Exploiting Partial FDD Reciprocity for Beam-Based Pilot Precoding and CSI Feedback in Deep Learning

Yu-Chien Lin*, Ta-Sung Lee, Fellow, IEEE, and Zhi Ding, Fellow, IEEE

Abstract—Massive MIMO systems can achieve high spectrum and energy efficiency in downlink (DL) based on accurate estimate of channel state information (CSI). Existing works have developed learning-based DL CSI estimation that lowers uplink feedback overhead. One often overlooked problem is the limited number of DL pilots available for CSI estimation. One proposed solution leverages temporal CSI coherence by utilizing past CSI estimates and only sending channel state information-reference symbols (CSI-RS) for partial arrays to preserve CSI recovery performance. Exploiting CSI correlations, FDD channel reciprocity is helpful to base stations with direct access to uplink CSI. In this work, we propose a new learning-based feedback architecture and a reconfigurable CSI-RS placement scheme to reduce DL CSI training overhead and to improve encoding efficiency of CSI feedback. Our results demonstrate superior performance in both indoor and outdoor scenarios by the proposed framework for CSI recovery at substantial reduction of computation power and storage requirements at UEs.

Index Terms—CSI feedback, FDD reciprocity, pilot placement, massive MIMO, deep learning.

I. INTRODUCTION

MUltiple-INPUT multiple-output (MIMO) technology and massive MIMO are vital to 5G and future generations of wireless systems for improvement of spectrum and energy efficiency. The power of massive MIMO hinges on accurate downlink (DL) channel state information (CSI) at the basestation gNodeB (gNB). Without the benefit of uplink/downlink channel reciprocity in time-division duplexing (TDD) systems, gNB of frequency-division duplexing (FDD) systems typically relies on user equipment (UE) feedback to acquire DL CSI. The extraordinarily large number of DL transmit antennas envisioned in millimeter wave or terahertz bands in future networks [1] places a tremendous amount of feedback burden on uplink (UL) resources such as bandwidth and power. As a result, CSI feedback reduction is crucial to widespread deployment of massive MIMO technologies in FDD systems.

The 3rd Generation Partnership Project (3GPP) recently released the features of Release 18 [2] which embraces artificial intelligence (AI) and machine learning (ML). Specifically, AI and ML are expected for enhancement of CSI feedback (e.g., overhead reduction and improved estimation accuracy). Since CSI in most environments has limited delay spread and can be viewed as sparse, CSI feedback by UEs can take advantage of such low dimensionality for CSI feedback compression. To extract CSI sparsity for improved feedback efficiency, the work [3] first proposed a deep autoencoder (AE) framework by deploying encoders and a decoder at UEs and the serving base station, respectively, for CSI compression and recovery. This and other related works have demonstrated significant performance improvement of CSI recovery with the use of deep learning AE [4], [5], [6].

In addition to AE for direct DL CSI feedback and recovery, recent works leveraged correlated channel information such as past CSI [7], [8], CSI of nearby UEs [9], and UL CSI [10], [11], [12], [13] to improve the recovery of DL CSI at base stations. Specifically, physical insights considering slow temporal variations of propagation scenarios, similar propagation conditions of similarly located UEs, and similarity of UL/DL radiowave paths reveal significant temporal, spatial, and spectral CSI correlations respectively. More strikingly, UL CSI is generally available at gNB in existing FDD wireless networks and is easier to utilize in practice. In addition, FDD reciprocity in magnitudes is not only shown from data generated by CSI models [11] but was also later verified in measurement [14]. In the works [11], [12], [13] using AE architectures, FDD magnitude reciprocity in angle and delay domains is introduced at the decoder to improve DL CSI recovery. Other related works also considered antenna array geometry to exploit the UL/DL angular reciprocity to improve DL CSI estimation in FDD wireless systems [15], [16]. The work [15] exploited UL/DL angular reciprocity in designing an adaptive dictionary learning for seeking the sparse representation of DL CSIs for feedback. The reciprocity is also utilized for directional training to enhance DL CSI estimation in [16].

Instead of CSI recovery, a related approach [17], [18], [19], [20] is to exploit FDD reciprocity and angular sparsity to directly determine precoding matrix for reducing...
feedback overhead. The authors [17] propose an AoD-adaptive subspace codebook framework for efficiently quantizing and feeding back DL CSI. The 5G (NR) supports Type I [18] and Type II [19] codebooks corresponding to low- and high-resolution beams, respectively. The optimum serving beam can be selected by feeding back a predetermined codebook with the largest response between the UE and gNB. Similarly, instead of feeding back predetermined codebook, another idea in [20] is for UE to feed back compressed singular vectors corresponding to the dominant singular values for precoding matrix optimization. According to the 3GPP Release 17 technical report [21], [22], industries have recently tested the eType II Precoding Subset (PS) codebook and found that utilizing DL/UL reciprocity of angle and delay can generally enhance performance.

Importantly, DL CSI estimation accuracy for UEs depends on factors such as channel fading properties and reference signal (RS) placement. In systems with large-scale array, more resources allocated to CSI-RS would improve accuracy but reduce spectrum efficiency, and practical systems often use sparse CSI-RS allocation. Only a few studies, such as [23], [24], and [25], have considered sparse CSI-RS availability in CSI feedback mechanism design. Reference [23] proposed a deep learning partial CSI feedback framework that reduces RS resource overhead by leveraging temporal CSI correlation. Reference [24] optimized DL pilot symbols based on UL CSI without reducing CSI-RS resources, but dynamic exchange of optimized pilot symbols between the gNB and UE is required, which is incompatible with the current use of predefined CSI-RS. Although the work [26] presents an AE-based UE-specific optimal pilot placement pattern, it suffers from the same issue of incompatibility with existing communication systems.

In this work, we aim to decrease DL CSI-RS and UL feedback overhead while maintaining DL CSI accuracy by utilizing available UL CSI. We create an efficient and reconfigurable deep learning beam-based CSI feedback framework through UL/DL angular reciprocity for FDD wireless systems. Our contributions are summarized as follows:

- We first propose a framework BSdualNet0 for selecting a number of significant beams to generate a low-dimensional CSI representation by exploiting the FDD UL/DL reciprocity in beam response magnitudes, leading to lower DL CSI training and UL feedback overhead.

- The new frameworks, BSdualNet and BSdualNet-MN, require no complex encoder at UEs and better utilize FDD reciprocity by not only feeding UL CSI magnitudes as deep learning inputs [11], [12], but also designing a beam-based precoding matrix according to high similarity of UL/DL beam response magnitudes.

- BSdualNet-FR, reduces DL CSI training overhead by reconfiguring CSI-RS placement to decrease pilot resource density or the number of time-frequency resources. It also includes an UL feedback compression module to further lower UL feedback overhead.

- The reduction of DL CSI training overhead in the framework can significantly lessen the computation and storage burdens related to the compression by the low cost UEs given the input size reduction of the compression module.

- Previous CSI feedback approaches based on AE only utilize FDD reciprocity at the decoder. In contrast, the proposed framework exploits FDD reciprocity not only at the decoder for better DL CSI recovery but also at the encoder for efficient pilot dimension reduction.

- The 3GPP considers the utilization of FDD reciprocity in precoder design and the enhancement of CSI feedback through AI as important options for the future, particularly in Release 18. We believe that our AI-powered UL-assisted CSI feedback frameworks, BSdualNet-FR, could be a good starting point to bridge these two 3GPP goals and pave the way for future research development in this area.

