Multiple Dimensional Scaling Hybrid Precoding in Millimeter Wave MIMO System

Jiaqi Su[1], Wenbin Zhang[1,2](✉), Liyan Zhang[1], Jinwei Huang[1], and Shaochuan Wu[1]

[1] Harbin Institute of Technology, Harbin 150001, People's Republic of China
zwbgxy1973@hit.edu.cn

[2] Science and Technology on Communication Networks Laboratory, Shijiazhuang 050050, Hebei, People's Republic of China

Abstract. In most wireless systems, the channel transmission loss of millimeter wave signals is often greater than that of traditional signals. Hybrid precoding can leverage large antenna arrays to compensate this loss and further improve the spectral efficiency of millimeter wave. This paper explores the sparsity of mmWave channel—only a few main paths are useful for the precoding procedure. This sparsity makes it possible to use a limited feedback in channel estimation. We propose a novel hybrid precoding scheme in this paper, which utilizes the main vector of the channel. We call this scheme Multiple Dimensional Scaling (MDS) hybrid precoding. On the one hand, comparing with many traditional hybrid precoding schemes with full channel information feedback, our scheme can decrease feedback overheads significantly. On the other hand, comparing with some hybrid precoding with limited channel feedback information, our scheme can improve spectral efficiency. Moreover, the simulation results show that the system spectral efficiency increases with the number of data streams significantly.

Keywords: Millimeter wave communications · Hybrid precoding · Limited feedback · Multiple Dimensional Scaling

1 Introduction

The principle of millimeter wave (mmWave) communication to achieve high data transmission rate is to utilize the potential available large bandwidth in the high frequency band [1–4]. However, because of the high frequency of mmWave and the small wavelength, the path loss is much higher than that of microwave band. The solution to compensate for the loss caused by the increase of frequency, is to

Supported by National Nature Science Foundation of China (NSFC) under Grant 62071148.

Q. Guo et al. (Eds.): WiSATS 2021, LNICST 410, pp. 392–404, 2022.
https://doi.org/10.1007/978-3-030-93398-2_38

use precoding technology and multi-antenna technology. Future cellular network base stations will deploy hundreds of antennas to form mmWave massive MIMO. However, the design of precoding matrix and combining matrix in mmWave MIMO system is not as simple as that in the low frequency scheme. It is mainly due to different hardware constraints, higher cost and higher power consumption. Therefore, adjusting the number of antennas and radio frequency (RF) chains, reducing hardware complexity and improving energy efficiency become the main problems in Massive MIMO technology [5–7]. Thus, research on the precoding scheme in mmWave MIMO systems has become a very important field.

In this context, analog-digital hybrid precoding is proposed, which achieves the purpose of reducing the number of RF chains by dividing the precoding process into analog domain and digital domain [2,3]. Digital precoding can achieve higher accuracy and analog precoding can reduce the originally high power consumption, which makes hybrid precoding have better performance. It can achieve a compromise and balance between hardware complexity and system performance. In analog-digital hybrid precoding, the transmitter needs to know the channel information. In FDD (frequency division duplex) mode, the mobile station estimates the downlink channel by receiving the pilot transmitted by the base station, and then transmits the channel information to the base station. In [2,3], the hybrid architecture under lower frequency is studied. In mmWave systems, the concept of hybrid precoding is similar to that of the lower frequency. In [8], considering the sparsity of mmWave channel, the precoder and combiner of mmWave system with large antenna array are combined to design the system. This precoding problem is transformed into a sparse reconstruction problem by utilizing the spatial sparsity of mmWave channels. In this paper, based on the principle of base tracking, a system is designed that can approximately achieve the optimal result under all-digital precoding. [9] proposes a low complexity reconstruction algorithm based on [8], which avoids the process of matrix inversion by querying the orthogonal codebook, thereby reducing the computational complexity. Similar to [8], in [10], the author proposes a greedy algorithm to design hybrid analog/digital precoders, which has lower complexity and is approximately orthogonal. The principle of the algorithm is to greedily select the RF beamforming vectors through the Gram-Schmidt orthogonalization process. Besides, [14] proposes a situation where only partial channel information is known. Some other heuristic algorithms that do not include the orthogonal matching pursuit process can also be seen in [12,13] to design hybrid precoding, which need to meet the condition of having perfect channel information at the transmitter. In [14] the combined channel matrix is fed back to the base station. The base station uses a simplified effective channel matrix with little training and feedback overhead to design the hybrid precoding. The performance analysis of the above mentioned algorithms is based on large dimensional regime and single-path channels.

This paper proposes an algorithm that can use limited channel feedback to achieve hybrid precoding, which compresses the channel information at the receiver and then feedback the compressed information to reduce the feedback overhead and power consumption. Then, the base station uses the compressed feedback information to design precoding. The simulation results show that although we have reduced the computational complexity, the performance of the algorithm in some cases has not been affected.

