

Exploiting Bi-Directional Channel Reciprocity in Deep Learning for Low Rate Massive MIMO CSI Feedback

Zhenyu Liu^{ID}, Lin Zhang^{ID}, and Zhi Ding^{ID}

Abstract—Channel state information (CSI) feedback is important for multiple-input multiple-output (MIMO) wireless systems to achieve their capacity gain in frequency division duplex mode. For massive MIMO systems, CSI feedback may consume too much bandwidth and degrade spectrum efficiency. This letter proposes a learning-based CSI feedback framework based on limited feedback and bi-directional reciprocal channel characteristics. The massive MIMO base station exploits the available uplink CSI to help recovering the unknown downlink CSI from low rate user feedback. We propose two deep learning architectures, DualNet-MAG and DualNet-ABS, to significantly reduce the CSI feedback payload based on the multipath reciprocity. DualNet-MAG and DualNet-ABS can exploit the bi-directional correlation of the magnitude and the absolute value of real/imaginary parts of the CSI coefficients, respectively. The experimental results demonstrate that our architectures bring an obvious improvement compared with the downlink-based architecture.

Index Terms—Massive MIMO, CSI feedback, multipath reciprocity, deep learning.

I. INTRODUCTION

MODERN wireless communication systems have made tremendous strides in utilizing the spatial diversity afforded by multiple-input multiple-output (MIMO) transceivers to improve radio link performance against poor channel conditions. In particular, massive MIMO systems have shown great promise for delivering high spectrum and energy efficiency for 5G wireless communication systems. To fully utilize the spatial diversity and multiplexing gains, however, massive MIMO transmitters must rely on sufficiently accurate channel state information (CSI). In a cellular network, this means that the base station (eNB or gNB) needs to acquire the downlink CSI in an accurate and timely manner.

In practice, gNBs in time division duplex (TDD) networks can rely on strong reciprocity between downlink and uplink

channels to estimate the downlink CSI. However, in frequency division duplex (FDD) systems, there often exists a weaker reciprocity in different frequencies. Consequently, gNB would require user equipment (UE) to provide downlink CSI feedback. For massive MIMO, such feedback can be substantial. Moreover, to cope with rapidly changing environment, UEs have to feed back CSI frequently, thereby consuming precious bandwidth resource.

The strong need for reducing massive MIMO CSI feedback in FDD systems has motivated a number of important studies. Generally, these works belong to three broad categories: 1) directly quantized CSI feedback; 2) compressed sensing (CS)-based feedback; 3) deep learning-based feedback. In directly quantized CSI feedback, UE quantizes its estimated downlink CSI into information bits for gNB [1], [2]. However, direct quantization requires too much uplink spectrum resource in massive MIMO. For CS-based feedback, it exploits the potential sparsity exhibited by massive MIMO CSI in certain transform domain to reduce feedback overhead [3], [4]. However, most CS-based approaches impose the strong channel sparsity and are often iterative, which causes additional delay.

The recent resurgence of artificial intelligence techniques offers some attractive alternatives. In fact, deep learning (DL) methods have already found applications in several aspects of wireless communications and networks, including resource allocation [5], channel estimation [6], signal classification [7], and low rate CSI feedback [8]. In particular, Wen *et al.* [8] developed a CSI feedback mechanism suitable for massive MIMO downlink and demonstrated effective improvement in downlink CSI reconstruction accuracy.

We aim to further investigate more effective downlink CSI feedback. We note that RF uplink and downlink channels in a wireless link are functions of physical environment such as multipaths and scatters. Even for FDD, existing works have demonstrated that certain of correlation exists between bi-directional channels. Thus, exploiting uplink CSI of FDD systems can potentially improve the estimation accuracy of downlink CSI. As shown in [9], the directional properties of radio channels are correlated in uplink and downlink FDD channels. Multipath delays in downlink and uplink frequency bands are same in [10]. In [11], it is illustrated that the correlation of shadowing effects also exists between uplink and downlink in FDD. Downlink channel covariance estimation can in fact benefits from its observed uplink covariance [12].

