Modern Wireless Networks

MIMO

UNIVERSITY AT ALBANY
State University of New York

IECE 574– Spring 2021

Prof. Dola Saha

SISO

$$y = \sqrt{\rho} \cdot h \cdot x + z$$

- y: Received Signal
- $\sqrt{\rho}$: Average SNR
- h: Channel Fading Coefficient
- x: Transmitted data with unity mean expected power
- z: independent, complex additive white Gaussian noise (AWGN)

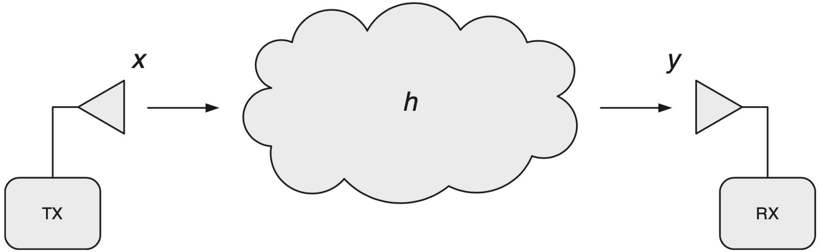
Capacity $$C(\text{bps/Hz}) = \log_2(1 + \rho \cdot |h|^2)$$

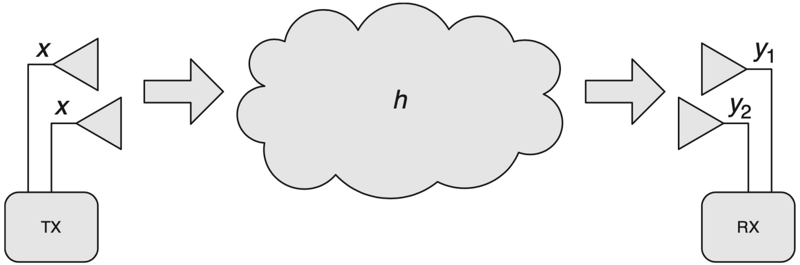
MIMO System

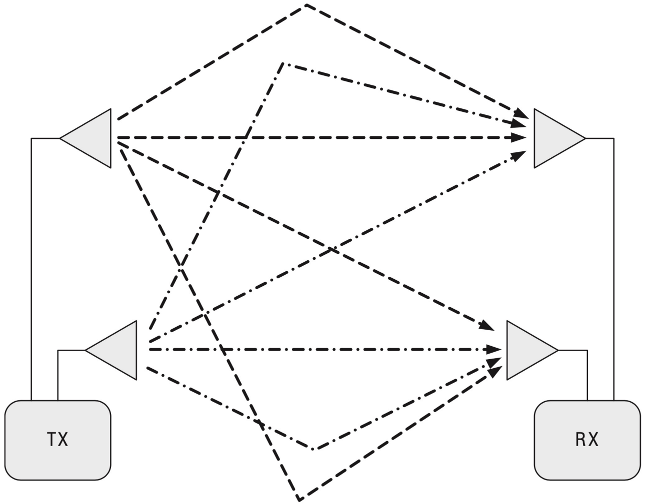
$$y_1 = \sqrt{\rho} \cdot h_1 \cdot x + z_1$$

$$y_2 = \sqrt{\rho} \cdot h_2 \cdot x + z_2$$

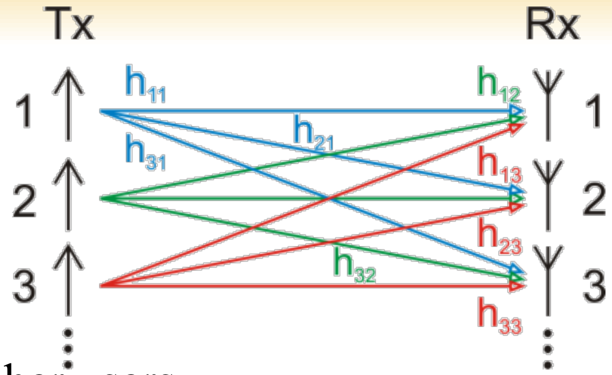
$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \sqrt{\rho} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} x + \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$







Categories of Multiple Antenna Tx & Rx



- Spatial Diversity
  - a number of different versions of the signal to be Tx/Rx
  - provides resilience against fading
- Interference suppression
  - uses the spatial dimensions to reject interference from other users
  - through the physical antenna gain pattern or through other forms of array processing such as linear precoding, postcoding, or interference cancellation
- Spatial multiplexing
  - allows multiple independent streams of data to be sent simultaneously in the same bandwidth, and hence is useful primarily for increasing the data rate

Spatial Diversity – Array Gain

- Coherently combines energy of each antenna (*channels can be correlated if LOS and closely spaced antenna*)
- Noise is uncorrelated and do not add coherently
- In correlated flat fading channel, received SNR increases linearly with the number of receive antennas, N_r

$y_i = h_i x + n_i = hx + n_i$, h is correlated flat fading channel

SNR at antenna i is $\gamma_i = |h^2|/\sigma^2$

Resulting Signal from all antennas $y = \sum_{i=1}^{N_r} y_i = N_r hx + \sum_{i=1}^{N_r} n_i$

Combined SNR is $\gamma = {|N_r h|^2}/{N_r \sigma^2} = \frac{N_r |h^2|}{\sigma^2} = N_r \gamma_i$

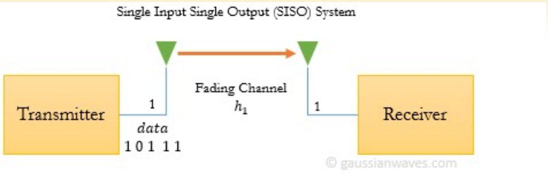
Spatial Diversity – Diversity Gain

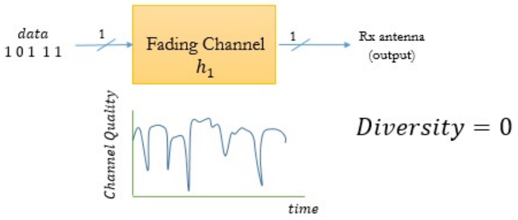
- Channel varies over space
- rms angular spread of a channel = θ_{rms}= statistical distribution of the angle of the arriving energy
- Dual of angular spread is coherence distance, D_c
- A coherence distance of d means that any physical positions separated by d have an essentially uncorrelated received signal amplitude and phase
- $D_C \approx .2\lambda/\theta_{rms}$, in Rayleigh fading, $D_C \approx 9\lambda/16\pi$
- coherence distance increases with the carrier wavelength λ, so higher-frequency systems have shorter coherence distances

