



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

By NURIA GONZÁLEZ-PRELCIC^{ID}, *Senior Member IEEE*, MUSA FURKAN KESKIN^{ID}, *Member IEEE*, OSSI KALTIOKALLIO^{ID}, *Member IEEE*, MIKKO VALKAMA^{ID}, *Fellow IEEE*, DAVIDE DARDARI^{ID}, *Senior Member IEEE*, XIAO SHEN^{ID}, *Student Member IEEE*, YUAN SHEN^{ID}, *Senior Member IEEE*, MURAT BAYRAKTAR, *Graduate Student Member IEEE*, AND HENK WYMEERSCH^{ID}, *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 (UE), and the sensing capability of the communication signal itself provide a new source of data that connects the physical and radio frequency (RF) 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 such as blockage and better support adaptation in dynamic environments as networks reconfigure. In this article, we provide a vision for integrated sensing and communication (ISAC) networks and an overview of how signal processing, optimization, and machine learning (ML) 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.

KEYWORDS | Distributed joint sensing and communication (S&C); integrated sensing and communications (ISACs); monostatic sensing; near-field ISAC; radio positioning; radio simultaneous localization and mapping (SLAM); reconfigurable intelligent surface (RIS)-aided localization; sensing-aided communication.

I. INTRODUCTION

Early work on integrated sensing and communication (ISAC) targeted the development of cooperation strategies to enable spectral coexistence between communication

Manuscript received 2 January 2024; revised 1 April 2024; accepted 21 April 2024. The work of Nuria González-Prelcic and Murat Bayraktar was supported in part by the National Science Foundation under Grant 2147955, in part by the Federal Agency and Industry Partners through the Resilient and Intelligent NextG Systems (RINGS) Program, in part by Samsung Research America, and in part by Toyota Motor North America. The work of Musa Furkan Keskin and Henk Wymeersch was supported in part by the Swedish Research Council under Grant VR 2022-03007; and in part by Hexa-X-II, part of the European Union's (EU) Horizon Europe Research and Innovation Programme, under Grant 101095759. The work of Ossi Kaltioikallio and Mikko Valkama was supported in part by the Academy of Finland under Grant 352754 and Grant 359095, and in part by the Business Finland through the "6G-ISAC" Project. The work of Davide Dardari was supported in part by the HORIZON-JU-SNS-2022 Project TIMES under Grant 101096307; and in part by the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on "Telecommunications of the Future" (Program "RESTART") co-funded by EU, under Grant PE00000001. The work of Xiao Shen and Yuan Shen was supported in part by the National Natural Science Foundation of China under Grant 62271285 and in part by Shanghai AI Laboratory. (Corresponding author: Nuria González-Prelcic.)

Nuria González-Prelcic and **Murat Bayraktar** are with the Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA 92093 USA (e-mail: ngprelcic@ucsd.edu; mbayraktar@ucsd.edu).

Musa Furkan Keskin and **Henk Wymeersch** are with the Department of Electrical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden (e-mail: furkan@chalmers.se; henkw@chalmers.se).

Ossi Kaltioikallio and **Mikko Valkama** are with the Unit of Electrical Engineering, Tampere University, 33014 Tampere, Finland (e-mail: ossi.kaltioikallio@tuni.fi; mikko.valkama@tuni.fi).

Davide Dardari is with the Department of Electrical, Electronic, and Information Engineering "Guglielmo Marconi," University of Bologna, 40126 Bologna, Italy, and also with WiLAB-CNIT, 40133 Bologna, Italy (e-mail: davide.dardari@unibo.it).

Xiao Shen and **Yuan Shen** are with the Department of Electronic Engineering and Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China (e-mail: shenx20@mails.tsinghua.edu.cn; shenyuan_ee@tsinghua.edu.cn).

Digital Object Identifier 10.1109/JPROC.2024.3397609

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, which motivated the initial designs of joint sensing and communication (S&C) systems [2]. Nowadays, the number of potential avenues to integrating S&C 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 S&C with motivations and benefits that go beyond conventional ones. 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 (ML) strategies to be exploited, or the S&C 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 S&C 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 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 article, we describe different frameworks for integrating S&Cs 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 S&C, 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 to 6G and beyond. For example, Hassanien et al. [8] described 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- 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, such as the impact of self-interference (SI), are not described. Technologies for sensor-aided communication, a new form of ISAC that may play a critical role in increasing resilience and adaptability in 6G networks, are not discussed in any of these surveys. Finally, work [3] includes a study of performance limits and tradeoffs, aspects of waveform design from a communication and radar perspective, and only a one-page description of receiver (RX) 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 RX that enable network sensing and sensor-aided communication, as this article does. In addition, the role of ML tools in the design of the physical layer of ISAC networks is hardly considered in the previous surveys.

The structure of the technical sections in this article 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 multistatic S&C that provide the sensing information directly from the fronthaul link, downlink (DL), uplink (UL), 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-terahertz (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 appropriate isolation between the transmitter (TX) and RX enables simultaneous transmission of the communication

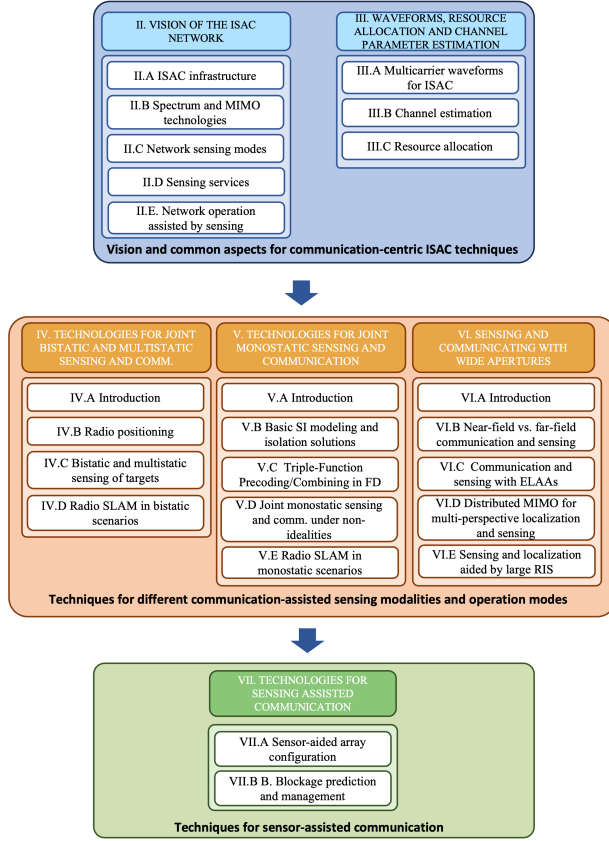


Fig. 1. Article's outline, covering the technical sections.

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, sensor-aided communication, where sensing information (e.g., user's position) can be leveraged to enhance the network operation, for example, significantly reducing the overhead associated with link configuration and reconfiguration or enabling early blockage detection. Throughout this article, 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 S&C. It offers S&C 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 S&C as part of an automated driving package. An ISAC network, and pertinent components of infrastructure, is 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 such as mmWave, there are low-power RISs 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 needs not to be terrestrial; ISAC also supports nonterrestrial components where the BSs are untethered to the ground in the form of satellites or unmanned aerial vehicles (UAVs). 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 UEs in the network. These sensors use a spectrum that is different from the communication spectrum. However, 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 bird's-eye view of the environment [12]. Sensing is also facilitated by reusing the communication signal for radio sensing. A network with ISAC reuses the existing communication waveforms possibly with more capable hardware (e.g., an 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 refer to the storage of past sensor data and communication performance data, which are collected over time. Computation refers to the training of ML 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 the core network. The computation capabilities will vary significantly at these

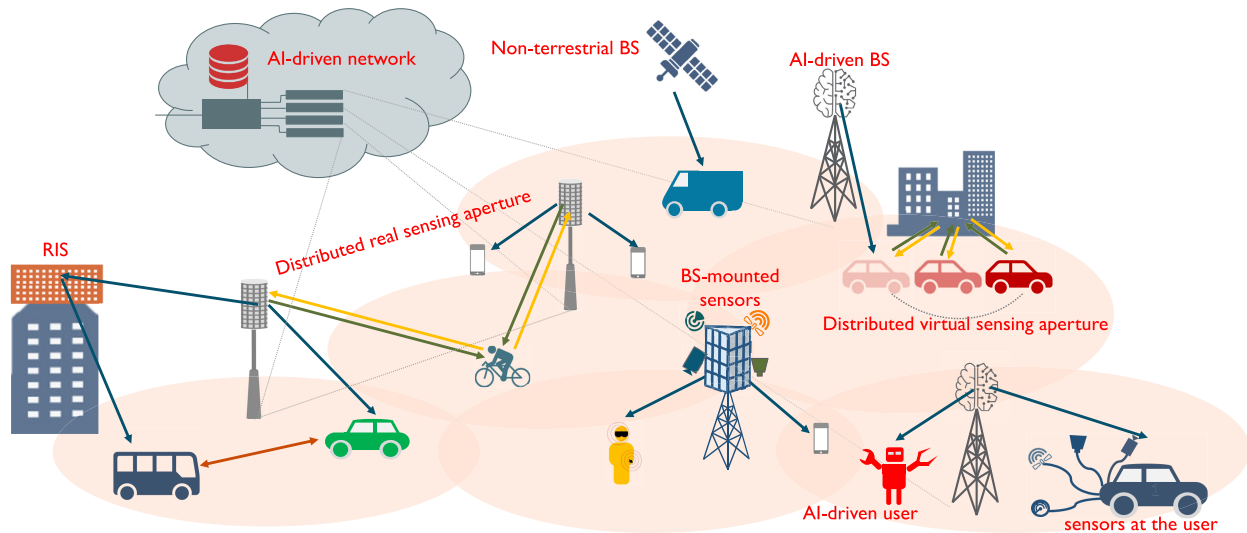


Fig. 2. Illustration of the different elements in the infrastructure of an ISAC network.

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 such as smart phones, watches, and glasses; low-capability connected devices such as smart meters and location tags; and high-capability connected devices such as 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 artificial intelligence (AI)-driven BS could use data to optimize how it balances S&C 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 S&C 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 is important from the sensing perspective.

- 1) *Frequency range 1 (FR1):* This band spans from around 400 MHz to 7 GHz. In this band, the bandwidths between 5 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.
- 2) *Frequency range 2 (FR2):* This band spans from around 24 to 70 GHz, with supported bandwidths ranging from 50 to 400 MHz. Due to the higher path loss, the 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 propagation, shadowing is more pronounced leading to fewer propagation paths and a more geometric channel.
- 3) *Upper mid-band:* This band lies between 7 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 tradeoff 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 due to the exploitation of extremely large arrays and increased bandwidth [13].

- 4) *Sub-THz*: The sub-THz bands span from 100 to 300 GHz. This band is envisioned for extremely high data rates in nearly static conditions. Because of large bandwidths and large arrays for fine, high-gain beams, resolution is expected to be high, but the 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 for 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 a wide variety of sensing services. Multiband 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 databases, rather than conventional stochastic channel models. Such a common dataset 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 RF chain per antenna element so that all the signal processing operations are performed digitally. 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 reduce the number of bits in the analog-to-digital converters (ADCs) and/or digital-to-analog converters (DACs). 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 S&C 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 a number of N_T transmit antennas and N_R receive antennas, the beamforming operation is

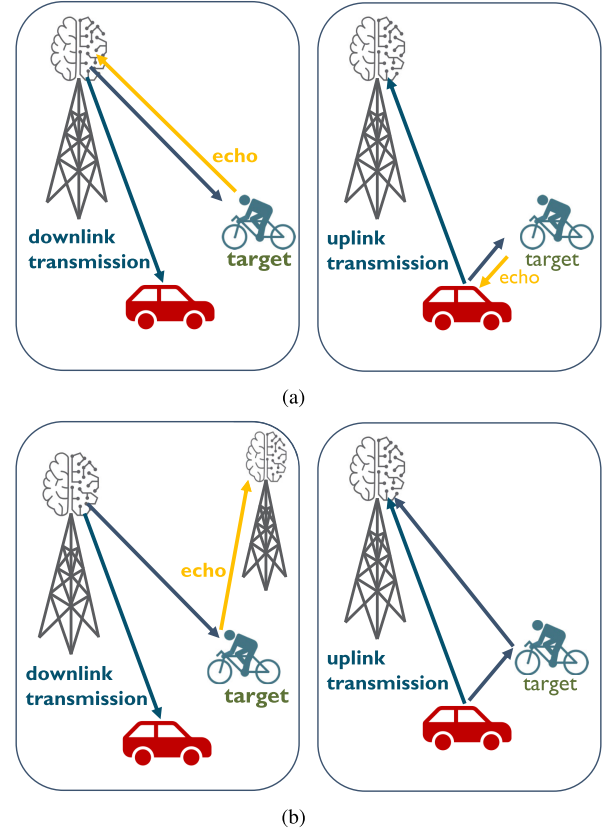


Fig. 3. Illustration of different network sensing modes.

(a) *Monostatic sensing*: the TX and RX are co-located, share a common clock, and are 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 TX and RX are physically separated. It can be performed in the UL or the DL, and the TX and RX can be two BSs, or one BS and one user.

represented by multiplication by a beamforming vector $\mathbf{f} \in \mathcal{F}^{N_T \times 1}$ at the TX and a weight beamforming vector $\mathbf{w} \in \mathcal{W}^{N_R \times 1}$ at the RX, with $\mathcal{F}$ and $\mathcal{W}$ the sets of possible phases at the TX and RX. 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 same set of phase shifters is used for the entire band. The design of analog beamformers for joint S&C 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 S&C metrics in addition to the phase shifter constraints and SI mitigation requirements.

A hybrid precoding architecture provides an interesting performance/complexity tradeoff. 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 $N_{\text{RF},R} < N_R$. In this way, the precoding matrix can be represented by $\mathbf{F} = \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}$, where $\mathbf{F}_{\text{RF}}$ is the analog precoding matrix and $\mathbf{F}_{\text{BB}}$ is the digital or baseband precoding matrix. Analogously, $\mathbf{W} = \mathbf{W}_{\text{RF}}\mathbf{W}_{\text{BB}}$, with $\mathbf{W}_{\text{RF}}$ the analog combiner and $\mathbf{W}_{\text{BB}}$ the digital combiner. The additional digital precoding/combining stage enables frequency-selective spatial processing. Moreover, the higher number of degrees of freedom (DoFs) in a hybrid design allows mutibeam solutions, making it suitable for multistream, multiuser, or multitarget scenarios. The S&C 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 high quantization noise. The investigation of specific designs and their performance in the context of ISAC systems is very limited [18], [19] and will not be further discussed in this article.

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 article.

1) *Sensing, Positioning, and 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 to a service. The received radio signals in general depend on the geometric state of the TX, RX, and the environment (e.g., radar sensing), though not all sensing services rely on this geometry (e.g., pollution monitoring). Hence, localization of connected users and passive objects relies on sensing information. In 3GPP, positioning refers to localizing UEs from UL, DL, 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 article, 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) *Monostatic, Bistatic, and Multistatic Sensing*: Sensing is conventionally broken down into three types, though many other forms of sensing exist, which are not covered in this article.

- 1) *Monostatic sensing*: As illustrated in Fig. 3(a), the TX and RX are co-located, share a common clock, and are 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 TX. Hence, the TX may be a BS or a mobile UE. In the former case, sensed objects are tracked, based on DL 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.
- 2) *Bistatic sensing*: The TX and RX are physically separate, as shown in Fig. 3(b). If the TX or RX 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 an SLAM problem in a global coordinate system. If both TX and RX 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.
- 3) *Multistatic sensing*: There are several TXs and/or several RXs, all physically separated. Pilot signals or some form of multiplexing is needed in case there are several TXs, 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 RXs.

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

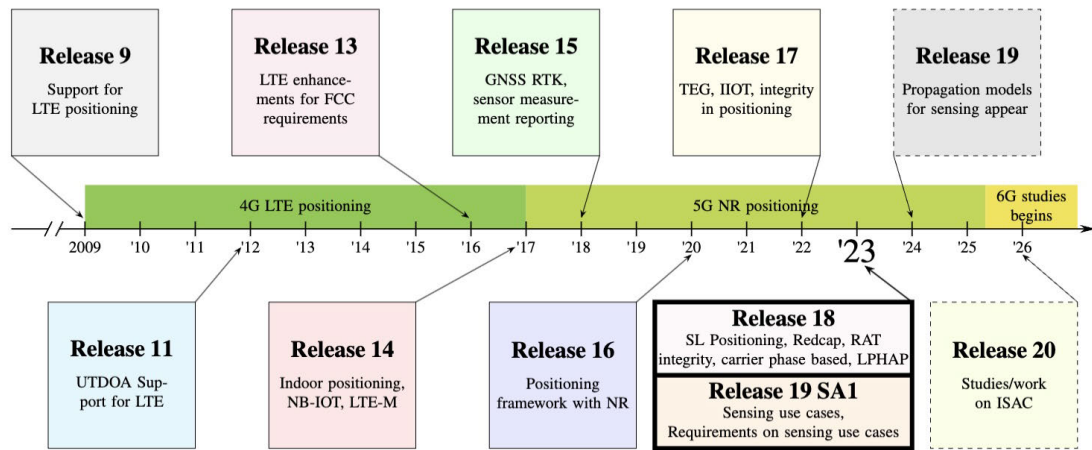


Fig. 4. Evolution of positioning across 3GPP releases and beginning of sensing. The dashed box shows the undecided release of 3GPP.

estimating UE's location has been built over many years. One aspect of this evolution is the change in the accuracy requirement. It started with positioning in release 9 long term evolution (LTE) networks, with the aim of meeting the regulatory requirement of 50-m accuracy in positioning a UE. The regulatory requirement was the mainstay for building the positioning specifications in nearly all LTE releases. Positioning signals, measurements, procedures, and architectures were specified to meet this requirement. The new radio (NR) in 5G supports a 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 the initial phases of standardization, selective use cases will be prioritized. A possible prioritization of use cases can be the following.

