Massive MIMO for High-Accuracy Target Localization and Tracking

Xiaolu Zeng, Feng Zhang, *Member, IEEE*, Beibei Wang, *Senior Member, IEEE*, and K. J. Ray Liu, *Fellow, IEEE*

***Abstract*—High-accuracy target localization and tracking have been widely used in the modern navigation system. However, most of the methods such as global positioning system (GPS) are highly dependent on time measurement accuracy, which prevents them from achieving high accuracy in practice. Time reversal (TR)-based technique has been shown to be able to achieve centimeter accuracy localization by fully utilizing the focusing effect brought by the massive multipaths naturally existing in a rich scattering environment such as indoor scenarios. By investigating a similar statistical property, this article develops a novel high-accuracy target localization method by using massive MIMO to provide massive signal components. We first observe that the statistical autocorrelation of the received energy physically focuses into a beam around the receiver exhibiting a sinc-like distribution in the far-field scenario. By leveraging such a distribution of the focusing beam, an effective way to estimate the relative moving speed of the target with respect to a single base station is proposed. We also obtain the absolute moving speed and subsequently track the target accurately by associating the speed estimation results and geometrical relationship of multiple stations. The theoretical analysis on the error in the speed and localization estimation validated by numerical simulation results show that the proposed system can achieve decimeter accuracy for target localization and tracking.**

***Index Terms*—Centimeter accuracy, MIMO, statistical electromagnetic, target localization and tracking.**

I. Introduction

TARGET localization and tracking have been of great interest over several decades because of their wide applications in navigation and many location-based services such as autonomous drivings [1], [2]. Furthermore, most of these localization requests emerge in urban areas where the global positioning system (GPS) [3] cannot offer good performance because the line-of-sight (LOS) signal between the GPS satellite and the terminal is easily to be blocked by obstacles such as tall buildings. As a result, it is imperative to seek for technologies that can provide high-accuracy localization in complex environments, such as dense urban areas under nonline-of-sight (NLOS) and multipath conditions [4].

Based on their principle, localization techniques can be classified into two categories, i.e., triangulation-based methods and fingerprinting-based methods. *Triangulation-based methods* consist of two steps. First, model-based parameters, such as the angle of arrivals (AOAs) [5], [6], time of arrivals (TOAs) [7], [8], or time difference of arrivals (TDOAs) of LOS signals [9], are measured at all access points (APs) or base stations (BSs). Then, the target location can be estimated by using triangulation/trilateration among all APs/BSs [10]. However, these methods cannot work well under the multipath effect and NLOS because of the unreliable parameter estimation. *Fingerprinting-based methods* first construct an offline database by collecting location related features, such as received signal strength (RSS) [11]–[13] and channel state information (CSI) [14]–[16], in the area of interest. Then, the same features are extracted from the online signals and compared with the offline database to obtain the location estimations. However, the overhead of establishing and updating the offline database also prevents these methods from being widely adopted [2].

More recently, massive MIMO has been gaining popularity in target localization because of its high angular resolution and degree of freedom [17]. This mainly benefits from the hundreds of antennas on the BS, which can enable narrow, highly directional, and high-gain beams by beamforming [18]. Similar to the localization methods without using massive MIMO, the existing massive MIMO-based localization methods can also be classified into the same two categories. The first is the triangulation-based methods in which many techniques, such as beamforming [19], multiple signal classification (MUSIC) [20], 2-D rotational invariance technique [21], and compressive sensing [22], [23], are explored on the base of massive MIMO systems. To reduce the prohibitive energy consumption and complexity increment caused by the massive antennas, high-efficient beam allocation/switching schemes [24], [25], AOAs estimations in beamspace [26], preenergy detections [27], as well as the combination of digital beamforming and analog techniques [28]–[30] have been considered. In the fingerprinting-based methods with massive MIMO [31]–[36], different matching techniques have been studied in comparing the online phase with the offline phase to estimate the target location, such as model-based similarity comparison [31], similarity learning by neural networks (NNs) [32], [33], support vector machines (SVMs) [35],

Manuscript received September 10, 2020; revised December 21, 2020; accepted January 6, 2021. Date of publication January 11, 2021; date of current version June 7, 2021. *(Corresponding author: Xiaolu Zeng.)*

Xiaolu Zeng was with the School of Electronic Engineering, Xidian University, Xi'an 71007, China. He is now with the Department of Electrical and Computer Engineering, University of Maryland at College Park, College Park, MD 20742 USA (e-mail: feiying3709@sina.com).

Feng Zhang was with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA, and also with Origin Wireless Inc., Greenbelt, MD 20770 USA. He is now with Amazon.com Inc., Seattle, WA, USA (e-mail: fzhang15@umd.edu).

Beibei Wang and K. J. Ray Liu are with Origin Wireless Inc., Greenbelt, MD 20770 USA, and also with the Department of Electrical and Computer Engineering, University of Maryland at College Park, College Park, MD 20742 USA (e-mail: bebewang@umd.edu; kjrliu@umd.edu).

Digital Object Identifier 10.1109/JIOT.2021.3050720

and Kernel-based methods [36]. Even though the localization accuracy is improved by leveraging the high range/angular resolution provided by massive antennas, most of existing massive MIMO-based localization methods still entail the same challenges as the traditional methods which do not use massive MIMO antennas, that is, the NLOS distortions and performance degradation in rich-scattering environment. This motivates us to design a high-accuracy localization system that is robust to environment dynamics while with good performance under multipath and NLOS conditions.

Inspired by the recent research on decimeter-accuracy indoor tracking [37]–[40] using time reversal focusing effect [41], [42], in this article, we propose a massive MIMO-based high-accuracy localization and tracking system by utilizing the focusing effect brought by the massive number of antennas. We first propose the definition of an important statistical variable, the strength of the autocorrelation function (ACFS) of the received signal in a massive MIMO system, to characterize the energy distribution of the focusing effect around a location of interest. Because the received signal in a massive MIMO system contains a large number of components due to the massive number of antenna elements and further reflections/scattering, it can be shown that the distribution of the ACFS exhibits a stationary sinc-like focusing beam[1] around the receiver in spatial domain regardless of the environment.

By leveraging the ACFS, we then develop an approach that can estimate the relative speed of the target with respect to a single BS. The absolute moving speed, moving direction/orientation, and location of the target can be further derived by jointly considering the relative speed estimation (ES) and geometrical relationship among multiple BSs. Different from [42], which needs an extra inertial sensor to estimate the moving direction because the energy distribution of the time reversal focusing effect shows the same trend along all the directions, the proposed system can estimate the moving speed/distance and direction simultaneously only based on the ACFS focusing beam that exhibits different distributions along different directions. This is because that in the proposed system, the massive number of the incident signal components reaches the receiver from the antennas/BS side, resulting in a directional focusing beam rather than a symmetrical focusing ball as shown in [42].

Based on the derivation of ACFS and how it can be used for speed and location estimation, we derive the theoretical expectation of the speed and location estimation errors, which are further verified by extensive simulations. It is shown that the proposed system can achieve decimeter-level accuracy for target localization and tracking in various scenarios, which outperforms three latest massive MIMO-based localization techniques [23], [31], [32].

In summary, the main contributions of this work are as follows.

1) We observed and proved that the statistical distribution of the ACFS of the received signal in a massive MIMO system exhibits a sinc-like beam pattern, because the received signal usually contains a large number of LOS and NLOS signal components.
2) Based on the distribution of the ACFS, we developed a target localization and tracking system that has robust performance in rich-scattering urban areas with NLOS. Because the proposed system only needs to calculate the ACFS of the received signal on the user side while the speed and location estimations are very straightforward according to the derived close-form expressions, the system enjoys a very low computation complexity and thus can be widely applied in real-time tracking and navigation applications with a stringent requirement on the latency.
3) We further derived the theoretical speed and localization error expectations of the proposed system and validated the theoretical performance analysis using extensive simulations.

The remainder of this article is organized as follows. In Section II, we elaborate on the signal model for the massive MIMO system followed by the derivation of the focusing beam. Then, Section III proposes an ES method by using the focusing beam of multiple distributed BSs. Section IV introduces the target localization system while Section V derives the theoretical speed and location estimation error expectations. Extensive numerical simulations are conducted to validate the performance of the proposed approach in Section VI. Finally, Section VII concludes this article.

II. Focusing Beam in Massive MIMO

In this section, we first introduce the background knowledge about the system model. Then, we elaborate on the signal model and derive the analytical distribution of the ACFS focusing beam in 5G massive MIMO communication systems.

A. Background Knowledge

Ultradense 5G BS Deployment: The 5G cellular network will be an ultradense cellular network, e.g., with a density of 40–50 BS/km^2, because that a massive number of antennas will be deployed on the BS [43], which means that every antenna's transmission power has to be greatly decreased compared to that of a 4G BS, leading to a smaller coverage area. Second, mmWave transmission is very likely to be adopted in 5G cellular networks and the signal decays much faster at mmWave frequency, which again will reduce the cell coverage and thus, denser BS deployment is needed. For example, the federal communications commission (FCC) in the USA issued a declaratory ruling that indicates that most of the 5G BSs are about 30 feet tall while the service range of each BS is about 400–500 feet or less in large crowded areas [44].

Far-Field Condition: As shown in Fig. 1, let H_B and L_{BR} denote the altitude of the BS and the horizontal distance between the BS and the receiver. A_e is the aperture of the antenna array A. Due to the ultradense 5G BS deployment, L_{BR} is about 80–200 m in practice [43]. In addition, A_e is less

[1]We use the term *focusing beam* rather than *beamforming* because we utilize the ACFS, a specific function of the received signal for positioning a target, and the distribution of the ACFS happens to exhibit a beam-shaped pattern. There is no "physical beamforming" that explicitly focuses a signal toward a receiver.

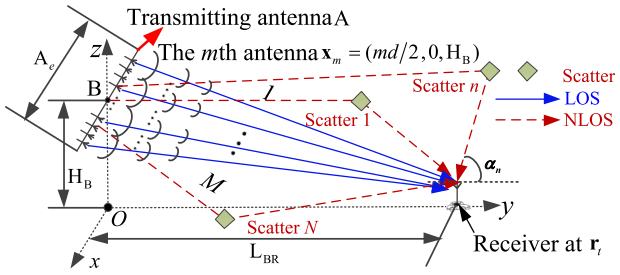


Fig. 1. Setup for a BS with massive MIMO antennas.

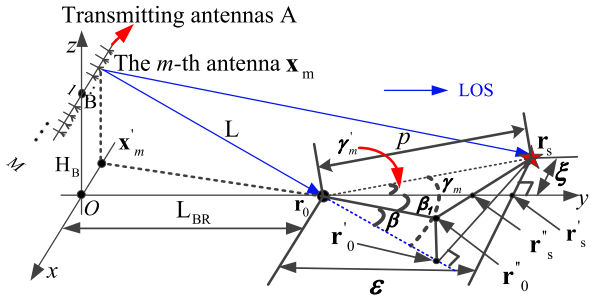


Fig. 2. Signal propagation geometry between $\mathbf{r}_0$ and $\mathbf{r}_s$.

than 2 m and $\mathrm{H_B}$ is about 10 m because of the antenna fabrication and installation requirements [44]. As a consequence, $\mathrm{L_{BR}} \geq 10\mathrm{H_B} \gg \mathrm{A}_e$ is the condition of the *far-field scenario* in this article and usually holds in the 5G networks [45], [46]. This is different from the *conventional far-field condition* in which $L_\mathrm{f} = 2\mathrm{A}_e^2/\lambda$ is the boundary between the Fresnel region and the Fraunhofer region [47] with λ as the wavelength of the signal.

B. Signal Model

As shown in Fig. 1, "B" denotes a BS equipped with M antennas, which communicate with the receiver "r" simultaneously. Note that practical measurements in [48]–[50] have validated that the LOS signal matches with the free space propagation model while the NLOS signal follows the Raleigh fading [51] in 5G massive MIMO system. As a result, in urban areas, the received signal consisting of both LOS and NLOS parts at baseband can be expressed as

$$\begin{aligned} y(t) &= y_L(t) + y_N(t) + n(t) \\ y_L(t) &= \sqrt{K_L} \sum_{m=1}^{M} \frac{\exp(j(k|\mathbf{x}_m\mathbf{r}_t| + \phi_m))}{4\pi|\mathbf{x}_m\mathbf{r}_t|} \\ y_N(t) &= \sqrt{K_N} \sum_{n=1}^{N} \exp[j(\omega_d t \cos\alpha_n + \phi_n)] \end{aligned} \tag{1}$$

where $y_L(t)$ and $y_N(t)$ denote the LOS and NLOS components and K_L and K_N are their corresponding power, $k = 2\pi/\lambda$ is the wave number, λ is the wave length, ω_d is the maximum radian Doppler frequency, $\mathbf{x}_m$ and $\mathbf{r}_t$ are the coordinates of the mth antenna and the receiver at time t, respectively, $|\mathbf{x}_m\mathbf{r}_t|$ denotes the Euclidean spatial distance between the mth antenna and the receiver, $n(t)$ represents the additive Gaussian noise, ϕ_m is the phase distortion of the mth LOS path signal, and α_n and ϕ_n are the AOA and phase distortion of the nth NLOS signal component.

