Manchester Metropolitan University

e-space

Manchester Metropolitan University's Research Repository

Please cite the Published Version

Wang, Jie, Gui, Guan, Ohtsuki, Tomoaki, Adebisi, Bamidele, Gacanin, Haris and Sari, Hikmet (2021) Compressive Sampled CSI Feedback Method Based on Deep Learning for FDD Massive MIMO Systems. IEEE Transactions on Communications, 69 (9). pp. 5873-5885. ISSN 0090-6778

DOI: https://doi.org/10.1109/TCOMM.2021.3086525

Publisher: IEEE

Version: Accepted Version

Downloaded from: https://e-space.mmu.ac.uk/633748/

Usage rights: © In Copyright

Additional Information: © 2021 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Enquiries:

If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines)

Compressive Sampled CSI Feedback Method Based on Deep Learning for FDD Massive MIMO Systems

Jie Wang, *Graduate Student Member, IEEE,* Guan Gui, *Senior Member, IEEE,* Tomoaki Ohtsuki, *Senior Member, IEEE*, Bamidele Adebisi, *Senior Member, IEEE*, Haris Gacanin, *Fellow, IEEE*, and Hikmet Sari, *Life Fellow, IEEE*

***Abstract*—Accurate downlink channel state information (CSI) is required to be fed back to the base station (BS) in frequency division duplexing (FDD) massive multiple-input multiple-output (MIMO) systems in order to achieve maximum antenna diversity and multiplexing. However, downlink CSI feedback overhead scales with the number of transceiver antennas, a major hurdle for practical deployment of FDD massive MIMO systems. To solve this problem, we propose a compressive sampled CSI feedback method based on deep learning (SampleDL). In SampleDL, the massive MIMO channel matrix is sampled uniformly in time/frequency dimension before being fed into neural networks (NNs), which will reduce the computational resource/time at user equipment (UE) as well as enhance the CSI recovery accuracy at the BS. Both theoretical analysis and normalized mean square errors (NMSE) results confirm the advantages of the proposed method in terms of time complexity and recovery accuracy. Besides, a suitable CSI feedback period is explored by link level simulations, which aims to further reduce the overhead of CSI feedback without degrading the communication quality.**

***Index Terms*—Channel state information, feedback, deep learning, frequency division duplexing, massive MIMO.**

I. Introduction

Massive multiple-input multiple-output (MIMO) is one of the key technologies to support the high spectrum efficiency in the fifth generation (5G) and beyond wireless communication systems [1]–[3]. The significant gains of massive MIMO technology rely on beamforming that requires accurate downlink channel state information (CSI) at the transmitters [4], [5]. In frequency division duplex (FDD) massive MIMO systems, receivers need to estimate the downlink CSI from the reference signals first and then feed them back to the transmitters, as channel reciprocity does not exist exactly between different frequency bands [6]. However, the CSI feedback overhead is restricted due to the limited uplink resources assigned to CSI feedback. In addition, the size of downlink CSI scales with the number of the antennas, which causes the CSI feedback overhead of FDD massive MIMO systems extremely high. This is a major hurdle for the practical deployment of FDD massive MIMO systems. Hence, it is crucial to provide accurate downlink CSI for transmitters with limited CSI feedback overhead. To reduce the overhead of CSI feedback, the core idea of the existing methods is that the receivers convey the downlink CSI to the transmitters using limited bits, also refers as limited feedback [6].

Traditional limited feedback methods are mainly based on codebook and compressive sensing (CS). In codebook-based methods [7]–[11], the feedback information is the binary index of the codeword chosen in the pre-defined codebook. In massive MIMO systems, however, the codebook size will increase exponentially with large number of antennas, which make codebook based look-up approach for CSI feedback infeasible. The CS-based methods [12]–[14], by utilizing the spatial and temporal correlations of channel, transform the correlated CSI into an uncorrelated sparse vector for feedback. However, these traditional methods need complicated iterative process and cause too long CSI feedback delays to satisfy the low latency requirements. Hence, new advanced methods are required to solve the technical challenge.

In recent years, inspired by the powerful ability of deep learning (DL), many DL-based method have been successfully developed in wireless communications [15]–[20]. To break through the performance of the conventional limited feedback methods, C.-K. Wen *et al*. [21] first proposed a DL-based limited CSI feedback method and pointed out that DL has potential to discover the inherent structure of CSI and thereby it is a good solution for limited CSI feedback. C.-K. Wen *et al*. [21] designed an autoencoder network (called CsiNet) by using convolution and fully connected layers for CSI compression and recovery. That paper [21] showed that the CsiNet recovers CSI with improved accuracy at the same compression ratios and performs multifold times faster than that of some CS-based approaches including an iterative thresholding algorithm [22] and the denoising-based approximate message passing algorithm [23]. Based on the CsiNet, a series of DL-based CSI feedback methods were proposed by considering more practical conditions [24]–[27]. Considering the temporal

This work was supported in part by the JSPS KAKENHI under Grant JP19H02142, Major Project of the Ministry of Industry and Information Technology of China under Grant TC190A3WZ-2, National Natural Science Foundation of China under Grant 61901228, the Summit of the Six Top Talents Program of Jiangsu under Grant XYDXX-010, the Program for High-Level Entrepreneurial and Innovative Team under Grant CZ002SC19001, the project of the Key Laboratory of Universal Wireless Communications (BUPT) of Ministry of Education of China under Grant KFKT-2020106. (*Corresponding authors: Guan Gui; Hikmet Sari*)

J. Wang and G. Gui, and H. Sari are with the College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China. (e-mail: {2018010223, guiguan, hikmet}@njupt.edu.cn).

T. Ohtsuki is with the Department of Information and Computer Science, Keio University, Yokohama 223-8521, Japan (e-mail: ohtsuki@ics.keio.ac.jp).

B. Adebisi is with the Department of Engineering, Faculty of Science and Engineering, Manchester Metropolitan University, UK (Email: b.adebisi@mmu.ac.uk).

H. Gacanin is with the Faculty of Electrical Engineering and Information Technology, RWTH Aachen University, Aachen, Germany (Email: harisg@ice.rwth-aachen.de).

correlation in time-varying channels, [24] extended CsiNet with three long short term memory (LSTM) layers to explore the temporal correlation between the current feedback and the previous CSI for further enhancing the reconstruction quality. Similarly, to utilize the temporal correlation in the time-varying channels, [26] introduced a convolutional LSTM (ConvLSTM) layer to CsiNet and achieved better accuracy performance than [24]. [25] explored the channel reciprocity between the uplink CSI and the downlink CSI to improve the CSI recovery quality. Considering the practical noisy feedback link, [27] proposed a specific noise extraction unit for codeword denoising. Problems including quantization, multiple compression ratios, and security in CSI feedback are first pointed out in [28], [29]. Different from the above related works considering CSI feedback as a stand-alone module, [30] proposed an innovative end-to-end precoding scheme for FDD downlink systems. In [30], the received pilots are mapped into feedback bits and the BS maps the feedback bits into precoding matrix directly. The joint design of channel feedback and precoding achieves outstanding performance and deserves further exploration. The aforementioned existing DL-based CSI feedback methods can work well in multiple-input single-output (MISO) systems. To the best of our knowledge, no paper has reported the same for massive MIMO systems. The main challenge is that the size of massive MIMO channel is large and the computational complexity of CSI compression increases significantly.

In this paper, we propose a compressive sampled CSI feedback method based on deep learning (SampleDL) for FDD massive MIMO systems. By exploiting the channel correlations in time/frequency dimension, the method combines sampling with neural networks (NNs) to compress the massive MIMO channel in three dimensions (i.e., time, frequency, and space). Specifically, the proposed method reduces the size of the massive MIMO channel by sampling in time/frequency dimension and then compress the sampled channel in space dimension by NNs. This approach reduces the computational resource/time at UE as well as enhances the CSI recovery accuracy at BS. We evaluate the performance of the proposed SampleDL method with the 5G new radio clustered delay line MIMO link-level fading channel (nrCDLChannel) model that follows 3GPP TR 38.901 specifications [31]. Our results demonstrate that the proposed SampleDL method outperforms the DL-based CSI compression method without sampling in terms of computational complexity and recovery accuracy. We also provide discussions on how often the receivers should perform feedback with considerable communication quality by link-level simulations in order to reduce the CSI feedback overhead.

II. System Model and Problem Formulation

Consider a single-cell FDD massive MIMO downlink system where the base station (BS) with N_t antennas and the user equipment (UE) with N_r antennas. The system adopts orthogonal frequency division multiplexing (OFDM) over N_f subcarriers. The received signal at n-th subcarrier is

$$\mathbf{y}_n = \mathbf{H}_n\mathbf{P}_n\mathbf{x}_n + \mathbf{w}_n \tag{1}$$

where $\mathbf{x}_n \in \mathbb{C}^{N_r}$ is the vector of transmitted data, $\mathbf{P}_n \in \mathbb{C}^{N_t \times N_r}$ is the precoding matrix at BS, $\mathbf{H}_n \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix between N_t transmit antennas and N_r receive antennas on the n-th subcarrier and $\mathbf{w}_n \in \mathbb{C}^{N_r}$ is the additive white Gaussian noise (AWGN) vector. In the FDD massive MIMO downlink system, the UE needs to convey the downlink CSI to the BS for designing $\mathbf{P}_n$ or other technologies, which is also known as CSI feedback. In this paper we consider nrCDLChannel model that conforms to 3GPP TR 38.901 specifications [31]. The downlink CSI matrix $\mathbf{H}$ of one time slot is a N_f-by-N_s-by-N_r-by-N_t complex array, where N_s denotes the number of OFDM symbols. The total number of feedback parameters is $N_fN_sN_rN_t$, which consume amounts of limited uplink bandwidth when N_r or N_t is large. In general, the downlink CSI feedback process at UE consists of three steps including compression, quantization and entropy encoding. The BS obtains the downlink CSI by implementing inverse operations. In this paper, we assume that the feedback link is error free and the entropy encoding/decoding process is omitted when train the neural network same as in [21] because it is lossless.

