Spherical Normalization for Learned Compressive Feedback in Massive MIMO CSI Acquisition

Zhenyu Liu, Mason del Rosario, Xin Liang, Lin Zhang, and Zhi Ding

***Abstract*—Massive multiple-input multiple-output (MIMO) systems have been highlighted as a key enabling technology for future communications networks. However, the efficacy of MIMO hinges on the availability of accurate channel state information (CSI). In frequency division duplex (FDD) massive MIMO systems, downlink CSI acquisition involves determining a sufficient amount of feedback compression to maintain high spectral efficiency. Prior works have highlighted the efficacy of deep convolutional neural networks (CNN) for learning an encoding scheme for downlink CSI. However, the prior works have approached the problem of channel estimation similarly to other domains in which CNNs have found success (e.g., computer vision). This paper proposes two techniques for tailoring CNN-based autoencoders to the problem of massive MIMO CSI recovery. The first technique is a power-based normalization of CSI (spherical normalization), which can optimize the input distribution and make the network more applicable to the commonly adopted accuracy metric. The second technique is an optimized CNN architecture for codeword efficiency. We propose networks, called SphNet and DualNet-Sph, which combine these techniques, and we demonstrate that these new networks outperform previously proposed deep learning networks for CSI feedback.**

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

I. Introduction

Massive multiple-input multiple-output (MIMO) technologies have the ability to increase the information capacity of wireless channels [1], making them the subject of focus for much research into future wireless communications networks. To realize the spectral efficiency and spatial diversity gains offered by massive MIMO, the base station must have accurate and frequently updated channel state information (CSI) to enable message precoding. In frequency division duplex (FDD) systems, gNB (or gNodeB) transmitters require user equipment (UE) to provide downlink CSI feedback frequently. Given the high antenna count in massive MIMO and the unpredictable nature of any given wireless channel, techniques for determining an appropriate compression ratio which maintains accurate CSI estimates at the base station are of critical importance.

Broadly speaking, research efforts in compressed CSI estimation have fallen in one of three categories. The first category, direct quantization, involves mapping continuous-valued downlink CSI at the user equipment (UE) to discrete levels. The resulting bit string is encoded and transmitted to the base station (eNB or gNB) [2], [3]. While encoding under direct quantization is time-efficient, the resulting compression often consumes too much uplink bandwidth to be practical. The second category, compressed sensing (CS), leverages the sparsity of massive MIMO CSI under a linear transformation (i.e., 2D Fourier Transform) [4], [5]. While these methods reduce feedback overhead, they predominantly involve strong assumptions of sparisty which are not always true in real channel environments and iterative solvers which increase the time to decode compressed downlink estimates.

The third category of CSI estimation, *deep learning*, involves the training of deep neural networks on large sets of CSI data. Recent work has highlighted the efficacy of deep learning in generating encoding schemes for CSI estimation. In [6], the authors utilize an autoencoder comprised of convolutional layers to learn an encoding of real and imaginary values of downlink CSI. Other authors use different feedback modalities to enhance estimation accuracy. In [7], the authors show that using uplink and downlink CSI's magnitude and absolute value can provide more reliable channel estimates than compressed real and imaginary downlink CSI values. In [8], the authors show the impact of temporal correlation between timeslots in increasing estimation accuracy at low compression ratios.

While the prior work has considered different feedback modalities, small effort has gone into adapting the objective function to channel data. Mean-squared error (MSE) is the typical optimization target for the training of deep neural networks, but the ultimate figure of merit for CSI estimation accuracy is normalized mean-squared error (NMSE). Normalization by the signal power is crucial when considering CSI, as the data span several orders of magnitude. In this work, we investigate the influence of power-based normalization in training convolutional neural networks (CNNs) for CSI feedback in massive MIMO systems, and the following contributions are claimed:

- **Contribution #1**: We propose a CSI feedback network **SphNet**, which is the optimization of the autoencoder-based CSI feedback neural network by
  1) Utilizing spherical CSI feedback by separating power from CSI matrices to produce a better data distribution and objective function for the network.
  2) Optimizing the autoencoder structure to produce more efficient CSI codewords and more accurate

Z. Liu is with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China, and also with the Department of Electrical and Computer Engineering, University of California at Davis, Davis, CA 95616 USA (e-mail: lzyu@bupt.edu.cn).