In this letter, we develop two enhanced DL network architectures, named as DualNet-MAG and DualNet-ABS, for estimating the downlink CSI based on low rate UE feedback by exploiting the multipath correlation between uplink and downlink channels in massive MIMO systems. Our basic principle is to utilize both the uplink CSI estimation available at the gNB and the low rate uplink feedback from the UE to recover the unknown downlink CSI. We apply both uplink CSI and downlink feedback to train the DL networks

Manuscript received November 13, 2018; revised January 24, 2019; accepted February 4, 2019. Date of publication February 11, 2019; date of current version June 19, 2019. The work of Z. Liu and L. Zhang was supported in part by the National Key Research and Development Program of China under Grant 2016YFB0100902, and in part by the China Scholarship Council. The work of Z. Ding was supported in part by the National Science Foundation under Grant 1702752. The associate editor coordinating the review of this paper and approving it for publication was C. Shen. (*Corresponding author: Lin Zhang.*)

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).

L. Zhang is with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: 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).

Digital Object Identifier 10.1109/LWC.2019.2898662

for decoding the efficient feedback codeword. The novelty of this letter is to establish that the bi-directional channel correlation can bring substantial performance improvement for downlink CSI estimation in FDD systems with limited uplink feedback.

II. SYSTEM MODEL

We examine a single-cell massive MIMO system, in which the gNB has $N_b \gg 1$ antennas and UEs have 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

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

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 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 conjugate transpose. The uplink received signal of the n -th subcarrier is given by

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

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 $\tilde{\mathbf{H}}_d = [\mathbf{h}_d^{(1)}, \dots, \mathbf{h}_d^{(N_f)}]^H \in \mathbb{C}^{N_f \times N_b}$ and $\tilde{\mathbf{H}}_u = [\mathbf{h}_u^{(1)}, \dots, \mathbf{h}_u^{(N_f)}]^H \in \mathbb{C}^{N_f \times N_b}$, respectively.

In this letter, we focus on downlink CSI feedback. Hence, we assume that $\tilde{\mathbf{H}}_d$ and $\tilde{\mathbf{H}}_u$ have been perfectly acquired by UE and gNB, respectively. Since CSI matrix $\tilde{\mathbf{H}}_d$ is $N_f \times N_b$, CSI feedback payload becomes huge for massive MIMO system with large N_b . To reduce feedback load, we take advantage of the important property that subchannel CSIs in uplink and downlink exhibit some common sparsity in the spatial delay domain [13]. We can transform a channel response matrix $\tilde{\mathbf{H}}_f$ from frequency domain to delay domain $\mathbf{H}_t$ using an IDFT, i.e.,

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

where $\mathbf{F}$ denotes the $N_b \times N_b$ unitary DFT matrix. After IDFT, most elements in the $N_f \times N_b$ matrix $\mathbf{H}_t$ are near zero except for the first $\hat{N}_f$ rows. Denote $\mathbf{H}_d$ and $\mathbf{H}_u$ as the first $\hat{N}_f$ rows of matrices after IDFT of $\tilde{\mathbf{H}}_d$ and $\tilde{\mathbf{H}}_u$, respectively. However, since $\hat{N}_f \times N_b$ is still a large number, further compression can start from the compressed CSI matrix $\mathbf{H}_d$ at the UE.

We apply an encoder and a decoder for CSI compression and reconstruction, respectively. Unlike existing works, our proposed architecture combines the downlink CSI for feedback with the available uplink CSI. Let $\hat{\mathbf{H}}_d$ be the reconstructed downlink CSI matrix. The encoder and decoder can be denoted, respectively, by

$$\mathbf{s} = f_{en}(\mathbf{H}_d) \quad (4)$$

$$\hat{\mathbf{H}}_d = f_{de}(\mathbf{s}, \mathbf{H}_u). \quad (5)$$

III. CORRELATION BETWEEN UPLINK AND DOWNLINK

To demonstrate the correlation between uplink and downlink, we generate CSI matrices using COST 2100 channel model [14]. A uniform linear array (ULA) is applied with half-wavelength spacing in an indoor environment with uplink and

