INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Downlink Learning-aided Precoding for Cell-Free networks Specially D esigned for beyond 5G Communication System

Swapnaja Deshpande[1], Mona Aggarwal[1]*, Pooja Sabherwal[1], Swaran Ahuja[1]

1 MDE Deptt, The NorthCap University, 122017, Gurgaon, India

OPEN ACCESS

Received: 21-11-2023

Accepted: 18-01-2024

Published: 22-03-2024

Citation: Deshpande S, Aggarwal M, Sabherwal P, Ahuja S (2024) Downlink Learning-aided Precoding for Cell-Free networks Specially D esigned for beyond 5G Communication System. Indian Journal of Science and Technology 17(13): 1292-1303. https://doi.org/10.17485/IJST/v17i13.2959

* **Corresponding author**.

ermonagarg24@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: With increase in number of access points in Cell-Free Massive Multiple Input Multiple Output (CFMM) system, spatial complexity of calculating precoding vector increases with traditional precoding schemes. So, to reduce computational complexity, network cost, and run time, a new deep learning -aided precoder is specially designed for CFMM system. The performance of downlink (DL) Cell-Free Massive Multiple Input Multiple Output (CFMM) system operating under Rayleigh fading channel model is analyze using new deep learning-based precoding scheme. **Methods:** We introduce new deep learning-aided precoder and an improved version of basic scalable pilot assignment algorithm which enhances system performance. We derive closed- form expression for average DL spectral efficiency (SE) for the proposed scheme, which is then compared with Minimum Mean Square Error Successive Interference Cancellation (MMSESIC), Regularized Zero Forcing (RZF), Least Square (LS) and Maximum Ratio (MR) combining techniques. We analyze the proposed scheme with perfect channel state information (CSI), instantaneous CSI, coherent transmission, non-coherent transmission, different pilot configuration. **Findings:** The spectral efficiency obtained with CFMM-MMSESIC and the proposed scheme is 3 bits/s/Hz, and 2.5bits/s/Hz respectively, which clearly indicates the difference of 20 % only. So, the proposed scheme gives near optimal results, and we can use simple linear combining techniques instead of complex non-linear combining techniques. Also, the proposed scheme with coherent data transmission gives performance curve, which is very close to MMSESIC scheme as desired. The performance of the proposed system is improved by reducing run time and computational complexity as compared to conventional linear precoding schemes. **Novelty:** The proposed precoder improves performance by reducing run time and computational complexity. Advanced pilot assignment algorithm enhances performance by reducing pilot contamination.

Keywords: Cell-Free Massive Multiple Input Multiple Output; Pilot Contamination; Precoding; Minimum Mean Square Error Successive Interference Cancellation; Maximum Ratio; Regularized Zero Forcing (RZF); Least Square (LS)

1 Introduction

Unlike optical wireless communication [1–3] where light is used to transmit data, wireless communication uses radio waves to transmit data. The concept of cell-free network in wireless communication, which is the core technology of sixth-generation (6G), is coined from small-cells (SC), distributed MIMO, and massive MIMO [4,5]. The CFMM systems reaps the benefits of network MIMO, small cells (SC) and massive MIMO [6–8]. In CFMM system processing unit is separate i.e. Central Processing unit (CPU) also known as edge cloud processor. Large number of access points (APs) are deployed in a specific area and are connected to CPU through fronthaul connection and CPUs are connected through backhaul connection. In conventional antenna system, antennas are deployed within a cell, but in CFMM system, antennas are deployed in a specific area, hence the name CFMM. The CFMM system provides improvement in spectral efficiency (SE), cell boundary performance and coverage capability. Two future research directions [9,10] and areas of concern for CFMM networks are addressed in this paper i.e. precoding [11] and pilot contamination. Using orthogonal pilot symbols for UL, CE is very common, but for densely populated cell-free networks it will create pilot contamination (PC) and use of non-orthogonal pilot symbol will affect SE adversely. So, there must be a trade-off between the two methods to enhance the system performance. There are many papers [12,13] which consider different methods taking PC into consideration such as pilot reuse, maximization of SINR, partitioning of cell into rings etc. Recently various pilot decontamination [14,15], pilot contamination [16,17] mitigation techniques are also investigated. Different pilot assignment methods are studied in paper [18]. In the proposed paper, we have used a modified version of scalable pilot assignment algorithm in which pilot symbols are assigned to UEs in a sequential manner, thereby improving system performance. Again, to improve the performance of the CFMM system, we can opt for maximizing sum-rate performance by choosing optimized precoding vector. This optimization of precoding vector is mathematically complex and time consuming. Also developing a model-based precoding optimizer algorithm is difficult task. Therefore, in this paper, we have designed a deep learning approach [19,20] for generating optimal matrix. In CFMM system spatial complexity increases with increase in number of APs, so with traditional precoding schemes, implementation cost of the network increases. To avoid aforementioned drawback and to enhance system performance deep neural network based precoder is designed for CFMM system, considering the effect of pilot contamination. As discussed in literature [21] ZF precoding outperforms CB technique, so we have also considered ZF precoding for comparison purpose. But it is proved that MMSE processing gives optimal results for CFMM system. So, we have considered MMSE with successive interference cancellation (SIC) processing as a benchmark for comparison purpose. Rician channel model is perfect for cell free networks for many use cases, but for rich scattering environment, Rayleigh fading channel model is perfect. Here in paper, the CFMM network is analysed under Rayleigh fading channel model. So, the main contribution of this paper is as follows:

• In this paper, we propose a deep learning aided precoder design for DL CFMM system under the assumption of independent Rayleigh fading. The proposed DNN network will reduce the run-time and computational complexity of the system compared to traditional precoding methods. For the proposed precoding scheme, we derive closed-form expression for an achievable data rate.

• We considered a modified version of scalable pilot assignment algorithm to mitigate PC in addition to assigning pilot symbols in sequential order.

