

Training-free Cost-efficient Compression for Massive MIMO Channel State Feedback

Yu-Chien Lin, Ta-Sung Lee, and Zhi Ding

Abstract—Acquiring downlink channel state information (CSI) at basestation (gNB) is crucial for optimizing performance in massive MIMO FDD systems. Deep learning (DL) architectures have shown successes in enabling UE-side CSI feedback and gNB-side recovery, but often lack flexibility and/or require volumes of customized training data for specific RF channel environments and compression ratios. This work proposes a new CSI feedback architecture called zero-replacement (ZR). ZR is free from customized training and can be directly applied to new and unseen channel scenarios without pre-training and/or customization. It is also scalable and simple to implement, making it suitable for practical massive MIMO wireless deployment. We further generalize a Select-ZR algorithm, which switches between different sparse transformation techniques to enhance recovery performance. Our numerical results demonstrate that both proposed ZR and Select-ZR algorithms achieve competitive CSI recovery accuracy and feedback efficiency across various channels against highly complex data-driven DL models.

Index Terms—Compressive feedback, model-free, massive MIMO, CSI recovery.

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) improves spectrum and energy efficiency in wireless systems, but requires accurate downlink channel state information (CSI) acquisition at the base station (BS) or gNodeB (gNB). In frequency-division duplexing (FDD) systems, downlink CSI acquisition depends on UE feedback, which can be costly due to the large number of parameters. Efficient compressive CSI feedback is crucial to conserve uplink bandwidth and UE power for practical deployment of massive MIMO in FDD wireless networks.

Cellular channel state information (CSI) has a limited delay spread (DS), which is a characteristic of radio physics. Efficient user equipment (UE) feedback can take advantage of this DS sparsity to compress CSI. One approach to efficient CSI compression and recovery is the use of a deep autoencoder framework, as demonstrated in [1]. This framework includes an encoder at the UE and a decoder at the serving gNB. Other related works have also demonstrated superior CSI

recovery or lightweight design using various autoencoder models, such as [2]–[5]. In addition to autoencoders, more recent works have utilized the underlying channel correlation to aid and improve the recovery of downlink CSI at BSs. These approaches include using previous CSI [2], [6], CSI of nearby UEs [7], and uplink CSI [8]–[10]. Further advances have focused on reducing model complexity and storage size to facilitate practical and low-cost deployment of DL-based CSI compressive feedback architecture in wireless networks, as demonstrated in [3], [5].

Deploying DL-based frameworks in wireless systems presents challenges for operators, such as the high cost of collecting training data required for DL optimization. Massive MIMO CSI data acquisition necessitates extensive field measurements, creating a practical challenge as stated in [11]. Wireless networks are typically deployed in a range of RF environments, necessitating the configuration of numerous DL models for CSI compression and recovery at various channel scenarios and compression ratios. These requirements impose a large memory burden on UEs, necessitate large training datasets, and increase costs, resulting in inflexibility and difficulties in practical deployment. Although transfer learning and online learning concepts [12], [13] have achieved modest success in reducing training costs, the implementation of multiple DL models still results in high hardware and power costs, particularly at the UE side, as channel bandwidth and antenna numbers increase in future wireless network generations. As a result, deploying DL models for compressive CSI feedback faces significant challenges.

In this paper, we propose a novel plug-and-play algorithm called zero-replacement (ZR) for efficient channel state information (CSI) feedback. Our approach is simpler, more scalable, and more flexible than previous methods, and does not require prior training. Our contributions are as follows:

- We develop a low-complexity, lightweight, scalable and unified analytical CSI feedback framework that can accommodate different propagation channel types and compression ratios without the need for prior training or fine-tuning.
- The ZR framework achieves competitive recovery performance compared to state-of-the-art DL-based methods.
- We design a modified run-length and Huffman encoding (mRLE) scheme for efficient compression of the unique 2D patterns of downlink CSI.
- Unlike DL-based models, our algorithm does not require the collection of training datasets, which can cause significant transmission overhead in practical systems

Y.-C. Lin and Z. Ding are with the Department of Electrical and Computer Engineering, University of California, Davis, CA, USA (e-mail: yclin, zding@ucdavis.edu).

T.-S. Lee is with the Institute of Communications Engineering, National Yang Ming Chiao Tung University, Taiwan (e-mail: tslee@nycu.edu.tw).

This work is based on materials supported by the National Science Foundation under Grants 2029027 and 2002937 (Lin, Ding) and by the National Science and Technology Council of Taiwan under grant NSTC 112-2218-E-A49-020, NSTC 112-2221-E-A49-079; in part by Qualcomm through a Taiwan University Research Collaboration Project. (Lee and Lin).

