SPACE TIME CODED MODULATION (STCM)

Reference: "Space-Time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction"
Vahid Tarokh, Nambi Sheshadri and A.R. Calderbank

IEEE Transactions on Information Theory, Vol. 44, No. 2, March 1998, pp. 744-765.

" A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation", Wittneben A., ICC' 1993, pp. 1630-1634.

**Principles of Communication Engineering*, Wozencraft and Jacobs
**Statistical Communication and Modulation Theory Part I*, H. L. Van Trees
- for Diversity Principles

A suitable Matrix Theory book

Background:

Consider the case involving both fading and AWGN, in which the signal ,s(t), represents the **single** transmission of one of M equally likely lowpass signals $\{s_i(t)\}$. The total received signal is therefore

$$r(t) = a\, s(t)\, \sqrt{2}\cos(w_0 t - q) + n(t)$$

where θ is a random variable with a probability density function p_θ that is uniform over the interval $[0,2\pi]$ and *a* is statistically independent of θ and

$$p_a(\alpha) = \frac{2a}{b} e^{-a^2/b} \quad a \geq 0$$

i.e., the transmission gain is Rayleigh distributed.

This equation is derived from the scattering model for a fading channel which leads to a received signal with a uniformly distributed phase and Rayleigh-distributed amplitude.

For the receiver we observe that if *a* were known, the optimum receiver would simply act as if the modulating signal set had been $\{as_i(t)\}$ and it can be shown that we have to determine that *i* for which,

$I_0(\frac{2aX_i}{N_0})\exp(-a^2E_i/N_0)$ is maximum. $I_0(x)$ is called the "zero-order modified Bessel function of the first kind".

Here E_I denotes the energy of $s_I(t)$ and

$$X_i = \left\{ \left[\int_{-\infty}^{\infty} r_c(t)s_i(t)dt \right]^2 + \left[\int_{-\infty}^{\infty} r_s(t)s_i(t)dt \right]^2 \right\}^{1/2}$$

$r_c(t) = s_i(t)\cos\theta + n_c(t)$ $\qquad$ $r_s(t) = s_i(t)\sin\theta + n_s(t)$

$n_c(t)$, $n_s(t)$ are statistically independent Gausssian processes each having a power spectrum that is uniform (with density $N_0/2$) over the rectangular frequency band [-W,W] occupied by $s_I(t)$. Since noise power outside the band occupied by the signals does not affect the performance of the optimum receiver, we may assume $n_c(t)$, $n_s(t)$ to be white.

X_i may be identified as the sampled envelope of a bandpass filter matched to

$s_i(t)\sqrt{2}\cos w_0 t$. When all E_i are equal, i=0,1,…M-1, the decision implied by this rule is the same regardless of the specific (positive) value assumed by *a*(random variable).

It can be seen that the probability of error for M=2 and orthogonal signals of energy E_s is given by

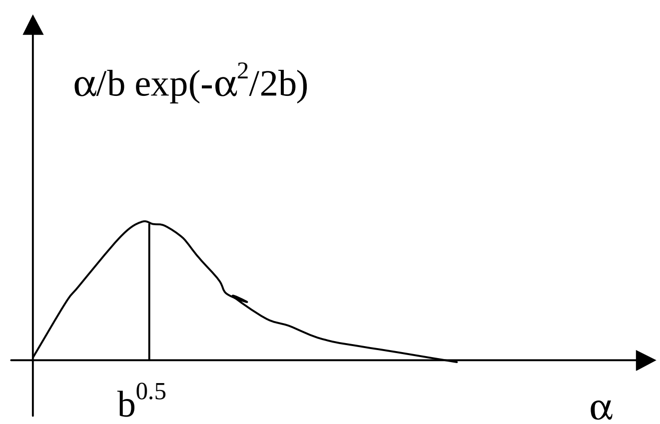
$$P[Error] = \overline{P[E\,|\,a]} = \overline{\frac{1}{2}\exp(-a^2E_s/2N_0)} = \frac{1}{2+\overline{E_s}/N_0}$$

where $\overline{E_s}$ =bE_s is the mean value of the received energy.