• The comparison of the proposed system with different precoding [22] methods such as MMSE-SIC, RZF and MR precoding method is analysed. The performance of the proposed scheme is evaluated for coherent and non-coherent DL data transmission, perfect and instantaneous channel state information (CSI), and for varying pilot lengths.

The rest of the paper is organized as follows. System model for downlink cell-free massive MIMO network is defined in section 2, which covers uplink training, pilot assignment, coherent downlink data transmission, and learning aided precoder. Performance analysis with numerical results is discussed in section 3 followed by conclusion in section 4.

2 Methodology

In this paper, we consider downlink (DL) analysis of CFMM system shown in Figure 1, with 'M' number of APs equipped with 'N' number of antennas and 'K' number of single antenna users (UEs). The system is supposed to operate in time division duplexing (TDD) mode. All APs and UEs are randomly distributed across a specified area considering $MN \gg K$. All APs are connected to central processing unit (CPU) via fronthaul links. The Rayleigh fading channel between UE 'K' and AP 'M', which captures the effect of large-scale and small-scale fading is given by:

$$h_{mk} \sim NC(0, R_{mk}) \tag{1}$$

Where R_{mk} is spatial correlation matrix having dimension $R_{mk} \in C^{N \times N}$. The large-scale fading coefficient $\beta_{mk} \triangleq \frac{t_r(R_{mk})}{N}$ indicates shadowing and pathloss. We assume that h_{mk} is an independent random variable for every $m = 1 \ldots . M$, $APs \ \ and \ \ K =$

$1 \ldots\ldots K$, UEs and channel realization h_{mk} in different coherence blocks are i.i.d. We analyse DL CFMM network with following assumptions:

• Channel Reciprocity: In TDD DL channel is reciprocal of UL channel. So, we assume DL channel to be exactly reciprocal of UL channel considering channel estimation error and pilot contamination.

• Assuming the channel to be constant during coherence interval and varies independently between coherence intervals i.e. block-fading channel model.

Let us consider a TDD frame with coherence block $\tau_c = \tau_p + \tau_{DL}$ where τ_c indicates length of coherence interval in samples, τ_p indicates length of pilot symbols in samples during pilot transmission in UL phase and τ_{DL} indicates length of DL data in samples during DL data transmission in DL phase.

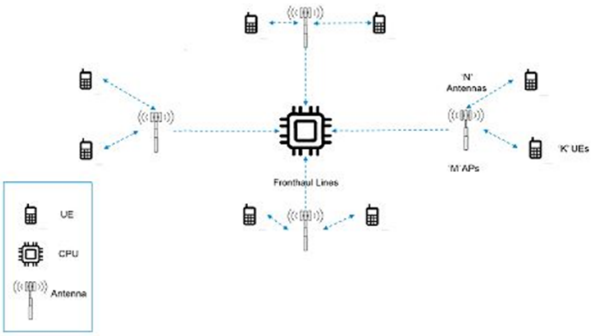

Fig 1. Cell-Free Massive MIMO Architecture

2.1 Uplink Training

Consider all UEs transmit their pilot signals, then the pilot signal received by $m^{th}AP$ is given as:

$$Y^p_{mk} = \ Y^p_{mk}\ \phi^H_K \ = \sum_{i=1}^{k} \sqrt{\rho_i \tau_P h_{im} \phi_K \phi^H_K} + \phi^H_K n^P_m \tag{2}$$

Where ρ_i is the UL pilot power of the i^{th} UE. $n^P_m \varepsilon C^{N \times \tau_p}$ is the received noise matrix i.e. $n^P_m \sim NC\left(0, \sigma^2\right)$ and σ^2 . Now to calculate estimated value of h_{im} take inner product of Y^p_m and ϕ_K , so that we have,

$$Y^p_{mk} = Y^p_{mk}\phi^H_K = \sum_{i=1}^{k} \sqrt{\rho_i \tau_P h_{im} \phi_K \phi^H_K} + \phi^H_K n^P_m \tag{3}$$

Here we assume that $K > \tau_p$ due to which each pilot sequence is shared by more than one UE, leading to so called pilot contamination. Also, we define subset S_K of UEs with same pilot sequence as UE K, so we can write (3) as:

$$Y^p_{mk} = \sum_{i \in S_K} \sqrt{\rho_i \tau_P h_{im} \phi_K \phi^H_K} + \phi^H_K n^P_m \tag{4}$$

Now if we observe (4), then it is clear that mutual interference is created due to sharing of same pilot sequence by UEs leading to so called pilot contamination. Due to pilot contamination, system performance is degraded similar to cellular massive MIMO. Also, channel estimation quality is affected and channel estimates becomes correlated. This correlation is directly proportional to number of AP antennas i.e. 'N'. Both factors will have negative impact on UE performance. In this paper, local CE is done at the serving AP. Now according to standard theory of MMSE estimation, we can calculate channel estimate $\widehat{h}_{mk}$ for K^{th} UE at m^{th}AP which belongs to the subset S_K is given as:

$$\widehat{h}_{mk} = \ \sqrt{\rho_k \tau_p R_{mk} \psi^{-1}_{mk} Y^p_{mk}} \tag{5}$$

where $\psi_{mk} = E\left\{Y_{mk^p}\left(Y_{mk}{}^p\right)\right\} = \sum_{i \in S_K} \rho_i \tau_p R_{im} + I_N$ is the correlation matrix of received signal (4). We can also write (5) as:

$$\hat{h}_{mk} = Z_{mk} Y^p_{mk} \tag{6}$$

where $Z_{mk} = \sqrt{\rho_k \tau_p R_{mk} \psi_{mk}^{-1}}$ and depends on channel statistics. So, we can have channel estimate by multiplying Y_{mk}^{p} with N × N matrix Z_{mk} of each UE served by AP m. The value of Z_{mk} is known prior to AP m and is calculated at AP m. This will reduce the computational complexity at AP m, to a significant extent.

