



# Article Deep Learning-Based Joint CSI Feedback and Hybrid Precoding in FDD mmWave Massive MIMO Systems

Qiang Sun <sup>1,2</sup>, Huan Zhao <sup>1</sup>, Jue Wang <sup>1,\*</sup> and Wei Chen <sup>1</sup>

- <sup>1</sup> School of Information Science and Technology, Nantong University, Nantong 226019, China; sungiang@ntu.edu.cn (Q.S.); 2010310019@stmail.ntu.edu.cn (H.Z.); chenwei0303@ntu.edu.cn (W.C.)
- <sup>2</sup> Nantong Research Institute for Advanced Communication Technologies (NRIACT), Nantong 226019, China
  - \* Correspondence: wangjue@ntu.edu.cn

**Abstract:** In this paper, we propose an end-to-end deep learning approach to realize channel state information (CSI) feedback and hybrid precoding for millimeter wave massive multiple-input multipleoutput systems in the frequency division duplexing mode. Different from conventional approaches that treat the CSI reconstruction and hybrid precoding as separate components, we propose a new end-to-end learning method bypassing the channel reconstruction phase, and design the hybrid precoders and combiners directly from the feedback codewords (a compressed version of the CSI). More specifically, we design a neural network composed of the CSI feedback and hybrid precoding. Experiment results show that our proposed network can achieve better performance than conventional hybrid precoding schemes that reserve channel reconstruction, especially when the feedback resources are limited.

Keywords: deep learning; massive MIMO; CSI feedback; hybrid precoding; millimeter wave



Citation: Sun, Q.; Zhao, H.; Wang, J.; Chen, W. Deep Learning-Based Joint CSI Feedback and Hybrid Precoding in FDD mmWave Massive MIMO Systems. *Entropy* **2022**, *24*, 441. https://doi.org/10.3390/e24040441

Academic Editor: Sergio Saponara

Received: 8 March 2022 Accepted: 20 March 2022 Published: 23 March 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

# 1. Introduction

Hybrid precoding is a promising technique for millimeter wave (mmWave) massive multiple-input multiple-output (MIMO) systems [1–5], thanks to its merit of reducing the number of RF chains while achieving a similar performance of fully digital architecture [1]. Many algorithms for hybrid precoding design have been proposed [2,3,6–10], e.g., the simultaneous orthogonal matching pursuit (SOMP) algorithm [2] and the manifold optimization based alternate minimization (MO-AltMin) algorithm [6]. In these works, it is critical for the base station (BS) to acquire accurate downlink channel state information (CSI) [11]. In frequency division duplexing (FDD) communication systems, however, it is challenging for the BS to acquire the downlink CSI since the uplink-downlink channel reciprocity does not hold. Hence, the user equipments (UEs) need to estimate the downlink CSI and report it to the BS through feedback links. Conventional CSI feedback schemes utilize techniques such as codebook design [12,13] and compressive sensing (CS) [14]. There are also some works that jointly optimize CSI feedback and hybrid precoding [15,16]. The authors in [15] proposed a two-stage approach based on long-term and instantaneous CSI to realize hybrid precoding design and reduce feedback overhead for FDD multiuser massive MIMO systems. The paper [16] investigated the performance of hybrid precoding based on quantized CSI feedback for multi-user massive MIMO systems. However, these approaches introduce an additional implementation cost and overhead to the system, especially when the number of users and the number of BS antennas are large.

In recent years, with the rapid development of deep learning technologies, many new approaches have been realized successfully in CSI feedback [11,17–21] and hybrid precoding [22–25]. As for the CSI feedback, a deep learning-based CSI compression and recovery scheme, named CsiNet, has been proposed in [11]. The CSI reconstruction accuracy of the CsiNet significantly outperforms existing CS algorithms [26]. Stemming from the CsiNet, more sophisticated architectures have been further developed to enhance the performance, e.g., CsiNet-LSTM [17], CsiNet+ [18]. The paper [20] further reduced the feedback overhead by feeding back bit streams instead of floating point numbers. Considering the noise in the practical feedback link, the authors in [21] proposed a denoising network to reduce the effect of noise on CsiNet. As for the hybrid precoding, the authors of [22] utilized the multi-layer perceptrons (MLP) to design the precoders. In [23,24], convolutional neural network (CNN) frameworks were proposed to estimate the analog precoders and combiners for the single-user and multi-user scenarios, respectively. The authors in [27] use the quantized received signal strength indicators for hybrid precoding design. In [28], the authors proposed a deep learning framework for hybrid precoding and channel estimation without instantaneous CSI feedback for mmWave massive MIMO systems.

Most of the aforementioned works realize CSI feedback and hybrid precoding in separate modules. However, such a separate design may not fully explore the capabilities and advantages that can be provided by end-to-end deep learning [29,30]. In this regard, several recent works focus on exploring end-to-end deep learning that bypasses different intermediate components of communication systems [29–32]. The authors in [29,32] proposed an end-to-end design that bypasses the channel estimation, and designs the hybrid precoders directly from the received pilots in TDD massive MIMO systems. Refs. [30,31] proposed the idea of bypassing channel reconstruction in FDD massive MIMO systems. In [30], considering a point-to-point MIMO system, the authors jointly train the DNNs at the transmitter (TX) and the receiver (RX) where the DNN in the receiver is used to map the pilot-aided signals into quantized vectors, and the DNN in the transmitter is used to map the quantized vectors into precoding vectors. In [31], the authors treats the end-to-end precoding design problem as a distributed source coding (DSC) problem and jointly designs the downlink pilot training, channel feedback, and precoding. However, Refs. [30,31] only consider the design of fully-digital precoding. The end-to-end design of CSI feedback and hybrid precoding that bypasses channel reconstruction in FDD mmWave massive MIMO is less understood in the literature and is still open for investigation.

