VersaBeam: Versatile Beamforming for Integrated Sensing and Communication over Commodity Wi-Fi

Yinghui He[1*] Mingming Xu[2*] Fu Xiao[2†] Jun Luo[1†]
[1]College of Computing and Data Science, Nanyang Technological University (NTU), Singapore
[2]School of Computer Science, Nanjing University of Posts and Telecommunications, China
Email: {yinghui.he, junluo}@ntu.edu.sg, {2020070137, xiaof}@njupt.edu.cn

***Abstract*—Reusing Wi-Fi communication packets for sensing purpose has been regarded as one of the most cost-effective ways to realize *integrated sensing and communication* (ISAC) on commodity Wi-Fi. However, the *channel state information* (CSI) measured from these packets can be heavily compromised by modern Wi-Fi beamforming protocols tailored primarily to maximize communication throughput, hence inadvertently affecting Wi-Fi sensing performance. To this end, we propose VersaBeam, a practical ISAC system actively leveraging beamforming to *simultaneously* achieve high-performance communication and sensing in commodity Wi-Fi. Instead of attempting to mitigate beamforming's negative impact on sensing, VersaBeam features a novel beamforming design to balance the demands of both sensing and communication. Additionally, we propose a user selection strategy to determine the sources for packet reuse and introduce a method to unify CSI measured from different users' packets, accommodating variations in beamforming direction. We implement a prototype of VersaBeam on commodity Wi-Fi devices and further demonstrate its efficacy through micro-benchmarking and practical experiments in gesture recognition.**

***Index Terms*—Wi-Fi sensing, wireless communication, beamforming, integrated sensing and communication (ISAC).**

I. Introduction

Over the past few decades, Wi-Fi devices have been widely deployed due to the low cost and high-performance communication capabilities [1], [2]. This widespread deployment has also attracted attention for capabilities beyond mere communication. Recent research has confirmed that Wi-Fi can support various sensing applications, including detection [3], [4], recognition [5]–[7], and estimation [8] by extracting channel state information (CSI) from Wi-Fi packets [9]. However, most Wi-Fi sensing proposals rely on specialized settings (e.g., dedicated packets) for sensing purpose. Consequently, they typically realize sensing and communication leveraging separate spectrum or temporal resources, thus significantly limiting efficiency for both. To this end, integrated sensing and communication (ISAC) has emerged, aiming to efficiently sense and communicate simultaneously for the future development of Wi-Fi technology [10]–[14].

To realize Wi-Fi ISAC, one of the most cost-effective methods is to directly extract CSI from ongoing Wi-Fi communication packets, eliminating the need for additional spectrum or temporal resources for sensing. Nevertheless, modern Wi-Fi designs aimed at enhancing communication can adversely affect sensing performance [15]. A notable example is communication-oriented beamforming: it enhances throughput and reliability through directional transmission for all Wi-Fi traffic compatible with beamforming [16]. When the communication direction differs from the sensing direction from the perspective of a Wi-Fi access point (AP), the power directed toward the sensing area is limited under the communication-oriented beamforming. Consequently, channel variations caused by target movements are often obscured by noise or other interference, resulting in a low signal-to-noise ratio (SNR) at a sensing receiver. Recently, SenCom [17], [18] employs signal processing techniques on received signals to mitigate the negative impact of beamforming on the sensing performance, passively addressing the challenge. Unfortunately, these techniques at the sensing receiver cannot fundamentally resolve the low SNR issue, as the noise and other interference are mixed with channel variations and cannot be totally eliminated, especially when the transmit power towards the sensing direction approaches zero.

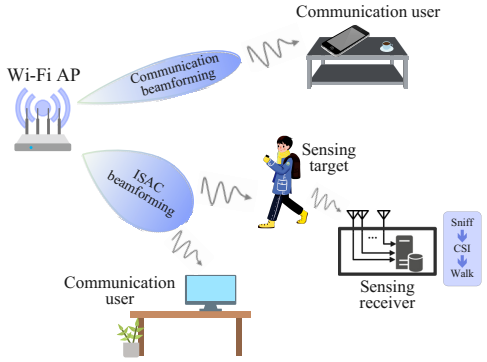


Fig. 1: VersaBeam actively leverages versatile beamforming to realize high-performance sensing and communication in practical Wi-Fi systems.

To overcome the limitation of the existing work, we propose VersaBeam, a practical ISAC system that actively leverages beamforming to realize ISAC and reusing communication packets for sensing in commodity Wi-Fi systems.[1] As shown in Figure 1, beamforming direction is jointly determined by both

* Both authors contributed equally to this research, which is done when Mingming Xu works as a visiting scholar at NTU, under the funding support offered by Jiangsu Provincial Government Scholarship.
† Fu Xiao and Jun Luo are both the corresponding authors.

[1]For Wi-Fi APs not allowing for tuning beamforming protocols, we can forge beamforming feedback to control the beamforming [19].

communication and sensing needs to achieve high performance of both functions. However, we still face three challenges in realizing this idea. First, the theoretical modeling of the relationship between beamforming and Wi-Fi sensing's SNR has not been adequately explored, which raises a barrier to further beamforming design. Secondly, we aim to enhance sensing performance with minimal impact on communication, a challenging balance since both functions are sensitive to beamforming direction. Thirdly, the ISAC-oriented beamforming matrix varies with user location, leading to inconsistencies in CSI and SNR measurements even with the same sensing targets, complicating the establishment of a reliable sensing deep neural model.

To address these challenges, we begin with a theoretical analysis of beamforming's impact on Wi-Fi sensing, confirmed through experimental results. To minimize the impact on communication caused by controlling beamforming, the communication SNR loss is restricted according to the table of Modulation Coding Scheme (MCS) index and related SNR threshold. We also develop an optimization framework that jointly considers sensing and communication performance to design an ISAC-oriented beamforming feedback. For the final challenge of inconsistent CSI and sensing SNR, we introduce a user selection strategy to ensure high sensing SNR under the requirement of the CSI sampling rate and then propose a method to unify CSI to address inconsistencies. In summary, we make the following major contributions:

- We propose VersaBeam as an effective ISAC solution in practical systems by actively leveraging the beamforming and reusing the communication packets.
- We conduct a mathematical analysis of beamforming's impact on Wi-Fi sensing, with findings supported by simulation results.
- We explore an optimization problem to design an ISAC-oriented beamforming strategy for enhancing sensing performance with minimal impact on communication.
- We propose a user selection strategy and a CSI unifying method for resolving the inconsistencies in CSI measured from different users' packets.
- We implement VersaBeam prototype with commodity Wi-Fi devices and extensive experiments confirm that VersaBeam achieves high-performance sensing with minimal impact on communication.

The rest of the paper is organized as follows. Section II explores the effect of beamforming on Wi-Fi sensing. Section III presents the design of VersaBeam in detail. Section IV and Section V explain VersaBeam's implementation and report the extensive evaluations on VersaBeam, respectively. Related works are briefly captured in Section VI, and the whole paper is concluded in Section VII.

II. Exploring Beamforming for ISAC

In this section, we first introduce the existing Wi-Fi beamforming basics and then establish the sensing SNR model under the beamforming. Finally, the impacts of beamforming on sensing and communication are analyzed with numerical results.

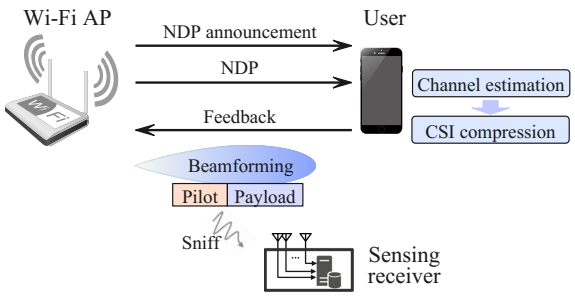


Fig. 2: Wi-Fi beamforming protocol and CSI collection of Wi-Fi sensing.

