Lecture 10: Universal Code Design for Optimal DMT
Theoretical Foundations of Wireless Communications

Ragnar Thobaben
CommTh/EES/KTH

Thursday, June 5, 2014 09:15-12:00, Conference Room SIP

Overview
Lecture 9: Diversity-multiplexing tradeoff
- How much diversity can we obtain if the rate increases linearly with log SNR, \( R = r \log \text{SNR} \).
- DMT for various slow Rayleigh fading channels
  - Scalar channels
  - Parallel channels
  - MISO channel
  - MIMO channels

Lecture 10: Design of DMT optimal Codes (Ch. 9.2)
1. Approximately Universal Schemes
2. Scalar Channels
3. Parallel Channels
4. MISO Channels
5. MIMO Channels

Notes
Approximately Universal Schemes

Outage and universal codes

- Achieve arbitrarily small error probabilities whenever the channel is not in outage (long block length, strong code).
- Deep-fade event dominates the error performance at high SNRs; it depends on the scheme and the channel.

Approximately universal schemes

- The deep-fade event depends only on the channel but not on the scheme.
- Approximately universal is sufficient to achieve the optimal DMT of the channel.

Scalar Channels

Why is uncoded QAM DMT-optimal?

- Error probability conditioned on the fading coefficient $h$:

  $$Q\left(\sqrt{\frac{\text{SNR}}{2}} |h|^2 d_{\text{min}}^2 \right), \quad \text{with } d_{\text{min}} \approx \frac{1}{2^{R/2}}.$$

  (normalized distance $d_{\text{min}}$)

- Consider the two cases:
  - $\sqrt{\text{SNR}} |h| d_{\text{min}} \gg 1$: separation of constellation points is much larger than the standard deviation of the noise.
  - $\sqrt{\text{SNR}} |h| d_{\text{min}} \approx 1$: separation of constellation points and standard deviation of the noise are in the same order of magnitude.

  $\Rightarrow$ Deep fade event: $|h|^2 < 2^R \frac{1}{\text{SNR}}$.

- Channel outage condition:

  $$\log(1 + |h|^2 \text{SNR}) < R \Leftrightarrow |h|^2 < 2^R - 1 \frac{1}{\text{SNR}}.$$

  $\Rightarrow$ For high SNR and high rates, the channel-outage condition and deep-fade condition coincide; i.e., typical errors occur only when the channel is in outage.
Scalar Channels

Comments
- The analysis is only conditioned on the realization $h$ but not on the distribution $p(h)$.
- QAM achieves the optimal diversity-multiplexing tradeoff under any channel statistics (not only Rayleigh); i.e., it is universal.
- Approximate universality of QAM depends only on the condition
\[ d_{\text{min}}^2 > \frac{1}{2R}, \]
any other constellation with this property is as well approximately universal.

Summary
- A scheme is approximately universal if it is in a deep fade only when the channel is in outage.
- Being approximately universal is sufficient to achieve the DMT of the channel.

Parallel Channels - Universal Design Criterion

- Consider $L$ parallel channels
\[ y_l = h_l x_l + w_l \]
with \begin{align*}
\text{i.i.d. AWGN } w_l &\sim \mathcal{CN}(0, 1), \\
\text{transmit power per sub-channel } P &\equiv \text{SNR}.
\end{align*}

- Coding: length-$L$ codewords across the $L$ parallel channels; rate $R$ bits/s/Hz per sub-channel.

- Channel outage condition:
\[ \sum_{l=1}^{L} \log(1 + |h_l|^2 \text{SNR}) < LR. \]

- Conditional pairwise error probability for codewords $x_A, x_B$:
\[ \Pr(x_A \rightarrow x_B | h) = Q\left( \sqrt{\frac{\text{SNR}}{2}} \sum_{l=1}^{L} |h_l|^2 |d_l|^2 \right), \]
with the $l$-th component of the normalized codeword difference
\[ d_l = \frac{1}{\sqrt{\text{SNR}}} (x_{A_l} - x_{B_l}). \]
Parallel Channels – Universal Design Criterion

- Goal: Find the worst-case error probability over all channels which are not in outage

\[
\min_{h_1, \ldots, h_L} \frac{\text{SNR}}{2} \sum_{l=1}^{L} |h_l|^2 |d_l|^2 \quad \text{subject to} \quad \sum_{j=1}^{L} \log(1 + |h_j|^2 \text{SNR}) \geq LR.
\]

→ Optimization problem!

