Leveraging Mutual Information in Stochastic CSI Analysis for Wi-Fi Sensing

Elena Tonini*, Renato Lo Cigno*
*DII, University of Brescia and CNIT, Italy.

***Abstract*—The latest generations of wireless networks are focusing on the integration of Joint Communication and Sensing capabilities to improve environment awareness and thus enhance services and user experience. One of the elements commonly used for sensing tasks is the Channel State Information (CSI) extracted at the receiver. Most studies exploit Artificial Intelligence (AI) to perform CSI-based sensing tasks, from localization to activity recognition and more. These studies have shown, at least in small, controlled experiments, exceptional capabilities, but have hardly ever offered interpretative models, insight on the sensing capabilities of the methods or fundamental results and bounds. Fortunately, studies are starting to appear that explore an analytic characterization of the CSI. This work falls into this latter category, proposing a quantitative method to measure the relationship between different CSI's. The approach leverages a quantized representation of the CSI amplitude and a reduced-complexity matrix representation that allows to compute a form of Mutual Information (MI) between CSI's. The MI is then used as a compact indicator of the similarity of two sets of CSI, potentially collected in different environments. The proposed method aims at providing a complement or, possibly, an alternative to AI-based sensing, improving the understanding of the CSI features exploited by AI. Results, encouraging though preliminary, are obtained on real-life situations and experiments that overall cover more than six hours of data collection, spanning across several months and amounting to over $800\,000$ CSI's.**

***Index Terms*—CSI Sensing; Joint Communication and Sensing; Stochastic Analysis; Mutual Information**

I. Introduction

The promise of enabling environment sensing in parallel to traditional communication through Joint Communication and Sensing (JCAS) is one of the aspects of future wireless networks that is drawing the most attention. The introduction of $78.125\,\mathrm{kHz}$ subcarrier spacing in both Wi-Fi and 5G, and increasingly larger bandwidth – especially in Wi-Fi, with channels up to $160\,\mathrm{MHz}$ in 802.11ax [1] and $320\,\mathrm{MHz}$ in 802.11be [2] –, promises to enhance JCAS as well as Multiple-Input Multiple-Output (MIMO) and Multi User-MIMO (MU-MIMO) performance.

JCAS, with all the different flavors and keywords it is dubbed as, is now a wide and varied field of research, encompassing pure device localization and tracking, mono and bistatic in-band radar, and many different methodologies and approaches. A key role in JCAS sensing applications is played by the Channel State Information (CSI), which can be described at a high level as the "footprint" that the propagation environment imprints on the signal. With studies on CSI sensing dating back to the early 2010s [3], [4], [5], the research area has lately started to encompass methodologies to exploit the extracted information to gain better insight into how the propagation environment affects the CSI. Multiple goals can now be achieved through CSI-based sensing, ranging from localization tasks [6], [7], to motion detection and gesture recognition [8], [9], [10], and more. Recent surveys are available to get the state of the art [11], [12].

Most proposed approaches use both supervised or unsupervised Artificial Intelligence (AI)-based techniques that leverage a Neural Network (NN) to extract knowledge from the CSI, recognize the environment where it was collected, and perform the desired sensing task. However, little is known on how the NN interprets and manipulates the CSI: although NNs yield impressively accurate results, it is difficult to pinpoint the properties of the CSI that reveal the environment the processed samples originate from, thus it is difficult, if not impossible, to generalize. Furthermore, unfortunately, very few proposals have proven to be robust, ambient independent, and replicable, which is a fundamental requirement to let JCAS exit from laboratories and become a real service.

All the advances in CSI-based sensing, we think, somewhat disregard two big areas: *i)* privacy protection, albeit some recent works started addressing the topic (for instance [13], [14], [15], [16], [17]); and *ii)* fundamental analysis and understanding of CSI structure and properties, which is the area this work fits into. Moreover, in many contexts, notably those regarding the Internet of Things (IoT) and all networking and applications outside the operator-based mainstream development, JCAS implementations may have resource constraints that can clash with the requirements of many deep learning approaches. To this day, however, even CSI sensing for IoT is mostly performed through centralized, cloud-based approaches [18]. Finally, if we consider real-time applications that require sensing tasks to rely on local computing at the edge, rather than remote data processing, careful resource management and algorithm simplicity are fundamental.

The main contribution of this study is to leverage the informative content of the CSI to implement a probability-based approach to ambient characterization that contributes to the understanding of how different CSI compare to one another and to what amount detailed and fine-grained CSI characterization is needed for some comparisons and tasks. Furthermore, we show how the technique can be applied to an easy scenario classification task exploiting a very large set of experimental data collected in realistic life scenarios: to this goal, we collect data for hours throughout several months of operation, without crafting extremely specialized experiments

ISBN: 978-3-903176-79-9

with tailored placement of transmitter and receiver and short-term (often just seconds) collections of CSI bursts for each scenario to be recognized, a methodology that hides the natural variability of the propagation environment.

II. Related Work

As mentioned, CSI-based sensing exploiting AI methods has proven its extreme versatility and accurate results, with works claiming the possibility of detecting breath [19], recognizing activities [20], counting people [21], and more. Some works like [22], however, analyze in depth its capabilities and limitations, highlighting the lack of fundamental results.

However accurate, the outcomes of sensing using NNs lack an explanation of which CSI components embed the pieces of information that support sensing. To tackle this issue, analytical methodologies of CSI analysis can be leveraged to shed light on the underlying links the CSI has with the physical environment. These approaches can be grouped into two main categories: sensing methods based on the interpretation of physical models and techniques based on stochastic or algebraic analysis. We refer the reader interested in AI-based sensing and its potentials to recent surveys and tutorials as [23], [12], and focus instead on the works that seek insights into the CSI characteristics and properties to establish theoretical or heuristic data-driven models to steer research.

The first category of studies encompasses works that exploit the Doppler effect and electromagnetic models. The authors of [8] tackle activity recognition tasks through AI-based Doppler analysis, while [24] exploits Doppler effects to ensure independence of the CSI samples from the environment before feeding them to an AI model trained for activity recognition. Interestingly, these works first manipulate the CSI's to extract explainable features, and then feed them to an AI tool for classification. The focus of [25] is instead on indoor speed estimation of a moving target: by computing the difference between consecutive CSI samples, the authors highlight the variations in the channel propagation paths, which generate the Doppler effects used to estimate speed.

The second category comprises of works, including this one, that represent the CSI through algebraic or stochastic models. These approaches do not exploit physical properties of the propagating signal, but rather propose interpretative models, accepting that they may not have a straightforward mapping onto physics. The approach to motion detection followed in [26] is based on the stochastic analysis of the electromagnetic field that impacts the CSI, assuming that the human body can be modeled as a reflector. Addressing intrusion detection with a different approach, [27] exploits the temporal analysis of CSI variations, and focuses on the CSI sampling process to reduce noise and outliers. The stochastic analysis of the electromagnetic field is also used in [28] to distinguish human activities characterized by movements at different speeds.

