IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON MOBILE MULTIMEDIA: METHODOLOGY AND APPLICATIONS

Received November 7, 2019, accepted November 24, 2019, date of publication December 6, 2019,
date of current version December 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2958060

Channel Estimation Capacity Enhancement for Multigroup Multicasting Multimedia Networks With DnCNN

TIANYI ZENG, YAFENG WANG, JUNYAO LI, AND SHUAI HOU
Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, China
Corresponding author: Yafeng Wang (wangyf@bupt.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61871049.

ABSTRACT In Time Domain Duplex (TDD) massive MIMO systems, multi-group multi-casting becomes a promising technology since it supports services of mass content distribution. Based on the nature of transmitting common message to groups of users simultaneously, there exists a rich literature discussing the resource allocation under various constraints. However, the practical acquisition of CSI has not been fully explored when the number of multi-groups is large and the band is narrow. The insufficient sounding reference signal resources lead to the limited Channel Estimation Capacity (CEC). Under this case, even with Multi-User (MU) channel estimation techniques, some users still cannot be estimated in-timely, which introduces degradation. Aiming at this problem, in this paper we provide a preliminary exploration on CEC enhancement. Based on Denoising Convolutional Neuron Network (DnCNN), which is recently proposed and has succeeded in image denoising, we propose MU-DnCNN Channel Estimation (M-DnCNN CE). M-DnCNN CE includes three parts. First, we modify the utilization of SRS sequences. Then we establish the feature maps and propose M-DnCNN to denoise the signals. Finally, a matched channel restoration method is provided. The practical 3-D MIMO channel model is utilized to evaluate the performance. Compared with DFT-based and conjugated separation methods, results show that the performance of M-DnCNN CE is robust and superior, and the CEC is remarkably improved on the premise of satisfying latency constraint.

INDEX TERMS Multi-user channel estimation, multi-group multi-casting, DnCNN, 3-D MIMO, mobile intelligence.

I. INTRODUCTION

With the rapid development of 5G technologies, such as network slice [1], physical layer coding [2], D2D [3]–[5], mobile edge caching [6], [7] and other IoT technologies [8]–[11], many streaming media application services emerge, and the services targeted at mass content distribution is expected to be popular [12]. Of particular interest is Multi-Casting (MCa), which refers to the case where multiple messages are transmitted simultaneously but each message is addressed to a group of users [13]. These users may require similar location-based information or some multimedia entertainments such as mobile TV, IPTV and other applications. Under this case, compared with the inefficient traditional unicast technologies that ignore the nature of traffic demand for common messages, Multi-Group Multi-Casting (MGMCa) becomes a promising technology [14]. The 3rd Generation Partnership Project (3GPP) is discussing details of corresponding techniques, such as evolved Multimedia Broadcast Multicast Service (eMBMS) and so on.

To enhance the system performance of MCa services, there exists a rich literature investigating the efficient resource allocation technologies, which includes the design of precoder, the group formation, the reduction of computational complexity and the combination with emerging techniques [15]–[25]. The constraints considered in those researches are in varied forms, such as per-cell power constraints, capacity of backhaul and single-cell or multi-cell scenarios. However, the most fundamental premise of these works is the knowledge of channel information, which can also be coined as imperfect or perfect Channel State Information at Transmitter (CSIT). More precisely, only in [15],

The associate editor coordinating the review of this manuscript and approving it for publication was Dapeng Wu.

[17], [19] the imperfect CSIT is researched, while the others tend to assume perfect CSIT and partial CSIT. Just as pointed in [14], the channels are assumed to be slow-fading, perfectly estimated and always available for Base Stations (BS). This assumption of CSIT is too ideal. Much work is remained to coordinate the this assumption with the practical acquisition of CSI, especially in following two cases.

First, the scenario of large amount of MCa groups has not been explored yet, which is the common case of imperfect CSIT. Actually, though perfect channel estimation is assumed, the accurate CSI acquired by BS is still limited and unable to support the MCa for vast quantity of MCa groups, because with the increasing number of antenna of user equipments, the resources of Sound Reference Signal (SRS) becomes limited [26]. Besides, if we consider the Doppler effect [22], this insufficiency of CSI (imperfect CSIT) will be aggravated. Note that this problem is different from the case of [15]. More precisely, some CSI errors of this problem are caused by channel aging, and the practical 3-D MIMO channels show different characteristics from Rayleigh channels, thus the assumption of Gaussian-distributed errors is questionable and the asymptotic theory may not be valid.

Second, for some multi-rate MCa transmission technologies, such as rate-splitting where the unicast and MCa exist at the same time, the individual CSI of each user is also needed. In addition, the sufficient individual CSI can help BS make better use of Degree of Freedom (DoF) with techniques such as Block Diagonalization (BD) precoder to serve multi-antenna users [21]. Under this case, letting intra-group users use the same pilot to reduce the overhead and suppress the pilot contamination [19], [27] is not a preferred scheme. However, it is a questionable problem whether the current Multi-User (MU) channel estimation technologies can support the estimation of a large amount of intra-group users.

One straightforward solution of two aforementioned cases is to enhance the Channel Estimation Capacity (CEC), which means to increase the number of estimated user on narrow band. Though Time Domain Duplex (TDD) massive MIMO systems have already facilitated the acquisition of CSI, directly increasing the number of MU to be estimated on a narrow band is not realistic. Given that the Constant Amplitude Zero-Autocorrelation (CAZA) SRS sequences are based on phase shift, conventional Discrete Fourier Transmission (DFT) based approaches will let desired taps overlap [28] and the Time Domain (TD) denoising will be serevely degrade, while the Conjugated Separation (CS) methods introduce the errors caused by the inequality of adjacent subcarriers. Thus, new technologies should be considered to enhance CEC.

Recently, Denoising Convolutional Neuron Network (DnCNN) becomes the promising technique to tackle with image denosing in Computer Vision (CV). In [29], DnCNN is firstly proposed to deal with the denoising tasks of gray and color image. By adopting the residual network structure proposed by K. He et al. [30], [31], a wider and deeper CNN can be implemented without concerning about the gradient explosion or vanishment. Given a CNN structure, the determination of the network weights can be formulated as non-convex optimization programming. Compared to the algorithms that include the manual feature extraction procedures, for example, the BM3D or C-BM3D where the block-matching is needed [32], the end-to-end DnCNN can learn the features by itself with gradient descent. Though the non-convex optimization of deep networks is mathematically intractable, DnCNN can offer the state-of-the-art performance.

Though estimating users with deep learning techniques has the potential to be hot topic in mobile intelligence, directly adopting DnCNN in MU channel estimation is not realistic because of three reasons. First, DnCNN is used to process real-numbers, while the Channel Frequency Responses (CFRs) are complex-numbers. Second, the MU CFRs are not matrix-wise data like images. Third and perhaps the most significant, the end-to-end latency of utilizing DnCNN to implement MU channel estimation and enhance CEC must satisfy the practical constraints of 5G.

To alleviate the problem of insufficient CSI in MGMCa, this paper focuses on MU CEC enhancement of narrow band based on DnCNN. Facing all of aforementioned challenges above, our contributions include:

1) Based on complex-number DnCNN which is introduced in [33], we provide a feasible MU CEC enhancement method of narrow band. The proposed method includes three parts: SRS modification, MU-DnCNN (M-DnCNN) denoising and channel restoration. In SRS restoration, we let MU use SRS generated from different SRS sequence. In M-DnCNN denoising, we first preprocess the CFRs and design the suitable network whose parameters are given by comprehensively considering the time complexity and performance. In channel restoration, we separate the MU channels based on the results outputed by M-DnCNN.
2) The practical 3-D MIMO channel model is utilized to evaluate the proposed method. The Normalized Mean Square Error (NMSE) performance is provided. Besides, the time complexity, operation time and other parameters are also evaluated to prove the feasibility to meet practical constraints.

Since here we aim to open the eyes of academia to implement MU channel estimation with DnCNN or alternative deep learning techniques, this paper only presents the preliminary exploration based on simple configurations such as narrow band and single cell. We emphasize that this issue is inspired by MGMCa scenario but not limited to this case.

The organization of this paper is as follows. In Section II, the related work including 3-D MIMO channel model, SRS sequence generation and DnCNN is presented. In Section III, we introduce the proposed M-DnCNN Channel Estimation (M-DnCNN CE). Section IV provides simulation results. Conclusions are given in Section V.

Notations: Capital and lowercase bold letters denote the matrix and vector, respectively. $\mathbf{x}(i)$ stands for the i^{th} element of vector $\mathbf{x}$, while $\mathbf{X}(i, j)$ denotes the element in the i^{th} row and i^{th} column of matrix $\mathbf{X}$. The $N \times N$ identity matrix is denoted as $\mathbf{I}_N$. $(\cdot)^{\text{H}}$ stands for Hermitian transpose. $\bar{x}$ denotes conjugated value of scalar x. $\|\cdot\|_2$ denotes l_2 norm of vectors.

II. REVIEW OF PREVIOUS WORK

The related work includes three parts, namely the channel model adopted in the simulation, the previous MU channel estimation methods and DnCNN.

