

Research Article

Transceiver Decoupling of Multivariate Symmetric Hybrid Precoding Based on 5G

Qiuliang Zhang,¹ Zhike Zhang,² Zhihui Wang ,¹ and Dong Yang¹

¹China Academy of Railway Sciences Corporation Limited, Beijing 100000, China

²China State Railway Group Co., Ltd., Beijing 100000, China

Correspondence should be addressed to Zhihui Wang; 52151724@ncepu.edu.cn

Received 13 January 2022; Accepted 16 February 2022; Published 26 March 2022

Academic Editor: Kalidoss Rajakani

Copyright © 2022 Qiuliang Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Mobile Internet will promote the continuous change of human interaction, leading to an increase in mobile traffic, so the demand for network bandwidth and data volume is rising rapidly, which is also one of the problems that 5G needs to solve. The mobile communication network of the railway system has the characteristics of high-speed user mobility, large-scale group mobility of users, high certainty of user mobile lines, and high QoS requirements for dispatching information. In order to improve the transmission reliability requirements of the railway system for wireless communication, a quick search method algorithm based on GMCS model to encode the number of each subinterval is proposed. Hybrid precoding is designed according to multivariate symmetry rules. The target beam is designed according to the GMCS model, and the hierarchical training beam is designed to minimize the mean square error between the training beam and the target beam as the objective function. Then, the fast search model based on beam overlap is extended to NLoS to solve the problem of misjudgment caused by multipath. In the simulation experiment, it proves that the search success rate of the research in this paper is 10% higher than that of the traditional algorithm. It improves the search speed and has obvious advantages in complexity. It can provide a dynamic reliable conversion mechanism for the railway communication environment, reduce the transmission power of the base station, and optimize the actual effect of uplink and downlink service requirements.

1. Introduction

Wireless communication technology has been closely integrated with people's lives. Mobile Internet technology has revolutionized the business model of traditional mobile communications and provides a brand-new user experience [1]. The Internet of Things has expanded the service scope of the mobile Internet, extending the communication between people to things and the intelligent interconnection between people and things. The mobile communication technology has been applied to all areas of society [2]. Literature [3] presents a Doppler frequency shift estimation and compensation algorithm based on position and precompensation for millimeter-wave HSR scene. This method calculates the Doppler frequency shift according to the position and speed of the train and precompensates before the signal occurs. However, this method needs to rely on additional hardware

equipment to provide high-precision position information. Reference [4] gives a Doppler shift cancellation method in millimeter-wave HSR mobile communication system. This method is mainly based on the assumption that the received signals of the head and tail relays on the top of the train have the same reverse frequency offset. The influence of Doppler shift is eliminated by simply multiplying the received signals of the head and tail relays. However, when the train passes through the base station, the Doppler frequency shift will decrease rapidly, and the time of the head and tail relays passing through the base station is different. Therefore, the performance of this method will become worse when the train handoff. Reference [5] gives a method for calculating Doppler frequency shift using channel estimation in millimeter-wave high-speed mobile scene. This method eliminates the influence of frequency shift through two-step processing of estimation and compensation, but this

method is based on more accurate channel estimation information. However, this method needs to pay a large computational complexity to obtain ideal estimation accuracy. In order to meet the requirements of special pilot segmentation, this method has strict restrictions on the length of pilot symbols so that it cannot be applied to the communication system with defined pilot structure.

For the hybrid beamforming structure of 5G communication, the analog beamforming depends on the relative position of the base station and the user end. Therefore, the position information is more used to simulate the beamforming part, with the parameters that the position information can provide to select the appropriate beam direction and establish the optimal beam channel so that a larger channel capacity can be obtained. In literature [6], the most direct idea is shown based on position information, considering the uniform linear antenna array (Uniform Linear Array (ULA)) and assuming that there is only a direct line of sight. Through the location information of the base station (BS) and the mobile terminal (mobile station (MS)), the distance, the angle of arrival (AoA), and the angle of departure (AoD) can be calculated. Then, AoA/AoD and distance are directly used as channel parameters to select the best codeword from the codebook based on maximizing the Signal-to-Noise Ratio (SNR). In literature [7], the use of location information in the high-speed rail scenario is studied. The high-speed rail environment represents the scene of a fixed trajectory, and the beam direction can be directly calculated by the angle information relative to the base station. Then, the power distribution is performed according to the beam direction. Since its trajectory is fixed, it can be calculated in advance. The performance of this method is very good in high-speed environment. In literature [8], the position information is also used to calculate the corresponding height angle and azimuth angle of the direct viewing path. Then, the calculated angle information is directly used to build the channel matrix as the result of channel estimation, which is applied to replacing the role of the reference signal. This method is relatively simple, but the accuracy of the estimated results is low, and other methods need to be used for further estimation, so the final cost has not been reduced too much. Reference proposed a coding management scheme based on ternary symmetric polynomials. The difference is that literature [9] is based on tree-ring model, which combines coding distribution with routing establishment process so as to reduce the energy consumption in the process of coding negotiation. Literature [10] not only designed the intercluster multihop routing algorithm according to the establishment of ternary symmetric polynomial but also introduced the network update parameters and update authentication number in order to ensure the security of the network in the coding update stage.

With the increasing popularity of mobile Internet, providing passengers with abundant mobile communication services and improving the ride experience have become another important task of the railway mobile communication system. Therefore, the number of wireless communication base stations deployed along railway, the density of base stations in central hubs and railway line dense areas, and the

energy consumption are all increasing. At the same time, railway lines are scattered, with many branches and diversified business types. Researching the energy-saving and emission reduction of wireless communication networks in railway application scenarios must consider various private networks, including railway mobile communication networks and public wireless communication networks. In addition, the research on energy-saving methods and mechanisms of mobile communication networks in railway scenes has important political, economic, and social significance.

Beamforming technology is a signal preprocessing technology based on antenna arrays. According to the spatial characteristics of the channel, the interference and diffraction principles of electromagnetic waves are used. By adjusting the phase or amplitude of each element in the antenna array, a beam with high directivity is formed, thereby obtaining a higher beamforming gain. The hybrid precoding algorithm based on the decoupling of the transceiver or the hybrid precoding algorithm based on the combination of the transceiver requires the transceiver to obtain the CSI in advance. Then, the signal can be precoded, and the accuracy of CSI will directly affect the performance of the precoding. However, in the millimeter-wave Massive MIMO system, the number of array antennas is very large, and the acquisition and feedback of CSI will become an important difficulty. Because the channel estimation technology has high complexity and the feedback ability of the feedback link is limited, when the transmitter obtains CSI, the channel may have changed, and the quality of the communication link cannot be guaranteed. Therefore, it is necessary to design a more efficient and faster CSI acquisition model. In existing research, considering factors such as system complexity, power consumption, and time delay, analog beamforming is an important selection model in the design of millimeter-wave Massive MIMO systems. In the analog beamforming structure, a connection can be established through beam search. The base station and the user transmit training beams multiple times to determine a suitable communication beam pair to establish a communication link. Compared with a mixed digital-analog structure, its beamforming technology does not require the sender and receiver to obtain CSI in advance, and the implementation cost and complexity are lower.

In order to meet the coverage reliability requirements of the railway system, the high frequency band system is more suitable for deployment in the railway scene dominated by the LoS track. Therefore, this paper mainly studies the 5G communication fast beam search model suitable for railway systems and extends the proposed model to general 5G communication. First, in order to reveal the difference between different search models, State Utilization Efficiency (SUE), Information Representation Efficiency (IRE), and Feedback Efficiency (FE) are introduced, and these three concepts are used to analyze the existing search models. The results show that the higher the SUE (IRE), the lower the search complexity. In 5G communication, it is deduced that Gray coding is used to encode the coding of each subinterval, which can make full use of the overlap state of the waveform, and then design a beam overlap search model based on Gray

mapping with 100% SUE and IRE. Next, it points out the problems of low reliability of 5G communication in the LoS scene and designs a beam search model suitable for the NLoS environment [11]. Compared with the model in Document 10, the proposed model has more sufficient theoretical analysis on the model of encoding the target beam, the hierarchical search design is more reasonable, and the search success rate is higher. Finally, the performance of the search model is simulated, and it is verified that the proposed model can obtain a higher search success rate with the lowest complexity in the railway system, which provides a reliable guarantee for railway communication system.

The main contributions of this paper are as follows:

- (1) According to the characteristics of high-speed movement, frequent switching, and traffic changing with train operation mechanism of railway mobile communication network, an application service energy-saving method is improved
- (2) For the analog beam forming structure, a fast search model suitable for LoS scenario and NLoS scenario is designed
- (3) Multivariate symmetric hybrid precoding is designed to realize the optimal combination of transceiver decoupling

The organization structure of this paper is as follows.

Section 1 introduces the research background and significance of this paper as well as main lines to increase the railway system wireless communication data rate and leads to beamforming technology.

Section 2 introduces the signal model, channel model, and optimization problems of the millimeter-wave Massive MIMO system.

Section 3 identifies the corresponding optimization problems in the system by modeling the special situation of the railway wireless communication system.

Section 4 is the combined design of hybrid precoding transceiver decoupling, the beam search in the NLoS scene has the problem of misjudgment caused by multipath compared with the LoS scene, and a search model suitable for the NLoS scene is designed.

Section 5 verifies the misjudgment probability and success rate of the model proposed in this paper through simulation, which proves the optimization effect of the research content of this paper.

Section 6 summarizes the research work of the paper and looks forward to the future research directions of beamforming technology in 5G systems.

2. Related Work

In this paper, the beamforming technology is studied in the order of digital precoding, analog beamforming, and hybrid precoding. The following describes the relevant content of each technical project.

2.1. Digital Precoding. According to different coding generation methods, the coding management scheme can be divided into symmetrical coding management scheme and nonstacking coding management scheme. Symmetric coding management scheme means that the encryption and decryption codes used by both communication parties are the same, and the codes can be calculated in their own ways. The characteristic of this scheme is that the computing overhead and storage overhead of the network are small.

In the traditional MIMO system, precoding is done by a digital encoder in the baseband part. Digital precoding is widely used in the 3rd-generation mobile communication system TD-SCDMA and the 4th-generation mobile communication system TD-LTE; the structure is shown in Figure 1.

In it, μ represents the original signal, N_s represents the number of data streams, ω represents the weight vector of the antenna array, N_{RF}^1 , N_{RF}^t , and N_t represent the number of Radio Frequency (RF) chains and the number of antennas, respectively, and r represents the processed received signal [12]. For the digital precoding structure, each antenna at the transmitter and receiver is connected to an independent RF link. The advantage of this structure is that it can control the amplitude and phase of the signal, has the highest flexibility, can achieve the optimal precoding effect, and also supports the parallel transmission of multiple data streams [13].