A specific mention goes to [29], which is, to the best of our knowledge, the first work to introduce the notion of mutual information between CSI's and to use information theory principles such as the Kullback–Leibler and Jensen–Shannon divergences to compare distributions of CSI, an approach we also follow in this work. These concepts are then used to train AI algorithms on small datasets, obtaining very promising results; unfortunately, the method is not tested on large, more varied datasets. An additional approach that goes beyond CSI interpretation but remains deserving of mention shifts the focus to improving CSI extraction quality, in an attempt to make the collected data environment-independent [30].

Finally, we mention previous works introducing CSI analysis and interpretation. In [31] we introduced a variation of the Hamming distance to quantify the differences between distinct environments, showing it is a viable approach, and [32] extended the work to encompass also mmWave frequencies.

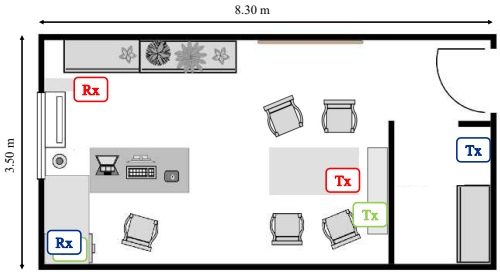


Figure 1. Office layout in the experimental testbed.

III. Experimental Setup and CSI Preprocessing

The experiments considered for this study were performed throughout several months. All the experiments have been carried out by the Advanced Networking Systems research group[1] of the University of Brescia.

A. Description of the Experimental Setup

The experimental space is a standard office used regularly (no 'ambient crafting'). The positions of the Tx and Rx are coupled and classified into three configurations, each associated to a different color in Fig. 1. In the red configuration, the transmitter is placed on a desk and the receiver is at the diagonally opposite corner of the room roughly at the same height. In the green configuration, the Tx is on a shelf about 2 m high, and the Rx is on a desk near the opposite wall. The blue configuration enables CSI analysis in Non-Line of Sight (NLoS) scenarios: it reuses the Rx position of the green setup, but the Tx is placed in a closet room behind a drywall wall.

To avoid any confusion, we will always use the term *scenario* to identify a situation that we wish to sense and isolate from any other. Different sessions of measurement within the same scenario are called *experiments*. Thus, different Tx and Rx placement are different scenarios, as well as an empty room is a different scenario from the same room with any number of people inside. Instead, two distinct one-minute collections of CSI samples captured in the empty room are two different experiments within the same scenario.

[1] https://ans.unibs.it

All the experiments are based on long-term measures with frames separated by several milliseconds (from 5 to 33) and are not short bursts of back-to-back frames. We avoided crafting experimental conditions that would only represent edge-cases in real-life sensing applications (e.g., a person standing still for several minutes in the middle of the Line of Sight (LoS)) and focused on real-world conditions. To this goal, different Tx/Rx positions were used to study the impact of device placement on the structure and behavior of CSI samples.

An overview of the conducted experiments is displayed in Tab. I. Each scenario is identified by the 'code' `XB.n` where `X` can be E (Empty office), S (people Sitting), or M (people Moving); `B` is the Bandwidth (20, 40 or 80 MHz); `n` is the number of people in the room. Multiple experiments within the same scenario are differentiated by appending the index of the experiment (e.g., /0,1,2) to the scenario code. The last column of the table specifies the positions of the Tx and Rx as red (R), green (G) or blue (B), referencing the colors used in Fig. 1 to identify the devices disposition. When two experimental configurations differ only for the Tx/Rx positions, the indexes of the experiments associated to each setup are indicated next to the letter identifying the devices placement. For example, 'B (1-3), G (4-9)' indicates that experiments indexed 1 to 3 were performed with the blue Tx/Rx disposition, while experiments 4 to 9 were instead carried out with the green one. In the experiments performed with the blue and green Tx/Rx dispositions, three receiving antennas were enabled, thus the number of available CSI samples triplicates in those cases, for a total of 118 80 MHz-bandwidth experiments.[2]

For all experiments, both Tx and Rx are equipped with the Wi-Fi Broadcom chipset BCM43684. The Tx is always an RT-AX82U Access Point (AP) from Asus. The receiver is also an RT-AX82U for experiments performed with the red Tx/Rx disposition and an RT-AX86U AP in all other cases. The CSI is extracted from the receiver Wi-Fi chipset using an extension of the Nexmon CSI Extractor tool, which was presented for the first time in [33] and later extended in collaboration with the SEEMOO research group at the University of Darmstadt.

B. CSI Preprocessing: Normalization and Quantization

The CSI of a Wi-Fi frame is a local estimation of the channel distortion measured at the receiver and it is represented as a complex number. Spanning over the N_{SC} subcarriers used for the Orthogonal Frequency Division Multiplexing (OFDM) modulation, the CSI associated to the k-th Wi-Fi frame is

$$\mathbf{C}(k,n) = A_{\mathbf{C}}(k,n) \cdot e^{j\angle \mathbf{C}(k,n)},$$

where $n \in [1, N_{\text{SC}}]$ is the normalized subcarrier index counting all and only the subcarriers effectively used in the modulation, i.e., subcarriers in the guard band and those suppressed in transmission are not considered.

The extraction tools normally represent the CSI in an arbitrary measure unit as IQ components of the received signal; the original extraction is based on a fixed-point representation of the Analog-to-Digital Converter (ADC) component of the Wi-Fi chipset that is unfortunately unknown to the extractor. The retrieved values are also affected by the Automatic Gain Control (AGC), whose effect cannot be controlled, and are then cast onto floating point numbers.

In this work, we only focus on the amplitude of the CSI's, thus each experiment is a *collection* $\mathbf{C}$ of $M_{\mathbf{C}}$ vectors $A_{\mathbf{C}}(k,\cdot)$ of size N_{SC}. The vector $A_{\mathbf{C}}(k,\cdot)$ is the amplitude of the k-th CSI, for $k \in [1, M_{\mathbf{C}}]$.

Considering the discrepancies in the extraction process due to different technologies, configurations, and devices, comparing CSI samples belonging to different experiments first requires a normalization of the collected data. A thorough analysis of this process can be found in [34], but the workflow can be summarized in three steps, which have to be applied to every experiment separately: i) Normalization of the CSI energy to compensate for the effect of the AGC; ii) Computation of the maximum and minimum $A_{\mathbf{C}}(\cdot,\cdot)$ values; iii) Mapping of the amplitudes of all CSI samples to the $[0,1]$ interval.