A. Wi-Fi Beamforming Basic

Beginning with the 802.11ac standard, singular value decomposition (SVD) beamforming, i.e., a typical communication-oriented beamforming technique, has been integrated into Wi-Fi systems [20]. This integration necessitates channel information between the Wi-Fi AP and the user. To facilitate it, a channel sounding protocol has been introduced into Wi-Fi systems. As shown in Figure 2, the protocol initiates with a null data packet (NDP) announcement transmitted by the Wi-Fi AP, followed by the broadcast of an NDP containing predefined pilots to the user. The user then estimates the channel between the AP and themselves by comparing the known pilots with the received signals. The measured CSI can be denoted by $\boldsymbol{H}_i \in \mathbb{C}^{M_i^{\mathrm{R}} \times M^{\mathrm{T}}}$, where M^{T} and M_i^{R} denote the numbers of antennas at the Wi-Fi AP and the i-th user, respectively.[2] Instead of transmitting $\boldsymbol{H}_i$ back to the Wi-Fi AP, resulting in high overhead, its right singular vectors, denoted by $\boldsymbol{V}_i \in \mathbb{C}^{M^{\mathrm{T}} \times M_i^{\mathrm{R}}}$, is fed back, and $\boldsymbol{V}_i$ is obtained using SVD, as

$$\boldsymbol{H}_i = \boldsymbol{U}_i \boldsymbol{\Sigma}_i \boldsymbol{V}_i^H, \tag{1}$$

where $(\cdot)^H$ is the operation of conjugate transpose, $\boldsymbol{\Sigma}_i \in \mathbb{C}^{M_i^{\mathrm{R}} \times M_i^{\mathrm{R}}}$ is a nonnegative real diagonal matrix, and $\boldsymbol{U}_i \in \mathbb{C}^{M_i^{\mathrm{R}} \times M_i^{\mathrm{R}}}$ is a unitary matrix. Note that $\boldsymbol{V}_i$ is unitary, i.e., satisfying $\boldsymbol{V}_i^H \boldsymbol{V}_i = \boldsymbol{I}$. The diagonal elements of $\boldsymbol{\Sigma}_i$, denoted by $\rho_{i,m}$, represent the channel gain of data transmission with beamforming, and the average SNRs derived from $\boldsymbol{\Sigma}_i$ are also fed back to the AP [20].

Upon receiving the beamforming feedback, the Wi-Fi AP performs data transmission using $\boldsymbol{V}_i$ as the beamforming matrix.[3] Specifically, the data packet contains two parts, i.e., pilot and payload. Specifically, the predefined pilot, denoted

[2]We ignore the subscript for subcarrier, and the proposed methods in this paper can be readily applied to the other subcarriers independently.

[3]There are two modes of beamforming in Wi-Fi systems, i.e., single-user multiple-in multiple-out (SU-MIMO) and multi-user multiple-in multiple-out (MU-MIMO). The primary distinction between two modes is the number of users served simultaneously [21]. Since the effect of beamforming on Wi-Fi sensing is similar in both modes, we mainly focus on the SU-MIMO mode, and the proposed method can be extended to the MU-MIMO mode as well.

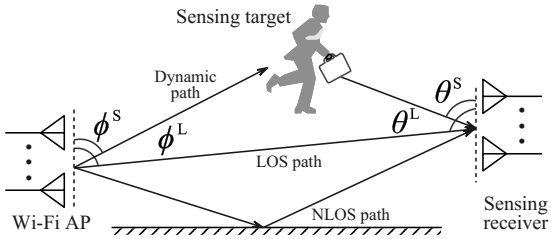


Fig. 3: Wireless channel model between the Wi-Fi AP and the sensing receiver with the sensing target.

by $\boldsymbol{S}_i^{\mathrm{p}} \in \mathbb{C}^{M_i^{\mathrm{R}} \times M_i^{\mathrm{R}}}$, is used for channel estimation and demodulating the subsequent payload. Let $\boldsymbol{s}_i^{\mathrm{D}} \in \mathbb{C}^{M_i^{\mathrm{R}} \times 1}$ denote the payload with the transmit power being $\mathbb{E}\left\{\boldsymbol{s}_i^{\mathrm{D}}(\boldsymbol{s}_i^{\mathrm{D}})^H\right\} = p\boldsymbol{I}/M_i^{\mathrm{R}}$. After undergoing beamforming matrix $\boldsymbol{V}_i$ and wireless channel $\boldsymbol{H}_i$, the received signal at the i-th user is given by:

$$\boldsymbol{y}_i = \boldsymbol{H}_i \boldsymbol{V}_i \boldsymbol{s}_i + \boldsymbol{n}_i, \tag{2}$$

where $\boldsymbol{n}_i$ is Gaussian noise.

B. Sensing with Beamforming

We now turn our focus to the CSI at the sensing receiver. As depicted in Figure 2, the sensing receiver measures the CSI between itself and the Wi-Fi AP by sniffing packets transmitted by the AP and comparing the received pilot with the predefined one, i.e., $\boldsymbol{S}_i^{\mathrm{p}}$. The received pilot of the sensing receiver at time t is

$$\boldsymbol{Y}_i^{\mathrm{S}}(t) = \boldsymbol{H}^{\mathrm{S}}(t)\boldsymbol{V}_i\boldsymbol{S}_i^{\mathrm{p}} + \boldsymbol{N}(t), \tag{3}$$

where $\boldsymbol{N}_i(t)$ is Gaussian noise and $\boldsymbol{H}^{\mathrm{S}}(t) \in \mathbb{C}^{M^{\mathrm{S}} \times M^{\mathrm{T}}}$ is the physical-world wireless channel between the Wi-Fi AP and the sensing receiver with M^{S} being the number of antennas at the sensing receiver. Consequently, the measured CSI at the sensing receiver becomes $\boldsymbol{H}^{\mathrm{S}}(t)\boldsymbol{V}_i + \boldsymbol{N}^{\mathrm{S}}(t)$, instead of $\boldsymbol{H}^{\mathrm{S}}(t)$, where $\boldsymbol{N}^{\mathrm{S}}(t) = \boldsymbol{N}(t)(\boldsymbol{S}_i^{\mathrm{p}})^{-1}$. Thus, the sensing performance would be influenced by the beamforming matrix $\boldsymbol{V}_i$. To fully understand this impact, it is necessary to develop a comprehensive sensing model and derive the sensing SNR expression under the beamforming.

As shown in Figure 3, we start by modeling the channel $\boldsymbol{H}^{\mathrm{S}}(t)$ as the linear combination of the line-of-sight (LOS) path, denoted by $\boldsymbol{H}^{\mathrm{S,L}}$, the dynamic path related to the sensing target, denoted by $\boldsymbol{H}^{\mathrm{S,D}}(t)$, and other non-line-of-sight (NLOS) paths, denoted by $\boldsymbol{H}^{\mathrm{S,N}}$. Specifically, for the dynamic path, the angle of arrival (AOA) is θ^{S} and the angle of departure (AOD) is ϕ^{S}. $\boldsymbol{H}^{\mathrm{S,S}}$ can be given as

$$\boldsymbol{H}^{\mathrm{S,D}}(t) = g^{\mathrm{S,D}}(t)\boldsymbol{a}^{\mathrm{S,R}}(\boldsymbol{a}^{\mathrm{S,T}})^H, \tag{4}$$

where $g^{\mathrm{S,S}}(t)$ is the attenuation that varies over time due to the movement of the sensing target, $\boldsymbol{a}^{\mathrm{S,R}} = \boldsymbol{a}(\theta^{\mathrm{S}}, M^{\mathrm{R}})$, and $\boldsymbol{a}^{\mathrm{S,T}} = \boldsymbol{a}(\theta^{\mathrm{S}}, M^{\mathrm{T}})$ with $\boldsymbol{a}(\theta, M) = [e^{-2j\pi\cos(\theta)(m-1)d/\lambda}] \in \mathbb{C}^{M\times 1}$ is the steering vector with d being the antenna spacing and λ being the wavelength, and it describes phase difference at different antennas. For the LOS path, $\boldsymbol{H}^{\mathrm{S,L}}$ can be similarly expressed as $g^{\mathrm{S,L}}\boldsymbol{a}^{\mathrm{L,R}}(\boldsymbol{a}^{\mathrm{L,T}})^H$, and the AOA and AOD are θ^{L} and ϕ^{L}, respectively. The measured CSI under the beamforming can be expressed as

$$\begin{aligned}\boldsymbol{H}^{\mathrm{S,B}}(\boldsymbol{V}_i, t) =& \boldsymbol{H}^{\mathrm{S}}(t)\boldsymbol{V}_i = g^{\mathrm{S,L}}\boldsymbol{a}^{\mathrm{L,R}}(\boldsymbol{a}^{\mathrm{L,T}})^H\boldsymbol{V}_i + \boldsymbol{H}^{\mathrm{S,N}}\boldsymbol{V}_i \\ &+ g^{\mathrm{S,D}}(t)\boldsymbol{a}^{\mathrm{S,R}}(\boldsymbol{a}^{\mathrm{S,T}})^H\boldsymbol{V}_i + \boldsymbol{N}^{\mathrm{S}}(t).\end{aligned} \tag{5}$$

