An Efficient Deep Learning Framework for Low Rate Massive MIMO CSI Reporting

Zhenyu Liu, Lin Zhang, and Zhi Ding

***Abstract*—Channel state information (CSI) reporting is important for multiple-input multiple-output (MIMO) wireless transceivers to achieve high capacity and energy efficiency in frequency division duplex (FDD) mode. CSI reporting for massive MIMO systems could consume large bandwidth and degrade spectrum efficiency. Deep learning (DL)-based CSI reporting integrated with channel characteristics has demonstrated success in improving CSI compression and recovery. To further improve the encoding efficiency of CSI feedback, we develop an efficient DL-based compression framework CQNet to jointly tackle CSI compression, codeword quantization, and recovery under the bandwidth constraint. CQNet is directly compatible with other DL-based CSI feedback works for further enhancement. We propose a more efficient quantization scheme in the radial coordinate by introducing a novel magnitude-adaptive phase quantization framework. Compared with traditional CSI reporting, CQNet demonstrates superior CSI feedback efficiency and better CSI reconstruction accuracy.**

***Index Terms*—Massive MIMO, FDD, CSI feedback, quantization, deep learning.**

I. Introduction

Massive multiple-input multiple-output (MIMO) systems have shown great promise in delivering high spectrum and energy efficiency for 5G and future wireless communication systems [1]. By utilizing a large number of antennas in the massive MIMO framework, gNB (or gNodeB) in 5G can achieve very high downlink throughput if sufficiently accurate downlink channel state information (CSI) is available at the gNB. Consequently, gNB needs to acquire the downlink CSI in an accurate and timely manner to fully utilize the spatial diversity and multiplexing gains.

In time division duplex (TDD) systems, gNB can leverage its uplink CSI as the estimation of its downlink CSI based on the well known reciprocity between downlink and uplink CSIs. In frequency division duplex (FDD) systems, however, uplink and downlink channels are in different frequency bands. Thus, it is difficult to only rely on uplink CSI to estimate the downlink CSI directly as the channel reciprocity becomes weak. Consequently, the gNB transmitter of FDD systems would require user equipment (UE) to provide certain CSI reporting about the downlink CSI. For massive MIMO systems, such feedback data can be substantial since the large number of antennas leads to very high CSI dimensionality [2]. Large bandwidth in high rate links further exacerbates the high feedback load.

To reduce the required bandwidth for CSI reporting in FDD systems, compressed sensing (CS)-based approaches can exploit the channel properties of low rank or sparsity to derive a compressed CSI representation for feedback. Two major correlation properties used in CS-based CSI feedback approaches include spatial CSI correlation [3], [4], [5] that stems from the limited scattering characteristics of signal propagation, and temporal CSI correlation [6] owing to Doppler effects. However, CS-based approaches still have some limitations. On the one hand, CS-based approaches require a strong channel sparsity condition which is not strictly held in some cases. On the other hand, CS algorithms are often iterative and computationally intensive during the decoding process, which may lead to long delays.

Deep learning (DL) is a powerful tool for exploring the underlying structures from large data sets, and has been widely used in computer vision and natural language processing [7]. It can play a helpful role in CSI acquisition when traditional methods generate limited performance. There has been a recent surge of successful applications to derive reliable CSI in massive MIMO systems for channel estimation [8] and low rate CSI feedback [9], [10], [11], [12], [13], [14].

To conserve feedback bandwidth and improve downlink CSI reconstruction accuracy for massive MIMO, the authors of [9] first developed a CSI compression and recovery mechanism using an autoencoder structure [15], and demonstrated better accuracy than CS-based methods in terms of downlink CSI reconstruction from limited UE feedback. Then, channel correlations have been further combined with the DL networks to enhance the CSI feedback performance. The work of [10] exploited the FDD bi-directional channel correlation. By jointly utilizing the available uplink CSI and low rate UE feedback in massive MIMO systems at gNB, downlink CSI reconstruction accuracy gain was shown over the DL architecture based only on UE feedback. Another work [11] proposed the use of a long-short time memory (LSTM) network [16] known as CsiNet-LSTM to exploit the temporal correlation of CSI. In [12], a new LSTM network further reduced the number of parameters to be trained while maintaining the CSI recovery accuracy. Besides, the DL network architectures have also been optimized to improve CSI feedback performance. In [13], a spherical normalization based CSI feedback framework was designed to optimize the input distribution and make the network more applicable to variational signal strength. In [14], a DL network was proposed to extract CSI features on multiple

The work of Z. Liu and L. Zhang was supported in part by the Major Science and Technology Special Projects under Grant 2018ZX03001024, the Industrial Internet Research Institute (Jinan) of Beijing University of Posts and Telecommunications under Grant 201915001, and the China Scholarship Council. The work of Z. Ding was supported in part by the National Science Foundation under Grant 1702752. (Corresponding Author: Lin Zhang)

Z. Liu and L. Zhang are with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: lzyu@bupt.edu.cn, zhanglin@bupt.edu.cn).

Z. Ding is with the Department of Electrical and Computer Engineering, University of California at Davis, Davis, CA 95616 USA (e-mail: zding@ucdavis.edu).

resolutions for massive MIMO systems.

In the aforementioned DL-based works, CSI feedback generally focused on the dimension reduction of CSI matrices, while the redundancy in the codeword can still exist owing to their data type. Common implementations of the DL networks comply with IEEE 754 standard [17]. Single precision, which is defined in this standard and used in CSI feedback [9], [10], [11], [14], is the most typical data type adopted in DL networks [18]. Although DL methods have delivered noticeable performance improvement in reducing the dimension of CSI matrices for massive MIMO, current CSI reporting works that use the single precision (32-bit) to encode feedback coefficients still consume too much bandwidth. To reduce the number of bits required by each codeword, low-bit quantization should take place in encoding after dimension compression.

Clearly, low-bit quantization for CSI codeword can degrade CSI reconstruction accuracy. Consequently, it is important to optimize codeword quantization together with dimension compression to maintain high CSI reconstruction accuracy without consuming excessive bandwidth in massive MIMO systems. A recent work [19] showed that encoding precision could be reduced with manageable accuracy loss in image classification. There are a few recent works involved in the quantization of CSI codewords. In [20], uniform quantization was used to discretize CSI codewords for the feedback. In [21], the μ-law quantization was utilized to encode the codeword after dimension compression. Owing to the non-differentiable property of quantization operation, the weights of encoder networks are hard to be updated during the training. Consequently, only the decoder network can be optimized after quantization, while the weights of the encoder networks are pretrained without the quantizer. However, neither μ-law quantization nor uniform quantization alone can efficiently handle the CSI feedback in low-bit quantization cases, which we shall demonstrate in the performance evaluation section of this work.

To achieve the end-to-end optimization for CSI feedback networks with the quantizer, [22] applied the independently and identically distributed (i.i.d) uniformly distributed noise to approximate the effect of quantization error during training. The authors then exploited entropy encoding to further reduce feedback bits. It was shown that the feedback noise in the training phase can unfortunately degrade the CSI feedback performance during the testing phase [21] because the deep neural networks (DNNs) need to build some redundancy to combat the uncertainty due to noise. Another recent method [23] proposed a specific quantization DNN to improve the CSI reporting performance. It first projected the dimension-compressed codeword to $(-1, 1)$ using the 'tanh' function, before exploiting the uniform quantization together by modifying the quantization gradient to 1 during back propagation. Owing to the range projection and gradient alteration, the portability in a new feedback network and the reuse of pretrained weights of the decoder without quantizer can be influenced. Besides, all codewords share the same quantization parameters, while each codeword within the dimension-compressed vector can have different importance.

In this paper, we develop an efficient CSI compression solution and design an end-to-end DL framework CQNet to jointly optimize CSI compression, codeword quantization, and recovery to improve CSI feedback accuracy and reduce CSI feedback payload. Our CQNet exploits a simple plug-in quantizer that fits nicely into existing DL-based CSI feedback works and can directly facilitate DL-based CSI feedback.

Our contributions in this paper are summarized as follows:

- Inadequately designed quantization module hinders the practical deployment of existing DL-based CSI feedback mechanisms. In this work, we propose a universal CSI compression framework (named CQNet) to jointly optimize CSI compression, codeword quantization, and recovery under limited bandwidth constraint. CQNet is compatible with existing CSI feedback frameworks by only inserting a specially designed quantizer module.
- We develop an efficient quantizer block, which can customize the quantization intervals for each element in a dimension-compressed vector. To further improve efficiency, the traditional μ-law compander is included in an enhanced type of quantization block. We demonstrate that adjusting the element-wise quantization stepsize in a pretrained dimension-compression CSI feedback network can achieve higher accuracy in comparison with retraining the decoder network for the quantized codewords.
- Departing from existing feedback frameworks to separately quantize in-phase component and quadrature component of the downlink CSI (e.g., [9], [11], [12], [14]), we develop a DL-based quantization solution for more efficient CSI encoding in radial coordinate (e.g., [10], [13]) and introduce a special DL-based magnitude-adaptive phase quantization framework. The new framework provides better flexibility of phase quantization and can regulate the weight of quantization entropy to achieve a trade-off between bandwidth and CSI recovery accuracy.
- We substantially integrate CQNet with the enhanced versions [13] of two previous DL-based solutions for CSI compression and feedback CsiNet [9] and DualNet-MAG [10] to demonstrate the simplicity and efficacy of CQNet in improving bandwidth efficiency. Test results demonstrate that the proposed framework can achieve good CSI feedback accuracy and robustness, with low overhead.
- We analyze the codeword quantization and reconstruction of DL-based CSI feedback mechanisms to investigate the underlying principles for the success of CQNet.

