Short review of the information theory

The amount of information in a random variable is measured by its entropy.
The entropy of a discrete random variable with the probability mass function $p(x)$ is defined as

$$H(X) \triangleq -\sum_x p(x) \log_2 (p(x))$$

The conditional entropy is the average information quantity which is needed to specify the input symbol x when the output symbol y is known.
The conditional entropy represents average amount of information that is lost in the channel and is called equivocation.

$$H(X|Y) \triangleq -\sum_x \sum_y p(x,y) \log_2 (p(x|y)) \text{ bit/symbol}$$

Part of the information transmitted over the channel is lost because of the noise in the channel.
The lost part is measured by the channel equivocation $H(X|Y)$
The average information flow is defined as mutual information between X and Y.

$$I(X;Y) \triangleq H(X) - H(X|Y) \text{ bit/symbol}$$

$$I(X;Y) = H(Y) - H(Y|X) = H(X) + H(Y) - H(X,Y)$$

The maximum value of $I(X;Y)$ is called channel capacity.

Channel capacity

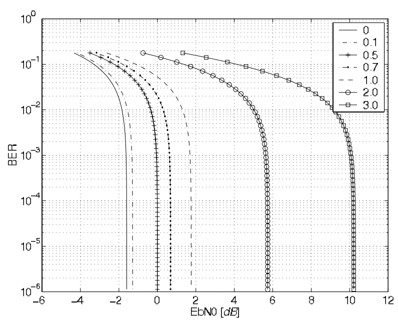


Figure 1: Channel capacity as function of data rate R_c

Shannon capacity:

Maximum achievable error free information flow in a given channel.
Assumes infinite long data sequence.
In classical form derived for a point to point connection.
For the real channel with infinite bandwidth

$$C = \frac{1}{2} \log \left[\left(1 + \frac{2R_cE_b}{N_0} \right) \right]$$

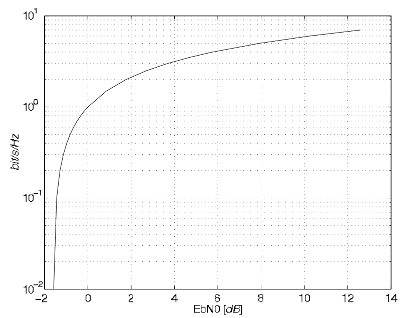


Figure 2: Spectral efficiency as $bit/s/Hz$

Practical capacity definition

Achievable BER for given physical realisation of the transmitter and receiver.

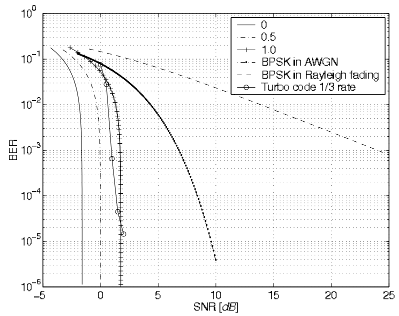


Figure 3: Probability of Bit Error Ratio for binary signals.

Capacity in fading channel

$$C(\nu) = W \log_2 (1 + \nu SNR)$$

ν exponentially distributed random variable.
Block fading channel. Flat Rayleigh fading with no dynamics.
Channel state information known only to the receiver.
The channel is described as the random variable. Distribution of the mutual information between the transmitted and receiver.

Ergodic capacity

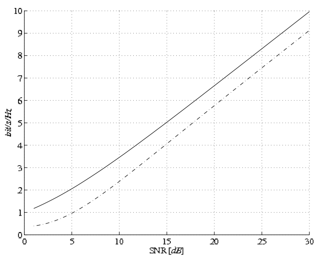


Figure 4: Channel capacity in Gaussian Noise and Rayleigh fading channel

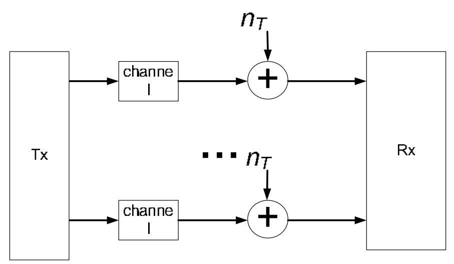
Average of the maximal value of the mutual information between the transmitter and receiver.
In random fading channel that can be calcualted as average over the Shannon capacities at each fading level - over the distribution of $C(\nu)$.

The mutual informtaion changes over the time. At each power level the channel can support different rate.

Outage capacity

We allow some error in case of severe fading. (outage condition)
Because the channel has random amplitude the capacity at certain time instant can be less than the transmitted data rate.
The outage capacity C_{outage} is associated with the probability Pb_{outage} that the instantaneous channel capacity is less than the transmitted data rate.
Outage capacity is the maximum data rate that with given transmission power can be transmitted if exclude the states where the system is in outage.
The probability of the system being in outage is constrained of being less or equal to Pb_{outage}.

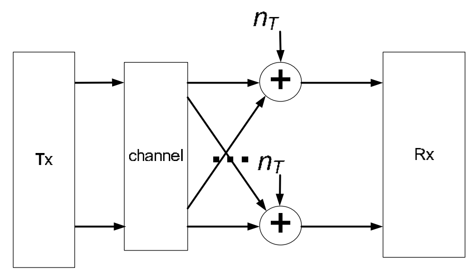
Capacity of independent parallel additive Gaussian noise channels



The capacity is calculated as the sum of the capacities of the individual channels

$$C = \sum_i \log_2 \left(1 + \frac{P_{tot}}{\sigma_\eta^2 n_T}\right) = \log_2 \left| I + \frac{P_{tot}}{\sigma_\eta^2 n_T}\mathbf{I}\right|$$

Non independent parallel additive Gaussian noise channels



The capacity is the logarithm of the determinant

$$I\left(S;R\right) = C = \log_2 \left| \mathbf{I_{n_R}} + \frac{P_{tot}}{\sigma_\varepsilon^2 n_T}\mathbf{HH}^H \right|$$

H Fourier transform of the channel impulse transform matrix.

- We assume a narrowband system.
  In the narrowband system the elements of $\mathbf{H}(f)$ can be assumed to be constants, equal to the channel response $g_{ij}(0)$.

H Describes correalation of between the tranmit and receive antennaes.

Capacity depends on the rank of the $\mathrm{HH}^{\mathbf{H}} matrix$.

$$\mathbf{H} = \mathbf{U}\cdot\mathbf{\Sigma}\cdot\mathbf{V}^H$$

Where:

U and **V** are orthogonal matrixes
$\mathbf{HH}^H = \mathbf{U}\cdot\sigma\cdot\mathbf{V}^H\mathbf{V}\cdot\sigma\cdot\mathbf{U}^H = \mathbf{U}\sigma^2\mathbf{U}^H$.
$|\mathbf{U}|\cdot|\mathbf{U}^H| = \mathbf{I}$

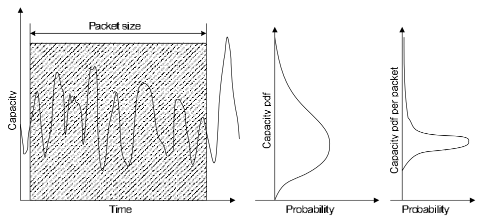
Fading channel

- The components of **H** are complex Gaussian iid with unit variance.
  - Channel amplitude at each receiver has Rayleigh distribution.
- Since channel matrix random mutual information also random.