A. 3-D MIMO CHANNEL MODEL

In order to make better use of spatial diversity, 3-D MIMO becomes a promising technology. In 3-D MIMO systems, the BS is equipped by a two-dimensional uniform rectangular array, which facilitates the BS to utilize the vertical dimension and enhance the accuracy of beams. In [34], the generation method of 3-D channel is provided. The calculation of channel coefficient between any pair of BS antenna and user antenna is given in [34]. With the channel coefficients, the TD wireless channels of can be obtained. By implementing DFT on TD Channel Impulse Responses (CIRs), the CFRs of frequency domain can be calculated.

B. EXISTING MU CHANNEL ESTIMATION METHODS

Numerous works with respect to channel estimation are provided in previous reaserches, but they are mainly focused on single-user channel estimation. Classical pilot-aided channel estimation methods, such as Least Square (LS), DFT and Minimum Mean Square Error (MMSE), are well studied in terms of estimation accuracy, time complexity and other practical constraints. By contrast, the researches of MU channel estimation are still needed to be explored. During estimating MU wireless channels, the received signals of BS should not only be denoised, but the superimposed signals be separated. Under this case, the phase shift of SRS sequences are of great help. With the characteristic that different users have equal inter-separation of phase, the MU channel estimation methods can be divided into two categories, namely the CS method and DFT-based approaches.

CS is an engineering solution of MU channel estimation. In CS method, it is proposed to denoise and separate the MU channels in frequency domain. Given that CS method is based on the phase difference caused by cyclic shift, the pilot points on the same subcarriers between different MUs are generated with the same base sequence but different phases. The phase difference is caused by cyclic shift indexes of antenna ports (which can be denoted with $n_{\text{SRS}}^{\text{cs},\hat{p}}$, and the details are provided in III-A). Under this case, LS method is the first step of CS method, which can help to eliminate the phase shift of the user to be estimated while shift the phases of other MU users with fixed proportion. After that, by averaging the signals on adjacent subcarriers, the estimated channel is obtained. Repeating the procedures for subcarriers of MU users, we can obtain the channels. However, CS method is just a preliminary MU estimation approach. CS method almost does nothing about denoising but separates the channels of MU. When the channels on a number of adjacent subcarriers are not very similar, the performance will suffer from two aforementioned kinds of errors together and degrade severely.

Different from the CS method, DFT-based approaches are proposed to estimate MU channels in TD. DFT-based methods are based on the property that the desired energy concentrates on several taps while the noise spreads the whole TD region. In Classical DFT-based (C-DFT) method, four steps are included, namely LS estimation, Inverse DFT (IDFT) transformation, extra-window denoising and intra-window denoising. With the help of the property that different users have different cyclic shift indexes, by implementing LS estimation and IDFT transformation we can let the desire taps of different MU users locate on different region. Extra-window denoising means to extract the desired taps of the user under estimation, while intra-window denoising stands for nulling the noise taps in the window with the denoising threshold, which can be calculated according to the assumed number of significant taps. The drawbacks of C-DFT are obvious. First, it fails to utilize the property of multiple antennas. Second, it did not discuss how to determine the number of significant taps, which may introduce severe degradation when the significant taps number are apparently different in practical wireless channels. In our previous work [28], the Enhanced DFT-based (E-DFT) method is proposed. E-DFT is based on that the number of significant taps between Line-of-Sight (LoS) and None-Line-of-Sight (NLoS) users are much different, and includes LoS/NLoS identification, the donoising of LoS/NLoS users and existence test. During the the donoising of LoS/NLoS users, the determination of significant tap number is introduced, and the existence test zeroing the taps using the spatial stationary property of 3-D channel model. However, even with E-DFT, it is not realistic to enhance the CEC capacity. More precisely, for the case of narrow band, the number of taps in the window is inverse proportional to the number of MU users. Under this case, the intra-window denoising will be invalid, which is the major challenge in CEC enhancement for MU chanel estimation.

C. DnCNN

DnCNN is a rising technology for image denoising. There are several image denoising techniques existing besides DnCNN, such as BM3D, LSSC and WNNM. By assuming nonlocal self-similarity (NSS) models to exploit image priors, these methods can achieve decent performance. However, they have two major drawbacks compared with DnCNN. First, the models involved in those methods are usually non-convex problems and the parameters are manually chosen, there always exists room to improve. Second and more important, the optimization of those methods is time-consuming and complex [29]. Given that the latency is a significant constraint in practical cases and the end-to-end learning of DnCNN is a

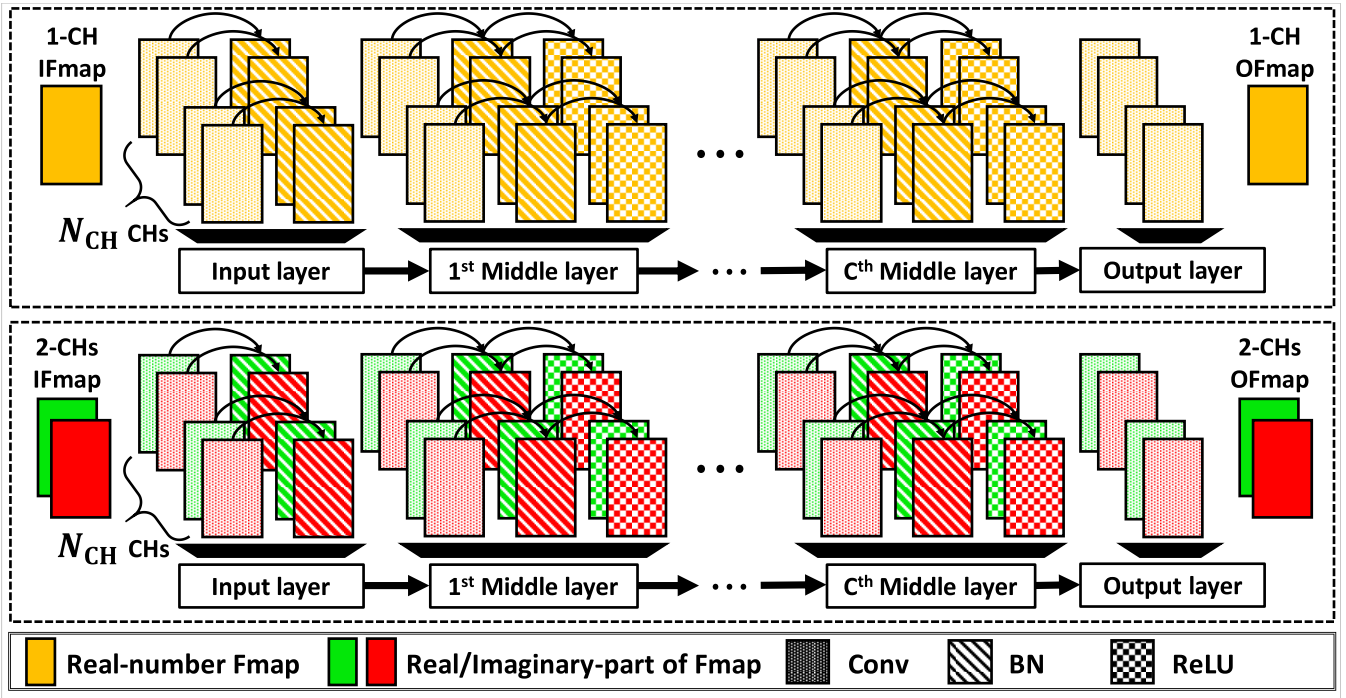


FIGURE 1. The pipelines of common real-number and complex-number DnCNNs.

natural advantage, DnCNN is more preferred to combine the channel estimation in wireless communications.

In deep learning based denoising, gradient descent is the key method to search the local optima for problems. The challenge of conventional deep learning based regression is the vanishing or exploding gradient. DnCNN solves this problem based on the researches in [30], [31]. There are three important components as follows.

• Convolutional layer (Conv). The convolution is done with block-wise and the unit of convolution is filter, which is equal to the channel number of the next layer. Note that the channel in DnCNN represents the dimension of data and is different from the wireless channel. To distinguish from the wireless channel, we coin the channel in DnCNN as 'CH'. Conv is used to extract high-layer features from the Input Feature maps (IFmaps). The results can be called with Output Feature maps (OFmaps). Conv finds the features in a larger area and reduces the complexity compared to full connection.

• Batch Normalization (BN). BN helps to resolve the gradient explosion or gradient vanishing. For a very deep network, the changes of amplitude will be accumulated with extracted features passing the layers. If the gradient is small, then only the weights of first several layers can be updated and the change of the layers afterwards can be neglected. On the other hand, if a large gradient exists in the first several layers, the weights of last several layers will fluctuates severely and even obtain invalid extreme values. BN solves this problem by normalizing the data over a batch of samples. Then for each layer, its IFmaps can be regarded as ''original samples''. Though it may sacrifice some features, BN guarantee the updating steps as valid values.

• Restricted Linear Unit (ReLU). ReLU introduces non-linearity. It filters the negative values while keeps the non-negative values unchanged. ReLU is the common used non-linearity unit for Convolutional Neuron Networks (CNNs).

