Robust Near-field Beam Tracking via Deep Q-network for THz Communications

Hyunwoo Park, Hyeonjin Chung, Andrea Conti*, Moe Z. Win†, and Sunwoo Kim

Department of Electronic Engineering, Hanyang University, Seoul, South Korea

(E-mail: {stark95, hyeonjingo, remero}@hanyang.ac.kr)

**Department of Engineering and CNIT, University of Ferrara, Via Saragat 1, 44122 Ferrara, Italy*

(E-mail: andrea.conti@unife.it)

†*Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139 USA*

(E-mail: moewin@mit.edu)

***Abstract*—This paper presents a robust near-field (NF) beam tracking algorithm for terahertz communications based on deep Q-network (DQN). Traditional NF beam tracking methods relying on mobility models are fatal in ultra-massive MIMO systems, where even the slightest error could result in beam tracking failures. Thus, the proposed algorithm aims to maintain a stable beamforming gain by tracking the mobile station through the analysis of received signals without requiring mobile dynamics. By utilizing DQN, the proposed algorithm strengthens its tracking capability from online experiences and updates the combining beam towards positions expected to maximize beamforming gain. Throughout simulations, we compare the proposed algorithm with the Bayesian filter-based NF beam tracking algorithm. The simulation results confirm the robustness of the proposed algorithm for NF beam tracking, especially for abrupt changes in mobile dynamics.**

***Index Terms*—Beam tracking, UM-MIMO, near-field, deep reinforcement learning, THz communications**

I. Introduction

As the carrier frequency continues to increase along with the generation of the mobile communications, ultra-massive MIMO (UM-MIMO) has been emerged as a solution for the severe path loss [1]. However, the problem of UM-MIMO is a large overhead for beam searching induced by high directivity of the antenna array. This makes beam tracking crucial in UM-MIMO systems since it is an effective technique for reducing large beam training overhead [2]–[4]. The beam tracking leverages the signal received from the mobile station (MS) and the previous angle information of the signal paths to estimate the beam steering angle. With this approach, the beam steering angle can be estimated without excessive beam training.

However, beam tracking in UM-MIMO systems requires considering near-field (NF) effect, which is crucial in systems with large antenna arrays [5]. When the NF effect exists, not only the angle but also the distance is needed for beam steering [5]. For this reason, the concept of NF beam tracking that tracks the beam steering angle and distance was first introduced in [6]. Subsequently, NF beam tracking algorithm based on extended Kalman filter (EKF) [7] with low complexity was introduced in [8]. Although EKF achieves low complexity through linearization, the performance degradation is inevitable for highly nonlinear systems, which is severe in NF UM-MIMO systems. The similar NF tracking research based on particle filter (PF) [9] was proposed for nonlinear UM-MIMO systems [10], but they confirmed that Bayesian filter [11] based trackings are vulnerable to mobility model mismatches. Accordingly, the performance of both EKF and PF-based beam tracking is vulnerable to errors and the randomness of the mobility model.

For robust NF beam tracking in practical mobility environments, we propose the NF beam tracking algorithm based on deep reinforcement learning to solve nonlinearity and the model dependency problem. To be more specific, we utilize deep Q-network (DQN) for NF beam tracking, which is shown to be effective for accurate and robust beam tracking in far-field conditions by its model-free characteristics [12], [13]. The proposed algorithm leverages DQN's ability to directly decide the best beam tracking action from the received signals by learning through online experiences. Here, the Q-network is trained using beamforming gain as a reward after each tracking action. With this approach, the proposed algorithm does not depend on the mobility model and enables online learning of NF beam tracking.

$Notations$: For a matrix $\mathbf{A}$, its transpose, and Hermitian transpose are respectively denoted as $\mathbf{A}^{\mathsf{T}}$ and $\mathbf{A}^{\mathsf{H}}$. The symbol $\imath$ represents the imaginary unit of complex numbers, ($\imath = \sqrt{-1}$). The notation $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the circularly symmetric complex Gaussian distribution of a random vector with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. $\mathbf{I}$ denotes the identity matrix. The notation $\mathrm{blkdiag}$ denotes the block diagonal matrix. $Re\{\cdot\}$ and $Im\{\cdot\}$ return real and imaginary parts of a complex values, respectively.