We let \((\cdot)^H, (\cdot)^T\) denote conjugate transpose and transpose operations, respectively. \((\cdot)^*\) denotes complex conjugate. The \(i\)-th column of \(N \times N\) identity matrix \(I\) is the unit vector \(e_i\).

II. SYSTEM MODEL

We consider a single-cell MIMO FDD link in which a gNB using a \(N_{H} \times N_{V}\) uniform planar array (UPA) with \(N_b = N_{V} \cdot N_{H}\) antennas communicates with single antenna UEs. In FDD systems, UEs estimate DL CSIs and feedback to the serving base station after quantization. Then, gNB recovers the DL CSI based on the feedback. For brevity, we use \(\hat{h}_{DL}\) and \(\hat{h}_{UL}\) in the following section to represent the estimated CSI obtained at UE and gNB, respectively. Focusing on a specific UE, the DL subband consists of \(K\) RBs within the bandwidth for CSI-RS placement. We assume channels within an RB to be under slow, flat and block fading. As shown in Fig. 1, there are \(N_t \times N_o\) time-frequency resource elements (REs) in a specific RB (\(N_t\) subcarriers and \(N_o\) OFDM symbols). Since the same processing procedures are applied for every RB, without loss of generality, we only discuss the processing in a single RB in this section.

A. DL CSI Recovery

Given that the gNB assigns \(N_b\) REs for DL CSI training for \(N_b\) antennas, the received signal vector \(y_{DL} \in \mathbb{C}^{N_b \times 1}\) at

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig1.png}
\caption{Resource block configuration. There are \(N_t \times N_o\) time-frequency REs in a RB (\(N_t\) subcarriers and \(N_o\) OFDM symbols). Pilots are allowed to be placed at those REs in designated region (surrounded by the black frame). There are \(I \) available REs for pilot placement in this illustration.}
\end{figure}
UE can be expressed as
\[ y_{DL} = S_{DL,N_b} \cdot h_{DL} + n_{DL}, \]
where \( h_{DL} = \text{vec}(H_{DL}) \in \mathbb{C}^{N_b \times 1} \) denotes the DL CSI vector whereas \( S_{DL,N_b} = \text{diag}(s_{DL}) \in \mathbb{C}^{N_b \times N_b} \) denotes the CSI-RS training symbol matrix which is diagonal matrix with diagonal entries of training symbols \( s_{DL}^{(n)} \neq 0, n = 1, \ldots, N_b \). \( n_{DL} \in \mathbb{C}^{N_b \times 1} \) denotes the additive noise. \( H_{DL} \in \mathbb{C}^{L \times N_b} \) denotes the DL CSI matrix before reshaping. With the assumption of perfect channel estimation, from known training symbols in \( S_{DL,N_b} \), the UE can estimate its DL CSI for feedback to gNB via \( h_{DL} \approx \hat{h}_{DL} \).

**B. BS Precoding and DL CSI Recovery**

Existing wireless systems [27], [28] have applied beamforming/precoding techniques to CSI-RS symbols for beam selection, DL CSI estimation, or resistance to attenuation in high frequencies. According to [29], we can find \( N_b \) orthogonal beams to construct an “orthogonal beam matrix” \( B = [b^{(1)} \ b^{(2)} \ldots \ b^{(N_b)}] \). Applying the OBM to the CSI-RS matrix \( S_{DL,N_b} \) in the digital beamforming module, the UE receives signals at different REs
\[ y_{DL} = S_{DL,N_b}B^T h_{DL} + n_{DL}. \]
From the orthogonality of the OBM, DL CSI can be reconstructed at the gNB from the BS quantized UE feedback \( \bar{g}_B = \tilde{Q}(\tilde{h}_{BS,DL} = B^TH_{DL}) \in \mathbb{C}^{N_b} \) via \( h_{DL} = B^T \bar{g}_B \in \mathbb{C}^{N_b} \) where \( \tilde{Q}(\cdot) \) denotes a differentiable quantizer. More details about \( \tilde{Q}(\cdot) \) can be found in [10].

Given the angular sparsity of DL CSIs, especially for DL CSIs in line-of-sight (LOS) scenarios, the beam space (BS) DL CSI \( \tilde{h}_{BS,DL} = B^T \bar{g}_B \) can be approximated as a \( L \)-sparse vector and thus DL CSI \( h_{DL} \) can be approximated according to the most significant \( L(L < N_b) \) beams as follows:
\[ \hat{h}_{DL} = B_{S}^T \bar{g}_{BS} \]
where \( B_S \in \mathbb{C}^{N_b \times L} \) and \( \bar{g}_{BS} \in \mathbb{C}^{L \times 1} \) respectively denote the significant beam matrix consisting of the steering vectors of the most significant \( L \) orthogonal beams, and the corresponding quantized beam responses.\(^1\) Relying on \( L \) significant beams, the gNB only need to assign \( L(<N_b) \) REs for CSI-RS in DL to reduce UL feedback. We denote the heuristic DL CSI approximation approaches as BS-DL and BS-UL when using DL CSI and UL CSI to obtain the significant beam matrix \( B_S \), respectively. Note that BS-DL is an ideal approach with the assumption that we already have perfect DL CSI.

Typically, the \( L \) significant beams could be found through beam training or direction finding [30], [31] by utilizing additional bandwidth and power resources. Fortunately, the FDD UL/DL reciprocity in magnitudes of angular CSI [11] can help gNB implement this beam selection process by relying on the available UL CSI at gNB.

---

\(^1\)According to [30], in propagation channels with low angular spread, only few significant beams contribute to most DL CSI energy in beam domain. This is also shown in Table I.
and zeros as initial BS DL CSI estimate according to the index set of the selected beams $\Omega_B$. The sparse map $M$ and local UL CSI magnitudes $|H_{UL}|$ form inputs to a deep learning network for estimating the missing elements in the sparse map for DL CSI refinement. The convolutional neural network (CNN) generates refined DL beam domain CSI $\hat{H}_{DL}$.

In this BS DL CSI recovery framework, the gNB assigns $L$ orthogonal beams to $L$ REs and recovers the full BS DL CSI based on the feedback of the $L$ beam responses from UEs via correlation between adjacent beam responses.

**B. BS Precoding and DL CSI Recovery**

We also develop a BS DL CSI recovery framework which assigns all orthogonal beams to $L$ REs ($L < N_b$). Instead of utilizing a single beam for each RE, a combination of weighted beams is applied. Let us denote an $L \times N_b$ beam merging matrix

$$T = \begin{bmatrix}
    t_1^T \\
    t_2^T \\
    \vdots \\
    t_L^T
\end{bmatrix} \in \mathbb{C}^{L \times N_b}, \quad t_i = \begin{bmatrix}
    t_{1,i} \\
    \vdots \\
    t_{N_b,i}
\end{bmatrix}. \tag{4}
$$

The received signal vector at UE is expressed as

$$y_{DL} = \begin{bmatrix}
    \sum_{i=1}^{N_b} t_{1,i} h_{DL}^T b_i^{(1)} s_{DL}^{(1)} + n_{DL} \\
    \sum_{i=0}^{N_b-1} t_{i+1} h_{DL}^T b_i^{(2)} s_{DL}^{(2)} \\
    \vdots \\
    \sum_{i=0}^{N_b-1} t_{L,i} h_{DL}^T b_i^{(L)} s_{DL}^{(L)}
\end{bmatrix} + n_{DL}.$$

$$= \begin{bmatrix}
    h_{DL}^T B_{\Omega_B}^{(1)} \Theta_{DL} \\
    h_{DL}^T B_{\Omega_B}^{(2)} \Theta_{DL} \\
    \vdots \\
    h_{DL}^T B_{\Omega_B}^{(L)} \Theta_{DL}
\end{bmatrix} + n_{DL}$$

$$= S_{DL,L}TB^T h_{DL} + n_{DL} = S_{DL,L} \Theta_{BS,DL} + n_{DL}, \tag{5}$$

where $T$ is used to reduce the required REs and to find a compact representation of DL CSI. $h_{BS,DL} = B^T h_{DL}$ denotes the DL CSI vector in beam domain. Since the recovery loss mainly attributes to the quantization and compression error instead of CSI estimation discrepancy, we adopt a common assumption without loss of generality in our benchmarks [3], [6], [8], [12] that UEs provide perfect channel estimation for simplicity. The raw and quantized response vectors of the merged beam responses are denoted by $g_{FB} = Th_{BS,DL}$ and $\bar{g}_{FB} = Q(g_{FB})$, respectively.

