MultiSense: Enabling Multi-person Respiration Sensing with Commodity WiFi

Youwei Zeng, Dan Wu, Jie Xiong, Jinyi Liu, Zhaopeng Liu, Daqing Zhang

MultiSense: Enabling Multi-person Respiration Sensing with Commodity WiFi

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

In recent years, we have seen efforts made to simultaneously monitor the respiration of multiple persons based on the channel state information (CSI) retrieved from commodity WiFi devices. Existing approaches mainly rely on spectral analysis of the CSI amplitude to obtain respiration rate information, leading to multiple limitations: (1) spectral analysis works when multiple persons exhibit dramatically different respiration rates, however, it fails to resolve similar rates; (2) spectral analysis can only obtain the average respiration rate over a period of time, and it is unable to capture the detailed rate change over time; (3) they fail to sense the respiration when a target is located at the "blind spots" even the target is close to the sensing devices.

To overcome these limitations, we propose MultiSense, the first WiFi-based system that can robustly and continuously sense the detailed respiration patterns of multiple persons even they have very similar respiration rates and are physically closely located. The key insight of our solution is that the commodity WiFi hardware nowadays is usually equipped with multiple antennas. Thus, each individual antenna can receive a different mix copy of signals reflected from multiple persons. We successfully prove that the reflected signals are linearly mixed at each antenna and propose to model the multi-person respiration sensing as a blind source separation (BSS) problem. Then, we solve it using independent component analysis (ICA) to separate the mixed signal and obtain the reparation information of each person. Extensive experiments show that with only one pair of transceivers, each equipped with three antennas, MultiSense is able to accurately monitor respiration even in the presence of four persons, with the mean absolute respiration rate error of 0.73 bpm (breaths per minute).

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

Additional Key Words and Phrases: Multi-person Respiration Sensing, WiFi Sensing, Channel State Information (CSI), Blind Source Separation

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; 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; Jie Xiong, College of Information and Computer Sciences, University of Massachusetts, Amherst, Amherst, USA, jxiong@cs.umass.edu; 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; 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; 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.

2474-9567/2020/9-ART102 $15.00
https://doi.org/10.1145/3411816

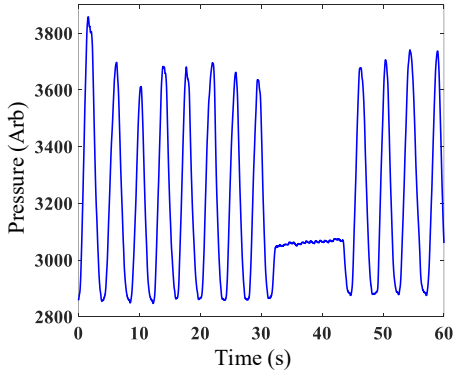

Fig. 1. A 60-second respiration pattern of a person, which is recorded by a respiration belt.

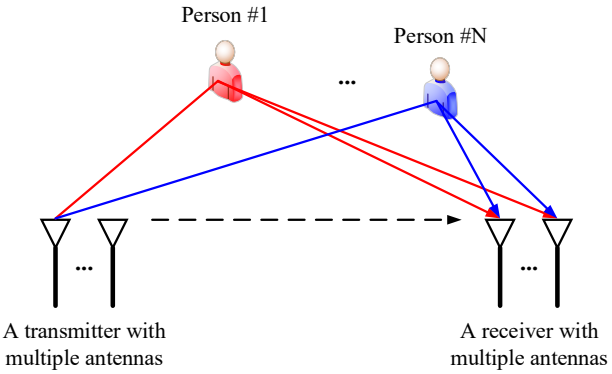

Fig. 2. A typical multi-person respiration sensing scenario of MultiSense, where both the transmitter and the receiver are equipped with multiple antennas.

ACM Reference Format:
Youwei Zeng, Dan Wu, Jie Xiong, Jinyi Liu, Zhaopeng Liu, and Daqing Zhang. 2020. MultiSense: Enabling Multi-person Respiration Sensing with Commodity WiFi. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 3, Article 102 (September 2020), 29 pages. https://doi.org/10.1145/3411816

1 INTRODUCTION

Accurate detection and monitoring of human respiration plays a crucial role in a wide range of applications in health care, such as sleep apnea detection [26] and sleep stage recognition [55]. The principle behind these applications is that the rich information extracted from human respiration can provide useful insights about the physiological state of an individual [15]. Two pieces of respiration-related information are widely used. One is the respiration pattern in time domain, which characterizes the detailed process of inhalation and exhalation over time [7]. Fig. 1 shows a 60-second respiration pattern measured by a commercial device (Neulog Respiration Monitor Belt logger sensor NUL-236 [12]) when a person breathes. We can tell from the figure that an apnea event occurs in the time interval from the 33rd to 43rd second. Another important information is the respiration rate, which is the frequency of breaths over a period of time, and it can be extracted from the respiration pattern. For example, by counting the number of peaks in Fig. 1, the respiration rate can be obtained as 12 bpm. Note that, if we apply the spectral analysis to the pattern, such as FFT (Fast Fourier Transform), we will get a wrong result of 15 bpm and also fail to capture the abnormal respiration in Fig. 1. Obviously, detailed respiration pattern over time can provide finer-grained information about the respiration process than just the respiration rate.

A wealth of literature exists in the field of human respiration detection and monitoring. Traditional solutions [25] usually require users to carry a device, which may cause discomfort and even affect the underlying physiological parameters being measured [39]. Other solutions [24] leverage cameras, but they do not work in non-line-of-sight (NLoS) scenario and also raise privacy concerns. The acoustic-based solutions [34] can achieve a high respiration rate accuracy, however, they are vulnerable against environmental noise and have a limited sensing range (around 1 m). RF-based solutions such as RFID system [18, 46] and FMCW radar [2, 49] need dedicated hardware, which are relatively expensive. Recently, WiFi-based sensing has gained lots of research interests, as WiFi has already been widely deployed in indoor environments and the CSI readings retrieved from commodity WiFi device can provide rich sensing information. With the availability of WiFi CSI, significant progress has been made in various

applications ranging from human activity recognition [35, 45], indoor localization [13, 27, 41] to respiration monitoring [4, 19–21, 33, 36, 37, 40, 48, 51–53, 55, 56].

Most previous work on WiFi-based respiration monitoring is dedicated to the single-person scenario, and it fails to extract respiration pattern in multi-person scenario, where the CSI is simultaneously affected by the chest movements of multiple persons. Existing approaches that monitor multi-person respiration mainly rely on spectral analysis of the CSI amplitude to extract multi-person respiration rates. However, according to our empirical study in Sec. 3, there are multiple obvious limitations associated with the existing approaches: (1) they fail to sense the respiration of a person located at the "blind spots"; (2) they fail to resolve the respiration rates when the targeted subjects have similar rates; (3) the spectrum-based approaches can only obtain the average respiration rate but not the fine-grained respiration pattern over time which is imperative to detect abnormal respiration (e.g., the apnea in Fig. 1). All the limitations make these multi-target sensing systems less practical and sometimes unuseful in real-life applications. Note that, the "blind spots" issue for single-person scenario has been addressed in [51, 52], however, the solution can not be directly applied in multi-person scenario as the underlying theory that enables elimination of "blind spots" is based on the assumption that there is only one breathing person. And no solution has been yet proposed for multi-person sensing.

In this paper, we introduce MultiSense, the first commodity WiFi based respiration sensing system capable of simultaneously extracting respiration patterns of multiple persons. The key enabler for MultiSense is the multiple antennas widely available nowadays at commodity WiFi devices. Through a theoretical analysis of how multiple persons' respiration affects the CSI readings, we prove that signals reflected from multiple persons are linearly mixed at each receiver antenna. The problem can thus be modeled as the well-known blind source separation (BSS) problem, which aims to recover the individual source signals from a given mixed signal without knowing how they are mixed together [1]. However, there are two challenges that prevent us from directly borrowing the solution from BSS: (1) time-varying phase offset exists in CSI readings due to time unsynchronization between the physically separated transmitter and receiver; (2) the mixed signal at the receiver is composed of not just signals reflected from the targets but also signals reflected from non-target static objects such as walls and direct path signal between the transceivers. The non-target background signals need to be removed.

To address the issues, we first propose a novel signal ratio method to cancel out the time-varying phase offset without distorting the linearity. We further measure the background CSI in advance when there is no target. Later when there are targets, the background CSI which corresponds to the signals reflected from non-target static objects is removed and only the components corresponding to the targets are kept. With the carefully processed WiFi CSI from multiple antennas, we can then model the multi-person respiration sensing as a BSS problem and solve it using the independent component analysis (ICA) method.

We design and implement MultiSense on commodity WiFi hardware to continuously recover respiration patterns of multiple persons. Fig. 2 presents a typical multi-person respiration sensing scenario of MultiSense, where both the transmitter and the receiver are equipped with multiple antennas. However, we want to point out that MultiSense does not require both the transmitter and receiver to have multiple antennas. As long as the receiver side is equipped with multiple antennas, the proposed method has no problem to work and nowadays it is very common for the WiFi access point to be equipped with multiple antennas. For example, to monitor two-person respiration, we can equip the receiver with three antennas and the transmitter with only one antenna. To validate the effectiveness and robustness of the proposed system, we conduct comprehensive experiments with different subjects in different environments. The underlying principle of our approach is quite different from that of the traditional spectrum-based approaches. Leveraging the BSS scheme, our approach can efficiently separate and recover respiration signals reflected from multiple persons as long as these signals are independent of each other, which has been validated through a set of experiments in [49].

The main contributions of the work can be summarized as follows:

- We propose a novel method to cancel out the time-varying phase offset without distorting the linear mixture of WiFi CSI. We believe this general method will benefit not just respiration sensing but a lot of other sensing applications.
- Based on both rigorous theoretical analysis and real-life experiments, we demonstrate that multi-person respiration sensing can be modeled as a BSS problem, which can be efficiently solved by the ICA method. Such idea will provide researchers a new perspective for contactless multi-person sensing.
- We implement MultiSense on commodity WiFi devices and evaluate it through extensive experiments. For the first time, MultiSense is able to extract respiration patterns with only a pair of transceivers even in the presence of four persons, outperforming the state-of-the-art approaches.

The remaining of the paper is organized as follows. Sec. 2 survey the state-of-the-arts. Sec. 3 employs experiments to visualize the limitations of existing approaches. Sec. 4 presents how the literature on BSS is utilized to extract respiration pattern of each person from WiFi CSI. Sec. 5 presents the detailed design and implementation of the prototype system MultiSense. Sec. 6 presents the experimental setup and evaluation results. Sec. 7 discusses the limitations and opportunities followed by a conclusion in Sec. 8.

2 RELATED WORK

Our work is related to the past work on RF-based respiration sensing, which can be grouped into the following two groups.

Radar-based Respiration Sensing. FMCW radar emits signals modulated with a linear frequency occupying a wide channel band (1-2 GHz) [31]. A basic feature of FMCW radar is the ability to measure the distance of the reflector from the device. In Vital-Radio [2], FMCW radar is employed to separate signals reflected from different persons based on the signal's propagation time, thus enabling multi-person respiration sensing. However, it fails when multiple persons are closely located since the time differences of the reflected signals are too small and it is extremely challenging to separate them. DeepBreath [49] demonstrates that the respiration signals of multiple individuals can be recovered from the received RF signals via ICA method. However, dedicated expensive hardware makes it less practical for everyday home usage. Also the large bandwidth (1.5 GHz in [49]) is not available on commodity WiFi hardware. In this work, we would like to employ cheap commodity WiFi hardware (with a limited bandwidth of 20/40 MHz) already pervasively deployed in the home environment for multi-person respiration sensing.

Besides the FMCW radar, other approaches include continuous-wave (CW) Doppler radar and ultra-wideband (UWB) pulse radar. CW Doppler radar emits a signal with a constant frequency and amplitude, and processes the received echo signal that is reflected off a moving target [8]. CW radar leans on a relatively simple design. However, the performance is easily affected by environmental noise and multipath reflections. UWB pulse radar emits short pulses, and receives the echo signals in the silent period [32]. In contrast to the CW Doppler radar, the large 1-2 GHz bandwidth enables it to eliminate interference and multipath [16]. However, such a wide bandwidth requires a precise control of pulse width and radar peak signal strength [14], which unavoidably increases the hardware complexity and cost.

WiFi-based Respiration Sensing. The WiFi-based approaches can be broadly grouped into two categories based on whether they are able to monitor respiration in the presence of multiple persons.

(1) Single-person respiration sensing. Wi-Sleep [20] is among the first few proposed sleep monitoring system that can extract rhythmic respiration patterns from WiFi signals in single-person scenario. The authors further improved the performance in [21] where both sleeping postures and abnormal respiration patterns are taken into consideration. These systems are all based on one observation that the CSI amplitude in time domain shows a clear sinusoidal-like pattern, corresponding to the single-person's respiration process. Wang *et al.* [33] introduce the Fresnel zone theory to explain why the "blind spots" issue occurs

for single-person respiration sensing. FullBreathe [51] demonstrates the complementarity between phase and amplitude for respiration sensing and propose to employ both amplitude and phase to remove the amount of "blind spots". The latest work – FarSense [52] pushes the respiration sensing range from room level (2-4 meters) to house level (8-9 meters) by utilizing a new metric called CSI ratio, which is defined as the ratio of CSI readings collected from two antennas hosted on one WiFi hardware. Solutions have been proposed in prior work [51, 52] to address the "blind spots" issue in single-person scenario. However, the proposed methods cannot be directly applied in multi-person scenario as the underlying theory that enables elimination of "blind spots" is based on the assumption that there is only one breathing person.