The following text structure of the paper is as follows. Section 2 briefly introduces the system model of mmWave MIMO hybrid precoding. Section 3 analyzes the hybrid precoding problem. Section 4 explains the proposed Multiple Dimensional Scaling hybrid precoding. Section 5 shows the results of simulation at the achievable rate. At last, the conclusion of the article is described in Sect. 6.

Notation: boldface uppercase, boldface lowercase, and lowercase letter $\mathbf{A}$, $\mathbf{a}$, a denote a matrix, vector, and scalar variable respectively. $\mathbf{A}^T$, $\mathbf{A}^*$ and $\mathbf{A}^{-1}$ represent the transpose, the conjugate transpose and the inverse of a matrix, respectively. $|\ |$ denotes absolute value, *tr()* denotes trace of a matrix, *diag* denotes a diagonal matrix.

2 System Model

2.1 System Model

First, Fig. 1 shows the mm Wave system model, where a base station (BS) with N_t antennas and N_{tRF} RF chains will establish communication relationship with a single mobile station (MS) with N_r antennas and N_{rRF} RF chains. The communication process between BS and MS utilizes N_s data streams, and the number N_s meets the condition that $N_s \leq N_{tRF} \leq N_t$ and $N_s \leq N_{rRF} \leq N_r$.

At the transmitter, we first use a $N_{tRF} \times N_s$ digital precoding matrix $\mathbf{F}_{BB}$ to precode the $N_s \times 1$ data symbols $\mathbf{s}$, then the symbols will be applied to $N_t \times N_{tRF}$ RF precoding $\mathbf{F}_{RF}$, such that $E[\mathbf{s}\mathbf{s}^*] = \frac{1}{N_s}\mathbf{I}_{N_s}$. Therefore, the transmitted discrete-time signal can be expressed by $\mathbf{x} = \mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{s}$. Since we use analog phase shifters to realize $\mathbf{F}_{RF}$, it is required to meet the condition that $({\mathbf{F}_{RF}}^{(i)}{\mathbf{F}_{RF}}^{(i)*})_{l,l} = \frac{1}{N_t}$, where $()_{l,l}$ represents the l-th diagonal element of a matrix. Due to the total power limitations of the transmitters, we must normalize $\mathbf{F}_{BB}$ to satisfy $\| \mathbf{F}_{BB}\mathbf{F}_{RF} \|_F^2 = N_s$.

We discuss a narrowband block-fading propagation channel as in [4–7] for simplicity, which generates a received signal

$$\mathbf{y} = \sqrt{\rho}\mathbf{H}\mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{s} + \mathbf{n} \tag{1}$$

where $\mathbf{y}$ denotes the $N_r \times 1$ received vector, $\mathbf{H}$ represents the $N_r \times N_t$ channel matrix which meets the condition that $E[\| \mathbf{H} \|_F^2] = N_t N_r$, ρ is the average received power, and $\mathbf{n}$ indicates the noise vector of i.i.d $\mathbb{CN}(0, \sigma_n^2)$. The received signal at the MS after combination can be represented as the following formula

$$\tilde{\mathbf{y}} = \sqrt{\rho}\mathbf{W}_{BB}^*\mathbf{W}_{RF}^*\mathbf{H}\mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{s} + \mathbf{W}_{BB}^*\mathbf{W}_{RF}^*\mathbf{n} \tag{2}$$

where $\mathbf{W}_{RF}$ denotes the $N_r \times N_{rRF}$ RF combining matrix, and $\mathbf{W}_{BB}$ represents the $N_{rRF} \times N_s$ baseband combining matrix. Like the transmitter, the receiver need to satisfy the power constrain such that $(\mathbf{W}_{RF}{}^{(i)}\mathbf{W}_{RF}{}^{(i)*})_{l,l} = \frac{1}{N_r}$. When the signal transmitted on the mmWave channel satisfies the Gaussian distribution, the system spectral efficiency can be expressed by

$$R = log_2(|\mathbf{I}_{N_s} + \frac{\rho}{N_s}\frac{\mathbf{R}_s}{\sigma_n^2\mathbf{W}_{BB}^*\mathbf{W}_{RF}^*\mathbf{W}_{RF}\mathbf{W}_{BB}}|) \tag{3}$$

where $\mathbf{R}_s = \mathbf{W}_{BB}^*\mathbf{W}_{RF}^*\mathbf{H}\mathbf{F}_{RF}\mathbf{F}_{BB} \times \mathbf{F}_{BB}^*\mathbf{F}_{RF}^*\mathbf{H}^*\mathbf{W}_{RF}\mathbf{W}_{BB}$.

