Simultaneous Localization and Communications With Massive MIMO-OTFS

Zijun Gong, *Member, IEEE*, Fan Jiang, *Member, IEEE*, Cheng Li, *Senior Member, IEEE*, and Xuemin Shen, *Fellow, IEEE*

***Abstract*— Next generation cellular network is expected to provide the simultaneous high-accuracy localization and ultra-reliable communication services, even in high mobility scenarios. To that end, the novel orthogonal time frequency space (OTFS) modulation has been developed as a promising physical-layer transmission technique, evident by the outstanding performance in terms of robustness against time-frequency selective fading over the orthogonal frequency division multiplexing (OFDM) counterpart. However, when OTFS meets massive multiple-input multiple-output (MIMO), the specific conditions, under which the delay-Doppler (DD) domain channel model holds, are not identified. In addition, the channel estimation and localization performance in such system is rarely studied. In this work, we target at these new challenges, and conduct comprehensive modelling, performance analysis, and algorithm design for massive MIMO-OTFS based simultaneous localization and communications. Specifically, we derive new channel models for the massive MIMO-OTFS system, which captures both time-frequency dispersion and spatial wideband effects. The specific conditions, under which the new models hold has been unveiled as well. Based on the new models, we establish the theoretical foundations for channel estimation and localization, by deriving the Cramér-Rao lower bounds of channel parameter and location estimation errors. Such bounds have been achieved with the newly designed low-complexity channel estimation and localization algorithms. Numerical simulations of the proposed framework with prevailing pulse functions are also conducted and the results validate the proposed designs and analysis.**

***Index Terms*— Simultaneous localization and communications, massive MIMO, OTFS, high mobility.**

I. Introduction

SIMULTANEOUS localization and communications (SLAC) becomes a new feature in the fifth-generation (5G) and beyond cellular networks. This technique can facilitate various applications, such as Internet-of-Things (IoTs) and autonomous driving [1]. The bridge between the localization and communications is the propagation channel, which is used for data transfer while carrying position-dependent information. Typical channel parameters with position dependence include the received signal strength (RSS), time of arrival (ToA), angle of arrival (AoA), Doppler shift, and others. The acquisition of these parameters is dependent on the physical layer signal design and processing of the communications systems.

The orthogonal frequency division multiplexing (OFDM) technology has been extremely successful in the evolution of mobile communications systems for the past three decades, due to its low-complexity implementation with fast Fourier transform (FFT) algorithms and efficiency to combat the inter-symbol-interference (ISI) over multipath fading channels by adding the cyclic prefix (CP). However, towards the evolution to the sixth-generation (6G) wireless communications, OFDM systems will experience the new challenges of strong Doppler effects [1], [2], [3]. For one thing, ultra-high data rate requirements in 6G use cases motivate the use of new radio frequencies over millimeter wave (mmWave) and terahertz (THz) bands where large bandwidth is available [4]. For another, the emerging applications, such as the mobile communications for high-speed railways (HSR), autonomous vehicles, on-board aircraft, low-earth orbit satellites (LEOs), unmanned aerial vehicles (UAVs), have critical challenges to maintain reliability in high-mobility environments [5], [6], [7], [8], [9]. The induced strong Doppler effects will destroy the orthogonality of the sub-channel over subcarriers, causing significant inter-carrier-interference (ICI) in OFDM systems [3].

To highlight the motivation on massive multiple-input multiple-output (MIMO) and orthogonal time frequency space (OTFS) for SLAC, we will review the related work as follows.

A. OTFS Versus OFDM

While the Doppler shift has been recognized to be harmful in OFDM systems, it can be utilized to separate multi-path propagation and mitigate multi-user interference when the user data is modulated in delay-Doppler (DD) domain [10], [11]. This enlightens a ground-breaking modulation technique known as OTFS, initially introduced in [12] and [13]. By establishing the signal transmission and receiving in the DD domain, we can develop the OTFS framework to spread each information symbol over the entire time-frequency (TF)

Manuscript received 9 January 2023; revised 26 June 2023; accepted 13 August 2023. Date of publication 9 October 2023; date of current version 22 November 2023. This work was supported by the National Natural Science Foundation of China under Grant 62201162. *(Corresponding author: Zijun Gong.)*

Zijun Gong was with NSERC PDF, University of Waterloo, Waterloo, ON N2L 3G1, Canada. He is now with the IoT Thrust, The Hong Kong University of Science and Technology (HKUST) (Guangzhou), Guangzhou 511458, China, and also with the Department of ECE, HKUST, Hong Kong, SAR, China (e-mail: gongzijun@ust.hk).

Fan Jiang was with the School of Information Technology, Halmstad University, 30118 Halmstad, Sweden. He is now with the Peng Cheng Laboratory (PCL), Shenzhen 518066, China (e-mail: jiangf02@pcl.ac.cn).

Cheng Li is with the School of Engineering Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada (e-mail: cheng_li_5@sfu.ca).

Xuemin Shen is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: sshen@uwaterloo.ca).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JSAC.2023.3322818.

Digital Object Identifier 10.1109/JSAC.2023.3322818

0733-8716 © 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.

domain [14], [15]. As a result, full TF diversity gain can be achieved to mitigate the doubly selective fading, enabling robust and reliable communications in highly dynamic and complex environments [16]. It has been widely demonstrated that OTFS outperforms OFDM over doubly-dispersive channels [17], [18], [19], [20], [21], [22], [23]. For example, OTFS modulation can be implemented with one CP for the entire TF block, thus significantly reducing the overhead compared to the OFDM counterpart [18]. Also, with the DD domain representation, OTFS systems enjoy much longer coherence time than OFDM systems, simplifying the channel estimation and data detection algorithm designs [19], [20], [21]. The conventional linear time invariant (LTI) channel model used for OFDM is actually the 0-th order Taylor expansion of the true wireless channel, while the DD domain channel model is close to the first-order Taylor approximation. As a result, the modelling error of the DD domain channel model accumulates over time much more slowly, leading to prolonged coherence time. It has been also reported that the OTFS modulation exhibits low peak-to-average power ratio (PAPR) and robustness to timing and frequency synchronization errors over OFDM techniques [22]. Nevertheless, a fair comparison between OFDM and OTFS modulations is difficult to conclude since the requirements and hardware differ from case to case [17].

B. Doubly-Dispersive Channels and TF Signalling

The underlying DD domain channel models are critical for the development of OTFS frameworks. The investigation of such doubly-dispersive channels dates back to Bello's benchmark work, where the mathematical fundamentals of *linear time-varying* (LTV) channels were derived [24]. Specifically, the LTV channels can be represented with multipath components, each associated with a path gain, propagation delay, and the Doppler shift caused by the mobility [24], [25], [26]. By doing that, the time-varying TF domain channel can be transformed to the LTV DD domain channel. The validity of such LTV channel holds assuming that the propagation environment remains unchanged during the entire TF block transmission, while the time-varying part is caused by the relative movement between the transmitter and receiver [3]. Besides the theoretical analysis, the measurement campaign with a time-domain channel sounder at 60 GHz for vehicle-to-infrastructure (V2I) communications has been conducted, where the sparse representation of such channel in DD domain has been demonstrated [27].

Prior to the OTFS work, several optimal TF signalling studies have been done on doubly-dispersive channels [28], [29], [30], [31]. By treating the received signal as a canonical decomposition of shifted versions of the transmit signal in DD domain, a DD domain RAKE receiver is designed in [28] to exploit the double dispersions. A non-orthogonal frequency division multiplexing system has been proposed in [29], employing a similar transceiver structure to OFDM systems by replacing the OFDM modulation/demodulation with *Heisenberg/Wigner* transforms. The general framework to account for the double dispersions based on short-time Fourier basis has been developed in [30], where the authors have shown that the optimal pulse should have the TF spread proportional to the channel spread. In general, these works focus on the transceiver designs in the TF domain rather than the DD domain, which differs from the OTFS framework. Additionally, when combining OTFS with multiple antenna systems, the spatial wideband effect should be properly considered in the design [32], [33].

C. The Implementation of OTFS

Generally, the OTFS modulation at transmitter can be implemented in two ways: an overlay of OFDM modulation by employing the inverse-symplectic finite Fourier transform (ISFFT) [34], [35], [36], and the direct implementation based on Zak transform [37], [38], [39]. In the first approach, the ISFFT-based pre-processing module is introduced before the OFDM modulation at transmitter, and the symplectic finite Fourier transform (SFFT)-based post-processing module is applied after the OFDM demodulation, respectively. With this approach, the OTFS implementation can reuse the OFDM transceiver structure; however, multiple CPs are required, degrading the overall spectrum efficiency [15]. On the other hand, the direct implementation based on continuous-time Zak-transform can map DD domain signal to the time-domain representation [39]. Specifically, the inverse discrete Zak-transform (IDZT) is shown to be equivalent to the sequential combination of ISFFT and OFDM modulation; and the discrete Zak-transform (DZT) is equivalent to the sequential combination of OFDM demodulation and SFFT. By using the DZT to replace the conventional ISFFT, computational complexity can be significantly reduced. In addition, the direct implementation can reduce the overhead of multiple CPs, thus improving the spectrum efficiency [38]. In general, the DZT-based implementation can simplify the DD domain derivation and analysis of the input-output to the doubly-dispersive channels.

Another important component associated with the OTFS implementation is the waveform design [23]. It is generally assumed that the transmit and receive window functions satisfy the bi-orthogonal property, such that the ICI and inter-symbol-interference (ISI) can be removed [26], [30]. In addition, the used window function will impact the PAPR performance [40]. It is demonstrated that the rectangular pulse outperforms the raised cosine (RC) pulse, and then Gaussian pulse in reducing the PAPR [22]. Moreover, since the pulse function behaves differently in delay and Doppler estimation accuracy, the impact of the pulse shapes on localization accuracy needs further investigation.

D. Channel Estimation and Data Detection in OTFS Systems

With the DD domain signal representation, the channel estimation in OTFS systems can be conducted in the DD domain [41], [42], [43]. The Cramér-Rao lower bound (CRLB) for DD domain channel estimation accounting for non-ideal pulse shaping in multiple use cases has been derived, used as benchmark for channel estimation accuracy [41]. Inspired by the sparsity of the DD domain channels,

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.

a set of compressed-sensing and sparse-Bayes learning (SBL) based approaches have been studied [44], [45], [46], [47], [48]. Specifically, a 3D-structured orthogonal matching pursuit algorithm has been exploited in [44] by formulating the corresponding channel estimation issue as a sparse signal recovery (SSR) problem. The SSR problem is solved with a two-dimensional off-grid SBL-based approach in [47], with hyper-parameter update for the conditional posterior distribution of the channels. By linking channel parameters to the location and moving speed of the user, a sensing-aided channel estimation work has been studied in [49]. In general, sparsity-inspired channel estimation methods can significantly reduce the overhead and complexity for channel estimation.

By establishing the DD domain channel and signal representation, linear data detection schemes can be directly applied with the knowledge of the channel matrix [50], [51], [52], [53]. Closed-form bit error rate (BER) performance using linear zero-forcing receivers has been reported in [50]. By noting that the linear minimum mean-square error (LMMSE) equalization matrix is sparse and quasi-banded, a low-complexity matrix inversion method is proposed in [51] to reduce the implementation complexity of the LMMSE detection to log-linear level. The sparsity and quasi-banded structure are further utilized in [53] to maintain low-complexity implementation of linear detection methods considering different pulse functions. The message passing based detection methods have been also studied in OTFS systems, which achieve better trade-off between the complexity and BER performance [54], [55]. The application of machine-learning based data detection methods to OTFS systems has been recently reported [56], [57]. To summarize, since the data detection requires the estimation of the channel, and the channel estimation can benefit from data detection, a joint channel estimation and data detection framework enjoys much improved performance [58], [59], [60], at the price of higher complexity.

E. OTFS for Sensing

Since OTFS is based on the doubly-dispersive channels, and both delay (interpreted as rang) and Doppler (equivalent to radial velocity/direction) parameters are critical sensing performance metrics, the application of OTFS to sensing framework has been studied recently [36], [61], [62], [63], [64]. Particularly, in [61], the OTFS sensing problem is solved with an iterative optimization method by searching the target in the predefined DD grids. Distinct from the DD domain processing, the authors in [36] proposed a time domain generalized likelihood ratio test (GLRT) detector, where the ICI and ISI are exploited to surpass the ambiguity barrier. Considering the advantages of OTFS over OFDM in both mobile communications and sensing, a universal waveform based on OTFS for integrated sensing and communications (ISAC) is generally preferred [62], [63], [64]. However, more studies on this topic are required.

F. Our Contributions

In this paper, we conduct channel modelling, performance analysis and algorithm design for SLAC based on massive MIMO-OTFS systems. The underlying assumptions of the conventional channel model are discussed in detail, and the spatial wideband effect is considered. Based on the newly derived signal model, comprehensive performance analysis is conducted, and low-complexity algorithms are designed. Our detailed contributions are summarized below.

- We have derived the massive MIMO-OTFS channel model by considering the ultra-wide bandwidth mmWave signals and large antenna arrays. Specifically, we prove that the product of frame length and bandwidth is confined by the reciprocal of the Doppler scaling factor. Also, we show that the spatial wideband effect cannot be ignored in channel modelling for massive MIMO systems with ultra-wide bandwidth.
- Based on the multi-path wideband channel model, the CRLBs are derived for channel estimation and localization results. A fundamental trade-off between bandwidth (delay estimation accuracy) and frame length (Doppler estimation accuracy) is unveiled. The impacts of multipath effect and spatial wideband effect on positioning accuracy are quantified and shown to be negligible with careful algorithm design.
- A low complexity channel estimation method based on FFT is proposed and shown to give the maximum likelihood estimate of the LoS (Line of Sight) channel parameters. Based on the estimated ToA, AoA and Doppler shift, the 3D positioning can be modelled as a non-linear weighted least-squares problem, and solved through Newton's method.
- Comprehensive numerical results are presented to validate the theoretical analysis and proposed designs. From the simulation results, we can clearly see how different factors contribute to positioning accuracy, including bandwidth, frame length, antenna array size, SNR (Signal to Noise Ratio), multi-path effect, spatial wideband effect, shape of the modulation pulse. These results will shed light on algorithm and system designs.

G. Organization and Notations

The rest of the paper is organized as follows. The massive MIMO-OTFS signal model with high mobility is introduced in Section II. After that, the theoretical foundations for channel estimation and localization are presented in Section III. Low-complexity channel estimation and localization algorithm are provided in Section IV. Following that, we present the numerical simulations and performance discussion in Section V. Finally, the conclusions are drawn in Section VI.

Throughout this paper, the uppercase bold letters represent matrices while the lowercase bold letters denote column vectors. $\mathbb{E}\{\cdot\}$ represents the mean of a random variable. For an arbitrary vector $\mathbf{a}$, $\mathbf{a}[n]$ is the n-th element. $\mathbf{A}[n,m]$ denotes the element lying on the n-th row and m-th column of matrix $\mathbf{A}$. $\mathbf{A}^{\mathrm{T}}$, $\mathbf{A}^{\mathrm{H}}$, and $\mathbf{A}^{*}$ indicate the transpose, Hermitian transpose, and element-wise conjugate of matrix $\mathbf{A}$. $\|\cdot\|$ denotes the L_2 norm of a vector, while $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm of a matrix. $\mathrm{tr}\{\mathbf{A}\}$ gives the trace of matrix $\mathbf{A}$. $\mathrm{diag}\{\mathbf{A}\}$ is a column vector containing the diagonal elements of $\mathbf{A}$,

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.

while $\mathrm{diag}\{\mathbf{a}\}$ constructs a diagonal matrix with elements of $\mathbf{a}$.