In detail, when the UE estimates the downlink CSI, the dimensionality reduction of $\mathbf{H}$ is done first. Then the UE quantizes the compressed downlink CSI with limited bits for further reducing the feedback overhead and practical transmitting. Let $f_C(\cdot)$ and $f_Q(\cdot)$ denote the compression function and quantization function, respectively. The final values of downlink CSI that the BS receives can be given as

$$\mathbf{H}_c^q = f_Q(f_C(\mathbf{H}, \Phi_1)) \tag{2}$$

where Φ_1 denotes the parameters of the compression function. When the BS receives $\mathbf{H}_c^q$, the dequantization and decompression will be used for recovering the original downlink CSI. Let functions $f_C^{-1}(\cdot)$ and $f_Q^{-1}(\cdot)$ denote the decompression and dequantization operations, respectively. The recovered downlink CSI at BS is obtained as

$$\widehat{\mathbf{H}} = f_C^{-1}(f_Q^{-1}(\mathbf{H}_c^q), \Phi_2) \tag{3}$$

where Φ_2 denotes the parameters of the decompression function. From (2) and (3), it can be observed that quantization errors and compression loss dominate the feedback errors. To avoid architectural change and parameter updating at the BS/UE under different number of quantization bits, this paper first focuses on the downlink CSI compression and decompression without considering the quantization errors by combining (2) and (3) using mean square errors (MSE) as

$$(\widehat{\Phi}_1, \widehat{\Phi}_2) = \arg\min_{\Phi_1,\Phi_2} \left\|\mathbf{H} - f_C^{-1}((f_C(\mathbf{H}, \Phi_1)), \Phi_2)\right\|_2^2. \tag{4}$$

After obtaining the well trained compression/decompression model, quantization are considered when evaluating the accuracy performance of the proposed method for downlink CSI compression and recovery.

III. The Proposed SampleDL Method

By exploiting the channel correlations in time, frequency and space dimensions, this paper proposes a CSI compression

and decompression method by combining sampling with deep learning, i.e., SampleDL method. The detailed procedure of the proposed SampleDL method is shown in Fig. 1. In this section, we present the proposed SampleDL method by three parts including CSI compression at UE, CSI recovery at BS and the structure of NNs used in SampleDL method.

A. CSI Matrix Compression at the UE

In order to conform to the processing structure of the NNs, the dimensions of $\mathbf{H}$ are first reshaped into $N_f \times N_s \times N_p$ without changing the inherent structure of the channel, where $N_p = N_r \times N_t$ denotes the number of transceiver antennas. Theoretical wireless channel models [33] have shown that the channel gain is highly correlated over coherence time and between adjacent subcarriers of a wide frequency band. Grouping of neighboring subcarriers is a method of CSI matrix compression in 802.11n/ac standards [5], [34]. Inspired by the prior knowledge of channel correlations in time and frequency dimensions, the first step of SampleDL method is to compress the CSI matrix from time/frequency dimension by sampling, see Fig. 1. The detailed sample operations include sampling N_{fs} elements per N_f subcarriers and sampling N_{ss} elements per N_s realizations, by which the UE compresses the full CSI matrix to the sampled CSI $\widetilde{\mathbf{H}}_s \in \mathbb{C}^{N_{fs} \times N_{ss} \times N_p}$ without any computation. The sampling ratio in the first step is defined as $R_s = R_{fs} \times R_{ss}$, where $R_{fs} = N_{fs}/N_f$ and $R_{ss} = N_{ss}/N_s$. One example of uniform sampling method is shown in Fig. 2.

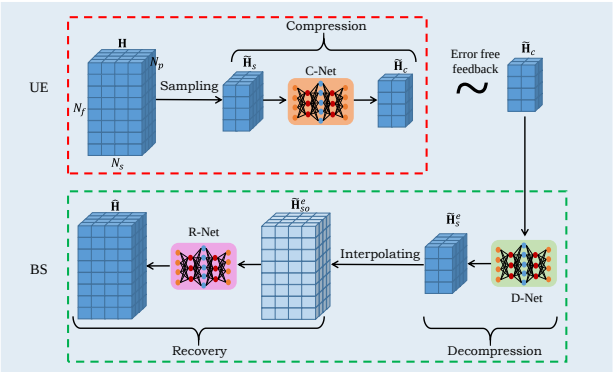

Fig. 1. The procedure of the SampleDL method where $N_f = 8$, $N_s = 4$, $N_r = 2$, $N_t = 2$, $R_{fs} = 1/2$, $R_{ss} = 1/2$, $R_c = 1/2$.

In the wireless multi-path channel model, channel gains of a set of antennas have a hidden relationship [35], [36]. Considering the channel correlation in space dimension, the second step of SampleDL method is to compress the sampled CSI from space dimension. The detailed compression operation is compressing N_p elements in space dimension of $\widetilde{\mathbf{H}}_s$ into N_{pc} elements by a compression NN (C-Net), by which the UE obtains the final compressed downlink CSI $\widetilde{\mathbf{H}}_c \in \mathbb{C}^{N_{fs} \times N_{ss} \times N_{pc}}$, see Fig. 1. The space compression ratio in the second step is defined as $R_c = N_{pc}/N_p$. After sampling and compression, the final number of feedback parameters is $N_{fs} \times N_{ss} \times N_{pc}$. The total compression ratio is defined as $R_T = R_s \times R_c$.

B. CSI Matrix Recovery at the BS

We assume that the UE feeds the compressed CSI $\widetilde{\mathbf{H}}_c$ back to the BS through the error free feedback link. The assumption is justifiable because the feedback link is usually protected using error correction coding and hence has a very low error probability [37]. When the BS receives the compressed CSI $\widetilde{\mathbf{H}}_c$, the first step is to recover $\widetilde{\mathbf{H}}_s$ by a decompression NN (D-Net). The recovered sampled downlink CSI is denoted as $\widetilde{\mathbf{H}}_s^e$. The second step is to interpolate $\widetilde{\mathbf{H}}_s^e$ with 0 to restore its dimensions to the original dimensions $N_f \times N_s \times N_p$ and put the interpolated CSI to refine NN (R-Net) to recover the real CSI $\mathbf{H}$. The interpolated $\widetilde{\mathbf{H}}_s^e$ is denoted as $\widetilde{\mathbf{H}}_{so}^e$ which can be regarded as a coarse CSI matrix. One example of the sampling and interpolating process is shown in Fig. 2. The R-Net recovers the missing information of $\widetilde{\mathbf{H}}_{so}^e$ by exploiting the channel correlations in time and frequency dimensions, and outputs the recovered downlink CSI $\widehat{\mathbf{H}}$.

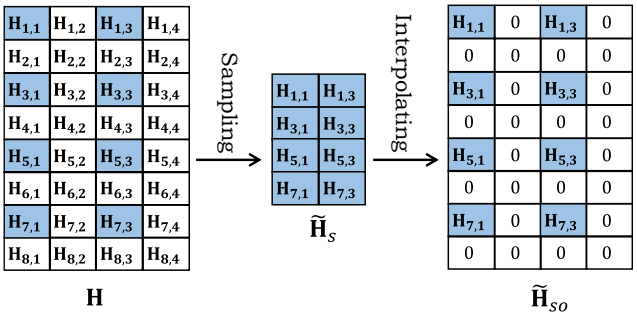

Fig. 2. The uniform sampling and interpolating process where $N_f = 8$, $N_s = 4$, $N_r = 1$, $N_t = 1$, $R_{fs} = 1/2$, $R_{ss} = 1/2$.

C. The Network Structure of C-Net, D-Net and R-Net

DL can discover the inherent structure of data and thereby provide solutions for our problems that aim to explore channel correlations in time, frequency and space dimensions. Csi-Net [21], the state-of-the-art DL solution for CSI feedback, is not suitable for processing the massive MIMO channel matrix in this paper. The reason is that the number of feedback codewords of the downlink CSI $\mathbf{H} \in \mathbb{C}^{N_f \times N_s \times N_r \times N_t}$ is 129024 in this paper, which is too large to be compressed by the fully connected layer that is used in Csi-Net. To resolve this problem, we apply 3D convolution layer (Conv3D) to the C-Net, D-Net and R-Net. The relationship between C-Net, D-Net and R-Net are shown in Fig. 1. C-Net and D-Net are trained together by end-to-end learning. The network structure of C-Net and D-Net is shown in Fig. 3 where different types of neural network layers are on different colors and some hyper parameter settings of them are shown. The C-Net consists of six Conv3D layers for feature extraction and one Conv3D layer for compression. Different from feature extraction layers, the strides of compression layer are $(R_c, 1, 1)$. Different space compression ratios can be achieved by adjusting the value of R_c. However, the strides of all feature extraction layers are $(1, 1, 1)$ for just learning the features of $\widetilde{\mathbf{H}}_s$ without changing the size of it. Through error free feedback, the output of the C-Net $\widetilde{\mathbf{H}}_c$ arrives at the BS and is fed into the D-Net to recover

the sampled downlink CSI $\widetilde{\mathbf{H}}_s$. The first layer of the D-Net is Conv3DTranspose layer for decompressing $\widetilde{\mathbf{H}}_c$ and obtaining the coarse $\widetilde{\mathbf{H}}_s$. Notably, the strides of Conv3DTranspose layer are same as that of compression layer in the C-Net. Following the decompression layer, there are three residual networks to refine the coarse $\widetilde{\mathbf{H}}_s$ and finally output the estimated sampled downlink CSI $\widetilde{\mathbf{H}}_s^e$ by exploiting the channel correlation in space dimension. Each residual network consists of three Conv3D layers. For all Conv3D layers in the C-Net and the D-Net, the activation function is leaky rectified linear unit (Leaky ReLU), the kernel size is $3 \times 3 \times 3$ and the parameter padding is set as "same" for keeping the size of the input data unchanging. We train the C-Net and the D-Net by using adaptive moment (Adam) algorithm. The loss function is the MSE function which is defined as

$$\mathcal{F}_1 = \left\| \widetilde{\mathbf{H}}_s - \widetilde{\mathbf{H}}_s^e \right\|_2^2 \tag{5}$$

where $\|\cdot\|_2$ denotes the Euclidean norm.

