

# Learning a Common Dictionary for CSI Feedback in FDD Massive MU-MIMO-OFDM Systems

PAVAN KUMAR GADAMSETTY <sup>1</sup> (Student Member, IEEE), K. V. S. HARI <sup>1</sup> (Fellow, IEEE),  
AND LAJOS HANZO <sup>2</sup> (Life Fellow, IEEE)

<sup>1</sup>Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru 560012, India

<sup>2</sup>School of Electronics and Computer Science, University of Southampton, SO17 1BJ Southampton, U.K.

CORRESPONDING AUTHOR: LAJOS HANZO (e-mail: lh@ecs.soton.ac.uk).

The work of Pavan Kumar Gadamsetty and K.V.S. Hari was supported by British Telecom India Research Centre (BTIRC), Indian Institute of Science, Bangalore.

The work of Lajos Hanzo was supported by Engineering and Physical Sciences Research Council under Projects EP/W016605/1, EP/X01228X/1, and EP/X01228X/1, and in part by European Research Council's Advanced Fellow Grant QuantCom Grant 789028.

**ABSTRACT** In a transmit preprocessing aided frequency division duplex (FDD) massive multi-user (MU) multiple-input multiple-output (MIMO) scheme assisted orthogonal frequency-division multiplexing (OFDM) system, it is required to feed back the frequency domain channel transfer function (FDCHTF) of each subcarrier at the user equipment (UE) to the base station (BS). The amount of channel state information (CSI) to be fed back to the BS increases linearly with the number of antennas and subcarriers, which may become excessive. Hence we propose a novel CSI feedback compression algorithm based on compressive sensing (CS) by designing a common dictionary (CD) to reduce the CSI feedback of existing algorithms. Most of the prior work on CSI feedback compression considered single-UE systems. Explicitly, we propose a common dictionary learning (CDL) framework for practical frequency-selective channels and design a CD suitable for both single-UE and multi-UE systems. A set of two methods is proposed. Specifically, the first one is the CDL-K singular value decomposition (KSVD) method, which uses the K-SVD algorithm. The second one is the CDL-orthogonal Procrustes (OP) method, which relies on solving the orthogonal Procrustes problem. The CD conceived for exploiting the spatial correlation of channels across all the subcarriers and UEs compresses the CSI at each UE, and upon reception reconstructs it at the BS. Our simulation results show that the proposed dictionary's estimated channel vectors have lower normalized mean-squared error (NMSE) than the traditional fixed Discrete Fourier Transform (DFT) based dictionary. The CSI feedback is reduced by 50%, and the memory reduction at both the UE and BS starts from 50% and increases with the number of subcarriers.

**INDEX TERMS** Wideband, frequency domain channel transfer function (FDCHTF), channel state information (CSI), compressive sensing (CS), massive MIMO, common dictionary learning (CDL), common dictionary (CD), orthogonal Procrustes (OP) problem, K-SVD algorithm.

## I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) systems constitute a promising enabling technique for 5 G/6 G cellular networks as a benefit of their substantial spatial multiplexing gain [1] in both time division duplex (TDD) and frequency division duplex (FDD) scenarios. At the base station (BS), combining the massive MIMO technology with orthogonal frequency-division multiplexing (OFDM)

is capable of transmitting multiple data symbols to multiple UEs on the same time-frequency resource block, resulting in increased system throughput [2]. In a multi-user (MU) massive MIMO-OFDM system, the knowledge of CSI<sup>1</sup> is needed at the BS to implement transmit precoding (TPC) for suppressing the co-channel interference (CCI) [3]. In FDD

<sup>1</sup>We use the terms CSI and FDCHTF interchangeably.

systems, due to the absence of channel reciprocity [4], the user equipment (UE) has to feed back the downlink (DL) frequency domain channel transfer function (FDCHTF) of each subcarrier to the BS. Feeding back the accurate CSI becomes more challenging with the increased number of antennas, subcarriers, and UEs [5].

The compression of high-dimensional CSI is essential for reducing the CSI feedback. The wireless channels can be represented in a sparse form in the spatial-frequency domain using ‘sparsifying’ bases termed as a dictionary [6]. In the compressive sensing (CS)-based feedback schemes [7], the original CSI is mapped to a sparse domain using a dictionary. The traditional choice of the dictionary is a fixed Discrete Fourier Transform (DFT) matrix. Then a random Gaussian measurement matrix is introduced to compress the sparse vector for feeding it back to the BS at a reduced rate. The sparse signal is then reconstructed at the BS using CS-based algorithms, such as the orthogonal matching pursuit (OMP) [8], basis pursuit (BP) [9] or covariance-assisted matching pursuit (CAMP) [10] procedures. The original CSI is then reconstructed by mapping the regenerated sparse signal back to the same dictionary used at the UE side.

The authors of [11] proposed a rotated version of the DFT basis to provide improved sparsity that results in reduced CSI mean-squared error (MSE) for a narrowband multi-user system with each UE having a single receive antenna. However, the proposed rotated basis does not exploit the antennas’ spatial correlation. The massive MIMO-OFDM channel between a multi-antenna UE and the BS can be represented by a matrix [12]. In such systems, the UE has to feed back the FDCHTF of each subcarrier, which results in huge feedback overhead. The FDCHTF feedback algorithm of a massive MIMO-OFDM single-UE system based on multidimensional compressive sensing theory using Tucker’s tensor decomposition model is developed in [13]. Briefly, Tucker’s tensor decomposition exploits the structure hidden in all the dimensions of the channel matrix and compresses it simultaneously in each dimension. The proposed scheme has a significant feedback reduction and hence improves the spectral efficiency. However, both the basis and measurement matrices should be learned. The authors of [14] introduced a recursive least squares dictionary learning algorithm (RLS-DLA) for CSI feedback. The proposed scheme achieves a substantial reduction in feedback requirements, however it requires the computation of large matrix inverses during the dictionary learning process.

Another line of work focused on designing non-dictionary-based methods for FDCHTF feedback [19]. In [20], an antenna grouping-based method was proposed for reducing the feedback overhead by grouping multiple correlated antenna elements into a single representative value. By considering a ray-based channel model, the authors of [21] and [22] designed an angle-of-departure (AoD) based adaptive subspace codebook for feedback compression. In [16] and [23] the authors exploited the low-rank characteristics of a large channel matrix for recovering the CSI at the BS.

Recent solutions include Deep Learning (DL) techniques conceived for CSI compression and recovery using so-called the Bi-LSTM [24], CsiNet-LSTM [25], DNNNet [26], CS-ReNet [27], and DCRNet [28] frameworks.<sup>2</sup> Additionally, the application of Deep unfolding techniques has also shown promising results, as demonstrated in [29], [30]. These techniques have better reconstruction performance than the conventional CS algorithms of [31], albeit at significantly increased computational complexity.

Massive MIMO-OFDM channels tend to be individually sparse and simultaneously share a common support set that typically exhibit joint sparsity in the time domain (TD) [32], which results in correlation among subcarriers in the frequency domain (FD). Since the DFT dictionary does not exploit spatial correlation across antenna arrays, we design a dictionary for massive MIMO-OFDM systems that can exploit the spatial correlation, hence achieving improved CSI reconstruction performance. The dictionary is generally learned from a training data set by relying on learning-based approaches [17], [18], [33], [34], [35]. The dictionaries learned have the potential to offer improved normalized mean squared error (NMSE) performance compared to fixed dictionaries, like the DFT-based one. In [15] a CS-based method was proposed, which exploited the spatial correlation among the antennas in a narrowband single-UE system using the K-SVD [36] algorithm. The method relies on learning the K-SVD dictionary from the training data set and on feeding back the K-SVD dictionary learned at the UE to the BS. Using this K-SVD dictionary, the CSI is compressed at the UE and reconstructed at the BS. The motivation for this K-SVD based dictionary is not only to reduce the CSI feedback, but also to reduce the NMSE of CSI reconstruction.

As the channel-induced dispersion is increased, the number of OFDM subcarriers also has to be increased to avoid an excessive performance degradation. Hence upon using the K-SVD algorithm of [15], the number of subcarrier K-SVD based dictionaries increases as the number of subcarriers increases. Handling ubiquitous subcarrier K-SVD based dictionaries is cumbersome in terms of memory management and feedback load. To circumvent this problem, we propose a novel common dictionary learning (CDL) technique, which can replace the requirement of individual subcarrier K-SVD dictionaries, leading to the concept of a common dictionary (CD). The CD effectively captures the channel characteristics of all the subcarriers and UEs, making it the optimal sparsifying dictionary for representing the channel’s sparsity in massive MIMO systems. Given the CD learned, compressive channel estimation techniques can be constructed for acquiring the CSI. A set of two methods having different pros and cons are proposed for CDL, namely the CDL-KSVD method and the CDL-orthogonal Procrustes (OP) [37] based method. These methods are detailed in Section III of the article. Again, our primary motivation is to reduce the CSI feedback overhead on the uplink as well as the memory requirement at

<sup>2</sup>For the expansion of these acronyms please refer to the relevant papers.

**TABLE 1. Comparing Our Contribution to the Existing Literature**

	[11]	[14]	[15]	[16]	[17]	[18]	Our
Massive MIMO architecture	✓	✓	✓	✓	✓	✓	✓
Spatially correlated channels	✓	✓	✓	✓			✓
Sparsity	✓	✓	✓		✓	✓	✓
Dictionary learning		✓	✓		✓	✓	✓
Feedback savings	✓	✓	✓	✓	✓	✓	✓
Memory savings at UE					✓		✓
Multi-user MIMO OFDM					✓		✓
UE mobility							✓

both the UE and the BS in FDD massive MU-MIMO-OFDM systems.

Main contributions of this article:

- 1) We proposed a novel CDL framework for learning a CD, mainly using the CDL-KSVD and the CDL-OP methods. In the CDL framework proposed for a multi-UE system, the CD conceived exploits the spatial correlation of the FD channels across all the subcarriers and the UEs. We demonstrate that this implementation improves the NMSE performance when compared to the existing methods.
- 2) In the CDL framework proposed for a single-UE system, the learning of CD is implemented at the UE. The UE sends only the CD to the BS in the uplink instead of all the subcarrier K-SVD dictionaries. This implementation reduces the dictionary feedback to the BS by a factor of  $N_c$  and also reduces the memory requirement by having a single CD at the UE and the BS.
- 3) We evaluate the proposed CD in the context of various system configurations and channel conditions in the face of UE mobility. The numerical results show a significant reduction in the NMSE of channel estimation and highlight the bit error rate (BER) performance of the channel estimates when using our learned dictionary. This corroborates the effectiveness of the CDL framework proposed over existing methods in wideband massive MIMO systems.

The remainder of this article is organized as follows. Section II presents the system model, CS procedure, and the motivation. In Section III, the proposed methods are discussed. Then the application of the proposed methods in wideband systems is discussed in Section IV. Our simulation results are provided in Section V to show the NMSE performance of the proposed method compared to state-of-the-art methods. Finally, in Section VI, our conclusions are given.

Table 1 presents the comparison of our contribution to the existing literature. Table 2 presents the list of acronyms. Table 3 presents the list of notations and in Table 4 the references for CSI compression techniques are given.

*Notations:* We use lower (upper) bold letters to denote column vectors (matrices) and super-scripts  $(\cdot)^{-1}$ ,  $(\cdot)^*$ ,  $(\cdot)^H$  to represent the inverse, complex conjugate and Hermitian operators respectively,  $\|\cdot\|_F$  denotes the Frobenius norm of

**TABLE 2. List of Acronyms**

Acronyms	Meaning
AOD	Angle-of-Departure
BER	Bit Error Rate
BS	Base Station
CD	Common Dictionary
CDL	Common Dictionary Learning
CFR	Channel Frequency Response
CIR	Channel Impulse Response
CS	Compressive Sensing
CSI	Channel State Information
DFT	Discrete Fourier Transform
DL	Deep Learning
FD	Frequency Domain
FDD	Frequency Division Duplex
FDCHTF	Frequency Domain Channel Transfer Function
GR-SVD	Golub-Reinsch SVD
K-SVD	K Singular Value Decomposition
MIMO	Multiple-Input Multiple-Output
MP	Matching Pursuit
MSE	Mean Squared Error
MU	Multi User
NMSE	Normalized Mean Squared Error
OF	Objective Function
OFDM	Orthogonal Frequency Division Multiplexing
OMP	Orthogonal Matching Pursuit
OP	Orthogonal Procrustes
QuaDRiGa	Quasi Deterministic Radio Channel Generator
RA	Receive Antenna
SU	Single User
TA	Transmit Antenna
TD	Time Domain
TDD	Time Division Duplex
TPC	Transmit Precoding
UE	User Equipment
VQC	Vector Quantization Codebook

a matrix;  $\otimes$  denotes the Kronecker product,  $\text{tr}(\cdot)$  is the trace of the matrix,  $\text{vec}(\cdot)$  operation returns a column vector by stacking all the columns of a matrix.

