THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Machine Learning for Integrated Sensing and Communication under Modeling Mismatch

José Miguel Mateos Ramos

Department of Electrical Engineering
Chalmers University of Technology
Gothenburg, Sweden, 2026

Machine Learning for Integrated Sensing and Communication under Modeling Mismatch

José Miguel Mateos Ramos
ISBN 978-91-8103-456-1

Acknowledgements, dedications, and similar personal statements in this thesis, reflect the author's own views.

© José Miguel Mateos Ramos 2026 except where otherwise stated.

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 5913
ISSN 0346-718X

Department of Electrical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg, Sweden
Phone: +46 (0)31 772 1000

Printed by Chalmers Digital Printing
Gothenburg, Sweden, August 2026

Machine Learning for Integrated Sensing and Communication under Modeling Mismatch

José Miguel Mateos Ramos
Department of Electrical Engineering
Chalmers University of Technology

Abstract

Integrated sensing and communication (ISAC) is envisioned as a key functionality of sixth-generation wireless systems, enabling spatial perception of the environment through the same hardware, spectrum, and waveforms used for data transmission. Realizing this vision is challenged by modeling mismatch between assumed signal models and real operating conditions, which conventional model-based signal processing struggles to address. This thesis studies two such forms of mismatch: (i) hardware impairments and (ii) complex propagation conditions such as non-line-of-sight. To address them, it investigates machine learning as a data-driven complement to model-based signal processing in ISAC, with contributions spanning supervised, semi-supervised, unsupervised, and self-supervised learning regimes across four appended papers.

Paper A proposes a model-based end-to-end learning framework that jointly optimizes the ISAC transmitter and sensing receiver under antenna-array hardware impairments, enabling supervised calibration through differentiable ISAC algorithms. The results show that learning the parameterized impairments outperforms standard model-based calibration while generalizing to unseen scenarios. Paper B reduces the labeling cost of this framework, showing that semi-supervised learning matches fully supervised performance with far less labeled data, while purely unsupervised learning falls short. Paper C closes this gap in the multi-target case, calibrating transmitter and receiver impairments with no labeled data via an approximation of the channel gradient, performing close to supervised calibration and generalizing better across signal-to-noise ratios.

Paper D removes the reliance on parametric line-of-sight channel models, addressing active positioning through a self-supervised channel charting (CC) framework augmented with a digital twin. By matching large-scale channel state information features, the proposed method outperforms the CC state of the art and remains robust to modeling mismatch and distribution shifts.

Keywords: Array calibration, channel charting, digital twin, integrated sensing and communication, machine learning, modeling mismatch, positioning.

List of Publications

This thesis is based on the following publications:

[A] **José Miguel Mateos-Ramos**, Christian Häger, Musa Furkan Keskin, Luc Le Magoarou, Henk Wymeersch, "Model-Based End-to-End Learning for Multi-Target Integrated Sensing and Communication Under Hardware Impairments". IEEE Transactions on Wireless Communications, vol. 24, no. 3, pp. 2574–2589, Mar. 2025.

[B] **José Miguel Mateos-Ramos**, Baptiste Chatelier, Christian Häger, Musa Furkan Keskin, Luc Le Magoarou, Henk Wymeersch, "Semi-Supervised End-to-End Learning for Integrated Sensing and Communications". IEEE International Conference on Machine Learning, Communications, and Networking (ICMLCN), Stockholm, Sweden, pp. 132–138, 2024.

[C] **José Miguel Mateos-Ramos**, Baptiste Chatelier, Luc Le Magoarou, Nir Shlezinger, Christian Häger, Henk Wymeersch, "Unsupervised End-to-End Array Calibration for Multi-Target Integrated Sensing and Communication". submitted to IEEE Transactions on Wireless Communications, Apr. 2026.

[D] **José Miguel Mateos-Ramos**, Frederik Zumegen, Henk Wymeersch, Christian Häger, Christoph Studer, "Positioning via Digital-Twin-Aided Channel Charting with Large-Scale CSI Features". submitted to IEEE Transactions on Wireless Communications, Nov. 2025.

Other publications by the author, not included in this thesis, are:

[E] **J. M. Mateos-Ramos**, C. Häger, M. F. Keskin, L. L. Magoarou, H. Wymeersch, "Model-driven end-to-end learning for integrated sensing and communication". *Proc. IEEE Int. Conf. Commun. (ICC)*, Rome, Italy, 2023.

[F] **J. M. Mateos-Ramos**, C. Häger, M. F. Keskin, L. L. Magoarou, H. Wymeersch, "Unsupervised learning for gain-phase impairment calibration in ISAC systems". *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, Hyderabad, India, 2025.

[G] S. Rivetti, **J. M. Mateos-Ramos**, Y. Wu, J. Song, M. F. Keskin, V. Yajnanarayana, C. Häger, H. Wymeersch, "Spatial Signal Design for Positioning via End-to-End Learning”. *IEEE Wireless Commun. Letters*, vol. 12, no. 3, pp. 525–529, Mar. 2023.

[H] H. Wymeersch, S. Saleh, A. Nimr, R. Halili, R. Berkvens, M. H. Moghaddam, **J. M. Mateos-Ramos**, A. Stavridis, S. Wänstedt, S. Barmpounakis, B. Priyanto, M. Beale, J. v. d. Beek, Z. Ye, M. Manalastas, A. Kousaridas, G. P. Fettweis, “Joint Communication and Sensing for 6G - A Cross-Layer Perspective”. *Proc. IEEE Int. Symposium Joint Commun. Sensing*, Leuven, Belgium, pp. 1–6, 2024.

[I] B. Chatelier, **J. M. Mateos-Ramos**, V. Corlay, C. Häger, M. Crussière, H. Wymeersch, L. L. Magoarou, “Physically parameterized differentiable MUSIC for DoA estimation with uncalibrated arrays”. *Proc. IEEE Int. Conf. Commun. (ICC)*, Montreal, Canada, 2025.

[J] M. C. Ilter, M. F. Keskin, **J. M. Mateos-Ramos**, C. Häger, M. Valkama, H. Wymeersch, “End-to-End Learning for RIS Profile Design and Channel Parameter Estimation under Pixel Failures”. *Proc. IEEE Vehicular Techn. Conf. (VTC2025-Spring)*, Oslo, Norway, 2025.

[K] H. Farhadi, B. Banerjee, R. Berkvens, N. N. Bhat, E. Bodji, D. Dampahalage, E. Eldeeb, J. Famaey, G. P. Fettweis, J. Jeong, D. Korpi, S. Kumar, Y. Lebrun, R. Legouable, **J. M. Mateos-Ramos**, N. H. Mahmood, M. H. Moghaddam, A. Nimr, N. Rajatheva, N. Rajapaksha, A. Stavridis, T. Svensson, H. Yu, L. Wilhelmsson, H. Wymeersch, “6G AI-Driven Air Interface — Hexa-X-II View”. *IEEE Commun. Magazine*, vol. 63, no. 10, pp. 118–125, Oct. 2025.

[L] V. Shatov, S. Schieler, C. Muth, **J. M. Mateos-Ramos**, I. Bizon, F. Euchner, S. Semper, S. t. Brink, G. Fettweis, C. Häger, H. Wymeersch, L. Schmalen, R. S. Thomä, N. Franchi, “Integrated Radio Sensing Capabilities for 6G Networks: AI/ML Perspective”. *IEEE Commun. Surveys Tutorials*, vol. 28, pp. 5081–5120, Feb. 2026.

Contents

Acknowledgments

I consider myself lucky to have been surrounded by such fantastic people during my PhD journey. First and foremost, I am deeply grateful to my supervisors, Christian Häger and Henk Wymeersch. Their guidance, availability, and passion for the work are in a class of their own. Between them they have taught me a myriad of skills, from getting unstuck in my day-to-day research to producing high-quality papers and presentations, building collaborations, and fostering a supportive working environment. I would also like to thank my co-supervisor Musa Furkan Keskin for many insightful discussions, late-night reviews, and for trusting me as his go-to person for machine learning in the early stages of my PhD.

My work would not have been possible without my collaborators. Thank you to Luc Le Magoarou, Baptiste Chatelier, and Nir Shlezinger for the exchanges on array calibration, machine learning, and optimization. A big thank you to Christoph Studer, Frederik Zumegen, and Sueda Taner for our conversations on channel charting, and to all the colleagues at ETH Zürich who made me feel at home during my visit. I am thankful to Stefan Schwarz, not only for the technical discussions but also for being so understanding of my hectic schedule while I was at his lab, and to everyone at TU Wien for the moments we shared.

Throughout my PhD I met incredible people who made the journey far more enjoyable. I want to thank Jinxiang and Yibo for mentoring me since my master's thesis and for welcoming me into the group; my former and current officemates Carl, Björn, Yu, Ziyong, Xin, and Georg for countless fikas and hangouts; and the rest of my colleagues in the Communication Systems group at Chalmers, with whom I have built lifelong memories. Special thanks to Yuhao and the party toastmasters for helping me arrange the PhD thesis preparations. I feel proud to have been part of the Communication Systems group led by Fredrik Brännström, whose leadership and ambition have created an excellent research environment.

I also want to thank my friends beyond Chalmers: Nerea, Pilar, Marta, and Octavio. Even though we often struggle to find time for a videocall, our time together is always wonderful. A wholehearted thank you to my family, who have supported and encouraged me in my studies and career and have always been there whenever I needed them.

Last but not least, words cannot fully express my feelings for my girlfriend Meagan. From the very beginning of our relationship, you have shown me

your devotion, where the distance that separates us has not been a barrier keeping us apart, but rather something to dream about being together. You have been my cornerstone and my emotional support, in both my good and bad times. I have no doubt that when I read these words again in the future, our bond will still feel stronger than ever.

Financial support

This work was supported, in part, by the Chalmers Artificial Intelligence (AI) Research Center Consortium (CHAIR), the National Academic Infrastructure for Supercomputing in Sweden (NAISS), the MSCA-IF grant 888913 (OTFS-RADCOM), the Swedish Foundation for Strategic Research (SSF) through Software Artificial Intelligence for Communications (SAICOM) under Grant FUS21-0004, the European Union's Horizon Europe Research and Innovation Program through Hexa-X under Grant 101015956 and Hexa-X-II under Grant 101095759, and the Swedish Research Council (VR) under Grants 2022-03007 and 2024-04390.

I would also like to acknowledge the Ericsson Research foundation, the Solveig and Karl G. Eliassons memorial and supplementary fund, and the Wilhelm and Martina Lundgrens science fund for supporting my research trips; and the Adlerbert Foreign Student Hospitality Foundation for supporting my stay in Sweden.

Acronyms

6G: sixth generation of wireless systems

ADI: antenna displacement impairment

ADM: angle–delay map

AoA: angle-of-arrival

AP: access point

AWGN: additive white Gaussian noise

BS: base station

CC: channel charting

CCE: categorical cross-entropy

CSI: channel state information

D-MIMO: distributed multiple-input multiple-output

DT: digital twin

GOSPA: generalized optimal sub-pattern assignment

GPI: gain-phase impairment

ISAC: integrated sensing and communication

LS: least squares

LoS: line-of-sight

MB-ML: model-based machine learning

MDE: mean distance error

ML: machine learning

MLE: maximum likelihood estimation

MLP:	multilayer perceptron
mMIMO:	massive multiple-input multiple-output
mmWave:	millimeter wave
MSE:	mean squared error
NLoS:	non-line-of-sight
NN:	neural network
OFDM:	orthogonal frequency-division multiplexing
OMP:	orthogonal matching pursuit
RF:	radio-frequency
RX:	receiver
SGD:	stochastic gradient descent
SL:	supervised learning
SNR:	signal-to-noise ratio
SSL:	semi-supervised learning
TX:	transmitter
UE:	user equipment
UL:	unsupervised learning
ULA:	uniform linear array

Part I

Overview

CHAPTER 1

Introduction

1.1 Background and Motivation

1.1.1 Integrated Sensing and Communication

Since their introduction to the public in the late 1970s [1], wireless communication technologies have become an essential tool in our daily activities. In the last few decades, we have experienced an exponential increase in the data traffic in wireless networks. As a matter of fact, the global mobile network data traffic has quadrupled from the last quarter of 2018 to the third quarter of 2025, and the number of mobile subscriptions is expected to reach 9.5 billion by 2031 [2]. From the service provider standpoint, communication networks have evolved according to new generations introduced approximately every 10 years. The goal of previous generations was to provide faster, more reliable, and lower-latency communications with a growing number of connected devices per square kilometer [3]. However, recent efforts for the upcoming sixth generation of wireless systems (6G) have broadened this perspective, shifting the focus beyond wireless communications [4].

In particular, 6G has been proposed to include sensing capabilities in the communication infrastructure, where the same transmitted signals used for communications can be leveraged to retrieve spatial information of the environment [5]. This dual-purpose functionality is known as integrated sensing and communication (ISAC) [6]. The goal of wireless sensing is to detect and locate objects in the environment of the transmitter (TX) and receiver (RX), ranging from a few meters (typically in the far-field region) to the cell-coverage scale. Embedding sensing in the communication infrastructure enables new functionalities that were absent in previous wireless generations, such as smart homes, autonomous vehicles, weather monitoring, and e-health [7], [8]. In addition, reusing the same signals for communication and sensing leads to hardware and energy savings [9], since dedicated sensing infrastructure can be avoided. Among the sensing tasks enabled by ISAC, estimating the position of targets in the environment is a particularly prominent example, and it is the central focus of the papers in this thesis.

Communication and sensing processing have traditionally relied on mathematical models, which have achieved remarkable performance in their respective domains. However, these models face fundamental limitations when applied to ISAC in next-generation wireless systems. Firstly, 6G is expected to operate at higher frequencies [10], where hardware impairments become more pronounced [11]. Although such impairments could in principle be incorporated into the mathematical models, characterizing them precisely is often analytically intractable, and the resulting ISAC algorithms can become prohibitively expensive. As a consequence, hardware impairments are typically neglected in the design of ISAC algorithms, introducing a mismatch between the assumed and the actual signal model. Secondly, mathematical models (particularly for sensing) rely on simplifying propagation assumptions, such as the existence of a line-of-sight (LoS) path between the TX and the RX [12], [13], [14]. Although some works incorporate additional effects such as non-line-of-sight (NLoS) propagation [15], [16] or interference from other TXs [17], accurately accounting for clutter, diffraction, and propagation under varying weather conditions, among other effects, remains extremely challenging.

Addressing the gap between modeling assumptions and the physical effects of next-generation networks is central to the deployment of ISAC, and, as discussed in the next subsection, the same challenge extends to a complementary scenario in which the device to be localized actively transmits its own signal.

1.1.2 Active Positioning

The positioning task underlying ISAC sensing assumes that the object to be localized does not take part in the transmission, but is rather illuminated by a base station (BS) that infers its position from the backscattered signal. In many wireless applications, however, the device to be localized actively transmits. We refer to this complementary paradigm as active positioning, where a user equipment (UE) emits a known pilot signal that is received by one or more access points (APs) deployed in the environment, and the network estimates the UE's spatial coordinates from the resulting channel state information (CSI). Since the same transmission simultaneously carries communication information addressed to the network, active positioning is itself an instance of ISAC, in which the sensing functionality is obtained as a by-product of ongoing communication.

Model-based active positioning relies on parametric channel models that link the CSI to the UE position through geometric quantities such as time-of-arrival or angle-of-arrival (AoA), typically under the assumption of a dominant LoS path between the UE and each AP. These assumptions seldom hold in the cluttered indoor environments that motivate most active-positioning deployments, where walls, furniture, and people block the LoS path and induce rich multipath propagation [18]. As in ISAC sensing, hardware impairments at the UE and APs further widen the gap between the assumed and the actual signal model. Both ISAC sensing and active positioning therefore share a common limitation: the mathematical models that underpin conventional signal processing struggle to capture the physical effects of next-generation wireless systems. These limitations motivate the use of data-driven approaches capable of learning directly from observations and adapting to the imperfections that mathematical models must either ignore or approximate.

1.1.3 Machine Learning for ISAC and Active Positioning

Machine learning (ML) has emerged as the dominant data-driven paradigm of the last decade, fueled by the advent of dedicated hardware such as the graphics processing unit and by the increasing availability of large datasets. ML algorithms have achieved excellent performance in a wide range of applications, including natural language processing [19], image generation [20], and protein structure prediction [21], and have surpassed human performance in image recognition [22] and in playing chess [23]. A common thread across these successes is that ML can extract useful structure from data even when an exact mathematical description of the underlying process is unavailable or too complex to manipulate analytically. This is precisely the regime in which next-generation wireless systems are expected to operate in some applications, given the hardware impairments and propagation conditions discussed in the previous section. In the following, we describe three concrete benefits that ML can offer for ISAC and active positioning, together with the trade-offs that they entail.