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

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

978-1-7281-7440-2/20/$31.00 ©2020 IEEE

CSI reconstruction accuracy.

- **Contribution #2**: We propose a CSI feedback network **DualNet-Sph**, which is the enhancement of DualNet-MAG that exploits bi-directional channel correlation [7] by implementing the changes outlined in **Contribution #1**.

We demonstrate enhanced CNN-based autoencoders, **Sph-Net** and **DualNet-Sph**, which have improved estimation accuracy for indoor and outdoor mobile networks.

This paper is organized as follows. Section II describes the massive MIMO system model adopted for this work. Section III describes the rationale for power-based spherical normalization. Section IV discusses architectural improvements for the CSI feedback network. Section V outlines how we implement these improvements in DualNet-MAG. Section VI presents our experimental results for the improved networks as they relate to CsiNet and DualNet-Mag. Section VII concludes the paper by highlighting some potential future directions.

II. System Model

In this paper, we consider a single-cell massive MIMO system, where the gNB is equipped with $N_b \gg 1$ antennas and each UE has a single antenna. Orthogonal frequency division multiplexing (OFDM) is adopted over N_f subcarriers, and the downlink received signal at the $m-$th subcarrier is

$$y_{d,m} = \mathbf{h}_{d,m}^H \mathbf{w}_{t,m} x_{d,m} + n_{d,m}, \quad (1)$$

where $\mathbf{h}_{d,m} \in \mathbb{C}^{N_b \times 1}$ denotes the channel vector of the $m-$th subcarrier, $\mathbf{w}_{t,m} \in \mathbb{C}^{N_b \times 1}$ denotes transmit precoding vector, $x_{d,m} \in \mathbb{C}$ is the transmitted symbol, and $n_{d,m} \in \mathbb{C}$ is the additive noise, and $(\cdot)^H$ represents the conjugate transpose. The uplink received signal of the $m-$th subcarrier is given by

$$y_{u,m} = \mathbf{w}_{r,m}^H \mathbf{h}_{u,m} x_{u,m} + \mathbf{w}_{r,m}^H \mathbf{n}_{u,m}, \quad (2)$$

where $\mathbf{w}_{r,m} \in \mathbb{C}^{N_b \times 1}$ denotes the received precoding vector, 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 $\tilde{\mathbf{H}}_u = \left[\mathbf{h}_{u,1}, ..., \mathbf{h}_{d,N_f}\right]^H \in \mathbb{C}^{N_f \times N_b}$, respectively.

With the downlink channel matrix $\tilde{\mathbf{H}}_d$, gNB can calculate the transmit precoding vector in each subcarrier. However, since the size of the CSI matrix $\tilde{\mathbf{H}}_d$ is proportional to N_f and N_b, CSI feedback payload becomes huge for massive MIMO system, which can greatly occupy the precious bandwidth.

To reduce the feedback overhead, we first exploit the sparsity of CSI in the delay domain [9]. Since the delay among multiple paths lies in a particularly limited period, CSI matrices exhibit sparsity in the delay domain. With the help of inverse discrete Fourier transform (IDFT), CSI matrix $\mathbf{H}_f$ in frequency domain can be transformed to be $\mathbf{H}_t$ in delay domain using

$$\mathbf{F}^H \mathbf{H}_f = \mathbf{H}_t, \quad (3)$$

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

Once the CSI matrix $\mathbf{H}_d$ has been estimated, further compression can start to reduce the redundancy in reporting downlink CSI with the help of deep neural networks at the UE. We consider two kinds of CSI feedback architectures including the downlink-only CSI feedback architecture and uplink-assisted downlink CSI feedback architecture.

