

Intelligent Beam Tracking in Radar-assisted MIMO-OFDM Communication Systems

Yijie Wang, Mengting Lou, Wanyun Qian, Yechao Bai, *Member, IEEE*, Lan Tang, *Member, IEEE*, and Ying-Chang Liang, *Fellow, IEEE*

Abstract—When the base station (BS) communicates with terminals via narrow beams, the BS should have beam tracking capability due to the uncertainty of the terminals' locations especially in the high mobility scenario. Unlike conventional beam tracking based on pilots, we investigate beam tracking in the framework of integrated communication and sensing. In the proposed architecture, the transmitting BS and the receiving BSs cooperate to complete target location and beam tracking. When the transmitting BS transmits to terminals, multiple receiving BSs apply extended kalman filter (EKF) or unscented kalman filter (UKF) to predict locations of terminals based on received echoes, and then the location information is fed back to the transmitting BS to assist in beam tracking. To improve the performance of beam tracking, we first analyze the posterior Cramer-Rao lower bound (PCRLB) of prediction performance under the first-order Taylor approximation. Then, we investigate the subcarrier and power allocation problem during tracking, which is essentially a sequential decision problem. We propose a novel algorithm that combines optimization and reinforcement learning (RL) to achieve optimal overall performance. Simulation results demonstrate the tracking performance and transmission rate of the proposed method, as well as its advantages compared to existing schemes.

Index Terms—beam tracking, EKF, PCRLB, resource allocation, reinforcement learning

I. INTRODUCTION

WITH the explosive growth of wireless communication traffic, there are urgent needs for higher transmission rate and greater bandwidth. Large-scale antenna array is used to improve the received signal-to-noise ratio (SNR) and hence the system performance. By applying beamforming on large-scale array, the energy of the signals transmitted by the base station (BS) directs to the communication users via narrow beams. Due to the hardware complexity and high cost of digital beamforming, current commercial large scale antenna arrays typically use analog or hybrid beamforming. In analog beamforming architecture, the beamforming will be carried out in the analog domain to decrease the number of radio frequency

(RF) chains [1]. However, with analog beamforming, beamformed transmission can only be done in one direction at a time. Hybrid beamforming serves as a compromise to fully digital beamforming with reduced flexibility and performance. Recently, the development of digital technology has promoted the development of mmWave digital array. NEC has developed its own 24-channel mmWave all-digital array [2] and a 64-channel all-digital multi-beam array for 5G mmWave applications has been realized in [3].

Except for beam alignment, when the users move, the BS should have the ability to quickly adjust the beam direction according to the changes of the users' locations, i.e., the ability of beam tracking. In [4], a new unmanned aerial vehicle (UAV) assisted millimeter wave (mmWave) beam training strategy was proposed, which constructed a training beam with flat-top characteristics. Beam training based on beam sweep is feasible in some static or slowly changing environment, but in mobile scenarios, a full scan in all directions is time-consuming and will cause heavy signaling overhead. Therefore, some technologies that did not require full beam scanning were proposed. The authors in [5] proposed two partial beam training schemes based on interlacing and selection probability, which significantly reduced the training cost at the expense of spectral efficiency. Aiming at the problem that high mobility users will consume a amounts of beam training resources, a special beam training strategy and a greedy channel estimation algorithm were proposed, which can greatly reduce the beam training cost while minimizing the channel estimation error [6]. In addition, compared with exhaustive search, introducing hierarchical codebook design into beam scanning is also an effective method to reduce the scanning cost [7], [8]. Some other methods to reduce the scanning frequency have been proposed. For instance, [9] proposed to extract the information needed to update the direction directly from the communication packets without spatial scanning during the ongoing data transmission. In [10], a candidate beam set that may contain the strongest beam pair from the perspective of probability was maintained, and then the scanning beam was selected from the candidate beam set, avoiding full beam scanning.

In the high frequency band, such as mmWave band, some features such as channel sparsity can be used to develop new technologies to implement beam tracking. In [11], the sparse scattering property of the mmWave channel was exploited to formulate the beam alignment problem as a sparse coding and phase-free decoding problem, and then a sparse bipartite graph coding algorithm was proposed to solve this problem. [12] used the sparse characteristics of the mmWave channel

Copyright (c) 2024 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

This work was supported by the National Key R&D Program of China (2023YFB2905601) and the National Natural Science Foundation of China (62072229, 62071220). (Corresponding author: Lan Tang)

Yijie Wang, Wanyun Qian, Yechao Bai and Lan Tang are with the School of Electronic Science and Engineering, Nanjing University, Nanjing 210093, P. R. China. (e-mail: tanglan@nju.edu.cn). Corresponding author: Lan Tang. Mengting Lou is with China Mobile Research Institute, Beijing 100053, China.

Ying-Chang Liang is with the Center for Intelligent Networking and Communications, University of Electronic Science and Technology of China, Chengdu, 611731, China (e-mail: liangyc@ieee.org).

in the angle-of-arrival/angle-of-departure (AoA-AoD) domain to model the beam alignment problem as a sparse vector estimation problem, and proposed a non-negative least squares algorithm to effectively find the strongest coupling AoA-AoD pair connecting user to the BS.

Most existing beam tracking schemes are implemented in the framework of communication protocol [13], [14]. During beam tracking, the signaling interaction between the BS and communication users will cause some delay, which will affect the accuracy of beam alignment in the high mobility scenario. Especially in the case of narrow beams, loss of connectivity due to a rapid degradation of established beam pairs, can be expected to occur more frequently. To reduce the influence of outdated channel information on beam tracking, in the recent research, the BS used Kalman filter (KF) to predict the location of the terminal for beam tracking [15]–[17]. However, in general, all of these aforementioned algorithms for beam tracking result in non-negligible communication overhead and delay. With the development of integrated communication-radar systems (ICRS), radar-assisted communication provides a new way for fast beam tracking. [18] proposed for the first time to predict users' positions and realize fast beam tracking using echo signals received by communication BS. To improve the performance of radar-assisted beam tracking, the best tradeoff between communication rate and tracking performance was obtained through power allocation.

In this paper, our goal is to maximize sum-rate of mobile terminals in the large-scale antenna array based mmWave OFDM communication systems by optimizing transmit parameters, including beamforming vectors, subcarrier and power allocation, according to terminal parameter information extracted from echoes. Inspired by the work in [18], in this paper, we realize fast multi-terminal beam tracking by using OFDM signals as dual-function radar-communication (DFRC) signals. Compared with beam tracking based on pilots, the main advantages of the beam tracking based on echo signals are low delay and zero additional overhead. To improve the prediction accuracy, we design the cooperative detection scheme, in which multiple receiving BSs convey echo data to a central processor (CP), CP uses EKF or UKF to predict the position and speed information of the terminals, and then terminals' position information is fed back to the transmitting BS for optimizing transmit parameters. To analyze the prediction accuracy, we derive the closed-form expression of posterior Cramer-Rao lower bound (PCRLB) when locations and velocities of terminals are time-varying [19]. During the tracking process, the subcarrier and power allocation at time t will affect the transmission rate at the current time and the prediction accuracy at the next time. Therefore, the resource allocation during tracking is essentially a sequential decision problem, which is solved by combining optimization and reinforcement learning to balance the present and future rewards. On the other hand, when using different spatial beams to track users in the same frequency band, as studied in [20], the receiving BS cannot associate the received beams with users. Existing accurate user association algorithms tend to have high complexity. In our scheme, each subcarrier is only allocated one communication user. When the receiving BSs know the subcarrier allocation, they can

naturally associate received beams in different subcarriers with users, without complex multi-terminal association algorithm. Our contributions can be summarized below:

- We propose a fast beam tracking scheme based on echo signals in a large-scale antenna array based mmWave OFDM system. CP applies EKF or UKF algorithm to extract user position information from the echo data from multiple receiving BSs and reports the position information to the transmitting BS to assist in the beam design and resource allocation.
- We analyze the prediction performance of the echo-based prediction scheme by calculating PCRLB. The subcarrier and power allocation at time t is optimized to capture the tradeoff between the transmission rate at the current time and prediction accuracy at the next time. We propose a novel learning frame to solve this sequential decision problem by combing convex optimization and reinforcement learning.
- The simulation results show the beam tracking and communication performance of the sensing assisted OFDM system with different system parameters. The comparison results with existing schemes demonstrate the advantages of the proposed scheme in balancing communication and tracking performance.

The organization of this paper is listed below. Section II introduces the system model, Section III presents the EKF and UKF algorithms, Section IV analyzes the PCRLB of estimation performance, Section V proposes the sequential resource allocation scheme which achieves the optimal trade-off between the communication and prediction performance, Section VI shows the simulation results, and Section VII concludes the paper.

II. SYSTEM MODEL

We consider an integrated mmWave communication-radar system, which consists of a transmitting BS with M_t antennas and R distributed receiving BSs with M_r antennas. Assume that there are K moving air communication terminals with single antenna, such as unmanned aerial vehicle (UAV). As shown in Fig. 1, in the considered ICRS, locations of terminals can be accurately predicted by central processor connected to R receiving BSs with received echoes, and then are used to assist information transmission. Specifically, the transmitting

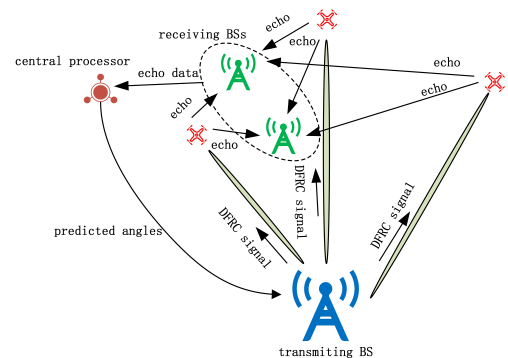


Fig. 1. Integrated communication-radar system for simultaneous information transmission and tracking.

BS sends OFDM signals to terminals via narrow beams, and the distributed receiving BSs receive echoes reflected by terminals. Then the receiving BSs send echo signals back to the central processor to predict the positions of terminals. Finally, the central processor reports the position information to the transmitting BS, which adjusts the beam direction to align with the moving terminals.

A. Communication Signal Model

We assume that transmitted baseband signals include Q OFDM symbols and each OFDM symbol has L subcarriers. Considering the flexibility of digital beamforming in adjusting the beam direction of different subcarriers, we adopt digital beamforming in each subcarrier. The baseband signals in the q th OFDM symbol can be denoted as

$$\mathbf{x}^q(t) = \sum_{k=1}^K \sum_{l=1}^L \rho_{k,l} \mathbf{w}_{k,l} b_{k,l}^q e^{j2\pi(l-1)\Delta f(t-(q-1)T_o)} \quad (1)$$

$$\xi(t - (q-1)T_o), (q-1)T_o \leq t \leq qT_o,$$

where $\mathbf{w}_{k,l}$ denotes the beamforming vector on the l th subcarrier to the terminal k , $\rho_{k,l}$ denotes the subcarrier allocation factor ($\rho_{k,l} = 1$ means the l th subcarrier is allocated to the k th terminal, otherwise, it is not allocated to the k th terminal), $b_{k,l}^q$ is the transmitted signal to terminal k on the l th subcarrier in the q th OFDM symbol, $\Delta f = 1/T$, T is the symbol period of OFDM, $T_o = T + T_{cp}$, T_{cp} is the length of cyclic prefix, and

$$\xi(t) = \begin{cases} 1, & t \in [0, T_o], \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

To avoid interference between terminals, each subcarrier is only allocated one communication terminal, i.e., $\sum_{k=1}^K \rho_{k,l} = 1, \forall l$. The transmitted signals in an OFDM block can be expressed by $\mathbf{x}(t) = \sum_{q=1}^Q \mathbf{x}^q(t), 0 \leq t \leq QT_o$. Then, the baseband signal $\mathbf{x}(t)$ needs to be upconverted to obtain the signal $\tilde{\mathbf{x}}(t) = Re(\mathbf{x}(t)e^{j2\pi f_c t})$ for transmission, where f_c denotes the carrier frequency.

We assume that the state of the k th terminal $\mathbf{s}_k[t]$ can be represented as $\mathbf{s}_k[t] = [d_{k,x}[t], d_{k,y}[t], v_{k,x}[t], v_{k,y}[t]]^T$, where $d_{k,x}[t]$ and $d_{k,y}[t]$ denote the horizontal and vertical coordinates of the k th terminal position, respectively, and $v_{k,x}[t]$ and $v_{k,y}[t]$ denote the horizontal and vertical velocity of the k th terminal, respectively. Assuming that the k th terminal is moving at an almost constant speed, the state evolution model of the k th terminal can be expressed as

$$\begin{cases} d_{k,x}[t] = d_{k,x}[t-1] + v_{k,x}[t-1]QT_o \\ d_{k,y}[t] = d_{k,y}[t-1] + v_{k,y}[t-1]QT_o \\ v_{k,x}[t] = \iota v_{k,x}[t-1] + \omega_{k,x}[t] \\ v_{k,y}[t] = \iota v_{k,y}[t-1] + \omega_{k,y}[t] \end{cases}, \quad (3)$$

where ι is the regression coefficient, $\omega_{k,x}[t]$ and $\omega_{k,y}[t]$ denote perturbation variables which obey the Gaussian distribution with zero mean and variance of $\sigma_{\omega_k}^2$. At time t , we use Q OFDM symbols to predict the positions and velocities of K terminals at time $t+1$.

