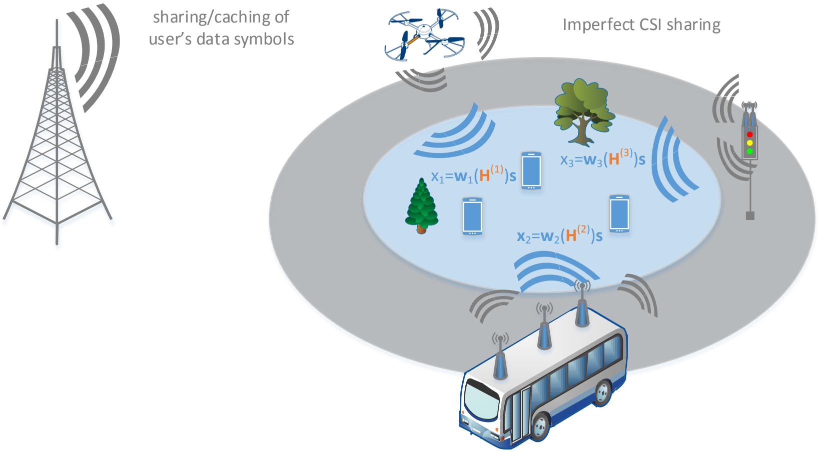
Degrees-of-Freedom Robust Transmission for the K-user Distributed Broadcast Channel

Presented by Paul de Kerret
Joint work with Antonio Bazco, Nicolas Gresset, and David Gesbert

ESIT 2017 in Madrid,
10/05/2017

Our Focus: Decentralized Broadcast Channel with Imperfect CSIT Sharing



Broadcast Channel (JP-CoMP, Network MIMO)

- Some simplifying asumptions:
  (i) K single-antenna TXs and K single-antennas RXs
  (ii) Perfect CSI at the RX
  (iii) Gaussian data symbols
  (iv) Block fading channel
- A key asumption: User's data symbols are available at all TXs

Broadcast Channel (JP-CoMP, Network MIMO)

- Some simplifying asumptions:
  (i) K single-antenna TXs and K single-antennas RXs
  (ii) Perfect CSI at the RX
  (iii) Gaussian data symbols
  (iv) Block fading channel
- A key asumption: User's data symbols are available at all TXs
- Received signal at user i

$$\overbrace{y_i}^{\text{Received signal at user i}} = \boldsymbol{h}_i^{\mathrm{H}}\mathbf{x} + \eta_i$$

Broadcast Channel (JP-CoMP, Network MIMO)

- Some simplifying asumptions:
  (i) K single-antenna TXs and K single-antennas RXs
  (ii) Perfect CSI at the RX
  (iii) Gaussian data symbols
  (iv) Block fading channel
- A key asumption: User's data symbols are available at all TXs
- Received signal at user i

$$y_i = \overbrace{\boldsymbol{h}_i^{\mathrm{H}}}^{\text{Channel from all TXs to user } i \ (1 \times K)} \mathbf{x} + \eta_i$$

Broadcast Channel (JP-CoMP, Network MIMO)

- Some simplifying asumptions:
  (i) K single-antenna TXs and K single-antennas RXs
  (ii) Perfect CSI at the RX
  (iii) Gaussian data symbols
  (iv) Block fading channel
- A key asumption: User's data symbols are available at all TXs
- Received signal at user i

$$y_i = \boldsymbol{h}_i^{\mathrm{H}} \overbrace{\mathbf{x}}^{\text{Multi-user multi-TX transmit signal } (K \times 1)} + \eta_i$$

with $\{\mathbf{x}\}_j$ sent from TX j

Broadcast Channel (JP-CoMP, Network MIMO)

- Some simplifying asumptions:
  (i) K single-antenna TXs and K single-antennas RXs
  (ii) Perfect CSI at the RX
  (iii) Gaussian data symbols
  (iv) Block fading channel
- A key asumption: User's data symbols are available at all TXs
- Received signal at user i

$$y_i = \boldsymbol{h}_i^{\mathrm{H}} \mathbf{x} + \overbrace{\eta_i}^{\text{Additive white Gaussian Noise } \mathcal{N}_{\mathbb{C}}(0,1)}$$

Broadcast Channel (JP-CoMP, Network MIMO)

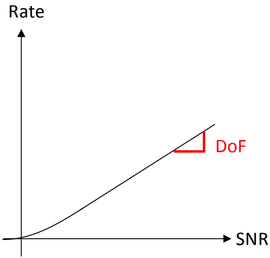
- Some simplifying asumptions:
  (i) K single-antenna TXs and K single-antennas RXs
  (ii) Perfect CSI at the RX
  (iii) Gaussian data symbols
  (iv) Block fading channel
- A key asumption: User's data symbols are available at all TXs
- Received signal at user i

$$y_i = \boldsymbol{h}_i^{\mathrm{H}} \mathbf{x} + \overbrace{\eta_i}^{\text{Additive white Gaussian Noise } \mathcal{N}_{\mathbb{C}}(0,1)}$$

- For a given transmit power P, let $\mathcal{C}(P)$ denote the sum capacity
- Our figure of merit will be the Degrees-of-Freedom:

$$\text{DoF} \triangleq \lim_{P \to \infty} \frac{\mathcal{C}(P)}{\log_2(P)}$$

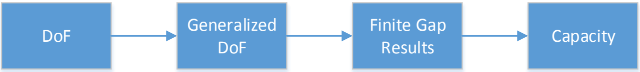
Is DoF Useful?



- First order approximation in the SNR

$$R^{\star} = \text{DoF} \log_2(SNR) + o(\log_2(SNR))$$

+ Closed form results
+ New insights and new paradigms
+ First step towards capacity



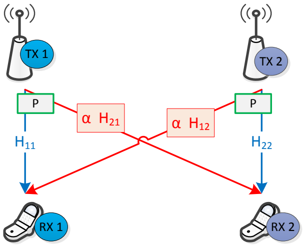
- Inaccurate if strong pathloss differences
- Results not always relevant at finite SNR

Very successful to discover new approaches/insights (MIMO [Telatar, 1999, ETC], IA [Cadambe and Jafar, 2008, TIT], delayed CSIT [Maddah-Ali and Tse, 2012, TIT],...)