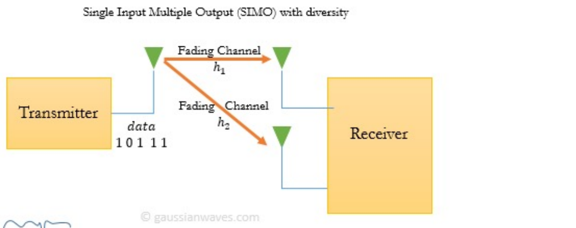
Spatial Diversity – Diversity Gain

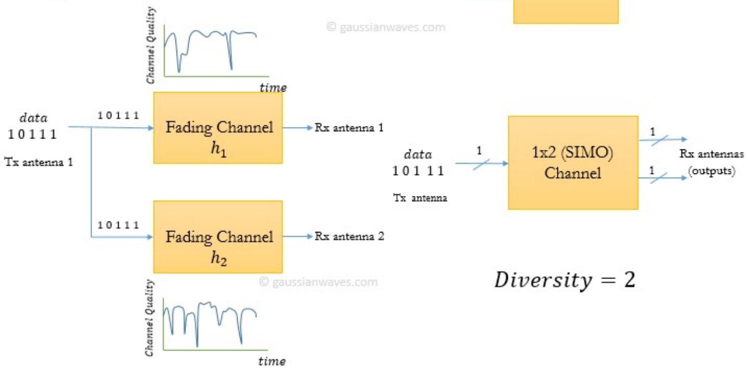
- If N_t transmit antennas and N_r receive antennas that are sufficiently spaced are added to the system
- the *diversity order* is $N_d = N_r N_t$
- N_d is the number of uncorrelated channel paths between the transmitter and receiver
- probability of all the N_d uncorrelated channels having low SNR is very small
- bit error probability improves dramatically

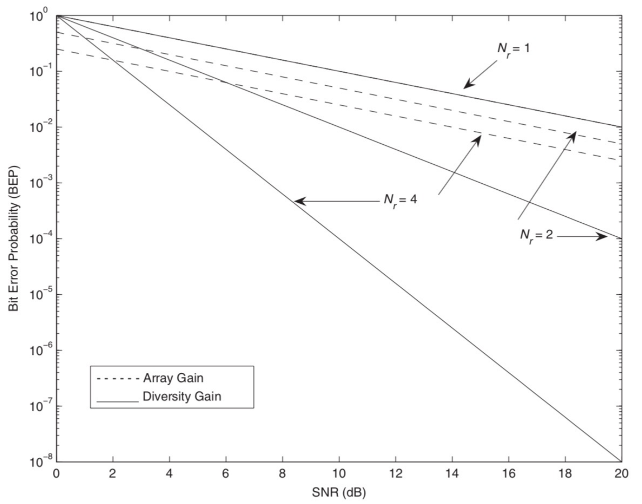
Spatial Diversity – Diversity

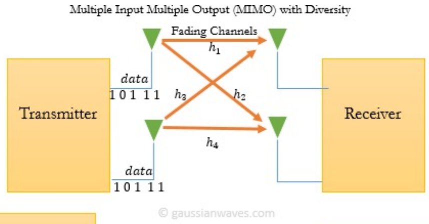


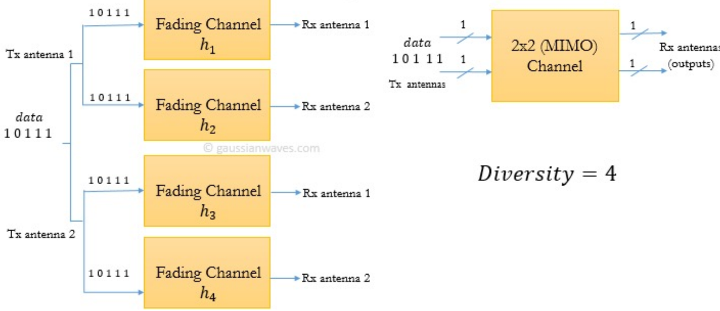












Sufficient spacing for the antennas is critical for increasing the system reliability

Benefits of Spatial Diversity

- Increased data rate
  - Antenna diversity increases SNR linearly
  - Receiver techniques increase capacity logarithmically wrt #antennas
  - data rate benefit rapidly diminishes as antennas are added
  - Multiple independent streams increase aggregate data rate
- Increased coverage or reduced Tx power
  - With only array gain, increase in SNR is $N_r \gamma_i$
  - Increase in SNR increases coverage range
  - transmit power can be reduced by $10log_{10}N_r dB$

Receive Diversity

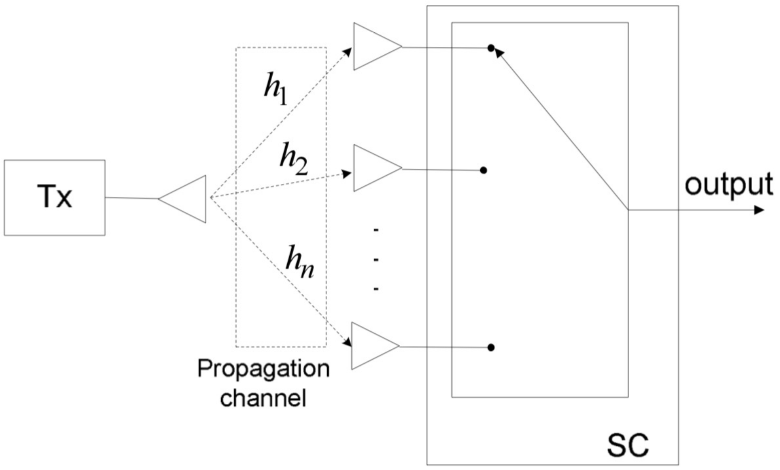
- Receive multiple streams and combine them
  - Selection Combining
  - Maximal Ratio Combining
  - Equal Gain Combining
  - Hybrid Combining

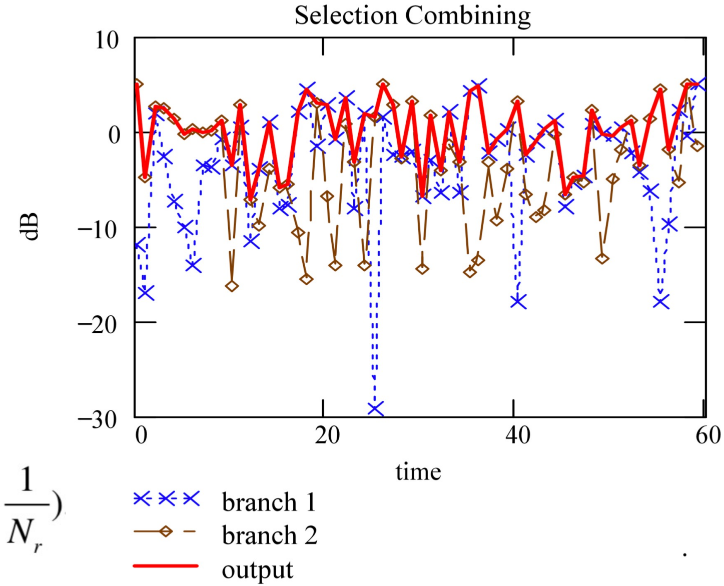
Selection Combining

- estimates the instantaneous strengths of each of the N_r streams and *selects the highest one*
- Since it ignores the useful energy on the other streams, SC is suboptimal
- Its simplicity and reduced hardware requirements make it attractive in many cases