In general, ϕ_m is caused by hardware imperfections, heterogeneity of the propagation medium and channel attenuations, etc. α_n and ϕ_n are mainly introduced by the reflection/absorption of the randomly distributed scatterers in a rich-scattering urban area. As a result, ϕ_m, α_n, and ϕ_n are not deterministic and can be assumed as i.i.d uniform distributions over $[-\pi, \pi)$ for $m = 1, 2, \ldots, M$ and $n = 1, 2, \ldots, N$ [51], [52], where N is the total number of NLOS signal components. In practice, the number of multipath N can vary from 10 to 100 in urban areas according to the practical measurements in New York City [45], [46].

C. Massive MIMO Focusing Beam

In the following, we explore the distribution of the focusing effect of massive MIMO in the far-field scenario by first deriving the ACF of the received signal and then the ACFS, which is inspired by TRRS in [42] and [53] but more robust to the randomness of signal distortions.

As shown in Fig. 2, a target moves from $\mathbf{r}_0$ at time t_0 to $\mathbf{r}_s$ at time t_s on the ground (xOy plane). Then, the ACF of the received signal between $\mathbf{r}_0$ and $\mathbf{r}_s$ is defined as

$$\eta_y(\mathbf{r}_0, \mathbf{r}_s) = \mathbb{E}\big[y(t_0)y^*(t_s)\big] \tag{2}$$

$$\approx \eta_{y_L} + \eta_{y_N} + \eta_n \tag{3}$$

where $\eta_n = \mathbb{E}[n(t_0)n^*(t_s)] = \sigma^2\mathbf{I}$, σ^2 is the power spectral density of the Gaussian noise $n(t)$. η_{y_L}, η_{y_N}, and η_n denotes the ACF of the LOS signal component $y_L(t)$, NLOS signal component $y_N(t)$, and the noise term $n(t)$, respectively.

Note that the independence among $y_L(t)$, $y_N(t)$ and $n(t)$ is assumed to obtain (3). Detailed derivations can be found in the Appendix. Next, we will derive η_{y_N} and then η_{y_L}, respectively.

1) ACF of NLOS Signal: According to [42] and [51], η_{y_N} can be written as

$$\begin{aligned} \eta_{y_N} &= \mathbb{E}\big[y_N(t_0)y_N^*(t_s)\big] \\ &= K_N \sum_{n=1}^{N}\sum_{i=1}^{N} \mathbb{E}_{\phi,\alpha}\{\exp[j(\omega_d t\cos\alpha_n + \phi_n) \\ &\quad \times \exp[j(\omega_d(t+\tau)\cos\alpha_i + \phi_i)\} \\ &= K_N N J_0(\omega_d\tau) = K_N N J_0(kp) \end{aligned} \tag{4}$$

where $\tau = t_s - t_0$, $\mathbb{E}_{\phi,\alpha}$ means taking expectation over ϕ and α, $J_0(\cdot)$ is the 0-order Bessel function, and p is the Euclidean distance between $\mathbf{r}_0$ to $\mathbf{r}_s$ and as shown in Fig. 2. We omit the details about the derivation of (4) because they are similar to that in [42] and [51].

2) ACF of LOS Signal: Similar to (4), the ACF of the LOS signal $y_L(t)$ between $\mathbf{r}_0$ and $\mathbf{r}_s$ is given by

$$\begin{aligned} \eta_{y_L} = \eta_{y_L}(\mathbf{r}_0, \mathbf{r}_s) = \mathbb{E}\big[y_L(t_0)y_L^*(t_s)\big] = K_L \\ \times \sum_{i=1}^{M}\sum_{m=1}^{M} \mathbb{E}_\phi\left\{\frac{\exp\big[j(k(|\mathbf{x}_i\mathbf{r}_0| - |\mathbf{x}_m\mathbf{r}_s|) + \phi_i - \phi_m)\big]}{(4\pi)^2|\mathbf{x}_i\mathbf{r}_0||\mathbf{x}_m\mathbf{r}_s|}\right\}. \end{aligned} \tag{5}$$

In the far-field scenario where $\{|\mathbf{x}_i\mathbf{r}_0|, |\mathbf{x}_m\mathbf{r}_s|\} > \mathrm{L_{BR}} \gg \mathrm{A}_e$ holds, $|\mathbf{x}_i\mathbf{r}_0|$ and $|\mathbf{x}_m\mathbf{r}_s|$ in the denominator of (5) can be approximated as the same for all elements, i.e., $|\mathbf{x}_i\mathbf{r}_0| \approx |\mathbf{x}_0\mathbf{r}_0|$ and $|\mathbf{x}_m\mathbf{r}_s| \approx |\mathbf{x}_0\mathbf{r}_s|$ because $(|\mathbf{x}_i\mathbf{r}_0| - |\mathbf{x}_m\mathbf{r}_s|)$ is usually magnitudes smaller than $|\mathbf{x}_i\mathbf{r}_0|$ and $|\mathbf{x}_m\mathbf{r}_s|$. We omit the denominator of (5) in the derivation for simplicity.

Next, we decompose (5) into two different cases, i.e., 1) $i = m$ and 2) $i \neq m$. Considering $i = m$, we have

$$\eta_{y_L}^{\mathrm{1st}} = K_L \sum_{m=1}^{M} \exp(jk(|\mathbf{x}_m\mathbf{r}_0| - |\mathbf{x}_m\mathbf{r}_s|)). \tag{6}$$

To compute $|\mathbf{x}_m\mathbf{r}_0| - |\mathbf{x}_m\mathbf{r}_s|$ in Fig. 2, the angle symbols are defined as $\angle\mathbf{r}_s\mathbf{r}_0\mathbf{r}_s' = \boldsymbol{\gamma}_m', \angle\mathbf{r}_s\mathbf{r}_0\mathbf{r}_0' = \boldsymbol{\gamma}_m, \angle\mathbf{r}_s'\mathbf{r}_0\mathbf{r}_0'' = \boldsymbol{\beta}_1, \angle\mathbf{r}_0''\mathbf{r}_0\mathbf{r}_0' = \boldsymbol{\beta}$, where $\mathbf{r}_0'$ lies in the extension line of $l_{\mathbf{x}_m\mathbf{r}_0}$ satisfying that $l_{\mathbf{x}_m\mathbf{r}_0'} \perp l_{\mathbf{r}_s\mathbf{r}_0'}$. $\mathbf{r}_0''$ is the projection of $\mathbf{r}_0'$ on the xOy plane. From the cosine theory, we have

$$|\mathbf{x}_m\mathbf{r}_s|^2 = (|\mathbf{x}_m\mathbf{r}_0| - p\cos\boldsymbol{\gamma}_m)^2 + p\sin\boldsymbol{\gamma}_m^2 \tag{7}$$

$$\cos\boldsymbol{\gamma}_m = \cos\boldsymbol{\beta}\cdot\cos(\boldsymbol{\beta}_1 + \boldsymbol{\gamma}_m'), \quad \cos\boldsymbol{\gamma}_m' = \boldsymbol{\epsilon}/p$$

$$\cos\boldsymbol{\beta} = \frac{\sqrt{\mathrm{L_{BR}^2} + x_m^2}}{\sqrt{\mathrm{L_{BR}^2} + \mathrm{H_B^2} + x_m^2}}, \quad \cos\boldsymbol{\beta}_1 = \frac{\mathrm{L_{BR}}}{\sqrt{\mathrm{L_{BR}^2} + x_m^2}}. \tag{8}$$

In the far-field scenario where $\mathrm{L_{BR}} \geq 10\mathrm{H_B} \gg x_m$, we can easily get the following approximations, i.e.,

$$\cos\boldsymbol{\beta} \approx 1, \quad \cos\boldsymbol{\beta}_1 \approx 1, \quad \boldsymbol{\beta}_1 \approx \boldsymbol{\beta} \approx 0, \boldsymbol{\gamma}_m' \approx \boldsymbol{\gamma}_m$$
$$\{p\sin\boldsymbol{\gamma}, p\cos\boldsymbol{\gamma}\} < p \ll \mathrm{L_{BR}} < |\mathbf{x}_m\mathbf{r}_0|. \tag{9}$$

In this case, we have

$$|\mathbf{x}_m\mathbf{r}_s|^2 \approx (|\mathbf{x}_m\mathbf{r}_0| - p\cos\boldsymbol{\gamma}_m)^2 \tag{10}$$

$$\Big||\mathbf{x}_m\mathbf{r}_0| - |\mathbf{x}_m\mathbf{r}_s|\Big| \approx p\cos\gamma_m = \frac{-L\boldsymbol{\epsilon} + x_m\boldsymbol{\xi}}{\sqrt{L^2 + x_m^2}}$$
$$\approx -\boldsymbol{\epsilon} + x_m\boldsymbol{\xi}/L \tag{11}$$

where $L = \sqrt{\mathrm{L_{BR}^2} + \mathrm{H_B^2} + x_m^2}, x_m = md/2$ while d denotes the inner element space in Array A. In addition, p, $\boldsymbol{\epsilon}$, and $\boldsymbol{\xi}$ represent the Euclidean spatial distance, range, and cross-range between $\mathbf{r}_0$ and $\mathbf{r}_s$ shown in Fig. 2. Substituting (11) into (6), $\eta_{y_L}^{\mathrm{1st}}$ can be rewritten as

$$\eta_{y_L}^{\mathrm{1st}} = K_L\exp(jk\boldsymbol{\epsilon})\sum_{m=1}^{M}\exp(jkx_m\boldsymbol{\xi}/L)$$
$$= K_{LM\exp(jk\boldsymbol{\epsilon})}\mathrm{sinc}\left(\frac{k\boldsymbol{\xi}\mathrm{A}_e}{2L}\right) \tag{12}$$

where $\mathrm{sinc}(t) = \sin(t)/t$.

Similar to (6), when $i \neq m$ in (5), we can get

$$\eta_{y_L}^{\mathrm{2nd}} = K_L \times \sum_{i=1}^{M}\sum_{\substack{m=1,\\ m\neq i}}^{M}\mathbb{E}_\phi\{\exp[j(k(|\mathbf{x}_i\mathbf{r}_0| - |\mathbf{x}_m\mathbf{r}_s|) + \phi_i - \phi_m)]\}. \tag{13}$$

The (i, m) pair of (13) is expressed as

$$\eta_{y_L}^{\mathrm{2nd}}(i,m) = \mathbb{E}_\phi\left\{\exp\underbrace{(\psi_{i,0} - \psi_{m,s})}_{\Psi_{im}} + \underbrace{(\phi_i - \phi_m)}_{\Phi}]\right\}$$
$$= \mathbb{E}_\phi\{\cos(\Psi_{im} + \Phi)\} + j\mathbb{E}_\phi\{(\sin(\Psi_{im} + \Phi))\} \tag{14}$$

where $\psi_{i,0} = k|\mathbf{x}_i\mathbf{r}_0|$, $\psi_{m,s} = k|\mathbf{x}_m\mathbf{r}_s|$.