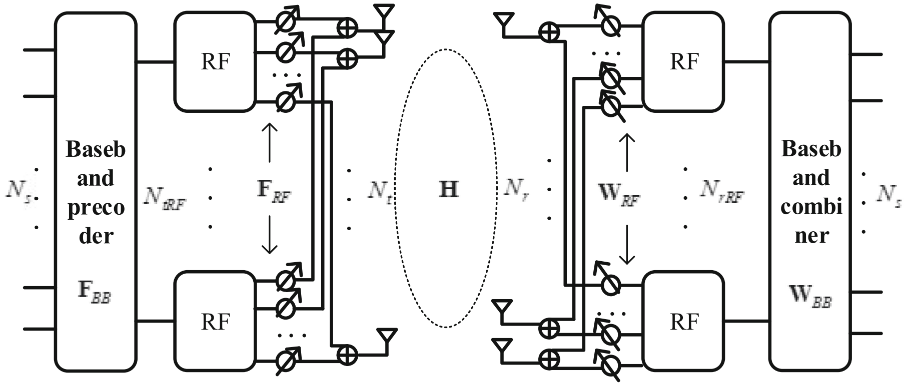

Fig. 1. A mmWave downlink system model with digital baseband precoding followed by a module using RF phase shifters to decrease radio frequency chains.

2.2 Channel Model

Millimeter wave transmission loses some degrees of freedom, which will cause limited spatial selectivity and limited scattering. We model a 2D channel through a widely used geometric mmWave channel model, which is a narrowband clustered channel expression in the light of the extended Saleh-Valenzuela model. The advantage of this model is that it enables us to accurately know the mathematical structure of the mmWave channel [8].

According to the narrowband clustered channel model, the channel matrix **H** is the set of all scattered transmission paths of N_{cl} scattering clusters, and each N_{ray} propagation paths of each scattering cluster is the composition of the channel matrix. Thus, the narrowband discrete-time channel matrix **H** can be expressed as the following formula

$$\mathbf{H} = \gamma \sum_{i,l} \alpha_{il}\mathbf{a}_r(\phi_{il}^r, \theta_{il}^r)\mathbf{a}_t^*(\phi_{il}^t, \theta_{il}^t) \tag{4}$$

where γ represents a normalization factor and equals to $\sqrt{\frac{N_r N_t}{N_{cl}N_{ray}}}$, α_{il} is the complex gain of i.i.d $\mathbb{CN}(0, \sigma_{\alpha,i}^2)$, which represents the gain of the lth ray in the ith scattering cluster, $\sigma_{\alpha,i}^2$ is the average power of the ith cluster. The average cluster powers satisfy the equation that $\sum_{i=1}^{N_{cl}} \sigma_{\alpha,i}^2 = \gamma$. For the lth ray in the ith scattering cluster, $\phi_{il}^r, \theta_{il}^r$, $\phi_{il}^t, \theta_{il}^t$ stand for its azimuth/elevation angles of arrival and departure respectively. The vector $\mathbf{a}_r(\phi_{il}^r, \theta_{il}^r)$ and $\mathbf{a}_t(\phi_{il}^t, \theta_{il}^t)$ denote the normalized receive and transmit array response vectors at an azimuth/elevation angle of $\phi_{il}^r, \theta_{il}^r$, $\phi_{il}^t, \theta_{il}^t$ respectively. We make an assumption that the azimuth and elevation angles of departure $\phi_{il}^t, \theta_{il}^t$, are distributed with a uniformly-random mean cluster angle of ϕ_i^t, θ_i^t and angular spread of σ_{ϕ^t}, σ_{θ^t} in the cluster i respectively. The azimuth and elevation angles of arrival $\phi_{il}^r, \theta_{il}^r$ are again randomly distributed with mean cluster angles of ϕ_i^r, θ_i^r and angular spreads σ_{ϕ^r}, σ_{θ^r}.

$\mathbf{a}_r(\phi_{il}^r, \theta_{il}^r)$ and $\mathbf{a}_t(\phi_{il}^t, \theta_{il}^t)$ are the receiving and transmitting antenna array response vectors of MS and BS respectively, and have nothing to do with the antenna element properties. The following two illustrative examples of generally-accepted antenna arrays can both be applied to the algorithms and simulation results in this paper. We give the conclusion that the array response vector for an N-element uniform linear array (ULA) on the y-axis can be expressed as the following formula

$$\mathbf{a}_{ULA}(\phi) = \frac{1}{N}[1, e^{jkdsin(\phi)}, ..., e^{j(N-1)kdsin(\phi)}]^T \tag{5}$$

where $k = \frac{2\pi}{\lambda}$, λ denotes the signal wavelength, and d represents the inter-element spacing. It should be pointed out that we do not consider θ in the discuss of a ULA, because the array's response remains constant in the elevation domain. Then, consider another antenna array distribution situation, which is a uniform planar array (UPA) that has W and H elements on the y and z axes respectively in the yz-plane, the array response vector is expressed by

$$\begin{aligned} \mathbf{a}_{UPA}(\phi, \theta) = \frac{1}{N}[1, ..., e^{jkd(msin(\phi)sin(\theta)+ncos(\theta))} \\ ..., e^{jkd((W-1)sin(\phi)sin(\theta)+(H-1)cos(\theta))}]^T \end{aligned} \tag{6}$$

where the parameters meet the condition that $0 \leq m \leq W$, $0 \leq n \leq H$, and $N = WH$. Through the above analysis, we can know that uniform planar arrays (UPA) have the following advantages, the UPA has smaller antenna array dimensions, has the ability to perform the beamforming process in the elevation domain, and packs more components in a reasonably sized antenna array. Therefore, it is of great importance to study uniform planar arrays in mmWave beamforming.

