# An overview on M-MIMO Hybrid Design with ML

Pin-Jing, Li

*ouo.ee11@nycu.edu.tw*
Advisor: Professor Carrson C. Fung

National Yang Ming Chiao Tung University

September 23, 2025

# Reference


[1] R. M. Dreifuerst and R. W. Heath Jr., "Massive MIMO in 5G: How Beamforming, Codebooks, and Feedback Enable Larger Arrays," *IEEE Communications Magazine*, vol. 61, no. 12, pp. 18–23, Dec. 2023, doi: 10.1109/MCOM.001.2300064.

[2] R. M. Dreifuerst and R. W. Heath Jr., "Neural Codebook Design for MIMO Network Beam Management," *IEEE Transactions on Wireless Communications*, vol. 24, no. 5, pp. 3909–3922, May 2025, doi: 10.1109/TWC.2025.3536290.

# CSI measurement and transmission

- **Reference Signals**: Downlink transmission for synchronization, channel estimation and handover.
  - **SSB**: narrowband, with more constraints but multiple transmission available.
  - **CSI-RS**: wideband, with more flexibility but with computation overhead.
- **Feedback**: the CSI report sent back from the UEs.
  - includes Channel Quality Indicator, Reference signal indicator, rank indicator and precoding matrix indicator.
  - With limited resourses counstraints and quantized CSI.

# Beam management

Beam management unify reference signals and feedback to design the beams accordingly.

1. **Initial Access**: initial synchronization of UEs with beamformed SSBs.
2. **Beam Reporting**: SSB CSI reports for selecting cells and perform coarse beam training
3. **Beam Refinement**: CSI-RS feedback to refine the beam design.

# Beam Management Configuration (1)

**System Setup:**

- $C$ base stations with overlapping coverage regions
- Each BS equipped with:
  - Planar antenna array of size

$$N_T^{\mathrm{RF}} = 2N_x^{\mathrm{RF}} \times N_y^{\mathrm{RF}}$$

  - Fully connected hybrid format with $b_{\mathrm{phase}}$-bit phase shifters
  - Connections to $2N_x \times N_y$ digital ports
- Users:
  - $U$ UEs in the scene where each has an $N_R$-element fully digital ULA
- Channels:
  - Multi-cell MU-MIMO OFDM model
  - Across $T$ time slots and $K$ OFDM symbols

# Beam Management Configuration (2)

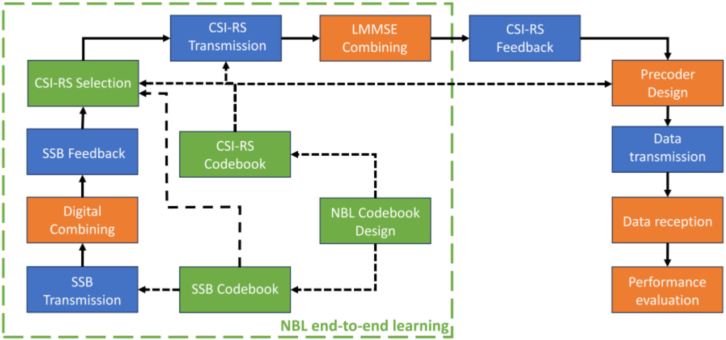


Figure: block diagram of the beam-management structure

# SSB initial access and beam reporting (1)

**SSB Transmission**

- Number of SSB frequency resources: $K^{\text{SSB}}$
- SSB timeslots: $T_i^{\text{SSB}}$
- Received SSB demodulation reference signal (DMRS):

$$\mathbf{y}_{c,u,t,k}^{\text{SSB}_i} = \frac{1}{\sqrt{KN_T}}\mathbf{H}_{c,u,t,k}\mathbf{f}_{c,i}s_{c,t,k}^{\text{DMRS}} + \sum_{c' \neq c}\mathbf{H}_{c',u,t,k}\mathbf{f}_{c',i}s_{c',t,k}^{\text{DMRS}} + \mathbf{n}_{u,t,k} \quad (1)$$

- Analog beamformer:

$$\mathbf{f}_{c,i} \in \mathbf{B}_c^{\text{SSB}}$$

  controlled from a single digital port.
- Digital combining at UE: maximizes received SSB reference signal.