Firstly, ML methods are less reliant on explicit mathematical models of the propagation environment and the transceiver chain. Conventional algorithms are derived from such models and therefore degrade when the underlying assumptions do not hold, for example when hardware impairments are not properly compensated and induce a mismatch between the assumed and the actual signal model [24], [25]. ML methods are typically trained by minimizing a cost function on observed data that already reflects these effects, so the learned parameters can absorb residual mismatches that are difficult to characterize analytically [26], [27]. This relaxation of explicit modeling does not eliminate prior assumptions, but rather displaces them. In particular, ML methods commonly require representative training data, which in the context of ISAC and positioning often takes the form of labeled samples such as the ground-truth positions of the targets or the known pilot signals received at the BS. The cost of acquiring these two types of samples is fundamentally different. Ground-truth target positions can only be obtained through dedicated measurement campaigns and are therefore costly and time-consuming to collect. Pilot symbols, in contrast, are known at the RX by design and incur no labeling cost, but increasing pilot density to enlarge the training set reduces the spectral efficiency available for data transmission. In either case, the resulting models may need to be retrained whenever the deployment conditions change significantly with respect to those seen during training.

Secondly, ML can provide computationally tractable approximations to estimation problems whose optimal solution scales poorly with the problem size. For instance, in multiple-input multiple-output communications, the maximum likelihood estimation (MLE) of the transmitted signal at the RX becomes intractable as the number of antennas grows [28]; and in radar sensing, the MLE of target range and velocity becomes intractable for an increasing number of targets [29]. ML offers tractable approximations to such estimators at the cost of a controlled performance loss [30], and dedicated hardware architectures can further reduce the energy consumption of the resulting inference [31], [32].

Thirdly, wireless networks comprise heterogeneous devices with diverse transmitted signals, data formats, and requirements [33]. Designing a separate model-based algorithm for each device class can require substantial engineering effort [34], a burden that is expected to grow with the increasing number of interconnected devices in future networks. ML can exploit correlations across data collected by different devices to provide unified estimation pipelines that aim to generalize across device classes and deployments [35]. The extent to which a single learned model can absorb such heterogeneity, however, remains an open question. Wireless data is highly configuration-dependent and tied to specific simulators or deployment setups, so the broad, foundation-model-style scaling that has succeeded in language and vision has yet to materialize in the wireless domain [36]. This suggests that purely ML pipelines are unlikely to replace model-based processing in the near term, motivating hybrid designs that retain the structure of model-based algorithms where it is reliable.

Inspired by these arguments and by the early results obtained in physical-layer communications [37], [38], [39], [40], ML has become one of the central pillars of the ongoing 6G research efforts, which envision the upcoming generation as the first ML-native wireless network [41]. As the discussion above illustrates, ML and model-based signal processing exhibit complementary strengths and weaknesses. This complementarity motivates the perspective adopted in this thesis, in which ML is harnessed as a data-driven complement to model-based signal processing for ISAC rather than as a replacement for it, as detailed in the next section.

1.2 Thesis Scope

This thesis investigates the use of ML as a data-driven complement to model-based signal processing in ISAC systems affected by hardware impairments and challenging propagation environments. The scope spans from supervised calibration under the availability of ground-truth target positions and pilot signals, to fully unsupervised calibration that eliminates ground-truth data entirely, and finally to model-agnostic positioning under NLoS conditions. The contributions are organized in the following four papers.

[Paper A] addresses the impact of hardware impairments on ISAC performance through an end-to-end model-based ML framework. Unlike prior works that optimize either the TX or the RX in isolation, [Paper A] proposes the joint learning of both the ISAC TX and sensing RX, enabling simultaneous multi-target detection and positioning under hardware impairments. The framework is trained in a supervised manner using the true positions of the targets and communication pilot messages as labeled data. Results show that the proposed approach significantly outperforms conventional model-based calibration and exhibits good generalization to unseen scenarios.

[Paper B] builds on [Paper A] and addresses a practical limitation of supervised learning (SL) in ISAC: collecting labeled target positions is costly and time-consuming in real sensing deployments. Working in a simplified single-target version of the scenario in [Paper A], [Paper B] investigates to what extent labeled data can be reduced. An unsupervised loss function driven solely by the received signal is proposed, together with semi-supervised learning (SSL) strategies that combine small amounts of labeled data with unlabeled data. SSL is observed to match fully SL performance while requiring substantially fewer labeled samples. However, relying exclusively on the unsupervised loss still falls short of supervised performance, motivating further investigation.

[Paper C] extends the scenario of [Paper B] to multiple targets and closes the performance gap identified in [Paper B] with a framework that is entirely free of labeled data. New unsupervised objective functions tailored to the multi-target setting are proposed, which outperform those of [Paper B]. Despite using no labels, the proposed framework performs close to SL and to a system that has perfect knowledge of the impairments.

[Paper D] takes a complementary direction to [Paper A]–[Paper C]. While those papers showcase effective ML-based ISAC calibration, they all rely on LoS propagation and parametric channel models that break down in NLoS environments. [Paper D] proposes a communication-centric ISAC framework for positioning that requires neither labeled data nor any explicit channel model, making it suitable for NLoS scenarios. The approach is based on channel charting (CC), a self-supervised technique that maps channel measurements to position estimates, augmented with a digital twin (DT) of the environment to provide coordinates in a true spatial reference system. Simulation experiments report improvements over the state of the art together with robustness to modeling mismatches and to shifts in the testing conditions.

1.3 Thesis Outline

The thesis is divided into two parts, where the first part lays the principal foundation components to introduce the appended papers in the second part. The first part is divided as follows: Chapter 2 presents the fundamentals of ISAC, including the integration paradigms, the wireless signal models adopted in this work, and the array calibration problem. Chapter 3 introduces the ML framework used throughout the thesis, covering the main learning paradigms and principles that underpin the calibration approaches developed later. Chapter 4 reviews CSI-based positioning, focusing on model-based localization, fingerprinting, and CC. Chapter 5 summarizes the papers of Part II of the thesis and highlights the contribution of the authors. Finally, Chapter 6 underlines the concluding remarks and future research directions.

1.4 Notation

The notation in the first part of the thesis is as follows. The sets of natural, real, and complex numbers are denoted as $\mathbb{N}$, $\mathbb{R}$, and $\mathbb{C}$, respectively. The absolute value is denoted as $|x|$. The imaginary unit is denoted as $\jmath = \sqrt{-1}$. Column vectors and matrices are denoted as boldface lowercase and uppercase letters, respectively. The all-ones vector is denoted as $\mathbf{1}$. The Euclidean norm of a vector and the Frobenius norm of a matrix are denoted as $\|\mathbf{a}\|_2$ and $\|\mathbf{A}\|_F$, respectively. The element-wise magnitude of a vector is denoted as $|\mathbf{a}|$. The two-argument arctangent is denoted as $\operatorname{atan2}(y, x)$. The transpose of a matrix $\mathbf{A}$ is denoted as $\mathbf{A}^\top$. The ith element of a vector is denoted as $[\mathbf{a}]_i$. The aggregations of two matrices $\mathbf{A} \in \mathbb{R}^{M\times N}$ and $\mathbf{B} \in \mathbb{R}^{M\times N}$ along the first and second dimensions are denoted as $[\mathbf{A}; \mathbf{B}] \in \mathbb{R}^{2M\times N}$ and $[\mathbf{A}\ \mathbf{B}] \in \mathbb{R}^{M\times 2N}$, respectively. The matrix vectorization operation is denoted as $\operatorname{vec}(\mathbf{A})$. The Hadamard product is denoted as $\odot$. Sets, intervals, and tuples are denoted by curly, square, and round brackets, respectively. The cardinality of a set $\mathcal{X}$ is denoted as $|\mathcal{X}|$. The probability that a random variable takes the value x is denoted as $\Pr(x)$. The gradient of a function $f(\mathbf{x})$ with respect to the variable $\mathbf{x}$ is denoted as $\nabla_{\mathbf{x}} f(\mathbf{x})$.

CHAPTER 2

Integrated Sensing and Communication

ISAC refers to any technology that combines sensing and communication functionalities to make an efficient use of the wireless channel, hardware, and resources (e.g., power) [6], [42]. The applications of ISAC range from technological developments to new use cases [43]. The former includes, among others, THz imaging, where distributed radio-frequency (RF) transceivers emit high-frequency signals to reconstruct an image of the environment [44]; or simultaneous localization and mapping, whose objective is to estimate the position of landmarks in the environment in addition to the trajectory of the sensing platform [45], [46]. New use cases include smart cities, where sensing capabilities can be leveraged to plan traffic, detect pedestrian crossings in dangerous areas, or enable autonomous driving [47]; DTs, which consist of digital replicas of the physical world that can be used for real-time planning and optimization [48], [49]; or physical-layer security, where the sensing capability is leveraged to enhance secrecy (e.g., via beamforming), authentication, and detection of attacks such as spoofing or jamming targeting legitimate users [50], [51].

2.1 ISAC History

Historically, ISAC research started in the radar community. One of the earliest implementations of ISAC used pulse interval modulation to embed communication information in radar pulses [52]. During the 1990s–2000s, various programs fostered the integration of communication information in radar signals [53], [54], [55]. In the 2010s, the introduction of orthogonal frequency-division multiplexing (OFDM) radar helped ISAC gain traction [56]. At the same time, in the communication landscape, massive multiple-input multiple-output (mMIMO) [57] and millimeter wave (mmWave) carrier frequencies [58] were proposed to increase the bandwidth and spatial diversity in communications and reduce the physical size of the antenna arrays.

These technologies (OFDM, mMIMO, and mmWave) bridged the gap between sensing and communication. On the one hand, sensing functionalities could be achieved with traditional communication waveforms [59], [60]. On the other hand, larger bandwidths and more antenna elements increased the range and angular resolution in sensing [61]. Thus, ISAC was formally defined in the early visions of 6G [62] and is now positioned as a cornerstone of its forthcoming specifications [8], [63]. ISAC is envisioned to provide hardware sharing, energy savings, and improved channel estimation, among other benefits [64], [65].

2.2 ISAC Paradigms

ISAC technologies can be categorized depending on the level of integration between sensing and communication [66], as represented in Fig. 2.1.

1. **Shared sites:** Sensing and communication transceivers are placed on the same site. For example, a communication BS can benefit from an attached camera that provides sensing information. In this case, the design of sensing and communication algorithms is completely independent and data integration occurs after individual processing of data modalities.

2. **Shared spectrum:** Sensing and communication transceivers operate in the same frequency bands and their design must manage their interference [67]. For instance, a sensing transceiver can design the waveforms to reduce the interference to coexisting communication users [68].

3. **Shared hardware platform:** The same RF hardware is shared for both

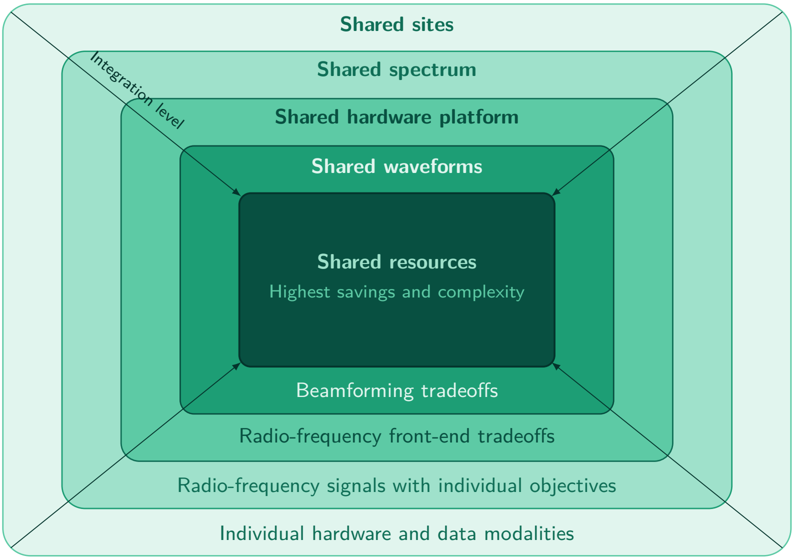


Figure 2.1: Integration levels in ISAC.

functionalities. For example, a BS that time-multiplexes sensing and communication functionalities while reusing the same antenna elements and some RF components such as amplifiers, mixers, or local oscillators. The main benefits of ISAC stem from saving hardware redundancies. The joint design must consider the interference, component linearity, dynamic range, and antenna isolation between sensing and communication functionalities [69].

4. **Shared waveforms:** Sensing and communication use the same transmit waveform. This allows the same signal-processing blocks to be reused for both functions. However, joint waveform design must balance the trade-off between sensing and communication waveform design. A common approach for waveform design is to maximize the performance of one of the functionalities while preserving a minimum quality of service for the other [70].

5. **Shared resources:** This is the tightest integration level. Hardware and resource allocation are completely shared. This level offers the

highest savings, but it implies more complexity between sensing and communication [71].

On the other hand, ISAC can be further classified depending on the goal of the system, affecting its design principles [72].

1. **Sensing-centric design:** The system is designed with a focus on sensing performance, and communication information is added as an extra functionality [73].
2. **Communication-centric design:** It endows communication systems with sensing capabilities, minimally modifying the communication infrastructure [74], [75].
3. **Joint design:** The ISAC system is designed to optimize both communication and sensing and obtain a performance trade-off between the two functionalities [76], [77], [78].

The scope of this thesis is to investigate level 5 of integration. Furthermore, in [Paper A]–[Paper C], we explore the joint design of an ISAC BS and the trade-offs between sensing and communication, while [Paper D] investigates a communication-centric scenario.

2.3 Wireless Signal Models

Most ISAC works rely on mathematical models that describe the propagation of RF signals in the environment. In the following section, we introduce the monostatic OFDM radar signal model that forms the basis of [Paper A]–[Paper C] and illustrate how hardware impairments create a mismatch between the assumed model and reality. In [Paper A]–[Paper C], we consider the monostatic ISAC scenario depicted in Fig. 2.2, where a BS uses a uniform linear array (ULA) of K antenna elements for both transmission and co-located sensing reception, while simultaneously serving a single-antenna communication RX. The transmitted signal is an OFDM waveform with S subcarriers and intercarrier spacing Δ_f.

2.3.1 Sensing Model

At most $T_{\max}$ targets are present in the sensing environment, each characterized by a complex channel gain α_t (encapsulating path loss and radar cross-section),

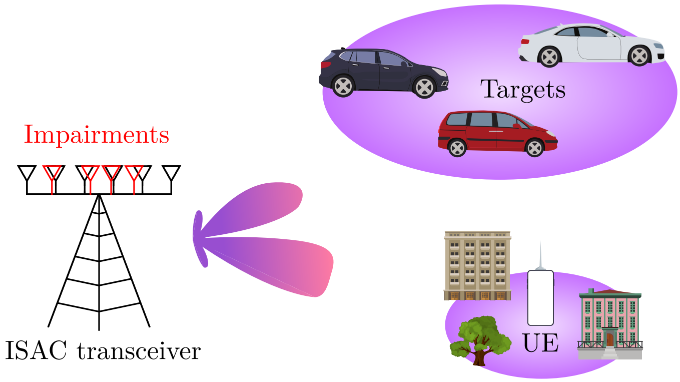


Figure 2.2: ISAC scenario considered for calibration in this thesis. The purple area around the targets and the communication RX corresponds to the prior information of their position.

angle θ_t, and range R_t. The corresponding round-trip delay is $\tau_t = 2R_t/c$, where c is the speed of light. The received sensing signal $\mathbf{Y} \in \mathbb{C}^{K \times S}$, collected across K antennas and S subcarriers, is [56], [79]

$$\mathbf{Y} = \frac{1}{\sqrt{S}} \sum_{t=1}^{T} \alpha_t \mathbf{a}_{\mathrm{rx}}(\theta_t) \mathbf{a}_{\mathrm{tx}}^{\top}(\varphi_t) \mathbf{f} [\mathbf{x} \odot \mathbf{d}(\tau_t)]^{\top} + \mathbf{W}, \tag{2.1}$$

where $T \in \{0, \ldots, T_{\max}\}$ is the number of targets, $\mathbf{a}_{\mathrm{tx}}(\varphi_t), \mathbf{a}_{\mathrm{rx}}(\theta_t) \in \mathbb{C}^K$ are the transmit and receive steering vectors of the ULA, $\mathbf{f} \in \mathbb{C}^K$ is the transmit beamformer, $\mathbf{x} \in \mathbb{C}^S$ collects the OFDM data symbols, $\mathbf{d}(\tau_t) = [1 \; e^{-\jmath 2\pi \Delta_f \tau_t} \; \cdots \; e^{-\jmath 2\pi (S-1) \Delta_f \tau_t}]^{\top}$, and $\mathbf{W}$ is additive white Gaussian noise (AWGN). In the monostatic scenario, the TX and RX share the same ULA, so $\theta_t = \varphi_t$ and $\mathbf{a}_{\mathrm{tx}}(\theta) = \mathbf{a}_{\mathrm{rx}}(\theta) = \mathbf{a}(\theta)$, with $[\mathbf{a}(\theta)]_k = e^{-\jmath 2\pi (k-(K-1)/2) d \sin(\theta)/\lambda}$ for $k = 0, \ldots, K-1$, where $d = \lambda/2$ is the antenna spacing and $\lambda = c/f_c$ is the carrier wavelength.