- 1) *Smart home/building intrusion detection:* Intrusion detection in buildings or surroundings of smart home.
- 2) *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.

- 3) *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.
- 4) *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, modeling of sensing targets in terms of their physical scattering surface as radar cross section (RCS) and modeling 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, and so on. 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 infrastructure 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 such as blockage [12], [26]. The huge amount of information that this type of network can collect is the basis for

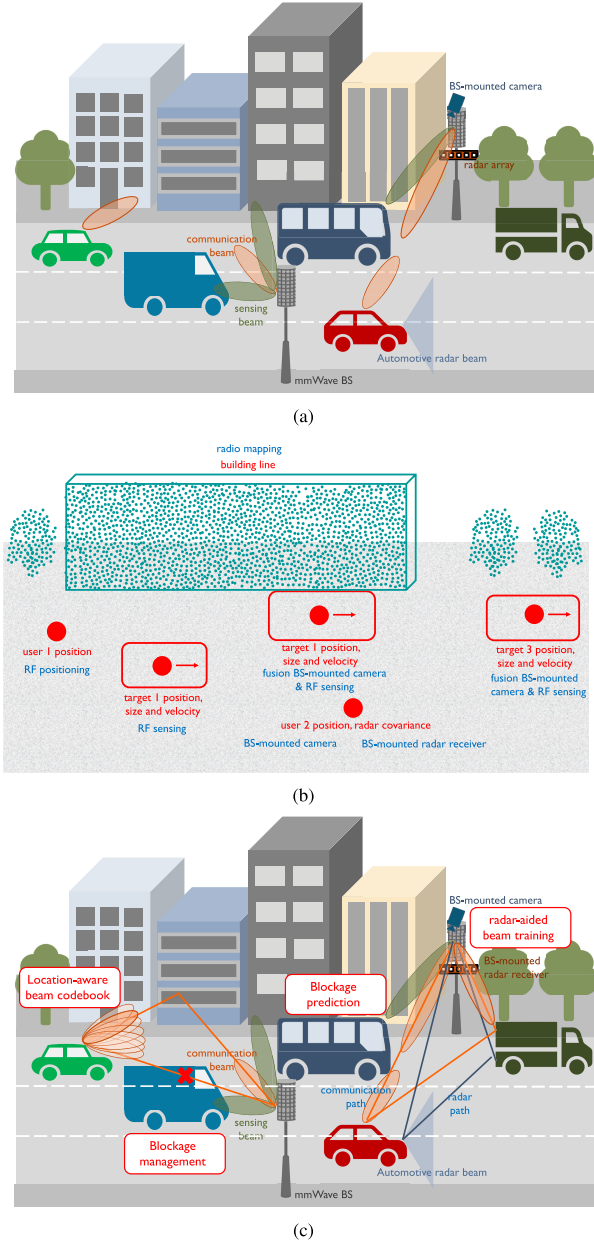


Fig. 5. Illustration of the sensing-assisted communication concept. (a) Joint sensing and vehicular communication system supported by a cellular network in an urban scenario. (b) ISAC map corresponding to the urban scenario in (a), showing the detected users, targets, and scatterers and the technology 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.

exploiting 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], [28], [29], [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 radio (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 such as type, size, and trajectory. 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, and so on. 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 with 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 RX deployed at the BS, and the information about the radar channel can be exploited as a prior to reducing 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 wireless local area network (WLAN)/Wi-Fi 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 a flexible and reconfigurable physical layer, in terms of multicarrier symbol durations while supporting also backward compatibility and coexistence of LTE and NR. Complementary filtering and windowing [38], [39] can also be added, either at the TX or the RX or both, in an essentially transparent manner [40], to enhance the waveform spectral containment. In addition, 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 TX, 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 a highly dynamic envelope is problematic from the power amplifier (PA) point of view, as the PA power efficiency is commonly improved if operating closer toward the saturating region. Such operation point, however, also implies highly nonlinear PA behavior, and 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 discrete Fourier transform (DFT)-spread OFDM (DFT-s-OFDM), where the precoding across the subcarriers helps to reduce the PAPR—especially with contiguous spectrum allocations. Such a DFT-s-OFDM approach is supported in the UL of LTE/LTE-Advanced and NR. OFDM is also known to be sensitive to oscillator phase noise (PN), carrier frequency offsets (CFOs), 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 one is the key aspect, both from the communications RX and the sensing RX perspectives. In addition, while the ordinary RX implementations build commonly on OFDM symbol-wise fast Fourier transform (FFT) processing, extending this to 2-D FFT/inverse FFT (IFFT) pairs over multiple symbol durations provides the 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 2-D 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 toward mmWave bands, the symbol durations and the corresponding CP lengths are reduced, and thus, this may become a more obvious limitation if not properly handled. In addition, the long symbol durations of OFDM waveforms may easily lead to intercarrier interference (ICI), calling for attention to devising OFDM-based ISAC and sensing systems. When properly handled, such a phenomenon can also be turned from a foe to a friend, and described and demonstrated later 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 CP. In addition, 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, e.g., [42], [43], [44], [45]. These aspects are also discussed further in Section III-C.

2) *Unified Communication/Localization/Sensing Signal Model With MIMO-OFDM*: In this section, 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 this article. Extending the model in [46] to the case of time-varying channels, the received signal $\mathbf{y}_{n,m} \in \mathbb{C}^{N_{\text{RF}}, R \times 1}$ 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 , $\mathbf{F}_{\text{BB}}[n, m] \in \mathbb{C}^{N_{\text{RF}}, T \times N_s}$ is the digital baseband precoder at subcarrier n and symbol m , $\mathbf{F}_{\text{RF}} \in \mathbb{C}^{N_T \times N_{\text{RF}}, T}$ is the analog RF precoding matrix applied in the time domain for the entire bandwidth, $\mathbf{W}_{\text{RF}} \in \mathbb{C}^{N_R \times N_{\text{RF}}, R}$ denotes the analog combining matrix at the RX, and $\mathbf{H}_{n,m} \in \mathbb{C}^{N_R \times N_T}$ is the channel at subcarrier n and symbol m , given by²

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

where $\mathbf{a}_T(\theta) \in \mathbb{C}^{N_T \times 1}$ and $\mathbf{a}_R(\phi) \in \mathbb{C}^{N_R \times 1}$ denote the array steering vectors at the TX and RX, respectively; and 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 CFO), 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 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., intersymbol) phase shifts represented by $e^{j2\pi m T_{\text{sym}} \nu_{\ell}}$ but also *fast-time* (i.e., intrasymbol) 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 considered 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 angle of departure (AoD), 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 m away from a BS, the angle estimation accuracy should be approximately 1°). 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, low-pass filtering after downconversion, and MF. In this case, the channel

¹The 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 (MF) through the incorporation of complex coefficients on a per-subcarrier, per-path basis, as shown in [46, eq. (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).

model in (2) has to be modified to introduce a time-domain function, which represents all the filtering stages that impact the baseband equivalent model [46]. This function contributes to the entanglement of the channel parameters and complicates its estimation. In summary, the channel estimation for joint S&C 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 structures 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 S&C can be classified into three main categories: based on maximum-likelihood estimation [51], [52], [53], [54], [55], exploiting compressed sensing [46], [56], [57], [58], or subspace-based estimators. Some types of techniques might be more suitable for a particular frequency band than others, as discussed in the following.

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 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 signal-to-noise ratio (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. In this way, the

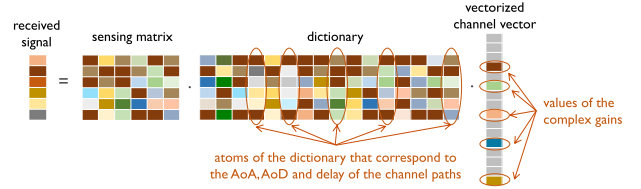


Fig. 6. Illustration of the channel estimation algorithms exploiting sparsity. The MIMO channel matrices are expanded in terms of a sparsifying dictionary and vectorized to create the sparse vector to be estimated. Identifying the AoD/AoA and delay boils down to the selection of the columns (atoms) in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

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], [57], [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 TX and at the RX evaluated on a grid for the AoD and the 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 arrays (ULAs) 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 $\mathbf{h}_{\text{vec}}$ 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 AoD, which contains the transmit steering vectors evaluated on a grid of potential AoDs; $\mathbf{A}_{\text{rx}}$ is the dictionary for the AoA, 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 $\mathbf{y}_{\text{td}}$, 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 $\mathbf{h}_{\text{vec}}$ 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 over-complete 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], [61], [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], [67], [68] design a new codebook for accurate angle estimation in a DL localization scenario. The following example shows the significantly better performance provided by the new design.

Example 1: We consider a 5G/6G DL 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 DL multiple-input-single-output (MISO) scenario with LOS-only propagation in a 2-D setup, where $N_T = 16$ and $N_R = 1$. In this case, using (1) and (2), the received signal at the single-antenna UE is given by

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

where $\mathbf{f}_m \in \mathbb{C}^{N_T \times 1}$ 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 the 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 $\mathbf{F} = [\mathbf{f}_0 \dots \mathbf{f}_{M-1}] \in \mathbb{C}^{N_T \times M}$ 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 $\mathbf{F}$.

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

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

- 2) *New codebook*: Recently proposed *directional/derivative* codebook [66], [67], [68] (similar to *sum/difference* beams used in monopulse radar [71]) given by

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

Here, $\dot{\mathbf{a}}_T(\theta) = \partial \mathbf{a}_T(\theta) / \partial \theta$ and $\{\theta_g\}_{g=1}^{2G}$ represent uniformly sampled grid points from $\mathcal{U}$, and each column of $\mathbf{F}^{\text{dir}}$ 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 $\dot{\mathbf{a}}_T(\theta)$ alongside standard directional beams $\mathbf{a}_T(\theta)$ is motivated by the need for the UE to detect subtle deviations around the intended direction θ and also supported by the theoretical Cramér–Rao lower bound (CRLB)

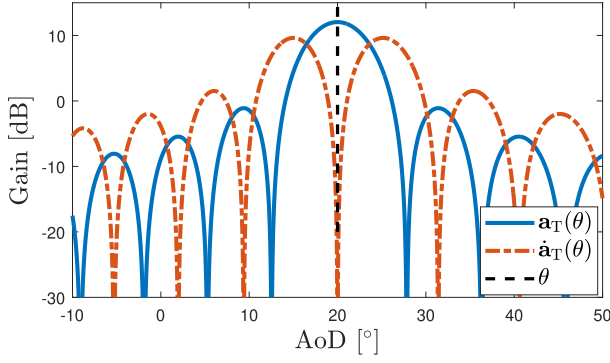


Fig. 7. 16-element ULA beam patterns of directional $\mathbf{a}_T(\theta)$ and derivative $\mathbf{a}_T'(\theta)$ beams for the UE located at an AoD $\theta = 20^\circ$ with respect to the BS. The directional beam ensures the necessary SNR for AoD estimation, whereas the derivative beam assists the UE 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.

analysis [66], [67]. The sharp curvature around θ in the beampattern of $\mathbf{a}_T'(\theta)$ 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 $\mathbf{F}^{\text{dir}}$ 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 $\text{SNR} = |\alpha|^2 N_T / \sigma^2 = 0$ dB, using the maximum-likelihood 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 $\mathbf{F}^{\text{dir}}$, 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 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 performances. Therefore, the design of resource allocation schemes is 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 recover 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

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

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

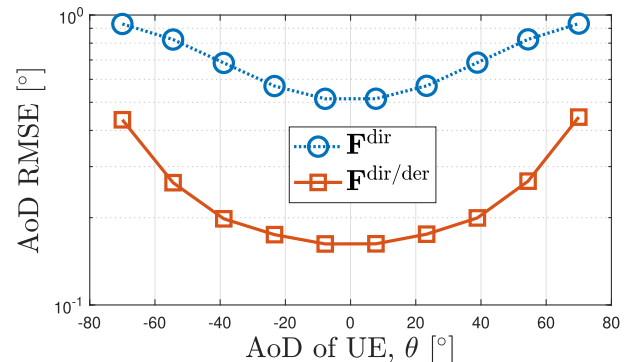


Fig. 8. AoD estimation performance with respect to the AoD of the UE, achieved by the considered codebooks in (7) and (8).

the Schur complement of submatrix in the original 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. 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.

- 1) 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

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

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 $\mathbf{F}_{\text{RF}}$ and $\mathbf{F}_{\text{BB}}$ at the TX, and the combining matrices $\mathbf{W}_c$ and $\mathbf{W}_s$ at the communication and sensing RXs. Note that the joint problem (12) is hard to solve due to the nonconvexity 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], [76], [77]. In addition, the above formulation reduces the resource allocation in localization and sensing networks with the communication constraints $\mathbf{h}_c$ removed, where more theoretical insights and 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, 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].

- 2) 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.,

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

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 $\mathbf{F}_{\text{RF}}$ and $\mathbf{F}_{\text{BB}}$ 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 twofold 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].

- 3) 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

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

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 and w_s denote the weights for S&C performance. 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, 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]. In addition, the objective function can refer to the total resource consumption of the ISAC system in the joint formulation, where the S&C performance is 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 TXs and RXs are separated. Sensing where TX and RX are co-located will be treated in Section V. When sensing is based on one TX and one RX, it is called bistatic [89], while sensing based on several TXs or RXs 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 multifaceted field. Before we delve into the technical aspects, we first provide a brief overview of the key concepts, as 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 sensing and multistatic sensing, which will be covered in Section IV-C, involve TXs and RXs 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 involving 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 3-D 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, 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 that 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 $\mathbf{p}_{\text{UE}}$ is the UE position, $\mathbf{p}_{\text{BS}}$ is the BS position, c is the speed of light, and τ_{bias} is the clock bias of the UE with respect to the same BS [97]. Similarly, the Doppler measurements are affected by the CFO between the UE and BS, and the angles at the UE side (i.e., AoA in DL or AoD in UL) depend on the unknown user orientation, which is a 3-D 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 (due 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 $\mathbf{y}_{\text{meas}}$ comprises the estimated angles, delays, and Dopplers; $\mathbf{x}_{\text{state}}$ comprises the 3-D UE location as well as any nuisance parameters (clock bias, CFO, and 3-D UE orientation); and $\mathbf{f}(\cdot)$ is a 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 nonconvex, due to the nonlinear relation between the measurements and the UE state, so heuristics/approximations/relaxations are employed 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:* While 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 positioning and the design of new positioning methods or technologies that can reduce the reliance on infrastructure is of great interest. These cases are called minimal problems/minimal solvers [104], in the sense that if measurements or technologies are removed, the problem

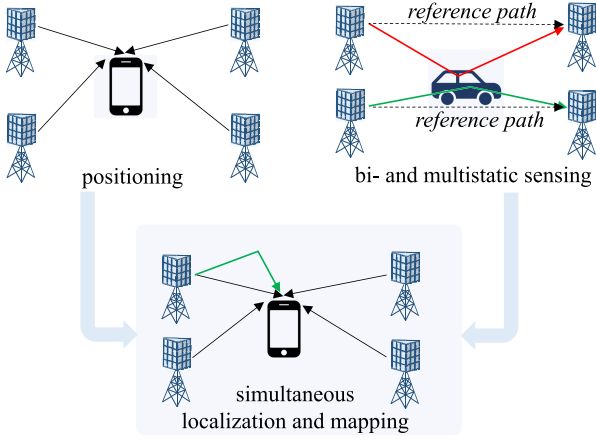


Fig. 9. Breakdown of bistatic and multistatic sensing. Operation is shown in DL but can equivalently occur in UL.

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 four 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 four observations of the form (15). These delay measurements are sufficient to determine the 3-D UE location and the 1-D 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 three (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.

- 1) *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 and 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.
- 2) *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 that they can be resolved. In particular, single-bounce NLoS paths are characterized by a single 3-D incidence point (IP). While this IP 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].

- 3) *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 is shown in Section IV-B8.
- 4) *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.
- 5) *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 two BSs, each equipped with a planar array, transmitting DL pilots to a UE. From the DL 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 two lines in 3-D, which is a unique point (see the left 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 two lines in 3-D away from the UE. Since the

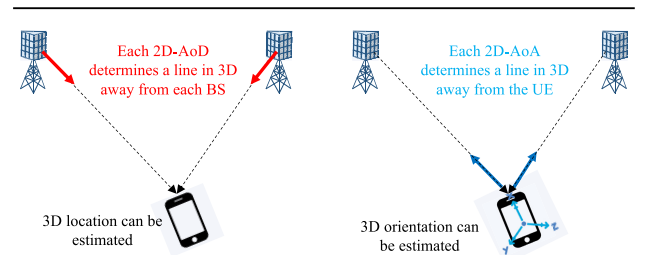


Fig. 10. Left: example of how a UE 3-D position can be estimated based on the 2D-AoD from two multi-antenna BSs. Right: once the 3-D UE location is known, a multi-antenna UE can use the 2D-AoA to determine its 3-D orientation.

Table 1 Examples of Minimal Configurations Needed to Solve the DL 3-D Localization Problem Without A Priori Knowledge of IP Locations

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)

two lines are parameterized by four parameters and the UE orientation has only three DoFs, the UE orientation can be uniquely determined (see the right 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 1. 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.