B. Beam Tracking Based On Sensing

Traditional beam tracking schemes in mmWave communications require extra pilot overhead. In our proposed beam

tracking scheme based on terminal location prediction, OFDM data blocks are used to track terminals and no additional pilots are needed. We denote the location and velocity of the k th terminal as $\mathbf{d}_k[t] = [d_{k,x}[t], d_{k,y}[t]]^T$ and $\mathbf{v}_k[t] = [v_{k,x}[t], v_{k,y}[t]]^T$, respectively, the location of the transmitting BS as $\mathbf{p}_t = [p_{t,x}, p_{t,y}]^T$, and the location of the r th receiving BS as $\mathbf{p}_r = [p_{r,x}, p_{r,y}]^T$. We assume that the terminals are UAVs, hence a simple ground-to-air channel model is adopted. The channels both from the transmitting BS to the terminals and from the terminals to the receiving BSs are line of sight (LoS) paths without multi-path components. Assuming that signals on each subcarrier are narrow-band signals, received signals at the r th receiving BS after down-conversion can be expressed by

$$\mathbf{y}_r(t) = \sum_{k=1}^K A_{r,k} e^{j2\pi f_{r,k} t} \mathbf{x}(t - \tau_{r,k}) + \boldsymbol{\eta}_r(t)$$

$$= \sqrt{M_t M_r} \sum_{q=1}^Q \sum_{l=1}^L \sum_{k=1}^K A_{r,k} e^{j2\pi f_{r,k} t} \mathbf{a}_r^l(\theta_{r,k}) (\mathbf{a}_t^l(\beta_k))^T \quad (4)$$

$$e^{j2\pi(l-1)\Delta f(t-(q-1)T_o - \tau_{r,k})} \xi(t - (q-1)T_o - \tau_{r,k})$$

$$\sum_{k'=1}^K \rho_{k',l} \mathbf{w}_{k',l} b_{k',l}^q + \boldsymbol{\eta}_r(t),$$

where

$$f_{r,k} = \frac{\mathbf{v}_k^T[t](\mathbf{p}_t - \mathbf{d}_k[t])}{\lambda \|\mathbf{p}_t - \mathbf{d}_k[t]\|_2} + \frac{\mathbf{v}_k^T[t](\mathbf{p}_r - \mathbf{d}_k[t])}{\lambda \|\mathbf{p}_r - \mathbf{d}_k[t]\|_2}, \quad (5)$$

$$\tau_{r,k} = \frac{1}{c} (\|\mathbf{p}_t - \mathbf{d}_k[t]\|_2 + \|\mathbf{p}_r - \mathbf{d}_k[t]\|_2), \quad (6)$$

$f_{r,k}$, $\tau_{r,k}$ and $A_{r,k}$ denote the Doppler frequency, delay and path loss for the channel between transmitting BS-terminal k -the r th receiving BS, respectively, $\theta_{r,k} = \arccos(\frac{d_{k,x}[t] - p_{r,x}}{\sqrt{(p_{r,x} - d_{k,x}[t])^2 + (p_{r,y} - d_{k,y}[t])^2}})$ is the angle of arrival (AoA) of the k th terminal at the r th receiving BS, $\beta_k = \arccos(\frac{d_{k,x}[t] - p_{t,x}}{\sqrt{(p_{t,x} - d_{k,x}[t])^2 + (p_{t,y} - d_{k,y}[t])^2}})$ is the angle of departure (AoD) of the k th terminal at the transmitting BS, $\mathbf{a}_t^l(\beta_k) = \frac{1}{\sqrt{M_t}} [1, e^{-j\frac{2\pi}{\lambda_t} d \cos \beta_k}, \dots, e^{-j\frac{2\pi}{\lambda_t} (M_t-1) d \cos \beta_k}]^T$ and $\mathbf{a}_r^l(\theta_{r,k}) = \frac{1}{\sqrt{M_r}} [1, e^{-j\frac{2\pi}{\lambda_r} d \cos \theta_{r,k}}, \dots, e^{-j\frac{2\pi}{\lambda_r} (M_r-1) d \cos \theta_{r,k}}]^T$ denote the normalized transmit steering vector and receive steering vector of the k th terminal on the l th subcarrier at the r th receiving BS respectively, $\lambda_l = \frac{c}{f_c + (l-1)\Delta f}$, c is the speed of light and $\boldsymbol{\eta}_r(t)$ is additive white Gaussian noise (AWGN) with zero mean and covariance matrix $\sigma^2 \mathbf{I}$ ($\mathbf{I}$ denotes the unit matrix).

With digital beamforming, the beamforming vector on each subcarrier can be determined separately. In order to maximize signal-to-noise ratio (SNR) at communication terminals, we set the beamforming vector $\mathbf{w}_{k',l} = (\mathbf{a}_t^l(\hat{\beta}_{k'}))^* \in \mathbb{C}^{M_t \times 1}$, where $\hat{\beta}_{k'}$ is the predicted AoD of terminal k at time t . The $\hat{\beta}_{k'}$ is obtained based on the tracking algorithm in Sec.III with measure information before time t . By using large-scale antenna array at the transmitting BS, steering vectors in different directions are approximately orthogonal, i.e., $|\mathbf{a}_t^l(\beta_k)^H \mathbf{a}_t^l(\beta_{k'})| \approx 0$ for $\beta_k \neq \beta_{k'}$ [18]. Therefore, when the prediction of AoD is sufficiently accurate, we can

isolate respective signals of different terminals at the receiving BSs. In this case, (4) can be simplified as

$$\begin{aligned} \mathbf{y}_r(t) &= \sqrt{M_t M_r} \sum_{q=1}^Q \sum_{l=1}^L \sum_{k=1}^K A_{r,k} e^{j2\pi f_{r,k} t} \mathbf{a}_r^l(\theta_{r,k}) \\ & (\mathbf{a}_t^l(\beta_k))^T (\mathbf{a}_t^l(\hat{\beta}_k))^* e^{j2\pi(l-1)\Delta f(t-(q-1)T_o - \tau_{r,k})} \rho_{k,l} b_{k,l}^q \\ & \xi(t - (q-1)T_o - \tau_{r,k}) + \boldsymbol{\eta}_r(t), \end{aligned} \quad (7)$$

When utilizing OFDM signals for detection, in order to eliminate the inter-block interference of OFDM signals during matched filtering, it is usually assumed that $\tau_{r,k} < T_{cp}$ [21]–[23]. However, in mmWave communication systems, the delay $\tau_{r,k}$ may be greater than T_{cp} . For example, in 5G NR standard, when $T_{cp} = 1.2\mu s$ and the detection distance is greater than 180m, the time delay will be larger than T_{cp} . To remove the interference between adjacent OFDM symbols during matched filtering in the case of $\tau_{r,k} > T_{cp}$, we need to know the value range of $\tau_{r,k}$, which can be determined by the predicted value of $\tau_{r,k}$ and corresponding prediction error (with the tracking algorithm in Sec. III, transmitting BS can predict $\tau_{r,k}$ with measurements before time t). Then, we can find a suitable time compensation $\hat{m}_{r,k}T_{cp}$ to make $0 < \tau_{r,k} - \hat{m}_{r,k}T_{cp} < T_{cp}$. As indicated in (7), when beams directed to different terminals are orthogonal, there is no inter-terminal interference on each subcarrier. In this case, if $\rho_{k,l} = 1$, at each receiver r , in the q th OFDM symbol, matched filtering on the l th subcarrier is performed to obtain

$$\begin{aligned} \mathbf{y}_r^{q,l} &= \frac{1}{T} \int_{(q-1)T_o}^{(q-1)T_o+T} \mathbf{y}_r(t + \hat{m}_{r,k}T_{cp}) e^{-j2\pi(l-1)\Delta f t} dt \\ &\approx \sqrt{M_t M_r} \sum_{k=1}^K \rho_{k,l} A_{r,k} b_{k,l}^q e^{j2\pi(l-1)\Delta f(\hat{m}_{r,k}T_{cp} - \tau_{r,k})} \\ & e^{j2\pi f_{r,k}(\hat{m}_{r,k}T_{cp} + (q-1)T_o + T/2)} \mathbf{a}_r^l(\theta_{r,k}) (\mathbf{a}_t^l(\beta_k))^T (\mathbf{a}_t^l(\hat{\beta}_k))^* \\ & + \boldsymbol{\eta}_r^{q,l} \\ & = \mathbf{h}_r^{q,l} + \boldsymbol{\eta}_r^{q,l}, \end{aligned} \quad (8)$$

where $\boldsymbol{\eta}_r^{q,l}$ is the AWGN with zero mean and covariance matrix $\frac{\sigma_c^2}{L} \mathbf{I}$. The approximation in (8) is because the phase rotation within one OFDM block due to the Doppler frequency can be approximated as a constant when $f_{r,k}T \ll 1$ [13].

When beams directed to different terminals are not orthogonal, we need to further consider the relative delays between any two terminals. If the maximum delay difference $\max_k \tau_{r,k} - \min_k \tau_{r,k}$, which can be estimated according to predicted values of $\{\tau_{r,k}\}_{k=1}^K$, is smaller than T_{cp} , we can always find a suitable time compensation $\hat{m}_r T_{cp}$ which makes $\min_k \tau_{r,k} - \hat{m}_r T_{cp} > 0$ and $\max_k \tau_{r,k} - \hat{m}_r T_{cp} < T_{cp}$ to avoid inter-block interference. Otherwise, we can group terminals based on the maximum relative delays between terminals, and terminals with the maximum relative delay less than T_{cp} are grouped together. Considering that the relative delays between terminals in different groups are typically much larger than the range resolution $\frac{1}{L\Delta f}$, Extensive Cancellation Algorithm (ECA) is used to eliminate mutual interference between terminals who are not in the same group [24]. After ECA, received signals in (4) can be approximately converted into $\mathbf{y}_r^i(t) = \sum_{k=1}^{|\mathcal{A}_i|} A_{r,k} e^{j2\pi f_{r,k} t} \mathbf{x}(t - \tau_{r,k}) + \boldsymbol{\eta}_r(t)$, $i = 1, \dots, I$,

where I is the number of group, $\mathcal{A}_i$ is the i th group of terminals with $\max_{k \in \mathcal{A}_i} \tau_{r,k} - \min_{k \in \mathcal{A}_i} \tau_{r,k} < T_{cp}$ and $|\mathcal{A}_i|$ is the size of $\mathcal{A}_i$. For each $\mathbf{y}_r^i(t)$, matched filtering on the l th subcarrier is performed with timing compensation $\hat{m}_{r,i}$ to obtain

$$\begin{aligned} \mathbf{y}_{r,i}^{q,l} &= \frac{1}{T} \int_{(q-1)T_o}^{(q-1)T_o+T} \mathbf{y}_r^i(t + \hat{m}_{r,i}T_{cp}) e^{-j2\pi(l-1)\Delta f t} dt \\ &\approx \sqrt{M_t M_r} \sum_{k=1}^{|\mathcal{A}_i|} \sum_{k'=1}^{|\mathcal{A}_i|} \rho_{k',l} A_{r,k} e^{j2\pi(l-1)\Delta f(\hat{m}_{r,i}T_{cp} - \tau_{r,k})} b_{k',l}^q \\ & \mathbf{a}_r^l(\theta_{r,k}) (\mathbf{a}_t^l(\beta_k))^T (\mathbf{a}_t^l(\hat{\beta}_{k'}))^* e^{j2\pi f_{r,k}(\hat{m}_{r,i}T_{cp} + (q-1)T_o + T/2)} \\ & + \boldsymbol{\eta}_{r,i}^{q,l} \\ & = \mathbf{h}_{r,i}^{q,l} + \boldsymbol{\eta}_{r,i}^{q,l}, \end{aligned} \quad (9)$$

where $\hat{m}_{r,i}$ is selected to ensure $\min_{k \in \mathcal{A}_i} \tau_{r,k} - \hat{m}_{r,i}T_{cp} > 0$ and $\max_{k \in \mathcal{A}_i} \tau_{r,k} - \hat{m}_{r,i}T_{cp} < T_{cp}$. For terminals in $\mathcal{A}_i$, we can get a RM_rQL -dimensional complex vector $\dot{\mathbf{y}}_i[t] = [(\mathbf{y}_{1,i}^{1,1})^T, \dots, (\mathbf{y}_{R,i}^{Q,L})^T]^T = \dot{\mathbf{h}}_i(\mathbf{s}_{\mathcal{A}_i}[t]) + \dot{\boldsymbol{\eta}}_i[t]$, where $\dot{\mathbf{h}}_i(\mathbf{s}_{\mathcal{A}_i}[t]) = [(\mathbf{h}_{1,i}^{1,1})^T, \dots, (\mathbf{h}_{R,i}^{Q,L})^T]^T$ and $\dot{\boldsymbol{\eta}}_i[t] = [(\boldsymbol{\eta}_{1,i}^{1,1})^T, \dots, (\boldsymbol{\eta}_{R,i}^{Q,L})^T]^T$. Define $\dot{\mathbf{y}}[t] = [\dot{\mathbf{y}}_1[t], \dots, \dot{\mathbf{y}}_I[t]]^T = \dot{\mathbf{h}}(\mathbf{s}[t]) + \dot{\boldsymbol{\eta}}[t]$, where $\dot{\mathbf{h}}(\mathbf{s}[t]) = [\dot{\mathbf{h}}_1(\mathbf{s}_{\mathcal{A}_1}[t]), \dots, \dot{\mathbf{h}}_I(\mathbf{s}_{\mathcal{A}_I}[t])]$ and $\dot{\boldsymbol{\eta}}[t] = [\dot{\boldsymbol{\eta}}_1[t], \dots, \dot{\boldsymbol{\eta}}_I[t]]^T$. For the convenience of later analysis, complex vectors $\dot{\mathbf{y}}[t]$, $\dot{\mathbf{h}}(\mathbf{s}[t])$ and $\dot{\boldsymbol{\eta}}[t]$, are arranged into a $2RM_rQLI$ -dimensional real vector $\mathbf{y}[t] = [\mathcal{R}(\dot{\mathbf{y}}[t]), \mathcal{I}(\dot{\mathbf{y}}[t])]^T = \mathbf{h}(\mathbf{s}[t]) + \boldsymbol{\eta}[t]$, where $\mathcal{R}(x)$ and $\mathcal{I}(x)$ denote the real and imaginary parts of x respectively, $\mathbf{h}(\mathbf{s}[t]) = [\mathcal{R}(\dot{\mathbf{h}}(\mathbf{s}[t])), \mathcal{I}(\dot{\mathbf{h}}(\mathbf{s}[t]))]^T$ and $\boldsymbol{\eta}[t] = [\mathcal{R}(\dot{\boldsymbol{\eta}}[t]), \mathcal{I}(\dot{\boldsymbol{\eta}}[t])]^T$.

The weighted sum rate of K terminals can be given by

$$\bar{R} = \frac{1}{L} \sum_{k=1}^K \sum_{l=1}^L \rho_{k,l} w_k \log \left(1 + \frac{|\mathbf{w}_{k,l}^H \mathbf{g}_{k,l}|^2 p_{k,l}}{\sigma_C^2} \right), \quad (10)$$

where $\mathbf{g}_{k,l} = \sqrt{M_t} A_k e^{j\varphi_{k,l}} \mathbf{a}_t^l(\beta_k)$ is the channel vector between transmitting BS and the k th terminal, A_k denotes the path loss from the transmitting BS to terminal k , $\varphi_{k,l}$ is the phase of the l th subcarrier channel of terminal k , σ_C^2 is the noise power at the communication receiver and w_k is the weight of terminal k for fairness.

III. BEAM TRACKING BASED ON KALMAN FILTER

In this section, we propose to use the EKF and UKF to track the communication terminals. In order to track terminal $k \in \mathcal{A}_i$, we extract the measure information corresponding to terminal k from $\{\mathbf{y}_{r,i}^{q,l}\}$ to form a real vector $\mathbf{y}_k[t] \in \mathbb{R}^{2RM_rQL_k \times 1}$ for terminal k at time t , where L_k denotes the number of subcarriers allocated to terminal k . Based on (3) and (9), the state evolution equation and measure equation can be expressed as

$$\begin{cases} \mathbf{s}_k[t] = \mathbf{F} \mathbf{s}_k[t-1] + \boldsymbol{\omega}_k[t], \\ \mathbf{y}_k[t] = \mathbf{h}_k(\mathbf{s}_k[t]) + \sum_{k' \neq k} \mathbf{h}_{k'}(\mathbf{s}_{k'}[t]) + \boldsymbol{\eta}_k[t], \end{cases} \quad (11)$$

where $\mathbf{F}$ can be expressed by

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & QT_o & 0 \\ 0 & 1 & 0 & QT_o \\ 0 & 0 & \iota & 0 \\ 0 & 0 & 0 & \iota \end{bmatrix}, \quad (12)$$

$\omega_k[t]$ is the process noise with zero-mean and covariance matrix $\mathbf{G}_{\omega_k} = \text{diag}(0, 0, \sigma_{\omega_k}^2, \sigma_{\omega_k}^2)$ and $\eta_k[t]$ is the measurement noise with zero-mean and covariance matrix $\mathbf{G}_{\eta_k} = \frac{\sigma^2}{2L} \mathbf{I}$. Note that when the number of array antennas is large, the multi-terminal interference during tracking is negligible, i.e., $\mathbf{h}_{k'}(\mathbf{s}_{k'}[t]) \approx \mathbf{0}$, as shown in (8). Otherwise, the echo signals from other terminals $\{\mathbf{h}_{k'}(\mathbf{s}_{k'}[t]), k' \neq k, k' \in \mathcal{A}_i\}$ will degrade the tracking performance. The measurement equation in (11) is nonlinear, so we can use EKF or UKF for tracking. The MSE of EKF has a recursive relationship, which can facilitate our performance optimization later. However, when the measurement equation is highly nonlinear, using UKF can obtain better tracking performance than EKF.

