6.962 Graduate Seminar in Communications

Week 8

Space-Time Modulation for Unknown Fading

- Bertrand M. Hochwald and Thomas L. Marzetta

presenter: Huan Yao

November 1, 2000

Motivation

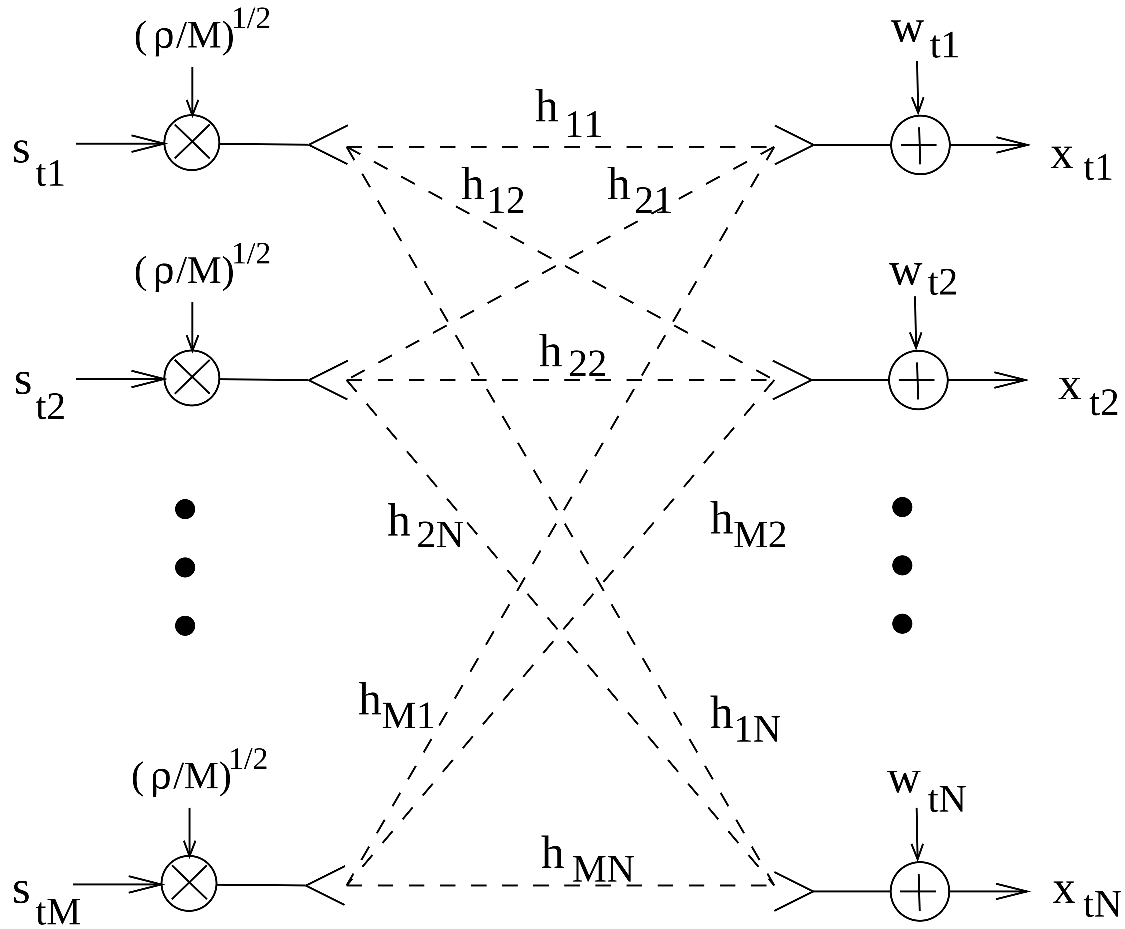
Coherent

- Assume channel is known.
- Need to train / track channel coefficients.
- More antennas, more coefficients.

Non-Coherent

- Channel unknown.
- Non-explicit channel estimation.
- A more general method.

System Model



- M transmitter antennas.
- N receiver antennas.
- Rayleigh flat-fading.
- Piecewise constant for T.
- Additive White Gaussian Noise.