**II. SYSTEM MODEL**

In this section, we first introduce the massive MU-MIMO-OFDM channel and the associated spatial correlation matrices

TABLE 3. List of Notations

$N_t(N_r)$	Number of TAs (RAs)
$L$	Number of channel taps in TD
$i$	Index of the channel tap
$N_c$	Number of subcarriers in FD
$l$	Index of the subcarrier
$K$	Number of UEs
$k$	Index of the UE
$V$	Velocity of the UE
$S$	Sparsity index
$n$	Frame index
$\Psi$	Dictionary
$\Phi$	Measurement (Sensing) matrix
$\Theta$	Product of a dictionary and $\Phi$
$N_g$	Number of rows in $\Phi$
$g$	Compression factor
$M'$	Total number of training channel vectors
$N$	Number of training channel vectors considered across each subcarrier
$UE_k$	$k$ -th user equipment
$\mathbf{R}_{BS}$	Spatial correlation matrix at the BS
$\mathbf{R}_{UE,k}$	Spatial correlation matrix at the $k$ -th UE
$\tilde{\mathbf{H}}_{n,k}$	TD channel matrix across the $n$ -th tap of the $k$ -th UE
$\check{\mathbf{H}}_{n,k}$	TD i.i.d channel matrix
$\mathbf{H}$	Concatenated channel matrix of $K$ UEs
$\mathbf{H}_k$	Channel matrix of the $k$ -th UE
$\hat{\mathbf{H}}_k$	Reconstructed channel matrix of the $k$ -th UE at the BS
$\mathbf{H}'$	Matrix of training channel vectors at the UE
$\tilde{\mathbf{H}}'$	Sparse channel matrix representation of $\mathbf{H}'$
$\hat{\mathbf{H}}'$	Matrix of reconstructed training channel vectors at the BS
$\tilde{\hat{\mathbf{H}}}'$	Sparse channel matrix representation of $\hat{\mathbf{H}}'$
$\mathbf{H}_{l,n,k}$	Channel matrix or FDCHTF across the $l$ -th subcarrier of the $k$ -th UE in $n$ -th frame
$\hat{\mathbf{H}}_{l,n,k}$	Reconstructed channel matrix across the $l$ -th subcarrier of the $k$ -th UE in $n$ -th frame
$\mathbf{h}_{l,n,k}$	Channel vector transformation of $\mathbf{H}_{l,n,k}$
$\tilde{\mathbf{h}}_{l,n,k}$	Sparse channel vector representation of $\mathbf{h}_{l,n,k}$
$\hat{\tilde{\mathbf{h}}}_{l,n,k}$	Channel vector transformation of $\hat{\mathbf{H}}_{l,n,k}$
$\tilde{\hat{\mathbf{h}}}_{l,n,k}$	Sparse channel vector representation of $\hat{\tilde{\mathbf{h}}}_{l,n,k}$
$\mathbf{h}_{c,l,n}^k$	Compressed channel vector across the $l$ -th subcarrier of the $k$ -th UE in $n$ -th frame
$\Psi_{DFT}$	DFT dictionary
$\Psi_{ksvd,l}^k$	K-SVD dictionary across the $l$ -th subcarrier of the $k$ -th UE
$\Psi_c$	Common dictionary
$\Delta_{saved}$	Total memory storage reduction
$\Upsilon$	Dictionary feedback reduction factor
$\mathcal{T}_{ksvd}$	Feedback required for sending $N_c$ subcarrier dictionaries
$\mathcal{T}_{com}$	Feedback required for sending $\Psi_c$
$\mathcal{T}_{saved}$	Total dictionary feedback reduction

TABLE 4. References for CSI Compression Techniques

	Single-UE	Multi-UE
Narrowband	CS basis [14], [31]	Rotated DFT [11] K-SVD dictionary [15] Bi-LSTM [24]
Wideband	Multidimensional CSI [17] CsiNet [38] CsiNet-LSTM [25] DNNNet [26]	CS-ReNet [27]

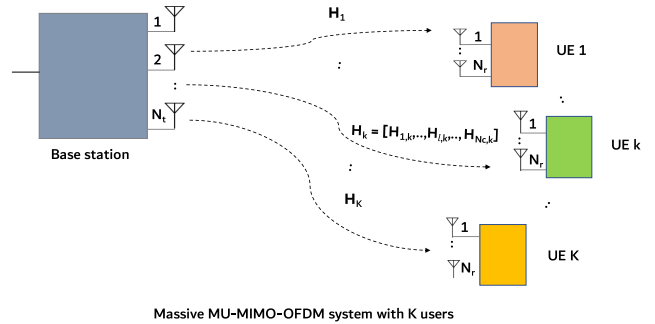


FIGURE 1. Overview of the considered massive MU-MIMO-OFDM system. Right:  $K$  UEs with  $N_r$  RAs each for  $k = 1$  to  $K$ ; Left: massive MIMO base station with  $N_t$  TAs.  $\mathbf{H}_k$  represents the complete FDCHTF of the  $k$ -th UE and  $\mathbf{H}_{l,k}$  represents the FDCHTF across the  $l$ -th subcarrier of the  $k$ -th UE.

at the BS and UEs. Furthermore, we conceive the CS-based channel reconstruction procedure of massive MIMO channels. Next, we highlight the dictionary learning algorithms available in the literature. Then, in the final sub-section we describe the motivation of the proposed CDL framework.

### A. THE MASSIVE MU-MIMO-OFDM CHANNEL

We consider a massive MU-MIMO-OFDM system using a uniform linear array (ULA) of  $N_t$  TAs at the BS,  $N_r$  RAs at all the  $K$  UEs and  $N_c$  subcarriers as shown in Fig. 1. For the  $k$ -th UE ( $k = 1$  to  $K$ ), consider a frequency-selective channel having  $L$  taps in the TD. Let  $\mathbf{H}_{l,k}$  represent the FDCHTF of the  $l$ -th subcarrier of the  $k$ -th UE given by

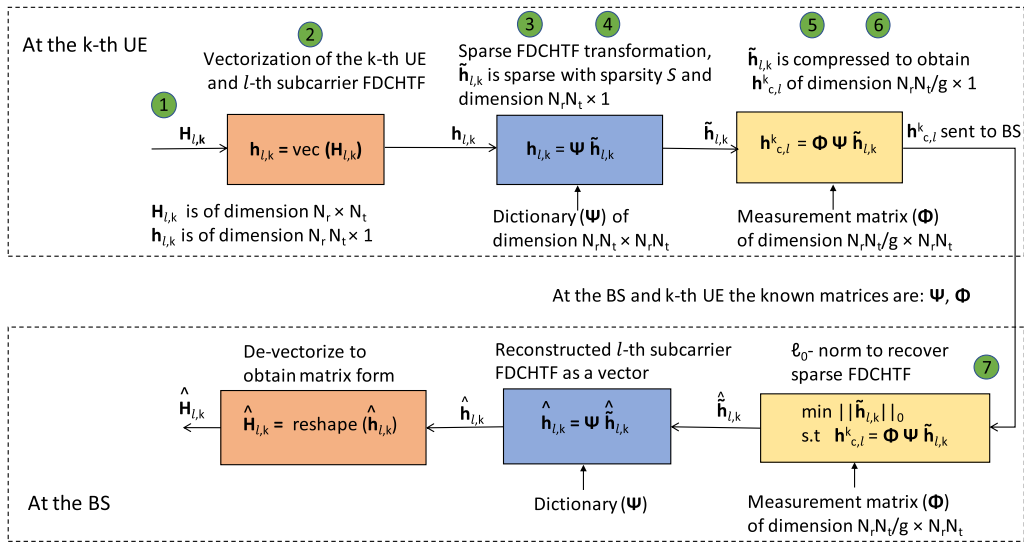
$$\mathbf{H}_{l,k} = \sum_{i=0}^{L-1} \tilde{\mathbf{H}}_{i,k} e^{-\frac{j2\pi il}{N_c}}, \quad (1)$$

where  $\tilde{\mathbf{H}}_{i,k} \in \mathbb{C}^{N_r \times N_t}$  is the  $i$ -th tap TD channel matrix. The tap coefficient  $\tilde{\mathbf{H}}_{i,k}(p, q)$  represents the channel impulse response (CIR) of the link spanning from the  $q$ -th BS antenna to the  $p$ -th UE antenna.

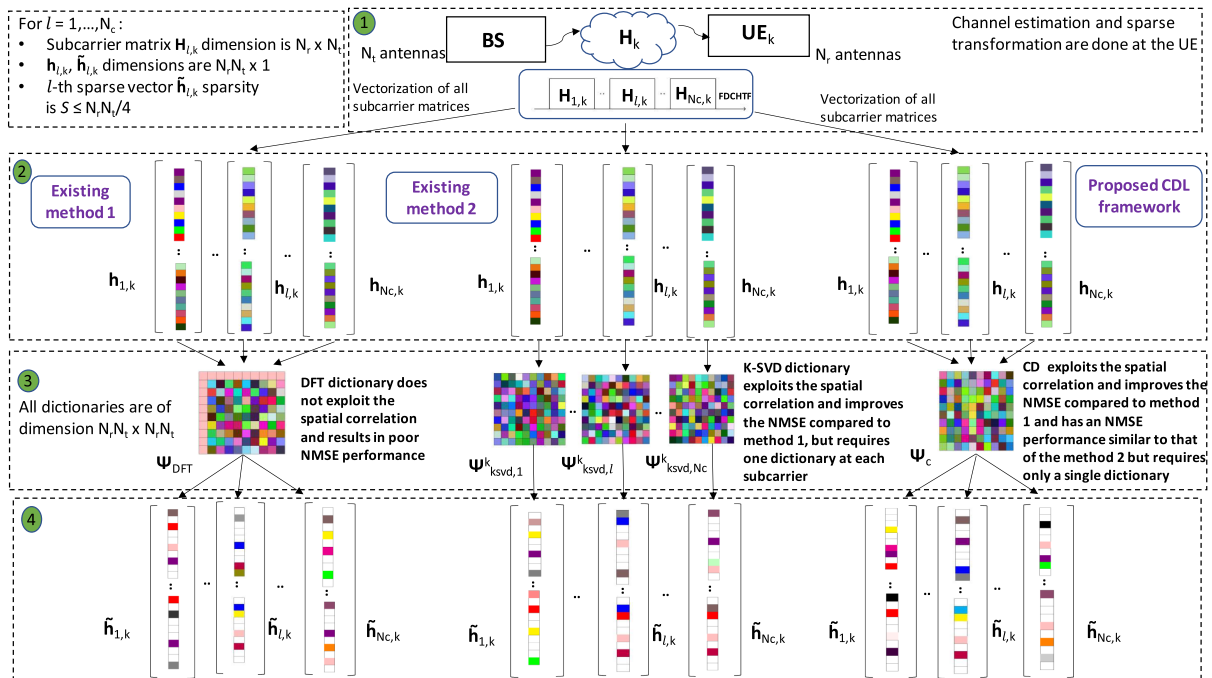
The spatial correlation of massive MIMO channels can be modeled by a Kronecker structure having separable transmit and receive correlation matrices [31], with  $\tilde{\mathbf{H}}_{i,k}$  given by

$$\tilde{\mathbf{H}}_{i,k} = \frac{1}{\sqrt{\text{tr}(\mathbf{R}_{UE,k})}} \mathbf{R}_{UE,k}^{\frac{1}{2}} \check{\mathbf{H}}_{i,k} \mathbf{R}_{BS}^{\frac{1}{2}}, \quad (2)$$

where  $\check{\mathbf{H}}_{i,k}$  is a  $N_r \times N_t$  matrix whose elements are independent and identically distributed (i.i.d.) complex zero-mean, unit variance, Gaussian random entries. Furthermore,  $\mathbf{R}_{BS}$  and



**FIGURE 2.** Overview of the considered CSI feedback compression scheme in the massive MU-MIMO-OFDM system. The FDCHTF across the  $l$ -th subcarrier of the  $k$ -th UE is compressed at the UE and reconstructed at the BS.



