# Integrated Sensing and Communication for Next generation Wireless Systems

by

Rifat Zabin
ID: 24266042

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
M.Sc. in Computer Science

Department of Computer Science and Engineering
BRAC University
January 2026

# Declaration

It is hereby declared that:

1. This thesis is my own original work carried out in partial fulfillment of the requirements for the degree at BRAC University.

2. The thesis does not contain material previously published or written by another person, except where such material has been appropriately cited and acknowledged.

3. Portions of this thesis have been submitted for publication in *IEEE Wireless Communication Letters* and are currently under review. These materials have not been accepted for publication at the time of submission of this thesis.

4. The thesis does not contain material that has been accepted for, or submitted toward, any other degree or academic qualification at any university or institution.

5. All sources of information and assistance have been properly acknowledged.

**Student's Full Name & Signature:**

**Rifat Zabin**
ID: 24266042
M.Sc Student
Dept of Computer Science and Engineering
BRAC University

# Approval

The thesis titled "Integrated Sensing and Communication for Next Generation wireless Systems"

**Submitted by**

Rifat Zabin (ID: 24266042)

Of Fall, 2025 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of M.Sc. in Computer Science on January 17, 2026.

**Thesis Examination Board:**

External Examiner:

**Dr. Kazi Ashrafuzzaman**
Professor, Department of Computer Science and Engineering
University of Chittagong
BSc in CSE, BUET
PhD in ECE, University of Calgary, Canada
E-mail: ashraf@cu.ac.bd

Internal Examiner:

**Dr. Md. Ashraful Alam**
Associate Professor
Department of Computer Science and Engineering
BRAC University
E-mail: ashraful.alam@bracu.ac.bd

Supervisor:

**Dr. Md. Golam Rabium Alam**
Professor, Department of Computer Science and Engineering
Brac University
E-mail: rabiul.alam@bracu.ac.bd

Program Coordinator:

**Dr. Md Sadek Ferdous**
Professor, Department of Computer Science and Engineering
Brac University
E-mail: sadek.ferdous@bracu.ac.bd

Chairperson:

**Sadia Hamid Kazi, Ph.D.**
Associate Professor, Department of Computer Science and Engineering
Brac University
E-mail: skazi@bracu.ac.bd

# Ethics Statement

This thesis investigates wireless sensing and communication using physical-layer channel measurements. The research does not involve invasive sensing, medical procedures, or the collection of personally identifiable information. Human activity data, where used, was collected with informed consent and anonymized prior to analysis. No visual, audio, or biometric data was recorded.

Public datasets were used in accordance with their stated usage policies. The research poses no risk to participants, and no formal ethical approval was required under institutional guidelines.

# Abstract

Wireless communication systems inherently interact with their surrounding physical environment, causing transmitted signals to carry information beyond the data they are intended to convey. This observation forms the basis of Integrated Sensing and Communication (ISAC), where sensing and communication are jointly supported using the same wireless infrastructure. Despite its promise, realizing ISAC in practical systems—particularly with commodity Wi-Fi—remains challenging due to high-dimensional channel representations, compute intensive channel sounding, hardware imperfection, and strong variability across environments. This thesis addresses these challenges by developing practical, learning-driven ISAC frameworks that explicitly exploit the structural properties of wireless channels across time, frequency, and space. The thesis is organized around two complementary thrusts. First, communication-aided sensing is realized through PULSE , a lightweight framework that transforms raw Channel Frequency Response (CFR) measurements into compact, physics-aware temporal representations for sensing inference and keeping it generalized across domains. Second, sensing-aided communication is enabled through ELF , a scalable channel feedback framework that reformulates multi-antenna channel reporting as a structural inference problem, by feeding back only a small number of representative subchannel embeddings. Experiments show that PULSE achieves over **99% sensing accuracy** across diverse activities while reducing the effective input dimensionality by approximately **85%** and maintaining low inference latency suitable for real-time edge deployment. PULSE generalizes to unseen environments and devices using as little as **5 seconds** worth of labeled data, outperforming state-of-the-art Wi-Fi sensing frameworks under domain shifts. On the other hand ELF reduces feedback overhead by up to **96%** relative to dense subcarrier reporting, while preserving near-identical communication reliability. Compared to standard IEEE 802.11ax explicit feedback, ELF achieves up to a **25× reduction** in feedback size across wide bandwidths, and its feedback cost remains effectively **independent of antenna count**, enabling scalable operation in large Multiple-Input Multiple-Output (MIMO) systems. Extensive simulations and real-world Wi-Fi testbed evaluations demonstrate that the proposed approaches achieve strong sensing accuracy, robust cross-domain generalization, and substantial reductions in communication overhead. Collectively, this thesis establishes a unified and deployable ISAC framework for commodity Wi-Fi systems, highlighting the role of channel structure in enabling efficient and reliable sensing and communication under real-world constraints.

**Keywords:** Channel Frequency Response (CFR); Channel State Information (CSI); Wi-Fi Sensing; Temporal Channel Evolution; Communication-Aided Sensing; Sensing-Aided Communication; Channel Sounding; Neural Channel Estimation; CFR Reconstruction; Multi-Band Channel Fusion; Activity Recognition; Environment Awareness; Machine Learning for Wireless Systems; Wi-Fi 6/6E; Next-Generation Wireless Networks.

# Acknowledgement

I would like to express my sincere gratitude to my advisor, Prof. Golam Rabiul Alam, for his continuous guidance, encouragement, and insightful discussions throughout the course of this research. His mentorship has been instrumental in shaping my understanding of wireless systems research and in guiding this thesis toward both technical depth and practical relevance.

I am also grateful to my thesis committee members for their valuable feedback, thoughtful suggestions, and time invested in reviewing this work. Their comments helped improve the clarity, rigor, and overall quality of the thesis. I am thankful to BRAC University for providing the academic resources, research facilities, and institutional support necessary to carry out this work.

Finally, I wish to express my deepest gratitude to the Almighty, and my family for their unwavering support, patience, and encouragement throughout this journey.

# Table of Contents

# List of Figures

# List of Tables

# Acronyms

**3GPP** 3rd Generation Partnership Project. 14, 54, 55, 68

**5G** Fifth-Generation. 3

**6G** Sixth-Generation. 3, 4

**BER** Bit Error Rate. x, 14, 44, 49, 54, 57, 62–64, 66, 68

**BS** Base Station. 13, 44, 46–50, 52, 55, 56, 66

**CDL** Clustered Delay Line. 14, 55, 57

**CFR** Channel Frequency Response. v, x, 9, 10, 12, 13, 15–18, 20–24, 26–35, 37–39, 42–45, 47–49, 53–59, 62, 67

**CNN** Convolutional Neural Network. 31, 34, 37, 38

**CSI** Channel State Information. 9, 10, 15–17, 24, 25, 28–30

**FFT** Fast Fourier Transform. 55, 57

**IoT** Internet of Things. 2, 3

**ISAC** Integrated Sensing and Communication. v, x, 2–12, 14, 15, 24, 27, 69

**LTF** Long Training Field. 21, 33

**MIMO** Multiple-Input Multiple-Output. v, x, 3, 6, 12–14, 17, 24, 25, 42, 44, 45, 47–49, 54, 55, 64, 65, 68

**mu-MIMO** Multi-user MIMO. 7, 8, 10, 18, 25, 26, 42

**NDP** Null Data Packet. 18, 21, 33

**OFDM** Orthogonal Frequency Division Multiplexing. 17, 20, 45, 49, 54, 55

**OFDMA** Orthogonal Frequency Division Multiple Access. 18

**RF** Radio Frequency. 14, 30

**RUs** Resource Units. 18

**SISO** Single Input Single Output. 17

**SNR** Signal to Noise Ratio. x, 62, 63

**SVD** Singular Value Decomposition. 21, 45

**UE** User Equipment. 13, 44, 47, 48, 50, 51, 53, 55, 56, 66

# Chapter 1

# Introduction

Wireless communication has become an essential part of modern life[1]. Today, wireless technologies—particularly Wi-Fi—form the backbone of indoor connectivity in homes, offices, educational institutions, and public spaces. The widespread availability of wireless infrastructure has enabled continuous internet access for billions of devices, supporting applications ranging from personal communication to enterprise networking and smart environments[1]. As wireless networks continue to expand in scale and capability, their role is no longer limited to simply delivering data between devices. Traditionally, wireless systems have been designed with a single goal: reliable and efficient data transmission. From this perspective, the wireless channel is viewed as a medium that introduces distortion due to multipath propagation, fading, mobility, and hardware imperfections[2][3]. These effects are treated as impairments that must be estimated and compensated through signal processing techniques such as channel estimation, equalization, and error correction.

However, extensive experimental evidence has shown that these channel variations are not random artifacts alone. Instead, they are strongly influenced by the physical environment in which communication takes place[4]. Human motion, object placement, and environmental geometry leave signatures on wireless signals as they propagate, reflect, scatter, and diffract. This observation has led to the realization that wireless signals inherently carry information about the surrounding physical world, in addition to the data they are intended to convey[5][6]. This insight forms the foundation of wireless sensing, an area that has gained significant attention in recent years. Prior work has demonstrated that Wi-Fi signals can be used for device-free sensing tasks such as human activity recognition, localization, and health monitoring by analyzing channel measurements obtained during normal communication operations[7][8].

At the same time, modern wireless communication systems are becoming increasingly complex. Recent Wi-Fi and cellular standards employ wide bandwidths, multiple antennas, beamforming, and multi-band operation to support high data rates and dense user populations[9]. While these technologies improve communication performance, they also place greater demands on accurate channel knowledge. In practical systems, channel estimation is often limited by sparse pilots, non-contiguous spectrum, hardware impairments, and dynamic environments, causing traditional analytical methods to struggle under real-world conditions. These parallel developments—wireless sensing on one hand and increasingly sophisticated communication systems on the other—have motivated a new design philosophy known as Integrated Sensing and Communication (ISAC), which seeks to unify sensing and communication within a single wireless system. In an ISAC framework, the same signals, hardware, and spectrum resources are jointly leveraged for data transmission and environmental perception, enabling sensing insights to improve communication efficiency and adaptability.

## 1.1 Integrated Sensing and Communication

The integration of sensing and communication (ISAC) has emerged as a unifying paradigm that seeks to jointly design wireless systems for data transmission and environmental perception. By treating sensing and communication as tightly coupled functionalities, ISAC enables wireless networks to reuse the same signals, hardware, and spectrum resources to simultaneously support connectivity and situational awareness.

### 1.1.1 ISAC in IoT Technologies

Internet of Things (IoT) refers to a large ecosystem of interconnected devices that are embedded into everyday environments and are capable of sensing, communication, and control[10]. These devices include sensors, actuators, smart appliances, wearable devices, and industrial machines. In traditional IoT systems, sensing and communication are treated as two separate functions[11]. Sensors first collect physical measurements such as temperature, motion, or light, and these measurements are then transmitted over a wireless network to a central server or controller for processing and decision-making. As IoT systems continue to expand, this separation between sensing and communication becomes increasingly inefficient. Deploying large numbers of dedicated sensors increases cost, power consumption, and system complexity[12]. In addition, many IoT environments—such as homes, offices, hospitals, and factories—already contain dense wireless connectivity through technologies like Wi-Fi, Bluetooth, and low-power wide-area networks. This observation has motivated the idea of using existing wireless communication signals themselves as a source of sensing information[13].

Integrated Sensing and Communication provides a natural solution to these challenges by enabling IoT systems to sense the environment using the same wireless signals that are used for data communication. When IoT devices communicate wirelessly, the transmitted signals interact with the physical environment through reflection, scattering, and attenuation[4]. These interactions are influenced by factors such as human movement, object placement, and environmental layout. By analyzing how the received signals change over time, ISAC-enabled IoT systems can infer meaningful information about their surroundings without requiring additional sensing hardware. One of the most important benefits of ISAC in IoT systems is energy efficiency. Many IoT devices are battery-powered or rely on energy harvesting, making power consumption a critical design constraint. Traditional sensing

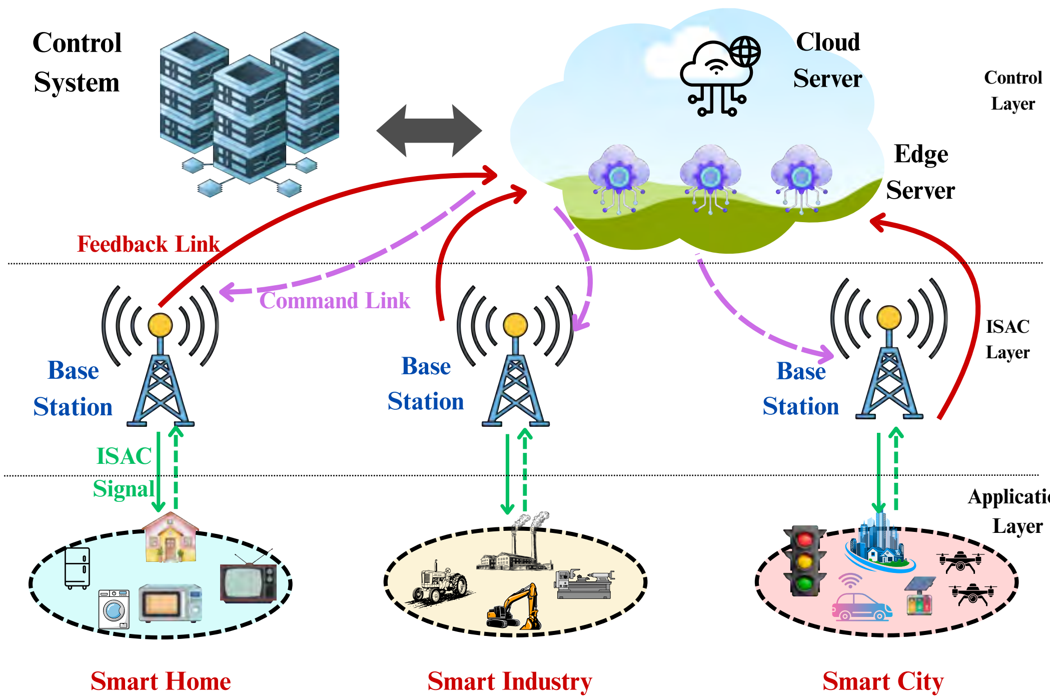



Figure 1.1: ISAC implementation in IoT network

approaches often require continuous sampling and signal processing, which can significantly reduce device lifetime. In contrast, ISAC allows sensing to be performed passively or opportunistically by reusing communication signals that are already transmitted for other purposes[14]. This reuse minimizes additional energy consumption and enables scalable sensing across large IoT deployments. Furthermore, ISAC enables IoT systems to become more context-aware and adaptive. Information inferred from wireless sensing can be used to guide communication decisions, trigger control actions, or adjust system behavior in real time[15]. For example, an IoT system may reduce communication activity when no human presence is detected or prioritize certain data transmissions based on sensed environmental conditions. By tightly integrating sensing and communication, ISAC supports the development of intelligent IoT systems that respond dynamically to changes in their physical environment rather than operating based on fixed assumptions.

### 1.1.2 ISAC in Cellular Networks

Cellular networks represent one of the earliest and most extensively studied platforms for Integrated Sensing and Communication, particularly in the context of fifth-generation (Fifth-Generation (5G)) and emerging sixth-generation (Sixth-Generation (6G)) wireless systems[5]. Modern cellular networks are designed to support high data rates, low latency, and reliable connectivity across large geographic areas. To achieve these goals, they employ advanced technologies such as massive multiple-input multiple-output MIMO, wide transmission bandwidths, high-frequency operation, and sophisticated beamforming techniques[16]. These same technologies also enable high-resolution sensing capabilities. Massive MIMO antenna arrays provide fine spatial resolution, while wide bandwidths offer precise temporal resolution. As a result, cellular base stations can observe detailed characteristics of the wireless channel, including angle, delay, and Doppler information[17]. By analyzing these channel properties, cellular networks can perform sensing tasks such as user localization, mobility tracking, environmental mapping, and vehicle detection, often using signal processing techniques similar to those used in radar systems[18].

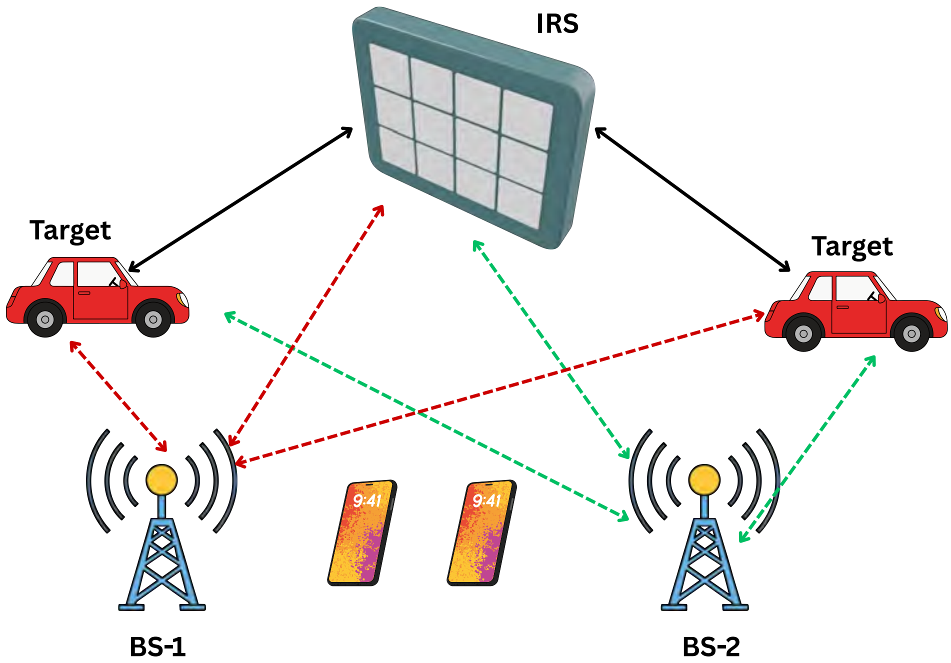



Figure 1.2: ISAC implementation in cellular network

In cellular ISAC systems, sensing is often performed at the infrastructure level. Base stations act as powerful sensing nodes that continuously monitor the environment while serving multiple users[19]. This network-centric sensing enables large-scale situational awareness and supports advanced applications such as intelligent transportation systems and autonomous driving. In addition, sensing information can be fed back into the communication system to improve performance[15]. For example, knowledge of user movement and environmental dynamics can help optimize beamforming, scheduling, and handover decisions. The integration of sensing into cellular networks is closely aligned with the vision of future wireless systems as intelligent and adaptive platforms. In 6G research, networks are expected to move beyond static configuration and become self-aware systems that continuously observe their environment, learn from data, and optimize their operation[20]. ISAC plays a key role in this vision by providing the sensory input needed for intelligent decision-making. Despite these advantages, cellular ISAC systems face practical limitations. The hardware required to support wideband sensing and large antenna arrays is expensive and consumes significant power[21]. Deploying such systems is often feasible only in large-scale or outdoor environments. Additionally, cellular infrastructure is centrally managed and may not provide the flexibility needed for fine-grained sensing in indoor or small-scale scenarios. These limitations highlight the need for complementary ISAC solutions based on more accessible wireless technologies, such as Wi-Fi.

### 1.1.3 ISAC in Wi-Fi Systems

Wi-Fi systems offer a highly practical and accessible platform for implementing Integrated Sensing and Communication (ISAC), particularly in indoor environments. Wi-Fi is one of the most widely deployed wireless technologies in the world, providing connectivity in homes, offices, campuses, and public spaces[22]. Wi-Fi access points are typically installed close to users and objects, resulting in dense spatial coverage that is well suited for detailed environmental sensing. Recent advancements in Wi-Fi standards have significantly enhanced the sensing capabilities of Wi-Fi systems. Modern Wi-Fi technologies support wider bandwidths, multiple antennas, and explicit channel measurements that capture detailed information about the wireless channel. Measurements such as Channel State Information and Channel Frequency Response describe how signals propagate across frequency and time, and they are sensitive to even small changes in the environment[23]. Human movement, object displacement, and changes in line-of-sight conditions can all cause measurable variations in these channel parameters[14]. One of the key strengths of Wi-Fi-based ISAC is its low cost and ease of deployment. Unlike cellular systems, Wi-Fi hardware is inexpensive and widely available, making it possible to deploy sensing applications at large scale without specialized equipment. Wi-Fi sensing can often be implemented through software updates or signal processing techniques applied to existing devices, which reduces deployment barriers and encourages experimentation. In addition, Wi-Fi operates in unlicensed spectrum, allowing rapid development and deployment without regulatory constraints. However, Wi-Fi was originally designed purely for communication, not sensing. As a result, Wi-Fi-based ISAC systems face unique challenges. Packet transmissions are irregular and driven by network traffic rather than sensing requirements[24]. Channel measurements may be noisy, incomplete, or vendor-specific, and hardware imperfections can introduce instability in sensing data. These factors make it difficult to extract reliable and generalizable sensing information directly from raw Wi-Fi measurements.
Addressing these challenges requires new approaches that are specifically tailored to the constraints of Wi-Fi systems. Instead of relying on idealized assumptions or dense measurements, Wi-Fi ISAC solutions must exploit the underlying structure of the wireless channel and focus on features that are robust to noise and variability[18]. Given its widespread adoption and practical relevance, Wi-Fi provides an ideal testbed for developing ISAC techniques that are both effective and deployable in real-world environments. For these reasons, this thesis focuses on the integration of sensing and communication in Wi-Fi systems, with the goal of demonstrating scalable and practical ISAC solutions using commodity hardware[8].

## 1.2 Research Motivation

Integrated Sensing and Communication (ISAC) is motivated by a simple but fundamental observation: the wireless channel is not only a communication medium, but also a physical phenomenon shaped by the environment. Every packet transmission interacts with walls, furniture, and human bodies through reflection, diffraction, and scattering. As a result, the channel response carries implicit information about the surrounding physical world. At the same time, modern wireless systems increasingly depend on accurate channel knowledge to enable advanced features such as beamforming, multi-user MIMO, and multi-band operation. These two realities suggest that sensing and communication should not be treated as isolated functionalities. Instead, they should be designed as complementary capabilities that reuse the same measurements, infrastructure, and resources whenever possible.

This thesis is motivated by the gap between what is theoretically possible in ISAC and what is practical in real deployments, especially with commodity Wi-Fi. As discussed in the challenges section, Wi-Fi sensing and learning-driven communication face instability from device imperfections, compute intensive heavyweight tasks, irregular measurements dictated by protocol behavior, and strong variability across environments. The variations in multipath components restrain ISAC from being consistent across different environments. These challenges do not imply that Wi-Fi is a poor platform for ISAC. Rather, they motivate a careful design philosophy: *Wi-Fi ISAC must be built by explicitly embracing the constraints of commodity hardware and real network behavior, instead of assuming idealized sensing schedules or perfectly calibrated measurements.* In particular, this thesis is motivated by the need to (i) exploit wireless channel structure for sensing purpose in realistic Wi-Fi settings, (ii) building up a domain adaptive learning driven intelligent wireless sensing system, (iii) reduce computational and energy burden of learning models so that they can be deployed on practical edge devices, and (iv) reduce communication overhead by sensing channel dynamics and channel structure more intelligently, rather than repeatedly reporting or retransmitting redundant information.

A key principle underlying this motivation is that wireless channels exhibit rich structure across frequency, space, and time. Even in challenging indoor environments, channel measurements are not arbitrary: they often show correlation across subcarriers, low-dimensional behavior in the spatial domain, and temporal consistency over short intervals. These properties provide an opportunity to redesign both sensing and communication pipelines. Instead of treating each packet as an isolated sample, or treating each subcarrier as an independent coefficient to estimate, one can treat channel measurements as structured objects that can be learned, compressed, aligned, and tracked. This shift in viewpoint directly supports both Wi-Fi sensing and learning-driven communication, because it encourages approaches that leverage the intrinsic regularities of the channel rather than relying on dense measurements or heavy computation.

### 1.2.1 A Unified View of Communication-Aided Sensing and Sensing-Aided Communication

A core motivation for Wi-Fi ISAC is that sensing and communication naturally enhance each other, but in different ways. This motivates two complementary directions: communication-aided sensing and sensing-aided communication. Both are necessary because they address different limitations of real wireless systems. **Communication-aided sensing** is motivated by the fact that Wi-Fi infrastructure is already widely deployed, continuously transmitting packets for connectivity. If sensing can be performed by reusing these existing communication signals and measurements, then sensing becomes scalable, low-cost, and easy to adopt. This is especially important in indoor environments where deploying dedicated sensors is expensive, intrusive, and difficult to maintain. Moreover, Wi-Fi operates in close proximity to people and objects, enabling rich sensing signatures without requiring line-of-sight cameras or specialized radar hardware. The motivation here is not merely to demonstrate that sensing is possible, but to make sensing *deployable*: sensing systems should work with commodity devices, tolerate measurement irregularity, and remain robust under interference and changing traffic patterns. In practice, this means sensing methods must handle sparse and non-uniform sampling, measurement noise, and device heterogeneity. Therefore, communication-aided sensing in Wi-Fi is motivated by the need for practical, privacy-preserving, and infrastructure-free sensing that can operate within the constraints of real networks.

