Fingerprint-Spectrum-Based Synchronization in Asynchronous Perceptive Mobile Networks

Xiao-Yang Wang[1,2], Shaoshi Yang[1], Mingyang Chen[2], Christos Masouros[2]
[1]School of Information and Communication Engineering, Beijing University of Posts and Telecommnuications, Beijing, China
[2]Department of Electronic and Electrical Engineering, University College London, London, UK
E-mails: wangxy_028@bupt.edu.cn, shaoshi.yang@bupt.edu.cn, uceemc4@ucl.ac.uk, c.masouros@ucl.ac.uk

***Abstract*—Asynchronous radio transceivers introduce significant range and velocity ambiguity, posing challenges for accurate positioning and velocity estimation in passive sensing perceptive mobile networks (PMNs). To address this issue, we propose the Cross-Multipath Cross-Correlation (CMCC) algorithm, a joint carrier frequency offset (CFO) and time offset (TO) estimation algorithm that supports passive sensing in both non-line-of-sight (NLOS) and line-of-sight (LOS) environments. The CMCC algorithm treats the delay-Doppler spectrum of signals reflected from static objects as an environment-specific "fingerprint spectrum" that exhibits a cyclic shift property as the CFO and/or TO change. By analyzing the number of cyclic shifts, CMCC efficiently estimates both CFO and TO. Additionally, we provide a simplified version of CMCC to reduce computational complexity. Simulation results demonstrate the performance advantages of our algorithms across various configurations.**

***Index Terms*—Integrated sensing and communications (ISAC), synchronization, time offset (TO), carrier frequency offset (CFO).**

I. Introduction

With the assistance of millimeter wave and terahertz frequencies, high-precision sensing [1], including range and velocity estimation, as well as imaging, can be performed in mobile networks [2], giving rise to the concept of perceptive mobile networks (PMNs). PMNs can be categorized into two distinct types: passive and active sensing [3]. In passive sensing, the remote radio unit (RRU) or user equipment (UE) estimates desired parameters by exploiting echo signals transmitted by other RRUs or UEs. In active sensing, the RRU performs sensing by processing signals transmitted by itself [4], [5]. Nevertheless, achieving high-performance active sensing in PMNs requires practical full-duplex technology, which is not yet sufficiently developed for implementation, at the time of writing [5]. Moreover, sensing in time-division duplex (TDD) or frequency-division duplex (FDD) modes, as described in [5], constrains the achievable communication rates [6]. Consequently, passive sensing emerges as a competitive alternative.

However, significant challenges remain for realizing passive sensing. Specifically, the geographically separated radio transceivers in passive sensing PMNs are naturally asynchronous [7], which causes non-negligible velocity and range sensing ambiguities [8]. To mitigate these ambiguities, both carrier frequency offset (CFO) and time offset (TO) should be compensated by designing synchronization schemes.

There has been a scarcity of contributions reported in terms of synchronization in PMNs [6]. Among them, the first one is based on cross-antenna cross-correlation (CACC) [9], where cross-correlation between signals received by different antennas is exploited to eliminate CFO and TO. However, the cross-correlation doubles the unknown parameters to be estimated, which results in high computational complexity. For reducing the computational complexity, the authors of [8] proposed the mirrored-CACC algorithm, which reconstructs the output of CACC and reduces the number of unknown parameters by half. However, both of the two algorithms are only applicable in LOS scenario. Another two contributions, namely FarSense [10] and MultiSense [11], applied the channel state information (CSI)-ratio-based CFO and TO synchronization algorithms to the non-line-of-sight (NLOS) scenario. Unfortunately, these can only synchronize the phase offset between transceivers, thus cannot estimate the delay and consequently cannot perform ranging for targets. Moreover, in [12], a family of CSI-ratio-based frequency synchronization schemes were proposed, which can only mitigate CFO. However, they are only effective in scenarios where the velocities of the targets are low and a single dynamic path exists.

To address the limitations of the state-of-the-art solutions, we propose the first viable joint CFO and TO estimation algorithm for asynchronous passive sensing in PMNs. Our algorithm is applicable to both single-antenna and multi-antenna transceivers operating in NLOS environments. In our approach, we exploit the presence of permanent or long-period static objects in the environment. Specifically, we leverage the delay-Doppler spectrum of signals reflected from these static objects, treating it as an environment-specific *fingerprint spectrum*. This spectrum exhibits a cyclic shift property in response to changes of CFO and TO. By performing cross-

This work was supported in part by the Beijing Municipal Natural Science Foundation (No. Z220004), in part by the Beijing Municipal Science & Technology Commission (No. Z221100007722036), in part by the National Key R&D Program of China (No. 2023YFB2904803), in part by the Guangdong Major Project of Basic and Applied Basic Research (No. 2023B0303000001), in part by the Fundamental Research Funds for the Central Universities (No. 2023ZCJH02), and in part by BUPT Excellent Ph.D. Students Foundation (No. CX2023238). *Corresponding author: Shaoshi Yang*.