A. Tracking based on EKF

With (12), the EKF is implemented to predict the state of the k th terminal at time t according to the following steps:

1) Terminal State Prediction:

$$\hat{\mathbf{s}}_k[t|t-1] = \mathbf{F} \hat{\mathbf{s}}_k[t-1|t-1], \quad (13)$$

where $\hat{\mathbf{s}}_k[t-1|t-1]$ denotes the estimated state of terminal k at time $t-1$ and $\hat{\mathbf{s}}_k[t|t-1]$ denotes the predicted state of terminal k at time t .

2) Mean Square Error (MSE) Matrix Prediction:

$$\mathbf{M}_k[t|t-1] = \mathbf{F} \mathbf{M}_k[t-1|t-1] \mathbf{F}^T + \mathbf{G}_{\omega_k}, \quad (14)$$

where $\mathbf{M}_k[t|t-1] = \mathbb{E}((\mathbf{s}_k[t] - \hat{\mathbf{s}}_k[t|t-1])(\mathbf{s}_k[t] - \hat{\mathbf{s}}_k[t|t-1])^T)$, $\mathbf{M}_k[t-1|t-1] = \mathbb{E}((\mathbf{s}_k[t-1] - \hat{\mathbf{s}}_k[t-1|t-1])(\mathbf{s}_k[t-1] - \hat{\mathbf{s}}_k[t-1|t-1])^T)$.

3) Kalman Gain Calculation:

$$\mathbf{K}_k[t] = \mathbf{M}_k[t|t-1] \mathbf{H}_k^T[t] (\mathbf{G}_{\eta_k} + \mathbf{G}_{\omega_k}^I + \mathbf{H}_k[t] \mathbf{M}_k[t|t-1] \mathbf{H}_k^T[t])^{-1}, \quad (15)$$

where $\mathbf{H}_k[t] = \frac{\partial \mathbf{h}_k(\mathbf{s}_k[t])}{\partial \mathbf{s}_k[t]} \big|_{\mathbf{s}_k[t] = \hat{\mathbf{s}}_k[t|t-1]}$, $\mathbf{G}_{\omega_k}^I$ is a diagonal matrix and its $2rmql$ th diagonal element is $M_t p_{k,l}$
 $|\sum_{k' \neq k} A_{r,k'} e^{-j \frac{2\pi}{\lambda_l} (m-1)d \cos \theta_{r,k'}} e^{j2\pi(l-1)\Delta f(\hat{m}_{r,i} T_{cp} - \tau_{r,k'})} e^{j2\pi f_{r,k'}(\hat{m}_{r,i} T_{cp} + (q-1)T_o + T/2)} (\mathbf{a}_t^l(\hat{\beta}_{k'})^T)^T (\mathbf{a}_t^l(\hat{\beta}_k))^*|^2$, $\mathbf{H}_k[t]$ is obtained by linearizing $\mathbf{h}_k(\mathbf{s}_k[t])$ near $\hat{\mathbf{s}}_k[t|t-1]$.

4) Terminal State Update:

$$\hat{\mathbf{s}}_k[t] = \hat{\mathbf{s}}_k[t|t-1] + \mathbf{K}_k[t] (\mathbf{y}_k[t] - \mathbf{h}_k(\hat{\mathbf{s}}_k[t|t-1])) \quad (16)$$

5) MSE Matrix Update:

$$\mathbf{M}_k[t] = (\mathbf{I} - \mathbf{K}_k[t] \mathbf{H}_k[t]) \mathbf{M}_k[t|t-1]. \quad (17)$$

Substituting Kalman gain in (15) into (17), we can get

$$\begin{aligned} \mathbf{M}_k[t] &= \mathbf{M}_k[t|t-1] - \mathbf{M}_k[t|t-1] \mathbf{H}_k^T[t] (\mathbf{G}_{\eta_k} + \mathbf{G}_{\omega_k}^I \\ &\quad + \mathbf{H}_k[t] \mathbf{M}_k[t|t-1] \mathbf{H}_k^T[t])^{-1} \mathbf{H}_k[t] \mathbf{M}_k[t|t-1] \\ &= (\mathbf{H}_k^T[t] (\mathbf{G}_{\eta_k} + \mathbf{G}_{\omega_k}^I)^{-1} \mathbf{H}_k[t] + \mathbf{M}_k^{-1}[t|t-1])^{-1}, \end{aligned} \quad (18)$$

where the second equation holds due to the matrix inversion equation in (34). Substituting (14) into (18) leads to

$$\begin{aligned} \mathbf{M}_k^{-1}[t] &= \mathbf{H}_k^T[t] (\mathbf{G}_{\eta_k} + \mathbf{G}_{\omega_k}^I)^{-1} \mathbf{H}_k[t] + \\ &\quad (\mathbf{F} \mathbf{M}_k[t-1|t-1] \mathbf{F}^T + \mathbf{G}_{\omega_k})^{-1}, \end{aligned} \quad (19)$$

which shows the recursive relationship between MSEs at different time in the EKF.

B. Tracking based on UKF

Similar to EKF, UKF is also built on the linear Kalman filter, so the state update and MSE matrix update are the same as Kalman filter, i.e.,

1) Terminal State Update:

$$\hat{\mathbf{s}}_k[t] = \hat{\mathbf{s}}_k[t|t-1] + \mathbf{K}_k[t] (\mathbf{y}_k[t] - \hat{\mathbf{y}}_k[t|t-1]), \quad (20)$$

2) MSE Matrix Update:

$$\mathbf{M}_k[t] = \mathbf{M}_k[t|t-1] - \mathbf{K}_k[t] \mathbf{M}_{k,yy}[t|t-1] \mathbf{K}_k^T[t], \quad (21)$$

where $\mathbf{K}_k[t] = \mathbf{M}_{k,sy}[t|t-1] \mathbf{M}_{k,yy}^{-1}[t|t-1]$ is the Kalman gain, $\mathbf{M}_{k,sy}[t|t-1]$ is the cross covariance matrix between the observation vector and the state, and $\mathbf{M}_{k,yy}[t|t-1]$ is the covariance matrix of the innovation process. Due to the nonlinear measure model, the prediction of observation $\hat{\mathbf{y}}_k[t|t-1]$, the terminal state prediction $\hat{\mathbf{s}}_k[t|t-1]$, MSE matrix prediction $\mathbf{M}_k[t|t-1]$ are calculated by numerical calculation based on sigma sampling points [25]. Specifically, $\hat{\mathbf{y}}_k[t|t-1]$, $\hat{\mathbf{s}}_k[t|t-1]$ and $\mathbf{M}_k[t|t-1]$ are calculated by (22), (27) and (28), respectively.

3) System Observation Prediction:

$$\hat{\mathbf{y}}_k[t|t-1] = \sum_{i=0}^8 w_i \mathbf{h}_k(\boldsymbol{\varsigma}_{k,i}[t-1]), \quad (22)$$

where $\boldsymbol{\varsigma}_{k,i}[t-1]$ denotes the i th generated sigma point for observation prediction of terminal k at time $t-1$. $\boldsymbol{\varsigma}_{k,i}[t-1]$ is given by

$$\boldsymbol{\varsigma}_{k,i}[t-1] = \begin{cases} \hat{\mathbf{s}}_k[t|t-1], i = 0 \\ \hat{\mathbf{s}}_k[t|t-1] + \tilde{\boldsymbol{\varsigma}}_{k,i}[t-1], i = 1, 2, 3, 4 \\ \hat{\mathbf{s}}_k[t|t-1] - \tilde{\boldsymbol{\varsigma}}_{k,i}[t-1], i = 5, 6, 7, 8 \end{cases}, \quad (23)$$

where $\tilde{\boldsymbol{\varsigma}}_{k,i}[t-1] = (\sqrt{(4+\kappa)\mathbf{M}_k[t|t-1]})_i^T$, κ is a hyper-parameter, $(\sqrt{\mathbf{M}})_i$ denotes the i th row of $\sqrt{\mathbf{M}}$ and $\sqrt{\mathbf{M}}$ is the Cholesky decomposition of $\mathbf{M}$. For $i = 0$, the weights of sigma points $w_0 = \frac{\kappa}{\kappa+4}$ and for $i = 1, \dots, 8$, $w_i = \frac{1}{2(\kappa+4)}$. With (22) and (27), $\mathbf{M}_{k,yy}[t|t-1]$ and $\mathbf{M}_{k,sy}[t|t-1]$ can be calculated by

$$\begin{aligned} \mathbf{M}_{k,yy}[t|t-1] &= \sum_{i=0}^8 w_i (\mathbf{h}_k(\boldsymbol{\varsigma}_{k,i}[t-1]) - \hat{\mathbf{y}}_k[t|t-1]) (\mathbf{h}_k(\boldsymbol{\varsigma}_{k,i}[t-1]) - \hat{\mathbf{y}}_k[t|t-1])^T + \mathbf{G}_{\eta_k} + \mathbf{G}_{\omega_k}^I, \end{aligned} \quad (24)$$

and

$$\begin{aligned} \mathbf{M}_{k,sy}[t|t-1] &= \sum_{i=0}^8 w_i (\mathbf{F} \boldsymbol{\chi}_{k,i}[t-1] - \hat{\mathbf{s}}_k[t|t-1]) (\mathbf{h}_k(\boldsymbol{\varsigma}_{k,i}[t-1]) - \hat{\mathbf{y}}_k[t|t-1])^T, \end{aligned} \quad (25)$$

respectively, where $\boldsymbol{\chi}_{k,i}[t-1]$ denotes the i th regenerated sigma point for state prediction of terminal k at time $t-1$. $\boldsymbol{\chi}_{k,i}[t-1]$ can be expressed by

$$\boldsymbol{\chi}_{k,i}[t-1] = \begin{cases} \hat{\mathbf{s}}_k[t-1|t-1], i = 0 \\ \hat{\mathbf{s}}_k[t-1|t-1] + \tilde{\boldsymbol{\chi}}_{k,i}[t-1], i = 1, 2, 3, 4 \\ \hat{\mathbf{s}}_k[t-1|t-1] - \tilde{\boldsymbol{\chi}}_{k,i}[t-1], i = 5, 6, 7, 8 \end{cases}, \quad (26)$$

where $\tilde{\boldsymbol{\chi}}_{k,i}[t-1] = (\sqrt{(4+\kappa)\mathbf{M}_k[t-1|t-1]})_i^T$.

4) Due to the linear state evolution equation, the terminal State Prediction is:

$$\hat{\mathbf{s}}_k[t|t-1] = \mathbf{F} \hat{\mathbf{s}}_k[t-1|t-1]. \quad (27)$$

5) MSE Matrix Prediction:

$$\mathbf{M}_k[t|t-1] = \sum_{i=0}^8 \mathbf{w}_i (\mathbf{F}\chi_{k,i}[t-1] - \hat{\mathbf{s}}_k[t|t-1]) (\mathbf{F}\chi_{k,i}[t-1] - \hat{\mathbf{s}}_k[t|t-1])^T + \mathbf{G}_{\omega_k}. \quad (28)$$

IV. ANALYSIS OF PCRLB DURING TRACKING

In this section, we derive PCRLB of predicted parameters in the multi-terminal tracking scenario, which will be used as performance metric of tracking performance in the resource allocation problem in the next section.

In our beam tracking problem, the parameter vectors $\{\mathbf{s}_k[t]\}_{k=1}^K$ in Equation (11) are Gaussian Markov stochastic processes. Therefore, we introduce PCRLB to analyze the parameters estimation performance.

Define $\mathbf{s}[t] = [\mathbf{s}_1^T[t], \dots, \mathbf{s}_K^T[t]]^T$, $\mathbf{S}[0:t] = [\mathbf{s}^T[0], \dots, \mathbf{s}^T[t]]^T$, $\mathbf{Y}[0:t] = [\mathbf{y}^T[0], \dots, \mathbf{y}^T[t]]^T$ and assume that the initial state $\mathbf{s}[0]$ has a known probability density function (PDF) $p(\mathbf{s}[0])$. Then $p(\mathbf{S}[0:t], \mathbf{Y}[0:t])$ can be expressed by

$$p(\mathbf{S}[0:t], \mathbf{Y}[0:t]) = p(\mathbf{s}[0]) \prod_{n=0}^t p(\mathbf{y}[n]|\mathbf{s}[n]) \prod_{m=0}^t p(\mathbf{s}[m]|\mathbf{s}[m-1]). \quad (29)$$