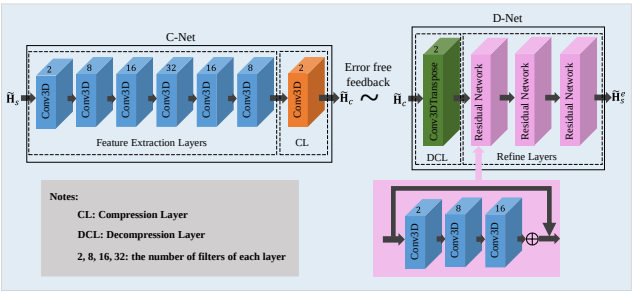

Fig. 3. The network structure of the C-Net and the D-Net.

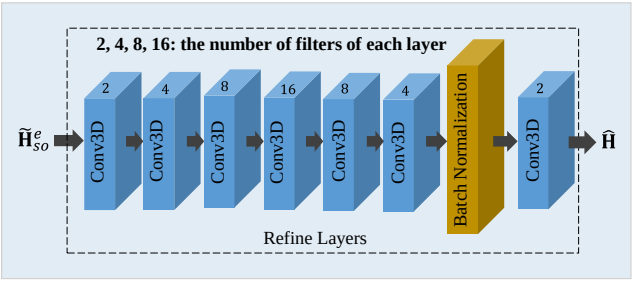

Fig. 4. The network structure of the R-Net.

After training the C-Net and the D-Net, the BS interpolates the output of D-Net $\widetilde{\mathbf{H}}_s^e$ with 0 to restore its dimensions to the original dimensions $N_f \times N_s \times N_p$ and use the interpolated CSI to train R-Net for recovering the original downlink CSI $\mathbf{H}$. The network structure of the R-Net is shown in Fig. 4. The R-Net only consists of seven Conv3D layers and one Batch Normalization layer for low complexity. Conv3D layers are used to learn the channel correlations in time/frequency dimension. Batch Normalization layer is used to prevent R-Net overfitting. The activation functions of the first six Conv3D layers are Leaky ReLU and the activation function of the last Conv3D layer is a linear function. The optimizer of R-Net is also Adam and the loss function via MSE is defined as

$$\mathcal{F}_2 = \left\| \mathbf{H} - \widehat{\mathbf{H}} \right\|_2^2 \tag{6}$$

where $\widehat{\mathbf{H}}$ is the output of the R-Net. In the proposed SampleDL method, the well trained C-Net, D-Net and R-Net under a specific compression ratio can work well with different quantization bits (which will be demonstrated by the experimental results in the next section). However, the C-Net, D-Net and R-Net need to be retrained for different compression ratios $\left\{R_T^1, R_T^2, \cdots, R_T^N\right\}$, where N is the number of different compression ratios. The detailed procedure of the SampleDL method is given in **Algorithm 1**.

Algorithm 1: The proposed SampleDL method.

Input: training data: D_{train}, testing data: D_{test}, quantization bits: B, different total compression ratios: $\left\{R_T^1, R_T^2, \cdots, R_T^N\right\}$, maximum epoches: E_{max}.

[Off-line training stage]:
for $r = R_T^1, R_T^2, \cdots, R_T^N$ **do**
  Randomly initialize the parameters of C-Net, D-Net, R-Net as Φ_{C-Net}^0, Φ_{D-Net}^0, and Φ_{R-Net}^0, respectively;
  for $i = 1, 2, \cdots, E_{max}$ **do**
    for $\mathbf{H}$ *in* D_{train} **do**
      Sample $\mathbf{H}$ and then obtain the $\widetilde{\mathbf{H}}_\mathbf{s}$;
      Update Φ_{C-Net}^i, Φ_{D-Net}^i via end-to-end training with $\widetilde{\mathbf{H}}_\mathbf{s}$;
    end
  end
  Output: $\Phi_{C-Net}^{r-final}$, $\Phi_{D-Net}^{r-final}$.
  for $i = 1, 2, \cdots, E_{max}$ **do**
    for $\mathbf{H}$ *in* D_{train} **do**
      Update Φ_{R-Net}^i via training with $\widetilde{\mathbf{H}}_\mathbf{so}^\mathbf{e}$ and $\mathbf{H}$;
    end
  end
  Output: $\Phi_{R-Net}^{r-final}$.
end
[Online testing stage]:
for $r = R_T^1, R_T^2, \cdots, R_T^N$ **do**
  for $\mathbf{H}$ *in* D_{test} **do**
    $\mathbf{H} \overset{Sample}{\rightarrow} \widetilde{\mathbf{H}}_\mathbf{s}$;
    $\widetilde{\mathbf{H}}_\mathbf{s} \overset{\text{C-Net}}{\rightarrow} \widetilde{\mathbf{H}}_\mathbf{c}$;
    Quantize $\widetilde{\mathbf{H}}_\mathbf{c}$ with B bits and then obtain $\widetilde{\mathbf{H}}_\mathbf{c}^\mathbf{q}$;
    $\widetilde{\mathbf{H}}_\mathbf{c}^\mathbf{q} \overset{\text{D-Net}}{\rightarrow} \widetilde{\mathbf{H}}_\mathbf{s}^\mathbf{e}$;
    $\widetilde{\mathbf{H}}_\mathbf{s}^\mathbf{e} \overset{\text{Interpolate}}{\rightarrow} \widetilde{\mathbf{H}}_\mathbf{so}^\mathbf{e}$;
    $\widetilde{\mathbf{H}}_\mathbf{so}^\mathbf{e} \overset{\text{R-Net}}{\rightarrow} \widehat{\mathbf{H}}$;
  end
  Output: $\widehat{\mathbf{H}}$ when total compression ratio is r.
end

IV. Experimental Results

This section describes the simulation parameter settings including 5G new radio (NR) link, channel model, training strategy, comparing methods and hyper parameter settings of NNs. Then we analyze the time complexity and accuracy performance of the proposed SampleDL method.

A. Parameter Settings

We evaluate the performance of the proposed SampleDL method on the nrCDLChannel model that conforms to 3GPP TR 38.901 [31]. The training data and testing data are generated by utilizing a link-level simulation for the 5G NR Release 15 of MATLAB, which simulates the physical downlink shared channel (PDSCH) throughputs of a 5G NR link [32]. The parameters of the nrCDLChannel model are set as Table I. By setting 80,000 different seeds, the link-level simulation generates 80,000 channel data. The velocities of UE of the 80,000 channel data are chosen uniformly and randomly in $\{4.8, 24, 40, 60\}$ km/h.

TABLE I
THE PARAMETER SETTINGS OF NRCDLCHANNEL.

Parameters	Values
Delay profile	"CDL-A"
Delay spread (s)	1.29×10^{-7}
Carrier frequency (Hz)	2.1×10^{9}
Velocity of UE (km/h)	$\{4.8, 24, 40, 60\}$
The size of transmit antenna array	$N_t = 32$
The size of receive antenna array	$N_r = 2$
The number of subcarriers	$N_f = 72$
The number of OFDM symbols	$N_s = 14$
Subcarrier spacing (Hz)	$F_s = 1.5 \times 10^4$
seeds	$[1, 80000]$

In the SampleDL method, the C-Net and the D-Net are trained together by end-to-end learning. After the C-Net and the D-Net has been well trained, the output of the D-Net is interpolated with 0 to restore their original dimensions and then constitute the data for training the R-Net. To verify the superiority of the proposed SampleDL method, we present two comparing methods. One comparing method is Non-sample method of which the whole procedure is shown in Fig. 5. The only difference between the Non-sample method and the SampleDL method is that the sampling and interpolating steps are removed in the Non-sample method, i.e. the original downlink CSI $\mathbf{H}$ is the input of the C-Net and the output of the D-Net is the input of the R-Net. The Non-sample method can accomplish different total compression ratios R_T by adjusting the strides (R_c, R_c^s, R_c^f) of compression layer in the C-Net. R_c denotes the space compression ratio same as in SampleDL method. R_c^s and R_c^f denote the compression ratio in time dimension and frequency dimension, respectively. The strides of decompression layer in the D-Net are set as same as the strides of compression layer in C-Net to recover the compressed CSI matrix to its original size. Another comparing method is channel reconstruction network (CRNet) in [38]. In order to apply the CRNet to the problem this paper want to solve, we replace the fully connected (FC) and 2D convolution layers in CRNet with Conv3D layer but keep its core design of decoder, i.e. two CRBlocks. The network structure of the CRNet in this paper is shown in Fig. 6, where $\{2, 4, 7, 8, 16, 32\}$ denotes the number of filters of each layer and $d \times h \times w$ denotes the kernel size of each layer. The design of the compression layer and the decompression layer in CRNet is the same as in the Non-sample method, which can accomplish different total compression ratios R_T by adjusting the strides (R_c, R_c^s, R_c^f).