- We introduce Select ZR, a dynamic-transform CSI feedback framework, which achieves better recovery performance by selecting the best sparse transformation for each specific scenario.

II. SYSTEM MODEL

A. Downlink CSI Preprocessing

We consider a single-cell MIMO FDD link in which a gNB with N_a antennas serves a plurality of single-antenna UEs. Following 3GPP technical specifications, sparse pilot symbols (CSI-RS) are distributed in frequency domain for downlink transmission. Assuming each subband contains N_f subcarriers with spacing of Δf and a pilot spacing of DR_f subcarriers, adjacent pilots are separated by $\text{DR}_f \cdot \Delta f$ Hz. We denote $\mathbf{h}_i \in \mathbb{C}^{M_f \times 1}$ as downlink CSI of the i -th gNB at M_f pilot positions. Let superscript $(\cdot)^H$ denote conjugate transpose. By collecting CSI of each gNB, a pilot sampled downlink CSI matrix $\tilde{\mathbf{H}}$ relates to the full downlink CSI matrix $\hat{\mathbf{H}}$ via

$$\tilde{\mathbf{H}} = \tilde{\mathbf{H}} \mathbf{Q}_{\text{DR}_f} = [\mathbf{h}_1 \ \mathbf{h}_2 \ \cdots \ \mathbf{h}_{N_a}]^H \in \mathbb{C}^{N_a \times M_f}, \quad (1)$$

where $\mathbf{Q}_{\text{DR}_f}$ is a downsampling matrix with pilot rate DR_f .

To reduce feedback overhead, we exploit physical angular and multipath delay sparsity of CSI by transforming full downlink CSI into angle-delay (AD) domain through discrete Fourier transform (DFT) or discrete cosine transform (DCT). We then truncate the insignificant near-zero elements in trailing delay indices as follows:

$$\mathbf{M} = \tilde{\mathbf{H}} \cdot \mathbf{F}_D \cdot \underbrace{\begin{bmatrix} \mathbf{I}_{N_t \times N_t} \\ \mathbf{0} \end{bmatrix}}_{\mathbf{T}} \in \mathbb{C}^{N_a \times N_t}, \quad (2)$$

where $\mathbf{F}_D \in \mathbb{C}^{M_f \times M_f}$ denotes a sparse transformation matrix such as IDFT or DCT matrix whereas $\mathbf{T}$ denotes delay domain truncation. Note that the matrix $\mathbf{T}$ may be controlled according to transformation and CSI properties. Matrix $\mathbf{T}$ in Eq. (2) is an example for DCT transformation that drops the last $M_f - N_t$ columns of $\tilde{\mathbf{H}} \cdot \mathbf{F}_D$ corresponding to long (but negligible) multipath delays.

B. DL Compression

Autoencoder has shown successes for CSI compression. An encoder at UE compresses its estimated downlink CSI for uplink feedback and a decoder at gNB recovers the estimated CSI according to the feedback from UE. Assuming negligible CSI elements at large delays, many have exploited convolutional and fully connected layers to compress and recover the truncated downlink pilot CSI via

$$\text{Encoder: } \mathbf{q} = f_{\text{en}}(\mathbf{M}), \quad (3)$$

$$\text{Decoder: } \hat{\mathbf{M}} = f_{\text{de}}(\mathbf{q}). \quad (4)$$

We note that the size of the codeword $\mathbf{q}$ for uplink feedback is determined by a specific compression ratio. However, in practical systems, multiple DL models are necessary to meet diverse uplink feedback requirements, resulting in computational and storage burdens, particularly for low-cost UEs. The decoder

then replaces the truncated downlink CSI $\hat{\mathbf{M}}$ via zero-padding to transform CSI back from BD domain to estimate downlink CSI matrix $\tilde{\mathbf{H}}$ in the subcarrier domain as follows:

$$\hat{\mathbf{H}} = \begin{bmatrix} \hat{\mathbf{M}} & \mathbf{0}_{N_a \times M_f - N_t} \end{bmatrix} \mathbf{F}_D^H \in \mathbb{C}^{N_a \times M_f}. \quad (5)$$

The CSI recovery accuracy can be measured by the normalized mean square error (NMSE) of the full downlink CSI:

$$\text{NMSE}(\hat{\mathbf{H}}, \tilde{\mathbf{H}}) = \sum_{d=1}^D \left\| \hat{\mathbf{H}}_d - \tilde{\mathbf{H}}_d \right\|_F^2 / \left\| \tilde{\mathbf{H}}_d \right\|_F^2, \quad (6)$$

where subscript d denotes the d -th random test.