Fast fading channel

Channel is assumed to be ergodic.
During the packet the fading visits all the possible channel states.
The mutual information is calculated as an average

$$I(S;R) = E\{I(S(t);R(t))\} = E\left\{\frac{1}{2}\log_2\left|\mathrm{I}_{n_R} + \frac{\rho(t)}{n_T}\mathrm{HH}^H\right|\right\}$$



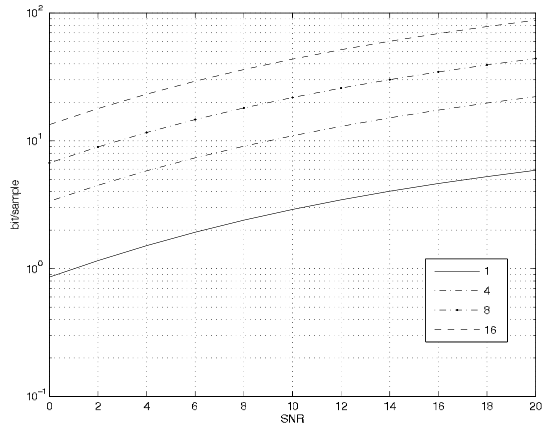


Figure 5: Increase of the Mutual information with the increase of the number of transmit and receive antennaes

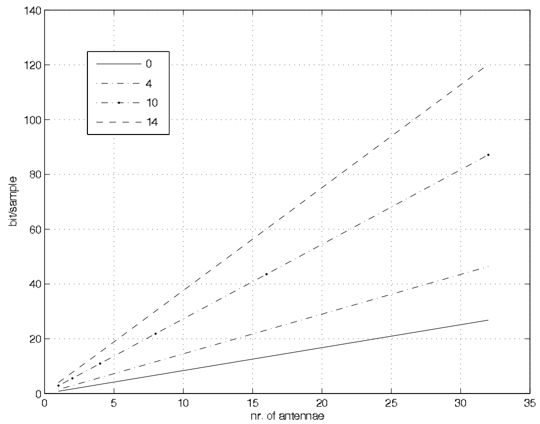


Figure 6: Increase of the mutual information for different values of SNR.

Distribution of the outage Pb

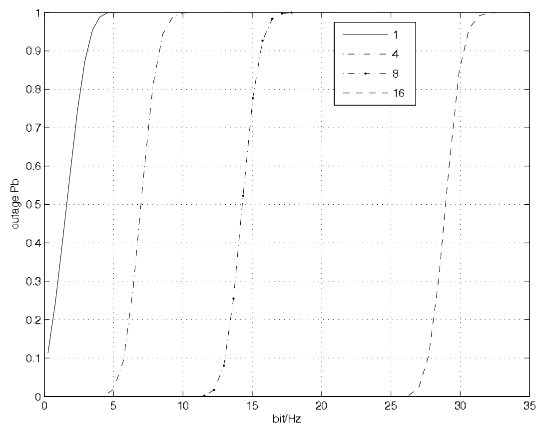


Figure 7: Cdf of the outage probability for different number of antennaes in the iid Gaussial distributed coefficients at $SNR = 6\ dB$

Slow fading channel

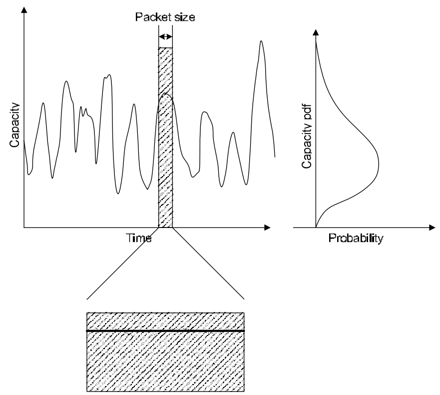
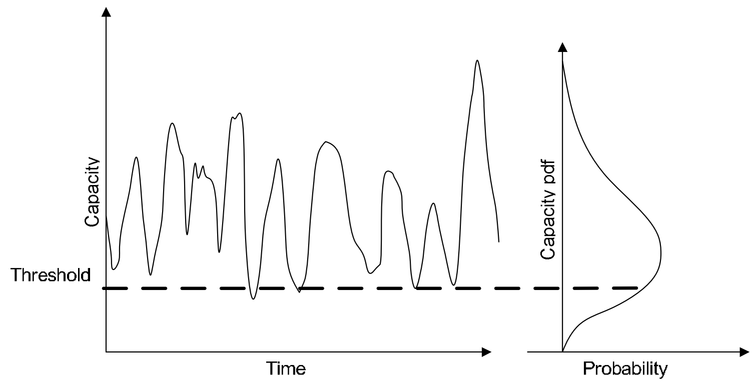
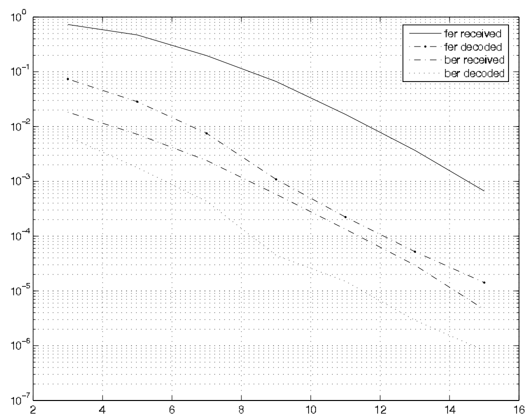


Figure 8: Quasi static channel



Because the channel amplitude changes over time, it migth happen that we attemt to send more than the channel capacity allows.
The system is in outage and the packet is with high probablity errornous.

Coding impact to BER and FER performance



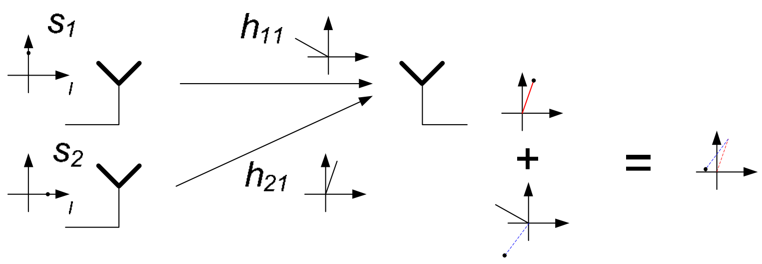
FER and BER before and after coding in an 2x2 system

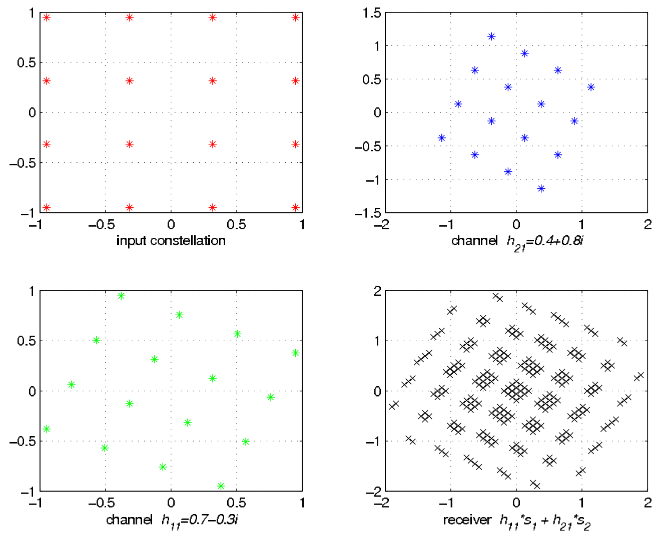
- If the system is in outage decoding introduces more errors

  The decoder converges with high probability to wrong codeword.
  Wrong codeword increase the BER.
  The frame that was errornous remains errornous.
- If the system is not in outage a strong coding allows to correct all the errors and the FER is improved.