2.2. Simulating Beamforming. Beamforming technology is a signal preprocessing technology based on antenna arrays. According to the spatial characteristics of the channel, with the help of the interference and diffraction principles of electromagnetic waves, a beam with high directivity is formed by adjusting the phase or amplitude of each element in the antenna array, thereby obtaining a higher beamforming gain [14]. What is more, the hybrid precoding algorithm based on the decoupling of the transceiver or the hybrid precoding algorithm based on the combination of the transceiver requires the transceiver to obtain the CSI in advance so that the signal can be precoded. Moreover, the accuracy of CSI will directly affect the performance of precoding [15]. However, in the millimeter-wave Massive MIMO system, the number of array antennas is very large, and the acquisition and feedback of CSI will become an important difficulty. Since the channel estimation technology has high complexity and the feedback ability of the feedback link is limited, when the transmitter obtains CSI, the channel may have changed, and the quality of the communication link cannot be guaranteed [16]. Therefore, it is necessary to design a more efficient and faster CSI acquisition model. In the existing research, considering factors such as system complexity, power consumption, and time delay, analog beamforming is an important choice model in the design of millimeter-wave Massive MIMO systems. Besides, in the analog beamforming structure, a connection can be established through beam search. The base station and the user transmit training beams multiple times to determine a suitable communication beam pair to establish a communication link. Compared with a mixed digital-analog structure, the beamforming technology under this structure does not

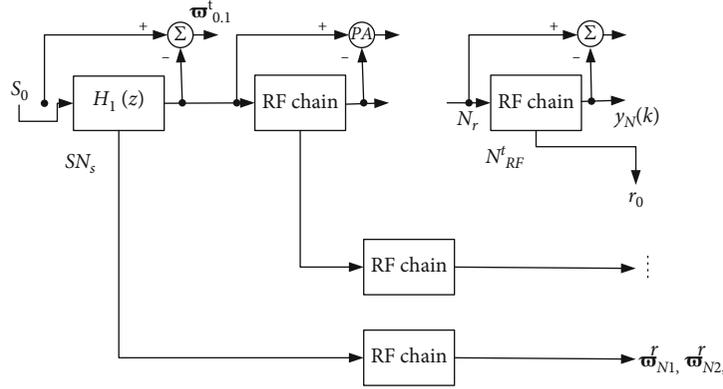


FIGURE 1: Digital precoding structure.

require the sender and receiver to obtain the CSI in advance, and the implementation cost and complexity are lower [17].

The structure of simulated beamforming is shown in Figure 2.

In this structure, the S terminal is the signal input, and PA is the power amplifier. All antennas share an RF chain, and the desired directional beam is obtained by controlling the weight of each antenna. The advantage of the analog beamforming structure is that the number of RF chains is small, which greatly reduces the hardware cost and complexity of the system. But the shortcomings of this structure are also obvious [18]. On the one hand, there is only one RF chain, which cannot take full advantage of the high degree of freedom of Massive MIMO. It can only support single data stream transmission and cannot be used in multiuser scenarios. On the other hand, because the structure can only control the phase of the signal and cannot adjust the amplitude, the performance is poor. In addition, the accuracy of the phase shifter in reality is also limited, which further limits the performance of the system.

2.3. Hybrid Precoding. Neither digital precoding nor analog beamforming can take full advantage of the Massive MEMO system. In order to improve the spectrum efficiency of the system and have the ability to transmit multiple data streams in parallel, the academia proposed a hybrid precoding structure. The hybrid precoding structure is divided into low-dimensional digital precoding and high-dimensional analog precoding. Connecting through a small number of RF chains can give full play to the advantages of the two. According to the characteristics of the channel, this structure first accurately matches the channel characteristics through low-dimensional digital precoding, so as to improve system performance and support multistream transmission. Then, in the high-dimensional analog precoding, the analog phase shifter is used to match the spatial characteristics of the signal in the analog domain to obtain the array gain, which effectively compensates for the high path loss of the millimeter wave [19].

A public asymmetric matrix and a confidential symmetric matrix are constructed by the server, and a symmetric matrix is constructed through the two matrices as the basic

matrix for establishing shared coding between network communication nodes. The program is mainly divided into the following stages.

2.3.1. Code Initialization. In the coding initialization stage of a network with n nodes, the server first needs to generate a matrix G with $(\lambda + 1)$ rows and columns as the public matrix and inform all nodes in the network. Secondly, the server also needs to generate a column of $(\lambda + 1)$ confidentiality symmetry matrix D and calculate the matrix $A = (G \cdot D)^T$. In it, $(G \cdot D)^T$ is the transpose matrix of $(G \cdot D)$. Finally, the server sends matrix A to each node in the network.

2.3.2. Code Establishment. After receiving the matrix distributed by the server, the nodes in the network obtain the matrix $K = A \cdot G$ by calculating. According to the definition of symmetric matrix, the transpose of symmetric matrix is equal to the symmetric matrix itself. Therefore, the formula $K = A \cdot G = (G \cdot D)^T \cdot G = G^T \cdot D^T \cdot G = G^T \cdot D \cdot T = G^T \cdot (G \cdot D)^T = K^T$ can be deduced. Therefore, K is also a symmetric matrix. The communication node i and node j in the network store the information of rows i and j of matrix K , respectively. At the beginning of communication establishment, node i and node j exchange their row and column information, respectively. After node i obtains the information of node j , it extracts the information of column j from the information of row i saved by itself as the communication code t . Similarly, node j extracts the information of line i from its own matrix information as communication code t . From the definition of symmetric matrix, the calculation $K_{ij} = K_{ji}$ is obtained.

3. Wireless Communication System Model and Corresponding Optimization Problems

The mobile communication system in the railway application scenario has diversified characteristics, such as high-speed movement (currently up to 500 km/h), high quality of service (QoS) requirements, severe electromagnetic environment, high certainty of user movement lines, and large-scale collective movement of users and load-bearing services. Therefore, the railway construction under 5G signal system

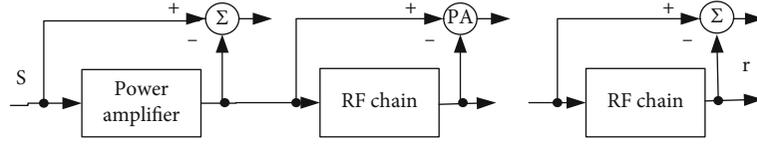


FIGURE 2: Simulated beamforming structure.

has a dedicated mobile communication network and a public mobile communication network, which usually have different functions and characteristics.

3.1. System Model. When the user needs to communicate, the base station and the user can obtain user location information. Then, the appropriate beam pair is selected by the algorithm to simulate beamforming, and the system uses a two-stage design method to perform hybrid beamforming. This paper uses a symmetrical beamforming millimeter-wave communication system, that is, both the transmitter and receiver use analog beamforming, as shown in Figure 3.

Assuming that the number of antennas of the transmitter is N_t , the number of antennas of the receiver is N_r , $N_t = N_r = P$, and the number of communication beams is Q . In order to improve the coverage of communication beams, Q is generally satisfied. These $Q = 2P$ communication beams cover the entire angular range of the transmitter and receiver, respectively, and then take AoD as an example for analysis (all analyses are applicable to AoA). It means that the communication beam actually divides the entire AoD interval into Q subintervals, and these subintervals are represented by $A_q (q = 0, 1, \dots, Q - 1)$.

The purpose of beam search is that in the scenario of unknown CSI, the communication parties use the codewords in the code book to perform multiple searches in the established training code book according to a certain search model, determine the subinterval to which the AoD belongs, and select the communication beam in this subinterval as a suitable communication beam to establish a communication link.

3.2. Optimization Problem. This paper takes the Training Beam Receiving Power (TBRP) of the receiver as the evaluation index, which is defined as

$$\text{TBRP} = y^H y. \quad (1)$$

The beam search is to search for the optimal codeword $f_{\text{BS,MS}}^W$ in the code book of the transceiver training beam. Then, the communication beam pair located in the corresponding subinterval of the codeword is regarded as the appropriate communication beam pair $A_{q(\text{opt})}^t, A_{q(\text{opt})}^r$ [20]. The code book of the transmitter is defined as A and the codebook of the receiver is defined as B , and then, the opti-

mization problem can be written as

$$\begin{cases} \arg \max \\ f_{\text{BS,MS}}^W \\ \text{s.t.} \left\{ \begin{array}{l} f_{\text{BS} \subset A} \\ w_{\text{MS}} \in B \end{array} \right\}. \end{cases} \quad (2)$$

4. Combined Design of Hybrid Precoding Transceiver Decoupling

The existing search models are simply divided into two categories, nonhierarchical search models and hierarchical search models. The nonhierarchical search model refers to traversal search. Among the hierarchical search models, the more classic ones include the search model of IEEE 802.15.3c, the search model of IEEE 802.11.ad, the search model based on SWR, the search model based on points, and the fast search model based on overlapping beam (FSSOB) [21].

4.1. Multivariate Symmetric Coding Design. According to the properties of symmetric polynomials, multivariate symmetric polynomials propose a group coding allocation mechanism. The mechanism includes an authentication node, which can randomly generate a set of sequences $\alpha_{ij}, i \geq 0, j \geq k$ for generating polynomial information. New nodes need to be authenticated before they can join the network. The authentication node will distribute the polynomial information to the authenticated nodes. In this mechanism, each node is required to have a unique ID. When two nodes establish communication, they need to exchange their respective ID and bring each other's ID into polynomials to calculate the communication code.

In the initial stage of code establishment, the authentication node will generate a bivariate polynomial of order k :

$$f(x, y) = \sum_{i=0}^k (a_i x^i y^i). \quad (3)$$

The authentication node randomly extracts multiple numbers from the data pool as polynomial coefficients $\alpha_{ij}, i \geq 0, j \geq k$, which meets the conditions. Therefore, the polynomial is a bivariate symmetric polynomial, which satisfies the following properties:

$$f(x, y) = f(y, x). \quad (4)$$

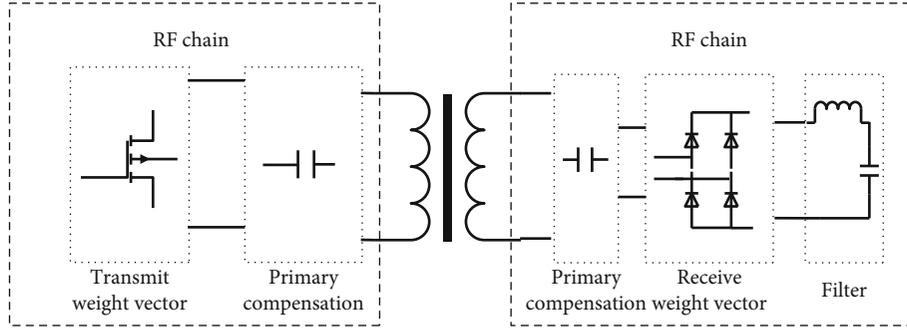


FIGURE 3: Simulated beamforming structure.

