On the Cost of CSI Acquisition in Large MIMO Systems

Giuseppe Durisi

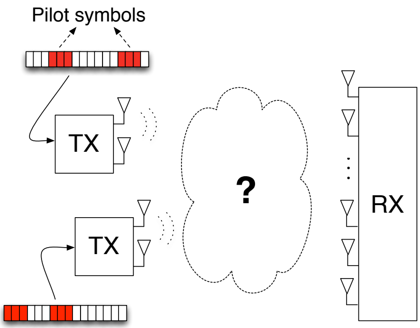
Chalmers, Sweden

June, 2013

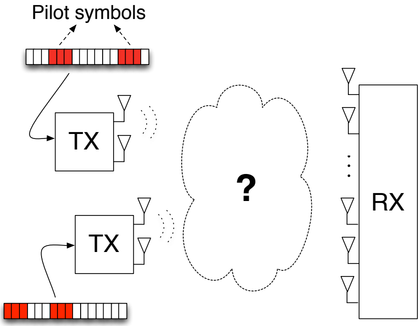
Joint work with **Wei Yang**, **Günther Koliander**, **Erwin Riegler**, **Franz Hlawatsch**, **Tobias Koch**, **Yury Polyanskiy**

Many thanks to **Ericsson Research Foundation**!

CSI acquisition limits large-MIMO gains



CSI acquisition limits large-MIMO gains

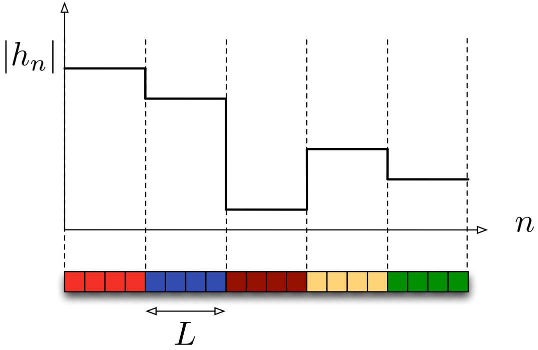


Capacity **in the absence of a priori** channel knowledge is the **ultimate limit** on the rate of reliable communication

Outline

1. Beyond the pre-log
2. Generic block-fading models
3. From asymptotics to finite-blocklength bounds

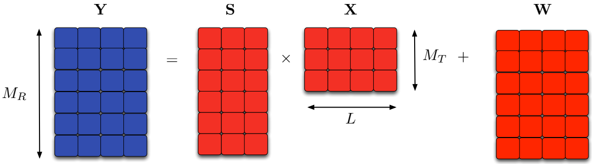
A simple channel model



Constant block-memoryless Rayleigh-fading channel

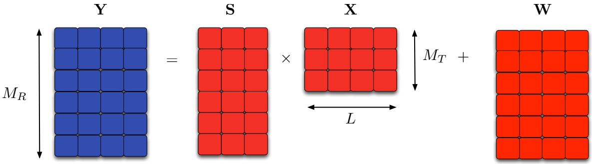
Coherence time is the bottleneck

- MIMO input-output relation



Coherence time is the bottleneck

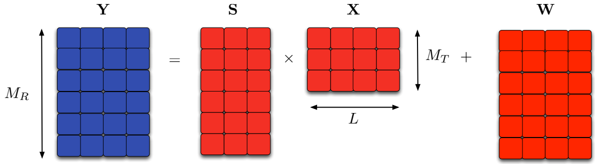
- MIMO input-output relation



- No closed-form expression available for $C(\rho)$

Coherence time is the bottleneck

- MIMO input-output relation

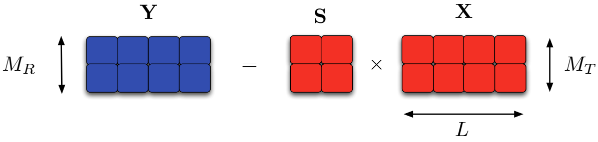


- No closed-form expression available for $C(\rho)$
- Pre-log [*Zheng & Tse, 2002*]

$$\chi = \lim_{\rho \to \infty} \frac{C(\rho)}{\log \rho} = M^* \left(1 - \frac{M^*}{L}\right)$$