To improve comparability of samples coming from different experiments and simultaneously allow to quantify the information embedded in each CSI, [31] suggests a method to select the appropriate number of bits for the quantization that ensure that enough information is maintained in the quantized amplitude to also represent random fluctuations. We follow the same approach, but we will explore how the reduction of

Table I
SUMMARY OF THE EXPERIMENTS WITH 802.11AX.

Scenario	# Frames per Exp.	Inter-frame (ms)	Tx/Rx Positions
E20.0/0…7	∼ 19000	∼33	R
E40.0/0,1,2	∼ 18000	∼33	R
E80.0/0	∼ 18000	∼33	R
E80.0/1…9	∼ 6000	10	B (1-3), G (4-9)
E80.0/10,11	∼ 12000	5	B
S20.1/0,1,2	∼ 19000	∼33	R
S40.1/0	∼ 18000	∼33	R
S40.2/0,1	∼ 18000	∼33	R
S40.3/0	∼ 18000	∼33	R
S40.4/0,1	∼ 18000	∼33	R
S40.5/0,1	∼ 18000	∼33	R
S80.1/0	∼ 18000	∼33	R
S80.1/1…8	∼ 6000	10	B (1-3), G (4-8)
S80.1/9,10,11	∼ 12000	5	B
S80.2/0,1	∼ 18000	∼33	R
S80.2/2,3,4	∼ 12000	5	B
S80.2/5,6	∼ 6000	10	B
S80.3/0	∼ 18000	∼33	R
S80.4/0	∼ 18000	∼33	R
S80.5/0	∼ 2700	∼22	R
M20.4/0…3	∼ 19000	∼33	R
M80.1/0…7	∼ 6000	10	B (0-2), G (3-7)
M80.1/8,9	∼ 12000	5	B

BW	# Experiments	# CSI samples	Tot. Duration
20 MHz	15	248483	∼ 150 min.
40 MHz	11	200313	∼ 110 min.
80 MHz	44	384480	∼ 107 min.

[2]In the 80 MHz experiments we identified an anomalous behavior of the subcarriers close to the central ones suppressed in transmission (from 511 to 515): in some CSI samples they have a very large amplitude that is not coherent with expectations. The percentage of CSI that display this behavior in any experiment is small, so they do not invalidate experiments. We do not have an explanation for this behavior, which can be caused by anomalies in transmission, reception or CSI extraction. The published dataset at https://zenodo.org/records/17225694 is complete, but in this work we do not consider the subcarriers with indices 500 to 524 to avoid biasing the results.

the number of bits used in this process influences the mutual information metric we introduce in this paper.

After normalization and quantization, $A_{\mathbf{C}}(k,\cdot)$ becomes a vector of N_{SC} integer numbers represented over q bits.

IV. A Mutual Information Metric

The quantized representation of a CSI introduced in Sect. III-B can be used to quantify the amount of information it embeds. We made a first attempt in this direction in [31], exploiting a notion of Hamming distance between the amplitudes of different CSI's. The results were encouraging, but a notion of distance is relatively crude. Here we want to explore a different direction, exploiting the notion of entropy and mutual information derived from information theory. The ideal method would be to compute a stochastic metric that indicates the probability that a specific CSI belongs to the family of CSI's characterizing a specific scenario. A direct approach, however, as discussed in [31], is numerically unfeasible given the dimensionality of the problem; thus, we seek a methodology that reduces the combinatorial dimension of the problem, trading off potential precision with tractability.

Let $\mathcal{A} = [0, 1, \ldots, 2^q - 1]$ be the generic alphabet of the subcarriers amplitude representation, and $\mathcal{A}_i$ the specific alphabet of CSI $A_{\mathbf{C}}(i,\cdot)$, i.e., alphabet $\mathcal{A}$ characterized by the CSI-specific amplitudes distribution.

Next, we build a normalized co-occurrence matrix $\mathbf{M}_{i,j}$ of size $|\mathcal{A}| \times |\mathcal{A}|$, where each element

$$\frac{\mathbf{M}_{i,j}(a,b)}{N_{\text{SC}}} \tag{1}$$

counts how many subcarriers $n \in [1, N_{\text{SC}}]$ satisfy the condition $A_{\mathbf{C}}(i,n) = a \wedge A_{\mathbf{C}}(j,n) = b$.

By construction $\mathbf{M}_{i,j}$ is a probability matrix because

$$\sum_{a\in\mathcal{A}_i, b\in\mathcal{A}_j} \mathbf{M}_{i,j}(a,b) = 1 \tag{2}$$

Furthermore, both marginal vectors

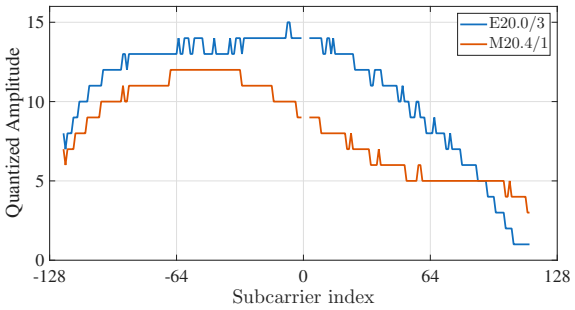
$$\mathbf{M}_i(a) = \sum_{b\in\mathcal{A}_j} \mathbf{M}_{i,j}(a,b) \;;\; \mathbf{M}_j(b) = \sum_{a\in\mathcal{A}_i} \mathbf{M}_{i,j}(a,b) \tag{3}$$

are probability vectors.

If $A_{\mathbf{C}}(i,\cdot) = A_{\mathbf{C}}(j,\cdot)$, only the elements on the main diagonal of $\mathbf{M}_{i,j}$ are non-null. Moreover, considering a perfectly equalized channel, all subcarriers have the same amplitude and only a single element $\mathbf{M}_{i,j}(a^\star, b^\star)$ is non-null and equal to 1; if $A_{\mathbf{C}}(i,\cdot) = A_{\mathbf{C}}(j,\cdot)$, then $a^\star = b^\star$.

$\mathbf{M}_{i,j}$ yields an empirical estimation of the similarity of $A_{\mathbf{C}}(i,\cdot)$ and $A_{\mathbf{C}}(j,\cdot)$ expressed as a joint probability matrix. Clearly, the reduction in numerical complexity implies some loss of information. In fact, by counting the occurrences of (a,b) amplitude pairs, the information on the subcarrier position is lost; however, this simplification allows to compute a mutual information metric between CSI's through $\mathbf{M}_{i,j}$.