II. System Model

We consider an uplink channel beam tracking for a 6G THz system, where a single antenna mobile station (MS) directly communicates with a UM-MIMO base station (BS). The BS employs uniform linear array (ULA) antenna with M elements and array-of-subarray (AoSA) architecture with N_{RF} RF chains, where $N \gg N_{\mathrm{RF}}$. The ULA is formed with N_{RF} subarrays, where each subarray has $M = N/N_{\mathrm{RF}}$ elements.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2023R1A2C3002890)

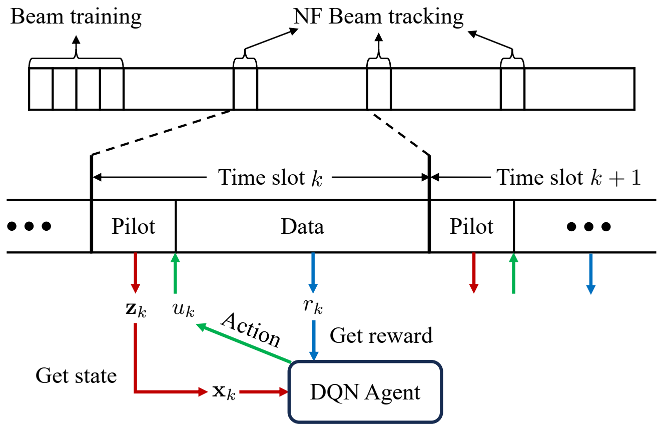

Fig. 1: NF beam tracking procedure in the time domain.

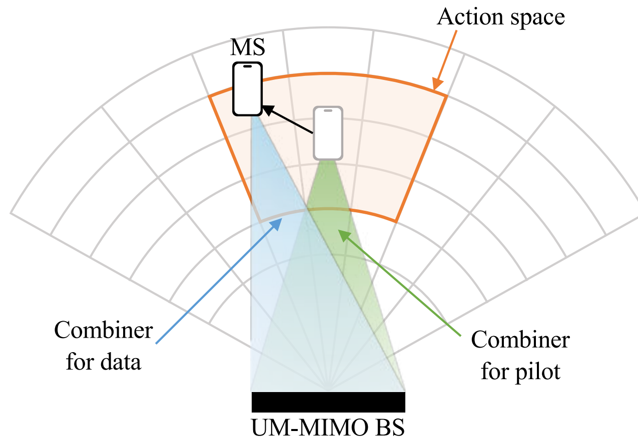

Fig. 2: Combiner and action space for NF beam tracking.

We define a beam tracking episode with K time slots, where a pilot signal transmission is performed at the beginning of each time slot.

The received signal $\mathbf{z}_k \in \mathbb{C}^{N_{\mathrm{RF}} \times 1}$ corresponding to the pilot signal b_k ($|b_k| = 1$) is expressed as

$$\mathbf{z}_k = \mathbf{W}_k^{\mathsf{H}} \mathbf{h}_k b_k + \mathbf{W}_k^{\mathsf{H}} \boldsymbol{\eta}_k, \tag{1}$$

where $\mathbf{W}_k \in \mathbb{C}^{N \times N_{\mathrm{RF}}}$ is the combining matrix, $\mathbf{h}_k \in \mathbb{C}^{N \times 1}$ is the channel vector, and $\boldsymbol{\eta}_k \sim \mathcal{CN}\left(\mathbf{0}_{N\times 1}, \sigma_\eta^2 \mathbf{I}_{N\times N}\right)$ is the additive white Gaussian noise with standard deviation σ_η. The combining matrix $\mathbf{W}_k$ follows

$$\mathbf{W}_k = \mathrm{blkdiag}\left(\mathbf{w}_{k\ 1}, \mathbf{w}_{k\ 2}, ..., \mathbf{w}_{k\ N_{\mathrm{RF}}}\right), \tag{2}$$

where $\mathbf{w}_{k\ t} \in \mathbb{C}^{M\times 1}$ is a subarray combining vector for $(t = 1, 2, ..., N_{\mathrm{RF}})$ The channel vector $\mathbf{h}_k$ between the user and the UM-MIMO BS can be expressed as