**Reference Signal Received Power (RSRP):**

- For UE $u$ with beam $i$ from BS $c$:

$$\mathrm{RSRP}_{c,i,u} = \sum_{k \in K^{\mathrm{SSB}}} \sum_{t \in T_i^{\mathrm{SSB}}} \left\| \mathbf{y}_{c,u,t,k}^{\mathrm{SSB}_i} \right\|^2$$

- feedback the quantized RSRP

$$\mathbf{p} = \left\{ \max_i \mathrm{RSRP}_{c,i,u} \right\}_{u=1}^{U}$$

- and the Best cell-beam index selection:

$$\mathbf{m}, \mathbf{b} = \left\{ \arg\max_{c,i} \mathrm{RSRP}_{c,i,u} \right\}_{u=1}^{U}$$

# CSI-RS Transmission and Beam Refinement (1)

**Refined Beam Training:**

- After SSB feedback, refined beam training is performed by transmitting CSI-RS using a **hierarchical codebook strategy**.
- CSI-RS occupy significantly larger resources and enable pilot-based channel estimation.
- Typical deployment:
  - Start with a large CSI-RS codebook ($\gg L_{\max}$).
  - Perform subset selection based on SSB feedback.
- Active CSI-RS codebook subset: $\mathbf{B}_c^{\mathrm{sub}}$

# CSI-RS Transmission and Beam Refinement (2)

**Precoder-Based Selection:**

- Instead of searching over rank-1 beam vectors, selection is done over **precoder matrices**.
- Allow possible multi-stream communication
- Codebook selection is by maximizing the cross-correlation between codebooks:

$$\mathbf{C} = \max_{s} \left\langle \mathbf{B}_{c}^{\mathrm{SSB}_i}, \mathbf{B}_{c,s}^{\mathrm{CSI\text{-}RS}_j} \right\rangle, \qquad \forall\, i,\, j. \tag{2}$$

# CSI-RS Transmission and Beam Refinement (3)

**Pilot transmission & received CSI-RS signal**

- Active CSI-RS subset at BS $c$: $\mathbf{B}_c^{\mathrm{sub}}$ with beam group size $B_g$.
- Training pilot symbols: $\mathbf{s}_{c,t,k}^{\mathrm{tr}} \in \mathbb{C}^{B_g \times 1}$.
- Received (pre-combining) $i$th CSI-RS transmission at UE $u$:

$$\mathbf{y}_{u,t,k}^{\text{CSI-RS}_i} = \mathbf{H}_{b_u,u,t,k}\,\mathbf{B}_{b_u,i}^{\mathrm{sub}}\,\mathbf{s}_{b_u,t,k}^{\mathrm{tr}} \tag{3}$$

$$+ \sum_{b' \neq b_u} \mathbf{H}_{b',u,t,k}\,\mathbf{B}_{b',i}^{\mathrm{sub}}\,\mathbf{s}_{b',t,k}^{\mathrm{tr}} + \mathbf{n}_{u,t,k}, \tag{4}$$

- $b_u$: serving BS selected at the SSB stage for UE $u$.
- $\mathbf{H}_{c,u,t,k} \in \mathbb{C}^{N_R \times N_T}$: channel from BS $c$ to UE $u$.
- $\mathbf{B}_{c,i}^{\mathrm{sub}} \in \mathbb{C}^{N_T \times B_g}$: CSI-RS precoder for the $i$th CSI-RS
- $\mathbf{n}_{u,t,k} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$.

# CSI-RS Transmission and Beam Refinement (4)

**LMMSE receive combining (per UE $u$)**

- Interference-plus-noise covariance:

$$\mathbf{R}_{u,t,k} = \left( \sum_{c} \mathbf{H}_{c,u,t,k} \mathbf{B}_{c,i}^{\text{sub}} \left( \mathbf{H}_{c,u,t,k} \mathbf{B}_{c,i}^{\text{sub}} \right)^{H} \right) + C N_T \sigma^2 \mathbf{I}. \tag{5}$$

- Desired effective channel (for chosen CSI-RS precoder index $i$ at serving BS $b_u$):

$$\mathbf{V}_{i,u,t,k} \triangleq \mathbf{H}_{b_u,u,t,k} \mathbf{B}_{b_u,i}^{\text{sub}}. \tag{6}$$

- LMMSE combiner:

$$\mathbf{W}_{u,t,k}^{\text{CSI-RS}} = \mathbf{R}_{u,t,k}^{-1} \mathbf{V}_{i,u,t,k}. \tag{7}$$

# CSI-RS Transmission and Beam Refinement (5)

SINR of the $(t, k)$th time-freq resources of the $i$th CSI-RS with the LMMSE combiner:

$$s_{u,t,k,n_r}^{\text{CSI-RS}_i} = \frac{(\mathbf{V}_{i,u,t,k})_{n_r}^H \, \mathbf{R}_{u,t,k}^{-1} \, (\mathbf{V}_{i,u,t,k})_{n_r}}{1 - (\mathbf{V}_{i,u,t,k})_{n_r}^H \, \mathbf{R}_{u,t,k}^{-1} \, (\mathbf{V}_{i,u,t,k})_{n_r}} \tag{8}$$