III. ZR CSI FEEDBACK FRAMEWORK

DL-based CSI feedback frameworks have shown promising results in compressing CSI at the UE and recovering it at the BS. However, customizing multiple DL models for different channel scenarios and compression levels makes them inflexible and difficult to apply in practical gNBs with different array sizes. This also poses a challenge for UEs to store multiple pre-trained DL models, hindering widespread deployment, especially for low-cost UEs. Training different DL configurations for various scenarios would require suitable channel models and large amounts of training data.

Our aim is to create a compressive CSI feedback algorithm that is simple, flexible, and widely applicable. It is worth noting that 2D CSI matrices bear resemblance to image data. CSI demonstrates distinct delay and angular sparsity properties that can be revealed by sinusoidal sparse transformations. To enable efficient CSI feedback, we capitalize on this similarity and adapt the entropy encoding technique used in JPEG compression. JPEG compression is renowned for its efficiency, universality, and cost-effectiveness. We customize this technique based on the sparsity distribution of CSI. The step-by-step procedures are summarized below:

A. Ordering Real/Imaginary CSI

Since the real and imaginary parts of the preprocessed DL CSI $\mathbf{M}$ exhibit similar sparsity distributions in the delay domain, this algorithm compresses their sparsity distributions together and feeds back the values separately. We first separate the processed DL CSI $\mathbf{M} \in \mathbb{C}^{N_a \times N_t}$ into real and imaginary parts, $\mathbf{M}_{\text{real}} = \text{Real}(\mathbf{M}) \in \mathbb{R}^{N_a \times N_t}$ and $\mathbf{M}_{\text{imag}} = \text{Imag}(\mathbf{M}) \in \mathbb{R}^{N_a \times N_t}$.

B. ZR Encoding

The goal of ZR encoding is to efficiently encode a large sparse matrix by extracting only the significant elements and representing their corresponding positions in a compact format. Figure 1 shows a toy example to implement the operation of the proposed ZR encoder with $CR = 5$ and the input matrices $\mathbf{M}_{\text{real}}$ and $\mathbf{M}_{\text{imag}}$ are of size 5×5 . The output bit stream is $[\Omega_B \ \Omega_m]$. We introduce ZR encoding in a step-by-step procedure as follows:

- *Significant Element Extraction:* $\mathbf{M}_{\text{real}}$ and $\mathbf{M}_{\text{imag}}$ are fed into *significant element extraction* block for sampling the top- $\lfloor N_a N_t / CR \rfloor$ significant elements $\mathbf{m}_{\text{real}} \in$

TABLE I
MODIFIED HUFFMAN CODING TABLE.

Bits (K)	Number of zeros (N_0)					Prefix (p)
0	0					0
1	1 (0)	EOS (1)				10
1	2 (0)	3 (1)				110
2	4(00)	5(01)	6(10)	7(11)		1110
3	8	9	...	14	15	11110
4	16	17	...	30	31	111110
5	32	33	...	62	63	1111110
6	64	65	...	126	127	11111110
7	128	129	...	254	255	111111110
8	256	257	...	510	511	1111111110
9	512	513	...	1022	1023	11111111110

$\mathbb{R}^{\lfloor N_a N_t / CR \rfloor}$ and $\mathbf{m}_{\text{imag}} \in \mathbb{R}^{\lfloor N_a N_t / CR \rfloor}$ for real and imaginary parts, respectively. Additionally, it also outputs a binary position map $\mathbf{B} \in \{0, 1\}^{N_a \times N_t}$, which contains only 1's and 0's corresponding to sampled and discarded elements, respectively.

- *Significant Element Quantization*: For effective binary representations, we express the downsampled sequences $\mathbf{m}_{\text{real}}$ and $\mathbf{m}_{\text{imag}}$ as bit stream Ω_m by μ -law companding entropy encoding with B bits.
- *mRLE of Binary Position Map*: To compactly represent the position map $\mathbf{B}$, we propose a modified RLE (mRLE) which is shown by the pseudo code (Alg. 1). The goal of mRLE is to find a symbol list Ω_S marking the numbers of consecutive zeros between ones in a back-and-forth scanning pattern shown in Figure 1. The final entry EOS in the symbol list Ω_S denotes "end of symbols" and also means that no more 1's in the remaining sequence.
- *Huffman Encoding of Binary Position Map*: In order to efficiently convert the symbol list Ω_S into a bit stream Ω_B , we have introduced a modified Huffman coding. As shown in the histogram of the symbol list Ω_S (Figure 3), the majority of cases involve only a few consecutive 0's between 1's. To reduce transmission costs, we have designed a modified Huffman coding presented in Table I. The frequently used symbols are presented in a compact bit stream, and prefix bit streams are designed for unambiguous decoding. Additionally, it should be noted that we represent EOS with a short bit stream (101) since it almost always appear in Ω_S for each DL CSI.