(2) Multi-person respiration sensing. In multi-person scenario, the CSI measurement is affected by multiple independent chest movements simultaneously. It is thus nontrivial to extract the respiration pattern of each individual from the entangled CSI. Inspired by the observation that the frequency of respiration coming from two persons may be still preserved if the CSI amplitude is transferred to frequency domain, Liu *et al.* [19] present the first attempt to estimate respiration rate in two-person scenario by analyzing the power spectral density (PSD) of CSI amplitude in frequency domain. However, based on the Fresnel zone theory, a breathing person at a certain location may have little effect on CSI amplitude. Wang *et al.* [33] point out that the "blind spots" issue greatly affects the performance of the method proposed in [19]. While these systems leverage CSI amplitude for respiration sensing, PhaseBeat [36] demonstrates that multi-person breathing rates can be estimated by applying root-MUSIC algorithm [28] to the CSI phase difference between two antennas. Based on the assumption that the observed CSI phase difference is the sum of multiple sinusoidal respiration patterns, TensorBeat [37] leverages the tensor decomposition technique to extract the respiration patterns of multiple persons from the CSI phase difference data. TR-BREATH [4] first projects WiFi CSI into the TRRS (Time-Reversal Resonating Strength) feature space and then applies root-MUSIC algorithm to estimate the respiration rates of multiple targets. These methods assume that the respiration rates of different people are distinct, and the performance degrades when different people have similar respiration rates. Furthermore, most of them are based on the spectral analysis, thus they can only obtain the average respiration rate but are not able to extract the detail respiration pattern over time. Inspired by the Fresnel zone theory [33, 53], Yang *et al.* [48] manage to monitor multi-person respiration by optimizing the deployment of WiFi transceivers so that each transceiver pair's transmission is only affected by one target. However, such method requires to know the accurate location of each person in advance and as long as one person changes the location, the system may not work.

Differently, with theoretical analysis and carefully processed WiFi CSI from multiple antennas, we successfully model the multi-person respiration sensing as a BSS problem. MultiSense is able to robustly and continuously extract not just the respiration rate but the detailed respiration patterns over time for multiple persons without a need of specific WiFi hardware deployment.

3 ILLUSTRATIVE EXAMPLE

In this section, to visualize the limitations of traditional approaches, we conduct a simple experiment in a multi-person scenario, where three breathing persons sit on the couch. The groundtruths of their respiration rates are {15 bpm, 18 bpm, 20 bpm}, each of which is collected with a commercial device (Neulog Respiration Monitor Belt logger sensor NUL-236 [12]). Here, we take the traditional approach presented in [19, 33] for example. First, the hampel filter and a moving average filter are applied to remove the outliers and high-frequency noise in the CSI amplitude. Second, FFT (Fast Fourier Transform) scheme is applied to convert the time-domain CSI amplitude to frequency domain. According to [19], for three persons, the spectrum is expected to show three strong peaks in frequency domain corresponding to the respiration rates of three persons.

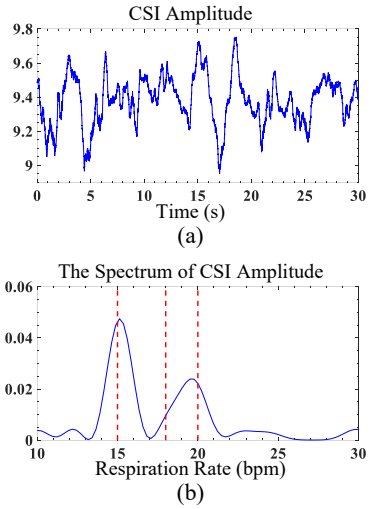

Fig. 3. The "blind spots" issue exists when using only CSI amplitude for multi-person respiration sensing.

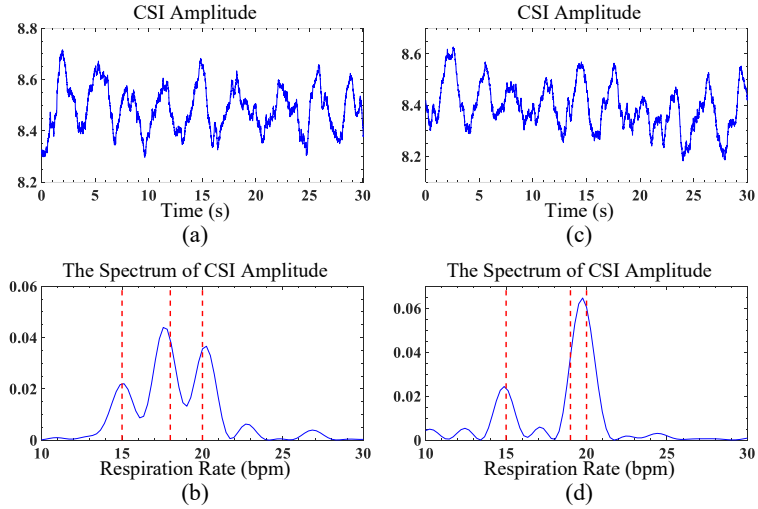

Fig. 4. The traditional approach that relies on spectral analysis on CSI amplitude loses the ability to resolve two close-by respiration rates {19 bpm, 20 bpm} while none of the multiple persons is at the "blind spots". The red dashed lines label the ground-truth respiration rates.

Limitation 1. Fig. 3 shows one result when the three persons breathe normally. Fig. 3 (a) is the CSI amplitude in time domain and Fig. 3 (b) is the corresponding plot in frequency domain. Note that we multiply the raw frequency value by a factor of 60 to convert the unit from hertz (Hz) to breaths per minute (bpm). The three ground-truth respiration rates are denoted as dashed vertical lines in the figures. From Fig. 3 (b), we can observe two strong peaks corresponding to the ground-truth respiration rates {15 bpm, 20 bpm}. However, we fail to detect respiration of the last person whose respiration rate is 18 bpm. This person is actually located at the so-called "blind spots" where his respiration can not be effectively sensed by using only CSI amplitude. This phenomenon can be explained by the Fresnel zone theory [33, 51, 53], which proves that a person located at the "blind spots" has very little effect on CSI amplitude and thus can hardly be detected.

Limitation 2. With this "blind spots" issue in mind, we ask the person whose respiration rate is 18 bpm to move to a location which is not a "blind spot" and breathe naturally. Fig. 4 (a) and Fig. 4 (b) show one of the experimental results. In Fig. 4 (b), we can observe three strong peaks that correspond to the three ground-truth respiration rates. Then, we ask the third person to slightly increase his respiration rate to 19 bpm while keeping all three persons' locations unchanged. The results are presented in Fig. 4 (c) and Fig. 4 (d). As shown in Fig. 4 (d), the two peaks corresponding to the ground-truth respiration rates {19 bpm, 20 bpm} fully merge into one peak. Thus, we can see that the traditional approach fails to resolve two respiration rates when they are too close to each other.

We obtain similar results with the method proposed in PhaseBeat [36], which adopts root-MUSIC algorithm to estimate the frequency components from the CSI phase difference. The experiments in this section demonstrate the two critical limitations of the prior multi-person respiration sensing systems that rely on spectral analysis of CSI amplitude or phase difference.

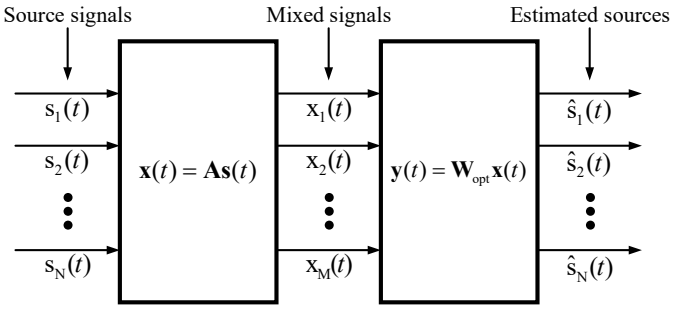

Fig. 5. Block diagram of blind source separation using ICA method.

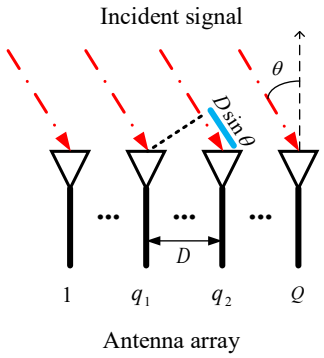

Fig. 6. Antenna array with an incident signal.

4 MODELLING OF WIFI-BASED MULTI-PERSON RESPIRATION SENSING

In this section, we show how the multi-person respiration sensing can be successfully modeled as a BSS problem based on CSI retrieved from commodity WiFi devices.

4.1 Preliminary

4.1.1 BSS and ICA. BSS (Blind Source Separation) refers to reconstructing unknown sources $\mathbf{s}(t)$ from given observations $\mathbf{x}(t)$, which result from the mixture of $\mathbf{s}(t)$ going through unknown propagation channels [1]. The BSS problem can be efficiently solved by a technique called ICA (Independent Component Analysis), if three assumptions are satisfied [30]: (1) the sources are non-Gaussian; (2) the sources are mutually statistically independent; (3) the mixture is of linear nature. Fig. 5 gives a detailed illustration of how ICA method works for BSS. As shown in the figure, the mixing process can be established as:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \tag{1}$$

where $\mathbf{A}$ is an $M \times N$ unknown mixing matrix and the components in $\mathbf{s}(t) = \left[s_1(t)\ s_2(t)\ \ldots\ s_N(t)\right]^T$ are N unknown source signals. ICA method attempts to find a $N \times M$ demixing matrix $\mathbf{W}_{opt}$ such that

$$\mathbf{y}(t) = \mathbf{W}_{opt}\mathbf{x}(t) \tag{2}$$

where $\mathbf{W}_{opt} = \hat{\mathbf{A}}^{-1}$ is an approximation of the inverse of the original mixing matrix $\mathbf{A}$ and $\mathbf{y}(t) = \hat{\mathbf{s}}(t)$ is an approximation of the sources [3]. The basic idea of ICA is to make the output signals $\mathbf{y}(t)$ as non-Gaussian as possible. In practice, iterative methods are used to maximize a given cost function that measures the non-Gaussianity, such as kurtosis and differential entropy [30].

The ICA method varies with the relationship between the number of the components in the observations (i.e., M) and that of independent components (i.e., N). In the case that $M > N$, the learning rule becomes much complex as the pseudo-inverse of matrix $\mathbf{A}$ is involved [38]. In the case that $M < N$, the independent components are not fully identifiable in most cases because $\mathbf{A}$ is non-invertible and the known information to the system is inadequate to represent the full space of the independent components [23]. For simplicity and without losing generality in this paper, we consider the case of the complete ICA where $M = N$.

4.1.2 WiFi CSI. WiFi CSI is a complex value to describe how the WiFi signals propagate from the transmitter to the receiver through multiple paths at a certain carrier frequency. Mathematically, CSI characterizes the linear

superposition of signals from all the paths, and it can be expressed as below:

$$x(f,t) = \sum_{i=1}^{L} A_i(t) e^{-j2\pi \frac{d_i(t)}{\lambda}} \tag{3}$$

where L is the total number of propagation paths, $A_i(t)$ is the amplitude attenuation factor, $d_i(t)$ is the length of the i^{th} propagation path and λ is the wavelength of carrier frequency f. Further, for commodity WiFi devices, since the physically separated transmitter and receiver are not time-synchronized, there is a time-varying phase offset $\phi(t)$ in each CSI sample as below [52]:

$$x(f,t) = e^{-j\phi(t)} \sum_{i=1}^{L} A_i(t) e^{-j2\pi \frac{d_i(t)}{\lambda}} \tag{4}$$

4.2 Effects of Multi-person Respiration on WiFi CSI of Multiple Antennas

First, we consider the effect of multi-person respiration on WiFi CSI with just one transmitting antenna and one receiving antenna. According to prior work [35], the propagation paths in Eq. 4 can be grouped into static paths and dynamic paths. The dynamic paths consist of the path reflected from each person's chest while the static paths consist of the line-of-sight (LoS) path and the paths reflected from static objects such as walls in the environment. Since the chest movement caused by human respiration is 5-12 mm [22], which is much smaller than the meter-level distance between the target and the devices, the signal amplitude of the reflection path can be considered as a constant. Based on the analysis above, the CSI in Eq. 4 can be rewritten as below:

$$x(f,t) = e^{-j\phi(t)} \left(\sum_{i=1}^{N} A_i e^{-j2\pi \frac{d_i(t)}{\lambda}} + b \right) \tag{5}$$

where N is the number of breathing persons, $d_i(t)$ is the path length of signal bouncing off the i^{th} human chest and b is the background static signal corresponding to static paths. When the i^{th} person breathes, the reflection path length $d_i(t)$ increases or decreases corresponding to the chest movement during the exhalation or inhalation, then $x(f,t)$ changes accordingly. For better illustration, we define the following term:

Definition 4.1. ***Respiration Signal*** of a person is defined as a complex number $e^{-j2\pi \frac{d(t)}{\lambda}}$, where $d(t)$ is the path length of signal reflected off his/her chest during respiration and λ is the wavelength.

The goal of multi-person respiration sensing is to extract the respiration signal $e^{-j2\pi \frac{d_i(t)}{\lambda}}$ $(i = 1, 2, ..., N)$ corresponding to the respiration of each individual from Eq. 5.

Next, we consider the scenario when multiple antennas are available at the WiFi transceivers. Assuming that we have P transmitting antennas and Q receiving antennas, we are able to obtain a CSI reading from each antenna pair. What's the relationship among the total M CSI readings where $M = PQ$? To answer this question, we need to first present two important observations. The prior work [51, 52] on single-person respiration sensing presents one key observation: for each person, when he/she moves a short distance (a few centimeters), although the reflection path length of each antenna pair changes, the length difference of reflection paths between two antenna pairs can be considered as a constant. Fig. 6 presents a brief illustration. Since the signal propagation path is much longer than the antenna spacing D, the path length difference between adjacent antennas can be expressed as $D \sin\theta$, where θ is the angle-of-arrival (AoA) of the incident signal. For millimeter-scale chest movement caused by respiration, the AoA of the signal reflected from human chest basically remains unchanged, thus the length difference of the reflection paths between two antennas can be considered constant. Another observation is that, for commodity WiFi devices, the time-varying phase offset in Eq. 5 is the same across different antenna pairs

since all the transmitting antennas share a same RF oscillator [13] and this is also the case for all the receiving antennas.