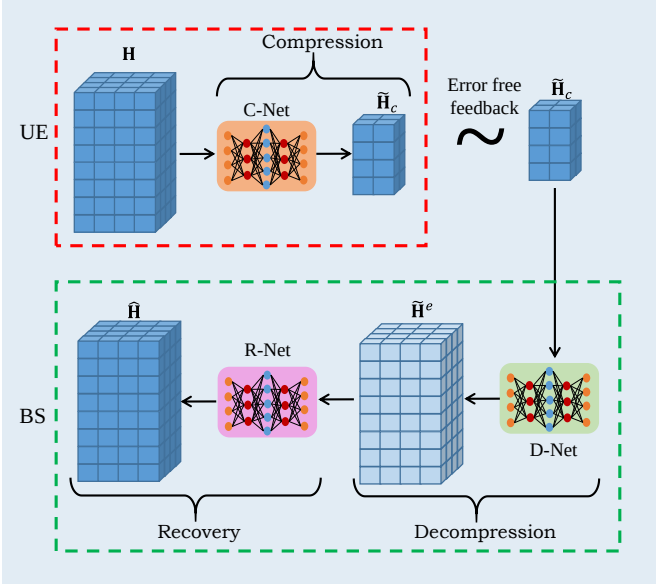

Fig. 5. The whole procedure of non-sample method, where $N_f = 8$, $N_s = 4$, $N_r = 2$, $N_t = 2$ and the strides size in compression/decompression layer is $(2, 2, 2)$.

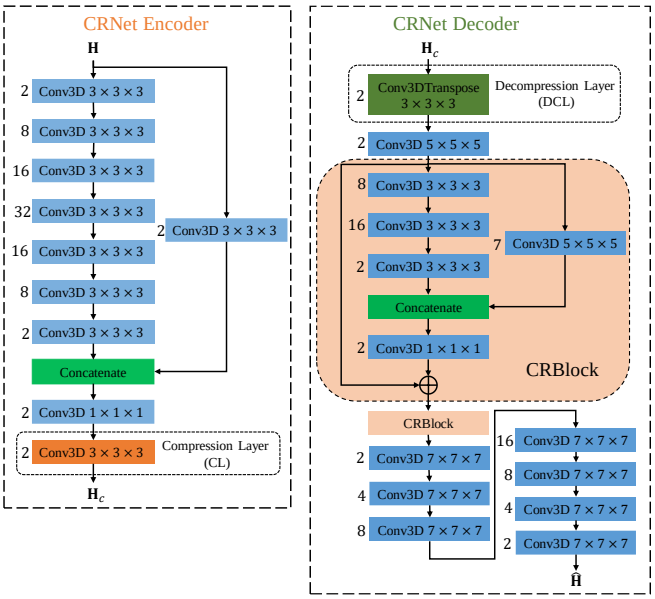

Fig. 6. The network structure of the CRNet.

The hyper parameter settings of NNs and training strategy in Non-sample method are the same as those in the proposed SampleDL method. The hyper parameter settings of NNs in SampleDL/Non-sample method are listed in Table II. In detail, if the validation loss does not decrease during 10 epochs, the learning rate drops to 0.5 times of the original learning rate in R-Net when the total compression ratio R_T is 1/84 and the learning rate drops to 0.8 times of the original learning rate in any other training process. Besides, if the validation loss

does not decrease during 25 epochs, the training process will be stopped. In Table II(a), the training data and the validating data are the first 40,000 data in the total 80,000 channel data, which also are the training data and the validating data for CRNet. In Table II(b), the training data and the validating data are the data from 40,001 to 70,000 in the total 80,000 channel data. The last 10,000 data in the total 80,000 channel data are used to test the performance of the three well trained methods, i.e., the proposed SampleDL method, the Non-sample method and CRNet. The hyper parameter settings for CRNet are different for different R_T. The initial learning rate is 0.0005 and the learning rate drops to 0.1 times of the original learning rate if the validation loss does not decrease during 5 epochs when $R_T = 1/84$. The initial learning rate is 0.0005 and the learning rate drops to 0.5 times of the original learning rate if the validation loss does not decrease during 5 epochs when $R_T = 1/168$. The initial learning rate is 0.001 and the learning rate drops to 0.5 times of the original learning rate if the validation loss does not decrease during 5 epochs when $R_T = 1/336$ and $R_T = 1/672$. Besides, the training process will be stopped if the validation loss does not decrease during 11 epochs for any R_T.

We explore the performance of the Non-sample method, CRNet and the proposed SampleDL method in different total compression ratios R_T. The parameter settings responding to the different R_T are listed in Table III. In detail, a uniform sampling method, see Fig. 2, is used in the proposed SampleDL method.

TABLE II
THE HYPER PARAMETER SETTINGS OF NNS IN SAMPLEDL/NON-SAMPLE METHOD.

(a) C-Net and D-Net.

Parameters	**Values**
Kernel size	$(3, 3, 3)$
Initial learning rate	0.001
Batch size	64
The number of training data	30000
The number of validating data	10000

(b) R-Net.

Parameters	**Values**
Kernel size	$(7, 7, 7)$
Initial learning rate	0.001
Batch size	64
The number of training data	25000
The number of validating data	5000

B. Time Complexity Analysis

The time complexity of the NNs or algorithms is one of important factors affecting the quality of wireless communications. The time complexity of NNs is dependent on the number of floating point operations (FLOPs). For one Conv3D layer, the number of FLOPs is

$$\text{FLOPs} = 2 \times C_i \times K_1 \times K_2 \times K_3 \times D_o \times H_o \times W_o \times C_o \quad (7)$$

TABLE III
THE PARAMETER SETTINGS CORRESPONDING TO DIFFERENT R_T.

(a) The proposed SampleDL method.

R_T	$1/84$	$1/168$	$1/336$	$1/672$
R_{fs}	$1/13$	$1/13$	$1/13$	$1/13$
R_{ss}	$1/4$	$1/4$	$1/4$	$1/4$
R_c	$1/2$	$1/4$	$1/8$	$1/16$

(b) The Non-sample method and the CRNet.

R_T	$1/84$	$1/168$	$1/336$	$1/672$
Strides size in CL and DCL	$(2, 7, 6)$	$(4, 7, 6)$	$(8, 7, 6)$	$(16, 7, 6)$

where C_i is the number of input tensor channels, C_o is number of output tensor channels, (K_1, K_2, K_3) is the kernel size, and D_o, H_o, W_o are the depth, height, width of the output tensor respectively. The total number of FLOPs of the SampleDL method, CRNet and the Non-sample method are shown in Table IV.

The FLOPs of the Non-sample method and the SampleDL method all consist of the FLOPs of the C-Net, the D-Net and the R-Net. As the total compress ratio increases, the size of the output tensor of the C-Net and that of the input tensor of the U-Net will become smaller, so the FLOPs will decrease. In the SampleDL method, the input of the C-Net and the output of the D-Net are sampled data of which the size is 1/42 of the original downlink CSI data. This will reduce computation resource at UE and BS. Results in Table IV show that the FLOPs of the SampleDL method is 6.2848788×10^9 less than that of Non-sample method, which will reduce the time complexity of the whole CSI feedback process. The total number of FLOPs of CRNet is 8.11562×10^7 less than that of Non-sample method, which is also large because of the high dimensions of the channel matrix.

TABLE IV
THE TOTAL NUMBER OF FLOPS OF THE SAMPLEDL METHOD, CRNET AND THE NON-SAMPLE METHOD.

R_T	Non-sample method	CRNet	SampleDL method
$1/84$	$2.14851132 \times 10^{10}$	$2.14039570 \times 10^{10}$	$1.52002344 \times 10^{10}$
$1/168$	$2.14850302 \times 10^{10}$	$2.14038740 \times 10^{10}$	$1.52001514 \times 10^{10}$
$1/336$	$2.14849887 \times 10^{10}$	$2.14038325 \times 10^{10}$	$1.52001099 \times 10^{10}$
$1/672$	$2.14849680 \times 10^{10}$	$2.14038117 \times 10^{10}$	$1.52000892 \times 10^{10}$

C. Recovery Accuracy Analysis

To analyze the recovery accuracy of the proposed SampleDL method, normalized MSE (NMSE) is used as the evaluation metric, which is defined as

$$\text{NMSE} = \mathbb{E}\left\{ \left\| \mathbf{H} - \widehat{\mathbf{H}} \right\|_2^2 / \left\| \mathbf{H} \right\|_2^2 \right\} \quad (8)$$

where $\mathbf{H}$ is the original downlink CSI and $\widehat{\mathbf{H}}$ is the recovered downlink CSI at the BS. The NMSE performance of the SampleDL method, CRNet and Non-sample method is shown

in Table V and the best NMSE performance results under the same compression ratio are presented in bold font.

TABLE V
THE NMSE (dB) PERFORMANCE OF THE SAMPLEDL METHOD AND THE NON-SAMPLE METHOD.

R_T	NMSE (dB) Non-sample method	NMSE (dB) CRNet	NMSE (dB) SampleDL method
1/84	−22.42	−22.63	**−31.28**
1/168	−18.95	−20.25	**−30.13**
1/336	−14.05	−18.22	**−29.09**
1/672	−11.05	−11.79	**−18.38**

The NMSE performance of CRNet is better than the Non-sample method, which demonstrates the advantage of the CRBlock module in the CRNet. However, Table. V shows that the SampleDL method outperforms CRNet and the Non-sample method at all the compression ratios. Especially, the NMSE performance of Non-sample method decreases significantly when R_T is 1/336. However, the NMSE performance of the SampleDL method still maintains high level when R_T is 1/336, which is better than that of Non-sample method with about 15.04 dB gains. Furthermore, the NMSE performance of the SampleDL method when R_T is 1/672 is comparable with that of the Non-sample method when its R_T is 1/168, which means that the SampleDL method can achieve the same downlink CSI feedback accuracy at four times lower cost than Non-sample method.

To apply the proposed SampleDL method to actual communication scenarios, quantization is introduced to the proposed downlink CSI feedback process, i.e. the compressed codewords are quantized by binary number before being transmitted to the BS. We only test the SampleDL method and the Non-sample method in the following parts due to space limitations. We use uniform quantization method in the SampleDL method and the Non-sample method, and the NMSE performance under different quantization bits in terms of different compression ratios is shown in Fig. 7 and Fig. 8. Fig. 7 shows that when the number of quantization bits is 8 the NMSE performance of the SampleDL method under 1/336 compression ratio degrades less than 2 dB compared with the performance without quantization, and the NMSE performance under the other three compression ratios degrades less than 1 dB compared with the performance without quantization. Simulation results in Fig. 8 also show that when the number of quantization bits is 8 the NMSE performance of Non-sample method under 1/84 and 1/672 compression ratios degrades less than 1 dB compared with the performance without quantization. The above results demonstrate that the well trained downlink CSI feedback methods without considering quantization in training stage is applicable in practical communication scenarios.