**FIGURE 3.** Overview of the massive MU-MIMO-OFDM system having a  $k$ -th UE with  $N_r$  RAs and a BS with  $N_t$  TAs.  $H_k$  represents the complete FDCHTF at the  $k$ -th UE and  $H_{l,k}$  represents the FDCHTF across the  $l$ -th subcarrier of the  $k$ -th UE. The FDCHTFs are vectorized and then undergo sparse transformation using the dictionary obtained from the existing methods 1, 2 of [15] and the proposed framework. The circled numbers and the notations are the same as in Fig. 2.

$\mathbf{R}_{UE,k}$  are the spatial correlation matrices at the BS and  $k$ -th UE, respectively.

The spatial correlation matrices are generated by Jakes' model often used in the literature, so the  $uv$ -th element of  $\mathbf{R}_{BS}$  and  $\mathbf{R}_{UE,k}$ , can be modeled by  $r_{uv} = J_0(2\pi d_{uv}/\lambda)$  where  $d_{uv}$  is the distance between the antennas  $u$  and  $v$ ,  $\lambda$  is the carrier wavelength and  $J_0(\cdot)$  denotes the zeroth-order Bessel function of first kind [39].

## B. COMPRESSIVE SENSING BASED CHANNEL RECONSTRUCTION

Fig. 2 shows the basic schematic of the FDCHTF compression and reconstruction across the  $l$ -th subcarrier of the  $k$ -th UE. More specifically, observe in Fig. 3 that at the  $k$ -th UE we present the FDCHTF view across all the subcarriers and its sparsification using the existing as well as the proposed methods in parallel. The remainder of

this section introduces each of the steps numbered in both figures.

① In between the BS and a  $k$ -th UE the complete channel frequency response matrix that includes all the  $N_c$  subcarrier channels is formed by stacking the channel matrices column-wise:

$$\mathbf{H}_k = [\mathbf{H}_{1,k}, \dots, \mathbf{H}_{l,k}, \dots, \mathbf{H}_{N_c,k}]. \quad (3)$$

② We assume that the  $k$ -th UE perfectly estimates its channel matrix  $\mathbf{H}_k$ , which should be shared with the BS through feedback. Instead of sending the FDCHTF of each subcarrier, the matrix  $\mathbf{H}_{l,k}$  is vectorized first into an  $N_r N_t \times 1$  column vector using the  $\text{vec}(\cdot)$  operation

$$\mathbf{h}_{l,k} = \text{vec}(\mathbf{H}_{l,k}). \quad (4)$$

③④ In practical systems, the UE has to compress the estimated channel vector  $\mathbf{h}_{l,k} \in \mathbb{C}^{N_r N_t \times 1}$  to avoid high feedback load. The wireless channel vector  $\mathbf{h}_{l,k}$  can be represented by a sparse vector [31] after a transformation

$$\mathbf{h}_{l,k} = \Psi \tilde{\mathbf{h}}_{l,k}, \quad (5)$$

where  $\tilde{\mathbf{h}}_{l,k}$  is the sparse representation of  $\mathbf{h}_{l,k}$ . The number of non-zero components of a sparse channel vector is called the sparsity or sparsity index, and it is denoted by  $S$ , while  $\Psi$  is a  $N_r N_t \times N_r N_t$  dictionary known to both the UE and the BS. A popular example of  $\Psi$  is the DFT matrix. Next, we introduce the measurement (sensing) matrix  $\Phi$ , which plays a crucial role in compressive sensing. The measurement (sensing) matrix defines the measurement process in CS, which influences the reconstruction quality and efficiency of the signal recovery algorithm. It is responsible for mapping the original high-dimensional signal to a lower-dimensional signal.

⑤⑥ To compress the channel vector  $\mathbf{h}_{l,k}$ , a measurement matrix  $\Phi \in \mathbb{C}^{N_g \times N_r N_t}$  ( $N_g \ll N_r N_t$ ) that satisfies the Restricted Isometry Property (RIP) [8], which facilitates sparse vector recovery relying on:

$$\mathbf{h}_{c,l}^k = \Phi \Psi \tilde{\mathbf{h}}_{l,k}, \quad (6)$$

where  $\mathbf{h}_{c,l}^k$  is the compressed channel vector with dimension  $N_g \times 1$ . Let us now define  $\Theta = \Phi \Psi$ .

⑦ Then the reconstruction of  $\mathbf{h}_{l,k}$  can be formulated as an  $\ell_0$ -norm minimization problem and the sparse vector  $\tilde{\mathbf{h}}_{l,k}$  can be obtained by solving

$$\min_{\tilde{\mathbf{h}}_{l,k}} \|\tilde{\mathbf{h}}_{l,k}\|_0 \quad \text{s.t.} \quad \mathbf{h}_{c,l}^k = \Theta \tilde{\mathbf{h}}_{l,k}. \quad (7)$$

Thus, instead of feeding back  $\mathbf{h}_{l,k}$ , the UE sends a low-dimensional vector  $\mathbf{h}_{c,l}^k$  to the BS for reducing the FDCHTF feedback. The BS reconstructs  $\hat{\mathbf{h}}_{l,k}$  from  $\mathbf{h}_{c,l}^k$ , where  $\hat{\mathbf{h}}_{l,k}$  represents the reconstructed  $\mathbf{h}_{l,k}$ . The reconstructed channel vector  $\hat{\mathbf{h}}_{l,k}$  at the BS is utilized for precoding during the data transmission stage. The precoder matrices employed at the BS are denoted as  $\mathbf{W}^s$  and  $\mathbf{W}_{op}^g$ , which correspond to the beamforming weights. These weights are obtained from the true channels and the channels estimated using CDL-OP dictionary, respectively, for a compression factor of  $g$ . Here,

the compression factor  $g$  is defined as  $g = \frac{N_r N_t}{N_g}$ , where  $N_g \times 1$  represents the dimension of the compressed channel vector  $\mathbf{h}_{c,l}^k$ .

### C. MOTIVATION FOR THE COMMON DICTIONARY LEARNING FRAMEWORK

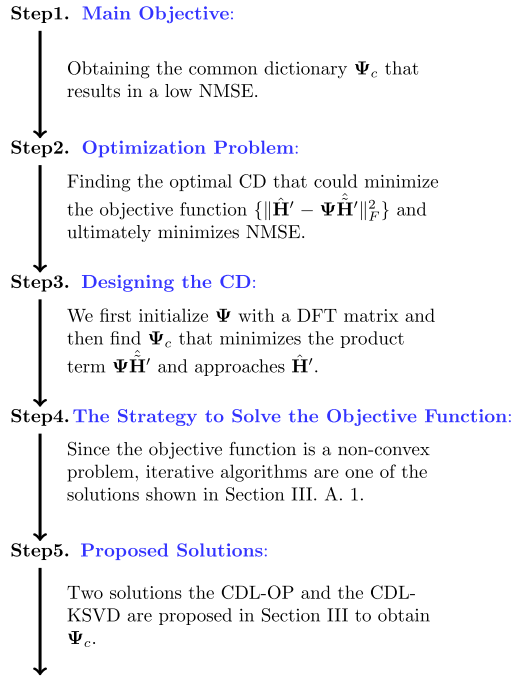
In CS-based feedback schemes, the traditional choice of the dictionary is a fixed DFT matrix, which does not exploit the spatial correlation between the antennas. The authors of [11] proposed a rotated version of the DFT dictionary for better exploiting the sparsity, resulting in reduced FDCHTF mean-squared error (MSE) for a narrowband multi-user system supporting single antenna UEs. But this rotated basis still failed to exploit the antenna's spatial correlation for improving the MSE further.

The authors of [15] have shown that a dictionary can be learned using the K-SVD algorithm for narrowband FDD massive SU-MIMO systems. This K-SVD dictionary learned exploits the spatial correlation between the antennas, and its FDCHTF reconstruction performance is improved compared to the fixed DFT dictionary. The proposed method requires FDCHTF and dictionary feedback to the BS. However, in practical communication systems, the channels are frequency-selective, and OFDM is a ubiquitous technique for such systems. In a massive MU-MIMO-OFDM system, to extend the idea of dictionary learning, it is necessary to feed back the FDCHTF and K-SVD based dictionary of each subcarrier of all UEs. Feeding back the entire FDCHTF  $\mathbf{H}_k$  of the  $k$ -th subcarrier will be a huge burden in the uplink. Another important issue is that substantial memory is required for saving all the  $N_c$  subcarrier dictionaries at both the UE and the BS. The dimension of each subcarrier dictionary is  $N_r N_t \times N_r N_t$ , hence the memory required to store  $N_c$  dictionaries is  $N_c (N_r N_t)^2$ .

To overcome these challenges, we propose the novel idea of a common dictionary, which can replace the requirement of individual subcarrier dictionaries. The CD is designed for exploiting the spatial correlation across all the subcarriers and UEs in the FD, hence improving the CSI reconstruction accuracy. The proposed CD reduces the CSI feedback load and memory requirement in both single and multi-UE systems. In particular, the feedback load is further reduced for a single-UE system by sending only a single CD from the UE to the BS. Hence, the proposed CDL framework reduces the CSI feedback and memory requirements, and we also study the NMSE performance compared to the DFT and subcarrier K-SVD dictionaries.

### III. PROPOSED COMMON DICTIONARY LEARNING FRAMEWORK

In this section, we detail the CDL framework proposed for a multi-UE system that constructs a dictionary from the estimated channel vectors and K-SVD based dictionaries of UEs. Before introducing the framework proposed, in Fig. 4 we provided a diagram showing the flow of the analysis described



**FIGURE 4.** Flow of the mathematical analysis.

in the article. This diagram guides the reader through the article.

## A. COMMON DICTIONARY LEARNING FRAMEWORK

The main goal of the proposed CDL framework is to construct a CD, denoted by  $\Psi_c$ , that can exploit the correlation of channels across all the UEs and BS, for improving the CSI reconstruction at the BS. The matrix of training channel vectors is denoted by  $\mathbf{H}'$ , which consists of  $M'$  channel vectors collected for  $N$  different frames across  $N_c$  subcarriers and  $K$  UEs. Then we have  $M' = N \times N_c \times K$ .

To elaborate further,  $\mathbf{H}'$  is structured as  $\mathbf{H}' = [\mathbf{H}'_1, \dots, \mathbf{H}'_k, \dots, \mathbf{H}'_K]$ , and each sub-matrix in  $\mathbf{H}'$  is represented as  $\mathbf{H}'_k = [\mathbf{H}'_{1,k}, \dots, \mathbf{H}'_{l,k}, \dots, \mathbf{H}'_{N_c,k}] \quad \forall k \in \{1, 2, \dots, K\}$ ,  $\forall l \in \{1, 2, \dots, N_c\}$ . Similarly,  $\mathbf{H}'_{l,k}$  is defined as  $\mathbf{H}'_{l,k} = [\mathbf{h}_{l_1,k}, \dots, \mathbf{h}_{l_n,k}, \dots, \mathbf{h}_{l_N,k}]$ , where  $\mathbf{h}_{l_n,k}$  represents the channel vector transformation of  $\mathbf{H}_{l_n,k}$  at the  $n$ -th MU-MIMO-OFDM frame (time-instant). We assume that the channel envelope remains constant for an OFDM frame and then changes for the successive frames, according to the vehicular velocity. Hence the consecutive frames are correlated.

The sparse representation of the matrix  $\mathbf{H}'$  is denoted as  $\tilde{\mathbf{H}}' = [\tilde{\mathbf{H}}'_1, \dots, \tilde{\mathbf{H}}'_k, \dots, \tilde{\mathbf{H}}'_K]$ , and each sub-matrix in  $\tilde{\mathbf{H}}'$  is represented as  $\tilde{\mathbf{H}}'_k = [\tilde{\mathbf{H}}'_{1,k}, \dots, \tilde{\mathbf{H}}'_{l,k}, \dots, \tilde{\mathbf{H}}'_{N_c,k}] \quad \forall k \in \{1, 2, \dots, K\}$ ,  $\forall l \in \{1, 2, \dots, N_c\}$ . Similarly  $\tilde{\mathbf{H}}'_{l,k} = [\tilde{\mathbf{h}}_{l_1,k}, \dots, \tilde{\mathbf{h}}_{l_n,k}, \dots, \tilde{\mathbf{h}}_{l_N,k}]$ , where  $\tilde{\mathbf{h}}_{l_n,k}$  denotes the sparse representation of the channel vector  $\mathbf{h}_{l_n,k}$ .