3 Problem Formulation

In order to improve the system spectral efficiency and reduce computational complexity, the goal of this paper is to propose a feasible hybrid precoding algorithm. In Sect. 2, we have introduced the system model in details. In order to make the subsequent research more simple, we assume that the optimal nearest neighbor decoding algorithm adopted by the receiver satisfies the following conditions. First, the received signal is N_s-dimensional, and secondly, it is transmitted by all-digital hardware. The above characteristics allow us to decouple the design of the transceiver and focus on the design of hybrid precoding. The amount of mutual information of the system after the hybrid precoding process can be maximized [10], which we define as

$$I(\mathbf{F}_{RF}, \mathbf{F}_{BB}) = log_2(|\mathbf{I}_{N_s} + \frac{\rho}{N_s \sigma_n^2} \mathbf{H}\mathbf{F}_{RF}\mathbf{F}_{BB} \times \mathbf{F}_{BB}^* \mathbf{F}_{RF}^* \mathbf{H}^*|) \tag{7}$$

where $SNR = \frac{\rho}{\sigma_n^2}$. Because the all-digital hardware design scheme is not feasible in practice, the design problem of hybrid precoding will become complicated and need to be reconsidered. However, the hybrid precoding design ideas proposed in this paper can directly construct a hybrid combining matrix. The design process of $\mathbf{W}_{RF}$ and $\mathbf{W}_{BB}$ is almost similar to the process of designing hybrid precoders. Due to the limitation of the length of the article, we will not repeat them here. We assume that the RF beamforming vector is directly obtained through the F_{RF} codebook, which meets the constraints of RF hardware design, then the maximum mutual information under the hybrid precoding model proposed in the article is given by the following formula

$$\begin{aligned} (\mathbf{F}_{RF}^*, \mathbf{F}_{BB}^*) &= \underset{\mathbf{F}_{RF}, \mathbf{F}_{BB}}{argmax}\, I(\mathbf{F}_{RF}, \mathbf{F}_{BB}) \\ s.t. \quad & \mathbf{F}_{RF} \in \mathscr{F}_{RF} \\ & \| \mathbf{F}_{BB}\mathbf{F}_{RF} \|_F^2 = N_s \end{aligned} \tag{8}$$

where F_{RF} represents the RF precoding domain, which is, the set of $N_t \times N_{tRF}$ matrices which have constant-magnitude elements.

4 Multiple Dimensional Scaling Hybrid Precoding

Generally, hybrid precoding systems need to obtain channel information to realize precoding design at the transmitting end. In large-scale antenna settings, channel information fed back from the receiver is huge, which will bring high feedback overhead and power consumption. Therefore, we propose a multi-dimensional scaling hybrid precoding scheme, which reduces the channel feedback by reducing the information dimension. The limited feedback scheme is shown in Fig. 2.

The commonly used dimension of information is tens of thousands. In addition, many computing methods involve distance calculation, and high-dimensional space will bring great trouble to distance calculation. When the dimension of information is high, it is no longer easy to calculate the inner product. As the number of antennas continues to increase, the computational complexity of the algorithm increases too. These computational obstacles in high dimensions are called "dimension disaster".

An important way to alleviate the "dimension disaster" is to dimension-reduction. The specific method is to transform the original high-dimensional

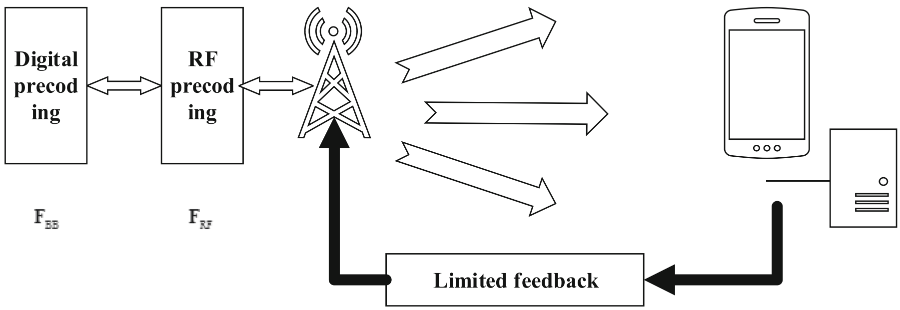

Fig. 2. A mmWave downlink system model with mobile station has a limited feedback channel to the base station.

attribute space into a low-dimensional "subspace" through some mathematical transformation. In this space, the sample density is greatly increased, and the distance calculation becomes easier. The principle of dimension-reduction is that: in many cases, the observed data samples are high-dimensional, however, only a low-dimensional distribution are closely related to task, that is, a low-dimensional embedding in high-dimensional space. As shown in Fig. 3, a three-dimensional graphic is simplified to a two-dimensional graphic. Sample points are easier to calculate in this low-dimensional subspace. But dimension-reduction also brings a defect, which means it loses some useful information of system and leads to the degradation of system performance.

