Exploring Multiple Antennas for Long-range WiFi Sensing

YOUWEI ZENG, Peking University, China
JINYI LIU, Peking University, China
JIE XIONG, University of Massachusetts Amherst, USA
ZHAOPENG LIU, Peking University, China
DAN WU, Peking University, China
DAQING ZHANG, Peking University, China and Telecom SudParis, France

Despite extensive research effort on contactless WiFi sensing over the past few years, there are still significant barriers hindering its wide application. One key issue is the limited sensing range due to the intrinsic nature of employing the weak target-reflected signal for sensing and therefore the sensing range is much smaller than the communication range. In this work, we address this challenging issue, moving WiFi sensing one step closer to real-world adoption. The key idea is to effectively utilize the multiple antennas widely available on commodity WiFi access points to simultaneously strengthen the target-reflected signal and reduce the noise. Although traditional beamforming schemes can help increase the signal strength, they are designed for communication and can not be directly applied to benefit sensing. To effectively increase the WiFi sensing range using multiple antennas, we first propose a new metric that quantifies the signal sensing capability. We then propose novel signal processing methods, which lay the theoretical foundation to support beamforming-based long-range WiFi sensing. To validate the proposed idea, we develop two sensing applications: fine-grained human respiration monitoring and coarse-grained human walking tracking. Extensive experiments show that: (i) the human respiration sensing range is significantly increased from the state-of-the-art 6-8 m to 11 m;[1] and (ii) human walking can be accurately tracked even when the target is 18 m away from the WiFi transceivers, outperforming the sensing range of the state-of-the-art by 50%.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile computing systems and tools**;

Additional Key Words and Phrases: WiFi Sensing, Channel State Information (CSI), Sensing-signal-to-noise Ratio (SSNR)

---

[1]The sensing range is calculated by adding the distances from the target to the transceivers and then dividing the summation by two.

Authors' addresses: Youwei Zeng, Key Laboratory of High Confidence Software Technologies (Ministry of Education), School of Electronics Engineering and Computer Science, Peking University, Beijing, China, ywzeng@pku.edu.cn; Jinyi Liu, Key Laboratory of High Confidence Software Technologies (Ministry of Education), School of Electronics Engineering and Computer Science, Peking University, Beijing, China, liujinyi@pku.edu; Jie Xiong, College of Information and Computer Sciences, University of Massachusetts Amherst, USA, jxiong@cs.umass.edu; Zhaopeng Liu, Key Laboratory of High Confidence Software Technologies (Ministry of Education), School of Electronics Engineering and Computer Science, Peking University, Beijing, China, liuzp@pku.edu.cn; Dan Wu, Key Laboratory of High Confidence Software Technologies (Ministry of Education), School of Electronics Engineering and Computer Science, Peking University, Beijing, China, dan@pku.edu.cn; Daqing Zhang, Key Laboratory of High Confidence Software Technologies (Ministry of Education), School of Electronics Engineering and Computer Science, Peking University, Beijing, China , Institut Mines, Telecom SudParis, Evry, France, dqzhang@sei.pku.edu.cn.
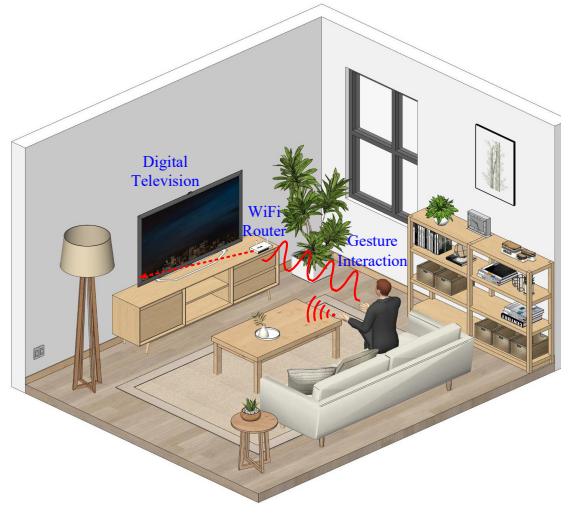

Fig. 1. In a smart home, the WiFi signals serve as a medium for not just data communication but also gesture interaction.

1 INTRODUCTION

In recent years, with the rapid development of wireless technologies, the role of WiFi signals has expanded from a sole communication medium to a non-intrusive human sensing tool. Through contact-free wireless sensing, a large variety of applications have been enabled, ranging from fine-grained respiration monitoring [11, 14, 25, 27, 33, 37–39, 45] and finger tracking [17, 32] to coarse-grained fall detection [19, 41], indoor localization/tracking [9], gait sensing [31, 34] and daily activity recognition [26]. Fig. 1 depicts a typical scene of WiFi sensing in a smart home environment. The digital TV is communicating with a WiFi router seamlessly via WiFi connections. In the meanwhile, the TV can "sense" the gestures of a human target on the sofa using the WiFi signals reflected from the human target, enabling contact-free and sensor-free interactions with the TV. The rationale behind this WiFi sensing is that the WiFi reflection signals vary with target movements. By analyzing the signal variations, we could obtain the target information such as hand gesture and respiration rate. The key advantage of this WiFi sensing is twofold. First, the target does not need to be equipped with any sensors, thus a more comfortable and friendly user experience is provided. Second, WiFi devices are widely available indoors, thus no additional infrastructure needs to be deployed.

While the idea of WiFi sensing is promising, researchers have identified a critical issue hindering the adoption of WiFi sensing in real world: the sensing range is very limited. Although the communication range of WiFi can reach up to tens of meters, the sensing range is just 6-8 m [42]. This is because while WiFi communication relies on both direct-path and reflection-path signals, WiFi sensing captures information only from the weak reflection signals and the subtle movement-induced signal variations can be easily buried in noise. In order to achieve reliable performance, target and WiFi transceivers need to be located in the same room, and the target needs to be close to the WiFi devices. However, these requirements cannot always be satisfied in real-world settings. For example, if the WiFi router is placed in the living room, it will be challenging to monitor a target sleeping in the bedroom. Therefore, the small sensing range greatly limits the real-world adoption of WiFi sensing. Recently, we have seen effort made to push the range limit of WiFi sensing. By exploiting the ratio of CSI (Channel State Information) readings between two adjacent antennas co-located at the same receiver, FarSense [39] cancels out

the common noise, pushing the respiration sensing range from 2-4 m to 6-8 m. Though promising, the achieved range is still not enough to cover a commonly seen house environment.

Different from previous work, in this paper, we propose to employ the antenna array widely available at WiFi access points to increase the sensing range. Specifically, we aim to simultaneously strengthen the weak target-reflected signal and reduce the noise. To achieve these two goals, we rely on two observations:

- The commodity WiFi access points nowadays are usually equipped with multiple antennas, each of which receives an independent copy of signal reflected from the target. If the signals received at multiple antennas are properly combined, the target-reflected signal can get strengthened.
- The CSI readings retrieved from different antennas at the same receiver share similar noises, i.e., the AGC (Automatic Gain Control) noise and time-varying phase offset. Therefore, these noises could be effectively canceled out by the signal division operation between adjacent antennas [16, 32, 39].

However, there is still a gap between our idea and a functional system. Two challenges need to be tackled before our idea can be applied to increase the sensing range:

- It is unclear how to optimally combine CSI readings collected from multiple antennas as there is no metric to quantify "how good" the CSI reading is for sensing. While traditional beamforming schemes help increase the signal strength for communication [13], we have no idea whether (and to what extent) they can help increase sensing range. Fundamentally, the performance of sensing and communication depends on different signal metrics. While signal-to-noise ratio (SNR) can well measure the quality of a communication channel, it can not be used to accurately quantify the sensing capability because the strong direct path signal does not contribute to sensing.
- Although signal division operation can reduce the common noise, the signal phase/amplitude which is critical for sensing can get distorted during this division process.

To tackle the aforementioned challenges, we first propose a metric called sensing-signal-to-noise ratio (SSNR), which is defined as the ratio of the target-reflected signal power and noise power to quantify the signal sensing capability. The signals arriving at the receiver are grouped into two categories [26], i.e., static signal and dynamic signal. The static signal consists of the direct path signal between the transmitter-receiver pair and the reflection signals from static objects such as a wall. The signal reflected from target is the dynamic signal. We show in this paper that the static signal actually has no effect on the sensing performance. To improve the WiFi sensing capability, we combine the received CSI readings from multiple antennas such that the dynamic components get constructively combined. Through rigorous mathematical derivation, we prove that the SSNR is maximized if we scale each antenna's target-reflected signal, align the signals to the same phase and then add them together. Unfortunately, we have no idea of the target-reflected signal at each antenna because it is mixed with the static signal. Moreover, the time-varying phase offset due to unsynchronization between transceivers makes it challenging to align the signal phase for optimal combination.

To address this issue, we propose a novel algorithm to accurately estimate the relative phase and magnitude between the target-reflected signals of two antennas, which can be utilized to extract the optimal weight applied to each CSI reading for combination. At last, to reduce the noise, we propose a signal processing technique by constructing a "target nulling signal" as a reference. The key here is that the "nulling signal" does not contain the dynamic signal component and therefore the division operation with respect to this "nulling signal" does not corrupt the phase/amplitude information which is critical for sensing. By dividing the optimally-combined signal with the reference signal, we can successfully eliminate the unwanted AGC noise and time-varying phase offset without corrupting the sensing information. In such a way, the weak target-reflected signal is greatly strengthened and the noise is also reduced, eventually increasing the WiFi sensing range.

To validate the proposed methods, we implement two sensing applications (i.e., fine-grained respiration monitoring and coarse-grained human tracking) on commodity WiFi devices and conduct extensive real-world

experiments. The range of human respiration sensing is significantly increased from the state-of-the-art 6-8 m to 11 m. Even with a 10 cm thick brick wall between the target and transceivers, the sensing range is still as large as 6 m. This is the first true through-wall respiration sensing with both transmitter/receiver in another room using commodity WiFi hardware. The state-of-the-art FarSense system [39] can only support respiration sensing with one of the transmitter/receiver placed in another room. For coarse-grained human walking tracking, accurate displacement and direction information can be obtained even when the target is 18 m away from the WiFi transceivers.

The main contributions of our work are summarized as follows.

- For the first time, we propose a metric called SSNR to quantify the signal sensing capability. This metric helps people understand the key factors affecting the performance of WiFi sensing and provides a general theoretical foundation for boosting the performance of WiFi sensing.
- By combining CSI signals from multiple antennas, we show both theoretically and experimentally the WiFi sensing range can be significantly increased. The proposed "target nulling" method makes sure the noise reduction process does not corrupt the sensing information on commodity WiFi.
- We build a real-time sensing system with commodity WiFi, and implement two applications including fine-grained human respiration sensing and coarse-grained human tracking. Comprehensive experiments in various settings demonstrate the effectiveness of our proposed methods.

The rest of this paper is organized as follows. Sec. 2 introduces the SSNR metric. Sec. 3 gives the mathematical derivation on how to combine the CSI signals of multiple antennas to maximize SSNR. Sec. 4 presents the implementation of strengthening target-reflected signal and reducing noise on commodity WiFi. Sec. 5 presents the evaluation. Sec. 6 surveys the related work. Sec. 7 discusses the limitations and opportunities followed by a conclusion in Sec. 8.

2 QUANTIFYING THE SENSING CAPABILITY OF CSI SIGNALS WITH SSNR

In this section, we present why the SSNR metric can be used to quantify the sensing capability of CSI signals. First, we introduce how WiFi CSI is utilized for sensing and analyze the factors affecting the signal sensing capability. Then, we present the definition of SSNR and explain why it can be used to quantify the sensing capability.

2.1 Preliminary

2.1.1 WiFi CSI. Given a pair of WiFi transceivers in a typical indoor environment, the WiFi signals propagate from the transmitter to the receiver through multiple paths, and the WiFi signals received at the receiver is a superposition of signals from all the paths. In consequence, the CSI at time t can be generally expressed as [15]:

$$H(f,t) = \sum_{l=1}^{L} a_l(f,t)e^{-j2\pi f \frac{d_l(t)}{c}} \tag{1}$$

where L is the number of propagation paths, f is the carrier frequency, c is the speed of light, $a_l(f,t)$ is the complex number representing the amplitude attenuation and initial phase and $d_l(t)$ is the length of the l-th path. These propagation paths can be grouped into static paths and dynamic paths [26], where the former does not change with time t and the latter changes with target movements. We further take the thermal noise into consideration and the CSI can then be denoted as:

$$H(f,t) = \sum_{i \in \Omega_s(t)} H_i(f,t) + \sum_{j \in \Omega_d(t)} H_j(f,t) + \varepsilon(f,t) \tag{2}$$
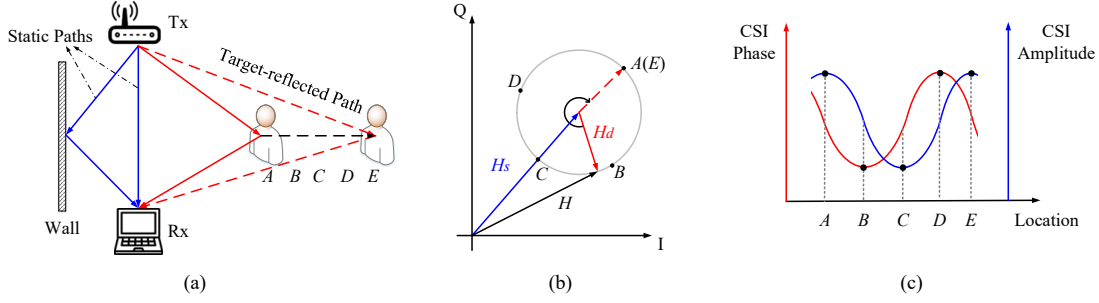
Fig. 2. Illustration of human sensing with WiFi CSI: (a) a target moves from location A to location E; (b) the vector representation of ideal CSI; (c) the CSI phase and amplitude with respect to different target locations.

where $\Omega_s(t)$ denotes the set of static paths, $\Omega_d(t)$ denotes the set of dynamic paths, $H_i(f, t)$ indicates the i-th static path, and $\varepsilon(f, t)$ is the thermal noise, which can be approximated as an additive white Gaussian noise (AWGN) with zero mean and a variance of $\sigma^2(f)$ [31, 43, 44].