Matrix Representation

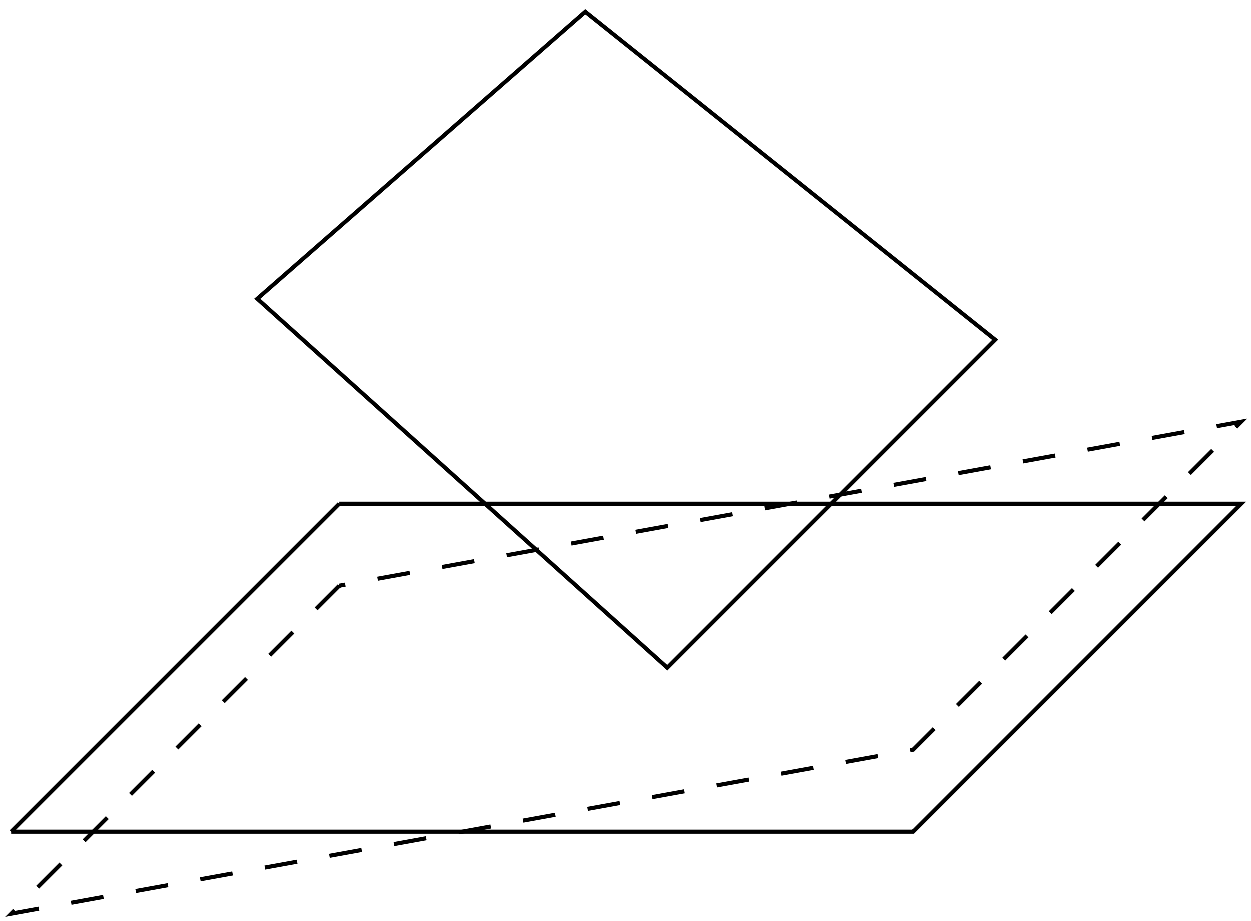
$$X = \sqrt{\frac{\rho}{M}} SH + W.$$

$$S = \begin{bmatrix} s_{11} & \cdots & s_{1M} \\ \vdots & \ddots & \vdots \\ s_{T1} & \cdots & s_{TM} \end{bmatrix}, \quad X = \begin{bmatrix} x_{11} & \cdots & x_{1N} \\ \vdots & \ddots & \vdots \\ x_{T1} & \cdots & x_{TN} \end{bmatrix},$$

$$H = \begin{bmatrix} h_{11} & \cdots & h_{1N} \\ \vdots & \ddots & \vdots \\ h_{M1} & \cdots & h_{MN} \end{bmatrix}, \quad W = \begin{bmatrix} w_{11} & \cdots & w_{1N} \\ \vdots & \ddots & \vdots \\ w_{T1} & \cdots & w_{TN} \end{bmatrix},$$

$$\frac{1}{M} \sum_{m=1}^{M} E(|s_{tm}|^2) = 1, \text{ for } \; t = 1, \ldots, T.$$