Denote $\hat{\mathbf{H}}_d$ as the reconstructed downlink CSI matrix. Define the encoding and decoding function as $f_e(\cdot)$ and $f_d(\cdot)$, respectively. For downlink-only CSI feedback architecture, the encoder network and decoder network can be denoted, respectively, by

$$\mathbf{s}_1 = f_{e,1}(\mathbf{H}_d), \quad (4)$$
$$\hat{\mathbf{H}}_d = f_{d,1}(\mathbf{s}_1). \quad (5)$$

For uplink-assisted downlink CSI feedback architecture, the encoder network and decoder network can be denoted, respectively, by

$$\mathbf{s}_2 = f_{e,2}(\mathbf{H}_d), \quad (6)$$
$$\hat{\mathbf{H}}_d = f_{d,2}(\mathbf{s}_2, \mathbf{H}_u). \quad (7)$$

III. Spherical CSI Feedback

As a powerful tool for exploring the underlying structures from large data sets, deep learning (DL) has been widely used in computer vision [10] and natural language processing [11]. Recently, DL has also been successfully applied to derive reliable downlink CSI in massive MIMO systems for channel estimation [12] and low rate CSI feedback [6]–[8] when traditional methods generate limited performance.

However, how to effectively apply DL techniques to exploit channel data properties and optimization objects remains an open research issue, as many existing DL based works mainly utilize the deep learning architectures and optimization functions from the computer vision area without customization for the channel data characteristics, which can degrade the performance of DL-based CSI feedback algorithms. In particular, the training data processing methods and loss functions for the computer vision may not be suitable for the training of CSI-related networks directly.

On the one hand, existing DL-based CSI feedback works mainly regard the 2D channel matrix as an image and the normalized feature scaling values of CSI matrix are utilized as the input to DL networks. However, owing to the influence of path loss, the distribution of CSI-related data is different from the distribution of image data.

Images are comprised of matrices of pixel values. For color images, they have a matrix of pixel values for each color channel, such as red, green, and blue. Pixel values are unsigned integers in the range between 0 and 255. By scaling pixel values to the range $[0, 1]$, there can be great benefit in preparing the image pixel values prior to training of the DL model.

However, different from images where all pixels are in same order of magnitude, the range of CSI data can be much larger than the image pixel values. Since the path loss decreases exponentially relative to the distance between gNB and UE [13], the CSIs of UE near the gNB and UE far from the gNB can differ by several orders of magnitude. Consequently, the DL network may not be as sensitive to the CSIs of UE far from the gNB as the CSIs of UE near the gNB.

On the other hand, existing DL-based CSI feedback works mainly used the loss function in image processing for the training of DL networks, e.g., mean square error (MSE), which can be defined as

$$\mathrm{MSE} = \frac{1}{n}\sum_{k=1}^{n} \|\mathbf{H}_d^k - \hat{\mathbf{H}}_d^k\|^2, \tag{8}$$

where k and n are, respectively, the index and total number of samples in the data set, and $\|\cdot\|$ is Frobenius norm. However, Normalized MSE (NMSE), which is defined as

$$\mathrm{NMSE} = \frac{1}{n}\sum_{k=1}^{n} \|\mathbf{H}_d^k - \hat{\mathbf{H}}_d^k\|^2 / \|\mathbf{H}_d^k\|^2, \tag{9}$$

is the common metric utilized to evaluate the accuracy of CSI estimation [12] and feedback [6]–[8]. By comparing the definition of MSE and NMSE, we can find that using MSE as the loss function can make DL networks pay more attention to the CSI matrices with the large power.

To deal with the above two problems, we propose a spherical CSI feedback architecture as shown in Fig. 1. The spherical CSI feedback architecture splits the downlink CSI matrix $\mathbf{H}_d^k$ into a power value p_k and a spherical downlink CSI matrix $\check{\mathbf{H}}_d^k$, where $p_k = \|\mathbf{H}_d^k\|$ is the power of CSI matrix and $\check{\mathbf{H}}_d^k = \mathbf{H}_d^k / \|\mathbf{H}_d^k\|$ is the spherical downlink CSI matrix. Since the power of CSI matrix is split from the original CSI matrix, the rest CSI matrix is in a spherical surface. Then the power and the spherical CSI matrix are fed back separately.