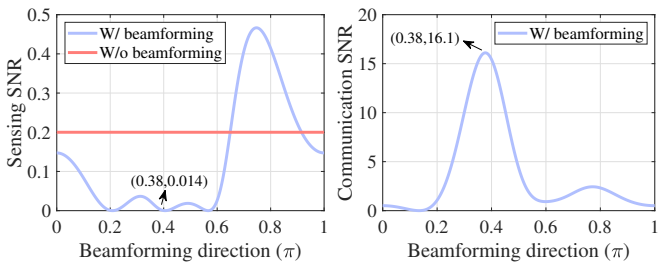
In order to characterize the effect of beamforming on sensing, we adopt the sensing SNR as the metric and it is defined as the power ratio of the dynamic path to the power sum of the LOS and other NLOS paths plus noise. According to [22], the power sum of other NLOS paths and noise is linearly proportional to that of the LOS path. Therefore, the sensing SNR can be calculated as

$$\eta^{\mathrm{S}}(\boldsymbol{V}_i) = \frac{\mathbb{E}\{|g^{\mathrm{S,D}}(t)|^2\}||(\boldsymbol{a}^{\mathrm{S,T}})^H\boldsymbol{V}_i||^2}{\xi||g^{\mathrm{S,L}}(\boldsymbol{a}^{\mathrm{L,T}})^H\boldsymbol{V}_i||^2 + b}, \tag{6}$$

where ξ and b are fixed parameters that reflect the power contributions of other NLOS paths and noise for a given pair of AP and receiver, and $\mathbb{E}\{\cdot\}$ denotes the operation of expectation. From Eqn. (6), it is evident that the sensing SNR is heavily influenced by the beamforming matrix $\boldsymbol{V}_i$. In fact, even worse, the sensing SNR could be reduced to zero if $\boldsymbol{V}_i$ is orthogonal to $\boldsymbol{a}^{\mathrm{S,T}}$. Thus, the sensing performance of SenCom [17] may be poor since it relies on communication-oriented beamforming.

C. Impact of Beamforming on Sensing and Communication

We can now analyze the impact of beamforming on sensing and communication with the help of simulation. For a more intuitive understanding, we adopt the phase shift beamforming [23] instead of SVD beamforming, despite its suboptimal performance for both sensing and communication, and there is only one communication user. Using this beamforming method, each column of $\boldsymbol{V}_1$ is equal to $\boldsymbol{a}(\theta, M)/\sqrt{M}$ with θ representing the beamforming direction. By setting $M^{\mathrm{T}} = 4$, $M_1^{\mathrm{R}} = 1$, $M^{\mathrm{S}} = 1$, $\phi^{\mathrm{L}} = \pi/2$, and $\phi^{\mathrm{S}} = 3\pi/4$, we illustrate the relationship between sensing SNR and beamforming direction in Figure 4a. For comparison, we also show sensing SNR without beamforming. It is evident that beamforming significantly impacts sensing, beneficially or detrimentally, depending on its direction. Specifically, when the beamforming direction aligns closer to the sensing direction ϕ^{S}, the sensing SNR improves with beamforming due to increased power allocation to the sensing direction. Conversely, when



(a) Sensing. (b) Communication.

Fig. 4: Impact of beamforming direction on (a) sensing and (b) communication.

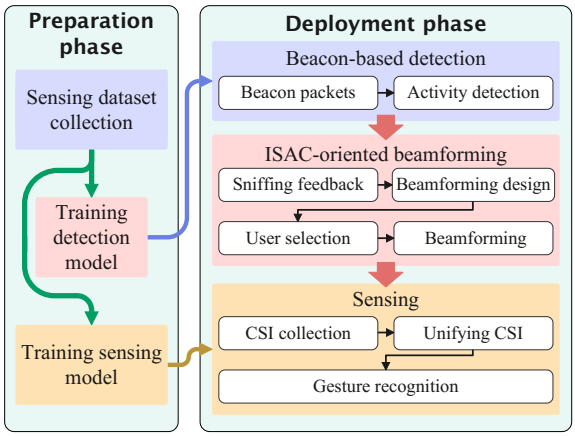


Fig. 5: Overview of VersaBeam.

the beamforming direction deviates from the sensing direction, sensing SNR decreases, potentially dropping to zero in three specific directions.

The impact of beamforming direction on communication is plotted in Figure 4b, where the channel between the Wi-Fi AP and the user is generated using the Rician fading channel model [24]. It can be seen that the impact of beamforming is also significant, and the beamforming direction should align with the communication direction so that the SNR can be maximized. However, from Figure 4, the desired beamforming directions of sensing and communication usually differ. For maximizing communication SNR, the direction of the communication-oriented beamforming should be 0.38π. However, in this case, sensing performance is seriously affected, i.e., sensing SNR almost being zero. This phenomenon indicates the poor sensing performance of the existing work (i.e., SenCom [17]) since it is limited by the communication-oriented beamforming. Meanwhile, this result highlights the challenge of balancing the requirements of both functionalities. Additionally, when there are multiple users, varying user locations necessitate different beamforming directions, resulting in diverse sensing SNR values and inconsistent CSI samples at the sensing receiver. In the next section, we will introduce VersaBeam to address these challenges.

III. The Design of VersaBeam

In this section, we first give the design overview of VersaBeam and then introduce the details.

A. Overview

As shown in Figure 5, VersaBeam aims at actively reusing communication packets for sensing with the proposed ISAC-oriented beamforming. Its operation is composed of two main phases: the preparation phase and the deployment phase.

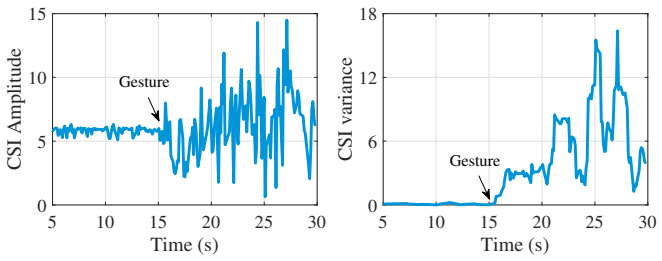
In the preparation phase, VersaBeam establishes a deep neural model for gesture recognition. Specifically, it first collects a sensing dataset that contains CSI samples from beacon packets and data packets under controlling the beamforming direction towards the sensing target. The CSI samples from beacon packets in the dataset are used for training a detection model in the deployment phase. Meanwhile, the CSI samples from data packets in the dataset are used for training a deep neural model for gesture recognition.

In the deployment phase, VersaBeam performs gesture recognition by measuring CSI from the communication packets. Specifically, VersaBeam first measures the CSI between the Wi-Fi AP and itself by sniffing the beacon packets without beamforming. Using these CSI samples and the pre-trained detection model in the preparation phase, VersaBeam determines whether there is a sensing target. When there is a sensing target, VersaBeam designs ISAC-oriented beamforming jointly considering the sniffed genuine communication-oriented beamforming feedback and the direction of the sensing target, and then appropriate users are selected as the source for packet reuse. Afterward, the designed beamforming is used to control the AP's behavior. Under the ISAC-oriented beamforming, CSI is measured from communication packets. VersaBeam unifies the CSI samples from different users' packets, and these CSI samples are further fed into the pre-trained deep neural model for gesture recognition.

B. Beacon-Based Detection

Since there may be no sensing target in the scenario, e.g., no human or the human being in the static state, continuously modifying existing communication-oriented beamforming into ISAC-oriented one would bring a high overhead. To avoid it, we aim to perform a motion detection first. Meanwhile, considering the potential negative effects of beamforming on sensing, we propose to use the beacon packets for detection. Beacon packets are utilized for periodically broadcasting the existence of the Wi-Fi AP without beamforming. Thus, the CSI measured from beacon packets is not affected by beamforming. The transmission frequency of beacon packets is low, around 10 Hz, they cannot support fine-granularity sensing applications, but they can be used for coarse-grained binary classification, i.e., detection.

Figure 6a presents the variation of CSI amplitude over time, and someone is making a specific gesture starting at 15 s. It can be observed that there is a variation of CSI without gesture, caused by the noise. With gesture, the variation of CSI over time becomes more obvious due to the dynamic path related to the sensing target. Thus, it is viable to detect the existence of the sensing target using the variance of CSI from beacon packets. Specifically, we first eliminate the effect of the LOS



(a) Amplitude. (b) Variance.

