Anchor Points Assisted Uplink Sensing in Perceptive Mobile Networks

Yanmo Hu, *Student Member, IEEE*, J. Andrew Zhang, *Senior Member, IEEE*,
Kai Wu, *Member, IEEE*, Weibo Deng, Y. Jay Guo, *Fellow, IEEE*

Abstract*—Uplink sensing in integrated sensing and communications (ISAC) systems, such as Perceptive Mobile Networks, is challenging due to the clock asynchronism between transmitter and receiver. Existing solutions typically require the presence of a line-of-sight path and the knowledge of the transmitter's location. In this paper, these requirements, we propose a novel uplink sensing scheme to relieve these requirements by introducing static anchor points for the first time. The scheme consists of two efficient algorithms. The first algorithm estimates the relative timing and carrier frequency offsets, with respect to a randomly selected reference snapshot. Its estimation performance is analyzed with closed-form bias and root mean squared error derived. The estimates from the first algorithm are then used to eliminate clock offsets, enabling the construction of angle-Doppler maps. Using the maps, the second algorithm is developed to identify anchor points and then locate the transmitter and dynamic targets. The impact of the locations of the transmitter and anchor points is also analytically illustrated. Extensive simulation results are provided, demonstrating the effectiveness of the proposed sensing scheme in practical 5G communication setups and its superiority over prior art in terms of noise resilience and asymptotic performances.

***Index Terms*—Integrated sensing and communication (ISAC), Perceptive mobile networks, uplink sensing, clock asynchronism, static anchor points, parameter estimation.**

I. INTRODUCTION

A. Motivations and Backgrounds

ISAC [1]–[4] enables the sharing of a majority of hardware and network infrastructure to integrate communication and sensing into one system, which currently serves as a candidate technology for 6G mobile networks. With the application of ISAC in mobile networks, the perceptive mobile networks (PMNs) [1] can be achieved using either uplink [5], [6] or downlink signals [7]. The user equipment (UE) and base station (BS) in uplink sensing can be regarded as the transmitter and receiver in a *bistatic sensing* [1], respectively. The major challenge in such a bistatic uplink sensing is that the UE and BS typically use their local oscillators, leading to clock asynchronism. Clock asynchronism may generate time-varying timing offset (TMO) and carrier frequency offset (CFO), causing ambiguity in delay and Doppler estimation and preventing from coherent processing of multiple discontinuous signals [1], [8]–[10].

Several techniques have been developed in the literature to resolve the clock asynchronism issue [8], [11]–[15]. Among them, the cross-antenna cross-correlation (CACC) [11] and cross-antenna signal ratio (CASR) [14], [15] exploit the fact that all antennas in an array share the same clock oscillator, and utilize cross-antenna techniques for suppressing clock offsets. CACC calculates the conjugate multiplication between signals from different receiving antennas, successfully removing TMO and CFO. However, the presence of multiple cross-product terms in CACC doubles the parameters to be estimated and incurs the spectrum image problem. CASR can effectively eliminate the clock offset by computing the signal ratio between antennas, which, however, can only sense a single moving target in general. Additionally, in [13], the SHARP scheme leverages the LOS path or a strong NLOS path signal to remove TMO and CFO. By applying the compressed sensing algorithm at each OFDM symbol, the delay and phase of the LOS/NLOS path signals are estimated. Then, they can be utilized to compensate the sensing signal. However, the SHARP scheme strongly depends on the quality of the LOS/NLOS path signal. Furthermore, all the methods mentioned above require the UE position for bistatic sensing.

A limited number of techniques have been reported to relax these requirements. Without requiring knowledge of the UE's position, recently, [16] locates dynamic targets by assuming that the TMO and CFO are noise-like and follow zero-mean Gaussian distribution. By employing the Kalman filter (KF), [16] proposes a KF-based channel state information (CSI) enhancer to suppress the noise-like TMO. Since the LOS from UE to BS is the shortest, [16] treats the smallest range value as the LOS path range and then locates the UE. It is worth noting that the Kalman-filter-based algorithm does not utilize the LOS path signal to eliminate the clock offsets. However, the performance of the scheme depends on the actual distribution of the TMO and CFO. Regarding the requirement for LOS path, [17] analytically shows that non-LOS (NLOS) paths alone can also provide sufficient information for locating UE if and only if the estimates for angle-of-arrival (AOA), angle-of-departure (AOD), and range are sufficiently accurate. Specifically, three or more NLOS paths are 'possible' for generating unambiguous estimation for UE's position and orientation. In [18], it is proved that the localizability of using NLOS paths is low under under the Boolean model in 5G mm-wave systems. Nevertheless, [19] proposes a unified framework assisted by the factor-graph technique [20], which has the capability of achieving accurate localization in NLOS scenarios. The framework proves that leveraging multi-source information has the potential of increasing the localizability and enhancing the localization

---

Yanmo Hu and Weibo Deng are with the Key Laboratory of Marine Environmental Monitoring and Information Processing, Ministry of Industry and Information Technology, and also with the School of Electronic and Information Engineering, Harbin Institute of Technology. Email: 19B905008@stu.hit.edu.cn; dengweibo@hit.edu.cn. *(Corresponding author: Weibo Deng.)*

J. A. Zhang, K. Wu and Y. J. Guo are with Global Big Data Technologies Centre, the University of Technology Sydney. Email: {Andrew.Zhang; kai.wu; Jay.Guo}@uts.edu.au.

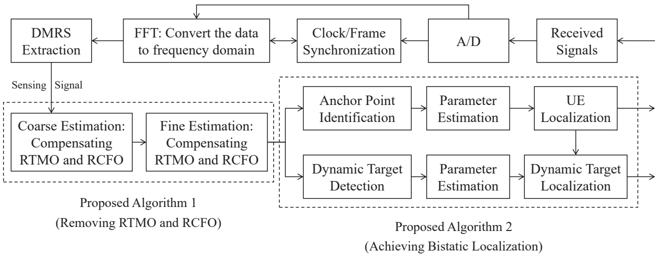


Fig. 1. The proposed uplink sensing scheme based on a typical mobile communication receiver diagram.

accuracy. Note that localization is an important application of bistatic sensing and we shall use the two terms interchangeably next.

Another potential method is to use anchor points, which are strong reflectors with known positions to the sensing receiver. For wireless localization, involving the localization of signal-emitting transmitters, it has been demonstrated that anchor points can efficiently reduce system costs and provide a solution for target localization [18], [21], [22], particularly in scenarios where a line-of-sight (LOS) path is absent. Anchor points can be classified into two groups: mobile and static anchor points. The mobile anchor points [21] are equipped with GPS, moving among the sensing area and periodically broadcasting their current positions to help the localization. The static anchor points [22] generally have known positions and are usually strong scatterers, thus are easy to be distinguished in the echo signals. Although anchor points have been demonstrated to be effective in assisting wireless localization, no work has yet to be reported for sensing in ISAC systems. For uplink sensing, significant challenges associated with clock asynchronism need to be overcome before signals and parameters from anchor points can be exploited.

B. Contributions and Organizations

In this paper, we consider the uplink sensing scenario with the following three challenges typically present simultaneously in PMNs:

1) The clock asynchronism between the UE and BS is present, which is practically inevitable and needs to be addressed for effective sensing. Otherwise, the sensing can suffer from ambiguities in the delay and Doppler estimation. Moreover, random phase offsets, as part of the clock asynchronism issue, can prevent coherent processing of multiple OFDM symbols;

2) There is no line-of-sight (LOS) path between the UE and BS. This can be quite common in practice, though many prior works rely on the LOS path;

3) The location of the UE is unknown to the BS. We note that bi-static sensing relies on the knowledge of a UE's position. Without the location, obtaining the absolute position of the target of interest becomes highly challenging.

Moreover, we note that our work considers the joint presence of the three challenges, which, to our knowledge, has not been studied yet. To address all these challenges effectively and efficiently, we develop an anchor-point-assisted novel sensing scheme for ISAC systems. The scheme will be presented by referring to a PMN involving a BS and UE, and it can also be applied to other networks. We assume the existence of several static anchor points near the BS. The proposed uplink sensing is implemented after extracting the demodulation reference signals (DMRS) from the received signals. Fig. 1 depicts how the proposed algorithms can fit into a typical mobile communication receiver. Algorithm 1 is designed to estimate relative TMO (RTMO) and relative CFO (RCFO) with respect to the values at the 0-th OFDM symbol. It includes two sub-algorithms: coarse estimation and refined estimation. Using the estimated RTMO and RCFO, Algorithm 1 compensates the TMO and CFO across multiple snapshots. Algorithm 2 obtains sensing parameters and achieves localization for UE and dynamic targets with the assistance of anchor points. Notably, our scheme does not require to know the UE's location or the existence of a line-of-sight (LOS) path between the BS and UE, thus addressing the limitations of existing technologies. This also makes our proposed scheme highly promising for practical uplink sensing in PMNs and general ISAC systems.

Our main contributions are summarized below:

- We propose a novel uplink sensing scheme that can address the clock asynchronism issue and does not rely on the presence of LOS path or UE's location information. In particular, we propose to estimate RTMO and RCFO first. Then, we introduce a novel idea of using anchor points to localize UE, which further facilitates accurate dynamic target localization.
- We propose a coarse estimation algorithm for RTMO and RCFO, which works efficiently when the power of the composite static path is larger than that of dynamic paths. Through the conjugate product across different OFDM symbols, we transform the clock asynchronism problem into a single-tone parameter estimation problem and then solve it using the maximum likelihood estimator (MLE). This algorithm works with both continuous or discontinuous CSI measurements either in the time or frequency domain.
- We also propose an iterative algorithm to obtain refined RTMO and RCFO estimation after compensating for clock offsets using the coarse estimates obtained earlier. This enables us to approximately construct the signal components of the composite static path, corresponding to $f_D = 0$. We further derive the bias and theoretical RMSE of the proposed estimators and identify the signal to noise ratio (SNR) regions where these estimates can have negligible impact on the subsequent parameter estimation for dynamic targets and anchor points. The theoretical results also demonstrate that Algorithm 1 is robust under wide ranges of RTMO and RCFO values.
- We develop an algorithm to extract sensing parameters. It is able to pinpoint anchor points, and locate both UE and dynamic targets. We analytically show that this algorithm works when the number of anchor points is not less than 2. We also analyze the feasibility of localization in relation to the locations of anchor points and UE.

The rest of this paper is organized as follows. Section II introduces the uplink sensing model without the LOS path. The

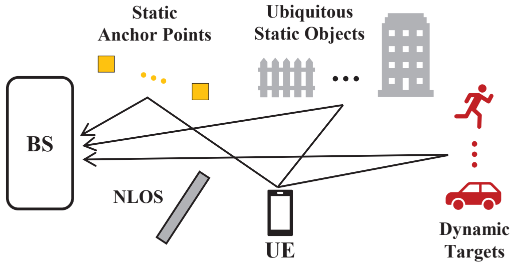


Fig. 2. The uplink sensing model without the LOS path.

RTMO and RCFO estimation algorithm is presented in Section III, including the coarse and refined estimation algorithms in Section III-A and III-B, respectively. The performance analysis is given in Section IV. Section V presents the localization algorithm for UE and dynamic targets. Simulation results are shown in Section VI, and conclusions are drawn in Section VII.

Notation: $\odot$ denotes the Hadamard product; $\circledast$ denotes the convolution; $\mathbf{A}^T$, $\mathbf{A}^*$ and $\mathbf{A}^H$ represent the transpose, conjugate and conjugate transpose of $\mathbf{A}$, respectively; $[\mathbf{a}]_n$ is the n-th element of vector $\mathbf{a}$, and $[\mathbf{A}]_{m,n}$ is the (m,n)-th element of matrix $\mathbf{A}$; $\angle\{a\}$ denotes the argument of the complex number; $\mathcal{F}_2$ is the two-dimensional discrete fourier transform; $\mathbf{I}_M \in \mathbb{R}^{M\times M}$ is identity matrix; $\mathbf{1}_{M\times 1} \in \mathbb{R}^{M\times 1}$ is the vector with all the elements are ones, while $\mathbf{0}_{M\times 1} \in \mathbb{R}^{M\times 1}$ is the vector filled by zeros. In particular, $\mathbf{1}_K \in \mathbb{R}^{K\times K}$ is the square matrix; $\mathbf{A}^\dagger = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T$ is the pseudo-inverse of $\mathbf{A}$; $\mathbf{e}_{M,m} \in \mathbb{R}^{M\times 1}$ is a vector with the m-th element equal to 1 and others equal to 0, where $m = 0, \cdots, M-1$.

II. Signal Model of Uplink Sensing

A. *Assumptions for Proposed Scheme*

The uplink sensing model is illustrated in Fig. 2. We consider an ISAC system with one *static* UE and one *static* BS. Unless the UE is carried by a high-speed vehicular platform, the UE's location change can be deemed negligible to the proposed sensing scheme which can work in a short time interval of tens to hundreds of milliseconds. The UE has M_t omnidirectional transmitter antennas, while the BS is equipped with an uniform linear array with M_r receiver antennas, and the antenna interval is half-wavelength. The uplink signal is received by the BS after being reflected by two types of objects: dynamic and static. We consider the typical case when most of reflected signals are from static objects. The number of dynamic objects is relatively low, but they are the targets of interest for sensing, such as human targets, cars, etc.

Similar to existing works on bistatic sensing [11], [12], we consider the existence of clock offset between the UE transmitter and BS receiver, which causes sensing ambiguity and prevents from coherent processing of discontinuous channel measurements [8]. However, unlike in most of these works, we consider *a more challenging scenario, where there is no LOS path between the UE and BS, and UE's location is unknown to the BS*. For a single BS where multiple antennas are closely located, multi-station-based techniques such as TDOA cannot be applied to solve the clock asynchronism problem. Hence, we emphasize the necessity of leveraging additional auxiliary information to resolve the clock offset in this scenario. Failing to do so could result in ambiguous range and Doppler estimates for the dynamic targets. The following two fundamental assumptions made in this paper can enable sensing in the absence of the LOS path between the UE and BS and without knowing the location of the UE:

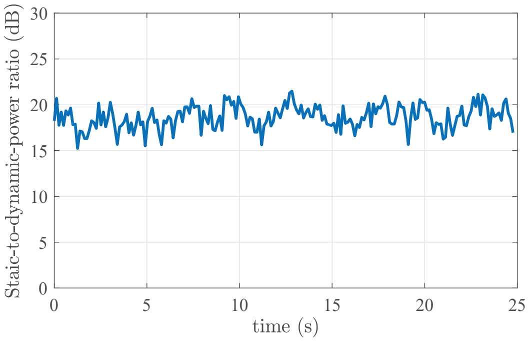


Fig. 3. Exemplified the static-to-dynamic-power ratio in WiFi system.

- Assumption One (A1): The total power of the electromagnetic wave reflected from the static objects is stronger than that of the dynamic targets;
- Assumption Two (A2): Assuming that the precise position of $L_a \geq 1$ static anchor points *close to* the BS is known, and signals reflected from them are strong and can be separated from other multipath signals by the measured AoA and power.

For A1, in many applications, there are often a large number of static objects/reflectors in the sensing environment, such as buildings and plants outdoors, and walls and furnitures indoors. They can lead to stronger reflected signals compared to dynamic targets, particularly when these dynamic targets are not very close to the transmitter or receiver. For example, in conventional radar sensing theory, where static clutter signals generally are much stronger than moving targets [23]. This assumption has been extensively observed in the WiFi sensing literature [24], [25], where WiFi sensing shares the commonality with the considered perceptive mobile network in the sense that the propagation channel can consist of myriads of static scatterers. To further validate the first assumption, we utilize the Intel 5300 NIC to detect two moving human targets in a meeting room and record the power variations of static and dynamic targets in a long time window of 25 seconds. As shown in Fig. 3, the power ratio between static and dynamic paths is consistently greater than 15 dB and can exceed 20 dB in quite a large portion of overall observation time. A1 provides a zero Doppler reference against the clock asynchronism influence, and is mainly used for estimating the RCFO and RTMO, as will be detailed in Section III.