II. Massive MIMO-OTFS Modelling With High Mobility

In this section, we will present the wireless channel model for massive MIMO-OTFS systems with high mobility user devices, i.e., vehicles. We will start with single-input-single-output (SISO) channel model. Then in the second sub-section, the Massive MIMO-OTFS channel model will be presented, and the underlying approximations/assumptions will be discussed. Due to the huge antenna array and bandwidth, the signal will experience the spatial wideband effect (beam squint).

A. OTFS Channel Modelling With SISO

We consider a point-to-point wireless channel with L multipath components, each associated with a path gain, delay, AoA, and Doppler shift. With a carrier frequency of f_c Hz, the equivalent baseband channel response in the time-delay domain, i.e., $h(t,\tau)$, is given as [65]

$$h(t,\tau) = \sum_{l=0}^{L-1} a_l e^{-j2\pi f_c \tau_l(t)} \delta(\tau - \tau_l(t)), \tag{1}$$

where a_l and $\tau_l(t)$ are the real-time complex path gain and propagation delay of the l-th path. Due to the Doppler effect, the propagation delays of different paths are changing over time. In a relatively short period, we can safely assume that the propagation delays change over time linearly:

$$\tau_l(t) \approx \tau_l + \rho_l t, \tag{2}$$

where τ_l is the zero-th order Taylor expansion of $\tau_l(t)$, and ρ_l quantifies how fast the propagation delay of the l-th path changes. Because ρ_l is proportional to the Doppler shift and indicates the strength of the Doppler effect, it is generally referred to as the *Doppler scaling factor*.

In OTFS modulation, suppose we are transmitting a 2D data frame $\mathbf{X} \in \mathbb{C}^{N\times M}$, with N time slots and M sub-carriers. The transmitted signal is

$$s(t) = \sum_{n,m} X_{n,m} g_t(t - nT) e^{j2\pi mF(t-nT)}, \tag{3}$$

where $X_{n,m}$ is the symbol on the (n,m)-th time-frequency slot, i.e., $X_{n,m} = \mathbf{X}[n,m]$. As we can see, symbols are modulated on different frequencies and shifted in time. The sub-carrier spacing is F Hz, while the symbol duration is T second. Generally we have $TF = 1$, and the signal can thus be rewritten as

$$s(t) = \sum_{n,m} X_{n,m} g_t(t - nT) e^{j2\pi mFt}. \tag{4}$$

At the receiver side, the received signal is

$$\begin{aligned} r(t) &= \int_\tau s(t-\tau) h(t,\tau) d\tau \\ &= \sum_l a_l e^{-j2\pi f_c(\tau_l + \rho_l t)} s(t - \tau_l - \rho_l t). \end{aligned} \tag{5}$$

In high-mobility scenarios, the Doppler effect is significant, i.e., $\rho_l \neq 0$, and it is impacting the received signal in two ways: first, the spectrum is shifted by $\nu_l = -\rho_l f_c$ for the l-th path, i.e., Doppler shift; second, the baseband signal $s(t)$ is scaled by a factor of $1 - \rho_l$ in time. For terrestrial communications with electromagnetic signals, the Doppler scaling factor is generally very small. For example, for high-speed train moving at 360 km/h (or 100 m/s), we have $|\rho_l| \leq \frac{1}{3}\times 10^{-6}$. In this case, the second impact of Doppler effect can be ignored for certain parameters. Specifically, suppose B is the total bandwidth of $s(t)$, and $|\rho_l t| \ll 1/B$ holds for any l and $t \in [0, NT]$, we approximately have

$$s(t - \tau_l - \rho_l t) \approx s(t - \tau_l). \tag{6}$$

Define $S = NT$ as the frame length in time, and we equivalently have

$$MN = BS \ll \min_l |1/\rho_l|. \tag{7}$$

When inequality (7) holds, we can safely ignore the scaling effect of $s(t)$, resulting from the Doppler effect. As we can see, the multiplication of frame lengths in time and frequency domains cannot be very large in OTFS. For example, consider high-speed railway system at 360 km/h, the product of M and N should be much smaller than 10^6. With a bandwidth of 10 MHz, the frame length should be no larger than 10 ms. We can also further increase the bandwidth to 100 MHz, and reduce the frame length to 1 ms. Apparently, there is a trade-off between total bandwidth and frame length, and we can adjust them based on our needs.

The above discussions are based on the time-delay channel response, while the DD domain channel modelling is almost exclusively used in OTFS related papers. The DD domain channel response is given as

$$h(\tau,\nu) = \sum_{l=0}^{L-1} \tilde{a}_l \delta(\tau - \tau_l, \nu - \nu_l), \tag{8}$$

with $\tilde{a}_l = a_l e^{-j2\pi(f_c\tau_l - \nu_l\tau_l)}$ being the equivalent channel gain, and $\nu_l = -\rho_l f_c$ is the Doppler shift of the l-th path. This leads to another representation of the received signal as

$$r(t) = \int_\nu \int_\tau s(t-\tau) h(\tau,\nu) e^{j2\pi\nu(t-\tau)} d\tau d\nu, \tag{9}$$

known as the Heisenberg transform. Furthermore, we have more explicit representation of the received signal as

$$\begin{aligned} r(t) &= \sum_l \tilde{a}_l e^{j2\pi\nu_l(t-\tau_l)} s(t - \tau_l) \\ &= \sum_l a_l e^{-j2\pi f_c \tau_l} e^{j2\pi\nu_l t} s(t - \tau_l). \end{aligned} \tag{10}$$

It should be noted that the result in (10) is only approximation of that in (5) by ignoring the scaling of the baseband signal resulting from the Doppler effect. Equivalently, the DD domain channel model in (8) is approximation of that in (1) by taking the first-order Taylor expansion in (2) and then ignoring the scaling effect on baseband signals. This approximation is only accurate when inequality (7) holds.

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.

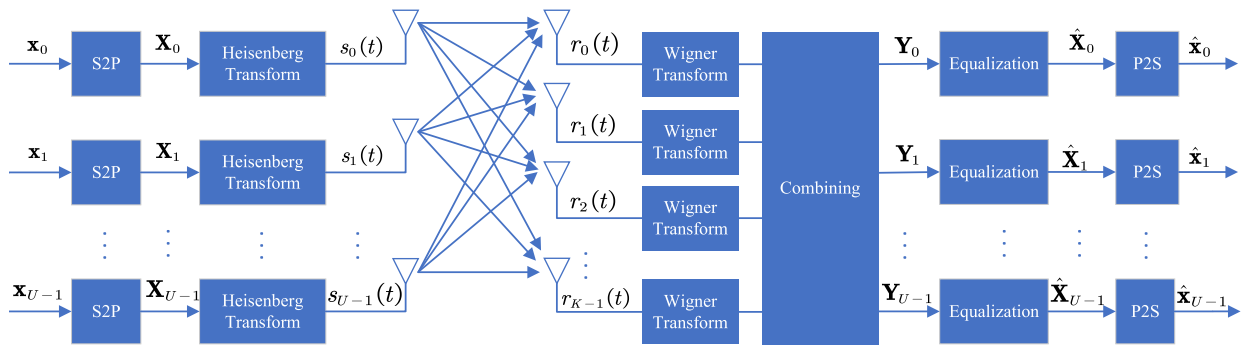


Fig. 1. Architecture of the multi-user MIMO-OTFS systems. S2P means serial-to-parallel, while P2S is for parallel-to-serial.

Suppose $g_r(t)$ is the receiving pulse, and we process the received signal with a matched filter:

$$\begin{aligned} Y(t,f) &= \int_{t'} r(t')g_r^*(t'-t)e^{-j2\pi f(t'-t)}dt' \\ &= \sum_{n,m} X_{n,m} \sum_l \tilde{a}_l e^{j2\pi(\nu_l+mF)(t-\tau_l)} \\ &\quad \cdot A(t-\tau_l-nT, f-\nu_l-mF), \end{aligned} \tag{11}$$

where $A(\tau,\nu)$ is the cross-ambiguity function of $g_t(t)$ and $g_r(t)$, given by

$$A(\tau,\nu) = \int_t g_t(t)g_r^*(t-\tau)e^{-j2\pi\nu(t-\tau)}dt. \tag{12}$$

Equation (11) is also referred to as the Wigner transform of the received signal.

$Y(t,f)$ is the received signal in time-frequency domain, and we can sample it as $Y_{n,m} = Y(nT, mF)$, given by

$$\begin{aligned} Y_{n,m} &= \sum_{n',m'} X_{n',m'} \sum_l \tilde{a}_l e^{j2\pi(\nu_l+m'F)(nT-\tau_l)} \\ &\quad \cdot A(nT-\tau_l-n'T, mF-\nu_l-m'F). \end{aligned} \tag{13}$$

For $n \neq n'$ and $m \neq m'$, the cross-ambiguity function is very small (i.e., bi-orthogonality), and the ISI in time-frequency domain can be ignored, resulting in

$$Y_{n,m} = X_{n,m} \sum_l \alpha_l e^{j2\pi nT\nu_l} e^{-j2\pi mF\tau_l}, \tag{14}$$

where α_l is given as

$$\alpha_l = a_l e^{-j2\pi f_c \tau_l} A(-\tau_l, -\nu_l). \tag{15}$$

The bi-orthogonality of the cross-ambiguity function lays a fundamental base for the OTFS framework. The validity of this assumption is dependent on the product of the delay spread and Doppler spread of the wireless channels. For terrestrial communications with electromagnetic signals, this product is generally smaller than 10^{-2}, and we can safely rely on the bi-orthogonality assumption.

As we can see, the Heisenberg transform at the transmitter side and the Wigner transform at the receiver side can help construct orthogonal sub-channels, similar to the OFDM. The ISI in time and frequency domains can be ignored, and we now have $M \times N$ *flat-fading* channels. The complex channel gain at the (n,m)-th slot will be

$$H_{n,m} = \sum_l \alpha_l e^{j2\pi nT\nu_l} e^{-j2\pi mF\tau_l}. \tag{16}$$

This is the SISO channel model, and we will extend the discussion to MIMO systems in the following sub-section.

B. Massive MIMO and OTFS

Similar to MIMO-OFDM systems, OTFS can also be combined with MIMO for spatial multiplexing, and the architecture is given in Fig. 1. This framework is very much like that of the MIMO-OFDM systems, and the major difference is that the IFFT/FFT are replaced by Heisenberg/Wigner transform, respectively. The base station is equipped with K antennas to simultaneously serve U single antenna users through space division multiple access. This structure dates back to decades ago, with the name *Weyl-Heisenberg systems*, or *non-orthogonal frequency division multiplexing (NOFDM)* [29]. With OTFS modulation, the doubly-dispersive channel is transformed to flat-fading sub-channels in time-frequency domain. On each sub-channel, multiple data streams are transmitted simultaneously through the MIMO technique. The data detection and beamforming techniques in the conventional MIMO systems can be directly utilized here. In practice, the BS assigns orthogonal pilot sequences to different users to avoid inter-user interference during uplink channel estimation. Therefore, we will focus on the single-user case in the following discussions. The pilot design for multi-user case will be investigated in sub-section IV-D.

Suppose the BS is equipped with a large uniform linear array (ULA), with K antennas. The antennas are critically spaced, i.e., the inter-antenna distance is $D = c/(2f_c) = \lambda_c/2$, with λ_c being the wavelength of the carrier. Then based on (10), we have the received signal at the k-th antenna of the BS as

$$r_k(t) = \sum_l a_l e^{-j2\pi f_c \tau_{l,k}} e^{j2\pi\nu_l t} s(t-\tau_{l,k}), \tag{17}$$

with $\tau_{l,k}$ denoting the propagation delay between the user antenna and the k-th antenna at the BS along the l-th path. When the distance between BS and user device is much larger

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.

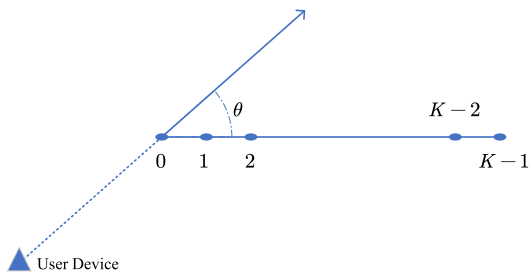


Fig. 2. Illustration of AoA.

than the antenna dimension at the BS, we have

$$\tau_{l,k} \approx \tau_l + kD\cos\theta_l/c, \tag{18}$$

where $\tau_l = \tau_{l,0}$, θ_l is the AoA of the k-th path, and c is speed of light. This angle is formed by the vector pointing from the 0-th antenna to the $(K-1)$-th antenna at the BS and the vector pointing from the user device to the BS, illustrated in Fig. 2.

If we define the spatial signature of the l-th path as

$$\omega_l = \pi\cos\theta_l, \tag{19}$$

the received signal at the k-th antenna can then be rewritten as

$$r_k(t) = \sum_l a_l e^{-j2\pi f_c\tau_l} e^{-jk\omega_l} e^{j2\pi\nu_l t} s(t-\tau_{l,k}). \tag{20}$$

Given that $KD \ll c/B$, the received signal at the k-th antenna can be further simplified as

$$r_k(t) = \sum_l a_l e^{-j2\pi f_c\tau_l} e^{-jk\omega_l} e^{j2\pi\nu_l t} s(t-\tau_l), \tag{21}$$

which means all the antennas are *seeing* the same symbol simultaneously. The propagation delays among different antennas only incur a phase delay, but not time delay in the baseband signal. This is a common assumption in MIMO-OTFS related work. Based on such an assumption, the channel gain between the user antenna and the k-th antenna at the BS for the (n,m)-th slot is $H_{k,n,m}$, given as

$$H_{k,n,m} = \sum_l \alpha_l e^{j2\pi nT\nu_l} e^{-j2\pi mF\tau_l} e^{-jk\omega_l}. \tag{22}$$

For the bandwidth satisfying $B \ll c/(KD)$, the propagation delay will only introduce a phase delay that grows linearly with antenna index, but no time delay.