The rest of the paper is organized as follows. Section II illustrates the system model. In Section III, we investigate the influence of feedback bandwidth on CSI reconstruction accuracy under a uniform quantization framework. In Section IV, the proposed DL-base joint dimension compression and codeword quantization framework CQNet is introduced in detail. Section V shows the simulation results and performance analysis. Finally, Section VI concludes this paper.

II. System Model

Consider a single-cell massive MIMO system, in which the gNB has $N_b \gg 1$ antennas and the UE has a single antenna.

The system applies orthogonal frequency division multiplexing (OFDM) over N_f subcarriers, for which the downlink received signal at the $n-$th subcarrier is

$$y_d^{(n)} = \mathbf{h}_d^{(n)H} \mathbf{w}_T^{(n)} x_d^{(n)} + n_d^{(n)}, \tag{1}$$

where $\mathbf{h}_d^{(n)} \in \mathbb{C}^{N_b \times 1}$ denotes the channel vector of the $n-$th subcarrier, $\mathbf{w}_T^{(n)} \in \mathbb{C}^{N_b \times 1}$ denotes the transmit beamformer, $x_d^{(n)} \in \mathbb{C}$ is the transmitted symbol, and $n_d^{(n)} \in \mathbb{C}$ denotes the additive noise. $(\cdot)^H$ denotes the conjugate transpose. With the downlink channel vector $\mathbf{h}_d^{(n)}$, gNB can calculate the transmit beamformer $\mathbf{w}_T^{(n)}$. The uplink received signal of the $n-$th subcarrier is given by

$$y_u^{(n)} = \mathbf{w}_R^{(n)H} \mathbf{h}_u^{(n)} x_u^{(n)} + \mathbf{w}_R^{(n)H} \mathbf{n}_u^{(n)}, \tag{2}$$

where $\mathbf{w}_R^{(n)} \in \mathbb{C}^{N_b \times 1}$ denotes the receive beamformer, and subscript u denotes uplink signals and noise, similar to (1). The downlink and uplink CSI matrices in the spatial frequency domain are denoted as $\tilde{\mathbf{H}}_d = \left[\mathbf{h}_d^{(1)}, ..., \mathbf{h}_d^{(N_f)}\right]^H \in \mathbb{C}^{N_f \times N_b}$ and $\tilde{\mathbf{H}}_u = \left[\mathbf{h}_u^{(1)}, ..., \mathbf{h}_u^{(N_f)}\right]^H \in \mathbb{C}^{N_f \times N_b}$, respectively.

To reduce the feedback overhead, we first exploit the property that CSI matrices exhibit some sparsity in the delay domain since the delay among multiple paths lies in a particularly limited period [24]. The CSI matrix $\mathbf{H}_f$ in frequency domain can be transformed to be $\mathbf{H}_t$ in delay domain using an inverse discrete Fourier transform (IDFT), i.e.,

$$\mathbf{F}^H \mathbf{H}_f = \mathbf{H}_t, \tag{3}$$

where $\mathbf{F}$ and $\mathbf{F}^H$ denote the $N_f \times N_f$ unitary DFT matrix and IDFT matrix, respectively. After IDFT of (3), most elements in the $N_f \times N_b$ matrix $\mathbf{H}_t$ are near zero except for the first Q_f rows. Therefore, we truncate the channel matrix to the first Q_f rows that are with distinct non-zero values, and utilize $\mathbf{H}_d$ and $\mathbf{H}_u$ to denote the first Q_f rows of matrices after IDFT of $\tilde{\mathbf{H}}_d$ and $\tilde{\mathbf{H}}_u$, respectively.

Downlink CSI of massive MIMO can be better compressed and reconstructed by utilizing the autoencoder based CSI feedback, where the encoder and decoder are designed for CSI dimension compression and reconstruction, respectively. Furthermore, auxiliary information including previously constructed CSI within channel coherence time [11], [12] and uplink CSI at the gNB [10] can be incorporated into the CSI decoder network to further reduce the redundancy in downlink CSI reporting by the UEs and improve the reconstruction accuracy of the time-varying downlink CSI.

To efficiently compress the downlink CSI in massive MIMO systems while compatible with the previous autoencoder based CSI feedback works, we design a modular DNN framework CQNet that can jointly optimize the dimension compression, quantization and CSI decoding process. The recovery module at the transmitter is correspondingly optimized for CSI recovery. As shown in Fig. 1, our CSI feedback framework for FDD downlink channel reconstruction consists of 3 modules: a dimension compression module, a quantizer module, and a recovery module. By modeling the encoder and decoder of the works focusing on dimension compression as the dimension compression module and recovery module respectively, the CQNet framework can be integrated with corresponding works directly [9], [10], [11], [12], [13], [14], etc. For the recovery module, available CSI within coherence time and uplink CSI at the gNB can be exploited as the auxiliary input to help reconstruct the downlink CSI. We take the enhanced versions [13] of CsiNet in [9] and DualNet-MAG in [10] as examples to jointly optimize dimension compression and codeword quantization of downlink CSI feedback. These corresponding new architectures are respectively named as CsiQnet and DualQnet.

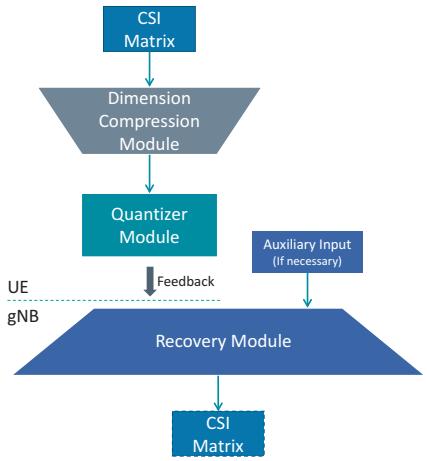


Fig. 1: CSI feedback framework.

We shall let $\hat{\mathbf{H}}_d$ denote the reconstructed downlink CSI matrix. We define the quantization function as $f_{\text{quan}}(\cdot)$. For CsiQnet, the dimension compression module, quantizer module, and recovery module can be denoted, respectively, by

$$\mathbf{s}_1 = f_{c,1}(\mathbf{H}_d), \tag{4}$$

$$\hat{\mathbf{s}}_1 = f_{\text{quan},1}(\mathbf{s}_1), \tag{5}$$

$$\hat{\mathbf{H}}_d = f_{r,1}(\hat{\mathbf{s}}_1). \tag{6}$$

For DualQnet, the dimension compression module, quantizer module, and recovery module can be denoted, respectively, by

$$\mathbf{s}_2 = f_{c,2}(\mathbf{H}_d), \tag{7}$$

$$\hat{\mathbf{s}}_2 = f_{\text{quan},2}(\mathbf{s}_2), \tag{8}$$

$$\hat{\mathbf{H}}_d = f_{r,2}(\hat{\mathbf{s}}_2, \mathbf{H}_u). \tag{9}$$

The optimization of downlink CSI compression and recovery method can be formulated as minimizing $\left\|\mathbf{H}_d - \hat{\mathbf{H}}_d\right\|^2$, where $\|\cdot\|$ denotes the Frobenius norm.

III. CSI Feedback in Low-Bit Quantization

The DL-based CSI feedback works including CsiNet in [9] and DualNet in [10] have demonstrated substantial performance gain in terms of downlink CSI feedback reduction and reconstruction accuracy. In [13], the performance of CsiNet and DualNet-MAG was enhanced by improving the network structures (named as CsiNet Pro and DualNet Pro) as well as the data preprocessing method. However, in addition to the benefit of downlink CSI compression, additional encoding of the compressed CSI feedback coefficients from the original float32 format can further reduce the downlink CSI feedback payload for massive MIMO systems drastically.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2020.2993626, IEEE Transactions on Communications

Toward this goal, we first evaluate the impact of codeword quantization bitwidth on CSI reconstruction accuracy by examining a simple uniform quantizer. We start from testing CsiNet Pro and DualNet Pro by simply adding a uniform quantizer between the encoder DNN and decoder DNN to assess the impact of quantization distortion.