The steering vector $\mathbf{a}(\theta_t)$ encodes angle θ_t as a linear phase progression across the K antennas, while $\mathbf{d}(\tau_t)$ encodes range $R_t = c\tau_t/2$ as a linear phase progression across the S subcarriers. Provided that the targets are separated by more than the angle–range resolution and lie within the unambiguous angle–range region of the array and waveform, each target produces a distinct signature in this two-dimensional spatial-frequency domain. The angle–range pairs (θ_t, R_t) can then be estimated from $\mathbf{Y}$ via the orthogonal matching pursuit (OMP) algorithm [80], [81] over a discrete angle–range grid, exploiting the knowledge of the transmitted data symbols $\mathbf{x}$ available at the co-located sensing RX in the monostatic setting.

2.3.2 Communication Model

The single-antenna communication RX observes the OFDM signal in the frequency domain as [79]

$$\mathbf{y}_c = \sum_{t=1}^{\tilde{T}} \tilde{\psi}_t \, \mathbf{a}_{\mathrm{tx}}^{\top}(\tilde{\theta}_t)\mathbf{f} \left[\mathbf{x} \odot \mathbf{d}(\tilde{\tau}_t)\right] + \mathbf{n}, \tag{2.2}$$

where $\mathbf{y}_c \in \mathbb{C}^S$, $\tilde{T}$ is the number of multipath components, and $\tilde{\psi}_t$, $\tilde{\theta}_t$, $\tilde{\tau}_t$ denote the complex channel gain, angle of departure, and delay of path t, respectively. We assume that a LoS link between the TX and the RX always exists, and we index it as $t = 1$; the remaining paths ($t > 1$) are NLoS reflections from scatterers. This assumption need not hold in general (e.g., under blockage the LoS component is absent), but it is consistent with the scenarios considered in [Paper A]–[Paper C]. AWGN $\mathbf{n}$ is added at the RX. Equation (2.2) can be written compactly as $\mathbf{y}_c = \mathbf{k} \odot \mathbf{x} + \mathbf{n}$, where the channel frequency response

$$\mathbf{k} = \sum_{t=1}^{\tilde{T}} \tilde{\psi}_t \, \mathbf{a}_{\mathrm{tx}}^{\top}(\tilde{\theta}_t)\mathbf{f} \, \mathbf{d}(\tilde{\tau}_t) \in \mathbb{C}^S \tag{2.3}$$

collects the effective complex gain across all subcarriers. In general, $\mathbf{k}$ is unknown at the RX and must be estimated from pilot transmissions. However, in the signal models of [Paper A]–[Paper C], we assume $\mathbf{k}$ to be perfectly known so as to isolate the effect of hardware impairments from channel estimation errors.

The models in (2.1)–(2.3) assume a nominal, perfectly calibrated steering vector $\mathbf{a}(\theta)$. In practice, hardware impairments in the antenna array, such as gain and phase errors or element displacement, distort the true steering vector away from its nominal form [24], [82]. This mismatch degrades angle–range estimation and communication performance when conventional model-based algorithms are applied, and is the central challenge addressed by [Paper A]–[Paper C].

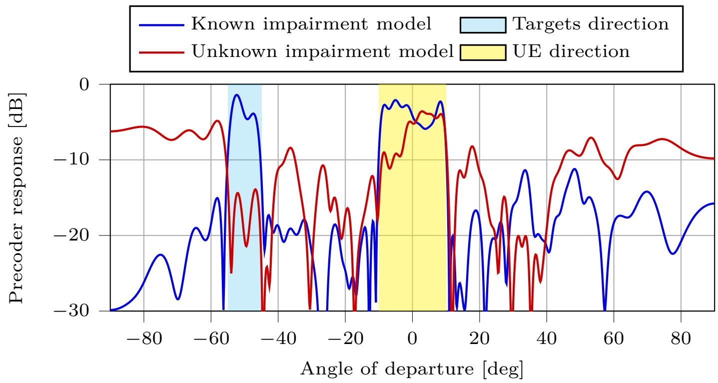


Figure 2.3: Precoder response at the TX. The colored areas correspond to the area where the targets and the communication RX lie.

2.3.3 Hardware Impairments

Across [Paper A]–[Paper C], the impairment model increases in generality: [Paper A] and [Paper B] consider antenna displacement impairments (ADIs) only in monostatic ISAC, while [Paper C] additionally incorporates gain-phase impairments (GPIs) and extends the scenario to two separate arrays.

In an ideal ULA centered at the origin, element k sits at $p_k^{\text{ideal}} = (k - (K-1)/2)\,\lambda/2$, producing the uniform phase ramp encoded in $\mathbf{a}_{\mathrm{x}}(\theta)$, for $\mathrm{x} \in \{\mathrm{tx}, \mathrm{rx}\}$. ADIs perturb each element to a true position $p_k \neq p_k^{\text{ideal}}$, breaking this uniformity and replacing $\mathbf{a}_{\mathrm{tx}}(\theta)$ and $\mathbf{a}_{\mathrm{rx}}(\theta)$ in (2.1) with the impaired steering vector

$$[\tilde{\mathbf{a}}_{\mathrm{x}}(\theta; \mathbf{p})]_k = e^{-\jmath 2\pi p_k \sin(\theta)/\lambda}, \quad k = 0, \ldots, K-1, \tag{2.4}$$

where $\mathbf{p} = [p_0, \ldots, p_{K-1}]^\top \in \mathbb{R}^K$. When the nominal $\mathbf{a}_{\mathrm{x}}(\theta)$ is used in place of $\tilde{\mathbf{a}}_{\mathrm{x}}(\theta; \mathbf{p})$, the inner products between the received signal $\mathbf{Y}$ and the assumed steering vectors may no longer peak sharply at the true target angles, causing mislocalized targets and introducing spurious detections. At the TX, the beamformer designed for $\mathbf{a}_{\mathrm{x}}(\theta)$ steers energy towards unintended directions, reducing the effective gain towards both the targets and the communication RX, as shown in Fig. 2.3.

[Paper C] further introduces GPIs in addition to ADIs, and models the transmit and receive arrays separately. The impaired steering vector model

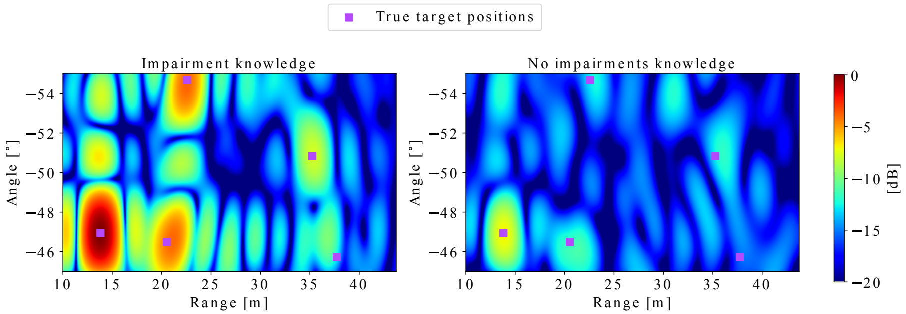


Figure 2.4: ADM at the sensing RX under ideal conditions (left) and hardware impairments (right). Each peak corresponds to a target; impairments shift the peaks from their true positions and introduce spurious lobes.

becomes

$$[\tilde{\mathbf{a}}_{\mathrm{x}}(\theta; \boldsymbol{\gamma}_{\mathrm{x}}, \mathbf{p}_{\mathrm{x}})]_k = \gamma_{\mathrm{x},k}\, e^{-\jmath 2\pi p_{\mathrm{x},k} \sin(\theta)/\lambda}, \tag{2.5}$$

where $\boldsymbol{\gamma}_{\mathrm{x}} = [\gamma_{\mathrm{x},0}, \ldots, \gamma_{\mathrm{x},K-1}]^\top \in \mathbb{C}^K$ and $\mathbf{p}_{\mathrm{x}} = [p_{\mathrm{x},0}, \ldots, p_{\mathrm{x},K-1}]^\top \in \mathbb{R}^K$. We assume that $p_{\mathrm{x},0} < \cdots < p_{\mathrm{x},K-1}$ and $|\gamma_{\mathrm{x},k}| \leq 1$ to ensure spatial ordering and no signal amplification, respectively.

The effect of hardware impairments on the sensing RX is illustrated in Fig. 2.4. The angle–delay map (ADM) is a two-dimensional representation of the matched-filter response of $\mathbf{Y}$ across a grid of candidate angles and delays. Under ideal conditions each target produces a well-defined peak at its true angle–range position. As shown in the figure, hardware impairments shift these peaks and introduce spurious lobes, ultimately hindering target detection and position estimation.

2.4 Array Calibration

Array calibration refers to the process of estimating the unknown hardware impairment parameters from received signals so that the beamformer and signal processing algorithms, which rely on the nominal steering vector $\mathbf{a}(\theta)$, can be corrected to operate with the true array response [83]. In the context of this thesis, we denote by $\boldsymbol{\phi}$ the vector of unknown impairment parameters. In the monostatic ADI model (2.4), $\boldsymbol{\phi} = \mathbf{p}$; or in the GPI and ADI model (2.5), $\boldsymbol{\phi} = [\boldsymbol{\gamma}_{\mathrm{tx}}; \boldsymbol{\gamma}_{\mathrm{rx}}; \mathbf{p}_{\mathrm{tx}}; \mathbf{p}_{\mathrm{rx}}]$. Calibration amounts to estimating $\boldsymbol{\phi}$ from a set of received signals. In conventional array processing, this is typically carried out in a dedicated calibration stage that relies on controlled inputs such as calibration sources at known angles, reference measurements obtained in an anechoic chamber, or probe signals injected through auxiliary hardware [83]. In this thesis, by contrast, we estimate $\boldsymbol{\phi}$ directly from the signals received during normal system operation, exploiting the communication and sensing waveforms already present in the ISAC link. Two broad families of methods exist for this task: model-based approaches, which exploit the known mathematical structure of the received signal to formulate an estimation problem, and data-driven approaches, which learn from data either a corrected downstream representation (e.g., a calibrated beamformer or ADM) or, as pursued in this thesis, the underlying impairment parameters $\boldsymbol{\phi}$ themselves. We describe both in the following subsections and discuss their applicability to the ISAC setting of this thesis.

2.4.1 Model-Based Calibration

To isolate the signal structure relevant for calibration, we focus for illustrative purposes on the sensing component of the ISAC system and consider a monostatic ISAC TX that uses a known constant-amplitude pilot waveform $\mathbf{x}$. Dividing the received sensing signal $\mathbf{Y}$ from (2.1) element-wise by the known pilot symbols x_s, and substituting the impaired steering vector (2.4)–(2.5), yields the pilot-normalized sensing signal $\tilde{\mathbf{Y}} \in \mathbb{C}^{K\times S}$ with model

$$\tilde{\mathbf{Y}} = \frac{1}{\sqrt{S}} \sum_{t=1}^{T} \alpha_t \, \tilde{\mathbf{a}}(\theta_t; \boldsymbol{\phi}) \, \mathbf{d}^\top(\tau_t) + \tilde{\mathbf{W}}, \tag{2.6}$$

where $\tilde{\mathbf{W}}$ is the AWGN after pilot normalization and α_t absorbs the transmit gain $\tilde{\mathbf{a}}^\top(\theta_t; \boldsymbol{\phi})\mathbf{f}$. The impairment vector $\boldsymbol{\phi}$ enters (2.6) nonlinearly through $\tilde{\mathbf{a}}(\theta_t; \boldsymbol{\phi})$, coupled with the unknown target parameters $\{(\theta_t, \tau_t, \alpha_t)\}_{t=1}^{T}$.

Conventional calibration methods exploit the signal structure of (2.6) to formulate the joint estimation of impairment and target parameters as an optimization problem. Since this joint problem is not identifiable in general, these methods rely on additional structural assumptions to resolve the ambiguity between $\boldsymbol{\phi}$ and the target parameters such as source noncoherence, angular-domain sparsity, or multiple independent snapshots. The direct formulation is a least squares (LS) problem,

$$\min_{\boldsymbol{\phi},\,\{\theta_t,\,\tau_t,\,\alpha_t\}_{t=1}^T} \left\| \tilde{\mathbf{Y}} - \frac{1}{\sqrt{S}} \sum_{t=1}^{T} \alpha_t \, \tilde{\mathbf{a}}(\theta_t; \boldsymbol{\phi}) \, \mathbf{d}^\top(\tau_t) \right\|_F^2 . \tag{2.7}$$

Since $\boldsymbol{\phi}$, $\{\theta_t\}$, and $\{\tau_t\}$ appear nonlinearly in (2.7), the problem is non-convex and its direct minimization is computationally expensive. A standard strategy is to alternate between estimating the target parameters $\{(\theta_t, \tau_t)\}$ for fixed $\boldsymbol{\phi}$, and updating $\boldsymbol{\phi}$ for fixed target parameters. Alternatively, a Bayesian framework that assigns a sparsity-promoting prior over the angular domain estimates the impairment and source parameters jointly via the expectation-maximization algorithm [84]. For GPIs (i.e., when $\boldsymbol{\phi}$ includes $\boldsymbol{\gamma}_{\mathrm{x}}$), gridless approaches based on atomic norm minimization [85] and methods based on low-rank and row-sparse matrix decomposition [86] have been proposed to estimate $\boldsymbol{\gamma}_{\mathrm{x}}$ jointly with the source directions.

When the assumed signal model is accurate and the signal-to-noise ratio (SNR) is sufficient, these model-based approaches offer, in some cases, well-understood convergence properties and Cramér-Rao lower bounds on the achievable calibration accuracy [84]. However, two specific conditions are not met in the monostatic ISAC setting of this thesis. First, most existing methods assume noncoherent sources [84], [86], [87] and break down when two or more sources are coherent, i.e., when one is a scaled and delayed copy of another. This is precisely the monostatic case: all target returns originate from the same transmitted waveform and are therefore coherent, which can cause subspace methods to fail. Second, most existing methods address a single type of impairment in isolation [85], [86], and extending them to the joint GPI and ADI model of (2.5) with separate TX and RX arrays requires significant reformulation.

2.4.2 Data-Driven Calibration

While the model-based approach in Sec. 2.4.1 exploits the analytical structure of the propagation model in (2.6) to derive estimators for $\boldsymbol{\phi}$, data-driven calibration estimates the impairment parameters directly from the received signals themselves, fitting the impaired array response to signals collected during system operation. The learning framework underlying these approaches, together with the way it reuses the conventional ISAC algorithms of Sec. 2.4.1

as its computational backbone, is the subject of Chapter 3.

A central practical question for data-driven calibration in deployed ISAC systems is what *reference information* the calibration procedure has access to. The most direct setting assumes a collection of received signals paired with the corresponding ground-truth target parameters $\{(\theta_t, \tau_t)\}$, obtained by placing calibration targets at precisely known angle–range positions. Such reference measurements provide strong guidance for estimating $\boldsymbol{\phi}$, but they are costly to acquire and often infeasible to maintain in dynamic sensing environments, in which the impairments themselves may drift over time. This motivates calibration approaches that reduce, or entirely remove, the need for reference measurements, relying instead on the structural properties of the received signal in (2.6).

[Paper A]–[Paper C] develop data-driven calibration for the monostatic ISAC setting of Sec. 2.3, with progressively weaker reference-measurement requirements and addressing the simultaneous calibration of both the TX and RX arrays.The specific estimation procedures, the way the conventional ISAC algorithms are repurposed as data-driven estimators, and the resulting contributions of each paper are presented in Chapter 3.

CHAPTER 3

Machine Learning

ML deals with learning, reasoning, and acting based on data without explicit specifications of the steps to achieve a concrete task [88]. In the context of ISAC, this implies that ML algorithms do not require an explicit mathematical description of the system, a property that makes them well suited to the hardware impairments and complex propagation environments that challenge model-based ISAC in next-generation wireless systems. This chapter presents the ML tools used in [Paper A]–[Paper D].

Sec. 3.1 introduces the general ML framework adopted throughout the thesis, formulated as a parametric function $f_{\boldsymbol{\theta}}$ whose parameters are optimized via mini-batch stochastic gradient descent (SGD), with particular emphasis on end-to-end learning across the full ISAC processing chain and on the generalization of the trained model under distribution shifts at inference. Sec. 3.2 distinguishes the three ML paradigms employed in [Paper A]–[Paper C]: (i) supervised, (ii) unsupervised, and (iii) semi-supervised. Sec. 3.3 introduces the multilayer perceptron (MLP), the neural network (NN) architecture used in [Paper D] when no analytical model of the input–output mapping is available. Finally, Sec. 3.4 presents model-based machine learning (MB-ML), where learnable parameters are embedded within conventional ISAC signal processing algorithms, as adopted in [Paper A]–[Paper C].

