS path (i.e., the delay τ_0 and the azimuth/elevation AoAs $\phi_{az,0}$, $\phi_{el,0}$) using (1) and estimate the UE position assuming known UE height. For channel estimation from (1), we employ two algorithms [131, Sec. IV-B]: (i) a matched filtering (MF) based method that performs correlation processing across frequency and spatial domains, and (ii) an ESPRIT-based super-resolution method.

We first investigate power delay profiles (PDPs) at sub-6 GHz and mmWave to provide an illustration of channel characteristics discussed in Sec. IV-B3 and Sec. IV-B4. Fig. 13 shows the path gains with respect to range at two distinct scenario instances, representing a rich and a sparse scattering environment (which arises from the presence or absence of large reflectors, such as buildings, between the BS and the UE). As expected, higher path gains are observed at sub-6 GHz. Furthermore, when the UE is located at a considerable distance from the BS, a more diverse multipath channel is formed due to reflections from surrounding buildings.

Fig. 14 demonstrates the positioning performances using the considered algorithms at sub-6 GHz and mmWave, considering two different bandwidths at mmWave. It is observed that

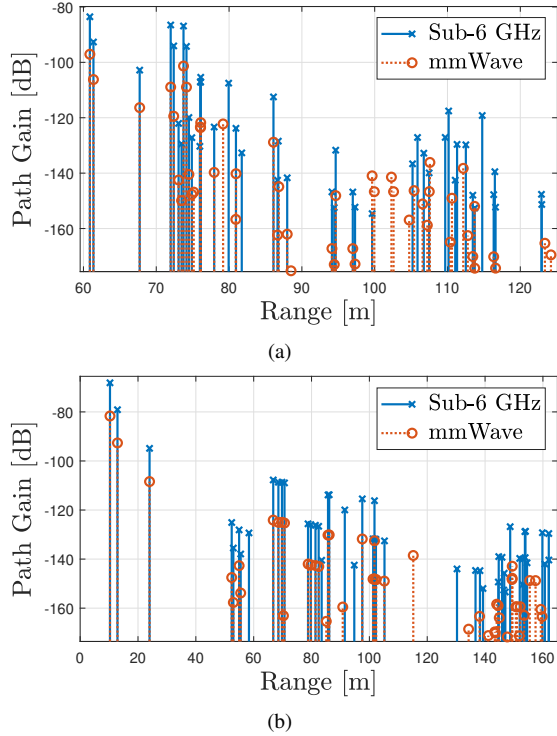


Fig. 13. Power delay profiles at sub-6 GHz and mmWave, obtained from the ray-tracing data at two different instances of the single-BS positioning scenario in an urban intersection: (a) the UE located at -60 m, leading to a rich multipath environment due to surrounding buildings, (b) the UE located at 5 m, leading to a small number of dominant paths.

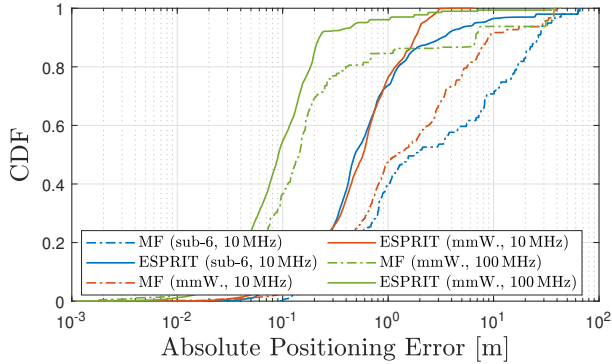


Fig. 14. Positioning performances of the MF and ESPRIT algorithms at sub-6 GHz and mmWave under different bandwidths, evaluated using the ray-tracing data. The parameters at sub-6 GHz are $f_c = 5.9$ GHz, $\Delta f = 60$ kHz, $M = 12$ and the BS array configuration 1×2 (2.54 cm), while the parameters at mmWave are $f_c = 28$ GHz, $\Delta f = 120$ kHz, $M = 24$ and the BS array configuration 1×6 (2.68 cm). The common parameters are set as follows: transmit power 10 dBm, the UE array configuration 1×1 , the noise PSD -174 dBm/Hz and the noise figure 8 dB.

ESPRIT significantly outperforms MF at both frequency bands and using different bandwidths, through its ability to resolve closely spaced paths, especially in dense multipath environments illustrated in Fig. 13. Moreover, under the same bandwidth utilization, both algorithms perform better at mmWave than at sub-6 GHz (despite larger path loss at mmWave). This can be attributed to the use of electrically large arrays at mmWave in the same physical footprint, leading to higher angular resolution and improved path resolvability. Increasing the bandwidth further enhances the performance,

which indicates the suitability and attractiveness of mmWave bands for positioning (in the sense of manifestation of the geometric nature of the channel via the use of electrically large arrays and access to large bandwidths). Therefore, in alignment with the explanations in Sec. IV-B3 and Sec. IV-B4, two key take-aways can be deduced from the ray-tracing based simulation results: (i) mmWave induces favorable channel characteristics for positioning, (ii) MF or correlation based processing suffers from poor resolution, leading to more than an order-of-magnitude degradation in accuracy compared to super-resolution approaches.

6) *Machine learning for positioning*: Positioning exploiting channel parameters and the geometry of the environment suffers from performance degradation in all types of scenarios and frequency bands. At sub-6 GHz frequencies, non-line of sight (NLOS) multipath acts as an interference that degrades the geometric localization performance in both indoor and outdoor urban environments. At mmWave frequencies, strategies like trilateration and triangulation are often not practical, since links to several BSs are required. Although at mmWave it is possible to estimate the position from the CSI of a single BS-UE link by exploiting the sparsity of the channel and the large arrays and bandwidths, high-accuracy localization requires very accurate channel estimates, which come at the cost of high pilot transmission overhead and high complexity. In addition, the sensitivity of geometric localization to channel estimation errors and impairments such as phase noise, residual clock offsets, array calibration errors, or beam squint may reduce the accuracy that can be achieved in practice.

CSI-fingerprinting positioning is an alternative to geometric approaches that can provide enhanced performance in LOS and NLOS scenarios [132]–[134]. Although it was initially proposed for indoor environments and WLANs [135], it has been successfully extended to outdoor scenarios and cellular networks [136]. It requires of an offline phase to create a database of fingerprints (one or more parameters associated to the propagation channel such as RSS or full CSI) and their associated locations. In conventional fingerprinting, the online phase is used to compare the fingerprint obtained in real time with the stored ones and infer the location exploiting algorithms such as K nearest neighbors [137], Horus [138] or RADAR [135] for example. Mathematically, this approach can be written as

$$g : \mathbf{F} \rightarrow \mathbf{p}_{\text{UE}} \quad (22)$$

where g is a mapping operation from the fingerprint \mathbf{F} to the user position \mathbf{p}_{UE} . The drawbacks of conventional fingerprinting come from (i) the requirement of permanently storing a large database and updating it periodically as the environment changes [139] (ii) the complexity of searching the whole database for every new position to be predicted.

ML-based fingerprinting can overcome these limitations. A large database will be collected for training the network offline, but it does not need to be stored or searched during online operation. During training, the network will learn the mapping g in (22), while in the online phase it will perform a regression operation to compute the position given the fingerprint. RSS is the most commonly used fingerprint in

conventional designs, but it can only capture coarse channel information, it is highly dependent on the device and it suffers from a high variability due to multipath. Recent works on ML-based fingerprinting for massive MIMO at sub-6GHz and mmWave MIMO exploit richer fingerprints based on the CSI which provide enhanced performance: the angle-delay domain channel power matrix [140], [141], a weighted average of CSI values over multiple antennas [142], the CSI per subcarrier [143], [144] or a decimated delay-domain CSI representation followed by autocorrelation to extract features invariant to the system impairments [145] to name a few. Alternatively, other recent works avoid the design of specific features by introducing the full CSI in time or frequency as the input to the deep network. This way, the first stage of the network itself extracts a suitable feature using the attention mechanisms in Transformer networks [146]–[148].

Most of the deep networks that have been designed for fingerprinting-based positioning approach the problem as a regression task. Designs based on convolutional neural network (CNN) architectures leverage image-like inputs and exploit convolutional layers' ability to extract features and relationships among adjacent data points [141], [143], [144], [149]–[154]. Long short-term memory (LSTM) networks have been also proposed to explore the correlation of CSI at different subcarriers [155]. To tackle the problem of outdated network weights in a dynamic environment, it is possible to use transfer learning [156], which enables the reconstruction of the fingerprinting database using the outdated fingerprints and a small number of new measurements. Hybrid approaches [100], [146] that combine model-based geometric localization and deep networks exploiting site-specific data can provide very high accuracy even with small training datasets, as shown in Example 4.

Position tracking can also be implemented exploiting ML. An obvious extension to the previously described approaches is to include a tracking stage that uses some kind of Bayesian filter (Kalman or extended Kalman for example) on the results of a ML-based fingerprinting approach [157]. However, recent designs replace the Bayesian filtering stage by a deep network specialized in tracking, so there is no need to build a mathematical evolution model—which may not hold in practice—to be exploited by the Bayesian filter. LSTM networks are common choices to implement tracking [158], but newer designs exploiting transformers and their attention mechanisms exhibit enhanced performance. For example, [147] proposes V-ChATNet for position tracking, an attention network that exploits the series of previous channel and position estimates to build the estimation for the new position. This design keeps the location error below 20cm for 95% of the time when evaluated with realistic vehicular channels generated by ray tracing.

Example 5 (ML-based positioning). *Consider a mmWave vehicular communication system operating with a hybrid MIMO architecture and UPAs at both ends. During initial access, the frequency selective mmWave channel between a vehicle and a single BS is estimated using MOMP [50] and some training symbols. The hybrid data/model-driven positioning system proposed in [100], [146] localizes the vehicles on the road exploiting three stages: 1) PathNet, a fully connected*

network that classifies the estimated channel paths as LOS, first NLOS or higher order; 2) a geometric localization algorithm that accounts for the clock offset and exploits the parameters of the LOS and first order paths extracted by PathNet to obtain an initial position estimation; and 3) ChanFormer, a Transformer network that exploits the concept of “attention” to evaluate which estimated paths are more credible and formulates the problem of position refinement as a classification task, by computing the probability of corresponding to the true location for a grid of points built around the initial position estimation. As illustrated in Fig. 15(a), the system achieves sub-meter accuracy localization for 95% of the users in channels with a LoS path when evaluated with ray tracing channels. In addition, the system provides sub-meter accuracy localization for 50% of the users in NLoS channels. During the tracking stage, a different attention network, V-ChATNet [147], further reduces the localization error, achieving 20cm accuracy for 95% of the users in a combination of LOS and NLOS channels, significantly outperforming Kalman Filtering based tracking, as shown in Fig. 15(b).

7) *Sidelink positioning*: Sidelink positioning is a technique which exploits sidelinks between UEs to transmit signals and perform measurements to enhance the positioning performance for V2X.

During the sidelink positioning process, the target UE aims to determine its own position under the assistance of the anchor UEs, which is further divided into three stages: the configuration stage, the signal transmission and measurement stage, and the position calculation stage [159]. First in the configuration stage, the target UE and anchor UEs are scheduled to form the sidelink positioning group for the following operations. Then in the signal transmission and measurement stage, the measurements are acquired by sending the positioning reference signals over the sidelinks between UEs. Finally in the position calculation stage, the ToA and AoA information obtained from link-level measurements are fused to calculate the absolute or relative position of the target UE, which can be performed either at the network location server or at the UE itself.

As sidelink positioning emerges as a promising technique to complement the traditional methods, e.g., global navigation satellite system (GNSS), in V2X, there have been extensive research on this area. In [114], the authors describe basic system architectures and key technologies for high-accuracy sidelink positioning. In [160], the authors present analysis of V2X sidelink positioning in sub-6 GHz, where a novel performance bound is derived to predict the positioning performance in the presence of severe multipath. In addition, the sidelinks in V2X are exploited for near-field localization in [161], where the fundamental positioning limits are determined in order to assess the possibility of the proposed scheme.

Furthermore, sidelink positioning can be viewed as the special case of the general cooperative localization paradigm, in which the unknown positions of multiple UEs are jointly inferred from the network measurements, and the performance gain comes from the exploitation of the relative position information provided by the UE-UE measurement links. Compared with conventional localization techniques, cooperative

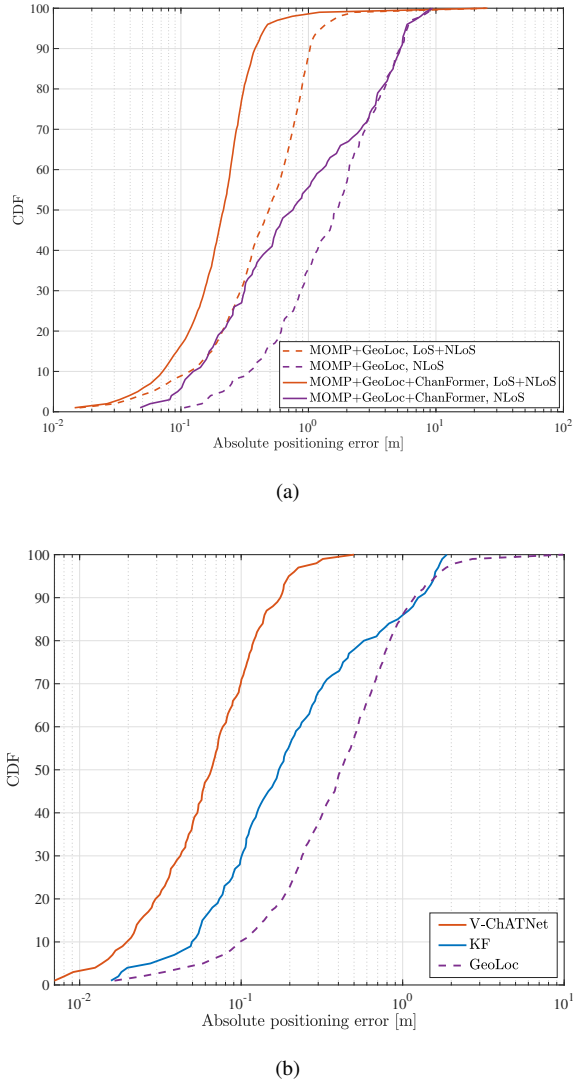


Fig. 15. Performance of the positioning strategies based on attention networks—ChanFormer for the initial access scenario and V-ChATNet for the tracking case—compared to positioning based only on channel parameter estimation and exploitation of the geometry of the environment (MOMP + GeoLoc) evaluated using the ray tracing setup described in [146], [147]: (a) Results for the initial access and position estimation scenario, where only one snapshot of the channel is estimated and leveraged for positioning; (b) Results for the tracking scenario, where series of channel and position estimates are exploited by both a Kalman Filter (KF) and V-ChATNet, while MOMP+GeoLoc only exploits previous channel estimates to reduce computational complexity in the channel parameter estimation stage.

localization demonstrates significant advantages in harsh propagation environments such as indoors and urban canyons, which attracts great attention from the research society in both theoretical and algorithmic aspects. As for the theoretical basis, the fundamental limits of localization accuracy in cooperative networks have been derived in [162], [163], where the structure of EFIM is revealed to characterize the information gain brought by UE cooperation. Then the spatio-temporal information coupling is investigated to quantify the cooperation efficiency, where the asymptotic error propagation laws are determined to provide guidelines for large-scale networks [164]–[166]. As for the operation strategies, efficient resource allocation schemes are developed from the perspectives of

convex programming and game theory, which can provide practical solutions for high-accuracy cooperative localization in both centralized and distributed schemes [167], [168].

8) *RIS-aided positioning*: A RIS is a programmable surface that can be used to control the reflection of radio waves by changing the electric and magnetic properties of the surface, in either a continuous or discrete way [169]. While there are different RIS technologies, their distinction is irrelevant for our purposes. RISs were originally devised to overcome LoS blockages, especially at FR2, by creating an additional path with high SNR. The most common kind of a RIS is a so-called passive reflective RIS, which creates an additional channel (in its more simple form)

$$\mathbf{H}_{n,m}^{\text{RIS}} = \alpha_m^{\text{RIS}} e^{-j2\pi n \Delta f \tau^{\text{RIS}}} e^{j2\pi m T_{\text{sym}} \nu^{\text{RIS}}} \mathbf{a}_{\text{R}}(\phi^{\text{RIS}}) \mathbf{a}_{\text{T}}^T(\theta^{\text{RIS}}), \quad (23)$$

where the only difference lies in the α_m^{RIS} , which is of the form [170]

$$\alpha_m^{\text{RIS}} = \alpha^{\text{T-RIS}} \alpha^{\text{RIS-R}} \mathbf{a}_{\text{RIS}}^T(\vartheta) \mathbf{\Omega}_m \mathbf{a}_{\text{RIS}}(\varphi) \quad (24)$$

in which $\alpha^{\text{T-RIS}}$ denotes the channel from transmitter to RIS, $\alpha^{\text{RIS-R}}$ is the channel from RIS to receiver, $\mathbf{a}_{\text{RIS}}(\varphi)$ is the RIS steering vector as a function of the AoA φ (from the receiver) and AoD ϑ (to the transmitter), and $\mathbf{\Omega}_m$ denotes the diagonal RIS control matrix, which represents the time-varying state of the RIS. The entries of $\mathbf{\Omega}_m$ have at most unit amplitude (since a passive RIS cannot amplify) and controllable phase. In communications $\mathbf{\Omega}_m$ can be set so that $\mathbf{a}_{\text{RIS}}^T(\vartheta) \mathbf{\Omega}_m \mathbf{a}_{\text{RIS}}(\varphi) \rightarrow M$, where M is the number of RIS elements, thereby providing an SNR scaling of up to M^2 . The value of M can be increased by making the RIS larger (leading to more power illuminated on the RIS, but also wavefront curvature effects appear (see Section VI-E) or by making the RIS denser (though the power illuminated on the RIS is constant and mutual coupling will play an important role). In addition to the considered passive reflective RIS, there are many other RIS variants, such as amplifying RIS (sometimes called active RIS), sensing RIS (sometimes called hybrid RIS), RISs that can simultaneously transmit (in the sense of passing through) and reflect (called STAR-RIS) [171], and RISs that allow non-diagonal control $\mathbf{\Omega}_m$ [172].

The role of RIS for positioning lies in the ability to provide (i) an additional position reference, similar to a BS, (ii) additional measurements as any other path (e.g., the AoA, AoD, delay, and Doppler in (23)), (iii) unique measurements of the AoA or AoD in (24) [173]. While the AoA and AoD in (24) are not jointly identifiable, typically one of the two angles is known since either the transmitter or the receiver is a BS. This leads to two practical issues: how can we know the RIS location and orientation [174] and how can the receiver separate the signal from the RIS with respect to all the uncontrolled multipath, including the LoS path? To address the first issue, RIS calibration methods have been devised, based either on transmissions between BSs or while simultaneously localizing the user. To address the second issue, which is especially relevant in a multi-RIS scenario, dedicated RIS control sequences have been designed that allow separation, e.g., such that $\sum_m \mathbf{\Omega}_m = \mathbf{0}$, so that under zero Doppler, the

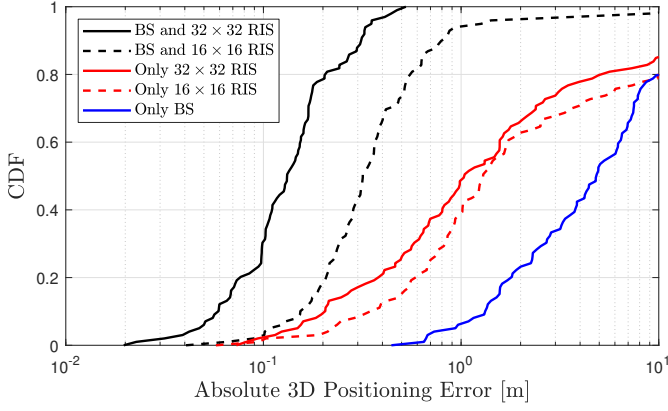


Fig. 16. RIS-aided 3D positioning performance with an 8×8 BS. Evaluated using the ray-tracing data from a smart factory environment with transmit power 20 dBm, center frequency $f_c = 60$ GHz, and bandwidth $B = 100$ MHz.

uncontrolled multipath can be recovered by adding the signals over time, for each transmit beam.

Example 6. Consider a point-to-point RIS-aided mmWave communication system where both the BS and the UE are equipped with planar arrays. In this case, the received signal at the UE contains contributions of the paths from two sources, i.e., from the environment and through the RIS, which can be expressed as

$$\mathbf{y}_{n,m}^{\text{RIS}} = \mathbf{W}^H (\mathbf{H}_{n,m} + \mathbf{H}_{n,m}^{\text{RIS}}) \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}[n, m] \mathbf{x}_{n,m} + \mathbf{z}_{n,m}. \quad (25)$$

In [175], it is shown that the sparse channel parameters from both sources can be simultaneously estimated by using a slightly modified version of the low-complexity multidimensional orthogonal pursuit matching (MOMP) algorithm proposed in [50], which estimates the parameters by using their own individual dictionaries instead of a Kronecker product of them. Note that the work in [175] considers the pulse shaping effects and clock offset, and estimates the time-domain channel while assuming that the LoS path between the BS and the RIS has been pre-estimated in the RIS calibration stage. The positioning stage contains two different cases: (1) LoS path from the BS and RIS exist, (2) LoS path from either BS or RIS exists. The first case can be solved with a simple linear equation whereas the second case can be solved via a linear system of equations constructed with LoS and NLoS paths of one source (i.e., BS or RIS) by leveraging the reflection properties of the indoor environment [50], [175]. Positioning error results with the described method obtained via ray tracing based simulations for an indoor environment are given in Fig 16. It can be seen that highly accurate positioning can be achieved when the LoS path exists for both BS and RIS. Specifically, %80 of the users have positioning error below 20 cm with a 32×32 RIS.

9) *Carrier phase positioning:* In general, the carrier phase of the incoming received signal depends explicitly on the propagation delay, as shown in Fig. 17. Hence, measuring carrier phase accurately, relative to a reference, allows carrying out delay estimation and thereon ranging with accuracies that are fractions of a wavelength. Especially in millimeter-wave

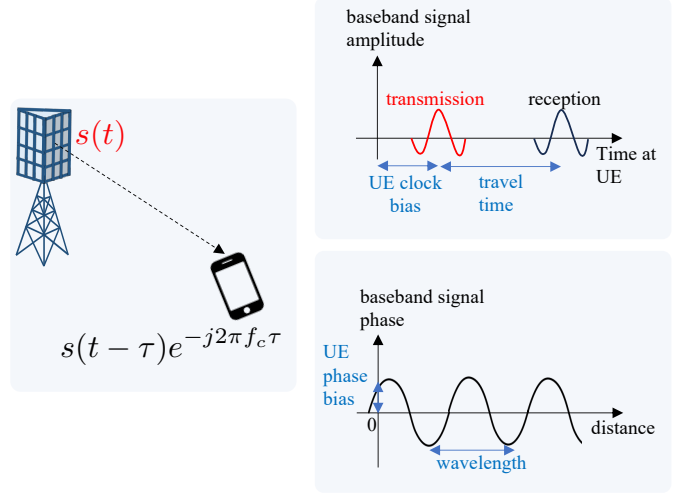


Fig. 17. Carrier phase positioning provides accurate distance information, but is subject to ambiguities equal to the wavelength.

networks, using carrier phase measurements can thus enable ranging accuracies in the order of few millimeters while in C-band networks a centimeter-level ranging accuracy is feasible.

In the context of localization and positioning, carrier phase based methods are conceptually known and broadly utilized in GNSS systems, particularly the GNSS-RTK approach, see for example [176] and the references therein. Additionally, and importantly, the utilization of carrier phase measurements is recently also considered in 3GPP cellular network standardization—particularly in the context of 5G-Advanced and corresponding enhanced positioning capabilities [177], [178].