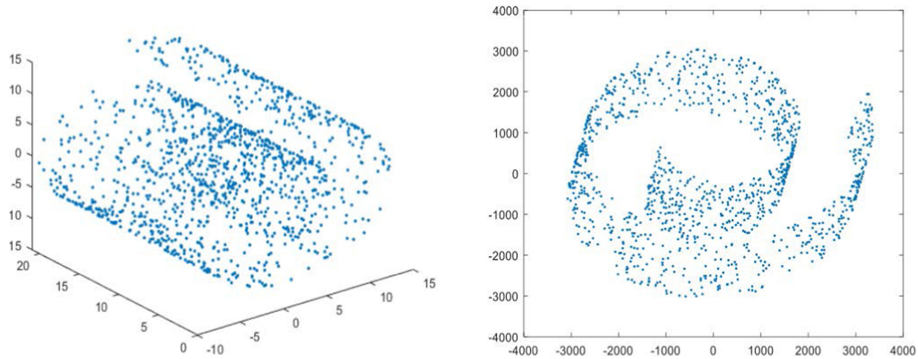

Fig. 3. A brief introduction to Multiple Dimensional Scaling (MDS), using MDS to compress data from 3D to 2D.

The starting point of this method—Multiple Dimensional Scaling (MDS) is to keep the distance between different points in the original data space unchanged as far as possible in the transformed space. Assuming m samples, it can be seen here as a channel vector whose distance matrix in the original channel space is

$\mathbf{D} \in \mathbb{R}^{m \times m}$, and it's ith row, jth column element $dist_{ij}$ is between the channel vector $\mathbf{h}_i$ and $\mathbf{h}_j$, that is

$$\parallel \mathbf{h}_i - \mathbf{h}_j \parallel = dist_{ij} \tag{9}$$

Our goal is to obtain the representation of samples in the d' dimension space $\mathbf{Z} \in \mathbb{R}^{d' \times m}$, $d' \leq d$, and the Euclidean distance of any two samples in the low-dimensional space equals the distance in the original space, that is, $\parallel \mathbf{z}_i - \mathbf{z}_j \parallel = dist_{ij}$. Let $\mathbf{B} = \mathbf{Z}^T\mathbf{Z} \in \mathbb{R}^{m \times m}$, which is the inner product matrix $\mathbf{B}$ of the reduced dimension sample $b_{ij} = \mathbf{z}_i^T\mathbf{z}_j$, we have

$$dist_{ij}^2 = \parallel \mathbf{z}_i \parallel^2 + \parallel \mathbf{z}_j \parallel^2 - 2\mathbf{z}_i^T\mathbf{z}_j \tag{10}$$

easily, we can obtain:

$$\sum_{i=1}^{m} dist_{ij}^2 = tr(\mathbf{B}) + mb_{jj} \tag{11}$$

$$\sum_{j=1}^{m} dist_{ij}^2 = tr(\mathbf{B}) + mb_{ii} \tag{12}$$

$$\sum_{i=1}^{m}\sum_{j=1}^{m} dist_{ij}^2 = 2mtr(\mathbf{B}) \tag{13}$$

where $tr(\mathbf{B}) = \sum_{i=1}^{m} \parallel \mathbf{z}_i \parallel^2$, let

$$dist_{i.}^2 = \frac{1}{m}\sum_{j=1}^{m} dist_{ij}^2 \tag{14}$$

$$dist_{.j}^2 = \frac{1}{m}\sum_{i=1}^{m} dist_{ij}^2 \tag{15}$$

$$dist_{..}^2 = \frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{m} dist_{ij}^2 \tag{16}$$

From the above formula, we can get that:

$$b_{ij} = \frac{1}{2}(dist_{ij}^2 - dist_{i.}^2 - dist_{.j}^2 + dist_{..}^2) \tag{17}$$

Thus, through keeping the distance matrix $\mathbf{D}$ unchanged before and after dimension reduction, we can obtain the inner product matrix $\mathbf{B}$.

