3GPP TSG-RAN WG1 #61 **R1-102630**
Montreal, Canada, May 10 – 14, 2010

Source: **Ericsson, ST-Ericsson**
Title: **Refinements of Feedback and Codebook Design**
Agenda Item: **6.3.4**
Document for: **Discussion and Decision**

1. Introduction

This contribution discusses feedback in support of SU-MIMO and MU-MIMO. In RAN1 #60, further progress on the long-standing feedback issue was made and it is now clear that RAN1's efforts should be spent on implicit type of precoder feedback as evident from

- *Implicit feedback (PMI/RI/CQI) is used also for Rel-10*
  - *UE spatial feedback for a subband represents a precoder (as constructed below)*
  - *CQI computed based on the assumption that eNodeB uses a specific precoder (or precoders), as given by the feedback, on each subband within the CQI reference resource*
    - *Note that a subband can correspond to the whole system bandwidth*

in the agreed way forward in [10] . The same way forward also nailed down the following structural properties of the precoder feed back

- *A precoder for a subband is composed of two matrices*
  - *The precoder structure is applied to all Tx antenna array configurations*
  - *Each of the two matrices belong to a separate codebook*
    - *The codebooks are for further study*
    - *The codebooks are known (or synchronized) at both the eNodeB and UE*
    - *Codebooks may or may not change/vary over time and/or different subbands*
  - *That is, two codebook indices together determine the precoder*
  - *One of the two matrices targets wideband and/or long-term channel properties*
  - *The other matrix targets frequency-selective and/or short-term channel properties*
  - *Note that a matrix codebook in this context should be interpreted as a finite enumerated set of matrices that for each RB is known to both UE and eNodeB.*
  - *Note that Rel-8 precoder feedback can be deemed as a special case of this structure*

Inspired by the agreed way forward, and also by the way forward on feedback refinements in [11] , this contribution describes an implicit feedback concept that encompasses SU-MIMO as well as MU-MIMO operation and at the same time manages to limit the feedback overhead.

2. Multi-Granular Feedback Covering SU-MIMO and MU-MIMO

The present type of implicit feedback of precoders can be used for SU-MIMO as well as MU-MIMO. Striving for finer spatial quantization however typically results in a high signaling overhead. To circumvent that, this section focuses on the most central points in the agreed way forward, i.e.,

- *A precoder for a subband is composed of two matrices*
  - *Each of the two matrices belong to a separate codebook*
  - *One of the two matrices targets wideband and/or long-term channel properties*
  - *The other matrix targets frequency-selective and/or short-term channel properties*

LG Elecs. Ex. 1011
LG Elecs. v. Pantech Corp.
IPR2023-01273 Page 1

and in more detail describes such a multi-granular precoder approach that offers finer spatial quantization while still limiting the feedback overhead. Such an approach would thus suit SU-MIMO and MU-MIMO operation without leading to unreasonable overhead for the common case of SU-MIMO.

Clearly, the decision says that a precoder matrix for a subband is a function of two other matrices. In other words, the overall recommended precoder $\mathbf{W}$ for a subband can be written as

$$\mathbf{W} = \mathbf{G}\left(\mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right)$$

where $\mathbf{G}(\cdot,\cdot)$ is some matrix function describing the mapping from the two constituent matrices $\mathbf{W}^{(1)}$, $\mathbf{W}^{(2)}$ to the overall recommended precoder $\mathbf{W}$. Instead of the Rel-8 PMI reporting, where only a single PMI describes a recommended precoder, the overall recommended precoder is determined from the feed back of two different PMIs, each corresponding to one of the two constituent precoder matrices. There are thus two separate codebooks, C_1 and C_2 and each constituent precoder matrix is selected from a codebook of its own, i.e.,

$$\mathbf{W}^{(1)} \in C_1, \quad \mathbf{W}^{(2)} \in C_2$$

As pointed out in the feedback decision, the purpose of the first precoder matrix $\mathbf{W}^{(1)}$ is to take care of wideband and/or long-term channel properties. This translates into designing the codebook C_1 for supporting a rather fine spatial resolution by exploiting the correlation properties of the channel. Since the correlation properties are stable over frequency and time, the frequency/time granularity of reporting $\mathbf{W}^{(1)}$ can be rather coarse without significant performance loss. On the other hand, as we will see in later section, it should be noted that exploiting time coherence is challenging since the CSI feedback no longer is self-contained within a subframe.

