PDSCH & PUSCH

DMRS for PDSCH and PUSCH

Since PDSCH and PUSCH are shared channels, there are lots of similarity in the channel processing. Also, the DMRS for both is lot similar.

The DMRS structure is shown above. As we see, the DMRS position varies in time and frequency axes. As we see in figure, DMRS occupies 6 x 3 = 18 REs in a PRB. The No. of PRBs taken by PDSCH / PUSCH, is same as the No. of PRBs taken by DMRS. Say for example, if PDSCH / PUSCH takes 100 PRBs, then DMRS will also be transmitted in all 100 PRBs.

We've already seen

- What is the functionality of DMRS.
- What is the benefit of transmitting more number of DMRS in time or frequency.
- Why do we transmit less number of DMRS.
- What is the Overhead

and so on...

In this section, we'll specifically understand DMRS for PDSCH and PUSCH, in the context of

- DMRS allocation in Frequency
- DMRS allocation in Time
- DMRS allocation in Space

Under DMRS allocation in Frequency, there are two types of configuration

① DMRS config type 1
  - where, the DMRS goes on alternative subcarriers (say 1, 3, 5, 7, 9, 11, ...) Fig. ①
  - most commonly used
  - being used in initial access

② DMRS config type 2
  - where, the DMRS goes in two consecutive subcarriers (say 1,2 ; 7,8 ; 13,14 ; 19,20) Fig ④
  - Overhead is less here (i.e) 4 out of 12 subcarriers carries DMRS. whereas in DMRS config type 1, 6 out of 12 carries DMRS.

How do we decide which DMRS config type to use?

- Config 1 has higher DMRS density (6/12) (More Overhead)
  config 2 has lower DMRS density (4/12) (Less overhead)

- When we want to have less Overhead, we can use config type 2.

- High density comes with <u>high reliability</u> (config type 1).
  If the channel is already good and <u>less reliability</u> is sufficient, then we can use config type 2.

- Multiple CDM groups used to multiplex DMRS from different layers (spatial multiplexing). In config type 1, since it is taking alternative subcarriers, in the first 6 we have one CDM group, and in the remaining 6 we have second CDM group. In config type 2, we can have 3 CDM groups.

Since there are more CDM groups in config type 2, we can have more no. of layers multiplexed. This is another advantage in config type 2.

- config type 1 is always used in initial access, since it is more reliable

---

The number of DMRS symbols in a slot can be 1, 2, 3 or 4. How do we decide how many DMRS symbols we need?

If the channel is changing really fast in time, (i) UE is moving very fast (ii) High doppler, then we want more reliable channel estimation. In that case, we need to have more DMRS in time. We can have upto 4 symbols.

Also, if we have FR2 deployment (mm wave), we assume that its more or less stationary channel (not varying very fast) and coverage is also very less. In this case, we require less DMRS symbols in time. Compared to FR1 or outdoor deployment.

More number of DMRS symbols also comes with more overhead.

Note that, DMRS and data can also interleave, as shown in figure. (i.e.) DMRS and Data can be multiplexed.

Apart from these, there is something called "Mapping Type", which is there for both PDSCH and PUSCH. There are two different mapping types:

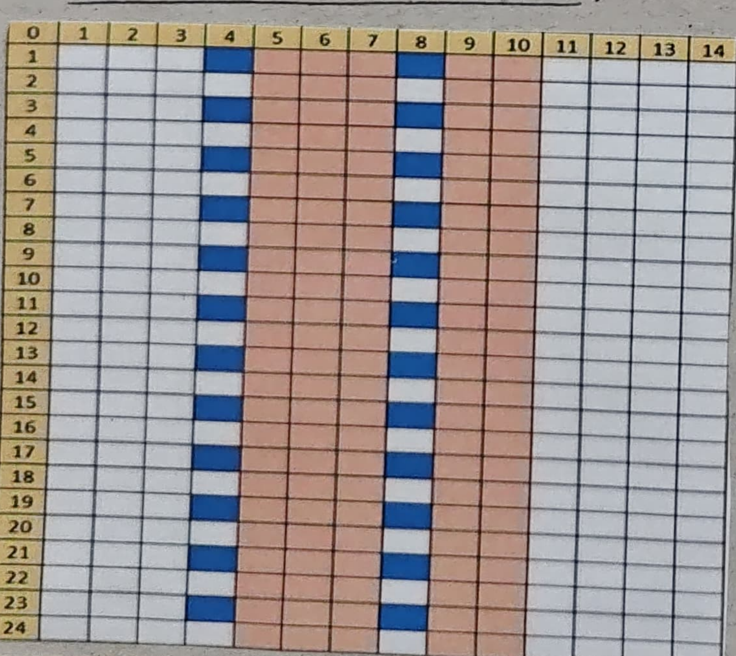
- Mapping Type A
- Mapping Type B

For Mapping Type A, the reference for DMRS is the "Start of Slot". For Mapping Type B, the reference for DMRS is the "Start of PDSCH or PUSCH".

In Mapping Type A, DMRS is always on the 2nd or 3rd symbol. In Mapping Type B, DMRS is always on the 1st symbol.

Note:
Mapping Type B scenario is used for mini-slot, where PDSCH can start anywhere in the slot, and it can take maximum 7 OFDM symbols, out of which 1 or 2 symbols are for DMRS.

· Front Loaded DMRS



As the name implies, the DMRS lies in the front of the data.

Sym. 4 to 10 → PDSCH/PUSCH allocation

DMRS is on Sym. 4 and 8.

The purpose of Front Loaded DMRS is that, at the receiver end, it is used to do Fast Channel Estimation to save time. How?

After receiving data at the receiver end, the channel estimation is performed using the DMRS at symbol 4, and by the time symbols 5, 6, 7 are received, the estimates may be applied on the PDSCH / PUSCH data in symbols 5, 6, 7.

Lets say, if the PDSCH / PUSCH starts from symbol 2, and the DMRS is at symbol 4, in this case, the data at symbols 2, 3 have to wait till the channel estimation is done using DMRS at symbol 4.

Summary :

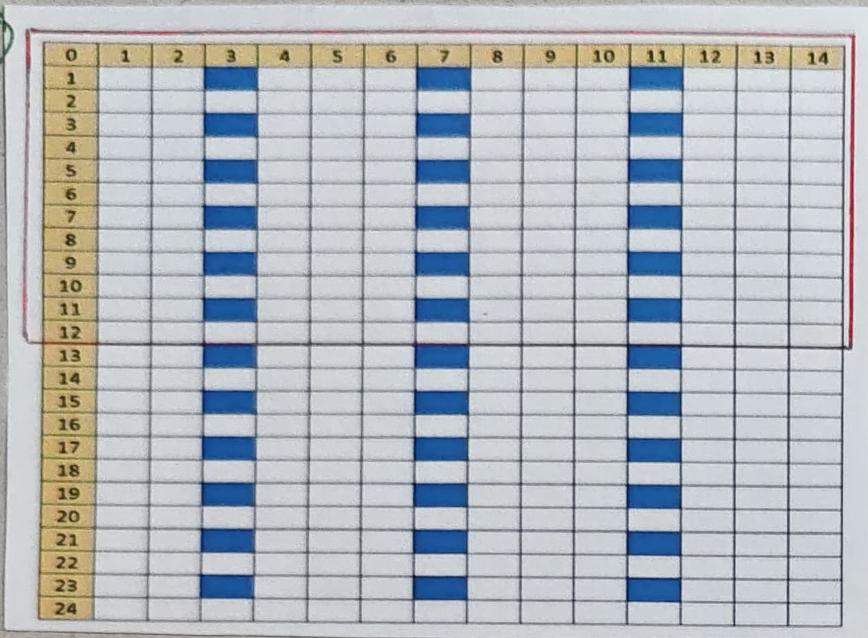
If we have the DMRS in front, then we can do the channel estimation, keep the estimates ready and start applying when the data comes.

Fig. ① Single Symbol DMRS → DMRS takes only one symbol

Fig ② Double symbol DMRS → DMRS takes two consecutive symbols

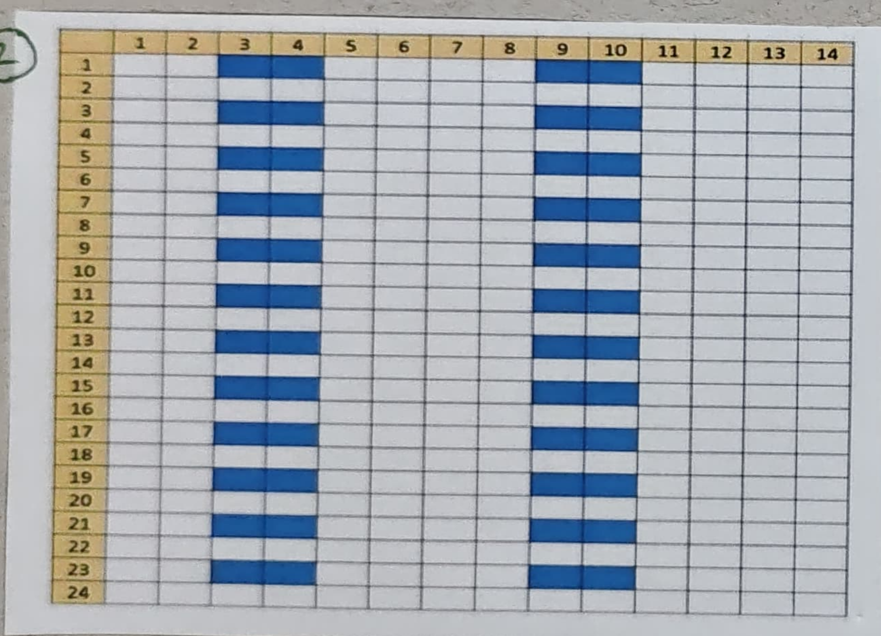
Use case :

In single symbol DMRS, we can have only limited number of layers multiplexed. (we'll see this again)



① If it is Single symbol DMRS, Config Type 1, then we can multiplex 4 layers per CDM group.

④ And, if it is config type 2, then we can multiplex 6 Layers or 6 antenna ports.

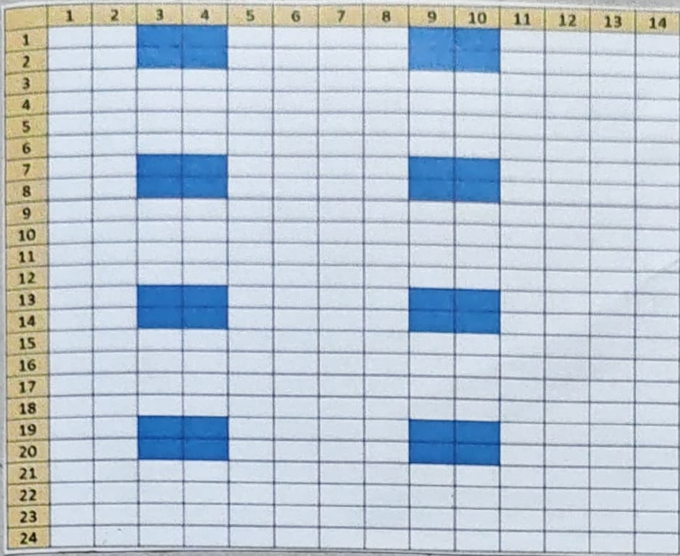


② If it is Double Symbol DMRS, Config Type 1, then we can multiplex 8 Layers or antenna ports.

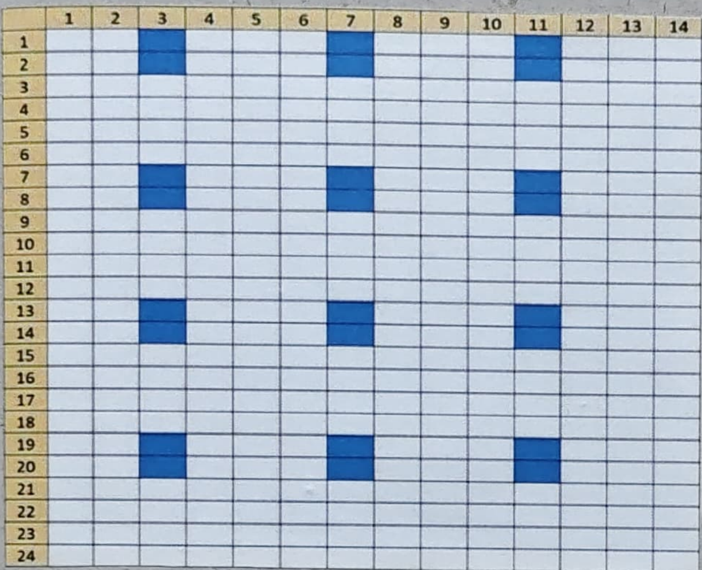
③ And, if it is config type 2, then we can multiplex 12 Layers or antenna ports.

So, basically, our orthogonal DMRS doubles when we have Double symbol DMRS. This concept highly relies on the channel estimation, where the channel remains same in the consecutive OFDM symbols.

(3)



(4)



This is all about ~~the~~ different frequency and time allocations of DMRS.

Lets now understand the DMRS signal generation. The DMRS generation and mapping is more or less same whether it is PDSCH or PUSCH.

Following snippet shows the sequence generation, which is PN sequence based and has QPSK modulation.

7.4.1.1 Demodulation reference signals for PDSCH

7.4.1.1.1 Sequence generation

QPSK

The UE shall assume the sequence $r(n)$ is defined by

$$r(n)=\overbrace{\frac{1}{\sqrt{2}}(1-2\cdot c(2n))+j\frac{1}{\sqrt{2}}(1-2\cdot c(2n+1))}.$$

PN Sequence

$$c_{\text{init}}=\left(2^{17}\left(N_{\text{symb}}^{\text{slot}}n_{s,f}^{\mu}+l+1\right)\left(2N_{\text{ID}}^{n_{\text{SCID}}}+1\right)+2N_{\text{ID}}^{n_{\text{SCID}}}+n_{\text{SCID}}\right)\bmod 2^{31}$$

The C_{init} is defined as above, which is based on

$l \rightarrow$ Symbol number

$n_{SCID} \rightarrow$ 0 or 1

$\rightarrow$ During Initial access, $n_{SCID} = \{0\}$. So we have only N^0_{ID}.

$\rightarrow$ For DCI 1_1, $n_{SCID} = \{0,1\}$. So we have N^0_{ID} and N^1_{ID}.

$\rightarrow$ This is an Scrambling ID, which can be used with cell specific identity. (Say, for some RRC messages or in initial access or SIB1). This can also be used as UE specific identity (provides option to have UE specific DMRS or broadcasted DMRS)

7.4.1.1.2 Mapping to physical resources

$$a_{k,l}^{(p,\mu)}=\beta_{\text{PDSCH}}^{\text{DMRS}}w_f(k')w_t(l')r(2n+k')$$

$$k=\begin{cases}4n+2k'+\Delta & \text{Configuration type 1}\\ 6n+k'+\Delta & \text{Configuration type 2}\end{cases}$$

$$k'=0,1$$

$$l=\bar{l}+l'$$

$$n=0,1,\ldots$$

The mapping of PDSCH / PUSCH DMRS is defined as above, where

$\beta^{DMRS}_{PDSCH} \rightarrow$ DMRS power, which can be tuned separately.

$w_f(k')$ and $w_t(l')$ $\rightarrow$ Factors to provide Orthogonality in time and frequency.

$r(2n+k')$ $\rightarrow$ Generated sequence

Table 7.4.1.1.2-1: Parameters for PDSCH DM-RS configuration type 1.

* 2 CDM groups

p	CDM group λ	Δ	$w_f(k')$		$w_t(l')$	
			$k'=0$	$k'=1$	$l'=0$	$l'=1$
1000	0	0	+1	+1	+1	+1
1001	0	0	+1	-1	+1	+1
1002	1	1	+1	+1	+1	+1
1003	1	1	+1	-1	+1	+1
1004	0	0	+1	+1	+1	-1
1005	0	0	+1	-1	+1	-1
1006	1	1	+1	+1	+1	-1
1007	1	1	+1	-1	+1	-1

Table 7.4.1.1.2-2: Parameters for PDSCH DM-RS configuration type 2.

* 3 CDM groups

p	CDM group λ	Δ	$w_f(k')$		$w_t(l')$	
			$k'=0$	$k'=1$	$l'=0$	$l'=1$
1000	0	0	+1	+1	+1	+1
1001	0	0	+1	-1	+1	+1
1002	1	2	+1	+1	+1	+1
1003	1	2	+1	-1	+1	+1
1004	2	4	+1	+1	+1	+1
1005	2	4	+1	-1	+1	+1
1006	0	0	+1	+1	+1	-1
1007	0	0	+1	-1	+1	-1
1008	1	2	+1	+1	+1	-1
1009	1	2	+1	-1	+1	-1
1010	2	4	+1	+1	+1	-1
1011	2	4	+1	-1	+1	-1

Tables define $w_f(k')$ and $w_t(l')$ for config type 1 and config type 2. Lets consider the CDM group 0 used by ports 1000 and 1001.

Note, for single symbol, we use $l'=0$. And for double symbol, we use both $l'=0$ and $l'=1$.

When $l'=0$, we get $w_t(l')=+1$ for both the ports (time domain).

And, in frequency domain, we get

$$w_f(k') = \begin{cases} +1, +1 & \text{for port 1000} \\ +1, -1 & \text{for port 1001} \end{cases}$$

This provides an orthogonality. (ie) [+1, +1] is orthogonal to [+1, -1].

How each DMRS group supports two orthogonal cover codes (OCCs) in 5G NR DMRS configuration type 1 (single-symbol case).

◆ **1. Why "orthogonal cover codes" (OCCs) are needed**

When multiple **DMRS ports** (for different MIMO layers) are transmitted **in the same time–frequency RE group**, they would interfere — unless they are made **orthogonal**.

To achieve this, 5G applies a short **orthogonal code sequence** (OCC) to each port's DMRS within a CDM (Code Division Multiplexing) group.

At the receiver, correlating with the same OCC retrieves the desired layer's DMRS while nulling out others — exactly like CDMA but within a small local RE set.

◆ **2. CDM (Code Division Multiplexing) group structure**

For **Configuration Type 1**, DMRS REs are arranged **every 6 subcarriers** per OFDM symbol.

- These 6 REs form **one CDM group**.
- There are **two such groups per RB (12 subcarriers)**.

Each CDM group (6 REs) can carry **2 orthogonal DMRS ports** via OCCs.

◆ **3. The Orthogonal Cover Codes (OCCs)**

The OCCs are **Walsh-type sequences** of length 2:

$$OCC_1 = [1,1]$$
$$OCC_2 = [1,-1]$$

These are applied **across two consecutive REs** in the time or frequency dimension (depending on DMRS mapping).

Because these codes are orthogonal:

$$[1,1].[1,-1]^T = 0$$

the receiver can separate the DMRS of one port from the other using correlation.

◆ **4. How 2 OCCs → 2 ports per CDM group**

Let's say one DMRS group = 6 subcarriers (k = 0, 1, 2, 3, 4, 5) within one OFDM symbol.

- **Port 0** uses OCC = [1, 1]
- **Port 1** uses OCC = [1, −1]

Each port's DMRS sequence is multiplied (spreading operation) by its OCC pattern. The two resulting DMRS sequences **occupy the same 6 subcarriers** but are *code-orthogonal*.

Thus, the UE can despread them to recover each one separately.

◆ **5. At the receiver (UE) side**

The UE knows which OCC each port uses.

To recover a given layer's DMRS:

1. Multiply the received signal by the **conjugate** of the corresponding OCC sequence.
2. Sum over the REs in the CDM group.
3. Due to orthogonality, interference from the other DMRS port cancels out (dot product = 0).

◆ 6. Mathematical Illustration (Simplified)

Let the DMRS sequences for ports 0 and 1 before OCC be (d_0) and (d_1).

After OCC application (length 2):

RE index	OCC(Port0)	OCC(Port1)	Transmitted signal
RE_1	$+1$	$+1$	$d_0 + d_1$
RE_2	$+1$	-1	$d_0 - d_1$

At the receiver:

- Multiply REs by [1, 1] → recover ($2d_0$)
- Multiply REs by [1, –1] → recover ($2d_1$)

Orthogonal separation achieved ✔

◆ 7. Summary — why 2 OCCs per group = 2 DMRS ports

Concept	Explanation
DMRS group	6 subcarriers (1 CDM group)
OCC set	Two orthogonal 2-length Walsh codes [1,1] and [1,–1]
Purpose	Multiplex 2 DMRS ports (for 2 MIMO layers) in same group
Result	Each group supports 2 ports → 2 groups × 2 = 4 ports per RB (config type 1)

☑ **In one line:**

Each 6-subcarrier DMRS group in 5G NR Configuration Type 1 uses two orthogonal 2-length Walsh cover codes ([1,1] and [1,–1]) to multiplex two DMRS ports within the same RE group — thus enabling up to 4 ports (2 CDM groups × 2 OCCs) for single-symbol DMRS.

And, considering the CDM group 0 used by ports 1004 and 1005, when $l' = 1$, we have

$$w_t(l') = -1$$, for both the ports.

So, we need the second symbol to have orthogonality. And, this is where the Double Symbol DMRS comes into picture, to have orthogonal DMRS.

The same concept applies to configuration type 2 as well, where we have 3 CDM groups.

Table 7.4.1.1.2-3: PDSCH DM-RS positions $\bar{l}$ for single-symbol DM-RS.

l_d in symbols	DM-RS positions $\bar{l}$							
	PDSCH mapping type A dmrs-AdditionalPosition				PDSCH mapping type B dmrs-AdditionalPosition			
	pos0	pos1	pos2	pos3	pos0	pos1	pos2	pos3
2	-	-	-	-	l_0	l_0		
3	l_0	l_0	l_0	l_0	-	-		
4	l_0	l_0	l_0	l_0	l_0	l_0		
5	l_0	l_0	l_0	l_0	-	-		
6	l_0	l_0	l_0	l_0	l_0	l_0, 4		
7	l_0	l_0	l_0	l_0	l_0	l_0, 4		
8	l_0	l_0, 7	l_0, 7	l_0, 7	-	-		
9	l_0	l_0, 7	l_0, 7	l_0, 7	-	-		
10	l_0	l_0, 9	l_0, 6, 9	l_0, 6, 9	-	-		
11	l_0	l_0, 9	l_0, 6, 9	l_0, 6, 9	-	-		
12	l_0	l_0, 9	l_0, 6, 9	l_0, 5, 8, 11	-	-		
13	l_0	l_0, l_1	l_0, 7, 11	l_0, 5, 8, 11	-	-		
14	l_0	l_0, l_1	l_0, 7, 11	l_0, 5, 8, 11	-	-		

This Table defines the No. of symbols (data) for PDSCH. Similar table is available for PUSCH as well.

- l_0 is the first symbol.
- For Mapping type A, l_0 can be 2 or 3.
- For Mapping type B, l_0 depends on when the PDSCH/PUSCH starts.
- Maximum No. of DMRS can be 4.

This table says that, when the No. of PDSCH / PUSCH data symbols are 14, and if it is Single symbol DMRS, and Additional positions are 3, in that case, the DMRS will be l_0, 5, 8, 11 where l_0 may be 2 or 3.

In Mapping Type B, since the max. No. of Symbols are 7, so the entries marked are not applicable. And since there are only 7 symbols, we don't need to have more than 2 DMRS symbols, so the entries marked are not applicable.

Table 7.4.1.1.2-4: PDSCH DM-RS positions $\bar{l}$ for double-symbol DM-RS.

l_d in symbols	DM-RS positions $\bar{l}$					
	PDSCH mapping type A			PDSCH mapping type B		
	dmrs-AdditionalPosition			dmrs-AdditionalPosition		
	pos0	pos1	pos2	pos0	pos1	pos2
<4				-	-	
4	l_0	l_0		-	-	
5	l_0	l_0		-	-	
6	l_0	l_0		l_0	l_0	
7	l_0	l_0		l_0	l_0	
8	l_0	l_0		-	-	
9	l_0	l_0		-	-	
10	l_0	l_0, 8		-	-	
11	l_0	l_0, 8		-	-	
12	l_0	l_0, 8		-	-	
13	l_0	l_0, 10		-	-	
14	l_0	l_0, 10		-	-	

This table is for Double Symbol DMRS. Let consider $l_0 = 2$, and the No. of PDSCH/PUSCH data symbols are 14, for this case, we have DMRS on 2, 3 and 10, 11.

Note: In PUSCH, we also have Transform Precoding: If Transform Precoding is enabled, then we don't have PN sequence based transmission. Instead, we have Zad off-chu based DMRS sequence generation. Here, the signal generation is little different, but the mapping remains the same. Instead of n_{ID} / n_{SCID}, we have group hopping / sequence hopping in case of Transform Precoding.