C. ZR Decoding

In ZR decoding, the first step is to separate Ω_B and Ω_m using a Huffman coding table. Next, we obtain $\mathbf{B}$ from Ω_B by reversing mRLE, and estimate $\hat{\mathbf{m}}_{\text{real}}$ and $\hat{\mathbf{m}}_{\text{imag}}$ from Ω_m using dequantization and μ -law decompression. Finally, we merge the binary position map $\mathbf{B}$ and the estimated real and imaginary significant elements $\hat{\mathbf{m}}_{\text{real}}$ and $\hat{\mathbf{m}}_{\text{imag}}$ to obtain $\hat{\mathbf{M}}_{\text{real}}$ and $\hat{\mathbf{M}}_{\text{imag}}$. We provide a detailed description of the above procedures in Algorithm 2.

D. Select ZR

We find that the pre-processing of DL CSI (i.e., sparse transformation) affects the final recovery performance for the ZR

Algorithm 1 Modified RLE

Require: $\mathbf{B}$, $\Omega_S = [\cdot]$

Ensure: Ω_S

```

i  $\leftarrow$  B-F-vec(B)  $\triangleright$  Vectorize in a back-forth way (Figure 1)
N0  $\leftarrow$  0  $\triangleright$  Initialize zero counter
for  $k = 1 : 1 : \text{length}(\mathbf{i})$  do
  if  $\mathbf{i}(k) = 1$  then
     $\Omega_S \leftarrow \{\Omega_S, N_0\}$   $\triangleright$  Record consecutive zeros
    N0  $\leftarrow$  0  $\triangleright$  Clear zero counter
  else
    N0  $\leftarrow$  N0 + 1  $\triangleright$  Accumulate zero counter
  end if
end for
if N0  $\geq 1$  then  $\Omega_S \leftarrow \{\Omega_S, \text{EOS}\}$   $\triangleright$  If no more 1's, marked with EOS
end if

```

Algorithm 2 ZR Decoding

Require: $\mathbf{b} = [\Omega_B \ \Omega_m]$

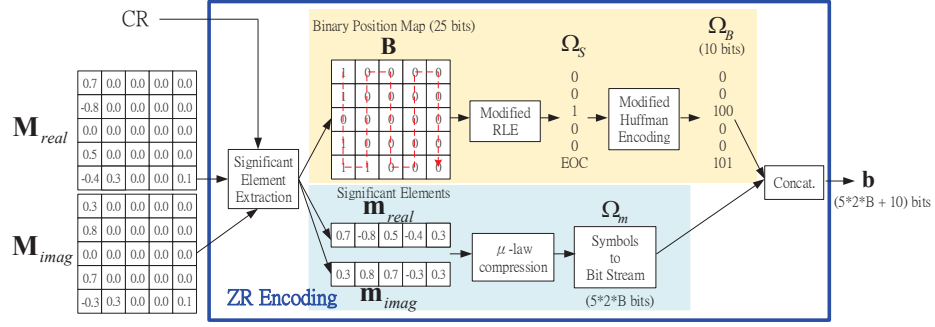
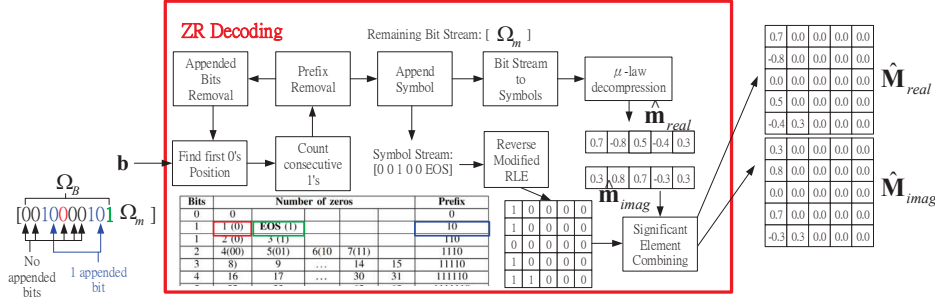
Ensure: $\hat{\mathbf{M}}_{\text{real}}$ and $\hat{\mathbf{M}}_{\text{imag}}$

```

flagEOS = False
 $\Omega_S = [\cdot]$ 
NS = 0  $\triangleright$  Initialize number of decoded symbols
while flagEOS = False do
  i  $\leftarrow$  Find index of the first 0 in b
  N1  $\leftarrow$  i - 1  $\triangleright$  = Prefix length + 1
  Remove prefix p = [11...10] from b  $\triangleright$  Prefix removal
  Find K by Table I according to p
  Find N0 by reading the first K bits in b by Table I
  Remove the K bits from b  $\triangleright$  Appended bits removal
   $\Omega_S \leftarrow [\Omega_S \ N_0]$   $\triangleright$  Append a symbol to symbol stream
  if N0  $\neq$  EOS then
    NS = NS + 1
  else
    flagEOS = True
  end if
end while
 $\mathbf{B} \leftarrow \text{mRLE}^{-1}(\Omega_S)$   $\triangleright$  By following Algorithm 1
 $\mathbf{b}_{\text{real}}, \mathbf{b}_{\text{imag}} \leftarrow \mathbf{b}$   $\triangleright$   $\mathbf{b} = \Omega_m$  since  $\Omega_B$  is removed from b
 $\hat{\mathbf{m}}_{\text{real}}, \hat{\mathbf{m}}_{\text{imag}} \leftarrow Q^{-1}(\mathbf{b}_{\text{real}}, \mathbf{b}_{\text{imag}})$ 
 $\hat{\mathbf{M}}_{\text{real}}, \hat{\mathbf{M}}_{\text{imag}} \leftarrow$  Significant element combining