To exemplify the representation we are proposing, Fig. 2 reports the amplitude of two CSI's drawn from E20.0/3 and M20.4/1 and quantized over $q = 4$ bits, alongside the resulting



2^q levels	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	$\mathbf{M}_j(b)$
0																	**0**
1				2	7												**9**
2					3	1											**4**
3						5											**5**
4						4											**4**
5						5											**5**
6						11	1										**12**
7						7	3										**10**
8						10		4									**14**
9						6	1	1	2								**10**
10						3	4		4	1							**12**
11							11			6	1						**18**
12							3	9			11						**23**
13								3	10		1	23	22				**59**
14									4	12	10	15	14				**55**
15											2						**2**
$\mathbf{M}_i(a)$	**0**	**0**	**0**	**2**	**10**	**52**	**23**	**17**	**20**	**19**	**25**	**38**	**36**	**0**	**0**	**0**	**242**

Figure 2. Amplitudes of two CSI samples from experiments E20.0/3 and M20.4/1 quantized over $q = 4$ bits and the relative matrix $\mathbf{M}_{i,j}(a,b)$; the rightmost column and bottom row are the marginal distributions and the bottom-right cell confirms that each CSI has $N_{\text{SC}} = 242$ valid subcarriers. To improve readability $\mathbf{M}_{i,j}(a,b)$ entries are not normalized, i.e., they are the pure number of occurrences.

matrix $\mathbf{M}_{i,j}(a,b)$. Even using a very coarse quantization to make $\mathbf{M}_{i,j}$ readable, it is clear that it contains well-structured information on the relationship between the two CSIs. Recall that rows and columns of $\mathbf{M}_{i,j}$ are not the N_{SC} index, but the quantization levels. Indeed, increasing q too much is detrimental: when $2^q \gg N_{\text{SC}}$, $\mathbf{M}_{i,j}$ becomes sparse with non-zero elements all equal to $1/N_{\text{SC}}$, making the Mutual Information (MI) (see Eqs. (4) and (5)) always maximal. To avoid this, we use small $q \in [6, 9]$ in Sect. VI.

With a little abuse of notation, we compute the entropy of the alphabets $\mathcal{A}_i, \mathcal{A}_j$ according to Shannon's definition:

$$\begin{aligned} H(\mathcal{A}_i) &= -\sum_{a\in\mathcal{A}_i} \mathbf{M}_i(a)\cdot\log_2[\mathbf{M}_i(a)], \\ H(\mathcal{A}_j) &= -\sum_{b\in\mathcal{A}_j} \mathbf{M}_j(b)\cdot\log_2[\mathbf{M}_j(b)], \\ H(\mathcal{A}_i,\mathcal{A}_j) &= -\sum_{a\in\mathcal{A}_i, b\in\mathcal{A}_j} \mathbf{M}_{i,j}(a,b)\cdot\log_2[\mathbf{M}_{i,j}(a,b)], \end{aligned} \tag{4}$$

and then the MI using its entropy-based formulation:

$$I(\mathcal{A}_i,\mathcal{A}_j) = H(\mathcal{A}_i) + H(\mathcal{A}_j) - H(\mathcal{A}_i,\mathcal{A}_j). \tag{5}$$

The abuse of notation lies, on one side, in defining the entropy of CSI's based on the matrix $\mathbf{M}_{i,j}$, which by construction disregards some properties of the CSI and is thus a (lossy) compressed representation. On the other side, it lies in defining

the entropy on a single CSI, and not on the entire population of CSI's in an experiment. Thus we have a mathematically correct formulation, but the interpretation through information theory may be non-trivial.

Discussing the properties of Eqs. (4) and (5), if the channel is perfectly equalized, then $H(\mathcal{A}_i) = 0$ because only one element $\mathbf{M}_i(a^\star) = 1$ and all others are 0. Also, a simple attenuation or amplification of the CSI does not yield any mutual information in this framework, as $H(\mathcal{A}_i) = H(\mathcal{A}_j) = H(\mathcal{A}_i, \mathcal{A}_j) = 0$. This is coherent with the intuition that a perfectly equalized channel does not bear information on the propagation environment, because in its corresponding physical modeling the "environment" is the empty space.

By definition, the MI is non-negative, therefore deriving its lower bound (0) is immediate. Deriving the upper bound is less intuitive, but not difficult, as follows. Conditional entropies are always non-negative and MI is symmetrical:

$$\begin{aligned} I(\mathcal{A}_i;\mathcal{A}_j) &= H(\mathcal{A}_j) - H(\mathcal{A}_j|\mathcal{A}_i) \leq H(\mathcal{A}_j) \\ I(\mathcal{A}_j;\mathcal{A}_i) &= H(\mathcal{A}_i) - H(\mathcal{A}_i|\mathcal{A}_j) \leq H(\mathcal{A}_i) \\ I(\mathcal{A}_j;\mathcal{A}_i) &= I(\mathcal{A}_i;\mathcal{A}_j) \leq \min\{H(\mathcal{A}_j), H(\mathcal{A}_i)\}\ \text{[bit]}. \end{aligned} \tag{6}$$

Since we quantize CSI amplitudes on $|\mathcal{A}| = 2^q$ values, the maximum MI we can obtain from $\mathbf{M}_{i,j}$ is q bits. This upper bound can only be reached for identical CSI's if the probability distribution of the 2^q quantized values is uniform, i.e., $\mathbf{M}_i(a) = \frac{1}{2^q} \quad \forall a \in [0, 2^q - 1]$. In this framework, the maximum entropy and thus the maximum MI of a CSI with itself can only be achieved if $N_{\text{SC}} = k \cdot 2^q$ with k integer, as this is the only condition that enables a uniform distribution over $\mathbf{M}_i(a)$. If, instead, $N_{\text{SC}} \neq k \cdot 2^q$, the entropy $H(\mathcal{A}_i)$ in Eq. (4) is strictly lower than the theoretical maximum.