With these two key observations, we start to investigate the relationship between CSI readings of two different antenna pairs. Without loss of generality, we pick any two CSI from the total M CSI readings and rewrite them as follows:

$$\begin{cases} x_1(f,t) = e^{-j\phi(t)}(\sum_{i=1}^{N} A_{1,i}e^{-j2\pi\frac{d_{1,i}(t)}{\lambda}} + b_1) \\ x_2(f,t) = e^{-j\phi(t)}(\sum_{i=1}^{N} A_{2,i}e^{-j2\pi\frac{d_{2,i}(t)}{\lambda}} + b_2) = e^{-j\phi(t)}(\sum_{i=1}^{N}(A_{2,i}e^{-j2\pi\frac{d_{2,i}(t)-d_{1,i}(t)}{\lambda}})e^{-j2\pi\frac{d_{1,i}(t)}{\lambda}} + b_2) \end{cases} \tag{6}$$

where $x_1(f,t)$ is the first CSI, $x_2(f,t)$ is the second CSI and $d_{2,i}(t) - d_{1,i}(t)$ is the path length difference between the two paths reflected by the i^{th} person for the two selected antenna pairs, which is a constant based on the first observation above. To simplify Eq. 6 for easier illustration, we employ $s_i(f,t)$, $a_{1,i}$ and $a_{2,i}$ to represent the terms: $e^{-j2\pi\frac{d_{1,i}(t)}{\lambda}} = s_i(f,t)$, $A_{1,i} = a_{1,i}$ and $A_{2,i}e^{-j2\pi\frac{d_{2,i}(t)-d_{1,i}(t)}{\lambda}} = a_{2,i}$. We can then simplify Eq. 6 as:

$$\begin{cases} x_1(f,t) = e^{-j\phi(t)}(\sum_{i=1}^{N} a_{1,i}s_i(f,t) + b_1) \\ x_2(f,t) = e^{-j\phi(t)}(\sum_{i=1}^{N} a_{2,i}s_i(f,t) + b_2) \end{cases} \tag{7}$$

where $s_i(f,t)$ is the respiration signal of the i^{th} person according to Def. 4.1. Interestingly, we can find that both CSI readings are a linear mixtures of $s_i(f,t)$ $(i = 1, 2, ..., N)$ and the background static signal.

The same analysis can be applied to the remaining $M - 2$ CSI readings, and they have the same form as Eq. 7. From this result, we can conclude that the signals reflected from multiple persons are linearly mixed at each receiving antenna. At last, we rewrite all M CSI readings in the form of matrix as below:

$$\mathbf{x}(f,t) = e^{-j\phi(t)}(\mathbf{A}\mathbf{s}(f,t) + \mathbf{b}) \tag{8}$$

where $\mathbf{x}(f,t) = \left[x_1(f,t)\ x_2(f,t)\ ...\ x_M(f,t)\right]^T$ represents the CSI readings of all antenna pairs, $\phi(t)$ is the time-varying phase offset, $\mathbf{A}$ is the $M \times N$ mixing matrix with each element $a_{i,j}$ in the i^{th} row and the j^{th} column, $\mathbf{s}(f,t) = \left[s_1(f,t)\ s_2(f,t)\ ...\ s_N(f,t)\right]^T$ represents the source respiration signals and $\mathbf{b} = \left[b_1\ b_2\ ...\ b_M\right]^T$ represents the background static signal.

4.3 The Challenge in Applying ICA to WiFi-based Multi-person Respiration Sensing

Obviously, whether the ICA method can be applied to separate multi-person respiration signals depends on whether the three assumptions in Sec. 4.1.1 are satisfied. As shown in [49], the respiration signals of different people are generally independent of each other and non-Gaussian. Unfortunately, by comparing Eq. 8 with Eq. 1, we find the CSI from commodity WiFi still suffers from an extra time-varying phase offset $\phi(t)$ and the background static signal $\mathbf{b}$. Thus, we cannot directly apply ICA to separate multi-person respiration signals with the raw CSI retrieved from commodity WiFi devices.

A straightforward thought is to cancel out the time-varying phase offset and remove the background static signal so that ICA method can be applied. One method was proposed in the state-of-the-art [52] that employs the ratio of CSI readings from two antennas to cancel out the phase offset. Unfortunately, such division operation distorts the linearity of the mixture. In other words, the ratio of two linear mixtures of $s_i(f,t)$ $(i = 1, 2, ..., N)$ cannot be represented by a linear mixture of $s_i(f,t)$ $(i = 1, 2, ..., N)$ any more as shown below:

$$\frac{x_2(f,t)}{x_1(f,t)} = \frac{e^{-j\phi(t)}(\sum_{i=1}^{N} a_{2,i}s_i(f,t) + b_2)}{e^{-j\phi(t)}(\sum_{i=1}^{N} a_{1,i}s_i(f,t) + b_1)} = \frac{\sum_{i=1}^{N} a_{2,i}s_i(f,t) + b_2}{\sum_{i=1}^{N} a_{1,i}s_i(f,t) + b_1} \neq \sum_{i=1}^{N} a_{0,i}s_i(f,t) + b_0 \tag{9}$$

where $a_{0,i}$ $(i = 1, 2, ..., N)$ and b_0 are complex constants. This can be easily proved by contradiction that the product of two linear mixtures are nonlinear. Thus, canceling out the time-varying phase offset without distorting the linearity for respiration sensing is actually a challenging task.

4.4 Making ICA Work for WiFi-based Multi-person Respiration Sensing

In this subsection, to make ICA work, we present our novel method to process the raw CSI readings retrieved from commodity WiFi. The process consists of two steps: (1) time-varying phase offset cancelation; (2) background static signal removal.

4.4.1 Time-varying Phase Offset Cancelation. As shown in Eq. 9, the linearity is distorted by the division operation while both the numerator and denominator are the linear mixtures of $s_i(f,t)$. Interestingly, we find that, if the denominator $x_1(f,t)$ to the left of Eq. 9 does not contain the respiration signals $s_i(f,t)$ $(i = 1, 2, ..., N)$, i.e., $x_1(f,t) = e^{-j\phi(t)}b_1$, the linearity can be kept as below:

$$\frac{x_2(f,t)}{x_1(f,t)} = \frac{e^{-j\phi(t)}(\sum_{i=1}^{N} a_{2,i}s_i(f,t) + b_2)}{e^{-j\phi(t)}b_1} = \frac{\sum_{i=1}^{N} a_{2,i}s_i(f,t) + b_2}{b_1} = \sum_{i=1}^{N} \frac{a_{2,i}}{b_1}s_i(f,t) + \frac{b_2}{b_1} \tag{10}$$

We call this CSI $x_1(f,t) = e^{-j\phi(t)}b_1$, the *reference CSI*. For convenience, we denote it as $x_{ref}(f,t) = e^{-j\phi(t)}\mu$, where $\phi(t)$ is the time-varying phase offset and μ is a complex-valued constant.

To obtain such a special reference CSI, we rely on the following insight. As shown Sec. 4.2, all M CSI readings contain the same time-varying phase offset, and each of them is the linear mixture of source respiration signals $s_i(f,t)$ $(i = 1, 2, ..., N)$. This enables us to null out the respiration signals by employing a weighted sum of all CSI readings. By carefully adjusting the weight of each CSI reading, the coefficient of each respiration signal $s_i(f,t)$ in the weighted sum may equal to 0. If we can do this, the weighted sum is exactly our wanted reference CSI. Mathematically, the weighted sum of all M CSI readings can be represented as the product of the transpose of a $M \times 1$ weight vector $\mathbf{g} = \begin{bmatrix} g_1 & g_2 & ... & g_M \end{bmatrix}^T$ and the $M \times 1$ observation matrix $\mathbf{x}(f,t)$ as below:

$$\begin{aligned} x_{sum}(f,t) &= \mathbf{g}^T\mathbf{x}(f,t) \\ &= \mathbf{g}^T e^{-j\phi(t)}(\mathbf{A}\mathbf{s}(f,t) + \mathbf{b}) \\ &= e^{-j\phi(t)}(\mathbf{g}^T\mathbf{A}\mathbf{s}(f,t) + \mathbf{g}^T\mathbf{b}) \end{aligned} \tag{11}$$

where $\mathbf{s}(f,t) = \begin{bmatrix} s_1(f,t) & s_2(f,t) & ... & s_N(f,t) \end{bmatrix}^T$ represents the source respiration signals, $\mathbf{g}^T\mathbf{A}$ is the corresponding $1 \times N$ coefficient matrix and $\mathbf{b} = \begin{bmatrix} b_1 & b_2 & ... & b_M \end{bmatrix}^T$ represents the background static signals. We attempt to find a $M \times 1$ weight vector $\mathbf{g}_{opt}$ such that

$$\mathbf{g}_{opt}^T\mathbf{A} = \mathbf{0} \tag{12}$$

where $\mathbf{0}$ is a $1 \times N$ zero vector. If Eq. 12 is satisfied, the weighted sum in Eq. 11 does not contain the respiration signals $s_i(f,t)$ $(i = 1, 2, ..., N)$. And the reference CSI can be obtained as $x_{ref}(f,t) = e^{-j\phi(t)}\mu$, where $\mu = \mathbf{g}_{opt}^T\mathbf{b}$. However, we have to answer two questions for this construction: (1) whether there exists the nontrivial solution $\mathbf{g}_{opt}$ for Eq. 12, as the trivial solution $\mathbf{g}_{opt} = \begin{bmatrix} 0 & 0 & ... & 0 \end{bmatrix}^T$ does not make sense; (2) how to obtain the nontrivial solution $\mathbf{g}_{opt}$.

Answer to question 1. Actually, by taking the transpose operation on both sides of Eq. 12, we have:

$$\mathbf{A}^T\mathbf{g}_{opt} = \mathbf{0} \tag{13}$$

where $\mathbf{A}^T$ is a $N \times M$ matrix and $\mathbf{0}$ is a $N \times 1$ zero vector. Obviously, Eq. 13 is a homogeneous system of linear equations that has N equations and M variables $(g_1, g_2, ..., g_M)$. Recall the well-known theorem in linear algebra that, a homogeneous system of linear equations in which there are more variables than equations must have some nontrivial solutions [6]. Thus, there must be some nontrivial solutions $\mathbf{g}_{opt}$ to satisfy Eq. 13, given that $M > N$. That's to say, to ensure that Eq.13 has nontrivial solutions, the number of antenna pairs needs to be larger than that of breathing persons. Fortunately, the commodity MIMO WiFi cards such as Intel 5300 [10] and

Atheros AR9580 [44] can support three antennas, which provides at most nine antenna pairs. Thus, theoretically, we are able to construct such reference CSI to support sensing of up to eight persons.

Answer to question 2. Unfortunately, since we have no knowledge about the detailed value of each element in $\mathbf{A}^T$, we are not able to directly obtain the nontrivial solution $\mathbf{g}_{opt}$ by solving Eq. 13. Instead, we turn it into an optimizing problem, and find $\mathbf{g}_{opt}$ that makes the weighted sum in Eq. 11 meet our expectation by maximizing or minimizing an objective function. Our selection of objective function is based on the following insight. If the $1 \times N$ coefficient matrix $\mathbf{g}^T\mathbf{A}$ in Eq. 11 equals to $\mathbf{0}$, $x_{sum}(f,t)$ will not contain any respiration information and its amplitude keeps constant over time. Thus, in the spectrum of its amplitude, very little energy is concentrated in the range of normal respiration rate (10-30 bpm [29]). And the CSI seems to be measured in a static environment. Thus we propose the following definition to act as the objective function:

Definition 4.2. ***Respiration Energy Ratio (RER)*** of a time series of CSI measurements is defined as the ratio of respiration energy to the overall energy in its amplitude.

The calculation of RER consists of two steps: (1) take FFT of the amplitude of $x_{sum}(f,t)$ over a time window of 30 seconds; (2) divide the energy sum of all FFT bins in the range of normal respiration rate (0.167-0.5 Hz) by the energy sum of all FFT bins. Then the genetic algorithm [9] is utilized to find $\mathbf{g}_{opt}$ by minimizing the objective function above. We choose the generic algorithm because it has the ability to avoid being trapped in a local optimal like traditional gradient-based methods that search from a single point. Once $\mathbf{g}_{opt}$ is obtained, we can construct the reference CSI based on Eq. 11, where $x_{ref}(f,t) = e^{-j\phi(t)}\mathbf{g}_{opt}^T\mathbf{b}$. Then, by the division operation, we cancel out the time-varying phase offset without distorting the linearity as below:

$$\frac{\mathbf{x}(f,t)}{x_{ref}(f,t)} = \frac{e^{-j\phi(t)}(\mathbf{A}\mathbf{s}(f,t)+\mathbf{b})}{e^{-j\phi(t)}\mu} = (\frac{\mathbf{A}}{\mu})\mathbf{s}(f,t) + \frac{\mathbf{b}}{\mu} \tag{14}$$

where $\mathbf{x}(f,t) = \left[x_1(f,t)\, x_2(f,t) \ldots x_M(f,t)\right]^T$ represents the raw CSI readings of all antenna pairs, $\mathbf{A}$ is the $M \times N$ unknown mixing matrix, $\mathbf{s}(f,t) = \left[s_1(f,t)\, s_2(f,t) \ldots s_N(f,t)\right]^T$ represents the source respiration signals, $\mathbf{b} = \left[b_1\, b_2 \ldots b_M\right]^T$ represents the background static signals and $\mu = \mathbf{g}_{opt}^T\mathbf{b}$ is the complex-valued constant.

