Windowing Optimization for Fingerprint-Spectrum -Based Passive Sensing in Perceptive Mobile Networks

Xiao-Yang Wang, Shaoshi Yang, *Senior Member, IEEE*,
Hou-Yu Zhai, Christos Masouros, *Fellow, IEEE*, J. Andrew Zhang, *Senior Member, IEEE*

Abstract—Perceptive mobile networks (PMN) have been widely recognized as a pivotal pillar for the sixth generation (6G) mobile communication systems. However, the asynchronicity between transmitters and receivers results in velocity and range ambiguity, which seriously degrades the sensing performance. To mitigate the ambiguity, carrier frequency offset (CFO) and time offset (TO) synchronizations have been studied in the literature. However, their performance can be significantly affected by the specific choice of the window functions harnessed. Hence, we set out to find superior window functions capable of improving the performance of CFO and TO estimation algorithms. We firstly derive a near-optimal window, and the theoretical synchronization mean square error (MSE) when utilizing this window. However, since this window is not practically achievable, we then test a practical "window function" by utilizing the multiple signal classification (MUSIC) algorithm, which may lead to excellent synchronization performance.

Index Terms—Integrated sensing and communication (ISAC), perceptive mobile network (PMN), orthogonal frequency-division multiplexing (OFDM), synchronization, window function.

I. INTRODUCTION

WITH the explosive growth of wireless tele-traffic, cellular networks are resorting to ultra-broadband communications technologies in the millimeter wave (mmWave) or terahertz (THz) bands [2], [3], which are naturally suited also

Xiao-Yang Wang, Shaoshi Yang and Hou-Yu Zhai are with the School of Information and Communication Engineering, Beijing University of Posts and Telecommnuications, the Key Laboratory of Universal Wireless Communications, Ministry of Education, and the Key Laboratory of Mathematics and Information Networks, Ministry of Education, Beijing 100876, China. Xiao-Yang Wang is also with the Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, UK (E-mail:{wangxy_028, 2hy, shaoshi.yang}@bupt.edu.cn).

Christos Masouros is with the Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, UK (E-mail: c.masouros@ucl.ac.uk).

J. Andrew Zhang is with the Global Big Data Technologies Centre, University of Technology Sydney, Sydney NSW 2007, Australia (E-mail: andrew.zhang@uts.edu.au).

for radar applications. With the assistance of mmWave and THz frequencies, high-precision sensing [4], [5], such as range and velocity estimation, human activity sensing [6], [7] and imaging can be performed in mobile networks [8], resulting in the perceptive mobile network (PMN) concept [9]–[12]. By exploiting the sensing capability, the PMN allows mutual benefits between the sensing and communication modules, such as an innovative framework that mutually enhances estimation accuracy and beamforming gain [5] with power allocation. PMN relies on a pair of distinct sensing types: passive and active sensing [11], [13]. In passive sensing, the remote radio unit (RRU) estimates the desired parameters by exploiting the reflected signals of the user equipments (UEs) [14] or other RRUs [11]. In active sensing mode, the RRU or UE works by processing the echo signals transmitted by itself [15]–[18]. Nevertheless, high-performance active sensing requires practical full-duplex technology, which is not mature enough to be implemented efficiently at the time of writing [16], [19]. Moreover, sensing in time-division duplex (TDD), such as [17], or frequency-division duplex (FDD), such as [16], [20], constrains the achievable communications rate [19]. As a result, passive sensing offers a competitive alternative.

However, challenges arise when aiming for high-precision passive sensing. The geographically separated transceivers of passive sensing systems inevitably use different oscillators, which leads to carrier frequency offset (CFO) and time offset (TO) between the transmitters and receivers [21]–[23]. Furthermore, the CFO and TO will result in nonnegligible velocity and range sensing ambiguity, which severely degrades the sensing performance [24]. To mitigate the ambiguity, both time and carrier frequency synchronization should be implemented [25]. Nevertheless, due to the systematic difference between sensing in PMN [21] and in bi-static radar [26], synchronization methods for traditional bi-static radar cannot be applied in PMN. Therefore, synchronization algorithms for passive sensing need to be carefully re-designed.

In passive sensing PMN systems, a limited number of synchronization methods have been reported [10], [27], [28]. The first method designed for passive sensing is the cross-antenna cross-correlation (CACC) algorithm [29], which exploits the cross-correlation between signals received at different antennas to extract TO and CFO. However, the cross-correlation doubles the number of parameters to be estimated, which results in excessive computational complexity. To reduce the complexity, the mirrored multiple signal classification (MUSIC) algorithm

was proposed in [24]. This concept halves the number of estimated parameters, thus decreasing the overall computational complexity. However, these two algorithms are unsuitable for non-line-of-sight (NLOS) or single-antenna scenarios. Another synchronization algorithm, namely FarSense [30], uses the ratio of the channel state information (CSI) measurements corresponding to different antennas to extract the CFO and the up-to-date value of TO during a period in single-target scenarios. However, this algorithm cannot be applied to multi-target scenarios. To address this issue, the authors extended, FarSense to MultiSense [31], where a filter is designed to separate signals reflected by different targets, enabling the estimation of the multi-pair values of TO and CFO. This modification allows MultiSense to operate in multi-target scenarios. Additionally, leveraging properties of the Mobius transform, the authors of [32] propose a family of CSI-ratio-based frequency synchronization schemes, which can only mitigate CFO, in scenarios where a single target exists and its velocity is low. To further characterize the impact of this kind of synchronization method on Doppler sensing, [33] derives the closed-from Cramér–Rao lower bound (CRLB) for Doppler sensing under CSI-ratio-based synchronization. However, the CSI-ratio based schemes in these works cannot support range estimation. Additionally, it is important to note that the aforementioned synchronization studies are conducted for multi-carrier passive sensing PMN based on existing mobile communication systems. Meanwhile, the authors of [34] propose a time and frequency synchronization method for cooperative multi-static imaging at a single frequency.

The general framework of the aforementioned synchronization schemes for multi-subcarrier PMNs involves extracting the TO and CFO by processing different signals received from different antennas with the aid of either cross-correlation methods, such as [24], [29], [38], or simple division operations, such as [30], [31]. However, this research framework *restricts applications to multi-antenna scenarios*. In response to this limitation, a framework achieving synchronization by processing signals from different paths is explored [35]–[37]. In principle, this framework does not require multiple receiving antennas. However, in [35], Doppler offset estimation at a single frequency is implemented by separating each dynamic and static path with the estimated angle of arrival (AOA), still necessitating a multi-antenna receiver. In [36], although TO estimation can be achieved in single-antenna systems by exploiting the slow-changing delay profile of the channel impulse response, CFO estimation relies on high-resolution angle of departure (AOD), which cannot be obtained in single-antenna systems. Moreover, the authors of [37] separate signals from multiple paths relying on a prior phase offset dictionary, which strongly depends on time delay, packet detection delay and carrier frequency offset. However, this work achieves synchronization by solving high-dimensional compressed sensing optimization problems with very high computational complexity. Overall, implementing these schemes in single-antenna PMNs is also challenging.

To further address synchronization in single-antenna systems, the cross-multipath cross-correlation (CMCC) synchronization algorithm in [39] explored a novel framework with

several key insights: 1) the delay-Doppler spectrum of the static signal component, namely the *fingerprint spectrum* [39], is uniquely identified by location distribution of static objects reflecting signals; 2) the fingerprint spectrum will cyclically shift, when the CFO and TO drift. As a benefit of these insights, CMCC achieves high-precision CFO and TO estimation in LOS/NLOS and single-target/multi-target scenarios[1]. For ease of comparison, we summarize the application scenarios for the aforementioned algorithms and CMCC in Table I. Concretely, in CMCC, the CFO and TO are acquired by locating the peak of the cross-correlation sequence of the fingerprint spectra at different instants of time. Since the fingerprint spectra are composed by the Fourier transform of the window functions utilized, the window function adopted naturally impacts the estimation performance [39]. However, the existing contributions employ standard window functions and there is an open research horizon for re-designing the windows to further enhance the synchronization performance.

To improve the synchronization performance of CMCC, we conduct theoretical analysis to acquire superior window functions. In a nutshell, this research can be divided into two main parts. The first part focuses on designing an asymptotically optimal window for minimizing the synchronization mean square error (MSE), and deriving the corresponding MSE. Concretely, the first part is broken down into three steps. Firstly, we find the fingerprint spectrum sequence corresponding to the asymptotically optimal window. Then, we derive the estimation MSE corresponding to the asymptotically optimal window. In the third step, we derive the asymptotically optimal window. However, since the acquisition of the asymptotically optimal window is practically unachievable, the second part of our work aims for designing superior practical window functions using a novel insight. Bearing in mind that the CMCC algorithm can be generalized to other bi-static integrated sensing and communication (ISAC) systems, the CMCC and this window function research may also be utilized in other kind of bi-static systems, such as that of [10]. Specifically, our key contributions are summarized as follows.

- We propose an asymptotically optimal fingerprint spectrum sequence, which achieves extremely low MSE in high-SNR scenarios. Moreover, based on the expression of the asymptotically optimal fingerprint spectrum, we derive the corresponding MSE. Both the fingerprint spectrum derived and MSE are formulated in Proposition 1.
- We derive the mathematical relationship between the fingerprint spectrum and the window utilized. Specifically, to achieve this in Proposition 2, we first derive the equation for the fingerprint spectrum sequence and the cross-correlation sequence of the window utilized. Secondly, we develop Proposition 3 to establish the relationship between the window utilized and the cross-correlation sequence. According to our derivation, achieving a specifically designed fingerprint spectrum relies

---

[1]Since the synchronization is implemented in the delay-Doppler domain, parameter estimation algorithms designed for orthogonal time frequency space (OTFS) modulation [40], such as the fractional delay and Doppler estimation [41] and the compressed-sensing based channel estimation [42], can be potentially used to further improve the synchronization performance.

TABLE I
COMPARISON BETWEEN STATE-OF-THE-ART CFO AND/OR TO ESTIMATION METHODS AND OUR SCHEME FOR ASYNCHRONOUS SENSING IN PMN

Methods	Propagation Environment	Number of Targets	Sensing Functions Supported	Number of Antennas	Number of Subcarriers
CACC [29], mirrored-MUSIC [24]	LOS	Multi-target	velocity, range	Multi-antenna	Multi-subcarrier
FarSense [30]	LOS / NLOS	Single-target	velocity	Multi-antenna	Multi-subcarrier
MultiSense [31]	LOS / NLOS	Multi-target	velocity	Multi-antenna	Multi-subcarrier
Schemes in [32]	LOS / NLOS	Single-target	velocity	Multi-antenna	Multi-subcarrier
Scheme in [35]	LOS	Multi-target	velocity	Multi-antenna	Single-subcarrier
Scheme in [34]	LOS / NLOS	Multi-target	cooperative imaging	Single-antenna, Multi-antenna	Single-subcarrier
JUMP [36]	LOS / NLOS	Multi-target	velocity, range	Multi-antenna	Multi-subcarrier
Our scheme, SHARP [37]	LOS / NLOS	Multi-target	velocity, range	Single-antenna, Multi-antenna	Multi-subcarrier

on many parameters, which we cannot acquire before sensing.

- We disclose that the estimation performance for fingerprint-spectrum-based synchronization improves as the mainlobe of the Fourier transform of the window used becomes sharper. Building on this finding, we test a method that employs a well-known super-resolution estimation algorithm, specifically the MUSIC algorithm, to generate a fingerprint spectrum. Due to the high resolution and inherently sharp lobe characteristics of this algorithm, the resulting spectrum's lobe is significantly narrower than that of traditional windows, such as the rectangular and Hamming windows. Numerical simulations confirm the superior synchronization performance of the bell-shaped window with the MUSIC algorithm (BS-W-MUSIC).

The remainder of this paper is organized as follows. In Section II, we describe the asynchronous MIMO-OFDM PMN system model and formulate the window function design problem. Then, we design practical window functions in Section III. In Section IV, numerical simulations are conducted to verify our analysis and to confirm the superior performance of the proposed window. Finally, our conclusions are offered in Section VI.

Notations: We use lower-case and upper-case bold-face letters to represent vectors and matrices, respectively. $\mathbf{A}^{\mathrm{T}}, \mathbf{A}^{*}, \mathbf{A}^{\mathrm{H}}$ and $\mathbf{A}^{-1}$, $\mathbf{A}^{\dagger}$ represent transpose, conjugate, conjugate transpose, inverse and the pseudo-inverse of $\mathbf{A}$, respectively; Moreover, $\mathrm{diag}(\mathbf{a}_1, \cdots, \mathbf{a}_n)$ is a block diagonal matrix whose diagonal blocks are $\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$. $\mathbf{A}[i,:]$ and $\mathbf{A}[:,j]$ represents the ith row and jth column of matrix $\mathbf{A}$; $\mathbf{a}[i]$ and $\mathbf{A}[i,j]$ represents the ith element of $\mathbf{a}$ and (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 t, respectively; $\mathbf{I}_N$ and $\mathbf{0}_{M \times N}$ are the $N \times N$ identity matrix and $M \times N$ zero matrix, respectively; $\mathrm{Round}(\alpha)$ and $\mathrm{E}(\cdot)$ represent the nearest integer to α and the expectation operator, respectively; $\lfloor \cdot \rfloor$, $\mathrm{Int}(\cdot)$ and $\mathrm{Frac}(\cdot)$ denotes the floor function, the integer part and the decimal parts of a real number; Finally, $\oplus$, $\odot$, $\circledast$ and $\mathrm{mod}(\cdot)$ are the right cyclic shift operator, the Hadamard product operator, the circular convolution operator, and the modulus operator, respectively.
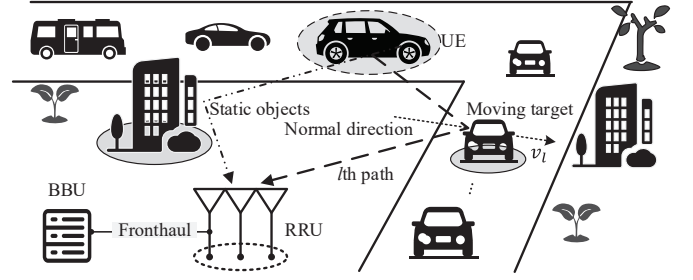
Fig. 1. System model.

II. SYETEM MODEL

The PMN system is depicted in Fig. 1, where a M_{r}-antenna RRU receives signals transmitted by the UE. The propagation channel consists of L paths. Let us define the propagation delay and the Doppler offset corresponding to the lth path as τ_l and $f_{\mathrm{D},l}$. Then, by estimating τ_l and $f_{\mathrm{D},l}$ for $l = 1, \cdots, L$, the RRU achieves both velocity measurement and ranging [9]. However, the existence of CFO, namely δ^f, and TO, namely δ^τ, between the UE and the RRU will introduce ambiguity in the estimation of the desired parameters: τ_l and $f_{\mathrm{D},l}$ for $l = 1, \cdots, L$ [43]. To mitigate the ambiguity, the whole data payload can be utilized as the sensing resource to extract both delay-Doppler and CFO/TO information, as described in [9], [10].

Specifically, we denote the carrier frequency and the subcarrier spacing in the system by f_c and Δf. Moreover, let us assume that there are a total of N_{c} OFDM subcarriers employed in the system and define c_n as the data modulated on the nth subcarrier. Moreover, we denote the number of antennas of the UE as M_{t} and the transmit precoder (TPC) of the UE as $\boldsymbol{\omega} \in \mathbb{C}^{M_{\mathrm{t}} \times 1}$. Therefore, we can represent the analog signal transmitted by the UE as

$$\mathbf{x}(t) = \left[e^{j2\pi f_c t} \sum_{n=0}^{N_{\mathrm{c}}-1} c_n e^{j2\pi n \Delta f t} \right] \boldsymbol{\omega}. \tag{1}$$

Furthermore, the channel impulse response (CIR) of lth path

can be formulated as[2]

$$\mathbf{H}_l(t) = a_l \delta[(1 - \frac{2v_l}{c})t - (\tau_l + \delta^\tau)]\mathbf{\Omega}_l^r \mathbf{\Omega}_l^t, \qquad (2)$$

where a_l and v_l are defined as the channel gain of the lth path and the velocity of the reflector projected onto the normal direction of the lth path, respectively. Moreover, $\delta(t)$, $\mathbf{\Omega}_l^r = [1, e^{j2\pi d \cos(\theta_l^r)/\lambda}, \cdots, e^{j2\pi(M_r-1)d\cos(\theta_l^r)/\lambda}]^T$, and $\mathbf{\Omega}_l^t = [1, e^{j2\pi d \cos(\theta_l^t)/\lambda}, \cdots, e^{j2\pi(M_r-1)d\cos(\theta_l^t)/\lambda}]$ are the impulse function, the received steering vector, and the transmitted steering vector of the signal propagated from the lth path, respectively. Furthermore, d, θ_l^r, θ_l^t, and λ are defined as the antenna spacing, the angle of arrival (AOA), the angle of departure (AOD) corresponding to the lth path, and the signal wavelength, respectively.

