Mobility-Aware Near-Field Beam Training in XL-MIMO Systems for High-Speed Railway

Wenhui Yi, Jiayi Zhang, *Senior Member, IEEE*, Zhe Wang, Huahua Xiao, and Bo Ai, *Fellow, IEEE*

Abstract—Near-field beam training is beneficial to avoid intricate direct channel estimation for extremely large-scale multiple-input-multiple-output (XL-MIMO) high-speed railway (HSR) systems. In this paper, we propose a mobility-aware near-field beam training (MNBT) scheme for the XL-MIMO system with orthogonal frequency-division multiplexing (OFDM) based on the periodicity and regularity of the train trajectory. The MNBT scheme can predict the beam direction and channel amplitude by utilizing the beam parameters obtained from the near-field graded beam training a few times initially, incorporating prior information of the HSR scenario such as the deterministic trajectory, motion model, and channel model. Based on the train's speed, we develop an inter-carrier interference (ICI) suppression algorithm to mitigate the impact of Doppler frequency offset (DFO) induced by high-speed movement. Furthermore, to prevent sudden degradation of beamforming performance due to error accumulation, we establish an error detection module that identifies and corrects prediction errors in real time based on the concavity of the beam focusing function concerning the time variable. Simulation results demonstrate that, compared to existing beam training schemes, the proposed MNBT scheme significantly reduces beam training overhead while maintaining competitive beamforming performance in HSR scenarios.

Index Terms—Extremely large-scale MIMO-OFDM, HSR networks, near-field, beam training.

I. Introduction

WITH the rapid development of artificial intelligence, big data, cloud computing, and fifth-generation communication, there are higher demands on railway communication [1]. Among these, extremely large-scale multiple-input multiple-output (XL-MIMO) technology shows promise for enhancing data transmission efficiency, thereby enhancing the reliability of train operations and passenger experience [2]. However, the deployment of large-scale antenna arrays in HSR environments poses challenges for channel estimation. These challenges can be primarily categorized into two aspects: first, the rapid fluctuations in channel state information (CSI) render

conventional estimation methods insufficient for real-time tracking of channel dynamics. Second, the exponential growth in dimension of the channel matrix associated with XL-MIMO elevates both computational complexity and pilot signal overhead, thereby compounding the difficulty of achieving precise channel estimation [3], [4].

Beam training is often employed to efficiently detect the CSI in XL-MIMO systems instead of channel estimation, thereby identifying the beam combination with the maximum received signal energy [5]. Specifically, [9] introduces a novel fast two-stage beam training method, significantly reducing the training cost of exhaustive search while maintaining beamforming performance. In [8], leveraging spatial non-stationarity, a contrast learning beam training method tailored to specific environments is devised to achieve higher beam gain with fewer code words. Further, in [7], provides design a fast wideband beam training scheme based on the near-field rainbow to suppress beam splitting at high frequencies. However, the existing beam training methods in XL-MIMO systems for HSR scenarios face challenges: the rapid time-variation of CSI necessitates real-time beam tracking, enforcing shorter training periods and thereby leading to excessive overheads.

In this letter, we propose the mobility-aware near-field beam training (MNBT) scheme for the downlink XL-MIMO HSR communication systems with orthogonal frequency-division multiplexing (OFDM). First, we construct a three-dimensional (3D) system model that considers a near-field channel model based on XL-MIMO-OFDM and the Doppler frequency offset (DFO) caused by the high-speed movement of the train. Then we describe the framework of the proposed MNBT algorithm, which calculates the train operation parameters with a few times of beam training compared with the conventional algorithm. It aims to predict the train operation status at each subsequent time point and constructs the inter-carrier interference (ICI) suppression matrix based on predictive speed to mitigate the impact of DFO. Finally, the simulation results indicate that the proposed MNBT method notably decreases training overhead during train operations while also achieving significant beam-focusing enhancement for data transmission in HSR scenarios.

II. System Model

This letter considers the downlink XL-MIMO-OFDM communication system based on the uniform planar array (UPA) shown in Fig. 1, where the majority of base stations (BS) are deployed on the track's sides to create cellular coverage [10]. We consider a transmitting BS located in the x-y plane with its center at the origin, consisting of $M = M_H \times M_V$ discrete point antennas. M_H and M_V are the number of transmitting antennas per row and column, respectively. The horizontal spacing and the vertical spacing between antennas
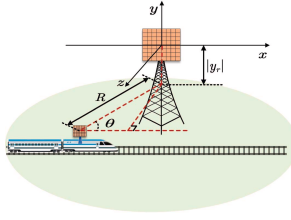
Fig. 1. XL-MIMO-enabled HSR communication systems.