An extra overhead due to a finer spatial resolution can be compensated for by simplifying the task of the second precoder via compressing the channel into fewer dimensions. Basically, $\mathbf{W}^{(1)}$ implements some form of beamforming based on correlation properties of the channel. The channel and $\mathbf{W}^{(1)}$ together form a new effective $N_R \times \tilde{N}_T$ channel were in many cases $\tilde{N}_T$ is substantially smaller than the number of antenna ports, N_T.

The second matrix, $\mathbf{W}^{(2)}$, tries to match the instantaneous properties of the effective channel, such as phase alignment for constructive combining of transmitted signals on receive side, or orthogonalization of the effective channel. As such it benefits from being reported often and in a frequency-selective manner. This way of splitting the reporting of the recommended precoder into two parts offers a way to increase the spatial resolution of the precoding while still limiting the feedback overhead. This is possible due to the dimension reduction limiting the needed codebook size for $\mathbf{W}^{(2)}$ and the low update rate needed for the correlation exploiting $\mathbf{W}^{(1)}$.

3. Multi-Granular Product Precoder Codebook Design

The way forward on feedback refinements in [11] roughly have three main points:

- A precoder is constructed as a matrix multiplication of two precoders
- Full PA utilization property for at least a subset of the precoders
- Grid of beams with at least a beam oversampling factor of 2

This section discusses these points and attempts to design codebooks in line with the assumptions.

3.1. Matrix Product Precoder Structure

A simple structure for the mapping function $\mathbf{G}(\cdot,\cdot)$ is to construct the overall precoder as a product of the constituent matrices. This is reminiscent of the proposals concerning adaptive codebook [3] as well as similar to the

product precoder proposal in [8] . In fact, many companies' proposals fall under this framework. The overall precoder is thus formed as

$$\mathbf{W} = \mathbf{W}^{(1)}\mathbf{W}^{(2)}$$

and it is now easy to see from

$$\mathbf{HW} = \mathbf{HW}^{(1)}\mathbf{W}^{(2)} = \left(\mathbf{HW}^{(1)}\right)\mathbf{W}^{(2)} = \mathbf{H}_{\text{eff}}\mathbf{W}^{(2)}$$

how the first, inner, $N_{\text{T}} \times \tilde{N}_{\text{T}}$ precoder matrix $\mathbf{W}^{(1)}$, serves to created a new effective and improved $N_{\text{R}} \times \tilde{N}_{\text{T}}$ channel matrix $\mathbf{H}_{\text{eff}}$ for the outer $\tilde{N}_{\text{T}} \times r$ precoder $\mathbf{W}^{(2)}$ to work on. Often, the number of virtual antennas $\tilde{N}_{\text{T}}$ is much smaller than the number of antenna ports N_{T}, offering a considerable dimension reduction and thus requiring a smaller codebook for the outer precoder.

Proposal

- *A precoder for a subband is obtained as a matrix multiplication of two precoder matrices*
- *The inner precoder* $\mathbf{W}^{(1)}$ *(i.e., closest to the channel) should typically be a tall matrix so that dimension reduction is offered.*

This structure is well-suited for efficiently supporting common antenna setups such as closely spaced cross-poles or closely spaced ULA. To see how correlations properties are exploited and dimension reduction achieved, consider the common case of an array of closely spaced cross-poles. The antennas can then be divided into two groups based on polarization and the corresponding channels are denoted $\mathbf{H}_{/}$ and $\mathbf{H}_{\backslash}$, respectively. Assuming for example that the inner precoder $\mathbf{W}^{(1)}$ has a block diagonal structure,

$$\mathbf{W}^{(1)} = \begin{bmatrix} \tilde{\mathbf{W}}^{(1)} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{W}}^{(1)} \end{bmatrix} \tag{1}$$