3.1 Machine Learning Framework

This thesis focuses on ML algorithms that operate through a *parametric function* $f_{\boldsymbol{\theta}}(\mathbf{x})$, with learnable parameters $\boldsymbol{\theta} \in \mathcal{O}$, which maps an input $\mathbf{x}$ to an estimated output $\hat{\mathbf{y}}$. The parameters $\boldsymbol{\theta}$ belong to a feasible set $\mathcal{O}$ and must be optimized to accurately perform a specific ISAC task, e.g., detecting and localizing radar targets under hardware impairments [Paper A]–[Paper C]. This optimization is carried out during a *training* phase over a dataset of observed samples, as described next. The choice of loss function used for training, discussed in Sec. 3.2, determines the ML type; and the structure of $f_{\boldsymbol{\theta}}$, discussed in Sections 3.3 and 3.4, determines the ML model.

3.1.1 Training

Given a dataset of N input–output pairs $\{(\mathbf{x}^{(i)}, \mathbf{y}^{(i)})\}_{i=1}^{N}$, the training objective is to find the parameters $\boldsymbol{\theta}^{\star}$ that minimize the average per-sample loss over the dataset:

$$\boldsymbol{\theta}^{\star} = \arg\min_{\boldsymbol{\theta}\in\mathcal{O}} \mathfrak{L}(\boldsymbol{\theta}), \qquad \mathfrak{L}(\boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^{N} \ell\Big(f_{\boldsymbol{\theta}}(\mathbf{x}^{(i)}),\, \mathbf{y}^{(i)}\Big)\,, \tag{3.1}$$

where $\ell(\cdot,\cdot)$ is a per-sample loss that depends on the estimated output $f_{\boldsymbol{\theta}}(\mathbf{x}^{(i)})$ and, depending on the ML type, on labeled data $\mathbf{y}^{(i)}$ or on side information derived from the observations (see Sec. 3.2). Note that $\mathfrak{L}(\boldsymbol{\theta})$ depends on $\boldsymbol{\theta}$ through $f_{\boldsymbol{\theta}}(\cdot)$.

In general, no closed-form solution exists for (3.1), hence $\boldsymbol{\theta}^{\star}$ is approximated through iterative first-order optimization. A common choice to solve (3.1) is gradient descent [89]. Since $\mathfrak{L}(\boldsymbol{\theta})$ is typically non-convex in $\boldsymbol{\theta}$, this procedure is not guaranteed to reach a global minimizer. Nevertheless, gradient-based optimization has proven empirically effective for training ML models [90] and is the framework adopted in [Paper A]–[Paper D]. Evaluating the gradient of $\mathfrak{L}(\boldsymbol{\theta})$, however, requires computing ℓ over all N samples per iteration, which becomes prohibitively expensive for large datasets. Mini-batch SGD addresses this by approximating the full gradient using a randomly sampled mini-batch

$\mathcal{B} \subseteq \{1, \ldots, N\}$ at each iteration, defining the mini-batch loss

$$\mathfrak{L}_{\mathcal{B}}(\boldsymbol{\theta}) = \frac{1}{|\mathcal{B}|} \sum_{i \in \mathcal{B}} \ell\Big(f_{\boldsymbol{\theta}}(\mathbf{x}^{(i)}),\, \mathbf{y}^{(i)}\Big)\,. \tag{3.2}$$

This reduces the computational cost per iteration and introduces stochastic noise that in some cases can help escape local minima [91]. Starting from an initial $\boldsymbol{\theta}^{(0)}$, the SGD update at iteration k is

$$\boldsymbol{\theta}^{(k)} = \boldsymbol{\theta}^{(k-1)} - \eta\, \nabla_{\boldsymbol{\theta}} \mathfrak{L}_{\mathcal{B}}(\boldsymbol{\theta}^{(k-1)}), \tag{3.3}$$

where $\eta > 0$ is the *learning rate* and a new mini-batch $\mathcal{B}$ is drawn randomly at each iteration.

For (3.3) to be applicable, $f_{\boldsymbol{\theta}}$ must be differentiable with respect to $\boldsymbol{\theta}$, so that the gradient $\nabla_{\boldsymbol{\theta}} \mathfrak{L}_{\mathcal{B}}$ can be computed via, for instance, backpropagation [92]. As a matter of fact, in [Paper A] and [Paper B], $f_{\boldsymbol{\theta}}$ corresponds to a conventional ISAC signal processing algorithm implemented as a differentiable computational graph, enabling gradient-based training. The structure of $f_{\boldsymbol{\theta}}$ for [Paper A]–[Paper D] is detailed in Sections 3.3 and 3.4.

3.1.2 End-to-End Learning

Conventional system design optimizes each processing block independently, which can only guarantee local optimality for each component in isolation. For example, ML has been applied individually to tasks such as channel estimation [93], signal detection [94], and beamformer design [95], each trained with a component-specific loss. End-to-end learning instead trains all components jointly by backpropagating a single system-level loss through the entire processing chain, allowing the TX to shape its output to facilitate the sensing and communication tasks at the RX, and the RX to be optimized for the signal it will actually observe during deployment [37]. This joint optimization has been shown to outperform independent block-wise training in communication systems [37], [96] and radar systems [97].

For the monostatic ISAC scenario of Sec. 2.3, the end-to-end chain spans the transmit beamformer $\mathbf{f}$ and the sensing and communication processing of the received signals $\mathbf{Y}$ in (2.1) and $\mathbf{y}_c$ in (2.2). In [Paper A]–[Paper C], end-to-end learning jointly optimizes $\mathbf{f}$ and the sensing RX, which share learnable parameters $\boldsymbol{\theta}$ representing the hardware impairments affecting both arrays. As discussed in Sec. 2.3, these impairments distort both the emitted beampattern at the TX and the steering vectors at the sensing RX, thus optimizing either component in isolation against a local objective does not necessarily translate into gains in the overall ISAC performance.

However, conventional ISAC signal processing algorithms may contain discrete selection and optimization steps that are not differentiable, obstructing gradient flow during backpropagation. In [Paper A] and [Paper B], differentiability is obtained by replacing such operations with smooth surrogates, enabling gradient-based end-to-end training through the full chain.

3.1.2.1 End-to-end Training Through the Channel

A fundamental requirement for end-to-end training is that the gradient of the system-level loss can flow back through every element of the processing chain, including the channel. Reaching the TX parameters via backpropagation requires the channel to be differentiable with respect to its inputs, a condition that holds only when a known differentiable channel model is available at training time. The wireless channel is a stochastic mapping whose distribution is generally unknown at the TX and is observable only through its inputs and outputs [98]. Even when a parametric channel model is available, the transmitted signal may pass through operations with no well-defined gradient, such as digital-to-analog conversion or discrete modulation, further obstructing gradient flow [98].

The available strategies split into *offline* training, performed before deployment using a fixed channel model as a stand-in for the real channel, and *in-situ* training, performed with the real channel during operation. Two offline variants have been proposed. When a known differentiable channel model is available, such as the AWGN or a Rayleigh fading model, it can be incorporated directly into the computational graph and gradients are computed analytically [37]. When the channel model is unknown or too complex to differentiate, an NN can be trained to approximate it from input–output measurements and used as a differentiable surrogate during backpropagation [99]. In both offline cases the channel model is fixed before TX optimization, whereas updating the TX shifts the distribution of signals entering the channel and can drive the model inaccurate at the operating point reached by the trained TX. The in-situ strategy, by contrast, avoids any differentiable channel model: the gradient of the expected loss with respect to the TX parameters is approximated using only a known perturbation distribution at the TX, without access to the channel gradient [98].

In [Paper C], the in-situ strategy is adopted. The transmit precoder is perturbed by a known random distribution, and the gradient of the expected loss is estimated using only the statistics of that perturbation instead of the channel gradient.

3.1.3 Inference

Once training is concluded, the SGD update in (3.3) returns a set of trained parameters $\hat{\boldsymbol{\theta}}$, which approximate the global minimizer $\boldsymbol{\theta}^{\star}$ in (3.1). During inference, $\hat{\boldsymbol{\theta}}$ is held fixed and the model computes $\hat{\mathbf{y}} = f_{\hat{\boldsymbol{\theta}}}(\mathbf{x})$ for a new, unseen input $\mathbf{x}$ in a single forward pass without any gradient computation.

A desirable property of the trained model is good *generalization*: the ability to maintain accurate predictions on inputs not seen during training, regardless of whether they follow the training distribution [100]. Generalization becomes particularly challenging under *distribution shift*, where the test distribution differs from the training distribution. In wireless systems, distribution shifts arise naturally. For instance, the propagation environment may change after deployment, users may operate at locations or under channel conditions not encountered during training, or system parameters may vary across deployments. A model that fails to generalize under such shifts will degrade in performance even if it performed well on training data, which is particularly problematic when collecting new labeled data to retrain is costly or impractical.

Generalization under distribution shift is therefore an important criterion for evaluating ML models in the context of this thesis. In [Paper A], generalization is assessed by testing on target angular sectors outside the range seen during training. In [Paper C], the model is trained at a fixed SNR and evaluated at different SNR levels. In [Paper D], distribution shifts are introduced by changing the height of the UE and by introducing new reflecting elements into the environment after the model has been deployed.

3.2 Types of Machine Learning

The form of the per-sample loss ℓ in (3.1) depends on what data is available during training. We distinguish three types of ML used in this thesis, each corresponding to a different assumption on the availability of labeled outputs $\mathbf{y}^{(i)}$.

3.2.1 Supervised Learning

In SL, a ground-truth label $\mathbf{y}^{(i)}$ is available for every input sample $\mathbf{x}^{(i)}$, and the per-sample loss measures the discrepancy between the estimated and labeled outputs [88]. The specific form of the loss depends on the nature of the output. For instance, measuring the discrepancy between an estimated target position and its ground-truth counterpart requires a vector-valued loss [Paper B], whereas measuring the discrepancy between an estimated set of target positions and a ground-truth set of unknown cardinality requires a set-valued loss [Paper A]. In the following, we describe two types of SL problems that arise in this thesis, each associated with different loss functions.

3.2.1.1 Regression

In regression tasks, the output contains quantitative values that can be ordered. When the output is a single vector, a standard choice is the mean squared error (MSE) loss [88]:

$$\ell^{\mathrm{MSE}}\left(\hat{\mathbf{y}}^{(i)}, \mathbf{y}^{(i)}\right) = \|\hat{\mathbf{y}}^{(i)} - \mathbf{y}^{(i)}\|_2^2. \tag{3.4}$$

For example, in [Paper B], MSE regression is used to estimate the position of a single radar target: $\hat{\mathbf{y}}^{(i)}$ is the estimated target position and $\mathbf{y}^{(i)}$ is the ground-truth position collected during a measurement campaign. In [Paper D], fingerprinting, one of the baselines to compare against our proposed approach, also relies on MSE regression, as discussed in Sec. 4.3.

The MSE loss assumes a fixed size of the output, e.g., a fixed number of targets. In [Paper A], however, the number of targets present in the environment is unknown and variable, so the estimated output $\hat{\mathbf{y}}^{(i)}$ is a set of positions rather than a single vector. Applying MSE in this case requires a fixed assignment between estimated and ground-truth targets, which is undefined when their cardinalities differ. Instead, [Paper A] employs the *generalized optimal sub-pattern assignment* (GOSPA) loss [101]. Let $\gamma > 0$, $0 < \mu \leq 2$ and $1 \leq p < \infty$. Let $\mathcal{Y} = \{\mathbf{y}_1, ..., \mathbf{y}_{|\mathcal{Y}|}\}$ and $\hat{\mathcal{Y}} = \{\hat{\mathbf{y}}_1, ..., \hat{\mathbf{y}}_{|\hat{\mathcal{Y}}|}\}$ be the finite subsets of $\mathbb{R}^2$ corresponding to the true and estimated target positions, respectively, with $0 \leq |\mathcal{Y}| \leq T_{\max}$, $0 \leq |\hat{\mathcal{Y}}| \leq T_{\max}$, and $T_{\max}$ the maximum number of targets. Let $d(\mathbf{y}, \hat{\mathbf{y}}) = \|\mathbf{y} - \hat{\mathbf{y}}\|_2$ be the distance between true and estimated positions, and $d^{(\gamma)}(\mathbf{y}, \hat{\mathbf{y}}) = \min(d(\mathbf{y}, \hat{\mathbf{y}}), \gamma)$, where γ is the cut-off distance. Let Π_n be the set of all permutations of $\{1, ..., n\}$ for any $n \in \mathbb{N}$ and any element $\pi \in \Pi_n$ be a sequence $(\pi(1), ..., \pi(n))$. For $|\mathcal{Y}| \leq |\hat{\mathcal{Y}}|$, the GOSPA loss function is defined as

$$\ell_p^{(\gamma,\mu)}(\mathcal{Y}, \hat{\mathcal{Y}}) = \left(\min_{\pi \in \Pi_{|\hat{\mathcal{Y}}|}} \sum_{i=1}^{|\mathcal{Y}|} d^{(\gamma)}(\mathbf{y}_i, \hat{\mathbf{y}}_{\pi(i)})^p + \frac{\gamma^p}{\mu}(|\hat{\mathcal{Y}}| - |\mathcal{Y}|) \right)^{\frac{1}{p}}. \tag{3.5}$$

If $|\mathcal{Y}| > |\hat{\mathcal{Y}}|$, $\ell_p^{(\gamma,\mu)}(\mathcal{Y}, \hat{\mathcal{Y}}) = \ell_p^{(\gamma,\mu)}(\hat{\mathcal{Y}}, \mathcal{Y})$.

3.2.1.2 Classification

In classification tasks, the output belongs to one of C discrete classes $\{c_j\}_{j=1}^{C}$. The parametric function outputs a probability vector $\hat{\mathbf{y}}^{(i)} \in [0,1]^C$, with $\mathbf{1}^\top \hat{\mathbf{y}}^{(i)} = 1$, where the jth element approximates the posterior probability $\Pr(c_j|\mathbf{x}^{(i)})$. The label $\mathbf{y}^{(i)}$ is the one-hot encoding of the true class [102, Ch. 5]. A standard loss for classification is the categorical cross-entropy (CCE) [103]:

$$\ell^{\mathrm{CCE}}\left(\hat{\mathbf{y}}^{(i)}, \mathbf{y}^{(i)}\right) = -\sum_{j=1}^{C} \left[\mathbf{y}^{(i)}\right]_j \log\left(\left[\hat{\mathbf{y}}^{(i)}\right]_j\right). \tag{3.6}$$

In [Paper A], the CCE loss is used for communication symbol detection at the RX: $\hat{\mathbf{y}}^{(i)}$ is the estimated symbol probability vector and $\mathbf{y}^{(i)}$ is the one-hot encoding of the true transmitted symbol.

3.2.2 Unsupervised Learning

In unsupervised learning (UL), no labeled outputs are available. The per-sample loss depends only on the estimated output and on side information derived from the structure of the observations, without access to ground-truth labels [88]. The key challenge in UL is constructing a loss ℓ^{UL} that is minimized when $f_{\boldsymbol{\theta}}$ performs the intended task accurately, despite the absence of labels. A natural approach is to design a *proxy loss* from observable quantities that serves as a surrogate for the supervised objective.

For example, when system parameters are unknown, a proxy loss can be constructed from the discrepancy between the received signal and its reconstruction under the current parameter estimate. In [Paper C], ADIs and GPIs

$\boldsymbol{\phi}$ of the transmit and receive array corrupt the OFDM signal and are unknown during deployment. The impairments are learned by minimizing the reconstruction error between the received signal matrix $\mathbf{Y}$ and its reconstruction $\hat{\mathbf{Y}}(\boldsymbol{\phi})$ under the estimated impairments and current target parameter estimates:

$$\ell^{\mathrm{UL}}(\boldsymbol{\phi}) = \|\mathbf{Y} - \hat{\mathbf{Y}}(\boldsymbol{\phi})\|_F^2. \tag{3.7}$$

This loss is minimized when $\boldsymbol{\phi}$ matches the true impairments, since only the correct estimate yields a reconstruction consistent with the received signal, requiring no labeled target positions.

3.2.3 Semi-Supervised Learning

SSL lies between SL and UL, incorporating a supervised loss over labeled samples alongside an unsupervised loss over the available data. This combination is particularly valuable when labeled samples are scarce or expensive to collect: using SL alone is constrained by the number of available labels, while UL alone may converge to solutions that minimize ℓ^{UL} without aligning with the intended task.

In the ISAC context of [Paper B], SSL is realized by first training $f_{\boldsymbol{\theta}}$ with ℓ^{MSE} over the small labeled subset to initialize the parameters in a meaningful region, and then optimizing with ℓ^{UL} over all available data. This strategy reduces the number of labeled position measurements required to reach performance comparable to fully supervised training.