Our goal is to find a beam merging matrix $T \in \mathbb{C}^{N_b \times L}$ and a mapping function $f_{re}$ for recovering the DL CSI based on the quantized feedback vector via the principle of

$$\arg\min_{T,\Theta_{BS}} ||B^* f_{re}(Q(Th_{BS,DL})) - h_{DL}||^2_2 \tag{6}$$

where $\Omega_{re}$ denotes the deep learning model parameters to be optimized. Following this principle, the detailed design and architecture of an UL CSI-aided feedback framework for DL CSI estimation will follow in the next section. Since the encoding function $Q(\cdot)$ [10] and each layer of the combining network are differentiable, the deep model parameters $\Omega_{re}$ can be optimized through backpropagation.
which reduces the required REs for CSI-RS of DL MIMO channels and reduces UL feedback overhead.

A. General Architecture

Consider a wireless communication system with $L$ REs assigned in each RB for CSI-RS placement. For CSI feedback reduction, we first design a beam merging matrix $T$ to match $N_b$ orthogonal beams with different weights to the $L$ REs that carry CSI-RS. We use a beam merging network that use UL CSI magnitudes in beam domain as inputs. Owing to the high correlation between magnitudes of UL and DL CSIs in beam domain, the beam merging network learn to assign suitable weights to orthogonal beams according to the corresponding BS UL CSI magnitudes $|\mathbf{B}^T \mathbf{h}_{UL}|$ that are locally available at gNB.

Next, we apply the beam merging matrix $T$ to $L$ CSI-RS symbols on the $L$ REs. The linear mapping matrix $T$ instead of a general or non-linear mapping function $f : \mathbb{C}^{N_b} \rightarrow \mathbb{C}^L$ for pilot dimension reduction provides the advantage of simpler implementation and easier decoupling of CSI-RS symbols. Consequently, the effective channels at UEs after CSI estimation would be the weighted sum of beam responses as estimate of the full CSI at downlink. Obtaining effective channels, the UE simply quantizes and feeds back the channel information to the gNB. The gNB recovers DL CSI by sending the quantized feedback and the known beam merging matrix $T$ into the proposed deep learning decoder network. For simplicity, Fig. 3 shows the general architecture of the proposed CSI feedback framework for a single UE, though the same principle applies for multiple UEs.

Unlike previous works, our new framework does not require any encoder at UE to store and compress full DL CSI. This is beneficial to UE devices with limited computation, storage, and/or power resources. Moreover, we reduce the DL overhead of CSI-RS and provide higher spectrum efficiency while previous frameworks require REs proportional to its transmit antennas. The number of required REs for DL CSI training in our framework heavily depends on the sparsity in beam domain. Since the beam sparsity increases with larger array, this would bring more benefits in reducing DL CSI overhead when considering a large-scale transmit antenna array.

B. BSdualNet

Fig. 4 shows the proposed CSI feedback framework, BSdualNet, in multi-user scenarios (i.e., $N$ UEs). BSdualNet consists of three learning networks at gNB which serve on distinct objectives:

- Beam merging network: it designs an matrix $T$ which is applied to DL CSI training for reducing the required REs and uplink feedback overhead while maintaining accurate CSI recovery. With the aids of partial FDD UL/DL reciprocity, the beam merging matrix $T$ transforms effective BS CSI at UEs into a compressive representation.
- Recovery network: it estimates the full BS CSI according to the quantized estimated beam responses from UEs.
- Combing network: it refines the magnitudes of DL BS CSI by using the known magnitudes of UL BS CSI based on partial FDD UL/DL reciprocity.

As shown in Fig. 5, we aggregate and reshape the magnitudes of BS UL CSIs of each UE $\mathbf{H}_{BS,UL}^{(i)} \in \mathbb{C}^{NH \times NV}$ into a tensor $[\mathbf{H}_{BS,UL}]_i \in \mathbb{C}^{NH \times NV \times N}$, which is sent to the beam merging network at gNB. The beam merging deep learning network (Fig. 5) consists of four $3 \times 3$ circular convolutional layers with 16, 8, 4, and 2 channels, respectively, to learn the importance of different orthogonal beams according to the spatial structures of UL beam domain CSI magnitudes. Given the circular characteristic of BS CSI matrices, we introduce circular
convolutional layers to replace traditional convolution. Subsequently, a fully connected (FC) layer with $2N_bL$ elements is included to generate desired dimension after reshaping (Recall that $T$ is a complex matrix with size of $N_b \times L$). After CSI estimation at UEs, the gNB receives the $N$ copies of quantized feedbacks from $N$ UEs and obtains quantized feedbacks $\vec{g}^{(i)}_{FB}$ in $\mathbb{C}^{2L}$, $i = 1, 2, \ldots, N$.

Now we focus on the network at gNB. For the $i$-th UE, we forward the received feedback $\vec{g}^{(i)}_{FB}$ to a FC layer with $2N_b$ elements. After reshaping the feedback data into a matrix of size $N_H \times N_V \times 2$, we use four $3 \times 3$ circular convolutional layers with 16, 8, 4, and 2 channels and activation functions to generate initial BS DL CSI estimate $\vec{H}^{(i)}_{\text{BS,DL},\text{ini}} = f_{re}(\vec{g}^{(i)}_{FB}) \in \mathbb{C}^{N_H \times N_V}$. Next, the gNB forwards the initial BS DL CSI estimate $\vec{H}^{(i)}_{\text{BS,DL},\text{ini}}$ together with the corresponding BS UL CSI magnitudes $|\vec{h}^{(i)}_{\text{BS,UL}}|$ to the combining network for final DL CSI estimation $\vec{H}^{(i)}_{\text{BS,DL}}$. The combining network $f_c$ uses $N_B$ residual blocks, each block contains the same design of circular convolutional layers and activation functions as the network for DL CSI recovery.