However, when the number of quantization bits is 4 the NMSE performance of the SampleDL method under all compression ratios degrades more than 10 dB compared with the performance without quantization. Fortunately, the downlink CSI recovery accuracy of the SampleDL method under 1/168 compression ratio when the number of quantization bits is 6 is better than the performance under 1/84 compression ratio when the number of quantization bits is 4, and the downlink CSI feedback overhead of the first scheme is less than the overhead of the second scheme. The simulation results of the Non-sample method in Fig. 8 show the same phenomenon. The phenomenon demonstrates that the downlink CSI recovery accuracy of the nrCDLChannel model is more sensitive to the number of quantization bits than to the compression ratio. It can give a guide to practical communication that increasing the compression ratio is a better way to reduce downlink CSI feedback overhead than reducing the number of quantization bits.

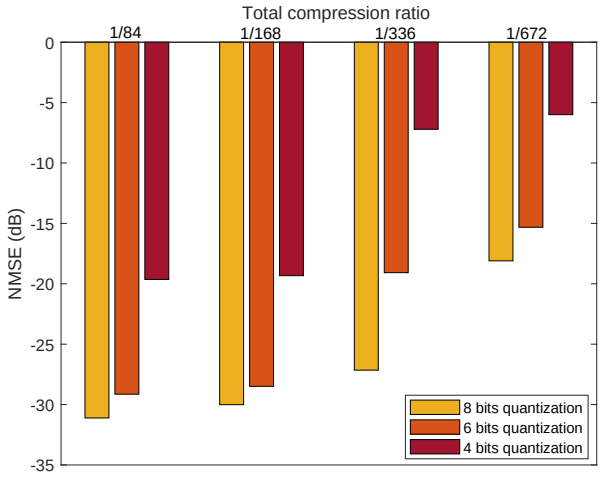

Fig. 7. The NMSE performance of SampleDL method under different quantization bits in terms of different compression ratios.

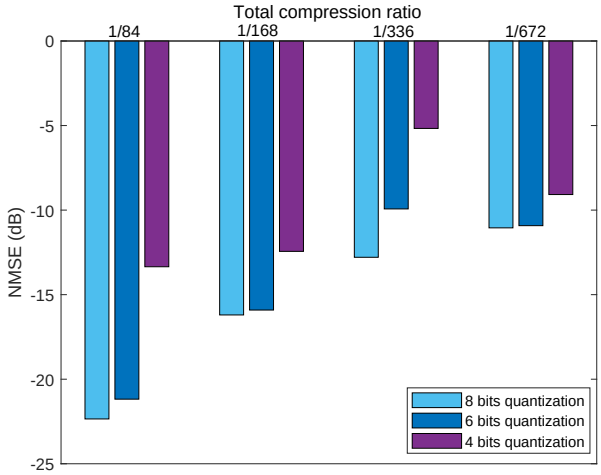

Fig. 8. The NMSE performance of Non-ample method under different quantization bits in terms of different compression ratios.

A generalizable NN architecture is important in practical applications. To evaluate the generalizability of SampleDL method for velocity of UE, the NMSE performance in terms of different moving speeds and different compression ratios are shown in Fig. 9. From Fig. 9 we can see that the fluctuation

of NMSE value is less than 1dB for all R_T when the velocity of UE is less than 60 km/h. This demonstrates that using the dataset generated under various moving speeds enable the generalizability of the SampleDL method for velocity of UE. When the velocity of UE is greater than 60 km/h, such as 80 km/h and 100 km/h, the NMSE performance of the SampleDL method is still considerable. Fig. 9 shows that expanding the range of velocity for generating dataset can improve the generalizability of NNs over a larger speed range.

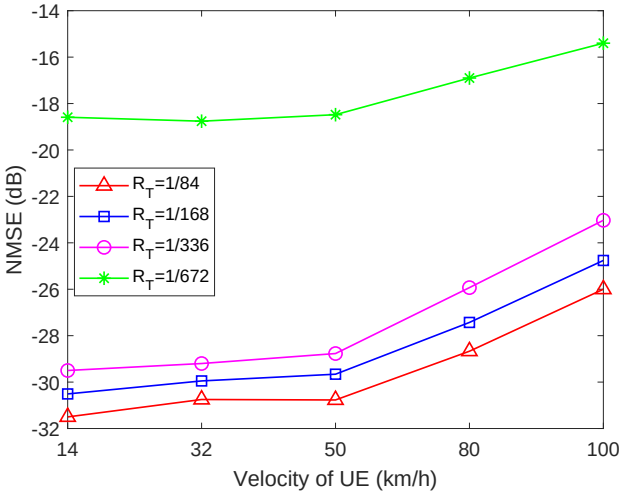

Fig. 9. The NMSE performance of SampleDL method under different velocities of UE in terms of different compression ratios.

D. Underlying Principles for the Success of SampleDL Method

The gains of the SampleDL method may come from two aspects. The first one is when the same NNs are used to process the data here, the NNs can achieve better performance for the data with small size than the data with large size. The second one is that the sampling process in time and frequency dimensions, which can be viewed as introducing the prior knowledge of channel correlations in the two dimensions to the SampleDL method, achieves better recovery performance than the feature learning when using Conv3D layer in the time and frequency dimensions. To explore how the data size contributes to the gains of the SampleDL method, the C-Net and D-Net are trained with sampled data and original data, respectively in term of the same data processing, where the C-Net just compresses the data in space dimension and the D-Net recovers the compressed data to the size as the input size. With the different space compression ratios R_c, the NMSE performance of the two data sizes is shown in Table VI. It can be observed that when the C-Net just compresses the data in space dimension, the achieved performance of the data with small size is better than the data with large size. The experimental results evaluate that dimension reduction of CSI matrix by sampling contributes to the gains of SampleDL method. However, when the compression ratio is too large, such as 1/16 in this experiment, the advantage of data with small size decreases. It is reasonable that when the compression ratio is too high, too much data information is lost to accurately recover both data with small size and data with large size.

To further evaluate the gains of dimensionality reduction of data due to sampling, experiments in terms of different data sizes and different sample methods were conducted. Similar to Table VI NMSE was computed as the performance metric. The experimental results are shown in Fig. 10. In Fig. 10, Orig(14,72) denotes the original downlink CSI data with size (64,14,72). First (7,36) denotes the sampled data with size (64,7,36) that samples the first 7 OFDM symbols and the first 36 subcarriers of the original downlink CSI data. Last (7,36) denotes the sampled data with size (64, 7, 36) that samples the last 7 OFDM symbols and the last 36 subcarriers of the original downlink CSI data. First (4,6) denotes the sampled data with size (64,4,6) that samples the first 4 OFDM symbols and the first 6 subcarriers of the original downlink CSI data. Mid (4, 6) denotes the sampled data with size (64,4,6) that samples the middle 4 OFDM symbols and the middle 6 subcarriers of the original downlink CSI data. Last (4,6) denotes the sampled data with size (64,4,6) that samples the last 4 OFDM symbols and the last 6 subcarriers of the original downlink CSI data. Uni (4,6) denotes the sampled data with size (64,4,6) with the uniform sample method as in Fig. 2. Experimental results show that the NMSE performance of data with different small sizes and different sampling methods are all better than the original data with size (64,14,72), which demonstrates the gains of dimensionality reduction of data when using the same NNs to process the data.

TABLE VI
THE NMSE PERFORMANCE OF THE C-NET AND THE U-NET FOR DIFFERENT DATA SIZES.

R_c	NMSE (dB) / (64,14,72)	NMSE (dB) / (64,4,6)
1/2	-30.25	$\mathbf{-35.06}$ (15.90% ↑)
1/4	-28.93	$\mathbf{-32.12}$ (11.03% ↑)
1/8	-23.46	$\mathbf{-29.17}$ (24.34% ↑)
1/16	-10.60	$\mathbf{-11.29}$ (6.51% ↑)

Besides, the performance of the same data size is different under different sampling methods. Fig. 10 shows that First (7,36) and Last (7,36) have different NMSE performance at same compression ratios. And data with size (4,6) under different sampling methods (First (4,6), Mid (4,6) and Last (4,6)) has the same phenomenon. The above results demonstrate that the gains of dimensionality reduction of data is related to the inherent structure of the data to some extent. From Fig. 10 we can also see that the performance of First (7,36) is better than the performance of four kinds of data with size (4,6) under the space compression ratio of 1/4. The reason is that reducing the dimensionality of data can not further improve the performance of the NN for some task when the NN is big enough for the current data size, such as (64,7,36). Similar to Table. VI, Fig. 10 also shows that when the compression ratio is too large, such as 1/672, the gains of dimensionality reduction of data decrease.