The product of the MIMO channel and the overall precoder can then be written as

$$\mathbf{HW} = \begin{bmatrix}\mathbf{H}_{/} & \mathbf{H}_{\backslash}\end{bmatrix}\mathbf{W}^{(1)}\mathbf{W}^{(2)} = \begin{bmatrix}\mathbf{H}_{/} & \mathbf{H}_{\backslash}\end{bmatrix}\begin{bmatrix} \tilde{\mathbf{W}}^{(1)} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{W}}^{(1)} \end{bmatrix}\mathbf{W}^{(2)} = \begin{bmatrix}\mathbf{H}_{/}\tilde{\mathbf{W}}^{(1)} & \mathbf{H}_{\backslash}\tilde{\mathbf{W}}^{(1)}\end{bmatrix}\mathbf{W}^{(2)} = \mathbf{H}_{\text{eff}}\mathbf{W}^{(2)}$$

As seen, $\tilde{\mathbf{W}}^{(1)}$ separately precodes each group of co-polarized and closely spaced antennas forming a smaller and improved effective channel $\mathbf{H}_{\text{eff}}$. If $\tilde{\mathbf{W}}^{(1)}$ corresponds to a beamforming vector, the effective channel would reduce to having only two virtual antennas, which reduces the needed size of the codebook used for the outer precoder $\mathbf{W}^{(2)}$ when tracking the instantaneous channel properties. In this case, instantaneous channel properties are to a large extent dependent upon the relative phase relation between the two orthogonal polarizations.

Multi-granular precoder feedback is also able to handle MU-MIMO in the light of that MU-MIMO primarily may have a need for reducing the spatial quantization errors while explicit versus implicit feedback type is of lesser importance. Since co-scheduled UEs should be separated by means of high correlation, $\mathbf{W}^{(1)}$ is responsible for the

UE separation and since the precoder matrix may be reported seldom and in a wideband fashion the signaling overhead can be limited even if the spatial resolution is increased in order to target MU-MIMO operation.

Observation

- *The same multi-granular product precoder feedback handles both SU-MIMO and MU-MIMO*
- *Multi-granular product precoder feedback keeps feedback overhead limited and thereby increases the chance of needing only a single feedback mode that addresses both SU-MIMO and MU-MIMO operation*

3.2. Full PA Utilization

In line with the way forward on feedback refinements in [11] , being able to ensure full PA utilization remains an important goal. The Rel-8 codebook design took extra care to make sure that the PAs can be fully utilized by designing according to the constant modulus property. Since PA resources are expensive, it remains important also for the Rel-10 to ensure that all PAs can be fully utilized even when the transmission is performed exactly according to the precoder report. Note that this does not necessarily imply that all precoders must allow full PA utilization. But a sufficiently large set of full PA utilization precoders need to be available so that it is possible to only use those and still get good performance.

The problem with not having a set of full PA utilization precoders is that this would effectively mandate the eNodeB to modify the precoder in order to fully utilize the PAs. Since the CQI is tied to the reported precoder, such mandatory precoder modification substantially increases the CQI errors, which can have a major impact on the performance and is thus not an attractive solution. For MU-MIMO this may be less of a problem since the CQIs anyway will have significant errors due to the hard to predict inter-user interference but for SU-MIMO departing from the present full PA utilization would be a substantial step backwards with respect to Rel-8.

Observation

- *Not having a set of full PA utilization precoders would mandate eNodeB to perform precoder modification with resulting CQI mismatch*
- *At least for SU-MIMO, full PA utilization remains an important goal*

Proposal

- *Codebooks for inner and outer precoder designed so that overall precoder fully utilizes all PAs, i.e., for the overall precoder it holds that*

$$\left[\mathbf{W}\mathbf{W}^*\right]_{mm} = \kappa, \quad m = 1, \cdots, N_{\mathrm{T}}$$

- *Introducing additional precoders not fulfilling full PA utilization FFS*

Note that the constant modulus property is a sufficient but not necessary condition for full PA utilization to hold. Thus enforcing constant modulus precoders pose a more restrictive design constraint than the requirement above.