The matrix $\mathbf{B}$ is decomposed into $\mathbf{B} = \mathbf{V}\Lambda\mathbf{V}^T$, in which the diagonal matrix $\Lambda = diag(\lambda_1, ..., \lambda_d)$, $\lambda_1 \geq \lambda_1 \geq, ..., \geq \lambda_d)$ is composed of eigenvalues, and V is the eigenvector matrix. Supposing that there are d^* non-zero eigenvalue which form a diagonal matrix $\Lambda_* = diag(\lambda_1, ..., \lambda_d^*)$, so that the corresponding eigenvector matrix can be represented as $\mathbf{V}_*$, then $\mathbf{Z}$ can be expressed as

$$\mathbf{Z} = \Lambda_*^{1/2}\mathbf{V}_*^T \in \mathbf{R}^{d^* \times m} \tag{18}$$

In order to reduce dimension effectively in practical applications, the distance after dimension reduction is often as close as possible to the distance in the original space, rather than strictly equal. In this case, a diagonal matrix consisting of the largest eigenvalues $\tilde{\Lambda} = diag(\lambda_1, ..., \lambda_{d'})$, $d' \leq d$ can be taken to represent the corresponding eigenvector matrix $\tilde{\mathbf{V}}$, which can be expressed as:

$$\mathbf{Z} = \tilde{\Lambda}^{1/2}\tilde{\mathbf{V}}^T \in \mathbb{R}^{d' \times m} \tag{19}$$

Now, we have the reduced channel vector $\mathbf{Z}$. We suppose that the receiver sends the reduced matrix $\mathbf{Z}$ as feedback to the transmitter. The transmitter can use the reduced information to design the precoding matrix of the transmitter. We can use the orthogonal matching pursuit to design the corresponding hybrid precoding. The algorithm is summarized in Algorithm 1. The following is procedure of this algorithm. $\mathbf{V}_*$ is first considered as the optimal precoding matrix. Then all the analog beam vectors $\mathbf{A}_t$ are projected to $\mathbf{V}_*$ in turn, where $\mathbf{A}_t = [\mathbf{a}_t(\phi^t_{i,l}, \theta^t_{i,l}), ..., \mathbf{a}_t(\phi^t_{N_{cl},N_{ray}}, \theta_{N_{cl},N^t_{ray}})]$. According to the maximum projection, the optimal analog precoding vector is found, and then the vector is set as the analog vector $\mathbf{F}_{RF(:,i)}$. After finding the principal eigenvector, the digital precoding matrix can be calculated by the least square method. Then the optimal analog precoding vectors are found in a loop, and the process will not stop until all the N_{tRF} analog vectors are found. After the N_{tRF} times iteration, the RF precoding matrix $\mathbf{F}_{RF}$ and the digital precoding matrix $\mathbf{F}_{BB}$ can be found. Finally, we need to normalize the digital precoding vector so that the normalized vector meets the power constraint

$$\mathbf{F}_{BB} = \sqrt{N_s}\frac{\mathbf{F}_{BB}}{\parallel \mathbf{F}_{RF}\mathbf{F}_{BB} \parallel_F} \tag{20}$$

5 Simulation Results

The parameter settings in the simulation process are as follows: we use the multipath extended Saleh-Valenzuela channel given in Sect. 2-B. Each channel has $N_{cl} = 8$ cluster and each cluster has $N_{ray} = 8$ scattering path. For simplicity, it is assumed that each cluster has equal power allocation. In the simulation, uniform linear array (ULA) is adopted, so only plane angle ϕ needs to be considered. Receiving angle and transmitting angle ϕ_r, ϕ_t are randomly uniformly distributed in $[0, \pi]$. All precoding schemes have the same SNR, which is defined as $SNR = \frac{\rho}{\sigma_n^2}$. Each simulation implements through 50 channel realizations. Receiver and transmitter arrays have $N_r = N_t = 32$ antennas and $N_{rRF} = N_{tRF} = 8$ RF chains.

Algorithm 1. Multiple dimensional scaling hybrid precoding scheme

Input: $d^* = N_s$, $\mathbf{H}$, $\mathbf{A}_t$
Output: $\mathbf{F}_{RF}$, $\mathbf{F}_{BB}$

compute matrix $\mathbf{D}$ based on (9)
compute matrix $\mathbf{B}$ based on (17)
decompose $\mathbf{B} = \mathbf{V}\Lambda\mathbf{V}^T$
we have $\mathbf{Z} = \Lambda_*^{1/2}\mathbf{V}_*^T \in \mathbb{R}^{d^* \times m}$
feed back $\mathbf{V}$ to BS for precoding
$\mathbf{F}_{RF} = []$
$\mathbf{F}_{res} = \mathbf{V}_*^T$
for $i \leq N_{tRF}$ **do**
  $\Psi = \mathbf{A}_t^* \mathbf{F}_{res}$
  $k = argmax_{l=1,\ldots,N_{cl}N_{ray}} (\Psi\Psi^*)_{(l,l)}$
  $\mathbf{F}_{RF} = [\mathbf{F}_{RF}|\mathbf{A}_t^{(k)}]$
  $\mathbf{F}_{BB} = (\mathbf{F}_{RF}^*\mathbf{F}_{RF})^{-1}\mathbf{F}_{RF}^*\mathbf{V}_*^T$
  $\mathbf{F}_{res} = \frac{\mathbf{V}_*^T - \mathbf{F}_{RF}\mathbf{F}_{BB}}{\|\mathbf{V}_*^T - \mathbf{F}_{RF}\mathbf{F}_{BB}\|_F}$
$\mathbf{F}_{BB} = \sqrt{N_s}\frac{\mathbf{F}_{BB}}{\|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F}$
return $\mathbf{F}_{RF}$, $\mathbf{F}_{BB}$