Fig. 6: CSI measured from beacon packets: (a) CSI amplitude over time and (b) CSI variance within a sliding window.

path and other NLOS paths by calculating the mean value of $\boldsymbol{H}^{\mathrm{S}}(t)$ over a sliding window t^{w} and subsequently subtracting this mean from $\boldsymbol{H}^{\mathrm{S}}(t)$ itself, as

$$\bar{\boldsymbol{H}}^{\mathrm{S}}(t) = \boldsymbol{H}^{\mathrm{S}}(t) - \underset{\text{over } t^{\mathrm{w}}}{\mathrm{Mean}}\{\boldsymbol{H}^{\mathrm{S}}(t)\}. \tag{7}$$

After that, the variance of $[\bar{\boldsymbol{H}}^{\mathrm{S}}(t)]_{m^{\mathrm{T}},m^{\mathrm{S}}}$ between the m^{T}-th transmit antenna and m^{S}-th receiving antenna is calculated over the sliding window, denoted by $\varpi^{\mathrm{S}}_{m^{\mathrm{T}},m^{\mathrm{S}}}(t)$. Figure 6b plots the variation of $\varpi^{\mathrm{S}}_{m^{\mathrm{T}},m^{\mathrm{S}}}(t)$ over time and we can observe that there is noticeable rise after motion. To further avoid false alarms, support vector machine (SVM) [25] is adopted with $\{\varpi^{\mathrm{S}}_{m^{\mathrm{T}},m^{\mathrm{S}}}, \forall m^{\mathrm{T}}, m^{\mathrm{S}}\}$ over all subcarriers being the input, since it is powerful enough to support binary classification with a low computational complexity.

C. ISAC-Oriented Beamforming

After detecting a motion, VersaBeam needs to reuse the communication packets for gesture recognition. According to the analysis in Section II-C, the existing communication-oriented beamforming may heavily compromise sensing performance. Thus, to solve this problem, we aim to propose a novel ISAC-oriented beamforming design method, and the beamforming matrix is denoted by $\boldsymbol{C}$. Let us first consider a simple case where the communication user is only equipped with one antenna.[4] The extension to the multi-antenna case will be presented based on this simple one.

For the single-antenna user, the transmit data at the AP is s. After precoding using beamforming matrix $\boldsymbol{c} \in \mathbb{C}^{M^{\mathrm{T}}\times 1}$ and undergoing wireless channel $\boldsymbol{h} \in \mathbb{C}^{1\times M^{\mathrm{T}}}$, the received signal can be given as

$$y = \boldsymbol{h}\boldsymbol{c}s + n, \tag{8}$$

where n is the noise with σ^2 variance. The resulting SNR can be expressed as

$$\eta^{\mathrm{C}}(\boldsymbol{c}) = \frac{p|\boldsymbol{h}\boldsymbol{c}|^2}{\sigma^2} \overset{\text{(a)}}{=} \frac{p|\rho|^2|\boldsymbol{v}^H\boldsymbol{c}|^2}{\sigma^2}, \tag{9}$$

where (a) holds since the SVD of $\boldsymbol{h}$ can be simplified as $\boldsymbol{h} = \rho\boldsymbol{v}^H$ for the single-antenna user. Meanwhile, using the genuine beamforming feedback $\boldsymbol{v}$, the maximal SNR is

$$\eta^{\mathrm{C}}(\boldsymbol{v}) = \frac{p|\rho|^2}{\sigma^2}. \tag{10}$$

According to Shannon's Theorem [26], the data rate should continuously increase with the SNR following a logarithmic function, i.e., $\log_2(1+\eta^{\mathrm{C}}(\boldsymbol{c}))$, theoretically. Actually, in practical Wi-Fi systems, as shown in Figure 7, the function from the SNR to the data rate follows a piecewise function and is not continuous, due to the finite MCS indices that define how many useful bits can be transmitted per unit time.[5] From the figure, we can observe that when the SNR is between the required SNRs of two sequential MCS indices, the data rate keeps a constant.[6] Motivated by this, it is realizable to adjust the beamforming with almost no communication performance loss by limiting the communication SNR loss. Let δ^{C} denote the SNR loss upper limit, that is

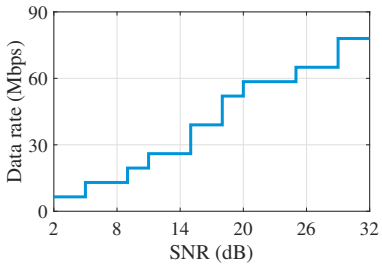


Fig. 7: The relationship between the data rate and SNR.

$$\eta^{\mathrm{C}}(\boldsymbol{v})/\eta^{\mathrm{C}}(\boldsymbol{c}) = 1/|\boldsymbol{v}^H\boldsymbol{c}|^2 \le \delta^{\mathrm{C}}. \tag{11}$$

δ^{C} is determined by the maximal SNR $\eta^{\mathrm{C}}(\boldsymbol{v})$. Specifically, we first find the MCS index according to the table of MCS and SNR, and then obtain δ^{C} by subtracting the required SNR threshold for the MCS index from the maximal SNR. For example, δ^{C} should be 4 dB when the maximal SNR is 24 dB since the required SNR threshold for MCS index 6 is 20 dB. However, VersaBeam cannot directly obtain the maximal SNR since VersaBeam and the user are generally at different locations. Fortunately, the beamforming feedback also contains the average SNR information that represents the SNR with beamforming, and thus we can directly use it for determining the upper limit of SNR loss, i.e., δ^{C}.

For the sensing performance, we have derived the expression, i.e., Eqn. (6), for sensing SNR in Section II-B. Since we aim to design an ISAC-oriented beamforming for improving sensing performance without compromising communication performance, the optimization problem can be formulated as

$$\max_{\boldsymbol{c}} \quad \eta^{\mathrm{S}}(\boldsymbol{c}) = \frac{\mathbb{E}\{||g^{\mathrm{S,D}}(t)||^2\}||(\boldsymbol{a}^{\mathrm{S,T}})^H\boldsymbol{c}||^2}{\xi||g^{\mathrm{S,L}}(\boldsymbol{a}^{\mathrm{L,T}})^H\boldsymbol{c}||^2 + b}, \tag{12a}$$

$$\text{s.t.} \quad \boldsymbol{c}^H\boldsymbol{c} = 1, \tag{12b}$$

$$|\boldsymbol{v}^H\boldsymbol{c}|^2 \ge 1/\delta^{\mathrm{C}}. \tag{12c}$$

In the above, the first constraint requires that $\boldsymbol{c}$ should be unitary for compatibility with commodity Wi-Fi systems since the existing communication-oriented beamforming matrix is obtained from SVD in Eqn. (1). The second constraint ensures no communication loss. This optimization problem can be solved using fractional programming [27]. In practice, to improve the sensing performance, the sensing receiver is placed near the sensing area and sensing target [13], [22]. For example, the sensing receiver could be the mobile phone of the sensing target. Thus, the AOD of the LOS path and dynamic path is almost the same, which means $\boldsymbol{a}^{\mathrm{S,T}} \approx \boldsymbol{a}^{\mathrm{L,T}}$, and the objective function is almost equivalent to $||(\boldsymbol{a}^{\mathrm{S,T}})^H\boldsymbol{c}||^2$. With

[4]The proposed design in this section is universal for all users, and thus we omit the subscript i.

[5]The table of MCS index and its required SNR can be found in the website: https://wlanprofessionals.com/mcs-table-and-how-to-use-it/.

[6]When the SNR is above the required SNR, the bit error rate is almost zero, and thus we do not discuss it in the paper.

the help of optimization theory [28], the closed-form optimal solution to Problem (12) can be derived as

$$\boldsymbol{c}^{\star}=\begin{cases}\boldsymbol{a}^{\mathrm{S,T}}/\sqrt{M^{\mathrm{T}}}, & \text{if } |\boldsymbol{v}^H\boldsymbol{a}^{\mathrm{S,T}}|^2\geq M^{\mathrm{T}}/\delta^{\mathrm{C}},\\ \left(\boldsymbol{v}+\sqrt{\delta^{\mathrm{C}}-1}\boldsymbol{v}^{\perp}\right)/\sqrt{\delta^{\mathrm{C}}}, & \text{otherwise,}\end{cases} \tag{13}$$

where $\boldsymbol{v}^{\perp}=\dfrac{\boldsymbol{a}^{\mathrm{S,T}}-\boldsymbol{v}^H\boldsymbol{a}^{\mathrm{S,T}}\boldsymbol{v}}{||\boldsymbol{a}^{\mathrm{S,T}}-\boldsymbol{v}^H\boldsymbol{a}^{\mathrm{S,T}}\boldsymbol{v}||}$ is orthogonal to $\boldsymbol{v}$. From the above solution, we can find that when the communication-oriented beamforming direction $\boldsymbol{v}$ is close to sensing direction $\boldsymbol{a}^{\mathrm{S,T}}$, i.e., $|\boldsymbol{v}^H\boldsymbol{a}^{\mathrm{S,T}}|^2\geq M^{\mathrm{T}}/\delta^{\mathrm{C}}$, we can directly use $\boldsymbol{a}^{\mathrm{S,T}}$ as the beamforming matrix; otherwise, we need to fine-tune $\boldsymbol{v}$ for enhancing the sensing performance with ensuring the communication performance.