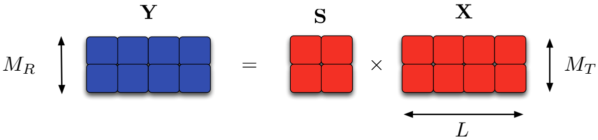
where $M^* = \min\{M_T, M_R, L/2\}$

The underlying geometry: $M_T = M_R = M$

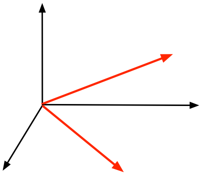


$$\chi = M\left(1 - \frac{M}{L}\right)$$

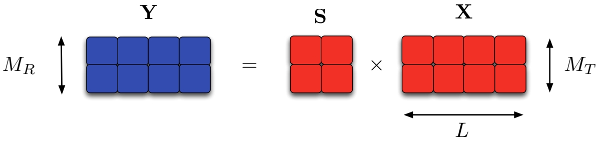
The underlying geometry: $M_T = M_R = M$



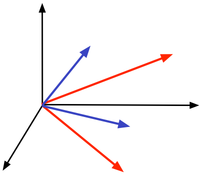
$$\chi = M\left(1 - \frac{M}{L}\right)$$



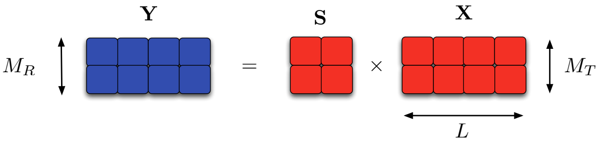
The underlying geometry: $M_T = M_R = M$



$$\chi = M\left(1 - \frac{M}{L}\right)$$

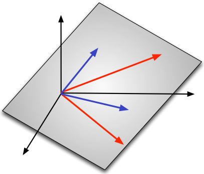


The underlying geometry: $M_T = M_R = M$



$$\chi = M\left(1 - \frac{M}{L}\right)$$

Communications on the Grassmannian manifold



Geometry suggests a signaling scheme

- **Uniform distribution** on the Grassmannian

$$\mathbf{X} = \sqrt{L\rho}\,\mathbf{U}$$

- $\mathbf{U}$: (truncated) **unitary** and **isotropically distributed**
- Unitary space-time modulation (**USTM**)

A conjecture

- Case $L \geq M_T + M_R$ ("small MIMO") [*Zheng & Tse (IT 2002)*]:

$$C(\rho) = R_{\mathrm{USTM}}(\rho) + o(1)$$

A conjecture

- Case $L \geq M_T + M_R$ ("small MIMO") [*Zheng & Tse (IT 2002)*]:

$$C(\rho) = R_{\mathrm{USTM}}(\rho) + o(1)$$

- **Conjecture** for $L < M_T + M_R$ ("large MIMO") [*Zheng & Tse (IT 2002)*]:

  - USTM **not** $o(1)$**-optimal**

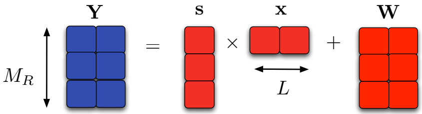
BSTM is the optimal distribution

[*Yang, Durisi, Riegler (JSAC 2013)*]

BSTM is $o(1)$-optimal when $L < M_T + M_R$ (large-MIMO)

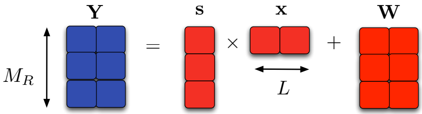
- $\mathbf{X} = \mathbf{D}\mathbf{U}$ with
- $\mathbf{U}$ i.d. and unitary
- $\mathbf{D}^2$ diagonal; contains the eigenvalues of a **complex matrix-variate beta distributed** matrix

Why is BSTM optimal? The SIMO case



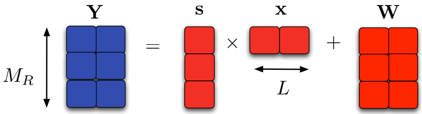
- Large MIMO $\Rightarrow L < 1 + M_R$

Why is BSTM optimal? The SIMO case



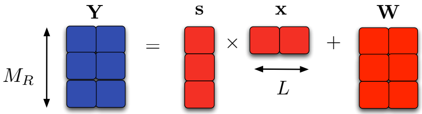
- Large MIMO $\Rightarrow L < 1 + M_R$
- USTM $\Rightarrow \mathbf{x}$ i.d., $\|\mathbf{x}\|^2 = L\rho$

