IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 16 November 2023, accepted 28 November 2023, date of publication 1 December 2023,
date of current version 14 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3338866

APPLIED RESEARCH

A Tutorial on Downlink Precoder Selection Strategies for 3GPP MIMO Codebooks

XIAOTIAN FU[1,2], (Member, IEEE), DIDIER LE RUYET[1], (Senior Member, IEEE), RAPHAEL VISOZ[3], (Senior Member, IEEE), VENKATESH RAMIREDDY[4], MARCUS GROSSMANN[4], MARKUS LANDMANN[4], AND WILMAR QUIROGA[3]

[1]CNAM CEDRIC/LAETITIA, 75141 Paris, France
[2]Ericsson R&D, 91300 Massy, France
[3]Orange Innovation, 92320 Chatillon, France
[4]Fraunhofer IIS, 98693 Ilmenau, Germany

Corresponding author: Raphael Visoz (raphael.visoz@orange.com)

ABSTRACT This paper deals with the evolution of downlink codebook based multiple-input multiple-output (MIMO) within the third generation partnership project (3GPP) through release 15 to 17. There exist already several tutorials on this topic. However, to the authors' knowledge, the comparison in terms of complexity performance trade-off between practical precoder selection strategies in frequency or delay domain has not been tackled so far in the literature. This paper describes with matrix formalism the two main codebook types specified within 3GPP, i.e., Type-I codebook (low resolution) and Type-II codebook (high resolution). The Rel. 17 port selection is also detailed as well as the multi-user MIMO (MU-MIMO) precoding strategy to be applied relying on Type-II codebook feedback. For the enhanced Type-II codebook, two main practical precoder selection strategies are detailed at the UE side (i) based on singular value decomposition per sub-band in the frequency domain, (ii) based on wideband singular decomposition in the delay domain. Monte Carlo simulations demonstrate that the delay domain strategy performance may suffer from spatial interference for single user MIMO high rank transmissions when the channel is both spatially correlated and frequency selective. On the other hand, the frequency domain strategy complexity increases linearly with the number of sub-bands while it is not the case for the delay domain selection strategy. As a result, the delay domain selection strategy is particularly relevant for Rel. 17 Type-II codebook port selection or for a low frequency selective channel with few significant consecutive delays.

INDEX TERMS Massive MIMO, codebook based MIMO, PHY abstractions, 5G, cross layer optimization, MIMO with limited feedback.

I. INTRODUCTION

The deployment of multiple antennas at both transmitter and receiver also known as multiple input multiple output (MIMO) technique, has been playing a significant role in LTE, LTE advanced and 5G new radio (NR) in terms of increasing the spectral efficiency of wireless communication networks. One of the important aspects of the MIMO techniques is to acquire the channel state information (CSI) at the transmitter as the precoder design requires the knowledge of the CSI at the transmitter. The so called precoder comprises a number of complex combining coefficients and maps the data of each spatial layer to the transmit antennas. It also aims at maximising the spectral efficiency of single user (SU)- or multi-user (MU)- MIMO by adapting the number of spatial layers to the rank of the channel and reducing the interference between the spatial layers. On the other hand, the CSI at receiver is easier to obtain by relying on reference signals known by the receiver.

In the literature there exist two main approaches for the CSI acquisition at the transmitter. The first one relies on channel reciprocity, where the channel between the transmitter and receiver is deduced from the channel between the receiver and the transmitter. This applies only to time division

The associate editor coordinating the review of this manuscript and approving it for publication was Stefan Schwarz.

duplex (TDD) systems. Although this has the advantage of providing high resolution CSI estimates, the provided CSI may be incomplete as the interference at the receiver cannot be estimated at the transmitter. The second one relies on the feedback from the receiver. The channel at the receiver is estimated based on the transmission of known reference signals from the transmitter. However due to the vast amount of uplink (UL) resources needed to feedback the full CSI, typically only limited CSI feedback is considered. Unlike the channel reciprocity approach, as the interference is known to the receiver, the precoder can be determined together with modulation and coding scheme (MCS) that maximizes the spectral efficiency. However, the optimal precoder feedback can still require vast amount of UL resources and must be consequently compressed. Such precoders and the compression techniques are being standardized in 3GPP in the form of Type-I and Type-II codebooks. The main goal of this paper is to give a tutorial on 3GPP based MIMO downlink codebooks and precoder selection strategies since Rel. 15.

In a typical downlink scenario, the base station (BS) transmits reference signals known to the user equipment (UE). The UE after performing channel estimation feeds back the compressed CSI to the BS. The feedback CSI can be explicit or implicit. For the explicit feedback, the channel matrix and/or covariance matrix is quantized and reported to the BS., whereas for the implicit feedback as in 3GPP codebooks, the CSI is composed of precoder matrix indicator (PMI), rank indicator (RI) and channel quality indicator (CQI) or index of the highest modulation and code scheme (MCS) [1], [2]. While explicit feedback can provide better scheduling flexibility compared to implicit feedback, only the implicit feedback has been considered in LTE, LTE advanced and 5G NR due to the amount of uplink resources required and backward compatibility [3], [4]. The core part of implicit CSI feedback is actually the precoder.

3GPP has standardized two codebooks, Type-I and Type-II, which define the CSI feedback for 5G New Radio (NR) DL MIMO transmission [5], [6]. Specifically, the UE first selects the precoder matrix following the standard based on the estimated DL channel, then reports the PMI, RI and CQI conditioned on the selected precoder matrix to the BS. Afterwards, the BS decodes the feedback and recovers the selected precoder matrix which can be used for the DL data transmission. Although both Type-I and Type-II codebooks are defined for 5G NR, the Type-I codebook inherits the LTE MIMO technology. Compared to the Type-I codebook, the Type-II codebook enables higher resolution of the CSI feedback but only supports the rank up to 2 [6] in Rel. 15. In addition, it suffers from an overhead which increases linearly proportional to the number of sub-bands. In Rel. 16, the enhanced Type-II codebook is introduced which exploits the correlation in the frequency domain and reduces the overhead by a factor of M/N_3, where M is the number of selected delays and N_3 is the number of subbands. In Rel. 17, partial reciprocity in both angle and delay domain is exploited to enhance the Rel. 16 port selection (PS) Type-II codebook. The Type-II regular and PS codebooks from Rel. 16 onwards support rank 4 transmission.

There exist already several tutorials on 3GPP codebooks [7]. However, to the authors's knowledge, the comparison in terms of complexity and performance between practical precoder selection strategies in either frequency or delay domain has not been tackled so far in the literature. It is important to emphasize that 3GPP NR standard only specifies the codebook and the report formats, the precoder selection as well as the link adaptation techniques are left to the UE implementation for the sake of competitive differentiation. Nonetheless, even if the techniques themselves are not specified, some minimum performances are required through UE conformance testing. This paper describes the two main codebook types with matrix formalism specified in the 3GPP, i.e., Type-I codebook (low resolution) and Type-II codebook (high resolution). The Rel. 17 port selection (PS) codebook is detailed in addition to the MU-MIMO precoding strategy to be applied relying on Type-II codebook. For the Type-II codebook, two main practical precoder selection strategies are described at the UE side (i) based on singular value decomposition per sub-band in the frequency domain, (ii) based on wideband singular decomposition in the delay domain.

The paper is organized as follows. In Section II, the state of the art precoder or beamforming methods and related work on 3GPP codebooks are summarized. In Section III, the system model is described. In Section IV, the Type-I codebook is explained and a practical precoder selection strategy is presented. In Section V, the Rel. 15 Type-II codebook is tackled together with a practical precoder selection strategy. In Section VI, the Rel. 16 enhanced Type-II codebook is described with two precoder selection alternatives one in the frequency domain and another in the delay domain. In Section VII, the link adaptation process is modeled with chase combining retransmissions. In Section VIII, performance of the different codebooks and selection strategies is compared. Finally, in Section VIII-A, MU-MIMO precoder design is explained while in Section IX, Rel. 17 Type-II PS codebook as well as the associated precoder selection strategy are described. Finally, Section X concludes the tutorial.

Notation: upper and lower boldface letters represent matrices and vectors. Let $\mathbf{A}$ be a matrix. $\mathbf{a}_j$, $\mathbf{a}^i$, and $a_{i,j}$ (or equivalently $[\mathbf{A}]_{i,j}$) denote the j-th column, the i-th row and the (i,j)-th element of $\mathbf{A}$. The superscripts $(\cdot)^\mathsf{T}$ and $(\cdot)^\mathsf{H}$ denote, respectively, transpose and Hermitian transpose operations. Moreover, $||\cdot||$ and $|\cdot|$ denote the vector and scalar norms. $\mathbf{a} \otimes \mathbf{b}$ denotes the Kronecker product of two vectors $\mathbf{a}$ and $\mathbf{b}$ and $eig(\mathbf{A})_l$ is the l-th eigenvector of the square matrix $\mathbf{A}$. The column vectorization of a matrix is represented by vec$(\cdot)$ and diag$(\mathbf{a})$ represents the generation of a diagonal matrix of the elements of the vector $\mathbf{a}$. The square identity matrix of dimension N is denoted by $\mathbf{I}_N$. The L-tuple $(a_i)_{i=0}^{L-1}$ stands for $(a_0, a_1, \cdots, a_{L-1})$. The null matrix of dimension $N \times L$ is denoted by $\mathbf{0}_{N\times L}$.

II. RELATED WORKS

A. BACKGROUND ON MIMO LIMITED FEEDBACK

The history of limited feedback in communication systems traces back to the early sixties [8], [9], [10]. The extension of limited feedback to multiple antenna wireless systems has recently received much attention [1]. One of the first proposed approach was to quantize the channel covariance matrix at the receiver and then feedback the statistical information to the transmitter [11], [12]. Different solutions for the quantization of the covariance matrix in order to maximize the average data rate have been proposed including vector quantization using the Lloyd algorithm [13], [14], [15] and random vector quantization [16].

Beamforming where the precoder is determined using rank one covariance matrix has been proposed in 1994 by Gerlach and Paulraj [17]. The early form of limited feedback is based on the quantization and reporting of either the channel or channel covariance information. One such method dealing with only per antenna phase quantization was proposed in [18]. Later, a different approach has been considered by many authors where the quantization and reporting of the beamforming vectors is considered instead of the channel or the channel covariance information. While the former approach is referred to as explicit feedback, the latter is referred to as implicit feedback. In this context of implicit feedback, it has been demonstrated that the quantization maximizing the rate must maximize the angular separation between the two closest line also known as the Grassmannian line packing problem [19], [20]. Codebook design based on Fourier concept has been introduced in [21] and [22] making the link between MIMO space-time code design problem and Grassmanian line packing.

While in beamforming only one data stream is sent spatially by projecting the data symbol onto a beamforming vector, linear precoding allows sending multiple data streams spatially [23]. For linear precoding, the concept of Grassmannian line packing beamforming codebooks has been extended to the case of M-dimensional subspace packing [24], [25]. Unitary rotation precoding with Givens rotations have been proposed in [26] and codebooks using Householder reflections have been used to generate a unitary matrix from a beamforming vector [27]. Several of these codebooks have been included in the 3GPP Rel. 8 and Rel. 10 [28].

In the previous works, an independent block-by-block fading channel is assumed. Different authors have considered temporally correlated channels that model the current channel realization dependent on the previous channel realizations. The exploitation of the temporal channel correlation allows to reduce the amount of feedback or improve the quality of the CSI at the transmitter. In [29], Banister and Zeidler have proposed a simple gradient sign algorithm for transmit antenna weight adaptation with feedback. In [30], the authors have proposed a switched codebook quantization where the codebook is dynamically chosen based on the channel distribution in order to exploit both spatial and temporal channel correlation. A beamforming codebook switching scheme, where the spherical cap codebook is selected from a super codebook containing various spherical cap codebooks with different centers and radii was considered in [31]. Progressive refinements of beamforming vectors using spherical cap codebook structure have also been studied in [32] and [33].

B. 3GPP CODEBOOKS FROM REL. 8 TO REL. 14

In order to meet growing demands of wireless data, Long Term Evolution (LTE) and LTE-Advanced (LTE-A) Releases 8 through 11 have introduced codebooks for SU-MIMO precoding. In the Rel. 8, two codebooks have been defined for the SU-MIMO precoding in downlink for two and four transmit antennas [34]. For two transmit antennas, a set of six vectors with dimension 2×1 and large chordal distance has been defined for rank 1 transmission. The codebooks for rank 2 with dimension 2×2 transmission are obtained by combining two vectors of the codebook for rank 1 transmission plus the 2×2 identity matrix. For four transmit antennas, the codebook has been designed in order to achieve a large chordal distance. The codebook has also a low computational complexity since it is obtained by selecting r columns among the sixteen 4×4 Householder matrices calculated from sixteen generating vectors. The elements of the generating vectors are taken from quaternary phase shift keying (QPSK) and 8-PSK constellations for reducing the storage requirement and limit the computational complexity. The codebook has a nested property since there exists at least one corresponding column subset in all the codebooks of the lower ranks.

In Rel. 10 (LTE-A), a dual-codebook approach has been introduced for a configuration with 8 transmit antennas in the downlink and two antenna setups have been considered: an uniformly linear array of eight uni-polarized antenna elements and an uniformly linear array of four dual-polarized antenna elements. The designed dual codebook is a double 4-bit codebook obtained via the multiplication of two precoding matrices $\mathbf{W} = \mathbf{W}_1\mathbf{W}_2$, where $\mathbf{W}_1$ targets long-term wideband channel information and $\mathbf{W}_2$ targets short term subband channel information [35].

In Rel. 13, the dual codebook structure was extended to encompass the elevation dimension with the so called full dimensional MIMO (FD-MIMO), i.e., some vertical antenna ports of a 2 Dimensional (2D) antenna array are dedicated to control the elevation dimension. The 2D antenna array is arranged as an uniform planar array (UPA).

Alternative CSI feedback paradigms have been proposed in [36]. In [36], rather than feeding back a set of CQI, RI and PMI, a reduced-spaced channel quantization and feedback scheme which exploits slow variation of the AoD distribution is proposed. Each UE is configured to feed back its associated quantized DL short-term channel.

C. 3GPP CODEBOOKS FROM REL. 15 TO REL. 17

In the first release of 5G, i.e., Rel. 15 [5], two NR MIMO codebooks have been introduced, namely, Type-I (low resolution CSI) and Type-II (high resolution CSI).

The 5G Type-I and Type-II codebooks inherit the dual stage precoding structure $\mathbf{W} = \mathbf{W}_1\mathbf{W}_2$ as well as the 2D antenna array of Rel. 13 FD-MIMO. Type-I codebook can be considered as a legacy of low resolution LTE MIMO codebooks. For both codebooks, $\mathbf{W}_1$ is based on IDFT beams. For a given layer, for the Type-I codebook, only one specific beam is selected from a group of beams, whereas for the Type-II codebooks, L beams are selected where $L \in \{2, 3, 4\}$ and then linearly combined [6]. In [37] the precoder selection strategies for Type-I and Type-II codebooks for rank 1 transmission are described, however, no explicit selection strategy for rank 2 or higher rank is presented. On the other hand, authors in [38] proposed a fully-connected neural network to find optimal Type-I precoders for MIMO transmission. Our previous work [39] focused on rank 2 precoder selecting strategies for the Type-II codebook which provides performance comparison of various selecting strategies. Authors of [40] proposed a selection strategy for the Type-II codebook without considering the critical phase pre-processing [41]. Reference [42] concentrates only on the joint wideband and SB amplitude quantization of the Type-II codebook.

Although the Type-II codebook provides substantial CSI feedback accuracy, the fact that it supports rank up to only 2 and has demanding feedback payload requirements motivates the conception of a further compressed CSI feedback. In papers [43], [44], and [45], frequency domain compression is proposed to apply on the top of the spatial domain compression in the Type-II codebook in order to obtain a compact and sparse representation in the beam and delay domain with a moderate complexity. Therefore, the enhanced Type-II codebook specified in Rel. 16 [46] supports frequency domain compression and only reports dominant compressed coefficients resulting in remarkably less feedback overhead than the Rel. 15 Type-II codebook. Moreover, the enhanced Type-II (eType-II) codebook allows higher quantization resolution for the reported amplitude and phase coefficients. Please note that the enhanced Type-II support rank values up to 4. In [47], an illustrative explanation of the eType-II codebook components and structure is presented. Though multi-stage amplitude quantization schemes are proposed in [44], [48], and [45], they were not accepted for the standardized enhanced Type-II codebook. Besides, authors in [45] elaborated the method to obtain the joint spatial and frequency domain precoder. Building on the Rel. 16 eType-II PS codebook, a further enhanced Type-II PS (FeType-II) codebook exploiting partial reciprocity in the delay and angle domain has been specified in Rel. 17. Table 1 recapitulates the MIMO codebook evolution with respect to 3GPP releases.

TABLE 1. Summary of MIMO codebook evolution with respect to 3GPP releases.

Release	3GPP MIMO codebook evolution
Rel. 8	1D antenna array Low resolution w. 4 antenna ports Up to 4 spatial layers Design based on chordal distance.
Rel. 10	1D antenna array Low resolution w. 8 antenna ports Up to 8 spatial layers Design based on $\mathbf{W}_1\mathbf{W}_2$
Rel. 13	2D antenna array (FD-MIMO) Low resolution w. 16 antenna ports Up to 8 spatial layers Design based on $\mathbf{W}_1\mathbf{W}_2$
Rel. 15	2D antenna array Type-I & Type-II w. 32 antenna ports Up to 8 spatial layers for Type-I Up to 2 spatial layers for Type-II Type-II PS based on angle reciprocity Design based on $\mathbf{W}_1\mathbf{W}_2$
Rel.16	eType-II introduced with FD compression eType-II extension to 4 spatial layers eType-II PS based on angle reciprocity
Rel. 17	FeType-II PS based on angle and partial delay reciprocity

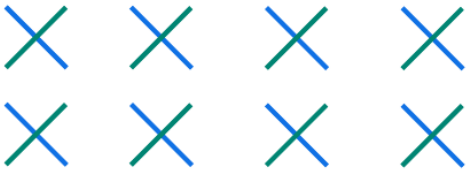


FIGURE 1. Illustration of the antenna port panel with $N_1 = 4$ and $N_2 = 2$.

III. SYSTEM MODEL

For 5G NR MIMO transmission, the BS is equipped with a dual polarized planar antenna with $P_{\text{TXRU}}/2$ Transceiver units (TXRUs) per polarization. The BS virtualizes the P_{TXRU} TXRUs into $P_{\text{CSI-RS}}$ antenna ports using a specification transparent precoding denoted $\mathbf{F}_{gNB}$ in the following. After the virtualization, the UE equipped with N_R receive antennas measure the channel from the Channel State Information Reference Signals (CSI-RSs) originating from a 2D antenna array comprising $P_{\text{CSI-RS}} = 2N_1N_2$ ports, where N_1 and N_2 are number of antenna ports in the horizontal and vertical dimensions respectively illustrated in Fig. 1. The received signal per resource element (RE), i.e., per subcarrier and per OFDM symbol, $\mathbf{y} \in \mathbb{C}^{N_R \times 1}$ at the UE side, after the spatial whitening of the noise-plus-interference can be written as

$$\mathbf{y} = \mathbf{HPx} + \mathbf{n} \tag{1}$$

where $\mathbf{H} \in \mathbb{C}^{N_R \times 2N_1N_2}$ is the DL channel matrix, $\mathbf{P} \in \mathbb{C}^{2N_1N_2 \times \nu}$ is the precoder matrix and $\mathbf{x} \in \mathbb{C}^{\nu}$ is the transmit signal vector with $\nu \geqslant 1$ being the number of spatial layers. For codebook based MIMO, the BS derives the transmit precoding matrix $\mathbf{P}$ based on the CSI report sent by the UE according to a pre-defined codebook. A typical CSI report consists of the precoder matrix $\mathbf{W} \in \mathbb{C}^{2N_1N_2 \times \nu}$, the RI value ν and the CQI calculated by the UE. In the following, it is assumed that $\mathbf{P} = \mathbf{W}$ for the SU transmission scenario.