3.3 Neural Networks

NNs are parametric functions $f_{\boldsymbol{\theta}}$ structured as compositions of layers, where each layer applies a learnable linear transformation followed by a nonlinear activation. This layered structure enables NNs to approximate complex input–output mappings that are difficult to characterize analytically [104, Ch. 4]. In this section, we describe the MLP, the NN architecture employed in [Paper D].

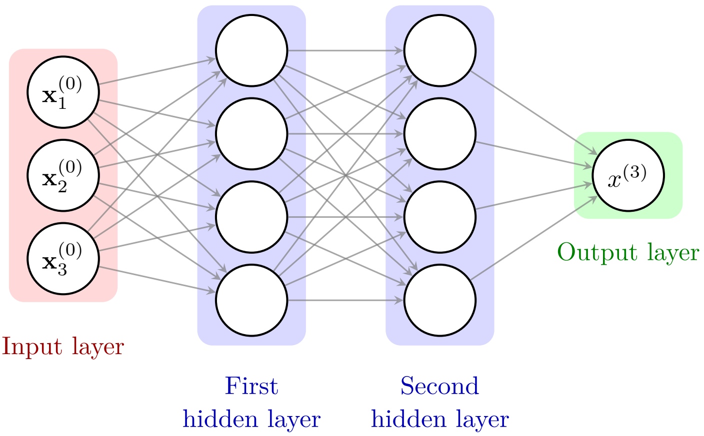


Figure 3.1: Example of an MLP.

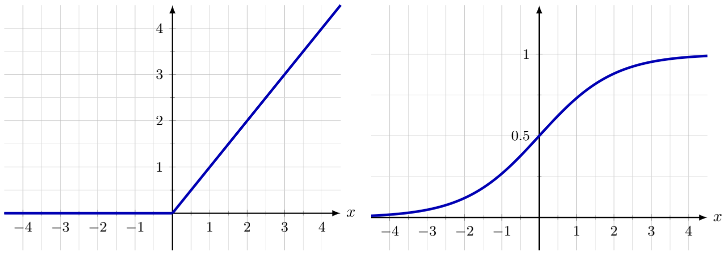
3.3.1 Multilayer Perceptron

The MLP consists of an input layer, an output layer, and one or more hidden layers, as illustrated in Fig. 3.1. The output of the lth layer is

$$\mathbf{x}^{[l]} = \sigma^{[l]}\left(\mathbf{W}^{[l]}\mathbf{x}^{[l-1]} + \mathbf{b}^{[l]}\right), \tag{3.8}$$

for $l = 1, \dots, L$, where L is the number of layers, $\mathbf{x}^{[0]}$ is the MLP input, $\sigma^{[l]}(\cdot)$ is the nonlinear *activation function* of the lth layer, and $\boldsymbol{\theta}^{[l]} = \left[\mathbf{W}^{[l]}\ \mathbf{b}^{[l]}\right]$ are its learnable parameters. The input and output dimensions are fixed by the application, while the dimension of each hidden layer corresponds to the number of neurons in that layer. Designing an MLP involves, among other choices, selecting the number of layers, neurons per layer, and activation functions.

Nonlinear activation functions allow the MLP to approximate complex input–output mappings that affine transformations alone cannot represent. Under mild regularity conditions and sufficient network size, MLPs are universal function approximators [105, Ch. 13]. Fig. 3.2 shows two representative examples of nonlinear activation functions. The rectified linear unit maintains a nonzero gradient for positive inputs, which can help mitigate the *vanishing gradient* problem [106]. The sigmoid activation constrains the output to the interval $[0, 1]$, making it suitable for the output layer when the target output



(a) Rectified linear unit. **(b)** Sigmoid.

Figure 3.2: Examples of nonlinear activation functions.

lies in a bounded range.

For classification tasks, the output of the MLP should form a probability vector with entries summing to one. This is achieved by the *softmax* activation, defined elementwise as

$$[\mathbf{y}]_i = \frac{\exp([\mathbf{z}]_i)}{\sum_{j=1}^{C} \exp([\mathbf{z}]_j)}, \tag{3.9}$$

where $\mathbf{z} \in \mathbb{R}^C$ is the pre-activation output of the last layer and $\mathbf{y} \in [0,1]^C$ is the resulting probability vector. This output can be used together with the CCE loss in (3.6) for classification tasks.

3.4 Model-Based Machine Learning

The ML types described in Sec. 3.2 characterize the training objective through the choice of loss ℓ, but place no restriction on the structure of the parametric function $f_{\boldsymbol{\theta}}$. For [Paper D], the MLP of Sec. 3.3 is a suitable choice, as no reliable analytical model for the mapping of interest is available. In [Paper A]–[Paper C], however, well-established ISAC signal processing algorithms exist for the transmit beamforming and sensing tasks, and these algorithms encode domain knowledge that a generic MLP would need to learn entirely from data. Rather than replacing these algorithms with a generic architecture such as the MLP, MB-ML introduces learnable parameters $\boldsymbol{\theta}$ at selected stages of the conventional algorithms, using their computation structure as the blueprint for $f_{\boldsymbol{\theta}}$ [107].

Compared to the MLP, MB-ML typically requires fewer training samples, since the algorithm structure already encodes domain knowledge that would otherwise need to be learned from data. Compared to the conventional algorithm, MB-ML can compensate for model mismatch by adjusting its learnable parameters, without requiring an explicit analytical characterization of the discrepancy. However, these benefits diminish if the assumed computation structure deviates substantially from the true underlying system, in which case the learnable parameters may not suffice to compensate for the mismatch.

A key design decision in MB-ML is the choice of parameterization. Some approaches substitute selected stages of the conventional algorithm with black-box NNs that learn the required mapping from data without any structural constraint [108]. When the unknown quantities appear explicitly in the algorithm's mathematical expressions, such as hardware impairment parameters that enter the signal model, the learnable parameters can instead be chosen to directly represent those physical variables within the algorithm, yielding a more compact representation that requires less training data and whose parameters retain physical meaning.

In [Paper A], the conventional ISAC algorithms for transmit beamforming and target sensing rely on ideal models of the antenna array; hardware impairments corrupt these models, causing the steering vectors used by both the beamformer and sensing RX to deviate from reality. Rather than replacing the affected components with black-box NNs, [Paper A] introduces the unknown impairments directly as the learnable parameters $\boldsymbol{\theta}$. Two levels of structure are explored: *dictionary learning* introduces a free matrix of steering vectors, offering flexibility at the cost of more learnable parameters, while *impairment learning* directly parameterizes the hardware impairments with a compact, physically interpretable set of parameters. Impairment learning is shown to outperform dictionary learning, as the more compact physical parameterization leads to faster convergence and better performance when the impairment structure is known. This motivates its adoption in [Paper B] and [Paper C], where the same parameterization is retained and the training regime is extended to semi-supervised and unsupervised settings, respectively.

In the calibration setting of [Paper A]–[Paper C], the learnable parameters $\boldsymbol{\theta}$ are the hardware impairment vector $\boldsymbol{\phi}$ of Sec. 2.4, and the parametric function $f_{\boldsymbol{\theta}}$ is formed by the conventional ISAC sensing algorithm instantiated with the current impairment estimate. Training follows the MB-ML loop of Sec. 3.1. At each iteration, a batch of received signals $\{\mathbf{Y}^{(i)}\}_{i\in\mathcal{B}}$ is processed by the sensing algorithm to produce target parameter estimates $\{\hat{\theta}_t^{(i)}, \hat{\tau}_t^{(i)}\}_{i\in\mathcal{B}}$, the loss ℓ

is evaluated against available reference information, and the gradient $\nabla_{\boldsymbol{\phi}} \ell$ is backpropagated through the algorithm to update $\boldsymbol{\phi}$ via mini-batch SGD. The nature of that reference information determines the learning paradigm. [Paper A] operates in a supervised regime, constructing ℓ from signal–position pairs $\{(\mathbf{Y}^{(i)}, \theta_t^{(i)}, \tau_t^{(i)})\}_{i \in \mathcal{B}}$ obtained from calibration sources at known locations. [Paper B] extends this to a semi-supervised setting, where a small labeled set is supplemented with unlabeled operational signals, reducing the number of reference measurements required. [Paper C] removes the need for position labels entirely by constructing ℓ from the structural properties of the signal model (2.6), yielding a fully unsupervised calibration procedure.

CHAPTER 4

CSI-Based Positioning

Chapters 2 and 3 established the ML tools and the monostatic ISAC signal model that underpin [Paper A]–[Paper C]. In those papers, the entities to be localized are passive targets that do not transmit any signal of their own: their position is recovered by an ISAC BS from the backscattered version of a self-emitted waveform, under the assumption of LoS propagation between the BS and the target. The present chapter considers a distinct but related problem, hereafter referred to as positioning, in which the entities to be localized are UEs that actively transmit pilot signals to a network of distributed APs, and whose position is inferred from the resulting uplink CSI. The classical approach to this problem mirrors the model-based estimation methodology of Chapter 2: an explicit channel model links the UE position to the observed CSI through the geometry of the LoS path, and the position is recovered by inverting that relation across the network of APs. This model-based positioning paradigm is reviewed in Sec. 4.2, where it serves as the conceptual and notational baseline for the rest of the chapter. Its applicability, however, rests on two assumptions that are difficult to meet in practice, i.e., that the propagation channel admits an accurate parametric description and that a LoS path is available between each UE and a sufficient number of APs.

The latter is particularly restrictive in the dense indoor deployments that motivate UE positioning, where walls, furniture, and other physical objects frequently obstruct the LoS path and give rise to NLoS propagation. Under such conditions, the geometric relations exploited by model-based estimators no longer hold reliably, motivating data-driven alternatives that learn the CSI-to-position mapping directly from measurements. The goal of this chapter is to introduce both families of positioning methods, culminating in the CC technique that underlies [Paper D].

4.1 The Positioning Problem

We formalize here the positioning problem for a distributed multiple-input multiple-output (D-MIMO) scenario. Consider a network of A APs deployed in the environment, each equipped with K antennas, and a set of UEs equipped with L antennas each. A multiple-access protocol schedules the UE transmissions so that the signal received at AP a at timestamp $t^{(n)}$ originates from a single UE located at the unknown position $\mathbf{x}^{(n)} \in \mathbb{R}^D$, with $D \in \{2, 3\}$. The UE transmits an OFDM pilot signal over S subcarriers known at all APs, and AP a estimates the corresponding CSI matrix $\mathbf{H}^{(a,n)} \in \mathbb{C}^{K \times S}$, which is corrupted by AWGN. Stacking the per-AP CSI matrices along the antenna dimension yields the aggregated CSI matrix

$$\mathbf{H}^{(n)} = \left[\mathbf{H}^{(1,n)}; \cdots ; \mathbf{H}^{(A,n)}\right] \in \mathbb{C}^{M \times S}, \qquad M = AK, \tag{4.1}$$

which constitutes the radio observation available to the network for positioning at timestamp $t^{(n)}$. The dataset of unlabeled observations collected over N transmissions is denoted by $\{\mathbf{H}^{(n)}, t^{(n)}\}_{n=1}^{N}$.

The positioning problem consists of designing an estimator

$$f \colon \mathbb{C}^{M \times S} \to \mathbb{R}^D, \qquad \hat{\mathbf{x}}^{(n)} = f\big(\mathbf{H}^{(n)}\big), \tag{4.2}$$

that maps the observed CSI matrix to an estimate $\hat{\mathbf{x}}^{(n)}$ of the UE position $\mathbf{x}^{(n)}$. Following [Paper D], we measure the quality of the estimator by two complementary metrics. The first is the mean distance error (MDE)

$$\mathrm{MDE} = \frac{1}{N} \sum_{n=1}^{N} \|\hat{\mathbf{x}}^{(n)} - \mathbf{x}^{(n)}\|_2, \tag{4.3}$$

which captures the average positioning accuracy. The second is the *95th-percentile distance error*, defined as the smallest value ρ_{95} such that at least 95% of the distance errors $\{\|\hat{\mathbf{x}}^{(n)} - \mathbf{x}^{(n)}\|_2\}_{n=1}^{N}$ are below ρ_{95}, which captures the tail behavior of the estimator.

The three positioning paradigms reviewed in the remainder of this chapter differ in how the estimator f in (4.2) is constructed and in what side information is assumed at design time. Sec. 4.2 first describes the model-based paradigm, which establishes the baseline for the chapter and makes explicit the LoS-assumption limitations that motivate the data-driven paradigms presented next. Sections 4.3 and 4.4 then present two such data-driven paradigms, fingerprinting and CC, which learn f from labeled and unlabeled CSI measurements, respectively.

4.2 Model-Based Positioning

Model-based positioning constructs the estimator f in (4.2) from the parametric channel model linking the per-AP CSI $\mathbf{H}^{(a,n)}$ to the UE position $\mathbf{x}^{(n)}$. After pilot normalization, the CSI at AP a admits the multipath decomposition [109]

$$\mathbf{H}^{(a,n)} = \sum_{t=1}^{T^{(a,n)}} \alpha_t^{(a,n)} \mathbf{a}(\theta_t^{(a,n)}) \mathbf{d}^\top(\tau_t^{(a,n)}) + \mathbf{W}^{(a,n)}, \tag{4.4}$$

where $T^{(a,n)}$ is the number of multipath components between the UE and AP a at timestamp $t^{(n)}$, $\alpha_t^{(a,n)} \in \mathbb{C}$ is the complex gain of path t, $\theta_t^{(a,n)}$ and $\tau_t^{(a,n)}$ are the corresponding AoA and propagation delay, and $\mathbf{W}^{(a,n)}$ collects the RX AWGN. Equation (4.4) is the uplink counterpart of the monostatic sensing model (2.1) of Chapter 2, with the round-trip replaced by a one-way link to an AP located at the known position $\mathbf{p}^{(a)} \in \mathbb{R}^D$.

The UE position enters (4.4) only through the LoS path ($t = 1$), whose delay and AoA satisfy

$$\tau_1^{(a,n)} = \|\mathbf{x}^{(n)} - \mathbf{p}^{(a)}\|_2 / c, \qquad \theta_1^{(a,n)} = \operatorname{atan2}\big([\mathbf{x}^{(n)} - \mathbf{p}^{(a)}]_2,\ [\mathbf{x}^{(n)} - \mathbf{p}^{(a)}]_1\big). \tag{4.5}$$

Model-based estimators exploit (4.5) in two steps: extract the LoS parameters $\{(\theta_1^{(a,n)}, \tau_1^{(a,n)})\}_{a=1}^{A}$ from $\{\mathbf{H}^{(a,n)}\}_{a=1}^{A}$, e.g., via subspace methods [110], OMP [80], [81], or atomic-norm minimization [85]; then invert the geometry through triangulation, trilateration, or time difference of arrival [111], [112]. Direct positioning methods instead search $\mathbb{R}^D$ by jointly maximizing a likelihood over $\{\mathbf{H}^{(a,n)}\}_{a=1}^{A}$ [113].

When the LoS path is resolvable and the arrays are calibrated, this approach yields an interpretable estimate from a single CSI snapshot with no training data. Under NLoS, however, the LoS parameters are unobservable, (4.5) cannot be enforced, and substituting an NLoS path introduces a bias that does not vanish with SNR. NLoS identification and mitigation techniques [111], [112] require either enough APs in LoS or accurate prior knowledge of the multipath structure, neither of which is generally available in dense indoor deployments. This limitation motivates the data-driven paradigms that follow, where the mapping f is learned directly from CSI and the multipath structure becomes a discriminative signature rather than a source of bias.

4.3 Fingerprinting

Fingerprinting addresses the positioning problem in a purely data-driven manner, bypassing the need for an explicit channel model. The underlying idea is that a UE at a given position generates a CSI matrix $\mathbf{H}^{(n)}$ whose structure is determined by the surrounding environment, including reflections, diffractions, and other multipath effects, and by the hardware of the APs and the UE. Although the relationship between $\mathbf{H}^{(n)}$ and the UE position $\mathbf{x}^{(n)}$ is too complex to be captured analytically in NLoS environments, it can be learned from a representative database of CSI–position pairs.

CSI fingerprinting operates in two phases [114], [115], as illustrated in Fig. 4.1. In the *offline training phase*, a labeled dataset $\{\mathbf{H}^{(n)}, \mathbf{x}^{(n)}\}_{n=1}^{N}$ is collected by recording the CSI measured at the APs while a UE traverses the deployment area at known positions. The estimator f in (4.2) is parameterized as a learnable function $g_{\boldsymbol{\theta}}\colon \mathbb{C}^{M\times S} \to \mathbb{R}^D$ with parameters $\boldsymbol{\theta}$, typically an NN. Identifying the estimated output with $g_{\boldsymbol{\theta}}(\mathbf{H}^{(n)})$ and the ground-truth label with $\mathbf{x}^{(n)}$, the supervised training objective for fingerprinting becomes a particular case of the MSE regression loss in (3.4), which yields the trained parameters $\hat{\boldsymbol{\theta}}$. In the *online inference phase*, the trained function is applied to each newly observed CSI matrix to produce the position estimate $\hat{\mathbf{x}}^{(n)} = g_{\hat{\boldsymbol{\theta}}}(\mathbf{H}^{(n)})$.