Specifically, if $2^q > N_{\text{SC}}$, we can write $2^q = m \cdot N_{\text{SC}} + r$, with $m = \lfloor \frac{2^q}{N_{\text{SC}}} \rfloor$, $0 < r < N_{\text{SC}}$. In this case, to maximize the entropy, values in $\mathcal{A}_i$ have probability

$$\mathbf{M}_i(a) = \begin{cases} \frac{1}{N_{\text{SC}}} & \text{for } N_{\text{SC}} \text{ values} \\ 0 & \text{for } 2^q - N_{\text{SC}} \text{ values} \end{cases} \tag{7}$$

Hence,

$$H_{max}(\mathcal{A}_i) = \log_2[N_{\text{SC}}] < \log_2[2^q] = q \tag{8}$$

Instead, if $2^q < N_{\text{SC}}$, we can write $N_{\text{SC}} = n \cdot 2^q + s$, with $n = \lfloor \frac{N_{\text{SC}}}{2^q} \rfloor$, $0 < s < 2^q$. To keep the distribution as close to uniform as possible, values in $\mathcal{A}_i$ will have probability

$$\mathbf{M}_i(a) = \begin{cases} \frac{n}{N_{\text{SC}}} & \text{for } 2^q - s \text{ values} \\ \frac{n+1}{N_{\text{SC}}} & \text{for } s \text{ values} \end{cases} \tag{9}$$

Hence,

$$\begin{aligned} H_{max}(\mathcal{A}_i) = &-s \cdot \frac{n+1}{N_{\text{SC}}} \cdot \log_2 \left[\frac{n+1}{N_{\text{SC}}} \right] \\ &- (2^q - s) \cdot \frac{n}{N_{\text{SC}}} \cdot \log_2 \left[\frac{n}{N_{\text{SC}}} \right] \\ &< \log_2[2^q] = q \end{aligned} \tag{10}$$

In general, since some subcarriers are suppressed, we never have $2^q = k \cdot N_{\text{SC}}$, always enforcing $H_{max}(\mathcal{A}_i) < q$.

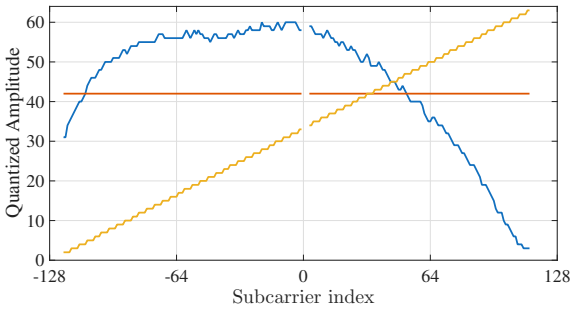


Figure 3. $A_{\mathbf{C}}(k, \cdot)$ from experiment E20.0/3 (blue) and two artificial CSI: maximal entropy (yellow) and linear channel (red) quantized over $q = 6$ bits.

To better grasp the characteristics of entropy and MI computed on $\mathbf{M}_{i,j}$, we focus in Fig. 3 on three different CSI quantized on $q = 6$ bits. In blue is a CSI from E20.0/3, in red an ideal-channel CSI and in yellow an artificial CSI that maximizes the entropy. Let us call $\mathcal{A}_b$, $\mathcal{A}_r$, and $\mathcal{A}_y$ their respective alphabets. We have $H(\mathcal{A}_b) = 5.08$, $H(\mathcal{A}_r) = 0$, and $H(\mathcal{A}_y) = 5.99$. As expected, free-space propagation bears no information, while a normal CSI has an entropy very close to the maximum, thus we can expect MI computed on $\mathbf{M}_{i,j}$ to give meaningful insight on CSI collected in different scenarios.

V. Reference CSI

To characterize scenarios, we need to identify a fingerprint representing all CSI that regard a scenario. The simplest choice is defining the reference as the average of $A_{\mathbf{C}}(k, \cdot)$ over an experiment $\mathbf{C}$:

$$A^\star_{\mathbf{C}}(\cdot) = \frac{1}{M_{\mathbf{C}}} \sum_{k=1}^{M_{\mathbf{C}}} A_{\mathbf{C}}(k, \cdot). \tag{11}$$

This way we can define the MI $I(\mathcal{A}_i; \mathcal{A}^\star_s)$ between an arbitrary CSI i and a reference for a specific scenario s, where $\mathcal{A}^\star_s$ is the specific alphabet computed on $A^\star_{\mathbf{C}s}$ of an experiment selected as representative of scenario s.

VI. Results

Since the complete dataset amounts to over 800 000 CSI samples and over 6 hours of data collection, only a subset of the experiments analysis can be presented. However, all results are tested on the whole dataset to guarantee general validity.

We first report the analysis of the distribution of the MI computed across different experiments. Focusing on 80 MHz bandwidth with $q = 9$ quantization bits, we randomly select E80.0/4 among the empty room collections as reference experiment and compute its $A^\star_{\mathbf{C}s}$. Figure 4 compares the MI distribution computed between several experiments of the green configuration and the reference; each point of the distribution is computed as $I(\mathcal{A}_i; \mathcal{A}^\star_E)$ with i spanning all the CSI's of the experiment under consideration. The MI is maximal and remarkably similar between empty-room experiments (blue) and decreases when a person is sitting in the room (red) or people are moving around (yellow).

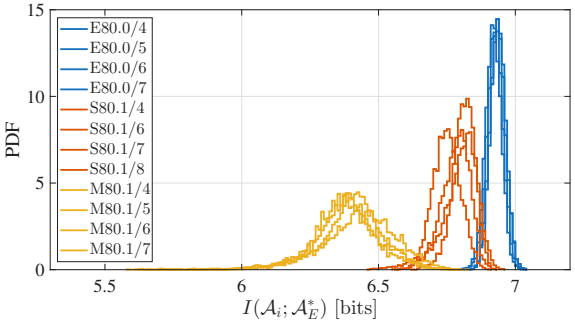


Figure 4. Distribution of the mutual information $I(\mathcal{A}_i;\mathcal{A}_E^\star)$, where $\mathcal{A}_E^\star$ refers to experiment E80.0/4 (top); $q = 9$ bits.

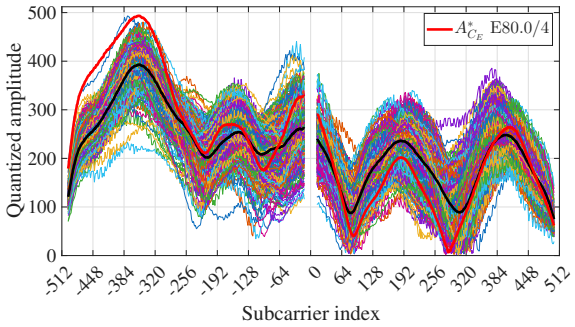


Figure 5. Representation of 10% of experiment M80.1/7 $A_{\mathbf{C}}$ with the relative $A_{\mathbf{C}}^\star$ in black. In red, $A_{\mathbf{C}}^\star$ computed on E80.0/4; $q = 9$ bits.