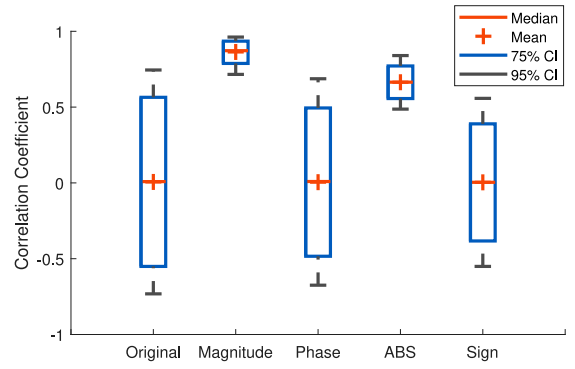


Fig. 1. Distribution of correlation coefficient between uplink and downlink CSI at various levels of Confidence Interval (CI).

downlink bands at 5.1 GHz and 5.3 GHz, respectively. We examine the CSI matrices in delay domain by evaluating their Pearson correlation coefficients. Since the CSI is complex, their real part and imaginary part correlations are evaluated.

As shown in Fig. 1, the correlation coefficients between uplink CSI and downlink CSI of the original (real/imaginary) format are rather erratic. This would seem to invalidate usefulness of the channel reciprocity in downlink CSI recovery. Nevertheless, a closer examination of the physics behind the CSI of different carriers reveals that in FDD, two different carrier frequencies can lead to random phase difference between them. We further recognize, based on FDD multipath channel models, that the magnitude in delay domain should exhibit much stronger correlation than phase.

Thus, we transform the CSI elements in the delay domain into polar coordinate to separately consider their magnitude and phase correlations. Fig. 1 demonstrates that the corresponding CSI magnitudes in uplink and downlink exhibit strong correlation whereas the corresponding phases have little correlation. In fact, by separating the signs of CSI's real and imaginary parts, Fig. 1 shows that the absolute values (ABS) of uplink and downlink channel coefficients are also positively correlated. However, their signs show little correlation.

These results demonstrate the practical reciprocal characteristics in the delay domain, clearly showing that the magnitudes of the CSI uplink/downlink coefficients are strongly correlated, whereas their phases and the signs of real/imaginary parts are not. These results motivate us to propose two DL network decoders for downlink CSI recovery. Specifically, UE should devote resources to encode the phases and the signs of downlink CSI for feedback, while gNB jointly utilizes UE's feedback and its own uplink CSI to decode the downlink CSI.

IV. DUALNET

We now propose two DL network architectures to reduce the CSI feedback payload of massive MIMO downlink.

A. DualNet-MAG

The DualNet-MAG utilizes the magnitude correlation to improve the efficiency of CSI feedback. Fig. 2 shows the general architecture of DualNet-MAG, in which magnitudes and phases of CSI are separately encoded and fed back.

After phase separation, the CSI magnitudes are fed into the DL network. The first encoder layer is a convolutional layer (CL) with batch normalization. This layer uses kernels with dimensions of 3×3 to generate one feature maps. We use the number of squares to denote the number of feature maps and use $X1 \times X2$ over a feature map to denote its size. After