**Sensing-aided communication** is motivated by the opposite direction: modern Wi-Fi communication performance is fundamentally constrained by overhead and imperfect channel knowledge. Advanced features such as beamforming and Multi-user MIMO (mu-MIMO) depend on accurate channel information, but obtaining that information repeatedly can consume significant airtime and control resources. As systems evolve toward larger antenna arrays, wider bandwidths, and more fragmented spectrum, the cost of acquiring, feeding back, and updating channel state increases. At the same time, channels are not arbitrary; they evolve in ways that reflect mobility and environmental dynamics. If the transmitter can sense when the channel has meaningfully changed, or infer missing channel information from structure and correlation, it can reduce unnecessary updates and focus resources where they matter. This motivates sensing-aided communication as a way to improve throughput and reliability not by adding more pilots or more feedback, but by *understanding channel behavior* and using that understanding to reduce overhead.

Crucially, these two directions should not be viewed as separate research tracks that happen to share a dataset. They represent two parallel ways of treating the channel as a structured physical object. Communication-aided sensing asks: ***What can we infer about the world from channel measurements that already exist for communication?*** Sensing-aided communication asks: ***What can we infer about the channel itself so that we can communicate more efficiently with fewer measurements and less overhead?*** Together, they motivate a unified ISAC perspective that is particularly relevant for Wi-Fi, where signals are abundant but measurements are constrained and protocol-driven.

### 1.2.2 Wi-Fi as a Practical and Deployable ISAC Platform

Wi-Fi is an appealing platform for ISAC because it is ubiquitous and operates primarily in indoor environments where sensing applications are most needed. However, the motivation to focus on Wi-Fi is not simply its popularity. The deeper motivation is that Wi-Fi represents the most realistic testbed for ISAC under practical constraints. Unlike controlled radar-style sensing systems, Wi-Fi transmissions are shaped by contention, scheduling, rate adaptation, and user traffic. Channel measurements arrive irregularly, may be partially available depending on the device and driver, and can be distorted by hardware impairments. Environmental variation is also more severe indoors, where multipath is rich and sensitive to small changes in geometry. This reality motivates a shift from idealized assumptions to deployable design. Many sensing approaches implicitly assume consistent sampling, stable hardware, and clean channel measurements. In practice, Wi-Fi breaks these assumptions. Therefore, a central motivation of this thesis is to develop ISAC methodologies that remain meaningful under Wi-Fi constraints: irregular packet timing, unstable phase behavior, limited access to raw channel information, and device heterogeneity. This also motivates investigating representations and processing strategies that are robust to device noise and protocol behavior. Rather than relying on fragile low-level signal details, practical Wi-Fi sensing and communication enhancement should exploit channel structure that remains stable across these variations.

Another motivation is scalability. In realistic deployments, Wi-Fi networks may serve many devices simultaneously, often using mu-MIMO and wide channels. Sensing approaches that require specialized firmware access, heavy per-packet computation, or dense measurement scheduling can become impractical at scale. Similarly, communication approaches that require frequent sounding or large feedback can reduce network efficiency. This motivates approaches that reduce both computational load and protocol overhead, so that ISAC benefits can be realized without disrupting network operation.

Finally, Wi-Fi is evolving toward explicit sensing support in standardization efforts. This trend motivates foundational research that can influence how sensing and communication will coexist in future Wi-Fi generations. Even before sensing is fully standardized, Wi-Fi already provides rich channel measurements through existing procedures such as channel sounding and feedback. A practical motivation of this thesis is therefore to exploit what is already available in commodity Wi-Fi today, while also guiding how Wi-Fi ISAC can be strengthened in the future.

### 1.2.3 Domain Generalization as a Core Design Requirement for Wi-Fi ISAC

A central motivation of this thesis is the challenge of *domain generalization* in real-world Wi-Fi sensing and learning-driven communication systems. In practical deployments, wireless environments are highly diverse: variations in room geometry, furniture layout, building materials, device placement, antenna orientation, and human behavior can significantly alter observed channel responses. Consequently, channel measurements collected in one environment often exhibit markedly different statistical properties in another, posing a major challenge for data-driven approaches. Many existing Wi-Fi sensing and communication enhancement systems are trained under controlled conditions, typically using data from a limited set of environments. While such models may achieve high accuracy during evaluation, their performance often degrades sharply when deployed in new environments or with new users. This degradation reflects the strong sensitivity of wireless channels to contextual factors such as multipath structure, interference, and hardware characteristics. As a result, models that rely on environment-specific features are difficult to deploy at scale.

Addressing domain generalization is essential given the realities of Wi-Fi operation. Wi-Fi networks are deployed for connectivity rather than sensing, and their environments evolve continuously as access points are repositioned, devices join or leave the network, and user behavior changes. Retraining models for every environmental change is impractical. A deployable Wi-Fi ISAC system must therefore extract channel representations that remain meaningful across environments, devices, and usage scenarios. While channel measurements vary across domains, they are governed by consistent physical propagation mechanisms such as reflection, diffraction, and scattering. This motivates a shift from learning superficial, environment-dependent patterns toward representations that capture transferable channel structure. Properties such as temporal consistency, frequency-domain correlation, and low-dimensional spatial structure often persist across environments, even as absolute channel values change. Exploiting these shared characteristics enables models that are more robust to domain shifts. Domain generalization is equally critical in sensing-aided communication, where inaccurate channel inference under changing conditions can degrade beamforming, rate adaptation, and feedback efficiency. Moreover, the cost of collecting labeled data across many environments and devices makes large-scale, environment-specific training impractical. These constraints motivate approaches that generalize from limited data and adapt gracefully without extensive retraining. For these reasons, this thesis treats domain variability as a first-class design consideration rather than an afterthought. By emphasizing channel structure, temporal evolution, and physically grounded representations, the proposed Wi-Fi ISAC frameworks aim to achieve robust, efficient, and deployable sensing and communication across diverse real-world environments.

### 1.2.4 Channel Structure as the Key to Reducing Model Complexity and Communication Overhead

A recurring challenge in both sensing and learning-driven communication is that naive learning solutions tend to become heavy. High-dimensional Channel State Information (CSI)/CFR tensors, multiple antenna pairs, and wide bandwidths encourage deep models with large memory footprints and high computation. These models may perform well in controlled experiments, but become difficult to deploy on real devices due to latency, energy consumption, and the need for frequent retraining under distribution shifts. Therefore, this thesis is motivated by the need to reduce the computational footprint of learning-based approaches while preserving robustness and accuracy. The key insight motivating this direction is that wireless channels are structured and often redundant. Across frequency, adjacent subcarriers are correlated because multipath delays are shared. Across antennas, spatial responses often lie in a lower-dimensional subspace because propagation paths are limited. Across time, short-term channel evolution is often smooth except during significant motion events. If learning models explicitly leverage these properties, then they can be made lighter: fewer parameters are needed to represent the channel, fewer samples are needed to infer missing information, and fewer updates are needed when the channel is stable. In other words, channel structure provides a path toward *efficient learning*, not only accurate learning.

This same structural view also motivates reducing communication overhead. Traditional channel acquisition often treats each update as an independent requirement: pilots are transmitted, channels are estimated, and feedback is returned. In practice, this can result in repeated reporting of information that has not meaningfully changed. If the system can sense channel evolution and determine when updates are truly needed, overhead can be reduced. Similarly, if the system can infer high-resolution channel information from sparse measurements using learned structural priors, then fewer pilots and fewer feedback bits are required. This is especially relevant in wideband and fragmented spectrum operation, where measuring every subcarrier in every band becomes expensive. Multi-band correlation further strengthens this motivation: channels at different frequency bands can share propagation paths, enabling inference across bands and reducing the need for exhaustive measurement.

In summary, the motivation of this thesis is rooted in efficiency and deployability. Wi-Fi ISAC should not rely on heavy models that assume ideal measurements, nor should it require excessive control overhead that reduces throughput. Instead, it should exploit the physical and statistical structure of wireless channels to build methods that are robust to real-world constraints, computationally efficient, and communication-efficient. These motivations naturally lead to the research objectives presented in the next section, where the thesis contributions are defined in terms of concrete thrusts, system designs, and methodological goals.

## 1.3 Problem Statement

Integrated Sensing and Communication (ISAC) aims to reuse the same wireless signals and hardware to simultaneously support environmental sensing and data communication. In commodity Wi-Fi, this goal is attractive because Wi-Fi infrastructure is ubiquitous and already exposes channel measurements (e.g., CFR/CSI) through standardized procedures such as channel sounding and feedback. However, directly using these measurements for ISAC in real deployments remains difficult because (i) measurements are *protocol-driven* rather than sensing-driven, (ii) commodity radios introduce *device-induced distortions* and irregular sampling, and (iii) the channel is high-dimensional and strongly *domain dependent* across environments, devices, and users. Meanwhile, modern Wi-Fi communication increasingly relies on accurate channel knowledge for mu-MIMO and wideband operation, but obtaining and reporting this channel state at scale creates substantial control overhead and user-side computational burden.
This thesis addresses a focused ISAC problem in commodity Wi-Fi:

> *How can commodity Wi-Fi channel measurements be transformed into structured, low-overhead representations that (i) enable reliable sensing under protocol-driven, imperfect measurements and domain shifts, and (ii) enable scalable multi-antenna communication by reducing channel feedback overhead without imposing heavy computation at the user device?*

Rather than pursuing waveform redesign, radar-style sensing, or MAC reconfiguration, the thesis focuses on *representation and inference*: treating Wi-Fi channel measurements as structured objects whose regularities across *time* and *frequency* can be exploited to make sensing and communication practical.

## Problem P1: Robust Sensing from Protocol-Driven and Imperfect Wi-Fi Measurements

Wi-Fi channel measurements are collected opportunistically based on traffic, contention, and scheduling. Consequently, samples can be sparse, bursty, and irregular in time. In addition, commodity hardware introduces distortions (e.g., phase instability, quantization artifacts, residual frequency offsets) that create fluctuations unrelated to physical motion. These effects violate common assumptions used in learning-based sensing, where channel tensors are treated as clean, uniformly sampled inputs. As a result, models trained under controlled conditions can misinterpret device-induced variations as environmental changes and fail when traffic patterns or devices change.

## Problem P2: Domain Shift across Environments, Devices, and Users

Indoor Wi-Fi channels are highly sensitive to room geometry, multipath structure, device placement, and human behavior. Channel statistics therefore change substantially across environments and stations, causing strong domain shifts that degrade sensing accuracy and learning-based inference. Frequent retraining is impractical due to the cost of data collection and labeling, especially for sensing tasks. A deployable Wi-Fi ISAC system must therefore learn representations that transfer across domains and adapt with minimal labeled data.

## Problem P3: Scalability Limits of Conventional Wideband MIMO Channel Feedback

In multi-antenna wideband Wi-Fi systems, channel feedback overhead grows rapidly with bandwidth and antenna dimension under conventional explicit sounding. This overhead consumes airtime and reduces effective throughput. Learning-based compression can reduce the number of feedback bits, but many approaches assume a neural encoder at the user device, shifting significant computational cost to battery-powered clients. Moreover, both conventional and learning-based pipelines often ignore frequency-domain redundancy and report information uniformly across subcarriers, even when many subchannels are highly correlated.

## Problem Scope and Thesis Focus

In summary, the thesis targets two concrete and coupled gaps in commodity Wi-Fi ISAC: (i) enabling robust, generalizable sensing from imperfect, protocol-driven channel measurements without heavy models, and (ii) enabling scalable communication by reducing wideband MIMO channel feedback overhead without introducing heavy user-side computation. The key hypothesis is that Wi-Fi channels exhibit exploitable structure and redundancy across time and frequency, and that explicitly leveraging this structure can yield deployable ISAC frameworks that remain robust under real-world constraints.

## 1.4 Research Objectives

Guided by the challenges and motivations discussed earlier, the objective of this thesis is to develop *concrete and deployable ISAC mechanisms for commodity Wi-Fi systems* under realistic operational constraints. Specifically, this thesis targets two narrowly defined and practically relevant instantiations of Wi-Fi ISAC: (i) temporal CFR-based human activity sensing using protocol-driven communication measurements, and (ii) low-overhead channel feedback for wideband Wi-Fi MIMO systems through structural sensing of the channel frequency response.
Rather than treating sensing and communication as isolated or idealized problems, the thesis adopts a unified perspective in which the wireless channel is viewed as a *structured physical object* whose temporal, frequency-domain, and spatial properties can be selectively exploited to support both functions efficiently. The overarching research objective of this thesis is therefore summarized as follows:

> *To design learning-driven Wi-Fi ISAC frameworks that enable (i) reliable temporal CFR-based sensing under irregular, protocol-driven measurements, and (ii) scalable MIMO channel feedback with reduced overhead, by explicitly exploiting channel structure while respecting the computational and protocol constraints of commodity Wi-Fi systems.*

To achieve this objective, the thesis is organized around two complementary but clearly scoped research thrusts, each corresponding to a specific ISAC mechanism rather than a general sensing or communication paradigm.

### 1.4.1 Thrust I: A Communication-Aided Sensing Approach via PULSE

The first research thrust focuses on *communication-aided sensing in the specific context of human activity recognition using Wi-Fi Channel Frequency Response (CFR).* The objective is to develop a novel Wi-Fi sensing framework, that robustly infer human activities from the temporal evolution of CFR measurements that are irregularly sampled, noisy, and shaped by Wi-Fi protocol behavior.

To this end, this thesis proposes PULSE , a sensing framework that explicitly prioritizes *temporal channel dynamics* over raw high-dimensional channel representations. The objectives of this thrust are threefold. First, to design sensing representations that capture physically meaningful temporal variations in the wireless channel while suppressing hardware- and protocol-induced artifacts. Second, to enable domain-adaptive activity sensing, where models trained in one environment generalize to new environments, users, and device placements with minimal labeled data. Third, to ensure that sensing can be performed with *lightweight computation suitable for edge devices*, without relying on dense CFR tensors or heavyweight neural architectures.

### 1.4.2 Thrust II: Sensing-Aided Communication via ELF

The second research thrust focuses on a *sensing-aided communication in the specific context of channel feedback for wideband Wi-Fi MIMO systems.* Particularly, the objective of this thrust is to address a precise and practical scalability bottleneck:

the rapid growth of channel feedback overhead and receiver-side computation as bandwidth and antenna counts increase.

This thesis introduces ELF , a framework that reinterprets channel feedback as a *structural sensing problem over the channel frequency response*, rather than as dense coefficient reporting or end-to-end autoencoder compression. The objectives of this thrust are threefold. First, to exploit redundancy across subcarriers by learning a latent representation in which only a small number of representative subchannels need to be reported. Second, to shift learning and reconstruction complexity away from user devices and toward the base station, ensuring that feedback mechanisms remain feasible on commodity clients. Third, to align channel reconstruction with communication performance metrics, ensuring that reduced feedback directly supports reliable precoding and data transmission.

### 1.4.3 Overall Contribution of the Thesis

Together, the two research thrusts establish a unified and coherent contribution toward practical Wi-Fi–based Integrated Sensing and Communication. The summery of contribution is listed as below:

- PULSE introduces a physics-aware temporal embedding framework that extracts temporal descriptors pertaining to channel dynamics from CFR, which preserves the discriminative structure while reducing input size by up to 85% for efficient edge sensing.
- PULSE proposes a unified temporal embedding based learning pipeline, where a lightweight 1D CNN jointly performs feature embedding and activity inference. Unlike conventional sensing approaches, this joint learning enhances discriminability and minimizes redundancy in temporal representations.
- PULSE integrates a supervised contrastive pretraining and prototype-based few-shot adaptation strategy, achieving strong cross-domain generalization from only a few seconds of labeled data.
- PULSE achieves **99.8% accuracy** in recognizing 20 activities within known environments and **95–96%** when tested in unseen domains, improving cross-domain performance by up to **9%** over SOTA approach– *BeamSense* [25].
- We propose ELF , a learning-based framework for scalable and low-overhead channel feedback in wideband Wi-Fi MIMO systems. Unlike conventional autoencoder-based methods, ELF removes the need for a heavy User Equipment (UE)-side encoder by shifting channel reconstruction entirely to the Base Station (BS), keeping UE-side processing lightweight.
- We introduce an *embedding-guided clustering* mechanism that selects a small set of representative subchannels and feeds back only their quantized latent embeddings. With a representative ratio of $K/F \approx 1/8$ (e.g., $K$=128 for $F$=1024 at 80 MHz), ELF achieves $\geq$**96%** feedback reduction while preserving near-dense reconstruction performance.

- We validate ELF in a Sionna simulation environment using 3rd Generation Partnership Project (3GPP) TR 38.901 Clustered Delay Line (CDL) channels across Wi-Fi-like bandwidths from **40–320 MHz** and MIMO dimensions up to 5×5. Results show that ELF provides up to 25× lower feedback than IEEE 802.11ax explicit feedback across bandwidths, while its feedback size remains approximately **constant across antenna configurations** (i.e., it does not grow with additional antenna pairs). At the practical operating point ($K$=128 at 80 MHz), ELF achieves ∼**34 MFLOPs** with inference latency below **0.5 ms/sample**, and maintains BER performance with **negligible degradation** compared to dense feedback.

- For reproducibility, we pledge to share our code base for PULSE at: https://github.com/rifatzabin/PULSE, and code base for ELF at: https://github.com/rifatzabin/ELF-Private

## 1.5 Thesis Organization

The following chapters of this thesis is organized to progressively develop the system designs required for practical ISAC system. The organization of this thesis is as follows: **Chapter 2** presents the literature review, covering the necessary background on Wi-Fi systems and Radio Frequency (RF) sensing, and summarizing state-of-the-art research relevant to this thesis. **Chapter 3** introduces PULSE and details the proposed communication-aided sensing framework, including system design and methodology. **Chapter 4** presents ELF and develops the sensing-aided communication framework for learning-driven feedback and adaptation in Wi-Fi MIMO systems. **Chapter 5** analyses the accuracy performance of both PULSE and ELF explicitly. **Chapter 6** concludes the thesis and outlines future research directions.

# Chapter 2

# Literature Review

## 2.1 Background

This section introduces the core technical foundations required to understand Wi-Fi-based ISAC. It first reviews how commodity Wi-Fi signals enable device-free sensing, then summarizes the evolution of IEEE 802.11 standards and the channel sounding mechanisms that expose CFR/CSI measurements used throughout this thesis.

### 2.1.1 Wi-Fi for Sensing

Wi-Fi has become one of the most pervasive wireless technologies in modern society that serves as the primary means of indoor Internet connectivity across homes, offices, campuses, hospitals, and public spaces [26][27][28][25]. The widespread adoption of Wi-Fi is driven by its low deployment cost, operation in unlicensed spectrum, and continual evolution through standardized IEEE 802.11 protocols. Economic projections estimate that the global Wi-Fi ecosystem will reach a value of approximately 4.9 trillion by 2025, highlighting both its technological and societal impact[29]. This extensive deployment makes Wi-Fi an attractive platform for applications that extend beyond traditional data communication.

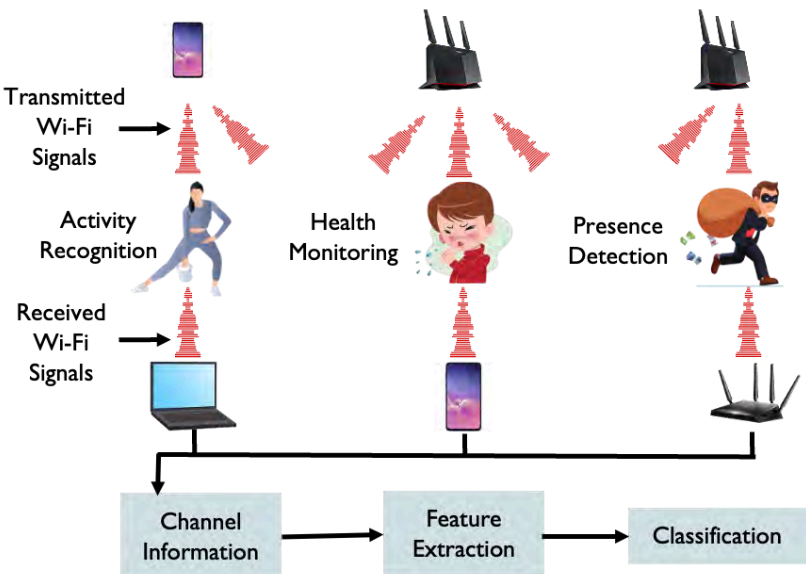



Figure 2.1: Sensing task leveraged by Wi-Fi networks

In recent years, Wi-Fi has emerged as a promising technology for device-free sensing. The goal is to infer information about the physical environment without requiring individuals to carry sensors or wearables. The fundamental principle behind Wi-Fi sensing is that wireless signals interact with their surroundings as they propagate. When Wi-Fi signals travel between a transmitter and a receiver, they undergo reflection, diffraction, and scattering due to walls, furniture, and human bodies present in the environment[30]. Any movement or change in the physical scene alters the multipath propagation characteristics of the wireless channel. As depicted in Fig 2.1, these changes are captured in channel measurements such as CSI or CFR, which are originally computed for communication purposes such as channel estimation and equalization. By analyzing variations in these channel measurements over time, Wi-Fi devices can infer a wide range of sensing information. Prior research has demonstrated the feasibility of using Wi-Fi signals for fine-grained indoor localization, where the position of a person or object is estimated based on signal distortions[31][32]. Wi-Fi sensing has also been applied to human activity recognition, enabling systems to distinguish between activities such as walking, sitting, standing, or gesturing. In healthcare-oriented applications, Wi-Fi sensing has been explored for monitoring respiration, detecting falls, and assessing general mobility patterns, all while preserving user privacy since no visual data is collected. Because Wi-Fi infrastructure is already densely deployed in indoor environments, sensing capabilities can often be enabled through software-level processing without requiring additional hardware installation. This significantly reduces cost and deployment complexity compared to traditional sensing modalities such as cameras or radar systems. Moreover, Wi-Fi sensing naturally operates in environments where vision-based approaches may fail due to occlusion, lighting conditions, or privacy concerns.
Despite these advantages, Wi-Fi sensing remains a challenging and evolving research area. A large portion of existing work focuses on single-subject sensing scenarios, where only one individual is present and performing activities at a time. While such settings simplify modeling and inference, they do not reflect realistic environments such as homes or offices, where multiple people may be present simultaneously. Multi-subject sensing introduces additional complexity because signal variations caused by different individuals overlap in time and space, making it difficult to disentangle individual contributions to the channel dynamics[27]. Another fundamental challenge in Wi-Fi sensing is its strong dependence on environment and subject characteristics. Wireless channels are highly sensitive to room layout, furniture placement, and background motion, which can vary significantly across deployments. In addition, different individuals interact with the wireless channel in different ways due to variations in body shape, movement speed, and activity execution style. As a result, sensing models trained in one environment or with one set of users often experience degraded performance when deployed in new settings. While some efforts have attempted to improve robustness and generalization, these approaches typically consider a limited set of activities or simplified sensing scenarios.

Recognizing the growing importance of Wi-Fi sensing, the IEEE has recently initiated efforts to formally support sensing applications within Wi-Fi standards[26]. The IEEE 802.11bf working group has been established to define requirements and mechanisms for Wi-Fi sensing, marking a significant step toward standardizing sensing functionality alongside communication. This development highlights the long-term relevance of Wi-Fi sensing and its potential role in future wireless systems.

The following sections will represent a clear conception about different Wi-Fi standards along with their underlying features, the general channel sounding procedure to obtain the corresponding CSI or CFR from the channel for sensing task, and the prior approaches from research community towards Wi-Fi based sensing.

### 2.1.2 Evolution of Wi-Fi

Wi-Fi is based on the IEEE 802.11 family of standards, which has evolved over more than two decades to meet the increasing demand for higher data rates, lower latency, and improved spectral efficiency. While early Wi-Fi standards were designed primarily for basic wireless data connectivity, successive generations have introduced physical-layer and medium-access innovations that unintentionally—but critically—enable wireless sensing.

#### Early Wi-Fi Standards: Foundations (802.11a/b/g)

The first IEEE 802.11 standard was ratified in 1997 and supported data rates up to 2 Mbps. Soon after, 802.11b introduced operation in the 2.4 GHz unlicensed band with data rates up to 11 Mbps, making Wi-Fi commercially viable. 802.11a and 802.11g followed, introducing Orthogonal Frequency Division Multiplexing (OFDM), which divides the channel into multiple orthogonal subcarriers. OFDM marked a foundational step toward sensing because it enabled subcarrier-level channel estimation, allowing the receiver to observe frequency-selective fading. However, these early standards used narrow channel bandwidths (20 MHz) and Single Input Single Output (SISO) transmission, limiting both communication capacity and sensing resolution.