For A2, the anchor points can be collected through the following two methods: 1) installing dedicated strong reflectors, such as the corner reflectors, which provide strong reflection at static positions; and 2) selecting several strong reflectors from existing static objects near the receiver. This can be achieved via pre-learning or long-period averaging the surrounding static objects using either uplink or downlink signals. For instance, lightning rods, as are common in many buildings, can serve for this purpose. This has been shown practically efficient in an existing work [26]. In scenarios where static objects are densely-spaced, anchor points can be obtained with ease. Note that we only consider passive reflective anchor points in this paper, although it is possible to extend our scheme to more advanced anchor points such as reconfigurable intelligent surface (RIS) with beam steering capabilities, which have been advocated for installation near BSs [27]. A2 is mainly used to remove a timing ambiguity to generate the estimates of the absolute locations of UE and dynamic targets, as will be discussed in Section V.

B. *Signal Model*

In what follows, we establish the signal model of uplink sensing based on the above assumptions. Our model is an abstract of 5G mobile signals. The demodulation reference signals (DMRS) for the physical uplink shared channel in 5G NR [28]–[30] have fixed and known signal values to the BS, which can be effectively used for uplink sensing. Suppose that the subcarrier interval is Δf and the number of subcarriers is N. Thus, the bandwidth is $B = N\Delta f$. Assume that the DMRS occupies the same subcarriers across OFDM symbols and it is periodic in the time domain, and the interval between two adjacent OFDM symbols is T_0. We use K to represent the number of used DMRS symbols within one coherent processing interval (CPI) for sensing. Since the clocks for BS and UE are not locked, transmissions can cause time-varying phase shift, TMO, and CFO to received DMRS symbols [1]. We use $e^{j\beta_k}$, $\tau_{o,k}$, and $f_{o,k}$ ($k = 0, \cdots, K-1$) to represent the phase shift, TMO, and CFO, respectively. Note that here the CFO is the small residual CFO after CFO estimation and compensation, which are typically applied in the time domain before signal processing in the frequency domain. Thus, the phase shift caused by the CFO can typically be approximated as a constant within one OFDM symbol. Moreover, it is worth noting that both static objects and dynamic targets are affected by the clock offsets in the same manner. However, since static targets are not affected by the Doppler frequencies, they can be utilized for eliminating the clock offset without affecting the dynamic targets' parameters, as will be detailed in Section III. Therefore, we separately treat these two types of targets in the work.

Assume that L targets are distributed in the space, with the propagation delay τ_ℓ, Doppler $f_{D,\ell}$, AOD φ_ℓ, and AOA θ_ℓ, where $\ell = 0, \cdots, L-1$ denotes the target index. We assume that the number of dynamic targets and static objects is L_d and L_s, respectively, and we have $L = L_d + L_s$. In this work, dynamic and static paths have non-zero and zero Doppler frequencies, respectively. For uplink sensing, the CSI for the n-th subcarrier ($n = 0, \cdots, N-1$) of the k-th OFDM symbol between the m_t-th antenna of UE and the m_r-th antenna of BS is given by [1], [8], [16]:

$$y_{n,k}^{(m_t,m_r)} = e^{j\beta_k} \sum_{\ell=0}^{L-1} \xi_\ell e^{-j2\pi n(\tau_\ell+\tau_{o,k})\Delta f} e^{j2\pi k(f_{D,\ell}+f_{o,k})T_0} \times e^{j\frac{2\pi}{\lambda} m_r d \sin\theta_\ell} e^{j\frac{2\pi}{\lambda} m_t d \sin\varphi_\ell} + z_{n,k}^{(m_t,m_r)}, \qquad (1)$$

where λ denotes the wavelength, ξ_ℓ denotes the value of the complex-valued reflection coefficient and the propagation attenuation, and $z_{n,k}^{(m)} \in \mathbb{C}$ is the additive white Gaussian noise matrix with $z_{n,k}^{(m)} \sim \mathcal{CN}\left(0, \sigma_n^2\right)$. The power attenuation of the NLOS path is $\frac{\lambda^2\sigma_x}{(4\pi)^3 R_T^2 R_R^2}$, where σ_x, R_T and R_R denote the reflecting factor, the transmitter-to-target range and the receiver-to-target range, respectively. Suppose that the noise is statistically independent at different subcarriers, OFDM symbols and receiver antennas.

In (1), $e^{j\beta_k}$ is the phase shift in the k-th OFDM symbol. This term has little influence on the communication, as it can be absorbed by channel estimation. While in sensing applications, $e^{j\beta_k}$ in different OFDM symbols must be equal or known, otherwise the ability of using complex signals to estimate the Doppler will be lost. However, $e^{j\beta_k}$ changes randomly across different OFDM symbols and is unknown both in BS and UE. Note that $e^{j\beta_k}$ in (1) is only dependent of the OFDM symbol index k, similar to the CFO term. In this paper, we combine the phase shift and CFO and eliminate their impact simultaneously.

Now, we rewrite (1) as:

$$y_{n,k}^{(m)} = e^{j\tilde{f}_{o,k}} \sum_{\ell=0}^{L-1} \xi_\ell e^{-j2\pi n(\tau_\ell+\tau_{o,k})\Delta f} e^{j2\pi k f_{D,\ell} T_0} a_\ell(m) + z_{n,k}^{(m)}, \qquad (2)$$

where we simplify the angle term for ease of derivation: $a_\ell(m) = e^{j\frac{2\pi}{\lambda} m_r d \sin\theta_\ell} e^{j\frac{2\pi}{\lambda} m_t d \sin\varphi_\ell}$ and $m = 0, \cdots, M_t M_r - 1$; and $\tilde{f}_{o,k} \triangleq \beta_k + 2\pi k f_{o,k} T_0$ combines the phase shift and the CFO, and $\tilde{f}_{o,k}$ is still regarded as the CFO.

III. ESTIMATION OF RELATIVE CFO AND RELATIVE TMO

In this section, the derivation is made only under the assumption of A1. We select reference values at a give time as the benchmarks to estimate the relative values of other CFO and TMO in this section, and then propose methods to estimate the reference values in Section V. Let $\tau_{o,0}$ and $\tilde{f}_{o,0}$ denote the reference values of TMO and CFO, respectively. The RTMO and RCFO values are thus defined by $\tau_{o,k}^R = \tau_{o,k} - \tau_{o,0}$ and $\tilde{f}_{o,k}^R = \tilde{f}_{o,k} - \tilde{f}_{o,0}$, $k = 0, \cdots, K-1$, respectively. Note that $\tau_{o,0}^R = 0$ and $\tilde{f}_{o,0}^R = 0$.

Based on A1, our main idea is to leverage the signal reflected from static objects and solve an optimization function to obtain the RTMO and RCFO. Note that the received signal in each received antenna is processed independently in this section.

A. *Estimation of RTMO and RCFO: Coarse Estimation*

Now, we collect the data from each subcarrier in the same OFDM symbol, i.e., $\mathbf{y}_k^{(m)} = \left[y_{0,k}^{(m)}, \cdots, y_{N-1,k}^{(m)}\right]^T \in \mathbb{C}^{N\times 1}$ with $k = 0, \cdots, K-1$, and construct the following vector based on (2):

$$\mathbf{y}_k^{(m)} = e^{j\tilde{f}_{o,k}} \operatorname{diag}\left\{\mathbf{t}_k\right\} \mathbf{\Gamma} \left[\boldsymbol{\xi} \odot \mathbf{d}_k \odot \mathbf{a}(m)\right], \qquad (3)$$

where $\mathbf{\Gamma} \in \mathbb{C}^{N\times L}$, $[\mathbf{\Gamma}]_{n,\ell} = e^{-j2\pi n \tau_\ell \Delta f}$ denotes the matrix containing all targets' delay;

$\mathbf{t}_k = \left[1, e^{-j2\pi\tau_{o,k}\Delta f}, \cdots, e^{-j2\pi(N-1)\tau_{o,k}\Delta f}\right]^T \in \mathbb{C}^{N\times 1}$ represents the TMO vector; $\boldsymbol{\xi} = [\xi_0, \cdots, \xi_{L-1}]^T \in \mathbb{C}^{L\times 1}$ denotes the vector of complex reflection coefficient; $\mathbf{d}_k = \left[e^{j2\pi k f_{D,0}T_0}, \cdots, e^{j2\pi k f_{D,L-1}T_0}\right]^T \in \mathbb{C}^{L\times 1}$ denotes the Doppler phase in the k-th OFDM block; and $\mathbf{a}(m) = [a_0(m), \cdots, a_{L-1}(m)]^T \in \mathbb{C}^{L\times 1}$.

Different from previous methods which employ the conjugate product operation between different antennas, such as the CACC [11] and the CSI-ratio method [14], the coarse estimation calculates the conjugate product between different OFDM symbols. Using the 0-th OFDM symbol as the reference, we now construct the following formula:

$$\begin{aligned}\varpi_{k,0}^{(m)} &= \mathrm{diag}\left\{\left(\mathbf{y}_0^{(m)}\right)^*\right\}\mathbf{y}_k^{(m)}\\ &= e^{j\left(\tilde{f}_{o,k}-\tilde{f}_{o,0}\right)}\mathrm{diag}\left\{\mathbf{t}_k \odot \mathbf{t}_0^*\right\}\\ &\quad\times \mathrm{diag}\left\{\boldsymbol{\Gamma}^*[\boldsymbol{\xi}\odot\mathbf{d}_0\odot\mathbf{a}(m)]^*\right\}\boldsymbol{\Gamma}\left[\boldsymbol{\xi}\odot\mathbf{d}_k\odot\mathbf{a}(m)\right]\\ &\approx e^{j\tilde{f}_{o,k}^R}\mathrm{diag}\left\{\mathbf{t}_k\odot\mathbf{t}_0^*\right\}\\ &\quad\times \mathrm{diag}\left\{\boldsymbol{\Gamma}^*[\boldsymbol{\xi}\odot\mathbf{a}(m)]^*\right\}\boldsymbol{\Gamma}\left[\boldsymbol{\xi}\odot\mathbf{a}(m)\right]\\ &= e^{j\tilde{f}_{o,k}^R}\mathrm{diag}\left\{\mathbf{t}_k^R\right\}\mathbf{h}^{(m)}, \end{aligned} \tag{4}$$

where $\varpi_{k,0}^{(m)} \in \mathbb{C}^{N\times 1}$, $k = 1, \cdots, K-1$ is the result of phase difference between the 0-th OFDM symbol and the k-th symbol, $\mathbf{t}_k^R \triangleq \left[1, e^{-j2\pi\tau_{o,k}^R\Delta f}, \cdots, e^{-j2\pi(N-1)\tau_{o,k}^R\Delta f}\right]^T \in \mathbb{C}^{N\times 1}$ contains RTMO values, $\mathbf{h}^{(m)} \triangleq \mathrm{diag}\left\{\boldsymbol{\Gamma}^*[\boldsymbol{\xi}\odot\mathbf{a}(m)]^*\right\}\boldsymbol{\Gamma}\left[\boldsymbol{\xi}\odot\mathbf{a}(m)\right] \in \mathbb{R}^{N\times 1}$ and each element in $\mathbf{h}^{(m)}$ is a positive and real number. Here, the approximation in the third step is based on A1.

The last line of (4) demonstrates that the initial phase and the frequency of $\varpi_{k,0}^{(m)}$ are the RCFO and RTMO in the k-th OFDM symbol, respectively. Herein, we normalize $\varpi_{k,0}^{(m)}$ in order to eliminate the impact from the amplitude $\mathbf{h}^{(m)}$:

$$\left[\tilde{\varpi}_{k,0}^{(m)}\right]_n = \left[\varpi_{k,0}^{(m)}\right]_n \Big/ \left|\left[\varpi_{k,0}^{(m)}\right]_n\right|. \tag{5}$$

where $n = 0, \cdots, N-1$. It is worth noting that (5) is a typical single-tone parameter estimation problem [31], and the MLE can be employed to estimate the unknown parameters. However, the noise component in (4) does not follow the Gaussian distribution. Here, we still hypothesize that the noise obeys Gaussian distribution in order to facilitate the application of MLE. Although the estimation is not sufficiently accurate, it is acceptable since we only find the coarse estimates. The approximated MLE formulas are given by:

$$\hat{\tau}_{o,k}^{R(m)} = \arg\max_{\tau_{o,k}^R}\left|\sum_{n=0}^{N-1} e^{j2\pi n\tau_{o,k}^R\Delta f}\left[\tilde{\varpi}_{k,0}^{(m)}\right]_n\right|, \tag{6a}$$

$$\hat{\tilde{f}}_{o,k}^{R(m)} = \angle\left\{\sum_{n=0}^{N-1} e^{j2\pi n\hat{\tau}_{o,k}^{R(m)}\Delta f}\left[\tilde{\varpi}_{k,0}^{(m)}\right]_n\right\}. \tag{6b}$$

Note that the fast Fourier transform (FFT) algorithm is conducive to reducing the computational complexity of (6a).

After performing (6) on the received signal at each antenna, we can obtain up to M_tM_r estimates of RTMO and M_tM_r estimates of RCFO. Note that to reduce complexity, we can compute and utilize M out of M_tM_r estimates, where M can be determined by the computational complexity. Thus, the final coarse estimates of RTMO and RCFO are given by:

$$\hat{\tau}_{o,k}^R = \frac{1}{M}\sum_{m=0}^{M-1}\hat{\tau}_{o,k}^{R(m)},\ \hat{\tilde{f}}_{o,k}^R = \frac{1}{M}\sum_{m=0}^{M-1}\hat{\tilde{f}}_{o,k}^{R(m)}, \tag{7}$$

where $M \le M_tM_r$. We set $\hat{\tau}_{o,0}^R = 0$ and $\hat{\tilde{f}}_{o,0}^R = 0$. It is worth noting that $\hat{\tau}_{o,k}^R$ and $\hat{\tilde{f}}_{o,k}^R$ are respectively equal to $\tau_{o,k}^R$ and $\tilde{f}_{o,k}^R$ when neither the dynamic targets nor the noise exists.

By utilizing (7), we can compensate the data in (2):

$$\begin{aligned}\tilde{y}_{n,k}^{(m)} &= e^{-j\hat{\tilde{f}}_{o,k}^R}e^{j2\pi n\hat{\tau}_{o,k}^R\Delta f}y_{n,k}^{(m)}\\ &= e^{j\left(\tilde{f}_{o,k}-\hat{\tilde{f}}_{o,k}^R+\tilde{f}_{o,0}-\tilde{f}_{o,0}\right)}e^{-j2\pi n\left(\tau_{o,k}-\hat{\tau}_{o,k}^R+\tau_{o,0}-\tau_{o,0}\right)\Delta f}\\ &\quad\times\sum_{\ell=0}^{L-1}\xi_\ell e^{-j2\pi n\tau_\ell\Delta f}e^{j2\pi k f_{D,\ell}T_0}a_\ell(m)\\ &= e^{j\Delta\tilde{f}_{o,k}^R}e^{-j2\pi n\Delta\tau_{o,k}^R\Delta f}\\ &\quad\times\sum_{\ell=0}^{L-1}\tilde{\xi}_\ell e^{-j2\pi n(\tau_\ell+\tau_{o,0})\Delta f}e^{j2\pi k f_{D,\ell}T_0}a_\ell(m), \end{aligned} \tag{8}$$

where $k = 0, \cdots, K-1$; $\Delta\tau_{o,k}^R = (\tau_{o,k}-\tau_{o,0}) - \hat{\tau}_{o,k}^R$ and $\Delta\tilde{f}_{o,k}^R = \left(\tilde{f}_{o,k}-\tilde{f}_{o,0}\right) - \hat{\tilde{f}}_{o,k}^R$ are the residual errors of coarse RTMO and RCFO, respectively, and are both small; $\tilde{\xi}_\ell = e^{j\tilde{f}_{o,0}}\xi_\ell$.