Fig. 7 shows an example of constructing the reference CSI. The raw CSI in Fig. 7 (a) is measured in a static environment without any moving targets. Obviously, the CSI amplitude keep nearly constant with a very small RER of 0.0020. The CSI in Fig. 7 (b) is measured in the presence of two breathing persons, whose ground-truth respiration rates are 14 bpm and 16 bpm, respectively. We can observe the large fluctuation in CSI amplitude and its RER is 0.6439, which is much larger than that in static environment. The reference CSI in Fig. 7 (c) is constructed by CSI readings of three antenna pairs in the two-person scenario above. By minimizing the RER, the reference CSI is obtained with little respiration information, which is comparable to that in static environment.

4.4.2 Background Static Signal Removal. To remove the background static signal $\frac{\mathbf{b}}{\mu}$ in Eq. 14, we measure CSI in the static environment without any moving objects:

$$\mathbf{x}'(f,t) \approx \left[e^{-j\phi(t)}b_1\, e^{-j\phi(t)}b_2 \ldots e^{-j\phi(t)}b_M\right]^T = e^{-j\phi(t)}\mathbf{b} \tag{15}$$

where $\mathbf{b}$ is the background static signal in Eq. 8. Note that the time-varying phase offset of CSI in static environment is different from that in multi-person scenario, as they are measured at two different timestamps. Thus, we use the same weight vector $\mathbf{g}_{opt}$ obtained in Sec. 4.4.1 to construct reference CSI in static environment as below:

$$x'_{ref}(f,t) = \mathbf{g}_{opt}^T\mathbf{x}'(f,t) \approx e^{-j\phi(t)}\mathbf{g}_{opt}^T\mathbf{b} = e^{-j\phi(t)}\mu \tag{16}$$

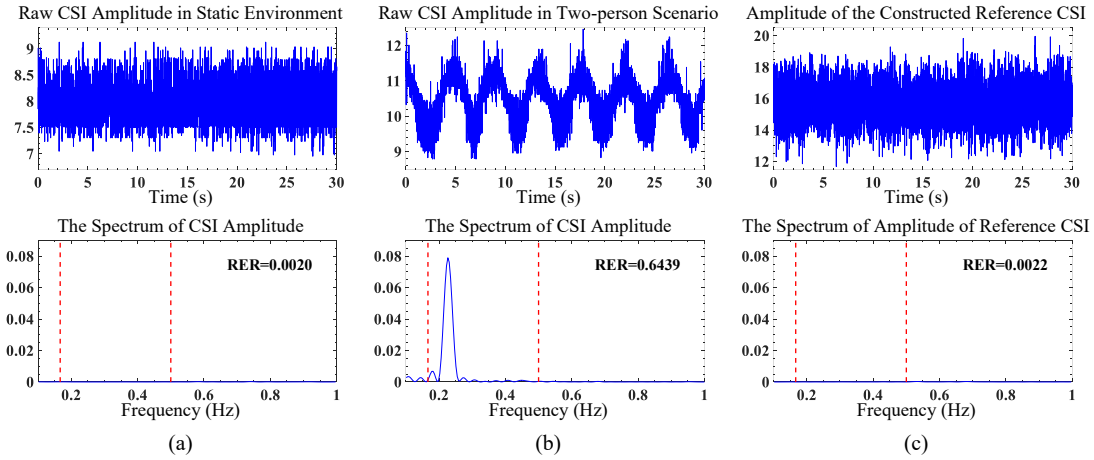

Fig. 7. The CSI amplitudes and the corresponding spectrums generated by FFT: (a) in a static environment; (b) in the presence of two breathing persons; (c) our constructed reference CSI for the scenario in (b). By minimizing the RER, we obtain the reference CSI comparable to the CSI in static environment, which meets our expectation. The red dashed lines in the figure label the range of normal respiration rate. And RER is the sum of all FFT bins in the range.

where $\mu = \mathbf{g}_{opt}^T \mathbf{b}$ is a complex-valued constant according to Sec. 4.4.1. Then, the background static signal in multi-person scenario can be approximated by the division operation as follows:

$$\frac{\mathbf{x}^{'}(f,t)}{x_{ref}^{'}(f,t)} \approx \frac{e^{-j\phi(t)}\mathbf{b}}{e^{-j\phi(t)}\mu} = \frac{\mathbf{b}}{\mu} \tag{17}$$

After these two steps, the processed CSI can be finally obtained as:

$$\mathbf{x}_{new}(f,t) = \frac{\mathbf{x}(f,t)}{x_{ref}(f,t)} - \frac{\mathbf{x}^{'}(f,t)}{x_{ref}^{'}(f,t)} \approx (\frac{\mathbf{A}}{\mu})\mathbf{s}(f,t) \tag{18}$$

which is exactly the linear mixture of each person's respiration signal that we want. Now the ICA method can finally be applied to the processed CSI for separating the source respiration signals.

Based on the literature on blind source separation by ICA, our approach can efficiently extract the source respiration signals as long as they are mutually independent of each other. Such condition holds true [49] even when two persons have similar respiration rates. This is because breathing is a natural physical activity and two persons' respiration processes have a extremely small chance to be fully synchronized over time (same rate and have peaks at exactly the same time). While the prior work can only present the average respiration rate, our approach can recover the detailed respiration patterns over time for each individual person, which can be utilized to detect abnormal respiration such as tachypnea, bradypnea and apnea. The reason why the "blind spots" do not exist in our system is illustrated in Sec. 5.3.2.

5 THE MULTISENSE SYSTEM

In this section, we will present the system workflow of MultiSense with a detailed introduction of each module. The sketch of the MultiSense system is shown in Fig. 8. It consists of four basic modules: Data Collection, Data Preprocessing, Multi-person Respiration Separation and Segment Matching.

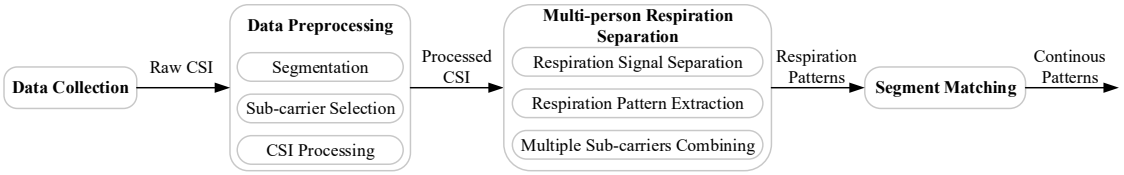

Fig. 8. The MultiSense system overview.

5.1 Data Collection

We collect CSI data from the receiver using the CSI tool [10], which reports the complex-valued CSI samples for each received packet. The Intel 5300 WiFi card in receiver is configured to run at a central frequency of 5.24 GHz with a sampling rate of 200 Hz for reporting CSI, and it provides CSI information on 30 sub-carriers. Unless specified otherwise, the transmitter and the receiver are both equipped with three antennas, which provides nine antenna pairs. Thus, at each timestamp, we can obtain a 9×30 CSI matrix. One thing to note, MultiSense is adaptive to the scenarios where either the transmitter or the receiver has less than three antennas, as long as the number of antenna pairs is larger than that of persons.[2]

5.2 Data Preprocessing

5.2.1 Segmentation. When there are large motions, MultiSense will not be able to accurately monitor respiration for multiple persons. First, the length difference of reflection paths between two antenna pairs cannot be considered as a constant since the angle-of-arrival of the reflected signal corresponding to a large motion changes. Therefore, the CSI readings of different antenna pairs are no longer linearly related to each other as before in Eq. 6. Second, the signal amplitude of the reflection path corresponding to a large motion cannot be considered as a constant. In such case, the mixing matrix $\mathbf{A}$ in Eq. 8 becomes time-varying. Thus, we need to segment the CSI streams when we detect large motions.

Here, we introduce the motion detector proposed in WiDetect [54] to flag periods of time when one of the targets moves. Based on a statistical model that comprehensively leverages all existing multipath components indoors for motion detection, WiDetect achieves almost zero false alarms for whole-home coverage using only a pair of transceivers. Then the time series of CSI readings can be splited into a series of stable periods. And the CSI measurements in each stable period is further segmented into a series of 30-second pieces, which are fed into the subsequent processing modules.

5.2.2 Sub-carrier Selection. As shown in Sec. 5.1, when both transmitter and receiver are equipped with three antennas, we get a 9×30 CSI matrix at each timestamp. Thus, the CSI measured in 30 seconds with a sampling rate of 200 Hz is a $9 \times 30 \times 6000$ tensor, as shown in Fig. 9 (a). Since each sub-carrier experiences different multipath fading and shadowing efffects, the CSI readings of different sub-carriers have varying sensitivity to subtle respiration sensing. Thus, we just include those so-called "good" ones for respiration sensing. The selection is based on the following insight. A low RER of a time series of CSI readings for a certain sub-carrier indicates a low possibility that it contains useful respiration information, hence it should be excluded.

As illustrated in Fig. 9, the selection consists of three steps. First, for the time series of CSI readings corresponding to each antenna pair and sub-carrier, we calculate its RER as in Sec. 4.4.1. Second, for each sub-carrier, we estimate

[2] As shown before, this condition is necessary for constructing the so-called reference CSI.

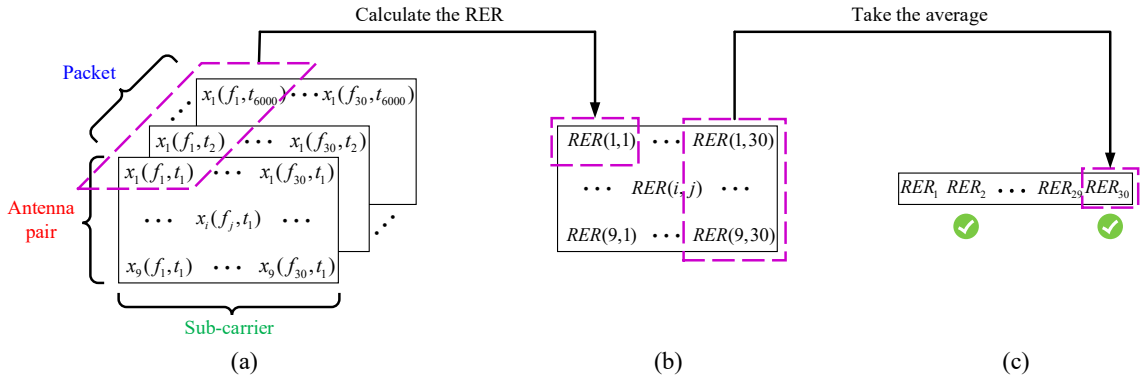

Fig. 9. Illustration of sub-carrier selection: (a) a $9 \times 30 \times 6000$ CSI tensor for the CSI measured in 30 seconds; (b) a 9×30 RER matrix; (c) a 1×30 matrix, whose element is the average of elements in corresponding column of (b).

its RER by averaging the RER of all antenna pairs. Third, let the maximal RER among all 30 sub-carriers be ε, we just select those sub-carriers whose RER is larger than 0.8ε. [3]

5.2.3 CSI Processing. For each selected sub-carrier, we process the CSI through two steps described in Sec. 4.4. In the presence of N breathing persons, we just utilize the CSI of $N + 1$ antenna pairs to construct reference CSI rather than that of all antenna pairs, as it helps to reduce unnecessary system overhead to some degree. Then, we divide the raw CSI of all antenna pairs by the reference CSI to cancel out time-varying phase offset.

Next, we estimate the background static signal and subtract it to obtain the final CSI. As demonstrated in [42], the background signal does not change in a short period of time as long as there is no significant change in the environment (e.g., the furniture moves). Thus, our system records the background signal at system startup as the initial estimation and updates the background signal estimation periodically to handle the changing multipath.

Initialize Estimation. According to Sec. 4.4.2, we processed the CSI data measured in static environment to obtain the background static signal. Theoretically, we need just one CSI sample for this step. In our implementation, we record 1-second CSI data (i.e., 200 samples) and further average the approximation result of all samples within one second to get a robust approximation.

Update Estimation. Basically, there are two kinds of movements that result in the change in multipath: one is the movements that induce static-path change (e.g., a piece of furniture is moved to another location) and the other is the movements that induce dynamic-path change (e.g., people are moving around). For the latter case, the moving subjects induce dynamic reflection components, and these signals are not considered as static background signal. If the surrounding people are far enough away from the transceivers, the impact of their motions on MultiSense are negligible. In the former case, as the furniture introduces a static reflection component, the background signal changes after the furniture movement. In the last segment (i.e., 30-second CSI data before the furniture movement), let the time-series CSI after phase offset cancelation be $u_1(t)$, and the background signal estimated be b_1. In the current segment, let the time-series of CSI after phase offset cancelation be $u_2(t)$. To estimate the current background signal b_2, we rely on the following insight. As described in Sec. 4.4, the CSI after phase offset cancelation consists of two parts: one is the background static signal and the other is the sum of respiration signals from all subjects. If the subjects breathe naturally and do not move around (i.e., they are semi-stationary), the mean of the respiration signal over a relatively large time window (multiple

[3]Using a large amount of data, we tested different coefficients of the threshold, ranging from 0.4 to 0.9 at a step size of 0.05. And we empirically chose 0.8ε for sub-carrier selection.