The authentication node distributes the polynomial to each node. Each node brings its own ID into the polynomial $f(x, y)$ and obtains $f(\text{ID}, y)$. Then, it saves the calculated polynomial $f(y) = f(\text{ID}, y)$ in the built-in information of the node. Therefore, even if the node is captured, the original polynomial cannot be obtained from the node.

4.2. Introduction to Existing Search Models. For the traversal search method, the training beam is the communication beam. When searching, the transmitter transmits Q communication beams in turn, and the receiver switches and receives among the Q receiving beams and records the strength of the received signal [22]. After Q searches, all the transceiver beam pair combinations can be traversed. Then, the receiver feeds back the beam number with the strongest signal to the transmitter to complete the search. The number of transmissions of the training beam is Q^2 , the number of feedbacks is 1, and the amount of feedback information is $\log_2 Q$ bits [23].

IEEE 802.15.3c uses a hierarchical search model, and the core idea is traversal search. The search model in 802.15.3c is divided into sector-level search and beam-level search. Some concepts are given, and then, the 3C training mechanism is described. For unification, the transmitter and receiver are used to replace device 1 and device 2 in the protocol [23]. The transmitter and receiver, respectively, contain $J^{(t)}$ and $J^{(r)}$ sectors, and the corresponding codewords are $[D_0^{(t)}, D_1^{(t)}, \dots, D_{J^{(t)}-1}^{(t)}]$ and $[D_0^{(r)}, D_1^{(r)}, \dots, D_{J^{(r)}-1}^{(r)}]$, respectively. The number of beams contained in each sector of the transmitter and receiver is $K^{(t)}$ and $K^{(r)}$, and the corresponding codewords are $[B_0^{(t)}, B_1^{(t)}, \dots, B_{K^{(t)}-1}^{(t)}]$ and $[B_0^{(r)}, B_1^{(r)}, \dots, B_{K^{(r)}-1}^{(r)}]$, respectively, where $J^{(t)}K^{(t)} = J^{(r)}K^{(r)} = Q$ [24].

Taking Figure 4 as an example, briefly introduce the training mechanism of 802.15.3c.

The transmitter has 4 sectors and the receiver has 8 sectors, so there are a total of 4 cycles at the sector level (cycle in the protocol). In the n cycle, the transmitter repeatedly transmits 8 training beams along the direction of $D_{n-1}^{(t)}$ [25]. Meanwhile, the receiver sequentially receives in 8 directions $D_0^{(r)}, D_1^{(r)}, \dots, D_7^{(r)}$. After 4 cycles, the signal energy of all combinations at the transceiver sector level can be determined [26]. According to the feedback from the

receiver, determine the best sector pair, such as 0 and 3, and then enter the beam-level search. The transmitter has 8 beams and the receiver has 4 beams, so the beam level has a total of 8 cycles [27]. In the n cycle, the transmitter repeatedly transmits the training beam 4 times along the direction of $B_{n-1}^{(t)}$. The receiver sequentially receives in 4 directions $B_0^{(r)}, B_1^{(r)}, B_2^{(r)}, B_3^{(r)}$. After 8 cycles, the signal energy of all combinations of the transceiver beam level can be determined. According to the feedback of the receiver, the optimal beam pair can be determined and the search can be completed. It can be seen that the number of sending training beams is $J^{(t)}J^{(r)} + K^{(t)}K^{(r)}$, the number of feedbacks is 2, and the amount of feedback information is $\log_2 J^{(t)} + \log_2 K^{(t)}$, which is $\log_2 Q$ bits [28].

The agreement points out that, in Sector-Level Sweep (SLS), the number of sectors at this stage needs to be determined. The calculation method of the number of sectors is the total number of transmitter sectors multiplied by the number of receiver antennas $\log_2 16 = 4$, and the value is set to the total number of sectors in the SLS stage [29]. Therefore, in the SLS phase, the transmitter transmits a total of $N_{\text{sector}}^{(t)}N_r$ sector-level training beams, and the receiver receives quasi-omnidirectional reception. After completing the search, the receiver determines the best transmitting sector based on the receiver's feedback. Similarly, in the multiple sector ID capture (MIDC), the transmitter transmits quasi-omnidirectionally, and the receiver uses different sectors for reception. After $N_{\text{sector}}^{(t)}N_r$ beam searches, the receiver determines the best receiving sector. The Beam Combining (BC) stage mainly pairs the beams in the sectors found above and performs a maximum of s^2 inspections (where $S \leq 7$) to find the optimal beam pair and feed it back to the transmitter [30]. Therefore, the number of transmissions of the training beam is $N_{\text{sector}}^{(t)}N_r + N_{\text{sector}}^{(t)}N_t + s^2$, the number of feedbacks is 2, and the amount of feedback information is $\log_2 N_{\text{sector}}^{(t)} + \log_2 N_{\text{beam}}' = \log_2 Q$ bits. The literature [31] points out that the BC stage is an optional stage, so the number of transmissions of the training beam can be modified to $N_{\text{sector}}^{(t)}N_r + N_{\text{sector}}^{(t)}N_t$ and satisfy $N_{\text{beam}}^t N_r \geq N_{\text{sector}}^{(t)}$.

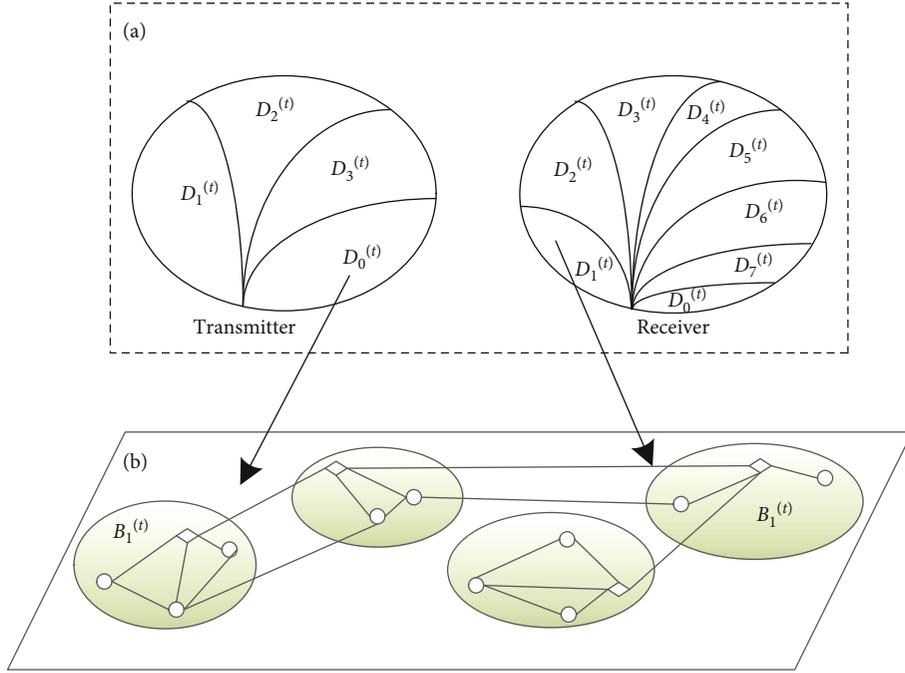


FIGURE 4: Schematic diagram of IEEE 802.15.3c training mechanism.

The SWR-based search model divides the search interval into two subintervals at each stage and uses two independent training beams for coverage, so it is also called a binary search model. In each stage, the transceiver can get all the beam combinations in this stage through 4 beam searches. Then, the receiver feeds back the transmit beam number with the highest signal strength to the transmitter and uses this subinterval as the beam search range of the next stage until the best communication beam is found. Therefore, the search process of binary search is divided into $\log_2 Q$ stages. In each stage, the transmitter needs to transmit 4 training beams, and the amount of feedback information is $\log_2 Q = 1$ bit. The number of transmissions of the training beam is $4 \log_2 Q$, and the total feedback amount is $\log_2 Q$ bits.

In the k subsearch model, each stage divides the search interval into k parts, and a total of $\log_k Q$ stages are required (assuming that $\log_k Q$ is a positive integer). In each stage, the transceiver passes $K^2 K$ beam training to determine the best transceiver beam for this stage and feeds it back to the transmitter as the search range for the next stage. Therefore, the number of transmissions of the training beam is $k^2 \log_k Q$, the number of feedbacks is $\log_k Q$, and the total feedback amount is $\log_2 K \times \log_k Q = \log_2 Q$ bits. Comparing k subsearch and binary search, it can be found that if k is set to 2 in k subsearch, k subsearch is binary search, that is, binary search is a special case of k subsearch. So, the subsequent analysis will be performed with the binary search as the representative of the k subsearch.

In the search process of FSSOB, each stage uses $\log_2(H+1)$ training beams ($H+1$ is an integer power of 2) to divide the search interval into Q subintervals. In theory, the FSSOB model requires a total of $\log_2 H$ stages, and each stage uses $\log_2 H$ bit binary to represent Q subintervals. So,

the total amount of feedback is $\log_2 H \times \log_2 Q = \log_2 Q$ bits. However, in the FSSOB scheme, Q and H do not satisfy the relationship of integer powers, and the number of search stages needs to be rounded up. Therefore, the actual number of times the training signal is sent is $\log_2^2(H+1) \lceil \log_2 Q \rceil$, and the actual feedback amount is $\log_2(H+1) \lceil \log_2 Q \rceil$, which is greater than $\log_2 Q$ bits. Taking $H=3$, $Q=8$ as a concrete example to illustrate. This algorithm requires the $\lceil \log_3 8 \rceil = 2$ stage, the number of training beams required for each stage is $\log_2^2(3+1) = 4$, and the actual feedback amount is $\lceil \log_3 8 \rceil \times \log_2^2(3+1) = 4$ bits, which is larger than the theoretical $\log_2 Q = \log_2 8 = 3$ bits.

According to the above analysis, for the Q communication beams, in the traversal search, 802.15.3c, 802.11.ad, binary search, and K subsearch models, their actual total feedback is $\log_2 Q$ bits, while the amount of total feedback in the FSSOB model is greater than $\log_2 Q$ bits. Besides, the number of transmissions of the training beam in different search models is different, that is, the complexity is different. In order to explore the essential differences of different search models, the concepts of SUE, IRE, and FE are introduced, and the SUE, IRE, and FE of the existing search models are analyzed.

4.3. Search Model Based on Beam Overlap in LoS Scenarios.

In this section, based on theoretical analysis, it is first deduced that if Gray mapping is used to encode the subinterval numbers of each stage, 100% SUE with the highest success rate can be obtained. Then, the target beam is determined according to the code of the subinterval number of each stage, and the training beam is designed by minimizing the mean square error between the training beam and the target beam. Finally, a search model based on beam overlap

is proposed, and the SUE, IRE, FE, and complexity of the model are analyzed.

