

Metasurface-Enabled Multi-Target WiFi Sensing

Yimiao Sun, Qunyan Zhou, Yulong Chen, Jiaming Gu,
Qiang Cheng, *Senior Member, IEEE*, Yuan He, *Senior Member, IEEE*

Abstract—As an emerging technology, WiFi sensing has garnered widespread attention in recent years. However, due to limitations in WiFi bandwidth and hardware capacity, multi-target sensing with WiFi remains an unresolved issue. In this paper, we present SLINGSHOT, a first-of-its-kind approach for multi-target WiFi sensing enabled by metasurface. SLINGSHOT exploits a metasurface’s ability of beam scanning to periodically scan among targets in a high frequency, so their sensing signals can be separated in the time domain. However, the WiFi-initiated and the metasurface-initiated signals will interfere with each other, which could significantly degrade the sensing accuracy. Moreover, the WiFi devices and the metasurface operate in a distributed manner, so the inherent clock offset and clock drift among them make it difficult to accurately separate the sensing signals of the targets. To address the above challenges, SLINGSHOT leverages the metasurface’s ability of phase shifting to cancel interfering signals and designs a passive clock synchronization scheme to synchronize the metasurface and WiFi devices. We implement SLINGSHOT and evaluate it under various settings. Results show that SLINGSHOT can sense up to 8 targets. Respectively compared to the state-of-the-art methods, SLINGSHOT has 6.75 % higher mean accuracy in activity recognition and 58.34 % lower mean error in respiration monitoring.

Index Terms—Metasurface, WiFi Sensing, Multi-Target

I. INTRODUCTION

Past decades have witnessed the rapid development of wireless sensing technologies [1]–[4]. Benefiting from the ubiquitous WiFi infrastructures, WiFi sensing has become the mainstream and attracts increasing attention [5]–[7]. One can estimate the location by mapping the WiFi Channel State Information (CSI) or Received Signal Strength Indicator (RSSI) with the corresponding location [8]–[10]. One can also recognize human activities by analyzing the Doppler frequency shift of reflected signals [11], [12].

In spite of the positive progress, most of the existing works can function well with a single target, but suffer severe performance degradation in multi-target sensing scenarios. The main reason is that the signals reflected by multiple targets superimpose at a WiFi Rx, which usually doesn’t

(Co-primary authors: Yimiao Sun and Qunyan Zhou; Corresponding authors: Yuan He and Qiang Cheng.) This work is supported by the National Natural Science Foundation of China (No. 62425207, No. 62225108, No. 62288101 and No. U21B2007), the Fundamental Research Funds for the Central Universities (2242022k60003), the Jiangsu Province Frontier Leading Technology Basic Research Project (BK20212002), the Jiangsu Provincial Scientific Research Center of Applied Mathematics (BK20233002), and the Jiangsu Science and Technology Research Plan (BK20243028).

Y. Sun, Y. Chen, J. Gu, Y. He are with the School of Software and BNRIst, Tsinghua University, Beijing, China. E-mail: {sym21, yl-chen22, gujm24}@mails.tsinghua.edu.cn, heyuan@tsinghua.edu.cn.

Q. Zhou and Q. Cheng are with the School of Information Science and Engineering, Southeast University, Nanjing, China. E-mail: {qyzhou, qiangcheng}@seu.edu.cn.

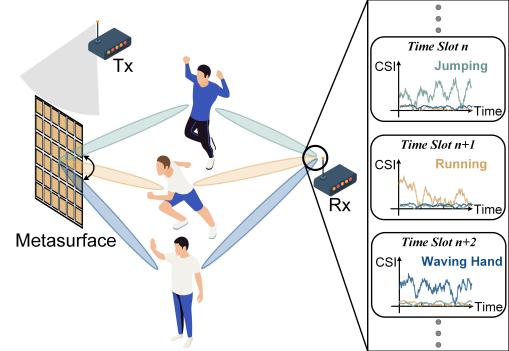


Figure 1: The high-level principle of SLINGSHOT.

have sufficient signal bandwidth and hardware capacity to distinguish those signals, and therefore the targets. However, there is a widespread demand for multi-target sensing, and even associating each target’s location with its sensing results, for example, location-aware applications like smart homes.

To deal with the above issue, some works exploit signal processing techniques like blind source separation [13] and virtual sample generation [11] to extract sensing information from the superimposed signals, assuming that the signals from different targets exhibit distinct features or that these features are known based on specific domain knowledge.

Another set of existing works proposes to increase the available sensing information. For instance, using multiple WiFi access points (APs) or a large antenna array for multi-target sensing [5], [12], [14], [15], however, at the cost of complex control over WiFi systems. Besides, the near-field domination effect is exploited to distinguish each target from the others by using its adjacent smartphones [16], [17], while the sensing range depends on the capacity of smartphones.

Recent works [18]–[20] exploit metasurface to perform beamforming for different targets to separate their sensing signals. Even so, these methods typically rely on dedicated incident signals (*i.e.*, single-tone [18], [20] or FMCW [19]), which aren’t compliant with WiFi standards.

We in this paper present SLINGSHOT, a novel metasurface-enabled approach for multi-target WiFi sensing. Leveraging its inherent capability of beamforming, a metasurface is employed to reflect the incident WiFi signals to a specified direction. When the direction of reflection keeps switching, the beam of metasurface periodically scans different directions at a very high frequency, namely beam scanning. In each beam scanning cycle, the targets will reflect the beams at different time, so the beams can be respectively mapped with the targets. After a WiFi Rx receives the reflected signals, the signals can be accordingly mapped to the beams, and in turn, the targets, as illustrated in Fig. 1. Thus, we can separate the signals in the

time domain, and achieve multi-target sensing.

To implement the above idea, we tackle the following critical challenges:

1) *Signal interference*: Incorporating a metasurface into a WiFi sensing system generates strong interference between the WiFi-initiated and the metasurface-initiated signals. From the perspective of a WiFi Rx, both types of signals contain information associated with the sensing targets, but they can't be effectively utilized when they are superimposed and may even degrade the sensing accuracy.

To mitigate such interference, the metasurface in SLINGSHOT keeps switching the beamforming signal between two opposite phase shifts, namely 0 to π , at a high rate. While the state of the interfering signals keeps stable in a short period, subtracting the signals of opposite phase shifts produces sensing signals with an enhanced signal-to-interference-plus-noise ratio (SINR).

2) *Clock synchronization*: It is important for a WiFi Rx to accurately map the received signals to the beams from the metasurface, because different beams generally correspond to different sensing targets. However, as the WiFi devices and the metasurface operate distributedly, their unsynchronized clocks may lead to the inconsistent mapping of the beams to the targets in different beam scanning cycles. When the metasurface's beam scans at a high frequency, this problem becomes very critical in ensuring the accuracy of sensing.

To deal with this problem, we search the values of the clock offset and the clock drift between the metasurface and the WiFi devices to find the correct mapping. This process is executed at the Rx side via numerical searching without the metasurface for complex computation.

Our contributions can be summarized as follows:

1) SLINGSHOT is the first-of-its-kind metasurface-enabled multi-target WiFi sensing approach. Based on the metasurface's beam scanning ability, SLINGSHOT translates multi-target sensing into the multi-signal separation problem in the time domain and efficiently resolves it at the WiFi Rx.

2) SLINGSHOT tackles the unique challenges in operating the sensing system of WiFi devices and metasurface. By incorporating the key designs like passive clock synchronization and interference mitigation, SLINGSHOT is robust and highly applicable under practical settings.

3) We implement SLINGSHOT, including the metasurface with aforementioned capacity of signal manipulation, and evaluate its performance under various settings. Results show that SLINGSHOT can sense up to 8 targets. Respectively compared to the state-of-the-art methods, SLINGSHOT has 6.75 % higher mean accuracy in activity recognition and 58.34 % lower mean error in respiration monitoring.

II. SLINGSHOT OVERVIEW

This section presents a brief overview of SLINGSHOT. A WiFi sensing system generally consists of two WiFi devices (Tx and Rx) and a metasurface. Compared to a WiFi device, a metasurface has very limited computational capacity. Therefore, the design of SLINGSHOT leverages the metasurface's ability in signal manipulation, while leaving most of the signal processing and computation on the WiFi Rx side.

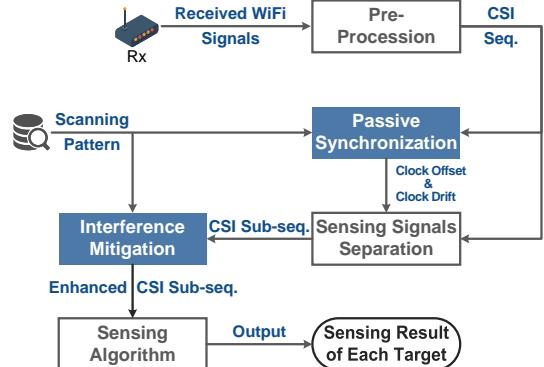


Figure 2: The workflow of SLINGSHOT on the WiFi Rx.

Fig. 2 illustrates the workflow of SLINGSHOT from the perspective of the WiFi Rx. All the WiFi signals are initially sent by the WiFi Tx. A portion of those signals incident on the metasurface will be manipulated and reflected. So the signals received by the WiFi Rx include the signals directly from the WiFi Tx, the signals only reflected by targets, and the signals reflected by the metasurface and targets.