In this paper, we investigate the joint design of CSI feedback and hybrid precoding for FDD massive MIMO systems. We propose a new neural network structure and an end-toend learning framework, which bypasses channel reconstruction and directly designs the hybrid precoders and combiners from the feedback codewords. Specifically, our proposed neural network consists of two parts: CSI feedback and hybrid precoding. In order to train the network, we generate the input-output pairs where the input is the channel matrices and the output is the hybrid precoders and combiners. The main contributions of this paper are summarized as follows:

- A new deep learning-based end-to-end method of joint CSI feedback and hybrid precoding for FDD massive MIMO systems is proposed. Differing from the existing works that jointly optimize CSI feedback and hybrid precoding by using traditional algorithms, we adopt end-to-end deep learning techniques to solve the problem. Meanwhile, our proposed method bypasses channel reconstruction and directly designs the hybrid precoders and combiners from the feedback codewords for FDD massive MIMO systems, which is different from prior works that treat the CSI reconstruction and hybrid precoding as separate components and has been less investigated in the latest end-to-end works;
- A new end-to-end neural network structure for FDD mmWave massive MIMO systems is proposed in this paper. It consists of two parts: CSI feedback and hybrid precoding. The former, realized by CNN, transforms the channel matrices into feedback codewords and the latter, realized by DNN, transforms feedback codewords into hybrid precoders and combiners;
- The simulation results illustrate that compared with conventional approaches, which
  reserve channel reconstruction, our proposed method can significantly reduce the
  feedback overhead and achieve better performance, especially when the feedback
  resources are limited.

**Notation 1.** We use  $(\cdot)^T$ ,  $(\cdot)^H$ ,  $(\cdot)^{-1}$ , which denote transpose, conjugate transpose and inverse, respectively.  $\mathbf{I}_N$  denotes the identity matrix whose size is  $N \times N$ .  $\binom{A}{B}$  represents the number of all combinations of B elements taken from A different elements.  $\Re\{\cdot\}$  and  $\Im\{\cdot\}$  denote the real and imaginary parts of a variable.  $\angle\{\cdot\}$  denotes the angle of complex quantity.  $\mathbb{E}\{\cdot\}$  denotes the statistical expectation.  $[\mathbf{Y}]_{i,j}$  denotes (*i*-th, *j*-th) element of matrix  $\mathbf{Y}$ .  $[\mathbf{Y}]_{:,j}$  denotes the determinant of matrix  $\mathbf{Y}$ .

#### 2. System Model

We consider a point-to-point FDD mmWave MIMO system in which the TX with  $N_T$  antennas serves the RX with  $N_R$  antennas.  $N_S$  is the number of the data streams to be transmitted. The TX and the RX are equipped with  $N_T^{RF}$  and  $N_R^{RF}$  RF chains such that  $N_S \leq N_T^{RF} \leq N_T$  and  $N_S \leq N_R^{RF} \leq N_R$ . This hybrid structure enables the TX to apply a baseband precoder  $\mathbf{F}_b \in \mathbb{C}^{N_T^{RF} \times N_S}$  to the transmit signal  $\mathbf{s} \in \mathbb{C}^{N_S}$  such that  $\mathbb{E}[\mathbf{ss}^H] = \mathbf{I}_{N_S} / N_S$ , followed by an analog precoder  $\mathbf{F}_a \in \mathbb{C}^{N_T \times N_T^{RF}}$ . The TX has the power constraint as  $\|\mathbf{F}_a\mathbf{F}_b\|_F^2 = N_S$ . Since the analog precoder  $\mathbf{F}_a$  is implemented using analog phase shifters, its elements have equal norm, i.e.,  $[[\mathbf{F}_a]_{:,i}[\mathbf{F}_a]_{:,i}^H]_{i,i} = N_T^{-1}$ . Therefore, the transmitted signal is given by  $\mathbf{x} = \mathbf{F}_a\mathbf{F}_b\mathbf{s}$ .

We consider a narrowband block-fading channel and the received signal can be written as

$$\tilde{\mathbf{y}} = \sqrt{\rho} \mathbf{H} \mathbf{F}_a \mathbf{F}_b \mathbf{s} + \mathbf{n},\tag{1}$$

where  $\tilde{\mathbf{y}} \in \mathbb{C}^{N_R \times 1}$ ,  $\rho$  is the average received power,  $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$  is the channel matrix,  $\mathbf{n} \in \mathbb{C}^{N_R \times 1}$  is the additive white Gaussian noise (AWGN) with  $\mathbf{n} \sim C\mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_{N_R})$ . The mmWave channel  $\mathbf{H}$ , which consists of  $N_c$  clusters with  $N_{ray}$  propagating rays [2], can be written as

$$\mathbf{H} = \gamma \sum_{i=1}^{N_c} \sum_{l=1}^{N_{ray}} \alpha_{il} \Lambda_R(\Theta_R^{(il)}) \Lambda_T(\Theta_T^{(il)}) \mathbf{a}_R(\Theta_R^{(il)}) \mathbf{a}_T^H(\Theta_T^{(il)}),$$
(2)