The Conv, BN and ReLU together form the middle layer of DnCNN. In [29], 20 layers are configured for blind denoising. Besides the middle layers, the input layer is consisted of Conv and BN, while only one Conv is used as the output layer. With this structure, DnCNN denoises the images by learning Gaussian noise rather than the content of images. The noise is the factor that more impacts the accuracy. This makes DnCNN an extendable denoising technique for the various desired information polluted by Gaussian noise.

As the signals in wireless communications are always complex-numbers, complex-number DnCNN is needed to process the wireless signals. Since the Cauchy-Riemann condition is strict, the complex-number is not differentiable anywhere. The method which separates the real-part and imaginary part of complex-number into adjacent CH, which is introduced in [33], becomes a simple but feasible way to handle this case. Besides, the complex-number BN and ReLU is introduced as well. The common real-number and complex-number DnCNNs are depicted in Fig. 1.

III. SYSTEM MODEL

A. SRS SEQUENCE GENERATION

According to [40], the generation of SRS sequence includes two parts, namely cyclic shift and the base sequence. The SRS

sequence generation can be described with

$$\mathbf{r}_{\hat{u},\hat{v}}^{\hat{p},n_{\mathrm{SRS}}^{\mathrm{cs}}}(m)=\exp(jm\alpha_{\hat{p},n_{\mathrm{SRS}}^{\mathrm{cs}}})\hat{\mathbf{r}}_{\hat{u},\hat{v}}(m),\quad 0\le m<M_{\mathrm{sc}}^{\mathrm{RS}} \tag{1}$$

where $\mathbf{r}$ and $\hat{\mathbf{r}}$ stand for SRS sequence and base sequence, respectively, $\hat{u}$ denotes the group number, $\hat{v}$ denotes the base sequence number within the group, $M_{\mathrm{sc}}^{\mathrm{RS}}$ represents the length of SRS, and for the $(n_{\mathrm{SRS}}^{\mathrm{cs}})^{\mathrm{th}}$ sequence and the antenna port $\tilde{p}$, the phase of cyclic shift $\alpha_{\hat{p},n_{\mathrm{SRS}}^{\mathrm{cs}}}$ is calculated with

$$\alpha_{\hat{p},n_{\mathrm{SRS}}^{\mathrm{cs}}}=2\pi\cdot n_{\mathrm{SRS}}^{\mathrm{cs},\hat{p}}/n_{\mathrm{SRS}}^{\mathrm{cs,max}}, \tag{2}$$

where the $n_{\mathrm{SRS}}^{\mathrm{cs,max}}$ is the number of cyclic shift and

$$n_{\mathrm{SRS}}^{\mathrm{cs},\hat{p}}=\left(n_{\mathrm{SRS}}^{\mathrm{cs}}+n_{\mathrm{SRS}}^{\mathrm{cs,max}}\hat{p}/N_{\mathrm{ap}}\right)\bmod n_{\mathrm{SRS}}^{\mathrm{cs,max}}, \tag{3}$$

where N_{ap} is the number of total antenna ports.

The calculation of base sequence $\hat{\mathbf{r}}$ can be given by

$$\hat{\mathbf{r}}(m)=\mathbf{x}_q\left(m\bmod N_{\mathrm{ZC}}^{\mathrm{RS}}\right),\quad 0\le n<M_{\mathrm{sc}}^{\mathrm{RS}}, \tag{4}$$

where $\mathbf{x}_q$ denotes the q^{th} root Zadoff-Chu sequence, and $N_{\mathrm{ZC}}^{\mathrm{RS}}$ is the largest prime number such that $N_{\mathrm{ZC}}^{\mathrm{RS}}<M_{\mathrm{sc}}^{\mathrm{RS}}$ and represents the length of $\mathbf{x}_q$, which can be calculated with

$$\mathbf{x}_q(m)=\exp\left(-j\frac{\pi qm(m+1)}{N_{\mathrm{ZC}}^{\mathrm{RS}}}\right),\quad 0\le m\le N_{\mathrm{ZC}}^{\mathrm{RS}}-1 \tag{5}$$

where q is given by

$$q=\lfloor\hat{q}+1/2\rfloor+\hat{v}\cdot(-1)^{\lfloor 2\hat{q}\rfloor}. \tag{6}$$

where $\hat{q}$ can be calculated with

$$\hat{q}=N_{\mathrm{ZC}}^{\mathrm{RS}}\cdot(\hat{u}+1)/31, \tag{7}$$

In this paper we let $\hat{p}=0$ to simplify the SRS sequence generation. The number of SRS sequences adopted is preferred to be a divisor of $n_{\mathrm{SRS}}^{\mathrm{cs,max}}$, and the phase of adjacent sequences are usually equally distributed.

B. THE SUPERIMPOSED RECEIVED SIGNALS

We consider the number of antenna of BS is N_{T}, assume the length of comb-type arranged SRS and the bandwidth to be estimated are M subcarriers. Let K denote the number of MUs. The received signals of BS can be given with

$$\mathbf{y}=\sum\nolimits_{k=1}^{K}\mathbf{h}_k\cdot\mathbf{S}_k+\mathbf{n}, \tag{8}$$

where both $\mathbf{y}$, $\mathbf{h}_k$ and $\mathbf{n}$ are in $M\times 1$ dimension and represent received signals, channel transform vector of user k and additive Gaussian noise whose variance is σ_{n}^2, respectively, and pilot matrix $\mathbf{S}_k\in\mathbb{C}^{M\times M}$ whose diagonal elements are pilot points standing for the SRS sequence for user k.

IV. M-DnCNN CHANNEL ESTIMATION

To enhance the channel estimation capacity on narrow band, in this section we provide the M-DnCNN CE.

A. OVERVIEW OF M-DnCNN CE

In classical DnCNN based image denoising, DnCNN can estimate element-level noise for one image, which is helpful to combine DnCNN with channel estimation. However, in MU channel estimation, especially on narrow band, one difficulty is to separate the received signals into different wireless channel of users. It is not realistic to restore MU channels with each CFR individually. One natural characteristic can be utilized is that a segment of adjacent subcarriers are likely to have similar complex-numbers. Benefit from this property, we can regard those CFRs are with same complex-number values, and establish equation set to separate the wireless channels of different users. The pipeline of proposed M-DnCNN CE is given in Fig. 2. But for the ease of understanding, we will illustrate the method with the order of M-DnCNN denoising, channel restoration and SRS modification. Note that the training and testing of M-DnCNN is separated. We train the M-DnCNN offline and estimate users by testing online. This separation makes it possible to estimate MUs on the premise of satisfying 5G end-to-end latency. The network is fine-tuned periodically because the fine tuning can already to meet the requirements of accuracy and generalization errors.

B. M-DnCNN DENOISING

Similar to the conventional methods, our M-DnCNN denoising starts from the LS estimation, but the different part is that we adopt the same SRS sequence to implement LS estimation. More precisely, we assume the pilot matrix $\mathbf{S}_{k_0}^{\mathrm{H}}$ as a constant matrix and k_0 is an integer between 1 and K. By learning the noise on each subcarrier and antenna, M-DnCNN can help to output the superimposed channels which is consisted of channels of different users.

We assume the parameters of M-DnCNN is Θ. We denote the LS estimation of n^{th} received signals using pilot matrix $\mathbf{S}_{k_0}$ as $\mathbf{h}_{n,k_0}^{\mathrm{LS}}$, and its rearranged IFmap as $\hat{\mathbf{H}}_{n,k_0}^{\mathrm{LS}}$. Similarly, we use $\hat{\mathbf{n}}_{n,k_0}^{\mathrm{F}}$ to stand for the flattened vector of OFmap which is rearranged from $\mathbf{n}_n\cdot\mathbf{S}_{k_0}^{\mathrm{H}}$. Let M and N_{B} represent the number of subcarriers and the batchsize, respectively, and loss function of M-DnCNN can be described with

$$\mathbb{L}(\Theta)=\frac{1}{M\cdot N_{\mathrm{T}}\cdot N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\left\|f(\Theta;\beta\hat{\mathbf{H}}_{n,k_0}^{\mathrm{LS}})-\beta\hat{\mathbf{n}}_{n,k_0}^{\mathrm{F}}\right\|_2^2, \tag{9}$$

where $f(\Theta;\beta\hat{\mathbf{H}}_{n,k_0}^{\mathrm{LS}})$ represents the training function whose results are in vector form, and the β is an amplification factor introduced to compensate the amplitude defects of samples and facilitate the setting of Learning Rate (LR). For simplicity, we omit k_0 in the following parts.