are denoted as d. The row and column index of the m-th antenna on the transmitter can be represented as (m_1, n_1), where $m_1 = \mathrm{mod}(m-1, M_H)$ and $n_1 = \lfloor (m-1)/M_H \rfloor$ are the horizontal and vertical indices, respectively, with $m \in \{1, \ldots, M\}$. In 3D coordinate space, the position of the m-th antenna on the transmitter is $\mathbf{r}_{t,m} = [x_{m_1}, y_{n_1}, 0]^T$, with $x_{m_1} = (m_1 - \frac{M_H-1}{2})d$ and $y_{n_1} = (n_1 - \frac{M_V-1}{2})d$ expressing the coordinates in the x-direction and y-direction, respectively.

Meanwhile, we consider a receiver on the mobile train as a relay station to communicate with the BS. The numbers of receiving antennas per row and per column are denoted as N_H and N_V, respectively. The receiver equipped with $N = N_H \times N_V$ discrete point antennas moves along the x-axis with a velocity denoted as v and the velocity is denoted as $\mathbf{v} = [v, 0, 0]^T$. The row and column index of the n-th antenna on the receiver can be represented as (m_2, n_2), where $m_2 = \mathrm{mod}(n-1, N_H)$ and $n_2 = \lfloor (n-1)/N_H \rfloor$ are the horizontal and vertical indices, respectively, with $n \in \{1, \ldots, N\}$.

When the center of the receiver moves to $[x_r, y_r, z_r]^T$, the position of the n-th antenna on the receiver is $\mathbf{r}_{r,n} = [x_{m_2} + x_r, y_{n_2} + y_r, z_r]^T$, with $x_{m_2} = (m_2 - \frac{N_H-1}{2})d$ and $y_{n_2} = (n_2 - \frac{N_V-1}{2})d$ expressing the relative positions in the x-direction and y-direction on the receive, respectively. As shown in Fig. 1, we can also express the position of the receiver as (θ, R), where $\theta = \arctan \frac{z_r}{-x_r}$ and $R = \sqrt{x_r^2 + z_r^2}$. Here, the distance between the centers of the transmitter and the receiver is denoted as $D = \sqrt{R^2 + y_r^2}$. The distance between the m-th transmitting antenna and the n-th receiving antenna is $|\mathbf{r}_n - \mathbf{r}_m| = r(m, n, \theta, R) = \sqrt{((x_{m_1} - x_r - x_{m_2})^2 + (y_{n_1} - y_r - y_{n_2})^2 + z_r^2)}$, and it can be also expressed as (1), shown at the bottom of the page, where the approximation (a) is the Taylor expansion $\sqrt{1+x} \approx 1 + \frac{x}{2}$. When $R \to \infty$, we can obtain $D \to \infty$ and $r(m, n, \theta, R) \to D$.

Considering the XL-MIMO-OFDM channel in the HSR scenario, there exists ICI between adjacent subcarriers. We define the light speed, the central carrier frequency, the central wavenumber, the bandwidth, the number of total subcarriers, the duration of one OFDM symbol, the coherence bandwidth and the coherence time as c, f_c, $k_c = \frac{2\pi f_c}{c}$, B, L_{total}, T_s, B_c and T_c, respectively. According to [11], we can analyze the performance by studying a single statistically representative coherence block. Considering a coherence block with L subcarriers, there are ICI effects among them. Given that the probability of communication relying on the line-of-sight (LoS) path between the transmitter and receiver

is significantly higher than that of non-LoS paths in HSR scenarios [3], this letter focuses solely on the LoS channel. The channel matrix $\mathbf{H} \in \mathbb{C}^{LM \times LN}$ is generated based on the scalar Green's function between the BS and the relay station, where the (m, n)-th element $[\mathbf{H}]_{m,n} = \mathbf{H}_{mn} \in \mathbb{C}^{L \times L}$ is the channel matrix between the m-th transmitting antennas and the n-th receiving antennas with L subcarriers. With $l \in \{1, 2, \ldots, L\}$ and $s \in \{1, 2, \ldots, L\}$ expressing the subcarrier index, the (l, s)-th element of $\mathbf{H}_{mn}$ can be denoted by h_{ls}.