When receiving such a mixture of signals, the Rx first pre-processes them to get the raw CSI sequence. Then the passive clock synchronization (Sec. IV) is executed on the Rx, which calculates *clock offset* and *clock drift* between the metasurface and the WiFi devices. Once the clocks of devices are synchronized, samples in the CSI sequence can be mapped to the corresponding beams. Then the CSI sequence will be separated into multiple sub-sequences *w.r.t.* the beamforming directions. After that, the interference mitigation scheme cancels the interference out from the CSI sub-sequences, getting their SINR enhanced (Sec. V). Finally, the CSI sub-sequences are fed into the sensing algorithm to output the sensing results of different directions (Sec. V-C).

III. METASURFACE DESIGN

In this section, we introduce the configuration of the metasurface used in SLINGSHOT, the basic principle of sensing signal separation with the metasurface and the analysis of the metasurface's sensing resolution.

A. Metasurface Configuration

Similar to most existing metasurface [21]–[25], SLINGSHOT adopts the 2D planar structure composed of periodically arranged sub-wavelength elements, called meta-atoms. Fig. 3(a) illustrates the schematic diagram of the proposed metasurface structure with an enlarged meta-atom. This meta-atom comprises two feeding lines interconnected with an irregular hexagon patch through two PIN diodes. As Fig. 3(b) shows, our proposed metasurface contains 8×8 meta-atoms and is sized 20 cm \times 22 cm. The working band of this metasurface is centered at 5.2 GHz, and its parameters are specified to $l_1=9.55$ mm, $l_2=7.26$ mm, $l_3=7.81$ mm, $l_4=9.30$ mm, $l_5=9.89$ mm, $l_6=6.04$ mm, $l_7=5.59$ mm, $l_8=10.24$ mm, and $p=25$ mm. The substrate material is F4B, whose relative permittivity $\epsilon_r = 2.65$ and dielectric loss tangent $\tan \sigma =$

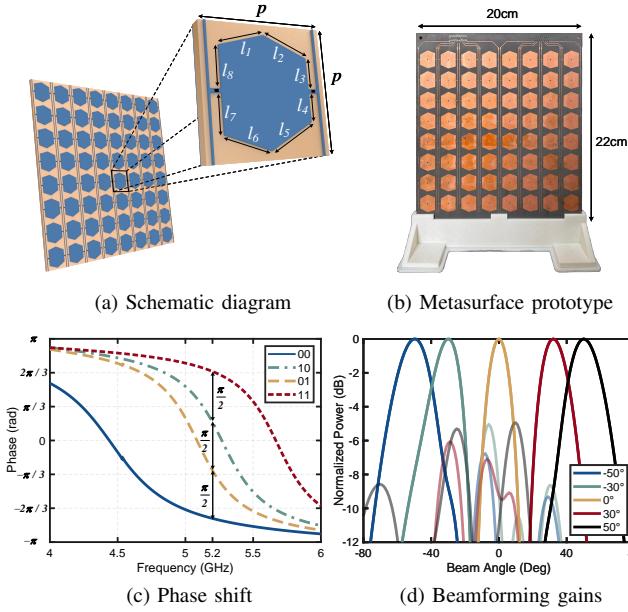


Figure 3: The metasurface design in SLINGSHOT.

0.001. A via hole is incorporated to penetrate the substrate, connecting the top and bottom layers in order to serve as the analog and digital ground. This via hole is arranged at the center of the substrate and the patch.

Our proposed metasurface has two crucial abilities to enable SLINGSHOT, phase shifting and beam scanning.

1) *Phase shifting*: The meta-atoms can introduce different phase shifts to the reflected signals, which is a key capability for beamforming and wavefront shaping. Positive-Intrinsic-Negative (PIN) diodes are usually incorporated into the meta-atom design to enable this capability. By controlling the on/off state of each PIN diode, the current distribution excited by the incident signals is restructured, thereby altering the phase of the reflected signals. In our design, two PIN diodes are used in each meta-atom to perform 2-bit phase shifting (*e.g.*, 0, $\frac{\pi}{2}$, π , $\frac{3\pi}{2}$). The metasurface's performance of phase shifting *w.r.t.* frequencies is shown in Fig. 3(c), where the phase gradient between adjacent states is around $\frac{\pi}{2}$ at 5.2 GHz.

2) *Beam scanning*: Based on the function of phase shifting, the metasurface can further implement beam scanning by adjusting the configuration of each meta-atom. For a metasurface with $M \times N$ meta-atoms, the phase shift introduced by a meta-atom at row m , column n can be denoted as $\Phi(m, n)$. Assume there is a target located in the direction of $D = (\theta, \varphi)$, where θ and φ are the elevation and azimuth angles from the metasurface to the target, so the far-field radiation pattern towards the target can be formulated as

$$F(\theta, \varphi) = \sum_{m=1}^M \sum_{n=1}^N \Psi(m, n, \theta, \varphi), \quad (1)$$

$$\Psi(m, n, \theta, \varphi) = e^{-i\{\Phi(m, n) + kd \sin \theta [(\bar{m} - \frac{1}{2}) \cos \varphi + (\bar{n} - \frac{1}{2}) \sin \varphi]\}}, \quad (2)$$

where k is the wave number and d is the distance between two meta-atoms [26].

Based on Eq. (1), the function of beamforming and beam scanning can be achieved by appropriately configuring and

changing the phase distribution across the metasurface. A quick validation is conducted to examine the proposed metasurface's ability of beamforming. Results in Fig. 3(d) demonstrate the obvious gain in each beam direction, indicating the enhancement of the sensing signals in these directions. By quickly configuring the metasurface to switch the beamforming direction, the metasurface can achieve fast beam scanning. In the implementation of SLINGSHOT, the metasurface's beam switching keeps a high frequency of 1000 Hz, the same as the frequency of Tx's sending packets, which ensures the continuous sensing for each target.

B. Sensing Signal Separation

By employing the beam scanning capability of the metasurface, SLINGSHOT can translate multi-target sensing into the multi-signal separation problem in the time domain. Supposing L targets locate within the sensing area, the composite signal received at the WiFi AP can be expressed as

$$s(t) = \sum_{i=1}^L \Gamma_i x(t - \tau_i) + n(t), \quad (3)$$

where Γ_i is the reflection coefficient of the i -th target, τ_i is its associated signal propagation delay and $n(t)$ is noise. In this case, it's hard for the AP to separate each target's sensing information from the superimposed received signal.

In contrast, a metasurface with beam scanning capability can impose a time-varying radiation pattern $F(\theta, \phi, t)$ to periodically steer its beam toward different sensing targets over a sensing period \mathcal{T} . Consequently, the received signal becomes

$$s(t) = \sum_{i=1}^L F(\theta_i, \phi_i, t) \Gamma_i x(t - \tau_i) + n(t). \quad (4)$$

As $F(\theta_i, \phi_i, t)$ attains its peak value only when the beam is directed at the i -th target, the received signal over a sensing period can be partitioned into L signal segments:

$$s_l(t) = s(t + (l-1)\mathcal{T}), \quad t \in [0, \mathcal{T}], \quad l = 1, \dots, L, \quad (5)$$

where $s_l(t)$ contains predominantly the reflection from the l -th target. Within each signal segment, the beamforming gain toward the intended target yields

$$s_l(t) = \Gamma_{i(l)} x(t - \tau_{i(l)}) + \delta \sum_{j \neq i(l)} \Gamma_j x(t - \tau_j) + n_l(t), \quad (6)$$

with $\delta \ll 1$ quantifying the suppressed sidelobe contributions from other targets. This time-domain signal separation forms the basis of metasurface-enabled multi-target sensing.

C. Sensing Resolution

To fully exploit the sensing capabilities of the proposed metasurface, we characterize its sensing resolution, the minimum distinguishable angular separation between adjacent targets. This metric significantly determines SLINGSHOT's capability to resolve multiple objects in space. The sensing resolution is derived from two fundamental factors and expressed as [27], [28]: $\theta_{res} = \sqrt{\theta_H^2 + \Delta\theta_Q^2}$, where θ_H is the

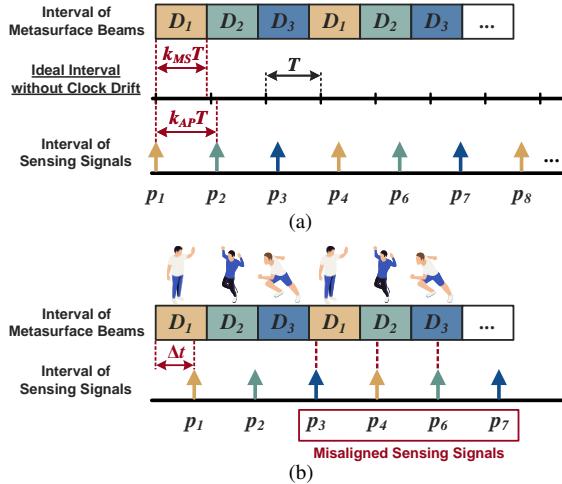


Figure 4: The illustration of the asynchronization between the metasurface and the WiFi devices. (a) The clock drift between the metasurface and the WiFi devices leads to biased operation intervals. (b) The clock offset (Δt) and the clock drift may lead to the wrong mapping between sensing signals and beams (red dashed lines).

intrinsic Half-Power Beam Width (HPBW) of the metasurface and $\Delta\theta_Q$ is the additional beam broadening induced by phase quantization effects. θ_H can be further denoted as

$$\theta_H \approx \frac{0.886\lambda}{\sqrt{(Md)^2 + (Nd)^2}} \cdot \frac{180^\circ}{\pi}, \quad (7)$$

where λ is the wavelength of incident signals. By substituting respective values into Eq. (7), we can obtain $\theta_H \approx 10.3^\circ$.