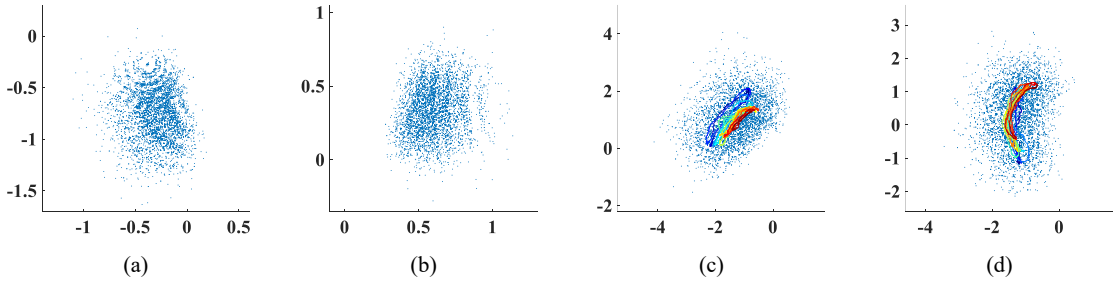

Fig. 10. An example of applying RobustICA to respiration signal separation in two-person scenario: (a) and (b) the processed CSI of two antenna pairs in complex plane; (c) and (d) the recovered two respiration signals and the corresponding denoised version in complex plane.

respiration cycles) is approximately constant, because the inhalation and exhalation induced signal variations will cancel each other out [49]. This constant value is only related to the subject and is independent of the furniture movement. Thus, we have: $\overline{u_1(t)} - b_1 \approx \overline{u_2(t)} - b_2$. Then, the current background signal b_2 can now be estimated as: $\overline{u_2(t)} - \overline{u_1(t)} + b_1$.

5.3 Multi-person Respiration Separation

In this subsection, for each selected sub-carrier, we apply ICA method to separate the source respiration signals and extract each person's respiration pattern from the complex-valued respiration signal. And a multiple sub-carriers combining scheme is further adopted to get the robust respiration patterns.

5.3.1 Respiration Signal Separation. For each sub-carrier, we obtain the time series of processed CSI from multiple antenna pairs. And the ICA method can be applied to separate respiration signals. As shown in Sec. 4.1.1, for simplicity and without loss of generality, we consider the case of the complete ICA where the number of antenna pairs equals to that of breathing persons. That is to say, in the case of N persons, we just take the processed CSI of N antenna pairs as the input of ICA. The antenna pair selection is based on the similar strategy as before in sub-carrier selection: for each selected sub-carrier, we just select the corresponding top N antenna pairs that have a larger RER among all antenna pairs.

Here, we adopt the RobustICA algorithm proposed in [50] to separate respiration signals. The algorithm performs independent component analysis by iteratively maximizing the kurtosis contrast function with algebraic optimal step size. RobustICA works for complex-valued signals and it shows a very high convergence speed measured in terms of source extraction quality versus numbers of operations. The output of RobustICA is the complex-valued respiration signals of N persons. Ideally, they are in the form of $s_i(f,t) = e^{-j2\pi \frac{d_i(t)}{\lambda}}$, where $d_i(t)$ is the path length of signal reflected off the i^{th} human chest. As the chest movement caused by human respiration is 5-12 mm [22], the locus of $s_i(f,t)$ in complex plane is a circular arc (part of a full circle) during respiration. What's more, when a person inhales, the reflection path length $d_i(t)$ decreases, then $s_i(f,t)$ rotates along the circular arc counterclockwise in complex plane.

Fig. 10 shows an example of respiration signal separation using RobustICA in two-person scenario for sub-carrier 7. As shown in the figure, the dots are the time series of samples and the colored lines are the denoised version after Savitzky-Golay filter which can effectively preserve the envelope of the raw waveform [5]. The ground-truth respiration rates are 19 bpm and 20 bpm. The processed CSI of two antenna pairs is presented in Fig. 10 (a) and Fig. 10 (b). And the output of RobustICA is presented in Fig. 10 (c) and Fig. 10 (d). We can observe that the locus of each person's respiration signals is a circular arc in complex plane, which meets our expectation.

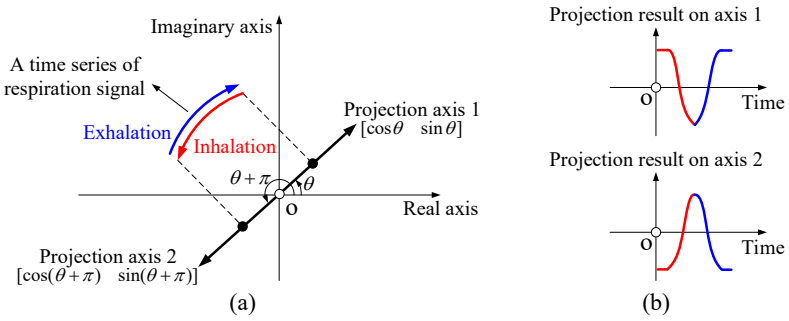

Fig. 11. An example of projecting a time series of respiration signal in complex plane: (1) the respiration signal and two opposite axes (i.e., axis 1 and axis 2) that maximize the variance; (b) projection results on two corresponding axes in (a).

5.3.2 Respiration Pattern Extraction. A Savitzky-Golay filter [5] is first applied to smooth the complex-valued respiration signal. Then, to extract respiration pattern from the smoothed respiration signal, we adopt the state-of-the-art method proposed in [8], which consists of the following two steps. First, as shown in Fig. 11 (a), we linearly combine the real/imaginary parts of the respiration signal $\mathbf{z}$ by projecting it on an axis $\left[\cos\theta \ \sin\theta\right]$ in complex plane as follows:

$$\mathbf{r}(\theta) = \left[\cos\theta \ \ \sin\theta\right] \left[\Re(\mathbf{z}) \ \Im(\mathbf{z})\right]^T \tag{19}$$

where $\Re(\mathbf{z})$ and $\Im(\mathbf{z})$ are the real/imaginary parts of respiration signal $\mathbf{z}$, respectively. Second, by varying θ from 0 to 2π, we select the one that maximizes the variance of $\mathbf{r}(\theta)$.

However, prior work ignores a problem that it fails to differentiate two opposite projection axes as their corresponding results have the same variance with $\mathbf{r}(\theta) = -\mathbf{r}(\theta + \pi)$, as shown in Fig. 11 (b). In real word, limited by the computer accuracy, the two projection results may have slightly different variances, which makes the method randomly select one of the two results in Fig. 11 (b). If the method proposed in [8] is applied directly, in some cases an ascending waveform corresponds to exhalation (e.g., projection result on axis 1 in Fig. 11 (b)), while in other cases it corresponds to inhalation (e.g., projection result on axis 2 in Fig. 11 (b)).

Here, we want MultiSense to be similar to the commercial respiration belt, that is, an ascending waveform always corresponds to the inhalation. To achieve this goal, we first apply the extended differentiate and cross-multiply (DACM) algorithm proposed in [17] to estimate the rotation direction of the respiration signal in complex plane. The algorithm computes a rough derivative to the arctangent-demodulated phase information, where a positive derivative exactly corresponds to the counterclockwise rotation direction and vice versa. Second, we apply PCA to obtain an initial respiration pattern. Third, we check whether the trend (upward or downward) of the initial respiration pattern matches the rotation direction (counterclockwise or clockwise) of the respiration signal. In the case that the respiration signal rotates counterclockwise (i.e., inhalation) but the waveform drops, we switch the initial projection axis to its opposite direction and obtain the corresponding respiration pattern.

Now, we illustrate why the "blind spots" do not exist in our system. If the respiration signal in Fig. 11 is projected on an axis $\left[\cos(\theta + \pi/2) \ \sin(\theta + \pi/2)\right]$, which is perpendicular to the circular arc, the projection result will have litter fluctuation. The subtle variation can easily be buried in noise without a detection. Obviously, for each respiration signal, different projection axes have different capabilities in terms of sensing respiration. As shown in Eq. 5, in multi-person scenario, the respiration signals of multiple persons are mixed together in the CSI. Traditional approaches simply project the complex-valued CSI into its amplitude, thus all respiration signals share the same project axis. However, a good projection axis for respiration signal of one person may be a

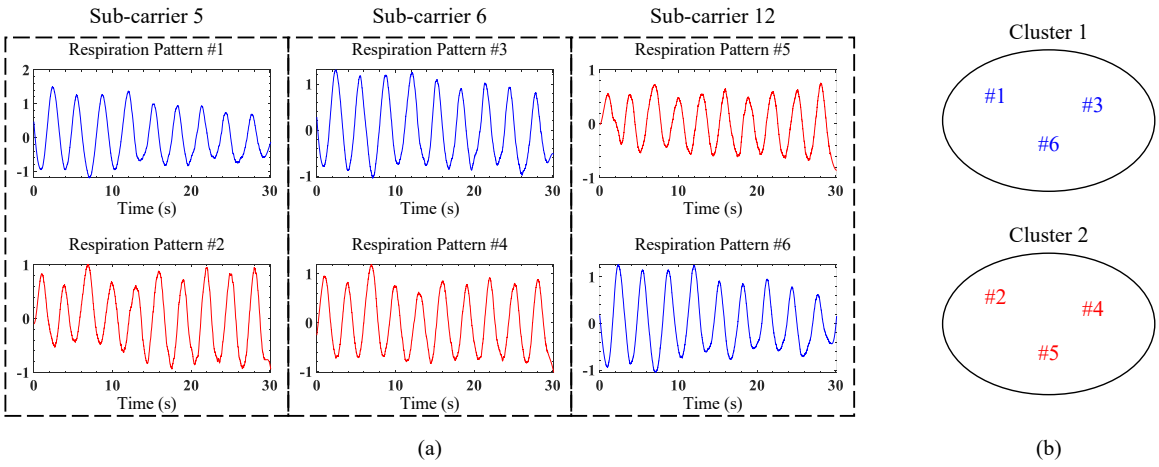

Fig. 12. An example of respiration pattern clustering in two-person scenario: (a) the total six normalized patterns from three sub-carriers (i.e., 5, 6, and 12); (b) each cluster contains three respiration patterns that belong to the same person.

bad one for that of another person, thus resulting in the "blind spots" issue. Different from the prior work that projects the mixed respiration signals into one axis, our approach separates the respiration signals of multiple persons and project each respiration signal on its unique best axis to extract the pattern, thus the "blind spots" do not exist.

5.3.3 Multiple Sub-carriers Combining. Up to now, for each sub-carrier, we obtain N respiration patterns that correspond to N persons. Given that there are K selected sub-carriers, we get $N \times K$ respiration patterns in total. As shown in Fig. 12 (a), we obtain six respiration patterns from three sub-carriers (i.e., 5, 6 and 12) in two-person scenario, where the ground-truth respiration rates are 19 bpm and 20 bpm. To combine the respiration patterns of K sub-carriers, we have to first pick up the respiration patterns that belong to the same person.

Obviously, by removing the ambiguity in Sec. 5.3.2, the respiration patterns of different sub-carriers that belong to the same person are now in-phase and very similar to each other. Thus, two patterns from different sub-carriers that have a higher correlation are more likely to belong to the same person. Here, we utilize the correlation-based K-means algorithm [11] to cluster the $N \times K$ respiration patterns into N clusters, where each cluster is the set of respiration patterns that belong to the same person. As shown in Fig. 12 (b), each cluster contains three elements that are highly correlated to each other. Then, for each cluster, we combine the contained respiration patterns by employing a weighted sum of them, where the weight of each pattern is its RER value. Mathematically, the final combined respiration pattern from the cluster C can be represented as:

$$\mathbf{r}_{msc} = \sum_{\mathbf{r} \in C} RER(\mathbf{r}) \times \mathbf{r} \tag{20}$$

where $\mathbf{r}$ is the element (i.e., the 30-second respiration pattern) in the cluster C and $RER(\mathbf{r})$ is the RER value of the time series of respiration pattern $\mathbf{r}$. Note that our system does not focus on extracting the respiration rate but the time series of respiration pattern. Here, we estimate the respiration rate by finding the peaks in the respiration pattern as in [51].

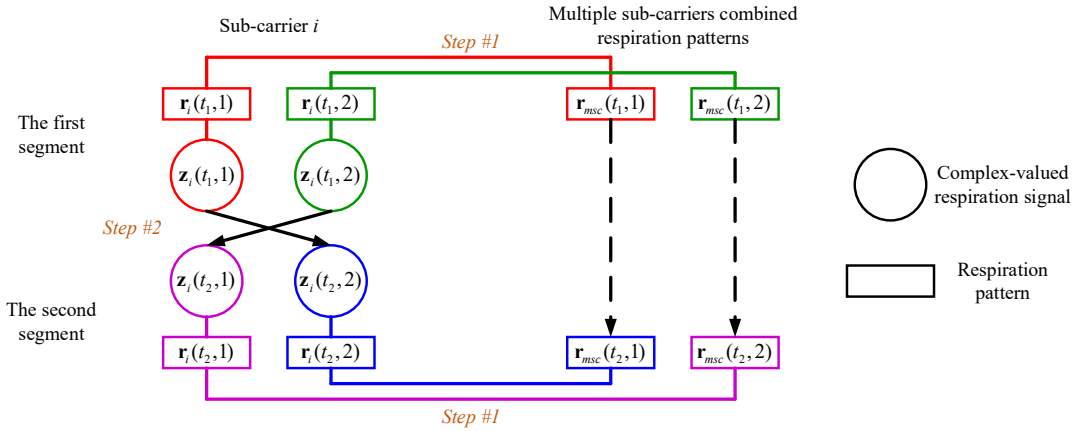

Fig. 13. An example of segment matching in two-person scenario. The matching problem can be solved by two steps: (a) for each segment, we match the respiration signals of a selected sub-carrier i with the MSC respiration patterns (e.g., $\mathbf{z}_i(t_1,1) \leftrightarrow \mathbf{r}_{msc}(t_1,1)$ and $\mathbf{z}_i(t_2,2) \leftrightarrow \mathbf{r}_{msc}(t_2,1)$); (b) we match the respiration signals of sub-carrier i across two segments (e.g., $\mathbf{z}_i(t_1,1) \leftrightarrow \mathbf{z}_i(t_2,2)$). Finally, we manage to match $\mathbf{r}_{msc}(t_1,1)$ with $\mathbf{r}_{msc}(t_2,1)$.