And h_{ls} expressed as (2), shown at the bottom of the page is the channel between the m-th transmitting antenna and the n-th receiving antenna affected by the ICI coefficients between the l-th and the s-th subcarrier [11], with $k_l = \frac{2\pi f_l}{c}$ being the wavenumber at subcarrier frequency $f_l = f_c + \frac{2l-L-1}{2T_s}$, and $\varepsilon_{nm} = \frac{-fT_s}{c} \frac{(\mathbf{r}_{r,n} - \mathbf{r}_{t,m}) \cdot \mathbf{v}}{r(m,n,\theta,R)}$ is the normalized DFO between the m-th transmitting antenna and the n-th receiving antenna. Here, we denote the average normalized DFO as $\varepsilon = \mathbb{E}\{\varepsilon_{nm}\}$. When D is larger than the array aperture, the channel matrix $\mathbf{H}$ can also expressed as $\mathbf{H} = \sqrt{MNL}\beta \mathbf{A} \mathbf{C}$, where $\beta = \frac{\exp(-j2\pi k_c D)}{4\pi D}$ is the complex path gain, $\mathbf{A} \in \mathbb{C}^{LM \times LN}$ is the near-field array response matrix with the (m, n)-th element $[\mathbf{A}]_{m,n} = \mathbf{A}_{mn} = \frac{1}{\sqrt{MNL}} \mathrm{diag}(e^{-jk_1(r(m,n,\theta,R)-D)}, \ldots, e^{-jk_L(r(m,n,\theta,R)-D)}) \in \mathbb{C}^{L \times L}$, and $\mathbf{C} = \mathrm{diag}(\mathbf{K}_1, \mathbf{K}_2, \ldots, \mathbf{K}_N) \in \mathbb{C}^{LN \times LN}$ is the ICI coefficients matrix with $\mathbf{K}_1 = \cdots = \mathbf{K}_N = \mathbf{K} \in \mathbb{C}^{L \times L}$ and $[\mathbf{K}]_{ls} = \frac{\sin(\pi(l-s+\varepsilon))}{L\sin(\frac{\pi}{L}(l-s+\varepsilon))} \exp(j\pi(1-\frac{1}{L})(l-s+\varepsilon))$.

Based on the XL-MIMO-enabled HSR channel in (2), the received signal $\mathbf{y} \in \mathbb{C}^{LN \times 1}$ at the receiver is given by

$$\mathbf{y} = \mathbf{H}^T \mathbf{W} \mathbf{s} + \mathbf{n}, \qquad (3)$$

where $\mathbf{W} \in \mathbb{C}^{LM \times LN}$ represents the beam focusing vector between the m-th transmitting antenna and the n-th receiving antenna, which is essentially a codeword, $\mathbf{s} \in \mathbb{C}^{LN}$ represents the transmitted pilot satisfying $\mathbb{E}\{\mathbf{s}\mathbf{s}^H\} = (P_t/LN)\mathbf{I}_{LN}$, P_t represents the downlink transmit power, $\mathbf{n} \in \mathbb{C}^{LN \times 1}$ is the additive white Gaussian noise (AWGN), and σ^2 is the noise power, which can be determined by considering the interplay among temperature, bandwidth, and the noise coefficient.

III. PREDICTIVE MNBT SCHEME

A. Near-Field Codewords Design

Near-field beam training necessitates beam search in both angular and distance domains, a departure from far-field beam training methods. To this end, a polar-domain near-field codebook is generated using frequency-independent phase shifts, as described in equation [14]

$$\mathcal{W} = \left\{ \mathbf{W}(\tilde{\theta}_{pq}, \tilde{R}_{pq}) \in \mathbb{C}^{LM \times LN} | p \in \mathcal{P}, q \in \mathcal{Q} \right\}, \quad (4)$$

where the (m, n)-th element of $\mathbf{W}(\tilde{\theta}_{pq}, \tilde{R}_{pq})$ is $\mathbf{W}_{mn}(\tilde{\theta}_{pq}, \tilde{R}_{pq}) = \frac{1}{\sqrt{MNL}} e^{jk_c(r(m,n,\tilde{\theta}_{pq},\tilde{R}_{pq})-D)}\mathbf{I}_L \in \mathbb{C}^{L \times L}$, $\mathcal{P} = \{1, \ldots, P\}$, $\mathcal{Q} = \{1, \ldots, Q\}$ with $\tilde{\theta}_{pq} = \arctan \frac{z_q}{-x_p}$

$$r(m, n, \theta, R) \overset{(a)}{\approx} D + \frac{R}{D}\cos\theta(x_{m_1} - x_{m_2}) + \frac{|y_r|}{D}(y_{n_1} - y_{n_2}) + \frac{(x_{m_1} - x_{m_2})^2 + (y_{n_1} - y_{n_2})^2}{2D}, \qquad (1)$$

$$h_{ls} = \frac{\exp(-jk_l r(m,n,\theta,R))}{4\pi r(m,n,\theta,R)} \frac{\sin(\pi(l-s+\varepsilon_{mn}))}{L\sin(\frac{\pi}{L}(l-s+\varepsilon_{mn}))} \exp\left(j\pi(1-\frac{1}{L})(l-s+\varepsilon_{mn})\right), \qquad (2)$$
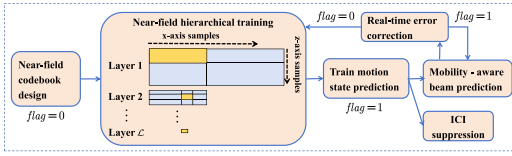
Fig. 2.   Proposed MNBT scheme.