- 1) *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 m [117]. Hence, if NLoS paths arrive within 15 m of the LoS path, these paths will merge and appear 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).
- 2) *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 that paths are well separated in the angle domain, as seen from the BS [118].
- 3) *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, multibounce 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 ML (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-m 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 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].

- 1) *Large bandwidth*: At FR2, significantly large bandwidths are available, up to 400 MHz, which corresponds to a distance resolution of less than 1 m. 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.
- 2) *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 toward 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].

- 3) *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 TX and RX, in part because shadowing is so severe [127]. Multibounce 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: 1) communication and positioning are more closely intertwined and 2) 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 2-D scenario shown in Fig. 11, with a BS that defines the coordinate system, a UE with unknown 2-D location $\mathbf{p}_{\text{UE}}$, 1-D clock bias τ_{bias} , and 1-D orientation ϕ_{UE} , and a scatter point with unknown 2D location $\mathbf{p}_{\text{SP}}$. From UL 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 $\phi_0 = \phi_{\text{UE}} + \theta_0 + \pi$, from which the UE orientation ϕ_{UE} can be solved. We also immediately find that the angle $\psi = \pi - (\phi_1 - \phi_0) - (\theta_1 - \theta_0)$. All the remaining angles follow 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 the UE ($\mathbf{v}_i$ from the AoD) so that the RX

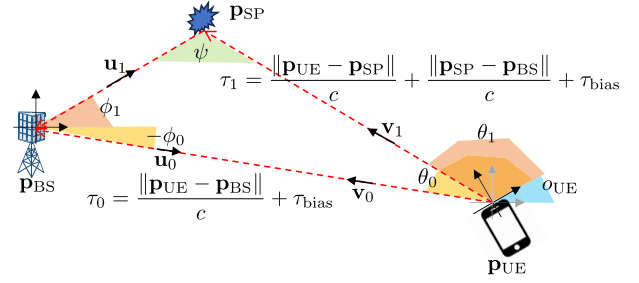


Fig. 11. Illustrative example of mmWave positioning with one BS by exploiting the uncontrolled multipath via a scatter point with unknown location.

position can be defined as

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

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

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

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

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

where $\mathbf{H}_i = [\mathbf{I}, -c\mathbf{v}_i]$, $\mathbf{x}_{\text{UE}} = [\mathbf{p}_{\text{UE}}^\top, \tau_{\text{bias}}]^\top$, and $\bar{\boldsymbol{\nu}}_i = \bar{\boldsymbol{\nu}}_i / \|\bar{\boldsymbol{\nu}}_i\|$. Now, the closed-form solution can be obtained by setting the cost function's gradient to zero and solving for $\mathbf{x}_{\text{UE}}$. It is important to note that for the LoS path ($i = 0$), we have $\mathbf{u}_0 = -\mathbf{v}_0$ and $\bar{\boldsymbol{\nu}}_0 = 0$. 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 nonexhaustive list of examples of single-BS positioning is provided in Table 2, one based on controlled multipath (from an 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 multiantenna BS or RIS. Note also that the examples in Table 2 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 Versus mmWave: Performance Analysis via Ray-Tracing Data: In this section, we carry

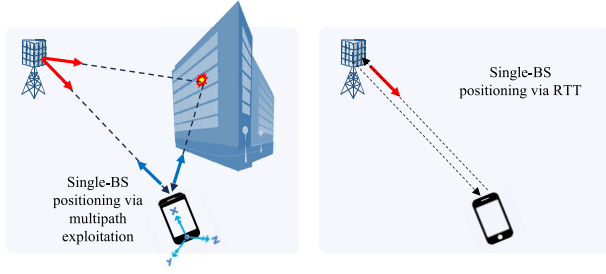
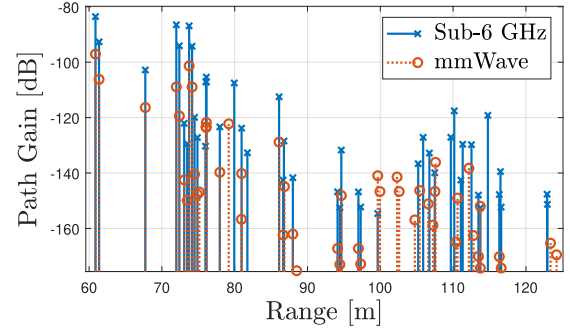


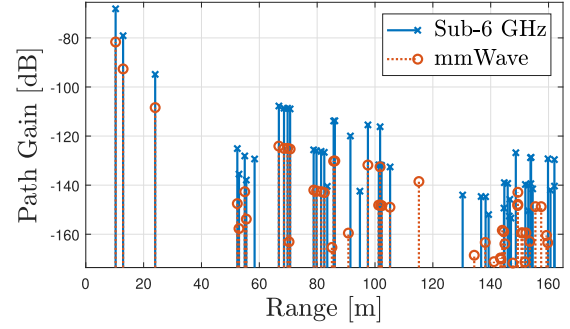
Fig. 12. Visualization of single-BS positioning approaches.

out a comparative performance analysis of positioning in sub-6 GHz and mmWave bands using realistic ray-tracing data obtained through the REMCOM Wireless InSite³ ray tracer [129]. In the ray tracer, we consider an urban intersection scenario involving: 1) a mobile UE (corresponding to a vehicle) crossing the intersection; 2) a BS (corresponding to a road-side unit (RSU) in vehicular settings [130]) located at the center of the intersection; and 3) 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 a 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



(a)



(b)

Fig. 13. PDPs 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) UE located at -60 m, leading to a rich multipath environment due to surrounding buildings. (b) UE located at 5 m, leading to a small number of dominant paths.

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]: 1) an MF-based method that performs correlation processing across frequency and spatial domains and 2) 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 Sections IV-B3 and IV-B4. Fig. 13 shows the path gains with respect to the 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 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

Table 2 Examples of Minimal Configurations for Single-BS Positioning

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 on intersection of sphere from RTT and line away from BS from 2D-AoA.	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		N/A

³Registered trademark.

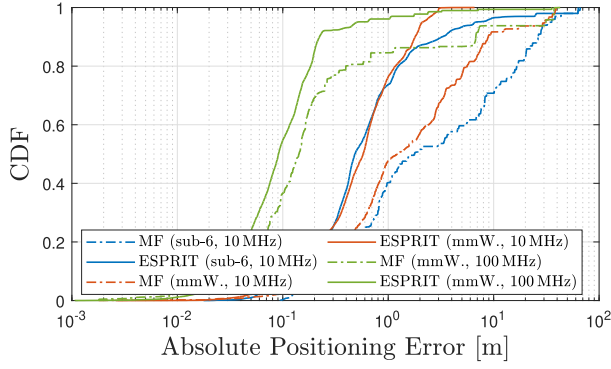


Fig. 14. Positioning performances of the MF and ESPRIT algorithms at sub-6 GHz and mmWave under different bandwidths, evaluated using the ray-tracing data. The parameters at sub-6 GHz are $f_c = 5.9$ GHz, $\Delta f = 60$ kHz, $M = 12$, and the BS array configuration 1×2 (2.54 cm), while the parameters 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.

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 Section IV-B3 and IV-B4, two key takeaways can be deduced from the ray-tracing-based simulation results: 1) mmWave induces favorable channel characteristics for positioning and 2) 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) *ML 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, NLOS multipath acts as an interference that degrades the geometric localization performance in both indoor and outdoor urban environments. At mmWave frequencies, strategies, such as 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 PN, 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], [133], [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 an offline phase to create a database of fingerprints (one or more parameters associated with 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 $\mathbf{p}_{\text{UE}}$. The drawbacks of conventional fingerprinting come from: 1) the requirement of permanently storing a large database and updating it periodically as the environment changes [139] and 2) 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, is highly dependent on the device, and suffers from a high variability due to multipath. Recent works on ML-based fingerprinting for massive MIMO at sub-6 GHz 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 sub-carrier [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. In this way, the first stage of the network itself extracts a suitable feature using the attention mechanisms in Transformer networks [146], [147], [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], [150], [151], [152], [153], [154]. Long short-term memory (LSTM) networks have also been 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 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 to exploit 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 an 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, Chen et al. [147] proposed 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 20 cm for 95% of the time when evaluated with realistic vehicular channels generated by ray tracing.

Example 5 (ML-Based Positioning): Consider an mmWave vehicular communication system operating with a hybrid MIMO architecture and uniform planar arrays (UPAs) at both ends. During initial access, the frequency-selective mmWave channel between a vehicle and a single-BS is estimated using multidimensional orthogonal pursuit matching (MOMP) [50] and some training symbols. The hybrid data/model-driven positioning system proposed in [100] and [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

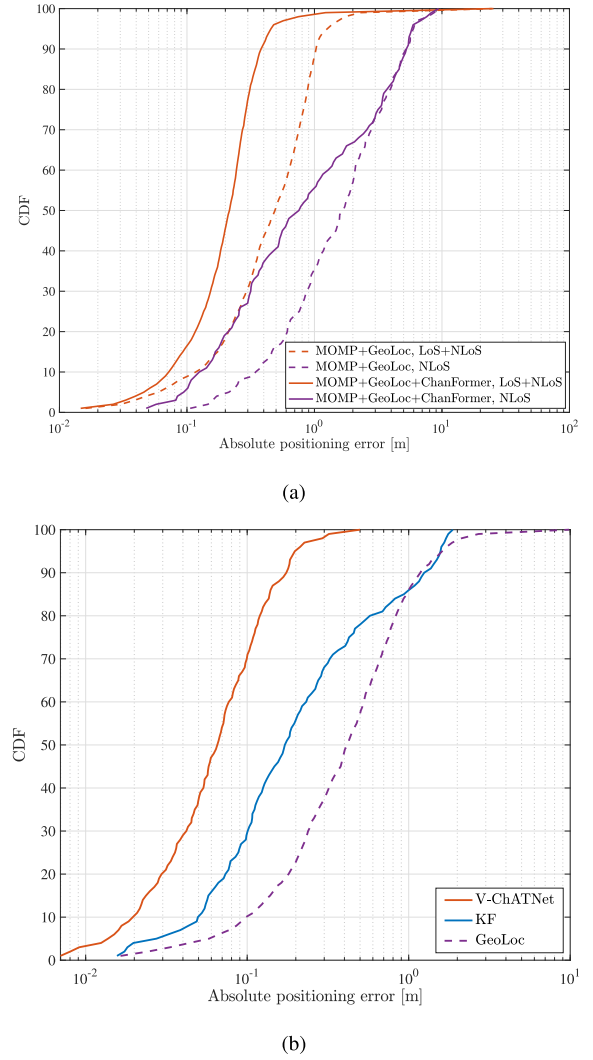


Fig. 15. Performance of the positioning strategies based on attention networks—ChanFormer for the initial access scenario and V-ChATNet for the tracking case—compared to positioning based only on channel parameter estimation and exploitation of the geometry of the environment (MOMP + GeoLoc) evaluated using the ray-tracing setup described in [146] and [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 a series of channel and position estimates is 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.

grid of points built around the initial position estimation. As illustrated in Fig. 15(a), the system achieves submeter accuracy localization for 95% of the users in channels with a LoS path when evaluated with ray-tracing channels. In addition, the system provides submeter 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 20-cm 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 that 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 PRSs over the sidelinks between UEs. Finally, in the position calculation stage, the ToA and the 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 has been extensive research in this area. Liu et al. [114] described basic system architectures and key technologies for high-accuracy sidelink positioning. Ge et al. [160] presented the 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 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] and [163], where the structure of EFIM is revealed to characterize the information gain brought by UE cooperation. Then, the spatiotemporal 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], [165], [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*: An 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 an 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 α_m^{RIS} , which is of the form [170]

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

in which $\alpha^{\text{T-RIS}}$ denotes the channel from TX to RIS; $\alpha^{\text{RIS-R}}$ is the channel from RIS to RX; $\mathbf{a}_{\text{RIS}}(\varphi)$ is the RIS steering vector as a function of the AoA φ (from the RX) and AoD ϑ (to the TX); and Ω_m denotes the diagonal RIS control matrix, which represents the time-varying state of the RIS. The entries of Ω_m have at most unit amplitude (since a passive RIS cannot amplify) and controllable phase. In communications, Ω_m can be set so that $\mathbf{a}_{\text{RIS}}^T(\vartheta) \Omega_m \mathbf{a}_{\text{RIS}}(\varphi) \rightarrow M$, 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 simultaneous transmit-and-reflect (STAR)-RIS) [171], and RISs that allow nondiagonal control Ω_m [172].

The role of RIS for positioning lies in the ability to provide: 1) an additional position reference, similar to a BS; 2) additional measurements as any other path [e.g., the AoA, AoD, delay, and Doppler in (23)]; and 3) 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 TX or the RX is a BS. This leads to two practical issues: how can we know the RIS location and orientation [174] and how can the RX 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 $\sum_m \Omega_m = 0$, so that under zero Doppler, the 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

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

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 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 preestimated in the RIS calibration stage. The positioning stage contains two different cases: 1) LoS path from the BS and RIS exists and 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 mmWave 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

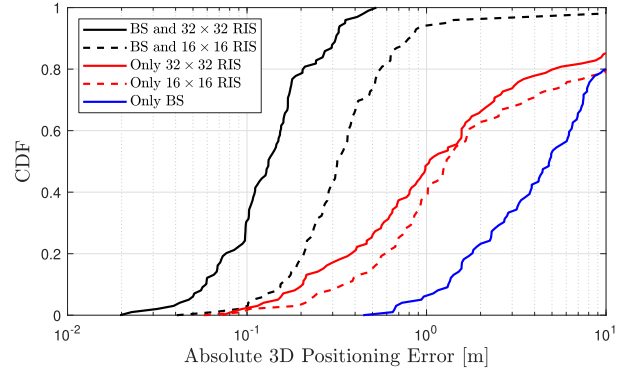


Fig. 16. RIS-aided 3-D positioning performance with an 8×8 BS. Evaluated using the ray-tracing data from a smart factory environment with transmit power 20 dBm, center frequency $f_c = 60$ GHz, and bandwidth $B = 100$ MHz.

GNSS-RTK approach, see, e.g., [176] and the references therein. In addition, and importantly, the utilization of carrier phase measurements is also recently 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 such a purpose.

An example of recent works in this area in the cellular system context covers, e.g., [181] and [182] focusing on

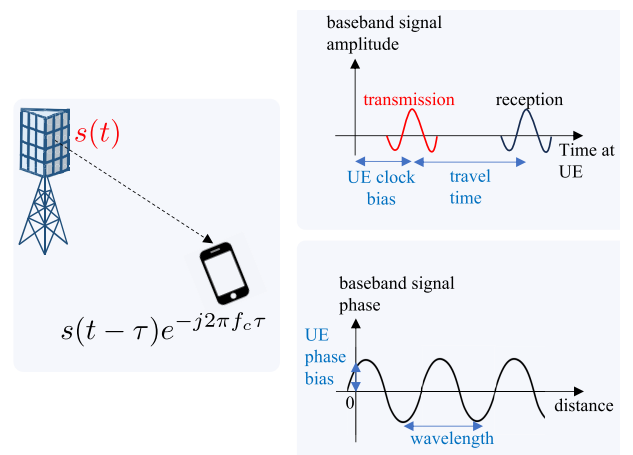


Fig. 17. Carrier phase positioning provides accurate distance information, but is subject to ambiguities equal to the wavelength.

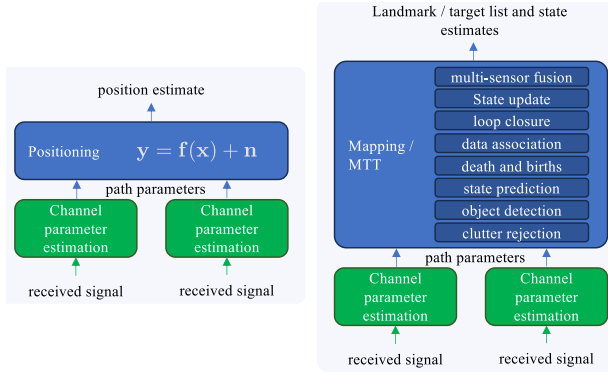


Fig. 18. Both positioning and sensing require channel parameter estimation. The subsequent processing for mapping and MTT is much more involved than for positioning.

estimation, synchronization, and positioning algorithms for different use cases. In addition, Talvitie et al. [108] described the methods of 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 Nonconnected Objects

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

1) *Fundamentals of Mapping and Target Tracking: Multisite 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 *multitarget 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, and Dopplers); and 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 nonconnected objects. The main difference

lies in the subsequent processing, as visualized in Fig. 18. The reasons are given as follows [183].

- 1) *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 detection (involving false alarms and missed detections) and estimation problem.
- 2) *Complicated object states:* In contrast to positioning, where the goal is to estimate the UE 3-D 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].
- 3) *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 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].
- 4) *Multisensor fusion:* When several RXs are processing the signals from one or more TXs, 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, this 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 (FOV) of the different RXs 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 [55], [96], [187], and [188]) since it is closely related to the minimal

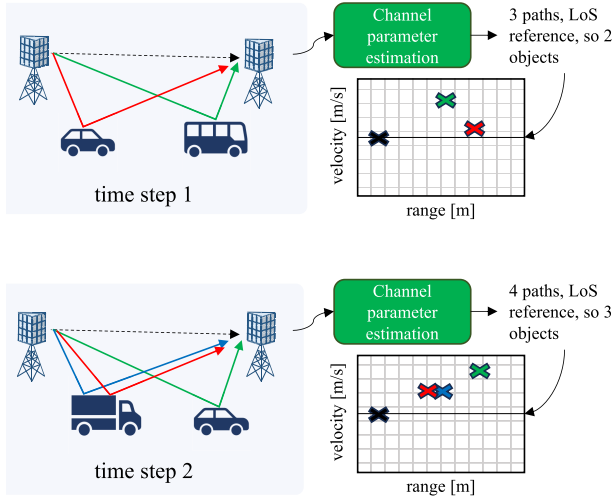


Fig. 19. Each measurement gives rise to a potential object. Objects may generate 0, 1, or several measurements at each time. Objects detected at time step 1 should be associated with measurements at time step 2.