However, in massive-MIMO and millimeter wave communications, both the antenna array size and bandwidth can be very large, and when c/B is comparable to the antenna array dimension, the delay in baseband signal can no longer be ignored, i.e., the spatial wideband effect [32], [33]. From another perspective, for huge bandwidth, different frequency components are seeing different steering vectors, which is referred to as "beam squint" in radar. In such cases, we have to consider the fact that $s(t-\tau_{l,k})$ cannot be approximated as $s(t-\tau_l)$. Specifically, we have

$$s(t-\tau_{l,k}) = \sum_{n,m} X_{n,m} g_t(t-\tau_{l,k}-nT) e^{j2\pi mF(t-\tau_{l,k})}. \tag{23}$$

For $KD \ll cT = c/(B/M) = Mc/B$, which can be easily satisfied, we have $g_t(t-\tau_{l,k}-nT) \approx g_t(t-\tau_l-nT)$, leading to

$$s(t-\tau_{l,k}) \approx \sum_{n,m} X_{n,m} g_t(t-\tau_l-nT) e^{j2\pi mF(t-\tau_l)} \cdot e^{-j2\pi mFkD\cos\theta_l/c}. \tag{24}$$

As we can see, there is an extra phase shift that is proportional to the product of frequency index and antenna index (i.e., m and k) in the last component at the right hand side of (24). By following similar procedure of the previous sub-section, the complex channel gain $H_{k,n,m}$ under the spatial wideband effect can be adjusted as

$$H_{k,n,m} = \sum_l \alpha_l e^{j2\pi nT\nu_l} e^{-j2\pi mF\tau_l} e^{-jk\omega_{l,m}}, \tag{25}$$

with $\omega_{l,m}$ given as

$$\omega_{l,m} = 2\pi(f_c + mF)D\cos\theta_l/c. \tag{26}$$

That is to say, the spatial signature, or steering vector equivalently, is changing over the sub-carriers. Comparing (22) and (25), we can see that if we don't consider the spatial wideband effect, a phase shift of $2\pi kmFD\cos\theta_l/c$ on the m-th sub-carrier is ignored for the l-th path and k-th antenna. Consider a total bandwidth of 100 MHz, with $f_c = 30$ GHz and $D = \lambda_c/2 = 5$ mm, $K = 100$, the ignored phase component can be as large as $2\pi KDB/c = \pi/3$, which is too significant to be ignored.

With the discrete-time equivalent baseband channel model, we are now ready to discuss the potential performance of such systems on localization in the next section. Before that, we will briefly summarize the assumptions of the channel model presented in this section as follows.

(1) Suppose each resource block has a size of B Hz $\times$ S seconds. Given that the user device moves at v, we have to guarantee that $BS \ll c/v$. For example, for a vehicular speed of 120 m/s, $c/v = 2.5 \times 10^6$.
(2) For the conventional massive MIMO model, we should have $KD \ll c/B$ or $B \ll c/(KD)$, if beam squint is ignored. By considering the spatial wideband effect in modelling and algorithm design, we can relax the constraint to $B/M \ll c/(KD)$, which can be easily satisfied and thus is more generally applicable.

III. Theoretical Foundations

In this section, we will analyze the CRLB of channel estimation and positioning error. Intuitively, the LoS path signal contains the desired position information of the mobile device, and the positioning accuracy is dependent on how accurate we can estimate the parameters of the LoS path signal, including propagation delay, Doppler shift, and AoA, i.e., τ_0, ν_0 and θ_0. The estimation errors of these three parameters are impacted by SNR, the multi-path effect, bandwidth, frame length, modulation pulses, etc. The estimation errors of these geometrical channel parameters will then impact positioning accuracy. Moreover, the estimation error of AoA is actually a function of AoA itself. Specifically, when AoA is close to

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.

$\pi/2$, the bearing estimation error will be minimized, while an AoA of 0 or π leads to infinitely large bearing error. Besides, for ToA and AoA based positioning, the positioning error also grows with ToA linearly. As we can see, how the estimation errors of channel parameters map to positioning errors is very complicated, which is exactly what we will attempt to reveal in this section. Specifically, we want to investigate the following problems.

- Increased frame length in time domain provides higher Doppler estimation accuracy, while larger bandwidth helps to estimate ToAs better. However, the product of total bandwidth and frame length is constrained by the physical channel, i.e., (7). Then, how should we optimize the frame length and bandwidth for positioning performance?
- The Doppler shift is generally believed to be a negative factor in wireless communications. However, Doppler shift also contains the position information of the mobile device. Therefore, it is possible that Doppler shift can actually help improve positioning accuracy.
- With a ULA, 2D positioning can be easily achieved by combining AoA and ToA measurements. However, by introducing Doppler shift, we will show that even 3D positioning becomes possible, given that the antenna array is properly installed at the BS.

By deriving the CRLB, we will be able to shed lights on these problems, and get some insights on practical system design.

A. Channel Estimation Error

Suppose we take P slots on time and Q slots on frequency for channel estimation and localization. That is to say, we insert pilots on (n_p, m_q)-th slots, with $n_p \in \{n_0, n_1, \cdots, n_{P-1}\}$ and $m_q \in \{m_0, m_1, \cdots, m_{Q-1}\}$. These pilot symbols are known at the BS. With these symbols, we can estimate the channel parameters. Each path has four parameters, the complex gain, propagation delay, Doppler shift, and AoA (or spatial signature, equivalently), i.e., α_l, τ_l, ν_l, ω_l. For positioning, we are only interested in the position-related parameters, and the amplitude of α_l has no significance. To decouple the channel gains from the other geometrical channel parameters, we further define ϕ_l as the phase of α_l, i.e., $\alpha_l = |\alpha_l| e^{j\phi_l}$ and $\boldsymbol{\theta}_l = [\tau_l, \nu_l, \omega_l, \phi_l]^{\mathrm{T}}$. Then we can put all the parameters from different paths together and represent them more concisely as $\boldsymbol{\theta} = [\boldsymbol{\theta}_0^{\mathrm{T}}, \boldsymbol{\theta}_1^{\mathrm{T}}, \cdots, \boldsymbol{\theta}_{L-1}^{\mathrm{T}}]^{\mathrm{T}}$, $\mathbf{a} = [|\alpha_0|, |\alpha_1|, \cdots, |\alpha_{L-1}|]^{\mathrm{T}}$ and $\tilde{\boldsymbol{\theta}} = [\boldsymbol{\theta}^{\mathrm{T}}, \mathbf{a}^{\mathrm{T}}]^{\mathrm{T}}$.

The measured channel matrix on the m_q-th sub-carrier is $\hat{\mathbf{H}}_{m_q}[k,p] = \hat{H}_{k,n_p,m_q}$, with the perfect CSI given as

$$H_{k,n_p,m_q} = \sum_l \alpha_l e^{j2\pi n_p T \nu_l} e^{-j2\pi m_q F \tau_l} e^{-jk\omega_l}. \tag{27}$$

Suppose the estimation errors of different channel gains are i.i.d. cyclically symmetric complex Gaussian random variables, with a variance of σ^2. For notional convenience, we concatenate the channel matrices on different sub-carriers, i.e., $\hat{\mathbf{H}} = \left[\hat{\mathbf{H}}_{m_0}, \hat{\mathbf{H}}_{m_1}, \cdots, \hat{\mathbf{H}}_{m_{Q-1}}\right]$, with the true channel matrix being $\mathbf{H} = \mathbb{E}\{\hat{\mathbf{H}}\}$. The likelihood of the channel estimate is then given as

$$p(\hat{\mathbf{H}}; \tilde{\boldsymbol{\theta}}) = (\pi\sigma^2)^{-KPQ} \exp\left\{-\|\hat{\mathbf{H}} - \mathbf{H}\|_{\mathrm{F}}^2/\sigma^2\right\}, \tag{28}$$

The log-likelihood function is

$$l_g(\hat{\mathbf{H}}; \tilde{\boldsymbol{\theta}}) = -KPQ\ln(\pi\sigma^2) - \|\hat{\mathbf{H}} - \mathbf{H}\|_{\mathrm{F}}^2/\sigma^2. \tag{29}$$

The gradient with respect of $\tilde{\boldsymbol{\theta}}$ is

$$\nabla_{\tilde{\boldsymbol{\theta}}} l_g = -\frac{2}{\sigma^2} \sum_{k,p,q} \underbrace{\Re\{\nabla_{\tilde{\boldsymbol{\theta}}} H_{k,n_p,m_q} (H_{k,n_p,m_q} - \hat{H}_{k,n_p,m_q})^*\}}_{l_{k,p,q}}. \tag{30}$$

Then the Fisher information matrix (FIM) concerning $\tilde{\boldsymbol{\theta}}$ is

$$\begin{aligned}\mathbf{F}_{\tilde{\boldsymbol{\theta}}} &= \mathbb{E}\left\{\nabla_{\tilde{\boldsymbol{\theta}}} l_g \nabla_{\tilde{\boldsymbol{\theta}}}^{\mathrm{T}} l_g\right\} \\ &= \frac{2}{\sigma^2} \sum_{k,p,q} \Re\left\{\nabla_{\tilde{\boldsymbol{\theta}}} H_{k,n_p,m_q} \nabla_{\tilde{\boldsymbol{\theta}}}^{\mathrm{H}} H_{k,n_p,m_q}\right\},\end{aligned} \tag{31}$$

where we implicitly used the fact that $\mathbb{E}\{l_{k,p,q} l_{k',p',q'}\} = 0$ for $k \neq k'$ or $p \neq p'$ or $q \neq q'$. Because $\Re\left\{\nabla_{\boldsymbol{\theta}} H_{k,n_p,m_q} \nabla_{\mathbf{a}}^{\mathrm{H}} H_{k,n_p,m_q}\right\} = \mathbf{0}$, we have

$$\mathbf{F}_{\tilde{\boldsymbol{\theta}}} = \mathrm{diag}\left\{\mathbf{F}_{\boldsymbol{\theta}}, \mathbf{F}_{\mathbf{a}}\right\}, \tag{32}$$

with $\mathbf{F}_{\boldsymbol{\theta}}$ and $\mathbf{F}_{\mathbf{a}}$ given as

$$\mathbf{F}_{\boldsymbol{\theta}} = \mathbb{E}\left\{\nabla_{\boldsymbol{\theta}} l_g \nabla_{\boldsymbol{\theta}}^{\mathrm{T}} l_g\right\}, \quad \mathbf{F}_{\mathbf{a}} = \mathbb{E}\left\{\nabla_{\mathbf{a}} l_g \nabla_{\mathbf{a}}^{\mathrm{T}} l_g\right\}. \tag{33}$$

The FIM concerning $\boldsymbol{\theta}$ can be rewritten as

$$\mathbf{F}_{\boldsymbol{\theta}} = \begin{bmatrix} \mathbf{F}_{0,0} & \mathbf{F}_{0,1} & \cdots & \mathbf{F}_{0,L-1} \\ \mathbf{F}_{0,1}^{\mathrm{T}} & \mathbf{F}_{1,1} & \cdots & \mathbf{F}_{1,L-1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{F}_{0,L-1}^{\mathrm{T}} & \mathbf{F}_{1,L-1}^{\mathrm{T}} & \cdots & \mathbf{F}_{L-1,L-1} \end{bmatrix}, \tag{34}$$

with $\mathbf{F}_{l,l'}$ defined as

$$\mathbf{F}_{l,l'} = \frac{2}{\sigma^2} \sum_{k,p,q} \Re\left\{\nabla_{\boldsymbol{\theta}_l} H_{k,n_p,m_q} \nabla_{\boldsymbol{\theta}_{l'}}^{\mathrm{H}} H_{k,n_p,m_q}\right\}. \tag{35}$$

The gradient is given as

$$\nabla_{\boldsymbol{\theta}_l} H_{k,n_p,m_q} = j a_{k,p,q,l} [-2\pi m_q F, 2\pi n_p T, k, 1]^{\mathrm{T}}, \tag{36}$$

with

$$a_{k,p,q,l} = \alpha_l e^{j2\pi n_p T \nu_l} e^{-j2\pi m_q F \tau_l} e^{-jk\omega_l}. \tag{37}$$

For $l = l'$ we have

$$\begin{aligned}&\Re\left\{\nabla_{\boldsymbol{\theta}_l} H_{k,n_p,m_q} \nabla_{\boldsymbol{\theta}_l}^{\mathrm{H}} H_{k,n_p,m_q}\right\} \\ &= |\alpha_l|^2 \underbrace{\begin{bmatrix} 4\pi^2 m_q^2 F^2 & -4\pi^2 m_q n_p & -2\pi k m_q F & -2\pi m_q F \\ -4\pi^2 m_q n_p & 4\pi^2 n_p^2 T^2 & 2\pi k n_p T & 2\pi n_p T \\ -2\pi k m_q F & 2\pi k n_p T & k^2 & k \\ -2\pi m_q F & 2\pi n_p T & k & 1 \end{bmatrix}}_{\mathbf{A}_{k,p,q}}.\end{aligned}$$

$\mathbf{F}_{l,l}$ is given as

$$\begin{aligned}\mathbf{F}_{l,l} &= \frac{2}{\sigma^2} \sum_{k,p,q} \Re\left\{\nabla_{\boldsymbol{\theta}_l} H_{k,n_p,m_q} \nabla_{\boldsymbol{\theta}_l}^{\mathrm{H}} H_{k,n_p,m_q}\right\} \\ &= \frac{2|\alpha_l|^2}{\sigma^2} \sum_{k,p,q} \mathbf{A}_{k,p,q}.\end{aligned} \tag{38}$$

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.

For $l \neq l'$, we have

$$\Re\left\{\nabla_{\boldsymbol{\theta}_l} H_{k,n_p,m_q} \nabla^{\mathrm{H}}_{\boldsymbol{\theta}_{l'}} H_{k,n_p,m_q}\right\} = \Re\{a_{k,p,q,l} a^*_{k,p,q,l'}\}\mathbf{A}_{k,p,q}, \tag{39}$$

and $\mathbf{F}_{l,l'}$ is thus given as

$$\mathbf{F}_{l,l'} = \sum_{k,p,q} \Re\{a_{k,p,q,l} a^*_{k,p,q,l'}\}\mathbf{A}_{k,p,q}. \tag{40}$$

The CRLB of $\boldsymbol{\theta}_0$ is thus given by[1]

$$\mathbf{R}_{\boldsymbol{\theta}_0} = \mathbf{F}^{-1}_{\boldsymbol{\theta}}[1:4, 1:4]. \tag{41}$$

As we can see, each reflected path will have an impact on the LoS path parameter estimation. The impact is dependent on various system parameters, such as bandwidth, frame length, antenna array size, pulse shape, etc. With huge bandwidth, frame length and antenna array size, the multi-path effect will eventually disappear. We can see this by comparing (38) and (40). In (38), all the components are summarized coherently, while in (40), $\Re\{a_{k,p,q,l} a^*_{k,p,q,l'}\}$ can be positive or negative, leading to much smaller elements in $\mathbf{F}_{l,l'}$.