2.2 Pilot Assignment

In this paper, we have considered $K > \tau_p$, and length of the pilot sequence is τ_p, therefore there are only τ_p orthogonal sequences. For number of UEs greater than τ_p, orthogonal pilot sequences are reused leading to so called pilot contamination (PC). Due to PC system performance is affected. Also, in TDD, CSI is obtained during UL training phase which is then further used in DL data transmission. So, to have optimal performance, accurate CSI is important. And PC is introduced in the system during UL channel estimation phase leading to increase in channel estimation error. Therefore, the two aspects first one PC and second one channel estimation error needs to be taken care of. Thus, the channel estimation error caused by PC, will affect spectral efficiency of UE leading to decrease in net throughput of the proposed CFMM system. In order to maximize the system throughput, PC and channel estimation error must be contained. To address PC issue, we have designed an advanced pilot assignment algorithm which is improved version of basic algorithm [15] is discussed below. A unique pilot assignment technique, which is the modified version of the technique mentioned in paper [15] is used for mitigating pilot contamination. Due to pilot contamination, quality of CE gets affected. Consider τ_p mutually orthogonal pilot sequences and pilot sequence of UE K is denoted by φ_k wherein ($||\varphi_k ||)2 = \tau_p$ and $\varphi_k \in$ C ($\tau_p \times 1$). Whenever a UE gains access into the network, a pilot is assigned to the UE depending on the distance between mobile UE and AP. The AP which is nearest to the UE, will serve that UE. If there are more than one APs near to the UE, then β large scale fading factor will decide the serving AP. In this way pilot assignment and AP selection is done through a small algorithm as follows:

1. Accessing UE selection: Here the UE gaining access to the network is selected using a specific process through which first of all, poor channel quality UE are assigned with pilot sequence and then other UEs are considered for pilot assignment. The channel quality of UEs is determined by large scale fading coefficient i.e., beta. This fading coefficient is calculated for all accessing UEs and the values are stored in set S as per descending order. Accordingly, the accessing UE, is selected for further process.

2. In 5G, accessing UE communicates with its neighbouring APs using either primary or secondary synchronization signals. Depending on the nearest distance measurement d_n a serving APAP_n is chosen by the accessing UE.

$$\text{Serving } (AP_n) = \arg(\min d_n) \tag{7}$$

3. If there are more than one equidistant APs in neighbourhood of the accessing UE, then the serving AP is chosen based on the large-scale fading coefficient, β. The serving AP_n will serve the accessing UE with all its N number of antennas. It will assign pilot p to the accessing UE using specific algorithm.

4. The serving AP_n will give information to all neighbouring APs that pilot p is being used by accessing UE. As a consequence of this, other pilot sequences will be considered for transmission apart from pilot p which will reduce PC to a significant extent. Finally, accessing UE will appoint serving AP, the one which is nearest to it and Serving AP will assign unused pilot p to the accessing UE. This will not only ensure reduced PC but also increases system performance to a greater extent as shown in the simulation results.

2.3 Coherent Downlink Data Transmission

In CFMM system, during DL coherent data transmission, all APs serve all the UEs in the same time frequency resources. That is with APs phase synchronised, different APs send the same signal to the same UE. During DL transmission, DL signal sent by $m^{th}AP$ to the $K^{th}UE$ is

$$S_{mk} = \sqrt{\rho_{mk} x_k} \tag{8}$$

where ρ_{mk} is the DL transmission power, and x_k is the data signal sent to the $K^{th}UE$.

Thus, the signal sent by $m^{th}AP$ is given as;

$$Y_{mk}^{p} = \sum_{i=1}^{k} W_{mk} S_{mk} \tag{9}$$

$Where W_{mk} \varepsilon C^{N \times M}$ $is the precoding vector designed by AP$m $for UE$k. $The signal received by K^{th}UE$ $is given by$;

$$Y_{mk}^{DL} = \sum_{l=1}^{M} h_{mk}^{H} W_{ml} S_{ml} + \sum_{i=1, i \neq k}^{k} \sum_{l=1}^{M} h_{ml}^{H} W_{il} S_{mk} + n_k \tag{10}$$

$$Y_{mk}^{DL} = \underbrace{\sum_{l=1}^{M} h_{mk}^{H} W_{ml} S_{ml}}_{\text{desired signal}} + \underbrace{\sum_{i=1, i\neq k}^{k} \sum_{l=1}^{M} h_{ml}^{H} W_{il} S_{mk}}_{\text{MUI}} + \underbrace{n_k}_{\text{noise}} \tag{11}$$

From (11) it is obvious that, the signal is divided into three parts, first term is the desired signal, second term indicates multi-user-interference (MUI) and last term is the noise signal received at the receiver. The performance of the CFMM system is evaluated considering DL SE [bits/s/Hz]. An achievable SE is obtained using hardening bound illustrated in[23]. To apply bounding technique, we rewrite above equation as:

$$Y_{mk}^{DL} = \underbrace{PG_k x_k}_{\text{Precoding Gain}} + \underbrace{PGU_k x_k}_{\substack{\text{Precoding Gain} \\ \text{Uncertainty}}} + \underbrace{\sum_{i\neq k}^{k} MU\ I_{ki} x_i}_{\text{Multi-user interference}} + n_{k\,\text{noise}} \tag{12}$$