The transmit vector $\mathbf{x}$ is normalized such that $\mathbb{E}\{\mathbf{xx}^{\mathrm{H}}\} = \mathbf{I}_{\nu}$. The precoding matrix $\mathbf{W}$ is normalized such that the transmit power P_T is equal to one, i.e., $P_T = \mathbb{E}\{\mathbf{x}^{\mathbf{H}}\mathbf{W}^{\mathrm{H}}\mathbf{Wx}\} = 1$. The averaged receive power per antenna port P_R assuming

$\mathbf{P} = \mathbf{I}_{2N_1N_2}$ is

$$P_R = \frac{1}{2N_1N_2N_R}\sum_{r=1}^{N_R}\sum_{t=1}^{2N_1N_2}\mathbb{E}\left\{|h_{r,t}|^2\right\} \quad (2)$$

$$= \sigma_h^2. \quad (3)$$

We normalize the channel such that $\sigma_h^2 = 1$. As a result, the average SNR per antenna port is the inverse of the noise power, i.e., $snr = 1/N_0$

In this work we assume a Linear Minimum Mean Square Error (LMMSE) receiver whose output is given by

$$\hat{\mathbf{x}} = \mathbf{F}^{\mathsf{H}}\mathbf{y} \quad (4)$$

where $\mathbf{F}^{\mathsf{H}}$ is the LMMSE equalizer filter and can be expressed as

$$\mathbf{F}^{\mathsf{H}} = \mathbf{W}^{\mathsf{H}}\mathbf{H}^{\mathsf{H}}\left[\mathbf{H}\mathbf{W}\mathbf{W}^{\mathsf{H}}\mathbf{H}^{\mathsf{H}} + \tfrac{1}{snr}\mathbf{I}\right]^{-1} \quad (5)$$

$$= \mathbf{W}^{\mathsf{H}}\mathbf{H}^{\mathsf{H}}\mathbf{R}^{-1} \quad (6)$$

where $\mathbf{R} = \mathbf{H}\mathbf{W}\mathbf{W}^{\mathsf{H}}\mathbf{H}^{\mathsf{H}} + \frac{1}{snr}\mathbf{I}$. The received signal then becomes:

$$\hat{\mathbf{x}} = \mathbf{W}^{\mathsf{H}}\mathbf{H}^{\mathsf{H}}\mathbf{R}^{-1}\mathbf{H}\mathbf{W}\mathbf{x} + \mathbf{W}^{\mathsf{H}}\mathbf{H}^{\mathsf{H}}\mathbf{R}^{-1}\mathbf{n} \quad (7)$$

The Signal to Interference plus Noise Ratio (SINR) measured at the output of the LMMSE receiver for layer l is expressed as follows [49]:

$$\gamma_l(\mathbf{H}, \mathbf{w}_l) = \frac{(\mathbf{H}\mathbf{w}_l)^{\mathsf{H}}\mathbf{R}^{-1}(\mathbf{H}\mathbf{w}_l)}{1 - (\mathbf{H}\mathbf{w}_l)^{\mathsf{H}}\mathbf{R}^{-1}(\mathbf{H}\mathbf{w}_l)} \quad (8)$$

where $\mathbf{w}_l$ is the l-th column of $\mathbf{W}$. Equivalently the above equation as function of snr is expressed as:

$$\gamma_l(\mathbf{H}, \mathbf{w}_l) = \frac{snr}{\left(\left[\mathbf{W}^{\mathsf{H}}\mathbf{H}^{\mathsf{H}}\mathbf{H}\mathbf{W} + \frac{1}{snr}\mathbf{I}\right]^{-1}\right)_{l,l}} - 1. \quad (9)$$

In the standard, the notion of subband (SB) is introduced where the allocated bandwidth is divided into several SBs and a SB comprises up to N_{SB} PRBs. As the channel variation is not significant in the frequency domain, the precoder can remain constant for N_{SB} PRBs of a SB for a given OFDM symbol. As a result, we resort to a "representative" channel per SB to derive the precoder per SB, denoted $\mathbf{H}(t)$. It can be the averaged channel over the SB or the channel value at the middle of the SB. For an SU-MIMO transmission, a modulated codeword (made of one or several LDPC Code blocks) is demutiplexed into ν spatial layers, each spatial layer is transmitted over N_3 SBs and several consecutive OFDM symbols which define the transmission time interval (TTI). In the following, we assume that the channel does not change within one TTI and is constant within a PRB. As a result, a codeword is spread over $N_{SB} \times \nu$ SINRs per SB t and per TTI. The SINR of PRB $j = 0, \cdots, N_{SB} - 1$ of SB $t = 0, \cdots, N_3 - 1$, and layer $l = 1, \cdots, \nu$ is denoted $\gamma_{l,j}^{\nu,t}$. For the sake of readability, we summarize the main mathematical symbols and their definitions introduced in this Section in Table 2.

TABLE 2. System model main mathematical symbols and their definitions.

Definition	Symbol
Number of horizontal co-polar antenna ports	N_1
Number of vertical co-polar antenna ports	N_2
Number of TXRU antenna ports at the BS	P_{TXRU}
Number of CSI-RS antenna ports	$P_{\text{CSI-RS}}$
Number of receive antenna ports	N_R
Number of SBs	N_3
Number of PRBs per SB	N_{SB}
Number of spatial layers	ν
LMMSE SINR for PRB j of SB t and layer $l \leq \nu$	$\gamma_{l,j}^{\nu,t}$
Precoding matrix applied to the CSI-RS	$\mathbf{F}_{gNB}$
Precoding matrix applied on the CSI-RS antenna ports	$\mathbf{P}$
Channel matrix representative of SB t (obtained from the CSI-RS)	$\mathbf{H}(t)$
Precoding matrix indicated in the CSI report	$\mathbf{W}$

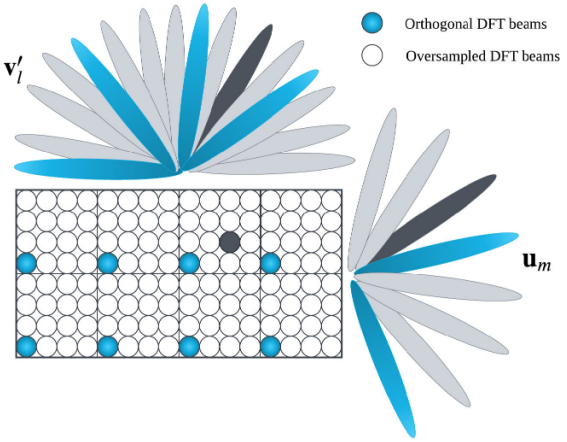


FIGURE 2. Illustration of the grid of beams.

IV. TYPE-I CODEBOOK

A. DESCRIPTION

In this Section, we only introduce Type-I single panel codebook up to rank 4 as defined in 3GPP Rel. 15. We also restrict our analysis to codebook mode 1 for rank 1 and rank 2 transmission, i.e., the higher layer parameter *codebookMode* is set to 1. To balance the trade-off between the CSI accuracy and the feedback overhead, both Type-I and Type-II codebooks consider IDFT based grid of beams in their precoder structures. Every IDFT beam $\mathbf{v}_{l,m} \in \mathbb{C}^{N_1N_2\times 1}$ in the grid of beams is built from the Kronecker product of a horizontal beam $\mathbf{v}'_l$ and a vertical beam $\mathbf{u}_m$, formulated as follows

$$\mathbf{v}_{l,m} = \mathbf{v}'_l \otimes \mathbf{u}_m \quad (10)$$

$$\mathbf{v}'_l = \left[1 \; e^{\frac{j2\pi l}{O_1N_1}} \; \cdots \; e^{\frac{j2\pi l(N_1-1)}{O_1N_1}}\right]^{\mathsf{T}} \quad (11)$$

$$\mathbf{u}_m = \begin{cases} \left[1 \; e^{\frac{j2\pi m}{O_2N_2}} \; \cdots \; e^{\frac{j2\pi m(N_2-1)}{O_2N_2}}\right]^{\mathsf{T}} & N_2 > 1 \\ 1 & N_2 = 1 \end{cases} \quad (12)$$

where O_1 and O_2 are the oversampling factors of the horizontal and vertical dimension, respectively, and

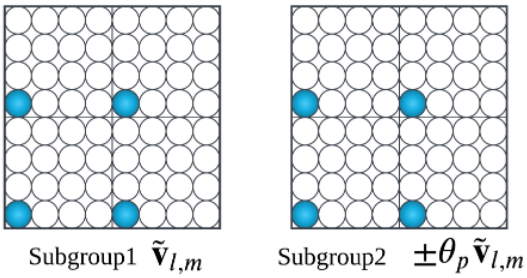


FIGURE 3. Illustration of the partitioning of the antenna ports into 2 subgroups for $N_1 = 4$, $N_2 = 2$, $O_1 = 4$ and $O_2 = 4$.

$l \in \{0, 1, \cdots, N_1O_1 - 1\}$ and $m \in \{0, 1, \cdots, N_2O_2 - 1\}$. Fig. 1. shows the grid of beams for $N_1 = 4$, $N_2 = 2$, $O_1 = 4$ and $O_2 = 4$. It can be seen that each beam in the grid points to a certain 3D spatial region. For instance, the black IDFT beam in the grid is the Kronecker product of $\mathbf{v}'_{10}$ and $\mathbf{u}_5$.

For rank 1, the Type-I precoder for the t-th SB ($t = 0, \ldots N_3 - 1$) is given by

$$\begin{aligned}\mathbf{W}_t^1 &= \frac{1}{\sqrt{P_{\text{CSI-RS}}}}\begin{bmatrix}\mathbf{v}_{l,m}\\ \varphi_{n_t}\mathbf{v}_{l,m}\end{bmatrix}\\ &= \frac{1}{\sqrt{P_{\text{CSI-RS}}}}\begin{bmatrix}\mathbf{v}_{l,m} & \mathbf{0}_{N_1N_2\times 1}\\ \mathbf{0}_{N_1N_2\times 1} & \mathbf{v}_{l,m}\end{bmatrix}\begin{bmatrix}1\\ \varphi_{n_t}\end{bmatrix}\\ &= \frac{1}{\sqrt{P_{\text{CSI-RS}}}}\mathbf{W}_1\mathbf{w}_2. \end{aligned} \tag{13}$$

The Type-I codebook for rank 1 transmission can be written as the product of a wideband beam matrix $\mathbf{W}_1 \in \mathbb{C}^{2N_1N_2\times 2}$ and a subband specific co-phasing vector $\mathbf{w}_2 \in \mathbb{C}^2$. More specifically, $\mathbf{W}_1$ is a block diagonal matrix comprising the IDFT vector $\mathbf{v}_{l,m}$ on the main diagonal. It remains unchanged across the N_3 subbands hence the name wideband. The IDFT vector $\mathbf{v}_{l,m}$ represents a beam selected from the grid of beams. The co-phasing vector $\mathbf{w}_2$ is specified for each t-th subband whose phase entry φ_{n_t} is selected from the set $\{1, j, -1, -j\}$. Similarly, the Type-I codebook for rank 2 transmission can be expressed as

$$\begin{aligned}\mathbf{W}_t^2 &= \frac{1}{\sqrt{2P_{\text{CSI-RS}}}}\begin{bmatrix}\mathbf{v}_{l,m} & \mathbf{v}_{l',m'}\\ \varphi_{n_t}\mathbf{v}_{l,m} & -\varphi_{n_t}\mathbf{v}_{l',m'}\end{bmatrix}\\ &= \frac{1}{\sqrt{2P_{\text{CSI-RS}}}}\mathbf{W}_1\mathbf{W}_2 \end{aligned} \tag{14}$$

where

$$\mathbf{W}_1 = \begin{bmatrix}\mathbf{v}_{l,m} & \mathbf{v}_{l',m'} & \mathbf{0}_{N_1N_2\times 1} & \mathbf{0}_{N_1N_2\times 1}\\ \mathbf{0}_{N_1N_2\times 1} & \mathbf{0}_{N_1N_2\times 1} & \mathbf{v}_{l,m} & \mathbf{v}_{l',m'}\end{bmatrix} \tag{15}$$

and

$$\mathbf{W}_2 = \begin{bmatrix}1 & 0\\ 0 & 1\\ \varphi_{n_t} & 0\\ 0 & -\varphi_{n_t}\end{bmatrix},$$

with φ_{n_t} being the co-phasing value for the t-th SB selected from the set $\{1, j\}$. Note that $\mathbf{v}_{l,m}$ and $\mathbf{v}_{l',m'}$ are (possibly) two distinct beams and are chosen from the grid of beams. Rules of the subscripts of these two beams are given in Table 5.2.2.2.1-3 of Release-15 [5].

For rank 3 and rank 4 transmission, the Type-I codebook structure depends on the number of antenna ports $P_{\text{CSI-RS}}$. When $P_{\text{CSI-RS}} < 16$, the Type-I precoder of rank 3 is given by

$$\begin{aligned}\mathbf{W}_t^3 &= \frac{1}{\sqrt{3P_{\text{CSI-RS}}}}\begin{bmatrix}\mathbf{v}_{l,m} & \mathbf{v}_{l',m'} & \mathbf{v}_{l,m}\\ \varphi_{n_t}\mathbf{v}_{l,m} & \varphi_{n_t}\mathbf{v}_{l',m'} & -\varphi_{n_t}\mathbf{v}_{l,m}\end{bmatrix}\\ &= \frac{1}{\sqrt{3P_{\text{CSI-RS}}}}\mathbf{W}_1\mathbf{W}_2 \end{aligned} \tag{16}$$

where $\mathbf{W}_1$ is given in Eq. (15) and

$$\mathbf{W}_2 = \begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ \varphi_{n_t} & 0 & -\varphi_{n_t}\\ 0 & \varphi_{n_t} & 0\end{bmatrix},$$

while for rank 4 it is given by

$$\begin{aligned}\mathbf{W}_t^4 &= \frac{1}{\sqrt{4P_{\text{CSI-RS}}}}\begin{bmatrix}\mathbf{v}_{l,m} & \mathbf{v}_{l',m'} & \mathbf{v}_{l,m} & \mathbf{v}_{l',m'}\\ \varphi_{n_t}\mathbf{v}_{l,m} & \varphi_{n_t}\mathbf{v}_{l',m'} & -\varphi_{n_t}\mathbf{v}_{l,m} & -\varphi_{n_t}\mathbf{v}_{l',m'}\end{bmatrix}\\ &= \frac{1}{\sqrt{4P_{\text{CSI-RS}}}}\mathbf{W}_1\mathbf{W}_2 \end{aligned} \tag{17}$$

where $\mathbf{W}_1$ is given in Eq. (15) and

$$\mathbf{W}_2 = \begin{bmatrix}1 & 0 & 1 & 0\\ 0 & 1 & 0 & 1\\ \varphi_{n_t} & 0 & -\varphi_{n_t} & 0\\ 0 & \varphi_{n_t} & 0 & -\varphi_{n_t}\end{bmatrix},$$

with φ_{n_t} being selected for the t-th SB from the set $\{1, j\}$. For $P_{\text{CSI-RS}} < 16$, the rank 3 and 4 precoders can be seen as an extension of the Type-I rank 2 codebook.

When $P_{\text{CSI-RS}} \geqslant 16$, the Type-I precoder of rank 3 is given by

$$\begin{aligned}\mathbf{W}_t^3 &= \frac{1}{\sqrt{3P_{\text{CSI-RS}}}}\begin{bmatrix}\tilde{\mathbf{v}}_{l,m} & \tilde{\mathbf{v}}_{l,m} & \tilde{\mathbf{v}}_{l,m}\\ \theta_p\tilde{\mathbf{v}}_{l,m} & -\theta_p\tilde{\mathbf{v}}_{l,m} & \theta_p\tilde{\mathbf{v}}_{l,m}\\ \varphi_{n_t}\tilde{\mathbf{v}}_{l,m} & \varphi_{n_t}\tilde{\mathbf{v}}_{l,m} & -\varphi_{n_t}\tilde{\mathbf{v}}_{l,m}\\ \varphi_{n_t}\theta_p\tilde{\mathbf{v}}_{l,m} & -\varphi_{n_t}\theta_p\tilde{\mathbf{v}}_{l,m} & -\varphi_{n_t}\theta_p\tilde{\mathbf{v}}_{l,m}\end{bmatrix}\\ &= \frac{1}{\sqrt{3P_{\text{CSI-RS}}}}\mathbf{W}_1\mathbf{W}_2 \end{aligned} \tag{18}$$

where

$$\mathbf{W}_1 = \begin{bmatrix}\tilde{\mathbf{v}}_{l,m} & \tilde{\mathbf{v}}_{l,m} & \mathbf{0}_{\frac{N_1N_2}{2}\times 1} & \mathbf{0}_{\frac{N_1N_2}{2}\times 1}\\ \theta_p\tilde{\mathbf{v}}_{l,m} & -\theta_p\tilde{\mathbf{v}}_{l,m} & \mathbf{0}_{\frac{N_1N_2}{2}\times 1} & \mathbf{0}_{\frac{N_1N_2}{2}\times 1}\\ \mathbf{0}_{\frac{N_1N_2}{2}\times 1} & \mathbf{0}_{\frac{N_1N_2}{2}\times 1} & \tilde{\mathbf{v}}_{l,m} & \tilde{\mathbf{v}}_{l,m}\\ \mathbf{0}_{\frac{N_1N_2}{2}\times 1} & \mathbf{0}_{\frac{N_1N_2}{2}\times 1} & \theta_p\tilde{\mathbf{v}}_{l,m} & -\theta_p\tilde{\mathbf{v}}_{l,m}\end{bmatrix} \tag{19}$$

and

$$\mathbf{W}_2 = \begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ \varphi_{n_t} & 0 & -\varphi_{n_t}\\ 0 & \varphi_{n_t} & 0\end{bmatrix},$$

while for rank 4 it is given by

$$\mathbf{W}_t^4 = \frac{1}{\sqrt{4P_{\text{CSI-RS}}}}\mathbf{W}_1\mathbf{W}_2 \tag{20}$$

where $\mathbf{W}_1$ is given in Eq. (19) and

$$\mathbf{W}_2 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ \varphi_{n_t} & 0 & -\varphi_{n_t} & 0 \\ 0 & \varphi_{n_t} & 0 & -\varphi_{n_t} \end{bmatrix},$$

with θ_p being selected from the set $\{1, e^{\frac{j\pi}{4}}, j, e^{\frac{j3\pi}{4}}\}$ and φ_{n_t} from the set $\{1, j\}$. Note that the vector $\tilde{\mathbf{v}}_{l,m} \in \mathbb{C}^{N_1N_2/2\times 1}$ is still an IDFT beam calculated as

$$\tilde{\mathbf{v}}_{l,m} = \tilde{\mathbf{v}}'_l \otimes \mathbf{u}_m \tag{21}$$

where vector $\mathbf{u}_m$ is generated according to Eq. (12) and vector $\tilde{\mathbf{v}}'_l \in \mathbb{C}^{N_1/2\times 1}$ is computed as

$$\tilde{\mathbf{v}}'_l = \begin{bmatrix} 1 & e^{\frac{j4\pi l}{O_1N_1}} & \cdots & e^{\frac{j4\pi l(N_1/2-1)}{O_1N_1}} \end{bmatrix}^{\mathsf{T}} \tag{22}$$

where $l \in \{0, 1, \cdots, \frac{N_1O_1}{2} - 1\}$. The beam $\tilde{\mathbf{v}}_{l,m}$ is half of the length of the beam $\mathbf{v}_{l,m}$. For $N_1 = 4$, $N_2 = 2$, $O_1 = 4$ and $O_2 = 4$, Fig. 3 depicts the partitioning of the antenna ports into 2 subgroups of equal size in the horizontal domain which correspond to the precoding vectors $\tilde{\mathbf{v}}_{l,m}$ and $\pm\theta_p\tilde{\mathbf{v}}_{l,m}$, respectively. The aforementioned beam arrangements is used in Eqs. (18) and (20) as part of beam matrix $\mathbf{W}_1$.

B. PRECODER SELECTION STRATEGY

The Type-I Codebook precoder selection strategy is comprised of two steps. In the first step, the wideband part of the precoder is selected. In the second step, the co-phasing vectors are determined for each SB. We describe the selection strategy for the rank 1 precoder given in Eq. (13) since it is similar for higher ranks. In the first step, the best IDFT beam is calculated using the wideband covariance matrix associated with each polarization. Here, we assume that the best IDFT beam is identical for both polarizations. the optimal beam $\mathbf{v}_{l^*,m^*}$ is selected as follows:

$$\mathbf{v}_{l^*,m^*} = \underset{\mathbf{v}_{l,m}\in\mathcal{B}}{\arg\max}\, \mathbf{v}_{l,m}^{\mathsf{H}}(\mathbf{R}_{11} + \mathbf{R}_{22})\mathbf{v}_{l,m}, \tag{23}$$

where $\mathcal{B}$ is the IDFT grid of beams, $\mathbf{R}_{11}$ and $\mathbf{R}_{22}$ are diagonal blocks of the wideband covariance matrix $\mathbf{R}$ given by

$$\begin{aligned}\mathbf{R} &= \frac{1}{N_3}\sum_{t=0}^{N_3-1}\mathbf{H}^{\mathsf{H}}(t)\mathbf{H}(t) \\ &= \begin{bmatrix}\mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{12}^{\mathsf{H}} & \mathbf{R}_{22}\end{bmatrix}.\end{aligned} \tag{24}$$

In the second step, the optimal co-phasing vector φ_{n_t} for each t-th SB is selected. We define the average mutual information for a given modulation Q over SB t as

$$\mathcal{I}_{Q,t}(\varphi_{n_t}) = \frac{1}{N_{SB}}\sum_{j=0}^{N_{SB}-1} I_Q\left(\gamma_{1,j}^{1,t}\right) \tag{25}$$

where the function $I_Q(\cdot)$ maps the SINR to the mutual information conditional on modulation Q and $\gamma_{1,j}^{1,t}$ is the SINR of PRB j of SB t. The cophasing phase $\varphi_{n_t} = e^{\frac{j\pi n_t}{2}}$ for SB t is obtained from

$$n_t = \underset{n\in\{0,1,2,3\}}{\arg\max}\, \mathcal{I}_{Q,t}(e^{\frac{j\pi n}{2}}). \tag{26}$$

V. TYPE-II CODEBOOK

A. DESCRIPTION

The Type-II codebook standardized in 3GPP Rel. 15 supports up to rank 2 ($\nu \leqslant 2$). The Type-II precoders for the t-th SB for rank 1 and rank 2 transmission are defined as follows:

$$\mathbf{W}_t^1 = \mathbf{w}_t^{(1)}, \tag{27}$$

$$\mathbf{W}_t^2 = \frac{1}{\sqrt{2}}\begin{bmatrix}\mathbf{w}_t^{(1)} & \mathbf{w}_t^{(2)}\end{bmatrix} \tag{28}$$

with

$$\mathbf{w}_t^{(l)} = \Lambda_t^{(l)}\begin{bmatrix}\sum_{i=0}^{L-1}\mathbf{v}_{m_1^{(i)}m_2^{(i)}}p_{l,i}^{(1)}p_{l,i,t}^{(2)}\varphi_{l,i,t} \\ \sum_{i=0}^{L-1}\mathbf{v}_{m_1^{(i)}m_2^{(i)}}p_{l,i+L}^{(1)}p_{l,i+L,t}^{(2)}\varphi_{l,i+L,t}\end{bmatrix} \tag{29}$$

where $L \in \{2, 3, 4\}$ is the number of IDFT beams and $\mathbf{v}_{m_1^{(i)}m_2^{(i)}}, i \in \{0, \ldots, L-1\}$ are the IDFT beams selected from the grid of beams introduced in the beginning of Section IV-A, and $\Lambda_t^{(l)}$ is the normalization factor given as

$$\Lambda_t^{(l)} = \frac{1}{\sqrt{N_1N_2\sum_{i=0}^{2L-1}\left(p_{1,i}^{(1)}p_{1,i,t}^{(2)}\right)^2}}. \tag{30}$$

The Type-II codebook is usually referred to as a high resolution codebook since for each layer the precoder combines $L \geq 2$ IDFT beams (contrary to the low resolution Type-I codebook where a single IDFT beam is supported per layer). The Type-II precoder for a given spatial layer and SB can be expressed in the form $\mathbf{W}_1\mathbf{w}_2$ where

$$\mathbf{W}_1 = \begin{bmatrix}\mathbf{v}_{m_1^{(0)}m_2^{(0)}} & \cdots & \mathbf{v}_{m_1^{(L-1)}m_2^{(L-1)}} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{v}_{m_1^{(0)}m_2^{(0)}} & \cdots & \mathbf{v}_{m_1^{(L-1)}m_2^{(L-1)}}\end{bmatrix} \tag{31}$$

of dimension $2N_1N_2 \times 2L$ defines the L selected wideband IDFT beams $\mathbf{v}_{m_1^{(i)}m_2^{(i)}}$ for $i \in \{0, \cdots, L-1\}$ of dimension N_1N_2 for both polarizations and $\mathbf{w}_2$ is a vector of Beam Combining Coefficients (BCCs) defined per SB and spatial layer. More specifically, for a given SB t and layer l, the precoder can be expressed as

$$\mathbf{w}_t^{(l)} = \Lambda_t^{(l)}\mathbf{W}_1\mathbf{W}_{C1}^{(l)}\mathbf{w}_{C2,t}^{(l)} \tag{32}$$

where

$$\mathbf{W}_{C1}^{(l)} = \text{diag}\left(\begin{bmatrix}p_{l,0}^{(1)} \cdots p_{l,2L-1}^{(1)}\end{bmatrix}\right)$$

and

$$\mathbf{w}_{C2,t}^{(l)} = \begin{bmatrix} p_{l,0,t}^{(2)}\varphi_{l,0,t} \\ \vdots \\ p_{l,L-1,t}^{(2)}\varphi_{l,L-1,t} \\ p_{l,L,t}^{(2)}\varphi_{l,L,t} \\ \vdots \\ p_{l,2L-1,t}^{(2)}\varphi_{l,2L-1,t} \end{bmatrix},$$

with $\mathbf{W}_{C1}^{(1)} \in \mathbb{C}^{2L\times 2L}$ the diagonal matrix comprising the quantized wideband amplitudes, $\mathbf{w}_{C2,t}^{(l)} \in \mathbb{C}^{2L}$ the vector of quantized phases and amplitudes of SB t and, finally, $\Lambda_t^{(l)}$ the normalization factor.