B. Asymptotic Analysis

From the previous sub-section, we can see that the estimation errors of position-related channel parameters (AoA, delay and Doppler shift associated with the LoS path) are dependent on frame length, bandwidth and antenna array size. When we increase the observation window in any of these three dimensions, the corresponding resolution will be higher, allowing us to more accurately isolate different paths and eliminate interference. Huge bandwidth, large antenna array and long frame length is no longer a dream for massive MIMO-OTFS on mmWave band. In such a case, the interference from the NLoS paths will eventually disappear, and we have

$$\mathbf{R}_{\boldsymbol{\theta}_0} \sim \mathbf{F}^{-1}_{0,0}. \tag{42}$$

To simplify the notation, we can define $\mathbf{A}$ as

$$\mathbf{A} = \sum_{k,p,q} \mathbf{A}_{k,p,q} = KQP \begin{bmatrix} 4\pi^2\overline{m_q^2}F^2 & -4\pi^2\overline{m_q n_p} & -2\pi\overline{k}\overline{m_q}F & -2\pi\overline{m_q}F \\ -4\pi^2\overline{m_q n_p} & 4\pi^2\overline{n_p}^2T^2 & 2\pi\overline{k}\overline{n_p}T & 2\pi\overline{n_p}T \\ -2\pi\overline{k}\overline{m_q}F & 2\pi k\overline{n_p}T & \overline{k^2} & \overline{k} \\ -2\pi\overline{m_q}F & 2\pi\overline{n_p}T & \overline{k} & 1 \end{bmatrix},$$

where $\bar{\cdot}$ denotes the average value of all the elements. When an integer N is very large, we have the following asymptotic results:

$$\sum_{n=0}^{N-1} n^L \sim N^{L+1}/(L+1). \tag{43}$$

Therefore, we have

$$\overline{k} \sim K/2, \quad \overline{k^2} \sim K^2/3. \tag{44}$$

Suppose m_q's and n_p's are uniformly sampled, and we have

$$\overline{m_q} \sim M/2, \quad \overline{n_p} \sim N/2, \quad \overline{m_q^2} \sim M^2/3, \quad \overline{n_p^2} \sim N^2/3. \tag{45}$$

Note that $MF = B$ and $NT = S$ (S denoting the frame length, i.e., spreading in time). Then we have the asymptotic limit of the FIM as

$$\mathbf{F}_{\boldsymbol{\theta}_0} \sim \frac{2|\alpha_0|^2}{\sigma^2} KPQ \begin{bmatrix} \mathbf{H}_0 & \mathbf{C} \\ \mathbf{C}^{\mathrm{T}} & \mathbf{H}_1 \end{bmatrix}, \tag{46}$$

with $\mathbf{H}_0$, $\mathbf{H}_1$ and $\mathbf{C}$ given as

$$\mathbf{H}_0 = \pi^2 \begin{bmatrix} 4B^2/3 & -BS \\ -BS & 4S^2/3 \end{bmatrix}, \quad \mathbf{H}_1 = \begin{bmatrix} K^2/3 & K/2 \\ K/2 & 1 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} -\pi KB/2 & -\pi B \\ \pi KS/2 & \pi S \end{bmatrix}. \tag{47}$$

For τ_0 and ν_0, the covariance matrix of estimation errors is lower bounded by $\mathbf{F}^{-1}_{\boldsymbol{\theta}_0}[1:2, 1:2]$,[2] which is approximately equal to

$$\mathbf{R}_{\tau_0,\nu_0} = \frac{\sigma^2}{2|\alpha_0|^2}\frac{1}{KPQ}\left(\mathbf{H}_0 - \mathbf{C}\mathbf{H}_1^{-1}\mathbf{C}^{\mathrm{T}}\right)^{-1}, \tag{48}$$

and the inverse of $\mathbf{H}_1$ is

$$\mathbf{H}_1^{-1} = \begin{bmatrix} \frac{12}{K^2} & -\frac{6}{K} \\ -\frac{6}{K} & 4 \end{bmatrix}. \tag{49}$$

Then we have

$$\mathbf{C}\mathbf{H}_1^{-1}\mathbf{C}^{\mathrm{T}} = \pi^2 \begin{bmatrix} B^2 & -BS \\ -BS & S^2 \end{bmatrix}, \tag{50}$$

leading to

$$\mathbf{H}_0 - \mathbf{C}\mathbf{H}_1^{-1}\mathbf{C}^{\mathrm{T}} = \pi^2 \begin{bmatrix} B^2/3 & 0 \\ 0 & S^2/3 \end{bmatrix}. \tag{51}$$

The CRLB of τ_0 and ν_0 is thus simplified as

$$\mathbf{R}_{\tau_0,\nu_0} = \frac{\sigma^2}{2|\alpha_0|^2}\frac{3}{\pi^2}\frac{1}{KPQ}\begin{bmatrix} 1/B^2 & 0 \\ 0 & 1/S^2 \end{bmatrix}. \tag{52}$$

The diagonal elements of $\mathbf{R}_{\tau_0,\nu_0}$ give us the CRLBs of estimation errors of propagation delay and Doppler shift. Now let's move to the AoA estimation error. Specifically, the covariance matrix of ω_0 and ϕ_0 is lower bounded by

$$\mathbf{R}_{\omega_0,\phi_0} = \frac{\sigma^2}{2|\alpha_0|^2}\frac{1}{KPQ}\left(\mathbf{H}_1 - \mathbf{C}^{\mathrm{T}}\mathbf{H}_0^{-1}\mathbf{C}\right)^{-1}. \tag{53}$$

The inverse of $\mathbf{H}_0$ is given as

$$\mathbf{H}_0^{-1} = \frac{3}{7\pi^2}\begin{bmatrix} \frac{4}{B^2} & \frac{3}{BS} \\ \frac{3}{BS} & \frac{4}{S^2} \end{bmatrix}. \tag{54}$$

The estimation error of AoA is lower bounded by the first diagonal element of $\mathbf{R}_{\omega_0,\phi_0}$. Through tedious by straightforward derivations, we can get the lower bound as

$$R_{\omega_0} = \frac{\sigma^2}{2|\alpha_0|^2}\frac{12}{K^3PQ} = \frac{\sigma^2}{|\alpha_0|^2}\frac{6}{K^3PQ}. \tag{55}$$

Remarks:

[1]For a scalar variable, the CRLB of estimation error is also a scalar, and it is the lower bound of the variance of any unbiased estimate of the variable. For a vector variable, the CRLB of the estimation error is a matrix. For example, consider an arbitrary unbiased estimate of $\boldsymbol{\theta}_0$ as $\hat{\boldsymbol{\theta}}_0$, while the CRLB is $\mathbf{R}_{\boldsymbol{\theta}_0}$. The CRLB gives the "lower bound" of covariance matrix in the sense that $\mathbb{E}\{(\hat{\boldsymbol{\theta}}_0 - \boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}}_0 - \boldsymbol{\theta}_0)^{\mathrm{T}}\} - \mathbf{R}_{\boldsymbol{\theta}_0}$ is positive semi-definite, i.e., $\mathbb{E}\{(\hat{\boldsymbol{\theta}}_0 - \boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}}_0 - \boldsymbol{\theta}_0)^{\mathrm{T}}\} \succeq \mathbf{R}_{\boldsymbol{\theta}_0}$.

[2]Throughout the paper, $\mathbf{A}^{-1}[r_a : r_b, c_a : c_b]$ denotes a sub-matrix of $\mathbf{A}^{-1}$, ranging from row r_a to r_b and column c_a to c_b.

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.

- BS gives the product of bandwidth and frame length, while $PQ = PF \times QT$ gives the product of bandwidth and time assigned for channel estimation/localization.
- The AoA estimation error is inversely proportional to the amount of time-frequency resources assigned for channel estimation/localization.
- The CRLBs of propagation delay and Doppler shift, i.e., R_{τ_0} and R_{ν_0}, are dependent on the total bandwidth and frame length, respectively. The product of their CRLBs is inversely proportional to the product of bandwidth and frame length:

$$R_{\tau_0} R_{\nu_0} \propto \frac{3}{\pi^2 KPQ(BS)^2}. \tag{56}$$

From here we can clearly see the trade-off between the estimation accuracy of ToA and Doppler shift. For a given resource block, i.e., fixed BS, we can increase B to improve the ToA estimation accuracy, but this will inevitably lead to the reduced estimation accuracy of Doppler shift. Note that both ToA and Doppler shift contribute to positioning accuracy, a *good* strategy should be identified to minimize the positioning error. For this purpose, we need to investigate the contributions of ToA and Doppler shift on positioning accuracy, covered in the following sub-section.

In the above analyses, the spatial wideband effect was ignored. To quantify the impact of the beam squint on positioning performance, we need to use the model in (25), and re-derive the CRLBs by following the previous steps. Specifically, the gradient in (36) should be revised as

$$\begin{aligned}&\nabla_{\boldsymbol{\theta}_l} H_{k,n_p,m_q} \\ &\quad = j a_{k,p,q,l} [-2\pi m_q F, 2\pi n_p T, k(1 + m_q F/f_c), 1]^{\mathrm{T}}.\end{aligned} \tag{57}$$

Elements in the third column/row of the original FIM should be revised accordingly, and we can get (58), as shown at the bottom of the next page, with $r = MF/f_c = B/f_c$ denoting the bandwidth-carrier frequency ratio, which is generally very small.

We can clearly see that $\mathbf{F}_{\boldsymbol{\theta}}^{(SW)} \succeq \mathbf{F}_{\boldsymbol{\theta}}$. That is to say, the spatial wideband effect actually brings more information concerning the channel parameters. However, the increment is negligible. For example, with $f_c = 30\,\mathrm{GHz}$, and $B = 10\,\mathrm{MHz}$, we have $r = 1/3000$. By considering the spatial wideband effect, we won't be able to extract more non-negligible information concerning the channel parameters. Nonetheless, ignoring the spatial wideband effect will induce huge biases, as we shall see in the numerical results.

C. Positioning Error

It is generally assumed that only the LoS (Line of Sight) path contains the target's position information. Therefore, our objective is to extract the position information of the mobile device from τ_0, ν_0, and ω_0. To simplify the notations, we define $\boldsymbol{\psi} = [\tau_0, \nu_0, \omega_0]^{\mathrm{T}}$.

For positioning, we need to build a coordinate system. Suppose the mobile device is located at $\mathbf{x} = [x, y, h]^{\mathrm{T}}$, with (x, y) denoting the 2D horizontal coordinate and h denoting the height. Similarly, $\mathbf{x}_{BS}$ denotes the 3D position of the BS. Suppose $\mathbf{v}$ is the 3D velocity of the mobile device, and we have the following results:

$$\begin{aligned}\tau_0 &= \|\mathbf{x} - \mathbf{x}_{BS}\|/c,\ \nu_0 = \frac{(\mathbf{x} - \mathbf{x}_{BS})^{\mathrm{T}}\mathbf{v}}{\|\mathbf{x} - \mathbf{x}_{BS}\| c} f_c, \\ \omega_0 &= \pi \cos\theta_0 = -\pi \frac{(\mathbf{x} - \mathbf{x}_{BS})^{\mathrm{T}}\mathbf{a}}{\|\mathbf{x} - \mathbf{x}_{BS}\|},\end{aligned} \tag{59}$$

where $\mathbf{a}$ is a unit vector pointing from the 0-th element to the $K-1$-th element of the BS antenna array. The second row of (59) explicitly gives the relationship between Doppler shift and velocity of the mobile device.

Without loss of generality, we assume the BS is located at the origin and the FIM concerning $\mathbf{x}$ is

$$\mathbf{F}_{\mathbf{x}} = \nabla_{\mathbf{x}}^{\mathrm{T}} \boldsymbol{\psi} \mathbf{F}_{\boldsymbol{\psi}} \nabla_{\mathbf{x}} \boldsymbol{\psi}, \tag{60}$$

with $\mathbf{F}_{\boldsymbol{\psi}}$ denoting the equivalent FIM (EFIM) concerning $\boldsymbol{\psi}$. For $\boldsymbol{\psi}$, the CRLB of estimation error is $\mathbf{R}_{\boldsymbol{\psi}} = \mathbf{R}_{\boldsymbol{\theta}_0}[1:3, 1:3]$, and the EFIM is $\mathbf{F}_{\boldsymbol{\psi}} = \mathbf{R}_{\boldsymbol{\psi}}^{-1}$.

The Jacobian in (60) is given as

$$\nabla_{\mathbf{x}} \boldsymbol{\psi} = \begin{bmatrix} \boldsymbol{\rho}_{\mathbf{x}}^{\mathrm{T}}/c \\ \mathbf{v}^{\mathrm{T}}(\mathbf{I} - \boldsymbol{\rho}_{\mathbf{x}}\boldsymbol{\rho}_{\mathbf{x}}^{\mathrm{T}}) f_c/d/c \\ \pi \mathbf{a}^{\mathrm{T}}(\mathbf{I} - \boldsymbol{\rho}_{\mathbf{x}}\boldsymbol{\rho}_{\mathbf{x}}^{\mathrm{T}})/d \end{bmatrix}, \tag{61}$$

with $d = \|\mathbf{x} - \mathbf{x}_{BS}\| = \|\mathbf{x}\|$ denoting the distance between the BS and mobile device, and $\boldsymbol{\rho}_{\mathbf{x}} = \mathbf{x}/d$.

When the antenna array at the BS is vertically placed, we will be able to achieve 3D positioning with a ULA, i.e., $\mathbf{a} = [0, 0, 1]^{\mathrm{T}}$. As a matter of fact, as long as $\mathbf{a}$ is not aligned with $\mathbf{v}$ or $\boldsymbol{\rho}_{\mathbf{x}}$, $\nabla_{\mathbf{x}}\boldsymbol{\psi}$ will have full rank, and we will be able to achieve 3D positioning with the measured τ_0, ν_0, and ω_0.

The CRLB of positioning error is

$$\mathbf{R}_{\mathbf{x}} = (\nabla_{\mathbf{x}}\boldsymbol{\psi})^{-1} \mathbf{R}_{\boldsymbol{\psi}} (\nabla_{\mathbf{x}}\boldsymbol{\psi})^{-\mathrm{T}}, \tag{62}$$

where $\mathbf{A}^{-\mathrm{T}} = \left(\mathbf{A}^{\mathrm{T}}\right)^{-1}$. Note that $\mathbf{R}_{\boldsymbol{\psi}}$ is diagonal, and we have the CRLB of positioning error as

$$\sigma_{\mathbf{x}}^2 = \mathrm{tr}\{\mathbf{R}_{\mathbf{x}}\} = \mathrm{tr}\left\{\mathbf{R}_{\boldsymbol{\psi}} \left(\nabla_{\mathbf{x}}\boldsymbol{\psi}\nabla_{\mathbf{x}}^{\mathrm{T}}\boldsymbol{\psi}\right)^{-1}\right\}. \tag{63}$$

Then, we define $\mathbf{D} = \mathrm{diag}\,[1/c, f_c/d/c, \pi/d]$ and $\mathbf{J} = [\boldsymbol{\rho}_{\mathbf{x}}, \mathbf{v}_t, \mathbf{a}_t]$ with

$$\begin{aligned}\mathbf{v}_t &= \mathbf{v}^{\mathrm{T}}(\mathbf{I} - \boldsymbol{\rho}_{\mathbf{x}}\boldsymbol{\rho}_{\mathbf{x}}^{\mathrm{T}}), \\ \mathbf{a}_t &= \mathbf{a}^{\mathrm{T}}(\mathbf{I} - \boldsymbol{\rho}_{\mathbf{x}}\boldsymbol{\rho}_{\mathbf{x}}^{\mathrm{T}}).\end{aligned} \tag{64}$$

Thus we have

$$\nabla_{\mathbf{x}}\boldsymbol{\psi} = \mathbf{D}\mathbf{J}^{\mathrm{T}}, \tag{65}$$

and

$$\nabla_{\mathbf{x}}\boldsymbol{\psi}\nabla_{\mathbf{x}}^{\mathrm{T}}\boldsymbol{\psi} = \mathbf{D}\mathbf{J}^{\mathrm{T}}\mathbf{J}\mathbf{D}, \tag{66}$$

with $\mathbf{J}^{\mathrm{T}}\mathbf{J}$ given by

$$\mathbf{J}^{\mathrm{T}}\mathbf{J} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \|\mathbf{v}_t\|^2 & \mathbf{v}_t^{\mathrm{T}}\mathbf{a}_t \\ 0 & \mathbf{v}_t^{\mathrm{T}}\mathbf{a}_t & \|\mathbf{a}_t\|^2 \end{bmatrix}. \tag{67}$$

The CRLB of positioning error can then be rewritten as

$$\begin{aligned}\sigma_{\mathbf{x}}^2 &= \mathrm{tr}\{\mathbf{R}_{\mathbf{x}}\} = \mathrm{tr}\left\{\mathbf{R}_{\boldsymbol{\psi}}\mathbf{D}^{-1}\left(\mathbf{J}^{\mathrm{T}}\mathbf{J}\right)^{-1}\mathbf{D}^{-1}\right\} \\ &= \mathrm{tr}\left\{\mathbf{R}_{\boldsymbol{\psi}}\mathbf{D}^{-2}\left(\mathbf{J}^{\mathrm{T}}\mathbf{J}\right)^{-1}\right\}.\end{aligned} \tag{68}$$

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.