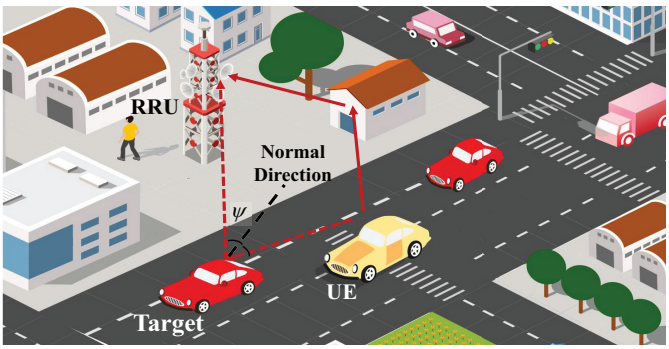


Fig. 1. System model of an asynchronous PMN.

correlation between fingerprint spectra captured at different time instants, our algorithm accurately estimates both CFO and TO. This paper constitutes a concise introduction of the synchronization algorithm module within the framework presented in our previous work [3].

Notations: $\mathbf{A}^{\mathrm{T}}$, $\mathbf{A}^{\mathrm{H}}$ and $\mathbf{A}^{-1}$ represent transpose, conjugate transpose and inverse of $\mathbf{A}$, respectively. $\mathrm{Re}(\mathbf{A})$, $\mathbf{A}[i,:]$, and $\mathbf{A}[i,j]$ stands for the real part, the ith row, and the (i,j)th element of $\mathbf{A}$, respectively. $a(t)$, $\mathbf{a}(t)$, and $\mathbf{A}(t)$ are the scalar function, the vector function, and the matrix function with respect to the scalar t, respectively. Moreover, $\mathrm{diag}(\mathbf{a}_1,\cdots,\mathbf{a}_n)$ is a block diagonal matrix whose diagonal blocks are $\{\mathbf{a}_1,\cdots,\mathbf{a}_n\}$. $\mathrm{Round}(\cdot)$ is the operator that takes the integer closest to a real number. $\lfloor\cdot\rfloor$ and $|\cdot|$ denote the floor function and the modulo operator, respectively. $\mathbf{I}_N$ is the N dimensional identity matrix.

II. System Model

We consider an asynchronous PMN relying on the multiple-input multiple-output orthogonal frequency division multiplexing (MIMO-OFDM) technique. As shown in Fig. 1, we assume that there are L propagation paths between the UE and the RRU, with L_{V} targets among these L paths. Moreover, the RRU is equipped with an M_{R}-element uniform linear array (ULA), while the UE is equipped with an M_{U}-element ULA. Particularly, the signals propagated via the L paths can be divided into two categories: the signals reflected by the targets and the signals reflected by other static objects in the environment, namely the clutter.

Let us consider the signal propagated from a particular path, e.g., the lth path, which undergoes reflection by the l_{V}th target. v_l represents the velocity projected onto the normal direction associated with the lth path. Moreover, we denote the TO and CFO caused by the separate oscillators of the RRU and the UE as τ_{o} and f_{o}, respectively.

As a result, the time-domain channel impulse response between the and the RRU is expressed as

$$\mathbf{H}(t)=\sum_{l=1}^{L} h_l\delta(t-\frac{2v_l}{c}t-\tau_{\mathrm{d},l}-\tau_{\mathrm{o}})\mathbf{a}^{\mathrm{T}}(M_{\mathrm{R}},\phi_l)\mathbf{a}(M_{\mathrm{U}},\theta_l), \tag{1}$$

where $\mathbf{a}(M_{\mathrm{R}},\phi_l) = [1,\cdots,e^{-\frac{j2\pi(M_{\mathrm{R}}-1)d}{\lambda}\sin\phi_l}]$ and $\mathbf{a}(M_{\mathrm{U}},\theta_l)=[1,\cdots,e^{-\frac{j2\pi(M_{\mathrm{U}}-1)d}{\lambda}\sin\theta_l}]$ are the receiving and transmitting steering vectors of the lth path, respectively. Here, d, λ, c and $\delta(\cdot)$ denote the antenna-element spacing, the signal wavelength, the speed of light and the Dirac delta function, respectively. In addition, h_l, ϕ_l, θ_l, and $\tau_{\mathrm{d},l}$ represent the channel gain, the direction of arrival (DOA), the angle-of-departure (AOD), and the time delay of the signal travelling along the lth path, respectively. Moreover, f_{c} is defined as the carrier frequency.

Furthermore, let us denote the OFDM symbol length as $T_{\mathrm{sym}}=N_{\mathrm{s}}T_{\mathrm{s}}$, where T_{s} and N_{s} represent the sampling period and the number of samples in each OFDM symbol. Moreover, the number of samples in the cyclic prefix (CP) is denoted by N_{cp}, and we define $N_{\mathrm{sub}}=N_{\mathrm{s}}-N_{\mathrm{cp}}$, which represents the number of subcarriers. Thus, the gth transmitted OFDM symbol of the UE upon excluding the CP is formulated as

$$\mathbf{p}_g(t)=e^{j2\pi f_{\mathrm{c}}t}\mathbf{w}^{\mathrm{T}}\sum_{u=0}^{N_{\mathrm{sub}}-1}\mathbf{x}_g[u]e^{j2\pi u\Delta f t}, \tag{2}$$

where $\mathbf{x}_g=[\mathbf{x}_g[0],\cdots,\mathbf{x}_g[N_{\mathrm{sub}}-1]]$ and $\mathbf{x}_g[u]$ represent the data symbol modulated on the uth subcarrier of the gth OFDM symbol, Δf is the subcarrier spacing. Moreover, $\mathbf{w}\in\mathbb{C}^{1\times M_{\mathrm{U}}}$ is the precoding vector.