Then, since CP converts linear convolution to circulant convolution, the received analog baseband signal of the gth OFDM symbol after removing the cyclic prefix (CP) can be formulated as (3)

$$\mathbf{y}_g(t) = e^{-j2\pi(f_c+\delta^f)t} \sum_{l=1}^{L}[\mathbf{H}_l(t) \circledast \mathbf{x}(t)]$$

$$= \sum_{l=1}^{L} \sum_{n=0}^{N_c-1} a_l c_n \mathbf{\Omega}_l^r \mathbf{\Omega}_l^t \boldsymbol{\omega} e^{-2\pi\delta^f t} e^{-j2\pi(f_c+n\Delta f)[\frac{2v_l}{c}t+(\delta^\tau+\tau_l)]}$$

$$\cdot e^{j2\pi n \Delta f t} + \mathbf{w}_g(t), \qquad (3)$$

where $\mathbf{w}_g(t)$ represents the complex additive white Gaussian noise (AWGN) process having zero mean and variance of σ^2 for any t.

Upon acquiring $\mathbf{y}_g(t)$, the sampler will digitize it into a discrete-time signal. To represent the discrete-time signal, we define the sampling interval and the length of an entire OFDM symbol as T_{sam} and T_{sym}, respectively. Then, the uth sample of the digitized $\mathbf{y}_g(t)$ can be formulated as

$$\mathbf{y}_{g,u} = \sum_{l=1}^{L} \sum_{n=0}^{N_c-1} e^{-j2\pi f_c[\frac{2v_l}{c}(uT_{\text{sam}}+(g-1)T_{\text{sym}}+(\delta^\tau+\tau_l)]}$$

$$\cdot e^{-j2\pi n\Delta f \frac{2v_l}{c}(uT_{\text{sam}}+(g-1)T_{\text{sym}})} e^{-2\pi\delta^f(uT_{\text{sam}}+(g-1)T_{\text{sym}})}$$

$$\cdot e^{-j2\pi n\Delta f(\delta^\tau+\tau_l)} e^{-j2\pi n\Delta f u T_{\text{sam}}} a_l c_n \mathbf{\Omega}_l^r \mathbf{\Omega}_l^t \boldsymbol{\omega} + \mathbf{w}_{g,u}, \qquad (4)$$

where $\mathbf{w}_{g,u}$ is defined as the complex AWGN vector with zero mean and covariance matrix $\sigma^2\mathbf{I}$, respectively. Noting that $[uT_{\text{sam}}+(g-1)T_{\text{sym}}]N_c\Delta f \frac{2v_l}{c}$ is always small, one can approximate $e^{-j2\pi n\Delta f \frac{2v_l}{c}[uT_{\text{sam}}+(g-1)T_{\text{sym}}]}$ by 1 [44]. Moreover, since $2\pi f_c \frac{2v_l}{c}T_{\text{sym}}$ and $2\pi T_{\text{sym}}\delta^f$ are also much smaller than 1, $e^{-j2\pi f_c[\frac{2v_l}{c}(uT_{\text{sam}}+(g-1)T_{\text{sym}})]}$ and $e^{-2\pi\delta^f[uT_{\text{sam}}+(g-1)T_{\text{sym}}]}$ can be approximated by $e^{-j2\pi f_c[\frac{2v_l}{c}(g-1)T_{\text{sym}}]}$ [44] and $e^{-2\pi\delta^f(g-1)T_{\text{sym}}}$, respectively.

By implementing these approximations, the gth OFDM, $\mathbf{Y}_g = [\mathbf{y}_{g,1}, \cdots, \mathbf{y}_{g,N_c}]$, can be further expressed as

$$\mathbf{Y}_g = \sum_{l=1}^{L} e^{-j2\pi f_c[(\frac{2v_l}{c}+\frac{\delta^f}{f_c})(g-1)T_{\text{sym}}]} e^{-j2\pi f_c(\delta^\tau+\tau_l)}$$

$$\cdot a_l \mathbf{\Omega}_l^r \mathbf{\Omega}_l^t \boldsymbol{\omega}\boldsymbol{\tau}_l \mathbf{D}\mathbf{F} + \mathbf{W}_g, \qquad (5)$$

[2]It's worth noting that in this work, the time-varying CIR of the lth path should be represented as $\mathbf{H}_l(t, \tau_l)$. However, given that τ_l is also a function of t, we simplify the representation to $\mathbf{H}_l(t)$.

where $\boldsymbol{\tau}_l$ is defined as $[1, \cdots, e^{-j2\pi(N_c-1)\Delta f(\delta^\tau+\tau_l)}]$, $\mathbf{D}$ represents $\text{diag}([c_1, \cdots, c_N])$, $\mathbf{F}$ is defined as the N_c-dimensional inverse fast Fourier transform (N_c-D IFFT) matrix, and $\mathbf{W}_g$ is defined as $[\mathbf{w}_{g,1}, \cdots, \mathbf{w}_{g,N_c}]$, respectively. To further estimate the desired parameters, namely the Doppler offset and time delay, a feasible method is to firstly compensate the received signal exploiting the demodulated communication data as

$$\breve{\mathbf{Y}}_g = \mathbf{Y}_g \mathbf{F}^H \mathbf{D}^{-1}. \qquad (6)$$

which is also exploited in [44]. Specifically, the process of utilizing data payload for target sensing is depicted in Fig. 2. The only connection between sensing synchronization studied in this paper and the communication module can be intuitively seen from the demodulated data flow. The data demodulated from the communication module is fed to the sensing module. Then, the data is first compensated and then used for our synchronization and target sensing. Given the unidirectional data flow and assuming there are no symbol demodulation errors in the flow, the performance of synchronization studied in this paper and the communication module will have no impact on each other.
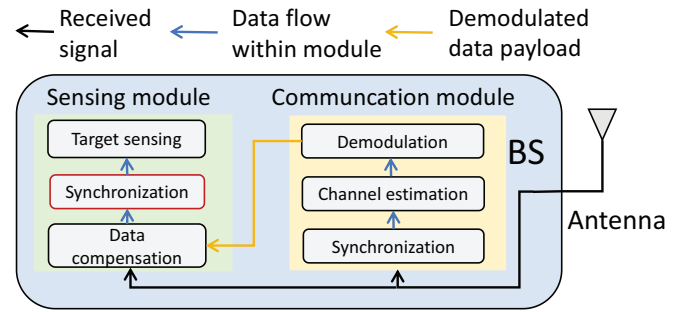
Fig. 2. Data processing flow at receivers in PMN systems.

Furthermore, let us denote the mth row of $\breve{\mathbf{Y}}_g$ as $\breve{\mathbf{y}}_{g,m}$. Then, we stack the received signals corresponding G OFDM symbols into $\mathbf{\Gamma}_m$ as $[(\breve{\mathbf{y}}_{1,m})^T, \cdots, (\breve{\mathbf{y}}_{G,m})^T]^T$, and assume that both the TO and the phase drift are constants within the packet. Specifically, $\mathbf{\Gamma}_m$ can be formulated as

$$\mathbf{\Gamma}_m = \sum_{l=1}^{L} \boldsymbol{\alpha}_l[m]\boldsymbol{\theta}_l \boldsymbol{\tau}_l + \breve{\mathbf{W}}_m, \qquad (7)$$

where $\boldsymbol{\alpha}_l = e^{-j2\pi f_c(\delta^\tau+\tau_l)}a_l\mathbf{\Omega}_l^r\mathbf{\Omega}_l^t\boldsymbol{\omega}$, $\boldsymbol{\theta}_l = [1, \cdots, e^{-j2\pi f_c(\frac{2v_l}{c}+\frac{\delta^f}{f_c})(G-1)T_{\text{sym}}}]^T$ and $\breve{\mathbf{W}}_m$ is the complex AWGN matrix, respectively.

III. WINDOWING FOR SYNCHRONIZATION AND FORMULATION OF OUR OPTIMIZATION PROBLEM

In this section, we firstly extend the synchronization algorithm of [39] by applying window functions for performance improvement. Then, we formulate an optimization problem to find the optimal window function.

As presented in [39], a two-dimensional Fourier transform (2D-FFT) is firstly applied to $\mathbf{\Gamma}_m \in \mathbb{C}^{G \times N_c}$. We then introduce the window functions during the process of the 2D-FFT. To represent the output of the 2D-FFT, $\boldsymbol{\psi}_G \in \mathbb{C}^{1 \times G}$ is defined as the G-element window function. As a result, the

output $\mathbf{F}_G^*\text{diag}(\boldsymbol{\psi}_G)\boldsymbol{\Gamma}_m\text{diag}(\boldsymbol{\psi}_{N_c})\mathbf{F}_{Nc}^*$ can be expressed as (8), where $\tilde{\mathbf{W}} = \mathbf{F}_G^*\breve{\mathbf{W}}_m\mathbf{F}_{N_c}^*$ is the AWGN matrix with zero mean and covariance matrix $\tilde{\sigma}^2\mathbf{I}$. Then we have

$$\mathbf{F}_G^*\text{diag}(\boldsymbol{\psi}_G)\boldsymbol{\Gamma}_m\text{diag}(\boldsymbol{\psi}_{N_c})\mathbf{F}_{Nc}^*$$
$$= \sum_{l=1}^L \boldsymbol{\alpha}_l[m][\mathbf{F}_G^*\text{diag}(\boldsymbol{\psi}_G)\boldsymbol{\theta}_l][\boldsymbol{\tau}_l\text{diag}(\boldsymbol{\psi}_{N_c})\mathbf{F}_{N_c}^*]+\tilde{\mathbf{W}}. \tag{8}$$

To reduce complexity in the 2D-FFT, it is possible to substitute $\boldsymbol{\Gamma}_m$ with its real part [39]. However, it is important to note that this operation leads to a half of the maximum unambiguous range.

It is important to note that the 2D-FFT can be substituted by alternative spectrum analysis methods, such as the MU-SIC and ESPRIT algorithms. Hence the 2D-FFT is adopted, because it makes closed-form performance analysis tractable. A detailed performance comparison between the MUSIC and 2D-FFT methods is provided later in section VI of this paper. Additionally, it is worth mentioning that the windows $\boldsymbol{\psi}_G$ and $\boldsymbol{\psi}_{N_c}$ used in the 2D-FFT can be of different types. For instance, $\boldsymbol{\psi}_G$ may be a rectangular window with G elements, while $\boldsymbol{\psi}_{N_c}$ could be a Hamming window with N_c elements. However, for the sake of brevity, we assume that $\boldsymbol{\psi}_G$ and $\boldsymbol{\psi}_{N_c}$ are of the same type in the following discussions.

To improve the estimation performance, zero-padding can be utilized [45]. In the following, we implement K times zero-padding. In such circumstances, (8) can be re-written as

$$\mathbf{F}_{KG}^*\text{diag}(\bar{\boldsymbol{\psi}}_{K,G})\bar{\boldsymbol{\Gamma}}_{K,m}\text{diag}(\bar{\boldsymbol{\psi}}_{K,N_c})\mathbf{F}_{KNc}^*=\tilde{\mathbf{W}}_K+\sum_{l=1}^L \boldsymbol{\alpha}_l[m]$$
$$[\mathbf{F}_{KG}^*\text{diag}(\bar{\boldsymbol{\psi}}_{K,G})\bar{\boldsymbol{\theta}}_{K,l}][\mathbf{F}_{KN_c}^H\text{diag}(\bar{\boldsymbol{\psi}}_{K,N_c})\bar{\boldsymbol{\tau}}_{K,l}^T]^T, \tag{9}$$

where we have $\bar{\boldsymbol{\psi}}_{K,G} = [\boldsymbol{\psi}_G, \mathbf{0}_{1\times(K-1)G}]$, $\bar{\boldsymbol{\psi}}_{K,N_c} = [\boldsymbol{\psi}_{N_c}, \mathbf{0}_{1\times(K-1)N_c}]$, $\bar{\boldsymbol{\theta}}_{K,l} = [\boldsymbol{\theta}_l^T, \mathbf{0}_{1\times(K-1)G}]^T$, $\bar{\boldsymbol{\tau}}_{K,l} = [\boldsymbol{\tau}_l, \mathbf{0}_{1\times(K-1)N_c}]$, and $\bar{\boldsymbol{\Gamma}}_{K,m}$ is defined as

$$\bar{\boldsymbol{\Gamma}}_{K,m} = \begin{bmatrix} \boldsymbol{\Gamma}_m & \mathbf{0}_{G\times(K-1)N_c} \\ \mathbf{0}_{(K-1)G\times N_c} & \mathbf{0}_{(K-1)G\times(K-1)N_c} \end{bmatrix}. \tag{10}$$

Moreover, $\tilde{\mathbf{W}}_K$ is defined as

$$\tilde{\mathbf{W}}_K = \mathbf{F}_{KG}^* \begin{bmatrix} \breve{\mathbf{W}}_m & \mathbf{0}_{G\times(K-1)N_c} \\ \mathbf{0}_{(K-1)G\times N_c} & \mathbf{0}_{(K-1)G\times(K-1)N_c} \end{bmatrix} \mathbf{F}_{KN_c}^*. \tag{11}$$

Then, we denote $\mathbf{F}_{KG}^*\text{diag}(\bar{\boldsymbol{\psi}}_{K,G})\bar{\boldsymbol{\Gamma}}_{K,m}\text{diag}(\bar{\boldsymbol{\psi}}_{K,N_c})\mathbf{F}_{KNc}^*$ by $\boldsymbol{\Xi}$, $[\mathbf{F}_{KG}^*\text{diag}(\bar{\boldsymbol{\psi}}_{K,G})\bar{\boldsymbol{\theta}}_{K,l}]$ by the row vector $\boldsymbol{\gamma}_{l,G}$, and $[\mathbf{F}_{KN_c}^H\text{diag}(\bar{\boldsymbol{\psi}}_{K,N_c})\bar{\boldsymbol{\tau}}_{K,l}^T]^T$ by the column vector $\boldsymbol{\gamma}_{l,N_c}^T$. Moreover, $\boldsymbol{\gamma}_{l,G}$ can also be expressed by the discrete-time Fourier transform (DTFT) of $\text{diag}(\bar{\boldsymbol{\psi}}_{K,G})\bar{\boldsymbol{\theta}}_{K,l}$, namely

$$\bar{\boldsymbol{\gamma}}_{l,G}(f) = \sum_{g=0}^{KG-1} \bar{\boldsymbol{\psi}}_{K,G}[g]\bar{\boldsymbol{\theta}}_{K,l}[g]e^{-j2\pi fg}. \tag{12}$$

Moreover, to associate the continuous DTFT output $\bar{\boldsymbol{\gamma}}_{l,G}$ and the discrete DFT output $\boldsymbol{\gamma}_{l,G}$, the grid size of $\boldsymbol{\gamma}_{l,G}$ has to be determined. For brevity, let us define T_R and F_R as the grid size in the delay-Doppler domain, respectively. F_R can be expressed as the reciprocal of the total time period, namely as $F_R = 1/(GT_{\text{sym}})$, while T_R as the reciprocal of the entire bandwidth, i.e. $T_R = 1/(N\Delta f)$, according to [44]. Due to the employment of zero-padding, the grid size

actually becomes $F_R = 1/(KGT_{\text{sym}})$ and $\frac{T_R}{K}$ [45]. Thus, we have $\boldsymbol{\gamma}_{l,G}[i] = \bar{\boldsymbol{\gamma}}_{l,G}(iF_R) = \bar{\boldsymbol{\gamma}}_{l,G}(\frac{i\Delta f}{KG})$ and $\boldsymbol{\gamma}_{l,N_c}^T[p] = \bar{\boldsymbol{\gamma}}_{l,N_c}(pT_R) = \bar{\boldsymbol{\gamma}}_{l,N_c}(\frac{pT_{\text{sam}}}{K})$. Similarly, the composition of $\boldsymbol{\gamma}_{l,N_c}^T$ is consistent with that of $\boldsymbol{\gamma}_{l,G}$. As a result, $\boldsymbol{\Xi}[i,p]$ can be expressed as

$$\boldsymbol{\Xi}[i,p]=\sum_{l=1}^L 2\pi^2\boldsymbol{\alpha}_l[m]\bar{\boldsymbol{\gamma}}_{l,N_c}\left(\frac{pT_{\text{sam}}}{K}\right)\bar{\boldsymbol{\gamma}}_{l,G}\left(\frac{i\Delta f}{KG}\right)+\tilde{\mathbf{W}}_K[i,p]. \tag{13}$$