$$\mathbf{h}_k = \mathbf{a}(\mathbf{p}_k) + \sum_{l=1}^{L} \alpha_l \mathbf{a}(\dot{\mathbf{p}}_l) \in \mathbb{C}^{N\times 1}, \tag{3}$$

where $\mathbf{p}_k = [d_k \cos\theta_k, d_k \sin\theta_k]^{\mathsf{T}}$ is the MS's position in the Cartesian coordinate, L is the number of non-line-of-sight (NLoS) path, and α_l is the path gain for l-th scatterer with the position $\dot{\mathbf{p}}_l$. $\mathbf{a}(\mathbf{p}) = [a_1(\mathbf{p}), a_2(\mathbf{p}), ..., a_N(\mathbf{p})]^{\mathsf{T}} \in \mathbb{C}^{N\times 1}$ is the array response vector, and $\dot{\mathbf{p}}_l$ $(l = 2, ..., L)$ is the position of l-th scatterer. The n-th element of array response vector $a_n(\mathbf{p})$ follows [14].

$$a_n(\mathbf{p}) = \frac{\lambda}{4\pi\|\mathbf{p}'_n - \mathbf{p}\|} \exp\left(-\imath \frac{2\pi}{\lambda}\|\mathbf{p}'_n - \mathbf{p}\|\right), \tag{4}$$

where λ is a wavelength of a carrier frequency and $\mathbf{p}'_n$ is the position of the n-th element of UM-MIMO array antenna.

III. NF Beam Tracking Algorithm

The NF beam tracking algorithm follows the following procedure as in Fig. 1. We assume that the initial position of the MS is estimated through NF beam training [8]. The BS with AoSA UM-MIMO tracks the MS with the received signal, and the combiner is adjusted to the estimated MS's position. Here, we employ a DQN agent in the BS for tracking an MS without mobility model and actual position information, relying solely on the received signals during the training.

The idea of the proposed algorithm is to track the change of angle and distance by making a decision to adjust the beam steering position where high RSS is anticipated. Therefore, it is imperative to devise a combiner for pilot signal transmission, formulate a state based on the received signal, define actions for tracking, and establish a reward for training the Q-network.

A. Combiner & State Design

For the robust tracking of abrupt change of MS's direction, the combiner at k-th time slot is formed based on the estimated MS position from $(k-1)$-th time slot. Consequently, the signal received at k-th time slot, $\mathbf{z}_k$, is measured, where each element is measured with the subarray combining vector focused toward the estimated position of MS at $(k-1)$-th time slot as

$$\begin{aligned}&\mathbf{w}_{k\ t}\\&= \left[a_{M(t-1)+1}(\hat{\mathbf{p}}_{k-1}), a_{M(t-1)+2}(\hat{\mathbf{p}}_{k-1}), ..., a_{Mt}(\hat{\mathbf{p}}_{k-1})\right]^{\mathsf{T}},\end{aligned} \tag{5}$$

where $\hat{\mathbf{p}}_{k-1}$ is the estimated position of the MS at $(k-1)$-th time slot. The state serves as the input to the Q-network and comprises information that the agent can utilize for NF beam tracking. Therefore, the state $\mathbf{x}_k \in \mathbb{C}^{2N_{\mathrm{RF}}\times 1}$ is constructed using information obtained from the received signal as

$$\mathbf{x}_k = \left[Re\{\mathbf{z}_k b_k^{\mathsf{H}}\}; Im\{\mathbf{z}_k b_k^{\mathsf{H}}\}\right]. \tag{6}$$

B. Action

The action adjusts the combining beam to track the moving MS and prevent NF beam tracking failure. To perform NF beam tracking using DQN with a discrete action space, the tracking space is discretized into a 2-dimensional angle-distance grid as in Fig. 2. The goal of the action is to select the grid closest to the actual position of the MS. With the inspiration from neighboring search-based beam tracking [8], [15], the action space is constructed using neighboring grids around the previously estimated position. While the performance of

TABLE I: Action u_k for NF beam tracking.