Given the nice properties of $I(\mathcal{A}_i;\mathcal{A}_E^\star)$ highlighted in Fig. 4, a classification based on $I(\mathcal{A}_i;\mathcal{A}^\star)$, with $\mathcal{A}^\star$ derived from an experiment representing a specific situation, might seem robust enough to solve many, if not most, of the interesting ambient sensing tasks. Unfortunately, this is not true, at least unless the situations are crafted at art, and Fig. 5 helps explaining why. When an experiment is not characterized by a stationary – in the sense of a system regime – situation (e.g., a specific movement repeated constantly during the experiment), but rather by a normal everyday life activity (e.g., people moving randomly in the room), its characterization through $A_{\mathbf{C}}^\star$ tends in the limit to the empty room $A_{\mathbf{C}}^\star$: all the CSI's of the experiment contain the empty room features plus – or, rather, distorted by – some randomness impressed by the specific situation 'snapshot' when the frame was transmitted. This is valid in general for every situation. Figure 5 reports roughly 600 $A_{\mathbf{C}}$ uniformly selected throughout experiment M80.1/7; we do not plot all $A_{\mathbf{C}}$ simply because the figure size in Mbytes would be too large to include. In black is the $A_{\mathbf{C}}^\star$ of the same experiment, compared with $A_{\mathbf{C}}^\star$ of experiment E80.0/4 in red. All CSI's are quantized over $q = 9$ bits. It is clear that the two averages are quite similar even though experiments are taken at different times, in very different scenarios; hence, conditioning $I(;)$ on experiments that were not performed in the empty-room scenario is meaningless.

We think this observation is important and influences any work on sensing based on CSI analysis. As future work we plan to study appropriate methodologies to pre-process CSI samples to remove the empty-room fingerprint to help classification and recognition. Unfortunately, simply subtracting $A_{\mathbf{C}}^\star$ of the empty room does not seem to yield reliable results.

To complete the analysis, Fig. 6 reports an evaluation conducted on the entire dataset of the average MI computed on $I(\mathcal{A}_i;\mathcal{A}_E^\star)$ measures. Each of the six plots reported in Fig. 6 refers to a group of experiments that make comparison and classification meaningful, i.e., same position of Tx/Rx and same bandwidth. In each group, one empty-room experiment is taken as reference together with its alphabet $\mathcal{A}_E^\star$; then, all MI $I(\mathcal{A}_i;\mathcal{A}_E^\star)$ are computed, splitting every experiment into 50 different batches of CSI's, so that confidence intervals can be computed to verify the reliability of results. This procedure is repeated for $q = 6, 7, 8, 9$ to analyze the impact of quantization granularity also as a function of the bandwidth: since it obviously changes N_{SC}, it influences the significance of MI, as discussed in Sect. IV. Each of the points in the plots reports the average MI and the confidence interval with a confidence level of 99%. Indeed, the confidence intervals are so small that they cannot be distinguished from the marks of the point estimate, indicating that the average MI measures are extremely stable even during long experiments.

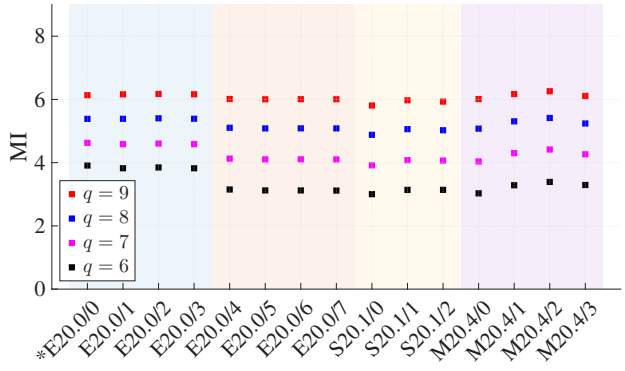
The most interesting result emerging from Fig. 6 is the choice of q confirming the theoretical analysis: a very fine-grained quantization makes the MI between any experiment and the empty-room reference nearly equal, making it impossible to gain enough information to tell scenarios apart. Switching to lower q, the scenarios become easier to discern, with the best outcomes coming from 80 MHz in LoS conditions. The lack of LoS impairs the 'cleanliness' of the MI metric: the empty room remains easy to identify, with more difficulties in telling apart the sitting and moving scenarios.

Finally, everyday, common alterations of the layout of the environment (e.g., desks or chairs disposition) may make the empty room appear as a different scenario when comparing experiments conducted on different weeks, due to the high sensitivity of the CSI samples to changes in the environment. This is what happens for the 20 MHz bandwidth case in Fig. 6, where the proposed simple metric lacks detailed environment classification capabilities; this is also due to the reduced amount of information carried by CSI collected on a 20 MHz channel, which only provides 242 subcarriers for the estimation of $\mathbf{M}_{i,j}$ compared to the 1024 of an 80 MHz channel. In a people detection system based on the proposed metric, a simple solution to this problem would be to periodically update the reference $A_{\mathbf{C}}^\star$ to prevent the information on the 'footprint' of the empty room from going stale and consequently hinder new MI computations.

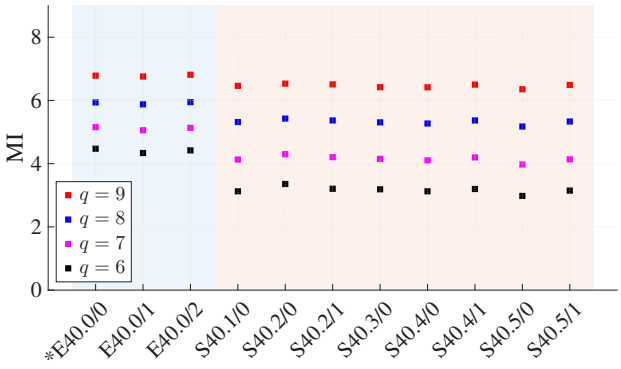
VII. Discussion and Future Work

This paper is not meant to draw conclusions, but to open a discussion and hopefully new research on CSI-based sensing to help JCAS exit academia and innovate telecommunications.

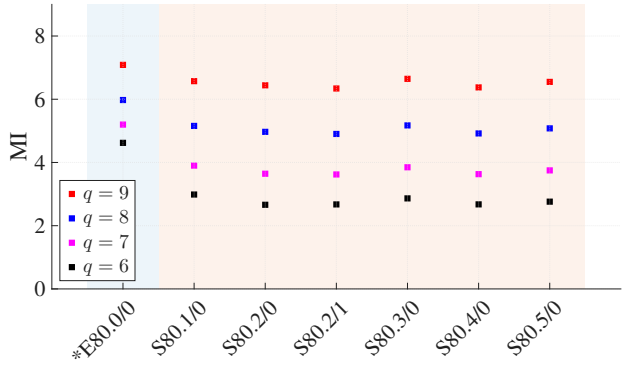
The work stems from a simple question: Can we characterize and understand the fundamental properties of the CSI