Let $\hat{\mathbf{S}}[0:t]$ represent the estimated state vector based on $\mathbf{Y}[0:t]$. According to the definition of PCRLB, the estimation error of $\mathbf{S}[0:t]$ is lower bounded by PCRLB, i.e.,

$$\mathbb{E}\{[\hat{\mathbf{S}}[0:t] - \mathbf{S}[0:t]][\hat{\mathbf{S}}[0:t] - \mathbf{S}[0:t]]^T\} \succeq \mathbf{C}(\mathbf{S}[0:t]) \triangleq \mathbf{J}^{-1}(\mathbf{S}[0:t]), \quad (30)$$

where $\mathbf{C}(\mathbf{S}[0:t])$ is the PCRLB matrix, and $\mathbf{J}(\mathbf{S}[0:t]) \in \mathbb{R}^{4K(t+1) \times 4K(t+1)}$ is the posterior Fisher information matrix (FIM) of $\mathbf{S}[0:t]$ and can be calculated as

$$[\mathbf{J}(\mathbf{S}[0:t])]_{ij} = -\mathbb{E}\left(\frac{\partial^2 \log p(\mathbf{S}[0:t], \mathbf{Y}[0:t])}{\partial([\mathbf{S}[0:t]]_i \partial([\mathbf{S}[0:t]]_j)}\right), \quad (31)$$

$$i, j = 1, \dots, 4K(t+1),$$

where $[\mathbf{S}[0:t]]_i$ denotes the i th element of $\mathbf{S}[0:t]$. Obviously, the inverse of the $4K \times 4K$ right-lower block of $\mathbf{J}^{-1}(\mathbf{S}[0:t])$ corresponds to the PCRLB $\mathbf{C}[t]$ of $\mathbf{s}[t]$. According to the inverse formula of the block matrix, for FIM $\mathbf{J}[t] = \mathbf{C}^{-1}[t]$, there exists the following recursive relation [26],

$$\begin{aligned} \mathbf{J}[t] &= \mathbf{G}_{\omega}^{-1} + \mathbb{E}\left(\left(\frac{\partial \mathbf{h}^T(\mathbf{s}[t])}{\partial \mathbf{s}[t]}\right) \mathbf{G}_{\eta}^{-1} \left(\frac{\partial \mathbf{h}(\mathbf{s}[t])}{\partial \mathbf{s}[t]}\right)\right) \\ &\quad - (\mathbf{G}_{\omega}^{-1})^T \hat{\mathbf{F}}(\mathbf{J}[t-1] + \hat{\mathbf{F}}^T \mathbf{G}_{\omega}^{-1} \hat{\mathbf{F}})^{-1} \hat{\mathbf{F}}^T \mathbf{G}_{\omega}^{-1} \\ &\approx \mathbf{G}_{\omega}^{-1} + \mathbf{H}^T[t] \mathbf{G}_{\eta}^{-1} \mathbf{H}[t] - (\mathbf{G}_{\omega}^{-1})^T \\ &\quad \hat{\mathbf{F}}(\mathbf{J}[t-1] + \hat{\mathbf{F}}^T \mathbf{G}_{\omega}^{-1} \hat{\mathbf{F}})^{-1} \hat{\mathbf{F}}^T \mathbf{G}_{\omega}^{-1} \\ &= \mathbf{H}^T[t] \mathbf{G}_{\eta}^{-1} \mathbf{H}[t] + (\hat{\mathbf{F}} \mathbf{J}^{-1}[t-1] \hat{\mathbf{F}}^T + \mathbf{G}_{\omega})^{-1} \end{aligned} \quad (32)$$

where $\hat{\mathbf{F}} = \text{diag}(\mathbf{F}, \dots, \mathbf{F})$, $\mathbf{G}_{\omega} = \text{diag}(\mathbf{G}_{\omega_1}, \dots, \mathbf{G}_{\omega_K})$, $\mathbf{G}_{\eta} = \text{diag}(\mathbf{G}_{\eta_1} + \mathbf{G}_1^T, \dots, \mathbf{G}_{\eta_K} + \mathbf{G}_K^T)$, $\mathbf{H}[t] = [\mathbf{H}_1[t], \dots, \mathbf{H}_K[t]] \in \mathbb{R}^{2RM_r QL \times 4K}$, $\tilde{\mathbf{H}}_k[t] = \frac{\partial \mathbf{h}_i(\mathbf{s}_{\mathcal{A}_i}[t])}{\partial \mathbf{s}_k[t]} \Big|_{\mathbf{s}_k[t] = \hat{\mathbf{s}}_k[t|t-1], k \in \mathcal{A}_i} \in$

$\mathbb{R}^{2RM_r QL \times 4}$. The second approximation holds by adopting the following first-order Taylor approximation,

$$\mathbf{h}(\mathbf{s}[t]) \approx \mathbf{h}(\hat{\mathbf{s}}[t|t-1]) + \frac{\partial \mathbf{h}}{\partial \mathbf{s}[t]} \Big|_{\mathbf{s}[t] = \hat{\mathbf{s}}[t|t-1]} (\mathbf{s}[t] - \hat{\mathbf{s}}[t|t-1]), \quad (33)$$

and the third equation holds due to the matrix inversion, i.e.,

$$\Sigma - \Sigma \mathbf{Z} (\mathbf{R} + \mathbf{Z}^T \Sigma \mathbf{Z})^{-1} \mathbf{Z}^T \Sigma = (\mathbf{Z} \mathbf{R}^{-1} \mathbf{Z}^T + \Sigma^{-1})^{-1}. \quad (34)$$

We split $\mathbf{J}[t]$ into two parts, i.e., $\mathbf{J}[t] = \mathbf{J}^1[t] + \mathbf{J}^2[t]$, where $\mathbf{J}^1[t] = \mathbf{H}^T[t] \mathbf{G}_{\eta}^{-1} \mathbf{H}[t]$ and $\mathbf{J}^2[t] = (\hat{\mathbf{F}} \mathbf{J}^{-1}[t-1] \hat{\mathbf{F}}^T + \mathbf{G}_{\omega})^{-1}$. $\mathbf{J}^1[t]$ can be further written as

$$\mathbf{J}^1[t] = \begin{bmatrix} \tilde{\mathbf{H}}_1^T[t] \mathbf{G}_{\eta}^{-1} \tilde{\mathbf{H}}_1[t] & \dots & \tilde{\mathbf{H}}_K^T[t] \mathbf{G}_{\eta}^{-1} \tilde{\mathbf{H}}_1[t] \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{H}}_1^T[t] \mathbf{G}_{\eta}^{-1} \tilde{\mathbf{H}}_K[t] & \dots & \tilde{\mathbf{H}}_K^T[t] \mathbf{G}_{\eta}^{-1} \tilde{\mathbf{H}}_K[t] \end{bmatrix}. \quad (35)$$

On the other hand, we can see that $\mathbf{J}^2[t]$ is a block diagonal matrix if $\mathbf{J}[t-1]$ is a diagonal matrix. Based on the expressions of $\mathbf{J}^1[t]$ and $\mathbf{J}^2[t]$, we have the following Lemma.

Lemma: When $|(\mathbf{a}_t^l(\beta_k))^H \mathbf{a}_t^l(\beta_{k'})| = 0$ for $\beta_k \neq \beta_{k'}$ and $l = 1, \dots, L$, $\mathbf{J}[t]$ is a block diagonal matrix, where the k th diagonal matrix $\mathbf{J}_k[t] = \tilde{\mathbf{H}}_k^T[t] \mathbf{G}_{\eta}^{-1} \tilde{\mathbf{H}}_k[t] + (\mathbf{F} \mathbf{J}_k^{-1}[t-1] \mathbf{F}^T + \mathbf{G}_{\omega_k})^{-1}$.

Proof: When $|(\mathbf{a}_t^l(\beta_k))^H \mathbf{a}_t^l(\beta_{k'})| = 0$ for $\beta_k \neq \beta_{k'}$, the interference between terminals will vanish, as given in (8). We divide all elements in $\mathbf{h}(\mathbf{s}[t])$ into some groups, and each group corresponds to a subcarrier. Since each subcarrier is allocated to only one terminal, each group in $\mathbf{h}(\mathbf{s}[t])$ only contains parameters of one terminal. Therefore, in $\mathbf{H}_k[t] = \frac{\partial \mathbf{h}_i(\mathbf{s}_{\mathcal{A}_i}[t])}{\partial \mathbf{s}_k[t]} \Big|_{\mathbf{s}_k[t] = \hat{\mathbf{s}}_k[t|t-1], k \in \mathcal{A}_i}$, only groups with subcarriers allocated to terminal k are nonzero and elements in other groups are zero. For different terminals k and k' , the nonzero terms of $\tilde{\mathbf{H}}_k[t]$ necessarily correspond to the zero terms of $\tilde{\mathbf{H}}_{k'}[t]$. Therefore, $\tilde{\mathbf{H}}_k^T[t] \mathbf{G}_{\eta}^{-1} \tilde{\mathbf{H}}_{k'}[t] = 0$ for $k \neq k'$. In this case, (35) will degenerate into a block diagonal matrix, i.e.,

$$\mathbf{J}^1[t] = \begin{bmatrix} \tilde{\mathbf{H}}_1^T[t] \mathbf{G}_{\eta}^{-1} \tilde{\mathbf{H}}_1[t] & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \tilde{\mathbf{H}}_K^T[t] \mathbf{G}_{\eta}^{-1} \tilde{\mathbf{H}}_K[t] \end{bmatrix}. \quad (36)$$

Therefore, $\mathbf{J}[t]$ is also a block diagonal matrix with the k th diagonal matrix

$$\mathbf{J}_k[t] = \underbrace{\tilde{\mathbf{H}}_k^T[t] \mathbf{G}_{\eta}^{-1} \tilde{\mathbf{H}}_k[t]}_{\mathbf{J}_{k,1}[t]} + \underbrace{(\mathbf{F} \mathbf{J}_k^{-1}[t-1] \mathbf{F}^T + \mathbf{G}_{\omega_k})^{-1}}_{\mathbf{J}_{k,2}[t]}. \quad (37)$$

The proof is completed. $\blacksquare$

When $\mathbf{J}[t]$ is a block diagonal matrix, PCRLB matrix of the k th terminal parameters estimation can be given by $\mathbf{C}_k[t] = \mathbf{J}_k^{-1}[t]$. It is observed that the recursive formula of $\mathbf{C}_k^{-1}[t]$ in (37) is the same as the that of $\mathbf{M}_k^{-1}[t|t]$ in (19). Supposing the initial state of the k th terminal $\mathbf{s}_k[0] \sim \mathcal{N}(\boldsymbol{\mu}_{sk}, \mathbf{C}_{sk})$, the initial posterior FIM of the k th terminal can be calculated as

$$\mathbf{J}_k[0] = -\mathbb{E}\left(\frac{\partial^2 \log p(\mathbf{s}_k[0])}{\partial^2(\mathbf{s}_k[0])}\right) = \mathbf{C}_{sk}^{-1}, \quad (38)$$

where $p(\mathbf{s}_k[0])$ denotes the priori PDF of $\mathbf{s}_k[0]$. Therefore, in EKF, when we choose the initial estimation as $\hat{\mathbf{s}}_k[0|0] = \boldsymbol{\mu}_{s_k}$, the MSE of EKF will be equal to $\mathbf{C}_k[t]$.

Next, we will derive the specific expression of $\mathbf{J}_{k,1}[t]$, which will facilitate our later optimization. Define a new parameter vector to be estimated $\boldsymbol{\xi}_k = [f_{1,k}, \theta_{1,k}, \beta_{1,k}, \tau_{1,k}, \dots, f_{R,k}, \theta_{R,k}, \beta_{R,k}, \tau_{R,k}]^T \in \mathbb{R}^{4R \times 1}$. With the help of the chain rule, $\mathbf{J}_{k,1}[t]$ can be expressed as

$$\begin{aligned} \mathbf{J}_{k,1}[t] &= \frac{\partial \mathbf{h}_i^T(\mathbf{s}_{A_i}[t])}{\partial \mathbf{s}_k[t]} \mathbf{G}_\eta^{-1} \frac{\partial \mathbf{h}_i(\mathbf{s}_{A_i}[t])}{\partial \mathbf{s}_k^T[t]} \Big|_{\mathbf{s}_k[t] = \hat{\mathbf{s}}_k[t|t-1]} \\ &= \left[\frac{\partial \mathcal{R}(\dot{\mathbf{h}}_i^T(\mathbf{s}_{A_i}[t]))}{\partial \mathbf{s}_k[t]}, \frac{\partial \mathcal{I}(\dot{\mathbf{h}}_i^T(\mathbf{s}_{A_i}[t]))}{\partial \mathbf{s}_k[t]} \right] \mathbf{G}_\eta^{-1} \\ &\quad \left[\frac{\partial \mathcal{R}(\dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t]))}{\partial \mathbf{s}_k^T[t]}, \frac{\partial \mathcal{I}(\dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t]))}{\partial \mathbf{s}_k^T[t]} \right] \Big|_{\mathbf{s}_k[t] = \hat{\mathbf{s}}_k[t|t-1]} \\ &= \frac{2L}{\sigma^2} \mathcal{R} \left(\frac{\partial \dot{\mathbf{h}}_i^H(\mathbf{s}_{A_i}[t])}{\partial \mathbf{s}_k[t]} \frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial \mathbf{s}_k^T[t]} \right) \Big|_{\mathbf{s}_k[t] = \hat{\mathbf{s}}_k[t|t-1]} \\ &= \mathbf{U} \mathbf{J}(\boldsymbol{\xi}_k[t]) \mathbf{U}^T \Big|_{\boldsymbol{\xi}_k[t] = \hat{\boldsymbol{\xi}}_k[t|t-1]}, \end{aligned} \quad (39)$$

where $\mathbf{U} = \frac{\partial \boldsymbol{\xi}_k}{\partial \mathbf{s}_k} \in \mathbb{R}^{4 \times 4R}$ is the Jacobian matrix, $\mathbf{J}(\boldsymbol{\xi}_k[t]) = \frac{2L}{\sigma^2} \mathcal{R} \left(\frac{\partial \dot{\mathbf{h}}_i^H(\mathbf{s}_{A_i}[t])}{\partial \boldsymbol{\xi}_k} \frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial \boldsymbol{\xi}_k^T} \right) \cdot \frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial \boldsymbol{\xi}_k}$ can be written as

$$\begin{aligned} \frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial \boldsymbol{\xi}_k} &= \left[\frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial f_{1,k}}, \frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial \theta_{1,k}}, \frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial \beta_{1,k}}, \right. \\ &\quad \left. \frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial \tau_{1,k}}, \dots, \frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial f_{R,k}}, \right. \\ &\quad \left. \frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial \theta_{R,k}}, \frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial \beta_{R,k}}, \frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial \tau_{R,k}} \right], \end{aligned} \quad (40)$$

where

$$\begin{aligned} \left[\frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial f_{r,k}} \right]_{rm_r,ql} &= j2\pi(\hat{m}_{r,k}T_{cp} + (q-1)T_o + T/2) \\ &\quad \sqrt{M_t} \rho_{k,l} A_{r,k} b_{k,l}^q e^{j2\pi(l-1)\Delta f(\hat{m}_{r,k}T_{cp} - \tau_{r,k})} e^{-j\frac{2\pi}{\lambda_l}(m-1)d\cos\theta_{r,k}} \\ &\quad (\mathbf{a}_{t1}^l(\beta_k))^T (\mathbf{a}_t^l(\hat{\beta}_k))^* e^{j2\pi f_{r,k}(\hat{m}_{r,k}T_{cp} + (q-1)T_o + T/2)}, \end{aligned} \quad (41)$$

$$\begin{aligned} \left[\frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial \theta_{r,k}} \right]_{rm_r,ql} &= j\frac{2\pi}{\lambda_l}(m-1)d\sin\theta_{r,k} \sqrt{M_t} \rho_{k,l} A_{r,k} b_{k,l}^q \\ &\quad e^{j2\pi(l-1)\Delta f(\hat{m}_{r,k}T_{cp} - \tau_{r,k})} e^{j2\pi f_{r,k}(\hat{m}_{r,k}T_{cp} + (q-1)T_o + T/2)} \\ &\quad e^{-j\frac{2\pi}{\lambda_l}(m-1)d\cos\theta_{r,k}} (\mathbf{a}_{t1}^l(\beta_k))^T (\mathbf{a}_t^l(\hat{\beta}_k))^*, \end{aligned} \quad (42)$$

$$\begin{aligned} \left[\frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial \beta_{r,k}} \right]_{rm_r,ql} &= \sqrt{M_t} \rho_{k,l} A_{r,k} b_{k,l}^q e^{-j\frac{2\pi}{\lambda_l}(m-1)d\cos\theta_{r,k}} \\ &\quad e^{j2\pi(l-1)\Delta f(\hat{m}_{r,k}T_{cp} - \tau_{r,k})} e^{j2\pi f_{r,k}(\hat{m}_{r,k}T_{cp} + (q-1)T_o + T/2)} \\ &\quad (\mathbf{a}_{t1}^l(\beta_k))^T (\mathbf{a}_t^l(\hat{\beta}_k))^*, \end{aligned} \quad (43)$$

$$\begin{aligned} \left[\frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial \tau_{r,k}} \right]_{rm_r,ql} &= -j2\pi(l-1)\Delta f \sqrt{M_t} \rho_{k,l} A_{r,k} b_{k,l}^q \\ &\quad e^{j2\pi(l-1)\Delta f(\hat{m}_{r,k}T_{cp} - \tau_{r,k})} e^{j2\pi f_{r,k}(\hat{m}_{r,k}T_{cp} + (q-1)T_o + T/2)} \\ &\quad (\mathbf{a}_{t1}^l(\beta_k))^T (\mathbf{a}_t^l(\hat{\beta}_k))^* e^{-j\frac{2\pi}{\lambda_l}(m-1)d\cos\theta_{r,k}}, \end{aligned} \quad (44)$$

and

$$\begin{aligned} \mathbf{a}_{t1}^l(\beta_k) &= \frac{1}{\sqrt{M_t}} [0, j\frac{2\pi}{\lambda_l} d \sin\beta_k e^{-j\frac{2\pi}{\lambda_l} d \cos\beta_k}, \dots, \\ &\quad j\frac{2\pi}{\lambda_l} d(M_t - 1) \sin\beta_k e^{-j\frac{2\pi}{\lambda_l} (M_t - 1) d \cos\beta_k}]^T. \end{aligned} \quad (45)$$