Interpretation of the system

figure 2x1 constellation





Multiple antenna systems

- Diversity
  - Compensate against channel unreliability.
  - Minimize the pairwise error probability between the codewords. Diversity can be interpreted as the distance between all the correct and errornous codewords.
  - Diversity equals to the product of the nonzero eigenvalues of the codeword difference matrixes.
  - Diversity order - the slope of the BER curve if plotted versus the average SNR in log-log scale.

- Spatial multiplexing
  - We can create independent channels and transmit more data.
  - Increase degrees of freedom.
  - How many parallel spatial channels between the transimitter and receiver can be created.
  - Spatial multiplexing gain r - increase of the data rate $R = r\log(SNR)\ \frac{bps}{Hz}$ compared to the single link data rate.

Diversity

If we are not separating different paths of the signal on radio interface they are producing fading.
Additional paths on radio interface create also diversity
For multiple antennas we can utilise spatial diversity.
Depending where the antennas are located the diversity is identified as:

- Receive diversity
- Transmit diversity
- Both

System description

b=a(:,len$_h + 1 : length(tr_seq) - len_h$);m$receivingantenna$
n$transmitantenna$
$Foreachreceivingantenna$,$1 \leq j \leq m$, we have

$$d_t^j = \sum_{i=1}^{n} h_{i,j} c_t^i \sqrt{E_s} + \eta_t^i$$

$h_{i,j}$ the path gain from the transmitting antenna i to receiver j.

Maximum likelihood combining

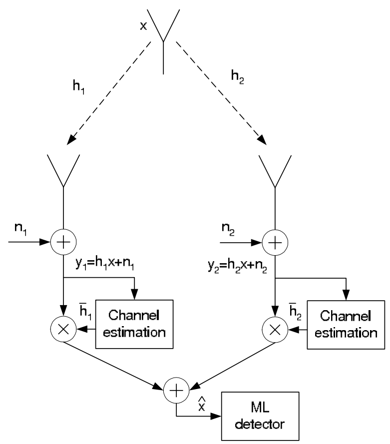


Figure 9: MRC technique using two receivers.

Error probability decays as SNR^{-2}

$$\begin{aligned} h_1 &= |h_1| e^{j\theta_1} \\ h_2 &= |h_2| e^{j\theta_2} \\ y_1 &= h_1 \cdot x + n_1 \\ y_2 &= h_2 \cdot x + n_2 \end{aligned}$$

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = x \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} + \begin{pmatrix} n_1 \\ n_2 \end{pmatrix}$$

$$\begin{aligned} \hat{x} &= \bar{h}_1 y_1 + \bar{h}_2 y_2 \\ &= \bar{h}_1 h_1 x + \bar{h}_1 n_1 + \bar{h}_2 h_2 x + \bar{h}_1 n_2 \\ &= \left(|h_1|^2 + |h_2|^2\right) x + \bar{h}_1 n_1 + \bar{h}_2 n_2 \end{aligned}$$

Transmit diversity

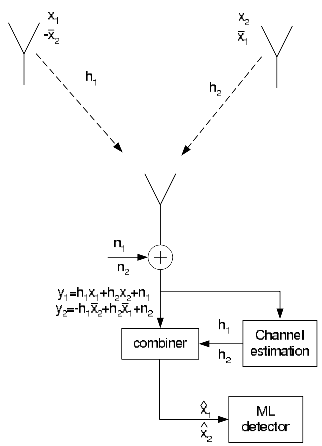


Figure 10: Example of transmit diversity.

Assume that we can consider only one sample
The transmitted signal is preweighted
The received sample can be written as

$$y_1 = h_1 \cdot w_1 x + h_2 \cdot w_2 x + n_1$$

$$y_1 = \frac{|h_1 \cdot w_1 + h_2 \cdot w_2|^2}{\sigma^2} \cdot E\{x^2\}$$

If the transmitter does not know the channel the weights do not depend on the channel.
In one symbol interval diversity can not be achieved.

Error probability in the system

Two codewords can be described as

$$c = \begin{pmatrix} c_1^1 & c_2^1 & \dots & c_l^1 \\ c_1^2 & c_2^2 & \dots & c_l^2 \\ \dots & \dots & \ddots & \vdots \\ c_1^n & c_2^n & \dots & c_l^n \end{pmatrix} \quad e = \begin{pmatrix} e_1^1 & e_2^1 & \dots & e_l^1 \\ e_1^2 & e_2^2 & \dots & e_l^2 \\ \dots & \dots & \ddots & \vdots \\ e_1^n & e_2^n & \dots & e_l^n \end{pmatrix}$$

The difference between the two codewords defines the errors

$$P(c \to e|\alpha_{i,j}) < e^{\left(\frac{E_s}{4N_0} d^2(c,e)\right)}$$

$$P(c \to e) \leqslant \prod_{j=1}^{m} \exp\left(-\Omega_j A(c,e)\Omega_j^H \frac{E_s}{4N_0}\right)$$

$d(c,e)$ is a difference between two codewords.
The error occures when the received signal is in the decision area of another codeword.

$$\begin{aligned} d^2(c,e) &= \sum_{t=1}^{l}\sum_{j=1}^{m}\left|\sum_{i=1}^{n} h_{i,j}\left(c_t^i - e_t^i\right)\right|^2 \\ &= \sum_{t=1}^{l}\sum_{j=1}^{m}\left(\sum_{i=1}^{n} h_{i,j}\left(c_t^i - e_t^i\right)\right)\overline{\left(\sum_{i'=1}^{n} h_{i',j}\left(c_t^{i'} - e_t^{i'}\right)\right)} \\ &= \sum_{j=1}^{m}\sum_{i=1}^{n}\sum_{i'=1}^{n} h_{i,j}\overline{h_{i',j}}\sum_{t=1}^{l}\left(c_t^i - e_t^i\right)\overline{\left(c_t^{i'} - e_t^{i'}\right)} \\ &= \sum_{j=1}^{m}\Omega_j A(c-e)\Omega_j^H \end{aligned}$$

$$\Omega = \left[\begin{array}{ccc} h_{1,j} & \dots & h_{n,j} \end{array}\right],\ A_{pq} = \sum_{t=1}^{l}(c_t^p - e_t^p)\overline{(c_t^q - e_t^q)}$$

The Hermitian matrix A can be decomposed $A(c-e) = V\Lambda V$
Channel coefficients can be replaced by $\beta = \Omega \cdot V$

$$\Omega_j A(c-e)\Omega_j^H = \sum_{i=1}^{n}\lambda_i\left|\beta_i, j\right|^2$$

Rayleigh fading channel

$$\int e^{jvx}p(x)dx$$

is the characteristic function.
Assume $|\beta|$ is complex normally distributed random variable with the characteristic function

$$\frac{1}{\left(1 + \frac{\lambda_i E_s}{4N_0}\right)}$$

$$P(c \to e) \leqslant \left(\frac{1}{\prod_{i=1}^{n}\left(1 + \frac{\lambda_i E_s}{4N_0}\right)}\right) \leqslant \left(\prod_{i=1}^{r}\lambda_i\right)^{-m}\left(\frac{E_s}{4N_0}\right)^{-rm}$$

Diversity is rm.
Coding advantage is comparable to $(\lambda_1\lambda_2 \dots \lambda_r)$.