#### MIMO and Channel Expansion: 802.11n

A major milestone occurred with IEEE 802.11n, which introduced MIMO transmission. By employing multiple antennas at the transmitter and receiver, 802.11n enabled spatial multiplexing and significantly increased throughput. It also introduced optional 40 MHz channels, doubling the spectral bandwidth.
From a sensing perspective, 802.11n was transformative. MIMO allowed Wi-Fi systems to observe spatial diversity, while wider bandwidth improved temporal resolution. More importantly, channel estimation became multi-dimensional, producing CSI or CFR matrices indexed by subcarrier, transmit antenna, and receive antenna. These multi-dimensional measurements form the basis of most modern Wi-Fi sensing systems.

#### Very High Throughput Wi-Fi: 802.11ac

IEEE 802.11ac extended Wi-Fi capabilities substantially by operating exclusively in the 5 GHz band and introducing 80 MHz and 160 MHz channels, along with

**256-QAM modulation**. It also standardized explicit beamforming through Null Data Packet (NDP) sounding and enabled mu-MIMO.
These features significantly enhanced sensing potential. Wider bandwidths increased delay resolution, while larger antenna arrays improved angular resolution. Explicit beamforming required frequent and structured channel sounding, resulting in high-quality CFR measurements. As a result, 802.11ac became one of the most widely used platforms for fine-grained indoor sensing, including activity recognition, localization, and gesture detection.

### High Efficiency Wi-Fi: 802.11ax (Wi-Fi 6)

IEEE 802.11ax, also known as Wi-Fi 6, shifted the design focus from peak throughput to spectral efficiency and dense deployment performance. It introduced Orthogonal Frequency Division Multiple Access (OFDMA), allowing multiple users to share subsets of subcarriers within the same transmission. While channel bandwidths remain similar to 802.11ac, OFDMA enables much finer granularity in frequency allocation.
For sensing, 802.11ax offers two important advantages. First, the division of channels into Resource Units (RUs) enables selective sensing over specific frequency regions. Second, denser channel sounding and improved synchronization increase temporal consistency. These features make 802.11ax particularly suitable for scalable sensing in multi-user and crowded environments.

### Toward Native Sensing Support: 802.11bf

Recognizing the growing importance of Wi-Fi sensing, IEEE established **Task Group 802.11bf**, which aims to explicitly support sensing functionality in future Wi-Fi standards. Unlike previous amendments—where sensing was an unintended byproduct—802.11bf seeks to define standardized mechanisms for sensing measurements, reporting, and coordination.

This development marks a paradigm shift: Wi-Fi is no longer viewed solely as a communication technology, but as a joint communication-and-sensing platform. The evolution from 802.11a to 802.11bf reflects a gradual transition from data delivery to environmental awareness. Table 2.1 demonstrate a comparison between major Wi-Fi standards in terms of their physical layer parameters.

Table 2.1: Comparison of Major Wi-Fi Standards and Their Physical-Layer Parameters

| **Standard** | **Year** | **Bandwidth** | **FFT Size** | **Subcarrier Spacing** | **Data** | **Pilot** | **Guard / Null** | **MIMO Support** |
|---|---|---|---|---|---|---|---|---|
| 802.11a/g | 1999 / 2003 | 20 MHz | 64 | 312.5 kHz | 48 | 4 | 12 | SISO |
| 802.11n | 2009 | 20 MHz | 64 | 312.5 kHz | 52 | 4 | 8 | Up to 4×4 |
| | | 40 MHz | 128 | 312.5 kHz | 108 | 6 | 14 | |
| 802.11ac | 2013 | 20 MHz | 64 | 312.5 kHz | 52 | 4 | 8 | Up to 8×8 (AP) |
| | | 40 MHz | 128 | 312.5 kHz | 108 | 6 | 14 | |
| | | 80 MHz | 256 | 312.5 kHz | 234 | 8 | 22 | |
| | | 160 MHz | 512 | 312.5 kHz | 468 | 16 | 44 | |
| 802.11ax | 2019 | 20 MHz | 256 | 78.125 kHz | 242 | 8 | 6 | Up to 8×8 MU-MIMO |
| | | 40 MHz | 512 | 78.125 kHz | 484 | 16 | 12 | |
| | | 80 MHz | 1024 | 78.125 kHz | 996 | 32 | 28 | |
| | | 160 MHz | 2048 | 78.125 kHz | 1992 | 64 | 56 | |
| 802.11bf | Ongoing | – | – | – | – | – | – | TBD |

### 2.1.3 Wi-Fi Channel Sounding Procedure

Modern Wi-Fi systems rely on accurate knowledge of the wireless channel to support advanced transmission techniques such as beamforming and multi-antenna spatial multiplexing. In IEEE 802.11ac and IEEE 802.11ax systems, this channel knowledge is obtained through an explicit *channel sounding* procedure. This procedure enables the transmitter to estimate how the wireless medium distorts transmitted signals across frequency, space, and time.

From a signal processing perspective, channel sounding allows the system to estimate the CFR, which describes how each frequency component of a transmitted signal is altered by the propagation environment as depicted in Fig. 2.2

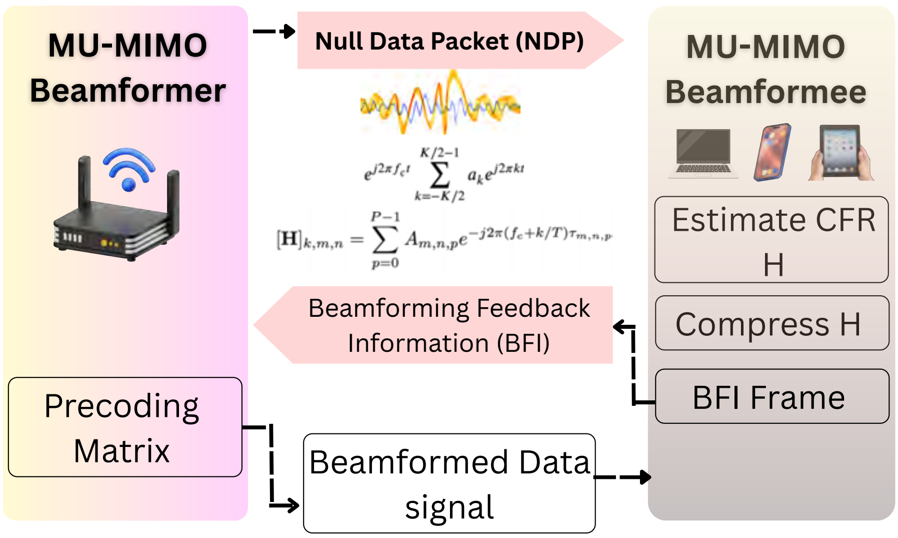



Figure 2.2: Wi-Fi channel sounding procedure

#### System Model and Channel Representation

Consider a Wi-Fi link where the transmitter is equipped with $M$ antennas and the receiver is equipped with $N$ antennas. The system operates over an OFDM waveform with $K$ subcarriers. For a given OFDM symbol, the baseband input–output relationship on the $k$-th subcarrier can be written as

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{n}_k, \quad k = 1, \ldots, K, \tag{2.1}$$

where

- $\mathbf{x}_k \in \mathbb{C}^M$ is the transmitted signal vector,
- $\mathbf{y}_k \in \mathbb{C}^N$ is the received signal vector,
- $\mathbf{H}_k \in \mathbb{C}^{N \times M}$ is the channel matrix at subcarrier $k$, and
- $\mathbf{n}_k$ represents additive noise.

Stacking the channel matrices across all subcarriers yields the full CFR tensor

$$\mathbf{H} \in \mathbb{C}^{K \times M \times N}. \tag{2.2}$$

Each element of this tensor captures both amplitude attenuation and phase rotation introduced by the wireless channel for a specific frequency and antenna pair.

### Null Data Packet (NDP) Transmission

To estimate the CFR, the transmitter—referred to as the *beamformer*—initiates channel sounding by transmitting a NDP. The NDP contains no user payload; instead, it consists entirely of known training symbols called Long Training Field (LTF). The LTF sequences are orthogonal across transmit antennas and known *a priori* to the receiver. This design allows the receiver—referred to as the *beamformee*—to isolate the contribution of each transmit antenna and estimate the corresponding channel coefficients. Importantly, the NDP is transmitted over the entire system bandwidth, enabling channel estimation across all active subcarriers.

### Channel Estimation at the Receiver

Upon receiving the NDP, the beamformee performs channel estimation by correlating the received LTFs with the known transmitted sequences. For each subcarrier $k$, the receiver computes an estimate $\hat{\mathbf{H}}_k$ of the channel matrix.
This estimation process yields a discrete-time sequence of CFR measurements

$$\{\hat{\mathbf{H}}(t)\}_{t=1}^{T}, \tag{2.3}$$

where $t$ indexes successive sounding instances. Each estimate reflects the instantaneous propagation conditions at the time of sounding.

Physically, these estimates encode the effects of multipath propagation, including reflections from walls and objects, diffraction around obstacles, and scattering caused by human motion.

### Feedback Compression and Quantization

Directly feeding back the full CFR tensor would incur excessive overhead, especially in wideband and multi-antenna systems. To address this, IEEE 802.11ac/ax specifies a feedback compression mechanism.
The receiver applies a matrix decomposition—typically a form of Singular Value Decomposition (SVD)—to extract the dominant spatial characteristics of the channel. The resulting representation is quantized and encoded into a beamforming feedback matrix, which is significantly smaller than the raw CFR. This feedback preserves the essential structure of the channel while reducing uplink overhead.

### Feedback Transmission and Channel Reconstruction

The compressed feedback is transmitted back to the beamformer using a dedicated beamforming feedback frame. Upon reception, the beamformer reconstructs an approximation of the channel and computes a precoding matrix $\mathbf{W}_k$ for each subcarrier. The transmitted signal is then pre-processed as

$$\mathbf{x}_k = \mathbf{W}_k \mathbf{s}_k, \tag{2.4}$$

where $\mathbf{s}_k$ denotes the data symbol vector. This precoding ensures that signals transmitted from different antennas combine constructively at the receiver.

**Temporal Evolution of the Channel**

Channel sounding is performed periodically to account for time-varying propagation conditions. As a result, the beamformer observes a temporal evolution of the CFR:

$$\mathbf{H}(t) = \mathbf{H}_{\text{static}} + \Delta\mathbf{H}(t), \tag{2.5}$$

where

- $\mathbf{H}_{\text{static}}$ represents quasi-static components (e.g., walls, furniture), and
- $\Delta\mathbf{H}(t)$ captures dynamic variations due to motion and environmental changes.

These temporal variations are typically treated as impairments in communication systems. However, from a sensing perspective, they constitute a rich source of physical information.

**Channel Sounding as a Sensing Enabler**

The CFR estimated during channel sounding is sensitive to changes in the physical environment. Human motion, object displacement, and posture changes alter the multipath structure of the channel, which manifests as measurable variations in $\mathbf{H}(t)$.

Consequently, the channel sounding procedure—originally designed to support beamforming, also provides a continuous stream of sensing measurements. By analyzing the temporal and spectral structure of the CFR, it becomes possible to infer environmental dynamics without modifying the Wi-Fi protocol or adding specialized sensing hardware. The following section provides detailed on this aspect.

### 2.1.4 CFR as a Sensing Primitive

The CFR estimated during the channel sounding procedure provides a detailed description of how wireless signals propagate through the environment. While its primary role in Wi-Fi systems is to enable reliable communication through beamforming and equalization, the CFR also serves as a powerful sensing primitive due to its direct dependence on the physical propagation environment.

From a physical-layer perspective, the wireless channel can be modeled as a superposition of multiple propagation paths. Each path corresponds to a reflected, diffracted, or scattered signal component caused by interactions with objects, walls, furniture, and human bodies. For a given transmit–receive antenna pair, the CFR at subcarrier $k$ and time $t$ can be expressed as

$$H_k(t) = \sum_{p=1}^{P} \alpha_p(t)\, e^{-j2\pi f_k \tau_p(t)}, \tag{2.6}$$

where $P$ denotes the number of dominant multipath components, $\alpha_p(t)$ represents the complex gain of the $p$-th path, $\tau_p(t)$ denotes its propagation delay, and $f_k$ is the frequency of the $k$-th subcarrier. Both $\alpha_p(t)$ and $\tau_p(t)$ depend on the geometry and material properties of the environment. Changes in the physical world directly affect these parameters. Human motion, for example, introduces time-varying reflections and shadowing effects, which modify the amplitudes and phases of the multipath components. As a result, even small movements lead to measurable perturbations in the CFR across frequency, antennas, and time.

When extended to a multi-antenna system with $M$ transmit antennas and $N$ receive antennas, the CFR becomes a tensor-valued representation

$$\mathbf{H}(t) \in \mathbb{C}^{K\times M\times N}, \tag{2.7}$$

capturing frequency-selective and spatially diverse channel responses. This high-dimensional representation encodes rich information about the environment, including spatial structure, motion dynamics, and temporal evolution.

From a sensing perspective, the key insight is that the CFR is not a static quantity. Instead, it evolves over time according to

$$\mathbf{H}(t) = \mathbf{H}_{\text{static}} + \mathbf{H}_{\text{dynamic}}(t), \tag{2.8}$$

where $\mathbf{H}_{\text{static}}$ corresponds to quasi-static elements such as walls and furniture, and $\mathbf{H}_{\text{dynamic}}(t)$ captures time-varying contributions caused by moving objects and human activities. While communication systems primarily aim to mitigate the effects of $\mathbf{H}_{\text{dynamic}}(t)$, sensing systems intentionally exploit these variations.
The frequency dimension of the CFR provides sensitivity to fine-grained delay changes, while the spatial dimension arising from multiple antennas enables angular and spatial discrimination. Additionally, the temporal evolution of the CFR reveals motion patterns, enabling the differentiation of activities with distinct dynamics, such as static postures, low-mobility gestures, and high-mobility actions. Importantly, CFR-based sensing does not require explicit knowledge of object locations or scene geometry. Instead, sensing is performed implicitly by learning or extracting patterns from how the channel responds to environmental changes. This makes CFR a versatile and device-free sensing primitive that operates without line-of-sight constraints or dedicated sensing hardware. Another practical advantage of CFR as a sensing primitive is its availability in commodity Wi-Fi systems. CFR estimates are routinely obtained during channel sounding for communication purposes, making them readily accessible without modifying the underlying protocol. This allows sensing functionality to be integrated into existing Wi-Fi deployments with minimal additional overhead.

In summary, the CFR provides a compact yet information-rich representation of the wireless propagation environment. Its sensitivity to multipath structure, spatial configuration, and temporal dynamics makes it a natural foundation for Wi-Fi-based sensing. By appropriately modeling and processing CFR measurements, it becomes possible to infer meaningful physical-world information, thereby enabling Integrated Sensing and Communication within standard Wi-Fi systems.

## 2.2 Related Works

In the last two decades, research community has emerged in the mighty sector of smart wireless technology and explored the theory of integrated sensing and communication. The dual role of the wireless channel as both a communication medium and a sensing signal has motivated a rapidly growing body of research at the intersection of wireless communications, signal processing, and machine learning. Building upon the understanding of CFR as a sensing primitive, existing works can be broadly categorized into two complementary directions: communication-aided sensing and sensing-aided communication. These two lines of research reflect different perspectives on how channel measurements can be exploited within ISAC systems.

### 2.2.1 Communication-Aided Sensing

Communication-aided sensing represents the earliest and most extensively studied paradigm in Wi-Fi sensing, where wireless communication signals—originally designed for data transmission—are repurposed to infer information about the surrounding physical environment[5]. The central premise is that human motion and object interactions alter the wireless propagation characteristics, and these alterations can be captured through channel measurements obtained during standard communication procedures. Early efforts in this domain relied primarily on coarse-grained metrics such as the Received Signal Strength Indicator (RSSI)[33][34][35][36][37][38]. While RSSI-based approaches demonstrated the feasibility of device-free sensing, their limited spatial and frequency resolution significantly constrained sensing accuracy and robustness. As Wi-Fi systems evolved toward multi-carrier and multi-antenna architectures, research attention shifted toward CSI, which provides fine-grained amplitude and phase measurements across subcarriers and antenna pairs. CSI-based sensing has since enabled a wide range of applications, including human activity and gesture recognition[39]-[40], respiration monitoring[32], localization[41], and health sensing[42]-[43]. Representative systems such as $WiAR$[44] demonstrated the ability to recognize complex human activities using CSI measurements extracted from IEEE 802.11n devices. Similarly, SignFi[45] leveraged CSI phase information to perform fine-grained gesture recognition. These works established CSI as a powerful sensing primitive due to its sensitivity to multipath dynamics.

Despite their success, CSI-based approaches suffer from fundamental practical limitations. CSI is computed at the physical layer and is not exposed by standard Wi-Fi interfaces. Consequently, most existing systems require firmware modifications or specialized extraction tools[46],[47], such as Linux CSI[23], Atheros CSI[48], Nexmon CSI[49], or AX-CSI[50]. These tools are hardware-specific, difficult to maintain, and often restricted to single-user MIMO operation, thereby preventing exploitation of multi-user spatial diversity. Furthermore, CSI extraction introduces additional processing overhead and energy consumption, limiting scalability in real-world deployments. To address environmental and subject variability, several works introduced learning-based generalization mechanisms. Widar 3.0[51] proposed a body velocity profile derived from CSI Doppler shifts to improve robustness across environments. OneFi[40] employed one-shot learning to recognize unseen gestures with minimal labeled Wi-Fi CSI data. Transfer learning–based systems such as Wi-Transfer[52] further demonstrated improved adaptability across sensing domains. However, these approaches still fundamentally rely on CSI extraction and remain constrained by its hardware and scalability limitations.

More recently, sensing approaches leveraging IEEE 802.11ac CSI have emerged. For example, SHARP[39] is an environment- and person-independent activity recognition framework that uses commercial WiFi devices. It utilized micro-Doppler signatures extracted from CSI to enhance domain generalization, while ReWiS[46] exploited spatial diversity by combining CSI from multiple receivers. Although these systems improved sensing robustness, they still require firmware-level access and cannot naturally support multi-user MIMO channel observations. An emerging alternative within communication-aided sensing is the use of beamforming feedback information (BFI), which is exchanged during standard Wi-Fi channel sounding procedures. Unlike CSI, BFI is transmitted over the air in compliance with IEEE 802.11 standards and can be captured without firmware modification. Recent studies have explored reconstructing BFI from compressed beamforming feedback to enable sensing tasks such as localization[53], respiration monitoring, and passive tracking[54]. For example, Bi-Directional Beamforming for WiFi Sensing, Respiratory Rate Monitoring Using WiFi Beamforming Feedback, and Device-Free Localization and Activity Recognition Using WiFi Beamforming Information demonstrated that BFI retains meaningful spatial and temporal information for sensing. However, most existing BFI-based systems reconstruct full channel matrices from compressed feedback, introducing additional computational complexity and latency. In contrast, the BeamSense[25] framework demonstrates that sensing can be directly performed on compressed beamforming feedback angles (BFAs), eliminating reconstruction overhead while simultaneously capturing spatially diverse channels from multiple users in mu-MIMO transmissions. This transition from CSI-centric to feedback-centric sensing marks a significant step toward scalable, standard-compliant communication-aided sensing.

Overall, prior work in communication-aided sensing has firmly established the feasibility of leveraging Wi-Fi communication signals for environmental perception. At the same time, the limitations of both BFI and CSI-based methods motivate the development of new sensing frameworks that align more closely with modern Wi-Fi standards—an evolution that directly informs the design philosophy adopted in this thesis. See Table 2.2 for a complete overview and a comprehensive comparison of communication aided sensing related previous works.

### 2.2.2 Sensing-Aided Communication

While communication-aided sensing exploits communication signals to infer environmental dynamics, sensing-aided communication adopts the reverse perspective by using sensed channel structure to enhance communication efficiency and reliability. In this paradigm, the wireless channel is treated not merely as an unknown impairment to be estimated, but as a structured physical object whose properties—such as sparsity, correlation, and temporal consistency—can be sensed and exploited to reduce overhead and improve link performance.
Early work in this direction focused on model-based channel inference, where physical properties of wireless propagation were explicitly leveraged to reduce estimation and feedback costs. In particular, compressed sensing–based channel estimation exploited sparsity in the delay and angular domains to reconstruct channels from a reduced number of pilots[55]. While this approach ensures accurate channel state

acquisition, it introduces non-negligible airtime overhead that grows with the number of antennas, users, and operating bandwidth[56],[57]. As a result, the practical throughput gains of mu-MIMO often fall short of theoretical expectations.

To overcome these limitations, recent research has shifted toward learning-based sensing-aided communication, where channel estimation and feedback reduction are formulated as data-driven inference problems. Instead of estimating channel coefficients independently on each subcarrier, neural models are trained to sense global channel structure across frequency, space, and time, enabling reconstruction of high-resolution CFR from sparse or compressed measurements. This line of work treats channel reconstruction as a sensing task, where the goal is to infer the latent propagation structure that generated the observed measurements.

A substantial body of work has investigated learning-based feedback compression techniques. Neural autoencoder architectures[58] have been proposed to compress estimated channel state information at the receiver and reconstruct it at the transmitter, thereby reducing feedback size while preserving precoding accuracy. Extensions of this idea[59] incorporate side information from the Wi-Fi protocol stack to enable online training and protocol-compatible compression. Other approaches adopt split neural network designs[60], placing lightweight compression modules at the receiver and more complex reconstruction models at the transmitter to reduce computational burden. Adaptive quantization schemes[61] have also been explored to dynamically trade off feedback precision against overhead. Despite their effectiveness in reducing control bits, these methods largely treat channel compression as a static encoding problem and remain agnostic to temporal channel dynamics.

In parallel, a smaller set of works has examined sounding rate adaptation, where the system decides when channel updates are necessary based on coarse channel statistics or throughput estimates. Representative approaches[62] trigger sounding when channel variance exceeds predefined thresholds or when estimated throughput degrades. These strategies rely on indirect indicators of channel quality and have primarily been evaluated through simulations, emulation platforms, or software-defined radios rather than commodity Wi-Fi hardware.

More recent efforts have highlighted the benefits of directly sensing channel evolution to guide sounding decisions. By leveraging instantaneous channel estimates derived from training packets and recent reception outcomes, these approaches enable more fine-grained and responsive adaptation to channel dynamics[56],[63]. A representative example is SHRINK[64], which senses short-term channel evolution at the station to adapt sounding behavior on commodity IEEE 802.11 devices. This sensing-driven view of channel sounding demonstrates that channel feedback overhead can be reduced not only by compressing feedback, but also by exploiting the intrinsic temporal structure of wireless channels.

Complementary to feedback reduction in the time domain, recent work has explored sensing-driven channel inference across frequency bands. These approaches[65], [66], [67] observe that channels measured at different frequency bands exhibit strong structural correlations due to shared propagation paths. Another work BandWeave[68] exemplifies this direction by leveraging cross-band structural correlations to reconstruct high-resolution channel representations from sparse multi-band observations. By learning these correlations, it becomes possible to infer high-resolution channel information in one band using sparse or partial measurements from another. This multi-band sensing perspective re-frames channel estimation as an inference problem over latent propagation structure rather than an isolated per-band estimation task. Such methods are particularly relevant in modern Wi-Fi systems operating over fragmented spectrum, where sensing-assisted reconstruction can significantly reduce pilot and feedback overhead.

Building on this insight, the sensing-aided communication paradigm adopted in this thesis goes beyond adjusting sounding frequency or compressing raw feedback. Instead, it leverages learned structural patterns in the channel frequency response to reconstruct high-resolution channel representations from sparse or imperfect measurements. This perspective treats channel reconstruction itself as a sensing problem—inferring latent propagation characteristics rather than explicitly reporting full channel state—thereby enabling more scalable and efficient feedback mechanisms suited for next-generation Wi-Fi systems operating over wide and fragmented spectrum. See Table 2.3 for a comprehensive overview and comparison of sensing aided communication related prior works.

### 2.2.3 Discussion and Research Gap

Taken together, existing work demonstrates the strong potential of using channel measurements for both sensing and communication enhancement. However, the two research directions have largely evolved in parallel, with limited integration between them. Communication-aided sensing systems often prioritize sensing accuracy without considering communication constraints, while sensing-aided communication approaches focus on link performance without exploiting the rich temporal dynamics induced by physical motion.

This separation motivates a unified perspective in which sensing and communication are treated as mutually reinforcing processes. By jointly considering the temporal evolution, structural properties, and physical interpretability of CFR measurements, it becomes possible to design ISAC frameworks that are both efficient and robust. Such a perspective forms the foundation for the approaches developed in this thesis, which aim to bridge the gap between communication-aided sensing and sensing-aided communication within practical Wi-Fi systems.