The purpose of having few different IDs (n_{ID}, n_{SCID}, group hopping, sequence hopping) is to minimize Inter-Cell-Interference

PDSCH / PUSCH : Physical Layer Processing

For the shared channel, the processing is almost same.
(a) The Physical Layer receives data from MAC Layer, to all the way it is transmitted, most of the modules remains the same. So, we'll study both PDSCH and PUSCH processing in this single section. And wherever there is differences, we'll point out.

In PDSCH, a single chain carries 2 Transport Blocks, whereas in PUSCH, a single chain carries 1 TB.

When we receive TB from MAC, in the Shared Channels, the TB size have a large range (can be very small or it can be very large). So, based on the size of TB, we decide which CRC to attach. If TB size > 3824, then we use CRC24A, else if TB size ≤ 3824, then we use CRC16. The larger the CRC size, the chances of detecting error is high. So, for a large TB, it is good to use large CRC.

After CRC attachment, we divide the TB into multiple Code Blocks (CB). This division helps in reducing the channel decoder complexity. How TB divided into multiple CBs?
First we find out LDPC Base Graph (BG). There are two BG. Basically, the Base Graph is a matrix that is used for channel encoding:

BG 1 → 46 x 68 matrix

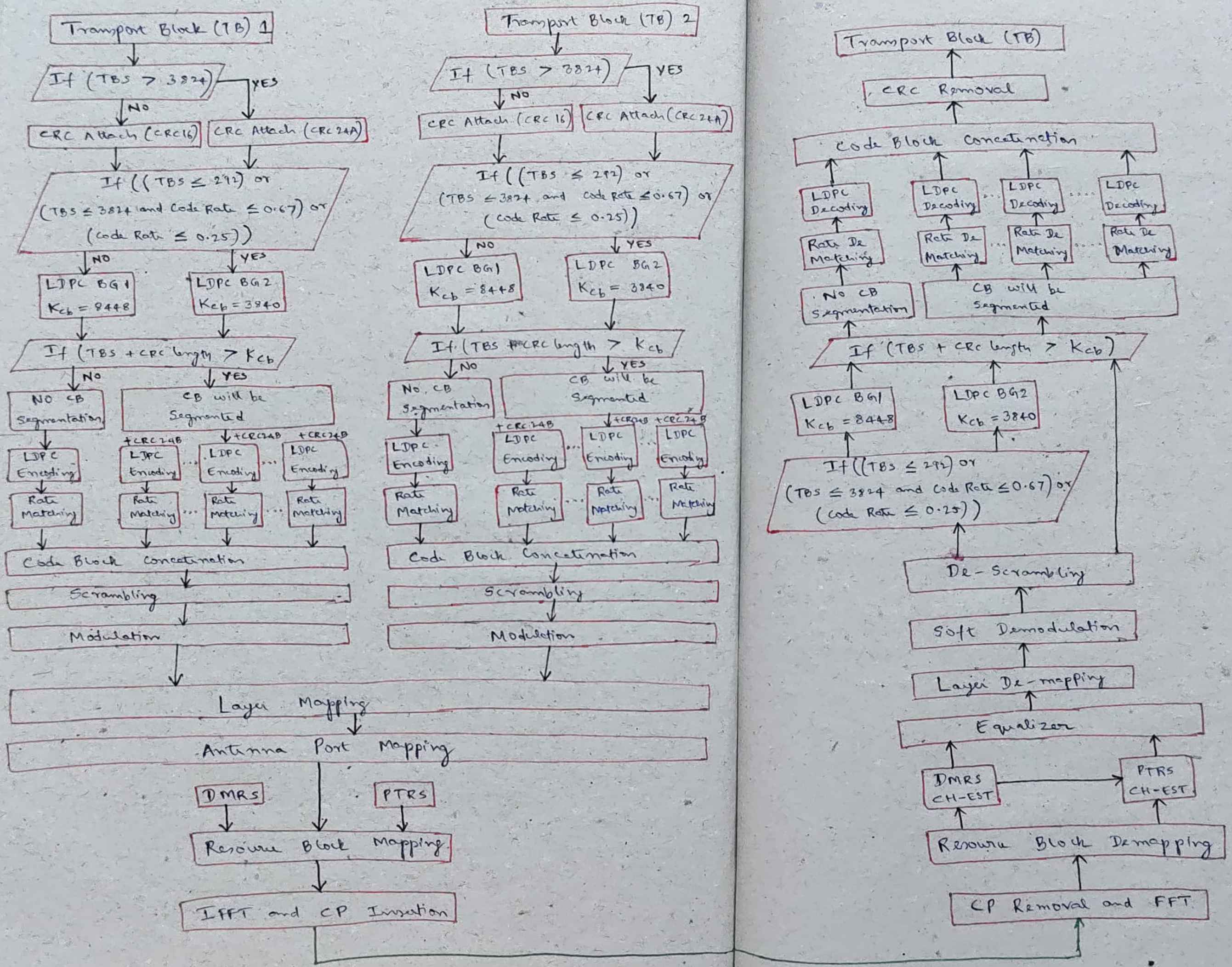
BG 2 → 42 x 52 matrix

BG1 is used for Larger TB, and

BG2 is used for shorter TB.

TRANSMITTER

RECEIVER



These matrices are expanded based on the parameter Z_c known as Lifting size, and becomes Low Density Parity Check matrix, where the no. of ones is very less compared to no. of zeros.

So, first we select the BG based on the following condition.

If ((TBS $\leq$ 292) or (TBS $\leq$ 3824 and $R \leq 0.67$) or ($R \leq 0.25$))

use BG 2 (Code Block size, $K_{cb} = 3840$)

Else

use BG 1 (code Block size, $K_{cb} = 8448$)

Having chosen the BG, now we do code block segmentation based on the following condition.

If (TBS + CRC (16 or 24) > K_{cb})

do codeblock segmentation

Else

No codeblock segmentation.

And, the No. of CBs can be calculated based on TBS and CB CRC size. Consider TB + CRC is 10,000 bits, BG 1 is chosen hence CB size is 8448, then

$$\text{No. of CBs} = \left\lceil \frac{10,000}{8448 - 24} \right\rceil = 2$$

(ii) the 10,000 bits will be divided into 2 codeblocks, and the size of each code block is basically half in this case (i) 5000 + 24 (CB CRC). Code Block size is 5024. So basically, each CB has equal length.

So, after CB segmentation, we add CRC24B to each codeBlock. If there is no CB segmentation, then we don't each CB CRC. The purpose of adding CB CRC is that, at the receiver end, for each CB, we can check whether the CB is successfully decoded or not, and the feedback (ACK/NACK) can be provided for each CB.

Now, each of the segmented CB undergoes LDPC encoding, which has standard input sizes. The input length can be calculated based on the formula

$$K_b \times Z_c \geq \text{CB size}$$

where, $K_b \rightarrow$ No. of systematic columns

$$= \begin{cases} 22 & (\text{for BG1}) \\ 10 & (\text{for BG2}) \end{cases}$$

$$\Rightarrow 22 \times Z_c \geq 5024$$

Table 5.3.2-1: Sets of LDPC lifting size Z

Set index (i_{LS})	Set of lifting sizes (Z)
0	{2, 4, 8, 16, 32, 64, 128, 256}
1	{3, 6, 12, 24, 48, 96, 192, 384}
2	{5, 10, 20, 40, 80, 160, 320}
3	{7, 14, 28, 56, 112, 224}
4	{9, 18, 36, 72, 144, 288}
5	{11, 22, 44, 88, 176, 352}
6	{13, 26, 52, 104, 208}
7	{15, 30, 60, 120, 240}

Here, the minimum value of Z_c for which the above condition is True is 240.

$$\Rightarrow \boxed{Z_c = 240}$$

<u>Note</u>: The K_b value for BG2 can be 10/9/8/6, based on the TB + CRC size.

$$\text{(i) } K_b \text{ (for BG2)} = \begin{cases} 10, & \text{if TB+CRC} > 640 \\ 9, & \text{if TB+CRC} > 560 \\ 8, & \text{if TB+CRC} > 192 \\ 6, & \text{elsewhere.} \end{cases}$$

So, the input to the LDPC encoder is $22 Z_c$ (for BG1) or $10 Z_c$ (for BG2). And the output from the LDPC encoder is $66 Z_c$ (for BG1) or $50 Z_c$ (for BG2).

(ie) The input and output lengths are fixed for both the Base Graphs.

The Code Rate is calculated as

$$\text{For BG1}, \quad R = \frac{22 Z_c}{66 Z_c} = \frac{1}{3}$$

$$\text{For BG2}, \quad R = \frac{10 Z_c}{50 Z_c} = \frac{1}{5}$$

Now, the 5024 bits CB is given to LDPC encoder with input size $22 \times 240 = 5280$. The remaining bits $(5280 - 5024 = 256)$, we add <u>filler bits</u>. (NULL bits)

Followed by LDPC encoder is the Rate Matching. For PDSCH, it is always Limited Buffer Rate Matching (LBRM). And for PUSCH, we can choose among LBRM or Full Buffer Rate Matching (FBRM).

After Rate Matching, all the Code Blocks are concatenated sequentially. (ie) The output of LDPC encoder is Rate matched according to the Resources allocated, and then Concatenated sequentially.

After concatenation, we do Scrambling (Randomize the bits). The Scrambling is done, so that, for other UEs or BS, the data looks like Noise. The Scrambling sequence is initialized differently for PDSCH and PUSCH.

For PDSCH $\Rightarrow$ $\boxed{n_{RNTI} * 2^{15} + q \cdot 2^{14} + n_{ID}}$

where, $q \rightarrow$ code Word Index ($\because$ PDSCH has 2 codewords)

For PUSCH $\Rightarrow$ $\boxed{n_{RNTI} * 2^{15} + n_{ID}}$

C_{init}

where, $q \rightarrow$ Not applicable ($\because$ PUSCH has only one codeword).

PUSCH (carries UCI bits) are scrambled differently.

(ii) The bits carried in PUSCH are placed in a way in order to maximize the Euclidean distance of the modulated symbols. So, they're scrambled differently.

The data sequence is scrambled according to C_{init}.

Here, n_{RNTI} is the RNTI.

$n_{ID} \rightarrow$ Scrambling ID

In case, we have transmission other than DCI-1-0, the RNTI is C-RNTI, MCS-RNTI or CS-RNTI, the n_{ID} is Scrambling ID.

Otherwise, the n_{ID} is cell ID.

After scrambling, the bits goes through the modulation. (QPSK / 16-QAM / 64-QAM / 256-QAM). for PDSCH. If it is PUSCH, and the Transform Precoding is enabled, then we have ($\frac{\pi}{2}$ BPSK / QPSK / 16-QAM / 64-QAM / 256-QAM). If Transform Precoding is not enabled in PUSCH, then we have (QPSK / 16-QAM / 64-QAM / 256-QAM).

After the modulation, we do Layer Mapping. As discussed earlier, there are two code words / Transport Blocks in PDSCH, and each TB is processed independently. Both the codewords (in case of PDSCH) or single codeword (in case of PUSCH), we map them to the Layers.

$$\text{Max. No. of Layers} = \begin{cases} 4 & (\text{PUSCH}) \\ 8 & (\text{PDSCH}) \end{cases}$$

The Mapping is very straight forward. Let say, these are the Symbols after the modulation.

0	1	2	3	4	5	6	7	8	9	10	....			

If the mapping is on Single Layer, then it is a one-to-one mapping.

L0	0	1	2	3	4	5	6	....

If there are Two Layers, then the mapping is done as follows.

L0	0	2	4	6	8	....
L1	1	3	5	7	9	....

If there are Three Layers, then the mapping is done as follows.

L0	0	3	6	9	....
L1	1	4	7	10	....
L2	2	5	8	11	....

If there are Four Layers, then the mapping is done as follows.

L0	0	4	8	12	....
L1	1	5	9	13	....
L2	2	6	10	14	....
L3	3	7	11	15	....

But, when we have more than Four Layers (ie) in PDSCH, lets say in case of Five Layers, Codeword 0 will get 2 Layers, and Codeword 1 will get 3 Layers.

L0	0	2	4	6	8	....	Codeword 0
L1	1	3	5	7	9	....	
L2	0	3	6	9	12	....	Codeword 1
L3	1	4	7	10	13	....	
L4	2	5	8	11	14	....	

If the number of Layers is six, then in this case, Codeword 0 will get 3 layers, and codeword 1 will get 3 layers.

L0	0	3	6	9	....	Codeword 0
L1	1	4	7	10	....	
L2	2	5	8	11	....	
L3	0	3	6	9	....	codeword 1
L4	1	4	7	10	....	
L5	2	5	8	11	....	

If the No. of Layers are Seven, then codeword 0 will get 3 layers and codeword 1 will get 4 layers.

L0	0	3	6	9	....	Codeword 0
L1	1	4	7	10	....	
L2	2	5	8	11	....	
L3	0	4	8	12	....	codeword 1
L4	1	5	9	13	....	
L5	2	6	10	14	....	
L6	3	7	11	15	....	

If there are 8 layers, then codeword 0 will get 4 layers and codeword 1 will get 4 layers.

L0	0	4	8	12	....	Codeword 0
L1	1	5	9	13	....	
L2	2	6	10	14	....	
L3	3	7	11	15	....	
L4	0	4	8	12	....	Codeword 1
L5	1	5	9	13	....	
L6	2	6	10	14	....	
L7	3	7	11	15	....	

Next to Layer Mapping is the Antenna Port mapping and RB mapping (along with DMRS and PTRS). We'll study this in next section.

Now, the receiver chain is exact opposite of the transmitter chain. (ie) whatever processing done at the Transmitter, we do exactly opposite at the receiver side.

We demap the DMRS from its location and do the channel estimation. The CH-EST data and the received data are fed to the Equalizer. The equalizer could be ZF based / MMSE based (depends on the implementation). After the Equalizer, they are still complex modulated symbols. The symbols are then undergoes Layer de-mapping, and combine the multiple Layers data into single Codeword.

If there are two codewords (PDSCH), there will be two parallel processings for each of the Codeword. (Codeword 0 and Codeword 1).

If there is a single codeword (PUSCH) (i.e) the No. of Layers are less than 4; then there'll be single codeword processing.

After Layer De-mapping, we do Soft-demodulation (instead of hard decision demodulation). So, the output of Soft-demodulation is not bits, but they are LLRs. (Log Likelihood Ratios).

Followed by, we do De-scrambling. And then, based on certain conditions, we find out ~~[illegible]~~ ~~[illegible]~~. the LDPC Base Graph at the receiver, and then we do code block segmentation, Rate de-matching and LDPC decoder processing.

Now, if there are multiple codeblocks, we concatenate into a single codeblock. Now, we remove the main TB-CRC (i.e) CRC 24A and finally get the TB at the receiver.

This is about the processing of PDSCH and PUSCH. Most of the modules are same whether it is DL or UL, though there were minor differences in the Physical Layer Chain.

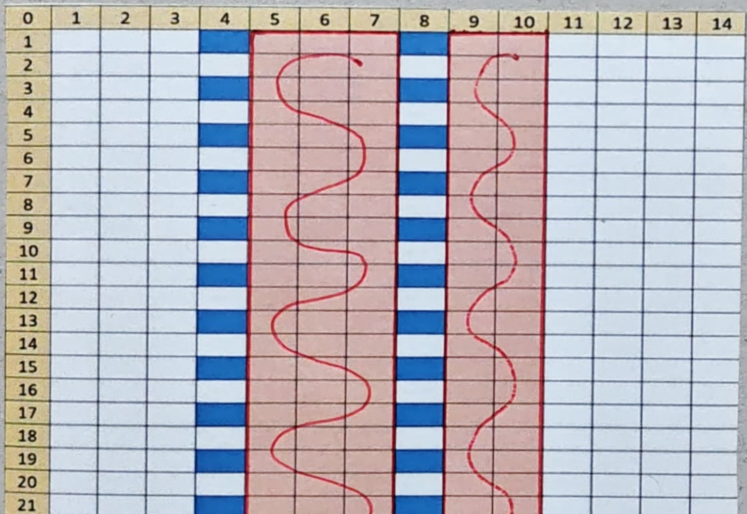
PDSCH/PUSCH : Time and Frequency domain Resource Allocation

In this section, lets understand how the data of PDSCH or PUSCH channels are being transmitted or being mapped to the Resource blocks.

We've already seen how the data (symbols) are mapped to Layers. Consider the data as below.

$$[d_n \ldots\ldots\ldots\ldots d_3 \; d_2 \; d_1 \; d_0]$$

This data is mapped to one of the Layers or one of the Antenna port. How this data is being mapped to the Subcarriers or OFDM symbols ?



The fundamental concept remains the same. (ie) We have the Start Symbol from where PDSCH/PUSCH is supposed to be transmitted, and the No. of Symbols. (in terms of Time).

In terms of frequency, we have the PRBs where the Shared channel data is supposed to be transmitted.

Now, the data that we have on the Antenna Ports, are mapped to the Resource block in increasing order, as shown in figure below.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1		1		37	61	85		121	145					
2		2	25	38	62	86	109	122	146					
3		3		39	63	87		123	147					
4		4	26	40	64	88	110	124	148					
5		5		41	65	89		125	149					
6		6	27	42	66	90	111	126	150					
7		7		43	67	91		127	151					
8		8	28	44	68	92	112	128	152					
9		9		45	69	93		129	153					
10		10	29	46	70	94	113	130	154					
11		11		47	71	95		131	155					
12		12	30	48	72	96	114	132	156					
13		13		49	73	97		133	157					
14		14	31	50	74	98	115	134	158					
15		15		51	75	99		135	159					
16		16	32	52	76	100	116	136	160					
17		17		53	77	101		137	161					
18		18	33	54	78	102	117	138	162					
19		19		55	79	103		139	163					
20		20	34	56	80	104	118	140	164					
21		21		57	81	105		141	165					
22		22	35	58	82	106	119	142	166					
23		23		59	83	107		143	167					
24		24	36	60	84	108	120	144	168					

In this particular example, we have 2 PRBs. So we map 24 complex samples in one symbol, and then we start mapping in the next symbol, and so on.

As we see, the data and DMRS are allowed to be in interleaved fashion.

So, the fundamental concept is, we are given what are the REs that can be used to transmit the shared channel data. The interleaving of data and DMRS may/may not be allowed.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1		1		25	49	73								
2		2		26	50	74								
3		3		27	51	75								
4		4		28	52	76								
5		5		29	53	77								
6		6		30	54	78								
7		7		31	55	79								
8		8		32	56	80								
9		9		33	57	81								
10		10		34	58	82								
11		11		35	59	83								
12		12		36	60	84								
13		13		37	61	85								
14		14		38	62	86								
15		15		39	63	87								
16		16		40	64	88								
17		17		41	65	89								
18		18		42	66	90								
19		19		43	67	91								
20		20		44	68	92								
21		21		45	69	93								
22		22		46	70	94								
23		23		47	71	95								
24		24		48	72	96								

In this case, two UEs (UE1 and UE2) are spatially multiplexed (i.e) they're in MIMO.

Here, the DMRS is represented in two colors and . We may consider is for one layer and is for another layer.

Also, we may consider as, is for UE1 and is for UE2. In this case, both and cannot be used for data transmission. So, in this case, data will go on Symbols 2, 4, 5 and 6.

Also, if any other REs are allocated to some other signals (say CSI-RS or PTRS), then these REs has to be excluded while transmitting PDSCH/PUSCH data or while mapping data to the REs allocated.

In summary, we leave all the REs that are not for PDSCH/PUSCH data. It could be DMRS/CSI-RS/PTRS, we leave all those REs, and wherever the data is supposed to be transmitted, we map it in the increasing order.

Table 5.1.2.1-1: Valid *S* and *L* combinations

PDSCH mapping type	Normal cyclic prefix			Extended cyclic prefix		
	S	L	S+L	S	L	S+L
Type A	{0,1,2,3} (Note 1)	{3,...,14}	{3,...,14}	{0,1,2,3} (Note 1)	{3,...,12}	{3,...,12}
Type B	{0,...,12}	{2,4,7}	{2,...,14}	{0,...,10}	{2,4,6}	{2,...,12}

Note 1: S = 3 is applicable only if *dmrs-TypeA-Position* = 3

PDSCH

Table 6.1.2.1-1: Valid *S* and *L* combinations

PUSCH mapping type	Normal cyclic prefix			Extended cyclic prefix		
	S	L	S+L	S	L	S+L
Type A	0	{4,...,14}	{4,...,14}	0	{4,...,12}	{4,...,12}
Type B	{0,...,13}	{1,...,14}	{1,...,14}	{0,...,12}	{1,...,12}	{1,...,12}

PUSCH

Tables show how the Start symbol (S) and the No. of symbols (L) that are allocated for PDSCH and PUSCH. As we see, there is no much difference among these tables.

WKT, there are two Mapping types. Mapping type B can be used for mini-slot type of applications. In Mapping type A, first DMRS is always on Symbol 2 or 3, so the PDSCH should start from symbol 0, 1, 2 or 3. It cannot start after that. Note that PDSCH should start from symbol 3 only when the first DMRS is on symbol 3.

The first DMRS is known as "dmrs-TypeA-position" which is provided to the UE in MIB.

Whereas, the mapping type B can start from anywhere from Symbols 0 to 12, but there are only 3 lengths defined (ie) 2, 4 or 7 for Normal CP and 2, 4 or 6 for Extended CP. This is something to remember.

In case of PUSCH, in mapping type A, PUSCH should start at only on symbol 0. Whereas in Mapping type B, the PUSCH can start anywhere (sym 0 to 13) and its length can be 1 to 14.

Table 5.1.2.1.1-2: Default PDSCH time domain resource allocation A for normal CP

PDSCH

Row index	dmrs-TypeA-Position	PDSCH mapping type	K_0	S	L
1	2	Type A	0	2	12
	3	Type A	0	3	11
2	2	Type A	0	2	10
	3	Type A	0	3	9
3	2	Type A	0	2	9
	3	Type A	0	3	8
4	2	Type A	0	2	7
	3	Type A	0	3	6
5	2	Type A	0	2	5
	3	Type A	0	3	4
6	2	Type B	0	9	4
	3	Type B	0	10	4
7	2	Type B	0	4	4
	3	Type B	0	6	4
8	2,3	Type B	0	5	7
9	2,3	Type B	0	5	2
10	2,3	Type B	0	9	2
11	2,3	Type B	0	12	2
12	2,3	Type A	0	1	13
13	2,3	Type A	0	1	6
14	2,3	Type A	0	2	4
15	2,3	Type B	0	4	7
16	2,3	Type B	0	8	4

Table 6.1.2.1.1-2: Default PUSCH time domain resource allocation A for normal CP

PUSCH

Row index	PUSCH mapping type	K_2	S	L
1	Type A	j	0	14
2	Type A	j	0	12
3	Type A	j	0	10
4	Type B	j	2	10
5	Type B	j	4	10
6	Type B	j	4	8
7	Type B	j	4	6
8	Type A	$j+1$	0	14
9	Type A	$j+1$	0	12
10	Type A	$j+1$	0	10
11	Type A	$j+2$	0	14
12	Type A	$j+2$	0	12
13	Type A	$j+2$	0	10
14	Type B	j	8	6
15	Type A	$j+3$	0	14
16	Type A	$j+3$	0	10

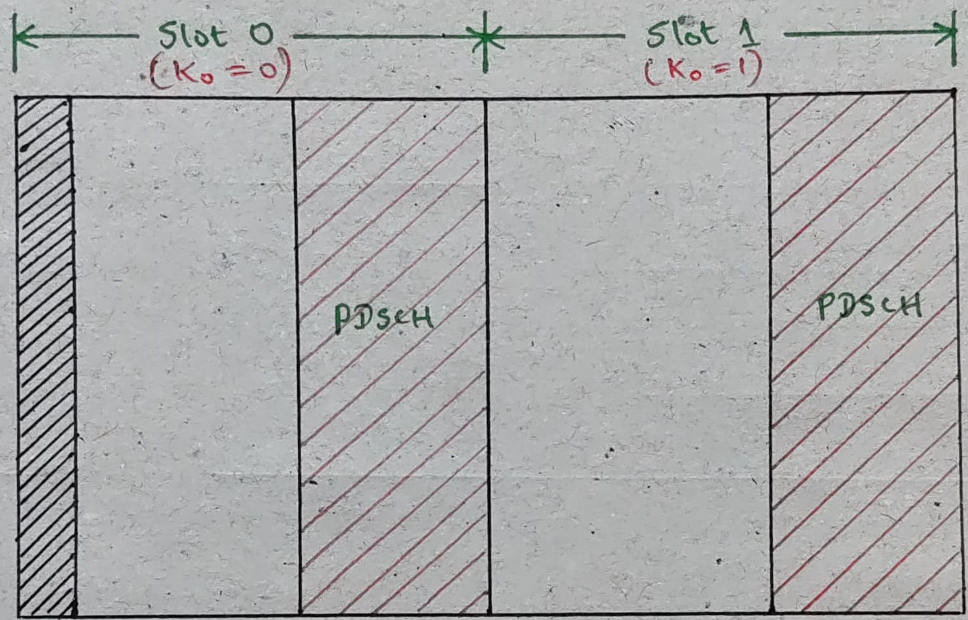
Above tables are predefined in Specification, which can be communicated to the receiver. The Receiver may have just the Row Index. Based on the row index, the receiver learns the start symbol (S), Length (L), and Slot offset (K_0) or (K_2)

Either, the UE can be told (i) Start Symbol (S) and Length (L) (or) (ii) Row index.