Because fingerprinting directly minimizes the squared positioning error on the training dataset, it usually attains the best positioning accuracy achiev-

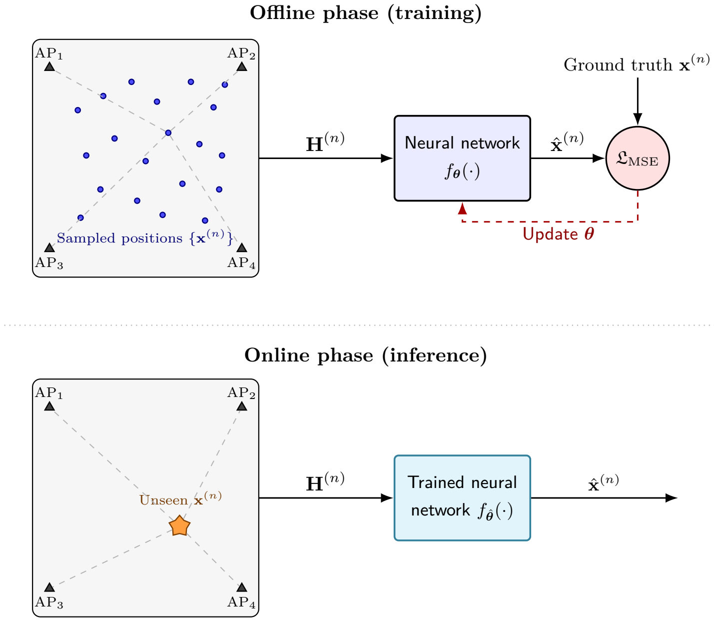


Figure 4.1: Operating principle of CSI fingerprinting, consisting of an offline training phase and an online inference phase.

able from the available labeled data in NLoS scenarios [116], [117], and is therefore commonly used as a benchmark and as an empirical lower bound on the positioning error of data-driven positioning methods. In [Paper D], fingerprinting is trained on the labeled ground-truth UE positions and used as the supervised reference against which the proposed label-free positioning estimators are compared, so that the gap between the two reflects the cost of operating without labeled CSI–position pairs. In addition, since no parametric channel model is assumed, fingerprinting is applicable under NLoS conditions, where the multipath structure induced by reflections and diffractions acts as a discriminative feature rather than as a source of bias, which has made it a popular technique for indoor positioning with WiFi CSI [114], [115].

The main limitation of fingerprinting is the cost of constructing the labeled dataset. Collecting CSI at a sufficiently dense set of reference positions throughout a real deployment area requires an extensive measurement campaign that must be repeated whenever the environment significantly changes, e.g., due to furniture rearrangements or the addition of new obstacles. Furthermore, as the number of labeled samples decreases, the trained function $g_{\hat{\boldsymbol{\theta}}}$ must extrapolate to spatial regions that were poorly covered during training, and the positioning accuracy degrades accordingly. The results reported in [Paper D] illustrate both effects: although fingerprinting attains the lowest positioning error when the entire labeled dataset is used, its accuracy degrades substantially as the fraction of retained labels is reduced, and it is eventually surpassed by self-supervised approaches that do not rely on labels at all. This dependence on labeled data motivates the CC framework described next.

4.4 Channel Charting

CC [118] addresses the positioning problem of Sec. 4.1 in a self-supervised manner: it parameterizes the estimator f in (4.2) as a learnable function $g_{\boldsymbol{\theta}}$ that is trained without any CSI–position pairs. As in fingerprinting, CC exploits the position-dependent structure of $\mathbf{H}^{(n)}$, which makes it applicable to NLoS scenarios where parametric channel models fail. Unlike fingerprinting, however, CC replaces the supervised MSE loss in (3.4) with self-supervised objectives that rely on side information about the relations between observations rather than on absolute position labels. Once trained, the resulting chart can be exploited for downstream applications such as pilot allocation [119], beam management [120], [121], and channel capacity prediction [122].

The general CC pipeline, illustrated in Fig. 4.2, is composed of three building blocks. The first block extracts *large-scale features* from the raw CSI, which ignores, for instance, phase shifts produced by poorly synchronized APs [123]. A typical choice is the amplitude-normalized vectorization $\mathbf{f} = |\mathbf{h}|/\|\mathbf{h}\|_2$ with $\mathbf{h} = \mathrm{vec}(\mathbf{H})$. The second block is a learnable charting function $g_{\boldsymbol{\theta}}$ that maps these features to the low-dimensional position estimates that constitute the channel chart. It is typically implemented as an NN, in contrast to classical non-parametric dimensionality-reduction techniques such as Sammon's mapping, which jointly embed the training samples but do not provide an explicit function evaluable on new CSI [118] and are therefore impractical for online positioning. The third block is a self-supervised loss that encourages the chart to preserve the geometry of the underlying UE positions. For instance, the Siamese loss of [124] passes pairs of CSI samples through two NNs with shared parameters and penalizes the mismatch between the pairwise distance in feature space

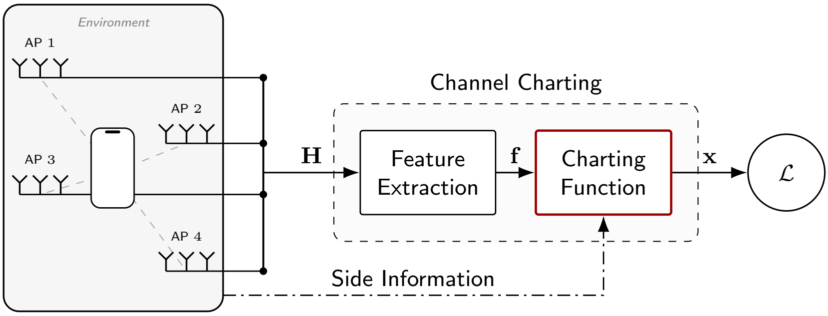
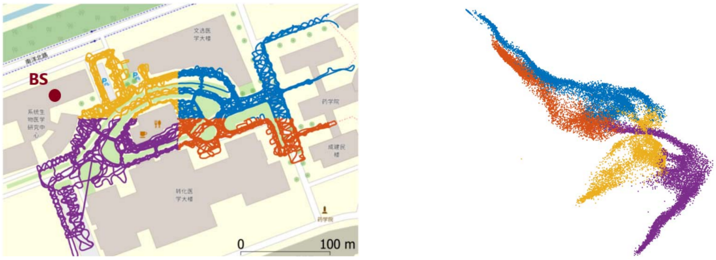


Figure 4.2: General CC block diagram.

and the pairwise distance in chart space, while [125] enforces that pairs of CSI samples whose normalized features are similar up to a phase rotation are mapped to nearby chart points. [Paper D] adopts the timestamp-based triplet loss of [126], which exploits the timestamps of the transmissions as side information.

The central limitation of CC for downstream positioning use is illustrated in Fig. 4.3. Although the local geometry of the chart reflects the relative arrangement of the UEs, the chart itself lies in an arbitrary coordinate system that is not aligned with the physical environment. Some positioning use cases require estimates in a global reference frame, e.g., for navigation [127], beam management [128], or handover decisions tied to physical regions [129]. The chart must therefore be transformed so that the estimated positions $\{\hat{\mathbf{x}}^{(n)}\}_{n=1}^{N}$ lie in the true coordinate system of the deployment, a property we refer to as *global geometry preservation*. The remainder of this section presents two complementary ways of preserving the global geometry: (i) a classical LS alignment based on a small set of anchor measurements collected in the real environment and (ii) a DT-aided approach in which simulated channel observations replace the anchor measurements. The latter underlies the framework proposed in [Paper D].



(a) True positions **(b)** Estimated positions

Figure 4.3: Example of CC using a triplet loss. The figures are based on [126].

4.4.1 Least-Squares Alignment

Arguably the most straightforward way to recover global geometry is to assume a small labeled dataset $\{\mathbf{x}^{(q)}, \mathbf{H}^{(q)}\}_{q=1}^{Q}$, with $Q \ll N$, that relates a few known positions to the corresponding CSI [130]. Passing the labeled CSI through the trained CC function yields a set of chart points $\{\hat{\mathbf{x}}^{(q)}\}_{q=1}^{Q}$ that serve as anchors to fit an affine transformation

$$\hat{\mathbf{A}}, \hat{\mathbf{b}} \;=\; \arg\min_{\mathbf{A},\mathbf{b}} \|\mathbf{A}\hat{\mathbf{X}} + \mathbf{b}\mathbf{1}^{\top} - \mathbf{X}\|_F^2, \tag{4.6}$$

where $\hat{\mathbf{X}}$ and $\mathbf{X}$ stack the estimated and ground-truth positions of the labeled samples, respectively. Every chart point is then mapped to the global coordinate system via $\tilde{\mathbf{x}}^{(n)} = \hat{\mathbf{A}}\,\hat{\mathbf{x}}^{(n)} + \hat{\mathbf{b}}$. An alternative approach avoids any UE measurement campaign by replacing the UE anchors with the APs themselves [131]. In that case, each AP position in the chart is estimated as the centroid of the chart points reporting the highest received power around it, and (4.6) is solved using the known AP coordinates. In both variants, the alignment is constrained to be affine, which limits the achievable accuracy whenever the chart is significantly distorted with respect to the true geometry or when only a few anchors are available. The DT-aided approach described next sidesteps both the affine restriction and the need for real labeled data by replacing the anchors with simulated channel observations.

4.4.2 Digital Twin Alignment

A DT is a virtual representation of a physical environment that mirrors its structure and behavior with sufficient fidelity to enable analysis, simulation, prediction, or decision-making about its physical counterpart [48]. In the context of wireless positioning, a DT models the propagation environment so that it can synthesize CSI for arbitrary TX–RX configurations through electromagnetic simulation, typically ray tracing. For the D-MIMO scenario considered in this thesis, a DT is specified by (i) the geometry and electromagnetic properties of the reflective surfaces, (ii) the position and orientation of the APs, and (iii) the RF parameters of the link. Given a set of predefined query positions $\{\tilde{\mathbf{x}}^{(p)}\}_{p=1}^{P}$ in the digital scenario, the DT returns the corresponding simulated CSI matrices $\{\tilde{\mathbf{H}}^{(p)}\}_{p=1}^{P}$. The pairs $\{\tilde{\mathbf{H}}^{(p)}, \tilde{\mathbf{x}}^{(p)}\}_{p=1}^{P}$ play the same role as the labeled dataset of Sec. 4.4.1, except that they are obtained without any physical measurement campaign and can be regenerated on demand whenever the environment is reconfigured in the digital model.

Existing DT-aided CC methods differ in the granularity of the simulated quantities used to anchor the chart, which trades off DT fidelity against robustness to model mismatch. *CSI-level* approaches treat the DT as a generator of synthetic fingerprints and train the positioning function on $\{\tilde{\mathbf{H}}^{(p)}, \tilde{\mathbf{x}}^{(p)}\}_{p=1}^{P}$ as if it were a labeled real dataset [132]. Their accuracy relies on a faithful reproduction of the small-scale multipath structure, which requires detailed scene modeling and is sensitive to small environmental perturbations such as furniture rearrangements. *Feature-level* approaches instead replace the fine-grained CSI with large-scale features that vary smoothly over space. [Paper D] follows this route by computing large-scale features in parallel from the simulated CSI on the DT side and from the observed CSI at the APs, and by adding an auxiliary loss to the CC loss that enforces consistency between the two. Large-scale features can be obtained from computationally efficient DTs [133] that approximate large-scale propagation quantities without resolving the full multipath response. Since large-scale features carry less information per sample, the auxiliary loss is typically combined with a local-geometry loss such as the timestamp triplet loss of [126]. [Paper D] develops this combination in detail and shows that, for an indoor D-MIMO deployment, the resulting framework yields global positioning accuracy competitive with labeled fingerprinting while remaining robust to mismatches in the DT geometry and to a distribution shift between training and test UE trajectories.

CHAPTER 5

Summary of included papers

This chapter provides a summary of the included papers.

5.1 Paper A

José Miguel Mateos-Ramos, Christian Häger, Musa Furkan Keskin, Luc Le Magoarou, Henk Wymeersch
Model-Based End-to-End Learning for Multi-Target Integrated Sensing and Communication Under Hardware Impairments
IEEE Transactions on Wireless Communications,
vol. 24, no. 3, pp. 2574–2589, Mar. 2025.
©2025 IEEE DOI: 10.1109/TWC.2024.3522667.

This paper studies multi-target ISAC under antenna-array hardware impairments at the transceiver, considering a monostatic OFDM sensing and multiple-input single-output (MISO) communication scenario. Conventional model-based ISAC processing relies on idealized steering-vector models and therefore degrades when residual array impairments break the assumed channel model, affecting both the detection and positioning of multiple targets and the symbol error rate at the communication RX. To address this mismatch, the paper proposes a model-based end-to-end learning framework that jointly optimizes the ISAC beamformer and the sensing RX from data. The key technical enabler is a novel differentiable formulation of the OMP algorithm that is suitable for multi-target angle–range estimation and allows gradients to flow through the RX during training. Building on the differentiable OMP, two model-based parameterization strategies are devised and trained in a supervised manner using true target positions and pilot symbols as labels: (i) learning a dictionary of steering vectors over a discrete angular grid and (ii) directly learning the parameterized hardware impairments while preserving the analytical structure of the steering vectors. The two approaches are compared against a baseline composed of LS beamforming, conventional OMP, and maximum-likelihood symbol detection. The results show that learning the parameterized hardware impairments achieves the highest target-detection probability, the best position-estimation accuracy, and the lowest symbol error rate, while requiring the fewest trainable parameters and generalizing well to angles and ranges unseen during training.

LLM and HW proposed the research topic. All authors contributed significantly to the problem formulation, follow-up of the results, and paper review. JMMR also contributed to the literature review, simulations, and solution development. CH also contributed to designing more realistic simulations.

5.2 Paper B

José Miguel Mateos-Ramos, Baptiste Chatelier, Christian Häger, Musa Furkan Keskin, Luc Le Magoarou, Henk Wymeersch
Semi-Supervised End-to-End Learning for Integrated Sensing and Communications
IEEE International Conference on Machine Learning, Communications, and Networking (ICMLCN),
Stockholm, Sweden, pp. 132–138, 2024.
©2024 IEEE DOI: 10.1109/ICMLCN59089.2024.10624785.

Paper B addresses end-to-end model-based learning for ISAC under antenna-array hardware impairments, with a focus on reducing the dependence on labeled data. The considered scenario is a monostatic ISAC transceiver with an impaired ULA that performs single-target detection and position estimation while simultaneously serving a single-antenna communication user, where the transmit precoder, the sensing RX, and the impairment vector are jointly learned in a differentiable model-based pipeline. Since acquiring ground-truth target positions in real sensing deployments is costly and time-consuming, the paper proposes an unsupervised loss given by the negative peak of the ADM of the received signal. Building on this loss, two semi-supervised strategies are studied that sequentially combine SL and UL, and they are compared against a model-aware baseline with perfect impairment knowledge, fully SL, and purely UL. Results show that SSL initialized with a short supervised stage matches fully supervised performance in both misdetection probability and ISAC trade-off curves while using around 1.2% of the labeled data, whereas relying on the unsupervised loss alone still leaves a gap to fully SL.

HW proposed the research topic. All authors contributed significantly to the problem formulation, follow-up of the results, and paper review. JMMR also contributed to the literature review, simulations, and solution development. BC also contributed to improving the code for the simulation results.

5.3 Paper C

José Miguel Mateos-Ramos, Baptiste Chatelier, Luc Le Magoarou, Nir Shlezinger, Christian Häger, Henk Wymeersch
Unsupervised End-to-End Array Calibration for Multi-Target Integrated Sensing and Communication
submitted to IEEE Transactions on Wireless Communications, Apr. 2026
©IEEE if accepted for publication.

Paper C addresses the joint calibration of the transmit and receive antenna arrays of an ISAC BS that simultaneously senses multiple targets and communicates with a single-antenna UE using OFDM signals. Both arrays are ULAs affected by GPIs and ADIs. The paper proposes an end-to-end MB-ML framework that calibrates these impairments without any labeled data, i.e., without collecting measurements from targets at known positions. At the TX, a differentiable precoder is designed to maximize the array response over the angular sectors of interest, while at the sensing RX target parameters are estimated through the OMP algorithm. Two unsupervised sensing losses are proposed and compared: (i) maximizing the peak of the ADM and (ii) minimizing the norm of the OMP residual once all estimated targets have been removed from the received signal. These are combined with a communication loss that maximizes the estimated signal energy at the UE, yielding an overall ISAC loss that exposes the sensing–communication trade-off through a single weighting hyperparameter. To enable end-to-end training without requiring the (in general unknown) gradient of the channel function, the framework introduces a gradient-free channel backpropagation scheme, applied for the first time in the ISAC context.