The beam broadening $\Delta\theta_Q$ defines the phase shift resolution of metasurfaces and originates from the 2-bit phase control of the metasurface. For discrete phase states with $\frac{\pi}{2}$ steps, the maximum phase error per meta-atom is bounded by $\frac{\pi}{4}$. Assuming uniform error distribution over $\delta\Phi \in [-\frac{\pi}{4}, +\frac{\pi}{4}]$, the phase error variance can be represented as

$$\langle\delta\Phi^2\rangle = \text{Var}(\delta\Phi) = \frac{(\frac{\pi}{4})^2}{3}. \quad (8)$$

This variance can be translated to angular broadening based on the following formulation

$$\Delta\theta_Q \approx \frac{\sqrt{\langle\delta\Phi^2\rangle}}{kd} \cdot \frac{180^\circ}{\pi}, \quad (9)$$

where k is the wave number of incident signals. By substituting respective values into Eq. (9), we can have $\theta_Q \approx 3.0^\circ$.

The resultant sensing resolution is $\theta_{res} = \sqrt{10.3^2 + 3.0^2} \approx 10.8^\circ$. This suggests that SLINGSHOT can reliably distinguish targets separated by around 10° , which is sufficient for typical indoor sensing scenarios. For sensing resolution enhancement in more fine-grained applications, one can increase the aperture (*i.e.*, more meta-atoms) of the metasurface to reduce θ_H and employ higher-bit (ideally continuous) phase control to minimize $\Delta\theta_Q$ (*i.e.*, increase the phase shift resolution).

IV. PASSIVE CLOCK SYNCHRONIZATION

Accurate separation of sensing signals necessitates the correct mapping of the beams and the sensing signals. However,

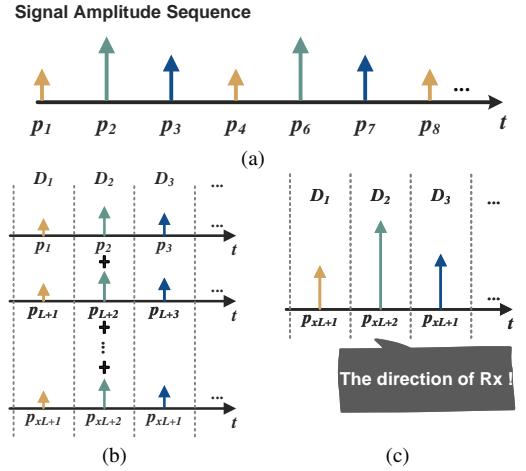


Figure 5: SINR calculation for a pair $(\Delta T, \Delta t)$. (a) The amplitude information of the WiFi signals is extracted to form a sequence. (b) The amplitude sequence is cut by a window and matched to different direction bins according to the beam scanning pattern of the metasurface. (c) The amplitude values in each direction bin are accumulated to identify the LoS signal.

the metasurface and the WiFi devices aren't in sync, which usually leads to biased mappings and wrong sensing results. In this section, we propose a passive clock synchronization scheme to address this issue.

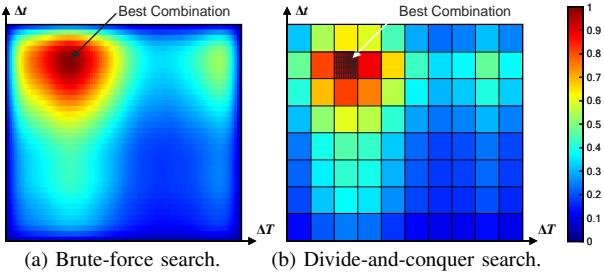
A. Problem Statement

As we mentioned in Sec. III, Tx's sending packets and the metasurface's beam switching keep a high frequency. However, the metasurface and the WiFi devices operate in a distributed manner and don't share the same system clock. Thus, the clock offset and clock drift between them may lead to the incorrect mapping of the beams and the targets, and in turn, the wrong sensing results, as Fig. 4(a) shows. Even though we set the same interval T of switching beams and sending WiFi signals, the clock drifts existing in the metasurface and WiFi devices will result in $T_{AP} = \alpha_{AP}T$, $T_{MS} = q\alpha_{MS}T$, $q \in N^*$, where $\alpha_{AP} \approx 1$ and $\alpha_{MS} \approx 1$ are clock drift coefficients of the WiFi devices and the metasurface, which lead to biased intervals T_{AP} and T_{MS} , respectively; q is the magnification, indicating the number of WiFi packets sent in each direction.

If we don't tackle this problem, they will accumulate over time. Moreover, the clock offset of the metasurface and the WiFi devices also influences the mapping relation. Affected by these two factors, the sensing signals could be mapped to the wrong beams and, consequently, wrong targets, as Fig. 4(b) illustrates. Thus, to ensure the right mapping and further accurate sensing results, we should obtain two values, the clock drift (denoted as ΔT) and the clock offset (denoted as Δt) between the metasurface and the WiFi devices.

B. Parameter Searching

1) *Basic idea:* To synchronize the metasurface and the WiFi devices, we propose a passive clock synchronization scheme based on searching for the values of the pair $(\Delta T, \Delta t)$.

Figure 6: Heatmaps of normalized $SINR_{LoS}$ in the searching space.

This process is executed at the Rx via numerical searching, without involving the metasurface for complex computation.

After receiving the sensing signals, we enumerate the value of the pair $(\Delta T, \Delta t)$ to obtain different potential mappings between the beams and the sensing signals. In each mapping, the amplitude values of the signals corresponding to the same beam are accumulated, then the highest value will be selected. The beam direction of the highest value is considered as the direction of the Rx, because the LoS (line-of-sight) signals between the metasurface and the Rx exhibit a stronger energy. Then, we calculate the SINR of the LoS signals based on the accumulated values. The pair leading to the highest SINR is considered the most accurate.

2) *Detailed algorithm:* As Fig. 5 shows, we first extract the amplitude information of the WiFi signals to obtain an amplitude sequence. Then, according to the beam scanning settings of the metasurface, this sequence will be cut by a window with the length of L and matched to L different direction bins, where L is the number of WiFi packets sent in a round of beam scanning ($L = 3$ in Fig. 5 for illustration). Then, we accumulate the amplitude values in each direction bin (depicted in Fig. 5(b)), as

$$A^l = \sum_{x=0}^{X-1} A_{xL+l}, l \in [1, L], \quad (10)$$

where A^l is the accumulation of values and X is the total number of values, in l -th direction bin, respectively.

Next, the direction bin bearing the highest accumulated value (denoted as A_{max}) is regarded as the direction of the Rx, as illustrated in Fig. 5(c). Finally, we calculate SINR of the LoS signals between Rx and metasurface in this pair as

$$SINR_{LoS} = \frac{A_{max}}{\sum A^j - A_{max}}. \quad (11)$$

Each pair of $(\Delta T, \Delta t)$ is associated with a specific value of $SINR_{LoS}$. Thus, we traverse possible values of the pair with a predefined searching step to identify the most accurate one that leads to the maximum $SINR_{LoS}$.

We conduct a preliminary experiment to validate our passive synchronization scheme. In this experiment, we set $T = 1$ ms and $q = 1$. The searching step is set to $10 \mu s$ for both ΔT and Δt , as a trade-off between accuracy and computational latency¹. Consequently, the range of Δt is established as

¹We set $q = 1$ in this preliminary experiment for illustration, while $q = 2$ in our practical settings based on our design in Sec. V. This difference doesn't affect the effectiveness of our algorithm, as long as $p \in N^*$.

[-0.5 ms, 0.5 ms]. Based on our experimental results and empirical observations, ΔT in our system consistently falls within [-0.2 ms, 0.2 ms]. Therefore, we designate this as the range for ΔT . Fig. 6(a) illustrates the heatmap of the normalized $SINR_{LoS}$ calculated based on different values of $(\Delta T, \Delta t)$. This result demonstrates that we can find the pair bearing the maximum $SINR_{LoS}$, indicating that the devices can be well synchronized with the value of the pair $(\Delta T, \Delta t)$.

C. Searching Space Reduction

Note that the clock drift will result in accumulated error over time, so the synchronization should be executed intermittently. In that case, the time-consuming brute-force searching is unacceptable for our time-sensitive system. To address this problem, we exploit the divide-and-conquer algorithm to significantly reduce the time cost.

Specifically, as for Fig. 6(a), we divide the searching space into 80 sub-spaces, all of which are sized 5×10 . Subsequently, we compute the $SINR_{LoS}$ value for the pair $(\Delta T, \Delta t)$ that is centered at each sub-space, and use this value to represent the value of the respective sub-space. This process yields an 8×10 matrix, as Fig. 6(b) shows. Following this, a coarse-grained search is conducted to select the sub-space with the maximum value of $SINR_{LoS}$ in this matrix. A fine-grained search is then performed within the selected sub-space to ultimately determine the pair $(\Delta T, \Delta t)$ exhibiting the globally maximum $SINR_{LoS}$. By doing so, the searching space can be reduced from 4000 (40×100) to 130 ($50 + 80$), saving nearly $30 \times$ time cost.