DoF and Pathloss – A short Parenthesis (1) –

Example

- 2-user IC, single-antenna nodes, $\alpha^2 = 10^{-12}$, $H_{i,j} \sim \mathcal{N}_{\mathbb{C}}(0,1)$



- DoF analysis: DoF $= 1$ [Etkin et al., 2008, TIT]

➥ Not the expected behaviour

Generalized DoF – A short Parenthesis (2) –

- With Generalized DoF, model the pathloss difference

$$\mathbb{E}[|H_{i,j}|^2] \doteq P^{-\gamma_{i,j}}$$

- Generalized DoF (GDoF) then defined as

$$\text{DoF}\left(\{\gamma_{i,j}\}_{i,j}\right) \triangleq \lim_{P\to\infty} \frac{C(P, \{\gamma_{i,j}\}_{i,j})}{\log_2(P)}$$

- Example continued: For $P = 20dB$,

$$\gamma_{1,2} = \gamma_{2,1} = 6$$

and

$$\text{DoF}\left(\{\gamma_{i,j}\}_{i,j}\right) = 2$$

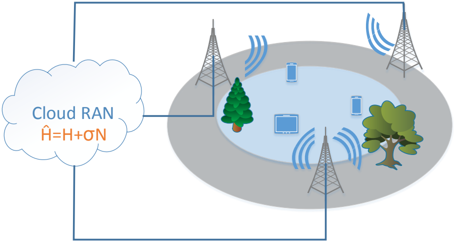
➥ Expected behaviour!

Remark

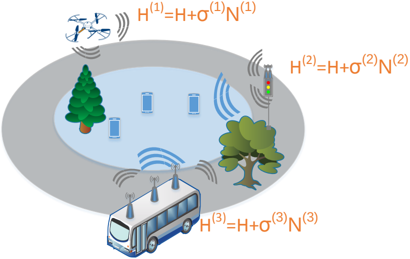
GDoF not discussed here but extension for the 2 users case in [Bazco et al., 2017, ISIT]

Centralized VS Distributed CSI

- **Centralized –TX Independent–: Conventional model**



- **Distributed –TX Dependent–: Our focus here**



Outline

DoF with Perfect CSIT

- All TXs have perfect knowledge of $\mathbf{H}$: Optimal DoF is $\text{DoF}^{\text{PCSI}} = K$
- DoF-optimal transmission scheme is Zero Forcing:

$$\mathbf{x} = \sqrt{P}\frac{\mathbf{H}^{-1}}{\|\mathbf{H}^{-1}\|_{\mathrm{F}}}\begin{bmatrix} s_1 \\ \vdots \\ s_K \end{bmatrix}$$

- Received signal is

$$y_i = \frac{\sqrt{P}}{\|\mathbf{H}^{-1}\|_{\mathrm{F}}} s_i + \eta_i$$

SNR scales in P: asymptotically possible to decode s_i with the rate $\log_2(P)$ bits

Remark

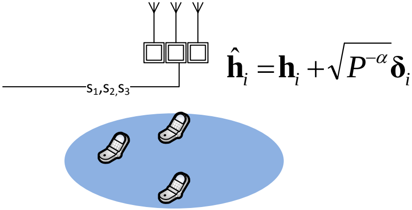
Importantly, $\mathbf{x}$ can also be chosen as $\mathbf{x} = \sum_{i=1}^{K} \sqrt{\frac{P}{K}} \frac{\boldsymbol{t}_i}{\|\boldsymbol{t}_i\|} s_i$ where the beamformer/precoder $\boldsymbol{t}_i \in \mathbb{C}^{K\times 1}$ is

$$\boldsymbol{t}_i = \Pi^{\perp}_{\boldsymbol{h}_1,,\boldsymbol{h}_{i-1},\boldsymbol{h}_{i+1},\dots,\boldsymbol{h}_K} \boldsymbol{h}_i$$

Outline

Imperfect CSIT in the Centralized Case

- Conventional high SNR parameterization



- $\alpha \in [0, 1]$ is called the CSIT quality exponent. Intuitively, equal to the ratio between the "available CSIT" over the "needed CSIT"
  - $\alpha = 0 \approx$ no CSIT
  - $\alpha = 1 \approx$ perfect CSIT
- Some practical motivation:
  - Quantization noise with VQ for $B >> 1$, with $B = \#$ quantization bits,

$$\sigma^2 \approx 2^{-\frac{B}{M-1}}$$

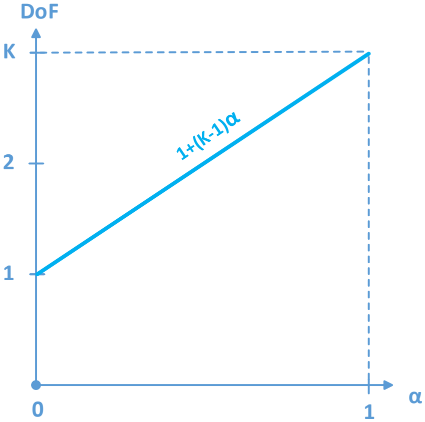
  - If $B = \alpha(M-1)\log_2(P), \alpha \in [0, 1]$

$$\sigma^2 \approx P^{-\alpha}$$

DoF Analysis of the Centralized Configuration

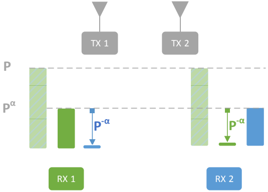
$$\text{DoF}^{\text{CCSI}}(\alpha) = 1 + (K-1)\alpha$$

- Outerbound recently proven in [Davoodi and Jafar, 2016, TIT]
- Achievable scheme in [Jindal, 2006, TIT][Hao et al., 2015, TCOM]



DoF-Optimal Scheme for the Centralized Case (1) [Jindal, 2006, TIT][Hao et al., 2015, TCOM]

- DoF-optimal scheme: Zero-Forcing (ZF) + Rate Splitting (RS)