where  $\gamma = \sqrt{N_T N_R / (N_c N_{ray})}$  is a normalization factor.  $\alpha_{il}$  is the complex channel gain of the *l*th propagation path in the *i*th scattering cluster.  $\phi_T^{(il)}(\theta_T^{(il)})$  and  $\phi_R^{(il)}(\theta_R^{(il)})$  represent the azimuth (elevation) angles of departure and arrival, respectively.  $\Theta_T^{(il)} = (\phi_T^{(il)}, \theta_T^{(il)})$  and  $\Theta_R^{(il)} = (\phi_R^{(il)}, \theta_R^{(il)})$  represent the angle of departure (AoD) and the angle of arrival (AoA), respectively.  $\Lambda_T(\Theta_T^{(il)})$  and  $\Lambda_R(\Theta_R^{(il)})$  denote the gains of transmit and receive antenna element that correspond to different AoDs and AoAs.  $\mathbf{a}_T(\Theta_T^{(il)}) \in \mathbb{C}^{N_T}$  and  $\mathbf{a}_R(\Theta_R^{(il)}) \in \mathbb{C}^{N_R}$ are the array response vectors at the TX and the RX. Considering the uniform planar array (UPA) with *U* elements on the *y*-axis and *V* elements on *z*-axis, the array response vector can be expressed as

$$\mathbf{a}(\phi,\theta) = \frac{1}{\sqrt{N}} [1, \dots, e^{j\frac{2\pi}{\lambda}d(u\sin(\phi)\sin(\theta) + v\cos(\theta))}, \dots, e^{j\frac{2\pi}{\lambda}d((U-1)\sin(\phi)\sin(\theta) + (V-1)\cos(\theta))}]^T, \quad (3)$$

where  $0 \le u < U$ ,  $0 \le v < V$  and N = UV.  $\lambda$  denotes the wavelength of mmWave and *d* is the space between adjacent antennas.

The decoded data streams  $\mathbf{y}$ , after being processed by analog and baseband combiners, can be written as

$$\mathbf{y} = \mathbf{W}_b^H \mathbf{W}_a^H \tilde{\mathbf{y}}$$
  
=  $\sqrt{\rho} \mathbf{W}_b^H \mathbf{W}_a^H \mathbf{H} \mathbf{F}_a \mathbf{F}_b \mathbf{s} + \mathbf{W}_b^H \mathbf{W}_a^H \mathbf{n},$  (4)

where  $\mathbf{W}_a \in \mathbb{C}^{N_R \times N_R^{RF}}$  and  $\mathbf{W}_b \in \mathbb{C}^{N_R^{RF} \times N_S}$  denote the analog and baseband combiners, respectively. Similar to  $\mathbf{F}_a$ ,  $\mathbf{W}_a$  is subject to  $[[\mathbf{W}_a]_{:i}[\mathbf{W}_a]_{:i}^{H}]_{i,i} = N_R^{-1}$ .

Our objective is to design the hybrid precoders and combiners  $F_a$ ,  $F_b$ ,  $W_a$ ,  $W_b$  at the TX so as to maximize the spectral efficiency of the system. Assuming that the symbol vector **s** follows Gaussian distribution, the spectral efficiency can be written as

$$R = \log_2 \left( \left| \mathbf{I}_{N_S} + \rho N_S^{-1} \mathbf{R}_n^{-1} \mathbf{W}_b^H \mathbf{W}_a^H \mathbf{H} \mathbf{F}_a \mathbf{F}_b \mathbf{F}_b^H \mathbf{F}_a^H \mathbf{H}^H \mathbf{W}_a \mathbf{W}_b \right| \right),$$
(5)

where  $\mathbf{R}_n = \sigma_n^2 \mathbf{W}_b^H \mathbf{W}_a^H \mathbf{W}_a \mathbf{W}_b$  is the covariance matrix of the noise after receive processing.

It is necessary for the TX to obtain the instantaneous CSI for optimal precoders and combiners design. For simplicity, we assume that the perfect downlink CSI has been obtained at RX via pilot-based training and only focus on the design of joint CSI feedback and hybrid precoding.

To reduce feedback overhead, the RX first compresses **H** into a *M*-dimensional codeword **c**, then feeds back **c** (other than **H**) to the TX. This is described as

$$\mathbf{c} = \mathcal{F}(\mathbf{H}),\tag{6}$$

where  $\mathbf{c} \in \mathbb{C}^{M \times 1}$  and  $\mathcal{F}(\cdot)$  represents the CSI compression scheme adopted at the RX.

The TX receives **c** and designs the downlink precoders and combiners accordingly. Note that, as illustrated in Figure 1, conventional hybrid precoding design schemes assume that the TX conducts design based on CSI feedback, i.e., CSI reconstruction from **c** is required, which may induce additional errors in this process. Differently, we design the precoders and combiners directly from **c** without requiring a CSI reconstruction process (the TX needs extra overhead to transmit the designed combiners to the RX). This is described as

$$\{\mathbf{F}_a, \mathbf{F}_b, \mathbf{W}_a, \mathbf{W}_b\} = \mathcal{P}(\mathbf{c}),\tag{7}$$

where  $\mathcal{P}(\cdot)$  denotes the hybrid precoding scheme.



**Figure 1.** (a) The architecture of the traditional hybrid precoding methods that reserve channel reconstruction. (b) Our proposed hybrid precoding method that bypasses channel reconstruction (assuming that the RX has obtained the perfect CSI through the pilot training in both architectures).

