MIMO Broadcast and Interference Channels towards 5G

Dirk Slock and Petros Elia

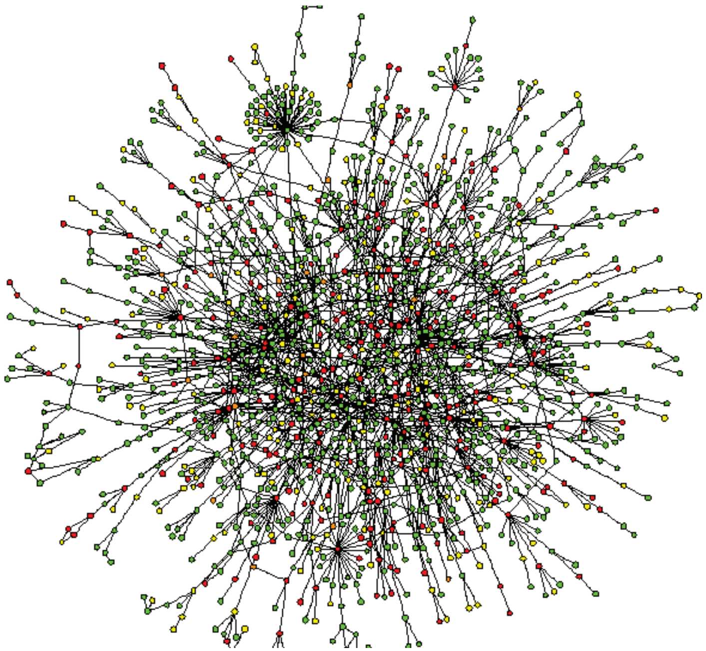
eurecom

Sophia Antipolis - France

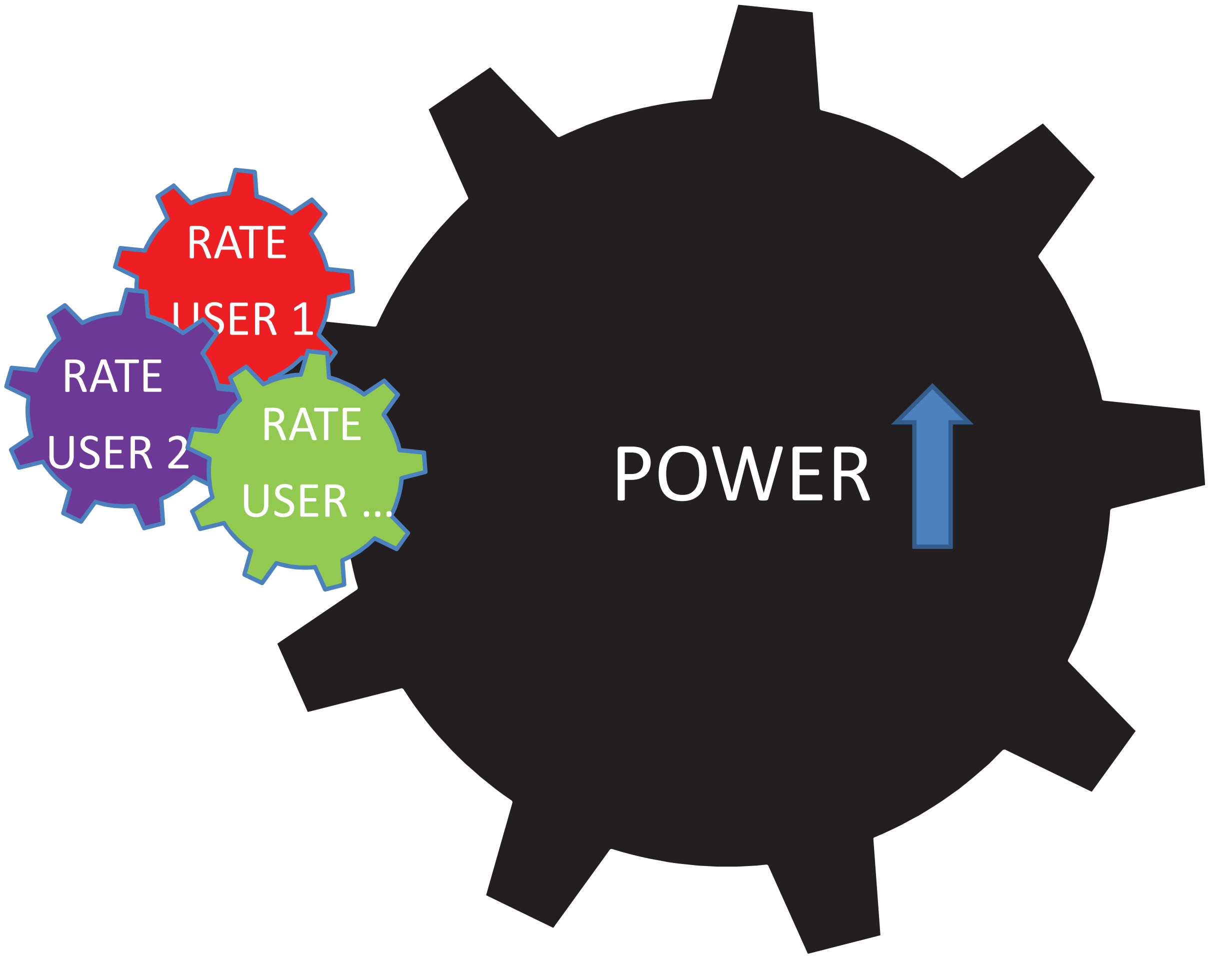
Wireless communications networks

Big Challenge:

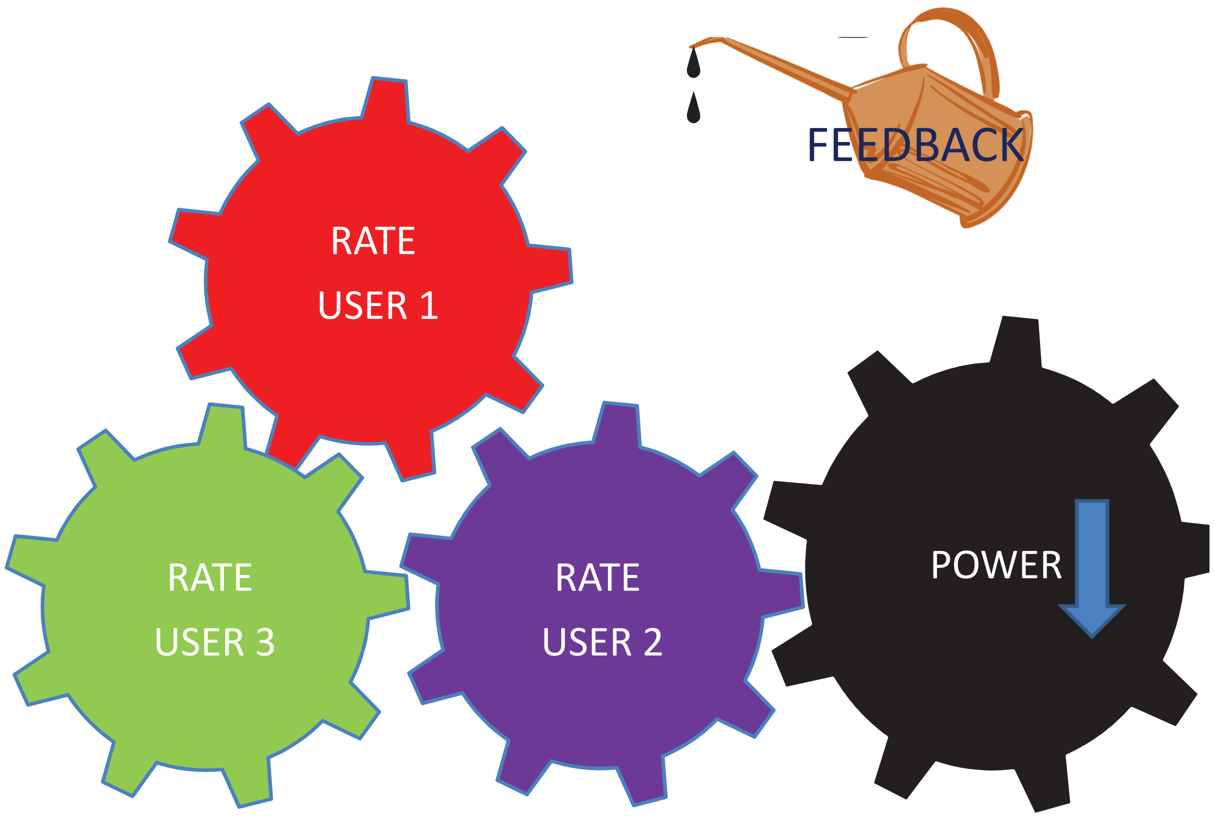
Efficient communication in $\geq$ 5G wireless networks



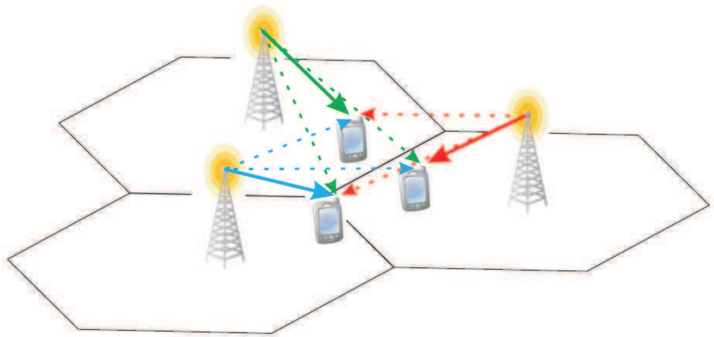
Distilled point-of-view in this presentation

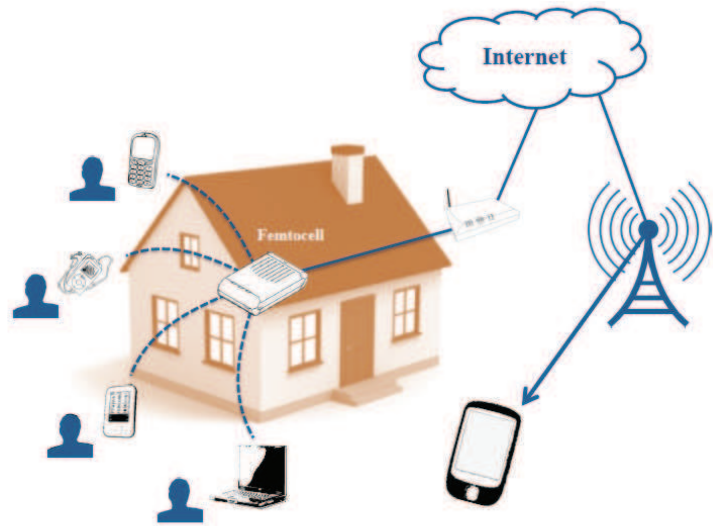


Distilled point-of-view of Part 1



Distilled point-of-view of Part 2





- Part 2: A system view and summary of many tools
  - ⋆ Many settings:
    - ∗ Coexistence of macrocells and small cells, especially when small cells are considered part of the cellular solution.
  - ⋆ Many candidate tools and measures of performance
    - ∗ Interference Alignment (IA) ....

Summary of tutorial - Part 1

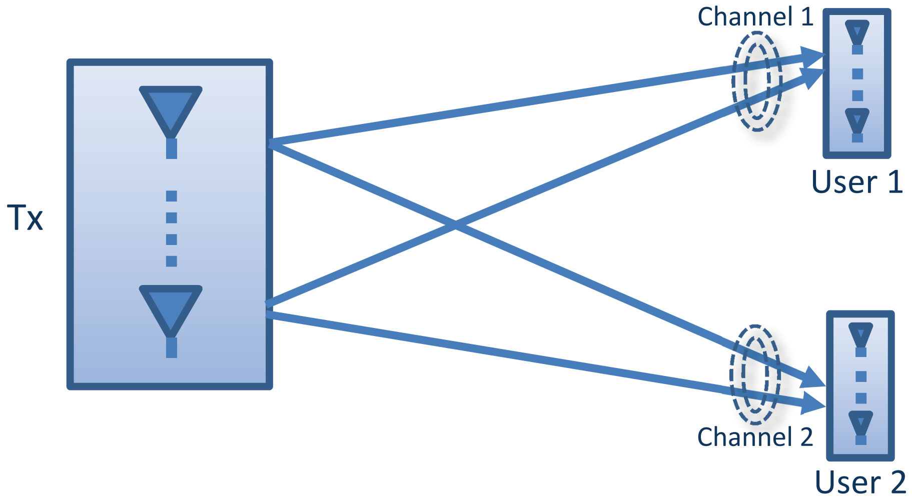
- Feedback in classical multiuser channels
  - Part 1-A
    - Motivation (Why feedback is important)
    - Basics - intro (Capacity, Degrees-of-Freedom)
    - New encoding/decoding/feedback tools
  - Part 1-B: Qualitative insight over a restricted setting
    - A unified exposition and a general framework
    - INSIGHT and answers to fundamental questions
    - Open problems

Summary of tutorial - Part 2

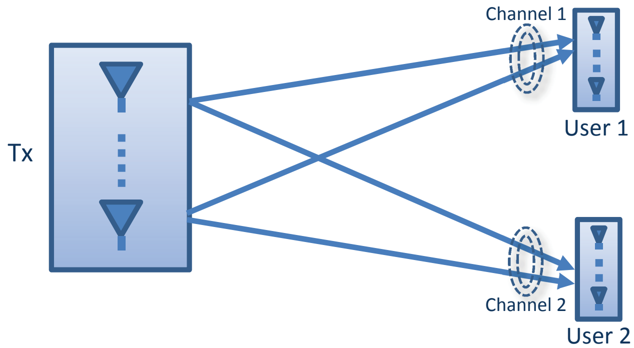
- Interference single cell
- Utility functions: SINR balancing
- Uplink/downlink(UL/DL) duality; BC,MAC; BF&DPC
- BC with user selection: DPC vs BF

- Interference multi-cell/HetNets: Interference Channel (IFC)
- Degrees of Freedom (DoF) and Interference Alignment (IA)
- Weighted Sum Rate (WSR) maximization and UL/DL duality
- Deterministic Annealing to find global max WSR
- Distributed Channel State Information at the Transmitter (CSIT) acquisition, netDoF
- Delayed CSIT, optimal handling of CSIT FB dead times
- Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels

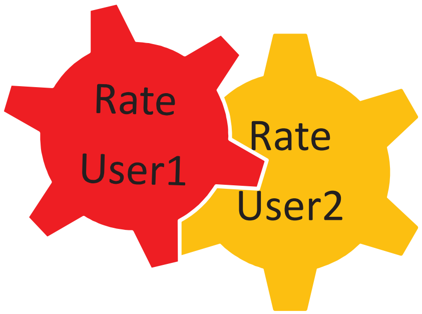
Typical multiuser scenario



Typical multiuser scenario: Interference



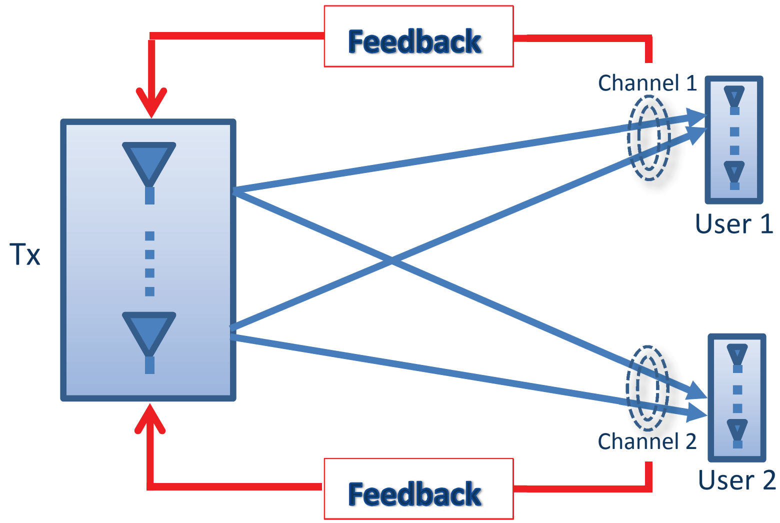
USERS INTERFERE, AND MUST SHARE THE PIE



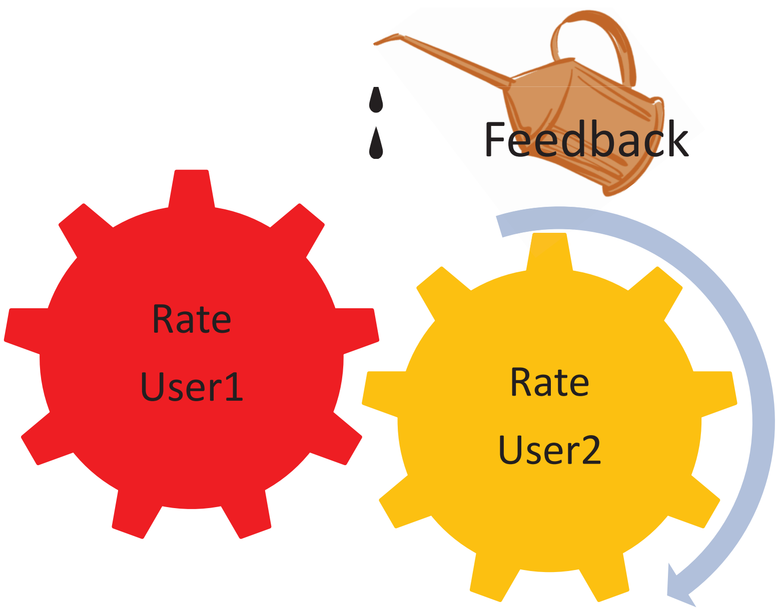
Communications with feedback

Feedback: notify transmitter of the channel state

Channel State Information at Transmitter (CSIT)

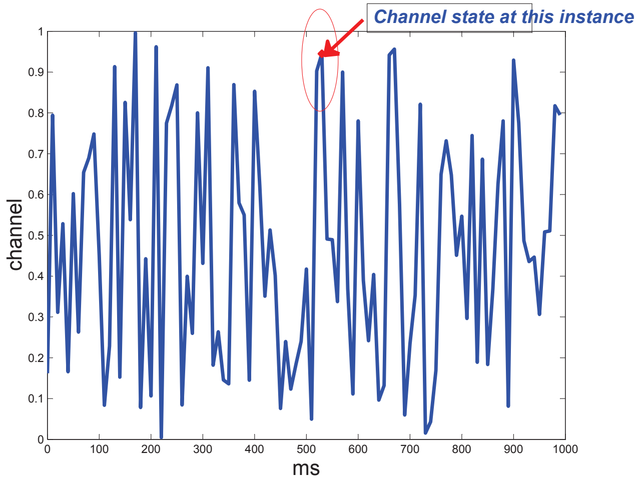


Feedback is crucial: Interference $\downarrow$ Rates $\uparrow$



Communications with feedback$_2$

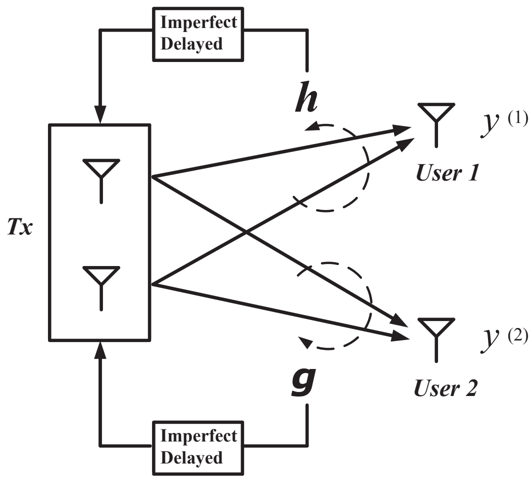
BUT! Feedback is hard to get

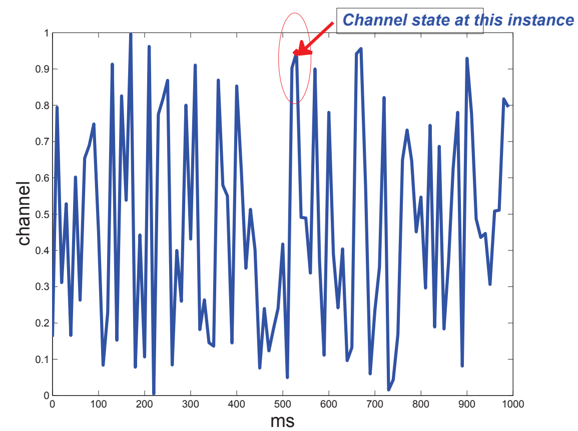


Long-standing challenge:
How to use imperfect feedback?

optimize (SNR Rate1 Rate2)

What is the source of this challenge?





- Transmit: $(\overbrace{\text{Inverse-channel}}^{\text{Feedback}} \times \text{Message}) \Rightarrow$ separates users' messages
  - ⋆ Channel × Inverse-channel × Message $\longrightarrow$ Message OK
- BUT, channel changes: Feedback can be imperfect, limited and delayed
  - ⋆ Channel × Approximately-inverse-channel × Message → Ⓡ‡♠H∅⅂ ≗

Massive gains from resolving challenge

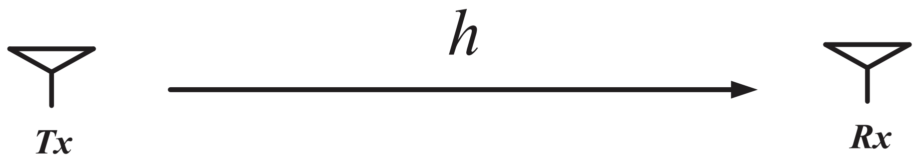
- No feedback: one user served at a time
- Perfect and immediate feedback: many users at a time

- Challenge: new algorithms that bridge gap

- Recent tools brought unprecedented excitement
  - ⋆ New insight sparked worldwide race to resolve challenge
  - ⋆ Much of work done after 2012

Quick summary of basics

Flat fading (single-input single-output) channel model



$$y_t = h_t x_t + z_t$$

- Ergodic (average) capacity $\quad \mathbb{E}_h[C(h_t)] : \mathbb{E}_h \log\big(1 + P|h_t|^2\big) \approx \log P$

- DEGREES OF FREEDOM d (HIGH SNR REGIME)

$$\text{CAPACITY } \approx d \cdot \ \log \text{SNR} \ = d \cdot \ \log P$$

(for large SNR)

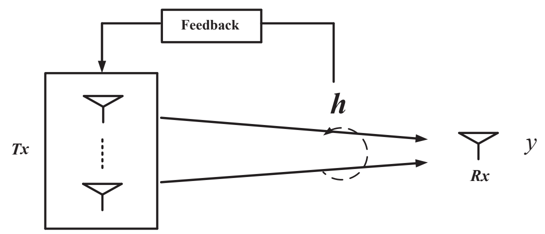
- Intuition: Number of dimensions available (seen) at a user

Single link (SISO) degrees of freedom

$$\text{DoF} = d \triangleq \lim_{P\to\infty} \frac{\text{Capacity}}{\log P} = \lim_{P\to\infty} \frac{\approx \log P}{\log P} = 1$$

$$\Rightarrow \text{ SISO: DoF } = 1$$

- Same holds for $n \times 1$ MISO (multiple input single output):



Importance of DoF

DoF increase means exponential power reductions

- Want to communicate at rate R
- Over 'system' with d DoF:

$$C \approx d \log_2 P$$

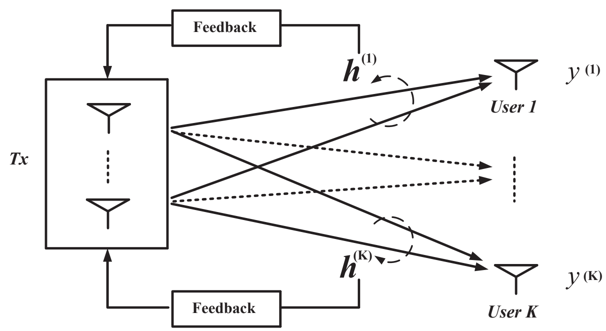
- Thus minimum power $P_{\min}$ so that

$$R \approx C \approx d \log_2 P_{\min}$$

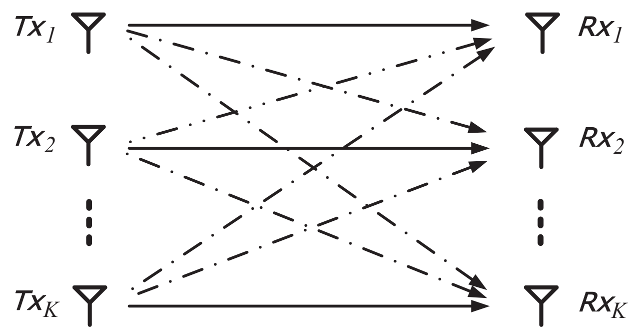
$$\Rightarrow P_{\min} \approx 2^{R/d}$$

Multiuser Channels suffer from interference

- Interference: users must share signal dimensions
  - ⋆ DoF reduction ⇒ Rates ↓, Power ↑,

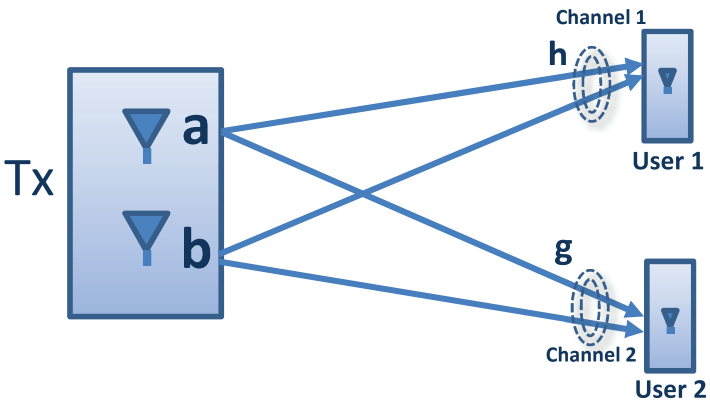


Multiuser Broadcast Channel



Multiuser Interference Channel
Multiuser X Channel

Example: interference in two-user MISO BC



- Let information symbol "a" for user 1 $\quad \mathbb{E}|a|^2 = P$
- Let information symbol "b" for user 2 $\quad \mathbb{E}|b|^2 = P$

Example: interference in two-user MISO BC$_1$

- No feedback $\Rightarrow$ transmit $\boldsymbol{x} = \begin{bmatrix} a \\ b \end{bmatrix}$

- User 1 receives:

$$y^{(1)} = \boldsymbol{h}^T \boldsymbol{x} + w = [\, h_1 \;\; h_2 \,] \begin{bmatrix} a \\ b \end{bmatrix} = h_1 a + \underbrace{h_2 b + w}_{\text{NOISE POWER } \approx P+1}$$

- User 1 treats $h_2 b$ as noise:

$$\text{average effective SNR} = \frac{\text{'signal' power}}{\text{'noise' power}} \approx \frac{P}{P+1} \approx \text{Constant}$$

- Received SNR does not increase with transmit power!

Example: interference in two-user MISO BC_2

- Thus maximum rate $R_{\max}$ does not increase with increasing transmit power

$$R_{\max} \approx \log\left(1 + \frac{P}{P+1}\right) = \text{constant}$$

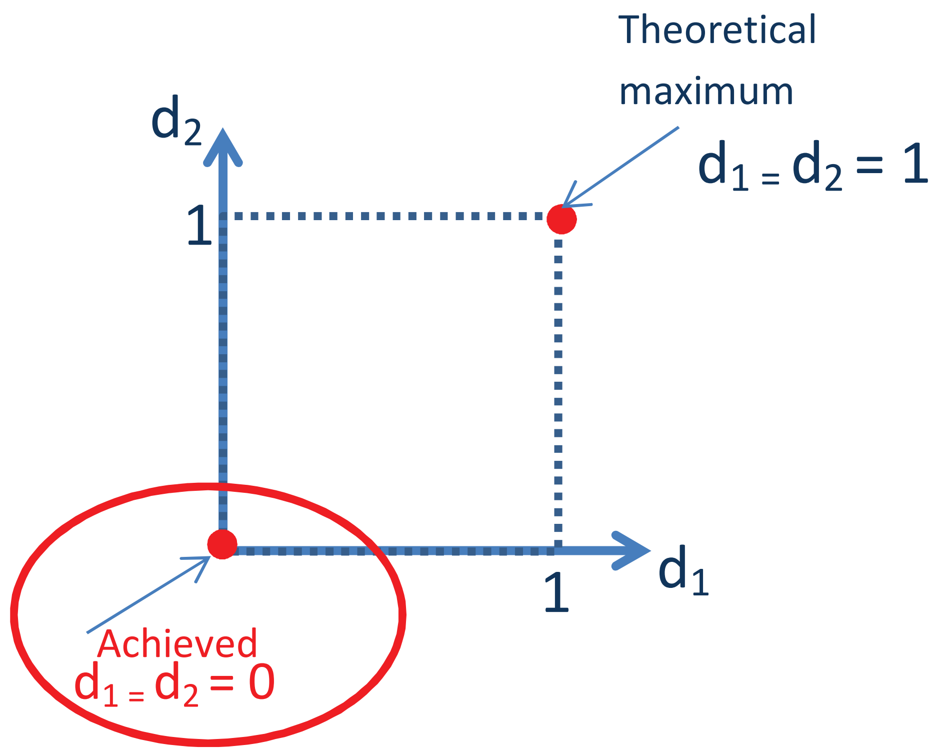
- Which means, zero DoF

$$d = \lim_{P \to \infty} \frac{R_{\max}}{\log P} = \lim_{P \to \infty} \frac{\text{constant}}{\log P} = 0$$

⋆ $\Rightarrow$ Massive damage from inter-user interference

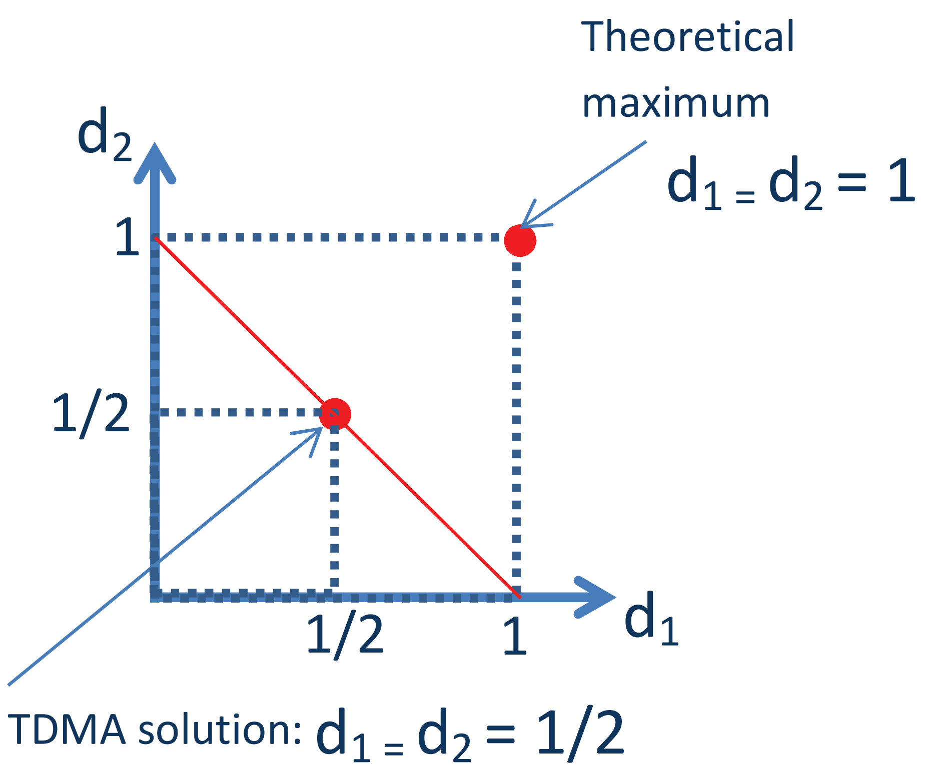
Example: interference in two-user MISO BC_3

Treating interference as noise



No-Feedback: Time division is DoF optimal

TDMA SOLUTION



Precoding with perfect feedback

- But what if we could feedback the channel state?

$$\mathbf{H} = \begin{bmatrix} \boldsymbol{h}^T \\ \boldsymbol{g}^T \end{bmatrix}$$

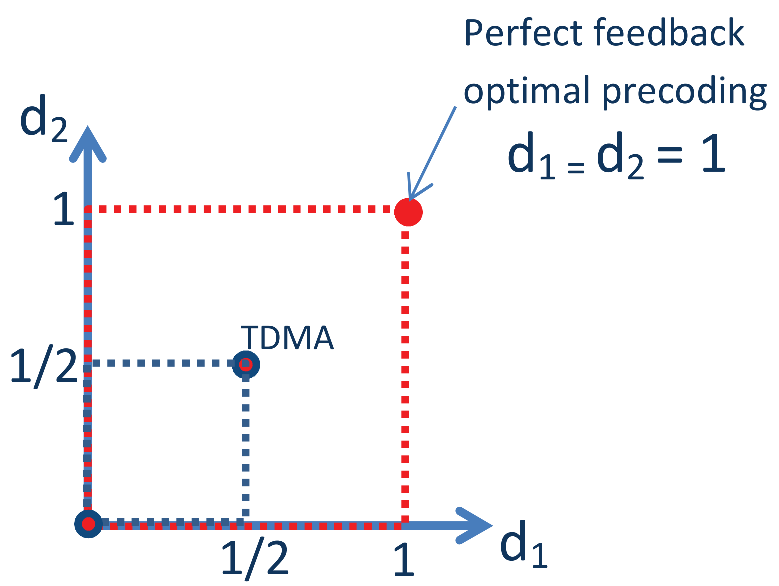
- Send $\mathbf{H}$ to the transmitter, and *precode*
- Instead of sending $\begin{bmatrix} a \\ b \end{bmatrix}$, now could send $\quad \boldsymbol{x} = \mathbf{H}^{-1} \begin{bmatrix} a \\ b \end{bmatrix}$.

$$\begin{bmatrix} y^{(1)} \\ y^{(2)} \end{bmatrix} = \mathbf{H}\boldsymbol{x} + \boldsymbol{z} = \mathbf{H} \overbrace{\mathbf{H}^{-1} \begin{bmatrix} a \\ b \end{bmatrix}}^{\boldsymbol{x}} + \boldsymbol{z} = \begin{bmatrix} a \\ b \end{bmatrix} + \boldsymbol{z}$$

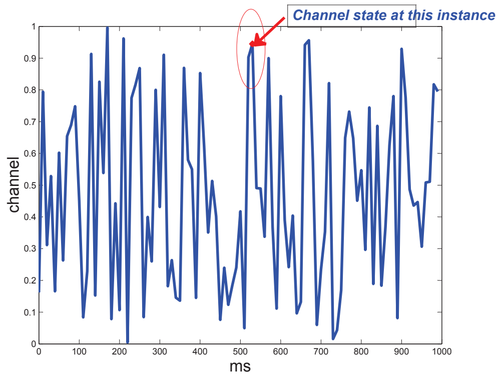
$$\begin{aligned} y^{(1)} &= a + z^{(1)} \qquad \text{user 1: } \ \text{DoF} = d_1 = 1 \\ y^{(2)} &= b + z^{(2)} \qquad \text{user 2: } \ \text{DoF} = d_2 = 1 \end{aligned}$$

Precoding with perfect feedback$_1$

- Precoding with perfect feedback allows for optimal DoF
- *channel state information at the transmitter* (CSIT) is important
  - ⋆ allows for separation of signals



But remember: perfect feedback is infeasible



- How to exploit predicted CSIT
- How to exploit delayed CSIT
- How to exploit imperfect CSIT
- How to minimize total amount of (delayed + current) feedback?
- How to achieve optimality even with feedback delays?
- How to utilize gradually arriving feedback?
- How much feedback quality and when?

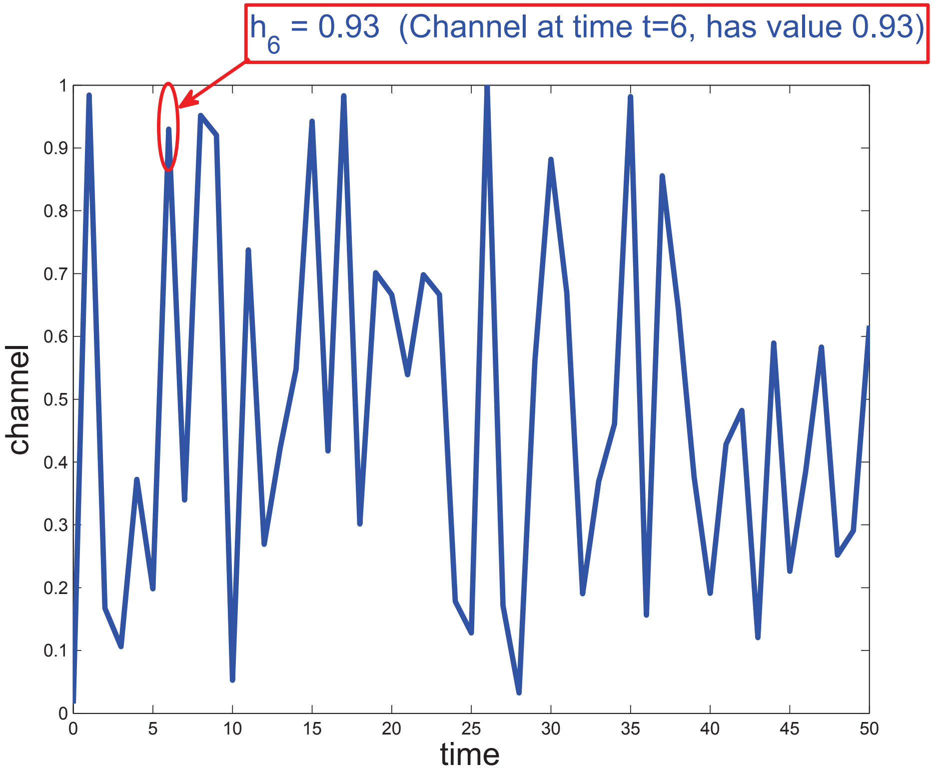
Toy examples for insight

Of course, the problem has randomness

Let us get some insight on the involved randomness

Let us look at some (simplistic) toy examples

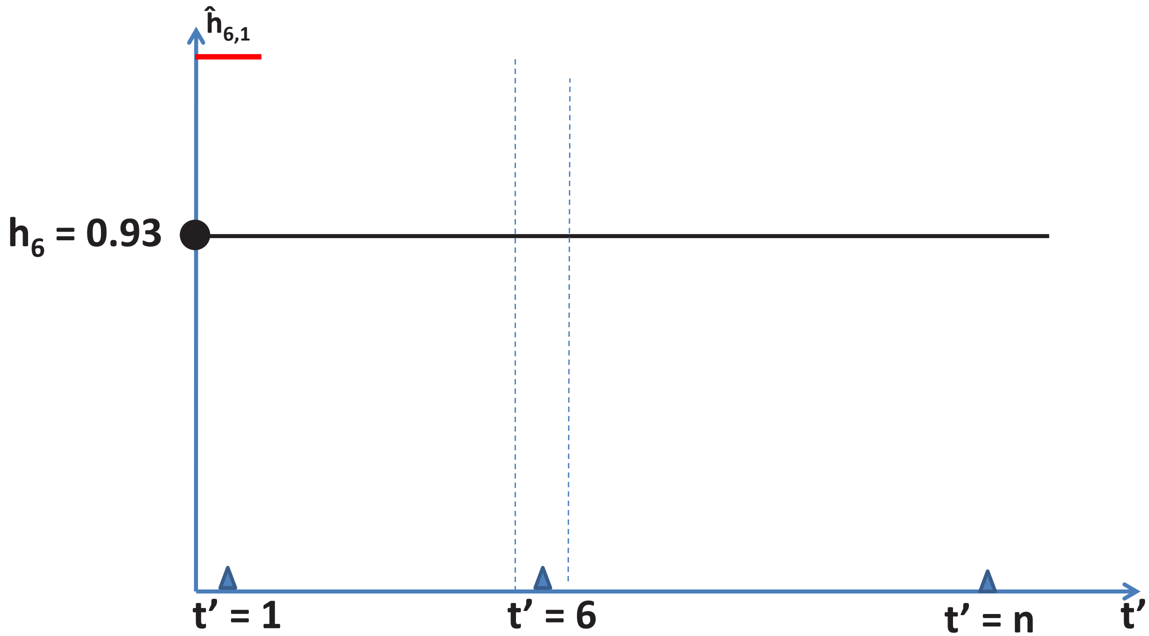
Progressive knowledge of channel



Progressive knowledge of channel

What do we know - at any point in time t' - about channel $\mathbf{h}_t$ (e.g **t=6**)?

$\hat{\mathbf{h}}_{6,t'}$ **t' = 1,2,3.....**

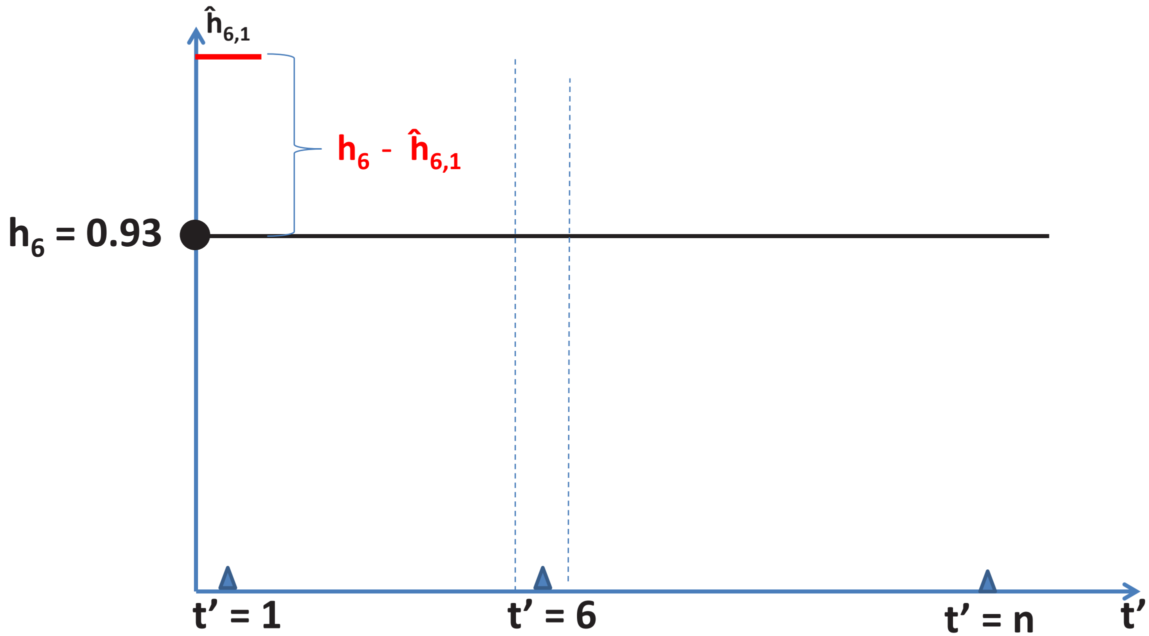


Knowledge at time t' = 1,2,3.....

No prediction at $t' = 1$ of $\boldsymbol{h}_6$

What do we know - at any point in time t' - about channel $\mathbf{h}_6$?

$\hat{\mathbf{h}}_{6,t'}$ **t' = 1,2,3.....**

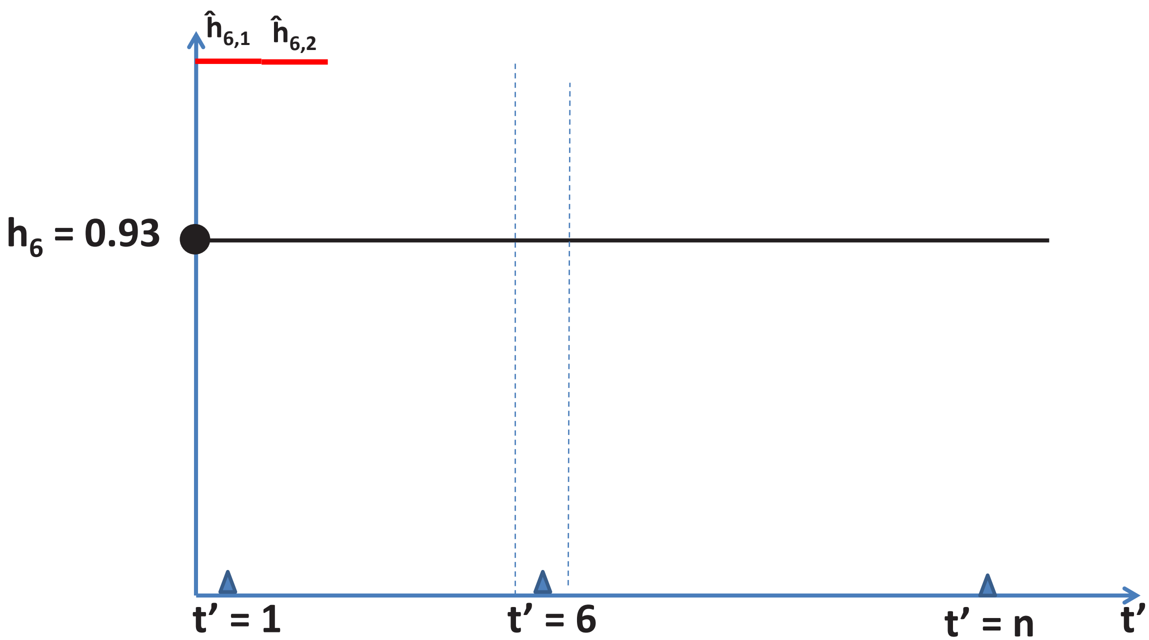


Knowledge at time t' = 1,2,3.....

Still no (of $\boldsymbol{h}_6$) prediction at $t' = 2$

What do we know - at any point in time t' - about channel $\mathbf{h}_6$?

$\hat{\mathbf{h}}_{6,t'}$ t' = 1,2,3.....

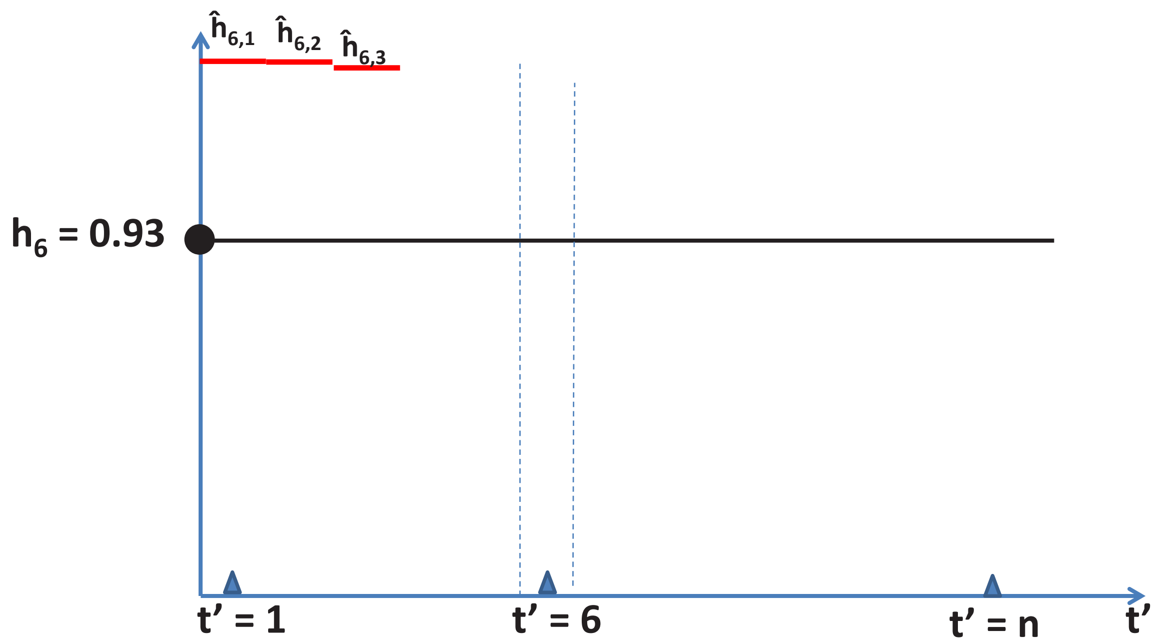


Knowledge at time t' = 1,2,3.....