For the multi-antenna user, the maximal SNR $\eta^{\mathrm{C}}(\boldsymbol{V})$ achieved by the genuine beamforming matrix can also be obtained from the beamforming feedback by averaging the average SNRs over all streams. δ^{C} can be determined using the same method for the single-antenna case. For the ISAC-oriented beamforming feedback, the solution in Eqn. (13) can be extended to the two-antenna case.[7] Specifically, since the transmit power for each stream is the same, the genuine beamforming matrix $\boldsymbol{V}=[\boldsymbol{v}_1,\boldsymbol{v}_2]\in\mathbb{C}^{M^{\mathrm{T}}\times 2}$ is equivalent to $\boldsymbol{V}'=[\boldsymbol{v}_1',\boldsymbol{v}_2']$ from the perspective of the communication performance, where

$$\boldsymbol{v}_1'=\frac{\boldsymbol{v}_1^H\boldsymbol{a}^{\mathrm{S,T}}\boldsymbol{v}_1+\boldsymbol{v}_2^H\boldsymbol{a}^{\mathrm{S,T}}\boldsymbol{v}_2}{\sqrt{|\boldsymbol{v}_1^H\boldsymbol{a}^{\mathrm{S,T}}|^2+|\boldsymbol{v}_2^H\boldsymbol{a}^{\mathrm{S,T}}|^2}}, \tag{14}$$

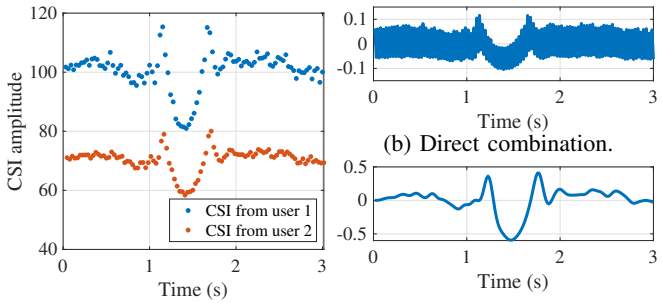
$$\boldsymbol{v}_2'=\frac{(\boldsymbol{a}^{\mathrm{S,T}})^H\boldsymbol{v}_2\boldsymbol{v}_1-(\boldsymbol{a}^{\mathrm{S,T}})^H\boldsymbol{v}_1\boldsymbol{v}_2}{\sqrt{|\boldsymbol{v}_1^H\boldsymbol{a}^{\mathrm{S,T}}|^2+|\boldsymbol{v}_2^H\boldsymbol{a}^{\mathrm{S,T}}|^2}}. \tag{15}$$

Since the $\boldsymbol{v}_1'$ and $\boldsymbol{v}_2'$ are the basis vectors in the plane spanned by $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, the beamforming direction is not influenced and the communication performance remains the same. Moreover, $\boldsymbol{v}_1'$ is structured to maximize the dot product with $\boldsymbol{a}^{\mathrm{S,T}}$ in the plane spanned by $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, and $\boldsymbol{v}_2'$ is orthogonal to $\boldsymbol{v}_1'$ and $\boldsymbol{a}^{\mathrm{S,T}}$. Thus, we only modify $\boldsymbol{v}_1'$ since the gain brought by modifying $\boldsymbol{v}_2'$ is meagre. The solution in Eqn. (13) can be applied by substituting $\boldsymbol{v}_1'$ for $\boldsymbol{v}$, and the ISAC-oriented beamforming matrix is $\boldsymbol{C}=[\boldsymbol{c}^{\star},\boldsymbol{v}_2']$. Note that $\boldsymbol{C}$ is unitary since $\boldsymbol{c}^{\star}$ is orthogonal to $\boldsymbol{v}_2'$. Accordingly, we only use the CSI under $\boldsymbol{c}^{\star}$, i.e., the first column of $\boldsymbol{H}^{\mathrm{S,B}}(\boldsymbol{C},t)$, for sensing.

D. User Selection

Sensing applications, such as gesture recognition, have CSI sampling rate requirements to avoid the missing of the key sensing information. However, only sniffing communication packets transmitted from the AP to one user may not be enough for supporting sensing applications since the packet arrival rate of one user may be lower than the requirement of CSI sampling rate. To address this issue, we can sniff packets transmitted to all active users. Let λ_i denote the communication packet arrival rate of the i-th user and λ^{S} denote the required CSI sampling rate. We aim to pick enough packets for sensing. Due to the different users' locations, the ISAC-oriented beamforming matrix designed in Sec. III-C is not the same, leading to varied sensing SNRs. The sensing SNR of the measured CSI at the sensing receiver is denoted by $\eta_i^{\mathrm{S},\star}$ when sniffing the communication packets of the i-th user. To satisfy the two requirements of the sensing, i.e., high sensing SNR and enough CSI samples, we first need to sort users by sensing SNR from highest to lowest. With the newly defined index, we pick out users sequentially from 1 to I until the CSI sampling rate requirement is met or the communications packets from all users have been utilized. Here, we do not directly use all CSI samples for sensing since a CSI with low sensing SNR may not be conducive to improving sensing performance.

[7]Here, we focus on the single-antenna and two-antenna cases since they are the most common configuration in practical. For other cases, the fractional programming mentioned before can be easily applied.



(a) CSI from different users. (c) Unified CSI.

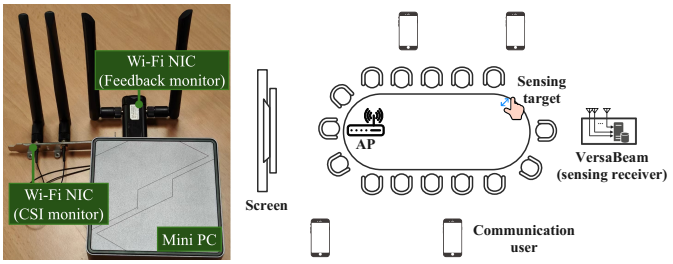
Fig. 8: Result of the proposed CSI unifying method: (a) CSI measured from different users' packets, (b) CSI after direct combination, and (c) CSI after unifying.

E. CSI Unifying

Although using the CSI measured from different users' packets can ensure enough CSI samples, the sensing SNRs of CSI samples are different, leading to a discontinuous CSI variation. Let $\boldsymbol{h}_i^{\mathrm{S,B}}(t)$ denote the CSI measured from i-th user's packets under the designed beamforming feedback. As shown in Figure 8a, the variations of $\boldsymbol{h}_1^{\mathrm{S,B}}(t)$ and $\boldsymbol{h}_2^{\mathrm{S,B}}(t)$ are different, reflected in the average and deviation. The former is related to the LOS path and other NLOS paths and the beamforming matrix, and the latter is caused by the different beamforming matrices across different users. When the CSI samples measured from two different users' packets are directly combined in chronological order, the mixed CSI amplitude is presented in Figure 8b after normalization. We can observe that the mixed CSI in Figure 8b varies frequently within the first second but the CSI measured from only one user in Figure 8a remains almost unchanged. Consequently, directly combining CSI samples from different users would lead to the wrong sensing result.

To address the above-mentioned problem, we propose a CSI unifying method. First of all, we aim to eliminate the difference in average by removing the LOS path and other NLOS paths from the measured CSI. Since they are static, we can calculate the mean value over time and subsequently subtract this mean from $\boldsymbol{h}_i^{\mathrm{S,B}}(t)$ packet, as

$$\bar{\boldsymbol{h}}_i^{\mathrm{S,B}}(t)=\boldsymbol{h}_i^{\mathrm{S,B}}(t)-\underset{\text{over } t}{\mathrm{Mean}}\{\boldsymbol{h}_i^{\mathrm{S,B}}(t)\}. \tag{16}$$



(a) Prototype. (b) Experiment layout.

Fig. 9: VersaBeam implementation: (a) hardware configurations and (b) experiment layout in a conference area.