Defining $b_{k,l}^q \rho_{k,l} \mathbf{\Gamma}_{k,l}^i[t] = [\frac{\partial \dot{\mathbf{h}}_i(\mathbf{s}_{A_i}[t])}{\partial \boldsymbol{\xi}_k}]((l-1)M_rQR + 1 : lM_rQR, i)|_{\boldsymbol{\xi}_k[t] = \hat{\boldsymbol{\xi}}_k[t|t-1]}$, $\mathbf{B}_{k,l}[t](i, j) = (\mathbf{\Gamma}_{k,l}^i[t])^H \mathbf{\Gamma}_{k,l}^j[t]$, the (z, z') -th element of $\mathbf{J}_{k,1}[t]$ can be written as

$$\begin{aligned} \mathbf{J}_{k,1}[t](z, z') &= \mathbf{U}(z, :) \mathbf{J}(\hat{\boldsymbol{\xi}}_k[t|t-1]) \mathbf{U}^T(:, z') \\ &= \sum_{i=1}^{4R} \sum_{j=1}^{4R} \sum_{l=1}^L \mathbf{U}(z, i) \mathbf{U}^T(j, z') p_{k,l} \rho_{k,l} \mathcal{R}[\mathbf{B}_{k,l}[t](i, j)] \\ &= \sum_{l=1}^L p_{k,l} \rho_{k,l} \tilde{\mathbf{B}}_{k,l}[t](z, z'), \end{aligned} \quad (46)$$

where $\tilde{\mathbf{B}}_{k,l}[t](z, z') = \frac{2L}{\sigma^2} \sum_{i=1}^{4R} \sum_{j=1}^{4R} \mathbf{U}(z, i) \mathbf{U}^T(j, z')$

$\mathcal{R}[\mathbf{B}_{k,l}[t](i, j)]$. For ease of representation, we define $\mathbf{D}_k[t] = (\mathbf{F} \mathbf{J}_k^{-1}[t-1] \mathbf{F}^T + \mathbf{G}_{\omega_k})^{-1}$. Then, the (z, z') -th element of $\mathbf{J}_k[t]$ can be expressed by

$$[\mathbf{J}_k[t]](z, z') = \sum_{l=1}^L p_{k,l} \rho_{k,l} \tilde{\mathbf{B}}_{k,l}[t](z, z') + \mathbf{D}_k[t](z, z'). \quad (47)$$

Notice that $\mathbf{J}_k[t-1]$ is independent of the transmission parameters at time t . Therefore, $\mathbf{D}_k[t](z, z')$ in (47) can be considered as a constant in the following optimization problem. The diagonal elements of $\mathbf{J}_k^{-1}[t]$ correspond the minimum MSE of estimation of $(d_{k,x}[t], d_{k,y}[t], v_{k,x}[t], v_{k,y}[t])$ respectively. However, the complex expression of $\mathbf{J}_k^{-1}[t](i, i)$ makes it difficult to be applied to the following optimization problem. Therefore, we take the 2×2 submatrix of the upper left corner of $\mathbf{J}_k[t]$ to calculate the lower bound of PCRLB of $(d_{k,x}[t], d_{k,y}[t])$. Letting $\mathbf{J}_{kd}[t] = \begin{bmatrix} \mathbf{J}_k[t](1, 1) & \mathbf{J}_k[t](1, 2) \\ \mathbf{J}_k[t](2, 1) & \mathbf{J}_k[t](2, 2) \end{bmatrix}$, it is easy to prove that $[\mathbf{J}_{kd}^{-1}[t]](i, i) \leq [\mathbf{J}_k^{-1}[t]](i, i)$ with the inverse of the block matrix, where equality holds when $\mathbf{J}_k[t]$ is a block diagonal matrix.

V. SUBCARRIER AND POWER ALLOCATION DURING TRACKING

In this section, we will perform subcarrier and power allocation to capture the optimal tradeoff between transmission rate and prediction accuracy. When optimizing the transmit parameters at time t , we have obtained the predicted locations and velocities of terminals with measure information before time t via EKF or UKF. With predicted locations, we can know $\hat{m}_{r,k}$ and $\hat{\beta}_k$ in Equation (9). Resource allocation among terminals also needs to know the channel information between the BS and terminals, which can be calculated according to predicted location information. Because the power and subcarrier allocation at time t influences transmission rate at time t and prediction accuracy at time $t+1$, thereby affecting the transmission rate at time $t+1$, we need to jointly optimize the subcarrier and power allocation in each stage to

maximize the transmission rate while satisfying the prediction accuracy requirement. In each stage, we consider the following optimization problem

$$\begin{aligned} & \max_{\{\rho_{k,l}, p_{k,l}\}_{k=1,\dots,K}^{l=1,\dots,L}} \bar{R} \quad (48) \\ \text{s.t.} \quad & \sum_{k=1}^K \sum_{l=1}^L \rho_{k,l} p_{k,l} \leq P_T, \quad (48a) \\ & p_{k,l} \geq 0, \rho_{k,l} \in \{0, 1\}, \forall k, l, \quad (48b) \\ & \sum_{k=1}^K \rho_{k,l} = 1, \forall l, \quad (48c) \\ & \mathbf{J}_{kd}^{-1}[t](1, 1) \leq \zeta_k[t], \forall k, \quad (48d) \\ & \mathbf{J}_{kd}^{-1}[t](2, 2) \leq \zeta_k[t], \forall k, \quad (48e) \end{aligned}$$

where $\zeta_k[t]$ denotes the PCRLB constraint of the k th terminal, and P_T is the total transmit power constraint. (48d) and (48e) denotes the requirements of location estimation accuracy at time t . It can be seen that (48d) and (48e) are non-convex constraints. Therefore, we make some transformations to the constraints. According to (48d), we have

$$\begin{bmatrix} \mathbf{J}_{kd}[t](1, 1) & \mathbf{J}_{kd}[t](1, 2) \\ \mathbf{J}_{kd}[t](2, 1) & \mathbf{J}_{kd}[t](2, 2) \end{bmatrix}_{1,1}^{-1} \leq \zeta_k[t] \quad (49)$$

$$\Rightarrow \frac{\mathbf{J}_{kd}[t](2, 2)}{\mathbf{J}_{kd}[t](1, 1)\mathbf{J}_{kd}[t](2, 2) - \mathbf{J}_{kd}[t](1, 2)\mathbf{J}_{kd}[t](2, 1)} \leq \zeta_k[t] \quad (50)$$

$$\Rightarrow \frac{\mathbf{J}_{kd}[t](2, 2)(\mathbf{J}_{kd}[t](1, 1)\zeta_k[t] - 1) - \mathbf{J}_{kd}[t](1, 2)\mathbf{J}_{kd}[t](2, 1)\zeta_k[t]}{\mathbf{J}_{kd}[t](1, 1)\mathbf{J}_{kd}[t](2, 2) - \mathbf{J}_{kd}[t](1, 2)\mathbf{J}_{kd}[t](2, 1)} \geq 0. \quad (51)$$

Constructing the matrix

$$\mathbf{L}(d_{k,x}[t]) = \begin{bmatrix} \mathbf{J}_{kd}[t](2, 2) & \mathbf{J}_{kd}[t](1, 2)\sqrt{\zeta_k[t]} \\ \mathbf{J}_{kd}[t](2, 1)\sqrt{\zeta_k[t]} & \mathbf{J}_{kd}[t](1, 1)\zeta_k[t] - 1 \end{bmatrix}, \quad (52)$$

it can be found that $\mathbf{L}(d_{k,x}[t])$ is a positive semidefinite matrix when (51) holds. Similarly, we construct

$$\mathbf{L}(d_{k,y}[t]) = \begin{bmatrix} \mathbf{J}_{kd}[t](1, 1) & \mathbf{J}_{kd}[t](1, 2)\sqrt{\zeta_k[t]} \\ \mathbf{J}_{kd}[t](2, 1)\sqrt{\zeta_k[t]} & \mathbf{J}_{kd}[t](2, 2)\zeta_k[t] - 1 \end{bmatrix}, \quad (53)$$

so that the constraint in (48e) can be converted to $\mathbf{L}(d_{k,y}[t]) \succcurlyeq 0$. Then, the optimization problem in (48) can be rewritten as

$$\begin{aligned} & \max_{\{\rho_{k,l}, p_{k,l}\}_{k=1,\dots,K}^{l=1,\dots,L}} \bar{R} \\ \text{s.t.} \quad & \sum_{k=1}^K \sum_{l=1}^L \rho_{k,l} p_{k,l} \leq P_T, \\ & p_{k,l} \geq 0, \rho_{k,l} \in \{0, 1\}, \forall k, l, \quad (54) \\ & \sum_{k=1}^K \rho_{k,l} = 1, \forall l, \\ & \mathbf{L}(d_{k,x}[t]) \succcurlyeq 0, \forall k, \\ & \mathbf{L}(d_{k,y}[t]) \succcurlyeq 0, \forall k. \end{aligned}$$

The problem in (54) is still a non-convex mixed integer programming problem. A conventional method is to solve (54) with alternating optimization. Specifically, for given $\{\rho_{k,l}\}$,

the problem is a semi-definite programming (SDP) problem and can be solved with convex optimization tool and for given $\{p_{k,l}\}$, the problem is a semi-definite integer programming and can be solved with optimization tool YALMIP. However, this method is only applicable to situations where the number of subcarriers is relatively small. When the number of subcarriers is large (e.g., $L \geq 64$), the solution by optimization tool YALMIP is time-consuming. Therefore, we propose a low-complexity solution to (54). Let $\tilde{p}_{k,l} = \rho_{k,l} p_{k,l}$, and relax $\rho_{k,l}$ to the real number in $(0, 1]$. To restrict $\rho_{k,l}$ to take values at the endpoints, we add the constraint

$$\rho_{k,l}(1 - \rho_{k,l}) \leq 0, \forall k, l. \quad (55)$$

Therefore, the problem (54) is transformed to

$$\max_{\{\rho_{k,l}, \tilde{p}_{k,l}\}_{k=1,\dots,K}^{l=1,\dots,L}} \bar{R}' \quad (56)$$

$$\text{s.t.} \quad \sum_{k=1}^K \sum_{l=1}^L \tilde{p}_{k,l} \leq P_T, \quad (56a)$$

$$\tilde{p}_{k,l} \geq 0, 0 < \rho_{k,l} \leq 1, \forall k, l, \quad (56b)$$

$$\rho_{k,l}(1 - \rho_{k,l}) \leq 0, \forall k, l, \quad (56c)$$

$$\sum_{k=1}^K \rho_{k,l} = 1, \forall l, \quad (56d)$$

$$\mathbf{L}(d_{k,x}[t]) \succcurlyeq 0, \forall k, \quad (56e)$$

$$\mathbf{L}(d_{k,y}[t]) \succcurlyeq 0, \forall k, \quad (56f)$$

where $\bar{R}' = \frac{1}{L} \sum_{k=1}^K \sum_{l=1}^L \rho_{k,l} \log \left(1 + \frac{|\mathbf{g}_{k,l} \mathbf{w}_{k,l}|^2 \tilde{p}_{k,l}}{\rho_{k,l} \sigma_C^2} \right)$ is a concave function on $\rho_{k,l}$ and $\tilde{p}_{k,l}$. Since $f(\rho_{k,l}) = \rho_{k,l}(1 - \rho_{k,l})$ in (56c) is a non-convex function, we utilize a first-order Taylor expansion at the point $\rho_{k,l}^{(j)}$ to obtain a convex upper bound function of $f(\rho_{k,l})$, i.e., $(\rho_{k,l}^{(j)})^2 + \rho_{k,l}(1 - 2\rho_{k,l}^{(j)})$, where j denotes the iteration index. By introducing a slack variable $\alpha_{k,l}$, the constraint (56c) is converted to

$$(\rho_{k,l}^{(j)})^2 + \rho_{k,l}(1 - 2\rho_{k,l}^{(j)}) \leq \alpha_{k,l}, \forall k, l. \quad (57)$$

Then, the problem in (56) can be solved by iteratively solving the following problem,

$$\max_{\{\rho_{k,l}, \tilde{p}_{k,l}, \alpha_{k,l}\}_{k=1,\dots,K}^{l=1,\dots,L}} \bar{R}' - \eta^{(j)} V \quad (58)$$

$$\text{s.t.} \quad \sum_{k=1}^K \sum_{l=1}^L \tilde{p}_{k,l} \leq P_T, \quad (58a)$$

$$\tilde{p}_{k,l} \geq 0, 0 < \rho_{k,l} \leq 1, \alpha_{k,l} \geq 0, \forall k, l, \quad (58b)$$

$$(\rho_{k,l}^{(j)})^2 + \rho_{k,l}(1 - 2\rho_{k,l}^{(j)}) \leq \alpha_{k,l}, \forall k, l, \quad (58c)$$

$$\sum_{k=1}^K \rho_{k,l} = 1, \forall l, \quad (58d)$$

$$\mathbf{L}(d_{k,x}[t]) \succcurlyeq 0, \forall k, \quad (58e)$$

$$\mathbf{L}(d_{k,y}[t]) \succcurlyeq 0, \forall k, \quad (58f)$$

where $\eta^{(j)}$ is the penalty coefficient at the j th iteration, and $V = \sum_{k=1}^K \sum_{l=1}^L \alpha_{k,l}$ denotes the total violation. In the $(j+1)$ th iteration, the values of $\{\rho_{k,l}^{(j+1)}\}$ are set to the solution of problem (58) at the j th iteration. $\eta^{(j)}$ is updated according

to $\eta^{(j+1)} = \min(\delta\eta^{(j)}, \eta_{max})$, where η_{max} is the maximum penalty coefficient, $\delta > 1$ is a constant to ensure $V \rightarrow 0$. The successive convex approximation algorithm for solving (58) is given in Algorithm 1.