Where PG_k is defined as precoding gain and is given by;

$$PG_k = \sum_{l=1}^{M} \sqrt{\rho_{kl}} E(\, h_{kl}^{H} W_{kl}) \tag{13}$$

2.4 Proposed Deep Feed Forward Neural Network Based Precoder

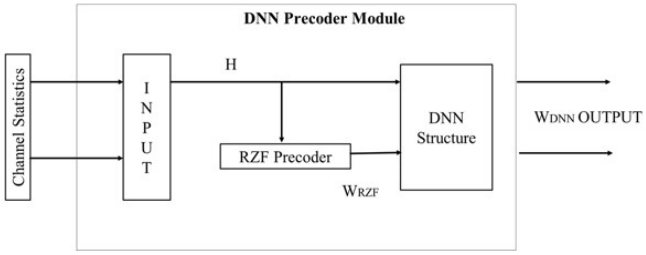

Fig 2. DNN Precoder module for the proposed CFMM system

Here in this section, we propose a new deep learning architecture for precoding of cell-free networks. Let Y = F (X; ψ) indicate a fully connected (FC) deep neural network (DNN). The DNN is supposed to process an input matrix X with a teachable parameter set ψ giving an output matrix Y. The input matrix X represents real and imaginary part of vector, and corresponding real and imaginary part of output vector Y is obtained by using DNN. For L no of DNN layers we have:

$$y_j = f\left(\sum_{j=1}^{L} w_j x_j + b\right) \tag{14}$$

Where y_j is the output, x_j is the input, w_j is the weight, b is the bias of j^{th} layer.

The weight and bias form a set of teachable parameters $\psi \triangleq (w, b, \forall L)$. The final output Y is thus obtained from (14). The block diagram of proposed DNN structure for generating precoding matrix is shown in Figure 2. The input given to the network is the channel state information i.e. channel matrix $H = H_1 \ldots\ldots H_k$, yielding output precoding matrix W. The proposed DNN network initially utilizes regularised zero forcing (RZF) precoder, which helps in extracting important features of optimal precoding matrices. Due to RZF precoder's simplicity, it used as a side information for proposed DNN network. Thus, from channel matrices as input and side information of $W_{RZ\,F}$, DNN precoding matrix W_{DNN} is obtained. The output layer will give DNN-based precoder solution

W_{DNN} which is given by $W_{DNN} = [W_{DNN1} \ldots\ldots W_{DNNk}]$. If we use W_{DNN}

as a precoding vector i.e. 'W' in (11), we will have SINR of

Thus, by using optimal precoder i.e. W_{DNN}, we can maximize or improve achievable sum-rate of K^{th}UE as;

$$\text{SINR}_k^{DNN} = \frac{\left|E\left\{\sum_{l=1}^{M} \sqrt{\rho_{kl}} h_{kl}^{H} W_{DNNk}\right\}\right|^2}{\sum_{i=1}^{k} E\left\{\left|\sum_{l=1}^{M} \sqrt{\rho_{ki}} h_{ki}^{H} W_{DNNk}\right|^2\right\} - \left|E\left\{\sum_{l=1}^{M} \sqrt{\rho_{kl}} h_{kl}^{H} W_{DNNk}\right\}\right|^2 + \sigma_{DL}^2} \tag{15}$$

Thus, by using optimal precoder i.e. W_{DNN} , we can maximize or improve achievable sum-rate of

$$R_{sumrate} = \frac{\tau_{DL}}{\tau_C}(1 + SINR_{DNNK}) \tag{16}$$

The same procedure is summarised in the algorithm (Figure 3) given below.

Algorithm 1 DNN based Precoding Algorithm for Cell-Free Networks

Require: Channel Matrix H,

environment simulator,

Ensure: Optimized Precoder

1.Initially number of iterations are started from zero onwards. Weights and precoder matrix are also considered empty and updated after each iteration.

2.Positionof UEs is randomly considered for generating training sequences.

3.Side information is provided by RZF precoder.

4.Construct the required DNN framework using above information.

5.Use the simulator to simulate Rayleigh wireless channel considering noise.

while Threshold ≥10−7 do

6.Train the DNN network.

7.Update precoding matrix after each iteration.

8.Obtain the bias between W and W_{DNN} .

9.Continue till convergence.

end while

10.Return

Fig 3. Algorithm 1 DNN based Precoding Algorithm for Cell-Free Networks

From comparison point of view DNN-based precoder is compared with other precoding techniques such as MMSE, RZF, MR precoder. The loss function used for training is mean square error (MSE) given by:

$$L\left(\sum_{j=1}^{L} W\ b\right) = \frac{1}{RT}\sum_{R=1}^{R}\sum_{t=1}^{T} \| \mathrm{W} - W_{DNN}\|^2 \tag{17}$$

where T is the training size, and R is the number of realizations used for training. Here we use trained DNN at the previous stage for initialization of the current DNN stage. Due to this training strategy, performance of the training is improved. In this case, we use weights obtained in previous stage to the loss function of the current stage. The process is repeated over multiple training stages until convergence. The DNN structure used for precoding is shown in Figure 4 and DNN hyper parameters are depicted in Table 1. We can generate data set by experimentally calculating channel matrices, but here data set generated by [24] is used for training. The training process involves updating of weights and biases for each layer, according to Xavier initialization method. The DNN is trained by Adam optimiser, exploiting stochastic gradient descent method. The input layer employs parametric rectified linear unit (PRelu), which is also used in initial expansion layers.