Since all layers and quantization function are differentiable, the BSdualNet is optimized via backpropagation and gradient descent to update the network parameters $\Theta_{bm}$, $\Theta_{re}$ and $\Theta_{c}$ of non-linear beam merging, recovery, and combining networks $f_{bm}$, $f_{re}$ and $f_c$:

$$\arg \min_{\Theta_{bm}, \Theta_{re}, \Theta_{c}} \left\{ \sum_{i=0}^{N-1} \left\| \hat{h}^{(i)}_{\text{BS,DL}} - h^{(i)}_{\text{BS,DL}} \right\|_2^2 \right\},$$

where $\hat{h}^{(i)}_{\text{BS,DL}} = \text{vec}(\hat{H}^{(i)}_{\text{BS,DL}})$ and $h^{(i)}_{\text{BS,DL}} = \text{vec}(h^{(i)}_{\text{BS,DL}})$ denote the vectorized estimated and original BS DL CSIs. $H^{(i)}_{\text{BS,DL}} = f_c(f_{re}(\vec{g}^{(i)}_{FB})), |\vec{h}^{(i)}_{\text{BS,DL}}| = Q(T(\hat{h}^{(i)}_{\text{BS,DL}}))$ denotes the quantized BS DL CSI and the beam merging matrix is given by $T = f_{bm}(H^{(1)}_{\text{BS,UL}}|, H^{(2)}_{\text{BS,UL}}|, \ldots, H^{(N)}_{\text{BS,UL}}|)$. Note that the superscript $(i)$ denotes the UE index, $H^{(i)}_{\text{BS,DL}}$ and $H^{(i)}_{\text{BS,UL}} \in \mathbb{C}^{N_H \times N_V}$ denote original BS DL and UL CSIs at the $i$-th UE.

C. BSdualNet-MN

In BSdualNet, the beam merging network provides a beam merging matrix $T$ to generate an efficient representation of the convoluted responses of all orthogonal beams. Although $T$ is optimized for the ease of decoupling individual beam responses, the decoder remains a blackbox such that the information within $T$ may not be fully exploited due to its indirect use. In this section, we would redesign the decoder by directly using the beam merging matrix $T$ to achieve better architectural interpretability and performance improvement.

Unlike the previous works that split the deployment of CSI encoder and decoder at UEs and gNB, respectively, our gNB knows the exact encoding and decoding processes in our framework. Thus, we can exploit the locally known beam merging matrix $T$ to decode the feedback more efficiently. To this end, we reformulate the problem of DL CSI recovery for $h^{(i)}_{\text{BS,DL}}$, $i = 0, \ldots, N - 1$ by seeking a minimum-norm solution to an under-determined linear system

$$y^{(i)}_{\text{DL}} = Th^{(i)}_{\text{BS,DL}} + n^{(i)}_{\text{DL}}, i = 0, \ldots, N - 1.$$
Interestingly, however, such channels are alternatively characterized by large coherence bandwidth because of the dominance of low-delay paths dominate [33]. This means that for such channels, it is not necessary to have high CSI-RS density in frequency domain.

In this section, a reconfigurable CSI feedback framework is described as a more flexible solution to reduce the number of pilots by selecting frequency reduction (FR) and beam reduction (BR) ratios. Instead of considering feedback of each RB independently, as discussed in Section II, joint UL feedback is performed, taking advantage of the large coherence bandwidth. Spectral coherence is leveraged to further reduce the UL feedback overhead by applying an AE network. The reconfiguration of CSI-RS placement and the design of a learning-based CSI feedback framework, BSdualNet-FR, are also elaborated.

A. Frequency Resource Reconfiguration

In modern wireless protocols, there are designated resource regions for CSI-RS placement [27]. Compatible with existing RS configurations, we can reduce the CSI-RS placement density along the frequency domain by a frequency reduction factor FR by placing pilots only at RB indices \( k = 1, 1 + FR, 1 + 2FR, \ldots, 1 + (K/FR - 1)FR \) as shown in Fig. 7. We can also further reduce the required REs by a beam reduction factor of BR (= round\((N_b/L)\)) by applying beam merging matrix \( T \) designed by using a three-dimensional (3-D) beam merging network with 3-D convolutional kernels as shown in Figs. 8 and 9. Jointly, the total REs for CSI-RS placement can be reduced by a factor of BR · FR. Thus, the total number of pilot REs becomes \( N_bK/(BR · FR) \).  

In previous proposed frameworks, the total REs for CSI-RS placement are reduced by a factor of BR. The total number of pilot REs for \( K \) RBs is \( N_bK/BR \).

![Resource Block (RB)](image)

Fig. 7. Illustration of pilot placement (a) before and (b) after reduction in frequency domain. Note that the color grids represent the designated REs in one of the pilot placement configurations defined in 5G specification [27]. The largest allowable \( L \) is 32 and FR can be 1 or 2 in legitimate pilot placement configurations. For example, \( FR = 2 \) means that one of every 2 consecutive RBs in the assigned bandwidth for CSI-RS is used for pilot placement. In this work, we assume \( FR \) can be any positive integer.

The DL received signal vector \( y_{DL}(i,k) \in \mathbb{C}^{L \times 1} \) at the i-th UE in the k-th RB can be expressed as

\[
y_{DL}(i,k) = S_{DL}(i,k) \cdot T_{H_{BS,DL}}(i,k) + h_{DL},
\]

where the superscript \((i,k)\) denotes the UE and RB indices, respectively. Following Section II, UE-i estimates beam response vectors \( g_{FB}^{(i,k)} \), \( k = 1, 1+FR, \ldots, 1+(K/FR-1)FR \) as a beam response matrix

\[
G_{FB}^{(i)} = \left[ g_{FB}^{(i,1)}, g_{FB}^{(i,1+FR)}, \ldots, g_{FB}^{(i,1+(K/FR-1)FR)} \right] \in \mathbb{C}^{L \times K/FR}
\]

where the estimates \( g_{FB}^{(i,k)} = T_{H_{BS,DL}}^{(i,k)} \in \mathbb{C}^{L} \) are based on pilots reduced by FR.

B. BSdualNet-FR

For further reduction of UL feedback overhead, we compress the beam responses \( G_{FB}^{(i)} \) by implementing a frequency compression module (FCM) similar to an AE. The FCM consists of an encoder at UE and decoder at gNB for CSI compression and recovery, respectively. The encoder consists of four \( 3 \times 3 \) circular convolutional layers with 16, 8, 4 and 2 channels. Subsequently, an FC layer with \( 2LK/(CR · FR) \) elements accounts for dimension reduction by a factor of \( CR_{eff} = BR · FR · CR \) after reshaping. CR_{eff} and CR respectively denote the effective and feedback compression ratios. The FC layer output is sent to a quantization module which uses a trainable soft quantization function as proposed in [10] to generate feedback codewords.

At the gNB, the codewords from different UEs are forwarded into the decoder network of the FMC to recover their
Fig. 9. Network design of BSdualNet-FR.

The BSdualNet-FR is optimized by updating the network parameters $\Theta_{\text{bm}}$, $\Theta_{\text{FMC,enc}}$, $\Theta_{\text{FMC,dec}}$ and $\Theta_{c}$ of the non-linear 3-D beam merging, FMC encoder/decoder, and combining networks $f_{\text{bm}}$, $f_{\text{FMC,enc}}$, $f_{\text{FMC,dec}}$, and $f_{c}$:

$$\arg \min_{\Theta_{\text{bm}}, \Theta_{\text{FMC,enc}}, \Theta_{\text{FMC,dec}}, \Theta_{c}} \left\{ \alpha \cdot \text{loss}_1 + (1 - \alpha) \cdot \text{loss}_2 \right\}$$

where hyperparameter $\alpha$ adjusts the weighting. Note that both $\text{loss}_1$ and $\text{loss}_2$ are differentiable.

It should be noted that the deep learning network used in this framework contains several hyperbolic tangent activation functions and a soft quantization function. These functions could lead to vanishing gradient for parameters in those layers. To address this issue, we propose a two-stage training scheme for optimizing the framework. In the first stage, we train the model by setting $\alpha = 1$ for $N_{\text{first}}$ epochs. During this stage, the combining network is frozen, while we focus on finding the optimal beam merging matrix and encoding/decoding networks. In the second stage, we update $\alpha$ to 0.1 and focus on refining the final estimates with the aid of UL CSI magnitudes. Using the elbow method [33], we found that $N_{\text{first}} = 30$ is typically sufficient to achieve a good tradeoff between performance and training time. This two-stage training scheme helps to ensure that the model is optimized for both accuracy and stability, resulting in a more effective deep learning framework for CSI estimation.