and $\tilde{R}_{pq} = \sqrt{x_p^2 + z_q^2}$ denote the pointing sample angle and the pointing sample distance, respectively. x_p and z_q represent the pointing sample on the x-axis and z-axis in the x-z plane within the coverage area with $x_p = \frac{p(x_{max}-x_{min})}{P}$ and $z_q = \frac{q(z_{max}-z_{min})}{Q}$, respectively.

Theorem 1: To evaluate the effectiveness of the codeword, we denote $g(\mathbf{W}(\tilde{\theta}_{pq}, \tilde{R}_{pq}), R, \theta)$ as the array gain when the receiver is located at (θ, R). Considering $g(\mathbf{W}(\tilde{\theta}_{pq}, \tilde{R}_{pq}), R, \theta) = \|\mathbf{A}^T \mathbf{W}(\tilde{\theta}_{pq}, \tilde{R}_{pq})\|_F$, the estimated physical location selected in sample location can be expressed as

$$\left(\hat{\theta}, \hat{R}\right) = \arg \max_{\left(\tilde{\theta}_{pq}, \tilde{R}_{pq}\right)} \left(g(\mathbf{W}(\tilde{\theta}_{pq}, \tilde{R}_{pq}), R, \theta) \right). \quad (5)$$

After that, we can generate a codeword $\mathbf{W}(\hat{\theta}, \hat{R})$ based on (4) with the location $(\hat{\theta}, \hat{R}) = (\theta, R)$ to serve the receiver and achieve a optimal array gain.

Proof: We consider a single subcarrier system with $L = 1$ and $k_1 = k_c$, which is not affected by ICI, simplifying calculations. The array gain achieved by $\mathbf{W}(\hat{\theta}, \hat{R})$ on an arbitrary receiver location (θ, R) is given by (6), at the bottom of the page, where $n_a \in \{1, \ldots, N\}$ and $n_b \in \{1, \ldots, N\}$ are the receiving antenna indexes. Since $|g(\mathbf{W}(\hat{\theta}, \hat{R}), R, \theta)| \leq 1$, it is obvious that $(\hat{\theta}, \hat{R}) = (\theta, R)$ is an optimal solution to maximize the array gain. For a multicarrier near-field system, the dimensions of the channel matrix and codeword matrix are expanded, and we consider that channels of different subcarriers correspond to the same codeword. Using the same steps above, we prove $(\hat{\theta}, \hat{R}) = (\theta, R)$. ∎

B. Mobility-Aware Near-Field Beam Training

Compared to the classical far-field beam training that determines a spatial angle, near-field beam training desires to obtain the receiver's location information (θ, R). Among the existing near-field beam training methods, exhaustive and hierarchical training are the most commonly used [7]. However, in HSR scenarios, CSI changes rapidly, necessitating continuous training when using traditional beam training schemes to locate the estimated physical location of maximum user-received power and obtain appropriate codewords. Excessive training iterations during this process can lead to unacceptable training overhead. To reduce the overhead, we propose a predictive MNBT scheme shown as Fig. 2.

Due to the stable speed and small curvature of the travel trajectory during train movement, we can consider the process of the train passing through a BS coverage area as uniform linear motion. Based on this, we can estimate the position information of the train at each interval with minimal training

Algorithm 1 Proposed MNBT Algorithm

Input: The time range when the train passes through the coverage area of the BS $[0, T]$, the time interval between each training session Δt, sampling range on the x-axis $[x_{\min}, x_{\max}]$, sampling range on the z-axis $[z_{\min}, z_{\max}]$, the layers number of codebook $\mathcal{L}$, layer sample vector of x-axis $\left[N_x^1, ..., N_x^{\mathcal{L}}\right]$, layer sample vector of z-axis $\left[N_z^1, ..., N_z^{\mathcal{L}}\right]$, training judgment value $flag = 0$.
1: **for** $t = 0, \Delta t, ... T$ **do**
2:    **if** $t > \Delta t$ **then**
3:      $flag = 1$, calculate the position $(\hat{\theta}, \hat{R})$ based on (9) and (10)
4:      **if** $(\hat{\theta}, \hat{R})$ doesn't satisfy (8) **then**
5:        $flag = 0$
6:      **end if**
7:    **end if**
8:    **if** $flag = 0$ **then**
9:      **for** $l = 1, 2, ...\mathcal{L}$ **do**
10:        Obtain $P = N_x^l$, $Q = N_z^l$, calculate the near-field codebook as (4), search the perfect position $(\hat{\theta}, \hat{R})$ to satisfy as (7), update the coverage range on the x-axis and z-axis according to $(\hat{\theta}, \hat{R})$
11:      **end for**
12:    **end if**
13:    $\theta_t = \hat{\theta}, R_t = \hat{R}$, obtain the best codeword $\mathbf{W}(\theta_t, R_t)$ at this moment as (4).
14: **end for**
Output: The best codewords correspond to each moment of train movement.