Define $v_t = \|\mathbf{v}_t\|$, $a_t = \|\mathbf{a}_t\|$, and $c_{v,a} = \mathbf{v}_t^{\mathrm{T}}\mathbf{a}_t/(v_t a_t)$. The diagonal elements of the inverse of $\mathbf{J}^{\mathrm{T}}\mathbf{J}$ are

$$\operatorname{diag}\left\{\left(\mathbf{J}^{\mathrm{T}}\mathbf{J}\right)^{-1}\right\} = \begin{bmatrix} 1 \\ \left(v_t^2(1-c_{v,a}^2)\right)^{-1} \\ \left(a_t^2(1-c_{v,a}^2)\right)^{-1} \end{bmatrix}. \tag{69}$$

At this point, we can obtain the CRLB of positioning error as

$$\sigma_{\mathbf{x}}^2 = \frac{\sigma^2}{2|\alpha_0|^2}\frac{3}{KPQ\pi^2}\left(\frac{c^2}{B^2} + \frac{d^2c^2}{f_c^2S^2v_t^2}\frac{1}{(1-c_{v,a}^2)} + \frac{4d^2}{K^2}\frac{1}{a_t^2(1-c_{v,a}^2)}\right). \tag{70}$$

Here $Sv_t/(c/f_c/2)$ is the equivalent length of the synthesized array through the Doppler effect. It is clear that as long as $\mathbf{v}_t$ and $\mathbf{a}_t$ are not aligned, i.e., $c_{v,a} \neq \pm 1$, we will be able to achieve 3D positioning with a ULA.

From the above discussions, we can see that the positioning error is not only dependent on system parameters, but also the vehicle's position and velocity. The system parameters include SNR, bandwidth, frame length, antenna array size. During system design, we can optimize system performance by carefully choosing these system parameters. However, the mobile device's position and velocity are out of our control. As a result, we need to consider the worst case during system design, so that the desired accuracy can be achieved no matter where the device is located in the covered area.

IV. Channel Estimation and Localization

A. Maximum Likelihood Estimate of the Channel Parameters

With L paths, there are totally $4L$ channel parameters, and the MLE (Maximum Likelihood Estimate) of the parameters can be obtained by solving the following problem:

$$\min_{\tilde{\alpha}_l,\tilde{\nu}_l,\tilde{\tau}_l,\tilde{\omega}_l} \sum_{k,p,q}\left|\sum_l \tilde{\alpha}_l e^{j2\pi n_p T\tilde{\nu}_l} e^{-j2\pi m_q F\tilde{\tau}_l} e^{-jk\tilde{\omega}_l} - \hat{H}_{k,n_p,m_q}\right|^2.$$

The above problem involves minimizing a quadratic function of $\tilde{\alpha}_l$'s. Therefore, for any given $\tilde{\nu}_l$, $\tilde{\tau}_l$ and $\tilde{\omega}_l$, we can find the corresponding optimal $\tilde{\alpha}_l$ to minimize the cost function. To simplify the notation, we define

$$\hat{\boldsymbol{\alpha}} = [\tilde{\alpha}_0, \tilde{\alpha}_1, \cdots, \tilde{\alpha}_{L-1}]^{\mathrm{T}}$$

$$\tilde{\mathbf{s}}_{k,p,q} = \begin{bmatrix} e^{j2\pi n_p T\tilde{\nu}_0} e^{-j2\pi m_q F\tilde{\tau}_0} e^{-jk\tilde{\omega}_0} \\ e^{j2\pi n_p T\tilde{\nu}_1} e^{-j2\pi m_q F\tilde{\tau}_1} e^{-jk\tilde{\omega}_1} \\ \vdots \\ e^{j2\pi n_p T\tilde{\nu}_{L-1}} e^{-j2\pi m_q F\tilde{\tau}_{L-1}} e^{-jk\tilde{\omega}_{L-1}} \end{bmatrix}^*.$$

Then we can rewrite the optimization problem as

$$\min_{\tilde{\alpha}_l,\tilde{\nu}_l,\tilde{\tau}_l,\tilde{\omega}_l} \sum_{k,p,q}\left|\tilde{\mathbf{s}}_{k,p,q}^{\mathrm{H}}\tilde{\boldsymbol{\alpha}} - \hat{H}_{k,n_p,m_q}\right|^2. \tag{71}$$

By taking the first-order derivative we can obtain

$$\sum_{k,p,q}\tilde{\mathbf{s}}_{k,p,q}\left(\tilde{\mathbf{s}}_{k,p,q}^{\mathrm{H}}\tilde{\boldsymbol{\alpha}} - \hat{H}_{k,n_p,m_q}\right) = \mathbf{0}. \tag{72}$$

Suppose $\tilde{\mathbf{S}} \in \mathcal{C}^{L\times KPQ}$ contains all the $\tilde{\mathbf{s}}_{k,p,q}$'s, while $\hat{\mathbf{h}} \in \mathcal{C}^{KPQ\times 1}$ contains all the $\hat{H}_{k,n_p,m_q}$, we have

$$\tilde{\mathbf{S}}\tilde{\mathbf{S}}^{\mathrm{H}}\tilde{\boldsymbol{\alpha}} = \tilde{\mathbf{S}}\hat{\mathbf{h}} \Rightarrow \tilde{\boldsymbol{\alpha}} = \left(\tilde{\mathbf{S}}\tilde{\mathbf{S}}^{\mathrm{H}}\right)^{-1}\tilde{\mathbf{S}}\hat{\mathbf{h}}. \tag{73}$$

Then the cost function is revised as

$$\begin{aligned}\min_{\tilde{\alpha}_l,\tilde{\nu}_l,\tilde{\tau}_l,\tilde{\omega}_l} &\sum_{k,p,q}\left|\tilde{\mathbf{s}}_{k,p,q}^{\mathrm{H}}\tilde{\boldsymbol{\alpha}} - \hat{H}_{k,n_p,m_q}\right|^2 \\ &= \min_{\tilde{\nu}_l,\tilde{\tau}_l,\tilde{\omega}_l}\left\|\tilde{\mathbf{S}}^{\mathrm{H}}\left(\tilde{\mathbf{S}}\tilde{\mathbf{S}}^{\mathrm{H}}\right)^{-1}\tilde{\mathbf{S}}\hat{\mathbf{h}} - \mathbf{h}\right\|^2 \\ &= \min_{\tilde{\nu}_l,\tilde{\tau}_l,\tilde{\omega}_l}\hat{\mathbf{h}}^{\mathrm{H}}\left(\mathbf{I} - \tilde{\mathbf{S}}^{\mathrm{H}}\left(\tilde{\mathbf{S}}\tilde{\mathbf{S}}^{\mathrm{H}}\right)^{-1}\tilde{\mathbf{S}}\right)\hat{\mathbf{h}},\end{aligned} \tag{74}$$

where $\mathbf{h} = \mathbb{E}\{\hat{\mathbf{h}}\}$ contains the true channel gains. Basically, we want to minimize the energy of $\hat{\mathbf{h}}$ on the null-space of $\tilde{\mathbf{S}}$. Equivalently, we have

$$\begin{aligned}&\arg\min_{\tilde{\nu}_l,\tilde{\tau}_l,\tilde{\omega}_l}\hat{\mathbf{h}}^{\mathrm{H}}\left(\mathbf{I} - \tilde{\mathbf{S}}^{\mathrm{H}}\left(\tilde{\mathbf{S}}\tilde{\mathbf{S}}^{\mathrm{H}}\right)^{-1}\tilde{\mathbf{S}}\right)\hat{\mathbf{h}}. \\ &\quad = \arg\max_{\tilde{\nu}_l,\tilde{\tau}_l,\tilde{\omega}_l}\hat{\mathbf{h}}^{\mathrm{H}}\tilde{\mathbf{S}}^{\mathrm{H}}\left(\tilde{\mathbf{S}}\tilde{\mathbf{S}}^{\mathrm{H}}\right)^{-1}\tilde{\mathbf{S}}\hat{\mathbf{h}}.\end{aligned} \tag{75}$$

Asymptotically, we have $\tilde{\mathbf{S}}\tilde{\mathbf{S}}^{\mathrm{H}} \sim KPQ\mathbf{I}$, i.e., the vectors corresponding to different paths are almost orthogonal for large bandwidth, antenna array size and long frame, as has been shown in the Appendix. As a result, the solution of the above optimization problem is very close to that of the following problem:

$$\max_{\tilde{\nu}_l,\tilde{\tau}_l,\tilde{\omega}_l}\hat{\mathbf{h}}^{\mathrm{H}}\tilde{\mathbf{S}}^{\mathrm{H}}\tilde{\mathbf{S}}\hat{\mathbf{h}} = \sum_l \max_{\tilde{\nu}_l,\tilde{\tau}_l,\tilde{\omega}_l}\left|\hat{\mathbf{h}}^{\mathrm{H}}\tilde{\mathbf{s}}_l\right|^2, \tag{76}$$

with $\mathbf{s}_l$ denoting the l-th column of $\tilde{\mathbf{S}}^{\mathrm{H}}$, containing all the complex channel gains attached to the l-th path. This can be done through FFT-based exhaustive search in the 3D space. The complexity of the problem in (76) is much lower than that of (74), because we can estimate the channel parameters of different paths sequentially, instead of jointly. For joint optimization, the complexity grows exponentially with the number of paths, while that of the sequential optimization only grows linearly. The fundamental reason for the possibility of reducing the complexity is rooted in the fact that the vectors

$$\mathbf{F}_{\boldsymbol{\theta}}^{(SW)} \sim \frac{2|\alpha_0|^2}{\sigma^2}KQP\begin{bmatrix} 4\pi^2B^2/3 & -\pi^2BS & -(1+2r/3)\pi KB/2 & -\pi B \\ -\pi^2BS & \pi^2S^2 & (1+r/2)\pi KS/2 & \pi S \\ -(1+2r/3)\pi KB/2 & (1+r/2)\pi KS/2 & (1+r^2/3+r)K^2/3 & (1+r)K/2 \\ -\pi B & \pi S & (1+r)K/2 & 1 \end{bmatrix} \tag{58}$$

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.

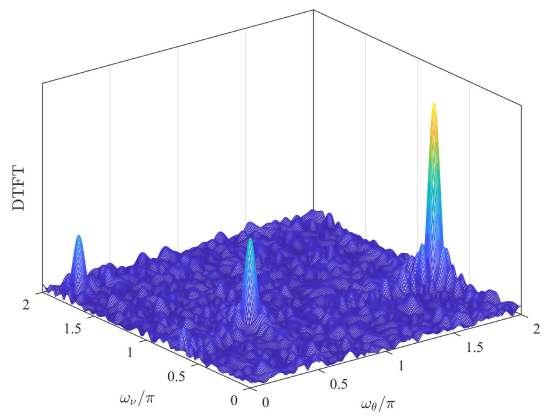


Fig. 3. Two dimensional DTFT.

corresponding to different paths are asymptotically orthogonal. For positioning, we are only interested in the LoS path, and the complexity is not even dependent on the number of paths. However, the sequential estimate of all the channel parameters will be necessary for mapping.

B. 3D Search With FFT

In this case we need to conduct 3D search through FFT, i.e., delay, Doppler shift and AoA. This can be done sequentially, and we should start from Doppler shift and AoA. The reason is that these two domains generally have higher resolutions, and multi-path interference can be better isolated. To start with, the measured channel matrix on the m_q-th sub-carrier is

$$\hat{\mathbf{H}}_{m_q} = \sum_{l=0}^{L-1} \alpha_{l,m_q} \mathbf{a}_{\theta_l} \mathbf{a}_{\nu_l}^{\mathrm{H}} + \mathbf{N}_{m_q} \tag{77}$$

with $\mathbf{N}_{m_q} \in \mathcal{C}^{K\times P}$ denoting the additive white Gaussian noise, and α_{l,m_q} given as

$$\alpha_{l,m_q} = \alpha_l e^{-j2\pi m_q F \tau_l}. \tag{78}$$

$\mathbf{a}_\theta$ and $\mathbf{a}_\nu$ are steering vectors given as

$$\mathbf{a}_\theta[k] = e^{-jk\pi\cos\theta}, \quad \mathbf{a}_\nu[p] = e^{-j2\pi n_p T\nu}. \tag{79}$$

Then we can estimate AoA and Doppler shift by analyzing the 2D spectrum of $\hat{\mathbf{H}}_{m_q}$ through 2D FFT, presented in [6]. Specifically, we can get the 2D DTFT (Discrete-Time Fourier Transform) of $\hat{\mathbf{H}}_{m_q}$, and the 2D FFT gives the sampled version of that. The 2D DTFT is illustrated in Fig. 3, where each path corresponds to a 2D sinc function.

Suppose different paths are well separated, and we are only interested in the LoS path. The 2D DFT is given as

$$\mathbf{H}_{\omega,m_q} = \mathbf{F}_K^{\mathrm{H}} \hat{\mathbf{H}}_{m_q} \mathbf{F}_P, \tag{80}$$

with $\mathbf{F}_K$ and $\mathbf{F}_P$ denoting the Fourier matrices of dimension K and P, respectively.