Note that SLINGSHOT performs passive synchronization at a fixed interval (e.g., every few seconds) to maintain continuous accuracy. Although clock drift may vary gradually with environmental factors (e.g., temperature), such variations typically occur over tens of minutes to hours. Thus, over intervals of a few seconds, the drift can be viewed as constant, and there is no need to perform the search at every time slot (i.e., 1 ms). Our evaluation in Sec. VI-C demonstrates no obvious improvement with searching at each time slot.

V. INTERFERENCE MITIGATION

After clock synchronization, sensing signals can be separated and correspond to different targets, respectively. However, the sensing signals may be heavily interfered. In SLINGSHOT, we exploit metasurface's phase shifting ability to mitigate the interference.

A. Signal Model

Considering that a target is in the beamforming direction, the beamforming signals (i.e., the sensing signals) are originally radiated from the Tx, then reflected by the metasurface and a target, and finally received by Rx. Thus, the beamforming signal contains the valuable sensing information of the target in the beamforming direction. We denote the beamforming signal as S_{MS} , whose energy [26], [29] is

$$P_{MS} = \frac{P_T G_T G_R G_u G_t p^2 \lambda^6 \cdot |\Gamma_t|^2 |\Gamma_{MS}|^2 F'(\theta_T, \varphi_T) F(\theta_t, \varphi_t)}{(4\pi)^6 r_1^2 r_2^2}, \quad (12)$$

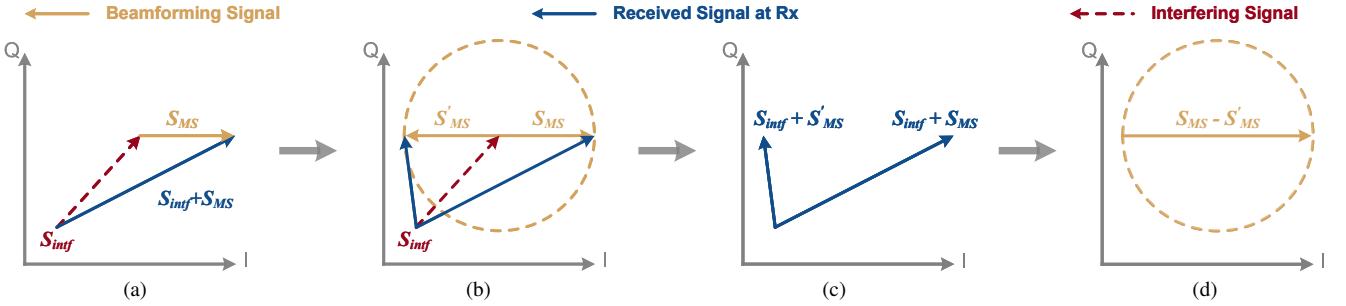


Figure 7: The phasor diagram of interference mitigation process in SLINGSHOT. (a) The received signal at Rx (blue line) is composed of the beamforming signal (S_{MS} in the yellow line) and the interfering signal (S_{intf} in the red line). (b) When metasurface performs fast phase shifting between 0 and π , S_{intf} can be viewed as relatively stable in such a short time, while the direction of S_{MS} in the I/Q diagram reverses to be S'_{MS} , leading to (c) different received signals at Rx. (d) By taking vector subtraction for the two received signals, the beamforming signal S_{MS} is doubled while the interfering signal S_{intf} is mitigated.

$$F'(\theta_T, \varphi_T) = \left(\sum_{m=1}^M \sum_{n=1}^N \frac{\Psi(m, n, \theta_T, \varphi_T)}{r_{n,m}^T} \right)^2, \quad (13)$$

where P_T is the transmission power of the Tx; G_T , G_R , G_u and G_t are the respective gains of the Tx, Rx, a meta-atom and the target; Γ_t and Γ_{MS} are the respective reflection coefficient of the target and the metasurface; λ is signal's wavelength; $F'(\theta_T, \varphi_T)$ and $F(\theta_t, \varphi_t)$ are metasurface's radiation pattern towards the Tx and the target, respectively; $r_{n,m}^T$, r_1 and r_2 are the respective distance between the Tx and each meta-atom, the metasurface and the target, and the target and the Rx.

As the beamforming operation concentrates signal energy to result in a higher P_{MS} , the target in the beamforming direction can dominate the channel variation. Therefore, its sensing information is easier to be distinguished compared to the targets in other directions. However, as Fig. 7(a) shows, the received signal S_{RX} is not only composed of the beamforming signal S_{MS} , but also the interfering signal S_{intf} . The main components of S_{intf} are the direct (S_{direct}) and reflected ($S_{multipath}$) signals between the Tx and Rx, *i.e.*, $S_{intf} \approx S_{direct} + S_{multipath}$.

B. Mitigating Interference

We exploit metasurface's ability of fast phase shifting to mitigate interfering signals. Our observation is that when the metasurface reverses states of all meta-atoms (*e.g.*, from 0 to π , and vice versa), the direction of S_{MS} reverses in the phasor diagram while maintaining the beamforming direction and gain unchanged. The reversed signal can be formulated as

$$S'_{MS} = -S_{MS}. \quad (14)$$

Accordingly, received signals before and after reversion are

$$S_{RX} = S_{MS} + S_{direct} + S_{multipath}, \quad (15)$$

$$S'_{RX} = S'_{MS} + S_{direct} + S'_{multipath}. \quad (16)$$

As S_{direct} is unchanged, subtracting S_{RX} and S'_{RX} will yield

$$\Delta S_{RX} = S_{RX} - S'_{RX} \quad (17)$$

$$= 2S_{MS} + (S_{multipath} - S'_{multipath}). \quad (18)$$

Our observation on the multipath signals ($S_{multipath}$ and $S'_{multipath}$) is two-fold. 1) With WiFi packet rates on the order

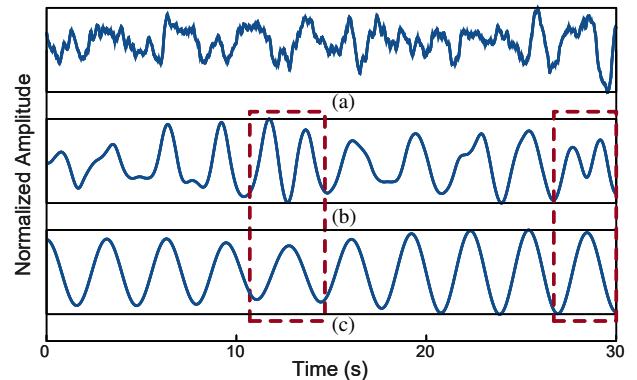


Figure 8: Respiration waveform recovery for one target in the multi-target sensing scenario. (a) Superimposed signal obtained without metasurface. (b) Respiration signal obtained with metasurface but without interference mitigation. (c) Enhanced respiration signal obtained with fast phase shifting to mitigate interfering signals.

of 1000 Hz, human motion is very slow, so the induced variations in multipath over the 1 ms reversion interval are very small. 2) The beamforming signal S_{MS} dominates the energy of the received signal, whereas multipath components are comparatively weak. In that case, the variation in the multipath signals $S_{multipath} - S'_{multipath}$ will exhibit little impact on the subtracted signal, which can be further represented as $\Delta S_{RX} = S_{RX} - S'_{RX} \approx 2S_{MS}$.

Thus, the beamforming signal containing the valuable sensing information is extracted from the received signal, with theoretically doubled energy, as Fig. 7(d) shows. Note that interfering signals experiencing multiple reflections (*e.g.*, signals reflected by the metasurface, the target and other targets consecutively) are ignored in SLINGSHOT, because they go through severe attenuation and only cause negligible impact on the sensing signals.

We conduct a proof-of-concept experiment to validate the proposed interference mitigation method. In the experiment, we try to monitor the respiration of 5 invited volunteers. Fig. 8(c) demonstrates the recovered respiration waveform of one volunteer after interference mitigation. It obviously contains much less interfering signals compared to that without using interference mitigation (Fig. 8(b)) and without metasurface (Fig. 8(a)). Note that although we can recover the respiration waveform without interference mitigation (Fig. 8(b)), wrong

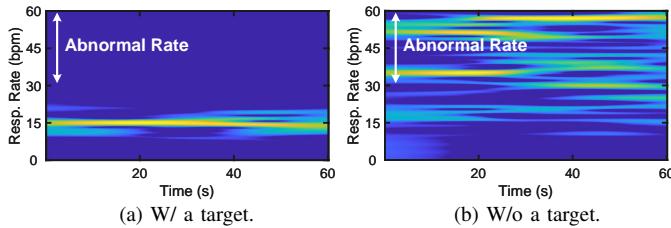


Figure 9: Respiration spectrogram obtained in respiration monitoring.

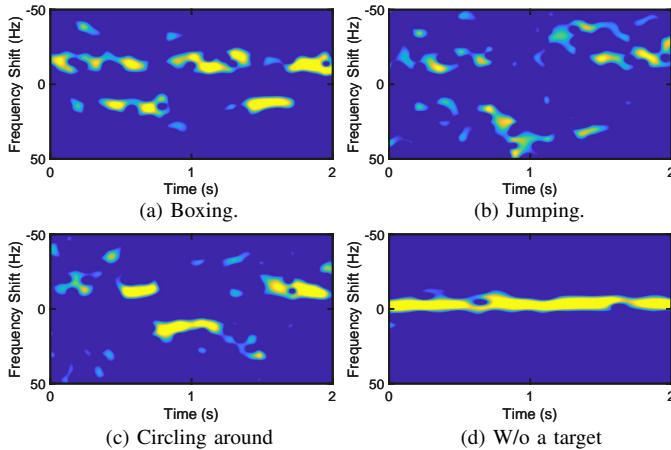


Figure 10: DFS of different activities.

sensing information may be obtained (in the red dashed boxes) which degrades the sensing accuracy.