A Subspace Picture



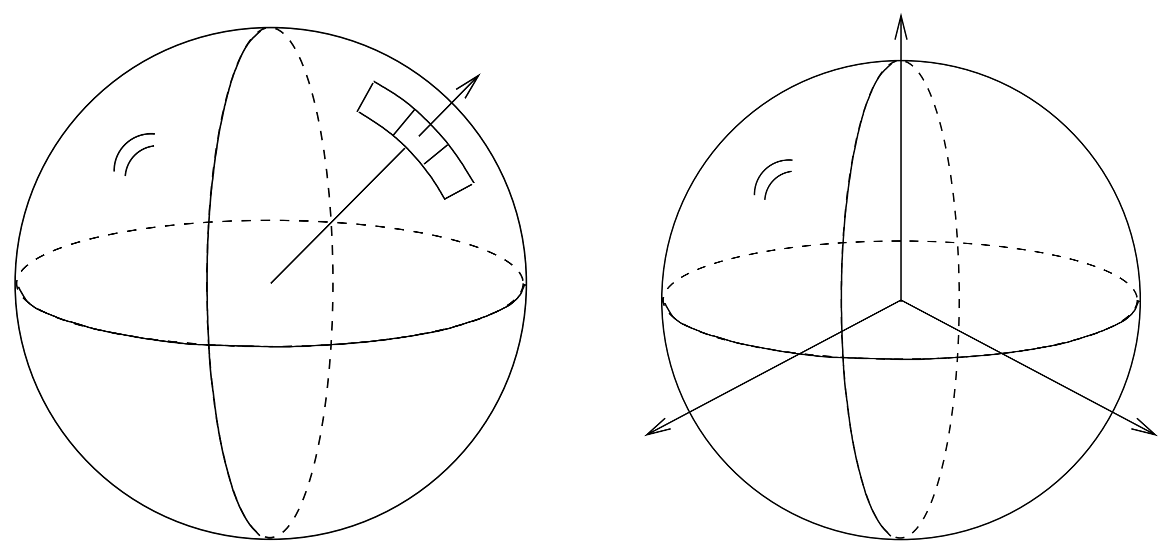
Channel Input/Output Probability

$$p(X|S) = \frac{\exp\left(-\mathsf{tr}\left\{\left[I_T + (\rho/M)SS^\dagger\right]^{-1} XX^\dagger\right\}\right)}{\pi^{TN}\det^N\left[I_T + (\rho/M)SS^\dagger\right]}$$

$$\begin{array}{rcl} H & \sim & I \\ SH & \sim & SS^\dagger \\ X = \sqrt{\frac{\rho}{M}}SH + W & \sim & I_T + (\rho/M)SS^\dagger \end{array}$$

- It is $SS^\dagger$ that matters in design.
- Don't need $M > T$.
- $p(X|S\Phi^\dagger) = p(X|S) \quad \forall$ unitary Φ

Isotropically Distributed (i.d.) unit vectors and unitary matrices



rotational symmetry: $p(\Phi) = p(\Theta^\dagger \Phi), \forall \Theta : \Theta^\dagger \Theta = I$

Capacity Achieving Distribution

$$p(X|S) = \frac{\exp\left(-\mathrm{tr}\left\{\left[I_T + (\rho/M)SS^\dagger\right]^{-1} XX^\dagger\right\}\right)}{\pi^{TN}\det^N\left[I_T + (\rho/M)SS^\dagger\right]}$$

$$I(X;S) = E_{X,S}\left\{\log\left(\frac{p(X|S)}{E_S\{p(X|S)\}}\right)\right\}$$

$$C(X;S) = \sup_{p(S)} I(X;S)$$

Theorem: The signal matrix that achieves capacity can always be factored as $S = \Phi V$, where Φ is an $T \times M$ isotropically distributed unitary matrix, and V is an independent $M \times M$ real, nonnegative, diagonal matrix.

A Quick Proof

First, show $S = \Phi V$ by performing SVD on S.

$$S = \Phi V \Psi^{\dagger}$$

$$SS^{\dagger} = (\Phi V \Psi^{\dagger})(\Psi V \Phi^{\dagger}) = (\Phi V)(\Phi V)^{\dagger}$$

Second, show Φ is i.d. by rotating S.

$$S' = \Theta S = \Theta \Phi V$$

Θ is i.d. $\implies \Theta\Phi$ is i.d.

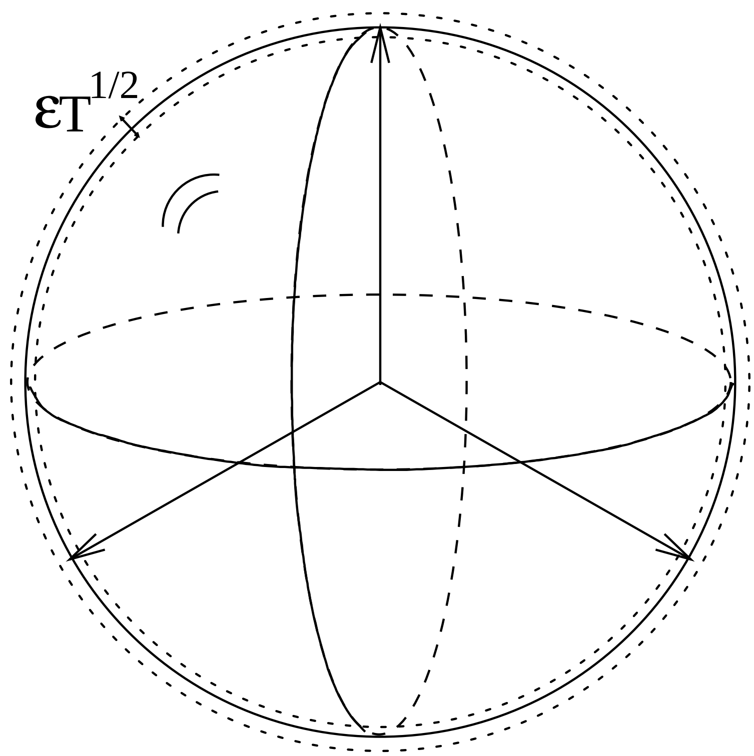
Capacity Achieving Distribution for V

for fixed M, as $T \to \infty$, that $V \to \sqrt{T} I_M$.