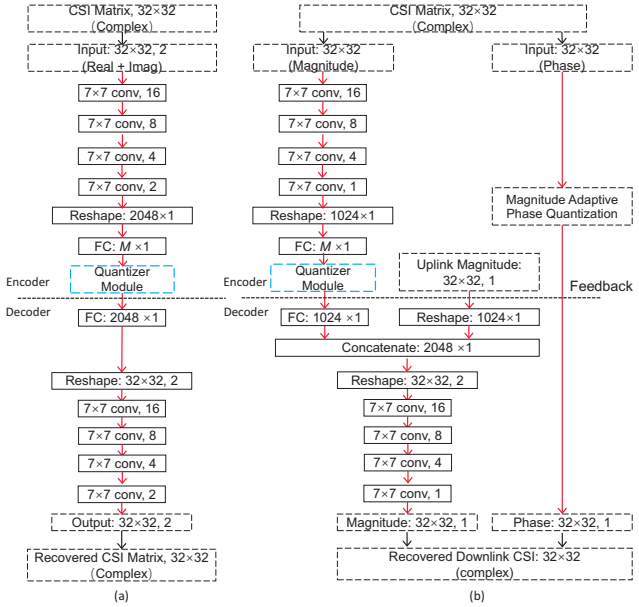
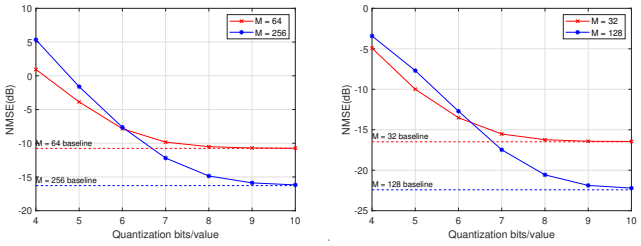


Fig. 2: Architectures of CsiNet Pro (a) and DualNet Pro (b). A quantizer module is inserted between the encoder DNN and decoder DNN.

As shown in Fig. 2 (a), CsiNet Pro utilizes an autoencoder architecture, where an encoder DNN acts as a dimension compression module and a corresponding decoder DNN is responsible for CSI reconstruction. Each complex CSI matrix is split into real and imaginary parts, rearranged into two feature maps of the input for the encoder. The encoder network contains four 7×7 convolutional (conv) layers with 16, 8, 4, 2 channels for feature extraction, and an M-unit fully connected (FC) layer for dimension compression. The decoder network consists of a fully connected layer for dimension decompression and four convolutional layers for CSI calibration. Compared with CsiNet, CsiNet Pro exploits more CNN kernels and layers in the encoder to extract the codewords that can contain more features of CSI matrices from the input.

DualNet Pro leverages the magnitude correlation between uplink and downlink to improve CSI feedback efficiency. As shown in Fig. 2 (b), DualNet Pro processes the magnitude and phase separately. After separation, the CSI magnitudes are sent to the encoder network including four 7×7 convolutional layers with 16, 8, 4, 1 channels and an M-unit fully connected layer for dimension compression. The gNB decoder uses the compressed codewords and the locally available uplink CSI magnitudes together to jointly decode downlink CSI. The received codewords are first mapped to their original length using a fully connected layer. The conjugation layer combines both downlink CSI and uplink CSI for the decoding. To save feedback bandwidth while limiting quantization error,



(a) CsiNet Pro (b) DualNet Pro

Fig. 3: CSI recovery accuracy at different quantization levels. M is the length of feedback vector after dimension compression.

a magnitude-adaptive phase quantization (MAPQ) is applied in which CSI coefficients with larger magnitude receive finer phase quantization, and vice versa.

Uniform quantization is simple and well known in practice. It is basically a rounding process, in which each sample value is rounded to the nearest value among a finite set of possible quantization levels. We can limit the amplitude of CSI coefficients between $[s_{\min}, s_{\max}]$. Let ℓ be the number of bits for amplitude quantization. Each CSI coefficient's amplitude can be uniformly quantized into 2^ℓ levels:

$$\hat{s} = \Delta \lfloor \frac{s}{\Delta} \rceil, \qquad \text{where} \quad \Delta = \frac{s_{\max} - s_{\min}}{2^\ell - 1}. \tag{10}$$

We include uniform quantization into the DL-based CSI feedback framework. Specifically, we first train CsiNet Pro and DualNet Pro without quantization in the initial end-to-end approach. Next, we apply uniform quantization on the compressed CSI coefficients, before sending them on the uplink to the decoder at the gNB. The $s_{\max}$ and $s_{\min}$, respectively, are the maximum and minimum settings using the pretrained model.

To evaluate the influence of codeword quantization on the CSI reconstruction accuracy, we use the COST 2100 channel model to generate CSI matrices [25]. A uniform linear array (ULA) of $N_b = 32$ transmit antennas is set up with half-wavelength spacing in an indoor environment with uplink and downlink bands at 5.1 GHz and 5.3 GHz, respectively. The bandwidth, N_f and Q_f are set to 20MHz, 1024 and 32, respectively. Fig. 3 shows the resulting NMSE for CsiNet Pro and DualNet Pro under different levels of quantization 2^ℓ as we vary ℓ from 4 to 10 bits. The lengths of vectorized input for CsiNet Pro and DualNet Pro before the dimension compression are 2048 and 1024, respectively. For CsiNet Pro, we set the length of compressed codeword vector to $M = 64$ and 256, respectively. For DualNet Pro, we set the length of compressed codeword vector to $M = 32$ and 128, respectively. Results from float32 serve as the baseline in CSI feedback.

A few observations can be made from the results of Fig. 3. First, high precision feedback of 32 bits per CSI value is unnecessary, as 9-bit uniform quantizer achieves nearly the same accuracy as float32 for both CsiNet Pro and DualNet Pro without retraining the neural networks for the quantized codewords. Second, both CsiNet Pro and DualNet Pro are more robust to quantization errors at lower compression. Finally, CSI reconstruction accuracy degrades with coarser

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2020.2993626, IEEE Transactions on Communications

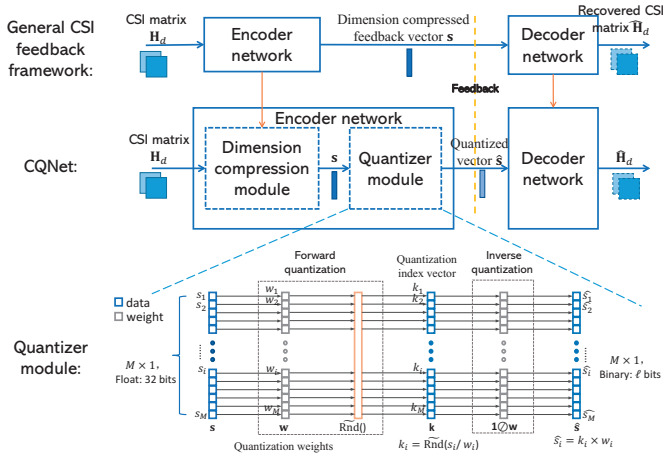


Fig. 4: Proposed CQNet. CQNet can optimize the codeword quantization of general CSI feedback directly by inserting a plug-in quantizer module which can customize element-wise quantization stepsizes for the feedback vector and improve quantization accuracy.

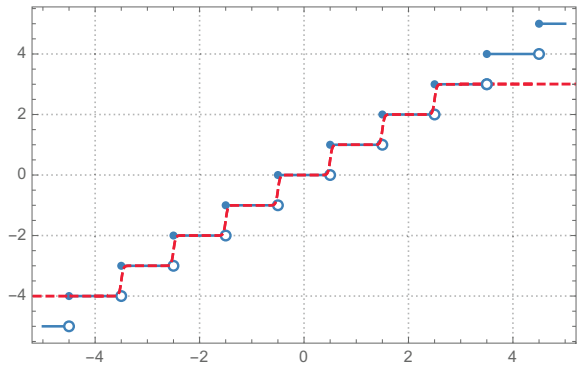


Fig. 5: Illustration of the approximate rounding function for 3-bit quantization: Blue line is the desired rounding function; Red line is the corresponding approximation function.

quantization. When ℓ drops below 7, the reconstruction experiences a clear degradation.

These test results motivate our study to design a more efficient and suitable quantization solution which can deliver high CSI reconstruction accuracy while using smaller ℓ to further improve the feedback bandwidth efficiency.

IV. CQNet

Optimum quantization depends on the distribution of data under quantization and the performance metric. For example, non-uniform quantizer such as the μ-law method can improve the signal-to-quantization noise ratio (SQNR) for lower power signals. Clearly, it is impractical to exhaustively test a huge number of encoding schemes for the downlink CSI coefficients generated by the encoder DNN in order to determine the best fit. Instead, we shall develop a novel DNN approach to learn and optimize the quantization intervals in order to improve the CSI recovery accuracy for the limited number of quantization levels. Furthermore, with the help of end-to-end optimization, the module for dimension compression and CSI recovery can also be refined.

A. Joint Compression and Quantization

We propose an end-to-end "CQNet", which can jointly optimize CSI dimension reduction, quantization and CSI reconstruction. As shown in Fig. 4, CQNet consists of an encoder DNN at the UE which includes a dimension compression module and a quantizer module, paired with a decoder network at the gNB. The compression module can adopt the encoder neural network of CsiNet Pro, DualNet Pro, or any other one for dimension compression. The decoder module can also be from a general CSI feedback framework. Dimension compression transforms the downlink CSI matrix $\mathbf{H}_d$ into a codeword vector $\mathbf{s}$ of dimension M which moves into the quantizer module. The quantizer module, parameterized by a trainable forward quantization weight vector $\mathbf{w}$, maps the unquantized $\mathbf{s}$ into an index vector $\mathbf{k}$ that corresponds to a quantized codeword $\hat{\mathbf{s}}$. The quantizer module can quantify the input vector from 32 bits (float32) to ℓ bits per compressed input value ($\ell \ll 32$). The gNB, upon reception of the quantized codeword on the uplink, can send the codeword to the decoder network.