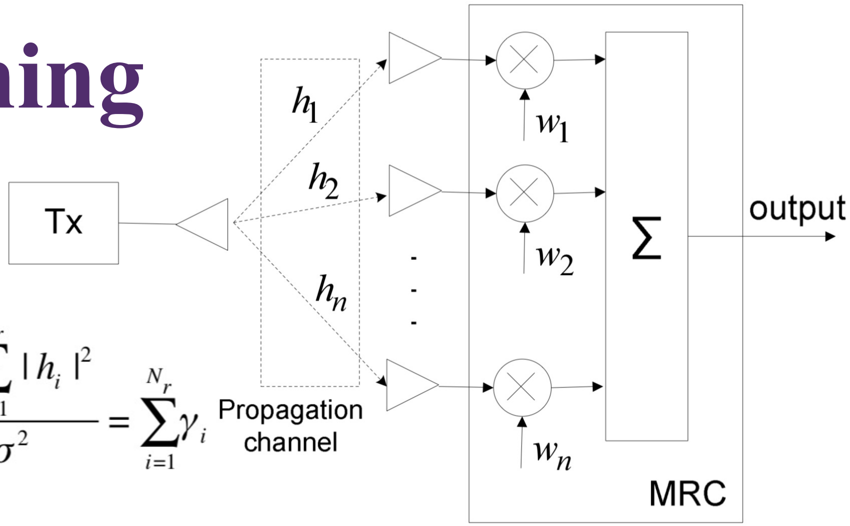
$$\bar{\gamma}_{sc} = \bar{\gamma}\sum_{i=1}^{N_r}\frac{1}{i},$$

$$= \bar{\gamma}\left(1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{N_r}\right)$$





Maximal Ratio Combining



$$\gamma_{MRC} = \frac{\varepsilon_x \sum_{i=1}^{N_r} |h_i|^2}{\sigma^2} = \sum_{i=1}^{N_r} \gamma_i$$

- use linear coherent combining of branch signals so that the output SNR is maximized
- Individual branch signal:

$$x_n = A \cdot h_n + \xi_n$$

- Output of the combiner:

$$x_{out} = \sum_{n=1}^{N} w_n x_n = \underbrace{A \sum_n w_n h_n}_{signal} + \underbrace{\sum_n w_n \xi_n}_{noise}$$

- coherent technique, i.e., signal's phase has to be estimated
- the output is a weighted sum of all branches
- Weights should be proportional to the branch SNRs
- Best performance
- Lot of circuitry for individual receivers

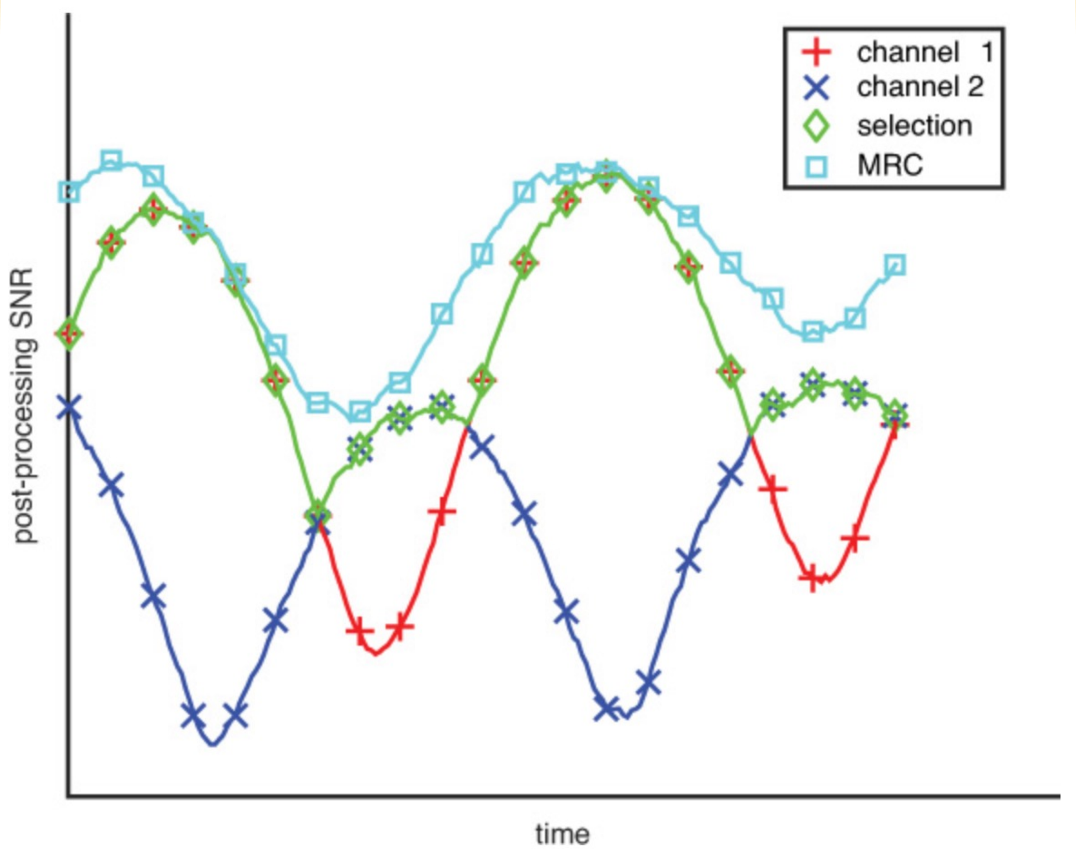
Equal Gain Combining

- MRC requires knowledge of the time-varying SNR on each branch
- corrects only the phase
- Simpler than MRC, easier to implement

$$\gamma_{\mathrm{EGC}} = \frac{\mathcal{E}_x \sum_{i=1}^{N_r} |h_i|^2}{N_r \sigma^2}$$

- Hybrid Combining
  - Combination of multiple of combining techniques

Comparing Receiver Diversity



Transmit Diversity

- signals sent from different transmit antennas interfere with one another
- processing is required at both the transmitter and the receiver
- goal is to achieve diversity while removing or attenuating the spatial interference
- used for the downlink of infrastructure-based systems
- Mobile stations may not need to use it due to size, power constraints
- Can be open loop or closed loop

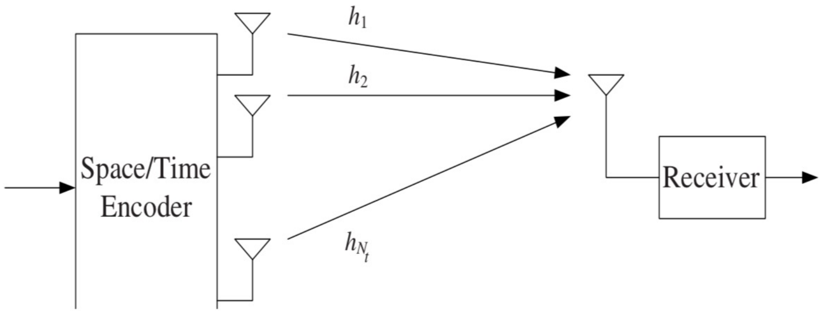
MRC at Transmit

$$y_1 = \sqrt{\rho} \cdot h_1 \cdot x + z_1$$
$$y_2 = \sqrt{\rho} \cdot h_2 \cdot x + z_2$$
$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \sqrt{\rho} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} x + \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

- The complication of transmit diversity
  - obtain the channel phase
  - the channel gain (for SC and MRC) at the transmitter

Open Loop Transmit Diversity

- Space Time Block Codes (STBC)
- Alamouti code is a type of STBC
- ease of implementation—linear at both the transmitter and the receiver



Alamouti Code

	Antenna	1	2
Time	0	s_1	s_2
	1	$-s_2^*$	s_1^*

- If two symbols to be transmitted

$$x_1[1] = u_1, x_2[1] = u_2 \qquad x_1[2] = -u_2^*, x_2[2] = u_1^*$$

- Channel remains constant over two symbols

$$h_1 = h_1[1] = h_1[2], h_2 = h_2[1] = h_2[2]$$

$$[y[1]\ \ y[2]] = [h_1\ \ h_2]\begin{bmatrix} u_1 & -u_2^* \\ u_2 & u_1^* \end{bmatrix} + [w[1]\ \ w[2]]$$