```

approach. The ZR encoder and decoder at UEs can evaluate DL CSI recovery performance before sending feedback to the gNB. This enables UEs to select the most efficient sparse transformation mechanism to enhance feedback efficiency. Thus, we provide an extension of the proposed CSI feedback framework, Select ZR, depicted in Figure 4. Assuming there are 2^P available sparse transformations for the ZR approach, we evaluate DL CSI recovery performance by applying all sparse transformations before feedback is sent to the gNB, and select the best one for feedback. To indicate the transformation chosen by the UE, an additional P -bit information is appended


 Fig. 1. Example of ZR encoding (Note that the numbers 0.0 represent negligible elements in matrix $\mathbf{M}$).

 Fig. 2. Example of ZR decoding (Note that the numbers 0.0 represent negligible elements in matrix $\mathbf{M}$).

Algorithm 3 Inverse mRLE

Require: Ω_S
Ensure: $\mathbf{B}$
 $\mathbf{q} = []$
for $k = 1 : 1 : \text{length}(\Omega_S)$ **do**
if $\Omega_S[k] \neq \text{EOS}$ **then**
 $\mathbf{q} \leftarrow [\mathbf{q}; \underbrace{[0 \dots 0 1]^T}_{\Omega_S[k]}]$
else
 $\mathbf{q} \leftarrow [\mathbf{q}; \underbrace{[0 \dots 0]^T}_{N_a N_t - \text{length}(\mathbf{q})}] \triangleright \text{The remaining are 0's}$
end if
end for
 $\mathbf{B} \leftarrow \text{Reshape}^\dagger(\mathbf{q}) \triangleright \text{Reshape } \mathbf{q} \text{ to a matrix with size of } N_a \times N_t \text{ with a back-and-forth pattern shown in Figure 1}$

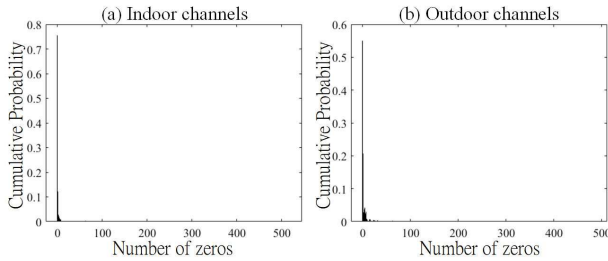
to the main feedback. By utilizing this approach, the DL CSI recovery performance can be significantly improved with only a few additional feedback bits.

IV. EXPERIMENTAL EVALUATIONS

A. Experiment Setup

Tests were conducted on indoor and outdoor channels using widely used channel model software, QuaTriGa and COST2100 [14], [15]. For the indoor setting, a circular cell with a 30 m radius and a gNB of height 20 m was used, while a 200 m radius circular cell was used for the outdoor environment. The simulators consider gNBs with an 8×4 UPA and 32-element ULA serving single-antenna UEs, respectively, with half-wavelength uniform spacing. The scenario features given in 3GPP TR 38.901 were followed, using $N_f = 1024$ subcarriers with 15K-Hz spacing and $M_f = 86$ pilots with a downsampling ratio of $\text{DR}_f = 12$ and assuming precise CSI estimates at the UEs. The NMSE metric (Eq. (6)) was used to assess performance.