To facilitate backpropagation during training, a soft quantizer can replace the non-differentiable quantizer function only during the training of the neural networks for the weights update. After training, a quantizer with strict rounding will determine the quantization index vector $\mathbf{k}$ and quantized codewords $\hat{\mathbf{s}}$ for uplink feedback. When the quantization bits/value is ℓ, the elements in $\mathbf{k}$ are denoted by $k_i \in \{-2^{\ell-1}, -2^{\ell-1}+1, ..., 2^{\ell-1}-1\}$. Since the corresponding elements in $\mathbf{k}$ and $\hat{\mathbf{s}}$ have a one-to-one mapping, either can be encoded for feedback. After receiving the quantized codeword on the uplink, the gNB will send it to the decoder network to recover the downlink CSI matrix $\hat{\mathbf{H}}_d$.

Different from previous works that share the same quantization parameters for all the input to the quantizer, CQNet customizes the quantization step length for each location of the dimension-compressed vector. We define the weights of the quantizer such at the quantization intervals as w_i for each element s_i in the compressed CSI vector $\mathbf{s}$. As illustrated in Fig. 4, the forward quantization stage can be implemented as a set of element-wise division filters which have only M parameters followed by a rounding function $\widetilde{\text{Rnd}}(\cdot)$. After forward quantization stage, an inverse quantization stage is added to reconstruct the approximated codeword $\hat{\mathbf{s}}$ using a vector $\mathbf{1} \oslash \mathbf{w}$ before feeding into the decoder module. Consequently, the quantizer module can be formulated as $\hat{\mathbf{s}} = \widetilde{\text{Rnd}}(\mathbf{s} \oslash \mathbf{w}) \odot \mathbf{w}$, where $\oslash$ and $\odot$ are Hadamard division and Hadamard product, respectively.

The rounding function can be viewed as an activation function of neurons acting on the outputs from element-wise multiplications. Since the ideal rounding function $\text{Rnd}(\cdot)$ has zero gradient almost everywhere and is otherwise non-differentiable, slow convergence may take place during back-propagation training. To mitigate this problem, we propose an

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2020.2993626, IEEE Transactions on Communications

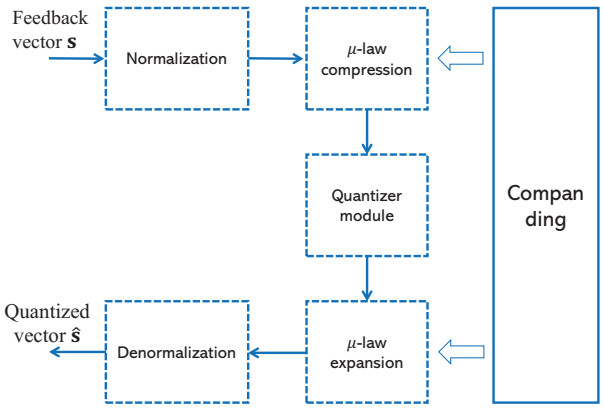


Fig. 6: Enhanced quantizer module using the μ-law compander. The quantizer module has been shown in Fig. 4.

approximate rounding function

$$\widetilde{\mathrm{Rnd}}(x,\ell,r)=\sum_{i=-2^{\ell-1}}^{2^{\ell-1}-1}\mathrm{sigmoid}[r(x-i-0.5)]-2^{\ell-1}, \quad (11)$$

which is differentiable and easy to implement. Note that $\widetilde{\mathrm{Rnd}}(\cdot)$ is a summation of sigmoid functions with parameters ℓ and r. Here, ℓ is the number of quantization bits for each element in feedback vector $\mathbf{s}$, whereas r controls the sigmoidal steepness. Fig. 5 presents the true and the approximate rounding functions using $\ell=3$ as an example. Bear in mind that the approximate $\widetilde{\mathrm{Rnd}}(\cdot)$ is only for training of the DNN weights. Upon the completion of training, the true rounding function is used.

Motivated by the work in [21] where μ-law quantization can achieve a better performance than uniform quantization after retraining the decoder network for the quantized codewords, we exploit the μ-law compander to formulate an enhanced quantizer module. The CQNet using the enhanced quantizer module is named as μ-CQNet. As shown in Fig. 6, the input is first normalized to the range $[-1,1]$ for μ-law companding. Then the μ-law compression and expansion can be done using the following equation and its inverse equation:

$$F(x)=\mathrm{sgn}(x)\frac{\ln(1+\mu\,|x|)}{\ln(1+\mu)}, x\in[-1,1]. \quad (12)$$

The output from the μ-law expansion is finally denomalized to the quantized vector $\hat{\mathbf{s}}$.

Then we take the CsiNet Pro and DualNet Pro as examples to show how to implement the CQNet framework with existing works that focused on the dimension compression of massive MIMO CSI matrices. CsiNet Pro [13] is an enhanced version of CsiNet [9]. The structure of CsiNet Pro has been shown in Fig. 2 (a). It utilizes the autoencoder architecture, where an encoder DNN acts as a compression module and a corresponding decoder DNN is responsible for CSI reconstruction. Each CSI matrix is split into real and imaginary parts, rearranged into two sets of encoder DNN input. The CQNet architecture that uses CsiNet Pro for CSI dimension reduction module and CSI recovery is named as CsiQnet. The dimension compression module is followed by the quantizer module shown in Fig. 4. The CsiQnet using the enhanced quantizer shown in Fig. 6 is named as μ-CsiQnet. The decoder network of CsiNet Pro is directly used in CsiQnet to decode the quantized codewords.

DualNet Pro [13] has been shown in Fig. 2 (b). It is an enhanced version of DualNet-MAG in [10]. DualNet Pro exploited the uplink-downlink CSI correlation to improve the CSI recovery accuracy. The bi-directional channel correlation between CSI magnitudes helps the decoder recover the downlink CSI magnitudes with better accuracy by leveraging the low-rate feedback codewords and locally available uplink CSI magnitudes. The CQNet architecture that uses DualNet Pro for CSI dimension reduction and CSI recovery is named as DualQnet. DualQnet uses the DualNet Pro's encoder network for dimension compression at the UE, and uses the DualNet Pro's decoder network at the gNB to recover the CSI. Quantizer module is inserted between the dimension compression module and decoder network. The DualQnet using the enhanced quantizer shown in Fig. 6 is named as μ-DualQnet.

We define the loss function of CsiQnet or DualQnet:

$$L(\hat{\mathbf{H}}_d,\mathbf{w})=L_m(\hat{\mathbf{H}}_d,\mathbf{H}_d)+\lambda L_{quan}(\mathbf{w}), \quad (13)$$

as a combination of mean square error (MSE) loss L_m and a quantization loss L_{quan}. L_{quan} is used for the regularization that accounts for quantization efficiency and convergence. One simple function for this purpose is $L_{quan}(\mathbf{w})=1/\left\|\mathbf{w}\right\|$. The training objective is to find the encoding and decoding parameters which can achieve the optimum CSI reconstruction accuracy given the specific quantization bits per value ℓ. Although adjusting λ can help limit the bandwidth after entropy encoding, we mainly focus on adjusting the ℓ in this paper to control the bandwidth since the entropy encoding highly relies on the input data distribution.

Through training based on a large MIMO CSI data set generated by using well known practical channel models such as COST 2100 model [25], CsiQnet and DualQnet can converge to optimized settings. During live downlink CSI feedback, both CsiQnet and DualQnet can generate more efficiently quantized codewords $\hat{\mathbf{s}}_i$ which can significantly improve the accuracy of CSI reconstruction at fixed bandwidth or bitwidth. CsiQnet and DualQnet enable more effective bandwidth usage with little CSI reconstruction accuracy loss. We also demonstrate that with the help of μ-law compander, μ-CsiQnet and μ-DualQnet can have better robustness than others.

To train the model, normalization is applied in both downlink and uplink CSI matrices. Adam optimizer is adopted to update the DL network parameters. The initial learning rate is set to 0.001. To accelerate the convergence speed of the training, we utilize the weights trained in CsiNet Pro and DualNet Pro to initialize the dimension compression modules and decoder networks of CsiQnet and DualQnet, respectively. Notice that DualQnet optimizes the magnitude feedback during joint training to minimize (13). For the separated phase feedback of compressed CSI, we shall design another MAPQ DNN in Section IV-B.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2020.2993626, IEEE Transactions on Communications

B. DL-based Phase Quantization

DualNet Pro utilizes the magnitude correlation of bi-directional CSIs in radial coordinate to reduce the amount of feedback for CSI magnitudes and improve CSI feedback efficiency. However, the weak phase correlation between uplink/downlink CSI requires the UE to efficiently quantize and encode all downlink CSI phases for feedback. However, it is well known that uniform phase quantization results in unnecessarily fine quantization at low magnitude and coarse quantization at high magnitude. Therefore, bandwidth efficiency phase quantization is an important issue to tackle in DualQnet.