Although the closed-form expression of $\boldsymbol{\Xi}$ has been obtained, it is not straightforward to see its shape. We can further reformulate the expression of $\boldsymbol{\Xi}$ as follows. According to (9), we know the term $[\mathbf{F}_{KG}^*\text{diag}(\bar{\boldsymbol{\psi}}_{K,G})\bar{\boldsymbol{\theta}}_{K,l}]$. Furthermore, $[\mathbf{F}_{KN_c}^H\text{diag}(\bar{\boldsymbol{\psi}}_{K,N_c})\bar{\boldsymbol{\tau}}_{K,l}^T]^T$ represents the Fourier transforms of the product of $\text{diag}(\bar{\boldsymbol{\psi}}_{K,G})$ and $\bar{\boldsymbol{\theta}}_{K,l}$, and $\text{diag}(\bar{\boldsymbol{\psi}}_{K,N_c})$ and $\bar{\boldsymbol{\tau}}_{K,l}^T$. According to the convolution theorem [45], namely the Fourier transform of a convolution between two sequences equals the product of the Fourier transforms of the two sequences, $[\mathbf{F}_{KG}^*\text{diag}(\bar{\boldsymbol{\psi}}_{K,G})\bar{\boldsymbol{\theta}}_{K,l}]$ and $[\mathbf{F}_{KN_c}^H\text{diag}(\bar{\boldsymbol{\psi}}_{K,N_c})\bar{\boldsymbol{\tau}}_{K,l}^T]^T$ can be re-formulated as

$$\boldsymbol{\gamma}_{l,G} = [(\mathbf{F}_{KG}^*\bar{\boldsymbol{\psi}}_{K,G}^T)\boxtimes(\mathbf{F}_{KG}^*\bar{\boldsymbol{\theta}}_{K,l})],$$
$$\boldsymbol{\gamma}_{l,N_c}^T = [(\mathbf{F}_{KN_c}^H\bar{\boldsymbol{\psi}}_{K,N_c}^T)\boxtimes(\mathbf{F}_{KN_c}^H\bar{\boldsymbol{\tau}}_{K,l}^T)]^T, \tag{14}$$

where $(\boxtimes)$ represents a matrix-wise convolution operator, defined as $[\mathbf{a}_1,\cdots,\mathbf{a}_i]\boxtimes[\mathbf{b}_1,\cdots,\mathbf{b}_i] = [\mathbf{a}_1\circledast\mathbf{b}_1,\cdots,\mathbf{a}_i\circledast\mathbf{b}_i]$. Here, $\mathbf{a}_i$ and $\mathbf{b}_i$ represent any column vectors of the same dimension. It is plansible that $\boldsymbol{\gamma}_{l,G}$ represents the discrete circular convolution sequence between the Fourier transform of the window function $\bar{\boldsymbol{\psi}}_{K,G}$ and the Fourier transforms of $\bar{\boldsymbol{\theta}}_{K,l}$. Moreover, since $\bar{\boldsymbol{\theta}}_{K,l}$ is the single-tone discrete sequence, $\bar{\boldsymbol{\gamma}}_{l,G}(f)$ can be naturally expressed as the cyclic-shift based Fourier transform of $\bar{\boldsymbol{\psi}}_{K,G}$. The symmetry axis of $\bar{\boldsymbol{\gamma}}_{l,G}(f)$ is at the exact frequency of the sequence $\bar{\boldsymbol{\theta}}_{K,l}$, namely at $f = \delta^f + f_{D,l}$. Similarly, the symmetry axis of $\bar{\boldsymbol{\gamma}}_{l,N_c}(f)$ is located at $f = \delta^\tau + \tau_l$.

Furthermore, by substituting (14) into (9) we obtain

$$\boldsymbol{\Xi} = \tilde{\mathbf{W}}_K + \sum_{l=1}^L \boldsymbol{\alpha}_l[m][(\mathbf{F}_{KG}^*\bar{\boldsymbol{\psi}}_{K,G}^T)\boxtimes(\mathbf{F}_{KG}^*\bar{\boldsymbol{\theta}}_{K,l})]$$
$$[(\mathbf{F}_{KN_c}^H\bar{\boldsymbol{\psi}}_{K,N_c}^T)\boxtimes(\mathbf{F}_{KN_c}^H\bar{\boldsymbol{\tau}}_{K,l}^T)]^T. \tag{15}$$

In the matrix $\boldsymbol{\Xi}$, the elements in each row and column are acquired by sampling from the same function $\bar{\boldsymbol{\gamma}}_{l,G}(f)$ and $\bar{\boldsymbol{\gamma}}_{l,N_c}(f)$, respectively. As a result, it is intuitive that $\boldsymbol{\Xi}$ is the sum of the Gaussian noise and the linear summation of L discrete 2-D sample matrices, namely $\boldsymbol{\gamma}_{l,G}\boldsymbol{\gamma}_{l,N_c}^T$ for $l = 1,\cdots,L$ [46], whose centers are at $[(\delta^f + f_{D,l})/F_R, (\delta^\tau + \tau_l)/T_R]$. Each of these discrete 2D sample matrices, $\boldsymbol{\gamma}_{l,G}\boldsymbol{\gamma}_{l,N_c}^T$, corresponds to signals reflected by objects. Among the L objects, we assume that there exist L_1 static objects in the environment. Since they are static, $f_{D,l}$ and τ_l corresponding to these L_1 objects are 0 and constant, respectively. Thus, according to (13), when the normalized CFO, δ^f, and TO, δ^τ, drift, these discrete 2D sample matrices corresponding to the L_1 static objects will consequently change. Let us consider the (i,p)th element of $\boldsymbol{\gamma}_{l,G}\boldsymbol{\gamma}_{l,N_c}^T$ as an example, which is given by $\bar{\boldsymbol{\gamma}}_{l,G}\left[\frac{i\Delta f}{KG}\right]\bar{\boldsymbol{\gamma}}_{l,N_c}\left[\frac{pT_{\text{sam}}}{K}\right]$. If the CFO and TO are

increased by $\Delta\delta^f$ and $\Delta\delta^\tau$ respectively, the transformed expression becomes $\bar{\gamma}_{l,G}\left[\frac{i\Delta f}{KG}+\Delta\delta^f\right]\bar{\gamma}_{l,N_c}\left[\frac{pT_{sam}}{K}+\Delta\delta^\tau\right]$. In this context, it is easy to see that the sampling points, corresponding to each sample of these sample matrices, in the 2D function $\bar{\gamma}_{l,G}(f)\bar{\gamma}_{l,N_c}(f)$ will undergo circular translation.

Moreover, it is natural that the symmetry points of these L_1 2D functions are $(\delta^f/F_R, (\delta^\tau+\tau_l)/T_R)$ because $f_{D,l}=0$ for these static objects, which means that these 2D functions are located at the same row index. Consequently, we have: $\bar{\gamma}_{l,G}(f)=\bar{\gamma}_{1,G}(f)$ and $\gamma_{l,G}=\gamma_{1,G}$ for $l=1,\cdots,L$. As a result, by locating the row index of the symmetry points of these L_1 2D functions, which is assumed as K_0, we can define $\boldsymbol{\beta}_n=\boldsymbol{\Xi}[K_0,:]$ as the *fingerprint spectrum*. The fingerprint spectrum sequence contains the location distribution information of these L_1 static objects in the environmental background. Essentially, the sequence may be deemed to be an identity code of the specific static environment. However, $\boldsymbol{\beta}_n$ is the fingerprint spectrum polluted by noise. For simplicity, we here define $\boldsymbol{\beta}\in\mathbb{C}^{1\times KN_c}$ as the pure fingerprint spectrum vector without noise and $\boldsymbol{\beta}$ can be formulated as

$$\boldsymbol{\beta}=\boldsymbol{\beta}_n-\tilde{\mathbf{W}}_K[K_0,:]. \tag{16}$$

Then, if the CFO increases by $\Delta\delta^f$ and the TO increases by $\Delta\delta^\tau$, the offset fingerprint spectrum contaminated by noise can be defined as $\boldsymbol{\beta}_{n,c}=\boldsymbol{\Xi}[K_0+\text{Round}(\Delta\delta^f/F_R),:]$. Moreover, we also define $\boldsymbol{\beta}_c$ as the offset fingerprint spectrum without noise. Therefore, by assuming that $\Delta\delta^f/F_R$ and $\frac{\Delta\delta^\tau}{T_R}$ are integers, the relationship between $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_c$ can be simply represented as[3]

$$\boldsymbol{\beta}_c[q]=\boldsymbol{\beta}\left[q+\frac{\Delta\delta^\tau}{T_R}\right]. \tag{17}$$

If $\Delta\delta^f/F_R$ and $\frac{\Delta\delta^\tau}{T_R}$ are not integers, we have

$$\boldsymbol{\beta}_c[q]=\sum_{l=1}^{L}\bar{\gamma}_{l,G}(\frac{K_0\Delta f}{KG}+\left[\text{Int}(\frac{\Delta\delta^f\Delta f}{KG})+\text{Fra}(\frac{\Delta\delta^f\Delta f}{KG})\right]\frac{KG}{\Delta f})$$
$$\cdot\bar{\gamma}_{l,N_c}(\frac{qT_{sam}}{K}+\left[\text{Int}(\frac{\Delta\delta^\tau T_{sam}}{K})+\text{Fra}(\frac{\Delta\delta^\tau T_{sam}}{K})\right]\frac{K}{T_{sam}}) \tag{18}$$

Moreover, since $\text{Fra}(\frac{\delta^f\Delta f}{KG})\frac{KG}{\Delta f}$ and $\text{Fra}(\frac{\Delta\delta^\tau T_{sam}}{K})\frac{K}{T_{sam}}$ are always less than $\frac{\Delta f}{KG}$ and $\frac{T_{sam}}{K}$, they are usually sufficiently small so that $\boldsymbol{\beta}_c[q]$ can be approximated as

$$\boldsymbol{\beta}_c[q]\approx\boldsymbol{\beta}\left[q+\frac{\Delta\delta^\tau}{T_{sam}}\right]=\sum_{l=1}^{L}\bar{\gamma}_{l,G}(\frac{K_0\Delta f}{KG}+\left[\text{Int}(\frac{\Delta\delta^f\Delta f}{KG})\right]$$
$$\frac{KG}{\Delta f})\cdot\bar{\gamma}_{l,N_c}(\frac{qT_{sam}}{K}+\left[\text{Int}(\frac{\Delta\delta^\tau T_{sam}}{K})\right]\frac{K}{T_{sam}}). \tag{19}$$

Equations (17) and (19) suggest that $\boldsymbol{\beta}_c$ may be deemed to be the cyclic-shift $\boldsymbol{\beta}$, regardless whether $\Delta\delta^f/F_R$ and $\frac{\Delta\delta^\tau}{T_R}$ are integers or not. As a result, the sliding window cross-correlation analysis between $\boldsymbol{\beta}_n\in\mathbb{C}^{1\times KN_c}$ and $\boldsymbol{\Xi}\in\mathbb{C}^{KG\times KN_c}$ can be implemented as (20) in order to find the number of cyclic shifts and consequently estimate the CFO

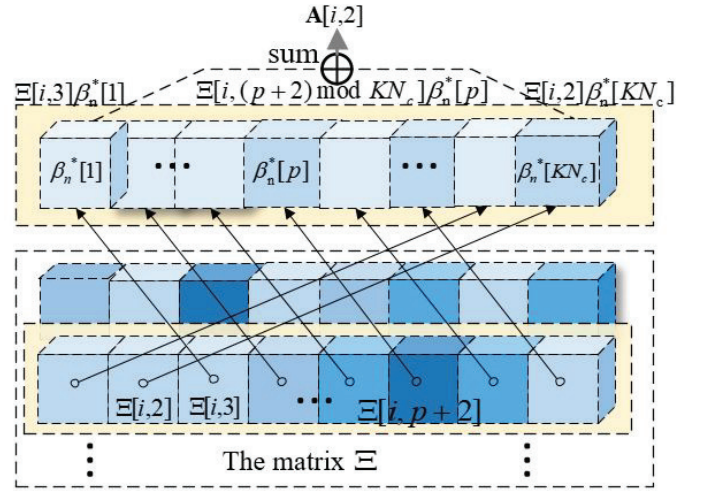[3]Assume the location of the UE does not change or changes very slowly.

Fig. 3. Demonstration of the sliding window cross-correlation in (20).

and TO. This is represented as

$$\left\{\frac{\Delta\hat{\delta}^f}{F_R},\frac{\Delta\hat{\delta}^\tau}{T_R}\right\}=\max_{i,q}\left|\sum_{p=1}^{KN_c}\frac{\boldsymbol{\Xi}[i,\bar{q}]\boldsymbol{\beta}_n^*[p]}{|\boldsymbol{\beta}_n^*|^2}\right|, \tag{20}$$
$$\text{for }i=1,\cdots,KG\text{ and }q=1,\cdots,KN_c,$$

where we have $\bar{q}=[(q+p)\bmod KN_c]$. Based on Equation (20), it is evident that $\frac{\Delta\hat{\delta}^f}{F_R}$ and $\frac{\Delta\hat{\delta}^\tau}{T_R}$ represent the relative CFO and TO with respect to a specific time instant. To obtain the absolute values of CFO and TO, it is necessary to establish an anchor with prior knowledge of velocity and range, as assumed in [24], [29]. For instance, this anchor could be a stationary UE in LOS scenarios or a static reflector in NLOS scenarios. However, in scenarios lacking an available anchor, the fingerprint spectrum can still be located through the following methods. 1) In scenarios with a priori UE, as outlined in [24], one can initially acquire the row index in LOS scenarios and subsequently utilize this index immediately upon transitioning from LOS to NLOS scenarios. Given the relatively minor change in CFO during such a brief period (the duration it takes for the scenario to shift from LOS to NLOS), the fingerprint spectrum can be accurately located. 2) In scenarios without any priori UE, one can identify the fingerprint spectrum by determining the row in $\boldsymbol{\Xi}$ with the highest power, as signals reflected by stationary objects typically exhibit greater prominence in the received signals. Moreover, in scenarios where part of static reflectors are prone to movement, it becomes crucial to frequently update the fingerprint spectrum and conduct (20). The frequent updating ensures that any changes resulting from the displacement of moving objects are captured continuously. Similarly, regular updates to (20) are necessary to estimate the CFO and TO under a slightly changing fingerprint spectrum, thereby maintaining estimation accuracy.

For simplicity, let us define $\sum_{p=1}^{KN_c}\boldsymbol{\Xi}[i,\bar{p}]\boldsymbol{\beta}_n^*[q]/|\boldsymbol{\beta}_n^*|^2$ as $\mathbf{A}[i,q]$, and by substituting (13) and $\boldsymbol{\beta}_n$ into (20), $\mathbf{A}[i,q]$ can

be further represented as

$$\mathbf{A}[i,q] = 4\pi^4 \frac{\boldsymbol{\gamma}_{1,G}[i]\boldsymbol{\gamma}_{1,G}[K_0]}{|\boldsymbol{\beta}_{\mathrm{n}}{}^*|^2} \sum_{l'=1}^{L}\sum_{l=1}^{L} \boldsymbol{\alpha}_l[m]\boldsymbol{\alpha}_{l'}[m]\boldsymbol{\rho}_{l,l'}[q] + \bar{\mathbf{W}},$$

(21)

where $\bar{\mathbf{W}}$ is the complex AWGN matrix with zero mean and variance $\bar{\sigma}^2$. Furthermore, $\boldsymbol{\rho}_{l,l'} \in \mathbb{C}^{1\times KN_c}$ is the circular convolution of the Fourier transform of the window function, $\boldsymbol{\gamma}_{l,N_c}$. Moreover, $\boldsymbol{\rho}_{l,l'}[q]$ can be formulated as

$$\boldsymbol{\rho}_{l,l'}[q] = \sum_{p=1}^{KN_c} \bar{\boldsymbol{\gamma}}_{l,N_c}[[(p+q) \bmod KN_c]T_{\mathrm{R}}]\bar{\boldsymbol{\gamma}}_{l',N_c}[pT_{\mathrm{R}}].$$

(22)

From (21) and (22), we find that $\mathbf{A}[i,q]$ is strongly coupled with $\boldsymbol{\gamma}_{l,N_c}^{\mathrm{T}}$ for $l=1,\cdots,L$, while $\boldsymbol{\gamma}_{l,G}$ only affects the signal-to-noise ratio (SNR) of $\mathbf{A}[i,q]$. Thus, the impact of k on the estimation performance is equivalent to that of the SNR. As a result, we here set the row index k as a constant $K_1 = \mathrm{Round}(\Delta\delta^f/F_{\mathrm{R}})$, which makes $\boldsymbol{\gamma}_G[K_1]\boldsymbol{\gamma}_G[K_0]$ as large as possible for simplicity.