The major steps of M-DnCNN training/testing include:

• Step 1: The pre-process of IFmaps/OFmaps. We think keeping the initial spatial-frequency dimension helps to learn the inner relationship and potential features. we separate the real/imaginary parts of $\mathbf{h}_n^{\mathrm{LS}}$ into different channel. and arrange the LS estimation on every antenna into the matrix $\hat{\mathbf{H}}_n^{\mathrm{LS}}$. More precisely, the dimension of $\hat{\mathbf{H}}_n^{\mathrm{LS}}$ is

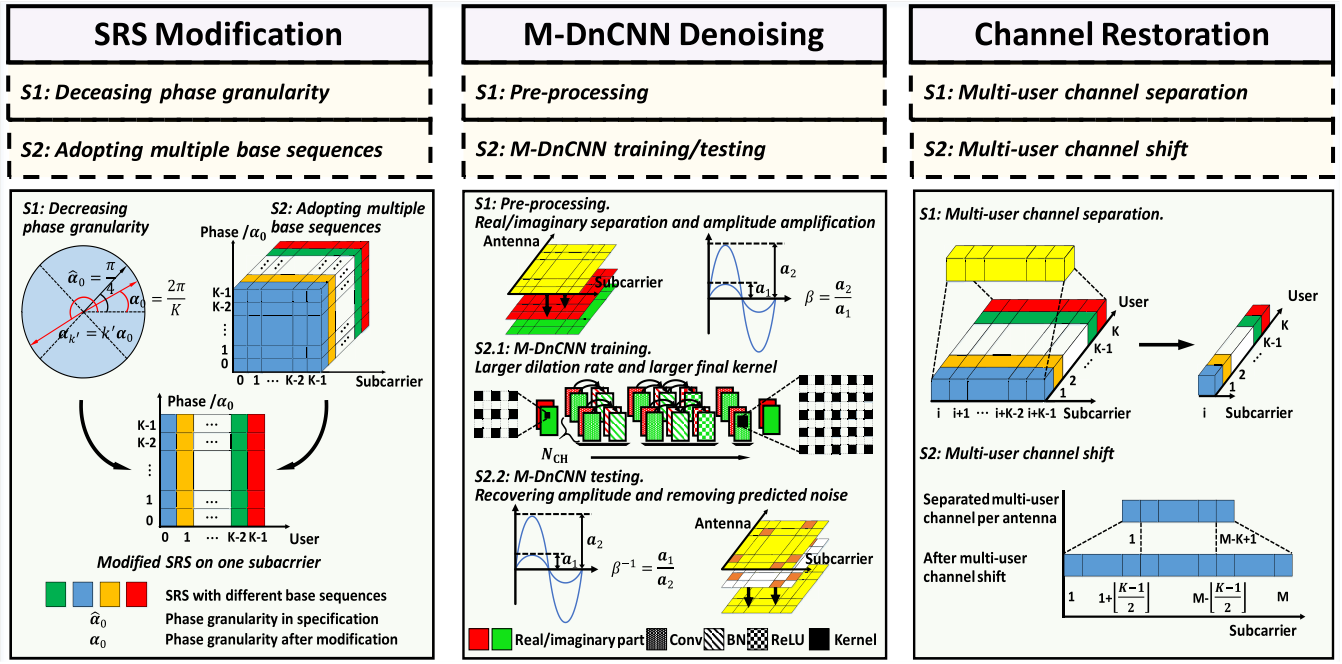


FIGURE 2. **The pipelines of M-DnCNN CE (from left to right).**

($N_{\mathrm{Tr}}, M, 2$), where N_{Tr} is the number of training antenna and the last dimension represents CH. Finally, multiplying amplification factor β, we obtain IFmaps/OFmaps.

• Step 2.1: The training of M-DnCNN. The structure of M-DnCNN is based on complex-number DnCNN, while some aspects are different: 1) We consider a larger dilation rate to enhance the receptive field. The utilization of dilation rate in convolution is inspired from semantic segmentation, and the dilation rate can be regarded as the inter-separation between the pixels of Fmaps. The reasons of adopting larger dilation rates are that the CFRs on narrow bands may have similar values, and a larger receptive field helps to learn the features from a larger area; 2) We use a larger kernel in the output layer for a better use of features extracted by middle layers. The deceasing LRs are adopted. LR means the amplitude that is used by the neurons to adjust their weights. If we adopt a very large LR, it exists high possibility that the weights cannot converge, and the results diverge from good solutions, while always adopting relatively small LR will cause a waste in training duration. From this sight, in order to prevent the oscillation near the minima, it is recommended to slow down the parameter updates by decreasing the LRs. For epoch n_{e}, the LR v_{LR} can be calculated with

$$v_{\mathrm{LR}} = \frac{v_0}{1 + n_{\mathrm{e}} \cdot a_{\mathrm{dr}}}, \tag{10}$$

where v_0 is the initial LR and a_{dr} denotes the decay rate.

• Step 2.2: The testing of M-DnCNN. The OFmaps of M-DnCNN are in vector form. Three operations are needed to obtain the denoised complex-number superimposed channels whose dimension is ($N_{\mathrm{T}}, M, 1$). First, we remove the amplification by dividing the OFmap with β. Then, the divided results are rearranged back into ($N_{\mathrm{T}}, M, 1$) dimension. Third, we minus the initial IFmap with the rearranged results.

Note that step 1 and 2.1 are for M-DnCNN training, while step 1 and 2.2 are for M-DnCNN testing.

C. CHANNEL RESTORATION

We denote the denoised spatial-frequency channels of n^{th} received signals as $\tilde{\mathbf{H}}_n$. The superimposed channel $\tilde{\mathbf{h}}_n^{(u_{\mathrm{T}})}$ on BS antenna u_{T}, which is the $u_{\mathrm{T}}^{\mathrm{th}}$ row of $\tilde{\mathbf{H}}_n$, can be given by

$$\tilde{\mathbf{h}}_n^{(u_{\mathrm{T}})} = \sum\nolimits_{k=1}^{K} \mathbf{h}_k^{(u_{\mathrm{T}})} \cdot \mathbf{S}_{k,k_0}^{\mathrm{M}} + \mathbf{e}^{(u_{\mathrm{T}})}, \tag{11}$$

where $\mathbf{e}^{(u_{\mathrm{T}})}$ is the denoising error whose MSE is equal to the loss, and $\mathbf{S}_{k,k_0}^{\mathrm{M}}$ represents the diagonal matrix, which can be calculated with the modified SRS sequence $\mathbf{S}_{k_0}^{\mathrm{M}}$ and $\mathbf{S}_k^{\mathrm{M}}$. Note that $\mathbf{S}_k^{\mathrm{M}}$ is in vector form. Given that all the operations below are the same for all BS antennas, thus we omit the superscript "(u_{T})" here for simplicity. Then for m^{th} subcarrier, we have

$$\tilde{\mathbf{h}}_n(m) = \mathbf{h}_{\mathrm{SI}}^m \cdot \mathbf{S}_m^{\mathrm{vec}} + \mathbf{e}(m), \tag{12}$$

where $\mathbf{h}_{\mathrm{SI}}^m = [h_1(m), h_2(m), \ldots, h_K(m)]$ stands for the MU channels on m^{th} subcarrier, and $\mathbf{S}_m^{\mathrm{vec}} = \left[s_1^{\mathrm{M}}(m)\bar{s}_{k_0}^{\mathrm{M}}(m), s_2^{\mathrm{M}}(m)\bar{s}_{k_0}^{\mathrm{M}}(m), \ldots, s_K^{\mathrm{M}}(m)\bar{s}_{k_0}^{\mathrm{M}}(m)\right]^{\mathrm{T}}$.

We roughly regard the CFRs on adjacent subcarriers are with the same value. For starting index m satisfying $m_K = m + K - 1 \le M$, we can establish equation set in matrix form with

$$\mathbf{h}_{\mathrm{SI}}^m \cdot \mathbf{S}_{(m,m_K)}^{\mathrm{Mat}} = \tilde{\mathbf{h}}_{(n,(m,m_K))}, \tag{13}$$

where

$$\mathbf{S}^{\mathrm{Mat}}_{(m,m_K)} = \left[\mathbf{S}^{\mathrm{vec}}_{m}, \mathbf{S}^{\mathrm{vec}}_{m+1}, \ldots, \mathbf{S}^{\mathrm{vec}}_{m_K}\right] = \begin{bmatrix} s^{\mathrm{M}}_1(m)\bar{s}^{\mathrm{M}}_{k_0}(m) & \cdots & s^{\mathrm{M}}_1(m_K)\bar{s}^{\mathrm{M}}_{k_0}(m_K) \\ \vdots & \ddots & \vdots \\ s^{\mathrm{M}}_K(m)\bar{s}^{\mathrm{M}}_{k_0}(m) & \cdots & s^{\mathrm{M}}_K(m_K)\bar{s}^{\mathrm{M}}_{k_0}(m_K) \end{bmatrix}, \quad (14)$$

and

$$\tilde{\mathbf{h}}_{(n,(m,m_K))} = [\tilde{\mathbf{h}}_n(m) + \mathbf{e}(m), \tilde{\mathbf{h}}_n(m+1) + \mathbf{e}(m+1), \ldots, \tilde{\mathbf{h}}_n(m_K) + \mathbf{e}(m_K)]. \quad (15)$$

Besides, if the denoising error is small, we can neglect $\mathbf{e}(m)$. Thus we can separate the channels for MUs with

$$\mathbf{h}^m_{\mathrm{SI}} = \tilde{\mathbf{h}}_{(n,(m,m_K))} \cdot \left(\mathbf{S}^{\mathrm{Mat}}_{(m,m_K)}\right)^{-1}. \quad (16)$$

Given that practical adjacent CFRs may be slightly different, another operation is done to lower MSE and the channel for user k' is given by

$$\tilde{\mathbf{h}}_{k'}(m) = \begin{cases} \mathbf{h}^1_{\mathrm{SI}}(k'), & m \le 1 + \lfloor \frac{K-1}{2} \rfloor \\ \mathbf{h}^K_{\mathrm{SI}}(k'), & m \ge K - \lfloor \frac{K-1}{2} \rfloor \\ \mathbf{h}^m_{\mathrm{SI}}(k'), & \text{otherwise} \end{cases} \quad (17)$$

Note that the sufficient and necessary condition of (16) is that $\mathbf{S}^{\mathrm{Mat}}_{(m,m_K)}$ must be invertible. However, the configurations in specifications cannot fulfill the non-singularity of $\mathbf{S}^{\mathrm{Mat}}_{(m,m_K)}$ when K is larger than $n^{\mathrm{cs,max}}_{\mathrm{SRS}}$. To solve this problem, SRS must be modified to match the proposed M-DnCNN CE.