This shows that minimum attainable error probability in communicating one of two equally likely orthogonal signals decreases only inversely with the transmitted energy.

This behavior is in marked contrast to the nonfading case, in which the error probability decreases exponentially with E_s.

The difference in performance is attributable to the fact that even when the *average* received energy on a fading channel is high there is still an appreciable probability that the actual energy received on any given transmission is quite small; that is, there is an appreciable probability of a "**deep fade**". This is quite clear in the plot of the Rayleigh pdf.



<u>Diversity Transmission</u>

The only efficient way to reduce error probability with a Rayleigh fading channel is to circumvent the high probability of a deep fade on a single transmission. This is accomplished by means of *diversity* transmission. The idea of diversity is simple: scattering channels of practical interest are characterized by the fact that the scattering elements move randomly with respect to one another as time goes on. So we can use

1. **Time diversity** – involves sending the same signal s(t) over and over again, say L times, in the hope that not all transmissions are subject to deep fades. So space successive transmissions in time in such a way that

the fading experienced by each transmission is statistically independent – one way to achieve this is to use an *interleaver*.

It can be shown that an upper bound for the error probability can be given by, assuming that there are L transmissions and $\{b_l\}$ are all equal with E_b *energy per message input bit* is divided equally among the transmissions,

$$E_s = E_b / L \quad h = \frac{\overline{E_s}}{N_0} = \frac{bE_b / L}{N_0} = \frac{\overline{E_b} / L}{N_0}$$

$$P[\varepsilon] \le \left[4 . \frac{1+h}{(2+h)^2} \right]^L = \exp\left[-\frac{\overline{E_b}}{N_0} (\ln 2) g(h) \right]$$

It follows that the probability of error may be made to decrease exponentially with $\frac{\overline{E_b}}{N_0}$ by means of diversity, even though the channel is subject to fading. It also shows that there is a minimum value to be attained with a certain value for L » $\frac{1}{3} . \frac{\overline{E_b}}{N_0}$.

$$P[\varepsilon] \le 2^{-0.215 \overline{E_b} / N_0}$$

For this value it is seen that compared to *nonfading* Gaussian channel with *unknown* phase where,

$$P[\varepsilon] = \frac{1}{2} \exp\left(-\frac{E_b}{2N_0}\right) \approx 2^{-0.72 E_b / N_0}$$

fading costs approximately 5 dB in signal energy.

The reason why an optimum value of L exists is simple: on the one hand, as L increases with the total energy held fixed, the average signal-to-noise ratio at the output of a bandpass filter matched to $s(t) \sqrt{2} \cos(w_0 t)$ decreases and the loss introduced by incoherent reception becomes larger. On the other hand, increasing L provides additional diversity and the decreases the probability that most of the transmissions are badly faded.

The optimum value of *L* reflects the best compromise between these two effects.

Other types of diversity techniques are:

2. **Frequency diversity** – justifies an assumption of statistical independence. Here there is only one receiver, but the lowpass signal simultaneously modulates *L* different carriers, each having a different frequency – *should consider implications in OFDM*

3. **Space Diversity** – single transmitted signal with *L* dispersed receivers. It is usually reasonable to assume transmission gains to different receivers will be statistically independent whenever the separation between sites is many carrier wave lengths.
   - Additional design freedom is introduced when signal diversity is considered.

Although diversity may be used to obtain a probability of error that decreases exponentially as $\frac{\overline{E_b}}{N_0}$ is increased, an arbitrarily small P[ε] can be obtained this way only by making $\frac{\overline{E_b}}{N_0}$ arbitrarily large. Just as in the unfaded case, **<u>coding</u>** provides a method of making P[ε] as small as we please without significant increase in transmitted energy per bit, provided that $\frac{\overline{E_b}}{N_0}$ exceeds a minimum threshold.

