Rate Splitting for MIMO Wireless Networks: A Promising PHY-Layer Strategy for 5G

Bruno Clerckx and **Hamdi Joudeh**

Dept. of Electrical and Electronic Engineering
Imperial College London

Montreal, Canada, VTC-Fall, September 2016

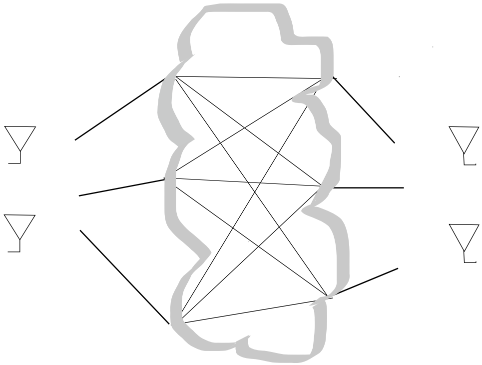
Presentation material downloadable
http://www.ee.ic.ac.uk/bruno.clerckx/Teaching.html

Contents

Introduction to MIMO Networks

Point-to-Point MIMO

- MIMO channel with M transmit and N receive antennas



- MISO ($M > 1, N = 1$), SIMO ($M = 1, N > 1$), SISO ($M = 1, N = 1$)

MIMO Benefits

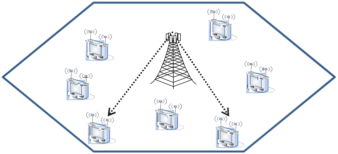
- Array Gain
  - SNR shift in rate or error probability
  - MRC combining (CSIR), MRC/matched beamforming (CSIT)

- Diversity Gain
  - Slope of error curve/outage probability vs SNR (at high SNR)
  - Measure of the reliability of the transmission in slow fading channels
  - MRC combining (CSIR), MRC/matched beamforming (CSIT), STC (no CSIT)

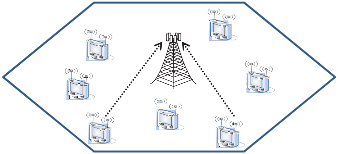
- Multiplexing Gain (also called Degrees of Freedom - DoF)
  - Slope of the achievable rate vs SNR (at high SNR)
  - Number of interference-free streams transmitted in parallel
  - $R \approx g_s \log_2(\rho)$ with $g_s \leq \min\{M, N\}$
  - Spatial Multiplexing/BLAST (no CSIT), Multiple eigenmode transmission with waterfilling power allocation (CSIT)

Multi-User MIMO

- Most systems are multi-user!
- How to deal with K users? Benefit of MIMO in a multi-user setting?
- MIMO Broadcast Channel (BC) and Multiple Access Channel(MAC)



(a) Broadcast Channel - Downlink



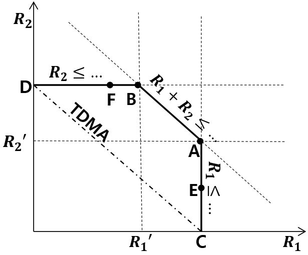
(b) Multiple Access Channel - Uplink

Differences between BC and MAC:

- Multiple independent additive noises in BC **vs** a single noise term in MAC.
- A single Tx power constraint in BC **vs** multiple Tx power constraints in MAC.
- The desired signal and the interference propagate through the same channel in the BC **vs** they propagate through different channels in the MAC.

Uplink Multi-User MIMO (MAC)

- Transmit independent streams and make use of SIC.
  - similar to Spatial Multiplexing for point-to-point MIMO.
- TDMA:
  - orthogonal resource allocation.
  - users never transmitting at the same time.

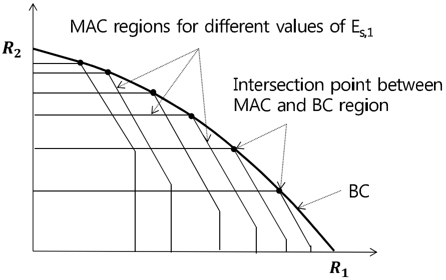


Two-user SISO MAC rate region

- SISO:
  - TDMA and SIC exploit a single DoF.
  - TDMA rate region is smaller than the one achievable with SIC.
- SIMO:
  - TDMA incurs a big loss compared to SIMO MAC (with SIC).
  - 1 DoF vs $\min \{N, K\}$ DoF.
- MIMO:
  - Gap between the TDMA and MIMO MAC rate regions decreases as M increases.

Downlink Multi-User MIMO (BC)

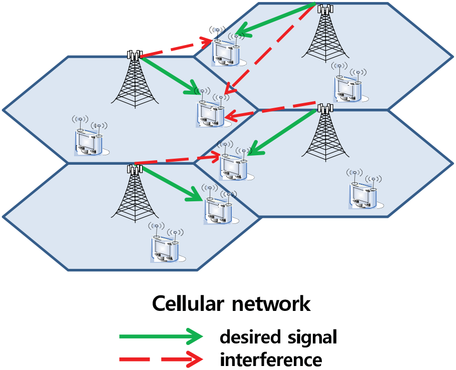
- Transmitter sends independent streams to multiple receivers.
- BC-MAC duality:
  - Reverse roles of transmitter and receivers (reciprocity).
  - Express the capacity region.
  - Use for design and optimization.



- SISO:
  - Degraded channel (users can be ordered according to strength).
  - Superposition coding and SIC achieve capacity region.
  - DPC can be used (transmitter side interference cancellation).
- MISO and MIMO:
  - Non-degraded channel in general (users cannot be ordered).
  - SIC degrades the channel (performance loss).
  - DPC necessary to achieved capacity region (in general).

Multi-Cell MIMO and CoMP

- Current wireless networks operate using a frequency reuse 1.
- Intra-cell (i.e. multi-user interference) + inter-cell (i.e. multi-cell) interference.
- Cell edge performance primarily affected by the inter-cell interference.

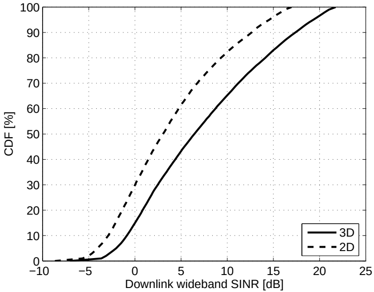


Wideband/long-term SINR

- Commonly evaluated by ignoring the effect of fading but only account for path loss and shadowing

$$SINR_{w,q} = \frac{\Lambda_{q,i}^{-1} P_i}{\sigma_{n,q}^2 + \sum_{j \neq i} \Lambda_{q,j}^{-1} P_j}.$$

- Provides a rough estimate of the network performance.
  - Function of major propagation mechanisms (path loss, shadowing, antenna radiation patterns,...), base stations deployment and user distribution.
- CDF of $SINR_{w,q}$ in a FR $= 1$ network with 2D and 3D antenna patterns in urban macro deployment.



Classical Inter-Cell Interference Mitigation

- Divide-and-conquer approach:
  - fragment the network area into small zones independently controlled from each other.
  - make use of error correction coding, link adaptation, frequency selective scheduling and MIMO in each of those zones.

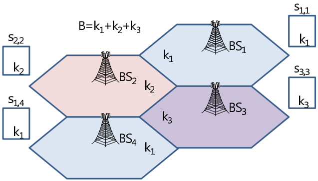


Figure: Frequency Reuse Partitioning.

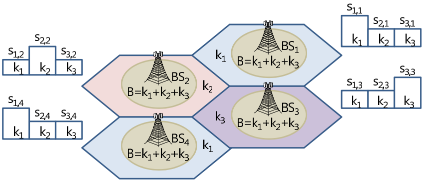
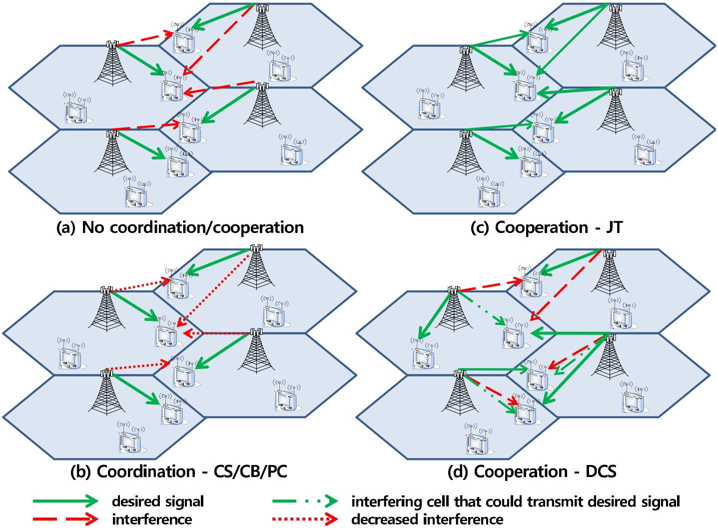


Figure: Static Fractional Frequency Reuse.

Multi-Cell MIMO: Coordination and Cooperation

- Jointly allocate resources across the whole network (and not for each cell independently) and use the antennas of multiple cells to improve the received signal quality at the mobile terminal and to reduce the co-channel interferences.
- Two categories: Coordination and Cooperation
- Coordination
  - No data sharing (user data is available at a single transmsitter)
  - CSI sharing
  - Modelled by an Interference Channel
- Cooperation
  - Data sharing (user data is available at multiple transmsitters)
  - CSI sharing
  - Modelled by BC (for Downlink) and MAC (for Uplink)

Multi-Cell MIMO: Coordination and Cooperation



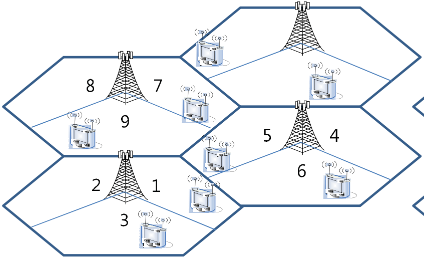
(a) No coordination/cooperation

(c) Cooperation - JT

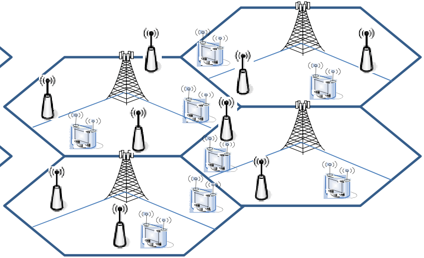
(b) Coordination - CS/CB/PC

(d) Cooperation - DCS

Network Deployments

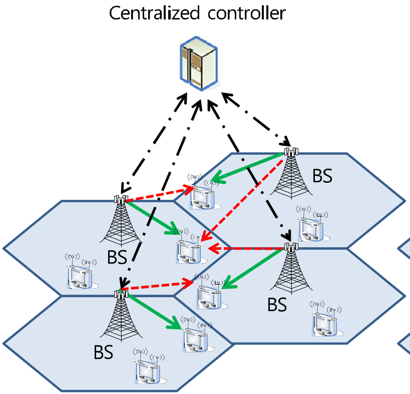


(a) Homogeneous network

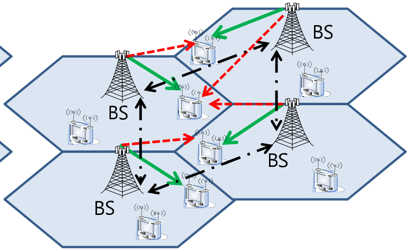


(b) Heterogeneous network

Distributed and Centralized Architecture



(a) Centralized architecture



(b) Distributed architecture

Massive MIMO

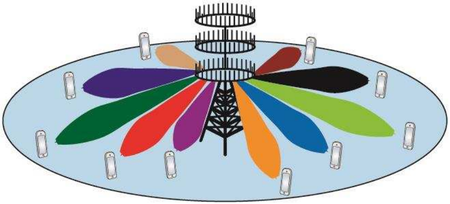


Figure: Downlink beamforming is centralized Massive MIMO deployment [1].