As a result, the MSE of the delay estimate in (20) is a function of $\boldsymbol{\psi}_{N_c}$:

$$f_{\mathrm{MSE}}(\boldsymbol{\psi}_{N_c}) = \sum_{q=1}^{KN_c} \mathrm{P}_{\{q,KN_c\}}(q - \frac{\Delta\delta^\tau}{T_{\mathrm{R}}})^2,$$

(23)

where $\mathrm{P}_{\{q,KN_c\}}$ is the probability that the qth element of the set $\{\mathbf{A}[K_1,q]|q=1,\cdots,KN_c\}$ is the maximum. According to the definition, $\mathrm{P}_{\{q,KN_c\}}$ can be computed as the probability distribution function (CDF) of a KN_c-D Gaussian random variable and it can be formulated as

$$\mathrm{P}_{\{q,KN_c\}} = \int_{\mathbb{V}} p(\mathbf{s},\tilde{\sigma}^2)\mathrm{d}\,\mathbf{b}$$

$$= \int_{-\infty}^{\infty} \underbrace{\int_{-\infty}^{b_q}\cdots\int_{-\infty}^{b_q}}_{KN_c-1} \frac{e^{-(\mathbf{b}-\mathbf{s})^T(\mathbf{b}-\mathbf{s})/2\tilde{\sigma}^2}}{(2\pi)^{\frac{d}{2}}\tilde{\sigma}^d}\mathrm{d}\,b_1\cdots\mathrm{d}\,b_{KN_c}\mathrm{d}\,b_q,$$

(24)

where $p(\mathbf{s},\tilde{\sigma}^2)$ is the probability of the multidimensional Gaussian function with mean $\mathbf{s}$ and covariance matrix $\tilde{\sigma}^2\mathbf{I}_{KN_c}$. Furthermore, $\mathbf{s}$ is the expectation of the K_1th row of $\mathbf{A}$. Note that A is the linear weighted sum of the circular convolution of $\boldsymbol{\gamma}_{N_c}$ according to (21) and (22), while $\boldsymbol{\gamma}_{N_c}$ is the Fourier transform of the selected window function. As a result, the window function $\boldsymbol{\psi}_{N_c}$ has a strong influence on the synchronization MSE. Moreover, $\mathbf{b}=(b_1,\cdots,b_{KN_c})$ in (24) belongs to the KN_c-element vector space $\mathbb{V}$, which satisfies:

$$b_q > b_i, \text{for } i \in \{1,\cdots,KN_c\} \text{ and } i \neq q.$$

(25)

In the following, we intend to optimize the window function to minimize $f_{\mathrm{MSE}}(\boldsymbol{\psi}_{N_c})$:

$$\min_{\boldsymbol{\psi}_{N_c}} f_{\mathrm{MSE}}(\boldsymbol{\psi}_{N_c}),$$

$$s.t. \quad \boldsymbol{\psi}_{N_c} \in \mathbb{C}^{1\times KN_c}.$$

(26)

IV. ASYMPTOTICALLY OPTIMIZED IDEAL WINDOW FUNCTION

We intend to find the solution of (26) by the following logic. Firstly, in Proposition 1, we derive an asymptotic optimal

solution of the following problem:

$$\min_{g(\boldsymbol{\psi})} f_{\mathrm{MSE}}(g(\boldsymbol{\psi})),$$

$$s.t. \quad g(\boldsymbol{\psi}) \in \mathbb{C}^{1\times KN_c}$$

$$\boldsymbol{\psi} \in \mathbb{C}^{1\times KN_c},$$

(27)

where $g(\boldsymbol{\psi})$ is defined as the function of the expectation of the K_1th row of $\mathbf{A}$, namely $\mathbf{s}$, with respect to the window function, $\boldsymbol{\psi}$. Moreover, Proposition 1 also derives the synchronization MSE corresponding to the asymptotic optimal solution. Then, based on the relation between $\mathbf{s} = g(\boldsymbol{\psi})$ and the window function $\boldsymbol{\psi}$, Proposition 2 formulates the cross-correlation of the Fourier transform of the window function utilized for any fingerprint spectrum. Finally, based on the cross-correlation derived, Proposition 3 specifies the approximately optimal $\boldsymbol{\psi}_{N_c}$ corresponding to any $\mathbf{s}$.

Concretely, Proposition 1 is described as:

Proposition 1. *If $\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}$ is an integer, $f_{\mathrm{MSE}}(\boldsymbol{\psi}_{N_c})$ is asymptotically minimized when the SNR tends to infinity and $\mathbf{s}$ satisfies*

$$\begin{cases} \mathbf{s}[q] = 1, q = \frac{\Delta\delta^\tau}{T_{\mathrm{R}}}, \\ \mathbf{s}[q] = 0, q \neq \frac{\Delta\delta^\tau}{T_{\mathrm{R}}}, \end{cases}$$

(28)

and the corresponding MSE is $f_{\mathrm{MSE}}(\boldsymbol{\psi}_{N_c}) = \sum_{q=1,q\neq\Delta\delta^\tau/T_{\mathrm{R}}}^{KN_c}$ $\left[q_{\mathrm{r}} - \frac{\Delta\delta^\tau}{T_{\mathrm{R}}}\right]^2 \int_{-\infty}^{\infty}\left[f_{\mathcal{N}}(b_q)\ Q(\frac{b_q-1}{\bar{\sigma}^2})\ Q^{(KN_c-2)}\ (b_q/\bar{\sigma}^2)\mathrm{d}\right]b_q,$ where $Q(\cdot)$ is the cumulative distribution function of the standard Gaussian random variable.

Proof: See Appendix A.

However, unlike the hypothesis in Proposition 1, $\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}$ can hardly be an integer in actual PMN systems. According to the criterion described in the proof of Proposition 1, to make $f_{\mathrm{MSE}}(\boldsymbol{\psi}_{N_c})$ as small as possible, one should assign as much weight as possible to the index closest to $\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}$ and as low weight as possible to indices far from $\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}$. As a result, the probability corresponding to $\mathrm{Round}(\frac{\Delta\delta^\tau}{T_{\mathrm{R}}})$ should be assigned as large weight as possible. Under this case, $\mathbf{s}_{\mathrm{ap}}$ is set as

$$\begin{cases} \mathbf{s}_{\mathrm{ap}}[q] = 1, q = \mathrm{Round}(\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}), \\ \mathbf{s}_{\mathrm{ap}}[q] = 0, q \neq \mathrm{Round}(\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}), \end{cases}$$

(29)

Similarly to the case where $\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}$ is an integer, $\mathbf{s}_{\mathrm{ap}}$ in (29) will be near-optimal, when $1/\tilde{\sigma}^2$ becomes large. Specifically, based on (39), under this case we further derive $\mathrm{P}_{\{q,KN_c\}}$ by dividing $q = 1,\cdots,KN_c$ into two categories, namely $q = \mathrm{Round}(\frac{\Delta\delta^\tau}{T_{\mathrm{R}}})$ and $q \neq \mathrm{Round}(\frac{\Delta\delta^\tau}{T_{\mathrm{R}}})$

$$\begin{cases} \int_{-\infty}^{\infty} f_{\mathcal{N}}(b_q/\bar{\sigma}^2)Q^{(KN_c-1)}(b_q/\bar{\sigma}^2)\mathrm{d}\,b_q, & q = \mathrm{Round}(\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}), \\ \int_{-\infty}^{\infty} f_{\mathcal{N}}(b_q)Q(\frac{b_q-1}{\bar{\sigma}^2})Q^{(KN_c-2)}(b_q/\bar{\sigma}^2)\mathrm{d}b_q, & q \neq \mathrm{Round}(\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}). \end{cases}$$

(30)

Furthermore, since expressions for q belong to two categories, we substitute (30) into (23). Then, f_{MSE} can be further expressed as (31), where q_{r} is defined as $\mathrm{Round}(\frac{\Delta\delta^\tau}{T_{\mathrm{R}}})$.

Although Proposition 1 describes the near-optimal $\mathbf{s}$ and the corresponding $f_{\mathrm{MSE}}(\boldsymbol{\psi}_{N_c})$, we have not obtained the expression of $\boldsymbol{\rho}_{l,l'}$ corresponding to the near-optimal $\mathbf{s}_{\mathrm{ap}}$. Essentially, $\mathbf{s}_{\mathrm{ap}}$ is constituted by the circular auto-correlation sequence of the discrete Fourier transform of the selected window

$$f_{\text{MSE}}(\boldsymbol{\psi}_{N_c}) = \left[q_{\text{r}} - \frac{\Delta \delta^\tau}{T_{\text{R}}} \right]^2 \left[\int_{-\infty}^{\infty} f_{\mathcal{N}}(b_q) Q^{(KN_c-1)}(\frac{b_{q_{\text{r}}}}{\bar{\sigma}^2}) \mathrm{d} \, b_{q_{\text{r}}} + \sum_{\substack{q=1, \\ q \neq q_{\text{r}}}}^{KN_c} \int_{-\infty}^{\infty} f_{\mathcal{N}}(b_q) Q(\frac{b_q - 1}{\bar{\sigma}^2}) Q^{(KN_c-2)}(\frac{b_q}{\bar{\sigma}^2}) \mathrm{d} \, b_q \right]. \tag{31}$$

functions, namely $\boldsymbol{\rho}_{l,l'}$ for $l = 1, \cdots, L$ and $l' = 1, \cdots, L$. Moreover, for any l and l', $\boldsymbol{\rho}_{l,l'}$ is a circular-shifted version of $\boldsymbol{\rho}_{l,l}$ and for any l we have: $\boldsymbol{\rho}_{1,1} = \boldsymbol{\rho}_{l,l}$ according to (22). Thus, we intend to derive the exact $\boldsymbol{\rho}_{1,1}$ in the following.

Then, Proposition 2 outlines the closed-form expression of the desired $\boldsymbol{\rho}_{1,1}$.

Proposition 2. *The $\boldsymbol{\rho}_{1,1}$ corresponding to any achievable $\mathbf{s}$ is formulated as*

$$\boldsymbol{\rho}_{1,1} = \mathbf{s} \breve{\mathbf{J}}_{\text{Inv}}, \tag{32}$$

where $\breve{\mathbf{J}}_{\text{Inv}}$ is defined in (62) of Appendix C.

Proof: See Appendix C.

However, although Proposition 2 derives $\boldsymbol{\rho}_{1,1}$, we still do not know the expression of $\boldsymbol{\psi}_{N_c}$. To find the window function corresponding to $\boldsymbol{\rho}_{1,1}$, we intend to further derive $\boldsymbol{\psi}_{N_c}$ based on the closed-form expression of $\boldsymbol{\rho}_{1,1}$, given by Proposition 2.

Proposition 3. *The $\bar{\boldsymbol{\psi}}_{N_c}$ corresponding to any achievable $\mathbf{s}$ must satisfy the condition*

$$\begin{cases} \bar{\boldsymbol{\psi}}_{K,N_c}[i] \sum_{j=1}^{N_c} \bar{\boldsymbol{\psi}}_{K,N_c}[j] \bar{\mathbf{E}}[i,j] = \boldsymbol{\rho}_{1,1}[i], \text{for } i = 1, \cdots, N_c, \\ \boldsymbol{\rho}_{1,1}[i] = 0, \text{ for } i = N_c + 1, \cdots, KN_c, \end{cases} \tag{33}$$

where $\bar{\mathbf{E}} = \mathbf{F}_{KN_c} \mathbf{E} \mathbf{F}_{KN_c}^{\text{H}}$ and $\mathbf{E}$ is defined as the anti-diagonal matrix, where the elements on the anti-diagonal are all 1.

Proof: See Appendix D.

According to (69), the value of the globally optimal $\boldsymbol{\psi}_{N_c}$ is strongly related to both $\boldsymbol{\phi}_m$ and to time delay of every paths. Moreover, $\boldsymbol{\phi}_m$ is composed of the complex gain of every propagation path, which is hard to acquire before successful sensing. Thus, even though there exists a globally optimal $\boldsymbol{\psi}_{N_c}$, we cannot practically find it since it is strongly time-delay-and-complex-gain-dependent, before we obtain the time delay and complex gain of all paths.

V. PRACTICAL WINDOW DESIGN BASED ON SUPER RESOLUTION ESTIMATION ALGORITHM

Even though the asymptotically optimal window cannot be practically deployed, a large number of feasible discrete windows belonging to $\mathbb{C}^{1 \times KN_c}$ can be selected. However, there are no mathematical criteria clarifying how to choose 'good' windows from the set of the massive variety of feasible windows $\boldsymbol{\psi}_{N_c}$.

To establish a criterion, we revisit Proposition 1. According to this proposition, the cross-correlation sequence of the fingerprint spectrum follows a complex Gaussian distribution with a mean of $\sum_{p=1}^{KN_c} \boldsymbol{\beta}[p] \boldsymbol{\beta}_{\text{c}}[p+q]$ for $q = 1, \cdots, KN_c$, and a variance of $\bar{\sigma}^2$. Additionally, when $\tilde{\sigma}^2$ is known, $\bar{\sigma}^2$ is determined by $|\boldsymbol{\beta}_{\text{n}}|^2$.
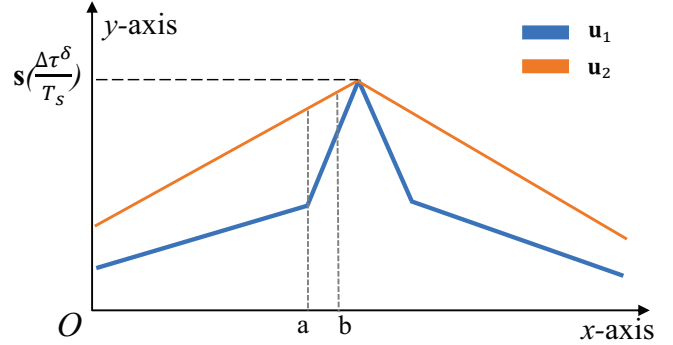

Fig. 4. $\mathbf{u}_1$ and $\mathbf{u}_2$ are two implementations of $\mathbf{s}$. $\mathbf{u}_1$ has a sharper mainlobe than $\mathbf{u}_2$.

In practical high-SNR scenarios, the noise power $\bar{\sigma}^2$ is significantly smaller than the difference between two consecutive elements on the mainlobe of the mean sequence, $|\boldsymbol{\beta}_{\text{n}}|^2 \mathbf{s}$, such as $|\boldsymbol{\beta}_{\text{n}}|^2 [\mathbf{s}(\text{Round}(\frac{\Delta \delta^\tau}{T_{\text{R}}})) - \mathbf{s}(\text{Round}(\frac{\Delta \delta^\tau}{T_{\text{R}}} - 1))]$. In such cases, if the mainlobe becomes narrower, although $|\boldsymbol{\beta}_{\text{n}}|^2$ may decrease, the change in $\bar{\sigma}^2$ is relatively smaller than the change in the difference between two adjacent elements. As a result, as shown in Fig. 4, the probability that $\mathbf{u}_2(a)$ is higher than $\mathbf{u}_2(b)$ will be lower than the probability that $\mathbf{u}_1(a)$ is higher than $\mathbf{u}_1(b)$. Consequently, the maximum will be located closer to the peak point of $\mathbf{s}$, when $\mathbf{s}$ has narrower lobe.

As (21) shows, $\mathbf{s}$ is the linear weighted sum of the cyclic shifted $\boldsymbol{\rho}_{1,1}$. Moreover, as (64) suggests, $\boldsymbol{\rho}_{1,1}$ can be represented as the convolution of the Fourier transforms of the windows utilized and the reversed sequences of the Fourier transforms, which is equivalent to the autocorrelation of the Fourier transform of the window used. In accordance with the definition of the autocorrelation, it can be viewed as the sliding multiply and sum operation for a certain sequence. Then, it is intuitive that the mainlobe sharpness of the cross-correlation, namely $\boldsymbol{\rho}_{1,1}$, is related to the mainlobe sharpness of the Fourier transform of the window utilized, given the usually high attenuation of the sidelobes. As a result, the sharper the mainlobe of the Fourier transform of the utilized window peak is, the better the estimation performance.

Based on the above insight, we aim for finding windows whose Fourier transform has sharp mainlobe to advance the estimation performance. However, there are no other traditional windows whose Fourier transform has narrower mainlobe width than that of the rectangular window. As a result, we find other ways and creatively utilize the super-resolution estimation algorithm (SREA) to generate suitable windows.