<u>Intuition:</u>

• No preference to any antenna.

• Since average energy is T, in the high dimension limit, most vectors lie in a thin shell of thickness $\epsilon\sqrt{T}$ and radius $\sqrt{T}$.



Deterministic Codebook Design

We want to design

$$S_1 = \Phi_1 V_1, S_2 = \Phi_2 V_2, \ldots, S_L = \Phi_L V_L,$$

where $L = 2^{RT}$ and $\Phi_l^\dagger \Phi_l = I$.

We can pick $V_1 = V_2 = \cdots = V_L = \sqrt{T}$.

We still need to find $\Phi_1, \Phi_2, \cdots, \Phi_L$.

Maximum Likelihood Non-Coherent Decoder

$$
\begin{aligned}
\Phi_{ml} &= \mathsf{argmax}_{\Phi_l=\Phi_1,\ldots,\Phi_L} tr\{X^\dagger \Phi_l \Phi_l^\dagger X\} \\
&= \mathsf{argmax}_{\Phi_l=\Phi_1,\ldots,\Phi_L} \sum_{i,j} \left| [\Phi^\dagger X]_{ij} \right|^2
\end{aligned}
$$

Assume lth signal is transmitted.

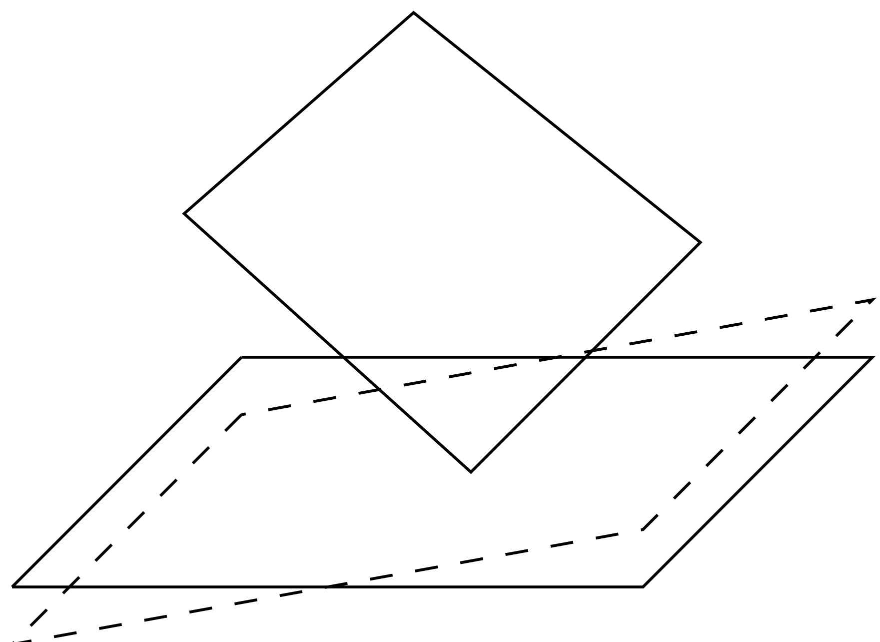
$$
\begin{aligned}
X &= \sqrt{\frac{\rho}{M}} S_l H + W \\
&= \sqrt{\frac{\rho T}{M}} \Phi_l H + \Phi_l \Phi_l^\dagger W + (W - \Phi_l \Phi_l^\dagger W) \\
&= \sqrt{\frac{\rho T}{M}} \Phi_l H + W_l^{in} + W_l^{out} \\
\Phi_l^\dagger X &= \sqrt{\frac{\rho T}{M}} H + \Phi_l^\dagger W_l^{in}
\end{aligned}
$$

The ML decoder $\Phi_{ml} = \text{argmax}_{\Phi_l} tr\{X^\dagger \Phi_l \Phi_l^\dagger X\}$

- Maximize energy in $\Phi_l^\dagger X$.

- Minimize energy in W_l^{out}.

- Find nearest valid subspace.



Channel Estimation

Channel estimation falls right out of the ML non-coherent decoder.

Assume lth signal is transmitted.

$$\begin{aligned} \hat{H}_l &= \left(\frac{\rho T}{M}\right)^{(-\frac{1}{2})} \Phi_l^{\dagger} X \\ &= H + \left(\frac{\rho T}{M}\right)^{(-\frac{1}{2})} \Phi_l^{\dagger} W_l^{in} \end{aligned}$$

This ML non-coherent decoder has an equivalent interpretation as a generalized likelihood ratio test (GLRT).

$$\Phi_{ml} = \mathsf{argmax}_{\Phi_l} \mathsf{argmax}_{H_l} \; p(X|\Phi_l, H_l)$$

Two-Signal Probability of Error

$$P_e \leq \frac{1}{2} \prod_{m=1}^{M} \left[\frac{1}{1 + \frac{(\rho T/M)^2 (1 - d_m^2)}{4(1+\rho T/M)}} \right]^N$$

where $1 \geq d_1 \geq \ldots \geq d_M \geq 0$ are the singular values of the $M \times M$ correlation matrix $\Phi_2^\dagger \Phi_1$.