Table 2.2: Overview of the communication aided sensing related works

| **Work** | **Wi-Fi Primitive** | **Target Application** | **Core Methodology** | **Key Limitations** |
|---|---|---|---|---|
| Hsieh *et al.* | RSSI | Device-free sensing | Statistical RSSI variation analysis | Coarse granularity, low robustness |
| Wang *et al.* | RSSI | Crowd estimation | RSSI fluctuation modeling | Sensitive to interference |
| Depatla *et al.* | RSSI | Crowd counting | RSSI-based spatial modeling | Poor scalability |
| Singh *et al.* | RSSI | Activity recognition | Machine learning on RSSI | Limited spatial resolution |
| WiAR | CSI | Human activity recognition | CNN on CSI amplitude | Environment dependent |
| SignFi | CSI | Gesture recognition | CSI phase exploitation | Sensitive to phase noise |
| Widar 3.0 | CSI | Cross-domain activity recognition | Doppler-based body velocity profile | Limited motion diversity |
| OneFi | CSI | One-shot gesture recognition | Meta-learning | Requires CSI extraction |
| Wi-Transfer | CSI | Domain adaptation | Transfer learning | High data dependency |
| SHARP | CSI | Environment-independent sensing | Micro-Doppler features | Firmware-level CSI access |
| ReWiS | CSI | Multi-receiver sensing | Spatial CSI fusion | Increased system complexity |
| Wu *et al.* | BFI | Localization | BFI reconstruction + learning | Reconstruction overhead |
| Liu *et al.* | BFI | Respiration monitoring | Beamforming feedback analysis | Limited scalability |
| BeamSense | BFA | Multi-user sensing | Direct learning on BFAs | Feedback resolution trade-off |
| **PULSE (Proposed)** | **CFR** | **Multi-user domain adaptive sensing** | **Learning CFR representation with temporal evolution** | |

Table 2.3: Overview of Sensing-Aided Communication Related Works

| Work | System | Objective | Core Methodology | Key Limitations |
|---|---|---|---|---|
| Zhu *et al.* | mmWave MIMO-OFDM | Channel structure inference | Gridless compressed sensing for delay/angle estimation | Strong sparsity assumptions, high sensitivity to noise |
| Wen *et al.* | MIMO-OFDM | Multipath parameter sensing | Multi-way compressive sensing exploiting tensor structure | Limited robustness in dynamic environments |
| CsiNet | Massive MIMO / Wi-Fi-like | CSI feedback compression | CNN-based autoencoder sensing spatial-frequency correlations | Static encoding, limited temporal awareness |
| LB-SciFi | IEEE 802.11ac | Feedback overhead reduction | Online learning-based CSI compression using protocol side information | Requires continuous retraining |
| SplitBeam | IEEE 802.11ac | Beamforming efficiency | Split neural inference across STA and AP | Increased system complexity |
| AIML-Quantization | Wi-Fi MU-MIMO | Feedback precision control | Adaptive ML-based quantization of feedback | No explicit channel dynamics modeling |
| MUTE | IEEE 802.11n/ac | Sounding rate sensing | Trigger sounding via channel variance thresholds | Coarse sensing indicators |
| Ma *et al.* | IEEE 802.11ac | Throughput-aware sounding | Sounding triggered by sensed throughput degradation | Indirect channel awareness |
| SHRINK | IEEE 802.11ac (COTS) | Temporal channel evolution sensing | Data-driven sensing of short-term CFR dynamics at STA | Focused on rate adaptation only |
| Li *et al.* | Multi-band Wi-Fi | Cross-band channel inference | Frequency-domain splicing for ranging | Limited generalization |
| Tian *et al.* | Multi-band OFDM | Super-resolution sensing | Multi-band ToF fusion | Application-specific |
| mmSplicer | mmWave multi-band | Channel reconstruction | Experimental multi-band channel splicing | Hardware-intensive |
| BandWeave | Wi-Fi multi-band MIMO | Structural channel sensing | Learning-based fusion of multi-band CFR structure | Requires multi-band availability |
| **ELF (Proposed** | **IEEE 802.11ac/ax** | **Low-overhead channel feedback** | **Embedding-guided subchannel sensing and selective feedback of representative embeddings** | **Performance tied to embedding quality** |

# Chapter 3

# PULSE: Communication-Aided Sensing

This chapter contains the complete design, methodology, and evaluation of PULSE , the proposed communication-aided Wi-Fi sensing framework. It begins by outlining the practical challenges that motivate PULSE , then details the end-to-end system pipeline—including temporal feature extraction from CFR, lightweight embedding and activity learning, and few-shot adaptation for domain generalization. The chapter concludes with a thorough experimental study on the *BeamSense* dataset, benchmarking sensing accuracy, cross-domain robustness, and computational efficiency against state-of-the-art baselines before presenting the final performance insights.

## 3.1 Challenges and Motivation

As discussed in the previous chapters, Wireless sensing has emerged as a promising paradigm for enabling non-intrusive and privacy-preserving perception of human activity and environmental dynamics. Unlike vision- or wearable-based sensing systems, wireless sensing leverages existing RF signals and infrastructure, allowing sensing to be performed without requiring the subject to carry devices or remain within line-of-sight of dedicated sensors. In Wi-Fi systems, sensing is typically achieved by observing variations in the wireless channel caused by human motion and interaction with the environment. These variations manifest through changes in multipath propagation, where reflections, scattering, and attenuation alter the channel response over time. Despite its strong potential, practical Wi-Fi sensing faces several fundamental challenges that limit its deployment in real-world systems—especially on resource-constrained edge devices. These challenges are tightly connected to how channel information is represented, processed, and learned.

### 3.1.1 High-Dimensional Channel Measurements and Edge Constraints

Modern Wi-Fi sensing relies heavily on CFR or CSI, which are inherently high-dimensional. A single Wi-Fi packet may contain channel measurements across hundreds of subcarriers and multiple antenna pairs. When collected over time, these measurements form large tensors that must be processed at fine temporal granularity to capture dynamic human activity.

This high dimensionality poses a major obstacle for real-time sensing on edge devices. Embedded platforms and commodity access points have limited computation capability, memory, and energy budgets. Processing raw CFR tensors using deep neural networks—particularly convolutional or attention-based models—can quickly become impractical. As a result, many existing systems either offload computation to powerful servers or aggressively reduce model complexity, often at the cost of sensing accuracy and robustness.

### 3.1.2 Limitations of Lightweight Wi-Fi Sensing Models

To address edge constraints, several lightweight Wi-Fi sensing models have been proposed. These approaches typically reduce overhead by simplifying the input representation, discarding parts of the channel information, or compressing features through heuristic selection. While such strategies lower computational cost, they often suffer from degraded reliability, especially in dynamic environments.

In practice, subtle human activities—such as low-mobility gestures, fine-grained motions, or slow gait patterns—produce weak perturbations in the channel that are easily masked when aggressive dimensionality reduction is applied. As a result, lightweight models may fail to distinguish between classes that differ primarily in temporal dynamics rather than spatial magnitude. This trade-off between efficiency and discriminability remains a central challenge in Wi-Fi sensing. Even recent approaches that aim to preserve discriminative power through adaptive subchannel selection or attention mechanisms often still rely on the full CFR as input and employ computationally expensive modules such as deep Convolutional Neural Network (CNN) stacks or spatial attention blocks. While effective in controlled settings, these designs remain ill-suited for deployment on low-power edge devices and do not fully resolve the efficiency–robustness tension.

### 3.1.3 Sensitivity to Environment and Motion Diversity

Another major challenge in Wi-Fi sensing is environment dependence. The wireless channel is highly sensitive to room geometry, furniture placement, building materials, and antenna configuration. As a result, sensing models trained in one environment often exhibit significant performance degradation when deployed in a new space or with different users.

This issue is exacerbated by the fact that many models implicitly rely on environment-specific spatial patterns in the CFR. When these patterns change, the learned representations no longer align with the sensing task. Some works attempt to mitigate this problem using phase sanitization, Doppler-based features, or domain adaptation techniques. While these approaches improve robustness to some extent, they often focus on specific types of motion or require complex preprocessing pipelines that increase computational overhead. Moreover, human activities exhibit diverse mobility characteristics. High-mobility activities generate strong and rapid channel variations that are easier to detect, whereas low-mobility activities induce smoother and more subtle temporal changes. Many sensing systems are biased toward high-mobility classes and struggle to maintain accuracy across a broad range of motion dynamics.

### 3.1.4 Underutilization of Temporal Channel Structure

A key observation motivating this work is that most Wi-Fi sensing approaches treat CFR primarily as a spatial–frequency object, emphasizing per-packet channel snapshots or spectrogram-like representations. While effective in some scenarios, this perspective underutilizes the rich temporal structure inherent in wireless channels.

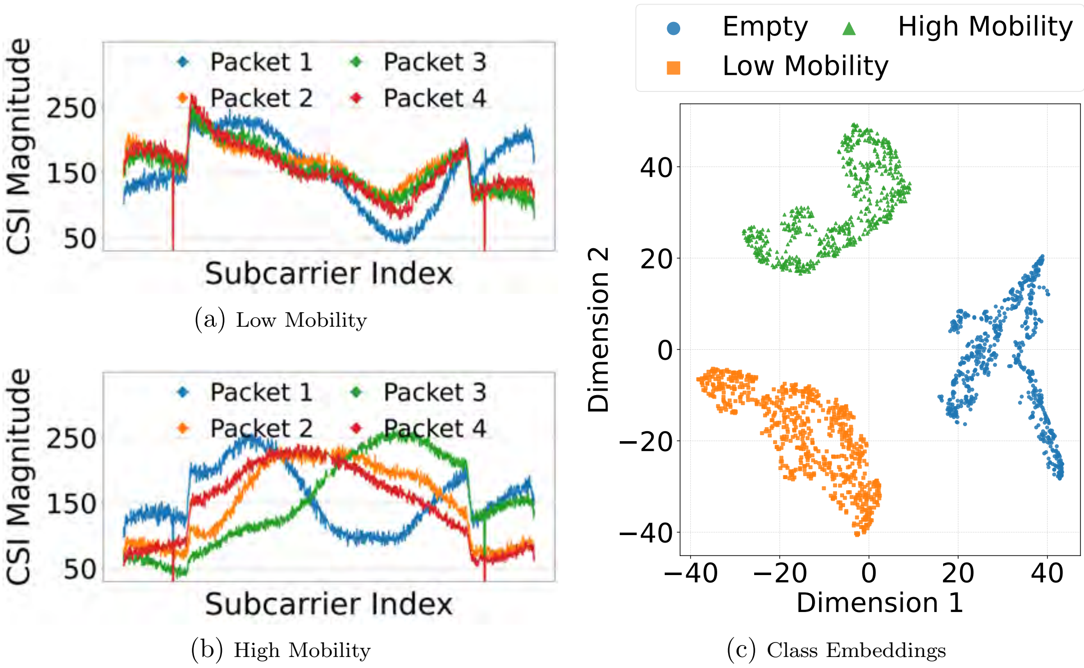



Figure 3.1: Temporal evolution of CFR for different classes of activities.

From a physical standpoint, human motion primarily affects how the channel evolves over time rather than instantaneously reshaping its full spatial structure. Temporal consistency, smooth transitions, abrupt fluctuations, and short-term stability patterns all carry meaningful information about activity type and intensity. However, when CFR is processed as a sequence of high-dimensional frames, these temporal signatures are often obscured by noise, redundant features, and heavy model architectures. This mismatch between the physical origin of sensing information and the way it is represented in learning models motivates a reevaluation of how temporal dynamics should be captured and exploited in Wi-Fi sensing.

### 3.1.5 Motivation for a Physics-Aware, Temporal-First Perspective

The challenges discussed above highlight a central gap in existing Wi-Fi sensing systems: the lack of representations that are simultaneously lightweight, discriminative, interpretable, and robust across environments. Addressing this gap requires moving away from brute-force processing of raw CFR tensors and toward representations that align more closely with the physics of wireless propagation.

In particular, if sensing-relevant information can be captured through a compact set of physics-aware temporal descriptors—reflecting how the channel evolves rather than its full instantaneous structure—then it becomes possible to significantly reduce input dimensionality and computational cost without sacrificing accuracy. Such a representation would naturally support edge deployment, improve robustness to environment changes, and enable efficient learning of discriminative patterns across diverse sensing tasks. This motivation directly leads to the design philosophy of PULSE , introduced in the following sections. PULSE is built on the premise that the temporal evolution of the wireless channel encodes sufficient and transferable information for sensing, and that extracting and learning from this temporal structure provides a principled solution to the efficiency, robustness, and generalization challenges identified in earlier chapters.

## 3.2 PULSE System Overview

In IEEE 802.11ac and IEEE 802.11ax Wi-Fi systems, channel knowledge at the transmitter is obtained through a standardized channel sounding procedure. During this process, the multi-antenna transmitter, referred to as the *beamformer*, periodically transmits NDP frames that contain known LTF. Each receiving device, known as the *beamformee*, uses these training fields to estimate the CFR, denoted as $\mathbf{H} \in \mathbb{C}^{K \times M \times N}$, where $K$ represents the number of subcarriers, and $M$ and $N$ denote the numbers of transmit and receive antennas, respectively. The estimated CFR characterizes the frequency-selective propagation between each transmit–receive antenna pair. After quantization, this information is fed back to the beamformer and is primarily used for precoding and beamforming in communication.

Beyond its traditional role in communication, the CFR inherently encodes rich information about the surrounding physical environment. Variations in multipath propagation caused by human motion, object interaction, and environmental dynamics manifest as structured changes in the CFR over time. As discussed in earlier chapters, these temporal variations are not arbitrary; rather, they reflect the underlying physical processes governing wireless signal propagation. Consequently, continuous CFR streams collected during normal Wi-Fi operation provide a powerful sensing signal that can be reused without introducing additional sensing-specific transmissions.

In this thesis, CFR streams obtained through standard Wi-Fi channel sounding serve as the raw input for communication-aided sensing. Building on this observation, we propose PULSE , a framework that transforms raw CFR measurements into compact, physically interpretable temporal embeddings that enable efficient and adaptive wireless sensing. Instead of operating directly on high-dimensional CFR matrices—whose size scales with bandwidth and antenna count—PULSE focuses on modeling how the wireless channel evolves over time. This temporal perspective provides a representation that is both more robust to environmental variability and substantially more efficient for real-time inference on resource-constrained devices. Figure 3.2 illustrates the overall architecture of the PULSE framework. It consists of three sequential stages. First, temporal features are extracted from CFR streams to capture motion-induced channel dynamics in a compact and physically meaning-

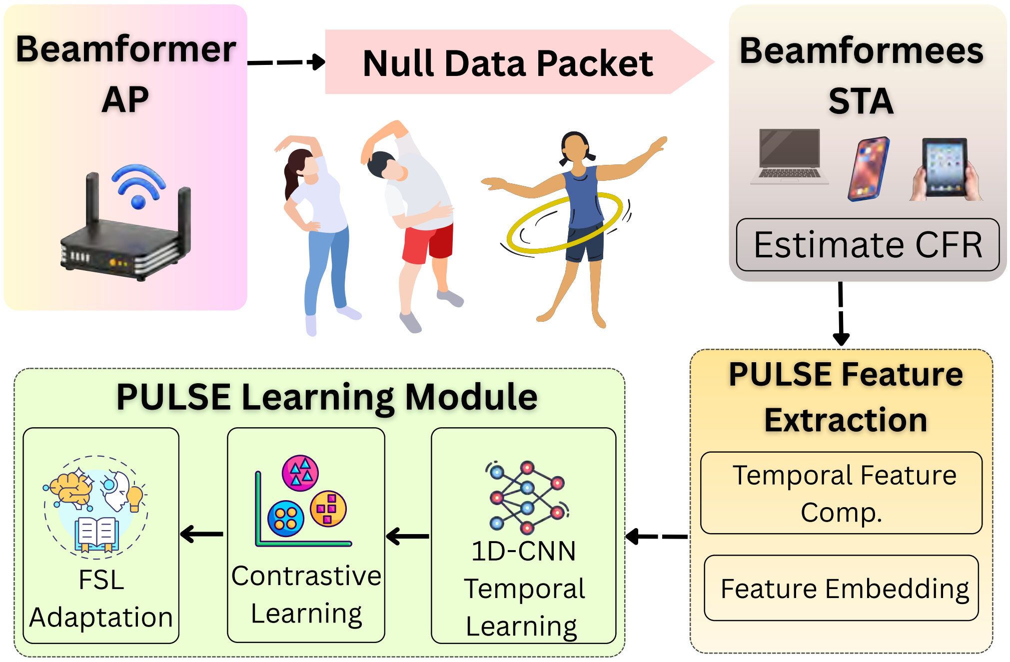



Figure 3.2: System Overview of the Proposed PULSE Wireless Sensing Framework

ful form. Second, these temporal features are mapped into a low-dimensional latent embedding space using a lightweight one-dimensional convolutional neural network (1D CNN), which jointly learns discriminative representations and performs activity inference. Finally, a few-shot domain adaptation mechanism is employed to improve robustness and generalization across unseen environments and subjects, enabling reliable sensing with minimal labeled data.

The following subsections describe each of these stages in detail, highlighting the design choices that allow PULSE to achieve high sensing accuracy while maintaining low computational complexity and strong cross-domain generalization.

### 3.2.1 Temporal Feature Extraction from CFR

The first stage of PULSE transforms raw Wi-Fi Channel Frequency Response (CFR) measurements into a compact set of temporally structured descriptors that characterize *how the wireless channel evolves over time*. This stage is essential because raw CFR measurements are (i) high-dimensional across subcarriers and antenna pairs, and (ii) heavily affected by hardware-induced distortions such as carrier frequency offset (CFO), sampling frequency offset (SFO), phase noise, and clock drift. These impairments often dominate the raw signal and can obscure the comparatively subtle motion-induced variations that are relevant for sensing.
Rather than attempting to reconstruct or sanitize the full channel, PULSE adopts a structure-aware strategy: it extracts *simple, interpretable, and stable* temporal statistics that preserve motion-driven channel dynamics while significantly reducing dimensionality and sensitivity to device artifacts. This design is consistent with both practical Wi-Fi constraints and the lightweight implementation of the PULSE system.

### CFR Acquisition and Representation

Wi-Fi channel sounding provides an estimate of the complex CFR

$$\mathbf{H}(t) \in \mathbb{C}^{K\times M\times N}, \qquad t = 1, \dots, T, \tag{3.1}$$

where $K$ is the number of OFDM subcarriers, $M$ and $N$ are the numbers of transmit and receive antennas, and $t$ indexes packet-level sounding instances. For sensing, PULSE operates on a *temporal stream* of CFR snapshots rather than treating each snapshot independently.

To ensure robustness across commodity hardware platforms, PULSE aggregates CFR measurements across antenna pairs prior to temporal analysis. For each subcarrier $k$, aggregated magnitude and phase traces are computed as

$$A_k(t) = \mathrm{Agg}_{m,n}(|H_{k,m,n}(t)|), \qquad \Phi_k(t) = \mathrm{Agg}_{m,n}(\angle H_{k,m,n}(t)), \tag{3.2}$$

where $\mathrm{Agg}(\cdot)$ denotes a lightweight aggregation operator (mean in the implementation).

To mitigate phase wrapping, phase sequences are unwrapped along the temporal dimension:

$$\widetilde{\Phi}_k(t) = \mathrm{unwrap}(\Phi_k(t)). \tag{3.3}$$

Importantly, PULSE focuses on temporal statistics and variations of phase rather than absolute phase values, which are known to be unstable on commodity Wi-Fi devices.

### Subchannel Grouping

Raw CFR consists of $F$ subchannels, many of which are strongly correlated due to shared multipath components. To reduce redundancy and computational cost, PULSE partitions the subchannel into $G$ contiguous frequency bands

$$\{\mathcal{K}_g\}_{g=1}^{G}, \qquad \bigcup_{g=1}^{G} \mathcal{K}_g = \{1, \dots, K\}, \tag{3.4}$$

with $|\mathcal{K}_g| \approx K/G$.

Band grouping provides dimensionality reduction while averaging out narrowband fading effects, allowing motion-induced trends that span multiple subcarriers to remain visible. The number of bands $G$ controls the trade-off between frequency resolution and computational efficiency.

### Temporal Statistical Descriptors

For each band $g$ and time index $t$, PULSE computes a set of lightweight temporal statistics that summarize both instantaneous frequency-domain behavior and short-term temporal evolution. Let

$$\mathbf{A}_g(t) = \{A_k(t)\}_{k\in\mathcal{K}_g}, \qquad \mathbf{\Phi}_g(t) = \{\widetilde{\Phi}_k(t)\}_{k\in\mathcal{K}_g}. \tag{3.5}$$

**Magnitude distribution descriptors.**

$$f_{1,g}(t) = \text{mean}\big(\mathbf{A}_g(t)\big), \tag{3.6}$$
$$f_{2,g}(t) = \text{std}\big(\mathbf{A}_g(t)\big), \tag{3.7}$$
$$f_{3,g}(t) = \text{median}\big(\mathbf{A}_g(t)\big), \tag{3.8}$$
$$f_{4,g}(t) = p_{75}\big(\mathbf{A}_g(t)\big) - p_{25}\big(\mathbf{A}_g(t)\big). \tag{3.9}$$

**Temporal magnitude variation.**

$$f_{5,g}(t) = f_{1,g}(t) - f_{1,g}(t-1). \tag{3.10}$$

**Phase evolution descriptors.**

$$f_{6,g}(t) = \text{mean}\big(\mathbf{\Phi}_g(t)\big), \tag{3.11}$$
$$f_{7,g}(t) = \text{std}\big(\mathbf{\Phi}_g(t)\big), \tag{3.12}$$
$$f_{8,g}(t) = f_{6,g}(t) - f_{6,g}(t-1). \tag{3.13}$$

**Energy descriptor.**

$$f_{9,g}(t) = \log\big(\varepsilon + \text{mean}\big(\mathbf{A}_g(t)^2\big)\big)\,, \tag{3.14}$$

where $\varepsilon$ is a small constant for numerical stability.
For each band $g$, these statistics form the feature vector

$$\mathbf{f}_g(t) = [f_{1,g}(t), \ldots, f_{9,g}(t)]^\top \in \mathbb{R}^9. \tag{3.15}$$

Concatenating across all $G$ bands yields

$$\mathbf{f}(t) = \left[\mathbf{f}_1(t)^\top, \ldots, \mathbf{f}_G(t)^\top\right]^\top \in \mathbb{R}^C, \qquad C = 9G. \tag{3.16}$$

Stacking over time produces the temporal feature matrix

$$\mathbf{F} = [\mathbf{f}(1), \mathbf{f}(2), \ldots, \mathbf{f}(T)] \in \mathbb{R}^{C\times T}. \tag{3.17}$$

**Temporal Windowing and Normalization**

To generate learning samples, PULSE segments $\mathbf{F}$ into fixed-length temporal windows. Let $W$ denote the window length and $S$ the stride. The $i$-th window is defined as

$$\mathbf{X}_i = \mathbf{F}(:,\ (i-1)S + 1 : (i-1)S + W) \in \mathbb{R}^{C\times W}. \tag{3.18}$$

Each window corresponds to a short temporal segment capturing localized motion patterns.
Finally, per-channel normalization is applied using training-set statistics:

$$\widetilde{\mathbf{X}}_i = (\mathbf{X}_i - \boldsymbol{\mu}) \oslash \boldsymbol{\sigma}, \tag{3.19}$$

where $\boldsymbol{\mu}, \boldsymbol{\sigma} \in \mathbb{R}^C$ are the channel-wise mean and standard deviation.

**Output of Stage I.** The output of this stage is a set of normalized temporal windows

$$\{\widetilde{\mathbf{X}}_i\}_{i=1}^{N}, \qquad \widetilde{\mathbf{X}}_i \in \mathbb{R}^{C\times W}, \tag{3.20}$$

which serve as input to the subsequent temporal embedding and classification module of PULSE .

### 3.2.2 Feature Embedding and Activity Learning with 1D CNN

After transforming raw CFR measurements into compact temporal feature representations, the next stage of PULSE focuses on learning discriminative temporal embeddings and performing activity inference. This stage is designed with a clear emphasis on efficiency, robustness, and deployability, reflecting the real-world constraints of Wi-Fi sensing systems discussed in earlier chapters. Rather than relying on deep or computationally intensive architectures, PULSE adopts a lightweight one-dimensional convolutional neural network (1D CNN) that operates directly on temporally structured features.

#### Design Rationale for Lightweight 1D CNNs

The choice of a 1D CNN backbone in PULSE is driven by both physical intuition and practical considerations. Wireless sensing using CFR fundamentally involves understanding how the channel evolves over time. Human motion, gestures, and activities manifest as temporal patterns—such as smooth fluctuations, abrupt transients, or periodic variations—in the extracted temporal features. One-dimensional convolutions along the temporal axis are therefore well-suited to capture such patterns efficiently.

Compared to two-dimensional CNNs or Transformer-based architectures, 1D CNNs offer several advantages for edge-level Wi-Fi sensing. First, they significantly reduce computational complexity and memory usage, making them suitable for deployment on resource-constrained devices such as embedded platforms and edge gateways. Second, convolutional filters naturally act as local temporal pattern detectors, enabling the model to learn motion primitives without requiring long-range attention mechanisms. Finally, 1D CNNs provide a favorable balance between expressiveness and interpretability: the learned filters often correspond to physically meaningful temporal variations in the channel, such as rapid phase changes or gradual amplitude trends.

By coupling the CNN with temporally structured and dimension-reduced inputs produced by the temporal feature extraction stage, PULSE avoids processing raw high-dimensional CFR tensors. This design ensures that the learning model focuses on sensing-relevant dynamics while maintaining low latency and energy consumption.