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, $p(\mathbf{m} | \mathbf{x}_{1:t}, \mathbf{y}_{1:t})$, 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]

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

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 (EKF) 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, $p(\mathbf{x}_{1:t}, \mathbf{m} | \mathbf{y}_{1:t})$, 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]

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

in which $p(\mathbf{x}_{1:t} | \mathbf{y}_{1:t})$ and $p(\mathbf{m} | \mathbf{x}_{1:t}, \mathbf{y}_{1:t})$ 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 $p(\mathbf{x}_{1:t} | \mathbf{y}_{1:t})$ and computing $p(\mathbf{m} | \mathbf{x}_{1:t}, \mathbf{y}_{1:t})$ is equivalent to solving (26). It is important to note that each particle represents a single UE trajectory and a unique map is associated with 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 $p(\mathbf{x}_{t'}, \mathbf{m} | \mathbf{y}_{t'})$, filtering approaches utilize measurements up to time t for approximating $p(\mathbf{x}_t, \mathbf{m} | \mathbf{y}_{1:t})$, and smoothing approaches perform batch processing to estimate the full SLAM posterior $p(\mathbf{x}_{1:T}, \mathbf{m} | \mathbf{y}_{1:T})$ in which $T \geq t$. The early works in conventional SLAM mainly considered the filtering problem, but due to inconsistency issues, the 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, the 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: 1) 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 and 2) 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

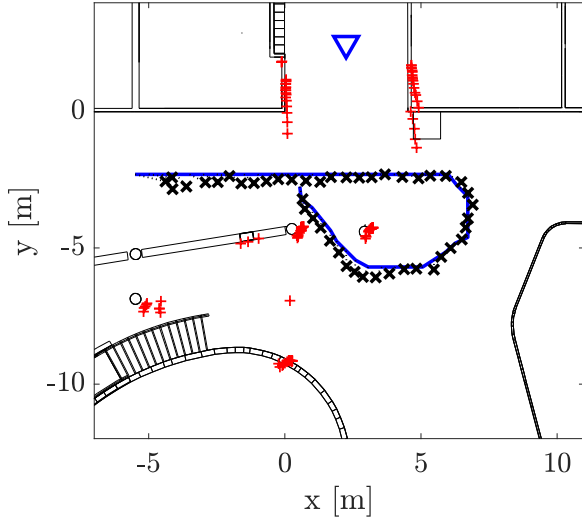


Fig. 20. Example bistatic radio SLAM performance obtained using an RBPF-PHD filtering solution. The TX position is illustrated with a blue triangle (∇), UE trajectory using a blue line (—), and estimated UE path and landmark locations with black crosses (X) and red crosses (+), respectively.

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 composed of a single BS that transmits PRSs that are received by the UE to jointly estimate its pose and map the surrounding environment. The experiment was conducted at the 60-GHz carrier frequency and using the 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° 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; 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 threefold. First, 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: 1) the landmark can be misdetected due to limitations in the RX and channel estimation routine; 2) clutter measurements and multibounce signals can generate false detections that are not inline with the measurement model; and 3) measurement ambiguities can lead to situations where a wrong measurement is associated with the wrong landmark.

V. TECHNOLOGIES FOR JOINT MONOSTATIC SENSING AND COMMUNICATION

A. Introduction

In this section, we address the monostatic sensing and SLAM paradigms in the spirit of cellular ISAC. We first briefly 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 the examples how precoding/beamforming optimization can contribute to SI suppression. Section V-D then looks into relevant impairments, most notably oscillator PN and Doppler-induced ICI, and shows how these can be turned from foe to friend in monostatic sensing. Finally, Section V-E presents monostatic SLAM, which is one potential application of monostatic sensing.

In general, different from the previously discussed bistatic and multistatic TX–RX arrangements, the *monostatic* case refers to a scenario where the transmitting and receiving entities are essentially collocated [194], [195], [196]. A principal illustration of such monostatic 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 a monostatic 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 bistatic or multistatic counterparts, the monostatic approach is appealing as the complete I/Q transmit waveform is directly known also to the RX. In addition, arranging for very accurate time and frequency synchronization between the TX and RX chains is clearly much more straightforward and can be resolved even at a hardware level.

The monostatic operation is, as such, well known from ordinary radars. Pulsed radars [194] operate based on time multiplexing between the exact transmit and receive time

periods. In the context of cellular systems, specifically in time-division duplexing (TDD) networks, the individual network nodes [UEs, gNBs, and the corresponding transmission and reception points (TRPs)] also operate based on dividing the TX and RX active periods in time. However, the individual active transmit periods in both UL and DL 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 RX simultaneous to transmitting, in order to measure the reflecting and scattering waves. This, in turn, means that such a sensor is essentially operating as an inband full-duplex (IBFD) transceiver, commonly also referred to as an STAR system [195], [196], [197].

One fundamental technical challenge in STAR systems is, in general, the SI or direct transmit–receive coupling [197], [198], [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 RX 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], [198], [199], [200], [201]. Importantly, sufficient amount of TX–RX isolation must be obtained already in the antenna or RF circuit domain prior to the RX low noise amplifier (LNA)—otherwise, the powerful SI will totally desensitize and block the LNA and thereon the sensing RX. In 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 simultaneously operating TX and RX, namely, sharing a set of antennas through circulators for both 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 commonly provide only isolation levels in the order of 20 dB, while the different sets of antennas allow for larger isolation levels by separating the antenna systems in space—especially in gNB type of entities as well as in UE with larger form factors such as vehicles. In addition, 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 RX antenna j , contributed by the TX subarray l , reads

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

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 the convolution, and the function $\phi_l^{\text{TX}}(\cdot)$ refers to a TX hardware model that can accommodate different impairments such as PA nonlinearity and oscillator PN. In addition, 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) and (29) show that both TX beamforming 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 trade-offs between SI suppression, target illumination capability, and achievable communication rate are described, e.g., in [202], [203], [204], and [205], and will also be discussed in Section V-C. Furthermore, the accommodation of the prominent hardware imperfections in (28) and (29) is of great importance. For one, as shown, e.g., in [200], [206], and [207], appropriate HW modeling allows to craft advanced digital SI cancellation solutions that can also suppress the nonlinear effects and distortion products imposed by TX PA, RX LNA, and other analog components. Such one is important when optimizing the RX dynamic range. What is more, as discussed and demonstrated concretely later in this section, proper modeling of hardware imperfections can turn them from a foe to a friend, when interpreted from the sensing point of view—this is

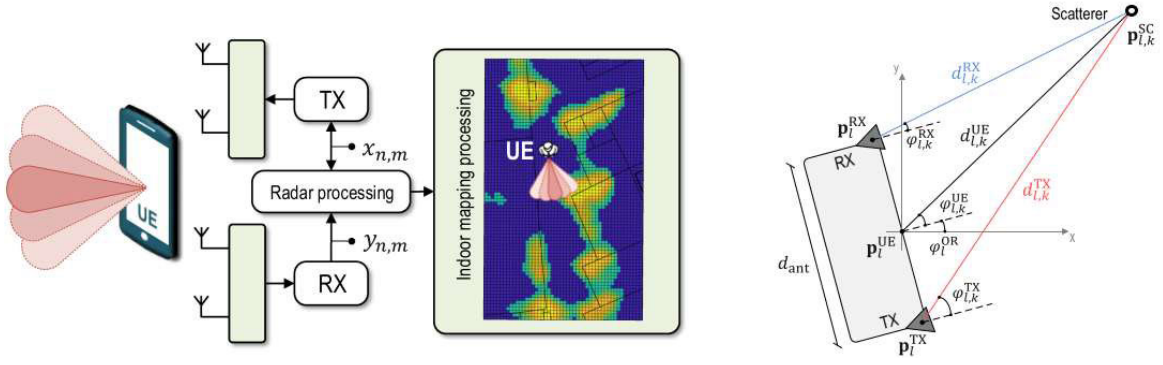


Fig. 21. Illustration of the monostatic sensing paradigm at UE end, together with the corresponding problem geometry.

particularly so for the oscillator PN as discussed further in Section V-D.

C. Triple-Function Precoding/Combining in FD Transceivers

In this section, we discuss the precoder and combiner design that enable the 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: 1) enabling DL communication; 2) providing high gain in the target direction; and 3) 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 $\mathbf{H}_{n,m}$, $\mathbf{H}_n^{\text{SI}} \in \mathbb{C}^{N_R \times N_T}$ 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

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

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

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

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

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

for $n_{\text{RF}} = 1, \dots, N_{\text{RF},R}$. 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

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

where $\mathcal{A}_T$ and $\mathcal{A}_R$ are the sets of feasible structures for the analog precoder and the analog combiner, respectively. 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. Islam et al. [203] maximized the data rate by assuming that a subset of the targets in the environment contribute to DL 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 maximizes 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 reduces the SI suppression capability. Finally, a convex relaxation-based algorithm is employed to design the combiner that minimizes 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 can 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 functions 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 comprises directional beams. Let us denote the directional codebooks for TX and RX by

$\mathbf{A}_{\text{tx}} = [\mathbf{a}_T(\theta_1) \dots \mathbf{a}_T(\theta_{M_T})] \in \mathbb{C}^{N_T \times M_T}$ and $\mathbf{A}_{\text{rx}} = [\mathbf{a}_R(\phi_1) \dots \mathbf{a}_R(\phi_{M_R})] \in \mathbb{C}^{N_R \times M_R}$, where M_T and M_R are the sizes of AoA set $\{\theta_i\}_1^{M_T}$ and AoD set $\{\phi_j\}_1^{M_R}$, 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 to design as $\bar{\mathbf{A}}_{\text{tx}} \in \mathbb{C}^{N_T \times M_T}$ and $\bar{\mathbf{A}}_{\text{rx}} \in \mathbb{C}^{N_R \times M_R}$. The coupling matrix that shows the residual SI between beam pairs can be defined as $\mathbf{C} = \bar{\mathbf{A}}_{\text{tx}}^H \mathbf{H}^{\text{SI}} \bar{\mathbf{A}}_{\text{rx}}$, where $\mathbf{H}^{\text{SI}} \in \mathbb{C}^{N_R \times N_T}$ is the frequency-flat SI channel. For a wideband SI channel, $\mathbf{H}^{\text{SI}}$ can be selected as the SI channel at the center frequency. The goal of 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 nonconvex. The TX and RX codebooks are designed with alternating minimization and convex relaxation combined with block-coordinate descent method by using $\bar{\mathbf{A}}_{\text{tx}}$ and $\bar{\mathbf{A}}_{\text{rx}}$ as initial codebooks in [209]. The beam training procedure for sensing is designed such that a TX–RX beam pair that corresponds 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 the 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 S&C 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 Nonidealities: Repurposing Challenges Into Benefits

In this section, we address the problem of joint monostatic S&C with OFDM waveform in the presence of nonidealities 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 nonidealities, namely, ICI and PN,

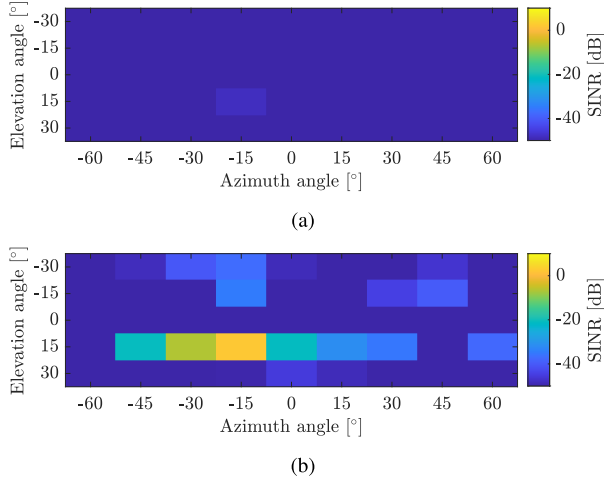


Fig. 22. Radar SINR at the FD BS equipped with 8×8 arrays that are separated by 10° for (a) initial codebook and (b) SI-aware codebook. The center frequency is 28 GHz, the bandwidth is 300 GHz, the transmit power is 20 dBm, and the SI-to-noise ratio is 80 dBm. The target is located at 80 m from the FD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

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 RXs.

1) *OFDM Monostatic Radar Sensing Without Nonidealities*: Consider a monostatic ISAC setup in Fig. 23, where a monostatic ISAC transceiver sends data symbols to a remote communications RX and collects the backscattered signals at the co-located radar RX for sensing the objects in the environment. Under an 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- and 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 MF (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/inverse 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 model (36) diminishes since intrasymbol Doppler-induced phase shifts become nonnegligible, destroying subcarrier orthogonality and leading to ICI. Under scenarios with high mobility and/or small Δf , (36) can be generalized to [49]

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

where $\mathbf{D}(\nu) = \text{diag}(1, e^{j2\pi(T/N)\nu}, \dots, e^{j2\pi(T(N-1)/N)\nu}) \in \mathbb{C}^{N \times N}$ encodes the intrasymbol (*fast time* in radar nomenclature) phase shifts as a function of Doppler ν . We note that intersymbol (*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 sidelobe 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, nonideal oscillators, PN should be considered in sensing. In this case, (36) becomes [216]

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

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 TX and the radar RX, as shown in Fig. 23, $\mathbf{W}(\tau_\ell)$ corresponds to a realization of a differential

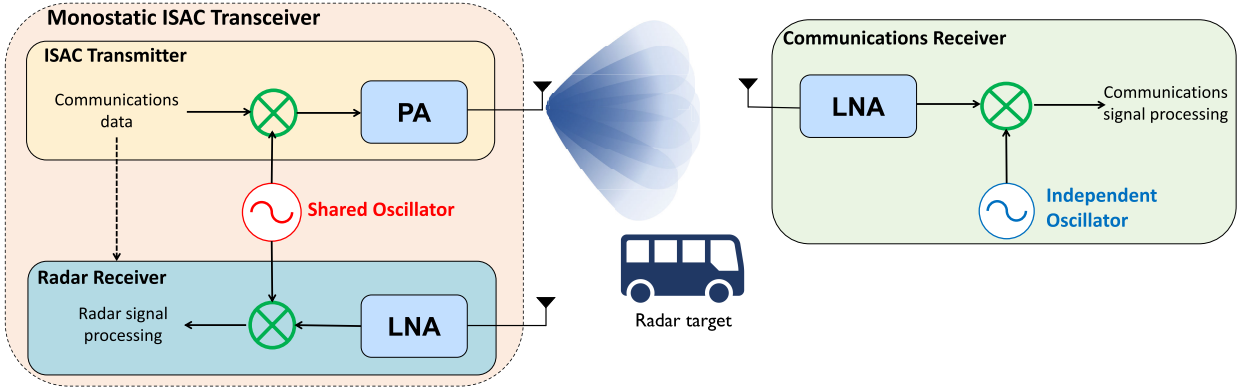


Fig. 23. Monostatic ISAC setup with a monostatic ISAC transceiver containing an ISAC TX and a radar RX co-located on shared hardware, and a communications RX located on a remote device. In contrast to the independent oscillator employed at the communications RX, the use of a shared oscillator in monostatic sensing implies that impairments, such as PN and 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 3 for further details.