Algorithm 1: Solution to Problem (58)

```

1 Set initial values:  $\{\rho_{k,l}^{(j)}\}, \eta^{(j)} > 0, \eta_{max}, \delta > 1, \varepsilon,$ 
   $j = 0;$ 
2 repeat
3   Solve (58) with CVX to obtain optimal solution
     $\{\rho_{k,l}^{j*}, \tilde{p}_{k,l}^*, \alpha_{k,l}^*\}$  and calculate  $\bar{R}^{(j)}$ ;
4   Set  $\rho_{k,l}^{(j+1)} = \rho_{k,l}^{j*}$  in (58c);
5   Set  $\eta^{(j+1)} = \min(\delta\eta^{(j)}, \eta_{max})$ ;
6   Set  $j = j + 1$ ;
7 until  $|\bar{R}^{(j+1)} - \bar{R}^{(j)}| < \varepsilon;$ 

```

In the problem (48), PCRLB constraints $\{\zeta_k[t]\}_{k=1}^K$ determine the impact of subcarrier/power allocation at time t on the transmission rates at current and future time. Specifically, increasing $\zeta_k[t]$ can improve the current transmission rate due to the expanded feasible set, but it may decrease the rate at the next time due to the reduced prediction accuracy. Therefore, to improve the overall performance during tracking, we proposed to optimize $\zeta_k[t]$ using reinforcement learning (RL).

In the framework of RL, in each time-step t , the agent (training is implemented in CP) selects an action a_t according to the policy π that is a probability distribution conditioned on the current state s_t , i.e., $a_t \sim \pi(\cdot|s_t)$. After executing the selected action a_t , the system state transfers to s_{t+1} according to the distribution $P(\cdot|s_t, a_t)$ (which is typically unknown), and the agent will receive a reward r_t fed back by the system. The objective of RL is to optimize π to maximize the expected average reward $\eta(\pi) = \mathbb{E}\{\sum_{t=0}^{T'} r_t|\pi\}$, in which T' is the total tracking time. In our problem, the definitions of state, action, and reward function are given as follows.

State: In each time-step t , the global state should have been the coalition of true states of all terminals. However, we do not have the knowledge of the real state but the predicted one $\hat{s}_k[t|t-1], \forall k$. Therefore, the (partially observed) state is defined as

$$s_t = (\hat{s}_1[t|t-1], \dots, \hat{s}_K[t|t-1]). \quad (59)$$

Action: The RL part only concentrates on the decision of PCRLB constraints $\zeta_k[t]$, so that the action in RL is defined as

$$a_t = (\zeta_1[t], \dots, \zeta_K[t]). \quad (60)$$

Reward: In each time t , the reward function is given by

$$r(s_t, a_t) = \bar{R}[t], \quad (61)$$

where $\bar{R}[t]$ is the weighted sum rate defined in Eq. (9). We will abbreviate $r(s_t, a_t)$ as r_t for simplicity.

The overall diagram of beam tracking with parameters optimization is given in Fig. 2. At each time-step t , the EKF/UKF tracking algorithm predicts the state s_t based on $\{\hat{s}_k[t-1|t-1]\}_{k=1}^K$. Then, the RL agent selects an action

a_t according to $\pi(\cdot|s_t)$. Based on the given s_t and a_t , the subcarrier/power allocation parameters $\{\rho_{k,l}[t], p_{k,l}[t]\}_{k=1, \dots, K, l=1, \dots, L}$ will be obtained by Algorithm 1. The transmitting BS transmits to terminals with optimized parameters. The achievable transmission rate r_t is used to update neural network and the echo signal $y[t]$ is used for predicting the next state s_{t+1} .

By interacting with the environment, the agent (CP) can collect the data segment $\{s_t, a_t, r_t, s_{t+1}\}_{t=0}^{T'}$ for updating the policy π (so as to maximize $\eta(\pi)$). Next, we will introduce the updating process. In this paper, we adopt the soft actor-critic (SAC) in [27] to train the policy. In SAC, the policy is parameterized by a tractable neural network with parameters θ . We still denote the policy as π for simplicity later, but it is implicitly parameterized by θ . Besides, a parameterized Q-function Q_ϕ (with parameters ϕ) is introduced, which is the approximation of the action-value function [28]. The policy can be trained by minimizing the following loss function [27],

$$J_\pi(\theta) = \mathbb{E}_{s_t \sim \mathcal{B}} \{ \mathbb{E}_{a_t \sim \pi} \{ \delta \log \pi(a_t|s_t) - Q_{\bar{\phi}}(s_t, a_t) \} \}, \quad (62)$$

where $\mathcal{B}$ is the experience replay buffer for storing historical data, and δ is a temperature factor to encourage the exploration of actions. Further, $Q_{\bar{\phi}}$ is the target Q-function introduced to stabilize the training process [29], whose parameters $\bar{\phi}$ are softly updated from ϕ .

Also, Q_ϕ needs to be updated to approach the real action-value function. Q_ϕ is updated by minimizing the following loss function,

$$J_Q(\phi) = \mathbb{E}_{(s_t, a_t, r_t, s_{t+1}) \sim \mathcal{B}} \{ (Q_\phi(s_t, a_t) - g_t)^2 \}, \quad (63)$$

and g_t is given by

$$g_t = r_t + (Q_{\bar{\phi}}(s_{t+1}, a_{t+1}) - \delta \log \pi(a_{t+1}|s_{t+1})), \quad (64)$$

where a_{t+1} is sampling from $\pi(\cdot|s_{t+1})$. We stress that since we consider the scenario with finite episode length, the discount factor is not applied in calculating g_t , as in [30].

With the definitions of loss functions, the parameters θ, ϕ and $\bar{\phi}$ are updated according to

$$\begin{aligned} \theta &\leftarrow \theta + \psi \nabla_\theta J_\pi(\theta) \\ \phi &\leftarrow \phi + \psi \nabla_\phi J_Q(\phi), \\ \bar{\phi} &\leftarrow \nu \phi + (1 - \nu) \bar{\phi} \end{aligned} \quad (65)$$

where $\psi \in (0, 1)$ is the learning rate, and $\nu \in (0, 1)$ is the moving average factor.

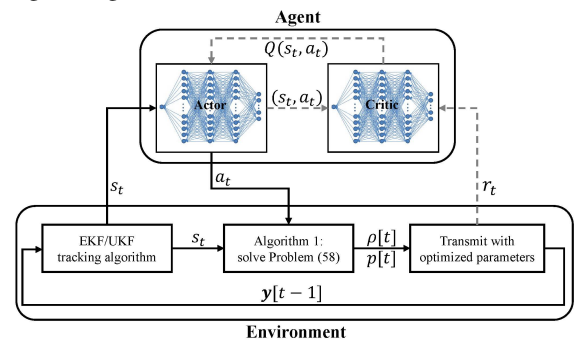


Fig. 2. Diagram of proposed algorithm.

In the algorithm architecture shown in Fig. 2, to maximize $\eta(\pi)$, PCRLB constraints are regarded as output action of

RL and $\{\rho_{k,l}[t], p_{k,l}[t]\}_{k=1,\dots,K, l=1,\dots,L}$ are optimized with successive convex approximation in Algorithm 1. We will explain why $\{\rho_{k,l}[t], p_{k,l}[t]\}_{k=1,\dots,K, l=1,\dots,L}$ are not explicitly set as the output action of RL. Firstly, it is difficult to map the output of the neural network into the set that satisfies the PCRLB constraints. Secondly, $\{\rho_{k,l}[t], p_{k,l}[t]\}_{k=1,\dots,K, l=1,\dots,L}$ is a high-dimensional action vector making it difficult to learn an ideal policy.

Algorithm 2: Beam tracking based on EKF (or UKF) with power and subcarrier allocation

```

1 Initiate policy parameters  $\theta$ , Q-function parameters  $\phi$ ,
  empty replay buffer  $\mathcal{B}$ ;
2 Initiate target Q-function parameters  $\bar{\phi} \leftarrow \phi$ ;
3 for iteration = 1, 2, ..., L do
4   Initiate  $\hat{\mathbf{s}}_k[0|0]$ ,  $\mathbf{M}_k[0|0]$ ,  $\mathbf{J}_k[0] = \mathbf{M}_k[0|0]$ ,  $\forall k$ ;
5   Predict the state  $\hat{\mathbf{s}}_k[1|0]$ ,  $\forall k$  using EKF or UKF
    ((13) or (27)) and get the global state  $\mathbf{s}_1$ ;
6   for  $t = 1, \dots$ , until the tracking process ends do
7     Select  $a_t$  according to  $\pi(\cdot|s_t)$ ;
8     Calculate  $\tilde{\mathbf{B}}_{k,l}[t]$  and  $\mathbf{D}_k[t]$  with (46) and (47),
      and estimate  $\hat{\xi}_k[t|t-1]$  using  $\hat{\mathbf{s}}_k[t|t-1]$ ,  $\forall k$ ;
9     function ResourceAllocation( $\hat{\xi}_k[t|t-1]$ ,
       $\tilde{\mathbf{B}}_{k,l}[t]$ ,  $\mathbf{D}_k[t]$ ,  $a_t$ ):
10      Execute Algorithm 1 to obtain the
        subcarrier/power allocation at time  $t$ ;
11    return  $\{p_{k,l}[t]\}$ ,  $\{\rho_{k,l}[t]\}$ 
12    Set beamformer  $\mathbf{w}_{k,l}[t] = (\mathbf{a}_t^l(\hat{\beta}_k[t|t-1]))^*$ ,
       $\forall k, l$  and allocate the power and subcarriers at
      time  $t$  according to the optimization results;
13    Get the reward  $r_t$  in (10);
14    CP collects echo signals from distributed
      receivers and predicts the state  $\hat{\mathbf{s}}_k[t+1|t]$ ,  $\forall k$ 
      with EKF or UKF ((13) or (27)) and get the
      global state  $\mathbf{s}_{t+1}$ ;
15    Store a data segment  $\{s_t, a_t, r_t, s_{t+1}\}$  into  $\mathcal{B}$ ;
16    Sample  $B$  group of data from  $\mathcal{B}$ ;
17    Calculate loss function  $J_\pi(\theta)$  and  $J_Q(\phi)$  using
      sampled data with (62) and (63);
18    Update  $\theta$ ,  $\phi$  and  $\bar{\phi}$  with (65);
19  end
20 end

```

VI. SIMULATION RESULTS

In this section, we present simulation results to demonstrate the communication and tracking performance in the multiuser MIMO-OFDM system. If not otherwise specified, we set $f_c = 26\text{GHz}$, $\Delta f = 120\text{KHz}$, $T_{cp} = 0.59\mu\text{s}$, $T_o = 8.92\mu\text{s}$, $M_r = 4$, the noise power spectral density of radar receiver and communication receiver are $n_0 = 2.5 \times 10^{-22}$ and $n'_0 = 10^{-18}$, the noise power in (4) is $\sigma^2 = n_0 L \Delta f$ and in (10) is $\sigma_C^2 = n'_0 \Delta f$. Besides, the transmitting BS is located at (0,0)m, and the receiving BSs are located at (0,30)m and (0,-30)m, respectively. The attenuation coefficients in (4) and (10) are set to $A_{r,k} = \sqrt{\frac{\lambda^2 RCS}{(4\pi)^3 (d_k^t)^2 (d_k^r)^2}}$ and $A_k = \sqrt{\frac{\lambda^2}{(4\pi)^2 (d_k^t)^{2.5}}}$, respectively, where d_k^t denotes the distance

between terminal k and the transmitting BS, d_k^r denotes the distance between terminal k and the r th receiving BS, and RCS is set to 0.1. In SAC, all networks have two fully-connected hidden layers, with each layer having 256 neurons, and the activation function is ReLU function. Besides, we set $\psi = 3 \times 10^{-4}$, $\mu = 0.005$, the size of experience replay buffer $\mathcal{B}$ is 1×10^6 and the mini-batch size is 256.

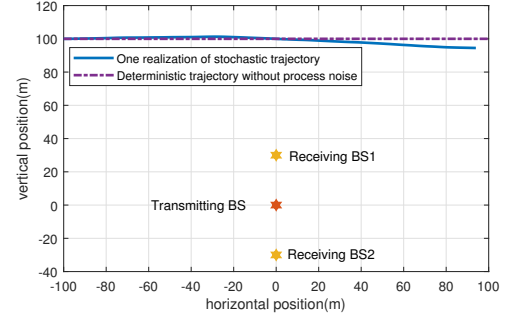


Fig. 3. The trajectory of the terminal with and without process noise.

A. Tracking and communication performance in single terminal case

In this subsection, we investigate the performance of tracking and communication when $K = 1$. We set $Q = 4$, $P_T = 5W$, $\sigma_{\omega_k}^2 = 0.03$. Because $|\mathbf{w}_{k,l}^H \mathbf{g}_{k,l}|^2$ in (10) is approximately equal for different l , we equalize the power between subcarriers, i.e., $p_{k,l} = P_T/L$. We set the initial real location and velocity of the terminal $\mathbf{s}[0] = [-100, 100, 20, 0]$ and the moving trajectory of the terminal follows the process equation in (3). In the considered scenario, the terminal first approaches the BSs and then moves away from them, as shown in Fig. 3. Since the moving trajectory of terminal is stochastic, the solid line in Fig. 3 is just a specific realization of random trajectory. We use EKF and UKF to predict the location of the terminal, and then determine the beamforming vectors on each subcarrier based on the predicted location. We set $\hat{\mathbf{s}}[0|0] = [-103, 103, 20, 0]$ and $\mathbf{M}[0|0] = \text{diag}(9, 9, 0, 0)$. Fig. 4(a) and Fig. 4(b) show the prediction performance of location and velocity of the terminal during the entire tracking process. We set $L = 64$, $M_t = 8, 16, 32, 64$. It can be seen that the prediction accuracy of location and velocity improves as the terminal approaches the BSs and deteriorates as it moves away from them. This occurs because echo signal power is inversely proportional to the distance between the terminal and BSs. Furthermore, increasing the number of transmit antennas helps to improve the tracking performance. Compared with EKF, UKF achieves better prediction accuracy due to the highly nonlinear measurement equation.

Fig. 4(c) shows the average communication rate between the transmitting BS and the terminal during the above tracking process with predicted beam directions. The transmission rates with perfect beamforming vectors, i.e., $\mathbf{w}_{k,l} = (\mathbf{a}_t^l(\beta_k))^*$ are also plotted as a benchmark. It can be seen that the rate loss due to prediction error is almost negligible.