- P_e is the lowest when $d_1 = \cdots = d_M = 0$.
- P_e is the highest when $d_1 = \cdots = d_M = 1$.

The $d_i' s$ corresponds to the cosines of a set of $principle$ $angles$ defined between the two subspaces Φ_1 and Φ_2.

Principle Angles

A set of *principle angles*, $\theta_1, \cdots, \theta_M$, can be defined between two M-dimensional subspaces Φ_1 and Φ_2 by

- Pick vectors $u_1 \in \Phi_1$, $v_1 \in \Phi_2$, s.t., $u_1^\dagger v_1$ is maximized, i.e., having the smallest angle in between. $u_1^\dagger v_1 = d_1 = \cos\theta_1$.

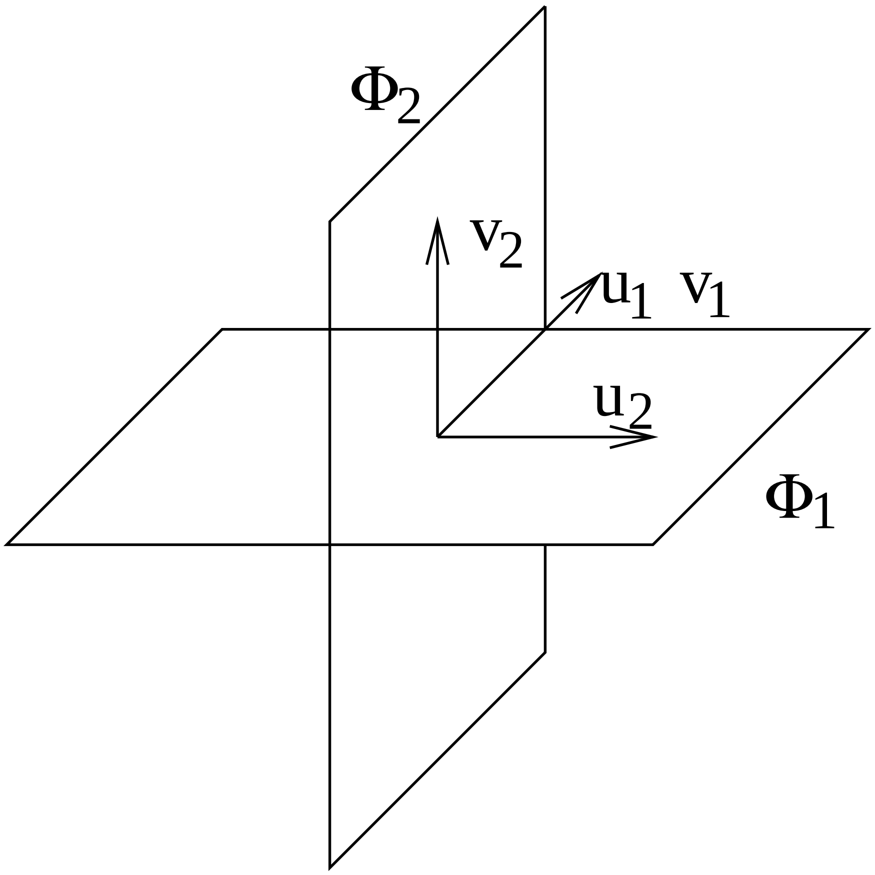
- Pick vectors $u_2 \in \Phi_1$, $v_2 \in \Phi_2$, s.t., $u_2 \perp u_1$ and $v_2 \perp v_1$, and $u_2^\dagger v_2$ is maximized.
  $u_2^\dagger v_2 = d_2 = \cos\theta_2$.

- $\vdots$

- Pick vectors $u_k \in \Phi_1$, $v_k \in \Phi_2$, s.t., $u_k \perp u_j$ and $v_k \perp v_j$ for $j = 1, \cdots, k-1$, and $u_k^\dagger v_k$ is maximized.
  $u_k^\dagger v_k = d_k = \cos\theta_k$.

- Continue until $k = M$.

Principle Angels Example



$$d_1 = 1 \quad d_2 = 0$$

$$\theta_1 = 0^\circ \quad \theta_2 = 90^\circ$$

Iterative Design Method for $M = 1$

We want to place L vectors in a T-dimensional space with maximum separation between the nearest pair.

1. Compute d_{max}, the maximum of the magnitudes of all $L(L-1)/2$ distinct inner products, and choose a pair of vectors whose inner product is d_{max}.

2. "Separate" the pair by moving each vector a small amount in opposite directions along the difference vector between the pair.