(self-correlated) PN process [217], which has *delay-dependent statistics* [216] (the so-called *range correlation effect* [218], [219], [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 S&Cs in ISAC systems when impairments, such as ICI and PN, are present. Specifically, in a monostatic sensing configuration, the radar RX shares the same oscillator with the ISAC TX. In contrast, the communications RX uses an independent oscillator. This implies that: 1) the ICI effect in (37) involves only target Doppler

without any CFO [49] and 2) 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 3 summarizes the differences between monostatic S&Cs 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: 1) resolving Doppler ambiguity of high-speed targets and 2) enhancing target

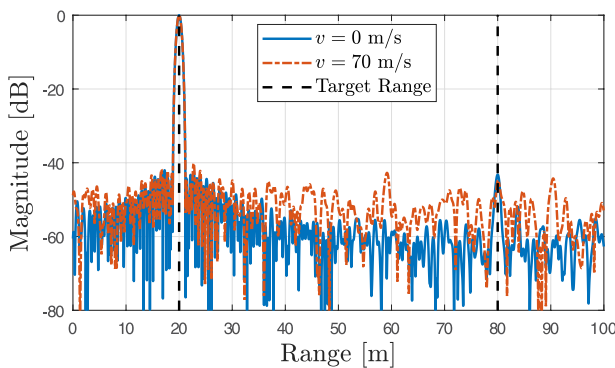


Fig. 24. Range spectrum in OFDM-based sensing under the impact of ICI. The OFDM parameters are $f_c = 28$ GHz, $\Delta f = 60$ kHz, $N = 4096$, and $M = 16$. The scenario includes two targets with the same velocity v , the ranges (20, 80) m, and the SNRs (25, -20) dB. $v = 70$ m/s represents the case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

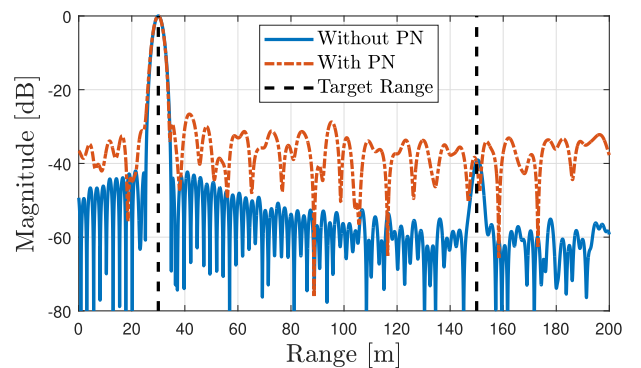


Fig. 25. Range spectrum in OFDM-based sensing under the impact of PN. The OFDM parameters are $f_c = 28$ GHz, $\Delta f = 120$ kHz, $N = 512$, and $M = 10$, while the oscillator is a free-running oscillator with the 3-dB bandwidth of 200 kHz. The scenario contains two targets with the ranges (30, 150)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

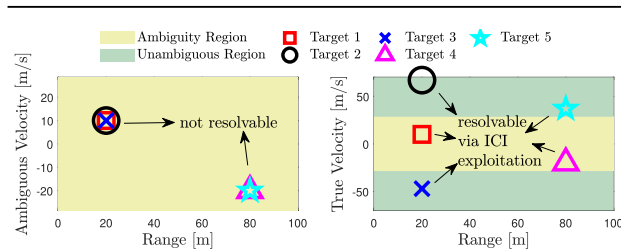
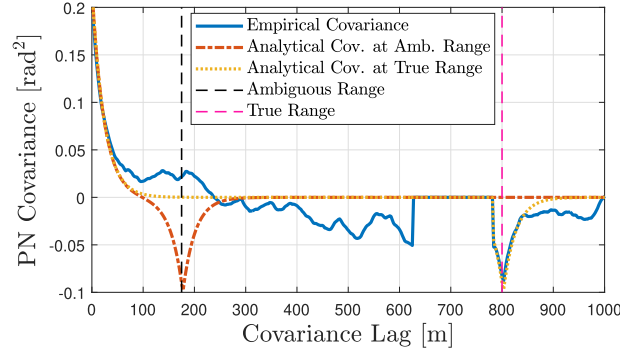
Table 3 Exploitation of Impairments in Joint Monostatic Sensing and Communications

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

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 faraway 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, Algorithm 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 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

**Fig. 26.** Exploitation of ICI in a multitarget 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 $N = 2048$.**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 of 1 MHz and the 3-dB bandwidth of 100 kHz.

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* 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: 1) classical SLAM algorithms that rely on the joint estimation of landmarks and user positions, as described earlier; and 2) algorithms based on relative pose estimation by

comparing measurements (scenes) collected in the last two time instants, $\mathbf{y}_n$ and $\mathbf{y}_{n-1}$. 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 have 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 sidelobes' 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- 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, 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 RX in space, whereas the latter consists in simultaneously transmitting multiple

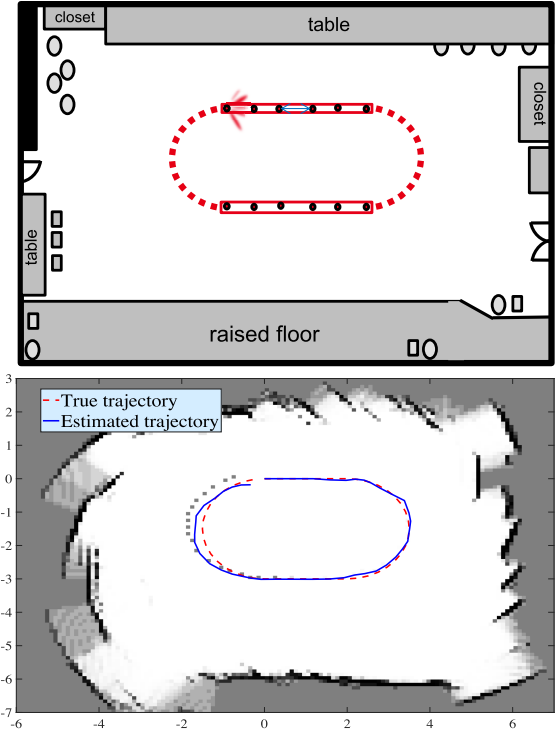


Fig. 28. Radio SLAM using measurements in the THz band. Top: sketch of the measured indoor scenario. Bottom: trajectory and map reconstruction using the MFM algorithm and an occupancy grid mapping method, respectively [225].

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], [230], [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], [233], [234]. Typically, these terms are applied when arrays consist of electrically small and densely packed elements.

B. Near-Field Versus 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], [234], [235], [236]. In fact, traditional wireless systems typically work at distances beyond the Fraunhofer distance, defined as $d_F = 2D^2/\lambda$ (far-field boundary),

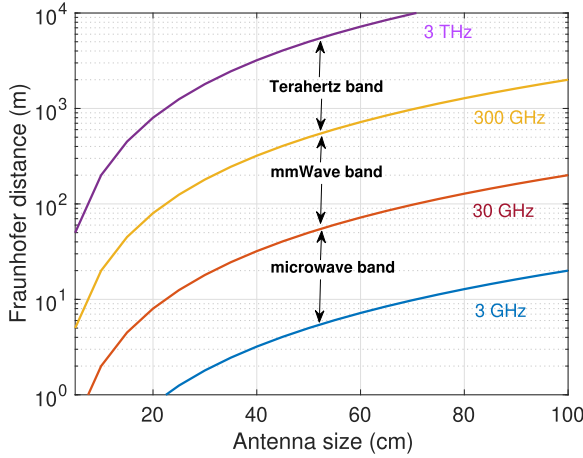


Fig. 29. Fraunhofer region boundaries as a function of the antenna size at different frequency bands. The area above each curve corresponds to the far-field region (Fraunhofer region), whereas the area below each curve corresponds to the near-field region.

where the electromagnetic 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 electromagnetic waveform is more informative than a plane wave, as it will be discussed in Section 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 electromagnetic 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 similar to an optical lens [239]. It not only allows for the control of multiuser

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 nonoverlapping 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 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 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 electromagnetic field approximately fills the aperture of the receiving antenna. In particular, the diffraction solid angle from the spot is $\Omega \approx (\lambda^2/a)$, 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

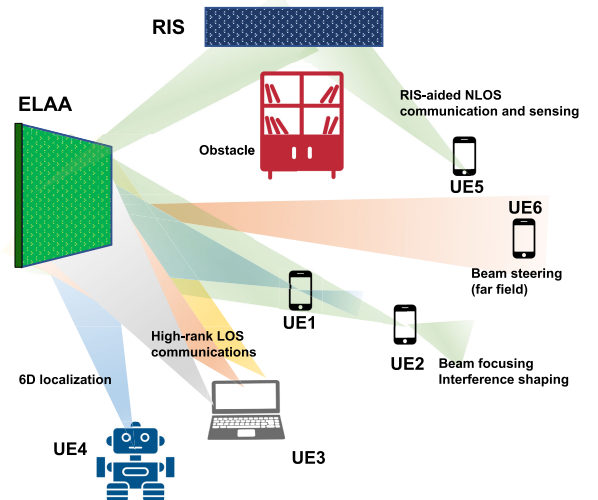


Fig. 30. Near-field communication and sensing. UE1 and UE2 can be discriminated even if they are at the same angle of view thanks to focusing (interference shaping). UE3 can establish a high-rank communication link (spatial multiplexing in LoS); the position and orientation of UE4 can be estimated (single-antenna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

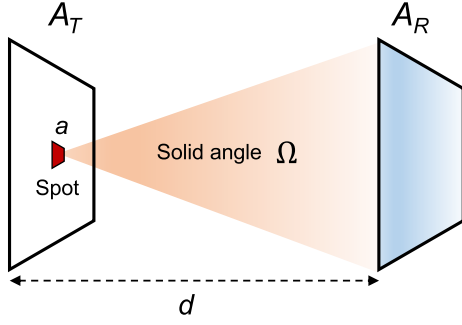


Fig. 31. Diffraction effect at a distance d of a small spot of area a composing the transmitting surface.

in strong LoS conditions. A more accurate theoretical bound on DoF valid also for asymptotically large ELAA modeled as an LIS can be found in [234]. Achieving capacity-approaching MIMO communications in the near field results in intricate array 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 electromagnetic wave [243], [244].

Example 8: Consider a simple 2-D 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

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

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 $(1+x)^{1/2} = 1 + (x/2) - (x^2/8) + o(x^2)$ leading to:

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

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 AoA θ can be estimated.

The case where the source employs an antenna array opens the door to 6-D 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 6-D 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 also possible 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 ISAC is still missing in the literature. In this direction, Wang et al. [245] derived 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. Dehkordi et al. [246] considered a multistatic 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 and 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.

- 1) *Near-field channel modeling:* The extremely large array aperture introduces spatial nonwide 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].

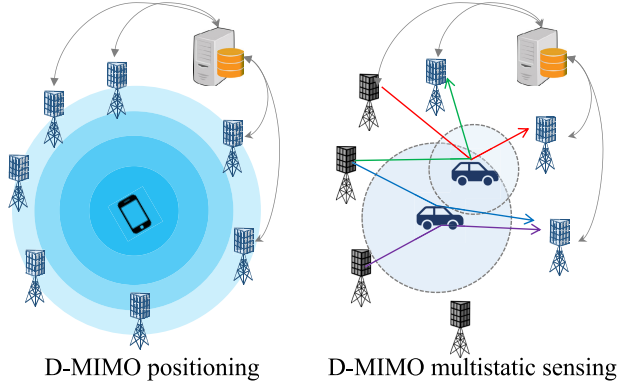


Fig. 32. D-MIMO for positioning (left) and multistatic sensing (right). Phase-coherent processing leads to wavefront curvature, considering all radio units (RUs) as part of a distributed array. Black RUs are TXs, while blue RUs are RXs.

- 2) *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, Cui and Dai [248] 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 position. In principle, only three parameters are sufficient for estimating the channel. Consequently, channel estimation aligns with the localization task, further intertwining S&C.
- 3) *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 Multiperspective 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 interelement spacing (see Fig. 32) [249]. Each antenna (or small array) is called a 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 UL pilot transmission, DL multiuser beamforming, and UL data transmission [252].

Typical D-MIMO assumes the phase synchronization among the RUs, enabling spatial focusing toward 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 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, model (24) changes to [259], [260], [261], [262]

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

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 3-D location $\mathbf{p}$, even when the LoS path between the UE and BS is blocked [261], [262], [263], [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

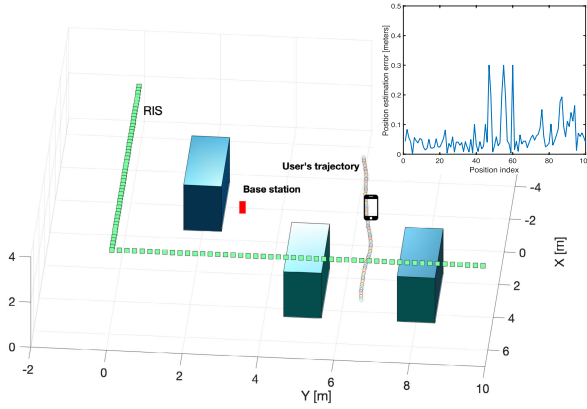


Fig. 33. RIS-assisted localization in a near-field NLoS scenario. The 28-GHz OFDM system with 250-MHz bandwidth. The error evolution along the trajectory is reported [259].

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. In addition, a narrowband version of the algorithm, leveraging the phase profile of the received signal, is proposed in [259]. For sensing, 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], [267], [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 the 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 5 s 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× throughput reduction when reconfiguring the beamforming to use an available NLOS path and a 4× 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 operations 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, 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 following, 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 $\mathcal{W}_L$ can be computed to account

for the estimation error in the angular direction of the user, denoted as $\Delta\phi_{\text{UE}}$. In 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

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

With this type of strategy, it is possible to achieve an overhead reduction of $(1 - |\mathcal{L}|/N_T)$ [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, such as 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 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 has been studied in [32], [280], and [281] via ray-tracing 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 vehicle-to-infrastructure (V2I) link with an active radar mounted at the BS [32]. An alternative design in [281] considers a passive radar RX 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

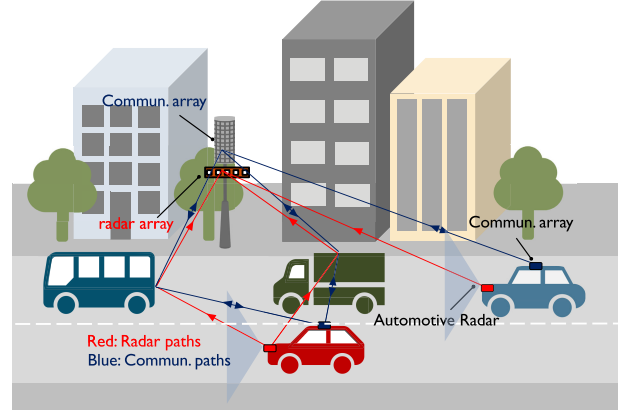


Fig. 34. Illustration of a vehicular communication system with radar-aided beam training. There is a similarity between the radar and communication main channel directions, which can be exploited to reduce beam training overhead in NLOS and multiuser scenarios.

the BS radar RX. Using this information, Ali et al. [281] showed 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 that 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 Fig. 34. The parameters of the radar and communication signals are specified in Table 4. The system is simulated with a number of four users on the road. Additional details of the urban environment simulated by ray tracing can be found in [30] and [281]. The considered performance metric is the sum rate, defined as $R_\Sigma = \sum_u R_u$, with the effective rate per user R_u

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

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

$$s_u = \sum_{k=1}^N \log_2(1 + \text{SINR}_u[k]) \quad (46)$$

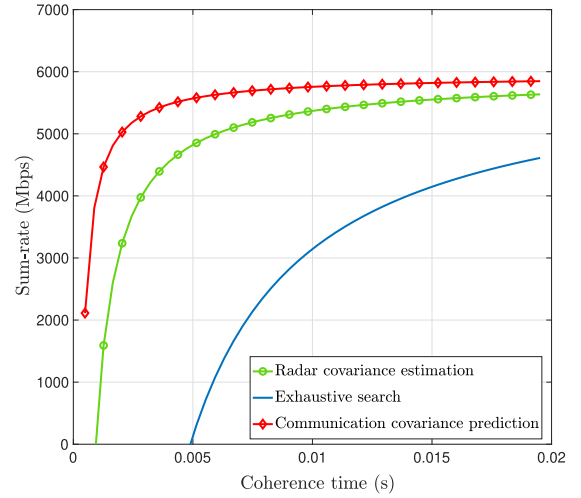
with N the number of subcarriers. The initial estimation of the radar covariance matrix at the radar arrays of the BS is performed using the multiuser separation and covariance

Table 4 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	μ s
Samples per chirp		1024	samples
Height of the vehicle radars		0.75	m

estimation strategies designed in [30] and [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 four beams around the directions suggested by radar. Fig. 35 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, Liu et al. [75] considered the beam tracking stage of a vehicular communication system, exploiting a BS operating as a radar to estimate 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, Yuan et al. [283] proposed 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, AoA, and path loss, to finally leverage the angular information for beam tracking as in [75].

**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).

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 3-D scene reconstruction [284]. 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 [285]. The collected images are used to reconstruct the 3-D 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 3-D environment contains spatial information such as the locations of the scatterers. Tian et al. [284] studied 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 by Alrabeiah et al. [286], where they 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 for 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 [287]. This approach not only provides the beam decisions but also predicts the duration of the beam coherence interval. The simulation results obtained by using 3-D modeling and 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, such as lidar or IMU, have also been leveraged to reduce beam training overhead [29], [288], [289], [290]. 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. The 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. The subsequent works introduced innovations that improve the performance in complex NLOS cases. For example, the approach in [288] 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 [290], 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 [291], [292], [293], [294].

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 path loss in the presence of blockage and specific studies describing the impact of a blocked link. For example, path loss for LOS and NLOS in urban areas can be characterized by different exponents [295] possibly with additional corner losses [296]. In terms of specific blocking objects, vehicles may incur 20 dB of attenuation [297], people 40 dB [298], and hand 20 dB [299]. The impact of the blockage and the duration depends critically on the mobility of the TX, RX, and blockage as well as the corresponding distances, with impact generally increasing with frequency [297]. It seems likely that 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 ISACs.

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 [300]. This work employs a centralized proactive BS selection algorithm, which is shown to be effective with experimental results. The work in [301] utilizes a dual-band BS supporting both sub-6-GHz and mmWave frequencies at the roadside unit of a vehicular communication system. In addition, 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-6-GHz 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 [302], a bimodal ML 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 [303] 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 [304] where an 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 3-D environment that can be used to trace the rays, referred to as ray casting, which are converted to the human blockage predictions. Similarly, the work in [305] proposes to convert the lidar

data to 3-D 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 [294], an experimental setup with a dual-band BS unit is also equipped with a camera and an 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 [306], an 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 [307] uses range-angle maps obtained with the radar at the roadside unit. The range-angle maps are utilized as the input of an 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, ML-based solutions are shown to be effective for converting the data obtained from the sensor(s) to blockage detection and/or prediction.