CSI-RS resource is selected based on the rate

$$\text{SE}_u^{\text{CSI-RS}_i} = \sum_{n_r=1}^{N_R} \log_2 \left( 1 + \frac{1}{T^{\text{CSI-RS}_i} \, K^{\text{CSI-RS}_i}} \sum_{t,k} s_{u,t,k,n_r}^{\text{CSI-RS}_i} \right) \tag{9}$$

the user $u$ served by cell $c$ select the corresponding CSI-RS resource

$$\hat{\mathbf{i}}_c = \left\{ \arg\max_i \; \text{SE}_u^{\text{CSI-RS},\, i} \right\}_{u \in U_c}. \tag{10}$$

# CSI-RS Transmission and Beam Refinement (6)

- $\hat{\mathbf{i}}_{c,u}$ determines the analog precoder (beam group) used by BS $c$ for UE $u$ during hybrid data transmission.
- Feedback is based on beamformed channels $\mathbf{H}_{b_u,u,t,k}\,\mathbf{B}^{\text{sub}}_{b_u,\hat{i}_{c,u}}$.
- Compared to SSB, CSI-RS sounding is multi-layer due to probing all $B_g$ ports.

# CSI-RS Feedback & Channel Estimation (1)

**Why SINR (from (18)) matters**

- Determines channel-estimation efficacy and guides MCS (modulation/coding) for data.
- Used as the preferred reporting metric for interference management (along with RSRP).

**Beamformed channel estimate**

$$\widehat{\mathbf{H}}_{u,t,k} \triangleq f\left(\mathbf{y}_{u,t,k}^{\text{CSI-RS}_i},\ \mathbf{s}_{c,t,k}^{\text{tr}}\right) \tag{21}$$

- We use $\widehat{\mathbf{H}}_{u,t,k}$ to denote the *beamformed* channel estimate (not the physical channel).
- Serving BS for UE $u$ is $b_u$; selected CSI-RS subset index is $\hat{i}_{b_u,u}$.

# CSI-RS Feedback & Channel Estimation (2)

- The estimate corresponds to the beamformed channel

$$\mathbf{H}_{b_u,u,t,k}\,\mathbf{F}_{b_u,\hat{i}_{b_u,u}}$$

  since the UE has no access to the analog beams.
- Beam management is therefore an **uninformed** operation at the UE (unknown analog beam choices).
- Given the CSI-RS timing/power limits at the UE, an **LMMSE** channel estimator is assumed.

**Frequency selectivity & feedback**

- The channel estimate is frequency-selective across subbands $b \in \{1, \ldots, S_B\}$; for notation simplicity, a single subband is shown.
- In the final step, $\widehat{\mathbf{H}}_{u,t,k}$ is **quantized** using a feedback codebook and reported to the BS.

# Channel Model (1)

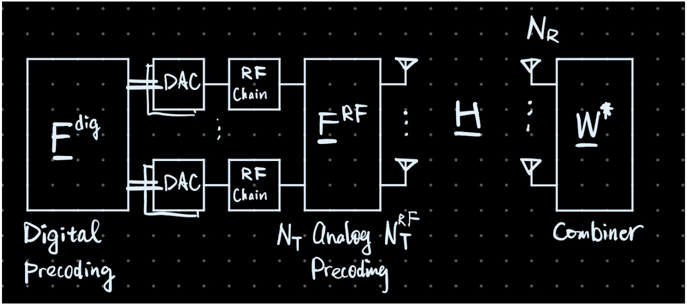


Figure: Hybrid MU-MIMO channel model

for a user $u$ at time-freq resource $t,\ k$

- Digital precoders are block-diagonal:

$$\mathbf{F}_{c,t,k}^{\text{dig}} \in \mathbb{C}^{N_T \times N_R}$$

## Channel Model (2)

- Analog precoders are frequency-flat beamformers:

$$\mathbf{F}_c^{\mathrm{RF}} \in \mathbb{C}^{N_T^{\mathrm{RF}} \times N_T}$$

- Combiner for user $u$: $\mathbf{W}_{u,t,k}^H$
- Received signal decomposition:

$$\begin{aligned}
\mathbf{D}_{u,t,k} &= \mathbf{H}_{c,u,t,k}\mathbf{F}_{c,t}^{\mathrm{RF}}\mathbf{F}_{c,t,k}^{\mathrm{dig}}\mathbf{s}_{c,t,k} && \text{(desired signal)} \quad (11)\\
\mathbf{I}_{u,t,k}^{\mathrm{intra}} &= \sum_{c' \neq c} \mathbf{H}_{c',u,t,k}\mathbf{F}_{c',t}^{\mathrm{RF}}\mathbf{F}_{c',t,k}^{\mathrm{dig}}\mathbf{s}_{c',t,k} && \text{(interference)} \quad (12)\\
\mathbf{y}_{u,t,k} &= \mathbf{W}_{u,t,k}^H\left(\mathbf{D}_{u,t,k} + \mathbf{I}_{c,u,t,k}^{\mathrm{intra}} + \mathbf{N}_{u,t,k}\right) && \text{(received signal)} \quad (13)
\end{aligned}$$

- Noise:

$$\mathbf{N}_{u,t,k} \sim \mathcal{CN}(0, \sigma^2)$$

i.i.d Gaussian, models thermal + hardware noise

# Data Transmission: Hybrid Precoding & ESSE (1)

- After SSB $\rightarrow$ CSI-RS $\rightarrow$ feedback, BS $c$ transmits data using the analog precoders selected from the CSI-RS codebook using the UE feedback $\widehat{\mathbf{H}}_u$.
- Analog beams from the CSI-RS codebook provide array gain
- *regularized zero-forcing* (RZF) digital precoder mitigates intra-cell interference.
- Cells operate independently

**Analog precoder (from selected CSI-RS beams)**

$$\mathbf{F}_c^{\mathrm{RF}} = \left[\, \mathbf{F}_{c,\hat{i}_{c,0}} \;\; \mathbf{F}_{c,\hat{i}_{c,1}} \;\; \cdots \;\; \mathbf{F}_{c,\hat{i}_{c,U_a}} \,\right]. \tag{22}$$

- $U_a$: number of actively scheduled users for BS $c$.
- $\hat{i}_{c,u}$: CSI-RS resource/beam index selected for UE $u$ at BS $c$.

**Per-user RZF digital precoder**

$$\mathbf{F}_{c,u,t,k} = \frac{\left( \sum_{i=0}^{U_a-1} \widehat{\mathbf{H}}^*_{i,t,k} \widehat{\mathbf{H}}_{i,t,k} + U N_T \, \mathbb{E}[N^2_{u,t,k}] \, \mathbf{I} \right)^{-1} \widehat{\mathbf{H}}^*_{u,t,k}}{\left\| \left( \sum_{i=0}^{U_a-1} \widehat{\mathbf{H}}^*_{i,t,k} \widehat{\mathbf{H}}_{i,t,k} + U N_T \, \mathbb{E}[N^2_{u,t,k}] \, \mathbf{I} \right)^{-1} \widehat{\mathbf{H}}^*_{u,t,k} \right\|}. \tag{23}$$

- $\widehat{\mathbf{H}}_{u,t,k} \in \mathbb{C}^{N_R \times r}$: beamformed channel estimate for UE $u$ (from CSI-RS stage).
- $N_T, N_R$: BS TX and UE RX antennas; $r =$ rank (streams sounded/served).

# Data Transmission: Hybrid Precoding & ESSE (3)

**Block-diagonal digital precoder**

$$\mathbf{F}_{c,t,k}^{\text{dig}} = \begin{bmatrix} \mathbf{F}_{c,0,t,k} & 0 & \cdots & 0 \\ 0 & \mathbf{F}_{c,1,t,k} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{F}_{c,U_a,t,k} \end{bmatrix}. \tag{24}$$

**Effective Sum Spectral Efficiency (ESSE)**

- Use the same LMMSE combiner as in (17); compute per-user SINR by (18) with $\mathbf{V}_{i,u,t,k} = \widehat{\mathbf{H}}_{u,t,k}$ and the data-phase precoders.
- Sum-rate over data resources only (remove BM time/frequency sets $T_{\text{BM}}, K_{\text{BM}}$):

$$\text{ESSE} = \sum_{u=1}^{U} \sum_{\substack{t \notin T_{\text{BM}} \\ k \notin K_{\text{BM}}}} \sum_{r=1}^{R} \log_2\big(1 + s_{u,t,k,r}\big). \tag{25}$$

# Data Transmission: Hybrid Precoding & ESSE (4)

- $s_{u,t,k,r}$: post-combiner SINR for UE $u$, stream $r$, at resource $(t, k)$.
- $R$: number of reported/used streams per UE in data phase.

**Notes**

- Beam management (through data transmission) is periodic; codebooks may be retained or updated between periods to match user/site dynamics.
- Learning can generate/update codebooks between periods and supports end-to-end differentiability through the explicit beam-management pipeline.