C. Implementation and Discussion

This section summarizes the proposed CSI feedback framework with a step-by-step algorithm to aid readers in better understanding the implementation process. Algorithm 1 summarizes the implementation steps of BSdualNet-FR. Compared to classic AE-based CSI feedback, this model offers several advantages:

- It should be noted that the deep learning network used in this framework contains several hyperbolic tangent activation functions and a soft quantization function. These functions could lead to vanishing gradient for parameters in those layers. To address this issue, we propose a two-stage training scheme for optimizing the framework. In the first stage, we train the model by setting $\alpha = 1$ for $N_{\text{first}}$ epochs. During this stage, the combining network is frozen, while we focus on finding the optimal beam merging matrix and encoding/decoding networks. In the second stage, we update $\alpha$ to 0.1 and focus on refining the final estimates with the aid of UL CSI magnitudes. Using the elbow method [33], we found that $N_{\text{first}} = 30$ is typically sufficient to achieve a good tradeoff between performance and training time. This two-stage training scheme helps to ensure that the model is optimized for both accuracy and stability, resulting in a more effective deep learning framework for CSI estimation.
• UE-friendly encoder: BSdualNet-FR offers a UE-friendly encoder that differs from other typical single-stage encoding models. Our model employs a two-stage encoding approach, beginning with compression of the pilot DL CSI during precoding and pilot downsampling at the gNB, followed by compression using an FMC encoder at UE. This approach offloads the potential computational and storage burden related to the encoder at resource limited UEs.

• Better DL CSI recovery with UL-assisted encoding: FDD reciprocity has been recognized for its potential in aiding DL CSI recovery or decoding, as seen in several recent works. However, our proposed framework is the first to explicitly leverage both FDD reciprocity and delay sparsity to encode DL CSI, to the best of our knowledge. This approach sets our framework apart from previous works and contributes to its improved efficiency and performance.

### Algorithm 1 Implementation of BSdualNet-FR

**Require:** known $BR$, $FR$, $CR$ at gNB and $N$ UEs

perfect UL CSIs in BS domain $\mathcal{H}^{(1)}_{BS,UL}, \mathcal{H}^{(2)}_{BS,UL}, \ldots, \mathcal{H}^{(N)}_{BS,UL}$ at gNB

**Ensure:** $\hat{H}^{(i)}_{BS,DL}$ at gNB

At gNB

- Design a beam merging matrix $T$ (by beam merging network) according to $BR$ and UL CSIs in BS domain of all UEs
  - Equivalently compressed by a factor of $BR$
- Apply $T$ to CSI-RS for pilot reduction
- Pilot downsampling with a factor of $FR$
  - Equivalently compressed by a factor of $FR$
- Pilot transmission

At the $i$-th UE

- Estimate beam response vectors $\mathbf{g}^{(i)}_{BS,DL}, k = 1, 1 + FR, \ldots, 1 + (K/FR - 1)FR$
- Feedback $f^{\text{FCM,en}}(\mathbf{g}^{(i)}_{BS,DL})$ by compressing and quantizing the beam response matrix $\mathbf{g}^{(i)}_{BS,DL}$ via FCM encoder
  - Compressed by a factor of $\frac{CR}{FR}$

At gNB

- Obtain $\hat{H}^{(i)}_{BS,DL,ini}$ by decompressing feedback $f^{\text{FCM,en}}(\mathbf{g}^{(i)}_{BS,DL})$ by FCM decoder
- Obtain final estimate $\hat{H}^{(i)}_{BS,DL}$ by refining $\hat{H}^{(i)}_{BS,DL,ini}$ via combining network according to the corresponding BS domain UL CSI $\mathbf{H}^{(i)}_{BS,UL}$.

It is worth noting that our proposed BSdualNet-FR offers a UE-friendly encoder that differs from other typical single-stage encoding models. Our model employs a two-stage encoding approach, beginning with compression of the pilot DL CSI during precoding and pilot downsampling at the gNB, followed by compression using an FMC encoder at UE. This approach offloads the potential computational and storage burden related to the encoder at resource limited UEs.

Better DL CSI recovery with UL-assisted encoding: FDD reciprocity has been recognized for its potential in aiding DL CSI recovery or decoding, as seen in several recent works. However, our proposed framework is the first to explicitly leverage both FDD reciprocity and delay sparsity to encode DL CSI, to the best of our knowledge. This approach sets our framework apart from previous works and contributes to its improved efficiency and performance.

### VI. Experimental Evaluations

#### A. Experiment Setup

In our numerical test, we consider both indoor and outdoor cases. Using channel model software, we position a gNB of height equal to 20 m at the center of a circular cell with a radius of 30 m for indoor and 200 m for outdoor environment. We equip the gNB with a $8 \times (N_H \times N_H)$ UPA for communication with single antenna UEs. UPA elements have half-wavelength uniform spacing. The number of residual blocks in the combining network is set to $N_B = 5$ throughout.

For our proposed model and other competing models, we set the number of epochs to 300 and 1500, respectively, with batch size of 200. We use a learning rate of 0.001 for the first 100 epochs before switching to $10^{-4}$ for the remaining epochs for our model. To evaluate performance, we generate several indoor and outdoor datasets using channel simulators, each containing $10^5$ random channels. We reserve one seventh of these channels as test data, and split the remaining into 2/3 for training and 1/3 for validation. For both indoor and outdoor scenarios, we use the QuaDRiGa simulator [34] with the scenario features given in 3GPP TR 38.901 Indoor and 3GPP TR 38.901 UMa at 5.1-GHz and 5.3-GHz, and 300 and 330 MHz of UL and DL with LOS paths, respectively. We assume UEs are capable of perfect channel estimation and use an antenna type of omni. We evaluate performance using normalized mean squared error (NMSE) as the performance metric

$$\frac{1}{ND} \sum_{d=1}^{D} \sum_{n=1}^{N} \left\| \mathbf{H}^{(i)}_{BS,DL,d} - \mathbf{H}^{(i)}_{BS,DL,d} \right\|^2_F / \left\| \mathbf{H}^{(i)}_{BS,DL,d} \right\|^2_F, \quad (17)$$

where the number $D$ and subscript $d$ denote the total number and index of channel realizations, respectively.

#### B. Determining Significant Beam Matrix $B_S$ Based on DL and UL CSIs

Fig. 11 illustrates the recovery performance of DL CSI by determining precoding matrix $B_S$ which consists of the $L$ significant beams selected according to CSI magnitudes in UL and DL beam domains, respectively. We call the two approaches as BS-UL and BS-DL, respectively. The modest difference in terms of CSI estimation error demonstrates the high correlation between CSI magnitudes in UL and DL beam domains (partial FDD reciprocity). Specifically, the $L$ dominant beams of UL and DL channels are highly correlated. Good CSI recovery performance requires sufficient number of beams $L$ or REs for CSI-RS.

To evaluate the beam sparsity for different array geometries, Table I demonstrates the average numbers and ratios of significant beams to recover $90\%$ of total CSI energy for UPA.
Fig. 10. Procedures of 3GPP eTypeII and the proposed BSdualNet CSI feedback.