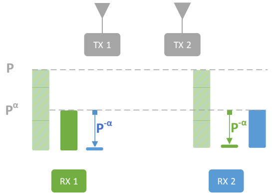
$$y_1 = \underbrace{\sqrt{P}\boldsymbol{h}_1^{\mathrm{H}} \begin{bmatrix}1\\0\end{bmatrix} s_0}_{\text{Common symbol} \doteq P} + \underbrace{\sqrt{P^{\alpha}}\boldsymbol{h}_1^{\mathrm{H}}\boldsymbol{t}_1^{\mathrm{ZF}} s_1}_{\text{private symbol} \doteq P^{\alpha}} + \underbrace{\sqrt{P^{\alpha}}\boldsymbol{h}_1^{\mathrm{H}}\boldsymbol{t}_2^{\mathrm{ZF}} s_2}_{\text{interference} \doteq P^{0}}$$

with $\boldsymbol{t}_i^{\mathrm{ZF}} = \frac{\Pi^{\perp}_{\boldsymbol{h}_{\bar{i}}}\boldsymbol{h}_i}{\|\Pi^{\perp}_{\boldsymbol{h}_{\bar{i}}}\boldsymbol{h}_i\|}$ and with

$$|\boldsymbol{h}_1^{\mathrm{H}}\boldsymbol{t}_2^{\mathrm{ZF}}|^2 = |\underbrace{\hat{\boldsymbol{h}}_1^{\mathrm{H}}\boldsymbol{t}_2^{\mathrm{ZF}}}_{0} + \sqrt{P^{-\alpha}}\delta_1^{\mathrm{H}}\boldsymbol{t}_2^{\mathrm{ZF}}|^2$$

$$= P^{-\alpha}|\delta_1^{\mathrm{H}}\boldsymbol{t}_2^{\mathrm{ZF}}|^2$$

DoF-Optimal Scheme for the Centralized Case (2) [Jindal, 2006, TIT][Hao et al., 2015, TCOM]



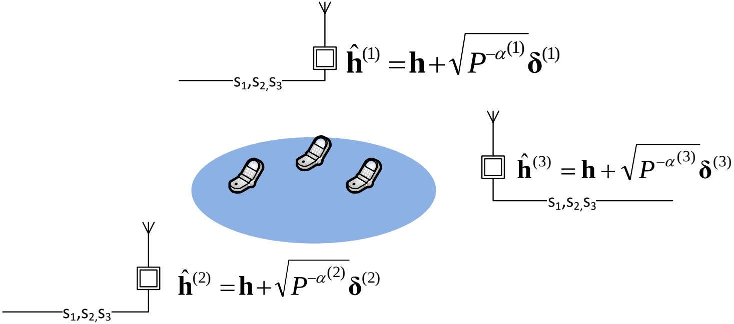
$$y_1 = \underbrace{\sqrt{P}\boldsymbol{h}_1^{\mathrm{H}}\begin{bmatrix}1\\0\end{bmatrix} s_0}_{\text{Common symbol } \doteq P} + \underbrace{\sqrt{P^\alpha}\boldsymbol{h}_1^{\mathrm{H}}\boldsymbol{t}_1^{\mathrm{ZF}} s_1}_{\text{private symbol } \doteq P^\alpha} + \underbrace{\sqrt{P^\alpha}\boldsymbol{h}_1^{\mathrm{H}}\boldsymbol{t}_2^{\mathrm{ZF}} s_2}_{\text{interference } \doteq P^0}$$

- Successive Decoding
  - Decode first s_0 with rate $(1-\alpha)\log_2(P)$ bits (SNR $\doteq P^{1-\alpha}$)
  - Decode then s_1 with rate $\alpha\log_2(P)$ bits (SNR $\doteq P^{\alpha}$)
- Sum DoF is $(1-\alpha)+K\alpha$

Outline

Distributed CSIT Configuration

- With imperfect CSIT sharing extends to



- CSIT configuration characterized by

$$1 \geq \alpha^{(1)} \geq \alpha^{(2)} \geq \ldots \geq \alpha^{(K)} \geq 0$$

Remark

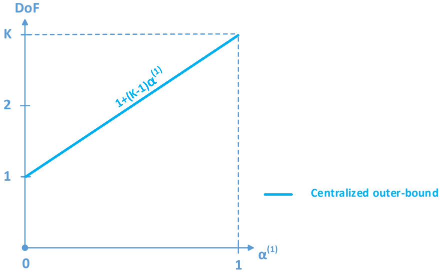
Arbitrary CSIT configuration

An Intuitive Outerbound [de Kerret and Gesbert, 2016, ISIT]

Theorem (The Centralized Outerbound)

$$\mathrm{DoF}^{\mathrm{DCSI}}(\boldsymbol{\alpha}) \leq 1 + (K-1) \underbrace{\max_{j \in \{1,\dots,K\}} \alpha^{(j)}}_{=\alpha^{(1)}}$$

- DoF upperbounded by DoF achieved by full CSIT exchange
- Having $\hat{\mathbf{H}}^{\alpha^{(1)}}, \dots, \dots, \hat{\mathbf{H}}^{\alpha^{(K)}}$ doesn't help over having just best CSIT $\hat{\mathbf{H}}^{\alpha^{(1)}}$



Conventional Zero Forcing [de Kerret and Gesbert, 2012, TIT]

- First idea: Use ZF (DoF-optimal for Centralized CSIT)

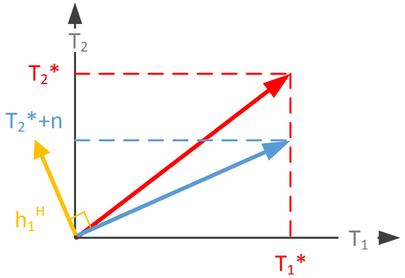
$$\mathrm{DoF}^{\mathrm{ZF}} = 1 + (K-1) \underbrace{\min_{j \in \{1,\dots,K\}} \alpha^{(j)}}_{=\alpha^{(K)}}$$

➥ Very inefficient!