- Alternative formulation: with \( Q_l = \text{SNR} |h_l|^2 |d_l|^2 \) we get

\[
\min_{Q_1 > 0, \ldots, Q_L > 0} \frac{1}{2} \sum_{l=1}^{L} Q_l \quad \text{subject to} \quad \sum_{j=1}^{L} \log(1 + \frac{Q_j}{|d_j|^2}) \geq LR.
\]

→ Solution similar to waterfilling!

- Worst case channel:

\[
|h|^2 = \frac{1}{\text{SNR}} \left[ \frac{1}{|d|^2} - 1 \right]^+, \]

with the Lagrange multiplier \( \lambda \) satisfying the non-outage constraint.

Parallel Channels – Universal Design Criterion

- Worst-case error probability

\[
Q \left( \frac{1}{2} \sum_{j=1}^{L} \left[ \frac{1}{\lambda} - |d_l|^2 \right]^+ \right),
\]

with \( \lambda \) satisfying

\[
\sum_{j=1}^{L} \left[ \log \left( \frac{1}{\lambda |d_l|^2} \right) \right]^+ = LR.
\]

- Universal design criterion :

\[
\sum_{j=1}^{L} \left[ \frac{1}{\lambda} - |d_l|^2 \right]^+.
\]
Parallel Channels
- Examples

No coding

- \( L \) independent constellations transmitted separately over the sub-channels.
- Worst case: two codewords \( x_A, x_B \) differ only in one symbol; i.e., there is only one non-zero \( d_i \).
- Universal design criterion evaluates to zero.

Repetition Code

- The same symbol is transmitted over all sub-channels.
- Example: \( L = 2 \) sub-channels, \( R = 2 \) bits/s/Hz.

\[
\begin{array}{cccc}
\circ & \circ & \circ & \circ \\
\bigtriangledown & \bigtriangledown & \bigtriangleup & \bigtriangleup \\
\bigtriangleup & \bigtriangleup & \bigtriangleup & \bigtriangleup \\
\end{array}
\]

With \( |d_{\text{min}}|^2 = 4/9 \), the universal design criterion evaluates to 8/3.

(D. Tse and P. Viswanath, Fundamentals of Wireless Communications.)

Notes

Parallel Channels
- Examples

Permutation Code

- The same constellation is used for the sub-channels but different symbol mappings.
- Symbols which are close to each other in one constellation are far apart in the other constellations.
- Example: \( L = 2 \) sub-channels, \( R = 2 \) bits/s/Hz.

\[
\begin{array}{cccc}
\circ & \circ & \circ & \circ \\
\bigtriangledown & \bigtriangledown & \bigtriangleup & \bigtriangleup \\
\bigtriangleup & \bigtriangleup & \bigtriangleup & \bigtriangleup \\
\end{array}
\]

(D. Tse and P. Viswanath, Fundamentals of Wireless Communications.)

The universal design criterion evaluates to 44/9.
Parallel Channels
- Universal Design Criterion at High SNR

- Evaluation of the universal design criterion is complicated.
- Simple bound by relaxing the non-negativity constraint:

$$L^2 R \left| \prod_{l=1}^{L} d_l \right|^{2/L} - \sum_{l=1}^{L} |d_l|^2.$$  

→ Lower bound on the universal design criterion.
→ Tight when the rate per sub-channel $R$ becomes large.
   (Deep water level in the waterfilling problem.)
- For good codes (i.e. large distance), the first term dominates the universal design criterion, and it can be approximated as

$$L^2 R \left| \prod_{l=1}^{L} d_l \right|^{2/L}.$$  

⇒ Choose codewords that maximize the pairwise product distance!

Parallel Channels
- Properties

- Pairwise typical error event (as before: argument of $Q(\sqrt{\cdot/2})$ is less than 1):

$$\text{SNR} \sum_{l=1}^{L} |h_l|^2 |d_l|^2 < 1.$$  

Approximate universality requires that
→ this event occurs if the channel is in outage;
→ this event does not occur if the channel is not in outage.
- Worst-case code design criterion should be greater than 1; i.e., at high SNR,

$$\left| \prod_{l=1}^{L} d_l \right|^{2/L} > \frac{1}{L^2 R}.$$  

⇒ Sufficient condition for achieving the optimal DMT of the parallel channel.
Parallel Channels
  - Properties

• Permutation codes can be designed to be approximately universal.

• Permutation code for \( L \) parallel channels
  - Fix the QAM constellation and define \( L - 1 \) permutations \( \pi_2, \ldots, \pi_L \).
  - A message \( m \) is encoded by symbol \( q \) on the first channel and will be encoded by the symbols \( \pi_2(q), \ldots, \pi_L(q) \) on the remaining \( L - 1 \) channels.