Matrix Theory- relevant definitions:

Let $A=(a_{ij})$ be a matrix in C_{mxn}. We denoteby R_i the *i*th row of the matrix. That is,

$R_i = [a_{i1}, a_{i2},\ldots, a_{ij},\ldots a_{in}]$ $i=1,2,\ldots,m$

The *j*th column of the matrix A is denoted by C_j. Thus

$$C_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \cdot \\ \cdot \\ a_{mj} \end{bmatrix} \quad j=1,2,\ldots,n$$

Note that the rows R_i, $i=1,2,..,$m and the columns C_j,$j=1,2,\ldots,n$ are in C_{1xn} and C_{mx1} *r*espectively. These *m* rows and *n* columns are also *called row vectors* and *columns vectors* of the matrix A.

Definition: If A=(a$_{ij}$) is in C_{mxn}, the *subspace of* C_{1xn} generated by the row vectors R_i, i= 1,2,…,m, is called the *row space* of A. We denote the row space of A by R(A). Thus

$$R(A) = \text{á}R_1, R_2,\ldots., R_m\,\text{ñ}$$

Similarly, the subspace of C_{mx1} generated by the *n* column vectors C_j,$j=1,2,\ldots,n$, is called the *column space* of A. The column space is denoted by $C(A)$, and thus

$$C(A) = \text{á}C_1, C_2,\ldots., C_n\,\text{ñ}$$

A given matrix A=(a_{ij}) is transformed to a row (column) equivalent matrix using the three elementary row (column) operations.

Theorem: Two row (column) – equivalent matrices have the same row (column) space.

e.g. Transform the following matrix into

i) a row- equivalent matrix,

ii) a column equivalent matrix (left as an exercise)

using the **Gauss-Jordan reduction** method.

$$A = \begin{bmatrix} 0 & 1 & 1 & -2 & 1 \\ 1 & 2 & 3 & -4 & 1 \\ 2 & 0 & 2 & 0 & -2 \end{bmatrix}$$

i) Applying elementary row operations on the rows of A we have

$$A \xrightarrow{R_{12}} \begin{bmatrix} 1 & 2 & 3 & -4 & 1 \\ 0 & 1 & 1 & -2 & 1 \\ 2 & 0 & 2 & 0 & -2 \end{bmatrix}$$

$$\xrightarrow{-2R_1+R_3} \begin{bmatrix} 1 & 2 & 3 & -4 & 1 \\ 0 & 1 & 1 & -2 & 1 \\ 0 & -4 & -4 & 8 & -4 \end{bmatrix}$$

$$\xrightarrow{-2R_2+R_1, 4R_2+R_3} \begin{bmatrix} \bar{1} & 0 & 1 & -0 & -1 \\ 0 & \bar{1} & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = E_R$$

This last matrix E_R is in **reduced row echelon** form and is row equivalent to the matrix A.

The matrix E_R has the leading 1's marked by $\bar{1}$. In E_R

(a) A row with all zero entries appears below any other row with nonzero entries.

(b) The first nonzero entry in each nonzero row is 1.

(c) The leading 1 in any nonzero row appears in a column to the right of any leading 1 in a preceding row.

(d) The leading 1 in each nonzero row is the only nonzero entry in its column.

Definition: The *row(column) rank* of a matrix A is the dimension of the row(column) space of A.

In the earlier example, the row rank of A is 2, and the column rank of A is 2 also. In fact, the row rank of A is equal to the number of nonzero rows in E_R, and the column rank of A is equal to the number of nonzero columns in E_C. Clearly the row space of A is

$$R(A) = \langle [1\ \ 0\ \ 1\ \ 0\ \ -1], [0\ \ 1\ \ 1\ \ -2\ \ 1] \rangle$$

Theorem: If $A=(a_{ij})$ is in C_{mxn}, the row rank of A is equal to its column rank. This common rank is called simply the *rank of the matrix* A, and is denoted by *r(A)*.