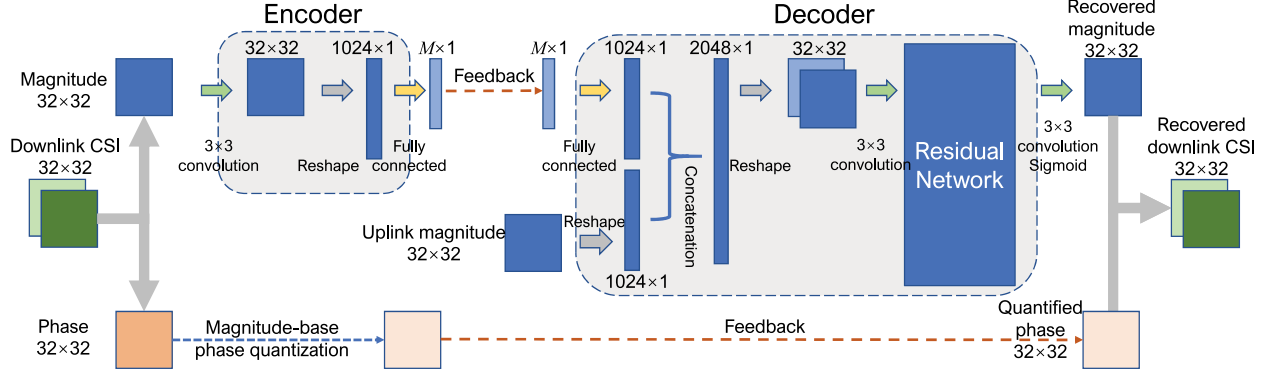


Fig. 2. Architecture of DualNet-MAG.

the CL, feature maps are reshaped into a vector. Then a fully connected layer generates the codeword $\mathbf{s}$ with the length M .

After the gNB receives the codeword $\mathbf{s}$, it uses the uplink CSI magnitudes to help recovering downlink CSI. The received $\mathbf{s}$ is first mapped into the original length using a fully connected layer. The uplink CSI matrices are reshaped to a vector before being conjugated with the modified codeword. Next, the conjugation layer is reshaped into 2 feature maps as an input to the residual network (RN). The RN in DualNet-MAG contains two residual blocks, which can effectively hand the vanishing gradient problem as proposed by [15]. Here the CL is used as the weight layer. The three CLs in each residual block utilize 3×3 kernels to generate 8, 16 and 2 feature maps, respectively. Zero padding is used to keep the size of feature maps at 32×32 . After RN, the next layer recovers the downlink CSI magnitudes. Finally, the magnitudes are combined with their corresponding phases as fully recovered CSI coefficients.

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. The loss function is the mean squared error (MSE), which is commonly used in CSI estimation and also works well here. Other differentiable loss functions can also be directly applied to DualNet.

Based on channel reciprocity, CSI magnitudes can be well compressed. However, the poor phase correlation makes it difficult to apply a similar DL structure to improve phase recovery. Thus, we let UE quantize and encode the downlink CSI phase for feedback. In polar coordinate, however, uniform phase quantization leads to unnecessarily fine quantization at low magnitude and coarse quantization for large magnitude. To avoid wasting feedback bandwidth, we limit the CSI quantization error by applying a magnitude dependent phase quantization (MDPQ) in which CSI coefficients with larger magnitude adopt finer phase quantization and vice versa. The gNB restores the quantified phase based on MDPQ using the recovered magnitude. Such MDPQ can keep the quantization error similarly small for all magnitudes. Note that the MDPQ codeword lengths vary with the magnitudes. Thus, the mean codeword length depends on the distribution of CSI magnitude. In our design, we estimate the cumulative distribution function (CDF) of magnitudes to determine the magnitudes corresponding to CDF value of 0.5, 0.7, 0.8, and 0.9, respectively. These four points divide the magnitudes into five ordered segments from small to large. We allocate 3, 4, 5, 6, 7 bits, respectively, to encode the phase of CSI with

magnitude inside the 5 segments. In our test, MDPQ codewords with the mean length of 4.1 bits can achieve the same MSE as a 6-bit uniform quantizer in both indoor and outdoor scenarios.

B. DualNet-ABS

Similar to DualNet-MAG, DualNet-ABS exploits the correlation between the uplink/downlink CSIs in terms of their real/imaginary absolute values to reduce CSI feedback. Unlike DualNet-MAG, DualNet-ABS separately feeds back the absolute values and signs of real/imaginary CSI components. Sending back uncorrelated sign information can be effective.

The architecture of DualNet-ABS is similar to DualNet-MAG. The difference lies in the number of feature maps and the size of fully connected layers. Since DualNet-ABS divides CSI matrix into real and imaginary parts, the number of feature maps and the size of fully connected layers are twice as large.