- Rewrite to find u_1, u_2:
- Channel needs to be estimated
- Eliminates spatial interference

$$\begin{bmatrix} y[1] \\ y[2]^* \end{bmatrix} = \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} w[1] \\ w[2]^* \end{bmatrix}$$

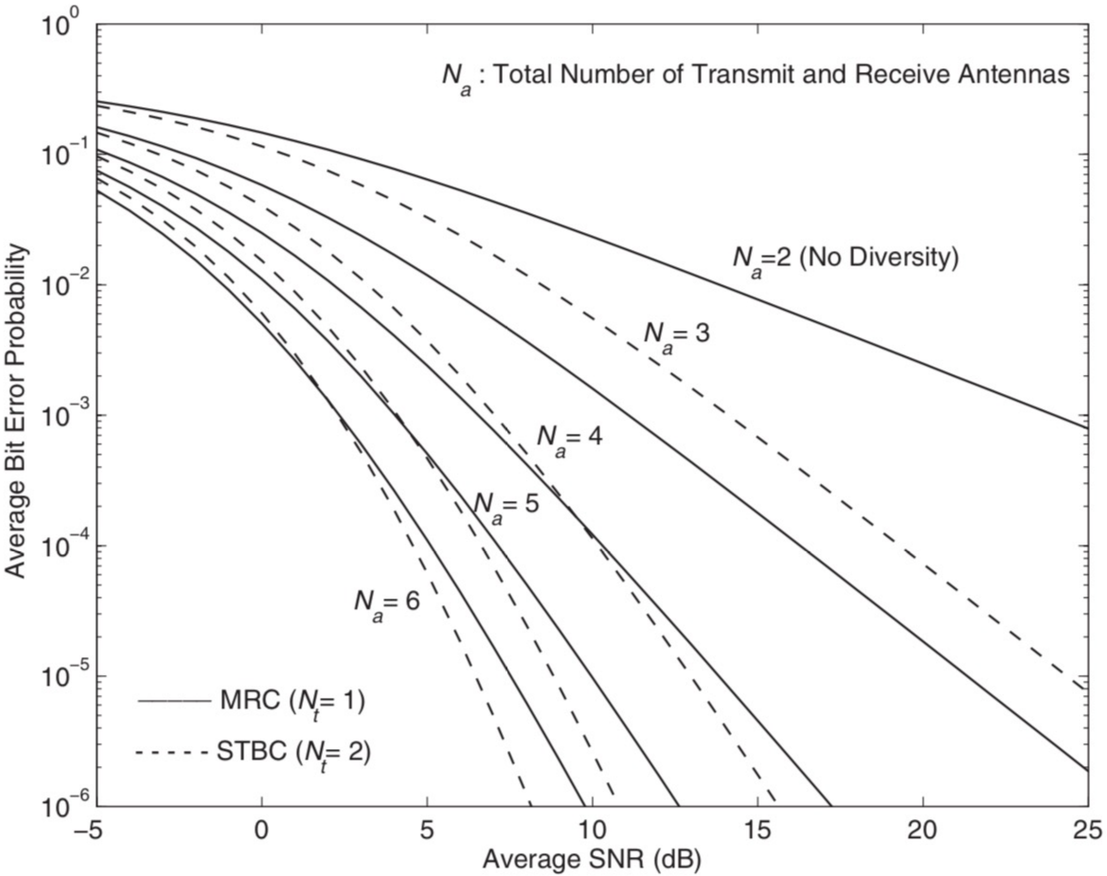
Columns are orthogonal

STBC in OFDM

- Owing to the flat-fading assumption, the STBC in an OFDM system is performed in the *frequency domain*, where each subcarrier experiences flat fading
- Space/time trellis codes introduce memory and achieve better performance (about 2dB) than orthogonal STBCs
- Trellis code decoding complexity $O\left(M^{min\{N_t,N_r\}}\right)$
- STBC complexity $O(min\{N_t, N_r\})$

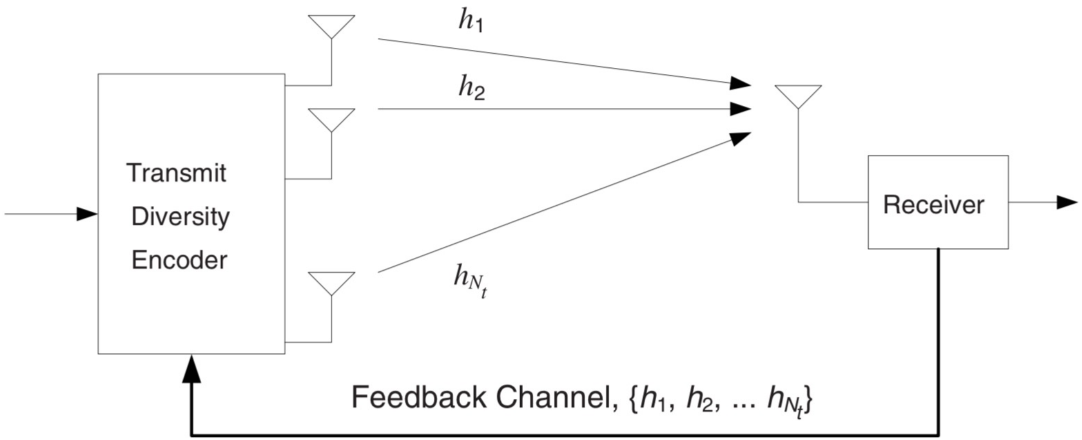
Alamouti STBC vs MRC

- Alamouti STBC outperforms MRC at high SNR owing to the diversity order
- MRC has better BEP performance than Alamouti STBC at low SNR owing to the array gain



Closed loop Transmit Diversity

- Feedback needs to be added to the system
- channel changes quickly in a highly mobile scenario
- closed-loop transmission schemes feasible primarily in fixed or low-mobility scenarios



Transmit Selection Diversity

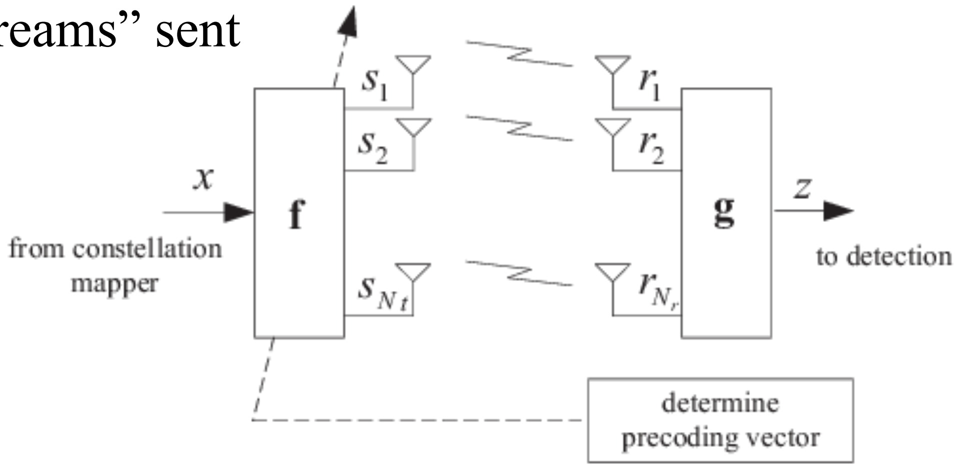
- A subset of all available antennas used
- Subset corresponds to the best channels between the transmitter and the receiver
- Advantages:
  - significantly reduced hardware cost and complexity
  - reduced spatial interference, since fewer transmit signals are sent
  - reaches $N_t N_r$ diversity order, even though only a subset of all antennas are used