Nullspace of a Matrix

Let $A=(a_{ij})$ be an mxn matrix. We denote by N(A) the set of all solutions of the homogeneous linear system of equations $Ax=0$.Thus,

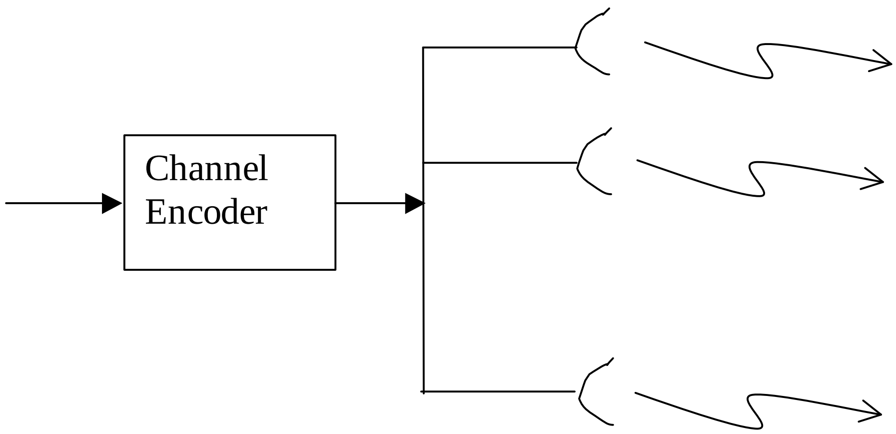
$N(A) = \{x | x \in C_{nx1} \text{ and } Ax = 0\}$.

It is easy to show that N(A) is indeed a subspace of C_{nx1}. The subspace N(A) is called the *nullspace* or *kernel* of matrix A, and the dimension of N(A) is called the *nullity* of A and is denoted by $n(A)$.

Theorem: Let $A=(a_{ij})$ be an mxn matrix of **rank** t. Then
(a) dim $N(A) = n(A) = n - t$
(b) dim $N(A^T)=m - t$

Performance Criteria

Consider a mobile communication system where the base station is equipped with *n* antennas and the mobile is equipped with *m* antennas.



Transmitter – Base station

The encoded data is divided in to *n* parallel streams to be fed to the antennas and transmitted simultaneously, each signal from a different antenna. All signals have the same transmission period T.

At the receiver, the signal d_t^j received by **antenna *j* at time *t*** is given by

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

h_t^j at tine t is modeled as independent samples of a zero-mean complex Gaussian random variable with variance $N_0/2$ per dimension.

The coefficient $\alpha_{i,j}$ is the path gain from transmit antenna *i* to receive antenna *j*.

It is assumed that these path gains $a_{i,j}$ $i=1,\ldots,n;j=1,2,\ldots,m$ are **constant during a frame** and vary from one frame to another (quasistatic flat fading).

Independent fade coefficients: Design criterion

Here we assume that the coefficients $\alpha_{i,j}$ are modeled as independent samples of complex Gaussian random variables with variance 0.5 per dimension. $\Rightarrow$ signals transmitted from different antennas undergo independent fades.

Let us also assume that each element of the signal constellation is contracted by a scale factor $\sqrt{E_s}$, chosen so that the average energy of the constellation elements is 1.

We consider the probability that a maximum-likelihood receiver decides erroneously in favor of a signal

$$\boldsymbol{e} = e_1^1 e_1^2 \ldots\ldots e_1^n \; e_2^1 e_2^2 \ldots\ldots e_t^i \ldots\ldots e_l^1 e_l^2 \ldots\ldots e_l^n$$

Assuming that

$$\boldsymbol{c} = c_1^1 c_1^2 \ldots\ldots c_1^n \; c_2^1 c_2^2 \ldots\ldots c_t^i \ldots\ldots c_l^1 c_l^2 \ldots\ldots c_l^n$$

Was transmitted.

The signal components at the receiver for these two sequences for jth antenna for the tth time interval are:

$\sum_{i=1}^{n} a_{i,j} c_t^i \sqrt{E_s}$, $\sum_{i=1}^{n} a_{i,j} e_t^i \sqrt{E_s}$ respectively. The fading coefficients $\boldsymbol{a}_{i,j}$ $i=1,\ldots,n; j=1,2,\ldots,m$ do not depend on t (assumed constant for a frame under quasi-static condition for fading).