Without loss of generality and for achieving high sensing performance with low overhead, in this paper, we assume that the data payload is used for sensing the target. Note that our following discussions are also applicable to the scenarios designating demodulation reference signals (DMRS) or synchronization signal blocks (SSB) as the sensing pilots. Then, since CP converts linear convolution to circulant convolution, the received analog signal excluding CP is $\mathbf{y}_g(t)=\mathbf{H}(t)\circledast\mathbf{p}_g(t)+\mathbf{z}_g^{\mathrm{T}}(t)$, which is further expressed as

$$\begin{aligned}\mathbf{y}_g(t)=&\sum_{l=1}^{L}\sum_{u=0}^{N_{\mathrm{sub}}-1}h_l e^{j2\pi f_{\mathrm{c}}\bar{t}}\mathbf{a}^{\mathrm{T}}(M_{\mathrm{R}},\phi_l)\mathbf{a}(M_{\mathrm{U}},\theta_l)\mathbf{w}^{\mathrm{T}}\\&\cdot\mathbf{x}_g[u]e^{j2\pi u\Delta f\bar{t}}+\mathbf{z}_g^{\mathrm{T}}(t),\end{aligned} \tag{3}$$

where "$\circledast$" denotes the circular convolution operator and $\bar{t}\triangleq t-\frac{2v_l}{c}t-\tau_{\mathrm{d},l}-\tau_{\mathrm{o}}$, and $\mathbf{z}_g(t)\in\mathbb{C}^{1\times M_{\mathrm{R}}}$ is the zero-mean complex-valued additive white Gaussian noise (AWGN) vector function satisfying $\mathrm{E}[\mathbf{z}_g^{\mathrm{H}}(t)\mathbf{z}_g(t)]=\sigma_0^2\mathbf{I}_{M_{\mathrm{R}}}$, respectively.

Next, $\mathbf{y}_g(t)\in\mathbb{C}^{M_{\mathrm{R}}\times 1}$ is demodulated to a baseband signal by multiplying it with the local carrier, which typically exhibits a frequency offset f_{o} relative to the transmitter, thus yielding $\bar{\mathbf{y}}_g(t)=\mathbf{y}_g(t)e^{-j2\pi(f_{\mathrm{c}}+f_{\mathrm{o}})t}$. Then, $\bar{\mathbf{y}}_g(t)$ is sampled at time instant nT_{s}. As a result, the analog signal $\bar{\mathbf{y}}_g(t)$ is transformed into a digital signal vector $\bar{\mathbf{y}}_g(nT_{\mathrm{s}}), n=1,\cdots,N_{\mathrm{sub}}$. For brevity, in what follows $\bar{\mathbf{y}}_g(nT_{\mathrm{s}})$ is rewritten as $\bar{\mathbf{y}}_g(n)$ and we approximately formulate it as

$$\begin{aligned}\bar{\mathbf{y}}_g(n)\approx&\sum_{l=1}^{L}\sum_{u=0}^{N_{\mathrm{sub}}-1}h_l\, e^{-j2\pi f_{\mathrm{c}}(\tau_{\mathrm{d},l}+\tau_{\mathrm{o}})}\, e^{-j2\pi(\xi_{\mathrm{D},l}+\,\xi_{\mathrm{o}})\frac{N_{\mathrm{cp}}}{N_{\mathrm{sub}}}}\\&\cdot e^{-j2\pi(\xi_{\mathrm{D},l}+\xi_{\mathrm{o}})(g-1)\frac{N_{\mathrm{s}}}{N_{\mathrm{sub}}}}\mathbf{a}^{\mathrm{T}}(M_{\mathrm{R}},\phi_l)\mathbf{a}(M_{\mathrm{U}},\theta_l)\mathbf{w}^{\mathrm{T}}\mathbf{x}_g[u]\\&\cdot e^{-j2\pi u\Delta f(\tau_{\mathrm{d},l}+\tau_{\mathrm{o}})}e^{j2\pi u\Delta f nT_{\mathrm{s}}}+\mathbf{z}_g^{\mathrm{T}}(n),\end{aligned} \tag{4}$$

where $\mathbf{z}_g^{\mathrm{T}}(n)$ is the complex-valued AWGN vector recorded at the nth time instant of sampling for the gth transmitted OFDM symbol. Additionally, $\xi_{\mathrm{o}}=N_{\mathrm{sub}}f_{\mathrm{o}}T_{\mathrm{s}}$ and $\xi_{\mathrm{D},l}=$

$2f_{\mathrm{c}}v_l N_{\mathrm{sub}}T_{\mathrm{s}}/c$ are the CFO and the Doppler offset normalized by the subcarrier spacing, respectively. Furthermore, we construct $\mathbf{Y}_g=[\bar{\mathbf{y}}_g(1),\cdots,\bar{\mathbf{y}}_g(N_{\mathrm{sub}})]$ as the OFDM samples received by all the M_{R} antennas of the RRU in the duration of a single OFDM symbol, and formulate it as

$$\mathbf{Y}_g=\sum\nolimits_{l=1}^{L} h_l e^{-j2\pi(\xi_{\mathrm{D},l}+\xi_{\mathrm{o}})\frac{N_{\mathrm{cp}}}{N_{\mathrm{sub}}}} e^{-j2\pi(\xi_{\mathrm{D},l}+\xi_{\mathrm{o}})(g-1)\frac{N_{\mathrm{s}}}{N_{\mathrm{sub}}}} e^{-j2\pi f_{\mathrm{c}}(\tau_{\mathrm{d},l}+\tau_{\mathrm{o}})} \cdot \mathbf{a}^{\mathrm{T}}(M_{\mathrm{R}},\phi_l)\mathbf{a}(M_{\mathrm{U}},\theta_l)\mathbf{w}^{\mathrm{T}}\boldsymbol{\tau}_l\mathbf{D}(\mathbf{x}_g)\mathbf{F}+\mathbf{Z}_g, \tag{5}$$