The L IDFT beams of $\mathbf{W}_1$ are defined by the indices $m_1^{(i)}$ and $m_2^{(i)}$ ($i \in \{1, \cdots, L-1\}$) whose values are

$$m_1^{(i)} = O_1 n_1^{(i)} + q_1, \tag{33}$$
$$m_2^{(i)} = O_2 n_2^{(i)} + q_2, \tag{34}$$

where $q_1 \in \{0, 1, \cdots, O_1 - 1\}$ and $q_2 \in \{0, 1, \cdots, O_2 - 1\}$ indicate the horizontal and vertical beam rotation indices, respectively. q_1 and q_2 are shared by the L selected IDFT beams which, thus, are orthogonal between each other. The entries $p_{l,i}^{(1)}, \forall i \in \{0, \cdots, 2L-1\}$ of the wideband amplitude matrix $\mathbf{W}_{C1}^{(1)}$ are drawn from the set $\mathcal{P}_1 = \left\{0, \sqrt{1/64}, \sqrt{1/32}, \sqrt{1/16}, \sqrt{1/8}, \sqrt{1/4}, \sqrt{1/2}, 1\right\}$. The SB amplitude values $p_{l,i,t}^{(2)}$ are all set to 1 if the higher layer parameter *subbandAmplitude* is set to 'false' otherwise they are drawn from the set $\mathcal{P}_2 = \left\{\sqrt{1/2}, 1\right\}$. The SB phase value set is $\Phi = \{e^{\frac{j2\pi n}{N_{PSK}}}, n \in \{0, 1, \cdots, N_{PSK} - 1\}\}$ where N_{PSK} is set by the higher layer parameter *phaseAlphabetSize* and can be either 4 or 8. Note that if the wideband amplitude value of a beam is equal to zero, all the SB coefficients associated with that beam are not reported in the CSI report.

The concatenated Type-II precoder across the N_3 SBs for layer l is of size $2N_1N_2 \times N_3$ and can be formulated as

$$\begin{aligned}\mathbf{W}^{(l)} &= \left[\mathbf{w}_0^l, \cdots, \mathbf{w}_{N_3-1}^l\right] \\ &= \mathbf{W}_1\mathbf{W}_{C1}^{(l)} \underbrace{\left[\mathbf{w}_{C2,0}^{(l)} \cdots \mathbf{w}_{C2,N_3-1}^{(l)}\right]}_{\mathbf{W}_{C2}^{(l)}:\, 2L\times N_3} \underbrace{\mathbf{\Lambda}^{(l)}}_{N_3\times N_3} \\ &= \mathbf{W}_1\mathbf{W}_{C1}^{(l)}\mathbf{W}_{C2}^{(l)}\mathbf{\Lambda}^{(l)}\end{aligned} \tag{35}$$

where $\mathbf{W}_{C2}^{(l)}$ comprises the quantized SB coefficient of all the N_3 SBs and $\mathbf{\Lambda}^{(l)} = \text{diag}\left(\left[\Lambda_0^{(l)} \cdots \Lambda_{N_3-1}^{(l)}\right]\right)$ is the diagonal matrix gathering the normalization factors of all the SB precoders. The t-th column of $\mathbf{W}^{(l)}$ is the precoder to be applied for t-th SB of the l-th layer.

B. PRECODER SELECTION STRATEGY

To find the best precoder among the Type-II codebook, the beam matrix $\mathbf{W}_1$ is selected first. It allows reducing the channel to an effective channel of dimension $N_R \times 2L$ per SB in the angle domain. Then, the BCC matrix $\mathbf{W}_2^{(l)} \in \mathcal{C}^{2L\times N_3}$ made of the unquantized BCCs of all SBs is derived conditional on $\mathbf{W}_1$. From the BCC matrix $\mathbf{W}_2^{(l)}$, after a phase pre-processing and amplitude normalization, the matrix $\mathbf{W}_{C1}^{(l)}$ is first obtained than $\mathbf{W}_{C2}^{(l)}$. The flow chart of the calculation of the Rel. 15 Type-II codebook is depicted in Fig. 4. The selection strategy is elaborated in five steps as follows.

Step 1. **Calculation of the beam matrix $\mathbf{W}_1$:** The beam selection aims at selecting the best L IDFT beams that maximize the channel sum power across both polarizations. Theoretically, this optimization problem can be formulated as

$$\left(\mathbf{v}_{m_1^{(i)},m_2^{(i)}}\right)_{i=0}^{L-1} = \underset{(\mathbf{v}_{m_i,l_i})_{i=0}^{L-1}}{\arg\max} \sum_{i=0}^{L-1} \mathbf{v}_{m_i,l_i}^{\mathsf{H}}(\mathbf{R}_{11} + \mathbf{R}_{22})\mathbf{v}_{m_i,l_i} \tag{36}$$

where each vector $\mathbf{v}_{m_i,l_i}$ belongs to the IDFT grid of beams $\mathcal{B}$ and where matrices $\mathbf{R}_{11}$ and $\mathbf{R}_{22}$ are the diagonal blocks of the wideband covariance matrix of the channel $\mathbf{R}$ given in Eq. (24). Nonetheless, to satisfy the rules of Eq. (24). in the standard, the L IDFT beam selection for the Type-II codebook should be carried out in an iterative manner. The selection algorithm for the L IDFT beams which considers the beam orthogonality property is described in Eqs. (33) and (34). There is no loss of optimality as the selected beams are mutually orthogonal.

Step 2. **Calculation of the BCC matrix $\mathbf{W}_2^{(l)}$:** The BCC matrix $\mathbf{W}_2^{(l)}$ comprising the unquantized BCCs of $2L$ beams and N_3 SBs has the following form

$$\begin{aligned}\mathbf{W}_2^{(l)} &= \begin{bmatrix} w_{2,0,0}^{(l)} & \cdots & w_{2,0,N_3-1}^{(l)} \\ \vdots & & \vdots \\ w_{2,2L-1,0}^{(l)} & \cdots & w_{2,2L-1,N_3-1}^{(l)} \end{bmatrix} \\ &= \left[\mathbf{w}_{2,0}^{(l)} \cdots \mathbf{w}_{2,N_3-1}^{(l)}\right]\end{aligned} \tag{37}$$

where $\mathbf{w}_{2,t}^{(l)} \in \mathbb{C}^{2L\times 1}$ is the BCC vector of the t-th SB and can be obtained by solving the following optimization problem knowing $\mathbf{w}_{2,t}^{(l')}$ for $l' < l$

$$\mathbf{w}_{2,t}^{(l)} = \underset{\mathbf{w}\in\mathbb{C}^{2L\times 1}\text{ s. t. } C_1 \text{ and } C_2}{\arg\max} \mathbf{w}^{\mathsf{H}}\mathbf{H}_A^{\mathsf{H}}(t)\mathbf{H}_A(t)\mathbf{w} \tag{38}$$

under the constraints $C_1 : \mathbf{w}^{\mathsf{H}}\mathbf{w} = 1$ and $C_2 : \mathbf{w}^{\mathsf{H}}\mathbf{w}_{2,t}^{(l')} = 0$ for $l' < l$ and where $\mathbf{H}_A(t)$ is the t-th SB effective channel of dimension $N_R \times 2L$ given as:

$$\mathbf{H}_A(t) = \mathbf{H}(t)\mathbf{W}_1. \tag{39}$$

Solving the optimization problem results in the SB precoder $\mathbf{w}_{2,t}^{(l)}$ that maximizes, for a given SB, the product between the channel, the matrix $\mathbf{W}_1$ and the SB precoder $\mathbf{w}_{2,t}^{(l)}$ while ensuring that the precoder for layer l is orthogonal to the ones selected previously for the layers 1 to $l-1$. The optimal solution to the above mentioned problem is

$$\mathbf{w}_{2,t}^{(l)} = eig(\mathbf{H}_A^{\mathsf{H}}(t)\mathbf{H}_A(t))_l \tag{40}$$

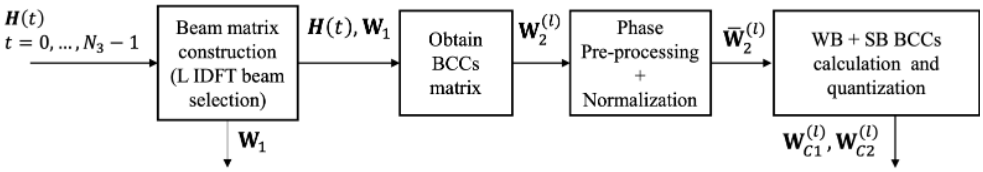


FIGURE 4. Flowchart showing the steps involved in the selection of the Rel. 15 Type-II precoder.

Algorithm 1 Beam Selection Algorithm

Input: $\mathbf{R}_{11}, \mathbf{R}_{22}, N_1, N_2, O_1$ and O_2.
Output: $\mathbf{v}_{m_1^{(0)}, m_2^{(0)}}, \cdots, \mathbf{v}_{m_1^{(L-1)}, m_2^{(L-1)}}$.
Initialization : $i = 0, \mathcal{N}_1 = \{0, \cdots, N_1 - 1\}$ and $\mathcal{N}_2 = \{0, \cdots, N_2 - 1\}$
while $i < L$ **do**
  if $i == 0$ **then**
    $m_1^{(0)}, m_2^{(0)} = \arg\max_{l,m} \mathbf{v}_{l,m}^{\mathsf{H}} (\mathbf{R}_{11} + \mathbf{R}_{22}) \mathbf{v}_{l,m}$
    $q_1 = m_1^{(0)} \mod O_1$
    $q_2 = m_2^{(0)} \mod O_2$
    $n_1^{(0)} = \lceil \frac{m_1^{(0)}}{O_1} \rceil$
    $n_2^{(0)} = \lceil \frac{m_2^{(0)}}{O_2} \rceil$
    $\mathcal{N}_1 = \mathcal{N}_1 \setminus \{n_1^{(0)}\}$
    $\mathcal{N}_2 = \mathcal{N}_2 \setminus \{n_2^{(0)}\}$
    $i = i + 1$
  else
    $n_1^{(i)}, n_2^{(i)} = \arg\max_{\substack{n_1 \in \mathcal{N}_1 \\ n_2 \in \mathcal{N}_2}} \mathbf{v}_{n_1 Q_1 + q_1, n_2 Q_2 + q_2}^{\mathsf{H}} (\mathbf{R}_{11} + \mathbf{R}_{22}) \mathbf{v}_{n_1 Q_1 + q_1, n_2 Q_2 + q_2}$
    $m_1^{(i)} = Q_1 n_1^{(i)} + q_1$
    $m_2^{(i)} = Q_2 n_2^{(i)} + q_2$
    $i = i + 1$
  end if
  Calculate $\mathbf{v}'_{m_1^{(i)}}$ based on Eq. (11).
  Calculate $\mathbf{u}_{m_2^{(i)}}$ based on Eq. (12).
  $\mathbf{v}_{m_1^{(i)}, m_2^{(i)}} = \mathbf{v}'_{m_1^{(i)}} \otimes \mathbf{u}_{m_2^{(i)}}$
end while

where $eig(\cdot)_l$ outputs the l-th eigenvector of the input square matrix in decreasing order of the associated eigenvector magnitudes. This solution is also called matrix projection method that shows better performance than other solutions as studied in [45]. Note that the obtained solution following Eq. (40) minimizes the interference between the spatial layers at the UE side since

$$\mathbf{w}_{2,t}^{(l)\mathsf{H}} \left(\mathbf{H}_A^{\mathsf{H}}(t) \mathbf{H}_A(t) \right) \mathbf{w}_{2,t}^{(l')} = 0 \text{ for all } l' \neq l. \tag{41}$$

Step 3. **Phase pre-processing and amplitude normalization:** As the SB precoders are calculated for each SB separately by eigen decomposition, phase pre-processing is essential [41] to alleviate the phase discontinuity between the SB precoders. Moreover, for the quantization, the BCC matrix needs to be normalized such that amplitude and phase of the BCCs associated with the strongest beam is equal to 1 and 0 respectively. Hence, the phase pre-processing and amplitude normalization function $\mathcal{N}(\cdot, \cdot)$ is applied to $\mathbf{W}_2^{(l)}$ in a column-wise manner as

$$\begin{aligned} \bar{\mathbf{W}}_2^{(l)} &= \left[\bar{\mathbf{w}}_{2,0}^{(l)} \cdots \bar{\mathbf{w}}_{2,N_3-1}^{(l)} \right] \\ &= \left[\mathcal{N}(\mathbf{w}_{2,0}^{(l)}, i'_l) \cdots \mathcal{N}(\mathbf{w}_{2,N_3-1}^{(l)}, i'_l) \right] \end{aligned} \tag{42}$$

where i'_l is the index of the strongest beam across both polarizations in average over the N_3 SBs for layer l. The index i'_l is obtained as

$$i'_l = \arg\max_i \sum_{t=0}^{N_3-1} |w_{2,i,t}^{(l)}|^2. \tag{43}$$

Specifically, the t-th column of matrix $\bar{\mathbf{W}}_2^{(l)}$ is computed by

$$\mathcal{N}(\mathbf{w}_{2,t}^{(l)}, i'_l) = \left[\frac{w_{2,0,t}^{(l)}}{w_{2,i'_l,t}^{(l)}} \cdots \frac{w_{2,i'_l,t}^{(l)}}{w_{2,i'_l,t}^{(l)}} \cdots \frac{w_{2,2L-1,t}^{(l)}}{w_{2,i'_l,t}^{(l)}} \right]^{\mathsf{T}} \tag{44}$$

As a result, matrix $\bar{\mathbf{W}}_2^{(l)}$ becomes

$$\bar{\mathbf{W}}_2^{(l)} = \begin{bmatrix} \frac{w_{2,0,0}^{(l)}}{w_{2,i'_l,0}^{(l)}} & \cdots & \frac{w_{2,0,t}^{(l)}}{w_{2,i'_l,t}^{(l)}} & \cdots & \frac{w_{2,0,N_3-1}^{(l)}}{w_{2,i'_l,N_3-1}^{(l)}} \\ \vdots & & \vdots & & \vdots \\ 1 & \cdots & 1 & \cdots & 1 \\ \vdots & & \vdots & & \vdots \\ \frac{w_{2,2L-1,0}^{(l)}}{w_{2,i'_l,0}^{(l)}} & \cdots & \frac{w_{2,2L-1,t}^{(l)}}{w_{2,i'_l,t}^{(l)}} & \cdots & \frac{w_{2,2L-1,N_3-1}^{(l)}}{w_{2,i'_l,N_3-1}^{(l)}} \end{bmatrix} \tag{45}$$

where the i'_l-th row is a all-one row.

Step 4. **Wideband coefficient calculation and quantization:** The amplitude $p_{l,i}^{(1)}$ takes the quantized value of the highest absolute value among the components of the i-th row of matrix $\bar{\mathbf{W}}_2^{(l)}$. Hence, the computation and quantization of the i-th wideband coefficient is given by

$$p_{l,i}^{(1)} = Q_{WB}\left(\max_t \left| \bar{w}_{2,i,t}^{(l)} \right| \right) \tag{46}$$

where $Q_{WB}(\cdot)$ is the quantization function that maps the absolute value of its entry to the closest value in the set $\mathcal{P}_1$. Note that there is no quantization error for the strongest beam associated coefficients since $\bar{w}_{2,i'_l,t}^{(l)}$ is always equal to 1. At the end of step 4, the matrix $\mathbf{W}_{C1}^{(l)}$ is obtained.

Step 5. **SB coefficient calculation and quantization:** After calculating $\bar{\mathbf{W}}_2^{(l)}$, matrix $\mathbf{D}^{(l)}$ comprising the differential amplitudes and phase values is obtained as

$$\mathbf{D}^{(l)} = \left(\mathbf{W}_{C1}^{(l)}\right)^{-1}\bar{\mathbf{W}}_2^{(l)} \tag{47}$$

Based on $\mathbf{D}^{(l)}$, the matrix $\mathbf{W}_{C2}^{(l)}$ is calculated such that $w_{C2,i,t}^{(l)} = p_{l,i,t}^{(2)}\varphi_{l,i,t}$ where $p_{l,i,t}^{(2)} = Q_{SB}(d_{i,t}^{(l)})$ and $\varphi_{l,i,t} = Q_P(d_{i,t}^{(l)})$. The function $Q_{SB}(\cdot)$ maps the absolute value of its entry to the closest value in the set $\mathcal{P}_2$ while $Q_P(\cdot)$ maps the phase of its entry to the closest value in the set Φ.

C. FEEDBACK QUANTITY ANALYSIS

We consider the case where the higher layer parameter *subbandAmplitude* is set to 'true', i.e., the CSI report includes quantized SB amplitudes. The CSI report containing the PMI of the Type-II codebook consists of two parts. The first part relates to wideband information (referred to as i_1 in the standard), it comprises the following indicators: a $\lceil \log_2(O_1O_2)\rceil$-bit indicator to indicate the selected rotation factor among O_1O_2 groups referred to as $(i_{1,1})$, a $\lceil \log_2 \binom{N_1N_2}{L}\rceil$-bit indicator to indicate the selected L IDFT beams from N_1N_2 IDFT beams referred to as $(i_{1,2})$, a $\lceil \log_2(2L)\rceil$-bit indicator to indicate the position of the strongest beam among the $2L$ beams for each l-th layer referred to as $i_{1,3,l}$ and $3\cdot(2L-1)$-bit indicator to indicate the wideband amplitudes of the remaining $2L-1$ beams referred to as $i_{1,4,l}$. Note that $i_{1,3,l}$ also indicates the position of a wideband amplitude value equal to 1. The second part relates to information per SB (referred to as i_2 in the standard) and can vary in size with respect to the number of non zero wideband amplitudes denoted $M_l \leq 2L$ in the standard. In the following, we assume the worst case in terms of payload size that is to say $M_l = 2L$. It comprises the following indicators: a $(K^{(2)}-1)$-bit indicator to indicate the subband amplitudes of the first $K^{(2)}$ beams for each SB and each l-th layer referred to as $i_{2,1,l}$, a $\log_2(N_{PSK})\cdot(K^{(2)}-1)$-bit indicator to indicate the phase values of the subband amplitudes of the first $K^{(2)}$ beams and a $2\cdot(2L-K^{(2)})$-bit indicator to indicate the phase values of the subband amplitudes of the remaining $2L-K^{(2)}$ beams for each SB and each l-th layer, respectively. The two latter indicators are referred to as $i_{2,2,l}$. Note that the SB amplitudes are not reported for the strongest beam indicated by $i_{1,3,l}$. The last $2L-K^{(2)}$ beams are determined based on the ordering with respect to the strongest beam. For values of L equal to 2,3, and 4, the values of $K^{(2)}$ are given by 4,4, and 6, respectively. Note that the values of N_{PSK} and L are given by the higher layer parameters *numberOfBeams* and *phaseAlphabetSize*, respectively.