Vague prediction (of $\boldsymbol{h}_6$) at time $t' = 3$ - high error

What do we know - at any point in time t' - about channel $\mathbf{h_6}$?

$\hat{\mathbf{h}}_{6,t'}$ **t' = 1,2,3…..**

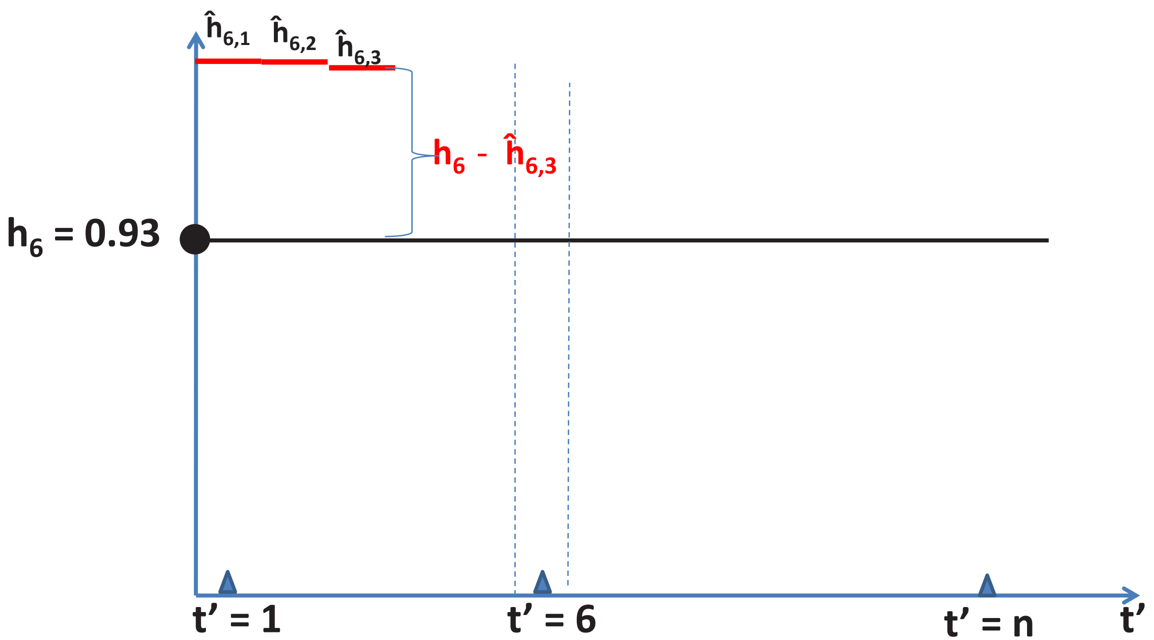


Knowledge at time t' = 1,2,3…..

Vague prediction (of $\boldsymbol{h}_6$) at time $t' = 3$ - high error[1]

What do we know - at any point in time t' - about channel $\mathbf{h}_6$?

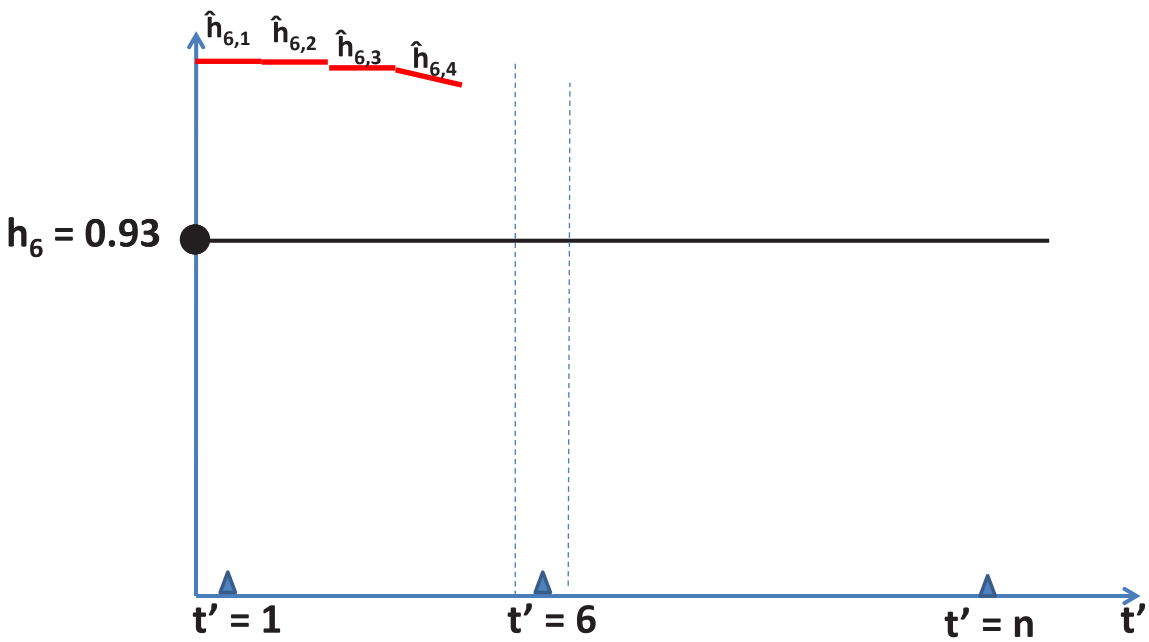
$\hat{\mathbf{h}}_{6,t'}$ **t' = 1,2,3.....**



Knowledge at time t' = 1,2,3.....

What do we know - at any point in time t' - about channel $\mathbf{h}_6$?

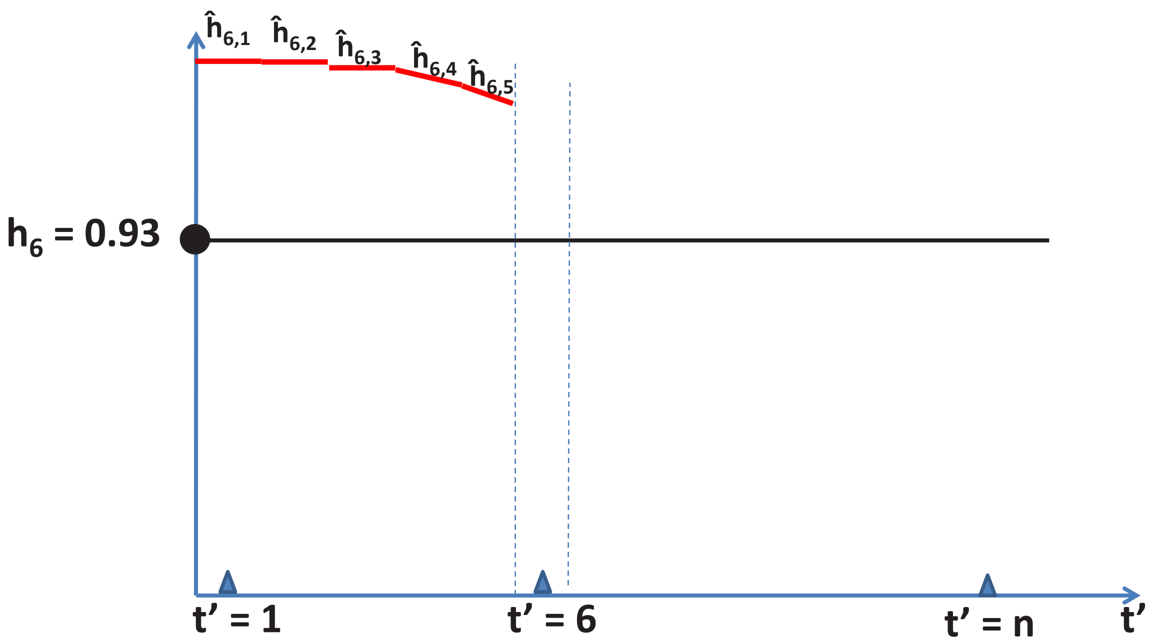
$\hat{\mathbf{h}}_{6,t'}$ **t' = 1,2,3.....**



Knowledge at time t' = 1,2,3.....

What do we know - at any point in time t' - about channel $\mathbf{h_6}$?

$\hat{h}_{6,t'}$ t' = 1,2,3.....

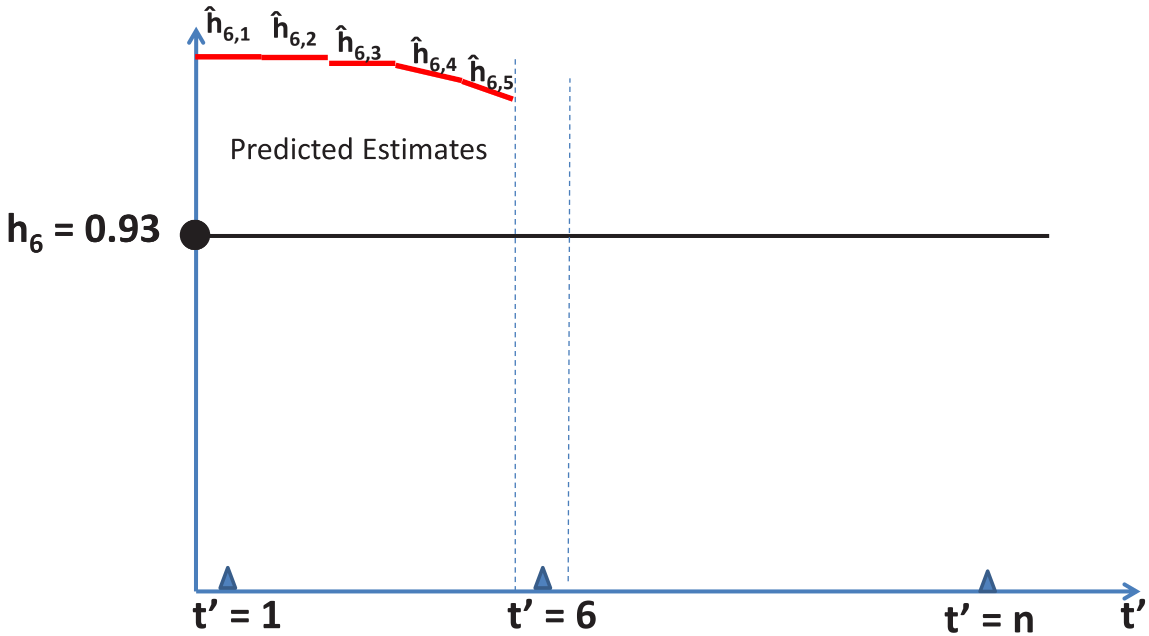


Knowledge at time t' = 1,2,3.....

These are the predicted estimates of $\boldsymbol{h}_6$

What do we know - at any point in time t' - about channel $\mathbf{h}_6$?

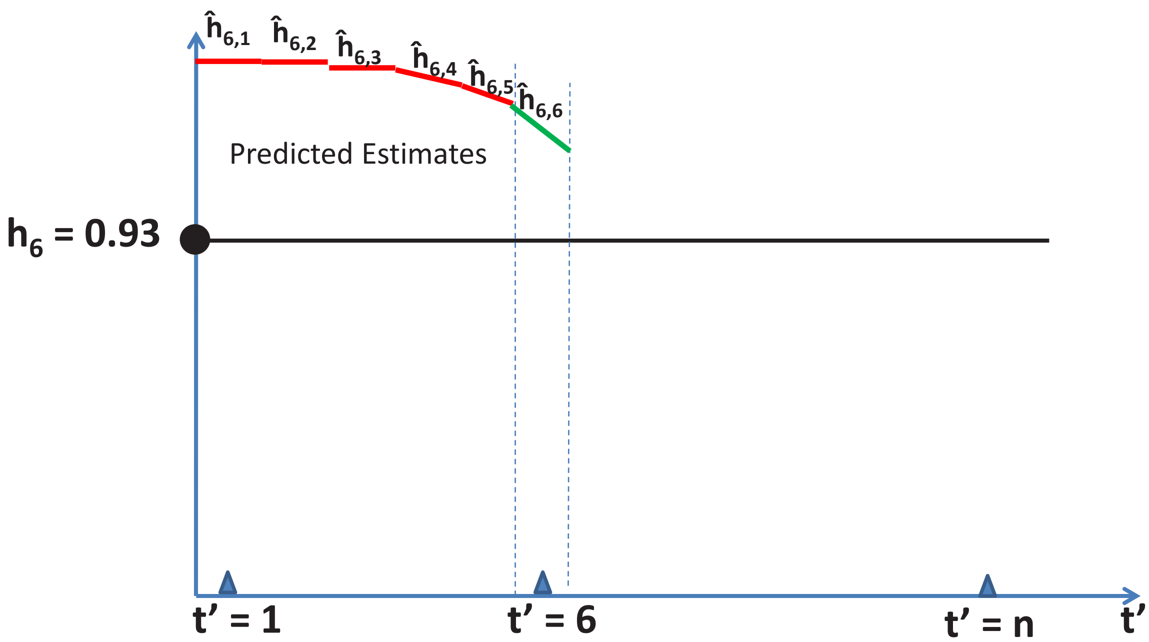
$\hat{\mathbf{h}}_{6,t'}$ t' = 1,2,3.....



Knowledge at time t' = 1,2,3.....

What do we know - at any point in time t’ - about channel $\mathbf{h_6}$?

$\hat{\mathbf{h}}_{6,t'}$ **t’ = 1,2,3…..**

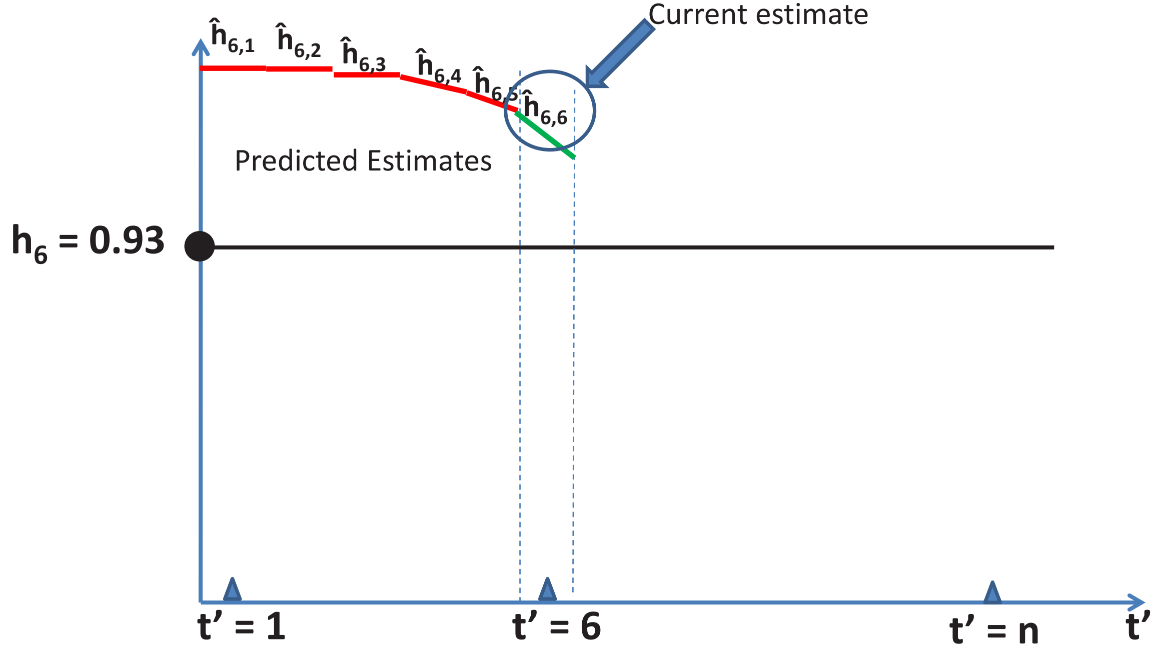


Knowledge at time t’ = 1,2,3…..

‘Current estimate’ of $\boldsymbol{h}_6$ at $t' = t = 6$

What do we know - at any point in time t’ - about channel $\mathbf{h}_6$?

$\hat{\mathbf{h}}_{6,t'}$ **t’ = 1,2,3.....**

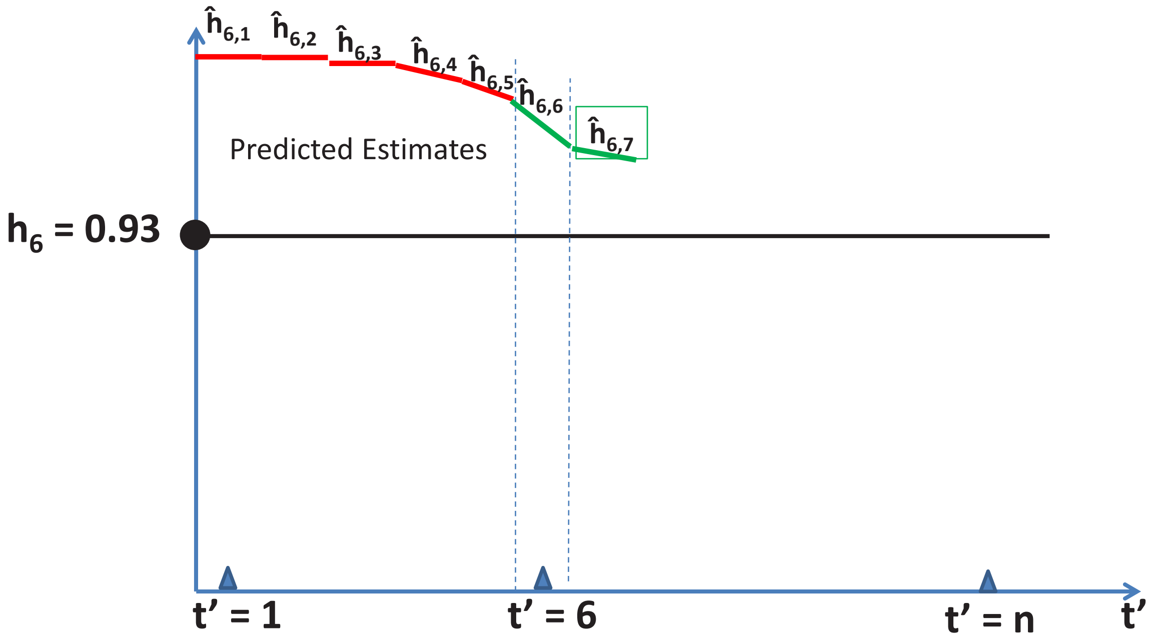


Knowledge at time t’ = 1,2,3.....

‘Delayed estimates’ at $t' > t = 6, \quad t' \leq n$

What do we know - at any point in time t’ - about channel $\mathbf{h_6}$?

$\hat{\mathbf{h}}_{6,t'}$ **t’ = 1,2,3…..**

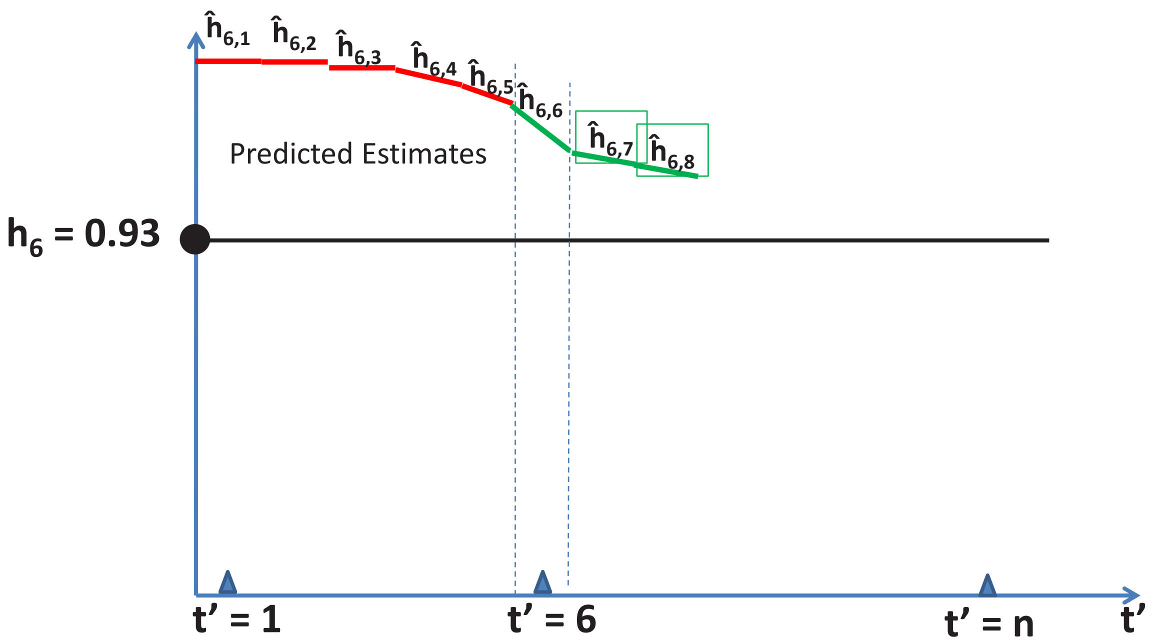


Knowledge at time t’ = 1,2,3…..

'Delayed estimate' at $t' > t = 6, \quad t' \leq n$

What do we know - at any point in time t' - about channel $\mathbf{h}_6$?

$\hat{\mathbf{h}}_{6,t'}$ **t' = 1,2,3.....**

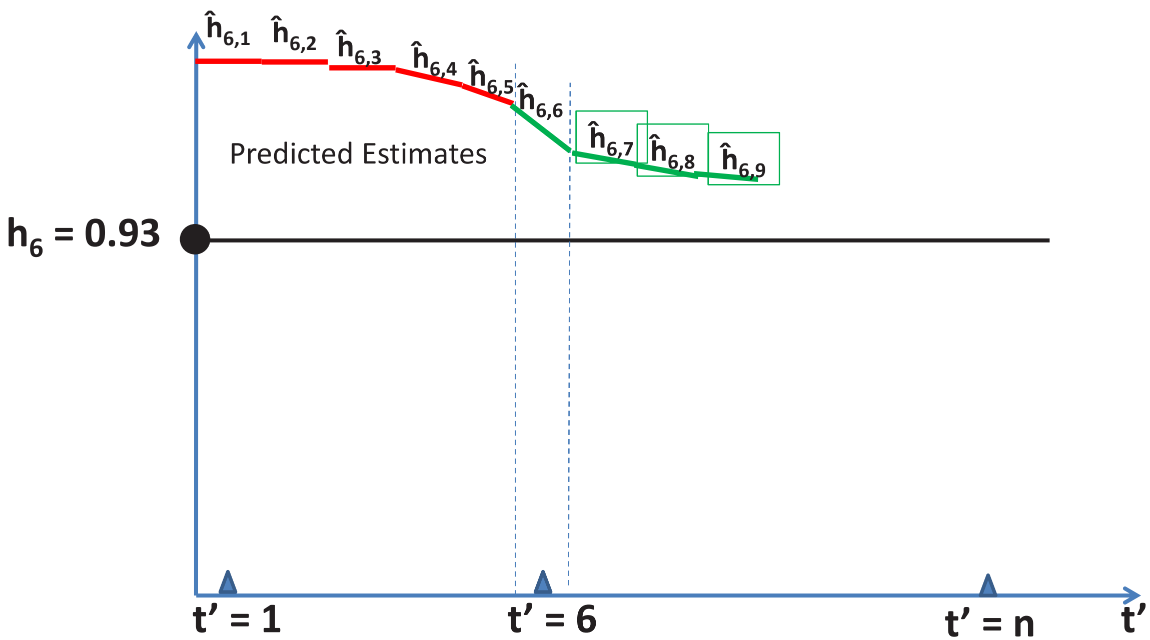


Knowledge at time t' = 1,2,3.....

'Delayed estimate' at $t' > t = 6, \quad t' \leq n$

What do we know - at any point in time t' - about channel $\mathbf{h}_6$?

$\hat{\mathbf{h}}_{6,t'}$ **t' = 1,2,3.....**

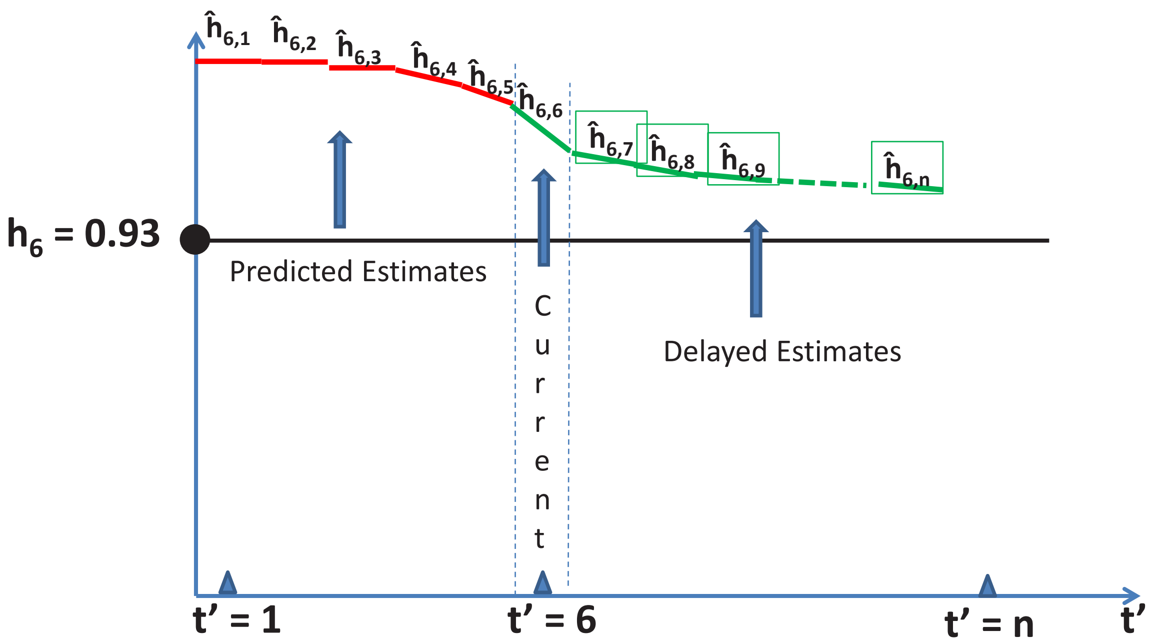


Knowledge at time t' = 1,2,3.....

'Delayed estimate' at $t' > t = 6, \quad t' \leq n$

What do we know - at any point in time t' - about channel $\mathbf{h}_6$?

$\hat{\mathbf{h}}_{6,t'}$ **t' = 1,2,3.....**

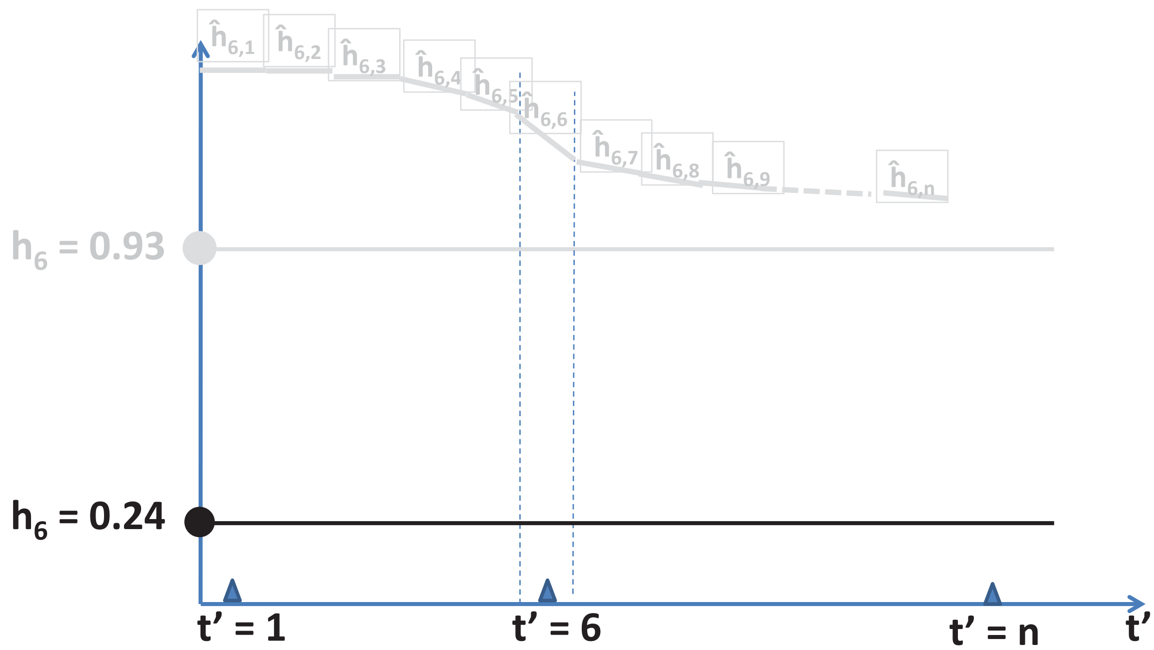


Knowledge at time t' = 1,2,3.....

And similarly another channel instance for $\boldsymbol{h}_6$

What do we know - at any point in time t' - about channel $\mathbf{h_6}$?

$\hat{\mathbf{h}}_{6,t'}$ **t' = 1,2,3.....**

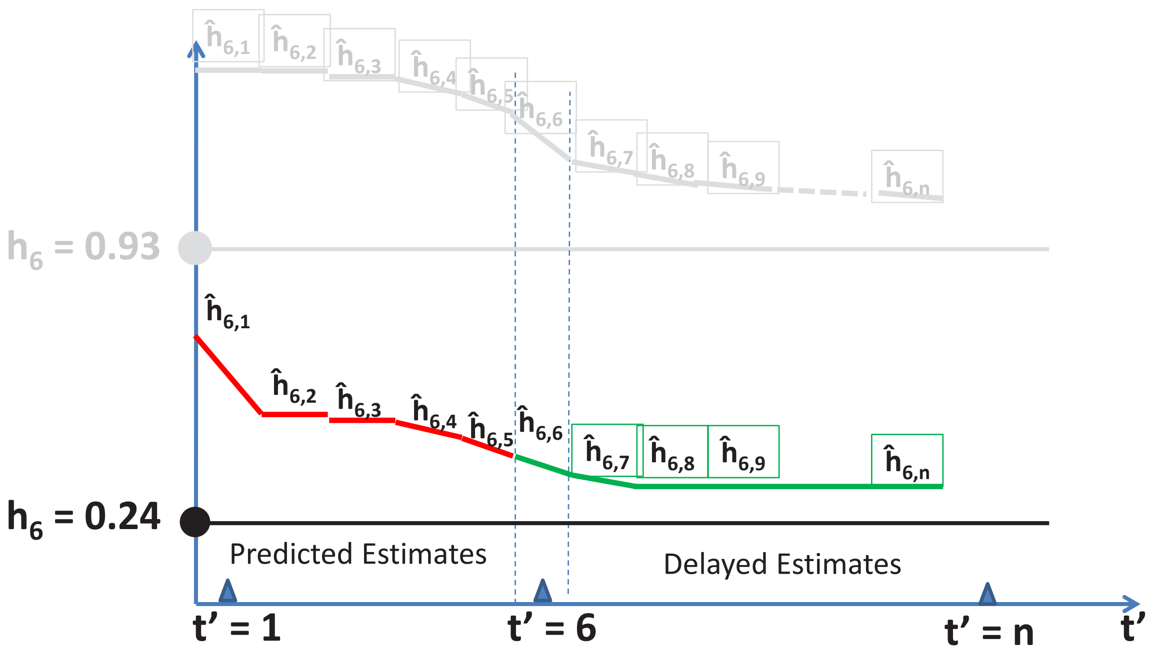


Knowledge at time t' = 1,2,3.....

And another CSIT estimate instance: $t' = 1 \to n$

What do we know - at any point in time t' - about channel $\mathbf{h}_6$?

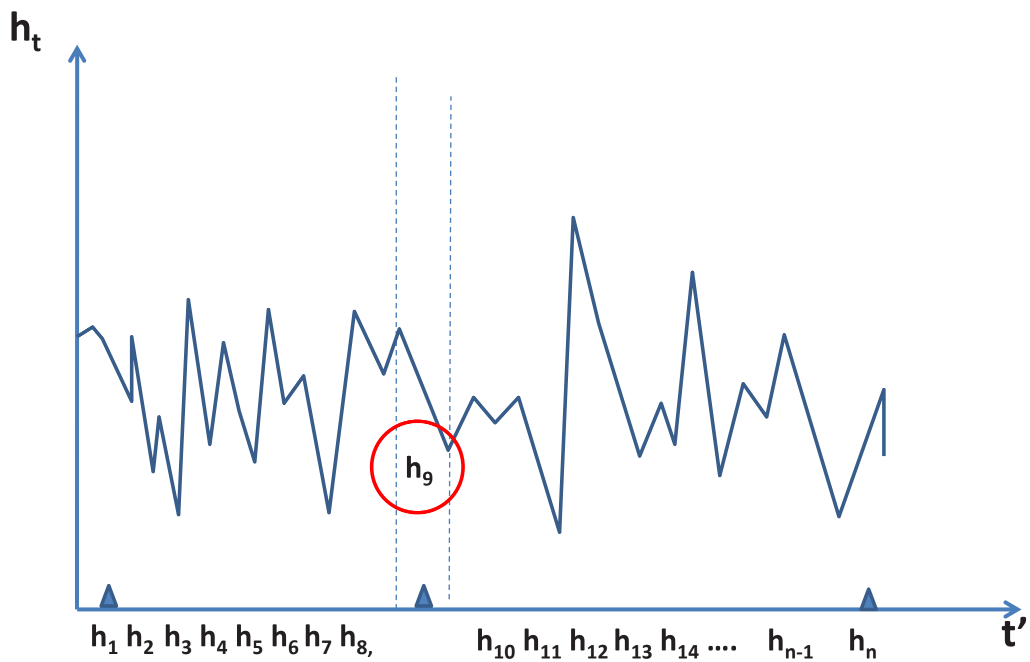
$\hat{\mathbf{h}}_{6,t'}$ **t' = 1,2,3.....**



Knowledge at time t' = 1,2,3.....

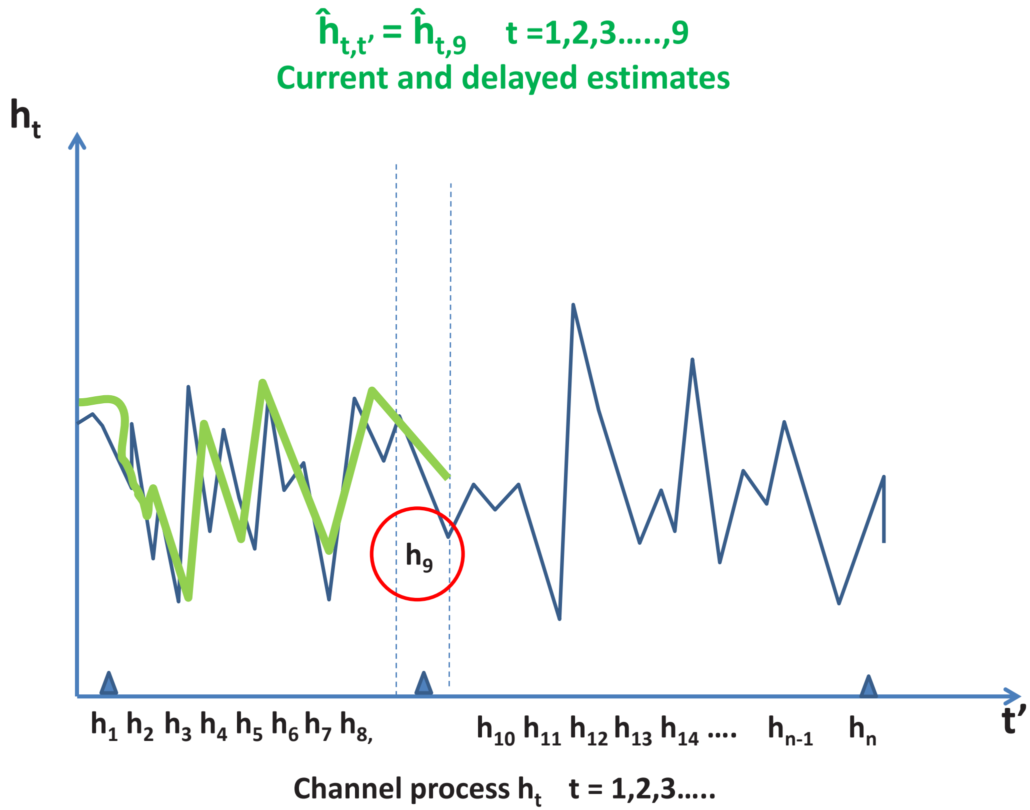
Yet another point of view - knowledge of channel process

What do we know at time t', about the channel process (say t'=9)



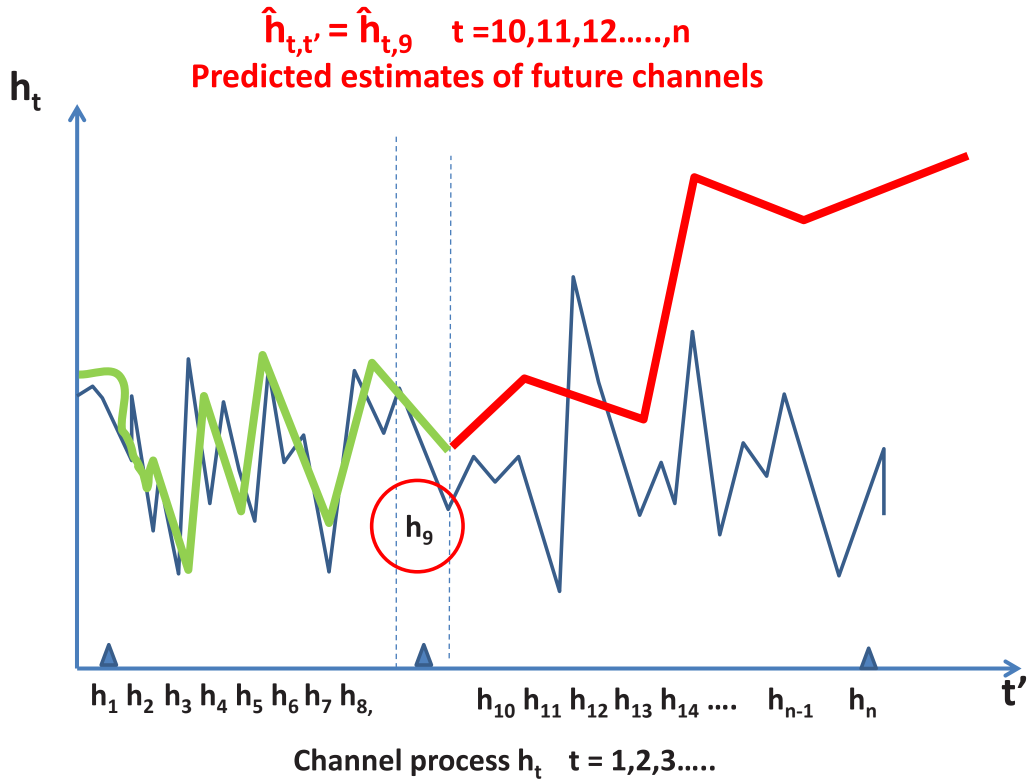
Channel process h_t t = 1,2,3.....

What do we know at time t', about the channel process (say t'=9)



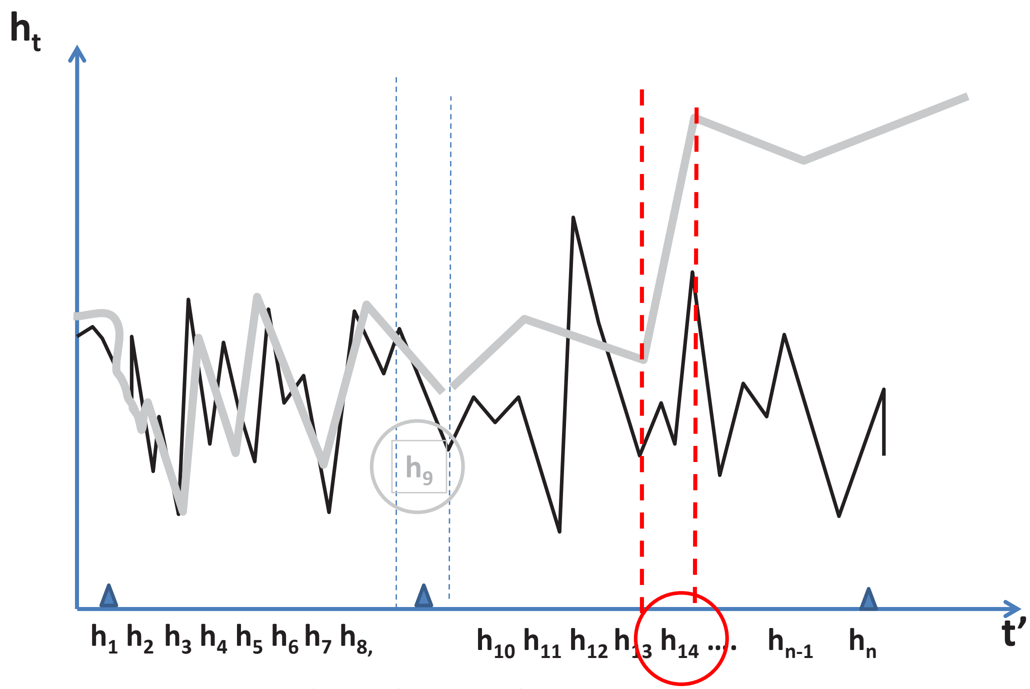
What we know at $t' = 9$, about future channels

What do we know at time t', about the channel process (say t'=9)



What is our knowledge at time $t' = 14$?

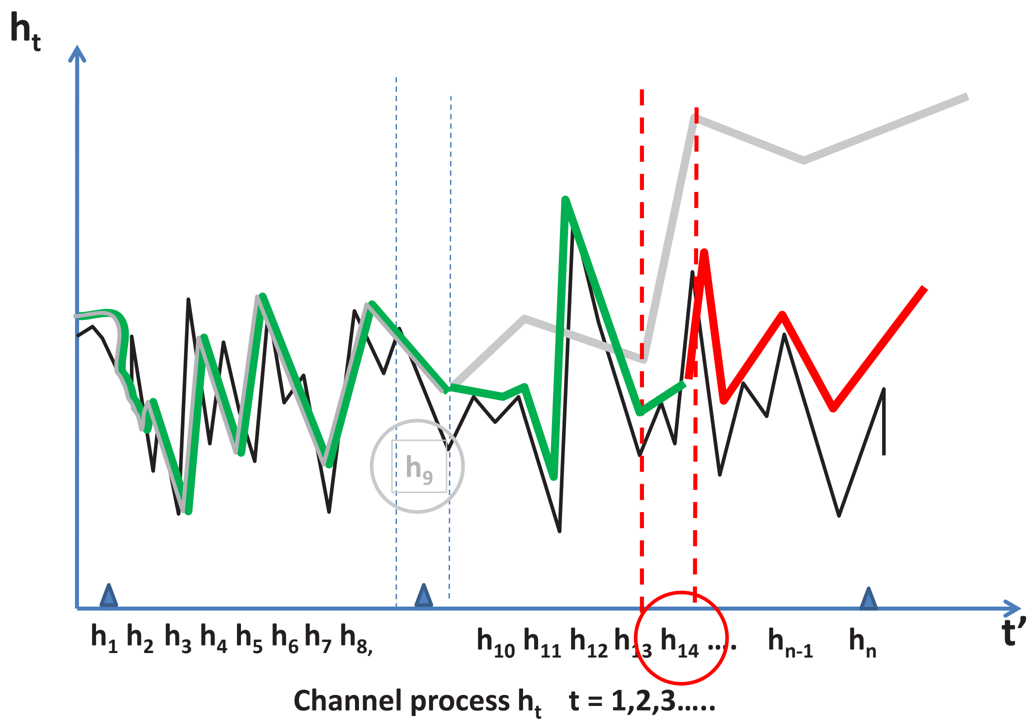
What do we know - at time t' = 14 – about the channel process?



Channel process h_t t = 1,2,3…..

Good for past, not so good for future

What do we know - at time t' = 14 – about the channel process?

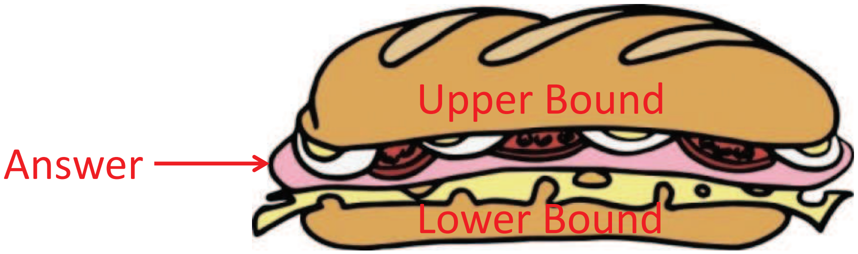


Tricks of the trade

Let us learn how to utilize different tools of the trade

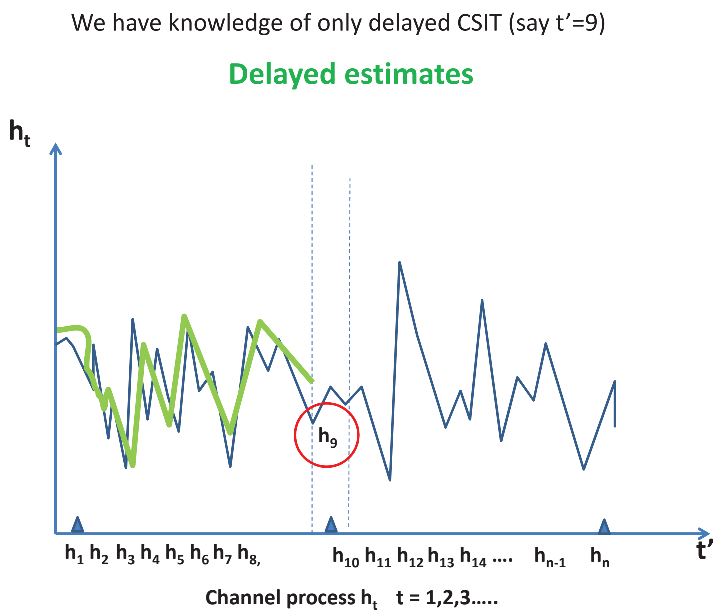
Answers in the form of:

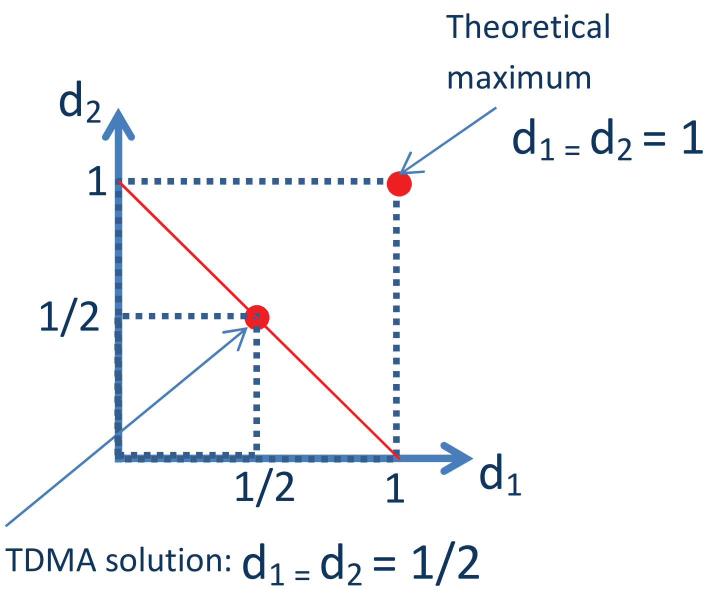
- Novel precoders/decoders that cleverly use feedback
- Information theoretic outer bounds (try to prove optimality)



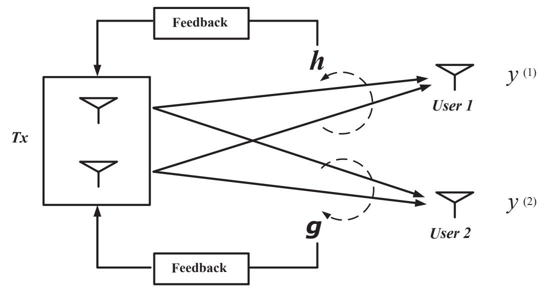
Delayed CSIT

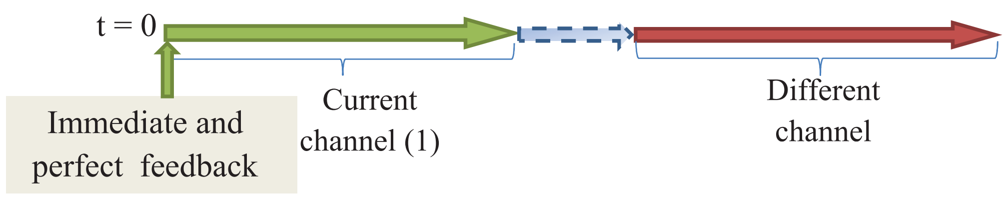
Tool: How to utilize delayed feedback?