where $\mathbf{F}$, $\mathbf{D}(\mathbf{x}_g)$, $\boldsymbol{\tau}_l$ and $\mathbf{Z}_g$ represent the inverse discrete Fourier transform (IDFT) matrix, $\mathrm{diag}(\mathbf{x}_g)$, $[1,\cdots,e^{-j2\pi(N_{\mathrm{sub}}-1)\Delta f(\tau_{\mathrm{d},l}+\tau_{\mathrm{o}})}]$ and $[\mathbf{z}_g^{\mathrm{T}}(1),\cdots,\mathbf{z}_g^{\mathrm{T}}(N_{\mathrm{sub}})]$, respectively.

III. The Fingerprint-Spectrum-Based Synchronization

Before implementing synchronization, we firstly compensate $\mathbf{Y}_g$ by $\mathbf{Y}_g\mathbf{F}^{-1}\mathbf{D}^{-1}(\mathbf{x}_g)$, as depicted in [13]. Then, let us define $\breve{\mathbf{y}}_{g,m}$ as the mth row of $\mathbf{Y}_g\mathbf{F}^{-1}\mathbf{D}^{-1}(\mathbf{x}_g)$, which is the gth compensated OFDM symbol received by the mth antenna, and stack $\breve{\mathbf{y}}_{g,m}$ as $\boldsymbol{\Gamma}_m=[\breve{\mathbf{y}}_{1,m}^{\mathrm{T}},\cdots,\breve{\mathbf{y}}_{G,m}^{\mathrm{T}}]^{\mathrm{T}}$. Then, $\boldsymbol{\Gamma}_m$ is formulated as

$$\boldsymbol{\Gamma}_m=\sum\nolimits_{l=1}^{L}\boldsymbol{\alpha}_l[m]\boldsymbol{\zeta}_l\boldsymbol{\tau}_l+\bar{\mathbf{Z}}_g, \tag{6}$$

where $\boldsymbol{\alpha}_l = h_l e^{-j2\pi f_{\mathrm{c}}(\tau_{\mathrm{d},l}+\tau_{\mathrm{o}})}\mathbf{a}^{\mathrm{T}}(M_{\mathrm{R}},\phi_l)\mathbf{a}(M_{\mathrm{U}},\theta_l)\mathbf{w}^{\mathrm{T}}$, and $\boldsymbol{\zeta}_l=[e^{-j2\pi\xi_l\frac{N_{\mathrm{cp}}}{N_{\mathrm{sub}}}},\cdots,e^{-j2\pi\xi_l(G-1)\frac{N_{\mathrm{s}}}{N_{\mathrm{sub}}}}e^{-j2\pi\xi_l\frac{N_{\mathrm{cp}}}{N_{\mathrm{sub}}}}]^{\mathrm{T}}$, where $\xi_l=\xi_{\mathrm{D},l}+\xi_{\mathrm{o}}$. Moreover, $\bar{\mathbf{Z}}=[\breve{\mathbf{z}}_{1,m}^{\mathrm{T}},\cdots,\breve{\mathbf{z}}_{G,m}^{\mathrm{T}}]^{\mathrm{T}}$, where $\breve{\mathbf{z}}_{g,m}$ is the mth row of $\hat{\mathbf{Z}}_g$.

Then, we implement two dimensional discrete Fourier transform (2D-DFT) to $\mathrm{Re}(\boldsymbol{\Gamma}_m)$ to reduce the complexity. Noting that $\mathrm{Re}(\boldsymbol{\Gamma}_m)$ has the same signal-to-noise ratio (SNR) as $\boldsymbol{\Gamma}_m$, this operation will not affect the estimation accuracy. Specifically, by defining $\tilde{\mathbf{Z}}=\mathbf{F}_G^{\mathrm{H}}\,\mathrm{Re}(\bar{\mathbf{Z}})\mathbf{F}_{N_{\mathrm{sub}}}^{\mathrm{H}}$, the 2D-DFT of $\mathrm{Re}(\boldsymbol{\Gamma}_m)$, namely $\bar{\mathbf{Y}}$, can be represented as

$$\bar{\mathbf{Y}}=\sum\nolimits_{l=1}^{L}\boldsymbol{\alpha}_l[m]\mathbf{F}_G^{\mathrm{H}}\boldsymbol{\zeta}_l(\mathbf{F}_{N_{\mathrm{sub}}}^{\mathrm{H}}\boldsymbol{\tau}_l^{\mathrm{T}})^{\mathrm{T}}+\tilde{\mathbf{Z}}. \tag{7}$$

