Permutation Codes: Achieving the Diversity Multiplexing Tradeoff

Pramod Viswanath

Joint work with Saurabh Tavildar

University of Illinois at Urbana-Champaign

April 27, 2004
Parallel Channel Model

\[ y_1 = h_1 x_1 + w_1 \]
\[ y_2 = h_2 x_2 + w_2 \]
\[ y_3 = h_3 x_3 + w_3 \]
Parallel Channel Model

\[ y_1 = h_1 x_1 + w_1 \]

\[ y_2 = h_2 x_2 + w_2 \]

\[ y_3 = h_3 x_3 + w_3 \]

*Time Diversity:* coding over time
Parallel Channel Model

\[
\begin{align*}
  y_1 &= h_1 x_1 + w_1 \\
  y_2 &= h_2 x_2 + w_2 \\
  y_3 &= h_3 x_3 + w_3
\end{align*}
\]

- **Time Diversity**: coding over time
- **Frequency Diversity**: coding over OFDM symbols
Parallel Channel Model

\[
\begin{align*}
    y_1 &= h_1 x_1 + w_1 \\
    y_2 &= h_2 x_2 + w_2 \\
    y_3 &= h_3 x_3 + w_3
\end{align*}
\]

- **Time Diversity**: coding over time
- **Frequency Diversity**: coding over OFDM symbols
- **Antenna Diversity**: coding for MIMO channel: D-BLAST
Communication Over Slow Fading Channel

- $L$ parallel sub-channels

- **Slow fading:** $h_1, \ldots, h_L$ random, but fixed over time

- **Correlated fading:** $h_1, \ldots, h_L$ jointly distributed

- **Coherent communication:** $h_1, \ldots, h_L$ known to the receiver
Communication Over Slow Fading Channel

- $L$ parallel sub-channels

- **Slow fading:** $h_1, \ldots, h_L$ random, but fixed over time

- **Correlated fading:** $h_1, \ldots, h_L$ jointly distributed

- **Coherent communication:** $h_1, \ldots, h_L$ known to the receiver

- Focus in this talk:
  
  Short block-length communication at high **SNR**
Rate and Probability of Error

◇ Tradeoff between rate $R$, and probability of error $P_e$

◇ **Outage:** Given a rate and SNR:

$$\min_{P_e} P \left( \{ h \mid I(x; y \mid h) < R \} \right)$$

◇ At high SNR, $P_e \geq$ probability of outage

◇ **Compound channel result:** The outage probability can be achieved with long block-length codes.

◇ Short block length **space time** codes - delay diversity, rotated QAM codes, space time trellis codes, space time turbo codes.
$R$ and $P_e: a Coarser Scaling$

- Coarser question: Zheng and Tse formulation
  - Rate $= r \log(\text{SNR})$
  - Probability of error $= \frac{1}{\text{SNR}^d}$

- Given $r$, find maximal $d = d^*(r)$

- i.i.d. Rayleigh fading MIMO channel
  - outage probability achieved by short block length Gaussian code.
Characterization of Channels in Outage

- **Outage**: Input distribution can be taken as i.i.d. Gaussian

\[
\text{outage} = \left\{ h \mid \sum_{i=1}^{L} \log \left( 1 + |h_i|^2 \text{SNR} \right) < r \log(\text{SNR}) \right\}
\]

- Outage condition independent of distribution on \( h \)

- **Outage curve**: \( \mathbb{P} (\text{outage}) = \text{SNR}^{-d_{\text{out}}(r)} \)

- \( \mathbb{P} (\text{outage}) \leq \mathbb{P}_e \Rightarrow d^*(r) \leq d_{\text{out}}(r) \)
Main Result

Setting:

- coherent communication over short block length at high SNR
- measure performance in terms of tradeoff curve
Main Result

Setting:
- coherent communication over short block length at high SNR
- measure performance in terms of tradeoff curve

Main Result:
- Space only codes universally achieve the outage curve
  * engineering value: code is robust to channel modeling errors
- Simple deterministic construction of permutation codes
\[ \mathbb{P}(\text{outage}) \leq \mathbb{P}_e \leq \mathbb{P}(\text{outage}) + \mathbb{P}(	ext{error}|\text{no-outage}) \]

- We want the second term to decay exponentially in SNR
  - Look at the union bound
  - Each pairwise error should decay exponentially in SNR