times [16]. Algorithm 1 gives the specific procedure of the predictive MNBT algorithm.

Firstly, when the train enters the coverage area of the BS, the BS conducts beam training at a time interval of Δt to gain the maximum downlink beam focusing gain as

$$\left(\hat{\theta}, \hat{R}\right) = \arg \max_{\left(\tilde{\theta}_{pq}, \tilde{R}_{pq}\right)} \|\mathbf{y}\|_F^2. \quad (7)$$

Based on (3) and (7), we can compute the convexity property of the function of beam training gain against time t under the condition of obtaining the optimal beam when the train is moving at a constant speed in a straight line.

Theorem 2: Considering that $\mathbf{y}_t = \mathbf{H}_t^T \mathbf{W}(\theta_t, R_t)\mathbf{s} + \mathbf{n}$ is the received signal at time t, and $\mathbf{H}_t$ is the channel matrix at time t, the function of beam focusing gain and time can be analyzed as a convex function, while the train passes the BS at high speed. So the focusing gain of t-time beam satisfies the expression as follows

$$\|\mathbf{y}_t\|_F^2 + \|\mathbf{y}_{t-2\Delta t}\|_F^2 \leq 2\|\mathbf{y}_{t-\Delta t}\|_F^2. \quad (8)$$

Proof: The proof is provided in the Appendix. ∎

We adopt the near-field hierarchical beam training method [14], which divides the codebook into $\mathcal{L}$ layers. Each layer is assigned a sample number, and the best codebook interval is determined through training. Subsequently, the codebook is adjusted to this interval before moving on to train the next layer. The process concludes upon training all $\mathcal{L}$ layers and obtaining the best codeword.

Then, we use the channel parameters obtained from the first two beam training to predict the motion state of the train after that, which is helpful to reduce the huge overhead caused by frequent beam training. We define the receiver position at time $t \in \{0, \Delta t, \ldots, T\}$ as (θ_t, R_t), so we can get the predicted

$$g\left(\mathbf{W}\left(\hat{\theta}, \hat{R}\right), R, \theta\right) = \frac{1}{MNL} \sqrt{\sum_{n_a=1}^{N} \sum_{n_b=1}^{N} \left(\sum_{m=1}^{M} e^{jk_c r\left(m, n_{column}, \hat{\theta}, \hat{R}\right) - jk_1 r(m, n_{row}, \theta, R)} \right)^2}, \quad (6)$$

result of the receiver speed expressed as

$$v = \frac{R_0 \cos\theta_0 - R_{\Delta t}\cos\theta_{\Delta t}}{\Delta t}, \quad (9)$$

and the position at time t is given by

$$(\theta_t, R_t) = \left(\text{atan}\left(\frac{R_0 \sin\theta_0}{R_0 \cos\theta_0 - vt} \right), \sqrt{(vt)^2 - 2vtR_0\cos\theta_0 + R_0^2} \right), \quad (10)$$

so we can obtain the codeword as (4). To avoid the error of beam prediction leading to the sudden decline of signal quality at a certain position, we use the concavity of the beam focusing gain function to monitor the error. If the predicted code word does not satisfy (8), the codeword needs to be retrained as (7).

Finally, the ICI interference suppression matrix $\mathbf{C}^H$ can be obtained by calculating the train status information. Adding the interference suppression module to the transmitter makes it easy to effectively suppress the ICI interference caused by high-speed movement without exchanging the train status information at the receiver and transmitter, which is mathematically shown as

$$\mathbf{y} = \mathbf{C}^H \mathbf{H}^T \mathbf{W}\mathbf{s} + \mathbf{n}. \quad (11)$$

This proposed beam training algorithm is theoretically extendable to multi-user scenarios [17], though it requires adaptive modifications tailored to multi-user channel characteristics and system requirements, specifically incorporating inter-user interference mitigation and multi-user scheduling mechanisms.