As we see, in Mapping Type A, the Start Symbol (S) can be either 2 or 3 and its corresponding Length (L) is given in the table. In Mapping Type B, the Length can be 2, 4, 7. This is in case of PDSCH.

Whereas in case of PUSCH, the Length (L) has huge variety here, and the Start Symbol (S) in Mapping type A is always 0, and in mapping type B, 2, 4, 8.

Now, K_0 and K_2 are the slot offsets



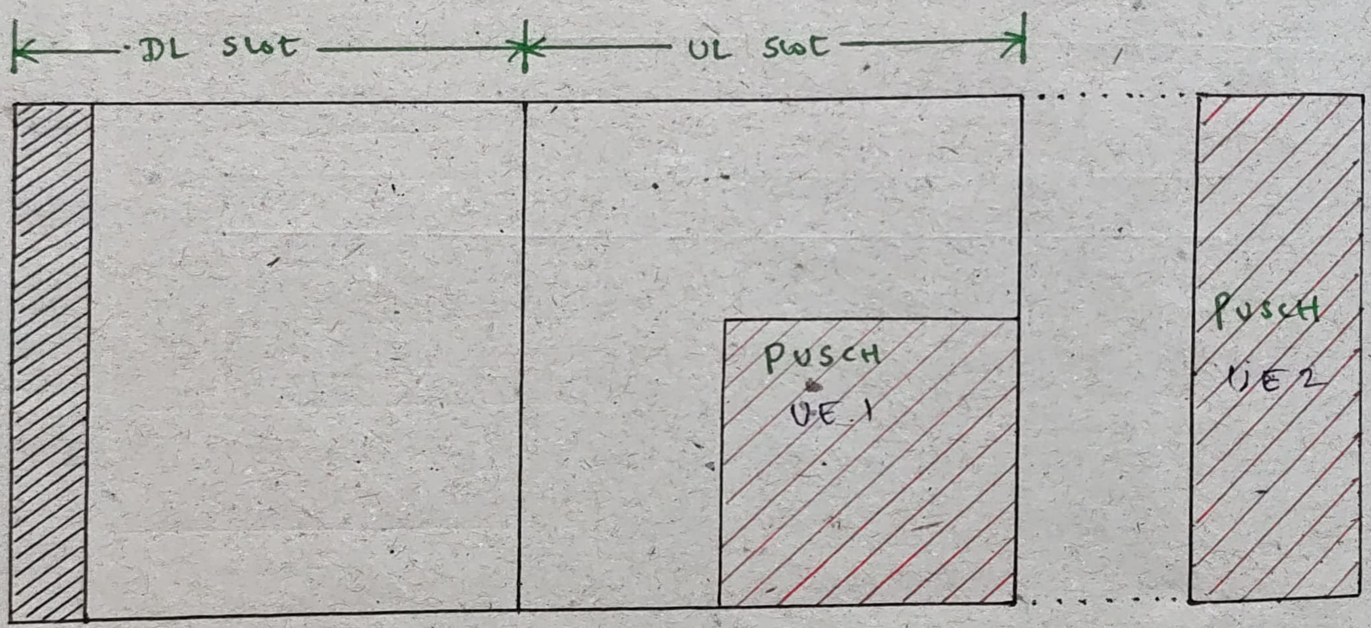
This is where the PDCCH carries DCI for that particular PDSCH / PUSCH. K_0 is the slot offset where the DCI and respective data (PDSCH) is present. If the DCI and PDSCH are in the same slot, then $K_0 = 0$. If the DCI is in One slot (slot 0) and PDSCH is in next slot (slot 1), in this case $K_0 = 1$. So, the K_0 basically tells the slot offset.

Now, we might have different numerologies of PDSCH and PDCCH, in terms of slot offset. So, the actual slot offset is

$$\left\lfloor \frac{n \cdot 2^{\mu\_PDSCH}}{2^{\mu\_PDCCH}} \right\rfloor + K_0$$

So, if PDSCH has higher numerology than PDCCH, then the slot offset can be more than the K_0.

Same analogy goes for the PUSCH as well.



For PUSCH, the slot offset is known as K_2. Here, the PUSCH comes next slot from where DCI is present. So, the slot offset $K_2 = 1$.

K_2 can be 0 as well, but that is for the case of mini-slot. (ie) DCI and PUSCH both in the same slot. Only in this case, $K_2 = 0$, but mostly $K_2 = 1$.

Lets say, there can be two PUSCH, for two UEs. So, for UE 1, $K_2 = 1$ and for UE 2, $K_2 = 2$.

This is for the case when the numerologies for PUSCH and PDCCH are same. If the numerologies are different, then the Slot offset will be communicated as

$$\left\lfloor \frac{n \cdot 2^{\mu\_PUSCH}}{2^{\mu\_PDCCH}} \right\rfloor + K_2$$

This is how the slot offset is being calculated if they are on different numerology.

There is another parameter in the Time domain resource allocation, known as SLIV. This is communicated to the UE, and this carries the Start symbol (S) and the No. of Symbols (L) allocated for PDSCH or PUSCH.

Either we pick S and L from the table, or as RRC message as SLIV (not as S and L).

if $(L-1) \leq 7$ then

$$SLIV = 14 \cdot (L-1) + S$$

else

$$SLIV = 14 \cdot (14 - L + 1) + (14 - 1 - S)$$

where $0 < L \leq 14 - S$, and

Lets say, $L = 7$ and $S = 3$.

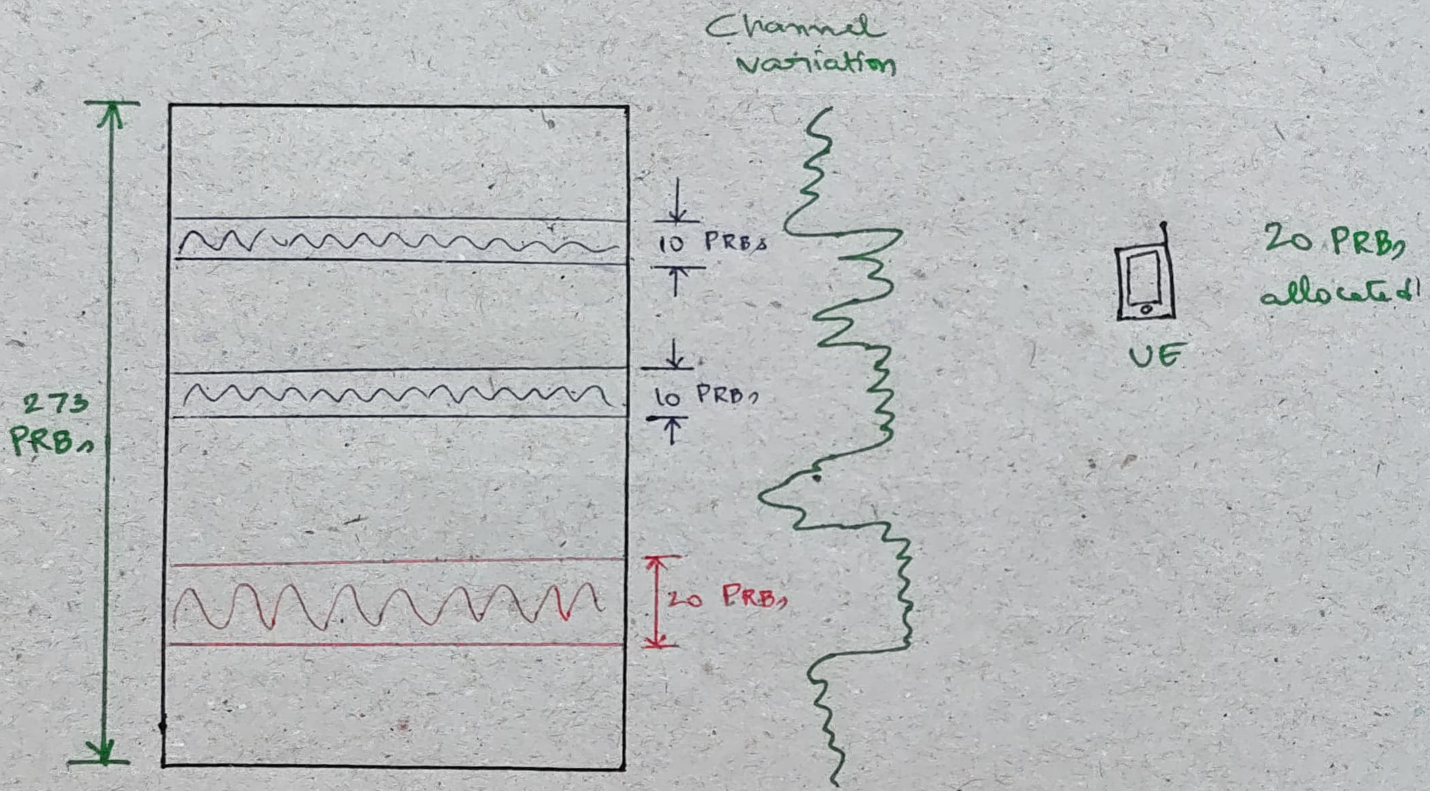
$L - 1 = 6$, which is ≤ 7.

Then in this case,

$$SLIV = 14 \cdot (6) + 3 = 87$$

This formula is used to communicate S and L (Saving the No. of bits during transmission). At the receiver end, UE will decode SLIV to find S and L, given $S + L \leq 14$.

Now, lets understand the Frequency domain Resource allocation of Shared channels. (ie) How the PRBs being allocated to a particular UE in frequency domain, for shared channels PDSCH / PUSCH.



Lets say, this is the complete bandwidth that we have (say 273 PRBs, which is the maximum), and a particular UE which is allocated 20 PRBs at a particular instant. So, where exactly the BS should schedule these 20 PRBs. Should all the PRBs allocated continuously, or could the PRBs allocated in parts (10 PRBs + 10 PRBs). (ie) How the frequency domain allocation should happen?

Ideally, the best way to allocate resources to UE is based on the channel conditions, since the channel is varying across frequency for a particular UE. So, BS should allocate the PRBs which are in good channel conditions for that UE. But it is practically not possible reason being UE is moving very fast. So, in high mobility case, we cannot get the feedback so fast to allocate those PRBs which are in good channel condition (Best SNR).

Also, for the low data rate services (very less data), informing the UE about the PRBs is costly (huge control overhead)

But, in general, to provide the frequency diversity, there are certain methods to allocate resources in frequency which is specified in 3GPP Specifications. There are two resource allocations.

① Resource allocation 0

② Resource allocation 1

Resource allocation 1 is straight forward, which allocates continuous PRBs. somewhere in the bandwidth.

Resource allocation 0 is group based transmission, where group of PRBs which are not continuous in the BWP or bandwidth allocated.

On top of these, we have "Interleaved VRB to PRB mapping". The Resource allocation type 0 and 1 is for both PUSCH and PDSCH. But the Interleaved mapping is only for PDSCH.

In Resource allocation type 1, we do Interleaved mapping to provide Frequency diversity. Whereas in Resource allocation type 0, we dont do this Interleaved mapping of VRB to PRB.

Bandwidth Part Size	Configuration 1	Configuration 2
1 – 36	2	4
37 – 72	4	8
73 – 144	8	16
145 – 275	16	16

The total number of RBGs (N_{RBG}) for a downlink bandwidth part i of size $N^{size}_{BWP,i}$ PRBs is given by

$N_{RBG} = \left\lceil \left(N^{size}_{BWP,i} + \left(N^{start}_{BWP,i} \bmod P \right) \right) / P \right\rceil$, where

- the size of the first RBG is $RBG^{size}_{0} = P - N^{start}_{BWP,i} \bmod P$.
- the size of last RBG is $RBG^{size}_{last} = \left(N^{start}_{BWP,i} + N^{size}_{BWP,i} \right) \bmod P$ if $\left(N^{start}_{BWP,i} + N^{size}_{BWP,i} \right) \bmod P > 0$ and P otherwise.
- the size of all other RBGs is P.

As already mentioned, Resource allocation Type 0 is a group based transmission. So, based on the BWP size, we group certain Resource blocks.

If BWP size $\leq$ 36, then we make group of 2 or 4 based on the configuration (P). Similarly if 145 $\leq$ BWP size $\leq$ 275 (Full bandwidth as BWP), then we can have group of 16 PRBs. So, as the BWP size increases, the PRB group size also increases.

So, as we see, this is not suitable for very small allocations. This is good for large allocations, to provide the diversity. How this diversity is being provided?

Based on the configuration / group size (P), we calculate the No. of groups.

$$N_{RBG} = \left\lceil \left(N^{size}_{BWP,i} + \left(N^{start}_{BWP,i} \bmod P \right) \right) / P \right\rceil$$

Lets say, the BWP start $N^{start}_{BWP,i} = 0$

BWP size $N^{size}_{BWP,i} = 36$

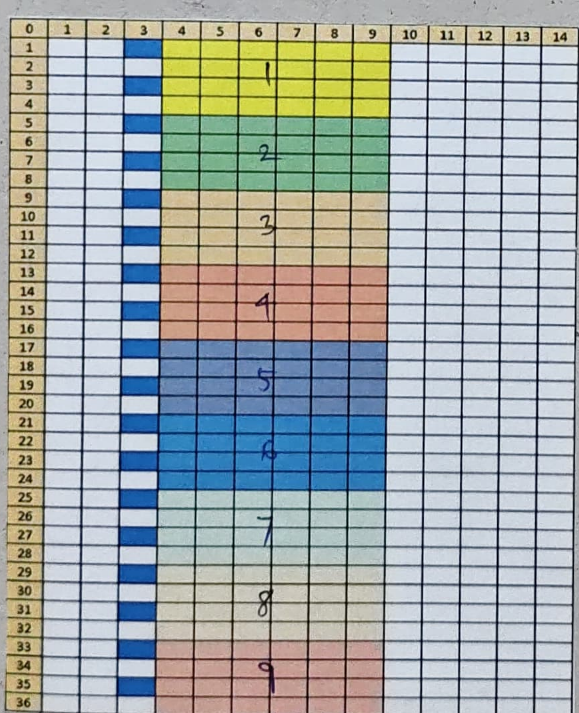
$P = 4$

$$N_{RBG} = \left\lceil \left(36 + (0 \bmod 0)\right) / 4 \right\rceil$$

$$= \frac{36}{4} = 9.$$

So, there are 9 groups. It it is not divisible by P (ie) 4 in this case, then all the RBGs may not have equal sizes. (size of the First and Last RBG may differ).

In this specific example, the size of all the RBGs is equal (ie) 4. But this may not be the case for all BWP size.

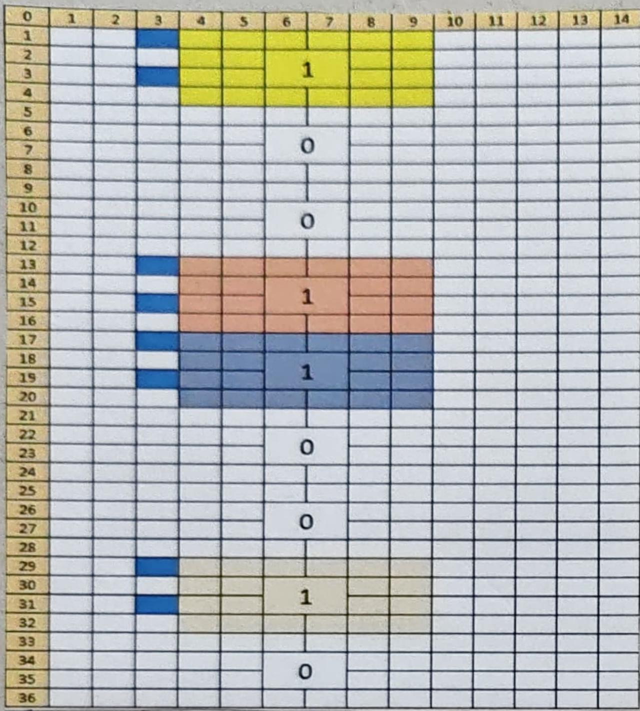


PRBs

Figure shows, how the 9 RBGs looks like.

Total 9 * 4 = 36 PRBs

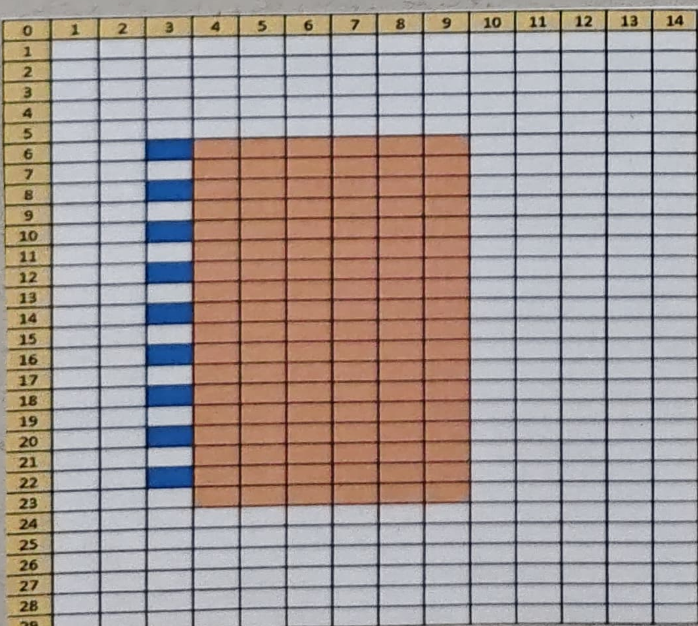
Lets say, a particular UE has scheduled allocation of 16 PRBs. How the Resource Allocation Type 0 is happen is shown in figure below.



PRBs

The UE can have 4 RBG groups which are not continuous allocated for transmission.

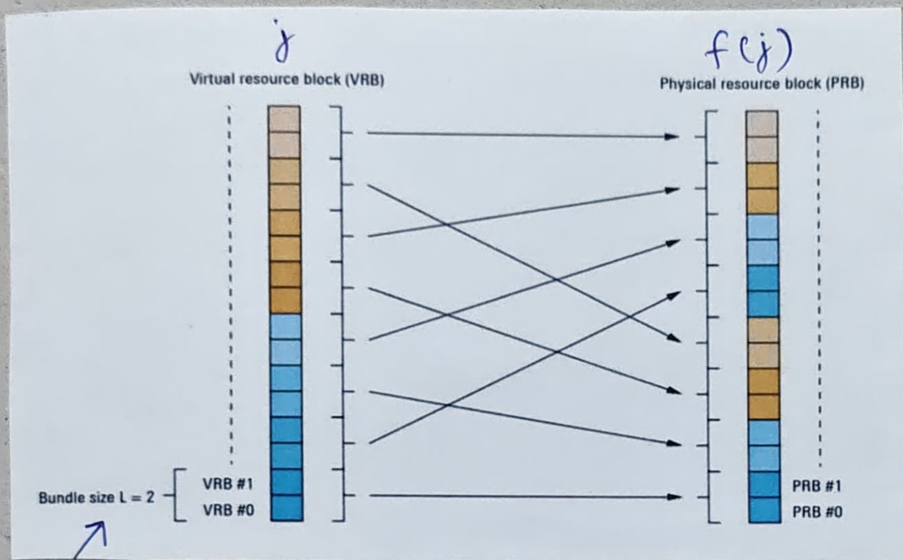
To indicate this to the UE, based on the RBG group we can have Bitmap "1 0 0 1 1 0 0 1 0", which says that, the first, fourth, fifth and eighth RBGs are allocated to that particular UE. This way, the Resource allocation type 0 provides frequency diversity. Here, which group to choose and which to avoid can also be done based on the feedback or channel conditions that are good at that instance for that particular UE.



PRBs

Now, as mentioned previously, allocation type 1 is very straight forward, and this is mostly used for its simplicity. We only need to convey what is the start PRB and No. of PRBs. In this case, the start PRB is 6, and the No. of PRBs is 18. So, PRBs from index 6 to 23 is allocated in continuous fashion.

Now, in this Resource allocation Type 1, we can still have frequency diversity by doing interleaved VRB to PRB mapping.



This interleaved mapping is very similar to allocation type 0, but not exactly same.

$$f(j) = rC + c$$
$$j = cR + r$$
$$r = 0,1,\dots,R-1$$
$$c = 0,1,\dots,C-1$$
$$R = 2$$
$$C = \lfloor N_{bundle} / R \rfloor$$

This is how interleaving is done here. Something similar we've also seen in PDCCH as well.

Lets say, the No. of bundles, $N_{bundle} = 8$

$R = 2$

$C = \left\lfloor \frac{8}{2} \right\rfloor = 4$

$\because$ Bundle size, $L = 2$, and

No. of Bundles, $N_{bundle} = 8$

$j \rightarrow$ VRB index

$f(j) \rightarrow$ PRB index

As shown in figure, the first and Last VRB bundle is directly mapped to the PRB bundles. And the remaining 6 VRB bundles are interleaved.

So, after the RB mapping, we do this kind of interleaving and then we transmit the data in interleaved fashion.

For PUSCH, there is no interleaved mapping and no VRB to PRB mapping. This is only for PDSCH.

PDSCH / PUSCH : MCS tables

Lets understand MCS tables for PDSCH and PUSCH. Why do we have different MCS indices ?

When the UE far away from the BS, or when the UE is inside a room with so many obstructions, the SNR would become low. (meaning channel is bad). So, at a lower SNR scenario, higher modulation orders cannot be used.

So, our end goal is to push maximum throughput. But the max. throughput should also go with good reliability. So, there is a trade-off between High throughput and Reliability. (ie) For higher modulation (say 256-QAM), throughput increases but reliability decreases.

To maintain good BLER (say 10%) (ie) not more than 10% of the packets should be lost or should have errors, we try to find out ideal MCS to use. At the same time, we also try to achieve high throughput. This is the reason we have varieties of MCS tables in 5G.

Table illustrates the structure of the MCS table. First column represents the MCS index, which goes from 0 to 31.

Second column represents the Modulation order.

MCS Index I_{MCS}	Modulation Order Q_m	Target code Rate R x [1024]	Spectral efficiency
0	2	120	0.2344
1	2	157	0.3066
2	2	193	0.3770
3	2	251	0.4902
4	2	308	0.6016
5	2	379	0.7402
6	2	449	0.8770
7	2	526	1.0273
8	2	602	1.1758
9	2	679	1.3262
10	4	340	1.3281
11	4	378	1.4766
12	4	434	1.6953
13	4	490	1.9141
14	4	553	2.1602
15	4	616	2.4063
16	4	658	2.5703
17	6	438	2.5664
18	6	466	2.7305
19	6	517	3.0293
20	6	567	3.3223
21	6	616	3.6094
22	6	666	3.9023
23	6	719	4.2129
24	6	772	4.5234
25	6	822	4.8164
26	6	873	5.1152
27	6	910	5.3320
28	6	948	5.5547
29	2	reserved	
30	4	reserved	
31	6	reserved	

Each row corresponds to a particular Modulation order, Code rate and Spectral Efficiency.

Modulation order, $Q_m = \begin{cases} 2 & \text{for QPSK} \\ 4 & \text{for 16-QAM} \\ 6 & \text{for 64-QAM} \\ 8 & \text{for 256-QAM} \end{cases}$

Code Rate, $R = \dfrac{\text{Information bits}}{\text{Total transmitted bits}}$

Spectral efficiency = Modulation Order x Code Rate.

The Spectral efficiency basically tells, how many bits we're going to transmit per second per hertz. (ie) bits / sec / Hz. This is for one Layer. If there are more no. of layers (say 8 Layer), then the spectral efficiency would be 8 times. So basically, the Spectral efficiency tells the maximum possible data that we can push with this MCS.

Lets say, the Spectral Efficiency value is 2.5703. and the Bandwidth is 100 MHz, then we can say that we can push 257.03 bits / sec.

* This is the maximum possible throughput without considering any overhead. But in practical scenario, even with OFDM, we have overheads like cyclic prefix, DMRS, PDCCH, ... which will bring down the throughput to a lower number. This is how the MCS tables are structured.

