Rate Splitting Approaches for Noncoherent Fading Channels

Adriano Pastore

Ecole Polytechnique Fédérale de Lausanne
IC/ISC/LINX Laboratory
Lausanne, Switzerland
adriano.pastore@epfl.ch

Workshop on Recent Advances in Network Information Theory
INSA Lyon
November 24, 2015
Outline

1. Noncoherent fading with LOS
   Capacity lower bounds via rate splitting
   A mismatched decoding perspective on rate splitting

2. Noncoherent fading without LOS
   Capacity lower bounds via rate splitting

3. The multiple-access channel
Outline

1 Noncoherent fading with LOS
   Capacity lower bounds via rate splitting
   A mismatched decoding perspective on rate splitting

2 Noncoherent fading without LOS
   Capacity lower bounds via rate splitting

3 The multiple-access channel
Channel model

\[ Y_t = (\tilde{H} + \tilde{H}_t)X_t + Z_t \quad (t \in \mathbb{Z}) \]

Assumptions:

- \( \{X_t\} \) are symbols from an i.i.d. Gaussian codebook
  \[ C = \{(X_1(m), \ldots, X_n(m)) : m \in [1:e^{nR}]\} \]
- \( \{\tilde{H}_t\} \) is complex zero-mean i.i.d. ergodic
- \( \{Z_t\} \) is complex zero-mean i.i.d. ergodic (not necessarily Gaussian)
- \( \{X_t\}, \{\tilde{H}_t\} \) and \( \{Z_t\} \) are mutually independent
- \( \text{T}_x \) and \( \text{Rx} \) know distributions but ignore realizations of \( \tilde{H}_t \) and \( Z_t \)

Channel coding theorem

The supremum of rates that support reliable communication is

\[ \lim_{n \to \infty} \frac{1}{n} I(X^n; Y^n) = I(X; Y) \]
Médard’s lower bound

\[ Y = (\tilde{H} + \hat{H})X + Z \]
\[ = \tilde{H}X + \hat{H}X + Z \]

"Worst-case noise bound" (Médard 2000)

\[ I(X; Y) \geq \log \left( 1 + \frac{|\tilde{H}|^2P}{\tilde{\sigma}^2P + \sigma^2} \right) \]
\[ \triangleq I_{LB} \]

The rate-splitting approach

Fix a power allocation $\mathbf{P} = (P_1, P_2)$ with $P_1 + P_2 = P$.

Split $X$ into independent Gaussian signals:

$$X = X_1 + X_2 \quad P = P_1 + P_2$$

$$Y = (\bar{H} + \tilde{H})(X_1 + X_2) + Z$$

**Chain rule**

$$I(X; Y) = I(X_1, X_2; Y) = I(X_1; Y) + I(X_2; Y|X_1)$$
The rate-splitting approach

1. "Decode" $X_1$ while treating effective noise as AWGN:

   $Y = \tilde{H}X_1 + \tilde{H}X_2 + \tilde{H}(X_1 + X_2) + Z$

   effective noise

   \[ I(X_1; Y) \geq \log \left( 1 + \frac{|\tilde{H}|^2 P_1}{|\tilde{H}|^2 P_2 + \tilde{\sigma}^2 P + \sigma^2} \right) \triangleq R_1(P) \]

2. Decode $X_2$ knowing $X_1$:

   \[ Y' = Y - \tilde{H}X_1 \]

   \[ = \tilde{H}X_2 + \tilde{H}(X_1 + X_2) + Z \]

   effective noise

   \[ I(X_2; Y|X_1) \geq \mathbb{E}_{X_1} \left[ \log \left( 1 + \frac{|\tilde{H}|^2 P_2}{\tilde{\sigma}^2 (|X_1|^2 + P_2) + \sigma^2} \right) \right] \triangleq R_2(P) \]
The rate-splitting approach

- Médard’s bound: $I_{LB}$
- Rate-splitting bound: $I_{LB,RS}(P) \triangleq R_1(P) + R_2(P)$

$$I_{LB,RS}(P) = R_1(P) + \text{E}_{X_1} \left[ \log \left( 1 + \frac{|H|^2P_2}{\tilde{\sigma}^2(X_1^2 + P_2) + \sigma^2} \right) \right]$$