The CDL optimization problem is formulated as:

$$\min_{\Psi, \tilde{\mathbf{H}}'} \{\|\mathbf{H}' - \Psi\tilde{\mathbf{H}}'\|_F^2\}$$

$$\begin{aligned} \text{s.t. } & \|\tilde{\mathbf{h}}_{l_n,k}\|_0 \leq S, \quad \forall n \in \{1, 2, \dots, N\}, \\ & \forall l \in \{1, 2, \dots, N_c\}, \quad \forall k \in \{1, 2, \dots, K\}. \end{aligned} \quad (8)$$

To solve the optimization problem in (8) we propose the following methods.

### 1) CDL-KSVD METHOD

In the CDL-KSVD method the training set  $\mathbf{H}'$  consists of the channel vectors of all the UEs. The training set is employed to learn the dictionary  $\Psi_c$  using the K-SVD algorithm [36]. The K-SVD algorithm has two stages: the sparse coding stage and the dictionary update stage. In the sparse coding stage, each column of  $\mathbf{H}'$  is sparsely represented using a dictionary. The dictionary update stage involves updating each column of  $\Psi$  with a dominant singular vector. As a result, the learned dictionary  $\Psi_c$  has unit-norm columns.

- *Sparse coding stage:* In the first stage of the K-SVD algorithm, the optimization problem is formulated as:

$$\begin{aligned} \min_{\tilde{\mathbf{H}}'} \{ & \|\mathbf{H}' - \Psi\tilde{\mathbf{H}}'\|_F^2 \} \\ \text{s.t. } & \|\tilde{\mathbf{h}}_{l_n,k}\|_0 \leq S, \quad \forall n \in \{1, 2, \dots, N\}, \\ & \forall l \in \{1, 2, \dots, N_c\}, \quad \forall k \in \{1, 2, \dots, K\}. \end{aligned} \quad (9)$$

To solve (9), we begin by initializing the matrix  $\Psi$  with a DFT dictionary. The next step involves finding the matrix  $\tilde{\mathbf{H}}'$  having a sparse representation, which is an  $\ell_0$  problem and it is carried out by using the OMP algorithm [8]. The objective of the OMP algorithm is to find a sparse representation of  $\mathbf{H}'$  using a small number of non-zero elements in the matrix  $\tilde{\mathbf{H}}'$ .

- *Dictionary update stage:* In the second stage of the K-SVD algorithm, the optimization problem is formulated as:

$$\min_{\Psi} \{\|\mathbf{H}' - \Psi\tilde{\mathbf{H}}'\|_F^2\}. \quad (10)$$

The solution to the problem posed in (10) is obtained by updating each column of the dictionary by computing a partial SVD of a matrix [6]. After the dictionary update stage, the dictionary  $\Psi$  gets updated to  $\Psi_c$ . It is to be noted that this method updates only the columns corresponding to sparse coefficients of the channel matrix  $\tilde{\mathbf{H}}'$ .

- Repeat sparse coding and dictionary update stages until the stopping criterion is met.

The main advantage of the CDL-KSVD method is that it captures the spatial correlation of the channel vectors. But its drawback is that it requires a partial SVD operation for each column update in the dictionary update stage, which is computationally expensive.

### 2) CDL-OP METHOD

The CDL-OP method stands for common dictionary learning - orthogonal Procrustes method. In this method we solve the orthogonal Procrustes problem to learn the CD [37]. The CD obtained is a square matrix with dimensions  $N_r N_t \times N_r N_t$ . The constraint imposed is that the columns of the dictionary

should be orthogonal. The optimization problem is formulated as:

$$\min_{\Psi} \{ \|\mathbf{H}' - \Psi \tilde{\mathbf{H}}'\|_F^2 \} \quad \text{s.t. } \Psi^H \Psi = \mathbf{I}, \quad (11)$$

where  $\Psi$  in (11) may be found explicitly by singular value decomposition (SVD).

$$\begin{aligned} \text{Let } \mathbf{C} &= \tilde{\mathbf{H}}' \mathbf{H}'^H \\ [\mathbf{U}, \Sigma, \mathbf{V}] &= \text{SVD}(\mathbf{C}) \\ \Psi &= \mathbf{V} \mathbf{U}^H. \end{aligned} \quad (12)$$

The resulting dictionary  $\Psi$  obtained in (12) is the CD ( $\Psi_c$ ). The main advantage of the CDL-OP method is that it captures the spatial correlation of the channel vectors, but at the cost of an SVD operation.

### B. COMMON DICTIONARY FOR WIDEBAND MULTI-UE SYSTEM

A wideband channel has a broader signal bandwidth than the coherence bandwidth. In this section, we first discuss the CDL framework conceived for a wideband system supporting  $K$  UEs, and then highlight the simplified scenario, where a single UE is present. Next, we will quantify the memory savings of storing only a single CD. Then in the final sub-section, we elaborate on the dictionary feedback reduction by sending only a single CD in a single-UE system.

In the multi-UE system, since no communication takes place among the  $K$  UEs, CDL is impossible at any UE. Consequently, the CDL is only feasible at the BS. For the CDL at the BS, we require the subcarrier dictionaries and the reconstructed channels (used as training channel vectors). Let the reconstructed sparse channel matrix at the BS be represented by  $\hat{\mathbf{H}}' = [\hat{\mathbf{H}}'_1, \dots, \hat{\mathbf{H}}'_k, \dots, \hat{\mathbf{H}}'_K]$ , where each sub-matrix in  $\hat{\mathbf{H}}'$  is represented as  $\hat{\mathbf{H}}'_k = [\hat{\mathbf{H}}'_{1,k}, \dots, \hat{\mathbf{H}}'_{l,k}, \dots, \hat{\mathbf{H}}'_{N_c,k}] \forall k \in \{1, 2, \dots, K\}$ , with  $\hat{\mathbf{H}}'_{l,k} = [\hat{\mathbf{h}}_{l_1,k}, \dots, \hat{\mathbf{h}}_{l_n,k}, \dots, \hat{\mathbf{h}}_{l_{N_c},k}]$ , containing  $N$  reconstructed sparse channel vectors of each subcarrier, i.e.,  $\forall l \in \{1, 2, \dots, N_c\}$ .

Let the reconstructed matrix of training channel vectors at the BS be represented by  $\hat{\mathbf{H}}' = [\hat{\mathbf{H}}'_1, \dots, \hat{\mathbf{H}}'_k, \dots, \hat{\mathbf{H}}'_K]$ , where each sub-matrix in  $\hat{\mathbf{H}}'$  is represented as  $\hat{\mathbf{H}}'_k = [\hat{\mathbf{H}}'_{1,k}, \dots, \hat{\mathbf{H}}'_{l,k}, \dots, \hat{\mathbf{H}}'_{N_c,k}] \forall k \in \{1, 2, \dots, K\}$ , with  $\hat{\mathbf{H}}'_{l,k} = [\hat{\mathbf{h}}_{l_1,k}, \dots, \hat{\mathbf{h}}_{l_{N_c},k}]$  when considering  $N$  reconstructed channel vectors for each subcarrier i.e.,  $\forall l \in \{1, 2, \dots, N_c\}$ . The total number of reconstructed training channel vectors in  $\hat{\mathbf{H}}'$  is  $M' = N \times N_c \times K$ .

Importantly, at this stage we have to consider  $\hat{\mathbf{H}}'$  instead of  $\mathbf{H}'$  and  $\hat{\mathbf{H}}'$  instead of  $\tilde{\mathbf{H}}'$  in (8). The optimization problem of finding  $\Psi_c$  in the multi-UE system is formulated as follows:

$$\begin{aligned} \min_{\Psi, \hat{\mathbf{H}}'} & \left\{ \|\hat{\mathbf{H}}' - \Psi \hat{\mathbf{H}}'\|_F^2 \right\} \\ \text{s.t. } & \|\hat{\mathbf{h}}_{n,k}\|_0 \leq S, \quad \forall n \in \{1, 2, \dots, N\}, \\ & \forall l \in \{1, 2, \dots, N_c\}, \quad \forall k \in \{1, 2, \dots, K\}. \end{aligned} \quad (13)$$

The single-UE system ( $K = 1$ ) is a special case of a multi-UE system. For the current system when compared to the multi-UE system of Fig. 5 there is no need to send the subcarrier dictionaries ( $\Psi_{ksvd,l}^k, \forall n \in \{1, 2, \dots, N\}, \forall l \in \{1, 2, \dots, N_c\}$ ) from the UEs to the BS. The CD is learned at the UE itself using one of the two proposed methods and then the UE sends  $\Psi_c$  to the BS. Now both the BS and the UE will start using  $\Psi_c$ . The total number of training channel vectors in  $\mathbf{H}'$  is  $M' = N \times N_c$ .

The optimization problem of finding  $\Psi_c$  is as follows:

$$\begin{aligned} \min_{\Psi, \hat{\mathbf{H}}'} & \left\{ \|\mathbf{H}' - \Psi \tilde{\mathbf{H}}'\|_F^2 \right\} \\ \text{s.t. } & \|\tilde{\mathbf{h}}_n\|_0 \leq S, \quad \forall n \in \{1, 2, \dots, N\}, \\ & \forall l \in \{1, 2, \dots, N_c\}. \end{aligned} \quad (14)$$

The CDL framework of our multi-UE system is outlined in Algorithm 1 and its block diagram is shown in Fig. 5. More specifically, in the Algorithm 1 we present the step-by-step procedure of the CDL framework. The remainder of this subsection introduces each of the steps numbered in the Fig. 5 and the corresponding steps in the algorithm.

① The BS sends the pilots to all the  $K$ -UEs in the system and each UE estimates its channels and follows the steps 4 to 10 in the algorithm for  $N$  frames and learns the K-SVD based subcarrier dictionaries.

② Each UE sends the  $N_c$  K-SVD dictionaries to the BS. For the next  $N$  frames each UE compresses the  $l$ -th subcarrier FDCHTF using the K-SVD dictionary ( $\Psi_{ksvd,l}^k$ ) and sends it to the BS. Then the BS reconstructs the  $l$ -th subcarrier FDCHTF with the aid of the same K-SVD dictionary. Using the reconstructed FDCHTFs and the K-SVD based subcarrier dictionaries of all the  $K$  UEs, the BS learns the  $\Psi_c$ . The ② procedure corresponds to the steps 14 to 19 in the algorithm.

③ After learning the CD at the BS, the BS sends the CD to all the  $K$ -UEs, the UEs and the BS will follow the steps 22 to 25 in the algorithm for FDCHTF compression and reconstruction using  $\Psi_c$ .