C. Sensing Algorithm

After interference mitigation, Rx can obtain the enhanced sensing signals in each beamforming direction. Therefore, different sensing algorithms can be applied to output the corresponding sensing results. Here, we demonstrate the sensing performance of SLINGSHOT w.r.t. two typical sensing tasks: respiration monitoring and activity recognition.

1) Respiration Monitoring: Leveraging the enhanced sensing signals, SLINGSHOT can estimate the respiration rates of the targets breathing simultaneously by using Variational Modal Decomposition (VMD) [17]. As Fig. 9(a) shows, the respiration rate of the target shows high SINR in the respiration spectrogram. Moreover, if there is no target in a direction, the spectrogram resembles that shown in Fig. 9(b), where many peaks exceed the normal range of a person's respiration rate (*i.e.*, 10-30 bpm [13]). This observation can be exploited for SLINGSHOT to detect the presence of targets.

2) *Activity Recognition*: Fig. 10(a)-(c) illustrates the spectrograms of Doppler frequency shift (DFS) when multiple targets perform different activities. It's obvious that the respective DFS shows higher SINR than the background noise, which can contribute to higher recognition accuracy (discussed in Sec. VI). SLINGSHOT also employs DFS to detect the targets, based on the fact that the DFS is nearly zero if there is no target, as Fig. 10(d) shows. Note that if a target keeps stable, we can borrow existing methods for discrimination [30], [31]. As a multi-target sensing approach, SLINGSHOT hasn't regarded it as a main concern. In the future, we may integrate these solutions into the design.

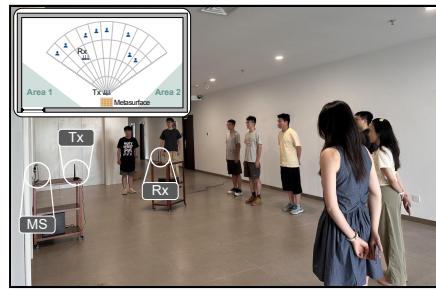


Figure 11: Experimental scenarios and deployment.

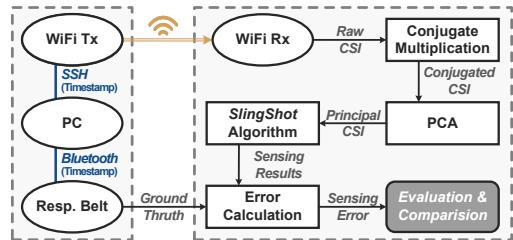


Figure 12: The procedure of data collection and processing.

VI. IMPLEMENTATION AND EVALUATION

A. Experimental Methodology

1) Experiment Setup: Our metasurface is connected to National Instruments (NI) PXIE-1082 [32] and the digital I/O module, which controls the on/off states of PIN diodes on metasurface to achieve phase shifting and beam scanning. In SLINGSHOT, the proposed metasurface is designed at 5.2 GHz and performs periodic beam scanning from -50° to $+50^\circ$, with 10° step (11 beamforming directions). The time interval of changing the beamforming direction is 2 ms, 1 ms for phase state 0 and 1 ms for phase state π .

The working band of SLINGSHOT is 5.19 GHz-5.21 GHz (*i.e.*, Channel 40) and we collect CSI data with PicoScenes [33]. We employ commercial AX210 NICs at both the Tx and Rx. PicoScenes running on the Tx configures the NIC to inject WiFi packets, while the NIC on the Rx is set to monitor mode to capture and parse those WiFi packets to extract raw CSI. The procedure of CSI processing is shown in Fig. 12. To address carrier frequency offset (CFO) and sampling frequency offset (SFO), we perform conjugate multiplication of the CSI from the Rx’s two antennas, and then apply principal component analysis (PCA) to extract the principal components of the CSI [12]. Then, the resulting CSI is fed into the SLINGSHOT algorithm for multi-target sensing.

2) *Sensing Tasks*.: We evaluate SLINGSHOT's ability of multi-target sensing with two realistic sensing applications, activity recognition and respiration monitoring.

For activity recognition, we first train a deep learning model for activity recognition, and then, we use this model to evaluate the performance of multi-target activity recognition. Specifically, we invite 8 volunteers to perform 6 types of daily human activities, including boxing (BX), standing up (SU), sitting down (SD), waving hand (WH), circling around (CA) and jumping (JP). Each volunteer individually repeats each type of activity for 200 times and completes an activity in 3 seconds. Accordingly, we collect and label 9600 CSI samples

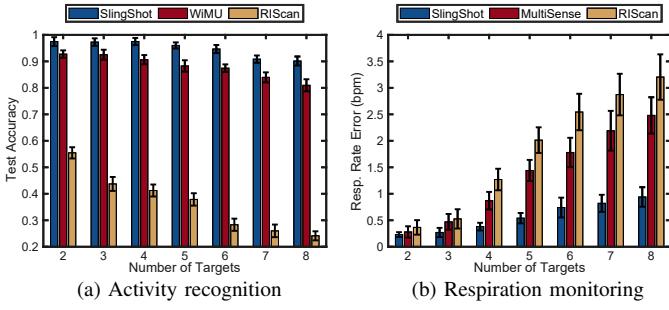


Figure 13: The overall performance of SLINGSHOT.

to build the dataset. Then, we augment the dataset size to 19200 by adding random noise into the data, which is a widely used method for data augmentation [34], [35]. We use the augmented dataset to train an RNN model for activity recognition. The model demonstrates a classification accuracy of 95.74%. Next, we invite multiple volunteers to simultaneously perform different activities and deploy SLINGSHOT to extract CSI samples corresponding to each target, as Fig. 11 shows. The extracted CSI samples are fed into the aforementioned RNN network to test the performance of SLINGSHOT’s on multi-target activity recognition.

For respiration monitoring, we invite volunteers to stand at different distances and beamforming directions to breathe simultaneously, and use Go Direct respiration belts [36] to obtain ground truth. We deploy SLINGSHOT to extract CSI samples corresponding to each target. Then, we employ VMD to extract the respiration waveform of each target. As Fig. 12 shows, the respiration belts are connected to the PC via Bluetooth, and the ground-truth results are synchronized with the WiFi sensing results using timestamps.

3) *Baselines*: We choose three state-of-the-art methods as baselines. In the case of activity recognition, we compare SLINGSHOT with WiMU [11]. Its core idea is to search for the possible combination of signals corresponding to multiple known activities, and then compare the detected signal with the combined signal to identify activities performed simultaneously. In the case of respiration monitoring, we compare SLINGSHOT with MultiSense [13]. With the input of the number of targets, MultiSense models the multi-target respiration sensing as a blind source separation problem, then uses independent component analysis to distinguish the respiration of each person. Besides, we also compare SLINGSHOT with RIScan [37] in both of the two sensing tasks. RIScan employs the metasurfaces and the beam scanning capability to localize multiple WiFi devices concurrently. We make our best effort to re-implement the baseline methods for a fair comparison.

B. Overall Performance

We first evaluate the general performance of SLINGSHOT when performing the two sensing tasks mentioned above in multi-target scenarios. Then, we demonstrate the accuracy of SLINGSHOT to identify targets.

1) *Activity Recognition*: We invite 8 volunteers to perform different activities, in different directions and at different distances with respect to the metasurface. None of the targets

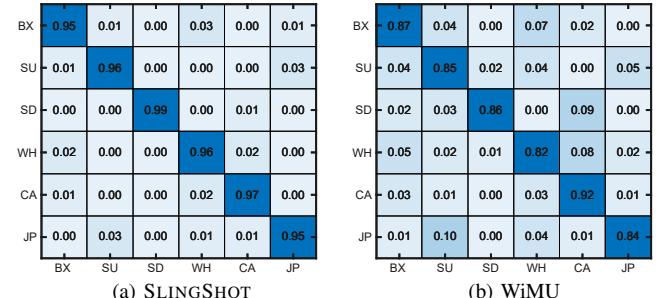


Figure 14: Confusion matrices on activity recognition.

is in the direction of Rx. Fig. 13(a) compares the test accuracy of SLINGSHOT with WiMU and RIScan when the number of targets in the sensing area increases from 2 to 8, with the comparison of confusion matrices in Fig. 14. Results show that the mean accuracy of SLINGSHOT is 94.84 %, which is 6.75 % higher than that of WiMU and 58.16 % higher than that of RIScan. RIScan’s high error stems from its assumption of zero clock drift between the WiFi APs and the metasurface. This assumption will severely misattribute the WiFi signal and metasurface beam to the wrong sensing target, and lead to significant sensing errors. Note that the accuracy of SLINGSHOT is higher than that of WiMU across all numbers of targets, and keeps the accuracy over 90 %, and even over 97 % when targets are less than 5. In comparison, the accuracy of WiMU is less than 90 % when there are over 5 targets. These results evidently confirm the effectiveness of SLINGSHOT in multi-target activity recognition. We note that the accuracy of SLINGSHOT declines obviously when there are 7 or 8 targets. The reason is that, in these cases, there are always 3 targets adjacent to each other among the 11 beamforming directions (from -50° to $+50^\circ$), which will lead to an obvious interference. This issue can be easily addressed by increasing meta-atoms in the metasurface.