3.3. Simultaneous Support of ULA and Cross-Pole

The type of antenna setup affects the channel properties, which in turn affects the design of the codebooks. As indicated by the way forward, two antenna setups appear to be more important than others – the closely spaced cross-pole and the co-polarized uniform linear array (ULA). Both theses types should be efficiently supported by the same codebook design.

3.3.1 Closely Spaced Cross-Pole

The closely spaced cross-pole is a common antenna array setup, both for 4 Tx as well as for 8 Tx. As indicated in (1), the antennas can then be divided into two separate groups depending on the polarization direction of the antenna.

The correlation is high among the channels within an antenna group while channels from different antenna groups fade in an independent manner, and to some extent with reduced cross-talk due to the use of orthogonal polarizations. Such an antenna setup thus creates quite pronounced channel properties, which are well-matched to a block diagonal design along the lines of

$$\mathbf{W} = \begin{bmatrix} \tilde{\mathbf{W}}^{(1)} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{W}}^{(1)} \end{bmatrix} \mathbf{W}^{(2)}$$

very well. The precoder on the diagonal, $\tilde{\mathbf{W}}^{(1)}$, is targeting a co-polarized antenna group. Since the correlation is high within the antenna group, it makes sense to use a grid of beam codebook implemented from DFT based precoder vectors. This is in line with the way forward which mentions that at least eight beams should be used. Eight beams corresponds to a beam oversampling factor of two which is the standard rule of thumb in beam design to ensure that the array gain is uniformly maintained over the sector. The outer precoder, $\mathbf{W}^{(2)}$, adjusts the relative phase shift between polarizations, at least for rank 1 and 2 where the 2 Tx Rel-8 codebook could be re-used. For rank 1, the precoder could for example be formed as

$$\mathbf{W} = \begin{bmatrix} \tilde{\mathbf{w}} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{w}} \end{bmatrix} \begin{bmatrix} 1 \\ \alpha \end{bmatrix}, \quad \alpha \in \{1, -1, j, -j\} \tag{2}$$

where the antenna group beam $\tilde{\mathbf{w}} \in G^{(1,2)}$ and

$$G^{(k,Q)} = \bigcup_{q=0}^{Q-1} G_q^{(k,Q)} \tag{3}$$

with $G_q^{(k,Q)}$ representing the set of all *k*-column columns subsets of the DFT based generator matrix $\mathbf{G}_q^{(Q)}$ having elements

$$\left[\mathbf{G}_q^{(Q)}\right]_{mn} = \exp\left(j \frac{2\pi}{N_{\mathrm{T}}/2} m \left(n + \frac{q}{Q} \right) \right), \quad q = 0, 1, \cdots, Q-1 \tag{4}$$

The rank 2 case would follow similarly as

$$\mathbf{W} = \begin{bmatrix} \tilde{\mathbf{w}} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{w}} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ \alpha & -\alpha \end{bmatrix}, \quad \alpha \in \{1, j\}, \quad \tilde{\mathbf{w}} \in G_q^{(1,2)} \tag{5}$$

3.3.2 Co-polarized Closely Spaced ULA

The rank 1 case in (2) can also be made to suit the closely spaced co-polarized ULA. Also in this case should grid of beams using DFT based precoders be supported for coping efficiently with the high correlation caused by closely spaced antennas. Again using the rule of thumb of an oversampling factor of two, sixteen beams are needed for the 8 Tx overall precoder $\mathbf{W}$. As we will see, this however does not imply that we need 16 new inner precoders $\mathbf{W}^{(1)}$ in addition to the 8 precoders in support of cross-polarized setups to support both the cross-pole and the ULA with the same design.