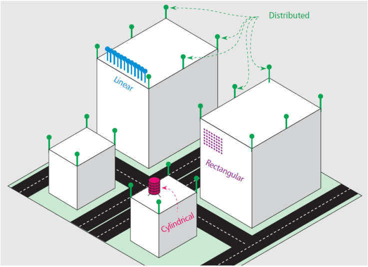
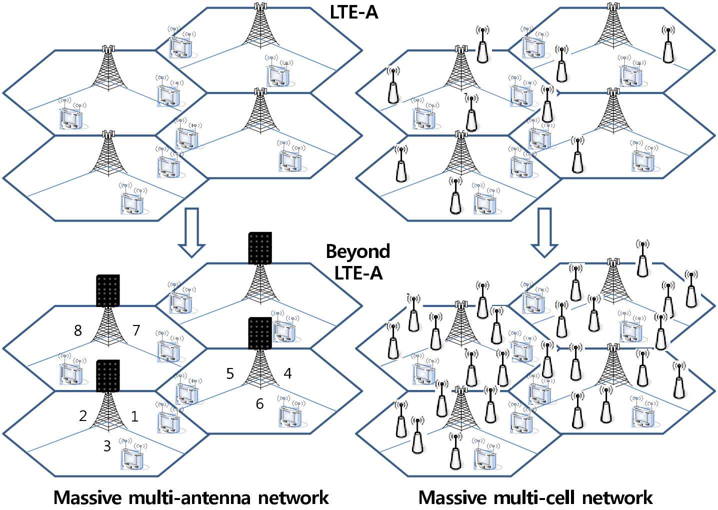


Figure: Various centralized and distributed massive MIMO deployments [2].

- Number of transmitting antennas at the transmitter is (massively) increased.
- Energy can be focused in very narrow beams (reduce multi-user interference).
- Simple precoder design based on matched beamforming (MRC).
- Simultaneously serve many users in the same resource block, simplified scheduling.
- With a massive number of antennas, comes a massive demand for CSIT.

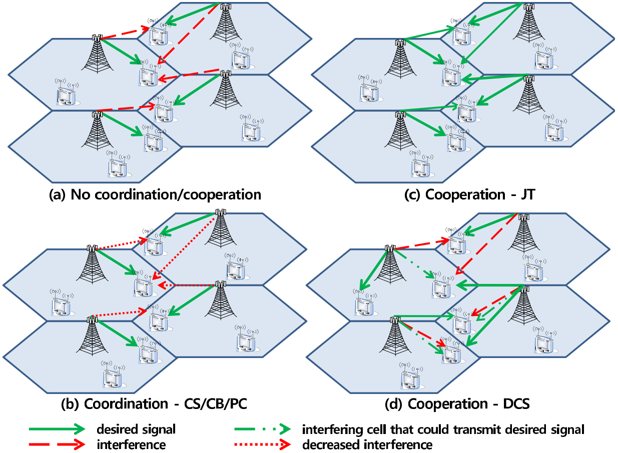
MIMO Networks: Single-user, Multi-user, Multi-cell, Massive, Network, Cooperative, Coordinated,...



Limitations of Current 4G and Emerging 5G Architecture

MIMO Networks: a central problem...the role of CSIT

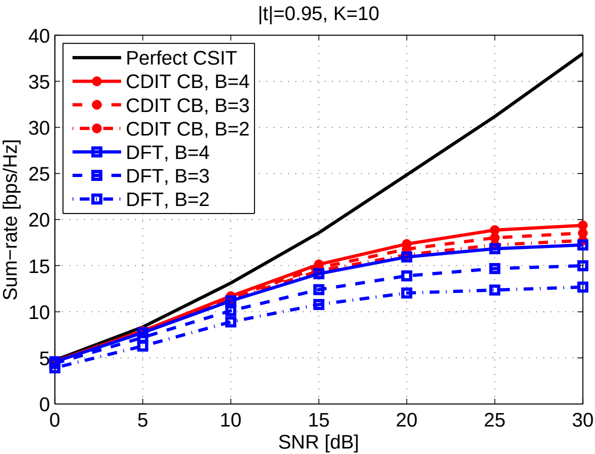
- MIMO Networks exploit more and more channel state information at the transmitter (CSIT)



- Performance crucially rely on accurate CSIT

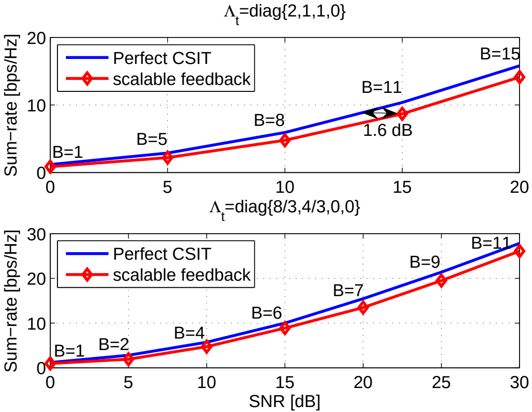
MU-MISO with Quantized Feedback - The Ceiling Effect

- MU-MISO with linear precoding and quantized feedback: the sum-rate saturates due to multi-user interference



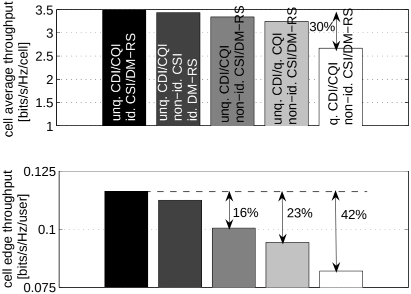
MU-MISO with Quantized Feedback - The Scaling Law

- Number of feedback bits necessary to maintain a rate loss of $\Delta\bar{R} \leq \log_2(\delta)$ bps/Hz per user
  - i.i.d. Rayleigh fading channels: $B \approx (n_t - 1)\log_2(P)$ [3].
  - spatially correlated Rayleigh fading channels $B \approx (r - 1)\log_2(P)$ (r the rank of the transmit correlation matrix)[4].

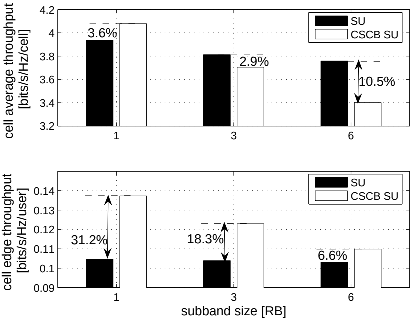


LTE-A SLS Observations

MU-MIMO



Coordinated Scheduling and Beamforming



Observations:

- Big loss due to imperfect CSIT.
- High CSIT accuracy is getting increasingly difficult to satisfy due to increasing number of antennas and access points in 5G (dense HetNet, Massive MIMO).

Motivation for a New Physical Layer

- MU-MISO with linear precoding and quantized feedback: the sum-rate saturates due to multi-user interference.
- Big loss as the CSIT accuracy decreases.
- High CSIT accuracy has become increasingly difficult to satisfy due to increasing number of antennas and access points in 5G (dense HetNet, Massive MIMO).
- So far, techniques designed for perfect CSIT applied to imperfect CSIT scenarios.
- Imperfect CSIT hardly avoidable.
- Wiser to design wireless networks from scratch accounting for imperfect CSIT and its resulting multi-user interference?

Motivation for a New Physical Layer

Information theoretic channel (e.g. MISO BC)

$\Downarrow$

Information theoretic limits (Capacity region)

$\Downarrow$

Communication scheme (e.g. DPC)

$\Downarrow$

Suboptimal scheme (Linear precoding)

$\Downarrow$

Signal processing (Precoder optimization)

$\Downarrow$

Imperfect CSIT (Robust optimization)

For example, robust optimization of $\mathbf{p}_1, \ldots, \mathbf{p}_K$ in

$$\mathbf{x} = \sum_{k=1}^{K} \mathbf{p}_k s_k.$$

BUT !!! The design is motivated by perfect CSIT to start with.

A Bottom-up Approach

Information theoretic channel (e.g. MISO BC with Imperfect CSIT)

$\Downarrow$

Information theoretic limits (Capacity region - unkown)

$\Downarrow$

Alternative information theoretic limits (DoF region)

$\Downarrow$

Communication scheme (Based on Rate-Splitting)

$\Downarrow$

Suboptimal scheme (Linear precoding)

$\Downarrow$

Signal processing (Precoder optimization)

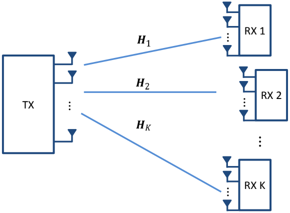
For example, optimizing $\mathbf{p}_c, \mathbf{p}_1, \ldots, \mathbf{p}_K$ in

$$\mathbf{x} = \mathbf{p}_c s_c + \sum_{k=1}^{K} \mathbf{p}_k s_k$$

where $\mathbf{p}_c s_c$ comes from Rate-Splitting.
Motivated by optimality in a DoF sense (multiplexing gain)

The MISO Broadcast Channel and Partial CSIT

System model



$$y_k(t) = \mathbf{h}_k^H(t)\mathbf{x}(t) + n_k(t)$$

- M transmit antennas and K single-antenna users $(M \geq K)$.
- Channel state (matrix): $\mathbf{H}(t) = [\mathbf{h}_1(t), \ldots, \mathbf{h}_K(t)]$.
- In each t, transmitter obtains the estimate $\widehat{\mathbf{H}}(t)$ (i.e. CSIT).
- Joint channel state: $\{\mathbf{H}(t), \widehat{\mathbf{H}}(t)\}$ with joint density $f_{\mathrm{H},\widehat{\mathrm{H}}}(\mathbf{H}, \widehat{\mathbf{H}})$.
- Stationary joint fading process:

$$\underbrace{\{\mathbf{H}(1), \widehat{\mathbf{H}}(1)\}}_{\text{state } 1}, \underbrace{\{\mathbf{H}(2), \widehat{\mathbf{H}}(2)\}}_{\text{state } 2}, \ldots, \underbrace{\{\mathbf{H}(t), \widehat{\mathbf{H}}(t)\}}_{\text{state } t}, \ldots$$

System model: Transmission and Linear precoding

- Transmission over T **random** states.
- Vector inputs: $W_1, \ldots, W_K \mapsto \mathbf{x}(1), \ldots, \mathbf{x}(T)$.
  - $T \gg 1$ (many random states): **Ergodic** transmission.
  - $T = 1$ (one random state): **Non-Ergodic** Transmission.

Linear precoding signal model:

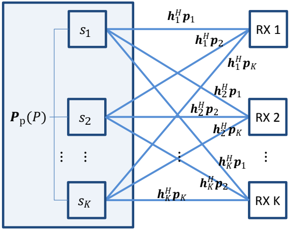
- Independent symbol streams: $W_1, \ldots, W_K \mapsto s_1, \ldots, s_K$.
- t is dropped for simplicity.
- Unity average power: $\mathbb{E}\{s_i s_k^*\} = 1$ if $i = k$, and 0 if $i \neq k$.
- Linear Precoding:

$$\mathbf{x} = \mathbf{p}_1 s_1 + \ldots + \mathbf{p}_K s_K.$$

- Average power constraint: $\sum_{k=1}^{K} \|\mathbf{p}_k\|^2 \leq P$.
- $\mathbf{P}_{\mathrm{p}} = [\mathbf{p}_1, \ldots, \mathbf{p}_K]$ can be adapted based on CSIT

$$\mathbf{P}_{\mathrm{p}}\big(\widehat{\mathbf{H}}(1)\big), \mathbf{P}_{\mathrm{p}}\big(\widehat{\mathbf{H}}(2)\big), \ldots, \mathbf{P}_{\mathrm{p}}\big(\widehat{\mathbf{H}}(T)\big).$$

System model: SINR and Rate

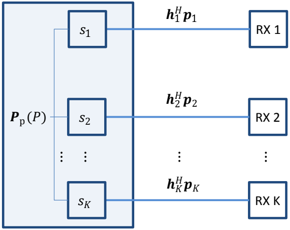


$$y_k = \overbrace{\mathbf{h}_k^H \mathbf{p}_k s_k}^{\text{desired signal}} + \overbrace{\mathbf{h}_k^H \sum_{i \neq k} \mathbf{p}_i s_i}^{\text{interference}} + \overbrace{n_k}^{\text{noise}}$$

- SINR (instantaneous): $\mathrm{SINR}_k = \frac{|\mathbf{h}_k^H \mathbf{p}_k|^2}{\sum_{i \neq k} |\mathbf{h}_k^H \mathbf{p}_i|^2 + \sigma_{\mathrm{n}}^2}$.
- Rate (instantaneous): $R_k = \log_2(1 + \mathrm{SINR}_k)$.
- Ergodic Rate (for $T \gg 1$): $\mathbb{E}\{R_k\}$.