Table 5.1.3.1-1: MCS index table 1 for PDSCH

MCS Index I_{MCS}	Modulation Order Q_m	Target code Rate R x [1024]	Spectral efficiency
0	2	120	0.2344
1	2	157	0.3066
2	2	193	0.3770
3	2	251	0.4902
4	2	308	0.6016
5	2	379	0.7402
6	2	449	0.8770
7	2	526	1.0273
8	2	602	1.1758
9	2	679	1.3262
10	4	340	1.3281
11	4	378	1.4766
12	4	434	1.6953
13	4	490	1.9141
14	4	553	2.1602
15	4	616	2.4063
16	4	658	2.5703
17	6	438	2.5664
18	6	466	2.7305
19	6	517	3.0293
20	6	567	3.3223
21	6	616	3.6094
22	6	666	3.9023
23	6	719	4.2129
24	6	772	4.5234
25	6	822	4.8164
26	6	873	5.1152
27	6	910	5.3320
28	6	948	5.5547
29	2	reserved	
30	4	reserved	
31	6	reserved	

Table 5.1.3.1-2: MCS index table 2 for PDSCH

MCS Index I_{MCS}	Modulation Order Q_m	Target code Rate R x [1024]	Spectral efficiency
0	2	120	0.2344
1	2	193	0.3770
2	2	308	0.6016
3	2	449	0.8770
4	2	602	1.1758
5	4	378	1.4766
6	4	434	1.6953
7	4	490	1.9141
8	4	553	2.1602
9	4	616	2.4063
10	4	658	2.5703
11	6	466	2.7305
12	6	517	3.0293
13	6	567	3.3223
14	6	616	3.6094
15	6	666	3.9023
16	6	719	4.2129
17	6	772	4.5234
18	6	822	4.8164
19	6	873	5.1152
20	8	682.5	5.3320
21	8	711	5.5547
22	8	754	5.8906
23	8	797	6.2266
24	8	841	6.5703
25	8	885	6.9141
26	8	916.5	7.1602
27	8	948	7.4063
28	2	reserved	
29	4	reserved	
30	6	reserved	
31	8	reserved	

For PDSCH, we have three such tables. As we see, Table 2 is the only table that has 256-QAM, highest spectral efficiency of 7.4063 bits/sec/Hz, and highest code Rate of $\frac{948}{1024}$. So, Table 2 is being used when UE and gNB both supports 256-QAM.

Table 1 doesn't have 256-QAM. The highest modulation is 64-QAM. So, this table is being used whenever the UE or the gNb does not support 256-QAM.

If we compare Table 3 with Table 1, the highest modulation order in both of them is 64-QAM, but the spectral efficiency is lower in Table 3 compared to Table 1. So, basically Table 3 is a low spectral efficiency table and the difference between two consecutive spectral efficiency is very low (Fine). So, Table 3 can be used for the Low code Rate scenario when we need highly reliable data transmission.

Table 5.1.3.1-3: MCS index table 3 for PDSCH

MCS Index I_{MCS}	Modulation Order Q_m	Target code Rate R x [1024]	Spectral efficiency
0	2	30	0.0586
1	2	40	0.0781
2	2	50	0.0977
3	2	64	0.1250
4	2	78	0.1523
5	2	99	0.1934
6	2	120	0.2344
7	2	157	0.3066
8	2	193	0.3770
9	2	251	0.4902
10	2	308	0.6016
11	2	379	0.7402
12	2	449	0.8770
13	2	526	1.0273
14	2	602	1.1758
15	4	340	1.3281
16	4	378	1.4766
17	4	434	1.6953
18	4	490	1.9141
19	4	553	2.1602
20	4	616	2.4063
21	6	438	2.5664
22	6	466	2.7305
23	6	517	3.0293
24	6	567	3.3223
25	6	616	3.6094
26	6	666	3.9023
27	6	719	4.2129
28	6	772	4.5234
29	2	reserved	
30	4	reserved	
31	6	reserved	

For URLLC transmission, we can choose MCS table 3.

These tables 1/2/3 can also be used for PUSCH, when Transform Precoding is disabled. When Transform Precoding is enabled, there are two more tables for PUSCH.

Note:

Table 2 can also be used when Transform Precoding is enabled.

MCS Tables for PUSCH :

Here also, Table 2 is lower spectral efficiency table. So, when Transform Precoding is enabled and we require reliable transmission, we use MCS table 2, Otherwise use MCS table 1.

Note : when both UE and gNB supports 256 QAM, use PDCCH MCS table 2.

Also, notice that, in all tables, we see last few rows are 'reserved'. These entries can be used for Retransmission purpose. So, whenever gNB mentions this MCS in the DCI, that means it is a retransmission data.

Table 6.1.4.1-1: MCS index table for PUSCH with transform precoding and 64QAM

MCS Index I_{MCS}	Modulation Order Q_m	Target code Rate R x 1024	Spectral efficiency
0	q	240/ q	0.2344
1	q	314/ q	0.3066
2	2	193	0.3770
3	2	251	0.4902
4	2	308	0.6016
5	2	379	0.7402
6	2	449	0.8770
7	2	526	1.0273
8	2	602	1.1758
9	2	679	1.3262
10	4	340	1.3281
11	4	378	1.4766
12	4	434	1.6953
13	4	490	1.9141
14	4	553	2.1602
15	4	616	2.4063
16	4	658	2.5703
17	6	466	2.7305
18	6	517	3.0293
19	6	567	3.3223
20	6	616	3.6094
21	6	666	3.9023
22	6	719	4.2129
23	6	772	4.5234
24	6	822	4.8164
25	6	873	5.1152
26	6	910	5.3320
27	6	948	5.5547
28	q	reserved	
29	2	reserved	
30	4	reserved	
31	6	reserved	

Table 6.1.4.1-2: MCS index table 2 for PUSCH with transform precoding and 64QAM

MCS Index I_{MCS}	Modulation Order Q_m	Target code Rate R x 1024	Spectral efficiency
0	q	60/q	0.0586
1	q	80/q	0.0781
2	q	100/q	0.0977
3	q	128/q	0.1250
4	q	156/q	0.1523
5	q	198/q	0.1934
6	2	120	0.2344
7	2	157	0.3066
8	2	193	0.3770
9	2	251	0.4902
10	2	308	0.6016
11	2	379	0.7402
12	2	449	0.8770
13	2	526	1.0273
14	2	602	1.1758
15	2	679	1.3262
16	4	378	1.4766
17	4	434	1.6953
18	4	490	1.9141
19	4	553	2.1602
20	4	616	2.4063
21	4	658	2.5703
22	4	699	2.7305
23	4	772	3.0156
24	6	567	3.3223
25	6	616	3.6094
26	6	666	3.9023
27	6	772	4.5234
28	q	reserved	
29	2	reserved	
30	4	reserved	
31	6	reserved	

Also notice the modulation order $Q_m = q$, which means, if

$q = 1 \Rightarrow \frac{\pi}{2}$ BPSK

$q = 2 \Rightarrow$ QPSK

$q = 1$, means 1 bit is mapped to 1 symbol.

$q = 2$, means 2 bit is mapped to 1 symbol.

In both the cases, the corresponding Target code Rate is divided by q. So, whether we use $\frac{\pi}{2}$ BPSK (or) QPSK, the spectral efficiency remains the same.

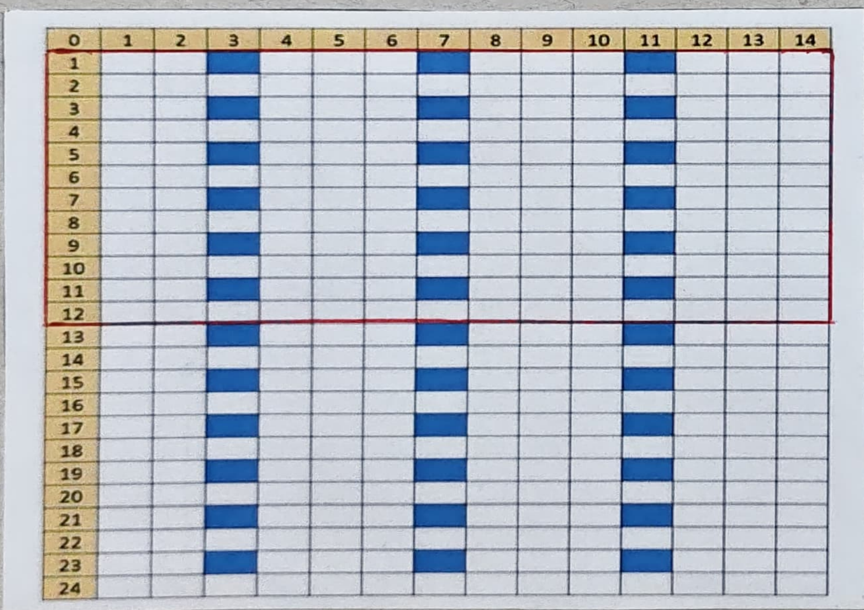
—— x ——

PDSCH/PUSCH : Transport Block (TB) size calculations

Lets understand the TB size calculation for PDSCH and PUSCH. The calculation is same for both the shared channels (DL or UL).

In 5G, TB size is not conveyed to the UE in the DCI. UE has to calculate the TB size based on the resources allocated, MCS and other parameters, based on a formula.

First, we calculate the resources allocated. (i.e.) No. of REs allocated to PDSCH or PUSCH in a PRB.



Total No. of REs in a PRB $\left.\right\} = 12 \times 14 = 168$

Out of 168, how many REs are allocated to the UE for the channel (UL or DL) at that instance?

$$= (12 \times \text{No. of Symbols}) - \underbrace{(\text{DMRS REs})} - (\text{Overhead i.e. PTRS/CSI-RS})$$

DMRS also covers REs that are allocated on other layers where we're not supposed to transmit the data. In this example, there are 3 DMRS symbols $(6 \times 3 = 18$ REs). Assuming data is not allowed to be transmitted on alternative REs, the DMRS overhead becomes $18 \times 2 = 36$ DMRS REs.

In this example, there is no CSI-RS, hence this overhead would be 0.

So, No. of REs allocated to the UE at this particular instance in a PRB. $= (12 \times 14) - 36 - 0$

(should not be > 156 out of 168 REs)

Now, Total No. of REs allocated (full allocation) (N_{RE}) = (No. of REs allocated in a PRB) * (Total No. of PRBs)

Having calculated N_{RE}, now we calculate the approximate TB size (intermediate information bits) N_{info}.

$$\boxed{N_{info} = N_{RE} * v * Q_m * R}$$

where, $v \rightarrow$ No. of layers allocated (1/2/3/4)

$Q_m \rightarrow$ Modulation Order (1/2/4/6/8)

$R \rightarrow$ Code Rate

Now, we got the approximate No. of bits before the channel coding at the transmitter. This is not exact TB size, coz there are certain rules to have the TB size (formula).

$$N'_{info} = \max\left(24, 2^n \left\lfloor \frac{N_{info}}{2^n} \right\rfloor\right)$$

$$N_{info} = \max\left(24, 2^n \cdot \left\lfloor \frac{N_{info}}{2^n} \right\rfloor\right), \text{ where } n = \max(3, \lfloor \log_2(N_{info}) \rfloor - 6).$$

WKT, in LDPC channel coding, if $N_{info} < 3824$, then there is no code block segmentation and only single code block would be transmitted. (for BG 2). So, when N_{info} is less than 3824, we use this formula.

$n \rightarrow$ step size (ie)

This tells that N'_{info} is multiple of 2^n. (8, 16, ..), with minimum value of $N'_{info} = 24$.

Table 5.1.3.2-2: TBS for $N_{info} \leq 3824$

Index	TBS	Index	TBS	Index	TBS	Index	TBS
1	24	31	336	61	1288	91	3624
2	32	32	352	62	1320	92	3752
3	40	33	368	63	1352	93	3824
4	48	34	384	64	1416		
5	56	35	408	65	1480		
6	64	36	432	66	1544		
7	72	37	456	67	1608		
8	80	38	480	68	1672		
9	88	39	504	69	1736		
10	96	40	528	70	1800		
11	104	41	552	71	1864		
12	112	42	576	72	1928		
13	120	43	608	73	2024		
14	128	44	640	74	2088		
15	136	45	672	75	2152		
16	144	46	704	76	2216		
17	152	47	736	77	2280		
18	160	48	768	78	2408		
19	168	49	808	79	2472		
20	176	50	848	80	2536		
21	184	51	888	81	2600		
22	192	52	928	82	2664		
23	208	53	984	83	2728		
24	224	54	1032	84	2792		
25	240	55	1064	85	2856		
26	256	56	1128	86	2976		
27	272	57	1160	87	3104		
28	288	58	1192	88	3240		
29	304	59	1224	89	3368		
30	320	60	1256	90	3496		

Once N'_{info} is calculated, we've to find the TB size from the above table, such that it is closest to N'_{info} and also greater than N'_{info}.

This is how we calculate TB size, if $N_{info} < 3824$.

If $N_{info} > 3824$, then TB size is calculated as below, since there are multiple code blocks involved here.

When $N_{info} > 3824$, TBS is determined as follows.

- quantized intermediate number of information bits $N'_{info} = \max\left(3840, 2^n \times round\left(\frac{N_{info}-24}{2^n}\right)\right)$, where $n = \lfloor \log_2(N_{info}-24) \rfloor - 5$ and ties in the round function are broken towards the next largest integer.
- if $R \le 1/4$

$$TBS = 8 \cdot C \cdot \left\lceil \frac{N'_{info}+24}{8 \cdot C} \right\rceil - 24, \text{ where } C = \left\lceil \frac{N'_{info}+24}{3816} \right\rceil$$

else

if $N'_{info} > 8424$

$$TBS = 8 \cdot C \cdot \left\lceil \frac{N'_{info}+24}{8 \cdot C} \right\rceil - 24, \text{ where } C = \left\lceil \frac{N'_{info}+24}{8424} \right\rceil$$

else

$$TBS = 8 \cdot \left\lceil \frac{N'_{info}+24}{8} \right\rceil - 24$$

end if

end if

Here also, we calculate N'_{info} using the formula

$$N'_{info} = \max\left(3840, 2^n \times round\left(\frac{N_{info}-24}{2^n}\right)\right)$$

Here, the minimum value of N'_{info} is 3840.
n represents the stepping size. So, 2^n would be the step size. '24' represents the CRC bits here.

So, if TB size is > 3824 and $R < 1/4$, then the code block size is $\lceil (N'_{info} + 24) / 3816 \rceil$

if TB size is > 3824 and $R > 1/4$, then the code block size is $\lceil (N'_{info} + 24) / 8424 \rceil$

Lets understand with an example.

$DMRS_{RE_s} = 24$

Overhead $= 0$,

Total PRBs $= 10$

No. of symbols $= 14$

No. of layers $= v = 1$

$Q_m = 2$

$R = \frac{308}{1024}$

$$\left.\begin{array}{l}\text{No. of REs in a PRB} \\ (N'_{RE})\end{array}\right\} = (12 \times 14) - 24 - 0 = 144$$

$$\left.\begin{array}{l}\text{Total No. of REs in} \\ \text{the full bandwidth} \\ (N_{RE})\end{array}\right\} = 144 \times \left(\begin{array}{c}\text{Total No.} \\ \text{of PRBs}\end{array}\right)$$

$$= 144 \times 10$$

$$= 1440$$

$$\left.\begin{array}{l}\text{Approximate TB size} \\ (N_{info})\end{array}\right\} = 1440 * v * Q_m * R$$

$$= 1440 \times 1 \times 2 \times \frac{308}{1024}$$

$$= 866.25 \leq 3824$$

$$n = \max\left(3, \lfloor \log_2 (N_{info}) \rfloor - 6\right)$$

$$= \max\left(3, \lfloor \log_2 866.25 \rfloor - 6\right)$$

$$= \max(3, 9 - 6)$$

$$= 3$$

$$N'_{info} = \max\left(24, 2^n \left\lfloor \frac{N_{info}}{2^n} \right\rfloor\right)$$

$$= \max\left(24, 2^3 \left\lfloor \frac{866.25}{2^3} \right\rfloor\right)$$

$$= 864.$$

$\Rightarrow$ TB size $= 888$ (Refer Table 5.1.3.2-2)

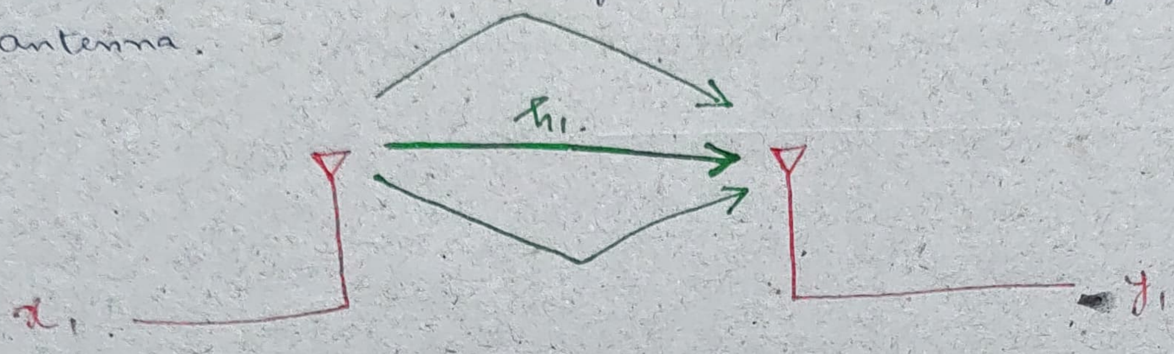
This is about TB size calculation. (u)

This is what UE does everytime it receives PDSCH

or when UE has to transmit PUSCH after receiving DCI.

MIMO : Basics of Multiple Antenna Systems.

Lets start with single transmit and single receive antenna.



where x_1 is the IQ sample transmitted

y_1 is the received IQ sample

x_1 is transmitted through the antenna, which will go over multiple paths and multiple copies are received at the receiver. Because of these multi paths, there'll be change in Amplitude and phase of the signal. This change in Amplitude and Phase is defined as channel h_1 (ie) h_1 is basically the channel, changing the phase of the transmitted signal x_1. So, the received signal y_1 is given by

$$\boxed{y_1 = h_1 x_1 + w_1}$$

where, $w_1 \rightarrow$ Hardware noise.

$h_1 \rightarrow$ channel that is estimated using DMRS

Now, at the receiver,

$h_1 \rightarrow$ channel estimated using DMRS

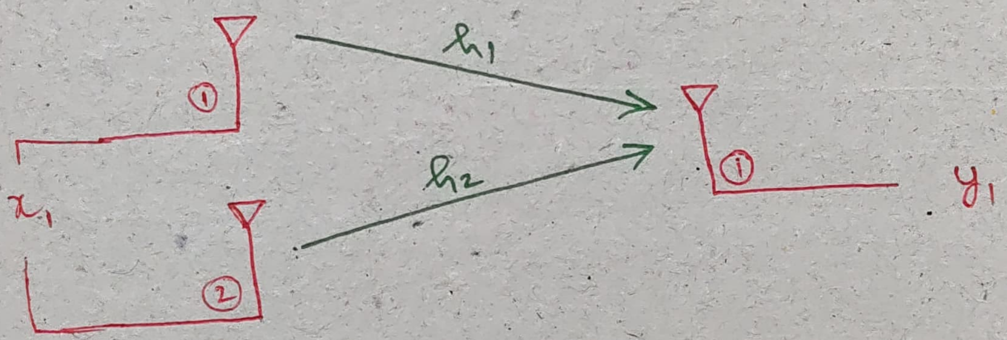
$y_1 \rightarrow$ received signal

$w_1 \rightarrow$ Noise

are known at receiver.

So, x_1 can be detected at the receiver. This particular system with single antenna at transmitter and receiver is known as SISO.

Lets have two transmit antennas and one receive antenna.



where $h_1 \rightarrow$ channel representing the phase and amplitude change between Tx. Antenna 1 and Rx. Antenna 1

$h_2 \rightarrow$ Channel representing the phase and amplitude change between Tx. Antenna 2 and Rx. Antenna 1

Now, we transmit the same signal x_1 through both the Tx. antennas. So, the received signal y_1 is given by

$$\boxed{y_1 = h_1 x_1 + h_2 x_1 + w_1}$$

Normally, if we transmit signal x_1 through two different antennas (without any change), the gain is very low. One of the advantage is that, if channel h_1 goes bad and cannot be decoded, we always have the second copy of x_1 in channel h_2 from the Tx. Antenna 2. and it can be decoded.

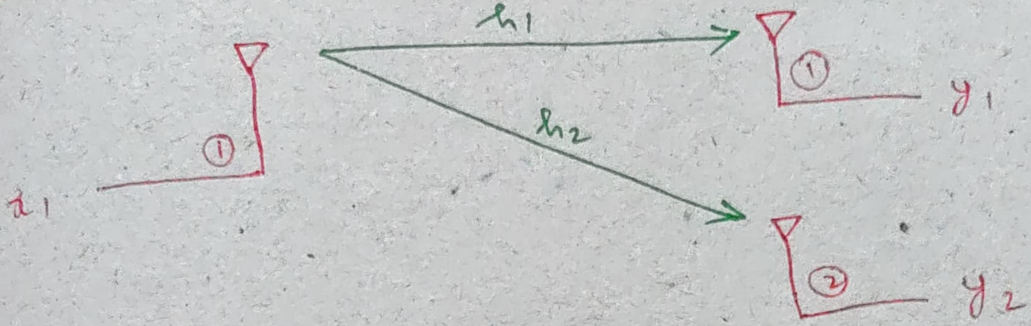
Normally, when we transmit the same signal through 2 antennas, we make some phase and amplitude change, which is basically known as <u>Precoding</u>. The signal x_1 undergoes Precoding using the <u>Precoding Matrix</u>, which <u>improves the signal quality</u> at the receiver. (i.e.) the Probability that we detect the right signal increased. This particular scenario is known as Transmit Diversity.

<u>Transmit Diversity</u> :

- Transmit the same data through multiple antennas

This particular system with two antennas at the transmitter and One Antenna at the receiver is known as MISO. (2T1R scenario)

Now, lets take another case, where we have Single Transmit Antenna and Multiple Receive Antennas.



The received signal at Rx. Antenna 1 is

$$y_1 = x_1 h_1 + w_1 \quad \text{———} \quad (1)$$

The received signal at Rx. Antenna 2 is

$$y_2 = x_1 h_2 + w_2 \quad \text{———} \quad (2)$$

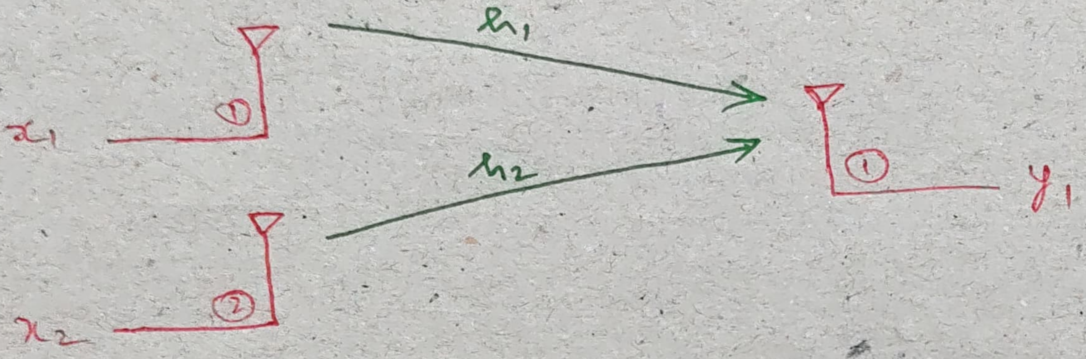
Here also, we receive two copies of the transmitted signal x_1. And the advantage is, if one of the path goes bad, we always have the second path. And, this scenario increases the signal quality.

The receiver uses equation (1) and (2) in many ways in the equalizer, to get the signal x_1.

Basically, we have two equations (1) and (2), and only one unknown x_1. So, using these two equations, we can find x_1. h_1 and h_2 ~~are~~ always estimated using the DMRS. This particular scenario is known as Receiver diversity.

And, this particular system with one Antenna at the transmitter and two antennas at the receiver is known as SIMO.

Now, lets take the previous case. (2T 1R scenario).
Now, instead of transmitting same information via both the transmit antennas ① and ②, we now transmit two different data x_1 and x_2 at same instance.



The received signal y_1 can be written as

$$y_1 = h_1 x_1 + h_2 x_2 + w_1$$

where, $h_1, h_2 \rightarrow$ estimated using DMRS

Assume $h_1 = 5$, $h_2 = 7$ and $y_1 = 17$ (typically complex values, for simplicity we assume real values)

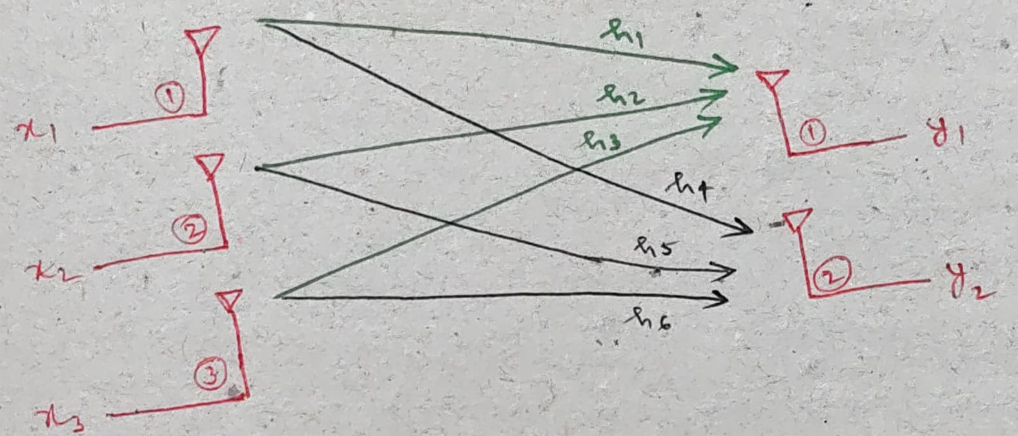
$$\Rightarrow 17 = 5x_1 + 7x_2 + \cancel{w_1}$$

Assume No noise

$$\Rightarrow 17 = 5x_1 + 7x_2$$

With one equation, we cannot find two unknowns, which means, when we have only one Rx. Antenna, we cannot transmit two different information using two different transmit antennas.