Since a DFT precoder of size N_{T} can be formed using two DFTs of size $N_{\mathrm{T}}/2$ combined with a size-2 butterfly, it can be shown that

$$\mathbf{W} = \begin{bmatrix} \tilde{\mathbf{w}} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{w}} \end{bmatrix} \begin{bmatrix} 1 \\ \alpha \end{bmatrix}, \quad \alpha \in \{1, -1, j, -j\}, \quad \tilde{\mathbf{w}} \in G_q^{(1,4)} \tag{6}$$

is enough to also create sixteen different DFT based size N_{T} precoders. Thus, if the beam $\tilde{\mathbf{w}}$ is chosen from sixteen DFT size $N_{\mathrm{T}}/2$ based beams, then that, in combination with the outer precoder $\mathbf{W}^{(2)}$, is enough to also create sixteen different DFT based size N_{T} precoders. Hence, some of the precoders for the cross-pole case are re-used for the co-pole. Although the oversampling factor is now $Q = 4$ for the antenna group precoder, this increase compared to the rule of thumb for beam oversampling provides an improvement of the spatial resolution that is beneficial for MU-MIMO but also for SU-MIMO. Rank 2 is similar to (5) except that also here sixteen different DFT size $N_{\mathrm{T}}/2$ based beams are used, i.e.,

$$\mathbf{W} = \begin{bmatrix} \tilde{\mathbf{w}} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{w}} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ \alpha & -\alpha \end{bmatrix}, \quad \alpha \in \{1, j\}, \quad \tilde{\mathbf{w}} \in G_q^{(1,4)} \tag{7}$$

The codebook design in (6) and (7) is referred to as a grid of beam (GOB) design and incurs a very limited overhead yet provides excellent performance for the cross-pole as well as the co-pole, as the simulation results will show. Therefore, the same design is suitable for both antenna configurations and the precoders for one configuration are re-used in the other so as to avoid as signaling overhead increase. The number of precoders for rank 1 and 2 are given in Table 1.

Observation

- *Same design of the two codebooks is suitable for ULA as well as cross-pole.*

Table 1: Codebook size per rank for each of the two precoders for the GOB design.

	$\#\mathbf{W}^{(1)}$	$\#\mathbf{W}^{(2)}$
Rank=1	16	4
Rank=2	16	2

4. Reporting Concept

Rel-8 supports two different types of CSI reports – aperiodic reports on PUSCH triggered by an UL grant and periodic reports on PUCCH. Since aperiodic reports are dynamically scheduled, rather large payloads and hence detailed reports can be supported. In contrast, PUCCH is efficient in terms of requiring limited resources but at the same time constitutes a very narrow data pipe for which considerably coarser reports are more appropriate.

Typically, not that many UEs are simultaneously receiving data in a subframe. But many UEs can be in an active state simultaneously. Letting all UEs in active state transmit detailed CSI reports would not be feasible from an overhead perspective, not to mention from a DRX point of view. Such an operation would be highly inefficient considering the few UEs that actually simultaneously receive data. Instead, coarse reports in time and frequency are fed back by utilizing PUCCH based reporting which requires relatively little resources and can thus sustain a large number of UEs provided the period is set long enough. This only gives a rough measure of the CSI so the coarse PUCCH based reports are therefore complemented with more detailed CSI reports a short while after the UE starts being scheduled. The Rel-8 CSI feedback concept is highly efficient and scalable with the number of active UEs in the cell and thus remains an attractive solution also for Rel-10.

Observation

- *Rel-8 aperiodic/periodic CSI feedback concept remains highly efficient also for Rel-10*
    - *Detailed CSI feedback on aperiodic CSI on PUSCH*
    - *Coarse CSI feedback on the narrow data pipe of PUCCH*

In Rel-8, coarse reporting on PUCCH refers to time as well as frequency. In other words, wideband reporting with a long period is typically used. The coarseness may in Rel-10 be extended also to the spatial domain by subsampling the precoder codebook(s) so that both precoders $\mathbf{W}^{(1)}$ and $\mathbf{W}^{(2)}$ fit in the same subframe. A simple approach would for example be to fix the second precoder $\mathbf{W}^{(2)}$ for PUCCH. With the above GOB design, this means only 4 bits for the wideband precoder $\mathbf{W}^{(1)}$ would be needed, thus ensuring that the coverage of the PUCCH reporting is as good as in Rel-8 for 4 Tx.