One of the most prominent technical challenges related to the use of carrier phase measurements is the so-called integer ambiguity problem [108], [179], [180]. This refers to the fact that the carrier phase is immune to any integer multiple of the wavelength, i.e., the amount of full wavelengths in the distance between the TX and RX entities cannot be directly measured. Practical ways to solve or relax the integer ambiguity problem stem from differential and/or double-differential measurements, that build on the concept of a reference device. In the cellular network context, a UE with a known and fixed reference position could serve for such purpose.

Example recent works in this area in the cellular system context covers, e.g., [181], [182] focusing on estimation, synchronization and positioning algorithms for different use cases. Additionally, [108] describes methods how cellular carrier phase measurements combined with appropriate Bayesian filtering can facilitate super-resolution and low-latency 6DoF tracking of 5G-empowered XR headsets without any additional sensors.

C. Bistatic and multistatic sensing of non-connected objects

In contrast to positioning, which involves determining the state of connected users, in this section, we will focus on sensing (detecting and localizing) non-connected objects (sometimes also called landmarks or targets, depending on the context). We refer back to Fig. 9.

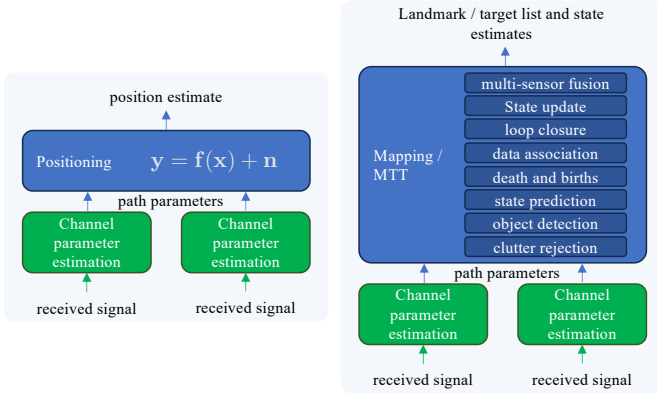


Fig. 18. Both positioning and sensing require channel parameter estimation. The subsequent processing for mapping and MTT is much more involved than for positioning.

1) *Fundamentals of mapping and target tracking, multi-site processing:* First, it is necessary to clarify the terminology: when the objects of interest are static, they are called landmarks. The corresponding sensing problem is called *mapping*. When the objects of interest are moving, they are called targets and the corresponding sensing problem is called *multi-target tracking (MTT)*. The reference to 'objects of interest' is due to there being other objects (e.g., ground reflections when tracking a moving vehicle), which generate measurements, but are not of interest to the application.

To separate positioning from mapping and MTT, let us first describe the commonalities (see Fig. 18): in both cases, processing involves two stages, wherein the first stage radio signals are sent and received, based on which LoS and NLoS paths are extracted, with associated channel parameters (delays, angles, Dopplers). In the second stage, those parameters are converted to a geometric state of the user/object. In terms of the differences, the first stage for positioning is generally focused on extracting the LoS path parameters, while in mapping and MTT, the LoS path only serves as a reference path, while the NLoS paths relate to the non-connected objects. The main difference lies in the subsequent processing, as visualized in Fig. 18. The reasons are as follows [183]:

- *Unknown number of objects:* The sensing system does not know a priori how many objects there are. This means that objects need to be first detected before they can be localized. This problem is also affected by background clutter, which is present but not of interest. Hence, sensing involves both a detection (involving false alarms and missed detections) and an estimation problem.
- *Complicated object states:* In contrast to positioning, where the goal is to estimate the UE 3D position, objects are characterized by a much more complicated state definition, which may include the velocity, the extent/shape of the object, and even the material type. Objects may also give rise to more than one measurement [184].
- *Unknown data association:* Another fundamental difference compared to positioning, is that the measurements carry no information about the object. Hence, measurements taken at some time t should be associated with

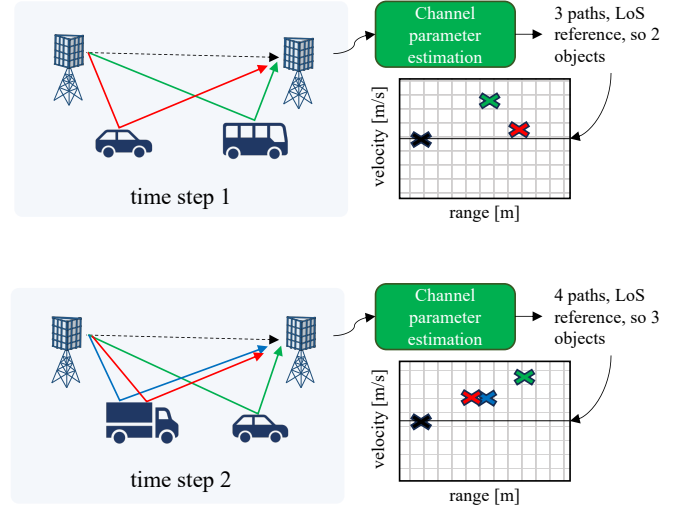


Fig. 19. Each measurement gives rise to a potential object. Objects may generate 0, 1, or several measurements at each time. Objects detected at time step 1 should be associated to measurements at time step 2.

objects detected at some earlier time $t' < t$. This type of combinatorial problem lies at the heart of all mapping and MTT methods. As time progresses the number of data associations grows quickly, which requires dedicated routines to mitigate computational complexity [185].

- *Multi-sensor fusion:* When several receivers are processing the signals from one or more transmitters, some form of fusion is needed to provide a consistent view of the landmarks or targets. The type of fusion depends on the level of coherence between the different BSs. When the receiving BSs are phase coherent, then can be interpreted as one large distributed sensor and processing can be performed on the aggregated received waveforms, as in distributed MIMO radar. Such processing involves sharing raw I/Q data with a central processing unit. In contrast, when the receiving BSs are not phase-synchronized, they act as independent observers, and fusion should occur after local channel parameter estimation. The different field of view of the different receivers makes this challenging [186].

The problems of an unknown number of objects, complicated object states, and unknown data association are visualized in Fig. 19 for a bistatic scenario.

D. Radio SLAM in bistatic scenarios

In this section, we will focus on bistatic Radio SLAM which combines aspects of positioning (see Section IV-B) and sensing (see Section IV-C) and it involves an UE determining its position, while simultaneously detecting and localizing landmarks, based on signals to/from the BSs. The Radio SLAM problem has gained widespread attention in the research field over the past years (see e.g., [55], [96], [187], [188]) since it is closely related to the minimal problems discussed in Section IV-B2 and solving it provides three important benefits. First, the NLoS paths provide additional information and can enhance positioning accuracy; second, reliance on infrastructure can

be reduced, enabling for example single-BS positioning in mmWave; third, localization is possible in mixed LoS/NLoS conditions, even in the absence of LoS.

The processing pipeline of Radio SLAM is similar to that of positioning/sensing (see Fig. 18). The key difference is that in the second stage, the channel parameter estimates are converted to a geometric state of both the UE and landmarks. Since Radio SLAM has many commonalities to positioning and sensing, we will first introduce the mapping problem which is closely related to the sensing problem and thereafter, we will present the overall Radio SLAM problem. In the following, mapping is used to refer to positioning M landmarks, $\mathbf{m}_{1:M}$, and a map is defined as $\mathbf{m} = [\mathbf{m}_1^\top, \mathbf{m}_2^\top, \dots, \mathbf{m}_M^\top]^\top$.

The mapping problem, commonly referred to as mapping with known poses in SLAM literature [189], aims to estimate the map posterior, $p(\mathbf{m} \mid \mathbf{x}_{1:t}, \mathbf{y}_{1:t})$, using the sequence of UE poses $\mathbf{x}_{1:t}$ and measurements $\mathbf{y}_{1:t}$. Since the UE trajectory is known, the landmarks are independent and the map posterior can be factorized as [189]

$$p(\mathbf{m} \mid \mathbf{x}_{1:t}, \mathbf{y}_{1:t}) = \prod_{i=1}^M p(\mathbf{m}_i \mid \mathbf{x}_{1:t}, \mathbf{y}_{1:t}). \quad (26)$$

The map posterior can be computed for example using Bayesian filtering and if the landmark density is approximated using a Gaussian, (26) can be efficiently estimated using M extended Kalman filters (EKFs) in parallel, one for each landmark. Landmark-based mapping approaches commonly decompose the physical environmental landmarks such as reflecting surfaces and scattering objects into parametric representations such as a point [187].

The objective of SLAM is to compute the joint posterior density of the UE trajectory and map, $p(\mathbf{x}_{1:t}, \mathbf{m} \mid \mathbf{y}_{1:t})$, given the measurements up to time t . An important characteristic of the SLAM problem is that by conditioning the map to the UE trajectory renders the landmark estimates conditionally independent. Exploiting this feature, the joint SLAM density can be factorized as [190]

$$p(\mathbf{x}_{1:t}, \mathbf{m} \mid \mathbf{y}_{1:t}) = p(\mathbf{x}_{1:t} \mid \mathbf{y}_{1:t}) p(\mathbf{m} \mid \mathbf{x}_{1:t}, \mathbf{y}_{1:t}) \quad (27)$$

in which $p(\mathbf{x}_{1:t} \mid \mathbf{y}_{1:t})$ and $p(\mathbf{m} \mid \mathbf{x}_{1:t}, \mathbf{y}_{1:t})$ are posterior of the UE and map, respectively. This factorization makes it natural to apply Rao-Blackwellized particle filter (RBPF) solutions, in which a particle filter (PF) is used to approximate $p(\mathbf{x}_{1:t} \mid \mathbf{y}_{1:t})$ and computing $p(\mathbf{m} \mid \mathbf{x}_{1:t}, \mathbf{y}_{1:t})$ is equivalent to solving (26). It is important to note that each particle represents a single UE trajectory and a unique map is associated to every particle. In Radio SLAM, the UE state consists of the pose and clock. The pose represents the position and orientation of the UE, whereas the clock represents the required parameters needed to synchronize the local clock of the UE to the network clock.

In literature, a large variety of solutions to SLAM are available and these can be classified as snapshot, filtering and smoothing approaches. The concrete difference between the approaches is the time horizon. Snapshot SLAM only considers observations at time t' for estimating $p(\mathbf{x}_{t'}, \mathbf{m} \mid \mathbf{y}_{t'})$, filtering approaches utilize measurements up to time t for approximating $p(\mathbf{x}_t, \mathbf{m} \mid \mathbf{y}_{1:t})$, while smoothing approaches

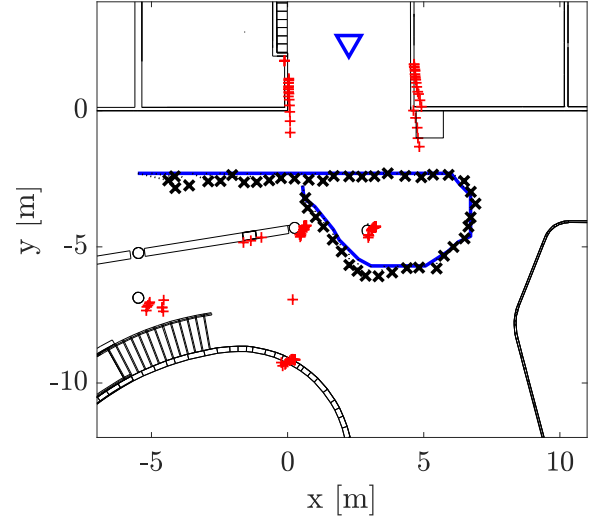


Fig. 20. Example bistatic Radio SLAM performance obtained using an RBPF-PHD filtering solution. The TX position is illustrated with (∇), UE trajectory using (\rightarrow), and estimated UE path and landmark locations with (\times) and ($+$), respectively.

perform batch processing to estimate the full SLAM posterior $p(\mathbf{x}_{1:T}, \mathbf{m} \mid \mathbf{y}_{1:T})$ in which $T \geq t$. The early works in conventional SLAM mainly considered the filtering problem but due to inconsistency issues, majority of the works nowadays consider the smoothing problem [191]. In bistatic Radio SLAM, the smoothing problem is yet unexplored and the other two approaches have been considered instead. The main differentiator of bistatic Radio SLAM is that the UE state is estimated in a global reference system that is defined with respect to the BS, whereas in conventional SLAM, estimation is typically performed in the local frame of the sensor (see also Section V-E). This has significant ramifications to the bistatic Radio SLAM problem and two prominent examples include: i) it is possible to estimate the UE state with respect to a global reference system just from single snapshot observation as presented in Section IV-B; ii) filter divergence can be identified by comparing the filtering solution to the snapshot solution. It is important to note that both snapshot and filtering approaches are relevant in bistatic Radio SLAM and the preferred choice depends on the overall system and application scenario. Snapshot SLAM is fundamentally important as it serves as a baseline for what can be done with radio signals alone, while filtering based SLAM methods are expected to improve the accuracy. The snapshot approach can be solved for example using numerical optimization methods [55], [192], whereas the filtering problem is commonly solved using Bayesian filtering [187], [188] or belief propagation on factor graphs [96], [109].

Example 7. In this example, we consider an indoor bistatic SLAM scenario composing of a single BS that transmits positioning reference signals (PRSs) which are received by the UE to jointly estimate its pose and map the surrounding environment. The experiment was conducted at 60 GHz carrier frequency and using 400 MHz PRS bandwidth. The BS and UE were equipped with 4×16 planar antenna arrays and

Sivers Semiconductors Evaluation Kits EVK06002 were used as the TX and RX entities. The TX and RX used 126 and 252 beams, respectively, which corresponds to a 180° field-of-view (FOV) for the BS and 360° FOV for the UE. Overall, the UE trajectory consisted of 45 measurements positions. The experimental scenario is illustrated in Fig. 20 and for further details, please see [193].

The considered SLAM problem is solved using an RBPF-probability hypothesis density (PHD) filter [187] that utilizes an optimal importance density (OID) approximation [188] to decrease the number of required particles. The performance of the algorithm is visualized in Fig. 20 and as illustrated, the estimated path closely follows the UE trajectory and the estimated landmarks coincide with the actual floor plan. In the experiment, the RMSE position, heading and synchronization error are: 0.55 m, 2.43° and 1.53 ns, respectively. The benefit of using a filtering approach is three fold. First of all, the filter can operate in mixed LoS/NLoS conditions and the posterior from the previous time step can be viewed as a regularization term which constrains the posterior update so that the system state is identifiable at every measurement position. Second, sequential processing of the measurements improves the accuracy. Third, the filter can inherently deal with the challenges of bistatic Radio SLAM which include: i) the landmark can be misdetected due to limitations in the RX and channel estimation routine; ii) clutter measurements and multi-bounce signals can generate false detections that are not inline with the measurement model; and iii) measurement ambiguities can lead to situations where a wrong measurement is associated to the wrong landmark.

V. TECHNOLOGIES FOR JOINT MONOSTATIC SENSING AND COMMUNICATION

A. Introduction

In this section, we address the mono-static sensing and SLAM paradigms in the spirit of cellular ISAC. We first shortly review the fundamentals and challenges of monostatic system scenarios. Then, in Section V.B, we present the basics of the related SI waveform modeling and discuss shortly the potential TX-RX isolation solutions, while Section V.C provides examples how precoding/beamforming optimization can contribute to SI suppression. Section V.D then looks into relevant impairments, most notably oscillator phase noise and Doppler-induced ICI, and shows how these can be turned from foe to friend in monostatic sensing. Lastly, section V.E presents monostatic SLAM, which is one potential application of monostatic sensing.

In general, different from the previously discussed bi-static and multi-static transmitter-receiver arrangements, the *mono-static* case refers to a scenario where the transmitting and receiving entities are essentially collocated [194]–[196]. A principal illustration of such mono-static sensing and mapping scenario is shown in Fig. 21, indicating also the opportunity for separate antenna systems at the TX and RX ends [89]. Such mono-static approach allows to turn the individual UEs or gNodeBs (gNBs) essentially into cellular radars, providing standalone situational awareness related to the surrounding

environment without relying on other network entities. Example representative use cases could be, e.g., in different vehicular or industrial systems where connected moving machines can extract and harness situational awareness in terms of the environment landmarks while also tracking their own coordinates relative to a reference point, by using their own transmit waveform as the illumination signal. Compared to the bi-static or multi-static counterparts, the mono-static approach is appealing as the complete I/Q transmit waveform is directly known also to the receiver. Additionally, arranging for very accurate time- and frequency-synchronization between the TX and RX chains is clearly much more straight-forward and can be resolved even at hardware level.

The mono-static operation is, as such, well-known from ordinary radars. Pulsed radars [194] operate on the basis of time-multiplexing between the exact transmit and receive time periods. In the context of cellular systems, specifically in TDD networks, the individual network nodes (UEs and gNBs and the corresponding transmission and reception points (TRPs)) also operate on the basis of dividing the TX and RX active periods in time. However, the individual active transmit periods in both uplink and downlink are commonly in the order of a millisecond – a period that is very long compared to pulsed radars, and does not facilitate measuring transmit waveform interaction with targets or landmarks at any meaningful distance. Hence, any cellular monostatic sensor must be able to execute the receiver simultaneous to transmitting, in order to measure the reflecting and scattering waves. This, in turn, means that such sensor is essentially operating as an inband full-duplex (IBFD) transceiver, commonly also referred to as a simultaneous transmit-and-receive (STAR) system [195]–[197].

One fundamental technical challenge in STAR systems is, in general, the *self-interference* (SI) or direct transmit-receive coupling [197]–[199]. As a concrete example, the effective isotropic radiated power (EIRP) of a macro gNB can be in the order of +70 dBm, or more, while the receiver thermal noise floor with e.g. 100 MHz channel bandwidth is around –90 dBm. Thus, from the sensing or radar function point of view, the direct TX-RX coupling acts as an extremely powerful target at a very short distance, and will mask everything else if not properly suppressed through antenna processing, active RF cancellation and baseband digital cancellation [197]–[201]. Importantly, sufficient amount of TX-RX isolation must be obtained already in the antenna or RF circuit domain prior to the receiver low noise amplifier (LNA) – otherwise, the powerful SI will totally desensitize and block the LNA and thereon the sensing receiver. In the following Section V.B, we provide further modeling of the SI waveform while also discussing the basic TX-RX isolation approaches.

B. Basic SI modeling and isolation solutions

We next shortly discuss the different potential solutions to arrange sufficient TX-RX isolation, as well as the related topic of SI channel and waveform modeling. There are, generally speaking, two alternative ways to arrange the antenna interface for the simultaneously operating transmitter and receiver, namely, sharing a set of antennas through circulators for both

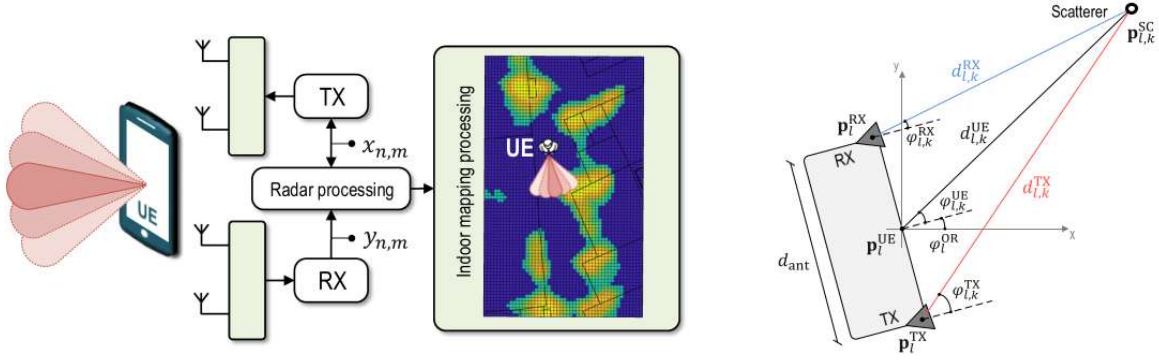


Fig. 21. Illustration of the monostatic sensing paradigm at UE end, together with the corresponding problem geometry.

TX and RX, or then utilizing different sets of physical antennas for TX and RX. One may consider the latter more feasible approach, as the circulators are commonly providing only isolation levels in the order of 20 dB while the different sets of antennas allow for larger isolation levels through separating the antenna systems in space—especially in gNB type of entities as well as in UE with larger form factors such as vehicles. Additionally, an importantly separate antenna systems allow harnessing beamforming optimization for TX-RX isolation in a more efficient manner. Hence, we also primarily assume in the following that separate antenna systems are utilized, though the RF and baseband digital canceller principles apply to both scenarios.

Considering first the basic SI channel and waveform modeling, let us denote the baseband precoded transmit waveform at transmit path l by $x_l(n)$, and further assume that subarray based hybrid beamforming is adopted at TX where each precoded signal is transmitted through a subarray of size M . Then, the complex baseband equivalent SI waveform at receiver antenna j , contributed by the TX subarray l , reads

$$x_{j,l}^{\text{RF,SI}}(n) = \sum_{i=1}^M h_{i,j}^{\text{SI},l}(n) * (f_{i,l} \phi_l^{\text{TX}}(x_l(n))), \quad (28)$$

where $h_{i,j}^{\text{SI},l}(n)$ denotes the physical SI channel from the antenna element i of the TX subarray l to RX antenna element j , $*$ denotes convolution, while the function $\phi_l^{\text{TX}}(\cdot)$ refers to a transmitter hardware model that can accommodate different impairments such as power amplifier (PA) nonlinearity and oscillator phase noise. Additionally, the beamforming weights of the TX subarray l are denoted by $f_{i,l}$.

Next, taking into account the contribution of all involved TX subarrays, say $l = 1, 2, \dots, L$, while incorporating also RX beamforming as well as RX hardware path impairments, the observable baseband SI waveform reads eventually

$$x^{\text{BB,SI}}(n) = \phi^{\text{RX}} \left(\sum_{j=1}^K w_j \sum_{l=1}^L x_{j,l}^{\text{RF,SI}}(n) \right), \quad (29)$$

where K denotes the array size at RX, w_j denotes the corresponding beamforming weight, and ϕ^{RX} refers to the RX path hardware model. This model is valid for an arbitrary RX (sub)array.

Importantly, the models in (28)-(29) show that both TX and RX beamforming contribute to the observable SI at RX. Thus, beamforming is one fundamental processing tool to facilitate or improve TX-RX isolation—and particularly the TX beamforming as it allows to control the SI waveform already at the RX LNA input as shown by (28). Beamforming optimization solutions that seek for controlled and favorable tradeoffs between SI suppression, target illumination capability and achievable communication rate are described, e.g., in [202]–[205], and will also be discussed in the following Section V.C. Furthermore, the accommodation of the prominent hardware imperfections in (28)-(29) is of great importance. For one, as shown e.g. in [200], [206], [207], appropriate HW modeling allows to craft advanced digital SI cancellation solutions that can suppress also the nonlinear effects and distortion products imposed by TX PA, RX LNA and other analog components. Such is important when optimizing the receiver dynamic range. And what is more, as discussed and demonstrated concretely later in this Chapter, properly modeling of hardware imperfections can turn them from a foe to a friend, when interpreted from the sensing point of view—this is particularly so for the oscillator phase noise as discussed further in the following in Section V.D.

C. Triple-function precoding/combining in full duplex transceivers

In this section, we discuss the precoder and combiner design that enable STAR operation for monostatic sensing with communication-centric FD transceivers. We focus on designs suitable for initial access and also tracking scenarios.

