

# Sensing in Bi-Static ISAC Systems with Clock Asynchronism: A Signal Processing Perspective

Kai Wu, *Member, IEEE*, Jacopo Pegoraro *Member, IEEE*, Francesca Meneghelli *Member, IEEE*,

J. Andrew Zhang, *Senior Member, IEEE*, Jesus O. Lacruz,

Joerg Widmer *Fellow, IEEE*, Francesco Restuccia *Senior Member, IEEE*,

Michele Rossi *Senior Member, IEEE*, Xiaojing Huang, *Senior Member, IEEE*

Daqing Zhang, *Fellow, IEEE*, Giuseppe Caire, *Fellow, IEEE*, and Y. Jay Guo, *Fellow, IEEE*

**Abstract**—Integrated Sensing And Communication (ISAC) has been identified as a pillar usage scenario for the impending 6G era. Bi-static sensing, a major type of sensing in ISAC, is promising to expedite ISAC in the near future, as it requires minimal changes to the existing network infrastructure. However, a critical challenge for bi-static sensing is clock asynchronism due to the use of different clocks at far-separated transmitters and receivers. This causes the received signal to be affected by time-varying random phase offsets, severely degrading, or even failing, direct sensing. Hence, to effectively enable ISAC, considerable research has been directed toward addressing the clock asynchronism issue in bi-static sensing. This paper provides an overview of the issue and existing techniques developed in an ISAC background. Based on the review and comparison, we also draw insights into the future research directions and open problems, aiming to nurture the maturation of bi-static sensing in ISAC.

**Index Terms**—Integrated sensing and communications, joint communications and sensing, bi-static sensing, clock asynchronism, cross-antenna cross-correlation, cross-antenna signal ratio

## I. INTRODUCTION

Communications-centric Integrated Sensing And Communication (ISAC) systems are expected to deliver a ubiquitous radio sensing network without notably affecting communication performance. These systems have sparked the rise of a novel technological paradigm that seamlessly integrates wireless sensing into next-generation mobile networks. This integration holds the potential to significantly enhance a diverse range of contemporary smart applications [1–3]. A prominent example is the implementation of ISAC in mobile networks, termed the Perceptive Mobile Network (PMN). Since its inception in 2017 [4], PMN has captured substantial attention from academia and industry alike. An important milestone occurred in June 2023 when the International Telecommunication Union (ITU) greenlighted the 6G vision framework, where the concept of ISAC has been identified as one of the six pivotal usage scenarios for the impending 6G era, also denoted by IMT-2030 [5].

This work is partially supported by the Australian Research Council under grant No. DP210101411.

This work was partially supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on “Telecommunications of the Future” (PE0000001 - program “RESTART”).

This work is supported in part by the National Science Foundation under grants CCF-2218845, ECCS-2229472, and ECCS-2329013, by the Office of Naval Research under N00014-23-1-2221 and by the Air Force Office of Scientific Research under grant FA9550-23-1-0261.

Many sensing approaches are possible in PMN [3], including mono-static sensing at a single base station, bi-static sensing between base stations, or between a mobile device – User Equipment (UE) – and a base station. Bi-static sensing configurations are particularly promising for seamlessly integrating sensing into communications. This approach boasts compelling advantages: It requires almost no change to the current network infrastructure, alleviating the formidable demands of challenging technologies such as full-duplex mobile transceivers. In particular, when the bi-static configuration includes a UE as the transmitter, sensing functions benefit from the spatial, spectral, and temporal diversities in user signals, thereby embracing the potency of networked and fused sensing methodologies. Moreover, performing sensing functionalities at the base station is more feasible, as these are equipped with heightened sensing capacities encompassing more advanced power, computational resources, and networking capabilities.

One critical challenge for bi-static sensing is clock asynchronism due to the bi-static ISAC setup. The bi-static transmitter (Tx) and sensing receiver (Rx), regardless of whether they are base stations or UEs, are spatially separated, and their clocks from the local oscillators are not locked and asynchronous. Clock asynchronism causes Timing Offset (TO), Carrier Frequency Offset (CFO), and random phase shifts across discontinuous transmissions, which hinder coherent processing of discontinuous channel measurements and introduce ambiguities into sensing results. This is a common problem in almost all communications-centric ISAC systems. Still, it is more prominent in mobile networks due to the typically wide separation of Txs and Rxs, the mobility of nodes, and the absence of line-of-sight connections. Should this problem be efficiently solved, sensing could be seamlessly realized in communication networks via software running in base stations with enhanced computing capabilities.

Considerable research has been directed toward bi-static sensing in ISAC systems to tackle this challenge. These efforts have led to some preliminary solutions, which were partially and conceptually reviewed in an early article [6]. However, these solutions have various limitations, and practically viable solutions are yet to be developed. In this article, we offer an all-encompassing and meticulous exploration of signal-processing techniques for compensation and removal of TO, CFO, and phase offsets in bi-static sensing. Specifically, this article encompasses ISAC signal processing architectures within contemporary and futuristic mobile networks, which may be applied to any bi-static scenario *without requiring co-*

operation or active signaling among nodes (also called *single-node* scenario). The manuscript is also devoted to thoroughly covering pertinent contributions from the existing open literature, coupled with the sharing of the authors' own reservoir of research experiences, outcomes, and forward-looking visions. Note that while most wireless sensing literature addresses the clock asynchronism challenge in Wi-Fi networks, the approaches can potentially be applied to mobile networks, given their similar physical-layer techniques. More detailed comparisons between the two networks will be provided in Section III. Overall, we hope that our work will contribute to the maturation of the ISAC technology in the near future.

The rest of this article is organized as follows. In Section II, we introduce the different phase offsets caused by clock asynchronism and establish the basic reference Channel Frequency Response (CFR) model used in this paper. A detailed review of three main approaches for eliminating such phase offsets is carried out in Section III, analyzing and comparing numerous works from the related literature. In Section IV, we generalize a technical framework and outline six key challenges for future research, focusing on the limitations of the reviewed methods. Finally, we give concluding remarks in Section V.

## II. CLOCK ASYNCHRONISM AND ITS MODELING IN BI-STATIC SENSING

### A. Offsets Caused by Clock Asynchronism

The basic clock is typically generated from the Local Oscillator (LO) to support various operations in a radio device. Clock asynchronism happens when the basic clocks of the Tx and Rx are not locked in phase or frequency. This introduces unwanted offsets in the received signals – including but not limited to CFO, TO, and phase offsets – as detailed next.

- *Carrier frequency offset.* CFO is mainly caused by the difference between the LOs of the transmitter and receiver. Since no two oscillators can generate the same frequency, the carrier frequencies generated by two independent oscillators will also differ, yielding CFO. Moreover, due to environmental factors such as temperature changes, supply voltage variations, and aging, oscillator frequency can drift slowly, leading to a slow-varying CFO between a pair of transmitters and receivers. It is typically estimated using, e.g., two segments of repeated signals in communication systems. This estimation can then be used to compensate for the CFO. However, the inevitable estimation error leads to residual CFO, which may become fast time-varying due to the variation of CFO estimation.

- *Timing offset.* TO results from the lack of synchronized time reference between the transmitter and receiver, as they use their own LOs to generate the necessary timing signals for transmitting and receiving signals. Due to independent clock sources, there can be an unknown shift or offset in the time perceived by the receiver compared to the actual transmission time. Such TO is almost unchanged during continuous transmission and reception; however, it may become time-varying during discontinuous transmission, particularly when a node

transits between transmission and reception. In addition, the change of synchronization point used in the receiver also introduces time-varying TO. The fine timing point may slightly vary in position due to, e.g., random noise, even if the channel is unchanged, particularly in Orthogonal Frequency Division Multiplexing (OFDM) systems with many subcarriers.

- *Phase offset.* The phase offset can be caused by transceiving electronic devices and the inherent phase noises of LOs. The factors contributing to these noises, such as thermal noise and flicker noise, are inherently rapid and stochastic, making phase offset fast-varying, possibly on a symbol basis.

Considering the different contributing factors to the three offsets, we can conclude that: CFO changes on a time scale of minutes to hours due to gradual factors like temperature changes and component aging, but residual CFO after compensation can change on a millisecond time scale; TO changes on a time scale of milliseconds to seconds, as it can be influenced both by slow-varying oscillator drift, transition between transmission and reception, and shift of fine-timing point; and phase offset changes on a time scale of microseconds as it results from rapid, intrinsic noise fluctuations within LO and electronic circuits.

Fig. I illustrates the phases of the Channel State Information (CSI) estimates over OFDM subcarriers and consecutive symbols for Long-Term Evolution (LTE) and Wi-Fi signals in static environments. The phases contain the effects of all offsets mentioned above. For LTE, which has a stringent timeslot structure, multiple frames (each containing 20 slots) are transmitted continuously. Each frame starts with a downlink slot, followed by uplink slots. The cell-specific reference signals are used for channel estimation, leading to 40 CSIs estimation vectors per frame (two symbols with reference signals in each slot). Each vector consists of CSI estimates over subcarriers. Four subcarriers are evenly selected over all resource elements for the illustration in Fig. I(a). The figure shows that the LTE CSI phase slowly varies over time during continuous transmission, but it suffers from rapid changes corresponding to the uplink-downlink transition. During continuous transmission, phase offset variations are slow as can be seen from the phases of each subcarrier; TO also varies slowly as can be seen from the slow variations of the relative phase differences between the four subcarriers. For Wi-Fi, random packet transmission is performed in the CSI extraction to emulate practical Wi-Fi scenarios. For clarity, two subcarriers (in different colors) are selected for illustration. We see that the CSI phase rapidly changes over timeslots and subcarriers, both in random manners, suggesting fast time-varying phase offset and TO. The results in Fig. I intuitively manifest the impact of the three offsets illustrated above on bi-static sensing signals.