Delayed vs. current CSIT in BLOCK FADING



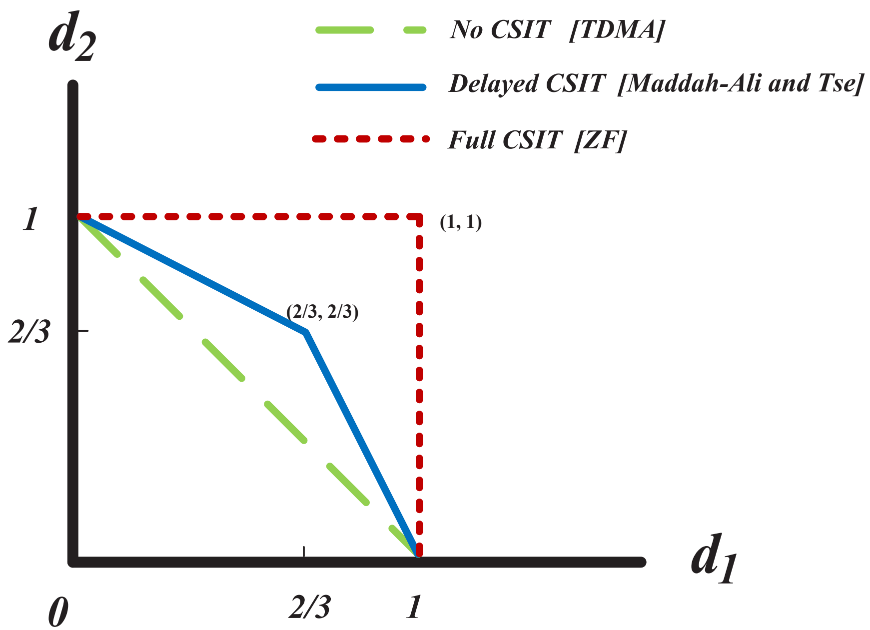


No current CSIT but perfect delayed CSIT

coherence block	1	2	3	4	$\cdots$
	$-$	$\boldsymbol{h}_1$	$\boldsymbol{h}_2$	$\boldsymbol{h}_3$	$\cdots$
	$-$	$\boldsymbol{g}_1$	$\boldsymbol{g}_2$	$\boldsymbol{g}_3$	$\cdots$

- *Theorem (Maddah-Ali and Tse): Optimal DoF*

$$d_1 = d_2 = 2/3$$



Maddah-Ali and Tse (MAT) scheme

- Tx sends symbols a_1, a_2 for user 1, and b_1, b_2 for user 2, in 3 channel uses
  - ⋆ WOLOG consider $T_{\text{coh}} = 1$ (unit coherence period)
  - ⋆ Duration $T = 3$: Tx sequentially sends vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3 \in \mathbb{C}^2$

- In the first two channel uses:

$$t = 1: \ \boldsymbol{x}_1 = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \qquad \begin{aligned} y_1^{(1)} &= \boldsymbol{h}_1^\top \boldsymbol{x}_1 + \text{noise} \\ y_1^{(2)} &= \boldsymbol{g}_1^\top \boldsymbol{x}_1 + \text{noise} \end{aligned}$$

$$t = 2: \ \boldsymbol{x}_2 = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \qquad \begin{aligned} y_2^{(1)} &= \boldsymbol{h}_2^\top \boldsymbol{x}_2 + \text{noise} \\ y_2^{(2)} &= \boldsymbol{g}_2^\top \boldsymbol{x}_2 + \text{noise} \end{aligned}$$

Maddah-Ali and Tse (MAT) scheme[1]

- Now - with delayed CSIT - Tx reconstructs $\boldsymbol{g}_1^\top \boldsymbol{x}_1$ and $\boldsymbol{h}_2^\top \boldsymbol{x}_2$

$$t = 3: \ \boldsymbol{x}_3 = \begin{bmatrix} \boldsymbol{h}_2^\top \boldsymbol{x}_2 + \boldsymbol{g}_1^\top \boldsymbol{x}_1 \\ 0 \end{bmatrix}, \quad \begin{array}{l} y_3^{(1)}/h_{3,1} = \boldsymbol{h}_2^\top \boldsymbol{x}_2 + \boldsymbol{g}_1^\top \boldsymbol{x}_1 + \text{noise} \\ y_3^{(2)}/g_{3,1} = \boldsymbol{h}_2^\top \boldsymbol{x}_2 + \boldsymbol{g}_1^\top \boldsymbol{x}_1 + \text{noise} \end{array}$$

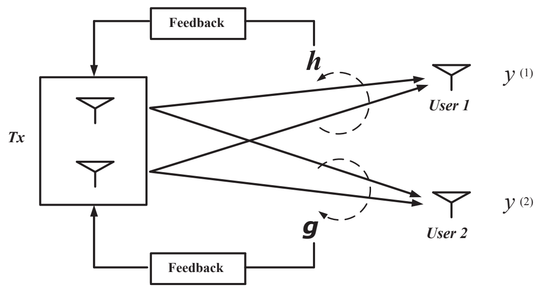
$$\tilde{\boldsymbol{y}}^{(1)} \triangleq \begin{bmatrix} y_1^{(1)} \\ y_3^{(1)}/h_{3,1} - y_2^{(1)} \end{bmatrix} = \underbrace{\begin{bmatrix} \boldsymbol{h}_1^\top \\ \boldsymbol{g}_1^\top \end{bmatrix}}_{2\times 2 \text{ MIMO}} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \text{noise}$$

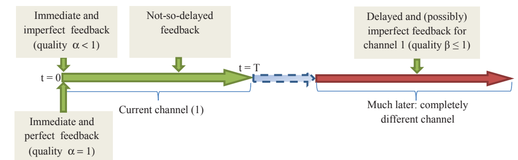
- Each user decodes two symbols in three timeslots: $d_1 = d_2 = 2/3$
- Scheme shown to be DoF optimal
- Insight: retrospective interference alignment in space and time, using delayed CSIT
  - ⋆ a.k.a. do the damage now, and fix it later

Feedback asymmetry: one user has more feedback

Tool: dealing with feedback asymmetry:
one user has more feedback

One user has more feedback: Maleki, Jafar and Shamai



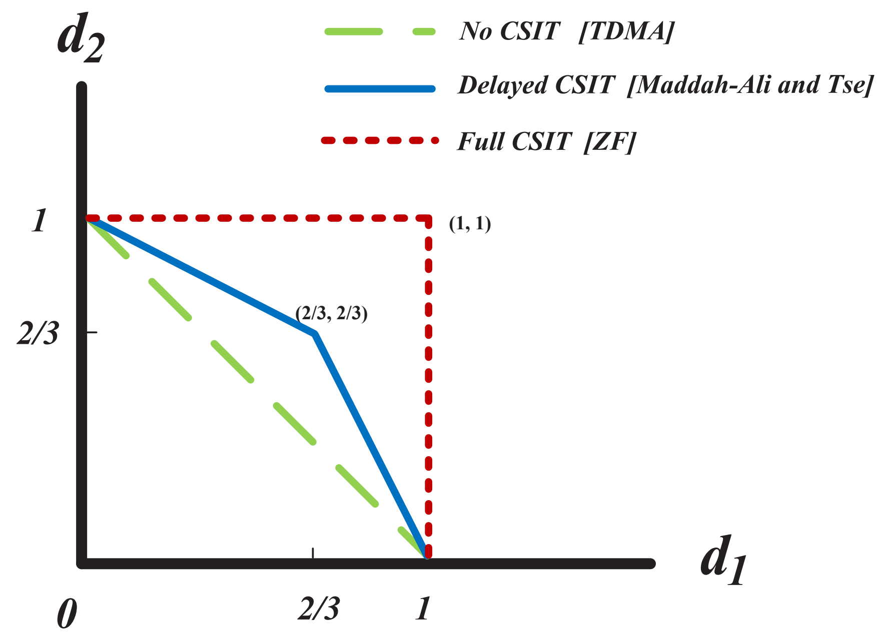


- Current CSIT for $\boldsymbol{h}_t$ (of 1st user): Perfectly and instantly known at Tx
- Delayed CSIT for $\boldsymbol{g}_t$ (of 2nd user): Perfectly known to Tx after coherence period passes

coherence block	1	2	3	4	$\cdots$
	$\boldsymbol{h}_1$	$\boldsymbol{h}_2$	$\boldsymbol{h}_3$	$\boldsymbol{h}_4$	$\cdots$
	—	$\boldsymbol{g}_1$	$\boldsymbol{g}_2$	$\boldsymbol{g}_3$	$\cdots$

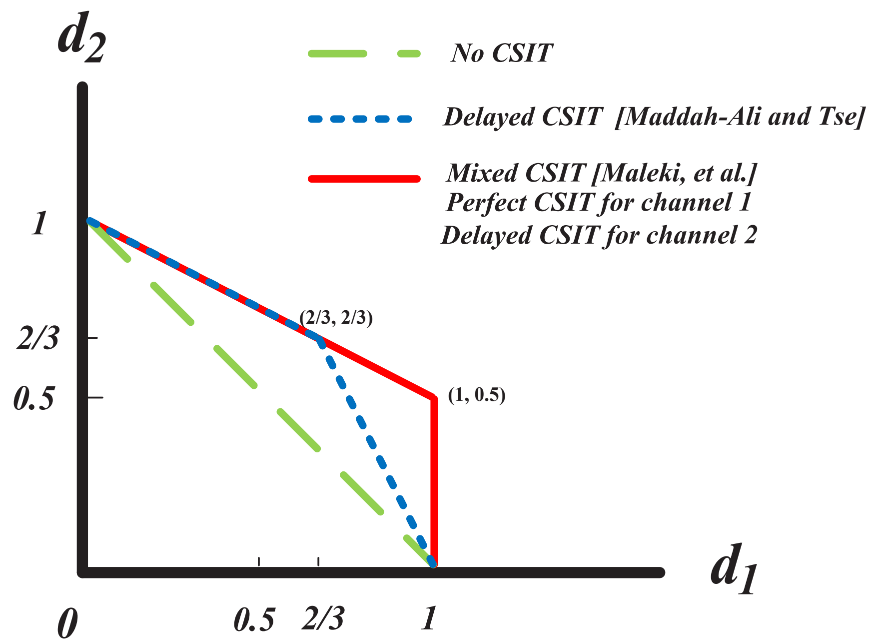
One user has more feedback: Maleki, Jafar and Shamai[1]

- Recall: if both users only had delayed feedback



One user has more feedback: Maleki, Jafar and Shamai[2]

- Now One user has delayed, the other had perfect
- *Theorem: Derived optimal DoF is* $d_1 = 1,\ d_2 = 1/2,$ *(sum DoF* $3/2 \geq 4/3$*)*



$d_1 = 1,\ d_2 = 1/2,$ (sum DoF $3/2$)

Maleki et al.: asymmetric scheme

- Tx sends symbols a_1, a_2 for user 1, and b for user 2, in 2 channel uses
  - ⋆ WOLOG: one channel use $=$ one coherence block
  - ⋆ Tx will sequentially send signal vectors $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{C}^2$
  - ⋆ note use of symbol $\perp \rightarrow$ (orthogonal)

$$\boldsymbol{x}_1 = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \boldsymbol{h}_1^{\perp} b, \quad \boldsymbol{x}_2 = \begin{bmatrix} \boldsymbol{g}_1^{\top} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \\ 0 \end{bmatrix} + \boldsymbol{h}_2^{\perp} b$$

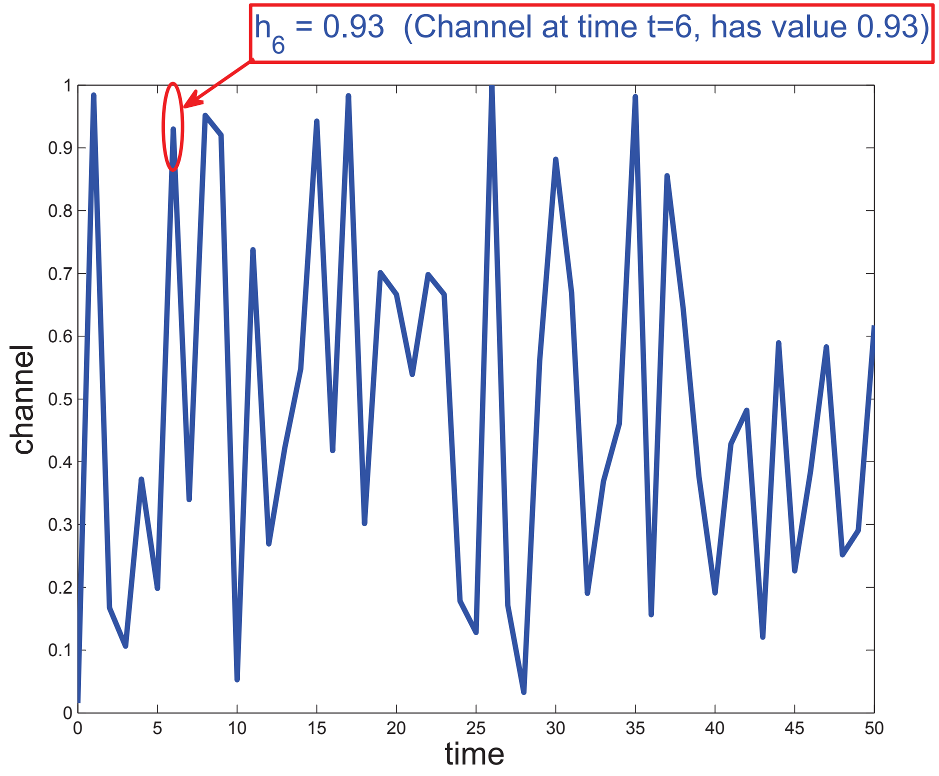
- Intuitions:
  - ⋆ Current CSIT can be used for instantaneous interference mitigation
  - ⋆ Delayed CSIT can be used for retrospective interference cancelation

Introducing feedback QUALITY considerations

Introducing feedback QUALITY considerations

Tool: how to exploit partial-feedback?

Introducing feedback QUALITY considerations$_1$

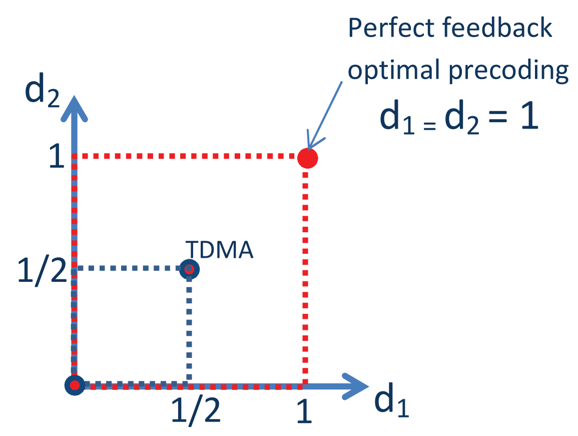


Introducing feedback QUALITY considerations$_2$

- Jindal et al., Caire et al: $\approx$ *"Optimal DoF does not need infinite number of feedback bits*
  - ⋆ Let $\hat{\boldsymbol{h}}_t$ be the <u>INSTANTANEOUS</u> estimate of channel $\boldsymbol{h}_t$
  - ⋆ Let $\hat{\boldsymbol{g}}_t$ be the <u>INSTANTANEOUS</u> estimate of channel $\boldsymbol{g}_t$
  - ⋆ Then if

$$\mathbb{E}[||\hat{\boldsymbol{h}}_t - \boldsymbol{h}_t||^2] \approx P^{-1}, \qquad \mathbb{E}[||\hat{\boldsymbol{g}}_t - \boldsymbol{g}_t||^2] \approx P^{-1}$$

  - ⋆ you can achieve the optimal DoF



Refining quality considerations

- Motivation: Note $\mathbb{E}[\|\hat{\boldsymbol{h}}_t - \boldsymbol{h}_t\|^2] \approx P^{-1}$ corresponds to sending about $\log P$ bits of feedback per scalar (rate distortion theory - not optimal)

- What if you cannot send so many bits?

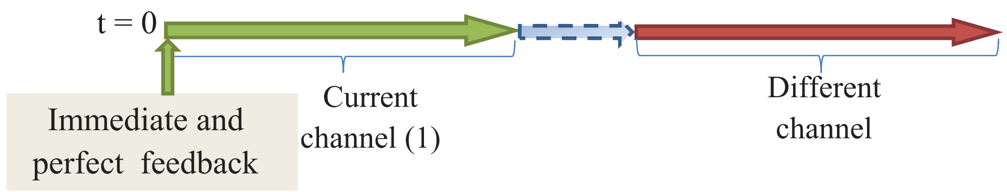
Kobayashi-Yang-Yi-Gesbert:
Current CSIT estimation errors with power $P^{-\alpha}$

- Current CSIT quality exponent

$$\alpha = -\lim \frac{\log \mathbb{E}[\|\hat{\boldsymbol{h}}_t - \boldsymbol{h}_t\|^2]}{\log P} = -\lim \frac{\log \mathbb{E}[\|\hat{\boldsymbol{g}}_t - \boldsymbol{g}_t\|^2]}{\log P}, \qquad \alpha : 0 \to 1$$

Combining current and delayed CSIT (Yang-Gesbert et al.)

- Perfect delayed CSIT + imperfect current CSIT



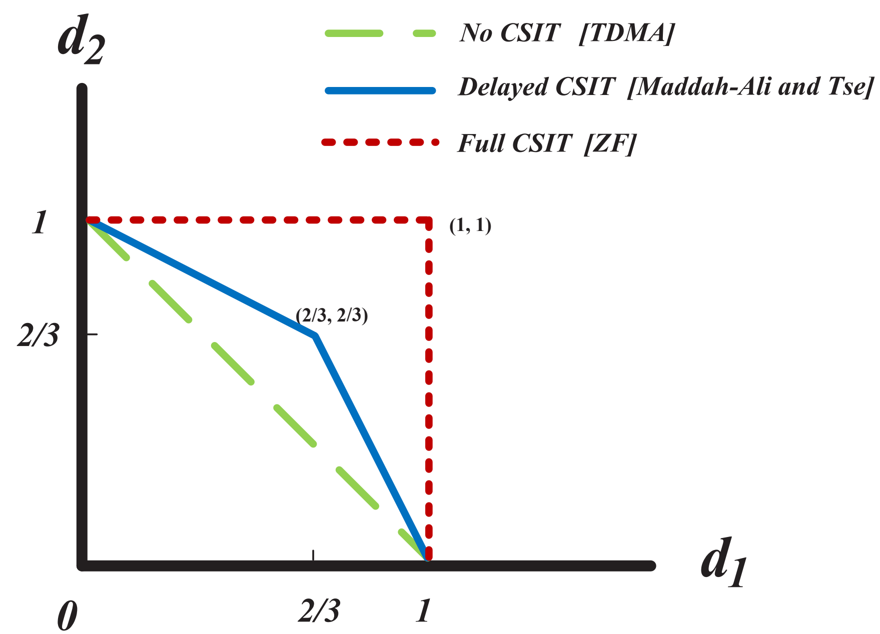
Coherence block	1	2	3	4	$\cdots$
Current estimates (quality α)	$\widehat{\boldsymbol{h}}_1, \widehat{\boldsymbol{g}}_1$	$\widehat{\boldsymbol{h}}_2, \widehat{\boldsymbol{g}}_2$	$\widehat{\boldsymbol{h}}_3, \widehat{\boldsymbol{g}}_3$	$\widehat{\boldsymbol{h}}_4, \widehat{\boldsymbol{g}}_4$	$\cdots$
Delayed estimates (exact) $\rightarrow$		$\boldsymbol{h}_1, \boldsymbol{g}_1$	$\boldsymbol{h}_2, \boldsymbol{g}_2$	$\boldsymbol{h}_3, \boldsymbol{g}_3$	$\boldsymbol{h}_4, \boldsymbol{g}_3$

- Current CSIT: PARTIAL instantaneous interference mitigation
- Delayed CSIT: retrospective interference management, at later time

Combining current and delayed CSIT (Yang-Gesbert et al.)[1]

RECALL: IF BOTH USERS ONLY HAD DELAYED FEEDBACK

$(\Rightarrow \alpha = 0)$

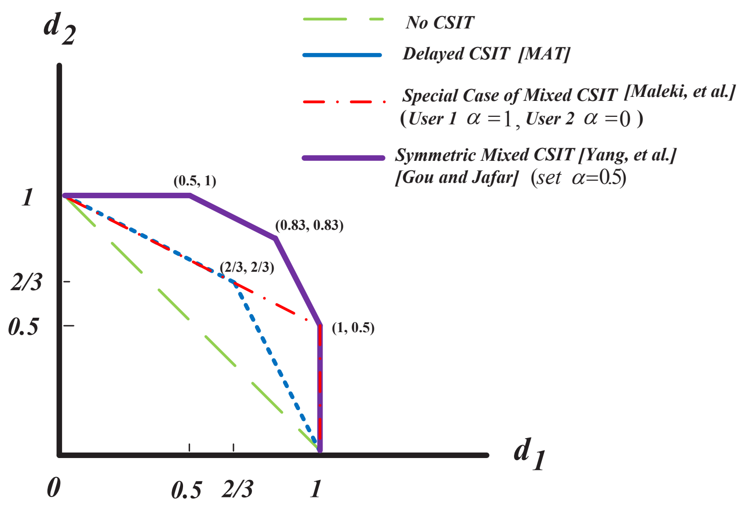


Perfect delayed, and imperfect current CSIT

Now each has delayed + imperfect current estimates
($\Rightarrow \alpha > 0$)

- *Theorem*[1]*: Perfect delayed CSIT and α-quality current CSIT, gives:*

$$d_1 = d_2 = \frac{2+\alpha}{3}$$



[1] Yang-Kobayashi-Yi-Gesbert, Gou-Jafar 2012

Perfect delayed, and imperfect current CSIT_1

- During phase 1 ($t = 1$), the transmitter sends ($\boldsymbol{u}_1 = \hat{\boldsymbol{g}}_1^{\perp},\ \boldsymbol{v}_1 = \hat{\boldsymbol{h}}_1^{\perp}$)

$$\boldsymbol{x}_1 = \underbrace{\overbrace{\boldsymbol{u}_1}^{\hat{\boldsymbol{g}}_1^{\perp}} a_1}_{\text{power } P,\ \text{rate prelog } r=1} + \underbrace{\overbrace{\boldsymbol{v}_1}^{\hat{\boldsymbol{h}}_1^{\perp}} b_1}_{P,\ r=1} + \underbrace{\overbrace{\boldsymbol{u}_1^{'}}^{\text{random}} a_1^{'}}_{P^{1-\alpha},\ r=1-\alpha} + \underbrace{\overbrace{\boldsymbol{v}_1^{'}}^{\text{random}} b_1^{'}}_{P^{1-\alpha},\ r=1-\alpha}$$

- Users receive

$$y_1^{(1)} = \boldsymbol{h}_1^{\mathsf{T}}\boldsymbol{u}_1 a_1 + \boldsymbol{h}_1^{\mathsf{T}}\boldsymbol{u}_1^{'} a_1^{'} + \overbrace{\underbrace{\tilde{\boldsymbol{h}}_1^{\mathsf{T}}\boldsymbol{v}_1 b_1 + \boldsymbol{h}_1^{\mathsf{T}}\boldsymbol{v}_1^{'} b_1^{'}}_{\text{power } P^{1-\alpha}}}^{\text{interference } \iota_1^{(1)}} + \text{noise},$$

$$y_1^{(2)} = \overbrace{\underbrace{\tilde{\boldsymbol{g}}_1^{\mathsf{T}}\boldsymbol{u}_1 a_1 + \boldsymbol{g}_1^{\mathsf{T}}\boldsymbol{u}_1^{'} a_1^{'}}_{\text{power } P^{1-\alpha}}}^{\text{interference } \iota_1^{(2)}} + \boldsymbol{g}_1^{\mathsf{T}}\boldsymbol{v}_1 b_1 + \boldsymbol{g}_1^{\mathsf{T}}\boldsymbol{v}_1^{'} b_1^{'} + \text{noise}.$$

Perfect delayed, and imperfect current CSIT$_2$

- At the end of phase 1. Reconstruct and quantize interference using delayed CSIT

$$\iota_1^{(1)} = \tilde{\boldsymbol{h}}_1^{\mathsf{T}} \boldsymbol{v}_1 b_1 + \boldsymbol{h}_1^{\mathsf{T}} \boldsymbol{v}_1' b_1', \quad \iota_1^{(2)} = \tilde{\boldsymbol{g}}_1^{\mathsf{T}} \boldsymbol{u}_1 a_1 + \boldsymbol{g}_1^{\mathsf{T}} \boldsymbol{u}_1' a_1'$$

- Phase 2, $t = 2, 3$, Tx sends c_t and extra a_t, b_t

$$\boldsymbol{x}_t = \underbrace{\boldsymbol{w}_t c_t}_{P,\quad r=1-\alpha} + \underbrace{\hat{\boldsymbol{g}}_t^{\perp} a_t}_{P^{\alpha},\quad r=\alpha} + \underbrace{\hat{\boldsymbol{h}}_t^{\perp} b_t}_{P^{\alpha},\quad r=\alpha}$$

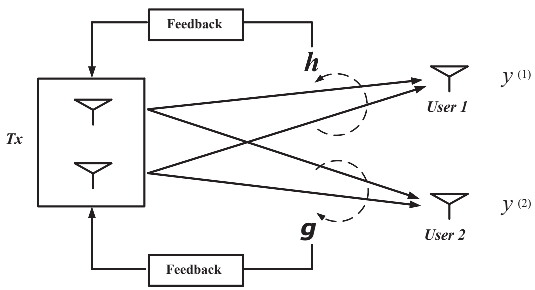
  - ⋆ Successive decoding: $c_t \to a_t$ at user 1, $c_t \to b_t$ at user 2
  - ⋆ Reconstructing approximate interference:$\{c_t\}_{t=2}^{3} \to \{\bar{\iota}_1^{(i)}\}_{t=1}^{2}$
  - ⋆ Go back to phase 1, and decode a_1, a_2 at user 1, and b_1, b_2 at user 2

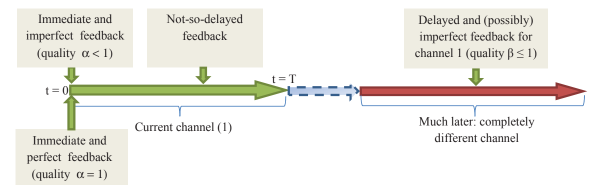
$$\begin{bmatrix} y_1^{(1)} - \bar{\iota}_1^{(1)} \\ \bar{\iota}_1^{(2)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}_1^{\mathsf{T}} \\ \boldsymbol{g}_1^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_1' \end{bmatrix} \begin{bmatrix} a_1 \\ a_1' \end{bmatrix} + \text{noise}$$

$$d_1 = d_2 = \frac{2+\alpha}{3}$$

Alternating CSIT - tool for offering symmetry

Alternating CSIT[2]
Feedback alternates from user to user





Time t	***1***	***2***	***3***	***4***	***5***	***6***	***7***	...
CSIT of channel h	***P***	***D***	***N***	***P***	***P***	***N***	***N***	...
CSIT of channel g	***D***	***P***	***N***	***N***	***N***	***P***	***P***	...

[2] Tandon-Jafar-Shamai-Poor 2012

Alternating CSIT - tool for offering symmetry[1]

- CSIT for each user's channel, at a specific time, can be either perfect (P), delayed (D) or not available (N)
  - ⋆ $I_1, I_2 \in \{P, D, N\}$
  - ⋆ For example, in a specific time: $I_1 = P,\ I_2 = D$
- $\lambda_{I_1 I_2}$ is the fraction of time associated with CSIT states I_1, I_2
  - ⋆ Symmetric assumption $\lambda_{I_1 I_2} = \lambda_{I_2 I_1}$
- $\lambda_P = \lambda_{PP} + \lambda_{PD} + \lambda_{PN}$
- $\lambda_D = \lambda_{DP} + \lambda_{DD} + \lambda_{DN}$
- *Theorem: Derived DoF*

$$d = \min\{\frac{2+\lambda_P}{3}, \frac{1+\lambda_P+\lambda_D}{2}\}$$

Symmetry gains

- Asymmetry: $\lambda_{PD} = 1 \quad \Rightarrow \quad d_1 + d_2 = 3/2$ (Maleki et al.)
  - ⋆ Instantaneous perfect CSIT for channel of user 1 $I_1 = P$
  - ⋆ Delayed CSIT for channel of user 2 $I_2 = D$

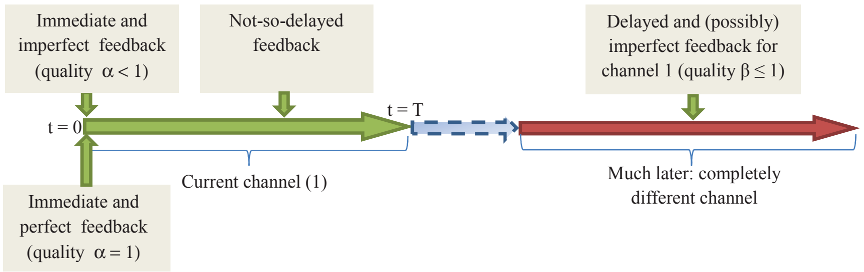
- Symmetry: $\lambda_{PD} = 0.5, \lambda_{DP} = 0.5$

$$\Rightarrow \quad d_1 + d_2 = 5/3 \geq 3/2$$

  - ⋆ Half of time $I_1 = P,\ I_2 = D$, other half $I_1 = D,\ I_2 = P$

- Same feedback cost, but symmetric provides gain $5/3 - 3/2$

Summary: Part 1-A



- No CSIT
- Perfect CSIT (precoding)
- Delayed CSIT-MAT (retrospective interference cancellation)
- Dealing with uneven feedback (Maleki et al.)
- Exploiting delayed and imperfect-quality current CSIT Yang et al. and Gou and Jafar
- Alternating CSIT to create symmetry (Tandon et al.)

Common themes of what we have seen

- Motivated by timeliness-and-quality considerations
- Timeliness and quality might be hard to get over limited feedback links
- Timeliness and quality affect performance
  - ⋆ Feedback delays and imperfections generally reduce performance
- A corresponding clear delay-and-quality question....

The fundamental question

How much feedback is necessary, and when, in order to achieve a certain performance?

Answering a broad range of practical questions

"Answering a broad range of practical performance-vs-feedback questions, up to a sublogarithmic *factor of P"*

WHAT WOULD ENGINEERS ASK?

- What is the role of MIMO in reducing feedback quality?
- When is delayed feedback necessary?
- When is predicted feedback necessary?
- What is better: less feedback early, or more feedback later?
- How to exploit feedback of imperfect quality?
- How to exploit feedback with predictions?
- How to exploit feedback with delayed information?
- How much feedback, where, and when, for a certain performance?

A unified performance-vs-feedback framework

Fundamental formulation of performance-vs-feedback problem

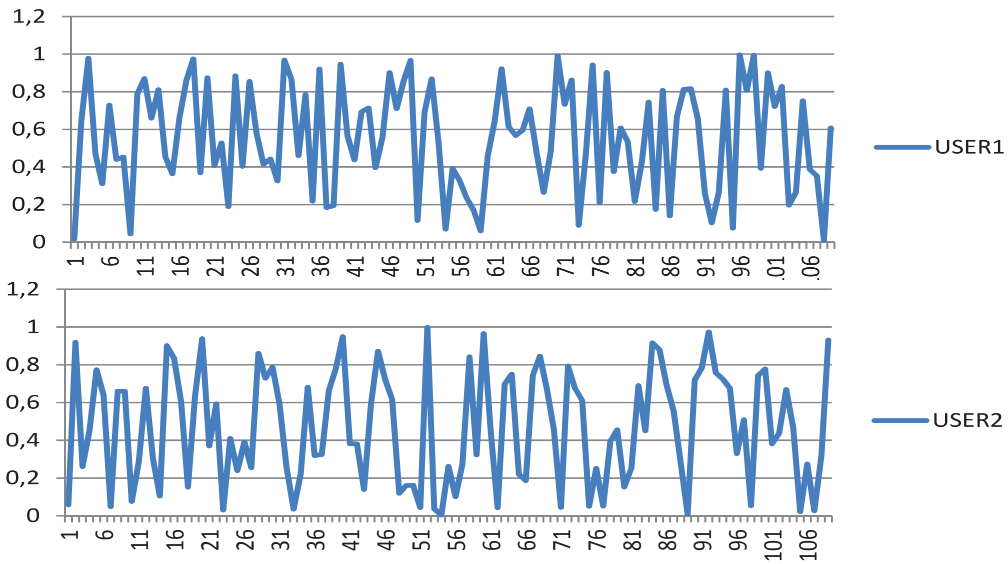
Fundamental formulation:step 1

STEP 1: COMMUNICATION OF DURATION n (n IS LARGE)

Fundamental formulation:step 2

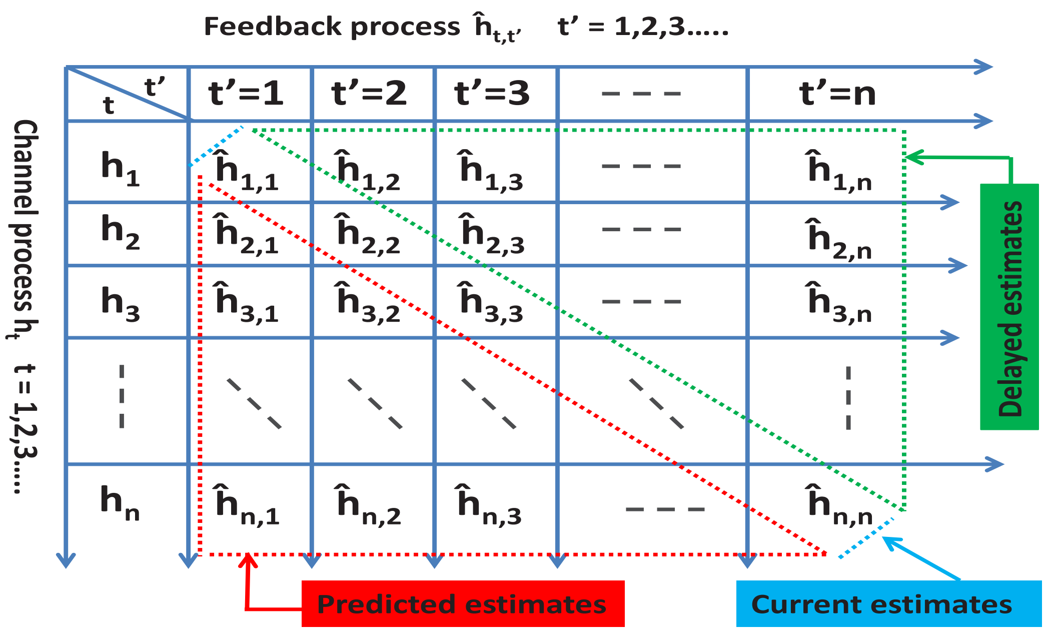
STEP 2: COMMUNICATION ENCOUNTERS AN ARBITRARY CHANNEL PROCESS

$$\begin{array}{lcccccc} \text{user } 1: & \boldsymbol{h}_1 & \boldsymbol{h}_2 & \boldsymbol{h}_3 & \cdots & \boldsymbol{h}_n \\ \text{user } 2: & \boldsymbol{g}_1 & \boldsymbol{g}_2 & \boldsymbol{g}_3 & \cdots & \boldsymbol{g}_n \end{array}$$



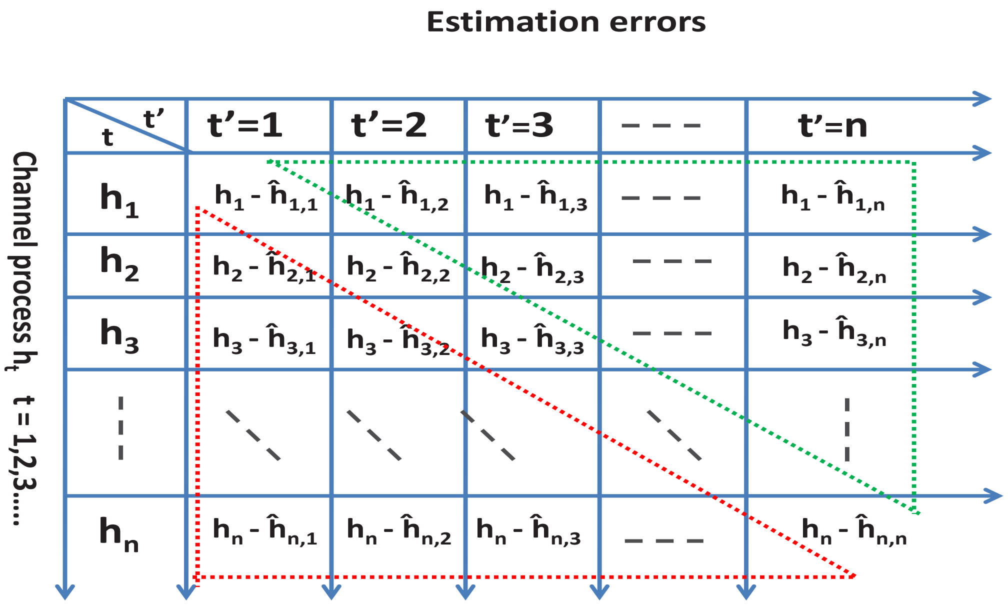
Step 3: An arbitrary feedback process

What do we know - at any time t'– about any channel $\mathbf{h}_t$?

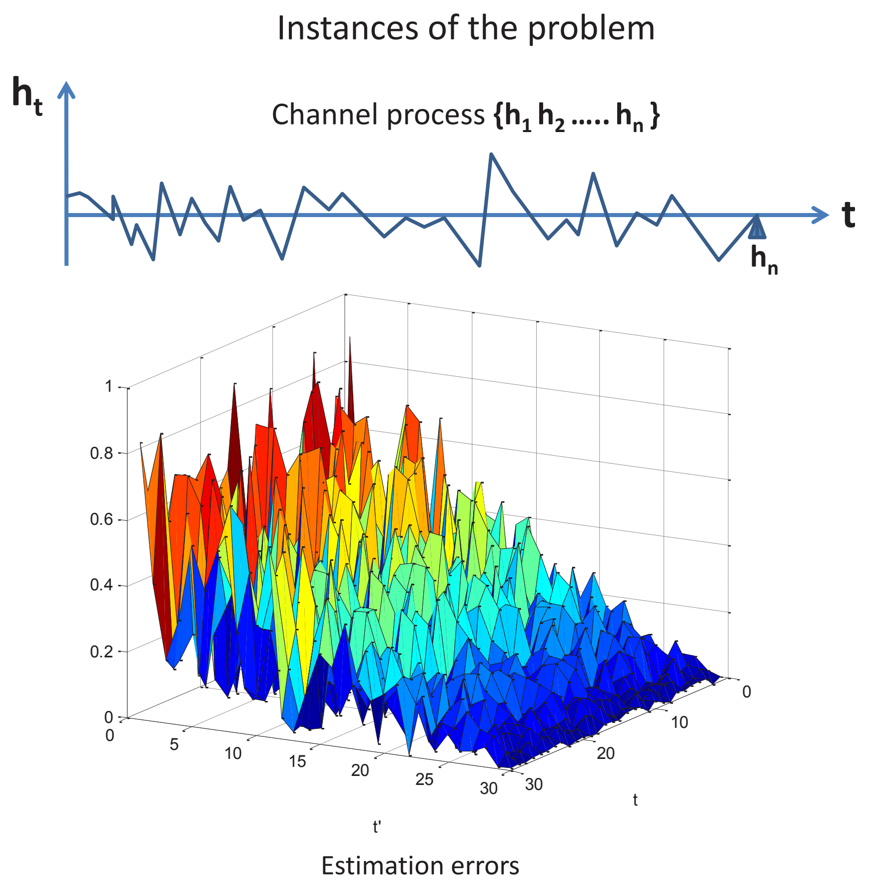


Fundamental formulation:step 4

Step 4: A 'primitive' measure of feedback 'goodness'

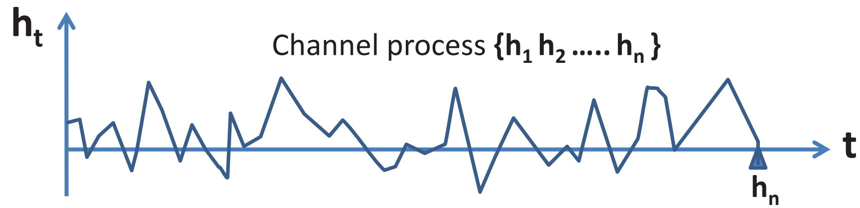


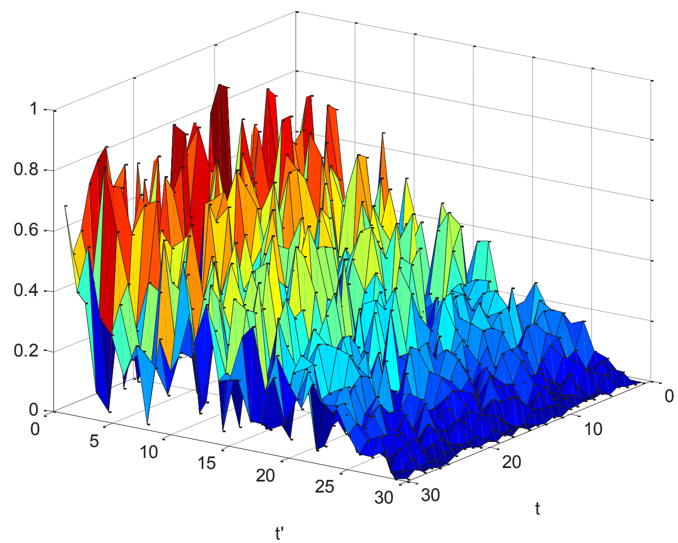
Remember the problem is random



Remember the problem is random$_1$

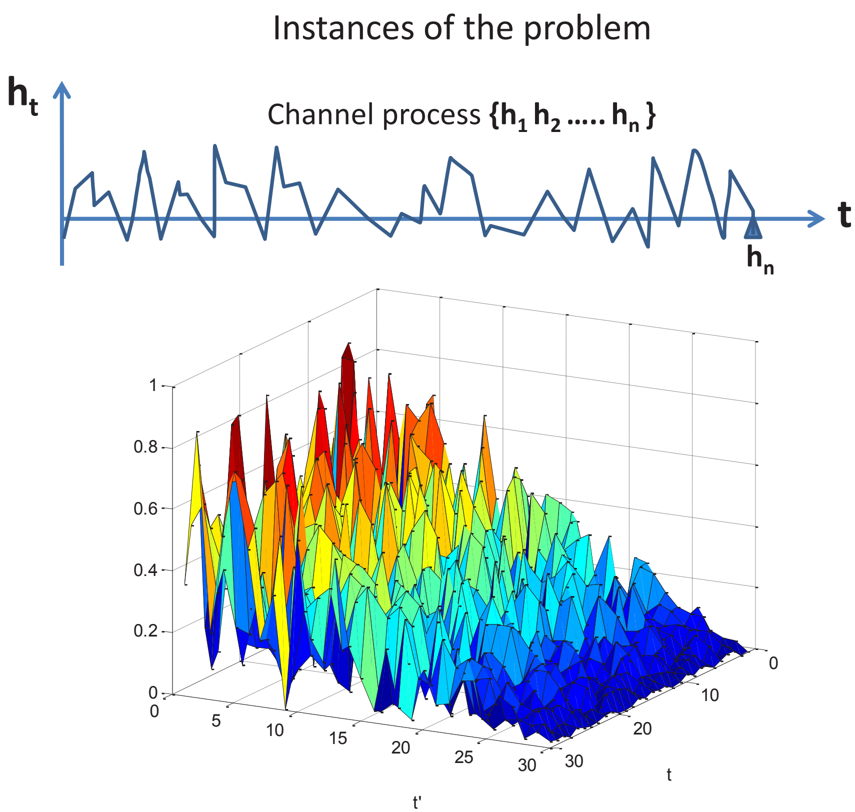
Instances of the problem





Estimation errors

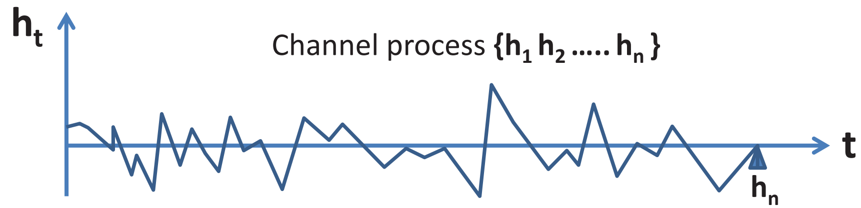
Remember the problem is random$_2$

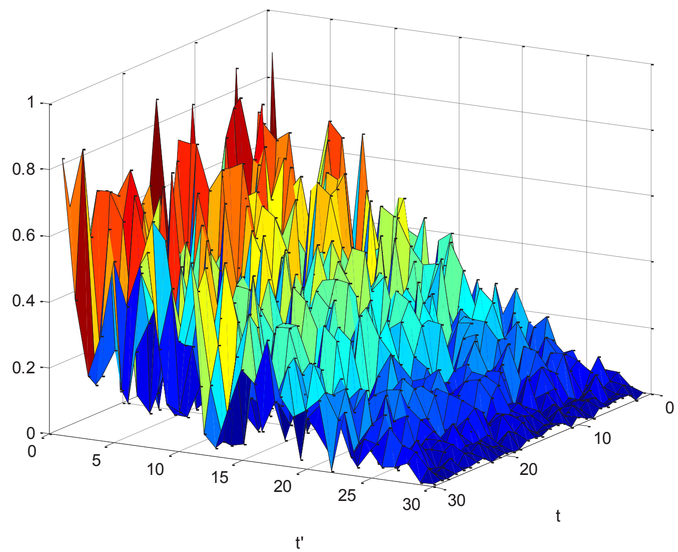


Estimation errors

Remember the problem is random$_3$

Instances of the problem

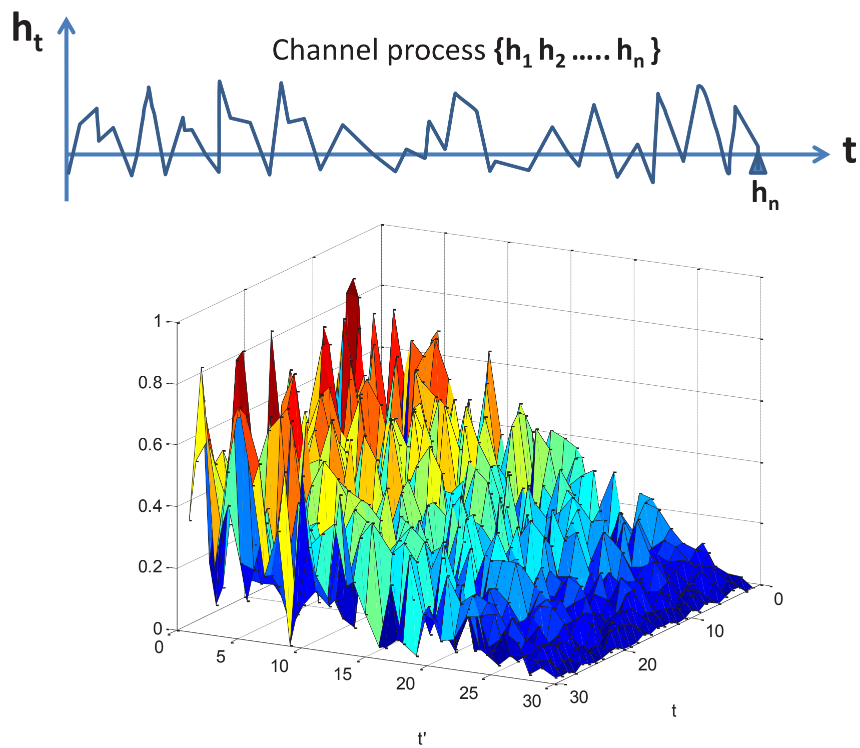




Estimation errors

Remember the problem is random$_4$

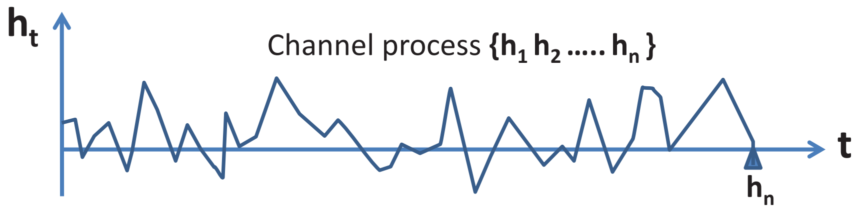
Instances of the problem

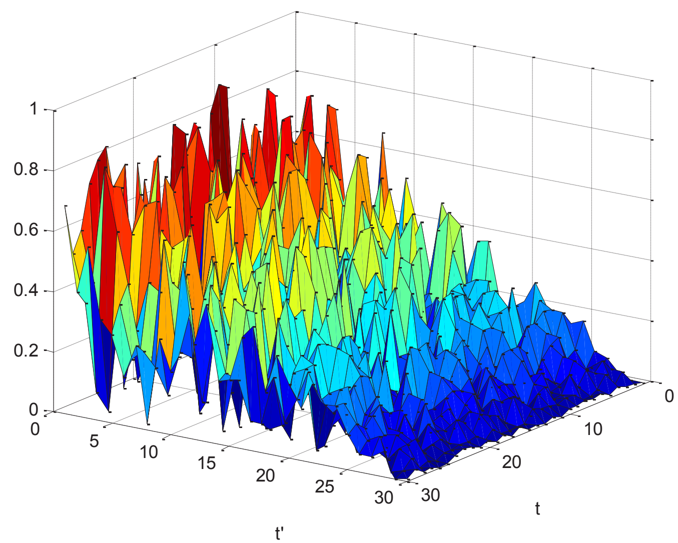


Estimation errors

Remember the problem is random$_5$

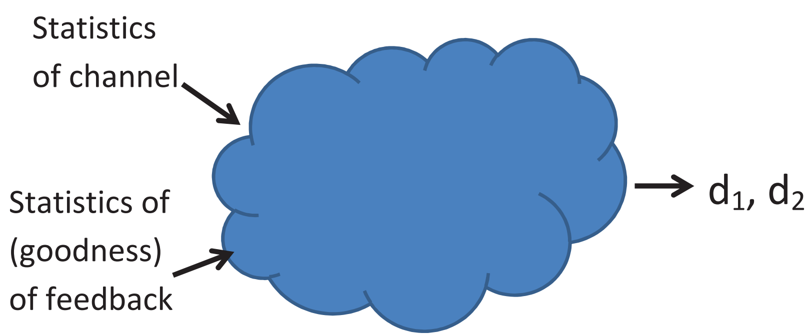
Instances of the problem

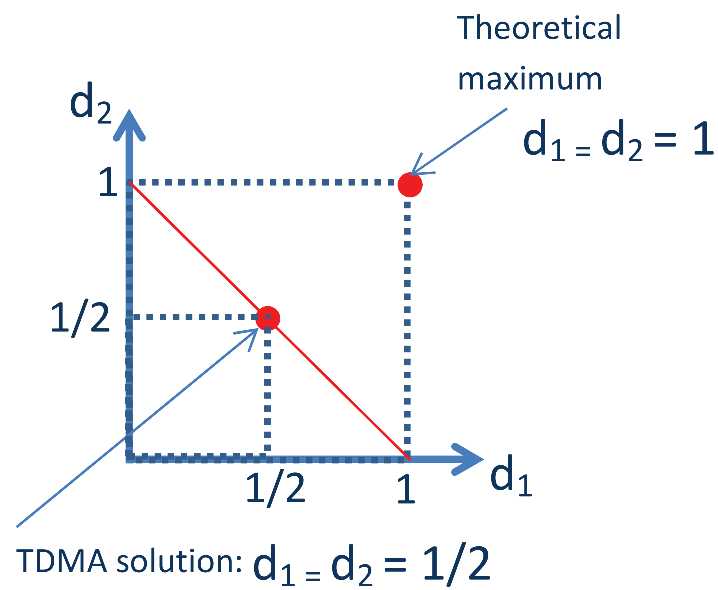




Estimation errors

Recall: performance in degrees-of-freedom (DoF)





$$d_i = \lim_{P\to\infty} \frac{R_i}{\log P}, \;\; i = 1, 2$$

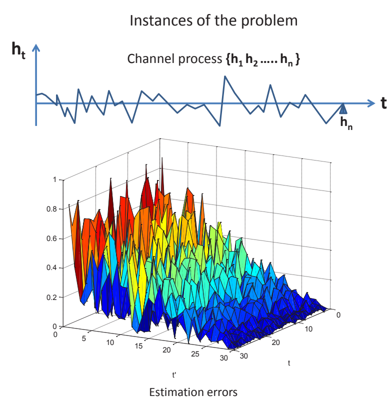
- (R_1, R_2): achievable rate pair $\quad R_i \approx d_i \log P$