Generally, we process $\boldsymbol{\Gamma}_m$ by replacing the traditional 2D-DFT by the SREA, which is reminiscent of the MUSIC algorithm. The outputs of the MUSIC algorithm are bell-shaped curves centered at specific coordinates. If the MUSIC
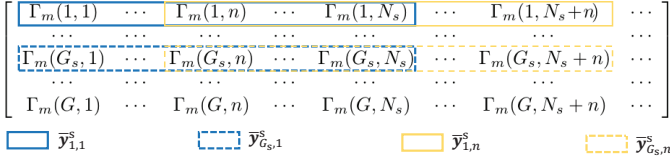
This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2024.3446610

IEEE TRANSACTIONS ON COMMUNICATIONS, ACCEPTED, AUG. 2024

9

$$\begin{bmatrix} \Gamma_m(1,1) & \cdots & \Gamma_m(1,n) & \cdots & \Gamma_m(1,N_s) & \cdots & \Gamma_m(1,N_s+n) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \Gamma_m(G_s,1) & \cdots & \Gamma_m(G_s,n) & \cdots & \Gamma_m(G_s,N_s) & \cdots & \Gamma_m(G_s,N_s+n) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \Gamma_m(G,1) & \cdots & \Gamma_m(G,n) & \cdots & \Gamma_m(G,N_s) & \cdots & \Gamma_m(G,N_s+n) & \cdots \end{bmatrix}$$

$\boxed{}\ \bar{\mathbf{y}}^{\mathrm{s}}_{1,1} \qquad \boxed{}\ \bar{\mathbf{y}}^{\mathrm{s}}_{G_s,1} \qquad \boxed{}\ \bar{\mathbf{y}}^{\mathrm{s}}_{1,n} \qquad \boxed{}\ \bar{\mathbf{y}}^{\mathrm{s}}_{G_s,n}$

Fig. 5. Frequency smoothing schemes. The matrix shown in this figure is $\mathbf{\Gamma}_m$. The smoothing vectors in the same color, such as $\bar{\mathbf{y}}^{\mathrm{s}}_{1,1}$ and $\bar{\mathbf{y}}^{\mathrm{s}}_{G_s,1}$, will be stacked into the same vector, $\bar{\mathbf{y}}_{1,1}$.

algorithm is regarded as a spectral analysis tool, just like the DFT, then the outputs of the MUSIC can be accordingly viewed as spectral components filtered by the "bell-shaped window"[4]. The process of using the MUSIC algorithm for $\mathbf{\Gamma}_m$ is very similar to that of utilizing the DFT for $\mathbf{\Gamma}_m$. Thus, the "bell-shaped window" is treated as a special window generated by the MUSIC algorithm. Given the generally sharp peak of the bell-shaped curve [47], the corresponding synchronization performance will be improved according to the analytical results in Section IV.

Concretely, to acquire a preferable "bell-shaped window", we first preprocess $\mathbf{\Gamma}_m$ by frequency-domain (FD) smoothing [44], [48] before utilizing the MUSIC algorithm. Specifically, we define $\bar{\mathbf{y}}^{\mathrm{s}}_{g,n}$ as the smoothing vector and its composition is shown in Fig. 5. Subsequently, we stack several smoothing vectors into $\bar{\mathbf{y}}_{g,n}$ as $[(\bar{\mathbf{y}}^{\mathrm{s}}_{g,n})^{\mathrm{T}}, \cdots, (\bar{\mathbf{y}}^{\mathrm{s}}_{g+G_s,n})^{\mathrm{T}}, \cdots, (\bar{\mathbf{y}}^{\mathrm{s}}_{g+(g_s-1)G_s,n})^{\mathrm{T}}]^{\mathrm{T}}$, where G_s is the index difference between two continuous vectors and g_{s} is the total number of smoothing vectors stacked into $\bar{\mathbf{y}}_{g,n}$. Moreover, N_s in Fig. 5 represents the length of the smoothing vectors.

After frequency-domain smoothing, we represent the covariance matrix

$$\mathbf{C} = \sum_{n=1}^{n_{\mathrm{s}}} \sum_{g=1}^{g_{\mathrm{s}}} \bar{\mathbf{y}}_{g,n} \bar{\mathbf{y}}^{\mathrm{H}}_{g,n}, \tag{34}$$

to implement the MUSIC algorithm. Then, by applying the eigenvalue decomposition to $\mathbf{C}$, we can obtain

$$\mathbf{C} = [\mathbf{U}_{\mathrm{s}}, \mathbf{U}_{\mathrm{n}}] \mathbf{\Lambda} [\mathbf{U}_{\mathrm{s}}, \mathbf{U}_{\mathrm{n}}]^{\mathrm{H}}, \tag{35}$$

where both $\mathbf{U}_{\mathrm{s}}$ and $\mathbf{U}_{\mathrm{n}}$ are the normalized unitary matrices and $\mathbf{\Psi}$ is a diagonal matrix, whose diagonal elements are the eigenvalues of $\mathbf{C}$. As a result, the metric composed by bell-shaped windows can be formulated as

$$\mathbf{B}_{\mathrm{s}}(\delta, \tau) = \frac{1}{\mathbf{a}^{\mathrm{H}}_{\delta,\tau} \mathbf{U}_{\mathrm{n}} \mathbf{U}^{\mathrm{H}}_{\mathrm{n}} \mathbf{a}_{\delta,\tau}}, \tag{36}$$

where we have $\mathbf{a}_{\delta,\tau} = \boldsymbol{\theta} \otimes \boldsymbol{\tau}$. Moreover, $\boldsymbol{\theta}$ and $\boldsymbol{\tau}$ are defined as $[1, \cdots, e^{-j2\pi g G_s \Delta f \tau}, \cdots, e^{-j2\pi g_{\mathrm{s}} G_s \Delta f \tau}]^{\mathrm{T}}$ and $[1, \cdots, e^{-j2\pi \delta(N_s-1)T_{\mathrm{sym}}}]^{\mathrm{T}}$, respectively. Since the center coordinates of the bell-shaped windows are related to δ and τ, we can estimate them by locating the peaks of the metric.

Compared with the conventional usage of the MUSIC algorithm, our scheme only exploits the comparatively sharp lobe of the output of the MUSIC algorithm to generate a superior

---

[4]The "bell-shaped window" here refers to the shape of output of the MUSIC algorithm when a single-tone signal is input. For convenience, we name this window as the bell-shaped window with the MUSIC algorithm (BS-W-MUSIC) in what follows.
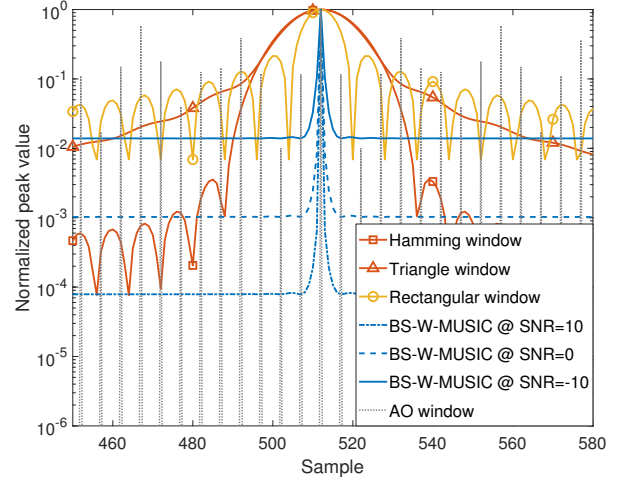

Fig. 6. Cross-correlation of the BS-W-MUSIC, Fourier transforms corresponding to different traditional window functions and the asymptotically optimal window. "AO window" represents correlation of our proposed asymptotically optimal window.

fingerprint spectrum derived from BS-W-MUSIC, rather than directly estimating the Doppler frequency shift and the time delay [44]. Then, by leveraging the fingerprint spectrum, the cross correlation in (20) is harnessed for estimating the CFO and TO. However, the conventional MUSIC algorithm is not directly applicable to CFO and TO estimation due to the associated clock asynchronism issue.

VI. NUMERICAL SIMULATIONS

In this section we perform numerical simulations for verifying the accuracy of our theoretical analysis and for evaluating the synchronization performance, when employing a bell-shaped window. In the simulations, we assume that the number of RRU and UE antennas is $M_{\mathrm{r}} = 64$ and $M_{\mathrm{t}} = 2$, respectively. The carrier frequency f_c of this system is assumed to be 28 GHz and the subcarrier spacing is set to 100 KHz, which is similar to [9], [24]. Moreover, we assume that the number of subcarriers is 128 and the length of CP is 16. Based on these parameters, the sampling interval can be simply calculated as 7.8125×10^{-8} s and the length of the whole OFDM symbol is 11.25 μs. The time length suggests that even though several hundreds of OFDM symbols are employed for sensing, it will take only a few milliseconds, corresponding to near-real-time perception.

In Fig. 6 we draw the auto-correlation sequence of the Fourier transforms of different windows, including the traditional windows, the bell-shaped window and the near-optimal window. Specifically, the length of the traditional windows is set to 128, which equals to the number of subcarriers. Moreover, a 25 times zero-padding is utilized, when applying the DFT to these windows, which means that 128×24 zeros will be appended to the windows before applying the DFT. For intuitive comparison, the bell-shaped window is generated by applying the MUSIC algorithm to a 128-subcarrier system. Furthermore, since the shape of peaks output by MUSIC is influenced by several factors, including the SNR and the
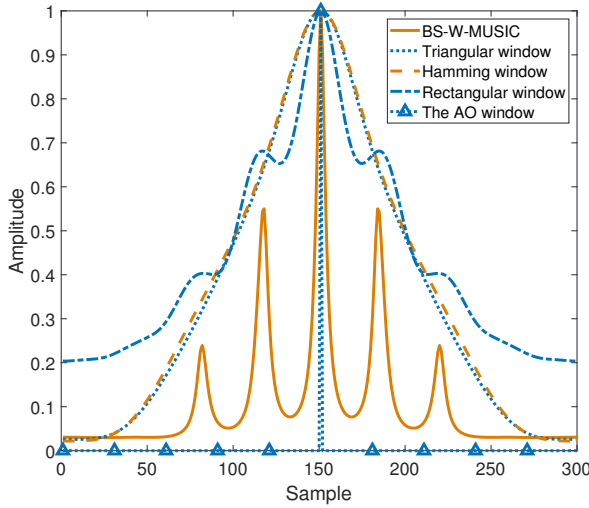
Fig. 7. The auto-correlation sequence of fingerprint spectra corresponding to different windows. The fingerprint spectra are generated by randomly setting 4 static objects in the environmental background. Moreover, the received signals reflected from different objects are randomly set as $[10, 4, 7, 1]$ dB and the propagation distances of these paths are randomly set as $[40, 50, 75, 110]$ meters. Moreover, the "The AO window" represents the auto-correlation sequence of fingerprint spectrum generated by the desired asymptotically optimal window.

number of antennas as well as snapshots [47], we evaluate the lobe width of bell-shaped window for diverse parameter combinations. Although the noise floor of the bell-shaped window varies when SNR and the number of snapshots increase, the width of the mainlobe of the window is invariably lower than that of all the traditional windows. Moreover, since the asymptotically optimal window shape is dependent on a specific scenario, we draw its correlation in the same scenarios as those in Fig. 7 for intuitive comparison. Observing in Figure 6 that both the mainlobe and the sidelobes associated with the 'AO window' are remarkably sharp, and the sidelobes are more prominent compared with those of the BS-W-MUSIC and conventional windows. Nevertheless, the amplitude of each sidelobe and the spacing between them in the 'AO window' have been carefully designed based on the gains and time delays of each path, as described in Proposition 2. This design allows for the cancellation of sidelobes, while preserving the mainlobe during the generation of the fingerprint spectrum (which involves the summation of multiple cyclic-shifted AO windows). As a result, despite the initially large sidelobes in the AO window, they disappear in the fingerprint spectrum. Conversely, for the BS-W-MUSIC and conventional windows, which have wider sidelobes, sidelobe cancellation can be hardly achieved. In such cases, the windows having sharper sidelobes may yield a fingerprint spectrum associated with narrower sidelobes, thereby leading to improved synchronization performance.

In Fig. 7, we draw the cross-correlation sequence of the fingerprint spectra generated by the traditional windows, bell-shaped windows and the asymptotically optimal window. As the figure shows, the fingerprint spectrum generated by the bell-shaped window exhibits sharper peaks than the mainlobe
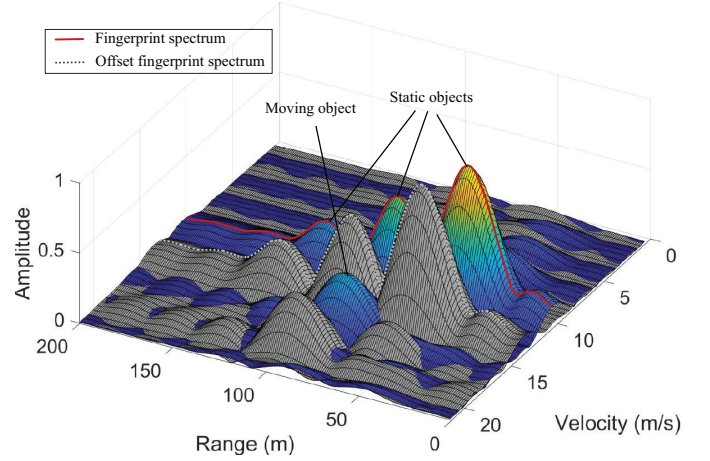
(a) The 3D spectrum corresponding to the rectangular window.
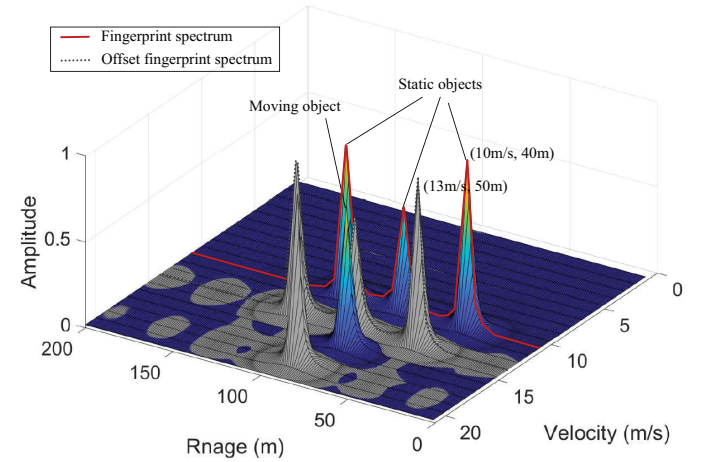
(b) The 3D spectrum corresponding to the proposed BS-W-MUSIC.

Fig. 8. The 3D spectrum implementations for multiple windows. The spectra are generated by randomly setting several static objects and dynamic object in the environmental with/without CFO and TO. Moreover, the spectrum corresponding to the rectangular window is generated by utilizing 25-fold zero-padding along the Range-axis and 5-fold zero-padding along the Velocity-axis. The blue 3D spectrum and the grey one represent the initial spectrum and the offset spectrum with CFO and TO drift. The set CFO and TO are equivalent to the Doppler frequency shift induced by a small velocity of 3 m/s, and the time delay resulting from a small distance of 10 m, respectively. The red outlines shown in both sub-figures are selected as the fingerprint spectrum sequence, while the dashed grey lines represent the offset fingerprint spectrum sequence.

of the traditional windows, while the peak of the asymptotically optimal window is the sharpest. Actually, the shape of "AO window" curve is formulated as (29). Moreover, to intuitively observe the shape of these fingerprint spectra, we also draw the 3D spectra shown in Fig. 8, and the fingerprint spectra utilized to draw Fig. 7 is actually selected from the row marked as red in Fig. 8. For brevity, only the fingerprint spectrum generated by the rectangular window is drawn in Fig. 8, due to its narrowest mainlobe among all the traditional windows. By observing all the different spectra in Fig. 8, it becomes evident that according to our analysis, the synchronization performance of the BS-W-MUSIC is better than that of traditional windows. Moreover, according to Fig. 8, the impact of CFO and TO on the fingerprint spectrum can be

This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2024.3446610

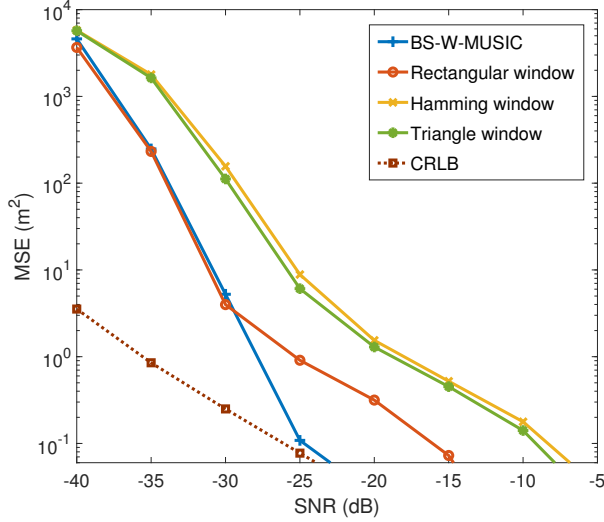IEEE TRANSACTIONS ON COMMUNICATIONS, ACCEPTED, AUG. 2024
11

Fig. 9. The synchronization MSE of different windows and the BS-W-MUSIC.
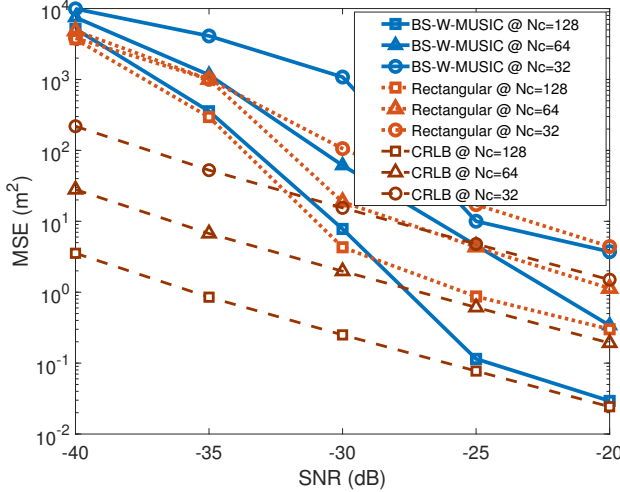
intuitively observed.