The results in Table V and Table VI show that the gains of data with small size were lower than the total gains of

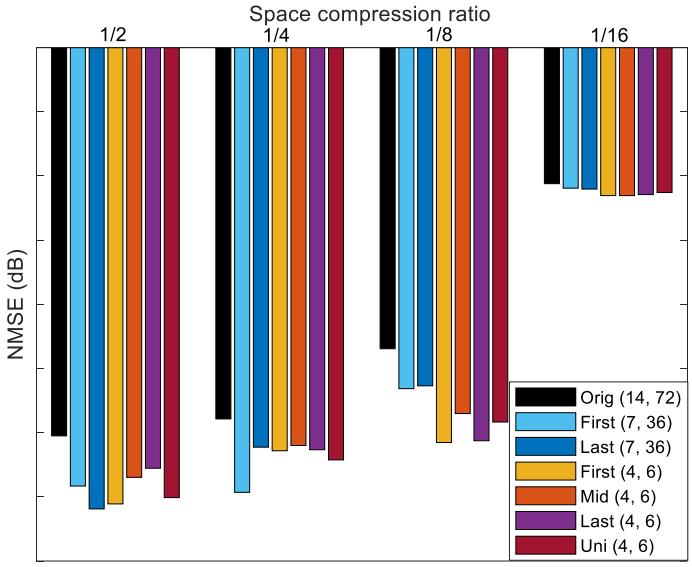

Fig. 10. The NMSE performance of data with different sizes in terms of different space compression ratios.

the SampleDL method, which demonstrates the gains of the second aspect above mentioned in the SampleDL method. To evaluate this gains, we conduct specific experiments where the Non-sample method compress the CSI matrix just from time/frequency dimension with 1/42 compression ratio, i.e. the strides size in compression layer and decompression layer is (1,7,6), and the SampleDL method directly interpolates the sampled CSI with 0 and then feeds them into the R-Net to recover the real CSI. The NMSE performance of the Non-sample method and the SampleDL method under the above conditions are -31.21 dB and -36.00 dB, respectively. It shows that the recovery of sampling process outperforms the feature learning of Conv3D layer in time/frequency dimension. In the Non-sample method, to achieve the compression ratio 1/42 in time/frequency dimension, the strides of compression layer in C-Net must be (1,7,6), which is limited compared with sampling process which is flexible. The above results verify the second reason why the SampleDL method outperforms the Non-sample method.

E. Exploring the Suitable CSI Feedback Period by Link-level Simulations

The longer the feedback period, the lower the feedback overhead, but too long CSI feedback period may decrease the communication quality significantly. To explore the suitable CSI feedback period and its influence factors, the link-level simulation experiments in terms of different CSI feedback periods and communication conditions are conducted. In detail, the link-level simulation implements the physical downlink shared channel (PDSCH), downlink shared channel (DL-SCH) and measures the PDSCH throughput of a 5G NR link. Fig. 11 shows the link-level simulation pipeline, in which the PDSCH demodulation reference signals (DM-RS), PDSCH phase tracking reference signals (PT-RS) and synchronization signal (SS) burst generation are omitted for clarity. In the link-level simulation, the nrCDLChannel model and perfect synchronization are adopted, and the downlink CSI fed back to the BS is used for precoding. The evaluation metric of communication quality in the link-level simulation is throughput [32], which is defined as

$$\text{Throughput} = \frac{B_t \times 10^{-6}}{F_n \times 10^{-2}} \text{ Mbps}, \tag{9}$$

where B_t denotes the practical data bits the link-level simulation transmits successfully, and F_n denotes the number of 10 ms frames of the link-level simulation. The well trained SampleDL and Non-sample networks based on the parameter settings in the part A are applied in the link-level simulation respectively, and the performance of the two methods are compared. The parameter settings of channel model in the link-level simulation are same as in Table I. The specific parameter settings of the link-level simulation are shown in Table VII. The length of the simulation in terms of the number of 10 ms frames is 1,000. The signal-to-noise-ratio (SNR) is defined per resource element (RE) at each UE antenna.

TABLE VII
THE SPECIFIC PARAMETER SETTINGS OF THE LINK-LEVEL SIMULATION.

Parameters	Values
Number of 10ms frames	1000
Modulation scheme	{64QAM, 256QAM}
Code rate	490/1024
SNR (dB)	5
Velocity of UE (km/h)	{4.8, 15}
Downlink CSI feedback period (ms)	{2, 5, 8, 10}

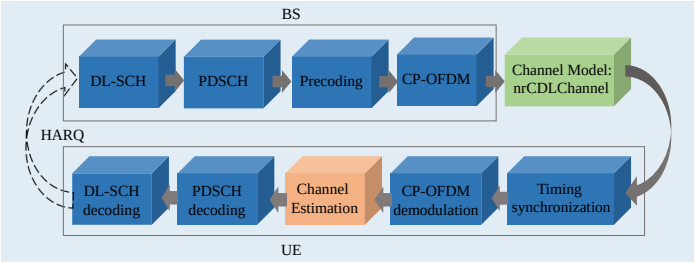

Fig. 11. The implemented pipeline of the link-level simulation.

Fig. 12 shows the throughput performance of the link-level simulation in terms of different CSI feedback periods (i.e., {5, 8, 10} ms) when the modulation scheme is 64QAM and the velocity of UE is 4.8 km/h. The SampleDL and Non-sample methods are applied in the simulation for feeding the downlink CSI back to BS. The upper bound performance is defined as the performance of the link level simulation when the BS obtains the perfect downlink CSI with zero feedback delay. From Fig. 12, it can be observed that the throughput can achieve approximately 95.2% of the upper bound at first three compression ratios by using the SampleDL method when the CSI feedback period is 5 ms. When the compression ratio is 1/672, the throughput still can achieve approximately 93.5% of the upper bound performance, which outperforms 3GPP Release 16 Type II codebook that achieves 89.34% of

the upper bound throughput when the CSI feedback period is 5 ms [39]. Compared with the SampleDL method, when using the Non-sample method for downlink CSI feedback, the throughput decreases slightly at the compression ratio 1/84 under 5 ms feedback period. Furthermore, the throughput decreases approximately 5% when the compression ratios are 1/168, 1/336 and 1/672 respectively. The above results demonstrate that when using the SampleDL method to feed downlink CSI back to BS, the CSI feedback period can be set as 5 ms at the above communication conditions, which reduce the feedback overhead by 5 times at cost of 5% throughput of the upper bound performance. However, when the CSI feedback period is made longer, such as 8 ms and 10 ms, the throughput decreases significantly under both of the SampleDL and the Non-sample methods, which cannot support the practical communications.

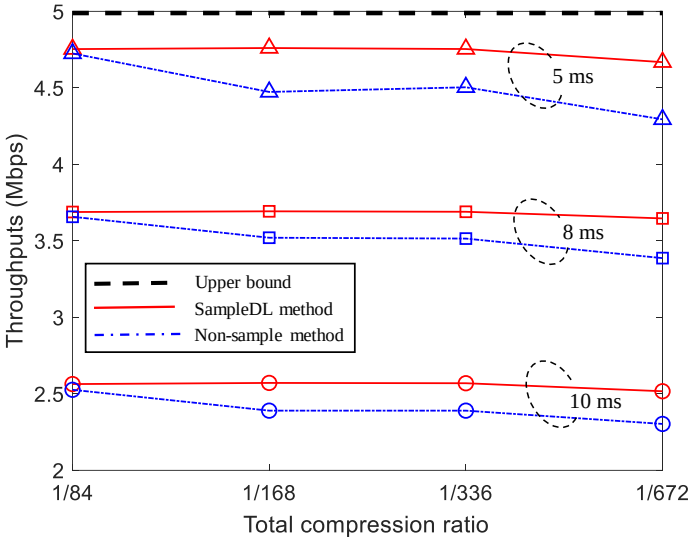

Fig. 12. The throughput performance of the link-level simulation in terms of different CSI feedback periods, different CSI feedback methods and different CSI compression ratios, respectively, when the modulation scheme is 64QAM and the velocity of UE is 4.8 km/h.

To explore the influence factors on the CSI feedback periods, link-level experiments in terms of different velocities of UE and different modulation schemes were conducted. Fig. 13 shows the experimental results when the modulation scheme is 64QAM and the velocity of UE is 15 km/h. The experimental results demonstrate that the throughput performance (just achieve approximately 32% of the upper bound performance) decreases significantly under 5 ms CSI feedback period when the velocity of UE is 15 km/h, which can not support the practical communications. When the CSI feedback period is made shorter, such as 2 ms, the throughput of the link-level simulation at all the CSI compression ratios by using the SampleDL method for downlink CSI feedback can achieve approximately 98% of the upper bound performance. The reason is that when the velocity of UE increases the CSI changes faster, which cause the CSI before 5 ms is outdated for current communications. Similar to Fig. 12, the throughput of the link-level simulation using the Non-sample method for downlink CSI feedback is lower than that of the link-level

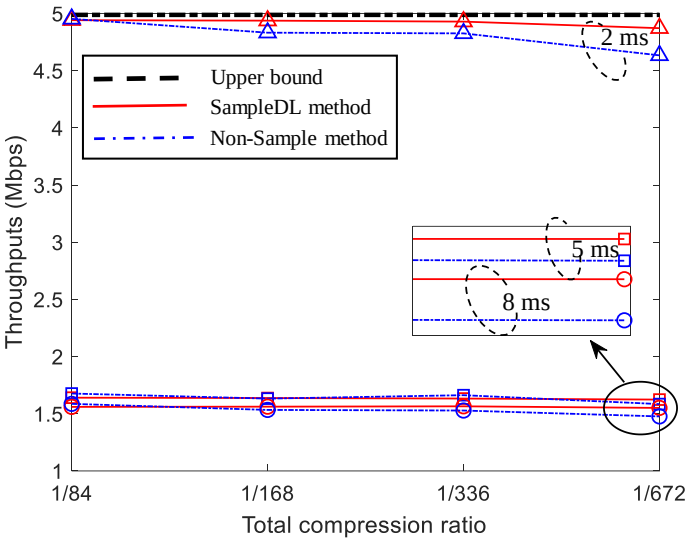

Fig. 13. The throughput performance of the link-level simulation in terms of different CSI feedback periods, different CSI feedback methods and different CSI compression ratios, respectively, when the modulation scheme is 64QAM and the velocity of UE is 15 km/h.