Brief notations, conventions and assumptions

Notation

$$\overbrace{\alpha_t^{(1)} \triangleq -\lim_{P\to\infty} \frac{\log \mathbb{E}[||\boldsymbol{h}_t - \hat{\boldsymbol{h}}_{t,t}||^2]}{\log P} \qquad \alpha_t^{(2)} \triangleq -\lim_{P\to\infty} \frac{\log \mathbb{E}[||\boldsymbol{g}_t - \hat{\boldsymbol{g}}_{t,t}||^2]}{\log P}}^{\text{Quality of }\underline{\text{CURENT}}\text{ CSIT for channel at time } t}$$

$$\overbrace{\beta_t^{(1)} \triangleq -\lim_{P\to\infty} \frac{\log \mathbb{E}[||\boldsymbol{h}_t - \hat{\boldsymbol{h}}_{t,t+\eta}||^2]}{\log P} \qquad \beta_t^{(2)} \triangleq -\lim_{P\to\infty} \frac{\log \mathbb{E}[||\boldsymbol{g}_t - \hat{\boldsymbol{g}}_{t,t+\eta}||^2]}{\log P}}^{\text{Quality of }\underline{\text{DELAYED}}\text{ CSIT for channel at time } t}, \quad \eta < \infty$$



Notation$_1$

$$\underbrace{\bar{\alpha}^{(1)} \triangleq \frac{1}{n}\sum_{t=1}^{n}\alpha_t^{(1)} \qquad \bar{\alpha}^{(2)} \triangleq \frac{1}{n}\sum_{t=1}^{n}\alpha_t^{(2)} \qquad \bar{\beta}^{(1)} \triangleq \frac{1}{n}\sum_{t=1}^{n}\beta_t^{(1)} \qquad \bar{\beta}^{(2)} \triangleq \frac{1}{n}\sum_{t=1}^{n}\beta_t^{(2)}}_{\text{Average of exponent sequences}}$$

$$\text{Quality range (WOLOG):} \quad 0 \leq \alpha_t^{(i)} \leq \beta_t^{(i)} \leq 1$$

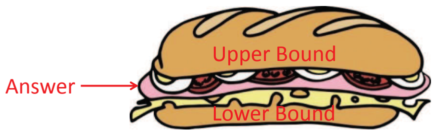
- Common conventions:
  - ⋆ Gaussian estimation errors
  - ⋆ Current estimate error is statistically independent of current and past estimates
  - ⋆ Wait for delayed-CSIT only for a finite amount of time
  - ⋆ Perfect and global knowledge of channel state information at receivers

Performance vs. CSIT timeliness and quality

- Results hold for general setting
  - ⋆ Communication over duration of n time slots: Channel $\left\{\boldsymbol{h}_t, \boldsymbol{g}_t\right\}_{t=1}^{n}$
  - ⋆ Feedback $\left\{\hat{\boldsymbol{h}}_{t,t'}, \hat{\boldsymbol{g}}_{t,t'}\right\}_{t,t'=1}^{n}$, of 'Goodness'

$$\left\{(\boldsymbol{h}_t - \hat{\boldsymbol{h}}_{t,t'}), (\boldsymbol{g}_t - \hat{\boldsymbol{g}}_{t,t'})\right\}_{t,t'=1}^{n}$$

- Answers in the form of bounds
  - ⋆ Novel precoders/decoders that try to lim-optimally use feedback
  - ⋆ Information theoretic outer bounds (try to prove optimality)



Magical reduction in difficulty of problem

Theorem: *(Chen-Elia 2013) The DoF region*

$$d_1 \leq 1, \quad d_2 \leq 1$$

$$2d_1 + d_2 \leq 2 + \bar{\alpha}^{(1)}$$

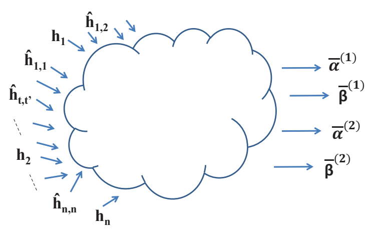
$$2d_2 + d_1 \leq 2 + \bar{\alpha}^{(2)}$$

$$d_1 + d_2 \leq \frac{1}{2}(2 + \bar{\beta}^{(1)} + \bar{\beta}^{(2)})$$

is achievable and is optimal for ... sufficiently good CSIT (To explain).

MAGICALLY, RESULT A FUNCTION OF JUST 4 STATISTICAL PARAMETERS!!!!

Complexity of the problem is captured by only 4 parameters



Specifically: Optimal DoF for sufficiently good delayed CSIT

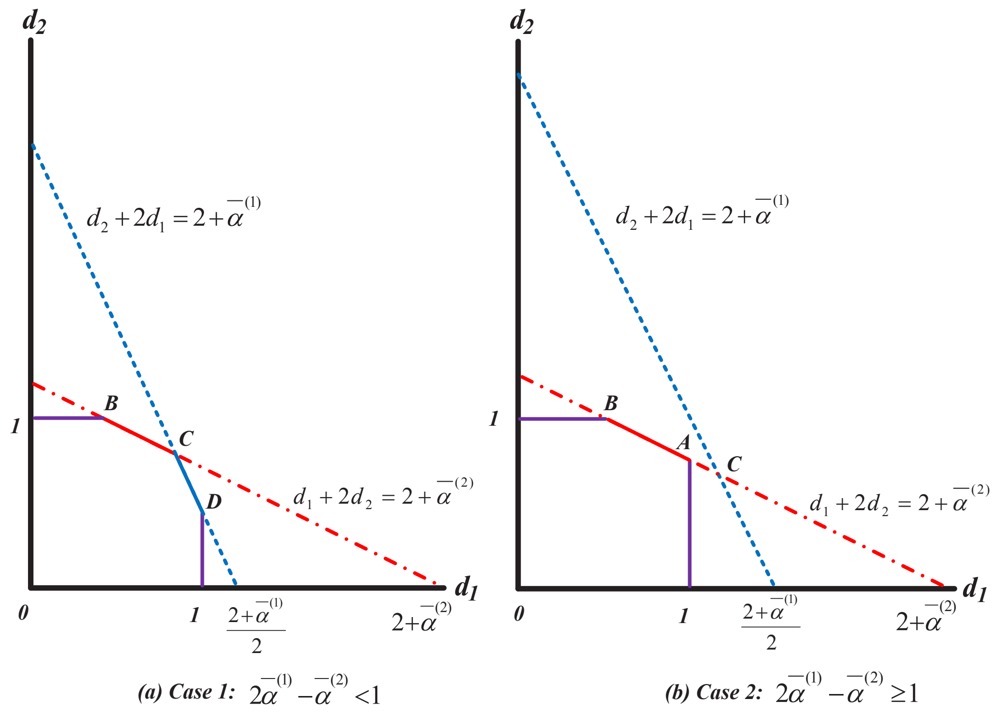
Theorem: *(Chen-Elia) The optimal DoF of the two-user MISO BC with a CSIT process* $\left\{\hat{\boldsymbol{h}}_{t,t'}, \hat{\boldsymbol{g}}_{t,t'}\right\}_{t=1,t'=1}^{n}$ *of quality* $\left\{(\boldsymbol{h}_t - \hat{\boldsymbol{h}}_{t,t'}), (\boldsymbol{g}_t - \hat{\boldsymbol{g}}_{t,t'})\right\}_{t=1,t'=1}^{n}$ *is given by*

$$d_1 \leq 1, \quad d_2 \leq 1$$
$$2d_1 + d_2 \leq 2 + \bar{\alpha}^{(1)}$$
$$2d_2 + d_1 \leq 2 + \bar{\alpha}^{(2)}$$

for any sufficiently good delayed-CSIT process such that

$$\min\{\bar{\beta}^{(1)}, \bar{\beta}^{(2)}\} \geq \min\{\frac{1 + \bar{\alpha}^{(1)} + \bar{\alpha}^{(2)}}{3}, \frac{1 + \min\{\bar{\alpha}^{(1)}, \bar{\alpha}^{(2)}\}}{2}\}$$

Specifically: Optimal DoF for sufficiently good delayed CSIT

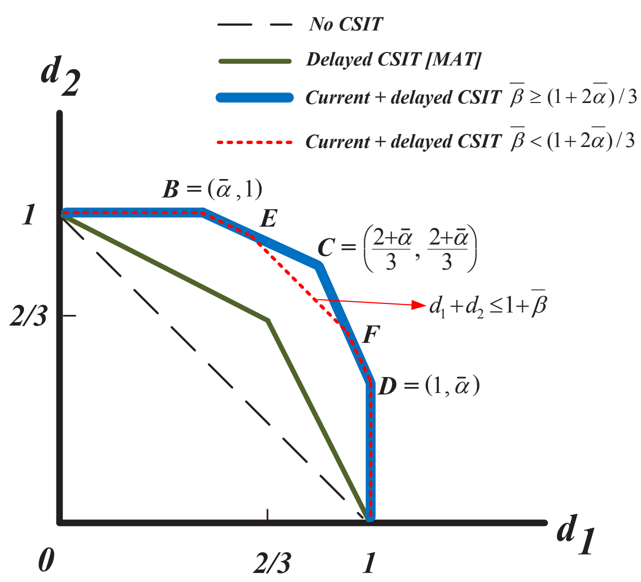


- *Optimal DoF regions for the two-user MISO BC with sufficiently good delayed CSIT.*

Symmetric case

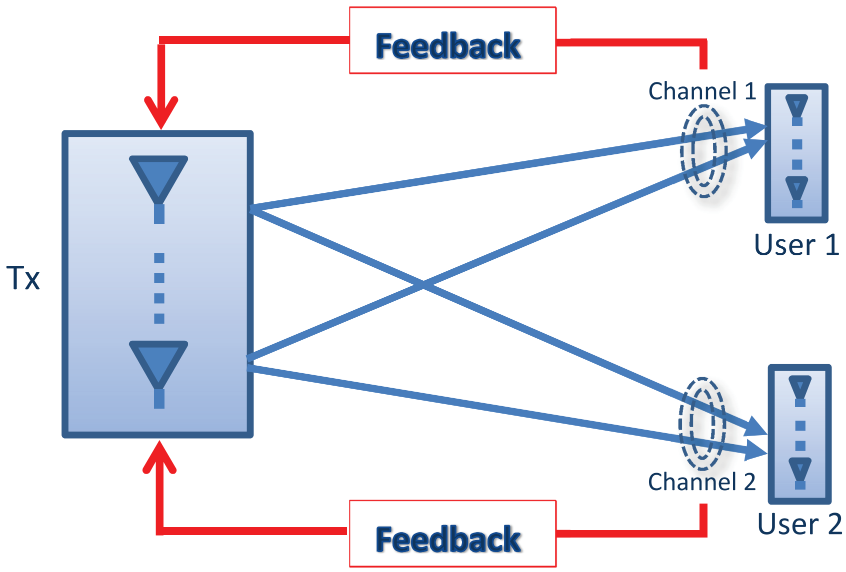
Users have similar long-term feedback capabilities

$$\bar{\alpha}^{(1)} = \bar{\alpha}^{(2)} = \bar{\alpha}$$
$$\bar{\beta}^{(1)} = \bar{\beta}^{(2)} = \bar{\beta}$$



MIMO BC

What if I have many transmit and receive antennas?



Theorem: *The optimal DoF region of the Two-user Symmetric $M \times (N, N)$ MIMO BC with sufficiently good delayed CSIT* [3]

$$d_1 + d_2 \leq \langle 2N \rangle'$$

$$d_1 \leq \langle N \rangle' ; \quad \frac{d_1}{\langle N \rangle'} + \frac{d_2}{\langle 2N \rangle'} \leq 1 + \frac{\langle 2N \rangle' - \langle N \rangle'}{\langle 2N \rangle'} \, \bar{\alpha}^{(1)}$$

$$d_2 \leq \langle N \rangle' ; \quad \frac{d_1}{\langle 2N \rangle'} + \frac{d_2}{\langle N \rangle'} \leq 1 + \frac{\langle 2N \rangle' - \langle N \rangle'}{\langle 2N \rangle'} \, \bar{\alpha}^{(2)}$$

[3] $\langle \bullet \rangle' = \min\{\bullet, M\}$. 'Sufficiently good delayed CSIT': $\min\{\bar{\beta}^{(1)}, \bar{\beta}^{(2)}\} \geq \min\{1, M - N', \frac{N(1+\bar{\alpha}^{(1)}+\bar{\alpha}^{(2)})}{\langle 2N \rangle' + N}, \frac{N(1+\min\{\bar{\alpha}^{(1)}+\bar{\alpha}^{(2)}\})}{\langle 2N \rangle'}\}$.

Theorem: *(Chen-Elia) The optimal DoF region of the Two-user Symmetric* $(M, M) \times (N, N)$ *IC with sufficiently good delayed CSIT, is*

$$d_1 + d_2 \leq \min\{2M, 2N, \max\{M, N\}\}$$

$$d_1 \leq \langle N \rangle' ; \quad \frac{d_1}{\langle N \rangle'} + \frac{d_2}{\langle 2N \rangle'} \leq 1 + \frac{\langle 2N \rangle' - \langle N \rangle'}{\langle 2N \rangle'} \, \bar{\alpha}^{(1)}$$

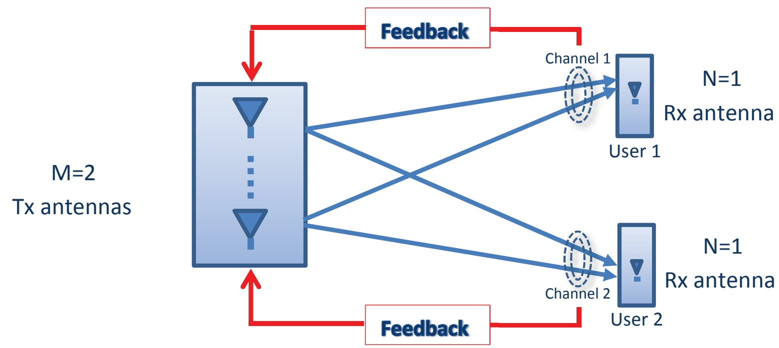
$$d_2 \leq \langle N \rangle' ; \quad \frac{d_1}{\langle 2N \rangle'} + \frac{d_2}{\langle N \rangle'} \leq 1 + \frac{\langle 2N \rangle' - \langle N \rangle'}{\langle 2N \rangle'} \, \bar{\alpha}^{(2)}$$

INSIGHT

Aim of asymptotic analysis is exactly this:
qualitative insight

Insight: more antennas for less CSIT quality

Can, having more receive antennas, allow for reduced feedback quality?

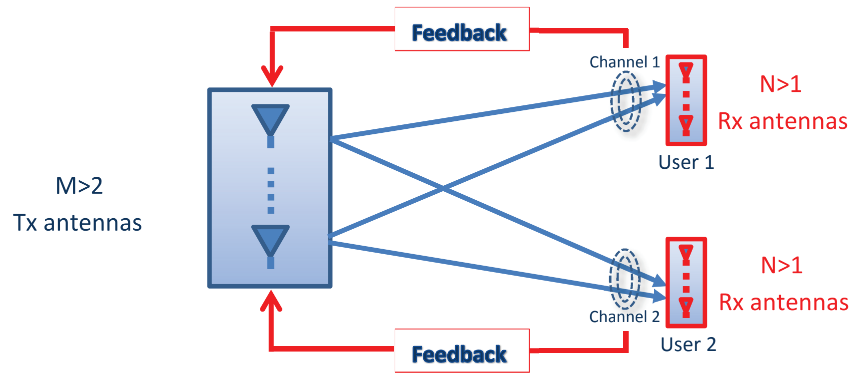


- Previous results show that, to achieve $d_1 = d_2 = 1$, we need constantly 'perfect' feedback.

$$\alpha_t^{(1)} = \alpha_t^{(2)} = 1, \ \forall t \quad \Rightarrow \bar{\alpha}^{(1)} = \bar{\alpha}^{(2)} = 1$$

More antennas for less CSIT quality

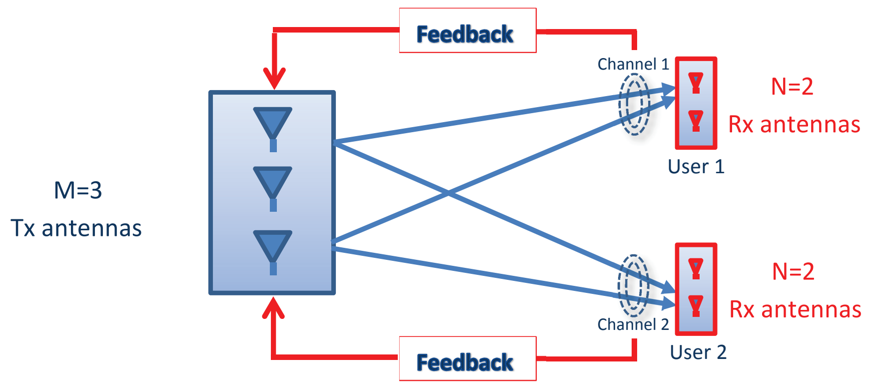
What if we have more antennas?



Corollary: *(Chen-Elia) A CSIT process with* $\bar{\alpha}^{(1)}+\bar{\alpha}^{(2)} \geq \min\{M, 2N\}/N$, *achieves the optimal sum-DoF associated to perfect feedback*[4].

[4] Interested in $M > N$ (recall that if $M \leq N$, then no CSIT is needed)

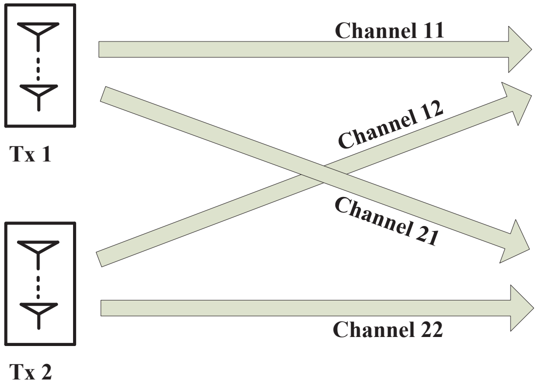
More antennas for less CSIT quality[1]



EXAMPLE: $M = 3, N = 2$

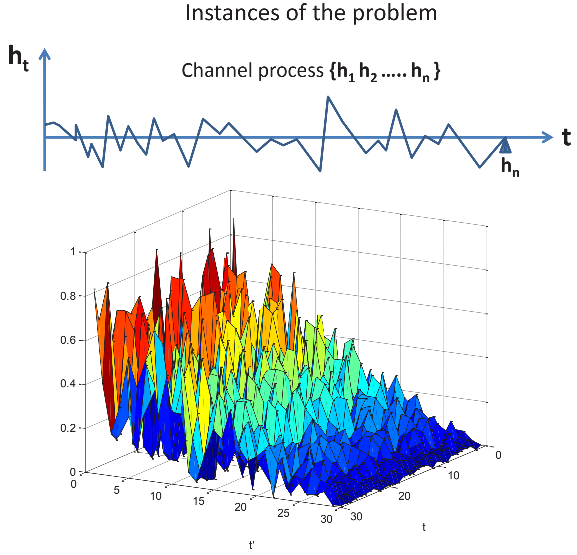
- Note: perfect CSIT ($\bar{\alpha}^{(1)} = \bar{\alpha}^{(2)} = 1$) gives optimal sum-DoF of 3
- BUT: same sum DoF with $\bar{\alpha}^{(1)} + \bar{\alpha}^{(2)} = 3/2$
  - ⋆ e.g. $\bar{\alpha}^{(1)} = \bar{\alpha}^{(2)} = 3/4$

Insight: MIMO-IC

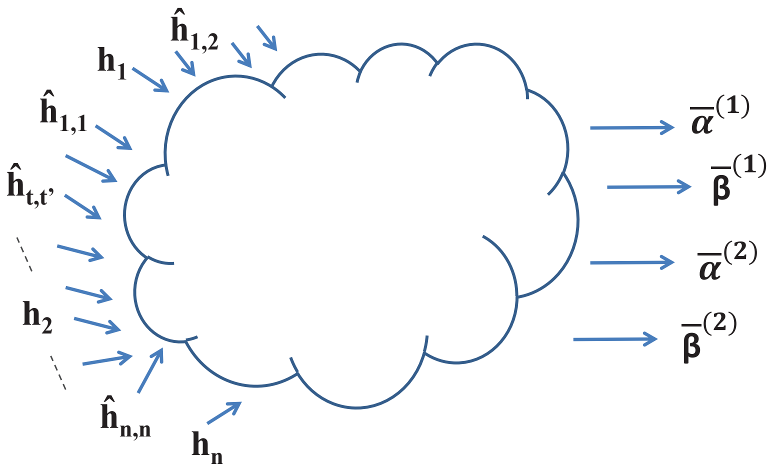


Corollary: *In the IC, no CSIT is needed for the direct links.*

Insight: reduced 'problem complexity'



Complexity of the problem is captured by only 4 parameters

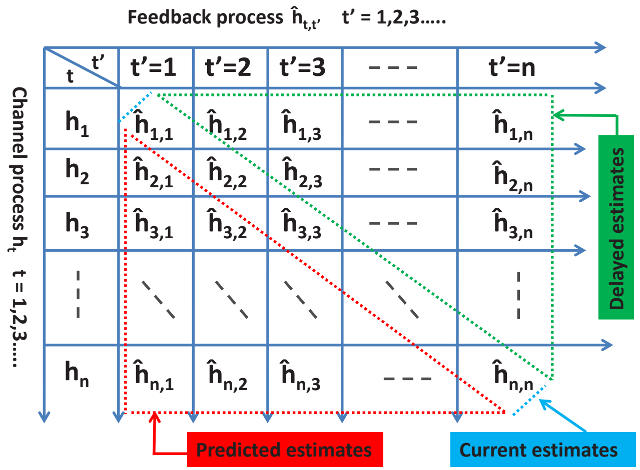


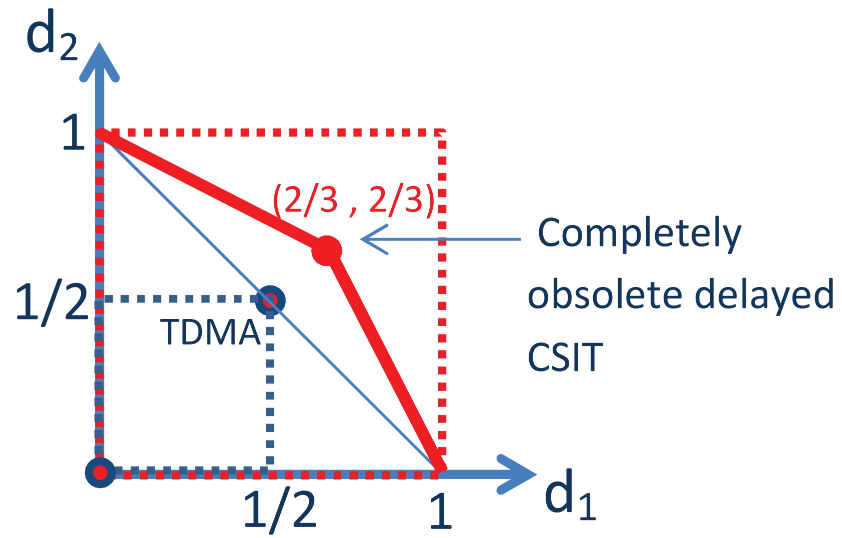
- Gaussianity $\Rightarrow$ Statistics of $\left\{(\boldsymbol{h}_t - \hat{\boldsymbol{h}}_{t,t'}), (\boldsymbol{g}_t - \hat{\boldsymbol{g}}_{t,t'})\right\}_{t,t'=1}^{n}$ captured by covariance matrix

$$Cov\left(vect(\left\{(\boldsymbol{h}_t - \hat{\boldsymbol{h}}_{t,t'}), (\boldsymbol{g}_t - \hat{\boldsymbol{g}}_{t,t'})\right\}_{t,t'=1}^{n})\right) \in \mathbb{C}^{2n^2 \times 2n^2}$$

- Diagonal entries of $Cov(\bullet)$ are $\left\{\frac{1}{M}\mathbb{E}[||\boldsymbol{h}_t - \hat{\boldsymbol{h}}_{t,t'}||^2], \frac{1}{M}\mathbb{E}[||\boldsymbol{g}_t - \hat{\boldsymbol{g}}_{t,t'}||^2]\right\}_{t,t'=1}^{n}$. Some of them are represented by the exponents
- But, the rest, plus the off-diagonal entries not used by scheme
- But, scheme meets outer bound that holds irrespective of these other entries
- $\Rightarrow$ exponents faithfully represent problem
- In the end only the four averages show up

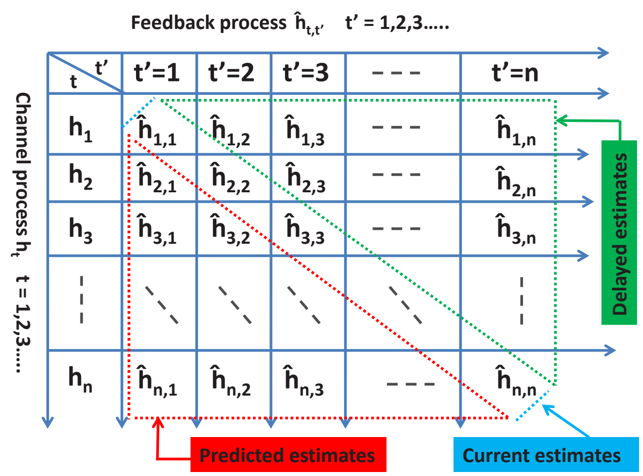
Theorem: *(Maddah-Ali and Tse) (Have seen). Completely obsolete feedback helps.*





Insight: predicted CSIT?

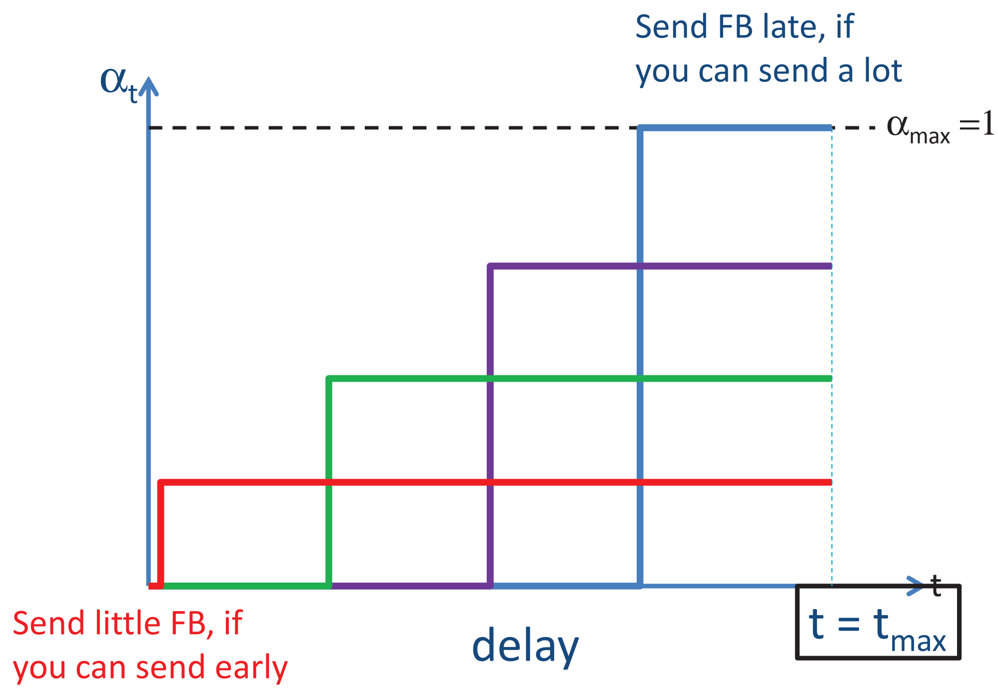
Corollary: *(Chen-Elia) There is no DoF gain in using predicted CSIT*[5],[6].



[5]For sufficiently good delayed CSIT. Same conclusion also holds based on inner bounds.

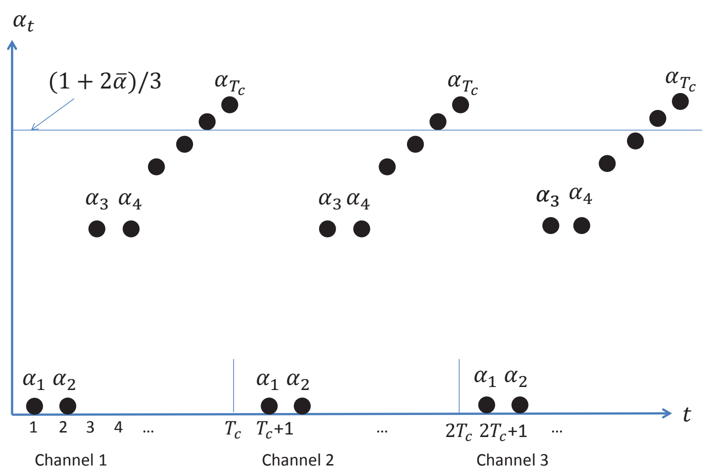
[6]No need to utilize predictions in shaping your current transmission.

Insight: Less feedback early, or more feedback later?



Periodic feedback in block fading - a useful tool

time	$t = 1$	$t = 2$	$t = 3$	$t = 4$	$\cdots$	$t = T$	$t > T$
quality exponent	$0 \leq \alpha_1$	α_2	α_3	α_4	$\cdots$	α_T	$\beta \leq 1$



Evolving CSIT: examples

EXAMPLE: How to achieve target DoF $d_1 = d_2 = d' = 7/9$?

- Sequence

$$\underbrace{\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_T}_{\substack{\text{Progressive feedback} \\ \text{during coherence period}}} \quad \leq \quad \underbrace{\beta}_{\substack{\text{Delayed feedback} \\ \text{after coherence period}}}$$

- Optimal (symmetric) DoF was given:

$$d = \frac{2 + \bar{\alpha}}{3}$$

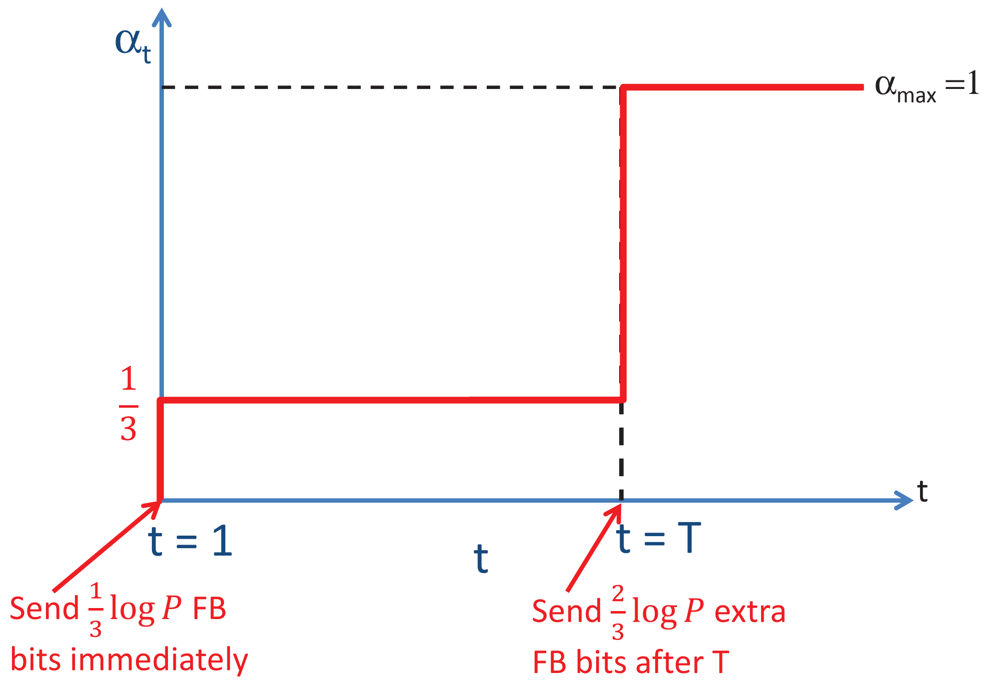
  ⋆ where $\bar{\alpha} = \text{average}(\alpha_1, \alpha_2, \cdots, \alpha_T)$

- Thus solve: We need

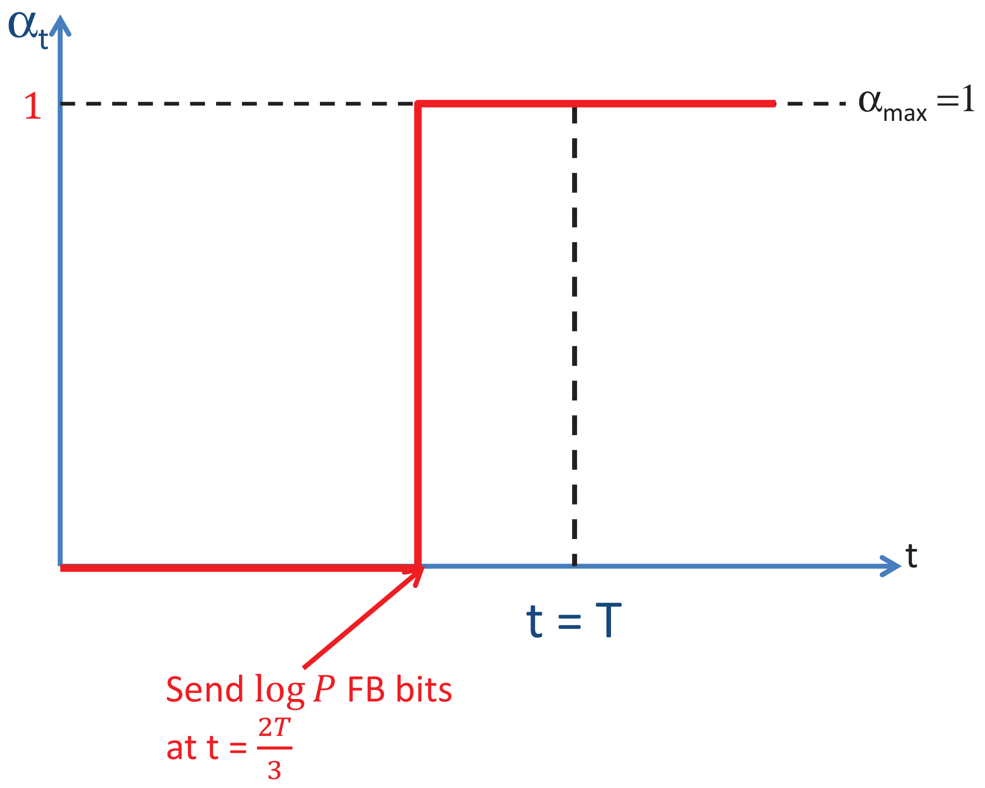
$$\bar{\alpha} \geq 3d' - 2 = 3 \cdot \frac{7}{9} - 2 = 1/3$$

- What are the feedback options?

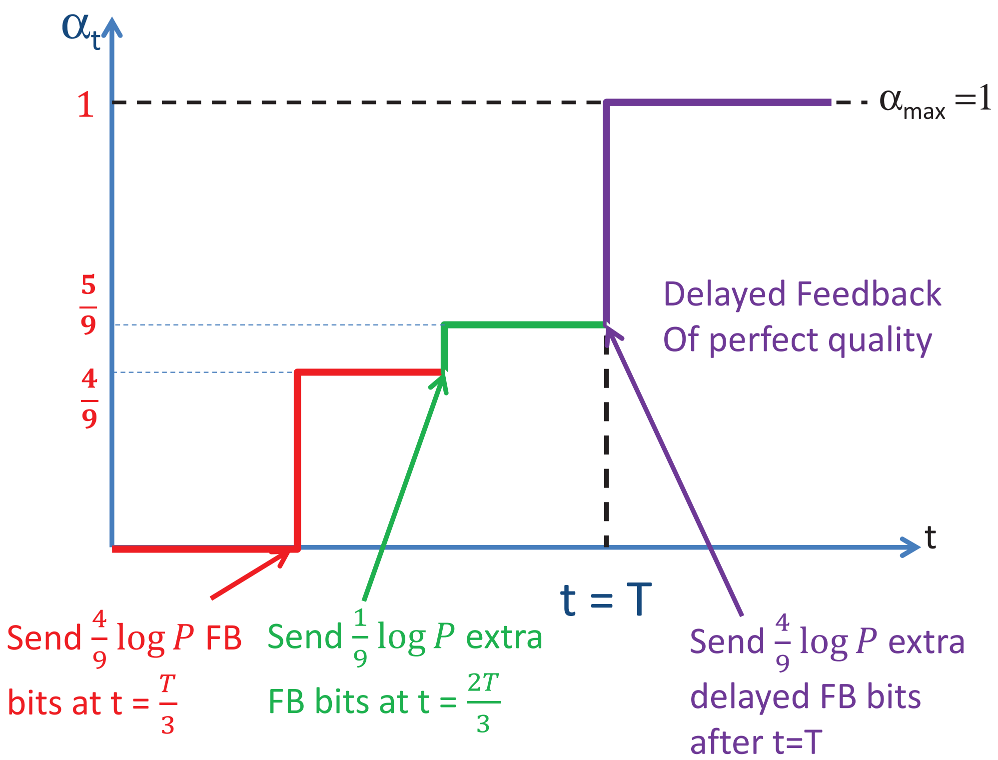
$\bar{\alpha} = 1/3$: Option 1



$\bar{\alpha} = 1/3$: Option 2



$\bar{\alpha} = 1/3$: Option 3

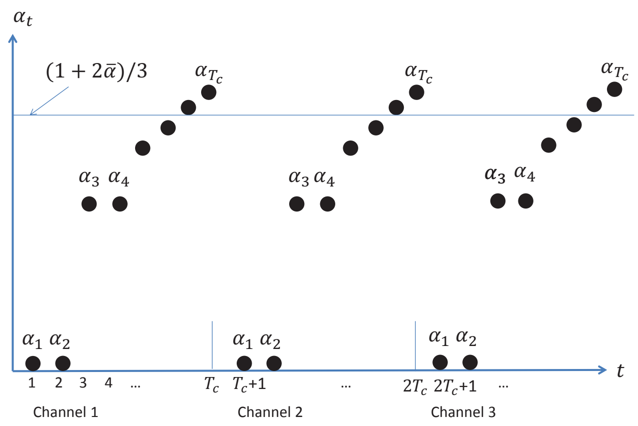


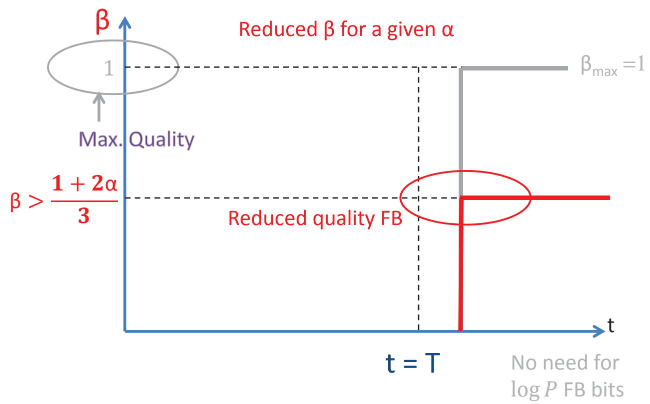
Insight: Reducing total feedback

HOW TO REDUCE TOTAL AMOUNT OF FEEDBACK?

- Must reduce delayed feedback quality (reduce β)

When is delayed feedback unnecessary?