In the following, we aim to jointly design  $\mathcal{F}(\cdot)$  in (6) and  $\mathcal{P}(\cdot)$  in (7), to maximize the spectral efficiency of the system (described in (5)). Such a problem is in general difficult to tackle with conventional optimization techniques. Alternatively, we seek a deep learning framework for handling this problem.

#### 3. Proposed Deep Learning Framework For CSI Feedback and Hybrid Precoding

In this section, we describe the details of our proposd deep learning framework for the joint design of CSI feedback and hybrid precoding. Then, we describe how to generate the training dataset. Figure 2 shows the architecture of our proposed neural network. It consists of a CSI feedback phase and a hybrid precoding phase. We build a CNN model to compress the channel matrices into the feedback codewords at the RX. At the TX, we build a DNN model to design the hybrid precoders and combiners from the feedback codewords. It is worth mentioning that in the deployment phase, the RX has stringent requirements on response latency and energy consumption, which is challenging for neural network design and provides a new research direction. We can use neural network quantization [33] to reduce model size and optimize the trade-off between accuracy and efficiency.



**Figure 2.** The architecture of the proposed neural network that represents the end-to-end CSI feedback and hybrid precoding.

## 3.1.1. CSI Feedback

As assumed, the RX has obtained perfect CSI and needs to compress the downlink channel matrix **H** into a *M*-dimensional codeword. This module is built with **H** as input, **c** as output, and with multiple-layer CNNs to realize the function  $\mathcal{F}(\cdot)$  in (6). The input of the proposed CNN framework, named **X**, are the real and imaginary parts of **H**, i.e.,  $[\mathbf{X}]_{:,:,1} = \Re{\{\mathbf{H}\}}$  and  $[\mathbf{X}]_{:,:,2} = \Im{\{\mathbf{H}\}}$ . The first and second layers are both convolutional layers with 64 filters to generate feature maps. The size of each filter is  $2 \times 2$  and the stride is 1. Batch normalization is introduced to each convolutional layer. Following the second layer, we use a fully connected layer with 1024 units. The rectified linear unit (ReLU) is adopted at the first three layers, where  $\text{ReLU}(x) = \max(x, 0)$ . Finally, a fully connected layer, whose size is  $M \times 1$ , is used to generate the codeword **c**.

#### 3.1.2. Hybrid Precoding

Under the assumption of an error-free feedback channel between the TX and the RX, the TX obtains the codewords **c** fed back from the RX and designs the hybrid precoders and combiners accordingly. We design a DNN model to realize the function  $\mathcal{P}(\cdot)$  in (7). The codeword **c** is the input vector and **z** is the output vector whose size is  $Q \times 1$ . Note that because  $\mathbf{F}_a$  and  $\mathbf{W}_a$  are analog precoder and combiner matrices, we only need to extract the angle information of the elements in  $\mathbf{F}_a$  and  $\mathbf{W}_a$ . The vector **z** is the vectorized version of  $\mathbf{F}_a$ ,  $\mathbf{F}_b$ ,  $\mathbf{W}_a$ ,  $\mathbf{W}_b$  and can be formed as

$$\mathbf{z} = [vec^{T}(\angle \mathbf{F}_{a}), vec^{T}(\angle \mathbf{W}_{a}), \Re(vec^{T}(\mathbf{F}_{b})), \Im(vec^{T}(\mathbf{F}_{b})), \Re(vec^{T}(\mathbf{W}_{b})), \Im(vec^{T}(\mathbf{W}_{b}))], (8)$$

where  $Q = N_T N_T^{RF} + N_R N_R^{RF} + 2N_T^{RF} N_S + 2N_R^{RF} N_S$ . The first and second layers of the DNN are both fully connected layers with 1024 units. The activation function ReLU and the dropout layer with 50% probability are placed after the first and second layers. The third layer is a fully connected layers with Q units, which is used to generate the vector  $\mathbf{z}$ .

#### 3.2. Dataset Generation

The dataset of the proposed network is denoted as  $\mathcal{D}$ , and a sample in  $\mathcal{D}$  is an inputoutput pair written as  $(\mathbf{X}, \mathbf{z})$ . In this paper, we need to design  $\mathbf{F}_a$ ,  $\mathbf{F}_b$ ,  $\mathbf{W}_a$ ,  $\mathbf{W}_b$  from  $\mathbf{H}$  to maximize the spectral efficiency. The optimization problem can be formulated as

$$\begin{array}{ll} \underset{\mathbf{F}_{a},\mathbf{F}_{b},\mathbf{W}_{a},\mathbf{W}_{b}}{\text{maximize}} & R\\ s.t. & \mathbf{F}_{a} \in \mathcal{F}_{a}, \\ & \mathbf{W}_{a} \in \mathcal{W}_{a}, \\ & ||\mathbf{F}_{a}\mathbf{F}_{b}||_{F}^{2} = N_{S}, \end{array}$$
(9)

where  $\mathcal{F}_a$  and  $\mathcal{W}_a$  are the sets including all the feasible candidates of analog precoders and combiners, respectively. It is difficult to obtain the optimal solution of (9). To solve the problem and obtain the sub-optimal solution,  $\mathcal{F}_a$  and  $\mathcal{W}_a$  need to be predefined. Because the analog precoder  $\mathbf{F}_a$  is related to the array responses  $\mathbf{a}_T(\Theta_T)$  [2],  $\mathcal{F}_a$  can be defined as