Perfect CSIT

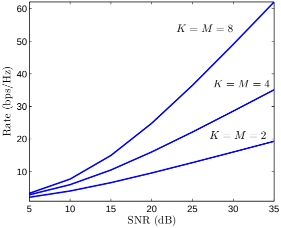
- Perfect CSIT: $\widehat{\mathbf{H}} = \mathbf{H}$.
- Zero-Forcing (ZF) precoding:
  - $\mathbf{P}_{\mathrm{p}} = \mathbf{H}\left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{B}$ where $\mathbf{B}$ is diagonal.
  - This yields: $\mathbf{p}_k \in \mathrm{null}\left([\mathbf{h}_1, \ldots, \mathbf{h}_{k-1}, \mathbf{h}_{k+1}, \ldots, \mathbf{h}_K]^H\right)$.



$$y_k = \mathbf{h}_k^H\mathbf{p}_k s_k + n_k$$

- Each user receives an interference-free stream.
- In other words, each user gets one full DoF.

Perfect CSIT:Degrees of Freedom (DoF)



- DoF: fraction of an interference-free stream's capacity as $P \to \infty$.
- Considering the Ergodic rate:

$$d_k = \lim_{P \to \infty} \frac{\mathbb{E}\{R_k\}}{\log_2(P)}.$$

- For MISO, we have $d_k \leq 1$ due to single-antenna receivers.
- Under perfect CSIT, ZF and equal power allocation achieves full DoF:

$$\sum_{k=1}^{K} d_k = K.$$

Imperfect CSIT

What happens when CSIT is imperfect?

Imperfect CSIT model:

$$\mathbf{H} = \widehat{\mathbf{H}} + \widetilde{\mathbf{H}}$$

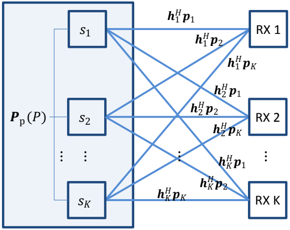
$$\mathbf{h}_k = \underbrace{\widehat{\mathbf{h}}_k}_{\text{estimate}} + \underbrace{\widetilde{\mathbf{h}}_k}_{\text{error}}$$

Estimate obtained through feedback or UL training [5].

- CSIT error power: $\mathbb{E}\left\{\|\widetilde{\mathbf{h}}_k\|^2\right\} = \sigma_{\mathrm{e},k}^2$.
- CSIT error scaling: $\alpha_k = \lim_{P\to\infty} -\frac{\log(\sigma_{\mathrm{e},k}^2)}{\log(P)}$
- It follows that: $\mathbb{E}\left\{\|\widetilde{\mathbf{h}}_k\|^2\right\} \sim P^{-\alpha_k}$.
- Assume: $\alpha_1, \ldots, \alpha_K = \alpha$.
  - $\alpha > 0$: CSIT improves with P (e.g. increasing number of feedback bit).
  - $\alpha = 0$: CSIT fixed with P (e.g. fixed number of feedback bit).
  - $\alpha = 1$: CSIT perfect in a DoF sense (as we see next).

Imperfect CSIT: Zero-Forcing

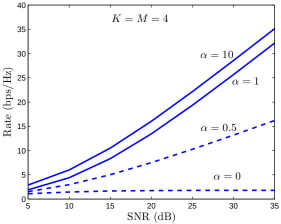
- ZF over the imperfect channel estimate:
  - $\mathbf{P}_{\mathrm{p}} = \widehat{\mathbf{H}}(\widehat{\mathbf{H}}^H\widehat{\mathbf{H}})^{-1}\mathbf{B}$.
  - This yields: $\mathbf{p}_k \in \mathrm{null}\left(\left[\widehat{\mathbf{h}}_1, \ldots, \widehat{\mathbf{h}}_{k-1}, \widehat{\mathbf{h}}_{k+1}, \ldots, \widehat{\mathbf{h}}_K\right]^H\right)$.



$$y_k = \overbrace{\mathbf{h}_k^H \mathbf{p}_k s_k}^{\text{desired signal}} + \overbrace{\widetilde{\mathbf{h}}_k^H \sum_{i \neq k} \mathbf{p}_i s_i}^{\text{residual interference}} + \overbrace{n_k}^{\text{noise}}$$

- Each user cannot enjoy an interference-free stream anymore.
- What happens to the DoF?

Imperfect CSIT: DoF loss



- ZF and equal power allocation: $\|\mathbf{p}_1\|^2 = \ldots = \|\mathbf{p}_K\|^2 = \frac{P}{K}$.

$$y_k = \overbrace{\mathbf{h}_k^H \mathbf{p}_k s_k}^{\text{desired signal } \sim P} + \overbrace{\tilde{\mathbf{h}}_k^H \sum_{i \neq k} \mathbf{p}_i s_i}^{\text{residual interference } \sim P^{1-\alpha}} + \overbrace{n_k}^{\text{noise} \sim P^0}$$

- Assume $\alpha \in [0, 1]$.
- $\mathrm{SINR}_k \sim P^\alpha$ from which $\mathbb{E}\{R_k\} = \log_2(P^\alpha) + O(1)$.
- $d_k = \alpha$ from which the sum DoF [3, 5]:

$$\sum_{k=1}^{K} d_k = K\alpha.$$

Imperfect CSIT: Interference

Perfect CSIT:

- Inter-user interference can be fully eliminated.
- Full DoF is achieved.

Partial CSIT with $\alpha \geq 1$:

- Inter-user interference can be reduced to the level of noise.
- No DoF loss.

Partial CSIT with $\alpha < 1$:

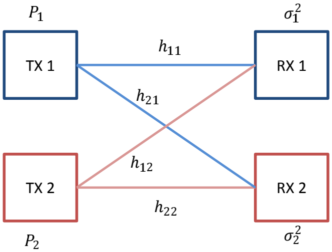
- Inter-user interference cannot be reduced to the level of noise.
- Treating interference as noise causes DoF loss.

If interference cannot be eliminated or reduced to noise level, why not decode it and remove it from the received signal (fully or in part)?

Let us first take a step back, and look at the 2-user Interference Channel (IC).

Fundamentals of Rate Splitting

Two-User Interference Channel (IC)



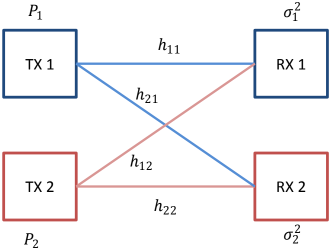
$$y_k = h_{k1}x_1 + h_{k2}x_2 + n_k$$

- Message W_k from TX-k to RX-k.
- Encoding: $W_k \mapsto x_k$.
- Decoding: $y_k \mapsto \widehat{W}_k$.

Symmetric setup:

- $|h_{11}|^2 = |h_{22}|^2 = |h_{\rm d}|^2$ and $|h_{12}|^2 = |h_{21}|^2 = |h_{\rm c}|^2$
- $P_1 = P_2 = P$ and $\sigma_1^2 = \sigma_2^2 = \sigma^2$

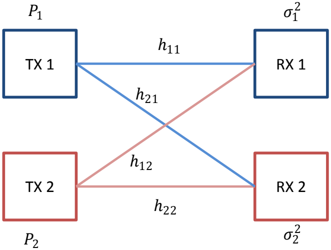
Two-User IC: Very weak interference



Very weak interference $|h_\mathrm{c}|^2 \ll |h_\mathrm{d}|^2$:

- Interference is so weak, it may be **treated as noise**.
- E.g. RX-1 decodes x_1 while treating x_2 as noise.
- $R_k \leq \log_2\left(1 + \frac{P|h_\mathrm{d}|^2}{\sigma^2 + P|h_\mathrm{c}|^2}\right)$.

Two-User IC: Strong interference



Strong interference $|h_\mathrm{c}|^2 > |h_\mathrm{d}|^2$:

- Interfering signal is stronger than desired signal, may as well **decode** it.
- E.g. RX-1 decodes both x_2 and x_1 (MAC).
- $R_k \leq \log_2 \left(1 + \frac{P|h_\mathrm{d}|^2}{\sigma^2}\right)$
- $R_1 + R_2 \leq \log_2 \left(1 + \frac{P|h_\mathrm{d}|^2 + P|h_\mathrm{c}|^2}{\sigma^2}\right)$ (comes from cross-decoding)

Two-User IC: Rate-Splitting

Weak interference $|h_{\mathrm{c}}|^2 < |h_{\mathrm{d}}|^2$ **(or general case)**:

- Not strong enough to **decode**, or weak enough to **treat as noise**.
- **Rate-Splitting**: part **decoded** by other and part **treated as noise**.
  - Split messages: $W_k \mapsto W_{k0}, W_{k1} \mapsto x_{k0}, x_{k1}$.
  - Split power: $P_k \mapsto P_{k0}, P_{k1}$.
  - RX-1 decodes x_{20} and x_1 (composed of x_{10}, x_{11}).
  - RX-2 decodes x_{10} and x_2 (composed of x_{20}, x_{21}).
- Reduces to **treat as noise** when $P_{10} = P_{20} = 0$.
  - i.e. $|W_{10}| = |W_{20}| = 0$.
  - $W_k \mapsto x_{k1}$.
- Reduces to **decode** interference when $P_{11} = P_{21} = 0$.
  - i.e. $|W_{11}| = |W_{21}| = 0$.
  - $W_k \mapsto x_{k0}$.
- **Bridges** the two in general [6].

The MISO-BC with imperfect CSIT revisited

Rate-Splitting for MISO-BC[7]:

- The general idea is to split messages.
- One part decoded by all, while the other treated as noise.

But!

- In what proportion are messages split?
- How much power to allocate?
- How to transmit each part?

Strategy:

- Private messages:
  - Parts which are treated as noise.
  - Received at the level of noise
- Common message(s):
  - Parts which are decoded by all.
  - Transmitted in a public manner.

MISO-BC: Parts to treat as noise (private messages)

Interference reduction through power control:

- Reduce allocated power to P^α.
- Note that $P^\alpha \leq P$ for $\alpha \in [0,1]$.
- Equal power allocation: $\|\mathbf{p}_1\|^2 = \ldots = \|\mathbf{p}_K\|^2 = \frac{P^\alpha}{K}$.

$$y_k = \overbrace{\mathbf{h}_k^H \mathbf{p}_k s_k}^{\text{desired signal } \sim P^\alpha} + \overbrace{\widetilde{\mathbf{h}}_k^H \sum_{i \neq k} \mathbf{p}_i s_i}^{\text{residual interference } \sim P^{\alpha-\alpha} = P^0} + \overbrace{n_k}^{\text{noise} \sim P^0}$$

- Interference is reduced to noise level $\sim P^0$.
- This also limits desired power $\sim P^\alpha$.
- DoF is maintained: $d_k = \alpha$ and $\sum_{k=1}^{K} d_k = K\alpha$.
- Only power levels (scalings) from 0 to α are occupied.
- The remaining power levels (α to 1) are freed for the other parts.

MISO-BC: Parts to decode (common message)

Superpose $W_\mathrm{c} \mapsto s_\mathrm{c}$ (with precoder $\mathbf{p}_\mathrm{c}$) to be decoded by all users.

$$\mathbf{x} = \mathbf{p}_\mathrm{c} s_\mathrm{c} + \sum_{k=1}^{K} \mathbf{p}_k s_k$$

where $\|\mathbf{p}_\mathrm{c}\|^2 = P - P^\alpha \sim P$ and $\|\mathbf{p}_1\|^2 = \ldots = \|\mathbf{p}_K\|^2 = \frac{P^\alpha}{K} \sim P^\alpha$

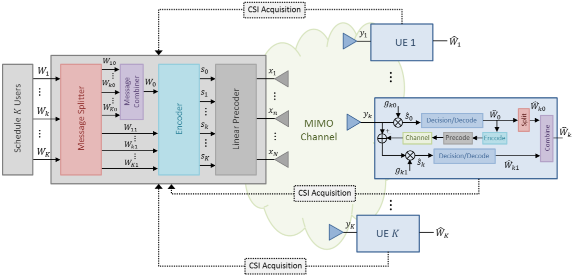
$$y_k = \overbrace{\mathbf{h}_k^H \mathbf{p}_\mathrm{c} s_\mathrm{c}}^{\sim P} + \overbrace{\mathbf{h}_k^H \mathbf{p}_k s_k}^{\sim P^\alpha} + \overbrace{\widetilde{\mathbf{h}}_k^H \sum_{i \neq k} \mathbf{p}_i s_i}^{\sim P^0} + \overbrace{n_k}^{\sim P^0}$$

- $\mathrm{SINR}_{\mathrm{c},k} \sim P^{1-\alpha}$ from which $\mathbb{E}\{R_{\mathrm{c},k}\} = \log_2(P^{1-\alpha}) + O(1)$.
- DoF of common message: $d_\mathrm{c} = 1 - \alpha$.
- SIC is used to remove s_c, as it is decoded by all.
- DoF of private messages is maintained: $d_k = \alpha$.
- Sum DoF is boosted: $d_\mathrm{c} + \sum_{k=1}^{K} d_k = (1-\alpha) + K\alpha$.