When layer size decreases gradually towards the desired output tanh activation function is used for more stable DNN models. It is given by

$$f(x) = \frac{e^{+x} - e^{-x}}{e^{+x} + e^{-x}} \tag{18}$$

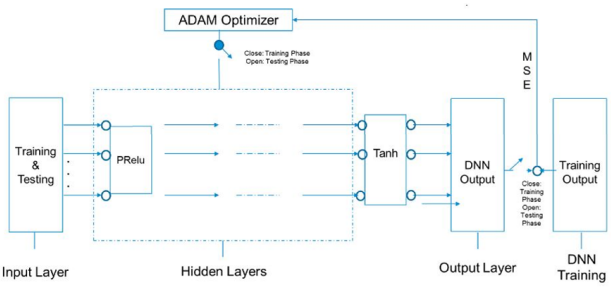

Fig 4. DNN Structure

Table 1. Simulation Specifications for DFF neural network

Parameter	Ranging values
Input Layer neurons	1248
Output Layer neurons	1248
Hidden Layer1neurons	1.6K
Hidden Layer2 neurons	1.2K
Hidden Layer neurons	800
Learning Rate	0.001
Optimizer	Adam
Batch Size	20
Gradient Descent Accuracy	10−8
Activation Function	tanh
Loss Function	MSE

Out of 250880 [24] realizations we have used 200000 only. For training data 140,000 realizations are used, and 30,000 for validation and data testing each respectively. The DFF neural network is trained for 900 iterations, with a learning rate of 0.001. The learning rate depicts how fast the DNN is able to change the weights and biases. For each iteration, mini batch size of 50 is considered, which indicates that after every 50 samples weights are updated. The number of epochs and batch size is selected after trial-and-error method. After training, performance is evaluated for test data set which gives us desired output. There are 3 hidden layers each with 1.6K,1.2K and 800 neurons respectively. After training the data, the DNN-based precoder is ready for testing.

3 Result and Discussion

In this section, we analyse performance of the proposed CFMM system under Rayleigh fading channel model. We consider a scenario in which the total coverage area for cell-free setting is 1 × 1Km2 area, with K single antenna UEs and M APs with N number of antennas. The large-scale fading coefficient is given by:

$$\beta_{mk} = 10^{\frac{PL_{mk} + Z_{mk}\sigma_{sh}}{10}} \tag{19}$$

Where Z_{mk} is a random variable with gaussian distribution, with σ_{sh} equal to 8db. PL_{mk} is the path loss for m^{th} AP and k^{th}UE, which is represented by three slope path loss model

given by:

$$PL_{mk} = \begin{cases} -P - 35log_{10}d_{mk} \ldots\ldots (if\ d_{mk} < 50m) \\ -P - 15log_{10}50 - 20log_{10}d_{mk} \ldots\ldots (if\ 10m < d_{mk} \leq 50m) \\ -P - 15log_{10}50 - 20log_{10}10 \ldots\ldots (if\ d_{mk} \leq 10m) \end{cases} \tag{20}$$

where P is 140.7db, which is a constant and depends on carrier frequency f_c, UE height is h_{UE} and AP height is h_{AP}, and d_{mk} is the distance between m^{th} AP and k^{th} UE. The APs are 10m above the UEs. We consider τ_{DL}= 190, τ_p = 10, τ_c = 200, channel bandwidth B = 20MHz, UL transmit power p_k = 100mW, DL transmit power p = 1W. We assume each UE transmit at full power and each AP operate at maximum transmit power. Each AP will serve UE with good channel condition as well as UE with worst channel condition by allocating power to UEs appropriately. This will guarantee non zero SE curve as shown in the simulation results. The performance analysis of the proposed CFMM system is evaluated considering two parameters time consumption and numerical results. These two aspects are discussed below. Time consumption parameter is used to compare time required to achieve desired solution. The conventional precoders are compared with proposed DNN-based precoder. The conventional methods require approximately 20ms to 30ms of time to achieve desire solution, while on the other hand DNN-based precoder is much faster consuming only 1ms to 2ms of time, saving approximately 90percent of run-time. So, this proposed method can be used for IoT applications. We know that training a particular DNN network is time consuming. But training is done only once and in offline mode. Once the network is trained, DNN framework will give desired output in few seconds.

3.1 Numerical Results

In this section, the performance of the proposed CFMM system is compared with linear and non-linear precoding techniques. The performance of the proposed system is also evaluated against different parameters i.e. perfect and instantaneous CSI, varying pilot lengths, coherent and non-coherent data transmission. In Figure 5 comparison of the proposed scheme with linear i.e. RZF, MRT and non-linear i.e. MMSESIC combining techniques is shown. It depicts CDF of DL SE for various schemes, considering M=400, N=1 and K=50 set-up. From the CDF curve, 90percent likely values for CFMM-MMSESIC, proposed scheme, CFMM-RZF, and CFMM-MRT are 3 bits/s/Hz, 2.5bits/s/Hz, 1.5 bits/s/Hz, 1.25 bits/s/Hz respectively. It is expected for the proposed scheme to give performance very close to MMSESIC combining technique. From CDF curves, it is clear that the proposed scheme outperforms RZF and MRT schemes and gives performance curve which is very close to MMSESIC i.e. non-linear combining technique. So, the proposed scheme gives near optimal results as expected. From graph it is also clear that we can use simple linear combining techniques instead of complex non-linear combining techniques to get near optimal results. This will further reduce the computational complexity of the proposed CFMM system.

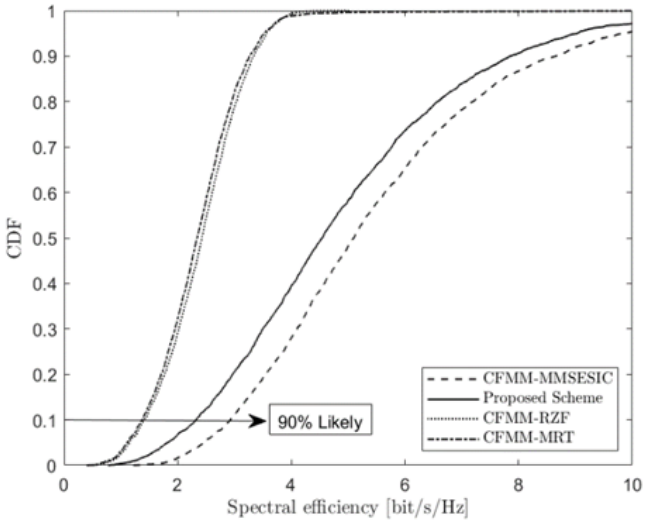

Fig 5. DLSE Comparison of the proposed scheme with MMSESIC, RZF and MRT combining techniques

Figure 6 analyses the CDF against the sum SE of proposed scheme verses RZF and MR precoding techniques for M=100, N=4 and K=50. The graph is plotted for different random UE locations. From the graph it is obvious that the proposed scheme gives better performance as compared to RZF and MR precoding techniques. The huge performance gap between the proposed scheme and RZF, MR scheme emphasizes the importance of using deep learning for precoding techniques.