D. SRS MODIFICATION

The current configuration in specification fails to support the case when $K > N^{\mathrm{cs,max}}_{\mathrm{SRS}}$. Besides, if we only use the SRS sequences obtained from the same base sequence by shifting the phases, it may sacrifice the latent features due to the inner-relationship between different sequences. Given that different base sequences are linearly independent and the phase granularity is limited, we propose to modify SRS as follows.

• When $K > N^{\mathrm{cs,max}}_{\mathrm{SRS}}$, we let $N^{\mathrm{cs,max}}_{\mathrm{SRS}} = K$. It is not be hard to find that the maximum estimated MU can be expanded to M. This can be seen the operation of decreasing phase granularity.

• We adopt different of base sequence to form a new SRS sequence. More precisely, here we only adopt different $\hat{u}$ while we set $\hat{v} = 0$ for simplicity. For user k', we first obtain his initial sequence of K points with

$$\mathbf{S}^{\mathrm{Ini}}_{k'} = \left[s^{\mathrm{M}}_{u_1,k'}(0), s^{\mathrm{M}}_{u_2,k'}(1), \ldots, s^{\mathrm{M}}_{u_K,k'}(K-1)\right], \quad (18)$$

where for $0 \le m < K$, $s^{\mathrm{M}}_{k'}(m)$ can be given by

$$s^{\mathrm{M}}_{u,k'}(m) = \mathbf{r}^{0,k'}_{u,0}(m) = \exp(j2\pi n\frac{k'}{K}) \cdot \hat{\mathbf{r}}_{u,0}(m), \quad (19)$$

where $0 \le n < K$ and u stands for the group number for user k'. Note that for any two users k' and k'', their group number $\hat{u}_{k'}$ and $\hat{u}_{k''}$ are unequal. Finally, we obtain $\mathbf{S}^{\mathrm{M}}_{k'}$ with

$$\mathbf{S}^{\mathrm{M}}_{k'} = \mathbf{S}^{\mathrm{Ini}}_{k'} \cdot \Sigma_{\mathrm{rep}}, \quad (20)$$

where $\Sigma_{\mathrm{rep}} \in \mathbb{C}^{K \times M}$ stands for the matrix which aims to repeat the initial sequence $n_{\mathrm{r}} = M/K$ times, and Σ_{rep} is given by $\Sigma_{\mathrm{rep}} = \left[\mathbf{I}^{(1)}_K, \mathbf{I}^2_K, \ldots, \mathbf{I}^{(n_{\mathrm{r}})}_K\right]$.

It can be easily found that the modifications above will not change the property of constant amplitude and zero-autocorrelation. Besides, we can also find that $\mathbf{S}^{\mathrm{Mat}}_{(m,m_K)}$ is invertible. This is due to the reason that different base sequences are linearly independent, and the matrix formed by a set of linearly independent vectors must be with full rank. Though using only one base sequence in (18) can also meet the requirements for the non-singularity of $\mathbf{S}^{\mathrm{Mat}}_{(m,m_K)}$, letting adjacent subcarriers adopting different pilot points generated with various group can dig the latent feature better.

Since it is a preliminary exploration in this paper, the aforementioned equations only show the simple case where K is a divisor of M. However, for the case where K is not a divisor of M, the solution can be given in the similar way. We can let the subcarriers whose number is equal to the maximum multiple of K adopt the above operations to enhance the CEC capacity. For the rest of subcarriers, if the steam splitting scenarios are considered, we can use reminder of subccariers to estimate the channel of high-rate users. For example, when $K = 11$ and $M = 24$, two subcarriers are left. We can use these two subcarriers to estimate two of users who require better performance. The design of this case can also comprehensively consider other prior information such as LoS/NLoS conditions. Note that the utilization of the rest of subcarriers may need to train other M-DnCNN models.

E. COMPLEXITY ANALYSIS

We utilize FLoating-point Operations Per Second (FLOPS) to evaluate the time complexity. For the fairness during the comparison, we consider the time complexity of estimating one sample, which means that we consider the time complexity of estimating the channels on N_{Tr} antennas for K MUs.

The time complexity of M-DnCNN CE consists of two parts: the time complexity of channel restoration and M-DnCNN denoising that includes Conv and BN. The time complexity of ReLU are neglected. In [35], the detailed algorithm of BN is provided. The procedures includes calculating mini-batch mean, mini-batch variance, normalize and scale and shift. The time complexity of BN T_{BN} can be given by

$$T_{\mathrm{BN}} \le \mathcal{O}\left((N_{\mathrm{IL}} + N_{\mathrm{ML}})MN_{\mathrm{Tr}}\right). \quad (21)$$

We use less-than-equal symbol in (21) because the operations of calculating mean and variance are shared. Inspired by [36], [37], the time complexity of Conv can be described with

$$T_{\mathrm{Conv}} = \mathcal{O}\left(MN_{\mathrm{Tr}}N^2_{\mathrm{CH}} \sum_{n=1}^{N_{\mathrm{IL}}+N_{\mathrm{ML}}+N_{\mathrm{OL}}} L^2_{\mathrm{Ker}}(n)\right), \quad (22)$$

TABLE 1. Time complexity of four MU channel estimation methods.

Method	FLOPS
CS	$\mathcal{O}\left(KMN_{\mathrm{Tr}}\right)$
C-DFT	$\mathcal{O}\left(N_{\mathrm{T}}(2MK\log M+M)\right)$
E-DFT	$\mathcal{O}\left(N_{\mathrm{T}}(2MK\log M+3M)\right)$
M-DnCNN CE	$\mathcal{O}\left(MN_{\mathrm{Tr}}N_{\mathrm{CH}}^2\sum_{n=1}^{N_{\mathrm{IL}}+N_{\mathrm{ML}}+N_{\mathrm{OL}}}L_{\mathrm{Ker}}^2(n)\right)$

where $L_{\mathrm{Ker}}(n)$ stand for the length of kernel of layer n. Then the time complexity of M-DnCNN denoising is given by

$$T_{\text{M-DnCNN}}=T_{\mathrm{Conv}}+T_{\mathrm{BN}}, \tag{23}$$

and it is not hard to find that $T_{\mathrm{BN}}\ll T_{\mathrm{Conv}}$, thus we have

$$T_{\text{M-DnCNN}}\approx\mathcal{O}\left(MN_{\mathrm{Tr}}N_{\mathrm{CH}}^2\sum_{n=1}^{N_{\mathrm{IL}}+N_{\mathrm{ML}}+N_{\mathrm{OL}}}L_{\mathrm{Ker}}^2(n)\right). \tag{24}$$

As the results of inversion for $\left(\mathbf{S}_{(m,m_K)}^{\mathrm{Mat}}\right)^{-1}$ are constant, they can be pre-calculated, thus we do not consider this part of time complexity. Then the time complexity of channel restoration in (16) and (17) can be calculated with

$$\begin{aligned}T_{\mathrm{CR}}&=\mathcal{O}\left(K^2(M-K+1)N_{\mathrm{Tr}}+KMN_{\mathrm{Tr}}\right)\\&\approx\mathcal{O}\left(K^2(M-K+1)N_{\mathrm{Tr}}\right).\end{aligned} \tag{25}$$

Given that T_{CR} is much smaller than $T_{\text{M-DnCNN}}$, we use $T_{\text{M-DnCNN}}$ to approximately calculate the time complexity of M-DnCNN CE. Based on this result, the typical complexities calculating structure matrix presented in [38] and time complexity in [28] and [39], the time complexity of four MU channel estimation approaches are provided in Table 1.

Though it seems that the time complexity of M-DnCNN CE is much larger than three conventional MU channel estimation methods, M-DnCNN operates on GPU and deep learning based platform and the operating time is acceptable. The corresponding operation information will be provided in section IV-C.