#### Embedding Learning and Classification

Let $\mathbf{X} \in \mathbb{R}^{N\times C\times W}$ denote the set of temporal windows obtained after preprocessing, where $N$ is the number of samples, $C$ is the number of temporal feature channels, and $W$ is the window length. These inputs are normalized using statistics computed exclusively from the training set to ensure consistent scaling across splits:

$$\boldsymbol{\mu} = \mathrm{mean}_{n,t}(\mathbf{X}_{\mathrm{train}}), \qquad \boldsymbol{\sigma} = \mathrm{std}_{n,t}(\mathbf{X}_{\mathrm{train}}), \tag{3.21}$$

and normalization is applied as $\widetilde{\mathbf{X}} = (\mathbf{X} - \boldsymbol{\mu})/\boldsymbol{\sigma}$.

To improve robustness against minor timing misalignments and channel noise, each temporal window undergoes lightweight data augmentation during training. Specifically, random circular time shifts are applied along the temporal axis, and small Gaussian perturbations $\mathcal{N}(0, \sigma^2_{\mathrm{noise}})$ are added to the feature values. These augmentations preserve the underlying temporal structure while encouraging invariance to phase drift, packet timing jitter, and minor environmental variations.

The normalized and augmented inputs are passed through a CNN encoder $f_\theta(\cdot)$, which consists of a small number of temporal convolutional blocks with residual connections, nonlinear activations, and dropout regularization. The encoder produces two outputs:

$$\mathbf{z}, \mathbf{h} = f_\theta(\widetilde{\mathbf{x}}), \tag{3.22}$$

where $\mathbf{h} \in \mathbb{R}^D$ denotes a compact temporal embedding that captures high-level channel dynamics, and $\mathbf{z} \in \mathbb{R}^M$ represents the class logits for $M$ activity classes. In this unified design, embedding learning and classification are performed jointly, allowing the model to learn representations that are directly aligned with the sensing task.

### Training Objective and Regularization

The CNN is trained using supervised learning within each environment. By default, PULSE employs the cross-entropy loss with optional label smoothing to reduce overconfidence and improve generalization:

$$\mathcal{L}_{\mathrm{CE}} = -\sum_{k=1}^{M} q_k \log p_k, \qquad p_k = \frac{e^{z_k}}{\sum_j e^{z_j}}, \tag{3.23}$$

where $q_k$ denotes the smoothed target distribution. For scenarios with class imbalance or where hard-to-classify samples are critical, a focal loss variant can be used:

$$\mathcal{L}_{\mathrm{Focal}} = \left(1 - e^{-\mathrm{CE}(\mathbf{z},y)}\right)^\gamma \mathrm{CE}(\mathbf{z}, y), \tag{3.24}$$

with focusing parameter $\gamma > 0$.

Regularization plays an important role in ensuring stable and transferable embeddings. Dropout layers within the CNN mitigate overfitting, while early stopping based on validation loss prevents unnecessary model complexity. Training proceeds until the validation performance stops improving for a predefined patience interval, at which point the best-performing model parameters are restored for evaluation.

Through this temporal embedding and learning strategy, PULSE produces compact, discriminative, and physically meaningful representations of CFR dynamics. These embeddings form the foundation for the subsequent few-shot adaptation mechanism, enabling robust sensing performance across environments with minimal additional data.

### 3.2.3 Few-Shot Adaptation for Domain Generalization

A central challenge in Wi-Fi sensing systems is their limited ability to generalize across environments, subjects, and deployment conditions. Wireless channels are highly sensitive to physical context: small changes in room geometry, furniture placement, device orientation, or human motion patterns can significantly alter the observed channel frequency response. As a result, models trained in one environment often experience noticeable performance degradation when applied to unseen environments or new users. Retraining a sensing model from scratch for every deployment is impractical, as it requires collecting large labeled datasets and incurs substantial computational cost.

To address this challenge, PULSE adopts a *few-shot adaptation* strategy that enables efficient and stable transfer from a source environment to a target environment using only a small amount of labeled data. Rather than relearning the sensing model entirely, PULSE focuses on learning *transferable temporal embeddings* that capture domain-invariant channel dynamics. These embeddings can then be adapted to a new domain using simple, non-parametric mechanisms, avoiding overfitting and reducing adaptation overhead.

The few-shot adaptation mechanism in PULSE consists of three tightly coupled components: source-domain pretraining with contrastive supervision, prototype-based inference in the target domain, and adaptive normalization to compensate for distribution shift across environments.

#### Source-Domain Pretraining

The first step toward robust domain generalization is to learn a feature representation that is discriminative for sensing tasks while remaining resilient to environmental variation. To this end, PULSE performs source-domain pretraining using a lightweight temporal encoder based on a 1D convolutional neural network. The encoder processes normalized temporal windows $\widetilde{\mathbf{x}} \in \mathbb{R}^{C\times W}$, extracted from CFR streams as described in the previous sections, and maps them into a compact latent embedding space:

$$\mathbf{e} = f_\theta(\widetilde{\mathbf{x}}), \qquad \mathbf{e} \in \mathbb{R}^D, \tag{3.25}$$

where $D$ denotes the embedding dimension and $f_\theta(\cdot)$ represents the encoder network parameterized by $\theta$.

Rather than training the encoder solely with a standard classification loss, PULSE employs a joint training objective that combines cross-entropy loss with supervised contrastive learning. The motivation for this design is twofold. First, cross-entropy loss encourages embeddings to be separable at the class level, ensuring good classification performance in the source environment. Second, supervised contrastive loss explicitly structures the embedding space by pulling samples of the same class closer together while pushing samples of different classes farther apart. This geometric organization of the embedding space improves robustness to domain shifts, as class-related structure is preserved even when absolute channel characteristics change.

The total pretraining objective is given by:

$$\mathcal{L}_{\text{pre}} = \mathcal{L}_{\text{CE}} + \lambda_{\text{sc}}\,\mathcal{L}_{\text{SupCon}}, \tag{3.26}$$

where $\lambda_{\text{sc}}$ balances the contribution of the supervised contrastive term. The supervised contrastive loss is defined as

$$\mathcal{L}_{\text{SupCon}} = -\frac{1}{|\mathcal{I}|} \sum_{i \in \mathcal{I}} \frac{1}{|\mathcal{P}(i)|} \sum_{p \in \mathcal{P}(i)} \log \frac{\exp\big(\langle \mathbf{z}_i, \mathbf{z}_p \rangle / \tau\big)}{\sum_{a \in \mathcal{A}(i)} \exp\big(\langle \mathbf{z}_i, \mathbf{z}_a \rangle / \tau\big)}, \tag{3.27}$$

where $\mathbf{z}$ denotes the $\ell_2$-normalized projection of $\mathbf{e}$, $\mathcal{P}(i)$ is the set of positive samples sharing the same label as sample $i$, $\mathcal{A}(i)$ denotes all samples in the batch except $i$, and $\tau$ is a temperature parameter controlling the sharpness of similarity weighting. By optimizing this objective, the encoder learns embeddings that reflect intrinsic temporal patterns associated with sensing classes rather than environment-specific artifacts. Importantly, this pretraining stage is performed only once on a source domain, after which the encoder parameters are fixed for deployment and adaptation.

### Prototype-Based Few-Shot Inference

Once the encoder is pretrained, PULSE performs domain adaptation in a new target environment using a small number of labeled samples per class. This setting is referred to as $K$-shot adaptation, where only $K$ labeled examples are available for each sensing class. The target-domain data are divided into a *support set* $\mathcal{S}$ and a *query set* $\mathcal{Q}$. The support set contains the limited labeled samples used for adaptation, while the query set is used for evaluation.
Instead of fine-tuning the encoder with backpropagation—which can be unstable with very limited data—PULSE adopts a prototype-based classification strategy. Each support sample is passed through the pretrained encoder to obtain its embedding. For each class $c$, a class prototype $\mathbf{P}_c$ is computed as the mean of the embeddings belonging to that class:

$$\mathbf{P}_c = \frac{\Big(\frac{1}{|\mathcal{S}_c|} \sum_{(\mathbf{x},y) \in \mathcal{S}_c} \mathbf{e}(\mathbf{x})\Big)}{\Big\| \frac{1}{|\mathcal{S}_c|} \sum_{(\mathbf{x},y) \in \mathcal{S}_c} \mathbf{e}(\mathbf{x}) \Big\|_2}. \tag{3.28}$$

These prototypes serve as compact class representatives in the embedding space. A query sample $\mathbf{x}_q$ is classified by comparing its embedding to each class prototype using cosine similarity:

$$\hat{y}(\mathbf{x}_q) = \arg\max_c \Big\langle \frac{\mathbf{e}(\mathbf{x}_q)}{\|\mathbf{e}(\mathbf{x}_q)\|_2},\ \mathbf{P}_c \Big\rangle. \tag{3.29}$$

This non-parametric inference mechanism has several advantages in the context of Wi-Fi sensing. First, it avoids overfitting, as no additional model parameters are learned during adaptation. Second, it is computationally efficient and suitable for edge deployment. Third, it naturally supports rapid adaptation when only a few labeled samples are available, making it practical for real-world deployments where collecting extensive labeled data is infeasible.

### Normalization Adaptation Across Domains

In addition to differences in channel dynamics, domain shifts often manifest as changes in signal scale and statistical distribution due to hardware imperfections,

calibration differences, or environmental attenuation. To mitigate this effect, PULSE incorporates an adaptive normalization strategy during few-shot adaptation.
Let $(\boldsymbol{\mu}_{\text{src}}, \boldsymbol{\sigma}_{\text{src}})$ denote the normalization statistics computed from the source-domain training data, and $(\boldsymbol{\mu}_{\text{sup}}, \boldsymbol{\sigma}_{\text{sup}})$ denote statistics computed from the target-domain support set. Rather than relying exclusively on either set of statistics, PULSE blends them using a convex combination:

$$\boldsymbol{\mu} = (1-\beta)\boldsymbol{\mu}_{\text{src}} + \beta\boldsymbol{\mu}_{\text{sup}}, \qquad \boldsymbol{\sigma} = (1-\beta)\boldsymbol{\sigma}_{\text{src}} + \beta\boldsymbol{\sigma}_{\text{sup}}, \tag{3.30}$$

where $\beta \in [0, 1]$ controls the degree of adaptation. A small $\beta$ emphasizes stability by favoring source-domain statistics, while a larger $\beta$ allows the model to better align with target-domain characteristics.
This normalization adaptation step plays a crucial role in stabilizing embeddings across domains. By aligning feature distributions before embedding extraction, PULSE reduces the impact of domain-dependent amplitude scaling and improves the reliability of prototype-based inference.

In summary, the few-shot adaptation mechanism in PULSE enables robust cross-domain sensing by combining contrastively pretrained temporal embeddings, lightweight prototype-based inference, and adaptive normalization. Together, these components allow PULSE to generalize effectively across environments and subjects using only a few seconds of labeled data, making it well-suited for practical Wi-Fi sensing deployments.

## 3.3 Experimental Dataset and Evaluation Benchmark

To rigorously evaluate the effectiveness, robustness, and deployability of the proposed PULSE framework, this thesis relies on a large-scale, real-world Wi-Fi sensing dataset rather than a controlled or synthetic collection. All experimental results presented in this chapter are obtained using the publicly available *BeamSense* dataset [25]. This dataset is specifically designed to support research on Wi-Fi-based sensing under realistic indoor conditions and therefore serves as a strong benchmark for validating the compatibility and practical relevance of PULSE .

The primary motivation for using the BeamSense dataset is twofold. First, it provides raw Wi-Fi channel measurements captured using commodity hardware operating under standard IEEE 802.11ac procedures, which closely aligns with the design assumptions of PULSE . Second, the dataset spans multiple environments, activities, and spatial viewpoints, making it well suited for evaluating temporal feature extraction and cross-domain generalization—two central objectives of the proposed framework.

### 3.3.1 Dataset Collection Setup

The BeamSense dataset was collected in three distinct indoor environments: a *kitchen*, a *living room*, and a *classroom*. These environments differ significantly in layout, size, furniture density, and surface materials, leading to diverse multipath propagation characteristics. Such diversity is essential for evaluating sensing frameworks intended for real-world deployment, where environmental conditions cannot be controlled or standardized. In each environment, three spatially separated Wi-Fi stations (STAs), denoted as *STA1*, *STA2*, and *STA3*, were deployed at different locations. This multi-STA setup introduces spatial diversity in channel observations, allowing the dataset to capture how the same activity manifests differently across multiple viewpoints. Importantly, each STA operates independently and captures its own channel measurements, enabling evaluation of sensing performance under both single-point and spatially diverse observations.

The data collection follows the IEEE 802.11ac mu-MIMO standard. The system operates on channel 153 at a center frequency of 5.77 GHz with an 80 MHz bandwidth. Channel measurements were obtained using modified Nexmon-enabled network interface cards, which expose *Channel Frequency Response* (CFR) information at the receiver. For each received frame, the CFR spans 242 subcarriers and captures the complex channel response of a $1 \times 1$ MIMO link, forming a tensor of size $1 \times 1 \times 242$. While the antenna configuration is simple, the wide bandwidth ensures rich frequency selectivity and sufficient temporal resolution to observe motion-induced channel dynamics.

### 3.3.2 Sensing Tasks and Activity Diversity

The dataset contains CFR recordings for twenty distinct human activities, covering a broad spectrum of motion dynamics. These activities include both highly dynamic motions, such as *jogging*, *walking*, *boxing*, and *hand waving*, as well as low-mobility or quasi-static activities, such as *reading a book*, *browsing a laptop*, *writing*, and *standing*. Additional fine-grained activities, including *brushing teeth*, *checking a wristwatch*, *phone calls*, and *drinking*, further increase the challenge by introducing subtle motion patterns that are easily masked by noise or multipath variations. Each activity was recorded for approximately 300 seconds per session, resulting in long continuous CFR traces. This extended duration enables temporal modeling over multiple time scales and allows the evaluation of both short-term discriminability and long-term stability. From the perspective of PULSE , this is particularly important because the framework focuses on extracting physics-aware temporal descriptors that capture how the channel evolves rather than relying on instantaneous CFR snapshots.

### 3.3.3 Relevance to PULSE Validation

The BeamSense dataset is well aligned with the goals of PULSE for several reasons. First, it provides access to raw CFR streams rather than preprocessed features, allowing PULSE to operate directly on channel measurements and validate its temporal feature extraction pipeline. Second, the presence of multiple environments and spatially separated STAs enables systematic evaluation of cross-domain generalization, where models trained in one environment are tested or adapted to another. This directly supports the few-shot adaptation strategy proposed in PULSE . Moreover, the dataset reflects realistic challenges encountered in Wi-Fi sensing, including device noise, multipath variability, and protocol-driven sampling irregularities. By validating PULSE on such a dataset, the experiments demonstrate not only accuracy under ideal conditions, but also robustness under practical constraints. Importantly, BeamSense has been used in prior state-of-the-art sensing systems, making it a recognized and credible benchmark within the community.

### 3.3.4 Baseline Methods for Comparison

To contextualize the performance of PULSE , comparisons are conducted against several representative sensing frameworks that operate on Wi-Fi channel measurements. These include *BeamSense* [25], which performs sensing using beamforming feedback and supports few-shot adaptation; *SignFi*[45], a CFR-based gesture recognition system that leverages phase information; and *OneFi* [40], a one-shot learning framework designed for cross-domain gesture recognition. These baselines collectively cover different design philosophies, ranging from raw CFR processing to few-shot learning, and provide a meaningful reference for evaluating the benefits of temporal modeling and lightweight embeddings in PULSE .

Overall, by leveraging the BeamSense dataset as a comprehensive experimental benchmark, this thesis ensures that the proposed PULSE framework is validated under realistic conditions and can be fairly compared against established Wi-Fi sensing approaches.

# Chapter 4

# ELF: Sensing-Aided Communication

This chapter contains the complete design, system model, and evaluation of ELF , the proposed sensing-aided channel feedback framework for scalable wideband Wi-Fi MIMO. It begins by identifying the key limitations of conventional channel estimation and feedback—including feedback overhead growth, UE-side computational burden, and redundancy-unaware reporting—then introduces the design philosophy of ELF . The chapter subsequently presents the end-to-end ELF pipeline, detailing UE-side normalization, latent projection, online representative subchannel selection, and quantized feedback formation, followed by BS-side dequantization and full-band CFR reconstruction. Finally, the chapter describes the standardized simulation environment and provides a comprehensive performance analysis, benchmarking BER, feedback overhead reduction, and computational efficiency across bandwidths, antenna configurations, and representative ratios.

## 4.1 Limitations of Conventional Channel Estimation

As discussed in the previous chapters, channel estimation and feedback form the backbone of multi-antenna wireless communication systems. In MIMO networks, the transmitter relies on accurate channel state information to design precoding matrices that spatially multiplex users and suppress interference. In Wi-Fi systems, this information is typically obtained through an explicit channel sounding and feedback procedure defined in IEEE 802.11 standards, where the receiver estimates the channel frequency response and reports a compressed version back to the BS. While this framework has proven effective for earlier Wi-Fi generations with small antenna configurations and moderate bandwidths, it encounters fundamental limitations as Wi-Fi evolves toward large-scale MIMO and ultra-wideband operation, as introduced in IEEE 802.11ax and beyond. These limitations are not merely implementation inefficiencies; rather, they reflect intrinsic scalability challenges in how channels are estimated, compressed, and fed back in conventional systems. In this section, we discuss these limitations in detail, focusing on feedback overhead growth, computational burden, and the increasing mismatch between traditional channel estimation pipelines and next-generation Wi-Fi requirements.

### 4.1.1 Explosive Growth of Feedback Overhead with Antenna Count and Bandwidth

In conventional Wi-Fi explicit sounding, the receiver estimates the channel across all OFDM subcarriers and antenna pairs. The resulting channel matrix is then compressed using SVD and Givens rotation-based parameterization, producing a set of quantized angles that describe the spatial characteristics of the channel. These angles are fed back periodically to the transmitter in beamforming report frames. A key limitation of this approach is that the feedback size grows rapidly as system dimensions increase. Specifically, the amount of feedback scales quadratically with the number of antennas and linearly with the transmission bandwidth. This means that when either antenna pair added or bandwidth is increased, the feedback overhead grows at a much faster rate than the achievable throughput gains promised by MIMO theory.

This scaling behavior becomes particularly problematic in modern Wi-Fi systems. For example, moving from a 2×2 MIMO configuration at 40 MHz to an 8×8 configuration at 160 MHz does not merely multiply the feedback size by a constant factor; instead, the combined effect of increased antennas and bandwidth leads to feedback payloads that are orders of magnitude larger. As demonstrated in recent experimental studies such as SHRINK, the feedback traffic in large MIMO Wi-Fi systems can reach tens of megabits per second when sounding is performed at standard intervals. In such scenarios, a non-negligible fraction of the available airtime is consumed solely by control traffic, reducing the net throughput available for data transmission. More importantly, this overhead grows faster than the spectral efficiency gains provided by adding antennas. As a result, the theoretical benefits of large MIMO arrays are increasingly offset by the cost of acquiring and maintaining accurate channel knowledge, leading to diminishing returns in practical deployments.

### 4.1.2 High Computational Burden at the User-End

Beyond feedback size, conventional channel estimation places a significant computational burden on the UE, i.e. the receiver, which is typically a user device such as a smartphone, laptop, or embedded wireless node. In the standard Wi-Fi feedback pipeline, the receiver must estimate the full CFR, perform matrix decompositions, compute multiple rotation angles, quantize them, and package the resulting information into feedback frames—all within tight timing constraints. These operations become increasingly expensive as antenna dimensions grow. Performing SVD or equivalent matrix factorizations across many subcarriers requires substantial processing power and memory bandwidth. When sounding is performed frequently, as required in dynamic or mobile environments, this computation must be repeated at millisecond-level intervals, further increasing the processing load. From a system design perspective, this poses a serious challenge. User devices are constrained by battery capacity, thermal limits, and real-time execution requirements. The computational cost of conventional channel feedback directly translates into higher energy consumption and longer processing delays at the receiver[64]. Experimental observations show that even with aggressive feedback compression, receiver-side processing remains a bottleneck, particularly for high-dimensional MIMO configurations.

This situation is fundamentally misaligned with the design goals of modern wireless networks, where complexity is expected to reside primarily at infrastructure nodes such as BS, not at user devices.

### 4.1.3 Heavyweight Encoder–Decoder Assumptions in Learning-Based Compression

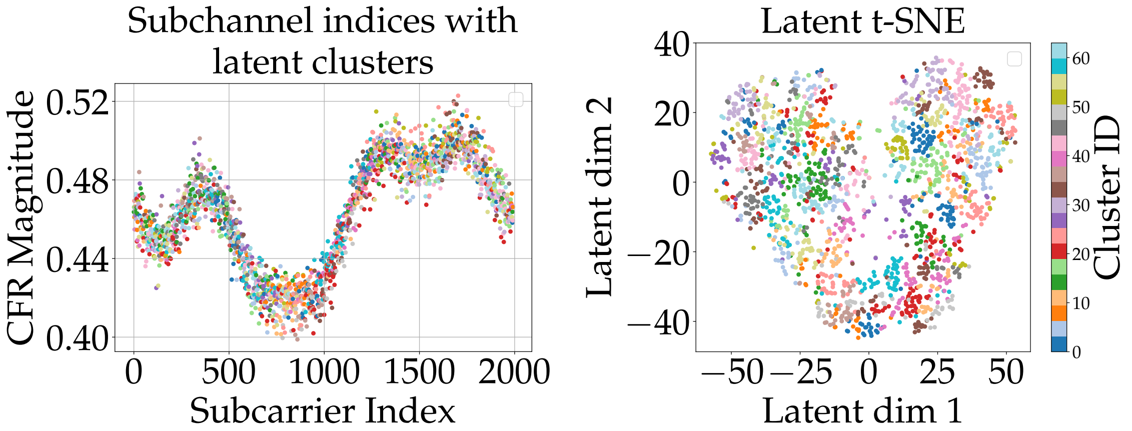
In response to the growing feedback overhead, recent research has explored learning-based channel compression techniques, often relying on autoencoder architectures to reduce the size of the reported channel information. While these approaches demonstrate promising compression performance in simulation, they introduce a new set of limitations when examined from a system-level perspective. Most learning-based schemes assume that the receiver runs a neural network encoder that processes the entire channel matrix and produces a compact latent representation. This latent vector is then fed back to the transmitter, where a corresponding decoder reconstructs the channel. Although this reduces the number of transmitted bits, it significantly increases the computational demands on the receiver.

In practice, such neural encoders require real-time inference over all subcarriers and antenna pairs, often involving multiple convolutional or fully connected layers. This makes them difficult to deploy on commodity Wi-Fi devices, which lack the specialized hardware acceleration available at infrastructure nodes. Additionally, the tight coupling between encoder and decoder architectures complicates system upgrades and model maintenance in real-world deployments. These heavyweight encoder–decoder designs effectively trade communication overhead for computational overhead at the receiver, rather than eliminating inefficiency altogether. Consequently, while learning-based compression reduces feedback size, it does so at the cost of increased complexity, energy consumption, and reduced deployment feasibility.

### 4.1.4 Lack of Adaptivity and Redundancy Awareness

Another fundamental limitation of conventional channel estimation is its lack of adaptivity to redundancy in the channel. Traditional feedback mechanisms treat all subcarriers and antenna pairs as equally important, reporting channel information uniformly across frequency and space. This design choice ignores the fact that real wireless channels exhibit strong correlation across neighboring subcarriers and spatial dimensions.

Practical experiments in prior literature show that many subcarriers carry highly redundant information and contribute little to changes in communication performance. Nevertheless, conventional feedback pipelines transmit updates for all subcarriers at fixed intervals, regardless of whether the channel has meaningfully changed. This results in unnecessary feedback traffic, i.e. communication overhead and inefficient use of control bandwidth. The inability to selectively prioritize informative channel components further exacerbates feedback overhead and limits scalability, especially in wide-band systems. We argue that overhead can be reduced by feeding back only the most informative subchannels. Our preliminary analysis as presented in Fig. 4.1




(a) Subchannel indices mapped to embedding clusters

(b) Embedding clusters obtained through K-means clustering

Figure 4.1: Analysis of the latent embeddings of the CFR.

shows that when projected into a latent embedding space, subchannels form clusters that reveal strong redundancy. This suggests that feeding back representatives from these clusters can preserve essential channel characteristics while avoiding redundant feedback.

## 4.2 ELF Design Philosophy

Taken together, these observations reveal a fundamental mismatch between conventional channel estimation pipelines and the requirements of modern Wi-Fi MIMO systems. The quadratic growth of feedback overhead, the heavy computational burden on user devices, the impracticality of encoder-heavy learning-based solutions, and the lack of redundancy awareness collectively limit the scalability and efficiency of existing designs.