Action u_k	Angle tracking	Distance tracking
1	$\hat{\theta}_k = \arccos\left(\cos\left(\hat{\theta}_{k-1}\right) - 1/N\right)$	$\hat{d}_k = \hat{d}_{k-1} - 2\delta_d$
2	$\hat{\theta}_k = \arccos\left(\cos\left(\hat{\theta}_{k-1}\right) - 1/N\right)$	$\hat{d}_k = \hat{d}_{k-1} - \delta_d$
$\vdots$	$\vdots$	$\vdots$
5	$\hat{\theta}_k = \arccos\left(\cos\left(\hat{\theta}_{k-1}\right) - 1/N\right)$	$\hat{d}_k = \hat{d}_{k-1} + 2\delta_d$
6	$\hat{\theta}_k = \arccos\left(\cos\left(\hat{\theta}_{k-1}\right) - 0.5/N\right)$	$\hat{d}_k = \hat{d}_{k-1} - 2\delta_d$
7	$\hat{\theta}_k = \arccos\left(\cos\left(\hat{\theta}_{k-1}\right) - 0.5/N\right)$	$\hat{d}_k = \hat{d}_{k-1} - \delta_d$
$\vdots$	$\vdots$	$\vdots$
10	$\hat{\theta}_k = \arccos\left(\cos\left(\hat{\theta}_{k-1}\right) - 0.5/N\right)$	$\hat{d}_k = \hat{d}_{k-1} + 2\delta_d$
$\vdots$	$\vdots$	$\vdots$
21	$\hat{\theta}_k = \arccos\left(\cos\left(\hat{\theta}_{k-1}\right) + 1/N\right)$	$\hat{d}_k = \hat{d}_{k-1} - 2\delta_d$
22	$\hat{\theta}_k = \arccos\left(\cos\left(\hat{\theta}_{k-1}\right) + 1/N\right)$	$\hat{d}_k = \hat{d}_{k-1} - \delta_d$
$\vdots$	$\vdots$	$\vdots$
25	$\hat{\theta}_k = \arccos\left(\cos\left(\hat{\theta}_{k-1}\right) + 1/N\right)$	$\hat{d}_k = \hat{d}_{k-1} + 2\delta_d$

NF beam tracking may vary depending on the action space, this paper focuses on the most basic form. In the proposed algorithm, the action space is defined within a 5 by 5 grid range around the previously estimated position. Each action follows the NF beam tracking procedure outlined in Table I, where $\hat{\theta}$ is the estimated angle, $\hat{d}$ is the estimated distance, and δ_d is the grid interval of the distance.

C. Reward

The reward r_k is utilized to evaluate action u_k given a state $\mathbf{x}_k$, aiding in the training of the Q-network. Since maintaining a high beamforming gain indicates a successful action, we use the beamforming gain as the reward. The beamforming gain for data can be expressed as

$$\zeta_k = \frac{\sqrt{M}\mathbf{a}(\hat{\mathbf{p}}_k)^{\mathsf{H}}\mathbf{h}_k}{\|\mathbf{a}(\hat{\mathbf{p}}_k)\|\|\mathbf{h}_k\|}. \tag{7}$$

Therefore, the reward r_k is measured as

$$r_k = \sqrt{M}\frac{|\mathbf{1}^{\mathsf{T}}\bar{\mathbf{z}}_k|}{\|\bar{\mathbf{z}}_k\|} \approx \zeta_k, \tag{8}$$

where $\mathbf{1} = [1, 1, ..., 1]^{\mathsf{T}} \in \mathbb{R}^{N_{\mathrm{RF}}\times 1}$ and $\bar{\mathbf{z}}_k$ is the received signal of the data combined with $\mathbf{W}_{k+1}$.

IV. Simulation Results and Discussions

In this section, we compare the performance of the proposed algorithm with the EKF-based NF beam tracking algorithm [8]. PF in [10] is not compared because the measurements are different since they use a digital combiner at the UM-MIMO BS.

A. Simulation Settings

The center of UM-MIMO BS is located at the origin, and the MS is moving within the area of NF with human mobility model [16], which has high uncertainty due to its command process. The carrier frequency is set as 100 GHz so that the wavelength equals 0.003 m. We consider the BS equipped with uniform linear array (ULA), which consists of 512 elements. The interval between adjacent antennas is set to half wavelength. The number of RF chains and subarrays are set to $N_{\mathrm{RF}} = 4$, where the number of subarray elements is $M = 128$. The path gain for NLoS follows $\alpha_l \sim \mathcal{CN}\left(0, 0.1^2\right)$ [17]. Simulation parameters are shown in Table II.