Spherical CSI feedback architecture can bring two advantages. First, since "MSE" is one of the default loss functions in the DL networks, the optimization of spherical downlink CSI recovery method can be formulated as minimizing the loss function, L_{MSE} defined as

$$\begin{aligned} L_{\mathrm{MSE}} &= \frac{1}{n}\sum_{k=1}^{n} \|\check{\mathbf{H}}_d^k - \hat{\check{\mathbf{H}}}_d^k\|^2 \\ &= \frac{1}{n}\sum_{k=1}^{n} \|\mathbf{H}_d^k/\|\mathbf{H}_d^k\| - \hat{\mathbf{H}}_d^k/\|\mathbf{H}_d^k\|\|^2, \\ &= \frac{1}{n}\sum_{k=1}^{n} \|\mathbf{H}_d^k - \hat{\mathbf{H}}_d^k\|^2 / \|\mathbf{H}_d^k\|^2, \end{aligned} \tag{10}$$

which is equivalent to the NMSE defined in equation (9). As a result, for CSI matrices in different orders of magnitude, they are equally important in training of neural networks.

Second, by separating the power, we can make the CSI matrices in approximately the same order of magnitude as the input to the DL networks. By normalizing the different features in similar ranges of values, gradient descents can converge more quickly [14]. Besides, spherical space can also limit the solution domain for the CSI recovery.

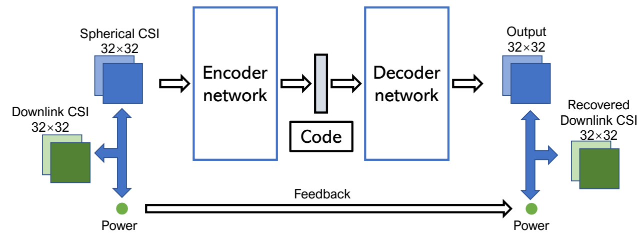


Fig. 1: Architecture of Spherical CSI Feedback

IV. Network Optimization

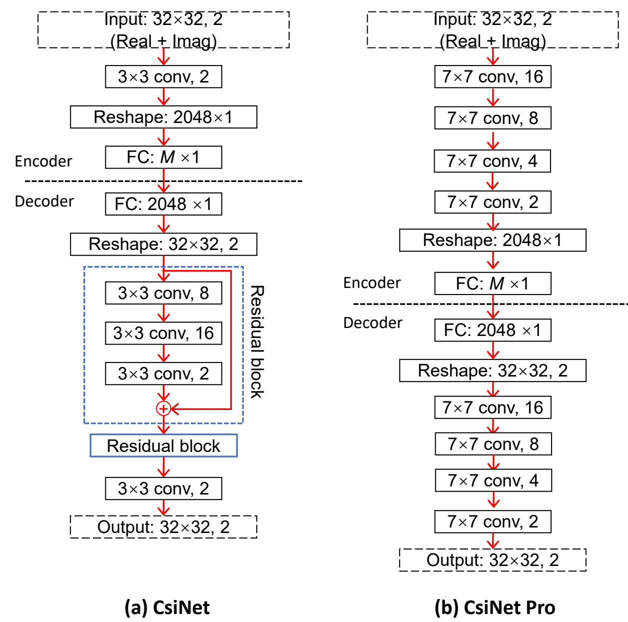


Fig. 2: Architecture of CsiNet and CsiNet Pro.

A DL-based CSI feedback framework named CsiNet was proposed in [6] to reduce massive MIMO downlink CSI feedback overhead. As shown in Fig. 2(a), CsiNet utilizes an encoder network which acts as a compression module and a corresponding decoder network which is responsible for CSI reconstruction. Treating CSI matrix as a virtual image, convolutional layer is used in both encoder and decoder to exploit its spatial and spectral correlation. Each CSI matrix is split into real and imaginary parts, rearranged into two channels as the input to the encoder. The encoder network consists of a 2-channel 3×3 convolutional layer and an $M\times1$ fully connected layer (FC) for dimension compression. The decoder consists of a FC for dimension decompression and two Residual blocks for CSI calibration. Although CsiNet shows substantial performance gain over some compressive sensing methods, the performance is still limited especially in the outdoor scenario.

To improve the performance of CsiNet, we propose an enhanced version of CsiNet, named CsiNet Pro.

On the one hand, CsiNet Pro utilizes a deeper encoder, which employs more convolutional layers to capture the features of CSI matrix. Autoencoder is utilized to learn a low-dimensional representation for a relatively high-dimensional dataset. By training the neural network to find the important features and ignore the noise details, an autoencoder can achieve good performance for dimensionality reduction. Consequently, the design of encoder for dimension compression is important. However, the encoder in [6]–[8] all utilized a simple encoder network, which consists of one convolutional layer and one fully connected layer. Different from above works, CsiNet Pro consists of 4 convolutional layers for feature extraction and 1 fully connected layer for dimension compression. Specifically, the 4 convolutional layers utilize the 7×7 kernel size to generate 16, 8, 4 and 2 feature maps, respectively (see Fig. 2(b)).