VI. ENHANCED TYPE-II CODEBOOK

A. DESCRIPTION

The enhanced Type-II (eType-II) codebook standardized in the 3GPP Rel. 16 [46] supports rank up to 4. The eType-II precoder of the t-th SB ($0 \leqslant t \leqslant N_3-1$) for rank 1, rank 2, rank 3, and rank 4 are given as

$$\mathbf{W}_t^1 = \mathbf{w}_t^{(1)}, \tag{48}$$

$$\mathbf{W}_t^2 = \frac{1}{\sqrt{2}}\left[\mathbf{w}_t^{(1)}\ \mathbf{w}_t^{(2)}\right], \tag{49}$$

$$\mathbf{W}_t^3 = \frac{1}{\sqrt{3}}\left[\mathbf{w}_t^{(1)}\ \mathbf{w}_t^{(2)}\ \mathbf{w}_t^{(3)}\right], \tag{50}$$

$$\mathbf{W}_t^4 = \frac{1}{2}\left[\mathbf{w}_t^{(1)}\ \mathbf{w}_t^{(2)}\ \mathbf{w}_t^{(3)}\ \mathbf{w}_t^{(4)}\right], \tag{51}$$

respectively, where

$$\mathbf{w}_t^{(l)} = \Lambda_t^{(l)}\begin{bmatrix}\sum_{i=0}^{L-1}\mathbf{v}_{m_1^{(i)},m_2^{(i)}}p_{l,0}^{(1)}\sum_{f=0}^{M_v-1}y_{t,l}^{(f)}p_{l,i,f}^{(2)}\varphi_{l,i,f}\\ \sum_{i=0}^{L-1}\mathbf{v}_{m_1^{(i)},m_2^{(i)}}p_{l,1}^{(1)}\sum_{f=0}^{M_v-1}y_{t,l}^{(f)}p_{l,i+L,f}^{(2)}\varphi_{l,i+L,f}\end{bmatrix}$$

is the eType-II precoder for the l-th layer. As the eType-II codebook considers frequency domain (FD) compression, to facilitate a deeper understanding of the codebook, we concatenate the l-th layer eType-II precoder $\mathbf{w}_t^{(l)}$ of all N_3 SBs as formulated in Eq. (52), as shown at the bottom of the next page. Matrix $\mathbf{W}_1$ is the same beam matrix as in Rel. 15 Type-II codebook, matrix $\tilde{\mathbf{W}}_{C1}^{(l)}$ comprises $2L$ polarization specific wideband amplitudes associated with the $2L$ IDFT beams across the two polarizations, matrix $\tilde{\mathbf{W}}_{C2}^{(l)}$ comprises the compressed/quantized SB coefficients, $\mathbf{W}_f^{(l)}$ is the FD compression matrix and $\mathbf{\Lambda}^{(l)} = diag\left(\left[\Lambda_0^{(l)}\ \cdots\ \Lambda_{N_3-1}^{(l)}\right]\right)$ is the normalization matrix with $\Lambda_t^{(l)} = \frac{1}{\sqrt{N_1N_2\gamma_{t,l}}}$ where $\gamma_{t,l}$ is the normalization factor given as

$$\gamma_{t,l} = \sum_{i=0}^{2L-1}\left(p_{l,\lfloor\frac{i}{L}\rfloor}^{(1)}\right)^2\left|\sum_{f=0}^{M_v-1}y_{t,l}^{(f)}p_{l,i,f}^{(2)}\varphi_{l,i,f}\right|^2. \tag{53}$$

The polarization specific wideband amplitude matrix $\tilde{\mathbf{W}}_{C1}^{(l)}$ has a diagonal form as shown below

$$\tilde{\mathbf{W}}_{C1}^{(l)} = diag\left(\left[\underbrace{p_{l,0}^{(1)}\cdots p_{l,0}^{(1)}}_{L\text{ items}}\underbrace{p_{l,1}^{(1)}\cdots p_{l,1}^{(1)}}_{L\text{ items}}\right]\right) \tag{54}$$

where $p_{l,0}^{(1)}$ and $p_{l,1}^{(1)}$ are the wideband amplitudes of the first and second polarization respectively. Note that the wideband amplitudes for the eType-II precoder is identical for all L beams in each polarization. This is in contrast to the different wideband amplitudes of the Type-II precoder in Eq. (29). Reference [50] shows that the differential quantization scheme standardized in Rel. 16 achieves better performance and overhead trade-off. The polarization specific amplitude values are drawn from the set $\tilde{\mathcal{P}}_1 = \left\{\frac{1}{\sqrt{128}}, \frac{1}{8192}^{1/4}, \frac{1}{8}, \frac{1}{2048}^{1/4}, \frac{1}{2\sqrt{8}}, \frac{1}{512}^{1/4}, \frac{1}{4}, \frac{1}{128}^{1/4}, \frac{1}{\sqrt{8}}, \frac{1}{32}^{1/4}, \frac{1}{2}, \frac{1}{8}^{1/4}, \frac{1}{\sqrt{2}}, \frac{1}{2}^{1/4}, 1\right\}$. Note that the quantized wideband amplitude of the polarization associated with the strongest beam is always equal to 1.

For better understanding, the eType-II precoder is compared with the Type-II precoder. From Eq. (35), the Type-II precoder for the l-th layer for all SBs is written as product of

the beam matrix $\mathbf{W}_1$, wideband amplitude matrix $\mathbf{W}_{C1}^{(l)}$ and SB coefficients matrix $\mathbf{W}_{C2}^{(l)} \in \mathbb{C}^{2L\times N_3}$ where the latter two matrices are obtained by quantizing the amplitude and phase of the coefficients of the BCC matrix $\mathbf{W}_2^{(l)} \in \mathbb{C}^{2L\times N_3}$. For the Type-II codebook all the $2LN_3$ quantized SB coefficients are reported as a part of the PMI feedback. Note that the feedback overhead increases proportionally with L or N_3 which was been the main motivation for the standardization of the Rel. 16 codebook. The Rel. 16 eType-II codebook effectively reduces the feedback overhead by exploiting the correlation between the coefficients of the BCCs matrix $\mathbf{W}_2^{(l)}$ in the frequency domain (FD) as demonstrated in Fig. 5a. Precisely, matrix $\mathbf{W}_2^{(l)} \in \mathbb{C}^{2L\times N_3}$ is compressed to matrix $\tilde{\mathbf{W}}_2^{(l)} \in \mathbb{C}^{2L\times M_v}$ ($M_v \ll N_3$) by the layer specific FD compression matrix $\mathbf{W}_f^{(l)} \in \mathbb{C}^{N_3\times M_v}$. The matrix $\mathbf{W}_f^{(l)}$ consists of M_v distinct DFT vectors of size N_3, expressed as

$$\mathbf{W}_f^{(l)} = \begin{bmatrix} y_{0,l}^{(0)} & \cdots & y_{M_v-1,l}^{(0)} \\ \vdots & & \vdots \\ y_{0,l}^{(N_3-1)} & \cdots & y_{M_v-1,l}^{(N_3-1)} \end{bmatrix}^{\mathsf{H}} = \begin{bmatrix} \mathbf{f}_{\tau_0}^{(l)} & \cdots & \mathbf{f}_{\tau_{M_v-1}}^{(l)} \end{bmatrix}^{\mathsf{H}} \tag{55}$$

where

$$\mathbf{f}_{\tau_m}^{(l)} = \begin{bmatrix} 1 & e^{-j\frac{2\pi\tau_m^{(l)}}{N_3}} & \cdots & e^{-j\frac{2\pi\tau_m^{(l)}(N_3-1)}{N_3}} \end{bmatrix}^{\mathsf{T}} \tag{56}$$

and $\tau_m^{(l)} \in \{0, 1, \cdots, N_3-1\}$. The number of distinct FD basis in matrix $\mathbf{W}_f^{(l)}$ is defined as

$$M_v = \left\lceil p_v \frac{N_3}{R} \right\rceil \tag{57}$$

in the standard. Note that in the specification the delay $\tau_m^{(l)}$ is denoted $n_{3,l}^{(m)}$. The compressed BCCs matrix $\tilde{\mathbf{W}}_2^{(l)}$ is calculated as

$$\tilde{\mathbf{W}}_2^{(l)} = \mathbf{W}_2^{(l)}\mathbf{W}_f^{(l)}. \tag{58}$$

Fig. 5 illustrates the sparsity of the compressed BCCs matrix $\tilde{\mathbf{W}}_2^{(l)}$ with $M_v = 3$. Therefore, the feedback overhead can be further reduced by reporting only the strongest K_l^{NZ} ($K_l^{NZ} \ll 2LM_v$) non-zero coefficients (NZCs) from $\tilde{\mathbf{W}}_2^{(l)}$. The location of the reported NZCs of the l-th layer is indicated by a bitmap matrix $\mathbf{K}_l$ of size $2L \times M_v$

$$\mathbf{K}_l = \begin{bmatrix} k_{l,0,0}^{(3)} & \cdots & k_{l,0,M_v-1}^{(3)} \\ \vdots & & \vdots \\ k_{l,2L-1,0}^{(3)} & \cdots & k_{l,2L-1,M_v-1}^{(3)} \end{bmatrix} \tag{59}$$

where $k_{l,i,f}^{(3)} \in \{0,1\} \forall i \in \{0,\cdots,2L-1\}, \forall f \in \{0,\cdots,M_v-1\}$) is an indicator for the entry in the i-th row and f-th column of the BCC matrix. If $k_{l,i,f}^{(3)} = 1$, the associated entry $p_{l,i,f}^{(2)}\varphi_{l,i,f}$ of the BCC matrix $\tilde{\mathbf{W}}_{C2}^{(l)}$ is reported after quantization. Moreover, the standard defines the limit on the number of reported NZCs for the eType-II codebook as

$$K_l^{NZ} = \sum_{i=0}^{2L-1}\sum_{f=0}^{M_v-1} k_{l,i,f}^{(3)} \leqslant K_0, \tag{60}$$

$$K^{NZ} = \sum_{l=1}^{v} K_l^{NZ} \leqslant 2K_0, \tag{61}$$

with

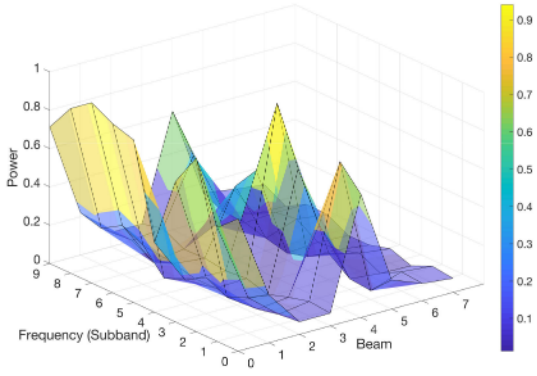
$$K_0 = \lceil \beta 2LM_1 \rceil. \tag{62}$$

The values for L, β and p_v in the standard are provided with respect to the higher layer parameter *paramCombination-r16* in Table 5.2.2.2.5-1 of Release-16 [46]. For the reported elements of the BCC matrix, the differential

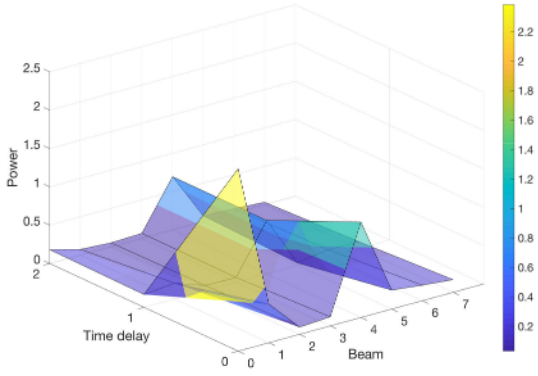
$$\underbrace{\begin{bmatrix} \mathbf{w}_0^{(l)} & \cdots & \mathbf{w}_{N_3-1}^{(l)} \end{bmatrix}}_{2N_1N_2\times N_3} = \underbrace{\mathbf{W}_1}_{2N_1N_2\times 2L} \underbrace{\begin{bmatrix} p_{l,0}^{(l)} & & & & & \\ & \ddots & & & & \\ & & p_{l,0}^{(l)} & & & \\ & & & p_{l,1}^{(l)} & & \\ & & & & \ddots & \\ & & & & & p_{l,1}^{(l)} \end{bmatrix}}_{\tilde{\mathbf{W}}_{C1}^{(l)}:\, 2L\times 2L}$$

$$\underbrace{\begin{bmatrix} p_{l,0,0}^{(2)}\varphi_{l,0,0} & \cdots & p_{l,0,M_v-1}^{(2)}\varphi_{l,0,M_v-1} \\ \vdots & & \vdots \\ p_{l,2L-1,0}^{(2)}\varphi_{l,2L-1,0} & \cdots & p_{l,2L-1,M_v-1}^{(2)}\varphi_{l,2L-1,M_v-1} \end{bmatrix}}_{\tilde{\mathbf{W}}_{C2}^{(l)}:\, 2L\times M_v} \underbrace{\begin{bmatrix} y_{0,l}^{(0)} & \cdots & y_{N_3-1,l}^{(0)} \\ \vdots & & \vdots \\ y_{0,l}^{(M_v-1)} & \cdots & y_{N_3-1,l}^{(M_v-1)} \end{bmatrix}}_{\mathbf{W}_f^{(l)\mathsf{H}}:\, M_v\times N_3} \mathbf{\Lambda}^{(l)}$$

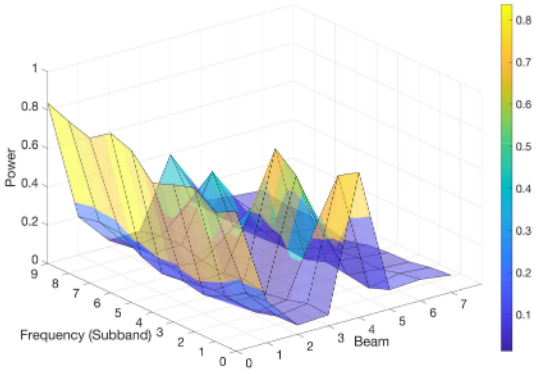
$$= \mathbf{W}_1\tilde{\mathbf{W}}_{C1}^{(l)}\tilde{\mathbf{W}}_{C2}^{(l)}\mathbf{W}_f^{(l)\mathsf{H}}\mathbf{\Lambda}^{(l)} \tag{52}$$



(a) Illustration of matrix $\mathbf{W}_2^{(l)}$.



(b) Illustration of matrix $\tilde{\mathbf{W}}_2^{(l)}$



(c) Illustration of matrix product $\tilde{\mathbf{W}}_2^{(l)}\mathbf{W}_f^{(l)\mathsf{H}}$

FIGURE 5. Illustration of the FD compression in the enhanced Type-II codebook.

amplitudes and phase values are drawn from set $\tilde{\mathcal{P}}_2 = \{\frac{1}{8\sqrt{2}}, \frac{1}{8}, \frac{1}{4\sqrt{2}}, \frac{1}{4}, \frac{1}{2\sqrt{2}}, \frac{1}{2}, \frac{1}{\sqrt{2}}, 1\}$ and $\tilde{\Phi} = \{e^{\frac{j2\pi n}{16}}, n \in \{0, 1, \cdots, 15\}\}$, respectively.

For the eType-II codebook, the BS receives $\tilde{\mathbf{W}}_{C1}^{(l)}\tilde{\mathbf{W}}_{C2}^{(l)}$ the quantized version of $\tilde{\mathbf{W}}_2^{(l)}$ with $2LM_v - K_l^{NZ}$ coefficients forced to zero and $\mathbf{W}_f^{(l)}$ as part of the PMI feedback and reconstructs the original BCC matrix as

$$\mathbf{W}_2^{(l)} \approx \tilde{\mathbf{W}}_2^{(l)}\mathbf{W}_f^{(l)\mathsf{H}} \approx \tilde{\mathbf{W}}_{C1}^{(l)}\tilde{\mathbf{W}}_{C2}^{(l)}\mathbf{W}_f^{(l)\mathsf{H}}. \tag{63}$$

Fig. 5c depicts the approximate matrix $\tilde{\mathbf{W}}_2^{(l)}\mathbf{W}_f^{(l)\mathsf{H}}$. Comparing Fig. 5a and Fig. 5c, it can be observed that the reconstructed BCC matrix retains most of the significant power from the matrix $\mathbf{W}_2^{(l)}$.

As illustrated in Fig. 6, every beam from the beam matrix $\mathbf{W}_1$ covers a 3D spatial region where more than one scatterer is located. Depending on the beam width, each beam may be associated with one or multiple delays. The eType-II precoder exploits the dominant delays of each beam to reduce the feedback overhead as the channel is sparse in the angle-delay domain.

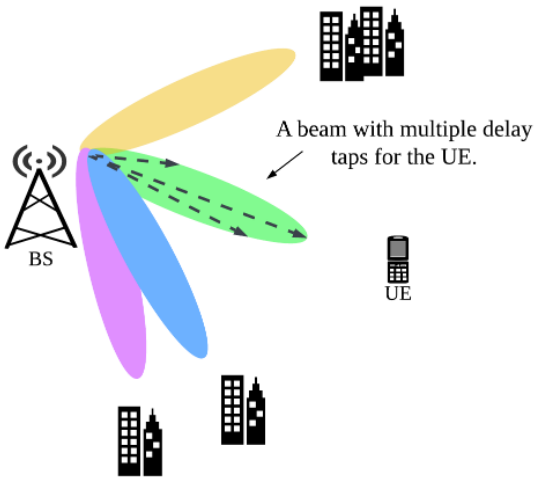


FIGURE 6. Illustration of a massive MIMO transmission in the angle-delay domain.

B. PRECODER SELECTION STRATEGY IN THE FREQUENCY DOMAIN

Fig. 7 shows the general procedure of the low-complexity FD selection strategy for the eType-II codebook. The

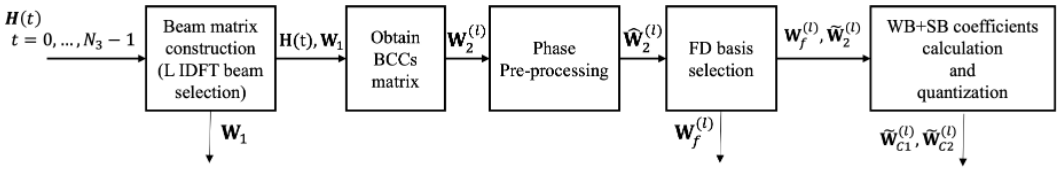


FIGURE 7. Flowchart showing the steps involved in the selection of the eType-II precoder using FD strategy.

FD selection strategy is elaborated in the following six steps:

Step 1. **Constructon of Beam matrix $\mathbf{W}_1$**: As the eType-II codebook shares the same beam matrix structure as the Type-II codebook, the L IDFT beams of matrix $\mathbf{W}_1$ are chosen based on Algorithm 1.

Step 2. **Calculation of the BCC matrix $\mathbf{W}_2^{(l)}$**: Similar to the Step 2 of the BCC matrix calculation of Type-II codebook explained in Section V-B, the t-th column of matrix $\mathbf{W}_2^{(l)}$ is obtained by the matrix projection method [45] given in Eq. (40).

Step 3. **Phase pre-processing**: In order to alleviate phase discontinuity between adjacent SBs and enhance performance [41], phase pre-processing function $\mathcal{F}_p(\cdot)$ is applied to every column of matrix $\mathbf{W}_2^{(l)}$ as

$$\begin{aligned}\hat{\mathbf{W}}_2^{(l)} &= \left[\hat{\mathbf{w}}_{2,0}^{(l)} \cdots \hat{\mathbf{w}}_{2,N_3-1}^{(l)}\right] \\ &= \left[\mathcal{F}_p(\mathbf{w}_{2,0}^{(l)}, i_l') \cdots \mathcal{F}_p(\mathbf{w}_{2,N_3-1}^{(l)}, i_l')\right] \end{aligned} \tag{64}$$

where i_l' is the index of the strongest beam across the two polarization in average over the N_3 SBs for layer l as defined in Eq. (43). The t-th column of matrix $\hat{\mathbf{W}}_2^{(l)}$ is calculated according to

$$\mathcal{F}_p(\mathbf{w}_{2,t}^{(l)}, i_l') = \left[\frac{w_{2,0,t}^{(l)} w_{2,i_l',t}^{(l)*}}{|w_{2,i_l',t}^{(l)}|} \cdots \frac{w_{2,2L-1,t}^{(l)} w_{2,i_l',t}^{(l)*}}{|w_{2,i_l',t}^{(l)}|}\right]^{\mathsf{T}}, \tag{65}$$

where $w_{2,i_l',t}^{(l)*}$ is the complex conjugate of $w_{2,i_l',t}^{(l)}$. Note that the phase pre-processing function $\mathcal{F}_p(\cdot)$ solely changes the phase of the input elements while the nomalization plus phase pre-processing function $\mathcal{N}(\cdot)$ changes both the phase and amplitude of the input elements.

Step 4. **Construction of FD compression matrix $\mathbf{W}_f$**: The FD compression matrix $\mathbf{W}_f^{(l)}$ comprises M_v FD DFT basis vectors, where the M_v DFT basis vectors are associated with the dominant delays associated with the BCC matrix for each layer l. As suggested in [45], the M_v DFT bases for FD compression are selected based on maximum power criterion in a sequential way as provided in Algorithm 2. After obtaining the FD compression matrix compressed BCCs matrix are calculated as

$$\tilde{\mathbf{W}}_2^{(l)} = \hat{\mathbf{W}}_2^{(l)} \mathbf{W}_f^{(l)}. \tag{66}$$

Step 5. **Calculation of Wideband coefficients and quantization**: The wideband coefficients $p_{l,0}^{(1)}$ and $p_{l,1}^{(1)}$ denote

Algorithm 2 FD Compression Basis Selection Algorithm

Input: $\hat{\mathbf{W}}_2^{(l)}, M_v$.
Output: $\mathbf{f}_{\tau_0}^{(l)}, \cdots, \mathbf{f}_{\tau_{M-1}}^{(l)}$.
Initialization : $m = 0$ and $\mathcal{N} = \{0, \cdots, N_3 - 1\}$.
while $m < M$ **do**
  $n^\star = \arg\max_{n \in \mathcal{N}} \left(\hat{\mathbf{W}}_2^{(l)} \mathbf{f}_n\right)^{\mathsf{H}} \left(\hat{\mathbf{W}}_2^{(l)} \mathbf{f}_n\right)$ where $\mathbf{f}_n$ is computed based on Eq. (56).
  $\tau_m^{(l)} = n^\star$
  $\mathcal{N} = \mathcal{N} \setminus \{n^\star\}$
  $m = m + 1$
  Calculate $\mathbf{f}_{\tau_m}^{(l)}$ according to Eq. (56).
end while

the maximum amplitude value of the coefficients of the first and second polarization, respectively. We define the strongest beam $i_l^* \in \{0, \cdots, 2L-1\}$ and strongest delay index $m_l^* \in \{0, \cdots, M_v - 1\}$ in the angle-delay domain for layer l as

$$(i_l^*, m_l^*) = \arg\max_{(i,m)} |\tilde{w}_{2,i,m}^{(l)}|. \tag{67}$$

In order to reduce the feedback related to the PMI, the standard specifies that the delays $\tau_m^{(l)}$ are remapped with respect to the strongest one $\tau_{m_l^*}^{(l)}$ such that $\tau_m^{(l)} = (\tau_m^{(l)} - \tau_{m_l^*}^{(l)}) \mod N_3$ in order to have the strongest delay at zero delay. Since a precoding vector is defined within a phase factor, it does not impact the performance (reference phase subtraction in each column of $\mathbf{W}_f^{(l)}$). The index m is also remapped with respect to m_l^* as $m = (m - m_l^*) \mod M_v$, such that the index of the strongest delay is $m_l^* = 0$ for each layer l after remapping. The latter operation corresponds to a permutation of the rows of $\mathbf{W}_f^{(l)}$. It ensures that the strongest coefficient of $\tilde{\mathbf{W}}_2^{(l)}$ belongs to its first column and that the zero delay is always selected among the M_v delays. The standard further stipulates that the polarization, in which the strongest beam is, should report its wideband coefficient always equal to 1. Therefore, the matrix $\tilde{\mathbf{W}}_2^{(l)}$ should be first normalized as follows:

$$\bar{\mathbf{W}}_2^{(l)} = \frac{\tilde{\mathbf{W}}_2^{(l)}}{|\tilde{w}_{2,i_l^*,m^*}^{(l)}|}. \tag{68}$$

Therefore, the two wideband coefficient are obtained by

$$p_{l,0}^{(1)} = Q_{WB}\big(\max_{i,m}(\bar{w}_{2,i,m}^{(l)})\big), \tag{69}$$

$$p_{l,1}^{(1)} = Q_{WB}\big(\max_{i,m}(\bar{w}_{2,i,m}^{(l)})\big), \tag{70}$$

where $Q_{WB}(\cdot)$ is the quantization function that maps the absolute value of its entry to the closest value in the set $\tilde{\mathcal{P}}_1$. Consequently, the diagonal matrix $\tilde{\mathbf{W}}_{C1}^{(l)}$ is built based on Eq. (54).