5.4 Segment Matching

After multi-person respiration separation in Sec. 5.3, we get N multiple sub-carriers combined (MSC) respiration patterns that correspond to N persons for each segment (i.e., 30-second CSI data). The last step is to stitch the respiration patters from two adjacent segments that belong to the same person. Take the two-person scenario for example. In the first segment, we have two respiration patterns $\mathbf{r}_{msc}(t_1,1)$ and $\mathbf{r}_{msc}(t_1,2)$. And in the next segment, we get another two respiration patterns $\mathbf{r}_{msc}(t_2,1)$ and $\mathbf{r}_{msc}(t_2,2)$. One problem that arises is that whether $\{\mathbf{r}_{msc}(t_1,1), \mathbf{r}_{msc}(t_2,1)\}$ or $\{\mathbf{r}_{msc}(t_1,1), \mathbf{r}_{msc}(t_2,2)\}$ belong to the same person. Obviously, the correlation between two respiration patterns is not a proper metric to differentiate the two cases since the respiration pattern of the same person may change across two segments.

Fig. 13 presents an example of segment matching in two-person scenario. As shown in the figure, we solve the matching problem by two steps: (1) for each segment, we match the respiration signals of a selected sub-carrier i with the MSC respiration patterns; (2) we match the respiration signals of sub-carrier i across two adjacent segments. As shown in the figure, if the two steps are achieved, we can finally match $\mathbf{r}_{msc}(t_1,1)$ with $\mathbf{r}_{msc}(t_2,1)$. Here, we choose sub-carrier i as the one that has the largest RER among all 30 sub-carriers (in Sec. 5.2.2) in the first segment. And we also include sub-carrier i in the second segment for multi-person respiration separation.

Step 1. Recall that, in Sec. 5.3, for each selected sub-carrier, each of its respiration signals is projected on an axis in complex plane to generate the corresponding respiration pattern. Then the patterns from all selected sub-carriers (including sub-carrier i) are clustered and combined to generate the final MSC respiration patterns. Thus, for each respiration signal of sub-carrier i, we know which MSC respiration pattern's combination process it participate in. In other word, we can directly match each respiration signal of sub-carrier i with the corresponding MSC respiration pattern that it participate in combination.

Step 2. For the second step, we rely on the following insight. For sub-carrier i, each column of the $M \times N$ mixing matrix ($\frac{\mathbf{A}}{\mu}$ in Eq. 18) is a $M \times 1$ vector, which represents how the respiration signal of a certain person contributes to the processed CSI from M antenna pairs. We know from Eq. 6 and Eq. 7 that, for sub-carrier i, each element of the $M \times 1$ vector is a complex number, which is mainly determined by two factors: (1) the amplitude of the path reflected off the person; (2) the angle-of-arrival of the signal reflected off the person. Obviously, the two

factors are both highly associated with his/her physical location, and they varies from person to person. Even if a person slightly changes his position, he/she is still in the same general location and hence the mixing matrix of his/her respiration just changes by a small amount. Thus, for sub-carrier i, two respiration signals from two adjacent segments that have the more similar $M \times 1$ mixing vector are more likely to belong to the same person. Here, we define the following similarity metric to measure the similarity of two complex-valued $M \times 1$ vectors $\mathcal{A}_1$ and $\mathcal{A}_2$:

$$\mathcal{L}(\mathcal{A}_1, \mathcal{A}_2) = \sum_{i=1}^{M} (1 - \frac{|\mathcal{A}_1(i) - \mathcal{A}_2(i)|}{|\mathcal{A}_1(i)| + |\mathcal{A}_2(i)|}) \tag{21}$$

where $\mathcal{A}_1(i)$ and $\mathcal{A}_2(i)$ are the complex-valued elements in the i^{th} row of $\mathcal{A}_1$ and $\mathcal{A}_2$, respectively. Obviously, a larger value of the similarity metric indicates a higher similarity between two mixing vectors. Recall that, in Sec. 5.3.1, in the presence of N persons, we just take the processed CSI of N antenna pairs as the input of ICA, thus we have $M = N$ in Eq. 21. One thing to note, for sub-carrier i, we select the same N antenna pairs across two adjacent segments.

Now, for sub-carrier i, we leverage the similarity metric in Eq. 21 to match its N respiration signals in the first segment with that in the second segment. After applying ICA to the processed CSI of N antenna pairs in the j^{th} segment ($j = 1, 2$), we obtain N respiration signals $(\mathbf{z}(t_j, 1), \mathbf{z}(t_j, 2), ..., \mathbf{z}(t_j, N))$ and N corresponding $N \times 1$ mixing vectors $(\mathcal{I}(t_j, 1), \mathcal{I}(t_j, 2), ..., \mathcal{I}(t_j, N))$. To represent the matching of respiration patterns across two segments, we introduce a term σ, which is the permutation of $(1, 2, ..., N)$. For example, there are six permutations for $N = 3$, namely, $(1, 2, 3)$, $(1, 3, 2)$, $(2, 1, 3)$, $(2, 3, 1)$, $(3, 1, 2)$ and $(3, 2, 1)$. Then, such a permutation σ indicates that the respiration signals of two segments are matched as follows:

$$\mathbf{z}(t_1, 1) \leftrightarrow \mathbf{z}(t_2, \sigma(1)), \mathbf{z}(t_1, 2) \leftrightarrow \mathbf{z}(t_2, \sigma(2)), ..., \mathbf{z}(t_1, N) \leftrightarrow \mathbf{z}(t_2, \sigma(N)) \tag{22}$$

Obviously, there are totally $N!$ possible permutations. For each possible permutation σ, based on Eq. 21, we define the following consistency metric to capture the likelihood that all N respiration signals in the first segment are correctly matched with that in the second segment:

$$\mathcal{L}_{all} = \sum_{i=1}^{N} \mathcal{L}(\mathcal{I}(t_1, i), \mathcal{I}(t_2, \sigma(i))) \tag{23}$$

where $\mathcal{I}(t_1, i)$ and $\mathcal{I}(t_2, \sigma(i))$ are the corresponding mixing vectors of respiration signals $\mathbf{z}(t_1, i)$ and $\mathbf{z}(t_2, \sigma(i))$, respectively. We go through all possible permutations and select the one that has the largest consistency metric. By the two steps above, we finally match N MSC respiration signals in the first segment with that in the second segment. By stitching the respiration patterns across two adjacent segments, we are able to obtain the continuous respiration patterns for continuous monitoring.

6 EVALUATION

In this section, we conduct comprehensive experiments to evaluate the performance of MultiSense with commodity WiFi devices. In Sec. 6.1, we describe the experiment setup. In Sec. 6.2, we zoom in and evaluate the two key components of MultiSense, i.e., recovered detailed respiration patterns of multiple persons and segment matching. In Sec. 6.3, we compare MultiSense with the state-of-the-art approaches in terms of respiration rate accuracy. In Sec. 6.4, we evaluate the robustness of MultiSense in challenging scenarios, i.e., the WiFi AP is located in another room. In Sec. 6.5, we further investigate the impact of ambient motion (i.e., the motion from others in the environment) on the performance of MultiSense.

(a) (b)

Fig. 14. The experimental setup in two scenarios: (a) all subjects sleep on a bed in the bedroom; (b) each subject sits on a couch or chair in the living room.

6.1 Experimental Setup

6.1.1 Devices. In all experiments, we use two GigaByte mini-PCs equipped with cheap Intel 5300 WiFi cards as transceivers. The receiver is configured to work under the monitor mode to capture packets from the transmitter. Unless specified otherwise, both the transmitter and receiver are equipped with three omnidirectional antennas. The frequency of WiFi channel is set to 5.24 GHz with a bandwidth of 20 MHz. And the transmitter sends 200 packets per second at a transmission power of 15 dBm. To obtain the ground-truth respiration of multiple persons, we ask each of them to wear a commercial Neulog respiration belt [12].

6.1.2 Environments. As shown in Fig. 14, we evaluate the performance of MultiSense in two scenarios: (1) multiple subjects sleep on a bed (1.9 m × 2.4 m) in a bedroom; (2) multiple subjects sit on a couch or chair in a living room (4 m × 7.5 m). Both environments have many furniture and electrical appliances, which create rich multipath. For the sleeping scenario in Fig. 14 (a), we mount the transceivers on the ceiling with a spacing of 2.7 m between them. And we can adjust the height of transceivers from 1.6 m to 2.8 m with the help of a hanger. For the sitting scenario in Fig. 14 (b), we place the transceivers on two tripods. As demonstrated in the prior work [19, 52], to better capture the millimeter-level chest movement caused by respiration, the antennas of transceivers are placed horizontally in the sleeping scenario and vertically in the sitting scenario.

6.1.3 Participants. We recruited 21 participants, i.e., 12 males and 9 females, aged from 14 to 57. Their natural respiration rates range from 14 bpm to 22 bpm. Some participants have dramatically different respiration rates and some have similar rates.

6.2 Evaluation of the Components of MultiSense

To the best of our knowledge, MultiSense is the first commodity WiFi based system that is capable of continuously monitoring the respiration patterns of multiple persons even they have similar respiration rates. In this subsection, we evaluate the two key components of our system, i.e., respiration pattern recovery and segment matching.

6.2.1 Experiment Settings. We conduct experiments in two different environments, as shown in Sec. 6.1.2. For both environments, we instruct the subjects to sit or lie at different locations with a distance between each other ranging from 5 cm to 150 cm. Specifically, for the living room environment, the subjects have the option to sit at just one side or both sides of the transmitter-receiver LoS path. In the bedroom environment we can adjust the height of transceivers with a fixed LoS path length of 2.7 m, while in the living room environment we can change not only the location but also the distance between the transceiver pair. In each environment, we conduct

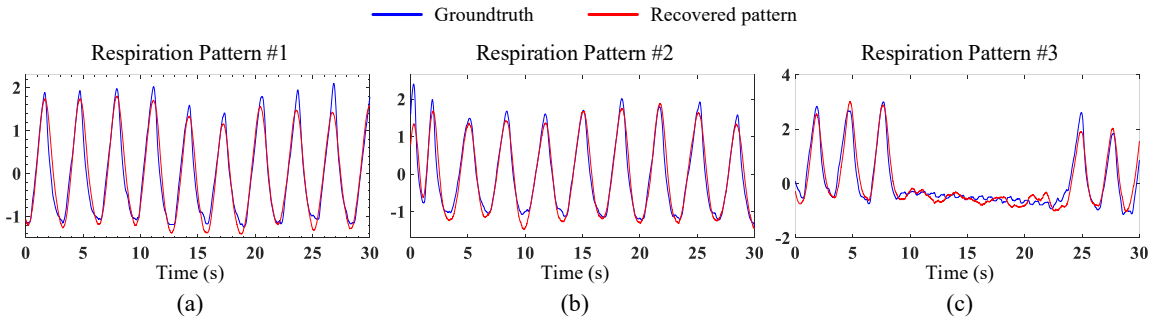

Fig. 15. An example of recovered respiration patterns and the ground-truth patterns in three-person scenario.

experiments by varying the subjects' locations with respect to the transceivers and clearly state in which part of the Fresnel zones each target is located. We measure the background CSI without any target in the environments.

In each environment above, the two-person, three-person, four-person and five-person scenarios are all evaluated. For simplicity, we use the term 1×3 MIMO to represent the transmitter is equipped with one antenna while the receiver is equipped with three antennas. In two-person scenario, we test MultiSense with 1×3 MIMO, 2×3 MIMO and 3×3 MIMO. And in three-person, four-person and five-person scenarios, we test MultiSense with 2×3 MIMO and 3×3 MIMO. [4] We collected a total of 483 hours of data for our experiments.

6.2.2 Evaluation of Respiration Pattern Recovery. Here, we compare the obtained respiration patterns with the ground-truths obtained from the respiration belt [12]. To quantitatively evaluate the sensed detailed respiration patterns over time, we adopt the metric of correlation coefficient to measure the similarity between the recovered patterns and the ground-truth patterns. Obviously, the optimal value for the correlation coefficient is 1, which would happen if the recovered pattern exactly matches the groundtruth pattern. A higher correlation value indicates a more accurate respiration pattern recovered. To clearly visualize the results, we show them from two aspects: one is the detailed respiration patterns in one of the above settings and the other is the average correlation coefficient under each of the various settings.

Respiration Pattern. Fig. 15 shows three recovered respiration patterns and the corresponding ground-truth patterns when three persons sit on the couch and MultiSense runs with 3×3 MIMO. While subject 1 is located at the boundary of Fresnel zone, subject 2 and subject 3 are located in the middle of the Fresnel zone. We see from Fig. 15 (a) and Fig. 15 (b) that MultiSense is able to separate two respiration patterns even when they are very similar to each other. Fig. 15 (c) shows that MultiSense can clearly capture irregular respiration patterns such as the apena that occurs in the time interval from the 9th to 22nd second. For all the experiments, no matter in which part of the Fresnel zones the targets are located, our system can always obtain clear respiration patterns, matching the groundtruth well. The experimental results show that the performance of MultiSense is not affected by target locations with respect to the transceivers.

Correlation Coefficient. Note that the sampling rate of the respiration belt is 100 Hz. Thus, to compute the correlation coefficient for each 30-second interval, we downsample the recovered respiration patterns from 200 Hz to 100 Hz. A natural question is what correlation value is considered good enough? We take the conclusion from [49] that an ideal recovery would show a correlation value higher than 0.9. Fig. 16 presents the average correlation coefficients in two-person, three-person, four-person and five-person scenarios with different numbers

[4] As shown in Sec. 4.4.1, to construct reference CSI in three-person scenario, we need at least four antenna pairs, which is inadequate with 1×3 MIMO.