### B. Signal Modelling

We now incorporate these offsets into a Multiple-Input and Multiple-Output (MIMO) Orthogonal Frequency Division Multiplexing (OFDM) ISAC signal model. We note

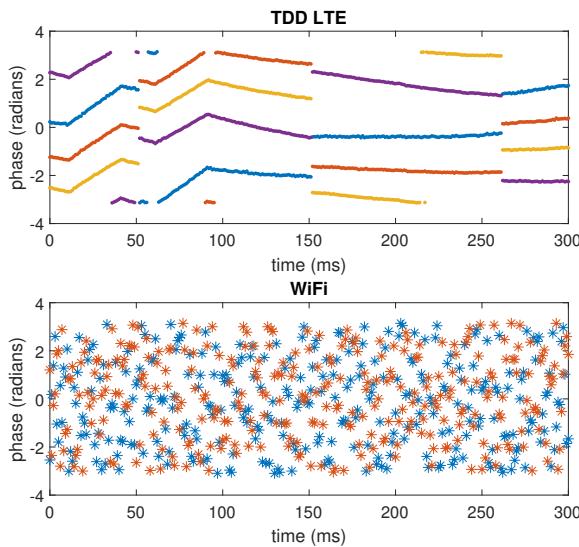


Figure 1: Time-varying phases of LTE's (top) and Wi-Fi's (bottom) CSIs extracted from the National Instrument MIMO testbed and a commercial Wi-Fi platform, respectively. Four different colors are for four equally spaced subcarriers.

that MIMO-OFDM has been extensively used in modern cellular networks, such as 4G and 5G. It is foreseeable to continue prevailing in 6G. Consider a two-node MIMO-OFDM communication system where the transmitting and receiving nodes are each equipped with a uniform linear array. The transmitter (receiver) array has  $M$  ( $N$ ) antennas with half-a-wavelength antenna spacing. Their steering vector can be given by  $\mathbf{a}(x, \alpha) = [1, e^{j\pi \sin(\alpha)}, \dots, e^{j(x-1)\pi \sin(\alpha)}]^T$ ,  $x = M$  or  $N$ , where  $\alpha$  can be the angle of departure (AoD) or arrival (AoA). For OFDM transmissions, the frequency band with bandwidth  $B$  is divided into  $S$  subcarriers. The subcarrier interval is then  $\Delta_f = B/S$  leading to an OFDM symbol period of  $T_O = T_0 + T_{cp}$ , where  $T_0 = 1/\Delta_f = S/B$  and  $T_{cp}$  is the period of the cyclic prefix.

Consider  $k = 1, \dots, K$  received OFDM symbols, equally spaced at an interval of  $T_s$ . Denote the TO and (residual) CFO as  $\tau_{o,k}$  and  $f_{o,k}$ , respectively. Moreover, we use  $\beta_k$ , a complex value, to reflect random phase offset and all other hardware imperfections (to sensing), such as Automatic Gain Control (AGC) and RF chain imbalance, etc. Note that all these nuisances depend on the relative asynchrony between the Tx and the Rx, hence they are specific to each Tx-Rx pair. Moreover, they are time-varying on an OFDM symbol level, as indicated by their subscript  $k$ . Assume there are  $L$  paths, where the  $l$ -th ( $l = 1, \dots, L$ ) path's complex path gain, propagation delay, Doppler frequency, AoA, and AoD are given by  $b_l$ ,  $\tau_l$ ,  $f_{D,l}$ ,  $\phi_l$  and  $\theta_l$ , respectively. The noise-free CFR matrix at the

$i$ -th subcarrier and  $k$ -th OFDM symbol can be expressed as

$$\mathbf{H}_{i,k} = \beta_k \sum_{l=1}^L b_l e^{-j2\pi i(\tau_l + \tau_{o,k})\Delta_f} e^{j2\pi(f_{D,l} + f_{o,k})kT_s} \mathbf{a}(N, \phi_l) \mathbf{a}^T(M, \theta_l) \quad (1)$$

$$= \underbrace{\beta_k e^{-j2\pi i\Delta_f \tau_{o,k}} e^{j2\pi f_{o,k}kT_s}}_{\xi_{i,k}} \sum_{l=1}^L b_l e^{-j2\pi i\Delta_f \tau_l} e^{j2\pi f_{D,l}kT_s} \mathbf{a}(N, \phi_l) \mathbf{a}^T(M, \theta_l). \quad (2)$$

Three assumptions have been made here: 1) All transmitting antennas share a common local oscillator, and all receiving antennas do so as well; 2) the intra-symbol phase change caused by Doppler frequency and CFO is negligible, and 3) the sensing parameters of the targets can be considered constant within a short coherent processing interval (CPI), such as over a period from several to tens of milliseconds [7]. Thus, their dependence on the symbol index  $k$  is suppressed in  $\mathbf{H}_{i,k}$  for brevity. All three assumptions are typical and widely used in most radar and ISAC work. In Eq. (2), several nuisance coefficients are combined into a single term, as denoted by  $\xi_{i,k}$ . The  $L$  paths in Eq. (2) can be grouped into *static* and *dynamic* paths, depending on whether the underlying scatterer is moving during the observation time period or not. Note that one observation period can consist of multiple OFDM slots or packets. For static paths, we have  $f_{D,l} = 0$ , and  $\tau_l, \theta_l, \phi_l$  remain constant over the whole observation period. For a dynamic path  $f_{D,l} \neq 0$ , the sensing parameters can be considered constant within a short processing interval, as done in Eq. (2). Over the observation period, the sensing parameters may change, as shown later in Section III-A and Fig. 3. These properties are key to some of the methods presented in Section III.

The offsets due to clock asynchronism generally degrade sensing. On the one hand, TO and CFO cause an ambiguity value in delay and Doppler frequency estimation, leading to ambiguity in ranging and speed estimation. On the other hand, the random phase offsets make it impossible to process multiple CSI measurements coherently directly; note that CSI power can be exploited for jointly processing multiple measurements in the temporal domain. Moreover, popular advanced sensing techniques from radar sensing, such as the micro-Doppler spectrogram [8], are also severely affected by CFO that destroys Doppler features. The micro-Doppler contains fine-grained information on targets, including multiple moving parts. It is a key component of many applications such as target recognition, human activity recognition, and person identification, among others [9]. We will use the quality of the micro-Doppler as one of the possible evaluation criteria and a means of comparison for CFO removal techniques.

### C. Challenges in Addressing Clock Asynchronism for Sensing

In communications systems, the aforementioned offsets due to clock asynchronism are compensated for using well-established algorithms [10, 11]. Moreover, there is no need to separate TO from the propagation delay for communications. The receiver determines the fine timing, and the remaining

timing offsets are included in the channel estimation and removed via channel equalization. For CFO, it can be estimated and compensated by using, e.g., two repeated sequences [11], and the impact of residual CFO, together with a small Doppler shift, may be ignored or mitigated via estimating and compensating their accumulated phase shift.

Unlike data communications, where the offsets can be treated effectively in general, mitigating the impacts of such offsets on sensing is challenging. For sensing, the TO has to be separated to get a clean estimate for propagation delays. The CFO and  $f_D$  also need to be separated for sensing, and this cannot be achieved via the conventional CFO estimation technique. The residual CFO may have a close value to the Doppler shifts to be estimated, causing large Doppler estimation errors if not handled properly. Therefore, addressing clock asynchronism is a critical challenge in bi-static sensing.

Several approaches can potentially address this problem [6], such as using a global GPS disciplined clock (GPSDO) or applying single-node-based signal processing techniques. A GPSDO with a timing accuracy up to 5.5 ns can cost 1000 in US dollars [6]. It also requires open-sky view to operate and its size is relatively large. Therefore, it cannot be practically integrated into every mobile device. In comparison, single-node-based processing can be more practical and promising in addressing the asynchronism challenge at a much lower cost, as it does not require additional devices or complex information exchange between multiple nodes. However, GPSDO is not without merits. In scenarios such as a large shopping mall, investing in a high-quality GPSDO may enable the distribution of a precise clock to many terminals via wires, improving clock stability for better bi-static sensing. This work mainly focuses on signal processing techniques for addressing the clock asynchronism issue. Next, we provide an in-depth overview of the techniques.

### III. SINGLE-NODE BI-STATIC SENSING TECHNIQUES

From the signal-processing perspective, addressing the clock asynchronism issue in bi-static sensing can be achieved by eliminating the contribution of  $\xi_{i,k}$  in Eq. (2). The elimination can be achieved via either cancellation without estimating the values or estimation followed by compensation; it can also be done in different domains, such as spatial and delay domains. Depending on how the “elimination” is achieved, we organize and present the techniques, as found in the literature, through three major types of methods, as summarized in Fig. 2.