Since ϕ_i and ϕ_m are uniformly distributed over $[-\pi, \pi)$, the probability density function (PDF) of $\Phi = (\phi_i - \phi_m)$ is

$$f_\Phi(\phi) = \begin{cases} 2\pi + \phi, & -2\pi \leq \phi \leq 0 \\ 2\pi - \phi, & 0 \leq \phi \leq 2\pi \\ 0, & \text{others.} \end{cases} \tag{15}$$

Then, the expectation in (14) can be reformulated as

$$\mathbb{E}_\phi\{\cos(\Psi_{im} + \Phi)\} = \int_{-2\pi}^{2\pi} f_\Phi(\phi)\cos(\Psi_{im} + \phi)d\phi$$
$$= \int_{-2\pi}^{2\pi} f_\Phi(\phi)\cos\Psi_{im}\cos\phi d\phi$$
$$- \int_{-2\pi}^{2\pi} f_\Phi(\phi)\sin\Psi_{im}\sin\phi d\phi. \tag{16}$$

Given $f_\Phi(\phi)$, we get that

$$\mathbb{E}_\phi\{\cos(\Psi_{im} + \Phi)\} = 0, \quad -2\pi \leq \phi \leq 0 \tag{17}$$

$$\mathbb{E}_\phi\{\cos(\Psi_{im} + \Phi)\} = 0, \quad 0 \leq \phi \leq 2\pi. \tag{18}$$

It is straightforward to obtain that

$$\mathbb{E}_\phi\{\cos(\Psi_{im} + \Phi)\} = 0, \mathbb{E}_\phi\{\sin(\Psi_{im} + \Phi)\} = 0$$
$$\eta_{y_L}^{\mathrm{2nd}}(i,m) = 0, \eta_{y_L}^{\mathrm{2nd}} = K_L \cdot \sum_{i=1}^{M}\sum_{\substack{m=1,\\ m\neq i}}^{M}\eta_{y_L}^{\mathrm{2nd}}(i,m) = 0. \tag{19}$$

Taking the summation of (12) and (19), η_{y_L} is given by

$$\eta_{y_L} = \eta_{y_L}^{\mathrm{1st}} + \eta_{y_L}^{\mathrm{2nd}} = K_{LM\exp(jk\boldsymbol{\epsilon})}\mathrm{sinc}\left(\frac{k\boldsymbol{\xi}\mathrm{A}_e}{2L}\right). \tag{20}$$

Given (4) and (20), the ACF of the received signal is

$$\eta_y(\mathbf{r}_0, \mathbf{r}_s) = \eta_y = \eta_{y_L} + \eta_{y_N} + \eta_N$$
$$= K_{LM\exp(jk\boldsymbol{\epsilon})}\mathrm{sinc}\left(\frac{k\boldsymbol{\xi}\mathrm{A}_e}{2L}\right) + K_N N J_0(kp) + \sigma^2\mathbf{I}. \tag{21}$$

D. ACFS of the Received Signal

In this section, we compute the ACFS of the received signal, i.e., the strength of ACF in (21) by first concluding and validating three properties of η_{y_L}, η_{y_N}, and η_{y_N} in (21).

Remark 1: η_{y_N} decays much faster than η_{y_L}.

Remark 2: At high SNRs, η_{y_N} is a constant term σ^2, which dose not impact the ACFS.

Remark 3: Given Remarks 1 and 2, ACFS of the received signal $y(t)$ is dominated by the ACFS of the LOS signal $y_L(t)$,

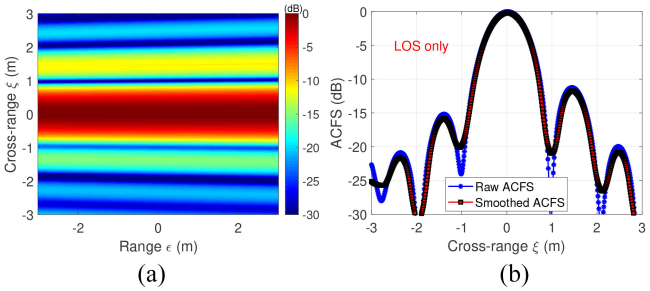


Fig. 3. ACFS distribution of LOS. (a) 2-D ACFS. (b) 1-D ACFS.

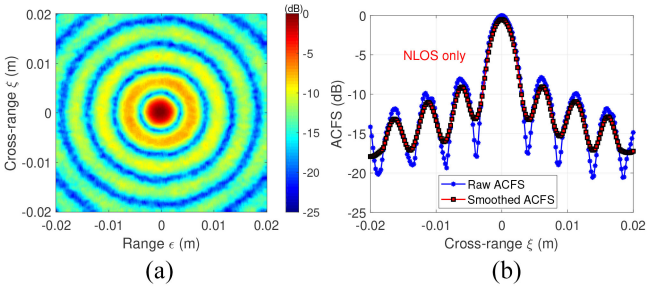


Fig. 4. ACFS distribution of NLOS. (a) 2-D ACFS. (b) 1-D ACFS.

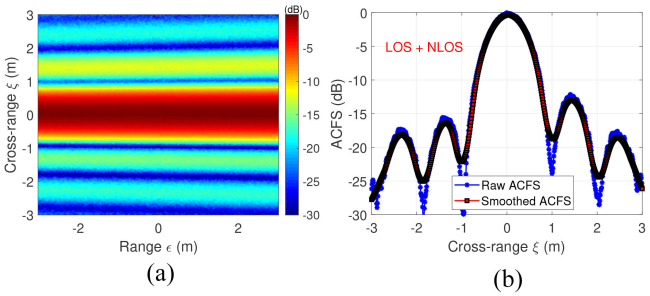


Fig. 5. ACFS distribution with both LOS and NLOS. (a) 2-D ACFS. (b) 1-D ACFS.

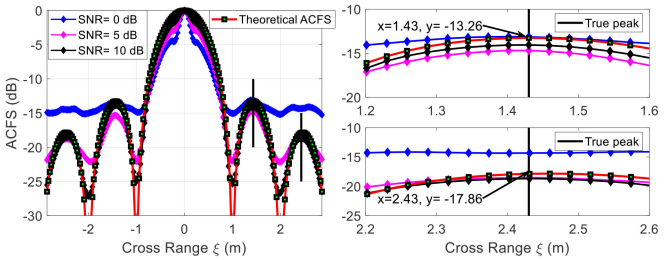


Fig. 6. ACFS distribution at different SNRs. [Theoretical ACFS is computed by (22) while others are simulated by (2).]

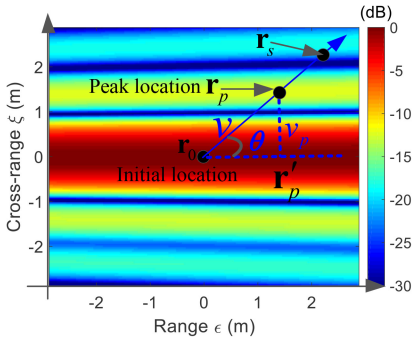


Fig. 7. Target moving illustration.

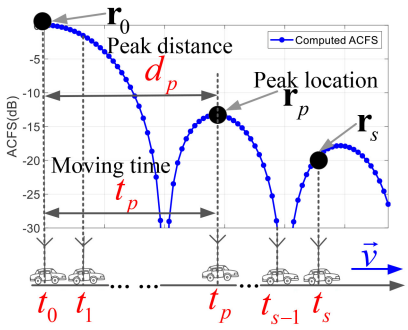


Fig. 8. ES illustration.

i.e., the normalized ACFS of $y(t)$ at two different locations $\mathbf{r}_0$ and $\mathbf{r}_s$ can be approximated by

$$\left|\eta_y(\mathbf{r}_0, \mathbf{r}_s)\right|^2 = \left|\eta_y\right|^2 \approx \left|\eta_{y_L}\right|^2 = \left|\text{sinc}\left(\frac{k\xi \text{A}_e}{2L}\right)\right|^2 \qquad (22)$$

which shows a focusing beam in spatial domain [see Fig. 5(a)].

To validate Remarks 1–3, we build a numerical simulation system using a massive MIMO antenna array with 100 elements at carrier frequency $f_0 = 28$ GHz. To be consistent with 5G small cell configurations, we set $H_B = 8$ m, $L_{\text{BR}} = 100$ m, $K_L = K_N$, and the multipath number N as an integer randomly selected between 10 and 100. Note that SNR is 10 dB in Figs. 3–5 while Fig. 6 explores the impact of SNR. The illustrations/definitions of target movement, peak distance p, and moving time t are given in Figs. 7 and 8.

Given the aforementioned parameters, theoretically, the peak distance of η_{y_L} is $p = 2.86L/k\text{A}_e \approx 1.432$ m, and $p \approx 0.61\lambda = 0.0061$ m corresponding to η_{y_N}, which match with our simulation results well in the positions of peaks and valleys as shown in Figs. 3(b) and 4(b). Moreover, Figs. 3 and 4 clearly show that the ACFS of the NLOS signal decays much faster than that of LOS, which validates Remark 1.

Equation (21) indicates that the constant term $\eta_n = \sigma^2\mathbf{I}$ does not impact η_y much at high SNR (i.e., σ^2 is much smaller compared with K_L and K_N). In Fig. 6, the theoretical ACFS according to (22) matches the ACFS directly computed by $\mathbb{E}[y(t_0)y^*(t_s)]$) when SNR $\geq$ 5 dB while it deviates a lot when SNR $\leq$ 0 dB. As a result, Remark 2 is verified. Given Remarks 1 and 2 and (21), it is straightforward to conclude that the ACFS of the received signal $y(t)$ is mainly dominated by the ACFS of the LOS signal especially when $K_{\text{LOS}} \geq K_{\text{NLOS}}$ in 5G massive MIMO system.[2] Thus, Remark 3 and (22) are verified. Fig. 5 shows the result when $K_{\text{LOS}} = K_{\text{NLOS}}$ and SNR $= 10$ dB, which also validates (22).

Note that when the target keeps moving, the line between the antenna center and $\mathbf{r}_0$ (i.e., $\overrightarrow{O\mathbf{r}_0}$) may not be perpendicular

[2]Practical measurements in [49] show that NLOS usually suffers an over 10-dB additional path loss than LOS signal in 5G massive MIMO system due to the greater traveling distance and absorption of corresponding scatterers.

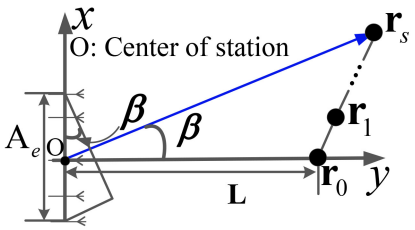


Fig. 9. Geometrical relationship between the target and BS when the target moves from $\mathbf{r}_0$ to $\mathbf{r}_s$.

to the line along which the antennas are deployed (i.e., $\overrightarrow{Ox}$). As shown in Fig. 9, $\overrightarrow{O\mathbf{r}_0} \perp \overrightarrow{Ox}$ does not hold when the target is moving. In this case, the effective aperture A_e in (22) should be replaced with $A_e \cos\beta$. Correspondingly, the distance L should be replaced by $L/\cos\beta$ [47].

III. Moving Speed and Direction Estimation

In this section, we first provide an overview of the ACFS-based tracking system. Then, we present a novel ACFS matching method to estimate the moving speed and direction simultaneously by leveraging the RF signals only. For description clarity, we define the range- and cross-range direction in Fig. 7 while the peak distance d_p and moving time t_p are illustrated in Fig. 8.

A. Overview of the ACFS-Based Tracking System

Consider that a target moves at a speed of v along the line joining $\mathbf{r}_0$ and $\mathbf{r}_s$ as shown in Fig. 7. The receiver is fixed on the target and keeps recording signals transmitted from the BS with a sample rate f_s. The proposed method estimates the moving trajectory of the receiver, i.e., the location of the receiver at time t_s can be estimated by

$$\mathbf{r}_s = \mathbf{r}_{s-1} + \Delta\mathbf{r}_s = \mathbf{r}_{s-1} + v\Delta t = \mathbf{r}_{s-1} + v_p \Delta t/\sin\theta \quad (23)$$

where $\mathbf{r}_{s-1}$ denotes the location of the receiver at t_{s-1}, while $\Delta t = 1/f_s$ denotes the sample period. The proposed system continuously searches for the peak location $\mathbf{r}_p$ of the computed ACFS, i.e., $|\mathbb{E}[y(t)y^*(t+\tau)|^2, t = t_0, t_1, \ldots$ It then estimates the consecutive v_p and θ, thus yielding the real-time tracking of a moving target.[3] In Fig. 7, we name v as the *absolute speed* while $v_p = v\sin\theta$ is the *projected speed*, which represents the projection of v along the cross-range direction (i.e., $\overrightarrow{\mathbf{r}'_p\mathbf{r}_p}$).

B. Projected Speed Estimation

As shown in (22), a moving target keeps receiving signals transmitted from the massive MIMO antennas on the BS. Then, the computed ACFS of the measured signal $y(t)$ at the receiver side is a sampled version of the theoretical ACFS $|\text{sinc}([k\boldsymbol{\xi}A_e/2L])|^2$, where $\boldsymbol{\xi}$ is the cross-range between $\mathbf{r}_0$ and $\mathbf{r}_s$ (see Figs. 2 and 7). As a result, we extract the first local peak of the theoretical ACFS ($|\text{sinc}([k\boldsymbol{\xi}A_e/2L])|^2$). The peak distance d_p in Fig. 8 is given by

$$d_p = 2.86L/kA_e. \quad (24)$$

[3]In case of outliers, popular smoothing techniques, such as moving average and local regression [54], can be further used to improve the robustness.