2.1.2 Human Sensing Using CSI. Fig. 2(a) presents a typical system setup for sensing, where a WiFi transceiver pair is placed at fixed locations. Apparently, the dynamic path signal is the signal reflected from the target. When the target moves for a short distance, the change in the magnitude of the target-reflected signal is very small and the magnitude can be regarded as a constant in a short time window [17]. After merging the signals of static paths into a constant complex value, we simplify the CSI representation as:

$$H(f, t) = H_s(f) + H_d(f, t) + \varepsilon(f, t) = H_s(f) + a(f)e^{-j2\pi f \frac{d(t)}{c}} + \varepsilon(f, t) \tag{3}$$

where $H_s(f)$ is the static component corresponding to the static paths, $H_d(f, t)$ is the dynamic component corresponding to the target-reflected path and $a(f)$ is a complex coefficient that represents both the amplitude attenuation and initial phase. This complex coefficient is mainly determined by the length of the reflection path and the size/material of the reflection surface.

As Eq. 3 is a bit abstract, we present a visual representation to intuitively show how the CSI changes with target movements in Fig. 2(b). This figure presents the vector representation of the ideal CSI, where there is no white noise $\varepsilon(f, t)$. When the target moves from A to E, the target-reflected path length increases and the dynamic component of the ideal CSI rotates clockwise in the complex I-Q plane accordingly. Apparently, when the reflection path length (RPL) changes one wavelength, the dynamic vector rotates for a full circle in the I-Q plane. In Fig. 2(c), we further present both the CSI phase and amplitude with respect to different target locations. We can observe that when the target moves away from the transceivers gradually, the CSI amplitude shows a continuous sinusoidal-like waveform with peaks and valleys. And when the RPL changes less than one wavelength, e.g., the target moves from A to B, the dynamic component rotates for less than one circle and the CSI amplitude is just a fragment of the sinusoidal-like waveform.

To sense the target movements, our focus is to deduce the RPL change from the CSI change. As shown in Eq. 3 and Fig. 2, when the RPL changes $\Delta d(t)$, the phase of dynamic component $H_d(f, t)$ changes $\Delta\theta = -2\pi f \Delta d(t)/c$. In this way, we can derive the RPL change from the CSI change as follows:

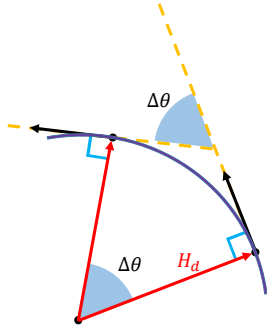$$\Delta d(t) = -\frac{\Delta\theta}{2\pi} \cdot \frac{c}{f} \tag{4}$$

Fig. 3. The radian change is measured by slope of vector on the tangential direction.
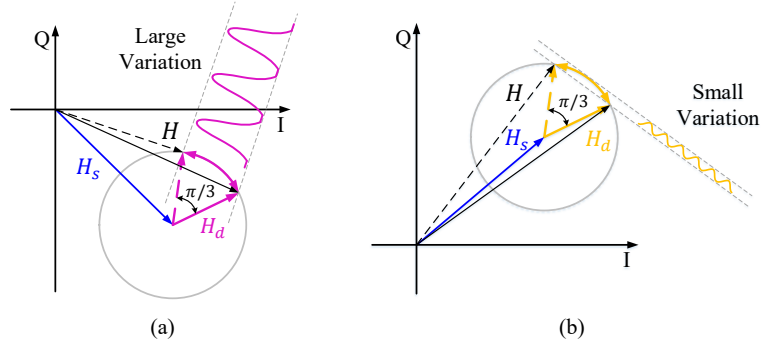
Fig. 4. Two time series of CSI signals that have the same dynamic component H_d but different static components H_s, and the corresponding CSI amplitude variations.
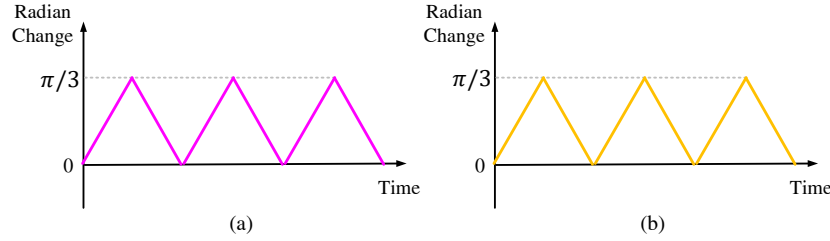

Fig. 5. The corresponding radian changes for the two time series of CSI signals in Fig. 4.

where $\Delta\theta$ is the radian change of the circular arc that the CSI signals generate in the I-Q plane. In practice, we can track the radian change by measuring the slope change of the tangential line at each position on the arc, as shown in Fig. 3.

2.2 Exploring the Factors Affecting CSI Sensing Capability

Observing from the formula of CSI in Eq. 3, it is apparent that there are three key factors that determine the CSI signal, namely $H_s(f)$, $H_d(f, t)$ and $\varepsilon(f, t)$. Let us explore these three possible influencing factors one by one as follows.

First, we study the effect of the non-target static signal $H_s(f)$. Fig. 4 presents two time series of CSI signals that have the same dynamic signal but different static signals. Apparently, the static signal determines the center of the circle in the I-Q plane. As shown in the figure, the same amount of dynamic component rotation (i.e., $\pi/3$) causes dramatically different CSI amplitude variations. The small variation in Fig. 4(b) can be easily buried in noise, thus resulting in a low sensing capability. In this way, previous work [17] claims that the static signal is a key determining factor for the sensing capability. However, we would like to argue that if we use the complex-valued CSI for sensing instead of using only the CSI amplitude, the static signal has no effect on the sensing capability. Note that the key point of contactless WiFi sensing is to deduce the RPL change from the CSI change. As shown in Eq. 4, there is a linear mapping from the radian change to the RPL change. Therefore, instead of employing CSI amplitude, we extract the radian change from complex-valued CSI for sensing. Fig. 5 presents the radian changes corresponding to the two CSI signals in Fig. 4. We can observe from the figure that the radian changes of
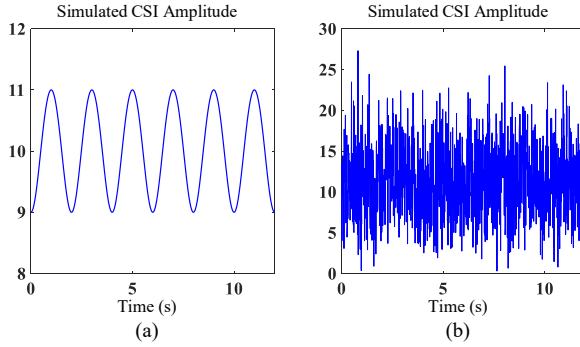
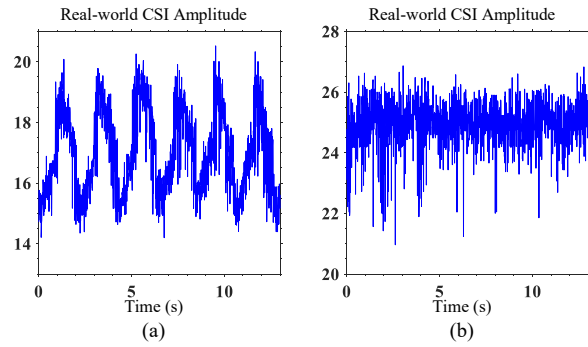Fig. 6. Two amplitude variations of simulated CSI signals: (a) without white noise; (b) with white noise.

Fig. 7. Raw CSI amplitudes in two cases when the metal plates moves: (a) the near case; (b) the far case.

two CSI signals are exactly the same and both of them have large variations, indicating a high sensing capability. The static signal considered "bad" when using CSI amplitude for sensing in Fig. 4, actually, does not affect the performance when we use the complex-valued CSI for sensing. Thus, we should not include the static signal when we quantify the sensing capability of CSI signals.

Next, we study the effect of the white noise $\varepsilon(f, t)$. As shown above, it is straightforward to deduce RPL change from the change of the complex CSI. However, when the noise term $\varepsilon(f, t)$ appears, deducing RPL change from CSI change may be difficult, sometimes even impossible. Fig. 6(a) presents the amplitude of simulated ideal CSI, where there is no white noise. Then, we add the white noise to the ideal CSI and generate the simulated CSI amplitude in Fig. 6(b). Obviously, the amplitude variation is overwhelmed by the large noise in Fig. 6(b), making it hard to sense target movements. Intuitively, a signal of high sensing capability should show a clear fluctuation pattern corresponding to target movements, which means it is easy to deduce RPL change accurately. Therefore, we should include the noise as a variable to quantify the sensing capability of CSI signals. It is obvious that the stronger the noise, the lower the signal sensing capability.

At last, we study the effect of the target-reflected signal $H_d(f, t)$. We conduct a simple experiment by employing a metal plate (ideal signal reflector) as the target reflecting off WiFi signals. With an LoS length of 6 m between the transmitter and receiver, we move a metal plate along the perpendicular bisector of the transmitter and receiver: (1) from 3 m to 3.24 m (the near case); (2) from 8 m to 8.18 m (the far case). For each case, the RPL changes exactly six wavelengths (for 5.24 GHz) and 6 peaks/valleys are thus expected to appear on the CSI amplitude waveform. Fig. 7 shows the raw CSI amplitude variation induced by the metal plate movement for each case. We can observe that the signal variation pattern is quite clear in Fig. 7(a). However, in the far case, the target-reflected signal is much weaker and the small plate-induced signal variation pattern is buried in noise, as shown in Fig. 7(b). Intuitively, if the noise is stronger than the target-reflected signal, the signal variation pattern caused by target movement will be overwhelmed by the noise, leading to a low sensing capability. Therefore, we should include the dynamic signal as a variable to quantify the sensing capability of CSI signals. It is obvious that the stronger the target-reflected signal, the higher the signal sensing capability.

2.3 The Definition of SSNR and Its Significance

Based on the analysis in Sec. 2.2, we know that the sensing capability of a CSI signal is determined by both the target-reflected signal component and the noise. Thus, we propose a metric called SSNR to quantify the sensing

Table 1. Simulation Parameters in Six Cases

	$|H_s(f)|$	$|a(f)|$	$\sigma(f)$
Case 1	10	1	5
Case 2	10	4	5
Case 3	10	1	1.25
Case 4	5	1	5
Case 5	20	1	5
Case 6	9.3	4	5

Table 2. The Corresponding SSNR and SNR in Table 1

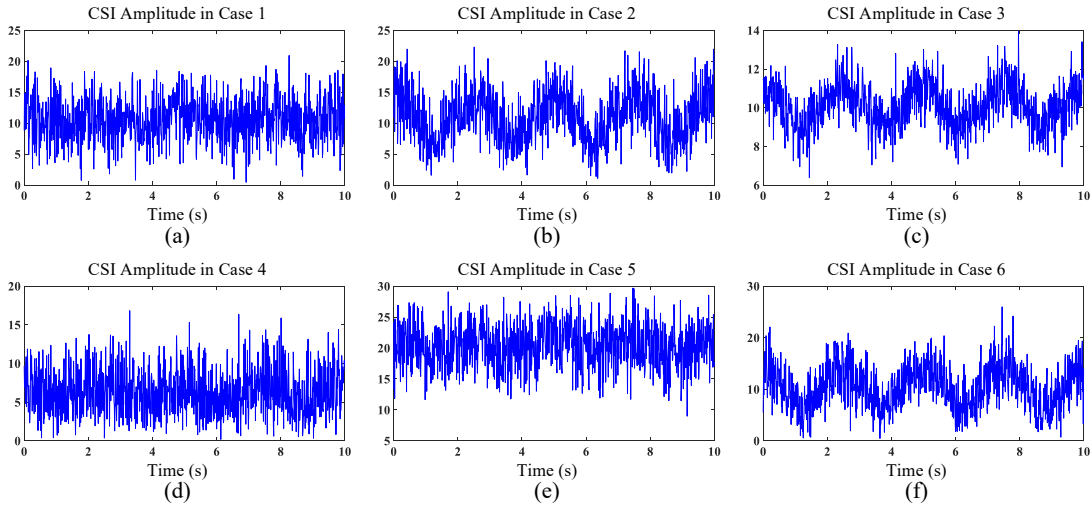|	SSNR (dB)	SNR (dB)
Case 1	-14.0	6.1
Case 2	-2.0	6.7
Case 3	-2.0	18.1
Case 4	-14.0	0.2
Case 5	-14.0	12.1
Case 6	-2.0	6.1

Fig. 8. Simulated CSI amplitude for each case in Table 1.

capability of a CSI signal, which is defined as the ratio of target-reflected signal power and noise power:

$$\gamma = \frac{P_{target}}{P_{noise}} = \frac{\mathbb{E}(|H_d(f,t)|^2)}{\mathbb{E}(|\varepsilon(f,t)|^2)} = \frac{|a(f)|^2 \mathbb{E}(|e^{-j2\pi f \frac{d(t)}{c}}|^2)}{\sigma^2(f)} = \frac{|a(f)|^2}{\sigma^2(f)} \tag{5}$$

where $\mathbb{E}$ refers to the expected value and $H_d(f,t)$ is the dynamic component corresponding to the target-reflected signal. We can see from Eq. 5 that a higher target-reflected signal power or a lower noise power results in a larger SSNR, and thus a higher sensing capability. Because wireless signal strength has a large range, it is often expressed using the logarithmic decibel scale. SSNR is thus expressed in decibels (dB) as $\gamma_{dB} = 10 \log_{10} \gamma$.