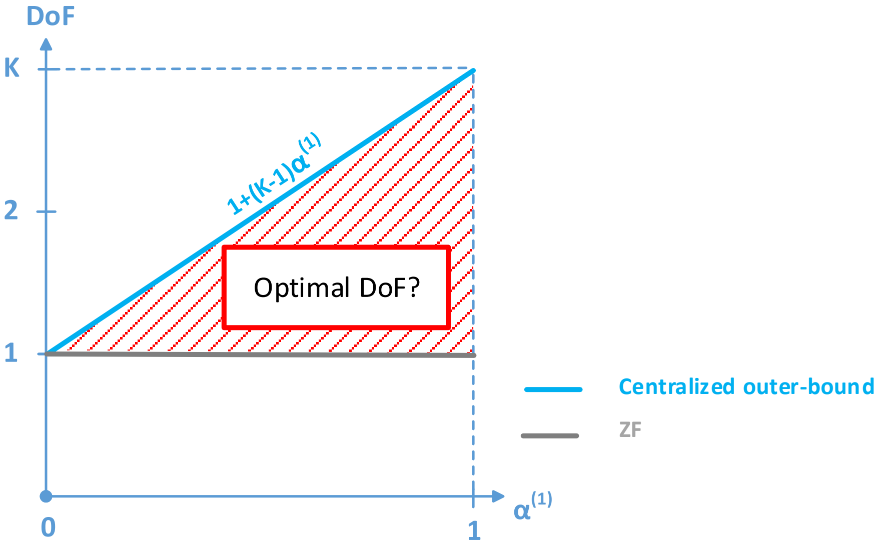
- Why? Goal is to design T_1 and T_2 such that

$$\boldsymbol{h}_1^{\mathrm{H}} \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} \approx 0, \qquad \text{(Zero Forcing constraint at RX 1)}$$

i.e., find a vector orthogonal to $\boldsymbol{h}_1^{\mathrm{H}}$



Problem Statement

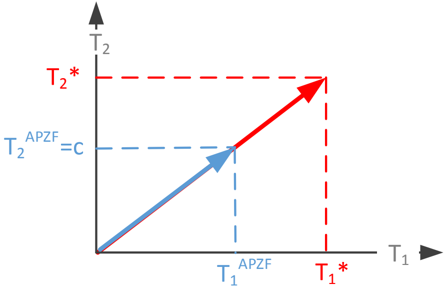


Outline

Active-Passive Zero-Forcing (AP-ZF) [de Kerret and Gesbert, 2012, TIT]

- Main Idea: Less informed TX generates interference, more informed TX removes it

$$\{(\boldsymbol{h}_1^{(1)})^{\mathrm{H}}\}_1 T_1 + \{(\boldsymbol{h}_1^{(1)})^{\mathrm{H}}\}_2 T_2 = 0 \rightarrow T_1 = -\frac{\{(\boldsymbol{h}_1^{(1)})^{\mathrm{H}}\}_2}{\{(\boldsymbol{h}_1^{(1)})^{\mathrm{H}}\}_1} T_2$$



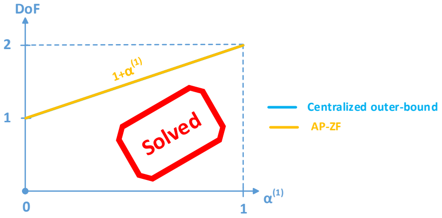
- Achieves the DoF

$$\mathrm{DoF}^{\mathrm{APZF}} = 1 + \alpha^{(1)}$$

Active-Passive Zero-Forcing (AP-ZF)

- Achieves the DoF

$$\mathrm{DoF}^{\mathrm{APZF}} = 1 + \alpha^{(1)}$$



Remark

In fact achieves also Generalized DoF [Bazco et al., 2017]

Outline

Generalization of AP-ZF?

- Problem: AP-ZF doesn't help much with more users

$$\text{DoF}^{\mathrm{APZF}} = 1 + (K-1)\alpha^{(K-1)}$$

➥ Need for a different approach

Main Idea

Exploit interference as side information: Interference useful for both the **interfered user** and the **desired user**

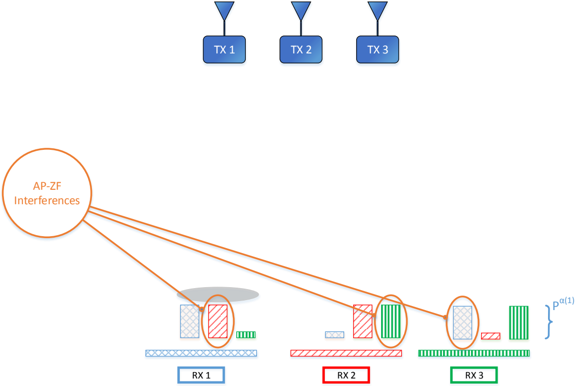
➥Analogy to the use of delayed CSIT [Maddah-Ali and Tse, 2012, TIT]

A Multi-layer Transmission Scheme [de Kerret and Gesbert, 2016, ISIT]

1. All TXs serve all users with power $P^{\alpha^{(1)}}$ using Active-Passive ZF

$$\mathbf{x} = \sqrt{P^{\alpha^{(1)}}} \sum_{i=1}^{K} \mathbf{T}_i^{\mathrm{APZF}} \mathbf{s}_i$$

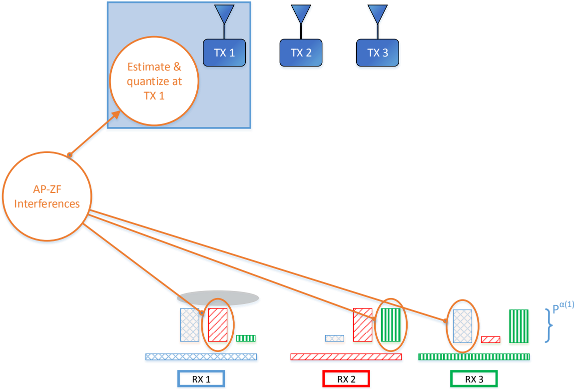
➥Generate interferences of power $P^{\alpha^{(1)}}$



A Multi-layer Transmission Scheme [de Kerret and Gesbert, 2016, ISIT]

2. TX 1 estimates and quantizes the interference terms before their generations

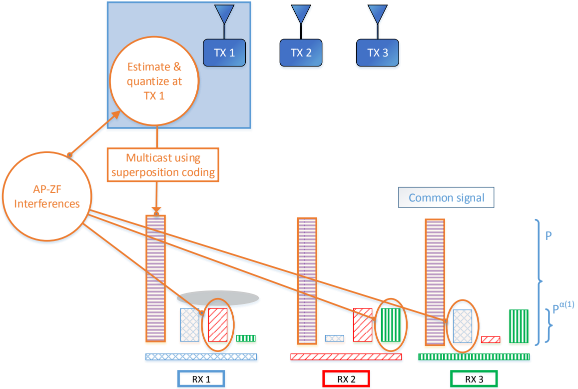
$$\mathbf{x} = \sqrt{P^{\alpha^{(1)}}} \sum_{i=1}^{K} \mathbf{T}_i^{\mathrm{APZF}} s_i$$