Fig. 5 demonstrates the prediction performance of location and velocity and communication rate versus the total

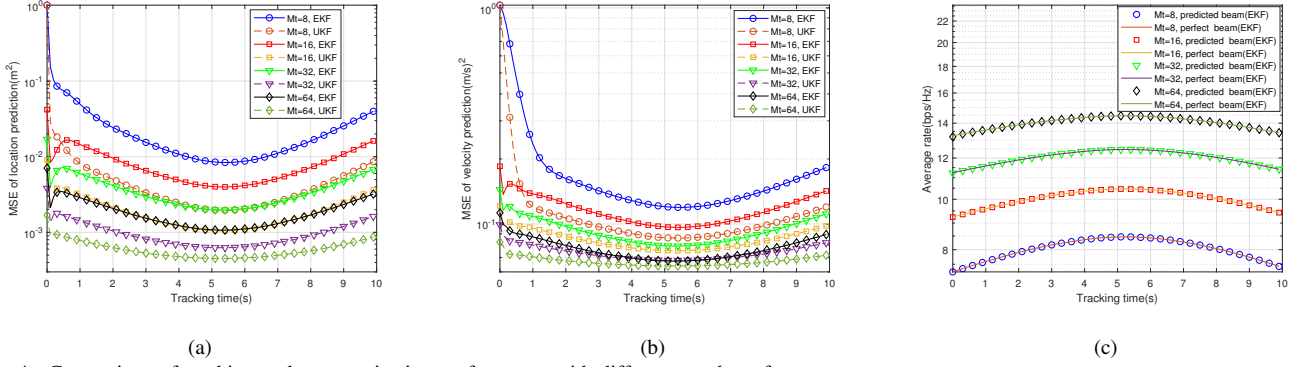


Fig. 4. Comparison of tracking and communication performance with different number of antennas.

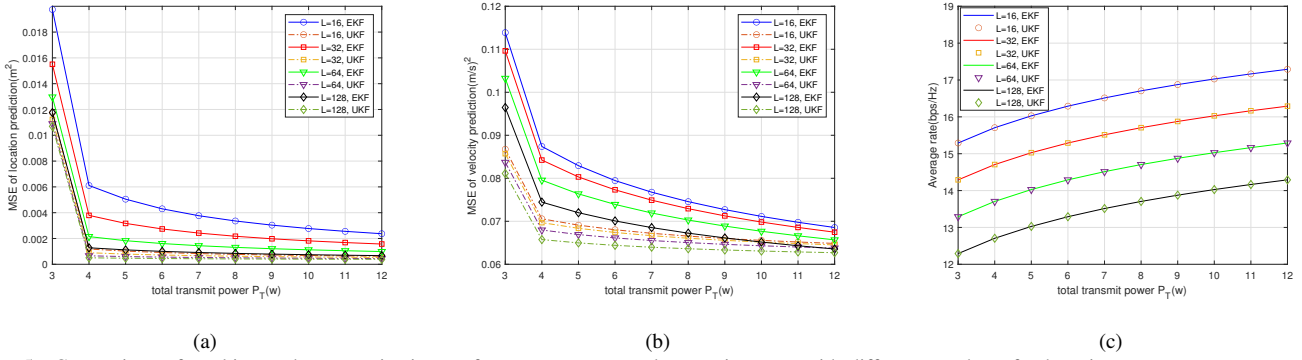


Fig. 5. Comparison of tracking and communication performance versus total transmit power with different number of subcarriers.

transmit power with different number of subcarriers. We set $M_t = 64$. It can be observed that the prediction accuracy and communication rate of the terminal improve with the increasing of total transmit power. As the power increases, the difference between MSEs decreases. Meanwhile, although UKF achieves better prediction performance than EKF, the difference shrinks as the power increases. The difference of prediction performance between UKF and EKF is not reflected in the communication performance, because the prediction errors of both are too small to have negligible impact on the communication performance. We can see that increasing the number of subcarriers improves the prediction accuracy, but decreases the spectral efficiency slightly.

B. Tracking and communication performance in multiple terminals case

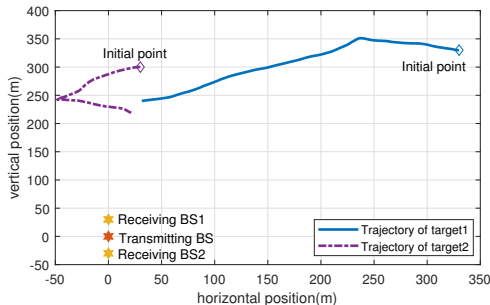


Fig. 6. The trajectories of two terminals with $\mathbf{s}_1[0] = [330, 330, -20, 0]$, $\mathbf{s}_2[0] = [30, 300, -20, 0]$, $\sigma_{\omega_1}^2 = \sigma_{\omega_2}^2 = 8$.

In this subsection, we investigate beam tracking for multiple terminals with optimal subcarrier/power allocation. We set $Q = 4$, $K = 2$, the initial states of the two terminals $\mathbf{s}_1[0] = [330, 330, -20, 0]$, $\mathbf{s}_2[0] = [30, 300, -20, 0]$, $\sigma_{\omega_1}^2 = \sigma_{\omega_2}^2 = 8$. Intuitively, we plot trajectories of two terminals in a tracking experiment in Fig. 6. As stated in Sec. II-B, during the entire tracking process, we need to keep judging whether the relative delay of the two terminals are smaller than T_{cp} . If not, ECA is used in the time domain to eliminate the mutual interference between terminals. In Fig. 7, we evaluate the PCRLB performance for location estimation with different transmit power. We set $M_t = 64$. The ordinate corresponds to the average PCRLB, which is obtained by averaging $\overline{PCRLB}$ in 1500 tracking experiments. In each tracking experiment, $\overline{PCRLB} = \frac{\sum_{t=1}^T (\mathbf{J}^{-1}[t](1,1) + \mathbf{J}^{-1}[t](2,2))}{2T}$ for location estimation (T is the tracking time). In each tracking experiment, for given t , the PCRLB constraints $\{\zeta_k[t]\}_{k=1}^K$ in (48) are determined according to the output of actor neural network in RL, and then we solve (56) to get $\rho_{k,l}[t]$, $p_{k,l}[t]$ and $\mathbf{J}^{-1}[t]$. Therefore, the vertical coordinate in Fig. 7 reveals the predictive performance of terminal states when maximizing communication rate under optimal PCRLB constraints. The curves with ‘exact PCRLB’ are plotted by calculating the diagonal elements of $\mathbf{J}^{-1}[t]$ in (32) with optimized transmit parameters. Lower bounds of PCRLB of location estimation are plotted by calculating the diagonal elements of $[\mathbf{J}_{kd}[t]]^{-1}$, i.e., ignoring the interference between multiple terminals. As the total transmit power increases, the PCRLB decreases

and the gap between PCRLB and its lower bound decreases as well. Additionally, the PCRLB of terminal 1 decreases significantly when L increases from 64 to 128, while the PCRLB of terminal 2 sometimes does not. This is because as L increases, terminal 2 may be allocated fewer resources.

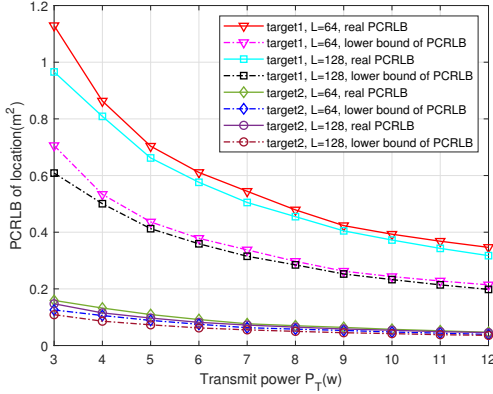


Fig. 7. Average PCRLB of location versus total transmit power with different number of subcarriers when $K = 2$.

Fig. 8 and Fig. 9 show the prediction accuracy and communication rate during tracking, respectively. We set $P_T = 5W$, $L = 64$. ‘MSE of location prediction’ in Fig. 8 is obtained by executing Algorithm 2 and calculating $\mathbf{M}_k[t|t-1]$ with (14). As shown in the measure equation in (11), there exists multi-terminal interference when $|(\mathbf{a}_t^l(\beta_k))^H \mathbf{a}_t^l(\beta_{k'})| \neq 0$ for $\beta_k \neq \beta_{k'}$. Compared with the ideal scenario without interference (ignoring interference item in (11)), the interference reduces the SNR at the radar receiver, thereby degrading the prediction performance. As the tracking time increases, the gap between MSEs with and without interference increases, mainly because the decrease in azimuth of the two terminals leads to increased interference. Obviously, increasing the number of transmit antennas can improve the prediction accuracy during tracking process due to a larger array gain. It is observed that although interference between echo signals can reduce prediction accuracy, its impact on transmission rate is negligible.

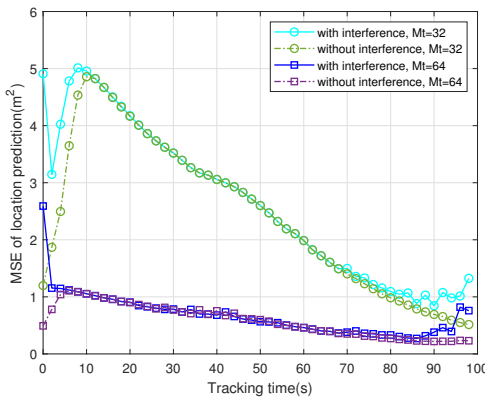


Fig. 8. Average MSEs of location prediction with different number of antennas when $K = 2$.

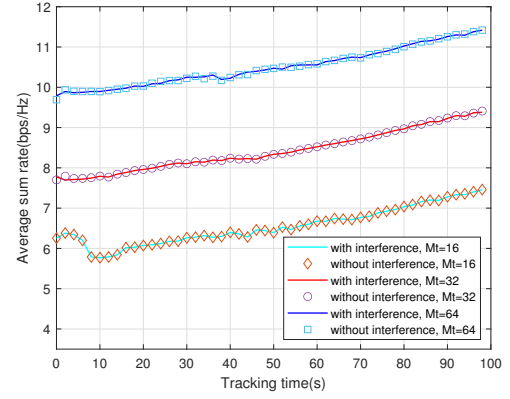


Fig. 9. Average sum rate with different number of antennas when $K = 2$.

C. Performance comparison of radar-assisted and pilot-based communications

In this subsection, we compare performance of sensing based and pilot based beam tracking schemes. In the time division duplex (TDD) system, beam tracking is usually accomplished by channel prediction based on pilots from terminals. In the pilot based comparison scheme, the terminal occupies N symbols in an OFDM block to transmit pilots to the transmitting BS, and then the transmitting BS estimates the location of terminal with EKF and adjusts the beam pointing to the terminal in the following data transmission phase. To ensure the fairness of the comparison, we set all parameters except the observed signals to be the same in the two schemes, including $M_t = 64$, $L = 64$, $Q = 14$, $P_T = 10W$, $\mathbf{s}[0] = [320, 300, 20, 0]$, and $\sigma_{\omega_k}^2 = 0.3$.

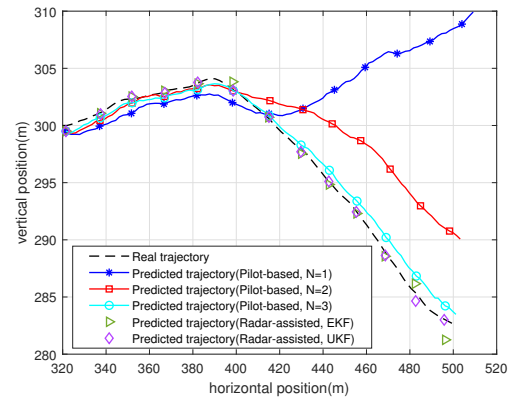


Fig. 10. Comparison of terminal trajectory tracking under radar-assisted and pilot-based methods.

Fig. 10 exhibits tracking trajectories of the terminal based on the radar-assisted beam tracking and pilot-assisted beam tracking in a tracking experiment with $K = 1$. As N decreases from 3 to 1, the pilot-based prediction trajectory gradually deviates from the real trajectory. Nevertheless, the radar-assisted predicted trajectory (EKF or UKF) is much closer to the real trajectory than pilot-based trajectory. This is because only a small number of OFDM symbols are used as pilot symbols in pilot-based scheme, while the sensing-assisted scheme uses all OFDM symbols in each block for both

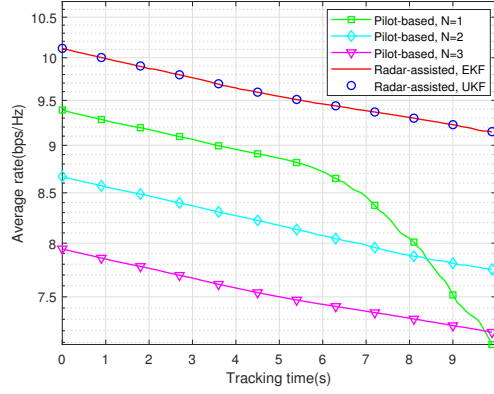


Fig. 11. Comparison of communication rate under radar-assisted and pilot-based methods.

prediction and communication. Besides, it can be observed that predicted trajectory of UKF is closer to the real trajectory than that of the EKF, which is in line with the results in Fig. 4.

In Fig. 11, average communication rates during the tracking process are plotted. Average transmission rate is obtained by averaging rates in 1500 tracking experiments. The average transmission rate gradually decreases because the terminal is heading gradually away from the BSs so that the signal power at the receiver gradually decreases. It can be seen that the rate with pilot-based method starts to drop sharply at the 7th second when $N = 1$. This is because the predicted location deviates significantly from the real location, which leads to the reduction of beamforming gain. Although both the sensing-assisted method and the pilot-based method with $N = 2$ and 3 can realize the beam tracking with low prediction error, the radar-assisted method can achieve higher rate because it avoids dedicated pilot overhead.

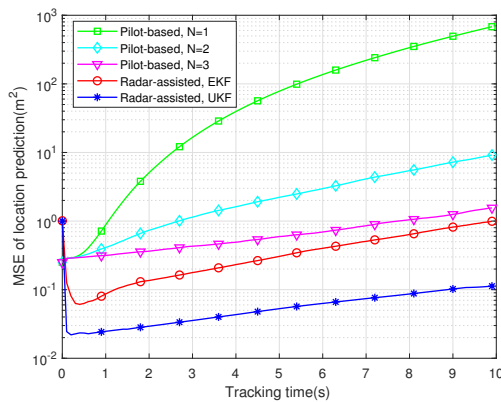


Fig. 12. Comparison of prediction performance of radar-assisted and pilot-based methods when $K = 2$.

In the multi-terminal scenario, we set $K = 2$, $\mathbf{s}_1[0] = [120, 100, 20, 0]$, $\mathbf{s}_2[0] = [300, 240, 20, 0]$. Fig. 12 and Fig. 13 show the MSE of predicted location and the average transmission rate of the radar-assisted method and the pilot-based method, respectively. It can be seen from Fig. 12 that when $N = 1$, the pilot-based method has a large prediction error and cannot realize the beam tracking. The sensing-

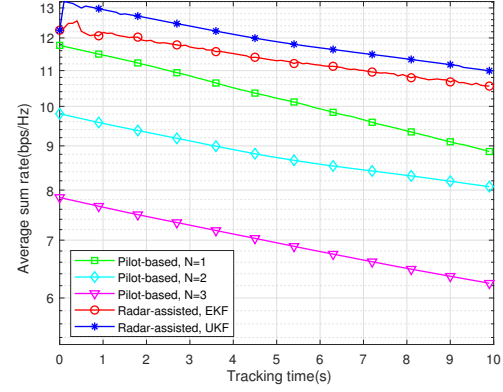


Fig. 13. Comparison of communication rate of radar-assisted and pilot-based methods when $K = 2$.

assisted method and the pilot-based method with $N = 2$ and 3 can realize the beam tracking. Since the terminals gradually move away from the BSs, the MSE of location prediction becomes larger and the transmission rate becomes smaller during the tracking process. Furthermore, unlike Fig. 11, the gap of transmission rates between UKF and EKF becomes more pronounced because the prediction accuracy will effect the resource allocation, which in turn puts a greater impact on transmission rate.