1) *Precoding/combining for tracking:* We focus on the tracking stage of the joint DL communication and sensing with an FD BS in a single UE setting. The three functions of the spatial filters at both the TX and the RX side can be stated as follows: (i) enabling DL communication, (ii) providing high gain in the target direction, (iii) and contributing to SI suppression. To provide simultaneous communication and sensing the system should operate with a multibeam precoder that illuminates the target angle and the multiple paths of the communication channel [195]. Considering a system operating in FR2 with a hybrid analog/digital architecture at both ends

of the FD BS and OFDM signaling, the received signal at the FD BS can be expressed as

$$\mathbf{y}_{n,m} = \mathbf{W}^H \mathbf{H}_{n,m} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}[n] \mathbf{x}_{n,m} + \underbrace{\mathbf{W}^H \mathbf{H}_n^{\text{SI}} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}[n] \mathbf{x}_{n,m}}_{\text{SI}} + \mathbf{z}_{n,m}, \quad (30)$$

where $\mathbf{H}_{n,m}, \mathbf{H}_n^{\text{SI}} \in \mathbb{C}^{N_{\text{R}} \times N_{\text{T}}}$ are the radar and SI channels, respectively. The entries of the SI channel are the coupling between each TX-RX antenna pair, as described previously. We assume that the FD BS is stationary, which is why the SI channel remains constant at different symbols. The second term in (30) should be suppressed with precoding/combining so that the LNAs are not saturated and the targets can be detected.

The precoders should also provide reliable DL communication and target gain for the radar operation. We can quantify the communication performance with the subcarrier dependent rate \mathcal{R}_n , which is a function of the precoders, the DL channel, and the combiners at the UE. The precoders and the combiners at the UE are designed separately, which is the common approach in communications. Moreover, we can quantify the sensing performance with the TX target gain per stream for a given target angle θ_{r} , which is expressed as

$$G_{\text{T},n_s,n}(\theta_{\text{r}}) = \left| \mathbf{a}_{\text{T}}^H(\theta_{\text{r}}) [\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}[n]]_{:,n_s} \right|^2, \quad (31)$$

for $n_s = 1, \dots, N_s$. The final objective of the precoders is to mitigate the SI that is equivalent to nulling the SI term in (30) as

$$\mathbf{W}^H \mathbf{H}_n^{\text{SI}} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}[n] = \mathbf{0}_{N_{\text{RF}}, \text{R} \times N_s}, \quad (32)$$

which is a joint task for the combiner. Aside from SI suppression, the goal of the combiner is to provide a high target gain at the n_{RF} -th RX RF chain that is expressed as

$$G_{\text{R},n_{\text{RF}}}(\theta_{\text{r}}) = \left| [\mathbf{W}]_{:,n_{\text{RF}}}^H \mathbf{a}_{\text{R}}(\theta_{\text{r}}) \right|^2, \quad (33)$$

for $n_{\text{RF}} = 1, \dots, N_{\text{RF}, \text{R}}$. The joint precoding/combining problem can be formulated as the maximization of the sum DL rate over all subcarriers under the target gain, SI suppression and hardware constraints [202], [203]. We can define TX and RX gain thresholds denoted by τ_{T} and τ_{R} , respectively. With the described formulation, we can write the overall optimization problem as

$$\begin{aligned} & \underset{\mathbf{W}, \mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}[n]}{\text{maximize}} && \sum_{n=0}^{N-1} \mathcal{R}_n \\ & \text{subject to} && G_{\text{T},n_s,n}(\theta_{\text{r}}) \geq \tau_{\text{T}}, \quad G_{\text{R},n_{\text{RF}}}(\theta_{\text{r}}) \geq \tau_{\text{R}}, \\ & && (32), \quad \mathbf{F}_{\text{RF}} \in \mathcal{A}_{\text{T}}, \quad \mathbf{W} \in \mathcal{A}_{\text{R}}, \\ & && \|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}[n]\|_{\text{F}}^2 = N_s, \end{aligned} \quad (34)$$

where \mathcal{A}_{T} and \mathcal{A}_{R} are the sets of feasible structures for the analog precoder and the analog combiner. For example, unit-modulus entries, which correspond to a phase shifter network, and the subarray architecture can be imposed on the analog precoder/combiner by using the defined sets. Finally, normalization is imposed on the precoders to satisfy the power constraint.

The problem in (34) is hard to solve due to the coupling between the precoders and the combiner, and the hardware constraints related to the analog precoder/combiner. Hence, alternating optimization can be adopted to decompose the problem into several subproblems, each of them responsible for one of the variables while the others are fixed. Furthermore, relaxation techniques such as convex approximations and semidefinite problems can be used to solve each subproblem. The work in [202] solves this problem by relaxing the cost function from rate to gain in the LoS angle of the users. Moreover, SI cancellation is satisfied with only the combiner by using the null-space projection (NSP) method which requires fine grained attenuators in addition to phase shifters. In [203], the authors maximize the data rate by assuming that a subset of the targets in the environment contribute to downlink communication. They utilize an analog SI cancellation method which reduces the SI. Then, the analog precoder and combiner are designed by using a DFT codebook. Finally, the residual SI is suppressed by the digital precoders, which also maximize the DL communication rate. In [208], a more general system model is considered where the DL channel can have multiple paths which do not necessarily coincide with the targets. This work utilizes generalized eigenvalue-based precoders that suppress the SI while maximizing the data rate. The precoders are coherently combined with the precoders that maximize the TX target gain. Then, a hybrid decomposition algorithm is used to obtain analog and digital precoders, which reduce the SI suppression capability. Finally, a convex relaxation-based algorithm is employed to design the combiner that minimize the residual SI while keeping the RX target gain above a threshold.

2) *Precoding/combining for initial access*: In this setting, the BS and UE have to establish communication for the first time. Thus, the precoders at the BS and the combiners at the UE should be designed such that they cover a wide angular region. This stage, in nature, is similar to the target discovery step of radar operation. Hence, the precoders and combiners at the BS that are used for initial access and channel estimation for communication are suitable for radar sensing. If we consider the triple-function precoding/combining for tracking, we could see that the spatial resources are shared between the communication and sensing, which would not be necessary for the initial access stage. In other words, the triple-function requirement for precoding/combining would be reduced to two-function as the communication and sensing requirement would be unified.

Initial access can be established with beam training for FR2 in 5G NR. The beams are selected from a codebook that comprise directional beams. Let us denote the directional codebooks for TX and RX by $\mathbf{A}_{\text{T}} = [\mathbf{a}_{\text{T}}(\theta_1) \dots \mathbf{a}_{\text{T}}(\theta_{M_{\text{T}}})] \in \mathbb{C}^{N_{\text{T}} \times M_{\text{T}}}$ and $\mathbf{A}_{\text{R}} = [\mathbf{a}_{\text{R}}(\phi_1) \dots \mathbf{a}_{\text{R}}(\phi_{M_{\text{R}}})] \in \mathbb{C}^{N_{\text{R}} \times M_{\text{R}}}$, where M_{T} and M_{R} are the sizes of AoA set $\{\theta_i\}_1^{M_{\text{T}}}$ and AoD set $\{\phi_j\}_1^{M_{\text{R}}}$, respectively. If these codebooks are used for sensing at the FD BS, we would observe high residual SI. Let us denote the SI-aware beam codebooks that we would like design as $\bar{\mathbf{A}}_{\text{T}} \in \mathbb{C}^{N_{\text{T}} \times M_{\text{T}}}$ and $\bar{\mathbf{A}}_{\text{R}} \in \mathbb{C}^{N_{\text{R}} \times M_{\text{R}}}$. The coupling matrix that shows the residual SI between beam pairs can be defined as $\mathbf{C} = \bar{\mathbf{A}}_{\text{R}}^H \mathbf{H}^{\text{SI}} \bar{\mathbf{A}}_{\text{T}}$, where $\mathbf{H}^{\text{SI}} \in \mathbb{C}^{N_{\text{R}} \times N_{\text{T}}}$ is the frequency-flat SI channel. For a wideband SI channel, \mathbf{H}^{SI} can be selected as the SI channel at the center frequency. The goal of

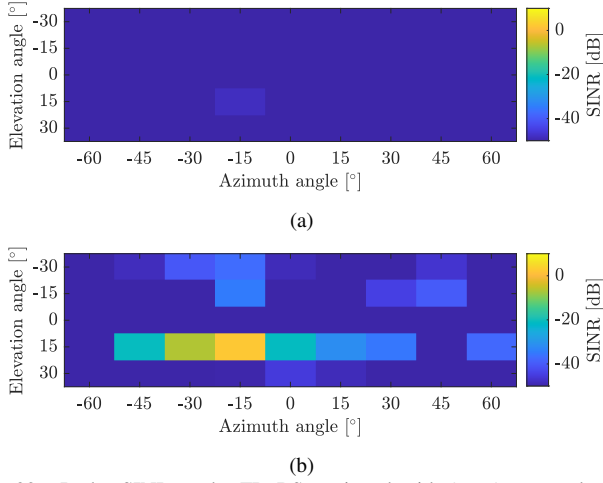


Fig. 22. Radar SINR at the FD BS equipped with 8×8 arrays that are separated by 10λ for: (a) the initial codebook, (b) the SI-aware codebook. The center frequency is 28 GHz, the bandwidth is 300 GHz, the transmit power is 20 dBm, and the SI-to-noise ratio is 80 dBm. The target is located at 80 m from the FD BS with azimuth angle -20° and elevation angle 15° .

the codebook design is to minimize the total coupling power, i.e., $\|\mathbf{C}\|_F^2$, while the beams maintain a high gain at the AoA and AoD grids. One other constraint is to have unit-modulus entries so that the beams can be realized with a purely analog architecture that comprises phase shifters, which makes the problem non-convex. The TX and RX codebooks are designed with alternating minimization, and convex relaxation combined with block-coordinate descent method by using $\bar{\mathbf{A}}_T$ and $\bar{\mathbf{A}}_R$ as initial codebooks in [209]. The beam training procedure for sensing is designed such that a TX-RX beam pair that correspond to the same angle on the AoD and AoA grids is employed for each symbol at the FD BS. We show the radar SINR obtained at the angular grid in a single point target scenario in Fig. 22. The FD BS is equipped with two 8×8 planar arrays that are separated by 10λ . The angular grid for both TX and RX codebooks is selected as $[-60^\circ, -45^\circ, \dots, 60^\circ] \times [-30^\circ, -15^\circ, \dots, 30^\circ]$. The results show that the SI-aware codebook efficiently suppresses SI and yields high radar SINR at target angle, while the initial codebook is corrupted by the SI, which makes the target detection impossible. The beams from the SI-aware codebook can also be used for the tracking stage by utilizing the beams from the codebook at the analog precoder/combiner design.

The described monostatic sensing scenarios for the tracking and initial access stages include only DL users for communication. Thus, the precoder at the FD BS is designed such that it serves for both sensing and communication while considering SI suppression. From a communication perspective, only half-duplex communication is considered. One of the most important challenges is to integrate UL users to enable FD communication which would significantly improve the spectrum efficiency of the overall system.

D. Joint monostatic sensing and communication under non-idealities: repurposing challenges into benefits

In this part, we address the problem of joint monostatic sensing and communication with OFDM waveform in the presence

of non-idealities, and demonstrate how such imperfections can be turned into an advantage for sensing. We begin by presenting the OFDM radar signal model and basic operations for range-Doppler detection/estimation. Then, we focus on two specific non-idealities, namely inter-carrier interference (ICI) and phase noise (PN), elaborating on their impacts on sensing as well as on how they can be exploited to improve sensing performance. For ease of exposition, we consider a single-input single-output (SISO) system at both radar and communication receivers.

1) *OFDM monostatic radar sensing without non-idealities:* Consider a monostatic ISAC setup in Fig. 23, where a monostatic ISAC transceiver sends data symbols to a remote communications receiver and collects the backscattered signals at the co-located radar receiver for sensing the objects in the environment. Under a SISO setup, the monostatic OFDM radar observations can be expressed, using (1) and (2), as

$$y_{n,m} = \sum_{\ell=0}^{L-1} \alpha_\ell e^{-j2\pi n \Delta f \tau_\ell} e^{j2\pi m T_{\text{sym}} \nu_\ell} x_{n,m} + z_{n,m}. \quad (35)$$

Stacking (35) over N subcarriers and M symbols, we obtain the frequency-time domain observations in compact matrix form as [44]

$$\mathbf{Y} = \mathbf{X} \odot \sum_{\ell=0}^{L-1} \alpha_\ell \mathbf{b}(\tau_\ell) \mathbf{c}^\top(\nu_\ell) + \mathbf{Z} \in \mathbb{C}^{N \times M}, \quad (36)$$

where $\mathbf{X} \in \mathbb{C}^{N \times M}$ with $[\mathbf{X}]_{n,m} = x_{n,m}$, $\mathbf{Y} \in \mathbb{C}^{N \times M}$ with $[\mathbf{Y}]_{n,m} = y_{n,m}$, and $\mathbf{b}(\tau) \in \mathbb{C}^{N \times 1}$ and $\mathbf{c}(\nu) \in \mathbb{C}^{M \times 1}$ denote, respectively, the frequency-domain and time-domain steering vectors, i.e., $[\mathbf{b}(\tau)]_n = e^{-j2\pi n \Delta f \tau}$ and $[\mathbf{c}(\nu)]_m = e^{j2\pi m T_{\text{sym}} \nu}$. To detect targets from (36) and estimate their delay-Doppler parameters, we first remove the impact of data symbols \mathbf{X} either via reciprocal filtering/zero-forcing (i.e., dividing \mathbf{Y} by \mathbf{X} element-wise) or matched filtering (i.e., multiplying \mathbf{Y} by the conjugate of \mathbf{X} element-wise) [3], [43], [210], [211]. Since $\mathbf{b}(\tau)$ and $\mathbf{c}(\nu)$ correspond to DFT/IDFT matrix columns on a uniformly sampled delay-Doppler grid, taking IDFT over the columns and DFT over the rows of the resulting observation matrix provides the delay-Doppler/range-velocity spectrum, from which target detection and parameter estimation can be performed [3], [43], [49], [212], e.g., via constant false alarm rate (CFAR) processing [194, Ch. 6.2.4]. Alternatively, super-resolution algorithms, such as MUSIC and ESPRIT, can be applied by harnessing the structure in $\mathbf{b}(\tau)$ and $\mathbf{c}(\nu)$ (e.g., shift-invariance property) [213], [214].

2) *OFDM monostatic radar sensing under ICI:* When dealing with high-speed targets and/or small Δf , the validity of the model (36) diminishes since intra-symbol Doppler-induced phase shifts become non-negligible, destroying subcarrier orthogonality and leading to ICI. Under scenarios with high-mobility and/or small Δf , (36) can be generalized to [49]

$$\mathbf{Y} = \sum_{\ell=0}^{L-1} \alpha_\ell \mathbf{F}_N \underbrace{\mathbf{D}(\nu_\ell)}_{\text{ICI}} \mathbf{F}_N^H (\mathbf{X} \odot \mathbf{b}(\tau_\ell) \mathbf{c}^\top(\nu_\ell)) + \mathbf{Z}, \quad (37)$$

where $\mathbf{D}(\nu) = \text{diag}\left(1, e^{j2\pi \frac{T}{N} \nu}, \dots, e^{j2\pi \frac{T(N-1)}{N} \nu}\right) \in \mathbb{C}^{N \times N}$ encodes intra-symbol (*fast-time* in radar nomenclature) phase

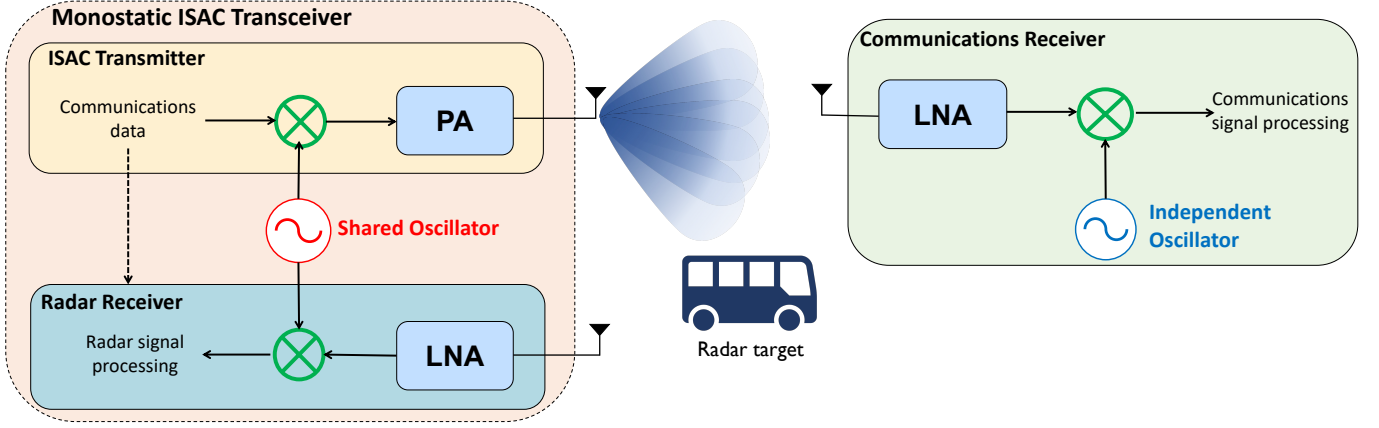


Fig. 23. Monostatic ISAC setup with a monostatic ISAC transceiver containing an ISAC transmitter and a radar receiver co-located on a shared hardware, and a communications receiver located on a remote device. In contrast to the independent oscillator employed at the communications receiver, the use of a shared oscillator in monostatic sensing implies that impairments, such as phase noise (PN) and intercarrier interference (ICI), carry geometric information on the parameters of radar targets (i.e., delay and Doppler, respectively) and thus can be turned into benefits for sensing. Please see Table III for further details.

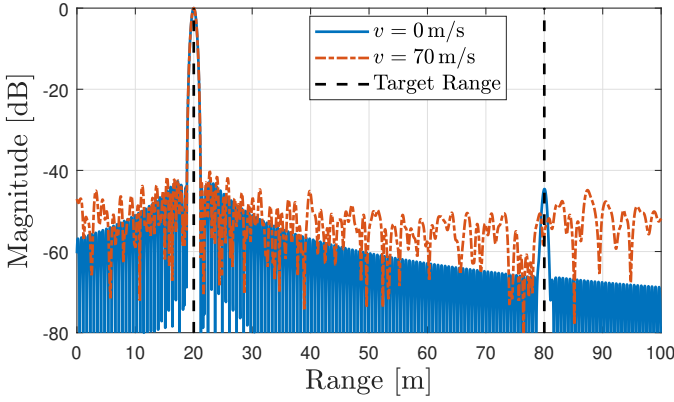


Fig. 24. Range spectrum in OFDM-based sensing under the impact of ICI. The OFDM parameters are $f_c = 28$ GHz, $\Delta f = 60$ kHz, $N = 4096$ and $M = 16$. The scenario includes 2 targets with the same velocity v , the ranges (20, 80) m and the SNRs (25, -20) dB, respectively. $v = 70$ m/s represents the case where two cars are approaching one another on a highway.

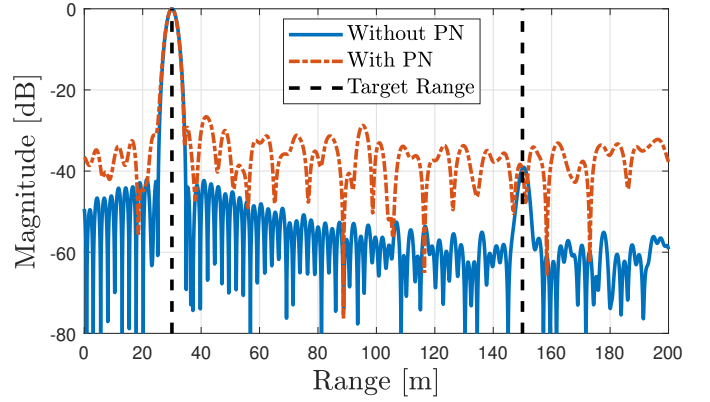


Fig. 25. Range spectrum in OFDM-based sensing under the impact of PN. The OFDM parameters are $f_c = 28$ GHz, $\Delta f = 120$ kHz, $N = 512$ and $M = 10$, while the oscillator is a free-running oscillator with the 3 dB bandwidth 200 kHz. The scenario contains 2 targets with the ranges (30, 150) m, the velocities (10, 10) m/s, and the SNRs (30, -10) dB, respectively.

shifts as a function of Doppler ν . We note that inter-symbol (*slow-time*) phase shifts are captured by $\mathbf{c}(\nu)$. For low velocities and/or large subcarrier spacing, the maximum phase progression in $\mathbf{D}(\nu)$ satisfies $2\pi T\nu \ll 2\pi$. In this case, $\mathbf{D}(\nu) \approx \mathbf{I}$ and (37) boils down to (36). To illustrate the impact of ICI on sensing, Fig. 24 shows the range spectrum of OFDM radar obtained via standard DFT/IDFT based processing, using typical 5G NR FR2 parameters in a high-mobility scenario. We observe a noticeable increase in side-lobe levels induced by ICI, leading to a masking effect on targets. This phenomenon poses a significant challenge for sensing as it could severely impede the detection of weaker targets [49], [215].

3) *OFDM monostatic radar sensing under PN*: For practical, non-ideal oscillators, PN should be taken into account in sensing. In this case, (36) becomes [216]

$$\mathbf{Y} = \sum_{\ell=0}^{L-1} \alpha_{\ell} \mathbf{F}_N \left[\underbrace{\mathbf{W}(\tau_{\ell})}_{\text{PN}} \odot \mathbf{F}_N^H \left(\mathbf{X} \odot \mathbf{b}(\tau_{\ell}) \mathbf{c}^T(\nu_{\ell}) \right) \right] + \mathbf{Z}, \quad (38)$$

where $\mathbf{W}(\tau_{\ell}) \in \mathbb{C}^{N \times M}$ represents the multiplicative PN matrix in the fast-time/slow-time domain, belonging to the ℓ^{th} target. Due to the use of a shared oscillator at the ISAC transmitter and the radar receiver, as shown in Fig. 23, $\mathbf{W}(\tau_{\ell})$ corresponds to a realization of a differential (self-correlated) PN process [217], which has *delay-dependent statistics* [216] (the so-called *range correlation* effect [218]–[220]). For the special case of an ideal oscillator, $\mathbf{W}(\tau_{\ell})$ degenerates to an all-ones matrix and (38) reverts to (36). Fig. 25 illustrates the impact of PN on the range spectrum of OFDM radar. Similar to ICI, PN distorts subcarrier orthogonality and reduces the dynamic range of the radar, which, in turn, degrades detection performance, particularly for weak targets.

4) *Exploitation of ICI and PN in monostatic sensing*: As illustrated in Fig. 23, a critical distinction emerges between monostatic radar sensing and communications in ISAC systems when impairments, such as ICI and PN, are present. Specifically, in a monostatic sensing configuration, the radar receiver shares the same oscillator with the ISAC transmitter. In contrast, the communications receiver uses an independent oscillator. This

TABLE III
EXPLOITATION OF IMPAIRMENTS IN JOINT MONOSTATIC SENSING AND COMMUNICATIONS

	Monostatic Sensing	Communications
Oscillator	Shared	Independent
ICI	Doppler	Doppler + CFO
PN	Delay-dependent statistics	Independent statistics

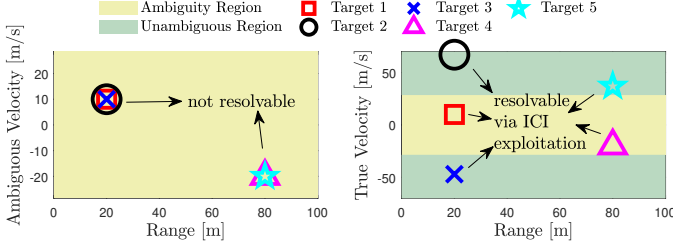


Fig. 26. Exploitation of ICI in a multi-target scenario to resolve Doppler ambiguity of high-mobility targets and introduce additional dimension for target resolvability, where $f_c = 60$ GHz, $B = 50$ MHz and $N = 2048$.

implies that (i) the ICI effect in (37) involves only target Doppler without any carrier frequency offset (CFO) [49], and (ii) the statistics of the PN in (38) are delay-dependent since this self-correlated PN process represents the difference between the original PN process and a time-shifted version, where the shift corresponds to the round-trip delay of the target [216], [221]. Table III summarizes the differences between monostatic sensing and communications in the face of impairments.