A Multi-layer Transmission Scheme [de Kerret and Gesbert, 2016, ISIT]

3. TX 1 then transmits them via a common data symbol at the same time as the private data symbols

$$\mathbf{x} = \sqrt{P}\begin{bmatrix} 1 \\ \mathbf{0}_{K-1} \end{bmatrix} s_0 + \sqrt{P^{\alpha^{(1)}}} \sum_{i=1}^{K} \mathbf{T}_i^{\mathrm{APZF}} \mathbf{s}_i$$



Signal Processing at TX 1

1. Interference estimation at TX 1:

$$\sqrt{P^{\alpha^{(1)}}}(\hat{\boldsymbol{h}}_1^{(1)})^{\mathrm{H}}\mathbf{T}_2^{\mathrm{APZF}}\boldsymbol{s}_2 = \sqrt{P^{\alpha^{(1)}}}(\hat{\boldsymbol{h}}_1^{(1)} + \sqrt{P^{-\alpha^{(1)}}}\boldsymbol{\delta}_1^{(1)})^{\mathrm{H}}\mathbf{T}_2^{\mathrm{APZF}}\boldsymbol{s}_2$$
$$= \sqrt{P^{\alpha^{(1)}}}\boldsymbol{h}_1^{\mathrm{H}}\mathbf{T}_2^{\mathrm{APZF}}\boldsymbol{s}_2 + \underbrace{\sqrt{P^{-\alpha^{(1)}}}(\boldsymbol{\delta}_1^{(1)})^{\mathrm{H}}\mathbf{T}_2^{\mathrm{APZF}}\boldsymbol{s}_2}_{O(1)}$$

➥TX 1 can compute DoF-perfect estimate of the interference terms!

2. Interference quantization: Use $\alpha^{(1)}\log_2(P)$ bits to quantize the signal scaling in $P^{\alpha^{(1)}}$
   ➥Quantization error scaling in P^0 [Cover and Thomas, 2006]
3. Transmit $3\alpha^{(1)}\log_2(P)$ bits to all users

Signal Processing at RX 1 (w.l.o.g.)

- User 1 has received

$$y_1 = \underbrace{\sqrt{P}\boldsymbol{h}_1^{\mathrm{H}} \begin{bmatrix} 1 \\ \mathbf{0}_{K-1} \end{bmatrix} s_0}_{\doteq P} + \underbrace{\sqrt{P^{\alpha^{(1)}}}\boldsymbol{h}_1^{\mathrm{H}}\mathbf{T}_1^{\mathrm{APZF}}\boldsymbol{s}_1}_{\doteq P^{\alpha^{(1)}}} + \underbrace{\sqrt{P^{\alpha^{(1)}}}\boldsymbol{h}_1^{\mathrm{H}}\mathbf{T}_2^{\mathrm{APZF}}\boldsymbol{s}_2}_{\doteq P^{\alpha^{(1)}}} + \underbrace{\sqrt{P^{\alpha^{(1)}}}\boldsymbol{h}_1^{\mathrm{H}}\mathbf{T}_3^{\mathrm{APZF}}\boldsymbol{s}_3}_{\doteq P^0}$$

- User 1 decodes s_0 and obtains then

$$\sqrt{P^{\alpha^{(1)}}}(\hat{\boldsymbol{h}}_1^{(1)})^{\mathrm{H}}\mathbf{T}_2^{\mathrm{APZF}}\boldsymbol{s}_2, \qquad \text{Useful: Remove interference}$$

$$\sqrt{P^{\alpha^{(1)}}}(\hat{\boldsymbol{h}}_2^{(1)})^{\mathrm{H}}\mathbf{T}_3^{\mathrm{APZF}}\boldsymbol{s}_3, \qquad \text{Useless (for RX 1)}$$

$$\sqrt{P^{\alpha^{(1)}}}(\hat{\boldsymbol{h}}_3^{(1)})^{\mathrm{H}}\mathbf{T}_1^{\mathrm{APZF}}\boldsymbol{s}_1, \qquad \text{Useful: Desired data}$$

- Achieved DoF: If $3\alpha^{(1)} \leq 1 - \alpha^{(1)}$, achieves DoF

$$\mathrm{DoF} = \underbrace{6\alpha^{(1)}}_{2\alpha^{(1)} \text{ per user}} + \underbrace{(1-\alpha^{(1)}) - 3\alpha^{(1)}}_{\text{multicast DoF after retransmitting interference}}$$

Weak CSIT Regime

Theorem

If $\max_{j\in\{1,\ldots,K\}} \alpha^{(j)} \leq \frac{1}{1+K(K-2)}$ *(weak CSIT regime)*,

$$\mathrm{DoF}^{\mathrm{DCSI}}(\boldsymbol{\alpha}) \geq 1 + (K-1) \max_{j\in\{1,\ldots,K\}} \alpha^{(j)}$$

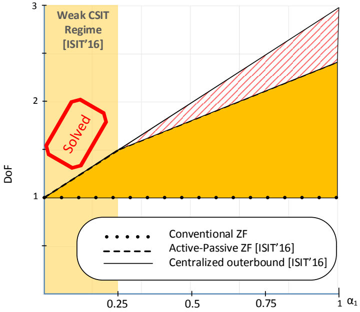
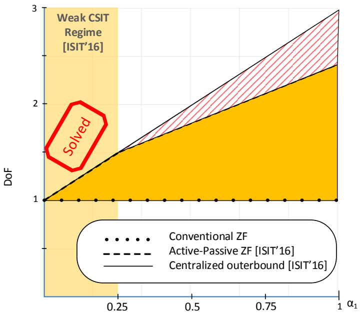


Figure: DoF as a function of $\alpha^{(1)}$ for $\alpha^{(2)} = \frac{2}{3}\alpha^{(1)}$ and $\alpha^{(3)} = 0$

Take Home Message

- DoF analysis allows to develop new schemes/insights with simple linear algebra
- Role of each TX adapts to the full multi-TX CSIT configuration
- Multi-layer transmission scheme: Estimate, Qantize & transmit interference at the most informed user
- Many extensions:
  - Developed a new Hierarchical Zero-Forcing to extend optimality region
  - Extend further?
  - Improving the centralized-outerbound
  - Going beyond the DoF



References I

Antonio. Bazco, P. de Kerret, D. Gesbert, and N. Gresset. Generalized Degrees-of-Freedom of the 2-User Case MISO Broadcast Channel with Distributed CSIT. to be published IEEE International Symposium on Information Theory (ISIT), 2017.