Remark 1. *In (8), $e^{j\tilde{f}_{o,0}}$ and $e^{-j2\pi n\tau_{o,0}\Delta f}$ are the residual terms after the compensation. Since $\tilde{f}_{o,0}$ is a constant and is independent of k, n and m, the term $e^{j\tilde{f}_{o,0}}$ has no influence on parameter (Doppler, delay, and angle) estimation. Thus, we can merge $e^{j\tilde{f}_{o,0}}$ with ξ_ℓ and define $\tilde{\xi}_\ell = e^{j\tilde{f}_{o,0}}\xi_\ell$. However, $e^{-j2\pi n\tau_{o,0}\Delta f}$ is dependent of n. It will lead to the delay ambiguity [1], and we will present how to remove its impact in Section V. In what follows, we discuss how to estimate $\Delta\tau_{o,k}^R$ and $\Delta\tilde{f}_{o,k}^R$.*

B. Estimation of RTMO and RCFO Errors: Refined Estimation

For ease of illustration, we reconstruct a set of new vectors along the time domain based on (8), i.e., $\tilde{\mathbf{y}}_n^{(m)} = \left[\tilde{y}_{n,0}^{(m)}, \cdots, \tilde{y}_{n,K-1}^{(m)}\right]^T \in \mathbb{C}^{K\times 1}$, $n = 0, \cdots, N-1$:

$$\tilde{\mathbf{y}}_n^{(m)} = \mathrm{diag}\left\{\Delta\tilde{\mathbf{f}}^R\odot\Delta\boldsymbol{\gamma}_n^R\right\}\mathbf{D}\left[\tilde{\boldsymbol{\xi}}\odot\tilde{\mathbf{s}}_n\odot\mathbf{a}(m)\right], \tag{9}$$

where $\mathbf{D} \in \mathbb{C}^{K\times L}$, $[\mathbf{D}]_{k,\ell} = e^{j2\pi k f_{D,\ell}T_0}$ represents the Doppler matrix composed of the Doppler phases of all targets; $\Delta\tilde{\mathbf{f}}^R = \left[1, e^{j\Delta\tilde{f}_{o,1}^R}, \cdots, e^{j\Delta\tilde{f}_{o,K-1}^R}\right]^T \in \mathbb{C}^{K\times 1}$ and $\Delta\boldsymbol{\gamma}_n^R = \left[1, e^{-j2\pi n\Delta\tau_{o,1}^R\Delta f}, \cdots, e^{-j2\pi n\Delta\tau_{o,K-1}^R\Delta f}\right]^T \in \mathbb{C}^{K\times 1}$ denote the vector of the estimation errors of coarse RCFO and RTMO, respectively; and $\tilde{\mathbf{s}}_n = \left[e^{-j2\pi n(\tau_0+\tau_{o,0})\Delta f}, \cdots, e^{-j2\pi n(\tau_{L-1}+\tau_{o,0})\Delta f}\right]^T \in \mathbb{C}^{L\times 1}$. The updated complex reflection coefficient vector is defined by $\tilde{\boldsymbol{\xi}} = e^{j\tilde{f}_{o,0}}\boldsymbol{\xi}$.

As aforementioned, the targets are composed of dynamic targets and static objects, with numbers L_d and L_s, respectively.

Equation (9) can thus be divided into two parts:

$$\begin{aligned}\tilde{\mathbf{y}}_n^{(m)} &= \operatorname{diag}\left\{\Delta\mathbf{f}^R \odot \Delta\boldsymbol{\gamma}_n^R\right\}\left[\mathbf{D}_d, \mathbf{D}_s\right]\begin{bmatrix}\tilde{\boldsymbol{\xi}}_d \odot \tilde{\mathbf{s}}_{d,n} \odot \mathbf{a}_d\left(m\right)\\ \tilde{\boldsymbol{\xi}}_s \odot \tilde{\mathbf{s}}_{s,n} \odot \mathbf{a}_s\left(m\right)\end{bmatrix}\\ &= \operatorname{diag}\left\{\Delta\mathbf{f}^R \odot \Delta\boldsymbol{\gamma}_n^R\right\}\mathbf{D}_d\left[\tilde{\boldsymbol{\xi}}_d \odot \tilde{\mathbf{s}}_{d,n} \odot \mathbf{a}_d\left(m\right)\right]\\ &+ \operatorname{diag}\left\{\Delta\mathbf{f}^R \odot \Delta\boldsymbol{\gamma}_n^R\right\}\mathbf{1}_{K\times L_s}\left[\tilde{\boldsymbol{\xi}}_s \odot \tilde{\mathbf{s}}_{s,n} \odot \mathbf{a}_s\left(m\right)\right], \quad (10)\end{aligned}$$

where the matrix or vector with subscripts 'd' and 's' belong to the dynamic targets and static objects, respectively. Note that the Doppler of static objects is zero, i.e., $\mathbf{D}_s = \mathbf{1}_{K\times L_s}$.

Although the phase of $\Delta\mathbf{f}^R \odot \Delta\boldsymbol{\gamma}_n^R$ is small, its impact cannot be overlooked. In (10), the static Doppler matrix is contaminated by $\Delta\mathbf{f}^R$ and $\Delta\boldsymbol{\gamma}_n^R$, and thus, the static object energy will diffuse into the non-zero Doppler regions, and the energy of zero Doppler frequency will be decreased. Our main idea is to utilize the Doppler characteristic of static objects and guarantee that the zero Doppler frequency possesses the maximum energy based on A1. The static object power can be expressed as:

$$\begin{aligned}\mathcal{P}_s &= \frac{1}{NK}\sum_{n=0}^{N-1}\sum_{k=0}^{K-1}\left|h_{s,n}^{(m)}\right|^2\\ &= \frac{1}{N}\sum_{n=0}^{N-1}\left|\frac{1}{K}\left(\Delta\mathbf{f}^R \odot \Delta\boldsymbol{\gamma}_n^R\right)^H \operatorname{diag}\left\{\Delta\mathbf{f}^R \odot \Delta\boldsymbol{\gamma}_n^R\right\}\mathbf{1}_{K\times 1}h_{s,n}^{(m)}\right|^2\\ &\approx \frac{1}{N}\sum_{n=0}^{N-1}\left|\frac{1}{K}\left(\Delta\mathbf{f}^R \odot \Delta\boldsymbol{\gamma}_n^R\right)^H \tilde{\mathbf{y}}_n^{(m)}\right|^2, \quad (11)\end{aligned}$$

where $h_{s,n}^{(m)} \triangleq \mathbf{1}_{1\times L_s}\left[\tilde{\boldsymbol{\xi}}_s \odot \tilde{\mathbf{s}}_{s,n} \odot \mathbf{a}_s\left(m\right)\right] \in \mathbb{C}$ denotes the static value at the n-th subcarrier of the m-th antenna. We now construct the cost function:

$$\mathcal{L}\left(\mathbf{c}, \boldsymbol{\kappa}\right) = \sum_{n=0}^{N-1}\left|\left(e^{j\mathbf{c}^T} \odot e^{-jn\boldsymbol{\kappa}^T}\right)\tilde{\mathbf{y}}_n^{(m)}\right|^2, \quad (12)$$

where $\mathbf{c}, \boldsymbol{\kappa} \in \mathbb{R}^{K\times 1}$, and it can be easily verified that $\max_{\mathbf{c},\boldsymbol{\kappa}}\mathcal{L}\left(\mathbf{c}, \boldsymbol{\kappa}\right) \approx NK^2\mathcal{P}_s$ when $\exp\left(j\mathbf{c}\right) = \Delta\mathbf{f}^R$ and $\exp\left(-jn\boldsymbol{\kappa}\right) = \Delta\boldsymbol{\gamma}_n^R$.

We have now constructed the relationship between $\mathcal{L}\left(\mathbf{c}, \boldsymbol{\kappa}\right)$ and static path power $\mathcal{P}_s$ based on A1, and $\Delta\mathbf{f}^R$ and $\Delta\boldsymbol{\gamma}_n^R$ can be estimated by maximizing $\mathcal{L}\left(\mathbf{c}, \boldsymbol{\kappa}\right)$. As aforementioned, the received signal at each received antenna is processed independently. We can formulate the following optimization problem to obtain $\Delta\mathbf{f}^R$ and $\Delta\boldsymbol{\gamma}_n^R$, which is energy-based:

$$\max_{\mathbf{c}^{(m)},\boldsymbol{\kappa}^{(m)}\in\mathbb{R}^{K\times 1}} \mathcal{L}\left(\mathbf{c}^{(m)}, \boldsymbol{\kappa}^{(m)}\right) \quad s.t.\ c_0^{(m)} = \kappa_0^{(m)} = 0. \quad (13)$$

Note that the cost function in (13) is nonlinear for both $\mathbf{c}^{(m)}$ and $\boldsymbol{\kappa}^{(m)}$, so that the closed-form solution does not exist. The Newton-Raphson method [32] can be utilized to maximize the cost function and calculate unknown vectors:

$$\begin{aligned}\mathbf{c}_1^{(m,q+1)} &= \mathbf{c}_1^{(m,q)} -\\ &\left[\frac{\partial\mathcal{L}\left(\mathbf{c}_1^{(m,q)}, \boldsymbol{\kappa}_1^{(m,q)}\right)}{\partial\mathbf{c}_1^{(m)}\partial\left(\mathbf{c}_1^{(m)}\right)^T}\right]^{-1}\frac{\partial\mathcal{L}\left(\mathbf{c}_1^{(m,q)}, \boldsymbol{\kappa}_1^{(m,q)}\right)}{\partial\mathbf{c}_1^{(m)}}\\ \boldsymbol{\kappa}_1^{(m,q+1)} &= \boldsymbol{\kappa}_1^{(m,q)} -\\ &\left[\frac{\partial\mathcal{L}\left(\mathbf{c}_1^{(m,q+1)}, \boldsymbol{\kappa}_1^{(m,q)}\right)}{\partial\boldsymbol{\kappa}_1^{(m)}\partial\left(\boldsymbol{\kappa}_1^{(m)}\right)^T}\right]^{-1}\frac{\partial\mathcal{L}\left(\mathbf{c}_1^{(m,q+1)}, \boldsymbol{\kappa}_1^{(m,q)}\right)}{\partial\boldsymbol{\kappa}_1^{(m)}}, \quad (14)\end{aligned}$$

where q denotes the iteration index, $\mathbf{c}_1^{(m)}, \boldsymbol{\kappa}_1^{(m)} \in \mathbb{R}^{(K-1)\times 1}$ are defined in $\mathbf{c}^{(m)} = \left[1, \left(\mathbf{c}_1^{(m)}\right)^T\right]^T$ and $\boldsymbol{\kappa}^{(m)} = \left[1, \left(\boldsymbol{\kappa}_1^{(m)}\right)^T\right]^T$, respectively. In particular, we set the initial values as $\mathbf{c}_1^{(m,0)} = \mathbf{0}_{(K-1)\times 1}$ and $\boldsymbol{\kappa}_1^{(m,0)} = \mathbf{0}_{(K-1)\times 1}$, since the RTMO and RCFO errors are small after the compensation in (8). Note that the Quasi-Newton methods can be exploited to approximate the inverse of the Hessian matrix.

After convergence, we obtain the estimates $\hat{\mathbf{c}}^{(m)} = \left[1, \hat{c}_1^{(m)}, \cdots, \hat{c}_{K-1}^{(m)}\right]^T \in \mathbb{C}^{K\times 1}$ and $\hat{\boldsymbol{\kappa}}^{(m)} = \left[1, \hat{\kappa}_1^{(m)}, \cdots, \hat{\kappa}_{K-1}^{(m)}\right]^T \in \mathbb{C}^{K\times 1}$. The terms $\Delta\tau_{o,k}^{R(m)}$ and $\Delta\tilde{f}_{o,k}^{R(m)}$ are related to $\hat{\kappa}_k^{(m)}$ and $\hat{c}_k^{(m)}$ via:

$$\Delta\hat{\tau}_{o,k}^{R(m)} = \frac{\hat{\kappa}_k^{(m)}}{2\pi\Delta f}, \quad \Delta\hat{\tilde{f}}_{o,k}^{R(m)} = \hat{c}_k^{(m)}. \quad (15)$$

Similar to (7), the refined estimates are obtained by the mean processing along m:

$$\begin{aligned}\Delta\hat{\tau}_{o,k}^R &= \frac{1}{M}\sum_{m=0}^{M-1}\Delta\hat{\tau}_{o,k}^{R(m)},\\ \Delta\hat{\tilde{f}}_{o,k}^R &= \frac{1}{M}\sum_{m=0}^{M-1}\Delta\hat{\tilde{f}}_{o,k}^{R(m)}, \quad (16)\end{aligned}$$

where M is the same as (7).

The estimates in (16) can then be used to compensate the residual errors in (8):

$$\begin{aligned}\tilde{\tilde{y}}_{n,k}^{(m)} &= e^{-j\Delta\hat{\tilde{f}}_{o,k}^R}e^{j2\pi n\Delta\hat{\tau}_{o,k}^R\Delta f}\tilde{y}_{n,k}^{(m)}\\ &\approx \sum_{\ell=0}^{L-1}\tilde{\xi}_\ell e^{-j2\pi n(\tau_\ell+\tau_{o,0})\Delta f}e^{j2\pi kf_{D,\ell}T_0}a_\ell(m). \quad (17)\end{aligned}$$

Only $\tau_{o,0}$ is unknown. We summarize the major steps of the RTMO and RCFO estimation algorithm in Algorithm 1.