Such distinctive properties of ICI and PN in monostatic sensing present opportunities to turn these traditionally detrimental effects into beneficial elements to improve sensing performance. As seen from (37), ICI brings additional Doppler information through $\mathbf{D}(\nu)$ on top of slow-time Doppler information carried by $\mathbf{c}(\nu)$. Due to the N times higher frequency of time-domain sampling in $\mathbf{D}(\nu)$ compared to $\mathbf{c}(\nu)$, we can derive an unambiguous velocity from $\mathbf{D}(\nu)$ that is N times greater. Hence, ICI can be exploited in two different ways: (i) resolving Doppler ambiguity of high-speed targets, and (ii) enhancing target resolvability by introducing an additional unambiguous Doppler dimension, allowing us to distinguish targets located in the same delay-Doppler-angle bin [49]. Fig. 26 shows an example range-velocity scenario illustrating how ICI can be turned into a benefit for sensing. Analogous to ICI, the delay-dependency of the PN statistics in (38) can be exploited to resolve range ambiguity of far-away targets since the range information conveyed through the statistics of $\mathbf{W}(\tau)$ is not subject to any ambiguity, unlike the one carried by $\mathbf{b}(\tau)$, which has a range ambiguity of $c/(2\Delta f)$ [216]. Fig. 27 provides an illustrative example of PN exploitation for resolving range ambiguity of a distant target by a covariance matching approach [216, Alg. 2]. We note that employing this exploitation strategy elevates the performance of a radar with PN even beyond that of an ideal radar without PN, effectively repurposing challenges into benefits for sensing.

E. Radio SLAM in monostatic scenarios

Classical monostatic SLAM originated in the field of robotics, where a mobile user (e.g., a robot) continuously

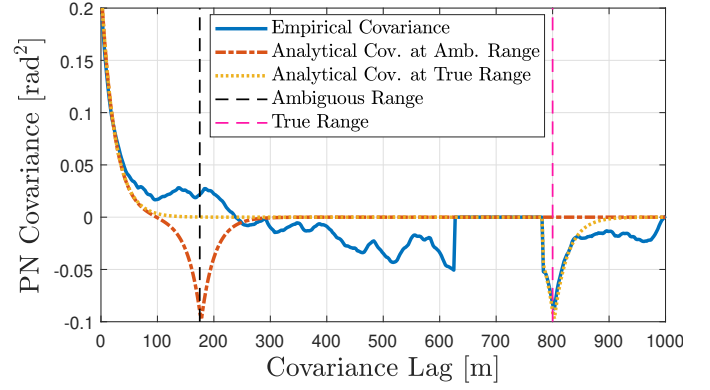


Fig. 27. Exploitation of PN via covariance matching to resolve range ambiguity of a target at 800 m, which appears at 175 m due to the maximum unambiguous range of 625 m. The OFDM parameters are $f_c = 28$ GHz, $\Delta f = 240$ kHz, $N = 256$ and $M = 10$, while the oscillator is a phase-locked loop (PLL) synthesizer with the loop bandwidth 1 MHz and the 3 dB bandwidth 100 kHz.

scans its surrounding environment using a laser device (lidar) or a camera. The primary goal is to detect specific features (*landmarks*) of the scenario and simultaneously estimate the mutual positions of the user and the landmarks. At first glance, this process may seem like a “chicken and egg” problem. Mapping, i.e., determining the position of landmarks, requires knowledge of the user’s position. Localization, i.e., determining the user’s position, requires knowledge of the map. However, this apparent dilemma can be generally resolved through appropriate algorithms. A crucial requirement for a monostatic SLAM algorithm to function effectively is that the user must be in motion to collect measurements from different perspectives. It is even more advantageous if the relative movements of the user can be measured through odometry or inertial devices. Surveys on the general SLAM problem can be found in [191], providing an overview of various techniques, including FastSLAM, GraphSLAM, and belief propagation SLAM.

In the classical setup, the (unknown) state of the system at the discrete time instant n consists of the state \mathbf{x}_n of the mobile user, encompassing factors such as its position, orientation, and speed, and the position of the landmarks \mathbf{m} representing the map. The objective is to determine the probability density function $p(\mathbf{x}_n, \mathbf{m} | \mathbf{y}_{1:n})$ (referred to as the *belief*) and deduce both \mathbf{x}_n and \mathbf{m} , where $\mathbf{y}_{1:n}$ denotes the set of measurements collected up to time instant n . To achieve real-time processing, Bayesian filtering approaches can be employed. In each iteration, $p(\mathbf{x}_n, \mathbf{m} | \mathbf{y}_{1:n})$ is computed, starting from the belief at time $n-1$ incorporating the new measurement \mathbf{y}_n and utilizing the statistical mobility model of the user, characterized by the probability density function $p(\mathbf{x}_n | \mathbf{x}_{n-1})$ [222].

The incorporation of ISAC functionalities into smartphones opens up the possibility of utilizing them as monostatic SLAM devices, employing radio signals for scenario scanning (Radio SLAM). This allows for infrastructure-less localization and the automatic generation of digital maps, all while safeguarding user privacy and minimizing energy consumption compared to lidar-based solutions that demand perfect visibility and manual operation [223]. Early investigations into Radio SLAM on handheld devices were previously referred to as *personal radar*

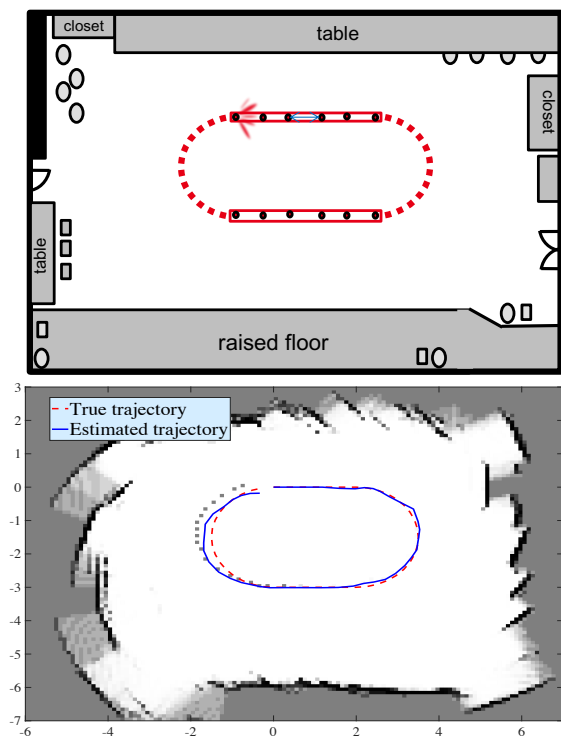


Fig. 28. Radio SLAM using measurements in the THz band. (top) Sketch of the measured indoor scenario; (bottom) Trajectory and map reconstruction using the MFM algorithm and an occupancy grid mapping method, respectively [225].

and were introduced in [224]. In Radio SLAM, landmarks are identified through specular reflections of signals emitted by the mobile user, which is assumed to be equipped with an antenna array.

From the perspective of signal processing algorithms, two main approaches are typically followed: i) Classical SLAM algorithms that rely on the joint estimation of landmarks and user positions, as described earlier; ii) Algorithms based on relative pose estimation by comparing measurements (scenes) collected in the last two time instants, \mathbf{y}_n and \mathbf{y}_{n-1} , respectively. In the latter approach, the goal is to estimate the position and orientation displacement of the mobile user, denoted as $\mathbf{d}_n = \mathbf{x}_n - \mathbf{x}_{n-1}$, at time n with respect to time instant $n-1$. A wide range of algorithms has been developed for this purpose, including the widely-used scan matching algorithm [226] and approaches borrowed from image processing, such as those utilizing the Fourier-Mellin transform [225], [227].

Nevertheless, the presence of multipath, diffuse reflections, and the side lobes' impact from the antenna array can introduce artifacts in the backscattered received signal, leading to the emergence of "ghost" landmarks. This complication renders the application of algorithms in Radio SLAM significantly more challenging compared to lidar-based or vision-based SLAM. Furthermore, at frequencies up to millimeter waves, specular reflections tend to dominate over diffuse reflections. Consequently, in a monostatic setup, the signal emitted by the user terminal is only reflected back to the user if it impinges the obstacle almost perpendicularly to its surface. Otherwise,

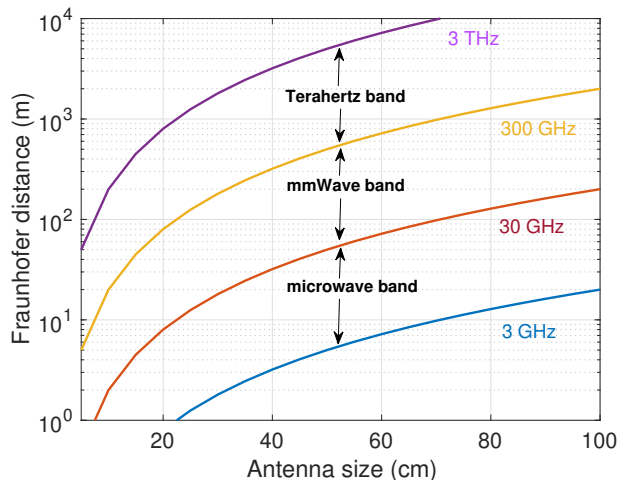


Fig. 29. Fraunhofer region boundaries as a function of the antenna size at different frequency bands. The area above each curve corresponds to the far-field region (Fraunhofer region), whereas the area below each curve corresponds to the near-field region.

the obstacle might remain invisible, posing a critical challenge to the SLAM process.

In this context, the capability to operate in the THz band is expected to play a crucial role in 6G systems. The wide bandwidth available in this range will result in high spatial resolution, and the feasibility of large antenna arrays will enable unprecedented angular resolution [228]. Moreover, at THz frequencies, the wavelength becomes comparable to the typical roughness of objects, potentially causing the diffuse component to dominate over the specular component. This aspect enhances object visibility even when the impinging angle of the signal is not perpendicular. Experimental investigations of Radio SLAM at THz frequencies are in their early stages. Initial results are reported in [225], where the performance of scan matching algorithms and a modified version of the Fourier-Mellin transform (MFM) is assessed using real-world THz radar measurements in an indoor environment. An example is provided in Fig. 28, where the performance of the MFM algorithm in terms of mapping and trajectory reconstruction is shown. The results were obtained starting from measurements taken at 300 GHz (lower THz frequency band) in a typical large office scenario by moving the measurement set up along a circular shape (red curve) and scanning the scene with a span of $\pm 90^\circ$ with respect to the direction of movement and a beamwidth of 18° . The obtained localization RMSE is 12 cm. More details on the algorithms and the measurement campaign can be found in [225].

VI. SENSING AND COMMUNICATING WITH WIDE APERTURES

A. Introduction

In contemporary wireless systems, antenna arrays are typically adopted to obtain beamforming and spatial multiplexing. The former involves directing the transmitted signal precisely toward its designated receiver in space, whereas the latter consists in simultaneously transmitting multiple data streams

to the same user or users located at different positions. These are the core functionalities of (massive) MIMO systems. As a natural evolution of the massive MIMO technology, extremely large-scale MIMO (XL-MIMO) further boosts the number of antennas by at least an order of magnitude, e.g., several hundreds or even thousands of antennas, thus obtaining electrically large aperture arrays (ELAAs) and unprecedentedly improving the spectral efficiency and spatial resolution for wireless communication and sensing [229]–[231]. In the following, we will refer to an ELAA as an antenna whose dimension D is much larger than the wavelength λ (electrically large antenna). Depending on the technology adopted and the way it is modeled, an ELAA may be dubbed as Holographic MIMO and large intelligent surface (LIS) [232]–[234]. Typically, these terms are applied when arrays consist of electrically small and densely packed elements.

B. Near-field vs. far-field communication and sensing

The exploitation of higher frequency bands, from millimeter waves to THz, foreseen in 6G networks and the availability of new antenna technologies for ELAAs opens new opportunities for communication and sensing [231], [233]–[236]. In fact, traditional wireless systems typically work at distances beyond the Fraunhofer distance, defined as $d_F = 2D^2/\lambda$ (*far-field boundary*), where the electromagnetic (EM) wavefront can be well approximated as being planar. Instead, using an ELAA, the wireless links are likely to operate at distances below d_F , corresponding to the radiating *near-field region*, where such an approximation no longer holds and, consequently, the wavefront impinging on the antenna is spherical [237]. While this aspect requires a revisitation of the classical channel models that in many cases are based on the far-field assumption and then may fail in the near field, at the same time, it provides the chance to improve the communication and sensing capabilities of the system [230], [238]. In Fig. 29, the Fraunhofer distance is plotted as a function of the antenna aperture size D for different frequency bands ranging from microwave to THz. As it can be noticed, at millimeter waves and beyond, the radiating near-field region might correspond to practical operating distances of several meters or hundred meters even with relatively physically small (but electrically large) antennas.

Regarding communication, a spherical EM waveform is more informative than a plane wave, as it will be discussed in Sec. VI-C. For sensing/positioning, incident spherical wavefronts embed not only angular information, as in the far-field regime, but also distance information. This property can be exploited to determine the position of a transmitting source by analyzing the phase profile of the received signal along the antenna aperture, as it will be detailed later [236].

C. Communication and sensing with ELAAs

The adoption of an ELAA provides improved flexibility in forming the beam and sensing the EM wave (see Fig. 30). While in the far field, only differently steered beams can be formed, thus providing user discrimination only at the angular level, in the near-field region the beam can be focused on a specific location similarly to an optical lens [239]. It not only allows for

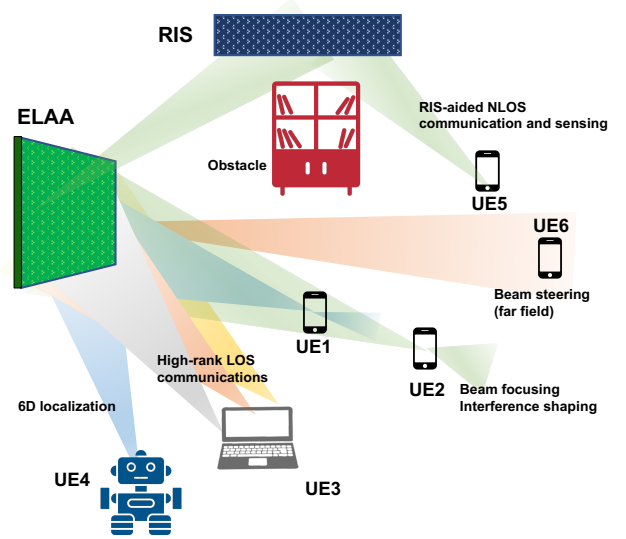


Fig. 30. Near-field communication and sensing: UE1 and UE2 can be discriminated even if they are at the same angle of view thanks to focusing (interference shaping); UE3 can establish a high-rank communication link (spatial multiplexing in LoS); the position and orientation of UE4 can be estimated (single-antenna 6D localization); UE5 can be localized/sensed in NLoS thanks to the RIS; UE6 is in the far-field region, then only rank 1 communication and beam steering can be realized in strong LoS conditions.

the control of multi-user interference (interference shaping) in terms of angular direction, similar to traditional beam steering but also provides control over interference in terms of distance thus allowing the discrimination of users seen with the same angle of view [240].

The near-field effect can be exploited to combat the multiplexing gain degradation caused by sparse multipath channels in MIMO links when operating in strong LoS conditions at high frequencies. In principle, it is possible to establish orthogonal channels (*communication modes*) by generating non-overlapping beams, each one focusing on different properly chosen locations of the receiving antenna (see UE3 in Fig. 30). A rough estimate of the number of communication modes, i.e., communication degrees of freedom (DoF), can be calculated starting from simple arguments of diffraction theory [241]. Consider two antennas, modeled as continuous surfaces for convenience, in parallax configuration with areas A_T and A_R , respectively, at distance d , as shown in Fig. 31. The smallest spot of area a we can use at the transmitting antenna will be the one from which the diffracted EM field approximately fills the aperture of the receiving antenna. In particular, the diffraction solid angle from the spot is $\Omega \approx \frac{\lambda^2}{a}$, where it must be $\Omega d^2 \approx A_R$ (full illumination). The number of distinct spots on A_R , i.e., communication modes, is

$$\text{DoF} \approx \frac{A_R}{a} = \frac{A_T A_R}{d^2 \lambda^2}. \quad (39)$$

The previous result confirms that with ELAA, i.e., $A_T/\lambda^2 \gg 1$, it is possible to obtain high-rank communications even in strong LoS conditions. A more accurate theoretical bound on DoF valid also for asymptotically large ELAA modeled as a LIS can be found in [234]. Achieving capacity-approaching MIMO communications in the near field results in intricate array

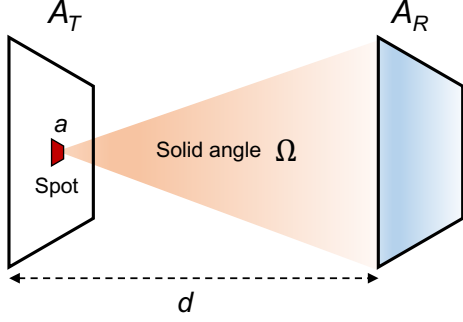


Fig. 31. Diffraction effect at distance d of a small spot of area a composing the transmitting surface.

phase profiles. As indicated in [242], these profiles can be effectively approximated by using a combination of simpler multiple-focused beams whose configuration depends on the geometry.

Regarding the localization task, as anticipated it is possible to localize a source or a target using only one ELAA if it is in the antenna's near-field region through the analysis of the phase profile of the impinging EM wave [243], [244].

Example 8. Consider a simple 2D scenario where a receiving linear ELAA located in the origin is composed of N elements deployed along the x -axis spaced apart of $\Delta = \lambda/2$ with coordinates $\mathbf{r}_n = [(n-1)\Delta, 0]$, for $n = 1, 2, \dots, N$. The size of the ELAA is $D = N\Delta$, with $N \gg 1$. A punctiform monochromatic RF source located at position $\mathbf{p} = [p_x, p_y]$ in the near field of the ELAA is present (e.g., an active node to be localized or a reflecting target in case of sensing). Denote with $d = \|\mathbf{p} - \mathbf{r}_1\|$ the distance between the first element of the ELAA and the source, and by θ the steering angle so that $\sin(\theta) = p_x/d$ and $\mathbf{p} = [d\sin(\theta), d\cos(\theta)]$. The complex baseband channel between the source and the ELAA can be modeled as

$$h_n = \beta_n e^{-j\frac{2\pi}{\lambda}d_n} \quad (40)$$

where β_n represents the channel amplitude and

$$d_n = \|\mathbf{p} - \mathbf{r}_n\| = d\sqrt{1 + \frac{(n-1)^2\lambda^2}{4d^2} - \frac{2p_x(n-1)\lambda}{2d^2}} \quad (41)$$

is the Euclidean distance between the source and the n th receiving antenna element. if $D < d$, then $\beta_n \simeq \beta_1 \forall n$ and we can use in (41) the following Taylor expansion $\sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + o(x^2)$ leading to

$$d_n \simeq d - \frac{(n-1)\lambda \sin(\theta)}{2} + \frac{(n-1)^2\lambda^2}{8d}. \quad (42)$$

The first term of (42) contributes in (40) with a constant phase shift which is not informative if $d > \lambda$ because of the 2π phase ambiguity and the need for the source to be perfectly synchronized with the antenna. The second and third terms indicate a parabolic behavior of the phase profile observed along the array which is a function of the steering angle and the distance, i.e., the position \mathbf{p} of the source, which can be therefore estimated. If $d > d_F$ (far-field condition), the third term contributes with a phase shift smaller than $\pi/8$ at the edge

of the antenna and hence becomes negligible. As a consequence, $d_n \simeq d - (n-1)\lambda \sin(\theta)/2$ and only the angle of arrival θ can be estimated.

The case where the source employs an antenna array opens the door to 6D positioning, i.e., the estimation of the source's position and orientation if located in the near-field region of the ELAA. (see UE4 in Fig. 30). The theoretical performance bound of 6D positioning has been derived in [243] for generic shapes of the antenna arrays and wideband signals.

It has to be remarked that, single-antenna positioning is possible also in the far field but that would require, in addition to the signal's AoA measurement, the exchange of wideband signals to estimate the distance between the source and the antenna through ToA measurements and a ranging protocol to cope with the lack of a common clock. On the contrary, with an ELAA, localization, and sensing can be obtained with a narrowband signal and lower latency because no ad-hoc synchronization procedures and time-base ranging protocols are needed. As a consequence, more resources become available for communication which allows for more efficient joint communication and sensing schemes.

A deep investigation of near-field integrated sensing and communication is still missing in the literature. In this direction, [245] derives the minimization of the Cramér-Rao bound for the near-field joint distance and angle sensing subject to the minimum communication rate requirement of each user. In the same work, both fully digital antennas and hybrid digital and analog antennas are investigated. The authors in [246] consider a multi-static OFDM-based ISAC scenario by proposing a low-complexity two-stage estimation process where first a rough estimate of the target is obtained assuming a far-field condition, then refined accounting for the correct near field model.

The primary challenges in modeling and designing joint communication and sensing functionalities using ELAAs can be summarized as follows: i) *Near-field channel modeling*: The extremely large array aperture introduces spatial non-wide sense stationary properties. Different regions of the array observe the propagation environment from varying perspectives, with diverse polarizations. This implies that regions may perceive signals transmitted along a specific propagation path with differing powers and polarizations, or signals from distinct propagation paths [247]. ii) *Channel estimation*: Achieving optimal transmission demands highly accurate channel estimation. Unfortunately, the vast number of elements constituting the ELAA generally entails estimating a considerable number of channel parameters, particularly in challenging propagation environments, leading to increased signaling overhead. In traditional MIMO systems, exploiting the sparsity of the angular-domain channel (e.g., through compressed sensing) simplifies this task. However, near-field propagation exhibits sparsity in the location domain due to different electromagnetic characteristics. Consequently, existing angular-domain-based algorithms are not directly applicable. For instance, in [248], the authors proposed a polar-domain representation of the channel that comprehensively captures near-field spherical wave characteristics. Differently, in cases of strong LoS, the channel is closely linked to geometry, specifically the source's

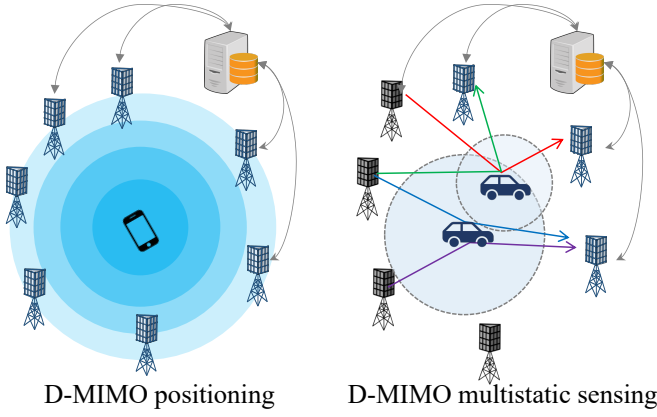


Fig. 32. D-MIMO for positioning (left) and multistatic sensing (right). Phase-coherent processing leads to wavefront curvature, considering all RUs as part of a distributed array. Black RUs are transmitters, while blue RUs are receivers.

position. In principle, only three parameters are sufficient for estimating the channel. Consequently, channel estimation aligns with the localization task, further intertwining sensing and communication. iii) *ELAA Modeling*: Depending on the technology adopted, if the ELAA comprises dense elements, mutual coupling effects between them cannot be overlooked and must be accurately modeled [230].