According to [14], we obtain the final expression for (k,n)th element of $\bar{\mathbf{Y}}$ as

$$\bar{\mathbf{Y}}[k,n]=\frac{1}{4}\sum\nolimits_{l=1}^{L}\Big\{\tilde{\boldsymbol{\alpha}}_l[m]\ \mathrm{S}_G((k-\xi_l\frac{N_{\mathrm{s}}}{N_{\mathrm{sub}}})f_{\mathrm{R}})\mathrm{S}_{N_{\mathrm{sub}}}((n-\frac{\tau_l}{T_{\mathrm{s}}})T_{\mathrm{R}}) + \tilde{\boldsymbol{\alpha}}_l^{\mathrm{H}}[m]\ \mathrm{S}_G((k+\ \xi_l\frac{N_{\mathrm{s}}}{N_{\mathrm{sub}}})f_{\mathrm{R}})\ \mathrm{S}_{N_{\mathrm{sub}}}((n+\frac{\tau_l}{T_{\mathrm{s}}})T_{\mathrm{R}})\Big\}+\ \tilde{\mathbf{Z}}[k,n], \tag{8}$$

where $\tau_l=\tau_{\mathrm{d},l}+\tau_{\mathrm{o}}$ and $\mathrm{S}_G(\cdot)$ is an aliased sinc function [14]. Moreover, f_{R} is the frequency resolution, which can be expressed as the reciprocal of the total sensing time, $1/(GT_{\mathrm{sym}})$ [13], while T_{R} is the time resolution and can be formulated as the reciprocal of the total sensing bandwidth, $1/(N_{\mathrm{sub}}\Delta f)$ [13]. In light of the high-attenuation side lobes of the aliased sinc function, $\bar{\mathbf{Y}}[k,n]$ can be approximately formulated as

$$\bar{\mathbf{Y}}[k,n]\approx\frac{1}{4}\sum_{l=1}^{L}\tilde{\boldsymbol{\alpha}}_l[m]\mathrm{S}_G((k-\xi_l\frac{N_{\mathrm{s}}}{N_{\mathrm{sub}}})f_{\mathrm{R}})\mathrm{S}_{N_{\mathrm{sub}}}[(n-\frac{\tau_l}{T_{\mathrm{s}}})T_{\mathrm{R}}]+\tilde{\mathbf{Z}}[k,n]. \tag{9}$$

For brevity, we here take the scenario with low-velocity UE as an example. The clutter channel response will exhibit a near-zero Doppler frequency shift and the part of spectrum corresponding to clutter will be centred around $\xi_{\mathrm{o}}N_{\mathrm{s}}Kf_{\mathrm{R}}/N_{\mathrm{sub}}$ in the DFT of $\mathrm{Re}(\boldsymbol{\Gamma}_m)$. Then, we can select the $K_{\mathrm{up}}=\mathrm{Round}(\xi_{\mathrm{o}}N_{\mathrm{s}}K/N_{\mathrm{sub}})$th row of $\bar{\mathbf{Y}}$ to be the "identification sequence". For brevity, we will refer to the sequence as the fingerprint spectrum in the following. Moreover, if the velocity of UE is large, even though the clutter channel response does not exhibit a near-zero Doppler frequency shift, K_{up} can still be determined with the known velocity. However, this is not the focus of this paper.

To investigate how the fingerprint spectrum evolves in response to the changes in CFO and TO, we assume that the CFO and TO increase by $\Delta\xi_{\mathrm{o}}$ and $\Delta\tau_{\mathrm{o}}$, respectively. Then, the nth element of the updated fingerprint spectrum $\bar{\mathbf{Y}}_{\mathrm{up}}[K_{\mathrm{up}},n]$ is formulated as

$$\begin{aligned}&\bar{\mathbf{Y}}_{\mathrm{up}}[K_{\mathrm{up}},n]\approx\\&\frac{1}{4}\ \mathrm{S}_G([K_{\mathrm{up}}-\xi_{\mathrm{o}}\frac{N_{\mathrm{s}}K}{N_{\mathrm{sub}}}-\mathrm{Int}(\Delta\xi_{\mathrm{o}}\frac{N_{\mathrm{s}}K}{N_{\mathrm{sub}}})-\mathrm{Frac}(\Delta\xi_{\mathrm{o}}\frac{N_{\mathrm{s}}K}{N_{\mathrm{sub}}})]f_{\mathrm{R}})\sum_{l=1}^{L}\\&\tilde{\boldsymbol{\alpha}}_l[m]\mathrm{S}_{N_{\mathrm{sub}}}([n-\frac{(\tau_{\mathrm{o}}+\tau_{\mathrm{d},l})K}{T_{\mathrm{s}}}-\mathrm{Int}(\frac{\Delta\tau_{\mathrm{o}}K}{T_{\mathrm{s}}})-\mathrm{Frac}(\frac{\Delta\tau_{\mathrm{o}}K}{T_{\mathrm{s}}})]T_{\mathrm{R}})\\&+\tilde{\mathbf{Z}}[K_{\mathrm{up}},n]\approx\bar{\mathbf{Y}}[K_{\mathrm{up}}+\mathrm{Round}(\Delta\xi_{\mathrm{o}}\frac{N_{\mathrm{s}}K}{N_{\mathrm{sub}}}),n+\mathrm{Round}(\frac{\Delta\tau_{\mathrm{o}}K}{T_{\mathrm{s}}})],\end{aligned} \tag{10}$$

where $\mathrm{Int}(\cdot)$ and $\mathrm{Frac}(\cdot)$ represent the integer and fractional part of a complex value, respectively.