Numerical results show that minimizing the OMP residual yields significantly better multi-target position estimation than maximizing the ADM response. The proposed unsupervised framework achieves sensing and communication performance close to a supervised calibration baseline and to a genie system with perfect knowledge of the impairments, and generalizes better than the supervised baseline to operating conditions (such as a different sensing SNR) unseen during training.

JMMR and HW proposed the research topic. All authors contributed significantly to the problem formulation, follow-up of the results, and paper review. JMMR also contributed to the literature review, simulations, and solution development.

5.4 Paper D

José Miguel Mateos-Ramos, Frederik Zumegen, Henk Wymeersch, Christian Häger, Christoph Studer
Positioning via Digital-Twin-Aided Channel Charting with Large-Scale CSI Features
submitted to *IEEE Transactions on Wireless Communications,* Nov. 2025
©IEEE if accepted for publication.

Paper D addresses the problem of self-supervised positioning in an uplink D-MIMO scenario with several UEs and APs. The proposed framework combines CC with a DT of the environment to estimate UE positions in true spatial coordinates, without requiring labeled positions. The main novelty is to align the observed and the simulated data through large-scale channel features rather than through fine-grained CSI. Training combines the timestamp-based CC loss, which preserves local geometry, and a novel DT-aided cosine-similarity loss that matches large-scale features computed from the observed and the simulated CSI. Several large-scale features are studied, including the received power, the angle-power profile, the delay-power profile, the channel covariance magnitude, and a truncated delay profile. A theoretical analysis establishes that, under a matched DT and a sufficiently rich feature, the proposed loss is minimized exactly at the true UE position.

Simulations in an indoor scenario based on ray-traced CSI show that the *Power*, *APP*, and *DPP* features outperform state-of-the-art label-free CC. Additional experiments evaluate the framework under DT modeling mismatches, namely, shifted AP positions, incorrect wall materials, and additional walls unaccounted for in the DT and under distribution shifts at inference time, such as a change in the UE height or the appearance of new reflecting objects in the environment. Under slight DT mismatches, DT-based fingerprinting outperforms the proposed approach by exploiting the full simulated CSI; however, the proposed approach is the most useful option under severe DT mismatches or when only a computationally efficient DT that provides large-scale features such as the received power is available.

CS proposed the research topic. All authors contributed significantly to the follow-up of the results and paper review. JMMR also contributed to the literature review, simulations, and solution development. FZ also contributed to the validation of the results.

CHAPTER 6

Concluding Remarks and Future Work

6.1 Conclusion

In this thesis, we have investigated ML as a data-driven complement to model-based signal processing for ISAC, addressing two fundamental sources of model mismatch that conventional signal processing struggles to handle: hardware impairments at the transceiver and complex propagation conditions such as NLoS environments. We have first shown that even small antenna-array hardware impairments can severely degrade sensing performance and, to a lesser extent, communication performance, and we have proposed a model-based end-to-end learning framework that jointly trains the ISAC TX and the sensing RX to learn parameterized impairments, outperforming both dictionary-based learning and standard model-based calibration while generalizing well to unseen scenarios. Motivated by the practical cost of acquiring labeled target positions, we have subsequently reduced the labeling burden of this supervised paradigm by introducing an unsupervised loss function driven solely by the received signal, together with a semi-supervised strategy that, when initialized from limited-data SL, has attained performance comparable to fully supervised training, although UL alone has not closed the gap. We have then addressed this remaining gap by proposing a gradient-free unsupervised framework that calibrates GPIs and ADIs at both the TX and the sensing RX without any labeled data, performing close to supervised calibration and to a system with perfect impairment knowledge, while exhibiting improved generalization across operating SNRs. Moving beyond the parametric LoS channel models that underlie the previous contributions, we have finally proposed a self-supervised positioning framework based on CC augmented with a DT of the environment, which has preserved both local and global geometry in true spatial coordinates without labeled data and has remained robust under DT modeling mismatches and distribution shifts at inference time. Taken together, these contributions have demonstrated that combining the inductive bias of model-based signal processing with the flexibility of ML enables ISAC systems to operate reliably under hardware impairments and challenging propagation conditions, with progressively reduced dependence on labeled data.

6.2 Future Work

The contributions of this thesis open several research avenues at the intersection of ML and ISAC signal processing, as outlined in the following.

1. **Beyond linear and single-cluster signal models:** Two extensions remain largely open. The first is the optimization of ISAC waveforms in the frequency domain, including power allocation and modulation format across subcarriers, with waterfilling as the communication-optimal limit. The second is the explicit treatment of hardware imperfections on the communication link itself, e.g., at the UE, rather than only at the ISAC BS. A further extension is to consider sensing scenarios with multiple target clusters, where beamforming cannot be reduced to a single angular sector and conventional baselines are not readily available.

2. **Continuous and online learning:** The methods developed in this thesis are trained offline on data collected under fixed operating conditions. Continuous or online training would allow the calibration and positioning functions to adapt to gradual changes in the environment and in the hardware impairments over time, without retraining from scratch whenever the operating conditions drift. Open questions in this direction include how to detect that adaptation is needed, how to balance

plasticity to recent data with stability of previously learned behavior, and how to perform this update with the limited computational and memory budget available at the deployed device.

3. **Tracking and uncertainty quantification:** The proposed frameworks produce point estimates of the quantities of interest and treat each measurement independently. Integration with tracking algorithms can exploit the temporal correlation of consecutive estimates to improve accuracy and smooth out occasional outliers. Equally important, principled uncertainty estimates would allow downstream modules to weigh each measurement according to its reliability, which is essential for sensor fusion, risk-aware control, and detection of anomalous operating conditions.

4. **Real-world experimental validation:** The contributions of this thesis are simulation-based, and a natural next step is to validate them on real hardware and in real environments. Real RF front-ends introduce mutual coupling, nonlinearities, and time-varying drifts that are not captured by the impairment parameterizations studied in this thesis, and real environments exhibit modeling mismatches that are difficult to emulate synthetically. Experimental campaigns would not only validate the proposed methods but also reveal which modeling assumptions need to be adjusted for the implementation of the proposed algorithms.

References

[1] D. T. Seymour and A. Shaheen, "History of wireless communication," *Review of Business Information Systems (RBIS)*, vol. 15, p. 37, Apr. 2011.

[2] Ericsson, "Ericsson mobility report," Tech. Rep., Nov. 2025, Available at: `https://www.ericsson.com/en/reports-and-papers/mobility-report/reports/november-2025`.

[3] X. Lin, J. Zhang, Y. Liu, and J. Kim, *Fundamentals of 6G Communications and Networking*. Springer, 2024.

[4] I. T. Union, "M.2160 : Framework and overall objectives of the future development of IMT for 2030 and beyond," ITU, Recommendation 2160-0, Nov. 2023.

[5] E. Semaan, E. Tejedor, R. K. Kochhar, S. Magnusson, and S. Parkvall, "6g spectrum – enabling the future mobile life beyond 2030," Ericsson, White Paper, Aug. 2024.

[6] Y. Cui, F. Liu, X. Jing, and J. Mu, "Integrating sensing and communications for ubiquitous IoT: Applications, trends, and challenges," *IEEE Network*, vol. 35, no. 5, pp. 158–167, 2021.

[7] N. Jabeen, H. Lei, A. Muhammad, A. Ali, Z. U. Khan, and G. Pan, "Localization in ISAC: A review," *IEEE Internet of Things Journal*, vol. 12, no. 22, pp. 46 526–46 552, 2025.

[8] 3rd Generation Partnership Project (3GPP), "TR 22.837: Study on Integrated Sensing and Communication," 3GPP TSG SA, Technical Report 22.837, 2023.

[9] X. Luo, Q. Lin, R. Zhang, H.-H. Chen, X. Wang, and M. Huang, "ISAC—A survey on its layered architecture, technologies, standardizations, prototypes, and testbeds," *IEEE Communications Surveys & Tutorials*, vol. 28, pp. 485–526, 2026.

[10] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proceedings of the IEEE*, vol. 109, no. 7, pp. 1166–1199, 2021.

[11] H. Wymeersch et al., "Cross-layer integrated sensing and communication: A joint industrial and academic perspective," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 6966–7015, 2025.

[12] R. Liu, M. Li, Q. Liu, and A. L. Swindlehurst, "Dual-functional radar-communication waveform design: A symbol-level precoding approach," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 6, pp. 1316–1331, 2021.

[13] O. Li et al., "Integrated sensing and communication in 6g a prototype of high resolution thz sensing on portable device," in *2021 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, 2021, pp. 544–549.

[14] H. Hua, J. Xu, and T. X. Han, "Optimal transmit beamforming for integrated sensing and communication," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 8, pp. 10 588–10 603, 2023.

[15] P. Tosi, M. Henninger, L. G. de Oliveira, and S. Mandelli, "Feasibility of non-line-of-sight integrated sensing and communication at mmwave," in *International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2024, pp. 331–335.

[16] Z. Li, F. Hu, Z. Ling, S. Song, and Q. Li, "Joint non-line-of-sight predictive beamforming and power allocation for isac-assisted vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 10, pp. 16 481–16 495, 2025.

[17] Y. Niu, Z. Wei, L. Wang, H. Wu, and Z. Feng, "Interference management for integrated sensing and communication systems: A survey," *IEEE Internet of Things Journal*, vol. 12, no. 7, pp. 8110–8134, 2025.

[18] A. Venus, E. Leitinger, S. Tertinek, and K. Witrisal, "A neural-enhanced factor graph-based algorithm for robust positioning in obstructed los situations," *IEEE Open Journal of Signal Processing*, vol. 5, pp. 29–38, 2024.

[19] A. Radford, K. Narasimhan, T. Salimans, I. Sutskever, et al., "Improving language understanding by generative pre-training," 2018.

[20] I. J. Goodfellow et al., "Generative adversarial nets," in *Advances in Neural Information Processing Systems*, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Weinberger, Eds., vol. 27, Curran Associates, Inc., 2014.

[21] J. Jumper et al., "Highly accurate protein structure prediction with alphafold," *Nature*, vol. 596, no. 7873, pp. 583–589, 2021.

[22] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, Eds., vol. 25, Curran Associates, Inc., 2012.

[23] D. Silver et al., "A general reinforcement learning algorithm that masters chess, shogi, and go through self-play," *Science*, vol. 362, no. 6419, pp. 1140–1144, 2018.

[24] H. Chen et al., "Modeling and analysis of ofdm-based 5G/6G localization under hardware impairments," *IEEE Transactions on Wireless Communications*, vol. 23, no. 7, pp. 7319–7333, 2024.

[25] U. Gustavsson et al., "On the impact of hardware impairments on massive MIMO," in *IEEE Globecom Workshops (GC Wkshps)*, 2014, pp. 294–300.

[26] K. Sankhe et al., "No radio left behind: Radio fingerprinting through deep learning of physical-layer hardware impairments," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 1, pp. 165–178, 2020.

[27] C.-Y. Wu, Y.-K. Lin, C.-K. Wu, and C.-H. Lee, “Deep learning-based end-to-end design for OFDM systems with hardware impairments,” *IEEE Open Journal of the Communications Society*, vol. 4, pp. 2468–2482, 2023.

[28] A. Mobasher, M. Taherzadeh, R. Sotirov, and A. K. Khandani, “A near-maximum-likelihood decoding algorithm for MIMO systems based on semi-definite programming,” *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 3869–3886, 2007.

[29] M. Mirabella, P. D. Viesti, A. Davoli, and G. M. Vitetta, “An approximate maximum likelihood method for the joint estimation of range and doppler of multiple targets in ofdm-based radar systems,” *IEEE Transactions on Communications*, vol. 71, no. 8, pp. 4862–4876, 2023.

[30] R. Livni, S. Shalev-Shwartz, and O. Shamir, “On the computational efficiency of training neural networks,” in *Advances in Neural Information Processing Systems*, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Weinberger, Eds., vol. 27, Curran Associates, Inc., 2014.

[31] P. A. Merolla et al., “A million spiking-neuron integrated circuit with a scalable communication network and interface,” *Science*, vol. 345, no. 6197, pp. 668–673, 2014.

[32] Y.-H. Chen, T. Krishna, J. S. Emer, and V. Sze, “Eyeriss: An energy-efficient reconfigurable accelerator for deep convolutional neural networks,” *IEEE Journal of Solid-State Circuits*, vol. 52, no. 1, pp. 127–138, 2017.

[33] M. A. Alsheikh, D. Niyato, S. Lin, H.-p. Tan, and Z. Han, “Mobile big data analytics using deep learning and apache spark,” *IEEE Network*, vol. 30, no. 3, pp. 22–29, 2016.

[34] M. Paolini and S. Fili, “Mastering analytics: How to benefit from big data and network complexity: An analyst report,” *RCR Wireless News*, vol. 20, no. 80, p. 120, 2017.

[35] C. Zhang, P. Patras, and H. Haddadi, “Deep learning in mobile and wireless networking: A survey,” *IEEE Commun. Surveys Tutorials*, vol. 21, no. 3, pp. 2224–2287, 2019.

[36] A. Djuhera et al., “Against the monolithic wireless world model: Why NextG needs composable and agentic intelligence,” 2026.

[37] T. O’Shea and J. Hoydis, “An introduction to deep learning for the physical layer,” *IEEE Trans. Cognitive Commun. Networking*, vol. 3, no. 4, pp. 563–575, 2017.

[38] Z. Qin, H. Ye, G. Y. Li, and B.-H. F. Juang, “Deep learning in physical layer communications,” *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 93–99, 2019.

[39] L. Dai, R. Jiao, F. Adachi, H. V. Poor, and L. Hanzo, “Deep learning for wireless communications: An emerging interdisciplinary paradigm,” *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 133–139, 2020.

[40] H. Kim, S. Oh, and P. Viswanath, “Physical layer communication via deep learning,” *IEEE J. Selected Areas Information Theory*, vol. 1, no. 1, pp. 5–18, 2020.

[41] ITU-R, “Framework and overall objectives of the future development of IMT for 2030 and beyond,” Int. Telecommun. Union, Geneva, Switzerland, Tech. Rep., Nov. 2023.

[42] F. Liu et al., “Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond,” *IEEE J. Selected Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, 2022.

[43] M. Matthaiou, O. Yurduseven, H. Q. Ngo, D. Morales-Jimenez, S. L. Cotton, and V. F. Fusco, “The road to 6G: Ten physical layer challenges for communications engineers,” *IEEE Commun. Mag.*, vol. 59, no. 1, pp. 64–69, Feb. 2021.

[44] F. Zhang et al., “Near-field ISAC for THz wireless systems,” *IEEE Network*, vol. 39, no. 6, pp. 54–61, 2025.

[45] Y. Ge et al., “Pilot-based end-to-end radio positioning and mapping for ISAC: Beyond point-based landmarks,” *preprint arXiv:2505.07402*, 2025.

[46] J. Yang, C.-K. Wen, X. Yang, J. Xu, T. Du, and S. Jin, “Multi-domain cooperative slam: The enabler for integrated sensing and communications,” *IEEE Wireless Communications*, vol. 30, no. 1, pp. 40–49, 2023.

[47] C.-X. Wang et al., "On the road to 6G: Visions, requirements, key technologies and testbeds," *IEEE Commun. Surveys & Tutorials*, vol. 25, no. 2, pp. 905–974, Feb. 2023.

[48] F. Tao, H. Zhang, A. Liu, and A. Y. C. Nee, "Digital twin in industry: State-of-the-art," *IEEE Trans. Ind. Inform.*, vol. 15, no. 4, pp. 2405–2415, Apr. 2019.

[49] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, "Characterising the digital twin: A systematic literature review," *CIRP Journal of Manufacturing Science and Technology*, vol. 29, pp. 36–52, 2020, ISSN: 1755-5817.

[50] N. Su, F. Liu, and C. Masouros, "Sensing-assisted eavesdropper estimation: An ISAC breakthrough in physical layer security," *IEEE Transactions on Wireless Communications*, vol. 23, no. 4, pp. 3162–3174, 2024.

[51] T. Matsumine, H. Ochiai, and J. Shikata, "Physical layer security for integrated sensing and communication: A survey," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 6690–6743, 2025.

[52] R. M. Mealey, "A method for calculating error probabilities in a radar communication system," *IEEE Trans. Space Electronics Telemetry*, vol. 9, no. 2, pp. 37–42, 1963.

[53] P. Hughes and J. Choe, "Overview of advanced multifunction rf system (AMRFS)," in *Proc. IEEE Int. Conf. Phased Array Systems Techn.*, Dana Point, CA, USA, 2000, pp. 21–24.