Proposal

- *Perform subsampling of the codebook(s) on PUCCH for maintaining the coverage of CSI on PUCCH which anyway can only support coarse reports*

In RAN1 #60bis, it was agreed that $\mathbf{W}^{(1)}$ and $\mathbf{W}^{(2)}$ can be transmitted in the same or different subframes on PUCCH while for PUSCH this issue was left for FFS. The idea would be to save overhead by transmitting $\mathbf{W}^{(1)}$ less seldom than $\mathbf{W}^{(2)}$ since $\mathbf{W}^{(1)}$ targets the long-term correlation properties of the channel.

The aperiodic on PUSCH can handle rather large payloads and the single reported wideband $\mathbf{W}^{(1)}$ requires much fewer bits than the per subband reporting of $\mathbf{W}^{(2)}$. Because of this, there appears to be little reason for only transmitting $\mathbf{W}^{(1)}$ in an aperiodic PUSCH report, particularly considering that it is expensive signaling-wise to devote an uplink grant for triggering such a coarse report.

The reporting on PUCCH can easily be handled in a robust manner using the above described concept of subsampling codebooks and thus transmit both $\mathbf{W}^{(1)}$ and $\mathbf{W}^{(2)}$ in the same subframe.

Proposal

- *Possible to report $\mathbf{W}^{(1)}$ and $\mathbf{W}^{(2)}$ in the same subframe using subsampling of codebooks to maintain control signaling coverage*

Reporting of $\mathbf{W}^{(1)}$ and $\mathbf{W}^{(2)}$ in different subframes on the other hand is challenging considering

- The risk for error propagation
- Complicated operation with respect to RI which is already reported in a separate subframe in Rel-8

Typically $\mathbf{W}^{(1)}$ would be updated at a slow rate, for example every 50-100 ms, while $\mathbf{W}^{(2)}$ would be updated considerably faster, around 10 ms. But if $\mathbf{W}^{(1)}$ is in error because of an error in receiving the PUCCH, 50-100 ms worth of PUCCH reports are automatically lost since the reports for $\mathbf{W}^{(2)}$ and CQI are dependent on first

receiving $\mathbf{W}^{(1)}$. Since the impact of this error event is substantial, $\mathbf{W}^{(1)}$ needs to be heavily protected so that the error event rarely occurs. The situation is a bit similar to the reporting of RI in Rel-8 which is also reported in a separate subframe and is heavily protected. A major difference is however that the RI is only two bits which allows a strong protection while the payload size for $\mathbf{W}^{(1)}$ could be substantially larger thus preventing a strong protection, particularly on the cell-edge where the CSI reports are most useful. Subsampling of the $\mathbf{W}^{(1)}$ codebook again provides a remedy since this allows the payload size for $\mathbf{W}^{(1)}$ to be brought down to a similar level as for Rel-8.

Observation

- *Reporting* $\mathbf{W}^{(1)}$ *and* $\mathbf{W}^{(2)}$ *in different subframes on PUCCH appears tricky because of error propagation and interaction with RI reporting*
    - *Would a report be dependent on three different subframes? – one for RI, one for* $\mathbf{W}^{(1)}$ *and one for* $\mathbf{W}^{(2)}$ *?*
- *Remedies include performing subsampling of* $\mathbf{W}^{(1)}$ *codebook so that it achieves similar coverage as RI in Rel-8*

5. SU-MIMO Simulation Results

To assess the performance of the GOB design, system level simulations for an urban macro environment have been conducted. Three different designs where evaluated; the GOB design, the Samsung codebook in [13] , and the Huawei codebook in [12] . In all cases, a CSI feedback report contains both $\mathbf{W}^{(1)}$ and $\mathbf{W}^{(2)}$, where the latter is reported per subband. The underlying assumption is that aperiodic CSI on PUSCH is being used to simultaneously transmit all the precoders. An overview of the reporting mechanism and associated overhead for the three schemes is given in Table 3. Both high and lower angular spread scenarios have been considered and the results are found in Table 4 and Table 5 for cross-pole and co-pole, respectively. It is seen that the GOB design outperforms the other two codebooks in almost all cases except for two cases where the Huawei and Samsung codebooks are marginally better in cell throughput but on the other hand performs worse on cell edge. From Table 3 we see that it compares favorably in terms of report sizes as well. This supports the thinking in the way forward on refined feedback of exploiting grid of beams using at least a beam oversampling factor of two.