Jensen's inequality:

$$\geq R_1(P) + \log \left( 1 + \frac{|H|^2P_2}{\tilde{\sigma}^2(P_1 + P_2) + \sigma^2} \right)$$

$$= I_{LB}$$

$$I(X; Y) \geq I_{LB,RS} \geq I_{LB}(P)$$
Infinitiesimal rate splitting

Fix a power allocation \( \mathbf{P} = (P_1, \ldots, P_L) \) with \( P_1 + \ldots + P_L = P \).

\[
I(X; Y) = \sum_{\ell=1}^{L} I(X_\ell; Y|X_1, \ldots, X_{\ell-1}) \\
\geq \sum_{\ell=1}^{L} \mathbb{E} \left[ \log \left( 1 + \frac{P_\ell |\bar{H}|^2}{\bar{\sigma}^2 \sum_{i<\ell} |X_i|^2 + \bar{\sigma}^2 P_\ell + (|\bar{H}|^2 + \bar{\sigma}^2) \sum_{i>\ell} P_i + \sigma^2} \right) \right]
\]

\[
\triangleq \sum_{\ell=1}^{L} R_\ell(\mathbf{P}) \\
= I_{LB,RS}(\mathbf{P})
\]
Infinitesimal rate-splitting

**Theorem**

The best rate-splitting bound is

\[ I_{LB}^* \triangleq \sup_{P: P_1 + \ldots + P_L = P} I_{LB,RS}(P) \]

\[ = \lim_{\max_{\ell} P_{\ell} \downarrow 0} I_{LB,RS}(P) \]

\[ = E_U \left[ \int_0^1 \frac{|\tilde{H}|^2}{(|\tilde{H}|^2 + \tilde{\sigma}^2)(1 - \lambda) + \tilde{\sigma}^2 U \lambda + \rho^{-1}} d\lambda \right] \]

\[ = E_U \left[ \frac{|\tilde{H}|^2}{\tilde{\sigma}^2 (U - 1) - |\tilde{H}|^2} \log \left( 1 + \frac{\tilde{\sigma}^2 (U - 1) - |\tilde{H}|^2}{|\tilde{H}|^2 + \tilde{\sigma}^2 + \rho^{-1}} \right) \right] \]

where \( U \) is unit-mean exponential.

Numerical results

Figure: Channel parameters: $|\tilde{H}|^2 = \tilde{\sigma}^2 = 1/2$. 
Nearest-neighbor decoding

Maximum-likelihood (matched) decoding:

\[
\hat{m} = \arg\max_{m \in [1:e^{nR}]} \prod_{t=1}^{n} p(Y_t|X_t(m))
\]

Nearest-neighbor (mismatched) decoding:

\[
\hat{m} = \arg\min_{m \in [1:e^{nR}]} \sum_{t=1}^{n} \left| Y_t - \bar{H}X_t(m) \right|^2
\]

**Generalized mutual information**

The GMI (associated to NND and to i.i.d. Gaussian codes) is the supremum of rates \( R \) such that \( \Pr\{m \neq \hat{m}\} \xrightarrow{n \to \infty} 0 \) and satisfies

\[
\text{GMI} \leq I(X; Y)
\]
Médard’s bound and NND

Theorem [Lapidoth and Shamai 2002]
The GMI under nearest-neighbor decoding and Gaussian codes satisfies
\[ \text{GMI} = I_{LB} \]

Can this insight be extended to rate splitting?

The naive extension: successive NND

\[\hat{m}_1 = \arg\min_{m_1 \in [1:e^{nR_1}]} \sum_{t=1}^{n} |Y_t - \bar{H}X_{1,t}(m_1)|^2\]

\[\hat{m}_2(\hat{m}_1) = \arg\min_{m_2 \in [1:e^{nR_2}]} \sum_{t=1}^{n} |Y_t - \bar{H}X_{1,t}(\hat{m}_1) - \bar{H}X_{2,t}(m_2)|^2\]