The peak of the LoS signal is located at $\omega_\nu = 2\pi(N/P)T\nu_0$ and $\omega_\theta = \pi\cos\theta_0$. The corresponding sinc function has four samples in the mainlobe. Suppose M/Q and N/P are both integers, and we have $m_q = qM/Q$ and $n_p = pN/P$. Thus, the indexes of the mainlobe samples are (l_θ, l_ν), $(l_\theta+1, l_\nu)$, $(l_\theta, l_\nu+1)$, $(l_\theta+1, l_\nu+1)$, with $l_\theta \in \{0,1,\cdots,K-1\}$ and $l_\nu \in \{0,1,\cdots,P-1\}$. Approximately, we have

$$\begin{aligned}
\mathbf{h}_{0,0}[m_q] &= \mathbf{H}_{\omega,m_q}[l_\theta, l_\nu] \\
\mathbf{h}_{0,1}[m_q] &= \mathbf{H}_{\omega,m_q}[l_\theta, l_\nu+1] \\
\mathbf{h}_{1,0}[m_q] &= \mathbf{H}_{\omega,m_q}[l_\theta+1, l_\nu] \\
\mathbf{h}_{1,1}[m_q] &= \mathbf{H}_{\omega,m_q}[l_\theta+1, l_\nu+1].
\end{aligned} \tag{81}$$

There must exist $\beta_\nu, \beta_\theta \in [0,1)$, so that $\omega_\theta = (l_\theta+\beta_\theta)\frac{2\pi}{K}$ and $\omega_\nu = (l_\nu+\beta_\nu)\frac{2\pi}{P}$. Then we approximately have

$$\begin{aligned}
\mathbf{h}_{0,1}[q] &\approx -\frac{\beta_\nu}{1-\beta_\nu} e^{-j\pi/P} \mathbf{h}_{0,0}[q] \\
\mathbf{h}_{1,0}[q] &\approx -\frac{\beta_\theta}{1-\beta_\theta} e^{-j\pi/K} \mathbf{h}_{0,0}[q] \\
\mathbf{h}_{1,1}[q] &\approx \frac{\beta_\nu\beta_\theta}{(1-\beta_\nu)(1-\beta_\theta)} e^{-j(\pi/P+\pi/K)} \mathbf{h}_{0,0}[q].
\end{aligned} \tag{82}$$

Meanwhile, we have

$$\mathbf{h}_{0,0}[q] = \mathbf{h}_{0,0}[0] e^{-j2\pi qMF/Q\tau_0} = \mathbf{h}_{0,0}[0] e^{-j2\pi B\tau_0 q/Q}, \tag{83}$$

which gives another complex sinusoid with a phase shift of $\omega_\tau = 2\pi B\tau_0/Q$, and we can employ IDFT to process the sequence $\mathbf{h}_{0,0}$ along the sub-carrier index and obtain $\mathbf{h}_\omega^{(0,0)}$, sampled from another sinc function with two samples in the mainlobe. Similarly, there must exist integer l_τ and $\beta_\tau \in [0,1)$, so that $\omega_\tau = (l_\tau+\beta_\tau)\frac{2\pi}{Q}$. The two samples in the mainlobe will be

$$\begin{aligned}
\mathbf{h}_\omega^{(0,0)}[l_\tau] &= \mathbf{h}_{0,0}[0]\sqrt{Q} e^{-j(Q-1)\beta_\tau\pi/N} \frac{\sin(\beta_\tau\pi)}{\beta_\tau\pi} \\
\mathbf{h}_\omega^{(0,0)}[l_\tau+1] &= \mathbf{h}_{0,0}[0] \frac{-\beta_\tau}{1-\beta_\tau} e^{-j\pi/Q}.
\end{aligned} \tag{84}$$

Then we get eight samples that contain most of the energy of the signal, which forms a tensor:

$$\mathcal{H}(n_\theta, n_\nu, n_\tau) = \left|\mathbf{h}_\omega^{(n_\theta,n_\nu)}[l_\tau+n_\tau]\right|, \tag{85}$$

with $n_\theta, n_\nu, n_\tau \in \{0,1\}$. The propagation delay, AoA and Doppler shift can all be estimated from these eight samples in very similar ways, and we will take the estimate of the propagation delay as an example. For any combination of n_ν and n_θ, we can estimate the fractional part of propagation delay (i.e., β_τ) as

$$\hat{\beta}_\tau^{(n_\nu,n_\theta)} = \frac{\mathcal{H}(n_\theta, n_\nu, 1)}{\mathcal{H}(n_\theta, n_\nu, 0) + \mathcal{H}(n_\theta, n_\nu, 1)}. \tag{86}$$

There are four combinations, and thus four estimates. We need to combine them properly, based on their accuracy. Specifically, we can combine them with the weights given by

$$a_{n_\nu,n_\theta} = \frac{\mathcal{H}(n_\theta, n_\nu, 0)^2 + \mathcal{H}(n_\theta, n_\nu, 1)^2}{\sum_{n_\theta,n_\nu,n_\tau} \mathcal{H}(n_\theta, n_\nu, n_\tau)^2}. \tag{87}$$

Then estimate of β_τ is updated as

$$\hat{\beta}_\tau = \sum_{n_\theta,n_\nu} a_{n_\nu,n_\theta} \hat{\beta}_\tau^{(n_\nu,n_\theta)}. \tag{88}$$

The ω_τ can be estimated as $\hat{\omega}_\tau = (l_\tau+\hat{\beta}_\tau)\frac{2\pi}{Q}$, and the propagation delay can thus be estimated as $\hat{\tau}_0 = Q\hat{\omega}_\tau/(2\pi B)$.

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.

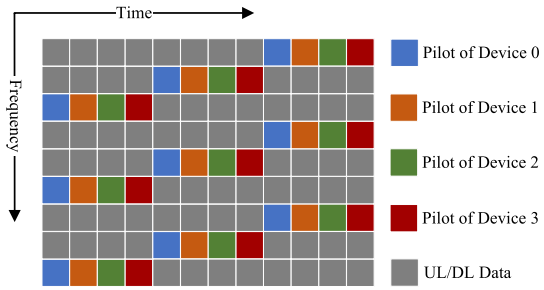


Fig. 4. One way to allocate resources for channel estimation and data transmission in multi-user MIMO-OTFS.

Similarly, we can estimate the AoA and Doppler shift. This algorithm was first proposed in [66]. Two applications of this algorithm for positioning and channel tracking are given in [6] and [67], respectively. Reference [6] actually extends the algorithm to 2D cases.

C. Localization

Suppose the geometrical channel parameters are estimated as $\hat{\tau}_0$, $\hat{\nu}_0$ and $\hat{\omega}_0$, respectively. The next step is to extract the position information. The non-linear weighted LS estimate of the mobile device's position can be obtained by solving the following optimization problem with Newton's method:

$$\min_{\hat{\mathbf{x}}} \frac{(\hat{\tau}_0 - \tau_0(\hat{\mathbf{x}}))^2}{\sigma_\tau^2} + \frac{(\hat{\nu}_0 - \nu_0(\hat{\mathbf{x}}))^2}{\sigma_\nu^2} + \frac{(\hat{\omega}_0 - \omega_0(\hat{\mathbf{x}}))^2}{\sigma_\omega^2}, \tag{89}$$

where $\tau_0(\hat{\mathbf{x}})$, $\nu_0(\hat{\mathbf{x}})$, $\omega_0(\hat{\mathbf{x}})$ denote the calculated values of delay, Doppler shift, and spatial signature, based on the estimated position. Their corresponding errors are at drastically different scales, and we need to normalize them with the corresponding variances, i.e., σ_τ^2, σ_ν^2 and σ_ω^2, respectively. The solution of the optimization problem gives the non-linear weighted least-squares estimate of the mobile device's position. It also happens to be the maximum likelihood estimate, given that the estimation errors of three geometrical channel parameters are uncorrelated zero-mean Gaussian random variables.

D. Multi-User Scenario

For multiple single-antenna users, suppose all the users are distinguishable in space, and we can allow them to share the same time-frequency resources through the space division multiple access. The received signal of the k-th antenna at the BS is

$$Y_{k,n,m} = \sum_{u=0}^{U-1} H_{k,n,m}^{(u)} X_{n,m}^{(u)} + N_{k,n,m}, \tag{90}$$

where $X_{n,m}^{(u)}$ and $H_{k,n,m}^{(u)}$ are the transmitted symbol of the u-th user at the (n,m)-th slot and the corresponding channel gain, while $N_{k,n,m}$ is the additive white Gaussian noise. Part of the time-frequency resources will be used for channel estimation. One possible way to do this is illustrated in Fig. 4

Without loss of generality, suppose there are four users and each user needs nine time-frequency slots for channel estimation. In such cases, to avoid inter-user interference, we need to assign orthogonal time-frequency resources to different users, and one strategy is shown in Fig. 4. In such cases, the channel estimation of different users are conducted independently and there is no mutual interference at all in the channel estimation stage. Following that, all the users will need to share the same time-frequency resources for data transmission through the spatial division multiplexing ("Uplink/Downlink Data" in Fig. 4, denoted in grey). Different beamforming/combining strategies will then be employed to suppress inter-user interference, such as the matched filter, decorrelator, or LMMSE detector [65].

V. Numerical Evaluations and Discussion

In this section, we will conduct simulations to verify the previously derived theoretical results. Consider $f_c = 30\,\text{GHz}$, and a mobile device mounted on a vehicle, with a maximum speed of 100 m/s, corresponding to a Doppler spread of 20 kHz. Suppose the delay spread is $0.5\,\mu$s, corresponding to a distance of 150 m. Based on previous discussions, we can take $T = 5\,\mu$s and $F = 200\,\text{kHz}$. In this case, we need to make sure that $MN = BS \ll 3 \times 10^6$, and we can afford $MN = 1 \times 10^5$. QPSK modulation is employed throughout the simulations. One percent of a resource block is utilized for channel estimation and positioning, i.e., $P/N = Q/M = 0.1$. The proposed system will be compared with existing work that ignores the spatial wideband effect and only uses square waveform for modulation.

A. Multi-Path Effect

In this part, we are going to evaluate the theoretical results and algorithm performance in multi-path environment. Consider five paths, one LoS path and four reflected paths. The Doppler shifts and AoAs of paths are uniformly distributed, while the propagation delays follow exponential distribution. The average power of reflected paths is ten times weaker than the LoS path. Fig. 5 shows the channel estimation errors and positioning errors for different K and SNR. The SNR is defined based on the channel measurement, i.e., $\hat{\mathbf{H}}$ in (28), given as $\mathbb{E}\{\|\mathbf{H}\|_F^2/(KPQ\sigma^2)\}$. For an arbitrary channel parameter x, the mean square error (MSE) of the estimated parameter (i.e., $\hat{x}$) is defined as $\mathbb{E}\{|\hat{x} - x|^2\}$. The root MSE (RMSE) is the root of the MSE.

In Fig. 5, the number of antennas at the BS varies in $\{32, 64, 128\}$. The total bandwidth is 200 MHz, and the frame length is 0.5 ms. As we can see, the estimation errors of Doppler shift and AoA are very close to the CRLB. The multi-path effect can almost be ignored. The reason is that the signals from different paths are almost orthogonal in the delay-Doppler-space domain, and the 3D FFT-based algorithm can identify the LoS signal and eliminate the others.

Based on the results, we can clearly see the advantages of the proposed work over existing work. First, for such wide bandwidth, the AoA cannot be accurately estimated if the beam squint is ignored, as we can see in the third and forth sub-figures, where "BSI" is short for "Beam Squint Ignored". When beam squint is ignored, the AoAs on different

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.

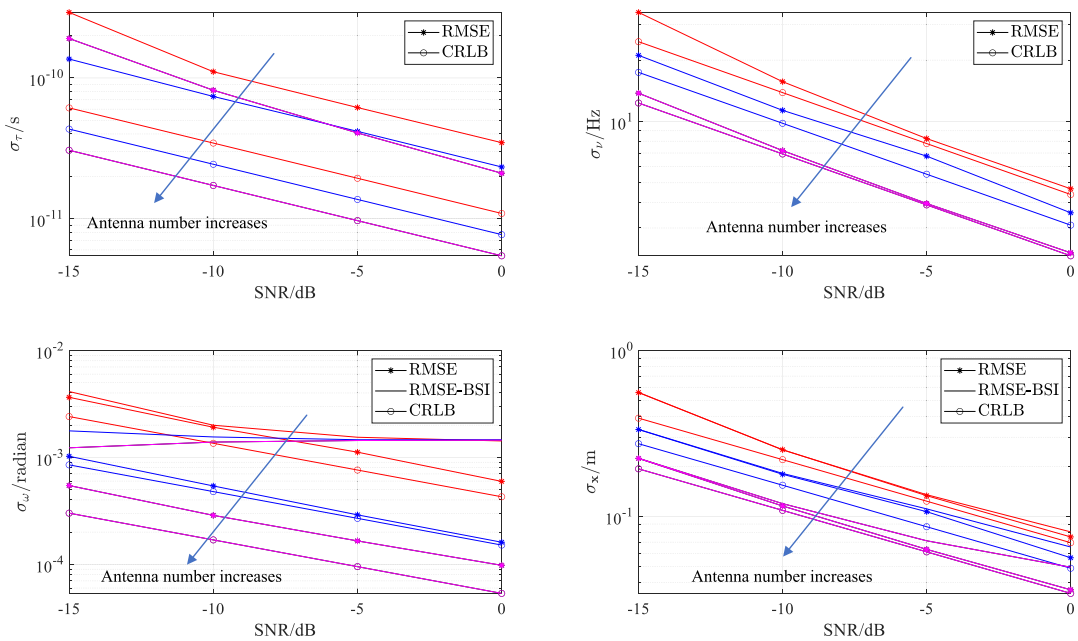


Fig. 5. Channel estimation error and positioning accuracy for different K and SNR.

sub-carriers will be estimated independently, and combined by taking the average. This will inevitably leads to a bias at the presence of beam squint. From the modelling, we can see that the spatial signatures on different sub-carriers are changing linearly over the sub-carrier index, and we can easily get the maximum likelihood estimate of the AoA by fitting the observed AoAs on different sub-carriers with a linear model. By doing so, the AoA estimation accuracy can be significantly improved. In terms of positioning error, we can see that the beam squint is gradually manifesting itself when SNR increases. Second, the theoretical results are very accurate, which verifies our previous analysis on how bandwidth (B) and frame length (S) impact positioning accuracy. For OFDM, the product of B and S is at the level of one hundred in highly dynamic environments. For example in LTE, one resource block has 12 sub-carriers and 7 OFDM symbols, which means the block size is only 84. For OTFS, the product of B and S can be enlarged by two orders of magnitude. The reason for such improvement is that the OTFS modulation is based on the DD domain channel model by partially considering the Doppler effect, while the OFDM is based on the LTI channel model which ignores the Doppler effect. Thus, the OTFS provides much better performance than OFDM.

From the simulation results, we can see that 3D positioning can be done with a ULA if the Doppler shift measurements are utilized. We also notice that the CRLB can serve as a metric for performance analysis for the chosen bandwidth, frame length, and antenna array size. Therefore, in the following sub-section, we will evaluate the impact of Doppler effect on positioning accuracy based on CRLB only.

B. Doppler Effect

As we already see, the Doppler shift provides extra position information concerning the vehicle. Besides, the Doppler effect helps to isolate different paths and thus suppress the interference. Then intuitively, when the vehicle moves faster we can get higher positioning accuracy, which is exactly what we can see in Fig. 6.

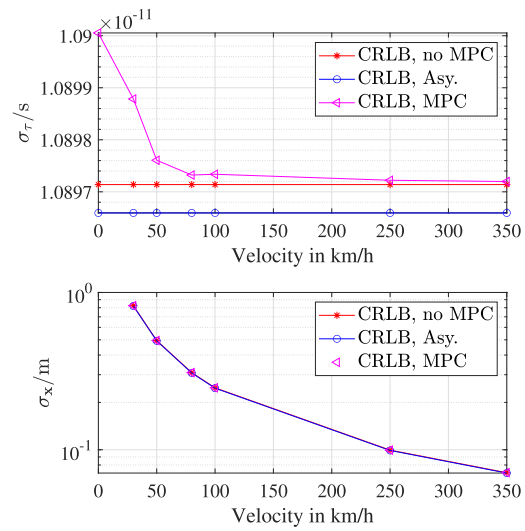


Fig. 6. Doppler effect on positioning accuracy.