So far, we have drawn the conclusion that the fingerprint spectrum approximately exhibits a *cyclic shift property in response to the changes of CFO and TO*. Specifically, if CFO and TO increase by $\Delta\xi_{\mathrm{o}}$ and $\Delta\tau_{\mathrm{o}}$, respectively, the fingerprint spectrum will cyclic shift $\mathrm{Round}(\Delta\xi_{\mathrm{o}}N_{\mathrm{s}}K/N_{\mathrm{sub}})$ and $\mathrm{Round}(\Delta\tau_{\mathrm{o}}K/T_{\mathrm{s}})$ respectively, in the direction of increasing the Doppler frequency shift and the delay.

By exploiting the change pattern between the fingerprint spectrum and the CFO/TO, the CFO/TO estimation algorithm is proposed. To begin with, we denote $\bar{\mathbf{Y}}[K_{\mathrm{up}},:]$ as the *original fingerprint spectrum*, $\boldsymbol{\zeta}$, in what follows for convenience. Then, since the updated fingerprint spectrum is the cyclic shift of the original fingerprint spectrum, cross-correlation between $\boldsymbol{\zeta}$ and $\bar{\mathbf{Y}}_{\mathrm{up}}$ can be performed to estimate the CFO and TO as

$$\{\Delta\hat{\xi}_{\mathrm{o}}\frac{N_{\mathrm{s}}K}{N_{\mathrm{sub}}},\frac{\Delta\hat{\tau}_{\mathrm{o}}K}{T_{\mathrm{s}}}\}=\arg\max_{k,q}\Big|\frac{\sum_{i=1}^{Q_{\mathrm{B}}}\bar{\mathbf{Y}}_{\mathrm{up}}[k,[(q+i)\bmod\ Q_{\mathrm{B}}]]\boldsymbol{\zeta}^{\mathrm{H}}[i]}{|\boldsymbol{\zeta}|\sqrt{\sum_{i=1}^{Q_{\mathrm{B}}}|\bar{\mathbf{Y}}_{\mathrm{up}}[k,[(q+i)\bmod\ Q_{\mathrm{B}}]]|^2}}\Big|, \tag{11}$$

where we have $k=1,\cdots,K_{\mathrm{B}}$ and $q=1,\cdots,Q_{\mathrm{B}}$, with $K_{\mathrm{B}}=\lfloor KG/2\rfloor$ and $Q_{\mathrm{B}}=\lfloor KN_{\mathrm{sub}}/2\rfloor$. As a result, by conducting (11), $\Delta\tau_{\mathrm{o}}K/T_{\mathrm{s}}$ and $\Delta\xi_{\mathrm{o}}N_{\mathrm{s}}K/N_{\mathrm{sub}}$ can be estimated.

However, the computational complexity of conducting the 2-D maximum likelihood (ML) search of (11) is high. To reduce the computational complexity, a simplified version of cross-multipath cross-correlation (CMCC), called simplified CMCC (S-CMCC), is designed by decomposing the 2-D ML search into two separate 1-D ML searches. The first 1-D ML search is dedicated to CFO estimation. Specifically, we aim to estimate the CFO by locating the row of $\bar{\mathbf{Y}}_{\mathrm{up}}$, which has the closest

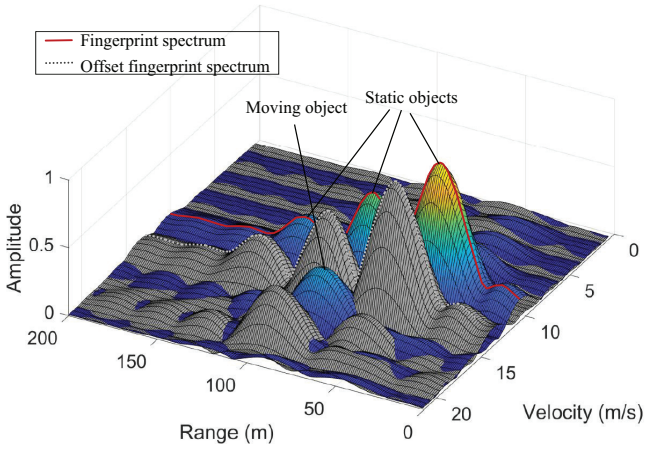


Fig. 2. An implementation of $\bar{\mathbf{Y}}$.

power to that of the original fingerprint spectrum.

$$\Delta\hat{\xi}_{\mathrm{o}} = \frac{N_{\mathrm{sub}}}{KN_{\mathrm{s}}} \arg \min_{k} \left| |\boldsymbol{\zeta}|^2 - \sum\nolimits_{n=0}^{Q_{\mathrm{B}}} |\bar{\mathbf{Y}}_{\mathrm{up}}[k, n \mod Q_{\mathrm{B}}]|^2 \right|. \tag{12}$$

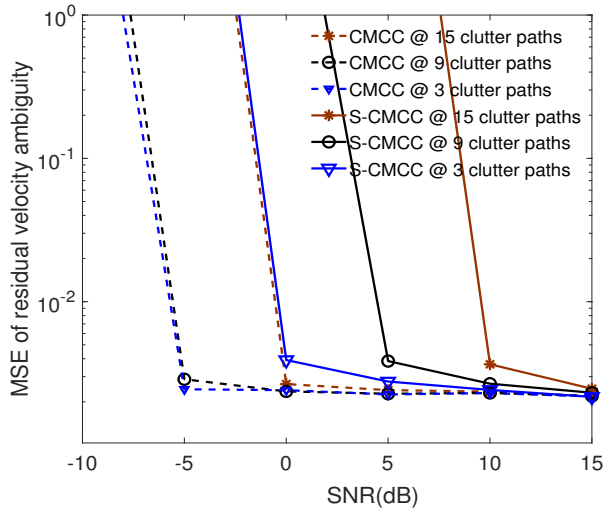
Upon completing the estimation of $\Delta\xi_{\mathrm{o}}$, another 1-D ML search can be performed to estimate $\Delta\tau_{\mathrm{o}}$. The estimate is given by