Lets take a different case where we have 3 Tx. Antennas and 2 Rx. Antennas, and we transmit three different data x_1, x_2 and x_3 at same instance.



$$y_1 = h_1 x_1 + h_2 x_2 + h_3 x_3 + w_1 \quad \text{——— (1)}$$

$$y_2 = h_4 x_1 + h_5 x_2 + h_6 x_3 + w_2 \quad \text{——— (2)}$$

where, $h_1, h_2, \ldots, h_6 \rightarrow$ estimated using DMRS

$y_1, y_2 \rightarrow$ known at the receiver

So, the unknowns x_1, x_2 and x_3 need to find out at the receiver using 2 equations

Note: using 2 equations, the maximum unknowns we can find is 2

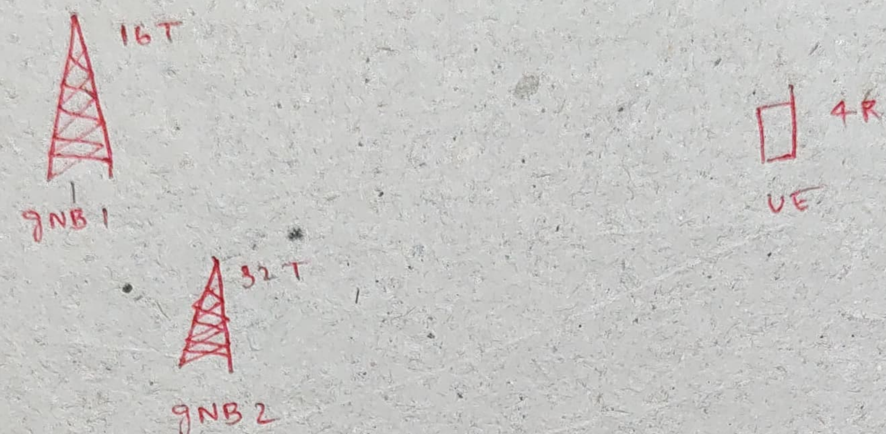
That means, in this case, at max we can transmit two different data. we cannot transmit 3 different data.

So, the No. of data streams / Layers are limited by both transmit and receive antennas.

$$\left.\begin{array}{l}\text{Max. data streams}\\ \text{or Layers}\end{array}\right\} = \min(n_t, n_r)$$

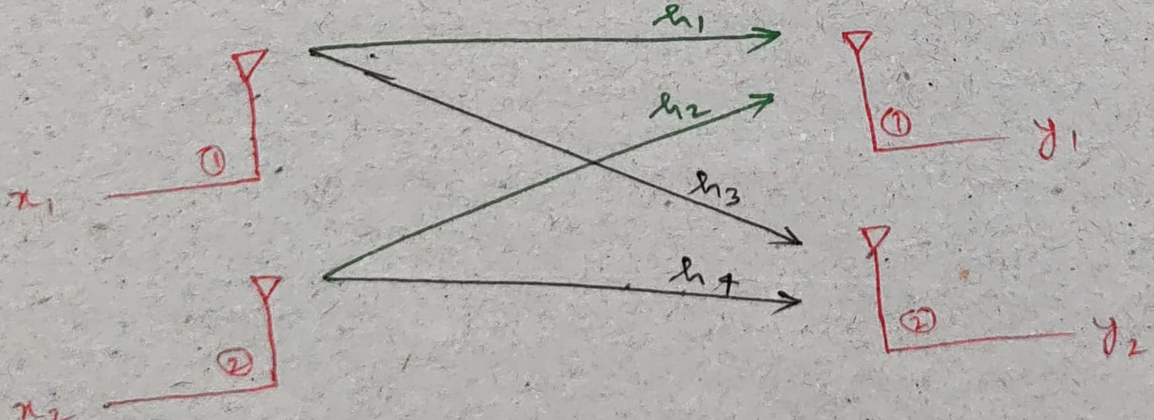
$n_t \rightarrow$ No. of Tx. Antenna

$n_r \rightarrow$ No. of Rx. Antenna



Means, even if we have a system with 16 or 32 antennas at gNB side, and the UE have 4 antennas, gNB can only transmit 4 layers in downlink. And, in order to transmit 8 layers in downlink, we need to have 8 receiver antennas at the UE side.

Lets take another case where we have 2T 2R system, means two Tx. Antenna and two Rx. Antenna.



$$y_1 = h_1 x_1 + h_2 x_2 + \cancel{w_1}^{\,0}$$

$$y_2 = h_3 x_1 + h_4 x_2 + \cancel{w_2}^{\,0}$$

$$\Rightarrow \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} h_1 & h_2 \\ h_3 & h_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Assume : $h_1 = 2$, $h_2 = 7$

$h_3 = 4$, $h_4 = 14$

$\Rightarrow y_1 = 2x_1 + 7x_2$ —— ①

$y_2 = 4x_1 + 14x_2$ —— ②

$\Rightarrow$ There is correlation between the signals received of Rx. Antenna 1 and Rx. Antenna 2.

(i.e) Equation ① and ② are identical.

① $\Rightarrow 2x_1 + 7x_2 = y_1$

② $\Rightarrow 2x_1 + 7x_2 = y_2/2$

And, there is no way we can find x_1 and x_2.

So, eventhough we have 2 transmit and 2 receive antennas, we cannot always transmit two different data streams or Layers, and is limited by the type of channel we have between the Tx. and Rx. Antennas.

If the channels are correlated (i.e) the equations are not independent (or) the rows of H matrix are not independent, then the No. of data streams or Layers that we can transmit at the same time is limited by the independent rows of the H matrix.

The MIMO matrix H has the elements from the channel. The No. of independent Rows (Rank of the Matrix) decides the maximum number of data streams or Layers that can be transmitted in parallel.

Note that, If two different Layers belonging to two different UEs, then it is called Multi-user MIMO (MU-MIMO). And, if two different Layers belonging to the single UE, then it is called Single-user MIMO (SU-MIMO).

But the concept here is, even though we have 2T2R system, we <u>cannot always</u> transmit two layers, because at the receiver, we may not be able to separate both the Layers. Since the channel is always changing, it might happen that, at one instance (t1), the channel coefficients are independent and we're able to transmit two layers. And the UE moves a little bit, due to which the channel may change in such a way that their coefficients are dependent, which in turn makes the Rank of the matrix goes down from 2 to 1, and hence we can only transmit in single layer.

<u>Note:</u>

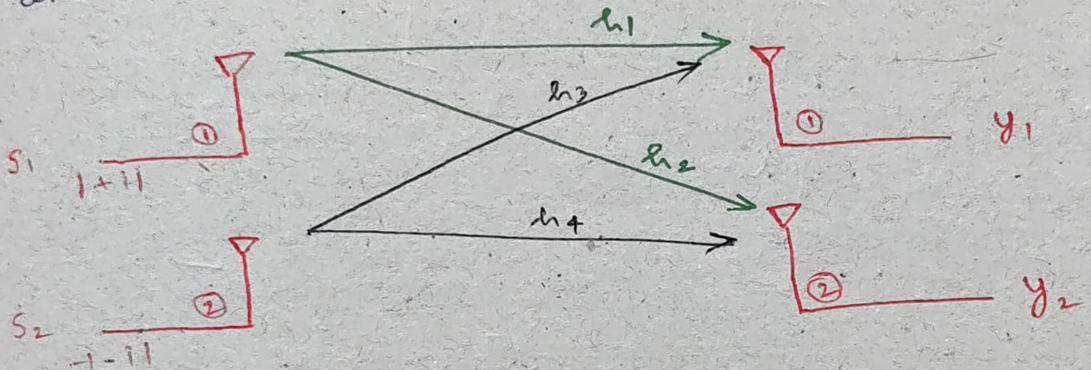
Transmitting multiple data streams at the same time is known as <u>Spatial Multiplexing</u>.

In current 5G scenario, even the gNB has 32 TR or 16 TR system, and at the UE side we normally have 4 antennas. So, at max, the No. of Layers would be 4. But we may not always have 4 layers, and sometime it may be limited by 3 layers or 2 layers or even just 1 layer, depending on the channel conditions.

PDSCH : Precoding for MIMO

Precoding for MIMO is a short term precoding, which is basically used to remove the channel effect. This is applicable for both PDSCH / PUSCH (UL and DL).

Lets assume a 2×2 system with two Tx. Antennas and two Rx. Antennas.



Assume, the bits to be transmitted is 0011..... we perform 'QPSK' modulation first. For sake of simplicity, we take first 4 bits. QPSK modulation maps 2 bits to a symbol. So,

$$00 \rightarrow (1+i1)\cancel{\frac{1}{\sqrt{2}}}$$

$$11 \rightarrow (-1-i1)\cancel{\frac{1}{\sqrt{2}}}$$

To simplify the math, lets remove the normalization factor. If we draw the constellation, both the symbols have same amplitude and in opposite quadrant.

$s_1 = 1 + i1$

$s_2 = -1 - i1$

$$y_1 = h_1 s_1 + h_3 s_2 + w_1$$

$$y_2 = h_2 s_1 + h_4 s_2 + w_2$$

$$\Rightarrow \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} h_1 & h_3 \\ h_2 & h_4 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \cancel{\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}}$$ for simplicity.

Assume, $H = \begin{bmatrix} h_1 & h_3 \\ h_2 & h_4 \end{bmatrix} = \begin{bmatrix} 0.5 + i\,0.7 & 0.3 + i\,0.8 \\ 0.9 - i\,0.1 & 1 \end{bmatrix}$

Amplitude of H $= \begin{bmatrix} \sqrt{(0.5)^2 + (0.7)^2} & \sqrt{(0.3)^2 + (0.8)^2} \\ \sqrt{(0.9)^2 + (-0.1)^2} & \sqrt{1^2 + 0^2} \end{bmatrix}$

$$= \begin{bmatrix} 0.86 & 0.85 \\ 0.90 & 1 \end{bmatrix}$$

Phase of H $= \begin{bmatrix} \tan^{-1}\left(\frac{0.7}{0.5}\right) & \tan^{-1}\left(\frac{0.8}{0.3}\right) \\ \tan^{-1}\left(\frac{-0.1}{0.9}\right) & \tan^{-1}(1) \end{bmatrix}$

$$= \begin{bmatrix} 54^\circ & 70^\circ \\ -6^\circ & 0^\circ \end{bmatrix}$$

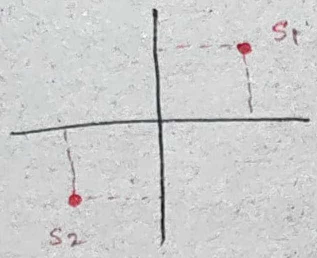
What does these Matrices say?

Lets say, the signal received at y_1, coming from Tx. Antenna 1, is modified by 0.86 amplitude and 54° degree phase shift. Similarly, for other channels as well.

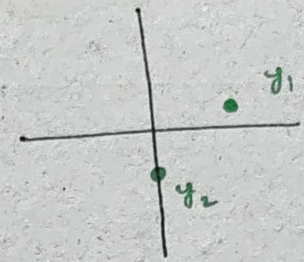
H matrix communicates, the amount of Amplitude and Phase changes that the signal undergoes, while the signal is transmitted at Tx. Antenna and received at Rx. Antenna.

$$\Rightarrow \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0.5 + i\,0.7 & 0.3 + i\,0.8 \\ 0.9 - i\,0.1 & 1 \end{bmatrix} \begin{bmatrix} 1+i1 \\ -1-i1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.3 + i\,0.1 \\ 0 - i\,0.2 \end{bmatrix}$$



Transmitted symbols



Received symbols

This shows the effect of the amplitude and phase change. (combined effect of signals coming from both the antennas) (ie) y_1 and y_2 has signals coming from both S_1 and S_2. (mixed up).

Normally, we use DMRS to calculate H (estimates of h_1, h_2, h_3 and h_4). With these estimates we do equalization (ie) remove the effect of H from

$$y = HS + w$$

at the receiver side...

But, how about removing this channel effect at the transmitter itself.? (ie) How can we remove this channel effect at the transmitter itself even before the signal S was transmitted?

For that, the transmitter should know H. (ie) the receiver is supposed to communicate H to the transmitter, and the transmitter should apply inverse of H to remove the channel effect. This is called as Precoding at the transmitter side.

(ie) UE sends SRS in the uplink. gNB assumes the channel in UL is same as the channel in DL. (aka) Reciprocity, and applies H^{-1} at the transmitter side.

(More detailed study in SRS section)

So, before transmitting the signal S, gNB multiplies the Precoding Matrix P with the signal S. There are many different ways to find out the Precoding Matrix P and they are proprietary. In Industry, we use ZF / MMSE / SVD precoding. These precodings are not defined in the specification, and the gNB vendors are supposed to use whatever precoding they want.

As the UE doesn't know that gNB has applied the precoding, so both the PDSCH and DMRS has to be precoded with the same Precoding matrix before they are transmitted. And at the receiver side, UE uses this DMRS to estimate the channel. So, basically, if Precoding is not applied on DMRS at transmitter side, the UE wont be able to find out the effect of Precoding and wont be able to estimate the channel for PDSCH.

Lets assume ZF precoding is applied at gNB. (i)

$$P = H^H (H H^H)^{-1}$$

Assuming $H = \begin{bmatrix} 0.5 + i0.7 & 0.3 + i0.8 \\ 0.9 - i0.1 & 1 \end{bmatrix}$

$$\Rightarrow P = \begin{bmatrix} 6.6 - i0.4 & -2.3 - i5.1 \\ -5.9 + i1 & 3.6 + i4.4 \end{bmatrix}$$

WKT, the signal to be transmitted is $S = \begin{bmatrix} 1+i1 \\ -1-i1 \end{bmatrix}$

Now, instead of transmitting S, we transmit $X = PS$.

$$\Rightarrow X = \begin{bmatrix} 4.25 + i\,13.7 \\ -6.2 - i\,13 \end{bmatrix}$$

Now, at the receiver, we receive

$$Y = HX + \omega$$

$$\Rightarrow Y = H(PS) + \omega$$

$$\Rightarrow Y = \cancel{H\left(H^H (HH^H)^{-1}\right)}^{\,1} S + \omega$$

$$\Rightarrow \boxed{Y = S + \omega}$$

Manual Calculation :

$$Y = HX + \omega$$

$$\Rightarrow Y = \begin{bmatrix} 0.5 + i\,0.7 & 0.3 + i\,0.8 \\ 0.9 - i\,0.1 & 1 \end{bmatrix} \begin{bmatrix} 4.25 + i\,13.7 \\ -6.2 - i\,13 \end{bmatrix} + \omega$$

$$\Rightarrow Y = \begin{bmatrix} 1 + i1 \\ -1 - i1 \end{bmatrix} + \omega$$

$$\Rightarrow Y = \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} + \omega$$

$$\Rightarrow Y = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} + \omega$$

$$\Rightarrow \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} + \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}$$

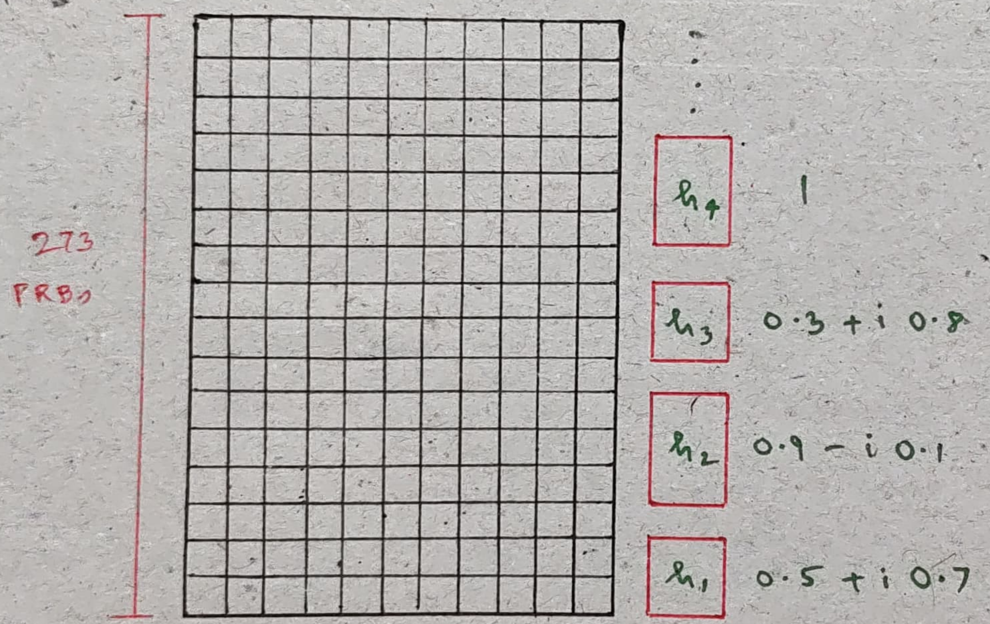
$$\Rightarrow \boxed{\begin{aligned} y_1 &= S_1 + \omega \\ y_2 &= S_2 + \omega \end{aligned}}$$

This type of Precoding is known as Precoding for MIMO.

This is a short term precoding, because the channel coefficients h_1, h_2, h_3, h_4 changes continuously. And hence, the Precoding Matrix P has to be updated continuously.

Same Precoding Matrix P has to be applied for both PDSCH and DMRS.

The channel H is not constant for all the 273 PRBs. The channel varies for different PRBs or group of PRBs.



Normally, the channel is frequency selective (ie) it varies across the subcarriers. So, we have to calculate channel H for the group of PRBs.

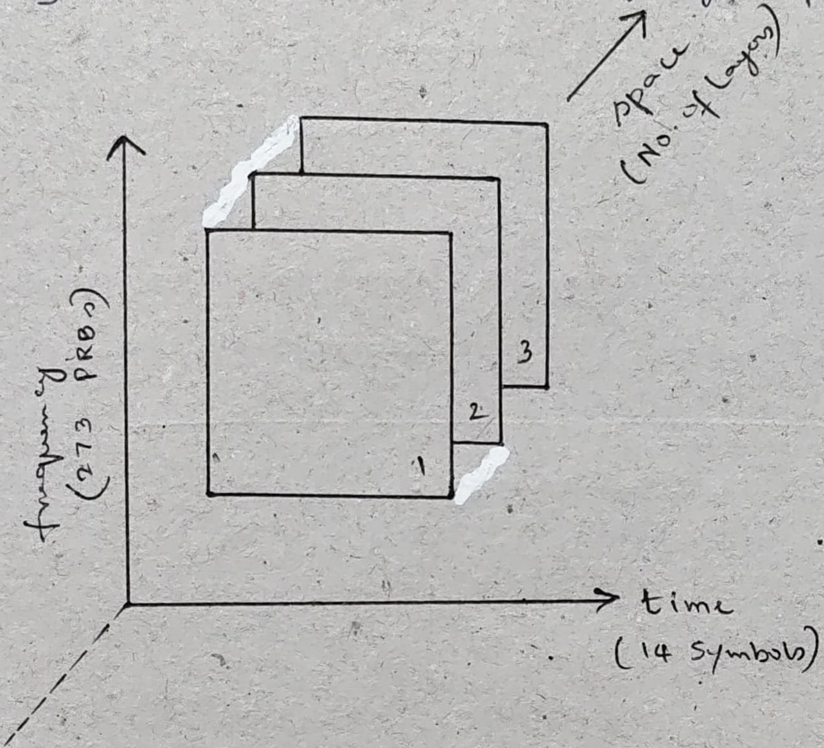
There is a concept of PRG. We group the PRBs that goes through the same channel, and use the Precoding Matrix for those group of PRBs, do the Precoding, and then transmit the data. So, the Precoding Matrix varies across the PRBs, and the quantum could be 1/2/8/16 PRB, ..., and it has to be updated frequently at a very fast pace (may be after 5 slot / 10 slot / so on...) based on how fast the channel is changing.

PDSCH / PUSCH : UE Multiplexing & Multi UE / TTI

Lets understand UE multiplexing for PDSCH and PUSCH.

Recap of PDSCH / PUSCH Resource allocation.

- 3 dimensions (time, frequency, space)



- We can multiplex UEs either in time / frequency / space
- There are 3 set of parameters defined

t → start symbol and total symbols

f → start PRB and total PRBs

→ Bitmap (if allocation type 0) for group of PRBs

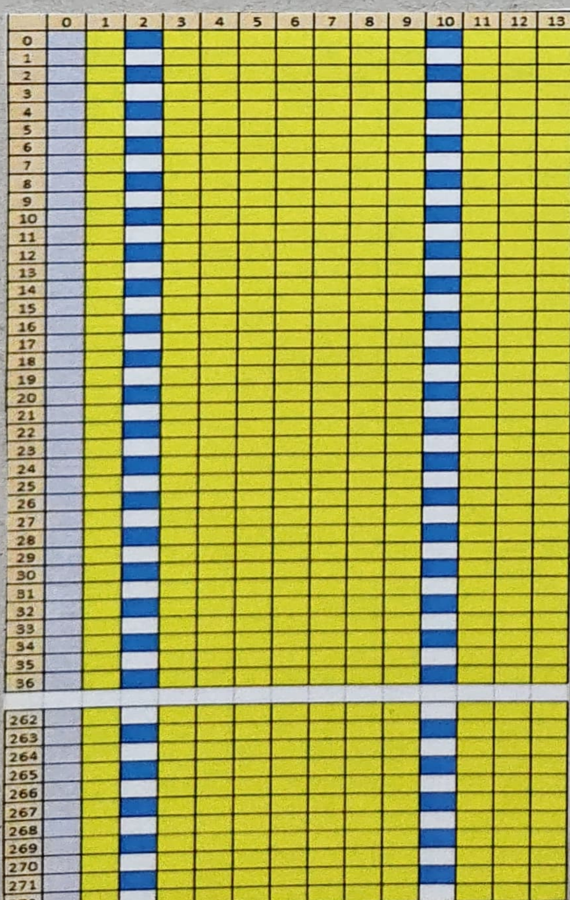
v → No. of Layers.

- Resource allocation is always Rectangular. And it cannot be like something as shown below.

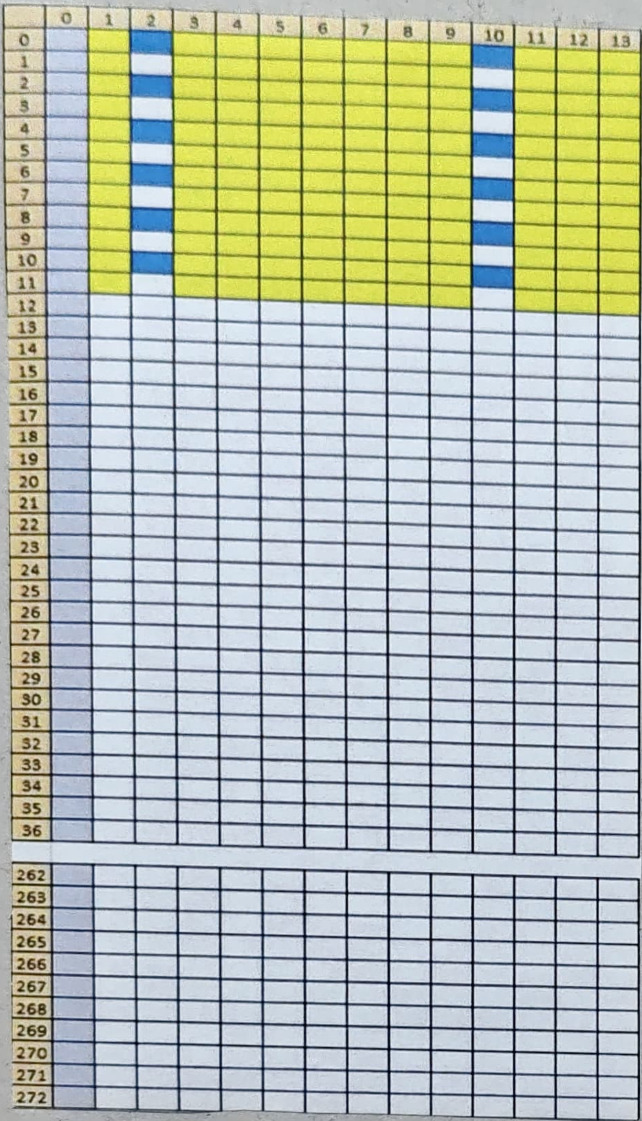
→ We will have start point and end point in time, start point and end point in frequency, and whatever applicable for single layer or antenna port remain the same for other layers as well.