Remark 2. *In this paper, the echo reflected from each target is regarded as a point source and there is only one echo for each object in delay, Doppler, and angle dimension, as illustrated in (1). In some practical applications, the micro-Doppler effect could be generated when a long coherent integration is performed [25]. We highlight that the proposed algorithm does not restrain the echo form and is also applicable to such challenging models with the micro-Doppler signatures. Our RTMO and RCFO estimation algorithm works as long as A1 is satisfied. In scenarios where dense static scatters are present, it is highly likely that the composite static paths has much higher power than dynamic paths, as validated in Fig. 3, which makes our proposed algorithms readily applicable.*

Algorithm 1 Proposed RTMO and RCFO Estimation Algorithm

```
Input: the received signal $y_{n,k}^{(m)}$, $m = 0, \cdots, M-1$, where $M \le M_t M_r$.
Output: the signal after RTMO and RCFO compensation $\tilde{\tilde{y}}_{n,k}^{(m)}$
for $m = 0$ to $M-1$ do
    Construct $\mathbf{y}_k^{(m)}$ using $y_{n,k}^{(m)}$. Calculate $\tilde{\boldsymbol{\varpi}}_{k,0}^{(m)}$ by (4) and (5), where $k = 1, \cdots, K-1$.
    Estimate the $\hat{\tau}_{o,k}^{R(m)}$ and $\hat{\tilde{f}}_{o,k}^{R(m)}$ by (6).
end for
Collect all of the estimates and perform (7) to calculate the coarse RTMO and RCFO estimates, i.e., $\hat{\tau}_{o,k}^{R}$ and $\hat{\tilde{f}}_{o,k}^{R}$. Utilize $\hat{\tau}_{o,k}^{R}$ and $\hat{\tilde{f}}_{o,k}^{R}$ to compensate $y_{n,k}^{(m)}$ and obtain $\tilde{y}_{n,k}^{(m)}$.
for $m = 0$ to $M-1$ do
    Set the initial values as $\mathbf{c}_1^{(m)} = \mathbf{0}_{(K-1)\times 1}$ and $\boldsymbol{\kappa}_1^{(m)} = \mathbf{0}_{(K)\times 1}$ and employ (14) to solve the optimization problem (13).
    Obtain $\Delta\tilde{f}_{o,k}^{R(m)}$ and $\Delta\tau_{o,k}^{R(m)}$ after convergence.
end for
Perform (16) to calculate the refined RTMO and RCFO estimates, i.e., $\Delta\hat{\tau}_{o,k}^{R}$ and $\Delta\hat{\tilde{f}}_{o,k}^{R}$.
Compensate $\tilde{y}_{n,k}^{(m)}$ by $\Delta\hat{\tau}_{o,k}^{R}$ and $\Delta\hat{\tilde{f}}_{o,k}^{R}$ and obtain $\tilde{\tilde{y}}_{n,k}^{(m)}$.
```

C. *Complexity of the Proposed Algorithm 1*

The proposed Algorithm 1 contains two sub-algorithms for coarse and fine estimation of the offsets. The key step of the coarse estimation is to perform the FFT to solve (6) at each OFDM symbol. Therefore, the computational complexity of the coarse estimation can be given by $\mathcal{O}\left\{MKN\log N\right\}$, where M, as defined in (7), is the number of RTMO and RCFO estimates utilized to enhance the estimation accuracy. The fine estimation obtains RTMO and RCFO via solving the optimization problem given in (13), where the number of unknown variables is $2\left(K-1\right)$. Though the Hessian matrix inversion is involved, the Quasi-Newton method can be exploited to substantially reduce the computational complexity, leading to a typical complexity of $\mathcal{O}\left\{MK^2\right\}$. Therefore, the overall complexity of Algorithm 1 is $\mathcal{O}\left\{\max\left\{MK^2, MKN\log N\right\}\right\}$.

IV. PERFORMANCE ANALYSIS OF ALGORITHM 1

In this section, we characterize the estimation errors for RTMO and RCFO, based on received signals from one antenna and $M=1$ in Algorithm 1.

A. *Derivation of RTMO and RCFO Estimation Errors*

The central part of our algorithm is the iteration algorithm in Section III-B since the calculation in Section III-A is only performed for finding initial values. In this subsection, we investigate the estimation performance of RTMO and RCFO for the proposed algorithm in Section III-B.

It is worth noting that there are three parts in $\tilde{\mathbf{y}}_n^{(m)}$ in (9): static, dynamic, and noise part. For ease of derivation, we redefine $\tilde{\mathbf{y}}_n^{(m)}$ based on A1:

$$\tilde{\mathbf{y}}_n^{(m)} \approx \mathrm{diag}\left\{\Delta\mathbf{f}^R \odot \Delta\boldsymbol{\gamma}_n^R\right\}\tilde{\mathbf{y}}_{s,n}^{(m)} + \mathbf{z}_n^{(m)}, \tag{18}$$

where $\tilde{\mathbf{y}}_{s,n}^{(m)} \in \mathbb{C}^{K\times 1}$ denotes the vector of the static part without the clock asynchronism.

Since c_0 and κ_0 are fixed in (13), we now define $\mathbf{z}_n^{\prime(m)}, \tilde{\mathbf{y}}_{s,n}^{\prime(m)} \in \mathbb{C}^{(K-1)\times 1}$ from $\tilde{\mathbf{y}}_{s,n}^{(m)} = \left[\tilde{y}_{s,n,0}^{(m)}, \left(\tilde{\mathbf{y}}_{s,n}^{\prime(m)}\right)^T\right]^T$, $\mathbf{z}_n^{(m)} = \left[z_{n,0}^{(m)}, \left(\mathbf{z}_n^{\prime(m)}\right)^T\right]^T$ to assist the derivation of estimation errors. Note that $\tilde{\mathbf{y}}_{s,n}^{(m)} = \mathbf{1}_{K\times 1}h_{s,n}^{(m)}$ based on (10). We have the following proposition.

Proposition 1. *The error of the vectors* $\mathbf{c}_1^{(m)} \in \mathbb{R}^{(K-1)\times 1}$ *and* $\boldsymbol{\kappa}_1^{(m)} \in \mathbb{R}^{(K-1)\times 1}$ *estimated by (14) is:*

$$\Delta\mathbf{c}_1^{(m)} \approx \left(\sum_{n=0}^{N-1}\left|h_{s,n}^{(m)}\right|^2\right)^{-1}\sum_{n=0}^{N-1}\mathrm{Im}\left\{\boldsymbol{\mu}_n^{(m)}\right\},$$

$$\Delta\boldsymbol{\kappa}_1^{(m)} \approx -\left(\sum_{n=0}^{N-1}n^2\left|h_{s,n}^{(m)}\right|^2\right)^{-1}\sum_{n=0}^{N-1}n\,\mathrm{Im}\left\{\boldsymbol{\mu}_n^{(m)}\right\}, \tag{19}$$

where $\boldsymbol{\mu}_n^{(m)} = \left(h_{s,n}^{(m)}\right)^*\left(\mathbf{z}_n^{\prime(m)} - z_{n,0}^{(m)}\mathbf{1}_{(K-1)\times 1}\right) \in \mathbb{C}^{(K-1)\times 1}$, *and* $h_{s,n}^{(m)}$ *is the value of the static part defined in (11).*

Proof: Please refer to Appendix A. ■

It is worth noting that Algorithm 1 generate biased estimates in the presence of dynamic targets. Fortunately, the denominator of (19) can be regarded as the static objects' power, i.e., $\mathcal{P}_s = \frac{1}{N}\sum_{n=0}^{N-1}\left|h_{s,n}^{(m)}\right|^2$, so the bias is small based on A1. Our algorithm can be considered to be approximately unbiased.

Remark 3. Estimation errors are irrelevant to RTMO and RCFO*: Note that* $\mathbf{z}_n^{(m)}$ *and* $h_{s,n}^{(m)}$ *in (19) are both independent of RTMO* ($\tau_{o,0}^R$) *and RCFO* ($\tilde{f}_{o,0}^R$)*. This indicates that the proposed algorithm will maintain the same estimation errors* $\Delta\mathbf{c}_1^{(m)}$ *and* $\Delta\boldsymbol{\kappa}_1^{(m)}$ *for any values of RTMO and RCFO, which shows the stability of the proposed algorithm to clock asynchronism.*

The estimation variance of RTMO and RCFO is given in Proposition 2.

Proposition 2. *For complex noise with equal variance of real and imaginary parts, i.e.,* $\mathrm{var}\left\{\mathrm{Im}\left\{\mathbf{z}_n^{(m)}\right\}\right\} = \mathrm{var}\left\{\mathrm{Re}\left\{\mathbf{z}_n^{(m)}\right\}\right\} = \mathbf{1}_{K\times 1}\sigma_n^2/2$ *. The estimation error variance is:*

$$\mathrm{var}\left\{\Delta\mathbf{c}_1^{(m)}\right\} \approx \sigma_n^2\left(\sum_{n=0}^{N-1}\left|h_{s,n}^{(m)}\right|^2\right)^{-1}\mathbf{1}_{(K-1)\times 1},$$

$$\mathrm{var}\left\{\Delta\boldsymbol{\kappa}_1^{(m)}\right\} \approx \sigma_n^2\left(\sum_{n=0}^{N-1}n^2\left|h_{s,n}^{(m)}\right|^2\right)^{-1}\mathbf{1}_{(K-1)\times 1}. \tag{20}$$

Proof: Please refer to Appendix B. ■

Note that $\sum_{n=0}^{N-1}n^2\left|h_{s,n}^{(m)}\right|^2 > \sum_{n=0}^{N-1}\left|h_{s,n}^{(m)}\right|^2$ in (19) and (20). The estimation for RTMO is generally more accurate than that for RCFO. Moreover, the RTMO accuracy is more sensitive to the value of N.

B. Influence of RTMO and RCFO Errors on Target Estimation

In this subsection, we use an example of FFT-based method [33] to estimate the sensing parameters and compare the sensing difference between the compensated signal $\tilde{\tilde{y}}_{n,k}^{(m)}$ in (17) and the signal without clock asynchronism. For the noise term, it is worth noting that the compensation in (8) and (17) only changes the noise phase, so the noise power in $\tilde{\tilde{y}}_{n,k}^{(m)}$ remains the same as the power of the original noise $z_{n,k}^{(m)}$. Therefore, we only need to consider the impact of RTMO and RCFO errors on the non-noise term:

$$\begin{aligned} \Delta\mathcal{F} &\approx \mathcal{F}_2\left\{e^{j\Delta c_k - jn\Delta\kappa_k}\tilde{y}_{s,n,k}^{(m)}\right\} - \mathcal{F}_2\left\{\tilde{y}_{s,n,k}^{(m)}\right\} \\ &= \frac{1}{NK}\mathcal{F}_2\left\{e^{j\Delta c_k - jn\Delta\kappa_k} - 1\right\} \circledast \mathcal{F}_2\left\{\tilde{y}_{s,n,k}^{(m)}\right\} \\ &\triangleq \mathcal{E}_\Delta \circledast \mathcal{F}_2\left\{\tilde{y}_{s,n,k}^{(m)}\right\}, \end{aligned} \tag{21}$$

where $\tilde{y}_{s,n,k}^{(m)}$ is defined from (18) with $\left[\tilde{\mathbf{y}}_{s,n}^{(m)}\right]_k = \tilde{y}_{s,n,k}^{(m)}$. For the influence of $\mathcal{E}_\Delta \circledast \mathcal{F}_2\left\{\tilde{y}_{s,n,k}^{(m)}\right\}$, we have the following proposition.

Proposition 3. *The influence of $\mathcal{E}_\Delta \circledast \mathcal{F}_2\left\{\tilde{y}_{s,n,k}^{(m)}\right\}$ can be neglected since its power is much lower than the noise power.*

Proof: Please refer to Appendix C. ■

To sum up, (17) can be directly used for target estimation since the RTMO and RCFO errors have ignorable influence in delay and Doppler. Moreover, the errors do not influence AOD and AOA estimation since the compensation in (17) is identical for the signal in different antennas.

V. LOCALIZATION OF UE AND DYNAMIC TARGETS

In this section, we present the proposed localization scheme for UE and dynamic targets assisted by the anchor points in the absence of the LOS path between UE and BS.

With the RTMO and RCFO estimated and compensated, conventional sensing techniques can be applied to estimate the parameters AOA (for dynamic targets only), AOD (for anchor points only), delay, and Doppler, almost like in a system without clock offset, except that the delay is relative to $\tau_{o,0}$. We first briefly describe how to estimate these parameters for dynamic targets and anchor points, and then elaborate how to estimate $\tau_{o,0}$ by exploiting the anchor points.

A. Parameter Estimation

For anchor points,we exploit the prior knowledge as assumed in A2. We first extract signals with $f_D = 0$, from which we construct the range-AOA spectrum initially. Since the AOA of each anchor point is precisely known according to A2, we extract the range spectrum based on the known AOA and record the *relative range* value (note that the range value is contaminated by $\tau_{o,0}$) corresponding to the highest energy point in each range spectrum. The recorded data belong to the anchor point if no extra static objects exist between the BS and the anchor point. Although interference may appear and impact the recognition of anchor points, we highlight that the interference is generally not constant, and the true anchor points can be effectively identified with high probability by multiple observations. In what follows, we do not consider the recognition failure case of anchor points.

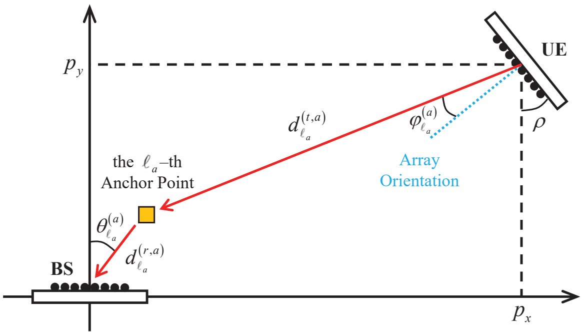


Fig. 4. Illustration of the parameters used for UE localization.

For dynamic target estimation, the static objects, regarded as interference, are able to be eliminated by Doppler filtering, such as the moving target detection (MTD) method [23]. Then, we perform the typical sensing algorithm to estimate the dynamic parameters, such as the MUSIC, ESPRIT, compressed sensing, FFT-based method, or tensor-based algorithms [1], [33]. Note that the estimated range is also contaminated by $\tau_{o,0}$.

We now define the estimated parameters as follows:

- Estimates of dynamic targets: $\hat{\boldsymbol{\theta}}_d = \left[\hat{\theta}_0^{(d)}, \cdots, \hat{\theta}_{L_d-1}^{(d)}\right]^T$ and $\hat{\tilde{\mathbf{R}}}_d = \left[\hat{\tilde{R}}_0^{(d)}, \cdots, \hat{\tilde{R}}_{L_d-1}^{(d)}\right]^T$ denote the AOA and the relative range of dynamic targets, respectively;
- Estimates of anchor points: $\hat{\boldsymbol{\varphi}}_a = \left[\hat{\varphi}_0^{(a)}, \cdots, \hat{\varphi}_{L_a-1}^{(a)}\right]^T$ and $\hat{\tilde{\mathbf{R}}}_a = \left[\hat{\tilde{R}}_0^{(a)}, \cdots, \hat{\tilde{R}}_{L_a-1}^{(a)}\right]^T$ denote the AOD and the relative range of anchor points, respectively.

where $\tilde{R}_{\ell_d}^{(d)}$ and $\tilde{R}_{\ell_a}^{(a)}$ denote the bistatic range contaminated by RTMO, i.e., $\tilde{R}_{\ell_d}^{(d)} = c\left(\tau_{o,0} + \tau_{\ell_d}\right)$ and $\tilde{R}_{\ell_a}^{(a)} = c\left(\tau_{o,0} + \tau_{\ell_a}\right)$. It is noted that we do not need to estimate the AOD of the dynamic targets for localization, and the AOA of anchor points is known as a priori.

B. Localization Algorithm

In this subsection, we utilize the parameters of L_a anchor points obtained in Section V-A to locate the UE position and estimate $\tau_{o,0}$ using the anchor point parameters, and subsequently locate each dynamic target. We underline that the array orientation of UE is unknown to the BS. As depicted in Fig. 4, we use ρ to describe the angle between the UE array and y-axis.