$\bar{\boldsymbol{h}}_i^{\mathrm{S,B}}(t)$ is equivalent to $g^{\mathrm{S,D}}(t)\boldsymbol{a}^{\mathrm{S,R}}(\boldsymbol{a}^{\mathrm{S,T}})^H\boldsymbol{c}_i^\star$ when we ignore noise according to the expression of $\boldsymbol{H}^{\mathrm{S,B}}(\boldsymbol{C},t)$. Thus, the difference in deviation is caused by the product of $\boldsymbol{c}_i^\star$ and $\boldsymbol{a}^{\mathrm{S,T}}$. Fortunately, since both $\boldsymbol{c}_i^\star$ and $\boldsymbol{a}^{\mathrm{S,T}}$ are known at VersaBeam, the difference in deviation can be directly eliminated by

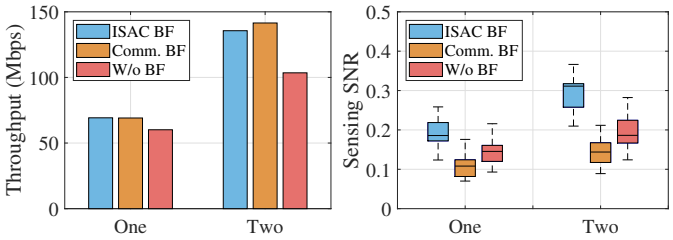
$$\hat{\boldsymbol{g}}_i^{\mathrm{S,D}}(t) = \bar{\boldsymbol{h}}_i^{\mathrm{S,B}}(t) / \left((\boldsymbol{a}^{\mathrm{S,T}})^H \boldsymbol{c}_i^\star\right). \tag{17}$$

$\hat{\boldsymbol{g}}_i^{\mathrm{S,D}}(t)$ from all selected users can be combined into one CSI sequence in chronological order. Since the CSI samples are measured from communications packets, they are not evenly distributed over time, and we apply the fitting-resampling method to maintain the uniformity of CSI required by most of sensing applications. Furthermore, we further apply the low-pass filter and normalization on the CSI sequence to suppress the noise. The filtered signal is presented in Figure 8c, and we can observe that it has a similar variation trace as $\boldsymbol{h}_i^{\mathrm{S,B}}$, which proves the effectiveness of the proposed unifying method. Moreover, the filtered signals over different receiving antennas and subcarriers are input into the deep neural model for gesture recognition.

IV. Implementation

In this section, we elaborate on the implementation of VersaBeam, as well as introduce the experiment setup.

Implementation: We use a mini PC with an R7-7840H AMD CPU and 16 GB RAM and two Wi-Fi NICs (Alfa AWUS036ACM [29] and Intel AX200 [30]) to form VersaBeam, as shown in Figure 9a. Both NICs work in monitor mode. Intel AX200 is used for measuring CSI from the received packets with PicoScenes [31], acting as the sensing receiver. To control the beamforming behavior of the Wi-Fi AP, we adopt the beamforming feedback forgery method proposed in [19]. Specifically, we first sniff the genuine beamforming feedback using Alfa AWUS036ACM with libpcap 1.10.3, then construct ISAC-oriented beamforming feedback following the method proposed in Section III-C using the mini PC with Eigen 3.3.7, and the constructed feedback is injected into the Wi-Fi systems using Alfa AWUS036ACM with libpcap 1.10.3. Moreover, the mini PC is also used for performing the beacon-based detection with Python and realizing gesture recognition with Python. Specifically, we adopt the ResNet as the neural network for gesture recognition and it is built upon PyTorch.



(a) Communication. (b) Sensing.

Fig. 10: The (a) communication and (b) sensing performance comparison with the single-antenna communication user and two-antenna communication user.

Experiment Setup: We first conduct micro-benchmark experiments in terms of the sensing performance (i.e., sensing SNR) and communication performance (i.e., throughput) in a video conference room and then conduct a real-world experiment of gesture recognition in three different environments, i.e., a conference room, an office room, and a dance room. Due to the page limit, we mainly introduce the setup in the conference room, with the other two rooms having similar setups. The layout of the conference room is shown in Figure 9b. The AP is Xiaomi Redmi Router AC2100 [32] with four antennas and operates at 5 GHz with 20 MHz bandwidth under the Wi-Fi 5 standard. There are four communication users randomly located in the room and the types of communication users include Realtek RTL8821CU with one antenna and Realtek RTL8812BU with two antennas. Iperf3 [33] is adopted for measuring the throughput between the AP and users. To show the average performance, the total test time for throughput is more than 8 hours. Meanwhile, 8 individuals are invited to perform six different gestures, namely circle, front-back, slide, star, wave, and zig-zag. The overall number of collected CSI samples in the conference room is 8,700.

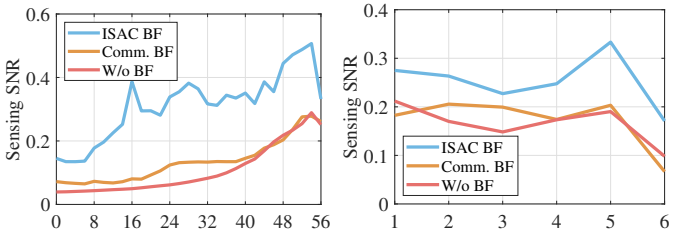
To show the superiority of the proposal, we consider two baselines: one is the scheme without beamforming (BF), and the other is the scheme with communication-oriented beamforming, i.e., SenCom [17].

V. Evaluations

We first present the experiment results focusing on throughput and sensing SNR, followed by real-world experiment results of gesture recognition to further verify the sensing performance of VersaBeam.

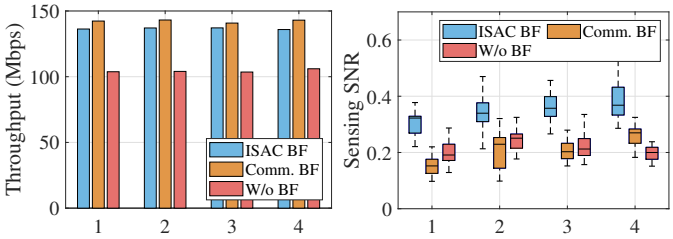
A. Communication and Sensing Performance

First, we present the communication and sensing performance with the single-antenna communication user and two-antenna communication user, as shown in Figure 10. Regarding the communication performance, compared to communication-oriented beamforming, the throughput loss with proposed beamforming is 4.1 % for the two-antenna user and almost 0 % for the one-antenna user, verifying the low impact on communication. The loss is caused by the inaccurate estimation of communication SNR and channel variation, and the estimation error and variation increase with the number of antennas. Thus, the throughput loss for the two-antenna user



(a) Subcarrier. (b) Gesture.

Fig. 11: The comparison of sensing SNR under different (a) subcarriers and (b) gestures.



(a) Communication. (b) Sensing.

Fig. 12: The impact of location on (a) communication performance and (b) sensing performance.

is higher than that for the one-antenna user. Moreover, the throughput achieved by the proposal is higher than that without beamforming. For the sensing performance, the proposed ISAC beamforming achieves the highest sensing SNR among the three schemes, further demonstrating the superiority of the proposed ISAC-oriented beamforming. The sensing SNR of the two-antenna user is higher than that of the single-antenna user. This can be explained as follows: the beamforming feedback is a matrix for a two-antenna communication user and a vector for a one-antenna communication user. Thus, with more antennas of the communication user, there are more degrees of spatial freedom for designing the beamforming feedback, which is beneficial for achieving higher sensing SNR. Additionally, the sensing SNR using the communication-oriented beamforming is lower than that without beamforming, highlighting the limitation of SenCom.

To fully explore the sensing performance of the proposed ISAC-oriented beamforming, we also plot the sensing SNRs over the 57 subcarriers (including all subcarriers available for sensing with bandwidth being 20 MHz) and 6 gestures in Figure 11. Although the sensing SNR varies with the subcarrier for the three schemes due to the multipath effect, the sensing SNR of the proposed remains the highest one since our proposed method designs the beamforming feedback for each subcarrier. Notably, the overall computation latency, including decompression, feedback matrix construction, and compression, is less than 1 ms, allowing real-time control of the AP's beamforming behavior. This latency could be further reduced via parallel computing. Meanwhile, the sensing SNR of the proposed ISAC-oriented beamforming is higher than the other two schemes across different gestures, verifying the universality of VersaBeam. This is because VersaBeam increases the transmit power towards the sensing target, making the variation of the sensing target more pronounced.

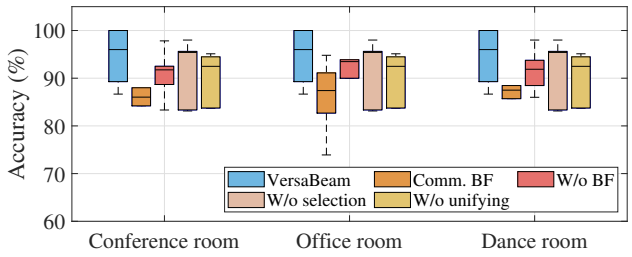


Fig. 13: The accuracy of the gesture recognition in three different environments.