Why is BSTM optimal? The SIMO case



- Large MIMO $\Rightarrow L < 1 + M_R$
- USTM $\Rightarrow \mathbf{x}$ i.d., $\|\mathbf{x}\|^2 = L\rho$
- BSTM $\Rightarrow \mathbf{x}$ i.d., $\frac{L-1}{\rho L M_R}\|\mathbf{x}\|^2 \sim$ Beta$(L-1, M_R+1-L)$

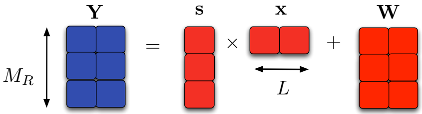
Why is BSTM optimal? The SIMO case



- Large MIMO $\Rightarrow L < 1 + M_R$
- USTM $\Rightarrow \mathbf{x}$ i.d., $\|\mathbf{x}\|^2 = L\rho$
- BSTM $\Rightarrow \mathbf{x}$ i.d., $\frac{L-1}{\rho L M_R}\|\mathbf{x}\|^2 \sim$ Beta$(L-1, M_R+1-L)$

$$I(\mathbf{x};\mathbf{Y}) = h(\mathbf{Y}) - h(\mathbf{Y} \,|\, \mathbf{x})$$

Why is BSTM optimal? The SIMO case



- Large MIMO $\Rightarrow L < 1 + M_R$
- USTM $\Rightarrow \mathbf{x}$ i.d., $\|\mathbf{x}\|^2 = L\rho$
- BSTM $\Rightarrow \mathbf{x}$ i.d., $\frac{L-1}{\rho L M_R}\|\mathbf{x}\|^2 \sim$ Beta$(L-1, M_R + 1 - L)$

$$I(\mathbf{x};\mathbf{Y}) = h(\mathbf{Y}) - h(\mathbf{Y}\,|\,\mathbf{x})$$

$$\approx h(\mathbf{s}\|\mathbf{x}\|) + 2(L - 1 - M_R)\,\mathbb{E}[\log\|\mathbf{x}\|] + \text{const}$$

Outline

The "generic" block-fading model

- Constant block-fading model for subchannel (r, t)

$$\mathbf{h}_{r,t} = \mathbf{1}_L \cdot s_{r,t}, \quad s_{r,t} \sim \mathcal{CN}(0, 1)$$

The "generic" block-fading model

- Constant block-fading model for subchannel (r, t)

$$\mathbf{h}_{r,t} = \mathbf{1}_L \cdot s_{r,t}, \quad s_{r,t} \sim \mathcal{CN}(0,1)$$

- A more accurate model for MIMO CP-OFDM systems

$$\mathbf{h}_{r,t} = \mathbf{z}_{r,t} \cdot s_{r,t}, \quad s_{r,t} \sim \mathcal{CN}(0,1)$$

$\mathbf{z}_{r,t} \in \mathbb{C}^L \Rightarrow$ Fourier transf. of power-delay profile

The “generic” block-fading model

- Constant block-fading model for subchannel (r, t)

$$\mathbf{h}_{r,t} = \mathbf{1}_L \cdot s_{r,t}, \quad s_{r,t} \sim \mathcal{CN}(0, 1)$$

- A more accurate model for MIMO CP-OFDM systems

$$\mathbf{h}_{r,t} = \mathbf{z}_{r,t} \cdot s_{r,t}, \quad s_{r,t} \sim \mathcal{CN}(0, 1)$$

$\mathbf{z}_{r,t} \in \mathbb{C}^L \Rightarrow$ Fourier transf. of power-delay profile

- We assume that $\{\mathbf{z}_{r,t}\}$ are generic

Generic $\{\mathbf{z}_{r,t}\}$ yield larger pre-log

[*Riegler, Koliander, Durisi, Hlawatsch (ISIT 2013)*]

- $\{\mathbf{z}_{r,t}\}$ generic and $M_R > \frac{M_T(L-1)}{L-T}$ with $M_T < L/2$

Generic $\{\mathbf{z}_{r,t}\}$ yield larger pre-log

[*Riegler, Koliander, Durisi, Hlawatsch (ISIT 2013)*]