REFERENCES

- [1] (2016). *Shared Spectrum Access for Radar and Communications (SSPARC)*. [Online]. Available: <https://www.darpa.mil/program/shared-spectrum-access-for-radar-and-communications>
- [2] G. C. Tavik et al., "The advanced multifunction RF concept," *IEEE Trans. Microw. Theory Techn.*, vol. 53, no. 3, pp. 1009–1020, Mar. 2005.
- [3] F. Liu et al., "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, Jun. 2022.
- [4] (2022). *Roadmap to 6G*. [Online]. Available: <https://roadmap.nextgalliance.org/>
- [5] (2021). *European Vision for the 6G Network Ecosystem*. [Online]. Available: <https://5g-ppp.eu/wp-content/uploads/2021/06/WhitePaper-6G-Europe.pdf>
- [6] (2023). *Industry Specification Group (ISG) Integrated Sensing and Communications (ISAC)*. [Online]. Available: <https://www.etsi.org/committee/2295-isac>
- [7] H. Wymeersch et al., "Joint communication and sensing for 6G—A cross-layer perspective," 2024, *arXiv:2402.09120*.
- [8] A. Hassanien, M. G. Amin, E. Aboutanos, 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, Sep. 2019.
- [9] K. V. Mishra, M. R. 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, Sep. 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, Jul. 2020.
- [11] F. Liu et al., "Seventy years of radar and communications: The road from separation to integration," *IEEE Signal Process. Mag.*, vol. 40, no. 5, pp. 106–121, Jul. 2023.
- [12] A. Ali, N. Gonzalez-Prelcic, R. W. Heath Jr., and A. Ghosh, "Leveraging sensing at the infrastructure for mmWave communication," *IEEE Commun. Mag.*, vol. 58, no. 7, pp. 84–89, Jul. 2020.
- [13] H. Holma, H. Viswanathan, and P. Mogensen, "Extreme massive MIMO for macro cell capacity boost in 5G-Advanced and 6G," Nokia, Espoo, Finland, Tech. Rep. CID210786, 2021.
- [14] M. A. Saeidi, H. Tabassum, and M.-S. Alouini, "Multi-band wireless networks: Architectures, challenges, and comparative analysis," *IEEE Commun. Mag.*, vol. 62, no. 1, pp. 80–86, Jul. 2023.
- [15] S. Aboagye et al., "Multi-band wireless communication networks: Fundamentals, challenges, and resource allocation," *IEEE Trans. Commun.*, early access, 2024. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/10438479>
- [16] R. W. Heath Jr., N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.
- [17] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic, and R. W. Heath Jr., "MIMO precoding and combining solutions for millimeter-wave systems," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 122–131, Dec. 2014.
- [18] P. Kumari, A. Mezghani, and R. W. Heath Jr., "A low-resolution ADC proof-of-concept development for a fully-digital millimeter-wave joint communication-radar," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Jan. 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 *Proc. IEEE Int. Conf. Commun. (ICC)*, Oct. 2022, pp. 3160–3165.
- [20] S. Kang et al., "Cellular wireless networks in the upper mid-band," 2023, *arXiv:2309.03038*.
- [21] S. Dwivedi et al., "Positioning in 5G networks," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 38–44, Nov. 2021.
- [22] *Feasibility Study on Integrated Sensing and Communication (Release 19)*, document TR 22.837, 3GPP 2023.
- [23] *Service Requirements for Integrated Sensing and Communication; Stage 1 (Release 19)*, document TR 22.837, 3GPP 2023.
- [24] B. Xiong et al., "Channel modeling for heterogeneous vehicular ISAC system with shared clusters," in *Proc. IEEE 98th Veh. Technol. Conf. (VTC-Fall)*, Oct. 2023, pp. 1–6.
- [25] Z. Zhang et al., "A general channel model for integrated sensing and communication scenarios," *IEEE Commun. Mag.*, vol. 61, no. 5, pp. 68–74, May 2023.
- [26] N. Gonzalez-Prelcic, A. Ali, V. Va, and R. W. Heath Jr., "Millimeter-wave communication

VIII. CONCLUSION

Cellular networks are integrating S&C functionalities. This integration enables new sensing services and also provides data that can be exploited to enhance the 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.

Acknowledgment

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 the 3rd Generation Partnership Project standardization for 6G. ■

- with out-of-band information," *IEEE Commun. Mag.*, vol. 55, no. 12, pp. 140–146, Dec. 2017.
- [27] V. Va, T. Shimizu, G. Bansal, and R. W. Heath Jr., "Online learning for position-aided millimeter wave beam training," *IEEE Access*, vol. 7, pp. 30507–30526, 2019.
- [28] V. Va, J. Choi, T. Shimizu, G. Bansal, and R. W. Heath Jr., "Inverse multipath fingerprinting for millimeter wave V2I beam alignment," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4042–4058, May 2018.
- [29] A. Klautau, N. González-Prelcic, and R. W. Heath Jr., "LiDAR data for deep learning-based mmWave beam-selection," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 909–912, Jun. 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 Trans. Veh. Technol.*, vol. 72, no. 6, pp. 7454–7468, Jun. 2023.
- [31] S. Bi, J. Lyu, Z. Ding, and R. Zhang, "Engineering radio maps for wireless resource management," *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 133–141, Apr. 2019.
- [32] N. González-Prelcic, R. Méndez-Rial, and R. W. Heath Jr., "Radar aided beam alignment in mmWave V2I communications supporting antenna diversity," in *Proc. Inf. Theory Appl. Workshop (ITA)*, Mar. 2016, pp. 1–7.
- [33] M. Mühlhäuser et al., "Street lamps as a platform," *Commun. ACM*, vol. 63, no. 6, pp. 75–83, May 2020, doi: [10.1145/3376900](https://doi.org/10.1145/3376900).
- [34] H. Holma, A. Toskala, and T. Nakamura, *5G Technology: 3GPP New Radio*, 1st ed. Hoboken, NJ, USA: Wiley, 2019.
- [35] M. Renfors, E. Kofidis, and F. Bader, *Orthogonal Waveforms and Filter Banks for Future Communication Systems*. Amsterdam, The Netherlands: Elsevier, Jan. 2017.
- [36] G. Fettweis, M. Krondorf, and S. Bittner, "GFDm-generalized frequency division multiplexing," in *Proc. IEEE 69th Veh. Technol. Conf.*, Apr. 2009, pp. 1–4.
- [37] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, no. 5, pp. 5–14, May 1990.
- [38] J. Yli-Kaakinen et al., "Efficient fast-convolution-based waveform processing for 5G physical layer," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1309–1326, Jun. 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 Trans. Signal Process.*, vol. 68, pp. 1213–1228, 2020.
- [40] T. Levanen, J. Pirsanen, K. Pajukoski, M. Renfors, and M. Valkama, "Transparent Tx and Rx waveform processing for 5G new radio mobile communications," *IEEE Wireless Commun.*, vol. 26, no. 1, pp. 128–136, Feb. 2019.
- [41] L. Gaudio, M. Kobayashi, G. Caire, and G. Colavolpe, "On the effectiveness of OTFS for joint radar parameter estimation and communication," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 5951–5965, Sep. 2020.
- [42] S. D. Liyanaarachchi, T. Riihonen, C. B. Barneto, and M. Valkama, "Optimized waveforms for 5G–6G communication with sensing: Theory, simulations and experiments," *IEEE Trans. Wireless Commun.*, vol. 20, no. 12, pp. 8301–8315, Dec. 2021.
- [43] C. Sturm and W. Wiesbeck, "Waveform design and signal processing aspects for fusion of wireless communications and radar sensing," *Proc. IEEE*, vol. 99, no. 7, pp. 1236–1259, Jul. 2011.
- [44] M. F. Keskin, V. Koivunen, and H. Wymeersch, "Limited feedforward waveform design for OFDM dual-functional radar-communications," *IEEE Trans. Signal Process.*, vol. 69, pp. 2955–2970, Apr. 2021.
- [45] M. Bićić and V. Koivunen, "Multicarrier radar-communications waveform design for RF convergence and coexistence," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 2019, pp. 7780–7784.
- [46] K. Venugopal, A. Alkhatieb, N. G. Prelcic, and R. W. Heath Jr., "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1996–2009, Sep. 2017.
- [47] P. Kumari, J. Choi, N. González-Prelcic, and R. W. Heath Jr., "IEEE 802.11ad-based radar: An approach to joint vehicular communication-radar system," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3012–3027, Apr. 2018.
- [48] NR Physical Channels and Modulation, document TS38.211, 3GPP, Sophia Antipolis, France, 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, Nov. 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)*, Aug. 2022, pp. 1002–1006.
- [51] P. Chen and H. Kobayashi, "Maximum likelihood channel estimation and signal detection for OFDM systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Apr. 2002, pp. 1640–1645.
- [52] R. Carvajal, J. C. Agüero, B. I. Godoy, and G. C. Goodwin, "EM-based maximum-likelihood channel estimation in multicarrier systems with phase distortion," *IEEE Trans. Veh. Technol.*, vol. 62, no. 1, pp. 152–160, Jan. 2013.
- [53] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. I. Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 434–450, Mar. 1999.
- [54] B. H. Fleury, P. Jourdan, and A. Stucki, "High-resolution channel parameter estimation for MIMO applications using the SAGE algorithm," in *Proc. Int. Zurich Seminar Broadband Commun. Access-Transmiss.-Netw.*, 2002, p. 30.
- [55] A. Shahmansoori, G. E. García, 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, Mar. 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 *Proc. IEEE Global Commun. Conf.*, Dec. 2014, pp. 3326–3331.
- [57] J. Rodríguez-Fernández, N. González-Prelcic, K. Venugopal, and R. W. Heath Jr., "Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2946–2960, May 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 Commun. Lett.*, vol. 8, no. 4, pp. 1103–1107, Aug. 2019.
- [59] J. Palacios and N. González-Prelcic, "Separable multidimensional orthogonal matching pursuit and its application to joint localization and communication at mmWave," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2023, pp. 1–6.
- [60] F. Jiang, Y. Ge, M. Zhu, and H. Wymeersch, "High-dimensional channel estimation for simultaneous localization and communications," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Dec. 2021, pp. 1–6.
- [61] F. Wen, H. C. So, and H. Wymeersch, "Tensor decomposition-based beamspace esprits algorithm for multidimensional harmonic retrieval," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 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 J. Sel. Topics Signal Process.*, vol. 15, no. 3, pp. 816–831, Apr. 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," in *Proc. Int. Conf. Commun. (ICC)*, Jun. 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 *Proc. 56th Asilomar Conf. Signals, Syst., Comput.*, Oct. 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 Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7121–7134, Sep. 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 Trans. Veh. Technol.*, vol. 71, no. 5, pp. 5558–5563, May 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 J. Sel. Topics Signal Process.*, vol. 16, no. 5, pp. 1141–1156, Aug. 2022.
- [68] S. Rivetti et al., "Spatial signal design for positioning via end-to-end learning," *IEEE Wireless Commun. Lett.*, vol. 12, no. 3, pp. 525–529, Mar. 2023.
- [69] S. H. Talisa, K. W. O'Haver, T. M. Comeriate, M. D. Sharp, and O. F. Somerlock, "Benefits of digital phased array radars," *Proc. IEEE*, vol. 104, no. 3, pp. 530–543, Mar. 2016.
- [70] I. Aykin and M. Krunic, "Efficient beam sweeping algorithms and initial access protocols for millimeter-wave networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2504–2514, Apr. 2020.
- [71] U. Nickel, "Overview of generalized monopulse estimation," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 21, no. 6, pp. 27–56, Jun. 2006.
- [72] M. Z. Win, W. Dai, Y. Shen, G. Christikos, and H. V. Poor, "Network operation strategies for efficient localization and navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1224–1254, Jul. 2018.
- [73] Y. Shen and M. Z. Win, "Fundamental limits of wideband localization—Part I: A general framework," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4956–4980, Oct. 2010.
- [74] R. Miller and C. Chang, "A modified Cramér–Rao bound and its applications (corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 398–400, May 1978.
- [75] F. Liu, W. Yuan, C. Masouros, and J. Yuan, "Radar-assisted predictive beamforming for vehicular links: Communication served by sensing," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7704–7719, Nov. 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 Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2755–2770, Apr. 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 Trans. Wireless Commun.*, vol. 23, no. 2, pp. 1158–1169, Feb. 2024.
- [78] W. Dai, Y. Shen, and M. Z. Win, "A computational geometry framework for efficient network localization," *IEEE Trans. Inf. Theory*, vol. 64, no. 2, pp. 1317–1339, Feb. 2018.
- [79] W. W. Li, Y. Shen, Y. J. Zhang, and M. Z. Win, "Robust power allocation for energy-efficient location-aware networks," *IEEE/ACM Trans. Netw.*, vol. 21, no. 6, pp. 1918–1930, Dec. 2013.
- [80] H. Zhao, N. Zhang, and Y. Shen, "Beamspace direct localization for large-scale antenna array systems," *IEEE Trans. Signal Process.*, 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 J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2085–2102, Jul. 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 Trans. Commun.*, vol. 70, no. 12, pp. 8196–8214, Dec. 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 Trans. Signal Process.*, vol. 67, no. 14, pp. 3818–3831, Jul. 2019.
 - [84] J. Mu, W. Ouyang, Z. Jing, B. Li, and F. Zhang, “Energy-efficient interference cancellation in integrated sensing and communication scenarios,” *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 1, pp. 370–378, Mar. 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 Trans. Inf. Theory*, vol. 69, no. 9, pp. 5723–5751, Sep. 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 Trans. Signal Process.*, vol. 64, no. 2, pp. 464–474, Jan. 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 Trans. Commun.*, vol. 71, no. 8, pp. 4957–4971, Aug. 2023.
 - [88] H. Yang et al., “Queue-aware dynamic resource allocation for the joint communication-radar system,” *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 754–767, Jan. 2021.
 - [89] L. Pucci, E. Paolini, and A. Giorgetti, “System-level analysis of joint sensing and communication based on 5G new radio,” *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2043–2055, Jul. 2022.
 - [90] A. Fascista et al., “Uplink joint positioning and synchronization in cell-free deployments with radio stripes,” in *Proc. IEEE Int. Conf. Commun. Workshops (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, Oct. 2019.
 - [92] A. Haimovich, R. Blum, and L. Cimini, “MIMO radar with widely separated antennas,” *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 116–129, Dec. 2007.
 - [93] H. Wymeersch and G. Seco-Granados, “Radio localization and sensing—Part I: Fundamentals,” *IEEE Commun. Lett.*, vol. 26, no. 12, pp. 2816–2820, Dec. 2022.
 - [94] S. E. Trellakis et al., “Localization as a key enabler of 6G wireless systems: A comprehensive survey and an outlook,” *IEEE Open J. Commun. Soc.*, vol. 4, pp. 2733–2801, 2023.
 - [95] İ. Güvenc, C.-C. Chong, F. Watanabe, and H. Inamura, “NLOS identification and weighted least-squares localization for UWB systems using multipath channel statistics,” *EURASIP J. Adv. Signal Process.*, vol. 2008, no. 1, pp. 1–14, Dec. 2007.
 - [96] H. Wymeersch et al., “5G mmWave downlink vehicular positioning,” in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 206–212.
 - [97] B. Camajori Tedeschini et al., “A feasibility study of 5G positioning with current cellular network deployment,” *Sci. Rep.*, vol. 13, no. 1, p. 15281, Sep. 2023.
 - [98] M. Koivisto et al., “Joint device positioning and clock synchronization in 5G ultra-dense networks,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2866–2881, May 2017.
 - [99] M. A. Nazari, G. Seco-Granados, P. Johannisson, and H. Wymeersch, “mmWave 6D radio localization with a snapshot observation from a single BS,” *IEEE Trans. Veh. Technol.*, vol. 72, no. 7, pp. 8914–8928, Jul. 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 *Proc. IEEE 12th Sensor Array Multichannel Signal Process. Workshop (SAM)*, Jun. 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 Process. Mag.*, vol. 22, no. 4, pp. 41–53, Jul. 2005.
 - [102] G. Wang, A. M.-C. So, and Y. Li, “Robust convex approximation methods for TDOA-based localization under NLOS conditions,” *IEEE Trans. Signal Process.*, vol. 64, no. 13, pp. 3281–3296, Jul. 2016.
 - [103] Y. Lu et al., “Bayesian filtering for joint multi-user positioning, synchronization and anchor state calibration,” *IEEE Trans. Veh. Technol.*, vol. 72, no. 8, pp. 10949–10964, Aug. 2023.
 - [104] S. Burgess, Y. Kuang, and K. Åström, “TOA sensor network self-calibration for receiver and transmitter spaces with difference in dimension,” *Signal Process.*, vol. 107, pp. 33–42, Feb. 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 Trans. Wireless Commun.*, vol. 18, no. 9, pp. 4254–4267, Sep. 2019.
 - [106] Y. T. Chan and K. C. Ho, “A simple and efficient estimator for hyperbolic location,” *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 1905–1915, Aug. 1994.
 - [107] H. Dun, C. C. J. M. Tiberius, and G. J. M. Janssen, “Positioning in a multipath channel using OFDM signals with carrier phase tracking,” *IEEE Access*, vol. 8, pp. 13011–13028, 2020.
 - [108] J. Talvitie, M. Säily, and M. Valkama, “Orientation and location tracking of XR devices: 5G carrier phase-based methods,” *IEEE J. Sel. Topics Signal Process.*, vol. 17, no. 5, pp. 919–934, Sep. 2023.
 - [109] E. Leitinger, F. Meyer, F. Hlawatsch, K. Witrisal, F. Tufvesson, and M. Z. Win, “A belief propagation algorithm for multipath-based SLAM,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 12, pp. 5613–5629, Dec. 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 Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6104–6117, Sep. 2016.
 - [111] H. Chen et al., “RISs and sidelink communications in smart cities: The key to seamless localization and sensing,” *IEEE Commun. Mag.*, vol. 61, no. 8, pp. 140–146, Aug. 2023.
 - [112] A. Elzanaty, A. Guerra, F. Guidi, and M.-S. Alouini, “Reconfigurable intelligent surfaces for localization: Position and orientation error bounds,” *IEEE Trans. Signal Process.*, 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 J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2011–2025, Jul. 2022.
 - [114] Q. Liu et al., “A highly accurate positioning solution for C-V2X systems,” *Sensors*, vol. 21, no. 4, p. 1175, Feb. 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 Netw.*, vol. 33, no. 3, pp. 96–101, May/Jun. 2019.
 - [116] H. Kim, K. Granström, L. Svensson, S. Kim, and H. Wymeersch, “PMBM-based SLAM filters in 5G mmWave vehicular networks,” *IEEE Trans. Veh. Technol.*, vol. 71, no. 8, pp. 8646–8661, Aug. 2022.
 - [117] R. Keating, M. Säily, J. Hultkonen, and J. Karjalainen, “Overview of positioning in 5G new radio,” in *Proc. 16th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2019, pp. 320–324.
 - [118] N. Garcia, H. Wymeersch, E. G. Larsson, A. M. Haimovich, and M. Coulon, “Direct localization for massive MIMO,” *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 2475–2487, May 2017.
 - [119] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, “Massive MIMO in real propagation environments: Do all antennas contribute equally?” *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 3917–3928, Nov. 2015.
 - [120] *Study on Channel Model for Frequencies From 0.5 to 100 GHz, Ver17.0.0*, document TR 38.901, 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 Process. Mag.*, vol. 40, no. 4, pp. 92–106, Jun. 2023.
 - [122] A. Shastri et al., “A review of millimeter wave device-based localization and device-free sensing technologies and applications,” *IEEE Commun. Surveys Tuts.*, vol. 24, no. 3, pp. 1708–1749, 3rd Quart., 2022.
 - [123] H. Chen, H. Sarieddeen, T. Ballal, H. Wymeersch, M. Alouini, and T. Y. Al-Naffouri, “A tutorial on terahertz-band localization for 6G communication systems,” *IEEE Commun. Surveys Tuts.*, vol. 24, no. 3, pp. 1780–1815, 3rd Quart., 2022.
 - [124] J. Zhang, X. Ge, Q. Li, M. Guizani, and Y. Zhang, “5G millimeter-wave antenna array: Design and challenges,” *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 106–112, Apr. 2017.
 - [125] H. Chen et al., “Modeling and analysis of OFDM-based 5G/6G localization under hardware impairments,” 2023, *arXiv:2301.01042*.
 - [126] D. Tubail, B. Cenklioglu, A. E. Canbilen, 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 Trans. Veh. Technol.*, vol. 72, no. 3, pp. 4063–4067, Mar. 2023.
 - [127] M. Shafi et al., “Microwave vs. millimeter-wave propagation channels: Key differences and impact on 5G cellular systems,” *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 14–20, Dec. 2018.
 - [128] T. S. Rappaport et al., “Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond,” *IEEE Access*, vol. 7, pp. 78729–78757, 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] *Technical Specification Group Radio Access Network; Study on Evaluation Methodology of New Vehicle-to-Everything (V2X) Use Cases for LTE and NR*, document TR 37.885, 3GPP, 2019.
 - [131] Y. Ge, M. Stark, M. Furkan Keskin, F. Hofmann, T. Hansen, and H. Wymeersch, “V2X sidelink positioning in FR1: Scenarios, algorithms, and performance evaluation,” 2023, *arXiv:2310.13753*.
 - [132] S. He and S.-H. G. Chan, “Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons,” *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 466–490, 1st Quart., 2016.
 - [133] R. Klus, J. Talvitie, J. Vinogradova, J. Torsner, and M. Valkama, “Machine learning based NLOS radio positioning in beamforming networks,” in *Proc. IEEE 23rd Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 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 *Proc. IEEE Global Commun. Conf.*, Dec. 2020, pp. 1–6.
 - [135] P. Bahl and V. N. Padmanabhan, “RADAR: An in-building RF-based user location and tracking system,” in *Proc. Conf. Comput. Commun. 19th Annu. Joint Conf. IEEE Comput. Commun. Soc.*, vol. 2, Jul. 2000, pp. 775–784.