Linear Diversity Precoding

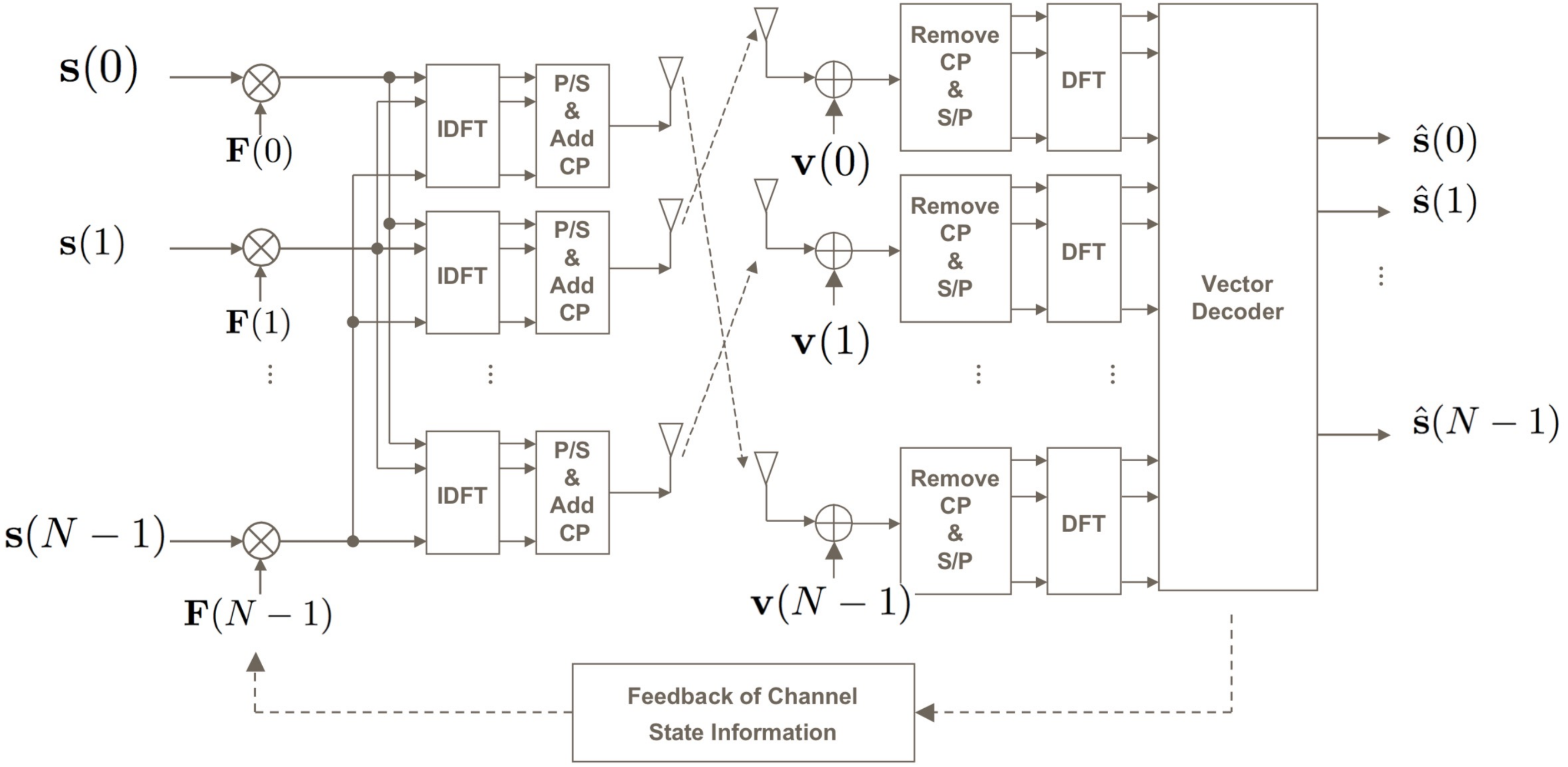
- general technique for improving the data rate by exploiting the CSI at the transmitter
- *diversity* precoding, a special case of linear precoding, where data rate is unchanged
- linear precoder at the transmitter and a linear postcoder at the receiver

Received Data Vector

- $z = Gy = G(HFx + n)$
  - M is the number of spatial data “streams” sent
  - Transmitted vector x is $M \times 1$
  - Received vector y is $N_r \times 1$
  - Postcoder matrix G is $M \times N_r$
  - Channel matrix H is $N_r \times N_t$
  - Precoder matrix F is $N_t \times M$
  - M = 1 is known as maximal ratio transmission (MRT)



Precoding in MIMO OFDM



Interference Cancellation Suppression

- Suppress undesired signals and/or enhance the power of the desired signal
- In MIMO, channel is multidimensional
  - the dimensions of the channel can be applied to null interference in a certain direction, while amplifying signals in another direction
  - Contrast to transmit diversity (statistical diversity of the total signal is increased)
- Types:
  - DOA-Based Beamsteering
  - Linear Interference Suppression: Complete Knowledge of Interference Channels

Beamsteering (Physically steering)

- Electromagnetic waves can be physically steered to create beam patterns at either the transmitter or the receiver
- Static pattern-gain beamsteering : called sectoring
  - Example: in a three-sector cell, a strong beam is projected over 120 degrees, while very little energy is projected over the remaining 240 degrees

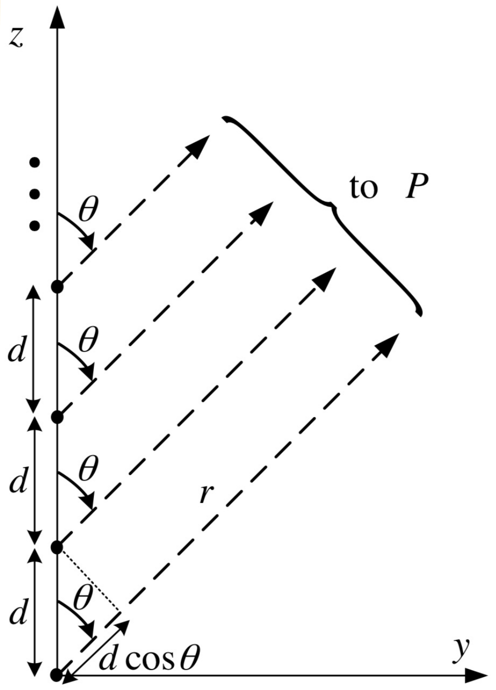
DOA based Beamsteering

- Incoming signal may consist of
  - desired energy + interference energy (other users or multipath)
- Signal processing techniques are used to identify angle of arrival (AoA) of these signals
  - MUSIC, ESPIRIT, JADE, MLE
- These AoAs are used by a beamformer to calculate weighting vector of the antenna elements

Uniform Linear Array

- wave at the first antenna element travels an additional distance of $d \sin \theta$ to arrive at the second element
- difference in propagation distance between the adjacent antenna elements results in arrival-time delay, $\tau = d/c \sin \theta$
- signal arriving at the second antenna can be expressed in terms of signal at the first antenna element

$$y_2(t) = y_1(t)\exp(-j2\pi f_c \tau),$$
$$= y_1(t)\exp(-j2\pi \frac{d\sin\theta}{\lambda}).$$