These limitations motivate a rethinking of channel estimation and feedback as a system-level problem, rather than a purely signal processing task. Specifically, there is a clear need for feedback mechanisms that reduce overhead without shifting excessive complexity to the receiver, and exploit redundancy in the channel. At the core of ELF is the principle that wireless channels exhibit substantial structure and redundancy across frequency, space, and time. Also that effective feedback mechanisms should exploit this structure rather than transmit full channel realizations. Instead of attempting to faithfully reconstruct the entire channel at the UE side or relying on complex encoders to compress it, ELF adopts a sensing-inspired view in which the channel is treated as a latent object that can be selectively observed. From this perspective, the objective of feedback is not to transmit all channel coefficients, but rather to convey a compact representation that preserves the information most relevant for downstream communication tasks such as precoding and detection.

This viewpoint naturally shifts the computational burden away from the UE device toward the BS, which is typically less constrained in terms of power, memory, and processing capability. By keeping UE-side operations lightweight and largely linear while allowing more expressive decoding and reconstruction at the access, ELF aligns with the asymmetric capabilities inherent in Wi-Fi networks. Moreover, this separation enables ELF to scale gracefully with increasing antenna counts and bandwidths, as the complexity of the feedback mechanism depends primarily on the number of selected channel representatives rather than the full channel dimensionality. In the next section, we describe how ELF addresses these challenges by selectively embedding and feeding back only the most informative components of the channel while shifting complexity to the access.

## 4.3 ELF System Model Design

This section presents the end-to-end system model of ELF , a low-overhead channel feedback framework that reduces feedback cost by (i) projecting the measured channel frequency response (CFR) into a compact latent space, (ii) selecting a small set of representative subchannels online at the user, (iii) quantizing only the selected latent representations for feedback, and (iv) reconstructing a full-band CFR at the BS using a learned decoder. The proposed design is aligned with commodity Wi-Fi MIMO systems, where the access point (AP) (or BS) requires accurate frequency-selective channel information for precoding, while the user device is constrained in terms of computation, power, and latency.
We first define the signal representation and notation used throughout this section. We then provide a high-level walkthrough of the ELF pipeline, followed by a detailed technical description of each processing stage implemented at the UE and the BS.

### 4.3.1 Signal Representation and Notation

We consider an OFDM-based MIMO system with $M$ transmit antennas at the BS and $N$ receive antennas at a station. Let $F$ denote the number of active OFDM subcarriers used for channel reporting after excluding guard and null tones. For each sounding instance, the station estimates a frequency-domain MIMO channel for every subcarrier $f \in \{1, \ldots, F\}$. The complex CFR at subcarrier $f$ is denoted by

$$\mathbf{H}_f \in \mathbb{C}^{N \times M}. \tag{4.1}$$

In the implementation, the CFR is represented in a real-valued tensor form by separating real and imaginary components and arranging them into feature vectors. Let

$$\mathbf{X} \in \mathbb{R}^{C \times T \times F} \tag{4.2}$$

denote the channel representation at the station, where $C$ denotes the number of per-time-step channel features, $T$ is a short temporal dimension associated with repeated observations or averaging, and $F$ is the number of subcarriers. For batch-based processing, the input tensor becomes

$$\mathbf{X}^{(b)} \in \mathbb{R}^{B \times C \times T \times F}, \tag{4.3}$$

where $B$ denotes the batch size.

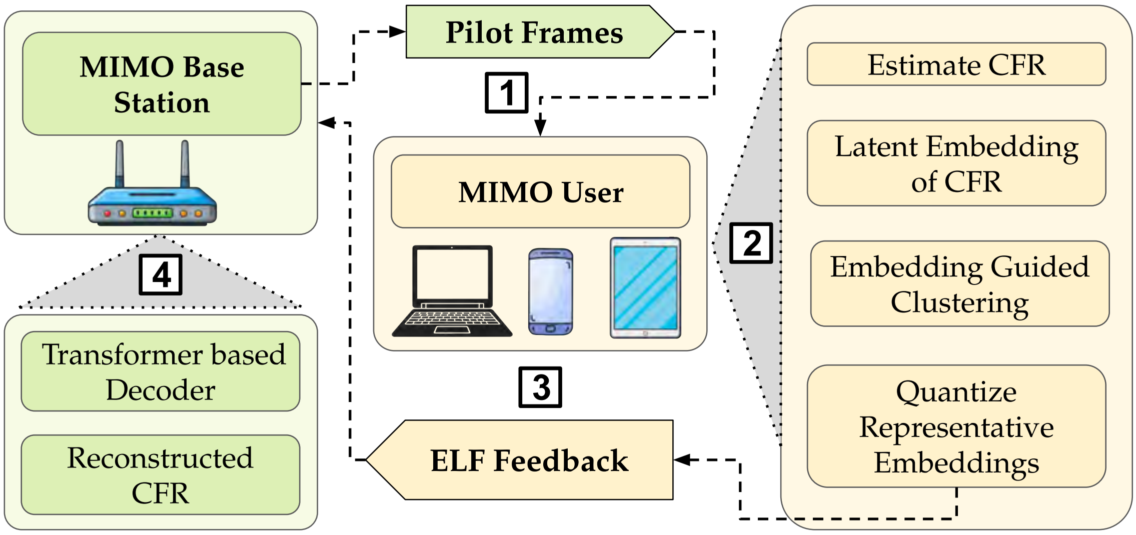



Figure 4.2: A walkthrough of ELF

A design choice in the current ELF implementation is to learn reconstruction over a time-averaged channel representation. The training target is defined as

$$\mathbf{Y} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{X}_{:,t,:} \in \mathbb{R}^{C \times F}, \tag{4.4}$$

which is later reshaped to $\mathbf{Y} \in \mathbb{R}^{F \times C}$ for convenience.

### 4.3.2 Walkthrough of ELF

The ELF feedback pipeline consists of four sequential stages distributed across the station and the BS as depicted in Fig. 4.2. First, the station estimates the CFR and forms the tensor $\mathbf{X}$. Instead of compressing and feeding back the full channel across all subcarriers, the station normalizes and projects each subcarrier into a low-dimensional latent space. An online selection procedure then identifies a small subset of representative subcarriers whose embeddings capture the essential structure of the channel.

Only the embeddings of these representatives, along with their indices and associated quantization parameters, are fed back to the BS. At the BS, the received embeddings are dequantized and passed through a learned decoder that reconstructs the channel representation across all subcarriers. The reconstructed channel is reshaped into the complex CFR tensor format required by the physical-layer processing chain, and communication performance is evaluated using a standard MIMO-OFDM pipeline with BER as the final metric. This design ensures that the feedback cost scales with the number of representatives rather than the full channel dimensionality, while shifting learning and reconstruction complexity to the BS.

### 4.3.3 UE-Side Processing: Normalization and Latent Projection

The UE -side processing begins by reshaping the input tensor $\mathbf{X}^{(b)} \in \mathbb{R}^{B\times C\times T\times F}$ into a per-subcarrier feature matrix. Let

$$\mathbf{X}_f \in \mathbb{R}^{F\times p}, \quad p = C \cdot T, \tag{4.5}$$

denote the reshaped matrix where each row corresponds to one subcarrier. In batch form, this becomes $\mathbf{X}_f^{(b)} \in \mathbb{R}^{B\times F\times p}$.

To reduce scale variability and stabilize downstream processing, ELF applies layer normalization independently to each subcarrier feature vector:

$$\tilde{\mathbf{X}}_{b,f,:} = \gamma \odot \frac{\mathbf{X}_{b,f,:} - \mu_{b,f}}{\sqrt{\sigma_{b,f}^2 + \epsilon}} + \beta, \tag{4.6}$$

where $\mu_{b,f}$ and $\sigma_{b,f}^2$ are the mean and variance across the feature dimension, $\gamma$ and $\beta$ are learned affine parameters, and $\epsilon$ is a small constant.

Each normalized feature vector is then projected into a $d$-dimensional latent embedding using a learned linear map:

$$\mathbf{z}_{b,f} = \tilde{\mathbf{X}}_{b,f,:}\mathbf{W} \in \mathbb{R}^d, \tag{4.7}$$

where $\mathbf{W} \in \mathbb{R}^{p\times d}$ is trained offline at the BS. The resulting latent tensor is

$$\mathbf{Z} \in \mathbb{R}^{B\times F\times d}. \tag{4.8}$$

### 4.3.4 UE-Side Processing: Online Representative Selection

At the user equipment (UE), feeding back latent embeddings for all $F$ subcarriers would result in prohibitively large overhead, especially for wideband Wi-Fi systems with large antenna arrays. To address this, ELF performs an online selection of a small subset of representative subcarriers whose latent embeddings capture the dominant structure of the frequency-selective channel.

Let $\tilde{\mathbf{X}}_{b,:,:} \in \mathbb{R}^{F\times p}$ denote the layer-normalized per-subcarrier feature matrix for the $b$-th sample, where $p = C \cdot T$. The goal of representative selection is to identify a set of $K \ll F$ subcarriers whose associated features are maximally informative while avoiding redundant feedback from highly correlated subcarriers.

**Low-Dimensional Projection for Clustering** Direct clustering in the original $p$-dimensional feature space is computationally expensive and impractical for execution at the UE. To enable lightweight selection, ELF first projects each subcarrier feature vector into a lower-dimensional clustering space. Specifically, a random orthogonal projection matrix

$$\mathbf{R} \in \mathbb{R}^{p\times d_c}, \quad d_c \ll p, \tag{4.9}$$

is generated offline and shared implicitly across devices via a fixed random seed. Each subcarrier feature is projected as

$$\mathbf{u}_{b,f} = \tilde{\mathbf{X}}_{b,f,:}\mathbf{R} \in \mathbb{R}^{d_c}, \tag{4.10}$$

forming a reduced representation $\mathbf{U}_b \in \mathbb{R}^{F\times d_c}$.

This projection preserves relative distances between subcarriers with high probability, while significantly reducing computational cost.

**Online Clustering Across Subcarriers** The projected features $\mathbf{U}_b$ are partitioned into $K$ clusters using a fixed-iteration $k$-means procedure executed independently for each sample:

$$\{\mathcal{C}_{b,1}, \ldots, \mathcal{C}_{b,K}\} = \mathrm{KMeans}(\mathbf{U}_b, K). \tag{4.11}$$

The number of clustering iterations is intentionally kept small to ensure predictable execution time at the UE.

**Medoid-Based Representative Selection** For each cluster $\mathcal{C}_{b,k}$, ELF selects a single representative subcarrier corresponding to the cluster medoid. Formally, the representative index is chosen as

$$\pi_{b,k} = \arg\min_{f \in \mathcal{C}_{b,k}} \sum_{f' \in \mathcal{C}_{b,k}} \|\mathbf{u}_{b,f} - \mathbf{u}_{b,f'}\|_2^2, \tag{4.12}$$

ensuring that the selected subcarrier corresponds to an actual measured channel realization rather than an abstract centroid.
Collecting the representatives across clusters yields the index vector

$$\boldsymbol{\pi}_b = \{\pi_{b,1}, \ldots, \pi_{b,K}\}, \quad \boldsymbol{\pi}_b \in \{1, \ldots, F\}^K. \tag{4.13}$$

The indices are sorted in ascending order prior to feedback to ensure deterministic ordering at the access point.

**Representative Latent Extraction** Using the selected indices, the corresponding latent embeddings are gathered from the latent tensor $\mathbf{Z} \in \mathbb{R}^{B \times F \times d}$ as

$$\mathbf{Z}^{\mathrm{rep}}_{b,i,:} = \mathbf{Z}_{b,\pi_{b,i},:}, \quad i = 1, \ldots, K, \tag{4.14}$$

forming the representative latent matrix

$$\mathbf{Z}^{\mathrm{rep}} \in \mathbb{R}^{B \times K \times d}. \tag{4.15}$$

This selection mechanism ensures that only a compact, non-redundant subset of the channel's latent structure is fed back, while remaining fully executable at the UE using simple linear operations and low-iteration clustering.

### 4.3.5 UE-Side Processing: Quantization and Feedback Formation

After selecting the representative latent embeddings, ELF applies a lightweight quantization procedure to further reduce feedback overhead while preserving reconstruction fidelity. This step is executed entirely at the UE and operates independently for each sample in the batch.
Let $\mathbf{Z}^{\mathrm{rep}}_b \in \mathbb{R}^{K \times d}$ denote the matrix of representative latent embeddings for the $b$-th sample, where $K$ is the number of selected subcarriers and $d$ is the latent dimension. Quantization is performed independently for each latent dimension to account for differences in dynamic range.

For each latent dimension $j \in \{1, \dots, d\}$, the minimum and maximum values across the selected representatives are computed as

$$z_{b,j}^{\min} = \min_{i \in \{1,\dots,K\}} Z_{b,i,j}^{\text{rep}}, \quad z_{b,j}^{\max} = \max_{i \in \{1,\dots,K\}} Z_{b,i,j}^{\text{rep}}. \tag{4.16}$$

To ensure numerical stability, the effective quantization range is clamped as

$$r_{b,j} = \max\left(z_{b,j}^{\max} - z_{b,j}^{\min}, \epsilon\right), \tag{4.17}$$

where $\epsilon$ is a small constant.
Each latent value is then normalized to the unit interval and uniformly quantized using $b_q$ bits:

$$q_{b,i,j} = \text{round}\left(\frac{Z_{b,i,j}^{\text{rep}} - z_{b,j}^{\min}}{r_{b,j}} \cdot (2^{b_q} - 1)\right), \tag{4.18}$$

where

$$q_{b,i,j} \in \{0, 1, \dots, 2^{b_q} - 1\}. \tag{4.19}$$

The quantized tensor

$$\mathbf{Q}_b \in \mathbb{Z}^{K \times d} \tag{4.20}$$

is stored using unsigned integer precision (e.g., `uint8` for $b_q \leq 8$), while the per-dimension minimum and maximum values,

$$\mathbf{z}_b^{\min}, \ \mathbf{z}_b^{\max} \in \mathbb{R}^d, \tag{4.21}$$

are stored using half-precision floating point to further reduce overhead.
In addition to the quantized embeddings, the station feeds back the indices of the selected subcarriers. Let

$$\boldsymbol{\pi}_b = \{\pi_{b,1}, \dots, \pi_{b,K}\} \tag{4.22}$$

denote the selected subcarrier indices. These indices are sorted in ascending order prior to transmission to ensure deterministic ordering at the access point and enable consistent latent placement during reconstruction.
The complete feedback packet for the $b$-th sample thus consists of the tuple

$$\mathcal{F}_b = \left(\mathbf{Q}_b, \ \mathbf{z}_b^{\min}, \ \mathbf{z}_b^{\max}, \ \boldsymbol{\pi}_b\right), \tag{4.23}$$

which is serialized and transmitted to the access point. Importantly, no additional metadata, side information, or channel statistics are included, ensuring that the feedback size scales linearly with $K$ and is independent of the total number of subcarriers $F$.
This quantization strategy preserves the relative geometry of the latent representations while introducing only bounded reconstruction error, and enables ELF to operate under strict feedback bandwidth constraints without imposing heavy computational requirements at the station.

### 4.3.6 Base Station Processing: Dequantization and Channel Reconstruction

Upon receiving the feedback packet $\mathcal{F}_b = (\mathbf{Q}_b, \mathbf{z}_b^{\min}, \mathbf{z}_b^{\max}, \boldsymbol{\pi}_b)$, the base station (BS) reconstructs a full-band channel representation from the compact, quantized feedback. Unlike the user equipment, the BS is not subject to strict computational or power constraints, allowing more expressive decoding operations to be performed centrally.

**Latent Dequantization** The first reconstruction step reverses the quantization applied at the UE. For each latent dimension $j \in \{1, \dots, d\}$, the dequantized latent representation is recovered as

$$\hat{z}_{b,i,j} = z_{b,j}^{\min} + \frac{q_{b,i,j}}{2^{b_q} - 1} \left( z_{b,j}^{\max} - z_{b,j}^{\min} \right), \tag{4.24}$$

where $q_{b,i,j}$ denotes the received quantized value. This operation yields a dequantized latent matrix

$$\hat{\mathbf{Z}}_b^{\text{rep}} \in \mathbb{R}^{K \times d}, \tag{4.25}$$

corresponding to the $K$ representative subcarriers selected at the UE.

**Frequency-Domain Interpolation via Learned Mixing** The objective of the decoder is to reconstruct latent embeddings for all $F$ subcarriers using only the $K$ representative embeddings. To this end, ELF employs a frequency-aware interpolation mechanism implemented as a lightweight multilayer perceptron (MLP).
Each subcarrier $f$ is assigned a normalized frequency position

$$\phi_f \in [-1, 1], \tag{4.26}$$

which serves as input to the decoder network. The network outputs a set of mixing coefficients

$$\mathbf{w}_f = \text{softmax}\big(\mathcal{D}_{\text{mix}}(\phi_f)\big) \in \mathbb{R}^K, \tag{4.27}$$

where $\mathcal{D}_{\text{mix}}(\cdot)$ is a two-layer MLP and the softmax ensures convex combination weights.
The reconstructed latent embedding at subcarrier $f$ is then computed as

$$\hat{\mathbf{z}}_{b,f} = \sum_{k=1}^{K} w_{f,k} \, \hat{\mathbf{Z}}_{b,k,:}^{\text{rep}}, \tag{4.28}$$

resulting in the full latent tensor

$$\hat{\mathbf{Z}} \in \mathbb{R}^{B \times F \times d}. \tag{4.29}$$

**Latent-to-Channel Mapping** Each reconstructed latent embedding is mapped back to the channel feature space using a learned linear projection:

$$\hat{\mathbf{y}}_{b,f} = \mathbf{A}\hat{\mathbf{z}}_{b,f} + \mathbf{c}, \tag{4.30}$$

where $\mathbf{A} \in \mathbb{R}^{C \times d}$ and $\mathbf{c} \in \mathbb{R}^C$ are decoder parameters learned during training.
Stacking across frequency yields

$$\hat{\mathbf{Y}} \in \mathbb{R}^{B \times F \times C}. \tag{4.31}$$

**Complex CFR Reconstruction** Finally, the reconstructed channel features are reshaped and separated into real and imaginary components to obtain the complex-valued channel frequency response. Since ELF reconstructs a time-averaged channel representation, the recovered CFR is replicated uniformly across the required temporal dimension to match the expected physical-layer input format:

$$\hat{\mathbf{H}} \in \mathbb{C}^{B \times \cdots \times T \times F}. \tag{4.32}$$

### 4.3.7 BER-Based Evaluation and End-to-End Validation

To evaluate whether the reconstructed channel produced by ELF  is suitable for practical communication, we embed the reconstructed CFR into a complete MIMO-OFDM physical-layer pipeline and measure end-to-end BER. This evaluation directly reflects the impact of feedback compression on communication reliability, rather than relying solely on representation-level distortion metrics.
The reconstructed channel $\hat{\mathbf{H}}$ is injected into a standards-compliant OFDM transmission chain, including precoding, pilot-based channel estimation, equalization, demodulation, and channel decoding. In our implementation, all physical-layer components follow 3GPP and IEEE-compliant models and are realized using the SIONNA simulation framework.
At the transmitter, encoded data symbols are mapped onto OFDM resource grids and precoded using the reconstructed CFR. The signal propagates through the reconstructed frequency-selective MIMO channel, with additive noise determined by the target signal-to-noise ratio. At the receiver, the channel is estimated using pilot symbols and equalized using linear minimum mean squared error (LMMSE) equalization, followed by demodulation and decoding.
The resulting bit error rate is computed as

$$\mathrm{BER} = \frac{1}{N_{\mathrm{bits}}} \sum_{i=1}^{N_{\mathrm{bits}}} \mathbb{I}\big(\hat{b}_i \neq b_i\big), \tag{4.33}$$

where $b_i$ and $\hat{b}_i$ denote the transmitted and decoded bits, respectively.
By evaluating BER using the reconstructed CFR, ELF  is assessed under the same conditions as conventional channel estimation and feedback schemes. This provides a system-level validation that the proposed sensing-aided feedback mechanism preserves communication performance despite aggressive reduction in feedback overhead.

## 4.4 Simulation Environment

To evaluate the proposed ELF  framework under realistic and reproducible conditions, we construct a fully controlled simulation environment based on standardized channel models and a standards-compliant physical-layer abstraction. All simulated channel realizations are generated using the SIONNA library. It is an open source library which provides differentiable, GPU-accelerated implementations of 3GPP-compliant wireless channel and OFDM system models. This choice enables accurate modeling of wideband MIMO channels while maintaining consistency with modern Wi-Fi and cellular system assumptions.

### 4.4.1 Simulation Software and Framework

All simulations are implemented using the SIONNA physical-layer simulation framework, which is built on top of *TensorFlow* and provides modular implementations of channel models, antenna arrays, and OFDM signal processing blocks. SIONNA is particularly well suited for this work because it supports *(i)* standardized stochastic channel models, *(ii)* scalable multi-antenna configurations, and *(iii)* tight integration with learning-based pipelines. The simulation scripts are executed on GPU-enabled systems, with *TensorFlow* configured to allow dynamic memory growth to ensure stable execution across different hardware setups. Random seeds are explicitly controlled at both the *TensorFlow* and SIONNA levels to guarantee reproducibility across simulation runs. Each generated channel realization is associated with a unique seed that is stored as part of the dataset metadata, enabling exact regeneration of individual samples when required.

### 4.4.2 Channel Model and Propagation Environment

We model the wireless propagation channel using the 3GPP TR 38.901 CDL model, which captures realistic multipath propagation effects observed in indoor and urban environments. Specifically, we employ CDL models A considering indoor environment, as the default configuration unless otherwise specified. For each channel realization, the *continuous-time channel impulse response* is generated using the CDL model and subsequently transformed into a frequency-domain representation using an OFDM-compatible transformation. The CFR is normalized to ensure consistent average channel power across realizations. To improve robustness and avoid overfitting to a narrow channel distribution, we introduce light stochastic augmentations in the frequency domain, including a random global phase rotation and a small linear frequency tilt. These augmentations preserve the physical structure of the channel while increasing variability across samples.

### 4.4.3 OFDM and Bandwidth Configuration

The simulated system follows an OFDM-based wideband transmission model consistent with modern Wi-Fi standards. The total system bandwidth is configurable and supports 40 MHz, 80 MHz, 160 MHz, and 320 MHz modes. The corresponding Fast Fourier Transform (FFT) sizes are selected accordingly to maintain realistic subcarrier spacing:

$$FFT \in \{512, 1024, 2048, 4096\}.$$

Unless otherwise specified, simulations are conducted at 80 MHz bandwidth, which represents a common operating point for IEEE 802.11ax systems.

Each OFDM frame consists of ten OFDM symbols, including pilot symbols placed at predefined indices to support channel estimation at the receiver. *Guard subcarriers* and *DC nulling* are applied to match practical OFDM resource grid configurations. The cyclic prefix length is fixed across simulations to ensure consistent handling of delay spread effects.

### 4.4.4 Antenna Configuration and MIMO Setup

To evaluate the scalability of ELF with respect to antenna dimension, we simulate multiple MIMO configurations by varying the number of transmit and receive antennas. The number of base station BS and user equipment UE antennas is configurable and includes square MIMO setups such as $2\times2$, $3\times3$, $4\times4$ and $5\times5$. Dual-polarized cross antennas are used at both the BS and UE, following the antenna patterns defined in TR 38.901. The antenna arrays are constructed using standardized array geometry models, with the number of rows and columns adjusted according to the total antenna count. This setup enables realistic modeling of spatial correlation, polarization effects, and antenna coupling, all of which directly influence the structure and redundancy of the resulting channel frequency response.

**Indoor Room Layout and Geometry** We evaluate ELF under an indoor room-layout setting in which the transmitter (**TX**) is fixed at a single location and six receivers (**RX**) are sampled across multiple spatial positions within each room, as illustrated in Fig. 4.3. We consider two representative indoor geometries: (i) a *Bedroom* of approximately 3.75 m × 3.5 m and (ii) a *Reading Room* with approximate dimensions 3.2 m × 3.45 m (with a non-rectangular boundary). For each geometry, the RX is placed at several distinct locations to emulate spatial diversity, resulting in TX–RX separations spanning roughly 1.5–2.0 m and 2.0–2.5 m, depending on the selected RX point. This configuration enables us to test reconstruction robustness across realistic indoor spatial variations while keeping the transmitter location fixed. A schematic of the indoor room layout used for channel generation is shown in Fig. 4.3.