[54] G. Tavik et al., "The advanced multifunction rf concept," *IEEE Trans. Microwave Theory Techniques*, vol. 53, no. 3, pp. 1009–1020, 2005.

[55] J. A. Molnar, I. Corretjer, and G. Tavik, "Integrated topside - integration of narrowband and wideband array antennas for shipboard communications," in *Proc. Military Commun. Conf. (MILCOM)*, Baltimore, MD, USA, 2011, pp. 1802–1807.

[56] 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, 2011.

[57] J. G. Andrews et al., "What will 5g be?" *IEEE J. Selected Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.

[58] T. S. Rappaport et al., “Millimeter wave mobile communications for 5G cellular: It Will Work!” *IEEE Access*, vol. 1, pp. 335–349, 2013.

[59] Y. L. Sit, C. Sturm, L. Reichardt, T. Zwick, and W. Wiesbeck, “The ofdm joint radar-communication system: An overview,” in *Proc. Int. Conf. Advances Satellite Space Communications (SPACOMM)*, Budapest, Hungary, 2011, pp. 69–74.

[60] C. Baquero Barneto et al., “Full-duplex OFDM radar with LTE and 5G NR waveforms: Challenges, solutions, and measurements,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 10, pp. 4042–4054, 2019.

[61] R. J. Sullivan, *Radar foundations for imaging and advanced concepts.* IET, 2004, vol. 2.

[62] W. Saad, M. Bennis, and M. Chen, “A vision of 6G wireless systems: Applications, trends, technologies, and open research problems,” *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, Oct. 2019.

[63] 3rd Generation Partnership Project (3GPP), “TS 22.137: Integrated Sensing and Communication,” 3GPP TSG SA, Technical Specification 22.137, 2023.

[64] D. K. P. Tan et al., “Integrated sensing and communication in 6G: Motivations, use cases, requirements, challenges and future directions,” in *Proc. 1st IEEE Int. Symp. Joint Commun. & Sens. (JC&S)*, Dresden, Germany, 2021, pp. 1–6.

[65] H. Wymeersch et al., “Integration of communication and sensing in 6G: A joint industrial and academic perspective,” in *Proc. 32nd IEEE Annu. Int. Symp. Personal Indoor Mobile Radio Commun. (PIMRC)*, Helsinki, Finland, 2021, pp. 1–7.

[66] H. Wymeersch et al., “Joint communication and sensing for 6G - a cross-layer perspective,” in *IEEE Int. Symposium Joint Commun. Sensing (JC&S)*, Leuven, Belgium, 2024, pp. 01–06.

[67] L. Zheng, M. Lops, Y. C. Eldar, and X. Wang, “Radar and communication coexistence: An overview: A review of recent methods,” *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 85–99, 2019.

[68] A. Aubry, A. De Maio, M. Piezzo, and A. Farina, “Radar waveform design in a spectrally crowded environment via nonconvex quadratic optimization,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 2, pp. 1138–1152, 2014.

[69] K. Wu, Y. Bigdeli, S. A. Keivaan, J. Deng, and P. Burasa, “Integrated sensing and communication (ISAC) transceiver: Hardware architectures, enabling technologies, and emerging trends,” *IEEE J. Selected Topics Electromagnetics, Antennas Propagation*, vol. 1, no. 1, pp. 37–64, 2025.

[70] W. Zhou, R. Zhang, G. Chen, and W. Wu, “Integrated sensing and communication waveform design: A survey,” *IEEE Open J. Commun. Society*, vol. 3, pp. 1930–1949, 2022.

[71] M. F. Keskin, V. Koivunen, and H. Wymeersch, “Limited feedforward waveform design for ofdm dual-functional radar-communications,” *IEEE Trans. Signal Processing*, vol. 69, pp. 2955–2970, 2021.

[72] Y. Cui, F. Liu, C. Masouros, J. Xu, T. X. Han, and Y. C. Eldar, “Integrated sensing and communications: Background and applications,” in *Integrated Sensing and Communications*, F. Liu, C. Masouros, and Y. C. Eldar, Eds. Singapore: Springer Nature Singapore, 2023, pp. 3–21, ISBN: 978-981-99-2501-8.

[73] D. Ma, N. Shlezinger, T. Huang, Y. Liu, and Y. C. Eldar, “FRaC: FMCW-based joint radar-communications system via index modulation,” *IEEE J. Selected Topics Signal Processing*, vol. 15, no. 6, pp. 1348–1364, 2021.

[74] L. Leyva, D. Castanheira, A. Silva, and A. Gameiro, “Hybrid beamforming design for communication-centric ISAC,” *IEEE Sensors J.*, vol. 24, no. 13, pp. 21 179–21 190, 2024.

[75] S. Naoumi, A. Bazzi, R. Bomfin, and M. Chafii, “High-resolution sensing in communication-centric ISAC: Deep learning and parametric methods,” *IEEE Journal on Selected Areas in Communications*, vol. 44, pp. 2201–2216, 2026.

[76] Y. Cui, F. Liu, W. Yuan, J. Mu, X. Jing, and D. W. Kwan Ng, “Optimal precoding design for monostatic ISAC systems: MSE lower bound and DoF completion,” in *IEEE Wireless Commun. Networking Conf. (WCNC)*, 2024, pp. 1–6.

[77] X. Liu, T. Huang, N. Shlezinger, Y. Liu, J. Zhou, and Y. C. Eldar, “Joint transmit beamforming for multiuser MIMO communications and MIMO radar,” *IEEE Trans. Signal Processing*, vol. 68, pp. 3929–3944, 2020.

[78] F. Liu, Y.-F. Liu, A. Li, C. Masouros, and Y. C. Eldar, “Cramér-rao bound optimization for joint radar-communication beamforming,” *IEEE Trans. Signal Processing*, vol. 70, pp. 240–253, 2022.

[79] M. F. Keskin, H. Wymeersch, and V. Koivunen, “MIMO-OFDM joint radar-communications: Is ICI friend or foe?” *IEEE J. Selected Topics Signal Process.*, vol. 15, no. 6, pp. 1393–1408, 2021.

[80] S. Mallat and Z. Zhang, “Matching pursuits with time-frequency dictionaries,” *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993, ISSN: 1053-587X.

[81] J. Lee, G.-T. Gil, and Y. H. Lee, “Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications,” *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2370–2386, Apr. 2016.

[82] T. Schenk, *RF imperfections in high-rate wireless systems: impact and digital compensation.* Springer Science & Business Media, 2008.

[83] V. Koivunen, M. F. Keskin, H. Wymeersch, M. Valkama, and N. González-Prelcic, “Multicarrier ISAC: Advances in waveform design, signal processing, and learning under nonidealities,” *IEEE Signal Process. Magazine*, vol. 41, no. 5, pp. 17–30, 2024.

[84] Z.-M. Liu and Y.-Y. Zhou, “A unified framework and sparse bayesian perspective for direction-of-arrival estimation in the presence of array imperfections,” *IEEE Trans. Signal Process.*, vol. 61, no. 15, pp. 3786–3798, 2013.

[85] P. Chen, Z. Chen, Z. Cao, and X. Wang, “A new atomic norm for DOA estimation with gain-phase errors,” *IEEE Trans. Signal Process.*, vol. 68, pp. 4293–4306, 2020.

[86] X.-Y. Wang, X.-P. Li, H. Huang, and H. C. So, “Robust DOA estimation with distorted sensors,” *IEEE Trans. Aerospace Electronic Systems*, vol. 60, no. 5, pp. 5730–5741, 2024.

[87] Y. Wang, L. Wang, J. Xie, M. Trinkle, and B. W.-H. Ng, “DOA estimation under mutual coupling of uniform linear arrays using sparse reconstruction,” *IEEE Wireless Commun. Letters*, vol. 8, no. 4, pp. 1004–1007, 2019.

[88] A. Lindholm, N. Wahlström, F. Lindsten, and T. B. Schön, *Machine Learning - A First Course for Engineers and Scientists*. Cambridge University Press, 2022.

[89] V. J. Reddi, *Introduction to Machine Learning Systems*. MIT Press, 2026.

[90] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[91] B. Kleinberg, Y. Li, and Y. Yuan, “An alternative view: When does SGD escape local minima?” In *Proceedings of the 35th International Conference on Machine Learning*, J. Dy and A. Krause, Eds., ser. Proceedings of Machine Learning Research, vol. 80, PMLR, Oct. 2018, pp. 2698–2707.

[92] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” *nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[93] M. Sattari, H. Guo, D. Gündüz, A. Panahi, and T. Svensson, “Full-duplex millimeter wave MIMO channel estimation: A neural network approach,” *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 2, pp. 1093–1108, 2024.

[94] W. Zhang, M. Feng, M. Krunz, and A. Hossein Yazdani Abyaneh, “Signal detection and classification in shared spectrum: A deep learning approach,” in *IEEE Conference on Computer Communications*, 2021, pp. 1–10.

[95] H. Huang, Y. Peng, J. Yang, W. Xia, and G. Gui, “Fast beamforming design via deep learning,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 1, pp. 1065–1069, 2020.

[96] F. A. Aoudia and J. Hoydis, “End-to-end learning for OFDM: From neural receivers to pilotless communication,” *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1049–1063, Aug. 2021.

[97] W. Jiang, A. M. Haimovich, and O. Simeone, "Joint design of radar waveform and detector via end-to-end learning with waveform constraints," *IEEE Trans. Aerospace Electron. Syst.*, vol. 58, no. 1, pp. 552–567, Aug. 2021.

[98] F. A. Aoudia and J. Hoydis, "Model-free training of end-to-end communication systems," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2503–2516, 2019.

[99] H. Ye, G. Y. Li, B.-H. F. Juang, and K. Sivanesan, "Channel agnostic end-to-end learning based communication systems with conditional GAN," in *IEEE Globecom Workshops (GC Wkshps)*, 2018, pp. 1–5.

[100] S. J. Prince, *Understanding Deep Learning*. MIT Press, 2023, `https://udlbook.github.io/udlbook/`.

[101] A. S. Rahmathullah, Á. F. García-Fernández, and L. Svensson, "Generalized optimal sub-pattern assignment metric," in *Proc. 20th IEEE Int. Conf. Inform. Fusion (Fusion)*, Xi'an, China, 2017, pp. 1–8.

[102] C. M. Bishop and H. Bishop, *Deep Learning: Foundations and Concepts*. Springer, 2024, `https://www.bishopbook.com/`.

[103] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, `http://www.deeplearningbook.org`.

[104] M. A. Nielsen, *Neural Networks and Deep Learning*. Determination Press, 2015, `http://neuralnetworksanddeeplearning.com/`.

[105] K. P. Murphy, *Probabilistic Machine Learning: An introduction*. MIT Press, 2022.

[106] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, G. Gordon, D. Dunson, and M. Dudík, Eds., ser. Proceedings of Machine Learning Research, vol. 15, Fort Lauderdale, FL, USA: PMLR, Nov. 2011, pp. 315–323.

[107] N. Shlezinger, Y. C. Eldar, and S. P. Boyd, "Model-based deep learning: On the intersection of deep learning and optimization," *IEEE Access*, vol. 10, pp. 115 384–115 398, 2022.

[108] N. Shlezinger, J. Whang, Y. C. Eldar, and A. G. Dimakis, "Model-based deep learning," *Proc. IEEE*, vol. 111, no. 5, pp. 465–499, 2023.

[109] M. F. Keskin, H. Wymeersch, and V. Koivunen, "MIMO-OFDM joint radar-communications: Is ICI friend or foe?" *IEEE J. of Select. Topics Signal Process.*, vol. 15, no. 6, pp. 1393–1408, Sep. 2021.

[110] E. Tsakalaki and J. Schäfer, "On application of the correlation vectors subspace method for 2-dimensional angle-delay estimation in multipath ofdm channels," in *2018 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2018, pp. 1–8.

[111] W. Xu, Z. Wang, and S. A. Zekavat, "An introduction to nlos identification and localization," *Handbook of Position Location: Theory, Practice, and Advances*, pp. 523–555, 2011.

[112] M. Ahadi, O. Esrafilian, F. Kaltenberger, and A. Malik, "TDoA-based self-supervised channel charting with nlos mitigation," *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 4, pp. 780–793, 2026.

[113] N. Garcia, H. Wymeersch, E. G. Larsson, A. M. Haimovich, and M. Coulon, "Direct localization for massive MIMO," *IEEE Transactions on Signal Processing*, vol. 65, no. 10, pp. 2475–2487, 2017.

[114] S. He and S.-H. G. Chan, "Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons," *IEEE Commun. Surveys Tut.*, vol. 18, no. 1, pp. 466–490, 2016.

[115] 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.

[116] J. Vieira, E. Leitinger, M. Sarajlic, X. Li, and F. Tufvesson, "Deep convolutional neural networks for massive mimo fingerprint-based positioning," in *Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017, pp. 1–6.

[117] A. Kirmaz, T. Şahin, D. S. Michalopoulos, and W. Gerstacker, "Time-based vs. fingerprinting-based positioning using artificial neural networks," in *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2023, pp. 1–6.

[118] C. Studer, S. Medjkouh, E. Gonultaş, T. Goldstein, and O. Tirkkonen, “Channel charting: Locating users within the radio environment using channel state information,” *IEEE Access*, vol. 6, pp. 47 682–47 698, Aug. 2018.

[119] L. Ribeiro, M. Leinonen, H. Al-Tous, O. Tirkkonen, and M. Juntti, “Channel charting aided pilot reuse for massive mimo systems with spatially correlated channels,” *IEEE Open Journal of the Communications Society*, vol. 3, pp. 2390–2406, 2022.

[120] L. L. Magoarou, T. Yassine, S. Paquelet, and M. Crussière, “Channel charting based beamforming,” in *2022 56th Asilomar Conference on Signals, Systems, and Computers*, 2022, pp. 1185–1189.

[121] P. Kazemi, H. Al-Tous, C. Studer, and O. Tirkkonen, “Channel charting assisted beam tracking,” in *IEEE Vehicular Technology Conference: (VTC2022-Spring)*, 2022, pp. 1–5.

[122] T. Kallehauge, A. E. Kalør, P. Ramírez-Espinosa, M. Guillaud, and P. Popovski, “Delivering ultra-reliable low-latency communications via statistical radio maps,” *IEEE Wireless Communications*, vol. 30, no. 2, pp. 14–20, 2023.

[123] S. Taner, V. Palhares, and C. Studer, “Channel charting in real-world coordinates with distributed MIMO,” *IEEE Trans. Wireless Commun.*, vol. 24, no. 9, pp. 7286–7300, Sep. 2025.

[124] E. Lei, O. Castañeda, O. Tirkkonen, T. Goldstein, and C. Studer, “Siamese neural networks for wireless positioning and channel charting,” in *57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2019, pp. 200–207.

[125] L. Le Magoarou, “Efficient channel charting via phase-insensitive distance computation,” *IEEE Wireless Communications Letters*, vol. 10, no. 12, pp. 2634–2638, 2021.

[126] P. Ferrand, A. Decurninge, L. G. Ordoñez, and M. Guillaud, “Triplet-based wireless channel charting: Architecture and experiments,” *IEEE J. Selected Areas Commun.*, vol. 39, no. 8, pp. 2361–2373, Aug. 2021.

[127] K. Muthineni, A. Artemenko, J. Vidal, and M. Nájar, "A survey of 5G-based positioning for industry 4.0: State of the art and enhanced techniques," in *Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, 2023, pp. 120–125.

[128] M. Arvinte, M. Tavares, and D. Samardzija, "Beam management in 5G NR using geolocation side information," in *Annual Conference on Information Sciences and Systems (CISS)*, 2019, pp. 1–6.

[129] A. I. Mbulwa, H. Tung Yew, A. Chekima, and J. A. Dargham, "Handover optimization framework for next-generation wireless networks: 5G, 5G-Advanced and 6G," in *IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)*, 2024, pp. 409–414.

[130] M. Stahlke, G. Yammine, T. Feigl, B. M. Eskofier, and C. Mutschler, "Indoor localization with robust global channel charting: A time-distance-based approach," *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 1, pp. 3–17, 2023.

[131] J. Pihlajasalo, M. Koivisto, J. Talvitie, S. Ali-Löytty, and M. Valkama, "Absolute positioning with unsupervised multipoint channel charting for 5G networks," in *IEEE Vehicular Techn. Conf.*, 2020, pp. 1–5.

[132] L. Cazzella, F. Linsalata, M. Maleki, D. Badini, M. Matteucci, and U. Spagnolini, "Chartwin: A case study on channel charting-aided localization in dynamic digital network twins," *arXiv preprint*, 2025, arXiv:2508.09055.

[133] F. A. Aoudia, J. Hoydis, and A. Keller, "Fast and differentiable radio maps with NVIDIA instant RM," NVIDIA, Tech. Rep., 2024, Available at: `https://developer.nvidia.com/blog/fast-and-differentiable-radio-maps-with-nvidia-instant-rm/`.