4.3.1. Coding Model Analysis. According to the analysis of the existing search model, although the SUE of the FSSOB model with the overlapped waveform state does not reach 100%, it is much higher than the other search models that do not use the overlapped state. Therefore, only by encoding the subinterval numbers in each stage according to a certain coding model as the FSSOB model can the target beam and the training beam in each stage overlap. Meanwhile, more states with fewer beams can be obtained to get 100% SUE.

In order to encode the subinterval number of each stage, it is necessary to determine the theoretically required number of bits of the encoding at this stage. In addition, since the target beam needs to be designed according to this code and there are two types of coverage of the target beam in each subinterval, namely, covering or not covering, which can be represented by two states of “1” and “0,” binary numbers are used to encode the subinterval numbers of each stage, and the number of binary numbers is also the number of target beams and training beams in this stage. Assume that the n stage contains $N_{\text{area}}^{(n)}$ subintervals and $N_{\text{area}}^{(n)}$ is an integer power of 2, that is, $N_{\text{area}}^{(n)} = 2B$ ($Q = 2P$ is a positive integer). Since the state of $N_{\text{area}}^{(n)}$ subintervals is fed back and the amount of information required in theory is $\log_2 N_{\text{area}}^{(n)} = B$ bits, the number of bits in the binary number must be $\log_2 N_{\text{area}}^{(n)} = B$ in order to obtain 100% FE. In addition, when the B -bit binary number is used to represent the $N_{\text{area}}^{(n)}$ subintervals, all the states of the stage are used in each stage, so 100% SUE can be obtained. In order to facilitate the analysis and understanding, the coding model is explained by taking 8 subintervals as an example, which will be generalized to the general situation.

(1) Coding Model of 8 Subintervals. According to the above analysis, the number of the 8 subintervals should be coded with $\log_2 8 = 3$ -bit binary numbers. These 3 binary numbers correspond to 3 target beams, and the number is also the number of training beams. When numbering, the corresponding binary number code is “1” for the interval covered by the target beam, and the corresponding binary number code is “0” for the uncovered interval. Therefore, according to the Binary Mapping Coding Scheme (BMCS), 8 subintervals can be coded as “000,” “001,” “010,” “011,” “100,” “101,” “110,” and “111.” The corresponding subinterval coverage is shown in Figure 5.

In Figure 5, target beams 1, 2, and 3, respectively, correspond to the first, second, and third positions in the serial number. Meanwhile, in this paper, the area covered by the target beam is defined as the main lobe, and the other areas are defined as side lobes. Therefore, the target beams 1, 2, and 3, respectively, have 1 main lobe, 2 main lobes, and 4 main lobes. In addition, different target beams have different gains in the corresponding coverage area, which is to clearly distinguish different target beams in the figure. In fact, the

gains of different target beams in the corresponding coverage areas are the same.

When performing beam search, the training beam with the criterion of minimizing the mean square error between the training beam and the target beam is first designed, and then, multiple beam searches are conducted between the transceivers. What is more, the receiving end judges the best AoD interval according to the strength of the received signal and feeds it back to the transmitter. Since the edge of the main lobe of the actual training beam is not steep enough, there is a possibility of misjudgment. For example, in Figure 5, if AoD is located at the right edge of the “001” subinterval, it is not only close to the right edge of training beam 3, but also close to the left edge of training beam 2. If the gain of the training beam 2 at this position is not reduced to a sufficiently low level, the receiver will misjudge the subinterval of the best AoD as “011.” Therefore, for the search model based on beam overlap, it is necessary to minimize the number of main lobe edges of all target beams in order to reduce the probability of misjudgment. In addition, it is necessary to analyze the coding performance of 3-BMCS.

In order to quantify the performance of different coding models, the Hamming Distance (HD) $d(x_i, x_j)$ in information theory is introduced, which represents the number of different characters at the corresponding positions of two strings of equal length. Analyzing the HD of 3-BMCS, the result is shown in Figure 6.

By comparing Figures 5 and 6, when the number $d(x_i, x_j)$ between the two adjacent subintervals in Figure 6 is n , there are n main lobe edges between the two adjacent subintervals in Figure 5. For example, if $d(x_i, x_j)$ between the subintervals “001” and “010” in Figure 6 is 2, the number of main lobe edges between the corresponding target beams in Figure 5 is also 2. Define the sum of the number of main lobe edges in all target beams as the total edge number (Total Edge (TE)). Since one main lobe corresponds to two edges, the half of TE is the number of main lobes. In Figure 5, the number of main lobes of the target beam is 7, and the TE is 14. In addition, as the number of binary digits increases, there will be a large HD between some adjacent codes in BMCS. This phenomenon is also known as the Hamming Cliff, which will greatly increase the probability of misjudgment by the search model.

Due to the fact that the edge of the main lobe in the actual training beam is not steep enough, it is necessary to minimize the number of main lobes in the target beam in order to reduce the probability of misjudgment by the search model, namely, the design criterion of the coding model is to minimize TE. Since the numbers of the 8 subintervals are different, the HD between adjacent subinterval numbers is at least 1. In other words, it is necessary to encode the number of each subinterval so that the HD between the numbers of two adjacent intervals can be as small as possible or even always equal to 1. After research, it is found that Gray Mapping Coding Scheme (GMCS) can meet the requirement. Additionally, the 3-digit Gray code is used to encode the numbers of the 8 subintervals, which is abbreviated as 3-

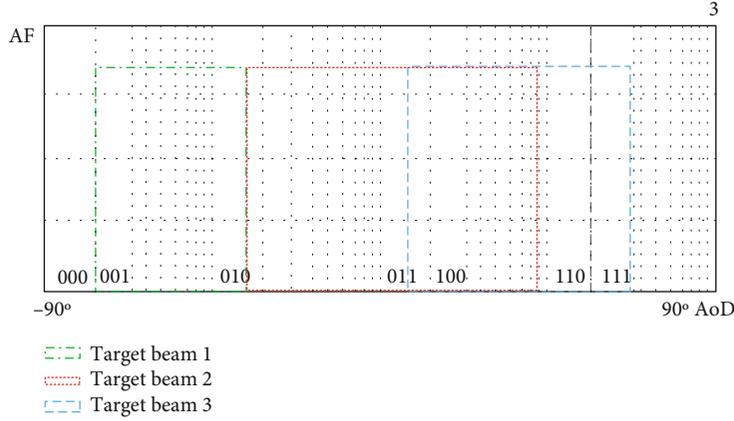


FIGURE 5: Target beam designed according to 3-BMCS.

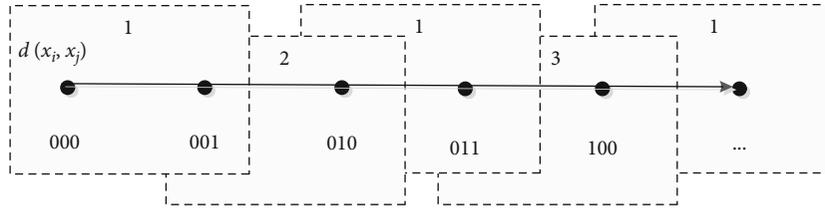


FIGURE 6: HD of 3-BMCS.

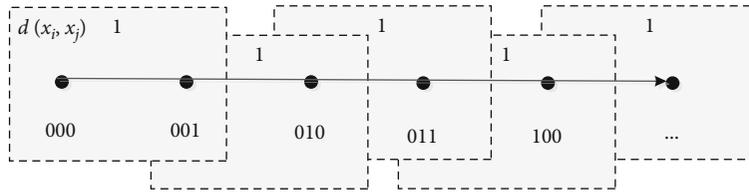


FIGURE 7: HD of 3-GMCS.

GMCS. Since in 3-GMCS, the HD between the numbers of two adjacent subintervals is always 1 and its TE is 8, which means that the number of main lobes of the three target beams is 4, only one of the target beams has two main lobes, and the other two target beams each have one main lobe. Next, 3-GMCS will be analyzed.

Since the start codewords of Gray mapping are different and different Gray codes will be obtained, there are many encoding forms. For 3-GMCS, when 8 codewords are connected end-to-end since in the angular space, 8 subintervals are connected end-to-end, there are 12 types that still meet the above rules. Through analysis, it can be found that the target beams designed based on these Gray codes are basically the same. Therefore, in the paper, the most representative 3-bit Gray code is taken as an example to encode the number of the subinterval, whose HD is shown in Figure 7.

3-GMCS is used to design the target beam, and the result is shown in Figure 8.

In Figure 8, both target beams 1 and 2 have only one main lobe, and target beam 3 has two main lobes, which is completely consistent with the above analysis.

Compared with 3-BMCS, 3-GMCS has the following advantages:

- (1) Reduce the possibility of misjudgment: the reason for the misjudgment of the search model is that the edge of the training beam is not steep enough. Therefore, the more main lobes of the training beam is, the greater the possibility of misjudgment will be. Since this model designs the training beam by approaching the target beam, the main lobe of the target beam is used for analysis here. From Figures 4–8 and Figures 4–11, it can be found that 3-BMCS has 7 main lobes, while 3-GMCS has only 4 main lobes, so 3-GMCS can effectively reduce the possibility of misjudgment. In addition, it can be seen from the HD of 3-BMCS and 3-GMCS that 3-GMCS can effectively reduce the possibility of multi-bit simultaneous misjudgment. For example, in Figures 4–8, if the AoD is located at the junction of the subintervals “101” and “110,” the second and third positions of the subinterval may be misjudged

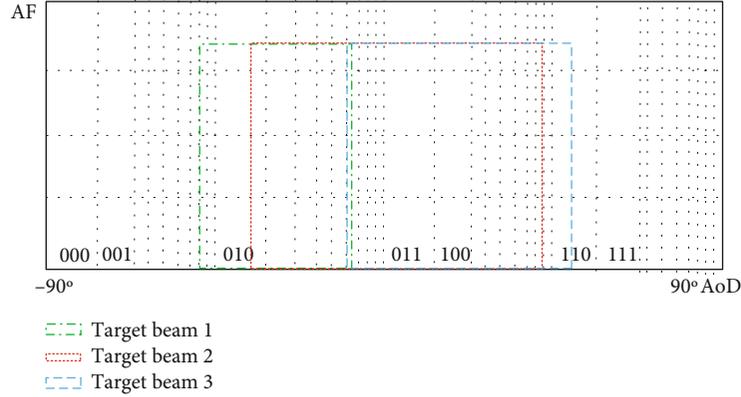


FIGURE 8: 3-GMCS design target beam.