D. Comparison mode

In this subsection, we compare two different radar-assisted beam tracking schemes in multiuser MIMO-OFDM systems. One is the single-stream frequency division scheme in this paper, where each subcarrier is allocated to only one terminal. Another is multi-stream space division scheme, where space division multiple access (SDMA) is adopted in each subcarrier. In SDMA scheme, the base band signal is $\mathbf{x}^q(t) = \sum_{k=1}^K \sum_{l=1}^L \mathbf{w}_{k,l} b_{k,l}^q e^{j2\pi(l-1)\Delta f t} \xi(t - (q-1)T_o)$, $(q-1)T_o \leq t \leq qT_o$. By letting $\mathbf{w}_{k,l} = (\mathbf{a}_t^l(\hat{\beta}_k))^*$, echo signals of terminal k after matched filtering are $\mathbf{y}_r^{k,q,l} = \sqrt{M_t M_r} A_{r,k} b_{k,l}^q e^{j2\pi(l-1)\Delta f (\hat{m}_{r,k} T_{cp} - \tau_{r,k})} \mathbf{a}_r^l(\theta_{r,k}) (\mathbf{a}_t^l(\hat{\beta}_k))^T e^{j2\pi f_{r,k} (\hat{m}_{r,k} T_{cp} + (q-1)T_o + T/2)} + \boldsymbol{\eta}_r^{k,q,l}$. We allocate the total transmit power equally among all subcarriers and in each subcarrier, transmit powers between users are allocated with the method in [18]. Once transmit powers in each subcarrier are determined, PCRLB formula in (32) can be applied to calculate PCRLB in the OFDM-SDMA scheme. In Fig. 14, curves with "Space Division" exhibit the relationship between sum rate and PCRLB during beam tracking process with optimal power allocation. We set $M_t = 64$, $L = 128$, $Q = 4$, $P_T = 5W$, $\sigma_{\omega_1}^2 = \sigma_{\omega_2}^2 = 0.1$.

We consider two cases with different initial states. In Case 1, $\mathbf{s}_1[0] = [330, 330, -20, 0]$ and $\mathbf{s}_2[0] = [30, 300, -20, 0]$, and in Case 2, $\mathbf{s}_1[0] = [110, 400, -20, 0]$ and $\mathbf{s}_2[0] = [85, 300, -20, 0]$. Fig. 14(a) and Fig. 14(c) show the learning curves in two cases respectively. Obviously, the average reward increases as the number of learning steps increases and converges at about 6×10^4 steps. This result verifies that, PCRLB constraints obtained by RL algorithm can make the system obtain better long-term performance.

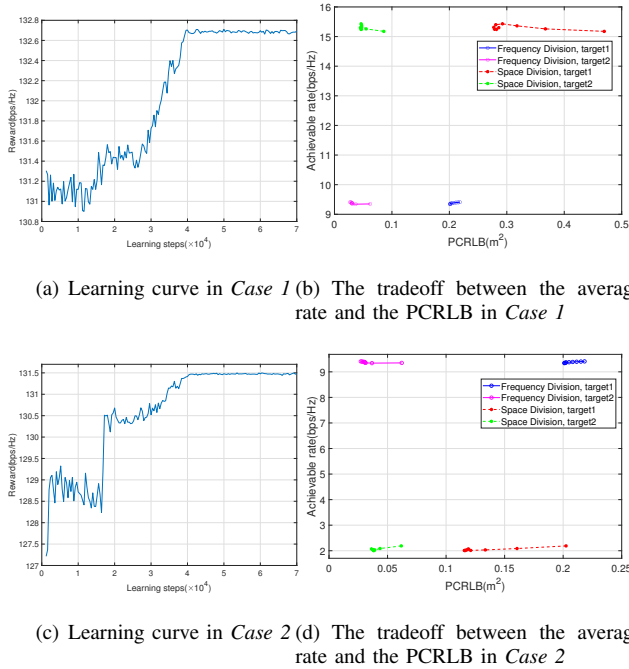


Fig. 14. The comparison of two different schemes in two cases.

Fig. 14(b) and Fig. 14(d) compare the tradeoff between the average sum rate and PCRLB of each terminal of two schemes in two cases respectively. In the first case, two terminals are far separated in AoD and in the second case, two terminals are close in AoD. In frequency division scheme, sum rate and PCRLB of each terminal during tracking are obtained by solving (48), in which $\zeta_k[t]s$ are determined by the learned policy $\pi(\cdot|s_t)$. Average sum rate and PCRLB are obtained by averaging sum rates and PCRLBs in 1500 tracking experiments. Fig. 14(b) illustrates that to maximize the sum rate, the PCRLB of terminal 1 gradually increases and the PCRLB of terminal 2 remains almost unchanged during the tracking process in the frequency division scheme. This is because terminal 2 is closer to the BSs than terminal 1, and the system tends to allocate more subcarriers to terminal 2 in order to achieve the maximum sum rate. Besides, in this case, the interference between users is small, so OFDM-SDMA can achieve higher rate than frequency division scheme. Due to low spatial separation, there is greater interference between terminals in Case 2. Fig. 14(d) shows that in the space division scheme, increased inter-terminal interference will reduce the sum rate significantly. In summary, the space division scheme has greater advantage in prediction performance since all subcarriers are used for beam tracking. However, the communication rate of the space division scheme is severely affected by interference between terminals. In contrast, when terminals' locations change, the frequency division scheme has better robustness in terms of transmission rate and prediction performance.

VII. CONCLUSION

In this paper, we propose a radar-assisted beam tracking scheme in a mmWave MIMO-OFDM system. In the closed-loop tracking process, CP uses EKF or UKF to predict

locations of terminals with echoes received by distributed receivers, and feeds back the information to the transmitting BS for optimizing transmit parameters. PCRLB of location estimation is derived as metric of estimation performance. Considering that the transmission parameters at the current time will simultaneously affect the transmission rate at the current time and the prediction of terminal locations at the next time, we propose to optimize the subcarrier/power allocation by combining RL and convex optimization to achieve optimal overall performance. Simulation results demonstrate that our proposed radar-assisted beam tracking scheme is able to obtain better tracking performance and higher transmission rate than the pilot-based scheme. Furthermore, the proposed single-stream frequency division scheme achieves better tradeoff performance than the multi-stream space division scheme when the interference between terminals cannot be ignored, for example, when the terminals are relatively close.

REFERENCES

- [1] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The next generation wireless access technology*. Academic Press, 2020.
- [2] N. Tawa, T. Kuwabara, Y. Maruta, M. Tanio, and T. Kaneko, "28 ghz downlink multi-user mimo experimental verification using 360 element digital aas for 5g massive mimo," in *2018 48th European Microwave Conference (EuMC)*, pp. 934–937, IEEE, 2018.
- [3] B. Yang, Z. Yu, J. Lan, R. Zhang, J. Zhou, and W. Hong, "Digital beamforming-based massive mimo transceiver for 5g millimeter-wave communications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 7, pp. 3403–3418, 2018.
- [4] W. Zhong, Y. Gu, Q. Zhu, L. Wang, X. Chen, and K. Mao, "A novel 3d beam training strategy for mmwave uav communications," in *2020 14th European Conference on Antennas and Propagation (EuCAP)*, pp. 1–5, 2020.
- [5] X. Sun, C. Qi, and G. Y. Li, "Beam training and allocation for multiuser millimeter wave massive mimo systems," *IEEE Transactions on Wireless Communications*, vol. 18, no. 2, pp. 1041–1053, 2019.
- [6] S. H. Lim, S. Kim, B. Shim, and J. W. Choi, "Efficient beam training and sparse channel estimation for millimeter wave communications under mobility," *IEEE Transactions on Communications*, vol. 68, no. 10, pp. 6583–6596, 2020.
- [7] Z. Xiao, T. He, P. Xia, and X.-G. Xia, "Hierarchical codebook design for beamforming training in millimeter-wave communication," *IEEE Transactions on Wireless Communications*, vol. 15, no. 5, pp. 3380–3392, 2016.
- [8] S.-G. Yoon and S. J. Lee, "Improved hierarchical codebook-based channel estimation for mmwave massive mimo systems," *IEEE Wireless Communications Letters*, vol. 11, no. 10, pp. 2095–2099, 2022.
- [9] J. Palacios, D. De Donno, and J. Widmer, "Tracking mm-wave channel dynamics: Fast beam training strategies under mobility," in *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, pp. 1–9, 2017.
- [10] C. Liu, L. Zhao, M. Li, and L. Yang, "Adaptive beam search for initial beam alignment in millimetre-wave communications," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 6, pp. 6801–6806, 2022.
- [11] X. Li, J. Fang, H. Duan, Z. Chen, and H. Li, "Fast beam alignment for millimeter wave communications: A sparse encoding and phaseless decoding approach," *IEEE Transactions on Signal Processing*, vol. 67, no. 17, pp. 4402–4417, 2019.
- [12] X. Song, S. Haghighatshoar, and G. Caire, "Efficient beam alignment for millimeter wave single-carrier systems with hybrid mimo transceivers," *IEEE Transactions on Wireless Communications*, vol. 18, no. 3, pp. 1518–1533, 2019.
- [13] D. Zhang, A. Li, M. Shirvanimoghaddam, P. Cheng, Y. Li, and B. Vucetic, "Codebook-based training beam sequence design for millimeter-wave tracking systems," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5333–5349, 2019.
- [14] S. Zhou, L. Chen, and W. Wang, "Attitude information aided mmwave beam tracking based on codebook," *IEEE Communications Letters*, vol. 26, no. 5, pp. 1160–1164, 2022.

- [15] S. Jayaprakasam, X. Ma, J. W. Choi, and S. Kim, "Robust beam-tracking for mmwave mobile communications," *IEEE Communications Letters*, vol. 21, no. 12, pp. 2654–2657, 2017.
- [16] S. Shaham, M. Ding, M. Kokshoorn, Z. Lin, S. Dang, and R. Abbas, "Fast channel estimation and beam tracking for millimeter wave vehicular communications," *IEEE Access*, vol. 7, pp. 141104–141118, 2019.
- [17] F. Liu, P. Zhao, and Z. Wang, "EKF-based beam tracking for mmwave mimo systems," *IEEE Communications Letters*, vol. 23, no. 12, pp. 2390–2393, 2019.
- [18] F. Liu, W. Yuan, C. Masouros, and J. Yuan, "Radar-assisted predictive beamforming for vehicular links: Communication served by sensing," *IEEE Transactions on Wireless Communications*, vol. 19, no. 11, pp. 7704–7719, 2020.
- [19] Y. Shen and M. Z. Win, "Fundamental limits of wideband localization—part i: A general framework," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4956–4980, 2010.
- [20] B.-n. Vo, M. Mallick, Y. Bar-Shalom, S. Coraluppi, R. Osborne, R. Mahler, and B.-t. Vo, "Multitarget tracking," *Wiley encyclopedia of electrical and electronics engineering*, no. 2015, 2015.
- [21] C. R. Berger, B. Demissie, J. Heckenbach, P. Willett, and S. Zhou, "Signal processing for passive radar using ofdm waveforms," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 1, pp. 226–238, 2010.
- [22] Y. Li, X. Wang, and Z. Ding, "Multi-target position and velocity estimation using ofdm communication signals," *IEEE Transactions on Communications*, vol. 68, no. 2, pp. 1160–1174, 2020.
- [23] C. Sturm and W. Wiesbeck, "Waveform design and signal processing aspects for fusion of wireless communications and radar sensing," *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1236–1259, 2011.
- [24] F. Colone, D. O'hagan, P. Lombardo, and C. Baker, "A multistage processing algorithm for disturbance removal and target detection in passive bistatic radar," *IEEE Transactions on aerospace and electronic systems*, vol. 45, no. 2, pp. 698–722, 2009.
- [25] S. Haykin, *Neural networks and learning machines*, 3/E. Pearson Education India, 2009.
- [26] P. Tichavsky, C. Muravchik, and A. Nehorai, "Posterior cramer-rao bounds for discrete-time nonlinear filtering," *IEEE Transactions on Signal Processing*, vol. 46, no. 5, pp. 1386–1396, 1998.
- [27] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *International conference on machine learning*, pp. 1861–1870, PMLR, 2018.
- [28] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
- [29] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, et al., "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.
- [30] T. Degris, M. White, and R. S. Sutton, "Off-policy actor-critic," *arXiv preprint arXiv:1205.4839*, 2012.



Mengting Lou is a member of technical staff in the Future Research Laboratory, China Mobile Research Institute. She received B.S. and M.S. degrees from Beijing Jiaotong University in 2015 and 2019, respectively. Her current research interests include wireless communications, integrated sensing and communication, virtual multi-input multi-output, sparse array and coordinated multi-point transmission.



Wanyun Qian received the B.S degree from School of Automation, Central South University, Hunan, China in 2020 and the M.S. degree from the school of Electronic Science and Engineering, Nanjing University, China in 2023. His research interests include multi-terminal beam tracking and resource allocation in integrated communication-radar systems.



Yechao Bai received the B.S. degree and the Ph.D. degree in electronic science and engineering from Nanjing University, Nanjing, China in 2005 and 2010, respectively. Since 2010, he has been with Nanjing University, Nanjing, China, where he is currently an associate professor with the School of Electronic Science and Engineering. His current research interests include array signal processing, target detection and parameter estimation.



Lan Tang received the PH.D degree in communication and information science from Southeast University, Nanjing, China, in 2009. Since 2009, she has joined the School of Electronic Science and Engineering, Nanjing University, China, where she is now an associate professor. In 2014, she was a visiting researcher in Columbia University, USA. Her current research interests include optimization and machine learning in wireless communications, and statistical signal processing applied for radar and communications.



Yijie Wang received the B.S degree in communication engineering from the Jilin University, Jilin, China, in 2022. He is currently pursuing the Ph.D with the Department of electronic science and engineering, Nanjing University, China. His current research interests include uav path planning, reinforcement learning, and multi-agent reinforcement learning.



Ying-Chang Liang (Fellow, IEEE) was a Professor with The University of Sydney, Australia, a Principal Scientist and a Technical Advisor with the Institute for Infocomm Research, Singapore, and a Visiting Scholar with Stanford University, USA. He is currently a Professor with the University of Electronic Science and Technology of China, China, where he leads the Center for Intelligent Networking and Communications. He is also a Deputy Director of the Artificial Intelligence Research Institute. His research interests include wireless networking and communications, cognitive radio, symbiotic networks, dynamic spectrum access, the Internet-of-Things, artificial intelligence, and machine learning techniques.