Let $\mathbf{M} = [\mathbf{M}_{i,j}]$ be the magnitudes of CSI matrix, $\mathbf{P} = [\mathbf{P}_{i,j}]$ be the phases of CSI matrix, $\hat{\mathbf{M}}$ be the recovered magnitudes of CSI matrix, and $\hat{\mathbf{P}}$ be the recovered phases of CSI matrix, respectively. We can write a matrix $e^{j\mathbf{P}}$ whose elements are $e^{j\mathbf{P}_{i,j}}$. The optimization objective of DualNet Pro is to minimize the MSE of recovered CSI matrices, i.e., $\mathbf{E}\{\left\|\mathbf{H}_d - \hat{\mathbf{H}}_d\right\|^2\} = \mathbf{E}\{\left\|\mathbf{M} \odot e^{j\mathbf{P}} - \hat{\mathbf{M}} \odot e^{j\hat{\mathbf{P}}}\right\|^2\}$, where $\odot$ denotes Hadamard matrix product. It is challenging to solve this problem since the recovered magnitude and phase influence the MSE jointly. To reduce the complexity of this problem, we consider the upper bound:

$$\begin{aligned}
&\left\|\mathbf{M} \odot e^{j\mathbf{P}} - \hat{\mathbf{M}} \odot e^{j\hat{\mathbf{P}}}\right\|^2 \\
&= \left\|\mathbf{M} \odot e^{j\mathbf{P}} - \mathbf{M} \odot e^{j\hat{\mathbf{P}}} + \mathbf{M} \odot e^{j\hat{\mathbf{P}}} - \hat{\mathbf{M}} \odot e^{j\hat{\mathbf{P}}}\right\|^2 \\
&\leqslant 2\left\|\mathbf{M} \odot e^{j\mathbf{P}} - \mathbf{M} \odot e^{j\hat{\mathbf{P}}}\right\|^2 + 2\left\|\mathbf{M} \odot e^{j\hat{\mathbf{P}}} - \hat{\mathbf{M}} \odot e^{j\hat{\mathbf{P}}}\right\|^2 \\
&= 2\left\|\mathbf{M} \odot (e^{j\mathbf{P}} - e^{j\hat{\mathbf{P}}})\right\|^2 + 2\left\|\mathbf{M} - \hat{\mathbf{M}}\right\|^2.
\end{aligned} \tag{14}$$

Consequently, the optimization goal can be relaxed to minimize

$$\mathbf{E}_{(\mathbf{M},\mathbf{P})}\left(\left\|\mathbf{M} \odot (e^{j\mathbf{P}} - e^{j\hat{\mathbf{P}}})\right\|^2\right) + \mathbf{E}_{\mathbf{M}}\left(\left\|\mathbf{M} - \hat{\mathbf{M}}\right\|^2\right).$$

DualQnet can minimize $\mathbf{E}_{\mathbf{M}}(\|\mathbf{M} - \hat{\mathbf{M}}\|^2)$. It is clear that the first part $\mathbf{E}_{(\mathbf{M},\mathbf{P})}(\|\mathbf{M} \odot (e^{j\mathbf{P}} - e^{j\hat{\mathbf{P}}})\|^2)$ represents phase quantization error amplified by the corresponding magnitude.

Therefore, to further reduce feedback bandwidth, CSI quantization error can be kept small by applying the magnitude-adaptive phase quantization (MAPQ) principle in which CSI coefficients with larger magnitude adopt finer phase quantization, and vice versa. After recovering the magnitude, gNB can restore the quantified phase based on MAPQ. Such MAPQ can keep the quantization error close for a range of magnitudes. Since quantization bits of phase vary with the magnitude, the expectation of quantization bits depends on the distribution of CSI magnitude. Thus, we need to allocate the number of phase quantization bits based on the distribution of CSI magnitude under limited average bitwidth. Such a problem typically become a mixed integer nonlinear programming problem as described in [26], which is NP-hard.

A heuristic quantization bit allocation solution was provided in [10] based on examining the data set. In this method, the cumulative distribution function (CDF) of CSI magnitudes is estimated to determine magnitude values corresponding to CDF value of 0.5, 0.7, 0.8, and 0.9, respectively. These four points divide the CSI magnitudes into five ordered segments from low to high. Accordingly, 3, 4, 5, 6, 7 phase quantization bits are allocated, respectively, to encode the CSI phases. This set of MAPQ codewords given in [10] can generate codewords of the mean length of 4.1 bits/value to achieve the same MSE as that obtained using a 6-bit uniform quantizer.

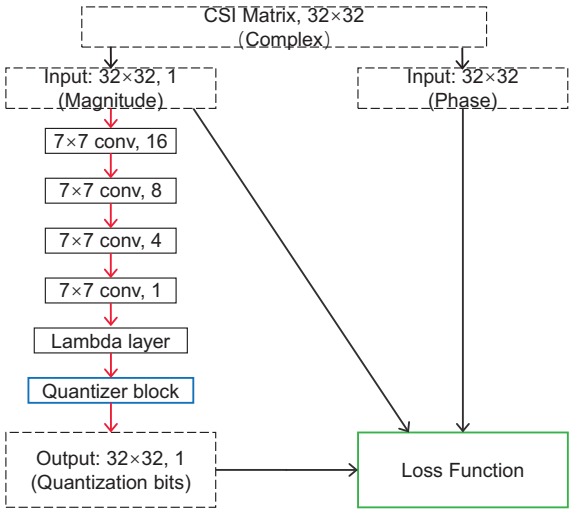


Fig. 7: Illustration of the magnitude-adaptive phase quantization.

However, the heuristic method of [10] is inflexible with respect to the segments. Given different mean bitwidth constraints, we have to determine different allocation based on heuristics. Thus, we propose a more flexible and general design method in this work.

We propose an innovative DNN to solve the bit allocation problem. This new framework PhaseQuan can optimize the allocation of MAPQ quantization bits for phase based on unsupervised learning. As shown in Fig. 7, PhaseQuan utilizes the magnitudes as the input to DNN and includes CSI phases, magnitudes and corresponding quantization bits in its loss function. The magnitude input passes through three 7×7 convolutional (conv) layers with 16, 8 and 4 channels, which explore the potential spatial correlation within the CSI matrix. The ensuing convolutional layer, lambda layer, and quantizer block are used to project the output of the last hidden layer to quantization bits. The convolutional layer first utilizes the "sigmoid" activation to project the input within $(0, 1)$. The lambda layer is defined as $log_2(\frac{1}{x+\epsilon})$, in which $\epsilon > 0$ is a small value to ensure non-zero denominator. Here, $\frac{1}{x+\epsilon}$ corresponds to the number of quantization intervals. Logarithm $log_2(\cdot)$ and the quantizer module can map the number of quantization intervals to the number of quantization bits.

To minimize $\mathbf{E}_{(\mathbf{M},\mathbf{P})}(\|\mathbf{M} \odot (e^{j\mathbf{P}} - e^{j\hat{\mathbf{P}}})\|^2)$ within the constraint on the quantization bitwidth, we adopt the entropy of phase as the optimization regularizer since more quantization bits lead to higher phase entropy. Thus, we propose a loss function

$$L(\mathbf{M}, \mathbf{P}, \mathbf{Y}) = L_m(\mathbf{M}, \mathbf{P}, \mathbf{Y}) + \lambda L_y(\mathbf{Y}), \tag{15}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2020.2993626, IEEE Transactions on Communications

where $\mathbf{Y}$ is a matrix of integer elements representing the number of quantization bits for the corresponding CSI matrix element, optimized by the PhaseQuan. The phase quantization error evaluation function is set to be

$$L_m(\mathbf{M},\mathbf{P},\mathbf{Y}) = \mathbf{E}_{(\mathbf{M},\mathbf{P})}\left(\left\|\mathbf{M}\odot e^{j(\hat{\mathbf{P}}-\mathbf{P})}\right\|^2\right),$$

in which $\hat{\mathbf{P}}$ is the quantized phase matrix. Additionally, define an entropy $H(\check{\mathbf{P}}_{i,j})$ for quantized phase $\check{\mathbf{P}}_{i,j}$ that corresponds to magnitude $\mathbf{M}_{i,j}$. In this paper, we assume the phase to be uniformly distributed over 2π [27]. Consequently, $H(\check{\mathbf{P}}_{i,j}) = \mathbf{Y}_{i,j}$. As a result, we use

$$L_y(\mathbf{Y}) = \mathbf{E}_{(\mathbf{M},\mathbf{P})}\left(\frac{1}{Q_f \times N_b}\sum_{i,j} H(\check{\mathbf{P}}_{i,j})\right)$$

as an entropy regularizer to reduce the number of quantization bits in phase feedback. Adjustable parameter λ value governs the trade-off between the quantization bits of phase and the reconstruction loss.

V. Performance Evaluation

A. Experiment Setup

We use the industry-grade COST 2100 model [25] to generate massive MIMO channels for both training and testing of our DNN architecture. The training set size is $70,000$ and testing set size is $30,000$. The values of epoch and batch size are set to 1000 and 200, respectively. We test two scenarios:

(a) indoor channels with 5.1 GHz uplink band and 5.3 GHz downlink center frequency.
(b) semi-urban outdoor channels with 850 MHz uplink band and 930 MHz downlink center frequency.

Uplink and downlink bandwidths of 20 MHz are selected for the indoor and outdoor scenarios.