- Let \( d = x(1) - x(2) \) be the difference codeword (space-only)

\[
\mathbb{P}_e (x(1) \rightarrow x(2) | h) \leq \exp \left( - \sum_{i=1}^{L} |h_i|^2 |d_i|^2 \right)
\]
Code Construction Criterion

◇ Worst case analysis over all realizations not in outage:

\[
\min_{h \notin \text{outage}} \left[ \sum_{i=1}^{L} |h_i|^2 |d_i|^2 \right]
\]

\[
h : \sum_{i=1}^{L} \log \left( 1 + \text{SNR}|h_i|^2 \right) > r \log \text{SNR}
\]
Code Construction Criterion

◊ Worst case analysis over all realizations not in outage:

\[
\min_{h \notin \text{outage}} \left[ \sum_{i=1}^{L} |h_i|^2 |d_i|^2 \right]
\]

\[
h : \sum_{i=1}^{L} \log (1 + \text{SNR}|h_i|^2) > r \log \text{SNR}
\]

◊ Outage condition independent of the distribution of \( h \)
  
  – Code construction is universal
  
  – Viewpoint taken by Köse and Wesel 03

◊ Contrast with the traditional analysis: average over the channel statistics
Code Construction Criterion

◇ **Worst case analysis** over all realizations not in outage:

\[
\min_{\mathbf{h} \not\in \text{outage}} \left[ \sum_{i=1}^{L} |h_i|^2 |d_i|^2 \right] > 1
\]

\[
\mathbf{h} : \sum_{i=1}^{L} \log (1 + \text{SNR}|h_i|^2) > r \log \text{SNR}
\]

◇ **Related** to the water-pouring problem

– Constraint function and objective function reversed

\[
|h_i|^2 = \left( \frac{1}{\lambda} - |d_i|^2 \right)^+
\]
Code Construction Criterion

◊ Turns out to be the **product distance**:

\[ |d_1|^2|d_2|^2 \cdots |d_L|^2 \geq \frac{SNR^L}{SNR^r} \]

◊ **Same** as the traditional criterion for i.i.d. Rayleigh fading
Tilted-QAM Codes

- Optimal Rotation Angle: $\theta = \frac{1}{2} \tan^{-1} (2)$
- Maximizes minimum product distance (Boutros and Viterbo ’98, Dayal and Varanasi, ’03)
- Achieves diversity-multiplexing tradeoff for Rayleigh fading channel (Yao and Wornell ’03)
**Tilted-QAM Codes**

- ML Decoding is not easy
- Encoding sensitive to rotation angle accuracy
- Hard to generalize beyond 2 diversity paths
Revisit Code Design Criterion

\[ |d_1|^2|d_2|^2 \cdots |d_L|^2 \geq \frac{\text{SNR}^L}{\text{SNR}^r} \]

- Nonzero product distance
  - Need each sub-channel to have all the information
  - So, alphabet size SNR^r
  - Can take it to be a QAM (Q) with SNR^r points (for each sub-channel)
Implications on Code Structure

- **Nonzero product distance**
  - Need each sub-channel to have **all the information**
  - So, alphabet size $\text{SNR}^r$
  - Can take it to be a $\text{QAM} (\mathbb{Q})$ with $\text{SNR}^r$ points

- **Mapping across sub-channels**
  - Each point in the QAM for any sub-channel should represent the entire codeword
  - So, code is $L - 1$ permutations of $\mathbb{Q}$
Repetition Coding

- the identity permutation

\[ L = 2, \text{ product distance} = \frac{\text{SNR}^2}{\text{SNR}^2 r} \]

- We want: product distance > \( \frac{\text{SNR}^2}{\text{SNR}^r} \)
Key Property of Permutations

- Two adjacent points should be mapped as far apart as possible
Random Permutation Codes

- Look at the ensemble of all permutation codes
  - huge number of permutations: \( (\text{SNR}^r!)^{L-1} \)
  - average product distance under appropriate measure

- There exists a permutation which satisfies the product distance

- Conclusion: A permutation code achieves universally the outage curve
An Example

- 16-point Permutation Code

- Minimum Product Distance: 0.4 (rotation code: 0.44)
Baseline for Complexity Analysis

◊ Complexity of QAM on a scalar channel
  – easy encoding
  – easy decoding
  – worst case complexity very small

◊ Ideal situation:

With $L$ parallel channels, complexity is $L$ times that of QAM
2 Sub-Channels

- Transmit $q \in \mathbb{Q}$ and $f(q) \in \mathbb{Q}$ over the two sub-channels

\[ q = \frac{\sqrt{\text{SNR}}}{\text{SNR}^{r/2}} (a + ib), \quad a, b \quad \text{integers} \]

\[ y_1 = h_1 \frac{\sqrt{\text{SNR}}}{\text{SNR}^{r/2}} (a + ib) + w_1 \]

\[ y_2 = h_2 \frac{\sqrt{\text{SNR}}}{\text{SNR}^{r/2}} f(a + ib) + w_2 \]
2-sub Channels

- Transmit \( q \in \mathbb{Q} \) and \( f(q) \in \mathbb{Q} \) over the two sub-channels

\[
q = \frac{\sqrt{\text{SNR}}}{\text{SNR}^{r/2}} (a + ib), \quad a, b \quad \text{integers}
\]

\[
y_1 = h_1 \frac{\sqrt{\text{SNR}}}{\text{SNR}^{r/2}} (a + ib) + w_1
\]

\[
y_2 = h_2 \frac{\sqrt{\text{SNR}}}{\text{SNR}^{r/2}} f(a + ib) + w_2
\]

- Look at permutations (\( \tilde{f} \)) of real and imaginary values

\[
\tilde{y}_1 = |h_1| \frac{\sqrt{\text{SNR}}}{\text{SNR}^{r/2}} (a + ib) + \tilde{w}_1
\]

\[
\tilde{y}_2 = |h_2| \frac{\sqrt{\text{SNR}}}{\text{SNR}^{r/2}} (\tilde{f}(a) + i\tilde{f}(b)) + \tilde{w}_2
\]
Effect of the Fading Channel

diamond Consider binary representation of integers $a$ and $b$
  - require $n = \frac{r}{2} \log_2 \text{SNR}$ bits

diamond Additive Gaussian noise very likely to move within neighboring integers
Effect of the Fading Channel

- Consider **binary** representation of integers $a$ and $b$
  
  - require $n = \frac{r}{2} \log_2 \text{SNR}$ bits

- Additive Gaussian noise very likely to move within neighboring integers

- Effect of the **multiplicative** channel: **distort the LSBs**

\[
|h_1| \frac{\sqrt{\text{SNR}}}{\text{SNR}^{r/2}} \approx 2^{-k_1} \quad \Rightarrow \quad k_1 \text{ LSBs of } a \text{ and } b \text{ lost}
\]

\[
|h_2| \frac{\sqrt{\text{SNR}}}{\text{SNR}^{r/2}} \approx 2^{-k_2} \quad \Rightarrow \quad k_2 \text{ LSBs } f(a) \text{ and } f(b) \text{ lost}
\]
Effect of the Fading Channel

- Consider binary representation of integers $a$ and $b$
  
  - require $n = \frac{r}{2} \log_2 \text{SNR}$ bits

- Additive Gaussian noise very likely to move to neighboring integers

- Effect of the multiplicative channel: distort the LSBs

\[
|h_1| \frac{\sqrt{\text{SNR}}}{\text{SNR}^{r/2}} \approx 2^{-k_1} \implies k_1 \text{ LSBs of } a \text{ and } b \text{ lost}
\]

\[
|h_2| \frac{\sqrt{\text{SNR}}}{\text{SNR}^{r/2}} \approx 2^{-k_2} \implies k_2 \text{ LSBs } f(a) \text{ and } f(b) \text{ lost}
\]

- No outage condition:

\[
|h_1||h_2| > \frac{\text{SNR}^{r/2}}{\text{SNR}} \implies k_1 + k_2 \leq n
\]
Bit Reversal Permutation

- **Bit reversal**: $\tilde{f}(a)$ is bit reversal of $a$

- **Encoding**: 
  - Same complexity as encoding a QAM

- **Decoding**: 
  - Use first sub-channel to determine MSBs of $a$ and $b$
  - Use second sub-channel to determine LSBs of $a$ and $b$
  - No outage condition means you recover all the bits
Product Distance and Bit Reversals

- **Product distance** not very good

- Example:

  \[
  a_1 = 1000000001 \\
  a_2 = 0111111110
  \]

  - distance \( a_1 - a_2 = 3 \)