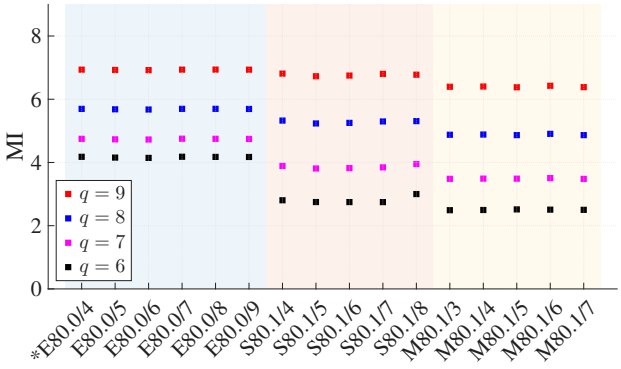
Red configuration, 20 MHz, LoS



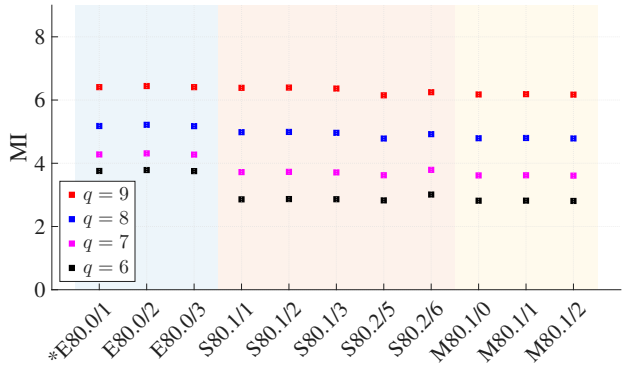
Red configuration, 40 MHz, LoS



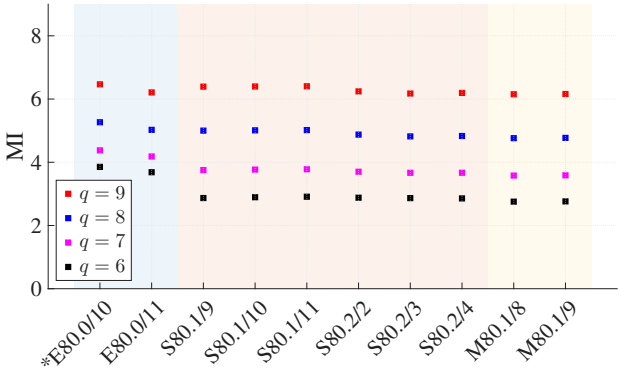
Red configuration, 80 MHz, LoS



Green configuration, 80 MHz, LoS



Blue configuration, 80 MHz, Non-LoS



Blue configuration, 80 MHz, Non-LoS

Figure 6. Average $I(\mathcal{A}_i; \mathcal{A}_E^\star)$ computed over all the experiments grouped by scenario and Tx/Rx position. The main features characterizing the group of experiments are reported under each plot. The reference experiment for each group is the first on the left marked with the $\star$.

that enable accurate AI-based sensing? To tackle this question, observing that approaches based on micro-propagation models like ray-tracing are extremely complex and often fail to generalize (e.g., because as frequency increases surfaces scatter and do not reflect), we started looking at stochastic and information theory models. This paper proposes a reduced-complexity model to measure mutual information between CSI's. The model is simple, prone to data-driven characterization, and computationally inexpensive. Early results show that the method is viable even with coarse quantization (fundamental to apply information theory results), with less quantization bits often yielding even better results, as discussed in light of information theory elements.

We think that ahead lies a vast area of research that can shed more light on CSI analysis foundations. Here we highlight some ideas that we want to explore ourselves, while also offering them for discussion to the community. First, we proposed a simple metric, but many others besides the matrix $\mathbf{M}_{i,j}$ we devised exist and just wait to be formalized and explored. Next, AI can be applied to the proposed metric, possibly yielding accurate classification alongside a stochastic interpretation. In this same line, it would be very interesting to apply successful AI methods from the literature to the dataset we share with the community, since this dataset is vast, varied, and collected in realistic conditions; similarly, it would be interesting to apply this method to other available

datasets. A further direction involves devising metrics based not only on the amplitude of single CSI subcarriers but on their full stochastic representation as a multi-dimensional stochastic process, including the correlation between the different measures and the time-frequency coherence. Finally, expanding the dataset and analysis to 160 and 320 MHz-wide channels with up to 4096 subcarriers in newer generations of Wi-Fi promises better accuracy and further insights. In this same line, extending the approach to 5G and 6G technologies where the CSI is not limited to preambles but is interspersed continuously in the wireless connection can shed light into differences between packet-oriented and circuit-oriented access technologies.

Acknowledgements

The authors wish to thank Prof. Marco Dalai for the useful discussions on Mutual Information properties and Prof. Francesco Gringoli for the technical support in CSI collection.

This work was partially supported by the European Union (EU) and the Italian Ministry for University and Research (MUR), National Recovery and Resilience Plan (NRRP), through the projects "PROSDS (CUP C89J24000270004) Spoke 4, Structural Project SUPER and EMBRACE (CUP E63C22002070006) Spoke 6, Structural Project SEXTET both part of RESTART (PE00000001);" "CSI-Future (Joint Communication and Sensing: CSI-Based Sensing for Future Wireless Networks) PRIN 2022 PNRR (CUP D53D23016040001);" and "Sustainable Mobility Center (MOST)", (CUP D83C22000690001), Spoke N^o 7, "CCAM, Connected networks and Smart Infrastructures".

References

[1] 802.11ax, "IEEE Standard for Information technology – Part 11: Wireless LAN MAC and PHY Specifications Amendment 1: Enhancements for High-Efficiency WLAN."

[2] 802.11be, "IEEE Draft Standard for Information technology – Part 11: Wireless LAN MAC and PHY Specifications Amendment: Enhancements for Extremely High Throughput (EHT)."

[3] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: Indoor Localization via Channel Response," *ACM Comp. Surv.*, vol. 46, no. 25, pp. 1–32, Dec. 2013.

[4] K. Wu, J. Xiao, Y. Yi, et al., "CSI-Based Indoor Localization," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 7, p. 1300–1309, Jul. 2013.

[5] F. Adib and D. Katabi, "See through walls with WiFi!" in *ACM Int. Conf. of the SIG on Data Communication (SIGCOMM)*, 8 2013, pp. 75–86.

[6] M. Abbas, M. Elhamshary, H. Rizk, et al., "WiDeep: WiFi-based Accurate and Robust Indoor Localization System using Deep Learning," in *IEEE Int. Conf. on Pervasive Computing and Communications (PerCom)*, 3 2019, pp. 1–10.

[7] F. Shi, W. Li, C. Tang, et al., "ML-Track: Passive Human Tracking Using WiFi Multi-Link Round-Trip CSI and Particle Filter," *IEEE Trans. on Mobile Comp.*, vol. 24, no. 6, pp. 5155–5172, 6 2025.