Fig. 10. The synchronization MSE of BS-W-MUSIC and the rectangular window while setting different number of subcarriers.

In Fig. 9, we evaluate the synchronization performance of different windows by locating the peaks of the autocorrelation sequence of the corresponding fingerprint spectrum using Monte Carlo simulations. We run thousands of Monte Carlo estimation trials to generate the synchronization MSE for different SNR, number and locations of static objects. Concretely, the average MSE is defined as

$$\text{MSE} = \frac{1}{N_\text{simu}} \sum_{n=1}^{N_\text{simu}} [R \times (\hat{l} - l_\text{R})]^2, \qquad (37)$$

where l_R and $\hat{l}$ are the ground truth and the estimated indices by using (20), respectively. Moreover, R is the propagation distance corresponding to the unit coordinate and N_simu represents the number of numerical simulation runs. In this

simulation, we employ 400 OFDM symbols each time for sensing and these symbols will approximately have a length of 4.5 ms in practical scenarios. This short time ensures the timeliness of perception. Additionally, to diversify the parameter configuration, we randomly select the number of static objects in the environment from the set of $\{2, 3, 4\}$. It is important to note that the vertical axis unit used to evaluate the peak estimation performance of the autocorrelation sequence of the fingerprint spectrum is square meters (m²). This choice is made because calculating the MSE as outlined in (23), based on the MSE of the range ambiguity induced by the TO estimation error, offers a more intuitive measure of the performance. Moreover, this unit applies to both Fig. 10 and Fig. 11.

As shown in Fig. 9, it is evident that the rectangular window has the best performance among the traditional windows, while the bell-shaped window outperforms the rectangular window. This phenomenon conforms to our insight. However, the performance of the bell-shaped window is better than that of the rectangular window, when the SNR is higher than -30 dB. This phenomenon appears, since the noise floor of the bell-shaped window increases with the SNR decreasing and consequently the peak location performance becomes worse than that of the rectangular window. Moreover, there is rapid drop in the curve for the proposed BS-W-MUSIC, when the SNR is in the range of $(-40, -30)$ dB. The drop emerges since the peak changes from inconspicuous at -40 dB to obvious at -30 dB, and that change makes the peak location more accurate. However, when the SNR becomes higher, the increase in SNR results in a diminishing enhancement of the positioning accuracy.

In Fig. 10, we compare the synchronization performance of the rectangular window and of the BS-W-MUSIC for different numbers of subcarriers. A pair of insights may be drawn from Fig. 10. The first is: both the performance of the rectangular window and of the bell-shaped window gets better with the number of subcarriers increasing, which is intuitive. As the CRLB we derived in [39] suggests: the larger the number of subcarriers, the lower the CRLB. Concretely, the CRLB matrix is formulted as $\text{CRLB} = \sigma^2 \left(\sum_{n=1}^{N_c} \mathbf{H}_n^\text{H} \mathbf{H}_n \right)^{-1}$, where $\mathbf{H}_n = [\frac{\partial \mathbf{\Gamma}_m}{\partial v_1}, \cdots, \frac{\partial \mathbf{\Gamma}_m}{\partial v_l}, \frac{\partial \mathbf{\Gamma}_m}{\partial \tau_1}, \cdots, \frac{\partial \mathbf{\Gamma}_m}{\partial \tau_l}]$. The second insight is: regardless of the specific subcarrier configuration, the performance of the bell-shaped window is better than that of the rectangular window, when the SNR achieves critical point, like -35dB for $N_c = 128$. However, the point will generally move to the left with the number of subcarriers increasing, because the MUSIC algorithms performs better, when the length of the steering vector is higher.

In Fig. 11, we evaluate the theoretical MSE of the asymptotically optimal window and compare it with that of the BS-W-MUSIC and the CACC synchronization method using numerical simulations, respectively. As Fig. 11 shows, for all the three methods, the MSE improves as the number of subcarriers increases. Moreover, the MSE drops rapidly when the SNR exceeds -10 dB for the asymptotically optimal window. However, under low-SNR conditions, the MSE of peak location degrades severely despite the impulse function's ex-
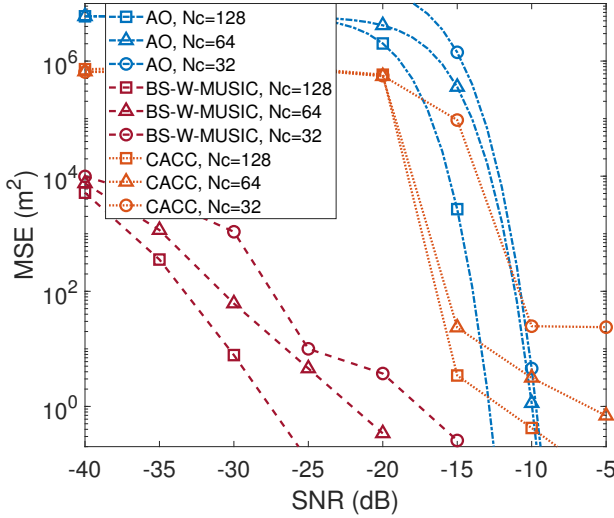
Fig. 11. The estimation MSE of the asymptotically optimal window, the BS-W-MUSIC, and the CACC synchronization method.

cellent peak location performance under high-SNR scenarios. Interestingly, the BS-W-MUSIC and CACC methods outperform the method of using the asymptotically optimal window, suggesting that a narrower mainlobe is not always better in low-SNR scenarios and wider mainlobe windows may perform better. Furthermore, the MSE of our BS-W-MUSIC method is lower than that of the CACC method. When considering Fig. 9 alongside this observation, it becomes apparent that the MSE performance of CACC is roughly equivalent to that of using the Hamming window. Another notable observation in Fig. 11 is that the rate of MSE decrease associated with the asymptotically optimal window is significantly faster than that of the BS-W-MUSIC and the CACC methods. This is due to the design of the asymptotically optimal window relying on the exact gains and delays of each propagation path, allowing for more known information to estimate the CFO and TO than when using traditional windows. However, it is important to note that it is infeasible to achieve the theoretical lower bound of the MSE for the asymptotically optimal window in practical systems.

VII. CONCLUSIONS

Window functions designed have been investigated for fingerprint-spectrum-based passive sensing tailored to PMNs. Concretely, we derived the expressions of an asymptotically optimal window and of the corresponding synchronization MSE. However, the asymptotically optimal window cannot be practically harnessed since it is strongly dependent on the specific time delays and gains corresponding to each propagation path, which are unknown before successful sensing. To circumvent this issue, we tested a practical bell-shaped window relying on super-resolution estimation algorithms and simulations have verified our theoretical analysis. To some degree, this work reveals the relationship between the synchronization performance and the window function utilized in the PMN systems, where sensing is achieved by spectrum

analysis. Future improvements of this work may be conducted by finding the exact globally optimal window or proposing a closed-form metric to accurately evaluate the synchronization performance of hundreds of windows.

APPENDIX A
PROOF OF PROPOSITION 1

According to (23), to minimize $f_{\mathrm{MSE}}(\boldsymbol{\psi}_{N_c})$, the decision variables, namely $\mathrm{P}_{\{q,KN_c\}}$ for $q = 1, \cdots, KN_c$, should be optimized. Based on the definition of $\mathrm{P}_{\{q,KN_c\}}$, namely (24), $\mathrm{P}_{\{q,KN_c\}}$ is only determined by the fingerprint spectrum, namely the K_1th row of $\mathbf{A}[K_1, q]$. Moreover, $\mathbf{A}[K_1, q]$ can be deemed to be a complex Gaussian variable with a mean of $\sum_{p=1}^{KN_c} \boldsymbol{\beta}[p]\boldsymbol{\beta}_{\mathrm{c}}[p+q]$ for $q = 1, \cdots, KN_c$ and variance of (refer to Appendix B)

$$\bar{\sigma}^2 \approx \frac{2\tilde{\sigma}^2}{|\boldsymbol{\beta}_{\mathrm{n}}^*|^2} + \frac{KN_c}{2|\boldsymbol{\beta}_{\mathrm{n}}^*|^4}\tilde{\sigma}^2. \tag{38}$$

Equation (38) suggests that if the noise power of the received signal is given, $\tilde{\sigma}^2$ will be fixed. Consequently, $\bar{\sigma}^2$ will be determined by the power of the fingerprint spectrum, $|\boldsymbol{\beta}_{\mathrm{n}}^*|^2$. Moreover, since the maximum value of $\mathbf{s}$ is normalized to 1, $|\boldsymbol{\beta}_{\mathrm{n}}^*|^2$ approximately equals the maximum element of $|\boldsymbol{\beta}_{\mathrm{n}}^*|^2\mathbf{s}$. Thus, the minimized MSE can be determined by finding the optimal $\mathbf{s}$. Specifically, we determine the optimal $\mathbf{s}$ in two steps.

Step 1: By assuming that $\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}$ is an integer, we first clarify that for any value of $\mathbf{s}(q)$ for $q \neq \frac{\Delta\delta^\tau}{T_{\mathrm{R}}}$, $f_{\mathrm{MSE}}(\boldsymbol{\psi}_{N_c})$ is minimized when $|\boldsymbol{\beta}_{\mathrm{n}}^*|^2\mathbf{s}[\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}]$ is maximized. Note that $\mathbf{s}[\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}] = 1$ in this case, according to the definition of $\mathbf{s}$.

Step 1.1: According to (38), the normalized noise power $\bar{\sigma}^2$ decreases upon increasing $|\boldsymbol{\beta}_{\mathrm{n}}^*|^2\mathbf{s}[\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}]$ for a fixed $\tilde{\sigma}^2$. Then, the probability $\mathrm{P}_{\{q,KN_c\}}$ for $q \neq \frac{\Delta\delta^\tau}{T_{\mathrm{R}}}$ will decrease, while $\mathrm{P}_{\{\Delta\delta^\tau/T_{\mathrm{R}},KN_c\}}$ will increase. To prove this, we represent $\mathrm{P}_{\{q,KN_c\}}$ as

$$\mathrm{P}_{\{q,KN_c\}} = \int_V p(\mathbf{s}, \tilde{\sigma}^2)\mathrm{d}\,\mathbf{b}$$
$$= \int_{-\infty}^{\infty} \frac{e^{-[b_q-\mathbf{s}[q]]^2/2\bar{\sigma}^2}}{\sqrt{2\pi}\bar{\sigma}} \left(\prod_{i=1,i\neq q}^{KN_c} \int_{-\infty}^{b_q} \frac{e^{-[b_i-\mathbf{s}[i]]^2/2\bar{\sigma}^2}}{\sqrt{2\pi}\bar{\sigma}}\mathrm{d}b_i\right)\mathrm{d}b_q$$
$$= \int_{-\infty}^{\infty} f_{\mathcal{N}}(b_q) \prod_{i=1,i\neq q}^{KN_c} Q(\frac{b_q-\mathbf{s}[i]}{\bar{\sigma}^2})\mathrm{d}\,b_q, \tag{39}$$

where $f_{\mathcal{N}}(\cdot)$ represents the probability density function of the standard normal distribution. From the formulation, it is observed that $\mathrm{P}_{\{q,KN_c\}}$ could be expressed as the integral of $\prod_{i=1,i\neq q}^{KN_c} Q(\frac{b_q-\mathbf{s}[i]}{\bar{\sigma}^2})$. For any value of $\mathbf{s}[q]$, $q \neq \frac{\Delta\delta^\tau}{T_{\mathrm{R}}}$, $\prod_{i=1,i\neq q}^{KN_c} Q(\frac{b_q-\mathbf{s}[i]}{\bar{\sigma}^2})$ will reduce when $|\boldsymbol{\beta}_{\mathrm{n}}^*|^2\mathbf{s}[\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}]$ increases and as a result, $\mathrm{P}_{\{q,KN_c\}}$ reduces. By contrast, for $q = \frac{\Delta\delta^\tau}{T_{\mathrm{R}}}$, $\prod_{i=1,i\neq q}^{KN_c} Q(\frac{b_q-\mathbf{s}[i]}{\bar{\sigma}^2})$ increases when $|\boldsymbol{\beta}_{\mathrm{n}}^*|^2\mathbf{s}[\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}]$ increases and consequently $\mathrm{P}_{\{\Delta\delta^\tau/T_{\mathrm{R}},KN_c\}}$ increases. Thus, $|\boldsymbol{\beta}_{\mathrm{n}}^*|^2\mathbf{s}[\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}]$ should be set as large as possible to minimize $f_{\mathrm{MSE}}(\boldsymbol{\psi}_{N_c})$.

Step 1.2: Upon confirming the optimal value of $|\boldsymbol{\beta}_{\mathrm{n}}{}^*|^2 \mathbf{s}[\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}]$, we then determine the optimal $\mathbf{s}[q]$, $q \neq \frac{\Delta\delta^\tau}{T_{\mathrm{R}}}$. Observe from (23), that $f_{\mathrm{MSE}}(\boldsymbol{\psi}_{N_c})$ can be deemed to be a linear weighted sum of elements in $\{(1 - \frac{\delta^\tau}{T_{\mathrm{R}}})^2, \cdots, (KN_c - \frac{\delta^\tau}{T_{\mathrm{R}}})^2\}$ and the weights are $\{\mathrm{P}_{\{1,KN_c\}}, \cdots, \mathrm{P}_{\{KN_c,KN_c\}}\}$. Moreover, it also holds that $\sum_{q=1}^{KN_c} \mathrm{P}_{\{q,KN_c\}} = 1$. As a result, to minimize $f_{\mathrm{MSE}}(\boldsymbol{\psi}_{N_c})$, we should assign as much weight as possible to the index closest to $\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}$ and as low weight as possible to indices far from $\frac{\Delta\delta^\tau}{T_{\mathrm{R}}}$. However, it is difficult to design the optimal $\mathbf{s}$ to minimize $f_{\mathrm{MSE}}(\boldsymbol{\psi}_{N_c})$ utilizing the complex integral of the Q function.

Step 2: Consequently, to find superior window functions, we aim for finding an approximate optimal solution, when the SNR tends to infinity. Specifically, we define $\mathbf{s}_{\mathrm{op}}$ as the undetermined optimal fingerprint spectrum and denote the approximate optimal fingerprint spectrum by $\mathbf{s}_{\mathrm{ap}}$. In the following, we prove that *when $\mathbf{s}_{\mathrm{ap}}$ satisfies*

$$\begin{cases} \mathbf{s}_{\mathrm{ap}}[q] = 1, q = \frac{\Delta\delta^\tau}{T_{\mathrm{R}}}, \\ \mathbf{s}_{\mathrm{ap}}[q] = 0, q \neq \frac{\Delta\delta^\tau}{T_{\mathrm{R}}}, \end{cases} \quad (40)$$

the MSE tends to approach that of the optimal fingerprint spectrum. Let us denote the probability corresponding to $\mathbf{s}_{\mathrm{ap}}$ and $\mathbf{s}_{\mathrm{op}}$ by $\mathrm{P}_{\{q,KN_c\}}^{\mathrm{op}}$ and $\mathrm{P}_{\{q,KN_c\}}^{\mathrm{ap}}$, respectively. Then, we have

$$f_{\mathrm{MSE}}(\mathbf{s}_{\mathrm{ap}}) - f_{\mathrm{MSE}}(\mathbf{s}_{\mathrm{op}}) = \sum_{q=1}^{KN_c} \Big[\mathrm{P}_{\{q,KN_c\}}^{\mathrm{ap}} - \mathrm{P}_{\{q,KN_c\}}^{\mathrm{op}} \Big](q - \frac{\Delta_\delta}{T_{\mathrm{R}}})^2, \quad (41)$$

where $\mathrm{P}_{\{q,KN_c\}}^{\mathrm{ap}} - \mathrm{P}_{\{q,KN_c\}}^{\mathrm{op}}$ can be represented as

$$\mathrm{P}_{\{q,KN_c\}}^{\mathrm{ap}} - \mathrm{P}_{\{q,KN_c\}}^{\mathrm{op}} = \int_{-\infty}^{\infty} f_{\mathcal{N}}(b_q) \Bigg[\prod_{i=1,i\neq q}^{KN_c} Q(\frac{b_q - \mathbf{s}_{\mathrm{ap}}(i)}{\bar{\sigma}^2}) - \prod_{i=1,i\neq q}^{KN_c} Q(\frac{b_q - \mathbf{s}_{\mathrm{op}}(i)}{\bar{\sigma}^2}) \Bigg] \mathrm{d}b_q. \quad (42)$$

With the SNR of the PMN system, namely $1/\tilde{\sigma}^2$, increasing, both $Q(\frac{b_q - \mathbf{s}_{\mathrm{ap}}(i)}{\bar{\sigma}^2})$ and $Q(\frac{b_q - \mathbf{s}_{\mathrm{op}}(i)}{\bar{\sigma}^2})$ decrease. Moreover, given any positive number ϵ, there exists an SNR, defined as $1/\tilde{\sigma}_\epsilon^2$, which satisfies

$$\Big| \prod_{i=1,i\neq q}^{KN_c} Q(\frac{b_q - \mathbf{s}_{\mathrm{ap}}(i)}{\bar{\sigma}^2}) - \prod_{i=1,i\neq q}^{KN_c} Q(\frac{b_q - \mathbf{s}_{\mathrm{op}}(i)}{\bar{\sigma}^2}) \Big| < \epsilon, \quad (43)$$

for any $q = 1, \cdots, KN_c$.