The comparison of the proposed scheme with MMSESIC with coherent and non-coherent data transmission is shown in Figure 7. It shows CDF of DL SE per UE. From 80 percent likely values on CDF curve, coherent data transmission is better than non-coherent data transmission. Also, the proposed scheme gives performance curve, which is very close to MMSESIC scheme

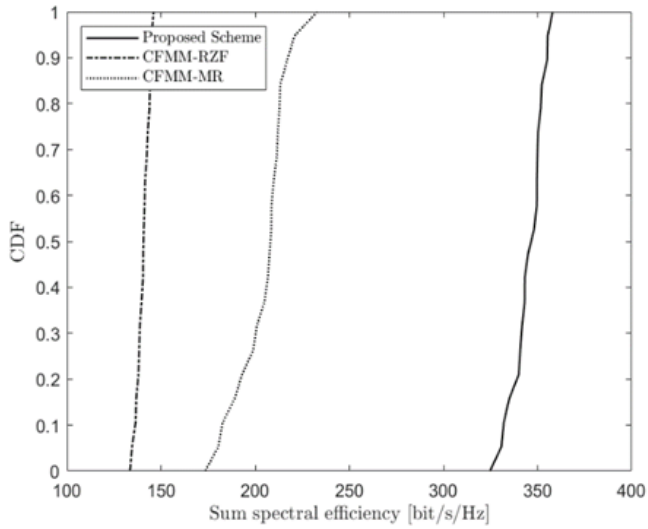

Fig 6. CDF of sum SE of the proposed scheme compared with MR and RZF combining techniques

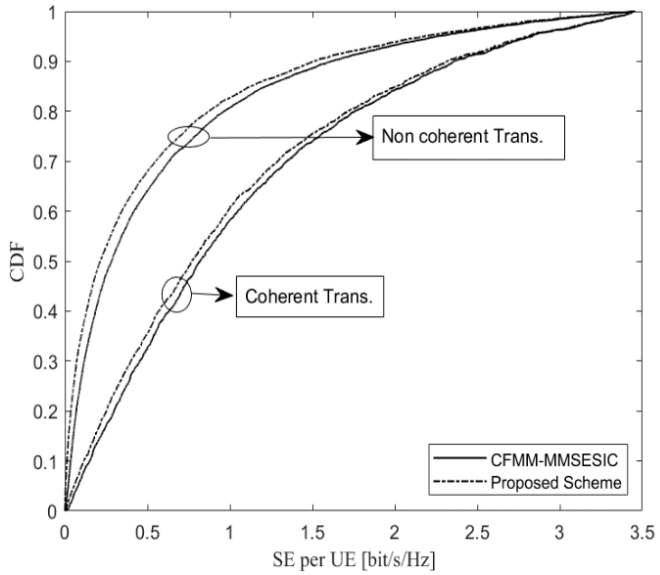

Fig 7. Comparison of the proposed scheme with MMSESIC with coherent and non-coherent transmission

as desired.

The comparison of the proposed scheme with LS combining with different pilot lengths for coherent transmission is depicted in Figure 8. It shows the CDF of SE per UE with pilot length 5 and 15 for M=100, N=4 and K=50. From 70 percent likely values of CDF curve, it is seen that the proposed scheme gives better performance than the other one, in both cases when pilot length is 5 and 15. It shows the impact of varying pilot lengths on the performance of the system. For varying pilot length, the performance gap between LS and the proposed system also changes accordingly. This fact proves that the performance of the system depends on pilot length. Increasing pilot length will improve estimation quality for LS combining technique by reducing pilot contamination. Thus, the performance loss of the system depends on the degree of pilot contamination. Performance loss is less when τ_p = 15 and it is more when τ_p= 5.

In Figure 9 the CDF of DL SE with perfect and instantaneous CSI is shown. The proposed scheme is compared with perfect CSI and instantaneous CSI. From CDF curve, 90 percent likely values for perfect CSI and instantaneous CSI are 1.15 bits/s/Hz,

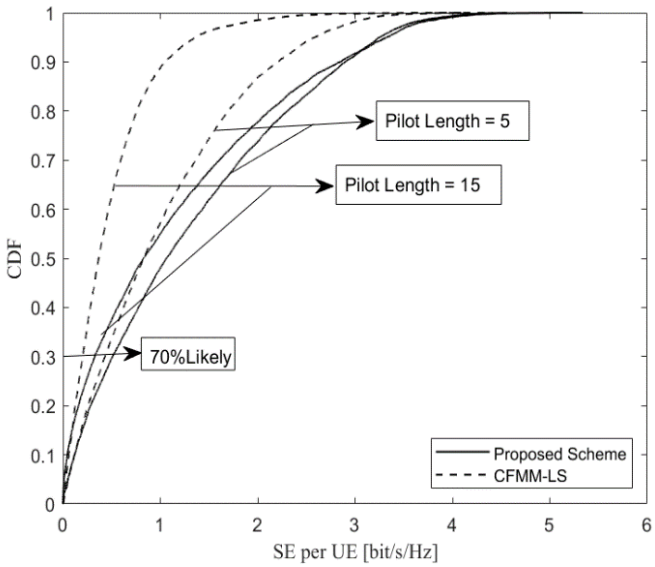

Fig 8. Comparison of the proposed scheme with LS combining technique with different pilot lengths