What remains is to load both parts (private and common) with user data.

MISO-BC: Rate-Splitting

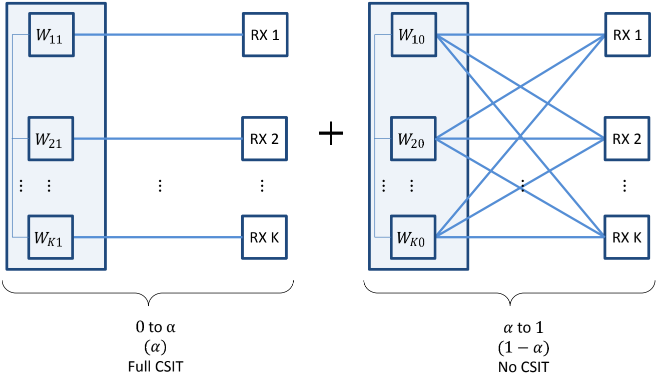
Instead of a new common message, s_c is loaded with part of user messages.



- Split message of user-1 : $W_1 \mapsto W_{10}, W_{11}$.
- Common part: $W_{10} \mapsto s_c$, decoded by all users but intended to users-1.
- Private part: $W_{11} \mapsto s_1$ decoded by user-1.
- $W_2, \ldots, W_K \mapsto s_2, \ldots, s_K$ decoded by corresponding users.

Splitting can be done for other (or all) users as in figure.

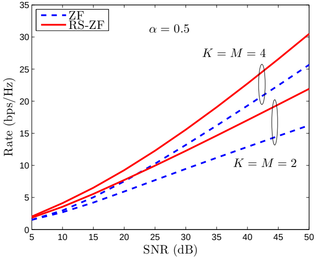
MISO-BC: Weighted sum interpretation



Decomposed into a weighted superposition of two networks

- Perfect CSIT.
  - Achieves sum DoF of K.
  - Weighted by α.
- No CSIT
  - Achieves sum DoF of 1.
  - Weighted by $1-\alpha$.

MISO-BC: DoF with RS



Proposition

In the K user MISO-BC with partial CSIT, sum DoF achieved by ZF is given by

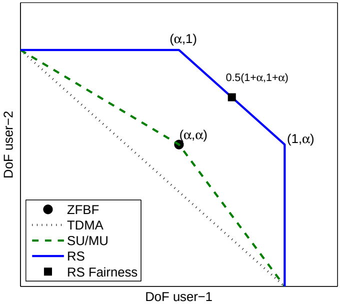
$$d_{\Sigma}^{\mathrm{ZF}} = K\alpha$$

while the sum DoF achieved by RS-ZF is given by

$$d_{\Sigma}^{\mathrm{RS}} = 1 + (K-1)\alpha.$$

Optimality of the RS DoF shown in [8].

MISO-BC: Two-User DoF region



- Assume splitting for user-1
  - user-1 DoF: $d_c + d_1 = (1-\alpha) + \alpha = 1$.
  - user-2 DoF: $d_2 = \alpha$.
- Time-sharing between splitting for user-1 and user-2.
- Compared to time-sharing between ZF and TDMA.

Sum-Rate enhancement and Feedback reduction

From DoF to rate analysis:

- So far we have looked at the DoF gains of RS ($P \to \infty$).
- Here we take a look at the rate performance (more general) [9].

To analyze the performance, we make specific **assumptions**.

- Two-user system.
- Partial CSIT due to limited feedback.
- Random Vector Quantization (RVQ):
  - B bits used to quantize $\mathbf{h}_k$ into $\widehat{\mathbf{h}}_k$.
  - Error $\widetilde{\mathbf{h}}_k$ is due to RVQ.
- Precoding:
  - ZF for private messages.
  - Random precoding for common message.
- Power allocation:
  - Power splitting factor: $t \in [0, 1]$.
  - Private message power: $P_1 = P_2 = \frac{Pt}{2}$.
  - Common message power: $P_c = P(1-t)$.

Sum-Rate enhancement

Proposition

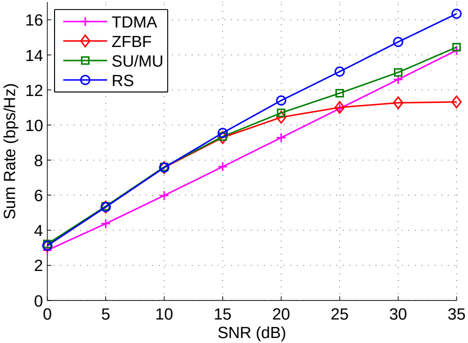
The sum-rate loss incurred by the RS-ZF with RVQ relative to the ZF with perfect CSIT is upper-bounded by

$$\Delta R_S^{eq}(t) \leq 2\epsilon(t) + 2\log_2\left(1+\frac{PtM}{2(M-1)}2^{\frac{-B}{M-1}}\right) - \log_2\left(1+\frac{P(1-t)}{2}e^{\kappa(t)}\right),$$

- Three terms:
  - 1st term $2\epsilon(t)$: rate loss due to the decrement of the power allocated to the private messages,
  - 2nd term (function of B): rate loss incurred by the ZF precoders (with power Pt) of RS with RVQ,
  - 3rd term: rate achieved by the common message.
- Taking $t = 1$:
  - The first and last term become zero.
  - Yields the rate loss incurred by the conventional ZFBF with RVQ [3].

Sum-Rate enhancement: Simulation results

- Sum-rate enhancement (slope gain and/or SNR gain) over ZF, TDMA, switching between TDMA/ZF (SU/MU).
- $M = 4$ antennas, $K = 2$ users, and $B = 15$ bits.



Feedback reduction

Proposition

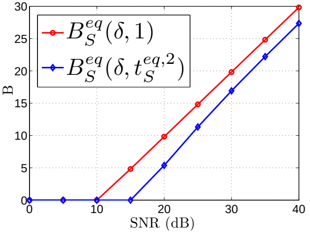
To achieve a sum rate loss $\log_2 \delta$ *bps/Hz relative to ZF with perfect CSIT, the number of feedback bits required by the RS is given by*

$$B = (M-1)\log_2 \frac{PM}{2(M-1)} - (M-1)\log_2 \left(\frac{\sqrt{\delta}\sqrt{1+\frac{P(1-t)}{2}e^{\kappa(t)}}}{t \cdot 2^{\epsilon(t)}} - \frac{1}{t} \right).$$

- Overhead reduction enabled by RS over ZF with RVQ

$$(M-1)\log_2 \frac{\frac{\delta}{2e} + \frac{e}{2} - 1}{\sqrt{\delta} - 1}$$

- Example: $M = 4$ and $\log_2 \delta = 6$bps/Hz



Precoder Optimization

Precoder Optimization

Recall that the RS (linearly precoded) signal model is:

$$\mathbf{x} = \mathbf{p}_c s_c + \sum_{k=1}^{K} \mathbf{p}_k s_k$$

- Precoding matrix: $\mathbf{P} = [\mathbf{p}_c, \mathbf{p}_1, \ldots, \mathbf{p}_K]$.
- Power constraint: $\operatorname{tr}\left(\mathbf{P}\mathbf{P}^H\right) \leq P$.
- So far we considered simple barely optimized designs (ZF, random).
- The choice of $\mathbf{P}$ influences $R_c, R_1, \ldots, R_K$.

Challenges

- Transmitter only known $\widehat{\mathbf{H}}$ and not $\mathbf{H}$.
- Instantaneous $R_c, R_1, \ldots, R_K$ not known by the transmitter.
- Transmission should be carried out at reliable (decodable) rates.

Long-term knowledge should be used to formulate problem and predict reliable rates, e.g. $f_{\mathrm{H},\widehat{\mathrm{H}}}(\mathbf{H}, \widehat{\mathbf{H}})$ and $f_{\mathrm{H}|\widehat{\mathrm{H}}}(\mathbf{H} \mid \widehat{\mathbf{H}})$.

Ergodic Sum-Rate Maximization

Assuming a short-term power constraint (for each $\widehat{\mathbf{H}}$), the ESR problem:

$$\max_{\mathrm{tr}\left(\mathbf{P}(\widehat{\mathbf{H}})\mathbf{P}(\widehat{\mathbf{H}})^H\right)\leq P} \min_i \mathbb{E}\left\{R_{c,i}\right\} + \sum_{k=1}^{K} \mathbb{E}\left\{R_k\right\}$$

Characterization in terms of partial CSIT

- Precoder: $\left\{\mathbf{P}(\widehat{\mathbf{H}})\right\}_{\widehat{\mathbf{H}}}$.
- The conditional Average Rates (ARs):

$$\bar{R}_{c,k} = \mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\left\{R_{c,k} \mid \widehat{\mathbf{H}}\right\} \quad \text{and} \quad \bar{R}_k = \mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\left\{R_k \mid \widehat{\mathbf{H}}\right\}$$

- Instantaneous measures capturing average uncertainty given $\widehat{\mathbf{H}}$.

	$\mathbf{H}(1), \widehat{\mathbf{H}}(1), \mathbf{P}(1)$	$\mathbf{H}(2), \widehat{\mathbf{H}}(2), \mathbf{P}(2)$	$\cdots$	$\mathbf{H}(T), \widehat{\mathbf{H}}(T), \mathbf{P}(T)$
Instantaneous:	$R_k(1), R_{c,k}(1)$	$R_k(2), R_{c,k}(2)$	$\cdots$	$R_k(T), R_{c,k}(T)$
Avg. (conditional):	$\bar{R}_k(1), \bar{R}_{c,k}(1)$	$\bar{R}_k(2), \bar{R}_{c,k}(2)$		$\bar{R}_k(T), \bar{R}_{c,k}(T)$

- ARs change from one state to another depending on $\widehat{\mathbf{H}}$.
- Can be used to optimize the instantaneous precoders.

Ergodic Sum-Rate Maximization: Average Rates

	$\mathbf{H}(1), \widehat{\mathbf{H}}(1), \mathbf{P}(1)$	$\mathbf{H}(2), \widehat{\mathbf{H}}(2), \mathbf{P}(2)$	$\cdots$	$\mathbf{H}(T), \widehat{\mathbf{H}}(T), \mathbf{P}(T)$
Instantaneous:	$R_k(1), R_{c,k}(1)$	$R_k(2), R_{c,k}(2)$	$\cdots$	$R_k(T), R_{c,k}(T)$
Avg. (conditional):	$\bar{R}_k(1), \bar{R}_{c,k}(1)$	$\bar{R}_k(2), \bar{R}_{c,k}(2)$		$\bar{R}_k(T), \bar{R}_{c,k}(T)$

- For given $\widehat{\mathbf{H}}$, ARs may not be achievable (e.g. higher than actual rates).
- However, ergodic rates are achievable (law of total expectation):

$$\mathbb{E}_{\mathrm{H},\widehat{\mathrm{H}}}\{R_{c,k}\} = \mathbb{E}_{\widehat{\mathrm{H}}}\left\{\bar{R}_{c,k}\right\} \quad \text{and} \quad \mathbb{E}_{\mathrm{H},\widehat{\mathrm{H}}}\{R_k\} = \mathbb{E}_{\widehat{\mathrm{H}}}\left\{\bar{R}_k\right\}$$

- Update precoder in each state, but code over a long sequence of states.