The distance is then

$$\left|\sum_{i=1}^{n} a_{i,j} \left(c_t^i - e_t^i\right)\sqrt{E_s}\right|^2, \ j=1,2,\ldots,m$$

The total distance, taking all antennas into account:

$$\sum_{j=1}^{m} \left|\sum_{i=1}^{n} a_{i,j} \left(c_t^i - e_t^i\right)\sqrt{E_s}\right|^2$$

Summing for time $t=1,\ldots,l$

$$\sum_{t=1}^{l} \sum_{j=1}^{m} \left|\sum_{i=1}^{n} a_{i,j} \left(c_t^i - e_t^i\right)\sqrt{E_s}\right|^2 .$$

Now, assuming ideal channel state information (CSI), the probability of transmitting ***c*** and deciding in favor of ***e*** at the decoder is upperbounded by

$$P(\mathbf{c} \rightarrow \mathbf{e} \mid a_{i,j}, i=1,2,\ldots,n, j=1,2,\ldots,m) \leq \exp\left(-d^2(\boldsymbol{c},\boldsymbol{e}) E_s / 4N_0\right)$$

where $N_0/2$ is the noise variance per dimension and

$$d^2(\mathbf{c},\mathbf{e}) = \sum_{j=1}^{m}\sum_{t=1}^{l}\left|\sum_{i=1}^{n} a_{i,j}(c_t^i - e_t^i)\right|^2$$ taking E_s out.

$\left|\sum_{i=1}^{n} a_{i,j}\left(c_t^i - e_t^i\right)\right|^2$ can be written as $\sum_{i=1}^{n}\sum_{i'=1}^{n} a_{i,j}\overline{a_{i',j}}\left(c_t^i - e_t^i\right)\overline{\left(c_t^{i'} - e_t^{i'}\right)}$

where $\bar{x}$ is the complex conjugate of x.

Thus $$d^2(\mathbf{c},\mathbf{e}) = \sum_{t=1}^{l}\sum_{j=1}^{m}\sum_{i=1}^{n}\sum_{i'=1}^{n} a_{i,j}\overline{a_{i',j}}(c_t^i - e_t^i)\overline{(c_t^{i'} - e_t^{i'})}$$

This can be rewritten as

$$d^2(\mathbf{c},\mathbf{e}) = \sum_{j=1}^{m}\sum_{i=1}^{n}\sum_{i'=1}^{n} a_{i,j}\overline{a_{i',j}}\sum_{t=1}^{l}(c_t^i - e_t^i)\overline{(c_t^{i'} - e_t^{i'})}$$

Setting $\Omega_j = (a_{1,j}, a_{2,j}, \ldots, a_{n,j})$ we can rewrite:

$$d^2(\mathbf{c},\mathbf{e}) = \sum_{j=1}^{m}\Omega_j A \Omega_j^*$$ Ω_j^* is the *transpose conjugate*

where $A_{pq} = \boldsymbol{x_p}.\boldsymbol{x_q}$ and $\boldsymbol{x_p} = (c_1^p\text{-}e_1^p, c_2^p\text{-}e_2^p, \ldots, c_l^p\text{-}e_l^p)$ for $1 \le p, q \le n$.

$A_{p,q} = \sum_{t=1}^{l}(c_t^p - e_t^p)\overline{(c_t^q - e_t^q)}$ and A is an nxn matrix, regardless of $l({}^{3}1)$.