- [136] Q. D. Vo and P. De, "A survey of fingerprint-based outdoor localization," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 491–506, 1st Quart., 2016.
- [137] J. Hu, D. Liu, Z. Yan, and H. Liu, "Experimental analysis on weight K -nearest neighbor indoor fingerprint positioning," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 891–897, Feb. 2019.
- [138] M. Youssef and A. Agrawala, "The Horus WLAN location determination system," in *Proc. 3rd Int. Conf. Mobile Syst. Appl. Services*. New York, NY, USA: Association for Computing Machinery, 2005, pp. 205–218, doi: [10.1145/1067170.1067193](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 J. Sel. Areas Commun.*, vol. 33, no. 11, pp. 2429–2438, Nov. 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 Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6134–6145, Jul. 2018.
- [141] C. Wu et al., "Learning to localize: A 3D CNN approach to user positioning in massive MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4556–4570, Jul. 2021.
- [142] J. Xiao, K. Wu, Y. Yi, and L. M. Ni, "FIFS: Fine-grained indoor fingerprinting system," in *Proc. 21st Int. Conf. Comput. Commun. Netw. (ICCCN)*, Jul. 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 Trans. Veh. Technol.*, vol. 66, no. 1, pp. 763–776, Jan. 2017.
- [144] B. Zhang, H. Sifaou, and G. Y. Li, "CSI-fingerprinting indoor localization via attention-augmented residual convolutional neural network," *IEEE Trans. Wireless Commun.*, vol. 22, no. 8, pp. 5583–5597, Aug. 2023.
- [145] P. Ferrand, A. Decurninge, and M. Guillaud, "DNN-based localization from channel estimates: Feature design and experimental results," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 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," 2023, [arXiv:2307.00167](https://arxiv.org/abs/2307.00167).
- [147] Y. Chen, N. González-Prelcic, T. Shimizu, H. Lu, and C. Mahabali, "Sparse recovery with attention: A hybrid data/model driven solution for high accuracy position and channel tracking at mmWave," in *Proc. IEEE 24th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Sep. 2023, pp. 491–495.
- [148] A. Salihu, S. Schwarz, and M. Rupp, "Attention aided CSI wireless localization," in *Proc. IEEE 23rd Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 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. 18066–18074, 2017.
- [150] M. M. Butt, A. Pantelidou, and I. Z. Kovács, "ML-assisted CSI positioning: Performance analysis and 5G architecture enhancements," *IEEE Open J. Veh. Technol.*, 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 Process. Lett.*, vol. 51, no. 1, pp. 487–514, Feb. 2020.
- [152] G. Kia, L. Ruotsalainen, and J. Talvitie, "A CNN approach for 5G mm wave positioning using beamformed CSI measurements," in *Proc. Int. Conf. Localization GNSS (ICL-GNSS)*, Jun. 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 *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 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 Process. Lett.*, vol. 29, pp. 269–273, 2022.
- [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 *Proc. IEEE 34th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 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 Things J.*, vol. 8, no. 5, pp. 3422–3437, Mar. 2021.
- [157] S. Shi, S. Sigg, L. Chen, and Y. Ji, "Accurate location tracking from CSI-based passive device-free probabilistic fingerprinting," *IEEE Trans. Veh. Technol.*, vol. 67, no. 6, pp. 5217–5230, Jun. 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, p. 5776, Aug. 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 Automot. Assoc., Munich, Germany, Tech. Rep. A-200118, 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 *Proc. 3rd Int. Symp. Joint Commun. Sens.*, 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 J. Sel. Areas Commun.*, vol. 42, no. 1, pp. 120–133, Jan. 2024.
- [162] Y. Shen, H. Wymeersch, and M. Z. Win, "Fundamental limits of wideband localization—Part II: Cooperative networks," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4981–5000, Oct. 2010.
- [163] M. Z. Win, Y. Shen, and W. Dai, "A theoretical foundation of network localization and navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1136–1165, Jul. 2018.
- [164] S. Mazuelas, Y. Shen, and M. Z. Win, "Spatiotemporal information coupling in network navigation," *IEEE Trans. Inf. Theory*, vol. 64, no. 12, pp. 7759–7779, Dec. 2018.
- [165] X. Shen, Y. Liu, and Y. Shen, "On the spatial information coupling in relative localization networks," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2021, pp. 1–6.
- [166] Y. Xiong, N. Wu, Y. Shen, and M. Z. Win, "Cooperative localization in massive networks," *IEEE Trans. Inf. Theory*, vol. 68, no. 2, pp. 1237–1258, Feb. 2022.
- [167] W. Dai, Y. Shen, and M. Z. Win, "Distributed power allocation for cooperative wireless network localization," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 1, pp. 28–40, Jan. 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 Trans. Signal Process.*, vol. 64, no. 24, pp. 6517–6532, Dec. 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. 116753–116773, 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 J. Sel. Topics Signal Process.*, vol. 16, no. 5, pp. 1125–1140, Aug. 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 Process.*, vol. 17, no. 4, Apr. 2023, Art. no. e12209.
- [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 J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2446–2460, Aug. 2023.
- [173] E. Björnson, H. Wymeersch, B. Matthieson, P. Popovski, L. Sanguinetti, and E. de Carvalho, "Reconfigurable intelligent surfaces: A signal processing perspective with wireless applications," *IEEE Signal Process. Mag.*, vol. 39, no. 2, pp. 135–158, Mar. 2022.
- [174] M. Mizmizi, D. Tagliaferri, D. Badini, and U. Spagnolini, "Target-to-user association in ISAC systems with vehicle-lodged RIS," *IEEE Wireless Commun. Lett.*, vol. 12, no. 9, pp. 1558–1562, Sep. 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. Adv. Wireless Commun. (SPAWC)*, Jul. 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 Sens.*, vol. 13, no. 18, p. 3642, Sep. 2021.
- [177] Study on Expanded and Improved NR Positioning (Release 18), document TR 38.859 V18.0.0, 3GPP, 2022. [Online]. Available: <http://www.3gpp.org>
- [178] Expanded and Improved NR Positioning Work Item Description (Release 18), document RP-223549, 3GPP, 2022. [Online]. Available: <http://www.3gpp.org>
- [179] P. J. G. Teunissen, "The least-squares ambiguity decorrelation adjustment: A method for fast GPS integer ambiguity estimation," *J. Geodesy*, vol. 70, nos. 1–2, pp. 65–82, Nov. 1995.
- [180] D. B. Cox 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, Sep. 2000.
- [181] L. Chen, X. Zhou, F. Chen, L.-L. Yang, and R. Chen, "Carrier phase ranging for indoor positioning with 5G NR signals," *IEEE Internet Things J.*, vol. 9, no. 13, pp. 10908–10919, Jul. 2022.
- [182] S. Fan, W. Ni, H. Tian, Z. Huang, and R. Zeng, "Carrier phase-based synchronization and high-accuracy positioning in 5G new radio cellular networks," *IEEE Trans. Commun.*, vol. 70, no. 1, pp. 564–577, Jan. 2022.
- [183] B.-N. Vo et al., "Multitarget tracking," in *Wiley Encyclopedia of Electrical and Electronics Engineering*. New York, NY, USA: Wiley, 2015.
- [184] K. Granström, M. Baum, and S. Reuter, "Extended object tracking: Introduction, overview and applications," 2016, [arXiv:1604.00970](https://arxiv.org/abs/1604.00970).
- [185] J. L. Williams and R. A. Lau, "Data association by loopy belief propagation," in *Proc. 13th Int. Conf. Inf. Fusion*, Jul. 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 Trans. Signal Process.*, 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 Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3782–3795, Jun. 2020.
- [188] O. Kaltioikallio et al., "Towards real-time radio-SLAM via optimal importance sampling," in *Proc. IEEE 23rd Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2022, pp. 1–5.
- [189] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics* (Intelligent Robotics and Autonomous Agents). Cambridge, MA, USA: 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 *Proc. AAAI Nat. Conf. Artificial Intell.*, Jul. 2002, pp. 593–598.
- [191] C. Cadena et al., "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, Jan. 2021.
- [193] E. Rastorgueva-Foi et al., “Millimeter-wave radio SLAM: End-to-end processing methods and experimental validation,” 2023, *arXiv:2312.13741*.
- [194] M. A. Richards, *Fundamentals of Radar Signal Processing*. New York, NY, USA: McGraw-Hill, 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, Feb. 2021.
- [196] C. Baquero Barneto et al., “Full-duplex OFDM radar with LTE and 5G NR waveforms: Challenges, solutions, and measurements,” *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 10, pp. 4042–4054, Oct. 2019.
- [197] K. E. Kolodziej, B. T. Perry, and J. S. Herd, “In-band full-duplex technology: Techniques and systems survey,” *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 7, pp. 3025–3041, Jul. 2019.
- [198] S. Khaledian, F. Farzami, B. Smida, and D. Erricolo, “Inherent self-interference cancellation for in-band full-duplex single-antenna systems,” *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 6, pp. 2842–2850, Jun. 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 Trans. Microw. Theory Techn.*, vol. 65, no. 12, pp. 5263–5273, Dec. 2017.
- [200] P. Pascual Campo, L. Anttila, D. Korpi, and M. Valkama, “Cascaded spline-based models for complex nonlinear systems: Methods and applications,” *IEEE Trans. Signal Process.*, vol. 69, pp. 370–384, 2021.
- [201] A. Kiayani et al., “Adaptive nonlinear RF cancellation for improved isolation in simultaneous transmit–receive systems,” *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 5, pp. 2299–2312, May 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, Dec. 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.*, May 2022, pp. 4673–4678.
- [204] S. D. Liyanaarachchi, C. B. Barneto, T. Riihonen, M. Heino, and M. Valkama, “Joint multi-user communication and MIMO radar through full-duplex hybrid beamforming,” in *Proc. 1st IEEE Int. Online Symp. Joint Commun. Sens.*, Feb. 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 *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.
- [206] P. P. Campo, D. Korpi, L. Anttila, and M. Valkama, “Nonlinear digital cancellation in full-duplex devices using spline-based Hammerstein model,” in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2018, pp. 1–7.
- [207] D. Korpi et al., “Advanced self-interference cancellation and multi-antenna techniques for full-duplex radios,” in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Nov. 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,” 2023, *arXiv:2311.14942*.
- [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 9th Int. Workshop Comput. Adv. Multi-Sensor Adapt. Process. (CAMSAP)*, Dec. 2023, pp. 1–5.
- [210] M. Braun, “OFDM radar algorithms in mobile communication networks,” Ph.D. dissertation, Dept. Inst. Commun. Eng., Karlsruhe Inst. Technol., Karlsruhe, Germany, 2014.
- [211] S. Mercier, S. Bidon, D. Roque, and C. Enderli, “Comparison of correlation-based OFDM radar receivers,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 6, pp. 4796–4813, Dec. 2020.
- [212] Z. Wei et al., “Integrated sensing and communication signals toward 5G-A and 6G: A survey,” *IEEE Internet Things J.*, vol. 10, no. 13, pp. 11068–11092, Jul. 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 Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11659–11672, Oct. 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 Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13112–13123, Nov. 2020.
- [215] G. Hakobyan and B. Yang, “A novel intercarrier-interference free signal processing scheme for OFDM radar,” *IEEE Trans. Veh. Technol.*, vol. 67, no. 6, pp. 5158–5167, Jun. 2018.
- [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 Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 9, pp. 1869–1884, Sep. 2006.
- [218] M. C. Budge and M. P. Burt, “Range correlation effects in radars,” in *Proc. Rec. IEEE Nat. Radar Conf.*, Apr. 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 Process. Mag.*, vol. 36, no. 5, pp. 60–70, Sep. 2019.
- [220] C. Aydogdu et al., “Radar interference mitigation for automated driving: Exploring proactive strategies,” *IEEE Signal Process. Mag.*, vol. 37, no. 4, pp. 72–84, Jul. 2020.
- [221] A. R. Chiriyath, B. Paul, and D. W. Bliss, “Joint radar-communications information bounds with clutter: The phase noise menace,” in *Proc. IEEE Radar Conf. (RadarConf)*, May 2016, pp. 1–6.
- [222] D. Dardari, P. Clossas, and P. M. Djurić, “Indoor tracking: Theory, methods, and technologies,” *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1263–1278, Apr. 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, Apr. 2020.
- [224] F. Guidi, A. Guerra, and D. Dardari, “Personal mobile radars with millimeter-wave massive arrays for indoor mapping,” *IEEE Trans. Mobile Comput.*, vol. 15, no. 6, pp. 1471–1484, Jun. 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 J. Sel. Topics Signal Process.*, vol. 17, no. 4, pp. 834–849, Jul. 2023.
- [226] W. Hess, D. Kohler, H. Rapp, and D. Andor, “Real-time loop closure in 2D LiDAR SLAM,” in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Stockholm, Sweden, May 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*. Berlin, Germany: Springer, 2010, pp. 151–161.
- [228] H. Sarieddeen, 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, May 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 Process.*, vol. 94, pp. 3–20, Jan. 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1051200419300776>
- [230] Z. Wang et al., “A tutorial on extremely large-scale MIMO for 6G: Fundamentals, signal processing, and applications,” *IEEE Commun. Surveys Tuts.*, early access, Jan. 2, 2024, doi: [10.1109/COMST.2023.3349276](https://doi.org/10.1109/COMST.2023.3349276).
- [231] E. Björnson, Y. C. Eldar, E. G. Larsson, A. Lozano, and H. V. Poor, “Twenty-five years of signal processing advances for multi-antenna communications: From theory to mainstream technology,” *IEEE Signal Process. Mag.*, vol. 40, no. 4, pp. 107–117, Jun. 2023.
- [232] C. Huang et al., “Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends,” *IEEE Wireless Commun.*, vol. 27, no. 5, pp. 118–125, Oct. 2020.
- [233] S. Hu, F. Rusek, and O. Edfors, “Beyond massive MIMO: The potential of positioning with large intelligent surfaces,” *IEEE Trans. Signal Process.*, vol. 66, no. 7, pp. 1761–1774, Apr. 2018.
- [234] D. Dardari, “Communicating with large intelligent surfaces: Fundamental limits and models,” *IEEE J. Sel. Areas Commun.*, 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 Commun.*, vol. 28, no. 2, pp. 106–113, Apr. 2021.
- [236] A. Elzanaty, A. Guerra, F. Guidi, D. Dardari, and M.-S. Alouini, “Towards 6G holographic localization: Enabling technologies and perspectives,” *IEEE Internet Things Mag.*, vol. 6, no. 3, pp. 138–143, Sep. 2023.
- [237] K. T. Selvan and R. Janaswamy, “Fraunhofer and Fresnel distances: Unified derivation for aperture antennas,” *IEEE Antennas Propag. Mag.*, vol. 59, no. 4, pp. 12–15, Aug. 2017.
- [238] D. Dardari and N. Decarli, “Holographic communication using intelligent surfaces,” *IEEE Commun. Mag.*, vol. 59, no. 6, pp. 35–41, Jun. 2021.
- [239] P. Nepa and A. Buffi, “Near-field-focused microwave antennas: Near-field shaping and implementation,” *IEEE Antennas Propag. Mag.*, vol. 59, no. 3, pp. 42–53, Jun. 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 Commun. Mag.*, vol. 61, no. 4, pp. 72–77, 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, 2000.
- [242] N. Decarli and D. Dardari, “Communication modes with large intelligent surfaces in the near field,” *IEEE Access*, vol. 9, pp. 165648–165666, 2021.
- [243] A. Guerra, F. Guidi, and D. Dardari, “Single-anchor localization and orientation performance limits using massive arrays: MIMO vs. beamforming,” *IEEE Trans. Wireless Commun.*, 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 Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3571–3586, Jun. 2021.
- [245] Z. Wang, X. Mu, and Y. Liu, “Near-field integrated sensing and communications,” *IEEE Commun. Lett.*, vol. 27, no. 8, pp. 2048–2052, Aug. 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,” 2023, *arXiv:2309.14778*.