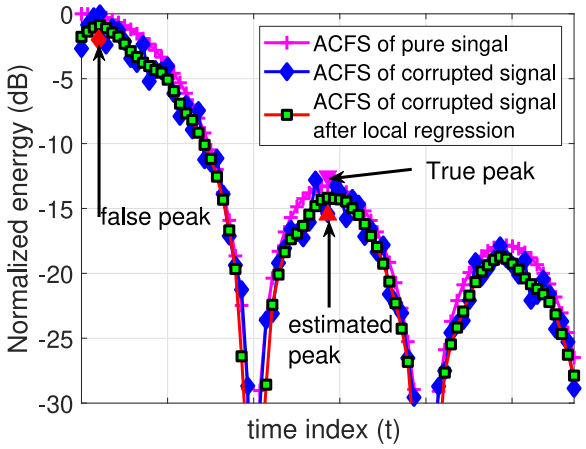


Fig. 10. Curve fitting by local regression.

Note that L denotes the distance between the BS center and the initial location. Similarly, we look for the first local peak of the computed ACFS of $y(t)$. Then, the moving time t_p can be estimated by

$$\hat{t}_p = \underset{\tau\in\{0,\Delta t,2\Delta t,\ldots,T_{\text{ACFS}}\}}{\arg\,\textbf{FindPeak}} \left\{ \left|\mathbb{E}[y(t_0)y^*(t_0+\tau)\right|^2 \right\} \quad (25)$$

where operation **FindPeak**$\{\bullet\}$ means looking for the first peak and T_{ACFS} is the time window length within which the first peak may fall in. Given d_p and $\hat{t}_p$, the *projected speed* estimation is expressed as $\hat{v}_p = d_p/\hat{t}_p$.

Note that in practice, we first apply a local regression [54] on the ACFS distribution curve to get rid of the spikes caused by noise or other distortions. Numerical simulation in Fig. 10 shows that when the signal is corrupted, it is difficult to find the true peak directly. However, after local regression, we can get a very good estimation of the true peak. Fig. 10 also shows that a false peak very close to the reference point ($t = 0$) may mislead the peak finding and thus induce large errors. However, this can be eliminated by using peak distance d_p at the previous time (which is known) as a new constraint. Specifically, the distance L between the BS center and receiver cannot change much during two adjacent measurements due to the high sample rate and limited moving velocity. As a result, d_p which is determined by L in (22) cannot change very much as well.

C. Absolute Speed and Moving Direction Estimation

In addition to the projected ES $\hat{v}_p$, this section introduces how to estimate the absolute speed and moving direction of the target in order to track a moving target continuously. We consider a practical multiple-BS case with one user/receiver and Q-based stations. For notation purpose, let $\hat{v}_{p,q}$ denote the projected speed estimated from BS q. Note that the absolute moving speed v of a moving target is unique and can be estimated by

$$v = \frac{\hat{v}_{p,q}}{\sin\theta_q}, q = 1, 2, \ldots, Q$$
$$\text{s.t.} \quad \theta_q + \theta_l = 180 - \Omega_{ql}, q \neq l, q, l \in \{1, 2, \ldots, Q\} \quad (26)$$

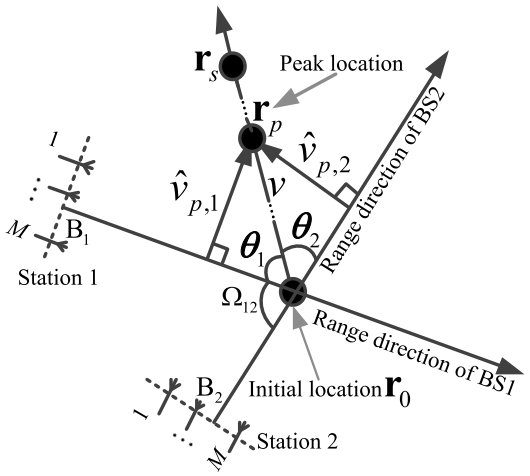


Fig. 11. Example scenario of two BSs.

where θ_q represents the angle between the moving direction ($\overrightarrow{\mathbf{r}_0\mathbf{r}_s}$) and the range direction ($\overrightarrow{\mathrm{B}_q\mathbf{r}_0}$) corresponding to station q centered at B_q (see Fig. 11). In (26), Ω_{ql} is the angle among the initial location $\mathbf{r}_0$, station B_l, and station B_q with vertex $\mathbf{r}_0$, which is known *a priori* since the location of the BSs and the initial location are easy to get in communication systems. Fig. 11 gives an example of two BSs, i.e., Q = 2. Then, (26) becomes

$$\begin{aligned} v &= \frac{\hat{v}_{p,1}}{\sin\theta_1} = \frac{\hat{v}_{p,2}}{\sin\theta_2} \\ \text{s.t.}\ \ &\theta_1 + \theta_2 = 180 - \Omega_{12}. \end{aligned} \tag{27}$$

In (27), $\hat{v}_{p,1}$ and $\hat{v}_{p,2}$ can be estimated by using the ACFS of the received signal as shown in Section III-B and Ω_{12} is known *a priori*. Thus. the *moving direction* θ_1 and θ_2 can be estimated by

$$\begin{cases} \hat{\theta}_1 = \arctan\left(\frac{\hat{v}_{p,1}\sin\Omega_{12}}{\hat{v}_{p,2}+\hat{v}_{p,1}\cos\Omega_{12}}\right), \\ \hat{\theta}_2 = \arctan\left(\frac{\hat{v}_{p,2}\sin\Omega_{12}}{\hat{v}_{p,1}+\hat{v}_{p,2}\cos\Omega_{12}}\right). \end{cases} \tag{28}$$

To improve the robustness and accuracy, we explore the different combining pairs of the BSs ($\mathrm{B}_q, \mathrm{B}_l, l \neq q$), if any. Similar to the ($\mathrm{B}_1, \mathrm{B}_2$) pair shown in Fig. 11, we can further get corresponding projected ESs ($\hat{v}_q, \hat{v}_l$) and moving direction estimations ($\hat{\theta}_q, \hat{\theta}_l$). Then, the *absolute speed* can be estimated by

$$\hat{v} = \frac{1}{Q}\sum_{q=1}^{Q}\frac{\hat{v}_{p,q}}{\sin\theta_q}, q = 1, 2, \ldots, Q. \tag{29}$$

IV. Target Localization

In this section, the location estimation is first calculated by integrating the consecutive moving speed and moving direction estimations. Then, the location estimations from different BS pairs ($\mathrm{B}_q, \mathrm{B}_l, l \neq q$) are fused to improve the robustness and accuracy. To have a high-level understanding of the algorithm, the architecture and main steps are summarized in Fig. 12.

Recalling the estimations of the absolute speed $\hat{v}$ and moving directions $\theta_q (q = 1, 2, \ldots, Q)$ from (28) and (29), as shown in Fig. 13, the new location $\mathbf{r}_{T_M}^{\mathrm{B}_q}$ estimated by station q in the local coordinate system $x_{B_q}Oy_{B_q}$ can be expressed as

$$\begin{cases} r_{T_M,x_{B_q}} = r_{0,x_{B_q}} - d_{T_M}\cos\theta_q \\ r_{T_M,y_{B_q}} = r_{0,y_{B_q}} + d_{T_M}\sin\theta_q \end{cases}, \quad q = 1, 2, \ldots, Q \tag{30}$$

where $d_{T_M} = \hat{v}T_M$, T_M is the updating window length, meaning that we update the location estimation every T_M seconds. $(r_{0,x_{B_q}}, r_{0,y_{B_q}})$ and $(r_{T_M,x_{B_q}}, r_{T_M,y_{B_q}})$ are the coordinates of the initial location $\mathbf{r}_0$ and the new location $\mathbf{r}_{T_M}^{\mathrm{B}_q}$ at the local coordinate system $x_{B_q}Oy_{B_q}$ shown in Fig. 13. We then transform the local coordinates of $\mathbf{r}_{T_M}^{\mathrm{B}_q}$ into the global Cartesian coordinate system xOy, which is denoted as $\mathbf{r}_{T_M}^{q}$ shown in the magenta color in Fig. 14. As a result, the coordinate $\mathbf{r}_{T_M}^{q} = (r_{T_Mx}^{q}, r_{T_My}^{q})$ can be calculated by

$$\begin{bmatrix} r_{T_Mx}^{q} \\ r_{T_My}^{q} \end{bmatrix} = \begin{bmatrix} r_{T_M,x_{B_q}} & r_{T_M,y_{B_q}} \\ r_{T_M,y_{B_q}} & -r_{T_M,x_{B_q}} \end{bmatrix}\begin{bmatrix} \cos\zeta_q \\ \sin\zeta_q \end{bmatrix} \tag{31}$$

where ζ_q is the angle between the global Cartesian coordinate system xOy and the local Cartesian coordinate system $x_{B_q}Oy_{B_q}$, and is known *a priori* in modern communication systems. Furthermore, we fuse the location estimations from different BSs, i.e.,

$$\mathbf{r}_{T_M} = \frac{1}{Q}\sum_{q=1}^{Q}\mathbf{r}_{T_M}^{q}, q = 1, 2, \ldots, Q \tag{32}$$

where $\mathbf{r}_{T_M}^{q} = (r_{T_Mx}^{q}, r_{T_My}^{q})$.

Once we get the global coordinates of the new location $\mathbf{r}_{T_M} = (r_{T_Mx}, r_{T_My})$, the distance between the qth station and the receiver/target can be updated by

$$L_{B_q\mathrm{R}}^{\mathrm{new}} = \sqrt{\left(\mathbf{r}_{T_Mx} - O_{x_q}\right)^2 + \left(\mathbf{r}_{T_My} - O_{y_q}\right)^2} \tag{33}$$

where (O_{x_q}, O_{y_q}) are the coordinates of the qth station center B_q at the global coordinate system. As a consequence, according to (22), the new peak distance $d_{p_{\mathrm{New}}}^{q}$ corresponding to the qth BS can be updated by

$$d_{p_{\mathrm{New}}}^{q} = 2.86L_{B_q\mathrm{R}}^{\mathrm{new}}/k\mathrm{A}_e. \tag{34}$$

In the next step, we take $\mathbf{r}_{T_M}$ as the new starting point to repeat the ACFS computation (based on the data measurements starting at the timestamp corresponding to $\mathbf{r}_{T_M}$), ES and localization process, thus getting a location estimation sequence $\mathbf{r}_{T_M}(t)$ representing the trajectory of the moving target.

V. Performance Analysis

In this section, we perform theoretical analysis about the expected error of the speed and location estimation using the proposed algorithm. Since the system estimates the time t_p corresponding to the first local peak of the computed ACFS (as in Section III-B), we first derive the distribution of the peak-location-error (PLE) measured by the distance that the estimated peak deviates from the true peak. The expected-error-of-speed-estimation (EES) and the expected-error-of-localization (EEL) are further derived on the base of the PLE distribution.

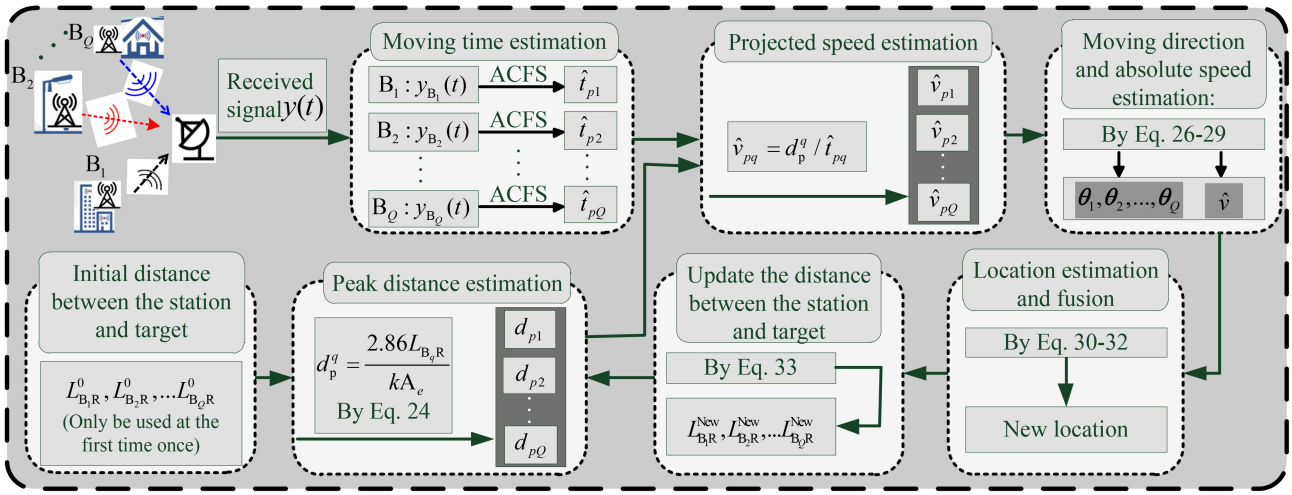


Fig. 12. Flow chart of the proposed target location system.