D. Distributed MIMO/cell-free massive MIMO for multi-perspective localization and sensing

A distributed multiple-input multiple-output (D-MIMO) system can be viewed as a massive MIMO system, where a large number of antennas are deployed over a wide area with irregular inter-element spacing (see Fig. 32) [249]. Each antenna (or small array) is called a radio unit (RU), several of which may be connected to a distributed unit (DU), which takes care of most of the baseband processing [250]. The DUs are connected to central unit (CU), which manages higher layers and coordinates the DUs. Various D-MIMO architectures correspond to different node topologies, different levels of local versus central processing, and different levels of synchronization [251]. The operation is otherwise similar to classical massive MIMO, involving uplink pilot transmission, downlink multi-user beamforming, and uplink data transmission [252].

Typical D-MIMO assumes phases synchronization among the RUs, enabling spatial focusing towards UEs [249], [253]. While for communication purposes, the locations of RUs are irrelevant, and phase stability during each coherence interval suffices, more stringent requirements exist for localization and sensing functions. In these cases, considering all D-MIMO RUs as a large array, the overall array steering vector needs to be known, which requires not only phase stability, but also precisely known locations of the RUs (within a fraction of the wavelength) [90], [254], [255]. Combined, this places UEs and objects in the near-field of a massive distributed array, while remaining in the far-field of individual elements [90]. Meeting these stringent requirements is likely only feasible at FR1 [256, Sec. 6.2.1], [257], whereas at higher frequencies (FR2 and

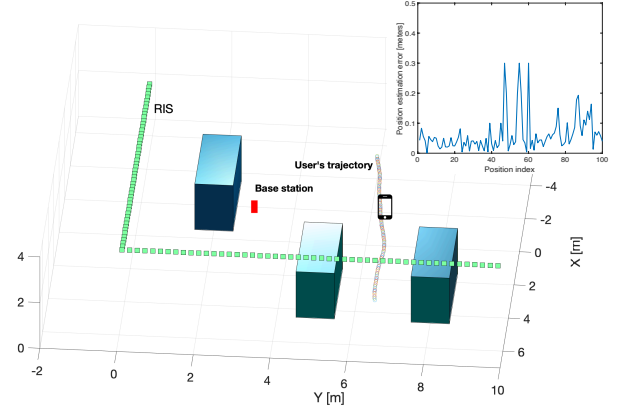


Fig. 33. RIS-assisted localization in a near-field NLoS scenario. 28 GHz OFDM system with 250 MHz bandwidth. The error evolution along the trajectory is reported [259].

above), classical time-coherent processing can be performed. At FR1, the operation of D-MIMO bears resemblance to carrier phase positioning (see Section IV-B9), with multistatic sensing (see Section IV) and with ELAA (see Section VI-C). The main differences lie in the large, dense, and distributed architecture of D-MIMO.

The main benefits of D-MIMO include more uniform coverage for communication, localization, and sensing [249], [255]. In addition, the large aperture provides enhanced multipath resolution and improved accuracy, though at the expense of high computational complexity [258].

E. Sensing and localization aided by large RIS

The use of RIS for localization was discussed in Section IV-B8, where the UE was in the far field of the RIS, providing additional delay and angle information. When the RIS becomes large and the UE is in the near-field region of the RIS, the model (24) changes to [259]–[262]

$$\alpha_m^{\text{RIS}} = \alpha^{\text{T-RIS}} \alpha^{\text{RIS-R}} \mathbf{a}_{\text{RIS}}^{\text{T}}(\mathbf{p}) \Omega_m \mathbf{a}_{\text{RIS}}(\varphi). \quad (43)$$

This means that a single-antenna UE, observing $y_m = \alpha_m^{\text{RIS}} x_m + z_m$ over a single subcarrier across different times m , can in principle estimate its 3D location \mathbf{p} , even when the LoS path between the UE and BS is blocked [261]–[264], or when some of the RIS elements are failing [265]. In (43), $\mathbf{a}_{\text{RIS}}(\mathbf{p})$ represents the near-field response vector as a function of the UE position. Fig. 33 illustrates a practical example of NLoS localization facilitated by RIS. In this scenario, a single-antenna UE navigates within an indoor setting in the presence of several obstacles, with a BS emitting a reference OFDM signal. A long linear RIS is deployed (green patches in the figure), allowing the UE to determine its location by analyzing only the signals reflected by the RIS. In this specific setup, as proposed in [259], the reflection coefficient of each element in the RIS matrix Ω_m dynamically varies during the transmission of the reference signal. This variation follows a predefined pattern designed to enable the UE to estimate the delay of the signal component reflected by each visible element of the RIS. Utilizing these BS-RIS-UE delay measurements, the UE can then compute its own

position. The inset plot demonstrates that, despite numerous obstacles partially shadowing the RIS, the localization error along the trajectory of the UE remains confined to 20 – 30 cm. Additionally, a narrowband version of the algorithm, leveraging the phase profile of the received signal, is proposed in [259]. For sensing, the model (43) can be used, for example, in a monostatic setup with a single-antenna BS monitoring an area and making use of the high spatial resolution of the large RIS [266]–[268]. By generating RIS configurations Ω_m that scan an area, targets can be detected and localized, even with limited bandwidth, from the observations at the BS, harnessing the paths from BS to RIS to target, back to RIS and back to the BS. The near-field imaging problem utilizing XL-MIMO antennas and RIS has been investigated in [269] where the design of the optimal illumination waveform and RIS configuration is addressed.

VII. TECHNOLOGIES FOR SENSING ASSISTED COMMUNICATION

MIMO communication operating at mmWave with large arrays and bandwidths provides the angular and delay resolution required for high accuracy localization and sensing, as we have discussed in previous sections. Resilient communication becomes more challenging, however, as cellular networks advance to higher carrier frequencies. On the other hand, fast adaptation of communication strategies is more difficult due to large MIMO arrays, higher bandwidths and challenging circuit designs. For example, initial access to configure the mmWave beams in 5G [270] can be up to 5s with a simple analog beamforming architecture, with more time expected at sub-THz frequencies [271]. On the other hand, there are fewer opportunities for high-rate communication due to the larger performance differential between LOS and NLOS links, more frequent blockage (smaller first-order Fresnel zone), higher penetration losses, increased scattering and less reflection [16]. For example, commercial 5G mmWave throughput measurements with a blocked LOS path show a $2\times$ throughput reduction when reconfiguring the beamforming to use an available NLOS path and a $4\times$ throughput reduction when the device switches to a lower frequency in 4G because no NLOS path is available [272].

Sensing-assisted network operation can mitigate the impact of these issues. For example, ISAC-maps can incorporate localization information about the environment and about the success of past communication configurations to design position dependent beam codebooks with a reduced number of beams, as explained in Section II-E and illustrated in Fig. 5(c). In addition, Fig. 5(c) shows how ISAC maps can also provide the position and velocity of potential blockers of the communication signal, so the network can proactively react to mitigate the impact of blockage. In this section, we describe specific approaches for adapting communication operation based on the exploitation of the ISAC map. In particular, we address the problems of sensing-aided array configuration and blockage prediction and management.

A. Sensor-aided array configuration

The earliest prior work on enhancing network operation in an ISAC setting considers the problem of mmWave MIMO beam training for initial access aided by some type of sensing information and past communication. The problem is to find the transmit beam \mathbf{f} in a codebook \mathcal{F} and the receive beam \mathbf{w} in a codebook \mathcal{W} to maximize a given performance metric, for example signal-to-noise ratio (SNR). The standard approaches are a brute-force search over pairs [273] or a hierarchical search [274] over increasingly refined codebooks. In both cases, search time generally grows with antenna size but it can be reduced by searching (intelligently designed) smaller codebooks. Achieving this reduction is especially important in vehicular settings, where the channel is highly dynamic and frequent antenna array reconfiguration is required. In the next paragraphs we will show examples of how sensing and learning can be used together to make communication more efficient.

1) *Position-aided beam training*: Initial work exploited position information obtained with a GPS [275] or a radar mounted at the BS [276] to reduce the size of the beam codebook. The basic idea consists of using the estimated direction of the user ϕ_{UE} suggested by its position to define a reduced set of beams to be tested [276]. This reduced beam codebook $\mathcal{W}_{\mathcal{L}}$ can be computed to account for the estimation error in the angular direction of the user, denoted as $\Delta\phi_{\text{UE}}$. This way, the reduced codebook will include the beams from an initial grid indexed by the index set \mathcal{L} such that $n \in \mathcal{L}$ if

$$\sin(\hat{\phi} - \Delta\phi) + 1 \leq \frac{2n}{N_{\text{T}}} \leq \sin(\hat{\phi} + \Delta\phi) + 1 + 2/N_{\text{T}}. \quad (44)$$

With this type of strategy, it is possible to achieve an overhead reduction of $(1 - |\mathcal{L}|/N_{\text{T}})$ [276]. Newer strategies that exploit location information obtained via radio positioning have also been proposed [277]. However, all these solutions are only feasible in LOS settings.

More elaborated approaches exploiting position and also ML can achieve a significant overhead reduction in both LOS and NLOS channels. For example, inverse fingerprinting learns a subset of location-dependent beam-pairs based on past measurements in similar locations, such that with high probability at least one of the vectors in the subset works well [27], [28], [278]. The recommendation algorithm is trained based on past measurements of the strength of different beam pairs as a function of location, made under different snapshots of the environment [28]. The recommendations may be refined using online learning, by further exploring the angular space to achieve better beam pointing [27]. Additional information may also be used, like the traffic density, to make more accurate predictions [278].

2) *Radar-aided beam configuration*: For NLOS settings and a hybrid architecture, it is possible to design the precoders and combiners from the full instantaneous channel state information (CSI) or from the spatial channel covariance information (also known as statistical CSI, partial CSI, or imperfect CSI). However, estimating either the channel or the spatial covariance introduces a significant overhead [46], [57], [279]. The spatial similarity between the radar and communication channels have been studied in [32], [280], [281] via ray tracing

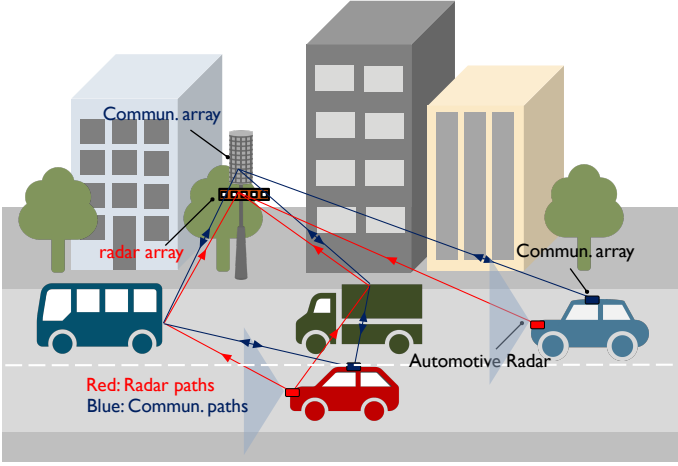


Fig. 34. Illustration of a vehicular communication system with radar-aided beam training. There is similarity between the radar and communication main channel directions, which can be exploited to reduce beam training overhead in NLOS and multiuser scenarios.

simulation and also experimental measurements. These studies show that the similarity in the main channel directions—in other words similarity in spatial covariance—between radar and communication channels operating at close but different mmWave frequencies is high, so that the radar covariance can be used as a prior for the communication covariance.

The first work on radar-aided spatial mmWave link configuration proposes the estimation of the spatial covariance in a V2I link with an active radar mounted at the BS [32]. An alternative design in [281] considers a passive radar receiver at the BS, which is listening to the automotive radar signals coming from the different vehicles on the road, as illustrated in Fig 34. Some of the vehicles are already connected to the BS, while others need to join the network. Even in the absence of the chirp reference signal used by the automotive radar in the vehicles, it is possible to estimate the radar covariance at the BS radar receiver. Using this information, [281] shows that communication overheads can be reduced by 77% in a realistic vehicular environment simulated by ray tracing. Though the results are promising, there is still a mismatch between radar and communication channels due to using different frequencies or different locations of radar modules and transceivers in the vehicle, as illustrated in Fig 34. Some subsequent work [30], [282] introduces multiple users into the environment and designs a deep learning strategy that translates the radar covariance into a communication covariance to compensate for these mismatches, achieving further overhead reductions which result in higher effective rates as shown in Example 9.

Example 9 (Deep learning based radar assisted link configuration). We consider a radar-aided communication system as the one illustrated in 34. The parameters of the radar and communication signals are specified in Table IV. The system is simulated with a number of 4 users on the road. Additional details of the urban environment simulated by ray tracing can be found in [30], [281]. The considered performance metric is the sum rate, defined as $R_\Sigma = \sum_u R_u$, with the effective rate

TABLE IV
SIMULATION OF A RADAR-AIDED MMWAVE VEHICULAR COMMUNICATION SYSTEM: PARAMETERS FOR THE COMMUNICATION SIGNALS AND THE AUTOMOTIVE RADAR SIGNALS.

Parameter	Symbol	Value	Units
Communication system			
Transmit power	P_c	30	dBm
Carrier frequency	f_c	73	GHz
Bandwidth	B	1	GHz
BS height		5	m
Vertical separation of arrays at the BS		10	cm
Distance BS to closest point on the road	d	10	m
Number of antennas at the BS		128	
Number of RF chains at the BS		1	
Number of antennas at the vehicle		16	
Number of arrays at the vehicle		4	
Number of RF chains at the vehicle		1	
Height of the communication arrays at the vehicle		1.6	m
Number of phase shifter bits	D	2	bits
Number of subcarriers	N	2048	
Subcarrier spacing	Δf	240	kHz
Cyclic prefix length	L_c	511	samples
Radar system			
Center frequency	f_r	76	GHz
Bandwidth	B_r	1	GHz
Transmit power		30	dBm
Number of antennas		128	
Chirp period		500	μs
Samples per chirp		1024	samples
Height of the vehicle radars		0.75	m

per user R_u

$$R_u = (1 - \frac{T_{train}}{T_{coh}}) \Delta f s_u, \quad (45)$$

with Δf the subcarrier spacing and s_u the spectral efficiency for the u -th link.

$$s_u = \sum_{k=1}^N \log_2 (1 + SINR_u[k]), \quad (46)$$

with N the number of subcarriers. The initial estimation of the radar covariance matrix at the radar arrays of the BS is performed using the multiuser separation and covariance estimation strategies designed in [30], [281]. A fully connected network trained with a set of 9600 samples and evaluated with a test set of 2400 samples is used to map the radar covariance to the communication covariance. Radar aided beam training is implemented using a reduced codebook composed of 4 beams around the directions suggested by radar. Fig. 34 shows the sum-rate for exhaustive search, radar covariance-aided beam training without compensating for mismatches, and communication covariance-aided beam training after ML-based mapping of the radar covariance to the communication covariance. The performance gain of radar-aided beam training strategies varies with the coherence time, but it is significant in all cases. The benefit of ML-based mismatch prediction is higher at smaller coherence times.

Other recent strategies for radar-aided communication exploit the radar information obtained with a joint radar and communication system to speed up beam training. For example, [283] considers the beam tracking stage of a vehicular communication system, exploiting a BS operating as a radar to estimate

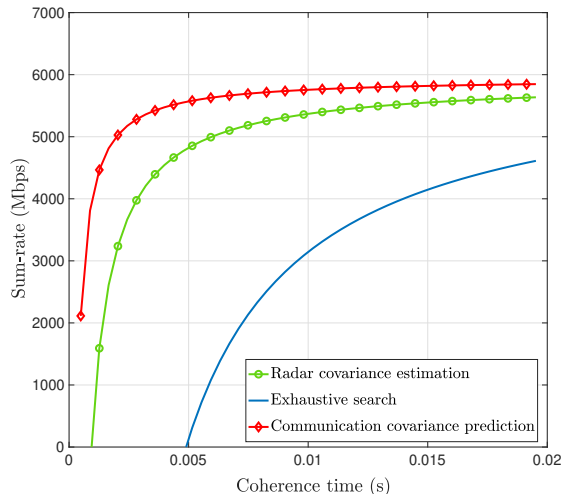


Fig. 35. Sum-rate versus coherence time for different beam training strategies including exhaustive search (blue), radar covariance-aided beam training without compensating for mismatches (green), and communication covariance-aided beam training after ML-based mapping of the radar covariance to the communication covariance (red).

the angle, distance and velocity of a given vehicle. The beamformers are then updated based on the predicted angle for the communication link, providing a relevant performance gain with respect to other benchmarks. In a similar setting, [284] proposes a variation of this idea, exploiting the delay and Doppler parameters estimated at the BS for the link corresponding to a particular vehicle. Then, a message passing algorithm based on factor graphs is derived to estimate the unknown range, speed, angle of arrival and path loss, to finally leverage the angular information for beam tracking as in [283].

3) *Vision-aided beam training*: The images obtained from cameras can be exploited to infer decisions related to beam selection for mmWave systems, which is motivated by the recent development of deep learning powered computer vision methods such as object detection and 3D scene reconstruction [285]. Moreover, the advancement of camera technology allows the adoption of high resolution cameras at both BSs and UEs at a low cost. The features extracted from the images may indicate the locations of the UEs and reflectors which would also help determine the existence of the LoS path. The images can be fused with other information, such as the position of the UEs, to enhance the beam selection performance. Furthermore, images can be useful to track the users in dynamic environments, which would be beneficial for beam prediction. Vision-aided beam training methods have the potential to reduce or even completely mitigate the beam training overhead incurred on the communication system.

One of the interesting works considers collecting multiple images in the vicinity of the mmWave BS [286]. The collected images are used to reconstruct the 3D environment which is the input of a deep neural network along with the position of the UE. The output of the network indicates the selected beam index. The takeaway from this work is that the reconstructed 3D environment contains spatial information such as the locations

of the scatterers. In [285], the authors study the beam prediction problem leveraging previous beam indices and corresponding images in a dynamic vehicular scenario where the roadside units have cameras to the BSs. The ray tracing-based simulation results show that several future beams can be accurately predicted with various deep learning architectures such as LSTM. Since these algorithms utilize deep learning methods, it is imperative to have a dataset that contains both communication channels and corresponding images. Although one such dataset is provided in [287], where the authors obtained the images of the scenes from ray tracing environments, creating datasets in realistic environments is an open challenge.

Another approach is to exploit the cameras mounted on the UEs, which is more suitable to vehicular settings. It is shown that it is possible to detect the surrounding objects and vehicles if the images obtained from the cameras mounted on an autonomous vehicle are exploited in [288]. This approach not only provides the beam decisions, but also predicts the duration of the beam coherence interval. The simulation results obtained by using a 3D modeling and a ray tracing software show that utilizing the images taken at the UE is more beneficial than leveraging the cameras at the BS.

4) *Exploiting LIDAR point clouds and multimodal information*: Sensors like LIDAR or IMU have also been leveraged to reduce beam training overhead [29], [289]–[291]. The exploitation of these sensors is particularly interesting in cellular networks supporting vehicles or robots. The different beam training strategies aided by LIDAR use the LIDAR point cloud to identify potential obstacles in the environment. Proposed designs create different features from the LIDAR point cloud and other potential data. For example, in [29], this feature is created from the positions of the BS and UE, the coverage area and the LIDAR point cloud. This information is the input to a convolutional network that decides about the channel state (LOS or NLOS) and recommends a reduced set of beams for beam training. Subsequent works introduced innovations that improve performance in complex NLOS cases. For example, the approach in [289] considers a curriculum learning strategy that trains with LOS measurements first, and then gradually introduces NLOS observations, resulting in an improved beam classification accuracy. In [291], the main novelty is the introduction of a federated learning framework such that a set of connected vehicles use their local LIDAR data to fine tune a shared neural network for beam selection. Since the same network can be utilized by vehicles that join the network later, the overhead is further reduced. LIDAR information is also commonly exploited in a multimodal fashion, combined with position information from GPS, images from cameras, or depth maps [292]–[295].

B. Blockage prediction and management

Wireless communication systems often experience blockages that could lead to unreliable communication links, even link failure. This problem is more pronounced at mmWave and THz bands where the LoS path is usually much stronger than the other paths. Work has been done to characterize the impact of blockage, especially at millimeter wave frequencies. There

have been many measurements of pathloss in the presence of blockage and specific studies describing the impact of a blocked link. For example, pathloss for LOS and NLOS in urban areas can be characterized by different exponents [296] possibly with additional corner losses [297]. In terms of specific blocking objects, vehicles may incur 20 dB of attenuation [298], people 40 dB [299], and hand 20 dB [300]. The impact the blockage and the duration depends critically on the mobility of the transmitter, receiver, and blockage as well as the corresponding distances, with impact generally increasing with frequency [298]. It seems likely the trend will continue for frequencies above 100 GHz.

In terms of blockage management, 5G NR primarily relies on high BS density to facilitate macro-diversity and increase the probability that at least one link is unblocked. Its operation is primarily reactive, e.g. repeating the beam training phase after a link failure. In particular, it does not make special accommodations for the impact of blockage despite the potential for harvesting a tremendous amount of environmental information that could be available in a network with integrated sensing and communications.

One of the earliest works on sensor-aided blockage management utilizes RGB and depth (RGB-D) cameras that are located close to the mmWave BSs to predict the human blockages [301]. This work employs a centralized proactive BS selection algorithm which is shown to be effective with experimental results. The work in [302] utilizes a dual-band BS supporting both sub-6GHz and mmWave frequencies at the roadside unit of a vehicular communication system. Additionally, the roadside unit is equipped with a camera to improve the blockage detection performance. The image feed provided by the camera and the CSI from the sub-6GHz system are used as the input of a deep neural network for blockage detection. Ray tracing-based simulation results show that the blockages can be detected with high accuracy. In [303], a bimodal machine learning algorithm that makes use of the temporal beam index sequence and the image feed from a camera equipped at the BS is used to predict future blockages with the help of a deep learning model that detects objects in the images. Furthermore, this work also describes a BS handover framework that is shown to be effective in ray tracing-based simulations. The work in [304] utilizes the GPS information at the vehicles and images provided by the camera at the mmWave BS to detect blockages. A feature extraction and a fusion network are used to process the vehicle location and camera images to detect the blockages at the edge.

LIDAR is another sensor that can provide physical information related to the environment, which can be useful for blockage prediction as shown in [305] where a LIDAR is used to scan an indoor environment. Trajectories of the users are predicted with an LSTM network. Then, the trajectories are used to construct the 3D environment that can be used to trace the rays, referred as ray casting, which are converted to the human blockage predictions. Similarly, the work in [306] proposes to convert the LIDAR data to 3D scenes. The created scene is used for the ray tracing simulations to obtain the path information between the BS and the UE, which is converted to an initial channel estimate and blockage detection

without any communication overhead. In [295], an experimental setup with a dual-band BS unit that is also equipped with a camera and a LIDAR. The authors use the camera images if the environment is bright, and the LIDAR output if it is dark. Object detection algorithms and a recurrent neural network are utilized for blockage prediction. The experimental results show that successful frequency handover is achieved with the described system. In [307], a LIDAR is equipped alongside a mmWave BS at the roadside unit of a vehicular communication system. The LIDAR data and power sequence, i.e., received power sequence obtained with the transmitted beams, are used as the inputs of a deep neural network to predict blockages. The experimental results verify the effectiveness of the considered approach.

Finally, the work in [308] uses range-angle maps obtained with the radar at the roadside unit. The range-angle maps are utilized as the input of a LSTM network to predict blockages. The authors show the effectiveness of the proposed approach with an experimental setup.

All this recent work shows that blockage detection and proactive blockage prediction can be achieved with the aid of sensors in mmWave systems. Moreover, machine learning-based solutions are shown to be effective for converting the data obtained from the sensor(s) to blockage detection and/or prediction.