IV. SIMULATION RESULTS AND DISCUSSIONS

In this section, we provide the simulations to demonstrate the performance of the predictive MNBT scheme for the HSR XL-MIMO-enabled system. We consider a downlink XL-MIMO system where both the transmitter and receiver equip a $M = N = 400$-element UPA with $M_H = M_V = 20$ and $N_H = N_V = 20$. Besides, we define downlink transmitter power $P_t = 20\,\text{dBm}$, the communication bandwidth $B = 20\,\text{MHz}$, carrier frequency $f_c = 2.1\,\text{GHz}$, wavelength $\lambda_c = \frac{c}{f_c}$, sample duration $T_s = 67\,\text{us}$, noise figure $N_F = 9\,\text{dB}$, the coverage radius of one BS $R = 2000\,\text{m}$, and the relative height of the transmitter and receiver $|y_r| = 8\,\text{m}$. It is defined that the UPA we used is a rectangle with a side length of $20\lambda_c$.

To quantity the beam training performance, we use the array gain shown as (5), the beam focusing gain shown as (7), and the average data rate defined [11]

$$Rate = \frac{1}{L}\sum_{l=1}^{L} \log_2\left(1 + \frac{P_{\text{desired}}}{\sigma^2 + P_{\text{int}}}\right), \quad (12)$$

where P_{desired} is the desired signal power and P_{int} is the ICI signal power. We compare the MNBT scheme with the perfect CSI [7], near-field hierarchical beam training scheme [14], far-field hierarchical beam training scheme [13] and exhaustive search scheme [15] to verify the proposed scheme can guarantee performance with a significant reduction in training overhead.

First of all, we compare the required overheads for different beam training schemes. We define the number of angle and distance samples as $U = 30$ and $S = 128$, respectively. Therefore, the overheads of the near-field and far-field exhaustive beam training schemes when the train runs within the coverage area of the BS for time t are $\frac{t}{\Delta t}US = 3840\frac{t}{\Delta t}$ and $\frac{t}{\Delta t}U = 30\frac{t}{\Delta t}$, respectively. We set the codebooks with $K_1 = 2$ levels for near-field hierarchical beam training and $\mathcal{L} = K_1$. Different

TABLE I
COMPARISON OF BEAM TRAINING OVERHEAD

Scheme	Overhead	Value
Near-field hierarchical scheme	$\frac{t}{\Delta t}\sum_{k=1}^{K_1} N_x^k N_z^k$	$180\frac{t}{\Delta t}$
Far-field hierarchical scheme	$K_2 K_3 \frac{t}{\Delta t}$	$24\frac{t}{\Delta t}$
Near-field exhaustive search scheme	$\frac{t}{\Delta t}US$	$3840\frac{t}{\Delta t}$
Far-field exhaustive search scheme	$\frac{t}{\Delta t}U$	$30\frac{t}{\Delta t}$
Proposed MNBT scheme $(F \ll \frac{t}{\Delta t})$	$\sum_{k=1}^{K_1} N_x^k N_z^k F$	$180F$

(a) Perfect CSI [9]   (b) Proposed MNBT

(c) Near-field hierarchi- (d) Far-field hierarchi-
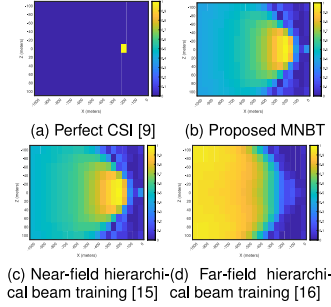cal beam training [15] cal beam training [16]

Fig. 3. Array gain for XL-MIMO-enabled HSR system against different beam training schemes.

levels of sub-codebooks have different sampling ranges and scales. In our simulations, we set $N_x^k = 30$ and $N_z^k = 3$ for $k = 1, 2$. So the overhead at t time is $\frac{t}{\Delta t}\sum_{k=1}^{K_1} N_x^k N_z^k = 180\frac{t}{\Delta t}$. Similarly, for the far-field hierarchical beam training, we set the level number of the angle space as $K_2 = 3$ and there are $K_3 = 8$ grids in every level, which gives rise to an overhead of $K_2 K_3 \frac{t}{\Delta t} = 24\frac{t}{\Delta t}$. For our predictive MNBT method, we denote the number of the near-field hierarchical training as F, which consists of starting two prediction training sessions and error correction training sessions, much smaller than the regular number of training sessions. Therefore, the overall overhead is $\sum_{k=1}^{K_1} N_x^k N_z^k F = 180F$.