On the other hand, CsiNet Pro also enhances its decoder from two perspective. First, CsiNet Pro uses a different normalization range and output activation function. Different from the pixel values of image which are non-negative, the real part and imaginary part of CSI values can be positive or negative. Consequently, normalizing the input data within the range $[0, 1]$ may ignore some features related to the negative values. Different from previous works which normalize the CSI values to the range $[0, 1]$ and use "sigmoid" or "ReLU" as the activation function of last layer, CsiNet Pro normalizes the CSI values to the range [-1, 1], and use "tanh" as its activation function of last layers. Second, CsiNet Pro does not utilize the residual block to reduce the number of convolutional layers in the decoder part. Residual blocks are proposed in [15] to handle the vanishing gradient problem and were adopted in [6]–[8]. However, since the neural networks for CSI feedback are not as deep as ResNet [15], the vanishing gradient problem hardly arises. As a result, we remove the residual blocks and utilize 4 convolutional layers for codewords decoding as shown in Fig. 2(b).

We further integrate the CsiNet Pro into the spherical CSI feedback framework shown in Fig. 1 to enhance the CSI recovery accuracy, and name this network as SphNet.

V. Enhanced DualNet-MAG

DualNet-MAG utilizes the magnitude correlation between the uplink and downlink to improve the efficiency of CSI feedback [7]. Although DualNet-MAG can achieve good performance with the help of channel correlation, the CSI reconstruction accuracy is still limited due to the influence of the signal strength attenuation especially in the outdoor case.

In this paper, we propose an enhanced DualNet-MAG named as DualNet-Sph. DualNet-Sph uses the spherical CSI feedback architecture, and optimizes encoder and decoder networks for further performance improvement.

As shown in Fig. 3, the power of CSI matrix is separated from downlink CSI matrix and fed back individually to reduce the impact of signal strength attenuation on the training of autoencoder. To exploit the uplink and downlink correlation in the magnitude, magnitudes and phases of CSI are then separately encoded and fed back.

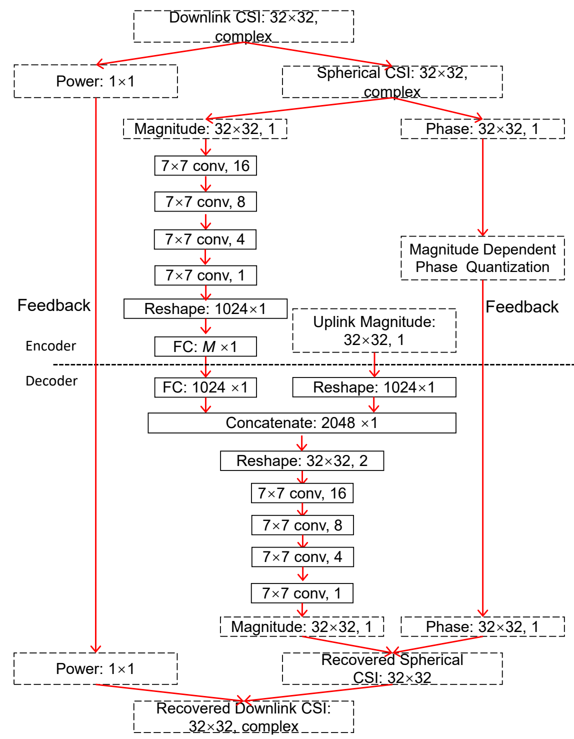
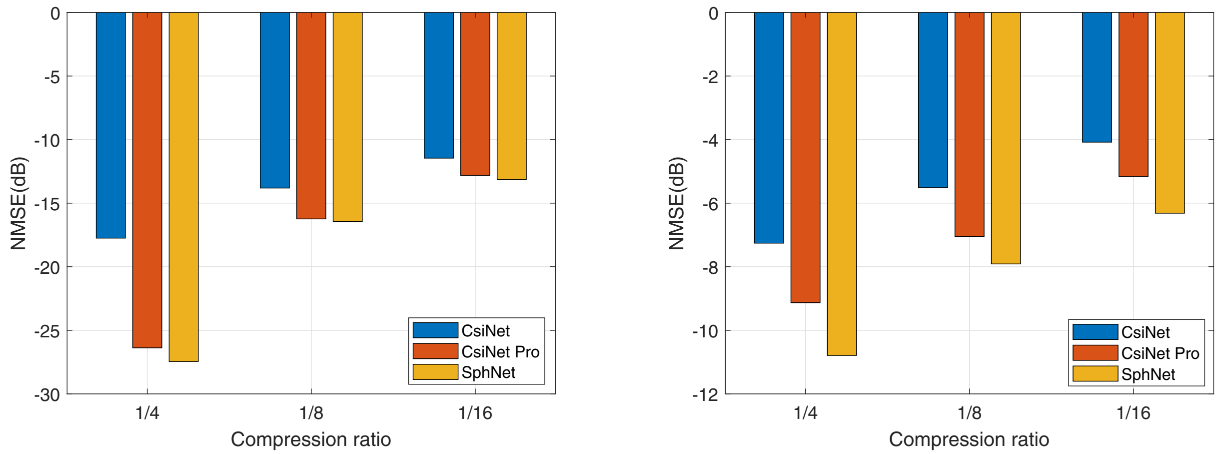