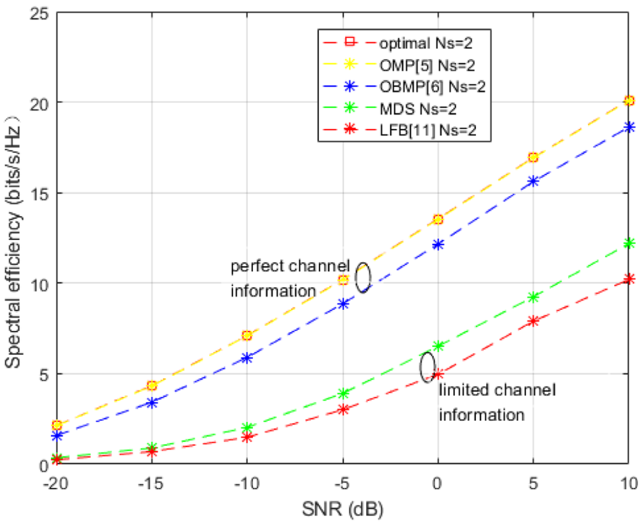

Fig. 4. When $N_s = 2$, performance comparison between Multiple Dimensional Scaling (MDS) hybrid precoding scheme and other schemes.

Experimental summary: from Fig. 4, it can be seen that under the same conditions and the number of data streams equals to two, the proposed Multiple Dimensional Scaling (MDS) hybrid precoding scheme outperforms the finite feedback scheme [13], and our scheme performs better than the finite feedback scheme [14] in the case of varying SNR.

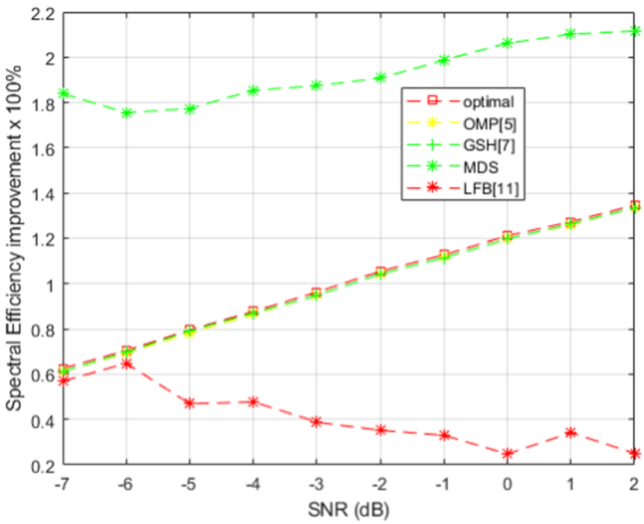

Fig. 5. When $N_s = 2$ and $N_s = 8$, performance improvement between Multiple Dimensional Scaling (MDS) hybrid precoding scheme and other schemes.

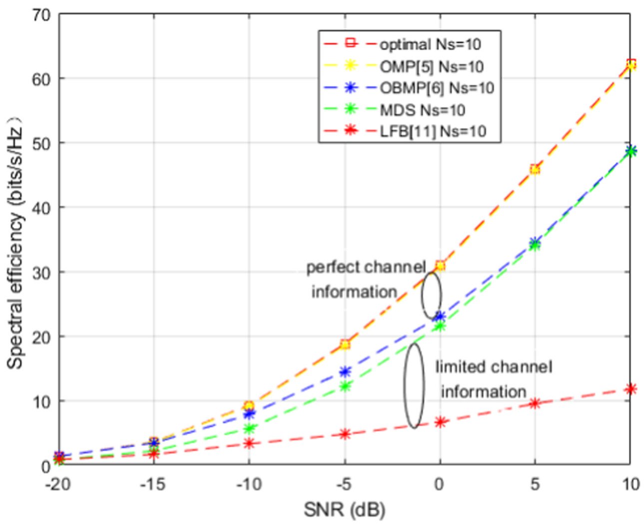

Fig. 6. When $N_s = 10$, $N_{tRF} = 16$, performance comparison between Multiple Dimensional Scaling (MDS) hybrid precoding scheme and other schemes.