VIII. CONCLUSION

Cellular networks are integrating sensing and communication functionalities. This integration enables new sensing services and also provides data which can be exploited to enhance network operation. This article has presented a vision of the ISAC cellular network and a comprehensive overview of methods for network sensing and also sensing-aided communication. Regarding technologies for network sensing, we have introduced strategies for bistatic and multistatic sensing—including positioning as a special case—, monostatic sensing, and also specific designs to operate with wide apertures. We have also described approaches for radio SLAM in bistatic and monostatic scenarios. In the overview of the opposite setting, network operation assisted by sensing, we have focused on the problems of sensor aided beam training for overhead reduction and blockage prediction and management. We have introduced multiple examples which illustrate the performance that these different technologies can achieve in practical scenarios. We have made the case that communication and sensing must be considered together and that this integration will be one of the key features of 6G.

IX. ACKNOWLEDGEMENTS

The authors would like to thank Dr. Satyam Dwivedi from Ericsson research for his invaluable feedback and discussions related to the evolution of sensing services in cellular networks and the current discussions on sensing use cases and prioritizations in the context of 3GPP standardization for 6G.

REFERENCES

- [1] "Shared Spectrum Access for Radar and Communications (SSPARC)," 2016. [Online]. Available: [\url{https://www.darpa.mil/program/shared-spectrum-access-for-radar-and-communications}](https://www.darpa.mil/program/shared-spectrum-access-for-radar-and-communications)
- [2] G. Tavik, C. Hilterbrick, J. Evins, J. Alter, J. Crnkovich, J. de Graaf, W. Habicht, G. Hrin, S. Lessin, D. Wu, and S. Hagewood, "The advanced multifunction RF concept," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 3, pp. 1009–1020, 2005.
- [3] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 6, pp. 1728–1767, 2022.
- [4] "Roadmap to 6G," 2022. [Online]. Available: [\url{https://roadmap.nextgalliance.org/}](https://roadmap.nextgalliance.org/)
- [5] "European vision for the 6G network ecosystem," 2021. [Online]. Available: [\url{https://5g-ppp.eu/wp-content/uploads/2021/06/WhitePaper-6G-Europe.pdf}](https://5g-ppp.eu/wp-content/uploads/2021/06/WhitePaper-6G-Europe.pdf)
- [6] "Industry specification group (ISG) integrated sensing and communications (ISAC)," 2023. [Online]. Available: [\url{https://www.etsi.org/committee/2295-isac}](https://www.etsi.org/committee/2295-isac)
- [7] H. Wymeersch, S. Saleh, A. Nimr, R. Halili, R. Berkvens, M. H. Moghaddam, J. M. Mateos-Ramos, A. Stavridis, S. Wänstedt, S. Barm-pounakis, B. Priyanto, M. Beale, J. van de Beek, Z. Ye, M. Manalastas, A. Kousaridas, and G. P. Fettweis, "Joint communication and sensing for 6G – a cross-layer perspective," 2024.
- [8] A. Hassanien, M. G. Amin, E. Aboutanios, and B. Himed, "Dual-function radar communication systems: A solution to the spectrum congestion problem," *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 115–126, 2019.
- [9] K. V. Mishra, M. B. Shankar, V. Koivunen, B. Ottersten, and S. A. Vorobyov, "Toward millimeter-wave joint radar communications: A signal processing perspective," *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 100–114, 2019.
- [10] D. Ma, N. Shlezinger, T. Huang, Y. Liu, and Y. C. Eldar, "Joint radar-communication strategies for autonomous vehicles: Combining two key automotive technologies," *IEEE Signal Process. Mag.*, vol. 37, no. 4, pp. 85–97, 2020.
- [11] F. Liu, L. Zheng, Y. Cui, C. Masouros, A. P. Petropulu, H. Griffiths, and Y. C. Eldar, "Seventy years of radar and communications: The road from separation to integration," *IEEE Signal Processing Magazine*, vol. 40, no. 5, pp. 106–121, 2023.
- [12] A. Ali, N. González-Prelcic, R. W. Heath, and A. Ghosh, "Leveraging sensing at the infrastructure for mmWave communication," *IEEE Communications Magazine*, vol. 58, no. 7, pp. 84–89, 2020.
- [13] H. Holma, H. Viswanathan, and P. Mogensen, "Extreme massive MIMO for macro cell capacity boost in 5G-Advanced and 6G," *Nokia, Tech. Rep.*, 2021.
- [14] M. A. Saeidi, H. Tabassum, and M.-S. Alouini, "Multi-band wireless networks: Architectures, challenges, and comparative analysis," *IEEE Communications Magazine*, vol. 62, no. 1, pp. 80–86, 2024.
- [15] S. Aboagye, M. A. Saeidi, H. Tabassum, Y. Tayyar, E. Hossain, H.-C. Yang, and M.-S. Alouini, "Multi-band wireless communication networks: Fundamentals, challenges, and resource allocation," *IEEE Transactions on Communications*, pp. 1–12, 2024.
- [16] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436–453, 2016.
- [17] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic, and R. W. Heath, "MIMO precoding and combining solutions for millimeter-wave systems," *IEEE Communications Magazine*, vol. 52, no. 12, pp. 122–131, 2014.
- [18] P. Kumari, A. Mezghani, and R. W. Heath, "A low-resolution adc proof-of-concept development for a fully-digital millimeter-wave joint communication-radar," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 8619–8623.
- [19] A. Kaushik, E. Vlachos, C. Masouros, C. Tsinos, and J. Thompson, "Green Joint Radar-Communications: RF Selection with Low Resolution DACs and Hybrid Precoding," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 3160–3165.
- [20] S. Kang, M. Mezzavilla, S. Rangan, A. Madanayake, S. B. Venkatakrishnan, G. Hellboug, M. Ghosh, H. Rahmani, and A. Dhananjay, "Cellular wireless networks in the upper mid-band," 2023.
- [21] S. Dwivedi, R. Shreevastav, F. Munier, J. Nygren, I. Siomina, Y. Lyazidi, D. Shrestha, G. Lindmark, P. Ernström, E. Stare *et al.*, "Positioning in 5G networks," *IEEE Communications Magazine*, vol. 59, no. 11, pp. 38–44, 2021.
- [22] 3GPP, "Feasibility study on integrated sensing and communication (release 19)," 3GPP, Tech. Rep. TR 22.837, 2023.
- [23] —, "Service requirements for integrated sensing and communication; stage 1 (release 19)," 3GPP, Tech. Rep. TR 22.837, 2023.
- [24] B. Xiong, Z. Zhang, Y. Ge, H. Wang, H. Jiang, L. Wu, and Z. Zhang, "Channel Modeling for Heterogeneous Vehicular ISAC System with Shared Clusters," in *2023 IEEE 98th Vehicular Technology Conference (VTC2023-Fall)*, 2023, pp. 1–6.
- [25] Z. Zhang, R. He, B. Ai, M. Yang, C. Li, H. Mi, and Z. Zhang, "A general channel model for integrated sensing and communication scenarios," *IEEE Communications Magazine*, vol. 61, no. 5, pp. 68–74, 2023.
- [26] N. González-Prelcic, A. Ali, V. Va, and R. W. Heath, "Millimeter-wave communication with out-of-band information," *IEEE Communications Magazine*, vol. 55, no. 12, pp. 140–146, 2017.
- [27] V. Va, T. Shimizu, G. Bansal, and R. W. Heath, "Online learning for position-aided millimeter wave beam training," *IEEE Access*, vol. 7, pp. 30 507–30 526, 2019.
- [28] V. Va, J. Choi, T. Shimizu, G. Bansal, and R. W. Heath, "Inverse multipath fingerprinting for millimeter wave V2I beam alignment," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4042–4058, 2018.
- [29] A. Klautau, N. González-Prelcic, and R. W. Heath, "LIDAR data for deep learning-based mmWave beam-selection," *IEEE Wireless Commun. Lett.*, vol. 8, pp. 909–912, 2019.
- [30] A. Graff, Y. Chen, N. González-Prelcic, and T. Shimizu, "Deep learning-based link configuration for radar-aided multiuser mmWave vehicle-to-infrastructure communication," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 6, pp. 7454–7468, 2023.
- [31] S. Bi, J. Lyu, Z. Ding, and R. Zhang, "Engineering radio maps for wireless resource management," *IEEE Wireless Communications*, vol. 26, no. 2, pp. 133–141, 2019.
- [32] N. González-Prelcic, R. Mendez-Rial, and R. W. Heath, "Radar aided beam alignment in MmWave V2I communications supporting antenna diversity," in *Proc. Inf. Theory Appl. Workshop (ITA)*, 2016, pp. 1–7.
- [33] M. Mühlhäuser, C. Meurisch, M. Stein, J. Daubert, J. Von Willich, J. Riemann, and L. Wang, "Street lamps as a platform," *Commun. ACM*, vol. 63, no. 6, p. 75–83, may 2020. [Online]. Available: <https://doi.org/10.1145/3376900>
- [34] H. Holma, A. Toskala, and T. Nakamura, Eds., *5G Technology: 3GPP New Radio*, 1st ed. Wiley, 2019.
- [35] M. Renfors, E. Kofidis, and F. Bader, Eds., *Orthogonal Waveforms and Filter Banks for Future Communication Systems*. United States: Elsevier Inc., Jan. 2017.
- [36] G. Fettweis, M. Krondorf, and S. Bittner, "GFDM - generalized frequency division multiplexing," in *VTC Spring 2009 - IEEE 69th Vehicular Technology Conference*, 2009, pp. 1–4.
- [37] J. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Communications Magazine*, vol. 28, no. 5, pp. 5–14, 1990.
- [38] J. Yli-Kaakinen, T. Levanen, S. Valkonen, K. Pajukoski, J. Pirskanen, M. Renfors, and M. Valkama, "Efficient fast-convolution-based waveform processing for 5G physical layer," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1309–1326, 2017.
- [39] J. Yli-Kaakinen, T. Levanen, A. Palin, M. Renfors, and M. Valkama, "Generalized fast-convolution-based filtered-OFDM: Techniques and application to 5G new radio," *IEEE Transactions on Signal Processing*, vol. 68, pp. 1213–1228, 2020.
- [40] T. Levanen, J. Pirskanen, K. Pajukoski, M. Renfors, and M. Valkama, "Transparent Tx and Rx waveform processing for 5G new radio mobile communications," *IEEE Wireless Communications*, vol. 26, no. 1, pp. 128–136, 2019.
- [41] L. Gaudio, M. Kobayashi, G. Caire, and G. Colavolpe, "On the effectiveness of OTFS for joint radar parameter estimation and communication," *IEEE Transactions on Wireless Communications*, vol. 19, no. 9, pp. 5951–5965, 2020.
- [42] S. D. Liyanarachchi, T. Riihonen, C. B. Barneto, and M. Valkama, "Optimized waveforms for 5G–6G communication with sensing: Theory, simulations and experiments," *IEEE Transactions on Wireless Communications*, vol. 20, no. 12, pp. 8301–8315, 2021.
- [43] C. Sturm and W. Wiesbeck, "Waveform design and signal processing aspects for fusion of wireless communications and radar sensing," *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1236–1259, July 2011.
- [44] M. F. Keskin, V. Koivunen, and H. Wymeersch, "Limited feedforward waveform design for OFDM dual-functional radar-communications," *IEEE Transactions on Signal Processing*, vol. 69, pp. 2955–2970, 2021.

- [45] M. Bică and V. Koivunen, "Multicarrier radar-communications waveform design for RF convergence and coexistence," in *Proc. IEEE ICASSP*, 2019, pp. 7780–7784.
- [46] K. Venugopal, A. Alkhateeb, N. González-Prelcic, and R. W. Heath, "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1996–2009, 2017.
- [47] P. Kumari, J. Choi, N. González-Prelcic, and R. W. Heath, "IEEE 802.11ad-based radar: An approach to joint vehicular communication-radar system," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 4, pp. 3012–3027, April 2018.
- [48] 3GPP, "NR physical channels and modulation," 3GPP TS38.211 V17.1.0, Sophia Antipolis, France, Tech. Rep., April 2022.
- [49] M. F. Keskin, H. Wymeersch, and V. Koivunen, "MIMO-OFDM joint radar-communications: Is ICI friend or foe?" *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 6, pp. 1393–1408, 2021.
- [50] J. Palacios, N. González-Prelcic, and C. Rusu, "Low complexity joint position and channel estimation at millimeter wave based on multidimensional orthogonal matching pursuit," in *Proc. 30th Eur. Signal Process. Conf. (EUSIPCO)*. IEEE, 2022, pp. 1002–1006.
- [51] P. Chen and H. Kobayashi, "Maximum likelihood channel estimation and signal detection for OFDM systems," in *2002 IEEE International Conference on Communications. Conference Proceedings. ICC 2002 (Cat. No.02CH37333)*, vol. 3, 2002, pp. 1640–1645 vol.3.
- [52] R. Carvajal, J. C. Aguero, B. I. Godoy, and G. C. Goodwin, "Em-based maximum-likelihood channel estimation in multicarrier systems with phase distortion," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 1, pp. 152–160, 2013.
- [53] B. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. Ingeman Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 3, pp. 434–450, 1999.
- [54] B. Fleury, P. Jourdan, and A. Stucki, "High-resolution channel parameter estimation for MIMO applications using the SAGE algorithm," in *2002 International Zurich Seminar on Broadband Communications Access - Transmission - Networking (Cat. No.02TH8599)*, 2002, pp. 30–30.
- [55] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, "Position and orientation estimation through millimeter-wave MIMO in 5G systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1822–1835, 2018.
- [56] J. Lee, G.-T. Gil, and Y. H. Lee, "Exploiting spatial sparsity for estimating channels of hybrid MIMO systems in millimeter wave communications," in *2014 IEEE Global Communications Conference*, 2014, pp. 3326–3331.
- [57] J. Rodríguez-Fernández, N. González-Prelcic, K. Venugopal, and R. W. Heath, "Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 5, pp. 2946–2960, 2018.
- [58] X. Wu, G. Yang, F. Hou, and S. Ma, "Low-complexity downlink channel estimation for millimeter-wave FDD massive MIMO systems," *IEEE Wireless Communications Letters*, 2019.
- [59] J. Palacios and N. González-Prelcic, "Separable multidimensional orthogonal matching pursuit and its application to joint localization and communication at mmWave," in *2023 IEEE Globecom Workshops (GC Wkshps)*, 2023, pp. 1–6.
- [60] F. Jiang, Y. Ge, M. Zhu, and H. Wymeersch, "High-dimensional channel estimation for simultaneous localization and communications," in *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, 2021, pp. 1–6.
- [61] F. Wen, H. C. So, and H. Wymeersch, "Tensor decomposition-based beamspace ESPRIT algorithm for multidimensional harmonic retrieval," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 4572–4576.
- [62] J. Zhang, D. Rakhimov, and M. Haardt, "Gridless channel estimation for hybrid mmWave MIMO systems via tensor-ESPRIT algorithms in DFT beamspace," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 3, pp. 816–831, 2021.
- [63] Y. Chen, N. González-Prelcic, T. Shimizu, H. Lu, and C. Mahabal, "Beamspace ESPRIT-D for Joint 3D Angle and Delay Estimation for Joint Localization and Communication at MmWave," *arXiv preprint*, 2024.
- [64] M. Bayraktar, C. Rusu, N. González-Prelcic, and C. J. Zhang, "Beam codebook design for joint initial access and localization in mmwave networks," in *2022 56th Asilomar Conference on Signals, Systems, and Computers*, 2022, pp. 919–924.
- [65] X. Ge, W. Shen, C. Xing, L. Zhao, and J. An, "Training beam design for channel estimation in hybrid mmwave mimo systems," *IEEE Transactions on Wireless Communications*, vol. 21, no. 9, pp. 7121–7134, 2022.
- [66] M. F. Keskin, F. Jiang, F. Munier, G. Seco-Granados, and H. Wymeersch, "Optimal spatial signal design for mmWave positioning under imperfect synchronization," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 5, pp. 5558–5563, 2022.
- [67] A. Fascista, M. F. Keskin, A. Coluccia, H. Wymeersch, and G. Seco-Granados, "RIS-aided joint localization and synchronization with a single-antenna receiver: Beamforming design and low-complexity estimation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 5, pp. 1141–1156, 2022.
- [68] S. Rivetti, J. Miguel Mateos-Ramos, Y. Wu, J. Song, M. F. Keskin, V. Yajnanarayana, C. Häger, and H. Wymeersch, "Spatial signal design for positioning via end-to-end learning," *IEEE Wireless Communications Letters*, vol. 12, no. 3, pp. 525–529, 2023.
- [69] S. H. Talisa, K. W. O'Haver, T. M. Comberiate, M. D. Sharp, and O. F. Somerlock, "Benefits of digital phased array radars," *Proceedings of the IEEE*, vol. 104, no. 3, pp. 530–543, 2016.
- [70] I. Aykin and M. Krunz, "Efficient beam sweeping algorithms and initial access protocols for millimeter-wave networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2504–2514, 2020.
- [71] U. Nickel, "Overview of generalized monopulse estimation," *IEEE Aerospace and Electronic Systems Magazine*, vol. 21, no. 6, pp. 27–56, 2006.
- [72] M. Z. Win, W. Dai, Y. Shen, G. Chrisikos, and H. Vincent Poor, "Network operation strategies for efficient localization and navigation," *Proceedings of the IEEE*, vol. 106, no. 7, pp. 1224–1254, 2018.
- [73] Y. Shen and M. Z. Win, "Fundamental limits of wideband localization—part I: A general framework," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4956–4980, 2010.
- [74] R. Miller and C. Chang, "A modified Cramér-Rao bound and its applications (corresp.)," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 398–400, 1978.
- [75] F. Liu, W. Yuan, C. Masouros, and J. Yuan, "Radar-assisted predictive beamforming for vehicular links: Communication served by sensing," *IEEE Transactions on Wireless Communications*, vol. 19, no. 11, pp. 7704–7719, 2020.
- [76] F. Liu, C. Masouros, A. Li, H. Sun, and L. Hanzo, "MU-MIMO communications with MIMO radar: From co-existence to joint transmission," *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, pp. 2755–2770, 2018.
- [77] H. Zhang, Y. Zhang, X. Liu, C. Ren, H. Li, and C. Sun, "Time allocation approaches for a perceptive mobile network using integration of sensing and communication," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2023.
- [78] W. Dai, Y. Shen, and M. Z. Win, "A computational geometry framework for efficient network localization," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 1317–1339, 2018.
- [79] W. W.-L. Li, Y. Shen, Y. J. Zhang, and M. Z. Win, "Robust power allocation for energy-efficient location-aware networks," *IEEE/ACM Transactions on Networking*, vol. 21, no. 6, pp. 1918–1930, 2013.
- [80] H. Zhao, N. Zhang, and Y. Shen, "Beamspace direct localization for large-scale antenna array systems," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3529–3544, 2020.
- [81] C. Ding, J.-B. Wang, H. Zhang, M. Lin, and G. Y. Li, "Joint MIMO precoding and computation resource allocation for dual-function radar and communication systems with mobile edge computing," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 7, pp. 2085–2102, 2022.
- [82] D. Xu, X. Yu, D. W. K. Ng, A. Schmeink, and R. Schober, "Robust and secure resource allocation for ISAC systems: A novel optimization framework for variable-length snapshots," *IEEE Transactions on Communications*, vol. 70, no. 12, pp. 8196–8214, 2022.
- [83] F. Wang, H. Li, and M. A. Govoni, "Power allocation and co-design of multicarrier communication and radar systems for spectral coexistence," *IEEE Transactions on Signal Processing*, vol. 67, no. 14, pp. 3818–3831, 2019.
- [84] J. Mu, W. Ouyang, Z. Jing, B. Li, and F. Zhang, "Energy-efficient interference cancellation in integrated sensing and communication scenarios," *IEEE Transactions on Green Communications and Networking*, vol. 7, no. 1, pp. 370–378, 2023.
- [85] Y. Xiong, F. Liu, Y. Cui, W. Yuan, T. X. Han, and G. Caire, "On the fundamental tradeoff of integrated sensing and communications under gaussian channels," *IEEE Transactions on Information Theory*, vol. 69, no. 9, pp. 5723–5751, 2023.