This decoding rule achieves a sum rate

\[
\sup_{\theta \leq 0} \{ T_1 \theta - \Lambda_1(\theta) \} + \sup_{\theta \leq 0} E_{X_1} \left[ T_2 \theta - \Lambda_2(\theta) \right] = I_{LB}
\]

where

\[T_1 = |\bar{H}|^2 P_2 + \tilde{\sigma}^2 P + \sigma^2\]

\[\Lambda_1(\theta) = \frac{(|\bar{H}|^2 + \tilde{\sigma}^2)P + \sigma^2}{1 - |\bar{H}|^2 P_1 \theta} \theta - \log(1 - |\bar{H}|^2 P_1 \theta)\]

\[T_2 = \tilde{\sigma}^2 P + \sigma^2\]

\[\Lambda_2(\theta) = \frac{|\bar{H}|^2 P_2 + \tilde{\sigma}^2(|X_1|^2 + P_2) + \sigma^2}{1 - |\bar{H}|^2 P_2 \theta} \theta - \log(1 - |\bar{H}|^2 P_2 \theta)\]
The naive approach: why does it fail?

We would wish to swap \( \sup \) and \( E_{X_1} \):

\[
I_{LB} = \sup_{\theta < 0} \{ T_1 \theta - \Lambda_1(\theta) \} + \sup_{\theta < 0} E_{X_1} \left[ T_2 \theta - \Lambda_2(\theta) \right]
\]

\[
\leq \sup_{\theta < 0} \{ T_1 \theta - \Lambda_1(\theta) \} + E_{X_1} \left[ \sup_{\theta < 0} \{ T_2 \theta - \Lambda_2(\theta) \} \right] = I_{LB,RS}(P)
\]
The smart approach: successive NND + multiplexing

Thresholds: \(0 = \xi_0 < \xi_1 < \xi_2 < \ldots < \xi_K = \infty\)

Time partitions: \(\mathcal{T}_k = \left\{ t \in \mathbb{N} : \xi_{k-1} \leq |X_1,t| < \xi_k \right\}\)

Time-multiplexing of \(K\) virtual users on the second layer:

\[
\sup_{\theta<0} \mathbb{E}_{X_1} \left[ T_2\theta - \Lambda_2(\theta) \right] \\
\xrightarrow{K \geq 1} \sum_{k=1}^{K} \sup_{\theta<0} \mathbb{E}_{X_1} \left[ T_2\theta - \Lambda_2(\theta) \bigg| \xi_{k-1} \leq |X_1| < \xi_k \right] \cdot \mu_{|X_1|} \left( [\xi_{k-1}; \xi_k) \right) \\
\xrightarrow{K \gg 1} \sum_{k=1}^{K} \sup_{\theta<0} \left\{ T_2\theta - \Lambda_2(\theta) \right\} \bigg|_{|X_1| = \xi_k} \cdot \mu_{|X_1|} \left( [\xi_{k-1}; \xi_k) \right) \\
\xrightarrow{K \to \infty} \int_{0}^{\infty} \sup_{\theta<0} \left\{ T_2\theta - \Lambda_2(\theta) \right\} \bigg|_{|X_1| = \xi} \cdot d\mu_{|X_1|} (\xi) \\
= \mathbb{E}_{X_1} \left[ \sup_{\theta<0} \left\{ T_2\theta - \Lambda_2(\theta) \right\} \right]
\]
The smart approach: in-layer multiplexing
Power allocation diagrams

<table>
<thead>
<tr>
<th></th>
<th>(2,1)</th>
<th>(2,2)</th>
<th>(2,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,1)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure**: Example of a power allocation diagram
Power allocation diagrams

\[
\begin{array}{|c|c|c|c|c|}
\hline
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) \\
\hline
(3,1) & & (3,2) & & \\
\hline
(2,1) & (2,2) & (2,3) & (2,4) & \\
\hline
(1,1) & & & & \\
\hline
\end{array}
\]

**Figure**: Example of a power allocation diagram for \( L = 4 \).
**Power allocation**

<table>
<thead>
<tr>
<th></th>
<th>(2,1)</th>
<th>(2,2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1,1)</td>
<td></td>
</tr>
</tbody>
</table>

*Figure: Allocation diagram for $L = 2$*
Power allocation

\[
\sum \text{GMI} = \frac{P_0}{P_1} - e^{-\xi} \left( 1 - e^{-\xi} \right)
\]
Outline

1. Noncoherent fading with LOS
   - Capacity lower bounds via rate splitting
   - A mismatched decoding perspective on rate splitting

2. Noncoherent fading without LOS
   - Capacity lower bounds via rate splitting

3. The multiple-access channel
Looseness of previous bounds

\[ Y = (\tilde{H} + \tilde{H})X + Z \]