- 1) Offset Cancellation Methods [12–18]: This approach is predominantly applied in the *spatial domain* [12–17], utilizing the fact that the clock asynchronism term,  $\xi_{i,k}$ , is common over antennas and can be mitigated through cross-antenna processing, including cross-correlation and signal ratio. Recently, an offset cancellation in the *delay domain* has been proposed [18], which does not require multiple receiving antennas;
- 2) Offset Estimation and Suppression [19–21]: These methods estimate the nuisance terms, namely  $\tau_{o,k}$  and  $f_{o,k}$  separately or their composite impact  $\xi_{i,k}$  as an integral, and then mitigate them to facilitate unambiguous sensing;

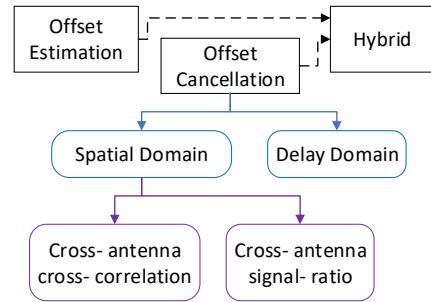


Figure 2: Classification of single node bi-static sensing techniques that tackle the clock asynchronism problem.

- 3) Hybrid of Offset Cancellation and Estimation [22]: This type of methods estimate and suppress the TO, while CFO is cancelled out similarly to point 1). Since the nuisance term,  $\xi_{i,k}$ , is common to all the propagation paths, delay-domain processing of the Channel Impulse Response (CIR) is used to obtain the TO efficiently through correlation. A reference static path is then used to cancel out the phase offsets without requiring the availability of multiple antennas at the receiver.

Each technique has specific features, advantages, and disadvantages, and their applicability depends on the sensing objectives. For example, some techniques may be suitable for obtaining estimates for some specific parameters, such as AoA and delay. Still, they may not be able to enable coherent signal processing over the temporal domains. Next, we review some representative works in the three types in more detail. Note that the different approaches will be analyzed based on the strategy used to tackle the clock asynchronism and estimate the multi-path propagation parameters, i.e., the paths’ complex path gain, propagation delay, Doppler frequency, AoA, and AoD, as described in the CFR model in Eq. (2). The target information regarding environmental position and movements can be obtained by applying further processing steps to the estimated propagation parameters. This processing can be independent of the estimation procedure. We refer interested readers to [22] for this aspect.

#### A. Offsets Cancellation Methods: Spatial-Domain Processing

- 1) *Cross-Antenna Cross-Correlation (CACC)* [12]: For clarity in reviewing the core idea of Cross-Antenna Cross-Correlation (CACC), we simplify  $\mathbf{H}_{i,k}$  in Eq. (1) by assuming a single transmitting antenna, i.e.,  $M = 1$ . Then the channel coefficient of the  $n$ -th antenna at the  $i$ -th sub-carrier and  $k$ -th OFDM symbol can be written into

$$H_{i,k,n} = \xi_{i,k} \sum_{l=1}^L b_l e^{-j2\pi i \Delta_f \tau_l} e^{j2\pi f_{D,l} k T_s} e^{jn\pi \sin(\phi_l)}. \quad (3)$$

By name, CACC takes the conjugate cross-correlation of the channel coefficients between two antennas. One antenna is considered as a *reference* and given index  $n = 0$  by convention. Let us consider the cross-correlation between  $H_{i,k,0}$ , i.e.,

$n = 0$ , with  $H_{i,k,n}$ , where  $n \neq 0$ . The CACC based on  $H_{i,k,n}$  can be calculated as

$$\begin{aligned} r_{i,k,n} &= H_{i,k,0} H_{i,k,n}^* \\ &\approx \sum_{l_1=1}^L \sum_{l_2=1}^L b_{l_1} b_{l_2}^* e^{-j2\pi i \Delta_f (\tau_{l_1} - \tau_{l_2})} \\ &\quad e^{j2\pi (f_{D,l_1} - f_{D,l_2}) k T_s} e^{-j\pi n \sin(\phi_{l_2})}, \end{aligned} \quad (4)$$

where the approximation is due to  $|\xi_{i,k}|^2 \approx 1$ . The phase offset is fully mitigated from  $r_{i,k,n}$ . However, its cross-products also introduce more unknown parameters by creating virtual paths. Numerous designs have been proposed in the literature to efficiently estimate the true target parameters from  $r_{i,k,n}$  [12–14].

In [12, 13], a strong Line-of-Sight (LoS) is assumed to be present, a single dynamic path is considered, and the  $L$  paths in Eq. (3) are divided into two groups: static paths and a single dynamic path. Then  $r_{i,k,n}$  in Eq. (4) has four main components:

- 1) Static paths on antenna  $n = 0$  times the conjugate of static paths on antenna  $n$ : in the presence of a strong LoS, this product is approximately constant over time and can be suppressed using a high-pass filter.
- 2) Dynamic path on the antenna  $n = 0$  times the conjugate of the dynamic path on the antenna  $n$ : This product is much weaker than the other three involving the strong LoS-dominant static path, and hence can be ignored.
- 3) Static paths on the antenna  $n = 0$  times the conjugate of the dynamic path on the antenna  $n$ : This product includes the dynamic path's sensing parameters, which have opposite signs to actual ones due to the conjugate operation. Its power is also dominated by the static path on antenna 0.
- 4) The dynamic path on antenna 0 times the conjugate of static paths on antenna  $n$ : This product reflects true sensing parameters with power dominated by the static paths on antenna  $n$ .

Based on the above four components' features, an add-minus sensing method (AMS) is proposed in [12]. As can be seen from the power features of the last two components above, adding a positive value to the antenna  $n$  can enhance the last component, and subtracting a positive value from antenna 0 can weaken the third component with the virtual target. This then facilitates the estimation of the actual target's sensing parameters based on CACC. Unlike AMS, the work in [13] proposes utilizing the other components listed above to mitigate unwanted terms. In particular, the authors of [13] reveal that, in the presence of a strong LoS path, the second-order cyclic differences of all cross-correlations over adjacent antennas can be used for estimating the terms related to the virtual targets in the first-order difference. The estimation can then be used to mitigate the impact of virtual targets in the first-order differences, facilitating the unambiguous Doppler frequency estimation using the root-MUSIC algorithm.

Different from the above work focusing on a single dynamic path, the work in [14] constructs new signals based on CACC

results in such a way that the constructed signals are spanned by the same number of bases as that of the dynamic paths. A mirrored MUSIC algorithm is then proposed to estimate the Doppler frequencies of multiple targets.

If multiple snapshots are exploited to generate AoA estimates only, we can compute the conventional spatial correlation matrix  $\frac{1}{K} \sum_{k=k_0}^{k_0+K-1} \mathbf{H}_{i,k} \mathbf{H}_{i,k}^H$ , where all offsets are canceled as CACC is implicitly applied. Thus, conventional spectrum analysis techniques such as MUSIC can be applied to obtain the estimates. However, it is important to note that the cross-correlation terms from  $K$  snapshots are averaged as the output, preventing the application of a bandpass filter to eliminate static components. Fortunately, as shown in [20], all static components can be estimated as one element in the MUSIC spectrum domain, as they appear as one constant vector in all snapshots.

2) *Cross-Antenna Signal Ratio (CASR)* [15]: Similar to CACC, Cross-Antenna Signal Ratio (CASR) utilizes the fact that the phase offset caused by asynchronism is common to all receiving antennas, the ratio between two antennas can cancel out the asynchronism factors, i.e.,  $\xi_{i,k}$  in Eq. (2), facilitating unambiguous sensing. Similar to how we obtain  $r_{i,k,n}$  in Eq. (4), the ratio between  $H_{i,k,n}$  and  $H_{i,k,0}$  provides the following CASR formulation

$$\rho_{i,k,n} = \frac{H_{i,k,n}}{H_{i,k,0}} = \frac{\sum_{l=1}^L b_l e^{-j2\pi i \Delta_f \tau_l} e^{j2\pi f_{D,l} k T_s} e^{j n \pi \sin(\phi_l)}}{\sum_{l=1}^L b_l e^{-j2\pi i \Delta_f \tau_l} e^{j2\pi f_{D,l} k T_s}}, \quad (5)$$

where  $\xi_{i,k}$  is canceled out, but the sensing parameters appear in the denominator. The latter change makes CASR a nonlinear model of sensing parameters, invalidating numerous conventional sensing methods based on a linear sensing model.

Assuming a single dynamic path and combining all static paths into a single term, the CASR formulation above can be simplified as

$$\rho_{i,k,n} = \frac{\tilde{b} e^{j n \pi \sin(\phi_l)} e^{j2\pi f_{D,l} k T_s} + \mathcal{A}}{\tilde{b} e^{j2\pi f_{D,l} k T_s} + \mathcal{B}}, \quad (6)$$

where  $\mathcal{A}$  and  $\mathcal{B}$  are independent of  $f_{D,l}$ . It is pointed in [15] that, if  $\mathcal{A} \tilde{b} e^{j2\pi f_{D,l} k T_s} - \mathcal{B} \tilde{b} e^{j n \pi \sin(\phi_l)} \neq 0$ ,  $\rho_{i,k,n}$  is the Möbius transformation of  $e^{j2\pi f_{D,l} k T_s}$ . Through analyzing the different impact of translation, complex inversion, and multiplication operations involved in a Möbius transformation, it is revealed that the CASR and  $e^{j2\pi f_{D,l} k T_s}$  span the same angular range in the complex domain. In the case of a single dynamic path, as considered in [15], the absolute CASR presents a sinusoidal-like oscillation and the peak-to-peak magnitude of the oscillation can be enhanced by rotating the CASR in the complex domain. After searching for the optimal rotation, the autocorrelation of the absolute CASR over symbols is used to estimate the dynamic path's Doppler frequency.