Average Sum Rate (ASR):

- The ESR can be lower-bounded by

$$\mathrm{ESR} = \min_i \mathbb{E}_{\widehat{\mathrm{H}}}\left\{\bar{R}_{c,i}\right\} + \sum_{k=1}^{K} \mathbb{E}_{\widehat{\mathrm{H}}}\left\{\bar{R}_k\right\} \geq \mathbb{E}_{\widehat{\mathrm{H}}} \overbrace{\left\{\min_i \bar{R}_{c,i} + \sum_{k=1}^{K} \bar{R}_k\right\}}^{\mathrm{ASR}}$$

- Inequality follows from moving expectation outside minimization.
- ASR can be maximized for each channel state.

Ergodic Sum-Rate Maximization: ASR problem

ASR problem:

$$\mathcal{R}_{\mathrm{RS}}(P):\begin{cases}\max\limits_{\bar{R}_c,\mathbf{P}} & \bar{R}_{\mathrm{c}}+\sum_{k=1}^{K}\bar{R}_k\\ \text{s.t.} & \bar{R}_{\mathrm{c},k}\geq\bar{R}_{\mathrm{c}},\ \forall k\in\mathcal{K}\\ & \mathrm{tr}\left(\mathbf{P}\mathbf{P}^H\right)\leq P\end{cases}$$

as opposed to the conventional (NoRS) formulation

$$\mathcal{R}(P):\begin{cases}\max\limits_{\mathbf{P}_{\mathrm{p}}} & \sum_{k=1}^{K}\bar{R}_k\\ \text{s.t.} & \mathrm{tr}\left(\mathbf{P}_{\mathrm{p}}\mathbf{P}_{\mathrm{p}}^H\right)\leq P.\end{cases}$$

- Stochastic optimization problem (due to expectations inside the ARs).
- Even a deterministic version is non-convex and very difficult.
- WMMSE approach can efficiently handle sum rate problems.

Sample Average Approximation (SAA) $\Rightarrow$ deterministic $\Rightarrow$ WMMSE

Ergodic Sum-Rate Maximization: SAA

- Monte-Carlo sample:

$$\mathbb{H}^{(M)} = \left\{ \mathbf{H}^{(m)} = \widehat{\mathbf{H}} + \widetilde{\mathbf{H}}^{(m)} \mid \widehat{\mathbf{H}}, m = 1, \ldots, M \right\}.$$

- Approximated ARs:

$$\bar{R}_{\mathrm{c},k}^{(M)} \triangleq \frac{1}{M} \sum_{m=1}^{M} R_{\mathrm{c},k}\left(\mathbf{h}_k^{(m)}\right) \quad \text{and} \quad \bar{R}_{k}^{(M)} \triangleq \frac{1}{M} \sum_{m=1}^{M} R_{k}\left(\mathbf{h}_k^{(m)}\right).$$

- Sampled ASR problem:

$$\mathcal{R}_{\mathrm{RS}}^{(M)}(P) : \begin{cases} \max\limits_{\bar{R}_{\mathrm{c}}, \mathbf{P}} & \bar{R}_{\mathrm{c}} + \sum_{k=1}^{K} \bar{R}_k^{(M)} \\ \text{s.t.} & \bar{R}_{\mathrm{c},k}^{(M)} \geq \bar{R}_{\mathrm{c}}, \ \forall k \in \mathcal{K} \\ & \mathrm{tr}\left(\mathbf{P}\mathbf{P}^H\right) \leq P. \end{cases}$$

  - Problem $\mathcal{R}_{\mathrm{RS}}^{(M)}(P)$ is deterministic.
  - From LLN, we have $\mathcal{R}_{\mathrm{RS}}^{(M)}(P) \to \mathcal{R}_{\mathrm{RS}}(P)$ as $M \to \infty$.
  - $\mathbb{H}^{(M)}$ can be generated from $f_{\mathrm{H}|\widehat{\mathrm{H}}}(\mathbf{H} \mid \widehat{\mathbf{H}})$.

Ergodic Sum-Rate Maximization: WMMSE approach

$$R = \log_2\left(1+\mathrm{SINR}\right) = \max_{g,u} \overbrace{\log_2(u) - u\mathrm{MSE}(g)}^{-\xi(g,u)} + 1$$

- g is an equalizer and $u > 0$ is a weight.
- At optimality we have:
  - $g = g^{\mathrm{MMSE}}$.
  - $\mathrm{MSE}(g^{\mathrm{MMSE}}) = \mathrm{MMSE}$.
  - $u = 1/\mathrm{MMSE}$.
  - $R = \log_2\left(1+\mathrm{SINR}\right) = -\log_2(\mathrm{MMSE})$.
- In many scenarios, $R(\mathbf{P})$ is non-concave in $\mathbf{P}$.
- Hence, something like $R_1(\mathbf{P}) + \ldots + R_K(\mathbf{P})$ is also non-concave.
- $-\xi(g,u,\mathbf{P})$ is concave in any variable while fixing the other two.
- This can be exploited in an Alternating Optimization (AO) manner.

Ergodic Sum-Rate Maximization: SAA-WMMSE algorithm

The sampled WMMSE problem:

$$\mathcal{A}_{\mathrm{RS}}^{(M)}(P): \begin{cases} \min\limits_{\bar{\xi}_{\mathrm{c}}, \mathbf{P}, \mathbf{U}, \mathbf{G}} & \bar{\xi}_{\mathrm{c}} + \sum_{k=1}^{K} \bar{\xi}_k^{(M)} \\ \text{s.t.} & \bar{\xi}_{\mathrm{c},k}^{(M)} \leq \bar{\xi}_{\mathrm{c}}, \ \forall k \in \mathcal{K} \\ & \mathrm{tr}\left(\mathbf{P}\mathbf{P}^H\right) \leq P \end{cases}$$

where $\mathbf{U}$ and $\mathbf{G}$ are sets of weights and equalizers.

Algorithm:

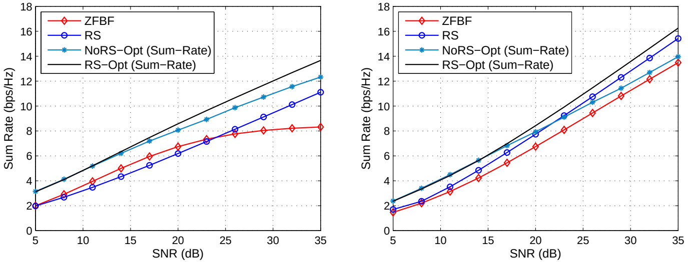
- Initialize: $n \leftarrow 0$, $\mathbf{P}^{[n]}$
- Repeat
  1. $n \leftarrow n + 1$
  2. $\mathbf{G} \leftarrow \mathbf{G}^{\mathrm{MMSE}}(\mathbf{P}^{[n-1]})$
  3. $\mathbf{U} \leftarrow \mathbf{U}^{\mathrm{MMSE}}(\mathbf{P}^{[n-1]})$
  4. $\mathbf{P}^{[n]} \leftarrow \arg \mathcal{A}_{\mathrm{RS}}^{(M)}(P, \mathbf{G}, \mathbf{U})$
- Until stopping criteria met

Steps: $\mathbf{G}$ and $\mathbf{U}$ are updated in closed-form. Updating $\mathbf{P}^{[n]}$ is a convex QCQP.

Convergence: the algorithm converges to a KKT point [14].

Ergodic Sum-Rate Maximization: Simulation results

- Optimized and non-optimized precoders with and without RS.
- $M = 2$ antennas and $K = 2$ users.

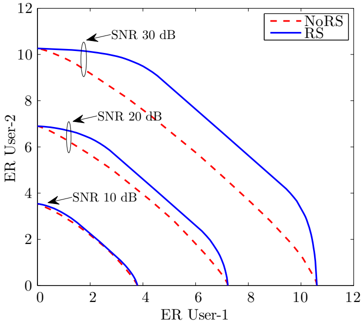


(a) $\sigma_e^2 = 0.063$ $(\alpha = 0)$ (b) $\sigma_e^2 = P^{-0.6}$ $(\alpha = 0.6)$

- RS performs better than NoRS.
- Optimized precoders give superior performances across all SNRs.

Ergodic Sum-Rate Maximization: Two-user ER region

- More generally, we can solve the Weighted ESR problem.
- Each user is assigned a weight (depending on importance).
- Common message is shared among users (as we see next for fairness).
- By varying the weights, we can obtain the 2-user ER region ($\alpha = 0.6$):



- Shows the ER trade-offs between the two users.

Robust Max-Min Fairness

Non-Ergodic transmission over $T = 1$ random state $\{\mathbf{H}, \widehat{\mathbf{H}}\}$.

- Assume bounded support of $f_{\mathrm{H}|\widehat{\mathrm{H}}}(\mathbf{H} \mid \widehat{\mathbf{H}}) \Rightarrow$ bounded CSIT errors.
- For kth user, CSIT errors bounded by sphere with radius δ_k:

$$\mathbb{H}_k = \left\{ \mathbf{h}_k \mid \mathbf{h}_k = \widehat{\mathbf{h}}_k + \widetilde{\mathbf{h}}_k, \|\widetilde{\mathbf{h}}_k\| \leq \delta_k \right\}$$

- For any $\mathbf{P}$, worst-case rates defined as:

$$\bar{R}_{\mathrm{c},k} = \min_{\mathbf{h}_k \in \mathbb{H}_k} R_{\mathrm{c},k}(\mathbf{h}_k) \quad \text{and} \quad \bar{R}_k = \min_{\mathbf{h}_k \in \mathbb{H}_k} R_k(\mathbf{h}_k).$$

- For given $\widehat{\mathbf{H}}$, transmission at worst-case rates is reliable (robust).

Rate-Splitting revisited: Sharing the common message

- $W_k \mapsto W_{k0}, W_{k1}$ for all $k \in \{1, \ldots, K\}$.
- $W_{10}, \ldots, W_{K0} \mapsto s_{\mathrm{c}}$.
- $W_{11}, \ldots, W_{K1} \mapsto s_1, \ldots, s_K$.

Robust Max-Min Fairness

$$\mathcal{R}_{\mathrm{RS}}(P):\begin{cases}\max\limits_{\bar{\mathbf{c}},\mathbf{P}} & \min\limits_{k\in\mathcal{K}}\ (\bar{R}_k+\bar{C}_k)\\ \text{s.t.} & \bar{R}_{\mathrm{c},k}\geq\sum_{i=1}^{K}\bar{C}_i,\ \forall k\in\mathcal{K}\\ & \bar{C}_k\geq 0,\ \forall k\in\mathcal{K}\\ & \mathrm{tr}\left(\mathbf{P}\mathbf{P}^H\right)\leq P.\end{cases}$$

where $\bar{\mathbf{c}}=[\bar{C}_1,\ldots,\bar{C}_M]$.

- Portion of the common message rate given to user k: $\bar{C}_k$.
- Sum of all portions: $\sum_{k=1}^{K}\bar{C}_k=\bar{R}_{\mathrm{c}}=\min_i\bar{R}_{\mathrm{c},i}$.
- Rate of user k: $\bar{R}_k+\bar{C}_k$ (private and common portions).

Classical (NoRS) problem formulated as:

$$\mathcal{R}(P):\begin{cases}\max\limits_{\mathbf{P}_{\mathrm{p}}} & \min\limits_{k\in\mathcal{K}}\ \bar{R}_k\\ \text{s.t.} & \mathrm{tr}\left(\mathbf{P}_{\mathrm{p}}\mathbf{P}_{\mathrm{p}}^H\right)\leq P.\end{cases}$$

Semi-infinite programs: finite variables and infinite constraints.

Robust Max-Min Fairness: Cutting-Set method

The ith sampled problem:

$$\mathcal{R}_{\mathrm{RS}}^{(i)}(P):\begin{cases}\max\limits_{\bar{R}_{\mathrm{t}},\bar{\mathbf{c}},\mathbf{P}} \bar{R}_{\mathrm{t}} \\ \text{s.t. } R_k\left(\mathbf{h}_k\right)+\bar{C}_k \geq \bar{R}_{\mathrm{t}}, \forall \mathbf{h}_k \in \mathbb{H}_k^{(i)}, k \in \mathcal{K} \\ R_{\mathrm{c},k}\left(\mathbf{h}_{\mathrm{c},k}\right) \geq \sum_{l=1}^{K} \bar{C}_l, \forall \mathbf{h}_{\mathrm{c},k} \in \mathbb{H}_{\mathrm{c},k}^{(i)}, k \in \mathcal{K} \\ \bar{C}_k \geq 0,\ \forall k \in \mathcal{K} \\ \mathrm{tr}\left(\mathbf{P}\mathbf{P}^H\right) \leq P.\end{cases}$$

- The ith sampled uncertainty regions: $\mathbb{H}_k^{(i)}, \mathbb{H}_{\mathrm{c},k}^{(i)} \subset \mathbb{H}_k$.
- Sampled for R_k and $R_{\mathrm{c},k}$ as messaged decoded independently.
- Large sampled regions would yield many constraints (need clever sampling).