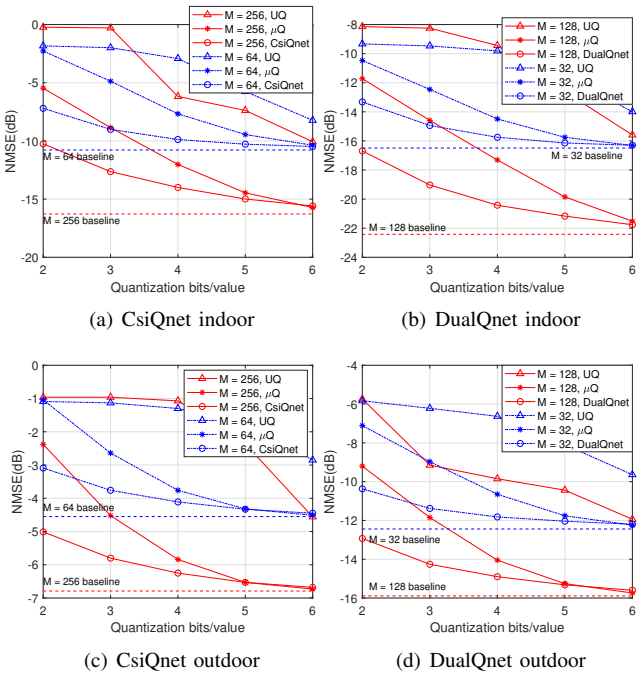
We place gNB at the center of a square area of lengths 20m for indoor coverage and 400m for outdoor coverage, respectively. We randomly position UEs within the coverage area. The gNB uses ULA with $N_b = 32$ antennas and $N_f = 1024$ subcarriers. After transforming the channel matrix $\mathbf{H}_f$ into the delay domain $\mathbf{H}_t$, only the first 32 rows are kept for feedback reporting due to sparsity. To evaluate the accuracy of CSI recovery, we use normalized MSE

$$\text{NMSE} = \frac{1}{n}\sum_{k=1}^{n}\|\mathbf{H}_d^k - \hat{\mathbf{H}}_d^k\|^2/\|\mathbf{H}_d^k\|^2, \qquad (16)$$

where k and n are the index and total number of samples in the testing set, respectively.

We compare CsiQnet and DualQnet with CsiNet Pro and DualNet Pro using other quantization methods, including:

- Uniform quantization (UQ) used in [20]. The decoder network is retrained after codeword quantization for fair comparison.
- μ-law non-uniform quantization (μQ) used in [21]. The decoder network is retrained after codeword quantization.
- JCnet in [23]. JCnet projects codewords into $(-1, 1)$ using the 'tanh' function, before exploiting UQ together by modifying the quantization gradient to 1 during back-propagation .



(a) CsiQnet indoor (b) DualQnet indoor

(c) CsiQnet outdoor (d) DualQnet outdoor

Fig. 8: CSI recovery comparison at different quantization levels among CQNet, UQ and μQ.

The performances of μ-CsiQnet and μ-DualQnet are also compared. We use the pretrained weights of CsiNet Pro and DualNet Pro for the initialization of dimension compression module and decoder network for all quantization methods. For the μ-law quantization, we use $\mu = 255$ in the companding function.

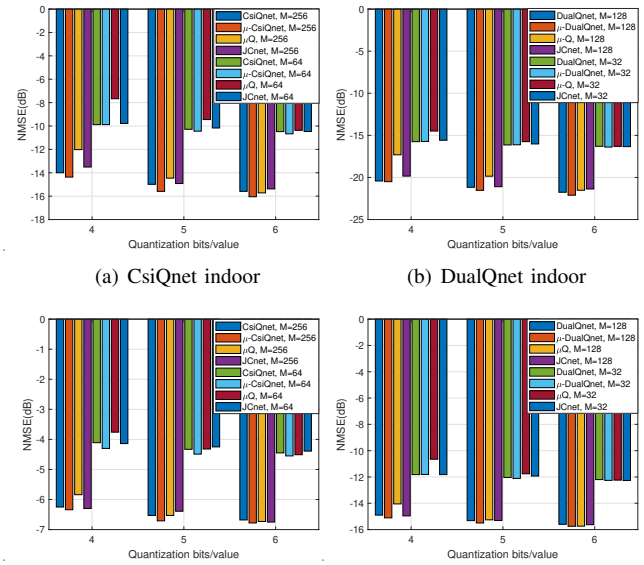
For the CsiQnet and CsiNet Pro, CSI matrix is divided into two (real and imaginary) channels as the input to the DNN. We compare the CSI reconstruction performance under the compressed dimension $M = 64$ and 256. For DualQnet and DualNet Pro, we compare the CSI reconstruction performance under the compressed dimension $M = 32$ and 128.

B. CSI Reconstruction Performance Evaluation

To provide clear demonstrations, we first compare the CSI reconstruction accuracy achieved by CsiQnet and DualQnet with the CSI accuracy of CsiNet Pro and DualNet Pro using UQ and μQ under five different bit-widths from 2 to 6. Single precision float32 result is given for a baseline comparison.

Fig. 8 compares the NMSE performance of our proposed CsiQnet and DualQnet with the corresponding networks using UQ and μQ. As shown in Fig. 8, CsiQnet and DualQnet both outperform corresponding networks using UQ in all bit-widths and μQ when the bit-widths are less than 6. Importantly, the accuracy gap grows with decreasing number of quantization bits/value. This result demonstrates that our end-to-end DNN framework can jointly integrate CSI compression and quantization with reconstruction to achieve performance superior to the approach of combining individually optimized modules.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2020.2993626, IEEE Transactions on Communications



(a) CsiQnet indoor (b) DualQnet indoor

(c) CsiQnet outdoor (d) DualQnet outdoor

Fig. 9: CSI recovery comparison at different quantization levels among CQNet, μ-CQNet, μQ and JCnet.

Fig. 8 also shows that non-uniform μQ generally delivers better performance than UQ.

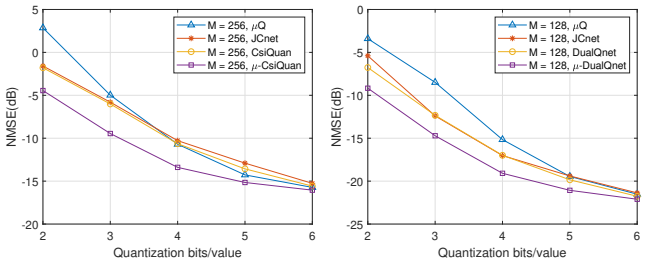
From the results in Fig. 8, we observe more robust performance when the compressed dimension M is relatively small. The NMSE degrades faster when M is large. The possible reason is that the lower dimension compression relies on principal components, while high-accuracy reconstruction further requires more detailed information of compressed vectors. Thus, a smaller feedback error can lead to a larger degradation in reconstruction accuracy when the compression is low.

We next examine the performance comparison of the CSI reconstruction accuracy among CQNet, μ-CQNet, μQ and JCnet under bitwidths of 4, 5, and 6 in Fig. 9. As shown in Fig. 9, μ-CsiQnet, and μ-DualQnet can achieve better performance than other schemes. Furthermore, the results from CsiQnet and DualQnet are marginally better than or comparable to JCnet. Typically, μQ in CsiNet Pro and DualNet Pro achieves lower CSI reconstruction accuracy than others.

C. Robustness Evaluation

To further test the robustness of CsiQnet and DualQnet, we consider the case when there is only one neural network trained for the codeword quantization and reconstruction but different bitwidths may be required due to the bandwidth changes. We select 6 quantization bits/value as an example to evaluate the robustness in the indoor case. Since μQ outperforms UQ clearly when quantization bits/value is 6, we evaluate the robustness of μQ, JCnet, CQNet and μ-CQNet without UQ for CsiNet Pro and DualNet Pro.

Considering a DNN previously trained for bitwidth of 6. If the UE is not assigned sufficient uplink bandwidth to transmit 6-bit codewords, we need to shorten the codewords into 2, 3, 4, or 5 bit codewords. Fig. 10 provides the resulting performance



(a) CsiQnet (b) DualQnet

Fig. 10: Robustness evaluation using the networks trained in 6 quantization bits/value CSI at fewer quantization bits/value cases.

of CsiQnet and DualQnet under different feedback bandwidths. As shown in Fig. 10, the performance of μ-CsiQnet and μ-DualQnet clearly outperforms CsiNet and DualNet-MAG using μQ, CQNet, and JCnet. Consequently, μ-CQNet framework can better encode CSI feedback in bandwidth limited scenarios.

D. Quantizer Optimization

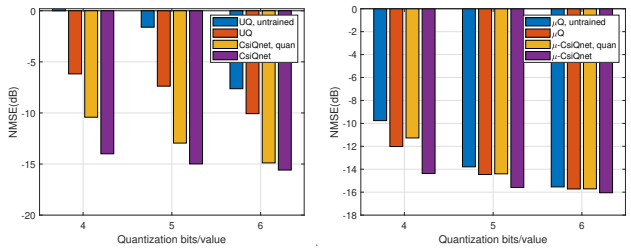
CQNet can achieve good CSI construction accuracy under the low-bit quantization through end-to-end optimization. Since CQNet consists of three modules including dimension compression module, quantizer module and recovery module, it is valuable to understand where the performance gain is from. For example, much training and storage overhead can be saved if the quantizer optimization can provide most of performance gains, since our quantizer only requires M parameters to be learned.

