I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) systems have demonstrated significant success in providing high spectrum and energy efficiency for 5G and future wireless communication systems. Considerable advantages are achieved by requiring sufficiently accurate channel state information (CSI) at the gNB (i.e., gNodeB). Owing to the well-known channel reciprocity, gNB can acquire downlink (DL) CSI by leveraging uplink (UL) CSI in time-division duplexing systems. However, as DL and UL channels are at different frequency bands in frequency-division duplexing (FDD) systems, we can only estimate DL CSI through feedback from UEs. However, because the feedback overhead requiring bandwidth to be transmitted increases proportionally with an increase in the number of antennas, reducing the feedback overhead becomes a critical task in FDD systems.

In recent years, deep learning techniques have been widely used in computer vision and natural language processing and have shown great promise in dimension reduction by extracting essential information from underlying structures in figures and signals. Therefore, deep learning techniques are regarded to be one of the most powerful tools for encoding and decoding CSI in massive MIMO FDD systems. The authors of [1], [2] proposed a CSI feedback deep neural network with an autoencoder structure whose encoder and decoder are deployed at UEs and base stations, respectively. This study and its variants [3], [4] have been demonstrated to outperform traditional compressive sensing-based methods [5]. Given that UEs must report CSI information to the gNB in bitstreams and cannot provide single-precision CSI information (i.e., 32 bits for each floating-point number), an efficient quantization module should be included in neural networks. Various quantization approaches have been designed for deep learning networks, such as uniform quantization, nonuniform quantization, and binary quantization [2], [6], [7].

Recently, researchers have observed that some magnitude-aided auxiliary information such as UL CSI magnitudes [8], past CSI magnitudes [3], [9] and CSI magnitudes of adjacent UEs [10] can help achieve more accurate DL CSI magnitude recovery at base stations. To fully capture the correlation between the side information and DL CSI magnitudes, advanced works have isolated CSI phase recovery from CSI magnitude recovery. These studies have utilized an isolated autoencoder to compress and recover the CSI magnitudes. Given that complex CSI matrix can be deemed as CSI phases weighted by CSI magnitudes, the authors in [8] design a magnitude-dependent phase quantization (MDPQ) to encode phases with different bits according to their magnitudes (or significance) with the intention of decreasing the overall MSE of...
complex CSI. They also provided an unsupervised deep-learning technique [11] to optimize the approach. The authors in [10] designed a deep learning model with a magnitude-dependent polar-phase (MDPP) loss function to automatically compress and quantize the significant CSI phases depending on the CSI magnitude. To the best of our knowledge, to leverage the side information for improving CSI magnitude recovery, existing works train two isolated deep learning models to estimate CSI magnitudes and phases, respectively, and combine them as complex CSI. However, to minimize the complex CSI MSE, the models for the CSI magnitude and phase should be jointly optimized. Moreover, the existing loss functions are not equivalent to the CSI MSE. Conventionally, the DL CSI matrix $\mathbf{H}_{DL}$ is estimated in the UE and feedback to the gNB. However, because the size of the CSI matrix $\mathbf{H}_{DL}$ is proportional to the number of subcarriers $N_f$ and the number of antennas $N_b$, the feedback overhead becomes a considerable burden in massive MIMO systems, thereby causing less bandwidth to be available to transmit data payload. To reduce the overhead, given the sparsity of the SFCSI matrix on the angle and delay domains, we can first apply the inverse discrete Fourier transform and DFT to the SFCSI matrix $\mathbf{H}_t$, and the transformed SFCSI matrix $\mathbf{H}_t$ can be represented as $\mathbf{H}_t = \mathbf{F}_D \mathbf{H}_A$. (3)

where $\mathbf{F}_D \in \mathbb{C}^{N_f \times N_f}$ and $\mathbf{F}_A \in \mathbb{C}^{N_b \times N_b}$ are the IFDT and DFT matrices, respectively. Most elements in the CSI matrix $\mathbf{H}_t$ are near zero, except for the first $Q_f$ rows. Therefore, we truncated the channel matrix into the first $Q_f$ rows that have distinct non-zero values and utilize $\mathbf{H}_{DL}$ and $\mathbf{H}_{UL}$ to denote the first $Q_f$ rows of the transformed matrices of $\mathbf{H}_{SF}^{DL}$ and $\mathbf{H}_{SF}^{UL}$, respectively. For simplicity, in the remainder of this letter, we denote $\mathbf{H}_{DL}$ as $\mathbf{H}$.