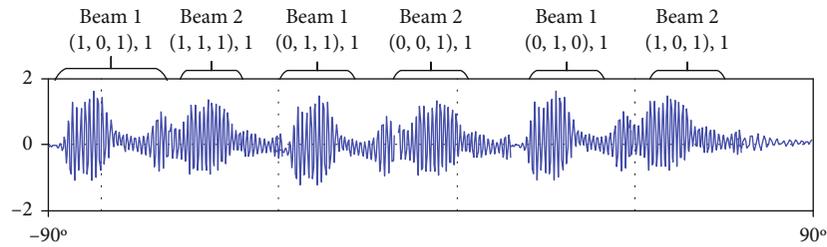


FIGURE 9: Schematic diagram of 2-GMCS target beam coverage.

at the same time. However, 3-GMCS will only cause a 1-bit misjudgment, which reduces the possibility of multiple-bit simultaneous misjudgments

- (2) Reduce the impact of misjudgments: compared with 3-BMCS, when an error occurs in 3-GMCS, the interval number obtained by the wrong judgment is actually adjacent to the optimal interval, which means that the suboptimal interval is found, and it can still provide relatively significant gain. For example, in Figures 4–10, when the interval is on the right edge of “011,” the situation mentioned above will exist. If the received power of target beam 3 does not reach the threshold, the receiver will judge the third bit of the interval as “0,” and the result of the erroneous judgment is “010,” which is next to the best interval “011.” It is suboptimal interval

(2) *Coding Model under General Conditions.* In each stage, the target beam has only one main lobe. For 16 subintervals, whether it is 4-GMCS (only one stage) or 2-GMCS (two stages), the number of times sending training signals is 4 times, and the amount of feedback information is 4 bits. The only difference is that the former only gives feedback once and the latter gives feedback twice.

For the case where the number of subintervals is greater than 16, the hierarchical thinking is adopted for design. Compared with Figures 4–11 and Figures 4–12, it can be seen that the main lobe of each target beam in 2-GMCS is

less. Therefore, when designing the hierarchical search model, 2-GMCS is the main method. If necessary, increase 3-GMCS at most once. Then, discuss the situation of $N_{\text{area}}^{(n)} = 2^B$ (B is a positive integer greater than 4):

- (1) When $B > 4$ and $B = 2c$ ($a(\theta)$ is a positive integer), $a(\theta)$ times 2-GMCS will be adopted
- (2) When $B > 4$ and $B = 2c + 1$, $c - 1$ times 2-GMCS and once 3-GMCS will be adopted

Therefore, the GMCS coding model can solve the coding problem of any subinterval number at each stage.

(3) *Application of GMCS in Different Communication Beam Scenarios.* In the above analysis, the design model of the GMCS coding model in different subintervals of each stage is analyzed. However, in actual communication, the beam search model in different communication beam scenarios is considered, which mainly includes the number of stages of the search model and the search model for each stage. Therefore, the GMCS model under different communication beam scenarios will be analyzed next. Denote the communication beam Q as $Q = 2M$ (M is a positive integer), and discuss the search models in different communication beam scenarios, as shown in Table 1.

It can be seen from the analysis results that the proposed search model based on GMCS is feasible regardless of the value of Q .

4.3.2. Design of Training Beam and Search Model

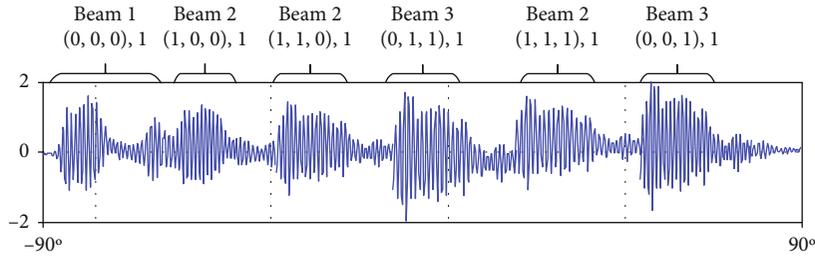


FIGURE 10: Schematic diagram of 3-GMCS target beam coverage.

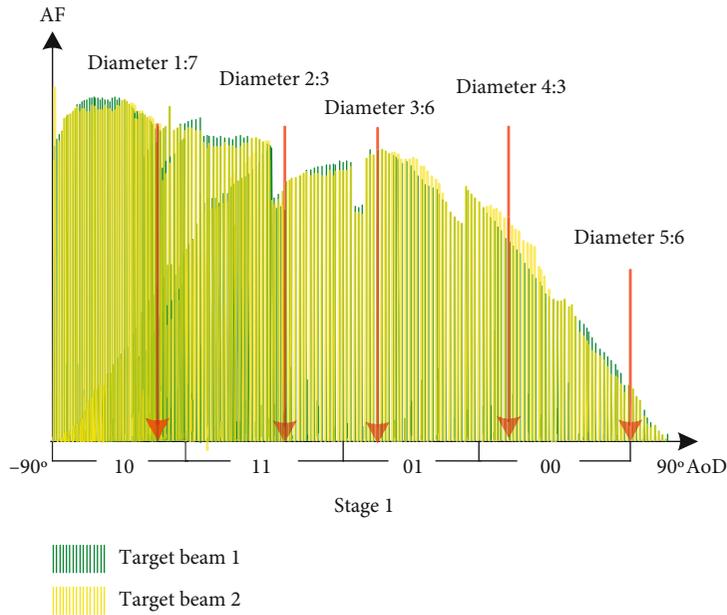


FIGURE 11: Schematic diagram of misjudgment due to multiple paths in the 2-GMCS model.

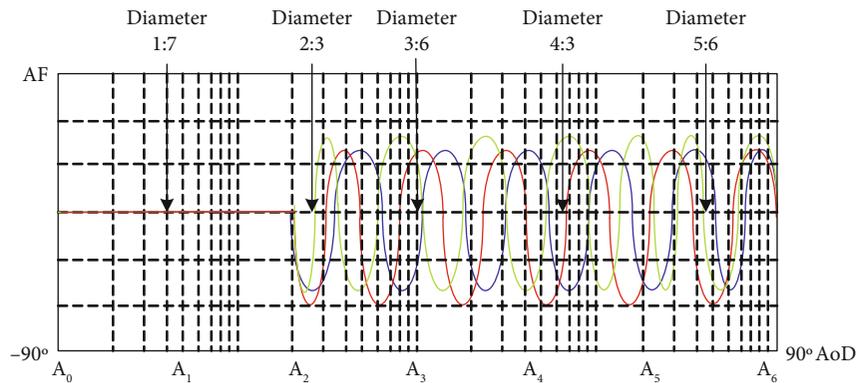


FIGURE 12: Schematic diagram of the 8-point search model for 5 paths.

(1) *Design of Training Beam.* The idea of designing a layered code book can be divided into two types according to the usage of the antenna. One is to increase the number of antennas layer by layer, and the other is to use all antennas on each layer. Both design ideas have their own advantages and disadvantages. The first idea is to control the number

of antennas in each layer by switching to generate training beams of different widths, whose advantage is that the training beams are better, but since the number of antennas in the first few layers is small, the transmission power is lower, reducing the coverage of the antenna and cause misjudgment by users who are too far away to receive the signal.

In addition, since each layer has to switch the antenna through hardware, the complexity of the system is increased. The advantage of the second model is simple operation and fixed transmission power, whose disadvantage is that the shape of the wide beam generated by the first few layers is not very ideal, reducing the search success rate of the search to a certain extent. After comprehensive consideration, the second design idea is adopted in the paper, that is, all antennas are used on each layer.

The number of antennas in each layer is determined, and then, the beam weight vector w of the training beam will be designed. Denote the beam response of the training beam as $AF(\theta)$, whose form is shown in (4). Moreover, the beam response $AF_{TB}(\theta)$ of the target beam can be obtained, and then, the target function will be established with the criterion of minimizing the mean square error between the training beam and the target beam, which is defined as follows.

$$f(w) = \int_{-\pi/2}^{\pi/2} |AF(\theta) - AF_{TB}(\theta)|^2 d\theta = \int_{-\pi/2}^{\pi/2} |w^H a(\theta) - AF_{TB}(\theta)|^2 d\theta. \quad (5)$$

Therefore, the design problem of the training beam is transformed into the following function.

$$\begin{aligned} & \arg_w \min f(w) \\ \text{s.t. } & |w_i| = 1, 0 \leq i \leq (P-1). \end{aligned} \quad (6)$$

MATLAB's own global search function (global search) is used to solve the problem, and the design results of the training beam will be given in the subsequent simulation.

(2) *Threshold Design.* In this search model, all zero states are used in each stage, that is, there are subintervals that are not covered by all training beams in each stage, such as $[C_{(0,0),1}]_1$. If the LoS path is in this interval, that is, there is no LoS path in the interval covered by training beams 1 and 2, and the power of training beams 1 and 2 received by the receiving end will be close. If the threshold is not set, it is impossible to judge whether the LoS path lies in $[C_{(0,0),1}]_1$ or $[C_{(0,0),1}]_1$. Therefore, a threshold is needed to be set in the paper to distinguish the situation where the LoS path is located within and outside the coverage of the training beam. In addition, since the received power is not only related to the path loss of each receiver but also related to the transmit power, the following settings are made for easy comparison.

First, the received power of the receiver plus the corresponding path loss is defined as the equivalent received power, namely, $\overline{pr}(dB) = P_r(dB) + PL_r(dB)$, to eliminate the influence of distance on the received power. Then, the ratio of the received power to the transmit power is defined as the relative received power, that is, $\beta = \overline{P}_r/P_t$, to eliminate the influence of the transmit power. In addition, the relative received power within a certain range can be used as the

TABLE 1: Search models in different communication beam scenarios.

M	2-GMCS execution times	3-GMCS execution times
$M = 2, Q = 4$	1	0
$M = 3, Q = 8$	0	1
$M = 4, Q = 16$	2	0
$M > 2, Q = 2^M (M = 2n)$	n	0
$M > 4, Q = 2^M (M = 2n + 1)$	$n - 1$	1

threshold to make judgments. More threshold analysis will be shown in subsequent simulations.

(3) *Fast Search Model Based on Gray Mapping.* Focusing on 2-GMCS, combined with Figures 4–14, and taking 16 communication beams as an example, the beam search is divided into 2 layers.