Therefore,

P(**c** ® **e** | $a_{i,j}$, *i=1,2,…..,n, j=1,2,…..m*)

$$\leq \prod_{j=1}^{m} \exp(-\Omega_j A(\mathbf{c},\mathbf{e})\Omega_j^* E_S / 4N_0)$$

Since $A(\boldsymbol{c},\boldsymbol{e})$ is Hermitian, i.e, $A=A^*$, there exists a unitary matrix $V(VV^* = I)$, and a real diagonal matrix D such that

$V\,\mathrm{A}(\mathbf{c},\mathbf{e})V^* = D$Þ $\mathrm{A}(\mathbf{c},\mathbf{e}) = V^*DV$

The rows $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\ldots,\boldsymbol{v}_n\}$ of V are a complete orthonormal basis of $\mathbf{C}^n$ given by the *eigenvectors* of *A*. Furthermore the diagonal elements of *D* are the *eigenvalues*

l_i, $i=1,2,\ldots,n$ of A including multiplicities.

According to the construction of $A(\boldsymbol{c},\boldsymbol{e})$, the matrix

$$B(\boldsymbol{c},\boldsymbol{e}) = \begin{bmatrix} e_1^1 - c_1^1 & e_2^1 - c_2^1 & . & . & . & e_l^1 - c_l^1 \\ e_1^2 - c_1^2 & e_2^2 - c_2^2 & . & . & . & e_l^2 - c_l^2 \\ . & & & & & . \\ . & & & & & . \\ . & & & & & . \\ e_1^n - c_1^n & e_2^n - c_2^n & . & . & . & e_l^n - c_l^n \end{bmatrix}$$

is clearly a square root of $A(\boldsymbol{c},\boldsymbol{e})$. i.e.,

$A(\boldsymbol{c},\boldsymbol{e}) = B(\boldsymbol{c},\boldsymbol{e}).\ B^*(\boldsymbol{c},\boldsymbol{e})$

Thus the *eigenvalues* of $A(\boldsymbol{c},\boldsymbol{e})$ are nonnegative (could be zero) real numbers. Next, we express $d^2(\mathbf{c},\mathbf{e})$ in terms of the *eigenvalues of the matrix* $A(\boldsymbol{c},\boldsymbol{e})$.

Let ($b_{1,j}, b_{2,j},\ldots,b_{n,j}$) = $\Omega_j V^*$, then

$$\Omega_j A(\boldsymbol{c},\boldsymbol{e})\, \Omega_j^* = \Omega_j V^* D V\, \Omega_j^* = \sum_{i=1}^{n} l_i \,|\, b_{i,j} \,|^2$$

Since $a_{i,j}$ are samples of a complex Gaussian random variable with mean $Ea_{i,j}$, let

$$\boldsymbol{K}^j = (Ea_{1,j}, Ea_{2,j}, \ldots\ldots, Ea_{n,j})$$

Since V is *unitary* $\Rightarrow$ $b_{i,j}$ are independent complex Gaussian random variables with variance 0.5 per dimension and with mean $\boldsymbol{K}^j.\ \boldsymbol{v}_{i.}$.

Let $K_{i,j} = |Eb_{i,j}|^2 = |\,\boldsymbol{K}^j.\ \boldsymbol{v}_i\,|^2$. Thus $|b_{i,j}|$ are independent Rician distributions with *pdf*

$$p(|\, b_{i,j} \,|) = 2\,|\, b_{i,j} \,| \exp(-|\, b_{i,j} \,|^2 - K_{i,j}) I_0(2\,|\, b_{i,j} \,| \sqrt{K_{i,j}}) \text{ for } |b_{I,j}| \geq 0$$

where $I_0(.)$ is the zero-order modified Bessel function of the first kind.

To get an upper bound on the average probability of error, we average

$$\prod_{j=1}^{m} \exp[-(E_s / 4N_0) \sum_{i=1}^{n} l_i \,|\, b_{i,j} \,|^2]$$

with respect to independent Rician distributions of $|b_{i,j}|$:

$$\text{P}(\boldsymbol{c}\text{® } \boldsymbol{e}) \le \prod_{j=1}^{m}\left(\prod_{i=1}^{n}\frac{1}{1+\frac{E_s}{4N_0}\,l_i}\exp\left[-\frac{K_{i,j}\frac{E_s}{4N_0}\,l_i}{1+\frac{E_s}{4N_0}\,l_i}\right]\right)$$

The Rayleigh fading case

In this case $Ea_{i,j} = 0$ for all i and j. Therefore the above inequality becomes

$$\text{P}(\boldsymbol{c}\text{® } \boldsymbol{e}) \le \left(\frac{1}{\prod_{i=1}^{n}(1+\frac{E_s}{4N_0}\,l_i)}\right)^{m}$$

Let *r* denote the rank of matrix *A*(*c*,*e*) or *A*, then the *kernel* of *A* has dimension *n-r* and exactly *n – r* eigenvalues of *A* are zero.