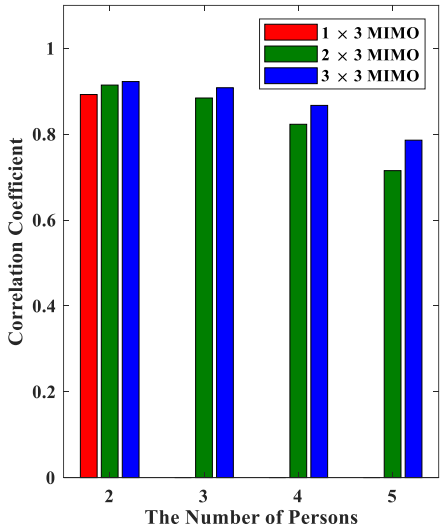

Fig. 16. The average correlation coefficient in each of the various settings.

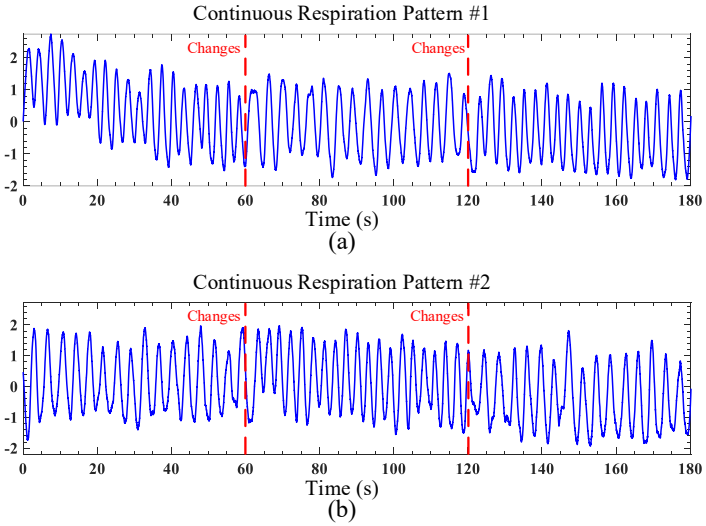

Fig. 17. An example of segment matching in two-person scenario. The ground-truth respiration rates in the first, second and third minute of subject 1 and subject 2 are {16 bpm, 20 bpm, 16 bpm} and {20 bpm, 16 bpm, 20 bpm}, respectively. The respiration rate of each subject changes at the 60th and 120th second.

of antennas. The figure shows that our MultiSense can always achieve a higher than 0.86 correlation coefficient in the presence of less than five persons when 3 x 3 MIMO is employed. We also notice that even with just 1×3 MIMO, MultiSense achieves a correlation coefficient of 0.892 in two-person scenario. We also observe that with the same number of antennas, the correlation coefficient slightly drops in the presence of more persons. The experimental results demonstrate the effectiveness and high accuracy of MultiSense in respiration pattern sensing in two-person, three-person and four-person scenarios. In theory, MultiSense is able to recover respiration patterns of eight persons with 3×3 MIMO. However, in our real-world experiments with 3×3 MIMO for five-person scenario, the correlation coefficient drops to 0.786. We believe the performance degradation with more than four persons is owing to the undermined linearity in signal mixture. As shown in Sec. 4.1.1, blind source separation using ICA relies on the assumption that the independent and non-Gaussian components are linearly mixed. However, in reality, the unavoidable hardware noise interferes with the signal separation process (e.g., the time-varying phase offset cancelation in Sec. 4.4.1), thus distorting the linearity of the mixture and accordingly degrading the performance of ICA.

6.2.3 Evaluation of Segment Matching. As the same in Sec. 6.2.2, we show the evaluation results from two aspects: one is the continuous respiration patterns for a period of three minutes in one of the settings described in Sec. 6.2.1 and the other is the overall accuracy of segment matching. We report the accuracy of segment matching by adopting the metric – the proportion of correct matches, which is the number of correct matches divided by the total number of matches. A match is considered correct if two segments that belong to the same person are matched while is incorrect if two segments that belong to different persons are matched together.

Respiration Pattern. Fig. 17 shows an example of segment matching when two subjects lie in the bed and MultiSense runs with 1×3 MIMO. The ground-truth natural respiration rates of subject 1 and subject 2 are

16 bpm and 20 bpm, respectively. In the first and the third minutes, both two subjects breath naturally. In the second minute, we ask subject 1 to increase his respiration rate to 20 bpm and ask subject 2 to decrease his respiration rate to 16 bpm. Based on our proposed method in Sec. 5.4, we need to identify the six respiration patterns belong to the same person and stitch them together. With this, we can combine six 30-second segments into one three-minute continuous pattern. As shown in Fig. 17 (a), the number of peaks in the first, second and third minute are 20, 16 and 20, respectively. This segment matching result obtains the respiration of subject 2. And the continuous pattern in Fig. 17 (b) corresponds to the respiration of subject 1. The respiration patterns of six segments are correctly stitched sequentially as expected. Note that, if we simply stitch two respiration patterns with similar respiration rates, we will get wrong results, i.e., one continuous pattern corresponds to a constant respiration rate of 16 bpm in three minutes and the other corresponds to that of 20 bpm in three minutes.

Accuracy. The overall average accuracy in two-person scenario is 98.5% while that in three-person scenario is 95.3%. And it drops to 90.7% in four-person scenario and 81.4% in five-person scenario. Unsurprisingly, the matching accuracy decreases with more persons. Note that when multiple persons keep moving, the consistency relationship between two adjacent segments in time degrades. However, we want to point out that the main focus of our work are those scenarios people are not moving all the time such as sleeping and watching TV.

6.3 Comparison with Previous Approaches

In this subsection, we compare MultiSense with four state-of-the-art WiFi-based multi-person respiration sensing approaches [4, 19, 36, 37] in terms of respiration rate accuracy.

6.3.1 *Baseline Approaches.*

- **TVS.**[5] TVS [19] uses the CSI amplitude for respiration sensing. TVS first utilizes the Hampel filter and a moving average filter to remove the outliers and high-frequency noise in the CSI amplitude. Then TVS employs a threshold-based method to select sub-carriers that have larger variance of CSI amplitude. At last, TVS takes FFT to transform the time-domain CSI amplitude to frequency domain. In the presence of N persons, N peaks are selected in the spectrum corresponding to the estimated respiration rates.
- **PhaseBeat.** PhaseBeat [36] uses the phase difference between CSI readings of two antennas for respiration sensing. DWT (Discrete Wavelet Transform) is first applied to remove high-frequency noises. Then PhaseBeat employs the root-MUSIC algorithm to estimate the frequency components in the denoised CSI phase difference.
- **TensorBeat.** TensorBeat [37] first removes the direct current component and high-frequency noises in CSI phase difference. Then, TensorBeat constructs the CSI tensor data using the calibrated CSI phase difference from different packets, sub-carriers and antenna pairs. Next, canonical polyadic decomposition is applied to obtain the periodical respiration signals. A peak detection method is further applied to estimate the respiration rates for multiple persons.
- **TR-BREATH.** TR-BREATH [4] first calculates the TRRS matrix using the calibrated CSI where the time-varying phase offset has been removed. Then, TR-BREATH performs temporal smoothing on the TRRS matrix to suppress the spurious estimates due to interference and noise. Next, root-MUSIC algorithm is applied to generate the respiration rate candidates. These candidates are further fed into the affinity propagation and cluster merging algorithms to obtain multi-person respiration rate estimation.

6.3.2 Experimental Results. We present the evaluation results from two aspects: one is the detection rate of each approach in two-person scenario versus the respiration rate difference between two persons, and the other is the mean absolute respiration rate error of each approach versus the number of persons. The detection rate is defined

[5]For convenience, we use the term "TVS" (Tracking Vital Signs) to represent the multi-person respiration sensing system proposed in [19].

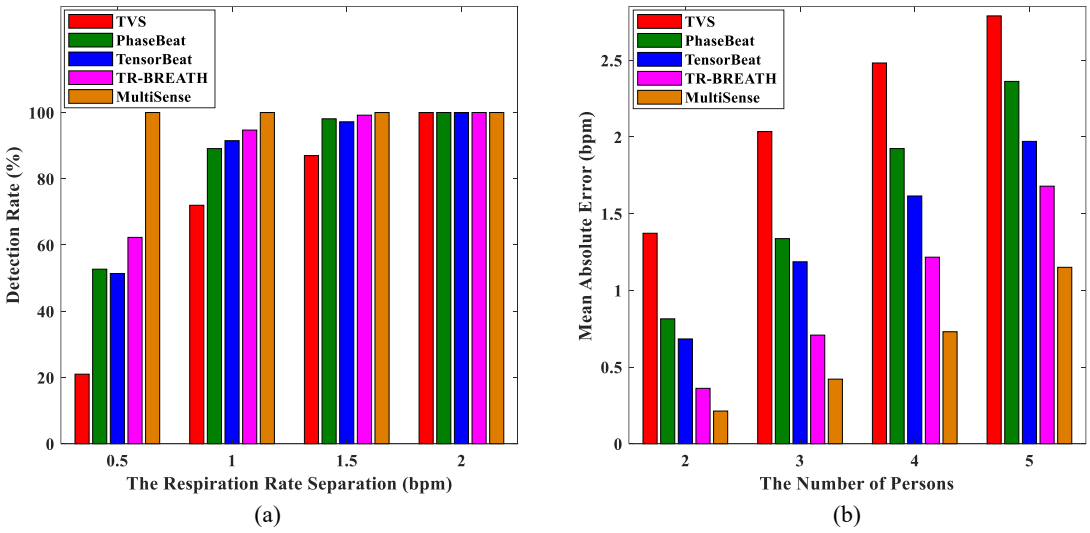

Fig. 18. The comparison with previous approaches from two aspects: (a) the detection rate versus the respiration rate separation in two-person scenario; (b) the mean absolute respiration rate error versus the number of persons.

as $\frac{N_{detected}}{N_{all}}$, where $N_{detected}$ is the number of CSI measurements the respiration rate is correctly estimated (i.e., the respiration rate error is less than 0.5 bpm), and N_{all} is the total number of collected CSI measurements [52].

Detection Rate. To investigate the impact of respiration rate difference of human subjects on the performance of each approach, we evaluate four respiration rate separations in two-person scenario, namely, 0.5 bpm, 1 bpm, 1.5 bpm and 2 bpm. We ask each subject to move to a location which is not a "blind spot" as described in Sec. 3. As shown in Fig. 18 (a), when two targets have dramatically different rates, all the systems work relatively well while MultiSense achieves the highest detection rate. However, when two targets have very similar respiration rates, MultiSense outperforms the other four systems.

Mean Absolute Respiration Rate Error. We plot the mean absolute errors of respiration rate for two-person, three-person, four-person and five-person scenarios in Fig. 18 (b). The figure shows that the mean absolute error of MultiSense is 0.21 bpm in two-person scenario and slightly increases to 0.42 bpm in three-person scenario. In contrast, in three-person scenario, TVS gets a mean absolute error of 2.04 bpm while PhaseBeat achieves an error of 1.33 bpm. The mean absolute errors for TensorBeat and TR-BREATH are 1.19 bpm and 0.71 bpm, respectively. A larger than 1 bpm error is usually unacceptable for real-life applications. In four-person scenario, only MultiSense achieves an error of less than 1 bpm, that is, 0.73 bpm. However, in five-person scenario, all approaches get errors larger than 1 bpm. Compared to our approach, there are two main reasons for the worse performance of the four baseline approaches. First, they fail to resolve two similar respiration rates. Second, they fail to estimate respiration rate of a person located at "blind spots". We also observe that in each scenario, PhaseBeat achieves a lower mean absolute error than TVS. We believe this is because CSI phase difference contains less noise than CSI amplitude and also the root-MUSIC algorithm is able to achieve a higher frequency resolution than FFT. The experimental results demonstrate that MultiSense outperforms the existing WiFi-based multi-person approaches in terms of respiration rate accuracy.

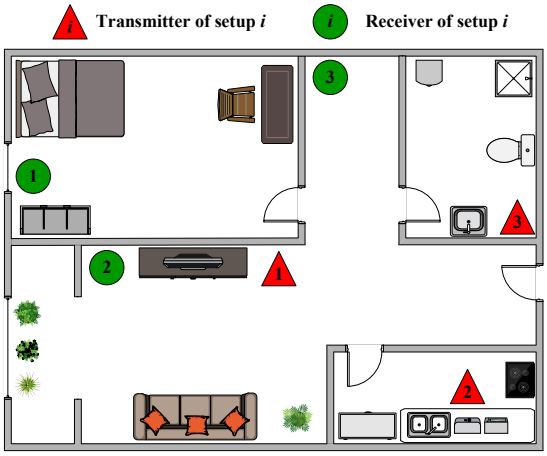

Fig. 19. The experiments for NLoS scenario are conducted in a typical home with three different transceiver setups, namely setup 1, 2 and 3.

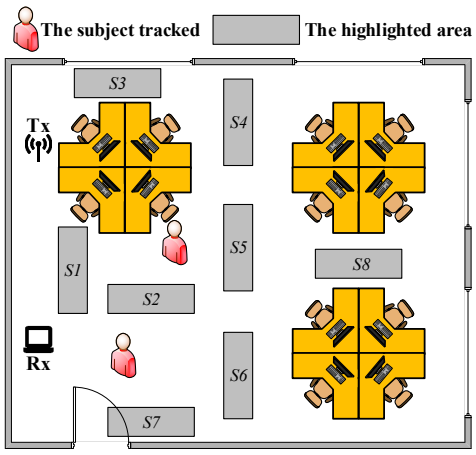

Fig. 20. Experiment settings for investigation of ambient motions. The subjects tracked breathe naturally while others walk in the highlighted area.

6.4 Performance in Challenging Scenarios

In this subsection, we evaluate MultiSense's performance in challenging scenarios, i.e., two subjects have NLoS path with the sensing devices. To evaluate the capability of MultiSense to recover respiration patterns in NLoS scenarios, we conduct experiments when the WiFi transmitter is placed in a different room from the receiver with a 8cm-thick wall in between. This is a realistic scenario in the home environment as the Wi-Fi access point can be located in a different room. As shown in Fig. 19, we conduct experiments in a typical home environment with three different transceiver setups, namely setup 1, 2 and 3. For each setup, the two subjects sit on the sofa or a chair.