In [16], three estimation methods are proposed to estimate the Doppler frequency of a single target based on the CSI ratio. The first method estimates the angular change of  $\rho_{i,k,n}$  over a time interval, say  $\Delta_t$ , and then performs the linear fitting to estimate the changing rate of the phase as the Doppler frequency. Note that, to estimate the angular change over time, the center of the arc/circle generated by  $\rho_{i,k,n}$

must be estimated first. The second method is built on the same periodicity between  $\rho_{i,k,n}$  and  $e^{j2\pi f_{D,i}kT_s}$ . It proposes to estimate the periodicity of  $\rho_{i,k,n}$  over  $k$  by searching for the time interval leading to the closest angles of  $\rho_{i,k,n}$ . The third method utilizes the periodicity of  $\rho_{i,k,n}$  and searches the time interval leading to the minimum cross-correlation between  $\rho_{i,k,n}$  and its delayed version.

When tested on the public Widar2.0 Wi-Fi dataset [23], the three Doppler estimation methods proposed in [16] outperform the IndoTrack based on CACC [12] in indoor scenarios, as illustrated in Fig. 3. The results in the figure are obtained based on the open dataset Widar [23]. The data used here is collected in a classroom with a person walking in a square track and fixed communication transmitter and receiver. Thus, the Doppler frequency is varying over time. We note that each estimation in the figure is obtained based on a limited number of consecutive symbols, where the sensing parameters are approximately constant. Thus, the signal model in Eq. (1) applies to each estimation. The main issue with CACC is the mirrored dynamic paths, which, as seen from Fig. 3, can lead to false yet symmetric moving path detection. The CASR can mitigate the issue, particularly in single-dynamic-path scenarios.

Recently, the performance bounds for Doppler sensing using the CASR method are provided in [24], via deriving closed-form CRLB expressions. The impacts of numerous system parameters on Doppler sensing are disclosed. Waveform optimization based on the insights is also conducted, achieving substantial performance improvement.

To utilize the CSI ratio for multi-target sensing, a Taylor series-based method is proposed in [17]. A general model with multiple dynamic targets is considered in the work. The second-order complex Taylor series of the CSI ratio is derived with a closed-form expression in terms of the vector of sensing parameters to be estimated. This reveals the unique patterns

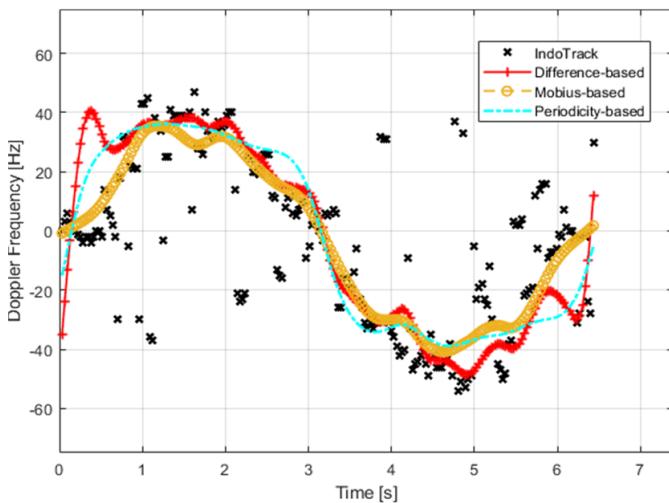


Figure 3: CASR-based Doppler estimation, in comparison with CACC-based estimation in a classroom setting [16], based on the Widar dataset [23]. As shown in [21], one can use the ground truth locations, along with timestamps, published in [23] for calculating the true Doppler frequencies.

embedded in the differences between CSI ratios over different snapshots, facilitating super-resolution multi-target Doppler sensing. In [19], instead, Independent Component Analysis (ICA) is used to separate the multiple targets contribution to the CFR, thus extending CASR to multitarget scenarios but entailing higher computational cost.

We remark that when more than one transmitting antennas are used, virtual antenna channels can be created at the sensing receiver side as in conventional MIMO radar processing. Referring to the channel matrix in Eq. (1), one can see that vectorizing the matrix can generate a new steering vector, which can be seen as the Kronecker product of the original steering vectors. Each entry of the vectorized channel vector can be seen as the CSI for a virtual channel. Nevertheless, we note that this spatial-domain processing does not change the fact that the nuisance caused by the random offsets, i.e.,  $\xi_{i,k}$  in Eq. (1), is still common to all entries in the CSI. This enables us to continue applying the offset elimination methods reviewed in this section to suppress the random offsets caused by clock asynchronism. However, sensing methods based on virtual channels after canceling offsets may need further investigation.

### B. Offsets Cancellation Methods: Delay-domain Processing Techniques

Delay-domain techniques remove TO and CFO by exploiting the fact that *they do not change in different signal propagation paths*, since they depend on the receiver hardware. Such techniques are, therefore, independent of the number of antennas at the transmitter or receiver devices, and they can work in Single Input Single Output (SISO) systems (where cross-antenna techniques are not applicable) or systems equipped with a single phased array. To see this, we simplify the CFR model for an OFDM system given in Eq. (2) using  $M = 1$  and  $N = 1$ . The channel gain for the  $k$ -th symbol at subcarrier  $i$  is written as

$$H_{i,k} = \xi_{i,k} \sum_{l=1}^L b_l e^{-j2\pi i \Delta_f \tau_l} e^{j2\pi f_{D,l} k T_s}. \quad (7)$$

The above model also holds if the transmitter and receiver are equipped with phased arrays of multiple antennas connected to the same RF chain. In this case, we consider the overall beamforming gain to be included in the coefficients  $b_l$ . Eq. (7) suggests that a viable method to compensate for the phase offsets is to use one of the propagation paths as a reference: The phase differences concerning the reference path only depend on the propagation path delay and Doppler shift since  $\xi_{i,k}$  cancels out. Hence, provided that the delay and Doppler of the reference path are known (e.g., because it is static so that the Doppler shift is absent and its delay can be estimated once using localization), estimation of the sensing parameters can be performed with low complexity algorithms. This idea is similar to the one behind spatial domain (i.e., cross-antenna) processing techniques discussed in Section III-A, but it does not require the use of multiple antennas, leveraging multipath resolution in the delay domain instead.

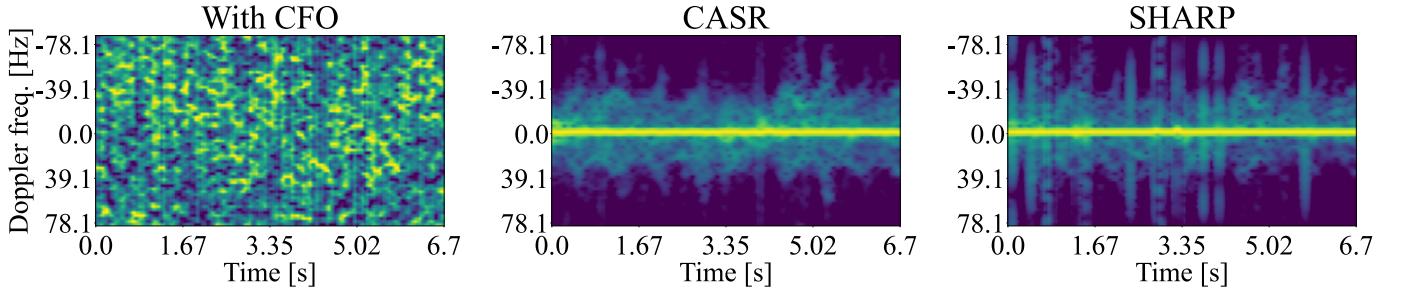


Figure 4: Example human micro-Doppler with 5 GHz carrier frequency (IEEE 802.11ac). In the leftmost plot, the spectrum is affected by CFO. In the other plots, the phase offset was removed using CASR [16] (middle) and SHARP [18] (right). Powers are normalized between 0 (blue) and 1 (yellow).

The CFR contribution of the reference path,  $\bar{H}_{i,k}$ , is expressed as  $\bar{H}_{i,k} = \xi_{i,k} b_1 e^{-j2\pi i \Delta_f \tau_1} e^{j2\pi f_{D,1} k T_s}$ , where we assume by convention that the reference path has index  $l=1$ . The offset-free CFR  $\hat{H}_{i,k}$  is obtained by multiplying the CFR in Eq. (7) by the complex-conjugate (to obtain phase differences) of the reference path's CFR, i.e.,

$$\hat{H}_{i,k} = H_{i,k} \bar{H}_{i,k}^* = b_1 \sum_{l=1}^L b_l e^{-j2\pi i \Delta_f (\tau_l - \tau_1)} e^{j2\pi (f_{D,l} - f_{D,1}) k T_s}. \quad (8)$$

In Eq. (8), the delay and Doppler parameters of path  $l$  have been replaced by the relative quantities  $\tau_l - \tau_1$  and  $f_{D,l} - f_{D,1}$ . However, unlike in spatial-domain techniques (see Section III-A) here, this does not increase the subsequent parameter estimation complexity for the following two reasons. First, both in LoS and Non-Line-of-Sight (NLoS) conditions, relative delays  $\tau_l - \tau_1$  can be easily used to track targets by exploiting the well-known bistatic radar geometry if the distance between Tx and Rx is known (see, e.g., [25]). Second, if the reference path used to compensate for CFO is *static*, we have  $f_{D,1} = 0$ , which means that the relative Doppler shifts in Eq. (8) become equal to the absolute ones. This allows for the direct estimation of the Doppler frequency, avoiding the ambiguity of relative frequency shift.