For DL-based models, we conducted training with a batch size of 200 for 1000 epochs, starting with a learning rate of 0.001 and reducing it to 5×10^{-4} after 300 epochs. We generated both indoor and outdoor datasets using channel simulators, each consisting of 100,000 randomly generated channels. We used one-seventh of the channels for testing and divided the remaining channels into two-thirds for training and one-third for validation. In Section IV.C, we compared different sparse transformations with ZR approaches in various


 Fig. 3. Accumulative ratio of symbols in symbol list Ω_S for (a) indoor channels at 5 GHz and (b) outdoor channels at 300 MHz generated by QuaTriGa channel simulator.

scenarios and evaluated their performance against the proposed method, Select ZR, which outperformed AE-based models in different scenarios. In Section IV.D, we conducted a comprehensive evaluation of our proposed method by comparing it with SOTAs and more conventional alternatives.

B. Different Sparse Transformations in Various Scenarios

The DCT and Discrete Wavelet Transform (DWT) are highly effective for compressing images, making them a popular choice for classic JPEG-based compression. In contrast, the DFT is commonly used in CSI compression to transform complex periodic features into low-dimensional subspaces. Figure 5 shows the NMSE performance of ZR with different sparse transformations (DCT, DFT, DCT+DWT, DFT+DWT) and Select ZR ($P = 2$) at various compression ratios (i.e., $CR = 4, 8, 16, 32$) in indoor and outdoor channels generated by QuaDRiGa and COST2100. The results demonstrate that no single sparse transformation performs best across all scenarios. Some sparse transformations exhibit better recovery performance at higher compression ratios but worse recovery at lower ratios compared to other alternatives. This highlights the advantage of using Select ZR, which can automatically select the most effective sparse transformation for CSI feedback before transmitting UL feedback.

Figure 6 displays the NMSE performance of AE-based approaches (CsiNet [1], CRNet [3], CsiNetPro [6]) and Select ZR at different BPPs across various channel scenarios. Although Select ZR may not always outperform DL-based approaches, it still delivers acceptable performance across different channel types without requiring any prior training. Notably, previous work [16] demonstrates that DL-based models are sensitive to the channel type and typically suffer significant performance degradation when applied to a new propagation scenario. In practice, operators must collect new data and customize the model for new channels through retraining, which poses a significant practical challenge. In contrast, our approach can be deployed directly without retraining, providing an acceptable CSI recovery performance. Furthermore, Table II compares the computational complexity and storage requirements of the entire encoding/decoding process, highlighting the computa-

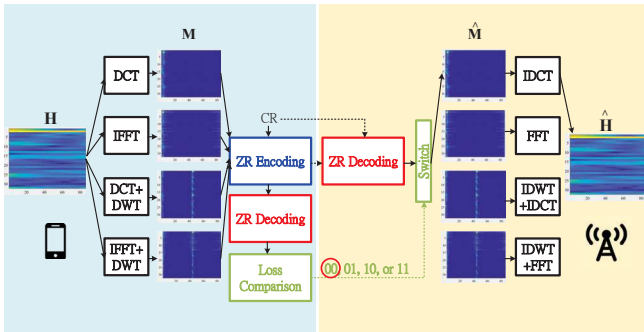


Fig. 4. Procedures of Select ZR. This example shows that we adopt 4 sparse transformations (i.e., $P = 2$).

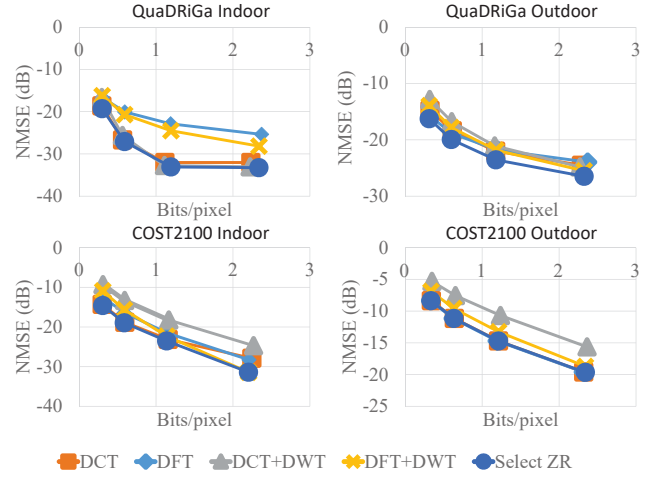


Fig. 5. Plot of NMSE versus bits/pixel for Select ZR and four different sparse transformations (DCT, DFT, DCT+DWT, DFT+DWT) used with ZR encoding in different test channels. Each curve shows the performance of ZR methods under different compression ratios, with the four anchor points on each curve corresponding to the use of compression ratios $CR = 4, 8, 16, 32$.