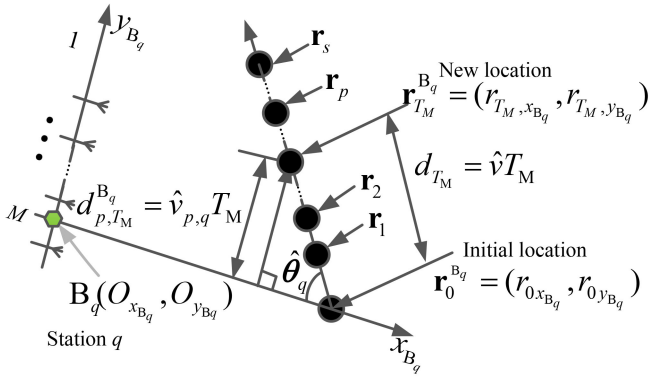


Fig. 13. Target localization at station q.

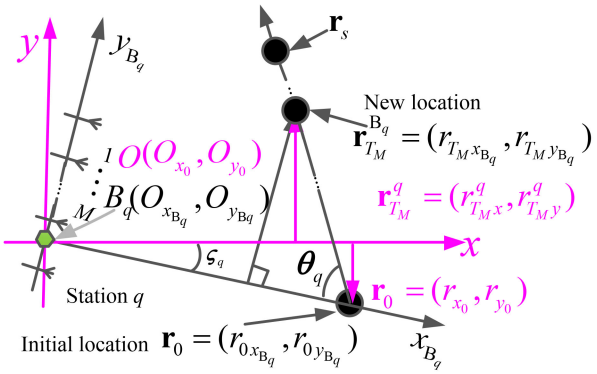


Fig. 14. Coordinates transformation.

A. Peak Location Error Distribution

To derive the PLE, we first introduce an intermediate variable peak prominence [55] as shown in Fig. 15, which indicates the relative height of a peak. In general, a larger prominence corresponds to a sharper peak and thus, the peak can be localized more accurately. Recalling (21) and Section II-D, the ACFS is given by

$$|\eta_y|^2 = \left| \frac{K_L M \mathrm{sinc}(k\boldsymbol{\xi} \mathrm{A}_e/2L) + K_N N J_0(kp) + \sigma^2}{K_L M + K_N N + \sigma^2} \right|^2. \quad (35)$$

The height p_{h} of the first local peak and height p_{v} of first local valley of (35) (see Fig. 15) can be expressed as

$$p_{\mathrm{h}} = \left| \frac{0.22K_L M + 0.01K_N N + \sigma^2}{K_L M + K_N N + \sigma^2} \right|^2$$
$$p_{\mathrm{v}} = \left| \frac{\sigma^2}{K_L M + K_N N + \sigma^2} \right|^2 \quad (36)$$

where we have $\mathrm{sinc}(k\boldsymbol{\xi_0}\mathrm{A}_e/2L) = 0.22$ and $J_0(kp) = 0.01$ while $\boldsymbol{\xi_0}$ is the first local peak location of (35) and $p = \sqrt{\boldsymbol{\xi}_0^2 + \boldsymbol{\epsilon}^2} \geq \boldsymbol{\xi}_0 = vt\sin\theta$ (see Fig. 7). Then, the prominence p_{pro} of the first local peak of (35) in the unit of decibel (dB) can be expressed as

$$p_{\mathrm{pro}} = 10\log_{10}[p_{\mathrm{h}} - p_{\mathrm{v}}]$$
$$= 10\log_{10}\left| \frac{0.22K_L M + 0.01K_N N + \sigma^2}{\sigma^2} \right|^2. \quad (37)$$

Note that SNR is defined as $\mathrm{SNR} = 20\log_{10}(K_L M/\sigma^2)$. Consequently, the prominence p_{pro} can be rewritten as

$$p_{\mathrm{pro}} = 10\log_{10}\left| \frac{0.21K_L M + 0.01K_N N + K_L M \cdot 10^{\frac{-\mathrm{SNR}}{20}}}{K_L M \cdot 10^{\frac{-\mathrm{SNR}}{20}}} \right|^2. \quad (38)$$

To have a better standing, Fig. 15 shows the peak prominence versus difference SNRs. It is clear that the prominence p_{pro} increases monotonically with the increment of SNR, thus improving the peak localization accuracy. Note that the system needs to estimate the moving distance of the target every T_M seconds as introduced in (30). As a result, we have to repeat the peak finding process for a large number of times to track a moving target. Then, by using the central limit theorem [56], the expectation of the PLE denoted by p_{err} can be assumed to follow a Gaussian distribution, i.e., the PDF of p_{err} can be expressed as

$$f(p_{\mathrm{err}}) = H(p_{\mathrm{pro}})\exp\left(-\frac{p_{\mathrm{err}}^2}{2G(p_{\mathrm{pro}})^2}\right) \quad (39)$$

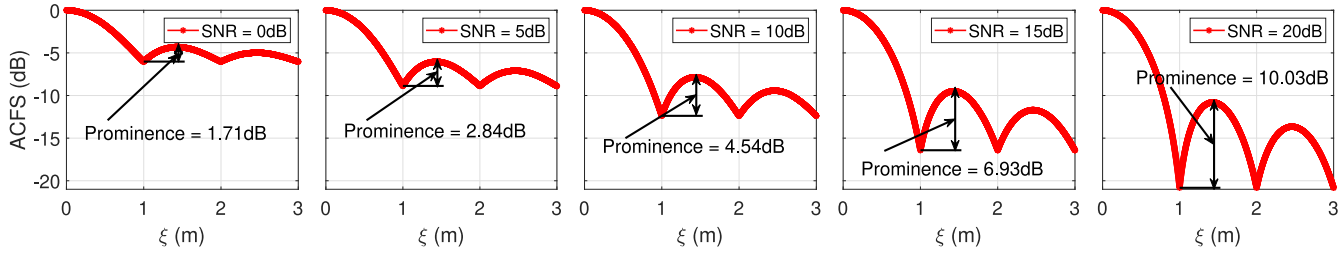


Fig. 15. Peak prominence versus SNRs.

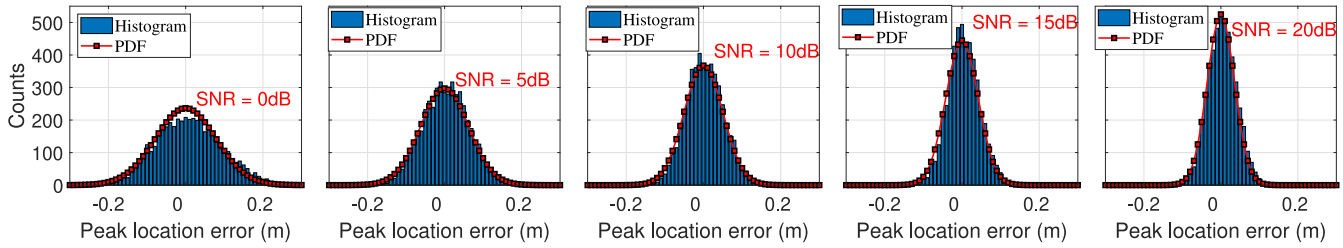


Fig. 16. Peak location error versus SNRs.

where $H(p_{\text{pro}})$ is the coefficient function while $G(p_{\text{pro}})$ denotes the standard deviation function (SDF). Since p_{err} decreases with the increment of SNR, $H(p_{\text{pro}})$ is a monotonically increasing function while $G(p_{\text{pro}})$ is a decreasing function. It is preferable that $H(p_{\text{pro}})$ grows slowly and $G(p_{\text{pro}})$ decreases slowly as their arguments increase. Here, we propose a pair of empirical approximations about $H(p_{\text{pro}})$ and $G(p_{\text{pro}})$ by 5000 Monte Carlo experiments, i.e.,

$$H(p_{\text{pro}}) = 186\sqrt{10\log_{10}(p_{\text{pro}})}$$
$$G(p_{\text{pro}}) = \frac{1}{10\sqrt{10\log_{10}(p_{\text{pro}})}}. \tag{40}$$

Fig. 16 shows that the PDF $f(p_{\text{err}})$ with $H(p_{\text{pro}})$ and $G(p_{\text{pro}})$ given in (40) can well approximate the distribution of PLE. Consequently, the expectation of $|p_{\text{err}}|$ can be calculated by

$$\mathbb{E}\{|p_{\text{err}}|\} = \int_{-\infty}^{\infty} |p_{\text{err}}| f(p_{\text{err}}) = 2G(p_{\text{pro}})^2 \cdot H(p_{\text{pro}}). \tag{41}$$

B. Expected Error of the Speed and Location Estimations

Similar to the scenario in Section III, a target is assumed to move at a speed of v along the line joining $\mathbf{r}_0$ and $\mathbf{r}_s$ and receive signals from Q nearby BSs. Given the estimation $\hat{t}_p = N\Delta t$ where N is the integer denoting the sample index in Section III-B, the expected moving-time-estimation error caused by $|p_{\text{err}}|$ at BS q can be expressed as

$$\hat{t}^q_{p_{\text{err}}} = \left(\Delta N + \left\lfloor \frac{\mathbb{E}\{|p_{\text{err}}|\}}{v\sin\theta_q \Delta t} \right\rfloor\right) \cdot \Delta t, q = 1, 2, \ldots, Q \tag{42}$$

where $\Delta N \in \{0, \pm 1\}$ represents the quantization error. As a result, the EES is given by

$$\hat{v}_{\text{err}} = \frac{1}{Q}\sum_{q=1}^{Q}\left|\frac{d^q_p}{\hat{t}^q_p \sin\theta_q} - \frac{d^q_p}{\sin\theta_q(\hat{t}^q_p \pm \hat{t}^q_{p_{\text{err}}})}\right|$$
$$= \frac{1}{Q}\sum_{q=1}^{Q}\frac{d^q_p \hat{t}^q_{p_{\text{err}}}}{\sin\theta_q \hat{t}^q_p(\hat{t}^q_p \pm \hat{t}^q_{p_{\text{err}}})} = \frac{1}{Q}\sum_{q=1}^{Q}\frac{\hat{v}^q \hat{t}^q_{p_{\text{err}}}}{(\hat{t}^q_p \pm \hat{t}^q_{p_{\text{err}}})} \tag{43}$$

TABLE I
PARAMETERS USED IN THE SIMULATIONS

Sample frequency f_s	28GHz
Number of antennas M	100
Coverage of 5G base station	200m
Signal-to-noise-ratio SNR	10dB
LOS and NLOS power	$K_L \geq K_N$
Number of NLOS signal N	Integer randomly selected from [10, 100]
Number of base station Q	2
Speed of the target v	within [5, 30]m/s

where $\hat{v}^q = \hat{v}^q_p / \sin\theta_q$ denotes the ES corresponding to the q*th* BS. Then, the EEL can be expressed as $\mathbf{r}^{\text{err}}_{T_M} = \hat{v}_{\text{err}} T_M$.

VI. SIMULATION RESULTS

In this section, simulations are conducted to evaluate the performance of the proposed method based on a 5G communication system. The default parameter used in the Monte Carlo experiments is listed in Table I, if not otherwise stated. In summary, six experiments are performed to evaluate the proposed approach: 1) overall performance; 2) speed and location estimation error; 3) impact of the number of antennas M; 4) impact of the sample rate; 5) impact of the SNR; and 6) comparison with existing works.