Step 6. **SB coefficients calculation and quantization**: After calculating $\tilde{\mathbf{W}}_{C2}^{(l)}$, matrix $\mathbf{D}^{(l)}$ comprising the differential amplitudes and phase values is obtained as

$$\mathbf{D}^{(l)} = \big(\tilde{\mathbf{W}}_{C1}^{(l)}\big)^{-1}\bar{\mathbf{W}}_{2}^{(l)}. \tag{71}$$

Note that $\mathbf{D}^{(l)}$ comprises K_l^{NZ} NZCs, where $K^{NZ} = \sum_{l=0}^{\nu-1} K_l^{NZ}$ as given in Eq. (60). Based on $\mathbf{D}^{(l)}$, the matrix $\tilde{\mathbf{W}}_{C2}^{(l)}$ and the associated bitmap indicator matrix $\mathbf{K}_l$ are calculated based on Algorithm 3 where the function $Q_{SB}(\cdot)$ maps the absolute value of its entry to the closest value in the set $\tilde{\mathcal{P}}_2$ and $Q_P(\cdot)$ maps the phase of its entry to the closest value in the set $\tilde{\Phi}$.

For the sake of the precoded channel rank preservation, we select the best $\lfloor \alpha_\nu K_0 \rfloor$ NZP coefficients for each layer l where $\alpha_\nu = 1$ for $\nu = 1$ and $\alpha_\nu = 2/\nu$ for $\nu = 2, 3, 4$.

Algorithm 3 Algorithm to Compute Matrices $\tilde{\mathbf{W}}_{C2}^{(l)}$ and $\mathbf{K}_l$

Input: $\mathbf{D}^{(l)}, M_\nu, L, K_0, \alpha_\nu$
Output: $\tilde{\mathbf{W}}_{C2}^{(l)}$ and $\mathbf{K}_l$
Initialization :
$\mathbf{K}_l = \mathbf{0}_{2L\times M}$
$S = \{0, \cdots, 2L-1\} \times \{0, \cdots, M-1\}$
for $j = 1 : \lfloor \alpha_\nu K_0 \rfloor$ **do**
  $(i^*, f^*) = \arg\max_{(i,f)\in S} \left(|d_{i,f}^{(l)}|\right)$
  $p_{l,i^*,f^*}^{(2)} = Q_{SB}(d_{i^*,f^*}^{(l)})$
  $\varphi_{l,i^*,f^*} = Q_P(d_{i^*,f^*}^{(l)})$
  $k_{l,i^*,f^*}^{(3)} = 1$
  $d_{i^*,f^*}^{(l)} = 0$
end for

C. PRECODER SELECTION STRATEGY IN THE DELAY DOMAIN

Fig. 7 shows the general procedure of the low-complexity selection strategy for the eType-II codebook in the delay-domain. The precoding matrix $\mathbf{W}^{(l)}$ for the l-th layer is written as the product of:

$$\mathbf{W}^{(l)} = \mathbf{W}_1\tilde{\mathbf{W}}_2^{(l)}\mathbf{W}_f^{(l)H}\Lambda^{(l)} \tag{72}$$

where $\mathbf{W}_1 \in \mathbb{C}^{2N_1N_2\times 2L}$ is the beam matrix, $\tilde{\mathbf{W}}_2^{(l)} \in \mathbb{C}^{2L\times M_\nu}$ is the compressed BCC matrix comprising coefficients in the angle-delay domain, $\mathbf{W}_f^{(l)H} \in \mathbb{C}^{M_\nu\times N_3}$ is the FD compression matrix and $\Lambda^{(l)} \in \mathbb{C}^{N_3\times N_3}$ is the normalization matrix. The selection strategy is elaborated in the following four steps:

Step 1. **Construction of beam matrix $\mathbf{W}_1$:** The delay domain selection strategy shares the same beam matrix structure as Type-II Codebook, therefore, the L IDFT beams are chosen based on Algorithm 1.

Step 2. **Selection of delays and construction of FD compression matrix $\mathbf{W}_f^{(l)}$:** We first determine the time domain representation $\check{\mathbf{H}}_A(m)$ of $\mathbf{H}_A(t)$ by calculating the IDFT of $\mathbf{H}_A(t)$ over all N_3 SBs. Let $\mathbf{U} \in \mathbb{C}^{N_3\times N_3}$ be the unitary DFT matrix and $\mathbf{U}_{N_R} = \mathbf{U}\otimes\mathbf{I}_{N_R} \in \mathbb{C}^{N_RN_3\times N_RN_3}$. We obtain

$$\begin{bmatrix} \check{\mathbf{H}}_A(0) \\ \check{\mathbf{H}}_A(1) \\ \vdots \\ \check{\mathbf{H}}_A(N_3-1) \end{bmatrix} = \mathbf{U}_{N_R}^H \begin{bmatrix} \mathbf{H}_A(0) \\ \mathbf{H}_A(1) \\ \vdots \\ \mathbf{H}_A(N_3-1) \end{bmatrix} \tag{73}$$

Based on the stacked matrix comprising $\check{\mathbf{H}}_A(m)$ for all N_3 delays, the M_ν strongest delays among the N_3 delays are selected such that the power is maximized across the $2L$ beams and N_R receive antennas following Algorithm 4. Based on the selected M_ν delays, the FD matrix $\mathbf{W}_f^{(l)H}$ is constructed. It should be noted that compared to the FD approach described in Section VI-B, the obtained FD matrix $\mathbf{W}_f^{(l)H}$ is independent of layer index (l) which provides room for additional feedback reduction.

Algorithm 4 Delays Selection

Input: $\check{\mathbf{H}}_A(m), M_\nu, N_3$
Output: $\tau_0, \tau_1, \cdots, \tau_{M_\nu-1}$
$\mathcal{N}_t = \{0, \cdots, N_3-1\}$
for $i = 0 : M_\nu - 1$ **do**
  $m^* = \arg\max_{m\in\mathcal{N}_t}\{tr(\check{\mathbf{H}}_A^H(m)\check{\mathbf{H}}_A(m))\}$
  $\tau_i = m^*$
  $\mathcal{N}_t = \mathcal{N}_t \setminus m^*$
end for

Step 3. **Calculation of Compressed BCC matrix $\tilde{\mathbf{W}}_2^{(l)}$:** The objective of this step is to find the coefficients of the matrix $\tilde{\mathbf{W}}_2^{(l)}$. Let us define the matrix $\mathbf{K} \in \mathbb{C}^{N_R\times 2M_\nu L}$ as

$$\mathbf{K} = \left[\check{\mathbf{H}}_A(\tau_0) \cdots \check{\mathbf{H}}_A(\tau_{M_\nu-1})\right]. \tag{74}$$

Its associated covariance matrix $\mathbf{C} \in \mathbb{C}^{2M_\nu L\times 2M_\nu L}$ is defined as

$$\mathbf{C} = \mathbf{K}^H\mathbf{K}. \tag{75}$$

The precoder for the l-th layer is obtained by the SVD on $\mathbf{C}$. The first ν singular vectors corresponds to the precoder for the ν layers. Let us define the l-th singular vector as $\boldsymbol{u}^{(l)}$, then the compressed BCC matrix in the delay domain is obtained by arranging the $2LM_\nu$ entries of $\boldsymbol{u}^{(l)}$ into the $2L \times M_\nu$ sized matrix $\tilde{\mathbf{W}}_2^{(l)}$, i.e., $\mathbf{u}^{(l)} = \text{vec}\left(\tilde{\mathbf{W}}_2^{(l)}\right)$. For example, for

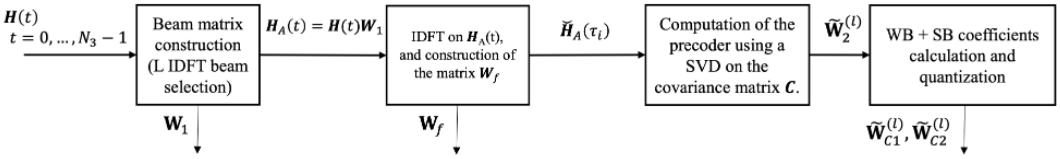


FIGURE 8. Flowchart showing the steps involved in the selection of the eType-II precoder using DD strategy.

$M_v = 2$, the first $2L$ entries of $\boldsymbol{u}^{(l)}$ corresponds to the first delay $m = 0$ and the remaining $2L$ entries corresponds to the second delay $m = 1$. The rationale behind this strategy is given in the Appendix.

Step 4. **Calculation of wideband and SB coefficients and quantization**: The structure of the non-quantized matrix $\tilde{\mathbf{W}}_2^{(l)}$ is the same as the one obtained with the FD strategy. Therefore, the matrix $\tilde{\mathbf{W}}_2^{(l)}$ can be written as $\tilde{\mathbf{W}}_2^{(l)} \approx \tilde{\mathbf{W}}_{C1}^{(l)} \tilde{\mathbf{W}}_{C2}^{(l)}$. The matrices with quantized entries $\tilde{\mathbf{W}}_{C1}^{(l)}$ and $\tilde{\mathbf{W}}_{C2}^{(l)}$ are obtained by using step 5 and 6 of the eType-II codebook FD selection strategy.

D. FEEDBACK QUANTITY ANALYSIS

The CSI report containing the PMI of the eType-II codebook consist of two parts similarly to type-II. The first part (referred to as i_1 in the standard) comprises the following indicators: a $\lceil \log_2(O_1 O_2) \rceil$-bit indicator to indicate the selected rotation factor among $O_1 O_2$ groups referred to as $i_{1,1}$, a $\lceil \log_2 \binom{N_1 N_2}{L} \rceil$-bit indicator to indicate the selected L IDFT beams from $N_1 N_2$ IDFT beams referred to as $i_{1,2}$, a $\left\lceil \log_2 \binom{N_3-1}{M_v-1} \right\rceil$-bit indicator to indicate the selected M_v delays from N_3 delays for each l-th layer referred to as $i_{1,6,l}$, a $\lceil \log_2(K^{NZ}) \rceil$-bit indicator to indicate the strongest coefficient for $v = 1$ and a $\lceil \log_2(2L) \rceil$-bit indicator to indicate the strongest coefficient for each l-th layer for $v > 1$ referred to as $i_{1,8,l}$, a $2LM_v$-bit indicator to indicate the NZC location for each l-th layer referred to as $i_{1,7,l}$. The second part (referred to as i_2 in the standard) comprise the following indicators: a 4-bit indicator to indicate the polarization specific wideband amplitude for each l-layer referred to as $i_{2,3,l}$, a $3 \cdot (K^{NZ} - v)$-bit indicator ($i_{2,4,l}$) and a $4 \cdot (K^{NZ} - v)$-bit indicator referred to as $i_{2,5,l}$ to indicate the SB amplitude and phase, respectively, of the K^{NZ} NZCs. Note that for $N_3 > 19$, the M_v delays for each l-th layer are selected from a layer common window comprising $2M_v$ adjacent delays instead of N_3 delays and are indicated by a $\left\lceil \log_2 \binom{2M_v-1}{M_v-1} \right\rceil$-bit indicator referred to as $i_{1,6,l}$. As the window is common for all v layers, the starting index of the window among N_3 indices is reported by a $\lceil \log_2(2M_v) \rceil$-bit indicator referred to as $i_{1,5}$. Note that since the strongest delay is the zero delay by construction as explained in Step. 5, the zero delay is always selected and is part of the common window which explains the bit size of $i_{1,6,l}$ and $i_{1,5}$. Please refer to the release [46] for more information on the feedback overhead indicators. Note that the standard does not allow optimizing the feedback when the FD compression matrix is the same for all layers (which can be the case for the DD strategy).

Precoder feedback overhead example: We compare the feedback overhead of the Type-I, Type-II and eType-II codebooks for a rank transmission. The simulation parameters are given in Table 3. As the number of FD basis is identical for all layers, for $\rho_v = 1/4$, the number of FD basis per layer is given by 3 i.e., $M_1 = M_2 = 3$. For $K_0 = 18$, ($K_l^{NZ} = \lfloor \alpha_v K_0 \rfloor$ where $\alpha_v = 1$ for $v = 1$ and $\alpha_v = 2/v$ for $v = 2, 3, 4$), the number of NZCs per layer $K_l^{NZ} = 18$ and results in a total of 36 NZCs across both layers. For $N_3 = 10$, the feedback overhead of Type-I codebook is only 20 bits, whereas the feedback overhead of type-II and eType-II codebooks is 543 bits and 327 bits, respectively. For a given L and M_v values, note that the feedback overhead of eType-II codebook increases with increasing K^{NZ} value. Among all three codebooks, it is obvious that Type-I results in the lowest feedback overhead however with a significant performance degradation compared to Type-II and eType-II codebooks.

E. COMPLEXITY ANALYSIS

For the BCC matrix calculation using the FD strategy, a single SVD is performed on the complex matrix of size $2L \times 2L$ of each subband, where the maximum rank is given by $\min(N_R, 2L)$. For $N_R \leq 2L$, the time complexity (TC) and space complexity (SC) in terms of total number of floating point operations (FLOPs) required to compute N_3 SVDs according to the truncated SVD approach [51] is approximately given by $O(N_3 \cdot X)$ and $O(N_3 \cdot Y)$, respectively. Here, $X = 4 \cdot (4L(N_R)^2 + (N_R)^3 + N_R + 2LN_R)$ and $Y = 4 \cdot (3(N_R)^2 + 3N_R + 4LN_R)$. In contrast, for the BCC matrix calculation using the DD strategy, a single SVD is performed on a complex matrix of size $2LM_v \times 2LM_v$ across all N_3 subbands, where the maximum rank is given by $\min(N_R, 2LM_v)$. For $N_R < 2LM_v$, the TC and SC in terms of total number of FLOPs required to compute a single SVD according to the truncated SVD approach [51] is approximately given by $O(4 \cdot (4LM_v(N_R)^2 + (N_R)^3 + N_R + 2LM_v N_R))$ and $O(4 \cdot (3(N_R)^2 + 3N_R + 4LM_v N_R))$, respectively.

For $(L, N_3, N_R) = (4, 10, 4)$, the total number of FLOPs for the FD strategy are given by $(TC, SC) = (14240, 4860)$, whereas for the DD strategy, for $(L, N_3, M_v, N_R) = (4, 10, 4, 4)$, the total number of FLOPs are given by $(TC, SC) = (4880, 1264)$. It can be observed that the TC and SC of the DD strategy represents approximately a three-fold

and four-fold reduction in the computational complexity compared to the FD strategy.

VII. LINK ADAPTATION PROCESS

Low complexity prediction methods of PHY performance, typically, block error rate (BLER) performance for a given SINR, are coined PHY abstractions. PHY abstraction allows avoiding the computationally complex simulations of physical layers functionalities such as equalization, demodulation and decoding. They are typically used for system level simulations or in the context of link adaptation for the determination of the PMI, RI and CQI at the UE level in order to maximize the achievable data rate for a given channel realization and target BLER [52], [53] [54], [55].

The first abstraction approach for the computation of the Effective γ is the Exponential Effective γ Mapping (EESM) [56], which is commonly used for linear receivers. The effective γ is obtained by performing the following non-linear averaging or compression of the P SB γ:

$$\gamma_{eff} = -\beta \ln\left(\frac{1}{P}\sum_{p=1}^{P} e^{-\frac{\gamma_p}{\beta}}\right) \tag{76}$$

where β is calibrated by means of link level simulations to fit the compression function to the AWGN BLER simulation results.

The second main abstraction approach is the Mutual Information Effective γ Mapping (MIESM) as described in [52], [57], [58], and [59]. The MIESM technique allows to compress the multiple γs into a single effective γ as follows:

$$\gamma_{eff} = I_Q^{-1}\left(\frac{1}{P}\sum_{p=1}^{P} I_Q(\gamma_p)\right) \tag{77}$$

where $I_Q(SNR)$ is the mutual information discrete-entries (modulation Q dependent) for an AWGN channel.

Using PHY abstraction, link level simulation can be divived into two phases:

- phase one: selection of the optimal precoding matrices over the different SBs as described in the previous sections.
- second phase: using the previously computed γ with the selected precoding matrices, selection of the MCS, RI and computation of the achievable throughput.

The algorithm for the determination of the MCS, RI and throughput in the second phase is given in Algorithm 5.

We have used the MCS index Table 2 for PDSCH (from table 5.1.3.1-2 of [5] up to QAM256) to compute the throughput. In the simulation results, the BLER target for the first transmission is fixed to 0.1.

3GPP defines two types of HARQ for LTE: Chase Combining (CC) and Incremental redundancy (IR). In this work, we will only consider Chase combining (CC). Chase Combining consists of retransmitting the same information and the receiver uses maximum ratio combining to combine the erroneous packet with the other retransmissions.

Algorithm 5 Algorithm to Select the *MCS* and Rank Indicator ν

Input: $\gamma_{s,j}^{\nu,t}$ $\forall(s,j) \in \{1,\cdots,\nu\} \times \{0,\cdots,N_3-1\}$
Output: $MCS^\star$, $\nu^\star$ where $Q^\star$ is the modulation size of $MCS^\star$

for $\nu = 1 : \nu_{max}$ **do**
  for $Q \in$ (QPSK, 16QAM, 64QAM, 256QAM) **do**
    Set $I_{Q,\nu} = 0$
    for $t = 0 : N_3 - 1$ **do**
      for $j = 0 : N_{SB} - 1$ **do**
        for $l = 1 : \nu$ **do**
          Calculate $I_{Q,\nu} = I_{Q,\nu} + I_Q(\gamma_{l,j}^{\nu,t})$
        end for
      end for
    end for
    $\gamma_{eff}^{\nu,Q} = I_Q^{-1}\left(\frac{1}{N_3 N_{SB}} I_{Q,\nu}\right)$
  end for
end for
Select $MCS^\star$, $\nu^\star$ to maximize the transmission rate while achieving $BLER_{MCS^\star}(\gamma_{eff}^{\nu^\star,Q^\star}) \leq 10\%$

Assuming that the radio channel is quasi-static and does not change between two retransmissions, the rank indicator and the precoding matrix remain unchanged. As a result, the SINR (per layer) of the retransmitted packet is identical to the SINR of the first transmission. In dB, the resulting SINR (after N_{RT} retransmission) after Chase combining for a given RI and precoding matrix, can be deduced according to the following relation:

$$\gamma_{l,j}^{\nu^\star,t} = \quad \gamma_{l,j}^{\nu^\star,t}(\text{first tx}) + 10\log_{10}(N_{RT}+1). \tag{78}$$

In this work, the maximum number of retransmissions is fixed at $N_{RT}\max$=3.

VIII. MONTE CARLO SIMULATION RESULTS

A first objective of this Section is to compare the link level performance of the different MIMO codebooks from Rel. 15 to Rel. 16. A second objective is to evaluate the performance of the FD and DD strategies as described earlier for the precoder selection for the Rel. 16 eType-II codebook. Obviously, the DD strategy is equivalent to the FD strategy for a non selective channel and presents the advantage of a much lower complexity. However, the respective behaviors of the two approaches for frequency selective and spatially correlated channels is not obvious and necessitates to resort to Monte Carlo simulations.