V. R. Cadambe and S. A. Jafar. Interference alignment and degrees of freedom of the K-user interference channel. *IEEE Trans. Inf. Theory*, 54(8): 3425–3441, Aug. 2008.

T. Cover and A. Thomas. *Elements of information theory*. Wiley-Interscience, Jul. 2006.

A. Gholami Davoodi and S. A. Jafar. Aligned image sets under channel uncertainty: Settling conjectures on the collapse of Degrees of Freedom under finite precision CSIT. *IEEE Trans. Inf. Theo.*, 62(10):5603–5618, Oct. 2016.

P. de Kerret and D. Gesbert. Degrees of freedom of the network MIMO channel with distributed CSI. *IEEE Trans. Inf. Theory*, 58(11):6806–6824, Nov. 2012.

P. de Kerret and D. Gesbert. Network MIMO: Transmitters with no CSI Can Still be Very Useful. In *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2016.

P. de Kerret, Bazco A., and D. Gesbert. Enforcing Coordination in Network MIMO with Unequal CSIT. In *Proc. IEEE Asilomar Conference on Signals, Systems and Computers (ACSSC)*, 2016a.

P. de Kerret, D. Gesbert, J. Zhang, and P. Elia. Optimally bridging the gap from delayed to perfect CSIT in the K-user MISO BC. In *Proc. IEEE Information Theory Workshop (ITW)*, 2016b.

R.H. Etkin, D.N.C. Tse, and H. Wang. Gaussian interference channel capacity to within one bit. *IEEE Trans. Inf. Theory*, 54(12):5534–5562, Dec. 2008.

C. Hao, Y. Wu, and B. Clerckx. Rate analysis of two-receiver MISO Broadcast Channel with finite rate feedback: A rate-splitting approach. *IEEE Trans. on Commun.*, 63(9):3232–3246, Sept. 2015.

N. Jindal. MIMO Broadcast Channels with finite-rate feedback. *IEEE Trans. Inf. Theory*, 52(11):5045–5060, Nov. 2006.

M.A. Maddah-Ali and D. Tse. Completely stale transmitter channel state information is still very useful. *IEEE Trans. Inf. Theory*, 58(7):4418–4431, Jul. 2012.

I. Emre Telatar. Capacity of multi-antenna Gaussian channels. *European Transaction on Communications*, 10:585–595, 1999.


thanks

Extension of the Weak CSIT Regime for $K = 3$ [de Kerret et al., 2016a, Asilomar]

- Improved scheme building on a new Hierarchical ZF precoding paradigm

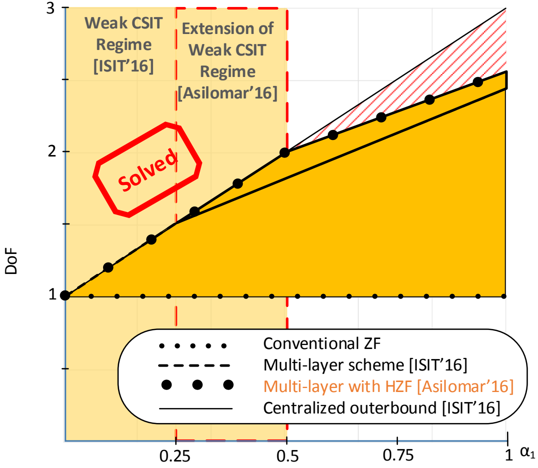


Figure: DoF as a function of $\alpha^{(1)}$ for $\alpha^{(2)} = \frac{2}{3}\alpha^{(1)}$ and $\alpha^{(3)} = 0$

Beyond the Weak CSIT Regime

Definition (Weak CSIT regime)

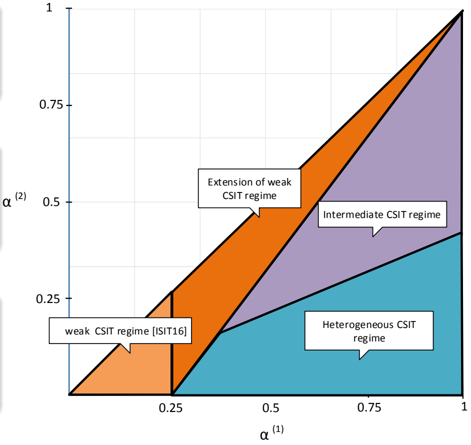
$$\alpha^{(1)} \leq \frac{1}{4} + \frac{3}{4}\alpha^{(2)}$$

Definition (Heterogeneous CSIT regime)

$$\alpha^{(1)} > \min\left(2\alpha^{(2)}, \frac{1}{4} + \frac{3}{4}\alpha^{(2)}\right)$$

Definition (Intermediate CSIT regime)

$$\frac{1}{4} + \frac{3}{4}\alpha^{(2)} < \alpha^{(1)} \leq 2\alpha^{(2)}$$



Achievable DoF [de Kerret et al., 2016a, Asilomar]

Theorem

In the 3-user MIMO BC with D-CSIT, it holds that

$$\mathrm{DoF}^{\mathrm{DCSI}}(\boldsymbol{\alpha}) \geq \begin{cases} 1 + 2\alpha^{(1)} & \textit{(Weak CSIT)} \\ \frac{3}{2}(1 + \alpha^{(2)}) & \textit{(Intermediate CSIT)} \\ 1+\alpha^{(1)} + \frac{3\alpha^{(1)}(1-\alpha^{(1)})+\alpha^{(2)}\left(5\alpha^{(1)}-3\alpha^{(2)}-1\right)}{9\alpha^{(1)}-8\alpha^{(2)}} & \textit{(Heterogeneous CSIT)} \end{cases}$$

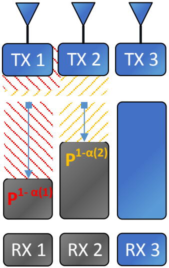
Can be achieved building on new precoding scheme: Hierarchical zero-forcing