2) *Respiration Monitoring*: We employ a similar methodology to evaluate the performance of SLINGSHOT in respiration monitoring. Fig. 13(b) compares the error of respiration rate estimation of SLINGSHOT with MultiSense and RIScan. The mean error of SLINGSHOT is 0.56 bpm, which is 58.34 % lower than that of MultiSense and 69.36 % lower than that of RIScan. We find that the mean error of SLINGSHOT is always lower no matter the number of the targets. Besides, SLINGSHOT demonstrates less than 1 bpm error even when 8 targets breathe simultaneously. Note that the error that is larger than 1 bpm is usually considered unacceptable [13], while the mean error of MultiSense exceeds this threshold when over 4 targets breathe simultaneously. Fig. 15 illustrates an example of respiration waveforms captured by SLINGSHOT and the corresponding groundtruth waveforms in an eight-target scenario. The result shows that the captured waveforms by SLINGSHOT (red dashed lines) closely track the true breathing waveforms (blue solid lines). Even though there are small temporal offsets, they do not significantly affect the accuracy of the estimated respiration rates.

3) *Target Detection*: Fig. 16 shows the accuracy of one-shot target detection in activity recognition (using DFS spectrogram) and respiration monitoring (using respiration spectro-

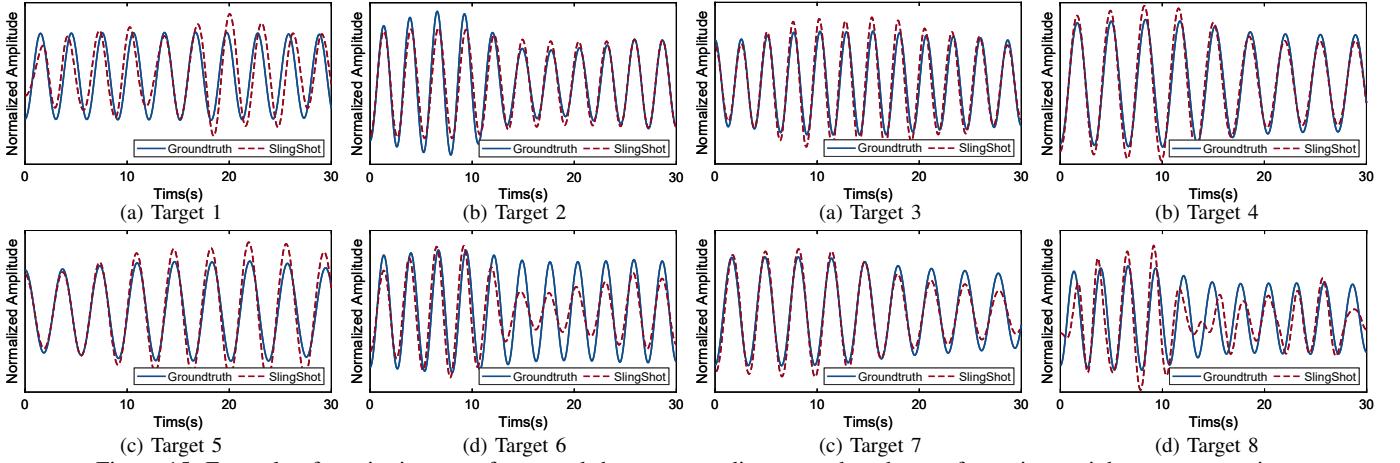


Figure 15: Example of respiration waveforms and the corresponding ground-truth waveforms in an eight-target scenario.

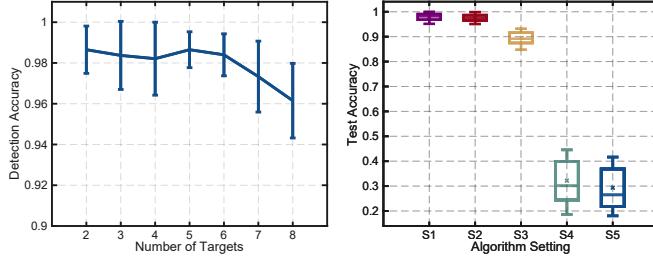


Figure 16: Target detection accuracy of SLINGSHOT.

Figure 17: Factorial experiment on the design of SLINGSHOT.

gram). When the number of targets increases from 2 to 8, the mean detection accuracy only slightly decreases from above 98 % to 96 %. The accuracy can be further improved to near 100 % by performing detection multiple times.

C. Factorial Experiment

In this section, we conduct a factorial experiment to analyze the impact of the two components, passive synchronization (*module 1*) and interference mitigation (*module 2*), on SLINGSHOT’s performance and their relative importance. The factorial experiment is conducted under 5 different settings: S1 (with both two modules and search at every time slot), S2 (with both two modules and search at a 3-second interval), S3 (only with *module 1* and search at a 3-second interval), S4 (only with *module 2*), and S5 (without any sub-module).

Fig. 17 reports the results of this factorial experiment under a four-target scenario. With both sub-modules employed, the mean accuracy of using the S1 and S2 is around 97.50 %. This suggests that there is no obvious improvement with searching at every time slot. In other words, SLINGSHOT can well synchronize the distributed devices with searching at a fixed interval. When SLINGSHOT doesn’t use the metasurface to mitigate interfering signals, the mean accuracy will decline to 89.26 % (S3) because of the decreased SINR. What’s worse, if sensing devices aren’t in sync, the mean accuracy decreases drastically to around 30 % (S4 and S5).

The above results indicate that interference mitigation has a more significant impact on system performance.

D. Impacting Factors

1) Distance between Targets: Although SLINGSHOT mitigates interfering signals by employing metasurface, some targets may also interfere with each other if they are very close. To examine the impact of distance between targets, we invite 2 volunteers to perform activities in different beamforming directions, where the angle difference between the two directions is from 10° to 50° . Results in Fig. 18(a) show that the accuracy of activity recognition is reduced to around 90 % if the angle difference is only 10° (*i.e.*, the two targets are in the adjacent beamforming directions). Otherwise, the accuracy will be over 95 %. The reason behind these results is that the beam width of the metasurface is approximately 10° , so the beamforming signal can’t well isolate two targets in the adjacent beamforming directions. To address this issue, we can simply increase the number of meta-atoms to reduce the beam width, which we leave for our future work.

2) Directions of Targets: As the beamforming angle increases from 0° to $\pm 50^\circ$, the beamforming gain of the metasurface may reduce [38], [39], which is implied in Fig. 3(d). We carry out experiments to see how the sensing performance is influenced by varying beamforming angles. 2 volunteers are invited to perform activities from $\pm 10^\circ$ to $\pm 50^\circ$, and an individual experiment is carried out in 0° . Fig. 18(b) reports the results that as the beamforming angle increases, the recognition accuracy decreases as the angle varies from 0° to -50° and $+50^\circ$. Note that the accuracy in all directions is above 90 %, indicating the robustness of SLINGSHOT.

3) Time-varying Propagation and Multipath: In real-world operational conditions, the sensing area may be subject to time-varying propagation due to moving persons. Also, different levels of multipath caused by indoor reflectors usually exist. To evaluate the performance and robustness of SLINGSHOT under such conditions, we conduct two sets of experiments. In the first set, 6 volunteers simultaneously perform activities while additional volunteers move within each of Area 1 or Area 2 (see Fig. 11), thereby creating a time-varying propagation environment. As shown in Fig. 18(c), the recognition accuracy declines from 94.67 % to 91.05 % or 90.37 % when moving persons are present in only one area or both of the two areas, respectively. In the second set, static

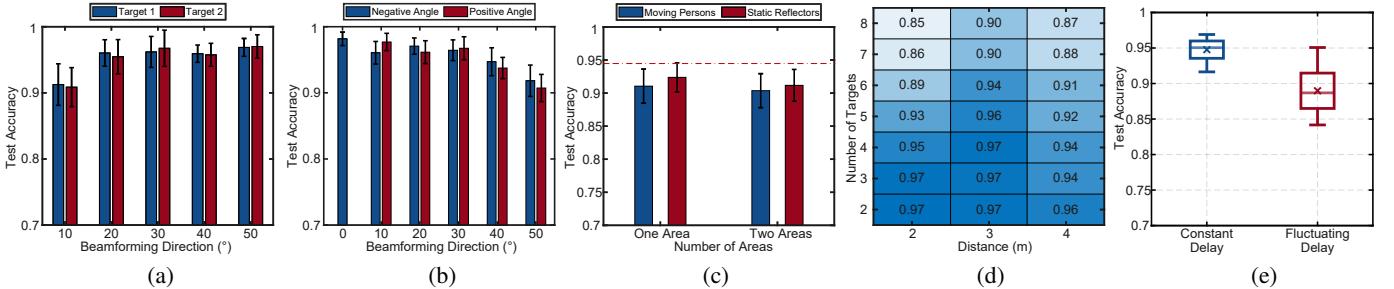


Figure 18: Impacting factors concerning the performance of SLINGSHOT. (a) Impact of the distance between targets. (b) Impact of the directions of targets. (c) Impact of the static reflectors and moving persons. (d) Impact of the distance between the metasurface and targets. (e) Impact of the WiFi packet delay.

reflectors (*i.e.*, desks and chairs) are placed in Area 1 or Area 2 to introduce different levels of multipath. Fig 18(c) shows that the recognition accuracy decreases to 92.38 % or 91.17 % with reflectors in only one area or both of the two areas, respectively. The results show that SLINGSHOT’s performance may be influenced by time-varying propagation and multipath, yet the sensing accuracy does not degrade drastically.