Constraints to the codes for achieving diversity

- The rank criterion: Maximize the diversity over all possible codeword pairs In order to achieve the full diversity the matrix B should have the full rank (assumes that the constellation does not change)
- The determinant criterion: The Determinant criterion maximize the coding advantage over all distinct codeword pairs c and e.
- The trace criterion: in the relay fading channel the distance between the possible codewords is maximized when the trace of the matrix $D(c-e)D(c-e)^H$ is as high as possible.

Some existing Space Time coding schemes

- BLAST schemes.
  Not all the schemes achieve the full diversity.
  Diversity is traded for the rate.
  Expects equal number of transmit and receive antennaes.
- Linear space time codes.
  Alamouti scheme.
  Good diversity properties.
  Not scalable to many antennaes.
  Simple decoding algorithm.
- Lattice codes.
  Allow full rate full diversity transmission.
  Simple (nonoptimal) decoding algorithm.
- Space time trellis codes

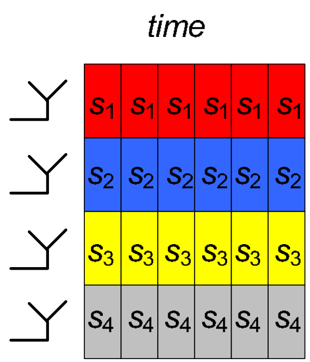
Transmitter side operations for qurantees the diversity

- ST codes → linear receiver.
- BLAST → successive cancellation receiver.
- Signal space codes and threaded allocation → Sphere decoder.

Bell Laboratories Layered Space-Time architecture (BLAST)

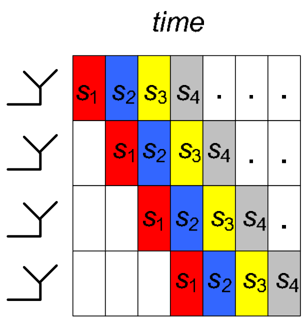
- Incoming data is separated into parallel streams.
- The streams are allocated to the transmit antennaes. Different allocation algorithms result in different type of BLAST's.
- Horizontally aloocation → HBLAST
  - One codeword per horizontal layer.

- Vertical allocation of streams → VBLAST



– Each stream allocated independently to each antennae
– No need for coding.
– No transmit diversity, diversity gain only over the receiver antennaes.

- Diagonal allocation of streams → DBLAST



  - Each symbol is transmitted over each antennaes.
  - Decoder cancels already received signals.
  - High diversity is traded to the rate.

- Mixed schemes. Combination of diagonal and vertical mapping
- Threaded layering. Before mapping the streams are transformed in order to increase resistance to fading and to retain the rate.

Receivers for BLAST

- Maximum Likelihood receiver

$$\hat{x} = \arg\min_{x \in C} |y - \mathbf{H}x|^2$$

- Zero Forcing Receiver

$$\hat{x} = \left(\mathbf{H}^H \mathbf{H}\right)^{-1} \mathbf{H}^H y$$

- MMSE receiver

$$\varepsilon^2 = E\left\{(x - \hat{x})^* (x - \hat{x})\right\}$$

$$\hat{x} = \left(\frac{1}{SNR}\mathbf{I}_m + \mathbf{H}^H \mathbf{H}\right)^{-1} \mathbf{H}y$$

- Successive interference cancellation.
- Sphere decoder.

Space Time Block Codes (STBC)

- If we can use only one time sampe withouth channel knowledge at the transmitter diversity cannot be achieved.
- The transmit diversity can be achieved via spreading the infoamtion over subsequent samples. (increase of the channel usages).
- The pupose of the space time coding is to maximize the transmitted information rate, at the same time as error probaility is minimized.
- Space Time Block Codes are the way of mapping a set of complex input samples (with size k) to the output matrix $\mathbf{X}$ with dimension $n \times p$ that is transmitted in column wise from p different antennas.

Types of STBC

- Linear
  The mapping of the input symbols is linear.
- Nonlinear
  The input symbols are first transmitted by non linear transformation to new set of symbols. Those symbols are mapped inot the matrix $\mathbf{X}$.
- Space time trellis coding
  The input data is encoded by convolutional encoder with p outputs.
- Delay diversity
  Special case of the trellis coding.
  The first antenna transmits the initial bit stream.
  The second transmits the same stream delayed by some symbol intervals

...

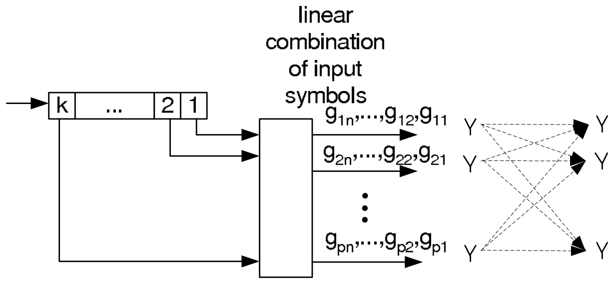


Figure 11: Linear STBC.

Alamouti scheme

A simple two transmitter based scheme with transmission matrix:

$$\mathbf{G}_2 = \begin{pmatrix} x_1 & x_2 \\ -\bar{x}_2 & \bar{x}_1 \end{pmatrix}$$

The matrix elemets are transmitted as following:

Time slot, T	antenna	
	Tx 1	Tx 2
1	x_1	x_2
2	$-\bar{x}_2$	$\bar{x}_1$

Using one receiver

$$y_1 = h_1 \cdot x_1 + h_2 \cdot x_2 + n_1$$
$$y_2 = h_1 \cdot \bar{x}_2 + h_2 \cdot \bar{x}_1 + n_2$$

$$\begin{aligned} \hat{x}_1 &= \bar{h}_1 y_1 + h_2 \bar{y}_2 \\ &= \bar{h}_1 h_1 x_1 + \bar{h}_1 h_2 x_2 + \bar{h}_1 n_1 - h_2 \bar{h}_1 x_2 + h_2 \bar{h}_2 x_1 + h_2 n_2 \\ &= \left(|h_1|^2 + |h_2|^2\right) x_1 + \bar{h}_1 n_1 + h_2 \bar{n}_2 \end{aligned}$$

$$\begin{aligned} \hat{x}_2 &= \bar{h}_2 y_1 - h_1 \bar{y}_2 \\ &= \bar{h}_2 h_1 x_1 + \bar{h}_2 h_2 x_2 + \bar{h}_2 n_1 + h_1 \bar{h}_1 x_2 - h_1 \bar{h}_2 x_1 - \bar{h}_1 n_2 \\ &= \left(|h_1|^2 + |h_2|^2\right) x_2 + \bar{h}_2 n_1 - \bar{h}_1 n_2 \end{aligned}$$