$$\mathcal{F}_a = \left\{ \mathbf{F}_a^{(1)}, \dots, \mathbf{F}_a^{(c_F)}, \dots, \mathbf{F}_a^{(C_F)} \right\},\tag{10}$$

where  $c_F = 1, 2, ..., C_F$ .  $C_F = \binom{N_{path}}{N_T^{RF}}$  is the number of the analog precoder candidates and  $N_{path} = N_c \times N_{ray}$ .  $\mathbf{F}_a^{(c_F)} = [\mathbf{a}_T(\Theta_T^{(1)}), ..., \mathbf{a}_T(\Theta_T^{(t)}), ..., \mathbf{a}_T(\Theta_T^{(N_T^{RF})})] \in \mathbb{C}^{N_T \times N_T^{RF}}$  is the candidate of  $\mathbf{F}_a$  in  $\mathcal{F}_a$ , where  $t = 1, ..., N_T^{RF}$ . Similarly, the set of feasible analog combiners  $\mathcal{W}_a$  can be defined as

$$\mathcal{W}_a = \left\{ \mathbf{W}_a^{(1)}, \dots, \mathbf{W}_a^{(c_W)}, \dots, \mathbf{W}_a^{(C_W)} \right\},\tag{11}$$

where  $c_W = 1, 2, ..., C_W$ .  $C_W = {N_{path} \choose N_R^{RF}}$  is the number of the analog combiner candidates.

 $\mathbf{W}_{a}^{(c_{W})} = [\mathbf{a}_{R}(\Theta_{R}^{(1)}), \dots, \mathbf{a}_{R}(\Theta_{R}^{(p)}), \dots, \mathbf{a}_{R}(\Theta_{R}^{(N_{R}^{RF})})] \in \mathbb{C}^{N_{R} \times N_{R}^{RF}}$  is the candidate of  $\mathbf{W}_{a}$  in  $\mathcal{W}_{a}$ , where  $p = 1, \dots, N_{R}^{RF}$ . Therefore, the optimization problem in (9) can be rewritten as

$$\begin{array}{ll} \underset{\hat{c}_{F,\hat{c}_{W}}}{\text{maximize}} & R \\ s.t. & \mathbf{F}_{a} \in \mathcal{F}_{a}, \\ \mathbf{W}_{a} \in \mathcal{W}_{a}, \\ \mathbf{F}_{b} = \left(\mathbf{F}_{a}^{H}\mathbf{F}_{a}\right)^{-1}\mathbf{F}_{a}^{H}\mathbf{F}^{opt}, \\ \mathbf{W}_{b} = \left(\mathbf{W}_{a}^{H}\mathbf{A}\mathbf{W}_{a}\right)^{-1}\left(\mathbf{W}_{a}^{H}\mathbf{A}\mathbf{W}^{opt}\right), \end{array}$$
(12)

where  $\mathbf{A} = \frac{\rho}{N_S} \mathbf{H} \mathbf{F}_a \mathbf{F}_b \mathbf{F}_a^H \mathbf{H}^H + \sigma_n^2 \mathbf{I}_{N_R}$  is the covariance of the array output in (1).  $\mathbf{F}^{opt}$ ,  $\mathbf{W}^{opt}$  represent the optimal fully-digital precoder and combiner that can be obtained from singular value decomposition (SVD) of  $\mathbf{H}$  [2,23].

To reduce the complexity, the problem (12) can be decomposed into the sub-problems of precoder and combiner designs. The precoder design problem (13) and combiner design problem (14) can be written as [23]

$$\begin{array}{l} \underset{\hat{c}_{F}}{\text{maximize}} \quad \log_{2} \left( \left| \mathbf{I}_{N_{S}} + \frac{\rho}{N_{S}\sigma_{n}^{2}} \left( \mathbf{W}^{opt^{H}} \mathbf{W}^{opt} \right)^{-1} \mathbf{W}^{opt^{H}} \mathbf{H} \mathbf{F}_{a} \mathbf{F}_{b} \mathbf{F}_{b}^{H} \mathbf{F}_{a}^{H} \mathbf{H}^{H} \mathbf{W}^{opt} \right| \right) \\ s.t. \qquad \mathbf{F}_{a} \in \mathcal{F}_{a}, \\ \mathbf{F}_{b} = \left( \mathbf{F}_{a}^{H} \mathbf{F}_{a} \right)^{-1} \mathbf{F}_{a}^{H} \mathbf{F}^{opt}, \end{array}$$
(13)

$$\begin{array}{ll} \underset{c_{W}}{\text{maximize}} & \log_{2} \left( \left| \mathbf{I}_{N_{S}} + \frac{\rho}{N_{S} \sigma_{n}^{2}} \left( \mathbf{W}_{b}^{H} \mathbf{W}_{a}^{H} \mathbf{W}_{a} \mathbf{W}_{b} \right)^{-1} \mathbf{W}_{b}^{H} \mathbf{W}_{a}^{H} \mathbf{H} \mathbf{F}^{opt} \mathbf{F}^{opt}^{H} \mathbf{H}^{H} \mathbf{W}_{a} \mathbf{W}_{b} \right| \right) \\ & s.t. & \mathbf{W}_{a} \in \mathcal{W}_{a}, \\ & \mathbf{W}_{b} = \left( \mathbf{W}_{a}^{H} \mathbf{A} \mathbf{W}_{a} \right)^{-1} \left( \mathbf{W}_{a}^{H} \mathbf{A} \mathbf{W}^{opt} \right). \end{array}$$
(14)

In this case, the Euclidean distance between the optimal fully-digital precoder (combiner) and the hybrid precoders (combiners) is minimized, which will maximize the spectral efficiency of hybrid precoding. Once we solve (13) and (14) and obtain  $\hat{c}_F$ ,  $\hat{c}_W$ ,  $\mathbf{F}_a$ ,  $\mathbf{F}_b$ ,  $\mathbf{W}_a$ ,  $\mathbf{W}_b$  can be constructed and the dataset  $\mathcal{D}$  can be generated.