**TABLE I**

| Beam Sparsity Evaluation for Different Array Geometries. Higher Sparsity Means Lower Ratio of Beams Required to Achieve 90% of Total Energy |
|---------------------------------------------------------------|
| Indoor                                                      |
| No. of antennas $(N_H \times N_V)$ | 32 (8 × 4) | 64 (16 × 4) | 128 (16 × 8) | 256 (32 × 8) |
| No. of beams $(> 90\%)$                                      | 11.688     | 19.43       | 31.45        | 55.398       |
| Ratio of beams $(> 90\%)$                                    | 0.36       | 0.30        | 0.25         | 0.22         |
| Outdoor                                                     |
| No. of antennas $(N_H \times N_V)$ | 32 (8 × 4) | 64 (16 × 4) | 128 (16 × 8) | 256 (32 × 8) |
| No. of beams $(> 90\%)$                                      | 8.23       | 13.3        | 18.71        | 32.9         |
| Ratio of beams $(> 90\%)$                                    | 0.26       | 0.21        | 0.15         | 0.13         |

with different antenna numbers. We see that larger array and lower angular spread (outdoor channels) lead to a higher beam sparsity. The proposed framework exploits beam sparsity to allocate a small number of required REs while maintaining recovery performance. Namely, more REs are saved for DL CSI training for large-scale arrays and channels with low angular spread. This shows the practical potential of such feedback framework in communications systems with large-scale arrays.

**C. Testing Different Numbers of Available REs**

We evaluate the performance of CSI recovery by adopting the proposed encoder-free CSI feedback frameworks, BSdualNet_0, BSdualNet and BSdualNet-MN. To test the efficacy without considering quantization, we first compare BSdualNet_0 with two heuristic approaches (denoted as BS-UL and BS-DL) that recover DL CSIs according to $L$ beam responses where the beams are selected according to the UL and DL CSI magnitudes, respectively. Note that BS-UL should serve as the lower bound of BSdualNet_0 since BSdualNet_0 is equivalent to refine the result of BS-UL with an additional combining network.

Figs. 12 (left) and (right) provide the NMSE performance for different number of available REs $L$ in an RB for BSdualNet_0, BS-UL and BS-DL in both indoor and outdoor scenarios, respectively. The results show that BSdualNet_0 delivers better performance than BS-UL and also BS-DL in outdoor scenario owing to the high spatial correlation in beam domain. Because of the high angle spread induced by the more complex multi-path environment in indoor scenarios, the combining network in BSdualNet_0 only marginally improve the recovery performance.

Figs. 13 (left) and (right) illustrate the NMSE performance for different number $L$ of REs within a RB for BSdualNet_0, BS-UL and BS-DL in both indoor and outdoor channels, respectively. We can observe the benefits of the beam merging matrix $T$ especially in outdoor cases. Furthermore, instead of using a convolution-layer based combining network, changing the combining function as a minimum-norm solution yields a significant performance improvement in both indoor and outdoor scenarios. Since minimum-norm solution directly
uses the beam merging matrix $\mathbf{T}$, it becomes more efficient to decouple the superposition of weighted beam responses by minimizing the MSE of DL CSIs.

**D. Performance for Different Numbers of UEs**

Similar to our beam merging matrix $\mathbf{T}$, measurement matrix in compressive sensing based frameworks [35], [36] also functions to shrink the dimension of original data and derive a better representation for their sparsity that can be easier to recover. To demonstrate the relative performance of the proposed frameworks, we also compare with two successful compressive approaches ISTA [35] and ISTA-Net [36]:

- **Iterative Shrinkage-Thresholding Algorithm (ISTA):**
  Its regularization parameter and maximum iteration number are set to 0.5 and 3000, respectively.

- **ISTA-Net:** The phase and epoch numbers are set to 5 and 1000, respectively. Please refer to [37] for the implementation of ISTA-Net in CSI feedback task.

Figs. 14 (left) and (right) provide the NMSE performance comparison for different numbers of UEs $N$ for $L = 8$ REs in a RB for BSdualNet, BSdualNet-MN, ISTA and ISTA-Net and under indoor and outdoor scenarios, respectively. From the results, we observe the clear performance degradation for BSdualNet and BSdualNet-MN as UE number grows. This is intuitive since it is difficult to find an optimum beam merging matrix for all active UEs. Fortunately, for most cases, the performance degradation tends to saturate after the UE number exceeds a certain number typically less than 10 for BSdualNet-MN.

Our tests show that both BSdualNet and BSdualNet-MN deliver better performance over ISTA and ISTA-Net under different UE numbers. Our heuristic insight is that measurement matrix in ISTA and ISTA-Net is unknown at recovery whereas the beam merging matrix is designed by the gNB and can be explicitly utilized by the recovery decoders of BSdualNet and BSdualNet-MN.

**E. Different CSI-RS Configurations and Compression Ratios**

We consider a 5.76 MHz subband (i.e., 32 RBs each of bandwidth 180K-Hz). Each codeword element uses 8 quantization bits. To comprehensively evaluate BSdualNet-FR, The two tables in Fig. 15 (a) and (b) provide the NMSE performance of BSdualNet-FR against different CSI-RS configurations and compression ratios in outdoor and indoor scenarios, respectively. We apply the same background color on results with the same pilot and feedback overhead reduction ratios.

Since outdoor channels generally exhibit stronger sparsity and larger delay spread respectively in beam and delay domains, we observe a slight performance degradation with BR increase as opposed to FR increase. Importantly, for $BR = 4$, there is a clear performance loss even when using the same pilot and feedback overhead reduction ratio. Despite the channel sparsity, with the use of half-wavelength antenna spacing (i.e., Nyquist sampling in spatial domain), the overly aggressive compression in beam domain cause too much information loss to recovery at the gNB. For indoor channels, we observe a slight performance degradation when increasing FR instead of BR because of larger angular and shorter delay spread of indoor CSI.

**F. Different Effective Compression Ratio $CR_{eff}$**

As benchmarks, we also compare BSdualNet-FR with CsiNet [3], CRNet [6], CsiNet-Pro [8] and another successful method DualNet-MP [12]. The newly proposed DualNet-MP also exploits FDD reciprocity by incorporating UL CSI...
Fig. 15. NMSE performance of BSdualNet-FR for different CSI-RS placement configurations in (a) indoor and (b) outdoor scenarios. (The results with the same effective compression ratio are denoted as the same color. The best performance at the same effective compression ratio is denoted by bold fonts with underline.)

Table II presents the comparison of NMSE for CsiNet, CRNet, CsiNet-Pro, DualNet-MP, and BSdualNet-FR under different values of effective compression ratio $C_{eff}$ in indoor and outdoor cases. Benefiting from the UL CSI magnitudes, both BSdualNet-FR and DualNet-MP can outperform CsiNet, CRNet and CsiNet-Pro in most cases. Interestingly, better utilization of UL CSI by BSdualNet-FR provides better performance than DualNet-MP. Although the performance gain becomes less impressive for higher $C_{eff}$, it is practically important to note the additional benefit of the BSdualNet-FR framework in reducing REs for DL CSI-RS by a factor of $BR \cdot FR$, which enables gNB to reconfigure the CSI-RS placement to enhance DL spectrum efficiency.