If the nonzero eigenvalues of *A* are l_1, l_2,…. l_r. Thus the inequality becomes assuming , $1+\frac{E_s}{4N_0}\,l_i \approx \frac{E_s}{4N_0}\,l_i$:

$$\text{P}(\boldsymbol{c}\text{® } \boldsymbol{e}) \le \left(\prod_{i=1}^{r} l_i\right)^{-m}\left(\frac{E_s}{4N_0}\right)^{-rm} = \left(\prod_{i=1}^{r} l_i^{1/r}\right)^{-rm}\left(\frac{E_s}{4N_0}\right)^{-rm}$$

Thus a diversity advantage of *mr* (from signal to noise ratio) and a coding advantage of (l_1. l_2…. l_r)$^{1/r}$ (due to properties of the coding sequences) is achieved. The coding advantage is an approximate measure of gain over an uncoded system having the same diversity advantage.

∗ l_1 . l_2…. l_r is the absolute value of the sum of determinants of **all** the *principal r* x *r cofactors of A*. **Also the *ranks* of *A* and *B* are equal.**

<u>Design Criterion for Rayleigh Space-Time Codes</u>

- **The rank Criterion:** In order to achieve the maximum diversity *mn*, the matrix *B(**c,e**)* has to be full rank for any codewords ***c*** and ***e***. If *B(**c,e**)* has minimum rank *r* over the set of two tuples of **distinct** codewords, then a diversity of *rm* is achieved.

- **The Determinant Criterion:** Suppose that a diversity benefit of *rm* is our target. The **minimum** of *r*th roots of the **sum** of determinants of all *r* x *r* principal cofactors of *A(**c,e**)= B(**c,e**) B*(**c,e**)* **taken over all pairs of 'distinct' codewords *c*** and ***e*** corresponds to the coding advantage. *r* is the rank of *A(**c,e**)*. Special attention in the design must be paid to make this **sum** as large as possible.
  If a diversity of *mn* is the design target, then the **minimum 'sum' taken over all pairs of distinct *c*** and ***e*** must be maximized.

Code construction for Quasi-Static Flat Fading Channels

- *at the beginning and the end of each frame, the encoder is required to be in the zero state.*

- ***at each time t, depending on the state of the encoder and the input bits, a transition branch is chosen****.*

if the label of the transition branch is $q_t^1, q_t^2, \ldots, q_t^n$, then transmit antenna i is used to send the constellation symbols q_t^i, i=1,2,…,n ***and all these transmissions are simultaneous****.*

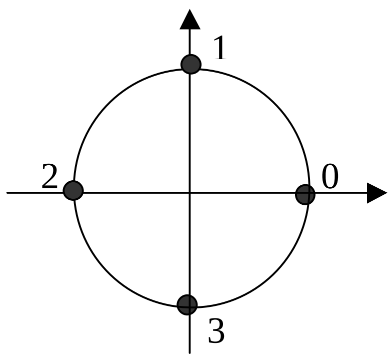
- *assuming that r_t^j is the received signal at antenna j at time t, the branch metric is given by*