Cutting-Set algorithm:

- For $i = 1, 2, \ldots,$ do:
  1. Optimization: solve $\mathcal{R}_{\mathrm{RS}}^{(i)}(P)$ (WMMSE approach).
  2. Pessimization: update $\mathbb{H}_k^{(i)}, \mathbb{H}_{\mathrm{c},k}^{(i)}$ for all $k \in \{1, \ldots, K\}$ with
     - $\mathbf{h}_{\mathrm{c},k}^{*} = \arg\min_{\mathbf{h}_{\mathrm{c},k} \in \mathbb{H}_k} R_{\mathrm{c},k}^{(i)}(\mathbf{h}_{\mathrm{c},k})$ and $\mathbf{h}_k^{*} = \arg\min_{\mathbf{h}_k \in \mathbb{H}_k} R_k^{(i)}(\mathbf{h}_k)$
- Until stopping criteria met

Convergence: to a KKT point [15].

Robust Max-Min Fairness: Simulation results

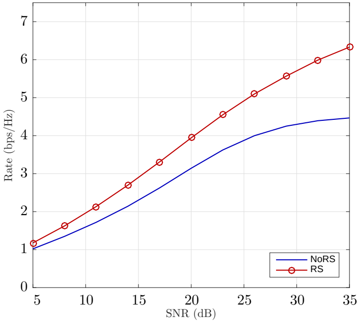
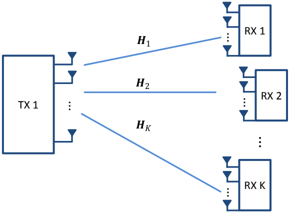


Figure: $K = M = 3$ and $\delta_1, \delta_2, \delta_3 = 0.1$.

- NoRS saturates due to non-scaling CSIT errors.
- RS avoids saturation and performs better across all SNRs.

Multiple receive antennas



$$\mathbf{y}_k(t) = \mathbf{H}_k^H(t)\mathbf{x}(t) + \mathbf{n}_k(t)$$

- MIMO-BC with M transmit antennas and $N_1, \ldots, N_K$ receive antennas.
- Symmetry for simplicity: $N_1, \ldots, N_K = N$ and $M \geq KN$.
- Linear precoding with symbols $\mathbf{s}_k = \left[s_k^{(1)}, \ldots, s_k^{(N)}\right]^T$:

$$\mathbf{x} = \sum_{k=1}^{K} \mathbf{P}_k \mathbf{s}_k$$

- DoF under perfect CSIT: $\sum_{k=1}^{K} d_k = NK$ (scaling by N)[10].

Multiple receive antennas

Common streams can be scaled up by N as well:

$$\mathbf{x} = \mathbf{P}_{\mathrm{c}}\mathbf{s}_{\mathrm{c}} + \sum_{k=1}^{K} \mathbf{P}_k \mathbf{s}_k$$

where $\mathbf{s}_{\mathrm{c}} = \left[s_{\mathrm{c}}^{(1)}, \ldots, s_{\mathrm{c}}^{(N)}\right]^T$ and $\mathbf{s}_k = \left[s_k^{(1)}, \ldots, s_k^{(N)}\right]^T$.

$$\mathbf{y}_k = \overbrace{\mathbf{H}_k^H \mathbf{P}_{\mathrm{c}}\mathbf{s}_{\mathrm{c}}}^{\sim P} + \overbrace{\mathbf{H}_k^H \mathbf{P}_k \mathbf{s}_k}^{\sim P^{\alpha}} + \overbrace{\widetilde{\mathbf{H}}_k^H \sum_{i \neq k} \mathbf{P}_i \mathbf{s}_i}^{\sim P^0} + \overbrace{n_k}^{\sim P^0}$$

- Each user has N received signal dimensions.
- Common streams achieve $d_{\mathrm{c}} = N(1-\alpha)$ by treating the rest as noise.
- After SIC, private streams achieved $d_k = N\alpha$ at each receiver.
- Sum DoF scaled by N:

$$d_{\mathrm{c}} + \sum_{k=1}^{K} d_k = N + N(K-1)\alpha$$

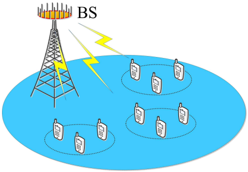
- Rate-Splitting can be used to fill common streams with private information.
- Things get more complicated in asymmetric scenarios [11, 12].

Massive MISO

Massive MIMO challenge: the huge demand for accurate CSIT.

The use of Rate-Splitting:

- The constraint: $R_c = \min_k \{R_{c,k}\}$.
- This highly reduces the gain when K is large.
- Channel statistics $\mathbf{R}_k$ can be further exploited.
- Large training and feedback overhead.



User grouping based on spatial correlation:

- Can be utilized with two-tier precoding to reduce the signalling overhead while maintaining the system performance [17, 18, 19]

$$\mathbf{x} = \sqrt{\frac{P}{K}} \sum_{g=1}^{G} \mathbf{B}_g \mathbf{W}_g \, \mathbf{s}_g,$$

- Users in g-th group share the same channel statistics: $\mathbf{R}_g$.
- $\mathbf{B}_g$: outer-precoding matrix based on channel statistics.
- $\mathbf{W}_g$: inner-precoding matrix designed based on short-term effective channel estimates: $\widehat{\overline{\mathbf{H}}}_g = \mathbf{B}_g^H \widehat{\mathbf{H}}_g$.

Massive MISO: Hierarchical Rate-Splitting (HRS)

- The overlap between the eigen-subspaces spanned by the groups' correlation matrices may give rise to inter-group interference while the intra-group interference cannot be eliminated due to the imperfect CSIT.

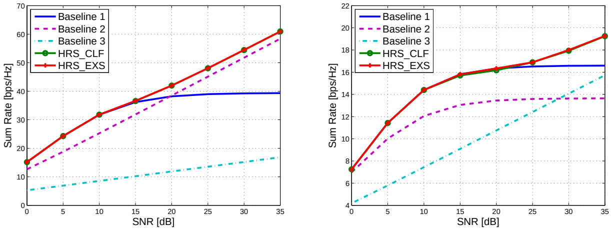
- **Hierarchical Rate-Splitting**: a hierarchy of common messages to combat the inter-group and intra-group interference in Massive MIMO

$$\mathbf{x} = \overbrace{\sqrt{P_{\mathrm{sc}}}\mathbf{w}_{\mathrm{sc}}\, s_{\mathrm{sc}}}^{\text{system common msg.}} + \sum_{g=1}^{G} \mathbf{B}_g \left(\overbrace{\sqrt{P_{cg}}\mathbf{w}_{cg}\, s_{cg}}^{\text{group common msg.}} + \overbrace{\sqrt{P_{gk}}\mathbf{W}_g\, \mathbf{s}_g}^{\text{private msgs.}} \right)$$

- System common msg. decoded by all users: for inter-group interference.
- Group common msg. decoded by group: for intra-group interference

Massive MISO: Simulation results

- HRS vs. various baselines under imperfect CSIT, $M = 100$, $K = 12, \tau^2 = 0.4$

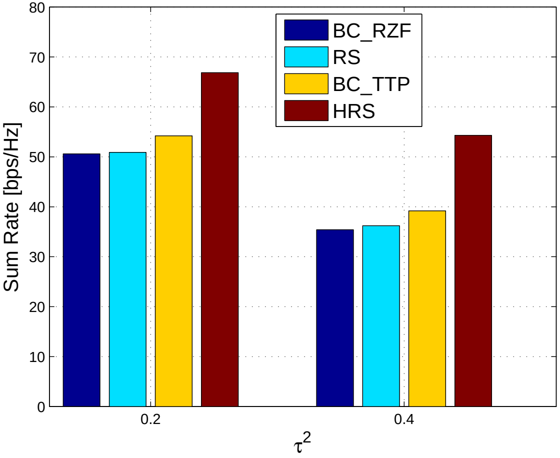


(a) disjoint eigen-subspace (b) overlapping eigen-subspace

- HRS behaves as two-tier BC at low to medium SNR and achieves a sum rate that increases with available transmit power.
- RS highly decreases the complexity of precoder design and scheduling at the expense of an increase in complexity of the encoders and decoders.

Massive MISO: Simulation results

- $M = 100,\ K = 12, \tau^2 = 0.4,\ SNR = 30dB,$ disjoint eigen-subspaces



Multi-Cell Coordination

MISO Interference Channel (IC):

- K pairs of TX-RX.
- Transmitting antennas: $M_k \geq K$ for all $k \in \{1, \ldots, K\}$.

$$y_k = \sum_{j=1}^{K} \mathbf{h}_{kj}^H \mathbf{x}_j + n_k.$$

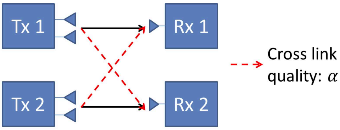
Perfect CSIT:

- ZF can be applied: $\mathbf{p}_j \in \text{null}\left([\mathbf{h}_{1j}, \ldots, \mathbf{h}_{j-1j}, \mathbf{h}_{j+1j}, \ldots, \mathbf{h}_{Kj}]^H\right)$.
- DoF: $\sum_{k=1}^{K} d_k = K$.

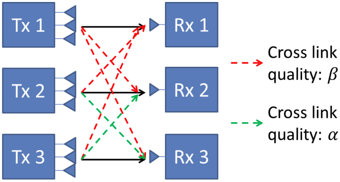
Imperfect CSIT:

- ZF applied over estimate: $\mathbf{p}_j \in \text{null}\left([\widehat{\mathbf{h}}_{1j}, \ldots, \widehat{\mathbf{h}}_{j-1j}, \widehat{\mathbf{h}}_{j+1j}, \ldots, \widehat{\mathbf{h}}_{Kj}]^H\right)$.
- DoF loss, e.g. $\sum_{k=1}^{K} d_k = K\alpha$.
- RS can be applied as in BC yielding: $d_c + \sum_{k=1}^{K} d_k = 1 + (K-1)\alpha$.
- TXs share common message (e.g. TDMA).
- Assume TX-1 transmits common message.

Multi-Cell Coordination: Rate-Splitting



(c) two-cell scenario [20]



(d) three-cell scenario [21]

Rate-Splitting:

$$y_k = \overbrace{\mathbf{h}_{k1}^H \mathbf{p}_c s_c}^{\sim P} + \overbrace{\mathbf{h}_{kk}^H \mathbf{p}_k s_k}^{\sim P^{a_k}} + \overbrace{\sum_{i \neq k} \widetilde{\mathbf{h}}_{ki}^H \mathbf{p}_i s_i}^{\sim P^{\max\{a_i - \alpha_{ki}\}_{i \neq k}}} + \overbrace{n_k}^{\sim P^0}$$

Two-cell:

- $a_1, a_2 = \alpha$.
- $d_1, d_2 = \alpha$ and $d_c = 1 - \alpha$.
- **sum**: $1 + \alpha$.

Three-cell ($\beta \geq \alpha$):

- $a_1 = \beta$ and $a_2, a_3 = \alpha$.
- $d_1 = \beta$, $d_2, d_3 = \alpha$ and $d_c = 1 - \beta$.
- **sum**: $1 + 2\alpha$.

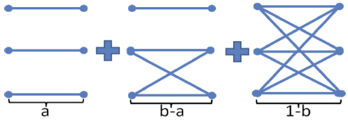
Higher CSIT (β) not utilized in Three-cell.