(Say, we may say PDSCH allocation is from Symbol 0 to 13 (all 14 symbols), PRB_10 to PRB_60. This time and frequency dimension remains same for all the layers allocated to that UE)

Lets visualize the allocation.

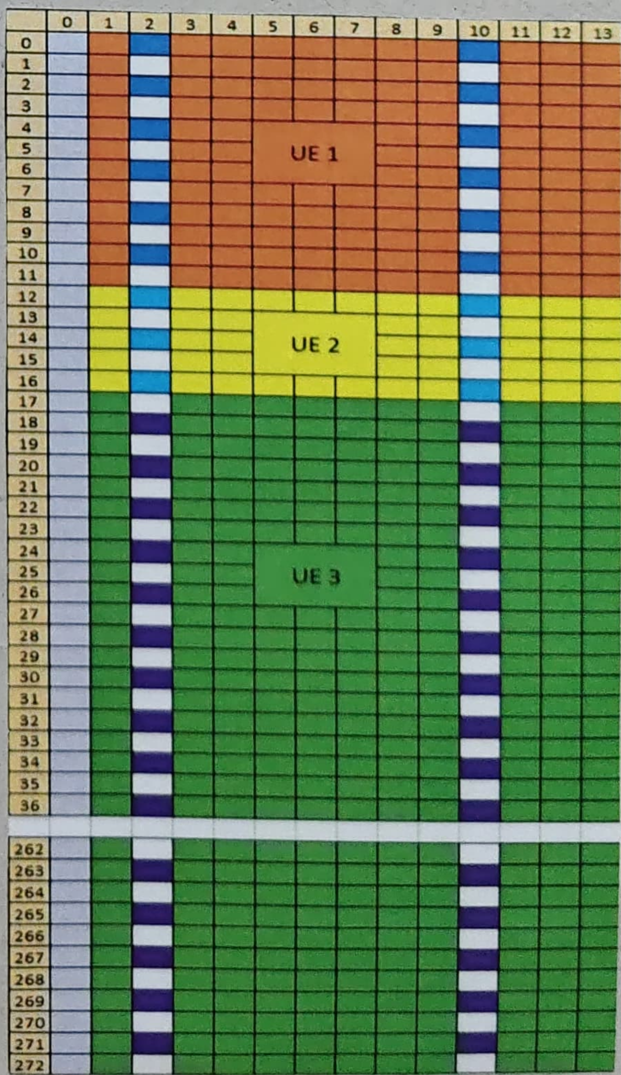


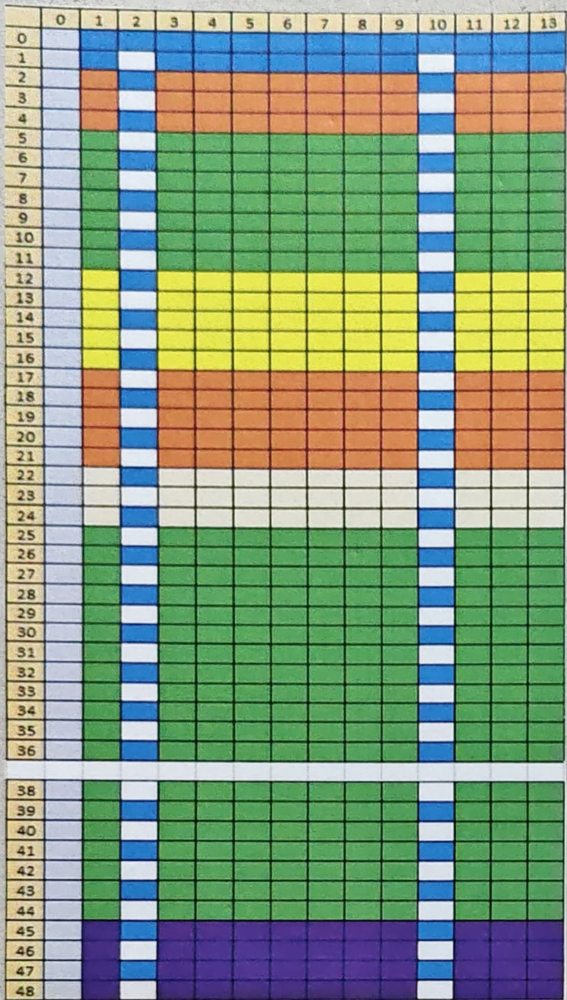
- Single UE (UE1) is allocated two DMRS on Symbols 2 and 10, and full 273 PRBs, for this particular slot.
- Allocation happens in the slot itself. It doesn't cross the slot boundary.



- This is also a single UE case, where the UE have to send less data, so the No. of PRBs are less.
- Rest of the resources are not allocated to any UE.
- Only a single UE is going in that particular slot, taking 12 PRB and full 13 symbols (Symbol 0 is for PDCCH)

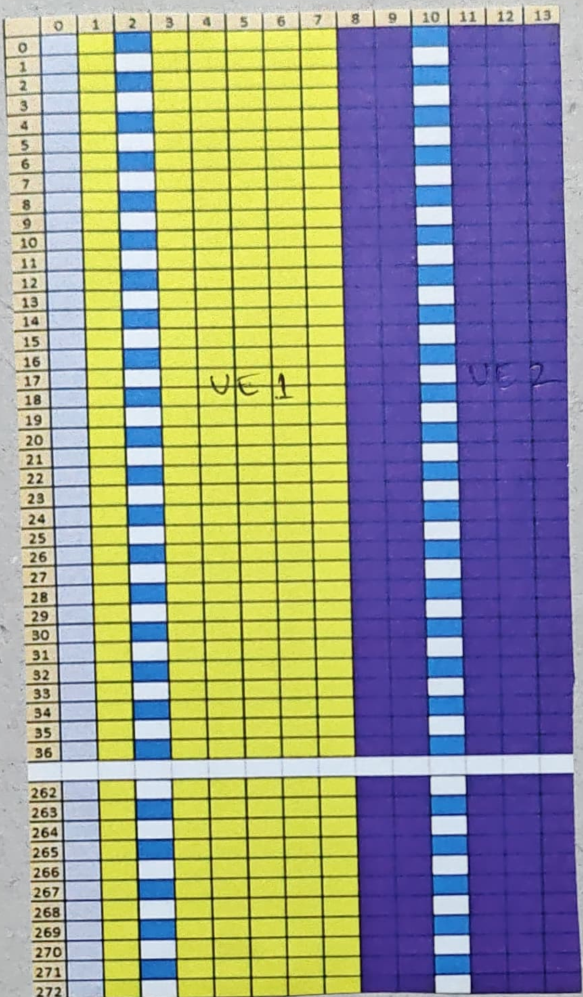
- In this case, where there are 3 UEs multiplexed across frequency.
- UE1 is allocated 12 PRBs
  UE2 is allocated 5 PRBs
  UE3 is allocated 256 PRBs
- This is called as "UEs multiplexed in frequency"





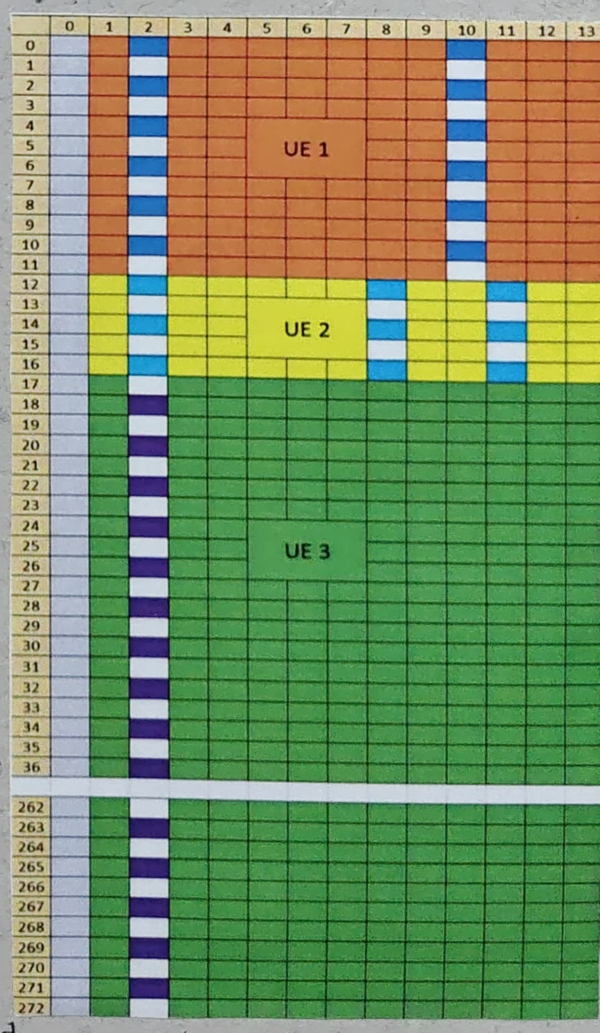
- In this case, more number of UEs are multiplexed.
- The allocation for single UE can be as low as 1 PRB
- In this case, 8 UEs are multiplexed in frequency within a single slot.
- As 1 PRB is the minimum allocation, we may have maximum 273 UEs multiplexed within a single slot. But practically it is not possible.
- The limit on No. of UEs allocated in a single slot depends on the PDCCH (ACK/NACK) and the PUCCH Capacity as well.

- In order to multiplex 273 UEs in a single slot, we also need 273 DCIs. As we cannot fit 273 DCIs as per specification even maximum number of symbols are given for PDCCH, multiplexing 273 UEs in a single slot is not possible.
- DCI capacity is discussed in PDCCH section (do refer).



- In this case, two UEs are multiplexed in time.
- In this case, UE 1 can use either mapping type A or mapping type B. But UE 2 has to use mapping type B (as defined in specification).
- UE 2 start symbol is symbol 8 and UE 1 start symbol is symbol 1.
- Practically, mostly, time multiplexing is not done. Mostly we do frequency multiplexing of the UEs.

- Though in the previous cases, it is shown, all UEs having the same DMRS symbols, but that is not the case always.
- Each UE can have different number of DMRS allocated.
- Like in this case, UE 1 having 2 DMRS, UE 2 having 3 DMRS, UE 3 having only one DMRS. So, it is independently scheduled and it depends on the UE requirement and channel condition.
- UE 3 is having more data and the channel condition is good, hence more no. of PRBs and single DMRS allocated.



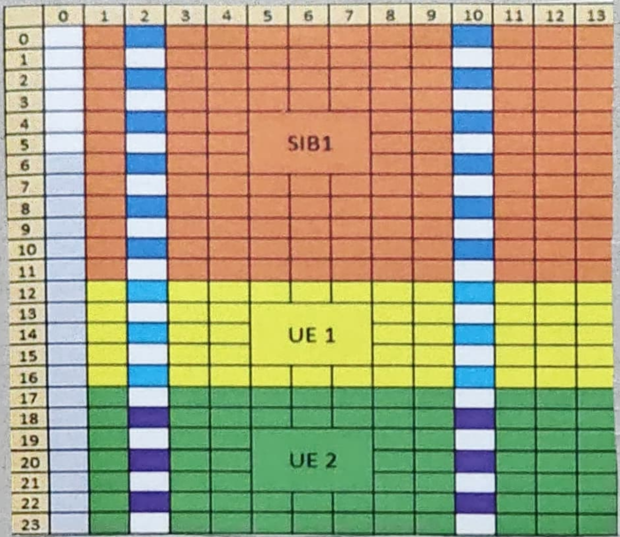
- UE 2 may be at bad channel condition and requires less data. So, less number of PRBs and 3 DMRS symbols allocated.
- UE 1 is somewhere in the middle.
- Each UE's data is processed separately and they may have different MCS as well (different modulation, encoding and TB size). (i) Each UE's processing is totally independent.
- The UEs that are multiplexed either in time / frequency / space doesn't know about the other UEs multiplexed in the same slot or not. For example, UE 3 does not know whether UE 1 and UE 2 also scheduled in the same slot.

As we know, in downlink we can have maximum 8 layers given to single UE, and in Uplink we can have maximum 4 layers.

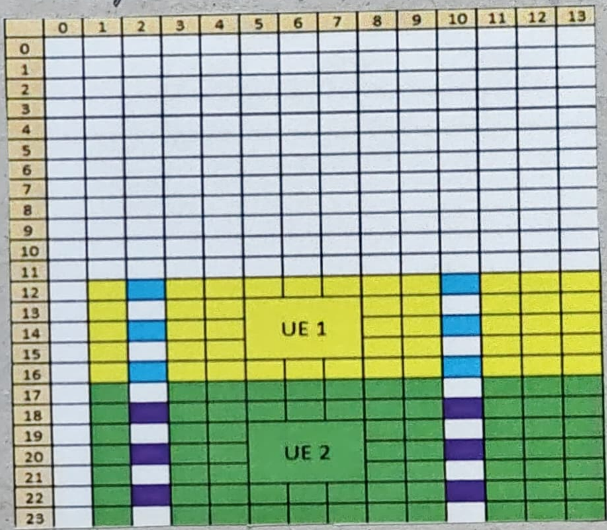
Following visualization shows

- SIB1 / any broadcast or any dedicated message taking 12 PRBs in Layer 1.
- UE1 is allocated 2 layers. So UE1 is going with two Antenna ports (port 0 and 1)
- UE2 is taking 4 layers. So UE2 data will go on Antenna ports 0, 1, 2 and 3. (the ports can be different also).

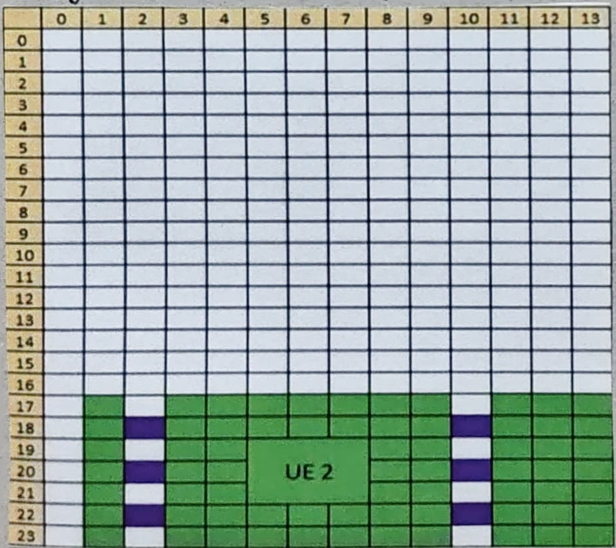
Layer 1 (Antenna Port 0)



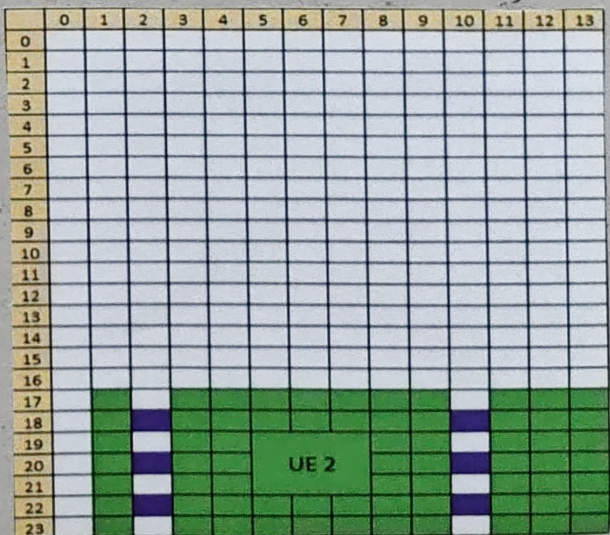
Layer 2 (Antenna port 1)



Layer 3 (Antenna port 2)



Layer 4 (Antenna port 3)



There is one more dimension where the UEs are multiplexed across space. (ie) UE1 is on one antenna port and UE2 is on another antenna port. This is known as MU-MIMO (discussed in next section).

Multiplexing multiple UEs in a single slot is called as "Multi UE per TTI"

The ground rule is, allocation is always rectangular, whether it is 1/2/4/8 layers.

UEs dont see each other; whether they are frequency/time/space multiplexed with different UEs. Each UEs has their own MCS / TB size / DMRS allocation.

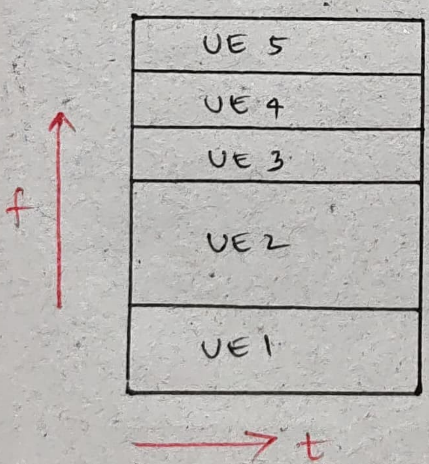
SU-MIMO and MU-MIMO

↳ Single-User MIMO ↳ Multi-User MIMO

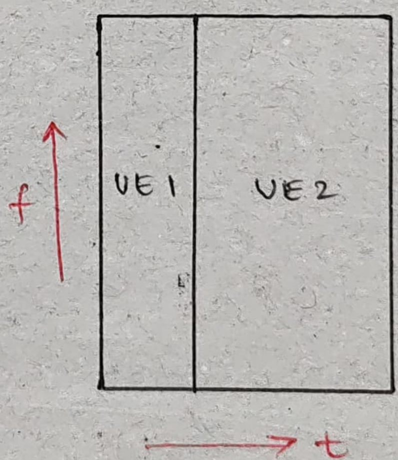
MIMO → Multiple Antennas at Transmitter and Receiver.

SU-MIMO → Serving single user, at the same time and frequency

In the previous section, we've seen how UEs are multiplexed in frequency and time.

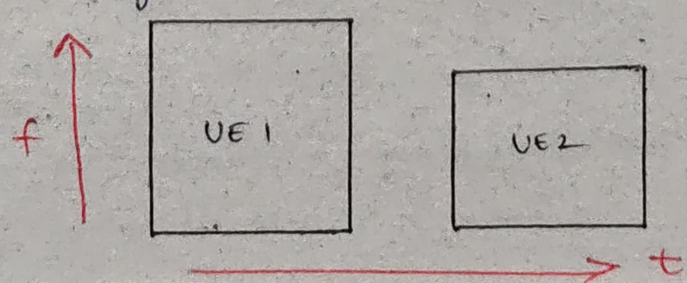


(a) Frequency Multiplexing of UEs



(b) Time Multiplexing of UEs.

And there is something called "Spatial Multiplexing" of UEs (or) MU-MIMO, where we provide same time and frequency resources to multiple UEs.



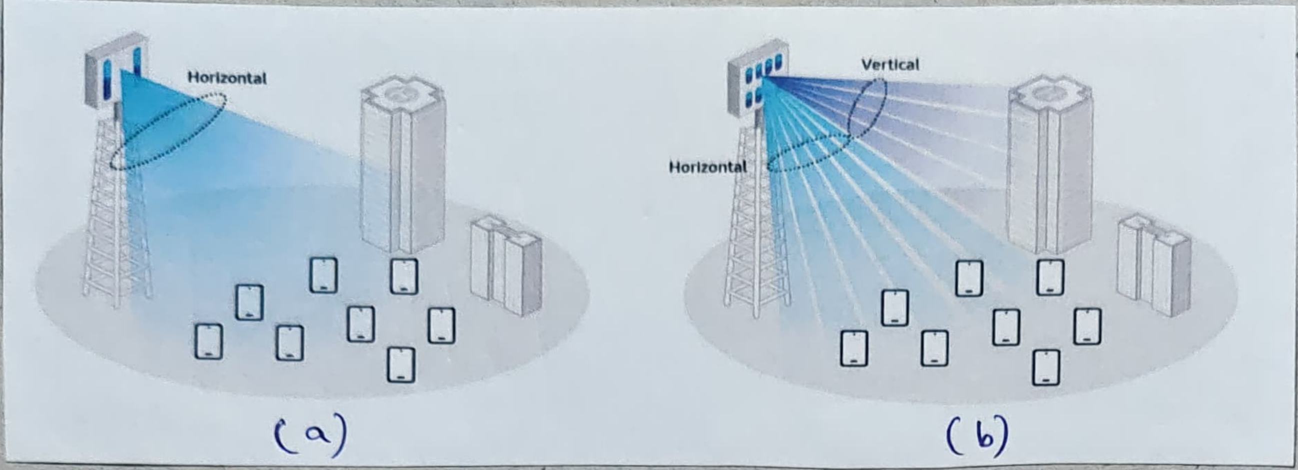
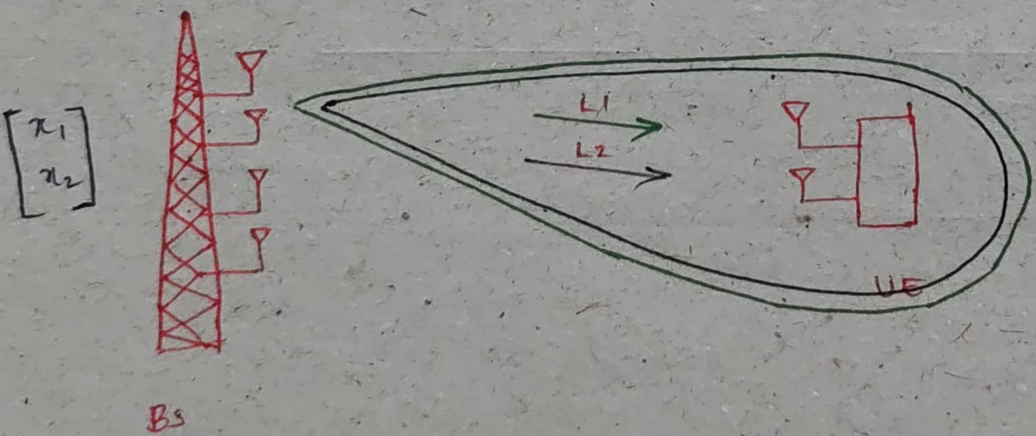


Figure (a) demonstrates, multiple UEs are served under single beam.

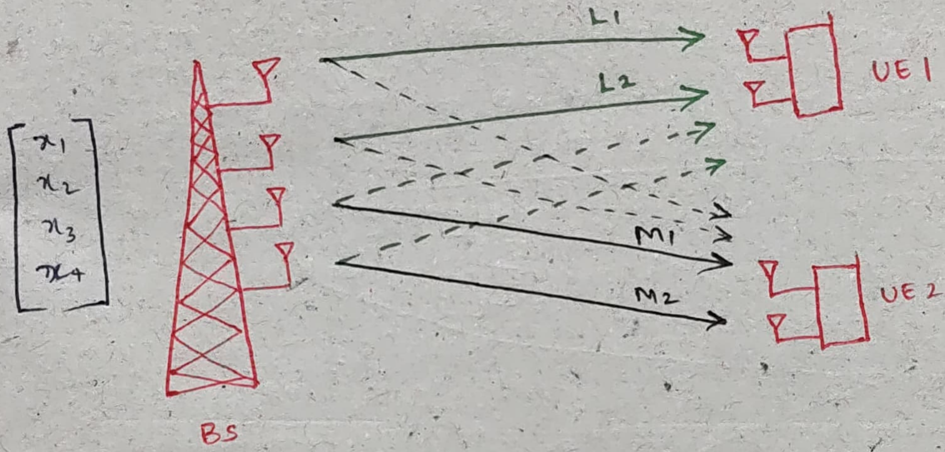
Figure (b) demonstrates, multiple UEs are served with different set of vertical and narrow beams. So, using multiple separate (non-Overlapping) beams, we can multiplex multiple UEs at the same time and frequency, to get higher gain or higher capacity. Higher gain implies higher overall throughput at gNB.

Lets understand this concept in detail.



Here, BS has 4 Antennas and UE has 2 Antennas. Using the 4 Antennas, the BS transmits <u>2 Layers</u> towards the UE. Here, Precoding is done at BS in such a way that the beams are formed towards the UE.

We've seen that, for data transmission in multiple layers, we want the channel to be uncorrelated. If the channel is correlated, then we cannot support multiple layers. So, based on the channel between these Tx. Antennas and Rx. Antennas, we find the Rank, which says maximum no. of Layers or data streams that can be transmitted.



Now, in this case, BS has 4 Antennas and two UEs each with 2 Antennas. Here, Precoding is done at BS in such a way that, the first 2 Antennas of BS forms 2 beams towards UE1, and the next 2 Antennas of BS forms 2 beams towards UE2. And of course, the beams towards UE1 will also be received by UE2. and vice versa.