Thus, $\mathbf{s}_{\mathrm{ap}}$ represented as (40) is the near-optimal solution, when $\Delta\delta^\tau/T_{\mathrm{R}}$ is an integer. Moreover, to characterize the synchronization MSE in this case, we can express $\mathrm{P}_{\{q,KN_c\}}$ for $q \in \{1, \cdots, KN_c\}$ and $q \neq \Delta\delta^\tau/T_{\mathrm{R}}$ as

$$\mathrm{P}_{\{q,KN_c\}} = \int_{-\infty}^{\infty} f_{\mathcal{N}}(b_q) Q(\frac{b_q - 1}{\bar{\sigma}^2}) Q^{(KN_c - 2)}(b_q/\bar{\sigma}^2) \mathrm{d}b_q. \quad (44)$$

Furthermore, the MSE in this case can be formulated as

$$f_{\mathrm{MSE}}(\boldsymbol{\psi}_{N_c}) = \sum_{\substack{q=1, \\ q\neq\Delta\delta^\tau/T_{\mathrm{R}}}}^{KN_c} \Bigg[q_{\mathrm{r}} - \frac{\Delta\delta^\tau}{T_{\mathrm{R}}} \Bigg]^2 \int_{-\infty}^{\infty} \Big[f_{\mathcal{N}}(b_q) Q(\frac{b_q - 1}{\bar{\sigma}^2}) Q^{(KN_c - 2)}(b_q/\bar{\sigma}^2) \mathrm{d} \Big] b_q. \quad (45)$$

However, since the $Q(\cdot)$ function is not in closed form, the infimum cannot be transformed into closed form.

APPENDIX B
PROOF OF THE NOISE $\bar{\sigma}^2$

According to (20), $\mathbf{A}[K_1, q]$ can be represented as

$$\mathbf{A}[K_1, q] = \frac{1}{|\boldsymbol{\beta}_{\mathrm{n}}{}^*|^2} \sum_{p=1}^{KN_c} (\boldsymbol{\beta}[p] + \tilde{\mathbf{W}}_K[K_0, p])(\boldsymbol{\beta}_{\mathrm{c}}[p + q]$$

$$+ \tilde{\mathbf{W}}_K[K_0, p + q]) = \frac{1}{|\boldsymbol{\beta}_{\mathrm{n}}{}^*|^2} \sum_{p=1}^{KN_c} (\boldsymbol{\beta}[p]\boldsymbol{\beta}_{\mathrm{c}}[p + q] \quad (46)$$

$$+ \boldsymbol{\beta}_{\mathrm{c}}[p + q]\tilde{\mathbf{W}}_K[K_0, p] + \boldsymbol{\beta}[p]\tilde{\mathbf{W}}_K[K_0, p + q]$$

$$+ \tilde{\mathbf{W}}_K[K_0, p]\tilde{\mathbf{W}}_K[K_0, p + q]).$$

It is plansible that the terms $\sum_{p=1}^{KN_c} \boldsymbol{\beta}[p]\boldsymbol{\beta}_{\mathrm{c}}[p + q]$, $\sum_{p=1}^{KN_c} \boldsymbol{\beta}_{\mathrm{c}}[p + q]\tilde{\mathbf{W}}_K[K_0, p]$, $\sum_{p=1}^{KN_c} \boldsymbol{\beta}[p]\tilde{\mathbf{W}}_K[K_0, p + q]$, and $\sum_{p=1}^{KN_c} \tilde{\mathbf{W}}_K[K_0, p]\tilde{\mathbf{W}}_K[K_0, p + q]$ are all independent. As a result, $\mathbf{A}[K_1, q]$ can be regarded as following a Gaussian distribution having the mean of

$$\mathrm{E}(\mathbf{A}[K_1, q]) = \frac{1}{|\boldsymbol{\beta}_{\mathrm{n}}{}^*|^2} \sum_{p=1}^{KN_c} \boldsymbol{\beta}[p]\boldsymbol{\beta}_{\mathrm{c}}[p + q]. \quad (47)$$

Moreover, the variance of the distribution is

$$\mathrm{D}(\mathbf{A}[K_1, q]) = \frac{1}{|\boldsymbol{\beta}_{\mathrm{n}}{}^*|^4} \Bigg[\mathrm{D}(\sum_{p=1}^{KN_c} \boldsymbol{\beta}_{\mathrm{c}}[p + q]\tilde{\mathbf{W}}_K[K_0, p]) +$$

$$\mathrm{D}(\sum_{p=1}^{KN_c} \boldsymbol{\beta}[p]\tilde{\mathbf{W}}_K[K_0, p + q]) + \mathrm{D}(\sum_{p=1}^{KN_c} \tilde{\mathbf{W}}_K[K_0, p]\tilde{\mathbf{W}}_K[K_0, p+q]) \Bigg]$$

$$= \frac{1}{|\boldsymbol{\beta}_{\mathrm{n}}{}^*|^4} \Bigg[|\boldsymbol{\beta}_{\mathrm{c}}{}^*|^2 \tilde{\sigma}^2 + |\boldsymbol{\beta}^*|^2 \tilde{\sigma}^2 + \frac{KN_c}{2}\tilde{\sigma}^2 \Bigg] \quad (48)$$

$$\approx \frac{2\tilde{\sigma}^2}{|\boldsymbol{\beta}_{\mathrm{n}}{}^*|^2} + \frac{KN_c}{2|\boldsymbol{\beta}_{\mathrm{n}}{}^*|^4}\tilde{\sigma}^2.$$

APPENDIX C
PROOF OF PROPOSITION 2

Reconsider $\mathbf{s}$ and re-formulate it as

$$\mathbf{s} = \boldsymbol{\phi}_m \mathbf{R}, \quad (49)$$

where $\boldsymbol{\phi}_m$ is represented by

$$\boldsymbol{\phi}_m = [\boldsymbol{\alpha}_1[m]\boldsymbol{\alpha}_1[m], \cdots, \boldsymbol{\alpha}_l[m]\boldsymbol{\alpha}_{l'}[m], \cdots, \boldsymbol{\alpha}_L[m]\boldsymbol{\alpha}_L[m]], \quad (50)$$

and $\mathbf{R} = [\boldsymbol{\rho}_{1,1}^{\mathrm{T}}, \cdots, \boldsymbol{\rho}_{l,l'}^{\mathrm{T}}, \cdots, \boldsymbol{\rho}_{L,L}^{\mathrm{T}}]^{\mathrm{T}}$ Moreover, as we described before, when KN_c is large enough, $\boldsymbol{\rho}_{l,l'}$ is approximately a circular-shifted version of $\boldsymbol{\rho}_{1,1}$ for any l and l', which can be formulated as

$$\boldsymbol{\rho}_{l,l'}[q] = \boldsymbol{\rho}_{1,1}[q \oplus \frac{\Delta_{l,l'}}{T_{\mathrm{sam}}}], \ \ q = 1, \cdots, KN_c. \quad (51)$$

Then, we rearrange the rows of the matrix $\mathbf{R}$ and the corresponding columns of $\boldsymbol{\phi}_m$ in the ascending order of $\frac{\Delta_{l,l'}}{T_{\mathrm{sam}}}$, while we define the rearranged matrix as $\tilde{\mathbf{R}}$ and $\tilde{\boldsymbol{\phi}}_m$. Thus, $\boldsymbol{\rho}_{1,1}$ should satisfy the following condition

$$\tilde{\boldsymbol{\phi}}_m \tilde{\mathbf{R}} = \mathbf{s}. \quad (52)$$

This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2024.3446610

IEEE TRANSACTIONS ON COMMUNICATIONS, ACCEPTED, AUG. 2024 14

To find the specific $\boldsymbol{\rho}_{1,1}$ satisfying (52), we first consider the structure of $\tilde{\mathbf{R}}$. Since each row of $\tilde{\mathbf{R}}$ is a right-circular-shifted version of $\boldsymbol{\rho}_{1,1}$, we can denote the number of right cyclic shifts corresponding to the nth row of $\tilde{\mathbf{R}}$ as p_n for $n = 1, \cdots, L^2$. Then, we define $\boldsymbol{\Lambda}_q$ as the location indication matrix of the qth element of $\boldsymbol{\rho}_{1,1}$ in $\tilde{\mathbf{R}}$ and $q = 1, \cdots, KN_c$. Concretely, $\boldsymbol{\Lambda}_q$ belongs to $\mathbb{C}^{L^2 \times KN_c}$ and all of its $(n+q, p_n)$th elements for $n = 0, \cdots, L^2 - 1$ are set as 1, while the other elements are set as 0. Furthermore, according to the definition, we have the following equation

$$\boldsymbol{\Lambda}_n = \boldsymbol{\Lambda}_{n-1}\mathbf{J}, \text{ for } n = 2, \cdots, KN_c, \tag{53}$$

where $\mathbf{J}$ is the circulant matrix:

$$\mathbf{J} = \begin{pmatrix} \mathbf{0}_{KN_c-1} & \mathbf{I}_{KN_c-1} \\ 1 & \mathbf{0}_{KN_c-1} \end{pmatrix}. \tag{54}$$

As a result, (52) can be reformulated as

$$\tilde{\boldsymbol{\phi}}_m\tilde{\mathbf{R}} = \tilde{\boldsymbol{\phi}}_m\boldsymbol{\Lambda}_1 \sum_{n=1}^{KN_c} \boldsymbol{\rho}_{1,1}(n)\mathbf{J}^{n-1}, \tag{55}$$

utilizing (53) and (54). Let us denote $\tilde{\boldsymbol{\phi}}_m\boldsymbol{\Lambda}_1$ as $\breve{\boldsymbol{\phi}}_m$. Then, the right-hand-side of (55) can be viewed as the circular convolution between $\breve{\boldsymbol{\phi}}_m$ and $\boldsymbol{\rho}_{1,1}$, which is $\breve{\boldsymbol{\phi}}_m \circledast \boldsymbol{\rho}_{1,1}$. Furthermore, the circular convolution can be reformulated as

$$\tilde{\boldsymbol{\phi}}_m\tilde{\mathbf{R}} = \breve{\boldsymbol{\phi}}_m \circledast \boldsymbol{\rho}_{1,1} = \boldsymbol{\rho}_{1,1} \circledast \breve{\boldsymbol{\phi}}_m = \boldsymbol{\rho}_{1,1} \sum_{n=1}^{KN_c} \breve{\boldsymbol{\phi}}_m[n]\mathbf{J}^{n-1} = \mathbf{s}. \tag{56}$$

For brevity, let us define $\breve{\mathbf{J}} = \sum_{n=1}^{KN_c} \breve{\boldsymbol{\phi}}_m[n]\mathbf{J}^{n-1}$, which is a circulant matrix. According to [49], a circulant matrix can be diagonalized by the discrete Fourier transform matrix and thus, $\breve{\mathbf{J}}$ can be represented as

$$\breve{\mathbf{J}} = \mathbf{F}\mathrm{diag}(\breve{\boldsymbol{\phi}}_{\mathrm{F}})\mathbf{F}^{\mathrm{H}}, \tag{57}$$

where $\breve{\boldsymbol{\phi}}_{\mathrm{F}} = \breve{\boldsymbol{\phi}}_m\mathbf{F}$. However, due to the randomness of $\breve{\boldsymbol{\phi}}_m$, $\breve{\boldsymbol{\phi}}_{\mathrm{F}}$ may have zero elements, implying that we cannot tell, whether $\breve{\mathbf{J}}$ is a full-rank matrix. As a result, we discuss the value of $\boldsymbol{\rho}_{1,1}$ in terms of whether $\breve{\mathbf{J}}$ is of full rank.

1) If $\breve{\mathbf{J}}$ is of full rank, we can simply write $\boldsymbol{\rho}_{1,1}$ as

$$\boldsymbol{\rho}_{1,1} = \mathbf{s}\breve{\mathbf{J}}^{-1}. \tag{58}$$

2) If $\breve{\mathbf{J}}$ is not full rank, we can obtain $\boldsymbol{\rho}_{1,1}$ by solving:

$$\boldsymbol{\rho}_{1,1}\mathbf{F}\mathrm{diag}(\breve{\boldsymbol{\phi}}_{\mathrm{F}})\mathbf{F}^{\mathrm{H}} = \mathbf{s}. \tag{59}$$

After some simple derivations, $\boldsymbol{\rho}_{1,1}$ can be expressed as

$$\boldsymbol{\rho}_{1,1} = \mathbf{s}\mathbf{F}\mathrm{diag}(\breve{\boldsymbol{\phi}}_{\mathrm{F}}^{\mathrm{Inv}})\mathbf{F}^{\mathrm{H}}, \tag{60}$$

where $\breve{\boldsymbol{\phi}}_{\mathrm{F}}^{\mathrm{Inv}}$ is defined as

$$\begin{cases} \breve{\boldsymbol{\phi}}_{\mathrm{F}}^{\mathrm{Inv}}[n]\breve{\boldsymbol{\phi}}_{\mathrm{F}}[n] = 1, \text{ for } \breve{\boldsymbol{\phi}}_{\mathrm{F}}[n] \neq 0, \\ \breve{\boldsymbol{\phi}}_{\mathrm{F}}^{\mathrm{Inv}}[n] = 0, \text{ for } \breve{\boldsymbol{\phi}}_{\mathrm{F}}[n] = 0. \end{cases} \tag{61}$$

Moreover, the full rank $\breve{\mathbf{J}}$ also applies to (61). Therefore, by defining $\breve{\mathbf{J}}^{\mathrm{Inv}}$ as

$$\breve{\mathbf{J}}^{\mathrm{Inv}} = \mathbf{F}\mathrm{diag}(\breve{\boldsymbol{\phi}}_{\mathrm{F}}^{\mathrm{Inv}})\mathbf{F}^{\mathrm{H}}, \tag{62}$$

we can finally represent $\boldsymbol{\rho}_{1,1}$ as

$$\boldsymbol{\rho}_{1,1} = \mathbf{s}\breve{\mathbf{J}}^{\mathrm{Inv}}, \tag{63}$$

regardless whether $\breve{\mathbf{J}}$ is of full rank or not.

APPENDIX D
PROOF OF THE PROPOSITION 3

Since the autocorrelation sequence of a discrete sequence is equivalent to the convolution of the sequence and its reversed version [50], we can reformulate $\boldsymbol{\rho}_{1,1}$ in (22) as

$$\boldsymbol{\rho}_{1,1} = [\bar{\boldsymbol{\psi}}_{K,N_c}\mathbf{F}_{KN_c}] \circledast \mathrm{Rev}[\bar{\boldsymbol{\psi}}_{K,N_c}\mathbf{F}_{KN_c}], \tag{64}$$

where $\mathrm{Rev}[\cdot]$ is the operator that inverts a discrete sequence.