Proposal

- *For 8 Tx rank 1 and 2, introduce support for GOB design as in (6) and (7)*
    - *Inline with way forward on feedback refinement and supports 16 beams for ULA and 16 beams for cross-pole*
    - $\mathbf{W}^{(1)}$ *block-diagonal with 4-element beamformer* $\tilde{\mathbf{w}}$ *taken from 16 DFT based precoders*
    - $\mathbf{W}^{(2)}$ *chosen from 2 Tx codebook as baseline*

$$\mathbf{W}^{(1)} = \begin{bmatrix} \tilde{\mathbf{w}} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{w}} \end{bmatrix}, \quad \mathbf{W}^{(2)} \in \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ j \end{bmatrix}, \begin{bmatrix} 1 \\ -j \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix} \right\}$$

    - *4 bits for wideband* $\mathbf{W}^{(1)}$ *and 2 and 1 bits per subband for* $\mathbf{W}^{(2)}$ *for rank 1 and 2, respectively*

Table 2: System level simulation assumptions.

Parameter	Assumption
Number of cells	57
Deployment model	Hex grid, 3 sector sites
Inter site distance	500 m
Average number of UEs per cell	10
Traffic model	Full buffer
UE speeds of interest	3 km/h
Bandwidth	10 MHz
Carrier frequency	2 GHz
Control OFDM symbols per RB pair	3
Max number of HARQ retransmissions	5
Channel model	SCME Urban Macro
Pathloss model	128,1 + 37,6 log10(R) dB, (R in km)
Transmit power	40 W
BS antenna configuration	8 Tx using two different alternatives: Alt 1. Four closely spaced ±45° cross-poles with 0.5 λ separation Alt 2. ULA with 0.5 λ separation and vertical polarization
UE antenna configuration	2 RX: cross-polarized 0°/90°, 0.5 λ separation
Receiver	MMSE with no inter-cell interference suppression
Scheduler	Proportional fair in time and frequency
ACK/NACK based outer loop link adaptation adjustment	Yes: target BLER=10%
Number of RBs per subband	8
Feedback CQI delay	6 ms
CQI reporting periodicity	5 ms

Table 3: Comparison of report sizes for one aperiodic report on PUSCH.

	Bits per PMI (rank1, rank 2)		Reporting type		Precoder report size in bits for 20 MHz (rank1, rank2)
	$\mathbf{W}^{(1)}$	$\mathbf{W}^{(2)}$	$\mathbf{W}^{(1)}$	$\mathbf{W}^{(2)}$	
GOB	4, 4	2, 1	wideband	subband	30, 17
Samsung	0, 0	4, 4	wideband	subband	52, 52
Huawei	4, 4	2, 2	wideband	subband	30, 30

Table 4: System level performance comparison of GOB design, Samsung codebook [13] and Huawei codebook [12] for cross-pole antenna setup. Relative gains compared to GOB design within parenthesis.

	Cross-pole (Alt 1)			
	15° angular spread		8° angular spread	
	Cell throughput [bps/Hz]	5-percentile user throughput [bps/Hz]	Cell throughput [bps/Hz]	5-percentile user throughput [bps/Hz]
GOB	2.7463 (0%)	0.0678 (0%)	2.7954 (0%)	0.069 (0%)
Samsung [13]	2.6989 (-1.73%)	0.0648 (-4.42%)	2.7263 (-2.47%)	0.065 (-5.80%)
Huawei [12]	2.7484 (0.08%)	0.0655 (-3.39%)	2.7377 (-2.06%)	0.0641 (-7.10%)

Table 5: System level performance comparison of GOB design, Samsung codebook [13] and Huawei codebook [12] for co-polarized antenna setup. Relative gains compared to GOB design within parenthesis.