In this subsection, we try to figure out how much benefit can be provided by just optimizing the quantizer module or the recovery module. We compare four schemes for CsiNet Pro and DualNet Pro in the indoor cases to check the performance gains:

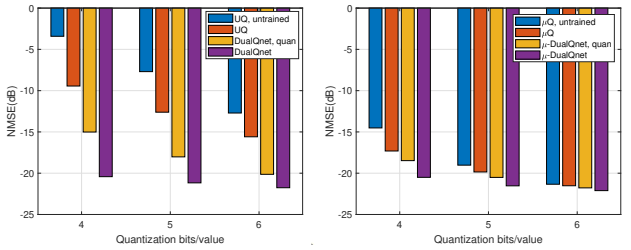
- UQ/μQ, untrained. After the quantization of the codeword, this scheme does not retrain the decoder network, so it provides the lower bound for performance optimization.
- UQ/μQ. This scheme retrains the decoder after the codeword quantization, thus corresponds to the gains from the optimization of the recovery module.
- CQNet/μ-CQNet, quan. This scheme only trains the parameters within the quantizer, thus corresponds to the gains from the optimization of the quantizer module. Note that, only M parameters are required to be learned.
- CQNet/μ-CQNet. The performance of this scheme is the upper bound through end-to-end optimization.

Fig. 11 shows the NMSE comparison of the above four schemes. The schemes using the μ-law compander are separated from the linear ones. As shown in Fig. 11, UQ/μQ without retraining the decoder has the worst performance. Furthermore, "CQNet/μ-CQNet, quan" generally outperforms CsiNet Pro and DualNet Pro using UQ and μQ with retrained decoders obviously, which means that more gains can be

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2020.2993626, IEEE Transactions on Communications



(a) CsiQnet, $M = 256$ (b) μ-CsiQnet, $M = 256$



(c) DualQnet, $M = 128$ (d) μ-DualQnet, $M = 128$

Fig. 11: CSI recovery comparison when optimizing the different module of the CSI feedback framework.

achieved by optimizing the quantizer than optimizing the decoder. It also illustrates that our designed quantizers are efficient by adjusting the element-wise quantization stepsize for the elements within a feedback vector.

E. Quantization Evaluation

In this subsection, we analyze the performance of our quantizer module, and try to go inside the DL networks to explore why CsiQnet and DualQnet outperform the UQ and μQ methods.

We first compare the performance of our approximate rounding function with the true rounding function during the testing phase. We show the MSE between the outputs of $\widetilde{\mathrm{Rnd}}(\cdot)$ and exact rounding in Numpy for various numbers of quantization levels. As shown in Table I, MSE generally is near -24 dB, negligible when compared with the rounded integers. We also compared the NMSE of CSI reconstruction when using $\widetilde{\mathrm{Rnd}}(\cdot)$ versus $\mathrm{Rnd}(\cdot)$. The differences are well below -0.15 dB.

It would be helpful for us to examine the effect of quantization error empirically to understand the performance of

TABLE I: MSE (dB) of the approximate rounding function

		Indoor		Outdoor	
	Bits	M = 64	M = 256	M = 64	M = 256
CsiQnet	2	-25	-25	-24	-24
	3	-24	-24	-24	-24
	4	-24	-24	-24	-24
	5	-24	-24	-24	-24
	6	-24	-24	-24	-24
	Bits	M = 32	M = 128	M = 32	M = 128
DualQnet	2	-25	-25	-25	-24
	3	-24	-24	-24	-24
	4	-24	-24	-24	-24
	5	-24	-24	-24	-24
	6	-24	-24	-24	-24

TABLE II: NMSQE (dB) of UQ, CsiQnet, μQ, and μ-CsiQnet.

Codeword dimension	Quantization bits/value	UQ	CsiQnet	μQ	μ-CsiQnet
M=64	5	-10.09	-23.36	-19.55	-21.66
	6	-16.22	-26.44	-25.69	-26.93
M=256	5	-3.89	-20.30	-19.20	-21.68
	6	-10.48	-24.36	-25.40	-26.96

TABLE III: The average entropy of quantized codeword (bits).

Codeword dimension	Quantization bits/value	UQ	CsiQnet	μQ	μ-CsiQnet
M=64	5	2.10	4.24	4.35	4.41
	6	3.06	4.77	5.37	5.38
M=256	5	1.38	3.86	4.20	4.38
	6	2.17	4.52	5.22	5.44

different CSI feedback methods under study. We measure the normalized mean square quantization error (NMSQE) defined as $\mathbf{E}[\frac{\|\mathbf{s}-\hat{\mathbf{s}}\|^2}{\|\mathbf{s}\|^2}]$. Consider 5- and 6-bit quantization per feedback value, respectively. We select CsiNet Pro with $M = 64$ and 256 in the indoor scenario when comparing the quantization errors of UQ, CsiQnet, μQ, and μ-CsiQnet. As shown in Table II, CsiQnet and μ-CsiQnet achieve lower NMSQE than UQ and μQ. The comparison clearly presents one reason on why CQNet and μ-CQNet can provide better CSI recovery accuracy.

We also evaluate the average entropy of quantized codeword from UQ, CsiQnet, μQ, and μ-CsiQnet. From the information theoretic perspective, higher entropy can lead to higher CSI recovery accuracy. As shown in Table III, CsiQnet and μ-CsiQnet deliver higher entropy, respectively, than UQ and μQ. This result provides an information theoretic interpretation on the advantages of CQNet and μ-CQNet. Furthermore, we observe that for 5-bit quantization, CsiQnet achieves better performance than μQ even with a smaller entropy. This result suggests that CQNet packs more useful information in its quantization codewords than μQ does.

Once the CSI feedback is quantized, entropy encoding can be exploited to further compress the CSI feedback [22]. Compared with estimating the entropy of quantized codewords in

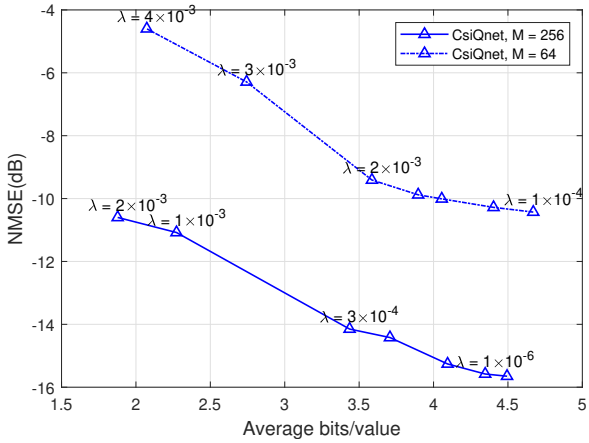
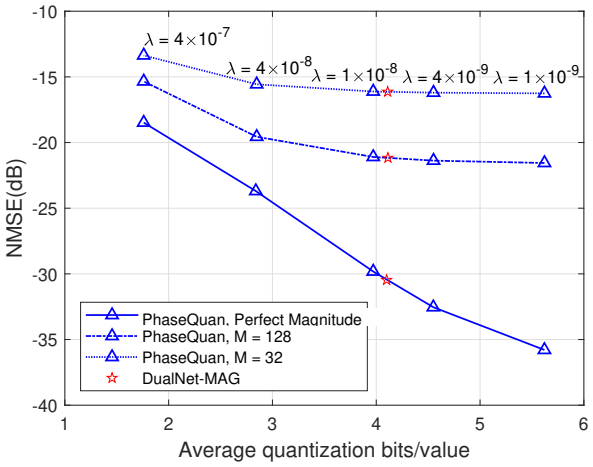
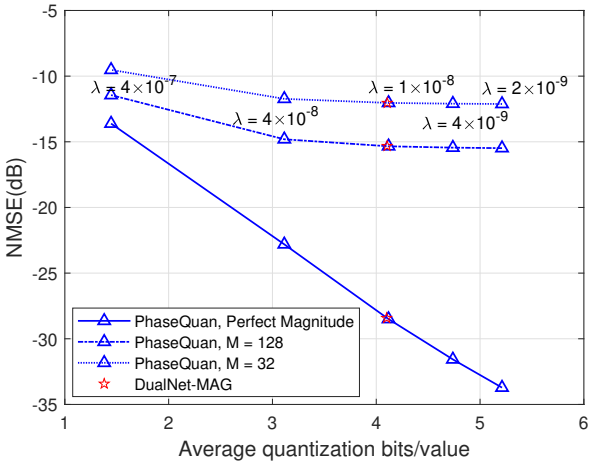


Fig. 12: Bitwidth adjustment in entropy encoding by controlling the quantization stepsize.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2020.2993626, IEEE Transactions on Communications



(a) Indoor



(b) Outdoor

Fig. 13: Quantization bits-NMSE trade-off under different λ.

the training process for backpropagation , our quantizer module enables a more efficient way in much lower complexity by controlling the quantization weight vector $\mathbf{w}$. We show an example of controlling the entropy of the quantized codewords by adjusting the λ in equation (13). We try the arithmetic coding [28], which is a simple and common entropy encoding to encode the quantized CSI coefficients. We consider CsiQnet with 6 bits per CSI value as examples. The NMSE and average required numbers of bits per CSI value after entropy encoding are shown in Fig. 12. As shown in Fig. 12, by adjusting the λ, average required numbers of bits per CSI value can be easily adjusted.

F. Phase Quantization

In order to flexibly allocate the phase quantization bits under different bandwidth limitation and reconstruction accuracy requirement, we train PhaseQuan under different λ values and illustrate the effect of λ on average phase quantization bits and reconstruction accuracy. Intuitively, smaller λ leads to higher reconstruction accuracy, though at a cost of higher feedback overhead. The converse also holds.