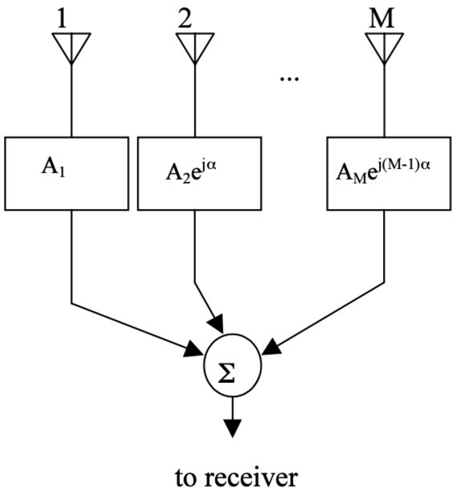
Uniform Linear Array

- For an antenna array with N_r elements all spaced by d , the resulting received signal vector is

$$(t) = [y_1(t)\; y_2(t)\; \dots\; y_{N_r}(t)]^T$$

$$= y_1(t)\underbrace{[1\;\; \exp(-j2\pi\frac{d\sin\theta}{\lambda})\; \dots \exp(-j2\pi(N_r-1)\frac{d\sin\theta}{\lambda})]^T}_{\mathbf{a}(\theta)},$$

a(θ) is the *array response vector*



Weight vector Calculation

- Example:
  - a three-element ULA with $d = \lambda/2$
  - desired signal is received at $\theta_1 = 0$, two interfering signals at $\theta_2 = \pi/3$ and $\theta_3 = -\pi/6$

$$\mathbf{a}(\theta_1) = [1 \;\; 1 \;\; 1]^T, \quad \mathbf{a}(\theta_2) = \left[1 \;\; e^{-j\frac{\sqrt{3}}{2}\pi} \;\; e^{-j\sqrt{3}\pi}\right]^T, \quad and \quad \mathbf{a}(\theta_3) = \left[1 \;\; e^{j\frac{\pi}{2}} \;\; e^{j\pi}\right]^T$$

- Objective:
  - The beamforming weight vector $\mathbf{w} = [w_1 \; w_2 \; w_3\,]^T$ should increase the antenna gain in the direction of the desired user while minimizing the gain in the directions of interferers.

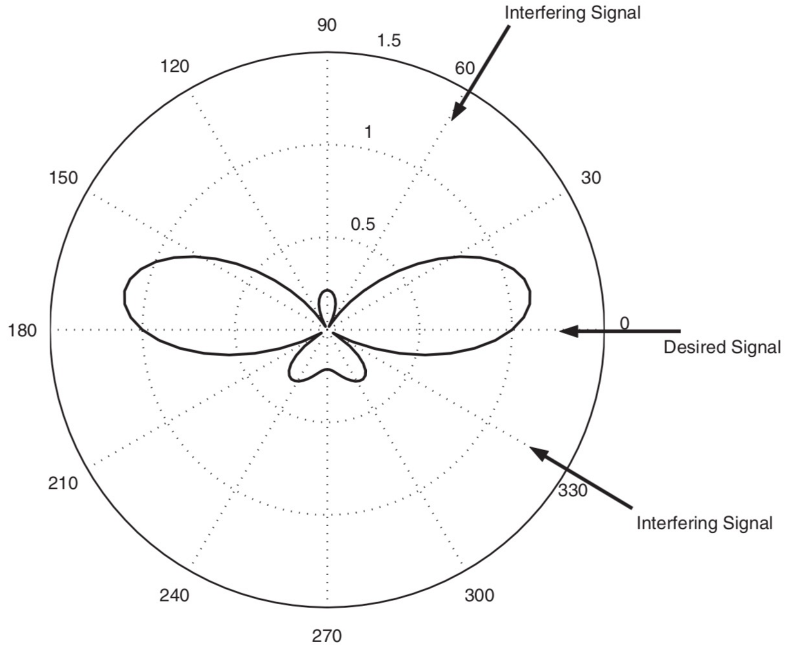
Weight vector Calculation

➢ weight vector **w** should satisfy the following

$$\mathbf{w}^*\left[\mathbf{a}(\theta_1)\ \mathbf{a}(\theta_2)\ \mathbf{a}(\theta_3)\right]=[1\ 0\ 0]^T$$

➢ Solution for weight vector

$$\mathbf{w}=\left[0.3034+j0.1966\ \ 0.3932\ \ 0.3034-j0.1966\right]^T$$



Null-steering Beamformer

- number of nulls is less than the number of antenna elements.
- the antenna gain is not maximized at the direction of the desired user
- trade-off between interference nulled and desired gain lost
- May exist several unresolved components coming from significantly different angles
- DOA-based beamformer is viable only in
  - LOS environments or
  - in environments with limited local scattering around the transmitter

Linear Interference Suppression

- Received signal vector

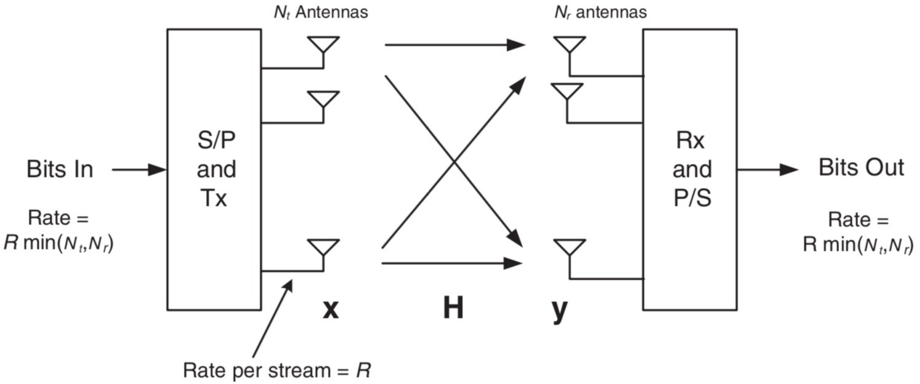
$$\mathbf{y} = \mathbf{H}\mathbf{w}_t x + \mathbf{H}_I x_I + \mathbf{n}$$

- where
  - $\mathbf{w}_t$ is the $N_t x 1$ weighting vector at the desired user's transmitter,
  - x is the desired symbol
  - $\mathbf{x}_I = [x_1\ x_2\ \ldots\ x_L]^T$ is the interference vector
  - $\mathbf{n}$ is the noise vector
  - $\mathbf{H}$ is the N_r x N_t channel gain matrix for the desired user
  - $\mathbf{H}_I$ is the N_r x L channel gain matrix for the interferers

Linear Interference Suppression

- With statistical knowledge of channel:
  - In order to maximize the output SINR at the receiver, joint optimal weighting vectors at both the transmitter and the receiver can be obtained
- This is termed optimum eigenbeamformer, or interference-aware beamforming, or optimum combiner (OC)
- interference-aware beamformer is conceptually similar to the linear diversity precoding
- difference is that the eigen-beamformer takes interfering signals into account

Spatial Multiplexing



- $N_t <= N_r$
- Split the incoming high rate-data stream into N_t independent data streams
- decoding N_t streams is theoretically possible when there exist at least N_t nonzero eigenvalues in the channel matrix, that is rank($\mathbf{H}$) $\geq N_t$
- *Assuming* that the streams can be successfully decoded, the nominal spectral efficiency is thus increased by a factor of N_t

Spatial Multiplexing: Key Points

- When the SNR is high, spatial multiplexing is optimal.
  - The capacity, or maximum data rate, grows as $\min(N_t, N_r) \log(1 + SNR)$ when the SNR is large.
- When the SNR is low, the capacity-maximizing strategy is to send a single stream of data using diversity precoding.
  - Although the capacity is much smaller than at high SNR, it still grows approximately linearly with $\min(N_t, N_r)$ since capacity is linear with SNR in the low-SNR regime.