Using multiple receivers

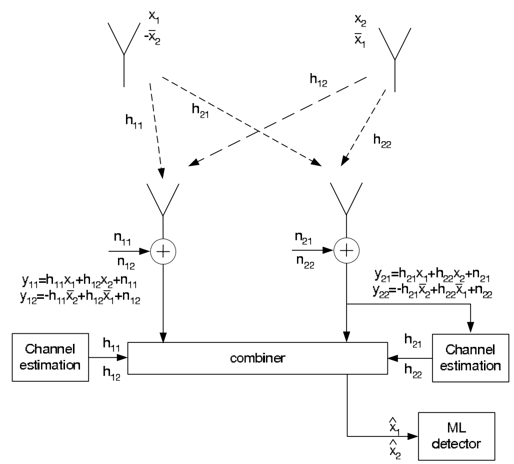


Figure 12: STBC with two transmitters two receivers.

$$y_{11} = h_{11} \cdot x_1 + h_{12} \cdot x_2 + n_{11} \quad y_{12} = -h_{11} \cdot \bar{x}_2 + h_{12} \cdot \bar{x}_1 + n_{12}$$
$$y_{21} = h_{21} \cdot x_1 + h_{22} \cdot x_2 + n_{21} \quad y_{22} = -h_{21} \cdot \bar{x}_2 + h_{22} \cdot \bar{x}_1 + n_{22}$$

$$\begin{aligned}\hat{x}_1 &= \bar{h}_{11} y_{11} + h_{12} \bar{y}_{12} + \bar{h}_{21} y_{21} + h_{22} \bar{y}_{22} \\ &= \sum_{i=1}^{q} \left(\bar{h}_{i1} y_{i1} + h_{i2} \bar{y}_{i2} \right) \\ &= \left(|h_{11}|^2 + |h_{12}|^2 + |h_{21}|^2 + |h_{22}|^2 \right) x_1 \\ &\quad + \bar{h}_{11} n_{11} + h_{12} \bar{n}_{12} + \bar{h}_{21} n_{21} + h_{22} \bar{n}_{22}\end{aligned}$$

$$\begin{aligned}\hat{x}_2 &= \bar{h}_{12} y_{11} - h_{12} \bar{y}_{12} + \bar{h}_{22} y_{21} - h_{21} \bar{y}_{22} \\ &= \sum_{i=1}^{q} \left(\bar{h}_{i2} y_{i1} - h_{i1} \bar{y}_{i2} \right) \\ &= \left(|h_{11}|^2 + |h_{12}|^2 + |h_{21}|^2 + |h_{22}|^2 \right) x_2 \\ &\quad + \bar{h}_{12} n_{11} - h_{11} \bar{n}_{12} + \bar{h}_{22} n_{21} - h_{21} \bar{n}_{22}\end{aligned}$$

Other Space Time Block codes

Example of rate 1/2 STBC

$$\mathrm{G}_3 = \begin{pmatrix} x_1 & x_2 & x_3 \\ -x_2 & x_1 & -x_4 \\ -x_3 & x_4 & x_1 \\ -x_4 & -x_3 & x_2 \\ \bar{x}_1 & \bar{x}_2 & \bar{x}_3 \\ -\bar{x}_2 & \bar{x}_1 & -\bar{x}_4 \\ -\bar{x}_3 & \bar{x}_4 & \bar{x}_1 \\ -\bar{x}_4 & -\bar{x}_3 & \bar{x}_2 \end{pmatrix} \qquad \mathbf{G}_4 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ -x_2 & x_1 & -x_4 & x_3 \\ -x_3 & x_4 & x_1 & -x_2 \\ -x_4 & -x_3 & x_2 & x_1 \\ \bar{x}_1 & \bar{x}_2 & \bar{x}_3 & \bar{x}_4 \\ -\bar{x}_2 & \bar{x}_1 & -\bar{x}_4 & \bar{x}_3 \\ -\bar{x}_3 & \bar{x}_4 & \bar{x}_1 & -\bar{x}_2 \\ -\bar{x}_4 & -\bar{x}_3 & \bar{x}_2 & \bar{x}_1 \end{pmatrix}$$

Example of rate 3/4 STBC for:
- 3 Tr antennas $\mathbf{H}_3$
- 4 Tr antennas $\mathbf{H}_4$

$$\mathbf{H}_3 = \begin{pmatrix} x_1 & x_2 & \frac{x_3}{\sqrt{2}} \\ -\bar{x}_2 & \bar{x}_1 & \frac{x_3}{\sqrt{2}} \\ \frac{\bar{x}_3}{\sqrt{2}} & \frac{\bar{x}_3}{\sqrt{2}} & \frac{(-x_1-\bar{x}_1+x_2-\bar{x}_2)}{2} \\ \frac{\bar{x}_3}{\sqrt{2}} & \frac{\bar{x}_3}{\sqrt{2}} & \frac{(x_2+\bar{x}_2+x_1-\bar{x}_1)}{2} \end{pmatrix}$$

$$\mathbf{H}_4 = \begin{pmatrix} x_1 & x_2 & \frac{x_3}{\sqrt{2}} & \frac{x_3}{\sqrt{2}} \\ -\bar{x}_2 & \bar{x}_1 & \frac{x_3}{\sqrt{2}} & -\frac{x_3}{\sqrt{2}} \\ \frac{\bar{x}_3}{\sqrt{2}} & \frac{\bar{x}_3}{\sqrt{2}} & \frac{(-x_1-\bar{x}_1+x_2-\bar{x}_2)}{2} & \frac{(-x_2-\bar{x}_2+x_1-\bar{x}_1)}{2} \\ \frac{\bar{x}_3}{\sqrt{2}} & -\frac{\bar{x}_3}{\sqrt{2}} & \frac{(x_2+\bar{x}_2+x_1-\bar{x}_1)}{2} & \frac{(x_1+\bar{x}_1+x_2-\bar{x}_2)}{2} \end{pmatrix}$$

MAP decoding for STBC

STBC decoder can output soft values.
These soft values can be used as input to the channel decoder.