Let (p_x, p_y), $d_{\ell_a}^{(t,a)}$ and $d_{\ell_a}^{(r,a)}$ denote the cartesian coordinates of the UE, the range between the UE and the ℓ_a-th anchor point, and the range between the ℓ_a-th anchor point and the BS, respectively, as shown in Fig. 4. The absolute bistatic range of the ℓ_a-th anchor point can be expressed as $c\tau_{\ell_a} = d_{\ell_a}^{(t,a)} + d_{\ell_a}^{(r,a)}$, so $\tilde{R}_{\ell_a}^{(a)} = c\tau_{o,0} + c\tau_{\ell_a} = c\tau_{o,0} + d_{\ell_a}^{(t,a)} + d_{\ell_a}^{(r,a)}$. We can further derive the following relationships using the parameters of anchor points:

$$\tilde{R}_{\ell_a}^{(a)} - \tilde{R}_0^{(a)} = \left(d_{\ell_a}^{(t,a)} + d_{\ell_a}^{(r,a)}\right) - \left(d_0^{(t,a)} + d_0^{(r,a)}\right), \tag{22a}$$

$$p_x = \cos\left(\rho - \varphi^{(a)}_{\ell_a}\right) d^{(t,a)}_{\ell_a} + \sin\left(\theta^{(a)}_{\ell_a}\right) d^{(r,a)}_{\ell_a}, \tag{22b}$$

$$p_y = \sin\left(\rho - \varphi^{(a)}_{\ell_a}\right) d^{(t,a)}_{\ell_a} + \cos\left(\theta^{(a)}_{\ell_a}\right) d^{(r,a)}_{\ell_a}, \tag{22c}$$

where $\theta^{(a)}_{\ell_a}$ denotes the AOA of the ℓ_a-th anchor point, and $d^{(r,a)}_{\ell_a}$ and $\theta^{(a)}_{\ell_a}$ are precisely known based on A2. The index in (22a) is $\ell_a = 1, \cdots, L_a - 1$, whereas the index in (22b) and (22c) is $\ell_a = 0, \cdots, L_a - 1$. It is worth noting that $\tau_{o,0}$ is effectively eliminated in (22a). We now define three types of variables based on (22):

- $g^{(R)}_{\ell_a} \triangleq \left(\tilde{R}^{(a)}_{\ell_a} - \tilde{R}^{(a)}_0\right) - \left[\left(d^{(t,a)}_{\ell_a} + d^{(r,a)}_{\ell_a}\right) - \left(d^{(t,a)}_0 + d^{(r,a)}_0\right)\right]$, $\ell_a = 1, \cdots, L_a - 1$;
- $g^{(p_x)}_{\ell_a} \triangleq p_x - \left[\cos\left(\rho - \varphi^{(a)}_{\ell_a}\right) d^{(t,a)}_{\ell_a} + \sin\left(\theta^{(a)}_{\ell_a}\right) d^{(r,a)}_{\ell_a}\right]$, $\ell_a = 0, \cdots, L_a - 1$;
- $g^{(p_y)}_{\ell_a} \triangleq p_y - \left[\sin\left(\rho - \varphi^{(a)}_{\ell_a}\right) d^{(t,a)}_{\ell_a} + \cos\left(\theta^{(a)}_{\ell_a}\right) d^{(r,a)}_{\ell_a}\right]$, $\ell_a = 0, \cdots, L_a - 1$.

We can thus obtain $3L_a - 1$ equations. Stacking all unknown parameters into a vector $\varsigma = \left[p_x, p_y, \rho, d^{(t,a)}_0, \cdots, d^{(t,a)}_{L_a-1}\right]^T \in \mathbb{R}^{(L_a+3)\times 1}$, we construct the following problem:

$$\mathbf{g}(\varsigma) = \left[g^{(R)}_1, \cdots, g^{(R)}_{L_a-1}, g^{(p_x)}_0, \cdots, g^{(p_x)}_{L_a-1}, g^{(p_y)}_0, \cdots, g^{(p_y)}_{L_a-1}\right]^T = \mathbf{0}_{(3L_a-1)\times 1}, \tag{23}$$

where $\mathbf{g}(\varsigma) \in \mathbb{R}^{(3L_a-1)\times 1}$. This problem can be solved if the number of unknown parameters is less than the number of equations, i.e., $L_a + 3 \leq 3L_a - 1$. We obtain:

$$L_a \geq 2. \tag{24}$$

Estimating ς from (23) is a nonlinear problem and there is no closed-form solution. We exploit the non-linear least square method to get the estimates and locate the UE. The method is based on the iteration algorithm:

$$\varsigma^{(q+1)} = \varsigma^{(q)} - \mathbf{H}^{\dagger}\left(\varsigma^{(q)}\right)\mathbf{g}\left(\varsigma^{(q)}\right), \tag{25}$$

where q represents the iteration index; $\mathbf{H}(\varsigma) \in \mathbb{R}^{(3L_a-1)\times(L_a+3)}$ denotes the Jacobi matrix of $\mathbf{g}\left(\varsigma^{(q)}\right)$ given in (26) at the top of the next page. In (26), $\mathbf{d}^{(t,a)} = \left[d^{(t,a)}_0, \cdots, d^{(t,a)}_{L_a-1}\right]^T \in \mathbb{C}^{L_a\times 1}$, $\mathbf{T}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, and $\boldsymbol{\varepsilon}^{(t,a)} = \left[\sin\left(\rho - \varphi^{(a)}_0\right), \cos\left(\rho - \varphi^{(a)}_0\right), \cdots, \sin\left(\rho - \varphi^{(a)}_{L_a-1}\right), \cos\left(\rho - \varphi^{(a)}_{L_a-1}\right)\right]^T \in \mathbb{C}^{2L_a\times 1}$. Note that the iterative algorithm given in (25) can be transferred to Kalman filtering in multi-observation case to enhance the localization accuracy for UE.

The absolute location of UE is thus estimated as:

$$(\hat{p}_x, \hat{p}_y) = \left([\hat{\varsigma}]_0, [\hat{\varsigma}]_1\right). \tag{27}$$

Then, we estimate the reference TMO, i.e., $\tau_{o,0}$, based on the relationship $\tilde{R}^{(a)}_{\ell_a} = c\tau_{o,0} + d^{(t,a)}_{\ell_a} + d^{(r,a)}_{\ell_a}$ given above (22):

$$\hat{\tau}_{o,0} = \frac{1}{cL_a}\sum_{\ell_a=0}^{L_a-1}\left[\hat{\tilde{R}}^{(a)}_{\ell_a} - \left(\left\|(\hat{p}_x, \hat{p}_y) - \left(p^{(\mathrm{ANP})}_{x,\ell_a}, p^{(\mathrm{ANP})}_{y,\ell_a}\right)\right\|_2 + d^{(r,a)}_{\ell_a}\right)\right], \tag{28}$$

Algorithm 2 Proposed UE and Dynamic Target Localization Algorithm

```
Input: The estimates: $\hat{\tilde{R}}^{(d)}_{\ell_d}$, $\hat{\theta}^{(d)}_{\ell_d}$, $\hat{\tilde{R}}^{(a)}_{\ell_a}$ and $\hat{\varphi}^{(a)}_{\ell_a}$, where $\ell_d = 0, \cdots, L_d - 1$ and $\ell_a = 0, \cdots, L_a - 1$. The known position of anchor points: $\left(p^{(\mathrm{ANP})}_{x,\ell_a}, p^{(\mathrm{ANP})}_{y,\ell_a}\right)$.
Output: $(\hat{p}_x, \hat{p}_y)$, $\left(\hat{x}^{(d)}_{\ell_d}, \hat{y}^{(d)}_{\ell_d}\right)$, where $\ell_d = 0, \cdots, L_d - 1$.
Initialization: $\varsigma^{(0)} = \left[0, 0, 0, \hat{\tilde{R}}^{(a)}_0, \cdots, \hat{\tilde{R}}^{(a)}_{L_a-1}\right]^T$.
while Unconvergence do
    Update $\mathbf{g}\left(\varsigma^{(q)}\right)$ and the Jacobi matrix $\mathbf{H}\left(\varsigma^{(q)}\right)$.
    Calculate $\varsigma^{(q+1)}$ by (25). $q = q + 1$.
end while
Obtain $(\hat{p}_x, \hat{p}_y)$ by (27).
Calculate $\hat{d}^{(r,d)}_{\ell_d}$ by (28) (29). Obtain $\left(\hat{x}^{(d)}_{\ell_d}, \hat{y}^{(d)}_{\ell_d}\right)$ by (30).
```

where $\left(p^{(\mathrm{ANP})}_{x,\ell_a}, p^{(\mathrm{ANP})}_{y,\ell_a}\right)$ represents the cartesian coordinates of the ℓ_a-th anchor point.

Finally, we are ready to estimate the absolute locations for the dynamic targets. According to (28), the absolute delay of the ℓ_d-th dynamic target is $\hat{\tau}_{\ell_d} = \frac{1}{c}\hat{\tilde{R}}^{(d)}_{\ell_d} - \hat{\tau}_{o,0}$. Let $d^{(t,d)}_{\ell_d}$ and $d^{(r,d)}_{\ell_d}$ denote the range between UE and dynamic target and the range between dynamic target and BS, respectively. We obtain $c\tau_{\ell_d} = d^{(t,d)}_{\ell_d} + d^{(r,d)}_{\ell_d}$. Using $\hat{\tau}_{\ell_d}$, $(\hat{p}_x, \hat{p}_y)$ and $\hat{\theta}^{(d)}_{\ell_d}$, the following equation with respect to the ℓ_d-th dynamic target holds:

$$c\hat{\tau}_{\ell_d} = \left\|\left(d^{(r,d)}_{\ell_d}\sin\hat{\theta}^{(d)}_{\ell_d}, d^{(r,d)}_{\ell_d}\cos\hat{\theta}^{(d)}_{\ell_d}\right) - (\hat{p}_x, \hat{p}_y)\right\|_2 + d^{(r,d)}_{\ell_d}. \tag{29}$$

We can obtain $\hat{d}^{(r,d)}_{\ell_d}$ from (29). The absolute location of the ℓ_d-th dynamic target is:

$$\left(\hat{x}^{(d)}_{\ell_d}, \hat{y}^{(d)}_{\ell_d}\right) = \left(\hat{d}^{(r,d)}_{\ell_d}\sin\hat{\theta}^{(d)}_{\ell_d}, \hat{d}^{(r,d)}_{\ell_d}\cos\hat{\theta}^{(d)}_{\ell_d}\right). \tag{30}$$

We summarize the major steps of the UE and dynamic target localization algorithm in Algorithm 2. Note that $g(\varsigma)$ equals zero when ς is the true value in a noise-free environment. Hence, our method is unbiased.

C. Complexity of the Proposed Algorithm 2

The proposed Algorithm 2 consists of three parts: 1) estimating the parameters of anchor points and dynamic targets; 2) locating the UE; and 3) locating dynamic targets. We individually analyze the complexity of these parts, and then provide the overall computational complexity. The computational complexity of the first part is mainly dominated by several FFTs involved. The estimation of delay and Doppler can be separated from estimating the AOA and AOD [34]. The computational complexity of estimating delay and Doppler can be given by $\mathcal{O}\left\{(L_a + L_d) N_{\Delta} K_{\Delta} \log N_{\Delta} K_{\Delta}\right\}$, and the complexity of estimating AOA and AOD can be depicted as $\mathcal{O}\left\{(L_a + L_d) M_{t,\Delta} M_{r,\Delta} \log N_{t,\Delta} K_{r,\Delta}\right\}$, where N_{Δ}, K_{Δ}, $M_{r,\Delta}$, and $M_{t,\Delta}$ denote the number of the grids used to

$$\mathbf{H}(\varsigma) = \begin{bmatrix} \mathbf{0}_{(L_a-1)\times 2} & \mathbf{0}_{(L_a-1)\times 1} & \left[-\mathbf{1}_{(L_a-1)\times 1}, \mathbf{I}_{L_a-1}\right] \\ -\mathbf{1}_{L_a\times 1} \otimes \mathbf{I}_2 & \operatorname{diag}\left\{\boldsymbol{\varepsilon}^{(t,a)}\right\}\left(\mathbf{d}^{(t,a)} \otimes [-1,1]^T\right) & \left(\mathbf{I}_{L_a} \otimes \mathbf{T}_2\right)\operatorname{diag}\left\{\boldsymbol{\varepsilon}^{(t,a)}\right\}\left(\mathbf{I}_{L_a} \otimes \mathbf{1}_{2\times 1}\right) \end{bmatrix}. \tag{26}$$

---

estimate delay, Doppler, AOA, and AOD, respectively. Since $M_{r,\triangle}$ and $M_{t,\triangle}$ are generally small compared to $N_\triangle$ and $K_\triangle$, the computational complexity of the first part of Algorithm 2 can be approximated by $\mathcal{O}\left\{(L_a + L_d) N_\triangle K_\triangle \log N_\triangle K_\triangle\right\}$. To locate a UE, we need to solve the optimization problem given in (25), where the number of the unknown variables is $L_a + 3$. Similar to complexity analysis for Algorithm 1, the computational complexity for solving (23) can be given by $\mathcal{O}\left\{L_a^2\right\}$. For the last part of Algorithm 2, locating L_d dynamic targets generally has a linear complexity, as given by $\mathcal{O}\left\{L_d\right\}$. Based on the above analysis, we can assert that the overall computational complexity of Algorithm 2 is quite low in general, as given by $\mathcal{O}\left\{(L_a + L_d) N_\triangle K_\triangle \log N_\triangle K_\triangle\right\}$.

D. Feasibility of UE Localization

In this subsection, we evaluate the impact of noise on the localization. Similar to the analysis in (33), the UE localization accuracy is described by first-order approximation:

$$\Delta\varsigma \approx \frac{\partial \varsigma}{\partial \mathbf{e}_{\text{all}}^T} \Delta \mathbf{e}_{\text{all}} = -\mathbf{H}^\dagger(\varsigma) \frac{\partial \mathbf{g}(\varsigma)}{\partial \mathbf{e}_{\text{all}}^T} \Delta \mathbf{e}_{\text{all}}, \tag{31}$$

where $\mathbf{e}_{\text{all}} = \left[\mathbf{R}_a^T, \boldsymbol{\varphi}_a^T\right]^T \in \mathbb{R}^{2L_a\times 1}$ denotes the parameters used for localization; $\Delta \mathbf{e}_{\text{all}}$ and $\Delta\varsigma \in \mathbb{R}^{(L_a+3)\times 1}$ represent the estimation error vector and the localization error vector, respectively. We provide the following proposition to illustrate the situation that needs to be averted.

Proposition 4. *The localization errors for the UE are infinite if the UE and all anchor points are on the same line.*

Proof: When UE and all anchor points are on the same line, the AOD of anchor points is equal, i.e., $\varphi \triangleq \varphi_0^{(a)} = \cdots = \varphi_{L_a-1}^{(a)}$. In this case, however, it is easy to verify that $\left|\mathbf{H}^T(\varsigma)\,\mathbf{H}(\varsigma)\right| = 0$, which means $\mathbf{H}^\dagger$ is infinite. Thus, the proposed algorithm fails while $\varphi_0^{(a)} = \cdots = \varphi_{L_a-1}^{(a)}$. ■

Essentially, the data is correlated when UE and all anchor points are on the same line, reducing the number of non-correlated equations, so our algorithm fails to locate the UE. The solution to this problem is to employ more anchor points and avoid any three anchor points collinearly.

VI. SIMULATION RESULTS

In this section, simulation results are presented to validate the proposed algorithms. We first describe the system setup, then evaluate the sensing performance.

A. System Setting

The carrier frequency is set to 3GHz, the wavelength is $\lambda = 0.1$m, the subcarrier spacing of OFDM signal is $\Delta f = 480$kHz, and the number of subcarrier is $N = 256$. Therefore, the bandwidth is $B = N\Delta f = 122.88$MHz, the maximum unambiguous range is $R_{\max} = c/\Delta f = 625$m. The interval between OFDM symbols is $T_0 = 62.5\mu$s, and there are $K = 10\text{ms}/T_0 = 160$ OFDM symbols used for sensing. The noise power is $\sigma_n^2 = k_B F_n T_{st} B = 4.92 \times 10^{-12}$W, where k_B is the Boltamann's constant, $F_n = 10$ is the receiver noise figure, and $T_{st} = 290$K is the standard temperature.

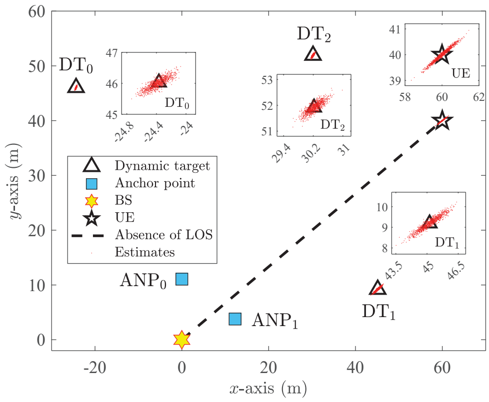


Fig. 5. Position of dynamic targets, BS, UE, and the localization results.