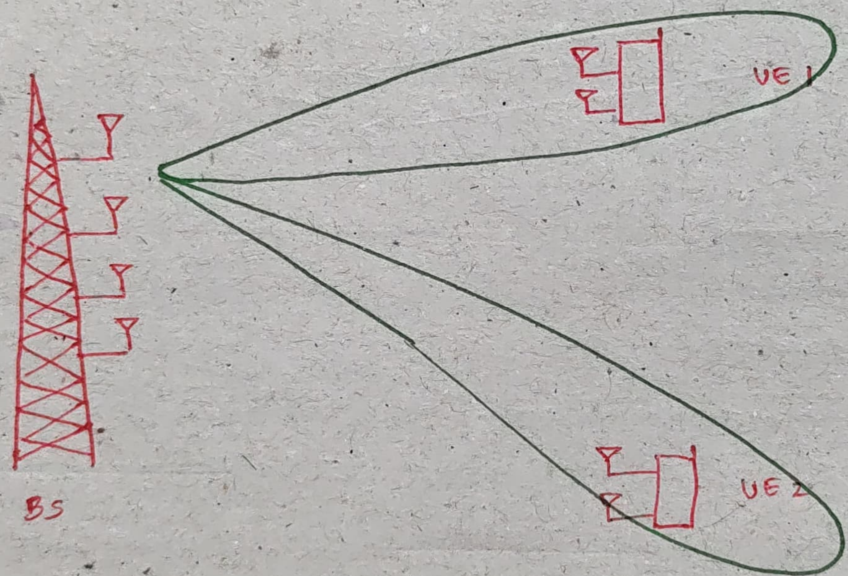
So, basically here, gNB transmits 4 layers ($L1, L_2, M_1, M_2$), using first 2 antennas for UE 1 and next 2 antennas for UE2, at the same time and frequency, using the concept of Spatial multiplexing. This is aka MU-MIMO.

The problem with this case is that the channel may not be uncorrelated to support 4 Layers. (ie) the Rank of the channel may be 2, so instead of 4 layers, it can support only 2 layers.

Note that, in this case, there is no specific precoding done, it is just direct mapping.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

This is known as Identity precoding, where the data is directly mapped to the BS antennas, to serve these two UEs. But the problem is, the channel may or may not be uncorrelated, so we may not get maximum gain.

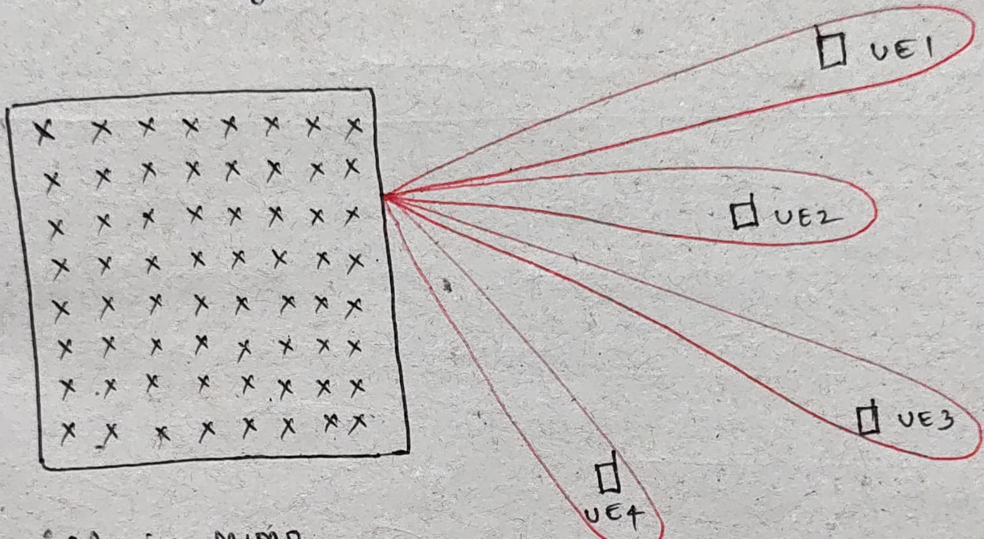


Now, in this case the two UEs are spatially apart. Here, using all the 4 antennas, the BS forms two beams, one towards UE1 and another towards UE2, making sure that the beams are narrow enough and doesn't overlap with each other. (i.e) basically the beams are well separated.

This way, we are kind of guaranteeing that, the channels between BS and UE1, and between BS and UE2 is different (close to orthogonal), and gives very high gain.

But, we can't form such narrow beams using just 4 antennas. We need more and more antennas. So, this MU-MIMO is mainly done using Massive-MIMO antennas.

Using massive-MIMO antennas, we can form multiple narrow beams towards multiple UEs. And these UEs needs to be spatially separated. This is called "User pairing".



Massive-MIMO Antenna

So, first job is, to find out the UEs that are spatially separated, using the CSI-RS transmitted on 32 antenna ports.

So, using CSI-RS, the BS finds out UEs that are well separated, and pair those UEs in the same time and frequency resources, and beamform the signals towards each and every paired UEs. The beams has to be narrow and does not overlap with each other. This gives the highest gain in the MU-MIMO.

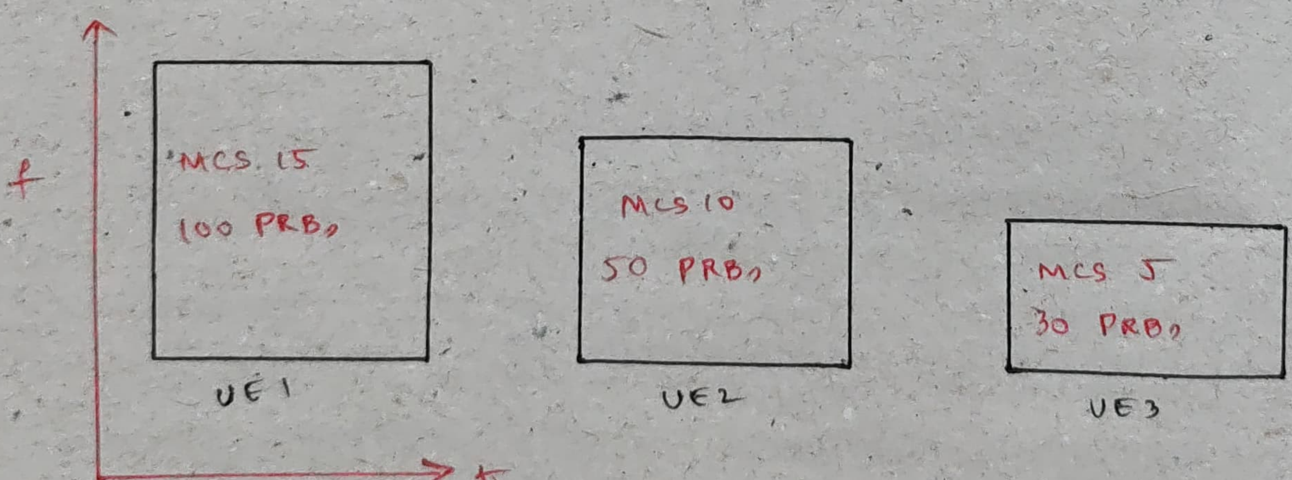
Note:
Since the beams are non-overlapping, we don't have to bother whether the DMRS are same or different. Even we have advantage that, we can use the same DMRS across all these 4 beams (Layers); and even the DMRS need not to be orthogonal to eachother as well.

This is why, in DMRS config Type 1, we can support 8 orthogonal ports. And in DMRS config Type 2, we can support 12 orthogonal ports.

Note 2:
When we have single beam serving multiple users, in this case, the DMRS has to be Orthogonal, even though the channel is uncorrelated.

So, with this concept, the BS can serve many more UEs and we are not limited by the orthogonality of the DMRS. So, this way, we can even do 16 or 24 layers, that are even practically possible as of today, and is showcased by Ericsson and Huawei.

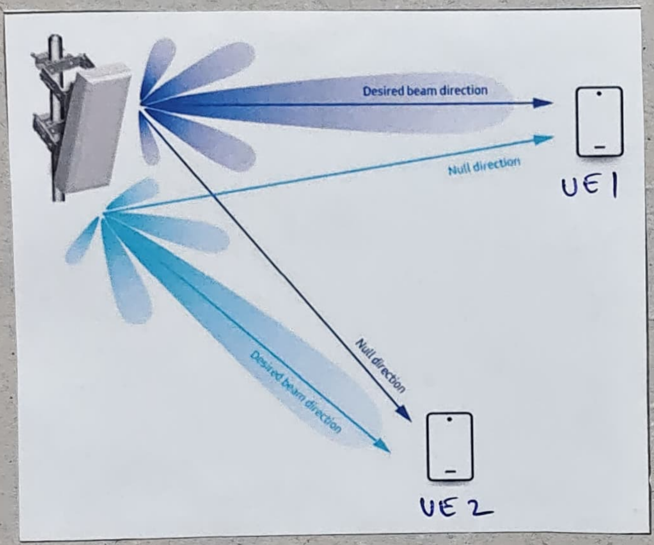
Lets now understand how resource allocation happens for the MU-MIMO.



Each UEs are allocated different number of PRBs, And Each UE may have different MCS assigned, But still they can be multiplexed.

Normally, we keep the resource allocation same across all the UEs and different MCS for each UEs. But, this may not be the case everytime.

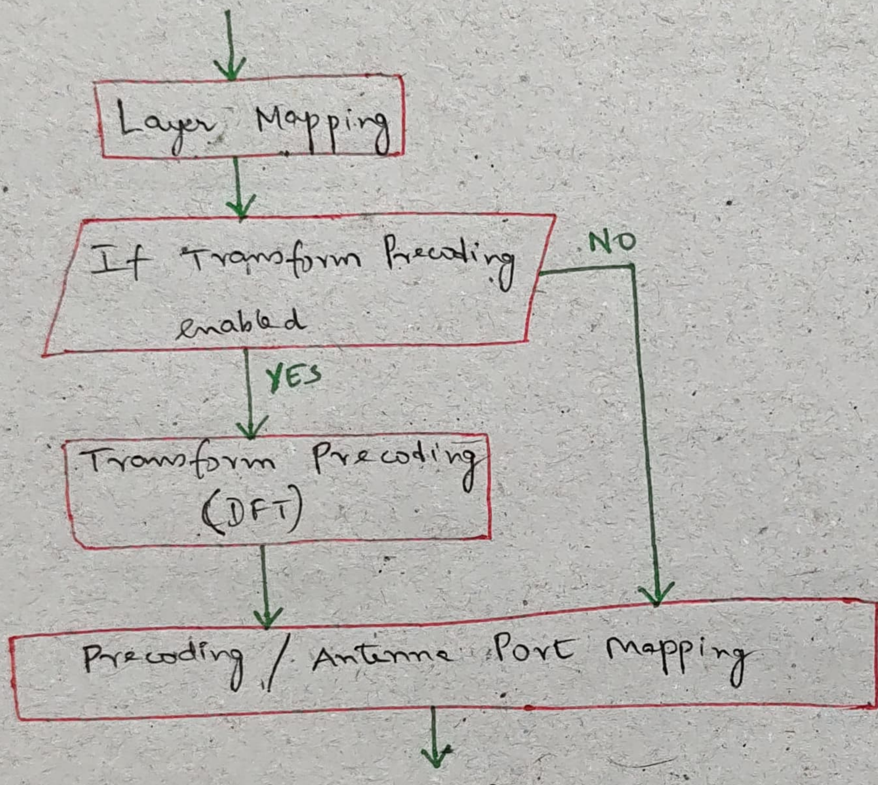
Now, when the UEs are multiplexed, they don't know whether there is another UE scheduled with the same resources.



- In this case, here UE 1 and UE 2 are multiplexed.
- Here, UE1 doesn't know that, at the same time and frequency, UE2 is also scheduled.
- Their transmission is completely independent, and they don't have any idea about the other UEs multiplexed with the same time and frequency resources.
- When we schedule in the MU-MIMO, each UE can have more than 1 layer. For example, UE1 may receive 4 layers and UE2 may receive 2 layers (depending on the channel condition between the respective UEs and BS).

PUSCH : DFT-s-OFDM | Transform Precoding

Lets understand DFT spread OFDM (or) Transform Precoding in PUSCH. We've seen the other case (i.e) when Transform precoding is not enabled, the Layer Mapped data directly goes to the Antenna Port mapping.



When Transform precoding is enabled, in this case, the Layer Mapped data gets precoded and then goes to Antenna Port mapping.

The Transform Precoding / DFT-s-OFDM is used to have Lower PAPR (we'll study this in next section). Here, in PUSCH, the Transform Precoding / DFT-s-OFDM is defined like this in the specification.

$$y^{(0)}(l \cdot M_{sc}^{PUSCH} + k) = \frac{1}{\sqrt{M_{sc}^{PUSCH}}} \sum_{i=0}^{M_{sc}^{PUSCH}-1} \tilde{x}^{(0)}(l \cdot M_{sc}^{PUSCH} + i)\, e^{-j\frac{2\pi ik}{M_{sc}^{PUSCH}}}$$

$$k = 0, \ldots, M_{sc}^{PUSCH} - 1$$

$$l = 0, \ldots, M_{symb}^{layer} / M_{sc}^{PUSCH} - 1$$

$$M_{RB}^{PUSCH} = 2^{\alpha_2} \cdot 3^{\alpha_3} \cdot 5^{\alpha_5}$$

This is basically, "take one symbol of PUSCH data and take the DFT of it" at the transmitter side. At the receiver side, we take IDFT. So, we are spreading the OFDM data across the symbol, by taking the DFT.

$\tilde{x}(0)$ → PUSCH data

k → Subcarrier index

i → Index of the PUSCH samples

M_{sc}^{PUSCH} → Total No. of Subcarriers in PUSCH

= 12 x No. of PRBs allocated for PUSCH

In case of DFT-spread-OFDM, the PRBs cannot be picked arbitrarily. So, the No. of PRBs that are used in DFT-s-OFDM case should follow

$$M_{RB}^{PUSCH} = 2^{\alpha_2} \cdot 3^{\alpha_3} \cdot 5^{\alpha_5}.$$

which means, we can't pick any other No. of PRBs. This is to make the DFT part easy. (multiple of 2/3/5).

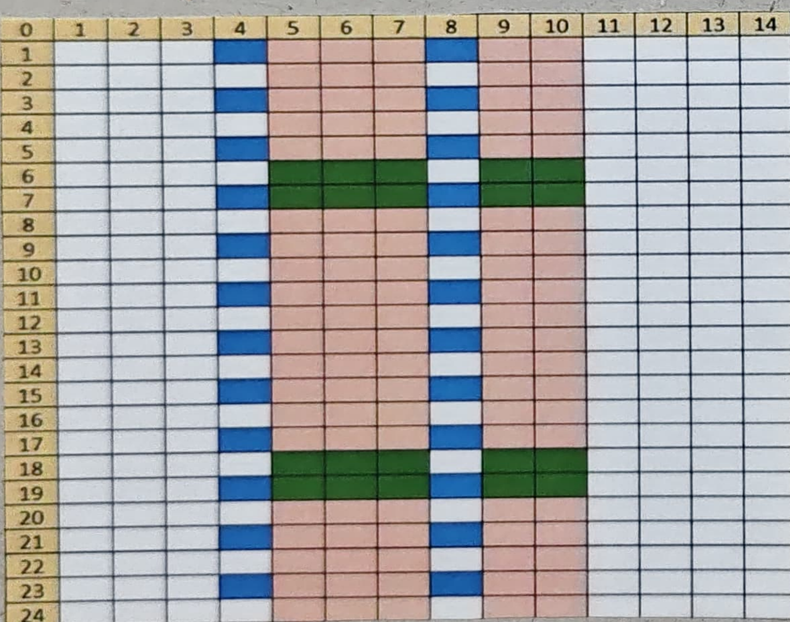
So, the DFT-s-OFDM for the first PUSCH symbol is

$$y^{(0)}\left((0\times 12)+k\right) = \frac{1}{\sqrt{12}} \sum_{i=0}^{11} \tilde{x}^{(0)}\left((0\times 12)+i\right) e^{-j\frac{2\pi i k}{12}}$$

$$\Rightarrow \boxed{y^{(0)}(k) = \frac{1}{\sqrt{12}} \sum_{i=0}^{11} \tilde{x}^{(0)}(i)\, e^{-j\frac{\pi i k}{6}}}$$

This is the DFT-s-OFDM for the first PUSCH symbol which goes to the Antenna port mapping.

- Assume, these are the two PRBs allocated for PUSCH.
- The Layer Mapped data (assume 120 modulated samples) are mapped to the respective symbols. (Symbols 5, 6, 7, 9, 10) (ie) 24 samples per symbol
- Now, we take 24 samples from each symbol and take DFT of it.



- In case of mmWave, PTRS is also transmitted along with the PUSCH.
- In this case, the Layer Mapped data (100 samples) are mapped to the respective symbols (ie) 20 samples per symbol. leaving the PTRS REs, and then take the DFT of it.

This is how the DFT-s-OFDM is followed.

Few Points to remember !

- In CP-OFDM case, we had 4 Layers, whereas in DFT-s-OFDM case, we have only 1 Layer defined. So, for higher throughput, CP-OFDM is the better choice. So, we disable the Transform Precoding.

- In case of DFT-s-OFDM, since PAPR is less, hence UE can push more power and the coverage is good. whereas, in case of CP-OFDM, the coverage is less and the UE can't push more power.

- In case of DFT-s-OFDM, there is no interleaving of DMRS with data. And only one DMRS config, is defined, since it is a single layer case.

PUSCH : MIMO, Codebook, beamforming and more

In this section, lets understand MIMO for Uplink (PUSCH), how it is done, Antenna placement on the cell phone, ...

UE → Anything that connects to the gNB.
(Mobile phone / Laptop / Customer Premises Equipment)

This section is about Mobile phones and MIMO related to it. The way we hold the mobile phone has a major impact on the UL signal transmission. For sub-6, it doesn't play a bigger role, but for mmWave communication the way our hand holds the phone plays a bigger role on signal transmission in Uplink.

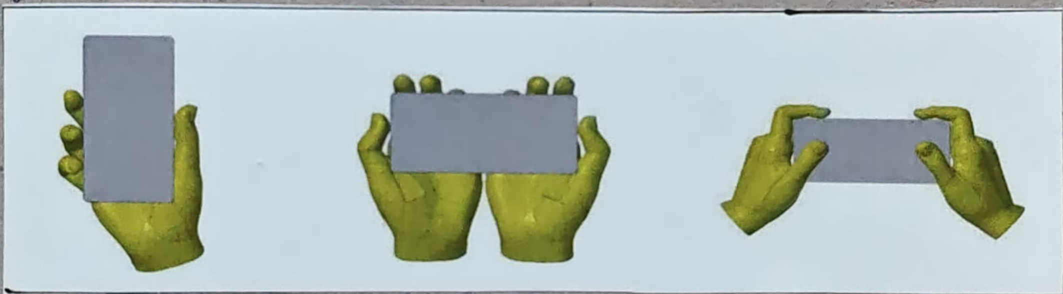


Figure shows the three different ways we normally hold the mobile phone. So, the way we place the antennas in Mobile phones should take care of the signal block due to our hands.

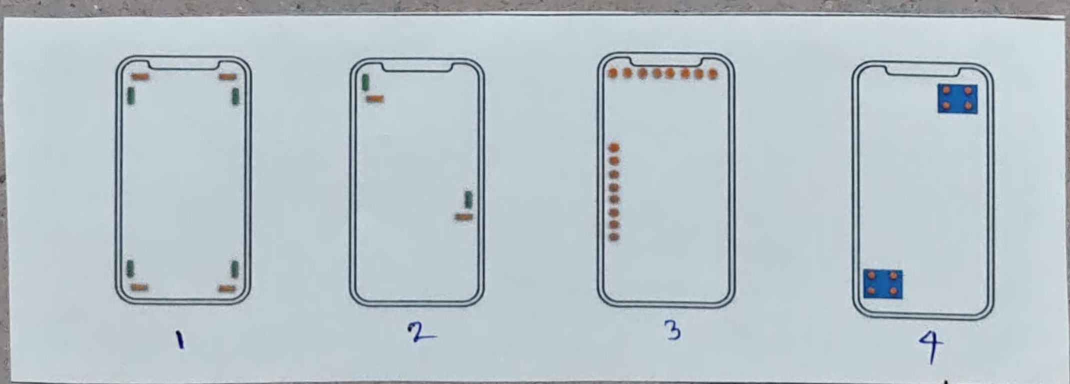


Figure shows the different way that the antennas can be placed on the mobile phones.

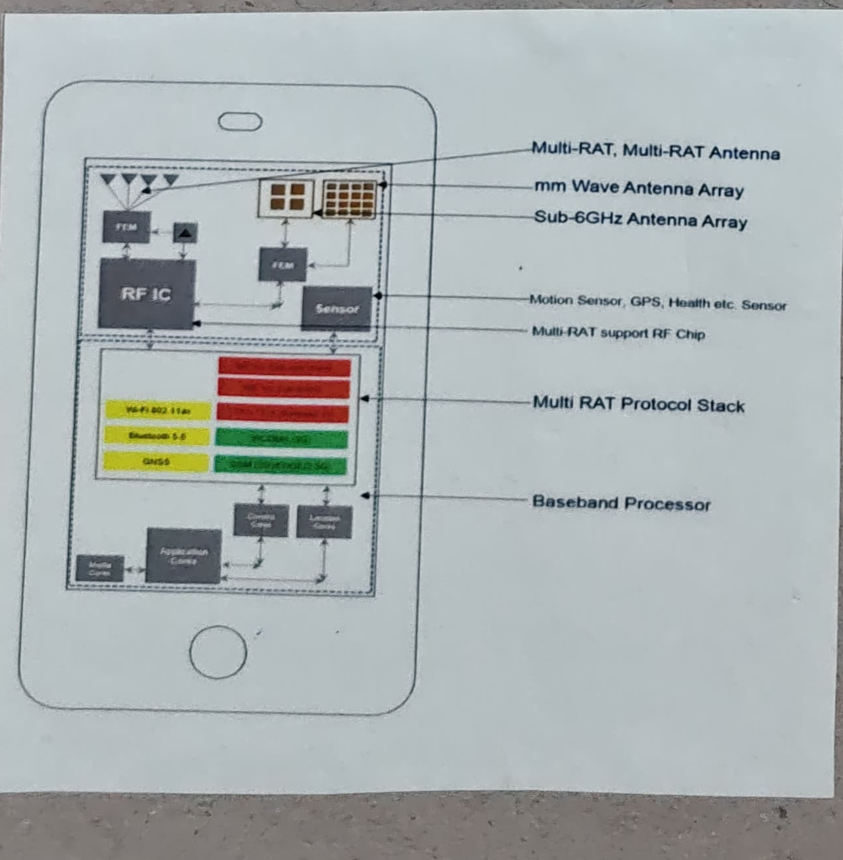
① Two cross pole antennas are placed on the 4 corners well apart from each other. In total, 8 antennas.

(2) Two pairs of cross pole antennas are placed. In total, there are 4 antennas to do the UL and DL MIMO.

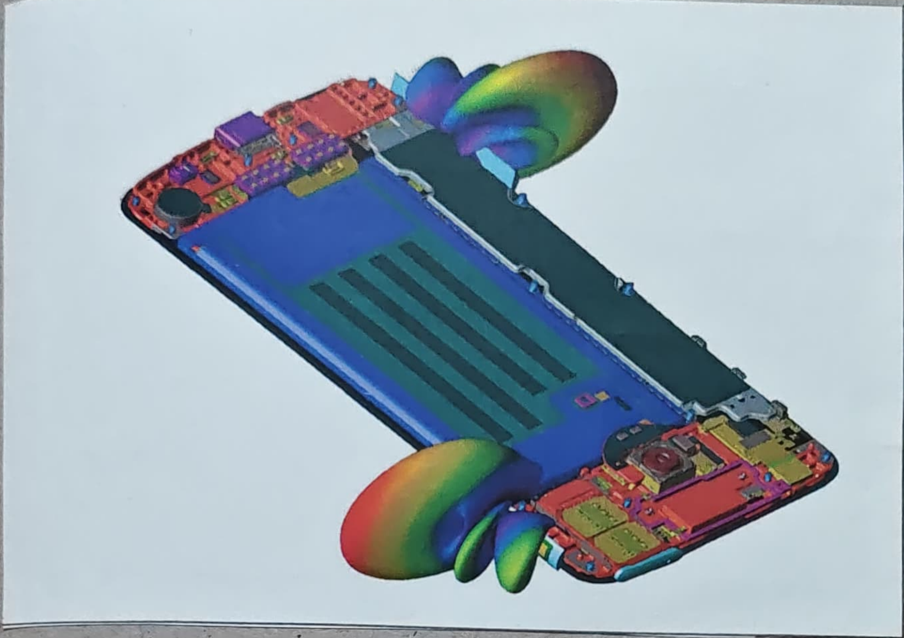
(3) Eight antennas on top, and eight antennas at the side. This is primarily for mmWave communication and can also be used for Sub-6 as well.