$P(x_1,\ldots,x_k|y_{11},\ldots,y_{qn}) = P(y_{11},\ldots,y_{qn}|x_1,\ldots,x_k) \cdot P(x_1,\ldots,x_k)$

Where $P(x_1,\ldots,x_k)$ is apriori information about transmitted symbols and can be provided by decoder for example.
In Rayleigh fading channel we have:

$$P(y_{11},\ldots,y_{qn}|x_1,\ldots,x_k) = \frac{1}{(\sigma\sqrt{2}\pi)^{qn}} e^{-\frac{1}{2\sigma^2}\sum_{l=1}^{q}\sum_{i=1}^{n}\left| y_{li} - \sum_{j=1}^{p} h_{lj} g j i \right|^2}$$

$$P(x_i|y_{11},\ldots,y_{qn}) = P(y_{11},\ldots,y_{qn}|x_i)P(x_i)$$

MAP example for the Alamouti code

$$
\begin{aligned}
&P(x_1,\ldots,x_k|y_{11},\ldots,y_{qn})\\
=\ & C\cdot\frac{1}{(\sigma\sqrt{2}\pi)^{qn}}e^{-\frac{1}{2\sigma^2}\sum_{l=1}^{q}\left[\left|y_{l1}-\sum_{j=1}^{p}h_{lj}gj1\right|^2+\left|y_{l2}-\sum_{j=1}^{p}h_{lj}gj2\right|^2\right]}\\
=\ & C'\cdot\exp\left\{-\frac{1}{2\sigma^2}\sum_{l=1}^{q}\left[|y_{l1}-h_{l1}x_1-h_{l2}x_2|^2+|y_{l2}+h_{l1}\bar{x}_2-h_{l2}\bar{x}_1|^2\right]\right\}
\end{aligned}
$$

By conditioning with the x_i probability and considering orthogonality of the code:

$$
P(x_1|y_{11},\ldots,y_{q2})=C'\cdot\exp\left\{-\frac{1}{2\sigma^2}\sum_{l=1}^{q}\left[|y_{l1}-h_{l1}x_1|^2+|y_{l2}-h_{l2}\bar{x}_1|^2\right]\right\}
$$

$$
=C''\cdot\exp\left\{-\frac{1}{2\sigma^2}\sum_{l=1}^{q}\left[\begin{array}{c} h_{l1}x_1\bar{y}_{l1}-\bar{h}_{l1}\bar{x}_1y_{l1}\\ -h_{l2}\bar{x}_1\bar{y}_{l2}-\bar{h}_{l2}\bar{x}_1y_{l2}\\ +|x_1|^2\sum_{i=1}^{2}|h_{li}|^2\end{array}\right]\right\}
$$

$$
P(x_1|y_{11},\ldots,y_{q2})=
$$

$$
C\cdot\exp\left\{-\frac{1}{2\sigma^2}\left[\left|\sum_{l=1}^{q}\left(\bar{h}_{l1}y_{l1}+h_{l2}\bar{y}_{l2}\right)-x_1\right|^2+\left(-1+\sum_{l=1}^{q}\sum_{i=1}^{2}|h_{li}|^2\right)|x_1|^2\right]\right\}
$$

Simulations results from the book. Pages 415-438.

Rotated constellation

- Increase of the constellation size.
- The same information transmitted in many time intervals.
- If to use the sphere decoder the decoding complexity is not increased signinificantly.

The point x of the roated constellation is obtained by applying the rotation matrix $\mathbf{M}$ to the input signal u.

$$
x=\mathbf{M}u
$$

The matrix $\mathbf{M}$ has to be selected such that it maximizes the minimum product distance

$$
d=\min_{s=\mathbf{M}(u-u'),u\neq u'}\prod_{j=1}^{m}|s_j|
$$

One possible policy is to use as $\mathbf{M}$ the lattice generating matrice.

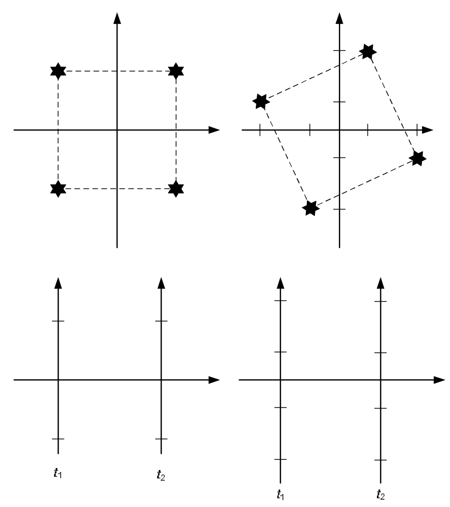


Figure 13: Example of the constellation rotation.

Allocation of the rotated symbols for achieving maximum diversity in MIMO

- The maximum diversity is achieved when the symbol is passed trough each possible channel.
- Rotation should have the dimension equal to the number of transmitters.
  - if we have m transmitters then we have $\mathbf{M} = m \times m$ size rotation matrix.
  - We group the bits into groups of m bits.
  - Convert each group by rotating with $\mathbf{M}$.
  - Map each symbol inside the group into different transmission antennae.

Full diversity full rate codes

Maximum amount of parallel rates that can be separated at the receiver is equal to amount of rank of the channel matrix.
If ralayigh fading iid channel it can be assumed to be equal to min $N_t x, M_t x$

- Split the information to min N_{tx}, M_{tx} parallel streams.
- Group and rotate the bits in each stream.
- Map each stream to the transmission antennaes so that the symbols do not overlap.
- The allocation of the bits to the antennae is called a thread.

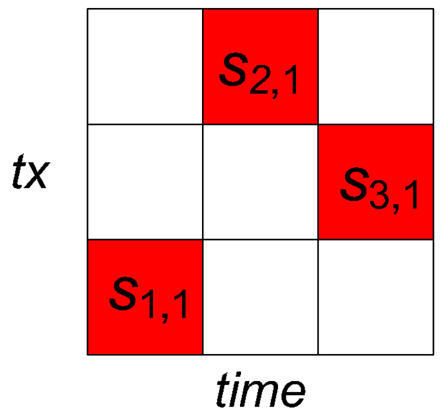


Figure 14: Example of the thread in the system with 3 transmition antennaes.

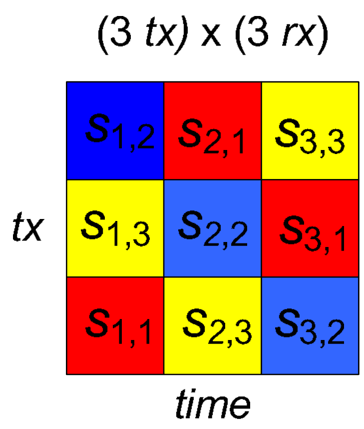


Figure 15: Allocation of the threads 3 × 3 system

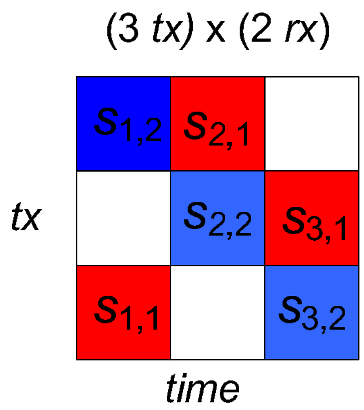


Figure 16: Allocation of the threads 3 × 2 system

- Transformation of the inoput bits $u \rightarrow$ for maintaining the rate inside the thread.
- Allocation of the bits to the thread $\rightarrow$ for achieving full diversity inside the stream.
- Allocation of different streams to different treads $\rightarrow$ for achieving full rate.
  - In order to maintain the rank criteria between the Space-Time code made by treading each thread has to be multiplied with an irrducible number (algebraic number).