#### 4. Implementation Details

In data generation, we generate the channel matrix **H** according to (2). We consider the UPA with  $N_T = 36$  and  $N_R = 36$  for the TX and RX, respectively. The number of the RF chains at the TX and the RX are both set as 4, i.e.,  $N_T^{RF} = N_R^{RF} = 4$ . For each channel matrix, the propagation environment is modeled with  $N_c = 4$  and  $N_{ray} = 4$  for each clusters with  $\sigma_{\Theta}^2 = 5^{\circ}$  for all transmit and receive azimuth and elevation angles, which are uniform and randomly selected from the interval  $[-60^{\circ}, 60^{\circ}]$  and  $[-30^{\circ}, 30^{\circ}]$ , respectively. The frequency is set as 28 GHz, and the antenna spacing is half the wavelength.

We implement the proposed neural network using MATLAB as a simulation environment. Notably, the channel matrices have been normalized before inputting the neural network. The typical mean squared error (MSE) between the label z and the actual output  $\hat{z}$  is computed as the loss function, which is described as

$$Loss = \frac{1}{J} \sum_{j=1}^{J} \|\mathbf{z}_{j} - \hat{\mathbf{z}}_{j}\|_{2}^{2},$$
(15)

where *J* denotes the size of the dataset. Stochastic gradient descent with momentum (SGDM) optimizer is used to reduce the loss and update the weight of the network. The epoch, batch size, and initial learning rate are set as 200, 400, and 0.0005, respectively. The learning rate is decreased after 20 epochs by a factor of 0.9.

#### 5. Experiment Results

In this section, we evaluate the spectral efficiency of the proposed neural network and compare the performance with the following benchmarks:

**Benchmark 1: SVD with perfect CSI [2]**: Considering that the TX has obtained the perfect CSI, the TX performs hybrid precoding using a fully-digital precoder and combiner, which can be obtained from the SVD of the channel matrix **H**. In this case, the upper bound of spectral efficiency can be obtained.

**Benchmark 2: MO-AltMin with perfect CSI** [6]: Given the perfect CSI at the TX, the TX performs hybrid precoding by using the MO-AltMin algorithm. The MO-AltMin algorithm is one of the alternate minimization hybrid precoding schemes. It is based on manifold optimization and has the best performance in [6].

**Benchmark 3: SOMP with perfect CSI** [2]: Given the perfect CSI at the TX, the SOMP, which is a greedy-based algorithm, is used by the TX to design the hybrid precoders and combiners.

**Benchmark 4: MO-AltMin with CsiNet** [11]: In this benchmark, no prior perfect CSI is initially assumed at the TX. The RX needs to feed the CSI back to the TX over finitecapacity links. To compare the performance of our proposed scheme that the TX designs the precoders and combiners from the codewords directly, and conventional schemes that the TX designs the precoders and combiners from the channel matrices reconstructed from the codewords, we implement a scheme that uses a deep learning approach to perform channel feedback and reconstruction, followed by a conventional hybrid precoding algorithm. We choose CsiNet, which is a classical CSI sensing and recovery mechanism, to realize the channel feedback and reconstruction. The TX uses MO-AltMin algorithm to design the hybrid precoders and combiners after reconstructing the channel matrices.

Figure 3 presents the spectral efficiency comparison of different schemes versus SNRs. The number of data streams  $N_S$  is set as 2 and the length of codewords M is set as 25. It can be observed that the spectral efficiency of all considered algorithms increases monotonically with increasing SNR. The SVD with perfect CSI has the best performance. We observe that

our proposed method can approach the performance of the MO-AltMin with perfect CSI, which means that our end-to-end neural network can effectively generate the precoders and combiners, which maximizes the spectral efficiency. In addition, our proposed scheme also has better performance than the MO-AltMin with CsiNet in the same codeword length, which verifies that our proposed end-to-end method can get better performance than the traditional separate design method in this situation. The SOMP with perfect CSI has the worst performance because it cannot select the optimal set of array responses from the dictionary.



**Figure 3.** Spectral efficiency versus SNRs for  $N_R = N_T = 36$ ,  $N_S = 2$ , M = 25.

We further investigate the performance of our proposed method and conventional hybrid precoding design approaches versus the length of codewords. In Figure 4, as the length of codewords increases, the spectral efficiency of our proposed scheme can gradually approach and eventually exceed the MO-AltMin with perfect CSI when M = 30. Note that the MO-AltMin with perfect CSI suffers from very high feedback overhead, which means that our proposed scheme has lower feedback overhead with similar performance. In addition, it can be observed that our proposed scheme outperforms the MO-AltMin with CsiNet in the same codeword length and the gap is significantly large when M is small, e.g., M = 5. This observation indicates the superior performance of the proposed end-to-end neural network approach for FDD mmWave massive MIMO systems in the case of very low CSI feedback overhead.