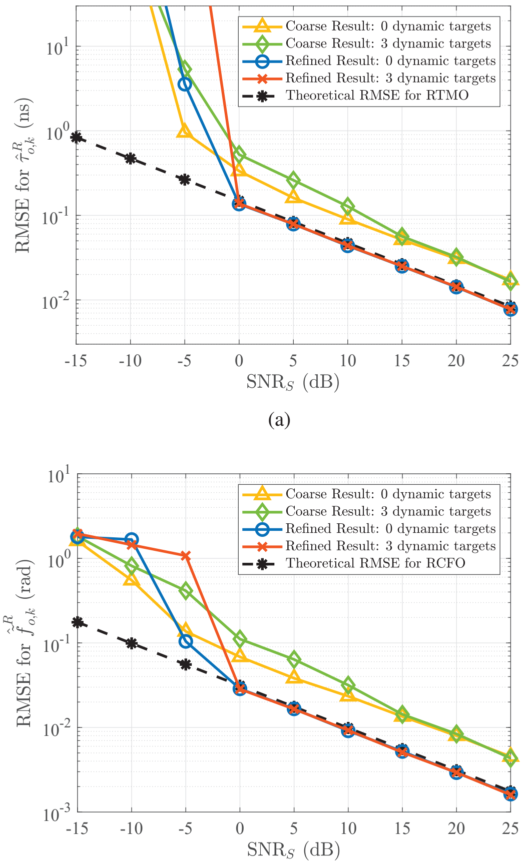
The sizes of antenna arrays of BS and UE are $M_r = 8$ and $M_t = 4$, respectively, and the antenna interval is set to be half-wavelength. Suppose that the positions of the static BS and UE are $(0\text{m}, 0\text{m})$ and $(60\text{m}, 40\text{m})$, respectively. The angle between the UE array and y-axis, as depicted in Fig. 4, is $\rho = 20°$. Moreover, three dynamic targets are distributed in the space, and the positions are $\text{DT}_0 = (-24.4\text{m}, 46.0\text{m})$, $\text{DT}_1 = (45.1\text{m}, 9.2\text{m})$, and $\text{DT}_2 = (30.2\text{m}, 51.9\text{m})$. The complex-valued reflection coefficient is set based on (1). Assume that the reflecting factor for dynamic target and the transmitted power are $\sigma_x = 1\text{m}^2$ and 0.1W, respectively.

In addition, $L_s = 15$ static objects are assumed to be randomly distributed in space. Based on the definition in Section IV, the averaged signal-to-noise ratio for static objects is defined as:

$$\text{SNR}_S = \frac{\mathcal{P}_s}{\sigma_n^2} = \frac{\sum\limits_{n=0}^{N-1}\left|h_{s,n}^{(m)}\right|^2}{N\sigma_n^2}. \tag{32}$$

Moreover, two anchor points are considered in this section, with positions of $\text{ANP}_0 = (0\text{m}, 11.1\text{m})$ and $\text{ANP}_1 = (12.3\text{m}, 3.8\text{m})$, as shown in Fig. 5. Assume that the reflecting factor for each anchor point is $\sigma_x = 2\text{m}^2$ [23]. In this case, the signal power of the static objects is $\text{SNR}_S = 4.13$dB, and is 11dB higher than that of the dynamic targets.

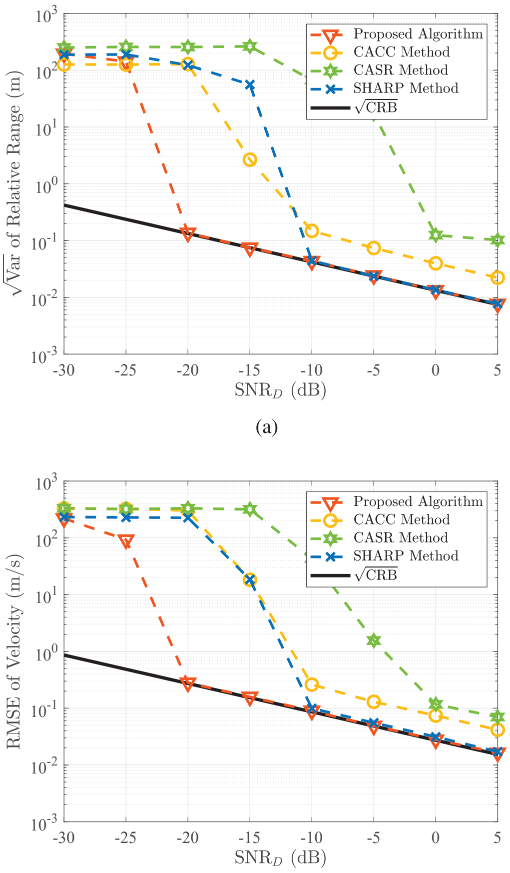
We use large CFO and TMO values in this section to verify Remark 3. The unambiguous TMO and CFO data is randomly distributed within $\left[\frac{0\text{m}}{c}, \frac{625\text{m}}{c}\right]$ and $[-\pi, \pi]$. In particular, we set $c\tau_{o,0} = 205.6$m.



(a)

(b)

Fig. 6. Results of Algorithm 1: the RMSEs of (a) $\hat{\tau}^R_{o,k}$ and (b) $\hat{\hat{f}}^R_{o,k}$ under different SNR$_S$ averaged over 1000 trials for each SNR$_S$ value.



(a)

(b)

Fig. 7. Estimation results of velocity and relative range versus SNR$_D$, where SNR$_D$ represents the power ratio between the dynamic paths and noise. Here, CACC, CASR and SHARP are benchmark methods, and CRB is derived from the signal model given in (1) with clock offsets suppressed.

B. Sensing Performance: Evaluation for Algorithm 1

First, we evaluate Algorithm 1 with $M = 4$. Fig. 6 illustrates the RMSE of $\hat{\tau}^R_{o,k}$ and $\hat{\hat{f}}^R_{o,k}$ versus SNR$_S$. In both figures, the refined estimation achieves a lower RMSE than the coarse estimation. In refined estimation, the RMSE is independent of the number of dynamic targets, which demonstrates the high efficacy of the refined estimator. In addition, it is worth noting that the RMSE of the refined estimate declines faster than the coarse estimate with the increment of SNR$_S$. This is because the noise in coarse estimation is no longer Gaussian distribution when the SNR is sufficiently high, referring to (5)-(6). It can be observed from Fig. 6 that the proposed Algorithm 1 works when SNR$_S \geq 0$dB. This indicates that our algorithm is effective when the signal power for static paths is only about 10dB higher than that for dynamic paths. From Fig. 6, we see that the coarse estimates exhibit better accuracy compared to the fine estimates when $-5\text{dB} \leq \text{SNR}_S \leq 0\text{dB}$. This is because the performance of the fine estimation algorithm is more dominant by the quality of the initialization than noises. Moreover, the initialization is provided by the coarse estimation whose accuracy is low in this SNR region. Finally, the RMSE of the refined estimate approaches to the theoretical RMSE provided in Proposition 2 for both $\hat{\tau}^R_{o,k}$ and $\hat{\hat{f}}^R_{o,k}$, validating the theoretical results.

Fig. 7 compares the performance of the proposed Algorithm 1 with three different methods: CACC [11], CASR [14], and SHARP [13]. Here, SNR$_D$ represents the SNR of the dynamic paths. The SNR of the static paths is set to SNR$_S$ = 12dB. Since Algorithm 1 produces a residual TMO, i.e., $\tau_{o,0}$ in (17), which is subsequently estimated by Algorithm 2, Algorithm 1 only provides a relative range estimation. Therefore, we evaluate the standard deviation of the relative range. Additionally, the Cramér-Rao bounds (CRBs) [32] of the parameters based on the signal model given in (1), yet with clock offsets suppressed, are plotted in Fig. 7. To suppress the impact of static path signal on dynamic target estimation, we filter the signal in Doppler domain to eliminate the static path signal after compensating for the RTMO and RCFO. Notably, the RMSE of the velocity and relative range estimates obtained by our proposed algorithm closely approach CRBs when SNR$_D \geq -20$dB, validating the effectiveness of the proposed Algorithm 1. Moreover, we see that the proposed method outperforms CACC and CASR in the whole SNR range considered. CACC mainly suffers from the mirror targets. For CASR, taking the ratio between signals over different antennas leads to a different signal model than

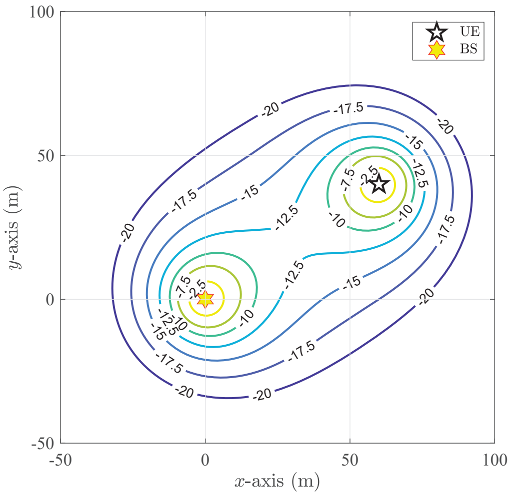
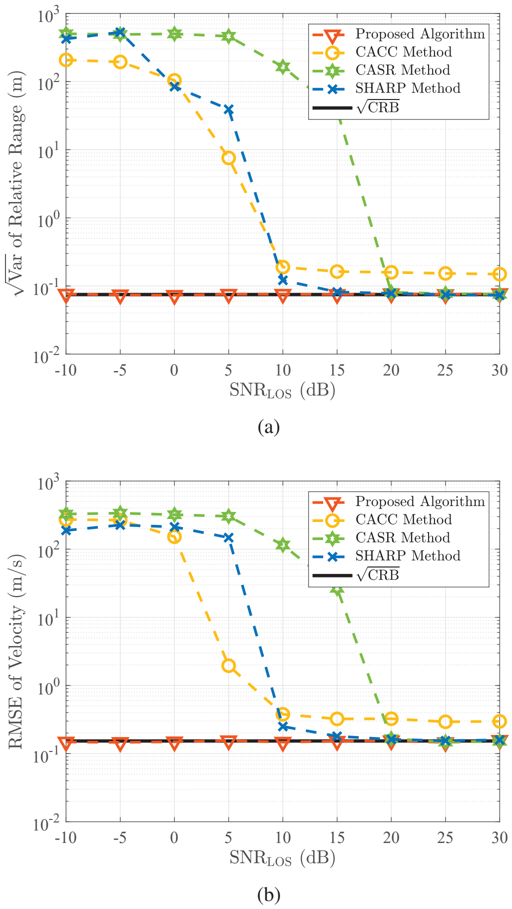


Fig. 8. SNR contour of a dynamic path when applying Algorithm 1.

(1). Therefore, the CRB plotted in Fig. 7 is not applicable to CASR, which explains the gap we see in Fig. 7 even in high SNR regions. We note that in a different yet relevant work [35], we have derived the CRB of CASR. In spite of the model difference, the proposed algorithm has better performance than CASR, especially when SNR_D is low. Comparing SHARP and our method, we see that they both approach the CRB in high SNR regions. However, as mentioned in Section I, SHARP requires a power-dominating path. This explains the superior noise resilience achieved by our design, as can be clearly seen in Fig. 7.

We note that multiplicative path attenuation can be incurred by the proposed Algorithm 1. Nevertheless, we note that the proposed algorithm embodies strong practicality in a typical 5G setup employed for simulations in Section VI-A. According to Fig. 7, the proposed algorithm still performs well when the power ratio between the dynamic paths and noise, i.e., SNR_D, is as low as -20dB. Therefore, we use this low SNR as a threshold to demonstrate the large sensing region covered by the proposed algorithm even with multiplicative attenuations. Fig. 8 illustrates the SNR contour of a single dynamic target on the two-dimensional space, given the parameters in Section VI-A. The power attenuation of the NLOS path is calculated as $\frac{\lambda^2 \sigma_x}{(4\pi)^3 R_T^2 R_R^2}$. The sensing region, which is $\mathrm{SNR}_D \geq -20$dB, can be approximated to an ellipse. In Fig. 8, this ellipse spans from -32m to 92m on the x-axis and from -34.5m to 74.5m on the y-axis. Therefore, our proposed Algorithm 1 is capable of detecting dynamic targets in a relatively large region, reflecting the practicality of the proposed algorithm.

Fig. 9 compares the proposed Algorithm 1 with three different methods under different SNRs of the LOS path, denoted as $\mathrm{SNR}_{\mathrm{LOS}}$. Although Algorithm 1 is designed for NLOS scenarios, it can also be applied to scenarios with a LOS path. In Fig. 9, we set $\mathrm{SNR}_S = 5$dB and $\mathrm{SNR}_D = -15$dB. The CACC method performs less effectively when $\mathrm{SNR}_{\mathrm{LOS}}$ is low, as this method utilizes the LOS path signal as a reference to eliminate clock offsets. For CASR and SHARP methods, their performances generally rely on the quality of the LOS/NLOS path signal, especially the static ones. In our settings, as $\mathrm{SNR}_{\mathrm{LOS}}$ decreases, both CASR and SHARP gradually deviate from the CRB. In contrast, the RMSEs of the proposed algorithm approach CRBs in a wide range of $\mathrm{SNR}_{\mathrm{LOS}}$. This is because the proposed algorithm utilizes the combination of all signal components with zero Doppler frequencies. When the LOS path is weak, the proposed algorithm can still leverage the strong composite static path to estimate the RTMO and RCFO.



(a)

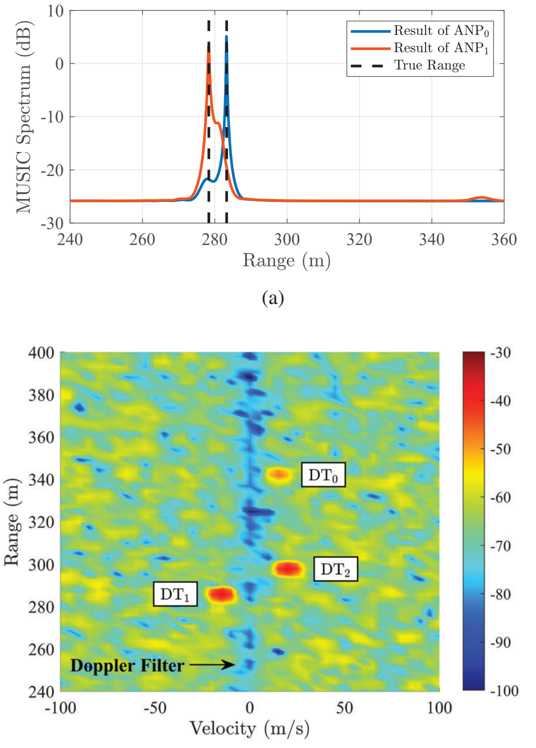
(b)

Fig. 9. Estimation results of velocity and relative range versus $\mathrm{SNR}_{\mathrm{LOS}}$, i.e., the power ratio between the LOS path and noise. Here, CACC, CASR and SHARP are benchmark methods, and CRB is derived from the signal model given in (1) with clock offsets suppressed.

C. Sensing Performance: Evaluation for Anchor Point Identification and Parameter Estimation

We first present the results for anchor point identification. The SNR of static objects is set to $\mathrm{SNR}_S = 4.13$dB. We obtain the $f_D = 0$ data from $\tilde{\tilde{y}}_{n,k}^{(m)}$ (17) at each receiving antenna, and perform the spatial-smoothing-based MUSIC algorithm to calculate the noise subspace about the range and AOA.