Hierarchical ZF with $K = 3$: Main Property

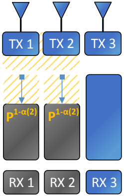
Lemma

Let $\boldsymbol{t}_3^{\mathrm{HZF}}$ be the HZF beamformer towards user 3 with average power P. Then:

$$|\boldsymbol{h}_1^{\mathrm{H}} \boldsymbol{t}_3^{\mathrm{HZF}}|^2 \dot{\leq} P^{1-\alpha^{(1)}}$$
$$|\boldsymbol{h}_2^{\mathrm{H}} \boldsymbol{t}_3^{\mathrm{HZF}}|^2 \dot{\leq} P^{1-\alpha^{(2)}}$$



compared with conventional ZF



Roadmap of Hierarchical ZF

1. Make CSIT hierarchical
2. Split precoding in layers
3. Design layer k to reduce interference at user k without reducing interference reduction already realized

(1) Make CSIT Hierarchical

Example

- Example for two transmitters TX1, TX2 with $\alpha^{(1)} \geq \alpha^{(2)}$
- Let $Q_{\alpha^{(2)}}$ be our Hierarchical quantizer using $\alpha^{(2)} \log_2(P)$ bits
- Let us define

$$\hat{\hat{\mathbf{H}}}^{(1)}_{\alpha^{(2)}} \triangleq Q_{\alpha^{(2)}}\left(\hat{\mathbf{H}}^{(1)}\right)$$
$$\hat{\hat{\mathbf{H}}}^{(2)}_{\alpha^{(2)}} \triangleq Q_{\alpha^{(2)}}\left(\hat{\mathbf{H}}^{(2)}\right)$$

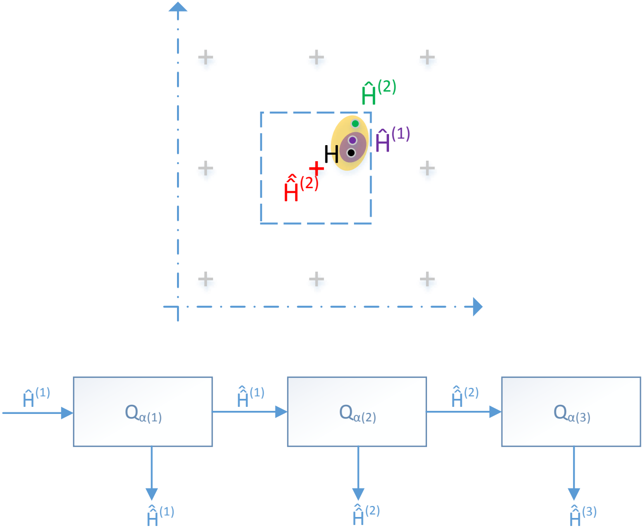
- Then, there exists a quantizer $Q_{\alpha^{(2)}}$ such that [de Kerret et al., 2016b, ITW] :

$$\lim_{P\to\infty} \Pr\left\{\hat{\hat{\mathbf{H}}}^{(1)}_{\alpha^{(2)}} = \hat{\hat{\mathbf{H}}}^{(2)}_{\alpha^{(2)}}\right\} = 1$$
$$\mathbb{E}\left[\|\hat{\hat{\mathbf{H}}}^{(j)}_{\alpha^{(2)}} - \mathbf{H}\|^2_{\mathrm{F}}\right] \dot{\leq} P^{-\alpha^{(2)}}, \qquad j = 1,2$$

➥ TX 1 can obtain $\hat{\hat{\mathbf{H}}}^{(2)}_{\alpha^{(2)}}$: CSIT configuration has been made hierarchical

➥ More generally, TX i knows what TX $i+1$ knows (post quantizing)

(1) Make CSIT Hierarchical



(2) Split Precoding in Layers

- $\boldsymbol{t}_3^{\mathrm{HZF}}$ aimed at user 3 decomposed as

$$\boldsymbol{t}_3^{\mathrm{HZF}} = \begin{bmatrix} t_3^{\mathrm{HZF}}(1) \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \{\boldsymbol{t}_3^{\mathrm{HZF}}(2)\}_1 \\ \{\boldsymbol{t}_3^{\mathrm{HZF}}(2)\}_2 \\ 0 \end{bmatrix} + \begin{bmatrix} \{\boldsymbol{t}_3^{\mathrm{HZF}}(3)\}_1 \\ \{\boldsymbol{t}_3^{\mathrm{HZF}}(3)\}_2 \\ \{\boldsymbol{t}_3^{\mathrm{HZF}}(3)\}_3 \end{bmatrix}$$

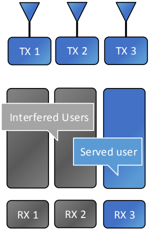
- e.g., TX 2 needs to be able to compute the 2th row:

$$\boldsymbol{t}_3^{\mathrm{HZF}} = \begin{bmatrix} t_3^{\mathrm{HZF}}(1) \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \{\boldsymbol{t}_3^{\mathrm{HZF}}(2)\}_1 \\ \{\boldsymbol{t}_3^{\mathrm{HZF}}(2)\}_2 \\ 0 \end{bmatrix} + \begin{bmatrix} \{\boldsymbol{t}_3^{\mathrm{HZF}}(3)\}_1 \\ \{\boldsymbol{t}_3^{\mathrm{HZF}}(3)\}_2 \\ \{\boldsymbol{t}_3^{\mathrm{HZF}}(3)\}_3 \end{bmatrix}$$

(3) Hierarchical ZF for $K = 3$

- First "layer" (at TX 1, TX 2 and TX 3)

$$\boldsymbol{t}_3^{\mathrm{HZF}}(3) = \lambda^{\mathrm{HZF}} \hat{\hat{\mathbf{H}}}^{(3)\mathrm{H}} \left(\hat{\hat{\mathbf{H}}}^{(3)} (\hat{\hat{\mathbf{H}}}^{(3)})^{\mathrm{H}} + \frac{1}{P} \mathbf{I}_3 \right)^{-1} \boldsymbol{e}_3$$