Due to that the proposed ISAC-oriented beamforming jointly considers directions of communication and sensing, i.e., $\boldsymbol{V}$ and $\boldsymbol{a}^{\mathrm{S,T}}$, the relative location between the sensing receiver and the communication user would influence the performance. To explore the impact, we vary the location of the communication user while fixing the location of the sensing receiver, and the result is shown in Figure 12. Note that the correlation between $\boldsymbol{V}$ and $\boldsymbol{a}^{\mathrm{S,T}}$ for four locations is 0.36, 0.43, 0.46, and 0.64, respectively. For the throughput in Figure 12a, it can be seen that the throughput of the ISAC-oriented beamforming approaches that of the communication-oriented beamforming under different locations. For the sensing SNR in Figure 12b, the ISAC-oriented beamforming achieves the highest sensing SNR. The sensing SNRs of both the ISAC-oriented and communication-oriented beamforming increase with the correlation. This result arises because the high correlation means a similar beamforming direction for the communication and sensing, and the transmit power used for sensing the target increases with correlation. These findings indicate that VersaBeam can realize high sensing performance with only low loss of communication performance.

B. Gesture Recognition

To verify the effectiveness of the proposed user selection strategy and CSI unifying method, we add two baselines: one is the scheme without user selection strategy, which randomly selects communication users for reusing, and the other is the scheme without CSI unifying, where the CSI samples measured from different users' packets are directly combined as introduced in Section III-E.

The accuracy of the gesture recognition in conference room, office room, and dance room is illustrated in Figure 13. The accuracy of VersaBeam is the highest among the three schemes. Specifically, the sensing SNR realized by VersaBeam is higher than these of the schemes without beamforming and with communication-oriented beamforming according to the experiment results in Section V-A and is also higher than the scheme without user selection since the packets with high sensing SNR are selected. Consequently, the accuracy of VersaBeam is higher as the high sensing SNR is beneficial to help the neural network extract features of each gesture and recognize different gestures. Moreover, without unifying CSI from different users' packets, the variation of the CSI over time caused by the gesture may be covered by the variation

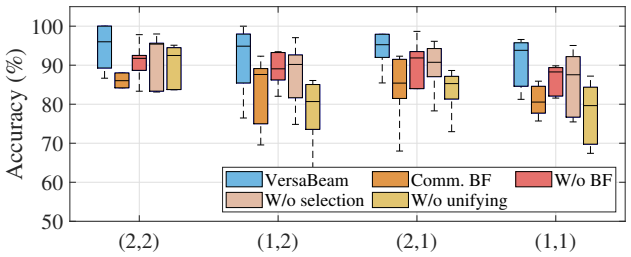


Fig. 14: The accuracy of the gesture recognition with different numbers of antennas, where (a, b) means that there are a (b) antennas at the communication (sensing) user.

caused by the different beamforming matrices for different users, and thus, the accuracy of this scheme is lower than that of VersaBeam. Besides, the results in three environments verify that VersaBeam can achieve outstanding performance in all scenarios.

To explore the influence of antennas' number at the communication user and sensing receiver, the accuracy under different antenna settings is plotted in Figure 14. It can be seen that the accuracy increases with the number of antennas at the communication user. The reason is that the antenna's number of the communication user determines the degrees of spatial freedom for designing the beamforming feedback, and the sensing SNR can be higher with more degrees of spatial freedom. Meanwhile, the accuracy is positively related to the number of antennas at the sensing receiver. With more antennas at the sensing receiver, spatial resolution of the receiver can be improved and more sensing information can be provided, beneficial for extracting features of different gestures. Moreover, even with one antenna at the communication user and sensing receiver, VersaBeam still can achieve high accuracy, further demonstrating the superiority of VersaBeam.

VI. Related Works

Our work is closely related to Wi-Fi sensing and beamforming design.

Wi-Fi sensing: Commodity Wi-Fi devices have been widely used to realize contactless human sensing due to their non-intrusive, device-free, light-needless, and pervasive advantages. The applications of Wi-Fi sensing including localization [14], [34], activity recognition [5], [35], person identification [36], [37], vital signs monitoring [38], and so on. To ensure high sensing performance, those works generally employ the transmitter to send exclusive sensing packets (that do not carry information) without beamforming, and the CSI extracted from these packets at the receiver with the help of CSI tools [9], [31], [39] is used for completing specific sensing task. Due to the usage of the exclusive sensing packets, the communication and sensing in the Wi-Fi system are realized via a time-division method, being low efficiency. To improve the efficiency, one solution is to extract CSI from communication packets (may transmitted via beamforming) without extra overhead for sensing. To avoid the impact of communication-oriented beamforming, [13], [40], [41] utilize the beamforming feedback information (BFI) transmitted in clear text to perform sensing instead of CSI. However, the sampling rate of BFI is 10$\sim$20 Hz, which may not be enough for fine-granularity sensing. To support fine-granularity sensing, SenCom [17] proposes a signal processing method to weaken the effect of communication-oriented beamforming. However, since the noise and interference are mixed with the useful signal, the signal processing method cannot fully eliminate the effect. To solve this issue, we propose VersaBeam that utilizes active ISAC-oriented beamforming for high-performance sensing and communication over commodity Wi-Fi systems.

Beamforming design: Different from the traditional beamforming design in communication systems that only aims to maximize the throughput of all active users, there are two objective functions, i.e., maximizing sensing and communication performance, in the beamforming design towards ISAC [11], [42]–[46]. For example, [42] uses signal-to-interference-plus-noise ratio (SINR) as the metric for sensing performance and maximizes it under the communication SINR requirement and total energy limitation. On the contrary, [43] proposes a beamforming design algorithm to maximize the communication SINR under the requirement of sensing SINR. Moreover, the trade-off between sensing and communication performance via weighting is studied in [44]. However, these works stay only on emulation platforms with limited real-world applicability. Worse, existing works do not consider the characteristic (i.e., the relationship between SNR and MCS) and requirement (i.e., beamforming matrix should be unitary one) of commodity Wi-Fi devices, being impractical. To address those problems, we propose VersaBeam that can directly be used in the existing Wi-Fi system without modification to realize ISAC.

VII. Conclusion

In this paper, we have introduced VersaBeam to explore the potential of the beamforming protocols to achieve high-performance ISAC in practical Wi-Fi systems. VersaBeam reuses the communication packets for sensing without overhead. To address the low sensing SNR caused by the existing communication-oriented beamforming, we propose an ISAC-oriented beamforming design based on the mathematical analysis of the beamforming on sensing SNR. Considering the different beamforming behaviors for different users, we also propose a user selection strategy and a CSI unifying method for improving sensing SNR of measured CSI and further eliminating the difference in CSI caused by different beamforming behaviors. The extensive evaluations with commodity Wi-Fi devices have evidently demonstrated that VersaBeam can realize high-performance sensing with minimal impact on communication performance.

Acknowledgement

This research is supported in part by MOE Tier 1 grant RG16/22, National Science Fund for Distinguished Young Scholars of China under grant No.62125203, and The Major Scientific Instruments and Equipments Development Project of National Natural Science Foundation of China under grant No.62427809.

REFERENCES

[1] B. Bellalta, L. Bononi, R. Bruno, and A. Kassler, "Next Generation IEEE 802.11 Wireless Local Area Networks: Current status, Future directions and Open Challenges," *Comput. Commun.*, vol. 75, pp. 1–25, Feb. 2016.

[2] E. Perahia and M. X. Gong, "Gigabit Wireless LANs: An Overview of IEEE 802.11ac and 802.11ad," *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 15, no. 3, pp. 23–33, Jul. 2011.

[3] S. Ji, Y. Xie, and M. Li, "SiFall: Practical Online Fall Detection with RF Sensing," in *Proc. ACM SenSys*, Nov. 2022, pp. 563–577.

[4] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, and S. Li, "RT-Fall: A Real-Time and Contactless Fall Detection System with Commodity WiFi Devices," *IEEE Trans. Mobile Comput.*, vol. 16, no. 2, pp. 511–526, Feb. 2016.

[5] F. Wang, S. Zhou, S. Panev, J. Han, and D. Huang, "Person-in-WiFi: Fine-grained Person Perception using WiFi," in *Proc. IEEE/CVF ICCV*, Oct. 2019, pp. 5452–5461.

[6] S. Ding, Z. Chen, T. Zheng, and J. Luo, "RF-Net: A Unified Meta-Learning Framework for RF-enabled One-Shot Human Activity Recognition," in *Proc. ACM SenSys*, Nov. 2020, pp. 517–530.

[7] Y. He, G. Yu, Y. Cai, and H. Luo, "Integrated Sensing, Computation, and Communication: System Framework and Performance Optimization," *IEEE Trans. Wireless Commun.*, vol. 23, no. 2, pp. 1114–1128, Feb. 2023.

[8] X. Liu, J. Cao, S. Tang, J. Wen, and P. Guo, "Contactless Respiration Monitoring Via Off-the-Shelf WiFi Devices," *IEEE Trans. Mobile Comput.*, vol. 15, no. 10, pp. 2466–2479, Oct. 2015.