- [86] A. R. Chiriyath, B. Paul, G. M. Jacyna, and D. W. Bliss, "Inner bounds on performance of radar and communications co-existence," *IEEE Transactions on Signal Processing*, vol. 64, no. 2, pp. 464–474, 2016.
- [87] B. Li, X. Wang, Y. Xin, and E. Au, "Value of service maximization in integrated localization and communication system through joint resource allocation," *IEEE Transactions on Communications*, vol. 71, no. 8, pp. 4957–4971, 2023.
- [88] H. Yang, Z. Wei, Z. Feng, C. Qiu, Z. Fang, X. Chen, and P. Zhang, "Queue-aware dynamic resource allocation for the joint communication-radar system," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 1, pp. 754–767, 2021.
- [89] L. Pucci, E. Paolini, and A. Giorgetti, "System-level analysis of joint sensing and communication based on 5G new radio," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 7, pp. 2043–2055, 2022.
- [90] A. Fascista, B. J. Deutschmann, M. F. Keskin, T. Wilding, A. Coluccia, K. Witrals, E. Leitinger, G. Seco-Granados, and H. Wymeersch, "Uplink joint positioning and synchronization in cell-free deployments with radio stripes," in *Proc. IEEE Int. Conf. Commun. (ICC) Workshops*, May, 2023, pp. 1330–1336.
- [91] N. Vukmirović, M. Erić, M. Janjić, and P. M. Djurić, "Direct wideband coherent localization by distributed antenna arrays," *Sensors*, vol. 19, no. 20, p. 4582, 2019.
- [92] A. M. Haimovich, R. S. Blum, and L. J. Cimini, "MIMO radar with widely separated antennas," *IEEE signal processing magazine*, vol. 25, no. 1, pp. 116–129, 2007.
- [93] H. Wymeersch and G. Seco-Granados, "Radio localization and sensing—part I: Fundamentals," *IEEE Communications Letters*, vol. 26, no. 12, pp. 2816–2820, 2022.
- [94] S. E. Trelvakis, A.-A. A. Boulogeorgos, D. Pliatsios, J. Querol, K. Ntontin, P. Sarigiannidis, S. Chatzinotas, and M. Di Renzo, "Localization as a key enabler of 6G wireless systems: A comprehensive survey and an outlook," *IEEE Open Journal of the Communications Society*, 2023.
- [95] İ. Güvenç, C.-C. Chong, F. Watanabe, and H. Inamura, "NLOS identification and weighted least-squares localization for UWB systems using multipath channel statistics," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, pp. 1–14, 2007.
- [96] H. Wymeersch, N. Garcia, H. Kim, G. Seco-Granados, S. Kim, F. Wen, and M. Fröhle, "5G mmWave downlink vehicular positioning," in *2018 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2018, pp. 206–212.
- [97] B. Camajori Tedeschini, M. Brambilla, L. Italiano, S. Reggiani, D. Vaccaroni, M. Alghisi, L. Benvenuto, A. Goia, E. Realini, F. Grec *et al.*, "A feasibility study of 5G positioning with current cellular network deployment," *Scientific reports*, vol. 13, no. 1, p. 15281, 2023.
- [98] M. Koivisto, M. Costa, J. Werner, K. Heiska, J. Talvitie, K. Leppänen, V. Koivunen, and M. Valkama, "Joint device positioning and clock synchronization in 5G ultra-dense networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 2866–2881, 2017.
- [99] M. A. Nazari, G. Seco-Granados, P. Johansson, and H. Wymeersch, "Mmwave 6D radio localization with a snapshot observation from a single BS," *IEEE Transactions on Vehicular Technology*, 2023.
- [100] Y. Chen, J. Palacios, N. González-Prelcic, T. Shimizu, and H. Lu, "Joint initial access and localization in millimeter wave vehicular networks: a hybrid model/data driven approach," in *2022 IEEE 12th Sensor Array and Multichannel Signal Processing Workshop (SAM)*. IEEE, 2022, pp. 355–359.
- [101] F. Gustafsson and F. Gunnarsson, "Mobile positioning using wireless networks: possibilities and fundamental limitations based on available wireless network measurements," *IEEE Signal processing magazine*, vol. 22, no. 4, pp. 41–53, 2005.
- [102] G. Wang, A. M.-C. So, and Y. Li, "Robust convex approximation methods for TDOA-based localization under NLOS conditions," *IEEE Transactions on Signal processing*, vol. 64, no. 13, pp. 3281–3296, 2016.
- [103] Y. Lu, O. Kaltiokallio, M. Koivisto, J. Talvitie, E. S. Lohan, H. Wymeersch, and M. Valkama, "Bayesian filtering for joint multi-user positioning, synchronization and anchor state calibration," *IEEE Transactions on Vehicular Technology*, 2023.
- [104] S. Burgess, Y. Kuang, and K. Åström, "TOA sensor network self-calibration for receiver and transmitter spaces with difference in dimension," *Signal Processing*, vol. 107, pp. 33–42, 2015.
- [105] X. Li, E. Leitinger, M. Oskarsson, K. Åström, and F. Tufvesson, "Massive MIMO-based localization and mapping exploiting phase information of multipath components," *IEEE transactions on wireless communications*, vol. 18, no. 9, pp. 4254–4267, 2019.
- [106] Y. T. Chan and K. C. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE transactions on signal processing*, vol. 42, no. 8, pp. 1905–1915, 1994.
- [107] H. Dun, C. C. Tiberius, and G. J. Janssen, "Positioning in a multipath channel using OFDM signals with carrier phase tracking," *IEEE Access*, vol. 8, pp. 13 011–13 028, 2020.
- [108] J. Talvitie, M. Säily, and M. Valkama, "Orientation and Location Tracking of XR Devices: 5G Carrier Phase-Based Methods," *IEEE Journal of Selected Topics in Signal Processing*, pp. 1–16, 2023.
- [109] E. Leitinger, F. Meyer, F. Hlawatsch, K. Witrals, 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.
- [110] C. Gentner, T. Jost, W. Wang, S. Zhang, A. Dammann, and U.-C. Fiebig, "Multipath assisted positioning with simultaneous localization and mapping," *IEEE Transactions on Wireless Communications*, vol. 15, no. 9, pp. 6104–6117, 2016.
- [111] H. Chen, H. Kim, M. Ammous, G. Seco-Granados, G. C. Alexandropoulos, S. Valaee, and H. Wymeersch, "RISs and sidelink communications in smart cities: The key to seamless localization and sensing," *IEEE Communications Magazine*, vol. 61, no. 8, pp. 140–146, 2023.
- [112] A. Elzanaty, A. Guerra, F. Guidi, and M.-S. Alouini, "Reconfigurable intelligent surfaces for localization: Position and orientation error bounds," *IEEE Transactions on Signal Processing*, vol. 69, pp. 5386–5402, 2021.
- [113] Z. Wang, Z. Liu, Y. Shen, A. Conti, and M. Z. Win, "Location awareness in beyond 5G networks via reconfigurable intelligent surfaces," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 7, pp. 2011–2025, 2022.
- [114] Q. Liu, P. Liang, J. Xia, T. Wang, M. Song, X. Xu, J. Zhang, Y. Fan, and L. Liu, "A highly accurate positioning solution for C-V2X systems," *Sensors*, vol. 21, no. 4, p. 1175, 2021.
- [115] Z. Li, K. Xu, H. Wang, Y. Zhao, X. Wang, and M. Shen, "Machine-learning-based positioning: A survey and future directions," *IEEE Network*, vol. 33, no. 3, pp. 96–101, 2019.
- [116] H. Kim, K. Granström, L. Svensson, S. Kim, and H. Wymeersch, "PMBM-based SLAM filters in 5G mmWave vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 8, pp. 8646–8661, 2022.
- [117] R. Keating, M. Säily, J. Hultkonen, and J. Karjalainen, "Overview of positioning in 5G new radio," in *2019 16th International Symposium on Wireless Communication Systems (ISWCS)*. IEEE, 2019, pp. 320–324.
- [118] N. Garcia, H. Wymeersch, E. G. Larsson, A. M. Haimovich, and M. Coulon, "Direct localization for massive MIMO," *IEEE Transactions on Signal Processing*, vol. 65, no. 10, pp. 2475–2487, 2017.
- [119] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, "Massive MIMO in real propagation environments: Do all antennas contribute equally?" *IEEE Transactions on Communications*, vol. 63, no. 11, pp. 3917–3928, 2015.
- [120] TR 38.901, "Study on channel model for frequencies from 0.5 to 100 GHz, ver17.0.0," *3GPP*, 2023.
- [121] M. Pesavento, M. Trinh-Hoang, and M. Viberg, "Three more decades in array signal processing research: An optimization and structure exploitation perspective," *IEEE Signal Processing Magazine*, vol. 40, no. 4, pp. 92–106, 2023.
- [122] A. Shastri, N. Valecha, E. Bashirov, H. Tataria, M. Lentmaier, F. Tufvesson, M. Rossi, and P. Casari, "A review of millimeter wave device-based localization and device-free sensing technologies and applications," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 3, pp. 1708–1749, 2022.
- [123] H. Chen, H. Sarieddeen, T. Ballal, H. Wymeersch, M.-S. Alouini, and T. Y. Al-Naffouri, "A tutorial on terahertz-band localization for 6G communication systems," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 3, pp. 1780–1815, 2022.
- [124] J. Zhang, X. Ge, Q. Li, M. Guizani, and Y. Zhang, "5G millimeter-wave antenna array: Design and challenges," *IEEE Wireless communications*, vol. 24, no. 2, pp. 106–112, 2016.
- [125] H. Chen, M. F. Keskin, S. R. Aghdam, H. Kim, S. Lindberg, A. Wolfgang, T. E. Abrudan, T. Eriksson, and H. Wymeersch, "Modeling and analysis of 6G joint localization and communication under hardware impairments," *arXiv preprint arXiv:2301.01042*, 2023.
- [126] D. Tubail, B. Cenikloglu, A. E. Canbilin, I. Develi, and S. S. Ikki, "The effect of hardware impairments on the error bounds of localization and maximum likelihood estimation of mm-Wave MISO-OFDM systems," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 3, pp. 4063–4067, 2022.

- [127] M. Shafi, J. Zhang, H. Tataria, A. F. Molisch, S. Sun, T. S. Rappaport, F. Tufvesson, S. Wu, and K. Kitao, "Microwave vs. millimeter-wave propagation channels: Key differences and impact on 5G cellular systems," *IEEE Communications Magazine*, vol. 56, no. 12, pp. 14–20, 2018.
- [128] T. S. Rappaport, Y. Xing, O. Kanhere, S. Ju, A. Madanayake, S. Mandal, A. Alkhateeb, and G. C. Trichopoulos, "Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond," *IEEE access*, vol. 7, pp. 78 729–78 757, 2019.
- [129] Remcom. Wireless InSite - 3D Wireless Prediction Software. Accessed: Oct 9, 2022). [Online]. Available: <https://www.remcom.com/wireless-insite-em-propagation-software>
- [130] 3GPP, "Technical specification group radio access network; study on evaluation methodology of new vehicle-to-everything (V2X) use cases for LTE and NR," TR 37.885, Technical Report 15.3.0, 2019.
- [131] Y. Ge, M. Stark, M. F. Keskin, F. Hofmann, T. Hansen, and H. Wymeersch, "V2X sidelink positioning in FR1: Scenarios, algorithms, and performance evaluation," *arXiv preprint arXiv:2310.13753*, 2023.
- [132] S. He and S.-H. G. Chan, "Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 466–490, 2016.
- [133] R. Klus, J. Talvitie, J. Vinogradova, J. Torsner, and M. Valkama, "Machine learning based NLOS radio positioning in beamforming networks," in *2022 IEEE 23rd International Workshop on Signal Processing Advances in Wireless Communication (SPAWC)*, 2022, pp. 1–5.
- [134] P. Wang, T. Koike-Akino, and P. V. Orlik, "Fingerprinting-based indoor localization with commercial MMWave WiFi: NLOS propagation," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, 2020, pp. 1–6.
- [135] P. Bahl and V. Padmanabhan, "RADAR: an in-building RF-based user location and tracking system," in *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)*, vol. 2, 2000, pp. 775–784 vol.2.
- [136] Q. D. Vo and P. De, "A survey of fingerprint-based outdoor localization," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 491–506, 2016.
- [137] J. Hu, D. Liu, Z. Yan, and H. Liu, "Experimental analysis on weight K -nearest neighbor indoor fingerprint positioning," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 891–897, 2019.
- [138] M. Youssef and A. Agrawala, "The Horus WLAN location determination system," in *Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '05. New York, NY, USA: Association for Computing Machinery, 2005, p. 205–218. [Online]. Available: <https://doi.org/10.1145/1067170.1067193>
- [139] B. Mager, P. Lundrigan, and N. Patwari, "Fingerprint-based device-free localization performance in changing environments," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 11, pp. 2429–2438, 2015.
- [140] X. Sun, X. Gao, G. Y. Li, and W. Han, "Single-site localization based on a new type of fingerprint for massive MIMO-OFDM systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 6134–6145, 2018.
- [141] C. Wu, X. Yi, W. Wang, L. You, Q. Huang, X. Gao, and Q. Liu, "Learning to localize: A 3D CNN approach to user positioning in massive MIMO-OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 7, pp. 4556–4570, 2021.
- [142] J. Xiao, K. Wu, Y. Yi, and L. M. Ni, "FIFS: Fine-grained indoor fingerprinting system," in *2012 21st International Conference on Computer Communications and Networks (ICCCN)*, 2012, pp. 1–7.
- [143] X. Wang, L. Gao, S. Mao, and S. Pandey, "CSI-based fingerprinting for indoor localization: A deep learning approach," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 1, pp. 763–776, 2017.
- [144] B. Zhang, H. Sifaou, and G. Y. Li, "CSI-fingerprinting indoor localization via attention-augmented residual convolutional neural network," *IEEE Transactions on Wireless Communications*, vol. 22, no. 8, pp. 5583–5597, 2023.
- [145] P. Ferrand, A. Decurninge, and M. Guillaud, "DNN-based localization from channel estimates: Feature design and experimental results," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, 2020, pp. 1–6.
- [146] Y. Chen, N. González-Prelcic, T. Shimizu, and H. Lu, "Learning to localize with attention: from sparse mmWave channel estimates from a single BS to high accuracy 3D location," *arXiv preprint arXiv:2307.00167*, 2023.
- [147] Y. Chen, N. González-Prelcic, T. Shimizu, H. Lu, and C. Mahabal, "Sparse recovery with attention: A hybrid data/model driven solution for high accuracy position and channel tracking at mmWave," in *IEEE 24th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2023, pp. 491–495.
- [148] A. Salihu, S. Schwarz, and M. Rupp, "Attention aided CSI wireless localization," in *2022 IEEE 23rd International Workshop on Signal Processing Advances in Wireless Communication (SPAWC)*, 2022, pp. 1–5.
- [149] H. Chen, Y. Zhang, W. Li, X. Tao, and P. Zhang, "ConFi: Convolutional neural networks based indoor Wi-Fi localization using channel state information," *IEEE Access*, vol. 5, pp. 18 066–18 074, 2017.
- [150] M. M. Butt, A. Pantelidou, and I. Z. Kovács, "ML-assisted UE positioning: Performance analysis and 5G architecture enhancements," *IEEE Open Journal of Vehicular Technology*, vol. 2, pp. 377–388, 2021.
- [151] J. Gante, G. Falcao, and L. Sousa, "Deep learning architectures for accurate millimeter wave positioning in 5G," *Neural Processing Letters*, vol. 51, no. 1, pp. 487–514, 2020.
- [152] G. Kia, L. Ruotsalainen, and J. Talvitie, "A CNN approach for 5G mmWave positioning using beamformed CSI measurements," in *2022 International Conference on Localization and GNSS (ICL-GNSS)*. IEEE, 2022, pp. 01–07.
- [153] X. Wang, M. Patil, C. Yang, S. Mao, and P. A. Patel, "Deep convolutional Gaussian processes for mmwave outdoor localization," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 8323–8327.
- [154] M. A. M. Sadr, J. Gante, B. Champagne, G. Falcao, and L. Sousa, "Uncertainty estimation via Monte Carlo dropout in CNN-based mmWave MIMO localization," *IEEE Signal Processing Letters*, vol. 29, pp. 269–273, 2021.
- [155] Z. Chen, Z. Zhang, Z. Xiao, C. Zhang, and Z. Yang, "CSI of each subcarrier is a fingerprint: Multi-carrier cumulative learning based positioning in massive MIMO systems," in *2023 IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2023, pp. 1–7.
- [156] Z. Gao, Y. Gao, S. Wang, D. Li, and Y. Xu, "CRISLoc: Reconstructable CSI fingerprinting for indoor smartphone localization," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3422–3437, 2021.
- [157] S. Shi, S. Sigg, L. Chen, and Y. Ji, "Accurate location tracking from csi-based passive device-free probabilistic fingerprinting," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 6, pp. 5217–5230, 2018.
- [158] Z. Zhang, M. Lee, and S. Choi, "Deep-learning-based wi-fi indoor positioning system using continuous csi of trajectories," *Sensors*, vol. 21, no. 17, 2021. [Online]. Available: <https://www.mdpi.com/1424-8220/21/17/5776>
- [159] T. Linget, "System architecture and solution development; high-accuracy positioning for C-V2X," *5G Automotive Association*, 2021.
- [160] Y. Ge, M. Stark, M. F. Keskin, F. Hofmann, T. Hansen, and H. Wymeersch, "Analysis of V2X sidelink positioning in sub-6 GHz," in *2023 IEEE 3rd International Symposium on Joint Communications & Sensing (JC&S)*. IEEE, 2023, pp. 1–6.
- [161] N. Decarli, A. Guerra, C. Giovannetti, F. Guidi, and B. M. Masini, "V2X sidelink localization of connected automated vehicles," *IEEE Journal on Selected Areas in Communications*, pp. 1–1, 2023.
- [162] Y. Shen, H. Wymeersch, and M. Z. Win, "Fundamental limits of wide-band localization—part ii: Cooperative networks," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4981–5000, 2010.
- [163] M. Z. Win, Y. Shen, and W. Dai, "A theoretical foundation of network localization and navigation," *Proceedings of the IEEE*, vol. 106, no. 7, pp. 1136–1165, 2018.
- [164] S. Mazuelas, Y. Shen, and M. Z. Win, "Spatiotemporal information coupling in network navigation," *IEEE Transactions on Information Theory*, vol. 64, no. 12, pp. 7759–7779, 2018.
- [165] X. Shen, Y. Liu, and Y. Shen, "On the spatial information coupling in relative localization networks," in *ICC 2021 - IEEE International Conference on Communications*, 2021, pp. 1–6.
- [166] Y. Xiong, N. Wu, Y. Shen, and M. Z. Win, "Cooperative localization in massive networks," *IEEE Transactions on Information Theory*, vol. 68, no. 2, pp. 1237–1258, 2022.
- [167] W. Dai, Y. Shen, and M. Z. Win, "Distributed power allocation for cooperative wireless network localization," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 1, pp. 28–40, 2015.
- [168] J. Chen, W. Dai, Y. Shen, V. K. N. Lau, and M. Z. Win, "Power management for cooperative localization: A game theoretical approach," *IEEE Transactions on Signal Processing*, vol. 64, no. 24, pp. 6517–6532, 2016.

- [169] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE access*, vol. 7, pp. 116 753–116 773, 2019.
- [170] K. Keykhosravi, M. F. Keskin, G. Seco-Granados, P. Popovski, and H. Wymeersch, "RIS-enabled SISO localization under user mobility and spatial-wideband effects," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 5, pp. 1125–1140, 2022.
- [171] J. He, A. Fakhreddine, and G. C. Alexandropoulos, "STAR-RIS-enabled simultaneous indoor and outdoor 3D localisation: Theoretical analysis and algorithmic design," *IET Signal Processing*, vol. 17, no. 4, p. e12209, 2023.
- [172] H. Li, S. Shen, and B. Clerckx, "Beyond diagonal reconfigurable intelligent surfaces: A multi-sector mode enabling highly directional full-space wireless coverage," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 8, pp. 2446–2460, 2023.
- [173] E. Björnson, H. Wymeersch, B. Matthiesen, P. Popovski, L. Sanguinetti, and E. de Carvalho, "Reconfigurable intelligent surfaces: A signal processing perspective with wireless applications," *IEEE Signal Processing Magazine*, vol. 39, no. 2, pp. 135–158, 2022.
- [174] M. Mizmizi, D. Tagliaferri, D. Badini, and U. Spagnolini, "Target-to-user association in ISAC systems with vehicle-lodged RIS," *IEEE Wireless Communications Letters*, 2023.
- [175] M. Bayraktar, J. Palacios, N. González-Prelcic, and C. J. Zhang, "Multidimensional orthogonal matching pursuit-based RIS-aided joint localization and channel estimation at mmWave," in *Proc. IEEE 23rd Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*. IEEE, 2022, pp. 1–5.
- [176] 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, 2021.
- [177] 3GPP, TR 38.859 V18.0.0, "Study on expanded and improved NR positioning (Release 18)," Dec. 2022. [Online]. Available: www.3gpp.org
- [178] 3GPP, RP-223549, "Expanded and Improved NR Positioning Work Item Description (Release 18)," Dec. 2022. [Online]. Available: www.3gpp.org
- [179] P. Teunissen, "The least-squares ambiguity decorrelation adjustment: A method for fast GPS integer ambiguity estimation," *Journal of Geodesy*, vol. 70, pp. 65–82, Nov. 1995.
- [180] D. B. Cox Jr. and J. D. W. Brading, "Integration of LAMBDA ambiguity resolution with Kalman filter for relative navigation of spacecraft," *NAVIGATION*, vol. 47, no. 3, pp. 205–210, 2000.
- [181] L. Chen *et al.*, "Carrier Phase Ranging for Indoor Positioning With 5G NR Signals," *IEEE Internet of Things Journal*, vol. 9, pp. 10 908–10 919, 2022.
- [182] S. Fan *et al.*, "Carrier Phase-Based Synchronization and High-Accuracy Positioning in 5G New Radio Cellular Networks," *IEEE Transactions on Communications*, vol. 70, no. 1, pp. 564–577, 2022.
- [183] B.-n. Vo, M. Mallick, Y. Bar-Shalom, S. Coraluppi, R. Osborne, R. Mahler, and B.-t. Vo, "Multitarget tracking," *Wiley encyclopedia of electrical and electronics engineering*, no. 2015, 2015.
- [184] K. Granström, M. Baum, and S. Reuter, "Extended object tracking: Introduction, overview and applications," *arXiv preprint arXiv:1604.00970*, 2016.
- [185] J. L. Williams and R. A. Lau, "Data association by loopy belief propagation," in *2010 13th International Conference on Information Fusion*. IEEE, 2010, pp. 1–8.
- [186] W. Yi, G. Li, and G. Battistelli, "Distributed multi-sensor fusion of PHD filters with different sensor fields of view," *IEEE Transactions on Signal Processing*, vol. 68, pp. 5204–5218, 2020.
- [187] H. Kim, K. Granström, L. Gao, G. Battistelli, S. Kim, and H. Wymeersch, "5G mmWave cooperative positioning and mapping using multi-model PHD filter and map fusion," *IEEE Transactions on Wireless Communications*, vol. 19, no. 6, pp. 3782–3795, 2020.
- [188] O. Kaltiokallio *et al.*, "Towards real-time Radio-SLAM via optimal importance sampling," in *Proc. IEEE SPAWC*, 2022, pp. 1–5.
- [189] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press, 2005.
- [190] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit, "FastSLAM: A factored solution to the simultaneous localization and mapping problem," in *Eighteenth National Conference on Artificial Intelligence*. USA: American Association for Artificial Intelligence, 2002, p. 593–598.
- [191] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," *IEEE Trans. Robot.*, vol. 32, no. 6, pp. 1309–1332, 2016.
- [192] F. Wen and H. Wymeersch, "5G synchronization, positioning, and mapping from diffuse multipath," *IEEE Wireless Commun. Lett.*, vol. 10, no. 1, pp. 43–47, 2021.
- [193] E. Rastorgueva-Foi, O. Kaltiokallio, Y. Ge, M. Turunen, J. Talvitie, B. Tan, M. F. Keskin, H. Wymeersch, and M. Valkama, "Millimeter-wave radio SLAM: End-to-end processing methods and experimental validation," 2023.
- [194] M. A. Richards, *Fundamentals of Radar Signal Processing*. Tata McGraw-Hill Education, 2005.
- [195] C. B. Barneto, S. D. Liyanaarachchi, M. Heino, T. Riihonen, and M. Valkama, "Full duplex radio/radar technology: The enabler for advanced joint communication and sensing," *IEEE Wireless Commun.*, vol. 28, no. 1, pp. 82–88, 2021.
- [196] C. Baquero Barneto, T. Riihonen, M. Turunen, L. Anttila, M. Fleischer, K. Stadius, J. Ryyänen, and M. Valkama, "Full-duplex OFDM radar with LTE and 5G NR waveforms: Challenges, solutions, and measurements," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 10, pp. 4042–4054, 2019.
- [197] K. E. Kolodziej, B. T. Perry, and J. S. Herd, "In-band full-duplex technology: Techniques and systems survey," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 7, pp. 3025–3041, 2019.
- [198] S. Khaledian, F. Farzami, B. Smida, and D. Erricolo, "Inherent Self-Interference Cancellation for In-Band Full-Duplex Single-Antenna Systems," *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 6, pp. 2842–2850, 2018.
- [199] Y. Liu, P. Roblin, X. Quan, W. Pan, S. Shao, and Y. Tang, "A Full-Duplex Transceiver With Two-Stage Analog Cancellations for Multipath Self-Interference," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 12, pp. 5263–5273, 2017.
- [200] P. Pascual Campo, L. Anttila, D. Korpi, and M. Valkama, "Cascaded spline-based models for complex nonlinear systems: Methods and applications," *IEEE Transactions on Signal Processing*, vol. 69, pp. 370–384, 2021.
- [201] A. Kiayani, M. Z. Waheed, L. Anttila, M. Abdelaziz, D. Korpi, V. Syrjälä, M. Kosunen, K. Stadius, J. Ryyänen, and M. Valkama, "Adaptive Nonlinear RF Cancellation for Improved Isolation in Simultaneous Transmit–Receive Systems," *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 5, pp. 2299–2312, 2018.
- [202] C. B. Barneto, T. Riihonen, S. D. Liyanaarachchi, M. Heino, N. González-Prelcic, and M. Valkama, "Beamformer design and optimization for joint communication and full-duplex sensing at mm-waves," *IEEE Trans. Commun.*, vol. 70, no. 12, pp. 8298–8312, 2022.
- [203] M. A. Islam, G. C. Alexandropoulos, and B. Smida, "Integrated sensing and communication with millimeter wave full duplex hybrid beamforming," in *Proc. IEEE Int. Conf. Commun.*, 2022, pp. 4673–4678.
- [204] S. D. Liyanaarachchi, C. Baquero Barneto, T. Riihonen, M. Heino, and M. Valkama, "Joint multi-user communication and MIMO radar through full-duplex hybrid beamforming," in *2021 1st IEEE International Online Symposium on Joint Communications & Sensing (JC&S)*, 2021, pp. 1–5.
- [205] C. B. Barneto, S. D. Liyanaarachchi, T. Riihonen, L. Anttila, and M. Valkama, "Multibeam design for joint communication and sensing in 5g new radio networks," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.
- [206] P. Pascual Campo, D. Korpi, L. Anttila, and M. Valkama, "Nonlinear digital cancellation in full-duplex devices using spline-based Hammerstein model," in *2018 IEEE Globecom Workshops (GC Wkshps)*, 2018, pp. 1–7.
- [207] D. Korpi, S. Venkatasubramanian, T. Riihonen, L. Anttila, S. Otewa, C. Icheln, K. Haneda, S. Tretyakov, M. Valkama, and R. Wichman, "Advanced self-interference cancellation and multi-antenna techniques for full-duplex radios," in *2013 Asilomar Conference on Signals, Systems and Computers*, 2013, pp. 3–8.
- [208] M. Bayraktar, N. González-Prelcic, and H. Chen, "Hybrid precoding and combining for mmWave full-duplex joint radar and communication systems under self-interference," *arXiv preprint arXiv:2311.14942*, 2023.
- [209] M. Bayraktar, C. Rusu, N. González-Prelcic, and H. Chen, "Self-interference aware codebook design for full-duplex joint sensing and communication systems at mmWave," in *Proc. IEEE Comput. Advances Multi-Sensor Adaptive Process. (CAMSAP)*, 2023, pp. 1–5.
- [210] M. Braun, "OFDM radar algorithms in mobile communication networks," *Karlsruher Institutes für Technologie*, 2014.
- [211] S. Mercier, S. Bidon, D. Roque, and C. Enderli, "Comparison of correlation-based OFDM radar receivers," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 6, pp. 4796–4813, 2020.