In the following, we detail SHARP, the delay-domain phase offsets compensation technique used in [18], for OFDM IEEE 802.11ac/ax systems operating on the 5 GHz Wi-Fi bands. Delay-domain processing is also used in [22], which includes a hybrid approach and will be presented in Section III-D.

In SHARP, the strongest propagation path is used as the reference by reasonably assuming that it corresponds to the LoS path and remains almost unchanged over a short acquisition time. The main challenge of applying delay-domain phase offset removal in sub-7 GHz systems is their low multipath resolution due to the small transmission bandwidth (typically 40 to 160 MHz). This makes it hard to separate the multiple propagation paths by simply applying inverse Fourier transformations over the time dimension. To address the challenge, [18] proposed a custom multipath resolution technique based on the compressive sensing-based optimization problem

$$\mathbf{r}_k = \arg \min_{\tilde{\mathbf{r}}} \|\mathbf{H}_k - \mathbf{T} \tilde{\mathbf{r}}\|_2^2 + \lambda \|\tilde{\mathbf{r}}\|_1, \quad (9)$$

where  $\mathbf{H}_k = [H_{0,k}, \dots, H_{S-1,k}]^T$  and  $\mathbf{T}$  is an  $(S \times L')$ -dimensional dictionary matrix with element  $(i, l)$  set to  $T_{i,l} = e^{-j2\pi i \tau_{l,tot}/T_O}$  and  $\tau_{l,tot}$  with  $l = 0, \dots, L' - 1$  is a set of *candidate* combined timing offsets, including the propagation delay,  $\tau_l$  and the TO,  $\tau_{o,k}$ .  $L' > L$  is the parameter that defines the grid of possible multi-path components included in the matrix  $\mathbf{T}$ . The vector  $\mathbf{r}_k$  (and, in turn,  $\tilde{\mathbf{r}}$ ) is  $L'$ -dimensional and represents the subcarrier-independent terms of the offset, including the CFO and the random phase offset, i.e., element  $l$  of  $\mathbf{r}_k$  is  $r_{k,l} = \beta_k e^{j2\pi f_{o,k} k T_s} b_l$ . Note that in our model in Eq. (2),  $b_l$  includes a carrier phase term,  $e^{-j2\pi f_c \tau_l}$  with  $f_c$  being the carrier frequency, which is instead made explicit in [18]. The position of the non-zero entries in vector  $\mathbf{r}_k$ , obtained by solving Eq. 9, indicate the estimated multi-path components out of the  $L'$  ones. The estimate of the multi-path decomposition of matrix  $\mathbf{H}_k$ , for each OFDM subcarrier  $i$ , is obtained by combining  $\mathbf{r}_k$  and the  $i$ -th row of  $\mathbf{T}$ , denoted by  $\mathbf{T}_i$ , through the Hadamard product, i.e.,  $\mathbf{X}_{i,k} = \mathbf{T}_i^T \circ \mathbf{r}_k$ . Finally,  $\hat{H}_{i,k}$  in Eq. (8) is obtained as  $\hat{H}_{i,k} = \sum_{l=1}^{L'} \mathbf{X}_{i,k} X_{i,k,0}^*$ , where  $X_{i,k,0}$  is the first element of vector  $\mathbf{X}_{i,k}$  and is associated with the strongest multipath component.

Example results on human micro-Doppler spectrogram extraction, comparing CASR and SHARP, are shown in Fig. 4. The results refer to an IEEE 802.11ac testbed with a 5 GHz carrier frequency.

### C. Offsets Estimation Methods

Different from the above methods using the spatial or time domain features to cancel the random phase offsets caused by the clock asynchronism, i.e.,  $\xi_{i,k}$  in (2), the designs in [19–21] seek to estimate and compensate the offset for high-performance Doppler sensing. The rationale can be seen from Eq. (4) – if only static paths are left in the denominator, the Doppler estimation would not suffer from the non-linearity caused by the CASR anymore.

A spatial-domain linear combination is adopted in [19, 20] to suppress the signals associated with the dynamic paths. The starting point of such designs is the transceiving relationship in the spatial domain, as embedded in the steering vectors in Eq. (2). More specifically,  $\mathbf{H}_{i,k}$  therein can be seen as a linear combination of  $\mathbf{a}(N, \phi_l) \mathbf{a}^T(M, \theta_l)$  which is the outer product of two steering vectors concerning transmitting and

receiving arrays. This outer product mainly conveys the spatial information of a path. Based on the matrix manipulation rule that  $\text{vec}(\mathbf{ABC}) = \mathbf{C}^T \otimes \text{Avec}(\mathbf{B})$ , where  $\otimes$  denotes the Kronecker product, we can vectorize  $\mathbf{H}_{i,k}$  and obtain

$$\begin{aligned} & \text{vec}(\mathbf{H}_{i,k}) \\ &= \underbrace{[\mathbf{a}(M, \theta_1) \otimes \mathbf{a}(N, \phi_1), \dots, \mathbf{a}(M, \theta_L) \otimes \mathbf{a}(N, \phi_L)]}_{\mathbf{A}} \mathbf{b}_{i,k}, \end{aligned} \quad (10)$$

where  $\mathbf{b}_{i,k}$  is a column vector collecting the coefficients of the  $L$  paths in Eq. (2). The above conversion is similar to conventional MIMO radar processing, where  $M$  Tx antennas and  $N$  Rx antennas can lead to  $MN$  virtual antennas. Using array signal processing theories, we see that it is theoretically feasible to identify the null space of dynamic paths. Projecting the CSI signals onto the null space can then facilitate the estimation of channel coefficients associated with the static path.

To achieve this, an optimization is performed in [19] to find the linear combination over virtual antennas that minimizes the ratio of respiration energy to the overall signal energy. The respiration energy is the signal energy in the frequency range caused by human respiratory movement, which can be determined based on system configurations, such as carrier frequency and the transceiver's geometric relationship. In [20], the MUSIC algorithm estimates the Angle of Arrivals (AoAs) of static and dynamic paths. It is shown in the work that, in the presence of a strong LoS path, the Bartlett beamformer would form a peak at the LoS path. This feature is then used to differentiate the LoS path's AoA from those of dynamic paths. The null space of the dynamic paths in the spatial domain is then constructed to extract the LoS path, which facilitates the estimation of the composite random phase caused by the asynchronism,  $\xi_{i,k}$ . Similarly, the null space method extracts each dynamic path by suppressing others, enabling the estimation of their Doppler frequencies.

If a strong LoS path exists and only the Doppler information is of interest, estimating and compensating for the asynchronism offsets can be quite effective by fully taking advantage of the LoS path. This is demonstrated in [21]. Treating the first path in Eq. (1) as the strong LoS path, then  $b_1 e^{-j2\pi i(\tau_1 + \tau_{o,k})\Delta_f}$  can be seen as a single-tone signal indexed by the sub-carrier index  $i$ . Estimating the frequency of this signal gives us an estimate of the composite delay with LoS path delay plus timing offset. We refer interested readers to [21, 26] for the details on the estimation method. It is based on DFT interpolation, a popular low-complexity frequency estimation method with off-grid accuracy. With the composite delay of the LoS path, i.e.,  $\tau_1 + \tau_{o,k}$ , estimated, signals can be coherently accumulated over the LoS path, and the phase of the high-Signal-to-Noise Ratio (SNR) accumulated signal can be used as the estimation of all other random phase offsets. This method applies to any antenna and symbol as seen from Eq. (1).

#### D. Hybrid Methods

In [22], a hybrid method, named JUMP, that combines TO estimation and suppression with CFO estimation was

proposed for a Millimeter-Wave (mmWave) system based on IEEE 802.11ay. mmWave systems offer a wide transmission bandwidth that, in turn, leads to higher multipath resolution than in sub-7 GHz ones. This makes mmWave systems the ideal application scenario for delay-domain techniques (see Section III-B) since multipath separation is easily achieved by detecting the peaks of the CIR. The latter is obtained as part of the communication protocol to perform channel equalization in Single Carrier (SC) systems (like IEEE 802.11ay), while it can be obtained via Inverse Discrete Fourier Transform (IDFT) in OFDM systems. Recalling that  $\xi_{i,k} = \beta_k e^{-j2\pi i \Delta_f \tau_{o,k}} e^{j2\pi f_{o,k} k T_s}$ , the CIR model corresponding to the CFR in Eq. (7) is

$$h_{n,k} = \beta_k \sum_{l=1}^L b_l e^{j2\pi(f_{D,l} + f_{o,k})k T_s} \delta_{n-\tau_l - \tau_{o,k}}, \quad (11)$$

where  $n$  is the index of delay-domain samples spaced by  $1/B$  and  $\delta_n$  is the Kronecker delta. Note that, in Eq. (11), using Kronecker delta functions is an approximation of the real CIR that neglects the impact of the non-ideal autocorrelation of pilot sequences and sampling points in SC systems, or the finite bandwidth in OFDM ones. We neglect such non-idealities as done in [22] to maintain the same notation.