Layer 1: the transmitter is in interval $AoD_t^{(0)} \Sigma[-90^\circ, 90^\circ]$, which transmits training beam 1 and training beam 2 in sequence, and the receiver is in interval $AoD_r^{(0)} \Sigma[-90^\circ, 90^\circ]$, where the quasi-omnidirectional mode is applied to receiving and recording the received power. If only when the transmitter transmits training beam 2, the received power will be higher than the threshold. The interval number of transmitter $AoD_t^{(1)}$ is $A_{q(\text{opt})}^{(t,1)} = "01"$, and $A_{q(\text{opt})}^{(t,1)}$ will be fed back to the transmitter. Then, the transmitter transmits the same training beam twice along the $AoD_t^{(1)}$ direction, and the receiver uses training beam 1 and training beam 2 to receive them in turn, and the received power is recorded at the same time. If only the received power of training beam 1 is greater than the threshold, the section number of receiver $AoD_r^{(1)}$ will be $A_{q(\text{opt})}^{(r,1)} = "10"$. Then, the next stage will be entered.

Layer 2: the transmitter transmits training beam 1 and training beam 2 in sequence in interval $AoD_r^{(1)} \Sigma[0^\circ, 45^\circ]$. Meanwhile, in section $AoD_r^{(1)} \Sigma[-90^\circ, -45^\circ]$, the receiver uses the quasi-omnidirectional mode to receive and records the received power. If the received power of the transmitter is higher than the threshold when the transmitter transmits training beam 1 and training beam 2, the section number of transmitter $AoD_t^{(2)}$ will be $A_{q(\text{opt})}^{(t,2)} = "11"$, and $A_{q(\text{opt})}^{(t,2)}$ will be fed back to the transmitter as $A_{q(\text{opt})}^{(t)}$. Then, the transmitter transmits the same training beam twice along the $AoD_t^{(2)}$ direction, and the receiver uses training beam 1 and training beam 2 to receive in turn. Meanwhile, the received power will be recorded. If the power received by using training beam 1 and training beam 2 is less than the threshold, the interval number of receiver $AoD_r^{(1)}$ will be $A_{q(\text{opt})}^{(r,1)} = "00"$, which is regarded as $A_{q(\text{opt})}^{(r)}$ and the search is completed.

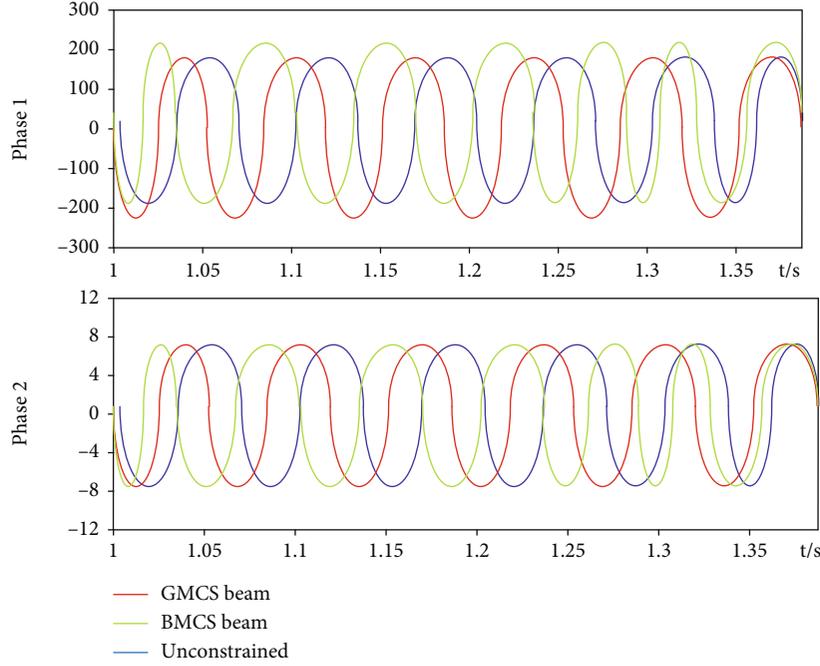


FIGURE 13: Target beam and training beam corresponding to 2-GMCS and 2-BMCS.

After the above two layers of search, the transmitter and receiver can use the best communication beam to communicate with $A_{q(\text{opt})}^{(t)}, A_{q(\text{opt})}^{(r)}$.

In order to facilitate the understanding of the proposed fast search model based on Gray mapping, given a general scenario, namely, $Q = 2^M$ (M is a positive integer), the pseudocode of the model is as follows.

4.4. Design of Target Beam. In order to design the training beam required by this model, the design method of the target beam needs to be given. According to the above analysis, only 2-GMCS and 3-GMCS are actually used, so only the target beams in these two cases need to be designed. The target beam of 2-GMCS is firstly analyzed. Taking QQ16 as an example, the coverage of the target beam of the 2-layer 2-GMCS at each stage is given, as shown in Figure 9.

The format in Figure 9 is $[C_{(m,n),d}]_l$, where $[m, n]$ represents the coverage of different target beams in the interval. The first bit represents target beam 1, and the second bit represents target beam 2; 0 represents no coverage, and 1 represents coverage. d represents $[m, n]$ the number $[-90^\circ + 4(d-1)w + w, -90^\circ + 4(d-1)W + 3W]$ of times such a combination appears in this layer, and l represents the number of layers.

It can be seen by observation that in the $[B_0^{(r)}, B_1^{(r)}, \dots, B_{K^{(r)}-1}^{(r)}]$ layer, 2 target beams can divide the layer area into $D = 2^l$ subintervals, the width of each subinterval is $W = \pi/D$, and the subintervals of each layer will be equally divided into $2^2 = 4$ parts in the next layer. Therefore, the AoD range covered by d target beam 1 of the l layer is $[-90^\circ + 4(d-1)w + w, -90^\circ + 4(d-1)W + 2W]$, and the AoD range covered

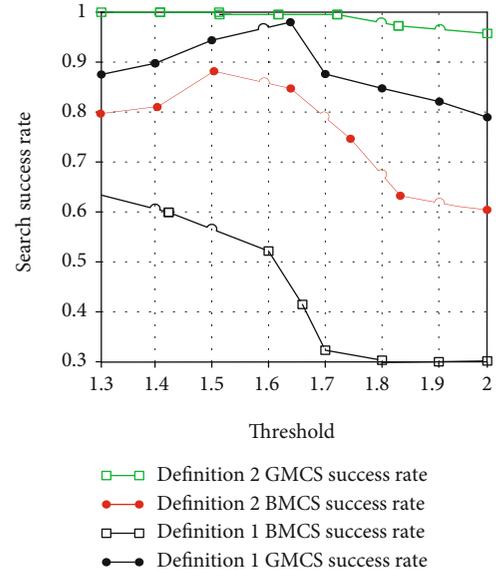


FIGURE 14: The relationship between the search success rate and the threshold of the two search models under 5G communication.

by d target beam 2 of the l layer is $[-90^\circ + 4(d-1)w + w, -90^\circ + 4(d-1)W + 3W]$.

Then, taking $\log_2 Q$ as an example give the coverage of the target beam of 3-GMCS at each stage, as shown in Figure 6.

The format in Figure 10 is $[C_{(x,y,z),d}]_l$, where (x, y, z) represents the coverage of different target beams in the interval, the first bit represents target beam 1, the second bit represents target beam 2, and the third bit represents target beam 3; 0 represents no coverage, 1 means coverage; d represents

the number of times that the combination of (x, y, z) appears in the layer, and l represents the number of layers.

Therefore, in the $[B_0^{(r)}, B_1^{(r)}, \dots, B_{K^{(r)}-1}^{(r)}]$ layer, 3 target beams can divide the layer interval into $D = 2^{3l}$ subintervals, the width of each subinterval is $W = \pi/D$, and the subintervals of each layer will be equally divided into $2^3 = 8$ parts in the next layer. The AoD covered by d target beam 1 of the l layer includes two parts: $-90^\circ + 8(d-1)W + W, -90^\circ + 8(d-1)W + 3W]$ and $-90^\circ + 8(d-1)W + 5W, -90^\circ + 8(d-1)W + 7W]$; the AoD range covered by d target beam 2 of the l layer is $90^\circ + 8(d-1)W + 2W, -90^\circ + 8(d-1)W + 6W]$, and the AoD range covered by d target beam 3 of the l layer is $90^\circ + 8(d-1)W + 2W, -90^\circ + 8(d-1)W + 6W]$.

When generating 2-GMCS and 3-GMCS target beams, it is only necessary to set the value of the AoD area covered by each target beam to 1 and the value of 0 for the uncovered AoD area. The beam response of the target beam is uniformly recorded as AF_{TB} , which is expressed as

$$AF_{TB}(\theta) = \frac{1, \theta \in [\alpha, \beta]}{0, \text{else}}, \quad (7)$$

where $[\alpha, \beta]$ represents the AoD area covered by the target beam.

For 2-GMCS and 3-GMCS, in addition to the above-mentioned target beams, it is also necessary to design target beams for each section of each layer to facilitate subsequent omnidirectional pattern search. When designing, each target beam only needs to cover an interval separately, that is, the beam response is 1 only in this interval.

4.5. Search Model Based on Beam Overlap in 5G Communication. The beam search model under 5G communication is studied. In order to make the model universal and widely used, it is necessary to study the beam search model under 5G communication. It should be pointed out that since this paper is for single-user single data stream research, Therefore, under 5G communication, only the strongest path needs to be found for communication, instead of multiple paths [32].

The width of the training beam based on the hierarchical search gradually narrows, and the path resolution capability gradually increases. Since there are multiple paths in 5G communication, if the original search model is still used and searched layer by layer, there will be a great probability of misjudgment. Taking 2-GMCS as an example briefly explain the phenomenon of misjudgment due to the existence of multiple paths, as shown in Figure 11.

Assuming there are 5 paths, the relative strength of the path is 7, 3, 6, 3, and 6, respectively (the relative strength here is only to indicate the relative strength of the path and has no actual physical meaning). In theory, the number of the strongest path that should be found is 1, and the corresponding interval is "10." However, since there are two paths 1 and 2 in the range covered by training beam 1, there are 3 paths 2, 3, and 4 in the range covered by training beam 2. During the search process, the paths within the coverage

TABLE 2: Number of clusters in the channel based on measurement and research in 5G communication.

	Urban micro cell (Umi)		Urban macro cell (Uma)	
	O2O	O2I	O2O	O2I
Cluster minimum	16	12	20	12
Cluster maximum	25	12	20	12
Carrier frequency	0.45-63 GHz		0.45-5 GHz	

of the training beam will overlap each other so that the sum of the relative intensities of the single coverage area of beam 1 is 7, the sum of the relative intensities of the overlapping coverage areas of beams 1 and 2 is 3, and the sum of the relative intensities of the separate coverage area of beam 2 is 9. It can be concluded that the interval number of the strongest path at this stage is the interval covered by beam 2 alone, that is, "01" instead of "10," which leads to misjudgment.