V. PERFORMANCE EVALUATION

A. Experiment Setup

For both training and testing, we randomly generate CSI matrices using the COST 2100 model [14]. Two scenarios are tested: (a) indoor channel with 5.1 GHz uplink and 5.3 GHz downlink bands; (b) semi-urban outdoor channel with 260 MHz uplink and 300 MHz downlink bands. In future works, we will extend the current work to more cases, such as pilot contamination in multi-cell coverage scenarios. We place gNB at the center of a square area of lengths 20 m for indoor coverage and 400 m for outdoor coverage, respectively. Both uplink and downlink bandwidths are 20 MHz. 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 into the delay domain, only the first 32 rows are kept due to sparsity. The training sample size is 70,000 and testing sample size is 30,000. The values of epoch and batch size are set to 600 and 200, respectively.

Normalized MSE (NMSE) is utilized to evaluate the accuracy of CSI recovery, which is defined as

$$\text{NMSE} = \frac{1}{n} \sum_{k=1}^n \|\mathbf{H}_d^k - \hat{\mathbf{H}}_d^k\|^2 / \|\mathbf{H}_d^k\|^2, \quad (6)$$

where k and n are, respectively, the index and total number of samples in the testing set. Compression ratio is defined as the ratio between total bits after the proposed encoding and original bits required by pure CSI quantization feedback.

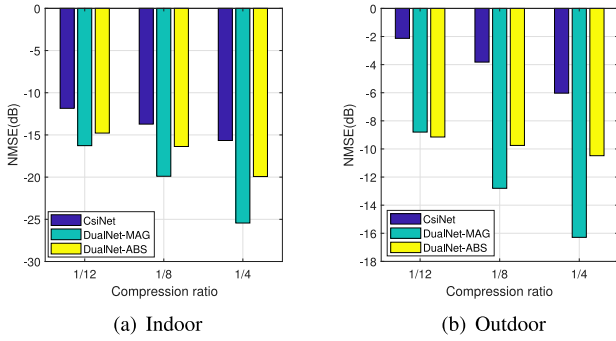


Fig. 3. NMSE comparison in different compression ratios.

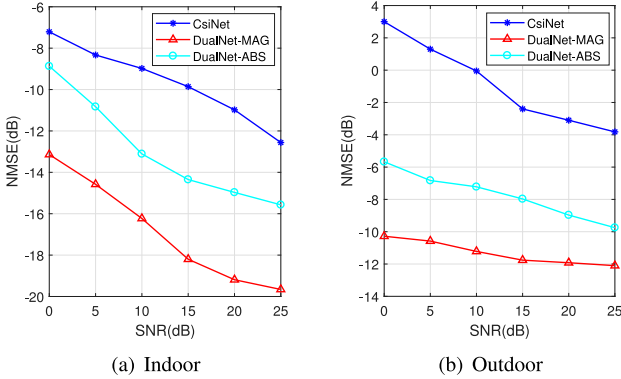


Fig. 4. CSI recovery comparison at different CSI error levels.

We compare the DualNet-MAG and DualNet-ABS with the CsiNet [8]. Recall that CsiNet uses a deep learning architecture for CSI feedback compression with only downlink CSI. Since the three network architectures and feedback modes differ, compression ratios should be defined separately. Let ξ be the codeword length for the real/imaginary part of one CSI coefficient. Let λ be the average codeword length for one quantified phase. For CsiNet, DualNet-MAG, and DualNet-ABS, compression ratios are calculated, respectively, as

$$\frac{M \cdot \xi}{2048 \cdot \xi}, \frac{M \cdot \xi + 1024 \cdot \lambda}{2048 \cdot \xi}, \text{ and } \frac{M \cdot \xi + 2048 \cdot 1}{2048 \cdot \xi}.$$

Since standard implementations of neural network typically use 32-bit floating-point for real values, ξ is set to 32 bits.