A. SIMULATION SETTING

The main simulation parameters are given in Table 3. We assume Perfect CSI at the receiver. We adopt a quasi-static channel assumption in which the channel does not change during a TTI but changes independently from one TTI to the other. It models low/medium mobility scenarios, typically, 3 km/h. The carrier frequency is 3.5 GHz and the

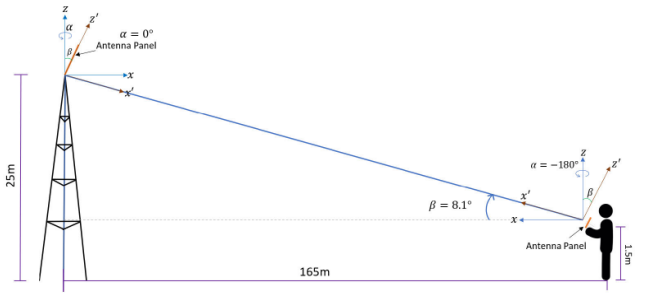


FIGURE 9. Illustration of the antenna panel orientation and the separation between the BS and the UE.

transmission bandwidth is 10 MHz or 50 PRBs with 5-PRBs in each SB. For each channel realization, Algorithm 5 is used to derive the MCS, PMI and rank of the transmission for a first transmission BLER target of 0.1. Based on the selected CSI parameters per channel realization, the normalized throughput or efficiency (bit per channel use) (bpcu) is averaged over 2000 channel realizations assuming a maximum of 3 chase re-transmissions. Moreover, an LMMSE detector is used at the receiver. The eType-II codebook parameter configuration is set to $paramCombination - r16 = 5$ for enhanced Type-II codebook, according to Table 5.2.2.2.5-1 in [46], $L = 4$, $p_v = 1/4$ and $\beta = 3/4$. The Type-II codebook parameter are $numberOfBeams$=4 (L=4), $phaseAlphabetSize$=8 (the phase per SB is coded over 3 bits or $N_{PSK} = 8$), $subbandAmplitude$='true' (the subband amplitude is reported). The Type-I for rank 1 and 2 is based on $codebookMode = 1$. Finally, for the link adaptation performance results no rank restriction and no subset restriction are considered, i.e., the maximum rank is 4. We assume that the BS is composed of a single panel with 96 antennas elements ($M = 12$ vertically $N = 8$ horizontally) per polarization. The inter-distance between antenna elements is $d_H = 0.5\lambda$ and $d_V = 0.5\lambda$ in the horizontal and vertical dimensions, respectively. Each of the 6 antenna elements in the vertical dimension are combined to form an antenna port resulting in the antenna port configuration $(N_1, N_2) = (8, 2)$. At the UE side, the dual polarized antenna port configuration considered is $(N_1, N_2) = (2, 1)$ with $d_H = 0.5\lambda$ where one antenna port is associated to an omni-directional single antenna element. The antenna port configuration considered at the BS and UE are are defined in 3GPP TS 38.901. The antenna panel orientation and the separation between the BS and the UE are depicted in Fig. 9. The height of the BS and the UE are 25 m and 1.5 m, respectively and are separated by a distance of 165 m. Our setting aims at having the two panels face each other which corresponds to a down-tilt of $\beta = 8.1°$ at the BS and a rotation of $\alpha = -180°$ around the z-axis and a inclination of $\beta = -8.1°$ at the UE.

B. MONTE CARLO SIMULATIONS WITH LINK ADAPTATION

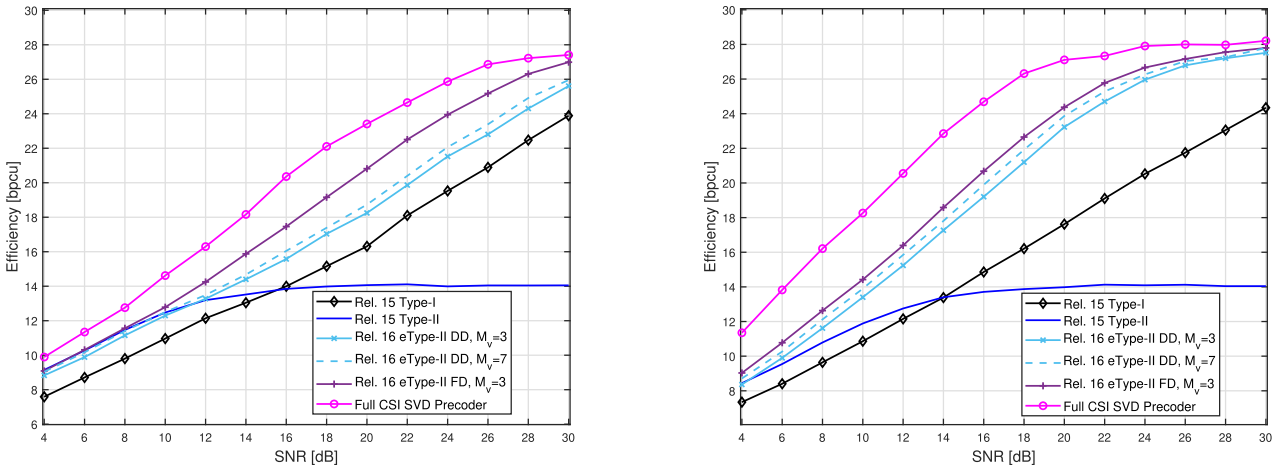
In this Section, we consider the link adaptation as defined in Algorithm 5. We compare the efficiency performance of Type-I, Type-II and eType-II codebooks using both FD and DD strategies. As a benchmark, we consider the spectral efficiency obtained using the uncompressed precoder resulting from the SVD of $\mathbf{H}(t)$ per SB which minimizes the inter-layer interference at the receiver side. In Fig. 10, it can be observed that the Rel. 16 eType-II codebook clearly outperforms the Rel. 15 Type-I codebook in terms of efficiency for both CDL-A and CDL-C channels. It can be also observed that the Rel. 15 Type-II codebook outperforms the Type-I codebook only for lower SNRs. It comes from the fact that the maximum supported rank is 2 for the Rel. 15 Type-II codebook, as a result, the efficiency is limited to 14.05 bpcu at higher SNRs. At lower SNRs, we can observe that Rel. 15 Type-II codebook has a similar performance as Rel. 16 eType-II codebook however at the expense of an higher feedback overhead. Comparing the FD and DD strategy for eType-II codebook, the FD strategy outperforms the DD strategy by 2.6 dB and 0.7 dB SNR gain for CDL-A and CDL-C channels at 20 bpcu efficiency for $M_v = 3$, respectively. Nevertheless, the DD strategy performs close to the FD strategy for CDL-C channels. Therefore, based on Fig. 10, we can deduce that the DD strategy suffers at high SNR for CDL-A. One reason could be that the precoder based on DD strategy does not mitigate the interference sufficiently between the spatial layers especially for high rank transmission. Moreover, it should be noted that increasing the delays from $M_v = 3$ to $M_v = 7$ for the DD strategy enhances the SNR only by 0.5 dB for both channels.

C. MONTE CARLO SIMULATIONS WITHOUT LINK ADAPTATION

In order to analyse the impact of the interference between the spatial layers for the DD and FD strategies, we compare their performance with Monte Carlo simulations without link adaptation. For this, we consider a fixed rank and MCS for all the channel realizations. To understand the performance behavior of the FD and DD strategy with respect to CDL-A and CDL-C channel models, we evaluate the full CSI SVD precoder with power control, where the power per layer per SB is optimized with respect to the Eigen values of $\mathbf{H}(t)$ following a water filling approach. Additionally, the wideband rank is adapted following Algorithm 5 keeping the MCS fixed to MCS 19. In addition, we also consider the aforementioned FD and DD strategy using a Genie-Aided (GA) receiver. The precoder selection algorithm is not changed but the receiver at the UE is considered to ideally cancel the interference between the spatial layers. The GA receiver is provided by a 'genie' the spatial layer interference which allows the receiver to perfectly cancel it. For all the other performance curves a Linear MMSE (LMMSE) receiver is always considered which mitigates the interference based on its covariance (the LMMSE can be viewed as the concatenation of a whitening filter on the interference plus noise covariance followed by a matched filter). The MCS is always fixed to 19 which corresponds to a 64-QAM modulation and a coding rate of 0.85. The number of delays is

TABLE 3. Simulation parameters.

Number of horizontal co-polar antenna ports at the BS	$N_1 = 8$
Number of vertical co-polar antenna ports at the BS	$N_2 = 2$
Number of DFT oversampling in the first dimension	$O_1 = 4$
Number of DFT oversampling in the second dimension	$O_2 = 4$
Number of antenna ports of the UE	$N_R = 4$
Number of SB	$N_3 = 10$
Number of PRBs per SB	$N_{SB} = 5$
Mechanical Downtilt	$\beta = 8.1°$
Modulation and coding scheme (MCS)	MCS table $5.1.3.1 - 2$ in [5]
Retransmission type	Chase
Max. number of retransmissions	N_{RT}max=3
Channel model	3GPP CDL-A and CDL-C
Detector at the receiver	LMMSE
eType-II codebook parameters	*paramCombination-r16*=5 ($L = 4$, $p_v = 1/4$, $\beta = 3/4$)
Type-II codebook parameters	*numberOfBeams*=4 *phaseAlphabetSize*=8 *subbandAmplitude* ='true'
Type-I codebook parameters	*codebookMode*=1



(a) CDL-A (b) CDL-C

FIGURE 10. Performance of Type-I, Type-II and eType-II codebooks with link adapatation.

fixed to $M_v = 3$ for the FD strategy while for the DD strategy both $M_v = 3$ and $M_v = 7$ are considered. The efficiency curves of Type-I, Type-II and eType-II codebooks using FD and DD strategy with and without GA receiver and full CSI SVD precoder are shown in Fig. 11 for a rank 2 transmission and CDL-A channel model. It can be observed that eType-II codebooks clearly outperforms the Type-I codebook. For an efficiency value of 6 bpcu, the Type-I codebook suffers with 3.4 dB SNR loss compared to the eType-II codebook FD strategy. The efficiency of the DD and FD selection strategies for the eType-II codebook are very close with a difference of only 0.35 dB for $M_v = 3$. For $M_v = 7$, the GA receiver with the DD strategy improves the SNR by 0.9 dB compared to the DD strategy without GA receiver. On the other hand, the GA receiver results in an SNR gain of only 0.3 dB for the FD strategy.

The efficiency of the aforementioned codebooks for a rank 3 transmission using CDL-A channels is shown in Fig. 12. For an efficiency of 10 bpcu, it can be observed that the FD strategy outperforms the DD strategy by 1.9 dB for $M_v = 3$. Moreover, the SNR difference between the FD and DD strategy widens compared to the SNR difference of rank 2. Interestingly, for an efficiency of 10 bpcu, the GA receiver with the DD strategy achieves an SNR gain of 6.6 dB with respect to the DD strategy using LMMSE receiver for $M_v = 7$.

The efficiency of the aforementioned codebooks for a rank 4 transmission using CDL-A channels is shown in Fig. 13.

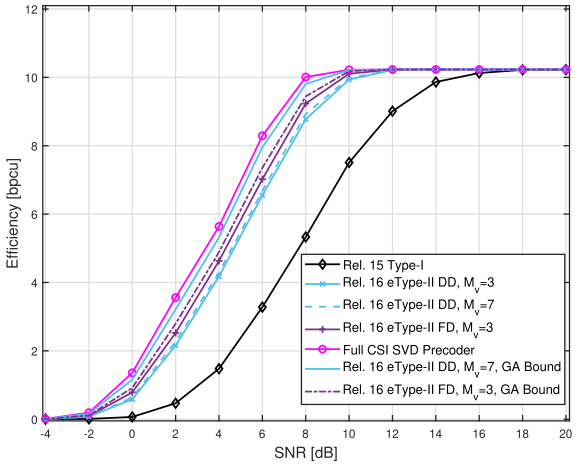


FIGURE 11. eType-II FD vs DD, CDL-A and rank 2.

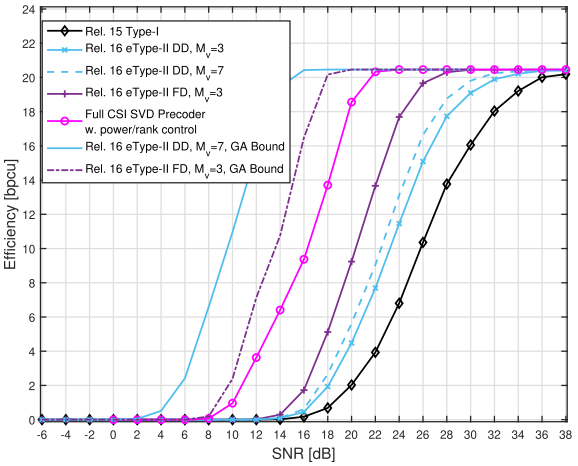


FIGURE 13. eType-II FD vs DD, CDL-A and rank 4.

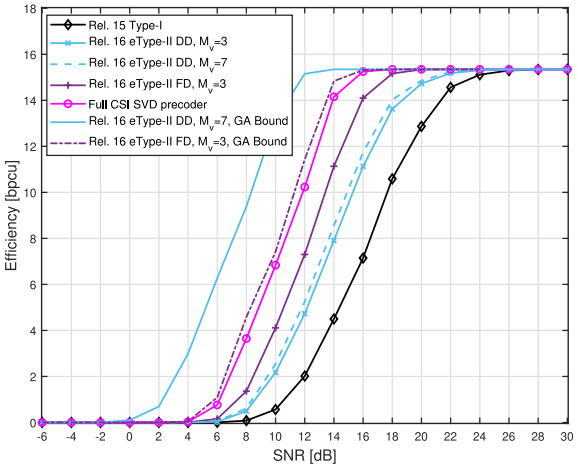


FIGURE 12. eType-II FD vs DD, CDL-A and rank 3.

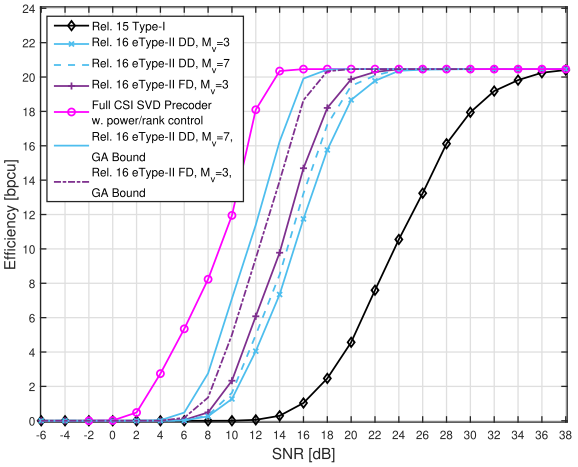


FIGURE 14. eType-II FD vs DD, CDL-C and rank 4.

For an efficiency of 14 bpcu, it can be observed that the FD strategy outperforms the DD strategy by 3.2 dB for $M_v = 3$. Similar to rank 2, the SNR difference between the FD and DD strategy widens compared to the SNR difference of rank 3. Interestingly, for an efficiency of 14 bpcu, the GA receiver for the DD strategy improves the performance by a staggering 13.2 dB with respect to the LMMSE receiver for $M_v = 7$. Therefore, it can be concluded that for rank 4 SU-MIMO transmission for the CDL-A channels, the DD precoder is not able to deal properly with the inter layer spatial interference. Form the above efficiency figures, the SNR gap widens with increasing rank. Therefore, it can be concluded that the interference plays a significant role in lowering the efficiency of the DD strategy for CDL-A channels

The efficiency of the aforementioned codebooks for a rank 4 transmission using CDL-C channels is shown in Fig. 14. The SNR difference between the FD and DD strategies is much lower than for the CDL-A channel. For an efficiency of 14 bpcu, the difference between the FD and DD strategy performance is only 1.4 dB for $M_v = 3$. Moreover, the GA receiver improves the performance of the DD for $M_v = 7$ and FD strategy for $M_v = 3$ by 3.3 dB and 1.7 dB, respectively. Based on the comparison between the full CSI SVD precoder and full CSI SVD precoder with power and rank control in Fig. 13 and Fig. 14, it can be concluded that CDL-A is more spatially correlated than CDL-C. Indeed, it can be observed that rank and power control together with a full CSI SVD precoder brings more gain for CDL-A channels than for CDL-C channels. It means that the eigenvalues of the channel per SB are more spread in average for CDL-A channels than for CDL-C channels. As result, it is observed that water filling in the spatial domain and rank adaptation is more beneficial for CDL-A to improve the spectral efficiency for a fixed MCS.

Note that the DD strategy performance would benefit more than the FD one from advanced receiver with Interference Cancellation (IC) capability at the UE. Their comparison under the assumption of IC receiver is an interesting topic for future study.

IX. MULTI-USER MIMO

Multi-user (MU)-MIMO is able to solve the problem of insufficient bandwidth in MIMO transmission as the BS transmits signals to multiple users using the same frequency resources simultaneously. In the downlink MU-MIMO scenario, transmission is broadcast and the challenge is to mitigate inter-user interference. Optimal precoding in MU-MIMO involves the interference subtraction technique also called dirty paper coding (DPC) [60] associated to user scheduling and power loading algorithm. However, due to its data dependency, DPC is too complicated to be implemented. Therefore, linear beamforming for downlink MU-MIMO transmission has attracted substantial interest. Regularized channel inversion proposed in [61] chooses beamforming vectors based on the inverse of the multiuser channel at the transmitter but only supports the UEs equipped with one receive antenna. Block diagonalization strategy [62], [63] is able to support multi-antenna UEs and multi-layer (data stream) MU-MIMO transmission while is limited by the constraint that the number of receive antennas is equal to the number of data stream. Coordinated beamforming where precoding and decoding process are applied at the BS and UEs respectively has much simpler structure and accommodates more general MU-MIMO scenarios [64]. In this section, we focus on the coordinated beamforming techniques for MU-MIMO which is also compatible with the standardized codebooks. An iterative approach named nullspace-directed SVD (Nu-SVD) is proposed in [65] for precoder and receiver design in downlink MU-MIMO transmission. In [66], the authors introduce a decomposition-based method to devise linear precoder and receiver where the precoder $\mathbf{P}$ is decomposed into $\mathbf{P} = \mathbf{P}_J\mathbf{P}_S$ with joint part $\mathbf{P}_J$ and separate part $\mathbf{P}_S$. The joint part can be obtained by either block-diagonalization method or successive LMMSE method while the separate part and receiver matrix are dependent on the matrix $\mathbf{P}_J$.

We can easily extend the downlink SU-MIMO transmission given in receive signal SU to a downlink MU-MIMO scenario. The same 2D antenna deployment with N_1 cross-polarization horizontal antenna ports and N_2 cross-polarization vertical antenna ports is applied at the BS. We consider a MU-MIMO transmission with K co-scheduled UEs. Please note that the step of user selection (scheduling) is not considered in our MU-MIMO transmission. Every co-scheduled UE is equipped with N_R antenna ports. In such MU-MIMO transmission, the received signal at the k-th UE is represented as

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{P}_k\mathbf{x}_k + \mathbf{H}_k \sum_{\substack{1\leqslant j\leqslant K \\ j\neq k}} \mathbf{P}_j\mathbf{x}_j + \mathbf{n}_k \tag{79}$$

where $\mathbf{H}_k \in \mathbb{C}^{N_R\times 2N_1N_2}$, $\mathbf{P}_k \in \mathbb{C}^{2N_1N_2\times\nu_k}$, $\mathbf{x}_k \in \mathbb{C}^{\nu_k}$ and $\mathbf{n}_k \in \mathbb{C}^{N_R}$ are the channel matrix, transmit precoder matrix, transmit signal vector and noise vector of the k-th co-scheduled UE respectively. $\mathbf{P}_j \in \mathbb{C}^{2N_1N_2\times\nu_j}$ and $\mathbf{x}_j \in \mathbb{C}^{\nu_j}$ are the transmit precoder matrix and transmit signal vector of the j-th UE respective. Summed transmission signal of all the UE j ($1 \leqslant j \leqslant K, j \neq k$), the second term $\mathbf{H}_k \sum_{\substack{1\leqslant j\leqslant K \\ j\neq k}} \mathbf{P}_j\mathbf{x}_j$, is taken as the multiuser interference of the k-th UE. With receiver matrix $\mathbf{F}_k^{\mathsf{H}} \in \mathbb{C}^{N_R\times\nu_k}$, the output of the k-th UE LMMSE receiver is given by

$$\begin{aligned}\hat{\mathbf{x}}_k &= \mathbf{F}_k^{\mathsf{H}}\mathbf{y}_k \\ &= \mathbf{F}_k^{\mathsf{H}}\mathbf{H}_k\mathbf{P}_k\mathbf{x}_k + \mathbf{F}_k^{\mathsf{H}}\mathbf{H}_k \sum_{\substack{1\leqslant j\leqslant K \\ j\neq k}} \mathbf{P}_j\mathbf{x}_j + \mathbf{F}_k^{\mathsf{H}}\mathbf{n}_k.\end{aligned} \tag{80}$$

Design of the MU-MIMO transmit precoder matrix and receiver matrix aims at mitigating the second term on the right of Eq. (80).

In the following, we illustrate the design of the receiver matrix $\mathbf{F}_k$ and the precoder matrix $\mathbf{P}_k$ $k \in \{1,\cdots,K\}$ for the classical Zero Forcing (ZF) criterion. Let us define the received vector $\mathbf{y}$ obtained by stacking the K vectors $\mathbf{y}_k$, $\mathbf{y} = \left[y_1^{\mathsf{T}} \cdots y_K^{\mathsf{T}}\right]^{\mathsf{T}} \in \mathbb{C}^{KN_R}$ and the overall precoder $\mathbf{P} = \left[\mathbf{P}_1 \cdots \mathbf{P}_K\right] \in \mathbb{C}^{2N_1N_2\times\nu}$ with $\nu = \sum_{i=1}^{K}\nu_i$ the total number of spatial layers. The received signal can be written as

$$\mathbf{y} = \begin{bmatrix}\mathbf{H}_1 \\ \vdots \\ \mathbf{H}_K\end{bmatrix}\left[\mathbf{P}_1 \cdots \mathbf{P}_K\right]\begin{bmatrix}\mathbf{x}_1 \\ \vdots \\ \mathbf{x}_K\end{bmatrix} + \begin{bmatrix}\mathbf{n}_1 \\ \vdots \\ \mathbf{n}_K\end{bmatrix} \tag{81}$$

Each channel matrix $\mathbf{H}_k$ can be factorized using a singular value decomposition as $\mathbf{H}_k = \mathbf{U}_k\Sigma_k\mathbf{V}_k^{\mathsf{H}}$ where $\mathbf{U}_k \in \mathbb{C}^{N_R\times N_R}$ is a unitary square matrix, $\Sigma_k = \mathrm{diag}(\sqrt{\lambda_{k,1}},\cdots,\sqrt{\lambda_{k,N_R}})$ is a square diagonal matrix of dimension N_R containing the ordered real positive eigenvalues such that $\sqrt{\lambda_{k,1}} \geq \sqrt{\lambda_{k,2}} \geq \cdots \geq \sqrt{\lambda_{k,N_R}}$, and $\mathbf{V}_k \in \mathbb{C}^{2N_1N_2\times N_R}$ is a unitary matrix. It yields Eq. (82), as shown at the bottom of the next page. The ZF is applied on the matrix $\mathbf{Z} = \begin{bmatrix}\mathbf{V}_1^{\mathsf{H}} \\ \cdots \\ \mathbf{V}_K^{\mathsf{H}}\end{bmatrix}$. It consists in pseudo inverting it. Let $\mathbf{M}^{ZF}$ be the pseudo inverse of $\mathbf{Z}$ such that $\mathbf{M}^{ZF} = \mathbf{Z}^{\mathsf{H}}\left(\mathbf{Z}\mathbf{Z}^{\mathsf{H}}\right)^{-1} = \left[\mathbf{M}_1^{ZF} \cdots \mathbf{M}_K^{ZF}\right]$. The precoding $\mathbf{P}_k$ corresponds to the ν_k first columns of $\mathbf{M}_k^{ZF}$, i.e., $\mathbf{P}_k = \left[\mathbf{m}_{k,1}^{ZF} \cdots \mathbf{m}_{k,\nu_k}^{ZF}\right] \in \mathbb{C}^{2N_1N_2\times\nu_k}$. By construction, the matrix $\mathbf{ZP} \in \mathbb{C}^{KN_R\times\nu}$ is an identity matrix $\mathbf{I}_{\nu\times\nu}$ to which is added $N_R - \nu_k$ null rows at each position $\sum_{i=1}^{k}\nu_i$ for all $k \in \{1,\cdots,K\}$. By selecting $\mathbf{F}_k$ as the first ν_k columns of $\mathbf{U}_k$, i.e., $\mathbf{F}_k = \mathbf{U}_k\begin{bmatrix}\mathbf{I}_{\nu_k\times\nu_k} \\ \mathbf{0}_{N_R-\nu_k\times\nu_k}\end{bmatrix}$, Eq. (80) becomes

$$\hat{\mathbf{x}}_k = \mathrm{diag}(\sqrt{\lambda_{k,1}},\cdots,\sqrt{\lambda_{k,\nu_k}})\mathbf{x}_k + \mathbf{w}_k \tag{83}$$

where $\mathbf{w}_k = \mathbf{F}_k^{\mathsf{H}}\mathbf{n}_k$ is a spatially white noise vector. Since the columns of the type-II precoder feedback by a given UE k can be considered as a good approximation of the first ν_k columns of $\mathbf{V}_k$ (corresponding to the best eigenvalues of the channel), the MU-MIMO precoder can assume that $\mathbf{Z}$ results from the stacking of the different precoders of the associated UEs. Note that type-I precoder is considered as not accurate enough for that approximation. Finally, the ZF power constraint $\mathrm{tr}\{\mathbf{PQP}^{\mathsf{H}}\} \leq P_{max}$, where $\mathbf{Q}$ is the

diagonal matrix made of the spatial layer powers, may induce the so called noise-enhancement that may be tempered by a regularized ZF version [67]. The performance analysis of the proposed approach of linear precoding for MU-MIMO is out of the scope of the paper. To the authors' knowledge the described MU-MIMO linear precoding approach is the best way to exploit the codebook based partial CSI per UE in a MU-MIMO framework.