#### 1) MEMORY REDUCTION CALCULATION

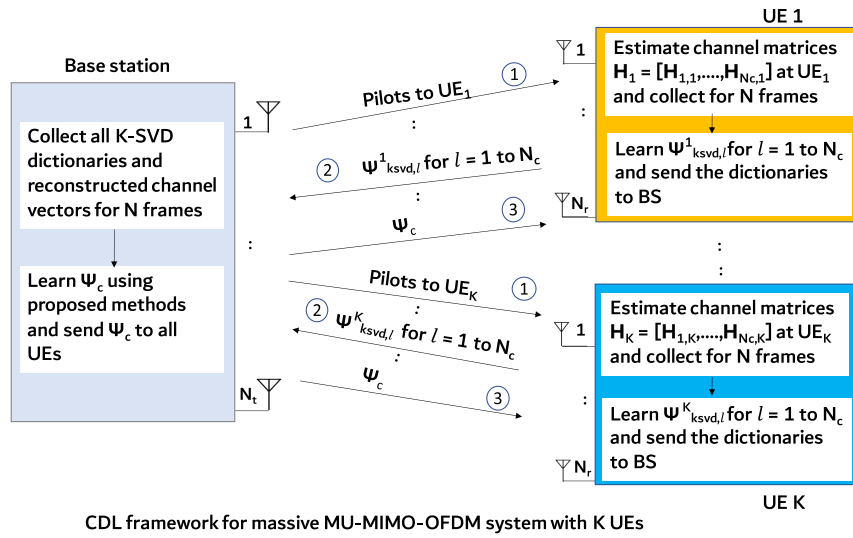
- The memory required to store each subcarrier dictionary  $\Psi_{ksvd,l}^k$  is  $(N_r N_t)^2, \forall l \in \{1, 2, \dots, N_c\}, \forall k \in \{1, 2, \dots, K\}$ .
- The total memory required to store  $N_c$  subcarrier dictionaries is  $N_c(N_r N_t)^2$ .
- The memory required to store  $\Psi_c$  is  $(N_r N_t)^2$ .
- Total memory storage reduction for a  $K$ -UE system =  $K \times ([\text{Memory required to store } N_c \text{ subcarrier dictionaries}] - [\text{Memory required to store } \Psi_c])$ , which is formulated as:

$$\Delta_{\text{saved}} = K(N_c - 1)(N_r N_t)^2. \quad (15)$$

- The total memory storage reduction for a single-UE is

$$\Delta_{\text{saved}} = (N_c - 1)(N_r N_t)^2. \quad (16)$$





**FIGURE 5.** Overview of the CDL framework for FDD massive MU-MIMO-OFDM system. Right:  $K$  UEs with  $N_r$  receive-antennas each; Left: massive MIMO base station with  $N_t$  transmit antennas; Center: For simplicity, only dictionary feedback is shown.

## 2) DICTIONARY FEEDBACK REDUCTION CALCULATION BY SENDING A CD IN A SINGLE-UE SYSTEM

- The dimension of each subcarrier K-SVD dictionary  $\Psi_{k_{svd},l}$  is  $N_r N_t \times N_r N_t, \forall l \in \{1, 2, \dots, N_c\}$ .
- The total dimension of  $N_c$  subcarrier dictionaries is  $N_c \times N_r N_t \times N_r N_t$ .
- The dimension of  $\Psi_c$  is  $N_r N_t \times N_r N_t$ .
- Total dictionary feedback reduction for a single-UE ( $\mathcal{T}_{saved}$ ) = [Feedback required for sending  $N_c$  subcarrier dictionaries ( $\mathcal{T}_{k_{svd}}$ )] – [Feedback required for sending  $\Psi_c$  ( $\mathcal{T}_{com}$ )], where we have

$$\mathcal{T}_{k_{svd}} = N_c (N_r N_t)^2$$

$$\mathcal{T}_{com} = (N_r N_t)^2$$

$$\mathcal{T}_{saved} = \mathcal{T}_{k_{svd}} - \mathcal{T}_{com} = (N_c - 1)(N_r N_t)^2.$$

- *Reduction in dictionary feedback:* We define the feedback reduction factor by ( $\Upsilon$ ):

$$\Upsilon = \frac{\mathcal{T}_{com}}{\mathcal{T}_{k_{svd}}} = \frac{1}{N_c} \quad (17)$$

For example, the dictionary feedback reduction factor in a single-UE system is provided in Table 5.

## 3) COMPUTATIONAL COMPLEXITY OF THE ALGORITHM

We calculate the computational complexity of the dictionary learning stage for both the CDL-OP and the CDL-KSVD methods.

### a) CDL-OP method:

- In (12), the SVD operation requires all the eigenvectors, resulting in a full SVD operation.
- The computation of a full SVD operation, specifically using the Golub-Reinsch SVD (GR-SVD) method, requires  $21(N_r N_t)^3$  floating-point operations (FLOPS) [6].

On the other hand, the Chan-SVD (R-SVD) method requires  $26(N_r N_t)^3$  FLOPS [6].

- Therefore, to update the dictionary, we require a computational complexity of  $\mathcal{O}[(N_r N_t)^3]$ .
- b) *CDL-KSVD method:*
  - In (10), updating each column of the dictionary requires an SVD operation. This SVD operation only requires the dominant eigenvector, resulting in a partial SVD computation.
  - The GR-SVD method requires  $14N_r N_t N_c^2 + 9N_c^3$  FLOPS [6], while the R-SVD method requires  $6N_r N_t N_c^2 + 20N_c^3$  FLOPS [6]. Here,  $N'_c$  represents the number of non-zero coefficients corresponding to the  $k$ -th row in  $\tilde{\mathbf{H}}'$ , and  $N'_c$  ranges from 0 to  $M'$ .
  - Therefore, to update the complete dictionary, a total of  $N_r N_t$  partial SVD operations are required.

c) For example, let us consider  $N_t = 64$ ,  $N_r = 1$ ,  $M' = 1600$ , and an average value of  $N'_c = M'/4$ . The computational complexity in FLOPS is provided in Table 6. We represent the number of FLOPS required for updating a single column in the dictionary by CDL-KSVD (min), and that imposed by updating all columns in the dictionary using CDL-KSVD (max).

## 4) CSI FEEDBACK CASE STUDY IN A SINGLE UE

The dimension of the compressed channel vector  $\mathbf{h}_{c,l}^k$  ( $\in \mathbb{C}^{N_g \times 1}$ ) sent from the UE in the uplink can be varied by adjusting the compression factor  $g$ . Specifically, the dimension is given by  $N_g = \frac{N_r N_t}{g}$ . By tuning the value of  $g$ , we can beneficially reduce the amount of CSI feedback required.

To quantify the feedback requirements, we introduce the variables  $\gamma_u$  and  $\gamma_c$  to represent the feedback for the non-dictionary and dictionary-based methods, respectively. In a non-dictionary based method without compression, the CSI

**Algorithm 1:** CDL Framework for  $K$  Users.

**Input:**  $\Psi_{DFT}$ , Pilots  
**Output:**  $\Psi_c$

- 1 **Variables:** Frame index  $n$ , UE index  $k$ , subcarrier index  $l$ .
- 2 **Implementation:**  $\forall k \in \{1, \dots, K\}, \forall l \in \{1, \dots, N_c\}$
- 3 **for**  $n = 1$  to  $N$  **do**
- 4     BS transmits pilot sequences for UE<sub>1</sub> to UE<sub>K</sub>.
- 5     The  $k$ -th UE estimates its channel matrix  $\mathbf{H}_{l_n,k}$  across  $l$ -th subcarrier from the pilots.
- 6     After transformatation of  $\mathbf{H}_{l_n,k}$  to vector  $\mathbf{h}_{l_n,k}$ ,  $k$ -th UE at  $l$ -th subcarrier computes its compressed vector  $\mathbf{h}_{c,l_n}^k$  using  $\mathbf{h}_{l_n,k}$ ,  $\hat{\mathbf{h}}_{l_n,k}$ ,  $\Psi_{DFT}$  and sends  $\mathbf{h}_{c,l_n}^k$  to BS.
- 7     BS reconstructs  $\hat{\mathbf{h}}_{l_n,k}$  and  $\hat{\mathbf{h}}_{l_n,k}$  using  $\Psi_{DFT}$  and  $\mathbf{h}_{c,l_n}^k$ . BS uses  $\hat{\mathbf{h}}_{l_n,k}$  for precoding.
- 8     In parallel  $k$ -th UE at  $l$ -th subcarrier collects training channel vectors  $\mathbf{h}_{l_n,k}$ 's to form matrix  $[\mathbf{h}_{l_1,k}, \dots, \mathbf{h}_{l_N,k}]$ .
- 9     **if**  $n = N$  **then**
- 10         The  $k$ -th UE learns  $\Psi_{ksvd,l}^k$  from  $[\mathbf{h}_{l_1,k}, \dots, \mathbf{h}_{l_N,k}]$  and  $\Psi_{DFT}$  using K-SVD algorithm and sends  $\Psi_{ksvd,l}^k$  (once) to BS.
- 11     **end**
- 12 **end**
- 13 **for**  $n = N + 1$  to  $2N$  **do**
- 14     Follow steps 4 and 5 for another  $N$  frames.
- 15     The  $k$ -th UE at  $l$ -th subcarrier computes compressed vector  $\mathbf{h}_{c,l_n}^k$  using  $\mathbf{h}_{l_n,k}$ ,  $\hat{\mathbf{h}}_{l_n,k}$ ,  $\Psi_{ksvd,l}^k$  and sends  $\mathbf{h}_{c,l_n}^k$  to BS.
- 16     BS reconstructs  $\hat{\mathbf{h}}_{l_n,k}$ ,  $\hat{\mathbf{h}}_{l_n,k}$  using  $\Psi_{ksvd,l}^k$ ,  $\mathbf{h}_{c,l_n}^k$ .
- 17     **if**  $n = 2N$  **then**
- 18         BS computes  $\Psi_c$  using one of the two methods from  $\hat{\mathbf{h}}_{l_n,k}$ 's,  $\hat{\mathbf{h}}_{l_n,k}$ 's and  $\Psi_{ksvd,l}^k$  of all UEs and sends  $\Psi_c$  to all the UEs.
- 19     **end**
- 20 **end**
- 21 **for**  $i > 2N$  **do**
- 22     Follow steps 4 and 5
- 23     The  $k$ -th UE starts using  $\Psi_c$  instead of  $\Psi_{ksvd,l}^k$  for step 15 and sends compressed vector  $\mathbf{h}_{c,l_n}^k$  to BS.
- 24     BS reconstructs  $\hat{\mathbf{h}}_{l_n,k}$  and  $\hat{\mathbf{h}}_{l_n,k}$  of  $k$ -th UE using  $\Psi_c$  and  $\mathbf{h}_{c,l_n}^k$  (obtained from step 23).
- 25     BS uses  $\hat{\mathbf{h}}_{l_n,k}$  of each UE for precoding.
- 26 **end**

fed back from the UE corresponds to  $N_c N_r N_t$  elements for one frame. For  $N'$  frames the CSI feedback will be  $\gamma_u = N' N_c N_r N_t$ . However, in a dictionary-based method associated with compression, the feedback is constituted by the CSI information having  $\frac{N_c N_r N_t}{g}$  elements for one frame, along with a one-time transmission of a dictionary with  $(N_r N_t)^2$  elements.

**TABLE 5.** Dictionary Feedback Reduction Factor in a single-UE System

$N_t$	$N_r$	$N_c$	$\mathcal{T}_{ksvd}$	$\mathcal{T}_{com}$	$\Upsilon$
64	1	4	16384	4096	1/4
64	1	32	131072	4096	1/32

**TABLE 6.** Computational Complexity in FLOPS

Method	GR-SVD	Chan-SVD
CDL-OP	$5.505 * 10^6$	$6.8157 * 10^6$
CDL-KSVD (min)	$7.1936 * 10^8$	$1.3414 * 10^9$
CDL-KSVD (max)	$4.6039 * 10^{10}$	$8.5852 * 10^{10}$

**TABLE 7.** CSI Feedback Savings Comparison Table for  $N' = 2^{10}$

$g$	$N_c$	$N_t$	$N_r$	$\Gamma = \frac{\gamma_c}{\gamma_u}$
2	32	64	1	0.5
2	64	64	2	0.5
4	32	64	1	0.252
4	64	64	2	0.252

So, for a total of  $N'$  frames the feedback is given by  $\gamma_c = (N_r N_t)^2 + \frac{N' N_c N_r N_t}{g}$ .