TABLE II: Simulation parameters

Parameter	Description	Value
N	Number of UM-MIMO antenna elements	512
N_{RF}	Number of RF chains and subarrays	4
M	Number of elements in a subarray	128
δ_d	Distance grid interval	0.01 m
L	Number NLoS path	2
K	Number of time slots in one episode	2000
-	Time slot duration	0.01 s
-	Maximum speed of MS	5 m/s
-	Discount factor	0.9
-	Learning rate	0.001
-	Size of memory	5,000
-	Target network update interval	100
-	Training interval	4
-	Mini-batch size	32
-	Number of training episode	20,000

B. Analysis on NF Beam Tracking Performance

The example of MS's trajectory and the NF Beam tracking results are shown in Fig. 3. The signal-to-noise ratio (SNR) is set to 10 dB. Two versions of EKF were compared with the proposed algorithm depending on the information of mobility model. The first version considers that the mobility model is known (indicated as EKF (known)) and the mobility model is applied to the EKF's prediction step. In the case of second version, the mobility model is not known (indicated as EKF (unknown)) and assumes that the Gaussian noise is added to the position at each time slot. When the mobility model is given, EKF shows good tracking performance at the beginning. However, it loses tracking at a point where dynamics abruptly change due to sharp curves. If the mobility model is not given, the performance of EKF deteriorates even earlier, and tracking fails almost from the beginning of the episode. In contrast, the proposed DQN-based NF beam tracking shows stable tracking until the end of the episode, although there are some grid-mismatch due to the discrete action space. The robustness of the sharp curves shows that the proposed algorithm does not rely on the information of dynamics, which is advantageous in practice, where the trajectory is unpredictable.

The beamforming gains for the same episode are shown in Fig. 4. The upper bound is the beamforming gain when perfect beam tracking is achieved, where the value is $\sqrt{M}$. Due to a very narrow beam for both the angle and distance of UM-MIMO, even a small position tracking error results in a large loss of beam gain. Although minor beam gain losses occur frequently due to grid-mismatch in the proposed algorithm, successful beam tracking is achieved with a stable

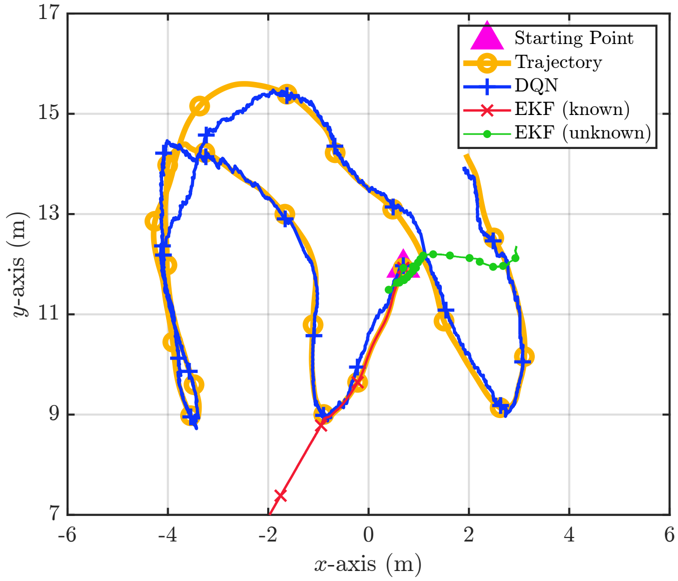

Fig. 3: NF beam tracking results of MS position.