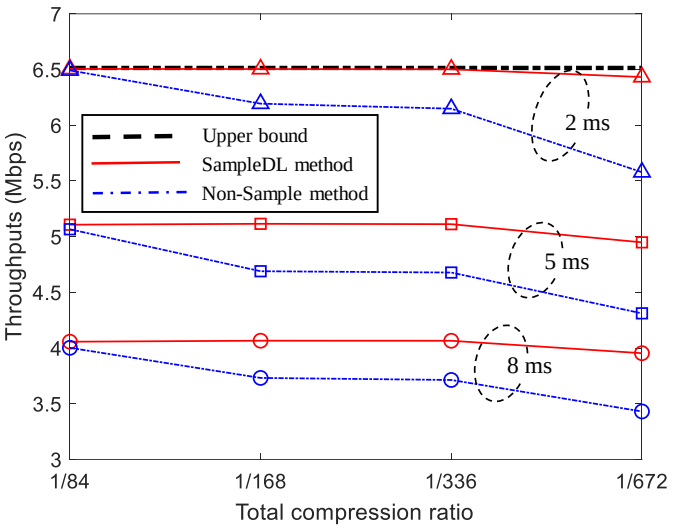

Fig. 14. The throughput performance of the link-level simulation in terms of different CSI feedback periods, different CSI feedback methods and different CSI compression ratios, respectively, when the modulation scheme is 256QAM and the velocity of UE is 4.8 km/h.

simulation using the SampleDL method. The above results show that when the velocity of UE increases, the CSI feedback period will be shorter and shorter to keep high throughput of the communications, which further increases the CSI feedback overhead. Fig. 14 shows the experimental results when the modulation scheme is 256QAM and the velocity of UE is 4.8 km/h, which demonstrates that when the communication quality requirements increase (such as increasing modulation order) the CSI feedback period will also be shorter and shorter to keep high throughput of the communications. The reason is that higher modulation order requires more accurate CSI, which can be supported by more frequent CSI feedback.

The above simulation results demonstrate that lengthening the CSI feedback period properly can reduce the downlink CSI feedback overhead and can be applied in practical wireless communication. However, the velocity of UE and the communication quality requirements will affect the setting of the CSI feedback period. How to trade off the communication quality and the CSI feedback period adaptively is left as our future work.

V. Conclusion

This paper proposed a SampleDL method for downlink CSI feedback in FDD massive MIMO systems. By combing compressive sampling with NNs, the proposed method outperforms the method without sampling, in terms of recovery accuracy and time complexity. Besides, we explored the suitable CSI feedback periods by link-level simulations to further reduce the CSI feedback overhead and guarantee the communication quality. The simulation results showed that lengthening the CSI feedback period properly is applicable in practical communications. However, too long CSI feedback period would decrease the communication quality significantly. In the future work, we will explore how to trade off the communication quality and the CSI feedback period adaptively to further reduce the CSI feedback overhead.

References

[1] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, 2014.

[2] T. L. Marzetta, "Massive MIMO: an introduction," *Bell Labs Tech. J.*, vol. 20, pp. 11–22, 2015.

[3] X. You, *et al*., "Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts," *Sci. CHINA Inf. Sci.*, vol. 64, article number: 110301, 2021.

[4] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, 2014.

[5] X. Xie, X. Zhang, and K. Sundaresan, "Adaptive feedback compression for MIMO networks," in *Proceedings of the Annual International Conference on Mobile Computing and Networking (MobiCom'13)*, pp. 477–488, 2013.

[6] D. J. Love, R. W. Heath Jr, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, 2008.

[7] K. Huang, R. W. Heath, and J. G. Andrews, "Limited feedback beamforming over temporally-correlated channels," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1959–1975, 2009.

[8] V. Raghavan, R. W. Heath and A. M. Sayeed, "Systematic codebook designs for quantized beamforming in correlated MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1298-1310, 2007.

[9] T. Shuang, T. Koivisto, H. Maattanen, K. Pietikainen, T. Roman and M. Enescu, "Design and evaluation of LTE-advanced double codebook," in *IEEE Vehicular Technology Conference (VTC Spring)*, 2011, pp. 1-5.

[10] K. Kim, T. Kim, D. J. Love and I. H. Kim, "Differential feedback in codebook-based multiuser MIMO systems in slowly varying channels," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 578-588, 2012.

[11] S. Schwarz, M. Rupp, and S. Wesemann, "Grassmannian product codebooks for limited feedback massive MIMO with two-tier precoding," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 1119-1135, 2019.

[12] P. H. Kuo, H. T. Kung, and P. A. Ting, "Compressive sensing based channel feedback protocols for spatially-correlated massive antenna arrays," in *IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 492–497, 2012.

[13] X. Rao and V. K. N. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3361–3271, 2014.

[14] Z. Gao, L. Dai, S. Han, I. Chih-Lin, Z. Wang, and L. Hanzo, "Compressive sensing techniques for next-generation wireless communications," *IEEE Wirel. Commun.*, vol. 25, no. 3, pp. 144–153, 2018.

[15] H. Ye, G. Y. Li and B. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 114–117, 2018.

[16] Y. Yang, F. Gao, Z. Zhong, B. Ai and A. Alkhateeb, "Deep Transfer Learning-Based Downlink Channel Prediction for FDD Massive MIMO Systems," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7485–7497, 2020.

[17] J. Wang, Y. Ding, S. Bian, Y. Peng, M. Liu and G. Gui, "UL-CSI data driven deep learning for predicting DL-CSI in cellular FDD systems," *IEEE Access*, vol. 7, pp. 96105–96112, 2019.

[18] H. Huang, Y. Peng, J. Yang, W. Xia, and G. Gui, "Fast beamforming design via deep learning," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 1065–1069, 2020.

[19] Y. Wang, *et al*., "Automatic modulation classification for MIMO systems via deep learning and zero-forcing equalization," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5688–5692, 2020.

[20] E. Balevi, A. Doshi and J. G. Andrews, "Massive MIMO channel estimation with an untrained deep neural network," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 2079–2090, 2020.

[21] C.-K. Wen, W. T. Shih, and S. Jin, "Deep learning for massive MIMO CSI feedback," *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 748–751, 2018.

[22] I. Daubechies, M. Defrise, and C. D. Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math*., vol. 57, no. 11, pp. 1413–1457, 2004.

[23] C. A. Metzler, A. Maleki and R. G. Baraniuk, "From denoising to compressed sensing," *IEEE Trans. Inf. Theory*, vol. 62, no. 9, pp. 5117–5144, 2016.

[24] T. Wang, C. K. Wen, S. Jin, and G. Y. Li, "Deep learning-based CSI feedback approach for time-varying massive MIMO channels," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 416–419, 2019.

[25] Z. Liu, L. Zhang, and Z. Ding, "Exploiting bi-directional channel reciprocity in deep learning for low rate massive MIMO CSI feedback," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 889–892, 2019.

[26] X. Li and H. Wu, "Spatio-temporal representation with deep neural recurrent network in MIMO CSI feedback," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 653–657, 2020.

[27] H. Ye, F. Gao, J. Qian, H. Wang and G. Y. Li, "Deep learning-based denoise network for CSI feedback in FDD massive MIMO systems," *IEEE Commun. Lett.*, vol. 24, no. 8, pp. 1742–1746, 2020.

[28] J. Guo, C. Wen, S. Jin, and G. Y. Li, "Convolutional neural network-based multiple-rate compressive sensing for massive MIMO CSI feedback: design, simulation, and analysis," *IEEE Trans. Wireless Commu.*, vol. 19, no. 4, pp. 2827–2840, 2020.

[29] Q. Liu, J. Guo, C. Wen, and S. Jin, "Adversarial attack on DL-based massive MIMO CSI feedback," *J. Commun. Netw.*, vol. 22, no. 3, pp. 230–235, 2020.

[30] F. Sohrabi, K. M. Attiah and W. Yu, "Deep learning for distributed channel feedback and multiuser precoding in FDD massive MIMO," *IEEE Trans. Wireless Commun.*, early access, doi: 10.1109/TWC.2021.3055202.

[31] 3GPP TR 38.901, "Study on channel model for frequencies from 0.5 to 100 GHz," *3rd Gener. Partnersh. Proj. Tech. Specif. Gr. Radio Access Network.*

[32] https://www.mathworks.com/help/5g/ug/nr-pdsch-throughput.html.

[33] D. Tse and P. Viswanath, "Fundamentals of wireless communications," *Cambridge University Press*, 2005.

[34] P. Pattanayak and P. Kumar, "SINR based limited feedback scheduling for MIMO-OFDM heterogeneous broadcast networks," in *Twenty Second National Conference on Communication (NCC)*, pp. 1–6, 2016.

[35] M. Alrabeiah and A. Alkhateeb, "Deep learning for TDD and FDD massive MIMO: mapping channels in space and frequency," in *53rd Asilomar Conference on Signals, Systems, and Computers*, pp. 1465–1470, 2019.

[36] P. Dong, H. Zhang and G. Y. Li, "Machine learning prediction based CSI acquisition for FDD massive MIMO downlink," in *IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, 2018.

[37] K. Huang, B. Mondal, R. W. Heath Jr. and J. G. Andrews, "WLC38-5: multi-antenna limited feedback for temporally-correlated channels: feedback compression," in *IEEE Global Communications Conference (GLOBECOM)*, pp. 1–5, 2006.

[38] Z. Lu, J. Wang, and J. Song, "Multi-resolution CSI feedback with deep learning in massive MIMO system," in *IEEE International Conference on Communications (ICC)*, pp. 1–6, 2020.

[39] L. Suarez, E. Dombrovsky, V. Lyashev and A. Sherstobitov, "CSI feedback compression with limited downlink pilots for 5G FDD-NR massive MIMO systems," in *IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, 2020.

Jie Wang (S'18) received the B.S. degree from the College of Automation & College of Artificial Intelligence, Nanjing University of Posts and Telecommunications (NJUPT), Nanjing, China, in 2015. She is currently pursuing the Ph.D. degree in the communication and information engineering from the same university. Her research interests include deep learning, channel estimation, and resource allocation and its applications in wireless communications.