• Example (\( L = 3 \)):

  Design the permutations \( \pi_2, \ldots, \pi_L \) to maximize the universal code design criterion.

Parallel Channels
  - Bit-Reversal Scheme

• Setup: \( L = 2 \) sub-channels, total rate \( 2R \) bits/s/Hz.

• No-outage condition:
  \[
  \frac{\log(1 + |h_1|^2 \text{SNR})}{\text{rate from sub-channel 1}} + \frac{\log(1 + |h_2|^2 \text{SNR})}{\text{rate from sub-channel 2}} > 2R.
  \]

• Assumptions:
  - Independent coding over the I and Q channels of the two channels; consider only the I channel (rate \( R \)).
  - The symbols \( x_i \) are labeled by the binary representation of the \( R \) bits.
  - The symbols \( x_i' \) are labeled in the reversed order.

• Intuitively (typical error event analysis), we can expect to be able to recover \( R_1 \) bits and \( R_2 \) bits over the respective channels, with
  \[
  |h_1|^2 \text{SNR} > 2^{2R_1}.
  \]

→ As long as \( R_1 + R_2 \geq R \) we can recover all bits, which is equivalent to the no-outage condition!
**MISO Channels - Conversion to Parallel Channels**

- Outage event for the $n_t \times 1$ MISO channel
  $$\log \left( 1 + \frac{|h|^2 \text{SNR}}{n_t} \right) < R.$$

- Special case $n_t = 2$: Alamouti scheme converts the MISO channel into a scalar channel with gain $|h|$ and power reduced by a factor 2.
  $\rightarrow$ Same outage behavior!

- What can be done for the general case $n_t > 2$?

- Possible solution: MISO channel as parallel channel.
  - Use one antenna at a time with a tradeoff-optimal parallel code.
    $\rightarrow$ Achieves the highest diversity gain $n_t$ for i.i.d. Rayleigh fading.
    $\rightarrow$ Tradeoff-optimal for channels with i.i.d. fading coefficients but not in general.

- Counterexample: channel with $h_1 \neq 0$ and $h_l = 0$ for $l \in \{2, \ldots, L\}$.
  - Channel outage probability: $p_{\text{out}} = \Pr(\log(1 + |h_1|^2 \text{SNR}) < R)$
  - Parallelized channel outage probability:
    $$p_{\text{parallel}} = \Pr(\log(1 + |h_1|^2 \text{SNR}) < n_t R).$$
  $\rightarrow$ Conversion to parallel channels is not optimal in this case!


**MISO Channels**

- **Universal Code Design Criterion**
  
  As above...
  
  \[ Pr\{X_A \rightarrow X_B|\mathbf{h}\} = Q\left(\frac{\|h^*(X_A - X_B)\|}{\sqrt{2}}\right) = Q\left(\sqrt{\frac{\text{SNR} \sum_{l=1}^{L} \left|\hat{h}_l^o\right|^2 \lambda_l^2}{2}}\right) \]

  - Worst-case over all channels not in outage
    \[
    \max_{\mathbf{h}, \|\mathbf{h}\|^2 \geq \frac{1}{\text{SNR}(2R - 1)}} Q\left(\frac{\|h^*(X_A - X_B)\|}{\sqrt{2}}\right) = Q\left(\sqrt{\frac{1}{2} \frac{1}{n_t}(2R - 1)}\right),
    \]
    with the smallest singular value \(\lambda_1\) of the normalized codeword difference
    \[
    \frac{1}{\sqrt{\text{SNR}}}(X_A - X_B).
    \]
    It aligns itself to the “weakest direction” of the codeword difference matrix.

  → **Design Criterion**
  
  - Maximize the minimum singular value of the codeword distance.
  - Ensure that all codeword pairs satisfy
    \[
    \lambda_2^2 > \frac{1}{n_t(2R - 1)} \approx \frac{1}{n_t 2R} \quad \text{(argument of } Q(\sqrt{1/2}) \text{ bigger than 1.)}
    \]

**MIMO Channels**

- **Universality of D-BLAST**
  
  - Channel model: slow fading MIMO channel,
    \[ y[m] = Hx[m] + w[m]. \]
  - Outage event for transmit covariance matrix \(K_x\)
    \[ \log \det(I_{n_t} + HK_x H^*) < R. \]
  - D-BLAST with MMSE-SIC converts the MIMO channel into \(n_t\) parallel sub-channels such that
    \[ \log \det(I_{n_t} + HK_x H^*) = \sum_{k=1}^{n_t} \log(1 + \text{SINR}_k). \]
    (Note: the SINR\(_1\), \ldots, SINR\(_n_t\) are correlated.)
MIMO Channels