By substituting (64) into (63), we arrive at

$$[\bar{\boldsymbol{\psi}}_{K,N_c}\mathbf{F}_{KN_c}] \circledast \mathrm{Rev}[\bar{\boldsymbol{\psi}}_{K,N_c}\mathbf{F}_{KN_c}] = \mathbf{s}\breve{\mathbf{J}}_{\mathrm{Inv}}. \tag{65}$$

Moreover, by utilizing the anti-diagonal matrix $\mathbf{E} \in \mathbb{C}^{KN_c \times KN_c}$ defined as

$$\mathbf{E} = \begin{bmatrix} 0 & \cdots & 1 \\ \vdots & 1 & \vdots \\ 1 & \cdots & 0 \end{bmatrix}, \tag{66}$$

we can further represent (65) as

$$[\bar{\boldsymbol{\psi}}_{K,N_c}\mathbf{F}_{KN_c}] \circledast [\bar{\boldsymbol{\psi}}_{K,N_c}\mathbf{F}_{KN_c}\mathbf{E}] = \mathbf{s}\breve{\mathbf{J}}_{\mathrm{Inv}}. \tag{67}$$

According to the time-domain convolution theorem [51], (67) can be converted into

$$\bar{\boldsymbol{\psi}}_{K,N_c} \odot [\bar{\boldsymbol{\psi}}_{K,N_c}\mathbf{F}_{KN_c}\mathbf{E}\mathbf{F}_{KN_c}^{\mathrm{H}}] = \mathbf{s}\breve{\mathbf{J}}_{\mathrm{Inv}}. \tag{68}$$

Moreover, the elements of $\mathbf{F}_{KN_c}\mathbf{E}\mathbf{F}_{KN_c}^{\mathrm{H}}$ are all constants for a constant K and we denote the matrix as $\bar{\mathbf{E}}$. Therefore, to ensure the validity of (68), the following conditions should be satisfied

$$\begin{cases} \bar{\boldsymbol{\psi}}_{K,N_c}[i]\sum_{j=1}^{N_c} \bar{\boldsymbol{\psi}}_{K,N_c}[j]\bar{\mathbf{E}}[i,j] = \boldsymbol{\rho}_{1,1}[i], \text{ for } i = 1, \cdots, N_c, \\ \boldsymbol{\rho}_{1,1}[i] = 0, \text{ for } i = N_c + 1, \cdots, KN_c, \end{cases} \tag{69}$$

which means that there are solutions for $\bar{\boldsymbol{\psi}}_{K,N_c}$ to ensure the validity of (28) when $\boldsymbol{\rho}_{1,1}[i] = 0$ for $i = N_c+1, \cdots, KN_c$ and meanwhile the solutions equal $\bar{\boldsymbol{\psi}}_{K,N_c}[i] \sum_{j=1}^{N_c} \bar{\boldsymbol{\psi}}_{K,N_c}[j]\bar{\mathbf{E}}[i,j]$ for $i = 1, \cdots, N_c$.

REFERENCES

[1] X.-Y. Wang, S. Yang, H.-Y. Zhai, C. Masouros, and J. A. Zhang, "Optimizing fingerprint-spectrum-based synchronization in integrated sensing and communications," in *Proc. 2024 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2024, accepted.

[2] X. Wang, L. Kong, F. Kong, F. Qiu, M. Xia, S. Arnon, and G. Chen, "Millimeter wave communication: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 1616–1653, Third Quarter. 2018.

[3] C. Han, A. O. Bicen, and I. F. Akyildiz, "Multi-ray channel modeling and wideband characterization for wireless communications in the terahertz band," *IEEE Transactions on Wireless Communications*, vol. 14, no. 5, pp. 2402–2412, May. 2015.

[4] K. Meng, Q. Wu, S. Ma, W. Chen, K. Wang, and J. Li, "Throughput maximization for UAV-enabled integrated periodic sensing and communication," *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 671–687, Jan. 2023.

This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2024.3446610

IEEE TRANSACTIONS ON COMMUNICATIONS, ACCEPTED, AUG. 2024
15

[5] 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.

[6] X. Li, Y. Cui, J. A. Zhang, F. Liu, D. Zhang, and L. Hanzo, "Integrated human activity sensing and communications," *IEEE Communications Magazine*, vol. 61, no. 5, pp. 90–96, May. 2023.

[7] Y. Ge, O. Kaltiokallio, H. Kim, F. Jiang, J. Talvitie, M. Valkama, L. Svensson, S. Kim, and H. Wymeersch, "A computationally efficient EK-PMBM filter for bistatic mmwave radio SLAM," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 7, pp. 2179–2192, Jul. 2022.

[8] 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.

[9] M. L. Rahman, J. A. Zhang, X. Huang, Y. J. Guo, and R. W. Heath, "Framework for a perceptive mobile network using joint communication and radar sensing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 3, pp. 1926–1941, Sep. 2019.

[10] J. A. Zhang, M. L. Rahman, K. Wu, X. Huang, Y. J. Guo, S. Chen, and J. Yuan, "Enabling joint communication and radar sensing in mobile networks–A survey," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 306–345, First Quarter. 2021.

[11] J. A. Zhang, M. L. Rahman, X. Huang, Y. J. Guo, S. Chen, and R. W. Heath, "Perceptive mobile networks: Cellular networks with radio vision via joint communication and radar sensing," *IEEE Vehicular Technology Magazine*, vol. 16, no. 2, pp. 20–30, Jun. 2020.

[12] 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.

[13] Y. Xiong, F. Liu, C. Yuanhao, W. Yuan, T. Han, and G. Caire, "On the fundamental tradeoff of integrated sensing and communications under gaussian channels," *IEEE Transactions on Information Theory*, vol. 69, no. 9, pp. 5723–5751, Sep. 2023.

[14] X. Tong, Z. Zhang, J. Wang, C. Huang, and M. Debbah, "Joint multi-user communication and sensing exploiting both signal and environment sparsity," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 6, pp. 1409–1422, Nov. 2021.

[15] A. Bazzi and M. Chafii, "Secure full duplex integrated sensing and communications," *IEEE Transactions on Information Forensics and Security*, vol. 19, pp. 2082–2097, Dec. 2023.

[16] 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.

[17] H. Hua, J. Xu, and T. X. Han, "Optimal transmit beamforming for integrated sensing and communication," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 8, pp. 10 588–10 603, Aug. 2023.

[18] A. Bazzi and M. Chafii, "On integrated sensing and communication waveforms with tunable PAPR," *IEEE Transactions on Wireless Communications*, vol. 22, no. 11, pp. 7345–7360, Nov. 2023.

[19] 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.

[20] Q. Shi, L. Liu, S. Zhang, and S. Cui, "Device-free sensing in OFDM cellular network," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 6, pp. 1838–1853, Jun. 2022.

[21] M. Morelli, C.-C. J. Kuo, and M.-O. Pun, "Synchronization techniques for orthogonal frequency division multiple access (OFDMA): A tutorial review," *Proceedings of the IEEE*, vol. 95, no. 7, pp. 1394–1427, Jul. 2007.

[22] R. Rogalin, O. Y. Bursalioglu, H. Papadopoulos, G. Caire, A. F. Molisch, A. Michaloliakos, V. Balan, and K. Psounis, "Scalable synchronization and reciprocity calibration for distributed multiuser MIMO," *IEEE Transactions on Wireless Communications*, vol. 13, no. 4, pp. 1815–1831, Apr. 2014.

[23] 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.

[24] 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.

[25] K. Ji, Q. Zhang, Z. Wei, Z. Feng, and P. Zhang, "Networking based ISAC hardware testbed and performance evaluation," *IEEE Communications Magazine*, vol. 61, no. 5, pp. 76–82, May. 2023.

[26] H. D. Griffiths and N. R. W. Long, "Television-based bistatic radar," *IEE Proceedings F: Communications Radar and Signal Processing*, vol. 133, no. 7, pp. 649–657, Dec. 1986.

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

[28] S. Lu, F. Liu, Y. Li, K. Zhang, H. Huang, J. Zou, X. Li, Y. Dong, F. Dong, J. Zhu, Y. Xiong, W. Yuan, Y. Cui, and L. Hanzo, "Integrated sensing and communications: Recent advances and ten open challenges," arXiv: 2305.00179v2 [cs.IT], Dec. 2023, doi: 10.48550/arXiv.2305.00179.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] 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: 2401.09064 [cs.IT], Jan. 2024, doi: 10.48550/arXiv.2401.09064.

[34] D. Tagliaferri, M. Manzoni, M. Mizmizi, S. Tebaldini, A. V. Monti-Guarnieri, C. M. Prati, and U. Spagnolini, "Cooperative coherent multi-static imaging and phase synchronization in networked sensing," arXiv: 2311.07212 [eess.SP], Nov. 2023, doi: 10.48550/arXiv.2311.07212.

[35] J. Zhao, Z. Lu, J. A. Zhang, S. Dong, and S. Zhou, "Multiple-target Doppler frequency estimation in ISAC with clock asynchronism," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 1, pp. 1382–1387, Jan. 2024.

[36] J. Pegoraro, J. O. Lacruz, T. Azzino, M. Mezzavilla, M. Rossi, J. Widmer, and S. Rangan, "Jump: Joint communication and sensing with unsynchronized transceivers made practical," *IEEE Transactions on Wireless Communications*, Feb. 2024, early access.

[37] 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 Transactions on Mobile Computing*, vol. 22, no. 10, pp. 6160–6175, Oct. 2023.

[38] W. Jiang, Z. Wei, S. Yang, Z. Feng, and P. Zhang, "Cooperation based joint active and passive sensing with asynchronous transceivers for perceptive mobile networks," *IEEE Transactions on Wireless Communications (early access)*, Jul. 2024.

[39] 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 (early access)*, Aug. 2024.

[40] X. Huang, A. Farhang, and R.-R. Chen, "Channel estimation and turbo equalization for coded OTFS and OFDM: A comparison," *IEEE Wireless Communications Letters*, vol. 12, no. 9, pp. 1613–1617, Sep. 2023.

[41] O. Zacharia and M. Vani Devi, "Fractional delay and Doppler estimation for OTFS based ISAC systems," in *Proc. 2023 IEEE Wireless Communications and Networking Conference (WCNC)*, Glasgow, UK, Mar. 2023, pp. 1–6.

[42] K. Meng, S. Yang, X.-Y. Wang, Y. Bu, Y. Tang, J. Zhang, and L. Hanzo, "Joint sparsity pattern learning based channel estimation for massive MIMO-OTFS systems," *IEEE Transactions on Vehicular Technology (early access)*, Mar. 2024.

[43] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.

[44] 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.

[45] A. V. Oppenheim, *Applications of Digital Signal Processing*. Englewood Cliffs, 1978.

[46] R. P. Singh and S. D. Sapre, *Communication Systems*. McGraw-Hill Education (India) Pvt Limited, 2008.

[47] A. Naha, A. K. Samanta, A. Routray, and A. K. Deb, "Determining autocorrelation matrix size and sampling frequency for MUSIC algorithm," *IEEE Signal Processing Letters*, vol. 22, no. 8, pp. 1016–1020, Aug. 2015.

[48] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "SpotFi: Decimeter level localization using WiFi," *ACM SIGCOMM Computer Communication Review*, vol. 45, no. 4, pp. 269–282, Aug. 2015.

[49] R. M. Gray, "Toeplitz and circulant matrices: A review," *Foundations and Trends® in Communications and Information Theory*, vol. 2, no. 3, pp. 155–239, Jan. 2006.

[50] D. Sundararajan, *Discrete Wavelet Transform: A Signal Processing Approach*. John Wiley & Sons, 2016.

[51] S. J. Orfanidis, *Introduction to Signal Processing*. Prentice-Hall, 1995.

Xiao-Yang Wang received the B.S. degree in electronics and information science and technology from Beijing University of Posts and Telecommunications (BUPT), China, in 2020. He is currently pursuing the Ph.D. degree in information and communication engineering with the School of Information and Communication Engineering, BUPT, and with the Key Laboratory of Universal Wireless Communications, Ministry of Education. Since Sep. 2023, he has been a visiting Ph.D. student with the Department of Electronic and Electrical Engineering, University College London, UK. His current research interest includes integrated sensing and communications (ISAC), time/frequency synchronization, and signal processing in distributed massive MIMO systems.

Shaoshi Yang (Senior Member, IEEE) received the B.Eng. degree in information engineering from Beijing University of Posts and Telecommunications (BUPT), China, in 2006, and the Ph.D. degree in electronics and electrical engineering from University of Southampton, UK, in 2013. From 2008 to 2009, he was a Researcher with Intel Labs China. From 2013 to 2016, he was a Research Fellow with the School of Electronics and Computer Science, University of Southampton. From 2016 to 2018, he was a Principal Engineer with Huawei Technologies Co., Ltd., where he made significant contributions to the products, solutions and standardization of 5G, wideband IoT, and cloud gaming/VR. He was a Guest Researcher with the Isaac Newton Institute for Mathematical Sciences, University of Cambridge. He is currently a Full Professor with BUPT. His research interests include 5G/5G-A/6G, massive MIMO, mobile ad hoc networks, distributed artificial intelligence, and cloud gaming/VR. He is a standing committee member of the CCF Technical Committee on Distributed Computing and Systems. He received Dean's Award for Early Career Research Excellence from University of Southampton in 2015, Huawei President Award for Wireless Innovations in 2018, IEEE TCGCC Best Journal Paper Award in 2019, IEEE Communications Society Best Survey Paper Award in 2020, CAI Invention and Entrepreneurship Award in 2023, CIUR Industry-University-Research Cooperation and Innovation Award in 2023, and the First Prize of Beijing Municipal Science and Technology Advancement Award in 2023. He was/is an Editor of *IEEE Transactions on Communications*, *IEEE Systems Journal*, *IEEE Wireless Communications Letters*, and *Signal Processing* (Elsevier). For more details of his research progress, please refer to https://shaoshiyang.weebly.com/.

Hou-Yu Zhai received the B.S. degree in communications engineering from Beijing University of Posts and Telecommunications (BUPT), China, in 2023. He is currently pursuing the Ph.D. degree in information and communication engineering with the School of Information and Communication Engineering, BUPT, and with the Key Laboratory of Universal Wireless Communications, Ministry of Education. His current research interest includes integrated sensing and communications (ISAC), distributed beamforming, and signal processing in distributed massive MIMO systems.

Christos Masouros (Fellow, IEEE) received the Diploma degree in electrical and computer engineering from University of Patras, Greece, in 2004, and the M.Sc. by research and Ph.D. degrees in electrical and electronic engineering from The University of Manchester, U.K., in 2006 and 2009, respectively. In 2008, he was a Research Intern at Philips Research Labs, U.K. Between 2009 and 2010, he was a Research Associate with The University of Manchester and a Research Fellow at Queen's University Belfast between 2010 and 2012. In 2012, he joined University College London as a Lecturer. He has held a Royal Academy of Engineering Research Fellowship between 2011 and 2016. Since 2019, he has been a Full Professor of signal processing and wireless communications with the Information and Communication Engineering Research Group, Department of Electrical and Electronic Engineering, and affiliated with the Institute for Communications and Connected Systems, University College London. His research interests lie in the field of wireless communications and signal processing with particular focus on green communications, large scale antenna systems, integrated sensing and communications (ISAC), interference mitigation techniques for MIMO, and multicarrier communications. He was the co-recipient of the 2021 IEEE SPS Young Author Best Paper Award, and the recipient of the Best Paper Awards in the IEEE GLOBECOM 2015 and IEEE WCNC 2019 conferences. He has been recognized as an Exemplary Editor for *IEEE Communications Letters* and as an Exemplary Reviewer for *IEEE Transactions on Communications*. He is a founding member and the Vice-Chair of the IEEE Emerging Technology Initiative on ISAC, the Vice Chair of the IEEE Special Interest Group on ISAC, and the Chair of the IEEE Special Interest Group on Energy Harvesting Communication Networks. He is an Editor of *IEEE Transactions on Communications*, *IEEE Transactions on Wireless Communications*, *IEEE Open Journal of Signal Processing*, and the Editor-at-Large of *IEEE Open Journal of the Communications Society*. He was an Associate Editor of *IEEE Communications Letters* and a Guest Editor for a number of special issues on *IEEE Journal of Selected Topics in Signal Processing* and *IEEE Journal on Selected Areas in Communications*.

J. Andrew Zhang (Senior Member, IEEE) received the B.Sc. degree from Xi'an JiaoTong University, China, in 1996, the M.Sc. degree from Nanjing University of Posts and Telecommunications, China, in 1999, and the PhD degree from the Australian National University, Australia, in 2004.

Currently, Dr. Zhang is a Professor in the School of Electrical and Data Engineering, University of Technology Sydney, Australia. He was a researcher with Data61, CSIRO, Australia from 2010 to 2016, the Networked Systems, NICTA, Australia from 2004 to 2010, and ZTE Corp., Nanjing, China from 1999 to 2001. Dr. Zhang's research interests are in the area of signal processing for wireless communications and sensing, with a focus on integrated sensing and communications. He has published more than 300 papers in leading Journals and conference proceedings, and has won 5 best paper awards for his work, including the best paper award in IEEE ICC 2013. He is a recipient of CSIRO Chairman's Medal and the Australian Engineering Innovation Award in 2012 for exceptional research achievements in multi-gigabit wireless communications.