Fig. 3 shows the array gain against different beam training schemes in XL-MIMO-enabled HSR systems with the speed of the train $v = 350\,\text{km/h}$ when the center of the receiver is at $[-200, y_r, z_r]^T$. In these heat maps, the brighter the color, the greater the array gain at this position. To show the beam pattern more clearly, it is worth noting that we utilize the rectangular coordinate system to present the array gains of the locations in three-dimensional space. Fig. 3(a) presents an ideal beam pattern, where the beam should focus on the target location. Fig. 3(b), Fig. 3(c) and Fig. 3(d) present the beam patterns after the predictive MNBT, the near-field hierarchical training, and the far-field hierarchical, respectively. It is shown that the array gain of the near-field training is much higher than the far-field training. The predictive MNBT method is only trained when the train enters the coverage area, and its performance with $F = 2$ at this position is close to that of conventional near-field hierarchical training. Therefore, our proposed algorithm can ensure satisfactory beam focusing gain while reducing overhead.

Fig. 4 shows the beam focusing gain against different positions relative to the target BS during train operation with different beam training schemes with $v = 350\,\text{km/h}$. For this XL-MIMO-enabled system, the near-field boundary is $2(\sqrt{2}M_H d + \sqrt{2}N_H d)^2 f_c/c = 1008\,\text{m}$, which is the Rayleigh distance. It can be seen that when the train runs to the near-field region, the near-field beam training methods considering the two dimensions of distance and angle are closer to the perfect
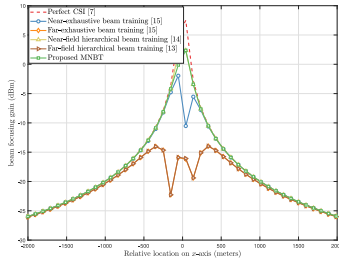
Fig. 4. Beam focusing gain during train movement under different beam training schemes.
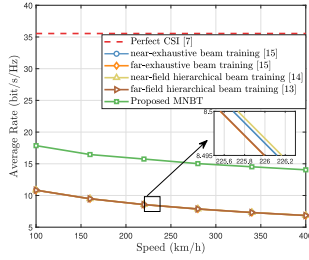

Fig. 5. The average rate for XL-MIMO-enabled HSR system against different speeds of the train with the $L = 8$ subcarriers in a coherence.

CSI than the far-field beam training methods considering only the angle region. Under the limited overhead, the performance of near-field hierarchical beam training is better than exhaustive beam training. The performance of the predictive MNBT method is close to that of the conventional near-field hierarchical beam training method, and its overhead is much lower than other beam training methods with $F = 5$ and $t = 41$ s in the whole process.

Fig. 5 shows the average rate against different train speeds in the XL-MIMO-OFDM systems with different beam training schemes when the receiver's center is at $[-200, y_r, z_r]^T$. The average rate decreases with the increase of the train speed, due to ICI caused by the DFO effect. In addition, our predictive MNBT method has a very obvious inhibition on ICI. Compared with the traditional beam training methods, the average rate is improved by about 63.6 % with the predictive MNBT scheme, which is closer to the perfect CSI without interference. It is shown that we can improve the system performance effectively by detecting train running state parameters and setting the ICI suppression module.

V. CONCLUSION

In this letter, we introduce a predictive MNBT scheme for the downlink XL-MIMO-OFDM HSR communication system. The proposed algorithm enables the calculation of parameters related to the train's operational state and incorporates an interference suppression module. Simulation results demonstrate that, compared to the classical beam training structure, the predictive MNBT scheme markedly reduces training overhead while ensuring communication quality.

APPENDIX: PROOF OF (8)

We consider a single subcarrier system with $L = 1$ and $k_1 = k_c$, which is not affected by ICI, simplifying calculations. The beam focusing gain function at time t is given by

$$\|\mathbf{y}_t\|_F^2 = \left\| \sqrt{MNL}\beta_t \mathbf{A}_t^T \mathbf{W}(\theta_t, R_t)\mathbf{s} + \mathbf{n} \right\|_F^2$$

$$\overset{(a)(b)}{\approx} P_t M^2 N^2 \frac{1}{16\pi^2\left(R_t^2 + y_r^2\right)}, \tag{12}$$

where β_t represents the complex path gain and $\mathbf{A}_t$ represents the near-field array response matrix. The approximation (a) is that AWGN can be ignored due to its time-invariant nature. The approximation of (b) is that the antenna gain is maximized as $\mathbf{A}_t^T \mathbf{W}(\theta_t, R_t) \to \mathbf{I}_N$ under the optimal codeword.