In Fig. 6, we present the RMSE of positioning results for typical vehicular speeds, i.e., {0, 30, 50, 80, 100, 250, 350} km/h. Consider $K = 32$ antennas at the BS, with an SNR of 0 dB. "MPC" is short for Multi-Path Components. From the simulation results, we can also see that the previous analyses are very accurate, including the exact CRLB with MPC in (41), the CRLB without MPC in (42), and the asymptotic CRLB without MPC in (70). This means the multipath effect is not a big threat for SLAC, given that the bandwidth and antenna array are large enough. The measurement noise dominates the positioning accuracy. However, for smaller bandwidth or

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.

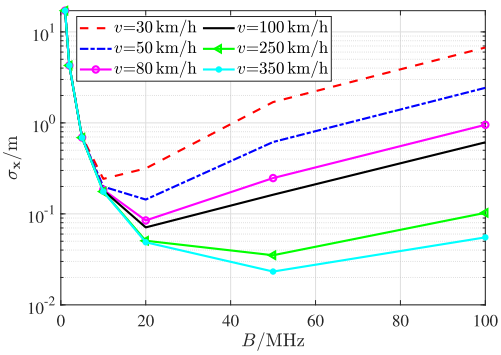


Fig. 7. Trade-off between B and S.

antenna array size, the MPC cannot be distinguished from the LoS signal, and the multipath interference will play a more important role in positioning error, as we will see in later discussions.

The results show that the positioning error will decrease when Doppler shift increases, as we expect from the theoretical analyses. For a vehicle speed of 30 km/h, the positioning error is at the level of one meter. However, if we increase the vehicle speed to 250 km/h, the positioning error will decrease to 0.1 m. There are two reasons for such improvement. First, when the vehicle moves, through spatial-temporal signal processing we can synthesize a virtual antenna array. When the vehicle moves faster, the size of the virtual array also increases proportionally, leading to higher spatial resolution and thus smaller positioning error. Second, the higher spatial resolution also helps to isolate different paths in space, leading to better estimate of delay. However, the first advantage dominates for the chosen parameters. Besides, we also notice that 3D positioning becomes impossible for a vehicular speed of zero.

From the estimation error of ToA, we can clearly see the impact of multi-path interference. For the two scenarios without multi-path interference, the estimation error of ToA is only dependent on the SNR. As a result, when the velocity changes, the estimation error is constant. However, with multi-path interference, both multi-path interference and noise are contributing to the estimation error. As the velocity increases, the multi-path interference can be better isolated and eliminated in the Doppler domain. As a result, the impact of multi-path interference will decrease, and eventually disappear. That is, the result will converge to the scenario without multi-path interference.

C. Trade-off Between S and B

As we have previously discussed, there is a trade-off between S and B. A larger S leads to higher Doppler shift estimation accuracy, while increased B leads to reduced ToA measurement error. In our case, both Doppler shift and ToA measurements contribute to the positioning result. Therefore, we need to figure out what are the optimal choices of B and S. The numerical results are presented in Fig. 7.

In the simulations, we fix the product of B and S, i.e., $BS = 10^5$. For a vehicle speed of 30 km/h, the positioning error decreases from around 20 meters to 0.25 meter when we increase bandwidth from 1 MHz to 10 MHz. However, if we further increase the bandwidth at this point, the positioning error will bounce back. This can be explained by the fact that positioning error is dependent on the estimation accuracy of both Doppler shift and ToA. When the product of B and S is fixed, increasing B will lead to increased accuracy in ToA estimate, but reduced accuracy in Doppler shift estimate. As a result, the positioning error will decrease first and rise at a certain point. For different vehicle speeds, we can see similar trends. Although the optimal B seems to vary over the vehicle speed, but 20 MHz can almost minimize the positioning error in all the situations.

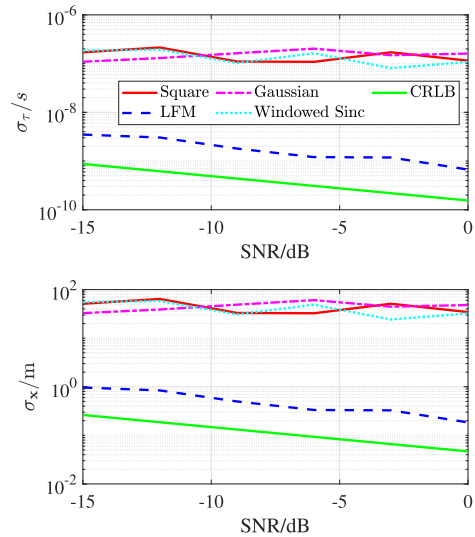


Fig. 8. Positioning errors of different modulation pulses.

D. Positioning Error of Different Pulse Shapes

One huge advantage of OTFS-based SLAC is that we can use a variety of pulse shapes for modulation. The potential impact of pulse shape on positioning accuracy can be clearly seen in (15). For different pulse shapes, the ambiguity function will be different. If the ambiguity function is very sensitive to delay and Doppler shift, then the corresponding complex gains of reflected paths will roll off faster and lead to reduced interference. In this sub-section, we will evaluate the performance of different pulse shapes. Specifically, we will consider four typical waveforms: the square waveform, the Gaussian waveform, the linear frequency modulated (LFM) waveform, and the windowed sinc waveform. Different pulse shapes provide different levels of multi-path resilience. For example, the square waveform is very sensitive to Doppler shift, while the LFM waveform has very high resolution in delay. From Fig. 6, we can see that multipath interference is not a problem for large bandwidth, antenna array and long frame length, and even square waveform works very well. However, small communications devices with limited bandwidth and antenna array size are not uncommon in the Internet of Things (IoT), and proper waveform can help us to further eliminate MPCs. Therefore, the bandwidth is chosen as 2 MHz for the the simulation results in Fig. 8 to represent strong multi-path

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.

environment, with a frame length of 1 ms and an array size of $K = 32$.

In Fig. 8, due to the much smaller bandwidth, frame length and array size, the MPC can no longer be well separated. From the figure, we can see that the positioning results are quite unstable due to very strong MPC, and the LFM waveform provides the best positioning accuracy in this case. As comparison, all the other waveforms have almost constant accuracy through the simulated SNR range, which means the multi-path interference dominates the estimation errors of channel parameters. This is evident from the results on propagation delay. The great performance of the LFM waveform comes from its ability to isolate different paths in the delay domain. For OFDM, the square waveform is employed for modulation, so that the sub-channels are orthogonal in frequency domain. As a result, this comparison shows the advantage of OTFS over OFDM, resulting from the flexible waveform design.

VI. Conclusion

In this paper, we conducted comprehensive signal modelling, performance analysis, algorithm design, and numerical evaluations for massive MIMO-OTFS based SLAC. Specifically, with a ULA at the BS, we showed that 3D positioning of a high-speed device is possible if we combine ToA, AoA, and Doppler shift measurements. The major discoveries lie in three aspects. First, we modelled the spatial wideband effect in massive MIMO-OTFS, and showed that this effect cannot be ignored in SLAC with ultra-wideband mmWave signals. Second, we showed that the product of bandwidth and frame length is confined by the reciprocal of Doppler scaling factor. Therefore, there is a trade-off between estimation accuracy of Doppler shift and propagation delay. The optimal parameters are derived. Third, the CRLB of positioning error is derived by considering a general multi-path model, and a low-complexity algorithm based on FFT is proposed to estimate the important parameters of the LoS path signal. Based on the analysis, we can see that the multi-path interference is not a problem for massive MIMO-OTFS based SLAC, given that the bandwidth, frame length and antenna array size are large enough. In strong multi-path environment with limited bandwidth, frame length and antenna array size, LFM signals provides better resolution in both delay and Doppler domains, and thus is a better choice than the conventional square waveform for SLAC.

Appendix A
Asymptotic Orthogonality of Different Rows in $\tilde{\mathbf{S}}$

To simplify the notations, suppose the three phase shifts of the l-th path are ω_a^l (in angle domain), ω_t^l (in Doppler domain) and ω_f^l (in delay domain), defined as

$$\omega_a^l = -\tilde{\omega}_l,\ \omega_t^l = 2\pi T\tilde{\nu}_l N/P,\ \omega_f = -2\pi F\tilde{\tau}_l M/Q. \quad (91)$$

Without loss of generality, we consider two paths, indexed by 0 and 1. The first column of $\tilde{\mathbf{S}}^T$ (or the first row of $\tilde{\mathbf{S}}$) is given by

$$\mathbf{s}_0 = \mathbf{a}_K(\omega_a^0) \otimes \mathbf{a}_P(\omega_t^0) \otimes \mathbf{a}_Q(\omega_f^0), \quad (92)$$

where $\otimes$ indicates the Kronecker product. $\mathbf{a}_N(\omega) = [1, e^{j\omega}, e^{j2\omega}, \cdots, e^{j(N-1)\omega}]^{\mathrm{T}}$ is a steering vector of dimension N. The second column of $\tilde{\mathbf{S}}^T$ is given by

$$\mathbf{s}_1 = \mathbf{a}_K(\omega_a^1) \otimes \mathbf{a}_P(\omega_t^1) \otimes \mathbf{a}_Q(\omega_f^1)). \quad (93)$$

When $\omega_a^0 \neq \omega_a^1$ or $\omega_f^0 \neq \omega_f^1$ or $\omega_t^0 \neq \omega_t^1$, the inner product of $\mathbf{s}_0$ and $\mathbf{s}_1$ is given by

$$\begin{aligned}\mathbf{s}_0^{\mathrm{H}}\mathbf{s}_1 &= \sum_{k,p,q} e^{jk(\omega_a^1-\omega_a^0)} e^{jp(\omega_t^1-\omega_t^0)} e^{jq(\omega_f^1-\omega_f^0)},\\ &= \left(\sum_k e^{jk(\omega_a^1-\omega_a^0)}\right)\cdot\left(\sum_p e^{jp(\omega_t^1-\omega_t^0)}\right)\\ &\quad\cdot\left(\sum_q e^{jq(\omega_f^1-\omega_f^0)}\right). \end{aligned} \quad (94)$$

Note that

$$\lim_{N\to\infty} \frac{\sum_n e^{jn\omega}}{N} = 0\ (\forall\omega \neq 0), \quad (95)$$

which holds for any integer N. Besides, we have $\left\|\mathbf{s}_0^{\mathrm{H}}\right\| = \left\|\mathbf{s}_1^{\mathrm{H}}\right\| = \sqrt{KPQ}$, leading to

$$\lim_{K,P,Q\to\infty} \frac{\mathbf{s}_0^{\mathrm{H}}\mathbf{s}_1}{\|\mathbf{s}_0^{\mathrm{H}}\|\|\mathbf{s}_1\|} = 0. \quad (96)$$

References

[1] L. Xiao, S. Li, Y. Qian, D. Chen, and T. Jiang, "An overview of OTFS for Internet of Things: Concepts, benefits, and challenges," *IEEE Internet Things J.*, vol. 9, no. 10, pp. 7596–7618, May 2022.

[2] R. Hadani and A. Monk, "OTFS: A new generation of modulation addressing the challenges of 5G," 2018, *arXiv:1802.02623*.

[3] W. Yuan et al., "New delay Doppler communication paradigm in 6G era: A survey of orthogonal time frequency space (OTFS)," 2022, *arXiv:2211.12955*.

[4] M. A. Uusitalo et al., "6G vision, value, use cases and technologies from European 6G flagship project Hexa-X," *IEEE Access*, vol. 9, pp. 160004–160020, 2021.

[5] J. Liu, Y. Shi, Z. M. Fadlullah, and N. Kato, "Space-air-ground integrated network: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2714–2741, 4th Quart., 2018.

[6] Z. Gong, C. Li, F. Jiang, and M. Z. Win, "Data-aided Doppler compensation for high-speed railway communications over mmWave bands," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 520–534, Jan. 2021.

[7] Y. Ma, G. Ma, N. Wang, Z. Zhong, and B. Ai, "OTFS-TSMA for massive Internet of Things in high-speed railway," *IEEE Trans. Wireless Commun.*, vol. 21, no. 1, pp. 519–531, Jan. 2022.

[8] Y. Su, Y. Liu, Y. Zhou, J. Yuan, H. Cao, and J. Shi, "Broadband LEO satellite communications: Architectures and key technologies," *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 55–61, Apr. 2019.

[9] H. Zhou, W. Xu, J. Chen, and W. Wang, "Evolutionary V2X technologies toward the Internet of Vehicles: Challenges and opportunities," *Proc. IEEE*, vol. 108, no. 2, pp. 308–323, Feb. 2020.

[10] M. Li, S. Zhang, F. Gao, P. Fan, and O. A. Dobre, "A new path division multiple access for the massive MIMO-OTFS networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 903–918, Apr. 2021.

[11] P. Raviteja, Y. Hong, E. Viterbo, and E. Biglieri, "Effective diversity of OTFS modulation," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 249–253, Feb. 2020.

[12] A. Monk, R. Hadani, M. Tsatsanis, and S. Rakib, "OTFS–orthogonal time frequency space," 2016, *arXiv:1608.02993*.

[13] R. Hadani et al., "Orthogonal time frequency space modulation," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2017, pp. 1–6.

[14] K. R. Murali and A. Chockalingam, "On OTFS modulation for high-Doppler fading channels," in *Proc. Inf. Theory Appl. Workshop (ITA)*, Feb. 2018, pp. 1–10.

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.

[15] S. K. Mohammed, "Derivation of OTFS modulation from first principles," *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 7619–7636, Aug. 2021.

[16] G. D. Surabhi, R. M. Augustine, and A. Chockalingam, "On the diversity of uncoded OTFS modulation in doubly-dispersive channels," *IEEE Trans. Wireless Commun.*, vol. 18, no. 6, pp. 3049–3063, Jun. 2019.

[17] L. Gaudio, G. Colavolpe, and G. Caire, "OTFS vs. OFDM in the presence of sparsity: A fair comparison," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4410–4423, Jun. 2022.

[18] P. Raviteja, Y. Hong, E. Viterbo, and E. Biglieri, "Practical pulse-shaping waveforms for reduced-cyclic-prefix OTFS," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 957–961, Jan. 2019.

[19] P. Raviteja, K. T. Phan, and Y. Hong, "Embedded pilot-aided channel estimation for OTFS in delay–Doppler channels," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4906–4917, May 2019.

[20] S. Wang, J. Guo, X. Wang, W. Yuan, and Z. Fei, "Pilot design and optimization for OTFS modulation," *IEEE Wireless Commun. Lett.*, vol. 10, no. 8, pp. 1742–1746, Aug. 2021.

[21] H. Qu, G. Liu, M. A. Imran, S. Wen, and L. Zhang, "Efficient channel equalization and symbol detection for MIMO OTFS systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 8, pp. 6672–6686, Aug. 2022.

[22] G. D. Surabhi, R. M. Augustine, and A. Chockalingam, "Peak-to-average power ratio of OTFS modulation," *IEEE Commun. Lett.*, vol. 23, no. 6, pp. 999–1002, Jun. 2019.