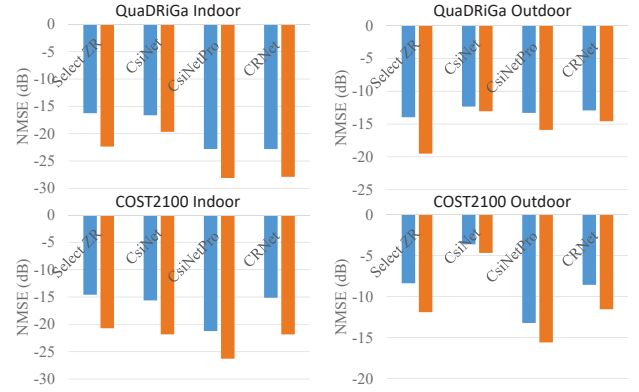


Fig. 6. NMSE performance for Select ZR and AE-based CSI feedback models in different test channels at two transmission costs (BPP = 0.375, 0.75).

tional and storage advantages of our approach over DL-based approaches¹.

C. Comparing with the state-of-the-arts (SOTAs)

To evaluate our method in a more comprehensive way, we compare it with several compressive sensing techniques such as Iterative Shrinkage Thresholding Algorithm (ISTA), Fast ISTA (FISTA) [17], and Least Absolute Shrinkage and

¹Select ZR is a rule-based approach and only needs to store the modified Huffman table.

TABLE II
COMPLEXITY AND STORAGE COMPARISON TO AE-BASED CSI FEEDBACK WHEN $CR = 16$ AND $P = 2$.

CR=16	Select ZR	CsiNet	CRNet	CsiNetPro
FLOPs	150K	5M	7M	68 M
Storage (32-bit elements)	21K	28M	28M	144M

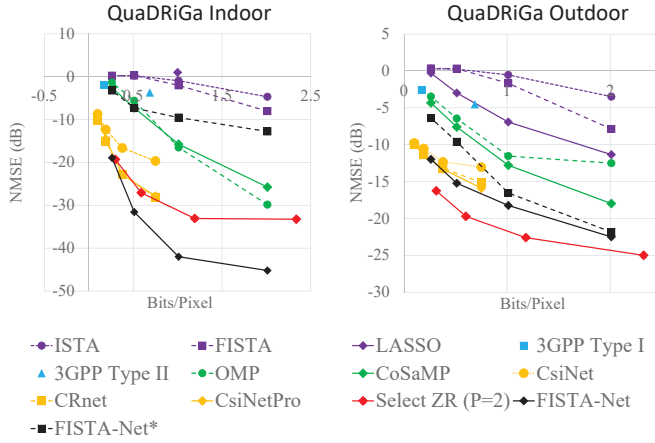


Fig. 7. Plot of NMSE versus bits/pixel for Select ZR and the SOTAs. Note that FISTA-Net* denotes the result when evaluating FISTA-Net with testing channels which are different from the training channels.

Selection Operator (LASSO [18]), as well as with orthogonal matching pursuit (OMP and CoSaMP [19]), 3GPP Type I [20], Type II [21], and a deep-unfolding DL framework [22].

Figure 7 illustrates the NMSE performance for Select ZR and the aforementioned SOTAs and conventional methods. Among the compared techniques, DL-based approaches (CsiNet, CRNet, CsiNetPro, FISTA-Net) and our proposed method demonstrate lower NMSE than other techniques. It is worth noting that DL-based approaches exhibit better performance in indoor scenarios, but perform worse in outdoor scenarios. We hypothesize that the reason behind this is related to the channel diversity in the delay domain. We find that DL models have difficulties in recovering data with diverse characteristics. In outdoor scenarios, the path delay profiles among different DL CSIs vary significantly. In contrast, Select ZR performs better than others in outdoor scenarios as the efficiency of our approach depends on delay sparsity and is barely related to the delay profile.

V. CONCLUSIONS

We propose ZR, a low-complexity, lightweight, scalable, and training-free CSI feedback framework for encoding and recovering downlink CSI in massive FDD MIMO wireless systems. It is free from customized training for different propagation channels at various compression ratios, and does not require high volumes of training data or multiple DL models for different RF channel environments and compression ratios. Unlike DL approaches, ZR can be directly applied to new and unseen channel scenarios without pre-training or customization. This flexible and scalable framework is simple to implement and amenable to broad deployment in practical massive MIMO wireless systems. Additionally, we provide an extension of ZR, Select ZR, for enhancing recovery performance by switching between different sparse transformation techniques. Numerical results demonstrate Select ZR's performance to be competitive with complex state-of-the-art

DL models such as CsiNet, CRNet, CsiNet-Pro, and FISTA-Net in most tested propagation channels. This new framework heralds a simple and easy-to-deploy CSI feedback approach that does not require a large dataset and can be rapidly deployed without any prior training.