Multi-Cell Coordination: Topological Rate-Splitting (TRS)

Can we do better for the three-cell example?[21]

	Tx1	Tx2	Tx3
Rx1		b	b
Rx2	b		a
Rx3	b	a	

(e) CSIT pattern



(f) Weighted-sum interpretation

Transmitted signals:

$$\mathbf{x}_1 = \underbrace{\overbrace{\mathbf{p}_1 s_1}^{\sim P^{\alpha}} + \overbrace{\mathbf{p}_{\mathrm{g}1} s_{\mathrm{g}1}}^{\sim P^{\beta}}}_{\bar{\mathbf{x}}_1} + \overbrace{\mathbf{p}_{\mathrm{c}} s_{\mathrm{c}}}^{\sim P^{1}}$$

$$\mathbf{x}_2 = \overbrace{\mathbf{p}_2 s_2}^{\sim P^{\alpha}} + \overbrace{\mathbf{p}_{\mathrm{g}2} s_{\mathrm{g}2}}^{\sim P^{\beta}}$$

$$\mathbf{x}_3 = \overbrace{\mathbf{p}_3 s_3}^{\sim P^{\alpha}}.$$

Received signals:

$$y_1 = \overbrace{\mathbf{h}_{11}^H \mathbf{p}_1 s_1}^{\sim P^{\alpha}} + \overbrace{\mathbf{h}_{11}^H \mathbf{p}_{\mathrm{g}1} s_{\mathrm{g}1}}^{\sim P^{\beta}} + \overbrace{\mathbf{h}_{11}^H \mathbf{p}_{\mathrm{c}} s_{\mathrm{c}}}^{\sim P^{1}} + \overbrace{\mathbf{h}_{12}^H \mathbf{x}_2 + \mathbf{h}_{13}^H \mathbf{x}_3 + n_1}^{P^0}$$

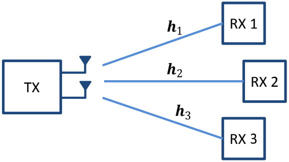
$$y_2 = \overbrace{\mathbf{h}_{22}^H \mathbf{p}_2 s_2}^{\sim P^{\alpha}} + \overbrace{\mathbf{h}_{22}^H \mathbf{p}_{\mathrm{g}2} s_{\mathrm{g}2}}^{\sim P^{\beta}} + \overbrace{\mathbf{h}_{21}^H \mathbf{p}_{\mathrm{c}} s_{\mathrm{c}}}^{\sim P^{1}} + \overbrace{\mathbf{h}_{21}^H \bar{\mathbf{x}}_1 + \mathbf{h}_{23}^H \mathbf{x}_3 + n_2}^{P^0}$$

$$y_3 = \overbrace{\mathbf{h}_{33}^H \mathbf{p}_3 s_3}^{\sim P^{\alpha}} + \overbrace{\mathbf{h}_{32}^H \mathbf{p}_{\mathrm{g}2} s_{\mathrm{g}2}}^{\sim P^{\beta}} + \overbrace{\mathbf{h}_{31}^H \mathbf{p}_{\mathrm{c}} s_{\mathrm{c}}}^{\sim P^{1}} + \overbrace{\mathbf{h}_{31}^H \bar{\mathbf{x}}_1 + \mathbf{h}_{32}^H \mathbf{p}_2 s_2 + n_3}^{P^0}.$$

$d_1, d_2, d_3 = \alpha$, $d_{\mathrm{g}1}, d_{\mathrm{g}2} = \beta - \alpha$ and $d_{\mathrm{c}} = 1 - \beta$.

sum: $1 + \beta + \alpha$ compared to $1 + 2\alpha$.

Overloaded systems



- Overloaded scenarios: $K > M$.
- Scheduling over orthogonal resource blocks (time/frequency).
- Serve at most M users at a time.
- Reduces to conventional MISO BC in each block.
- With perfect CSIT, achieves DoF M in each block.

Consider a scenario where some user have little or no CSIT:

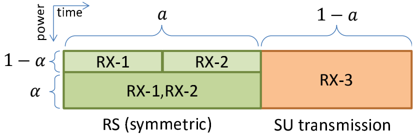
- IoT with many devices.
- Low-power sensor-like receivers.
- Can be served using the common message in the RS scheme [22].

Overloaded systems: Three-User example

- System: $M = 2$ antennas and $K = 3$ users.
- CSIT: $\alpha_1 = \alpha_2 = \alpha$ and $\alpha_3 = 0$.

Scheduling approach:

- The two users with CSIT are served together (fraction a of the time).
- User with no CSIT is served separately (fraction $1 - a$ of the time).



CSIT part:

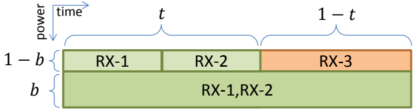
- RS used with sum DoF $(1 + \alpha)a$.
- DoF per user: $\frac{1+\alpha}{2}a$.
- Loss due to resource sharing.

No CSIT part:

- SU transmission.
- DoF: $1 - a$.

Overloaded systems: A RS approach

- Instead of scheduling, serve all together through one RS scheme.
- 2 private streams s_1 and s_2, and one common stream s_c.
- Users with CSIT: served through s_1 and s_2 and part of s_c.
- User with no CSIT: served using s_c.



Assume power control of private messages $(b \leq \alpha)$:

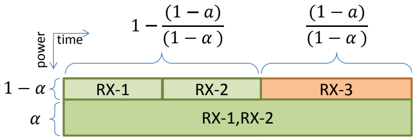
- DoF of RX-1 and RX-2: $b + \frac{(1-b)t}{2}$.
- DoF of RX-3: $(1-b)(1-t)$.

Assume we want to maintain DoF $1-a$ for RX-3.

What do RX-1 and RX-2 achieve?

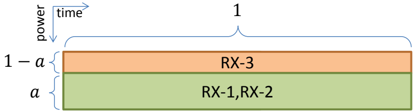
Overloaded systems: A RS approach

Case $a > \alpha$:



- RX-1 and RX-2: $\frac{\alpha+a}{2}$.
- RX-3: $1-a$.
- $\frac{\alpha+a}{2} > \frac{1+\alpha}{2}a$ for any $a < 1$.
- User-1 and user-2 achieve a gain.

Case $a \leq \alpha$:



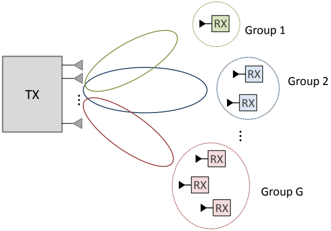
- RX-1 and RX-2: a.
- RX-3: $1-a$.
- $a > \frac{1+\alpha}{2}a$ for any $\alpha < 1$.
- User-1 and user-2 achieve a gain.

Multigroup multicast beamforming

Users clustered into groups depending on content demand.

- K users grouped into $\mathcal{G}_1, \ldots, \mathcal{G}_G$.
- One message for each group: $W_1, \ldots, W_G$.
- Classical beamforming:

$$\mathbf{x} = \sum_{g=1}^{G} \mathbf{p}_g s_g.$$



Achieving max-min fairness (perfect CSIT):

$$\mathcal{R}(P): \begin{cases} \max\limits_{\mathbf{P}_\mathrm{p}} & \min\limits_{g \in \{1,\ldots,G\}} \min\limits_{i \in \mathcal{G}_g} R_i \\ \text{s.t.} & \sum\limits_{g=1}^{G} \|\mathbf{p}_g\|^2 \leq P. \end{cases}$$

- **Overloaded scenarios**: M is not enough for interference nulling.
- Rate saturation (even with perfect CSIT) due to inter-group interference.

Multigroup multicast beamforming: Overloaded scenarios

- Assume that: $|\mathcal{G}_1| \leq |\mathcal{G}_2| \leq \ldots \leq |\mathcal{G}_G|$.
- To place each beam in the null space of all unintended groups, we need:

$$M \geq 1 + K - |\mathcal{G}_1|.$$

- Otherwise, DoF loss is experienced [16].

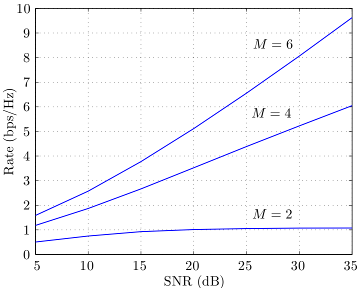


Figure: $K = 6$ users, $G = 3$ groups, $|\mathcal{G}_1| = 1$, $|\mathcal{G}_2| = 2$ and $|\mathcal{G}_3| = 3$.

Multigroup multicast beamforming: Rate-Splitting

- Split group messages: $W_g \mapsto W_{g0}, W_{g1}$.
- Common message: $W_{10}, \dots, W_{G0} \mapsto s_{\mathrm{c}}$.
- Designated messages: $W_{11}, \dots, W_{G1} \mapsto s_1, \dots, s_G$.
- Beamforming:

$$\mathbf{x} = \mathbf{p}_{\mathrm{c}} s_{\mathrm{c}} + \sum_{g=1}^{G} \mathbf{p}_g s_g.$$

- s_{c} decoded by all users, while s_g decoded by users in $\mathcal{G}_g$.

Achieving max-min fairness with RS:

$$\mathcal{R}_{\mathrm{RS}}(P): \begin{cases} \max\limits_{\mathbf{c},\mathbf{P}} & \min\limits_{g \in \{1,\dots,G\}} \left(C_g + \min\limits_{i \in \mathcal{G}_g} R_i \right) \\ \text{s.t.} & R_{\mathrm{c},k} \geq \sum_{g=1}^{G} C_g, \forall k \in \mathcal{K} \\ & C_g \geq 0, \forall g \in \{1,\dots,G\} \\ & \|\mathbf{p}_{\mathrm{c}}\|^2 + \sum\limits_{g=1}^{G} \|\mathbf{p}_g\|^2 \leq P \end{cases}$$

Multigroup multicast beamforming: Simulation results

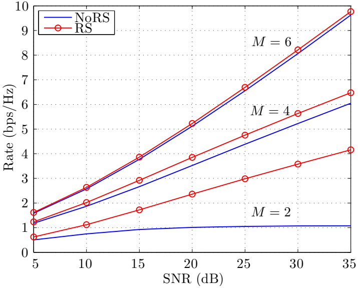


Figure: $K = 6$ users, $G = 3$ groups, $|\mathcal{G}_1| = 1$, $|\mathcal{G}_2| = 2$ and $|\mathcal{G}_3| = 3$.

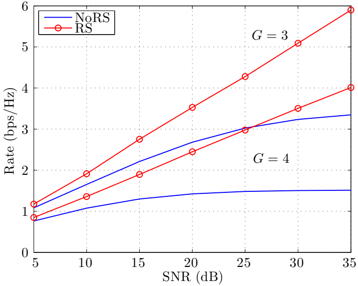
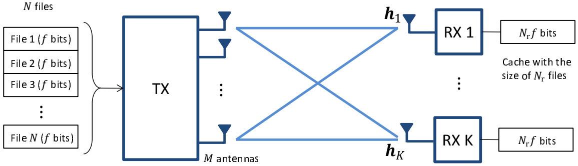


Figure: $M = 4$ antennas, $|\mathcal{G}_g| = 2$ users per group, $G = 3$ and 4.

RS is powerful in combating inter-group interference in overloaded scenarios.

Cache aided transmission



- N distinct files at the transmitter: $W_1, \ldots, W_N$.
- Length of each file: f bits.
- Each receiver has a cache (memory).
- Cache size: $N_\mathrm{r} f$ bits (or N_r) files.
- **Normalization**: $f = \log_2(P)$. Hence: 1 file $\Leftrightarrow$ 1 DoF.

Caching: Proactively fill caches with data to reduce the load in peak times.

- Placement phase:
  - Off-peak times.
  - Fill caches with data.
- Delivery phase:
  - Peak times.
  - Deliver files requested by users.

Cache aided transmission: Examples

Example: No caching

- Assume $N_\mathrm{r} = 0$ (or no caching is carried out).
- Each receiver demands a distinct file (in the worst case).
- TX needs to send K messages $\Rightarrow$ standard BC transmission.
- **Perfect CSIT**: No problem as each user gets 1 DoF.
- Deliver N_d files: $N_\mathrm{d} = \mathrm{DoF} \times T$ (normalized time).

Example: Uncoded caching

- **Placement**: each user caches the same $\frac{N_\mathrm{r}}{N} f$ bits of each file.
- **Delivery**: each user requires the remaining $(1 - \frac{N_\mathrm{r}}{N}) f$ bits.
- **Perfect CSIT**: 1 DoF per user, K DoF in total.
- **Delivery time**: $T = 1 - \frac{N_\mathrm{r}}{N}$ (less than 1 file required per user).
- **Imperfect CSIT**: Less DoF per user $\Rightarrow$ longer delivery time.