In order to effectively distinguish multipath, it is necessary to reduce the probability of misjudgment as much as possible. The K subsearch model is adopted in the first layer, that is, the entire interval is equally divided into K parts in the first-layer transceivers, and a traversal search is performed to determine the interval where the strongest energy path is located. The larger the K , the stronger the resolution, the lower the probability of misjudgment caused by multipath, and the higher the complexity of the search. Therefore, considering comprehensively, let the number of partitions in the first layer K be the smallest integer power of 2 not less than the number of multipaths, where K is not less than the number of multipaths to ensure resolution and K is an integer power of 2. It is to ensure the execution of the subsequent search model. For example, if the number of multipaths is 5, an 8-point search model can be used, as shown in Figure 12.

In the METIS (Mobile and Wireless Communications Enablers for the Twenty-Two (2020) Information Society) project established by the European Union in 2012, based on extensive measurement activities and analysis, computer simulation is combined to develop a new channel model. The number of clusters in different scenarios in the report is listed, as shown in Table 2.

The base station spacing in Umi (urban micro cell) is less than 1 km, and the base station antenna is located at the height of the roof. The distance between base stations in Uma (urban macro cell) is about 3 km, and the base station antenna is higher than the roof height. O2O (outdoor to outdoor) and O2I (outdoor to indoor) are two link topologies, which are outdoor to outdoor and outdoor to indoor, respectively. Therefore, refer to the data to design the appropriate K value for the first-level search.

After the K subsearch of the first layer is completed, the misjudgment caused by multipath can be avoided with great probability. Then, the search model in 5G communication can be used for subsequent searches. In order to facilitate the understanding of the proposed search model under 5G

Input: Q, M . The initial search interval of the transceiver is $\text{AoD}_t^{(0)} \Sigma[-90^\circ, 90^\circ]$, $\text{AoD}_r^{(0)} \Sigma[-90^\circ, 90^\circ]$

1. if $M = 2n$ (n is a positive integer) do
2. for $i = 1 : n$ do
3. The transmitter transmits training beams 1 and 2 in sequence according to 2-GMCS in the corresponding $\text{AoD}_t^{(i-1)}$ interval. The receiver adopts the quasi-omnidirectional mode to receive in the $\text{AoA}_r^{(i-1)}$ interval. The receiver compares the received power with the threshold to find the stage $A_{q(\text{opt})}^{(t,i)}$ and feeds it back to the transmitter, and the transmitter updates the search interval $\text{AoD}_t^{(i)}$ of the next stage
4. The transmitter transmits the same training beam twice along the $\text{AoD}_t^{(i)}$ direction, and the receiver uses training beams 1 and 2 to receive within the range of $\text{AoA}_r^{(i-1)}$. The receiver finds the optimal receiving beam number $A_{q(\text{opt})}^{(r,i)}$ at this stage by comparing the received power with the threshold and updates it to the search interval $\text{AoD}_r^{(i)}$ of the next stage
5. end for
6. else
7. Perform Step 3 and Step 4 ($n - 1$) times in sequence, and the output is marked as $[\alpha, \beta]$, $\text{AoA}_r^{(n-1)}$
8. The transmitter transmits training beams 1, 2, and 3 in sequence according to 3-GMCS in the corresponding AoD_t^{n-1} interval. The receiver adopts the quasi-omnidirectional mode to receive in the AoA_r range. Meanwhile, the receiver finds this stage $A_{q(\text{opt})}^{(t)}$ by comparing the received power with the threshold and feeds it back to the transmitter. The corresponding search interval is AoD_t^n
9. The transmitter transmits the same training beam three times along the AoD_t^n direction, and the receiver uses training beams 1, 2, and 3 to receive within the range of AoD_r^{n-1} . The receiver finds the optimal receiving beam number $A_{q(\text{opt})}^{(r)}$ at this stage by comparing the received power with the threshold and completes the search
10. end if

Output: $A_{q(\text{opt})}^{(t)}$, $A_{q(\text{opt})}^{(r)}$

ALGORITHM 1: The fast beam search model based on GMCS in the LoS scenario.

communication, the pseudocode of the model is given as follows.

5. Simulation Results and Analysis

This paper considers a single transmitter. In a 5G communication with a single receiver, both the transmitter and the receiver use 8 half-wave uniform antenna arrays, and use the N ($N = 16$)-phase codebook to generate 16 communication beams. In 5G communication, the number of antennas at the transceiver end is 16, and a 32-phase codebook is used to generate 32 communication beams. The number of multipaths in the channel is 5, and the angles and strengths of different paths are random. Using the S-V channel model, AoDs at the transmitter and AoAs at the receiver are uniformly distributed on $[-\pi/2, \pi/2]$. The carrier frequency of the system is 28 GHz, the bandwidth is 100 MHz, and the path loss factors in 5G communication and 5G communication are 1.73 and 3.83, respectively. In the simulation, AoDs at the transmitting end and AoAs at the receiving end are randomly generated to generate the corresponding channel matrix. Then, the best communication beam pair is found through beam search at the transceiver end. All simulation results in this paper are the average value of 5000 random channels.

5.1. Target Beam and Training Beam Simulation Results. For the scenario of 8 antennas and 16 communication beams, a two-stage 2-GMCS is used to design the training beam, and the design result of a two-stage 2-BMCS is also given, as shown in Figure 13.

For the target beams of the two models in Figure 13, the solid line represents the generated training beam. The design results of 2-BMCS are put on the $[-3\pi/2, -\pi/2]$ interval, and both are in the $[-\pi/2, \pi/2]$ interval. The target beam and training beam of the first stage, it can be seen that 2-GMCS training beam 1 and training beam 2 and 2-BMCS training beam 2 are the same, because the intervals covered by these training beams are consistent, that is, the continuous $\pi/2$ interval. The target beam 1 of the 2-BMCS has two main lobes, and the training beam is not steep enough at the edge of the main lobe. The target beam and training beam of the second stage show that the 2-GMCS-based training beam has a very good effect, the main lobe edge is steep, and the side lobe energy is very low. The training beam effect based on 2-BMCS is poor. The two main lobes of the training beam 1 are asymmetric in the corresponding coverage interval, and the side lobes have high energy, which will affect the success rate of the beam search.

5.2. Search Success Rate. This paper gives two definitions of the search success rate. The first is to find the optimal communication beam, and the second is to find the optimal communication beam or suboptimal communication beam. In the simulation, through 5000 random simulations, the communication beams searched each time are compared with the actual optimal or suboptimal communication beams to calculate the search success rate. In practical applications, for specific scenarios, the model requires a large number of simulations to obtain the optimal threshold range. Then, in the subsequent search of this scene, it can be directly applied without setting the threshold again.

Input: Q, M, N_{ray} , the initial search interval of the transceiver $\text{AoD}_t^{(0)} \in [-90^\circ, 90^\circ]$
 Output: $A_{q(\text{opt})}^{(t)}, A_{q(\text{opt})}^{(r)}$

1. Begin
2. The value of K is determined according to N_{ray} , and a codebook based on K ($K = 2^m$) score search is designed
3. The transceiver searches through all codebook combinations, and the receiver records the corresponding power. Through K^2 searches, it is determined that the region of the strongest path corresponding to the transceiver is $\text{AoD}_t^{(1)}$ and $\text{AoA}_r^{(1)}$
4. Algorithm 1 is executed to find $A_{q(\text{opt})}^{(t)}$ and $A_{q(\text{opt})}^{(r)}$
5. End

ALGORITHM 2: The paper swarm hybrid precoding algorithm [33].

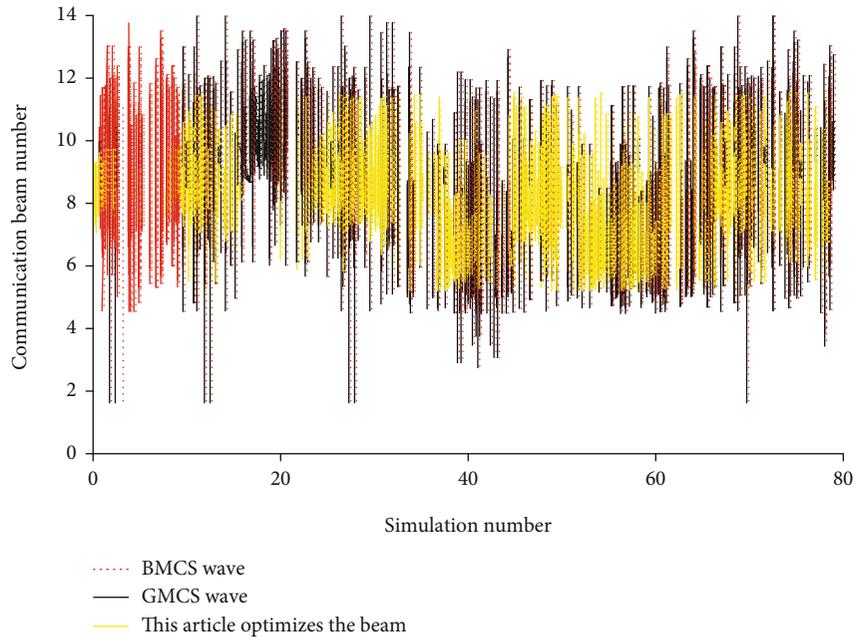


FIGURE 15: Partial search results of two search models under 5G communication.

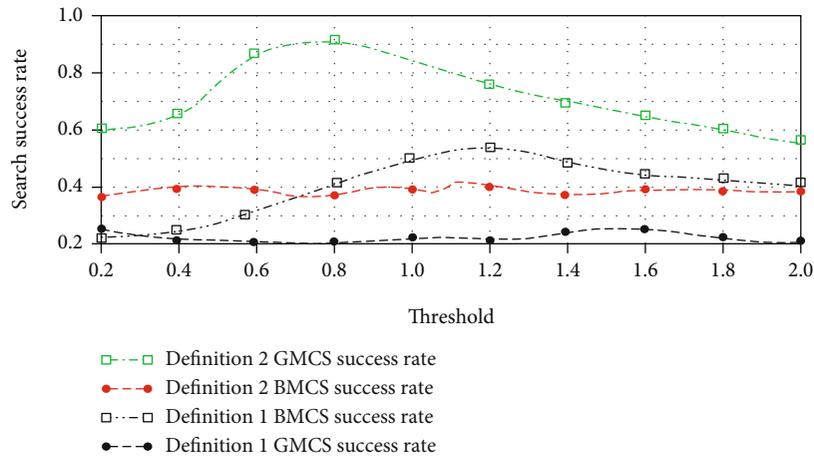


FIGURE 16: The relationship between the search success rate and the threshold of the two search models under 5G communication.

The simulation results of the proposed model under 5G communication are given. Assuming that the number of antennas at the transceiver end is 8 and the number of communication beams is 16, the entire search model is divided into two-stage 2-GMCS (2-BMCS). The relationship between the search success rate and the threshold of the two coding models is given, as shown in Figure 10.