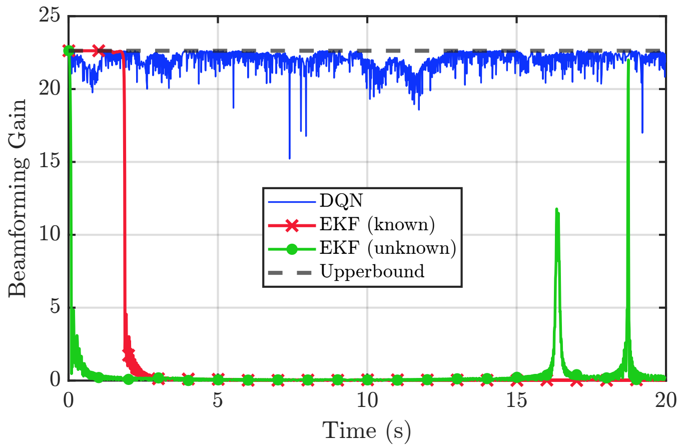

Fig. 4: Comparisons of the beamforming gains.

beamforming gain over a long period of time compared to EKF.

C. Performance Analysis of NF Beam Tracking

To investigate the position tracking performance of NF beam tracking algorithms, the RMSE of position tracking is compared in Fig. 5. The result averages over 10,000 episodes, and the SNR is set to 10 dB. As in the previous tracking results, EKF (known) generally shows accurate tracking in the beginning, but tracking performance decreases significantly over time due to the increase in uncertainty caused by sharp curves. Also, when the tracking fails, the estimated position tends to move in a straight line, resulting in even higher RMSE than EKF (unknown). In contrast, the RMSE of the proposed algorithm is less than 1.7 m for the whole 20 seconds, which implies that a low tracking RMSE can be expected under abrupt dynamics, except for the occasional tracking failure case.

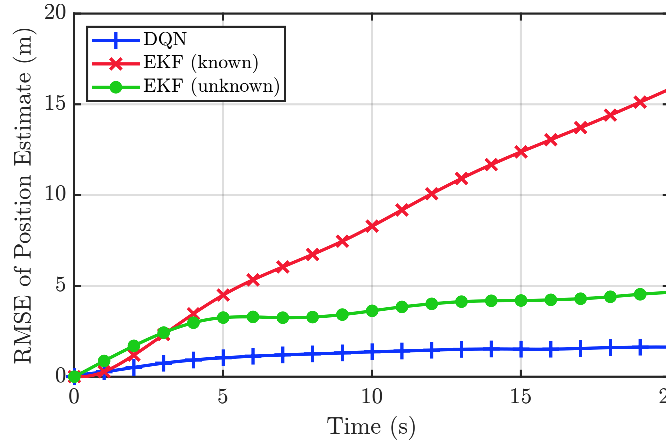

Fig. 5: RMSE of position tracking over time.

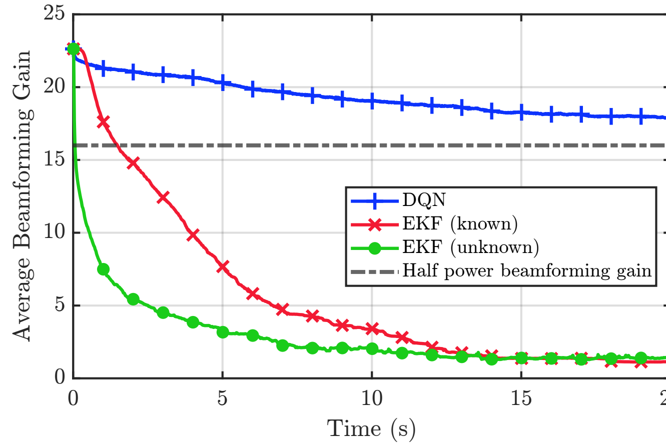

Fig. 6: Averaged beamforming gains over time.

The ultimate goal of beam tracking is to maintain beamforming gain so that beam training is not required for a long period of time. Average beamforming gains over time are compared in Fig 6. For more than 20 seconds, the average beamforming gain of the proposed algorithm exceeds 17, which is higher than the half-power beamforming gain of perfect beam tracking($=\sqrt{M/2}$). However, the average beamforming gain of EKF gets lower than 16 before 2 seconds, even with the mobility model, meaning that frequent beam training should be performed for NF mobile environments. These results show that the proposed algorithm is effective in reducing beam training overhead in practice.