Fig. 3: Architecture of DualNet-Sph.

After phase separation, the CSI magnitudes are fed into the encoder network for dimension compression, while the uplink CSI magnitudes is utilized in the decoder network to help recover downlink CSI from the encoded codewords. Similar to the CsiNet Pro, the encoder of DualNet-Sph consists of 4 convolutional layers for feature extraction and 1 fully connected layer for dimension compression. The 4 convolutional layers exploit the 7×7 kernel size to generate 16, 8, 4 and 1 feature maps, respectively. In the decoder network, uplink CSIs are utilized to help recover downlink CSI from the received codewords. Specifically, the received codewords are first mapped into the original length using a fully connected layer. The uplink CSI matrices are reshaped to a vector and concatenated with the decompressed codewords. After concatenation, the output is reshaped into 2 feature maps as an input to 4 convolutional layers for downlink CSI magnitude recovery.

DualNet-Sph employs the same phase feedback method as DualNet-MAG. Owing to the poor correlation in phase, it is difficult to apply a similar DL structure to improve phase recovery accuracy. Thus, to avoid wasting feedback bandwidth, the downlink CSI phase is quantized and encoded for feedback. To limit the quantization error given limited bandwidth, a magnitude dependent phase quantization (MDPQ) is applied, where CSI coefficients with larger magnitude adopt finer phase quantization and vice versa. The details can be found in [7].

Finally, the magnitudes are combined with their correspond-



(a) Indoor (b) Outdoor

Fig. 4: NMSE (lower is better) comparison in different compression ratios for downlink-based CSI feedback. **CsiNet** is the previously proposed CNN-based autoencoder [6], **CsiNet-Pro** is the optimized network based on Fig. 2, and **SphNet** is the network which uses spherical normalization (see Fig. 1).

ing phases, and then recovered spherical CSIs are combined with their powers as fully recovered CSI coefficients.

VI. Experimental Evaluation

We assess the CSI feedback performance evaluation for two different massive MIMO scenarios using the COST 2100 model [16]:

1) An **indoor** environment using downlink 5.3GHz and uplink 5.1GHz, 0.1 m/s UE mobility, gNB at center of square area of length 20m
2) An **outdoor** environment using downlink 930MHz and uplink 850MHz, 1 m/s UE mobility, gNB at center of square area of length 400m

For all experiments, we utilize $N_b = 32$ antennas at the gNB and a single antenna at the UE. The gNB uses antennas arranged in a uniform linear array (ULA) with half-wavelength spacing. We use $N_f = 1024$ subcarriers and truncate the delay-domain CSI matrix to include the first $R_d = 32$ rows. For each environment, we generate a large dataset of 10^5 samples and split the dataset into $7 \cdot 10^4$ and $3 \cdot 10^4$ samples for training and testing sets, respectively. For training, we utilize Adam with a learning rate of 10^{-3} and a batch size of 200. Each network was trained for 1000 epochs using MSE (as Eq. (10)) as the loss function, and the figure of merit used to compare all networks was NMSE (as Eq. (9)).