In Fig. 5, when we increase the number of data streams from $N_s = 2$ to $N_s = 8$, we compare the performance and capability of various schemes. The formula in Fig. 5 is as follows: $imp = (SE(N_s = 8) - SE(N_s = 2))/SE(N_s = 2) \times 100\%$. From Fig. 5, we can see that when the number of data streams increases, the performance of the MDS precoding scheme increases the most, nearly 200%, while the performance of other schemes increases slightly more than 100%, and the performance of the limited feedback scheme [11] even decreases slightly.

In Fig. 6, when we continue to increase the number of data streams to $N_s = 10$, we get the simulation results of Fig. 6. We can see that when the number of data streams is large, our scheme always performs better than the limited feedback proposal [11], and is close to some full feedback schemes, so our scheme is more suitable for systems with large number of data streams.

6 Conclusion

This paper explores the sparsity of mmWave channel—only a few main paths are useful to help the precoding procedure. We study a limited feedback scenario for mmWave massive MIMO system, we propose a hybrid precoding scheme called Multiple Dimensional Scaling (MDS) hybrid precoding. Instead of the most paper which uses perfect channel information, our scheme only feeds back main vector of the channel which have a significant decrease of feedback overhead. Compared with other limited feedback scheme, our plan has a good performance. When the data streams improve, our scheme improves most of them. Our scheme always performs better than the limited feedback scheme, and is close to some full feedback schemes, and when the system increases the number of data streams, the system performance improves significantly. Therefore, in conclusion, the proposed scheme in this paper can be applied to the limited feedback scheme, and is more suitable for systems with large number of data streams. However, the future communication technology is mainly affected by two factors, namely spectrum efficiency and energy efficiency. This article mainly studies and simulates the spectrum efficiency of our proposed scheme. Therefore, research on energy efficiency of our scheme will be carried out in the future.

References

1. Andrews, J.G., et al.: What will 5G be? IEEE J. Sel. Areas Commun. **32**(6), 1065–1082 (2014)
2. Heath, R.W., González-Prelcic, N., Rangan, S., Roh, W., Sayeed, A.M.: An overview of signal processing techniques for millimeter wave MIMO systems. IEEE J. Sel. Topics Signal Process. **10**(3), 436–453 (2016)
3. Rusek, F., et al.: Scaling up MIMO: opportunities and challenges with very large arrays. IEEE Signal Process. Mag. **30**(1), 40–60 (2013)
4. Chen, C., Tsai, C., Liu, Y., Hung, W., Wu, A.: Compressive sensing (CS) assisted low-complexity beamspace hybrid precoding for millimeter-wave MIMO systems. IEEE Trans. Signal Process. **65**(6), 1412–1424 (2017)
5. Zhang, W., Xia, X., Fu, Y., Bao, X.: Hybrid and full-digital beamforming in mmWave massive MIMO systems: a comparison considering low-resolution ADCs. China Commun. **16**(6), 91–102 (2019). https://doi.org/10.23919/JCC.2019.06.008
6. Zhang, Y., Du, J., Chen, Y., Li, X., Rabie, K.M., Kharel, R.: Near-optimal design for hybrid beamforming in mmWave massive multi-user MIMO systems. IEEE Access **8**, 129153–129168 (2020). https://doi.org/10.1109/ACCESS.2020.3009238

7. Xue, X., Wang, Y., Yang, L., Shi, J., Li, Z.: Energy-efficient hybrid precoding for massive MIMO mmWave systems with a fully-adaptive-connected structure. IEEE Trans. Commun. **68**(6), 3521–3535 (2020). https://doi.org/10.1109/TCOMM.2020.2979139
8. Ayach, O.E., Rajagopal, S., Abu-Surra, S., Pi, Z., Heath, R.W.: Spatially sparse precoding in millimeter wave MIMO systems. IEEE Trans. Wirel. Commun. **13**(3), 1499–1513 (2014)
9. Hung, W.L., Chen, C.H., Liao, C.C., et al.: Low-complexity hybrid precoding algorithm based on orthogonal beamforming codebook. In: IEEE Workshop on Signal Processing Systems. IEEE (2015)
10. Alkhateeb, A., Heath, R.W.: Gram schmidt based greedy hybrid precoding for frequency selective millimeter wave MIMO systems. In: IEEE International Conference on Acoustics. IEEE (2016)
11. Alkhateeb, A., El Ayach, O., Leus, G., et al.: Channel estimation and hybrid precoding for millimeter wave cellular systems. IEEE J. Sel. Topics Signal Process. **8**(5), 831–846 (2014)
12. Sohrabi, F., Yu, W.: Hybrid digital and analog beamforming design for large-scale MIMO systems. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brisbane, Australia, April 2015
13. Chen, C.-E.: An iterative hybrid transceiver design algorithm for millimeter wave MIMO systems. IEEE Wirel. Commun. Lett. **4**(3), 285–288 (2015)
14. Alkhateeb, A., Leus, G., Heath, R.W., Jr.: Limited feedback hybrid precoding for multi-user millimeter wave systems. IEEE Trans. Wirel. Commun. **14**(11), 6481–6494 (2014)