- $\{\mathbf{z}_{r,t}\}$ generic and $M_R > \frac{M_T(L-1)}{L-T}$ with $M_T < L/2$
- Then

$$\chi_{\text{gen}} = M_T \left(1 - \frac{1}{L}\right)$$

Generic $\{\mathbf{z}_{r,t}\}$ yield larger pre-log

[*Riegler, Koliander, Durisi, Hlawatsch (ISIT 2013)*]

- $\{\mathbf{z}_{r,t}\}$ generic and $M_R > \frac{M_T(L-1)}{L-T}$ with $M_T < L/2$
- Then

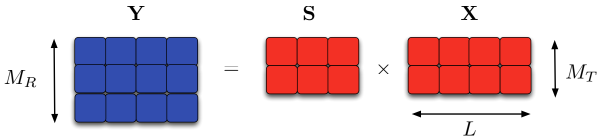
$$\chi_{\text{gen}} = M_T\left(1 - \frac{1}{L}\right)$$

- Compare with constant block-fading model

$$\chi_{\text{const}} = M_T\left(1 - \frac{M_T}{L}\right)$$

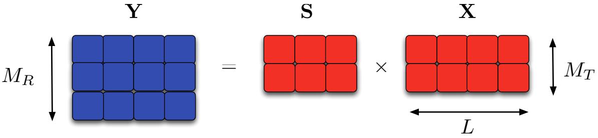
Intuition behind pre-log increase: $M_R = 3, M_T = 2, L = 4$

- Constant block-fading: $\chi_{\text{const}} = M_T \left(1 - \frac{M_T}{L}\right) = 1$

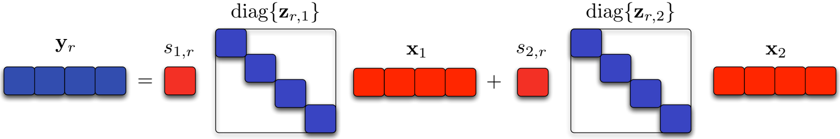


Intuition behind pre-log increase: $M_R = 3, M_T = 2, L = 4$

- Constant block-fading: $\chi_{\text{const}} = M_T\left(1 - \frac{M_T}{L}\right) = 1$



- Generic block-fading: $\chi_{\text{gen}} = M_T\left(1 - \frac{1}{L}\right) = \frac{3}{2}$



Outline

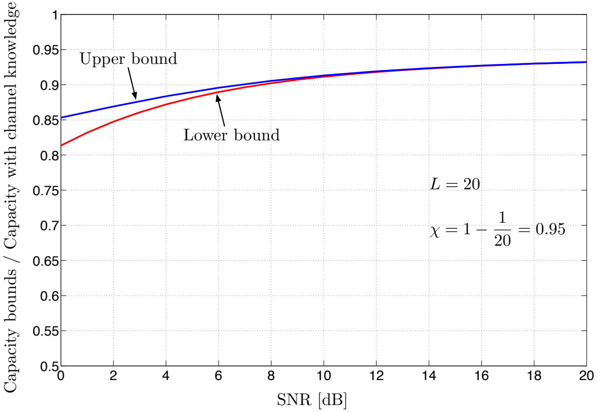
Lost in "asymptotia"?

Lost in "asymptotia"?

- capacity characterizations up to $o(1)$ yield tight bounds ☺
- pre-log sensitive to small changes in the channel model ☹

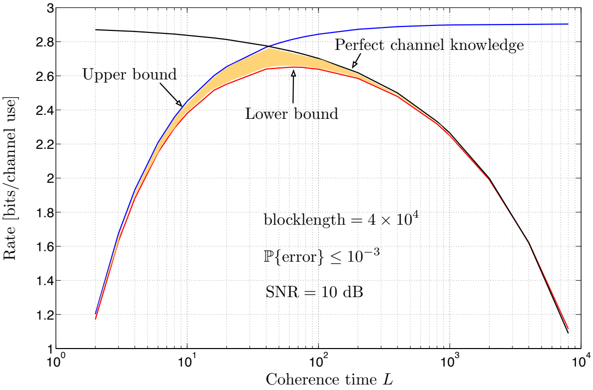
From asymptotia to tight bounds

[*Yang, Durisi, Koch, Polyanskiy (ITW 2012)*]