4) *Distance between Metasurface and Targets*: We evaluate the impact of the distance between the metasurface and the targets on the recognition accuracy. The distance is set to 2 m, 3 m and 4 m in each group of experiments. Results Fig. 18(d) show that the mean accuracy at the distance of 2 m declines heavily as the number of targets increases. This is because more targets at this distance will cause a very small interval between targets. Thus, it is more likely to result in mutual interference and low accuracy. Things get better when the distance is set to 3 m where the mean accuracy where the mean accuracy across different numbers of targets is over 90 %. However, as the distance further increases to 4 m, the recognition accuracy declines. The reason is that the distance is too long to ensure sufficient SINR.

5) *WiFi Packet Delay*: The transmissions of dedicated WiFi sensing devices usually exhibit less variability and uncertainty than those of devices operating in live networks. However, they could still be affected by Carrier-Sense Multiple Access with Collision Avoidance (CSMA/CA) and random backoff, which may lead to fluctuating packet delays. SLINGSHOT’s design is robust to such fluctuation: following the method discussed in Sec. IV-B and Eq. (10), we can use each packet’s timestamp to map each packet to the corresponding direction bin. Then similarly, we can compute the average amplitude for each bin to find the direction of the Rx, and map each WiFi packet to each sensing target.

We conduct experiments under fluctuating packet delays between 1 ms to 10 ms and report the results in Fig. 18(e). Results show that in a six-target scenario, the mean accuracy decreases from 94.67% under constant delay to 88.91% under fluctuating delay. Although the performance is affected, the accuracy does not drop drastically.

VII. DISCUSSION

In this section, we discuss practical issues concerning the applicability and efficacy of SLINGSHOT.

1) *Phase-Shift switching time of the metasurface*: The phase-shift switching time of our proposed metasurface is under 1 μ s, which is less than 1 % of the interval between consecutive phase shifts. This sub- μ s latency sufficiently supports high-frequency periodic beam scanning and has a negligible impact on SLINGSHOT’s performance.

2) *Co-Located Sensing Targets*: Multiple targets may be co-located within the same beam in practical scenarios, which could lead to wrong sensing results or cause some targets’ sensing information to be lost. Two further enhancements to SLINGSHOT can mitigate this problem. First, increasing the number of meta-atoms and employing higher-bit phase control will improve sensing resolution (*i.e.*, reducing the beam width). Second, deploying multiple metasurfaces to perform beam scanning from different angles to separate targets.

VIII. RELATED WORK

The distinctive feature of SLINGSHOT lies in performing multi-target WiFi sensing with the metasurface. We briefly summarize existing works in the related fields, metasurface-assisted wireless sensing and multi-target WiFi sensing.

1) *Metasurface-Assisted Wireless Sensing*: Metasurfaces are typically designed as 2D artificial structures, which can flexibly and accurately manipulate wireless signals [21], [22], [25], [26], [29]. Recently, the promise of integrating metasurface into IoT technologies has been proven [23], [37], [39]–[41]. Many works offer novel solutions for wireless sensing, spanning WiFi, mmWave and acoustic sensing. Metasight [42] proposes a passive metasurface in the mmWave band to help detect objects in occluded regions. AMS [43] enhances AoA estimation accuracy by using a passive acoustic metasurface to significantly boost the beamforming gain. RFlens [39] uses a metasurface to resteer WiFi signals to sense the respiration rate of the target in NLoS scenarios. MetaSense [41] achieves gesture recognition with a single frequency based on a dedicated metasurface.

Different from the existing works that focus on novel metasurface designs or control schemes, SLINGSHOT addresses the coordination between WiFi sensing devices and the metasurface as well as the multi-signal separation in the time domain for multi-target sensing. Moreover, SLINGSHOT is compatible with the most widely used metasurface designs (*e.g.*, 2-bit phase shifting) and control mechanism (*i.e.*, beam scanning), thereby ensuring broad applicability.

2) *Multi-Target WiFi Sensing*: As WiFi sensing technologies gain increasing attention, many works try to adopt different methods for multi-target sensing [5], [11], [13], [16], [17], mainly including signal processing based, distributed WiFi devices based, and metasurface based methods.

Signal processing based methods. Traditional methods exploit signal processing techniques to achieve multi-target sensing. MultiSense [13] models the multi-target respiration sensing as a blind source separation problem to separate the reparation information of each target. WiMU [11] searches the possible combination of signals corresponding to known gestures and then compares the detected signal with the combined signal to identify gestures performed simultaneously. Even though these methods can extract the sensing signals from the superimposed signals, they assume that the sensing signals from different targets exhibit distinct features or that these features are known based on some specific domain knowledge.

Distributed WiFi devices based methods. To accurately map sensing signals with targets, much of the existing works try to increase the available sensing by using multiple distributed WiFi devices. For instance, exploiting distributed WiFi APs or APs with large antenna arrays for multi-target sensing [5], [12], [14], [15], but at the cost of complex control over WiFi systems. Besides, MUSE-Fi [16] and M²-Fi [17] exploit the near-field effect to enable multi-target sensing with the distributed smartphones. Specifically, they assume that every target is close to a smartphone, so the near-field channel variation caused by the target will significantly overwhelm variations caused by others. However, the sensing range of these methods significantly depends on the smartphones.

Metasurface based methods. Recent works [18]–[20], [37] propose to leverage metasurfaces to dynamically manipulate the wireless channel for multi-target sensing. The work closest to ours is RIScan [37], which employs two metasurfaces and their beam scanning capability to localize multiple WiFi devices concurrently. Compared to RIScan, the core contribution of SLINGSHOT is twofold. 1) *Sensing algorithm*: SLINGSHOT achieves accurate target detection and activity sensing without imposing targets to carry any devices. By contrast, RIScan does not provide detection capability and requires computing sensing results at each user device. 2) *Metasurface control strategy*: SLINGSHOT ensures a robust coordination between WiFi devices and the metasurface without relying on a centralized controller. However, RIScan imposes the WiFi devices and the metasurface to be under the same controller, which would constrain scalability and may be difficult to implement with commercial WiFi devices.

Besides, MetaBreath [18] and Li *et al.* [20] exploit metasurface to generate harmonic signals with different frequencies for different targets to distinguish their sensing signals. However, both of the approaches require the incident signals to be single-tone signals, which aren't compliant with the WiFi standards. MetaPhys [19] requires the incident signals to be FMCW signals, which are also not compliant with the WiFi standards. Compared with the aforementioned works, SLINGSHOT is fully compliant with the WiFi standards.

IX. CONCLUSION

We presented SLINGSHOT, the first-of-its-kind WiFi multi-target sensing technique that decently separates the sensing targets in the time domain. SLINGSHOT innovates a multi-target sensing technique in that it fully exploits the fast phase shifting and beam scanning capabilities of the metasurface to enhance the sensing signals of each target and achieves the passive clock synchronization between the metasurface and the WiFi APs. Through extensive experiments, we demonstrate that SLINGSHOT offers better performance than state-of-the-art solutions when performing two representative sensing tasks, activity recognition and respiration monitoring. We believe that SLINGSHOT has significant potential to support ubiquitous multi-target sensing across a variety of applications.

REFERENCES

- [1] Y. He, J. Liu, M. Li, G. Yu, J. Han, and K. Ren, “Sencom: Integrated sensing and communication with practical wifi,” in *Proc. of ACM MobiCom*, 2023.
- [2] Y. Zeng, D. Wu, J. Xiong, E. Yi, R. Gao, and D. Zhang, “Farsense: Pushing the range limit of wifi-based respiration sensing with csi ratio of two antennas,” in *Proc. of ACM IMWUT*, 2019.
- [3] S. Shi, Y. Xie, M. Li, A. X. Liu, and J. Zhao, “Synthesizing wider wifi bandwidth for respiration rate monitoring in dynamic environments,” in *Proc. of IEEE INFOCOM*, 2019.
- [4] C. Wu, X. Huang, J. Huang, and G. Xing, “Enabling ubiquitous wifi sensing with beamforming reports,” in *Proc. of ACM SIGCOMM*, 2023.
- [5] S. Tan, L. Zhang, Z. Wang, and J. Yang, “Multitrack: Multi-user tracking and activity recognition using commodity wifi,” in *Proc. of ACM CHI*, 2019.
- [6] E. Yi, F. Zhang, J. Xiong, K. Niu, Z. Yao, and D. Zhang, “Enabling wifi sensing on new-generation wifi cards,” in *Proc. of ACM IMWUT*, 2024.
- [7] Y. He, J. Liu, M. Li, G. Yu, and J. Han, “Forward-compatible integrated sensing and communication for wifi,” *IEEE JSAC*, 2024.
- [8] Y. Xie, J. Xiong, M. Li, and K. Jamieson, “md-track: Leveraging multi-dimensionality in passive indoor wi-fi tracking,” in *Proc. of ACM MobiCom*, 2019.
- [9] E. Soltanaghaei, A. Kalyanaraman, and K. Whitehouse, “Multipath triangulation: Decimeter-level wifi localization and orientation with a single unaided receiver,” in *Proc. of ACM MobiSys*, 2018.
- [10] E. Soltanaghaei, A. Dongare, A. Prabhakara, S. Kumar, A. Rowe, and K. Whitehouse, “Tagfi: Locating ultra-low power wifi tags using unmodified wifi infrastructure,” in *Proc. of ACM IMWUT*, 2021.
- [11] R. H. Venkatnarayanan, G. Page, and M. Shahzad, “Multi-user gesture recognition using wifi,” in *Proc. of ACM MobiSys*, 2018.
- [12] Y. Zheng, Y. Zhang, K. Qian, G. Zhang, Y. Liu, C. Wu, and Z. Yang, “Zero-effort cross-domain gesture recognition with wi-fi,” in *Proc. of ACM MobiSys*, 2019.
- [13] Y. Zeng, D. Wu, J. Xiong, J. Liu, Z. Liu, and D. Zhang, “Multisense: Enabling multi-person respiration sensing with commodity wifi,” in *Proc. of ACM IMWUT*, 2020.
- [14] C. R. Karanam, B. Korany, and Y. Mostofi, “Tracking from one side – multi-person passive tracking with wifi magnitude measurements,” in *Proc. of ACM/IEEE IPSN*, 2019.
- [15] R. H. Venkatnarayanan, M. Shahzad, S. Yun, C. Vlachou, and K.-H. Kim, “Leveraging polarization of wifi signals to simultaneously track multiple people,” in *Proc. of ACM IMWUT*, 2020.
- [16] J. Hu, T. Zheng, Z. Chen, H. Wang, and J. Luo, “Muse-fi: Contactless multi-person sensing exploiting near-field wi-fi channel variation,” in *Proc. of ACM MobiCom*, 2023.
- [17] J. Hu, H. Jiang, T. Zheng, J. Hu, H. Wang, H. Cao, Z. Chen, and J. Luo, “M²-fi: Multi-person respiration monitoring via handheld wifi devices,” in *Proc. of IEEE INFOCOM*, 2024.
- [18] D. Xia, L. Guan, H. Liu, Y. Mu, X. Wang, J. Han, Y. Shi, and L. Li, “Metabreath: Multitarget respiration detection based on space-time-coding digital metasurface,” *IEEE TMTT*, 2023.
- [19] Z. Li, T. Jin, D. Guan, and H. Xu, “Metaphys: Contactless physiological sensing of multiple subjects using ris-based 4-d radar,” *IEEE IoTJ*, 2023.
- [20] X. Li, J. W. You, Z. Gu, Q. Ma, J. Zhang, L. Chen, and T. J. Cui, “Multiperson detection and vital-sign sensing empowered by space-time-coding reconfigurable intelligent surfaces,” *IEEE IoTJ*, 2024.