A. Overall Performance

Assume the SNR is 10 dB, Fig. 17 depicts the simulation result when the target moves with a variable acceleration from $t = 0$ s to $t = 2$ s, a constant speed from $t = 2$ s to $t = 4$ s and a variable deceleration from $t = 4$ s to $t = 5$ s. Table II shows 12 different moving situations, including different initial speeds, accelerations, decelerations, and turning angles to

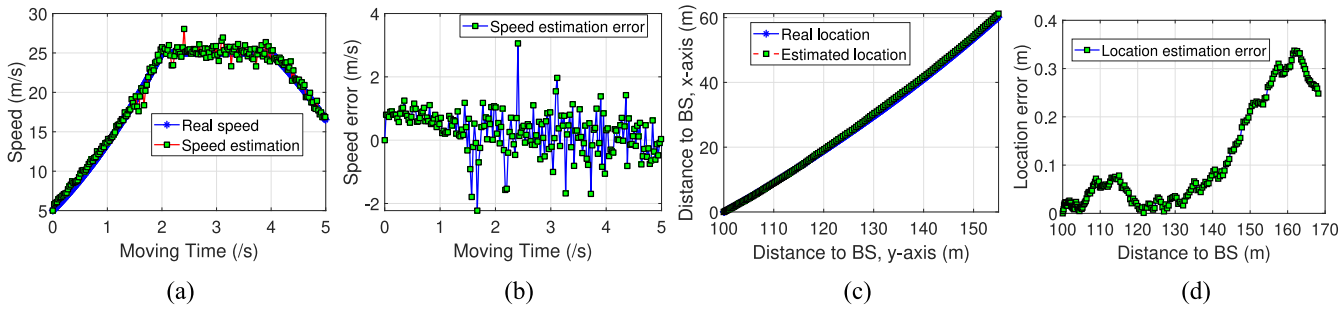


Fig. 17. Speed and location estimation results with variable speeds (directions of x-axis and y-axis are shown in Fig. 2). (a) ES. (b) ES error. (c) Location estimation. (d) Location estimation error.

TABLE II
MEAN VELOCITY AND LOCATION ESTIMATION ERROR IN DIFFERENT SITUATIONS

Situation	Velocity		Angle		RMSE of velocity estimation error	RMSE of location estimation error
	Initial (m/s)	Acceleration (m/s^2)	Initial ($\circ$)	Acceleration ($\circ/s^2$)	(cm/s)	(cm)
1	10	0 [1]	40	0 [2]	44.32	5.45
2	10	0	45	2 [2]	45.61	10.36
3	10	0	50	6	40.67	9.44
4	20	0	45	0	57.89	18.81
5	20	0	60	6	69.63	11.28
6	30	0	40	0	78.21	19.33
7	30	0	65	3	81.41	15.23
8	5	6 [1]	45	0	47.93	14.21
9	5	6 (2) [3]	40	0	44.04	10.66
10	5	6 (2) [3]	50	2	35.47	24.89
11	5	7 (3) [3]	45	3	47.87	13.77

[1] Acceleration of the velocity being 0 means the velocity is constant while nonzero means it is changing as the target is moving.
[2] Acceleration of the angle being 0 means direction angle is constant, i.e., the trajectory is a straight while nonzero means the direction is changing, which indicates a curved trajectory.
[3] The number in the bracket () means the acceleration is also changing during the process.

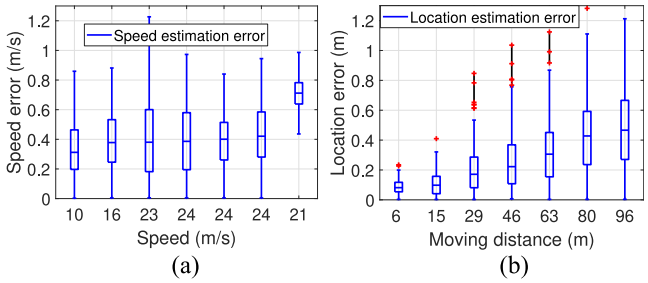


Fig. 18. Estimation error distribution. (a) ES error. (b) Location estimation error.

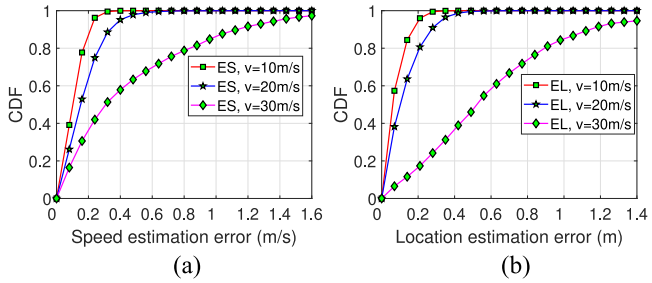


Fig. 19. CDF of estimation error. (a) ES error. (b) Location estimation error.

further verify the proposed method. For example, in situation "11," the target starts moving at a speed of 5 m/s, acceleration of 7 m/s^2, and angle [i.e., θ_1 in Fig. 11(a)] of 45°. Moreover, the acceleration is also varying with a rate of 3 m/s^2. Similarly, the angle θ_1 is changing at a rate of 3° per second to create a curved trajectory. Overall, we can conclude that our method can track the moving object with decimeter or even better centimeter accuracy in different scenarios.

B. Speed and Location Estimation Error

To evaluate the error of ES and error of localization (EL), we conduct 1000 independent Monte Carlo simulations in which the target moves at variable speeds along a curved trajectory. The SNR is fixed at 10 dB and the error distribution is shown in Fig. 18. Then, the empirical cumulative distribution function (CDF) corresponding to different velocities ($v = 10, 20, 30$ m/s) is given in Fig. 19. As illustrated in

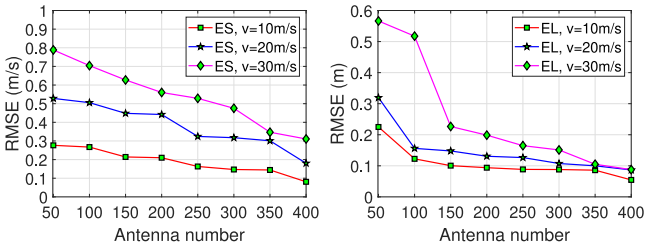


Fig. 20. Estimation error versus antenna number. (a) ES error. (b) Location estimation error.

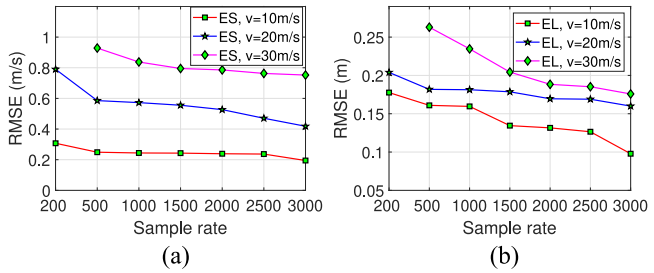


Fig. 21. Estimation error versus sample rate. (a) ES error. (b) Location estimation error.

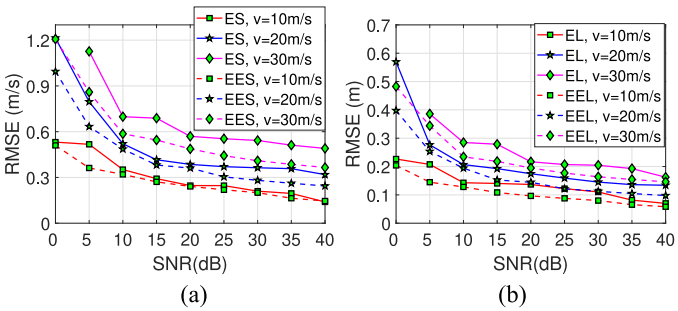


Fig. 22. Estimation error versus SNRs. (a) ES error. (b) Location estimation error.

Fig. 18, our method achieves high-accuracy ES results with a median error of 0.4 m/s. Fig. 19 indicates that the median ES errors are about 0.18, 0.26, and 0.45 m/s while the corresponding location errors are about 0.06, 0.12, and 0.53 m. Moreover, when $v \leq$ 20m/s (45 mph), the 80 percentile of the ES error is within 0.25 m/s while the location estimation error is less than 0.2 m. Therefore, our method has a promising performance in urban areas where $v \leq 20$ m/s generally holds and there are strong NLOS. Moreover, Fig. 18(b) shows that the location estimation error accumulates at a moderate rate as the object moves continuously. This is mainly because the estimation of the next location is highly dependent on the previous location estimation result, which causes accumulative errors. Another possible reason is that SNR of the received signal drops with the target moving toward cell boundary. In the future, locations of the nearby BSs may be used to mitigate the accumulative error, which we leave for future work.

C. Impact of the Number of Antennas

Fig. 20 shows the root mean-square error (RMSE) of the speed and location estimation versus different antenna number M. Evidently, both the speed and location estimation accuracy are improved with the increment of M. Specifically, when M is less than 100, it may not work well when the velocity is too high (e.g., $v = 30$ m/s). However, our system can localize the target within 0.3 m error when M is no less than 100. This is because that as M increases, we can harvest more signal components and thus get more accurate ACFS estimation. The performance starts to saturate when $M \geq 200$. Note that when M approaches to 400, the location estimation error can be as low as 8 cm.

D. Impact of the Sample Rate

Fig. 21 further explores impacts of the sample rate on our method. In general, higher sample rate improves the estimation accuracy. For a fixed sample rate, the object moving at a higher speed suffers from a worse location accuracy than that moving at a lower speed, which is consistent with (42). Moreover, Fig. 21 demonstrates that the minimum sample rate is a moderate value about 500 Hz. Further reducing the sample rate may lead to worse ES error or even failure especially for high-speed objects, e.g., $v = 30$ m/s. When the sample rate further increases, such as 3000 Hz, the ES error is no greater than 0.6 m/s while the minimum location estimation error is about 10.8 cm.

E. Impact of SNR and the Number of Stations

Fig. 22 explores the system performance versus different SNRs with $M = 100$ and a sample rate 1000 Hz. The EES and EEL are shown in dotted lines with corresponding markers. Clearly, the system is seriously impacted by noise when SNR is less than 10 dB and does not work when the target moves at 30 m/s if SNR further decreases to 0 dB. Moreover, the object with a higher speed shows a worse estimation error than that with a lower speed. However, in all the scenarios, our method works well when SNR is no less than 10 dB, which is easy to meet in a typical communication system.

F. Impact the Number of Stations

A unique feature of the proposed method is that it jointly explores the directional ACFS distribution of the received signal and the geometric relationship among multiple BSs to estimate the moving direction of a target without any further information. We investigate how the performance would change with the number of BSs. As shown in Fig. 23, the performance is improved by fusing the information from more surrounding stations. However, the proposed system can achieve very good location accuracy with only two stations. In practice, users can flexibly select the number of stations according to the requirements of system latency, complexity, and accuracy for real applications.

G. Comparison With Existing Works

In this section, we compare the proposed method with corresponding existing works in the aspect of speed, direction,

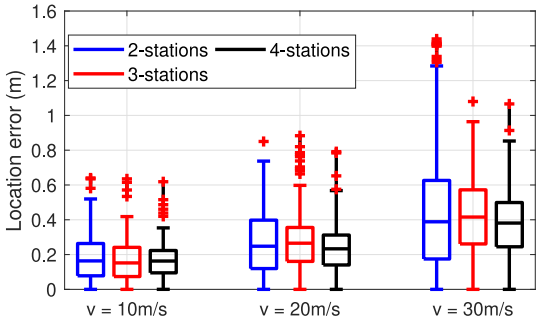


Fig. 23. Location error versus the number of stations.

TABLE III
COMPUTATIONAL COMPLEXITY COMPARISONS

Algorithms	Computational complexity
DiSouL	$O\{f_s^2/T_{\mathrm{M}}(QL+\sum_q N_q)^{3.5}\}$
Proposed	$O\{(Qf_s+f_s+Q^2+7Q)f_s/T_{\mathrm{M}}\}$
New notations	
T_{M}	output estimated location every T_{M} seconds.
L	number of location grid in candidate area
N_q	number of NLOS components at station q

location estimation accuracy, and complexity. To simulate a typical localization and tracking scenario, we assume a moving target that continues recording signals transmitted from two surrounding 5G massive MIMO BSs. The two stations are 200 m away from each other and equipped with $M = 100$ antennas on each station. Considering the rich-scattering urban environment, we randomly choose N (within 10–100) NLOS components[4] while the impinging angles follow uniform distributions over $(-\pi, \pi]$. The target starts to move at a speed of 5 m/s, accelerates 2 s, then keeps constant speed for 2 s, and finally decelerates until end. In total, the target moves 80 ms away from the starting point.[5]

Speed Estimation: Fig. 24(a) compares the ES performance of the proposed system with the existing SenSpeed [57], WiFiDetect [58], and GPS [59] methods. Clearly, our method outperforms the benchmark algorithms in accuracy. Specifically, SenSpeed estimates the speed by assuming that the error of the integrated acceleration accumulates linearly over time, which dose not always hold in practice. WiFiDetect considers only one dominant NLOS signal and the GPS method relies on the LOS signal greatly, which is vulnerable to NLOS distortions in urban rich-scattering environments. However, by exploring the ACFS of the received signal, the proposed method treats all the LOS and NLOS signal components as a whole and thus improves the accuracy.