V. SIMULATION RESULTS

A. SYSTEM CONFIGURATION

The experiments are carried out with Keras whose backend is Theano, which runs on a PC with Intel (R) Xeon (R) E5-2620 v2 CPU (2.10GHz) and a Nvidia Titan X GPU (12G). The system parameters are provided in Table 1.

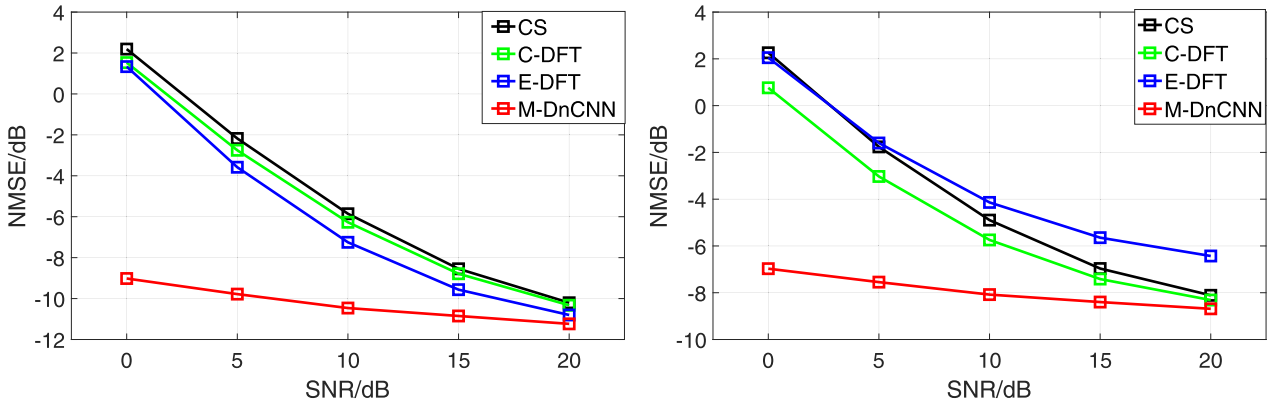
The generation of training and testing datasets includes two steps. First, we drop the users using 3-D channel model specified by 3GPP to generate a large amount of wireless channels, which we call as original datasets. Then we randomly combine K users to obtain training samples and add the noise according to the total received signal and the SNR. The second step is repeated until the number of samples is the same as configured size of training dataset. We take an example for the generation of training dataset when SNR is 0 dB. When K is 12, the number of single-user wireless channels in original dataset is 12000. By randomly combining non-repetitive K users we can obtain 1000 received signals. Then we add the noise according to 0 dB SNR and obtain 1000 training samples. The operations above are repeated for 50 times to obtain the training dataset that contains 50000 samples. As for testing dataset, we only need to repeat the aforementioned two steps for 2 times.

TABLE 2. Simulation parameters.

Parameter	Assumption
Channel model	3D-UMa
The number of BS antenna N_{T}	256
The number of training antenna N_{Tr}	16
The number of user antenna	1
The inter-antenna separation	$\lambda/2$
The number of subcarrier M	24
OFDM symbol	1
User speed	3 Km/h
The direction of user mobility	Horizontal forward
2-D distance	[35, 150]m
The distribution of user	Uniformly distributed
The number of multi-user K	8, 12
SNR γ	0,5,10,15,20 dB
The size of original dataset	$K=8$:8000, $K=12$:12000
The size of training dataset	50000
The size of testing dataset	2000
The number of input layer N_{IL}	1
The number of middle layerN_{ML}	9
The number of output laterN_{OL}	1
The number of CH (N_{CH})	64
Dilation rate	2
Kernel size	Final kernel: 5×5, others: 3×3
Initial LR v_0	0.1
Decay rate a_{dr}	0.05
Amplification factor β	100
Batch size N_{B}	100
The number of epoch	150

We emphasize the following four points.

1) Since 3-D channel model fails to support the updating of user location, we generate the original datasets by dropping many times. More precisely, we drop 100 users each time, and then we drop 80/120 times to obtain 8000/12000 samples for original dataset. In our preliminary exploration, we consider a relatively stationary case, which means that the mobility of vehicles and pedestrians will not impact channels too much. In this way, the channels in original datasets can nearly meet the generalization of training datasets.
2) As the selection of MUs in MU channel estimation is based on resource allocation schemes in MGMCa, for any one user, he can be estimated with any other users together. Thus the combination of channels in original datasets is closer to practical cases.
3) The addition of noise is from the perspective of BS and the superimposed channels of MUs are desired signals. Besides, the Gaussian noise is added independently on each subcarrier according to the noise power. Though we want the SNR to be a single value for a training/testing dataset, the SNR of the samples is dispersive. This characteristic is good for practicability and we do not bother to train on multiple SNRs.



(a) $K = 8$ (b) $K = 12$

FIGURE 3. The NMSE performance of four methods. The CFR shift operation of (17) is done for CS and $n_{\mathrm{SRS}}^{\mathrm{cs,max}} = K$. The significant tap number $M^{\mathrm{s}} = 1$. k_1, k_2, γ, β and α in [36] are configured as 0.75, 0.6, 0, 0.25 and 1.25, respectively.

4) The number of BS antenna is configured as 256 when generating the original samples. We choose 16 adjacent antennas from 256 total antennas to establish IFmap and OFmaps. More precisely, we select antennas whose indexes are $1-16$ for simplicity. As long as the selected antennas are adjacent, the indexes will not impact the performance.

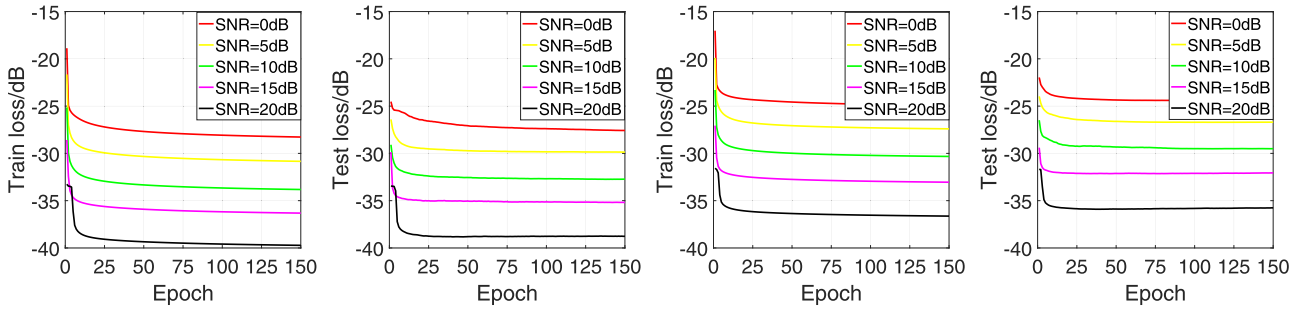
B. PERFORMANCE ANALYSIS

The NMSE performance of four channel estimation approaches are provided in Fig. 3, and the results of two MU configurations are given in subfigures Fig. 3-(a) and Fig. 3-(b). The estimation results are collected from every antenna and user. More precisely, each point in Fig. 3 is the mean of $2000\times16\times8(12) = 192000(288000)$ NMSE results. The simulation parameters of comparison methods are given in the caption of Fig. 3. The configurations are based on our simulation experience and we have empirically determined that it produces good solutions. Note that $\mu_3 = 0$ means the existence test for taps in E-DFT are canceled, because the M/K is too small to have the taps containing pure noise.

From Fig. 3, E-DFT shows a better performance than C-DFT when $K = 8$, since it can distinguish the LoS and NLoS users and provide a better TD denoising for LoS users. When K is increased, both the NMSE results of C-DFT and E-DFT degrade, but the performance of E-DFT degrades more severely. This is mainly due to the reason that its LoS and NLoS identification becomes more inaccurate and the mismatch between the denoising method and LoS/NLoS conditions may zero the desired taps. CS method shows the worst NMSE results when $K = 8$, because it merely separates the MU channels based on the LS estimation results while the denoising ability of LS estimation is limited. The good part of CS method is the robustness when SNR is small, and a larger K only impacts the NMSE performance when noise is small. This degradation is due to the fact that the CFRs of the adjacent subcarriers are severely unequal for some NLoS users, and the averaging between more subcarriers will aggravate this problem.