Figure 4. Spectral efficiency versus the length of codewords *M*.

Finally, we compare the computational complexity for our proposed method and different benchmarks. Table 1 shows that our proposed method has much lower running

time than other benchmarks. It means that our proposed method can be executed with a relatively lower overhead and is more suitable for practical scenarios.

Table 1. Comparison of Computational Complexity

Methods	Running Time
Proposed method	0.0046 s
MO-AltMin with perfect CSI	1.2999 s
MO-AltMin with CsiNet	1.3007 s

#### 6. Conclusions

In this paper, we consider the joint design of CSI feedback and hybrid precoding for FDD massive MIMO systems. We propose a new deep learning-based end-to-end method that bypasses channel reconstruction and directly designs the hybrid precoders and combiners from the feedback codewords for FDD massive MIMO systems. We propose a new neural network that jointly optimizes CSI feedback and hybrid precoding. In order to train the network, we generate the input-output pairs, where the input is the channel matrices and the output is the hybrid precoders and combiners. Numerical results indicate the ability of the proposed network in reducing the feedback overhead and boosting the system performance in terms of spectral efficiency, especially in the case of the limited feedback resources. Future research directions include some other transmission modules, e.g., the downlink pilot transmission and quantized CSI feedback. Moreover, the performance of our proposed method in terms of energy efficiency is another promising direction for future works.

Author Contributions: Conceptualization, Q.S. and H.Z.; methodology, Q.S.; software, Q.S. and H.Z.; validation, Q.S.; formal analysis, Q.S., H.Z. and J.W.; investigation, Q.S.; resources, Q.S.; data curation, Q.S.; writing—original draft preparation, Q.S.; writing—review and editing, Q.S., H.Z., J.W., and W.C.; visualization, Q.S.; supervision, Q.S.; project administration, Q.S.; funding acquisition, Q.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by National Nature Science Foundation of China under Grant 61971467, the Key Research and Development Program of Jiangsu Province of China under Grant BE2021013-1, the Scientific Research Program of Nantong under Grant JC2021020, Natural Science Research Program of Nantong Vocational University under Grant 21ZK01.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

### References

- Sohrabi, F.; Yu, W. Hybrid Digital and Analog Beamforming Design for Large-Scale Antenna Arrays. *IEEE J. Sel. Top. Signal Process.* 2016, 10, 501–513. [CrossRef]
- 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. 2014, 13, 1499–1513. [CrossRef]
- Rusu, C.; Mèndez-Rial, R.; González-Prelcic, N.; Heath, R.W. Low Complexity Hybrid Precoding Strategies for Millimeter Wave Communication Systems. *IEEE Trans. Wirel. Commun.* 2016, 15, 8380–8393. [CrossRef]
- Zhang, J.; Björnson, E.; Matthaiou, M.; Ng, D.W.K.; Yang, H.; Love, D.J. Prospective Multiple Antenna Technologies for Beyond 5G. IEEE J. Sel. Areas Commun. 2020, 38, 1637–1660. [CrossRef]
- González-Coma, J.P.; Suárez-Casal, P.; Castro, P.M.; Castedo, L. FDD channel estimation via covariance estimation in wideband massive MIMO systems. *Sensors* 2020, 20, 930. [CrossRef] [PubMed]
- Yu, X.; Shen, J.C.; Zhang, J.; Letaief, K.B. Alternating Minimization Algorithms for Hybrid Precoding in Millimeter Wave MIMO Systems. *IEEE J. Sel. Top. Signal Process.* 2016, 10, 485–500. [CrossRef]