Subsequently, to further reduce the feedback overhead, the DL CSI matrix $\mathbf{H}$ is compressed and quantized as codewords by the encoder at the UE for recovery at the gNB. The recovered DL CSI matrix can be expressed as $\hat{\mathbf{H}} = f_{de}(f_{en}(\mathbf{H}))$. (4)

where $f_{en}(\cdot)$ and $f_{de}(\cdot)$ denote the encoding and decoding operations.

III. Magnitude-aided CSI feedback Framework

We aim to recover the DL CSI based on the received codewords encoded by the UEs using the decoder at the gNB. The training process optimizes the encoder–decoder pair by minimizing the discrepancy between the true and estimated DL CSI matrices (i.e., MSE of DL CSI). Most studies [1]–[3], [9] decomposed DL CSI matrices into real and imaginary parts to be tackled by real-value deep neural networks, as shown in Fig. 1 (a). However, in recent studies [8]–[10], CSI magnitudes were isolated from CSI phases and fed into an isolated encoder–decoder pair only for CSI magnitude recovery.

In this study, we develop a learning-based CSI feedback framework based on limited feedback, CSI magnitude, and phase-side information. We also redesign a loss function, termed the sinusoidal magnitude-dependent phase (SMDP), that corresponds to the MSE of the DL CSI. The contributions of this study are as follows.

- To the best of our knowledge, with magnitude-aided auxiliary inputs, the framework is the first attempt to jointly optimize magnitude and phase recovery models for minimizing overall complex CSI MSE.
- With regard to the execution of the framework, one of the sinusoidal phase inputs can be replaced with a one-bit sign matrix for reducing almost half of the feedback overheads. Pruning of the sign matrix can further reduce feedback overhead, and the overall required overhead is provided.

The remainder of this letter is organized as follows. Section II introduces the system model and the general CSI feedback principle. Section III explains the general principle and provides an example of the proposed framework. Section IV describes the previous approaches and the proposed loss function for phase recovery. The numerical results are presented in Section IV. Finally, concluding statements are presented in Section V.

II. System Model

In this study, we consider a single-cell massive MIMO FDD scenario, where gNB is equipped with $N_b \gg 1$ and each UE has a single antenna. An orthogonal frequency-division modulation waveform was adopted with $N_f$ subcarriers. The DL received signal at the $k$-th subcarrier is

$$y_{DL}(k) = \mathbf{h}_{DL}(k) \mathbf{w}_T(k) x_{DL}(k) + n_{DL}(k)$$

where $\mathbf{h}_{DL}(k) \in \mathbb{C}^{N_b \times 1}$ denotes the channel vector at the $k$-th subcarrier, and $\mathbf{w}_T(k) \in \mathbb{C}^{N_f \times 1}$ denotes the transmit precoding vector 1. $x_{DL}(k) \in \mathbb{C}$ and $n_{DL}(k) \in \mathbb{C}$ are the DL transmit signal and additive white Gaussian noise at the $k$-th subcarrier. $(\cdot)^H$ denotes the conjugate transpose. The UL received signal is

$$y_{UL}(k) = \mathbf{h}_{UL}(k) x_{UL}(k) + n_{UL}(k) \in \mathbb{C}^{N_f \times 1}$$

(2)

where $\mathbf{h}_{UL}(k) \in \mathbb{C}^{N_b \times 1}$ is the UL channel vector, and the subscript UL denotes the UL signals and noise, similar to Eq. (1). DL and UL channel vectors can be represented as DL and UL spatial-frequency channel state information (SFCSI) matrices $\mathbf{H}_{SF}^{DL} = [\mathbf{h}_{DL}(1), ..., \mathbf{h}_{DL}(N_f)]^H \in \mathbb{C}^{N_f \times N_b}$ and $\mathbf{H}_{UL} = [\mathbf{h}_{UL}(1), ..., \mathbf{h}_{UL}(N_f)]^H \in \mathbb{C}^{N_f \times N_b}$, respectively.