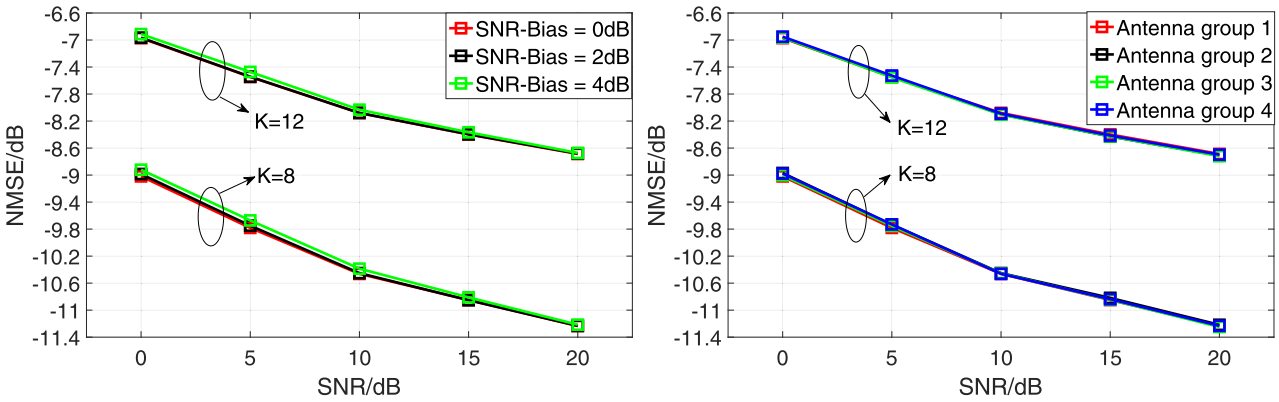
Compared with three classical methods, the proposed M-DnCNN CE realizes a great improvement for both $K = 8$ and $K = 12$ cases, and when SNR is 0 dB the performance gaps increase to 10 dB and 8 dB, respectively. This enhancement comes from the good denoising ability of M-DnCNN. Though the mathematical description of received signals for MU channel estimation is a set of wireless channels multiplied by different SRS sequences, the received signals can be regarded as the wireless channels for one ''virtual'' single user multiplied by one of SRS sequences. Actually, this idea has already been embodied in the preprocessing step of M-DnCNN where a fixed SRS sequence is adopted. Then the samples in training datasets can be seen as a huge amount wireless channels of virtual single users distributed in the cell. Given that we consider a relatively stationary case, the users whose locations are near may have similar wireless channels. M-DnCNN is an end-to-end network that can automatically find the users that may have similar channels. Benefit from big amount of data and the zero-mean property of noise, the M-DnCNN can predict the noise on each subcarrier. This utilization of similar channels can also be explained with theory of Wiener filter based denoising. The similar channels from different neighbor users just like the channels of one user collected from other OFDM symbols. The M-DnCNN can multiply those channels with different weights to restore the desired channels. Though M-DnCNN still faces the same problem as CS, its denoising ability achieves the great improvement. Please note that the assumption of stationary is used to describe the degree of stability in spatial domain rather than time domain.

According to the simulation results in Fig. 3, the NMSE of conventional methods is too large when SNR is small, thus the estimation accuracy can hardly support the transmission afterwards. Actually, if the comb-type arrangement and classical methods are adopted, the number of MU is usually



(a) The train loss when $K = 8$ (b) The test loss when $K = 8$ (c) The train loss when $K = 12$ (d) The test loss when $K = 12$

FIGURE 4. **The loss curves for different SNR and number of MU.**



(a) The NMSE performance considering different SNR bias (b) The NMSE performance for different antenna groups

FIGURE 5. **The NMSE performance considering different SNR bias or antenna groups.**

set no larger than 4 users. Just as depicted in Fig. 3, three comparison methods fail to estimate MUs that is not less than 8 users. M-DnCNN CE can guarantee that NMSE is always smaller than -6.5 dB, thus we can say that M-DnCNN CE achieves the CEC enhancement and the CEC is tripled.

The loss curves for different SNR and number of MU are depicted in Fig. 4. The loss given in (9) can be regarded the mean square errors between predicted results and noise-free LS-estimated results. From Fig. 4, M-DnCNN converge with no over-fitting, thus the configurations of epoch and LR are suitable. We emphasize that the degradation in Fig. 4-(c) and Fig. 4-(d) can be alleviated by increasing the depth of M-DnCNN. However, by comprehensively considering the time complexity, we think the current setting of M-DnCNN is good enough to meet the requirements, thus it does not need to make the network deeper or wider.

C. FEASIBILITY DISCUSSION

In this subsection, we will firstly discuss the performance considering the mismatch of SNR. In practice, the SNR of the samples to be estimated may not match the trained model, i.e., the SNR of sample may be 2 dB but the model is trained with the dataset whose SNR is 0 dB. Then we will discuss the performance considering the mismatch of different antennas. In our configurations, we select 1-16 antennas to train the models for simplification. However, for massive MIMO systems, the amount of antennas equipped by BS is larger than 16. Here we will evaluate whether the model trained with 1-16 antennas is suitable for the denoising of other antennas, i.e., 17-32 antennas and so on. Finally, to prove that the proposed M-DnCNN can meet the practical latency constraint, we provide the relevant operation time.

The NMSE performance considering different SNR bias or antenna group is provided in Fig. 5. The definition of SNR bias is that the range of the difference between the SNR of samples and that of model, and the difference is assumed to be uniformly distributed. For example, when SNR bias is 2 dB, for any model that is trained with the dataset whose SNR is γ, the SNRs of samples range for $[\gamma - 1, \gamma + 1]$ dB. For the ease of illustrations, we use antenna group to describe different antennas. For any antenna group n_{AG}, it stands for the samples which is generated with the CFRs on the antennas whose indexes range for $[n_{\mathrm{AG}} \times 16 - 15, n_{\mathrm{AG}} \times 16]$. Please note that SNR-bias 0 dB and antenna group 1 is the evaluation configuration of M-DnCNN in Fig. 3, and the SNR bias and antenna group are individually discussed.

TABLE 3. Operation information : T_E is offline training time per epoch; T_{CE} is online testing time per sample.

	Unit	T_E/ s	T_{CE}/ ms
$K = 8$	673664	106	0.536
$K = 12$	673664	108	0.542

According to Fig. 5-(a), the SNR-bias has little impact on NMSE performance. This is mainly due to the fact that the SNRs of training samples are subjected to a distribution rather than a single value. Though the existence of SNR-bias will slightly increase the range of the SNR distribution of samples, the models can still cover most of cases. From Fig. 5-(b), there is a negligible performance difference among the cases of different antenna groups. This is because that the TD channels of different antenna groups have similar features, and our training strategy that using CFRs on antenna group 1 to train the model has already caught sufficient latent features to predict high-precision results.

The operation information for the M-DnCNN training and testing are provided in Table 3. From Table 3, the utilization of GPU makes it possible to finish online M-DnCNN CE on the permise of satisfying practical 1 ms constraint. Though the number of MU are different, the network structure and the number of neurons are the same in M-DnCNN. Therefore, the operation time are similar. We emphasize that the training may be time-consuming, but fine-tuning will much faster because of less epochs. Besides, the operation time is further decreased with the development of equipment in the future.

However, we have to frankly state the underlying challenges. The collection of samples are based on 3-D channel model. Though it is supported by 3GPP, there still exists a gap between the generated and real-world channels. We think that training and testing datasets should include the samples collected in practice. This can help to fit our method with the cases of real world. However, the aim of this paper is to provide a new way to enhance the CEC capacity, and good feasibility has been shown for the simplified case. The results of Fig. 5 shows that the proposed M-DnCNN CE is robust even when the samples are slightly polluted, and using the feature maps generated with the CFRs on 16 antennas is already capable of M-DnCNN based massive MIMO denoising. In addition, the operation time T_{CE} is less than 1 ms, which satisfies practical latency constraints, thus the proposed M-DnCNN CE shows good feasibility.

VI. CONCLUSION

Facing the insufficient channel information caused by limited SRS resources that is a crucial problem faced by MGMCa, in this paper we propose M-DnCNN CE to enhance CEC for narrow band. Three parts are included in the M-DnCNN CE, which is SRS modification, M-DnCNN denoising and channel restoration. We validate the feasibility of M-DnCNN CE with 3D-MIMO channel model specified by 3GPP. The results show that M-DnCNN CE realizes the great improvement of NMSE performance especially when noise is large compared to three conventional methods, and the CEC is even tripled compared to the configurations of specification. The operation time, SNR-bias and different antenna groups are also discussed. We find that the configurations is capable of estimating massive antennas. Besides, M-DnCNN satisfies practical latency constraint and shows good robustness.

However, the design methodology is still in its infancy and of exploratory nature, there are still some aspects remained to be studied. First, does there exist the optimal configuration for M-DnCNN that can realize the near-accurately denoising? Second, given that the variations of adjacent CFRs between LoS and NLoS users are different, can we use LoS/NLoS conditions to clean the training datasets and further enhance the MU channel estimation performance? Third, in proposed SRS modification, more base sequences are utilized for one BS, thus a system-level pilot allocation scheme should be investigated to avoid pilot contamination. Fourth, the proposal of M-DnCNN CE is inspired by the problem of insufficient CSI in MGMCa, but M-DnCNN based CEC enhancement should not be limited to this case. Then how to improve M-DnCNN to lend itself to various scenarios? The aforementioned problems are interesting topics, and we will resolve them in the future work.

ACKNOWLEDGMENT

We gratefully acknowledge reviewers who made helpful suggestions.