Guan Gui (M'11–SM'17) received the Ph.D. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2012. From 2009 to 2014, he joined the Tohoku University as a research assistant as well as a postdoctoral research fellow, respectively. From 2014 to 2015, he was an Assistant Professor in the Akita Prefectural University. Since 2015, he has been a professor with Nanjing University of Posts and Telecommunications, Nanjing, China. His recent research interests include artificial intelligence, deep learning, non-orthogonal multiple access, wireless power transfer, and physical layer security.

Dr. Gui has published more than 200 IEEE Journal/Conference papers and won several best paper awards, e.g., ICC 2017, ICC 2014 and VTC 2014-Spring. He received the IEEE Communications Society Heinrich Hertz Award in 2021, the Elsevier Highly Cited Chinese Researchers in 2020, the Member and Global Activities Contributions Award in 2018, the Top Editor Award of IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY in 2019, the Outstanding Journal Service Award of KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEM in 2020, the Exemplary Reviewer Award of IEEE COMMUNICATIONS LETTERS in 2017. He was also selected as for the Jiangsu Specially-Appointed Professor in 2016, the Jiangsu High-level Innovation and Entrepreneurial Talent in 2016, the Jiangsu Six Top Talent in 2018, the Nanjing Youth Award in 2018. He is serving or served on the editorial boards of several journals, including IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEICE Transactions on Communications, Physical Communication, Wireless Networks, IEEE ACCESS, Journal of Circuits Systems and Computers, Security and Communication Networks, IEICE Communications Express, and KSII Transactions on Internet and Information Systems, Journal on Communications. In addition, he served as the IEEE VTS Ad Hoc Committee Member in AI Wireless, Executive Chair of IEEE VTC 2021-Fall, Vice Chair of IEEE WCNC 2021, TPC Chair of PHM 2021, Symposium Chair of WCSP 2021, General Co-Chair of Mobimedia 2020, TPC Chair of WiMob 2020, Track Chairs of IEEE VTC 2020-Spring, ISNCC 2020 and ICCC 2020, Award Chair of IEEE PIMRC 2019, and TPC member of many IEEE international conferences, including GLOBECOM, ICC, WCNC, PIRMC, VTC, and SPAWC. He is an IEEE Senior Member.

Tomoaki Ohtsuki (SM'01) received the B.E., M.E., and Ph. D. degrees in Electrical Engineering from Keio University, Yokohama, Japan in 1990, 1992, and 1994, respectively. From 1994 to 1995 he was a Post Doctoral Fellow and a Visiting Researcher in Electrical Engineering at Keio University. From 1993 to 1995 he was a Special Researcher of Fellowships of the Japan Society for the Promotion of Science for Japanese Junior Scientists. From 1995 to 2005 he was with Science University of Tokyo. In 2005 he joined Keio University. He is now a Professor at Keio University. From 1998 to 1999 he was with the department of electrical engineering and computer sciences, University of California, Berkeley. He is engaged in research on wireless communications, optical communications, signal processing, and information theory. Dr. Ohtsuki is a recipient of the 1997 Inoue Research Award for Young Scientist, the 1997 Hiroshi Ando Memorial Young Engineering Award, Ericsson Young Scientist Award 2000, 2002 Funai Information and Science Award for Young Scientist, IEEE the 1st Asia-Pacific Young Researcher Award 2001, the 5th International Communication Foundation (ICF) Research Award, 2011 IEEE SPCE Outstanding Service Award, the 27th TELECOM System Technology Award, ETRI Journal's 2012 Best Reviewer Award, and 9th International Conference on Communications and Networking in China 2014 (CHINACOM '14) Best Paper Award. He has published more than 205 journal papers and 415 international conference papers.

He served as a Chair of IEEE Communications Society, Signal Processing for Communications and Electronics Technical Committee. He served as a technical editor of the IEEE Wireless Communications Magazine and an editor of Elsevier Physical Communications. He is now serving as an Area Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and an editor of the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS. He has served as general-co chair, symposium co-chair, and TPC co-chair of many conferences, including IEEE GLOBECOM 2008, SPC, IEEE ICC 2011, CTS, IEEE GLOBECOM 2012, SPC, IEEE ICC 2020, SPC, IEEE APWCS, IEEE SPAWC, and IEEE VTC. He gave tutorials and keynote speeches at many international conferences including IEEE VTC, IEEE PIMRC, IEEE WCNC, and so on. He was Vice President and President of the Communications Society of the IEICE. He is a senior member and a distinguished lecturer of the IEEE, a fellow of the IEICE, and a member of the Engineering Academy of Japan.

Bamidele Adebisi (M'09–SM'15) received his Bachelor's degree in electrical engineering from Ahmadu Bello University Zaria, Nigeria, in 1999, and his Masters degree in advanced mobile communication engineering and Ph.D. degree in communication systems from Lancaster University, United Kingdom, in 2003 and 2009, respectively. He was a senior research associate with the School of Computing and Communication, Lancaster University, from 2005 to 2012. He joined Manchester Metropolitan University in 2012, where he is currently a Full Professor (Chair) of intelligent infrastructure systems. He is the current Vice Chair of IEEE TC-PLC; was General Chair, IEEE ISPLC'18, UK; Co-Chair, 6th IEEE Int'l Conference on Smart Grid Communications, 2015, Miami, US, etc. He is a Panel Member of the UK Engineering and Physical Sciences Research Council (EPSRC) Peer Review College, and an EU H2020 Expert Reviewer/ rapporteur. He has been part of multi-partner, multi-country, multi-million pounds projects as PI and Co-I. One of his projects with an SME received the 2020 UK Best Knowledge Transfer Partnership Project of the Year Awards. He has published over 140 peer-review papers and given several talks/panel discussions in the research areas of Internet of Things, smart cities, smart grids, communication systems and cyber physical systems. Bamidele is a Fellow of IET, a Fellow of Higher Education Academy and a Chartered Engineer.

Haris Gacanin (SM'13–F'20) received his Dipl.-Ing. degree in Electrical engineering from the University of Sarajevo in 2000. In 2005 and 2008, respectively, he received MSc and Ph.D. from Tohoku University in Japan. He was with Tohoku University from 2008 until 2010 first as Japan Society for the Promotion of Science (JSPS) postdoctoral fellow and later, as an Assistant Professor. He joined Alcatel-Lucent Bell (now Nokia Bell) in 2010 as a Physical-layer Expert and later moved to Nokia Bell Labs as Department Head. Since April 2020, he joined RWTH Aachen University. He is a head of the Chair for Distributed Signal Processing and co-director of the Institute for Communication Technologies and Embedded Systems.

His professional interests are related to broad areas of digital signal processing and artificial intelligence with applications in wireless communications. He has 200+ scientific publications (journals, conferences and patents) and invited/tutorial talks. He is a fellow of IEEE. He was a Distinguished Lecturer of IEEE Vehicular Technology Society and an Associate Editor of IEEE COMMUNICATIONS MAGAZINE, while he served as the editor of IEICE Transactions on Communications and IET Communications. He acted as a general chair and technical program committee member of various IEEE conferences. He is a recipient of several Nokia innovation awards, IEICE Communications Society Best Paper Award in 2021, IEICE Communication System Study Group Best Paper Award (joint 2014, 2015, 2017), The 2013 Alcatel-Lucent Award of Excellence, the 2012 KDDI Foundation Research Award, the 2009 KDDI Foundation Research Grant Award, the 2008 JSPS Postdoctoral Fellowships for Foreign Researchers, the 2005 Active Research Award in Radio Communications, 2005 Vehicular Technology Conference (VTC 2005-Fall) Student Paper Award from IEEE VTS Japan Chapter and the 2004 Institute of IEICE Society Young Researcher Award.

Hikmet Sari (F'95-LF'20) received the engineering diploma and the Ph.D. degree from ENST, Paris, France, and the Habilitation degree from the University of Paris XI. From 1980 to 2002, he held various research and management positions at Philips Research Laboratories, SAT, Alcatel, Pacific Broadband Communications, and Juniper Networks. From 2003 to 2016, he was a Professor and the Head of the Telecommunications Department, Suplec, and a Chief Scientist at Sequans Communications. He is currently a Professor with the Nanjing University of Posts and Telecommunications (NJUPT). Dr. Sari's distinctions include the Andre Blondel Medal in 1995, the Edwin H. Armstrong Achievement Award in 2003, the Harold Sobol Award in 2012, as well as election to the Academia Europaea (Academy of Europe) and the Science Academy of Turkey in 2012. He was the Chair of the Communication Theory Symposium of ICC 2002, a Technical Program Chair of ICC 2004, a Vice General Chair of ICC 2006, a General Chair of PIMRC 2010, a General Chair of WCNC 2012, an Executive Chair of WCNC 2014, a General Chair of ICUWB 2014, a General Co-Chair of IEEE BlackSeaCom 2015, a Technical Program Chair of EuCNC 2015, an Executive Co-Chair of ICC 2016, a General Co-Chair of ATC 2016, an Executive Chair of ICC 2017, a General Co-Chair of ATC 2018, and a General Co-Chair of PIMRC 2019. He also chaired the Globecom and ICC Technical Content (GITC) Committee from 2010 to 2011, and was the Communications Society (ComSoc) Vice President for Conferences from 2014 to 2015, the Director for Conference Operations of ComSoc from 2018 to 2019, and the Vice President for Conferences of the IEEE France Section from 2017 to 2019. He is currently serving as a General Chair of WCNC 2021 and an Executive Co-Chair of GLOBECOM 2023. He served as an Editor for the IEEE Transactions on Communications from 1987 to 1991, an Associate Editor for the IEEE Communications Letters from 1999 to 2002, and a Guest Editor for several special issues of the IEEE Journal on Selected Areas in Communications, European Transactions on Telecommunications (ETT), and other journals. He served as a Distinguished Lecturer of ComSoc from 2001 to 2006, a member of the IEEE Fellow Evaluation Committee from 2002 to 2007, and a member of the IEEE Awards Committee from 2005 to 2007.