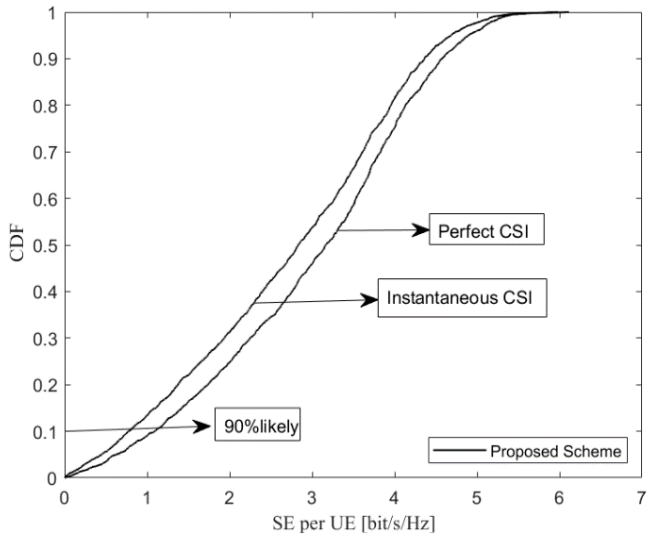

Fig 9. Comparison of the proposed scheme with perfect CSI and instantaneous CSI

and 0.85 bits/s/Hz respectively. From graph, performance of the proposed scheme is better with perfect CSI as compared to instantaneous CSI. But the performance gap between the two is very small due to the fact that, for DNN based precoder perfect CSI is not required. Even without knowing the perfect CSI it gives better performance as shown in the Figure 9.

The comparison of average DL SE of the proposed scheme with MMSE, MR and RZF techniques is seen in Figure 10. Here average DL SE is considered for different number of APs. Performance increases with increase in number of APs as seen from Figure 10. From the result it is obvious that the proposed scheme outperforms other techniques with a huge performance gap re-emphasizing the importance of using DNN-based precoder. Thus, from results it is clear that DNN based precoder increases the system performance to a greater extent.

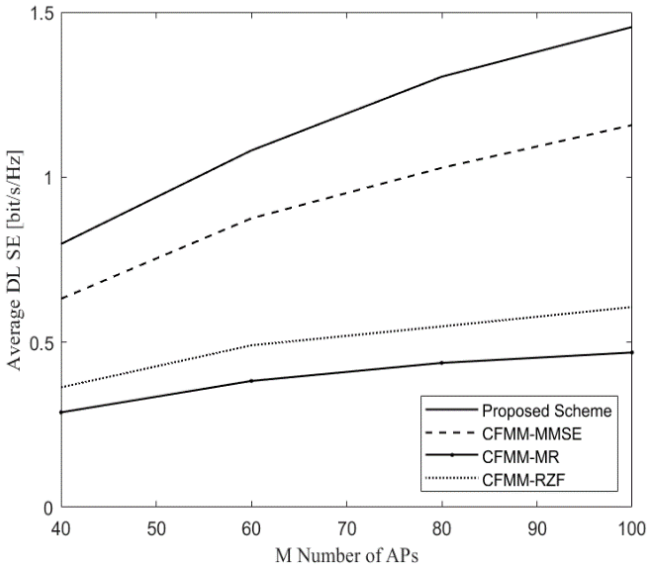

Fig 10. Average DL SE of the proposed scheme compared with MMSE, MR and RZF combining

4 Conclusion

This paper developed a new framework for DL CFMM system under Rayleigh fading channel model incorporating learning aided precoding with enhanced basic pilot assignment algorithm. First, we developed deep learning- based precoder, enhancing the system performance by reducing fronthaul requirement, since perfect CSI is not required. DNN-based precoder is trained using previous readings so it gives optimal results even in the absence of perfect CSI. Secondly an improved version of scalable basic pilot assignment algorithm is developed to mitigate pilot contamination, thereby enhancing the system performance. The simulation result shows that:

- From the CDF curve, 90percent likely values for CFMM-MMSESIC, proposed scheme, CFMM-RZF, and CFMM-MRT are 3 bits/s/Hz, 2.5bits/s/Hz, 1.5 bits/s/Hz, 1.25 bits/s/Hz respectively. It clearly indicates, that we can use the proposed system instead of complex CFMM-MMSESIC for near optimal results.
- Sum spectral efficiency curve depicts that the proposed scheme gives better performance as compared to RZF and MR precoding techniques.
- With coherent data transmission, the proposed scheme gives performance curve, which is very close to MMSESIC scheme as desired.
- The performance of the proposed system depends on pilot length.
- From CDF curve, 90 percent likely values for perfect CSI and instantaneous CSI are 1.15 bits/s/Hz, and 0.85 bits/s/Hz respectively, which depicts better performance of the proposed scheme with perfect CSI as compared to instantaneous CSI.
- The average DL SE of the proposed scheme increases with increase in number of APs.

With near optimal results obtained for the proposed scheme, one important aspect which is power allocation, can be considered as future work.

References

1) Khanna H, Aggarwal M, Ahuja S. Statistical characteristics and performance evaluation of FSO links with misalignment fading influenced by correlated sways. *AEU - International Journal of Electronics and Communications*. 2018;85:118–125. Available from: https://doi.org/10.1016/j.aeue.2017.12.032.
2) Sharma R, Aggarwal M, Ahuja S. Performance Analysis of Indoor FSO Communication Systems under Receiver Mobility. *2016 International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE)*. 2016;p. 652–657. Available from: https://doi.org/10.1109/ICMETE.2016.61.
3) Aggarwal M, Garg P, Puri P, Sharma PK. Performance analysis of optical wireless communication system with a decode and forward relay. *2014 International Conference on Signal Processing and Integrated Networks (SPIN)*. 2014;p. 333–337. Available from: https://doi.org/10.1109/SPIN.2014.6776973.