B. Accuracy and Robustness Comparison

We compare the CSI accuracy achieved by CsiNet, DualNet-MAG, and DualNet-ABS under three compression ratios (CRs) of 1/12, 1/8 and 1/4. Fig. 3 provides the NMSE results for both indoor and outdoor channels. Leveraging uplink CSI for massive MIMO, DualNet-MAG and DualNet-ABS clearly outperform CsiNet in both scenarios. In Fig. 3a, DualNet-MAG of CR = 1/12 outperforms CsiNet of CR = 1/4. Note that DualNet-MAG consistently outperforms DualNet-ABS in indoor scenario, since CSI magnitudes exhibit stronger correlation between uplink and downlink over the absolute values of real/imaginary parts. Fig. 3b shows that for the outdoor channel, both DualNet-MAG and DualNet-ABS are much more superior. Although DualNet-MAG outperforms DualNet-ABS in most cases, DualNet-ABS can achieve a smaller CR since the total space for the phases is twice as large as the space for the signs. Moreover, when the MSE is large (e.g., CR = 1/12 in

outdoor), DualNet-MAG performs worse since the low accuracy of recovered magnitudes further degrades the accuracy of phases in MDPQ.

To further test the robustness of CsiNet, DualNet-MAG, and DualNet-ABS, we examine the reconstruction performance in the presence of channel estimation error by the UE. We include an additive random white Gaussian error to the CSI at the UE encoder. We test different levels of errors in terms of signal-to-noise ratio (SNR). Fixing the CR to 1/8, we compare the massive MIMO CSI recovery performance for both indoor and outdoor channels. As the results in Fig. 4 show, both DualNet-MAG and DualNet-ABS remain robust and superior.

VI. CONCLUSION

This letter presents two DL-based low rate CSI feedback solutions for massive MIMO wireless communications. Recognizing and utilizing the multipath reciprocity between uplink and downlink CSI in the delay domain, our DL-based architectures can significantly improve the feedback efficiency and the recovery accuracy of downlink CSI. The proposed DualNet-MAG and DualNet-ABS represent viable practical solutions in both indoor and outdoor wireless networks.

REFERENCES

- [1] B. Makki and T. Eriksson, "On hybrid ARQ and quantized CSI feedback schemes in quasi-static fading channels," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 986–997, Apr. 2012.
- [2] H. Shirani-Mehr and G. Caire, "Channel state feedback schemes for multiuser MIMO-OFDM downlink," *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2713–2723, Sep. 2009.
- [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, Jun. 2014.
- [4] M. E. Eltayeb, T. Y. Al-Naffouri, and H. R. Bahrami, "Compressive sensing for feedback reduction in MIMO broadcast channels," *IEEE Trans. Commun.*, vol. 62, no. 9, pp. 3209–3222, Sep. 2014.
- [5] H. Ye and G. Y. Li, "Deep reinforcement learning for resource allocation in V2V communications," in *Proc. IEEE ICC*, May 2018, pp. 1–6.
- [6] 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.
- [7] T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 168–179, Feb. 2018.
- [8] C.-K. Wen, W.-T. Shih, and S. Jin, "Deep learning for massive MIMO CSI feedback," *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 748–751, Oct. 2018.
- [9] K. Hugi, K. Kalliola, and J. Laurila, "Spatial reciprocity of uplink and downlink radio channels in FDD systems," in *Proc. COST TD*, Espoo, Finland, vol. 273, no. 2, May 2002, p. 66. [Online]. Available: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.492.8048>
- [10] Y. Liang and F. P. S. Chin, "Downlink channel covariance matrix (DCCM) estimation and its applications in wireless DS-CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 2, pp. 222–232, Feb. 2001.
- [11] S. S. Szyszczkiewicz, H. Yanikomeroglu, and J. S. Thompson, "On the feasibility of wireless shadowing correlation models," *IEEE Trans. Veh. Technol.*, vol. 59, no. 9, pp. 4222–4236, Nov. 2010.
- [12] A. Decurvinge, M. Guillaud, and D. T. M. Stock, "Channel covariance estimation in massive MIMO frequency division duplex systems," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, San Diego, CA, USA, Dec. 2015, pp. 1–6.
- [13] 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.
- [14] L. Liu et al., "The cost 2100 MIMO channel model," *IEEE Wireless Commun.*, vol. 19, no. 6, pp. 92–99, Dec. 2012.
- [15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE CVPR*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.