X. MIMO PORT SELECTION CODEBOOKS

The Type-II port selection codebook was introduced in Rel. 15 5G NR [5] and has been enhanced in Rel. 16 [46] and further enhanced in Rel. 17 [68]. The PS codebooks until Rel. 16 exploits the reciprocity of the radio channel only in the spatial domain. The BS determines the angle information of dominant channel path components from UL channel measurements based on sounding reference signals (SRSs) and uses it for beamforming the P_{TXRU} TXRUs in the DL. For the Rel. 15 Type-II and Rel. 16 eType-II PS codebooks, each beamformed antenna port is associated with a specific spatial direction. Upon receiving the beamformed reference signals in the DL on $P_{\text{CSI-RS}}$ antenna ports, L consecutive beamformed ports are selected by the UE per polarization. The L selected ports are identical for both polarizations of the $P_{\text{CSI-RS}}$ antenna ports. Contrary to the Regular Type-II codebooks, the matrix $\mathbf{W}_1$ comprises L port-selection vectors per polarization, where the port-selection vector $\boldsymbol{e}_i$ is the i-th column of an identity matrix and selects a single port out of the $P_{\text{CSI-RS}}$ ports. Other than $\mathbf{W}_1$, the precoder calculation and coefficient quantization are kept identical to the regular Type-II CB counterparts of each release. Furthermore, the PS codebooks in Rel. 15 and Rel. 16 supports rank up to two and four, respectively.

A further enhancement of the Rel. 16 Type-II PS codebook was specified in Rel. 17. This codebook exploits the reciprocity of the UL and DL channels in both angle and delay domains of the channels especially in FDD scenarios [69]. The BS extracts the angle and delay information of dominant path components of the radio channel from UL MIMO channel measurements and uses it to beamform the TXRUs or antenna ports in the angle and delay domains. Unlike TDD scenarios, the UL and DL reciprocity may not exist in FDD systems. It was shown in [70] that full reciprocity in the delay domain may not always be satisfied. From the channel sounding measurement results presented in [70] for the 3.75 GHz center frequency and 100 MHz bandwidth, the reciprocity in the angle domain for the UL and DL channels has been observed to hold well, whereas the reciprocity of the delays for the dominant beamforming angles mostly differs by either one or two channel taps. Therefore, it is concluded that only partial reciprocity for the delays can be assumed for the FDD systems.

Assuming full reciprocity of the angles and partial reciprocity of the delays in the UL and DL channels, one possible method of BS beamforming using the angle and delay information determined from the UL channels is described in the following. Furthermore, based on the beamformed CSI-RS, a precoder selection strategy for Rel. 17 FeType-II PS codebook is presented.

A. ANGLE- AND DELAY-BEAMFORMING AT BS

In general, beamforming at the BS is specification-transparent and no particular method is specified by the 3GPP standard. In this contribution, a method for BS-side beamforming exploiting angle and delay information in FDD systems is described for the sake of completeness.

Based on the UL channel measurements at the BS, A_1 dominant angles and D_1 delays per angle are estimated. The estimated angles and delays are mapped to the entries of spatial IDFT codebooks of size $P_{\text{TXRU}} \times O_A P_{\text{TXRU}}$ and delay IDFT codebook of size $N_3 \times O_D N_3$, respectively. Here, O_A and O_D are the oversampling factors of the respective codebooks, P_{TXRU} is the number of TXRU antenna ports across both polarizations and N_3 is the number of SBs. Since BS side beamforming is not specified, high-resolution estimation of the uplink channel can be used and realized by considering higher values for O_A and O_D. The spatial and delay DFT vectors are denoted by $\boldsymbol{v}_i \in \mathbb{C}^{\frac{P_{\text{TXRU}}}{2}\times 1}$ and $\boldsymbol{f}_i(t) = [f_{i,0}(t), \ldots, f_{i,D_1-1}(t)] \in \mathbb{C}^{1\times D_1}$ respectively, where $i = 0, \ldots, A_1 - 1$. $f_{i,j}(t)$ denotes the t-th entry of the i-th spatial j-th delay DFT vectors. Note that for each spatial DFT vector $\boldsymbol{v}_i$, D_1 associated delay DFT vectors are used for beamforming, where D_1 takes any integer value. Note that for simplicity, the angle information is assumed to be identical across both polarizations, whereas the delays can be either identical or different across polarizations. Moreover, the number of delays D_1 are assumed to be the same across all A_1 angles.

To beamform the CSI-RS with the estimated A_1 IDFT vectors associated with the spatial beams and D_1 IDFT vectors associated with the D_1 delays per spatial beam, the CSI reference signals of the t-th SB across all antenna ports are multiplied with matrix $\boldsymbol{v}_i \boldsymbol{f}_i(t) \in \mathbb{C}^{\frac{P_{\text{TXRU}}}{2}\times D_1}, \forall i$. For the t-th SB, the full beamforming matrix $\mathbf{F}_{gNB}(t) \in \mathbb{C}^{P_{\text{TXRU}}\times P_{\text{CSI-RS}}}$ for the two polarizations of the antenna ports can then be written as

$$\mathbf{F}_{gNB}(t) = \begin{bmatrix} \boldsymbol{v}_0 \ldots \boldsymbol{v}_{A_1-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{v}_0 \ldots \boldsymbol{v}_{A_1-1} \end{bmatrix} \begin{bmatrix} \mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{bmatrix} \tag{84}$$

$$\mathbf{y} = \begin{bmatrix} \mathbf{U}_1\Sigma_1 & & 0 \\ & \ddots & \\ 0 & & \mathbf{U}_K\Sigma_K \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^{\mathsf{H}} \\ \vdots \\ \mathbf{V}_K^{\mathsf{H}} \end{bmatrix} \begin{bmatrix} \mathbf{P}_1 \cdots \mathbf{P}_K \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_K \end{bmatrix} + \begin{bmatrix} \mathbf{n}_1 \\ \vdots \\ \mathbf{n}_K \end{bmatrix} \tag{82}$$

where

$$\mathbf{D}_q = \begin{bmatrix} \boldsymbol{f}_0(t) & 0 & \dots & 0 \\ 0 & \boldsymbol{f}_1(t) & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \boldsymbol{f}_{A_1-1}(t) \end{bmatrix},$$

with $q = 1, 2$ the polarization index and $P_{\text{CSI-RS}} = 2A_1D_1$.

B. PRECODER CALCULATION BASED ON ANGLE- AND DELAY-BEAMFORMED CSI-RS

The FeType-II PS codebook standardized in 3GPP Rel. 17 also supports rank up to 4. The FeType-II PS precoder of the t-th SB has the same structure as Rel. 16 eType-II codebook as

$$\mathbf{w}_t^{(l)} = \Lambda_t^{(l)} \begin{bmatrix} \sum_{i=0}^{L-1} \mathbf{b}_i p_{l,0}^{(1)} \sum_{f=0}^{M_v-1} y_{t,l}^{(f)} p_{l,i,f}^{(2)} \varphi_{l,i,f} \\ \sum_{i=0}^{L-1} \mathbf{b}_i p_{l,1}^{(1)} \sum_{f=0}^{M_v-1} y_{t,l}^{(f)} p_{l,i+L,f}^{(2)} \varphi_{l,i+L,f} \end{bmatrix}. \tag{85}$$

The precoder for all N_3 SBs can be written as

$$\underbrace{\begin{bmatrix} \mathbf{w}_0^{(l)} & \cdots & \mathbf{w}_{N_3-1}^{(l)} \end{bmatrix}}_{2N_1N_2 \times N_3} = \mathbf{W}_1^{PS} \tilde{\mathbf{W}}_{C1}^{(l)} \tilde{\mathbf{W}}_{C2}^{(l)} \mathbf{W}_f^{(l)\mathsf{H}} \mathbf{\Lambda}^{(l)} \tag{86}$$

where $\mathbf{W}_1^{PS}$ is the port selection matrix of size $P_{\text{CSI-RS}} \times 2L$ comprising $L = \frac{K_1}{2}$ port selection vectors $\boldsymbol{b}_i \in \mathbb{C}^{N_1N_2}$ per polarization which belong to the set of columns of the identity matrix of dimension N_1N_2. The matrices $\tilde{\mathbf{W}}_{C1}^{(l)}$, $\tilde{\mathbf{W}}_{C2}^{(l)}$, $\mathbf{W}_f^{(l)}$, and $\mathbf{\Lambda}^{(l)}$ are similar to the Rel. 16 eType-II precoder described in VI-A and are derived from a port combining coefficient (PCC) matrix comprising port combining coefficients. Based on the received beamformed CSI-RS, the estimated DL channel $\mathbf{H}(t) \in \mathbb{C}^{N_R \times P_{\text{CSI-RS}}}$ across all N_3 SBs can be expressed as

$$\mathbf{H}(t) = \mathbf{H}_B(t)\mathbf{F}_{gNB}(t) \tag{87}$$

where $\mathbf{H}_B(t)$ is the non-precoded DL channel of dimension $N_R \times P_{\text{TXRU}}$ on the t-th SB. Note that $\mathbf{H}_B(t)$ cannot be measured by the UE since there is no CSI-RS directly associated with the TXRUs or antenna ports. The precoder selection strategy is elaborated in the following five steps:

Step 1. **Calculation of PS matrix $\mathbf{W}_1^{PS}$:** In the first step, $2L$ ports out of $P_{\text{CSI-RS}}$ ports are selected from the estimated DL channel $\mathbf{H}(t)$. Based on the $2L$ selected ports, the port selection matrix $\mathbf{W}_1^{PS} \in \mathbb{C}^{P_{\text{CSI-RS}} \times 2L}$ comprising L identical port-selection vectors across both polarizations is determined. The $2L$ ports are determined, for example, by summing up the power of the beamformed channel $\mathbf{H}(t)$ across all N_3 SBs. Here, each port selection vector $\boldsymbol{b}_i$ per polarization comprises a single one and $\frac{P_{\text{CSI-RS}}}{2} - 1$ zeros.

Step 2. **Calculation of a window comprising N strongest delays:** Following Eq. (39), the effective DL channel $\mathbf{H}_A(t) = \mathbf{H}(t)\mathbf{W}_1 \in \mathbb{C}^{N_R \times 2L}$ for all N_3 SBs is calculated. Then the time domain representation $\check{\mathbf{H}}_A(m)$ of $\mathbf{H}_A(t)$ is determined by calculating the IDFT of $\mathbf{H}_A(t)$ over all N_3 SBs as described in Eq. (73). As the UL and DL channel delays are not always reciprocal and differ by either one or two taps, a DFT matrix $\mathbf{W}_f$ can be used at the UE side like in Rel. 16 eType-II Regular and PS CBs to account for the delay mismatch. As the delay mismatch seems to be within a small range, a reduced size DFT matrix comprising N, $N \leq N_3$, delay DFT vectors is considered, where $N \in \{2, 4\}$. Therefore, a window comprising N delays from the N_3 delays are selected. The window of size N is selected by summing up the power of N consecutive delay-domain channels from $\check{\mathbf{H}}_A(m)$ associated with indices $\{m, m+1, \dots, m+N-1\}$. The detailed algorithm describing the selection of the window comprising N strongest delays is presented in Algorithm 6.

Algorithm 6 Window Selection

Input: $\check{\mathbf{H}}_A(m)\, m = 0, \cdots, N_3 - 1, N, N_3$
Output: $\mathcal{N}_t$
selection of the window of size N:
$\mathcal{N}_0 = \{0, \cdots, N_3 - N\}$
$m_0 = \arg\max_{m \in \mathcal{N}_0} \{\sum_{i=m}^{m+N-1} tr(\check{\mathbf{H}}_A^H(i)\check{\mathbf{H}}_A(i))\}$
$\mathcal{N}_t = \{m_0, \cdots, m_0 + N - 1\}$

Step 3. **Calculation of PCC matrix $\mathbf{G}_2^{(l)}$:** The PCC matrix $\mathbf{G}_2^{(l)} \in \mathbb{C}^{2L \times N}$ for each layer is calculated by performing an SVD on the covariance matrix $\mathbf{C} = \mathbf{K}^H\mathbf{K} \in \mathbb{C}^{2LN \times 2LN}$ as described in Step 3 of Section VI-C with

$$\mathbf{K} = \begin{bmatrix} \check{\mathbf{H}}_A(m_0) & \cdots & \check{\mathbf{H}}_A(m_0 + N - 1) \end{bmatrix},$$

i.e., $\mathbf{u}_l = \text{vec}\big(\mathbf{G}_2^{(l)}\big)$ where $\mathbf{u}_l$ is the $l-$th eigenvector of $\mathbf{C}$

Step 4. **Construction of FD compression matrix $\mathbf{W}_f^{(l)}$:** The best M_v columns of $\mathbf{G}_2^{(l)}$ are selected to obtain $\tilde{\mathbf{W}}_2^{(l)} \in \mathbb{C}^{2L \times M_v}$ which correspond to the M_v strongest delays for layer l. As a byproduct, it yields the FD compression matrix $\mathbf{W}_f^{(l)} \in \mathbb{C}^{N_3 \times M_v}$ which becomes layer dependent. The selection of the M_v delays is based on Algorithm 7.

Algorithm 7 Delay Selection per Layer

Input: $\mathbf{G}_2^{(l)}, M_v, N$
Output: $\tilde{\mathbf{W}}_2^{(l)}, \mathbf{W}_f^{(l)}, \tau_0^{(l)}, \cdots, \tau_{M_v-1}^{(l)}$
$\mathcal{N}_t = \{m_0, \cdots, m_0 + N - 1\}$
for $i = 0 : M_v - 1$ **do**
  $m^* = \arg\max_{m \in \mathcal{N}_t} \{\mathbf{g}_{2,m-m_0+1}^{(l)H} \mathbf{g}_{2,m-m_0+1}^{(l)}\}$
  $\tau_i^{(l)} = m^*$
  $\mathcal{N}_t = \mathcal{N}_t \backslash m^*$
end for
$\tilde{\mathbf{W}}_2^{(l)} = \begin{bmatrix} \mathbf{g}_{2,\tau_0-m_0+1}^{(l)} & \cdots & \mathbf{g}_{2,\tau_{M_v-1}-m_0+1}^{(l)} \end{bmatrix}$
$\mathbf{W}_f^{(l)} = \begin{bmatrix} \mathbf{f}_{\tau_0}^{(l)} & \cdots & \mathbf{f}_{\tau_{M_v-1}}^{(l)} \end{bmatrix}$

Step 5. **Quantization:** Based on the reduced size PCC matrix $\tilde{\mathbf{W}}_2^{(l)}$, the wideband amplitudes and SB coefficients are calculated and quantized using Step 5 and Step 6 of Section VI-B resulting in the matrices $\tilde{\mathbf{W}}_{C1}^{(l)}$ and $\tilde{\mathbf{W}}_{C2}^{(l)}$,

respectively. The K_l^{NZ} non-zero combining coefficients are determined using Algorithm 3.

Based on the matrices $\tilde{\mathbf{W}}_{C1}^{(l)}$, $\tilde{\mathbf{W}}_{C2}^{(l)}$ and $\mathbf{W}_f^{(l)}$, the BS reconstructs the PCC matrix in the frequency domain for the l-th layer and N_3 SBs as

$$\mathbf{W}^{(l)} = \mathbf{W}_1^{PS}\tilde{\mathbf{W}}_{C1}^{(l)}\tilde{\mathbf{W}}_{C2}^{(l)}\mathbf{W}_f^{(l)H} \tag{88}$$

where the t-th column $\mathbf{w}_t^{(l)}$ denotes the composite precoder associated with the t-th SB. Finally, the precoder of the t-th SB and l-th layer to be applied on the TXRU antenna ports, following the UE feedback, can be expressed as

$$\mathbf{F}_{\text{TXRU}}^{(l)}(t) = \mathbf{F}_{gNB}(t)\mathbf{w}_t^{(l)}. \tag{89}$$

In contrast to the Rel. 15 Type-II and Rel. 16 eType-II codebooks, where the precoder calculation is highly computationally complex and requires N_3 SVDs for the N_3 SBs, the precoder calculation of the Rel. 17 FeType-II PS codebook requires only one SVD for all N_3 SBs, thereby drastically reducing the computational complexity at the UE. This is possible due to the delay beamforming at the BS side. Compared to the Regular codebooks, the PS codebooks of Rel. 15 and Rel. 16 result in a sub-optimal performance due to the restriction of the beamformed antenna port selection for the port selection matrix $\mathbf{W}_1$. As the selected beamformed ports need to be always adjacent, the performance degradation is significant compared to the Regular codebooks. However, as the beamformed ports can be freely selected for the Rel. 17 FeType-II PS codebooks, it is expected that the performance is similar to that of the Rel. 16 Regular eType-II codebook.

XI. CONCLUSION

The paper describes with matrix formalism the two main codebook types specified within 3GPP, i.e., Type-I codebook (low resolution) and Type-II codebook (high resolution) up to Rel. 17. For the enhanced Type-II codebook two main practical precoder selection strategies are detailed at the UE side (i) based on singular value decomposition per sub-band in the frequency domain, (ii) based on wideband singular decomposition in the delay domain. Monte Carlo simulations demonstrate that both strategies yield good performance. The delay domain strategy performance may suffer from spatial interference for single user MIMO high rank transmissions when the channel is both spatially correlated and frequency selective. However, the frequency domain strategy complexity increases linearly with the number of sub-bands while it is not the case for the delay domain selection strategy. As a result, the delay domain selection strategy is particularly relevant for Rel. 17 codebook Type-II port selection or for a low frequency selective channel with few significant consecutive delays.

APPENDIX. DELAY DOMAIN PRECODER SELECTION STRATEGY DERIVATION

The matrix $\tilde{\mathbf{W}}_2^{(l)} \in \mathbb{C}^{2L\times M}$ can be expressed column wise as $\left[\mathbf{u}_0^{(l)}\ \mathbf{u}_1^{(l)}\ \cdots\ \mathbf{u}_{M-1}^{(l)}\right]$ where $\mathbf{u}_m^{(l)} \in \mathbb{C}^{2L}$ and the matrix $\mathbf{W}_f^{(l)H}$ is made of the best delays obtained in Algorithm 4. As a result, the matrix product $\tilde{\mathbf{W}}_2^{(l)}\mathbf{W}_f^{(l)H}$ is developed in Eq. (90), as shown at the bottom of the page.