Direction Estimation: Fig. 24(b) shows that the proposed system can achieve 1.8° direction estimation accuracy, with improvement than the existing DOA approaches, i.e., Capon [60], ESPRIT [21], and MUSIC [20]. This mainly benefits from the ACFS which has been proved to be tolerable to NLOS distortions. On the contrary, most DOA approaches are highly dependent on the time measurement accuracy of the LOS signal, which is easily to be distorted/impacted by NLOS signals in practice.

Location Estimation: Fig. 24(c) demonstrates the location accuracy of the proposed method and the state-of-the-art techniques, including DiSouL [23], Conv-fingerprint [31], and DNN-fingerprint [32]. From Fig. 24(c), the proposed method can achieve less than 0.2-m error with the percentile ≥95%. However, the 95% percentile estimation error of the DNN-fingerprint is 1.2 m while both Conv-fingerprint and DiSouL cannot offer ≥95% confidence with less than 2-m error. Overall, our method is more robust because it explores the statistical ACFS of the received signal, which is very stable in 5G massive MIMO systems regardless of the environment. However, fingerprint-based methods may suffer from fingerprint mismatch issue due to the change of the wireless propagation environment, thus degrading the accuracy. In dynamic environments, DiSouL works even worse because there are two hyperparameters in the model, which are sensitive to the environment changes.

Complexity: Considering that complexity is very important for real-time tracking and navigation applications with a stringent requirement on the latency, Table III compares the complexity between our algorithm and the state-of-the-art DiSouL [23] approach. In DiSouL, the main computation comes from solving the second-order cone program (SOCP) problem, which is about $(QL+\sum_q N_q)^{3.5}$ in a single snapshot [23]. Our computation is mainly caused by ACFS computation, 1-D peak searching in (25) and the location computation from (28) to (32), which are Qf_s, f_s, and Q^2+7Q, respectively. Since $L \geq Q$, our method is much cheaper than DiSouL.

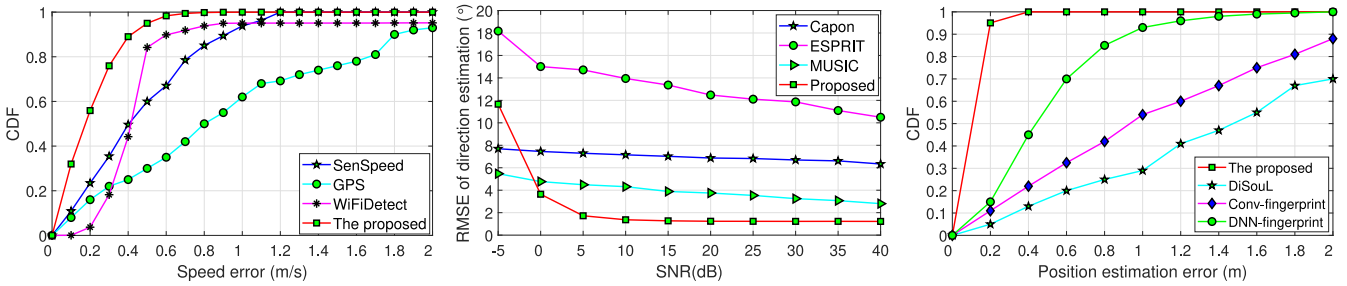
Note that we do not compare with Conv-fingerprint [31] and DNN-fingerprint [32] here because they are all training-based methods in which the overhead in map construction stage is hard to be quantified. To improve accuracy, fingerprint-based methods usually require a lot of training and updating to construct the offline map, which leads to a prohibitive overhead especially in a dynamic environment.

VII. CONCLUSION

This article proposes a high-accuracy target location method based on the 5G massive MIMO system. We first prove the existence of a sinc-like focusing beam in a massive MIMO system by computing the statistical autocorrelation of the received signal in the far-filed scenario. Based on the focusing beam, an ES algorithm is then proposed by jointly using the relative ESs with respect to multiple BSs. Give an initial point, we develop a target localization method by further using the geometrical relationships between multiple BSs. Theoretical error analysis and extensive numerical simulations show that our method can achieve decimeter localization accuracy by

[4]Many practical measurements in the New York City [45], [46] validated that the number of dominate NLOS signal usually varies from 10 to 100.

[5]After moving 80 m, the target is about 180 m from one of the two stations, which is close to the cell boundary of the station. In this case, we need to switch to closer BSs, and detailed cell switching procedure is out of the scope of this article.



(a) (b) (c)

Fig. 24. Performance comparison. (a) ES error. (b) Direction estimation error. (c) Location estimation error.

computing ACFS of the received signal, which outperforms many prior works in accuracy and cost.

APPENDIX

In this Appendix, we prove (3), i.e., $\eta_y(\mathbf{r}_0, \mathbf{r}_s) = \mathbb{E}[y(t_0)y^*(t_s)] \approx \eta_{y_L} + \eta_{y_N} + \eta_n$. Mathematically, the ACF of the received signal $y(t)$ given by

$$\begin{aligned}
\eta_y(\mathbf{r}_0, \mathbf{r}_s) &= \mathbb{E}[y(t_0)y^*(t_s)] = \eta_{\mathrm{LL}^*} + \eta_{\mathrm{LN}^*} + \eta_{\mathrm{Ln}^*} \\
&\quad + \eta_{\mathrm{NL}^*} + \eta_{\mathrm{NN}^*} + \eta_{\mathrm{Nn}^*} + \eta_{nL^*} + \eta_{nN^*} + \eta_{nn^*} \\
\eta_{\mathrm{LL}^*} &= \mathbb{E}[y_L(t_0)y_L^*(t_s)],\ \eta_{\mathrm{LN}^*} = \mathbb{E}[y_L(t_0)y_N^*(t_s)] \\
\eta_{\mathrm{Ln}^*} &= \mathbb{E}[y_L(t_0)n^*(t_s)],\ \eta_{\mathrm{NL}^*} = \mathbb{E}[y_N(t_0)y_L^*(t_s)] \\
\eta_{\mathrm{NN}^*} &= \mathbb{E}[y_N(t_0)y_N^*(t_s)],\ \eta_{\mathrm{Nn}^*} = \mathbb{E}[y_N(t_0)n^*(t_s)] \\
\eta_{nL^*} &= \mathbb{E}[n(t_0)y_L^*(t_s)],\ \eta_{nN^*} = \mathbb{E}[n(t_0)y_N^*(t_s)] \\
\eta_{nn^*} &= \mathbb{E}[n(t_0)n^*(t_s)] = \sigma^2\mathbf{I}. \qquad (44)
\end{aligned}$$

Referring to our derivations from (14) to (19), it is easy to obtain

$$\begin{aligned}
\eta_{\mathrm{LN}^*} &= \sum_{i=1}^{M}\sum_{n=1}^{N}\eta_y(\mathrm{LN}^*)_{i,n} = \sum_{i=1}^{M}\sum_{n=1}^{N} \\
&\quad \mathbb{E}_\phi\left\{\exp\left[j(\psi_{i,0} - \omega_d t_s \cos\alpha_n + \phi_i - \phi_n)\right]\right\} = 0 \\
\eta_y(\mathrm{Ln}^*) &= \mathbb{E}[y_{LOS}]\mathbb{E}[n^*(t_s)] = 0 \qquad (45)
\end{aligned}$$

given the fact that ϕ_i and ϕ_n are assumed to be uniform distribution over $(-\pi, \pi]$ and the signal components y_L and y_N are independent with noise. Similarly, we can get $\eta_y(\mathrm{NL}^*) = 0, \eta_y(\mathrm{Nn}^*) = 0, \eta_y(\mathrm{nL}^*) = 0, \eta_y(\mathrm{nN}^*) = 0$. As a consequence, $\eta_y(\mathbf{r}_0, \mathbf{r}_s)$ can be simplified as

$$\begin{aligned}
\eta_y(\mathbf{r}_0, \mathbf{r}_s) &= \eta_y(\mathrm{LL}^*) + \eta_y(\mathrm{NN}^*) + \eta_y(\mathrm{nn}^*) \\
&= \eta_{y_L} + \eta_{y_N} + \eta_n. \qquad (46)
\end{aligned}$$

Thus, the proof is completed.

REFERENCES

[1] H. Wang, L. Wan, M. Dong, K. Ota, and X. Wang, "Assistant vehicle localization based on three collaborative base stations via SBL-based robust DOA estimation," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5766–5777, Jun. 2019.

[2] S. Kuutti, S. Fallah, K. Katsaros, M. Dianati, F. Mccullough, and A. Mouzakitis, "A survey of the state-of-the-art localization techniques and their potentials for autonomous vehicle applications," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 829–846, Apr. 2018.

[3] J. G. McNeff, "The global positioning system," *IEEE Trans. Microw. Theory Techn.*, vol. 50, no. 3, pp. 645–652, Mar. 2002.

[4] J. Wang, J. Luo, S. J. Pan, and A. Sun, "Learning-based outdoor localization exploiting crowd-labeled WiFi hotspots," *IEEE Trans. Mobile Comput.*, vol. 18, no. 4, pp. 896–909, Apr. 2019.

[5] X. Zeng, M. Yang, B. Chen, and Y. Jin, "Estimation of direction of arrival by time reversal for low-angle targets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 6, pp. 2675–2694, Dec. 2018.

[6] M. Guo, Y. D. Zhang, and T. Chen, "DOA estimation using compressed sparse array," *IEEE Trans. Signal Process.*, vol. 66, no. 15, pp. 4133–4146, Apr. 2018.

[7] A. Alessandrini, F. Mazzarella, and M. Vespe, "Estimated time of arrival using historical vessel tracking data," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 1, pp. 7–15, Jan. 2019.

[8] F. Shang, B. Champagne, and I. N. Psaromiligkos, "A ML-based framework for joint TOA/AOA estimation of UWB pulses in dense multipath environments," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5305–5318, Oct. 2014.

[9] G. Wang, W. Zhu, and N. Ansari, "Robust TDOA-based localization for IoT via joint source position and NLOS error estimation," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8529–8541, Oct. 2019.

[10] Y. Qi, H. Kobayashi, and H. Suda, "Analysis of wireless geolocation in a non-line-of-sight environment," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 672–681, Mar. 2006.

[11] N. Patwari, A. O. Hero, M. Perkins, N. S. Correal, and R. J. O'Dea, "Relative location estimation in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 51, no. 8, pp. 2137–2148, Aug. 2003.

[12] H. C. So and L. Lin, "Linear least squares approach for accurate received signal strength based source localization," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 4035–4040, Aug. 2011.

[13] R. W. Ouyang, A. K. Wong, and C. Lea, "Received signal strength-based wireless localization via semidefinite programming: Noncooperative and cooperative schemes," *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1307–1318, Mar. 2010.

[14] X. Tian, S. Zhu, S. Xiong, B. Jiang, Y. Yang, and X. Wang, "Performance analysis of Wi-Fi indoor localization with channel state information," *IEEE Trans. Mobile Comput.*, vol. 18, no. 8, pp. 1870–1884, Aug. 2019.

[15] Q. Song, S. Guo, X. Liu, and Y. Yang, "CSI amplitude fingerprinting-based NB-IoT indoor localization," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1494–1504, Jun. 2018.

[16] C. Hsieh, J. Chen, and B. Nien, "Deep learning-based indoor localization using received signal strength and channel state information," *IEEE Access*, vol. 7, pp. 33256–33267, 2019.

[17] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, "Position and orientation estimation through millimeter-wave MIMO in 5G systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1822–1835, Mar. 2018.

[18] X. Xia, K. Xu, D. Zhang, Y. Xu, and Y. Wang, "Beam-domain full-duplex massive MIMO: Realizing co-time co-frequency uplink and downlink transmission in the cellular system," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 8845–8862, Oct. 2017.

[19] A. Guerra, F. Guidi, and D. Dardari, "Single-anchor localization and orientation performance limits using massive arrays: MIMO vs. beamforming," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5241–5255, Aug. 2018.