Conventionally, the DL CSI matrix $\mathbf{H}_{DL}$ is estimated in the UE and feedback to the gNB. However, because the size of the CSI matrix $\mathbf{H}_{DL}$ is proportional to the number of subcarriers $N_f$ and the number of antennas $N_b$, the feedback overhead becomes a considerable burden in massive MIMO systems, thereby causing less bandwidth to be available to transmit data payload. To reduce the overhead, given the sparsity of the SFCSI matrix on the angle and delay domains, we can first apply the inverse discrete Fourier transform and DFT to the SFCSI matrix $\mathbf{H}_t$, and the transformed SFCSI matrix $\mathbf{H}_t$ can be represented as $\mathbf{H}_t = \mathbf{F}_D \mathbf{H}_A$. (3)

where $\mathbf{F}_D \in \mathbb{C}^{N_f \times N_f}$ and $\mathbf{F}_A \in \mathbb{C}^{N_b \times N_b}$ are the IFDT and DFT matrices, respectively. Most elements in the CSI matrix $\mathbf{H}_t$ are near zero, except for the first $Q_f$ rows. Therefore, we truncated the channel matrix into the first $Q_f$ rows that have distinct non-zero values and utilize $\mathbf{H}_{DL}$ and $\mathbf{H}_{UL}$ to denote the first $Q_f$ rows of the transformed matrices of $\mathbf{H}_{SF}^{DL}$ and $\mathbf{H}_{SF}^{UL}$, respectively. For simplicity, in the remainder of this letter, we denote $\mathbf{H}_{DL}$ as $\mathbf{H}$.

Subsequently, to further reduce the feedback overhead, the DL CSI matrix $\mathbf{H}$ is compressed and quantized as codewords by the encoder at the UE for recovery at the gNB. The recovered DL CSI matrix can be expressed as $\hat{\mathbf{H}} = f_{de}(f_{en}(\mathbf{H}))$. (4)

where $f_{en}(\cdot)$ and $f_{de}(\cdot)$ denote the encoding and decoding operations.

III. Magnitude-aided CSI feedback Framework

We aim to recover the DL CSI based on the received codewords encoded by the UEs using the decoder at the gNB. The training process optimizes the encoder–decoder pair by minimizing the discrepancy between the true and estimated DL CSI matrices (i.e., MSE of DL CSI). Most studies [1]–[3], [9] decomposed DL CSI matrices into real and imaginary parts to be tackled by real-value deep neural networks, as shown in Fig. 1 (a). However, in recent studies [8]–[10], CSI magnitudes were isolated from CSI phases and fed into an isolated encoder–decoder pair only for CSI magnitude recovery.

1 gNB can calculate each precoding vector at each subcarrier with the DL CSI matrix.
to fully exploit the multipath correlation between DL CSI magnitudes and auxiliary magnitude information (AMI). Another encoder–decoder pair is trained for the CSI phase recovery. The general architecture of such studies is illustrated in Fig. 1 (b).

A. Basic Principle of the Proposed Framework

Our proposed general framework is shown in Fig. 1 (c), having the magnitude, phase branches, and a combining network. Each branch contains an encoder-decoder pair and yields an initial estimation of the CSI magnitude and phase. At the UE, the DL CSI is first decomposed into the CSI magnitudes and phases. Subsequently, the CSI phases are further decomposed as a cosine (or sine) and one-bit sign matrices, serving as auxiliary phase information and direct feedback to the gNB. The codewords are generated by the encoders at the UE based on the CSI magnitudes and cosine matrix and decompressed as initial estimates of CSI magnitudes and phases based on AMI and API at the gNB. Finally, the combining network is applied to merge the two estimates into a complex CSI matrix and refine it.

Our goal is to optimize the framework with the minimum complex CSI MSE. Even for different CSI phases, we would not obtain a different CSI MSE while having the same amount of CSI magnitude errors. Hence, the CSI magnitude branch can be optimized without considering the CSI phases. However, to minimize the complex CSI MSE, the CSI phases are not equally important. Specifically, the CSI is equivalent to the CSI phases weighted by the CSI magnitudes. Hence, the CSI magnitude is considered when optimizing the CSI phase branch and the combining network. Thus, the training scheme is designed as illustrated in Fig. 2. As the magnitude branch can be independently optimized, the training of the framework is completed in two stages. In the first stage, the parameters in the CSI magnitude branch were pre-trained for magnitude estimation. Subsequently, in the second stage, the CSI phase branch and the combining network are optimized with the help of the magnitude branch, while the parameters of the magnitude branch are fixed.