In mmWave case, since the carrier frequencies are large, the antenna size can be very small and can be placed together. So, we can put many number of antennas on a small place.

(4) Two Antenna panels, each having 4 antennas, are placed at two corners. In total, we have 8 antennas, using which we can do UL and DL MIMO. As per Release 15, UE is allowed to have 4x4 MIMO even though it has 8 antennas. Out of 2 antenna panels, UE can use one antenna panel to do the 4x4 MIMO.



- Figure shows two kind of antenna panels.
- The sub-6 antenna panel has <u>4 antennas</u>
- The mmWave antenna panel has <u>16 antennas.</u>



- Figure shows the visualization of Beamforming.
- UE might have 4 antennas placed at two corners.
- At a time, one panel gets activated based on the feedback from the gNB, which says, which among the two beams is good at that particular instance, and it depends on how the Mobile phone is placed.

Antenna Coherence.

Antennas can be coherent or non-coherent.

Coherent Antennas means, the Antennas are Phase aligned. (i.e) we can provide different phase to different Antennas. So, we can control the phase and we can have consistent phase relation between the signals that we are transmitting.

But in non-coherent antennas, the antennas are not phase aligned. So, we cannot control the phase between two antennas.

Normally, the non-coherent antennas are cheaper and coherent antennas are costlier. Coherent antennas can be used for Advanced beamforming and MIMO case, whereas Non-coherent antennas can be used for Spatial multiplexing, MIMO and all.

So, the UE antennas are categorized into 3 types

1. Non-coherent antennas
2. Partially coherent antennas
3. Fully coherent antennas

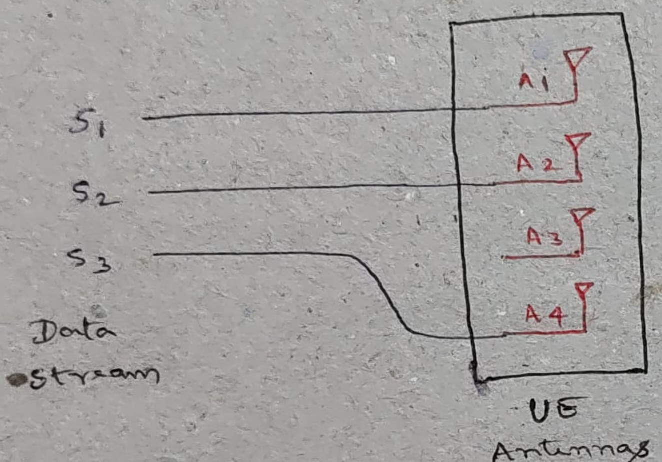


Figure shows the non-coherent case, where the data stream is directly mapped to the UE antennas. Here, there is no phase alignment, just 1-to-1 mapping.

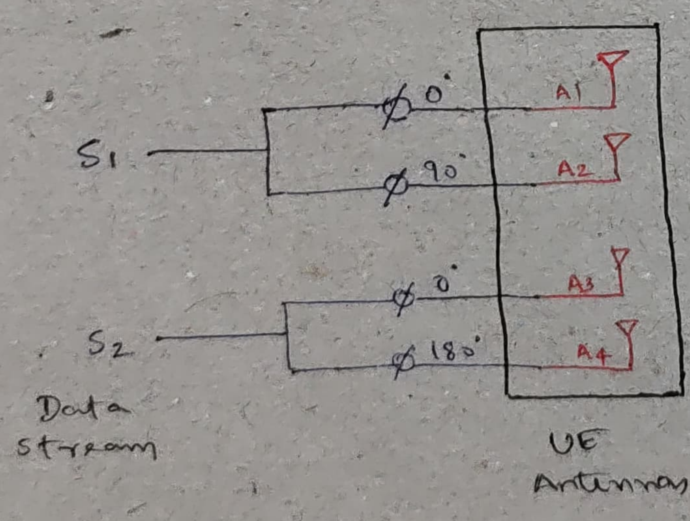


Figure shows the partially coherent case, where a pair of antennas are coherent. (ie) Antennas A_1 and A_2 are coherent, and Antennas A_3 and A_4 are coherent. But, these two pairs are not coherent.

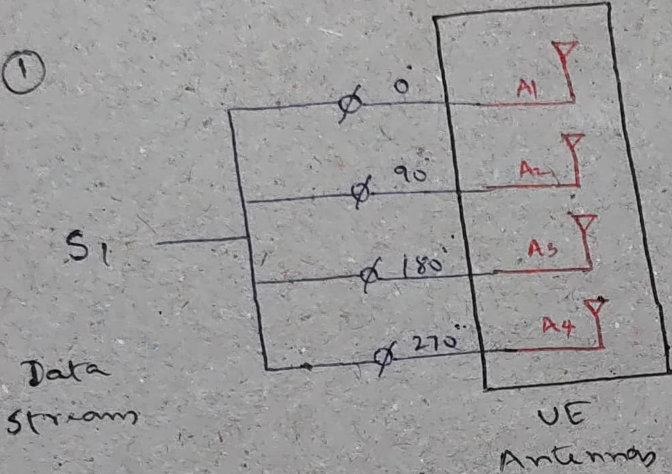
Here, the Data stream S_1 is transmitted across two antennas with 0° phase shift and 90° phase shift.

Also, the Data stream S_2 is transmitted across two antennas with 0° and 180° phase shifts.

And, we transmit same data stream using two antennas.

---

①



Figures ① and ② shows the fully coherent case, where all the antennas A1, A2, A3 and A4 are in coherence with each other. (Phase aligned).

Figure ① shows, Single Data stream S1 is transmitted on all 4 antennas with different phase shifts.

②

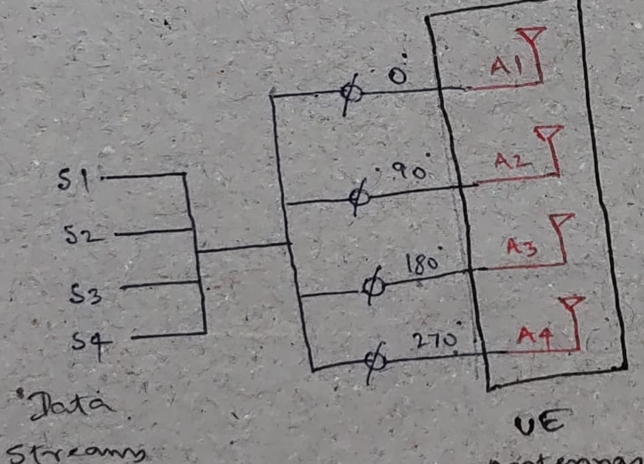


Figure ② shows, multiple data streams S1, S2, S3, S4 transmitted on all 4 antennas. Here, each stream data (say S1) will be transmitted on all the Antennas A1, A2, A3 and A4. with different phase shifts.

As the antennas at the UE side can be coherent/ Non-coherent / partially coherent, and the way the antennas are placed in UE, there are different codebooks defined in Specification.

. In this section, we'll see Codebook based MIMO. And, in the next section, we'll see non-codebook based MIMO.

Lets say, there are 2 Antenna panels at UE. For the UE to send signal, first it needs to find out which antenna panel works best, at that particular moment, for signal transmission. For that, UE sends SRS on both the panels, and gNB finds out which panel have better signal quality, and gNB signal back that (say panel 2) has good quality signals transmitted. From now onwards, UE sends SRS on panel 2 along with its Antenna capabilities. Then the gNB signals back the Rank, No. of antennas in the panel 2 to use: (ie) basically gNB signals the TPMI value (Transform Precoding Matrix Indicator).

Lets understand the table, given in specification.

Table 6.3.1.5-1: Precoding matrix W for single-layer transmission using two antenna ports.

TPMI index	W (ordered from left to right in increasing order of TPMI index)							
0 – 5	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\j\end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-j\end{bmatrix}$	-	-
	1	2	3	4	5	6		

In case of single-layer transmission using 1 antenna port, it'd be One-to-one mapping and No precoding involved.

Whereas, in case of Single-layer transmission using 2 Antenna ports, how signal transmission happens?

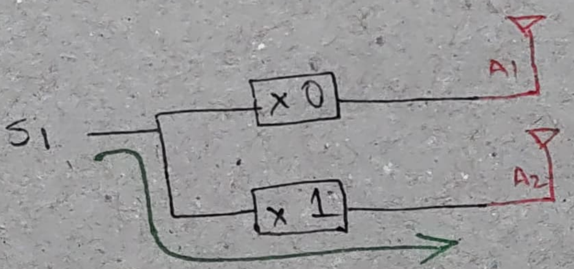
As we see, in this Table, the first two Precoding matrices are used for Non-coherent. And the rest of the Precoding matrices can be used either for non-coherent / partially coherent / fully coherent.

Note: When the antennas are fully coherent, they can be used either as non-coherent or partially coherent.

So, if the UE signals that it has non-coherent antennas, then it is supposed to use the Precoding Matrix 1 or 2, (cannot use the remaining 4 matrices)

And, if the UE signals that it has coherent antennas, then it can use any of the 6 precoding matrices. (ie) The gNB signals back which one to use.

Assume, the gNB signals back to use Precoding matrix 2. (ie) $W = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. ← Send S1 only on A1



In this case, S1 is mapped to A2 (not transmitted on A1).

Note: The Amplitude factors $\left(\frac{1}{\sqrt{2}} \text{ or } \frac{1}{2} \text{ or } \frac{1}{2\sqrt{2}}\right)$ makes sure that the total power is 1.

In case of single-layer transmission using 4 Antenna, the Precoding Matrices 1, 2, 3 and 4 are used for Non-coherent case. Say, if the gNB signals back to use Precoding Matrix 3 (ie) $W = \frac{1}{2}\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ then S1 is mapped to A3 as shown below.

Send S1 only on A3

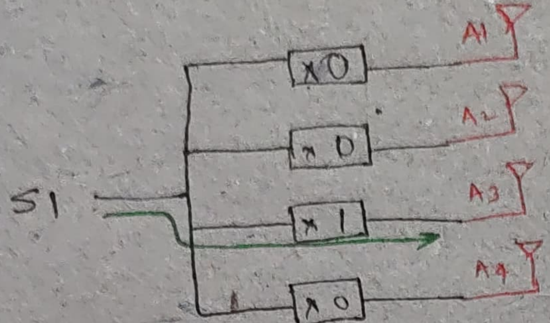
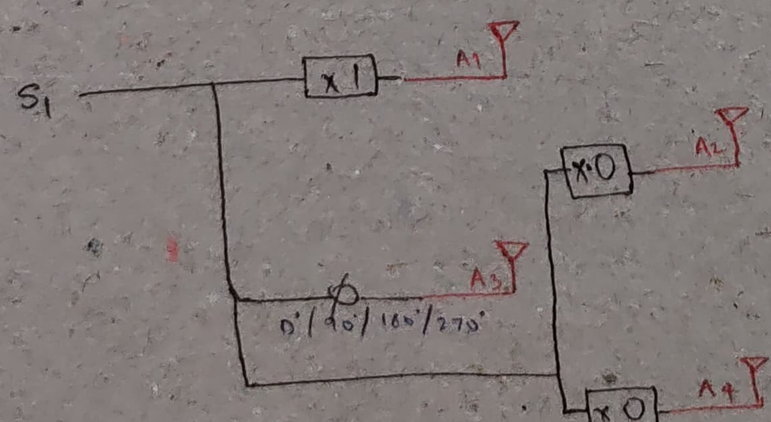


Table 6.3.1.5-2: Precoding matrix W for single-layer transmission using four antenna ports with transform precoding enabled.

TPMI index	W (ordered from left to right in increasing order of TPMI index)							
0 – 7	$\frac{1}{2}\begin{bmatrix}1\\0\\0\\0\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}0\\1\\0\\0\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}0\\0\\1\\0\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\0\\1\\0\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\0\\-1\\0\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\0\\j\\0\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\0\\-j\\0\end{bmatrix}$
8 – 15	$\frac{1}{2}\begin{bmatrix}0\\1\\0\\1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}0\\1\\0\\-1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}0\\1\\0\\j\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}0\\1\\0\\-j\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\1\\1\\-1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\1\\j\\j\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\1\\-1\\1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\1\\-j\\-j\end{bmatrix}$
16 – 23	$\frac{1}{2}\begin{bmatrix}1\\j\\1\\j\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\j\\j\\1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\j\\-1\\-j\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\j\\-j\\-1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\-1\\1\\1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\-1\\j\\-j\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\-1\\-1\\-1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\-1\\-j\\j\end{bmatrix}$
24 – 27	$\frac{1}{2}\begin{bmatrix}1\\-j\\1\\-j\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\-j\\j\\-1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\-j\\-1\\j\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1\\-j\\-j\\1\end{bmatrix}$	-	-	-	-

The next 8 matrices are used for Partially coherent case. When the UE signals gNB that it has partially coherent antennas, the gNB signals back one among these 8 matrices.

WKT, partially coherent means, a pair of two antennas are coherent.



(Figure applicable for first 4 Precoding matrices W)

Here, as per the Precoding matrices, A_1 and A_3 are coherent ; A_2 and A_4 are coherent ; But these two pairs are not coherent with each other.

Figure illustrates, single layer data is sent on two antennas A1 and A3 (both are partially coherent). out of 4 available antennas.

All the remaining matrices are for fully coherent case, when different phase combinations are given in the table.

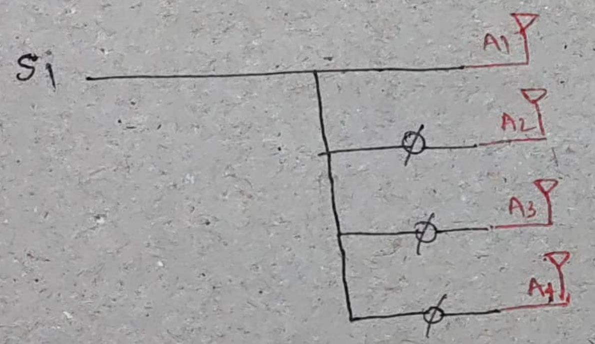


Table 6.3.1.5-4: Precoding matrix W for two-layer transmission using two antenna ports with transform precoding disabled.

TPMI index	W (ordered from left to right in increasing order of TPMI index)			
0 – 2	$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1 & 1\\j & -j\end{bmatrix}$	

As we know, for single layer transmission, Transform Precoding is enabled. For two layer and four layer transmission, Transform Precoding is disabled.

In case of 2 layer transmission using 2 Antennas, the Precoding Matrix $W = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$ is used for Non-coherent case, and the remaining two matrices are used for coherent case.

Table 6.3.1.5-5: Precoding matrix W for two-layer transmission using four antenna ports with transform precoding disabled.

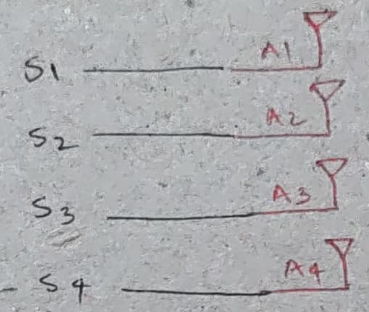
TPMI index	W (ordered from left to right in increasing order of TPMI index)			
0 – 3	$\frac{1}{2}\begin{bmatrix}1&0\\0&1\\0&0\\0&0\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1&0\\0&0\\0&1\\0&0\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1&0\\0&0\\0&0\\0&1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}0&0\\1&0\\0&1\\0&0\end{bmatrix}$
4 – 7	$\frac{1}{2}\begin{bmatrix}0&0\\1&0\\0&0\\0&1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}0&0\\0&0\\1&0\\0&1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1&0\\0&1\\1&0\\0&-j\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1&0\\0&1\\1&0\\0&j\end{bmatrix}$
8 – 11	$\frac{1}{2}\begin{bmatrix}1&0\\0&1\\-j&0\\0&1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1&0\\0&1\\-j&0\\0&-1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1&0\\0&1\\-1&0\\0&-j\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1&0\\0&1\\-1&0\\0&j\end{bmatrix}$
12 – 15	$\frac{1}{2}\begin{bmatrix}1&0\\0&1\\j&0\\0&1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}1&0\\0&1\\j&0\\0&-1\end{bmatrix}$	$\frac{1}{2\sqrt{2}}\begin{bmatrix}1&1\\1&1\\1&-1\\1&-1\end{bmatrix}$	$\frac{1}{2\sqrt{2}}\begin{bmatrix}1&1\\1&1\\j&-j\\j&-j\end{bmatrix}$
16 – 19	$\frac{1}{2\sqrt{2}}\begin{bmatrix}1&1\\j&j\\1&-1\\j&-j\end{bmatrix}$	$\frac{1}{2\sqrt{2}}\begin{bmatrix}1&1\\j&j\\j&-j\\-1&1\end{bmatrix}$	$\frac{1}{2\sqrt{2}}\begin{bmatrix}1&1\\-1&-1\\1&-1\\-1&1\end{bmatrix}$	$\frac{1}{2\sqrt{2}}\begin{bmatrix}1&1\\-1&-1\\j&-j\\-j&j\end{bmatrix}$
20 – 21	$\frac{1}{2\sqrt{2}}\begin{bmatrix}1&1\\-j&-j\\1&-1\\-j&j\end{bmatrix}$	$\frac{1}{2\sqrt{2}}\begin{bmatrix}1&1\\-j&-j\\j&-j\\1&-1\end{bmatrix}$		

When there are 2 layers, and 4 antennas case, then the first 6 matrices are used for non-coherent case. (i) These 6 precoding matrix helps to select which two antennas to select, out of 4, for transmission in two layers. The next 8 matrices are used for partially coherent case. And the last 8 matrices are used for fully coherent case.

Table 6.3.1.5-7: Precoding matrix W for four-layer transmission using four antenna ports with transform precoding disabled.

TPMI index	W (ordered from left to right in increasing order of TPMI index)			
0 – 3	$\frac{1}{2}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$	$\frac{1}{2\sqrt{2}}\begin{bmatrix}1&1&0&0\\0&0&1&1\\1&-1&0&0\\0&0&1&-1\end{bmatrix}$	$\frac{1}{2\sqrt{2}}\begin{bmatrix}1&1&0&0\\0&0&1&1\\j&-j&0&0\\0&0&j&-j\end{bmatrix}$	$\frac{1}{4}\begin{bmatrix}1&1&1&1\\1&-1&1&-1\\1&1&-1&-1\\1&-1&-1&1\end{bmatrix}$
4	$\frac{1}{4}\begin{bmatrix}1&1&1&1\\1&-1&1&-1\\j&j&-j&-j\\j&-j&-j&j\end{bmatrix}$			

Similarly for 4 layers transmission with 4 antennas case, $W = \frac{1}{2}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$ is used for non-coherent case, where each stream is directly mapped to the antennas.



Whereas, the next 2 matrices is used for partially coherent case.

And, the last 2 matrices is used for fully coherent case, where each stream will go into each and every antenna with different phase.

PUSCH : Non-Codebook based MIMO

Lets understand how the UE can do MIMO without the codebook.

Non-codebook based MIMO →	UE needs to find its own codebook, Apply it on the PUSCH, and then transmit the signal.

What are the things that are required to do the non-codebook based MIMO?

(i) When UE has multiple panels, it needs to find out which panel to use.

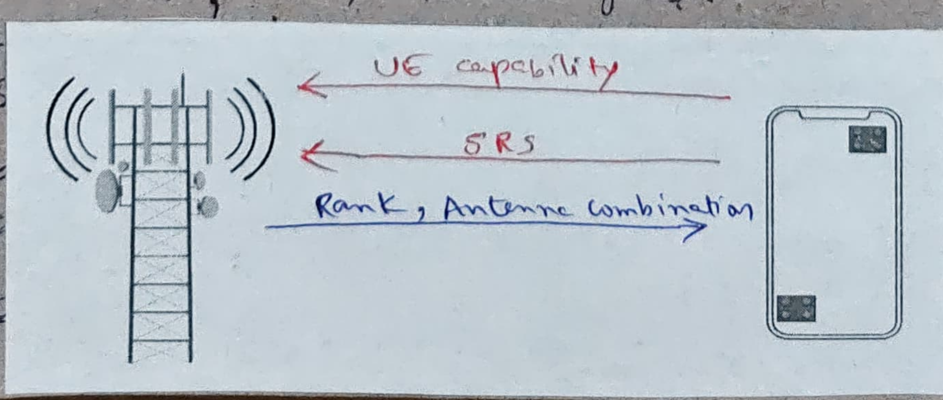
(ii) UE needs to find out the Precoding weights.

(iii) Maximum number of layers, that gNB should know, coz gNB is the one scheduling it.

(iv) Antenna combination.

These are the things that UE or gNB should know, when they want to do non-codebook based MIMO.

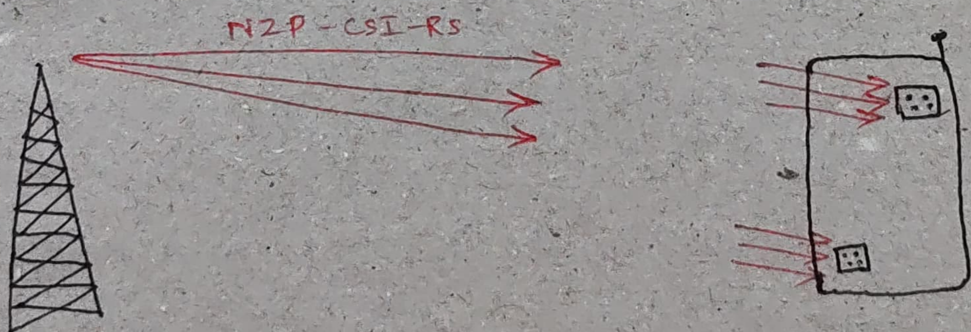
Even before that, UE signals to gNB, whether it supports codebook / non-codebook based MIMO or not in the UE capability message (UL). (ii) UE sends the Rank, available Antennas, Non-codebook based MIMO supported or not, etc., to the gNB.



When the gNB schedules PUSCH for a particular UE, it should know the channel condition (i.e. How many layers & which antenna combination is most suitable). For them to be known by gNB, UE need to signal SRS towards gNB • to find out No. of layers and Antenna combination.

For each antenna, there is SRS associated with it. So, UE will send SRS using all the antennas. For example, UE sends SRS from its 4 different antennas, gNB does channel estimation, and found that "Antenna ① and ③ works best, and maximum Rank = 1". (Rank = 1 ⇒ No. of layers = 1) So, gNB schedules single layer on two Antennas, which is communicated to the UE when the PUSCH is scheduled.

Now, UE also needs to find out, which Antenna Panel to use. Lets say, here there are two panels, and UE needs to find out the Precoding weights.



So, to find out the Precoding weights and Antenna panels, UE uses the NZP-CSI-RS (Non-Zero-Power) transmitted from gNB. The BS transmits CSI-RS, received by both the panels of UE. The UE estimates the channel using CSI-RS and find out which Antenna panel works best and also finds the Precoding coefficients.

Lets say for example, UE found that Antenna panel 2 works best and also found the Precoding Matrix W.

Now, the BS already signaled UE that, "No. of Layer / Rank = 1, use Antenna ① and ③". So, UE schedules Single Layer data with Antenna ① and ③.

$$\begin{bmatrix} x \\ 0 \\ y \\ 0 \end{bmatrix} \text{[Single Layer data]}$$

Here, the precoding weights x and y are calculated from CSI-RS. Hence, the UE transmits the Single Layer data on Antenna ① and ③ in Uplink.

This is the non-codebook based MIMO.

Can we use <u>PUSCH-DMRS</u> instead of <u>SRS</u>?

Yes! Since both carries the same information, PUSCH-DMRS can be used when we are moving from higher rank to lower rank, or when the rank is not changing. But PUSCH-DMRS cannot be used when we are moving from Lower rank to higher rank.

Lets say for example, UE is transmitting currently on all four Layers (Rank = 4) on all 4 antennas.

Now, the gNB estimates the channel using PUSCH-DMRS, and found the channel is not good. Hence gNB signals UE to switch back to two layers (Rank = 2).

So, the UE starts sending PUSCH data using two layers (Rank = 2).

Now, again the gNB estimates the channel using PUSCH - DMRS, and found the channel is not good. Hence gNB signals UE to switch back to Single Layer (Rank = 1).

So, the UE starts sending PUSCH data using single layer (Rank = 1).

Now, at this stage, there is no way for the UE or gNB to find out whether it can move from Rank 1 to Rank 2/3/4. Because, the gNB should know the information about other antennas in UE. So for that, UE is supposed to send SRS on all 4 Antennas, and gNB estimates the Rank and Antenna combination, and signals back to the UE.

So, DMRS can be used instead of SRS, only to move from higher rank to lower rank. And UE must use SRS for the best estimate in the uplink.