[20] L. Hachad, O. Cherrak, H. Ghennioui, F. Mrabti, and M. Zouak, "DOA estimation based on time-frequency music application to massive mimo systems," in *Proc. Int. Conf. Adv. Technol. Signal Image Process. (ATSIP)*, 2017, pp. 1–5.

[21] A. Hu, T. Lv, H. Gao, Z. Zhang, and S. Yang, "An ESPRIT-based approach for 2-D localization of incoherently distributed sources in massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 996–1011, Oct. 2014.

[22] Z. Marzi, D. Ramasamy, and U. Madhow, "Compressive channel estimation and tracking for large arrays in mm-Wave picocells," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 514–527, Apr. 2016.

[23] N. Garcia, H. Wymeersch, E. G. Larsson, A. M. Haimovich, and M. Coulon, "Direct localization for massive MIMO," *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 2475–2487, May 2017.

[24] J. Wang, H. Zhu, L. Dai, N. J. Gomes, and J. Wang, "Low-complexity beam allocation for switched-beam based multiuser massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8236–8248, Dec. 2016.

[25] W. Wu, D. Liu, X. Hou, and M. Liu, "Low-complexity beam training for 5G millimeter-wave massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 361–376, Jan. 2020.

[26] T. Lv, F. Tan, H. Gao, and S. Yang, "A beamspace approach for 2-D localization of incoherently distributed sources in massive MIMO systems," *Signal Process.*, vol. 121, pp. 30–45, Apr. 2016.

[27] F. Guidi, A. Guerra, D. Dardari, A. Clemente, and R. D'Errico, "Joint energy detection and massive array design for localization and mapping," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1359–1371, Mar. 2017.

[28] K. Wu, W. Ni, T. Su, R. P. Liu, and Y. J. Guo, "Robust unambiguous estimation of angle-of-arrival in hybrid array with localized analog subarrays," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2987–3002, May 2018.

[29] X. Liu *et al.*, "Beam-oriented digital predistortion for 5G massive MIMO hybrid beamforming transmitters," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 7, pp. 3419–3432, Jul. 2018.

[30] S. A. Shaikh and A. M. Tonello, "Localization based on angle of arrival in EM lens-focusing massive MIMO," in *Proc. IEEE 6th Int. Conf. Consum. Electron. Berlin (ICCE-Berlin)*, 2016, pp. 124–128.

[31] X. Sun, X. Gao, G. Y. Li, and W. Han, "Single-site localization based on a new type of fingerprint for massive MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6134–6145, Jul. 2018.

[32] X. Sun, C. Wu, X. Gao, and G. Y. Li, "Fingerprint-based localization for massive MIMO-OFDM system with deep convolutional neural networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 10846–10857, Nov. 2019.

[33] J. Vieira, E. Leitinger, M. Sarajlic, X. Li, and F. Tufvesson, "Deep convolutional neural networks for massive MIMO fingerprint-based positioning," in *Proc. IEEE 28th Annu. Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–6.

[34] V. Savic and E. G. Larsson, "Fingerprinting-based positioning in distributed massive MIMO systems," in *Proc. IEEE 82nd Veh. Technol. Conf. (VTC-Fall)*, Sep. 2015, pp. 1–5.

[35] A. Chriki, H. Touati, and H. Snoussi, "SVM-based indoor localization in wireless sensor networks," in *Proc. 13th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jun. 2017, pp. 1144–1149.

[36] J. Yan, L. Zhao, J. Tang, Y. Chen, R. Chen, and L. Chen, "Hybrid kernel based machine learning using received signal strength measurements for indoor localization," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2824–2829, Mar. 2018.

[37] C. Chen, Y. Chen, Y. Han, H. Lai, F. Zhang, and K. J. R. Liu, "Achieving centimeter-accuracy indoor localization on WiFi platforms: A multi-antenna approach," *IEEE Internet Things J.*, vol. 4, no. 1, pp. 122–134, Feb. 2017.

[38] K. J. R. Liu and B. Wang, *Wireless AI: Wireless Sensing, Positioning, IoT, and Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2019.

[39] B. Wang, Q. Xu, C. Chen, F. Zhang, and K. J. R. Liu, "The promise of radio analytics: A future paradigm of wireless positioning, tracking, and sensing," *IEEE Signal Process. Mag.*, vol. 35, no. 3, pp. 59–80, May 2018.

[40] C. Wu, F. Zhang, B. Wang, and K. J. R. Liu, "EasiTrack: Decimeter-level indoor tracking with graph-based particle filtering," *IEEE Internet Things J.*, vol. 7, no. 3, pp. 2397–2411, 2020.

[41] Z. Wu, Y. Han, Y. Chen, and K. J. R. Liu, "A time-reversal paradigm for indoor positioning system," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1331–1339, Apr. 2015.

[42] F. Zhang, C. Chen, B. Wang, H. Lai, Y. Han, and K. J. R. Liu, "WiBall: A time-reversal focusing ball method for decimeter-accuracy indoor tracking," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 4031–4041, Oct. 2018.

[43] X. Ge, S. Tu, G. Mao, C. Wang, and T. Han, "5G ultra-dense cellular networks," *IEEE Wireless Commun.*, vol. 23, no. 1, pp. 72–79, Feb. 2016.

[44] Federal Communications Commission (FCC). *City Commission Agenda MEMO/5G Small Cell Infrastructure in Manhattan*. Accessed: Mar. 12, 2019. [Online]. Available: https://cityofmhk.com/DocumentCenter/View/53891/Item-7B-Local-Standards-for-Small-Cell-Infrastructure

[45] Y. Azar *et al.*, "28 GHz propagation measurements for outdoor cellular communications using steerable beam antennas in New York City," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2013, pp. 5143–5147.

[46] A. I. Sulyman, A. T. Nassar, M. K. Samimi, G. R. Maccartney, T. S. Rappaport, and A. Alsanie, "Radio propagation path loss models for 5G cellular networks in the 28 GHZ and 38 GHZ millimeter-wave bands," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 78–86, Sep. 2014.

[47] W. Croswell, *Antenna Theory, Analysis, and Design*. San Francisco, CA, USA: Harper and Row, 1982.

[48] A. Taira *et al.*, "Performance evaluation of 44GHz band massive mimo based on channel measurement," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, 2015, pp. 1–6.

[49] O. Martinez, J. A. Nielsen, E. De Carvalho, and P. Popovski, "An experimental study of massive MIMO properties in 5G scenarios," *IEEE Trans. Antennas Propag.*, vol. 66, no. 12, pp. 7206–7215, Dec. 2018.

[50] J. Chen, X. Yin, X. Cai, and S. Wang, "Measurement-based massive MIMO channel modeling for outdoor LOS and NLOS environments," *IEEE Access*, vol. 5, pp. 2126–2140, 2017.

[51] C. Xiao, Y. R. Zheng, and N. C. Beaulieu, "Novel sum-of-sinusoids simulation models for Rayleigh and Rician fading channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3667–3679, Dec. 2006.

[52] C. Chen, Y. Chen, Y. Han, H. Lai, and K. J. R. Liu, "Achieving centimeter-accuracy indoor localization on WiFi platforms: A frequency hopping approach," *IEEE Internet Things J.*, vol. 4, no. 1, pp. 111–121, Feb. 2017.

[53] F. Zhang, C. Chen, B. Wang, and K. J. R. Liu, "Wispeed: A statistical electromagnetic approach for device-free indoor speed estimation," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 2163–2177, Jun. 2018.

[54] C. William, "Robust locally weighted regression and smoothing scatterplots," *J. Amer. Stat. Assoc.*, vol. 74, no. 368, pp. 829–836, 1979.

[55] Manhattan Local News. *Topographic Prominence*. Accessed: Aug. 25, 2020. [Online]. Available: https://en.wikipedia.org/wiki/Topographic_prominence

[56] G. A. Brosamler, "An almost everywhere central limit theorem," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 104. Cambridge, U.K.: Cambridge Univ. Press, 1988, pp. 561–574.

[57] J. Yu *et al.*, "SenSpeed: Sensing driving conditions to estimate vehicle speed in urban environments," *IEEE Trans. Mobile Comput.*, vol. 15, no. 1, pp. 202–216, Jan. 2016.

[58] J. Wang, J. Tong, Q. Gao, Z. Wu, S. Bi, and H. Wang, "Device-free vehicle speed estimation with WiFi," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8205–8214, 2018.

[59] A. Gorski, "Understanding GPS performance in urban environments," in *Geospatial Analysis*, Jan. 2011. [Online]. Available: https://www.agi.com/missions/analysis/understanding-gps-performance-in-urban-environment

[60] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.

Xiaolu Zeng received the B.S. degree from Harbin Institute of Technology, Harbin, China, in 2014, and the Ph.D. degree from the School of Electronic Engineering, Xidian University, Xi'an, China, in 2020.

From September 2017 to June 2020, he is a visiting student with the Department of Electrical and Computer Engineering, University of Maryland at College Park, College Park, MD, USA, where he is currently a Postdoctoral Research Associate with the Department of Electrical and Computer Engineering. His research interests include target detection, tracking and localization, wireless sensing, 5G millimeter wave communication and applications, intelligent recognition and classification, and advanced driver assistance systems.

Feng Zhang (Member, IEEE) received the B.S. and M.S. degrees from the Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei, China, in 2011 and 2014, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Maryland at College Park, College Park, MD, USA, in December 2018.

He was with Origin Wireless, Inc., Greenbelt, MD, USA, as a Principal Scientist from 2018 to 2020. He is currently with Amazon.com, Inc., Seattle, WA, USA. His research interests include wireless sensing, statistical signal processing, and wireless indoor localization.

Dr. Zhang was a recipient of the 2020 IEEE INTERNET OF THINGS JOURNAL Best Paper Award.

Beibei Wang (Senior Member, IEEE) received the B.S. degree (Hons.) in electrical engineering from the University of Science and Technology of China, Hefei, China, in 2004, and the Ph.D. degree in electrical engineering from the University of Maryland at College Park, College Park, MD, USA, in 2009.

She was with the University of Maryland at College Park as a Research Associate from 2009 to 2010 and Qualcomm Research and Development, Shanghai, China, from 2010 to 2014. Since 2015, she has been with Origin Wireless Inc., Greenbelt, MD, USA, where she is currently the Vice President of Research. She is also affiliated with the University of Maryland at College Park. She has coauthored *Cognitive Radio Networking and Security: A Game-Theoretic View* (Cambridge University Press, 2010) and *Wireless AI: Wireless Sensing, Positioning, IoT, and Communications* (Cambridge University Press, 2019). Her research interests include Internet of Things, mobile computing, wireless sensing and positioning, and communications and networking.

Dr. Wang is a recipient of the 2020 IEEE INTERNET OF THINGS JOURNAL Best Paper Award, the 2015 IEEE Signal Processing Society Overview Paper Award, and several research and invention awards from the University of Maryland. She has served on the editorial board of IEEE SIGNAL PROCESSING LETTERS, IEEE INTERNET OF THINGS JOURNAL, and IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS.

K. J. Ray Liu (Fellow, IEEE) is a Distinguished University Professor and a Distinguished Scholar-Teach with the University of Maryland at College Park, College Park, MD, USA, where he is also a Christine Kim Eminent Professor of Information Technology. He leads the Maryland Signals and Information Group conducting research encompassing broad areas of information and communications technology with recent focus on wireless AI for indoor tracking and wireless sensing.

Dr. Liu was the recipient of two IEEE Technical Field Awards: the 2021 IEEE Fourier Award for Signal Processing and the 2016 IEEE Leon K. Kirchmayer Graduate Teaching Award, the IEEE Signal Processing Society 2009 Technical Achievement Award, the IEEE Signal Processing Society 2014 Society Award, and over a dozen of best paper/invention awards. Recognized by Web of Science as a Highly Cited Researcher. He also received teaching and research recognitions from University of Maryland, including university-level Invention of the Year Award, and college-level Poole and Kent Senior Faculty Teaching Award, Outstanding Faculty Research Award, and Outstanding Faculty Service Award, all from A. James Clark School of Engineering. As the founder of Origin Wireless, his inventions won the 2017 CEATEC Grand Prix and three CES Innovation Awards, including CES Best of Innovation in 2021. He is a Fellow of AAAS and U.S. National Academy of Inventors. He is the IEEE President-Elect in 2021. He was an IEEE Vice President, Technical Activities, and a member of IEEE Board of Director as Division IX Director. He has also served as a President of IEEE Signal Processing Society, where he was a Vice President—Publications and the Editor-in-Chief of *IEEE Signal Processing Magazine*.