  - Bit reversals are the **same**
Product Distance and Bit Reversals

- **Product distance** not very good

- Example:

  \[
  a_1 = 1000000001 \\
  a_2 = 0111111110
  \]

  - distance \( a_1 - a_2 = 3 \)
  - Bit reversals are the **same**

- **Modification** of bit reversal works

  - Alternate bit flipping
  - Insert boundaries to avoid carry-over problem
Arbitrary Fading MIMO Channel

\[ y = Hx + w \]

- Entries of \( H \) have a joint distribution.
  - slow fading, coherent
Universal code construction criterion

- Extend methodology to MIMO channel:

\[ \lambda_1 \lambda_2 \ldots \lambda_{\min(n_r,n_t)} \geq \frac{\text{SNR}^\text{min}(n_r,n_t)}{\text{SNR}} \]

- \( \lambda_1 \leq \ldots \leq \lambda_{n_t} \) are eigenvalues of \( DD^\dagger \), \( D := X(1) - X(2) \)

- Need this criterion for every pair of codewords

  - then code universally achieves the outage curve
Insights

- Universal code construction criterion:

\[ \lambda_1 \lambda_2 \ldots \lambda_{\min(n_r,n_t)} \geq \frac{\text{SNR}^{\min(n_r,n_t)}}{\text{SNR}^r} \]

- Compare with traditional criterion for i.i.d. Rayleigh fading (Tarokh, Seshadri and Calderbank, 98)

\[ \lambda_1 \lambda_2 \ldots \lambda_{n_t} \]

- Similar
  - Also the product of eigenvalues of \( D \)

- Different as well
  - only look at product of \( \min(n_r, n_t) \) smallest eigenvalues
Insights

- Universal code construction criterion:

\[ \lambda_1 \lambda_2 \ldots \lambda_{\min(n_r,n_t)} \geq \frac{\text{SNR}^{\min(n_r,n_t)}}{\text{SNR}^r} \]

- Compare with traditional criterion for i.i.d. Rayleigh fading

\[ \lambda_1 \lambda_2 \ldots \lambda_{n_t} \]

- Similar
  - Also the product of eigenvalues of D

- Different as well
  - only look at product of \( \min (n_r, n_t) \) smallest eigenvalues

- Subsumption: if a code does well for \( n_t \times n_r \) channel, then it will do well for \( n_t \times k \) channel if \( k \leq n_r \)
D-BLAST Architecture

- D-BLAST scheme for 2 transmit antennas:
  \[ X = \begin{bmatrix} 0 & a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 & 0 \end{bmatrix} \]

- Let \( D = X(0) - X(1) \)

\[ D = \begin{bmatrix} 0 & \tilde{a}_2 & \tilde{b}_2 & \tilde{c}_2 \\ \tilde{a}_1 & \tilde{b}_1 & \tilde{c}_1 & 0 \end{bmatrix} \]

- Universal criterion same as product distance criterion for each stream

\[ \det(DD^\dagger) \geq |\tilde{a}_1|^2|\tilde{a}_2|^2 \geq \frac{\text{SNR}^2}{\text{SNR}^r} \]
Using Permutation Codes

- D-BLAST
  - converts MIMO channel to a parallel channel
  - can use permutation codes

- D-BLAST preserves mutual information
  - outage curve can be achieved
  - small bootstrap overhead

- Encoding/decoding complexity same as permutation codes
Universal Codes for the $2 \times 2$ channel

- Tilted QAM codes for the $2 \times 2$ channel
- Satisfies universal code design criterion
- So it trades off universally
  - Strengthening of earlier result for iid Rayleigh channel (Yao and Wornell, 03)
- Decoding complexity high
- No easy generalization
Universal Codes for the MIMO Channel

- **Open question**: construct universal codes for general MIMO channel

- **Compound channel** result guarantees they exist for long block lengths

- For short block lengths, simpler but non-universal, codes may exist
V-Blast with ML decoding

diamond Send QAM independently across antennas (space only code)
Tradeoff curve for i.i.d. Rayleigh fading

Diamond: space only V-Blast optimal for \( r \geq 1 \) on 2 × 2 Rayleigh channel
Tradeoff curve for i.i.d. Rayleigh fading

- **Diversity**
- **Multiplexing Rate**

Optimal tradeoff

V-BLAST with ML decoding

- Space only V-BLAST optimal for $r \geq n - 1$ on $n \times n$ Rayleigh channel