- *Corollary: Having delayed-CSIT quality $\beta \geq \frac{1+2\bar{\alpha}}{3}$ is equivalent to having perfect delayed CSIT.*

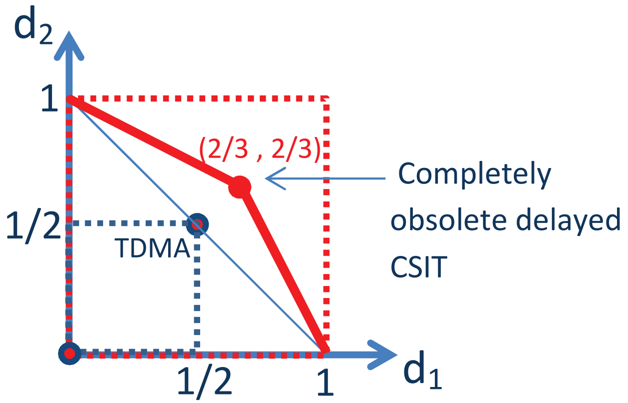
- *Corollary: When $\alpha_T \geq \frac{1+2\bar{\alpha}}{3}$, there is no need for any delayed CSIT, i.e., do not send feedback after the end of the coherence block.*

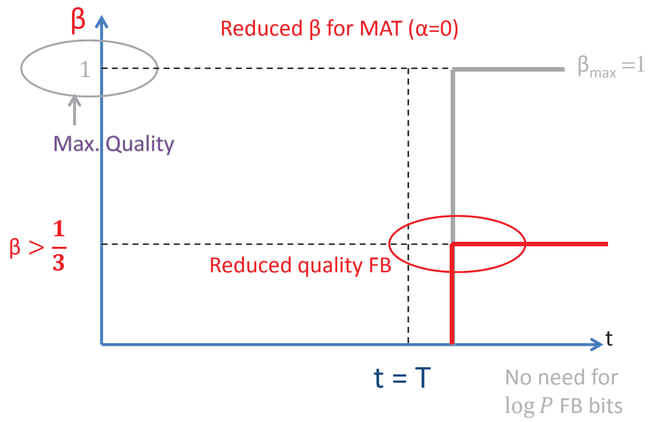
Corollary: MAT with fewer bits

EXAMPLE:

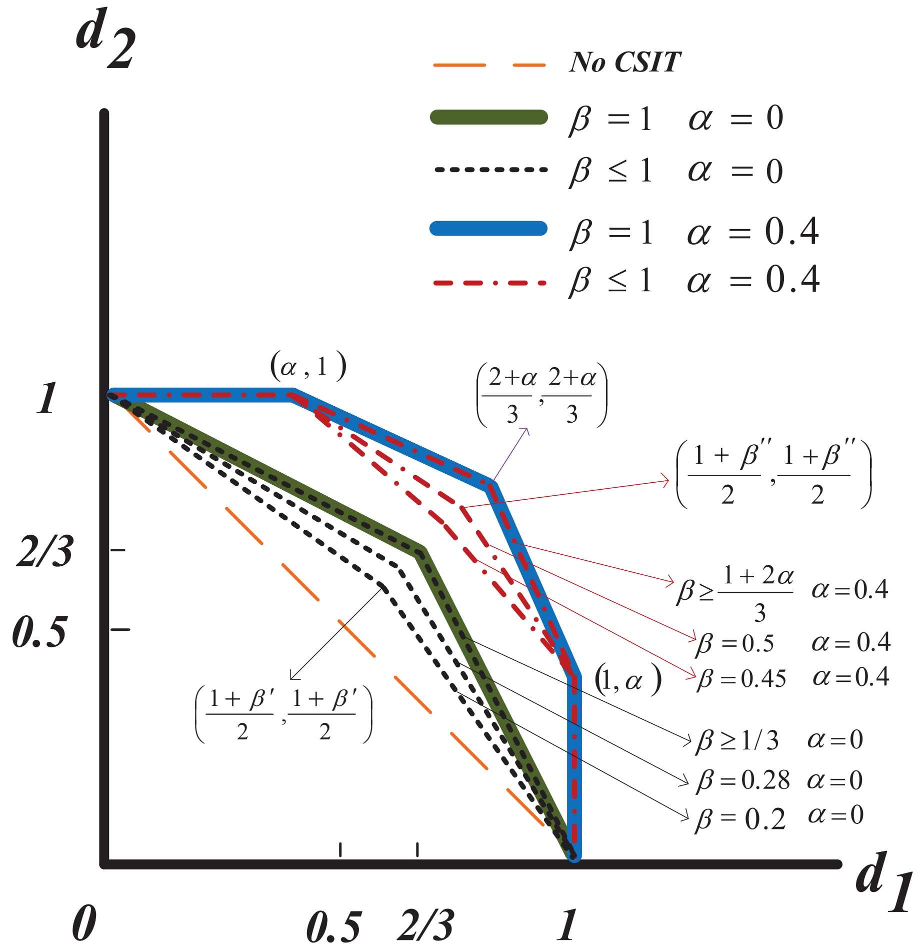
- Can we get the MAT $d = 2/3$, with $< \log P$ (current + delayed) feedback bits?
  - ⋆ I.e., with imperfect delayed feedback

- *Corollary (Chen-Elia): MAT case (originally $\beta = 1, \alpha = 0$):*
  $\beta = 1/3$ suffices to achieve the optimal region ($d_1 = d_2 = 2/3$)





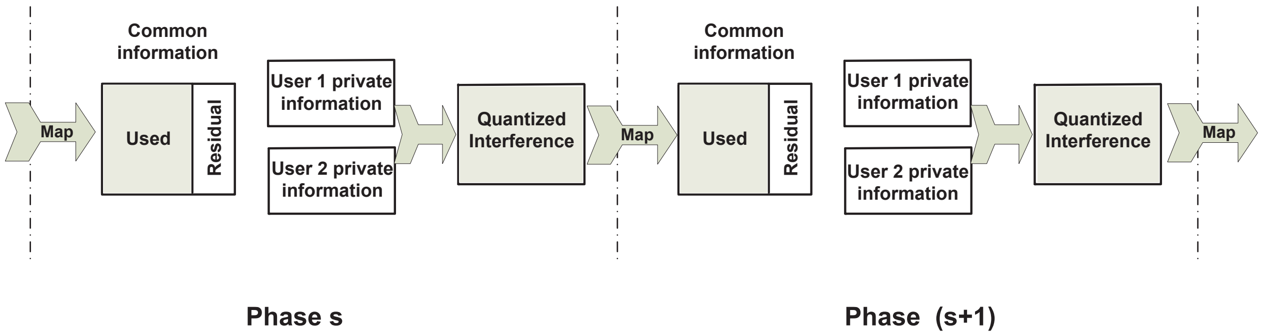
Insight: feedback flow



Universal encoding-decoding scheme

Schemes exploit imprecise, delayed or premature feedback

Block Markov type schemes

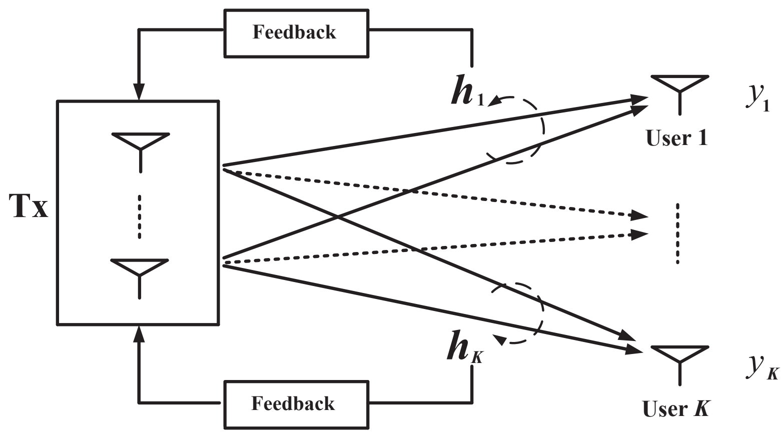


Encoding and decoding phase-Markov scheme:

- Accumulated quantized interference bits of phase s, can be broadcasted to both users inside the common information symbols of the next phase
- while also a certain amount of common information can be transmitted to both users during phase s, which will then help resolve the accumulated interference of phase $(s-1)$.
- All parameters (power and rate allocation, etc) are functions of the (declared) quality exponents

Similar channel model: K-user MISO BC

K-user MISO BC
A wide range of open problems



K-user MISO BC with only delayed feedback

What we know:

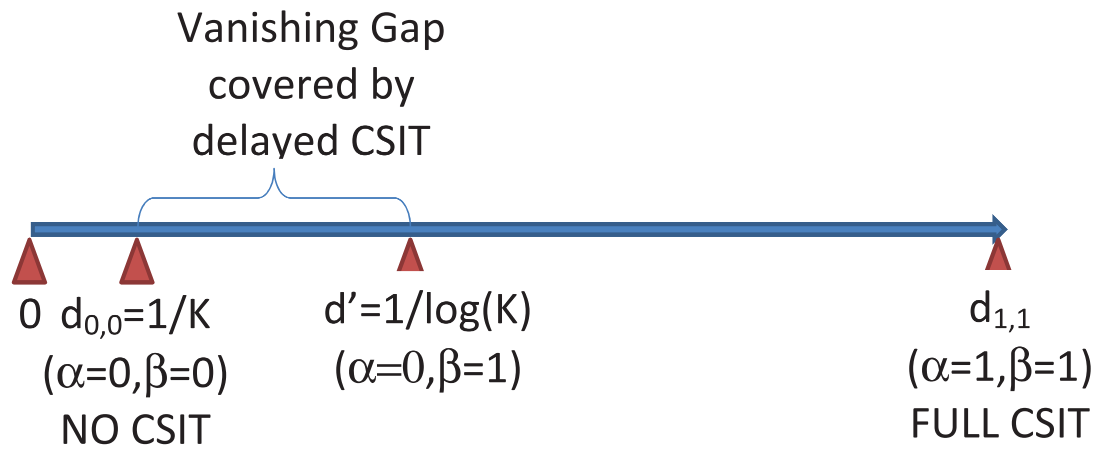
Theorem: *(Maddah-Ali and Tse) The optimal sum-DoF* $d_\Sigma \triangleq \sum_{k=1}^{K} d_k$ *of the* K*-user MISO BC with delayed feedback, takes the form*

$$d_{MAT} \triangleq \frac{K}{1 + \frac{1}{\min\{2,M\}} + \frac{1}{\min\{3,M\}} + \cdots + \frac{1}{\min\{K,M\}}}$$

Corollary 1 *(Maddah-Ali and Tse) When* $M \geq K \to \infty$ *then*

$$d_{MAT} \approx \frac{K}{\ln K}$$

K-user BC with only delayed feedback



Glass half-full or half-empty

- Recall that no feedback gives $d_\Sigma = 1$
- Recall that perfect feedback gives $d_\Sigma = K$
- Good news:

$$d_{\text{MAT}} \approx \frac{K}{\ln K} >> 1 \quad \text{(scales with } K\text{)}$$

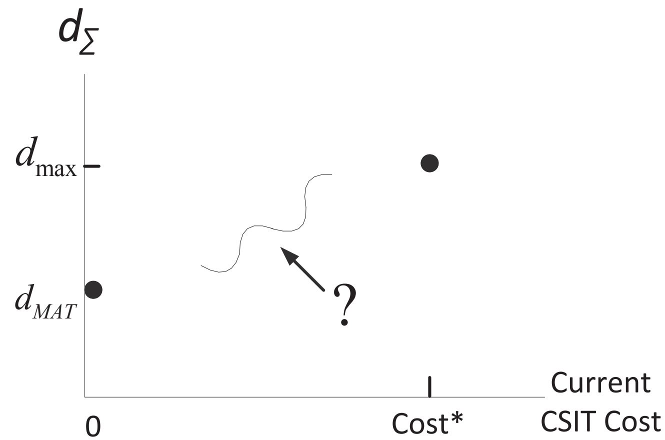
- Bad news:

$$\frac{d_{\text{MAT}}}{K} \approx \frac{1}{\ln K} \to 0 \quad \text{(vanishing per user DoF)}$$

K-user BC with only delayed feedback[1]

K-USER PROBLEM LARGELY OPEN

- Strong need for understanding role of current feedback



  ⋆ [Tandon et al. 12] [Lee and Heath 12]

- Strong need for outer bounds [Tandon et al. 12][Chen-Yang-Elia 13]

Same problem formulation

- Communication of duration n (n is large)
- An arbitrary channel fading process (random)

$$\left\{\boldsymbol{h}_{k,t}\right\}_{k=1,\ t=1}^{K\quad n}$$

- An arbitrary feedback process (CSIT)

$$\left\{\hat{\boldsymbol{h}}_{k,t,t'}\right\}_{k=1,\ t=1,\ t'=1}^{K\quad n\quad n}$$

  - $\hat{\boldsymbol{h}}_{k,t,t'}$: CSIT estimate at any time t', of channel $\boldsymbol{h}_{k,t}$ (at time t)

- A 'primitive' measure of feedback 'goodness'

$$\left\{(\boldsymbol{h}_{k,t} - \hat{\boldsymbol{h}}_{k,t,t'})\right\}_{k=1,\ t=1,\ t'=1}^{K\quad n\quad n}$$

New outer bound (DoF Region)

For general setting: general channel process (large duration n), general feedback process

Theorem: *[DoF region outer bound] (Chen-Elia): The DoF region of the K-user $M \times 1$ MISO BC with a general CSIT feedback process, is outer bounded as*

$$\sum_{k=1}^{K}\frac{d_{\pi(k)}}{\min\{k,M\}} \leq 1+\sum_{k=1}^{K-1}\left(\frac{1}{\min\{k,M\}}-\frac{1}{\min\{K,M\}}\right)\bar{\alpha}^{(\pi(k))}$$

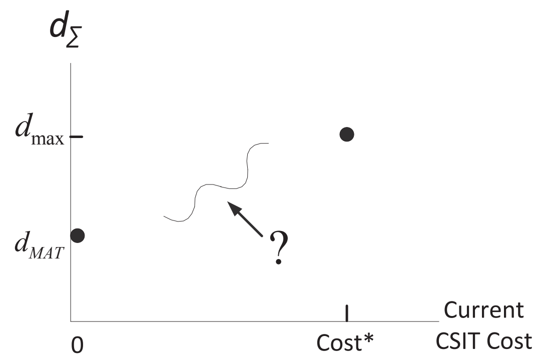
$$d_k \leq 1, \quad k=1,2,\cdots,K$$

New outer bound (Sum DoF)

Corollary: *[Sum DoF outer bound] For the K-user $M \times 1$ MISO BC, the sum DoF is outer bounded as*

$$d_{\Sigma} \leq d_{MAT} + \left(1 - \frac{d_{MAT}}{\min\{K, M\}}\right) \sum_{k=1}^{K} \bar{\alpha}^{(k)}$$

Current CSIT cost vs sum DoF



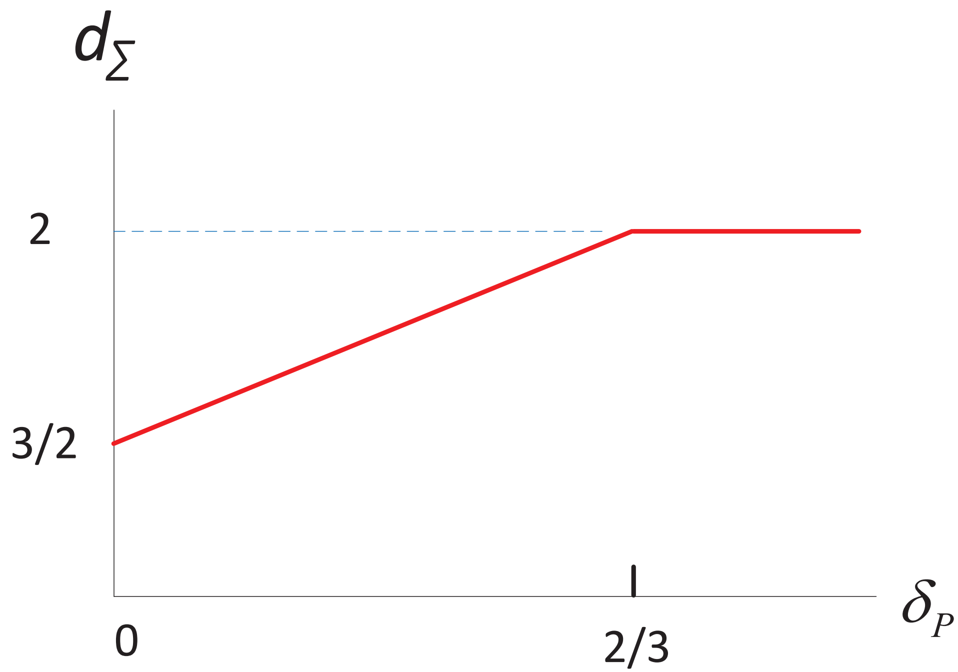
What is the current CSIT cost for a certain $d_\Sigma \in [d_{\text{MAT}}, d_{\max}]$?

- E.g, for the case with $M = 2, K = 3$ ($d_{\text{MAT}} = \frac{3}{2}, d_{\max} = 2$)
  - What is the current CSIT cost for $d_\Sigma = \frac{7}{4}$?
  - What is the current CSIT cost for $d_\Sigma = \frac{5}{3}$?

Current CSIT cost vs sum DoF

Theorem: *[Optimal cases, d_Σ vs $\bar{\alpha}$] For the K-user MISO BC with $M \geq K$ or with $M = 2, K = 3$, and given a current CSIT cost $\bar{\alpha}$, the optimal sum DoF is characterized as*

$$d_\Sigma = d_{MAT} + \left(K - \frac{K d_{MAT}}{\min\{K, M\}}\right) \min\left\{\bar{\alpha}, \frac{\min\{K, M\}}{K}\right\}$$

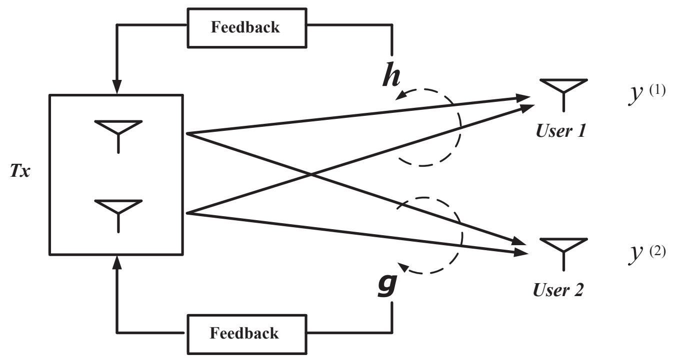


Optimal sum DoF (d_Σ) vs. $\bar{\alpha} =: \delta_p$ for the MISO BC with $M = 2, K = 3$

($\bar{\alpha} = 1/3$ for $d_\Sigma = \frac{7}{4}$) and ($\bar{\alpha} = 2/9$ for $d_\Sigma = \frac{5}{3}$)

Global CSIR

GLOBAL CHANNEL STATE INFORMATION AT RECEIVERS
(GLOBAL CSIR)



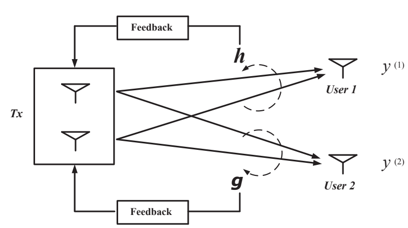
- Global CSIR: A user must know the channels of the other users

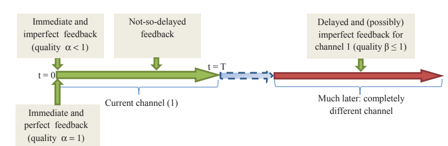
The challenge of global CSIR

GREAT CHALLENGE IN DISTRIBUTING PERFECT GLOBAL CSIR (see Kobayashi-Caire ISIT 2012)

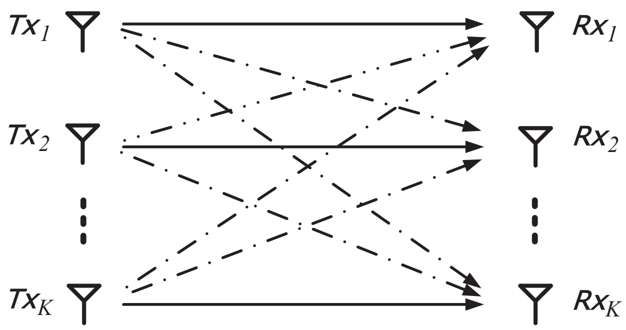
- Training and limited-capacity/limited-reliability feedback links
- Challenge extreme as number of users increases
  - ⋆ How to use imperfect-quality and delayed global CSIR in difference interference settings?
- Problem: Achilles' heel of delayed-CSIT approaches

CONSIDER IMPERFECT AND DELAYED GLOBAL CSIR





Interference alignment with delayed and imperfect-quality CSIT



A PROMISING DIRECTION:
INTERFERENCE ALIGNMENT WITH DELAYED FEEDBACK

- Interference alignment (IA) [Cadambe and Jafar 08] $d = 1/2$
  - ⋆ "Each user gets half of the cake"
  - ⋆ Powerful tool but!
  - ⋆ Global and perfect CSIT is required for IA

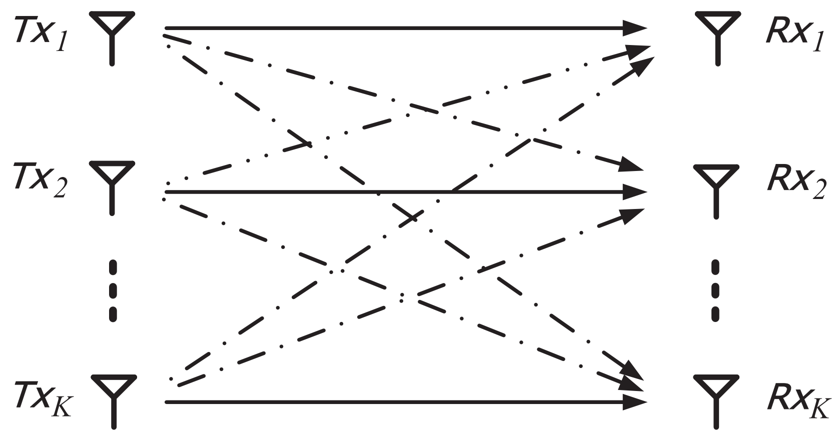
SOME EXISTING APPROACHES

- With delayed CSIT 3×3 SISO IC can achieve $\frac{9}{8}$ sum DoF
  - ⋆ [Maleki, Jafar and Shamai 11]
- With delayed CSIT 3×3 SISO IC can achieve $\frac{36}{31}$ sum DoF
  - ⋆ [Abdoli, Ghasemi and Khandani 11]

A MAIN OPEN PROBLEM:

$K \times K$ SISO (MIMO) IC WITH IMPERFECT AND DELAYED CSIT

Extension to X channel



X channel: Each transmitter has a message to be communicated with each receiver

- With perfect global CSIT, $M \times N$ SISO X channel has sum DoF $\frac{MN}{M+N-1}$
  - ⋆ [Cadambe and Jafar 2009]
  - ⋆ Example: $2 \times 2:$ sum Dof $= \frac{4}{3}$

Extension to X channel[1]

- With delayed CSIT 2×2 SISO X channel can achieve $\frac{8}{7}$ sum DoF
  - ⋆ [Maleki et al. 2011]
- With delayed CSIT 2×2 SISO X channel can achieve $\frac{6}{5}$ sum DoF
  - ⋆ [Ghasemi, Motahari and Khandani 11]
- With delayed CSIT 3×3 SISO X channel can achieve $\frac{5}{4}$ sum DoF
  - ⋆ [Ghasemi, Motahari and Khandani 11]

THE MAIN OPEN PROBLEM:

WHAT IS OPTIMAL DOF FOR $K \times K$ SISO X CHANNEL WITH DELAYED AND IMPERFECT CSIT?

Bibliography

Bibliography

[1] A. E. Gamal, "The feedback capacity of degraded broadcast channels," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 379–381, Apr. 1978.

[2] M. A. Maddah-Ali, "On the degrees of freedom of the compound MIMO broadcast channels with finite states," Oct. 2009, available on arXiv:0909.5006v3.

[3] M. A. Maddah-Ali and D. N. C. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4418 – 4431, Jul. 2012.

[4] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Capacity results for binary fading interference channels with delayed CSIT," Jan. 2013, submitted to *IEEE Trans. Inf. Theory*, available on arXiv:1301.5309.

[5] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley-Interscience, 2006.

[6] D. Tse and P. Viswanath, *Fundamentals of wireless communication.* Cambridge University Press, 2005.

[7] N. Lee and R. W. Heath Jr., "Not too delayed CSIT achieves the optimal degrees of freedom," in *Proc. Allerton Conf. Communication, Control and Computing*, Oct. 2012.

[8] C. Vaze and M. Varanasi, "The degree-of-freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5254 – 5374, Aug. 2012.

[9] C. S. Vaze and M. K. Varanasi, "The degrees of freedom region of two-user and certain three-user MIMO broadcast channel with delayed CSI," Dec. 2011, submitted to *IEEE Trans. Inf. Theory*, available on arXiv:1101.0306.

[10] Y. Lejosne, D. Slock, and Y. Yuan-Wu, "Degrees of freedom in the MISO BC with delayed-CSIT and finite coherence time: A simple optimal scheme," in *Proc. IEEE Int. Conf. on Signal Processing, Communications and Control (ICSPCC)*, Aug. 2012.

[11] C. Hao and B. Clerckx, "Imperfect and unmatched CSIT is still useful for the frequency correlated MISO broadcast channel," in *Proc. IEEE Int. Conf. Communications (ICC)*, Jun. 2013.

[12] 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. 2845 – 2866, Jun. 2010.

[13] A. Adhikary, H. C. Papadopoulos, S. A. Ramprashad, and G. Caire, "Multi-user MIMO with outdated CSI: Training, feedback and scheduling," Sep. 2011, available on arXiv:1109.6371.

[14] M. Kobayashi and G. Caire, "On the net DoF comparison between ZF and MAT over time-varying MISO broadcast channels," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Jul. 2012.

[15] M. J. Abdoli, A. Ghasemi, and A. K. Khandani, "On the degrees of freedom of three-user MIMO broadcast channel with delayed CSIT," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Jul. 2011.

[16] M. J. Abdoli and A. Ghasemi and A. K. Khandani, "On the degrees of freedom of K-user SISO interference and X channels with delayed CSIT," Nov. 2011, submitted to *IEEE Trans. Inform. Theory*, available on arXiv:1109.4314v2.

[17] A. Ghasemi, A. S. Motahari, and A. K. Khandani, "On the degrees of freedom of X channel with delayed CSIT," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Jul. 2011.

[18] A. Ghasemi, A. S. Motahari, and A. K. Khandani, "Interference alignment for the MIMO interference channel with delayed local CSIT," Feb. 2011, available on arXiv:1102.5673v1.

[19] A. Ghasemi, A. S. Motahari, and A. K. Khandani, "On the degrees of freedom of X channel with delayed CSIT," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2011.

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

[21] G. Caire, N. Jindal, and S. Shamai, "On the required accuracy of transmitter channel state information in multiple antenna broadcast channels," in *Proc. Allerton Conf. Communication, Control and Computing*, Nov. 2007.

[22] A. Lapidoth, S. Shamai, and M. A. Wigger, "On the capacity of fading MIMO broadcast channels with imperfect transmitter side-information," in *Proc. Allerton Conf. Communication, Control and Computing*, Sep. 2005.

[23] H. Weingarten, S. Shamai, and G. Kramer, "On the compound MIMO broadcast channel," in *Proc. Inf. Theory and App. Workshop (ITA)*, Jan. 2007.

[24] T. Gou, S. A. Jafar, and C. Wang, "On the degrees of freedom of finite state compound wireless networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3286 – 3308, Jun. 2011.

[25] T. Gou and S. Jafar, "Optimal use of current and outdated channel state information: Degrees of freedom of the MISO BC with mixed CSIT," *IEEE Communications Letters*, vol. 16, no. 7, pp. 1084 – 1087, Jul. 2012.

[26] C. Huang, S. A. Jafar, S. Shamai, and S. Vishwanath, "On degrees of freedom region of MIMO networks without channel state information at transmitters," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 849 – 857, Feb. 2012.

[27] S. Jafar and A. Goldsmith, "Isotropic fading vector broadcast channels: The scalar upper bound and loss in degrees of freedom," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 848 – 857, Mar. 2005.

[28] S. A. Jafar, "Blind interference alignment," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 3, pp. 216 – 227, Jun. 2012.

[29] S. A. Jafar, "Topological interference management through index coding," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 529 – 568, Jan. 2014.

[30] H. Maleki, S. Jafar, and S. Shamai, "Retrospective interference alignment over interference networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 3, pp. 228 – 240, Mar. 2012.

[31] R. Tandon, S. A. Jafar, S. Shamai, and H. V. Poor, "On the synergistic benefits of alternating CSIT for the MISO broadcast channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4106 – 4128, Jul. 2013.

[32] J. Xu, J. G. Andrews, and S. A. Jafar, "Broadcast channels with delayed finite-rate feedback: Predict or observe?" *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1456 – 1467, Apr. 2012.

[33] O. Shayevitz and M. Wigger, "On the capacity of the discrete memoryless broadcast channel with feedback," , to appear in *IEEE Trans. Inf. Theory* 2013.

[34] M. Kobayashi, S. Yang, D. Gesbert, and X. Yi, "On the degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Jul. 2012.

[35] P. de Kerret, X. Yi, and D. Gesbert, "On the degrees of freedom of the K-user time correlated broadcast channel with delayed CSIT," Jan. 2013, available on arXiv:1301.2138.

[36] 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. 315 – 328, Jan. 2013.

[37] X. Yi, S. Yang, D. Gesbert, and M. Kobayashi, "The degrees of freedom region of temporally-correlated MIMO networks with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 494 – 514, Jan. 2014.

[38] J. Chen and P. Elia, "Toward the performance vs. feedback tradeoff for the two-user MISO broadcast channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8336 – 8356, Dec. 2013.

[39] J. Chen and P. Elia, "Can imperfect delayed CSIT be as useful as perfect delayed CSIT? DoF analysis and constructions for the BC," in *Proc. Allerton Conf. Communication, Control and Computing*, Oct. 2012.

[40] J. Chen and P. Elia, "Degrees-of-freedom region of the MISO broadcast channel with general mixed-CSIT," May 2012, *available on arXiv:1205.3474v1*.

[41] J. Chen and P. Elia, "How much and when to feedback?" *(invited paper), Information Theory Applications (ITA 2013), La Jolla, California*.

[42] J. Chen, R. Knopp and P. Elia, "Interference alignment for achieving both full DOF and full diversity in the broadcast channel with delayed CSIT," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Jul. 2012.

[43] J. Chen and P. Elia, "Degrees-of-freedom region of the MISO broadcast channel with general mixed-CSIT," in *Proc. Inf. Theory and App. Workshop (ITA)*, Feb. 2013, (also available on arXiv:1205.3474, May 2012).

[44] J. Chen and P. Elia, "MISO broadcast channel with delayed and evolving CSIT," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2013.

[45] J. Chen and P. Elia, "MIMO BC with imperfect and delayed channel state information at the transmitter and receivers," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2013.

[46] J. Chen, S. Yang, and P. Elia, "On the fundamental feedback-vs-performance tradeoff over the MISO-BC with imperfect and delayed CSIT," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2013.

[47] J. Chen and P. Elia, "Symmetric two-user MIMO BC and IC with evolving feedback," Jun. 2013, available on arXiv:1306.3710.

Thank you

Thank you

MIMO Broadcast and Interference Channels towards 5G

Dirk T.M. Slock

with the help of

Francesco Negro, Irfan Ghauri[†], Yohan Lejosne, Yi Yuan[#]

Mobile Communication Department, EURECOM

[†](formerly) Intel Mobile Communications, Sophia Antipolis

[#]Orange Labs, Issy-les-Moulineaux

WCNC2014, Istanbul, April 06, 2014

Some Lessons Learned in Wireless Com

- SDMA (Spatial Division Multiple Access) why did it not take off in the early 90's?
  - No cross-layer design (proper scheduling) at that time.
  - Only feedback was for Power Control.
- At the start of 3G+ activities: it was said that no new PHY development was required, only integration of existing systems. What happened? A lot of PHY work! Dimensions of multi-antenna, multi-user and increasing bandwidth (equalization) were underestimated.
- Wireless standardization starts with the PHY layer. Should become more crosslayer though.
- User Selection: $\Rightarrow$ diversity, simplified transceiver designs.
- Channel Feedback: the return of analog transmission?
- Smart phones: location information everywhere.

Outline

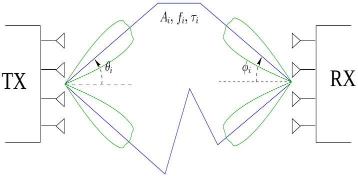
- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - uplink/downlink(UL/DL) duality; SU MIMO,BC,MAC; BF&DPC
  - BC with user selection: DPC vs BF
  - Max WSR, UL/DL duality, CSIT: perfect, partial, LoS
- interference multi-cell/HetNets: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
  - distributed Channel State Information at the Transmitter (CSIT) acquisition, netDoF
  - Delayed CSIT, optimal handling of CSIT FB dead times
  - Finite Rate of Innovation (FRoI)/Basis Expansion Model (BEM) channel models
  - Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels (IBC)

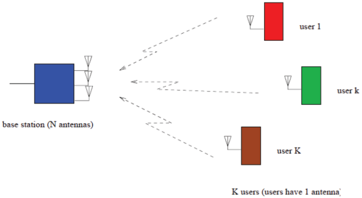
Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - uplink/downlink(UL/DL) duality; SU MIMO,BC,MAC; BF&DPC
  - BC with user selection: DPC vs BF
  - Max WSR, UL/DL duality, CSIT: perfect, partial, LoS
- interference multi-cell/HetNets: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
  - distributed Channel State Information at the Transmitter (CSIT) acquisition, netDoF
  - Delayed CSIT, optimal handling of CSIT FB dead times
  - Finite Rate of Innovation (FRoI)/Basis Expansion Model (BEM) channel models
  - Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels (IBC)

SDMA considerations

- Whereas single user (SU) MIMO communications represented a big breakthrough and are now integrated in a number of wireless communication standards, the next improvement is indeed multi-user MIMO (MU MIMO).
- This topic is nontrivial as e.g. illustrated by the fact that standardization bodies were not able to get an agreement on the topic until recently to get it included in the LTE-A standard.
- MU MIMO is a further evolution of SDMA, which was THE hot wireless topic throughout the nineties.





MU MIMO key elements

- SDMA is a suboptimal approach to MU MIMO, with transmitter precoding limited to linear beamforming, whereas optimal MU MIMO requires Dirty Paper Coding (DPC).
- Channel feedback has gained much more acceptance, leading to good Channel State Information at the Transmitter (CSIT), a crucial enabler for MU MIMO, whereas SDMA was either limited to TDD systems (channel CSIT through reciprocity) or Covariance CSIT. In the early nineties, the only feedback that existed was for slow power control.
- Since SDMA, the concepts of multiuser diversity and user selection have emerged and their impact on the MU MIMO sum rate is now well understood. Furthermore, it is now known that user scheduling allows much simpler precoding schemes to be close to optimal.

MU MIMO key elements (2)

- Whereas SU MIMO allows to multiply transmission rate by the spatial multiplexing factor, when mobile terminals have multiple antennas, MU MIMO allows to reach this same gain with single antenna terminals.
- Whereas in SU MIMO, various degrees of CSIT only lead to a variation in coding gain (the constant term in the sum rate), in MU MIMO however CSIT affects the spatial multiplexing factor (= Degrees of Freedom (DoF)) (multiplying the log(SNR) term in the sum rate).

CSIT considerations

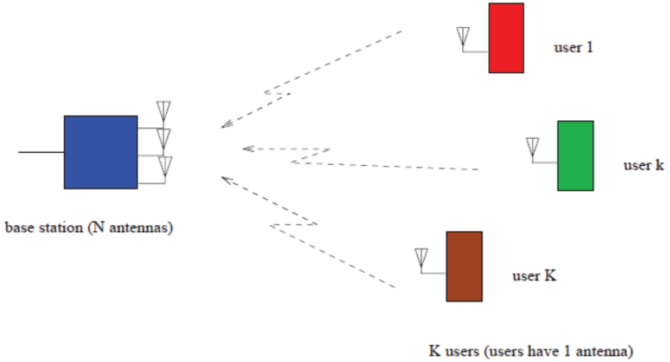
In the process attempting to integrate MU-MIMO into the LTE-A standard, a number of LTE-A contributors had recently become extremely sceptical about the usefulness of the available MU-MIMO proposals. The isue is that they currently do MU-MIMO in the same spirit as SU-MIMO, i.e. with feedback of CSI limited to just a few bits! However, MU-MIMO requires very good CSIT! Some possible solutions:

- Increase CSI feedback enormously (possibly using analog transmission).
- Exploit channel reciprocity in TDD (electronics calibration issue though).
- Limit MU-MIMO to LOS users and extract essential CSIT from DoA or location information.

SDMA system model

- Rx signal at user k:

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_k, \qquad k = 1, 2, \ldots, K$$



Utility Functions

- single user (MIMO) in Gaussian noise: Gaussian signaling optimal (avg. power constr.)
- rate stream k : $R_k = \ln(1 + \text{SINR}_k)$
- SINR balancing: $\max_{BF} \min_k \text{SINR}_k/\gamma_k$ under Tx power P, fairness
- related: min Tx power under $\text{SINR}_k \geq \gamma_k$ GREEN
- max Weighted Sum Rate (WSR): $\max_{BF} \sum_k u_k R_k$, given P

  weights u_k may reflect state of queues (to minimize queue overflow)

  weights also allow to vary orientation of normal to Pareto boundary of rate region and hence to explore whole Pareto boundary if rate region convex
  Pareto boundary: cannot increase an R_k without decreasing some R_i.

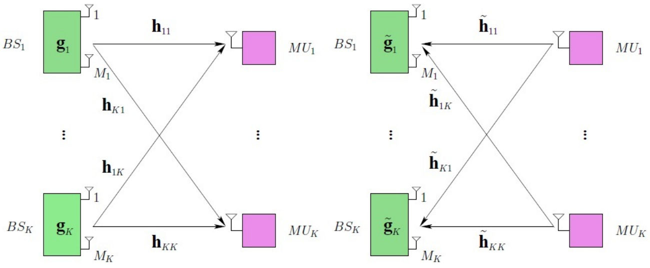
Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - uplink/downlink(UL/DL) duality; SU MIMO,BC,MAC; BF&DPC
  - BC with user selection: DPC vs BF
  - Max WSR, UL/DL duality, CSIT: perfect, partial, LoS
- interference multi-cell/HetNets: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
  - distributed Channel State Information at the Transmitter (CSIT) acquisition, netDoF
  - Delayed CSIT, optimal handling of CSIT FB dead times
  - Finite Rate of Innovation (FRoI)/Basis Expansion Model (BEM) channel models
  - Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels (IBC)

MISO Interference Channel

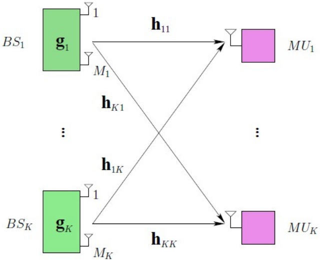
- K pairs of multiantenna Base Station (BS) and single antenna Mobile User (MU)
- BS number k is equipped with M_k antennas
- $\mathbf{g}_k$ ($\tilde{\mathbf{g}}_k$) is the beamformer (RX filter) applied at the k-th BS in DL (UL) transmission
- y_k is Rx signal at the k-th MU in the DL phase,
  $\tilde{r}_k$ is output of Rx filter at the k-th BS in the UL phase:

$$y_k = \mathbf{h}_{kk}\mathbf{g}_k s_k + \sum_{\substack{l=1 \\ l\neq k}}^{K} \mathbf{h}_{kl}\mathbf{g}_l s_l + n_k \qquad \tilde{r}_k = \tilde{\mathbf{g}}_k \tilde{\mathbf{h}}_{kk}\tilde{s}_k + \sum_{\substack{l=1 \\ l\neq k}}^{K} \tilde{\mathbf{g}}_k \tilde{\mathbf{h}}_{kl}\tilde{s}_l + \tilde{\mathbf{g}}_k \tilde{n}_k$$



a) MISO DL　　b) SIMO UL

- MISO DL IFC



- The SINR for the DL channel is:

$$SINR_k^{DL} = \frac{p_k \mathbf{g}_k^H \mathbf{h}_{kk}^H \mathbf{h}_{kk} \mathbf{g}_k}{\sum_{l \neq k} p_l \mathbf{g}_l^H \mathbf{h}_{kl}^H \mathbf{h}_{kl} \mathbf{g}_l + \sigma^2}$$

p_k is the TX power at the k-th BS.

- Imposing a set of DL SINR constraints at each mobile station: $SINR_k^{DL} = \gamma_k$ we obtain in matrix notation:

$$\mathbf{\Phi p} + \boldsymbol{\sigma} = \mathbf{D}^{-1}\mathbf{p}$$

with:

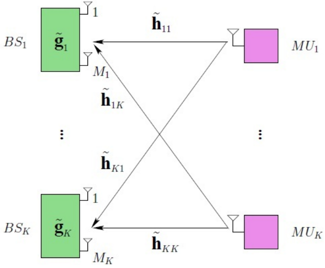
$$[\mathbf{\Phi}]_{ij} = \begin{cases} \mathbf{g}_j^H \mathbf{h}_{ij}^H \mathbf{h}_{ij} \mathbf{g}_j, & j \neq i \\ 0, & j = i \end{cases}$$

$$\mathbf{D} = \mathrm{diag}\{\frac{\gamma_1}{\mathbf{g}_1^H \mathbf{h}_{11}^H \mathbf{h}_{11} \mathbf{g}_1}, \ldots, \frac{\gamma_K}{\mathbf{g}_K^H \mathbf{h}_{KK}^H \mathbf{h}_{KK} \mathbf{g}_K}\}.$$

- We can determine the TX power solving w.r.t. $\mathbf{p}$ obtaining:

$$\mathbf{p} = (\mathbf{D}^{-1} - \mathbf{\Phi})^{-1} \boldsymbol{\sigma} \tag{1}$$

- SIMO UL IFC



- Assuming that $\tilde{\mathbf{h}}_{ij} = \mathbf{h}_{ji}^H$ and $\tilde{\mathbf{g}}_i = \mathbf{g}_i^H$
  the SINR for the UL channel can be written as:

$$SINR_k^{UL} = \frac{q_k \mathbf{g}_k^H \mathbf{h}_{kk}^H \mathbf{h}_{kk} \mathbf{g}_k}{\mathbf{g}_k^H (\sum_{l \neq k} q_l \mathbf{h}_{lk}^H \mathbf{h}_{lk} + \sigma^2 \mathbf{I}) \mathbf{g}_k}$$

q_k represents the Tx power from the k-th MS.

- Imposing the same SINR constraints also in the UL: $SINR_k^{UL} = \gamma_k$ it is possible to rewrite that constraints as:

$$\tilde{\boldsymbol{\Phi}}\mathbf{q} + \boldsymbol{\sigma} = \mathbf{D}^{-1}\mathbf{q}$$

with:

$$[\tilde{\boldsymbol{\Phi}}]_{ij} = \begin{cases} \mathbf{g}_i^H \mathbf{h}_{ji}^H \mathbf{h}_{ji} \mathbf{g}_i, & j \neq i \\ 0, & j = i \end{cases}$$

$$\mathbf{D} = \text{diag}\{\frac{\gamma_1}{\mathbf{g}_1^H \mathbf{h}_{11}^H \mathbf{h}_{11} \mathbf{g}_1}, \ldots, \frac{\gamma_K}{\mathbf{g}_K^H \mathbf{h}_{KK}^H \mathbf{h}_{KK} \mathbf{g}_K}\}.$$

- The power vector can be found as:

$$\mathbf{q} = (\mathbf{D}^{-1} - \tilde{\boldsymbol{\Phi}})^{-1}\boldsymbol{\sigma} \qquad (2)$$

- Comparing the definition we can see that $\tilde{\mathbf{\Phi}} = \mathbf{\Phi}^T$. This implies that there exists a duality relationship between the DL MISO and UL SIMO IFCs.
- We can extend the results for UL-DL duality for MAC/BC [Schubert & Boche'04] to the MISO/SIMO IFC:

Targets $\gamma_1, \ldots, \gamma_K$ are jointly feasible in UL and DL if and only if the spectral radius ρ of the weighted coupling matrix satisfies $\rho(\mathbf{D\Phi}) < 1$.

Both UL and DL have the same SINR feasible region under a sum-power constraint, i.e., target SINRs are feasible in the DL if and only if the same targets are feasible in the UL:

$$\sum_i q_i = \mathbf{1}^T\mathbf{q} = \sigma\mathbf{1}^T(\mathbf{D}^{-1} - \mathbf{\Phi}^T)^{-T} = \sigma\mathbf{1}^T(\mathbf{D}^{-1} - \mathbf{\Phi})^{-1} = \sum_i p_i \quad (3)$$

- Using this results it is possible to extend some BF design techniques used in the BC [Schubert & Boche'04] to the MISO IFC:
  - Max-Min SINR (SINR Balancing)
  - Power minimization under SINR constraints

- Even though the sum power constraint is analytically attractive such constraint is not enough in a practical IFC where each user is subject to a per user power constraint.

- The BF design problem now becomes:

$$\begin{aligned} &\min_{\mathbf{g}_1,\dots,\mathbf{g}_K} \textstyle\sum_{k=1}^{K} \mathbf{g}^H \mathbf{g} && \\ &\text{s.t.} \quad \mathbf{g}_k^H \mathbf{g}_k \leq P_k && k = 1,\dots,K \\ &\quad SINR_k^{DL} = \frac{\mathbf{g}_k^H \mathbf{h}_{kk}^H \mathbf{h}_{kk} \mathbf{g}_k}{\sum_{l \neq k} \mathbf{g}_l^H \mathbf{h}_{kl}^H \mathbf{h}_{kl} \mathbf{g}_l + \sigma_k^2} \geq \gamma_k && k = 1,\dots,K \end{aligned}$$

P_k represents the maximum Tx power for user k.

- The Lagrangian of the DL optimization problem is:

$$\mathcal{L}(\lambda_i, \mu_i, \mathbf{g}_i) = \sum_{i=1}^{K} \mathbf{g}^H \mathbf{g} + \sum_{i=1}^{K} \mu_i [\mathbf{g}_i^H \mathbf{g}_i - P_i] + \sum_{i=1}^{K} \lambda_i [-\frac{1}{\gamma_i} \mathbf{g}_i^H \mathbf{h}_{ii}^H \mathbf{h}_{ii} \mathbf{g}_i + \sum_{l \neq i} \mathbf{g}_l^H \mathbf{h}_{il}^H \mathbf{h}_{il} \mathbf{g}_l + \sigma_i^2]$$

- λ_k Lagrange multiplier of the k-th SINR constraint
- μ_k Lagrange multiplier associated to the Tx power constraint at user k.
- The Lagrange Dual problem is:

$$\begin{aligned} &\max_{\lambda_1,\ldots,\lambda_K,\mu_1,\ldots,\mu_K,} \textstyle\sum_{k=1}^{K} \lambda_k \sigma_k^2 - \displaystyle\sum_{i=1}^{K} \mu_i P_i \\ &\text{s.t.} \quad -\tfrac{\lambda_k}{\gamma_k} \mathbf{h}_{kk}^H \mathbf{h}_{kk} + \textstyle\sum_{l \neq k} \lambda_l \mathbf{h}_{lk}^H \mathbf{h}_{lk} + (\mu_k + 1)\mathbf{I} \succeq 0 \quad k = 1, \ldots, K \end{aligned} \tag{4}$$

Strong duality holds between the original problem and the Lagrange dual.

- A phase rotation in the optimal BF vectors does not influence the SINRs
- $\mathbf{g}_k e^{j\phi_k}$ we choose ϕ_k s.t $\mathbf{h}_{kk}\mathbf{g}_k \in \mathbb{R}$
- The SINR constraint for the k-th user reads:

$$(1+\frac{1}{\gamma_k})\mathbf{g}_k^H\mathbf{h}_{kk}^H\mathbf{h}_{kk}\mathbf{g}_k \geq \sum_{i=1}^{K}\mathbf{g}_i^H\mathbf{h}_{ki}^H\mathbf{h}_{ki}\mathbf{g}_i + \sigma_k^2 \quad\longmapsto\quad (1+\frac{1}{\gamma_k})|\mathbf{h}_{kk}\mathbf{g}_k|^2 \geq \|\mathbf{H}_k\mathbf{G}\ \sigma_k\|^2$$

where $\mathbf{H}_k = [\mathbf{h}_{k1},\ldots,\mathbf{h}_{kK}]$ and $\mathbf{G} = diag\{\mathbf{g}_1,\ldots,\mathbf{g}_K\}$

- The original problem now becomes:

$$\begin{aligned} \min_{\mathbf{g}_1,\ldots,\mathbf{g}_K} \quad & \textstyle\sum_{k=1}^{K}\mathbf{g}^H\mathbf{g} \\ \text{s.t.} \quad & \mathbf{g}_k^H\mathbf{g}_k \leq P_k \qquad\qquad k=1,\ldots,K \qquad (5) \\ \sqrt{1+\tfrac{1}{\gamma_k}}\,\mathbf{h}_{kk}\mathbf{g}_k \geq & \|\mathbf{H}_k\mathbf{G}\ \sigma_k\| \quad k=1,\ldots,K. \end{aligned}$$

- The SINR constraint becomes a convex SOCP constraint.