While the position of the receiver's center is $[x_r, y_r, z_r]^T$, we take the second derivative of (12) as follows

$$\frac{\partial^2 \|\mathbf{y}_t\|_F^2}{\partial t^2} = \frac{\partial}{\partial t}\left(\frac{-2tv^2 - 2R_0 v \cos\theta_0}{16\pi^2(R_t^2 + y_t^2)^2}\right)$$

$$= \frac{v^2\left(3x_r^2 - y_r^2 - z_r^2\right)}{8\pi^2\left(R_t^2 + y_r^2\right)^3}. \tag{13}$$

When $x_r \in \left(-\infty, -\sqrt{\frac{y_r^2 + z_r^2}{3}}\right) \cup \left(\sqrt{\frac{y_r^2 + z_r^2}{3}}, +\infty\right)$, $\frac{\partial^2 \|\mathbf{y}_t\|_F^2}{\partial t^2} \geq 0$ and (12) is a convex function. Because in the high-speed mobile environment, the probability of $x_r \in \left(-\sqrt{\frac{y_r^2 + z_r^2}{3}}, \sqrt{\frac{y_r^2 + z_r^2}{3}}\right)$ is very small, for example, where the coverage radius is 2000 m, $|y_r|$ is 8 m, and $|z_r|$ is 3m, the probability of $x_r \in \left(-\sqrt{\frac{y_r^2 + z_r^2}{3}}, \sqrt{\frac{y_r^2 + z_r^2}{3}}\right)$ is only 0.25%. Therefore, (12) can be directly analyzed as a convex function.

REFERENCES

[1] J. Zhang et al., "Prospective multiple antenna technologies for beyond 5G," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1637–1660, Aug. 2020.
[2] W. Yi et al., "Performance analysis of extremely large-scale MIMO systems for high-speed railway," *IEEE Trans. Veh. Technol.*, to be published.
[3] B. Ai et al., "5G key technologies for smart railways," *Proc. IEEE*, vol. 108, no. 6, pp. 856–893, Jun. 2024.
[4] D. He et al., "Channel measurement, simulation, and analysis for high-speed railway communications in 5G millimeter-wave band," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 10, pp. 3144–3158, Oct. 2018.
[5] K. Chen et al., "Beam training and tracking for extremely large-scale MIMO communications," *IEEE Trans. Wireless Commun.*, vol. 23, no. 5, pp. 5048–5062, May 2024.
[6] Z. Wang et al., "A tutorial on extremely large-scale MIMO for 6G: Fundamentals, signal processing, and applications," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 3, pp. 1560–1605, 3rd Quart., 2024.
[7] M. Cui et al. "Near-field rainbow: Wideband beam training for XL-MIMO," *IEEE Trans. Wireless Commun.*, vol. 22, no. 6, pp. 3899–3912, Jun. 2023.
[8] X. Zhang et al., "Environment-specific beam training for extremely large-scale MIMO systems via contrastive learning," *IEEE Commun. Lett.*, vol. 27, no. 10, pp. 2638–2642, Oct. 2023.
[9] Y. Zhang et al., "Fast near-field beam training for extremely large-scale array," *IEEE Wireless Commun. Lett.*, vol. 11, no. 12, pp. 2625–2629, Dec. 2022.
[10] Y. Liu et al., "Adaptive non-uniform hybrid beamforming for mmWave train-to-ground communications in high-speed railway scenarios," *IEEE Trans. Veh. Technol.*, vol. 72, no. 7, pp. 9085–9098, Jul. 2023.
[11] J. Zheng et al., "Cell-free massive MIMO-OFDM for high-speed train communications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 10, pp. 2823–2839, Oct. 2022.
[12] Y. Liu et al., "Near-field communications: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, early access, Oct. 14, 2024, doi: 10.1109/COMST.2024.3475884.
[13] S. Noh et al., "Multi-resolution codebook and adaptive beamforming sequence design for millimeter wave beam alignment," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5689–5701, Sep. 2017.
[14] X. Wei et al., "Codebook design and beam training for extremely large-scale RIS: Far-field or near-field?" *China Commun.*, vol. 19, no. 6, pp. 193–204, Jun. 2022.
[15] C. Wu et al., "Two-stage hierarchical beam training for near-field communications," *IEEE Trans. Veh. Technol.*, vol. 73, no. 2, pp. 2032–2044, Feb. 2024.
[16] M. Golchoubian et al., "Pedestrian trajectory prediction in pedestrian-vehicle mixed environments: A systematic review," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 11, pp. 11544–11567, Nov. 2023.
[17] Z. Wang et al., "Optimal bilinear equalizer for cell-free massive MIMO systems over correlated Rician channels," *IEEE Trans. Signal Process.*, early access, Mar. 11, 2025, doi: 10.1109/TSP.2025.3547380.