V. Conclusion

This paper presents the NF beam tracking algorithm based on DQN for THz UM-MIMO communications. For the NF beam tracking which must track both angle and distance, the randomness of mobility is crucial. The proposed algorithm maintains a stable beamforming gain by tracking the mobile station through analysis of received signal data without requiring mobile dynamics. By utilizing DQN, the proposed algorithm strengthens its tracking capability from experiences and updates the combining beam towards positions expected to maximize beamforming gain. Simulation results confirmed the robustness and high beamforming gain of the proposed algorithm for NF beam tracking, especially for abrupt mobile dynamics.

REFERENCES

[1] A. Faisal, H. Sarieddeen, H. Dahrouj, T. Y. Al-Naffouri, and M.-S. Alouini, "Ultramassive MIMO systems at terahertz bands: Prospects and challenges," *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 33–42, Oct. 2020.

[2] V. Va, H. Vikalo, and R. W. Heath, "Beam tracking for mobile millimeter wave communication systems," in *Proc. IEEE Global Conf. Signal Inf. Process.*, Dec. 2016, pp. 743–747.

[3] S. Jayaprakasam, X. Ma, J. W. Choi, and S. Kim, "Robust beam-tracking for mmWave mobile communications," *IEEE Commun. Lett.*, vol. 21, no. 12, pp. 2654–2657, Dec. 2017.

[4] L. Yang and W. Zhang, "Beam tracking and optimization for UAV communications," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5367–5379, Nov. 2019.

[5] M. Cui, Z. Wu, Y. Lu, X. Wei, and L. Dai, "Near-field MIMO communications for 6G: Fundamentals, challenges, potentials, and future directions," *IEEE Commun. Mag.*, vol. 61, no. 1, pp. 40–46, Sep. 2023.

[6] C. You, Y. Zhang, C. Wu, Y. Zeng, B. Zheng, L. Chen, L. Dai, and A. L. Swindlehurst, "Near-field beam management for extremely large-scale array communications," *arXiv preprint arXiv:2306.16206*, 2023.

[7] P. Zarchan and H. Musoff, *Fundamentals of Kalman Filtering: A Practical Approach*. Amer. Inst. Aeronaut. Astronaut. (AIAA), 2005, vol. 208.

[8] K. Chen, C. Qi, C.-X. Wang, and G. Y. Li, "Beam training and tracking for extremely large-scale MIMO communications," *IEEE Trans. Wireless Commun.*, vol. 23, no. 5, pp. 5048–5062, May 2024.

[9] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House, Dec. 2003.

[10] A. Guerra, F. Guidi, D. Dardari, and P. M. Djurić, "Near-field tracking with large antenna arrays: Fundamental limits and practical algorithms," *IEEE Trans. Signal Process.*, vol. 69, pp. 5723–5738, Aug. 2021.

[11] S. Särkkä and L. Svensson, *Bayesian filtering and smoothing*. Cambridge university press, May 2023, vol. 17.

[12] H. Park, J. Kang, S. Lee, J. W. Choi, and S. Kim, "Deep Q-network based beam tracking for mobile millimeter-wave communications," *IEEE Trans. Wireless Commun.*, vol. 22, no. 2, pp. 961–971, Feb. 2023.

[13] H. Park, J. H. Jeon, H. Chung, and S. Kim, "DQN-based joint adaptive beamwidth control and beam tracking for mmWave communications," in *proc. IEEE Statistical Signal Process. W. (SSP)*, Jul. 2023, pp. 473–477.

[14] Q. C. Li, G. Wu, and T. S. Rappaport, "Channel model for millimeter-wave communications based on geometry statistics," in *proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2014, pp. 427–432.

[15] B. Ning, Z. Chen, Z. Tian, C. Han, and S. Li, "A unified 3D beam training and tracking procedure for terahertz communication," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2445–2461, Apr. 2022.

[16] Tong Liu, P. Bahl, and I. Chlamtac, "Mobility modeling, location tracking, and trajectory prediction in wireless ATM networks," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 6, pp. 922–936, Aug. 1998.

[17] C. Han, Y. Wang, Y. Li, Y. Chen, N. A. Abbasi, T. Kürner, and A. F. Molisch, "Terahertz wireless channels: A holistic survey on measurement, modeling, and analysis," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 3, pp. 1670–1707, Third Quart. 2022.