Utilizing the known AOA of two anchor points, we obtain two range MUSIC spectrums with respect to the two AOAs, shown in Fig. 10(a). As can be seen, in addition to the anchor points, there are other peaks in the range MUSIC spectrum,



(a)

(b)

Fig. 10. Results of dynamic target and anchor point estimation after RTMO and RCFO compensation, with $c\tau_{o,0} = 205.6$m, exemplified by SNR$_S = 4.13$dB. (a) Result of identifying anchor points in range MUSIC spectrum; (b) Result after compensating and Doppler-domain filtering.

which belong to the static objects. According to A2, the power of the anchor points is higher than that of the static objects, so the SNR of the anchor points' signal is higher. When estimating the range using the MUSIC algorithm, a higher SNR will produce better orthogonality, leading to a higher peak in MUSIC spectrum. Thus, we select the peaks with the maximum amplitude in Fig. 10(a) as the anchor points, as shown in Fig. 10(a). After identifying through MUSIC spectrum, we estimate the range and AOD of each anchor point, and the results are given in Table I. The AOD estimates are unbiased, while the range estimates are offset by the reference TMO, i.e., $c\tau_{o,0} = 205.6$m.

Next, we remove static signals using a Doppler filter with transfer function $H(z) = \frac{1-0.9z^{-1}}{1-z^{-1}}$, as described in Section V-A, and the range-Doppler spectrum result is shown in Fig. 10(b). It can be observed that the signals with zero Doppler are entirely eliminated, and the dynamic targets are all retained. Then, we estimate the range and AOA of dynamic targets using the DFT-based algorithm, and the bias and RMSE results are given in Table I, averaging over 1000 trials. Note that the dynamic targets and anchor points have the same biased values of range, i.e., $c\tau_{o,0} = 205.6$m. This is used as the basis for removing $c\tau_{o,0}$ using Algorithm 2. In addition, the deviation of the estimated AOA and AOD is less than 10^{-4} degree, and the deviation of unbiased range estimate (without considering the reference TMO $c\tau_{o,0} = 205.6$m) is less than 1mm.

TABLE I
ESTIMATION RESULTS OF DYNAMIC TARGETS AND ANCHOR POINTS

	ANP$_0$	ANP$_1$	DT$_0$	DT$_1$	DT$_2$
Bistatic range (bias, mm) (Add $c\tau_{o,0}$)	< 1	< 1	< 1	< 1	< 1
AOD (bias, degree)	$< 10^{-4}$	$< 10^{-4}$	—	—	—
AOA (bias, degree)	—	—	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
Bistatic range (RMSE, mm)	5.26	5.58	27.99	9.96	12.03
AOD (RMSE, degree)	0.06	0.06	—	—	—
AOA (RMSE, degree)	—	—	0.19	0.29	0.09

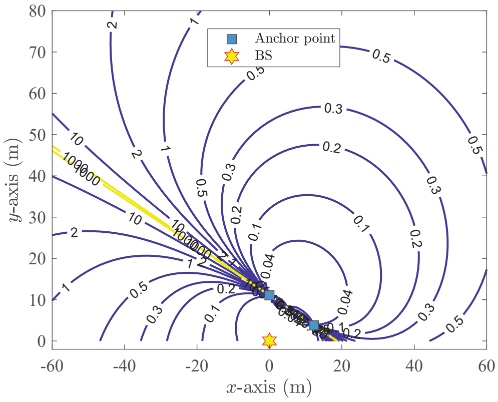


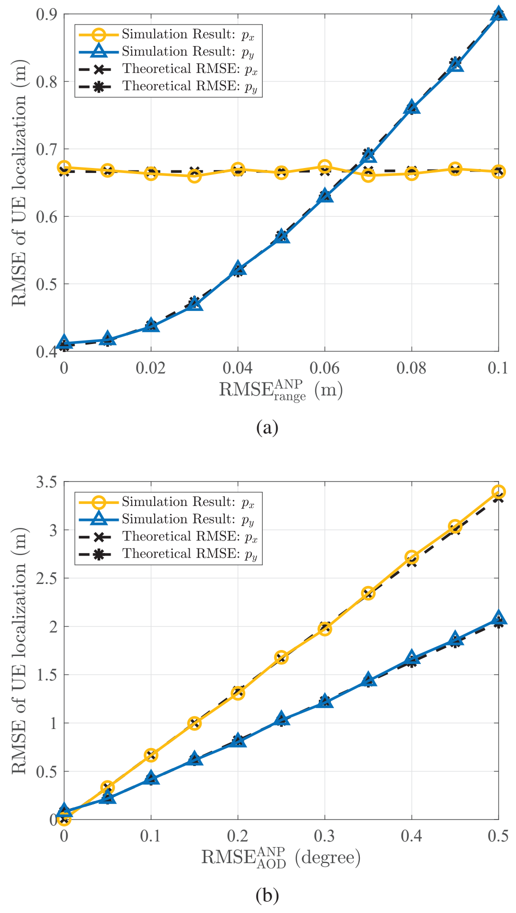
Fig. 11. Localization accuracy of the UE at different locations using two anchor points. Parameters are provided in Table I and the result is given in meters.

D. Sensing Performance: Evaluation for Algorithm 2

Next, Algorithm 2 is evaluated with the assumption that all the anchor points are successfully identified in the range-AOA spectrum. For the first evaluation, we use the output of Algorithm 1 (given in Table I) as the input of Algorithm 2 to locate UE and dynamic targets. After performing 1000 trials, the location results of three dynamic targets and one UE are ploted in Fig. 5, where each red dot represents the location estimate in each independent trial. Clearly, all the estimate points fell precisely on the correspondence UE and dynamic targets in Fig. 5. These results clearly demonstrate that the proposed model in Section V is able to effectively overcome the impact of $\tau_{o,0}$ as well as the absence of LOS.

To evaluate the impact of UE's location on the localization accuracy, we vary its locations and present the results in Fig. 11, using two anchor points. Note that the localization accuracy for the UE is defined as $\sqrt{\mathrm{var}\{\hat{p}_x\} + \mathrm{var}\{\hat{p}_y\}}$. Clearly, the localization accuracy is infinite once the location of UE is collinear with two anchor points, which validates Proposition 4. Employing a third anchor point is able to solve this problem. Moreover, the farther away from the anchor points, the lower the localization accuracy. In practical applications, the accuracy of the area of interest can be improved by selecting the anchor points with optimal positions.

Fig. 12 shows the RMSE results of UE localization compared with theoretical RMSEs. The theoretical RMSEs are given by (31). In particular, all the estimation results are unbiased. Both figures show that the simulated results match the theoretical RMSEs well under different RMSE$^{\mathrm{ANP}}_{\mathrm{range}}$ and RMSE$^{\mathrm{ANP}}_{\mathrm{AOD}}$.



(a)

(b)

Fig. 12. RMSE results of UE localization under different RMSEs of anchor point. (a) RMSE results under different $\text{RMSE}^{\text{ANP}}_{\text{range}}$, with $\text{RMSE}^{\text{ANP}}_{\text{AOD}} = 0.1°$; (b) RMSE results under different $\text{RMSE}^{\text{ANP}}_{\text{AOD}}$, with $\text{RMSE}^{\text{ANP}}_{\text{range}} = 0.01\text{m}$.

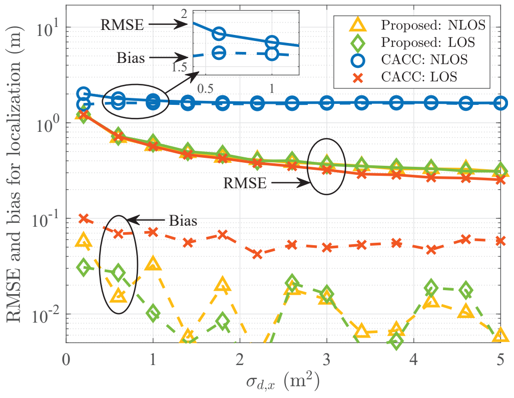


Fig. 13. Localization performance of CACC and the proposed scheme.

Compared with Fig. 12(b), it can also be seen that the localization RMSE is more sensitive to the RMSE $^{\text{ANP}}_{\text{AOD}}$ values of anchor points. Interestingly, the UE localization RMSE is nearly zero when RMSE $^{\text{ANP}}_{\text{AOD}}$ equals zero, which means a more accurate AOD estimation for anchor points leads to more precise UE localization.

Fig. 13 compares the localization performance between CACC and the proposed algorithms (cascading Algorithm 1 and Algorithm 2) for different reflecting factors of DT1. The simulation setting is described in Section VI-A. For the LOS scenario, we added a LOS path signal into the echo and the power attenuation follows $\frac{\lambda^2}{(4\pi R_{\text{LOS}})^2}$, where R_{LOS} denotes the LOS path range. Fig. 13 shows that CACC and the proposed scheme achieve similar RMSEs and are both unbiased in the LOS case. Thus, our method achieves nearly the same localization accuracy without requiring to know the position of the UE. In NLOS scenario, the proposed scheme still has the same accuracy as in LOS scenario and is also unbiased. However, the CACC method fails with significant bias in the NLOS scenario since no LOS path serves as the reference path for localization.

VII. CONCLUSION

We have presented a novel anchor-point assisted uplink sensing scheme for ISAC systems with clock asynchronism between transmitter and receiver. Our scheme can be applied to more general scenarios where the transmitter location is unknown to the receiver and there is no dominating LOS path between them. The scheme consists of two main algorithms. Algorithm 1 estimates RTMO and RCFO with respect to those at a reference snapshot. Algorithm 2 pinpoints the anchor points from the static objects, locates the UE, and then locates dynamic targets. We analytically show that two anchor points are generally sufficient, and three non-collinear anchor points can guarantee to work. The simulation results demonstrate that our proposed scheme can effectively locate the dynamic targets and UE, without requiring the LOS path or knowing the UE position, and is very promising for practical applications.

Using anchor points, the proposed scheme can evidently extend the sensing field-of-view of the BS, especially in urban areas where line-of-sight paths are often blocked by buildings and plants. We note that the positions of the selected anchor points have significant influence on the proposed uplink sensing scheme. Though we have identified and revealed the theoretical impact of the locations, optimizing them is not solved yet, and can be an interesting future work. Moreover, the proposed sensing scheme can be extended to multi-static scenarios with more than one sensing stations. In such scenarios, factor graph-based localization scheme can be resorted for fusing multi-observation data, potentially enhancing the localization and tracking of dynamic targets in uplink sensing.

APPENDIX A
PROOF OF PROPOSITION 1

Take the RCFO as an example. The formula $\mathbf{g}_c = \frac{\partial}{\partial \mathbf{c}_1^{(m)}}\mathcal{L}\left(\mathbf{c}_1^{(m)}, \boldsymbol{\kappa}_1^{(m)}\right) = \mathbf{0}_{(K-1)\times 1}$ holds while (14) converges. For ease of derivation, we define $\mathbf{w}_c^{(m)} = \exp\left(j\mathbf{c}_1^{(m)}\right) \in \mathbb{C}^{(K-1)\times 1}$, $\mathbf{w}_\kappa^{(m)} = \exp\left(j\boldsymbol{\kappa}_1^{(m)}\right) \in \mathbb{C}^{(K-1)\times 1}$, and $\mathbf{w}_n^{(m)} = \mathbf{w}_c^{(m)} \odot \left(\mathbf{w}_\kappa^{(m)}\right)^{-n} \in \mathbb{C}^{(K-1)\times 1}$. Based on A1, each element in $\tilde{\mathbf{y}}_{s,n}^{(m)}$ is large. The impact of $\mathbf{z}_n^{(m)}$ on $\mathbf{w}_c^{(m)}$ can be expressed by the first-order approximation of $\mathbf{w}_c^{(m)}$ in the point $\tilde{\mathbf{y}}_{s,n}^{\prime(m)}$, where $\tilde{\mathbf{y}}_{s,n}^{\prime(m)} \in \mathbb{C}^{(K-1)\times 1}$ is defined from $\tilde{\mathbf{y}}_{s,n}^{(m)} = \left[\tilde{y}_{s,n,0}^{(m)}, \left(\tilde{\mathbf{y}}_{s,n}^{\prime(m)}\right)^T\right]^T$:

$$\begin{aligned}\Delta\mathbf{w}_c^{(m)} &\approx \sum_{n=0}^{N-1} \frac{\partial \mathbf{w}_c^{(m)}}{\partial\left(\tilde{\mathbf{y}}_{s,n}^{\prime(m)}\right)^T}\mathbf{z}_n^{(m)} \\ &\approx -\left[\frac{\partial \mathbf{g}_c}{\partial\left(\mathbf{w}_c^{(m)}\right)^T}\right]^{-1} \sum_{n=0}^{N-1} \frac{\partial \mathbf{g}_c}{\partial\left(\tilde{\mathbf{y}}_{s,n}^{\prime(m)}\right)^T}\mathbf{z}_n^{(m)}, \end{aligned} \quad (33)$$

where the implicit differentiation is used.

The derivation of $\Delta\mathbf{w}_\kappa^{(m)}$ is identical to the process for deriving (33), and we can get the same result as (33). With $\tilde{\mathbf{y}}_{s,n}^{(m)} = \mathbf{1}_{K\times 1} h_{s,n}^{(m)}$, we obtain four partial derivative results:

$$\begin{aligned}\frac{\partial \mathbf{g}_c}{\partial\left(\tilde{\mathbf{y}}_{s,n}^{\prime(m)}\right)^T} &= j\left(h_{s,n}^{(m)}\right)^* \left[\mathbf{1}_{(K-1)\times 1}, \mathbf{1}_{K-1} - K\mathbf{I}_{K-1}\right] \\ \frac{\partial \mathbf{g}_\kappa}{\partial\left(\tilde{\mathbf{y}}_{s,n}^{\prime(m)}\right)^T} &= -jn\left(h_{s,n}^{(m)}\right)^* \left[\mathbf{1}_{(K-1)\times 1}, \mathbf{1}_{K-1} - K\mathbf{I}_{K-1}\right] \\ \frac{\partial \mathbf{g}_c}{\partial\left(\mathbf{w}_c^{(m)}\right)^T} &= -j\sum_{n=0}^{N-1}\left|h_{s,n}^{(m)}\right|^2\left(\mathbf{1}_{K-1} - K\mathbf{I}_{K-1}\right) \\ \frac{\partial \mathbf{g}_\kappa}{\partial\left(\mathbf{w}_\kappa^{(m)}\right)^T} &= -j\sum_{n=0}^{N-1}n^2\left|h_{s,n}^{(m)}\right|^2\left(\mathbf{1}_{K-1} - K\mathbf{I}_{K-1}\right). \end{aligned} \quad (34)$$

Substituting (34) into (33) and using $\left(\mathbf{1}_{K-1} - K\mathbf{I}_{K-1}\right)^{-1} = -K^{-1}\left[\mathbf{1}_{K-1} + \mathbf{I}_{K-1}\right]$ yields:

$$\Delta\mathbf{w}_c^{(m)} \approx \frac{\sum\limits_{n=0}^{N-1}\boldsymbol{\mu}_n^{(m)}}{\sum\limits_{n=0}^{N-1}\left|h_{s,n}^{(m)}\right|^2}, \quad \Delta\mathbf{w}_\kappa^{(m)} \approx \frac{-\sum\limits_{n=0}^{N-1}n\boldsymbol{\mu}_n^{(m)}}{\sum\limits_{n=0}^{N-1}n^2\left|h_{s,n}^{(m)}\right|^2}, \quad (35)$$

where $\boldsymbol{\mu}_n^{(m)} = \left(h_{s,n}^{(m)}\right)^*\left(\mathbf{z}_n^{\prime(m)} - z_{n,0}^{(m)}\mathbf{1}_{(K-1)\times 1}\right) \in \mathbb{C}^{(K-1)\times 1}$.