The bitwidth-NMSE trade-off under different λ is shown in Fig. 13. We use the phase quantization method in DualNet-MAG as the baseline, and evaluate 3 cases of magnitude knowledge for CSI reconstruction. We consider (a) perfect CSI magnitude; (b) DualQnet after dimension compression using 5 quantization bits per value with $M = 128$, and (c) DualQnet using 5 quantization bits per value with $M = 32$. As shown in Fig. 13, the performances of PhaseQuan are comparable to the baseline with better flexibility by adjusting λ. NMSE decreases with increasing quantization bitwidth. With a large λ value, the DNN tries to reduce the number of quantization bits, which in turn degrades NMSE. To find the suitable λ for a given NMSE, we can first select several candidate values of λ to train the PhaseQuan as reference anchors. By interpolating λ according to user's requirements in terms of CSI reconstruction accuracy and the available feedback bandwidth, we can obtain the phase quantization parameters under these constraints.

On the other hand, with lower accuracy in magnitude, the influence of quantization bits becomes weaker. This means that we can save bits in phase quantization according to the magnitude accuracy. For example, in the indoor cases when $M = 32$, phase quantization using 2.8 bits/value and 4.1 bits/value can generate similar NMSE. In the outdoor cases when $M = 32$, phase quantization using 3.1 bits/value and 4.1 bits/value can generate similar NMSE. In future works, we should jointly optimize the compression and encoding of magnitude and phase.

VI. Conclusions

Recent successes of DL in achieving more efficient CSI feedback for massive MIMO systems in FDD deployment strongly motivate further investigations of bandwidth-efficient encoding of the compressed CSI coefficients. In this paper, we propose a novel and end-to-end DL-based CSI feedback framework CQNet to jointly optimize dimension compression, codeword quantization, and CSI recovery for massive MIMO wireless transceivers. We integrate CQNet with two DL-based CSI feedback mechanisms, and demonstrate clear savings of feedback bandwidth and improved CSI reconstruction accuracy over massive MIMO wireless links. CQNet offers superior performance using few quantization bits with little loss of CSI reconstruction accuracy. We further present a DL-based CSI phase encoder for the CSI feedback framework (DualNet Pro) that exploits bi-directional channel correlation and improve the flexibility to manage the trade-off between CSI reconstruction accuracy and feedback bandwidth.

References

[1] 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, February 2014.

[2] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan 2013.

[3] 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. 3261–3271, June 2014.

[4] Z. Lv and Y. Li, "A channel state information feedback algorithm for massive MIMO systems," *IEEE Commun. Lett*, vol. 20, no. 7, pp. 1461–1464, July 2016.

[5] Z. Gao, L. Dai, Z. Wang, and S. Chen, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6169–6183, Dec 2015.

[6] H. Son and Y. Cho, "Analysis of compressed CSI feedback in MISO systems," *IEEE Wireless Commun. Lett*, vol. 8, no. 6, pp. 1671–1674, Dec 2019.

[7] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[8] H. He, C. Wen, S. Jin, and G. Y. Li, "Deep learning-based channel estimation for beamspace mmwave massive MIMO systems," *IEEE Wireless Commun. Lett*, vol. 7, no. 5, pp. 852–855, Oct 2018.

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

[10] 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, June 2019.

[11] T. Wang, C. 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, April 2019.

[12] C. Lu, W. Xu, H. Shen, J. Zhu, and K. Wang, "MIMO channel information feedback using deep recurrent network," *IEEE Commun. Lett*, vol. 23, no. 1, pp. 188–191, Jan 2019.

[13] Z. L. L. Z. Z. Liu, M. del Rosario and Z. Ding, "Spherical normalization for learned compressive feedback in massive MIMO CSI acquisition," *accepted by 2020 IEEE ICC Workshops*, 2020.

[14] Z. Lu, J. Wang, and J. Song, "Multi-resolution CSI feedback with deep learning in massive MIMO system," *arXiv preprint arXiv:1910.14322*, 2019.

[15] L. Theis, W. Shi, A. Cunningham, and F. Huszár, "Lossy image compression with compressive autoencoders," *International Conference on Learning Representations (ICLR) 2017*, April 2017.

[16] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[17] "IEEE standard for floating-point arithmetic," *IEEE Std 754-2008*, pp. 1–70, Aug 2008.

[18] J. Johnson, "Rethinking floating point for deep learning," *arXiv preprint arXiv:1811.01721*, 2018.

[19] N. Wang, J. Choi, D. Brand, C. Chen, and K. Gopalakrishnan, "Training deep neural networks with 8-bit floating point numbers," in *Advances in neural information processing systems*, 2018, pp. 7675–7684.

[20] Y. Jang, G. Kong, M. Jung, S. Choi, and I. Kim, "Deep autoencoder based CSI feedback with feedback errors and feedback delay in FDD massive MIMO systems," *IEEE Wireless Commun. Lett*, vol. 8, no. 3, pp. 833–836, June 2019.

[21] 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 Commun.*, pp. 1–1, 2020.

[22] Q. Yang, M. B. Mashhadi, and D. Gündüz, "Deep convolutional compression for massive MIMO CSI feedback," in *2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP)*, Oct 2019, pp. 1–6.

[23] C. Lu, W. Xu, S. Jin, and K. Wang, "Bit-level optimized neural network for multi-antenna channel quantization," *IEEE Wireless Commun. Lett*, pp. 1–1, 2019.

[24] R. H. Jr and A. Lozano, *Foundations of MIMO communication*. Cambridge University Press, 2018.

[25] L. Liu, C. Oestges, J. Poutanen, K. Haneda, P. Vainikainen, F. Quitin, F. Tufvesson, and P. D. Doncker, "The COST 2100 MIMO channel model," *IEEE Wireless Commun.*, vol. 19, no. 6, pp. 92–99, December 2012.

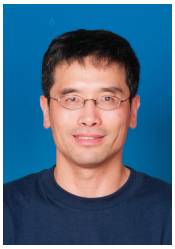
[26] P. Belotti, C. Kirches, S. Leyffer, J. Linderoth, J. Luedtke, and A. Mahajan, "Mixed-integer nonlinear optimization," *Acta Numerica*, vol. 22, pp. 1–131, 2013.

[27] P. . Kildal and K. Rosengren, "Correlation and capacity of MIMO systems and mutual coupling, radiation efficiency, and diversity gain of their antennas: simulations and measurements in a reverberation chamber," *IEEE Commun. Mag*, vol. 42, no. 12, pp. 104–112, Dec 2004.

[28] I. Witten, R. Neal, and J. Cleary, "Arithmetic coding for data compression," *Communications of the ACM*, vol. 30, no. 6, pp. 520–540, 1987.

Zhenyu Liu (S'18) received his B.S. degree from the Henan University in 2014. After that, he joined the School of Information and Communication Engineering in Beijing University of Posts and Telecommunications to pursue the Ph.D. degree. From 2017 to 2019, he was a visiting Ph.D. student with the Department of Electrical and Computer Engineering, University of California at Davis. His research interests include deep learning for wireless communications and vehicular communications.

Lin Zhang (M'09) LIN ZHANG received the B.S. and Ph.D. degrees from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 1996 and 2001, respectively. He is the Dean of School of Information and Communication Engineering, BUPT, Beijing, China. He was a Postdoctoral Researcher with the Information and Communications University, Daejeon, Korea, from December 2000 to December 2002. He went to Singapore and held a Research Fellow position with Nanyang Technological University, Singapore, from January 2003 to June 2004. He joined BUPT in 2004 as a Lecturer, then an Associate Professor in 2005, and a Professor in 2011. He served the university as the Director of Faculty Development Center and the Deputy Dean of Graduate School. He has authored more than 120 papers in referenced journals and international conferences. His current research interests include ultra-wideband bio-radar imaging and vital signal detection, AI driven information processing and Internet of Vehicles.



Zhi Ding (S'88-M'90-SM'95-F'03) is a Professor of Electrical and Computer Engineering at the University of California, Davis. He received his Ph.D. degree in Electrical Engineering from Cornell University in 1990. From 1990 to 2000, he was a faculty member of Auburn University and later, University of Iowa. Prof. Ding has held visiting positions in Australian National University, Hong Kong University of Science and Technology, NASA Lewis Research Center and USAF Wright Laboratory. Prof. Ding has active collaboration with researchers from universities in Australia, Canada, China, Finland, Hong Kong, Japan, Korea, Singapore, and Taiwan.

Dr. Ding is a Fellow of IEEE and has been an active member of IEEE, serving on technical programs of several workshops and conferences. He was associate editor for IEEE Transactions on Signal Processing from 1994-1997, 2001-2004, and associate editor of IEEE Signal Processing Letters 2002-2005. He was a member of technical committee on Statistical Signal and Array Processing and member of technical committee on Signal Processing for Communications (1994-2003). Dr. Ding was the General Chair of the 2016 IEEE International Conference on Acoustics, Speech, and Signal Processing and the Technical Program Chair of the 2006 IEEE Globecom. He was also an IEEE Distinguished Lecturer (Circuits and Systems Society, 2004-06, Communications Society, 2008-09). He served on as IEEE Transactions on Wireless Communications Steering Committee Member (2007-2009) and its Chair (2009-2010). Dr. Ding is a coauthor of the text: Modern Digital and Analog Communication Systems, 5th edition, Oxford University Press, 2019. Prof. Ding received the IEEE Communication Society's WTC Award in 2012.