We define  $\Gamma$  as the CSI feedback ratio, which is calculated as the ratio between  $\gamma_c$  and  $\gamma_u$ . If  $\Gamma < 1$ , it indicates that the value of  $\gamma_c$  is lower than  $\gamma_u$ , resulting in a saving in CSI feedback.

$$\Gamma = \frac{\gamma_c}{\gamma_u} = \frac{(N_r N_t)^2 + \frac{N' N_c N_r N_t}{g}}{N' N_c N_r N_t} = \frac{N_r N_t}{N' N_c} + \frac{1}{g} \quad (18)$$

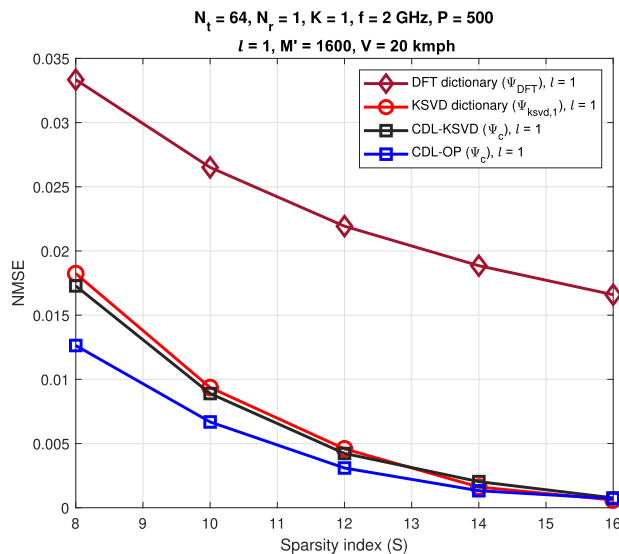
For example in Table 7 we consider a scenario associated with  $N' = 2^{10}$  and vary the values of  $N_c$ ,  $N_r$ , and  $g$ . Using the formula given in 18, we demonstrate significant reductions in CSI feedback. The level of compression applied to the channel vector  $\mathbf{h}_{c,l}^k$  depends on the sparsity parameter  $S$ . As per the lower bound, the number of elements in  $\mathbf{h}_{c,l}^k$  must satisfy  $N_g > 2S$  [9]. In Fig. 13, we illustrate the impact of the compression factor  $g$  on the BER vs. signal-to-noise ratio (SNR) performance for both the non-dictionary and dictionary-based methods.

#### IV. NUMERICAL RESULTS

In this section, we provide the simulation results for characterising the NMSE performances of the DFT dictionary ( $\Psi_{DFT}$ ), of the individual K-SVD based subcarrier dictionaries ( $\Psi_{ksvd,l}^k$ ) and of the proposed CD ( $\Psi_c$ ). Using the DFT dictionary as the initial reference dictionary, we obtain the individual K-SVD based subcarrier dictionaries. Then  $\Psi_c$  is learned by the proposed methods. For NMSE calculations, each dictionary is used for reconstructing the  $P$  channel vectors of each subcarrier at the BS.

The NMSE of the reconstructed channel is used as a performance metric defined as

$$\text{NMSE} = \frac{1}{P} \sum_{i=1}^P \frac{\|\hat{\mathbf{h}}_{l_n,k} - \mathbf{h}_{l_n,k}\|_2^2}{\|\mathbf{h}_{l_n,k}\|_2^2}, \quad (19)$$



**FIGURE 6.** Single-UE system is considered, where the NMSE performance comparison among all the dictionaries is carried out for  $N_t = 64$ ,  $N_r = 1$  and  $N_c = 32$ . Subcarrier 1 is considered for comparison and the UE velocity is 20 kmph.

where  $\hat{\mathbf{h}}_{l,n,k}$  is the reconstructed channel vector and  $\mathbf{h}_{l,n,k}$  is the original one.

### A. SIMULATION SETTINGS

The simulations are carried out for a massive MIMO-OFDM system having  $N_t = 64$ , antenna spacing of  $d = \lambda/15$ , carrier wavelength  $\lambda$ , operating at a carrier frequency of  $f_c = 2$  GHz. Furthermore, we have a communication bandwidth of  $B = 20$  MHz,  $K = 1$  and  $K = 3$  UEs,  $N = 50$ ,  $P = 500$  test channel vectors,  $N_c = 32$  and  $M = N_r N_t / 2$ . For  $M = 32$ , all the existing and the proposed dictionaries reduce the CSI feedback by 50% in the uplink. For all the experiments, the subcarrier K-SVD based dictionaries are learned for  $N = 50$  from each subcarrier.

For experiments in the multi-UE system, the channels are generated using a Quasi Deterministic Radio Channel Generator (QuaDRiGa) [40], [41] for the three UEs having velocities of  $V = 10, 15$  and 20 kmph. For experiments in the single-UE system, the channels are taken from the UE of  $V = 20$  kmph. The channel update rate (CUR) considered to generate channels in QuaDRiGa is 10 ms. The QuaDRiGa simulation platform is recommended by 3GPP (3rd Generation Partnership Project) for designing and simulating wireless communications systems.

The main motivation for the CDL is not only to reduce the CSI feedback but also to reduce the CSI reconstruction NMSE. The benefit of the proposed methods in terms of the NMSE performance has to be studied. Hence we have conducted experiments for determining which of the proposed methods will best replace the DFT dictionary and subcarrier K-SVD based dictionaries in both single-UE and multi-UE systems. Since we have considered  $N_c = 32$ , it is not feasible to show the CD performance across all the subcarriers.

In the single-UE system, we have considered subcarriers  $l = 1$  or 8, since the channel gains of these two subcarriers are relatively low over the period of time. In the multi-UE system, we have considered the first subcarrier  $l = 1$  of all the three UEs to evaluate the NMSE performance as a function of sparsity and compared the proposed methods' CD performance to the DFT and K-SVD dictionaries in the literature.

In OFDM systems, the need for subcarrier K-SVD based dictionaries increases with the number of subcarriers. The FDCHTFs of each subcarrier are considered to be independent in a wideband OFDM system, but this is only realistic for extremely long CIRs. Hence in the proposed system we assume having realistic correlation among subcarriers in the FD. This correlation among the subcarriers is captured by the CD using one of the two proposed methods. Our CDL procedure may also be extended to larger  $N_c$ .

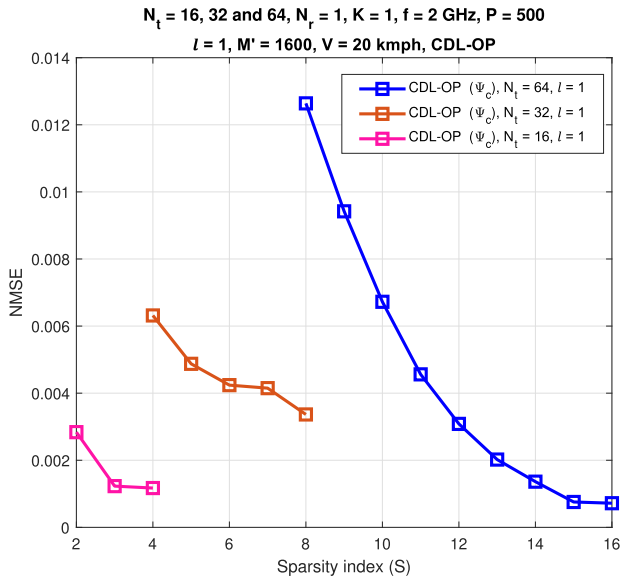
In all the simulation results, for each subcarrier it can be observed that the NMSE decreases as the sparsity increases. This is because in the sparse vector transformation (5) the sparse vector  $\hat{\mathbf{h}}$  picks many columns in the dictionary for higher sparsities, which in turn helps to solve the optimization problem (7) by minimizing the distance between  $\mathbf{h}$  and  $\hat{\mathbf{h}}$ . Hence, higher sparsity will improve the reconstruction performance.

*Experiment 1:* In this single-UE experiment, we study the NMSE performance across a particular subcarrier of the UE using the CD learned from the proposed methods and existing methods as a function of sparsity. The proposed methods' CD is learned across all the subcarriers of the UE. For this experiment we consider a massive SU-MIMO-OFDM system having  $N_c = 32$  subcarriers and a UE having  $N_r = 1$  RA and moving with a velocity of 20 kmph.

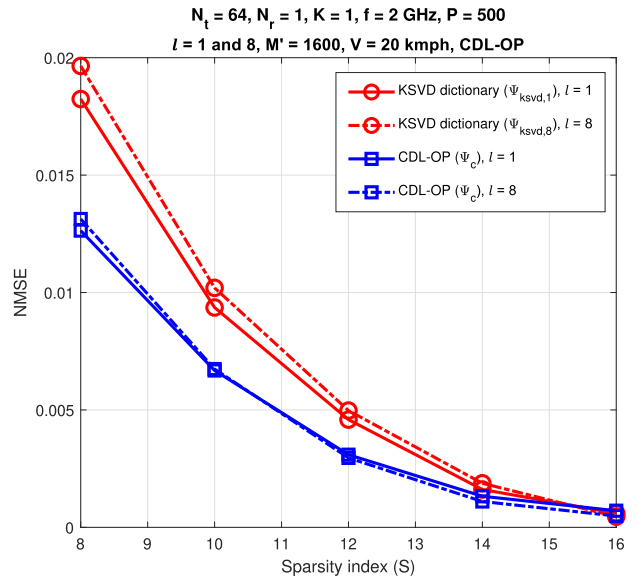
In Fig. 6, we consider the subcarrier  $l = 1$  to study the NMSE performance of all the dictionaries. For a particular sparsity index of  $S = 8$ , the NMSE of the CDL-KSVD dictionary is  $1.7 \times 10^{-2}$ , of the CDL-OP dictionary is  $1.3 \times 10^{-2}$ , of the subcarrier K-SVD based dictionary is  $1.8 \times 10^{-2}$ , and of the DFT dictionary is  $3.3 \times 10^{-2}$ . The CDL-OP dictionary has the lowest NMSE at  $S = 8$ . For all other sparsities, the CDL-OP and the CDL-KSVD method dictionaries perform similarly, and the NMSE values are close to those of the subcarrier K-SVD dictionary. All the proposed methods' CD exhibit better performance than the DFT dictionary.

In Fig. 7, we consider a scenario where the BS has a different number of antennas, namely  $N_t = 16, 32$ , and 64. We consider the subcarrier  $l = 1$  to study the NMSE performance of the CDL-OP dictionaries learned for different  $N_t$  values. We observe that the NMSE value increases with the number of antennas at the BS.

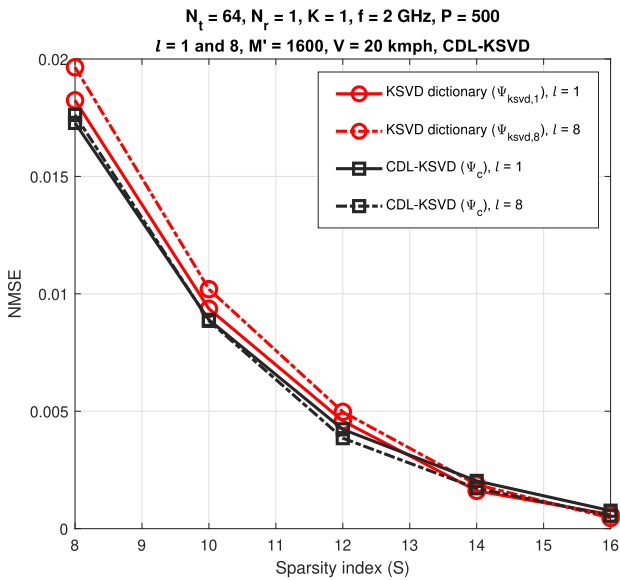
*Experiment 2:* In this single-UE experiment, we study the NMSE performance across different subcarriers of the UE employing the CD learned by the proposed methods and the individual subcarrier K-SVD based dictionaries as a function of sparsity. For this experiment we consider a massive SU-MIMO-OFDM system having  $N_c = 32$  subcarriers. The UE is



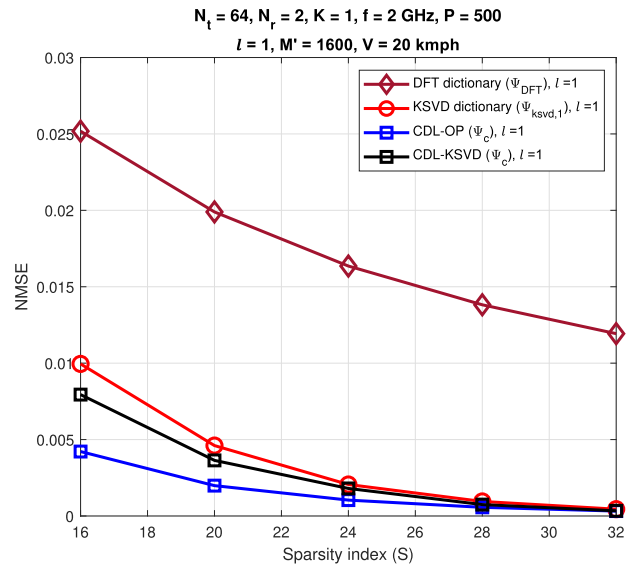
**FIGURE 7.** Single-UE system is considered, where the NMSE performance comparison among the CDL-OP dictionaries is carried out for  $N_t = 16, 32$  and  $64, N_r = 1$  and  $N_c = 32$ . Subcarrier 1 is considered for comparison and the UE velocity is 20 kmph.