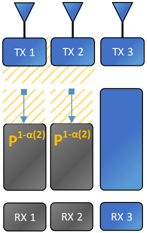
(3) Hierarchical ZF for $K = 3$

- First "layer" (at TX 1, TX 2 and TX 3)

$$\boldsymbol{t}_3^{\mathrm{HZF}}(3) = \lambda^{\mathrm{HZF}} \hat{\hat{\mathbf{H}}}^{(3)\mathrm{H}} \left(\hat{\hat{\mathbf{H}}}^{(3)} (\hat{\hat{\mathbf{H}}}^{(3)})^{\mathrm{H}} + \frac{1}{P} \mathbf{I}_3 \right)^{-1} \boldsymbol{e}_3$$

- Second "layer" (at TX 1 and TX 2)

$$\boldsymbol{t}_3^{\mathrm{HZF}}(2) = -\hat{\hat{\mathbf{H}}}^{(2)\mathrm{H}}_{[1:2,1:2]} \left(\hat{\hat{\mathbf{H}}}^{(2)}_{[1:2,1:2]} \hat{\hat{\mathbf{H}}}^{(2)\mathrm{H}}_{[1:2,1:2]} + \frac{1}{P} \mathbf{I}_2 \right)^{-1} \hat{\hat{\mathbf{H}}}^{(2)}_{[1:2,1:3]} \boldsymbol{t}_3^{\mathrm{HZF}}(3)$$



(3) Hierarchical ZF for $K = 3$

- First "layer" (at TX 1, TX 2 and TX 3)

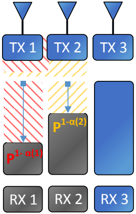
$$\boldsymbol{t}_3^{\mathrm{HZF}}(3) = \lambda^{\mathrm{HZF}} \hat{\hat{\mathbf{H}}}^{(3)\mathrm{H}} \left(\hat{\hat{\mathbf{H}}}^{(3)} (\hat{\hat{\mathbf{H}}}^{(3)})^{\mathrm{H}} + \frac{1}{P}\mathbf{I}_3 \right)^{-1} \boldsymbol{e}_3$$

- Second "layer" (at TX 1 and TX 2)

$$\boldsymbol{t}_3^{\mathrm{HZF}}(2) = -\hat{\hat{\mathbf{H}}}^{(2)\mathrm{H}}_{[1:2,1:2]} \left(\hat{\hat{\mathbf{H}}}^{(2)}_{[1:2,1:2]} \hat{\hat{\mathbf{H}}}^{(2)\mathrm{H}}_{[1:2,1:2]} + \frac{1}{P}\mathbf{I}_2 \right)^{-1} \hat{\hat{\mathbf{H}}}^{(2)}_{[1:2,1:3]} \boldsymbol{t}_3^{\mathrm{HZF}}(3)$$

- Third "layer" (at TX 1)

$$\boldsymbol{t}_3^{\mathrm{HZF}}(1) = -\hat{\hat{H}}^{(1)\mathrm{H}}_{1,1} \left(|\hat{\hat{H}}^{(1)}_{1,1}|^2 + \frac{1}{P} \right)^{-1} \hat{\hat{\boldsymbol{h}}}^{(1)\mathrm{H}}_1 \left(\begin{bmatrix} \boldsymbol{t}_3^{\mathrm{HZF}}(2) \\ 0 \end{bmatrix} + \boldsymbol{t}_3^{\mathrm{HZF}}(3) \right)$$



(3) Hierarchical ZF for $K = 3$

- First "layer" (at TX 1, TX 2 and TX 3)

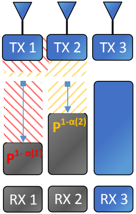
$$\boldsymbol{t}_3^{\mathrm{HZF}}(3) = \lambda^{\mathrm{HZF}} \hat{\hat{\mathbf{H}}}^{(3)\mathrm{H}} \left(\hat{\hat{\mathbf{H}}}^{(3)} (\hat{\hat{\mathbf{H}}}^{(3)})^{\mathrm{H}} + \frac{1}{P} \mathbf{I}_3 \right)^{-1} \boldsymbol{e}_3$$

- Second "layer" (at TX 1 and TX 2)

$$\boldsymbol{t}_3^{\mathrm{HZF}}(2) = -\hat{\hat{\mathbf{H}}}^{(2)\mathrm{H}}_{[1:2,1:2]} \left(\hat{\hat{\mathbf{H}}}^{(2)}_{[1:2,1:2]} \hat{\hat{\mathbf{H}}}^{(2)\mathrm{H}}_{[1:2,1:2]} + \frac{1}{P} \mathbf{I}_2 \right)^{-1} \hat{\hat{\mathbf{H}}}^{(2)}_{[1:2,1:3]} \boldsymbol{t}_3^{\mathrm{HZF}}(3)$$

- Third "layer" (at TX 1)

$$\boldsymbol{t}_3^{\mathrm{HZF}}(1) = -\hat{\hat{H}}^{(1)\mathrm{H}}_{1,1} \left(|\hat{\hat{H}}^{(1)}_{1,1}|^2 + \frac{1}{P} \right)^{-1} \hat{\hat{\boldsymbol{h}}}^{(1)\mathrm{H}}_{1} \left(\begin{bmatrix} \boldsymbol{t}_3^{\mathrm{HZF}}(2) \\ 0 \end{bmatrix} + \boldsymbol{t}_3^{\mathrm{HZF}}(3) \right)$$

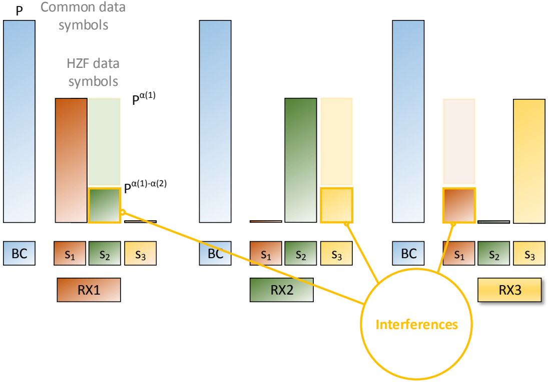


Main Intuition of the Proof

➥ Does not increase interference at user 2 when reducing interference at user 1 because

$$|\boldsymbol{t}_3^{\mathrm{HZF}}(1)|^2 \dot{\leq} P^{1-\alpha^{(2)}}$$

Transmission Scheme



- Less interference bits to convey in second layer
- More information bits can be squeezed in first layer