From Eq. (11), one can see that the effect of TO is a *shift* of the CIR by  $\tau_{o,k}$ , while the CFO affects the phase of the CIR. [22] presents a hybrid algorithm that first estimates and compensates for the TO, then cancels out the CFO using a delay-domain approach. The two processing steps are detailed next.

- 1) *TO estimation and compensation:* The multiple propagation paths appear as peaks in the magnitude of the estimated CIR at the receiver, which can be detected, e.g., by defining a suitable threshold or by more advanced adaptive methods [22]. Then, the first detected path, having the lowest delay, is the LoS path to be used as a reference. The TO appears as a common shift of all propagation paths, so it can be removed by estimating the LoS path peak location, equal to  $\tau_1 + \tau_{o,k}$ , and applying an opposite shift to the CIR by convolving it with  $\delta_{n+\tau_1 + \tau_{o,k}}$ .
- 2) *CFO cancellation:* This second step is based on the delay-domain approach. The phase of a static reference path, i.e., having  $f_{D,1} = 0$ , is extracted. Such phase contains the nuisance due to  $f_{o,k}$ , and  $\angle \beta_k$ , so the CFO and random phase offset are removed from the CIR by computing  $\tilde{h}_{n,k} = \tilde{h}_{n,k} e^{-j\angle \tilde{h}_{1,k}}$ , where  $\tilde{h}_{1,k}$  is the TO-free CIR resulting from point 1).

The resulting CIR allows estimating relative delays and Doppler frequencies with low complexity algorithms. As an example, in [22], this is done by applying peak detection to the CIR magnitude to obtain  $\tau_l - \tau_1$  with  $l = 1, \dots, L$  from the locations of the peaks. Then,  $f_{D,l}$  are estimated by performing a Discrete Fourier Transform (DFT) over a window of  $K$  CIR samples. The Doppler frequencies correspond to the peaks in the Doppler spectrum, computed as the squared magnitude of the DFT.

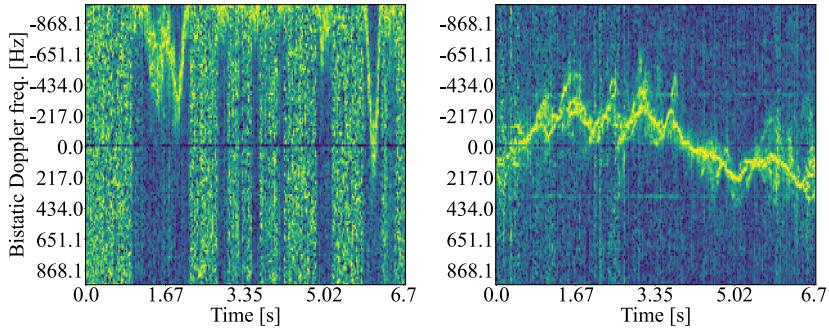


Figure 5: Example human micro-Doppler with 60 GHz carrier frequency (IEEE 802.11ay). On the left, the micro-Doppler spectrum is affected by CFO. On the right, we apply JUMP [22] to remove it. Powers are normalized between 0 (blue) and 1 (yellow).

Example results showing the effect of CFO compensation on a human micro-Doppler spectrogram are shown in Fig. 5. The experimental data was obtained using a Field Programmable Gate Array (FPGA)-based IEEE 802.11ay system with 60 GHz carrier frequency.

The above offsets compensation technique is then extended to the case where the LoS path is not always available, e.g., because it is temporarily blocked by an obstacle. The approach relies on the assumption that the channel contains multipath reflections on static objects, which are slowly time-varying compared to the packet transmission rate, which is a common assumption for radar and ISAC. The technique leverages the similarity of the CIR magnitude profiles across subsequent packets and performs the following steps

- 1) Estimation of the *relative* TO between subsequent CIR estimates, defined as  $\tau_o^r \triangleq \tau_{o,k} - \tau_{o,k-1}$ . This can be done, e.g., by locating the correlation peak for different candidate shift values between the CIR amplitude profiles  $|h_{n,k}|$  and  $|h_{n,k-1}|$ .
- 2) Compensation of the relative TO by shifting  $h_{n,k}$  by  $\tau_o^r$ . When applied sequentially to subsequent CIR estimates for  $k = 0, \dots, K - 1$ , this step incrementally removes the TO.

Thanks to 1), any static path can be used as a reference since the CIR samples across time share the same timing reference. Selecting a reference path different from the LOS may be challenging when CFO is present since the distinction between static paths and dynamic ones cannot be made based on the corrupted Doppler shift. In [22], this is solved by applying a tracking algorithm to the multipath components of the CIR to estimate the location of the scatterer in the 2D Cartesian plane after TO has been compensated for. Static paths are identified as the ones whose position does not change significantly across time, and the reference path is selected as the static path with the highest amplitude.

To initialize the algorithm, [22] assumes that the sensing operation starts in a LoS condition, where the TO can be removed, or that the LoS becomes available at least once during the system operation. Under this assumption, the compensation of the *relative* TOs between subsequent packets by steps 1)-2) allows maintaining the LoS-based timing reference even when transitioning to NLoS conditions for intermittent time intervals. Note that the above technique depends on the fact that transitions between LoS and NLoS conditions are slow

compared to variations of the CIR magnitude profile. This makes it appealing for indoor scenarios, in which it is likely to have multiple static propagation paths (e.g., from reflections on the walls). Conversely, due to its sensitivity to abrupt changes in the multipath profile, this technique may not be suitable for highly dynamic outdoor scenarios.

#### E. Comparisons and Remarks

Table I compares all methods reviewed above regarding their requirements and features, where OTA is short for “over the air”. The three types are stated at the beginning of Section III, where the first type is further divided into two sub-types, with the spatial-domain methods, under the type I(S) in the table reviewed in Section III-A, and the delay-domain method, under I(D), reviewed in Section III-B. In Table I, the three domains are mainly referred to as where the offsets are addressed. As mentioned earlier, most works are predominantly performed in the spatial domain. However, some recent works incorporate other domains to relieve the reliance on multiple antennas and explore the degrees of freedom available.

Moreover, not all methods can sense multiple targets concurrently moving in the environment. Specifically, CACC and delay-domain techniques support multiple targets, subject to having enough bandwidth and/or AoA resolution to resolve the corresponding multipath components. Conversely, CASR is mainly applicable to single-target scenarios and requires special processing to separate the multiple contributions, as shown in [17] [19].

As a sensing-related application, positioning or localization in 5G and beyond networks is emerging as a hot research topic [28] [29]. In such applications, timing offset is of particular interest, as the time of arrival is typically a key sensing parameter used in localization algorithms. In [28], a time difference of arrival is proposed to counteract timing offsets. In [29], a novel maximum likelihood estimator is developed to estimate timing offset and other sensing parameters, such as multi-path angles. While the timing offset can be estimated using these approaches, they do not sufficiently treat the frequency offset and random phase. In bi-static sensing, the methods reviewed in Sections III-B to III-D may be further applied on top of these approaches [28] [29]; alternatively, it can also be interesting to extend the methods [28] [29] with the other two types of offsets taken into account.

Table I: Comparison of different bi-static sensing Techniques (In the Processing Domain column, S, D, and F stand for “Spatial,” “Delay,” and “Frequency,” respectively).

Type	Methods	Strong LoS	Multiple Targets	Single Antenna	Processing Domain	Doppler-AoA Only	OTA
I(S)	Indotrack [12]	✓	✗	✗	S	✓	✓
	WiDFS [13]	✓	✗	✗	S	✓	✓
	CACC [14]	✓	✓	✗	S	✓	✗
	FarSense [15]	✗	✗	✗	S	✓	✓
	CASR [16]	✗	✗	✗	S	✓	Widar
	CASR [17]	✗	✓	✗	S	✓	WiDFS
I(D)	SHARP [18]	✗	✗	✓	D	✗	✓ [27]
	multiSense [19]	✗	✓	✗	S, F	✓	✓
II	[20]	✓	✓	✗	S	✓	✓
	[21]	✓	✓	✓	F	✓	Widar
III	JUMP [22]	✗	✓	✓	D	✗	✓

We remark that the ideas behind the methods reviewed in this section are promising to be applicable in the 6G context for addressing clock offsets, although quite a number of them are developed in the Wi-Fi sensing background. This is because, regardless of network differences, the MIMO-OFDM waveform is used by both cellular and Wi-Fi networks, and all the reviewed methods can apply to a general MIMO-OFDM setup. However, it is noteworthy that, when applying these methods to the 6G network, adaptions/changes may be necessary due to the potentially different features of 6G networks compared with Wi-Fi, as highlighted below.