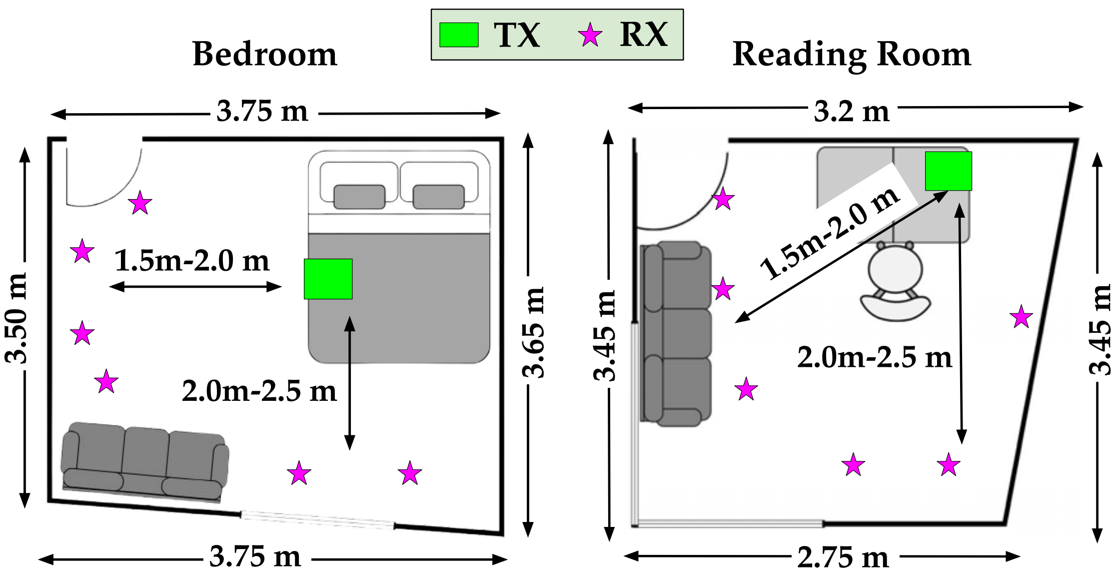



Figure 4.3: Simulated indoor room layout and node placement.

### 4.4.5 Dataset Generation and Storage

Channel realizations are generated in batches and stored in a sharded dataset format to enable efficient streaming during training and evaluation. Each shard contains a small number of samples (often a single sample per shard), allowing fine-grained access without loading the entire dataset into memory. For each sample, the complex-valued channel frequency response CFR is stored by packing the real and imaginary components into a compact floating-point representation. In addition to the channel data, extensive metadata is stored alongside each sample. This metadata includes the CDL model type, delay spread, carrier frequency, antenna configuration, bandwidth, FFT size, user mobility parameters, and random seed. This design ensures that every simulated channel realization is fully self-describing and traceable, supporting reproducibility and detailed post hoc analysis.

### 4.4.6 Relevance to ELF Evaluation

The simulation environment is intentionally designed to stress the scalability limitations of conventional channel feedback mechanisms. Wide bandwidths, large antenna arrays, and frequency-selective channels collectively produce high-dimensional CFRs that are costly to estimate, compress, and feed back using traditional methods. By generating such channels under controlled yet realistic conditions, the simulation environment provides a rigorous testbed for evaluating the effectiveness of ELF in reducing feedback overhead while preserving end-to-end communication performance. In the subsequent sections, this simulation environment is used consistently across training, CFR reconstruction, and BER evaluation, ensuring that all performance gains attributed to ELF arise from its system design rather than artifacts of the simulation setup.

# Chapter 5

# Performance Analysis

This chapter presents an extensive performance evaluation of the proposed PULSE and ELF frameworks. We analyze the sensitivity to key design parameters, compare it against state-of-the-art Wi-Fi sensing approaches, study the impact of spatial diversity across sensing stations, evaluate robustness under domain shifts, and examine computational efficiency and model compactness. All experiments related to PULSE are conducted using the *BeamSense* dataset under consistent settings to ensure fair comparison. All experiments related to ELF are conducted in the customized simulated environment shown in Fig. 4.3.

## 5.1 Performance Analysis of PULSE

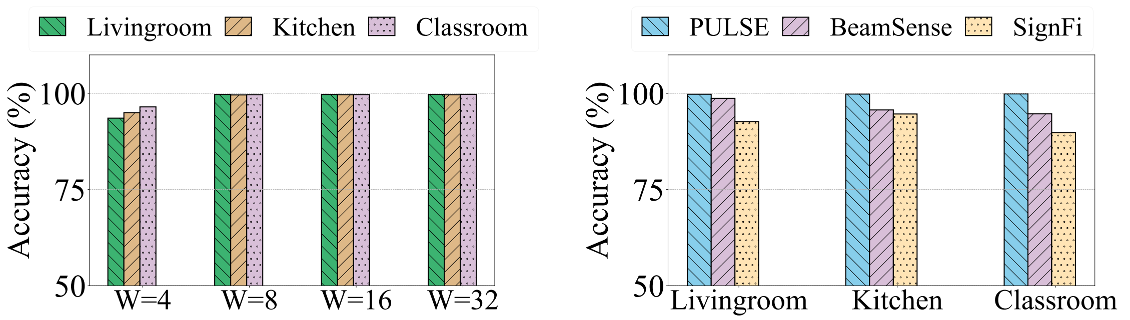
### 5.1.1 Performance of PULSE with Different Temporal Windows

In communication-aided sensing framework PULSE , the temporal window length $W$ determines the number of consecutive CFR frames aggregated into a single sensing instance. A short window may fail to capture sufficient temporal dynamics, while an excessively long window increases latency and computational overhead without proportional gains. To evaluate this trade-off, we vary the temporal window length as $W \in \{4, 8, 16, 32\}$ and report classification accuracy across three indoor environments: living room, kitchen, and classroom.

The results are illustrated in Figure 5.1a

Across all environments, increasing the window size from $W = 4$ to $W = 8$ leads to a noticeable improvement in accuracy—by approximately 6.16% in the living room, 4.64% in the kitchen, and 3.16% in the classroom. This improvement indicates that a moderate temporal context is necessary to capture motion-induced channel variations reliably.

However, further increasing $W$ beyond 8 yields diminishing returns. For example, accuracy saturates around 99.72%–99.76% for larger windows while incurring higher inference latency. Based on this analysis, $W = 8$ is selected as the default configuration, as it achieves near-optimal performance with minimal delay, making it suitable for real-time edge sensing applications.




(a) Impact of temporal window length on PULSE accuracy

(b) Performance comparison of PULSE with baseline models

Figure 5.1: PULSE Accuracy analysis under varying temporal window $W$ and baseline comparison using the selected configuration $W = 8$

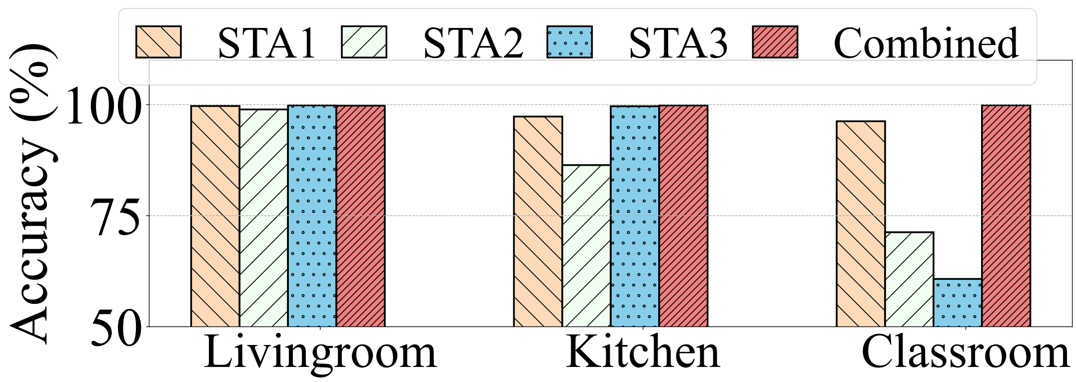



Figure 5.2: PULSE accuracy comparison of individual against combined Wi-Fi stations across three indoor environments

### 5.1.2 Comparison with State-of-the-Art Approaches

We next compare PULSE against two representative state-of-the-art Wi-Fi sensing systems: *BeamSense* and *SignFi*. All methods are evaluated under identical experimental conditions across the three indoor environments. As shown in Fig. 5.1b, PULSE consistently outperforms both baselines. Specifically, PULSE achieves classification accuracies of 99.78%, 99.82%, and 99.86% in the living room, kitchen, and classroom environments, respectively. In comparison, BeamSense attains 98.72%, 95.69%, and 94.67%, while SignFi achieves 92.64%, 94.64%, and 89.76%.

The accuracy gains of PULSE —up to 5.2% over BeamSense and nearly 10% over SignFi—stem from its ability to focus on physics-aware temporal dynamics rather than static CFR representations. By embedding structured temporal descriptors, PULSE captures motion patterns more effectively while avoiding unnecessary high-dimensional processing.

### 5.1.3 Impact of Spatial Diversity Across Wi-Fi Stations

Wi-Fi sensing performance is highly sensitive to the relative position and orientation of sensing devices. Individual stations may experience poor channel visibility depending on placement and line-of-sight conditions.

Fig. 5.2 evaluates the sensing accuracy of individual stations (STA1, STA2, and STA3) and their fusion across different environments. Results show that while single-station performance can degrade severely—e.g., STA2 and STA3 achieve only 86.39% and 60.76% accuracy in the kitchen and classroom, respectively—the fused multi-station configuration consistently achieves near-perfect accuracy exceeding 99.8%.
These results demonstrate that PULSE effectively integrates spatially diverse channel observations, compensating for unfavorable station orientations. The framework remains robust even when individual sensing nodes provide weak or noisy information, highlighting its suitability for practical multi-device deployments.

### 5.1.4 Generalization to Unseen Environments and Stations

To evaluate domain generalization, we test PULSE under cross-environment and cross-station scenarios. In these experiments, the model is trained in one environment and evaluated in unseen environments using few-shot adaptation.

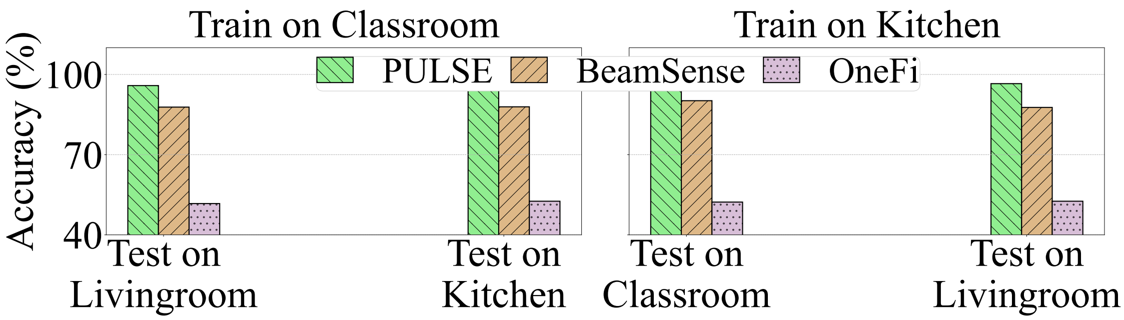



Figure 5.3: Comparison of PULSE performance across environments with baseline BeamSense, and SignFi.

As illustrated in Fig. 5.3, when trained in the classroom and tested in the living room and kitchen, PULSE achieves average accuracies of 95.84% and 95.05%, respectively. This represents an improvement of more than 7.1% over BeamSense and over 40% relative to OneFi. Training on kitchen dataset and testing on the data from other unseen environments, yield similar accuracy, improving the performance of PULSE by up to 9% over its nearest competitor BeamSense.
Cross-station generalization results, depicted in Fig. 5.4, follow the same pattern. These results validate the effectiveness of PULSE 's temporal embedding design and few-shot adaptation mechanism in real-world deployment scenarios.

PULSE used contrastive learning with few-shots leveraging only 5s worth of data from unseen environment, the baseline framework BeamSense takes 15s worth of data per activity to fine-tune to new environment. Thus, PULSE not only generalizes effectively across unseen domains but also achieves breakthrough sensing performance under extreme data constraints.

### 5.1.5 Computational Efficiency and Model Compactness

Table 5.1 compares the computational requirements of PULSE with BeamSense and SignFi. While BeamSense requires nearly 20 billion FLOPs and SignFi approx-

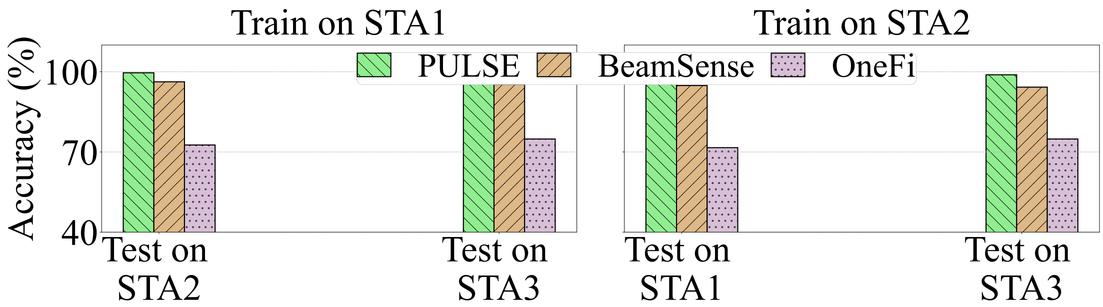



Figure 5.4: Comparison of PULSE performance across Wi-Fi stations with baseline BeamSense, and SignFi.

imately 1.7 billion FLOPs, PULSE performs the same sensing task with only 22.41 million FLOPs.

Table 5.1: Analysis of computational efficiency of PULSE

| Approach | FLOPs(M) | Params(M) | Inference (ms/samp) |
|---|---|---|---|
| **PULSE** | 22.41 | 1.47 | 0.012 |
| BeamSense | 20000 | 4.25 | 22.0 |
| SignFi | 1700 | 2.10 | 0.62 |
| BLTHAT | 0.127 | 1.39 | 0.19 |

This represents a reduction of over 900× compared to BeamSense and over 75× compared to SignFi. Despite its lightweight nature, PULSE maintains competitive parameter count (1.47 million) and achieves an exceptionally low inference latency of 0.012 ms per sample. To ensure a fair comparison against compact neural architectures, we additionally include *BLTHAT*, a recently proposed lightweight CSI-based sensing model for edge deployment. Despite being designed as a lightweight CSI-based model, *BLTHAT* remains approximately 5.7× more computationally expensive than PULSE . It further maintains a competitive parameter count (1.47 M) and achieve an exceptionally low inference latency of 0.012 ms/sample, compared to the SOTA baselines- 0.19 ms for *BLTHAT*, 22 ms for *BeamSense* and 0.62 ms for *SignFi*. This efficiency enables real-time operation on edge devices with limited computational resources.

Table 5.2 further analyzes scalability with respect to temporal window length $W$ and sub-channel grouping parameter $k$. As $W$ increases from 8 to 64, FLOPs scale linearly from 22.41M to 179.26M, while inference latency remains below 0.025 ms. In contrast, varying $k$ from 4 to 32 introduces only marginal increases in computational cost, confirming that temporal modeling dominates complexity. With the configuration $W = 8$ and $k = 8$, PULSE reduces the input dimensionality from $W \times F \times 2$ (raw CFR) to $W \times C = 8 \times 72$, resulting in an 85.12% reduction in feature dimensionality. This compact representation preserves sensing-relevant information while significantly lowering computational and memory overhead.

Table 5.2: Computational Efficiency of PULSE  Under Varying Window and sub-channel Configurations

| Window $W$ ($K=8$) | FLOPs(M) | Params(M) | Inference (ms/samp) |
|---|---|---|---|
| 8 | 22.41 | 1.47 | 0.012 |
| 16 | 44.82 | 1.47 | 0.012 |
| 32 | 89.63 | 1.47 | 0.014 |
| 64 | 179.26 | 1.47 | 0.023 |
| **Sub-channel Group $k$ ($W=8$)** | **FLOPs(M)** | **Params(M)** | **Inference (ms/samp)** |
| 4 | 21.89 | 1.43 | 0.011 |
| 8 | 22.41 | 1.47 | 0.012 |
| 16 | 23.44 | 1.53 | 0.012 |
| 32 | 25.50 | 1.66 | 0.012 |

### 5.1.6 Result Discussion

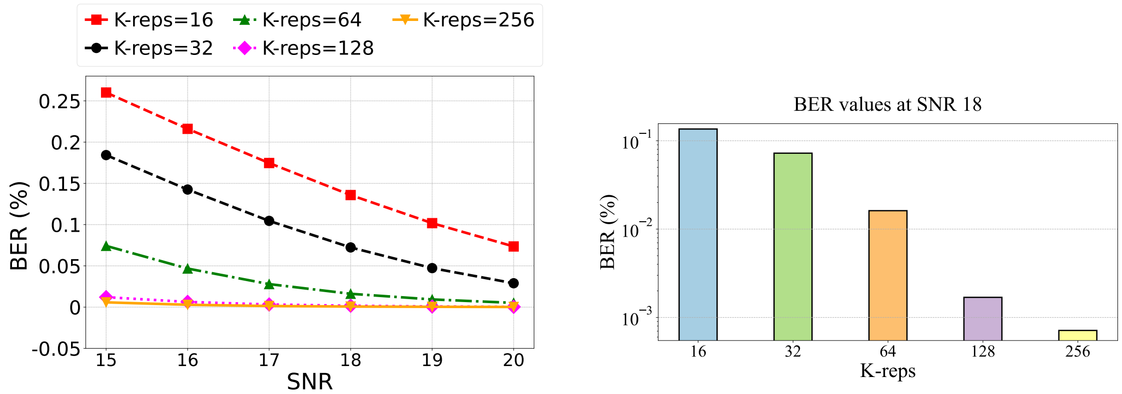
The performance statistics imply that, PULSE is a lightweight domain-adaptive sensing framework that learns the physics-aware temporal embeddings, extracted from Wi-Fi CFR. Unlike the existing approaches where the CFR is directly used as the input the tensor for Learning model, PULSE extracts temporal descriptors and embeds them through a light-weight 1D convolutional network that jointly learns discriminative representations and sensing semantics. The framework achieves over 99% accuracy while reducing 85% of the input tensor dimensionality and maintaining low inference latency, demonstrating suitability for real-time edge inference. Furthermore, through supervised contrastive pretraining and few-shot adaptation, PULSE generalizes effectively to unseen domains using only 5 s worth of labeled data, outperforming the state-of-the-art frameworks.

## 5.2 Performance Analysis of ELF

### 5.2.1 Impact of the Number of Representative Subchannels

We begin the performance evaluation of ELF  by analyzing the effect of the number of representative subchannels, denoted by $K$, on the end-to-end BER. Since BER depends jointly on channel quality and system noise, we first identify an optimum SNR operating point that allows meaningful comparison across different values of $K$ without degrading the BER performance.

Figure 5.5a shows the BER performance of ELF  across a range of SNR values from 15 dB to 20 dB for multiple values of $K$. At low SNR values, the system performance is dominated by noise, and the BER performance remains sub-per. At higher SNR values, while BER continues to decrease, the overall simulation cost increases due




(a) BER performance across a range of SNR (b) BER performance across different K-reps

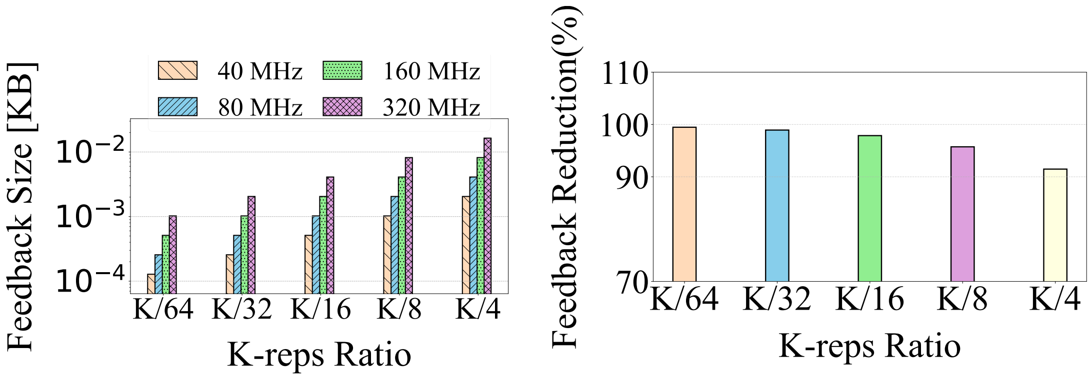
Figure 5.5: Optimum SNR value and corresponding BER performance at different number of representative subchannels at 80MHz BW

to longer decoding times and larger numbers of successfully decoded frames. Based on this trade-off, we select an SNR of 18 dB as the optimum operating point for subsequent analysis. This SNR corresponds to a regime where the BER is sufficient to reflect meaningful reconstruction quality. Moreover, 18 dB represents a practical mid-to-high SNR range for IEEE 802.11ax deployments, making it a suitable choice for system-level evaluation.

With the SNR fixed at 18 dB, we next examine how BER varies as a function of the number of representative subchannels $K$. Figure 5.5b illustrates the BER achieved by ELF at 80 MHz bandwidth for different values of $K$. As expected, increasing $K$ consistently improves BER performance, since more representative subchannels provide a richer description of the underlying frequency-selective channel. However, the improvement is not linear. When $K$ increases from 16 to 64, the BER drops by more than an order of magnitude, indicating that a small number of representatives is insufficient to capture the frequency-domain structure of the channel. Increasing $K$ further to 128 yields a substantial additional reduction in BER, pushing the system into a regime where communication reliability approaches that of much denser feedback. In contrast, increasing $K$ from 128 to 256 provides only marginal BER improvement, despite doubling the feedback size and associated processing cost to be analyzed next. This behavior reveals a clear performance–overhead trade-off. For the 80 MHz IEEE 802.11ax configuration considered here, $K = 128$ emerges as an effective operating point that balances feedback efficiency and communication reliability. Notably, this corresponds to approximately **one-eighth of the total number of active subcarriers**, demonstrating that ELF can reconstruct a full-band channel with high fidelity while feeding back only a small fraction of the original channel information.

### 5.2.2 Feedback Overhead Analysis with K-reps Ratio

We next analyze how the feedback size of ELF scales with the ratio of representative subchannels $K$ under different system bandwidths. Figure 5.6a illustrates the total feedback size as a function of the $K$-to-subcarrier ratio for bandwidths ranging from 40 MHz to 320 MHz.




(a) Channel feedback size across different K-reps ratio

(b) Percentage of channel feedback reduction

Figure 5.6: Channel feedback size reduction analysis across different K-reps ratio for different channel BW.

Across all bandwidths, the feedback size grows approximately linearly with the number of representatives. For example, at 80 MHz bandwidth, the feedback size increases from $1.28 \times 10^{-4}$ KB when using $K/64$ representatives to $2.05 \times 10^{-3}$ KB when using $K/4$, corresponding to a $16\times$ increase. Similar linear scaling trends are observed for 40 MHz, 160 MHz, and 320 MHz bandwidths, indicating that the dominant contributor to feedback overhead is the number of selected subchannels rather than the absolute system bandwidth. For a fixed representative ratio, increasing the bandwidth results in a proportional increase in feedback size due to the larger total number of subcarriers. At a representative ratio of $K/8$, the feedback size grows from approximately $1.02 \times 10^{-3}$ KB at 40 MHz to $1.64 \times 10^{-2}$ KB at 320 MHz, reflecting the fourfold increase in bandwidth. Nevertheless, even at the largest bandwidth considered, the absolute feedback size remains on the order of only a few tens of bytes.

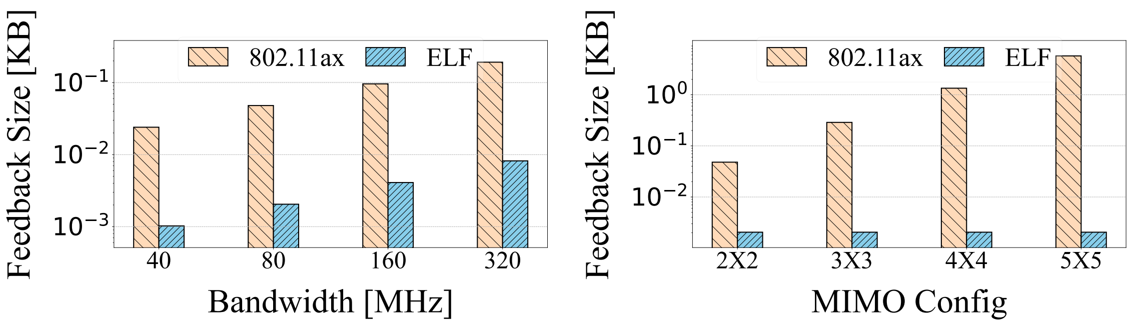
When combined with the BER results presented earlier, these findings highlight a favorable operating regime for ELF . In particular, the operating point of $K/8$ (e.g., $K$-reps = 128 at 80 MHz with total 1024 subcarriers) achieves a strong reduction in feedback size while preserving sufficient channel information for reliable communication. As shown in Fig. 5.6b, this configuration yields approximately 96% feedback reduction compared to dense subcarrier reporting. More aggressive compression ratios ($K/16$–$K/64$) increase the reduction to 98%–99%, but lead to noticeable BER degradation, while larger ratios such as $K/4$ reduce the feedback savings to about 92% with only marginal BER improvement.

This confirms that $K/8$ provides a well-balanced trade-off between feedback reduction efficiency and communication reliability, enabling ELF to scale effectively across wide bandwidths without excessive feedback overhead.

### 5.2.3 Comparison with Traditional IEEE 802.11ax Channel Feedback

We next compare the feedback overhead of ELF against the standard IEEE 802.11ax explicit channel feedback across different channel bandwidths and MIMO configurations. Figure 5.7 summarizes the feedback size for both schemes, while the corresponding numerical values are reported in the accompanying data files.