- [21] L. Zhang, X. Q. Chen, S. Liu, Q. Zhang, J. Zhao, J. Y. Dai, G. D. Bai, X. Wan, Q. Cheng, G. Castaldi *et al.*, “Space-time-coding digital metasurfaces,” *Nature Communications*, vol. 9, no. 1, 2018.
- [22] L. Chen, W. Hu, K. Jamieson, X. Chen, D. Fang, and J. Gummesson, “Pushing the physical limits of iot devices with programmable metasurfaces,” in *Proc. of USENIX NSDI*, 2021.
- [23] X. Li, C. Feng, F. Song, C. Jiang, Y. Zhang, K. Li, X. Zhang, and X. Chen, “Protego: Securing wireless communication via programmable metasurface,” in *Proc. of ACM MobiCom*, 2022.
- [24] X. Li, C. Feng, X. Wang, Y. Zhang, Y. Xie, and X. Chen, “Rf-bouncer: A programmable dual-band metasurface for sub-6 wireless networks,” in *Proc. of USENIX NSDI*, 2023.
- [25] K. W. Cho, M. H. Mazaheri, J. Gummesson, O. Abari, and K. Jamieson, “mmwall: A steerable, transfective metamaterial surface for nextg mmwave networks,” in *Proc. of USENIX NSDI*, 2023.
- [26] T. J. Cui, M. Q. Qi, X. Wan, J. Zhao, and Q. Cheng, “Coding metamaterials, digital metamaterials and programmable metamaterials,” *Light: Science & Applications*, vol. 3, no. 10, 2014.
- [27] R. C. Hansen, *Phased array antennas*. John Wiley & Sons, 2009.
- [28] D. K. Cheng *et al.*, *Field and wave electromagnetics*. Pearson Education India, 1989.
- [29] T. J. Cui, S. Liu, and L. Zhang, “Information metamaterials and metasurfaces,” *Journal of materials chemistry C*, vol. 5, no. 15, 2017.
- [30] Y. Gu, J. Chen, K. He, C. Wu, Z. Zhao, and R. Du, “Wifileaks: Exposing stationary human presence through a wall with commodity mobile devices,” *IEEE TMC*, 2023.
- [31] B. Korany and Y. Mostofi, “Counting a stationary crowd using off-the-shelf wifi,” in *Proc. of ACM MobiSys*, 2021.
- [32] N. Instruments, “Ni pxie-1082,” <https://www.ni.com/en-us/support/model.pxie-1082.html>, 2024, accessed: 2024-10.
- [33] Z. Jiang, T. H. Luan, X. Ren, D. Lv, H. Hao, J. Wang, K. Zhao, W. Xi, Y. Xu, and R. Li, “Eliminating the barriers: Demystifying wi-fi baseband design and introducing the picoscenes wi-fi sensing platform,” *IEEE IoTJ*, vol. 9, no. 6, 2022.
- [34] H. Huang and S. Lin, “Widet: Wi-fi based device-free passive person detection with deep convolutional neural networks,” in *Proc. of ACM MSWIM*, 2018.
- [35] J. Zhang, F. Wu, W. Hu, Q. Zhang, W. Xu, and J. Cheng, “Wienhance: Towards data augmentation in human activity recognition using wifi signal,” in *Proc. of IEEE MSN*, 2019.
- [36] Vernier, “Go direct respiration belt,” <https://www.vernier.com/product/go-direct-respiration-belt/>, 2024, accessed: 2024-10.
- [37] C. Li, Q. Huang, Y. Zhou, Y. Huang, Q. Hu, H. Chen, and Q. Zhang, “Riscan: Ris-aided multi-user indoor localization using cots wi-fi,” in *Proc. of ACM SenSys*, 2023.
- [38] J. C. Liang, Q. Cheng, Y. Gao, C. Xiao, S. Gao, L. Zhang, S. Jin, and T. J. Cui, “An angle-insensitive 3-bit reconfigurable intelligent surface,” *IEEE TAP*, vol. 70, no. 10, 2021.
- [39] C. Feng, X. Li, Y. Zhang, X. Wang, L. Chang, F. Wang, X. Zhang, and X. Chen, “Rflens: Metasurface-enabled beamforming for iot communication and sensing,” in *Proc. of ACM MobiCom*, 2021.
- [40] L. Li, H. Ruan, C. Liu, Y. Li, Y. Shuang, A. Alù, C.-W. Qiu, and T. J. Cui, “Machine-learning reprogrammable metasurface imager,” *Nature Communications*, vol. 10, no. 1, 2019.
- [41] G. Lan, M. F. Imani, Z. Liu, J. Manjarrés, W. Hu, A. S. Lan, D. R. Smith, and M. Gorlatova, “Metasense: Boosting rf sensing accuracy using dynamic metasurface antenna,” *IEEE IoTJ*, vol. 8, no. 18, 2021.
- [42] T. Woodford, K. Qian, and X. Zhang, “Metasight: High-resolution nlos radar sensing through efficient metasurface encoding,” in *Proc. of ACM SenSys*, 2023.
- [43] Y. Zhang, Y. Wang, L. Yang, M. Wang, Y.-C. Chen, L. Qiu, Y. Liu, G. Xue, and J. Yu, “Acoustic sensing and communication using metasurface,” in *Proc. of USENIX NSDI*, 2023.

Yimiao Sun is currently a PhD student in Tsinghua University. He received his B.E. degree in the University of Electronic Science and Technology of China (UESTC). His research interests include mobile computing and wireless sensing.



Qunyan Zhou received the B.E. degree in electronic science and technology from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2021. He is currently pursuing a Ph.D. degree with the State Key Laboratory of Millimeter Waves, Southeast University, Nanjing. His current research focuses on space-time-coding digital metasurface.



Yulong Chen received his B.S. and M.E. degrees in the School of Software from Tsinghua University in 2022 and 2025, respectively. He is currently a software engineer at China Merchants Bank.



Jiaming Gu is currently a PhD student in Tsinghua University. He received his B.E. and B.S. degrees from the Weiyang College at the same university. His research interests include mobile computing and wireless sensing.



Qiang Cheng (Senior Member, IEEE) received the B.S. and M.S. degrees from the Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu, China, in 2001 and 2004, respectively, and the Ph.D. degree from Southeast University, Nanjing, in 2008. In 2008, he joined the State Key Laboratory of Millimeter Waves, Southeast University. He is the chief professor of Southeast University, deputy director of the State Key Laboratory of Millimeter Wave, winner of the National Natural Science Foundation Outstanding Youth Fund, Young Scholar of Changjiang Scholars Award Program, and the first batch of Outstanding Youth Fund in Jiangsu Province. His research work was selected as the second prize of the National Natural Science Award in 2014 and 2018, the First Prize of the Natural Science Award of the Ministry of Education in 2011, and the Top Ten Scientific and Technological Progress of Chinese Universities in 2021.



Yuan He (Senior Member, IEEE) is a professor in the School of Software and BNRIst of Tsinghua University. He received his B.E. degree from the University of Science and Technology of China, his M.E. degree in the Institute of Software, Chinese Academy of Sciences, and his PhD degree in Hong Kong University of Science and Technology. His research interests include wireless networks, Internet of Things, pervasive and mobile computing. He is a senior member of IEEE and a member of ACM.