$$\Delta\hat{\tau}_{\mathrm{o}} = T_{\mathrm{s}}/K \arg\max_{q} \left| \frac{\sum_{i=1}^{Q_{\mathrm{B}}} \boldsymbol{\zeta}^{\mathrm{H}}(i) \bar{\mathbf{Y}}_{\mathrm{up}}[\Delta\hat{\xi}_{\mathrm{o}} N_{\mathrm{s}} K/N_{\mathrm{sub}}, (i+q) \bmod Q_{\mathrm{B}}]}{|\boldsymbol{\zeta}| \sqrt{\sum_{i=1}^{Q_{\mathrm{B}}} |\bar{\mathbf{Y}}_{\mathrm{up}}[\Delta\hat{\xi}_{\mathrm{o}} N_{\mathrm{s}} K/N_{\mathrm{sub}}, (i+q) \bmod Q_{\mathrm{B}}]|^2}} \right|. \tag{13}$$

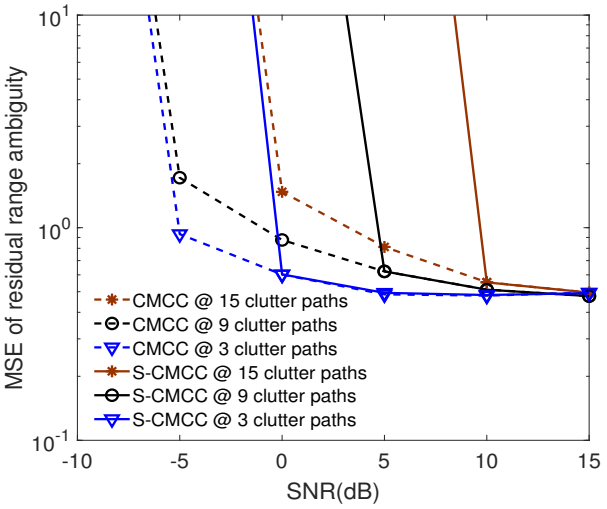
IV. Numerical Simulations

In this section, simulation results are provided to evaluate the performance of the proposed scheme. Unless otherwise stated, we assume that the number of antenna-elements on RRU and each UE is $M_{\mathrm{R}} = 64$ and $M_{\mathrm{U}} = 2$, respectively. According to [15], the transmit power of UE is set as 25 dBm. In addition, the carrier frequency f_{c} is set as 28 GHz, while N_{sub} and N_{cp} are set as 128 and 16, respectively. Furthermore, in order to ensure good sensing performance, we assume that the subcarrier spacing Δf is as large as 100 kHz. Moreover, the multipath signals are simulated by generating signals reflected from the static and the moving objects, while the static objects are randomly set around each target by adopting the cluster channel model.

In Fig. 2, we present a realization of $\bar{\mathbf{Y}}$ in an asynchronous clutter-rich PMN. To enhance sensing resolution, we apply 25 times and 5 times zero-padding for the range estimation and the velocity estimation, respectively. The blue and grey 3D spectra represent the original and the updated delay-Doppler spectrum. The CFO and TO in the simulation are equivalent to a Doppler frequency shift caused by a velocity of 3 m/s and a time delay resulting from a distance of 10 m, respectively. The red lines in Fig. 2 are selected as the fingerprint spectrum sequence, while the dashed grey lines represent the updated fingerprint spectrum sequence. As depicted in Fig. 2, the updated fingerprint spectrum is the original fingerprint spectrum



(a) SNR vs MSE of residual velocity ambiguity



(b) SNR vs MSE of residual range ambiguity

Fig. 3. MSE of residual velocity and range ambiguity vs SNR with {3, 9, 15} clutter paths and simplified/unsimplified estimation algorithm.

with cyclic shifts. Thus, we can estimate both CFO and TO by determining the number of shifts of the fingerprint spectrum.

In Fig. 3, we evaluate the synchronization performance of the proposed CMCC and S-CMCC algorithm in practical scenarios with different number of clutter path. As presented in the figure, CMCC has much better performance than S-CMCC in all scenarios. This phenomenon emerges mainly since the 1-D ML CFO estimation in S-CMCC will definitely performs worse than the 2D ML estimation in CMCC, particularly when the SNR is small. Another notable phenomenon is, the synchronization performance degrades when the number of clutter path increases for both CMCC and S-CMCC. The underlying reason for this phenomenon is that, at a constant SNR, the power distribution on the fingerprint spectrum tends to become more uniform with an increasing number of clutter path (assuming the static objects reflecting clutter are

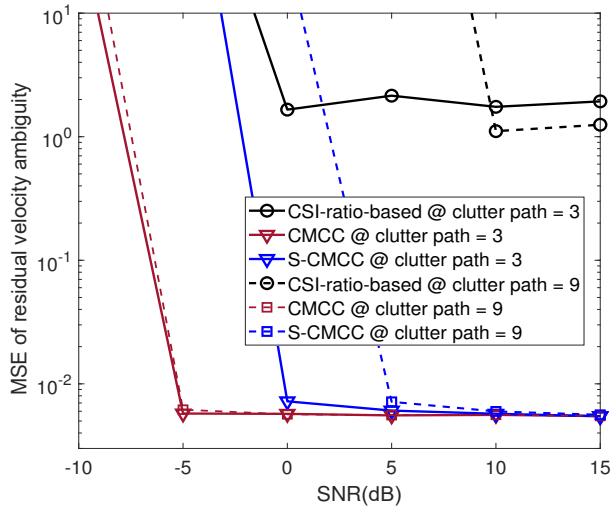


Fig. 4. Comparison of MSE for velocity estimation in asynchronous systems.