**FIGURE 9.** Single-UE system is considered, where the NMSE performance of the K-SVD dictionary and the CDL-OP dictionary are used for  $N_t = 64, N_r = 1$  and  $N_c = 32$ . Subcarriers 1 and 8 are considered for comparison and the UE velocity is 20 kmph.



**FIGURE 8.** Single-UE system is considered, where the NMSE performance of the K-SVD dictionary and the CDL-KSVD dictionary are used for  $N_t = 64, N_r = 1$  and  $N_c = 32$ . Subcarriers 1 and 8 are considered for comparison and the UE velocity is 20 kmph.



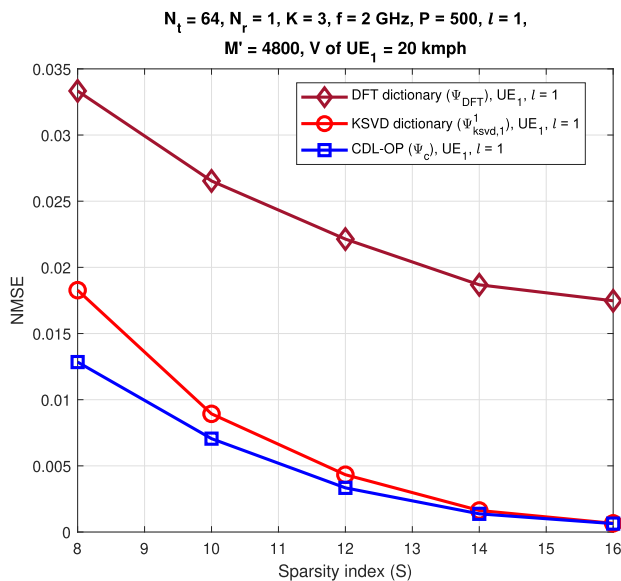
**FIGURE 10.** Single-UE system with multiple UE antennas is considered, where the NMSE performance comparison among all the dictionaries is carried out for  $N_t = 64, N_r = 2$  and  $N_c = 32$ . Subcarrier 1 is considered for comparison and the UE velocity is 20 kmph.

equipped with  $N_r = 1$  (and 2) RAs and is moving at a velocity of 20 kmph.

In Fig. 8, the  $\Psi_c$  employed for NMSE characterization is learned from the CDL-KSVD method and in Fig. 9, the  $\Psi_c$  employed for NMSE characterization is learned from the CDL-OP method. Observe from Figs. 8 and 9, for subcarriers  $l = 1$  and 8 at all the sparsity index values, the NMSE values are close to each other, but the CDL-OP method is the best among all the three methods in terms of the NMSE performance attained. Both the CDL-KSVD and CDL-OP

methods rely on the SVD operation and learn the CD from the channels estimated at the UE. Consequently, the NMSE reconstruction results shown in Figs. 8 and 9 exhibit a high degree of similarity.

We have carried out a simulation and presented the results in Fig. 10, where we specifically focus on the scenario where  $N_r > 1$ , indicating the presence of multiple receive antennas. By considering this scenario, we ensure that our analysis is not limited to a specific number of receive antennas. It is



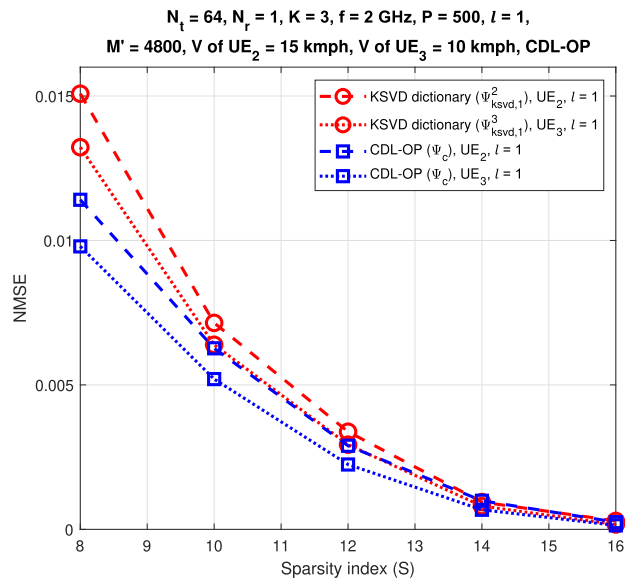
**FIGURE 11.** Multi-UE system is considered, where the NMSE performance comparison of the DFT, the K-SVD and the CDL-OP methods is carried out for  $K = 3$ ,  $N_t = 64$ ,  $N_r = 1$  and  $N_c = 32$ . Subcarrier 1 of  $UE_1$  is considered for comparison. The velocities of the three UEs are 10, 15, and 20 kmph.

observed that all the proposed methods' CD exhibit better NMSE performance than the DFT dictionary.

*Experiment 3:* In this multi-UE experiment, we initially study the NMSE performance of the dictionary generated using the CDL-OP method and existing methods for a particular UE's subcarrier as a function of sparsity. Then we study the NMSE performance of the CDL-OP dictionary across different UEs for a particular subcarrier. The CDL-OP dictionary is learned across all the subcarriers and all the  $K$  UEs. For the experiment, we consider a massive MU-MIMO-OFDM system having  $N_c = 32$  subcarriers and  $K = 3$  UEs each associated with  $N_r = 1$  RA.

The wireless channels are generated for three UEs having velocities of  $V = 10, 15$ , and  $20$  kmph using a QuaDRiGa simulator. Let us assume that a UE changes its velocity to that of another UE, but experiences different channel characteristics. In that case, there is no need to learn a new dictionary for that particular UE. The CD has already captured all the three UE channel characteristics, and this procedure can be extended to larger  $K$  and  $N_c$  values.

In Fig. 11, we consider the first UE and subcarrier  $l = 1$  to study the NMSE performance of the dictionaries from the full set of  $N_c = 32$  subcarriers and  $K = 3$  UEs. For a particular sparsity of  $S = 8$ , the NMSE value of the CDL-OP dictionary is  $1.3 \times 10^{-2}$ , of the subcarrier K-SVD based dictionary is  $1.8 \times 10^{-2}$ , and of the DFT dictionary is  $3.3 \times 10^{-2}$ . We observe that all the proposed methods' CD have better NMSE performance than the DFT dictionary. Among the proposed methods, the CDL-KSVD method exhibits the poorest NMSE performance. Therefore, we are not pursuing this method any further in the simulations.



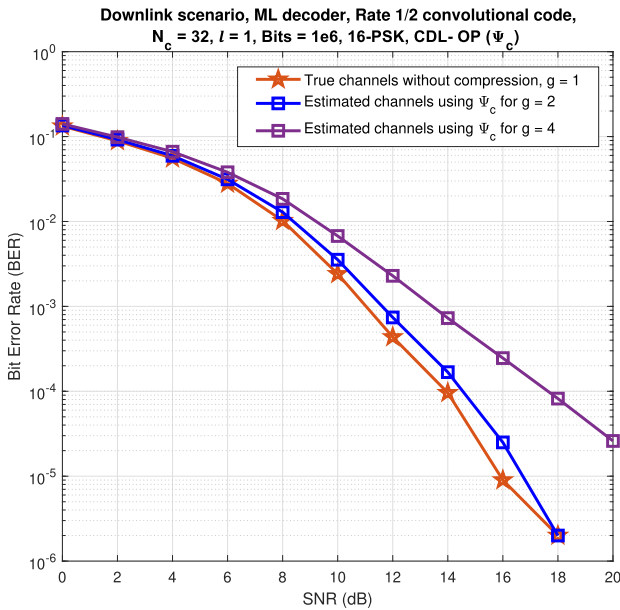
**FIGURE 12.** In a multi-UE system, the NMSE performance of the K-SVD dictionary and CDL-OP method's dictionary for  $K = 3$ ,  $N_t = 64$ ,  $N_r = 1$  and  $N_c = 32$ . Subcarrier 1 of  $UE_2$  and  $UE_3$  are considered for comparison. Velocities of three UEs are 10, 15, and 20 kmph.

In Fig. 12, the  $\Psi_c$  employed for NMSE characterization is learned from the CDL-OP method. Observe from the Fig. 12, for subcarrier  $l = 1$  of  $UE_2$  and  $UE_3$  at almost all the sparsity index values, the NMSE values for the proposed CDL-OP method's dictionary is better than subcarrier K-SVD based dictionaries, because the CDL-OP dictionary is learned from the estimated channels of K-SVD based dictionaries.

*Experiment 4:* In this single-UE experiment, we study the BER performance at the UE as a function of the SNR. The data symbols are transmitted from the BS to the UE in two ways: a) Using the true channel estimates without compression and b) Using the channel estimates with compression that are obtained from the CDL-OP dictionary.

To elaborate further on the performance of our proposed framework, we analyze the BER in a downlink scenario. In Fig. 13 we evaluate the BER using two sets of channels: the true uncompressed channels and the channels estimated using the CDL-OP dictionary on subcarrier 1. The precoder matrices employed at the BS are denoted as  $\mathbf{W}^g = \text{diag}\left(\frac{\mathbf{h}_1^H}{\|\mathbf{h}_1\|}\right)$

and  $\mathbf{W}_{op}^g = \text{diag}\left(\frac{\hat{\mathbf{h}}_{1,op}^H}{\|\hat{\mathbf{h}}_{1,op}\|}\right)$ , which correspond to the weights obtained from the true channels and to the channels estimated using CDL-OP dictionary, respectively. Let  $\mathbf{x}$  represents the modulated symbol vector. Explicitly, in our BER analysis, we harness a half-rate convolutional encoder having the generator sequences of  $G = [101, 111]$  and the resultant bits are 16-PSK modulated for generating the symbol vector  $\mathbf{x}$ . The received signal in the case of true uncompressed channels at the UE can be represented as  $y = \mathbf{h}_1^T \mathbf{W}^g \mathbf{x} + n$ , where  $n$  is the additive noise. Similarly the received signal for channels estimated using the CDL-OP dictionary can be represented as  $y_{op} = \mathbf{h}_1^T \mathbf{W}_{op}^g \mathbf{x} + n$ .



**FIGURE 13.** In a single-UE system with 64 TAs at the BS and a UE with 1 RA, we examine the BER performance of channel estimates using the CDL-OP methods’ dictionary. We consider two scenarios:  $N_g = N_r N_t / 2$  and  $N_g = N_r N_t / 4$ . For comparison, we assume uncompressed channel estimates for subcarrier 1. The UE has a velocity of 20 kmph.

The CDL-OP dictionaries used in this experiment are learned at a sparsity of 16 for  $g = 2$  and at a sparsity of 8 for  $g = 4$ . As we increase the compression factor  $g$  to reduce the CSI feedback from the UE, we observe a BER performance erosion. The lower bound shown in Fig. 13 represents the BER performance when the BS employs uncompressed CSI. Notably, as depicted in Fig. 13, the BER of the UE, recorded for  $g = 2$  when utilizing the CDL-OP dictionary approaches the lower bound.

### V. SUMMARY AND CONCLUSIONS

We proposed a novel CDL framework for reducing the FD-CHTF feedback and memory requirement of the UE and BS. The framework is more beneficial for the UE, which is usually resource-constrained, and the savings can be significant. For the simulations, the channels are generated using the QuaDRiGa simulator. In a multi-UE system of three UEs and in a single-UE system, all the proposed method’s dictionaries have better NMSE performance than the DFT dictionary. The CDL-OP dictionary performs better than the CDL-KSVD dictionary for all the sparsities. Hence the CDL-OP method can be beneficially employed for CDL to compress the CSI and improve the CSI reconstruction performance. In terms of the computational complexity, the CDL-OP method requires only a single SVD operation to learn the CD, while the CDL-KSVD method requires an SVD operation for learning each column in the CD. So the CDL-OP method has lower computational complexity than the CDL-KSVD method. To minimize the impact imposed on the BER performance, it is important to choose an appropriate compression ratio  $g$ . In our case, a compression ratio of  $g = 2$  is considered to

have a modest impact on the BER performance. By selecting a suitable compression ratio, the system can strike a balance between reducing the amount of feedback, while maintaining a satisfactory BER performance.