The optimal solution of the dual problem is also optimal for the original

- The Lagrange dual of the DL beamforming problem can be rewritten as an equivalent UL optimization problem:

$$\max_{\lambda_1,\ldots,\lambda_K,\mu_1,\ldots,\mu_K,} \sum_{k=1}^{K} \lambda_k \sigma_k^2 - \sum_{i=1}^{K} \mu_i P_i$$
$$SINR_k^{UL} = \frac{\lambda_k \tilde{\mathbf{g}}_k^H \mathbf{h}_{kk}^H \mathbf{h}_{kk} \tilde{\mathbf{g}}_k}{\tilde{\mathbf{g}}_k^H (\sum_{l\neq k} \lambda_l \mathbf{h}_{lk}^H \mathbf{h}_{lk} + \eta_k \mathbf{I}) \tilde{\mathbf{g}}_k} \leq \gamma_k \quad k = 1, \ldots, K \qquad (6)$$

- The dual Tx power λ_k and the dual noise power $\eta_k = 1 + \mu_k$ are to be optimized.
- The optimal UL Rx filter is: $\tilde{\mathbf{g}}_k = (\sum_{l\neq k} \lambda_l \mathbf{h}_{lk}^H \mathbf{h}_{lk} + \eta_k \mathbf{I})^{-1} \mathbf{h}_{kk}^H \lambda_k$
- At the optimum the SINR constraints must be satisfied with equality.
- The optimal DL BFs are given:

$$\mathbf{g}_k = \sqrt{\beta_k}\tilde{\mathbf{g}}_k \qquad \boldsymbol{\beta} = \mathbf{D}^{-1}\boldsymbol{\sigma} \quad [\mathbf{D}]_{ij} = \begin{cases} \frac{1}{\gamma_i} \tilde{\mathbf{g}}_i^H \mathbf{h}_{ii}^H \mathbf{h}_{ii} \tilde{\mathbf{g}}_i & i = j \\ -\tilde{\mathbf{g}}_j^H \mathbf{h}_{ij}^H \mathbf{h}_{ij} \tilde{\mathbf{g}}_j & i \neq j \end{cases} \qquad (7)$$

MISO DL BF Design Algorithm

Algorithm 1 Beamformer Design via UL-DL duality

Initialize: $i = 0$, $\lambda_k^{(0)} = 1, \mu_k^{(0)} = 1, \forall k = 1, \ldots, K$

repeat

$i = i + 1$

For $k = 1, \ldots, K$ find the UL receiver filter as

$$\tilde{\mathbf{g}}_k^{(i)} = \Big(\sum_{l \neq k}^{K} \lambda_l^{(i-1)} \mathbf{h}_{lk}^H \mathbf{h}_{lk} + \eta_k^{(i-1)} \mathbf{I}\Big)^{-1} \mathbf{h}_{kk}^H \lambda_k^{(i-1)}$$

Determine $\lambda_k^{(i)}$ as: $\lambda_k^{(i)} = \gamma_k \frac{\tilde{\mathbf{g}}_k^{(i)H} (\sum_{l \neq k} \lambda_l^{(i-1)} \mathbf{h}_{lk}^H \mathbf{h}_{lk} + \eta_k^{(i-1)} \mathbf{I}) \tilde{\mathbf{g}}_k^{(i)}}{\tilde{\mathbf{g}}_k^{(i)H} \mathbf{h}_{kk}^H \mathbf{h}_{kk} \tilde{\mathbf{g}}_k^{(i)}}$

Determine the optimal DL BF $\mathbf{g}_k^{(i)}$ using (7)

Update the matrix $\mathbf{M}^{(i)} = diag\{\mu_1^{(i)}, \ldots, \mu_K^{(i)}\}$ using the subgradient method with step size $t^{(i)}$

$$\mathbf{M}^{(i)} = [\mathbf{M}^{(i-1)} + t^{(i)} \mathbf{Q}^{(i)}]_+$$

where $\mathbf{Q}^{(i)} = diag\{\mathbf{g}_1^{(i)H} \mathbf{g}_1^{(i)}, \ldots, \mathbf{g}_K^{(i)H} \mathbf{g}_K^{(i)}\} - diag\{P_1, \ldots, P_K\}$

until convergence

Tx determination and UL/DL Duality (BC/MAC)

- beautifully explained in [ViswanathTse:T-ITaug03]
- start from SU MIMO channel H w stream Tx & Rx filters G, F and SINRs: $(FHG)^H = G^H H^H F^H$, UL/DL duality for any filters and SINRs, same power feasibility and sum power constraint (and SU MIMO: Gaussian signaling)
- SIMO MAC (Multiple Access Channel) (MU UL) = special case of SU MIMO with $G = I_K$,
  MAC SR=$\ln\det(I + HDH^H)$, $\text{tr}\{D\} = P$, D = diagonal
  Rx = stripping (successive interference cancellation and LMMSE)
- MISO BC (Broadcast Channel) (MU DL) = special case of SU MIMO with $F = I_K$,
  duality (same rates, SINRs) for BC/MAC with same Tx/Rx filters and same (sum) power constraint

Tx determination and UL/DL Duality (BC/MAC) (2)

- Costa: $y = x + s + v$, s known to Tx, has same capacity as $y = x + v$, Dirty Paper Coding (DPC)
- Costa rate region of MISO BC = rate region of SIMO MAC w stripping = MISO rate region lower bound
- Sato upper bound: rate region of BC is upper bounded by that of corresponding SU MIMO (Rx's cooperate)
- Observe: difference between "corresponding" MISO BC and SU MIMO: consider SU MIMO with spatially colored noise covariance matrix, only its diagonal elements count in MISO BC.

  Can show that there exists a noise covariance matrix for which coperation between Rx's does not help (via UL/DL relation).

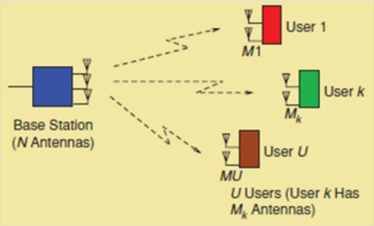
  Hence: Costa lower bound reaches Sato upper bound and hence BC rate region = MAC rate region with sum power constraint.
- Can be immediately extended to MIMO BC and MIMO MAC.
- DPC in "practice": Tomlinson-Harashima (TH), Vector Precoding (VP = vector TH)

Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - uplink/downlink(UL/DL) duality; SU MIMO,BC,MAC; BF&DPC
  - BC with user selection: DPC vs BF
  - Max WSR, UL/DL duality, CSIT: perfect, partial, LoS
- interference multi-cell/HetNets: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
  - distributed Channel State Information at the Transmitter (CSIT) acquisition, netDoF
  - Delayed CSIT, optimal handling of CSIT FB dead times
  - Finite Rate of Innovation (FRoI)/Basis Expansion Model (BEM) channel models
  - Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels (IBC)

MIMO BroadCast

- MIMO BC = Multi-User MIMO Downlink
- N_t transmission antennas.
- K users with N_k receiving antennas.
- Assume perfect CSI
- Possibly multiple streams/user d_k.
- Power constraint P
- Noise variance $\sigma^2 = 1$.
- $\mathbf{H}_k$ the MIMO channel for user k.



$$\begin{aligned}\mathbf{F}_k\mathbf{y}_k &= \mathbf{F}_k\mathbf{H}_k \textstyle\sum_{i=1}^{K} \mathbf{G}_i\mathbf{s}_i + \mathbf{F}_k\mathbf{z}_k \\ &= \underbrace{\mathbf{F}_k\mathbf{H}_k\mathbf{G}_k\mathbf{s}_k}_{\text{useful signal}} + \underbrace{\sum_{i=1, i\neq k}^{K} \mathbf{F}_k\mathbf{H}_k\mathbf{G}_i\mathbf{s}_i}_{\text{inter-user interference}} + \underbrace{\mathbf{F}_k\mathbf{z}_k}_{\text{noise}}\end{aligned}$$

System Model (2)

- Rx signal: $\mathbf{y}_k = \mathbf{H}_k\mathbf{x} + \mathbf{z}_k = \mathbf{H}_k \sum_{i=1}^{K} \mathbf{G}_i\mathbf{s}_i + \mathbf{z}_k$
- $$\underbrace{\mathbf{F}_k}_{d_k\times N_k} \underbrace{\mathbf{y}_k}_{N_k\times 1} = \underbrace{\mathbf{F}_k}_{d_k\times N_k} \underbrace{\mathbf{H}_k}_{N_k\times N_t} \sum_{i=1}^{K} \underbrace{\mathbf{G}_i}_{N_t\times d_i} \underbrace{\mathbf{s}_i}_{d_i\times 1} + \underbrace{\mathbf{F}_k}_{d_k\times N_k} \underbrace{\mathbf{z}_k}_{N_k\times 1}$$
- [Christensen etal:T-WC08]: use of linear receivers in MIMO BC is not suboptimal (full CSIT, // SU MIMO): can prefilter $\mathbf{G}_k$ with a $d_k \times d_k$ unitary matrix to make interference plus noise prewhitened channel matrix - precoder cascade of user k orthogonal (columns)

User Selection Motivation

- Optimal MIMO BC design requires DPC, which is significantly more complicated than BF.
- User selection allows to
  - improve the rates of DPC
  - bring the rate of BF close to those of DPC
- Optimal user/stream selection requires selection of optimal combination of N_t streams: too complex. Greedy user/stream selection (GUS): select one stream at a time $\Rightarrow$ complexity $\approx$ N_t times the complexity of selecting one stream ($K \gg N_t$).
- Multiple receive antennas cannot improve the sum rate prelog. So what benefit can they bring?
  Of course: cancellation of interference from other transmitters (spatially colored noise): not considered here.

Zero-Forcing (ZF)

- ZF-BF $\mathbf{F}_{1:i}\mathbf{H}_{1:i}\mathbf{G}_{1:i} =$

$$\begin{bmatrix} \mathbf{F}_1 & 0 \cdots 0 \\ 0 & \mathbf{F}_2 \ddots \vdots \\ \vdots & \ddots 0 \\ 0 \cdots & 0\ \mathbf{F}_i \end{bmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_i \end{bmatrix} [\mathbf{G}_1\ \mathbf{G}_2 \cdots \mathbf{G}_i] = \begin{bmatrix} \mathbf{F}_1\mathbf{H}_1\mathbf{G}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{F}_2\mathbf{H}_2\mathbf{G}_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{F}_i\mathbf{H}_i\mathbf{G}_i \end{bmatrix}$$

- ZF-DPC (modulo reordering issues) $\mathbf{F}_{1:i}\mathbf{H}_{1:i}\mathbf{G}_{1:i} =$

$$\begin{bmatrix} \mathbf{F}_1 & 0 \cdots 0 \\ 0 & \mathbf{F}_2 \ddots \vdots \\ \vdots & \ddots 0 \\ 0 \cdots & 0\ \mathbf{F}_i \end{bmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_i \end{bmatrix} [\mathbf{G}_1\ \mathbf{G}_2 \cdots \mathbf{G}_i] = \begin{bmatrix} \mathbf{F}_1\mathbf{H}_1\mathbf{G}_1 & 0 & \cdots & 0 \\ * & \mathbf{F}_2\mathbf{H}_2\mathbf{G}_2 & & \vdots \\ \vdots & & \ddots & 0 \\ * & \cdots & * & \mathbf{F}_i\mathbf{H}_i\mathbf{G}_i \end{bmatrix}$$

- BF-style selection, DPC-style selection: as if it's going to be used in BF/DPC

Stream Selection Criterion from Sum Rate

- At high SNR, both
  - optimized (MMSE style) filters vs. ZF filters
  - optimized vs. uniform power allocation

  only leads to $\frac{1}{\text{SNR}}$ terms in rates.
- At high SNR, the sum rate is of the form

$$\underbrace{N_t}_{\text{DoF}} \log(\text{SNR}/N_t) + \underbrace{\sum_i \log\det(\mathbf{F}_i \mathbf{H}_i \mathbf{G}_i)}_{\text{constant}} + O(\frac{1}{\text{SNR}}) + \underbrace{O(\log\log(\text{SNR}))}_{\text{noncoherent Tx}}$$

for properly normalized ZF Rx $\mathbf{F}_i$ and ZF Tx $\mathbf{G}_i$ (BF or DPC).

MIMO BC Greedy Stream Selection (GUS) Criteria

One can focus on either

(i) (the constant term in) the sum rate at high SNR,

(ii) the sum rate at any SNR of the associated ZF transceiver designs with uniform power loading,

(iii) the sum rate at any SNR of the associated ZF transceiver designs with waterfilling,

(iv) the sum rate at any SNR of optimized transceiver designs.

At high SNR, (i) is the analysis of interest.
More variations could be considered, e.g. regularized ZF as an intermediate between ZF and optimized transceiver designs.

State of the Art: MISO BC

- [TuBlum:COMlet99]: Gram-Schmidt channel orthogonalization with pivoting (DPC-style GUS)
- [DimicSidiropoulos:T-SP05]: introduced proper BF-style GUS, large K analysis DPC-style GUS, simulations. Matrix inversion lemma for bordered matrices, in order to lower complexity of BF-style GUS.
- [YooGoldsmith:06]: analyze BF, but w pseudo-BF-style GUS: SUS (semi-orthogonal) = DPC-style GUS + inner product constraints (limiting size of pool of users for selection). Show that for BF-SUS, as for DPC-US,

$$\lim_{K\to\infty} \frac{SR}{N_t \log(1 + \frac{P}{N_t} \log K)} = 1$$

- [WangLoveZoltowski:08]: small refinement of [YooGoldsmith:06], has more constraints.

State of the Art: MIMO BC

- [SunMcKay:10]: transforms MIMO to MISO. [WangLoveZoltowski:08] type analysis. Pseudo-BF-style GUS (SUS). Analysis for use in DPC and in BF. Analysis only shows effect of antennas in higher-order terms.
- [Jindal:TWC08]: single stream MIMO BC, use of Rx antennas to minimize quantization error (for FB) on resulting virtual channel. Emphasis on partial CSIT (and CSIR) with (G)US.
- [HungerJoham:ciss09]: obtain the high SNR SR offset between BF and DPC, without user selection. They extend the [Jindal:isit05] analysis from MISO to MIMO. [Hassibi] also.
- [Utschick:06] SESAM: proper DPC-style GUS for MIMO case (extension of [TuBlum] from MISO to MIMO).

State of the Art: MIMO BC (2)

- [GuthyUtschick:1209]: propose a BF-style GUS for MIMO-BC-BF. In the style of predecessors, they only adapt the Rx of the new stream to be added. They replace the proper geometric average of the stream channel powers by its harmonic average: $1/\text{tr}\{diag((HH^H)^{-1})\}$, which leads to a generalized eigenvector solution for the Rx filter (minFrob algo). Can be simplified to a classical eigenvector problem: the LISA algo = SESAM algo. No asymptotic analysis.
- [GuthyUtschick:T-SP0410]: same greedy approaches are proposed now for max WSR, without user selection.
- [reference in Yu?]: introducing more than one sweep in GUS.
- [Christensen:TW08]: working per stream is equivalent to working per user.
- many other references on MIMO BF design for max (W)SR

Asymptotic Regimes

Assume for simplicity all $N_k = N_r$. Possible asymptotic regimes for analyzing US:

(0) $K \to \infty$, M, N finite.

(1) $M \to \infty$, $N = \alpha M$, α, K finite. In this regime, one lets $M, N \to \infty$ and then one introduces selection mechanisms.

(2) $M \to \infty$, $N = \alpha M$, α, K finite. In this regime, one first introduces selection mechanisms and then one lets $M, N \to \infty$.

(3) $M \to \infty$, $K = \beta M$, β, N finite.

Regime (0) is the classical asymptotic regime for the analysis of the effect of stream selection. However, the results of this analysis are only relevant when K is extremely large.

Asymptotic Regimes (2)

The behavior of various stream selection mechanisms is more interesting when there is some redundancy allowing selection, but not too much. In order to benefit from simplified asymptotic analytical results, consider $M \to \infty$. Some room for selection is then obtained when $KN \to \infty$ also, but at the same rate as M with $KN > M$.

One way for $KN \to \infty$ is to keep K finite and let $N \to \infty$ as in single-user MIMO asymptotic analysis. The correct selection analysis corresponds to regime (2), whereas regime (1) is a simplified approximation of (2), and is considered in [GuthyUtschickHonig:isit10]. Indeed, in this case all user channels behave identically asymptotically, and hence the selection process becomes very simple.

Another way for $KN \to \infty$ is to keep N finite and let $K \to \infty$. Such analysis would also encompass the MISO case.

From MISO to MIMO: Rx design

Design of receivers:

(i) Fixed unitary receiver. For i.i.d. channels, any fixed unitary Rx is equivalent, hence can choose an identity matrix, in which case

$$\text{stream selection} = \text{Rx antenna selection.}$$

The analysis of the corresponding algorithms is very simple since in this case K users MIMO with N Rx antennas is equivalent to MISO with KN users.

(ii) Rx for user k is optimized only in function of its channel $\mathbf{H}_k$. E.g. [SunMcKay:T-SP10]: singular modes of $\mathbf{H}_k$. Rx fixed $\Rightarrow$ K user MIMO $= K\,N$ user MISO, but virtual MISO channels no longer i.i.d.

(iii) Optimized receivers (Rx).

MISO DPC-style GUS

- GUS: Greedy User Selection
  Gram-Schmidt orthogonalize $\mathbf{h}_k$ w.r.t. those of already selected users and choose user with maximum residual norm (matched to DPC).
- MISO: $h_k = \mathbf{H}_k^H$, k_i = user selected at stage i, $H_i = h^H_{k_{1:i}}$.
-
$$\det(H_i H_i^H) = \prod_{j=1}^{i} \| P^{\perp}_{h_{k_{1:j-1}}} h_{k_j} \|^2$$
- at stage i: $\quad k_i = \arg\max_k \| P^{\perp}_{h_{k_{1:i-1}}} h_k \|^2$
- Introduce ϕ_i = angle between h_{k_i} and $h_{k_{1:i-1}}$ $\Rightarrow$ can write $\| P^{\perp}_{h_{k_{1:i-1}}} h_{k_i} \|^2 = \| h_{k_i} \|^2 \sin^2 \phi_i$.

MISO BF-style GUS

$$k_i \text{ maximizes } (\det(\mathrm{diag}\{(H_i H_i^H)^{-1}\}))^{-1} =$$
$$\|P^\perp_{h_{k_{1:i-1}}} h_{k_i}\|^2 \prod_{j=1}^{i-1}(\|P^\perp_{h_{k_{1:i-1}\setminus k_j}} h_{k_j}\|^2 - \frac{|h^H_{k_i} P^\perp_{h_{k_{1:i-1}\setminus k_j}} h_{k_j}|^2}{\|P^\perp_{h_{k_{1:i-1}\setminus k_j}} h_{k_i}\|^2})$$

For sufficiently large K, the BF-style user selection process will lead to the selection of channel vectors that are close to being mutually orthogonal. As a result we can write up to first order the contribution of stream i to the sum rate offset

$$\begin{aligned} &\|P^\perp_{h_{k_{1:i-1}}} h_{k_i}\|^2 \prod_{j=1}^{i-1} \sin^2\phi_{ij} \approx \|P^\perp_{h_{k_{1:i-1}}} h_{k_i}\|^2 \sin^2\phi_i \\ &= \|h_{k_i}\|^2 \sin^4\phi_i = \|P^\perp_{h_{k_{1:i-1}}} h_{k_i}\|^4 / \|h_{k_i}\|^2 \,. \end{aligned} \tag{8}$$

DPC offset is $\|P^\perp_{h_{k_{1:i-1}}} h_{k_i}\|^2 = \|h_{k_i}\|^2 \sin^2\phi_i$ = certain compromise between max $\|h_{k_i}\|^2$ and min $\cos^2\phi_i$. In the case of BF, $\|h_{k_i}\|^2 \sin^4\phi_i$ leads to a similar compromise, but with more emphasis on orthogonality.

Role of Rx antennas?

- Different distributions of ZF between Tx and Rx give different ZF channel gains! If Rx ZF's k streams, hence Tx only has to ZF $M-1-k$ streams! So, number of possible solutions (assuming $d_k \equiv 1$):

$$\prod_{k=1}^{M}\left(\sum_{i=0}^{N_k-1}\frac{(M-1)!}{k!(M-1-k)!}\right)$$

  for each user, Rx can ZF k between 0 and N_k-1 streams, to choose among $M-1$.
  Explains non-convexity of MIMO SR at high SNR.
- ZF by Rx can alternatively be interpreted as IA by Tx (Rx adapts Rx-channel cascades to lie in reduced dimension subspace).
- SESAM (and all existing MIMO stream selection algorithms): assumes that all ZF is done by Tx only. Hence, Rx can be a MF, matched to channel-BF cascade.

Concluding Remarks MIMO BC User Selection

- Introduced new MISO BC BF-style GUS criterion/interpretation.
- Extension to MIMO BC with receiver design.
- Plenty of room for asymptotic analysis of transient regime of stream selection.
- Here, did not touch upon CSIT FB issues, user preselection schemes to reduce pool size etc.
- Joint Tx-Rx ZF (IA) provides more opportunities (but hence also larger search space and complexity).

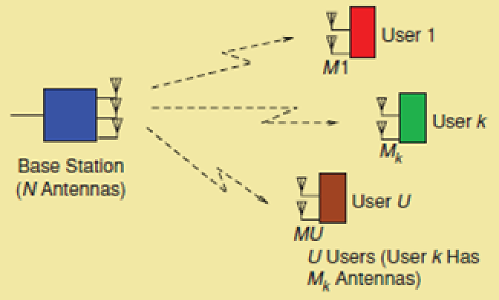
Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - uplink/downlink(UL/DL) duality; SU MIMO,BC,MAC; BF&DPC
  - BC with user selection: DPC vs BF
  - Max WSR, UL/DL duality, CSIT: perfect, partial, LoS
- interference multi-cell/HetNets: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
  - distributed Channel State Information at the Transmitter (CSIT) acquisition, netDoF
  - Delayed CSIT, optimal handling of CSIT FB dead times
  - Finite Rate of Innovation (FRoI)/Basis Expansion Model (BEM) channel models
  - Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels (IBC)

Max WSR Tx BF Design Outline

- max WSR Tx BF design with perfect CSIT
  - using WSR - WSMSE relation
  - from difference of concave to linearized concave
  - MIMO BC: local optima, deterministic annealing
- Gaussian partial CSIT
- max EWSR Tx design w partial CSIT
- Line of Sight (LoS) based partial CSIT
- max EWSR Tx design with LoS based CSIT

MIMO Broadcast Channel (BC) with Linear Tx/Rx



- The $N_k \times 1$ received (Rx) signal at user k is

$$y_k = H_k \, \mathbf{g}_k \, x_k + \sum_{i=1,\neq k}^{K} H_k \, \mathbf{g}_i \, x_i + v_k$$

- x_i is the unit variance scalar Gaussian signal for user i,
- channel H_k has size $N_k \times M$,
- v_k is additive white noise $v_k \sim \mathcal{CN}(0, \sigma^2_{v,k} I_{N_k}), \;\; \sigma^2_{v,k} = 1$.
- $K \leq M$ users, each with a single stream and Transmit (Tx) BeamFormer (BF) $\mathbf{g}_k$.

Max Weighted Sum Rate (WSR)

- Weighted sum rate (WSR)

$$WSR = WSR(\mathbf{g}) = \sum_{k=1}^{K} u_k \ \ln \frac{1}{e_k}$$

where $\mathbf{g} = \{\mathbf{g}_k\}$, the u_k are rate weights

- MMSEs $e_k = e_k(\mathbf{g})$

$$\frac{1}{e_k} = 1 + \mathbf{g}_k^H H_k^H R_{\overline{k}}^{-1} H_k \mathbf{g}_k = (1 - \mathbf{g}_k^H H_k^H R_k^{-1} H_k \mathbf{g}_k)^{-1}$$
$$R_k = R_{\overline{k}} + H_k \mathbf{g}_k \mathbf{g}_k^H H_k^H \ , \ R_{\overline{k}} = \sum_{i \neq k} H_k \mathbf{g}_i \mathbf{g}_i^H H_k^H + I_{N_k} \ ,$$

R_k, $R_{\overline{k}}$ = total, interference plus noise Rx cov. matrices resp.

- MMSE e_k obtained at the output $\widehat{x}_k = f_k^H y_k$ of the optimal (MMSE) linear Rx

$$f_k = R_k^{-1} H_k \mathbf{g}_k \ .$$

From max WSR to min WSMSE

- For a general Rx filter f_k we have the MSE $e_k(f_k, \mathbf{g})$
  $= (1 - f_k^H H_k \mathbf{g}_k)(1 - \mathbf{g}_k^H H_k^H f_k) + \sum_{i \neq k} f_k^H H_k \mathbf{g}_i \mathbf{g}_i^H H_k^H f_k + ||f_k||^2$
  $= 1 - f_k^H H_k \mathbf{g}_k - \mathbf{g}_k^H H_k^H f_k + \sum_i f_k^H H_k \mathbf{g}_i \mathbf{g}_i^H H_k^H f_k + ||f_k||^2.$
- The $WSR(\mathbf{g})$ is a non-convex and complicated function of $\mathbf{g}$. Inspired by [Christensen:TW1208], we introduced [Negro:ita10],[Negro:ita11] an augmented cost function, the Weighted Sum MSE, $WSMSE(\mathbf{g}, f, w)$

$$= \sum_{k=1}^{K} u_k (w_k \; e_k(f_k, \mathbf{g}) - \ln w_k) + \lambda (\sum_{k=1}^{K} ||\mathbf{g}_k||^2 - P)$$

  where λ = Lagrange multiplier and P = Tx power constraint.
- After optimizing over the aggregate auxiliary Rx filters f and weights w, we get the WSR back:

$$\min_{f,w} WSMSE(\mathbf{g}, f, w) = -WSR(\mathbf{g}) + \overbrace{\sum_{k=1}^{K} u_k}^{\text{constant}}$$

From max WSR to min WSMSE (2)

- Advantage augmented cost function: alternating optimization $\Rightarrow$ solving simple quadratic or convex functions

$$\min_{w_k} WSMSE \;\Rightarrow\; w_k = 1/e_k$$

$$\min_{f_k} WSMSE \;\Rightarrow\; f_k = (\sum_i H_k \mathbf{g}_i \mathbf{g}_i^H H_k^H + I_{N_k})^{-1} H_k \mathbf{g}_k$$

$$\min_{\mathbf{g}_k} WSMSE \;\Rightarrow$$

$$\mathbf{g}_k = (\textstyle\sum_i u_i w_i H_i^H f_i f_i^H H_i + \lambda I_M)^{-1} H_k^H f_k u_k w_k$$

- UL/DL duality: optimal Tx filter g_k of the form of a MMSE linear Rx for the dual UL in which λ plays the role of Rx noise variance and $u_k w_k$ plays the role of stream variance.

From WSR to WSSINR

- The WSR can be rewritten as

$$WSR = WSR(\mathbf{g}) = \sum_{k=1}^{K} u_k \ \ln(1 + \mathrm{SINR}_k)$$

where $1 + \mathrm{SINR}_k = 1/e_k$ or for general f_k :

$$\mathrm{SINR}_k = \frac{|f_k H_k \mathbf{g}_k|^2}{\sum_{i=1,\neq k}^{K} |f_k H_k \mathbf{g}_i|^2 + ||f_k||^2} \quad .$$

- WSR variation

$$\partial WSR = \sum_{k=1}^{K} \frac{u_k}{1 + \mathrm{SINR}_k} \ \partial \mathrm{SINR}_k$$

interpretation: variation of a weighted sum SINR (WSSINR)

- The BFs obtained: same as for WSR or WSMSE criteria. But this interpretation shows: WSR = optimal approach to the SLNR or SJNR heuristics. WSSINR approach = [KimGiannakis:IT0511] below.

Optimal Lagrange Multiplier λ

- (bisection) line search on $\sum_{k=1}^{K} ||\mathbf{g}_k||^2 - P = 0$ [Luo:SP0911].
- Or updated analytically as in [Negro:ita10],[Negro:ita11] by exploiting $\sum_k \mathbf{g}_k^H \frac{\partial WSMSE}{\partial \mathbf{g}_k^*} = 0$.
- This leads to the same result as in [Hassibi:TWC0906]: λ avoided by reparameterizing the BF to satisfy the power constraint: $\mathbf{g}_k = \sqrt{\frac{P}{\sum_{i=1}^{K} ||\mathbf{g}_i^{'}||^2}}\, \mathbf{g}_k^{'}$ with $\mathbf{g}_k^{'}$ now unconstrained

$$\mathrm{SINR}_k = \frac{|f_k H_k \mathbf{g}_k^{'}|^2}{\sum_{i=1,\neq k}^{K} |f_k H_k \mathbf{g}_i^{'}|^2 + \frac{1}{P}||f_k||^2 \sum_{i=1}^{K} ||\mathbf{g}_i^{'}||^2} \ .$$

- This leads to the same Lagrange multiplier expression obtained in [Christensen:TW1208] on the basis of a heuristic that was introduced in [Joham:isssta02] as was pointed out in [Negro:ita10].

[KimGiannakis:IT0511]

- Let $Q_k = \mathbf{g}_k \mathbf{g}_k^H$ be the transmit covariance for stream k $\Rightarrow$

$$WSR = \sum_{k=1}^{K} u_k [\ln\det(R_k) - \ln\det(R_{\overline{k}})]$$

w $R_k = H_k(\sum_i Q_i) H_k^H + I_{N_k}$, $R_{\overline{k}} = H_k(\sum_{i\neq k} Q_i) H_k^H + I_{N_k}$.

- Consider the dependence of WSR on Q_k alone:

$$WSR = u_k \ln\det(R_{\overline{k}}^{-1} R_k) + WSR_{\overline{k}}, \; WSR_{\overline{k}} = \sum_{i=1,\neq k}^{K} u_i \ln\det(R_{\overline{i}}^{-1} R_i)$$

where $\ln\det(R_{\overline{k}}^{-1} R_k)$ is concave in Q_k and $WSR_{\overline{k}}$ is convex in Q_k. Since a linear function is simultaneously convex and concave, consider the first order Taylor series expansion in Q_k around $\widehat{Q}$ (i.e. all $\widehat{Q}_i$) with e.g. $\widehat{R}_i = R_i(\widehat{Q})$, then

$$WSR_{\overline{k}}(Q_k, \widehat{Q}) \approx WSR_{\overline{k}}(\widehat{Q}_k, \widehat{Q}) - \mathrm{tr}\{(Q_k - \widehat{Q}_k)\widehat{A}_k\}$$

$$\widehat{A}_k = - \left. \frac{\partial WSR_{\overline{k}}(Q_k, \widehat{Q})}{\partial Q_k} \right|_{\widehat{Q}_k, \widehat{Q}} = \sum_{i=1,\neq k}^{K} u_i H_i^H (\widehat{R}_{\overline{i}}^{-1} - \widehat{R}_i^{-1}) H_i$$

[KimGiannakis:IT0511] (2)

- Note that the linearized (tangent) expression for $WSR_{\overline{k}}$ constitutes a lower bound for it.
- Now, dropping constant terms, reparameterizing $Q_k = \mathbf{g}_k \mathbf{g}_k^H$ and performing this linearization for all users,

$$WSR(\mathbf{g}, \widehat{\mathbf{g}}) = \sum_{k=1}^{K} u_k \ln(1 + \mathbf{g}_k^H H_k^H \widehat{R}_{\overline{k}}^{-1} H_k \mathbf{g}_k) - \mathbf{g}_k^H (\widehat{A}_k + \lambda I)\mathbf{g}_k + \lambda P \,.$$

The gradient of this concave WSR lower bound is actually still the same as that of the original WSR or of the WSMSE criteria! Allows generalized eigenvector interpretation:

$$H_k^H \widehat{R}_{\overline{k}}^{-1} H_k \mathbf{g}_k = \frac{1 + \mathbf{g}_k^H H_k^H \widehat{R}_{\overline{k}}^{-1} H_k \mathbf{g}_k}{u_k} (\widehat{A}_k + \lambda I)\mathbf{g}_k$$

or hence $\mathbf{g}_k^{'} = V_{max}(H_k^H \widehat{R}_{\overline{k}}^{-1} H_k, \widehat{A}_k + \lambda I)$
which is proportional to the "LMMSE" $\mathbf{g}_k$,
with max eigenvalue $\sigma_k = \sigma_{max}(H_k^H \widehat{R}_{\overline{k}}^{-1} H_k, \widehat{A}_k + \lambda I)$.

- Again, [KimGiannakis:IT0511] BF:

$$\mathbf{g}_k^{'} = V_{max}(H_k^H \widehat{R}_{\overline{k}}^{-1} H_k, \sum_{i=1,\neq k}^{K} u_i H_i^H (\widehat{R}_{\overline{i}}^{-1} - \widehat{R}_i^{-1}) H_i + \lambda I)$$

- This can be viewed as an optimally weighted version of SLNR (Signal-to-Leakage-plus-Noise-Ratio) [Sayed:SP0507]

$$SLNR_k = \frac{||H_k \mathbf{g}_k||^2}{\sum_{i\neq k} ||H_i \mathbf{g}_k||^2 + \sum_i ||\mathbf{g}_i||^2 / P} \text{ vs}$$

$$SINR_k = \frac{||H_k \mathbf{g}_k||^2}{\sum_{i\neq k} ||H_k \mathbf{g}_i||^2 + \sum_i ||\mathbf{g}_i||^2 / P}$$

- SLNR takes as Tx filter

$$\mathbf{g}_k^{'} = V_{max}(H_k^H H_k, \sum_{i\neq k} H_i^H H_i + I)$$

[KimGiannakis:IT0511] Interference Aware WF

- Let $\sigma_k^{(1)} = \mathbf{g}_k^{'H} H_k^H \widehat{R}_{\overline{k}}^{-1} H_k \mathbf{g}_k^{'}$ and $\sigma_k^{(2)} = \mathbf{g}_k^{'H} \widehat{A}_k \mathbf{g}_k^{'}$.
- The advantage of this formulation is that it allows straightforward power adaptation: substituting $\mathbf{g}_k = \sqrt{p_k}\, \mathbf{g}_k^{'}$ yields

$$WSR = \lambda P + \sum_{k=1}^{K} \{ u_k \ln(1 + p_k \sigma_k^{(1)}) - p_k(\sigma_k^{(2)} + \lambda) \}$$

which leads to the following interference leakage aware water filling

$$p_k = \left(\frac{u_k}{\sigma_k^{(2)} + \lambda} - \frac{1}{\sigma_k^{(1)}} \right)^+ .$$

- For a given λ, $\mathbf{g}$ needs to be iterated till convergence.
- And λ can be found by duality (line search):

$$\min_{\lambda \geq 0} \max_{\mathbf{g}} \lambda P + \sum_k \{ u_k \ln\det(R_{\overline{k}}^{-1} R_k) - \lambda p_k \} = \min_{\lambda \geq 0} WSR(\lambda).$$

High/Low SNR Behavior

- At high SNR, max WSR BF converges to ZF solutions with uniform power

$$\mathbf{g}_k^H = f_k H_k P^{\perp}_{(fH)^H_{\overline{k}}} / ||f_k H_k P^{\perp}_{(fH)^H_{\overline{k}}}||$$

  where $P^{\perp}_{\mathbf{X}} = I - P_{\mathbf{X}}$ and $P_{\mathbf{X}} = \mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H$ projection matrices
  $(fH)_{\overline{k}}$ denotes the (up-down) stacking of $f_i H_i$ for users $i = 1, \ldots, K$, $i \neq k$.
- At low SNR, matched filter for user with largest $||H_k||_2$ (max singular value)

Local Optima

- MIMO: distribute ZF between Tx and Rx, yielding different gains ($|f_k H_k \mathbf{g}_k|$)! If the Rx ZF's n streams, then the Tx only has to ZF $M-1-n$ streams! # possible solutions :

$$\prod_{k=1}^{K} \left(\sum_{n=0}^{N_k-1} \frac{(M-1)!}{n!(M-1-n)!} \right)$$

i.e., for each user, the Rx can ZF n between 0 and $N-1$ streams, to choose among $M-1$.
- These different ZF solutions are the possible local optima for max WSR at infinite SNR. By homotopy [Negro:ita11] this remains the number of max WSR local optima as the SNR decreases from infinity. As the SNR decreases further, a stream for some user may get turned off until only a single stream remains at low SNR. Hence, the number of local optima reduces as streams disappear at finite SNR.
- As a corollary, in the MISO case, the max WSR optimum is unique, since there is only one way to perform ZF BF.

Mean and Covariance Gaussian CSIT

- Mean information about the channel can come from channel feedback or reciprocity, and prediction, or it may correspond to the non fading (e.g. LoS) part of the channel (note that an unknown phase factor $e^{j\phi}$ in the overall channel mean does not affect the BF design).
- Covariance information may correspond to channel estimation (feedback, prediction) errors and/or to information about spatial correlations. The separable (or Kronecker) correlation model (for the channel itself, as opposed to its estimation error or knowledge) below is acceptable when the number of propagation paths N_p becomes large ($N_p \gg MN$) as possibly in indoor propagation.
- Given only mean and covariance information, the fitting maximum entropy distribution is Gaussian.

Mean and Covariance Gaussian CSIT (2)

- Hence consider

$$\text{vec}(H) \sim \mathcal{CN}(\text{vec}(\overline{H})\,, C_t^T \otimes C_r) \text{ or } H = \overline{H} + C_r^{1/2}\,\widetilde{H}\,C_t^{1/2}$$

where $C_r^{1/2}$, $C_t^{1/2}$ are Hermitian square-roots of the Rx and Tx side covariance matrices

$$\begin{aligned} \mathrm{E}\,(H-\overline{H})(H-\overline{H})^H &= \mathrm{tr}\{C_t\}\; C_r \\ \mathrm{E}\,(H-\overline{H})^H(H-\overline{H}) &= \mathrm{tr}\{C_r\}\; C_t \end{aligned}$$

and the elements of $\widetilde{H}$ are i.i.d. $\sim \mathcal{CN}(0,1)$. A scale factor needs to be fixed in the product $\mathrm{tr}\{C_r\}\mathrm{tr}\{C_t\}$ for unicity.

- In what follows, it will also be of interest to consider the total Tx side correlation matrix

$$R_t = \mathrm{E}\;H^H H = \overline{H}^H\overline{H} + \mathrm{tr}\{C_r\}\,C_t\ .$$

- Gaussian CSIT model could be considered an instance of Ricean fading in which the ratio $\mathrm{tr}\{\overline{H}^H\overline{H}\}/(\mathrm{tr}\{C_r\}\mathrm{tr}\{C_t\}) =$ Ricean factor.

Max Expected WSR (EWSR)

- scenario of interest: perfect CSIR, partial (LoS) CSIT
- Imperfect CSIT $\Rightarrow$ various possible optimization criteria: outage capacity,.... Here: expected weighted sum rate
  $\mathrm{E}_{\,H} WSR(\mathbf{g}, H) =$

$$EWSR(\mathbf{g}) = \mathrm{E}_{\,H} \sum_k u_k \ln(1 + \mathbf{g}_k^H H_k^H R_{\overline{k}}^{-1} H_k \mathbf{g}_k)$$

  perfect CSIR: optimal Rx filters f_k (fn of aggregate H) have been substituted: $WSR(\mathbf{g}, H) = \max_f \sum_k u_k(-\ln(e_k(f_k, \mathbf{g})))$.
- At high SNR we get:

Theorem

Sufficiency of Incomplete CSIT for Full DoF in MIMO BC *In the MIMO BC with perfect CSIR, it is sufficient that for each of the K users* $\operatorname{rank}(R_{t,k}) \leq N_k$ *and that the BS knows any vector* $h_k \in \operatorname{Range}(R_{t,k})$ *(as long as the resulting vectors* h_k *are linearly independent) in order for ZF BF to produce* $\min(M, K)$ *interference free streams (degrees of freedom (DoF)).*

Max EWSR by Stochastic Approximation

- In [Luo:spawc13] a stochastic approximation approach for maximizing the EWSR was introduced: replace statistical average by sample average (samples of H get generated according to its Gaussian CSIT distribution in a Monte Carlo fashion), and one iteration of the min WSMSE approach gets executed per term added in the sample average.
- Some issues: in this case the number of iterations may get dictated by a sufficient size for the sample average rather than by a convergence requirement for the iterative approach.
- Another issue is that this approach converges to a local maximum of the EWSR. It is not immediately clear how to combine this stochastic approximation approach with deterministic annealing.
- Below: various deterministic approximations and bounds for the EWSR, which can then be optimized as in the full CSI case.

EWSR Lower Bound: EWSMSE

- $EWSR(\mathbf{g})$: difficult to compute and to maximize directly. [Negro:iswcs12] much more attractive to consider $\mathrm{E}_H e_k(f_k, \mathbf{g}, H)$ since $e_k(f_k, \mathbf{g}, H)$ is quadratic in H. Hence optimizing $\mathrm{E}_H WSMSE(\mathbf{g}, f, w, H)$.

$$\min_{f,w} \mathrm{E}_H WSMSE(\mathbf{g}, f, w, H)$$
$$\geq \mathrm{E}_H \min_{f,w} WSMSE(\mathbf{g}, f, w, H) = -EWSR(\mathbf{g})$$

or hence $EWSR(\mathbf{g}) \geq -\min_{f,w} \mathrm{E}_H WSMSE(\mathbf{g}, f, w, H)$.

- So now only a lower bound to the EWSR gets maximized, which corresponds in fact to the CSIR being equally partial as the CSIT.

$$\mathrm{E}_H e_k = 1 - 2\Re\{f_k^H \overline{H}_k \mathbf{g}_k\} + \sum_{i=1}^K f_k^H \overline{H}_k \mathbf{g}_i \mathbf{g}_i^H \overline{H}_k^H f_k$$
$$+ f_k^H R_{r,k} f_k \sum_{i=1}^K \mathbf{g}_i^H R_{t,k} \mathbf{g}_i + ||f_k||^2.$$

$\Rightarrow$ signal term disappears if $\overline{H}_k = 0$! Hence the EWSMSE lower bound is (very) loose unless the Rice factor is high, and is useless in the absence of mean CSIT.

EWSR Upper Bound: WSESINR

- Using the concavity of ln(.), we get

$$EWSR(\mathbf{g}) \leq \sum_{k=1}^{K} u_k \ln(1 + \mathrm{E}_{H_k} \mathrm{SINR}_k(\mathbf{g}, H_k)) \, .$$

Max ES-EI-NR Approach

- Consider the approximation

$$\mathrm{E}_{\,H} \ln(1 + \mathrm{SINR}_k) \approx \ln(1 + \frac{\mathrm{E}\; S}{\mathrm{E}\; I + N}) \;\; .$$

  This can be solved as easily as min (E)WSMSE!
  However, here the $\widetilde{H}_k$ part in the signal gets also counted in the signal power, unlike in the EWSMSE criterion where it gets ignored.
- Approximation becomes exact in Massive MIMO, $M \to \infty$.
- Rewrite WSR (level of Rx signal iso Rx output)

$$WSR = \sum_{k=1}^{K} u_k [\ln\det(\widetilde{R}_k) - \ln\det(\widetilde{R}_{\overline{k}})]$$

  w/ $\widetilde{R}_k = (\sum_i Q_i) H_k^H H_k + I_M$, $\widetilde{R}_{\overline{k}} = (\sum_{i \neq k} Q_i) H_k^H H_k + I_M$.
- Can apply [KimGiannakis:IT0511], replacing $\widetilde{R}_k$, $\widetilde{R}_{\overline{k}}$ by E $\widetilde{R}_k$, E $\widetilde{R}_{\overline{k}}$, and hence $H_k^H H_k$ by $R_{t,k}$ and expressions of the form $H_k^H R^{-1} H_k$ by $R_{t,k} R^{-1}$.

Large MIMO Asymptotics Refinement

- SU MIMO asymptotics from [Loubaton:IT0310],[Taricco:IT0808] (in which both $M, N \to \infty$, which tends to give more precise approximations when M is not so large) for a term of the form $\ln\det(QH^HH+I)$ correspond to replacing $H_k^H H_k$ in the $\widetilde{R}_k$ and $\widetilde{R}_{\overline{k}}$ with a kind of $R_{t,k}$ with a different weighting of the $\overline{H}_k^H\overline{H}_k$ and $C_{t,k}$ portions, of the form $R'_{t,k} = a_k C_{t,k} + \overline{H}_k^H \mathbf{B}_k \overline{H}_k$ for some scalar a_k and matrix $\mathbf{B}_k$ that depends on $C_{r,k}$.
- For the general case of Gaussian CSIT with separable (Kronecker) covariance, get

$$
\begin{aligned}
&\mathrm{E}_{\,H} \ln\det(I + HQH^H) \\
&= \max_{z,w}\left\{\ln\det\begin{bmatrix} I + wC_r & \overline{H} \\ -Q\overline{H}^H & I + zQC_t \end{bmatrix} - zw\right\}.
\end{aligned}
$$

$\max_{z,w}$ interpretation is new.

Large MIMO Asymptotics Refinement (2)

- Simpler case: zero channel means $\overline{H}_k = 0$ and no Rx side correlations $C_r = I$, and with per user Tx side correlations $C_t \leftarrow C_k$, the EWSR w large MIMO asymptotics:

$$
\begin{aligned}
& EWSR = \\
& \sum_{k=1}^{K} \Big\{ u_k \max_{z_k, w_k} \left[\ln\det(I + z_k G G^H C_k) + N_k \ln(1 + w_k) - z_k w_k \right] \\
& - u_k \max_{z_{\overline{k}}, w_{\overline{k}}} \left[\ln\det(I + z_{\overline{k}} G_{\overline{k}} G_{\overline{k}}^H C_k) + N_k \ln(1 + w_{\overline{k}}) - z_{\overline{k}} w_{\overline{k}} \right] \Big\}
\end{aligned}
$$

where $G = [\mathbf{g}_1 \cdots \mathbf{g}_K]$ and $G_{\overline{k}}$ is the same as G except for column $\mathbf{g}_k$. Can be maximized by alternating optimization.

Other possible WSR Approximations

- Absorbing the Mean in the Covariance:
  Replacing $\overline{H}_k$ by 0 and $C_{t,k}$ by $R_{t,k}$ as suggested in [deFrancisco:asilo05] for SU MIMO leads to one simplification. Other simplifications can be obtained by either absorbing the noise term in the "Rayleigh" channel part of the interference or vice versa.
- Improvements upon ESEINR
  E.g. apply E_{H} (only) to the explicit (quadratic) appearances of H in $\frac{\partial WSR}{\partial \mathbf{g}_k}$, replacing terms like the MSE e_k and $R_{\overline{k}}$ by their mean. This approach acknowledges that the Rx contains the channel matched filter as factor and applies the second order statistics to the resulting quadratic appearances of the channel.
- Higher-Order Taylor Series Expansions
  E.g. go to the next (second) order term in the Taylor series expansion of the log as in [MartinOttersten:SP0704].

Location Aided Partial CSIT LoS Channel Model

- Assuming the Tx disposes of not much more than the LoS component information, model

$$H = h_r\, h_t^H(\theta) + \widetilde{H}^{'}$$

  where θ is the LoS AoD and the Tx side array response is normalized: $||h_t(\theta)||^2 = 1$.
- Since the orientation of the MT is random, model the Rx side LoS array response h_r as vector of i.i.d. complex Gaussian

$$\begin{array}{ll} h_r \;\text{ i.i.d. } \sim \mathcal{CN}(0, \frac{\mu}{\mu+1}) & \text{and} \\ \widetilde{H}^{'} \;\text{ i.i.d. } \sim \mathcal{CN}(0, \frac{1}{\mu+1}\frac{1}{M}) & \text{, independent of } h_r, \end{array}$$

  where the matrix $\widetilde{H}$ represents the aggregate NLoS components.

Location Aided Partial CSIT LoS Channel Model (2)

- Note that

$$\mathrm{E}\ ||H||_F^2 = \mathrm{E}\ \mathrm{tr}\{H^H H\} =$$
$$||h_t(\theta)||^2\ \mathrm{E}\ ||h_r||^2 + \mathrm{E}\ ||\widetilde{H}^{'}||_F^2 = \tfrac{\mu N}{\mu+1} + \tfrac{N}{\mu+1} = N\,,$$

$(\mathrm{E}\ ||h_r\, h_t^T(\theta)||_F^2)/(\mathrm{E}\ ||\widetilde{H}^{'}||_F^2) = \mu =$ a Rice factor.