Next, we perform the first-order approximation with respect to phase to calculate $\Delta\mathbf{c}_1^{(m)}$ and $\Delta\boldsymbol{\kappa}_1^{(m)}$. Note that the phase of $\mathbf{w}_c^{(m)}$ and $\mathbf{w}_\kappa^{(m)}$ is small after the coarse compensation. We can obtain the following results:

$$\begin{aligned}\Delta\mathbf{c}_1^{(m)} &= \angle\left\{\mathbf{w}_c^{(m)} + \Delta\mathbf{w}_c^{(m)}\right\} - \angle\left\{\mathbf{w}_c^{(m)}\right\} \approx \operatorname{Im}\left\{\Delta\mathbf{w}_c^{(m)}\right\} \\ \Delta\boldsymbol{\kappa}_1^{(m)} &= \angle\left\{\mathbf{w}_\kappa^{(m)} + \Delta\mathbf{w}_\kappa^{(m)}\right\} - \angle\left\{\mathbf{w}_\kappa^{(m)}\right\} \approx \operatorname{Im}\left\{\Delta\mathbf{w}_\kappa^{(m)}\right\}. \end{aligned} \quad (36)$$

Substituting (35) into (36), proposition 1 is thus proved.

APPENDIX B
PROOF OF PROPOSITION 2

Use $\operatorname{var}\left\{\Delta\mathbf{c}_1^{(m)}\right\}$ as an example. We define two real numbers first: $\bar{h}_{s,n}^{(m)} = \operatorname{Re}\left\{h_{s,n}^{(m)}\right\}$ and $\breve{h}_{s,n}^{(m)} = \operatorname{Im}\left\{h_{s,n}^{(m)}\right\}$. Substituting $\bar{h}_{s,n}^{(m)}$ and $\breve{h}_{s,n}^{(m)}$ into (19) yields: $\operatorname{Im}\left\{\boldsymbol{\mu}_n^{(m)}\right\} = \bar{h}_{s,n}^{(m)}\operatorname{Im}\left\{\mathbf{z}_n^{(m)}\right\} - \breve{h}_{s,n}^{(m)}\operatorname{Re}\left\{\mathbf{z}_n^{(m)}\right\} - \bar{h}_{s,n}^{(m)}\operatorname{Im}\left\{z_{n,0}^{(m)}\right\}\mathbf{1}_{(K-1)\times 1} + \breve{h}_{s,n}^{(m)}\operatorname{Re}\left\{z_{n,0}^{(m)}\right\}\mathbf{1}_{(K-1)\times 1}$. Then, $\operatorname{var}\left\{\Delta\mathbf{c}_1^{(m)}\right\}$ is derived in (37), given at the top of the next page. The term $\operatorname{var}\left\{\Delta\boldsymbol{\kappa}_1^{(m)}\right\}$ can be similarly derived. Proposition 2 is thus proved.

APPENDIX C
PROOF OF PROPOSITION 3

In (35), $\mathcal{E}_\Delta$ can be expressed by $\Delta w_{c,k}^{(m)}$ and $\Delta w_{\kappa,k}^{(m)}$, where $\Delta w_{c,k}^{(m)} = \left[\Delta\mathbf{w}_c^{(m)}\right]_k$ and $\Delta w_{\kappa,k}^{(m)} = \left[\Delta\mathbf{w}_\kappa^{(m)}\right]_k$:

$$\begin{aligned}\mathcal{E}_\Delta &\approx \frac{1}{NK}\mathcal{F}_2\left\{\left(1+\Delta w_{c,k}^{(m)}\right)\left(1+\Delta w_{\kappa,k}^{(m)}\right)^{-n} - 1\right\} \\ &\approx \frac{1}{NK}\mathcal{F}_2\left\{\Delta w_{c,k}^{(m)} - n\Delta w_{\kappa,k}^{(m)}\right\} \\ &= \frac{1}{NK}\left[\mathcal{F}_\tau\{1\}\mathcal{F}_{f_D}\left\{\Delta w_{c,k}^{(m)}\right\} - \mathcal{F}_\tau\{n\}\mathcal{F}_{f_D}\left\{\Delta w_{\kappa,k}^{(m)}\right\}\right] \\ &\approx \frac{1}{K}\delta(\tau)\left[\mathcal{F}_{f_D}\left\{\Delta w_{c,k}^{(m)}\right\} - \frac{N-1}{2}\mathcal{F}_{f_D}\left\{\Delta w_{\kappa,k}^{(m)}\right\}\right], \end{aligned} \quad (38)$$

where $\delta(0) = 1$ and $\delta(\tau) = 0$, $\tau \neq 0$, and $\mathcal{F}_\tau\{1\} = N\delta(\tau)$ and $\mathcal{F}_\tau\{n\} = \sum_{n=0}^{N-1} n e^{j2\pi n\tau\Delta f} \approx \frac{(N-1)N}{2}\delta(\tau)$. Substituting (35) into (38) yields:

$$\begin{aligned}&\operatorname{var}\{\mathcal{E}_\Delta\} \\ &= \frac{\delta(\tau)}{K}\sigma_n^2\left[\frac{1}{\sum\limits_{\tilde{n}=0}^{N-1}\left|h_{s,n}^{(m)}\right|^2} + \left(\frac{N-1}{2}\right)^2\frac{1}{\sum\limits_{\tilde{n}=0}^{N-1}n^2\left|h_{s,n}^{(m)}\right|^2}\right] \\ &\approx \frac{\delta(\tau)}{K}\frac{\sigma_n^2}{\sum\limits_{\tilde{n}=0}^{N-1}\left|h_{s,n}^{(m)}\right|^2}\left(1+\frac{3}{4}\right) = \frac{1}{NK}\frac{7\sigma_n^2}{4\mathcal{P}_s}\delta(\tau), \end{aligned} \quad (39)$$

where $\mathcal{P}_s = \frac{1}{N}\sum_{n=0}^{N-1}\left|h_{s,n}^{(m)}\right|^2$ denotes the static object power, and $\operatorname{var}\left\{\mathcal{F}_{f_D}\left\{z_{n,k}^{(m)}\right\}\right\} = K\sigma_n^2$.

Next, we substitute (39) into (21) and use the standard deviation to express $\mathcal{E}_\Delta$. Then, $\mathcal{E}_\Delta \circledast \mathcal{F}_2\left\{\tilde{y}_{s,n,k}^{(m)}\right\}$ is given by:

$$\begin{aligned}\mathcal{E}_\Delta \circledast \mathcal{F}_2\left\{\tilde{y}_{s,n,k}^{(m)}\right\} &= \sqrt{\frac{1}{NK}\frac{7\sigma_n^2}{4\mathcal{P}_s}}\delta(\tau) \circledast \left[NKY_s(\tau)\delta(f_D)\right] \\ &= \sqrt{\frac{1}{NK}\frac{7\sigma_n^2}{4\mathcal{P}_s}}NKY_s(\tau), \end{aligned} \quad (40)$$

where $\delta(f_D) = 1$ and $\delta(f_D) = 0$, $f_D \neq 0$, and $\mathcal{F}_2\left\{\tilde{y}_{s,n,k}^{(m)}\right\} = NKY_s(\tau)\delta(f_D)$. Similar to (40), the noise term is given by $\mathcal{F}_{noise} = \mathcal{F}_2\left\{z_{n,k}^{(m)}\right\} = \sqrt{NK}\sigma_n$. Finally, we compute the ratio between $\mathcal{E}_\Delta \circledast \mathcal{F}_2\left\{\tilde{y}_{s,n,k}^{(m)}\right\}$ and $\mathcal{F}_{noise}$ to prove Proposition 3:

$$\frac{\mathcal{E}_\Delta \circledast \mathcal{F}_2\left\{\tilde{y}_{s,n,k}^{(m)}\right\}}{\mathcal{F}_{noise}} = \sqrt{\frac{7}{4}}\frac{Y_s(\tau)}{\sqrt{\mathcal{P}_s}}. \quad (41)$$

Note that $\mathcal{E}_\Delta \circledast \mathcal{F}_2\left\{\tilde{y}_{s,n,k}^{(m)}\right\}$ is much lower than $\mathcal{F}_{noise}$ when the static power is sufficiently large. Proposition 3 is thus proved.

$$\mathrm{var}\left\{\Delta\mathbf{c}_1^{(m)}\right\} = \frac{\sigma_n^2}{2}\left(\sum_{n=0}^{N-1}\left|h_{s,n}^{(m)}\right|^2\right)^{-2}\mathbf{1}_{(K-1)\times 1}\sum_{n=0}^{N-1}\left[2\left(\bar{h}_{s,n}^{(m)}\right)^2+2\left(\breve{h}_{s,n}^{(m)}\right)^2\right] = \sigma_n^2\left(\sum_{n=0}^{N-1}\left|h_{s,n}^{(m)}\right|^2\right)^{-1}\mathbf{1}_{(K-1)\times 1}. \quad (37)$$

REFERENCES

[1] 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 Commun. Surveys Tuts.*, vol. 24, no. 1, pp. 306–345, 2022.

[2] S. Li, Z. Liu, Y. Zhang, Q. Lv, X. Niu, L. Wang, and D. Zhang, "Wiborder: Precise Wi-Fi based boundary sensing via through-wall discrimination," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 4, no. 3, sep 2020. [Online]. Available: https://doi.org/10.1145/3411834

[3] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, 2022.

[4] 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. [Online]. Available: https://doi.org/10.1145/3351279

[5] C. Li, N. Raymondi, B. Xia, and A. Sabharwal, "Outer bounds for a joint communicating radar (comm-radar): The uplink case," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 1197–1213, 2022.

[6] C. R. Berger, B. Demissie, J. Heckenbach, P. Willett, and S. Zhou, "Signal processing for passive radar using OFDM waveforms," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 1, pp. 226–238, 2010.

[7] Z. Abu-Shaban, X. Zhou, T. Abhayapala, G. Seco-Granados, and H. Wymeersch, "Performance of location and orientation estimation in 5G mmwave systems: Uplink vs downlink," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2018, pp. 1–6.

[8] 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 Commun. Mag.*, pp. 1–7, 2022.

[9] W. Yuan, N. Wu, B. Etzlinger, Y. Li, C. Yan, and L. Hanzo, "Expectation-maximization-based passive localization relying on asynchronous receivers: Centralized versus distributed implementations," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 668–681, 2019.

[10] G. Wen, H. Song, F. Gao, Y. Liang, and L. Zhu, "Target localization in asynchronous distributed MIMO radar systems with a cooperative target," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4098–4113, 2022.

[11] K. Qian, C. Wu, Y. Zhang, G. Zhang, and Y. Liu, "Widar2.0: Passive human tracking with a single Wi-Fi link," *Proc. 16th Annu. Int. Conf. Mobile Syst., Appl. Serv.*, pp. 350–361, 2018. [Online]. Available: https://doi.org/10.1145/3210240.3210314

[12] 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 Process.*, vol. 69, pp. 1287–1300, 2021.

[13] F. Meneghello, D. Garlisi, N. D. Fabbro, I. Tinnirello, and M. Rossi, "SHARP: Environment and person independent activity recognition with commodity IEEE 802.11 access points," *IEEE Trans. on Mobile Comput.*, vol. 22, no. 10, pp. 6160–6175, 2023.

[14] 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. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 3, no. 3, 2019.

[15] D. Wu, Y. Zeng, R. Gao, S. Li, Y. Li, R. C. Shah, H. Lu, and D. Zhang, "WiTraj: Robust indoor motion tracking with WiFi signals," *IEEE Trans. Mobile Comput.*, vol. 22, no. 5, pp. 3062–3078, 2023.

[16] X. Chen, Z. Feng, J. A. Zhang, X. Yuan, and P. Zhang, "Kalman filter-based sensing in communication systems with clock asynchronism," *IEEE Trans. Commun.*, vol. 72, no. 1, pp. 403–417, 2024.

[17] R. Mendrzik, H. Wymeersch, G. Bauch, and Z. Abu-Shaban, "Harnessing NLOS components for position and orientation estimation in 5G millimeter wave MIMO," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 93–107, 2019.

[18] C. E. OLone, H. S. Dhillon, and R. M. Buehrer, "Single-anchor localizability in 5G millimeter wave networks," *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 65–69, 2020.

[19] L. Yang, N. Wu, B. Li, W. Yuan, and L. Hanzo, "Indoor localization based on factor graphs: A unified framework," *IEEE Internet Things J.*, vol. 10, no. 5, pp. 4353–4366, 2023.

[20] C. Xu, Y. Shi, J. Wan, and S. Duan, "Uncertainty-constrained belief propagation for cooperative target tracking," *IEEE Internet Things J.*, vol. 9, no. 19, pp. 19 414–19 425, 2022.

[21] K.-F. Ssu, C.-H. Ou, and H. Jiau, "Localization with mobile anchor points in wireless sensor networks," *IEEE Trans. Veh. Technol.*, vol. 54, no. 3, pp. 1187–1197, 2005.

[22] M. Kolakowski, V. Djaja-Josko, and J. Kolakowski, "Static LiDAR assisted UWB anchor nodes localization," *IEEE Sensors J.*, vol. 22, no. 6, pp. 5344–5351, 2022.

[23] M. I. Skolnik, *Radar Handbook, Third Edition*. New York, USA: The McGraw-Hill Companies, 2008.

[24] B. Tan, K. Woodbridge, and K. Chetty, "A real-time high resolution passive WiFi doppler-radar and its applications," in *Proc. Int. Radar Conf.*, 2014, pp. 1–6.

[25] H. Sun, L. G. Chia, and S. G. Razul, "Through-wall human sensing with WiFi passive radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 57, no. 4, pp. 2135–2148, 2021.

[26] J. Liu, K. Wu, T. Su, and J. A. Zhang, "Practical frequency-hopping MIMO joint radar communications: Design and experiment," 2023.

[27] Y. He, Y. Cai, H. Mao, and G. Yu, "RIS-assisted communication radar coexistence: Joint beamforming design and analysis," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2131–2145, 2022.

[28] 3GPP TS 38.211, Physical channels and modulation, V15.2.0, January 2023.

[29] X. Lin, J. Li, R. Baldemair, J.-F. T. Cheng, S. Parkvall, D. C. Larsson, H. Koorapaty, M. Frenne, S. Falahati, A. Grovlen, and K. Werner, "5G New Radio: Unveiling the essentials of the next generation wireless access technology," *IEEE Commun. Stand. Mag.*, vol. 3, no. 3, pp. 30–37, 2019.

[30] H. Yin, N. Li, J. Guo, J. Zhu, and X. She, "NR coverage enhancements for PUSCH," *IEEE Commun. Mag.*, vol. 60, no. 7, pp. 36–42, 2022.

[31] D. Rife and R. Boorstyn, "Single tone parameter estimation from discrete-time observations," *IEEE Trans. Inf. Theory*, vol. 20, no. 5, pp. 591–598, 1974.

[32] S. M. Kay, *Fundamentals of Statistical Signal Processing Volume I: Estimation Theory*, (Prentice-Hall Signal Processing Series). Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

[33] L. Zheng and X. Wang, "Super-resolution delay-Doppler estimation for OFDM passive radar," *IEEE Trans. Signal Process.*, vol. 65, no. 9, pp. 2197–2210, 2017.

[34] S. M. Patole, M. Torlak, D. Wang, and M. Ali, "Automotive radars: A review of signal processing techniques," *IEEE Signal Process. Mag.*, vol. 34, no. 2, pp. 22–35, 2017.

[35] 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.