[23] Md. N. Hossain, Y. Sugiura, T. Shimamura, and H.-G. Ryu, "Waveform design of low complexity WR-OTFS system for the OOB power reduction," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, Apr. 2020, pp. 1–5.

[24] P. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun.*, vol. COM-11, no. 4, pp. 360–393, Dec. 1963.

[25] R. Hadani et al., "Orthogonal time frequency space (OTFS) modulation for millimeter-wave communications systems," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2017, pp. 681–683.

[26] R. Hadani et al., "Orthogonal time frequency space modulation," 2018, *arXiv:1808.00519*.

[27] H. Groll et al., "Sparsity in the delay-Doppler domain for measured 60 GHz vehicle-to-infrastructure communication channels," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2019, pp. 1–6.

[28] A. M. Sayeed and B. Aazhang, "Joint multipath-Doppler diversity in mobile wireless communications," *IEEE Trans. Commun.*, vol. 47, no. 1, pp. 123–132, Jan. 1999.

[29] W. Kozek and A. F. Molisch, "Nonorthogonal pulseshapes for multicarrier communications in doubly dispersive channels," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1579–1589, Oct. 1998.

[30] K. Liu, T. Kadous, and A. M. Sayeed, "Orthogonal time–frequency signaling over doubly dispersive channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2583–2603, Nov. 2004.

[31] T. Strohmer and S. Beaver, "Optimal OFDM design for time-frequency dispersive channels," *IEEE Trans. Commun.*, vol. 51, no. 7, pp. 1111–1122, Jul. 2003.

[32] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, "Spatial- and frequency-wideband effects in millimeter-wave massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 66, no. 13, pp. 3393–3406, Jul. 2018.

[33] Y. Song, Z. Gong, Y. Chen, and C. Li, "Efficient channel estimation for wideband millimeter wave massive MIMO systems with beam squint," *IEEE Trans. Commun.*, vol. 70, no. 5, pp. 3421–3435, May 2022.

[34] P. Raviteja, E. Viterbo, and Y. Hong, "OTFS performance on static multipath channels," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 745–748, Jun. 2019.

[35] Z. Wang, X. Chen, and X. Ning, "BER analysis of integrated WFRFT-OTFS waveform framework over static multipath channels," *IEEE Commun. Lett.*, vol. 25, no. 3, pp. 754–758, Mar. 2021.

[36] M. F. Keskin, H. Wymeersch, and A. Alvarado, "Radar sensing with OTFS: Embracing ISI and ICI to surpass the ambiguity barrier," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2021, pp. 1–6.

[37] J. Zak, "Finite translations in solid-state physics," *Phys. Rev. Lett.*, vol. 19, no. 24, pp. 1385–1387, Dec. 1967, doi: 10.1103/PhysRevLett.19.1385.

[38] F. Lampel, A. Alvarado, and F. M. J. Willems, "Orthogonal time frequency space modulation: A discrete Zak transform approach," 2021, *arXiv:2106.12828*.

[39] S. K. Mohammed, "Time-domain to delay-Doppler domain conversion of OTFS signals in very high mobility scenarios," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 6178–6183, Jun. 2021.

[40] P. Wei, Y. Xiao, W. Feng, N. Ge, and M. Xiao, "Charactering the peak-to-average power ratio of OTFS signals: A large system analysis," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 3705–3720, Jun. 2022.

[41] J. Pan, "Cramer–Rao low bound of channel estimation for orthogonal time frequency space modulation system," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 9646–9658, Oct. 2021.

[42] A. Pfadler, T. Szollmann, P. Jung, and S. Stanczak, "Leakage suppression in pulse-shaped OTFS delay-Doppler-pilot channel estimation," *IEEE Wireless Commun. Lett.*, vol. 11, no. 6, pp. 1181–1185, Jun. 2022.

[43] M. Mohammadi, H. Q. Ngo, and M. Matthaiou, "Cell-free massive MIMO meets OTFS modulation," *IEEE Trans. Commun.*, vol. 70, no. 11, pp. 7728–7747, Nov. 2022.

[44] W. Shen, L. Dai, J. An, P. Fan, and R. W. Heath, "Channel estimation for orthogonal time frequency space (OTFS) massive MIMO," *IEEE Trans. Signal Process.*, vol. 67, no. 16, pp. 4204–4217, Aug. 2019.

[45] Y. Liu, S. Zhang, F. Gao, J. Ma, and X. Wang, "Uplink-aided high mobility downlink channel estimation over massive MIMO-OTFS system," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 9, pp. 1994–2009, Sep. 2020.

[46] F. Liu, Z. Yuan, Q. Guo, Z. Wang, and P. Sun, "Message passing-based structured sparse signal recovery for estimation of OTFS channels with fractional Doppler shifts," *IEEE Trans. Wireless Commun.*, vol. 20, no. 12, pp. 7773–7785, Dec. 2021.

[47] Z. Wei, W. Yuan, S. Li, J. Yuan, and D. W. K. Ng, "Off-grid channel estimation with sparse Bayesian learning for OTFS systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7407–7426, Sep. 2022.

[48] Y. Song, Z. Gong, Y. Chen, and C. Li, "Tensor-based sparse Bayesian learning with intra-dimension correlation," *IEEE Trans. Signal Process.*, vol. 71, pp. 31–46, 2023.

[49] W. Yuan, S. Li, Z. Wei, J. Yuan, and D. W. Kwan Ng, "Bypassing channel estimation for OTFS transmission: An integrated sensing and communication solution," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, Mar. 2021, pp. 1–5.

[50] P. Singh, K. Yadav, H. B. Mishra, and R. Budhiraja, "BER analysis for OTFS zero forcing receiver," *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2281–2297, Apr. 2022.

[51] S. Tiwari, S. S. Das, and V. Rangamgari, "Low complexity LMMSE receiver for OTFS," *IEEE Commun. Lett.*, vol. 23, no. 12, pp. 2205–2209, Dec. 2019.

[52] G. D. Surabhi and A. Chockalingam, "Low-complexity linear equalization for OTFS modulation," *IEEE Commun. Lett.*, vol. 24, no. 2, pp. 330–334, Feb. 2020.

[53] P. Singh, S. Tiwari, and R. Budhiraja, "Low-complexity LMMSE receiver design for practical-pulse-shaped MIMO-OTFS systems," *IEEE Trans. Commun.*, vol. 70, no. 12, pp. 8383–8399, Dec. 2022.

[54] L. Xiang, Y. Liu, L.-L. Yang, and L. Hanzo, "Gaussian approximate message passing detection of orthogonal time frequency space modulation," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 10999–11004, Oct. 2021.

[55] Z. Yuan, F. Liu, W. Yuan, Q. Guo, Z. Wang, and J. Yuan, "Iterative detection for orthogonal time frequency space modulation with unitary approximate message passing," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 714–725, Feb. 2022.

[56] Z. Zhou, L. Liu, J. Xu, and R. Calderbank, "Learning to equalize OTFS," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7723–7736, Sep. 2022.

[57] Y. K. Enku et al., "Two-dimensional convolutional neural network-based signal detection for OTFS systems," *IEEE Wireless Commun. Lett.*, vol. 10, no. 11, pp. 2514–2518, Nov. 2021.

[58] W. Yuan, S. Li, Z. Wei, J. Yuan, and D. W. K. Ng, "Data-aided channel estimation for OTFS systems with a superimposed pilot and data transmission scheme," *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 1954–1958, Sep. 2021.

[59] M. Li, S. Zhang, Y. Ge, F. Gao, and P. Fan, "Joint channel estimation and data detection for hybrid RIS aided millimeter wave OTFS systems," *IEEE Trans. Commun.*, vol. 70, no. 10, pp. 6832–6848, Oct. 2022.

[60] Y. Liu, Y. L. Guan, and G. D. González, "Near-optimal BEM OTFS receiver with low pilot overhead for high-mobility communications," *IEEE Trans. Commun.*, vol. 70, no. 5, pp. 3392–3406, May 2022.

[61] L. Gaudio, M. Kobayashi, G. Caire, and G. Colavolpe, "On the effectiveness of OTFS for joint radar parameter estimation and communication," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 5951–5965, Sep. 2020.

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.

[62] S. K. Dehkordi, L. Gaudio, M. Kobayashi, G. Colavolpe, and G. Caire, "Beam-space MIMO radar with OTFS modulation for integrated sensing and communications," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2022, pp. 509–514.

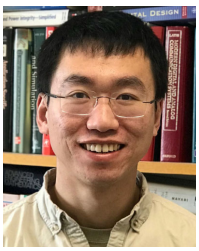
[63] Q. Wang, A. Kakkavas, X. Gong, and R. A. Stirling-Gallacher, "Towards integrated sensing and communications for 6G," in *Proc. 2nd IEEE Int. Symp. Joint Commun. Sens. (JC&S)*, Mar. 2022, pp. 1–6.

[64] K. Wu, J. A. Zhang, X. Huang, and Y. J. Guo, "OTFS-based joint communication and sensing for future industrial IoT," *IEEE Internet Things J.*, vol. 10, no. 3, pp. 1973–1989, Feb. 2023.

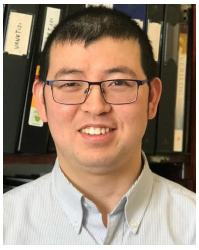
[65] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[66] B. G. Quinn, "Estimating frequency by interpolation using Fourier coefficients," *IEEE Trans. Signal Process.*, vol. 42, no. 5, pp. 1264–1268, May 1994.

[67] Z. Gong, C. Li, F. Jiang, and J. Zheng, "AUV-aided localization of underwater acoustic devices based on Doppler shift measurements," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2226–2239, Apr. 2020.



Zijun Gong (Member, IEEE) received the B.Eng. and M.Eng. degrees from the Harbin Institute of Technology (HIT), Harbin, China, in 2013 and 2015, respectively, and the Ph.D. degree from the Memorial University of Newfoundland, St. John's, NL, Canada, in 2021. From May 2021 to December 2021, he was with the University of Waterloo as a Post-Doctoral Researcher. He is currently an Assistant Professor with the IoT Thrust in Information Hub, The Hong Kong University of Science and Technology (GZ). He is also an affiliate Assistant Professor with the ECE Department, HKUST, Hong Kong. His research interests lie in statistical signal processing and optimization, including channel estimation in massive MIMO, OTFS modulation, millimeter wave communications, mmWave radar, radio propagation modeling, localization of WSN, and localization of underwater targets and devices. He received the Best Paper Award at the IEEE GLOBECOM'17, Singapore, in December 2017.



Fan Jiang (Member, IEEE) received the B.E. degree in communication engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2010, the M.E. degree in communication and information systems from Southeast University (SEU), Nanjing, China, in 2013, and the Ph.D. degree from the Memorial University of Newfoundland (MUN), St. John's, Canada, in 2018. He held post-doctoral research positions with the Department of Electrical Engineering, Chalmers University of Technology, the Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology (MIT), and the Faculty of Engineering and Applied Science, MUN, after graduation. He has been an Associate Researcher with the Peng Cheng Laboratory (PCL), Shenzhen, China, since July 2022, and was an Assistant Professor with Halmstad University, Sweden, before the current appointment. His research interests include signal processing for wireless communications, reconfigurable intelligent surface technology, and integrated sensing and communications framework. He has received the Post-Doctoral Fellowship from Natural Science and Engineering Research Council (NSERC) of Canada in 2019, The Governor General's Gold Medal at MUN in 2018, and the Best Paper Award at the IEEE GLOBECOM in 2017.

Cheng Li (Senior Member, IEEE) received the B.Eng. and M.Eng. degrees from the Harbin Institute of Technology, Harbin, China, in 1992 and 1995, respectively, and the Ph.D. degree in electrical and computer engineering from Memorial University, St. John's, NL, Canada, in 2004. He is currently a Full Professor with the School of Engineering Science, Simon Fraser University. His research interests include wireless communications and networking, mobile computing, underwater communications and networks, the Internet of Things (IoT), and broadband communication networks. He is also a registered Professional Engineer in Canada and a Senior Member of the IEEE Communications Society, IEEE Computer Society, IEEE Ocean Engineering Society, and IEEE Vehicular Technology Society. He was a recipient of the Best Paper Award at the International Conference on Ad Hoc Networks (ADHOCNETS) 2018, the IEEE GLOBECOM 2017, and the IEEE International Conference on Communications (ICC) 2010. He is an Associate Editor of IEEE INTERNET OF THINGS JOURNAL, *IEEE Network Magazine*, IEEE SYSTEM JOURNAL, and *China Communications*. He was the General Co-Chair of the ICNC'22, AICON'19, WINCOM'19, and WINCOM'17. He was the TPC Co-Chair of the ICNC'20, ADHOCNETS'19, ADHOCNETS'17, MSWiM'14, MSWiM'13, WINCOM'11, and QBSC'10. He has served as the Co-Chair for various technical symposia or tracks of many international conferences, including IEEE ICC, GLOBECOM, ICNC, VTC, and IWCMC.

Xuemin (Sherman) Shen (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Rutgers University, New Brunswick, NJ, USA, in 1990. He is currently an University Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research focuses on network resource management, wireless network security, the Internet of Things, 5G and beyond, and vehicular networks.

He is also a registered Professional Engineer of Ontario, Canada, an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, a Royal Society of Canada Fellow, a Chinese Academy of Engineering Foreign Member, and a Distinguished Lecturer of the IEEE Vehicular Technology Society and Communications Society. He received the "West Lake Friendship Award" from Zhejiang Province in 2023, President's Excellence in Research from the University of Waterloo in 2022, the Canadian Award for Telecommunications Research from the Canadian Society of Information Theory (CSIT) in 2021, the R. A. Fessenden Award in 2019 from IEEE, Canada, Award of Merit from the Federation of Chinese Canadian Professionals (Ontario) in 2019, James Evans Avant Garde Award in 2018 from the IEEE Vehicular Technology Society, Joseph LoCicero Award in 2015 and Education Award in 2017 from the IEEE Communications Society (ComSoc), and Technical Recognition Award from Wireless Communications Technical Committee (2019) and AHSN Technical Committee (2013). He has also received the Excellent Graduate Supervision Award in 2006 from the University of Waterloo and the Premier's Research Excellence Award (PREA) in 2003 from the Province of Ontario, Canada. He serves/served as the General Chair for the 6G Global Conference 23 and ACM Mobihoc'15, the Technical Program Committee Chair/Co-Chair for IEEE Globecom 24, IEEE Globecom 16, IEEE Globecom 07, IEEE Infocom 14, and IEEE VTC 10 Fall, and the Chair for the IEEE ComSoc Technical Committee on Wireless Communications. He is also the President of the IEEE ComSoc. He was the Vice President of Technical & Educational Activities, the Vice President of Publications, a Member-at-Large on the Board of Governors, the Chair of the Distinguished Lecturer Selection Committee, and a member of IEEE Fellow Selection Committee of the ComSoc. He served as the Editor-in-Chief for the IEEE IoT JOURNAL, IEEE NETWORK, and *IET Communications*.

Authorized licensed use limited to: University of Waterloo. Downloaded on January 09,2024 at 18:35:26 UTC from IEEE Xplore. Restrictions apply.