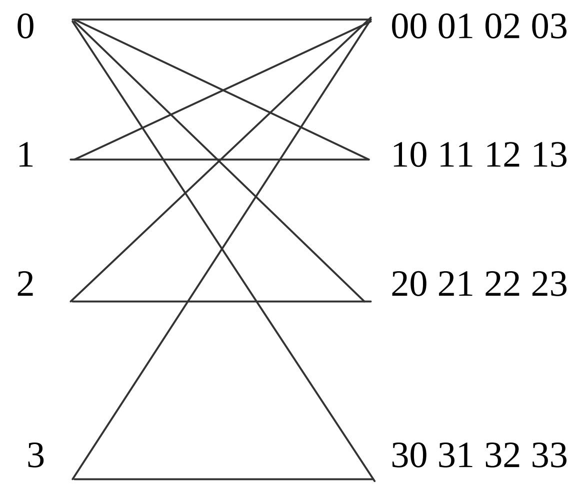
  $$\sum_{j=1}^{m}\left| r_t^j - \sum_{i=1}^{n} a_{i,j} q_t^i \right|^2$$ *The Viterbi algorithm is used to compute the path with the lowest metric.*

- *trellis codes mentioned above are **space-time** codes as they combine spatial and temporal diversity techniques.*

- *If a space-time trellis code guarantees a diversity advantage of r for the quasistatic flat fading channel model (given one receive antenna), then it is an **r-space-time trellis code**.*

Example: (2 transmit antennas, 1 receive antenna – *r = 2* here*)*





2 space-time trellis code, 4-PSK, 4 states, <u>**2 bit/s/Hz**</u>

Transmission Rate

If the diversity advantage is *nm* then the transmission rate is at most *b* bit/s/Hz with a 2^b signal constellation.
*4-PSK, 8-PSK, 16 QAM constellations will be upper bounded by 2,3,4 bit/s/Hz respectively.

Geometrical Uniformity

If a space time code is geometrically uniform, then the performance is independent of the transmitted codeword.

Algebraic Structure

4-PSK, signal points are labeled in Z_4 (integer ring of modulo 4). The edge label x_1x_2 ; x_1 *transmitted from antenna 1,* x_2 *transmitted from antenna 2*

(b_k , a_k) *input binary bits.*

The output signal pair $x_1^k\, x_2^k$ *at time* k is given by

$$(x_1^k,\ x_2^k) = b_{k-1}(2,0) + a_{k-1}(1,0) + b_k(0,2) + a_k(0,1) \qquad (\text{mod } 4)$$
$$= \left((2b_{k-1} + a_{k-1}) \text{ mod } 4,\ (2b_k + a_k) \text{ mod } 4 \right)$$

The states are ($b_{k-1}\, a_{k-1}$).

For the diversity advantage to be 2, it is required that the rank of $B(\mathbf{c},\mathbf{e})$ must be 2.

To compute the coding advantage, we need to find codewords **c** and **e** such that the determinant

$$\det\left(\sum_{k=1}^{l}(e_k^1 - c_k^1, e_k^2 - c_k^2)^*(e_k^1 - c_k^1, e_k^2 - c_k^2)\right)$$

is minimized.

We can assume that **c** is the all zero codeword since code is geometrically uniform.

j^x *gives the 4-PSK symbols.*

The **design rules:**

1. Transitions departing from the same state differ in the second symbol
2. Transitions arriving at the same state differ in the first symbol

4-PSK , 8 states:

$$(x_1^k,\ x_2^k) = a_{k-2}\,(2,2) + bk_{-1}\,(2,0) + a_{k-1}\,(1,0) + b_k(0,2) + a_k(0,1)$$

16 states:

$$(x_1^k,\ x_2^k) = b_{k-2}(0,2) + a_{k-2}\,(2,0) + b_{k-1}\,(2,0) + a_{k-1}\,(1,2) + b_k(0,2) + a_k(0,1)$$

- The constraint length of an *r*-space time trellis code is at least $r-1$.
- If b is the transmission rate, the trellis complexity is at least $2^{b(r-1)}$.

Extensions

- Multilevel Codes
- Applications in OFDM
- ST Turbo Codes

Issues:

- Performance in frequency selective channels- equalization
- Multiuser detection- performance in the presence of MAI
- Comparison to other coding schemes
- Space time processing *vs* Space time codes