- In fact the only parameter additional to the LoS AoD θ is μ.
- So, this is a case of zero mean CSIT and Tx side covariance CSIT

$$R_t = \mathrm{E}\ \ H^H H = \frac{\mu N}{\mu+1} h_t(\theta) h_t^H(\theta) + \frac{N}{\mu+1}\frac{1}{M} I_M\ .$$

Location Aided LoS ZF BF

- For ZF BF, the BS shall use for user k a spatial filter $\mathbf{g}_k = \sqrt{p_k}\,\mathbf{g}_k^{'}$ such that $\mathbf{g}_k^{'} = \mathbf{g}_k^{''}/||\mathbf{g}_k^{''}||$

$$\mathbf{g}_k^{''} = P^{\perp}_{h_{t,\bar{k}}} h_{t,k}$$

  where $h_{t,\bar{k}} = [h_{t,1} \cdots h_{t,k-1}\, h_{t,k+1} \cdots h_{t,K}]$.
- And uniform power distribution $p_k = P/K,\ k = 1, \ldots, K$.
- The $\mathbf{g}_k^{''}$ can also be computed from

$$\mathbf{g}^{''} = [\mathbf{g}_1^{''} \cdots \mathbf{g}_K^{''}] = h_t (h_t^H h_t)^{-1}\ , \quad h_t = [h_{t,1} \cdots h_{t,K}]\,.$$

LoS based Partial CSIT: Beyond ZF

- Go beyond the asymptotics of high SNR and high Ricean factor: even if the Tx ignores the multipath and the Rx can handle it, it would be better to have a multipath aware Tx design. Note that the Ricean factor μ satisfies uplink/downlink (UL/DL) reciprocity, even in a FDD. Solution: previous partial CSIT design.

- Absorbing the Rayleigh Component in the Noise

$$\begin{aligned} y_k &= H_k \textstyle\sum_{i=1}^K \mathbf{g}_i \, x_i + v_k \\ &= h_{r,k} \, h_{t,k}^H \textstyle\sum_{i=1}^K \mathbf{g}_i \, x_i + \widetilde{H}_k^{'} \textstyle\sum_{i=1}^K \mathbf{g}_i \, x_i + v_k \, . \end{aligned}$$

From MIMO to equivalent SIMO with same SINR (or ESINR):

$$y_k = \sqrt{\frac{\mu_k}{\mu_k+1}} h_{t,k}^H \sum_{i=1}^K \mathbf{g}_i \, x_i + \frac{1}{\sqrt{(\mu_k+1)M}} \widetilde{h}_k^H \sum_{i=1}^K \mathbf{g}_i \, x_i + v_k^{'} \, .$$

or also

$$y_k = h_{t,k}^H \sum_{i=1}^K \mathbf{g}_i \, x_i + v_k$$

with noise var $\sigma_{v,k}^2 + \dfrac{P}{(\mu_k+1)M} = \sigma_{v,k}^2 (1 + \dfrac{\mathrm{SNR}_k}{(\mu_k+1)M})$ and

$$\mathrm{SNR}_{eff,k} = \frac{\mu_k \mathrm{SNR}_k}{\mu_k + 1 + \mathrm{SNR}_k / M}$$

which is now a deterministic MISO BC model.

Ricean Model Specific Approximations (2)

- Absorbing the Noise in the Rayleigh Component
  effectively replacing $H_k^H H_k$ by $C_{t,k}$, again resulting in a deterministic WSR scenario.

Preliminary Simulation

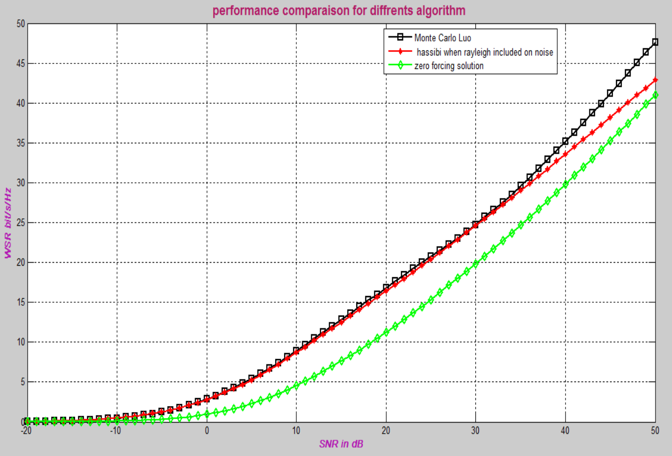


Figure : EWSR vs SNR for $K = M = N_k = 4$ with Rice factor $\mu = 10$. [Luo:spawc13] stochastic approximation, ZF on the LoS component, optimized deterministic BF design when the Rayleigh part is absorbed in the noise.

Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - uplink/downlink(UL/DL) duality; SU MIMO,BC,MAC; BF&DPC
  - BC with user selection: DPC vs BF
  - Max WSR, UL/DL duality, CSIT: perfect, partial, LoS
- interference multi-cell/HetNets: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
  - distributed Channel State Information at the Transmitter (CSIT) acquisition, netDoF
  - Delayed CSIT, optimal handling of CSIT FB dead times
  - Finite Rate of Innovation (FRoI)/Basis Expansion Model (BEM) channel models
  - Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels (IBC)

Past Instances of Non-circular Symbol Constellations

- SAIC for real signal constellations

  application to GSM: GMSK $\approx$ filtered modulated BPSK

- space-time coding

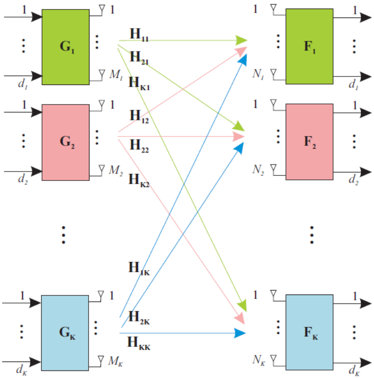
  Alaouti scheme, = special case of linear dispersion space-time codes

- turbo receivers

  due to channel coding and bit to symbol mapping, reconstructed interfering symbols are typically non-circular (at order 2)

MIMO IFC Introduction

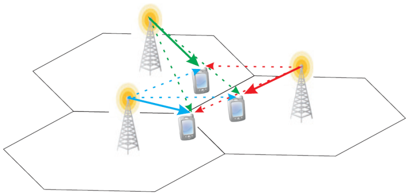
- Interference Alignment (IA) was introduced in [Cadambe,Jafar 2008]
- The objective of IA is to design the Tx beamforming matrices such that the interference at each non intended receiver lies in a common interference subspace
- If alignment is complete at the receiver simple Zero Forcing (ZF) can suppress interference and extract the desired signal
- In [SPAWC2010] we derive a set of interference alignment (IA) feasibility conditions for a K-link frequency-flat MIMO interference channel (IFC)
- $d = \sum_{k=1}^{K} d_k$



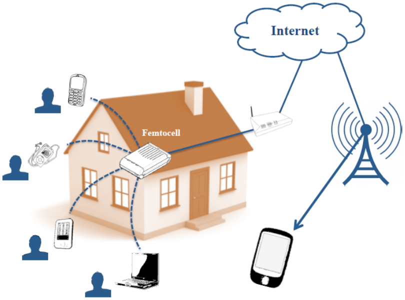
MIMO Interference Channel

Possible Application Scenarios

- Multi-cell cellular systems, modeling intercell interference.
  Difference from Network MIMO: no exchange of signals, "only" of channel impulse responses.



- HetNets: Coexistence of macrocells and small cells, especially when small cells are considered part of the cellular solution.



Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - uplink/downlink(UL/DL) duality; SU MIMO,BC,MAC; BF&DPC
  - BC with user selection: DPC vs BF
  - Max WSR, UL/DL duality, CSIT: perfect, partial, LoS
- interference multi-cell/HetNets: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
  - distributed Channel State Information at the Transmitter (CSIT) acquisition, netDoF
  - Delayed CSIT, optimal handling of CSIT FB dead times
  - Finite Rate of Innovation (FRoI)/Basis Expansion Model (BEM) channel models
  - Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels (IBC)

Why IA?

- The number of streams (degrees of freedom (dof)) appearing in a feasible IA scenario correspond to prelogs of feasible multi-user rate tuples in the multi-user rate region.
  Max Weighted Sum Rate (WSR) becomes IA at high SNR.
- Noisy IFC: interfering signals are not decoded but treated as (Gaussian) noise.
  Apparently enough for dof.
- Lots of recent work more generally on rate prelog regions: involves time sharing, use of fractional power.

Some State of the Art

- perfect CSI
  - signal space interference alignment (joint Tx/Rx ZF)
    - Interference Alignment (IA) [Khandani etal]
    - [CadambeJafar:IT08] IA can get $K/2$ DoF in t/f selective SISO (hence MIMO) via symbol extension
    - noisy MIMO w/o symbol extension, IA feasibility [Santamaria etal:isit12],[RuanLau:isit12]
      Tx/Rx design: IA, max per stream SINR, max WSR
    - ergodic IA: group channel realizations H_1, H_2 s.t. $\text{offdiag}(H_2) = -\text{offdiag}(H_1)$, $\text{diag}(H_2) = \text{diag}(H_1)$
  - signal scale IA (Tx rational numbers, diophantine equation)

Noisy MIMO IFC: Some State of the Art

- IA: alternating ZF algorithm [Jafar etal: globecom08],[Heath etal: icassp09].
- IA feasibility: - $K = 2$ MIMO: [JafarFakhereddin:IT07]
  - [Yetis,Jafar:T10], [Slock etal:eusipco09,ita10, spawc10]
  - 3xNxN, 3xMxN: [BreslerTse:arxiv11]
- max WSR: single stream/link
  - approximately: max SINR [Jafar etal: globecom08]
  - eigenvector interpretation of WSR gradient w.r.t. BF: starting [Honig,Utschick:asilo09]
  - added DA-style approach in [Honig,Utschick:allerton10]
- max WSR: multiple streams/link
  - [Slock etal:ita10] application of [Christensen etal:TW08] from MIMO BC
  - further refined in [Negro etal:allerton10], independently suggested use of DA, developed in [Negro etal:ita11]

Various IA Flavors

- *linear* IA [GouJafar:IT1210], also called *signal space* IA, only uses the spatial dimensions introduced by multiple antennas.
- *asymptotic* IA [CadambeJafar:IT0808] uses symbol extension (in time and/or frequency), leading to (infinite) symbol extension involving diagonal channel matrices, requiring infinite channel diversity in those dimensions. This leads to infinite latency also. The (sum) DoF of asymptotic MIMO IA are determined by the *decomposition* bound [WangSunJafar:isit12].
- *ergodic* IA [NazerGastparJafarVishwanath:IT1012] explains the factor 2 loss in DoF of SISO IA w.r.t. an interference-free Tx scenario by transmitting the same signal twice at two paired channel uses in which all cross channel links cancel out each other. Ergodic IA also suffers from uncontrolled latency but provides the factor 2 rate loss at any SNR. The DoF of ergodic MIMO IA are also determined by the decomposition bound [LejosneSlockYuan:icassp14].

IA as a Constrained Compressed SVD

- $F_k^H: \ d_k \times N_k, \ \ H_{ki}: \ N_k \times M_i, \ \ G_i: \ M_i \times d_i \qquad F^H H G =$

$$\begin{bmatrix} F_1^H & 0 & \cdots & 0 \\ 0 & F_2^H & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & F_K^H \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} & \cdots & H_{1K} \\ H_{21} & H_{22} & \cdots & H_{2K} \\ \vdots & & \ddots & \vdots \\ H_{K1} & H_{K2} & \cdots & H_{KK} \end{bmatrix} \begin{bmatrix} G_1 & 0 & \cdots & 0 \\ 0 & G_2 & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & G_K \end{bmatrix} = \begin{bmatrix} F_1^H H_{11} G_1 & 0 & \cdots & 0 \\ 0 & F_2^H H_{22} G_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & F_K^H H_{KK} G_K \end{bmatrix}$$

  F^H, G can be chosen to be unitary for IA

- per user vs per stream approaches:

  IA: can absorb the $d_k \times d_k$ $F_k^H H_{kk} G_k$ in either F_k^H (per stream LMMSE Rx) or G_k or both.

  WSR: can absorb unitary factors of SVD of $F_k^H H_{kk} G_k$ in F_k^H, G_k without loss in rate $\Rightarrow F^H H G =$ diagonal.

Interference Alignment: Feasibility Conditions (1)

- To derive the existence conditions we consider the ZF conditions

$$\underbrace{\mathbf{F}_k^H}_{d_k \times N_k} \underbrace{\mathbf{H}_{kl}}_{N_k \times M_l} \underbrace{\mathbf{G}_l}_{M_l \times d_l} = \mathbf{0} \,, \quad \forall l \neq k$$

$$\text{rank}(\mathbf{F}_k^H \mathbf{H}_{kk} \mathbf{G}_k) = d_k \,, \quad \forall k \in \{1, 2, \ldots, K\}$$

- rank requirement $\Rightarrow$ SU MIMO condition: $\quad d_k \leq \min(M_k, N_k)$
- The total number of variables in $\mathbf{G}_k$ is
  $d_k M_k - d_k^2 = d_k(M_k - d_k)$
  Only the subspace of $\mathbf{G}_k$ counts, it is determined up to a $d_k \times d_k$ mixture matrix.
- The total number of variables in $\mathbf{F}_k^H$ is
  $d_k N_k - d_k^2 = d_k(N_k - d_k)$
  Only the subspace of $\mathbf{F}_k^H$ counts, it is determined up to a $d_k \times d_k$ mixture matrix.

Interference Alignment: Feasibility Conditions (2)

- A solution for the interference alignment problem can only exist if the total number of variables is greater than or equal to the total number of constraints i.e.,

$$\sum_{k=1}^{K} d_k(M_k - d_k) + \sum_{k=1}^{K} d_k(N_k - d_k) \geq \sum_{i\neq j=1}^{K} d_i\, d_j$$

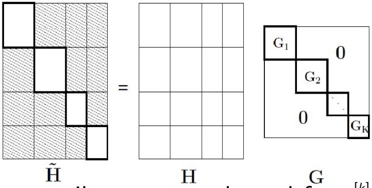
$$\Rightarrow \; \sum_{k=1}^{K} d_k(M_k + N_k - 2d_k) \geq (\sum_{k=1}^{K} d_k)^2 - \sum_{k=1}^{K} d_k^2$$

$$\Rightarrow \; \sum_{k=1}^{K} d_k(M_k + N_k) \geq (\sum_{k=1}^{K} d_k)^2 + \sum_{k=1}^{K} d_k^2$$

- In the symmetric case: $d_k = d$, $M_k = M$, $N_k = N$:
  $d \; \leq \; \frac{M+N}{K+1}$
- For the $K = 3$ user case ($M = N$): $\quad d \; = \; \frac{M}{2}$.
  With 3 parallel MIMO links, half of the (interference-free) resources are available!
  However $d \leq \frac{1}{(K+1)/2} M < \frac{1}{2} M$ for $K > 3$.

Interference Alignment: Feasibility Conditions (3)

The main idea of our approach is to convert the alignment requirements at each RX into a rank condition of an associated interference matrix $\mathbf{H}_I^{[k]}$=[$\mathbf{H}_{k1}\mathbf{G}_1, \ldots \mathbf{H}_{k(k-1)}\mathbf{G}_{(k-1)}, \mathbf{H}_{k(k+1)}\mathbf{G}_{(k+1)}, \ldots \mathbf{H}_{kk}\mathbf{G}_K$], that spans the interference subspace at the k-th RX (the shaded blocks in each block row).Thus the dimension of the Interference subspace must satisfy rank$(\mathbf{H}_I^{[k]}) = r_I^{[k]} \leq N_k - d_k$



The equation above prescribes an upperbound for $r_I^{[k]}$ but the nature of the channel matrix (full rank) and the rank requirement of the BF specifies the following lower bound $r_I^{[k]} \geq \max_{I \neq k}(d_I - [M_I - N_k]_+)$. Imposing a rank $r_I^{[k]}$ on $\mathbf{H}_I^{[k]}$ implies imposing $(N_k - r_I^{[k]})(\sum_{\substack{I=1 \\ I \neq k}}^{K} d_I - r_I^{[k]})$ constraints at RX k. Enforcing the minimum number of constraints on the system implies to have maximum rank: $r_I^{[k]} \leq \min(d_{tot}, N_k) - d_k$

Interference Alignment: Feasibility Conditions (4)

- [BreslerTse:arxiv11]: counting equations and variables not the whole story!
- appears in very "rectangular" ($\neq$ square) MIMO systems
- example: $(M, N, d)^K = (4, 8, 3)^3$ MIMO IFC system comparing variables and ZF equations: $d = \frac{M+N}{K+1} = \frac{4+8}{3+1} = \frac{12}{4} = 3$ should be possible
- supportable interference subspace dim. $= N - d =$ 8-3 $= 5$
- however, the 2 interfering 8×4 cross channels generate 4-dimensional subspaces which in an 8-dimensional space do not intersect w.p. 1 !
- hence, the interfering 4×3 transmit filters cannot massage their 6-dimensional joint interference subspace into a 5-dimensional subspace!
- This issue is not captured by # variables vs # equations: $d = \frac{M+N}{K+1}$ only depends on $M + N$: $(5, 7, 3)^3$, $(6, 6, 3)^3$ work.

Feasibility Linear IA

- We shall focus here on linear IA, in which the spatial Tx filters align their various interference terms at a given user in a common subspace so that a Rx filter can zero force (ZF) it. Since linear IA only uses spatial filtering, it leads to low latency.
- The DoF of linear IA are upper bounded by the so-called *proper bound* [Negro:eusipco09], [Negro:spawc10], [YetisGouJafarKayran:SP10], which simply counts the number of filter variables vs. the number of ZF constraints.
- The proper bound is not always attained though because to make interference subspaces align, the channel subspaces in which they live have to sufficiently overlap to begin with, which is not always the case, as captured by the so-called *quantity bound* [Tingting:arxiv0913] and first elucidated in [BreslerCartwrightTse:allerton11], [BreslerCartwrightTse:itw11], [WangSunJafar:isit12].
- The transmitter coordination required for DL IA in a multi-ce"

I and Q components: IA with Real Symbol Streams

- Using real signal constellations in place of complex constellations, transmission over a complex channel of any given dimension can be interpreted as transmission over a real channel of double the original dimensions (by treating the I and Q components as separate channels).
- This doubling of dimensions provides additional flexibility in achieving the total DoF available in the network.
- Split complex quantities in I and Q components:

$$\mathbf{H}_{ij} = \begin{bmatrix} \mathrm{Re}\{\mathbf{H}_{ij}\} & -\mathrm{Im}\{\mathbf{H}_{ij}\} \\ \mathrm{Im}\{\mathbf{H}_{ij}\} & \mathrm{Re}\{\mathbf{H}_{ij}\} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} \mathrm{Re}\{\mathbf{x}\} \\ \mathrm{Im}\{\mathbf{x}\} \end{bmatrix}$$

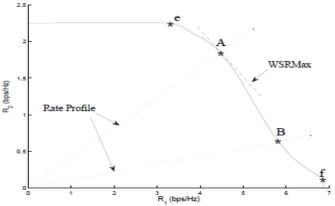
- Example: GMSK in GSM: was considered as wasting half of the resources, but in fact unknowingly anticipated interference treatment: 3 interfering GSM links can each support one GMSK signal without interference by proper joint Tx/Rx design! (SAIC: handles 1 interferer, requires only Rx design).

Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - uplink/downlink(UL/DL) duality; SU MIMO,BC,MAC; BF&DPC
  - BC with user selection: DPC vs BF
  - Max WSR, UL/DL duality, CSIT: perfect, partial, LoS
- interference multi-cell/HetNets: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
  - distributed Channel State Information at the Transmitter (CSIT) acquisition, netDoF
  - Delayed CSIT, optimal handling of CSIT FB dead times
  - Finite Rate of Innovation (FRoI)/Basis Expansion Model (BEM) channel models
  - Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels (IBC)

From IA to Optimized IFC's

- from Interference Alignment (=ZF) to max Sum Rate (SR) for the "Noisy IFC".
- to vary the point reached on the rate region boundary: SR $\rightarrow$ Weighted SR (WSR)
- problem: IFC rate region not convex $\Rightarrow$ multiple (local) optima for WSR (multiple boundary points with same tangent direction)
- solution of [CuiZhang:ita10]: WSR $\rightarrow$ max SR under rate profile constraint: $\frac{R_1}{\alpha_1} = \frac{R_2}{\alpha_2} = \cdots = \frac{R_K}{\alpha_K}$: $K-1$ constraints.
  Pro: explores systematically rate region boundary.
  Con: for a fixed rate profile,
  bad links drag down good links.
  $\Rightarrow$ stick to (W)SR
  (monitoring global opt issues).
  Note: multiple WSR solutions
  $\Leftrightarrow$ multiple IA solutions.



MWSR: Maximum Weighted Sum Rate (WSR)

The received signal at the $k-$th receiver is:

$$\mathbf{y}_k = \mathbf{H}_{kk}\mathbf{G}_k\mathbf{x}_k + \sum_{\substack{l=1 \\ l\neq k}}^{K} \mathbf{H}_{kl}\mathbf{G}_l\mathbf{x}_l + \mathbf{n}_k$$

Introduce the interference plus noise covariance matrix at receiver k: $\mathbf{R}_{\bar{k}} = \mathbf{R}_{nn} + \sum_{l\neq k} \mathbf{H}_{kl}\mathbf{G}_l\mathbf{G}_l^H\mathbf{H}_{kl}^H$.
The WSR criterion is

$$\mathcal{R} = \sum_{k=1}^{K} \mathrm{u}_k \log\det(\mathbf{I} + \mathbf{G}_k^H\mathbf{H}_{kk}^H\mathbf{R}_{\bar{k}}^{-1}\mathbf{H}_{kk}\mathbf{G}_k) \tag{9}$$

$$\text{s.t.} \quad \mathrm{Tr}\{\mathbf{G}_k^H\mathbf{G}_k\} \leq P_k$$

This criterion is highly non convex in the Tx BFs $\mathbf{G}_k$.

MWSR: Tx Filter Optimization

- [JohamUtschick:02] MMSE and scale factor heuristics
- [Hassibi] MISO BC Tx optimization, iterative algorithm without convergence proofs
- [EldarShamai:06] Tx optimization for fixed Rx, SOCP (Second-Order Cone Programming)
- [ChristensendeCarvalhoCioffi:T-WCdec08]: I-MMSE inspired WSR-WSMSE relation (similar gradient)
- ....
- [Hassibi] turns out to be MISO special case of MIMO algo below
- [Hassibi] or here: even if only MISO (only Tx), need to jointly optimize Tx, Rx, weights

MWSR: Maximum Weighted Sum Rate

Weighted sum rate expression in terms of Tx, Rx filter:

$$\mathcal{R} = \sum_{k=1}^{K} \mathrm{u}_k R_k = \sum_{k=1}^{K} \mathrm{u}_k \log\det(\mathbf{I}_{d_k} + \mathbf{F}_k \mathbf{H}_{kk} \mathbf{G}_k \mathbf{G}_k^H \mathbf{H}_{kk}^H \mathbf{F}_k^H [\mathbf{F}_k \mathbf{R}_{\bar{k}} \mathbf{F}_k^H]^{-1})$$

where $\mathbf{R}_{\bar{k}}$ denotes the interference plus noise covariance matrix at receiver k:

$$\mathbf{R}_{\bar{k}} = \mathbf{R}_{vv} + \sum_{l \neq k} \mathbf{H}_{kl} \mathbf{G}_l \mathbf{G}_l^H \mathbf{H}_{kl}^H$$

The MMSE Rx filter at user k is given as:

$$\mathbf{F}_k = \mathbf{G}_k^H \mathbf{H}_{kk}^H (\mathbf{H}_{kk} \mathbf{G}_k \mathbf{G}_k^H \mathbf{H}_{kk}^H + \mathbf{R}_{\bar{k}})^{-1} \tag{10}$$

The MMSE covariance matrix for the k-th user, using a MMSE Rx filter, can be written as:

$$\mathbf{E}_k = \mathbb{E}[(\mathbf{F}_k \mathbf{y}_k - \mathbf{d}_k)(\mathbf{F}_k \mathbf{y}_k - \mathbf{d}_k)^H] = (\mathbf{I} + \mathbf{G}_k^H \mathbf{H}_{kk}^H \mathbf{R}_{\bar{k}}^{-1} \mathbf{H}_{kk} \mathbf{G}_k)^{-1}$$

MWSR (2)

With the expression of the MMSE covariance matrix given before it is possible to express the WSR in terms of $\mathbf{E}_k$:

$$\mathcal{R} = \sum_{k=1}^{K} \mathrm{u}_k \log\det(\mathbf{I} + \mathbf{G}_k^H \mathbf{H}_{kk}^H \mathbf{R}_{\bar{k}}^{-1} \mathbf{H}_{kk} \mathbf{G}_k) = \sum_{k=1}^{K} \mathrm{u}_k \log\det(\mathbf{E}_k^{-1})$$

We want to derive the Tx filters to maximize the WSR subject to a Tx power constraint or ,equivalently, minimize the following:

$$\sum_{k=1}^{K} -\mathrm{u}_k \log\det(\mathbf{I} + \mathbf{G}_k^H \mathbf{H}_{kk}^H \mathbf{R}_{\bar{k}}^{-1} \mathbf{H}_{kk} \mathbf{G}_k) \tag{11}$$

$$\text{s.t.} \quad \mathrm{Tr}\{\mathbf{G}_k^H \mathbf{G}_k\} = P_k$$

MWSR (3)

To solve the previous optimization problem we need consider the following Lagrangian:

$$J(\{\mathbf{G}_k, \lambda_k\}) = \sum_{k=1}^{K} -\mathrm{u}_k \log\det(\mathbf{I} + \mathbf{G}_k^H \mathbf{H}_{kk}^H \mathbf{R}_{\bar{k}}^{-1} \mathbf{H}_{kk} \mathbf{G}_k) + \lambda_k (\mathrm{Tr}\{\mathbf{G}_k^H \mathbf{G}_k\} - P_k)$$

Now deriving the Lagrangian w.r.t. the Tx filter $\mathbf{G}_k$ we obtain:

$$\frac{\partial \mathrm{J}(\{\mathbf{G}_k, \lambda_k\})}{\partial \mathbf{G}_k^*} = 0$$

$$-\mathrm{u}_k \mathbf{H}_{kk}^H \mathbf{R}_{\bar{k}}^{-1} \mathbf{H}_{kk} \mathbf{G}_k \mathbf{E}_k + \sum_{l \neq k} \mathrm{u}_l \mathbf{H}_{lk}^H \mathbf{R}_{\bar{l}}^{-1} \mathbf{H}_{ll} \mathbf{G}_l \mathbf{E}_l \mathbf{G}_l^H \mathbf{H}_{ll}^H \mathbf{R}_{\bar{l}}^{-1} \mathbf{H}_{lk} \mathbf{G}_k + \lambda_k \mathbf{G}_k = 0$$

Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - uplink/downlink(UL/DL) duality; SU MIMO,BC,MAC; BF&DPC
  - BC with user selection: DPC vs BF
  - Max WSR, UL/DL duality, CSIT: perfect, partial, LoS
- interference multi-cell/HetNets: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
  - distributed Channel State Information at the Transmitter (CSIT) acquisition, netDoF
  - Delayed CSIT, optimal handling of CSIT FB dead times
  - Finite Rate of Innovation (FRoI)/Basis Expansion Model (BEM) channel models
  - Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels (IBC)

Deterministic Annealing

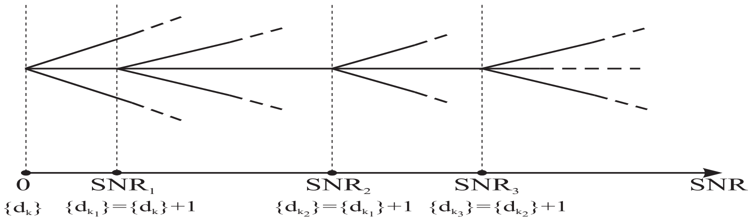
- In non-convex optimization, in order to avoid getting trapped in local minima/maxima, several heuristic approaches have been proposed.
- In analogy with the physical annealing process, Simulated Annealing (SA) has been proposed in optimization theory for non-convex problems.
- In SA the problem is optimized using a sequence of random moves, the magnitude of which depends on a temperature parameter that gets gradually cooled down.
- Deterministic annealing (DA) is inspired by the same principle but neither the cost function nor the initializations are random.
- The basic principle of DA is that the global optimum of the problem at the next temperature value is in the region of attraction of the solution of the problem at the previous temperature.

Deterministic Annealing (2)

- As in physical systems, also in an optimization problem it can happen that cooling down the temperature leads to phase transitions and hence several possible local optima may appear.
- A phase transition is characterized by a critical temperature that manifests itself by the Hessian of the cost function becoming singular. A stationary point evolves into a non stable point
- In our problem the cost function is the Weighted Sum Rate (WSR), a highly non convex function, and the annealing parameter is the noise variance, $t \propto \sigma^2$.
- Interestingly, in WSR maximization for the K-user MIMO IFC, we can associate phase transitions to the activation of an additional stream for a particular user.

Deterministic Annealing vs Homotopy

- DA is about optimization of a cost function. In the process, we are tracking what we hope to be the gobal optimum.
- Tracking of extrema, the roots of the KKT conditions, is actually called a homotopy method.
- So DA, in going from one phase transition to the next while tracking the (appropriate) extremum, is a homotopy method.



Homotopy Methods

- Homotopy is used to find the roots of a non linear system of equations $\mathcal{F}(x) = 0$.
- Homotopy transformation is such that it starts from a trivial system $\mathcal{G}(x)$, with known solution, and it evolves towards the target system $\mathcal{F}(x)$ via continuous deformations according to the homotopy parameter t:

$$\mathcal{H}(x, t) = (1 - t)\,\mathcal{G}(x) + t\mathcal{F}(x)$$

- Predicting the solution at the next value of $t^{(i+1)} = t^{(i)} + \Delta t$ is called Euler prediction phase
- Once we have a solution at $t^{(i+1)}$ it is possible to refine the estimate using a Newton correction phase for fixed t.
- A property of Homotopy continuation method for the solution of system of equation is that the number of solutions in the target system is at most equal to the number of solution in the trivial system
- The number of solutions along the trajectory remain constant as long as the Jacobian (w.r.t. x and t jointly) is full rank

Homotopy Applied to IA

- Homotopy method can be applied to the IA problem. In particular we can define an homotopy deformation that starts from a trivial system and arrive to the target problem that we want to solve: K-user MIMO IFC: use instead of SNR as temperature, a scale factor for the channel excess singular values: $H_{ji} = \sum_{k=1}^{d} \sigma_{jik} \mathbf{u}_{jik} \mathbf{v}_{jik}^H + t \sum_{k=d+1} \sigma_{jik} \mathbf{u}_{jik} \mathbf{v}_{jik}^H$
- IA Jacobian still full rank if reduce $\text{rank}(H_{ji})$ to $\max(d_j, d_i)$.
- Finding a trivial starting system is easy: e.g. rank 1 channel system.

$$\sigma_{ji1} \mathbf{f}_j^H \mathbf{u}_{ji1} \mathbf{v}_{ji1}^H \mathbf{g}_i = 0$$

- After a coordination phase, where each user decides how is going to suppress a particular stream, the solution of the IA problem is easy.
- Once we define the starting point we describe the homotopy deformation that varies increasing the rank of the channel

Alternative Zero Forcing Approach to IA

- The interpretation of IA as joint transmit-receiver liner zero forcing can be easily understood considering the trivial rank-one MIMO IFC for the case of all $d_k = 1$.
- The channel between Tx i and Rx j can be represented using its SVD decomposition:

$$\mathbf{H}_{ji} = \sigma_{ji}\mathbf{u}_{ji}\mathbf{v}_{ji}^H$$

- The IA (ZF) condition for link $i - j$ can be written as

$$\sigma_{ji}\mathbf{f}_j^H\mathbf{u}_{ji}\mathbf{v}_{ji}^H\mathbf{g}_i = 0$$

- Two configurations are possible:

$$\mathbf{f}_j^H\mathbf{u}_{ji} = 0 \quad \text{or} \quad \mathbf{v}_{ji}^H\mathbf{g}_i = 0$$

- Either the Tx or the Rx suppresses one particular interfering stream
- Homotopy here not suggested for computing IA solutions, but for counting number of solutions.

Homotopy Applied to IA (3)

- Another possible description of the same problem can be done using the expansion of the IA conditions up to the first order:

$$(\mathbf{F}_j^H + d\mathbf{F}_j^H)(\mathbf{H}_{ji} + d\mathbf{H}_{ji})(\mathbf{G}_i + d\mathbf{G}_i) = 0$$

- Expanding the products above and considering only the terms up to first order we get

$$\mathbf{F}_j^H \mathbf{H}_{ji} d\mathbf{G}_i + d\mathbf{F}_j^H \mathbf{H}_{ji} \mathbf{G}_i = -\mathbf{F}_j^H d\mathbf{H}_{ji} \mathbf{G}_i$$

- IA: n joint bilinear equations. Overall number of solutions upper bounded by 2^n (again: $2 =$ either Tx or Rx side). However, equations structured $\Rightarrow$ number of solutions less. Consider e.g. $K = 3$, all $N_k = M_k = N$: in this case all solutions are known analytically and correspond to selecting $d = N/2$ out of N eigenvectors: $\dfrac{N!}{(\frac{N}{2}!)^2} \ll 2^n = 2^{(1.5\,N)^2 - 1.5\,N}$ solutions.

Homotopy applied to WSR backwards: decreasing SNR

- First consider reduced rank channels and apply DA to WSR for SNR decreasing from infinity (IA).
- IA $\Rightarrow$ every IA solution corresponds to a distribution of interference zeroing between Tx and Rx.
- Homotopy in decreasing SNR $\Rightarrow$ can interpret all WSR extrema as different distributions of Tx and Rx roles. Phase transitions $\Rightarrow$ some stream gets turned off $\Rightarrow$ reduces a number of local maxima.
- For a given SNR, homotopy in channel rank allows for a similar interpretation of WSR extrema in original system at any SNR.
- At high SNR, number of WSR extrema $=$ number of IA solutions.
  All WSR extrema (at any SNR) are local maxima: Hessian negative definite.

Back to Forward DA for WSR

Algorithm 2 MWSR Algorithm for MIMO IFC

---

set (SNR) $t = 0$
Fix an initial set of precoding matrices $\mathbf{G}_k$, $\forall \in k = \{1, 2 \ldots K\}$
repeat
  set $t^{(i+1)} = t^{(i+1)} + \delta t$
  set $n = 0$
  Try augmenting $\mathbf{G}^{(t)}$ for one extra stream in any link.
  repeat
    $n = n + 1$
    Given $\mathbf{G}_k^{(n-1)}$ compute $\mathbf{F}_k^{(n)H}$ and $\mathbf{W}_k^{(n)}$, $\forall k$
    Given $\mathbf{F}_k^{(n)H}$, $\mathbf{W}_k^{(n)}$, compute $\mathbf{G}_k^{(n)}$ $\forall k$
  until convergence
until target SNR is reach

Back to Forward DA for WSR

- MWSR iterations for fixed SNR point are equivalent to Newton correction phase in homotopy methods
- MWSR iterations at increased SNR are equivalent to Euler prediction step in homotopy methods
- To find a good initialization at SNR$=0$ we should study the WSR expansion up to second order in the SNR. The first order term only depends on the useful signal part hence MF are optimal. The second order contribution depends by both useful signal and interference
- When we iterate MWSR algorithm Tx and Rx filter are initialized as MF, after one iteration we optimize up to second order the WSR. This because at each iteration we find the MMSE filter with the other filter being MF.

Concluding Remarks DA

- Deterministic Annealing = Homotopy + Phase Transitions
- Annealing MIMO channel singularity at high SNR $\Rightarrow$ counting number of IA solutions and interpreting each as a different distribution of ZF roles between Tx's and Rx's
- Annealing down in SNR $\Rightarrow$ all WSR local maxima correspond 1-to-1 to continuations of IA solutions
- Annealing up in SNR for Max WSR:
  - global solution known at low SNR: one stream per link, with MF for Tx and Rx
  - alternating WSR maximization leads to simple subproblems for updates of Tx, Rx and weights
  - at any temperature increase, test for phase splitting = introduction of a new stream
  - its Tx and Rx filters are again (colored noise) MF, introduced Jammer WF algo for optimal power redistribution, alternating max algo tracks correct global maximum since WSR is convex up to second order in power variations
  - resulting DA algorithm is perhaps only known structured solution for finding the global Max WSR

Outline

- interference single cell: Broadcast Channel (BC)
    - utility functions: SINR balancing, (weighted) sum rate (WSR)
    - uplink/downlink(UL/DL) duality; SU MIMO,BC,MAC; BF&DPC
    - BC with user selection: DPC vs BF
    - Max WSR, UL/DL duality, CSIT: perfect, partial, LoS
- interference multi-cell/HetNets: Interference Channel (IFC)
    - Degrees of Freedom (DoF) and Interference Alignment (IA)
    - Weighted Sum Rate (WSR) maximization and UL/DL duality
    - Deterministic Annealing to find global max WSR
    - distributed Channel State Information at the Transmitter (CSIT) acquisition, netDoF
    - Delayed CSIT, optimal handling of CSIT FB dead times
    - Finite Rate of Innovation (FRoI)/Basis Expansion Model (BEM) channel models
    - Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels (IBC)

IFC CSI Acquisition Outline

- Centralized CSIT Acquisition
- Distributed CSIT Acquisition
- Channel Feedback & Output Feedback
- DoF optimization as a function of coherence time

State of the Art on MIMO IFC w Partial CSI

- MISO BC (MU-MISO DL) w CSIT acquisition: [KobayashiCaireJindal:IT10]
- TDD MISO BC w CSIT acquisition: [SalimSlock:JWCN11]
- Space-Time Coding for Analog Channel Feedback: [ChenSlock:isit08]
- [NegroShenoySlockGhauri: eusipco09]: TDD MIMO IFC IA iterative design via UL/DL duality and TDD reciprocity
- Interference Alignment with Analog CSI Feedback: [ElAyachHeath:Milcom10]
  Centralized approach: BS's are connected to a central unit gathering all CSI, performing BF computations and redistributing BF's.
- [Jafar:GLOBECOM10] Blind IA
- [MaddahAliTse:allerton10] Delayed CSIT approach for $K = 2$ MISO BC
- [VazeVaranasi:submIT] DoF region for MIMO IFC w FB
- [SuhTse:IT11] GDoF for IFC with feedback

Key Points

- Distributed approach: no other connectivity assumed than the UL/DL IFC. FB over reversed IFC
- "distributed" = "duplicated" (decentralized)
- A distributed approach does not have to be iterative. It can be done with a finite overhead (finite prelog loss) and finite SNR loss compared to full CSI, even as SNR $\rightarrow \infty$. Hence of interest compared to non-coherent (no/outdated CSIT) IFC approaches.
- Distributed ($\mathcal{O}(K^2)$) requires more FB than Centralized ($\mathcal{O}(K)$).
- centralized/decentralized IFC CSIT estimation (only exchange of data at temporal coherence variation rate), vs NW-MIMO/CoMP (exchange of data at symbol/sample rate)
- Multiple Rx antennas $\Rightarrow$ Rx training also crucial!
- TDD vs FDD, depends on distributed/centralized.
- Channel FB vs Output feedback (OFB)
- "Practical" scheme far from unique

Signal Structure w Partial CSI

- Perfect CSI: Rx signal at the $k-$th receiver :

$$\mathbf{y}_k = \sum_{i=1}^{K}\sum_{m=1}^{d_i}\mathbf{H}_{ki}\mathbf{g}_{i,m}\,x_{i,m} + \mathbf{v}_k$$

  Estimate stream (k,n):

$$\widehat{x}_{k,n} = \mathbf{f}_{k,n}\mathbf{H}_{kk}\mathbf{g}_{k,n}\,x_{k,n} + \sum_{i=1}^{K}\sum_{m\neq n}\mathbf{f}_{k,n}\mathbf{H}_{ki}\mathbf{g}_{i,m}\,x_{i,m}\mathbf{f}_{k,n}\mathbf{v}_k$$

- Imperfect CSI: $\underbrace{\widehat{\widehat{\mathbf{f}}}_{k,n}}_{\text{est. at Rx } k}$ $\underbrace{\mathbf{H}_{ki}}_{\text{true}}$ $\underbrace{\widehat{\mathbf{g}}_{i,m}}_{\text{est. at Tx } i}$
- signal of interest in direct link:

$$\widehat{\widehat{\mathbf{f}}}_{k,n}\mathbf{H}_{kk}\,\widehat{\mathbf{g}}_{k,n} = \underbrace{\widehat{\widehat{\mathbf{f}}}_{k,n}\widehat{\mathbf{H}_{kk}\,\widehat{\mathbf{g}}_{k,n}}}_{\text{known to Rx}} + \underbrace{\widehat{\widehat{\mathbf{f}}}_{k,n}\widetilde{\mathbf{H}_{kk}\,\widehat{\mathbf{g}}_{k,n}}}_{\text{put in interf.}}$$

3 Partial CSI Rate Analysis Approaches (1)

1. Bound loss of partial CSI ergodic rate to full CSI ergodic rate.

$$\text{e.g.} \quad \mathcal{R}_k^{PCSI}(\rho) \le \left(1 - \frac{T_{overhead}}{T}\right) \mathcal{R}_k^{FCSI}(\rho/\alpha_k)$$

-

$$\begin{aligned} \widehat{\widehat{\mathbf{f}}}_{k,n}\, \mathbf{H}_{ki}\, \widehat{\mathbf{g}}_{i,m} &= (\mathbf{f}_{k,n} + \widetilde{\widetilde{\mathbf{f}}}_{k,n})\, \mathbf{H}_{ki}\, (\mathbf{g}_{i,m} + \widetilde{\mathbf{g}}_{i,m}) \\ &= \mathbf{f}_{k,n}\, \mathbf{H}_{k,i}\, \mathbf{g}_{i,m} + 3 \text{ error terms} \end{aligned}$$

3 Partial CSI Rate Analysis Approaches (2)

2. Bound loss of partial CSI ergodic rate to full CSI ergodic rate for case of channel pdf = that of the estimated channel: provides closer bounds, but requires ergodic rate expressions with different channel statistics.

   - $$\begin{aligned} \widehat{\widehat{\mathbf{f}}}_{k,n}\,\mathbf{H}_{ki}\,\widehat{\mathbf{g}}_{i,m} &= (\widehat{\mathbf{f}}^{(i)}_{k,n} + \widetilde{\widehat{\mathbf{f}}}^{(i)}_{k,n})\,(\widehat{\mathbf{H}}^{(i)}_{ki} + \widetilde{\mathbf{H}}^{(i)}_{ki})\,\widehat{\mathbf{g}}_{i,m} \\ &= \widehat{\mathbf{f}}^{(i)}_{k,n}\widehat{\mathbf{H}}^{(i)}_{ki}\widehat{\mathbf{g}}_{i,m} + 3 \text{ error terms} \end{aligned}$$

3 Partial CSI Rate Analysis Approaches (3)

3. High SNR ρ rate asymptote: $\mathcal{R} = a\log(\rho) + b + \mathcal{O}(1/\rho)$
   a: multiplexing gain (prelog, dof), b: rate offset a, b independent of:
   - MMSE regularization (MMSE-ZF filters suffice)
   - optimized WF (uniform WF suffices)
   - LMMSE channel estimation (becomes deterministic estimation)
   -

$$
\begin{aligned}
\widehat{\widehat{\mathbf{f}}}_{k,n}\,\mathbf{H}_{ki}\,\widehat{\mathbf{g}}_{i,m} &= (\widehat{\mathbf{f}}^{(i)}_{k,n} + \widetilde{\widehat{\mathbf{f}}}^{(i)}_{k,n})\,(\widehat{\mathbf{H}}^{(i)}_{ki} + \widetilde{\mathbf{H}}^{(i)}_{ki})\,\widehat{\mathbf{g}}_{i,m} \\
&= \underbrace{\widehat{\mathbf{f}}^{(i)}_{k,n}\widehat{\mathbf{H}}^{(i)}_{ki}\widehat{\mathbf{g}}_{i,m}}_{=0} + \mathbf{f}_{k,n}\,\widetilde{\mathbf{H}}^{(i)}_{ki}\,\mathbf{g}_{i,m} + \widetilde{\widehat{\mathbf{f}}}^{(i)}_{k,n}\,\mathbf{H}_{ki}\,\mathbf{g}_{i,m}
\end{aligned}
$$

High SNR Rate Analysis

- Asymptote $\mathcal{R} = a \log(\rho) + b$ permits meaningful optimization for finite (but high) SNR, and may lead to more than minimal FB.
- At very high SNR ρ, only rate prelog a (dof) counts. Its maximization requires FB to be minimal (channel just identifiable).
- At moderate SNR, finding an optimal compromise between estimation overhead and channel quality will involve a properly adjusted overhead. However, the overhead issue is not the only reason for a possibly diminishing multiplexing gain a as SNR decreases, also reducing the number of streams $\{d_k\}$ may lead to a better compromise (as for full CSI).
- The rate offset b is already a non-trivial rate characteristic even in the full CSI case. b may increase as the number of streams decreases, due to reduced noise enhancement.

Unification Stationary & Block Fading

- Doppler Spectrum is bandlimited to $1/T$ ($1/D$ in figure)
- Nyquist's Theorem : downsampling possible with factor T
- Vectorize channel coefficients over T, matrix spectrum of rank 1, MIMO prediction error of rank 1.
- Hence channel coefficient evolution during current "coherence period" T is along a single basis vector, plus prediction from past.
- Block fading: basis vector = rectangular window and prediction from the past = 0

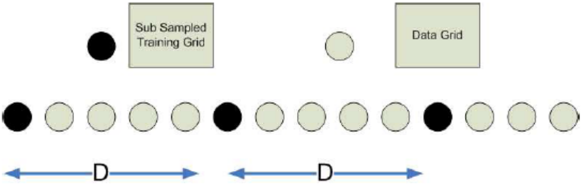
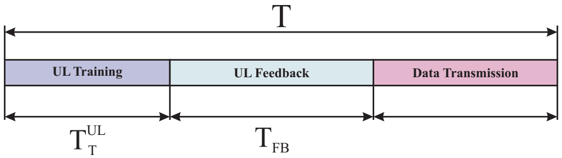


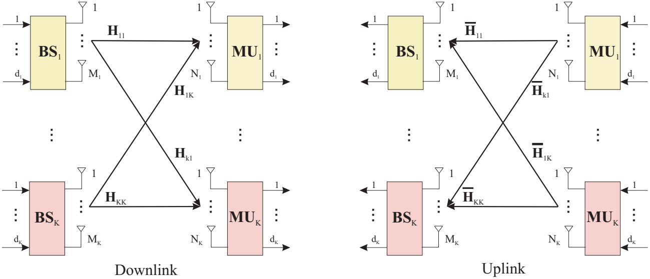
Figure 1: Subsampling Grid.

Centralized Approach



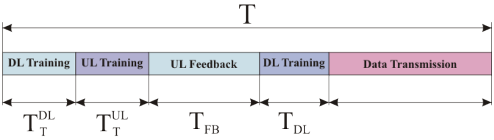
- Proposed by Heath [Milcom10,arxiv]
- The authors extrapolate the single antenna case, where only the estimate of the overall ch-BF gain and associated SINR is required
- In the MIMO IFC Rx not only needs to estimate the ch-BF cascade but also the I+N covariance matrix
- Not trivial. Training length similar as for the BF determination (order K) is required.
- Rate analysis of type 1.

FDD Communication



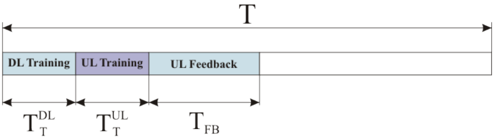
- We Assume FDD transmission scheme
- Downlink channel matrix $\mathbf{H}_{ki}$ from BS_i to MU_k
- Uplink channel matrix $\overline{\mathbf{H}}_{ik}$ from MU_k to BS_i
- Analyze both centralized and distributed approaches.

Transmission Phases



- We consider a block fading channel model with Coherence time interval T
- The general channel matrix $\mathbf{H}_{ik} \sim \mathcal{N}(0, \mathbf{I})$
- To acquire the necessary CSI at BS and MU side several training and feedback phases are necessary
- Hence a total overhead of T_{ovrhd} channel usage is dedicated to BS-MU signaling
- Only part of the time $T_{data} = T - T_{ovrhd}$ is dedicated to real data transmission

Uplink Feedback Phase



- After the UL and DL training phases each device knows all channels directly connected to it
- To compute the Tx beamformers, complete IFC channel knowledge is required
- Each MU feeds back its channel knowledge (CFB) using *Analog Feedback*
- Two different approaches are possible:
  (a) Centralized Processing
  (b) Distributed Computation

(a) A Central Controller acquires complete CSI and computes all the BF, and disseminates this information.