To demonstrate the benefits of DL spectrum efficiency, we use achievable rate as another performance metric. We allow gNB to choose its precoder $f = \hat{h}_{DL}$ for maximizing DL transmission gain. According to 5G NR specification, we assume 32 REs (for 32 antenna ports) among all 168 REs in each RB for CSI-RS transmission and we adopt a frequency reduction rate $FR$ to lower CSI-RS placement density. We can define the achievable rate in each RB as follows:

$$R = \gamma \cdot E[\log_2(1 + \frac{\|\hat{h}_{DL}\|^2}{\|\hat{h}_{DL}\|^2 + \frac{\|\hat{h}_{UL}\|^2}{N_0}})](\text{bit/s/Hz}),$$  \hspace{1cm} (18)

where $\hat{h}_{DL}$ and $\hat{h}_{UL}$ respectively denote the estimated and original DL CSIs. $N_0$ denotes the ambient noise level. The quantity

$$\gamma = \frac{K \cdot 168 \cdot P - K \cdot 32/(FR \cdot BR)}{168P - 32/(FR \cdot BR)}$$

$$= \frac{K \cdot 168 \cdot P}{168P}$$

denotes the effective ratio of REs being used for data transmission, where $P$ is the sparsity of CSI-RS placement in terms of slots.\(^3\) Fig. 16 shows the achievable rate of all alternatives under different signal-to-noise ratios (SNRs) for both indoor and outdoor scenarios. We observe that BSdualNet outperforms other compared approaches in terms of DL achievable rate although its NMSE performance may not always prevail.

\(^3\)One out of every $P$ slots is assigned for CSI-RS placement.
Fig. 16. DL achievable rate under different SNRs for (a) indoor and (b) outdoor scenarios. Note that we consider a CSI-RS sparsity $P = 4$ and $CR_{eff} = 32$ ($(FR = 8, BR = 1, CR = 4)$ is for indoor channels whereas $FR = 2, BR = 2, CR = 8$ is for outdoor channels) in the test results.

This is due to the effect of saving REs for DL signaling to avoid bandwidth waste for data transmission.

**G. Complexity: FLOPs and Parameters**

Most UEs have stronger memory, computation, and power constraints. The system design favors light-weight and simpler encoders for deployment at UEs. In comparison with the baseline CsiNet Pro and DualNet, Table III shows dimension reduction in frequency and beam domains and smaller input size of our encoder/decoder architecture. BSdualNet-FR provides significant reduction in terms of FLOPs and the number of model parameters. Similarly, if the total reduction factor $FR \cdot BR \geq 4$, BSdualNet-FR shows lower storage requirement than those light-weight models CsiNet and CRNet.

DL-based CSI feedback can enhance feedback accuracy and reduce overheads but may require more computational resources and a larger model, which can impose some burden on UEs with limited resources. To examine tradeoffs, [38] provides a figure that clearly shows the computational and storage costs and NMSE performance of DL-based SOTAs. Following this illustration, we present Fig. 17 and 18 to also demonstrate the NMSE performance versus computation (FLOPs) of the NNs deployed at UEs for indoor and outdoor scenarios, respectively, setting CR to 16. With a two-stage compression, most of the computational tasks are performed at gNB in BSdualNet CSI feedback, reducing the burden on UEs. Although DL-based CSI feedback approaches generally perform better at the cost of more computations, BSdualNet outperforms them and requires much less computational resources at UEs with its efficient two-stage encoding.

**TABLE III**

**COMPARISON OF PARAMETERS (PARAS) AND FLOPS AT ENCODER**

| $CR_{eff}$ | CsiNet | CRNet | CsiNet-Pro | DualNet-MP | BSdualNet-FR |
|------------|--------|-------|------------|------------|--------------|
|            | PARAs  | FLOPs | PARAs      | FLOPs      | PARAs        |
|            |        |       |            |            |              |
| 4          | 2.8M   | 1.1M  | 2.8M       | 1.2M       | 4.3M         |
|            |        |       |            |            | 11.1M        |
| 8          | 1.4M   | 0.56M | 1.4M       | 0.68M      | 3.8M         |
|            |        |       |            |            | 10.56M       |
| 16         | 0.7M   | 300K  | 0.7M       | 420K       | 1.9M         |
|            |        |       |            |            | 10.3M        |
| 32         | 350K   | 170K  | 350K       | 290K       | 950K         |
|            |        |       |            |            | 10.2M        |
|            |        |       |            |            | 490 K        |

| $CR_{eff}$ | FLOPs | PARAs |
|------------|-------|-------|
| 32         | 950K  | 10.2M |

**H. Imperfect CSI Estimation**

The errors in acquiring DL CSI in FDD systems can arise from both DL CSI estimation at UE and feedback/recovery.
In previous experiments, we assumed perfect CSI estimation at UE. In this subsection, we consider the effect of imperfect model-based CSI estimation [39] on the proposed BSdualNet CSI feedback framework. Assuming the communication link is subject to block-fading channels, the received signal at the i-th fading block, as described in [39], can be expressed as:

$$y_i = X_i^H h_i + n_i,$$  \hfill (19)

where $X_i \in \mathbb{C}^{N_t \times L}$ is the training pilot matrix. The signal model becomes identical to our signal model (5) by letting $X_i^H = S_{DL,L} T_i$ and $h_i = B^T h_{DL}$. Note that the $T_i$ is a function of UL CSI at the i-th fading block. The DL CSI follows a Gauss-Markov distribution according to

$$h_0 = R^{\frac{1}{2}} g_0,$$  \hfill (20)

$$h_i = \eta h_{i-1} + \sqrt{1-\eta^2} R^{\frac{1}{2}} g_i, \ i \geq 1,$$  \hfill (21)

where $R$ is a spatial correlation matrix, $g_i$ is an i.i.d. random vector according to $CN(0, I)$ for all $i$, and $\eta$ is a temporal correlation coefficient. In the following experiments, we consider three cases of DL CSI estimation:

- **MMSE CSI estimation**: $\hat{h}_i = \rho X_i (I + X_i^H R X_i)^{-1} y_i$ with assumption that spatial correlation matrix $R$ and temporal correlation coefficient $\eta$ are known at UEs.
- **Model-based CSI estimation** [39]: it applies MMSE estimator and Kalman filter to refine DL CSI estimation over time. Assume spatial correlation matrix $R$ and temporal correlation coefficient $\eta$ are known at UEs.
- **Perfect CSI estimation**

Fig. 19 shows the NMSE performance versus fading block index for the indoor and outdoor scenarios when $CR_{eff} = 16$, incorporating imperfect model-based CSI estimation [39] with the proposed BSdualNet CSI feedback. The results indicate that the model-based CSI estimation provides better NMSE performance over time by leveraging the CSI temporal correlation, but it never achieves the performance of perfect CSI estimation. This suggests that the errors in CSI estimation dominate the errors from CSI feedback, and better CSI estimation should be a primary focus for further performance improvement.

On the other hand, if perfect CSI estimation can be achieved, we should focus on developing more efficient CSI feedback approaches. Therefore, it is crucial to have both CSI estimation and feedback approaches to achieve accurate DL CSI acquisition at the gNB.

### VII. Conclusion

This work presents a new deep learning framework for CSI estimation in massive MIMO downlink. Leveraging UL CSI estimate to reduce its CSI-RS resources, the gNB designs a beam merging matrix based on UL channel magnitude information to transform DL CSI observation at UEs into a lower dimensional representation that is easier for feedback and recovery. We further develop an efficient minimum-norm CSI recovery network to improve recovery accuracy. Our new framework does not deploy training deep learning models at UEs, thereby lowering UE complexity and power consumption. We achieve further reduction of DL CSI training and feedback overhead, by introducing a reconfigurable CSI-RS placement. Test results demonstrate significant improvement of CSI recovery accuracy and reduction of both DL CSI training and UL feedback overheads.