- [212] 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. 11 068–11 092, 2023.
- [213] Y. Liu, G. Liao, Y. Chen, J. Xu, and Y. Yin, "Super-resolution range and velocity estimations with OFDM integrated radar and communications waveform," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11 659–11 672, 2020.
- [214] J. B. Sanson, P. M. Tomé, D. Castanheira, A. Gameiro, and P. P. Monteiro, "High-resolution delay-Doppler estimation using received communication signals for OFDM radar-communication system," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11, pp. 13 112–13 123, 2020.
- [215] G. Hakobyan and B. Yang, "A novel intercarrier-interference free signal processing scheme for OFDM radar," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 6, pp. 5158–5167, 2017.
- [216] M. F. Keskin, H. Wymeersch, and V. Koivunen, "Monostatic sensing with OFDM under phase noise: From mitigation to exploitation," *IEEE Trans. Signal Process.*, vol. 71, pp. 1363–1378, 2023.
- [217] A. Demir, "Computing timing jitter from phase noise spectra for oscillators and phase-locked loops with white and $1/f$ noise," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 53, no. 9, pp. 1869–1884, Sep. 2006.
- [218] M. C. Budge and M. P. Burt, "Range correlation effects in radars," in *The Record of the 1993 IEEE National Radar Conference*, April 1993, pp. 212–216.
- [219] M. Gerstmaier, A. Melzer, A. Onic, and M. Huemer, "On the safe road toward autonomous driving: Phase noise monitoring in radar sensors for functional safety compliance," *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 60–70, 2019.
- [220] C. Aydogdu, M. F. Keskin, G. K. Carvajal, O. Eriksson, H. Hellsten, H. Herbertsson, E. Nilsson, M. Rydstrom, K. Vanas, and H. Wymeersch, "Radar interference mitigation for automated driving: Exploring proactive strategies," *IEEE Signal Processing Magazine*, vol. 37, no. 4, pp. 72–84, 2020.
- [221] A. R. Chiriyath, B. Paul, and D. W. Bliss, "Joint radar-communications information bounds with clutter: The phase noise menace," in *2016 IEEE Radar Conference (RadarConf)*, May 2016, pp. 1–6.
- [222] D. Dardari, P. Closas, and P. M. Djuric, "Indoor tracking: Theory, methods, and technologies," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 4, pp. 1263–1278, April 2015.
- [223] G. Pasolini, A. Guerra, F. Guidi, N. Decarli, and D. Dardari, "Crowd-based cognitive perception of the physical world: Towards the Internet of senses," *Sensors*, vol. 20, no. 9, p. 2437, 2020.
- [224] F. Guidi, A. Guerra, and D. Dardari, "Personal mobile radars with millimeter-wave massive arrays for indoor mapping," *IEEE Transactions on Mobile Computing*, vol. 15, no. 6, pp. 1471–1484, 2016.
- [225] M. Lotti, G. Pasolini, A. Guerra, F. Guidi, R. D'Errico, and D. Dardari, "Radio SLAM for 6G systems at THz frequencies: Design and experimental validation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 17, no. 4, pp. 834–849, July 2023.
- [226] W. Hess, D. Kohler, H. Rapp, and D. Andor, "Real-time loop closure in 2D LIDAR SLAM," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 1271–1278.
- [227] P. Checchin, F. Gérossier, C. Blanc, R. Chapuis, and L. Trassoudaine, "Radar scan matching SLAM using the Fourier-Mellin transform," in *Field and Service Robotics*. Springer, 2010, pp. 151–161.
- [228] H. Saeed, N. Saeed, T. Y. Al-Naffouri, and M.-S. Alouini, "Next generation terahertz communications: A rendezvous of sensing, imaging, and localization," *IEEE Commun. Mag.*, vol. 58, no. 5, pp. 69–75, 2020.
- [229] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, "Massive MIMO is a reality. what is next?: Five promising research directions for antenna arrays," *Digital Signal Processing*, vol. 94, pp. 3–20, 2019, special Issue on Source Localization in Massive MIMO. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1051200419300776>
- [230] Z. Wang, J. Zhang, H. Du, D. Niyato, S. Cui, B. Ai, M. Debbah, K. B. Letaief, and H. V. Poor, "A tutorial on extremely large-scale MIMO for 6G: Fundamentals, signal processing, and applications," 2023.
- [231] E. Björnson, Y. C. Eldar, E. G. Larsson, A. Lozano, and H. V. Poor, "Twenty-five years of signal processing advances for multiantenna communications: From theory to mainstream technology," *IEEE Signal Processing Magazine*, vol. 40, no. 4, pp. 107–117, 2023.
- [232] C. Huang, S. Hu, G. C. Alexandropoulos, A. Zappone, C. Yuen, R. Zhang, M. D. Renzo, and M. Debbah, "Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends," *IEEE Wireless Communications*, vol. 27, no. 5, pp. 118–125, 2020.
- [233] S. Hu, F. Rusek, and O. Edfors, "Beyond massive MIMO: The potential of positioning with large intelligent surfaces," *IEEE Transactions on Signal Processing*, vol. 66, no. 7, pp. 1761–1774, 2018.
- [234] D. Dardari, "Communicating with large intelligent surfaces: Fundamental limits and models," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2526–2537, Nov 2020.
- [235] N. Shlezinger, G. C. Alexandropoulos, M. F. Imani, Y. C. Eldar, and D. R. Smith, "Dynamic metasurface antennas for 6G extreme massive MIMO communications," *IEEE Wireless Communications*, vol. 28, no. 2, pp. 106–113, 2021.
- [236] A. Elzanaty, A. Guerra, F. Guidi, D. Dardari, and M.-S. Alouini, "Towards 6G holographic localization: Enabling technologies and perspectives," *IEEE Internet of Things Magazine*, pp. 1–7, 2023.
- [237] K. T. Selvan and R. Janaswamy, "Fraunhofer and Fresnel distances: Unified derivation for aperture antennas," *IEEE Antennas and Propagation Magazine*, vol. 59, no. 4, pp. 12–15, 2017.
- [238] D. Dardari and N. Decarli, "Holographic communication using intelligent surfaces," *IEEE Communications Magazine*, vol. 59, no. 6, pp. 35–41, June 2021.
- [239] P. Nepa and A. Buffi, "Near-field-focused microwave antennas: Near-field shaping and implementation," *IEEE Antennas and Propagation Magazine*, vol. 59, no. 3, pp. 42–53, 2017.
- [240] H. Zhang, N. Shlezinger, F. Guidi, D. Dardari, and Y. C. Eldar, "6G wireless communications: From far-field beam steering to near-field beam focusing," *IEEE Comm. Magazine*, pp. 1–6, Apr 2023.
- [241] D. A. B. Miller, "Communicating with waves between volumes: evaluating orthogonal spatial channels and limits on coupling strengths," *Appl. Opt.*, vol. 39, no. 11, pp. 1681–1699, Apr 2000. [Online]. Available: <http://ao.osa.org/abstract.cfm?URI=ao-39-11-1681>
- [242] N. Decarli and D. Dardari, "Communication modes with large intelligent surfaces in the near field," *IEEE Access*, vol. 9, pp. 165 648–165 666, 2021.
- [243] A. Guerra, F. Guidi, and D. Dardari, "Single-anchor localization and orientation performance limits using massive arrays: MIMO vs. beamforming," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5241–5255, Aug 2018.
- [244] F. Guidi and D. Dardari, "Radio positioning with EM processing of the spherical wavefront," *IEEE Transactions on Wireless Communications*, vol. 20, no. 6, pp. 3571–3586, June 2021.
- [245] Z. Wang, X. Mu, and Y. Liu, "Near-field integrated sensing and communications," *IEEE Communications Letters*, vol. 27, no. 8, pp. 2048–2052, 2023.
- [246] S. K. Dehkordi, L. Pucci, P. Jung, A. Giorgetti, E. Paolini, and G. Caire, "Multi-static Parameter Estimation in the Near/Far Field Beam Space for Integrated Sensing and Communication Applications," *arXiv e-prints*, p. arXiv:2309.14778, Sep. 2023.
- [247] Z. Yuan, J. Zhang, Y. Ji, G. F. Pedersen, and W. Fan, "Spatial non-stationary near-field channel modeling and validation for massive MIMO systems," *IEEE Transactions on Antennas and Propagation*, vol. 71, no. 1, pp. 921–933, 2023.
- [248] M. Cui and L. Dai, "Channel estimation for extremely large-scale MIMO: Far-field or near-field?" *IEEE Transactions on Communications*, vol. 70, no. 4, pp. 2663–2677, 2022.
- [249] Ö. T. Demir, E. Björnson, L. Sanguinetti *et al.*, "Foundations of user-centric cell-free massive MIMO," *Foundations and Trends® in Signal Processing*, vol. 14, no. 3–4, pp. 162–472, 2021.
- [250] D. Löschbrand, M. Hofer, L. Bernadó, S. Zelenbaba, and T. Zemen, "Towards cell-free massive MIMO: A measurement-based analysis," *IEEE Access*, vol. 10, pp. 89 232–89 247, 2022.
- [251] 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.
- [252] S. Elhoushy, M. Ibrahim, and W. Hamouda, "Cell-free massive MIMO: A survey," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 492–523, 2022.
- [253] E. G. Larsson and J. Vieira, "Phase calibration of distributed antenna arrays," *IEEE Communications Letters*, vol. 27, no. 6, pp. 1619–1623, 2023.
- [254] A. Sakhnini, S. De Bast, M. Guenach, A. Bourdoux, H. Sahli, and S. Pollin, "Near-field coherent radar sensing using a massive MIMO communication testbed," *IEEE Transactions on Wireless Communications*, vol. 21, no. 8, pp. 6256–6270, 2022.
- [255] H. Chen, M. F. Keskin, A. Sakhnini, N. Decarli, S. Pollin, D. Dardari, and H. Wymeersch, "6G localization and sensing in the near field: Fundamentals, opportunities, and challenges," *arXiv preprint arXiv:2308.15799*, 2023.

- [256] M. E. Leinonen *et al.*, "Initial radio models and analysis towards ultra-high data rate links in 6G," Hexa-X project Deliverable D2.2, 2021. [Online]. Available: <https://hexa-x.eu/deliverables/>
- [257] G. Callebaut, L. Liu, T. Eriksson, L. Van der Perre, O. Edfors, and C. Fager, "6G radio testbeds: Requirements, trends, and approaches," *arXiv preprint arXiv:2309.06911*, 2023.
- [258] M. Manzoni, D. Tagliaferri, S. Tebaldini, M. Mizmizi, A. V. Monti-Guarnieri, C. M. Prati, and U. Spagnolini, "Wavefield networked sensing: Principles, algorithms and applications," *arXiv preprint arXiv:2305.10333*, 2023.
- [259] D. Dardari, N. Decarli, A. Guerra, and F. Guidi, "LOS/NLOS near-field localization with a large reconfigurable intelligent surface," *IEEE Transactions on Wireless Communications*, vol. 21, no. 6, pp. 4282–4294, June 2022.
- [260] Z. Abu-Shaban, K. Keykhosravi, M. F. Keskin, G. C. Alexandropoulos, G. Seco-Granados, and H. Wymeersch, "Near-field localization with a reconfigurable intelligent surface acting as lens," in *ICC 2021 - IEEE International Conference on Communications*, 2021, pp. 1–6.
- [261] O. Rinchì, A. Elzanaty, and M.-S. Alouini, "Compressive near-field localization for multipath RIS-aided environments," *IEEE Communications Letters*, vol. 26, no. 6, pp. 1268–1272, 2022.
- [262] C. Ozturk, M. F. Keskin, H. Wymeersch, and S. Gezici, "RIS-aided near-field localization under phase-dependent amplitude variations," *IEEE Transactions on Wireless Communications*, vol. 22, no. 8, pp. 5550–5566, 2023.
- [263] M. Rahal, B. Denis, K. Keykhosravi, B. Uguen, and H. Wymeersch, "RIS-enabled localization continuity under near-field conditions," in *2021 IEEE 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2021, pp. 436–440.
- [264] M. Rahal, B. Denis, K. Keykhosravi, M. F. Keskin, B. Uguen, and H. Wymeersch, "Constrained RIS phase profile optimization and time sharing for near-field localization," in *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, 2022, pp. 1–6.
- [265] C. Ozturk, M. F. Keskin, V. Sciancalepore, H. Wymeersch, and S. Gezici, "RIS-aided localization under pixel failures," *arXiv preprint arXiv:2302.04436*, 2023.
- [266] S. Buzzi, E. Grossi, M. Lops, and L. Venturino, "Foundations of MIMO radar detection aided by reconfigurable intelligent surfaces," *IEEE Transactions on Signal Processing*, vol. 70, pp. 1749–1763, 2022.
- [267] E. Grossi, H. Taremi-Zadeh, and L. Venturino, "Radar target detection and localization aided by an active reconfigurable intelligent surface," *IEEE Signal Processing Letters*, pp. 1–5, 2023.
- [268] S. Buzzi, E. Grossi, M. Lops, and L. Venturino, "Radar target detection aided by reconfigurable intelligent surfaces," *IEEE Signal Processing Letters*, vol. 28, pp. 1315–1319, 2021.
- [269] G. Torcolacci, A. Guerra, H. Zhang, F. Guidi, Q. Yang, Y. C. Eldar, and D. Dardari, "Holographic imaging with XL-MIMO and RIS: Illumination and reflection design," 2023. [Online]. Available: <https://arxiv.org/abs/2312.11102>
- [270] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, "A tutorial on beam management for 3GPP NR at mmWave frequencies," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 173–196, 2019.
- [271] M. Polese, J. M. Jornet, T. Melodia, and M. Zorzi, "Toward end-to-end, full-stack 6G terahertz networks," *IEEE Communications Magazine*, vol. 58, no. 11, pp. 48–54, 2020.
- [272] A. Narayanan, E. Ramadan, J. Carpenter, Q. Liu, Y. Liu, F. Qian, and Z.-L. Zhang, "A first look at commercial 5G performance on smartphones," in *Proceedings of The Web Conference 2020*, ser. WWW '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 894–905. [Online]. Available: <https://doi.org/10.1145/3366423.3380169>
- [273] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Transactions on Communications*, vol. 61, no. 10, pp. 4391–4403, 2013.
- [274] J. Wang, Z. Lan, C.-W. Pyo, T. Baykas, C.-S. Sum, M. A. Rahman, R. Funada, F. Kojima, I. Lakkis, H. Harada, and S. Kato, "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," in *GLOBECOM 2009 - 2009 IEEE Global Telecommunications Conference*, 2009, pp. 1–6.
- [275] N. Garcia, H. Wymeersch, E. G. Ström, and D. Slock, "Location-aided mm-wave channel estimation for vehicular communication," in *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2016, pp. 1–5.
- [276] A. Ali, N. González-Prelcic, and A. Ghosh, "Millimeter wave V2I beam-training using base-station mounted radar," in *2019 IEEE Radar Conference (RadarConf)*, 2019, pp. 1–5.
- [277] 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. 55 513–55 525, 2020.
- [278] K. Satyanarayana, M. El-Hajjar, A. A. M. Mourad, and L. Hanzo, "Deep learning aided fingerprint-based beam alignment for mmWave vehicular communication," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 11, pp. 10 858–10 871, 2019.
- [279] S. Park, A. Ali, N. González-Prelcic, and R. W. Heath, "Spatial channel covariance estimation for hybrid architectures based on tensor decompositions," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 1084–1097, 2020.
- [280] A. Graff, A. Ali, and N. González-Prelcic, "Measuring radar and communication congruence at millimeter wave frequencies," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, 2019, pp. 925–929.
- [281] A. Ali, N. González-Prelcic, and A. Ghosh, "Passive radar at the roadside unit to configure millimeter wave vehicle-to-infrastructure links," *IEEE Trans. Veh. Technol.*, vol. 69, pp. 14 903–14 917, 2020.
- [282] Y. Chen, A. Graff, N. González-Prelcic, and T. Shimizu, "Radar aided mmWave vehicle-to-infrastructure link configuration using deep learning," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 01–06.
- [283] F. Liu, W. Yuan, C. Masouros, and J. Yuan, "Radar-assisted predictive beamforming for vehicular links: Communication served by sensing," *IEEE Transactions on Wireless Communications*, vol. 19, no. 11, pp. 7704–7719, 2020.
- [284] W. Yuan, F. Liu, C. Masouros, J. Yuan, D. W. K. Ng, and N. González-Prelcic, "Bayesian predictive beamforming for vehicular networks: A low-overhead joint radar-communication approach," *IEEE Trans. Wireless Commun.*, vol. 20, pp. 1442–1456, 2021.
- [285] Y. Tian, G. Pan, and M.-S. Alouini, "Applying deep-learning-based computer vision to wireless communications: Methodologies, opportunities, and challenges," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 132–143, 2021.
- [286] W. Xu, F. Gao, S. Jin, and A. Alkhateeb, "3D scene-based beam selection for mmWave communications," *IEEE Wireless Commun. Lett.*, vol. 9, pp. 1850–1854, 2020.
- [287] M. Alrabeiah, A. Hredzak, Z. Liu, and A. Alkhateeb, "ViWi: A deep learning dataset framework for vision-aided wireless communications," in *Proc. IEEE 91st Veh. Technol. Conf. (VTC2020-Spring)*, 2020, pp. 1–5.
- [288] W. Xu, F. Gao, X. Tao, J. Zhang, and A. Alkhateeb, "Computer vision aided mmWave beam alignment in V2X communications," *IEEE Trans. Wireless Commun.*, vol. 22, pp. 2699–2714, 2023.
- [289] M. Zecchin, M. B. Mashhadi, M. Jankowski, D. Gunduz, M. Kountouris, and D. Gesbert, "LIDAR and position-aided mmWave beam selection with non-local CNNs and curriculum training," *IEEE Trans. Veh. Technol.*, vol. 71, pp. 2979–2990, 2022.
- [290] A. Ali, J. Mo, B. L. Ng, V. Va, and J. C. Zhang, "Orientation-assisted beam management for beyond 5G systems," *IEEE Access*, vol. 9, pp. 51 832–51 846, 2021.
- [291] M. B. Mashhadi, M. Jankowski, T.-Y. Tung, S. Kobus, and D. Gunduz, "Federated mmWave beam selection utilizing LIDAR data," *IEEE Wireless Commun. Lett.*, vol. 10, pp. 2269–2273, 2021.
- [292] M. Dias, A. Klautau, N. González-Prelcic, and R. W. Heath, "Position and LIDAR-aided mmWave beam selection using deep learning," in *Proc. IEEE 20th Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*, 2019, pp. 1–5.
- [293] W. Xu, F. Gao, J. Zhang, X. Tao, and A. Alkhateeb, "Deep learning based channel covariance matrix estimation with user location and scene images," *IEEE Trans. Commun.*, vol. 69, pp. 8145–8158, 2021.
- [294] B. Salehi, G. Reus-Muns, D. Roy, Z. Wang, T. Jian, J. Dy, S. Ioannidis, and K. Chowdhury, "Deep learning on multimodal sensor data at the wireless edge for vehicular network," *IEEE Trans. Veh. Technol.*, vol. 71, pp. 7639–7655, 2022.
- [295] T. Zhang, J. Liu, and F. Gao, "Vision aided beam tracking and frequency handoff for mmWave communications," in *Proc. IEEE Int. Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, 2022, pp. 1–2.
- [296] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Transactions on Communications*, vol. 63, no. 9, pp. 3029–3056, 2015.
- [297] A. Karttunen, A. F. Molisch, S. Hur, J. Park, and C. J. Zhang, "Spatially consistent street-by-street path loss model for 28-GHz channels in micro cell urban environments," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7538–7550, Nov. 2017.

- [298] M. Boban, D. Dupleich, N. Iqbal, J. Luo, C. Schneider, R. Müller, Z. Yu, D. Steer, T. Jämsä, J. Li, and R. S. Thomä, "Multi-band vehicle-to-vehicle channel characterization in the presence of vehicle blockage," *IEEE Access*, vol. 7, pp. 9724–9735, 2019.
- [299] G. R. MacCartney, S. Deng, S. Sun, and T. S. Rappaport, "Millimeter-wave human blockage at 73 GHz with a simple double knife-edge diffraction model and extension for directional antennas," in *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*, Sep. 2016, pp. 1–6.
- [300] V. Raghavan, L. Akhondzadeh-Asl, V. Podshivlov, J. Hulten, M. A. Tassoudji, O. H. Koymen, A. Sampath, and J. Li, "Statistical blockage modeling and robustness of beamforming in millimeter-wave systems," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 7, pp. 3010–3024, 2019.
- [301] Y. Oguma, R. Arai, T. Nishio, K. Yamamoto, and M. Morikura, "Proactive base station selection based on human blockage prediction using RGB-D cameras for mmWave communications," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2015, pp. 1–6.
- [302] M. Alrabeiah, A. Hredzak, and A. Alkhateeb, "Millimeter wave base stations with cameras: Vision-aided beam and blockage prediction," in *Proc. IEEE 91st Veh. Technol. Conf. (VTC2020-Spring)*, 2020, pp. 1–5.
- [303] G. Charan, M. Alrabeiah, and A. Alkhateeb, "Vision-aided 6G wireless communications: Blockage prediction and proactive handoff," *IEEE Trans. Veh. Technol.*, vol. 70, pp. 10 193–10 208, 2021.
- [304] G. Reus-Muns, B. Salehi, D. Roy, T. Jian, Z. Wang, J. Dy, S. Ioannidis, and K. Chowdhury, "Deep learning on visual and location data for V2I mmWave beamforming," in *Proc. 17th Int. Conf. Mobility, Sens., Netw. (MSN)*, 2021, pp. 559–566.
- [305] D. Marasinghe, N. Rajatheva, and M. Latva-aho, "LiDAR aided human blockage prediction for 6G," in *Proc. IEEE Global Commun. Conf. Workshops (GC Wkshps)*, 2021, pp. 1–6.
- [306] M. Nerini and B. Clerckx, "Overhead-free blockage detection and precoding through physics-based graph neural networks: LIDAR data meets ray tracing," *IEEE Wireless Commun. Lett.*, vol. 12, pp. 565–569, 2023.
- [307] S. Wu, C. Chakrabarti, and A. Alkhateeb, "Proactively predicting dynamic 6G link blockages using LiDAR and in-band signatures," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 392–412, 2023.
- [308] U. Demirhan and A. Alkhateeb, "Radar aided proactive blockage prediction in real-world millimeter wave systems," in *Proc. IEEE Int. Conf. Commun.*, 2022, pp. 4547–4552.