B. DualNet-MAG-PHA

Without loss of generality, we herein demonstrate a framework with an example, called DualNet-MAG-PHA. As shown in Fig. 3, each complex CSI matrix is split into magnitude and exponential phase parts (i.e., $H$ is decomposed into $|H|$ and $\angle H$). Similar to the network in [11], the CSI magnitudes are sent to the magnitude encoder network, including four $7 \times 7$ convolutional layers with 16, 8, 4, and 1 channels. Three leaky ReLU and one ReLU activation function follow the four layers. Subsequently, a fully connected (FC) layer with $\text{Round}(\text{CR}_\text{MAG}Q/\alpha_b)$ elements is connected for dimension reduction after reshaping. $\text{CR}_\text{MAG}$ denotes the compression ratio in the magnitude branch. The output of the FC layer is then fed into the quantization module, called the sum-of-sigmoid (SSQ), which is designed in CQNet [11]. The gNB magnitude decoder utilizes quantized codewords and locally available UL CSI magnitudes as AMI to jointly decode the DL CSI magnitudes. The magnitude network is optimized by updating the network parameters $\Theta_\text{MAG}$ to minimize the MSE in magnitude:

$$\arg\min_{\Theta_{\text{en,MAG}},\Theta_{\text{de,MAG}}} \left\{ \| |\hat{H}| - |H| \|_F^2 \right\}$$

(5)

where

$$|\hat{H}| = f_{\text{de,MAG}}(f_{\text{en,MAG}}(|H|), \Theta_{\text{en,MAG}}, \Theta_{\text{de,MAG}}, H_{\text{UL}}),$$

(6)

the subscripts en, de, UL, and MAG of the $f(\cdot)$ denote the encoder, decoder, UL, and magnitude branch, respectively. $\Theta$ denotes network parameters.

For the exponential phase part, instead of feeding the original CSI phases directly (i.e., $\angle H$), the CSI phases are further decomposed into a cosine matrix and a sign matrix comprising only 1-bit entries representing a positive or a negative sign. The explanation for this is discussed in the next section. The cosine matrix is
fed into the phase encoder, which is similar to the magnitude encoder for compression and quantization, and the Round(CRPHA, Qf)Nb-element codewords are obtained. CRPHA denotes the compression ratio in the phase branch. The tanh activation functions are used after each convolutional layer in the phase encoder as they can provide high-order nonlinearity to capture the underlying features of significant phases associated with large magnitudes. Subsequently, the quantized codewords $f_{en,PHA}(\text{Cos}, \Theta_{en,PHA})$ and feedback sign matrix $A$ are fed into the phase decoder with a tanh activation function as the last layer to range the value of the estimated DL CSI cosine matrix $\hat{\text{Cos}}$ within $[-1, 1]$.

The estimated CSI magnitude matrix $\hat{\text{H}}$ and cosine matrix $\hat{\text{Cos}}$ are then fed into the combining network to obtain the DL CSI matrix. By applying the Pythagorean trigonometric identity, we can obtain the corresponding sine matrix based on the cosine and sign matrices. Then, combined with the estimated magnitude matrix, initial estimates of the real and imaginary parts of the DL CSI matrix can be obtained and further refined by four convolutional layers as the final estimates of the DL CSI matrix.

$$\arg\min_{\Theta_{en,PHA}, \Theta_{de,PHA}, \Theta_{C}} \left\{ \| \hat{\text{H}} - \text{H} \|^2_F \right\},$$

$$\hat{\text{H}} = f_{C}(|\hat{\text{H}}|, \hat{\text{Cos}}, \Theta_{C}),$$

$$\hat{\text{Cos}} = f_{de,PHA}(f_{en,PHA}(\text{Cos}, \Theta_{en,PHA}), A, \Theta_{de,PHA})$$

where the subscripts PHA, C of $f(\cdot)$, and $\Theta$ denote the encoder, decoder, phase branch, and combining network, respectively.

Lastly, in such a framework, although we exploit two loss functions during training, it is equivalent to optimizing the whole framework only following Eq. (7) (i.e., minimizing the overall CSI MSE) because the training of the magnitude branch can be optimized without the help of CSI phases.

IV. PRIOR WORKS AND LOSS FUNCTION REDESIGN

The most common loss function to evaluate the CSI discrepancy is the CSI MSE. However, for introducing AMI to boost the CSI magnitude estimation, it is not trivial to render the entire framework to minimize the CSI MSE exactly. In this section, we explore why the existing and possible approaches cannot fully meet the seemingly easy goal, and a sinusoidal magnitude-dependent phase (SMDP) loss function is proposed to achieve this goal.