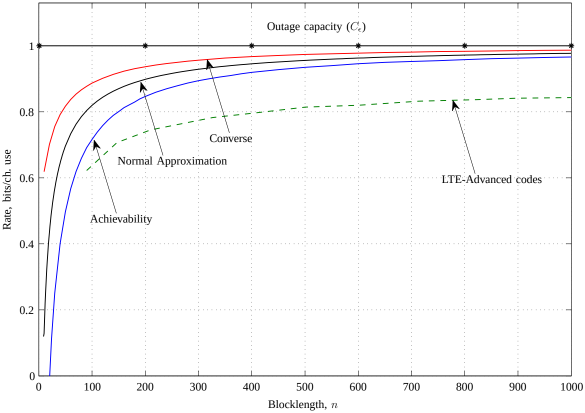
From asymptotia to tight bounds

[*Yang, Durisi, Koch, Polyanskiy (ITW 2012)*]



From asymptotia to tight bounds

[*Yang, Durisi, Koch, Polyanskiy (ISIT 2013)*]



Zero dispersion

- AWGN channel [*Polyanskiy, Poor, Verdú (IT 2010)*]

$$R^*_{\mathrm{awgn}}(n,\epsilon) = C_{\mathrm{awgn}} - \sqrt{\frac{V}{n}}Q^{-1}(\epsilon) - \mathcal{O}\left(\frac{\log n}{n}\right)$$

Zero dispersion

- AWGN channel [*Polyanskiy, Poor, Verdú (IT 2010)*]

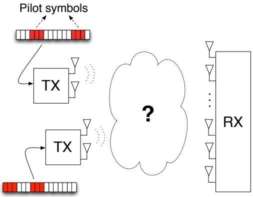
$$R^*_{\text{awgn}}(n,\epsilon) = C_{\text{awgn}} - \sqrt{\frac{V}{n}}Q^{-1}(\epsilon) - \mathcal{O}\left(\frac{\log n}{n}\right)$$

- SISO quasi static [*Yang, Durisi, Koch, Polyanskiy (ISIT 2013)*]

$$\{R^*_{\mathsf{csirt}}(n,\epsilon), R^*_{\mathsf{no}}(n,\epsilon)\} = C_\epsilon - 0\cancel{\sqrt{\frac{1}{n}}} - \mathcal{O}\left(\frac{\log n}{n}\right)$$

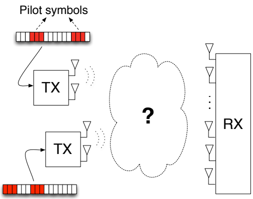
Summary

Capacity without a-priori CSI



Summary

Capacity without a-priori CSI

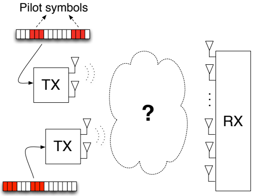


Too conservative estimates?

- USTM $\Rightarrow$ BSTM
- $M\left(1 - M/L\right) \Rightarrow M\left(1 - 1/L\right)$

Summary

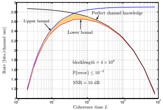
Capacity without a-priori CSI



Too conservative estimates?

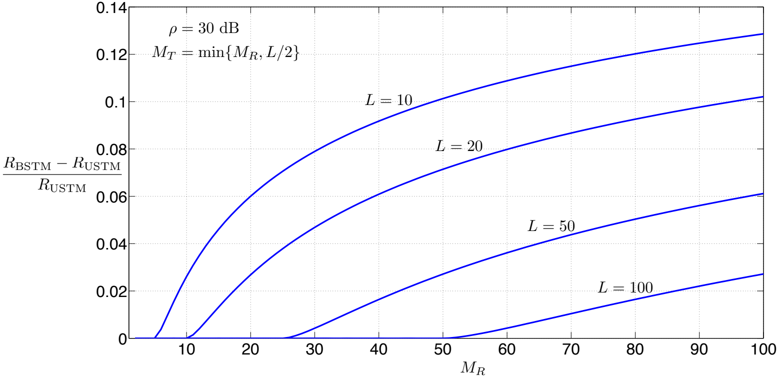
- USTM $\Rightarrow$ BSTM
- $M\,(1 - M/L) \Rightarrow M\,(1 - 1/L)$

From asymptotia to finite blocklength



Backup Slides

Gain of BSTM over USTM for large-MIMO systems



Achievability for finite blocklength