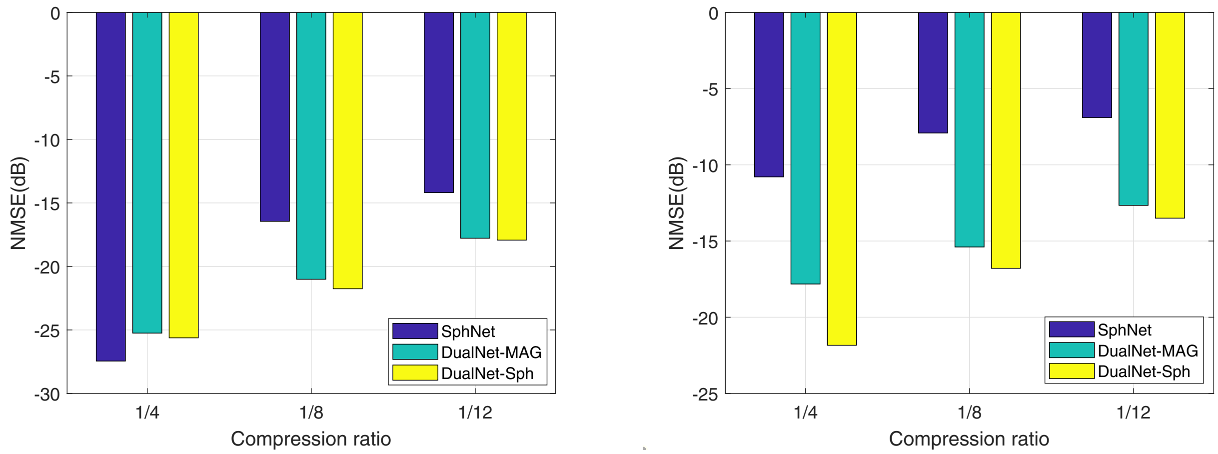
A. Spherical CSI Feedback and Network Architecture Optimization on CsiNet

We compare the performance of CsiNet, CsiNet Pro, and SphNet under three different compression ratios, $\frac{1}{4}$, $\frac{1}{8}$, and $\frac{1}{16}$. In [6], the authors demonstrate superior performance of CsiNet over three compressed sensing (CS) methods, namely LASSO ℓ_1 solver [17], TVAL3 [18], and BM3D-AMP [19], so it is reasonable to assume that any gains over CsiNet subsume gains made over these CS-based methods.

Fig. 4(a) shows the performance of the different networks for the indoor environment. In all cases, CsiNet Pro yields lower NMSE than vanilla CsiNet, and SphNet outperforms both networks. At a compression ratio of $\frac{1}{4}$, SphNet and CsiNet Pro outperform CsiNet by nearly 10 dB. In particular, we can notice that SphNet can achieve a more obvious performance improvement than CsiNet Pro in the outdoor environments (see Fig. 4(b)). The reason is that outdoor environment has a larger coverage range, which can make the power of CSI matrices span more orders of magnitude than the indoor case and degrade the CSI feedback performance. Consequently, by separating the power and using the spherical CSI feedback, SphNet can achieve noticeable performance gains.

B. Spherical CSI Feedback for Bi-directional Correlated DualNet

We also perform the previously described experiments using DualNet-Sph (see Fig. 3), the only discrepancy being that we vary use compression ratios $\frac{1}{4}$, $\frac{1}{8}$, and $\frac{1}{12}$. Fig. 5 shows the NMSE performance for DualNet-Sph, DualNet-MAG [7], and SphNet. For both the indoor and outdoor environments, the magnitude-based feedback networks have improving performance relative to SphNet as compression increases, except the case that the compression ratio is $\frac{1}{4}$ in the indoor scenarios. SphNet can achieve better performance than DualNet-MAG since phase quantization can result in the additional recovery error, especially since only 4 bits/phase are utilized for phase quantization on average. When the accuracy of magnitude recovery is relatively low, error from phase quantization can be neglected compared with the error from the magnitude. However, SphNet and the magnitude of DualNet can achieve higher accuracy compared with phase quantization when the compression ratio is $\frac{1}{4}$ in the indoor scenarios, which make the performance of DualNet relatively worse.