	Co-pole (Alt 2)			
	15° angular spread		8° angular spread	
	Cell throughput [bps/Hz]	5-percentile user throughput [bps/Hz]	Cell throughput [bps/Hz]	5-percentile user throughput [bps/Hz]
GOB	2.6047 (0%)	0.072 (0%)	2.5758 (0%)	0.0718 (0%)
Samsung [13]	2.5835 (-0.81%)	0.0687 (-4.58%)	2.5916 (0.61%)	0.0699 (-2.65%)
Huawei [12]	2.5227 (-3.15%)	0.0643 (-10.69%)	2.5598 (-0.62%)	0.0638 (-11.14%)

6. Conclusions

Based on the discussion and simulation results above about feedback type and codebook design issues for SU-MIMO and MU-MIMO we propose the following:

- *A precoder for a subband is obtained as a matrix multiplication of two precoder matrices*
- *The inner precoder* $\mathbf{W}^{(1)}$ *(i.e., closest to the channel) should typically be a tall matrix so that dimension reduction is offered.*
- *Codebooks for inner and outer precoder designed so that overall precoder fully utilizes all PAs, i.e., for the overall precoder it holds that*

$$\left[\mathbf{W}\mathbf{W}^{*}\right]_{mm} = \kappa, \quad m = 1, \cdots, N_{\mathrm{T}}$$

- *Introducing additional precoders not fulfilling full PA utilization FFS*
- *Perform subsampling of the codebook(s) for maintaining the coverage of CSI on PUCCH which anyway can only support coarse reports*

- *Possible to report* $\mathbf{W}^{(1)}$ *and* $\mathbf{W}^{(2)}$ *in the same subframe using subsampling of codebooks to maintain control signaling coverage*

- *For 8 Tx rank 1 and 2, introduce support for GOB design as in (6) and (7)*
  - *Inline with way forward on feedback refinement and supports 16 beams for ULA and 16 beams for cross-pole*
  - $\mathbf{W}^{(1)}$ *block-diagonal with 4-element beamformer* $\tilde{\mathbf{w}}$ *taken from 16 DFT based precoders*
  - $\mathbf{W}^{(2)}$ *chosen from 2 Tx codebook as baseline*

$$\mathbf{W}^{(1)} = \begin{bmatrix} \tilde{\mathbf{w}} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{w}} \end{bmatrix}, \quad \mathbf{W}^{(2)} \in \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ j \end{bmatrix}, \begin{bmatrix} 1 \\ -j \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix} \right\}$$

  - *4 bits for wideband* $\mathbf{W}^{(1)}$ *and 2 and 1 bits per subband for* $\mathbf{W}^{(2)}$ *for rank 1 and 2, respectively*

7. References

[1] R1-095102, "Way forward on UE feedback".

[2] R1-093997, "Implicit Feedback in Support of Downlink MU-MIMO", Texas Instruments.

[3] R1-094695, "Extension to Rel. 8 PMI feedback by adaptive codebook", Huawei.

[4] TR36.814 V1.2.2, "Further Enhancements for E-UTRA: Physical Layer Aspects".

[5] R1-100117, "Further Discussions on Feedback Framework for SU/MU-MIMO", Samsung.

[6] R1-100794, "Proposal for Rel-10 Feedback Framework", Samsung, Texas Instruments, LG Electronics, NEC, RIM.

[7] R1-100798, "Proposal for Rel-10 Feedback Framework", Ericsson, ST-Ericsson, Huawei, Panasonic, Texas Instruments.

[8] R1-100251, "Extensions to Rel-8 type CQI/PMI/RI feedback using double codebook structure", Huawei.

[9] R1-101741, "On DL Single-Cell MU-MIMO", Ericsson, ST-Ericsson.

[10] R1-101683, "Way Forward for Rel-10 Feedback Framework".

[11] R1-102560, "Refinement on UE Feedback", Ericsson, NTT DoCoMo, Texas Instruments, Nokia, Nokia Siemens Networks, LG Electronics, ST-Ericsson, AT&T, Sharp, Mitsubishi Electric, Huawei.

[12] R1-101959, "Further Results of DL 8Tx Codebook", Huawei.

[13] R1-102199, "Design and Performance Evaluation of 8 Tx Codebook", Samsung.