4) Xu S, Cao Y, Li C, Wang D, Yang L. Spanning Tree Method for Over-the-Air Channel Calibration in 6G Cell-Free Massive MIMO. *IEEE Transactions on Wireless Communications*. 2023;22(8):5567–5582. Available from: https://doi.org/10.1109/TWC.2023.3235355.
5) Wang CXX, You X, Gao X, Zhu X, Li Z, Zhang C, et al. On the Road to 6G: Visions, Requirements, Key Technologies, and Testbeds. *IEEE Communications Surveys & Tutorials*. 2023;25(2):905–974. Available from: https://doi.org/10.1109/COMST.2023.3249835.
6) Murthy KC. Performance evaluation of chaotic spreading codes in massive MIMO OFDM system. *Indian Journal of Science and Technology*. 2020;13(42):4374–4385. Available from: https://doi.org/10.17485/IJST/v13i42.2027.
7) Matta JCP, Siddiah P. A Modified OMP Algorithm with Reduced Feedback Overhead for Massive MIMO System. *Indian Journal of Science and Technology*. 2021;14(33):2663–2670. Available from: https://doi.org/10.17485/IJST/v14i33.1442.
8) Abose TA, Olwal TO, Hassen MR. Hybrid Beamforming for Millimeter Wave Massive MIMO under Multicell Multiuser Environment. *Indian Journal of Science and Technology*. 2022;15(20):1001–1011. Available from: https://doi.org/10.17485/IJST/v15i20.114.
9) Obakhena HI, Imoize AL, Anyasi FI, Kavitha KVN. Application of cell-free massive MIMO in 5G and beyond 5G wireless networks: a survey. *Journal of Engineering and Applied Science*. 2021;68(1):1–41.
10) Apiyo A, Izydorczyk J. A Survey of NOMA-Aided Cell-Free Massive MIMO Systems. *Electronics*. 2024;13(1):231–231. Available from: https://doi.org/10.3390/electronics13010231.
11) Han L, Ang LW, Palaniappan S. Linear digital precoding technology in massive multiple input multiple output wireless communication system. *Wireless Networks*;5:1–7. Available from: https://doi.org/10.1007/s11276-023-03482-7.
12) Liu H, Zhang J, Jin S, Ai B. Graph Coloring Based Pilot Assignment for Cell-Free Massive MIMO Systems. *IEEE Transactions on Vehicular Technology*. 2020;69(8):9180–9184. Available from: https://doi.org/10.1109/TVT.2020.3000496.
13) Deshpande S, Sharma P, Aggarwal M, Ahuja S. Mitigating pilot contamination in Rician faded massive MIMO 5G systems using enhanced zero forcing precoding and ring partitioning. *Physical Communication*. 2021;49:101467–101467. Available from: https://www.sciencedirect.com/science/article/abs/pii/S1874490721002044?via%3Dihub#:~:text=https%3A//doi.org/10.1016/j.phycom.2021.101467.
14) Victor CM, Mvuma AN, Mrutu SI. Multi-input fully CNN for joint pilot decontamination and symbol detection in 5G massive MIMO. *IET Communications*. 2023;17(16):1899–1906. Available from: https://doi.org/10.1049/cmu2.12670.
15) Polegre AA, Sanguinetti L, Armada AG. Pilot Decontamination Processing in Cell-Free Massive MIMO. *IEEE Communications Letters*. 2021;25(12):3990–3994. Available from: https://doi.org/10.1109/LCOMM.2021.3118890.
16) Ayidh A, Sambo A, Imran Y, A M. Mitigation pilot contamination based on matching technique for uplink cell-free massive MIMO systems. *Scientific Reports*. 2022;12(1):16893–16893. Available from: https://doi.org/10.1038/s41598-022-21241-0.
17) Zheng J, Zhang J, Bjornson E, Ai B. Cell-Free Massive MIMO with Channel Aging and Pilot Contamination. *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*. 2020;p. 1–6. Available from: https://doi.org/10.1109/GLOBECOM42002.2020.9322468.
18) Misso A, Kissaka M, Maiseli B. Exploring pilot assignment methods for pilot contamination mitigation in massive MIMO systems. *Cogent Engineering*. 2020;7(1):1831126–1831126. Available from: https://doi.org/10.1080/23311916.2020.1831126.
19) Jo S, Lee J, So J. Deep learning-based massive multiple-input multiple-output channel state information feedback with data normalisation using clipping. *Electronics Letters*. 2021;57(3):151–154. Available from: https://doi.org/10.1049/ell2.12080.
20) Lin WYY, Chang THH, Tseng SMM. Deep Learning-Based Cross-Layer Power Allocation for Downlink Cell-Free Massive Multiple-Input–Multiple-Output Video Communication Systems. *Symmetry*. 1968;15(11):1968–1968. Available from: https://doi.org/10.3390/sym15111968.
21) Interdonato G, Karlsson M, Bjornson E, Larsson EG. Local Partial Zero-Forcing Precoding for Cell-Free Massive MIMO. *IEEE Transactions on Wireless Communications*. 2020;19(7):4758–4774. Available from: https://doi.org/10.1109/TWC.2020.2987027.
22) Han L, Ang LW, Palaniappan S. Linear digital precoding technology in massive multiple input multiple output wireless communication system. *Wireless Networks*;5:1–7. Available from: https://doi.org/10.1007/s11276-023-03482-7.
23) ¨ornson EB, Sanguinetti L. Scalable cell-free massive MIMO systems. *IEEE Trans Commun*;68. Available from: https://doi.org/10.1109/TCOMM.2020.2987311.
24) Ha AL, Van Chien T, Nguyen TH, Choi W, Van Duc Nguyen. Deep Learning-Aided 5G Channel Estimation. *15th International Conference on Ubiquitous Information Management and Communication (IMCOM)*. 2021. Available from: https://ieeexplore.ieee.org/document/9377351.