For all the NLoS scenario experiments, the average correlation coefficient between recovered respiration pattern and groundtruth is 0.906, and the mean absolute error of respiration rate is 0.48 bpm. The results indicate the MultiSense has no problem to reliably recover respiration pattern even in the challenging NLoS scenario when there is a wall between the transmitter and receiver. For two persons, the mean absolute error of respiration rate in NLoS scenario is higher than that in LoS scenario (0.21 bpm) in Sec. 6.3. This is because the signal reflected off each target becomes even weaker after penetrating the wall.

6.5 Impact of Ambient Motion

To study the effect of ambient motion, i.e., the motion from others in the environment, we conduct additional experiments in two-person scenario in a large office room (7.5 m × 9 m), as shown in Fig. 20. In the experiments, we vary the distance between each subject and the LoS path from 1.5 m to 3.5 m at a step size of 0.5 m. [5] And We further vary the distance between the transmitter and receiver from 2.5 m to 5.5 m at a step size of 0.5 m. Besides the two subjects under respiration monitoring, we ask others to walk randomly in the eight highlighted areas in Fig. 20 (i.e., the areas indicated by $S1$-$S8$).

The experimental results show that MultiSense can tolerate ambient motions as long as the transceivers are far from these motions. For instance, given that the subjects tracked are 2.5 m away from the transceivers with a

[5]Here, the distance between a target and a transceiver pair is defined as the average distance from the target to the transmitter and the receiver [52].

LoS path length of 3.5 m, MultiSense gets a mean absolute respiration rate error of 0.23 bpm when the distance between the motion and the transceivers is larger than 4 m. In this case, if the motions occur at the distance of 3 m to the transceivers, the error increases to 2.27 bpm. To certain degree, when the distance between the motion and the transceivers increases, the performance is improved. We also notice that a shorter LoS path length helps to increase the tolerance of MultiSense to the ambient motions that occur at the same distance to the transceivers.

7 LIMITATION AND DISCUSSION

In this section, we present and discuss several limitations of this work.

7.1 Prior Knowledge of the Number of Persons

When we perform blind source separation using ICA method, the number of persons is required as the input. An inherent characteristic of ICA is that it cannot identify the actual number of source signals in general [30]. Basically, there are two ways to address the challenging issue when applying the ICA method. One is to employ a people counting system to estimate the number of targets in the environment, e.g., the WiFi-based system Wi-Count [47]. The other solution consists of two steps: (1) take different numbers as the input of ICA and generate multiple candidate results by performing ICA with each number; (2) select one from the candidate results based on a certain strategy, e.g., by maximizing the RER of the separated source signals.

7.2 Matching the Respiration Pattern to Each Person

In Sec. 5.4, for continuous monitoring, we manage to stitch multiple respiration patterns belong to the same person. However, we still have no idea whom each continuous respiration pattern belongs to. On one hand, even if we do not know whom each pattern belongs to, as long as one irregular respiration pattern is identified, we can still detect an abnormal respiration event. On the other hand, a potential method is to employ multi-dimensional information such as AoA (Angle of Arrival), ToF (Time of Flight) and Doppler shift [43] to locate each individual target in the environment.

8 CONCLUSION

This paper presents MultiSense, the first commodity WiFi based system to simultaneously recover detailed respiration patterns of multiple persons even they have similar respiration rates and are physically closely located. Different from prior work, MultiSense can extract not only the respiration rates but also the detailed respiration patterns over time, which is imperative to track abnormal respiration. Extensive experiments show that MultiSense is highly accurate with a mean absolute respiration rate error of 0.73 bpm even in the presence of four persons with only a pair of Wi-Fi transceivers.

The theory behind MultiSense is that the multi-person respiration sensing can be modeled as a BSS problem, which can be efficiently solved by ICA method. We believe such an idea provides a new perspective for contactless multi-person sensing. In addition, we propose a novel signal ratio method to cancel out the time-varying phase offset without distorting the linear mixture of WiFi CSI. We believe the proposed method contributes to not only fine-grained WiFi sensing, but also other wireless sensing scenarios where there is a lack of tight time synchronization between transmistter and receiver.

ACKNOWLEDGMENTS

This research is supported by the National Key Research and Development Plan (No. 2016YFB1001200), the PKU-Baidu Fund (No. 2019BD005) and Peking University Information Technology Institute (Tianjin Binhai).

REFERENCES

[1] Frédéric Abrard and Yannick Deville. 2005. A time–frequency blind signal separation method applicable to underdetermined mixtures of dependent sources. *Signal processing* 85, 7 (2005), 1389–1403.

[2] Fadel Adib, Hongzi Mao, Zachary Kabelac, Dina Katabi, and Robert C Miller. 2015. Smart homes that monitor breathing and heart rate. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 837–846.

[3] Magnus Borga and Hans Knutsson. 2001. A canonical correlation approach to blind source separation. *Report LiU-IMT-EX-0062 Department of Biomedical Engineering, Linkping University* (2001).

[4] Chen Chen, Yi Han, Yan Chen, Hung-Quoc Lai, Feng Zhang, Beibei Wang, and KJ Ray Liu. 2017. TR-BREATH: Time-reversal breathing rate estimation and detection. *IEEE Transactions on Biomedical Engineering* 65, 3 (2017), 489–501.

[5] Jin Chen, Per Jönsson, Masayuki Tamura, Zhihui Gu, Bunkei Matsushita, and Lars Eklundh. 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote sensing of Environment* 91, 3-4 (2004), 332–344.

[6] Ward Cheney and David Kincaid. 2009. Linear algebra: Theory and applications. *The Australian Mathematical Society* 110 (2009).

[7] Vivian L Clark and James A Kruse. 1990. Clinical methods: the history, physical, and laboratory examinations. *Jama* 264, 21 (1990), 2808–2809.

[8] 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.

[9] David E Godberg. 1989. Genetic algorithms in search. *Optimization, and Machine Learning* (1989).

[10] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2011. Tool release: Gathering 802.11 n traces with channel state information. *ACM SIGCOMM Computer Communication Review* 41, 1 (2011), 53–53.

[11] John A Hartigan and Manchek A Wong. 1979. Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28, 1 (1979), 100–108.

[12] Wern Kam, Waleed Soliman Mohammed, Gabriel Leen, Mary O'Keeffe, Kieran O'Sullivan, Sinead O'Keeffe, and Elfed Lewis. 2017. Compact and low-cost optical fiber respiratory monitoring sensor based on intensity interrogation. *Journal of Lightwave Technology* 35, 20 (2017), 4567–4573.

[13] 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.

[14] Joshua Chong Yue Lai, Ying Xu, Erry Gunawan, Eric Chern-Pin Chua, Arash Maskooki, Yong Liang Guan, Kay-Soon Low, Cheong Boon Soh, and Chueh-Loo Poh. 2010. Wireless sensing of human respiratory parameters by low-power ultrawideband impulse radio radar. *IEEE Transactions on Instrumentation and Measurement* 60, 3 (2010), 928–938.

[15] Ronald Ley. 1994. Breathing and the psychology of emotion, cognition, and behavior. In *Behavioral and psychological approaches to breathing disorders*. Springer, 81–95.

[16] Changzhi Li, Victor M Lubecke, Olga Boric-Lubecke, and Jenshan Lin. 2013. A review on recent advances in Doppler radar sensors for noncontact healthcare monitoring. *IEEE Transactions on microwave theory and techniques* 61, 5 (2013), 2046–2060.

[17] Feng Lin, Chen Song, Yan Zhuang, Wenyao Xu, Changzhi Li, and Kui Ren. 2017. Cardiac scan: A non-contact and continuous heart-based user authentication system. In *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*. 315–328.

[18] Chen Liu, Jie Xiong, Lin Cai, Lin Feng, Xiaojiang Chen, and Dingyi Fang. 2019. Beyond Respiration: Contactless Sleep Sound-Activity Recognition Using RF Signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–22.

[19] 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*. 267–276.

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

[21] Xuefeng Liu, Jiannong Cao, Shaojie Tang, Jiaqi Wen, and Peng Guo. 2015. Contactless respiration monitoring via off-the-shelf WiFi devices. *IEEE Transactions on Mobile Computing* 15, 10 (2015), 2466–2479.

[22] C Lowanichkiattikul, M Dhanachai, C Sitathanee, S Khachonkham, and P Khaothong. 2016. Impact of chest wall motion caused by respiration in adjuvant radiotherapy for postoperative breast cancer patients. *SpringerPlus* 5, 1 (2016), 144.

[23] Wei Lu and Jagath C Rajapakse. 2003. Eliminating indeterminacy in ICA. *Neurocomputing* 50 (2003), 271–290.

[24] Carlo Massaroni, Daniel Simões Lopes, Daniela Lo Presti, Emiliano Schena, and Sergio Silvestri. 2018. Contactless monitoring of breathing patterns and respiratory rate at the pit of the neck: A single camera approach. *Journal of Sensors* 2018 (2018).

[25] L Robert Mogue and Boerje Rantala. 1988. Capnometers. *Journal of clinical monitoring* 4, 2 (1988), 115–121.

[26] Rajalakshmi Nandakumar, Shyamnath Gollakota, and Nathaniel Watson. 2015. Contactless sleep apnea detection on smartphones. In *Proceedings of the 13th annual international conference on mobile systems, applications, and services*. 45–57.

[27] Kun Qian, Chenshu Wu, Zheng Yang, Yunhao Liu, and Kyle Jamieson. 2017. Widar: Decimeter-level passive tracking via velocity monitoring with commodity Wi-Fi. In *Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. 1–10.

[28] Bhaskar D Rao and KV Sl Hari. 1989. Performance analysis of root-MUSIC. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 37, 12 (1989), 1939–1949.

[29] Alejandro Rodríguez-Molinero, Leire Narvaiza, Jorge Ruiz, and César Gálvez-Barrón. 2013. Normal respiratory rate and peripheral blood oxygen saturation in the elderly population. *Journal of the American Geriatrics Society* 61, 12 (2013), 2238–2240.

[30] V David Sánchez A. 2002. Frontiers of research in BSS/ICA. *Neurocomputing* 49, 1-4 (2002), 7–23.

[31] Andrew G Stove. 1992. Linear FMCW radar techniques. In *IEE Proceedings F (Radar and Signal Processing)*, Vol. 139. IET, 343–350.

[32] James D Taylor. 2018. *Ultra-wideband radar technology*. CRC press.

[33] 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.

[34] Tianben Wang, Daqing Zhang, Yuanqing Zheng, Tao Gu, Xingshe Zhou, and Bernadette Dorizzi. 2018. C-FMCW based contactless respiration detection using acoustic signal. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2018), 1–20.

[35] 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.

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

[37] Xuyu Wang, Chao Yang, and Shiwen Mao. 2017. TensorBeat: Tensor decomposition for monitoring multiperson breathing beats with commodity WiFi. *ACM Transactions on Intelligent Systems and Technology (TIST)* 9, 1 (2017), 1–27.

[38] Lu Wei and Jagath C Rajapakse. 2000. A neural network for undercomplete independent component analysis.. In *ESANN*. Citeseer, 123–128.

[39] Stefan Wiesner and Ziv Yaniv. 2007. Monitoring patient respiration using a single optical camera. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2740–2743.

[40] Chenshu Wu, Zheng Yang, Zimu Zhou, Xuefeng Liu, Yunhao Liu, and Jiannong Cao. 2015. Non-invasive detection of moving and stationary human with wifi. *IEEE Journal on Selected Areas in Communications* 33, 11 (2015), 2329–2342.

[41] Chenshu Wu, Feng Zhang, Yusen Fan, and KJ Ray Liu. 2019. RF-based inertial measurement. In *Proceedings of the ACM Special Interest Group on Data Communication*. 117–129.

[42] Zhung-Han Wu, Yi Han, Yan Chen, and KJ Ray Liu. 2015. A time-reversal paradigm for indoor positioning system. *IEEE Transactions on Vehicular Technology* 64, 4 (2015), 1331–1339.

[43] Yaxiong Xie, Jie Xiong, Mo Li, and Kyle Jamieson. 2019. mD-Track: Leveraging multi-dimensionality for passive indoor Wi-Fi tracking. In *The 25th Annual International Conference on Mobile Computing and Networking*. 1–16.

[44] Yaxiong Xie, Yanbo Zhang, Jansen Christian Liando, and Mo Li. 2018. SWAN: Stitched Wi-Fi ANtennas. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*. 51–66.

[45] 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.

[46] Yanni Yang, Jiannong Cao, and Xiulong Liu. 2019. ER-Rhythm: Coupling Exercise and Respiration Rhythm Using Lightweight COTS RFID. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 4 (2019), 1–24.

[47] Yanni Yang, Jiannong Cao, Xuefeng Liu, and Xiulong Liu. 2018. Wi-count: Passing people counting with COTS WiFi devices. In *2018 27th International Conference on Computer Communication and Networks (ICCCN)*. IEEE, 1–9.

[48] Yanni Yang, Jiannong Cao, Xuefeng Liu, and Kai Xing. 2018. Multi-person sleeping respiration monitoring with COTS WiFi devices. In *2018 IEEE 15th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*. IEEE, 37–45.

[49] Shichao Yue, Hao He, Hao Wang, Hariharan Rahul, and Dina Katabi. 2018. Extracting multi-person respiration from entangled rf signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (2018), 86.

[50] Vicente Zarzoso and Pierre Comon. 2009. Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size. *IEEE Transactions on neural networks* 21, 2 (2009), 248–261.

[51] 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.

[52] 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.

[53] Daqing Zhang, Hao Wang, and Dan Wu. 2017. Toward centimeter-scale human activity sensing with Wi-Fi signals. *Computer* 50, 1 (2017), 48–57.

[54] 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.

[55] 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* (2019).

[56] 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.