The formula of SSNR in Eq. 5 is fairly simple and straightforward. However, it is no less significant. It contributes as an important supplement to WiFi sensing theory. Although a lot of effort has been devoted to WiFi sensing in the last few years, little attention has been paid to formally define a metric to quantify the sensing capability of CSI signals. It also provides a theoretical foundation in improving sensing performance. For example, if we increase the static signal power which is beneficial for communication, it does not help improve the sensing capability.
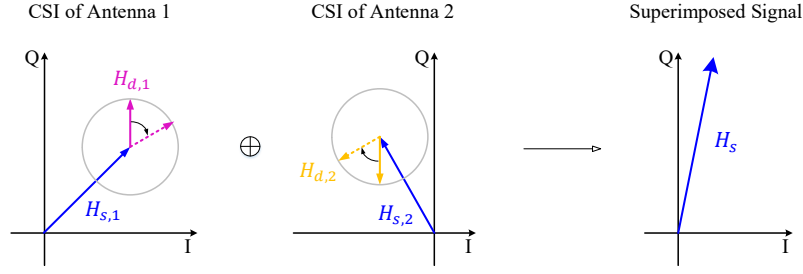
Fig. 9. Directly superimposing the CSI signals of two antennas may null out the target-reflected signal, thus resulting in a signal of low sensing capability.

2.4 Verification

In this subsection, we conduct simulations to verify that SSNR rather than SNR is the proper metric to quantify the signal sensing capability. Here, we generate CSI samples for 10 seconds with a sampling rate of 100 Hz, which means there are 1000 samples in total. The phase of the dynamic component $H_d(f, t)$ is set to evenly rotate for 8π clockwise within 10 seconds. Therefore, we ought to observe $8\pi/2\pi = 4$ peaks/valleys in the CSI amplitude waveform. As shown in Eq. 3, there are three key parameters that determine the generated CSI samples, namely $H_s(f)$, $a(f)$ and $\sigma(f)$. In this way, we carefully design six cases in our simulation verification with different values of $\{|H_s(f)|, |a(f)|, \sigma(f)\}$, as shown in Table 1. As SNR is defined as the ratio of signal power and noise power [23], the SNR of a CSI signal can be expressed as:

$$\beta = \frac{P_{signal}}{P_{noise}} = \frac{\mathbb{E}(|H_s(f) + H_d(f,t)|^2)}{\mathbb{E}(|\varepsilon(f,t)|^2)} = \frac{|H_s(f)|^2 + |a(f)|^2}{\sigma^2(f)} \tag{6}$$

Based on Eq. 5 and Eq. 6, for each case in Table 1, we calculate and present the corresponding SSNR and SNR values in Table 2. Here, both the SSNR and SNR values are expressed in dB. The simulation results are presented in Fig. 8, where each sub-figure is the CSI amplitude waveform corresponding to one case. We can observe from the figures that, the CSI amplitude has a high sensing capability in Case 2, Case 3 and Case 6, while it has a low sensing capability in Case 1, Case 4 and Case 5. By comparing the simulation results with the corresponding SNR and SSNR values, we could draw the following conclusion: the value of SSNR is closely related to the signal's sensing capability while the value of SNR is not. In summary, our proposed SSNR metric is a proper metric which can quantify the signal sensing capability.

3 LEVERAGING MULTIPLE ANTENNAS TO IMPROVE SSNR

In this section, we mathematically show how to leverage multiple antennas to improve the SSNR of CSI signals.

3.1 Combining CSI Signals from Different Antennas

In modern WiFi card, MIMO (Multiple Input Multiple Output) techniques are used to exploit spatial diversity and multiplexing for higher reliability and data throughput [5]. This enables us to obtain multiple CSI readings from different antenna pairs. For example, given N_T transmitting antennas and N_R receiving antennas, we are able to retrieve $M = N_T \times N_R$ CSI readings at one timestamp in total. A natural question is how to make the best use of them. The simplest method is to use the antenna with the best signal (i.e., the one with the largest SSNR) for sensing. It helps with reliability because it is unlikely for all the signals to be bad. However, this method does not utilize those signals which are still good for sensing.

A better method is to combine the CSI signals from all the antennas together. However, this cannot be done by simply adding the signals, which may cancel out the target-reflected signal we want, as shown in Fig. 9. Instead, we combine CSI signals from all the antennas by a weighted sum. To derive the optimal weight for combination, we first present two observations:

(1) When the target moves for a short distance, the length difference of target-reflected path between two adjacent antennas can be considered as a constant [32, 39].
(2) The noise term $\varepsilon(f, t)$ in Eq. 2 is identically and independently distributed across different antennas with the variance $\sigma^2(f)$ [31, 43, 44].

With the observations above, we can rewrite the CSI signals from two antenna pairs as:

$$\begin{cases} H_1(f, t) = H_{s,1}(f) + a_1(f)e^{-j2\pi f \frac{d_1(t)}{c}} + \varepsilon_1(f, t) \\ H_2(f, t) = H_{s,2}(f) + a_2(f)e^{-j2\pi f \frac{d_2(t)}{c}} + \varepsilon_2(f, t) = H_{s,2}(f) + a_2(f)e^{-j2\pi f \frac{d_2(t)-d_1(t)}{c}} e^{-j2\pi f \frac{d_1(t)}{c}} + \varepsilon_2(f, t) \end{cases} \tag{7}$$

where $d_2(t) - d_1(t)$ is the length difference of target-reflected path between two antennas, which is a constant based on the first observation above. To simplify Eq. 7 for easier illustration, we employ $A_1(f)$ and $A_2(f)$ to represent the terms: $a_1(f) = A_1(f)$ and $a_2(f)e^{-j2\pi f \frac{d_2(t)-d_1(t)}{c}} = A_2(f)$. We can then simplify Eq. 7 as:

$$\begin{cases} H_1(f, t) = H_{s,1}(f) + A_1(f)e^{-j2\pi f \frac{d_1(t)}{c}} + \varepsilon_1(f, t) \\ H_2(f, t) = H_{s,2}(f) + A_2(f)e^{-j2\pi f \frac{d_1(t)}{c}} + \varepsilon_2(f, t) \end{cases} \tag{8}$$

where $e^{-j2\pi f \frac{d_1(t)}{c}}$ is a unit circle that rotates clockwise when the target-reflected path $d_1(t)$ increases. The above analysis can be applied to the remaining CSI signals, and they have the same form as Eq. 8. At last, we rewrite all M CSI signals in the vector form as below:

$$\mathbf{H}(f, t) = \mathbf{H}_s(f) + \mathbf{A}(f)e^{-j2\pi f \frac{d_1(t)}{c}} + \boldsymbol{\varepsilon}(f, t) \tag{9}$$

where $\mathbf{H}(f, t) = \begin{bmatrix} H_1(f, t) H_2(f, t) ... H_M(f, t) \end{bmatrix}^T$ represents the CSI signals of all antenna pairs, $\mathbf{H}_s(f) = \begin{bmatrix} H_{s,1}(f, t) H_{s,2}(f) ... H_{s,M}(f) \end{bmatrix}^T$ represents the static signal, $\mathbf{A}(f) = \begin{bmatrix} A_1(f) A_2(f) ... A_M(f) \end{bmatrix}^T$ represents the complex coefficient of target-reflected signal and $\boldsymbol{\varepsilon}(f, t) = \begin{bmatrix} \varepsilon_1(f, t) \varepsilon_2(f, t) ... \varepsilon_M(f, t) \end{bmatrix}^T$ represents the noise.

Let $\mathbf{W} = \begin{bmatrix} w_1 w_2 ... w_M \end{bmatrix}^T$ denote the normalized complex weight, i.e., $||\mathbf{W}||_2 = 1$. Here, $||\mathbf{W}||_2$ is the 2-norm of vector $\mathbf{W}$, that is, $||\mathbf{W}||_2 = \sqrt{\sum_{i=1}^{M} |w_i|^2}$. Then, the combined CSI signal of all antennas can be expressed as:

$$H_{comb}(f, t) = \mathbf{W}^T \mathbf{H}(f, t) = \mathbf{W}^T \mathbf{H}_s(f) + \mathbf{W}^T \mathbf{A}(f)e^{-j2\pi f \frac{d_1(t)}{c}} + \mathbf{W}^T \boldsymbol{\varepsilon}(f, t) \tag{10}$$

3.2 Maximizing SSNR of the Combined Signal

Now, let us derive the optimal weight that maximizes SSNR of the combined signal. Based on Eq. 5, the SSNR of the combined signal can be expressed as:

$$\gamma_{comb} = \frac{P_{target}}{P_{noise}} = \frac{\mathbb{E}(|\mathbf{W}^T \mathbf{A}(f)e^{-j2\pi f \frac{d_1(t)}{c}}|^2)}{\mathbb{E}(|\mathbf{W}^T \boldsymbol{\varepsilon}(f, t)|^2)} = \frac{|\mathbf{W}^T \mathbf{A}(f)|^2}{\sigma^2(f)} \tag{11}$$

We put the details of the derivation in appendix A, and summarize the final result as above. Interestingly, the noise of combined signal still has the power of $\sigma^2(f)$. By the Cauchy-Schwarz inequality [2], we have:

$$|\mathbf{W}^T \mathbf{A}(f)|^2 \leq ||\mathbf{W}||_2^2 \cdot ||\mathbf{A}(f)||_2^2 \tag{12}$$

The equality is achieved when $\mathbf{W}$ is linearly proportional to $\mathbf{A}^*(f)$, i.e., $\mathbf{W} = m\mathbf{A}^*(f)$, where the operation $(\cdot)^*$ indicates the complex conjugate and m is any non-zero complex number [2]. Because $||\mathbf{W}||_2 = 1$, we have
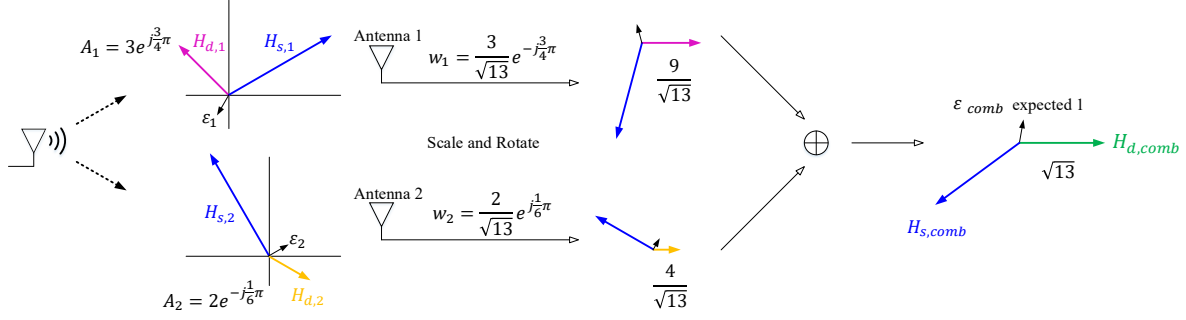
Fig. 10. An example of optimal CSI signal combination for two-antenna case. The complex coefficients of target-reflected signals for these two antennas are $3e^{j\frac{3}{4}\pi}$ and $2e^{-j\frac{1}{6}\pi}$, respectively. According to Eq. 13, the optimal weight is $\left[\frac{3}{\sqrt{13}}e^{-j\frac{3}{4}\pi} \quad \frac{2}{\sqrt{13}}e^{j\frac{1}{6}\pi}\right]^T$.

$|m| = 1/||\mathbf{A}(f)||_2$. Therefore, the optimal weight that maximizes the SSNR of combined signal can be expressed as:

$$\mathbf{W}_{opt} = e^{j\theta_{ref}}\frac{\mathbf{A}^*(f)}{||\mathbf{A}(f)||_2} \tag{13}$$

where θ_{ref} is any real number, whose value is generally taken as 0. By substituting Eq. 12 into Eq. 11, we get the maximum of SSNR as below:

$$\max(\gamma_{comb}) = \frac{||\mathbf{W}||_2^2 \cdot ||\mathbf{A}(f)||_2^2}{\sigma^2(f)} = \sum_{i=1}^{M}\frac{|A_i(f)|^2}{\sigma^2(f)} \tag{14}$$

Obviously, the maximum of SSNR is exactly the sum of SSNRs of all antenna pairs. By optimally combining the CSI signals, the target-reflected signal is greatly strengthened while the noise keeps the same power so that a CSI signal of higher sensing capability is obtained to support longer-range sensing.

3.3 The Meaning of the Optimal Combining Weight

Now, we explain the meaning of the optimal weight $\mathbf{W}_{opt}$ that maximizes the SSNR of combined signal. For better understanding, we rewrite Eq. 13 as below:

$$w_{i,opt} = e^{j\theta_{ref}}\frac{A_i^*(f)}{\sqrt{\sum_{i=1}^{M}|A_i(f)|^2}} \tag{15}$$

where $w_{i,opt}$ is the i-th element in the optimal weight vector $\mathbf{W}_{opt}$, corresponding to the i-th antenna pair $(i = 1, 2, ..., M)$. We can observe from Eq. 15 that:

(1) The target-reflected signal of each antenna is scaled by its magnitude. This gives less weight to a signal that has a larger fraction of noise, so that the effect of noise is not amplified. In this way, the "good" antennas will boost the signal while the "bad" antennas will help attenuate the noise since their noise terms are independent.

(2) The complex coefficient of target-reflected signal at each antenna, that is, $A_i(f)$, is aligned to the same phase θ_{ref} after multiplying it by $w_{i,opt}$; then, the power of the target-reflected signals will combine coherently.