REFERENCES

[1] D. Wu, Z. Zhang, S. Wu, J. Yang, and R. Wang, "Biologically inspired resource allocation for network slices in 5G-enabled Internet of Things," *IEEE Internet Things J.*, to be published, doi: 10.1109/JIOT.2018.2888543.
[2] J. Dai, K. Niu, and J. Lin, "Iterative Gaussian-approximated message passing receiver for MIMO-SCMA system," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 3, pp. 753–765, Jun. 2019.
[3] Z. Zhang and L. Wang, "Social tie-driven content priority scheme for D2D communications," *Inf. Sci.*, vol. 480, pp. 160–173, Apr. 2019.
[4] Z. Li, J. Chen, and Z. Zhang, "Socially aware caching in D2D enabled fog radio access networks," *IEEE Access*, vol. 7, pp. 84293–84303, 2019.
[5] P. Zhang, X. Kang, X. Li, Y. Liu, D. Wu, and R. Wang, "Overlapping community deep exploring based relay selection method towards multi-hop D2D communication," *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1357–1360, Oct. 2019.
[6] D. Wu, B. Liu, Q. Yang, and R. Wang, "Social-aware cooperative caching mechanism in mobile social networks," *J. Netw. Comput. Appl.*, vol. 149, Jan. 2020, Art. no. 102457, doi: 10.1016/j.jnca.2019.102457.
[7] D. Wu, Q. Liu, H. Wang, Q. Yang, and R. Wang, "Cache less for more: Exploiting cooperative video caching and delivery in D2D communications," *IEEE Trans. Multimedia*, vol. 21, no. 7, pp. 1788–1798, Jul. 2019.
[8] Z. Li, Y. Jiang, Y. Gao, D. Yang, and L. Sang, "On buffer-constrained throughput of a wireless-powered communication system," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 2, pp. 283–297, Feb. 2019.
[9] P. Zhang, X. Kang, D. Wu, and R. Wang, "High-accuracy entity state prediction method based on deep belief network toward IoT search," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 492–495, Apr. 2019.
[10] D. Wu, H. Shi, H. Wang, R. Wang, and H. Fang, "A feature-based learning system for Internet of Things applications," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1928–1937, Apr. 2019.
[11] Z. Li, H. Liu, and R. Wang, "Service benefit aware multi-task assignment strategy for mobile crowd sensing," *Sensors*, vol. 19, p. 4666, Oct. 2019.
[12] Y. C. B. Silva and A. Klein, "Linear transmit beamforming techniques for the multigroup multicast scenario," *IEEE Trans. Veh. Technol.*, vol. 58, no. 8, pp. 4353–4367, Oct. 2009.
[13] M. Alodeh, D. Spano, A. Kalantari, C. Tsinos, D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Symbol-level and multicast precoding for multiuser multiantenna downlink: A state-of-the-art, classification and challenges," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 2, pp. 295–308, 2nd Quart., 2018.

[14] R. O. Afolabi, A. Dadlani, and K. Kim, "Multicast scheduling and resource allocation algorithms for OFDMA-based systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 240–254, 1st Quart., 2013.
[15] M. Sadeghi and C. Yuen, "Multi-cell multi-group massive MIMO multicasting: An asymptotic analysis," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)* San Diego, CA, Dec. 2015, pp. 1–6.
[16] Z. Xiang, M. Tao, and X. Wang, "Coordinated multicast beamforming in multicell networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 12–21, Jan. 2013.
[17] G.-W. Hsu, B. Liu, H.-H. Wang, and H.-J. Su, "Joint beamforming for multicell multigroup multicast with per-cell power constraints," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 4044–4058, May 2017.
[18] G. Dartmann, X. Gong, and G. Ascheid, "Cooperative beamforming with multiple base station assignment based on correlation knowledge," in *Proc. VTC*, 2010, pp. 1–5.
[19] Z. Xiang, M. Tao, and X. Wang, "Massive MIMO multicasting in noncooperative cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1180–1193, Jun. 2014.
[20] O. Tervo, L.-N. Tran, H. Pennanen, S. Chatzinotas, M. Juntti, and B. Ottersten, "Energy-efficient coordinated multi-cell multi-group multicast beamforming with antenna selection," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2017, pp. 1209–1214.
[21] D. Senaratne and C. Tellambura, "Beamforming for physical layer multicasting," in *Proc. IEEE WCNC*, Mar. 2011, pp. 176–1781.
[22] H. Joudeh and B. Clerckx, "Sum-rate maximization for linearly precoded downlink multiuser MISO systems with partial CSIT: A rate-splitting approach," *IEEE Trans. Commun.*, vol. 64, no. 11, pp. 4847–4861, Nov. 2016.
[23] A. Z. Yalcin, M. Yuksel, and I. Bahceci, "Downlink MU-MIMO with QoS aware transmission: Precoder design and performance analysis," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 969–982, Feb. 2019.
[24] J. Montalban, P. Scopelliti, M. Fadda, E. Iradier, C. Desogus, P. Angueira, M. Murroni, and G. Araniti, "Multimedia multicast services in 5G networks: Subgrouping and non-orthogonal multiple access techniques," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 96–103, Mar. 2018.
[25] M. Sadeghi, L. Sanguinetti, R. Couillet, and C. Yuen, "Reducing the computational complexity of multicasting in large-scale antenna systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2963–2975, May 2017.
[26] T. Zeng, Y. Chang, M. Hu, and Y. Zhang, "CSI-RS based joint grouping and scheduling scheme with limited SRS resources," in *Proc. PIMRC*, Sep. 2018, pp. 1–6.
[27] H. Yang, T. L. Marzetta, and A. Ashikhmin, "Multicast performance of large-scale antenna systems," in *Proc. IEEE Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2013, pp. 604–608.
[28] M. Xiang, Y. Chang, and T. Zeng, "Channel estimation for 3D MIMO system based on LOS/NLOS identifcation," *IET Commun.*, vol. 13, no. 7, pp. 898–904, 2019.
[29] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
[30] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
[31] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 630–645.
[32] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
[33] C. Trabelsi, O. Bilaniuk, Y. Zhang, D. Serdyuk, S. Subramanian, J. F. Santos, S. Mehri, N. Rostamzadeh, Y. Bengio, and C. J. Pal, "Deep complex networks," 2017, *arXiv:1705.09792*. [Online]. Available: https://arxiv.org/abs/1705.09792
[34] *Study 3D Channel Model for LTE (Release 12)*, document 36.873, V12.3.0, Third-Generation Partnership Project, 3GPP TS, Jan. 2018.
[35] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.
[36] T. Zeng, Y. Chang, Q. Zhang, M. Hu, and J. Li, "CNN based LOS/NLOS identification in 3D massive MIMO systems," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2491–2494, Dec. 2018.
[37] K. He and J. Sun, "Convolutional neural networks at constrained time cost," *CoRR*, vol. abs/1412.1710, pp. 1–8, Dec. 2014.
[38] I. Gohberg and V. Olshevksy, "Complexity of multiplication with vectors for structured matrices," *Linear Algebra Appl.*, vol. 202, pp. 163–192, Apr. 1994.
[39] X. Xiong, B. Jiang, X. Gao, and X. You, "DFT-based channel estimator for OFDM systems with leakage estimation," *IEEE Commun. Lett.*, vol. 17, no. 8, pp. 1592–1595, Aug. 2013.
[40] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels Modulation (Release 15)*, document 3GPP TS 36.211, V15.7.0, Third-Generation Partnership Project, Sep. 2019.

TIANYI ZENG received the B.S. degree in communication engineering from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2015. He is currently pursuing the Ph.D. degree in information and communication engineering with the Wireless Theories and Technologies (WT&T) Laboratory. His research interests include resource allocation, channel estimation, and deep learning-based signal processing for wireless communication systems. He currently focuses on resolving the problem of limited SRS resources in massive MIMO systems.

YAFENG WANG received the B.Sc. degree from Baoji University, in 1997, the M.Eng. degree from the University of Electronic Science and Technology of China, in 2000, and the Ph.D. degree from the Beijing University of Posts and Telecommunications, in 2003. In 2008, he was a Visiting Scholar with the Faculty of Engineering and Surveying, University of Southern Queensland, Australia.

He is currently a Professor of electronic engineering with the School of Information and Telecommunications, Beijing University of Posts and Telecommunications. He leads the Broadband Mobile Communication Engineering Laboratory, which is one of Zhongguancun Science Park Open Laboratories. He has authored or coauthored more than 100 peer-reviewed journal and conference papers. His research mainly focuses on wireless communications and information theory.

JUNYAO LI received the B.S. degree in communication engineering from the Beijing University of Posts and Telecommunications(BUPT), Beijing, China, in 2017, where he is currently pursuing the master's degree in information and communication engineering with the Wireless Theories and Technologies Laboratory. His two articles have been published at the WCNC Conference and the *Journal of Beijing University of Posts and Telecommunications*, respectively. His current research interests include LOS/NLOS identification, massive MIMO, channel estimation, and reference signal pattern design.

SHUAI HOU received the B.S. degree in communication engineering from Hebei University, Baoding, China, in 2016. He is currently pursuing the Ph.D. degree with the Wireless Theories and Technologies Laboratory, Beijing University of Posts and Telecommunications, Beijing, China. His research interests include massive MIMO, channel estimation, and sparse signal processing.