**Lower bounds on** \( I(X; Y) \) (reminder)

\[
I_{LB} = \log \left( 1 + \frac{|\tilde{H}|^2 P}{\tilde{\sigma}^2 P + \sigma^2} \right)
\]

\[
I_{LB,RS}^* = \mathbb{E}_U \left[ \int_0^1 \frac{|\tilde{H}|^2 P}{(1 - \lambda)(|\tilde{H}|^2 + \tilde{\sigma}^2)P + \lambda\tilde{\sigma}^2 UP + \sigma^2} \, d\lambda \right]
\]

For \( \tilde{H} = 0 \) we get \( I_{LB} = I_{LB,RS}^* = 0 \) but we know \( I(X; Y) > 0 \)
Failure of Médard’s bounding approach

Médard’s bound (reminder)

\[ I_{LB} = \log \left( 1 + \frac{|\bar{H}|^2 P}{\tilde{\sigma}^2 P + \sigma^2} \right) \leq I(X; Y) \]

Proof:

\[ I(X; Y) = h(X) - h(X|Y) \]
\[ = \log(\pi e P) - h(X - \hat{X}(Y)|Y) \]
\[ \geq \log(\pi e P) - h(X - \hat{X}(Y)) \]
\[ \geq \log(P) - \log(\text{var}(X - \hat{X}(Y))) \]
\[ = \log(P) - \log(P + \text{var}(\hat{X}(Y)) - 2 \text{Re}\{E[X*\hat{X}(Y)]\}) \]

Choosing \( \hat{X}(Y) = \alpha Y \) does not do the trick since \( E[X^*\hat{X}(Y)] = \alpha \bar{H} P \)

This approach fails for \( \bar{H} = 0 \)
A fix: biased signalling + non-linear input estimator

Suggestion: \[
\begin{aligned}
X &\sim \mathcal{N}_c(\sqrt{P_0}, P_1) \quad \text{with} \quad P_0 + P_1 = P \\
\hat{X}(Y) &= \alpha |Y|^2
\end{aligned}
\]

\[
E[X^* \hat{X}(Y)] = \alpha \tilde{\sigma}^2 P_1 \sqrt{P_0}
\]

After optimizing \(\alpha\):

**Lower bound on** \(I(X; Y)\) **for noncoherent fading without LOS**

\[
I(X; Y) \geq \log \left( \frac{\text{var}(|Y|^2)}{\text{var}(|Y|^2) - \tilde{\sigma}^4 P_0 P_1} \right)
\]
Theorem

The optimal power tradeoff is attained at a ratio

\[
\frac{P_0}{P_1 \bigg|_{\text{opt}}} = \frac{\eta(\text{snr}) - \sqrt{\eta(\text{snr})(\eta(\text{snr}) - \text{snr}^2 \kappa_{\tilde{H}})}}{\text{snr}^2 \kappa_{\tilde{H}}}
\]

where

- \( \kappa_A = \text{E}[|A|^4]/\text{E}[|A|^2]^2 \) denotes the kurtosis of \( A \)
- \( \text{snr} \triangleq P/\sigma^2 \)
- \( \eta(\text{snr}) \triangleq (2\kappa_{\tilde{H}} - 1)\text{snr}^2 + 2\text{snr} + \kappa_Z - 1 \)
Bias/signal power tradeoff

\[ \lim_{P/\sigma^2 \to 0} \frac{P_0}{P_1}_{\text{opt}} = \frac{1}{2} \]

\[ \lim_{P/\sigma^2 \to \infty} \frac{P_0}{P_1}_{\text{opt}} = 2 - \frac{1}{\kappa \bar{H}} - \sqrt{\left(2 - \frac{1}{\kappa \bar{H}}\right)\left(1 - \frac{1}{\kappa \bar{H}}\right)} \]
Infinitesimal rate splitting