3.4 An Illustrative Example

Fig. 10 depicts the optimal combination operation graphically for CSI signals at two antennas. In this example, the target-reflected signals of two antennas (i.e., $H_{d,1}$ and $H_{d,2}$) have magnitudes of $\{3, 2\}$ and phases of $\{\frac{3}{4}\pi, -\frac{1}{6}\pi\}$, respectively. With expected noise power 1, the two signals correspond to SSNRs of 9 and 4. As shown in the figure, each antenna's signal is first multiplied by the magnitude of the target-reflected signal (i.e., 2 and 3 respectively) and then normalized by the square sum (i.e., $\sqrt{13}$) of the magnitudes. Then, the two signals are rotated to the same direction ($\theta_{ref} = 0$ in this example) and added together. The resulted target-reflected signal $H_{d,comb}$ has a magnitude of $\sqrt{13}$ and the normalized weighted sum of noise ε_{comb} still has expected power 1. Thus, by coherently combining target-reflected signals from different antennas, the optimal combination output has the expected SSNR of 13, which is the sum of SSNRs at these two antennas. In this example, if we only align the signals without scaling them, the weight for combination will be $\left[\frac{1}{\sqrt{2}}e^{-j\frac{3}{4}\pi} \quad \frac{1}{\sqrt{2}}e^{j\frac{1}{6}\pi} \right]^T$. Then, the resulted target-reflected signal has a magnitude of $5/\sqrt{2}$, indicating an SSNR of $25/2$. Obviously, although this combination helps improve SSNR, it is not the optimal one ($25/2 < 13$).

4 IMPLEMENTATION ON COMMODITY WIFI

In this section, we first present the representation of the CSI reading obtained from commodity WiFi. Then, we present how to strengthen the target-reflected signal. Next, we show how to reduce the noise on commodity WiFi without corrupting the sensing information. At last, we conduct benchmark experiments to validate the effectiveness of the proposed methods.

4.1 The CSI Reading from Commodity WiFi

For a pair of commodity WiFi transmitter and receiver, the measured CSI is affected not only by the signal propagation paths between transceivers but also by the hardware [15]. On one hand, the AGC in RF chain usually results in impulse and burst noise in CSI amplitude [47]. On the other hand, the carrier frequency difference between the transmitter and receiver leads to a time-varying phase offset in each CSI sample, which distorts the phase variation caused by target movements [6, 24]. In this way, the CSI readings obtained from commodity WiFi can be expressed as [16]:

$$\widetilde{H}(f, t) = \delta(t)e^{-j\phi(t)}H(f, t) \tag{16}$$

where $\delta(t)$ is the AGC noise, $\phi(t)$ is the time-varying phase offset and $H(f, t)$ refers to the CSI value in Eq. 3. Two interesting observations were identified by previous works: (1) for signal amplitude, the AGC noise at each antenna is similar at the same receiver [16, 32, 39]; (2) for signal phase, since different antennas on the same receiver share the same oscillator, the random phase offsets caused by carrier frequency offset (CFO) and sampling frequency offset (SFO) are identical [6, 9]. As shown in Eq. 3 and Eq. 5, the SSNR metric is defined with only the thermal noise taken into consideration. Since the real-world CSI measurements also contain the AGC noise and time-varying phase offset, the SSNR cannot be directly applied to real-world raw CSI measurements. Fortunately, we will show in Sec. 4.4 that the AGC noise and time-varying phase offset can be effectively cancelled out by a division operation. In this way, the SSNR metric can be applied to the processed CSI to guide how to improve the signal quality for sensing purpose.

4.2 Strengthening Target-reflected Signal

Based on Eq. 13, the optimal weight for constructive combination can be directly obtained from the complex coefficient $A_i(f)$ of the target-reflected signal at each antenna. Unfortunately, we have no idea of the complex coefficient $A_i(f)$. It is hard to directly estimate $A_i(f)$ from the CSI readings for two reasons: (i) the static signal

and target-reflected signal are mixed together in the CSI signal; and (ii) the time-varying phase offset distorts the phase of the target-reflected signal.

Fortunately, based on Sec. 3.3, we find that the optimal weight for combination actually scales the target-reflected signal at each antenna and aligns them to the same phase reference. Therefore, we instead propose to estimate the relative magnitude and phase between the target-reflected signals of two antennas, which can be then utilized to derive the optimal weight for combination. We start from a simplified two-antenna case to combine the CSI signals before generalizing it to the realistic, multi-antenna case.

4.2.1 Two-antenna Case. Let the complex number $k_{i,j} = \Delta p e^{j\Delta q}$ denote the relative magnitude Δp and phase Δq of target-reflected signals at the j-th antenna with respect to the i-th antenna. Based on the definition, we have $k_{i,j} = A_j(f)/A_i(f)$, where $A_i(f)$ and $A_j(f)$ denote the complex coefficients of target-reflected signal at the i-th and j-th antenna, respectively. According to Sec. 3, the optimal weight for combining CSI signals at these two antennas can be expressed as:

$$\mathbf{W}_{opt} = r \cdot \left[1 \ k_{i,j}^* \right]^T \tag{17}$$

where the factor $r = 1/\sqrt{1 + |k_{i,j}^*|^2}$ normalizes the vector so that its 2-norm equals to 1.

To estimate $k_{i,j}$, we perform the opposite operation of strengthening target-reflected signal process to destructively combine the target-reflected signals received at the i-th and the j-th antennas. We term this process as "target nulling". Let α be a complex number, the weighted sum of signals at these two antennas is thus given by:

$$\widetilde{H}_{comb}(f,t) = \widetilde{H}_i(f,t) + \alpha \widetilde{H}_j(f,t)$$

$$= \delta(t)e^{-j\phi(t)}[\underbrace{(H_{s,i}(f,t) + \alpha H_{s,j}(f,t))}_{\text{static signal combination}} + \underbrace{(A_i(f) + \alpha A_j(f)) \cdot e^{-j2\pi f \frac{d_1(t)}{c}}}_{\text{target-reflected signal combination}} + \underbrace{(\varepsilon_i(f,t) + \alpha \varepsilon_j(f,t))}_{\text{thermal noise combination}}] \tag{18}$$

where $\widetilde{H}_i(f,t)$ and $\widetilde{H}_j(f,t)$ are the CSI signals obtained at the i-th and the j-th antennas on commodity WiFi, respectively; $A_i(f) + \alpha A_j(f)$ is the complex coefficient of target-reflected signal in the weighted sum signal. If we successfully null out the target-reflected signal, we will have $A_i(f) + \alpha A_j(f) = 0$, that is, $\alpha = -A_i(f)/A_j(f)$. In this way, the complex number $k_{i,j}$ representing the relative magnitude and phase can be obtained as $k_{i,j} = -1/\alpha$.

Next, we explain how to find a complex number α that can null out the target-reflected signal. Here, we formulate it as an optimizing problem. Given a complex number α, we define the corresponding objective function as follows:

Definition 4.1. **Target-reflected Signal Energy Ratio (TSER)** is defined as the ratio of target motion energy to the overall energy in the frequency spectrum.

Specifically, the calculation of TSER consists of three steps: (i) perform FFT (Fast Fourier Transform) operation on the amplitude of the weighted signal sum over a time window; (ii) calculate the energy percentage of each bin with respect to the energy of all the bins; (iii) sum up the energy percentage for those bins in a certain frequency range corresponding to a specified activity, e.g., the frequency range for respiration is 0.1-0.5 Hz. Then the genetic algorithm [29] is utilized to find the optimal weight α by minimizing the objective function above (i.e., the TSER). The genetic algorithm is a metaheuristic inspired by the process of natural selection with the ability of not being trapped at a local optimal. In our implementation, the input of the algorithm is two time series of CSI readings corresponding to two antennas and the output is the optimal α. The genetic algorithm works as follows:

(1) Initialization. We randomly create a group of N potential α ($N = 40$ in our implementation).
(2) Evaluation. We calculate the TSER of each α. A lower TSER indicates a better α.
(3) Selection. We select those α with a low TSER and eliminate the rest. In our implementation, among the 40 α, we keep 20 of them and eliminate the rest 20.

(4) Crossover and mutation. The 20 α kept are called parents and we slightly change each of them to generate a new (child) α. The newly generated 20 α and the kept 20 α form the new group of N potential α.

(5) Termination. We stop if the relative change in the TSER value is less than a threshold (10^{-4} in our implementation). Otherwise, we return to step 2.

4.2.2 Multi-antenna Case. We now consider the realistic case where M antennas are present. In Sec. 4.2.1, we have shown how to implement the idea for two-antenna case. To extend the idea to work for M antennas, we first choose one random antenna as the reference antenna and then let each of the remaining $M - 1$ antennas scale and align with respect to the reference antenna. In this way, the M antenna case is converted to $M - 1$ two antenna cases. In our implementation, we take the antenna with the largest amplitude variance as the reference antenna. For each two-antenna case, we execute the process elaborated in Sec. 4.2.1 to estimate $k_{i,j}$. The optimal weights for combining CSI signals at these M antennas are expressed as:

$$\mathbf{W}_{opt} = r \cdot \left[\underbrace{k_{i,1}^* \; k_{i,2}^* \; ... \; k_{i,i-1}^*}_{i-1} \; 1 \; \underbrace{k_{i,i+1}^* \; ... \; k_{i,M}^*}_{M-i} \right]^T \tag{19}$$

where the factor r normalizes the vector and the i-th antenna is the reference antenna. The CSI signal after strengthening can be denoted as:

$$\begin{aligned} \widetilde{H}_{enh}(f,t) &= \mathbf{W}_{opt}^T \widetilde{\mathbf{H}}(f,t) \\ &= \delta(t) e^{-j\phi(t)} \mathbf{W}_{opt}^T \mathbf{H}(f,t) \\ &= \delta(t) e^{-j\phi(t)} (H_{s,enh}(f,t) + A_{enh}(f) e^{-j2\pi f \frac{d_1(t)}{c}} + \varepsilon_{enh}(f,t)) \end{aligned} \tag{20}$$

where $\widetilde{\mathbf{H}}(f,t) = \left[\widetilde{H}_1(f,t) \; \widetilde{H}_2(f,t) \; ... \; \widetilde{H}_M(f,t) \right]^T$ represents the CSI signals of all antennas.

4.3 Reducing Noise in the CSI Signal

In Sec. 4.2, after multi-antenna signal combination, the target-reflected signal containing the target information is greatly strengthened. However, the AGC noise $\delta(t)$ and the time-varying phase offset $\phi(t)$ still exist, as shown in Eq. 20. On one hand, the AGC noise may distort the target movement-induced signal variation. On the other hand, due to the time-varying phase offset, the phase information of the composite signal is unusable. Existing work [39] calculates the ratio of CSI signals between two antennas, which is termed CSI ratio, to reduce the AGC noise and remove the time-varying phase offset. It works because although the AGC noise and phase offset change over time, they are similar for signals retrieved from two antennas at the same receiver [6, 9, 32, 39].

However, during this process of signal ratio calculation, the sensing information (signal phase and amplitude) can be corrupted. This is because the signal received at each antenna contains the dynamic component and therefore a division operation between two signals both containing dynamic components can corrupt the amplitude/phase information of the dynamic signal. This corruption does not affect the periodicity of the CSI signals, thus previous work on respiration sensing [39] can accurately estimate the respiration rate with CSI ratio. However, it does affect the radian change of the circular arc that the CSI signals generate in the I-Q plane, thus resulting in an RPL change estimation that is not sufficiently accurate. The RPL change estimation is critical for applications involving tracking target movements, such as finger motion tracking.

Fig. 11 presents an example of the side effect of the signal ratio method. Fig. 11(a)-Fig. 11(b) show two CSI signals retrieved from two antennas at the same receiver, i.e., $H_1(f,t)$ and $H_2(f,t)$. As the target movement causes the RPL to change for $\lambda/4$ (λ is the wavelength), the radian change of the circular arc is $\pi/2$ for each antenna. Fig. 11(c) presents the ratio of CSI signals between these two antennas, that is, $H_1(f,t)/H_2(f,t)$. Interestingly, the
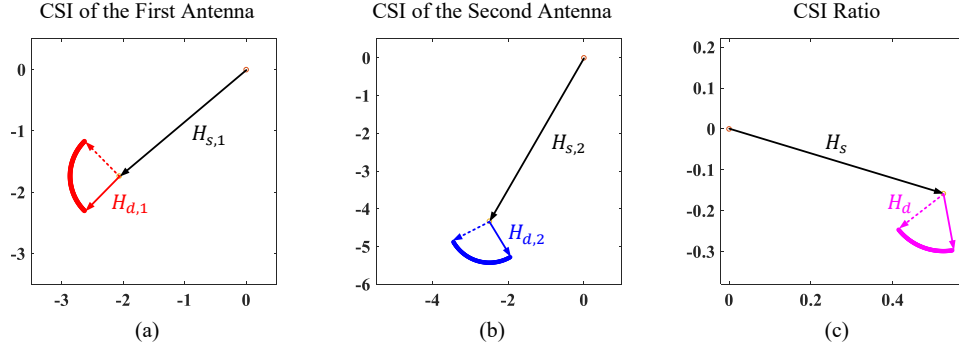
Fig. 11. The signal ratio method corrupts the signal amplitude and phase. In this example, given the dynamic component of CSI for each antenna rotates $\pi/2$, the dynamic component of CSI ratio between these two antennas rotates just $\pi/3$, where the expected radian change is $\pi/2$.

radian change of this CSI ratio is about $\pi/3$, indicating a RPL change of $\lambda/6$ which is not accurate.[2] We can see that the signal ratio method can corrupt the sensing information and affect the sensing accuracy. Therefore, we need to find another reference signal without the dynamic component to address the AGC noise and phase offset issue without corrupting the information for sensing. Next, we explain how to accomplish this. Our objective is to find a reference signal which contains the noise but not the dynamic signal component. This reference signal can then serve as the denominator during the signal division operation to cancel the unwanted AGC noise and time-varying phase offset without corrupting the sensing information.