The horizontal axis represents the threshold, and the vertical axis represents the search success rate of different models. According to Figure 10, the following conclusions can be drawn:

- (1) Based on the GMCS and BMCS search models, the search success rate under definition 2 is higher than the search success rate under definition 1, and the optimal threshold points under the two definitions are not necessarily the same
- (2) Under the two definitions, the GMCS-based search success rate is close to 100%, which is much higher than that of the BMCS-based search model. This is caused by the excessive number of main lobes of the BMCS training beam and the unsatisfactory actual shape
- (3) The search model based on BMCS is more susceptible to the influence of the threshold than the search model based on GMCS. Because BMCS-based training beams are not symmetrical in the coverage interval, multiple training beams will be affected when the threshold changes

Furthermore, the thresholds of the two models are both set as the optimal thresholds, and simulations are carried out for 5000 times. Next, 80 simulations are taken out to observe the search results of the two models, as shown in Figure 15.

In Figure 15, the horizontal axis represents the simulation serial number, which is randomly selected in all simulations and then output according to the selected order. The vertical axis represents the number of the communication beam.

For the two models, when the searched point overlaps with the blue box, it means that the optimal communication beam has been found. The search model based on BMCS seeks out 1, but no suitable communication beam is found. It can also be seen from the figure that the success rate of the BMCS-based search model is lower than that of the GMCS-based search model.

Finally, the simulation results of the proposed model under 5G communication are given. When searching, the 8-point search model is used to perform the first-level search to avoid misjudgments caused by multipath as much as possible, and 2-GMCS (2-BMCS) is used for the second-level search. The relationship between the search success rate and the threshold of the two coding models is given, as shown in Figure 16.

Comparing with Figure 11, it can be seen that the search success rate of the two search models under 5G communication has dropped significantly. There are two reasons for this

phenomenon. On the one hand, the 8-point search model cannot completely avoid the misjudgment phenomenon caused by multipath, that is, there are still multiple paths located in a training beam at the same time. According to the simulation data, the misjudgment probability of 8-point search is 10% on average. On the other hand, due to the relatively large number of communication beams, the width of the second-layer training beam is relatively narrow. Therefore, the probability of obtaining the suboptimal beams on the left and right sides of the optimal beam will be greatly increased, which also makes the success rate of the search model based on GMCS much higher than that of the search model based on BMCS.

6. Conclusion

This paper mainly focuses on the beamforming technology in the millimeter-wave Massive MIMO of the railway system and considers the design of a low-complexity, approximately optimal beamforming model for the railway system. In the digital-analog hybrid architecture, the main research is the hybrid precoding algorithm with decoupling of transceivers and the hybrid precoding algorithm with transceivers, which can provide a variety of compromise models between performance and complexity. In the analog beamforming structure, a multilayer beam search model based on the main lobe overlap of Gray mapping, which is suitable for single-path and multipath scenarios, is proposed; it can obtain a higher search success rate with the lowest complexity. The proposed model provides a certain theoretical basis for the application of beamforming technology in millimeter-wave Massive MIMO systems. The algorithm in this paper needs a certain amount of training data to achieve, but in practice, the training data cannot be obtained directly, and it is difficult to measure artificially. The follow-up study can use incremental SVM to evaluate each classification result and add it to the training data, so as to achieve the goal of classification with a very small number of samples.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authors' Contributions

The authors of the manuscript "Transceiver Decoupling of Multivariate Symmetric Hybrid Precoding Based on 5G" declare the following contributions to the creation of the manuscript: Qiuliang Zhang—conceptualization, resources, and methodology; Zhike Zhang—supervision and project administration; Zhihui Wang—original draft and writing—review and editing; Dong Yang—resources and review.

References

- [1] W. Zou, M. Cui, and C. Guo, "Fast beam search for massive MIMO based on mainlobe overlapping state of training beam," *IEEE Access*, vol. 7, pp. 66007–66019, 2019.
- [2] K. N. R. S. Prasad, E. Hossain, and V. K. Bhargava, "Energy efficiency in massive MIMO-based 5G networks: opportunities and challenges," *IEEE Wireless Communications*, vol. 24, no. 3, pp. 86–94, 2017.
- [3] S. A. Busari, K. M. Huq, S. Mumtaz, L. Dai, and J. Rodriguez, "Millimeter-wave massive MIMO communication for future wireless systems: a survey," *IEEE Communication Surveys and Tutorials*, vol. 20, no. 2, pp. 836–869, 2018.
- [4] Z. Hou, Y. Zhou, L. Tian, J. Shi, Y. Li, and B. Vucetic, "Radio environment map-aided Doppler shift estimation in LTE railway," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 5, pp. 4462–4467, 2017.
- [5] C. Zhang, G. Wang, R. He, and Y. Zou, "A Doppler shift estimator for millimeter-wave communication systems on high-speed railways," in *2018 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 227–231, Beijing, China, 2018.
- [6] G. Yu, L. Tian, J. Zhang, and C. Wang, "The method to implement 5g channel model with spatial consistency," in *2018 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 736–740, Beijing, China, 2018.
- [7] M. Peter, K. Sakaguchi, S. Jaeckel et al., "Measurement campaigns and initial channel models for preferred suitable frequency ranges," *Deliverable D2*, vol. 1, p. 160, 2016.
- [8] P. Kyosti, I. Carton, A. Karstensen, W. Fan, and G. F. Pedersen, "Frequency dependency of channel parameters in urban LOS scenario for mmwave communications," in *2016 10th European Conference on Antennas and Propagation (EuCAP)*, pp. 1–5, Davos, Switzerland, 2016.
- [9] Y. Hu, H. Li, Z. Chang, and Z. Han, "Scheduling strategy for multimedia heterogeneous high-speed train networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3265–3279, 2017.
- [10] N. Sun, Y. Zhao, L. Sun, and Q. Wu, "Distributed and dynamic resource management for wireless service delivery to high-speed trains," *IEEE Access*, vol. 5, pp. 620–632, 2017.
- [11] G. A. Akpakwu, B. J. Silva, G. P. Hancke, and A. M. Abu-Mahfouz, "A survey on 5G networks for the Internet of Things: communication technologies and challenges," *IEEE Access*, vol. 6, pp. 3619–3647, 2018.
- [12] K. Venugopal, A. Alkhateeb, N. Gonzalez Prelcic, and R. W. Heath, "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1996–2009, 2017.
- [13] E. Koyuncu, X. Zou, and H. Jafarkhani, "Interleaving channel estimation and limited feedback for point-to-point systems with a large number of transmit antennas," *IEEE Transactions on Wireless Communications*, vol. 17, no. 10, pp. 6762–6774, 2018.
- [14] Z. Xiao, P. Xia, and X. Xia, "Channel estimation and hybrid precoding for millimeter-wave MIMO systems: a low-complexity overall solution," *IEEE Access*, vol. 5, pp. 16100–16110, 2017.
- [15] M. Kokshoorn, H. Chen, P. Wang, Y. Li, and B. Vucetic, "Millimeter wave MIMO channel estimation using overlapped beam patterns and rate adaptation," *IEEE Transactions on Signal Processing*, vol. 65, no. 3, pp. 601–616, 2017.
- [16] Y. Wang and W. Zou, "Low complexity hybrid precoder design for millimeter wave MIMO systems," *IEEE Communications Letters*, vol. 23, no. 7, pp. 1259–1262, 2019.
- [17] Y. Wang and W. Zou, "Hybrid precoding design for millimeter wave systems with the partially-connected structure," *IET Communications*, vol. 14, no. 4, pp. 561–567, 2020.
- [18] W. Huang, Q. Si, and M. Jin, "Alternating optimization based low complexity hybrid precoding in millimeter wave MIMO systems," *IEEE Communications Letters*, vol. 24, no. 3, pp. 635–638, 2020.
- [19] B. Srinivas, K. Mawatwal, D. Sen, and S. Chakrabarti, "An iterative semi-blind channel estimation scheme and uplink spectral efficiency of pilot contaminated one-bit massive MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 7854–7868, 2019.
- [20] E. Nayebi and B. Rao, "Semi-blind channel estimation for multiuser massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 66, no. 2, pp. 540–553, 2018.
- [21] A. Liu, V. K. N. Lau, M. L. Honig, and R. F. Compressive, "Compressive RF training for massive MIMO with channel support side information," *IEEE Transactions on Wireless Communications*, vol. 18, no. 7, pp. 3628–3641, 2019.
- [22] C. Hu, X. Wang, L. Dai, and J. Ma, "Partially coherent compressive phase retrieval for millimeter-wave massive MIMO channel estimation," *IEEE Transactions on Signal Processing*, vol. 68, pp. 1673–1687, 2020.
- [23] 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.
- [24] G. Yue, D. Yu, H. Qiu, K. Guan, L. Yang, and Q. Lv, "Measurements and ray tracing simulations for non-line-of-sight millimeter-wave channels in a confined corridor environment," *IEEE Access*, vol. 7, pp. 85066–85081, 2019.
- [25] E. Vlachos, G. C. Alexandropoulos, and J. Thompson, "Wideband MIMO channel estimation for hybrid beamforming millimeter wave systems via random spatial sampling," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 5, pp. 1136–1150, 2019.
- [26] M. Y. Kang, J. Y. Lee, and S.-C. Kim, "Analysis of channel characteristics for outdoor 28 GHz millimeter wave channel," in *2019 25th Asia-Pacific Conference on Communications (APCC)*, pp. 388–391, Ho Chi Minh City, Vietnam, 2019.
- [27] J. M. García-Loygorri, C. Briso, and A. P. Yuste, "Saleh-Valenzuela modelization and clustering information for a mmwave intra-train channel," in *2019 13th European Conference on Antennas and Propagation (EuCAP)*, pp. 1–5, Krakow, Poland, 2019.
- [28] Z. Xiao, H. Dong, L. Bai, P. Xia, and X. G. Xia, "Enhanced channel estimation and codebook design for millimeter-wave communication," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 9393–9405, 2018.
- [29] Q. Qin, L. Gui, P. Cheng, and B. Gong, "Time-varying channel estimation for millimeter wave multiuser MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 9435–9448, 2018.
- [30] S. Bazzi, S. Stefanatos, L. Le Magoarou et al., "Exploiting the massive MIMO channel structural properties for minimization of channel estimation error and training overhead," *IEEE Access*, vol. 7, pp. 32434–32452, 2019.

- [31] Y. Liao, Y. Hua, and Y. Cai, "Deep learning based channel estimation algorithm for fast time-varying MIMO-OFDM systems," *IEEE Communications Letters*, vol. 24, no. 3, pp. 572–576, 2020.
- [32] Z. Wang, M. Li, Q. Liu, and A. L. Swindlehurst, "Hybrid precoder and combiner design with low-resolution phase shifters in mmWave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 2, pp. 256–269, 2018.
- [33] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, 2014.