We conclude by highlighting the differences between the multi-UE and single-UE systems as follows:

- 1) In the multi-UE system, the memory is reduced by a factor of  $N_c$  at the UE and  $KN_c$  at the BS by having only a single CD instead of multiple subcarrier dictionaries. For  $N_t = 64$  and  $N_c = 32$ , the memory required for storing the CD at each UE is reduced by a factor of 32 and the CSI feedback is reduced by a factor of two.
- 2) In the single-UE system, the memory is reduced by a factor of  $N_c$  at both the UE and BS, the dictionary feedback is also reduced by a factor of  $N_c$  in the uplink. This is achieved by sending only a single CD instead of  $N_c$  subcarrier dictionaries to the BS. For  $N_t = 64$  and  $N_c = 32$ , the memory is reduced by a factor of 32 for storing only a CD, and the dictionary feedback is also reduced by a factor of 32. Finally, the CSI feedback is reduced by a factor of two.
- 3) In the multi-UE system, the CD ( $\Psi_c$ ) generated using the CDL-OP has a better NMSE performance to that of the individual subcarrier K-SVD dictionaries  $\Psi_{k_{svd,l}}^k \forall l \in [1, N_c], \forall k \in [1, K]$ .
- 4) In the single-UE system, the CD ( $\Psi_c$ ) generated using the CDL-OP method has a better NMSE performance to that of the individual subcarrier K-SVD dictionaries  $\Psi_{k_{svd,l}} \forall l \in [1, N_c]$ .

### REFERENCES

- [1] E. Björnson, J. Hoydis, and L. Sanguinetti, “Massive MIMO has unlimited capacity,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 574–590, Jan. 2018.
- [2] S. Yang and L. Hanzo, “Fifty years of MIMO detection: The road to large-scale MIMOs,” *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1941–1988, Fourthquarter 2015.
- [3] R. Shafin and L. Liu, “Multi-cell multi-user massive FD-MIMO: Downlink precoding and throughput analysis,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 487–502, Jan. 2019.
- [4] Z. Zhong, L. Fan, and S. Ge, “FDD massive MIMO uplink and downlink channel reciprocity properties: Full or partial reciprocity?,” in *Proc. IEEE Glob. Commun. Conf.*, 2020, pp. 1–5.
- [5] Z. Zhang, X. Wang, K. Long, A. V. Vasilakos, and L. Hanzo, “Large-scale MIMO-based wireless backhaul in 5G networks,” *IEEE Wireless Commun.*, vol. 22, no. 5, pp. 58–66, Oct. 2015.
- [6] P. K. Gadamsetty and K. V. S. Hari, “A fast dictionary learning algorithm for CSI feedback in massive MIMO FDD systems,” in *Proc. Nat. Conf. Commun.*, 2023, pp. 1–6.
- [7] S. S. Yilmaz and B. Özbek, “Compressive sensing based low complexity user selection for massive MIMO systems,” in *Proc. IEEE 91st Veh. Technol. Conf.*, 2020, pp. 1–5.
- [8] J. Wen, Z. Zhou, J. Wang, X. Tang, and Q. Mo, “A sharp condition for exact support recovery with orthogonal matching pursuit,” *IEEE Trans. Signal Process.*, vol. 65, no. 6, pp. 1370–1382, Mar. 2017.
- [9] X.-J. Liu, S.-T. Xia, and F.-W. Fu, “Reconstruction guarantee analysis of basis pursuit for binary measurement matrices in compressed sensing,” *IEEE Trans. Inf. Theory*, vol. 63, no. 5, pp. 2922–2932, May 2017.
- [10] H. Ge, L. Wang, J. Wen, and J. Xian, “An RIP condition for exact support recovery with covariance-assisted matching pursuit,” *IEEE Signal Process. Lett.*, vol. 26, no. 3, pp. 520–524, Mar. 2019.
- [11] L. Lu, G. Y. Li, D. Qiao, and W. Han, “Sparsity-enhancing basis for compressive sensing based channel feedback in massive MIMO systems,” in *Proc. IEEE Glob. Commun. Conf.*, 2015, pp. 1–6.

- [12] L. Hanzo et al., *OFDM and MC-CDMA for Broadband Multi-User Communications, WLANs and Broadcasting*. Hoboken, NJ, USA: Wiley, 2005.
- [13] P. Cheng and Z. Chen, "Multidimensional compressive sensing based analog CSI feedback for massive MIMO-OFDM systems," in *Proc. IEEE 80th Veh. Technol. Conf.*, 2014, pp. 1–6.
- [14] W. Zeng, Y. He, B. Li, and S. Wang, "Sparsity learning-based CSI feedback for FDD massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 585–588, Mar. 2021.
- [15] Z. Lv and Y. Li, "A channel state information feedback algorithm for massive MIMO systems," *IEEE Commun. Lett.*, vol. 20, no. 7, pp. 1461–1464, Jul. 2016.
- [16] S. Qiu, D. Gesbert, D. Chen, and T. Jiang, "A covariance-based hybrid channel feedback in FDD massive MIMO systems," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8365–8377, Dec. 2019.
- [17] H. Xie and N. González-Prelcic, "Dictionary learning for channel estimation in hybrid frequency-selective mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7407–7422, Nov. 2020.
- [18] Y. Zhao, Y. Teng, A. Liu, and V. Lau, "Two-timescale joint UL/DL dictionary learning and channel estimation in massive MIMO systems," in *Proc. IEEE Glob. Commun. Conf.*, 2022, pp. 5408–5413.
- [19] F. Kulsoom, A. Vizziello, H. N. Chaudhry, and P. Savazzi, "Joint sparse channel recovery with quantized feedback for multi-user massive MIMO systems," *IEEE Access*, vol. 8, pp. 11046–11060, 2020.
- [20] B. Lee, J. Choi, J.-Y. Seol, D. J. Love, and B. Shim, "Antenna grouping based feedback compression for FDD-based massive MIMO systems," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3261–3274, Sep. 2015.
- [21] W. Shen, L. Dai, B. Shim, Z. Wang, and R. W. Heath, "Channel feedback based on AoD-adaptive subspace codebook in FDD massive MIMO systems," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5235–5248, Nov. 2018.
- [22] G. Huang, A. Liu, and M.-J. Zhao, "Two-stage adaptive and compressed CSI feedback for FDD massive MIMO," *IEEE Trans. Veh. Technol.*, vol. 70, no. 9, pp. 9602–9606, Sep. 2021.
- [23] Z. Wei, H. Li, H. Liu, B. Li, and C. Zhao, "Randomized low-rank approximation based massive MIMO CSI compression," *IEEE Commun. Lett.*, vol. 25, no. 6, pp. 2004–2008, Jun. 2021.
- [24] Y. Liao, H. Yao, Y. Hua, and C. Li, "CSI feedback based on deep learning for massive MIMO systems," *IEEE Access*, vol. 7, pp. 86810–86820, 2019.
- [25] 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 Commun. Lett.*, vol. 8, no. 2, pp. 416–419, Apr. 2019.
- [26] H. Ye, F. Gao, J. Qian, H. Wang, and G. Y. Li, "Deep learning-based denoise network for CSI feedback in FDD massive MIMO systems," *IEEE Commun. Lett.*, vol. 24, no. 8, pp. 1742–1746, Aug. 2020.
- [27] P. Liang, J. Fan, W. Shen, Z. Qin, and G. Y. Li, "Deep learning and compressive sensing-based CSI feedback in FDD massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 9217–9222, Aug. 2020.
- [28] S. Tang, J. Xia, L. Fan, X. Lei, W. Xu, and A. Nallanathan, "Dilated convolution based CSI feedback compression for massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 71, no. 10, pp. 11216–11221, Oct. 2022.
- [29] Z. Hu, G. Liu, Q. Xie, J. Xue, D. Meng, and D. Gündüz, "A learnable optimization and regularization approach to massive MIMO CSI feedback," *IEEE Trans. Wireless Commun.*, early access, 2023, doi: [10.1109/TWC.2023.3275990](https://doi.org/10.1109/TWC.2023.3275990).
- [30] Z. Cao, J. Guo, C.-K. Wen, and S. Jin, "Deep-unfolding-based bit-level CSI feedback in massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 12, no. 2, pp. 371–375, Feb. 2023.
- [31] P.-H. Kuo, H. T. Kung, and P.-A. Ting, "Compressive sensing based channel feedback protocols for spatially-correlated massive antenna arrays," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2012, pp. 492–497.
- [32] Y. Barbotin, A. Hormati, S. Rangan, and M. Vetterli, "Estimation of sparse MIMO channels with common support," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3705–3716, Dec. 2012.
- [33] L. Yang, J. Fang, and H. Li, "Sparse bayesian dictionary learning with a Gaussian hierarchical model," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2016, pp. 2564–2568.
- [34] K. Zhang, E. L. Xu, H. Zhang, Z. Feng, and S. Cui, "Data driven automatic modulation classification via dictionary learning," *IEEE Wireless Commun. Lett.*, vol. 7, no. 4, pp. 586–589, Aug. 2018.
- [35] M. Nazzal, M. A. Aygül, A. Görçin, and H. Arslan, "Dictionary learning-based beamspace channel estimation in millimeter-wave massive MIMO systems with a lens antenna array," in *Proc. 15th Int. Wireless Commun. Mobile Comput. Conf.*, 2019, pp. 20–25.
- [36] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
- [37] J. C. Gower et al., *Procrustes Problems*, vol. 30. London, U.K.: Oxford Univ. Press, 2004.
- [38] 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.
- [39] M. Ettorre, S. C. Pavone, M. Casaletti, and M. Albani, "Experimental validation of bessell beam generation using an inward Hankel aperture distribution," *IEEE Trans. Antennas Propag.*, vol. 63, no. 6, pp. 2539–2544, Jun. 2015.
- [40] Z. F. et al., "Channel simulation and validation by quadriga for suburban microcells under 6 GHz," in *Proc. 12th Int. Symp. Antennas, Propag. EM Theory*, 2018, pp. 1–4.
- [41] F. Burkhardt, S. Jaeckel, E. Eberlein, and R. Prieto-Cerdeira, "Quadriga: A MIMO channel model for land mobile satellite," in *Proc. 8th Eur. Conf. Antennas Propag.*, 2014, pp. 1274–1278.



**PAVAN KUMAR GADAMSETTY** (Student Member, IEEE) received the B.E. degree in electronics and communication engineering from Andhra University, Visakhapatnam, India, in 2011, and the M.Tech. degree in electrical engineering from the Indian Institute of Technology (IIT), Kanpur, India, in 2014. He is currently working toward the Ph.D. degree in electrical communication engineering from the Indian Institute of Science, Bangalore, India. From 2014 to 2018, he was with Cisco Systems, Bangalore, India. His research interest includes channel estimation techniques, precoding techniques for massive MIMO systems, and emerging technologies in wireless communication.



**K. V. S. HARI** (Fellow, IEEE) received the Ph.D. degree in systems science from the University of California San Diego, La Jolla, CA, USA. He is currently a Professor with the Department of ECE, Indian Institute of Science (IISc), Bangalore, India. He was a Visiting Faculty with Stanford University, Stanford, CA, USA, and Affiliate Professor with the KTH-Royal Institute of Technology, Stockholm, Sweden. He is a co-author of the IEEE 802.16 standard on wireless channel models and has conducted drone-ground wireless channel modelling experiments. His research interests include signal processing and deep learning with applications to 5G wireless communications, indoor positioning, dual function radar and communication systems, autonomous navigation, neuroscience, and affordable MRI systems. He also leads the British Telecom India Research Centre (BTIRC), IISc. He was the Chair, Standardisation Committee, Telecom Standards Development Society, India. He was an Editor of EURASIP's Signal Processing and is currently the Editor-in-Chief (Electrical Sciences) of *Sadhana*, the *Journal of the Indian Academy of Sciences* published by Springer. He is the Fellow of the Indian National Academy of Engineering, Indian National Science Academy. He was on the Board of Governors, IEEE Signal Processing Society as VP-Membership.



**LAJOS HANZO** (Life Fellow, IEEE) received Doctorates degree (with Hons.) from the Technical University of Budapest, Budapest, Hungary, in 2009, and Edinburgh University, Edinburgh, U.K., in 2015, respectively. He was the recipient of the IEEE Eric Sumner Technical Field Award. He is also the Foreign Member of the Hungarian Science-Academy, Fellow of the Royal Academy of Engineering (FREng), IET, and EURASIP.