To achieve this, we apply the "target nulling" method described in Sec. 4.2.1. This "target nulling" process destructively combines the target-reflected signals received at two antennas, which is exactly what we need for constructing the reference signal. After finding the optimal complex number α in Eq. 18, we can quickly obtain the "target nulling" reference signal that does not contain the target-reflected signal. We denote this signal as:

$$\widetilde{H}_{ref}(f, t) = \delta(t)e^{-j\phi(t)}(H_{s,ref} + \varepsilon_{ref}(f, t)) \tag{21}$$

Finally, to reduce the unwanted AGC noise and time-varying phase offset, we divide the strengthened signal by the reference signal using Eq. 20 and Eq. 21 as follows:

$$
\begin{aligned}
H_{fin}(f, t) &= \frac{\widetilde{H}_{enh}(f, t)}{\widetilde{H}_{ref}(f, t)} \\
&= \frac{\delta(t)e^{-j\phi(t)}(H_{s,enh}(f, t) + A_{enh}(f)e^{-j2\pi f \frac{d_1(t)}{c}} + \varepsilon_{enh}(f, t))}{\delta(t)e^{-j\phi(t)}(H_{s,ref} + \varepsilon_{ref}(f, t))} \\
&\approx \frac{H_{s,enh}}{H_{s,ref}} + \frac{A_{enh}(f)}{H_{s,ref}}e^{-j2\pi f \frac{d_1(t)}{c}} + \frac{\varepsilon_{enh}(f, t)}{H_{s,ref}}
\end{aligned}
\tag{22}
$$

Here, we can safely ignore the thermal noise term $\varepsilon_{ref}(f, t)$ in the denominator, since it is orders of magnitude smaller than the static signal $H_{s,ref}$. From Eq. 22, we can see that the AGC noise and time-varying phase offset are effectively eliminated. The remaining part only depends on the target-reflected path length $d_1(t)$ and some constants (i.e., the static signal), as well as the thermal noise. As shown in Eq. 22, during the process of multi-antenna signal combination, the ToF (Time of Flight) and DFS (Doppler Frequency Shift) information are kept

---

[2]This phenomenon can be explained by the mathematical analysis of CSI ratio in [39]. Here, we omit it for brevity.
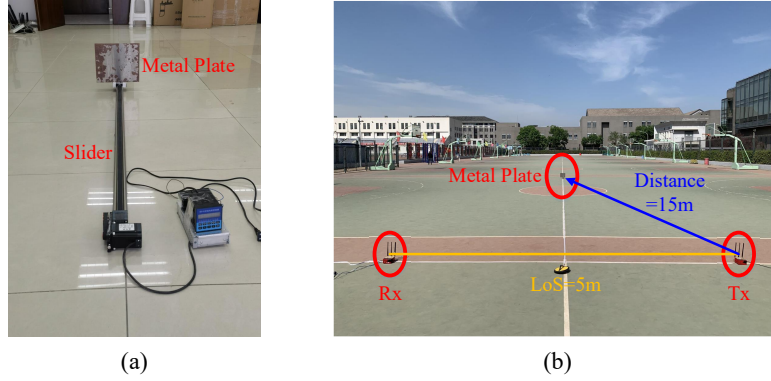
Fig. 12. Experimental setup: (a) the slider with a metal plate; (b) the environment.

while the phase information which is critical for AoA (Angle of Arrival) estimation gets distorted. Thus, to estimate AoA, signals before multi-antenna combination should be used. By performing the two processing steps in Sec. 4.2 and Sec. 4.3, the dynamic component that contains target information gets strengthened and the noise is well reduced without corrupting the sensing information. The processed signal thus has a higher sensing capability, which serves as the theoretical basis for long-range WiFi sensing.

4.4 Understanding the Effects of the Proposed Methods

In this subsection, we first conduct simulations to verify the effectiveness of our proposed methods for two-antenna combination case. For each of the two antennas, we generate CSI samples for 10 seconds at a sampling rate of 100 Hz. Each CSI sample is composed of three parts: static signal $H_s(f, t)$, target-reflected signal $H_d(f, t)$ and thermal noise $\varepsilon(f, t)$. For CSI samples, we create a total of 500 combinations of static signal $H_s(f, t)$, target-reflected signal $H_d(f, t)$ and thermal noise $\varepsilon(f, t)$.

We compare the SSNR achieved by our proposed method with the theoretic optimal SSNR. To quantitatively evaluate the method, we adopt the metric of non-completion rate, which is defined as $\frac{|\gamma_{opt} - \gamma_{our}|}{\gamma_{opt}} \times 100\%$, in which γ_{opt} is the theoretic optimal SSNR and γ_{our} is the SSNR achieved by our method. A lower non-completion rate indicates a more effective SSNR improvement. The simulation results show that the median non-completion rate is 2.8% while the 90th percentile non-completion rate is 7.7%. The simulation results demonstrate the effectiveness of the proposed method.

Then, we conduct benchmark experiments to visually show the effects of our proposed methods in sensing the movements of a target far away from the WiFi transceivers. We setup a pair of GigaByte mini-PCs each equipped with an Intel 5300 card as the transceivers. Both the transmitter and receiver have three omnidirectional antennas. The central frequency is set to 5.24 GHz, which corresponds to a wavelength of 5.73 cm. A metal plate with a size of 28 cm × 25 cm is adopted as the target reflecting off WiFi signals. We use a metal plate rather than a human subject in this experiment because we need to precisely control the speed and displacement of the target which is not feasible with a human subject. As shown in Fig. 12(a), we mount the plate on a linear motion slider to precisely control its movements. As shown in Fig. 12(b), the LoS path between the transceivers is set as 5 m, and the slider is placed at a distance of 3 m-15 m to the transceivers at a step size of 1 m. At each distance, we control the plate to move a certain distance to ensure that the RPL changes exactly 8λ, indicating the theoretical number of peaks/valleys on the CSI amplitude waveform is $8\lambda/\lambda = 8$.
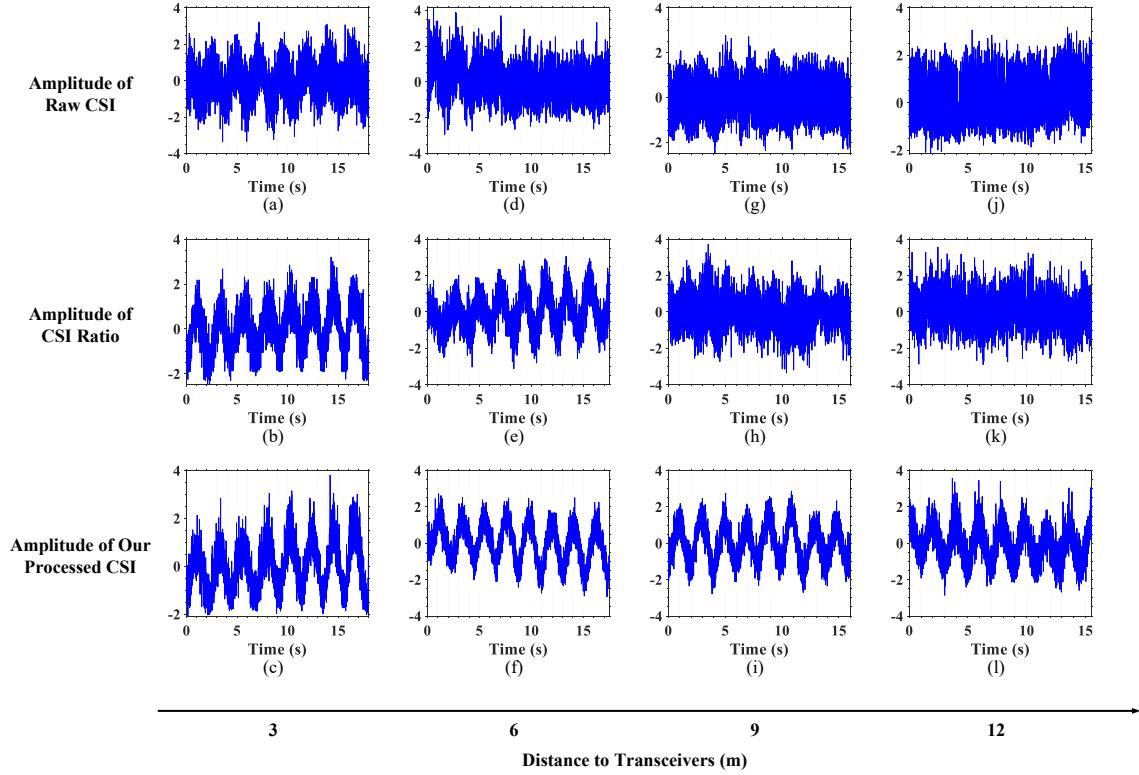
Fig. 13. The amplitude variations caused by the metal plate movement at 4 different distances to the transceivers.

Fig. 13 shows the normalized amplitudes at different distances. Three types of amplitudes are presented in the figure: (1) the amplitude of the raw CSI, (2) the amplitude of the CSI ratio, and (3) the amplitude of our processed CSI. At a distance of 3 m, all the amplitudes exhibit clear fluctuations corresponding to the target movement. Among these three clear amplitude patterns, our method achieves the clearest fluctuation. When the distance is increased to 6 m, the amplitude of raw CSI is buried by noise, and we can hardly extract the target movement information. However, the amplitudes of CSI ratio and our processed CSI still have clear fluctuations which match the ground truths. At the distance of 12 m, only the amplitude of our processed CSI shows clear patterns. The experiment results demonstrate the promises of our proposed method in terms of increasing the sensing range.

5   APPLICATIONS AND EVALUATION

In this section, building upon our proposed methods, we implement two applications on commodity WiFi: (1) long-range respiration sensing; and (2) long-range human tracking. We conduct comprehensive experiments to evaluate the performance of these two sensing applications. The system consists of a pair of WiFi transceivers equipped with cheap Intel 5300 WiFi cards. Unless specified otherwise, both the transmitter and receiver are equipped with three omnidirectional antennas. We term this antenna configuration as "3 × 3 AC". The frequency of WiFi channel is set to 5.24 GHz with a bandwidth of 20 MHz. We use the CSI Tool developed by Halperin [5] to collect CSI samples on the receiver at the sampling rate of 200 Hz. The WiFi receiver is connected to a data

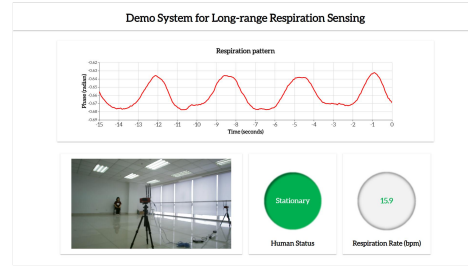Fig. 14. Respiration monitor belt logger sensor.

Fig. 15. The web-based GUI for long-range respiration sensing.

processing laptop with an Intel Xeon E3-1505M v6 CPU and 8 GB RAM, which processes the CSI readings with MATLAB in real time.

5.1 Long-range Respiration Sensing

In this subsection, we first present the implementation of a real-time respiration sensing system. Then, we detail the data collection and the evaluation metrics. After that, we validate the effectiveness of respiration sensing in challenging scenarios, where the target is far away from the transceivers and behind the walls. We also discuss the impact of various factors on system performance, including the number of antennas, environment changes and surrounding interference. Finally, we compare the performance of the proposed system with two state-of-the-art systems: PhaseBeat [27] and FarSense [39].

5.1.1 System Implementation. We first collect the CSI readings from multiple antennas using the CSI tool [5], which provides CSI information on 30 sub-carriers. Then, we adopt the motion detection method proposed in [43] to flag periods of time when the target moves and divide the reminding CSI samples into a series of stable periods. Next, we process the CSI data of stable periods with a 30-second sliding window in real time. For each sub-carrier, our proposed signal processing methods are adopted to strengthen target-reflected signal and reduce noise. We apply the Savitzky-Golay filter [20] to smooth the complex-valued CSI signal and further apply the PCA (Principle Component Analysis) method [3] to extract respiration pattern from the processed CSI. At last, we estimate respiration rate from the recovered respiration pattern. Our system can accurately detect abnormal respiration patterns for events such as apnea and further obtain the respiration rate. We achieve this goal by first computing the variance of the recovered respiration pattern in a short sliding window (2 seconds in our implementation) and then using a threshold-based method to identify an apnea event. If the variance drops below the threshold, an apnea event is detected. Once the abnormal respiration patterns are detected, we utilized the peak detection method with fake peak removal [11] to estimate respiration rate, and refine it by averaging the estimates from multiple antennas. For normal respiration patterns, we employ autocorrelation [30] to estimate the respiration rate and refine it by combining the autocorrelation result of multiple sub-carriers as in [39]. We obtain the ground-truth respiration information using a respiration monitor belt logger sensor NUL-236 [1], as shown in Fig. 14. We develop a GUI (Graphical User Interface) to visualize the human status (stationary or non-stationary), the respiration pattern and the corresponding respiration rate, as shown in Fig. 15.

5.1.2 Data Set and Evaluation Metrics. To fully evaluate the performance of respiration sensing, we recruit 15 volunteers to participate in the experiments over three weeks. These volunteers include 9 males and 6 females, aged from 18 to 52. We conduct experiments when each volunteer breathes, including both normal breathing and abnormal breathing. We conduct experiments in different environments detailed in Sec. 5.1.3. We vary the
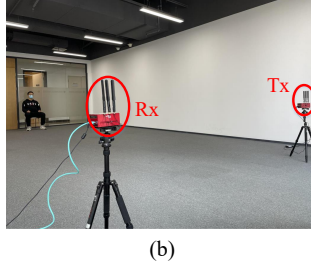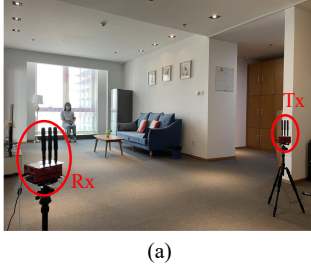
Fig. 16. The experimental setup in two LoS scenarios: (a) a typical home; (b) an empty room.
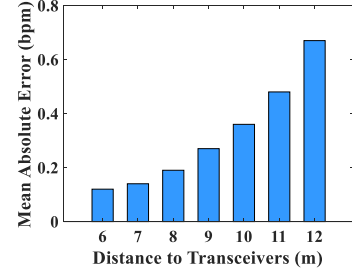
Fig. 17. The mean absolute error of respiration rate versus distance to transceivers in LoS scenarios.

distance between the subject and the transceivers from 2 m to 12 m at a step size of 1 m. We also vary the distance between the transmitter and receiver from 2.5 m to 7.5 m at a step size of 1 m.