# Neural Codebook Design and Training Strategy

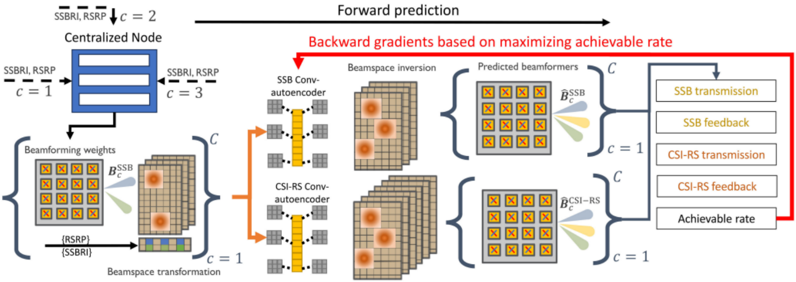


Figure: Neural Codebook Design

# Model Input: Beamspace Representation

We first transform from the array responses $N_1$ dimension $\mathbf{a}_{N_1}(\boldsymbol{\theta})$ to the beamspace sampling dimension $N_2$

$$\boldsymbol{\theta}_{N_2} = \frac{1}{\pi}\left[0, 1, \ldots, N_2 - 1\right]^{\mathsf{T}}, \tag{26}$$

$$\mathbf{U}_{N_1,N_2} \triangleq \left[\, \mathbf{a}_{N_1}(\theta_0),\, \ldots,\, \mathbf{a}_{N_1}(\theta_{N_2-1}) \,\right] \in \mathbb{C}^{N_1 \times N_2}, \tag{27}$$

Beamspace representation allows arbitrary sizing to sampling dimension $N_{x,O} \geq N_X, N_{y,O} \geq N_Y$

$$\widehat{\mathbf{O}}_{c,i}^{\mathrm{BSC}} = \mathbf{U}^{H}_{N_X^{\mathrm{RF}},\, N_{x,O}}\, \mathbf{f}_{c,i}\, \mathbf{U}_{N_Y^{\mathrm{RF}},\, N_{y,O}}, \qquad \forall\, c,\ i,\ \mathbf{f}_{c,i} \in \mathcal{B}_c^{\mathrm{SSB}}. \tag{28}$$

**Beamspace cube as a stack of slices**

$$\mathbf{O}_c^{\mathrm{BSC}} = \left\{\, \widehat{\mathbf{O}}_{c,i}^{\mathrm{BSC}} \,\right\}_{i=0}^{L_{\max}-1}. \tag{29}$$

# Loss function

**Target labels (max achievable SE per UE $u$):**

$$\mathbf{U}_u \, \mathbf{S}_u \, \mathbf{V}_u^H \;=\; \mathbf{H}_{c,u,t,k} \qquad \text{(SVD of the beamformed channel)} \tag{30}$$

$$r_u \;=\; \sum_{n_r=1}^{N_R} \log_2\!\left(1 + \frac{S_{u,n_r}^2}{\sigma^2}\right). \tag{31}$$

- $\mathbf{H}_{c,u,t,k} \in \mathbb{C}^{N_R \times N_T}$: channel from BS $c$ to UE $u$ at $(t,k)$.
- $\mathbf{U}_u \in \mathbb{C}^{N_R \times N_R}$, $\mathbf{V}_u \in \mathbb{C}^{N_T \times N_T}$: unitary; $\mathbf{S}_u = \mathrm{diag}(S_{u,1}, \ldots, S_{u,\min(N_R,N_T)})$ with $S_{u,n_r} \geq 0$.
- $\sigma^2$: noise variance
- $r_u$: the maximum achievable spectral efficiency (bits/s/Hz).
- 

**Training loss:**

$$\mathcal{L}(\mathbf{r}, \hat{\mathbf{r}}) \;=\; \frac{1}{U} \sum_{u=1}^{U} \big(r_u - \hat{r}_u\big)^2. \tag{32}$$

# Takeaways

Reflecting on the proper modeling of the robust design

- **Physical Impairments**
  - Thermal noise during pilot reception (our current channel mismatch model)
  - Inter-cell and intra-cell interference in SSB/CSI-RS
  - Doppler spread and mobility resulting staled CSI
- **System / Standardization Constraints**
  - Quantized and delayed feedback (PMI, CQI, RI)
  - Limited pilot/codebook resources (SSB coarse, CSI-RS finite subset)
  - UE knowledge: only beamformed channel, no direct access to the physical channel
- **Algorithmic / Modeling Limitations**
  - LMMSE estimation suboptimal under UE power/timing limits
  - Rich FR1 vs. sparse FR2 propagation environments