– Universality of D-BLAST

- Outage probability (by using approximately universal parallel channel codes with D-BLAST and MMSE-SIC at a rate $R = r \log \text{SNR}$ bits/s/Hz per stream)

$$\Pr \left( \sum_{k=1}^{n_t} \log(1 + \text{SINR}_k) < R \right).$$

- Outage probability considering the initialization loss for $n$ interleaved streams

$$\Pr \left( \sum_{k=1}^{n} \log \det(I_{n_t} + H_k H_k^H) < r \log \text{SNR} \right).$$

- It follows by comparison with the MIMO outage probability

$$\rho_{\text{out}}^{\text{D-BLAST}}(r \log \text{SNR}) = \rho_{\text{out}}^\text{mimo} \left( \frac{r(n + n_t - 1)}{n} \log \text{SNR} \right).$$

→ Achieves the universally tradeoff curve for $n \to \infty$, but it is strictly sub-optimal for finite $n$.

→ Improved tradeoff performance by joint ML decoding of the streams.

MIMO Channels

Example: (2 × 2) MIMO Rayleigh fading channel.

- $n = 2$ streams with sub-codewords

$$x^{(1)} = [x_A^{(1)}, x_B^{(1)}]$$

$$x^{(2)} = [x_A^{(2)}, x_B^{(2)}]$$

- Transmission over antennas and time:

$$\begin{bmatrix}
0 & x_B^{(1)} & x_B^{(2)} \\
x_A^{(1)} & x_A^{(2)} & 0
\end{bmatrix}$$

- D-BLAST with ML decoding is DMT optimal for $0 \leq r \leq 1$. 

Notes
### MIMO Channels

- **Universal Code Design Criterion**

  - Similar to the MISO case: the worst-case channel aligns itself in the weakest directions afforded by a codeword pair difference matrix.
  
  - MIMO channel: \( n_{\text{min}} \) directions.
  
  - Code design criterion: maximize the
    
    \[
    \lambda_1 \lambda_2 \ldots \lambda_{n_{\text{min}}},
    \]
    
    with the \( n_{\text{min}} \) smallest singular values of the normalized codeword difference matrices.

  \[\rightarrow \text{For } n_t \leq n_r, \text{ this is equivalent to the determinant criterion.}\]

  - As for parallel channels, an approximately universal code is characterized by
    
    \[
    |\lambda_1 \lambda_2 \ldots \lambda_{n_{\text{min}}}|^2/n_{\text{min}} > \frac{1}{n_{\text{min}}2^R/n_{\text{min}}}.\]

  - Relaxed criterion (sufficient to guarantee that a code achieves the optimal DMT):
    
    \[
    |\lambda_1 \lambda_2 \ldots \lambda_{n_{\text{min}}}|^2/n_{\text{min}} > c \frac{1}{n_{\text{min}}2^R/n_{\text{min}}}, \text{ for some constant } c > 0.
    \]

### MIMO Channels

- **Interesting Conclusion**

  A code that satisfies the relaxed criterion for an \( n_t \times n_t \) MIMO channel is approximately universal for an \( n_t \times n_r \) MIMO channel for every value of the number of receive antennas \( n_r \).

Using the relaxed criterion to verify the D-BLAST performance

- Normalized codeword difference matrix
  
  \[
  D = \begin{bmatrix}
  0 & d_{(1)}^{(1)} & d_{(2)}^{(1)} \\
  d_{(1)}^{(2)} & d_{(2)}^{(1)} & 0
  \end{bmatrix},
  \]

  with the normalized pairwise difference codewords \((d_{(1)}^{(i)}, d_{(2)}^{(i)})\) of a universal parallel code satisfying

  \[
  |d_{(1)}^{(i)} d_{(2)}^{(i)}| > \frac{1}{42^R}, \quad i = 1, 2.
  \]

  - Then,
    
    \[\lambda_1^2 \lambda_2^2 = \det(DD^*) = |d_{(1)}^{(1)} d_{(1)}^{(2)}|^2 + |d_{(2)}^{(1)} d_{(2)}^{(2)}|^2 + |d_{(2)}^{(2)} d_{(1)}^{(1)}|^2 > \frac{1}{42^R}.\]

  \[\rightarrow \text{D-BLAST is optimal!}\]