We adopt two metrics to quantify the performance of respiration sensing: one is the mean absolute respiration rate error and the other is the detection rate. The mean absolute error is calculated as $\sum_{i=1}^{N} |b_{est} - b_{gt}|/N$, where N is the total number of CSI measurements, b_{est} is the estimated respiration rate and b_{gt} is the ground-truth respiration rate. The detection rate at a certain distance is defined as N_{detect}/N, where N_{detect} is the number of measurements whose respiration rate estimation is close to the ground truth (i.e., $|b_{est} - b_{gt}| < 0.5$). In this way, we define the sensing range of a respiration sensing system as the maximal distance where the detection rate is higher than 95%.

5.1.3 Effectiveness of Respiration Sensing. We now evaluate the respiration sensing performance with varying distances and through walls.

(1) **Long-range sensing.** Fig. 16 presents the two LoS scenarios for evaluating long-range respiration sensing: one is a typical home and the other is an empty room (11.7 m × 10.8 m). Fig. 17 plots the mean absolute error of respiration rate as a function of distance from 6 m to 12 m. The results show that the error is only 0.12 bpm (breaths per minute) at 6 m, and slightly increases to 0.67 bpm when the subject is 12 m away from the transceivers. When the distance from the target to the transceivers increases, the mean absolute error also increases. We can see that the error is still less than 0.5 bpm when the distance is 11 m. Therefore, our system can reliably sense human respiration even when the target is 11 m away from the transceivers.

(2) **Through-wall sensing.** We further study the sensing performance when there are walls between the target and the transceivers. Fig. 18 presents the two NLoS (Non-Line-of-Sight) scenarios for evaluating long-range respiration sensing, where the target is located in a different room from the transceivers: one is a typical home and the other is a meeting room. In each environment, we have three transceiver setups. Fig. 19 shows the mean absolute respiration sensing error versus the target-transceiver distance. We find that when there is a wall in between, we can still sense respiration at a distance of 6 m. At the same distance, the mean absolute error in NLoS scenarios is larger than that in LoS scenarios. This is because the signal reflected off the target becomes even weaker after penetrating the wall twice in NLoS scenarios. The results indicate that our system can reliably sense human respiration even in the challenging scenario with a wall in between.
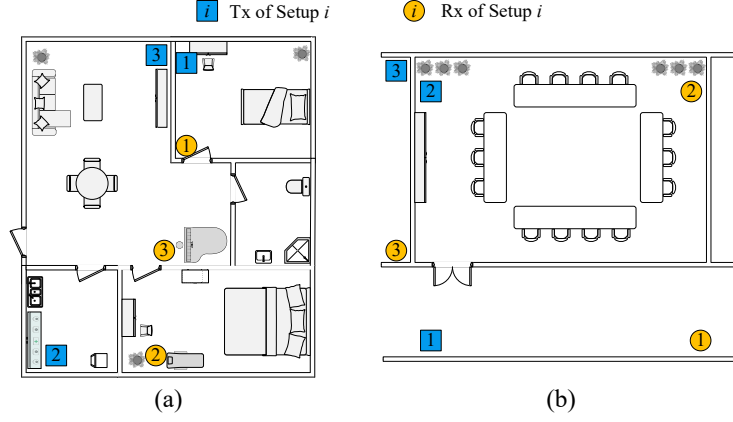
Fig. 18. The experimental setup in two NLoS scenarios: (a) a typical home;
(b) a meeting room. In each environment, we have 3 transceiver setups.
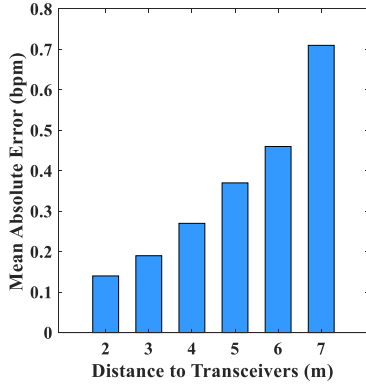

Fig. 19. The mean absolute error of respiration rate versus distance to transceivers in NLoS scenarios.
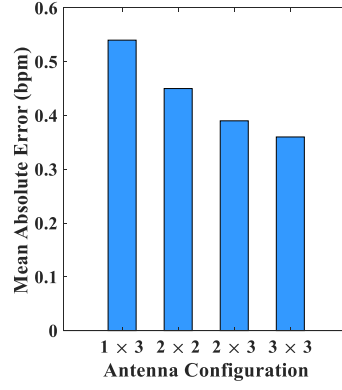
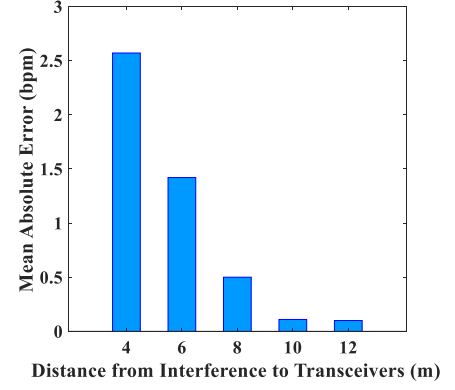Fig. 20. The mean absolute error in each antenna configuration.

Fig. 21. The mean absolute error versus distance from interference to transceivers.

5.1.4 Impact of Various Factors. We further study the impact of various factors on system performance:

(1) **The number of antennas.** In each environment above, we test the system with different antenna configurations, namely 1×3 AC, 2×2 AC, 2×3 AC, 3×3 AC. Fig. 20 presents the mean absolute error with different numbers of antennas at a distance of 10 m from the target to the transceivers. When 1×3 AC is adopted, the mean absolute error is larger than 0.5 bpm. We can observe that when more antennas are available, a lower respiration rate error is achieved. This is because more antennas can help strengthen the weak target-reflected signal more. The proposed system still performs well even the target is 10 m away from the transceivers with 3×3 AC.

(2) **The environment changes.** To see the impact of environment changes, we move the furniture randomly to change the multipath. Specifically, in the experiment, we fix the locations of the transceivers and move the refrigerator, which is an ideal signal reflector and it introduces a significant static component. We
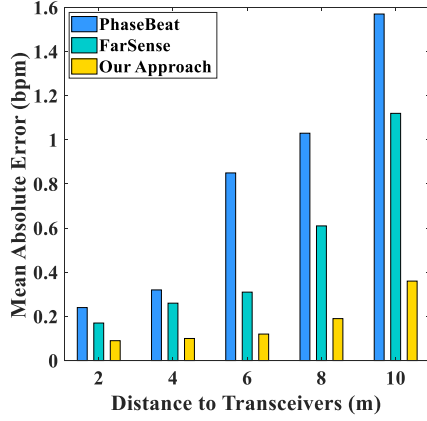
Fig. 22. The mean absolute error of each approach versus target-transceiver distance.
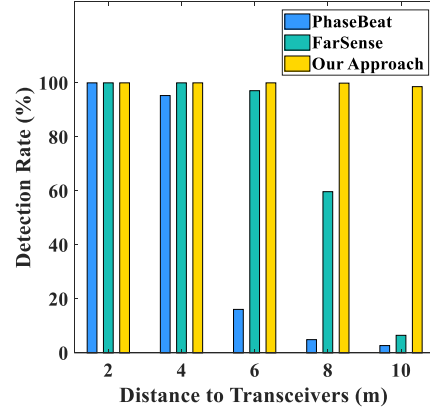

Fig. 23. The overall detection rate of each approach versus target-transceiver distance.

change the environment by moving the refrigerator to 10 different positions. At each refrigerator position, we collect 100 CSI measurements when the target is located 6-8 m away from the transceivers. The results show that the detection rates are 100% for all the environment changes. We can see that the environment changes actually do not affect the system performance.

(3) **The surrounding interference.** Here, we study the impact of interference from surroundings on respiration sensing. In the experiment, we ask another human subject to walk randomly, serving as the interference source. We vary the distance between the interference source and the transceivers from 4 m to 12 m at a step size of 2 m while a target breathes naturally, who is 4 m away from the transceivers. Fig. 21 plots the mean absolute error of respiration rate with respect to the distance between the interference and the transceivers. The experiment results show that if the interference is close to transceivers (less than 8 m), the system performance can be severely degraded. When the interference is far away from the transceivers, the system can still accurately sense the target respiration. The surrounding interference has a small influence on the proposed method when the interference is far away from the transceivers. The reason is two-fold. First, when the interference source is far away, the interference signal is weaker than the target-reflected signal. Second, the proposed beamforming method finds the optimal multi-antenna combination weight to strengthen the target-reflected signal while the interference signal is randomly combined.

5.1.5 Comparison with State-of-the-art Approaches. We compare the proposed system with two following state-of-the-art systems in terms of single-person respiration sensing accuracy:

- **PhaseBeat.** PhaseBeat [27] exploits the phase difference for respiration sensing. It first applies the DWT (Discrete Wavelet Transform) scheme to remove high-frequency noise, and then employs the peak detection method to estimate the respiration rate.
- **FarSense.** FarSense [39] exploits the ratio of CSI signals between two antennas for sensing. It first applies Savitzky-Golay filter to smooth the data, and then extracts the respiration pattern by combining the amplitude and phase information of the CSI ratio. Next, it fuses the autocorrelation result of multiple sub-carriers to estimate the respiration rate.

We plot the mean absolute respiration rate errors of the three systems at different distances in Fig. 22. We can see that when the target is close to the transceivers, all the systems work relatively well while our system achieves the lowest error. At the distance of 6 m, the mean absolute errors for PhaseBeat and FarSense increase
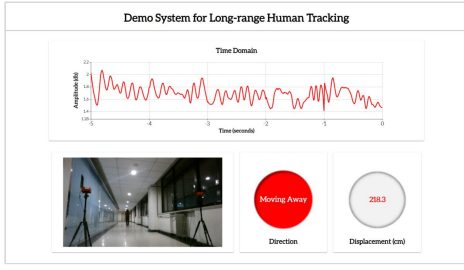
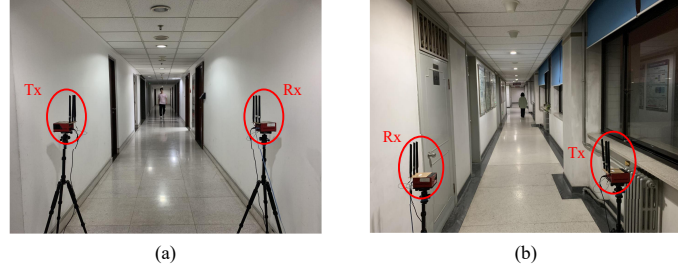Fig. 24. The web-based GUI for long-range displacement estimation.

(a) (b)

Fig. 25. The experiment setup in two long corridors: (a) the length is 30 m; (b) the length is 35 m.

significantly to 0.85 bpm and 0.31 bpm while our system can still achieve a low error rate of 0.12 bpm. At a large distance of 10 m, the errors for PhaseBeat and FarSense are larger than 1 bpm. Note that an error rate lower than 1 bpm is the common requirement for respiration sensing systems. In contract, our system still achieves an error less than 0.4 bpm.

Fig. 23 presents the overall detection rates of the three approaches when the target is at different distances with respect to the transceivers. The figure shows that our approach achieves a high detection rate of 98.6% even when the target is 10 m away from the transceivers, while the detection rates of PhaseBeat and FarSense drop to 2.7% and 6.5%, respectively. Obviously, at a greater distance, both PhaseBeat and FarSense can hardly detect the human respiration, leading to very low detection rates. Based on the definition of sensing range, we get the sensing range for PhaseBeat (4-6 m), FarSense (6-8 m) and our system (larger than 10 m). The experiment results demonstrate the effectiveness of our system in terms of increasing the sensing range without sacrificing the accuracy.

5.2 Long-range 1D Human Tracking

In this subsection, we first present the implementation of a real-time 1D human tracking system. Here, we deploy one pair of transceivers to perform 1D displacement measurement of human walking. We then detail the data collection and the evaluation metrics. After that, we validate the effectiveness of human tracking in the long corridors. We also discuss the impact of various factors, including the number of antennas, the sampling rate and the walking speed. At last, we compare the system performance with the state-of-the-art system IndoTrack [9].

5.2.1 System Implementation. We first adopt the motion detection method proposed in [43] to determine whether there is human walking, and further segment the walking activity from the time-series of CSI signals. For each walking segment, we process the CSI data with a 0.3-second sliding window in real time. Then, for each sub-carrier, the proposed signal processing methods are adopted to strengthen target-reflected signal and reduce noise. Next, we use a Savitzky-Golay filter to smooth the complex-valued CSI and apply Principal Component Analysis (PCA) method on the smoothed CSI data over different sub-carriers. At last, we track the radian change of the circular arc that the CSI signals generate in the I-Q plane. Based on Eq. 4, we extract the walking direction and the RPL change. Specifically, a positive radian change indicates the target is moving towards the transceivers, and a negative radian change indicates the target is moving away. As the LoS path length is short and the target is far away from the transceivers, the target displacement can be approximated as $\Delta d/2$, where Δd is the RPL change. The ground truth of the displacement is measured with a tapeline placed along the perpendicular bisector of the transmitter and receiver. As shown in Fig. 24, we develop a GUI to visualize human walking in real time.
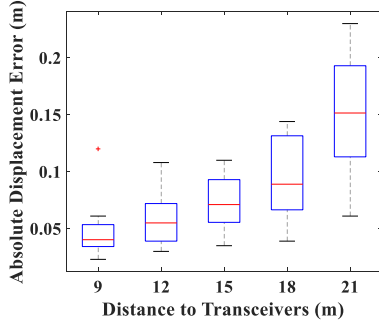
Fig. 26. The absolute displacement error versus target-transceiver distance.
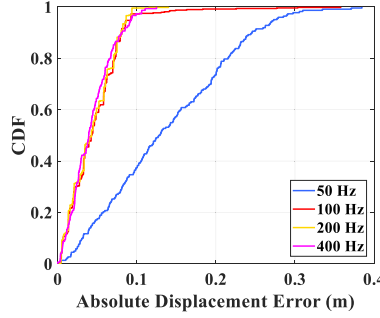

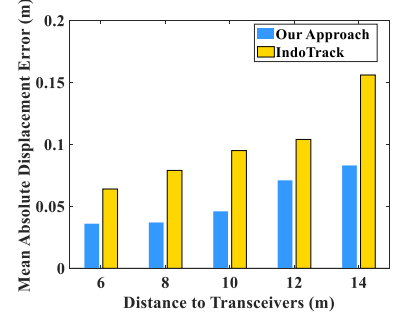Fig. 27. The CDF of absolute displacement error versus sampling rate.