Sphere decoder

- High number of possible constellation points?
- Optimal decoder requires evaluation of the distance from all possible constellation points $x \in C$ to the received points y.
- Suboptimal decoder selects only few of the nearest points.
- Efficient algorithm for finding the nearest points to a vector is Sphere deocder.

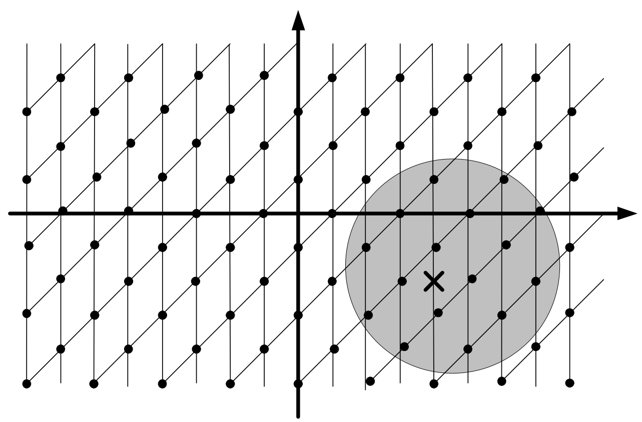


Figure 17: Points considered in a sphere decoder.

Sphere decoder enumerates all the points of the lattice that fall inside of the sphere.

The received point at moment t

$$y_t = \mathbf{HM}u_t + \nu_t$$

Maximum Likelihood detection in the MIMO channel

$$\hat{x} = \arg\min_{x \in C} |y - \mathbf{B}x|^2$$

where

$\hat{x}$ is most likely received constellation point.
x describes the possible constellation points.
y is the received signal value
$\mathbf{B}$ describes the channel matrix that has real values.

The complex channel matrix $\mathbf{H}$ is transferred to the matrix with real elements

$$\mathbf{B} = \begin{bmatrix} \mathrm{Re}\{\mathbf{H}\} & -\mathrm{Im}\{\mathbf{H}\} \\ \mathrm{Im}\{\mathbf{H}\} & \mathrm{Re}\{\mathbf{H}\} \end{bmatrix}$$

The sphere decoder attempts to find all the points inside the n dimensional spere C_0

For finding the distance of the received point from the possible constellation point one has to solve matrix equation

$$|y - Bx|^2$$

The solving of the equation can be siplified if we express the matrix **B** in the trianqualar form.

$$\mathbf{B} = [\mathbf{Q}, \mathbf{Q}'] \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix}$$

where **R** is an upper triangular matrix.

$$|y - Bx|^2 \leqslant C_0$$

$$\left| [\mathbf{Q}, \mathbf{Q}']^T y - \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix} x \right|^2 \leqslant C_0$$

$$\left| \mathbf{Q}^T y - \mathbf{R}x \right|^2 \leqslant C_0 - \left| (\mathbf{Q}')^T y \right|^2$$

$$|y' - \mathbf{R}x|^2 \leqslant C_0'$$

Since matrix R is upper triangular the equation can be simplified

$$\sum_{j=1}^{m} \left| y' - \sum_{j_1=j}^{m} r_{j,j_1} x_{j_1} \right|^2 \leqslant C_0'$$

The porobability of the point is described by the distance

$$d^2(y, \mathbf{B}x) = \sum_{j=1}^{m} \left| y' - \sum_{j_1=j}^{m} r_{j,j_1} x_{j_1} \right|^2$$

- Select all the points inside one sphere with Radius C.
- The sphere decoder implements the tree search.
- Identify the possbile constellation points coordinate by coordinate. $\mathbf{x} = x_1, x_2, .., x_n$.

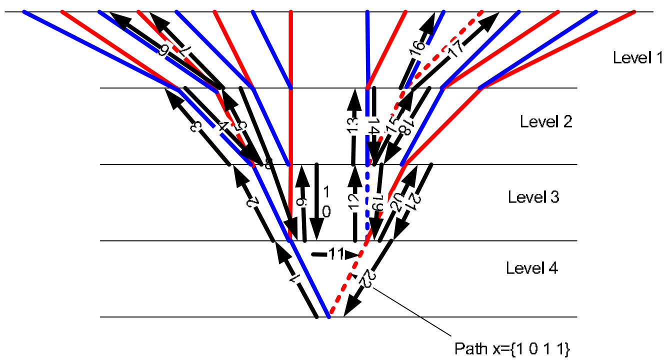
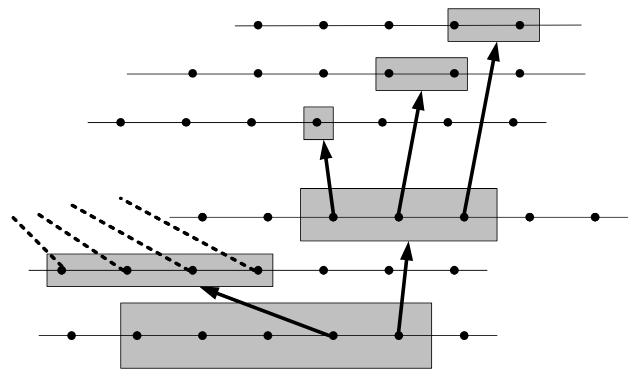


Figure 18: The search path of an spehre decoder.

We go trough all the branches and attempt to find the closest point to the received point.

If in some branch at some level the distance exceeds the distance of the sphere C_0 we drop further explanation of that branch and contiunue with the neighbouring branch.



Soft values from the sphere decoder

- If to use turbo decoder after the sphere decoder one needs soft values.
- Straight forward way is to use as soft values the marginal probabilities for each bit.
- Simplification: to use only the probabilities of the symbols with the maximum value.

  If the bit has representation of only one of the two binary values, for example only the paths where $x_i = 1$

  We have to evaluate new path where this bit has also alternative value, a constellation point where $x_i = 0$.

  The soft value can be approximated with the difference of the probabilities of the constellation points where $\frac{p(x_i=0)}{p(x_i=1)}$

Mapping threaded code for sphere decoding (system description)

Input information symbol vector for the thread j

$$\mathbf{u}_j = \begin{pmatrix} u_1, & ..., & u_M \end{pmatrix}^T$$

The symbols for the thread j

$$\gamma_j(\mathbf{u}_j) = \phi_j \mathbf{s}_j = \phi_j \mathbf{M} \mathbf{u}_j$$

Mapping the thread is described by a mapping function

$$T_{M,L,R}$$

where

M stands for the number of transmit antennaes.

L stands for the amount of layers (streams).

R describes the nuber of symbols transmitted per layer per cahnnel use.

Mapping threaded code for sphere decoding

We transfer the system to a new system into form

$$y = \mathrm{HM}\mathbf{u} + w$$

where rmHM represents a new channel matrix with the nm equations and lm unknows.

The received signal at the dector output is

$$Y = HT_{M,L,R} + W$$

Where Y is $m \times L$ matrix describing all the symbols received during the block.