For a given CSIT quality, caching reduces delivery time.

For fixed delivery time, caching relaxes CSIT (as less DoF is needed).

Cache aided transmission: Coded caching

Assume we want to maintain $T = 1 - \frac{N_r}{N}$ with $\alpha < 1$.

- $\alpha < 1$ yields $\mathrm{DoF} < 1$ per user.
- Not possible with uncoded caching.

Coded caching:

- Instead of caching the same $\frac{N_r}{N} f$ bits of each file, something more clever is done.
- The whole file library is placed in the collective (global) cache memory of users.
- Data available at one user may be required by some other user (side information).
- This creates coded **multicasting opportunities** despite distinct requests.

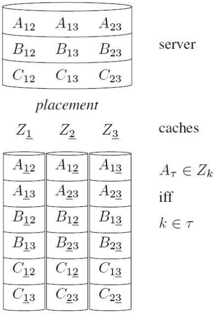


Figure: Coded placement example $N = K = 3$ [27].

- Common data mixed with private data to exploit partial CSIT.

common data + private data $\Rightarrow$ only partial CSIT needed.

Cache aided transmission: Coded caching

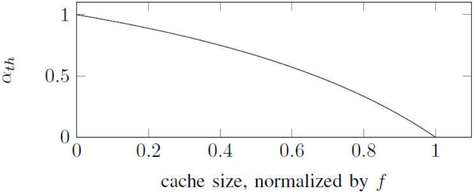


Figure: Required alpha to achieve $T = 1 - \frac{N_r}{N}$ for $K = N = 2$ [27].

- Coded caching scheme for MISO-BC proposed in [27].
- Scheme depends on transmitting common + private streams.

$$\mathbf{x} = \mathbf{p}_c s_c + \sum_{k=1}^{K} \mathbf{p}_k s_k$$

- Common message used for coded multicasting + some private parts.
- Multicasting opportunities created $\Rightarrow$ CSIT requirements relaxed.
- RS is included in the machinery used to achieve such gains.

Rate-Splitting in 5G

Standardization Issues

- (H/T)RS is a generalized strategy
  - Conventional SU/MU-MIMO and CoMP as special cases.
  - new SU/MU/RS mode switching in 5G depending on the SNR and the CSIT quality.
- New transmission mode indicator (DCI format)
  - Inform the Tx mode and the relevant information required for demodulation
- New signaling from BS to UEs
  - Number and type of messages (common/private)
  - Modulation and coding scheme of all common/private message
  - Information about whether common message is intended for the user or not
  - Transmit power of each message.
- CSI feedback mechanisms and signaling
  - Knowledge about the CSIT accuracy to allocate power to the common and private messages, e.g. computed by a UE and reported back to the BS.
  - Scheduling and Tx strategy decided based on CSIT accuracies from all users in all subbands.
  - CSI reporting on PUCCH and PUSCH.
  - Some CSIT patterns lead to a higher DoF than others [23].

From LTE Rel. 13/14 to RS

- NAICS in Rel-12
  - Network-Assisted Interference Cancellation and Suppression.
  - Providing knowledge about interfering transmissions at the receivers.
  - Allows the use of more advanced receivers (joint decoding, SIC).
- Downlink Multiuser Superposition Transmission (MUST)
  - Study item approved (3GPP).
  - Non-Orthogonal Multiple Access (NOMA).
  - Uses superposition coding at the transmitter.
  - Relies on SIC at the receivers.
  - Metric of interest: sum-rate, fairness, delays etc.

The machinery required for RS is already being studied, discussed and developed.

- RS schemes can utilize such developments (or the other way around!).
- RS can fit in nicely.
- RS complements other schemes, and vice versa.

Rate-Splitting in 5G

Conclusions and Future Challenges

- 4G and current 5G candidates rely on private message transmissions
  - Such a strategy is only motivated in the presence of perfect CSIT
  - Apply techniques designed for perfect CSIT to imperfect CSIT
- A more efficient strategy is to rely on the superposed transmission of common and private messages (through rate-splitting).
  - For long kept in the realm of SISO IC but is actually useful for a wide range of scenarios.
  - Motivated by information theory for the realistic scenario of imperfect CSIT.
  - Benefits in terms of spectral and energy efficiencies, reliability, CSI feedback overhead reduction.
- RS through the transmission of common and private message would have fundamental changes in the design of PHY and Lower MAC.
  - A gold mine of research problems for academia and of standard specification issues for industry
- The standardization of rate-splitting can leverage 3GPP current study/work items

Future Challenges: A gold mine of research problems

Fundamental Limits

- DoF region for K-user MISO BC with imperfect CSIT [8].
- Capacity region of K-user MISO BC with imperfect CSIT: DPC + RS?
- DoF region for MIMO BC with imperfect CSIT [10, 11, 12].
- DoF region of overloaded MISO BC with imperfect CSIT [22].
- DoF region for MISO IC with imperfect CSIT [21]. TRS?
- DoF region for MIMO IC with imperfect CSIT [11]. RS + IA?
- Interplay between RS and coded caching.

Optimization

- Ergodic sum-rate maximization for BC [14].
- Robust Max-Min Fairness for BC [15].
- RS beamforming optimization for other types of channels.

PHY challenges

- Finite SNR rate analysis [9].
- Energy efficiency of RS-based transmission.
- Space-time/frequency RS [24, 25, 9].
- RS with multi-carrier transmissions.
- RS with non-linear precoding.
- Diversity (and BER) performance of RS-based strategies.

Future Challenges

PHY challenges (continued)

- RS for Multigroup Multicast [16].
- RS/HRS for Massive MIMO [13].
- RS as a way to combat pilot contamination.
- RS to mitigate hardware impairments.
- RS in higher frequency bands operation (e.g. millimeter-wave).
- RS-based network MIMO.
- Coordination/cooperation among distributed antennas in homogeneous and heterogeneous network deployments.
- RS in half-duplex relay.
- RS in full duplex.
- RS and NOMA/MUST [22].
- RS and superposition of multicast and unicast messages.
- RS and physical layer security.

PHY/MAC challenges

- User pairing and scheduling of common and private messages.
- RS design with Quality of Experience (QoE) and traffic constraints.

Performance Analysis

- Performance analysis of RS using stochastic geometry.

Future Challenges

Standardization

- Link and system-level evaluations of RS.
- MIMO receiver implementation.
- Transmission schemes/mode.
- CSI feedback mechanisms.
- Downlink and uplink signaling.

References I

[1] E. Bjornson, E. G. Larsson and T. L. Marzetta, “Massive MIMO: ten myths and one critical question,” IEEE Communications Magazine, vol. 54, no. 2, pp. 114-123, February 2016.

[2] E. G. Larsson, O. Edfors, F. Tufvesson and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” IEEE Communications Magazine, vol. 52, no. 2, pp. 186-195, February 2014.

[3] N. Jindal, “MIMO Broadcast Channels With Finite-Rate Feedback,” Trans. on Info Theory, vol. 52, no. 11, pp. 5045-5060, Nov. 2006.

[4] B. Clerckx, G. Kim, and S. Kim, “MU-MIMO with channel statistics-based codebooks in spatially correlated channels,” in IEEE Global Communications Conf. (Globecom), Nov 2008, pp. 15.

[5] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, Multiuser MIMO achievable rates with downlink training and channel state feedback, IEEE Trans. Inf. Theory, vol. 56, no. 6, pp. 28452866, Jun. 2010.

[6] A. El Gamal and Y.-H. Kim, *Network information theory*, Cambridge university press, 2011.

[7] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi, Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT, IEEE Trans. Inf. Theory, vol. 59, no. 1, pp. 315328, Jan. 2013.

[8] A.G. Davoodi, S.A. Jafar, “Aligned Image Sets under Channel Uncertainty: Settling a Conjecture by Lapidoth, Shamai and Wigger on the Collapse of Degrees of Freedom under Finite Precision CSIT,” arXiv:1403.1541.

[9] C. Hao, Y. Wu and B. Clerckx, “Rate analysis of two-receiver MISO broadcast channel with finite rate feedback: A rate-splitting approach,” IEEE Trans. on Comm., vol. 63, no. 9, pp. 3232-3246, Sept. 2015.

[10] J. Chen and P. Elia, Symmetric two-user MIMO BC with evolving feedback, in Inf. Theory and Applications Workshop (ITA), 2014, San Diego, CA, Feb 2014, pp. 15.

[11] C. Hao, B. Rassouli and B. Clerckx, “Achievable DoF Regions of MIMO Networks with Imperfect CSIT,” submitted to IEEE Trans. on Info Theory, arXiv:1603.07513.

[12] B. Yuan, S.A. Jafar, “Elevated Multiplexing and Signal Space Partitioning in the 2 User MIMO IC with Partial CSIT,” IEEE SPAWC 2016.

[13] M. Dai, B. Clerckx, D. Gesbert and G. Caire, “A Rate Splitting Strategy for Massive MIMO with Imperfect CSIT,” IEEE IEEE Trans. on Wireless Comm., vol. 15, no. 7, pp. 4611-4624, July 2016.

[14] H. Joudeh and B. Clerckx, “Sum-Rate Maximization for Linearly Precoded Downlink Multiuser MISO Systems with Partial CSIT: A Rate-Splitting Approach,” IEEE Trans. on Comm., accepted for publication, 2016, arXiv:1602.09028.

References II

[15] H. Joudeh and B. Clerckx, "Robust Transmission in Downlink Multiuser MISO Systems: A Rate-Splitting Approach," IEEE Trans. on Sig Proc, accepted for publication, 2016, arXiv:1602.04345.

[16] H. Joudeh and B. Clerckx, "A Rate-Splitting Strategy for Max-Min Fair Multigroup Multicasting," IEEE SPAWC 2016.

[17] C. Lim, T. Yoo, B. Clerckx, B. Lee, and B. Shim, "Recent trend of multiuser MIMO in LTE-advanced," IEEE Commun. Mag., vol. 51, no. 3, pp. 127135, March 2013.

[18] A.Adhikary, J.Nam, J.Ahn and G.Caire. "Joint spatial division and multiplexing - the large-scale array regime," IEEE TIT, Oct. 2013.

[19] J. Chen and V. Lau, "Two-tier precoding for FDD multi-cell massive MIMO time-varying interference networks," IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 12301238, June 2014.

[20] X. Yi, D. Gesbert, S. Yang, and M. Kobayashi, On the dof of the multiple-antenna time correlated interference channel with delayed csit, in 2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), Nov 2012, pp. 15661570.

[21] C. Hao and B. Clerckx, "MISO Networks with Imperfect CSIT: A Topological Rate Splitting Approach," submitted to IEEE Trans on comm., arXiv:1602.03768.

[22] E. Piovano, H. Joudeh and B. Clerckx, "Overloaded MU-MISO Transmission with Imperfect CSIT," 2016 Asilomar Conference on Signals, Systems, and Computers.

[23] B. Rassouli, C. Hao and B. Clerckx, "DoF Analysis of the MIMO Broadcast Channel with Alternating/Hybrid CSIT," IEEE Trans. on Info Theory, vol. 62, no. 3, pp. 1312-1325, Mar. 2016.

[24] R. Tandon, S. Jafar, S. S. Shitz, and H. Poor, On the synergistic benefits of alternating CSIT for the MISO broadcast channel, IEEE Trans. Inf. Theory., vol. 59, no. 7, pp. 41064128, Jul. 2013.

[25] J. Chen and P. Elia, Optimal DoF region of the two-user MISO-BC with general alternating CSIT, in Proc. Asilomar Conf. Signals, Syst. Comput., Pacific Grove, CA, USA, Nov. 2013, pp. 18601864.

[26] B. Clerckx, H. Joudeh, C. Hao, M. Dai and B. Rassouli, "Rate Splitting for MIMO Wireless Networks: A Promising PHY-Layer Strategy for LTE Evolution," IEEE Communications Magazine.

[27] J. Zhang, F. Engelmann and P. Elia, "Coded caching for reducing CSIT-feedback in wireless communications," Allerton, Monticello, IL, 2015, pp. 1099-1105.