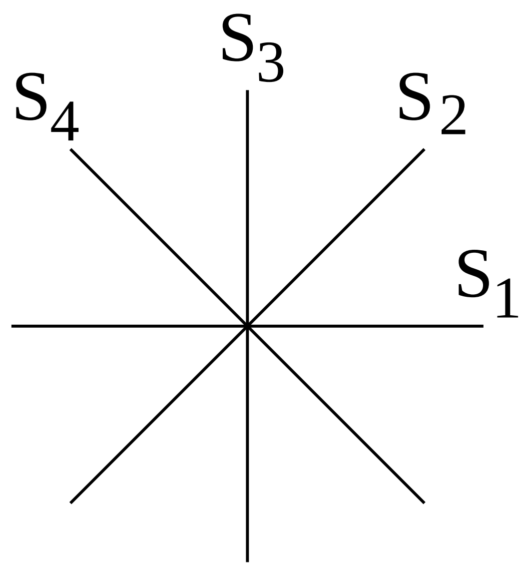
3. Renormalize the pair, if needed.

4. Repeat Steps 1-3 until d_{max} stops decreasing.

Deterministic Codebook Design Example : $M = 1$

$T = 2 \quad L = 2^2 = 4$
$d_{max} = 0.707 \quad \theta_{min} = 45°$

$T = 5 \quad L = 2^5 = 32$
$d_{max} = 0.515 \quad \theta_{min} = 59°$



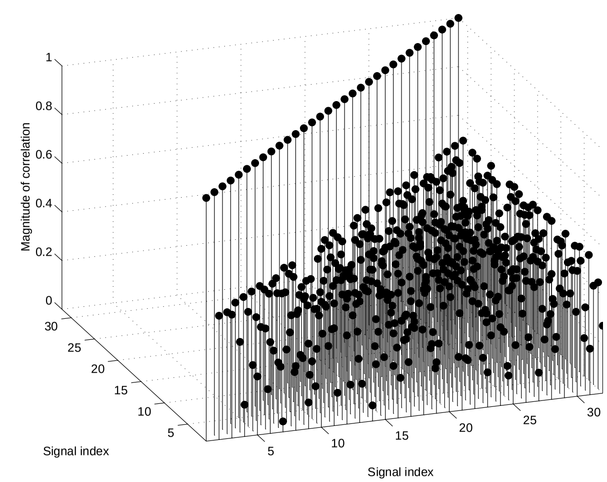


Figure : Magnitudes of correlations between $\Phi_1, \ldots, \Phi_{32}$ for $T = 5$. The diagonal entries with value 1.0 represent each signal correlated with itself.

Deterministic Codebook Design for $M > 1$

For each pair of Φ_l and $\Phi_{l'}$, there is a whole set of $(d_1, \ldots, d_M)$. To determine which pair is the "closest", we can use

- The Chernoff Bound :

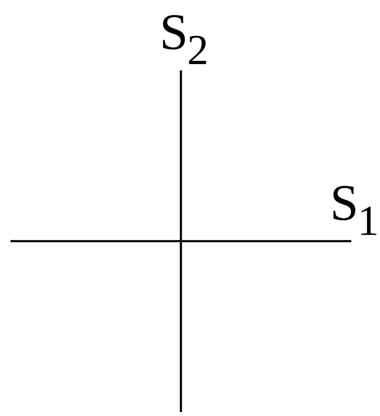
$$P_e \leq \frac{1}{2} \prod_{m=1}^{M} \left[\frac{1}{1 + \frac{(\rho T/M)^2 (1 - d_m^2)}{4(1 + \rho T/M)}} \right]^N$$

- The Frobenius Norm:

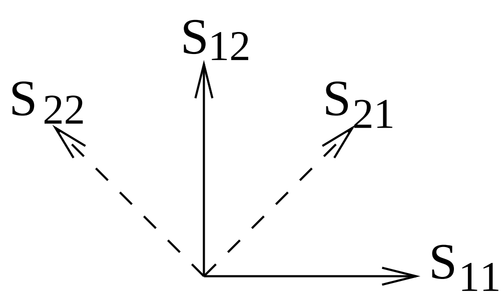
$$\left\| \Phi_l^\dagger \Phi_{l'} \right\| = \sqrt{\frac{d_1^2 + \ldots + d_M^2}{M}}$$

Example: $T = 2$, $L = 2$, $M = 1$ vs. $M = 2$

$M = 1$: Two orthogonal 1-D subspaces,
$d_1 = 0$, $\theta_1 = 90^\circ$



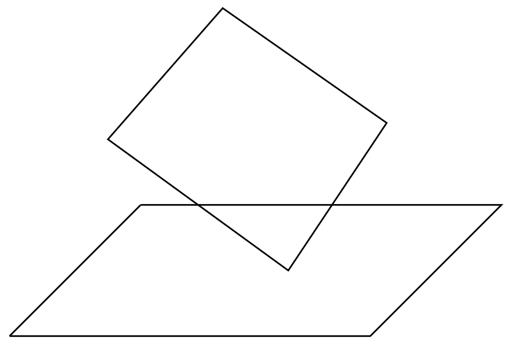
$M = 2$: Two 2-D subspaces entirely overlapping,
$d_1 = d_2 = 1$, $\theta_1 = \theta_2 = 0^\circ$



We don't want to use $M = T$.