(b) Each BS acquires complete CSI to compute all the BF, then uses only it own BF.

Uplink Feedback Phase: Centralized Processing

- The received symbol vector received at each BS is sent to the Central Controller for the estimation of DL channels. Staking all the received symbols together we get:

$$\overline{\mathbf{Y}} = \sqrt{P_{FB}} \underbrace{\begin{bmatrix} \overline{\mathbf{H}}_{11} & \dots & \overline{\mathbf{H}}_{1K} \\ \vdots & \ddots & \vdots \\ \overline{\mathbf{H}}_{K1} & \dots & \overline{\mathbf{H}}_{KK} \end{bmatrix}}_{M \times N} \underbrace{\begin{bmatrix} \widehat{\mathbf{H}}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{H}}_2 & \dots & \mathbf{0} \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \dots & \widehat{\mathbf{H}}_K \end{bmatrix}}_{N \times KM} \underbrace{\begin{bmatrix} \mathbf{\Phi}_1 \\ \vdots \\ \mathbf{\Phi}_K \end{bmatrix}}_{KM \times T_{FB}} + \underbrace{\begin{bmatrix} \mathbf{V}_1 \\ \vdots \\ \mathbf{V}_K \end{bmatrix}}_{\overline{\mathbf{V}}}$$

where $N = \sum_i N_i$ and $M = \sum_i M_i$

- To satisfy the identifiability condition the minimum CFB length is

$$T_{FB} \geq \frac{N \times M}{\sum_i \min\{N_i, M_i\}} \propto K$$

- To extract the i-th feedback contribution we use LS estimate based on the UL channel estimate $\widehat{\overline{\mathbf{H}}}_{ik}$

Uplink Feedback Phase: Centralized Processing

$$\overline{\mathbf{Y}}\mathbf{\Phi}_i = \sqrt{P_{FB}} \underbrace{\begin{bmatrix} \overline{\mathbf{H}}_{i1} \\ \vdots \\ \overline{\mathbf{H}}_{iK} \end{bmatrix}}_{\overline{\mathbf{H}}_i} \widehat{\mathbf{H}}_i + \overline{\mathbf{V}}\mathbf{\Phi}_i$$

- Using the UL channel estimate the LS estimator is: $\overline{\mathbf{H}}_i^{LS} = P_{FB}^{-\frac{1}{2}} (\widehat{\overline{\mathbf{H}}}_i^H \widehat{\overline{\mathbf{H}}}_i)^{-1} \widehat{\overline{\mathbf{H}}}_i^H$

$$\widehat{\widehat{\mathbf{H}}}_i = \widehat{\mathbf{H}}_i + P_{FB}^{\frac{1}{2}} \overline{\mathbf{H}}_i^{LS} \widetilde{\overline{\mathbf{H}}}_i \widehat{\mathbf{H}}_i + \overline{\mathbf{H}}_i^{LS} \overline{\mathbf{V}}\mathbf{\Phi}_i = \mathbf{H}_i - \underbrace{\widetilde{\mathbf{H}}_i + P_{FB}^{\frac{1}{2}} \overline{\mathbf{H}}_i^{LS} \widetilde{\overline{\mathbf{H}}}_i \widehat{\mathbf{H}}_i + \overline{\mathbf{H}}_i^{LS} \overline{\mathbf{V}}\mathbf{\Phi}_i}_{\widetilde{\widehat{\mathbf{H}}}_i}$$

- The estimate of the CFB can be written in function of the true DL channel $\mathbf{H}_i$ plus the estimation error $\widetilde{\widehat{\mathbf{H}}}_i$

$$\mathrm{Cov}(\widetilde{\widehat{\mathbf{H}}}_i | \widehat{\overline{\mathbf{H}}}_i) = \sigma^2_{\widetilde{\mathbf{H}}_i} \mathbf{I} + [(\sigma^2_{\widehat{\mathbf{H}}_i} \sigma^2_{\widetilde{\overline{\mathbf{H}}}_i}) + \frac{\sigma^2}{P_{FB}}] (\widehat{\overline{\mathbf{H}}}_i^H \widehat{\overline{\mathbf{H}}}_i)^{-1}$$

- The estimation error is then distributed as $\mathcal{N}(0, \sigma^2_{\widetilde{\widehat{\mathbf{H}}}_i})$

Uplink Feedback Phase: Distributed Processing

- The received symbols at BS_k can be described as follows

$$\overline{\mathbf{Y}}_k = \sqrt{P_{FB}}\underbrace{\begin{bmatrix} \overline{\mathbf{H}}_{k1} & \dots & \overline{\mathbf{H}}_{kK} \end{bmatrix}}_{M_k \times N} \underbrace{\begin{bmatrix} \widehat{\mathbf{H}}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{H}}_2 & \dots & \mathbf{0} \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \dots & \widehat{\mathbf{H}}_K \end{bmatrix}}_{N \times KM} \underbrace{\begin{bmatrix} \mathbf{\Phi}_1 \\ \vdots \\ \mathbf{\Phi}_K \end{bmatrix}}_{KM \times T_{FB}} + \mathbf{V}_k$$

- To satisfy the identifiability condition the minimum CFB length is

$$T_{FB} \geq \frac{N \times M}{\min_i\{M_i, N_i\}} \propto K^2$$

- To extract the i-th feedback contribution at BS_k we use LS estimate based on the UL channel estimate $\widehat{\overline{\mathbf{H}}}_{ki}$

Uplink Feedback Phase: Distributed Processing

$$\overline{\mathbf{Y}}_k \mathbf{\Phi}_i = \sqrt{P_{FB}} \overline{\mathbf{H}}_{ki} \widehat{\mathbf{H}}_i + \mathbf{V}_k \mathbf{\Phi}_i$$

- Using the UL channel estimate the LS estimator is: $\overline{\mathbf{H}}_{ki}^{LS} = P_{FB}^{-\frac{1}{2}} (\widehat{\overline{\mathbf{H}}}_{ki}^{H} \widehat{\overline{\mathbf{H}}}_{ki})^{-1} \widehat{\overline{\mathbf{H}}}_{ki}^{H}$

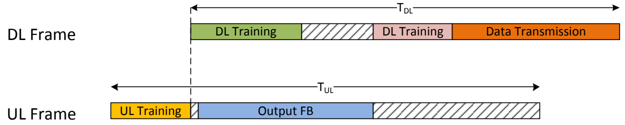
$$\widehat{\widehat{\mathbf{H}}}_i = \widehat{\mathbf{H}}_i + P_{FB}^{\frac{1}{2}} \overline{\mathbf{H}}_{ki}^{LS} \widetilde{\overline{\mathbf{H}}}_{ki} \widehat{\mathbf{H}}_i + \overline{\mathbf{H}}_{ki}^{LS} \mathbf{V}_k \mathbf{\Phi}_i = \mathbf{H}_i - \underbrace{\widetilde{\mathbf{H}}_i + P_{FB}^{\frac{1}{2}} \overline{\mathbf{H}}_{ki}^{LS} \widetilde{\overline{\mathbf{H}}}_{ki} \widehat{\mathbf{H}}_i + \overline{\mathbf{H}}_{ki}^{LS} \mathbf{V}_k \mathbf{\Phi}_i}_{\widetilde{\widehat{\mathbf{H}}}_i}$$

- The estimate of the CFB can be written in function of the true DL channel $\mathbf{H}_i$ plus the estimation error $\widetilde{\widehat{\mathbf{H}}}_i$

$$\mathrm{Cov}(\widetilde{\widehat{\mathbf{H}}}_i | \widehat{\overline{\mathbf{H}}}_{ki}) = \sigma^2_{\widetilde{\mathbf{H}}_i} \mathbf{I} + [(\sigma^2_{\widehat{\mathbf{H}}_i} \sigma^2_{\widetilde{\overline{\mathbf{H}}}_{ki}}) + \frac{\sigma^2}{P_{FB}}] (\widehat{\overline{\mathbf{H}}}_{ki}^{H} \widehat{\overline{\mathbf{H}}}_{ki})^{-1}$$

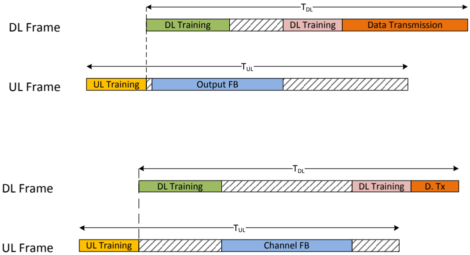
- The estimation error is then distributed as $\mathcal{N}(0, \sigma^2_{\widetilde{\widehat{\mathbf{H}}}_i})$

Output Feedback



- Each MU feeds back to all BS the noiseless version of its received signal using un-quantized feedback: Output FB (OFB).
- In FDD systems UL and DL transmission can take place at the same time
- T_{UL} represents the UL coherence Time
- T_{DL} represents the DL coherence Time
- OFB phase can start one time instant after the beginning of the DL training phase

Output Feedback



- Output FB allows us to reduce the overhead due to CSI exchange
- In channel FB each MU has to wait the end of the DL training phase before being able to FB DL channel estimates
- For easy of exposition we consider $M_i = N_t \; \forall i$, $N_i = N_r \; \forall i$ where $N_t \geq N_r$

Importance of CSIR

- CSIR is usually neglected
- Some schemes for arbitrary time-varying channels assume that Rxs know all channel matrices at all time: impossible to realize in practice
- An additional DL training phase is required to build the Rx filters

In the end: Sum Rate (high SNR)

- SR

$$\mathcal{R}^{PCSI} = \sum_{k,n} \underbrace{(1 - \frac{\sum T_i}{T})}_{\substack{\text{reduced data} \\ \text{channel uses}}} \ln(|\mathbf{f}_{kn}\mathbf{H}_{kk}\mathbf{g}_{kn}|^2 \underbrace{\rho/(1 + \sum_i \frac{b_{kni}}{T_i}))}_{\text{SNR loss}} ,$$

$$T_i \geq T_{i,min}$$

Assume $b_{kni} = b_i$ for what follows.

- Fixing $\sum_i T_i = T_{ovrhd}$, optimal $T_i = T_{ovrhd} \sqrt{b_i}/(\sum_i \sqrt{b_i})$.
- Optimizing over T_{ovrhd} now

$$T_{ovrhd} = \frac{\sqrt{T} \, (\sum_i \sqrt{b_i})}{\sqrt{\mathcal{R}^{FCSI}}}$$

Sum Rate at even higher SNR: DoF

- DoF optimization as function of Coherence Time
- full CSI: maximize $d_{tot} = \sum_k d_k$
- with CSI acquisition, larger MIMO systems need more training/FB, hence at short coherence times, the MIMO dimensions and number of streams may need to be reduced
- symmetric systems $(N, N, d)^K$: $\quad N \geq \frac{d}{2}(K+1)$
- DL time overhead for CFB as:

$$T_{ovrhd} = T_T^{DL} + T_{FB} + T_{DL} = \begin{cases} dK(K+2) & \text{(Centr.)} \\ \frac{Kd}{2}((K+1)^2+2) & \text{(Distr.)} \end{cases}$$

- DoF term in Sum Rate:

$$\max_d J(d) = \max_d (1 - \frac{T_{ovrhd}}{T}) Kd \log SNR$$

DoF Optimization (2)

- $\frac{\partial J}{\partial d} = 0 \;\Rightarrow\; d^* = \begin{cases} \frac{T}{2K(K+2)} & \text{(Centr.)} \\ \frac{T}{K[(K+1)^2+2]} & \text{(Distr.)} \end{cases}$
- Hence

$$d \leq \min\left\{ d^*, \frac{2N}{K+1} \right\}$$

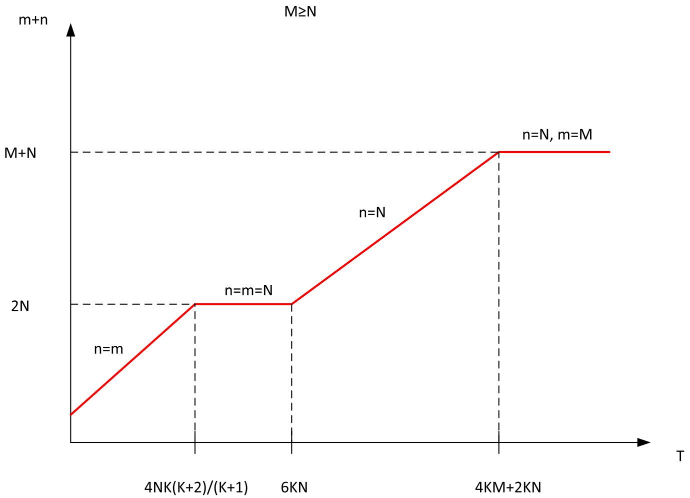
$$\Rightarrow \quad d = \begin{cases} \min\left\{ \frac{T}{2K(K+2)}, \frac{2N}{K+1} \right\} & \text{(Centr.)} \\ \min\left\{ \frac{T}{K[(K+1)^2+2]}, \frac{2N}{K+1} \right\} & \text{(Distr.)} \end{cases}$$

DoF Optimization (3)

- if
$$T \geq \frac{4NK(K+2)}{K+1} = 2T_{ovrhd}$$
then number of streams $d = \frac{2N}{K+1}$, kept at its maximum.
- otherwise shrink d and number of active antennas
$$n \geq \frac{d^*(K+1)}{2}$$
with a consequent reduction of the time overhead for CSI acquisition.
- similar analysis for output feedback

$(M, N, d)^K$ case

- Assuming $M \geq N$.
- Evolution of number of active transmit m and receive n antennas as a function of coherence time T.



TDD vs FDD

- Usually TDD transmission scheme is used to simplify the DL CSI acquisition at the BS side
- BS_k learns the DL channel $\mathbf{H}_{ik}$, $\forall i$ through reciprocity
- MU_i do not need to feedback $\mathbf{H}_{ik}$ to BS_k but this channel is required at $BS_{j \neq k}$
- In Distributed Processing reciprocity does NOT help in reducing channel feedback overhead $\Longrightarrow$ TDD almost equivalent to FDD
- In Centralized Processing reciprocity makes channel feedback NOT required

Further Optimizing DoF

- data Tx stage (as good as perfect CSI):
  - Can FB increase DoF with perfect CSIT?
    According to [HuangJafar:IT09] and [VazeVaranasi:ITsubm11]
    NO for $K = 2$ MIMO IFC; $K > 2$ is OPEN.
  - If not in general, then use of OFB is mainly (only) for CSI acquisition, not for augmenting DoF in presence of CSIT
- CSI acquisition stages:
  - Optimize number of streams/number of active antennas for small T: if less channel to learn then more time to Tx data, even if on reduced number of streams
  - Instead of going from $K = 1$ to full K immediately, could gradually increase number of interfering links (and their CSI acquisition) from 1 to K.
  - When T gets too short: delayed CSIT approaches.
    Optimal combination: do delayed CSIT during training dead times.
  - A single (the largest) MIMO link can start transmitting right away w/o CSIT (possibly w/o CSIR also).

Perspectives

- When (analog) channel FB is of extended (non-minimal) duration, BF's can get computed and some DL transmission could start while FB is still going on. No need to wait until all CSI is gathered before transmission can get started.
- rate constants: partial CSI Tx/Rx design, diversity issues (optimized IA)
- optimization of training duration/power
- can OFB increase dof w perfect CSIT for $K \geq 3$?
- need to handle CSIR also in delayed CSIT approaches
- users with different coherence times
- full duplex operation (2-way communications)
- minimum reciprocity:
  coherence times equal on UL and DL, feasible dof same on UL and DL
- real IFC system: doubly selective

Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - uplink/downlink(UL/DL) duality; SU MIMO,BC,MAC; BF&DPC
  - BC with user selection: DPC vs BF
  - Max WSR, UL/DL duality, CSIT: perfect, partial, LoS
- interference multi-cell/HetNets: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
  - distributed Channel State Information at the Transmitter (CSIT) acquisition, netDoF
  - Delayed CSIT, optimal handling of CSIT FB dead times
  - Finite Rate of Innovation (FRoI)/Basis Expansion Model (BEM) channel models
  - Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels (IBC)

Outline: back to Full DoF with DCSIT

- noisy MIMO IC w/o symbol extension
  centralized/distributed CSI FB - Tx/RX design
  netDoF concept
- Part 1: only accounting for feedback delay $\Rightarrow$ delayed CSIT
  - T_{fb} arbitrary: MAT
  - $T_{fb} < T_c$ case: [YangKobayashiGesbertYi:isit12/asilo12]
  - stationary-block fading unification: FRoI/BEM
- Part 2: no DoF loss from CSIT delay
  - ST-ZF
  - Foresighted Channel FB
- Part 3: FRoI filter optimization
  - single basis function
  - multiple basis functions

Some State of the Art

- perfect CSI
  - signal space interference alignment (joint Tx/Rx ZF)
    - Interference Alignment (IA) [Khandani etal]
    - [CadambeJafar:IT08] IA can get $K/2$ DoF in t/f selective SISO (hence MIMO) via symbol extension
    - noisy MIMO w/o symbol extension, IA feasibility [Santamaria etal:isit12],[RuanLau:isit12]
      Tx/Rx design: IA, max per stream SINR, max WSR
    - ergodic IA: group channel realizations H_1, H_2 s.t. $\text{offdiag}(H_2) = -\text{offdiag}(H_1)$, $\text{diag}(H_2) = \text{diag}(H_1)$
  - signal scale IA (Tx rational numbers, diophantine equation)
- delayed (perfect) CSI
  - MAT, retrospective IA, blind IA, others (see further)
- imperfect delayed CSI
  - CSI acquisition, training, feedback

Sum Rate at very high SNR: DoF

- At very high SNR: prelog dominates: Degrees of Freedom (DoF)
- netDoF concept :
  account for loss of DoF due to overheads:
  - forward training (common + dedicated)
  - forward dead time (during FB)
  - reverse link FB
- netDoF in literature:
  - netDoF only picked up also by [SuhTse:ita12/isit12]: duplex netDoF, only accounts for reverse link FB though
  - whereas [CaireKobayashi] netDof only account for (common only, not dedicated) training

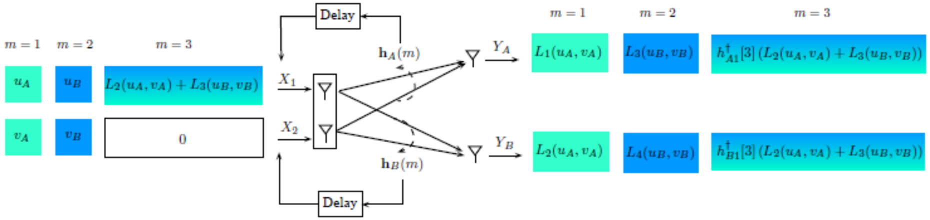
Part I

Only degradation from perfect CSI: CSIT is Delayed

- schemes

MAT: Maddah-Ali & Tse scheme

- (perfect) CSIT available only after FB delay T_{fb}
  (T_{fb} taken as unit of time in number of following schemes)
- channel correlation over T_{fb} arbitrary, possibly zero
- perfect overall CSIR assumed
- MISO BC (Broadcast Channel) exposed on next slide
  [MaddahAliTse:allerton08]
  MISO IC very similar
  some extensions to MIMO [VazeVaranasi:isit11]



MAT scheme for MISO BC with $M = K = 2$.

- for MISO BC with $M = K$, MAT allows to reach a multiplexing gain of
$$\frac{K}{1 + \frac{1}{2} \cdots \frac{1}{K}} = \frac{KD}{Q} \quad (\approx \frac{K}{\ln K})$$
with no *current* CSIT at all. Here $\{D, Q\} \in \mathbb{N}^2$ such that $\frac{1}{1+\frac{1}{2}\cdots\frac{1}{K}} = \frac{D}{Q}$, where D is the least common multiple of $\{1, 2, \cdots, K\}$ and $Q = DH_K$ with $H_K = \sum_{k=1}^{K} \frac{1}{k}$.
- allows per user Tx of D symbols in Q channel uses.

[YangKobayashiGesbertYi:isit12/asilo12]: extending MAT to $T_{fb} < T_c$

- assume channel piecewise constant over T_{fb}, for the rest (cyclo)stationary
- exploit $T_{fb} < T_c$ (= coherence period = 1 / Doppler BW)
- focus on temporal correlation of one channel coefficient h (enough for DoF considerations):

  channel FB: estimate and error: $h = \widehat{\widehat{h}} + \widetilde{\widetilde{h}}\,,\ \ \frac{\sigma^2_{\widetilde{\widetilde{h}}}}{\sigma^2_{\widehat{\widehat{h}}}} = \mathcal{O}(\frac{1}{\rho})$

  At Tx, on basis of $\widehat{\widehat{h}}$, channel prediction over T_{fb} and prediction error: $h = \widehat{h} + \widetilde{h}\,,\ \ \frac{\sigma^2_{\widetilde{h}}}{\sigma^2_{\widehat{h}}} = \mathcal{O}(\rho^{-(1-\frac{T_{fb}}{T_c})})$
- Attain sumDoF $= 2(1 - \frac{T_{fb}}{3T_c}) = 2(\frac{2}{3}\frac{T_{fb}}{T_c} + 1 - \frac{T_{fb}}{T_c})$
- Mostly MISO (BC or IC). Limited to $K = 2$. FB every T_{fb}.
- They also consider: imperfect CSIT (apart from delayed), DoF region.

Outline

- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - uplink/downlink(UL/DL) duality; SU MIMO,BC,MAC; BF&DPC
  - BC with user selection: DPC vs BF
  - Max WSR, UL/DL duality, CSIT: perfect, partial, LoS
- interference multi-cell/HetNets: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
  - distributed Channel State Information at the Transmitter (CSIT) acquisition, netDoF
  - Delayed CSIT, optimal handling of CSIT FB dead times
  - Finite Rate of Innovation (FRoI)/Basis Expansion Model (BEM) channel models
  - Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels (IBC)

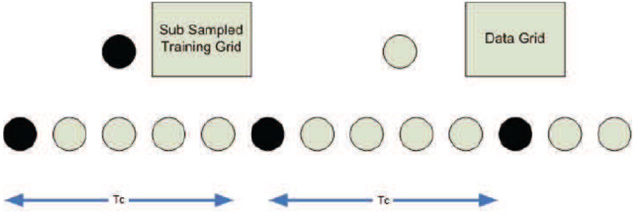


Figure : Subsampling and polyphase representation of a bandlimited channel coefficient signal.

- already introduced in [Salim:PhD08]
- for DoF considerations: sufficient to focus on any scalar channel coefficient separately (to be optimal at finite SNR, treat all correlated channel coefficients jointly).
- Assume the channel coefficient h_k has a Doppler spectrum strictly bandlimited to $1/T_c$.
- Assume for a moment $T = T_c$ to be an integer number of symbol periods.

Finite Rate of Innovation (FRoI)/ Basis Expansion Model (BEM)

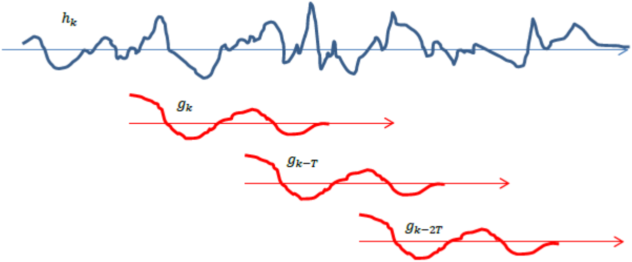


Figure : Subsampling and reconstruction from basis functions.

- block fading: $g_k = \begin{cases} 1 & , \; k = 0, 1, \ldots, T-1 \\ 0 & , \; \text{elsewhere} \end{cases}$
- stationary bandlimited (BL): $g_k = \text{sinc}(\pi k/T) = \dfrac{\sin(\pi k/T)}{\pi k/T}$

FRoI/BEM (2)

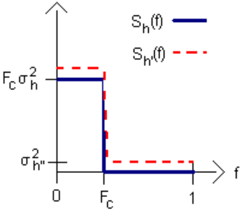
- BEM: [Grenier:PhD-ENST82], [TsatsanisGiannakis:96]
- FRoI: [VetterliMarzilianoBlu:TSP02], [VetterliKovecevicGoyal:FoundationsofSignalProcessing'13]
  FRoI: finite # parameters/sample $= 1/T$, here; linear FRoI
- Filterbank with a single subband. Synthesis filter g_k, analysis filter f_k.

$$a_n = \sum_k f_k \, h_{nT-k}$$

$$h_{nT+i} = \sum_{l=0}^{L-1} a_{n-l} \, g_{lT+i} \; , \quad i = 0, 1, \ldots, T-1 \; .$$

- Perfect reconstruction for BL: $g_k * f_k = \mathrm{sinc}(\pi k/T)$.
  E.g. $g_k = \mathrm{sinc}(\pi k/T)\,,\; f_k = \delta_{k0}$.
- Case of causal g_k, $f_k = g^*_{-k}$, $(g_k * g^*_{-k})_{k=nT} = \delta_{n0}$: reconstructed signal = least-squares projection on subspace of BL signals.
  Requires $f_k = g^*_{-k}$ (matched filter) to be non-causal!
  Impractical for channel feedback (both g_k & f_k causal).

Noisy BL signals



- $S(f) = \dfrac{\widetilde{\sigma}^2}{|A(f)|^2} = \widetilde{\sigma}^2 \, |B(f)|^2$
- $\widetilde{\sigma}^2 = e^{\int \ln S(f)\, df} = (T\sigma_h^2)^{1/T} \sigma_v^{2(1-1/T)}$ (high SNR)
- $|A(f)|^2 = \dfrac{\widetilde{\sigma}^2}{S(f)} = \begin{cases} (\frac{T\sigma_h^2}{\sigma_v^2})^{-(1-1/T)} & f \in [0, 1/T] \\ (\frac{T\sigma_h^2}{\sigma_v^2})^{1/T} & f \in [1/T, 1] \end{cases},$

  $||A||^2 = 1/T(\frac{T\sigma_h^2}{\sigma_v^2})^{-(1-1/T)} + (1-1/T)(\frac{T\sigma_h^2}{\sigma_v^2})^{1/T} \to \infty$
- $|B(f)|^2 = 1/|A(f)|^2$, $B =$ monic causal spectral factor

  $||B||^2 = 1/T(\frac{T\sigma_h^2}{\sigma_v^2})^{1-1/T} + (1-1/T)(\frac{T\sigma_h^2}{\sigma_v^2})^{-1/T} \to \infty$

Model details in existing works: no exact BL model anywhere!

- block fading: MAT, MAT-ZF, [ChenElia], [LeeHeath] etc
- [YangKobayashiGesbertYi:isit12/asilo12] :
  2 possible interpretations
  - block fading over T_{fb} + BL between blocks
  - really BL: but need to predict over T_{fb} every sample $\Rightarrow$ FB every sample!
- [KobayashiCaire:isit12]:
  block fading over T_c + BL between blocks
- FRoI: behaves like block fading, but closer to reality
  * finite length basis functions (cannot predict from ∞ past)
  * as a result: effect of noise remains $O(\sigma_v^2)$
- real channels are not BL: Doppler shifts are time-varying!
- BL still interesting: only FRoI that is stationary
  (FRoI = cyclo-stationary in general)

Delayed CSIT

- can assume block fading model henceforth, block length = coherence period T_c
- CSIT feedback delay T_{fd}
- block T_c can be slit into 2 parts:
  - $0 \leq t < T_{fd}$: the current channel state is unknown to the transmitter
  - $T_{fd} \leq t < T_c$: the transmitter has full CSI
- idea: use two different techniques within each block, the MAT scheme when the current channel state is unknown and then ZF for $t \geq T_{fd}$. Both techniques have been proven to be optimal in terms of multiplexing gain in their respective settings.
- We first review the multiplexing gains achievable with these schemes.

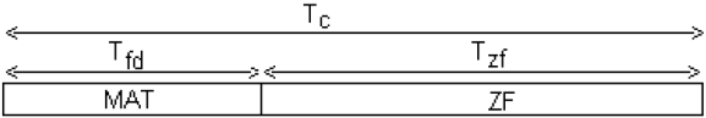
MISO BC/IC Zero Forcing BF (ZF)

- with full CSIT: full DoF can be achieved with ZF
- transmitter uses a pseudo inverse of the channel as precoder thereby zero-forcing all inter-user interferences
- ZF only $\Rightarrow$ allows to transmit 1 symbol per channel use in the second part of each block and nothing in the first part, yielding ergodic DoF

$$\mathrm{DoF}(\mathrm{ZF}_K) = K\mathrm{DoF}(\mathrm{ZF}_1) = K\left(1 - \frac{T_{fd}}{T_c}\right).$$

MAT-ZF scheme

- idea: essentially perform ZF and superpose MAT only during the dead times of ZF.
- Applicable to any MMO IBC, consider MISO (BC) here for details.
- For that purpose we consider Q blocks of T_c symbol periods and split each block into two parts as in the Figure.
- The first part, the dead times of ZF, spans T_{fd} symbol periods and the second part, the $T_c - T_{fd}$ remaining symbols.
- We use the first part of each block to perform the MAT scheme T_{fd} times in parallel.
- During the second part of each block, ZF is performed.



DoF MAT-ZF

Theorem

The sum DoF for the MAT-ZF$_K$ scheme is

$$DoF(MAT\text{-}ZF_K) = K\left(1 - \frac{(Q-D)T_{fd}}{QT_c}\right).$$

Proof.

Per user, in QT_c channel uses (Q coherence periods), the ZF portion transmits $Q(T_c - T_{fd})$ symbols, whereas the MAT scheme transmits DT_{fd} symbols. □

Optimality DoF MAT-ZF

Theorem

The MAT-ZF$_K$ scheme is optimal in terms of sum multiplexing gain i.e., for any transmission scheme ψ_K for the MISO BC with K users,

$$DoF(\psi_K) \leq DoF(MAT\text{-}ZF_K)\ .$$

Proof.

The MAT-ZF$_K$ approach decomposes the channel with feedback delay into two orthogonal parts: the ZF part in which CSIT is perfect, and the MAT part with delayed CSIT.

In the ZF part, the relative portion of which is maximal, ZF allows to obtain the DoF of the full CSIT case.

In the MAT part, the MAT scheme has been shown to maximize DoF for the case of delayed CSIT with block size equal to T_{fd}. □

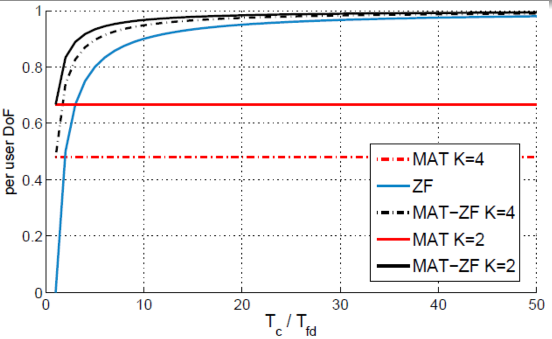


Figure : Per user DoFs as a function of T_c/T_{fd} for $K \in \{2, 4\}$.

- The DoF being an increasing function of T_c/T_{fd}, and the coherence time being a fixed parameter of the channel, the feedback delay should be reduced to its minimum in order to improve the multiplexing gain.
- We can already notice that for $K = 2$ the gap between MAT-ZF and pure ZF is larger than for $K = 4$ hinting that the gain due to the optimal combining of MAT and ZF could be decreasing with the number of users.

Part II

Schemes without DoF loss due to CSIT delay

- spatiotemporal ZF (ST-ZF) [LeeHeath:allerton12,asilomar12]
- FRoI/BEM models and increase FB sampling rate
  Foresighted Channel Feedback (FCFB)

Getting to full DoF: ST-ZF

- MISO BC/IC: [LeeHeath:allerton12/asilo12]
- ingredients:
  - symbol extension (t-variation required): space-time ZF precoding
  - due to CSIT delay, transmit fewer symbols per user
  - but make up by overloading, to get full sumDoF
  - send M symbols to $K = M + 1$ users over $M + 1$ T_c's
- scheme also valid for stationary fading due to stationary/block fading equivalence

ST-ZF example, MISO BC $M = 2$

$$\begin{bmatrix} \mathbf{y}[1] \\ \mathbf{y}[6] \\ \mathbf{y}[8] \end{bmatrix} = \begin{bmatrix} \mathbf{H}[1] & 0 & 0 \\ 0 & \mathbf{H}[6] & 0 \\ 0 & 0 & \mathbf{H}[8] \end{bmatrix} \begin{bmatrix} I_2 & I_2 & I_2 \\ \mathbf{V}^{(1)}[6] & \mathbf{V}^{(2)}[6] & \mathbf{V}^{(3)}[6] \\ \mathbf{V}^{(1)}[8] & \mathbf{V}^{(2)}[8] & \mathbf{V}^{(3)}[8] \end{bmatrix} \begin{bmatrix} \mathbf{s}^{(1)}] \\ \mathbf{s}^{(2)}] \\ \mathbf{s}^{(3)}] \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{H}[1] & \mathbf{H}[1] & \mathbf{H}[1] \\ \mathbf{H}[6]\mathbf{V}^{(1)}[6] & \mathbf{H}[6]\mathbf{V}^{(2)}[6] & \mathbf{H}[6]\mathbf{V}^{(3)}[6] \\ \mathbf{H}[8]\mathbf{V}^{(1)}[8] & \mathbf{H}[8]\mathbf{V}^{(2)}[8] & \mathbf{H}[8]\mathbf{V}^{(3)}[8] \end{bmatrix} \begin{bmatrix} \mathbf{s}^{(1)} \\ \mathbf{s}^{(2)} \\ \mathbf{s}^{(3)} \end{bmatrix}$$

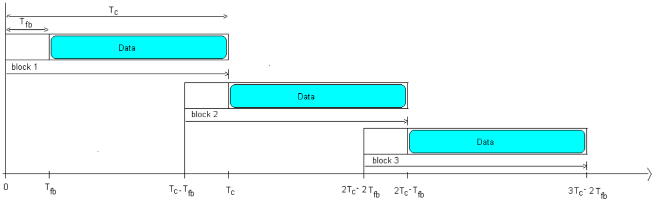
For user i, at time $n \in \{6, 8\}$ we have
$y^{(i)}[1] - y^{(i)}[n] = \sum_{k=1}^{3} \left(\mathbf{h}^{(i)}[1] - \mathbf{h}^{(i)}[n]\mathbf{V}^{(k)}[n]\right) \mathbf{s}^{(k)}$
so the interferences are aligned if

$$\mathbf{h}^{(i)}[1] - \mathbf{h}^{(i)}[n]\mathbf{V}^{(k)}[n] = \mathbf{0}, \ \forall i \neq k.$$

then user i can decode from square mixture

$$\begin{bmatrix} y^{(i)}[1] - y^{(i)}[6] \\ y^{(i)}[1] - y^{(i)}[8] \end{bmatrix} = \begin{bmatrix} \mathbf{h}^{(i)}[1] - \mathbf{h}^{(i)}[n]\mathbf{V}^{(i)}[n] \\ \mathbf{h}^{(i)}[1] - \mathbf{h}^{(i)}[n]\mathbf{V}^{(i)}[n] \end{bmatrix} * \mathbf{s}^{(i)}$$

ZF w Foresighted Channel Feedback (FCFB)



- key idea: stationarity

  if a FRoI/BEM model is good enough, then so is any shifted version!
- FRoI/BEM model allows to predict CSIT over coherence period $T = T_c$.
- Overlap basis functions by FB delay T_{fb}, then have CSIT all the time!
- works for any multi-user system (BC, IC, MAC etc)

Beyond FRoI: Predictive Rate Distortion

- FRoI is one way to get certain rate (DoF) for a distortion of $O(\sigma_v^2)$ (noise level)
- more generally: predictive R-D theory
  requires (new) channel models
- related work:

  [GoldsmithEldar:ita13]: filter f not causal or optimized
  [SilvaDerpichOstergaard:ita13] (and refs): causal R-D

NetDoFs [isit13]

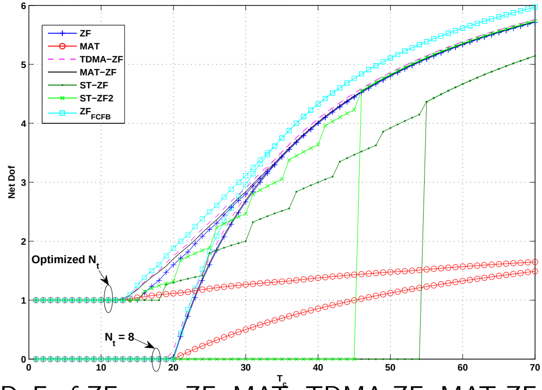


Figure : NetDoF of ZF_{FCFB}, ZF, MAT, TDMA-ZF, MAT-ZF, ST-ZF and TDMA and their optimized variants for $N_t = 8,\ T_{fb} = 3$ as a fn of T_c.

Due to enormous CSIR distribution overhead, MAT needs enormous coherence time T_c to reach its ideal DoF.

Concluding Remarks

- DoF in multi-user systems accounting for (channel) feedback are extremely sensitive to channel model.
- All this argues for shrinking the Feedback delay as much as possible: in FDD, feedback delay can be shrunk to roundtrip delay! **Immediate Feedback**.

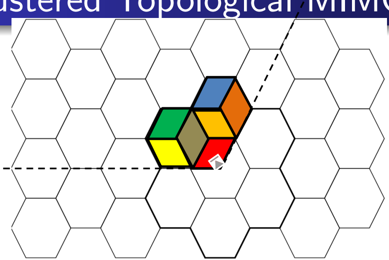
Outline

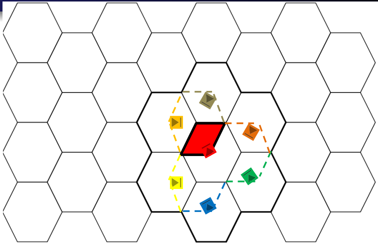
- interference single cell: Broadcast Channel (BC)
  - utility functions: SINR balancing, (weighted) sum rate (WSR)
  - uplink/downlink(UL/DL) duality; SU MIMO,BC,MAC; BF&DPC
  - BC with user selection: DPC vs BF
  - Max WSR, UL/DL duality, CSIT: perfect, partial, LoS
- interference multi-cell/HetNets: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and UL/DL duality
  - Deterministic Annealing to find global max WSR
  - distributed Channel State Information at the Transmitter (CSIT) acquisition, netDoF
  - Delayed CSIT, optimal handling of CSIT FB dead times
  - Finite Rate of Innovation (FRoI)/Basis Expansion Model (BEM) channel models
  - Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels (IBC)

Reduced CSIT and Decoupled Tx/Rx Design

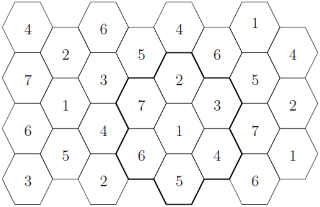
- for IA to apply to cellular: overall Tx/Rx design has to decompose so that the CSIT required is no longer global and remains bounded regardless of the network size.
- simplest case : local CSIT : a BS only needs to know the channels from itself to all terminals. In the TDD case : reciprocity. The local CSIT case arises when all ZF work needs to be done by the Tx: $d_{c,k} = N_{c,k}, \forall c, k$. The most straightforward such case is of course the MISO case: $d_{c,k} = N_{c,k} = 1$. It extends to cases of $N_{c,k} > d_{c,k}$ if less than optimal DoF are accepted. One of these cases is that of reduced rank MIMO channels.
- reduced CSIT [Lau:SP0913]: variety of approaches w reduced CSIT FB in exchange for DoF reductions.
- incomplete CSIT [deKerretGesbert:TWC13]: min some MIMO IC optimal DoF can be attained with less than global CSIT. Only occurs when M and/or N vary substantially so that subnetworks of a subgroup of BS and another subgroup of terminals arise in which the numbers of antennas available are just enough to handle the interference within the subnetwork.
- Massive MIMO leads to exploiting covariance CSIT, which will tend to have reduced rank and allows decoupled approaches.

Clustered Topological MIMO IBC





- attenuation ⇒ "banded" channel matrix (e.g. first tier)
- sectoring ⇒ "triangular" channel matrix (spatially causal)
- cover w clusters: treat one cluster as a IBC + ZF to neighboring Rx antennas
- cell or sector numbering: not for frequency reuse but for pilot (DL)/FB(UL) reuse



IA Feasibility Reduced Rank MIMO IBC

- The ZF from BS j to MT (i,k) requires

$$F_{i,k}^H H_{i,k,j} G_{j,n} = F_{i,k}^H \mathbf{B}_{i,k,j} A_{i,k,j}^H G_{j,n} = 0$$

which involves $\min(d_{i,k}d_{j,n}, d_{i,k}r_{i,k,j}, r_{i,k,}d_{j,n})$ constraints to be satisfied by the $(N_{i,k}-d_{i,k})d_{i,k}/(M_j-d_{j,n})d_{j,n}$ variables parameterizing the column subspaces of $F_{i,k}/G_{j,n}$.

- **IA feasibility singular MIMO IC with Tx/Rx decoupling**

$$F_{i,k}^H \mathbf{B}_{i,k,j} = 0 \text{ or } A_{i,k,j}^H G_{j,n} = 0\ .$$

This leads to a possibly increased number of ZF constraints $r_{i,k,j}\min(d_{i,k}, d_{j,n})$ and hence to possibly reduced IA feasibility. ZF of every cross link now needs to be partitioned between all Txs and Rxs, taking into account the limited number of variables each Tx or Rx has. The main goal of this approach however is that it leads to Tx/Rx decoupling.

Massive MIMO & Covariance CSIT

In massive MIMO, the Tx side channel covariance matrix is very likely to be (very) singular even though the channel response H may not be singular:

$$\text{rank}(C^t_{i,k,j} = A_{i,k,j}A^H_{i,k,j}) = r_{i,k,j}\,,\; A_{i,k,j} : M_j \times r_{i,k,j}$$

Let $P_{\mathbf{X}} = \mathbf{X}(\mathbf{X}^H\mathbf{X})^{\#}\mathbf{X}^H$ and $P^{\perp}_{\mathbf{X}}$ be the projection matrices on the column space of $\mathbf{X}$ and its orthogonal complement resp. Consider now a massive MIMO IBC with C cells containing K_i users each to be served by a single stream. The following result states when this will be possible.

Theorem

Sufficiency of Covariance CSIT for Massive MIMO IBC *In the MIMO IBC with (local) covariance CSIT, all BS will be able to perform ZF BF if the following holds*

$$||P^{\perp}_{A_{\overline{i,k},j}} A_{i,k,j}|| > 0\,,\ \forall i, k, j$$

where $A_{\overline{i,k},j} = \{A_{n,m,j},\, (n,m) \neq (i,k)\}$.

Massive MIMO & Covariance CSIT (2)

These conditions will be satisfied w.p. 1 if $\sum_{i=1}^{C}\sum_{k=1}^{K_i} r_{i,k,j} \leq M_j\,,\ j=1,\ldots,C$. In that case all the column spaces of the $A_{i,k,j}$ will tend to be non-overlapping. However, the conditions could very well be satisfied even if these column spaces are overlapping, in contrast to what [Gesbert:arxiv1013],[Caire:arxiv0912] appear to require. In Theorem 4, we assume that all ZF work is done by the BS. However, if the MT have multiple antennas, they can help to a certain extent.

Theorem

Role of Receive Antennas in Massive MIMO IBC *If MT* (i,k) *disposes of* $N_{i,k}$ *antennas to receive a stream, it can perform rank reduction of a total amount of* $N_{i,k}-1$ *to be distributed over* $\{r_{i,k,j}\,,\ j=1,\ldots,C\}$.

Such rank reduction (by ZF of certain path contributions) facilitates the satisfaction of the conditions in Theorem 4.

FIR IA for Asynchronous FIR Frequency-Selective IBC

FIR frequency-selective channels : OFDM : assumes that the same OFDM is used by synchronized BS. In HetNets, this may not be the case. Then FIR Tx/Rx filters may be considered. We get in the z-domain:

$$F_{i,k}(z)H_{i,k,j}(z)G_{j,n}(z) = 0\ ,\ (i,k) \neq (j,n)\ ,$$

If we denote by L_F, L_H, L_G the length of the 3 types of filters, then in a symmetric configuration, the proper conditions become

$$\begin{aligned} & KC\left[d(ML_G - d) + d(NL_F - d)\right] \geq \\ & \qquad KC(KC-1)d^2(L_H + L_G + L_F - 2) \\ & \Rightarrow d \leq \frac{ML_G + NL_F}{(KC-1)(L_H + L_G + L_F - 2) + 2} \leq \frac{\max\{M,N\}}{KC-1} \end{aligned}$$

where the last inequality can be attained by letting L_G or L_F tend to infinity. Unless $M \gg N$, this represents reduced DoF compared to the frequency-flat case ($d \leq (M+N)/(KC+1)$).

Alternatively, the double convolution by both Tx and Rx filters can be avoided by considering most of the decoupled approaches above, leading to more traditional equalization configurations, with equal DoF possibilities for frequency-selective as for frequency-flat cases.

5G Perspectives

- multi-user multi-cell interference management: theoretical possibilities, but (global) CSIT
  - FB delay $\Rightarrow$ channel prediction and channel Doppler models crucial
  - analog channel FB?
  - FDD: immediate channel FB
  - distributed : yes but watch for fast fading
- Massive MIMO simplifications: separating fast and slow fading channel components
- mmWave (beamforming, bandwidth), spectrum aggregation, full duplex radio
- beyond classical cellular:
  - HetNets (macro/small):
  - wireless/self backhauling

A Few References

- Y. Lejosne, D. Slock, and Y. Yuan-Wu, "Ergodic Interference Alignment for MIMO Interference Networks," in *Proc. ICASSP*, Florence, Italy, May 2014.
- T. Liu and C. Yang, "Genie Chain and Degrees of Freedom of Symmetric MIMO Interference Broadcast Channels," Sep. 2013, [Online]. Available: http://arxiv.org/abs/1309.6727.
- F. Negro, I. Ghauri, and D. T. M. Slock, "Deterministic Annealing Design and Analysis of the Noisy MIMO Interference Channel," in *Proc. IEEE Information Theory and Applications workshop (ITA)*, San Diego, CA, USA, Feb. 2011.
- F. Negro, M. Cardone, I. Ghauri, and D. T. M. Slock, "SINR Balancing and Beamforming for the MISO Interference Channel," in *Proc. IEEE PIMRC*, Toronto, Canada, Sept. 2011.
- F. Negro, D. Slock, and I. Ghauri, "On the Noisy MIMO Interference Channel with CSI through Analog Feedback," in *Int'l Symp. Communications, Control and Sig. Proc. (ISCCSP)*, Rome, Italy, May 2013.
- Y. Lejosne, D. Slock, and Y. Yuan-Wu, "Foresighted Delayed CSIT Feedback for Finite Rate of Innovation Channel Models and Attainable NetDoFs of the MIMO Interference Channel," in *Proc. Wireless Days*, Valencia, Spain, Nov. 2013.
- S. Jafar, "Topological Interference Management Through Index Coding," *IEEE Trans. Info. Theory*, Jan. 2014.
- L. Ruan, V. Lau, and M. Win, "The Feasibility Conditions for Interference Alignment in MIMO Networks," *IEEE Trans. Signal Processing*, Apr. 2013.
- M. Razaviyayn, L. Gennady, and Z. Luo, "On the Degrees of Freedom Achievable Through Interference Alignment in a MIMO Interference Channel," *IEEE Trans. Signal Processing*, Feb. 2012.
- X. Rao, L. Ruan, and V. Lau, "CSI Feedback Reduction for MIMO Interference Alignment," *IEEE Trans. Signal Processing*, Sep. 2013.
- P. de Kerret and D. Gesbert, "Interference Alignment with Incomplete CSIT Sharing," *IEEE Trans. Wireless Comm.*, 2013.
- Y. Lejosne, M. Bashar, D. Slock, and Y. Yuan-Wu, "Decoupled, Rank Reduced, Massive and Frequency-Selective Aspects in MIMO Interfering Broadcast Channels," in *Proc. ISCCSP*, Athens, Greece, May 2014.
- M. Bashar, Y. Lejosne, D. Slock, and Y. Yuan-Wu, "MIMO Broadcast Channels with Gaussian CSIT and Application to Location based CSIT," in *Proc. ITA*, San Diego, CA, USA, Feb. 2014.