randomly distributed). Then, a more uniform power distribution essentially diminishes the distinctiveness of a fingerprint spectrum, resulting in a reduction of the correlation peak in (11) and (13). This, in turn, reduces the mean squared error (MSE) of synchronization. Moreover, as shown in the figure, an obvious MSE error floor emerges when the SNR surpasses a certain threshold. This phenomenon is caused by the 2D-DFT estimation, which involves identifying the coordinates of the largest peak in the 2D discrete grid spectrum, as depicted in Fig. 2. Since the coordinates of the largest peaks typically do not exactly align with the grid, inevitable estimation errors occur, resulting in the observed MSE error floor. The error floor can be mitigated by improving the resolution of 2D-DFT or using super-resolution estimation algorithms.

In Fig. 4, we evaluate the synchronization performance of the proposed CMCC and the CSI-ratio based schemes with varying numbers of clutter paths. It's important to note that the state-of-the-art CSI-ratio based scheme does not explicitly obtain the CFO; hence, the MSE of the CFO estimate cannot be used as the evaluation metric. Specifically, the CSI-ratio based scheme does not directly estimate the CFO but instead mitigates its impact on received signals by utilizing the CSI ratio to estimate velocity. In contrast, the proposed CMCC algorithm directly estimates both the CFO and TO, then compensates for the offsets in the estimated range and velocity. Therefore, we use the MSE of the estimated velocity under asynchronous systems to characterize synchronization performance. As shown in the figure, regardless of the number of clutter paths, our CMCC scheme consistently outperforms the CSI-ratio based scheme. However, both schemes exhibit a degradation in performance with an increase in the number of clutter paths. Overall, the simulations demonstrate the superior performance of CMCC synchronization.

V. Conclusions

To synchronize transceivers in asynchronous PMNs, we propose the joint CFO and TO synchronization algorithm suitable for NLOS scenarios, CMCC, along with its simplified version, S-CMCC. Simulation results demonstrate that our algorithm achieves significantly improved synchronization performance, with performance notably influenced by the number of clutter paths. Future improvements can be pursued by exploring the correlation between window functions and synchronization performance.

References

[1] K. Meng, Q. Wu, W. Chen, and D. Li, "Sensing-assisted communication in vehicular networks with intelligent surface," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 1, pp. 876–893, Jan. 2024.

[2] 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 Journal on Selected Areas in Communications*, vol. 40, no. 6, pp. 1728–1767, Jun. 2022.

[3] X.-Y. Wang, S. Yang, J. Zhang, C. Masouros, and P. Zhang, "Clutter suppression, time-frequency synchronization, and sensing parameter association in asynchronous perceptive vehicular networks," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 10, Oct. 2024, DOI:10.1109/JSAC.2024.3414581.

[4] Z. Wei, Y. Wang, L. Ma, S. Yang, Z. Feng, C. Pan, Q. Zhang, Y. Wang, H. Wu, and P. Zhang, "5G PRS-based sensing: A sensing reference signal approach for joint sensing and communication system," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 3, pp. 3250–3263, Mar. 2023.

[5] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, "Joint radar and communication design: Applications, state-of-the-art, and the road ahead," *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3834–3862, Jun. 2020.

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

[7] X.-Y. Wang, S. Yang, T.-H. Yuan, H.-Y. Zhai, J. Zhang, and L. Hanzo, "High-performance low-complexity hierarchical frequency synchronization for distributed massive MIMO-OFDMA systems," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 9, pp. 12 343–12 348, Sep. 2023.

[8] Z. Ni, J. A. Zhang, X. Huang, K. Yang, and J. Yuan, "Uplink sensing in perceptive mobile networks with asynchronous transceivers," *IEEE Transactions on Signal Processing*, vol. 69, pp. 1287–1300, Feb. 2021.

[9] X. Li, D. Zhang, Q. Lv, J. Xiong, S. Li, Y. Zhang, and H. Mei, "IndoTrack: Device-free indoor human tracking with commodity Wi-Fi," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, p. 1–22, Sep. 2017.

[10] 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," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 3, pp. 1–26, Sep. 2019.

[11] Y. Zeng, D. Wu, J. Xiong, J. Liu, Z. Liu, and D. Zhang, "MultiSense: Enabling multi-person respiration sensing with commodity WiFi," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, no. 3, pp. 1–29, Sep. 2020.

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

[13] Y. Liu, G. Liao, Y. Chen, J. Xu, and Y. Yin, "Super-resolution range and velocity estimations with OFDM integrated radar and communications waveform," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11 659–11 672, Oct. 2020.

[14] A. V. Oppenheim, J. R. Buck, and R. W. Schafer, *Discrete-Time Signal Processing*. Upper Saddle River, NJ, USA: Prentice Hall, 2001.

[15] "5G NR user equipment (UE) radio transmission and reception," TS 38.101, V15.2.0, 3GPP, Jul. 2018.