Example: $T = 4$, $L = 2$, $M = 2$ vs. $M = 3$

$M = 2:$ $d_1 = d_2 = 0$, $\theta_1 = \theta_2 = 90^\circ$.



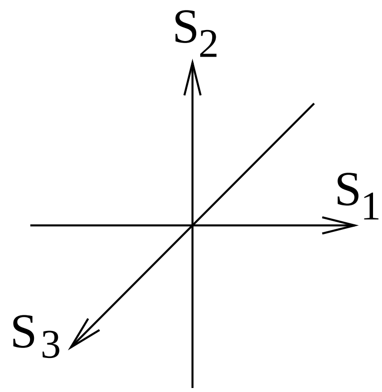
$M = 3:$ The intersection of two 3-D subspaces in 4-D space is a 2-D plane. $3 + 3 - 4 = 2$

$$d_1 = d_2 = 1 \qquad d_3 = 0$$
$$\theta_1 = \theta_2 = 0^\circ \quad \theta_3 = 90^\circ$$

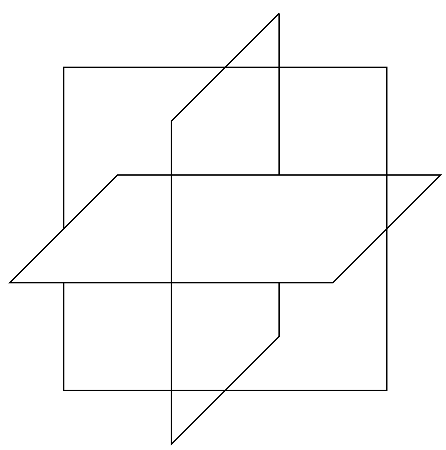
We don't want to use $M > \frac{T}{2}$.

Example: $T = 3$, $L = 3$, $M = 1$ vs. $M = 2$

$M = 1$: Between every pair, $d_1 = 0$ and $\theta_1 = 90°$.



$M = 2$: $d_1 = 1$ and $d_2 = 0$, $\theta_1 = 0°$ and $\theta_2 = 90°$.



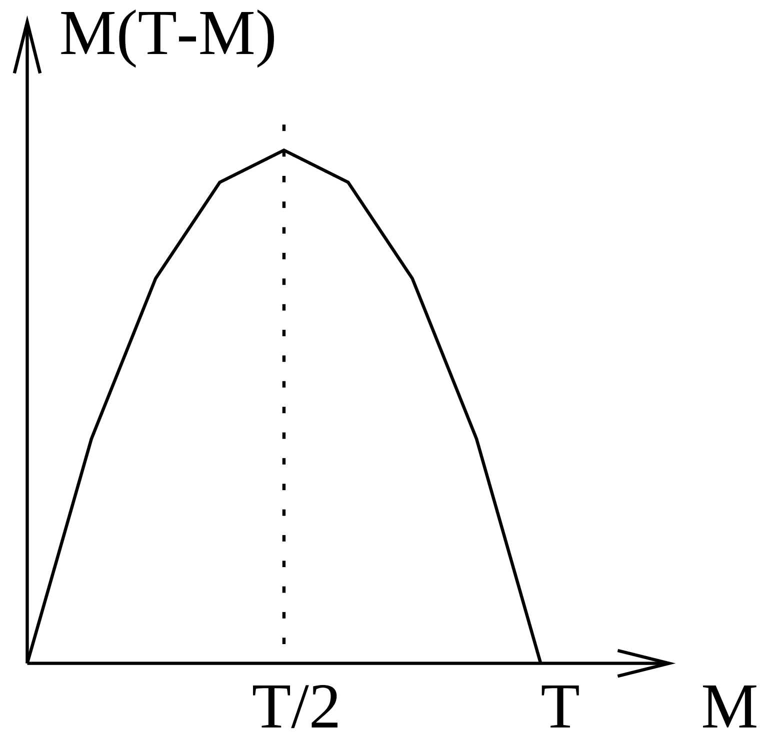
They have the same Chernoff's bound, equally good.

What M is Optimal?

We want to maximize $M(T-M)$. $M = \lfloor \frac{T}{2} \rfloor$ is optimal.

Interpretation: (training view)

- Need time M to solve for channel coefficients.
- Has time $(T - M)$ for data transmission.
- Capacity $\propto M$ when channel is known.



Systematic Design of Unitary Space-Time Codebook

How do we design good codebook efficiently when $L = 2^{RT}$ is large?

$$\Phi_l = \Theta^{l-1}\Phi_1,\ l = 1, \cdots, L$$

where Θ is a $T \times T$ diagonal unitary matrix such that $\Theta^L = I_T$, and Φ_1 is a unitary matrix, $\Phi_1^\dagger \Phi_1 = I_T$.