The received vector Y can be described as the sum of the symbols from eahc tread weighted with corresponding channel amplitude and summed the gaussian noise W

$$= \sum_{j=1}^{L} H_j \phi_j \mathrm{diag}(s_{j1}, ..., s_{jM}) + W$$

The matrix H can be modified by stacking its columns

$$\mathrm{H}_j = (\mathrm{diag}(h_{j1}), ..., \mathrm{diag}(h_{j1}))^T$$

For example for the thread j in 2×2 system

$$\mathrm{H}_j = \begin{bmatrix} h_{11,j} & h_{12,j} \\ h_{21,j} & h_{22,j} \end{bmatrix} \Rightarrow \begin{bmatrix} h_{11,j} & 0 \\ 0 & h_{12,j} \\ h_{21,j} & 0 \\ 0 & h_{22,j} \end{bmatrix}$$

The received signal y can be expressed

$$y = \sum_{j=1}^{L} \mathrm{H}_j \phi^{\frac{L-1}{M}} \mathbf{M} u_j + w$$

Where y is the vector generated by stacking the columns of Y under each other.

Similarly from W is generated the vector w.

After rearranging

$$y = \left(\mathrm{H}_1, ..., \phi^{\frac{L-1}{M}} \mathrm{H}_L\right) I_L \otimes \mathbf{Mu} + w$$
$$= \mathrm{HM}\mathbf{u} + w$$

where

$$\mathrm{M} = I_L \otimes \mathbf{M}$$

Space Time Trellis Coding (STTC)

STTC incorporte jointly
- channel coding,
- modulation,
- transmit diversity,
- optimal receiver diversity.

Example of STTC

The STTC transmits symbols $x_{k,1}$ and $x_{k,2}$ over the transmit antennas $Tx1$ and $Tx2$.
The output symbols are generated from the input data

$$x_{k,1} = 0 \cdot d_{k,1} + 0 \cdot d_{k,2} + 1 \cdot d_{k-1,1} + 2 \cdot d_{k-1,2}$$
$$x_{k,2} = 1 \cdot d_{k,1} + 2 \cdot d_{k,2} + 0 \cdot d_{k-1,1} + 0 \cdot d_{k-1,2}$$

The 4-State trellis codes

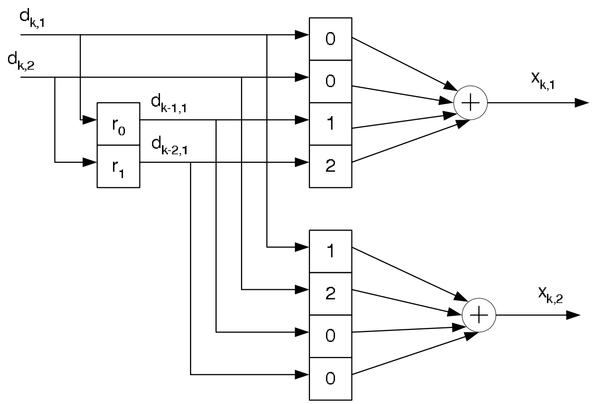


Figure 19: 4-state 4PSK STTC encoder.

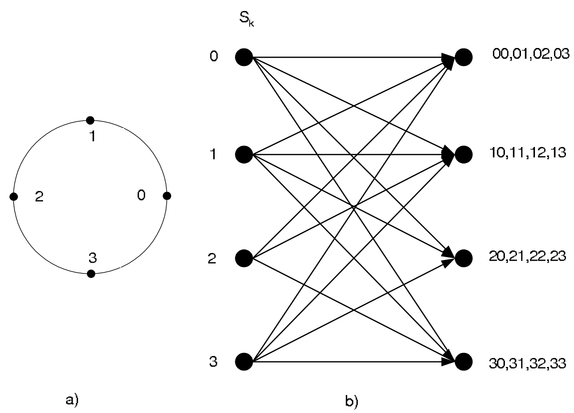


Figure 20: a) The 4PSK constellation points b) 4-state ST trellis.

Example

Input	k	Input bits $(d_{k,1};d_{k,2})$	shift register $(d_{k-1,1};d_{k-1,2})$	state S_k	tranmitted symbols $(x_{k,1};x_{k,2})$
00101101	0	–	00	0	–
001011	1	10(1)	00	0	01
0010	2	11(3)	10	1	12
00	3	01(2)	11	3	32
	4	00(0)	01	2	20
	5	–	00	0	–

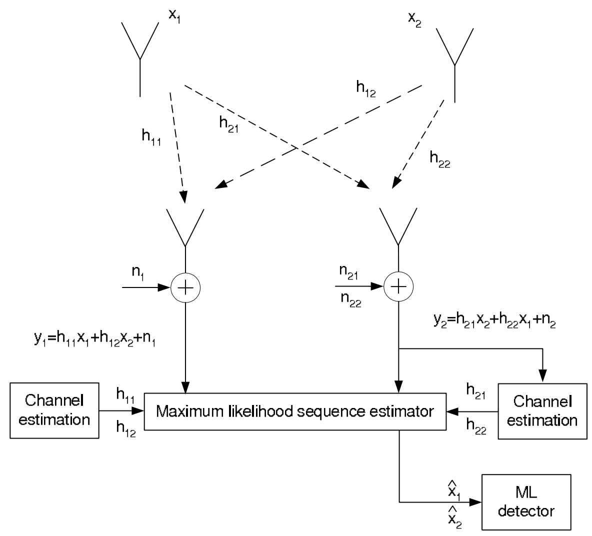


Figure 21: 4 state 4PSK STTC system.

Space Tome Trellis Decoder

At any time instant we have at receiver:

$$y_1 = h_{11}x_1 + h_{12}x_2 + n_1$$
$$y_2 = h_{21}x_1 + h_{22}x_2 + n_2$$

Where $h_{11}, h_{12}, h_{21}, h_{22}$ represents corresponding complex time-domain channel transfer factors.
The receiver applies the Viteri-algorithm based maximum likelihood sequence estimator for the code trellis

- Finds the branch metric for every transition in the trellis,
- Runs the Viterbi algorithm to find the ML path in the trellis.

For each trellis transitions we have two estimated transmit symols,$\hat{x}_1$,$\hat{x}_2$
The branch metrics is

$$BM = |y_1 - h_{11}\hat{x}_1 - h_{12}\bar{x}_2|^2 + |y_2 - h_{21}\hat{x}_1 - h_{22}\hat{x}_2|^2$$
$$= \sum_{i=1}^{2}\left|y_i - \sum_{j=1}^{2} h_{ij}\hat{x}_j\right|^2$$

The generalised form for p transmitters and q receivers is:

$$BM = \sum_{i=1}^{p}\left|y_i - \sum_{j=1}^{q} h_{ij}\hat{x}_j\right|^2$$

STTC Complexity

- The complexity of TC is expressed as number of trellis transmissiotions per information data bit.
- The complexity expresses as function of number of trellis states in STTC decoder.
- Number of trellis transitions leaving of each state is equivalent to 2^{BPS}

  BPS - the number of transmitted bits per modulation symbol.

$$comp\{STT\} = \frac{2^{BPS} \times \text{No. of States}}{BPS} = 2^{BPS-1} \times \text{No. of States}$$