### References

[1] C. Lin, S. Lin, and E. Blasch, “TULVCAN: Terahertz ultra-broadband learning vehicular channel-aware networking,” in Proc. IEEE INFOCOM Workshop, May 2021, pp. 1–6.

[2] X. Lin, “An overview of 5G advanced evolution in 3GPP release 18,” 2022, arXiv:2201.01358.

[3] 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.

[4] Y. Sun, W. Xu, L. Liang, N. Wang, G. Y. Li, and X. You, “A lightweight deep network for efficient CSI feedback in massive MIMO systems,” IEEE Wireless Commun. Lett., vol. 10, no. 8, pp. 1840–1844, Aug. 2021.

[5] S. Ji and M. Li, “CLNet: Complex input lightweight neural network designed for massive MIMO CSI feedback,” IEEE Wireless Commun. Lett., vol. 10, no. 10, pp. 2318–2322, Oct. 2021.

[6] Z. Lu, J. Wang, and J. Song, “Multi-resolution CSI feedback with deep learning in massive MIMO system,” in Proc. IEEE Int. Conf. Commun. (ICC), Jun. 2020, pp. 1–6.

[7] 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, Apr. 2020.
[8] Z. Liu, M. D. Rosario, and Z. Ding, “A Markovian model-driven deep learning framework for massive MIMO CSI feedback,” IEEE Trans. Wireless Commun., vol. 21, no. 2, pp. 1214–1228, Feb. 2022.

[9] J. Guo, X. Yang, C.-K. Wen, S. Jin, and G. Ye Li, “Deep learning-based CSI feedback and cooperative recovery in massive MIMO,” arXiv:2003.03303.

[10] 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, Aug. 2020.

[11] 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, Jun. 2019.

[12] Y. Lin, Z. Liu, T. Lee, and Z. Ding, “Deep learning phase compression for MIMO CSI feedback by exploiting FDD channel reciprocity,” IEEE Wireless Commun. Lett., vol. 10, no. 10, pp. 2200–2204, Oct. 2021.

[13] Y. Liu and O. Simeone, “Learning how to transfer from uplink to downlink via hyper-recurrent neural network for FDD massive MIMO,” IEEE Trans. Wireless Commun., vol. 21, no. 10, pp. 7975–7989, Oct. 2022.

[14] Z. Zhong, L. Fan, and S. Ge, “FDD massive MIMO uplink and downlink channel reciprocity properties: Full or partial reciprocity?” in Proc. IEEE GLOBECOM, Dec. 2020, pp. 1–5.

[15] Y. Ding and B. D. Rao, “Dictionary learning-based sparse channel representation and estimation for FDD massive MIMO systems,” IEEE Trans. Wireless Commun., vol. 17, no. 3, pp. 5437–5451, Aug. 2018.

[16] X. Zhang, L. Zhong, and A. Sabharwal, “Directional training for FDD massive MIMO,” IEEE Trans. Wireless Commun., vol. 17, no. 8, pp. 5183–5197, Aug. 2018.

[17] W. Shen, L. Dai, B. Shim, Z. Wang, and R. W. Heath Jr., “Channel feedback based on AoD-adaptive subspace codebook in FDD massive MIMO,” IEEE Trans. Commun., vol. 66, no. 11, pp. 5235–5248, Nov. 2018.

[18] On NR Type I Codebook, document 3GPP TSG RAN WG1 No. 88, R1-1702205, Intel, 2021.

[19] Type II CSI Reporting, document 3GPP TSG RAN WG1 No. 89, R1-1707962, Samsung, 2021.

[20] M. Chen, J. Guo, C.-K. Wen, S. Jin, G. Ye Li, and A. Yang, “Deep learning-based implicit CSI feedback in massive MIMO,” arXiv:2105.10100.

[21] Discussion on Field Measurements and Evaluation Assumptions for FDD CSI, document 3GPP TSG RAN WG1 No. 102-e, R1-2006414, Huawei, HiSilicon, 2020.

[22] Summary of CSI Enhancements for MTRP and FDD, document 3GPP TSG RAN WG1 No. 103-e, 2020.

[23] Y. Lin, T. Lee, and Z. Ding, “Deep learning for partial MIMO CSI feedback by exploiting channel temporal correlation,” in Proc. 55th Asilomar Conf. Signals, Syst., Comput., Oct. 2021, pp. 345–350.

[24] J. Guo, C. Wen, and S. Jin, “CAnet: Uplink-aided downlink channel acquisition in FDD massive MIMO using deep learning,” IEEE Trans. Commun., vol. 70, no. 1, pp. 199–214, Jan. 2022.

[25] M. B. Mashhadi and D. Gündüz, “Pruning the pilots: Deep learning-based pilot design and channel estimation for MIMO-OFDM systems,” IEEE Trans. Wireless Commun., vol. 20, no. 10, pp. 6315–6328, Oct. 2021.

[26] Y. Yamada and T. Ohtsuki, “Autoencoder-based pilot pattern design for CDL channels,” in Proc. 27th Asia Pacif. Conf. Commun. (APCC), Oct. 2022, pp. 76–80.

[27] NR: Physical Channels and Modulation, document TS 38.211, Version 16.6.0, 3rd Generation Partnership Project (3GPP), Jun. 2020.

[28] G. Morozov, A. Davydov, and V. Sergeev, “Enhanced CSI feedback for FD-MIMO with beamformed CSI-RS in LTE-A pro systems,” in Proc. VTC-Fall, Sep. 2016, pp. 1–5.

[29] R. L. Haupt, “Array beamforming,” in Timed Arrays: Wideband and Time Varying Antenna Arrays. Hoboken, NJ, USA: Wiley, 2015, pp. 78–94.

[30] C. Lin, W. Kao, S. Zhan, and T. Lee, “BxNet: A deep learning-based beam selection method for mmWave communications,” in Proc. VTC-Fall, Sep. 2019, pp. 1–6.

[31] Beam Management, document 3GPP TR 38.802, Version 16.6.0, 3GPP, Sep. 2017.

[32] Y.-C. Lin, T.-S. Lee, and Z. Ding, “Exploiting partial FDD reciprocity for beam based pilot precoding and CSI feedback in deep learning,” 2022, arXiv:2205.14558.

[33] W. Debaenst, A. Feys, I. Cuiñas, M. G. Sánchez, and J. Verhaevert, “RMS delay spread vs. coherence bandwidth from 5G indoor radio channel measurements at 3.5 GHz band,” Sensors, vol. 20, no. 3, p. 750, Jan. 2020.

[34] S. Jaekel, 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 Propag., vol. 62, no. 6, pp. 3242–3256, Jun. 2014.

[35] A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” SIAM J. Imag. Sci., vol. 2, no. 1, pp. 183–202, Jan. 2009.

[36] J. Zhang and B. Ghanem, “ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing,” in Proc. IEEE CVPR, Jul. 2019, pp. 1832–1840.

[37] J. Guo, L. Wang, F. Li, and J. Xue, “CSI feedback with model-driven acquisition in FDD massive MIMO using deep learning,” IEEE Commun. Lett., vol. 26, no. 3, pp. 547–551, Mar. 2022.

[38] J. Guo, C. Wen, and S. Jin, “CAnet: Uplink-aided downlink channel acquisition in FDD massive MIMO using deep learning,” IEEE Trans. Commun., vol. 68, no. 10, pp. 6315–6328, Oct. 2021.

[39] Y. Yamada and T. Ohtsuki, “Autoencoder-based pilot pattern design for CDL channels,” in Proc. 27th Asia Pacif. Conf. Commun. (APCC), Oct. 2022, pp. 76–80.