[8] W. Li, M. J. Bocus, C. Tang, et al., "On CSI and Passive Wi-Fi Radar for Opportunistic Physical Activity Recognition," *IEEE Trans. on Wireless Comm.*, vol. 21, no. 1, pp. 607–620, 1 2022.

[9] T. Nakamura, M. Bouazizi, et al., "Wi-Fi-Based Fall Detection Using Spectrogram Image of Channel State Information," *IEEE Internet of Things Jou.*, vol. 9, no. 18, pp. 17 220–17 234, 2022.

[10] H. Abdelnasser, M. Youssef, and K. A. Harras, "WiGest: A ubiquitous WiFi-based gesture recognition system," in *IEEE Conf. on Computer Comm. (INFOCOM)*, 4 2015, pp. 1472–1480.

[11] R. C. Shit, S. Sharma, D. Puthal, et al. "Ubiquitous Localization (UbiLoc): A Survey and Taxonomy on Device Free Localization for Smart World," *IEEE Comm. Surveys & Tutorials*, vol. 21, no. 4, pp. 3532–3564, 2019.

[12] F. Miao, Y. Huang, Z. Lu, et al., "Wi-Fi Sensing Techniques for Human Activity Recognition: Brief Survey, Potential Challenges, and Research Directions," *ACM Comp. Surv.*, no. 5, 1 2025.

[13] Y. Qiao, O. Zhang, W. Zhou, et al., "PhyCloak: Obfuscating Sensing from Communication Signals," in *13th USENIX Conf. on Networked Systems Design and Implementation (NSDI'16)*, 3 2016, pp. 685–699.

[14] M. Cominelli, F. Gringoli, and R. Lo Cigno, "Non Intrusive Wi-Fi CSI Obfuscation Against Active Localization Attacks," in *16th IFIP/IEEE Conf. on Wireless On-demand Network Systems and Services (WONS)*, 3 2021, pp. 87–94.

[15] P. Staat, S. Mulzer, S. Roth, et al., "IRShield: A countermeasure against adversarial physical-layer wireless sensing," in *2022 IEEE Symp. on Security and Privacy*, 5 2022, pp. 1705–1721.

[16] M. Cominelli, F. Gringoli, and R. Lo Cigno, "AntiSense: Standard-compliant CSI obfuscation against unauthorized Wi-Fi sensing," *Computer Communications*, vol. 185, pp. 92–103, 3 2022.

[17] J. Hu, H. Jiang, S. Chen, et al., "WiShield: Privacy Against Wi-Fi Human Tracking," *IEEE Jou. on Selected Areas in Comm.*, vol. 42, no. 10, pp. 2970–2984, 10 2024.

[18] J. Yang, X. Chen, H. Zou, et al., "EfficientFi: Toward Large-Scale Lightweight WiFi Sensing via CSI Compression," *IEEE Internet of Things Jou.*, vol. 9, no. 15, pp. 13 086–13 095, 2022.

[19] N. Bao, J. Du, C. Wu, et al., "Wi-breath: A wifi-based contactless and real-time respiration monitoring scheme for remote healthcare," *IEEE Jou. of Biomedical and Health Informatics*, vol. 27, no. 5, pp. 2276–2285, 2023.

[20] J. Ding and Y. Wang, "WiFi CSI-Based Human Activity Recognition Using Deep Recurrent Neural Network," *IEEE Access*, vol. 7, pp. 174 257–174 269, 12 2019.

[21] Z. Guo, F. Xiao, B. Sheng, et al., "Twcc: A robust through-the-wall crowd counting system using ambient wifi signals," *IEEE Trans. on Vehicular Technology*, vol. 71, no. 4, pp. 4198–4211, 2022.

[22] M. Cominelli, F. Gringoli, and F. Restuccia, "Exposing the CSI: A Systematic Investigation of CSI-based Wi-Fi Sensing Capabilities and Limitations," in *IEEE Int. Conf. on Pervasive Comp. and Comm. (PerCom)*, 3 2023, pp. 81–90.

[23] C. Chen, G. Zhou, and Y. Lin, "Cross-Domain WiFi Sensing with Channel State Information: A Survey," *ACM Comput. Surv.*, vol. 55, no. 11, 2 2023.

[24] F. Meneghello, D. Garlisi, N. D. Fabbro, et al., "Sharp: Environment and person independent activity recognition with commodity ieee 802.11 access points," *IEEE Trans. on Mobile Comp.*, vol. 22, no. 10, pp. 6160–6175, 10 2023.

[25] W. Li, R. Gao, J. Xiong, et al., "Wifi-csi difference paradigm: Achieving efficient doppler speed estimation for passive tracking," in *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 8, no. 2, pp. 1–29, 2024.

[26] F. Zhang, C. Wu, B. Wang, et al., "WiDetect: Robust Motion Detection with a Statistical Electromagnetic Model," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 3, no. 3, 9 2019.

[27] J. Huang, Y. Tian, X. Hu, Z. Li, "A Novel WiFi-Based Temporal Precision Detection System With Enhanced CSI Sampling," *IEEE Wireless Comm. Letters*, vol. 14, no. 1, pp. 128–132, 1 2025.

[28] F. Zhang, C. Chen, et al., "WiSpeed: A Statistical Electromagnetic Approach for Device-Free Indoor Speed Estimation," *IEEE Internet of Things Jou.*, vol. 5, no. 3, pp. 2163–2177, Jun. 2018.

[29] B. Chen, J. Wang, Y. Lv, et al., "Device-Free Wireless Sensing With Few Labels Through Mutual Information Maximization," *IEEE Internet of Things Jou.*, vol. 11, no. 6, pp. 10 513–10 524, 3 2024.

[30] Y. He, J. Liu, M. Li, et al., "SenCom: Integrated Sensing and Communication with Practical WiFi," in *29th ACM Annual Int. Conf. on Mobile Comp. and Networking (MobiCom)*, 10 2023.

[31] E. Tonini, F. Gringoli, R. Lo Cigno, and M. Cominelli, "Towards a Quantitative Analysis of CSI for AI/ML Based Sensing," in *IEEE Wireless Comm. and Networking Conf. (WCNC)*, 3 2025.

[32] J. Angjo, E. Tonini, F. Dressler, and R. Lo Cigno, "mmWave CSI-based Sensing:a Feasibility Study," in *31st ACM Annual Int. Conf. on Mobile Comp. and Networking (MobiCom 2025), Poster*, 11 2025.

[33] M. Schulz, D. Wegemer, and M. Hollick. Nexmon: The C-based Firmware Patching Framework, 5 2017. Available: https://nexmon.org

[34] E. Tonini, "Statistical Analysis to Support CSI-Based Sensing Methods," Master's thesis, University of Brescia, arXiv:2411.03203, 9 2024.