As mentioned in the previous section, to leverage the AMI, in existing works, the CSI magnitude is separated from the CSI phases and fed into an isolated model for CSI magnitude recovery at the gNB. They utilized the MSE of CSI magnitudes as a loss function to optimize the model for CSI magnitude recovery. The choice of the loss function is reasonable because the MSE of the CSI magnitudes is uncorrelated with the CSI phases. Thus, we only explore whether the existing and possible approaches for phase recovery are equivalent to minimizing the complex CSI MSE.

In [8], we designed magnitude-dependent phase quantization (MDPQ), which assigns more bits to significant phases corresponding to greater magnitudes, as well as provided an unsupervised deep learning model [11] to determine the hyperparameters of MDPQ. This can sometimes achieve the goal by properly setting the hyperparameters.

Some authors have attempted exploiting pure deep neural networks to conduct CSI phase compression and recovery. It is intuitive to design a naive loss function as follows:

$$\text{LossNaive} = \text{MSE}_{\text{CSI}} = \| \text{H} - \hat{\text{H}} \|^2_F,$$

$$= \| |\text{H}| \circ \cos(\angle \text{H}) - |\hat{\text{H}}| \circ \cos(\angle \hat{\text{H}}) \|^2_F + \| |\text{H}| \circ \sin(\angle \text{H}) - |\hat{\text{H}}| \circ \sin(\angle \hat{\text{H}}) \|^2_F.$$ (10)

As shown in Eq. (10), to calculate the loss between the true and estimated CSI, sinusoidal functions are required to act as activation functions. However, the presence of infinitely many and shallow local minima of sinusoidal functions causes training difficulties [12].

Fig. 3. Network architecture of DualNet-MAG-PHA.
Recently, the authors in [10] designed a network with an MDPP loss function to avoid training difficulties and reconstruct the original CSI phases. The loss function is

$$\text{Loss}_{\text{MDPP}} = \text{MSE}_{\text{MDPP}} = \|H - \hat{H}\|_F^2$$  \hspace{1cm} (11)

where $H$ and $\hat{H}$ denote the true and estimated phases in radians, respectively. The original polar-phase discrepancy was weighted by the true CSI magnitude. To do so, the network will pay more attention for capturing the underlying features of the critical phases associated with greater magnitudes. However, although the loss function can effectively help compress significant phase components, the loss function is not equivalent to our function can effectively help compress significant phase components, the loss function is not equivalent to our final goal, i.e., the MSE of complex CSI.

To achieve this and to avoid the training problem of Eq. (10), we propose the SMDP loss function by replacing the phase inputs (i.e., $\angle H$) with sinusoidal inputs (i.e., $\text{Sin}$ and $\text{Cos}$).

$$\text{MSE}_{\text{SMDP}} = \|H - \hat{H}\|_F^2$$

$$= \|H \circ \text{Cos} - |\hat{H}| \circ |\text{Cos}|\|_F^2$$

$$+ \|H \circ \text{Sin} - |\hat{H}| \circ \text{Sin}|_F^2$$ \hspace{1cm} (12)

$$\text{Sin} = A \circ (1 - \text{Cos} \circ \text{Cos}),$$

$$\text{Cos} = f_{de, PHA}(f_{en, PHA}(\text{Cos}, \Theta_{en, PHA}), \Theta_{de, PHA}),$$  \hspace{1cm} (15)

where $A$ is the sign matrix that encodes the sign of the sine matrix using one bit for each entry. Using Pythagorean trigonometric identity, we can save almost half the bandwidth by transmitting the sign matrix instead of the compressed sine matrix. Moreover, similar to the CSI phases, most entries in the sign matrix are insignificant. In practice, we can prune the sign matrix before transmitting. That is, we only transmit partial entries of the sign matrix associated with large magnitudes in a sign ratio $R_s$ to further reduce feedback overhead. The total phase feedback overhead (in bits) is summarized as follows:

$$B_{\text{SMDP}} = CR_{\text{PHA}}(K_{\text{PHA}}Q_fN_b + R_sQ_fN_b)/(\text{bits}),$$  \hspace{1cm} (16)

where $K_{\text{PHA}}$ denotes the number of quantization bits for each entry of the compressed cosine matrix $f_{en, PHA}(\text{Cos}, \Theta_{en, PHA})$. We use Eq. (5) as the loss function during the first training stage. In the second training stage, we used Eqs. (12) as the loss function to build an end-to-end learning architecture.