Consider allocations \( \mathbf{P} = (P_0, P_1, \ldots, P_L) \) with

\[
P_0 + P_1 + \ldots + P_L = P
\]

\[
X_0 + X_1 + \ldots + X_L = X
\]

with \( X_0 = \sqrt{P_0} = \text{const.} \)

\[
I(X; Y) = I(X_0; Y) + \sum_{\ell=1}^{L} I(X_\ell; Y | X_1, \ldots, X_{\ell-1}) \\
\geq 0 + \sum_{\ell=1}^{L} \mathbb{E} \left[ \log \left( \frac{\text{var}(|Y|^2 | X_\ell)}{\text{var}(|Y|^2 | X_\ell) - \tilde{\sigma}^4 P_\ell |X_\ell|^2} \right) \right]
\]

\[
\approx \sum_{\ell=1}^{L} \mathbb{E} \left[ \frac{\tilde{\sigma}^4 P_\ell |X_\ell|^2}{\text{var}(|Y|^2 | X_\ell)} \right]
\]

\[
\xrightarrow{L \to \infty} \int_{0}^{1} \mathbb{E} \left[ \frac{\tilde{\sigma}^4 |X_\lambda|^2}{\text{var}(|Y|^2 | X_\lambda)} \right] P \, d\lambda
\]
Infinitesimal rate splitting

Written out in full:

\[ I_{LB,RS}^* = \text{snr}^2 \int_0^1 \int_0^\infty \frac{\lambda \nu}{D(\lambda \nu, 1 - \lambda)} e^{-\nu} \, d\nu \, d\lambda \]

where

\[ D(x, y) = \text{snr}^2 (\kappa_H - 1)x^2 + \text{snr}^2 (2\kappa_H - 1)y(2x + y) + 2\text{snr}(x + y) + \kappa_Z - 1 \]

and

- \( \kappa_A = \frac{E[|A|^4]}{E[|A|^2]^2} \) denotes kurtosis
- \( \text{snr} \triangleq \frac{P}{\sigma^2} \)
Numerical results

\[ I(X; Y) \text{ (Rayleigh fading + AWGN)} \]

\[ I_{LB,RS} \]

\[ I_{LB} \]
Outline

1. Noncoherent fading with LOS
   Capacity lower bounds via rate splitting
   A mismatched decoding perspective on rate splitting

2. Noncoherent fading without LOS
   Capacity lower bounds via rate splitting

3. The multiple-access channel
Multiple-access channel

\[ Y_t = (\hat{H}_1 + \tilde{H}_{1,t})X_{1,t} + (\tilde{H}_2 + \hat{H}_{2,t})X_{2,t} + Z_t \quad (t \in \mathbb{Z}) \]

**MAC rate region**

For fixed \((P_1, P_2)\) the rate region is the convex hull around

\[
\begin{bmatrix}
I(X_1; Y|X_2) & I(X_2; Y)
\end{bmatrix}
\quad \text{and} \quad
\begin{bmatrix}
I(X_1; Y) & I(X_2; Y|X_1)
\end{bmatrix}
\]

An inner bound is obtained using lower bounds

\[
I_{LB,RS}^*(X_1; Y|X_2) \leq I(X_1; Y|X_2)
\]
\[
I_{LB,RS}^*(X_2; Y) \leq I(X_2; Y)
\]
\[
I_{LB,RS}^*(X_1; Y) \leq I(X_1; Y)
\]
\[
I_{LB,RS}^*(X_2; Y|X_1) \leq I(X_2; Y|X_1)
\]
Multiple-access channel with LOS

Figure: MAC with $P_1/\sigma^2 = P_2/\sigma^2 = 1$ and $|\bar{H}_1|^2 = |\bar{H}_2|^2 = \tilde{\sigma}_1^2 = \tilde{\sigma}_2^2 = 1/2$
Multiple-access channel without LOS

Figure: MAC with $P_1/\sigma^2 = P_2/\sigma^2 = 10$dB and $\bar{\sigma}_1^2 = \bar{\sigma}_2^2 = \frac{1}{2}$
Bounds put together

Figure: $| \tilde{H} |^2 = \alpha$ and $\tilde{\sigma}^2 = 1 - \alpha$. SNR is $P/\sigma^2 = 0$dB.
Conclusion

- With LOS
  - Rate splitting improves Médard’s bound
  - Improved bound achievable with a modified NND scheme
- Without LOS
  - Modified Médard-type lower bound with biased signalling
  - Rate splitting allows to
    - remove the signal bias
    - greatly improve the bound
    - come close to exact MI for Rayleigh AWGN
- Both bounds applicable to MAC

Open questions:
- Non-LOS achievable with modified NND scheme?
- Does multiplexing allow analysis of non-Gaussian codes?

Thanks!