(a) Indoor

(b) Outdoor

Fig. 5: NMSE (lower is better) comparison in different compression ratios for bi-directional correlation-based CSI feedback. **SphNet** incorporates spherical normalization, **DualNet-MAG** incorporates bidirectional reciprocity [7], and **DualNet-Sph** incorporates both techniques.

Similar to the observations in Section VI-A, for the outdoor network, the influence of spherical CSI feedback is apparent. The improvement in NMSE for DualNet-Sph relative to DualNet-MAG is substantial, exceeding -20 dB at a compression ratio of $\frac{1}{4}$.

VII. Discussion

Based on the results highlighted in Sections VI-A and VI-B, the impact of power-based spherical normalization and network architecture optimization on CSI recovery accuracy is salient. The proposed enhancements yield improved CSI recovery accuracy for all compression ratios. In this work, we demonstrate enhancements for the training and structure of CNN-based autoencoders for CSI estimation for a single timeslot. Other efforts have utilized temporal correlation between timeslots to increase estimation accuracy [8]. Combining temporal feedback and the techniques outlined in this work could yield further performance gains.

References

[1] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of mimo channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, June 2003.

[2] B. Makki and T. Eriksson, "On hybrid arq and quantized csi feedback schemes in quasi-static fading channels," *IEEE Transactions on Communications*, vol. 60, no. 4, pp. 986–997, April 2012.

[3] H. Shirani-Mehr and G. Caire, "Channel state feedback schemes for multiuser mimo-ofdm downlink," *IEEE Transactions on Communications*, vol. 57, no. 9, pp. 2713–2723, Sep. 2009.

[4] X. Rao and V. K. N. Lau, "Distributed compressive csit estimation and feedback for fdd multi-user massive mimo systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3261–3271, June 2014.

[5] M. E. Eltayeb, T. Y. Al-Naffouri, and H. R. Bahrami, "Compressive sensing for feedback reduction in mimo broadcast channels," *IEEE Transactions on Communications*, vol. 62, no. 9, pp. 3209–3222, Sep. 2014.

[6] C. Wen, W. Shih, and S. Jin, "Deep learning for massive mimo csi feedback," *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 748–751, Oct 2018.

[7] Z. Liu, L. Zhang, and Z. Ding, "Exploiting Bi-Directional Channel Reciprocity in Deep Learning for Low Rate Massive MIMO CSI Feedback," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 889–892, 2019.

[8] T. Wang, C. Wen, S. Jin, and G. Y. Li, "Deep learning-based csi feedback approach for time-varying massive mimo channels," *IEEE Wireless Communications Letters*, vol. 8, no. 2, pp. 416–419, April 2019.

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

[10] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017, pp. 1800–1807.

[11] C. dos Santos and M. Gatti, "Deep convolutional neural networks for sentiment analysis of short texts," in *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*. Dublin, Ireland: Dublin City University and Association for Computational Linguistics, Aug. 2014, pp. 69–78. [Online]. Available: https://www.aclweb.org/anthology/C14-1008

[12] S. Gao, P. Dong, Z. Pan, and G. Y. Li, "Deep learning based channel estimation for massive mimo with mixed-resolution adcs," *IEEE Communications Letters*, vol. 23, no. 11, pp. 1989–1993, Nov 2019.

[13] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[14] Y. LeCun, L. Bottou, G. Orr, and K. Müller, "Efficient backprop," in *Neural networks: Tricks of the trade*. Springer, 2012, pp. 9–48.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[16] 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 Communications*, vol. 19, no. 6, pp. 92–99, December 2012.

[17] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004.

[18] C. Li, W. Yin, H. Jiang, and Y. Zhang, "An efficient augmented Lagrangian method with applications to total variation minimization," *Computational Optimization and Applications*, vol. 56, no. 3, pp. 507–530, 2013.

[19] C. A. Metzler, A. Maleki, and R. G. Baraniuk, "From denoising to compressed sensing," *IEEE Transactions on Information Theory*, vol. 62, no. 9, pp. 5117–5144, Sep. 2016.