Fig. 28. The comparison with previous work: the mean absolute displacement error versus distance to transceivers.

5.2.2 Data Set and Evaluation Metrics. In the experiment, we recruit 12 volunteers aged between 18 and 45, with a height range of 155 cm to 184 cm, and a weight range of 43 kg to 106 kg. We first evaluate the system performance at different distances. As shown in Fig. 25, the distance between the subject and the transceivers is varied from 5 m to 25 m at a step size of 1 m. At each distance, we ask the volunteer to move for 1 m, including moving towards and away from the transceivers.

We evaluate the system performance from three perspectives: absolute displacement error, direction accuracy and computational cost. The absolute displacement error is defined as $|d_{est} - d_{gt}|$, where d_{est} is the estimated target displacement and d_{gt} is the ground truth. Here, we define the sensing range of a human tracking system as the maximal distance where the mean absolute displacement error is less than 0.1 m (given the ground truth target displacement is 1 m). The direction accuracy is calculated as N_{est}/N, where N_{est} is the number of CSI measurements whose estimated direction matches the ground truth and N is the total number of collected CSI measurements. Note that, as we perform 1D displacement estimation, there are only two possible moving directions: (1) moving towards the transceivers; and (2) moving away from the transceivers. The computational cost refers to the processing time of the system, that is, the time interval between receiving the CSI data and outputting the displacement/direction estimates.

5.2.3 Performance of 1D Displacement Estimation. We evaluate the displacement estimation accuracy with varying distances.

(1) **Absolute displacement error.** Fig. 26 shows the absolute displacement errors at different distances. At the distance of 9 m, the median error is just 4.0 cm. The median error is as small as 8.9 cm even when the target is 18 m away from the transceivers. We can observe that the absolute displacement error increases when the distance from target to transceivers increases. The median error increases to 15.6 cm at the distance of 21 m, which is still relatively accurate.

(2) **Direction accuracy.** The experiment results show that for all volunteers, our system can reliably detect the walking direction at an accuracy of 99.3% even at the distance of 18 m. The computational cost is as small as 0.6 second. The direction accuracy is as high as 100% when the target-transceiver is less than 16 m. This demonstrates that our system can accurately track human walking from far away.

Table 3. The Computation Time and Speed-up of the Two Algorithms

Sampling Rate	Samples	Doppler-MUSIC	Our Algorithm	Speed-up
100 Hz	1500	2.31 s	0.44 s	5.3×
200 Hz	3000	4.85 s	0.83 s	5.8×
400 Hz	6000	9.74 s	1.34 s	7.3×

5.2.4 Impacts of Various Factors. We now study the impact of various factors on the displacement estimation performance.

(1) **The number of antennas.** At the distance of 12-18 m, we test four different antenna configurations, namely 1×3 AC, 2×2 AC, 2×3 AC, 3×3 AC. The results show that the lowest mean absolute displacement error and the highest direction accuracy are achieved when 3×3 AC is adopted. This is consistent with the result of long-range respiration sensing. We skip the figures here for conciseness.

(2) **The sampling rate.** At the distance of 10 m, we test the system performance under four different sampling rates, namely 50 Hz, 100 Hz, 200 Hz and 400 Hz. The absolute displacement errors with different sampling rates are shown in Fig. 27. We see no obvious difference when the sampling rate is greater than 100 Hz. The absolute displacement error is quite low when the sampling rate is greater than 100 Hz, but becomes much larger when the sampling rate is reduced to 50 Hz. The main reason is that this sampling rate is no longer sufficient to accurately characterize the signal fluctuation induced by walking. According to the Fresnel zone model [25], the frequency of the CSI signal variation induced by human walking is between 10 Hz and 40 Hz. According to the Nyquist sampling theorem, a sampling rate of 80 Hz is required to fully capture this walking information. In practice, we set the sampling rate as 200 Hz to cover a wider walking speed range.

(3) **The walking speed.** We evaluate the system performance under three different walking speeds: (1) slow speed (less than 1 m/s); (2) normal speed (1-1.5 m/s); and (3) fast speed (1.5-3 m/s). From the experiment results, we do not observe obvious differences in the absolute displacement errors among all three speeds. Therefore, our system is able to track human at different walking speeds.

5.2.5 Comparison with the State-of-the-art Approaches. We now compare our system with the state-of-the-art human tracking system IndoTrack [9]. This work utilizes the conjugate product of CSI signals received at two adjacent antennas and proposes the Doppler-MUSIC algorithm to extract Doppler speed information of a moving target from the conjugate product.

(1) **Absolute displacement error.** Fig. 28 presents the mean absolute displacement errors of the two systems with different target-transceiver distances. As can be seen from the figure, our system performs better than IndoTrack at all the distances. Based on the definition of sensing range, our system achieves a sensing range of 18 m, outperforming that achieved by IndoTrack (10-12 m). With the processed CSI that has a higher sensing capability than the conjugate product, we increase the sensing range by more than 50%.

(2) **Computational cost.** The algorithm in our system is more computationally efficient than the Doppler-MUSIC algorithm and is capable of running in real time. We use a 15-second CSI data as input, and compare the processing time of the two algorithms at different sampling rates. We run MATLAB code of both algorithms on a DELL Precision 5520 laptop that has an Intel Xeon E3-1505M v6 CPU. The running time and acceleration are listed in Tab. 3. We can observe that the computation of our algorithm takes much less time, with roughly 7× speed-up compared with the Doppler-MUSIC algorithm.
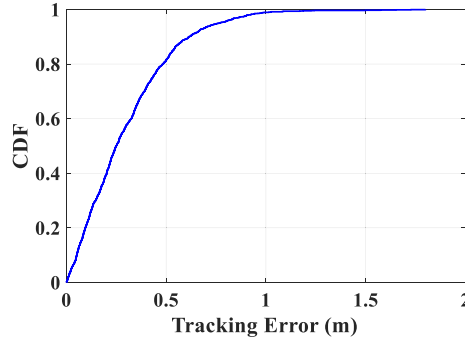
Fig. 29. The CDF of overall tracking error.

5.3 Long-range 2D Human Tracking

In this subsection, we first present the implementation of a real-time 2D human tracking system. We then detail the data collection and the evaluation metrics. After that, we validate the effectiveness of human tracking in indoor environments. We also discuss the impact of various factors including different environments, different trajectories and varying walking speeds.

5.3.1 System Implementation. We deploy one WiFi transmitter and two receivers to perform 2D human tracking. In the experiment, these three devices are fixed at the three vertexes of a rectangular sensing area. At each WiFi receiver, we calculate the corresponding reflection path length (RPL) change as illustrated in Sec. 5.2.1. We use the method proposed in [9] to estimate the initial location of the target. Based on the target's initial location and the RPL changes at the two receivers, we estimate the target's locations and accordingly trajectories.

5.3.2 Data Set and Evaluation Metrics. We evaluate the tracking performance in two indoor environments: one is an empty room with a size of 11.7 m × 10.8 m and the other is a meeting room with a size of 8 m × 6.5 m with furniture and electrical appliances. We obtain the groundtruth of the walking trajectory by placing markers on the floor and using three video cameras to record the walking process. We design trajectories with various lengths, directions and shapes (such as line, curve, rectangular, circle and fold line). We recruit 12 volunteers to walk along the trajectories over a period of two weeks to evaluate the tracking performance. We collect a total of 300 trajectories for each volunteer. We report the tracking error to show the performance of our system. For each trajectory, the error is calculated as the distance between the obtained target location and the groundtruth.

5.3.3 Performance of 2D Human Tracking. Fig. 29 shows the CDF of target tracking error across all trajectories. We can see that the median tracking error is 0.25 m while the 90th percentile error is 0.62 m. Compared with the state-of-the-art IndoTrack [9], which achieves a median tracking error of 0.35 m, our system achieves a higher accuracy for 2D human tracking.

5.3.4 Impact of Various Factors. We further study the impact of various factors on human tracking.

(1) **The environment.** In the empty room, the achieved median error is 0.22 m while in the meeting room, the median error is 0.28 m. Even though there are more objects in the meeting room creating richer multipath, our system achieves a similar tracking performance. These results demonstrate that the static multipath has little impact on our system's performance.

(2) **The trajectory diversity.** We choose four typical shapes of trajectories to show the impact of trajectory diversity: two Z-shaped paths (vertical & horizontal), rectangle and circle. The median tracking error of

the four trajectories are 0.27 m, 0.31 m, 0.19 m and 0.20 m respectively. The results show that the proposed system performs well under different shapes of trajectories.

(3) **The walking speed.** We also conduct experiments to evaluate the influence of walking speed. We evaluate three different levels of speeds: (1) slow speed (less than 1 m/s); (2) normal speed (1-1.5 m/s); and (3) fast speed (1.5-3 m/s). The median tracking error for the three walking speeds are 0.24 m, 0.26 m and 0.30 m respectively. We observe a slightly increasing error with the speed increase. Overall, the proposed system achieves consistent high tracking accuracies under different human walking speeds.

6 RELATED WORK

In recent years, various contactless human sensing techniques have been developed, such as camera-based [18, 40], acoustic-based [21, 28] and radar-based approaches [4, 7, 36]. Camera-based approaches do not work well in poor lighting conditions and may raise privacy concerns. Acoustic-based approaches have a small sensing range (1-2 meters) and they are vulnerable to the interferences from environment (e.g., music in the gym and traffic sound). Radar-based approaches can achieve a relatively high sensing accuracy, but they require dedicated sensing devices which are quite costly. Different from these approaches, WiFi-based approaches utilize the devices that are ubiquitous in indoor environments, such as WiFi router, television and refrigerator. In this section, we survey the most relevant literature on contactless human sensing with WiFi CSI. According to the base signal utilized for sensing, it can be divided into five categories as follows.

- **CSI amplitude.** CSI amplitude is widely used in WiFi-based human sensing applications [11, 14, 31, 34?], even though it contains relatively large noise. Wang *et al.* [26] proposed the CSI-speed model to quantify the correlation between CSI amplitude and human movement speed. The Fresnel zone model was introduced in [25] to explain how the user location and orientation affect the respiration sensing performance. An interesting observation is that if we use only CSI amplitude for sensing, the performance will vary with the static signal, as shown in Fig. 4.
- **CSI phase.** Due to the presence of CFO and SFO, the phase information of CSI signal obtained from commodity WiFi is especially noisy, which cannot be directly used for sensing. Efforts have been made in [6, 35, 46] to calibrate the CFO and SFO. However, as shown in [32, 37], these methods do not work for fine-grained sensing, such as the millimeter-level respiration monitoring.
- **CSI phase difference.** While the raw phase information itself is not directly usable due to the phase offset, the phase difference between two antennas is quite stable for sensing, since the phase offset keeps the same across two antennas at the same receiver [6]. Compared to the raw CSI amplitude information, the phase difference has a higher sensing capability. This information has been successfully applied to enable various applications, such as intrusion detection [8] and respiration sensing [27].
- **CSI conjugate product.** While previous works use either amplitude or phase information for sensing, recent works [22, 37] propose to employ both amplitude and phase for sensing. WiDance [22] and IndoTrack [9] use conjugate product of CSI signals from two receiver antennas to cancel out phase offset and extract Doppler speed information of a moving human target. FullBreathe [37] demonstrates the complementarity between phase and amplitude for respiration sensing, and proposes to employ both amplitude and phase of CSI conjugate product to remove the amount of "blind spots" identified by [25]. Unfortunately, the conjugate multiplication operation amplifies the noise in amplitude, since the amplitude of CSI conjugate product equals to the product of CSI amplitudes at two antennas.
- **CSI ratio.** FarSense [39] employs the ratio of CSI signals from two antennas for respiration sensing, where the noise is mostly canceled out by the division operation to increase the sensing range. This signal ratio method is also adopted in [32] to enable fine-grained finger tracking. However, during the division process, the phase/amplitude information which is critical for sensing may get corrupted, which thus compromises

the sensing capability. Besides, FarSense does not take the best use of multiple antennas, as it just selects two antennas to participate in the division operation.

7 LIMITATION AND DISCUSSION

In this work, we focus on utilizing multiple antennas to increase WiFi sensing range in single-target scenarios. We briefly discuss the limitation and future direction of our work below.

7.1 Applicability to Other Wireless Sensing Systems

The proposed multi-antenna based method can be applied to other RF sensing systems to increase the sensing range as long as multiple antennas are available at the receiver. Similar idea has been adopted recently by LoRa sensing [42] to increase the range and we believe the proposed method can also be applied to acoustic systems equipped with multiple microphones to boost the limited sensing range.

7.2 Exploring Multiple Sub-carriers to Enhance Sensing Performance

In 802.11 a/g/n/ac WiFi standards, OFDM (Orthogonal Frequency Division Multiplexing) technique is employed to efficiently utilize the wireless channel by dividing the channel into smaller sub-carriers [10]. This enables us to obtain rich CSI readings for each sub-carrier. In this work, we utilize the spatial diversity of multiple antennas to increase the SSNR for longer sensing range. We believe similar idea that utilizes the frequency diversity could also be applied to increase the SSNR by combining CSI signals from different sub-carriers. For devices equipped with just one single antenna such as Google Nexus5 smartphone [12], this sub-carrier based method can be a feasible approach to increase the sensing range.