Spatial Multiplexing: Key Points

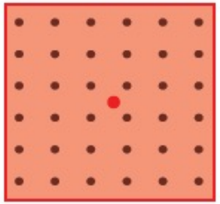
- Both of these cases are superior in terms of capacity to space-time coding, where the data rate grows at best logarithmically with N_r
- The average SNR of all N_t streams can be maintained without increasing the total transmit power relative to a SISO system
  - each transmitted stream is received at $N_r \geq N_t$ antennas and hence recovers the transmit power penalty of N_t due to the array gain.
- Note: even a single low eigenvalue in the channel matrix can dominate the error performance.

Open Loop Spatial Multiplexing

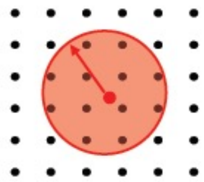
- **Optimal Receiver:**
  - Maximum likelihood: finds input symbol most likely to have resulted in received vector
  - Exponentially complex with # of streams and constellation size
- **Sphere Decoder:**
  - Only considers possibilities within a sphere of received symbol.
    - If minimum distance symbol is within sphere, optimal, otherwise null is returned

$$\hat{x} = \arg\min_{x} |y - Hx|^2$$

ML Decoding



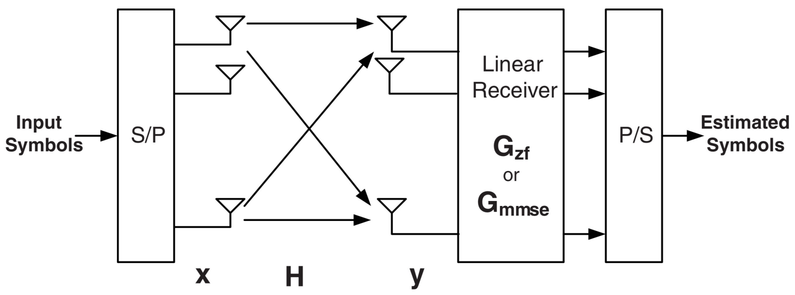
Sphere Decoding



$$\hat{x} = \arg\min_{x:|Q^H y - Rx| < r} |Q^H y - Rx|^2$$

Linear Detectors : Zero Forcing Detector

- sets the receiver equal to the inverse of the channel $\mathbf{G}_{zf} = \mathbf{H}^{-1}$



$$\mathbf{G}_{zf} = (\mathbf{H}^*\mathbf{H})^{-1}\mathbf{H}^*$$

$$\hat{\mathbf{x}} = \mathbf{G}_{zf}\mathbf{y} = \mathbf{G}_{zf}\mathbf{H}\mathbf{x} + \mathbf{G}_{zf}\mathbf{n} = \mathbf{x} + (\mathbf{H}^*\mathbf{H})^{-1}\mathbf{H}^*\mathbf{n}.$$

- zero-forcing detector removes the spatial interference from the transmitted signal
- As $\mathbf{G}_{zf}$ inverts eigenvalues of H, poor subchannels can severely amplify noise
- Not practical in interference-limited MIMO (LTE)

Linear Detectors : MMSE Receiver

- MMSE receiver attempts to strike a balance between spatial-interference suppression and noise enhancement by minimizing the distortion

$$\mathbf{G}_{mmse} = \arg\min_{\mathbf{G}} \mathrm{E} \| \mathbf{G}\mathbf{y} - \mathbf{x} \|^2$$

$$\Downarrow$$

$$\mathbf{G}_{mmse} = (\mathbf{H}^*\mathbf{H} + \frac{\sigma_z^2}{P_t}\mathbf{I})^{-1}\mathbf{H}^*$$

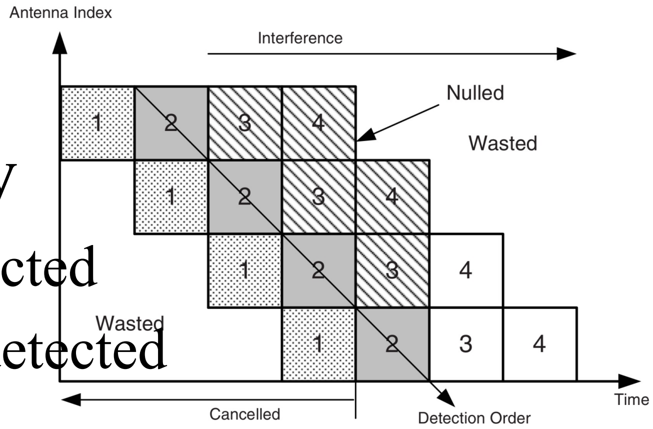
- As the SNR grows large, the MMSE detector converges to the ZF detector
- At low SNR, it prevents the worst eigenvalues from being inverted

Interference Cancellation: BLAST

- *Bell labs LAyered Space-Time* (BLAST) : invented and prototyped in Bell Labs
- BLAST consists of parallel "layers" supporting multiple simultaneous data streams
- The layers (substreams) in BLAST are separated by interference-cancellation techniques that decouple the overlapping data streams
- two most important techniques are
  - the original *diagonal BLAST* (D-BLAST)
  - its subsequent version, *vertical BLAST* (V-BLAST)

D-BLAST

- in each layer's data is transmitted in a *diagonal* of space and time
  - groups the symbols into "layers" that are then coded in time independently of the other layers
  - these layers are then cycled to the various transmit antennas in a cyclical manner
- one layer decoded at a time
- Each successive layer is detected by
  - nulling the layers that have not yet been detected
  - canceling the layers that have already been detected

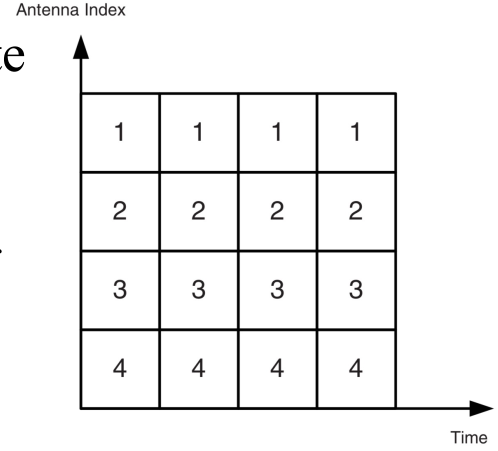


D-BLAST Pros & Cons

- Pro: each symbol stream achieves diversity
  - in time via coding and
  - in space by it rotating among all the antennas
- Cons:
  - Decoding process is iterative and complex
  - wastes space/time slots at the beginning and end of a D-BLAST block

V-BLAST

- each antenna transmits an independent symbol stream—for example, QAM symbols
- different techniques can be used at the receiver to separate the various symbol stream from one another
  - Including ZF, MMSE
  - the strongest symbol stream is detected, using a ZF or MMSE receiver
  - subtracted out from the composite received signal
- Pros:
  - ordered successive interference cancellation lowers the block error rate by a factor of ten relative to a purely linear receiver
- Cons:
  - error propagation when initial layers are detected incorrectly leads to huge penalty
  - depends on high SNR (not available in cell edge)