V. Experimental Evaluations

A. Experiment Setup

For the indoor channel, we used the industry-grade COST 2100 model [13] to generate channels at 5.1-GHz UL and 5.3-GHz DL bands. We also evaluated our method under the second dataset generated by the QuaDRiGa software, satisfying 3GPP TR 38.901 v15.0.0 [14]. The urban microcell (UMi) scenario at 2 and 2.1 GHz of UL and DL bands, respectively, with non-line-of-sight paths are considered. The number of cluster paths was set as 13. Both the UL and DL bandwidths were 20 MHz. We placed the gNB with a height of 20 m at the center of a circular area with a radius of 20 m for indoor coverage and 200 m for outdoor coverage. The number of gNB antennas is $N_b = 64$, and each UE is equipped with a single antenna. The inter-antenna spacing is a half wavelength, that is, $d = f_c/2c$, where $f_c$ and $c$ are the carrier frequency and light speed, respectively. For each trained model, the number of epochs and batch size were set to 1,000 and 200, respectively.

The performance metric is the normalized MSE of DL CSI shown as follows.

$$\text{NMSE} = \sum_{d=1} D \left\| H_{\text{SF}, DL,d} - \hat{H}_{\text{SF}, DL,d} \right\|^2_F / \left\| H_{\text{SF}, DL} \right\|^2_F ,$$  \hspace{1cm} (17)

$$\hat{H}_{\text{SF}, DL,d} = \left\{ (F_H^T H_{\text{DL},d} F_A) T 0_{N_b \times (N_f - Q_f)} \right\}^T ,$$  \hspace{1cm} (18)

where the number $D$ and subscript $d$ denote the total number and index of channel realizations, respectively. Instead of evaluating the estimated DL CSI matrix $\hat{H}_{\text{DL},d}$, we evaluate the estimated SFCSI matrix $\hat{H}_{\text{SF}, DL,d}$ that can be obtained by reversing the Fourier processing and padding zero matrix $0_{N_b \times (N_f - Q_f)}$.

In the following subsection, we examine the performance improvement in phase recovery by adopting the proposed loss function. Thus, we trained DualNet-MAG-PHA with the same core layer design for magnitude recovery but with different methods to reconstruct the CSI phases in phase compression ratios $CR_{\text{PHA}} = 1/8$ and $1/16$:

- **SMDP**: the network architecture follows DualNet-MAG-PHA. The sign ratios $R_s$ and $K_{\text{PHA}}$ were set as $[0.25, 0.125]$ and $[8, 8]$ bits for $CR_{\text{PHA}} = [1/8, 1/16]$, respectively.
- **MDPQ**: it follows the magnitude branch of DualNet-MAG-PHA. The design assigns $[0, 0, 0, 3, 7]$ and $[0, 0, 0, 0, 5]$ bits for $CR_{\text{PHA}} = [1/8, 1/16]$, respectively, to encode the CSI phases corresponding to $[0, 0.5, 0.7, 0.8, 0.9]$ of the cumulative distribution function of magnitude.
- **Naive MSE**: it basically follows DualNet-MAG-PHA but, instead of the cosine matrix, original phases

\footnote{All following alternate approaches consume 1.2 and 0.625 bits/phase entry}
are fed into the phase encoder and cosine and sine functions are appended as the last layer of the phase decoder. The loss function for phase reconstruction is given by Eq. (9). $K_{PHA}$ was set to 8 bits.

- MDPP (unwrapped and wrapped): the network architecture is the same as (c) while the loss function for phase reconstruction is Eq. (10). $K_{PHA}$ was set to 10 bits. With wrapped phases, the loss function for phase reconstruction is as follows:

$$\text{Loss}_{\text{wrapped}}^\text{MDPP} = \min(\Delta \phi, \Delta \phi + 2\pi, \Delta \phi - 2\pi),$$

$$\Delta \phi = \frac{1}{2} \left\| \angle H - \angle \hat{H} \circ |H| \right\|_F^2.$$ (19)

B. Different Phase Compression Designs

To demonstrate the superiority of the proposed SMDP loss function, we applied different phase reconstruction approaches to DualNet-MAG-PHA for different phase compression ratios $CR_{PHA}$. Figs. 4 (a) and (b) show the NMSE performance for different phase reconstruction approaches under indoor and outdoor scenarios and different compression rates. As expected, DualNet-MAG-PHA encounters training difficulties when using a naive loss function. As for adopting MDPP loss functions, DualNet-MAG-PHA performs better than the naive loss function. Although DualNet-MAG-PHA performs better when using MDPO than the MDPP loss function, the bit-assignment rule should be carefully determined to achieve a satisfactory result. Finally, DualNet-MAG-PHA using the proposed SMDP loss function has an approximately 4-dB improvement in terms of NMSE performance.