7.3 Multi-subject Scenarios

In a more general case, there are multiple subjects in the environment and their movements cause changes in the corresponding reflection signals. If we would like to sense one subject, the dynamic reflection signals from the other subjects that we are not interested in will be viewed as interference. In this case, we need to consider not just the noise but also the undesired interference when we define the SSNR metric. While white noise is easy to be modelled, dynamic human interference can be much more complicated, making multi-subject sensing a challenging yet exciting research direction.

8 CONCLUSION

While contactless human sensing using WiFi signals is very promising due to its non-intrusive and pervasive nature, the small sensing range hinders it from real-world adoption. In this paper, we effectively utilize multiple antennas widely available on commodity WiFi hardware to strengthen the target-reflected signal and reduce the noise at the same time, significantly increasing the WiFi sensing range. We believe the proposed methods contribute to not just WiFi sensing but also other wireless sensing modalities.

REFERENCES

[1] 2017. Respiration Monitor Belt logger sensor NUL-236. https://neulog.com/respiration-monitor-belt/.
[2] Rajendra Bhatia and Chandler Davis. 1995. A Cauchy-Schwarz inequality for operators with applications. *Linear algebra and its applications* 223 (1995), 119–129.

[3] Amy Diane Droitcour et al. 2006. *Non-contact measurement of heart and respiration rates with a single-chip microwave doppler radar.* Ph.D. Dissertation. Citeseer.

[4] Unsoo Ha, Salah Assana, and Fadel Adib. 2020. Contactless seismocardiography via deep learning radars. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking.* 1–14.

[5] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2010. 802.11 with multiple antennas for dummies. *ACM SIGCOMM Computer Communication Review* 40, 1 (2010), 19–25.

[6] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. 2015. Spotfi: Decimeter level localization using wifi. In *ACM SIGCOMM Computer Communication Review*, Vol. 45. ACM, 269–282.

[7] Changzhi Li, Julie Cummings, Jeffrey Lam, Eric Graves, and Wenhsing Wu. 2009. Radar remote monitoring of vital signs. *IEEE Microwave Magazine* 10, 1 (2009), 47–56.

[8] Shengjie Li, Xiang Li, Kai Niu, Hao Wang, Yue Zhang, and Daqing Zhang. 2017. Ar-alarm: An adaptive and robust intrusion detection system leveraging csi from commodity wi-fi. In *International conference on smart homes and health telematics.* Springer, 211–223.

[9] Xiang Li, Daqing Zhang, Qin Lv, Jie Xiong, Shengjie Li, Yue Zhang, and Hong Mei. 2017. IndoTrack: Device-free indoor human tracking with commodity Wi-Fi. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–22.

[10] Zhuqi Li, Yaxiong Xie, Longfei Shangguan, Rotman Ivan Zelaya, Jeremy Gummeson, Wenjun Hu, and Kyle Jamieson. 2019. Towards programming the radio environment with large arrays of inexpensive antennas. In *16th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 19).* 285–300.

[11] Jian Liu, Yan Wang, Yingying Chen, Jie Yang, Xu Chen, and Jerry Cheng. 2015. Tracking vital signs during sleep leveraging off-the-shelf wifi. In *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing.* ACM, 267–276.

[12] Jinyi Liu, Youwei Zeng, Tao Gu, Leye Wang, and Daqing Zhang. 2021. WiPhone: Smartphone-based Respiration Monitoring Using Ambient Reflected WiFi Signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 1 (2021), 1–19.

[13] Liang Liu, Rui Zhang, and Kee-Chaing Chua. 2014. Multi-antenna wireless powered communication with energy beamforming. *IEEE Transactions on Communications* 62, 12 (2014), 4349–4361.

[14] Xuefeng Liu, Jiannong Cao, Shaojie Tang, and Jiaqi Wen. 2014. Wi-Sleep: Contactless sleep monitoring via WiFi signals. In *Real-Time Systems Symposium (RTSS), 2014 IEEE.* IEEE, 346–355.

[15] Yongsen Ma, Gang Zhou, and Shuangquan Wang. 2019. WiFi sensing with channel state information: A survey. *ACM Computing Surveys (CSUR)* 52, 3 (2019), 1–36.

[16] Kai Niu, Fusang Zhang, Yuhang Jiang, Jie Xiong, Qin Lv, Youwei Zeng, and Daqing Zhang. 2019. WiMorse: A contactless morse code text input system using ambient WiFi signals. *IEEE Internet of Things Journal* 6, 6 (2019), 9993–10008.

[17] Kai Niu, Fusang Zhang, Jie Xiong, Xiang Li, Enze Yi, and Daqing Zhang. 2018. Boosting fine-grained activity sensing by embracing wireless multipath effects. In *Proceedings of the 14th International Conference on emerging Networking EXperiments and Technologies.* ACM, 139–151.

[18] Katsunori Ohnishi, Atsushi Kanehira, Asako Kanezaki, and Tatsuya Harada. 2016. Recognizing activities of daily living with a wrist-mounted camera. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 3103–3111.

[19] Sameera Palipana, David Rojas, Piyush Agrawal, and Dirk Pesch. 2018. FallDeFi: Ubiquitous fall detection using commodity Wi-Fi devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2018), 1–25.

[20] William H Press and Saul A Teukolsky. 1990. Savitzky-Golay smoothing filters. *Computers in Physics* 4, 6 (1990), 669–672.

[21] Kun Qian, Chenshu Wu, Fu Xiao, Yue Zheng, Yi Zhang, Zheng Yang, and Yunhao Liu. 2018. Acousticcardiogram: Monitoring heartbeats using acoustic signals on smart devices. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications.* IEEE, 1574–1582.

[22] Kun Qian, Chenshu Wu, Zimu Zhou, Yue Zheng, Zheng Yang, and Yunhao Liu. 2017. Inferring motion direction using commodity wi-fi for interactive exergames. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.* 1961–1972.

[23] Rishad Ahmed Shafik, Md Shahriar Rahman, and AHM Razibul Islam. 2006. On the extended relationships among EVM, BER and SNR as performance metrics. In *2006 International Conference on Electrical and Computer Engineering.* IEEE, 408–411.

[24] Deepak Vasisht, Swarun Kumar, and Dina Katabi. 2016. Decimeter-level localization with a single WiFi access point. In *13th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 16).* 165–178.

[25] Hao Wang, Daqing Zhang, Junyi Ma, Yasha Wang, Yuxiang Wang, Dan Wu, Tao Gu, and Bing Xie. 2016. Human respiration detection with commodity wifi devices: do user location and body orientation matter?. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing.* 25–36.

[26] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2015. Understanding and modeling of wifi signal based human activity recognition. In *Proceedings of the 21st annual international conference on mobile computing and networking.* 65–76.

[27] Xuyu Wang, Chao Yang, and Shiwen Mao. 2017. PhaseBeat: Exploiting CSI phase data for vital sign monitoring with commodity WiFi devices. In *Distributed Computing Systems (ICDCS), 2017 IEEE 37th International Conference on.* IEEE, 1230–1239.

[28] Yanwen Wang, Jiaxing Shen, and Yuanqing Zheng. 2020. Push the Limit of Acoustic Gesture Recognition. *IEEE Transactions on Mobile Computing* (2020).

[29] Darrell Whitley. 1994. A genetic algorithm tutorial. *Statistics and computing* 4, 2 (1994), 65–85.

[30] J Wise. 1955. The autocorrelation function and the spectral density function. *Biometrika* 42, 1/2 (1955), 151–159.

[31] Chenshu Wu, Feng Zhang, Yuqian Hu, and KJ Ray Liu. 2020. Gaitway: Monitoring and recognizing gait speed through the walls. *IEEE Transactions on Mobile Computing* (2020).

[32] Dan Wu, Ruiyang Gao, Youwei Zeng, Jinyi Liu, Leye Wang, Tao Gu, and Daqing Zhang. 2020. Fingerdraw: Sub-wavelength level finger motion tracking with WiFi signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 1 (2020), 1–27.

[33] Dan Wu, Daqing Zhang, Chenren Xu, Hao Wang, and Xiang Li. 2017. Device-free WiFi human sensing: From pattern-based to model-based approaches. *IEEE Communications Magazine* 55, 10 (2017), 91–97.

[34] Yang Xu, Wei Yang, Jianxin Wang, Xing Zhou, Hong Li, and Liusheng Huang. 2018. WiStep: Device-free step counting with WiFi signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2018), 1–23.

[35] Nan Yu, Wei Wang, Alex X Liu, and Lingtao Kong. 2018. Qgesture: Quantifying gesture distance and direction with wifi signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1 (2018), 1–23.

[36] Shichao Yue, Yuzhe Yang, Hao Wang, Hariharan Rahul, and Dina Katabi. 2020. Bodycompass: Monitoring sleep posture with wireless signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 2 (2020), 1–25.

[37] Youwei Zeng, Dan Wu, Ruiyang Gao, Tao Gu, and Daqing Zhang. 2018. FullBreathe: Full Human Respiration Detection Exploiting Complementarity of CSI Phase and Amplitude of WiFi Signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 148.

[38] Youwei Zeng, Dan Wu, Jie Xiong, Jinyi Liu, Zhaopeng Liu, and Daqing Zhang. 2020. MultiSense: Enabling multi-person respiration sensing with commodity wifi. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (2020), 1–29.

[39] Youwei Zeng, Dan Wu, Jie Xiong, Enze Yi, Ruiyang Gao, and Daqing Zhang. 2019. Farsense: Pushing the range limit of wifi-based respiration sensing with csi ratio of two antennas. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–26.

[40] Chenyang Zhang and Yingli Tian. 2012. RGB-D camera-based daily living activity recognition. *Journal of computer vision and image processing* 2, 4 (2012), 12.

[41] Daqing Zhang, Hao Wang, Yasha Wang, and Junyi Ma. 2015. Anti-fall: A non-intrusive and real-time fall detector leveraging CSI from commodity WiFi devices. In *International Conference on Smart Homes and Health Telematics*. Springer, 181–193.

[42] Fusang Zhang, Zhaoxin Chang, Kai Niu, Jie Xiong, Beihong Jin, Qin Lv, and Daqing Zhang. 2020. Exploring lora for long-range through-wall sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 2 (2020), 1–27.

[43] Feng Zhang, Chenshu Wu, Beibei Wang, Hung-Quoc Lai, Yi Han, and KJ Ray Liu. 2019. WiDetect: Robust motion detection with a statistical electromagnetic model. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–24.

[44] Feng Zhang, Chenshu Wu, Beibei Wang, Min Wu, Daniel Bugos, Hangfang Zhang, and KJ Ray Liu. 2019. Smars: sleep monitoring via ambient radio signals. *IEEE Transactions on Mobile Computing* 20, 1 (2019), 217–231.

[45] Fusang Zhang, Daqing Zhang, Jie Xiong, Hao Wang, Kai Niu, Beihong Jin, and Yuxiang Wang. 2018. From fresnel diffraction model to fine-grained human respiration sensing with commodity wi-fi devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1 (2018), 1–23.

[46] Jincao Zhu, Youngbin Im, Shivakant Mishra, and Sangtae Ha. 2017. Calibrating time-variant, device-specific phase noise for COTS WiFi devices. In *Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems*. 1–12.

[47] Yiwei Zhuo, Hongzi Zhu, Hua Xue, and Shan Chang. 2017. Perceiving accurate CSI phases with commodity WiFi devices. In *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*. IEEE, 1–9.

A APPENDIX

In this appendix, we derive the SSNR of combined CSI signal. For the numerator in Eq. 11, we have:

$$\mathbb{E}(|\mathbf{W}^T\mathbf{A}(f)e^{-j2\pi f\frac{d_1(t)}{c}}|^2) = |\mathbf{W}^T\mathbf{A}(f)|^2\mathbb{E}(|e^{-j2\pi f\frac{d_1(t)}{c}}|^2) = |\mathbf{W}^T\mathbf{A}(f)|^2 \tag{23}$$

For the denominator in Eq. 11, we have:

$$\begin{aligned}\mathbb{E}(|\mathbf{W}^T\boldsymbol{\varepsilon}(f,t)|^2) &= \mathbb{E}((\mathbf{W}^T\boldsymbol{\varepsilon}(f,t))(\mathbf{W}^T\boldsymbol{\varepsilon}(f,t))^*)\\ &= \mathbb{E}((w_1\varepsilon_1(f,t)... + w_M\varepsilon_M(f,t))(w_1^*\varepsilon_1^*(f,t) + ... + w_M^*\varepsilon_M^*(f,t)))\\ &= \mathbb{E}(\sum_{i=1}^{M}|w_i|^2|\varepsilon_i(f,t)|^2 + \sum_{i=1}^{M}\sum_{j=1,j\neq i}^{M}w_iw_j^*\varepsilon_i(f,t)\varepsilon_j^*(f,t))\end{aligned} \tag{24}$$

where the operation $(\cdot)^*$ indicates the complex conjugate. Given that $\boldsymbol{\varepsilon}(f, t)$ is the identically and independently distributed AWGN, we have:

$$\mathbb{E}(|\varepsilon_i(f, t)|^2) = \sigma^2(f) \tag{25}$$

$$\forall i \neq j, \mathbb{E}(\varepsilon_i(f, t)\varepsilon_j^*(f, t)) = 0 \tag{26}$$

Then we can simplify Eq. 24 as:

$$\mathbb{E}(|\mathbf{W}^T\boldsymbol{\varepsilon}(f, t)|^2) = \sigma^2(f) \sum_{i=1}^{M} |w_i|^2 = \sigma^2(f) \tag{27}$$

Based on Eq. 23 and Eq. 27, we have:

$$\gamma_{comb} = \frac{|\mathbf{W}^T\mathbf{A}(f)|^2}{\sigma^2(f)} \tag{28}$$