- The 6G base station may be equipped with the so-called massive MIMO arrays, which can be efficiently applied to achieve great spatial resolution and separate multiple objects to be sensed. Millimeter wave systems will likely be hybrid arrays with analog beamforming before their digital counterparts. This means that the cross-antenna techniques [12–17] will need to be extended to deal with *cross-beam* signals. Moreover, analog beamforming may be designed to facilitate better solutions to addressing clock offsets.
- 6G is expected to have a much larger frequency bandwidth than Wi-Fi, leading to more degrees of freedom in the time-frequency-domain resources. These can further improve some of the methods reviewed in this section. For example, the large bandwidth can be enjoyed by SHARP [18] to facilitate much higher resolution in the range domain, hence addressing clock offset issues more effectively. As another example, the waveform design in the time-frequency domain can be resorted to, as performed in [24], for better offset estimation performance.
- 6G networks may be deployed in both indoor and outdoor. For indoor applications, the channels will be similar to WiFi channels, and techniques validated on a WiFi platform can be expected to work effectively on 6G networks. The 6G outdoor channels can be substantially different from WiFi ones. This makes channel-dependent methods, such as [19] in indoor WiFi scenarios, not directly applicable to 6G. However, the idea may be borrowed. For example, instead of using the respiratory movement for estimating the optimal combination over antennas (see Section III-C for more details), the varying channels caused by a car’s movement in a rural area can be relied on in 6G outdoor sensing to achieve similar

goals.

#### IV. FUTURE RESEARCH DIRECTIONS AND OPEN PROBLEMS

We have seen promising advances in addressing the clock asynchronism in bi-static ISAC. However, admittedly, many open problems still need more attention from the academia and industry communities.

The discussion on the techniques above enables us to formulate a comprehensive framework, laying the groundwork for future research at an elevated level. This framework, depicted in Fig. 6, is two-dimensional. One dimension encompasses signal processing methods, including reference signal construction, offset estimation, and offset cancellation via cross-correlation and signal ratio. The second dimension pertains to the domains where signal processing is primarily applied, with the typical space and delay domains. These methods have the flexibility to be implemented individually or in combination across one or more domains. While existing works, as explored in Section III, have substantiated feasibility and demonstrated potential, a thorough examination of the advantages and disadvantages of these methods and their application domains reveals opportunities for further refinement and exploration. A summary of the examination is as follows:

- Constructing a reference signal can simplify the subsequent sensing parameter estimation, but it will inevitably introduce errors;
- Offset estimation enables flexible signal processing. However, it also faces estimation errors, which could be larger than those in reference signal construction;
- Offset cancellation exploits signal structures and removes offsets without introducing errors. However, it typically leads to more parameters to be estimated (signal cross-correlation) or a nonlinear problem (signal ratio);
- Spatial-domain processing well retains independence of paths in delay and Doppler domain, but multiple receiving chains are required, which may not be cost-effective, particularly for millimeter wave systems;
- Delay-domain processing only requires a single receive antenna, but obtaining CIR is sometimes challenging in OFDM systems when measurements are only available at part of the subcarriers. In addition, each CIR may be the superposition of multiple paths when delays are off-grid.

Therefore, a proper combination can be designed by jointly considering the respective advantages and disadvantages of

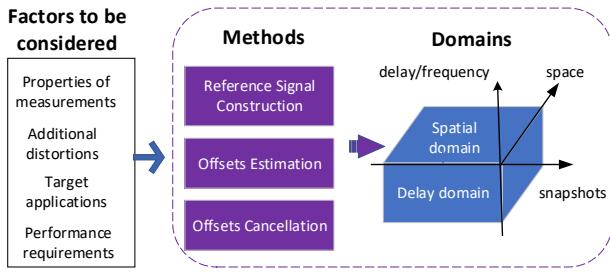


Figure 6: A generalized framework for solutions to clock asynchronism in bi-static sensing.

these methods and domains, together with additional signal and system factors, as shown in Fig. 6. By delineating the two-dimensional landscape of signal processing methods and their application domains, novel combinations and innovations may be achieved. Within the framework, some more specific research problems are elaborated next.

1) *Construction of a reference path*: The efficient and reliable construction of a reference path is an interesting open research challenge in delay-domain processing techniques to enable phase offset compensation. [22] proposed to apply tracking algorithms to maintain an estimate of several candidate reference paths, including first-order reflections on static objects. However, the sensing accuracy depends on the signal quality received from the reference path and degrades with low SNR.

2) *Sensing of multiple targets*: To the best of our knowledge, only two works [19] [22] have experimentally validated phase offsets removal techniques (hybrid, delay-domain and CASR, respectively) on data collected with multiple sensing targets concurrently moving in the environment. Cross-antenna techniques do not work out of the box in such cases, and techniques to separate the contribution of the different subjects are required, as detailed in Section III-A2. Reference path-based techniques are not affected by multiple targets. However, obtaining a reliable reference path when multiple targets are moving, possibly causing occlusion to the LoS, is extremely challenging and deserves further attention from the research community.

3) *Mobility of the nodes*: Performing ISAC with mobile nodes is a challenging problem in the presence of phase offsets. Notably, the assumption that the transmitter and receiver are static is critical to all three phase offset removal techniques. To our knowledge, only [30] has presented a preliminary dynamic CFR model for Wi-Fi sensing, but it does not deal with phase offsets. Possible approaches to tackle this problem may come from the passive radar literature, where radars mounted on moving platforms are well-studied, e.g., [31].

4) *Coherently combine multiple receivers*: To fully exploit the potential of pervasive ISAC, algorithms that combine the information from multiple, phase-synchronized receivers need to be developed. This would boost the system's sensing resolution, overcoming the limitation imposed by the bandwidth and enabling wavelength-level sensing accuracy. To enable this, algorithms that effectively remove the different phase

offsets between each transmitter-receiver must be investigated to allow the coherent fusion of the data obtained at each receiver. A preliminary study on this subject can be found in [32].

5) *Joint spatial-delay-frequency-domain processing*: From Table I we see that existing designs generally utilize one or two but not all three domains (spatial, delay, and frequency) for addressing the clock asynchronism. This contributes to the restrictions of existing methods, e.g., having a strong LoS, large bandwidth, or a single dynamic path, that are often associated with which domain we use for phase offsets compensation. For instance, many spatial and frequency domain methods require strong LoS, which is unnecessary for the delay-domain methods. However, the performance of the delay domain method may be subject to the bandwidth available, while the other domains are not. Jointly using the three domains can allow us to exploit and combine the advantages of different approaches, promising to remove the restrictions suffered by most existing works.

6) *Open dataset for bi-static sensing research and development*: While most works reviewed in Section III perform OTA experiments to validate their methods, custom configurations and hardware platforms are often heterogeneous and do not follow unified standards. For this purpose, only a few publicly available open datasets, such as those mentioned in Table II are available. Moreover, most existing works and experiments are based on Wi-Fi protocols/platforms. Other wireless systems, such as mobile networks, have been validated in a few works, such as [33]. Therefore, a critical aspect for advancing bi-static sensing further is the availability of a comprehensive open dataset based on different communications protocols in various yet typical sensing scenarios that can be used to assess and compare the performance of the proposed approaches.

## V. CONCLUDING REMARKS

In this survey, we have provided a first review of the clock asynchronism issue in the bi-static sensing against the highly popular ISAC background. We started by establishing the MIMO-OFDM signal model, which is widely used in modern communications systems. We then modeled and explained the impact of random offsets caused by the clock asynchronism issue on bi-static sensing. Moreover, we reviewed three major solutions to the issue, as differentiated by whether the offsets are canceled, estimated and suppressed, or treated in a hybrid way. The basic signal processing principles of the reviewed techniques are highlighted, and some interesting results are demonstrated. After comparing all reviewed methods, we drew insights into potential research directions and open problems. We hope our first survey paper on this topic can foster more research activities on maturing bi-static sensing in ISAC.

## REFERENCES

- [1] A. Graff, Y. Chen, N. González-Prelcic, and T. Shimizu, "Deep learning-based link configuration for radar-aided multiuser mmwave vehicle-to-infrastructure communication," *IEEE Trans. Veh. Techn.*, vol. 72, no. 6, pp. 7454–7468, 2023.
- [2] K. V. Mishra, M. Bhavani Shankar, V. Koivunen, B. Ottersten, and S. A. Vorobyov, "Toward millimeter-wave joint radar communications: A signal processing perspective," *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 100–114, 2019.

[3] J. A. Zhang, M. L. Rahman, K. Wu, X. Huang, Y. J. Guo, S. Chen, and J. Yuan, "Enabling joint communication and radar sensing in mobile networks—a survey," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 306–345, 2022.

[4] J. A. Zhang, A. Cantoni, X. Huang, Y. J. Guo, and R. W. Heath, "Framework for an innovative perceptive mobile network using joint communication and sensing," in *2017 IEEE 85th VTC Spring*, 2017, pp. 1–5.

[5] A. Kumar, "ITU approves 6G vision framework," <https://telecom.economictimes.indiatimes.com/news/policy/itu-approves-6g-vision-framework/101184232>, accessed: 2023-08-27.

[6] J. A. Zhang, K. Wu, X. Huang, Y. J. Guo, D. Zhang, and R. W. Heath, "Integration of radar sensing into communications with asynchronous transceivers," *IEEE Communications Magazine*, vol. 60, no. 11, pp. 106–112, 2022.