C. Different Core Layer Designs

To examine the appropriate core layer designs of DualNet-MAG-PHA that can efficiently capture the underlying features of CSI phases, we provide a performance evaluation using FC layers and convolutional layers, denoted as DNN and CNN, respectively, for the core network, and using SSQ [11] and binary-level quantization (BLQ) as the quantization module. We consider the phase compression ratio of $CR_{PHA} = 1/8$. For SSQ, we assign $K_{PHA} = 8$ bits for each codeword. That is, there are 8-bit $CR_{PHA}Q_fN_b = 128$ codewords transmitted to the gNB. In contrast, there are 1-bit $K_{PHA}CR_{PHA}Q_fN_b = 1024$ codewords when applying BLQ.

Figs. 5 (a) and (b) show the NMSE performance for network architectures with different core layer designs. For both indoor and outdoor scenarios, DualNet-MAG-PHA performs better when adopting SSQ and CNN. This implies two possible reasons. The first is that, different from BLQ, SSQ is differentiable such that it is easier to be trained. The second is that there are still structural features of CSI phases in the angle-delay domain that could be extracted. Moreover, there are 860 parameters of CNN as compared with approximately 8M parameters of DNN. According to the results, we can largely conclude that the preferable core layer design is a combination of SSQ and CNN.

VI. Conclusions

In this letter, we propose a learning-based CSI feedback framework based on limited feedback and magnitude-aided information for massive MIMO FDD systems. In contrast to previous works, our proposed framework with a proposed loss function enables end-to-end learning to jointly optimize the CSI magnitude and phase recovery performance. Simulations reveal that, compared to competitive alternatives for phase recovery, the proposed loss function can significantly improve the overall CSI recovery in both indoor and outdoor scenarios. Finally, the combination of SSQ and CNN was verified to be the preferable core layer design for phase recovery.
References

[1] 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, 2018.
[2] 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., vol. 19, no. 4, pp. 2827–2840, 2020.
[3] 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, 2019.
[4] Z. Qin, H. Ye, G. Y. Li, and B. F. Juang, “Deep Learning in Physical Layer Communications,” IEEE Wireless Commun., vol. 26, no. 2, pp. 93–99, 2019.
[5] P. Kuo, H. T. Kung, and P. Ting, “Compressive Sensing Based Channel Feedback Protocols for Spatially-Correlated Massive Antenna Arrays,” in 2012 IEEE Wireless Commun. and Netw. Conf. (WCNC), 2012, pp. 492–497.
[6] 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, 2019.
[7] J. Jang, H. Lee, S. Hwang, H. Ren, and I. Lee, “Deep Learning-Based Limited Feedback Designs for MIMO Systems,” IEEE Wireless Commun. Lett., vol. 9, no. 4, pp. 558–561, 2020.
[8] 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.
[9] Z. Liu, M. Rosario, and Z. Ding, “A Markovian Model-Driven Deep Learning Framework for Massive MIMO CSI Feedback,” arXiv preprint arXiv:2009.09468, 2020.
[10] J. Guo, X. Yang, C. Wen, S. Jin, and G. Li, “DL-based CSI Feedback and Cooperative Recovery in Massive MIMO,” arXiv preprint arXiv:2003.03303, 2020.
[11] Z. Liu, L. Zhang, and Z. Ding, “An Efficient Deep Learning Framework for Low Rate Massive MIMO CSI Reporting,” IEEE Trans. Commun., vol. 68, no. 8, pp. 4761–4772, 2020.
[12] G. Parascandolo, H. Huttunen, and T. Virtanen, “Taming the Waves: Sine as Activation Function in Deep Neural Networks,” 2017.
[13] 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, 2012.
[14] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, “QuaDRiGa: A 3-D Multi-Cell Channel Model with Time Evolution for Enabling Virtual Field Trials,” IEEE Trans. Antennas and Propag., vol. 62, no. 6, pp. 3242–3256, 2014.