REFERENCES

- [1] C. Wen, W. Shih, and S. Jin, "Deep Learning for Massive MIMO CSI Feedback," *IEEE Wirel. Commun. Lett.*, vol. 7, no. 5, pp. 748–751, 2018.
- [2] T. Wang, C.-K. 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, 2019.
- [3] Z. Lu, J. Wang, and J. Song, "Multi-resolution CSI Feedback with Deep Learning in Massive MIMO System," in *IEEE Intern. Conf. Communications (ICC)*, 2020, pp. 1–6.
- [4] Q. Yang, M. B. Mashhadi, and D. Gündüz, "Deep Convolutional Compression For Massive MIMO CSI Feedback," in *IEEE Intern. Workshop Mach. Learning for Signal Process. (MLSP)*, 2019, pp. 1–6.
- [5] S. Ji and M. Li, "CLNet: Complex Input Lightweight Neural Network Designed for Massive MIMO CSI Feedback," *IEEE Wirel. Commun. Lett.*, vol. 10, no. 10, pp. 2318–2322, 2021.
- [6] Z. Liu, M. Rosario, and Z. Ding, "A Markovian Model-Driven Deep Learning Framework for Massive MIMO CSI Feedback," *IEEE Trans. Wirel. Commun.*, 2021, early access.
- [7] J. Guo *et al.*, "DL-based CSI Feedback and Cooperative Recovery in Massive MIMO," *arXiv preprint arXiv:2003.03303*, 2020.
- [8] Z. Liu, L. Zhang, and Z. Ding, "An Efficient Deep Learning Framework for Low Rate Massive MIMO CSI Reporting," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4761–4772, 2020.
- [9] —, "Exploiting Bi-Directional Channel Reciprocity in Deep Learning for Low Rate Massive MIMO CSI Feedback," *IEEE Wirel. Commun. Lett.*, vol. 8, no. 3, pp. 889–892, 2019.
- [10] Y.-C. Lin, Z. Liu, T.-S. 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, 2021.
- [11] Y. L. J. Guo, C.-K. Wen, "Overview of Deep Learning-based CSI Feedback in Massive MIMO Systems," *arXiv preprint arXiv:2206.14383*, 2022.
- [12] Y. Yang, F. Gao, Z. Zhong, B. Ai, and A. Alkhateeb, "Deep Transfer Learning-Based Downlink Channel Prediction for FDD Massive MIMO Systems," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7485–7497, 2020.
- [13] J. Z. *et al.*, "Downlink CSI Feedback Algorithm With Deep Transfer Learning for FDD Massive MIMO Systems," *IEEE Trans. Cogn. Commun.*, vol. 7, no. 4, pp. 1253–1265, 2021.
- [14] S. Jaeckel *et al.*, "QuaDRiGa: A 3-D Multi-Cell Channel Model with Time Evolution for Enabling Virtual Field Trials," *IEEE Trans. Antennas and Propag.*, vol. 62, no. 6, pp. 3242–3256, 2014.
- [15] L. Liu *et al.*, "The COST 2100 MIMO Channel Model," *IEEE Wirel. Commun.*, vol. 19, no. 6, pp. 92–99, 2012.
- [16] Z. Lin and Z. Ding, "Training Deep Learning Models for Massive MIMO CSI Feedback with Small Datasets in New Environments," *arXiv preprint arXiv:2211.14785*, 2022.
- [17] A. Beck and N. Teboulle, "A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems," *Society for Industrial and Applied Mathematics*, vol. 2, no. 1, p. 183–202, Mar. 2009.
- [18] I. Daubechies, M. Debrise, and C. D. Mol, "An Iterative Thresholding Algorithm for Linear Inverse Problems with a Sparsity Constraint," *Comm. Pure and Applied Math.*, vol. 75, p. 1412–1457, 2004.
- [19] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.
- [20] Intel, "On NR Type I Codebook," TSG RAN WG1 No. 88 R1-1702205, 2021.
- [21] Samsung, "Type II CSI Reporting," TSG RAN WG1 No. 89 R1-1707962, 2021.
- [22] J. Guo, L. Wang, F. Li, and J. Xue, "CSI Feedback With Model-Driven Deep Learning of Massive MIMO Systems," *IEEE Wirel. Commun. Lett.*, vol. 26, no. 3, pp. 547 – 551, 2022.