[7] J. A. Zhang, F. Liu, C. Masouros, R. W. Heath, Z. Feng, L. Zheng, and A. Petropulu, "An overview of signal processing techniques for joint communication and radar sensing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 6, pp. 1295–1315, 2021.

[8] V. C. Chen, F. Li, S.-S. Ho, and H. Wechsler, "Micro-Doppler effect in radar: phenomenon, model, and simulation study," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 1, pp. 2–21, 2006.

[9] K. V. Mishra, M. B. Shankar, V. Koivunen, B. Ottersten, and S. A. Vorobyov, "Toward millimeter-wave joint radar communications: A signal processing perspective," *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 100–114, 2019.

[10] W.-C. Liu, T.-C. Wei, Y.-S. Huang, C.-D. Chan, and S.-J. Jou, "All-digital synchronization for SC/OFDM mode of IEEE 802.15. 3c and IEEE 802.11ad," *IEEE Transactions on Circuits and Systems*, vol. 62, no. 2, pp. 545–553, Oct 2014.

[11] T. M. Schmidl and D. C. Cox, "Low-overhead, low-complexity burst synchronization for OFDM," in *Proceedings of ICC/SUPERCOMM'96-International Conference on Communications*, vol. 3. IEEE, 1996, pp. 1301–1306.

[12] X. Li, D. Zhang, Q. Lv, J. Xiong, S. Li, Y. Zhang, and H. Mei, "Indotrack: Device-free indoor human tracking with commodity Wi-Fi," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 3, sep 2017.

[13] Z. Wang, J. A. Zhang, M. Xu, and Y. J. Guo, "Single-target real-time passive Wifi tracking," *IEEE Transactions on Mobile Computing*, vol. 22, no. 6, pp. 3724–3742, 2023.

[14] Z. Ni, J. A. Zhang, X. Huang, K. Yang, and J. Yuan, "Uplink sensing in perceptive mobile networks with asynchronous transceivers," *IEEE Trans. on Signal Processing*, pp. 1–1, 2021.

[15] 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," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 3, no. 3, sep 2019.

[16] X. Li, J. A. Zhang, K. Wu, Y. Cui, and X. Jing, "CSI-ratio-based Doppler frequency estimation in integrated sensing and communications," *IEEE Sensors Journal*, vol. 22, no. 21, pp. 20 886–20 895, 2022.

[17] Z. Ni, J. A. Zhang, K. Wu, and R. P. Liu, "Uplink sensing using CSI ratio in perceptive mobile networks," *IEEE Transactions on Signal Processing*, vol. 71, pp. 2699–2712, 2023.

[18] F. Meneghelli, D. Garlisi, N. Dal Fabbro, I. Tinnirello, and M. Rossi, "SHARP: Environment and Person Independent Activity Recognition with Commodity IEEE 802.11 Access Points," *IEEE Transactions on Mobile Computing*, 2022.

[19] Y. Zeng, D. Wu, J. Xiong, J. Liu, Z. Liu, and D. Zhang, "Multisense: Enabling multi-person respiration sensing with commodity Wifi," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 4, no. 3, sep 2020.

[20] J. Zhao, Z. Lu, J. A. Zhang, S. Dong, and S. Zhou, "Multiple-target Doppler frequency estimation in ISAC with clock asynchronism," *IEEE Transactions on Vehicular Technology*, pp. 1–6, early access, 2023.

[21] K. Wu, J. A. Zhang, X. Huang, and Y. J. Guo, "A low-complexity CSI-based Wifi sensing scheme for LoS-dominant scenarios," in *ICC 2023 - IEEE International Conference on Communications*, 2023, pp. 2747–2752.

[22] J. Pegoraro, J. O. Lacruz, T. Azzino, M. Mezzavilla, M. Rossi, J. Widmer, and S. Rangan, "JUMP: Joint communication and sensing with Unsynchronized transceivers Made Practical," *IEEE Transactions on Wireless Communications*, 2024.

[23] K. Qian, C. Wu, Y. Zhang, G. Zhang, Z. Yang, and Y. Liu, "Widar2.0: Passive human tracking with a single Wi-Fi link," in *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 350–361.

[24] Y. Hu, K. Wu, J. A. Zhang, W. Deng, and Y. J. Guo, "Performance bounds and optimization for CSI-ratio based bi-static Doppler sensing in ISAC systems," *arXiv preprint arXiv:2401.09064*, 2024.

[25] P. Sameczyński, K. Abratkiewicz, M. Plotka, T. P. Zieliński, J. Wszołek, S. Hausman, P. Korbel, and A. Ksieżyk, "5G network-based passive radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–9, 2021.

[26] K. Wu, J. A. Zhang, X. Huang, and Y. J. Guo, "Accurate frequency estimation with fewer DFT interpolations based on Padé approximation," *IEEE Trans. Veh. Techn.*, vol. 70, no. 7, pp. 7267–7271, 2021.

[27] F. Meneghelli, N. D. Fabbro, D. Garlisi, I. Tinnirello, and M. Rossi, "A CSI Dataset for Wireless Human Sensing on 80 MHz Wi-Fi Channels," *IEEE Communications Magazine*, vol. 61, no. 9, pp. 146–152, June 2023.

[28] J. Palacios, N. González-Prelicic, and C. Rusu, "Low complexity joint position and channel estimation at millimeter wave based on multidimensional orthogonal matching pursuit," in *2022 30th European Signal Processing Conference (EUSIPCO)*. IEEE, 2022, pp. 1002–1006.

[29] M. A. Nazari, G. Seco-Granados, P. Johannesson, and H. Wymeersch, "Mmwave 6D radio localization with a snapshot observation from a single BS," *IEEE Transactions on Vehicular Technology*, 2023.

[30] J. Liu, W. Li, T. Gu, R. Gao, B. Chen, F. Zhang, D. Wu, and D. Zhang, "Towards a Dynamic Fresnel Zone Model to Wifi-based Human Activity Recognition," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 7, no. 2, pp. 1–24, 2023.

[31] P. Wojaczek, F. Colone, D. Cristallini, and P. Lombardo, "Reciprocal-filter-based STAP for passive radar on moving platforms," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 2, pp. 967–988, 2018.

[32] D. Tagliaferri, M. Manzoni, M. Mizzini, S. Tebaldini, A. V. Monti-Guarnieri, C. M. Prati, and U. Spagnolini, "Cooperative Coherent Multistatic Imaging and Phase Synchronization in Networked Sensing," *arXiv preprint arXiv:2311.07212*, 2023.

[33] Y. Feng, Y. Xie, D. Ganeshan, and J. Xiong, "LTE-based pervasive sensing across indoor and outdoor," in *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems*, ser. SenSys '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 138–151. [Online]. Available: <https://doi.org/10.1145/3485730.3485943>

**Kai Wu** (M'21) is a lecturer with the Global Big Data Technologies Centre (GBDTC), School of Electrical and Data Engineering (SEDE), University of Technology Sydney (UTS), Sydney, NSW 2007, Australia. (kai.wu@uts.edu.au)

**Jacopo Pegoraro** (M'23) is a postdoctoral researcher and lecturer in the Department of Information Engineering (DEI) at the University of Padova, Italy. (pegawairoja@dei.unipd.it)

**Francesca Meneghelli** (M'22) is an assistant professor with the Department of Information Engineering (DEI) at the University of Padova, Italy. (francesca.meneghelli.1@unipd.it)

**J. Andrew Zhang** (M'04-SM'11) is a Professor in the School of Electrical and Data Engineering, University of Technology Sydney, Australia. (andrew.zhang@uts.edu.au)

**Jesus O. Lacruz** is a Research Engineer at IMDEA Networks, Spain since 2017. (jesusomar.lacruz@imdea.org)

**Joerg Widmer** (F'20) is Research Professor and Research Director of IMDEA Networks in Madrid, Spain. (Joerg.Widmer@imdea.org)

**Francesco Restuccia** (SM'21) is an Assistant Professor of Electrical and Computer Engineering at Northeastern University, United States. (f.restuccia@northeastern.edu)

**Michele Rossi** (SM'13) is a full professor at the Department of Information Engineering (DEI) and the Department of Mathematics of the University of Padova. (michele.rossi@unipd.it)

**Xiaojing Huang** (M'99-SM'11) is currently a Professor of Information and Communications Technology with the School of Electrical and Data Engineering and the Program Leader for Mobile Sensing and Communications with the Global Big Data Technologies Center, University of Technology Sydney (UTS), Australia. (xiaojing.huang@uts.edu.au)

**Daqing Zhang** is a Professor with Peking University, China and IP Paris, France. He is a Fellow of IEEE and Member of Academy of Europe. (daqing.zhang@telecom-sudparis.eu)

**Giuseppe Caire** (S'92 – M'94 – SM'03 – F'05) is an Alexander von Humboldt Professor with the Faculty of Electrical Engineering and Computer Science at the Technical University of Berlin, Germany. (caire@tu-berlin.de)

**Y. Jay Guo** (F'14) is a Distinguished Professor and the funding Director of Global Big Data Technologies Centre (GBDTC) at the University of Technology Sydney (UTS), Australia. He is the founding Technical Director of the New South Wales (NSW) Connectivity Innovation Network (CIN). (jay.guo@uts.edu.au)