- [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 Trans. Antennas Propag.*, vol. 71, no. 1, pp. 921–933, Jan. 2023.
- [248] M. Cui and L. Dai, "Channel estimation for extremely large-scale MIMO: Far-field or near-field?" *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2663–2677, Apr. 2022.
- [249] Ö. T. Demir, E. Björnson, and L. Sanguinetti, "Foundations of user-centric cell-free massive MIMO," *Found. Trends Signal Process.*, vol. 14, nos. 3–4, pp. 162–472, 2021.
- [250] D. Löschenbrand, M. Hofer, L. Bernadó, S. Zelenbaba, and T. Zemen, "Towards cell-free massive MIMO: A measurement-based analysis," *IEEE Access*, vol. 10, pp. 89232–89247, 2022.
- [251] Z. H. Shaik, E. Björnson, and E. G. Larsson, "MMSE-optimal sequential processing for cell-free massive MIMO with radio stripes," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7775–7789, Nov. 2021.
- [252] S. Elhoushy, M. Ibrahim, and W. Hamouda, "Cell-free massive MIMO: A survey," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 1, pp. 492–523, 1st Quart., 2022.
- [253] E. G. Larsson and J. Vieira, "Phase calibration of distributed antenna arrays," *IEEE Commun. Lett.*, vol. 27, no. 6, pp. 1619–1623, Jun. 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 Trans. Wireless Commun.*, vol. 21, no. 8, pp. 6256–6270, Aug. 2022.
- [255] H. Chen et al., "6G localization and sensing in the near field: Features, opportunities, and challenges," 2023, *arXiv:2308.15799*.
- [256] M. E. Leinonen. (2021). *Initial Radio Models and Analysis Towards Ultra-high Data Rate Links in 6G*. Hexa-X project Deliverable D2.2. [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," 2023, *arXiv:2309.06911*.
- [258] M. Manzoni et al., "Wavefield networked sensing: Principles, algorithms and applications," 2023, *arXiv:2305.10333*.
- [259] D. Dardari, N. Decarli, A. Guerra, and F. Guidi, "LOS/NLOS near-field localization with a large reconfigurable intelligent surface," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4282–4294, Jun. 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 *Proc. IEEE Int. Conf. Commun.*, Jun. 2021, pp. 1–6.
- [261] O. Rinchi, A. Elzanaty, and M.-S. Alouini, "Compressive near-field localization for multipath RIS-aided environments," *IEEE Commun. Lett.*, vol. 26, no. 6, pp. 1268–1272, Jun. 2022.
- [262] C. Ozturk, M. F. Keskin, H. Wymeersch, and S. Gezici, "RIS-aided near-field localization under phase-dependent amplitude variations," *IEEE Trans. Wireless Commun.*, vol. 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 *Proc. IEEE 22nd Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Sep. 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 *Proc. IEEE 95th Veh. Technol. Conf. (VTC-Spring)*, Aug. 2022, pp. 1–6.
- [265] C. Ozturk, M. Furkan Keskin, V. Sciancalepore, H. Wymeersch, and S. Gezici, "RIS-aided localization under pixel failures," 2023, *arXiv:2302.04436*.
- [266] S. Buzzi, E. Grossi, M. Lops, and L. Venturino, "Foundations of MIMO radar detection aided by reconfigurable intelligent surfaces," *IEEE Trans. Signal Process.*, vol. 70, pp. 1749–1763, 2022.
- [267] E. Grossi, H. Taremdadeh, and L. Venturino, "Radar target detection and localization aided by an active reconfigurable intelligent surface," *IEEE Signal Process. Lett.*, vol. 30, pp. 903–907, 2023.
- [268] S. Buzzi, E. Grossi, M. Lops, and L. Venturino, "Radar target detection aided by reconfigurable intelligent surfaces," *IEEE Signal Process. Lett.*, vol. 28, pp. 1315–1319, 2021.
- [269] G. Torcolacci et al., "Holographic imaging with XL-MIMO and RIS: Illumination and reflection design," 2023, *arXiv: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 Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 173–196, 1st Quart., 2019.
- [271] M. Polese, J. M. Jornet, T. Melodia, and M. Zorzi, "Toward end-to-end, full-stack 6G terahertz networks," *IEEE Commun. Mag.*, vol. 58, no. 11, pp. 48–54, Nov. 2020.
- [272] A. Narayanan, E. Ramadan, J. Carpenter, Q. Liu, F. Qian, and Z. L. Zhang, "A first look at commercial 5G performance on smartphones," in *Proc. ACM Web Conf. New York, NY, USA: Association for Computing Machinery*, 2020, pp. 894–905, doi: [10.1145/3366423.3380169](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 Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013.
- [274] J. Wang et al., "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," in *Proc. IEEE Global Telecommun. Conf.*, Nov. 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 *Proc. IEEE 17th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2016, pp. 1–5.
- [276] A. Ali, N. González-Prelcic, and A. Ghosh, "Millimeter wave V2I beam-training using base-station mounted radar," in *Proc. IEEE Radar Conf. (RadarConf)*, Apr. 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. 55513–55525, 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 Trans. Veh. Technol.*, vol. 68, no. 11, pp. 10858–10871, Nov. 2019.
- [279] S. Park, A. Ali, N. González-Prelcic, and R. W. Heath Jr., "Spatial channel covariance estimation for hybrid architectures based on tensor decompositions," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1084–1097, Feb. 2020.
- [280] A. Graff, A. Ali, and N. González-Prelcic, "Measuring radar and communication congruence at millimeter wave frequencies," in *Proc. 53rd Asilomar Conf. Signals, Syst., Comput.*, Nov. 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, no. 12, pp. 14903–14917, Dec. 2020.
- [282] Y. Chen, A. Graff, N. González-Prelcic, and T. Shimizu, "Radar aided mmWave vehicle-to-infrastructure link configuration using deep learning," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2021, pp. 1–6.
- [283] 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, no. 3, pp. 1442–1456, Mar. 2021.
- [284] 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.
- [285] W. Xu, F. Gao, S. Jin, and A. Alkhateeb, "3D scene-based beam selection for mmWave communications," *IEEE Wireless Commun. Lett.*, vol. 9, no. 11, pp. 1850–1854, Nov. 2020.
- [286] 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. (VTC-Spring)*, May 2020, pp. 1–5.
- [287] 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, no. 4, pp. 2699–2714, Apr. 2023.
- [288] 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, no. 3, pp. 2979–2990, Mar. 2022.
- [289] 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. 51832–51846, 2021.
- [290] 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, no. 10, pp. 2269–2273, Oct. 2021.
- [291] M. Dias, A. Klautau, N. González-Prelcic, and R. W. Heath Jr., "Position and LiDAR-aided mmWave beam selection using deep learning," in *Proc. IEEE 20th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2019, pp. 1–5.
- [292] 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, no. 12, pp. 8145–8158, Dec. 2021.
- [293] B. Salehihiouei et al., "Deep learning on multimodal sensor data at the wireless edge for vehicular network," *IEEE Trans. Veh. Technol.*, vol. 71, no. 7, pp. 7639–7655, Jul. 2022.
- [294] T. Zhang, J. Liu, and F. Gao, "Vision aided beam tracking and frequency handoff for mmWave communications," in *Proc. IEEE Conf. Comput. Commun. Workshops*, May 2022, pp. 1–2.
- [295] 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 Trans. Commun.*, vol. 63, no. 9, pp. 3029–3056, Sep. 2015.
- [296] 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 Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7538–7550, Nov. 2017.
- [297] M. Boban et al., "Multi-band vehicle-to-vehicle channel characterization in the presence of vehicle blockage," *IEEE Access*, vol. 7, pp. 9724–9735, 2019.
- [298] 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 *Proc. IEEE 84th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2016, pp. 1–6.
- [299] V. Raghavan et al., "Statistical blockage modeling and robustness of beamforming in millimeter-wave systems," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 7, pp. 3010–3024, Jul. 2019.
- [300] 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)*, Dec. 2015, pp. 1–6.
- [301] 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.*, May 2020, pp. 1–5.
- [302] G. Charan, M. Alrabeiah, and A. Alkhateeb, "Vision-aided 6G wireless communications: Blockage prediction and proactive handoff," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 10193–10208, Oct. 2021.
- [303] G. Reus-Muns et al., "Deep learning on visual and location data for V2I mmWave beamforming," in *Proc. Int. Conf. Mobility Sens. Netw. (MSN)*, Dec. 2021, pp. 559–566.
- [304] D. Marasinghe, N. Rajatheva, and M. Latva-aho, "LiDAR aided human blockage prediction for 6G," in *Proc. IEEE Globecom Workshops*, Dec. 2021, pp. 1–6.
- [305] 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, no. 3, pp. 565–569, Mar. 2023.
- [306] 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.
- [307] U. Demirhan and A. Alkhateeb, "Radar aided proactive blockage prediction in real-world millimeter wave systems," in *Proc. IEEE Int. Conf. Commun.*, Mar. 2022, pp. 4547–4552.

ABOUT THE AUTHORS

Nuria González-Prelcic (Senior Member, IEEE) received the Ph.D. degree (honors) from the University of Vigo, Vigo, Spain, in 2000.

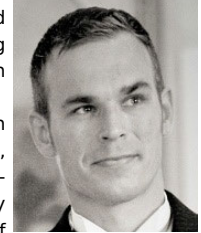
She previously held faculty positions at the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC, USA, from 2020 to 2023; and the Signal Theory and Communications Department, University of Vigo, from 2000 to 2020. She also held visiting positions at The University of Texas at Austin, Austin, TX, USA; and The University of New Mexico, Albuquerque, NM, USA. She was the Founding Director of the Atlantic Research Center for Information and Communication Technologies (atlanTtic), University of Vigo, from 2008 to 2017. She has been a Professor at the Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA, USA, since January 2024. Her main research interests include signal processing and machine learning for wireless communications, with a focus on multiple-input-multiple-output (MIMO) processing for millimeter-wave (mmWave) communication, joint sensing and communication, sensor-aided communication, signal processing under hardware impairments, vehicular communication, and multiantenna technology for LEO satellite communication. She has published more than 150 articles in these areas, including a highly cited tutorial on signal processing for mmWave MIMO published in *IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING*, which has received the 2020 IEEE SPS Donald G. Fink Overview Paper Award, and a paper pioneering the idea of enabling automotive radar with a Wi-Fi waveform that won the 2022 IEEE Vehicular Technology Society Best Vehicular Electronics Paper Award.

Dr. González-Prelcic has been an Editor of *IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS* and *IEEE TRANSACTIONS ON COMMUNICATIONS*.



Ossi Kaltiokallio (Member, IEEE) received the Ph.D. degree in electrical engineering from Aalto University, Espoo, Finland, in 2017.

He has held a postdoctoral fellow position at the University of Utah, Salt Lake City, UT, USA; Aalto University; and Tampere University, Tampere, Finland. He is currently a Senior Research Fellow with the Unit of Electrical Engineering, Tampere University. His current research interests include the intersection of statistical signal processing and wireless networking for improving radio-based positioning and sensing technologies.



Mikko Valkama (Fellow, IEEE) received the M.Sc. (Tech.) and D.Sc. (Tech.) degrees (honors) from Tampere University of Technology, Tampere, Finland, in 2000 and 2001, respectively.

In 2003, he was with the Communications Systems and Signal Processing Institute, San Diego State University (SDSU), San Diego, CA, USA, as a Visiting Research Fellow. He is currently a Full Professor and the Head of the Unit of Electrical Engineering, Tampere University, Tampere. His research interests include radio communications, radio localization, and radio-based sensing, with particular emphasis on 5G and 6G mobile radio networks.



Davide Dardari (Senior Member, IEEE) has been a Research Affiliate at Massachusetts Institute of Technology, Cambridge, MA, USA. He is currently a Full Professor at the University of Bologna, Bologna, Italy. He has published more than 280 technical articles and played several important roles in various national and European projects. His research interests include wireless communications, smart radio environments, localization techniques, and distributed signal processing.



Dr. Dardari is a Senior Member of the Editorial Board of *IEEE Signal Processing Magazine*. He received the IEEE Aerospace and Electronic Systems Society's M. Barry Carlton Award in 2011 and the IEEE Communications Society's Fred W. Ellersick Prize in 2012. He was the Chair of the Radio Communications Committee and a Distinguished Lecturer of the IEEE Communications Society from 2018 to 2019. He was the Co-General Chair of the 2011 IEEE International Conference on Ultra-Wideband and a Co-Organizer of the IEEE International Workshop on Advances in Network Localization and Navigation (ANLN)—ICC 2013–2016 editions. He served as an Editor for *IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS* from 2006 to 2012 and a guest editor for several journals.

Musa Furkan Keskin (Member, IEEE) received the B.S., M.S., and Ph.D. degrees from the Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey, in 2010, 2012, and 2018, respectively.

He is currently a Research Specialist with the Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden. His current research interests include integrated sensing and communications, reconfigurable intelligent surface (RIS)-aided localization and sensing, and hardware impairments in beyond 5G/6G systems.



Xiao Shen (Student Member, IEEE) received the B.E. degree in electronic engineering from Tsinghua University, Beijing, China, in 2020, where he is currently working toward the Ph.D. degree at the Department of Electronic Engineering.

His research interests include integrated sensing and communication, network localization, and statistical inference.



Murat Bayraktar (Graduate Student Member, IEEE) received the B.S. and M.S. degrees in electrical and electronics engineering from Middle East Technical University (METU), Ankara, Turkey, in 2018 and 2021, respectively. He is currently working toward the Ph.D. degree in electrical engineering at the University of California San Diego, La Jolla, CA, USA.



He has been a Graduate Student Researcher with the University of California San Diego since 2024. He was a Ph.D. student and a Graduate Research Assistant at North Carolina State University, Raleigh, NC, USA, from 2021 to 2024. His current research interests include signal processing for wireless communications, with a particular focus on millimeter-wave (mmWave) systems, sparse methods for channel estimation, integrated sensing and communication, wireless localization, and full-duplex systems.

Yuan Shen (Senior Member, IEEE) received the B.E. degree in electronic engineering from Tsinghua University, Beijing, China, in 2005, and the S.M. and Ph.D. degrees in electrical engineering and computer science from Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 2008 and 2014, respectively.

He is currently a Full Professor with the Department of Electronic Engineering, Tsinghua University. His research interests include network localization and navigation, integrated sensing and control, and multiagent systems.

Dr. Shen's papers have received the IEEE ComSoc Fred W. Ellersick Prize and several best paper awards from IEEE conferences. He has served as the TPC Symposium Co-Chair for IEEE ICC and IEEE Global Communications Conference several times. He was the Elected Chair of the IEEE ComSoc Radio Communications Committee from 2019 to 2020. He is an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, and *China Communications*.



Henk Wymeersch (Fellow, IEEE) received the Ph.D. degree in electrical engineering/applied sciences from Ghent University, Ghent, Belgium, in 2005.

He was a Postdoctoral Researcher with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, USA, from 2005 to 2009. He is currently a Professor of communication systems with the Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden. His current research interests include the convergence of communication and sensing, in a 5G and beyond 5G context.

Dr. Wymeersch is a Senior Member of the Editorial Board of *IEEE Signal Processing Magazine*. He served as an Associate Editor for IEEE COMMUNICATION LETTERS from 2009 to 2013 and IEEE TRANSACTIONS ON COMMUNICATIONS from 2016 to 2018. He has been serving as an Associate Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS since 2013. From 2019 to 2021, he was a Distinguished Lecturer of the IEEE Vehicular Technology Society.