For a fixed $2 \times 2$ MIMO configuration, the feedback size of IEEE 802.11ax increases linearly with bandwidth, from 0.024 KB at 40 MHz to 0.192 KB at 320 MHz. In contrast, the feedback size of ELF remains significantly lower, increasing only from 0.001024 KB to 0.008192 KB over the same bandwidth range. This corresponds to a reduction factor of upto 25×, demonstrating that ELF effectively decouples feedback overhead from wideband channel dimensionality.




(a) Feedback comparison with varying BW

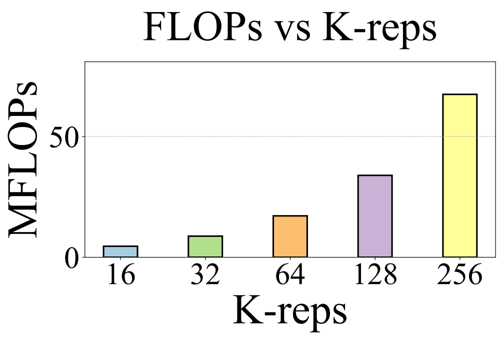
(b) Feedback comparison with different MIMO config.

Figure 5.7: Feedback size comparison between ELF and traditional IEEE 802.11ax Wi-Fi standard with varying BW and MIMO configurations.

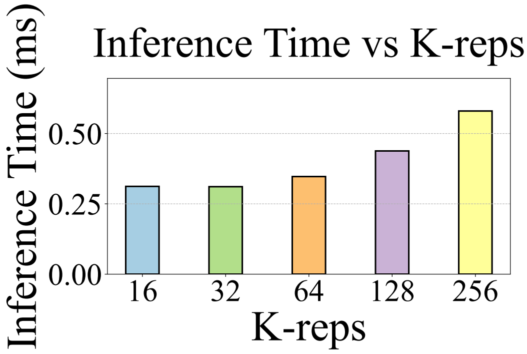
A similar trend is observed when scaling the number of antennas at a fixed bandwidth of 80 MHz. With IEEE 802.11ax feedback, the overhead grows rapidly with the MIMO dimension, increasing from 0.048 KB for $2 \times 2$ to 5.76 KB for $5 \times 5$ MIMO[64]. In contrast, the feedback size of ELF remains constant at approximately 0.002048 KB across all evaluated MIMO configurations. This behavior highlights a key advantage of ELF : feedback overhead is determined primarily by the number of selected representatives rather than the number of antenna pairs. Overall, these results confirm that conventional IEEE 802.11ax feedback mechanisms scale poorly with both bandwidth and antenna dimension, whereas ELF achieves consistently low and predictable feedback overhead. By exploiting frequency-domain redundancy and selective representation, ELF enables scalable MIMO operation without the exponential increase in feedback cost that characterizes standard feedback procedures.

### 5.2.4 Computational Complexity and Inference Latency

We next evaluate the computational cost of ELF in terms of floating-point operations (FLOPs) and inference latency as a function of the number (or ratio) of representative subchannels $K-$reps, considering an 80 MHz bandwidth and a $2 \times 2$ MIMO configuration. Figure 5.8 summarizes the corresponding FLOPs and average inference time per sample.
As $K$-reps increases, the computational complexity grows approximately linearly. Specifically, the FLOPs increase from 4.57 MFLOPs at $K$-reps = 16 to 33.93 MFLOPs at $K$-reps = 128, and further to 67.49 MFLOPs at $K$-reps = 256. A similar trend is observed in inference latency: the average per-sample latency remains around




(a) Computational efficiency of ELF in terms of FLOPS




(b) Computational efficiency of ELF in terms of inference latency

Figure 5.8: Computational efficiency and compactness analysis of ELF .

0.31–0.35 ms for $K$-reps $\leq 64$ (i.e. one sixteenth of total subcarriers) increases to 0.44 ms at $K$-reps $= 128$ (i.e. one eighth of total subcarriers), and exceeds 0.58 ms at $K$-reps $= 256$ (i.e. one fourth of total subcarriers). This growth in latency is accompanied by a corresponding drop in throughput, from over 3200 samples/s at $K = 16$ to approximately 1723 samples/s at $K = 256$.

These results show that using more representative subchannels beyond a certain point provides only small performance improvements, while significantly increasing the computational cost. In particular, $K$-reps $= 128$ in 80 MHz BW, emerges as a favorable operating point: it keeps the inference time below 0.5 ms, limits the computational load to approximately 34 MFLOPs, and maintains high throughput, while already achieving strong BER performance and substantial feedback reduction, as shown in previous analysis. Consequently, selecting one-eighth of the total subchannels as representatives enables ELF to remain lightweight and suitable for practical deployment without sacrificing communication reliability.

### 5.2.5 Discussion

The results show that ELF achieves a fundamentally different operating point by exploiting frequency-domain redundancy and shifting reconstruction complexity to the BS. Selecting approximately one-eighth of the active subcarriers consistently preserves near-dense feedback performance, revealing substantial redundancy in wideband Wi-Fi channels that conventional feedback mechanisms overlook. Unlike IEEE 802.11ax explicit feedback, whose overhead scales poorly with bandwidth and antenna count, the feedback cost of ELF depends primarily on the number of representative subchannels, enabling up to 96% overhead reduction without degrading communication reliability. UE-side processing remains lightweight, while decoding and interpolation are handled centrally at the BS, with measured FLOPs and latency confirming practical deployability. Overall, these results validate a sensing-oriented view of channel feedback, positioning ELF as a scalable and standards-compatible solution for next-generation wideband Wi-Fi systems.

# Chapter 6

# Conclusion

## 6.1 Review of Contributions

This thesis investigated how the wireless channel, traditionally treated solely as a medium for data transmission, can be systematically repurposed as a rich sensing resource alongside communication. Motivated by the observation that conventional Wi-Fi systems treat the wireless channel primarily as an instantaneous communication object, this work demonstrated that the channel also embodies rich physical structure that can be harnessed to enable scalable, low-overhead sensing and feedback mechanisms.

The core contributions of the thesis are realized through two complementary frameworks: PULSE and ELF . Together, they establish a unified perspective in which communication signals assist sensing, and sensing principles in turn improve communication efficiency.

### 6.1.1 Communication-Aided Sensing via PULSE

The first part of this thesis introduced PULSE , a communication-aided wireless sensing framework that exploits the temporal evolution of Wi-Fi CFR for human activity recognition. Rather than operating directly on high-dimensional channel tensors, PULSE adopts a physics-aware, temporal-first design that captures magnitude dynamics, phase evolution, and energy variation through compact temporal descriptors. This representation **reduces the input dimensionality by over 85%** while retaining sensing-relevant information. The extracted descriptors are processed using a lightweight 1D convolutional network that jointly learns discriminative embeddings and activity semantics. To address domain variability, PULSE further integrates supervised contrastive learning and prototype-based few-shot adaptation, enabling rapid adaptation to unseen environments, devices, and users. With only a few seconds of labeled data PULSE successfully recognize activity classes with over **95% performance accuracy**, surpassing other SOTA approaches by upto **9%**.

Extensive evaluation on the real-world *BeamSense* dataset demonstrates that PULSE consistently achieves over **99% activity recognition accuracy across multiple environments and user stations**, while maintaining low inference latency suitable for real-time edge deployment. Compared to existing Wi-Fi sensing approaches, PULSE achieves orders-of-magnitude lower computational complexity **(only 22.41MFLOPs)** without sacrificing performance. These results validate that explicitly modeling temporal channel structure is sufficient to enable accurate, robust, and deployable wireless sensing under practical Wi-Fi constraints.

### 6.1.2 Sensing-Aided Communication via ELF

The second part of this thesis presented ELF , a sensing-aided channel feedback framework that addresses the fundamental scalability limitations of conventional Wi-Fi channel estimation and feedback. As Wi-Fi systems evolve toward wider bandwidths and larger MIMO configurations, standard feedback mechanisms incur rapidly growing overhead and impose significant computational burden on user devices, limiting practical deployment.

ELF reformulates channel feedback as a selective sensing problem over a structured frequency-domain channel rather than full-channel reporting. By projecting per-subcarrier channel features into a latent space and exploiting frequency-domain redundancy, ELF enables the user to identify and feed back only a small set of representative subchannels. These representatives are quantized using lightweight operations and transmitted to the base station, where a learned decoder reconstructs the full-band channel. Extensive simulations based on 3GPP-compliant channel models show that selecting **approximately one-eighth of the active subchannels** achieves more than **96%** reduction in feedback overhead with negligible BER degradation across wide bandwidths. Compared to IEEE 802.11ax explicit feedback, **ELF achieves up to** $25\times$ **reduction in feedback size** and, critically, maintains nearly constant overhead as the number of antenna pairs increases. Computational analysis further confirms that ELF remains lightweight, with **inference latency well below 0.5 ms with around 34MFLOPs** at practical operating points. These results demonstrate that sensing-inspired feedback—based on redundancy awareness and selective representation—can substantially reduce communication overhead while preserving reliability, positioning ELF as a scalable and deployable solution for next-generation wideband Wi-Fi systems.

## 6.2 Future Research Directions

The work presented in this thesis opens several promising avenues for future research.

- Extending both PULSE and ELF to cell-free MIMO network settings would be a natural next step. In such scenarios, spatial diversity introduces additional structure and redundancy that could further improve sensing robustness and feedback efficiency in both terrestrial networks (TN) and non-terrestrial networks (NTN).

- A key direction for future work is the testbed realization of the proposed frameworks. While PULSE has been validated on real-world datasets, an end-to-end testbed implementation would enable evaluation under real-time and hardware constraints. Similarly, extending ELF from emulation to a hardware-based Wi-Fi testbed would allow direct assessment of feedback integration, latency, and deployment feasibility.

- Online and adaptive optimization of $k$-reps could be implemented in ELF to allow more efficient channel estimation frameworks to leverage communication. Moreover, continual learning and self-supervised adaptation in PULSE may further reduce the reliance on labeled data and manual calibration.

- Finally, deploying the proposed frameworks on real hardware platforms and evaluating them in large-scale testbeds would provide valuable insights into system-level trade-offs, latency constraints, and energy efficiency. Such deployments would help bridge the gap between simulation-based evaluation and practical adoption in next-generation wireless systems.

## 6.3 Closing Remarks

In conclusion, this thesis demonstrated that communication channels can be treated as a primitive for sensing tasks. By exploiting the inherent structure of wireless channels, it is possible to design ISAC systems that are both efficient and robust. The frameworks proposed in this work—PULSE and ELF —represent concrete steps toward this vision and highlight the potential of physics-aware, representation-driven approaches for the future of wireless networks.

# Bibliography


[1] S. R. Theodore, *Wireless communications: principles and practice.* CAMBRIDGE University Press, 2024.

[2] A. F. Molisch, *Wireless communications: from fundamentals to beyond 5G.* John Wiley & Sons, 2022.

[3] K. Pahlavan and P. Krishnamurthy, "Evolution and impact of wi-fi technology and applications: A historical perspective," *International Journal of Wireless Information Networks*, vol. 28, no. 1, pp. 3–19, 2021.

[4] W. Xi et al., "Electronic frog eye: Counting crowd using wifi," in *IEEE INFOCOM 2014-IEEE Conference on Computer Communications*, IEEE, 2014, pp. 361–369.

[5] F. Liu et al., "Integrated sensing and communications: Toward dual-functional wireless networks for 6g and beyond," *IEEE journal on selected areas in communications*, vol. 40, no. 6, pp. 1728–1767, 2022.

[6] S. Kumar, *Wireless Communication-the fundamental and advanced concepts.* River Publishers, 2022.

[7] R. N. Dave and D. K. Patel, "Efficient human activity recognition using deep unfolding in 802.11 bf wi-fi sensing," in *2025 17th International Conference on COMmunication Systems and NETworks (COMSNETS)*, IEEE, 2025, pp. 557–565.

[8] B. Kellogg, V. Talla, and S. Gollakota, "Bringing gesture recognition to all devices," in *Proceedings of the 11th USENIX Conference on Networked Systems Design and Implementation*, ser. NSDI'14, Seattle, WA: USENIX Association, 2014, pp. 303–316, ISBN: 9781931971096.

[9] N. C. Luong et al., "Advanced learning algorithms for integrated sensing and communication (isac) systems in 6g and beyond: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, 2025.

[10] D. C. Nguyen et al., "6g internet of things: A comprehensive survey," *IEEE Internet of Things Journal*, vol. 9, no. 1, pp. 359–383, 2021.

[11] M. Lombardi, F. Pascale, and D. Santaniello, "Internet of things: A general overview between architectures, protocols and applications," *Information*, vol. 12, no. 2, p. 87, 2021.

[12] R. Medeiros, S. Fernandes, and P. G. Queiroz, "Middleware for the internet of things: A systematic literature review," *Journal of Universal Computer Science*, vol. 28, no. 1, p. 54, 2022.

[13] J. Wilson and N. Patwari, "Radio tomographic imaging with wireless networks," *IEEE Transactions on Mobile Computing*, vol. 9, no. 5, pp. 621–632, 2010.

[14] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proceedings of the 19th annual international conference on Mobile computing & networking*, 2013, pp. 27–38.

[15] 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.

[16] E. G. Larsson, "Massive synchrony in distributed antenna systems," *IEEE Transactions on Signal Processing*, vol. 72, pp. 855–866, 2024.

[17] A. Hassanien, M. G. Amin, E. Aboutanios, and B. Himed, "Dual-function radar communication systems: A solution to the spectrum congestion problem," *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 115–126, 2019.

[18] J. A. Zhang et al., "Enabling joint communication and radar sensing in mobile networks—a survey," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 306–345, 2021.

[19] A. Zhang, M. L. Rahman, X. Huang, Y. J. Guo, S. Chen, and R. W. Heath, "Perceptive mobile networks: Cellular networks with radio vision via joint communication and radar sensing," *IEEE Vehicular Technology Magazine*, vol. 16, no. 2, pp. 20–30, 2020.

[20] Z. Zhang et al., "6g wireless networks: Vision, requirements, architecture, and key technologies," *IEEE vehicular technology magazine*, vol. 14, no. 3, pp. 28–41, 2019.

[21] A. Kaushik et al., "Toward integrated sensing and communications for 6g: Key enabling technologies, standardization, and challenges," *IEEE Communications Standards Magazine*, vol. 8, no. 2, pp. 52–59, 2024.

[22] K. Wu, J. Xiao, Y. Yi, D. Chen, X. Luo, and L. M. Ni, "Csi-based indoor localization," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 7, pp. 1300–1309, 2012.

[23] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11 n traces with channel state information," *ACM SIGCOMM computer communication review*, vol. 41, no. 1, pp. 53–53, 2011.

[24] Z. Wang, B. Guo, Z. Yu, and X. Zhou, "Wi-fi csi-based behavior recognition: From signals and actions to activities," *IEEE Communications Magazine*, vol. 56, no. 5, pp. 109–115, 2018.

[25] K. F. Haque, M. Zhang, F. Meneghello, and F. Restuccia, "Beamsense: Rethinking wireless sensing with mu-mimo wi-fi beamforming feedback," *Computer Networks*, vol. 258, p. 111 020, 2025.

[26] K. F. Haque, M. Zhang, F. Meneghello, and F. Restuccia, "Si-fi: Learning the beamforming feedback for simultaneous multi-subject sensing," *Computer Networks*, p. 111 907, 2025.

[27] K. F. Haque, M. Zhang, and F. Restuccia, "Simwisense: Simultaneous multi-subject activity classification through wi-fi signals," in *2023 IEEE 24th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, IEEE, 2023, pp. 46–55.

[28] H. Liu, L. Xi, W. Wang, F. Zhang, and Z. J. Haas, "Openfi: Open-set wifi human sensing via virtual embedding confidence-aware," *IEEE Transactions on Network Science and Engineering*, 2024.

[29] W.-F. Alliance, *The economic value of wi-fi: A global view (2018 and 2023)*, 2021.

[30] S. Tan, Y. Ren, J. Yang, and Y. Chen, "Commodity wifi sensing in ten years: Status, challenges, and opportunities," *IEEE Internet of Things Journal*, vol. 9, no. 18, pp. 17 832–17 843, 2022.

[31] M. Abbas, M. Elhamshary, H. Rizk, M. Torki, and M. Youssef, "Wideep: Wifi-based accurate and robust indoor localization system using deep learning," in *2019 IEEE International Conference on Pervasive Computing and Communications (PerCom*, IEEE, 2019, pp. 1–10.

[32] Y. Zeng, D. Wu, J. Xiong, J. Liu, Z. Liu, and D. Zhang, "Multisense: Enabling multi-person respiration sensing with commodity wifi," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, no. 3, pp. 1–29, 2020.

[33] C.-F. Hsieh, Y.-C. Chen, C.-Y. Hsieh, and M.-L. Ku, "Device-free indoor human activity recognition using wi-fi rssi: Machine learning approaches," in *2020 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-Taiwan)*, IEEE, 2020, pp. 1–2.

[34] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of wifi signal based human activity recognition," in *Proceedings of the 21st annual international conference on mobile computing and networking*, 2015, pp. 65–76.

[35] M. Zhang, Z. Fan, R. Shibasaki, and X. Song, "Domain adversarial graph convolutional network based on rssi and crowdsensing for indoor localization," *IEEE Internet of Things Journal*, vol. 10, no. 15, pp. 13 662–13 672, 2023.

[36] S. Depatla and Y. Mostofi, "Crowd counting through walls using wifi," in *2018 IEEE international conference on pervasive computing and communications (PerCom)*, IEEE, 2018, pp. 1–10.

[37] P. Ssekidde, O. Steven Eyobu, D. S. Han, and T. J. Oyana, "Augmented cwt features for deep learning-based indoor localization using wifi rssi data," *Applied Sciences*, vol. 11, no. 4, p. 1806, 2021.

[38] N. Singh, S. Choe, and R. Punmiya, "Machine learning based indoor localization using wi-fi rssi fingerprints: An overview," *IEEE access*, vol. 9, pp. 127 150–127 174, 2021.

[39] F. Meneghello, D. Garlisi, N. Dal Fabbro, I. Tinnirello, and M. Rossi, "Sharp: Environment and person independent activity recognition with commodity ieee 802.11 access points," *IEEE Transactions on Mobile Computing*, vol. 22, no. 10, pp. 6160–6175, 2022.

[40] R. Xiao, J. Liu, J. Han, and K. Ren, "Onefi: One-shot recognition for unseen gesture via cots wifi," in *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems*, 2021, pp. 206–219.

[41] C. Shi et al., "Environment-independent in-baggage object identification using wifi signals," in *2021 IEEE 18th international conference on mobile ad hoc and smart systems (MASS)*, IEEE, 2021, pp. 71–79.

[42] M. Muaaz, A. Chelli, M. W. Gerdes, and M. Pätzold, "Wi-sense: A passive human activity recognition system using wi-fi and convolutional neural network and its integration in health information systems," *Annals of Telecommunications*, vol. 77, no. 3, pp. 163–175, 2022.

[43] Y. Ge et al., "Contactless wifi sensing and monitoring for future healthcare-emerging trends, challenges, and opportunities," *IEEE Reviews in Biomedical Engineering*, vol. 16, pp. 171–191, 2022.

[44] L. Guo et al., "Wiar: A public dataset for wifi-based activity recognition," *IEEE Access*, vol. 7, pp. 154 935–154 945, 2019.

[45] Y. Ma, G. Zhou, S. Wang, H. Zhao, and W. Jung, "Signfi: Sign language recognition using wifi," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 1, pp. 1–21, 2018.

[46] N. Bahadori, J. Ashdown, and F. Restuccia, "Rewis: Reliable wi-fi sensing through few-shot multi-antenna multi-receiver csi learning," in *2022 IEEE 23rd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, IEEE, 2022, pp. 50–59.

[47] B. Bloessl, M. Segata, C. Sommer, and F. Dressler, "An ieee 802.11 a/g/p ofdm receiver for gnu radio," in *Proceedings of the second workshop on Software radio implementation forum*, 2013, pp. 9–16.

[48] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity wifi," in *Proceedings of the 21st Annual international conference on Mobile Computing and Networking*, 2015, pp. 53–64.

[49] F. Gringoli, M. Schulz, J. Link, and M. Hollick, "Free your csi: A channel state information extraction platform for modern wi-fi chipsets," in *Proceedings of the 13th International Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization*, 2019, pp. 21–28.

[50] F. Gringoli, M. Cominelli, A. Blanco, and J. Widmer, "Ax-csi: Enabling csi extraction on commercial 802.11 ax wi-fi platforms," in *Proceedings of the 15th ACM Workshop on Wireless Network Testbeds, Experimental Evaluation & CHaracterization*, 2022, pp. 46–53.

[51] Y. Zheng et al., "Zero-effort cross-domain gesture recognition with wi-fi," in *Proceedings of the 17th annual international conference on mobile systems, applications, and services*, 2019, pp. 313–325.

[52] Y. Fang, B. Sheng, H. Wang, and F. Xiao, "Witransfer: A cross-scene transfer activity recognition system using wifi," in *Proceedings of the ACM turing celebration conference-China*, 2020, pp. 59–63.

[53] C. Wu, X. Huang, J. Huang, and G. Xing, "Enabling ubiquitous wifi sensing with beamforming reports," in *Proceedings of the ACM SIGCOMM 2023 Conference*, 2023, pp. 20–32.

[54] T. Kanda, T. Sato, H. Awano, S. Kondo, and K. Yamamoto, "Respiratory rate estimation based on wifi frame capture," in *2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC)*, 2022, pp. 881–884. DOI: 10.1109/CCNC49033.2022.9700721.

[55] E. Perahia and R. Stacey, *Next generation wireless LANs: 802.11 n and 802.11 ac*. Cambridge university press, 2013.

[56] X. Ma, Q. Gao, J. Wang, V. Marojevic, and J. H. Reed, "Dynamic sounding for multi-user mimo in wireless lans," *IEEE Transactions on Consumer Electronics*, vol. 63, no. 2, pp. 135–144, 2017.

[57] E. Reshef and C. Cordeiro, "Future directions for wi-fi 8 and beyond," *IEEE Communications Magazine*, vol. 60, no. 10, pp. 50–55, 2022.

[58] C.-K. Wen, W.-T. Shih, and S. Jin, "Deep learning for massive mimo csi feedback," *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 748–751, 2018.

[59] P. K. Sangdeh, H. Pirayesh, A. Mobiny, and H. Zeng, "Lb-scifi: Online learning-based channel feedback for mu-mimo in wireless lans," in *2020 IEEE 28th International Conference on Network Protocols (ICNP)*, IEEE, 2020, pp. 1–11.

[60] N. Bahadori, Y. Matsubara, M. Levorato, and F. Restuccia, "Splitbeam: Effective and efficient beamforming in wi-fi networks through split computing," in *2023 IEEE 43rd International Conference on Distributed Computing Systems (ICDCS)*, IEEE, 2023, pp. 864–874.

[61] Z. Zuo, Y. Zhong, Y. Li, J. Zhang, and X. Ge, "Aiml-enhanced adaptive quantization for channel sounding feedback in wireless communications," in *2024 IEEE Wireless Communications and Networking Conference (WCNC)*, IEEE, 2024, pp. 1–6.

[62] O. Bejarano, E. Magistretti, O. Gurewitz, and E. W. Knightly, "Mute: Sounding inhibition for mu-mimo wlans," in *2014 Eleventh Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, IEEE, 2014, pp. 135–143.

[63] S. Su, W.-T. Tan, X. Zhu, R. Liston, H. Wildfeuer, and B. Aazhang, "Motion-aware optimizations for downlink mu-mimo in 802.11 ax networks," *IEEE Transactions on Network and Service Management*, vol. 18, no. 4, pp. 4088–4102, 2021.

[64] K. Rumman, F. Meneghello, K. F. Haque, F. Gringoli, and F. Restuccia, "Shrink: Reducing mimo feedback overhead in wi-fi with dynamic data-driven channel sounding," in *Proceedings of the Twenty-sixth International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*, 2025, pp. 11–20.

[65] X. Li, Z. Tian, Z. Li, and K. Liu, "A ranging solution with multi-channel splicing in the single frequency band," in *2023 IEEE 11th Asia-Pacific Conference on Antennas and Propagation (APCAP)*, IEEE, 2023, pp. 1–2.

[66] Z. Tian, Y. Ren, Z. Li, and X. Cheng, "Super-resolution time-of-flight estimation for ranging via multi-band splicing," in *GLOBECOM 2023-2023 IEEE Global Communications Conference*, IEEE, 2023, pp. 4656–4661.

[67] S. Dimce, A. Zubow, and F. Dressler, "Mmsplicer: Toward experimental multi-band channel splicing at mmwave frequencies," in *IEEE INFOCOM 2024-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, IEEE, 2024, pp. 1–2.

[68] K. F. Haque, F. Meneghello, J. Ashdown, and F. Restuccia, "Bandweave: Enhanced channel estimation in mimo networks with multi-band fusion," in *IEEE INFOCOM 2026-IEEE Conference on Computer Communications*, IEEE, 2026, pp. 1–10.