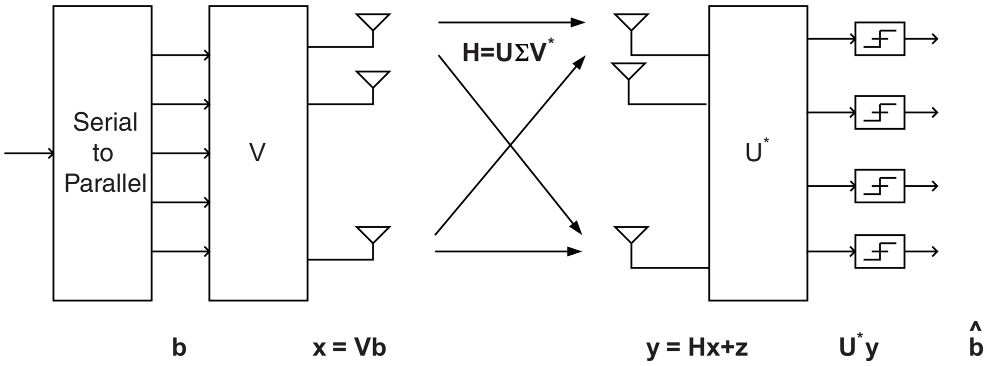


Closed Loop Spatial Multiplexing

- The advantage of channel knowledge
- SVD Precoding and Postcoding
  - Channel expressed as singular-value decomposition (SVD, or generalized eigenvalue decomposition)
  - **U** and **V** are complex unitary matrices, ***Σ*** is a diagonal matrix of singular values (non-negative real numbers)

Impractical, but promising results compared to open loop approach

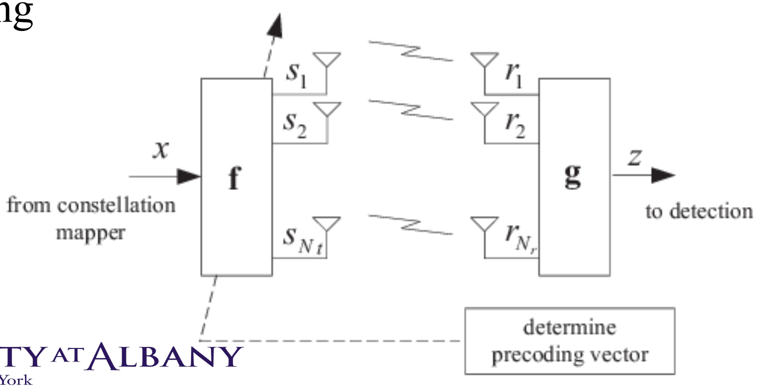
complexity of finding the SVD of an $N_t x N_r$ matrix is on the order of $O(N_r, N_t^2)$ if $N_r >= N_t$



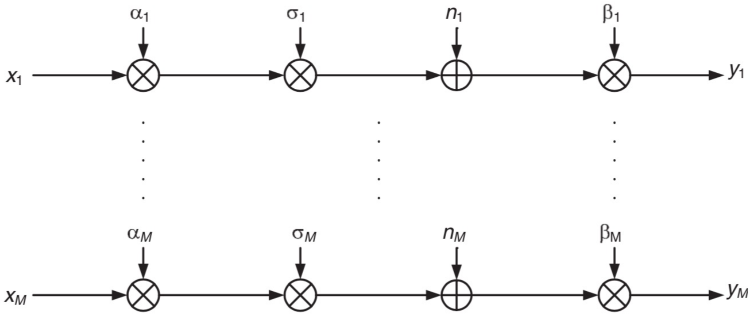
Chapter 7 of Tse & Vishwanath book

Linear Precoding and Postcoding

- decomposes the MIMO channel into a set of parallel subchannels
- the precoder and the postcoder can be jointly designed based on
  - information capacity, error probability, detection MSE, or received SNR
- precoder weights are used to maximize the total capacity by distributing more transmission power to subchannels with larger gains and less to the others - waterfilling



- $z = Gy = G(HFx + n)$
  - M is the number of spatial data "streams" sent
  - Transmitted vector x is $M\times 1$
  - Received vector y is $N_r\times 1$
  - Postcoder matrix G is $M\times N_r$
  - Channel matrix H is $N_r\times N_t$
  - Precoder matrix F is $N_t\times M$
  - $1 \le \text{M} \le \min(N_r, N_t)$



$$y_i = \alpha_i \sigma_i \beta_i x_i + \beta_i n_i, \quad i = 1, \cdots, M$$

σ_i are the singular values of $\mathbf{H}$,

Channel Estimation for MIMO OFDM

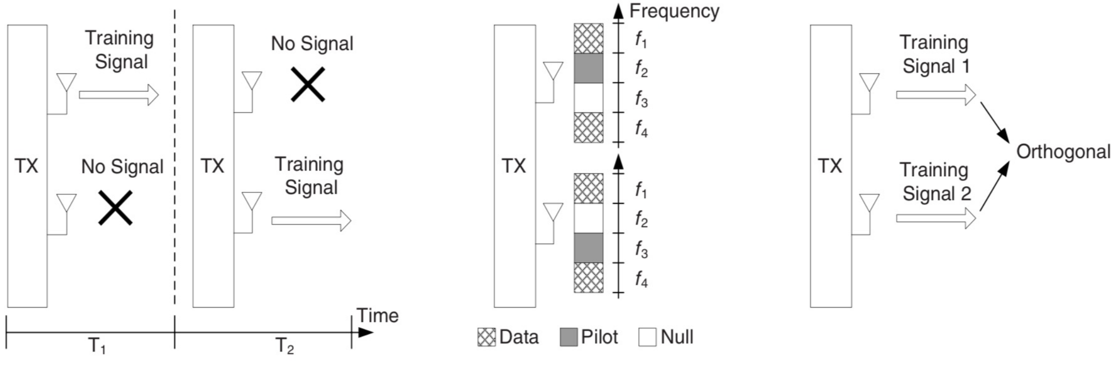
- Channel estimation required
  - At the receiver in order to
    - coherently detect the received signal
    - for diversity combining
    - spatial-interference suppression
  - At the transmitter
    - For closed loop MIMO
- Types:
  - Training based – known symbols (preambles, pilots) transmitted, reliable, mostly used
  - Blind – no training, no overhead, low convergence speed, lower estimation accuracy

Training Symbols

- Two ways to transmit training symbol:
  - Preambles : send a certain number of training symbols prior to the user data symbols
  - Pilot tones : insert a few known (time, frequency, phase, amplitude) pilot symbols among the subcarriers
- Channel estimation typically done by using
  - the preamble for synchronization and initial channel estimation
  - the pilot tones for tracking the time-varying channel in order to maintain accurate channel estimates

Pilot Insertion Patterns

- received signal at each antenna is a superposition of the signals transmitted from N_t transmit antennas
- the training signals for each transmit antenna should not interfere with one another
- Independent: orthogonality achieved in time domain, requires N_t training signal times
- Scattered: orthogonality achieved in frequency domain
- Orthogonal: orthogonality achieved using orthogonal codes



(a) Independent Pattern (b) Scattered pattern (c) Orthogonal pattern