- 7. Lyu, S.; Wang, Z.; Gao, Z.; He, H.; Hanzo, L. Lattice-Based mmWave Hybrid Beamforming. *IEEE Trans. Commun.* 2021, 69, 4907–4920. [CrossRef]
- Huang, Y.; Liu, C.; Song, Y.; Yu, X. Near-optimal hybrid precoding for millimeter wave massive MIMO systems via cost-efficient Sub-connected structure. *IET Commun.* 2020, 14, 2340–2349. [CrossRef]
- Sun, Y.; Gao, Z.; Wang, H.; Shim, B.; Gui, G.; Mao, G.; Adachi, F. Principal Component Analysis-Based Broadband Hybrid Precoding for Millimeter-Wave Massive MIMO Systems. *IEEE Trans. Wirel. Commun.* 2020, 19, 6331–6346. [CrossRef]
- Feng, C.; Shen, W.; An, J.; Hanzo, L. Weighted Sum Rate Maximization of the mmWave Cell-Free MIMO Downlink Relying on Hybrid Precoding. *IEEE Trans. Wirel. Commun.* 2021. [CrossRef]
- Wen, C.K.; Shih, W.T.; Jin, S. Deep Learning for Massive MIMO CSI Feedback. *IEEE Wirel. Commun. Mag.* 2018, 7, 748–751. [CrossRef]
- Shen, W.; Dai, L.; Shim, B.; Wang, Z.; Heath, R.W. Channel Feedback Based on AoD-Adaptive Subspace Codebook in FDD Massive MIMO Systems. *IEEE Trans. Commun.* 2018, 66, 5235–5248. [CrossRef]
- Nair, S.S.; Bhashyam, S. Hybrid Beamforming in MU-MIMO Using Partial Interfering Beam Feedback. *IEEE Commun. Lett.* 2020, 24, 1548–1552. [CrossRef]
- Kuo, P.H.; Kung, H.T.; Ting, P.A. Compressive sensing based channel feedback protocols for spatially-correlated massive antenna arrays. In Proceedings of the 2012 IEEE Wireless Communications and Networking Conference (WCNC), Paris, France, 1–4 April 2012; pp. 492–497.
- Almradi, A.; Matthaiou, M.; Xiao, P.; Fusco, V.F. Hybrid Precoding for Massive MIMO With Low Rank Channels: A Two-Stage User Scheduling Approach. *IEEE Trans. Commun.* 2020, 68, 4816–4831. [CrossRef]
- 16. Zhao, Y.; Xu, W.; Xu, J.; Jin, S.; Wang, K.; Alouini, M.S. Analog Versus Hybrid Precoding for Multiuser Massive MIMO With Quantized CSI Feedback. *IEEE Commun. Lett.* 2020, *24*, 2319–2323. [CrossRef]
- Wang, T.; Wen, C.K.; Jin, S.; Li, G.Y. Deep Learning-Based CSI Feedback Approach for Time-Varying Massive MIMO Channels. IEEE Wirel. Commun. Mag. 2019, 8, 416–419. [CrossRef]
- Guo, J.; Wen, C.K.; Jin, S.; Li, G.Y. Convolutional Neural Network-Based Multiple-Rate Compressive Sensing for Massive MIMO CSI Feedback: Design, Simulation, and Analysis. *IEEE Trans. Wirel. Commun.* 2020, 19, 2827–2840. [CrossRef]
- Jin, Y.; Zhang, J.; Jin, S.; Ai, B. Channel Estimation for Cell-Free mmWave Massive MIMO Through Deep Learning. *IEEE Trans. Veh. Technol.* 2019, 68, 10325–10329. [CrossRef]
- Guo, J.; Li, X.; Chen, M.; Jiang, P.; Yang, T.; Duan, W.; Wang, H.; Jin, S.; Yu, Q. AI enabled wireless communications with real channel measurements: Channel feedback. *J. Commun. Inf. Netw.* 2020, *5*, 310–317.
- Ye, H.; Gao, F.; Qian, J.; Wang, H.; Li, G.Y. Deep Learning-Based Denoise Network for CSI Feedback in FDD Massive MIMO Systems. *IEEE Commun. Lett.* 2020, 24, 1742–1746. [CrossRef]
- Huang, H.; Song, Y.; Yang, J.; Gui, G.; Adachi, F. Deep-Learning-Based Millimeter-Wave Massive MIMO for Hybrid Precoding. IEEE Trans. Veh. Technol. 2019, 68, 3027–3032. [CrossRef]
- Elbir, A.M. CNN-Based Precoder and Combiner Design in mmWave MIMO Systems. *IEEE Commun. Lett.* 2019, 23, 1240–1243. [CrossRef]
- 24. Elbir, A.M.; Papazafeiropoulos, A.K. Hybrid Precoding for Multiuser Millimeter Wave Massive MIMO Systems: A Deep Learning Approach. *IEEE Trans. Veh. Technol.* 2020, 69, 552–563. [CrossRef]
- Zhang, J.; Zhang, J.; Ng, D.W.K.; Jin, S.; Ai, B. Improving Sum-Rate of Cell-Free Massive MIMO With Expanded Compute-and-Forward. *IEEE Trans. Signal Process.* 2022, 70, 202–215. [CrossRef]
- Jin, Y.; Zhang, J.; Huang, C.; Yang, L.; Xiao, H.; Ai, B.; Wang, Z. Multiple Residual Dense Networks for Reconfigurable Intelligent Surfaces Cascaded Channel Estimation. *IEEE Trans. Veh. Technol.* 2022, 71, 2134–2139. [CrossRef]
- Hojatian, H.; Nadal, J.; Frigon, J.F.; Leduc-Primeau, F. Unsupervised Deep Learning for Massive MIMO Hybrid Beamforming. IEEE Trans. Wirel. Commun. 2021, 20, 7086–7099. [CrossRef]
- Elbir, A.M. A Deep Learning Framework for Hybrid Beamforming Without Instantaneous CSI Feedback. *IEEE Trans. Veh. Technol.* 2020, 69, 11743–11755. [CrossRef]
- 29. Attiah, K.M.; Sohrabi, F.; Yu, W. Deep learning for channel sensing and hybrid precoding in TDD massive MIMO OFDM systems. *arXiv* 2021, arXiv:2011.10709.
- Jang, J.; Lee, H.; Hwang, S.; Ren, H.; Lee, I. Deep Learning-Based Limited Feedback Designs for MIMO Systems. *IEEE Wirel.* Commun. Mag. 2020, 9, 558–561. [CrossRef]
- Sohrabi, F.; Attiah, K.M.; Yu, W. Deep Learning for Distributed Channel Feedback and Multiuser Precoding in FDD Massive MIMO. *IEEE Trans. Wirel. Commun.* 2021, 20, 4044–4057. [CrossRef]
- 32. Attiah, K.M.; Sohrabi, F.; Yu, W. Deep Learning Approach to Channel Sensing and Hybrid Precoding for TDD Massive MIMO Systems. In Proceedings of the 2020 IEEE Globecom Workshops (GC Wkshps), Taipei, Taiwan, 7–11 December 2020; pp. 1–6.
- Jin, Q.; Yang, L.; Liao, Z. AdaBits: Neural Network Quantization with Adaptive Bit-Widths. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 14–19 June 2020.