[9] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool Release: Gathering 802.11n Traces with Channel State Information," *ACM SIGCOMM CCR*, vol. 41, no. 1, pp. 53–53, Jan. 2011.

[10] Z. Chen, T. Zheng, C. Hu, H. Cao, Y. Yang, H. Jiang, and J. Luo, "ISACoT: Integrating Sensing with Data Traffic for Ubiquitous IoT Devices," *IEEE Commun. Mag.*, vol. 61, no. 5, pp. 98–104, May 2022.

[11] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated Sensing and Communications: Toward Dual-Functional Wireless Networks for 6G and Beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, Jun. 2022.

[12] F. Meneghello, C. Chen, C. Cordeiro, and F. Restuccia, "Toward Integrated Sensing and Communications in IEEE 802.11bf Wi-Fi Networks," *IEEE Commun. Mag.*, vol. 61, no. 7, pp. 128–133, Jul. 2023.

[13] J. Hu, T. Zheng, Z. Chen, H. Wang, and J. Luo, "MUSE-Fi: Contactless MUti-person SEnsing Exploiting Near-field Wi-Fi Channel Variation," in *Proc. ACM MobiCom*, Oct. 2023, pp. 1–15.

[14] X. Li, H. Wang, Z. Chen, Z. Jiang, and J. Luo, "UWB-Fi: Pushing Wi-Fi towards Ultra-wideband for Fine-Granularity Sensing," in *Proc. ACM MobiSys*, Jun. 2024, pp. 42–55.

[15] D. Zhang, D. Wu, K. Niu, X. Wang, F. Zhang, J. Yao, D. Jiang, and F. Qin, "Practical Issues and Challenges in CSI-based Integrated Sensing and Communication," in *Proc. IEEE ICC Workshops*, May 2022, pp. 836–841.

[16] E. Perahia and R. Stacey, *Next Generation Wireless LANs: 802.11 n and 802.11 ac*. Cambridge, U.K.: Cambridge university press, 2013.

[17] Y. He, J. Liu, M. Li, G. Yu, J. Han, and K. Ren, "SenCom: Integrated Sensing and Communication with Practical WiFi," in *Proc. ACM MobiCom*, Oct. 2023, pp. 1–16.

[18] Y. He, J. Liu, M. Li, G. Yu, and J. Han, "Forward-Compatible Integrated Sensing and Communication for WiFi," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 9, pp. 2440–2456, Sep. 2024.

[19] M. Xu, Y. He, X. Li, J. Hu, Z. Chen, F. Xiao, and J. Luo, "Beamforming made Malicious: Manipulating Wi-Fi Traffic via Beamforming Feedback Forgery," in *Proc. ACM MobiCom*, Nov. 2024, pp. 1–15.

[20] I. S. Association *et al.*, "IEEE Std 802.11-2016, IEEE Standard for Local and Metropolitan Area Networks—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," 2016.

[21] G. Redieteab, L. Cariou, P. Christin, and J.-F. Helard, "SU/MU-MIMO in IEEE 802.11ac: PHY+MAC Performance Comparison for Single Antenna Stations," in *Proc. IEEE Wireless Telecommun. Symp.*, Apr. 2012, pp. 1–5.

[22] X. Wang, K. Niu, J. Xiong, B. Qian, Z. Yao, T. Lou, and D. Zhang, "Placement Matters: Understanding the Effects of Device Placement for WiFi Sensing," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 6, no. 1, pp. 1–25, Mar. 2022.

[23] V. Venkateswaran and A.-J. van der Veen, "Analog Beamforming in MIMO Communications with Phase Shift Networks and Online Channel Estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4131–4143, Aug. 2010.

[24] M. Kang and M.-S. Alouini, "Capacity of MIMO Rician channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 112–122, Jan. 2006.

[25] V. Jakkula, *Tutorial on Support Vector Machine (SVM)*. Washington, DC, USA: School of EECS, Washington State University, 2006.

[26] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge university press, 2005.

[27] K. Shen and W. Yu, "Fractional Programming for Communication Systems—Part I: Power Control and Beamforming," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2616–2630, May 2018.

[28] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge university press, 2004.

[29] ALFA Network, "AWUS036ACM - Long Range Dual-Band AC1200 Wireless USB 3.0 Wi-Fi Adapter," https://www.alfa.com.tw/products/awus036acm?variant=39477234597960, 2024, online; accessed: 16 July 2024.

[30] Intel Corporation, "Intel Wi-Fi 6 AX200 Gig+ Specifications," https://www.intel.com/content/www/us/en/products/sku/189347/intel-wifi-6-ax200-gig/specifications.html, 2024, online; accessed: 16 July 2024.

[31] Z. Jiang, T. H. Luan, X. Ren, D. Lv, H. Hao, J. Wang, K. Zhao, W. Xi, Y. Xu, and R. Li, "Eliminating the Barriers: Demystifying Wi-Fi Baseband Design and Introducing the PicoScenes Wi-Fi Sensing Platform," *IEEE Internet Things J.*, vol. 9, no. 6, pp. 4476–4496, Mar. 2021.

[32] Xiaomi, "Xiaomi Redmi Router AC2100," https://www.mi.com/rm2100, 2024, online; accessed: 16 July 2024.

[33] Iperf Project, "Iperf3: A TCP, UDP, and SCTP Network Bandwidth Measurement Tool," 2014, version 3. [Online]. Available: https://iperf.fr/

[34] Z. Chen, G. Zhu, S. Wang, Y. Xu, J. Xiong, J. Zhao, J. Luo, and X. Wang, "M3: Multipath Assisted Wi-Fi Localization with a Single Access Point," *IEEE Trans. Mobile Comput.*, vol. 20, no. 2, pp. 588–602, Mar. 2019.

[35] R. Xiao, J. Liu, J. Han, and K. Ren, "OneFi: One-Shot Recognition for Unseen Gesture via COTS WiFi," in *Proc. ACM SenSys*, Nov. 2021, pp. 206–219.

[36] F. Wang, J. Han, F. Lin, and K. Ren, "WiPIN: Operation-Free Passive Person Identification Using Wi-Fi Signals," in *Proc. IEEE GLOBECOM*, Dec. 2019, pp. 1–6.

[37] J. Liu, C. Xiao, K. Cui, J. Han, X. Xu, and K. Ren, "Behavior Privacy Preserving in RF Sensing," *IEEE Trans. Depend. Sec. Comput.*, vol. 20, no. 1, pp. 784–796, Jan.-Feb. 2022.

[38] S. Shi, Y. Xie, M. Li, A. X. Liu, and J. Zhao, "Synthesizing Wider WiFi Bandwidth for Respiration Rate Monitoring in Dynamic Environments," in *Proc. IEEE INFOCOM*, Apr. 2019, pp. 181–189.

[39] Y. Xie, Z. Li, and M. Li, "Precise Power Delay Profiling with Commodity WiFi," in *Proc. ACM MobiCom*, Sep. 2015, pp. 53–64.

[40] C. Wu, X. Huang, J. Huang, and G. Xing, "Enabling Ubiquitous WiFi Sensing with Beamforming Reports," in *Proc. ACM SIGCOMM*, Sep. 2023, pp. 20–32.

[41] X. Li, J. Hu, and J. Luo, "Efficient Beamforming Feedback Information-Based Wi-Fi Sensing by Feature Selection," *IEEE Wireless Commun. Lett.*, vol. 13, no. 9, pp. 2347–2351, Sep. 2024.

[42] Y. He, Y. Cai, G. Yu, and K.-K. Wong, "Joint Transceiver Design for Dual-Functional Full-Duplex Relay Aided Radar-Communication Systems," *IEEE Trans. Commun.*, vol. 70, no. 12, pp. 8355–8369, Dec. 2022.

[43] Y. He, Y. Cai, H. Mao, and G. Yu, "RIS-Assisted Communication Radar Coexistence: Joint Beamforming Design and Analysis," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2131–2145, Jul. 2022.

[44] X. Liu, T. Huang, N. Shlezinger, Y. Liu, J. Zhou, and Y. C. Eldar, "Joint Transmit Beamforming for Multiuser MIMO Communications and MIMO Radar," *IEEE Trans. Signal Process.*, vol. 68, pp. 3929–3944, 2020.

[45] H. Hua, J. Xu, and T. X. Han, "Optimal Transmit Beamforming for Integrated Sensing and Communication," *IEEE Trans. Veh. Technol.*, vol. 72, no. 8, pp. 10 588–10 603, Aug. 2023.

[46] Z. He, W. Xu, H. Shen, D. W. K. Ng, Y. C. Eldar, and X. You, "Full-Duplex Communication for ISAC: Joint Beamforming and Power Optimization," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 9, pp. 2920–2936, 2023.