Each column t of the matrix $\tilde{\mathbf{W}}_2^{(l)}\mathbf{W}_f^{(l)H} \in \mathbb{C}^{2L\times N_3}$ corresponds to the precoder to be applied to $\mathbf{H}_A(t)$. Stacking the precoded channel of each SB t in a vector of dimension N_3N_R, it yields the vector $\mathbf{h}_c$ defined as

$$\mathbf{h}_c = \begin{bmatrix} \mathbf{H}_A(0)(\sum_{i=0}^{M-1}\mathbf{u}_i^{(l)}) \\ \mathbf{H}_A(1)(\sum_{i=0}^{M-1}\mathbf{u}_i^{(l)}e^{j\frac{2\pi\tau_i}{N_3}}) \\ \vdots \\ \mathbf{H}_A(N_3-1)(\sum_{i=0}^{M-1}\mathbf{u}_i^{(l)}e^{j\frac{2\pi(N_3-1)\tau_i}{N_3}}) \end{bmatrix}. \tag{91}$$

$$\begin{aligned}\tilde{\mathbf{W}}_2^{(l)}\mathbf{W}_f^{(l)H} &= \left[\mathbf{u}_0^{(l)}\ \mathbf{u}_1^{(l)}\ \cdots\ \mathbf{u}_{(M-1)}^{(l)}\right]\begin{bmatrix} 1 & e^{j\frac{2\pi\tau_0}{N_3}} & \cdots & e^{j\frac{2\pi(N_3-1)\tau_0}{N_3}} \\ 1 & e^{j\frac{2\pi\tau_1}{N_3}} & \cdots & e^{j\frac{2\pi(N_3-1)\tau_1}{N_3}} \\ \vdots & \vdots & & \vdots \\ 1 & e^{j\frac{2\pi\tau_{M-1}}{N_3}} & \cdots & e^{j\frac{2\pi(N_3-1)\tau_{M-1}}{N_3}} \end{bmatrix} \\ &= \left[\sum_{i=0}^{M-1}\mathbf{u}_i^{(l)}\ \sum_{i=0}^{M-1}\mathbf{u}_i^{(l)}e^{j\frac{2\pi\tau_i}{N_3}}\ \cdots\ \sum_{i=0}^{M-1}\mathbf{u}_i^{(l)}e^{j\frac{2\pi(N_3-1)\tau_i}{N_3}}\right]\end{aligned} \tag{90}$$

$$\check{\mathbf{h}}_c = \begin{bmatrix} \check{\mathbf{H}}_A(\tau_0) \\ \mathbf{0}_{N_R(\tau_1-\tau_0-1)\times 2L} \\ \check{\mathbf{H}}_A(\tau_1) \\ \mathbf{0}_{N_R(\tau_2-\tau_1-1)\times 2L} \\ \vdots \\ \check{\mathbf{H}}_A(\tau_{M-1}) \\ \mathbf{0}_{N_R(N_3-\tau_{M-1}-1)\times 2L} \end{bmatrix}\mathbf{u}_0^{(l)} + \begin{bmatrix} \check{\mathbf{H}}_A(\tau_1) \\ \mathbf{0}_{N_R(\tau_2-\tau_1-1)\times 2L} \\ \check{\mathbf{H}}_A(\tau_2) \\ \mathbf{0}_{N_R(\tau_3-\tau_2-1)\times 2L} \\ \vdots \\ \check{\mathbf{H}}_A(\tau_0) \\ \mathbf{0}_{N_R(\tau_1-\tau_0-1)\times 2L} \end{bmatrix}\mathbf{u}_1^{(l)} + \cdots + \begin{bmatrix} \check{\mathbf{H}}_A(\tau_{M-1}) \\ \mathbf{0}_{N_R(N_3-\tau_{M-1}-1)\times 2L} \\ \check{\mathbf{H}}_A(\tau_0) \\ \mathbf{0}_{N_R(\tau_1-\tau_0-1)\times 2L} \\ \vdots \\ \check{\mathbf{H}}_A(\tau_{M-2}) \\ \mathbf{0}_{N_R(\tau_{M-1}-\tau_{M-2}-1)\times 2L} \end{bmatrix}\mathbf{u}_{M-1}^{(l)} \tag{94}$$

It can be expanded as:

$$\mathbf{h}_c = \sum_{i=0}^{M-1} \begin{bmatrix} \mathbf{H}_A(0) \\ \mathbf{H}_A(1)e^{j\frac{2\pi\tau_i}{N_3}} \\ \vdots \\ \mathbf{H}_A(N_3-1)e^{j\frac{2\pi(N_3-1)\tau_i}{N_3}} \end{bmatrix} \mathbf{u}_i^{(l)}. \quad (92)$$

Let $\check{\mathbf{h}}_c$ be the delay domain representation of the vector $\mathbf{h}_c$, i.e.,

$$\check{\mathbf{h}}_c \in \mathbb{C}^{N_3 N_R} = \mathbf{U}_{N_R}^H \mathbf{h}_c. \quad (93)$$

Developing equation Eq. (93), based on the shift theorem stating that a delay corresponds to a linear phase term in the frequency domain, yields Eq. (94), as shown at the bottom of the previous page, where the channel is assumed to be negligible outside the selected delays.

The first precoded MIMO channel coefficient in time (corresponding to zero delay) can be written as:

$$\left[\check{\mathbf{H}}_A(\tau_0)\ \check{\mathbf{H}}_A(\tau_1)\ \cdots\ \check{\mathbf{H}}_A(\tau_{M-1})\right]\mathbf{u}^{(l)} \quad (95)$$

where $\mathbf{u}^{(l)} = \text{vec}\left(\tilde{\mathbf{W}}_2^{(l)}\right)$. Clearly, $\mathbf{u}^{(l)}$ being the l^{th} eigenvector of $\mathbf{C}$ given in Eq. (75) maximizes the power of the precoded channel in the delay domain.

ACKNOWLEDGMENT

The authors would like to thank Thierry Clessienne from Orange Innovation. Without his kind advice and constant support, this tutorial project would not have happened. They also want to thank Rita Ibrahim from Orange Innovation for the insightful discussions on multi-user MIMO precoding.

REFERENCES

[1] D. J. Love, R. W. Heath Jr., V. K. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[2] C. Lim, T. Yoo, B. Clerckx, B. Lee, and B. Shim, "Recent trend of multiuser MIMO in LTE-advanced," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 127–135, Mar. 2013.

[3] B. Clerckx and C. Oestges, *MIMO Wireless Networks: Channels, Techniques and Standards for Multi-Antenna, Multi-User and Multi-Cell Systems*. New York, NY, USA: Academic, 2013.

[4] B. Özbek and D. Le Ruyet, *Feedback Strategies for Wireless Communication Systems*. New York, NY, USA: Springer, Oct. 2013, p. 332. [Online]. Available: https://hal.science/hal-01126396, doi: 10.1007/978-1-4614-7741-9.

[5] *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; NR; Physical Layer Procedures for Data (Release 15)*, document TS 38.214, Version 15.14.0, 3GPP, Sep. 2021.

[6] *WF on Type I and II CSI Codebooks*, document R1-1709232, 3GPP, 2017. [Online]. Available: https://www.3gpp.org/ftp/tsg_ran/WG1_RL1/TSGR1_97/Docs

[7] H. Jin, K. Liu, G. Lee, E. J. Farag, M. Zhang, D. Zhu, L. Zhang, E. Onggosanusi, M. Shafi, and H. Tataria, "Massive MIMO evolution towards 3GPP release 18," 2022, *arXiv:2210.08218*.

[8] C. E. Shannon, "Channels with side information at the transmitter," *IBM J. Res. Develop.*, vol. 2, no. 4, pp. 289–293, Oct. 1958.

[9] J. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback–I: No bandwidth constraint," *IEEE Trans. Inf. Theory*, vol. IT-12, no. 2, pp. 172–182, Apr. 1966.

[10] G. Turin, "Signal design for sequential detection systems with feedback," *IEEE Trans. Inf. Theory*, vol. IT-11, no. 3, pp. 401–408, Jul. 1965.

[11] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2632–2639, Sep. 2001.

[12] R. S. Blum, "MIMO with limited feedback of channel state information," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 4, Apr. 2003, p. IV-89.

[13] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1423–1436, Oct. 1998.

[14] V. Lau, Y. Liu, and T.-A. Chen, "On the design of MIMO block-fading channels with feedback-link capacity constraint," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 62–70, Jan. 2004.

[15] J. C. Roh and B. D. Rao, "Transmit beamforming in multiple-antenna systems with finite rate feedback: A VQ-based approach," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1101–1112, Mar. 2006.

[16] A. D. Dabbagh and D. J. Love, "Feedback rate-capacity loss tradeoff for limited feedback MIMO systems," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2190–2202, May 2006.

[17] D. Gerlach and A. Paulraj, "Adaptive transmitting antenna arrays with feedback," *IEEE Signal Process. Lett.*, vol. 1, no. 10, pp. 150–152, Oct. 1994.

[18] R. W. Heath Jr. and A. Paulraj, "A simple scheme for transmit diversity using partial channel feedback," in *Proc. 32nd Asilomar Conf. Signals, Syst. Comput.*, vol. 2, Nov. 1998, pp. 1073–1078.

[19] K. Kiran Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2562–2579, Oct. 2003.

[20] D. J. Love, R. W. Heath Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2735–2747, Oct. 2003.

[21] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 1962–1973, 2000.

[22] D. J. Love and R. W. Heath Jr., "Equal gain transmission in multiple-input multiple-output wireless systems," *IEEE Trans. Commun.*, vol. 51, no. 7, pp. 1102–1110, Jul. 2003.

[23] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1051–1064, May 2002.

[24] D. J. Love and R. W. Heath Jr., "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2967–2976, Aug. 2005.

[25] J. C. Roh and B. D. Rao, "Efficient feedback methods for MIMO channels based on parameterization," *IEEE Trans. Wireless Commun.*, vol. 6, no. 1, pp. 282–292, Jan. 2007.

[26] M. A. Sadrabadi, A. K. Khandani, and F. Lahouti, "Channel feedback quantization for high data rate MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3335–3338, Dec. 2006.

[27] B. Varadarajan, E. Onggosanusi, A. Dabak, and R. Chen, "Nested codebook design for MIMO precoders," in *Proc. 42nd Asilomar Conf. Signals, Syst. Comput.*, Oct. 2008, pp. 723–727.

[28] J. Wang, M. Wu, and F. Zheng, "The codebook design for MIMO precoding systems in LTE and LTE-A," in *Proc. 6th Int. Conf. Wireless Commun. Netw. Mobile Comput. (WiCOM)*, Sep. 2010, pp. 1–4.

[29] B. C. Banister and J. R. Zeidler, "A simple gradient sign algorithm for transmit antenna weight adaptation with feedback," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1156–1171, May 2003.

[30] B. Mondal and R. W. Heath Jr., "Channel adaptive quantization for limited feedback MIMO beamforming systems," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4717–4729, Dec. 2006.

[31] R. Samanta and R. W. Heath Jr., "Codebook adaptation for quantized MIMO beamforming systems," in *Proc. 39th Asilomar Conf. Signals, Syst. Comput.*, 2005, pp. 376–380.

[32] R. W. Heath Jr., T. Wu, and A. C. K. Soong, "Progressive refinement of beamforming vectors for high-resolution limited feedback," *EURASIP J. Adv. Signal Process.*, vol. 2009, no. 1, pp. 1–13, Dec. 2009.

[33] T. Kim, D. J. Love, and B. Clerckx, "MIMO systems with limited rate differential feedback in slowly varying channels," *IEEE Trans. Commun.*, vol. 59, no. 4, pp. 1175–1189, Apr. 2011.

[34] G. Berardinelli, T. B. Sørensen, P. Mogensen, and K. Pajukoski, "SVD-based vs. release 8 codebooks for single user MIMO LTE-A uplink," in *Proc. IEEE 71st Veh. Technol. Conf.*, May 2010, pp. 1–5.

[35] T. Shuang, T. Koivisto, H.-L. Maattanen, K. Pietikainen, T. Roman, and M. Enescu, "Design and evaluation of LTE-advanced double codebook," in *Proc. IEEE 73rd Veh. Technol. Conf. (VTC Spring)*, May 2011, pp. 1–5.

[36] E. Onggosanusi, Y. Li, M. S. Rahman, Y.-H. Nam, J. Zhang, J.-Y. Seol, and T. Kim, "Reduced space channel feedback for FD-MIMO," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 3873–3878.

[37] B. Mondal, V. Sergeev, A. Sengupta, and A. Davydov, "5G-NR (new radio) CSI computation algorithm and performance," in *Proc. 52nd Asilomar Conf. Signals, Syst., Comput.*, Oct. 2018, pp. 1068–1071.

[38] T. Akyıldız and T. M. Duman, "Search-free precoder selection for 5G new radio using neural networks," in *Proc. IEEE Int. Black Sea Conf. Commun. Netw. (BlackSeaCom)*, May 2020, pp. 1–6.

[39] X. Fu, D. L. Ruyet, R. Visoz, and T. Clessienne, "On the performance analysis of different selecting strategies for type-II codebook," in *Proc. Joint Eur. Conf. Netw. Commun. 6G Summit (EuCNC/6G Summit)*, Jun. 2022, pp. 83–88.

[40] C.-C. Tsai, T.-Y. Yeh, W.-H. Chou, W.-C. Pao, and J.-Y. Pan, "A low complexity PMI selection scheme for 3GPP 5G NR FDD systems," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Dec. 2021, pp. 1917–1922.

[41] *Phase Preprocessing for Type-II CSI Enhancement*, document R1-1906348, 3GPP, 2019. [Online]. Available: https://www.3gpp.org/ftp/tsg_ran/WG1_RL1/TSGR1_97/Docs

[42] H. Miao, M. D. Mueck, and M. Faerber, "Amplitude quantization for type-2 codebook based CSI feedback in new radio system," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2018, pp. 1–9.

[43] L. Suárez, N. Ryabov, V. Lyashev, and A. Sherstobitov, "DFT based beam-time delay sparse channel representation for channel state information (CSI) compression in 5G FDD massive MIMO systems," in *Proc. IEEE Int. Black Sea Conf. Commun. Netw. (BlackSeaCom)*, Jun. 2018, pp. 1–5.

[44] R. Ahmed, F. Tosato, and M. Maso, "Overhead reduction of NR type II CSI for NR release 16," in *Proc. 23rd Int. ITG Workshop Smart Antennas (WSA)*, Apr. 2019, pp. 1–5.

[45] Z. Liu, S. Sun, Q. Gao, and H. Li, "CSI feedback based on spatial and frequency domains compression for 5G multi-user massive MIMO systems," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Aug. 2019, pp. 834–839.

[46] *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; NR; Physical Layer Procedures for Data (Release 15)*, document TS 38.214, Version 16.2.0, 3GPP, Jun. 2020.

[47] Y. Zhang, W. Zeng, H. Sun, Y. Kim, and D. Zhang, "Encoding of enhanced type II channel state information," U.S. Patent 11 063 649, Jul. 13, 2021.

[48] A. Hindy, U. Mittal, and T. Brown, "CSI feedback overhead reduction for 5G massive MIMO systems," in *Proc. 10th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2020, pp. 0116–0120.

[49] T. Abe and G. Bauch, "Effective SINR computation for maximum likelihood detector in MIMO spatial multiplexing systems," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2009, pp. 1–5.

[50] *CSI Enhancement for MU-MIMO Support*, document R1-1903343, 3GPP, 2019. [Online]. Available: https://www.3gpp.org/ftp/tsg_ran/WG1_RL1/TSGR1_96/Docs

[51] X. Li, S. Wang, and Y. Cai, "Tutorial: Complexity analysis of singular value decomposition and its variants," 2019, *arXiv:1906.12085*.

[52] E. Ohlmer and G. Fettweis, "Link adaptation in linearly precoded closed-loop MIMO-OFDM systems with linear receivers," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2009, pp. 1–6.

[53] J.-B. Landre, Z. El Rawas, R. Visoz, and S. Bouguermouh, "Realistic performance of LTE: In a macro-cell environment," in *Proc. IEEE 75th Veh. Technol. Conf. (VTC Spring)*, May 2012, pp. 1–5.

[54] I. Latif, F. Kaltenberger, N. Nikaein, and R. Knopp, "Large scale system evaluations using PHY abstraction for LTE with OpenAirInterface," in *Proc. 6th Int. Conf. Simul. Tools Techn.*, 2013, pp. 24–30.

[55] E. Lukashova, F. Kaltenberger, R. Knopp, and C. Bonnet, "Phy layer abstraction for SU-MIMO LTE system employing parallel interference-aware detection," in *Proc. 1st Int. Workshop Link- Syst. Level Simulations (IWSLS)*, Jul. 2016, pp. 1–6.

[56] E. Tuomaala and H. Wang, "Effective SINR approach of link to system mapping in OFDM/multi-carrier mobile network," in *Proc. 2nd Asia–Pacific Conf. Mobile Technol., Appl. Syst.*, 2005, p. 5.

[57] K. Sayana, J. Zhuang, and K. Stewart, "Short term link performance modeling for ML receivers with mutual information per bit metrics," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2008, pp. 1–6.

[58] A. M. Cipriano, R. Visoz, and T. Salzer, "Calibration issues of PHY layer abstractions for wireless broadband systems," in *Proc. IEEE 68th Veh. Technol. Conf.*, Sep. 2008, pp. 1–5.

[59] X. Lei and Z. Jin-bao, "An accurate mean mutual information computational approach of link performance abstraction," in *IEEE 10th Int. Conf. SIGNAL Process.*, Oct. 2010, pp. 1552–1556.

[60] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.

[61] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication—Part I: Channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195–202, Jan. 2005.

[62] L.-U. Choi and R. D. Murch, "A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 20–24, Jan. 2004.

[63] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[64] D. Gesbert, M. Kountouris, R. W. Heath Jr., C.-B. Chae, and T. Salzer, "Shifting the MIMO paradigm," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 36–46, Sep. 2007.

[65] Z. Pan, K.-K. Wong, and T.-S. Ng, "Generalized multiuser orthogonal space-division multiplexing," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1969–1973, Nov. 2004.

[66] B. Bandemer, M. Haardt, and S. Visuri, "Linear MMSE multi-user MIMO downlink precoding for users with multiple antennas," in *Proc. IEEE 17th Int. Symp. Pers., Indoor Mobile Radio Commun.*, Sep. 2006, pp. 1–5.

[67] F. Kaltenberger, D. Gesbert, R. Knopp, and M. Kountouris, "Performance of multi-user MIMO precoding with limited feedback over measured channels," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2008, pp. 1–5.

[68] *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; NR; Physical Layer Procedures for Data (Release 15)*, document TS 38.214, Version 17.3.0, 3GPP, Sep. 2022.

[69] H. Huawei, *Discussion on Field Measurements and Evaluation Assumptions for FDD CSI*, document R1-2006414, 2020.

[70] F. H. Fraunhofer IIS, *On FDD Channel Reciprocity in Real-World Scenarios*, document R1-2006998, 2020.

XIAOTIAN FU (Member, IEEE) received the B.Sc. degree in automation from Southeast University, Nanjing, China, in 2017, the M.Sc. degree in telecommunications and networks from Conservatoire National des Arts et Métiers (CNAM), Paris, France, in 2018, and the Ph.D. degree in telecommunications and networks from HESAM University, Paris, in 2022. She is currently a Concepts Researcher, focusing on simulations and research relating to satellite communications. Her major research interests include NTN, MIMO, NOMA, and AI/ML for future wireless networks.

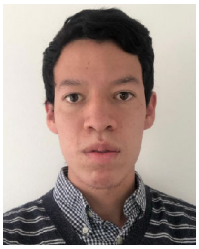
DIDIER LE RUYET (Senior Member, IEEE) received the Engineering degree in electrical engineering, the M.Sc. degree in physical systems and metrology, the Ph.D. degree in communication from Conservatoire National des Arts et Métiers (CNAM), Paris, France, in 1994 and 2001, respectively, and the Habilitation à diriger des recherches degree from Paris XIII University, in 2009. From 1988 to 1996, he was a Senior Member of the Technical Staff with SAGEM Defense and Telecommunication, Paris. He joined the Signal and Systems Laboratory, CNAM, as a Research Assistant, in 1996, where he was an Assistant Professor, from 2002 to 2010. Since 2010, he has been a Full Professor with CNAM, CEDRIC Research Laboratory. From 2016 to 2019, he was the Deputy Director of the CEDRIC. He has published about 200 papers in refereed journals and conference proceedings and published five books in the area of digital communication. He has been involved in several national and European projects dealing with multicarrier transmission techniques and multi-antenna transmission. His main research interests include the areas of digital communications and signal processing, including channel coding, detection and estimation algorithms, filter bank-based multi-carrier communication, and multi-antenna transmission. He served as a Technical Program Committee Member for major IEEE ComSoc and VTS conferences, the General Chair for the Ninth Edition of the ISWCS 2012 Conference, and the Co-Chair for the ISWCS 2013 and 2019 editions.

RAPHAEL VISOZ (Senior Member, IEEE) received the Ph.D. degree in digital communications from École nationale Supérieure des télécommunications (ENST), Paris, France, in February 2002, and the Accreditation degree (Habilitation à Diriger des Recherches, Supervise Research) from University Pierre et Marie Curie, Paris 6, France, in September 2006. Since November 1997, he has been working for Orange Laboratories in the field of 3/4/5 generation mobile radio systems. Since 2015, he has been a 3GPP RAN1 Delegate for Orange. He is the author or coauthor of about 90 papers in refereed journals and conference proceedings and owns more than 50 patents in the field of digital communications. His research interests include network information theory, PHY/MAC cross-layer optimization mechanisms, multi-antenna technology (MIMO systems), iterative decoding on graphs, and channel estimation. He was elevated to an Emeritus Member of SEE, in 2021.

VENKATESH RAMIREDDY received the M.Sc. degree in electrical engineering from the Ilmenau University of Technology, in 2016, where he is currently pursuing the Ph.D. degree with the Department of Electrical Engineering. Since then, he has been with the Fraunhofer Institute for Integrated Circuits IIS. Since 2018, he has been a Standardization Delegate for NR in 3GPP RAN WG1. His current research interests include multi-dimensional precoding and codebook design for multi-antenna systems.

MARCUS GROSSMANN received the Dipl.-Ing. (M.S.) and Dr.-Ing. degrees in electrical engineering from the Ilmenau University of Technology, in 2004 and 2012, respectively. He joined the Fraunhofer Institute for Integrated Circuits IIS, in 2011, and since he has been working on multi-antenna processing and mmWave communications. Since 2016, he has been a Standardization Delegate for NR in 3GPP RAN WG1.

MARKUS LANDMANN received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the Ilmenau University of Technology, Germany, in 2001 and 2008, respectively. Until 2009, he was a Research Assistant with the Ilmenau University of Technology and the Tokyo Institute of Technology on wireless propagation, channel modeling, and array signal processing. Until 2018, he was responsible for the testing and development of satellite and terrestrial-based communication systems (2G–5G) with the Fraunhofer IIS in the Facility for Over Air Research and Testing. Since 2018, he has been a Chief Scientist of the EMS Department being the Co-Department Head, since 2022, and he is responsible for the strategic topics related to 5G standardization.



WILMAR QUIROGA is currently pursuing the double degree in electronic engineering with the National University of Colombia, and Télécom Paris, Palaiseau, France, in cellular networks and the Internet of Things. He was an Intern in the areas of automation and industrial control with TechnipFMC, Bogotá, Colombia, in 2021, and recently with Orange Innovation, Chatillon, France, in 2023, in the field of evaluation of the performance of MIMO techniques. His research interests include wireless networks, network automation, and machine learning.