- Any Φ_l is unitary.

$$\Phi_l^\dagger \Phi_l = \Phi_1^\dagger \Theta^{-(l-1)} \Theta^{l-1} \Phi_1 = \Phi_1^\dagger \Phi_1 = I_T$$

- The correlation matrix

$$\Phi_i^\dagger \Phi_j = \Phi_1^\dagger \Theta^{-(i-1)} \Theta^{j-1} \Phi_1 = \Phi_1^\dagger \Theta^{(j-i) \text{ mod } L} \Phi_1$$

- For every matrix Φ_l, the set of correlation matrices it forms with the others are

$$\{\Phi_1^\dagger \Theta \Phi_1, \Phi_1^\dagger \Theta^2 \Phi_1, \cdots, \Phi_1^\dagger \Theta^{L-1} \Phi_1\}$$

Design of Θ

$$\Theta^L = I_T \Longrightarrow [\Theta]_{tt} = e^{i\frac{2\pi}{L}u_t},\ t = 1, \cdots, T,$$

where the u_t are integers, $0 \leq u_1, \cdots, u_T \leq L - 1$.

We only need to search over $\begin{pmatrix} T + L - 1 \\ T \end{pmatrix}$ sets of $\{u_t\}$ to minimize a cost criterion, such as Chernoff's bound or Frobenius norm.

example : For $T = 4$ and $L = 4$

$$\begin{pmatrix} T + L - 1 \\ T \end{pmatrix} = \begin{pmatrix} 7 \\ 4 \end{pmatrix} = \frac{7 \cdot 6 \cdot 5 \cdot 4}{4 \cdot 3 \cdot 2 \cdot 1} = 35$$

- Choice of Φ_1 is less critical.

Systematic Design Example

$T = 4$, $M = 2$, $L = 4$, and $\Phi_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}$.

Criterion: Minimize the worst case Frobenius norm

$$\delta = \max_{1 \le l < l' \le L} \left\| \Phi_l^\dagger \Phi_{l'} \right\| = \max_{1 \le l < l' \le L} \sqrt{\frac{d_1^2 + \ldots + d_M^2}{M}}$$

Result: $\{u_t\} = 0, 0, 1, 2$, $\Theta = \text{diag}(1, 1, i, -1)$.

	Φ_2	Φ_3	Φ_4
	$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ i & i \\ -1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & -1 \\ 1 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -i & -i \\ -1 & 1 \end{bmatrix}$
$d_1(\theta_1)$	$\frac{1}{\sqrt{2}}(45^\circ)$	$1(0^\circ)$	$\frac{1}{\sqrt{2}}(45^\circ)$
$d_2(\theta_2)$	$0(90^\circ)$	$0(90^\circ)$	$0(90^\circ)$
$\left\| \Phi_l^\dagger \Phi_1 \right\|$	$1/2$	$\frac{1}{\sqrt{2}}$	$1/2$

If I use a different $\Phi_1 \ldots$

$T = 4$, $M = 2$, $L = 4$, and $\Phi_1 = \begin{bmatrix} 1 & 1 \\ 1 & i \\ 1 & -1 \\ 1 & -i \end{bmatrix}$.

Result: $\{u_t\} = 0, 1, 2, 3$, $\Theta =$ diag$(1, i, -1, -i)$.

$$\begin{array}{cccc} & \Phi_2 & \Phi_3 & \Phi_4 \\ & \begin{bmatrix} 1 & 1 \\ i & -1 \\ -1 & 1 \\ -i & -1 \end{bmatrix} & \begin{bmatrix} 1 & 1 \\ -1 & -i \\ 1 & -1 \\ -1 & i \end{bmatrix} & \begin{bmatrix} 1 & 1 \\ -i & 1 \\ -1 & 1 \\ i & 1 \end{bmatrix} \\ d_1(\theta_1) & 1(0^\circ) & 0(90^\circ) & 1(0^\circ) \\ d_2(\theta_2) & 0(90^\circ) & 0(90^\circ) & 0(90^\circ) \\ \left\| \Phi_l^\dagger \Phi_1 \right\| & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{array}$$

4×4 FFT matrix $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix}$

Summary

- Described a non-coherent model for a multiple antenna system, assuming flat-fading and piecewise constant channel, $X = \sqrt{\frac{\rho}{M}} SH + W$.

- Structure of capacity achieving signal, $S = \Phi V$.

- Maximum likelihood decoder, $\Phi_{ml} = \text{argmax}_{\Phi_l} tr\{X^\dagger \Phi_l \Phi_l^\dagger X\}$.

- Two signal probability of error, P_e related to $d_m = \cos\theta_m$.

- Design examples with $M = 1$ and $M > 1$.

- $M = \lfloor \frac{T}{2} \rfloor$ is optimal.

- Systematic design method, $\Phi_l = \Theta^{l-1}\Phi_1$.